Molecular characterisation of the thermostability and catalytic properties of enzymes from hyperthermophiles

Promotor: **dr. W.M. de Vos**

hoogleraar in de Microbiologie

Co-promotor: **dr. J. van der Oost**

universitair docent bij het laboratorium voor Microbiologie

Molecular Characterisation of the Thermostability and Catalytic Properties of Enzymes from Hyperthermophiles

Joyce Helen Geraldine Lebbink

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van de Wageningen Universiteit,
dr. C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 19 november 1999
des namiddags te vier uur in de Aula.

Voor mijn ouders

The research described in this thesis was funded by contracts BIOT-CT93-0274 (Biotechnology of Extremophiles) and BIO 4-CT96-0488 (Extremophiles as Cell Factories) of the European Union.

J.H.G. Lebbink - *Molecular characterisation of the thermostability and catalytic properties of enzymes from hyperthermophiles - 1999*
Dutch: 'Moleculaire karakterisatie van de thermostabiliteit en katalytische eigenschappen van enzymen uit hyperthermofielen'

Thesis Wageningen University - With summary in Dutch – 200 p.
ISBN: 90-5808-093-5

Cover: Crystals of *Thermotoga maritima* glutamate dehydrogenase N97D/G376K

Theses

---

1. Disulfide bridges are indeed common in proteins from hyperthermophiles.
This thesis page 11;
Chi, Y.-I., Martinez-Cruz, L.A., Jancarik, J., Swanson, R.V., Robertson, D.E., Kim, S.-H. Crystal structure of the β-glucosidase from the hyperthermophile *Thermospaera aggregans*: insights into its activity and thermostability. *FEBS Letters* **445**, 375-383, 1999;
Cacciapuoti, G., Fusco, S., Caiazzo, N., Zappia, V. Porcelli, M. Heterologous expression of 5'-methylthioadenosine phosphorylase from the archaeon *Sulfolobus solfataricus*: Characterisation of the recombinant protein and involvement of disulfide bonds in thermophilicity and thermostability. *Protein expression and purification* **16**, 125-135, 1999;
Singleton, M.R., Isupov, M.N., Littlechild, J.A. X-ray structure of pyrrolidone carboxyl peptidase from the hyperthermophilic archaeon *Thermococcus litoralis*. *Structure* 7, 237-244, 1999;
Hopfner, K.-P., Eichinger, A.E., Engh, R.A., Laue, F., Ankenbauer, W., Huber, R., Angerer, B. Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl. Acad. Sci. USA* **96**, 3600-3605, 1999.

---

2. The definition of salt-bridges as electrostatic interactions between oppositely charged amino acids that also form hydrogen bonds, as opposed to ion-pairs in which no hydrogen bonds are present, is extremely useful and deserves to be generally adopted.
This thesis;
Marqusee, S. & Baldwin, R. L. Helix stabilisation by Glu⁻...Lys⁺ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci. U.S.A* **84**, 8898-8902, 1987.

---

3. More is needed than a fundamentally different cell membrane to make an archaeon hyperthermophilic.
Priem, H. Het regent leven. *Natuur en techniek* 2, 68-77, 1999.

---

4. Discussions about protein thermostability among biochemists, biophysicists, microbiologists and structural biologists would benefit much from a more precise definition of the kind of stability they are referring to, i.e. thermodynamic stability, kinetic stability or resistance towards thermal inactivation.

---

5. Because the general antibacterial and antifungal agent triclosan is a specific inhibitor of the enzyme enoyl-acyl carrier protein reductase in the last step of the fatty-acid synthase cycle, this compound should be classified as an antibiotic and therefore should not be added to soap, toothpaste, cosmetics, carpets, plastic kitchenware, toys and any other product.
McMurry, L.M., Oethinger, M., Levy, S.B. Triclosan targets lipid synthesis. *Nature* **394**, 531-532, 1998.
Levy, C.W., Roujeinikova, A., Sedelnikova, S., Baker, P.J., Stuitje, A.R., Slabas, A.R., Rice, D.W. Rafferty, J.B. Molecular basis of triclosan activity. *Nature* **398**, 383-384, 1999.

---

6. Assigning compensations based on lacking notifications concerning mugs of hot coffee, little dogs in microwaves and objects in mirrors that are closer than they appear, is a grave injustice to common sense.

7. One does not catch a cold from an open window.

8. Only a mathematical impossibility: multiplying joy by sharing it.
Pythagoras, Greek philosopher, 6 century BC.

9. I think therefore I am. The opposite is rather desirable.

10. Time stays and we pass by.

11. The more complicated a problem, the more people think they know the answer.

12. A good book devours its reader.

Theses belonging to the thesis 'Molecular characterisation of the thermostability and catalytic properties of enzymes from hyperthermophiles'
Joyce Lebbink
Wageningen, 19 november 1999

Stellingen

1. Disulfide-bruggen komen inderdaad regelmatig voor in eiwitten van hyperthermofielen.
Dit proefschrift pagina 11;
Chi, Y.-I., Martinez-Cruz, L.A., Jancarik, J., Swanson, R.V., Robertson, D.E., Kim, S.-H. Crystal structure of the β-glucosidase from the hyperthermophile *Thermospaera aggregans*: insights into its activity and thermostability. *FEBS Letters* **445**, 375-383, 1999;
Cacciapuoti, G., Fusco, S., Caiazzo, N., Zappia, V. Porcelli, M. Heterologous expression of 5'-methylthioadenosine phosphorylase from the archaeon *Sulfolobus solfataricus*: Characterisation of the recombinant protein and involvement of disulfide bonds in thermophilicity and thermostability. *Protein expression and purification* **16**, 125-135, 1999;
Singleton, M.R., Isupov, M.N., Littlechild, J.A. X-ray structure of pyrrolidone carboxyl peptidase from the hyperthermophilic archaeon *Thermococcus litoralis*. *Structure* **7**, 237-244, 1999;
Hopfner, K.-P., Eichinger, A.E., Engh, R.A., Laue, F., Ankenbauer, W., Huber, R., Angerer, B. Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl. Acad. Sci. USA* **96**, 3600-3605, 1999.

2. De definitie van zoutbruggen als electrostatische interacties tussen aminozuren met tegenovergestelde lading die tevens waterstofbruggen vormen, in tegenstelling tot ionparen waarin geen waterstofbruggen aanwezig zijn, is buitengewoon handig en verdient een ruime acceptatie.
Dit proefschrift;
Marqusee, S. & Baldwin, R. L. Helix stabilisation by $Glu^{-}...Lys^{+}$ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci. U.S.A* **84**, 8898-8902, 1987.

3. Er is meer nodig dan een fundamenteel andere opbouw van het celmembraan om een archaeon hyperthermofiel te maken.
Priem, H. Het regent leven. *Natuur en techniek* **2**, 68-77, 1999.

4. Discussies over eiwitstabiliteit tussen biochemici, biophysici, microbiologen en structuurbiologen, zouden veel winnen bij een meer preciese definitie van de soort stabiliteit die bedoeld wordt, namelijk thermodynamische stabiliteit, kinetische stabiliteit of resistentie tegen inactivatie door hoge temperatuur.

5. Omdat het antibacteriële en antischimmel-middel triclosan een specifieke remmer is van het enzym enoyl-acyl dragend eiwit reductase in de laatste stap van de vetzuursynthese cyclus, moet deze verbinding geclassificeerd worden als antibioticum en dient daarom niet toegevoegd te worden aan zeep, tandpasta, cosmetica, tapijt, plastic keukenartikelen, speelgoed of welk ander produkt dan ook.
McMurry, L.M., Oethinger, M., Levy, S.B. Triclosan targets lipid synthesis. *Nature* **394**, 531-532, 1998.
Levy, C.W., Roujeinikova, A., Sedelnikova, S., Baker, P.J., Stuitje, A.R., Slabas, A.R., Rice, D.W. Rafferty, J.B. Molecular basis of triclosan activity. *Nature* **398**, 383-384, 1999.

6. Het toekennen van schadevergoedingen wegens ontbreken van waarschuwingen aangaande bekers warme koffie, hondjes in magnetrons en objecten in spiegels die dichterbij zijn dan ze lijken, doet groot onrecht aan het gezond verstand.

7. Men wordt niet verkouden van een open raam.

8. Slechts wiskundig onmogelijk: vreugde vermenigvuldigen door haar te delen.
Vrij naar **Pythagoras**, Grieks wijsgeer, 6e eeuw voor Christus.

9. Ik denk dus ik ben. Het tegenovergestelde is buitengewoon wenselijk.

10. De tijd blijft en wij gaan voorbij.

11. Hoe moeilijker een probleem, hoe meer mensen het antwoord menen te weten.

12. Een goed boek verslindt de lezer.

Stellingen behorende bij het proefschrift: *'Molecular characterisation of the thermostability and catalytic properties of enzymes from hyperthermophiles'*
Joyce Lebbink
Wageningen, 19 november 1999

Dankwoord

Dit dankwoord, alhoewel helemaal voorin geplaatst en niet zonder reden, is het laatste wat ik moet schrijven voor dit boekje af is, en het is tevens het belangrijkste en het moeilijkste. Want mijn naam mag dan wel op de omslag staan, zonder een *heleboel* andere mensen was dit boekje er nooit geweest. En het is onmogelijk te beschrijven wat jullie hebben betekent voor mijn onderzoek, mijn wetenschappelijke vorming en mijn persoontje gedurende de afgelopen vijf jaar. Ik ga dan nu ook gruwelijk te kort schieten.

Willem, als mijn promotor en begeleider heb jij het meeste invloed gehad op mijn ontwikkeling als wetenschapper. Van een goede sturing en sterke motivatie in het begin, naar een steeds grotere vrijheid naarmate het onderzoek vorderde. Altijd razendsnel en efficient, altijd betrokken en op de hoogte. Mijn dank is enorm.

John, buitengewoon inspirerend en positivistisch denkend, jouw vermogen om alles te kunnen relativeren is buitengewoon en heeft me door menig chagrijnig moment heengetrokken. En ook van jouw geleerd; tijd voor werk, tijd voor bier. Enorm bedankt voor alle discussies en de gezelligheid.

Bedankt, paranimfen Thijs en Leon, voor al jullie hulp, tips en ideeën, voor de toffe samenwerking, de kritiek en de vele emails toen, nu en straks nog.

Ans, voor de lol, de roddels, het gegiechel, de uitjes, alle tips en de vele sekwenties. Hauke, voor je betrokkenheid. Ineke, eveneens voor de sequenties, alle roddels en nog meer gegiechel. Servé, Arjen, Corné, Johan, Don, voor de gezelligheid en gedachtenwisselingen en samenwerking. Wilfried, voor al je hulp gedurende de eerste jaren (blauwe en gele epjes, rode auto's), en Rik, voor je begeleiding en jouw rol in dit alles, ver voordat het allemaal begon.

Alle studenten, Thijs, Leon en Johan (ja, alweer en net als ik blijven hangen), Michaël, Mark, Peter, Eveline en Sander, PGO 1996 en 1998; bedankt voor de samenwerking, het enthousiasme, de interesse, de gein, het gebakken spekje en de zandloper.

Bedankt, Frits, voor de razendsnelle reparaties en de onderonsjes. Nees, voor alle noodzakelijke en moeilijke transacties. Wim, voor hulp met eigenzinnige computers en printers. Ria, Jannie, Renée, Ans en Francis voor de keuken boven (vooral de koffie en de ineens weer witte mok), de keuken beneden, de bieb en de administratie. Alle andere Micro's, dankjulliewel.

Stefan, many and even more thanks for introducing me into the wonderful world of protein crystallography. For the time you made for me when you were awfully busy finishing your own thesis (I am so sorry for the Dutch virus!). For your hospitality (and yours, Susanne !) when I came to work at CSB three years ago. For the very fruitful collaboration on GDH, many many discussions, emailed versions of papers, nice beers at nice conferences and many a laugh.

Rudolf, also for a very fruitful collaboration on GDH, for letting me visit your lab three years ago, for lots of fun and the dance on Capri.

Valerio and Roberta, for more fruitful collaborations, not only on GDH, and for your hospitality, for introducing me to yet another world, that of complicated and fascinating biophysics. And even more thanks for your love and

care during my stay in Roma, most beautiful of all cities; for the jokes, cappuccinos, the countryside weekend, the diners. And Roberto, thank you so much for your hospitality and charms and care.

David, Kitty, Tim, Linda and Pat, for two nice stays in Sheffield, for your help, nice discussions and yet again a succesful collaboration on GDH, for the barbecues and the very English excursion to the ruined castle in the mist and the King's Arms (or the Queen's ?) were we enjoyed game pie and beer.

Marco, Maria and Mosè in Napoli, for the endless talks on LacS and CelB, a nice London weekend with Indian Curries, Jacked Potatoes and Dinosaurs, and for the LacS coordinates (also Laurence !). Jürgen and Georg in Freiburg, for the interesting collaboration on the structure of CelB (and many mutants). Ana and Helena in Oeiras for a nice collaboration on compatible solutes. All those European colleagues in the Biotechnology of Extremophiles and Extremophiles as Cell Factories networks; for wonderful meetings in wonderful places, for stiff and inspiring discussions, for lots of fun. The European Union, for funding my project.

Collega's aan de overkant, Willem, Walter, Michel en Carlo op Biochemie, Leo op Erfelijkheidsleer, enorm bedankt voor de interesse, discussies en de mogelijkheid tot het doen van metingen.

My colleagues in my new lab in Sweden; Rudolf (again!) for your trust in me and your support, for the fascinating new projects; Gudrun (again so sorry for the Dutch virus !), Winfried, Heiko, Mikael, Niklas, Xiaofeng, Bin, Jordi, Andrej, Assen, Gioia, for making me feel very welcome and very much at home, for support during the last awful moments of getting this ready for printing, for many laughs and tea-times and the Great Canoeing Event.

Rob en Nathalie, Andrien en Bertil, Nicole en Bouke, Jürgen, Guido, Sander en Bart, Droef 63-ers, meppers, wavers, chocolademousse en Fawlty Towers toetjesnachters, Gothic Party Ouwelullen, Oud en Nieuw feestbeesten, Zaaijer-hangers. Voor de foto's-van-jullie compilatie zodat deze Sjoislich Aio zich niet zou opsluiten op het lab. Dikkie-Dick, voor die avonden dat ik mezelf wél opsloot en dan bij jou kwam eten. Peter, lachende keukenprins, voor de enerverende gebeurtenissen in jouw leven. Pat en Sas, voor roeibootjes en picknicks. De windsurf- en snowboard clan; Bart, Anja, Roel, Pascal, Ronald, Anne-Claire, Koert, Emile, Loes, Winfried, Michael, Anouk, Mariska, Tom, Toine, Yvonne; voor het bevriezen in de Grevelingen of bovenop een gletsjer, voor supergaaf en helemaal tof, voor een hoop herrie en voor het passen op mijn vent.

Onontbeerlijk voor de nodige rust en onrust, vooral toen de proefjes en het boekje écht afmoesten; Stephanie, Rick, Sascha en Aimée, Piet en Els, Pa en Ma.

En Marcel, vriend en maatje, dankjewel voor je vertrouwen in mij. Dankjewel dat je het met mij uithoudt, eerst onuitstaanbaar achter de compu om dit af te krijgen en nu onuitstaanbaar ver weg.

Pa en Ma, dit is voor jullie.

Joyce

Contents

Chapter 1

General Introduction

1. Hyperthermophilic Archaea and Bacteria

Microbial life is thriving in habitats that are considered highly inhospitable to man. The isolation of so-called 'extremophiles' from boiling geysers, geothermally heated sediments, acid mudholes, hypersaline inland lakes and below Antarctic ice-floors, has vastly expanded our conceptions of which environments are able to sustain microbial life. The wide variety of phenotypes of these exotic microorganisms, most of which are members of the domain of the Archaea, is listed in Table 1. Studying the mechanisms by which extremophiles are able to survive and reproduce, is adding much to our current knowledge of fundamental biological issues. Furthermore, these microorganisms may be a valuable source of industrially relevant biocatalysts and compounds. Current knowledge is increasing rapidly and many aspects like genetics, physiology and biochemistry have recently been reviewed in special journal issues of FEMS Microbiological Reviews (Antranikian *et al.*, 1996), Cell (Lewin, 1997), and Advances in Biochemical Engineering/Biotechnology (Antranikian, 1998). Moreover, a novel journal has recently been established that is entirely dedicated to these extremophiles (Horikoshi, 1997). Considerable attention is given to the specific adaptations of structure and function of macromolecules in extremophiles. As a consequence, a variety of three-dimensional crystal and solvent structures of proteins from extremophiles has been determined that allows for the elucidation of the molecular basis of protein stability and the mechanisms of enzyme catalysis under extreme conditions. In addition, within a time span of less than 4 years, the complete genomes from 20 organisms have been published, among which 5 genomes from extremophiles (TIGR, 1999). More than 90 genomes (8 from extremophiles) are currently under study and will be finished within a few years. Complete genomes allow comparative analysis of protein amino acid composition, components of metabolic pathways and transcriptional, translational and repair mechanism and provide phylogenists with enormous amounts of new data. Together with DNA chip or array technologies, complete genomes enormously expand our ability to study gene regulation, protein function and genetic disorders. Comparative genomics, in concert with comparative biochemistry and crystallography, will be powerful tools in the identification of characteristics that enable an organism to survive

Table 1: Classification of archaeal (extremophilic) phenotypes (Danson & Hough, 1998)

Thermophiles	Grow optimally between 55 and 80°C
Hyperthermophiles	Grow optimally between 80 and 113°C
Psychrophiles	Require temperatures below 20°C for growth
Halophiles	Grow optimally at high salt concentrations; internal concentrations isotonic with exterior
Methanogens	Reduce C1 and C2 compounds (anaerobically) to methane. Mesophilic, (hyper)thermophilic and halophilic members
Acidophiles	Grow optimally below pH 4
Alkaliphiles	Grow optimally above pH 9
Barophiles	Grow optimally up to 50 MPa pressure

and reproduce efficiently in extreme environments.

Hyperthermophiles are organisms that are able to grow optimally above 80°C. The upper temperature of life has recently been set on 113°C with the isolation of *Pyrolobus fumarii* (Blöchl *et al.*, 1997). Most of the hyperthermophiles belong to the domain of the Archaea, together with the extreme halophiles and the methanogens (Woese *et al.*, 1990). Two hyperthermophilic bacterial genera have been described, namely *Thermotoga* sp. and *Aquifex* sp. In the original 'universal tree of life', based on 16S/18S rRNA sequence comparisons, all hyperthermophiles cluster at the base of the tree and many of their branches are short (Figure 1) (Woese *et al.*, 1990). Hyperthermophiles therefore seem to be the closest relatives to our last universal common ancestor. This resulted in the hypothesis that this last universal common ancestor might have been a hyperthermophile and life might have evolved at high temperatures (Stetter, 1994). This would mean that what we now tend to call 'adaptations' to high temperatures, are in fact the originally evolved mechanisms, and that mesophily and consequently, psychrophily, should be regarded as adaptations to low temperatures.

Comparison of the rRNA tree with phylogenetic trees based on protein sequences, often results in conflicting topologies. These conflicting gene histories can be explained by differences in evolutionary rates of the particular macromolecules as well as lateral gene transfer (Forterre, 1998; Brown & Doolittle, 1997). Taking these processes into consideration, a 'genetic annealing model' has been proposed, in which the last universal common ancestor is no longer a primitive organism, but rather a community of primitive cells

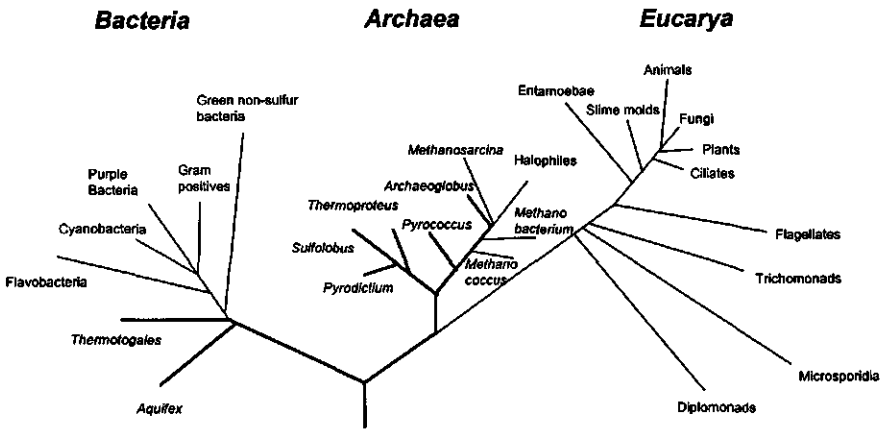


Figure 1: Evolutionary relationships among living organisms in the three-domain model based on 16S/18S rRNA sequences according to Woese (modified from Woese *et al.*, 1990; Stetter, 1996). Rooting of the tree is derived from paralogous pairs of protein sequences Ef-Tu/EF-G and F- and V-ATPases (Iwabe *et al.*, 1989; Gogarten *et al.*, 1989). Hyperthermophilic genera are depicted as thick lines.

(progenotes) that survives and evolves as a biological unit (Woese, 1998). In this community a high mutation rate and high level of lateral gene transfer initially defined evolution, in contrast to vertical (genealogical) inheritance. Ultimately, the ancestors of the three primary domains evolved from this entity (Woese, 1998).

An alternative hypothesis for the evolutionary relationship among living organisms has recently been suggested (Gupta, 1998). This hypothesis is based on (i) multiple alignments of protein sequence data, (ii) entire genome sequences of several prokaryotic and eukaryotic organisms, and (iii) integration of signature sequences and several morphological features (Gupta, 1998). In the proposed model the primary division within prokaryotes is not between Archaea and Bacteria, but rather between prokaryotes that have a single cell membrane (monoderms; including all archaebacteria and gram-positive bacteria) and prokaryotes that are surrounded by an inner cytoplasmic membrane and an outer membrane (diderms; all 'true' gram-negative bacteria). The eukaryotic cell may have originated from a fusion between a thermoacidophilic archaebacterium and a gram-negative bacterium, with the eucaryal genome arising as a chimera with distinct archaeal and bacterial features (Gupta, 1998). In an alternative theory, eukaryotes are thought to have evolved through symbiotic association of a respiring and molecular hydrogen generating eubacterium with an anaerobic, hydrogen-dependent, autotrophic archaeon, as is suggested by comparative biochemistry of the energy metabolism (Martin & Muller, 1998).

While both the three-domain and the chimeric model suggest that the last common universal ancestor could have been a hyperthermophile and that life may have evolved at high temperatures, several authors have argued against this notion (Forterre, 1995; Forterre, 1998; Miller & Lazcano, 1998). Reverse gyrase, an enzyme involved in regulating DNA topology, seems only to be present in hyperthermophiles, but is most probably the result of a fusion of already highly evolved helicase and topoisomerase modules. If reverse gyrase is also a *prerequisite* for hyperthermophily, the ancestors from which its constituting modules were derived, were probably not hyperthermophiles (Forterre, 1995; Forterre, 1998). The hyperthermophilic origin of the last universal common ancestor is based on the rooting of the universal tree; however, this rooting in the prokaryotic branch is still under debate (Forterre *et al.*, 1993; Forterre, 1997; Doolittle & Brown, 1994). Further considerations indicate that hyperthermophily is not in all cases an ancestral phenotype, but may have evolved separately during evolution. For example, based on RNA polymerase sequences, the hyperthermophilic bacterium *Aquifex pyrophilus* is not located at the base of the universal tree, but is placed among the Gram-negative bacteria, below the ε-subdivision of the proteobacteria (Klenk, 1994). The complete genome sequence of *Aquifex aeolicus*, however, has as yet not yielded a statistically significant placement of the *Aquifex* lineage (Deckert *et al.*, 1998). It did reveal that *Aquifex* contains a two-fold higher number of 'archaeal' homologues than other bacteria (Aravind *et al.*, 1998). It has therefore been suggested that *Aquifex* may have acquired its hyperthermophilic phenotype by the acquisition of genes from hyperthermophilic Archaea via horizontal gene transfer (Aravind et *al.*, 1998). Further evidence for independent evolution of hyperthermophily has been based on the mechanism of cell membrane stabilisation. While bacterial

Thermotogales as well as Archaea have developed *analogous* mechanisms to stabilise their membranes, i.e. the formation of covalently-linked ether lipids, these mechanisms are not *homologous* since the lipids have opposite stereochemistry and different chemistry (Langworthy & Pond, 1986; Forterre, 1996). Even between different members of the hyperthermophilic Bacteria, different patterns of lipid composition have been recognised (Langworthy & Pond, 1986). Also on the protein level there is indication that secondary adaptation to thermophily in Bacteria may have evolved twice independently. In many DNA polymerases widely distributed in the three domains of life, 3' to 5' exonuclease activity, residing either in subunits or in homologous domains, is present and therefore appears to be very ancient (Forterre *et al.*, 1994). Taq DNA polymerase from the extreme thermophilic bacterium *Thermus aquaticus* and a DNA polymerase from *Bacillus stearothermophilus* have lost this 3' to 5' exonuclease activity that is present in their mesophilic homologues. The crystal structures of these enzymes suggest that the exonuclease activity has been sacrificed in order to achieve thermostability (Korolev *et al.*, 1995; Kiefer *et al.*, 1997). In Taq polymerase the DNA-binding crevice in the exonuclease domain has been filled with hydrophobic residues, in the *Bacillus* enzyme this crevice has been blocked by a particular polypeptide loop (Korolev *et al.*, 1995; Kiefer *et al.*, 1997). The different nature of the structural adaptations suggest that these occured independent of each other in a DNA polymerases present in a mesophilic ancestor (Forterre, 1998). Further insight into the origin of hyperthermophilicity and hyperthermostability can only be obtained by multidisciplinary approaches combining sequence-based molecular phylogenies with comparative biochemistry and structural and molecular biology.

2. Protein stability

The term 'protein stability' refers to the resistance of a protein towards influences such as heat or denaturants that may affect its structural integrity and/or its biological function. High temperatures or high concentrations of denaturant may cause unfolding of the native protein, which means loss of quaternary, tertiary and/or secondary structure. The loss of correct positioning of certain residues in three-dimensional space means that the protein is no longer able to carry out its biological function. This *denaturation* is usually cooperative and, in theory, reversible. Polypeptides are furthermore prone to chemical modifications, which are often accelerated under extreme conditions such as high temperature or low pH. These modifications may result in irreversible loss of activity or *inactivation* (Klibanov, 1983a; Klibanov, 1983b; Klibanov & Mozhaev, 1978; Mozhaev & Martinek, 1982). A variety of inactivating processes have been described that usually follow (and not precede) unfolding. These include proteolysis, loss of essential cofactors or prosthetic groups, aggregation and precipitation, formation of refolding intermediates in kinetically trapped conformations, peptide bond hydrolysis, cysteine oxidation, deamidation and breakage of disulphide bridges (Fágáin, 1997). Both types of *in vivo* protein stability can be depicted in the following relation:

$$N \overset{K}{\leftrightarrow} U \overset{k}{\rightarrow} I \qquad (1)$$

in which N represents the native, U the unfolded and I the irreversibly denatured or inactivated form of the protein, K (= k_1 / k_{-1}) is the equilibrium constant for the reversible N ↔ U transition, and k is the rate constant for the irreversible U → I reaction.

Thermodynamic or conformational stability is described by the Gibbs free energy of unfolding ($\Delta^U_N G$). This is the difference in free energy between the native and the denatured state of the protein. Providing that reversibility exists and a two-state, cooperative transition between the native and the unfolded protein occurs, the unfolding free energy may be calculated according to

$$\Delta^U_N G = G^U - G^N = \Delta^U_N H - T\,\Delta^U_N S \qquad (2)$$

and

$$\Delta^U_N G = -RT \ln K \qquad (3)$$

in which $\Delta^U_N H$ and $\Delta^U_N S$ are the enthalpy and entropy differences between the native and the unfolded protein, respectively, T is the absolute temperature and *R* is the gas constant. The requirement of reversibility implies the complete absence of irreversible denaturation or inactivation, or a sufficiently low rate constant so that relation (1) reduces to

$$N \overset{K}{\leftrightarrow} U \qquad (4)$$

Techniques that allow determination of N and U, like UV spectroscopy, fluorescence or circular dichroism, or techniques like differential scanning microcalorimetry that determine the denaturation enthalpy and the change in heat capacity upon protein thermal unfolding, allow calculation of thermodynamic parameters and thereby the conformational stability of a certain protein. Conformational stability may be enhanced by lowering the free energy of the native state as well as by raising the free energy of the unfolded state. Furthermore, an enzyme may be stabilised by lowering the rate constant for unfolding (k_1), i.e. by raising the activation energy barrier (Gibbs free energy for activation). This dependence of stability on the activation energy is referred to as kinetic stability and may be studied using techniques that allow quantitative determination of the disappearance of N.

Irreversible inactivation is a kinetically controlled process and is described by

$$-d[A]/dt = k\,[A] \qquad (5)$$

in which A is activity, t is time and *k* is the rate constant for inactivation. Enzyme inactivation is followed by determining the remaining catalytic activity after

progressive incubation at high temperature or incubation with denaturants. If the loss of activity follows first-order kinetics, or pseudo first-order kinetics, it may be described according to

$$A_t = A_0 * e^{-kt} \quad (6)$$

and the half-life for inactivation is defined as ln 2 / *k*.

The thermodynamic stability of a protein is determined by the marginal net free energy difference of a large number of stabilising and destabilising interactions. This net difference in free energy is typically in the range of 5-17 kcal/mol, which is equivalent to the energy of only a few hydrophobic interactions, ion-pairs or hydrogen bonds (Jaenicke, 1991; Pfeil, 1986). A hypothetical stability curve for a protein from a mesophilic organism is shown in Fig. 3 (curve a), in which the free energy is plotted against temperature. The hyperbolic nature of the curve implies that proteins do not only denature at high temperatures (T_m, or melting temperature) but also show cold denaturation (T_c). For thermostable proteins the free energy curve may have been shifted either vertically (i.e. increase in maximum free energy; curve b), horizontally (i.e. increase in melting temperature; curve c) or may have been flattened (i.e. raise in melting temperature but a *decrease* in maximum free energy; curve d). Examples of non-enzymatic proteins that have been studied in this respect include rubredoxin from *P.furiosus*, ferredoxin from *T.maritima* and histones from the hyperthermophiles *Methanothermus fervidus* and *Pyrococcus* strain GB-3a, which are stabilised by a combination of (b) and (c) (Hiller *et al.*, 1997; Pfeil *et al.*, 1997; Li *et al.*, 1998). The DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus* has a raised melting temperature at the expense of maximum free energy (d) (Knapp *et al.*, 1996). Examples of thermodynamically studied enzymes include aspartate aminotransferase from *Sulfolobus solfataricus*, which is 15 kJ/mol more stable at 25°C than the cytosolic aspartate aminotransferase from pig heart, and the trimeric adenylate kinase from *S.acidocaldarius* for which a maximum intrinsic free energy of 130 kcal/mol-trimer at 32°C is reported (Arnone *et al.*, 1997; Backmann *et al.*, 1998).

Studies on the stability of proteins from hyperthermophiles are currently limited and restricted to the examples described above, since they suffer from several practical considerations. Enzymes often allow easy determination of irreversible loss of catalytic function (which is the most important parameter for the microbial cell and for biotechnological purposes) by spectrophotometric quantification of substrates or reaction products. However, enzymes often do not reversibly unfold and are therefore not accessible to thermodynamic analysis (Klump *et al.*, 1992; Jaenicke, 1996; Pappenberger *et al.*, 1997; Jaenicke & Bohm, 1998). Studies on large model enzymes are therefore limited to description of kinetic stability. Small, monomeric proteins or isolated domains, in comparison with enzymes, more often denature reversibly (Knapp et *al.*, 1996; Perl *et al.*, 1998; Li *et al.*, 1998). However, due to their small size, the contribution of large structural features and domain/subunit interfaces cannot be

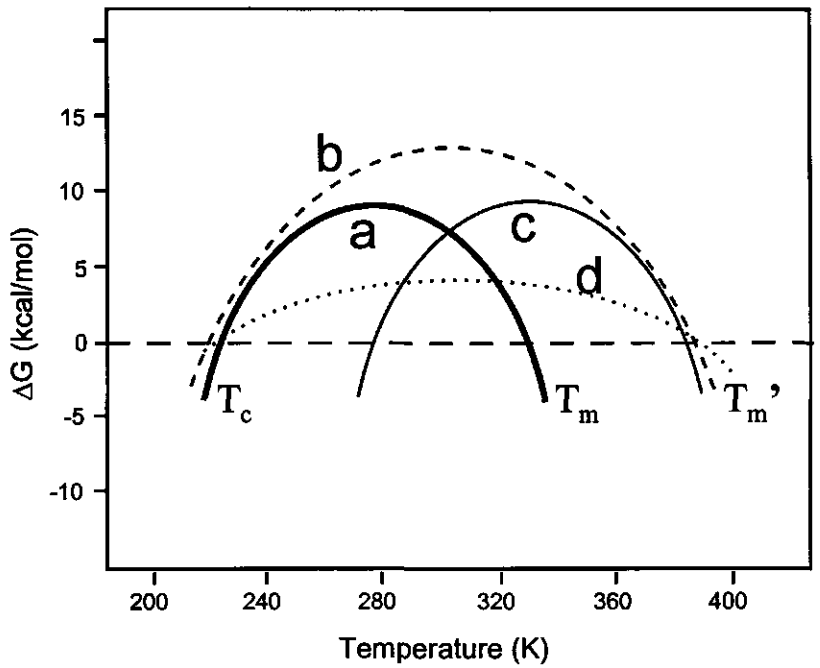


Figure 3: Hypothetical free energy temperature profiles from (a) proteins from mesophiles and (b-d) thermostable proteins. T_m and T_m' are the melting temperatures of the proteins from mesophiles and thermostable proteins, respectively. T_c is the temperature at which cold denaturation occurs (modified from Jaenicke & Bohm, 1998).The temperature dependance of ΔG is defined by a modified Gibbs Helmholtz equation (Becktel & Schellman, 1987):

$$\Delta G^{U}_{N}(T) = \Delta^{U}_{N}H(T_m)-T\Delta^{U}_{N}S(T_m)+ \Delta^{U}_{N}C_p((T-T_m)-T\ln(T/T_m)) \quad (7)$$

in which ΔG^{U}_{N} is the difference in free energy, $\Delta^{U}_{N}H$ the difference in enthalpy, $\Delta^{U}_{N}S$ the difference in entropy and $\Delta^{U}_{N}C_p$ the difference in heat capacity between the native and the denatured state, and T and T_m are the absolute temperature and the melting temperature, respectively.

studied in these systems. In this thesis, the main interest is focused on large structural features in model enzymes that do not unfold reversibly, and the role of these features will therefore mainly be studied with respect to kinetics of thermal inactivation. Thermal- and denaturant-induced unfolding will be determined in order to compare wild-type and mutant variants of the enzymes, without concomitant thermodynamic analysis of the obtained unfolding curves.

3. Molecular adaptations to high temperatures

It has been well established that, in their respective operational temperature ranges, enzymes from psychrophiles, mesophiles and hyperthermophiles are in so-called corresponding states (Jaenicke, 1991). This

means that the flexibility of the enzymes has been tuned to the desired temperature such that optimal catalysis is possible without the interference of denaturation and inactivation. Consequently, while enzymes from mesophiles are good catalysts around room temperature, hyperthermostable enzymes are rigid at this temperature and conformational flexibility is too low for efficient catalysis. On the other hand, at high temperatures, thermostable enzymes reach optimal conformational flexibility, while enzymes from mesophiles have long been inactivated and denatured because of thermal motions that exceed the capacity of the interactions to maintain an intact enzyme structure. Since all proteins, independent of their origin, consist exclusively of the 20 canonical natural amino acids, this intrinsic enzyme flexibility and stability in psychro-, meso- and thermostable proteins is solely determined by the amino acid composition and the interaction between them in the native enzyme structure. Regarding this intrinsic adaptation, one must, however, not overlook the temperature dependance of chemical equilibria (Danson *et al.*, 1996). A consequence of this is, for example, a large downfield shift in pK_a (as much as 2.5 pH units on going from 25 to 100°C) for the basic side chains of histidine, lysine and arginine; in contrast, the carboxylate side chains of glutamate and aspartate are only slightly affected. All these charged residues are often directly involved in enzymatic catalysis as well as in stabilising the native enzyme structure. It is known that pK_a values can be fine-tuned by interactions with other amino acids and by the existing electrostatic field, and it is obvious that this fine-tuning must be achieved in different ways in enzymes from mesophiles and (hyper)thermophiles (Danson et al., 1996).

Proteins may furthermore be stabilised by extrinsic factors. In the cytosol, the natural substrates and cofactors may help to achieve and maintain the native protein structure (Risse *et al.*, 1992). Many hyperthermophiles have been reported to accumulate certain molecules, which despite their high intracellular concentrations, do not interfere with metabolic processes, and are therefore referred to as compatible solutes (for a recent review see da Costa *et al.*, 1998). On the one hand, compatible solutes are thought to be stabilising by being preferentially excluded from the protein surface (Arakawa & Timasheff, 1985). The preferential hydration of the protein surface forces a compact conformation, opposes an increase in surface area, and thereby favours the native state since unfolding usually results in an increase in surface area (Arakawa & Timasheff, 1985). On the other hand, compatible solutes have been proposed to interact specifically with the protein surface, and the stabilisation of a protein by a certain compatible solute may be a balance between its interaction with both water and protein (Carpenter *et al.*, 1992; Ramos *et al.*, 1997). Another extrinsic factor is the presence and action of molecular chaperones. Chaperones are proteins that facilitate folding of polypeptide chains and prevent aggregation of unfolded proteins (Hartl, 1996). They are required under normal physiological conditions and their production is often induced in response to heat-shock or other stress-factors (Hartl, 1996). High levels of heat-shock 'thermosomes' have been reported in many hyperthermophiles as heat-stress response (Trent *et al.*, 1990; Minuth *et al.*, 1998; Holden & Baross, 1993).

In this thesis, the adaptation of enzyme flexibility and stability to different temperatures will be the central issue. This chapter therefore continues with a description of the factors that determine intrinsic protein stability.

Primary structure analysis

Correlations between protein thermostability and amino acid composition include a decrease in chain flexibility, a decrease in content of chemically labile residues, an increase in average hydrophobicity, and an increase in frequency of aromatic residues (Menendez-Arias & Argos, 1989; Vihinen, 1987; Zwickl *et al.*, 1990; Eggen et *al.*, 1994; Vieille & Zeikus, 1996). The increase in isoleucine content in hyperthermostable glutamate dehydrogenases (GDH) has been correlated with the ability of this residue to adopt additional rotamer conformations, and hence may result in increased packing efficiency (Britton *et al.*, 1995; Chapter 4 of this thesis). The thermolabile amino acid residue cysteine seems to be almost completely absent in many hyperthermophile proteins (Eggen *et al.*, 1993; Voorhorst *et al.*, 1995; Vieille & Zeikus, 1996). They are only conserved in redox proteins, where they serve to ligand the redox-centres, as for example in aldehyde ferredoxin oxidoreductase and glyceraldehyde-3-phosphate ferredoxin oxidoreductase (Chan et *al.*, 1995; van der Oost et *al.*, 1998). Disulfide bridges have been suggested to be important stabilising factors in moderately stable proteins (Matsumura, 1989). In hyperthermophiles,

Table 2: Relative amino acid compositions (in percentages) of mesophiles and hyperthermophiles. [1]Include *Haemophilus influenza, Helicobacter pylori, Escherichia coli* and *Synechosystis*. [2]Include *Aquifex aeolicus* and *Methanococcus jannaschii*. (Modified from Deckert et *al.*, 1998).

Amino acid	Mesophiles[1]	Hyperthermophiles[2]
Ala	8.42	5.72
Cys	1.06	1.03
Asp	5.01	4.92
Glu	6.37	9.15
Phe	4.37	4.67
Gly	6.90	6.58
His	2.09	1.49
Ile	6.64	8.89
Lys	6.00	9.88
Leu	10.79	9.98
Met	2.43	2.13
Asn	4.59	4.42
Pro	4.13	3.73
Gln	4.51	1.74
Arg	4.67	4.38
Ser	5.95	4.63
Thr	5.11	4.14
Val	6.62	7.39
Trp	1.15	0.82
Tyr	3.10	4.23
Charged residues (DEKRH)	24.11	29.84
Polar/uncharged residues (GSTNQYC)	31.15	26.79
Hydrophobic residues (LMIVWPAF)	44.74	43.36

however, these kind of interactions have only been reported in the TATA-binding protein from *P.woesei* (Dedecker *et al.*, 1996) and proteins in which they serve a functional role as for example in the protein disulfide oxidoreductase from *P.furiosus* (Ren *et al.*, 1998). The existence of this latter protein, however, suggests that disulfide bridges may be more common in hyperthermophiles than currently thought.

Often, however, the above-mentioned correlations do not hold (Vieille & Zeikus, 1996). An example is the reduction in glycine residues correlating with decreased chain flexibility, versus an increase in glycine residues in hyperstable proteins that reduce strained conformations which are present in mesophilic homologues (Macedo-Ribeiro *et al.*, 1997). Until recently, surveys dealt with relatively small sample sizes, and a concise re-evaluation proved several of these early predictions to be statistically insignificant (Bohm & Jaenicke, 1994). The availability of many complete genome sequences, among which several from hyperthermophiles, has allowed a survey with a large data set (Table 2) (Haney *et al.*, 1998; Deckert *et al.*, 1998). From these studies, it is concluded that proteins in hyperthermophiles on average contain higher levels of charged residues and lower levels of polar, uncharged residues. Residues with a higher side chain volume and higher average hydrophobicity appear to be preferred. Unlike asparagine, glutamine is much discriminated against, although both residues are prone to deamidation at high temperatures. A difference in number of cysteine residues is, remarkably, not detected in this analysis.

Several reservations against these kinds of analyses must however be made. Protein amino acid composition is known to reflect to some extent the underlying G+C content of the genome (Sueoka, 1961; Lobry, 1997). This effect may even exceed environmentally determined adaptations to such extent that highly unlikely clustering in phylogenetic trees occurs (Benachenhou-Lahfa et *al.*, 1994). Additionally, it has recently been calculated that amino acid synthesis at high temperatures, under conditions that prevail in hydrothermal vents, is often thermodynamically favoured, to such an extent that the synthesis of about half of the amino acids generates energy (Amend & Shock, 1998). In contrast, in 25°C sea water, synthesis of all amino acids requires energy. Although this prediction has not been corroborated with database analysis as yet, it cannot be excluded that this may have an effect on the amino acid composition of proteins in organisms that thrive in these habitats. Finally, the ultimate factor determining whether a certain amino acid is present in a protein, is not its chemical stability at high temperatures as isolated amino acid, but its stability and functional role in the context of its surroundings in the protein, i.e. whether it is surface-exposed or buried and protected against degradation. Therefore, the analysis of three-dimensional structures, albeit restricted to much smaller data sets than primary structure analysis, is expected to provide valuable information, as described below.

Structural adaptations

The elucidation of many solvent and crystal structures from hyperthermostable enzymes and their comparison with structures from homologous enzymes from mesophiles (and, in some cases, psychrophiles), has

Table 3: Proposed stabilising strategies that have been identified in hyperthermostable enzymes from the archaeal genera *Pyrococcus* and *Sulfolobus*, and the bacterial genera *Thermotoga* and *Aquifex*.

Enzyme	Organism	Proposed stabilising strategies	Reference
Glutamate dehydrogenase	*P.furiosus*	Extensive ion-pair networks at domain and subunit interfaces Isoleucine clusters	Yip *et al.*, 1995
Citrate synthase	*P.furiosus*	Increased compactness Intimate association of subunits Increased number ion-pair networks Decreased number and burial of thermolabile residues	Russell *et al.*, 1997
Aldehyde oxidoreductase	*P.furiosus*	Small solvent exposed surface area Large number of ion-pairs and buried atoms	Chan *et al.*, 1995
Ornithine carbamoyl Transferase	*P.furiosus*	Oligomerisation with increased hydrophobicity at subunit interfaces	Villeret *et al.*, 1998
TATA-binding protein	*P.woesei*	Increased number of surface electrostatic interactions Increased compactness Disulphide bridge formation	Dedecker *et al.*, 1996
Indole-3-glycerol phosphate synthase	*S.solfataricus*	Increased number of salt-bridges Improved helix-capping Improved hydrophobic interactions	Hennig *et al.*, 1995
β-glycosidase	*S.solfataricus*	Large surface-exposed ion-pair networks Solvent-filled hydrophilic cavities	Aguilar *et al.*, 1997
Superoxide dismutase	*S.acidocaldarius*	Increased number of intersubunit ion-pairs and hydrogen bonds Decrease of solvent accessible hydrophobic surface area Increased burial of hydrophobic residues	Knapp *et al.*, 1999
Glutamate dehydrogenase	*T.maritima*	Ion-pair networks at domain interfaces Increased hydrophobic interactions at subunit interfaces	Knapp *et al.*, 1997
Glyceraldehyde-3-phosphate dehydrogenase	*T.maritima*	Increased number of salt-bridges	Korndorfer *et al.*, 1995
Phosphoribosyl anthranilate isomerase	*T.maritima*	Increased hydrophobicity Increased number of salt-bridges	Hennig *et al.*, 1997
Superoxide dismutase	*A.pyrophilus*	Increased compactness of tetramer Ion-pairs in small networks Increase in buried surface area	Lim *et al.*, 1997

revealed a variety of features that might contribute to the increased stability of hyperthermophile enzymes (see Table 3 for an overview and references). Proposed stabilising mechanisms include an increase in the number of hydrogen bonds, charged-neutral hydrogen bonds and ion-pairs, as well as the arrangement of ion-pairs in large networks. Electrostatics can furthermore be optimised by increasing the α-helix-dipole stabilisation. This electrostatic optimisation fits well with the increase in charged residues found in the hyperthermophiles *Methanococcus jannaschii* and *Aquifex aeolicus* for which the complete genomes have been sequenced (Bult *et al.*, 1996; Deckert *et al.*, 1998). Furthermore, Van der Waals interactions seem to be optimised by improved packing of the hydrophobic core, resulting in a decreased number and total volume of cavities, as well as in a lower surface-to-volume ratio. Additionally, a lower surface-to-volume ratio may be achieved by shortening or deletion of flexible surface loops. Alternatively, such loops have been strengthened by rigidifying amino acid substitutions. Additional proposed mechanisms include reduction of conformational strain, an increase in aromatic interactions, and stabilisation of α-helices by introduction of residues with enhanced secondary structure propensity. A relative increase in polar surface area and decrease in hydrophobic surface area has been reported, as well as tightening of contacts at interdomain and intersubunit interfaces and formation of higher association states. A reduction in the number of thermolabile residues or their burial in the protein interior protects against thermal inactivation by covalent destruction. Finally, the presence of solvent-filled, hydrophilic cavities may confer resilience, i.e. the ability to return to the original state after being stretched or bent, allowing larger fluctuations of the polypeptide without complete unfolding occurring. This higher resilience is also conferred by the large ion-pair networks that are mentioned earlier. These networks have been found in many of the hyperthermostable enzyme structures described to date, and will be discussed in detail below.

Ion-pairs and ion-pair networks

In proteins, the N-terminal amino group and side chains of histidine, arginine and lysine may be positively charged. The C-terminal carboxyl group and side chains of aspartate and glutamate may be negatively charged. When oppositely charged groups are within 4 Å distance, the favourable energetic interaction that is formed, is called an ion-pair (Barlow & Thornton, 1983). When the charged groups also form hydrogen bonds, these particular ion-pairs are preferably referred to as salt-bridges in order to distinguish between these different forms of electrostatic interactions (Marqusee & Baldwin, 1987). Unfortunately, this distinction is not consequently followed throughout literature.

A role for ion-pairs in protein stabilisation was already postulated some twenty years ago when Perutz described additional salt-bridges in the structures from ferredoxin and glyceraldehyde-3-phosphate dehydrogenase from thermophilic bacteria (Perutz, 1978; Perutz & Raidt, 1975). Since then, however, several studies have produced controversial results. A buried salt-bridge in lysozyme contributes 3-5 kcal/mol to the free energy of stabilisation (Anderson

et al., 1990), while a buried salt-bridge triad in Arc repressor can be replaced by a more stable hydrophobic core (Waldburger *et al.*, 1995). Surface salt-bridges in barnase and lysozyme stabilised by only 0.98-1.25 and 0.1-0.25 kcal/mol (Horovitz *et al.*, 1990; Dao-pin *et al.*, 1991). It was argued that the entropic cost of localisation of solvent-exposed charged groups on the surface of a protein largely offsets the interaction energy expected from the formation of a defined salt-bridge (Dao-pin *et al.*, 1991; Hendsch & Tidor, 1994). In networks, however, each additional ion-pair interaction requires the desolvation and localisation of only one side chain. Cooperativity was indeed shown in barnase where two salt-bridges in a triad reduce each other's energy by 0.77 kcal/mol (Horovitz *et al.*, 1990). Furthermore, networks are often located at interfaces or in clefts on the protein surface, where the folding of the protein has already provided part of the entropic cost (Dao-pin *et al.*, 1991). Networks allow a higher flexibility which increases the entropy of these arrangements without diminishing the total electrostatic energy (Knapp, 1996; Aguilar *et al.*, 1997). While at room temperature none or only marginal effects of salt-bridges were reported (Horovitz *et al.*, 1990; Dao-pin *et al.*, 1991; Hendsch & Tidor, 1994), several considerations indicate that at high temperatures they may play a crucial role. The dielectric constant decreases with temperature, which results in an increased electrostatic energy upon formation of an ion-pair or salt bridge. Furthermore, hydration free energies are markedly reduced at high temperatures and ion-pair formation is therefore favoured. Because this reduction in hydration free energy is larger for charged side chains than for hydrophobic isosteres, ion-pairs at high temperatures may be more stable than hydrophobic interactions (Elcock & McCammon, 1998; Elcock, 1998). The effect of temperature on pK_a values for positively charged residues (see above) is difficult to assess. A decrease in pK_a would result in a higher number of side chains being neutral, which can no longer participate in ion-pair formation; however, the actual pK_a of a residue at high temperature in a large ion-pair network, may be fine-tuned to an unknown extent by its interactions with other amino acids and the existing electrostatic field.

The theoretical considerations described above indicate that large ion-pair networks at high temperatures may play an important role in stabilisation of proteins, and this is reflected in the high abundance of these features in three-dimensional structures of enzymes from hyperthermophiles (Table 3). Recent biochemical and mutagenesis studies are supporting this. Surface-located, isolated ion-pairs do not seem to contribute to the thermostability of *T.maritima* glyceraldehyde-3-phosphate dehydrogenase (Tomschy *et al.*, 1994). However, the removal of a central arginine from an exposed four-residue charge cluster that connects C- and N-terminal parts of this enzyme, leads to accelerated thermal denaturation (Pappenberger *et al.*, 1997). A kinetic role for electrostatic interactions in thermostabilisation was reported for glutamate dehydrogenase (GDH), rubredoxin and methionine aminopeptidase from the hyperthermophile *P.furiosus* (Chiaraluce *et al.*, 1997; Cavagnero *et al.*, 1998; Ogasahara *et al.*, 1998). Analysis of large ion-pair networks in GDH from mesophiles and hyperthermophiles shows fragmentation of these structural features upon going down along the temperature spectrum (Chapter 4 of this thesis). The reconstruction of networks by introducing missing charged residues into less

stable, very homologous enzymes, either or not accompanied with second-site mutations, results in significant increases in half-life values for inactivation and apparent melting temperatures (Vetriani *et al.*, 1998; Rahman *et al.*, 1998). All these examples show that ion-pair networks enhance the kinetic stability of hyperthermostable enzymes. To date, all enzymes that have been demonstrated to contain large ion-pair networks do not unfold reversibly, and therefore the thermodynamic contribution of ion-pair networks in these model systems cannot be analysed and awaits the availability of reversible model systems that still contain these structural features.

4. Protein engineering

Protein engineering is a multidisciplinary approach in which the amino acid sequence of a certain protein is changed in order to create a new variant that possesses improved or novel properties. It is a potent technique to study structure-function and structure-stability relationships in proteins. A myriad of genetic techniques that enable the creation of altered proteins, has been developed, which allows to choose the approach that will best serve the goal of each individual experiment. For the studies described in this thesis, different approaches have been used which will be shortly introduced in the following paragraphs, together with successful examples of their applications in areas related to the research described in this thesis.

Domain swapping

Domain swapping is the exchange of functional domains of homologous enzymes, resulting in the formation of interspecies hybrid enzymes. This approach has successfully been used to confer a certain property of one enzyme on the other, such as thermostability, substrate specificity or cofactor specificity (Mas *et al.*, 1986; Kataoka *et al.*, 1994; Numata *et al.*, 1995). The construction of hybrid enzymes between glyceraldehyde-3-phosphate dehydrogenase from the thermophilic archaeon *Methanothermus fervidus*, and the mesophile *Methanobacterium bryantii*, revealed an important role for a C-terminal fragment that is involved in interdomain contacts, in conferring thermostability (Biro *et al.*, 1990). Often, hybrid enzymes display characteristics that are intermediate between those of the parent enzymes, as in the case of the temperature optima, pH optima and substrate affinities for family 3 glycosyl hydrolases from *Agrobacterium tumefaciens*, *Cellvibrio gilvus* and their hybrids (Singh & Hayashi, 1995).

Site-directed mutagenesis

In order to study structure-function or structure-stability relationships, site-directed mutagenesis is an appropriate tool since usually one or several amino acids are replaced. Using this approach, much progress has been achieved in obtaining insight in stabilisation mechanisms of proteins (Matthews, 1993a; Matthews, 1993b), folding and unfolding mechanisms (Matouschek *et al.*, 1990; Matouschek *et al.*, 1989; Itzhaki *et al.*, 1995) and active site architecture and catalytic mechanisms (Fersht, 1985; van der Oost *et al.*, 1992; Voorhorst *et al.*,

1995). The successful design of site-directed mutations usually requires knowledge of the three-dimensional crystal or solvent structure of the protein under study. Alternatively, a structure may be available for homologous proteins, and information on the structure of the protein under study may be achieved by investigation of multiple sequence alignments and/or homology modelling. Properties of a certain protein may be changed by substitutions to residues that are present in homologous enzymes displaying the desired characteristic. Alternatively, they may be rationally designed based on existing knowledge or hypotheses, eventually aided by the use of computer algorithms. With respect to protein hyperthermostability, these methods have been successfully applied in the construction of hyperthermostable variants of a thermolysin-like protease and the streptococcal protein Gβ1 domain (van den Burg *et al.*, 1998; Malakauskas & Mayo, 1998). The thermolysine-like protease was stabilised by 8 mutations, of which a relative large portion concerned 'rigidifying' substitutions such as Gly to Ala, Ala to Pro, as well as the introduction of a disulfide bridge (van den Burg *et al.*, 1998). An objective computer algorithm was used in the design of the 7-fold mutant variant of the Gβ1 domain, which resulted in a 4.3 kcal/mol enhancement in thermodynamic stability at 50°C by optimising core packing, increasing hydrophobic surface area, more favourable helix dipole interactions and secondary structure improvement (Malakauskas & Mayo, 1998).

Directed evolution using random mutagenesis and *in vitro* recombination

Darwinian evolution involves the repeated operation of three processes: mutation, selection and amplification (Figure 4). Application of these processes in *in vitro* directed evolution, generally starts with the generation of a library of randomly mutated genes. Using appropriate screening or selection procedures, gene products that show improvement with respect to a desired property, are identified. The corresponding genes of these selected variants are amplified, and subjected to further cycles of mutation and screening. In this way beneficial mutations are accumulated throughout as many cycles as one desires (Kuchner & Arnold, 1997).

Random mutations may be introduced using chemical mutagenesis, UV irradiation, mutator strains, poisoned or mismatching nucleotides or error-prone PCR (Kuchner & Arnold, 1997). Error-prone PCR is a relatively recent development. This method makes use of the lack of proofreading capacity of several DNA polymerases; a practical advantage is that the mutagenic rate can be easily controlled (e.g. Cadwell & Joyce, 1994; Leung *et al.*, 1989).

A significant advance in the ability to mimic the natural mechanisms of evolution, is the development of the DNA-shuffling method for *in vitro* DNA recombination (Stemmer, 1994a). Starting from a randomly mutagenised gene pool, or from a collection of homologous genes, this approach will create libraries containing combinations of mutations, or shuffled chimerae of homologous genes (Figure 4). The *in vitro* recombination method rapidly accumulates beneficial mutations, removes deleterious mutations, and is able to evolve desired phenotypes at a much higher evolutionary rate than complementary approaches without DNA shuffling are able to do (Stemmer, 1994b). This approach has successfully been used to engineer a cefotaxime-specific β-lactamase (Stemmer, 1994b), to evolve a β-fucosidase from a β-

galactosidase (Zhang *et al.*, 1997), to create a shuffled library of murine with human interleukin-1β genes (Stemmer, 1994a) and to increase moxalactamase activity by family shuffling (Crameri *et al.*, 1998).

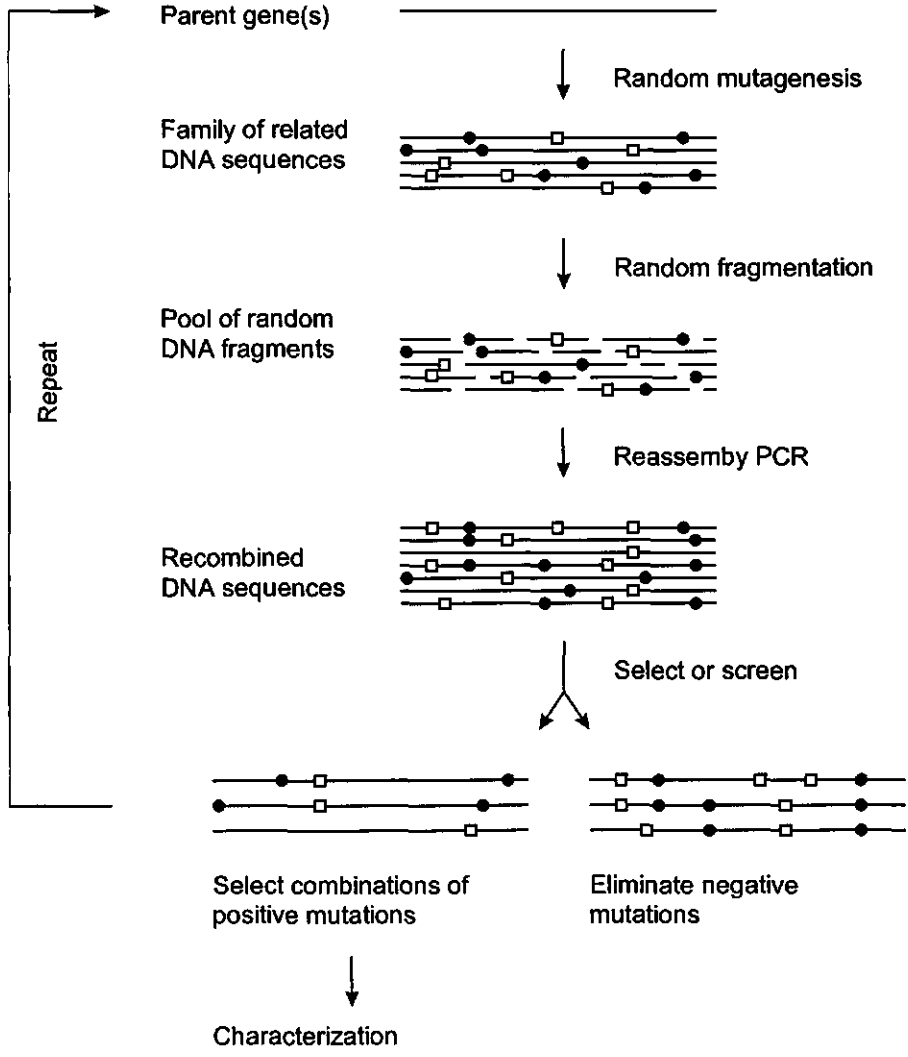


Figure 4: Flow scheme of *in vitro* recombination by DNA shuffling. A single gene or a pool of genes is randomly mutagenised and the resulting mutated genes are fragmented. In a PCR reaction, fragments recombine because they serve as primer for one another and elongate. In this way combinations of the initial mutations are present in this pool. Beneficial mutations accumulate and deleterious mutations are eliminated after multiple rounds of shuffling and selection or screening. Adapted from Zhang *et al.*, 1997.

5. Central metabolism in hyperthermophilic Bacteria and Archaea

Many hyperthermophiles show fermentative growth on a variety of peptides and carbohydrates. Among the best-studied hyperthermophilic Archaea is *Pyrococcus furiosus*, while the best studied representative of the hyperthermophilic Bacteria is *Thermotoga maritima*. These two organisms will be used here to discuss hyperthermophile metabolism as well as to introduce the two model enzymes that have been central in the research described in this thesis, namely β-glucosidase and glutamate dehydrogenase (GDH).

The anaerobic, obligate heterotrophic bacterium *T.maritima* was isolated from geothermally heated sea floor sediments at Vulcano in Italy (Huber *et al.*, 1986). This organism grows optimally at 80°C and ferments peptide mixtures and a variety of oligosaccharides (Schröder *et al.*, 1994). *T.maritima* is able to grow on a much larger variety of sugars than many of the hyperthermophilic Archaea, among which starch, glycogen, maltose, lactose, sucrose as well as the monosaccharides xylose, glucose and galactose (Bragger *et al.*, 1989; Huber et al., 1986). Many oligosaccharide-cleaving enzymes have been found to be present in this organism. These include α-amylase, xylanases, 4-α-glucanotransferase, maltosyl transferase, β-galactosidase, α-galactosidase and β-glucosidase (Liebl *et al.*, 1997; Winterhalter *et al.*, 1995; Liebl *et al.*, 1992; Meissner & Liebl, 1998; Gabelsberger *et al.*, 1993; Liebl *et al.*, 1998). Glucose fermentation proceeds mainly via the classical, unmodified Embden-Meyerhof pathway and for approximately 15% via the classical Entner-Doudoroff pathway (Schröder et al., 1994; Selig *et al.*, 1997). Carbohydrates are being converted to acetate, H_2, CO_2, lactate and alanine (Schröder, 1994; Ravot *et al.*, 1996) (Figure 5).

P.furiosus was isolated as well from geothermally heated marine sediments at Vulcano, Italy (Stetter, 1982; Fiala & Stetter, 1986). The organism grows optimally around 100°C on proteins or protein hydrolysates, on a variety of polysaccharides like starch, glycogen and laminarin, on the disaccharides maltose and cellobiose, and, although with low efficiency, also on the monosaccharide glucose (Fiala & Stetter, 1986; Schäfer and Schönheit, 1992; Kengen *et al.*, 1994; Gueguen *et al.*, 1997; Driskill *et al.*, 1999). For the degradation of polysaccharides, *P.furiosus* excretes several enzymes that cleave α-(1,4)-, α-(1,6)-, β-(1,3)- and β-(1,4)-glycosidic bonds such as α-amylase, pullulanase, amylopullulanase and two endoglucanases (Brown *et al.*, 1990; Brown *et al.*, 1993; Koch *et al*, 1990; Laderman *et al.*, 1993; Gueguen *et al.*, 1997; Bauer *et al.*, 1999). Oligosaccharides containing 1 to 6 sugar units are transported into the cell, probably by specific transporters since genome database sequences reveal the presence of ABC-type and H^+/sugar symporters (Horlacher *et al.*, 1998; Verhees and van der Oost, pers. comm.). Growth on glucose has only recently been reported and requires periodic addition by fed-batch procedure to avoid excessive caramellisation and Maillard reactions (Driskill *et al.*, 1999). Oligosaccharides are converted to monosaccharides by the action of an α-glucosidase and a β-glucosidase (Costantino *et al.*, 1990; Kengen *et al.*, 1993). Glucose fermentation proceeds via a modified Embden-Meyerhof (EM) pathway (Kengen *et al.*, 1994; de Vos *et al.*, 1998) (Figure 5). The main fermentation products are acetate, H_2, CO_2 and alanine (Kengen & Stams, 1994).

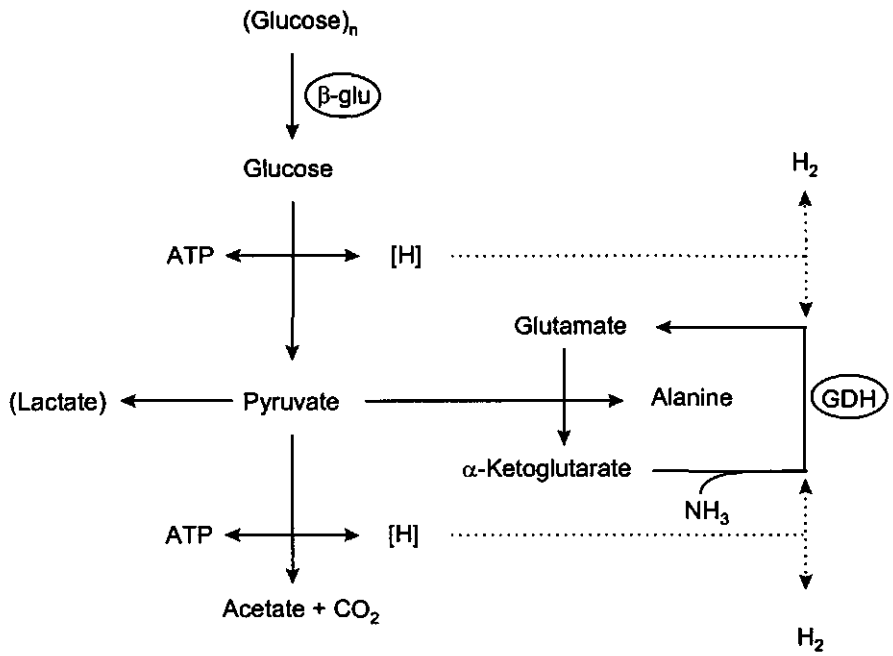


Figure 5: Schematic drawing of proposed sugar metabolism in *T.maritima* and *P.furiosus* with special attention for the role of the two enzymes that have been the research models described in this thesis. β-glu = β-glucosidase, GDH = glutamate dehydrogenase, (lactate) = only for *T.maritima*, [H] = reduction equivalent.

Reduction equivalents that are generated during glucose fermentation in *P.furiosus* are transferred to ferredoxin and subsequently to the nicotinamide cofactor NADP (Ma *et al.*, 1993). *P.furiosus* is then able to dispose these reducing equivalents via three different ways (Figure 5). The bifunctional enzyme sulfhydrogenase, reduces protons to H_2, or elemental sulfur and polysulfide to H_2S (Bryant & Adams, 1989; Ma & Adams, 1994; Ma *et al.*, 1993). Alternatively, the NADPH is used by GDH to fix ammonia by aminating α-ketoglutarate to glutamate (Consalvi *et al.*, 1993; Kengen *et al.*, 1994). The amino-group is further passed on to pyruvate via an alanine aminotransferase, and alanine is excreted into the medium. The alanine/acetate ratio was found to depend on the hydrogen partial pressure and the presence of S^0 in the medium, indicating that *P.furiosus* is able to shift its metabolism, depending on the redox potential of the available electron acceptor (Kengen & Stams, 1994). Alanine formation has been described for many other organisms, among which other members from the *Thermococcales* (*T.profundus, T.celer* and *T.stetteri* ; Kobayashi *et al.*, 1995; Kengen *et al.*, 1996), hyperthermophilic bacteria of the order *Thermotogales* including *T.maritima* (Ravot *et al.*, 1996) as well as the early branching eucaryote *Giardia lamblia* (Edwards *et al.*, 1989). The operation of the GDH shuttle is proposed here for *T.maritima*, based on the reported alanine formation and the observation that *T.maritima* GDH is NADH-

dependent instead of NADPH-dependent (Kort *et al.*, 1997), and hence can function as an electron sink for reducing equivalents that are transferred to NAD during glucose fermentation in this organism. It has been proposed that maintenance of the intracellular redox balance by excretion of alanine is an ancestral metabolic characteristic (Ravot *et al.*, 1996).

6. Aims and outline of the thesis

This thesis describes studies on the molecular determinants of stability and activity of hyperthermostable enzymes. The aims of these studies were (i) to develop hypotheses concerning enzyme thermostabilisation strategies, (ii) to validate these and other emerging hypotheses, and (iii) to identify and characterise factors that determine enzyme catalysis and substrate recognition at low and high temperatures. The basic concepts of protein stability have been presented in this introductory chapter, as well as the current knowledge on hyperthermostability, thermoactivity and factors that are governing these characteristics.

Because the hyperthermophilic phenotype is found both in Archaea and in Bacteria, and may have distinct phylogenetic histories in these organisms, we decided to study this characteristic in representatives of both domains. In this way we would be able to discriminate between general and species- or domain-specific variations. We selected *P.furiosus* as model hyperthermophilic archaeon, since it is well studied and genome sequences from *Pyrococcus furiosus* and two other *Pyrococcus* species either have been or shortly will be elucidated (Kawarabayasi *et al.*, 1998; Heilig *et al.*, 1998; TIGR, 1999). Many enzymes from *P.furiosus* have been isolated, characterised and for several of these, crystal structures are available. Furthermore, corresponding genes have been functionally overexpressed in *E.coli*, enabling to address the above-mentioned research aims by a protein engineering approach. *T.maritima* is the best-studied hyperthermophilic bacterium. Also from this hyperthermophile many enzymes have been isolated and characterised, several three-dimensional structures of these enzymes are available, corresponding genes have been overexpressed in *E.coli*, and its genome sequence is available (Nelson *et al.*, 1999). *T.maritima* was therefore chosen as representative of the hyperthermophilic Bacteria.

We focused on two key metabolic enzymes, a β-glucosidase (CelB) which is responsible for supplying glucose monomers into the central metabolic pathway, and GDH which plays an essential role in the newly discovered pathway for removal of reduction equivalents (Figure 5). Both enzymes are very well studied, not only in *P.furiosus* and *T.maritima*, but also in other hyperthermophilic and mesophilic Archaea, Bacteria, as well as Eucarya. Three-dimensional structures have been elucidated of GDH from both model organisms (Yip *et al.*, 1995; Knapp *et al.*, 1997) as well as a three-dimensional model for *P.furiosus* CelB (see below). This allows for extensive comparative analyses of these enzymes regarding the molecular determinants of stability and activity at high temperature.

Glutamate dehydrogenase (GDH) is introduced in Chapter 2 of this thesis. A full description of this model enzyme is given, including its phylogeny, three-dimensional structure, and a comparison of homologous GDHs in many hyperthermophiles.

The subunits of the hexameric GDH are composed of a substrate binding domain and a cofactor binding domain, which are separated by a cleft in which the active site is located. The role of each domain in conferring thermoactivity and thermostability was studied by exchanging them between the GDHs of the hyperthermophilic archaeon *P.furiosus* and the mesophilic bacterium *C.difficile*. The construction of the hybrid GDHs and biochemical characterisation is described in Chapter 3.

GDH from *P.furiosus* is the first enzyme in which large ion-pair networks were identified (Yip *et al.*, 1995). In order to assess whether this is a more general stabilising strategy which is not only employed in GDH from *P.furiosus*, but also in enzymes in other hyperthermophilic Archaea and even Bacteria, an homology-based modelling study was carried out using GDH amino acid sequences derived from species spanning a wide spectrum of optimal growth temperatures. This study is described in Chapter 4.

The results obtained by the structural and homology modelling studies described above, suggest that ion-pair networks are major determinants of enzyme hyperthermostability. If this were the case, it would be possible to rationally engineer ion-pair networks into enzymes that lack these features and display a much lower thermostability. In order to study the role of large ion-pair networks and to test their universal applicability, we attempted to engineer the networks that are found in the hyperthermostable *P.furiosus* GDH into the less thermostable GDH from *T.maritima*. The successful introduction of a six-residue network into the flexible hinge region, and a 16-residue network into the subunit interface of *T.maritima* GDH, as well as biochemical and structural characterisation of the wild-type and the mutant enzymes, is described in Chapters 5 and 6, respectively.

The second model enzyme, β-glucosidase (CelB) from *P.furiosus*, is described in Chapter 7. In addition, optimised heterologous gene expression systems and the development of a directed evolution procedure for CelB are presented. The random mutagenesis procedure is validated by the isolation and characterisation of a CelB variant that contains an amino acid substitution in the active site.

A three-dimensional model for CelB, based on 3.3 Å X-ray diffraction data, is described in Chapter 8. This model was used, together with that of the 6-phospho-β-galactosidase from *Lactococcus lactis*, to design mutations that should increase the activity of CelB towards phosphorylated sugars. Characterisation of these mutant CelB enzymes is presented.

Low-temperature activity and substrate specificity of CelB was studied in Chapter 9, using the random CelB library described in Chapter 7. Screening of this library at room temperature on an artificial substrate resulted in the isolation and characterisation of mutants with increased activity on this substrate and the natural substrate cellobiose.

Three-dimensional structure analysis showed that also in family I of glycosyl hydrolases, ion-pair networks may play an important role in

determining thermostability (Aguilar et al., 1997). This hypothesis was investigated and described in Chapter 10, and involved biochemical analysis of wild-type CelB and mutants with deletions in the C-terminal region, supposedly involved in ion-pair network formation.

Chapter 11 contains a summary of obtained results and a comparison with literature data that deal with similar and complementary aspects of enzyme thermostability and catalytic function at low and high temperatures. This chapter concludes with a careful evaluation of the conclusions against the original aims as described here.

References

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the beta-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Amend, J. P. & Shock, E. L. (1998). Energetics of amino acid synthesis in hydrothermal ecosystems. *Science* **281**, 1659-1662.

Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. (1990). pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403-2408.

Antranikian, G., Konings, W.N. & de Vos, W.M. (1996). First International Congress on Extremophiles. *FEMS Microbiol. Rev.* **18**.

Antranikian, G. (1998). Biotechnology of Extremophiles. *Adv. Biochem. Engng. Biotechnol.* **61**.

Arakawa, T. & Timasheff, S. N. (1985). The stabilisation of proteins by osmolytes. *Biophys. J.* **47**, 411-414.

Aravind, L., Tatusov, R. L., Wolf, Y. I., Walker, D. R. & Koonin, E. V. (1998). Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. *Trends Genet.* **14**, 442-444.

Arnone, M. I., Birolo, L., Pascarella, S., Cubellis, M. V., Bossa, F., Sannia, G. & Marino, G. (1997). Stability of aspartate aminotransferase from *Sulfolobus solfataricus*. *Prot. Engng.* **10**, 237-248.

Backmann, J., Schafer, G., Wyns, L. & Bonisch, H. (1998). Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* **284**, 817-833.

Barlow, D. J. & Thornton, J. M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867-885.

Bauer, M. W., Driskill, L. E., Callen, W., Snead, M. A., Mathur, E. J. & Kelly, R. M. (1999). An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyses beta-1,4 bonds in mixed-linkage (1->3),(1->4)-beta-D-glucans and cellulose. *J. Bacteriol.* **181**, 284-90.

Becktel, W. J. & Schellman, J. A. (1987). Protein stability curves. *Biopolymers* **26**, 1859-1877.

Benachenhou-Lahfa, N., Labedan, B. & Forterre, P. (1994). PCR-mediated cloning and sequencing of the gene encoding glutamate dehydrogenase from the archaeon *Sulfolobus shibatae*: identification of putative amino-acid signatures for extremophilic adaptation. *Gene* **140**, 17-24.

Biro, J., Fabry, S., Dietmaier, W., Bogedain, C. & Hensel, R. (1990). Engineering thermostability in archaebacterial glyceraldehyde-3-phosphate dehydrogenase. Hints for the important role of interdomain contacts in stabilising protein conformation. *FEBS Lett.* **275**, 130-134.

Blochl, E., Rachel, R., Burggraf, S., Hafenbradl, D., Jannasch, H. W. & Stetter, K. O. (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of Archaea, extending the upper temperature limit for life to 113 degrees C. *Extremophiles* 1, 14-21.

Bohm, G. & Jaenicke, R. (1994). Relevance of sequence statistics for the properties of extremophilic proteins. *Int. J. Pept. Prot. Res.* **43**, 97-106.

Bragger, J. M., Daniel, R. M., Coolbear, T., Morgan, H. W. (1989). Very stable enzymes from extremely thermophilic archaebacteria and eubacteria. *Appl. Microbiol. Biotechnol.* **31**, 556-561.

Britton, K. L., Baker, P. J., Borges, K. M., Engel, P. C., Pasquo, A., Rice, D. W., Robb, F. T., Scandurra, R., Stillman, T. J. & Yip, K. S. (1995). Insights into thermal stability from a comparison of the glutamate dehydrogenases from *Pyrococcus furiosus* and *Thermococcus litoralis*. *Eur. J. Biochem.* **229**, 688-695.

Brown, J. R. & Doolittle, W. F. (1997). Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Molecul. Biol. Rev.* **61**, 456.

Brown, S. H., Costantino, H.R., Kelly, R.M. (1990). Characterisation of amylolytic enzyme activities associated with the hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Appl. Environ. Microbiol.* **56**, 1985-1991.

Brown, S. H., Kelly, R. M. (1993). Characterisation of amylolytic enzymes, having both α-1,4 and α-1,6 hydrolytic activity, from the thermophilic archaea *Pyrococcus furiosus* and *Thermococcus litoralis*. *Appl. Environ. Microbiol.* **59**, 2614-2621.

Bryant, F. O. & Adams, M. W. (1989). Characterisation of hydrogenase from the hyperthermophilic archaebacterium, *Pyrococcus furiosus*. *J. Biol. Chem.* **264**, 5070-5079.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M., Weidman, J. F., Fuhrmann, J. L., Venter, J. C. & et al. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058-1073.

Burg, B. van den, Vriend, G., Veltman, O. R., Venema, G. & Eijsink, V. G. H. (1998). Engineering an enzyme to resist boiling. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2056-2060.

Cadwell, R. C. & Joyce, G. F. (1994). Mutagenic PCR. *PCR Meth. Appl.* **3**, 136-140.

Carpenter, J. F., Arakawa, T. & Crowe, J. H. (1992). Interactions of stabilising additives with proteins during freeze-thawing and freeze-drying. *Dev. Biol. Stand.* **74**, 225-239.

Cavagnero, S., Debe, D. A., Zhou, Z. H., Adams, M. W. W. & Chan, S. L. (1998). Kinetic role of electrostatic interactions in the unfolding of hyperthermophilic and mesophilic rubredoxins. *Biochemistry* **37**, 3369-3376.

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. & Rees, D. C. (1995). Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* **267**, 1463-1469.

Chiaraluce, R., Schwerdtfeger, R. M., Scandurra, R., Antranikian, G. & Consalvi, V. (1997). Acid-induced disassembly of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* occurs below pH 2.0. *Eur. J. Biochem.* **247**, 224-30.

Consalvi, V., Chiaraluce, R., Politi, L., Pasquo, A., De Rosa, M. & Scandurra, R. (1993). Glutamate dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*: studies on thermal and guanidine-dependent inactivation. *Biochim. Biophys. Acta* **1202**, 207-15.

Costa, M. S. da., Santos, H., Galinski, E.A. (1998). An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv. Biochem. Engng. Biotechnol.* **61**, 117-153.

Costantino, H. R., Brown, S. H. & Kelly, R. M. (1990). Purification and characterisation of an alpha-glucosidase from a hyperthermophilic archaebacterium, *Pyrococcus furiosus*, exhibiting a temperature optimum of 105 to 115 degrees C. *J. Bacteriol.* **172**, 3654-3660.

Crameri, A., Raillard, S. A., Bermudez, E. & Stemmer, W. P. C. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288-291.

Danson, M. J. & Hough, D. W. (1998). Structure, function and stability of enzymes from the Archaea. *Trends Microbiol.* **6**, 307-314.

Danson, M. J., Hough, D. W., Russell, R. J., Taylor, G. L. & Pearl, L. (1996). Enzyme thermostability and thermoactivity. *Prot. Engng.* **9**, 629-630.

Dao-pin, S., Nicholson, H., Baase, W. A., Zhang, X. J., Wozniak, J. A. & Matthews, B. W. (1991). Structural and genetic analysis of electrostatic and other interactions in bacteriophage T4 lysozyme. *CIBA Found. Symp.* **161**, 52-62.

Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. & Swanson, R. V. (1998). The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353-358.

Dedecker, B. S., Obrien, R., Fleming, P. J., Geiger, J. H., Jackson, S. P. & Sigler, P. B. (1996). The crystal structure of a hyperthermophilic archaeal TATA box binding protein. *J. Mol. Biol.* **264**, 1072-1084.

Doolittle, W. F. & Brown, J. R. (1994). Tempo, mode, the progenote, and the universal root. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6721-6728.

Driskill L.E., Bauer, M. W., Kelly, R. M. (1999). Relationship between glycosyl hydrolase inventory and growth physiology of the hyperthermophile *Pyrococcus furiosus* on carbohydrate-based media. *Appl. Environ. Microbiol.* **65**, 893-897.

Edwards, M. R., Gilroy, F. V., Jimenez, B. M. & O'Sullivan, W. J. (1989). Alanine is a major end product of metabolism by *Giardia lamblia*: a proton nuclear magnetic resonance study. *Mol. Biochem. Parasitol.* 37, 19-26.

Eggen, R. I. L., Geerling, A. C. M., Waldkötter, K., Antranikian, G. & de Vos, W. M. (1993). The glutamate dehydrogenase-encoding gene of the hyperthermophilic archaeon *Pyrococcus furiosus*: sequence, transcription and analysis of the deduced amino acid sequence. *Gene* **132**, 143-148.

Eggen, R. I. L., Geerling, A. C. M., Voorhorst, W. G. B., Kort, R. & de Vos, W. M. (1994). Molecular and comparative analysis of the hyperthermostable *Pyrococcus furiosus* glutamate dehydrogenase and its gene. *Biocatalysis* **11**, 131-141.

Elcock, A. H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489-502.

Elcock, A. H. & McCammon, J. A. (1998). Electrostatic contributions to the stability of halophilic proteins. *J. Mol. Biol.* **280**, 731-748.

Fagain, C. O. (1997). Stabilising protein function. Springer Verlag and Landes Bioscience, Berlin Germany.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D.M., Wilkinson, A. J., Blow, D.M., Brick, P., Carter, P., Waye, Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* **314**, 235-238.

Fiala, G. & Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* **145**, 56-61.

Forterre, P., Bergerat, A., Gadelle, D. (1994). Evolution of DNA topoisomerases and DNA polymerases: a perspective from Archaea. *System. Appl. Microbiol.* **16**, 746-758.

Forterre, P. (1995). Thermoreduction, a hypothesis for the origin of prokaryotes. *C. R. Acad. Sci.* **318**, 415-422.

Forterre, P. (1996). A hot topic: the origin of hyperthermophiles. *Cell* **85**, 789-792.

Forterre, P. (1997). Archaea: what can we learn from their sequences? *Curr. Opin. Genet. Dev.* **7**, 764-770.

Forterre, P. (1998). Were our ancestors actually hyperthermophiles ? Viewpoint of a devil's advocate. *In: Thermophiles, the keys to molecular evolution and the origin of life ? Wiegel, J. & Adams, M. W. W., eds.* Taylor & Francis Inc., Philadelphia, USA.

Forterre, P., Benachenhou-Lafha, N. & Labedan, B. (1993). Universal tree of life. *Nature* **362**, 795.

Gabelsberger, J., Liebl, W., Schleifer, K.-H. (1993). Purification and properties of recombinant beta-glucosidase of the hyperthermophilic bacterium *Thermotoga maritima*. *Appl. Microbiol. Biotechnol.* **40**, 44-52.

Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, T., Oshima, T. & et al. (1989). Evolution of the vacuolar H^+-ATPase: implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6661-6665.

Gueguen, Y., Voorhorst, W. G. B., van der Oost, J. & de Vos, W. M. (1997). Molecular and biochemical characterisation of an endo-beta-1,3-glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **272**, 31258-31264.

Gupta, R. S. (1998). Protein phylogenies and signature sequences: A reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. *Microbiol. Mol. Biol. Rev.* **62**, 1435-1491.

Haney, P. J., Badger, J.H., Buldak, G.L., Woese, C.R. and Olsen, G.J. (1998). Large scale statistical analyses identify amino acid replacements repeatedly utilised in the adaptations between closely related mesophilic and thermophilic proteins. *International Conference on Thermophiles 1998, Brest, France.*

Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. *Nature* **381**, 571-580.

Heilig, R., Barbe, V., Duprat, S., Robert, C., Vico, V., Petit, J., Wincker, P., Brottier, P., Lecompte, O., Querellou, J., Erauso, G., Prieur, D., Pocs, O., Thierry, J., Forterre, P., Weissenbach, J., Zivanovich, Y., Saurin, W. (1998). Complete genomic sequence of the hyperthermophilic archaeon *Pyrococcus abyssi*. *International Conference on Thermophiles, Brest, France.*

Hendsch, Z. S. & Tidor, B. (1994). Do salt bridges stabilise proteins? A continuum electrostatic analysis. *Prot. Sci.***3**, 211-226.

Hennig, M., Darimont, B., Sterner, R., Kirschner, K. & Jansonius, J. N. (1995). 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure* **3**, 1295-306.

Hennig, M., Sterner, R., Kirschner, K. & Jansonius, J. N. (1997). Crystal structure at 2.0 Angstrom resolution of phosphoribosyl anthranilate isomerase from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability. *Biochemistry* **36**, 6009-6016.

Hiller, R., Zhou, Z. H., Adams, M. W. W. & Englander, S. W. (1997). Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200 degrees C. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11329-11332.

Holden, J. F. & Baross, J. A. (1993). Enhanced thermotolerance and temperature-induced changes in protein composition in the hyperthermophilic archaeon ES4. *J. Bacteriol.* **175**, 2839-2843.

Horikoshi, K. (1997). *Extremophiles.*

Horlacher, R., Xavier, K. B., Santos, H., DiRuggiero, J., Kossmann, M. & Boos, W. (1998). Archaeal binding protein-dependent ABC transporter: molecular and biochemical analysis of the trehalose/maltose transport system of the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Bacteriol.* **180**, 680-689.

Horovitz, A., Serrano, L., Avron, B., Bycroft, M. & Fersht, A. R. (1990). Strength and cooperativity of contributions of surface salt bridges to protein stability. *J. Mol. Biol.* **216**, 1031-1044.

Huber, R., Langworthy, T. A., König, H., Thomm, M., Woese, C. W., Sleytr, U. B. & Stetter, K. O. (1986). *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch. Microbiol.***144**, 324-333.

Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. (1995). The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254**, 260-288.

Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S. & Miyata, T. (1989). Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9355-9359.

Jaenicke, R. (1991). Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* **202**, 715-728.

Jaenicke, R. (1996). Glyceraldehyde-3-phosphate dehydrogenase from *Thermotoga maritima*: strategies of protein stabilisation. *FEMS Microbiol. Rev.* **18**, 215-224.

Jaenicke, R. & Bohm, G. (1998). The stability of proteins in extreme environments. *Curr. Opinion Struct. Biol.* **8**, 738-748.

Kataoka, K., Takada, H., Tanizawa, K., Yoshimura, T., Esaki, N., Ohshima, T. & Soda, K. (1994). Construction and characterisation of chimeric enzyme consisting of an amino-terminal domain of phenylalanine dehydrogenase and a carboxy-terminal domain of leucine dehydrogenase. *J. Biochem. Tokyo* **116**, 931-936.

Kawarabayasi, Y. S. M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, O. R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Kikuchi, H. et al. (1998). Complete sequence and gene organisation of the genome of a hyperthermophilic archaebacterium, *Pyrococcus horikoshii* OT3. *DNA Res.* **5**, 55-76.

Kengen, S. W. M., de Bok, F. A., van Loo, N. D., Dijkema, C., Stams, A. J. & de Vos, W. M. (1994). Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. *J. Biol. Chem.* **269**, 17537-17541.

Kengen, S. W. M., Luesink, E. J., Stams, A. J. M. & Zehnder, A. J. (1993). Purification and characterisation of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**, 305-312.

Kengen, S. W. M. & Stams, A. J. M. (1994). Formation of L-alanine as a reduced end product in carbohydrate fermentation by the hyperthermophilic archaeon *Pyrococcus furiosus*. *Archiv. Microbiol.* **161**, 168-175.

Kengen, S. W. M., Stams, A. J. M. & de Vos, W. M. (1996). Sugar metabolism of hyperthermophiles. *FEMS Microbiol. Rev.* **18**, 119-137.

Kiefer, J. R., Mao, C., Hansen, C. J., Basehore, S. L., Hogrefe, H. H., Braman, J. C. & Beese, L. S. (1997). Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution. *Structure* **5**, 95-108.

Klenk, H. P., Palm, P., Zillig, W. (1994). DNA-dependent RNA polymerases as phylogenetic marker molecules. *System. Appl. Microbiol.* **16**, 638-647.

Klibanov, A. M. (1983a). Approaches to enzyme stabilisation. *Biochem. Soc. Trans.* **11**, 19-20.

Klibanov, A. M. (1983b). Stabilisation of enzymes against thermal inactivation. *Adv. Appl. Microbiol.* **29**, 1-28.

Klibanov, A. M. & Mozhaev, V. V. (1978). On the mechanism of irreversible thermoinactivation of enzymes and possibilities for reactivation of "irreversibly" inactivated enzymes. *Biochem. Biophys. Res. Commun.* **83**, 1012-1017.

Klump, H., Di Ruggiero, J., Kessel, M., Park, J. B., Adams, M. W. & Robb, F. T. (1992). Glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*. Thermal denaturation and activation. *J. Biol. Chem.* **267**, 22681-22685.

Knapp, S. (1996). On the stability of proteins from hyperthermophiles: a structural and thermodynamic study. *PhD Thesis*, Karolinska Institute, Stockholm, Sweden.

Knapp, S., de Vos, W. M., Rice, D. & Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3.0 Å resolution. *J. Mol. Biol.* **267**, 916-932.

Knapp, S., Kardinahl, S., Hellgren, N., Tibbelin, G., Schäfer, G. & Ladenstein, R. (1999). Refined crystal structure of a superoxide dismutase from the hyperthermophilic archaeon *Sulfolobus acidocaldarius* at 2.2 angstrom resolution. *J. Mol. Biol.* **285**, 689-702.

Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B. & Ladenstein, R. (1996). Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus*. *J. Mol. Biol.* **264**, 1132-1144.

Kobayashi, T., Higuchi, S., Kimura, K., Kudo, T. & Horikoshi, K. (1995). Properties of glutamate dehydrogenase and its involvement in alanine production in a hyperthermophilic archaeon, *Thermococcus profundus*. *J. Biochem. Tokyo* **118**, 587-592.

Koch, R., Zablowski, P., Spreinat, A., Antranikian, G. (1990). Extremely thermostable amylolytic enzyme from the archaebacterium *Pyrococcus furiosus*. *FEMS Microbiol. Lett.* **71**, 21-26.

Korndörfer, I., Steipe, B., Huber, R., Tomschy, A. & Jaenicke, R. (1995). The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J. Mol. Biol.* **246**, 511-521.

Korolev, S., Nayal, M., Barnes, W. M., Di Cera, E. & Waksman, G. (1995). Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5-Å resolution: structural basis for thermostability. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9264-9268.

Kort, R., Liebl, W., Labedan, B., Forterre, P., Eggen, R. I. L. & de Vos, W. M. (1997). Glutamate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: molecular characterisation and phylogenetic implications. *Extremophiles* **1**, 52-60.

Kuchner, O. & Arnold, F. H. (1997). Directed evolution of enzyme catalysts. *Trends Biotechnol.* **15**, 523-530.

Laderman, K. A., Davis, B. R., Krutzsch, H. C., Lewis, M. S., Griko, Y. V., Privalov, P. L. & Anfinsen, C. B. (1993). The purification and characterisation of an extremely thermostable alpha-amylase from the hyperthermophilic archaebacterium *Pyrococcus furiosus*. *J. Biol. Chem.* **268**, 24394-24401.

Langworthy, T. A., Pond, J. L. (1986). Archaebacterial ether lipids and chemotaxonomy. *In: Archaebacteria '85. (O. Kandler and W. Zillig, Eds.) Gustav Fischer, New York.*, 278-285.

Leung, D. W., Chen, E., Goeddel, D.V. (1989). A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1**, 11-15.

Lewin, B. (1997). *Cell* **89**.

Li, W. T., Grayling, R. A., Sandman, K., Edmondson, S., Shriver, J. W. & Reeve, J. N. (1998). Thermodynamic stability of archaeal histones. *Biochemistry* **37**, 10563-10572.

Liebl, W., Feil, R., Gabelsberger, J., Kellermann, J. & Schleifer, K. H. (1992). Purification and characterisation of a novel thermostable 4-alpha-glucanotransferase of *Thermotoga maritima* cloned in *Escherichia coli*. *Eur. J. Biochem.* **207**, 81-88.

Liebl, W., Stemplinger, I. & Ruile, P. (1997). Properties and gene structure of the *Thermotoga maritima* alpha-amylase AmyA, a putative lipoprotein of a hyperthermophilic bacterium. *J. Bacteriol.* **179**, 941-948.

Liebl, W., Wagner, B. & Schellhase, J. (1998). Properties of an alpha-galactosidase, and structure of its gene galA, within an alpha- and beta-galactoside utilisation gene cluster of the hyperthermophilic bacterium *Thermotoga maritima*. *Syst. Appl. Microbiol.* **21**, 1-11.

Lim, J. H., Yu, Y. G., Han, Y. S., Cho, S., Ahn, B. Y., Kim, S. H. & Cho, Y. (1997). The crystal structure of an Fe-superoxide dismutase from the hyperthermophile *Aquifex pyrophilus* at 1.9 Å resolution: structural basis for thermostability. *J. Mol. Biol.* **270**, 259-274.

Lobry, J. R. (1997). Influence of genomic G - C content on average amino-acid composition of proteins from 59 bacterial species. *Gene* **205**, 309-316.

Ma, K. & Adams, M. W. (1994). Sulfide dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*: a new multifunctional enzyme involved in the reduction of elemental sulfur. *J. Bacteriol.* **176**, 6509-6517.

Ma, K., Schicho, R. N., Kelly, R. M. & Adams, M. W. (1993). Hydrogenase of the hyperthermophile *Pyrococcus furiosus* is an elemental sulfur reductase or sulfhydrogenase: evidence for a sulfur-reducing hydrogenase ancestor. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5341-5344.

Macedo-Ribeiro, S., Darimont, B. & Sterner, R. (1997). Structural features correlated with the extreme thermostability of 1[4Fe-4S] ferredoxin from the hyperthermophilic bacterium *Thermotoga maritima*. *Biol. Chem.* **378**, 331-336.

Malakauskas, S. M. & Mayo, S. L. (1998). Design, structure and stability of a hyperthermophilic protein variant. *Nat. Struct. Biol.* **5**, 470-475.

Marqusee, S. & Baldwin, R. L. (1987). Helix stabilization by $Glu^{-}...Lys^{+}$ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci. U.S.A* **84**, 8898-8902.

Martin, W. & Muller, M. (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37-41.

Mas, M. T., Chen, C. Y., Hitzeman, R. A. & Riggs, A. D. (1986). Active human-yeast chimeric phosphoglycerate kinases engineered by domain interchange. *Science* **233**, 788-790.

Matouschek, A., Kellis, J. T. Jr., Serrano, L., Bycroft, M. & Fersht, A. R. (1990). Transient folding intermediates characterised by protein engineering. *Nature* **346**, 440-445.

Matouschek, A., Kellis, J. T. Jr., Serrano, L. & Fersht, A. R. (1989). Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122-126.

Matsumura, M., Signon, G., Matthews, B. W. (1989). Substantial increase of protein stability by multiple disulfide bonds. *Nature* **342**, 291-293.

Matthews, B. W. (1993a). Structural and genetic analysis of protein structure and folding. *Curr. Opinion Struc. Biol.* **3**, 589-593.

Matthews, B. W. (1993b). Structural and genetic analysis of protein structure. *Annu. Rev. Biochem.* **62**, 139-160.

Meissner, H. & Liebl, W. (1998). *Thermotoga maritima* maltosyltransferase, a novel type of maltodextrin glycosyltransferase acting on starch and malto-oligosaccharides. *Eur. J. Biochem.* **258**, 1050-1058.

Menendez-Arias, L. & Argos, P. (1989). Engineering protein thermal stability. Sequence statistics point to residue substitutions in α-helices. *J. Mol. Biol.* **206**, 397-406.

Miller, S. L., Lazcano, A. (1998). Facing up to chemical realities: Life did not begin at the growth temperature of hyperthermophiles. *In: Thermophiles, the keys to molecular evolution and the origin of life ?* Wiegel, J. and Adams, M.W.W, eds. Taylor and Francis, Philadelphia, USA., 127-131.

Minuth, T., Frey, G., Lindner, P., Rachel, R., Stetter, K. O. & Jaenicke, R. (1998). Recombinant homo- and hetero-oligomers of an ultrastable chaperonin from the archaeon *Pyrodictium occultum* show chaperone activity in vitro. *Eur. J. Biochem.* **258**, 837-845.

Mozhaev, V. V. & Martinek, K. (1982). Inactivation and reactivation of proteins (enzymes). *Mol. Biol. Mosk.* **16**, 676-694.

Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A., McDonald, L., Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M., Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A., Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D., Eisen, J.A., Fraser, C.M., et al. (1999). Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323-329.

Numata, K., Muro, M., Akutsu, N., Nosoh, Y., Yamagishi, A. & Oshima, T. (1995). Thermal stability of chimeric isopropylmalate dehydrogenase genes constructed from a thermophile and a mesophile. *Prot. Engng.* **8**, 39-43.

Ogasahara, K., Lapshina, E. A., Sakai, M., Izu, Y., Tsunasawa, S., Kato, I. & Yutani, K. (1998). Electrostatic stabilisation in methionine aminopeptidase from hyperthermophile *Pyrococcus furiosus*. *Biochemistry* **37**, 5939-5946.

van der Oost, J., Schut, G., Kengen, S. W. M. Hagen, W. R., Thomm, M., de Vos, W. M. (1998). The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycolytic regulation. *J. Biol. Chem.* **273**, 28149-28154.

van der Oost, J., Lappalainen, P., Musacchio, A., Warne, A., Lemieux, L., Rumbley, J., Gennis, R. B., Aasa, R., Pascher, T., Malmstrom, B. G. & et al. (1992). Restoration of a lost metal-binding site: construction of two different copper sites into a subunit of the *E. coli* cytochrome o quinol oxidase complex. *EMBO J.* **11**, 3209-3217.

Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676-683.

Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R. & Schmid, F. X. (1998). Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nature Struct. Biol.* **5**, 229-235.

Perutz, M. F. (1978). Electrostatic effects in proteins. *Science* **201**, 1187-1191.

Perutz, M. F. & Raidt, H. (1975). Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature* **255**, 256-259.

Pfeil, W. (1986). Unfolding of proteins. In: Thermodynamic data for Biochemistry and Biotechnology (Hinz, H.-J., ed). *Springer Verlag, Berlin, Heidelberg, New York, Tokyo*, 349-376.

Pfeil, W., Gesierich, U., Kleemann, G. R. & Sterner, R. (1997). Ferredoxin from the hyperthermophile *Thermotoga maritima* is stable beyond the boiling point of water. *J. Mol. Biol.* **272**, 591-596.

Rahman, R., Fujiwara, S., Nakamura, H., Takagi, M. & Imanaka, T. (1998). Ion pairs involved in maintaining a thermostable structure of glutamate dehydrogenase from a hyperthermophilic archaeon. *Biochem. Biophys. Res. Comm.* **248**, 920-926.

Ramos, A., Raven, N. D. H., Sharp, R. J., Bartolucci, S., Rossi, M., Cannio, R., Lebbink, J., van der Oost, J., de Vos, W. M. & Santos, H. (1997). Stabilisation of enzymes against thermal stress and freeze drying by mannosylglycerate. *Appl. Environ. Microbiol.* **63**, 4020-4025.

Ravot, G., Ollivier, B., Fardeau, M. L., Patel, B. K., Andrews, K. T., Magot, M. & Garcia, J. L. (1996). L-alanine production from glucose fermentation by hyperthermophilic members of the domains Bacteria and Archaea: a remnant of an ancestral metabolism? *Appl. Environ. Microbiol.* **62**, 2657-2659.

Ren, B., Tibbelin, G., de Pascale, D., Rossi, M., Bartolucci, S. & Ladenstein, R. (1998). A protein disulfide oxidoreductase from the archaeon *Pyrococcus furiosus* contains two thioredoxin fold units. *Nature Struct. Biol.* 5, 602-611.

Risse, B., Stempfer, G., Rudolph, R., Schumacher, G. & Jaenicke, R. (1992). Characterisation of the stabilising effect of point mutations of pyruvate oxidase from *Lactobacillus plantarum*: protection of the native state by modulating coenzyme binding and subunit interaction. *Prot. Sci.* **1**, 1710-1718.

Russell, R. J. M., Ferguson, J. M. C., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). The Crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus Furiosus* at 1.9 Angstrom resolution. *Biochemistry* **36**, 9983-9994.

Schäfer, T. & Schönheit., P. (1992). Maltose fermentation to acetate, CO_2 and H_2 in the anaerobic hyperthermophilic archaeon *Pyrococcus furiosus*: evidence for the operation of a novel sugar fermentation pathway. *Arch. Microbiol.* **158**, 188-202.

Schröder, C., Selig, M., Schonheit, P. (1994). Glucose fermentation to acetate, CO_2 and H_2 in the anaerobic hyperthermophilic eubacterium *Thermotoga maritima*: involvement of the Emden-Meyerhof pathway. *Arch. Microbiol.* **161**, 460-470.

Selig, M., Xavier, K. B., Santos, H. & Schonheit, P. (1997). Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic Archaea and the bacterium *Thermotoga*. *Arch. Microbiol.* **167**, 217-232.

Singh, A. & Hayashi, K. (1995). Construction of chimeric beta-glucosidases with improved enzymatic properties. *J. Biol. Chem.* **270**, 21928-21933.

Stemmer, W. P. (1994a). DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10747-10751.

Stemmer, W. P. (1994b). Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-391.

Stetter, K. O. (1982). Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105°C. *Nature* **300**, 258-260.

Stetter, K. O. (1994). The lesson of Archaebacteria. *Nobel symposium, New York, Columbia University Press*, 143-151.

Stetter, K. O. (1996). Hyperthermophilic prokaryotes. *FEMS Microbiol. Rev.* **18**, 149-158.

Sueoka, N. (1961). Correlation between base composition of deoxyribonucleic acid and amino acid composition of protein. *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1141-1149.

TIGR. (1999). http://www.tigr.org/tdb/tdb.html. *TIGR Database (TDB).*

Tomschy, A., Bohm, G. & Jaenicke, R. (1994). The effect of ion pairs on the thermal stability of D-glyceraldehyde 3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *Prot. Engng.* **7**, 1471-1478.

Trent, J. D., Osipiuk, J. & Pinkau, T. (1990). Acquired thermotolerance and heat shock in the extremely thermophilic archaebacterium *Sulfolobus* sp. strain B12. *J. Bacteriol.* **172**, 1478-1484.

Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K. S. P. , Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. (1998). Protein thermostability above 100°C: A key role for ionic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12300-12305.

Vieille, C. & Zeikus, J. G. (1996). Thermozymes: Identifying molecular determinants of protein structural and functional stability. *Trends Biotechnol.* **14**, 183-190.

Vihinen, M. (1987). Relationship of protein flexibility to thermostability. *Prot. Engng.* **1**, 477-480.

Villeret, V., Clantin, B., Tricot, C., Legrain, C., Roovers, M., Stalon, V., Glansdorff, N. & van Beeumen, J. (1998). The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerisation in enzyme stability at extremely high temperatures. *Proc. Natl. Acad. Sci. U.S.A.***95**, 2801-2806.

Voorhorst, W. G. B., Eggen, R. I.L., Luesink, E. J. & de Vos, W. M. (1995). Characterisation of the *celB* gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *J. Bacteriol.* **177**, 7105-7111.

de Vos, W. M., Kengen, S. W. M., Voorhorst, W. G. B. & van der Oost, J. (1998). Sugar utilisation and its control in hyperthermophiles. *Extremophiles* **2**, 201-205.

Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995). Are buried salt bridges important for protein stability and conformational specificity?. *Nat. Struct. Biol.* **2**, 122-128.

Winterhalter, C., Heinrich, P., Candussio, A., Wich, G. & Liebl, W. (1995). Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium *Thermotoga maritima*. *Mol. Microbiol.* **15**, 431-444.

Woese, C. (1998). The universal ancestor. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6854-6859.

Woese, C. R., Kandler, O. & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576-4579.

Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D. & Rice, D. W. (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336-346.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V. & et al. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3**, 1147-1158.

Zhang, J. H., Dawes, G. & Stemmer, W. P. C. (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4504-4509.

Zwickl, P., Fabry, S., Bogedain, C., Haas, A. & Hensel, R. (1990). Glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaebacterium *Pyrococcus woesei*: characterisation of the enzyme, cloning and sequencing of the gene, and expression in *Escherichia coli*. *J Bacteriol* **172**, 4329-4338.

Chapter 2

Glutamate dehydrogenase from hyperthermophilic Bacteria and Archaea: determinants of thermostability and catalysis at extremely high temperatures.

Joyce H.G. Lebbink, Servé W.M. Kengen, John van der Oost and Willem M. de Vos.

Reprinted with permission from
Journal of Molecular Catalysis B: Enzymatic **7**, 133-145, 1999.

Abstract

Insight in the molecular mechanisms determining the extreme intrinsic thermostability of enzymes isolated from hyperthermophilic Archaea and Bacteria, is increasing rapidly as more comparative studies on their amino acid sequences, biochemical characteristics and three-dimensional structures are reported. In order to test the hypotheses arising from these data, protein engineering strategies have been applied to mesophilic and thermostable glutamate dehydrogenases (GDH) from different prokaryotic sources, followed by biochemical and structural characterisation of the engineered enzymes. This review aims to provide an overview of (i) the state of the art on biochemical and structural characterisation of thermostable GDHs, (ii) the construction and properties of hybrid GDHs obtained by domain swapping between GDHs from the mesophilic bacterium *Clostridium difficile* and the hyperthermophilic archaeon *Pyrococcus furiosus*, and (iii) the elucidation of the role of large ion-pair networks in conferring stability to GDHs from hyperthermophiles by the introduction of ion-pair networks into GDH from *Thermotoga maritima*.

Introduction

The remarkable ability of hyperthermophilic Bacteria and Archaea to grow optimally at temperatures around the boiling point of water, has raised questions on how these microorganisms have stabilized their macromolecules, and notably, their enzymes. Comparative studies of amino acid sequence data, biochemical features and an increasing number of three-dimensional structures from homologous enzymes from mesophiles and (hyper)thermophiles, have revealed a large variety of molecular adaptations that might contribute to an elevated intrinsic stability. Proposed stabilizing features range from a decreased amount of thermolabile amino acid residues, improved hydrogen bonding patterns and hydrophobic packing, lowering the surface-to-volume ratio and multimerization, to an increased number of ion-pairs and the presence of large ion-pair networks (reviewed by [1]). To test the validity of these adaptation mechanism on the one hand, and generate novel enzymes with altered properties on the other hand, protein engineering approaches have been applied, followed by biochemical and structural characterisation of the wild-type and engineered enzymes.

Among the enzymes used as model for studying adaptation mechanisms, glutamate dehydrogenase (GDH) is a suitable candidate because it is well studied, catalyzes an important reaction, and is present in all three domains of life [2,3]. A large number of primary sequences and several three-dimensional structures from GDHs obtained from organisms with a wide range of optimal

Abbreviations: GDH, glutamate dehydrogenase; GdmCl, guanidinium chloride

Table 1: Comparison of GDHs from mesophilic and hyperthermophilic Bacteria and Archaea.

Organism	T_{growth} (°C)	% aa identity	cofactor specificity	T_{opt} (°C)	$t_{1/2}$ (h/°C)	T_m (°C)	T_s (M)	3D	References
P. furiosus	100	100	NADP>NAD	> 90	12 / 100	113	6.1	2.2 Å	[4,10,16,19,21]
P. endeavori	98	96	NADP	nd	10.5 / 100	113			[22]
P. sp. KOD1	95	83	NADP>NAD	80	4 / 100				[24,52]
P. islandicum	100	nd	NAD	90	> 2 / 100				[23,51]
A. fulgidus	83	nd*	NADP	> 95	2.3 / 100				[27,53]
T. litoralis	88	87	NADP	> 95	2 / 98				[25,45]
T. profundus	80	nd	NADP	85	1 / 90				[26,54]
S. solfataricus	80	43	NADP>NAD	70	15 / 80		1.8		[17,18,55]
T. maritima	80	55	NADP<NAD	75	1.8 / 85	93	3.5	3.0 Å	[5,11,50]
C. difficile	37	52	NAD	60	5.1 / 65		1.5		[6,15]
C. symbiosum	37	37	NAD	nd	0.3 / 52	55		1.9 Å	[8,9,11,56,57]

T_{growth}:optimal temperature for growth of the organism. % aa identity: identity with *P.furiosus* GDH. T_{opt}: temperature at which optimal enzymatic activity occurs. $t_{1/2}$: time in which 50% enzymatic activity is retained after incubation at given temperature. T_m: melting temperature as determined by differential scanning microcalorimetry or thermal denaturation monitored by circular dichroism. T_s: concentration of guanidinium chloride in M in which 50% of enzyme has retained native conformation. 3D: resolution of the available crystal structure. nd = not determined. *GDH has been isolated from *A.fulgidus* strain 7324 (DSM 8774) [27], no GDH is present in *A.fulgidus* strain VC-16 (DSM 4304) of which the complete genome sequence has been determined [59].

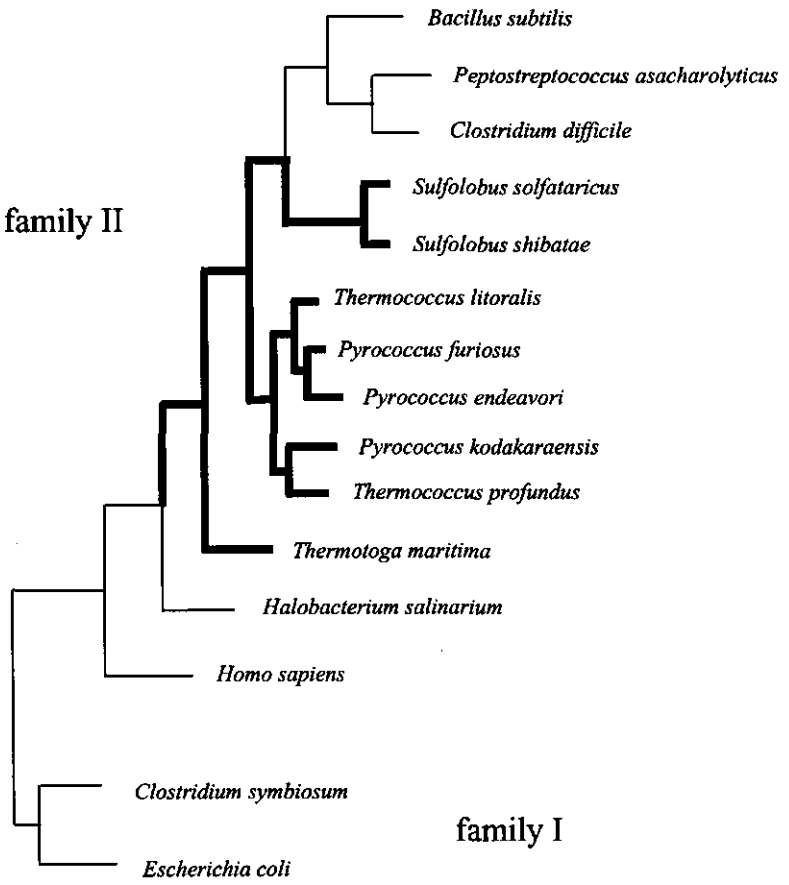


Figure 1: Schematic phylogenetic tree for GDH amino acid sequences, redrawn and modified from [5] with additions taken from [7] and [24] showing the division of GDHs into two families. Branchings corresponding to GDHs from hyperthermophilic organisms are drawn in *bold*. For accession numbers see [5,7,24].

growth temperatures are available (Table 1). Furthermore, the *gdh* genes from several organisms have been functionally overexpressed in *E.coli*, facilitating the analysis of their expression products and allowing for a variety of mutagenesis approaches [4,5,6]. GDH couples carbon and nitrogen metabolism in the cell by the oxidative deamination of L-glutamate to α-ketoglutarate and ammonia, accompanied by the reduction of the cofactor NAD^+ or $NADP^+$. Phylogenetic analysis of GDH primary sequences revealed that GDHs can be organised in two enzyme families (Fig. 1) [5,7]. In family I, exclusively sequences from bacterial GDHs are present, in family II, GDH sequences from Bacteria as well as Archaea

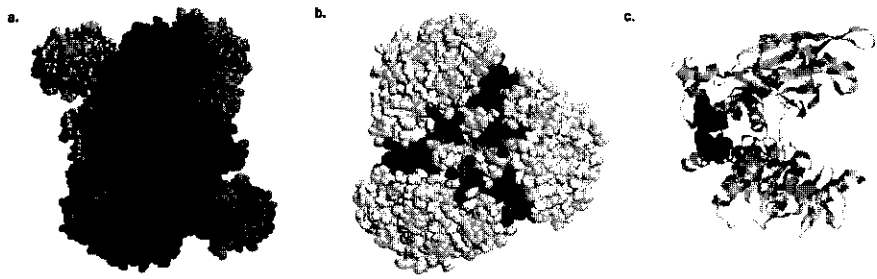


Figure 2: Hexameric structure of *P.furiosus* GDH (a) and a view of the trimer interface (b). The 18-residue ion-pair networks are highlighted in black, the pairs of 6-residue networks in grey. (c) the composition of each subunit in two separate domains. The upper, N-terminal domain binds the substrate in the cleft in between the two domains, the lower C-terminal domain binds the cofactor and is separated by a flexible hinge region from the first domain, allowing rotation of this domain resulting in opening and closure of the active site during catalysis. The five-residue hinge ion-pair network is highlighted in black. The figure was generated using coordinates of the *P.furiosus* GDH [10] and the programme RASMOL [58].

and Eukarya are found. In general, GDH is a multimeric enzyme consisting of six identical subunits (Fig. 2) [8]. The hexamer is composed of two trimers that are stacked upside down on top of each other. Each subunit contains an N-terminal, substrate-binding domain which forms almost all of the intersubunit interactions, and a C-terminal domain binding the cofactor $NAD(P)^+$. During catalysis the cleft in between the two domains, in which the active site is located, is opened and closed by a rotation of the cofactor binding domain with respect to the substrate binding domain [9]. This movement is thought to be mediated by residues and interactions in the so-called hinge region connecting the two domains. These interactions are therefore assumed to play an important role in determining activity and stability of the enzyme [10,11].

This overview describes biochemical and structural characteristics of available thermostable GDHs (summarized in Table 1), and the characterisation of mutant enzymes that have been constructed in order to study the mechanisms that are responsible for the thermo-activity and -stability of this well-studied class of enzymes (summarized in Table 2).

Thermostable GDHs

One of the most extensively studied GDHs is the enzyme from the mesophilic bacterium *Clostridium symbiosum*, that grows optimally at 37°C (Table 1). *C.symbiosum* GDH clusters among other bacterial GDH sequences in family I (Fig. 1). Its crystal structure was determined at 1.9 Å resolution and the complexes with the cofactor NAD and glutamate have been solved [8,9,12]. These different structures showed a rotation of the cofactor-binding domain with

Table 2: Overview of mutations and their effect on stability and activity of different enzymes that have been described in the text.

Enzyme	Mutation	Description of mutation and effect	Reference
C.symbiosum GDH	Phe187Asp	Mutation at subunit interface results in inactive, dimeric GDH. Activity requires communication along 3-fold axis.	[13]
	Asp165His	Mutation of general base inactivates and results in dimer: Communication between active sites and subunit interfaces.	[14]
P.furiosus GDH *C.difficile* GDH	Hybrid GDHs containing subunit/ cofactor binding domains from *P.furiosus*/*C.difficile* GDH.	Thermoactivity and stability towards thermal inactivation are determined by weakest or most flexible part of protein. Unlinking of stability towards temperature and chemical denaturants.	[15]
T4 lysozyme Barnase Arc repressor	Introduction of ion-pairs at protein surface or buried in interior	Surface exposed salt-bridges contribute only marginally to protein stability due to entropic cost of side-chain localization. Cooperativity in network: sharing of entropic costs. Buried salt-bridges may contribute considerably to stability because entropic cost has already been provided for during protein folding. Buried salt-bridges may be replaced by more stabilising hydrophobic clusters.	[40-43]
T. litoralis GDH	Thr138Glu Thr138Glu/Asp157Thr	Addition of glutamate to charge cluster at subunit interface is destabilizing; second-site mutation results in mutant enzyme more stable than wild-type GDH.	[47]
T.maritima GAPDH	Removal of charges from surface-exposed or buried charge clusters.	Removal of peripheral ion-pairs has no effect. Disruption of 4-residue ion-pair network leads to accelerated thermal denaturation.	[48,49]
T.maritima GDH	Asn97Asp Gly376Lys Asn97Asp/Gly376Lys	Introduction of 6-residue ion-pair network into flexible hinge region does not stabilize GDH but has significant effects on thermoactivity and substrate affinity.	[50]

respect to the substrate-binding domain upon ligand binding, indicating hinge bending during catalysis. As a consequence, the active site which is located in the cleft between the two domains will close upon substrate and/or cofactor binding, thereby creating an adequate hydrophobic environment for hydride transfer during catalysis [9]. Construction of a dimeric form of *C.symbiosum* GDH by substitution of a phenylalanine residue for an aspartate at the dimer-interface resulted in an inactive enzyme, indicating that catalytic activity requires subunit interaction along the three-fold symmetry axis [13]. Mutation of the general base of the catalytic reaction, aspartate 165 to histidine in the glutamate binding site, almost completely inactivates the enzyme. Moreover, this substitution results in the formation of a dimeric conformation instead of a hexamer [14]. This again indicates communication between the active site and the subunit interfaces. *C.symbiosum* GDH is not a thermostable enzyme and has a melting temperature of only 55°C [11]. Therefore, *C.symbiosum* GDH serves as mesophilic model enzyme in many studies dealing with GDH thermostability.

The gene encoding GDH from the mesophilic bacterium *C.difficile* has been cloned, sequenced and overexpressed in *E.coli* [6]. Interestingly, its expression product shows a remarkable thermostability with a half-life of activity at 65°C of 5.1 hours and is optimally active around 60°C [15]. Amino acid sequence homology comparison showed that *C.difficile* GDH does not belong to family I with the other bacterial GDHs, but clusters with the eukaryal and archaeal sequences in family II (Fig. 1).

The gene coding for GDH from the hyperthermophilic bacterium *Thermotoga maritima* (optimum temperature for growth 80°C) was cloned and overexpressed in *E.coli* and the recombinant enzyme was characterized [5]. The enzyme showed optimal activity at 75°C [5] and a melting temperature of 93°C [11]. Whereas several archaeal GDHs prefer NADP over NAD, the *T.maritima* enzyme uses NAD in favor of NADP; while other bacterial GDHs only use the non-phosphorylated cofactor. Interestingly, in the phylogenetic tree *T.maritima* GDH does not cluster with the other bacterial GDHs but is found in between that of halophilic and hyperthermophilic Archaea (Fig. 1) [5].

GDH from the hyperthermoacidophile *Sulfolobus solfataricus* (optimal temperature for growth 80°C) displays a half-life of activity of 15 hours at 80°C (0.2 mg/ml protein) [16]. This stability was shown to decrease at lower protein concentrations and increased in the presence of guanidinium chloride. This suggests that the inactivation is probably caused by the exposure to the solvent of hydrophobic regions that leads to inactive aggregates, whose formation is efficiently prevented by GdmCl [17].

GDH from the hyperthermophilic archaeon *Pyrococcus furiosus* is the most stable GDH described up to now with a half-life of activity at 100°C of 12 hours [18] and a melting point of 113°C [19]. In contrast to the thermostability of the *S.solfataricus* enzyme, that increased upon higher protein concentrations, the stability of the *P.furiosus* GDH appears to be independent of protein concentration, suggesting differences in the mechanism of stabilisation between the two enzymes. The *P.furiosus gdh* gene was cloned, characterized and its expression was studied [3,4]. Efficient overexpression was achieved under control of the phage lambda P_L promoter that requires expression under heat-

shock conditions. This resulted in identification of GDH hexamers that have a ten-fold lower specific activity than GDH isolated from *P.furiosus* [15]. A heat-incubation of the cell-extract resulted in the formation of fully active hexamers, indicating that enzymes from hyperthermophiles may need a temperature close to the *in vivo* growth optimum of the organism in order to fold correctly. A translational fusion of the PCR amplified *P.furiosus gdh* gene with the phage T7 system in *E.coli* resulted in the production of GDH comprising up to 15% of total cell protein [20]. 50% of this GDH was formed in an inactive monomeric conformation and could also be fully activated by heat-treatment. *In vitro* refolding of *P.furiosus* GDH occurs via structured monomers to the formation of higher association states with a tertiary structure different from that of the native enzyme [21]. Again heat-treatment at 70°C is required for the hexamer to acquire activity.

Thermostable GDHs have furthermore been isolated from *P.endeavori* (formerly known as ES4, [22]), *Pyrobaculum islandicum* [23], *Pyrococcus* sp. KOD1 [24], *Thermococcus litoralis* [25], *Thermococcus profundus* [26] and *Archaeoglobus fulgidus* [27]. The GDH from *P.endeavori* is highly homologous to *P.furiosus* GDH (96%, Table 1) and has a similar half-life of activity and an identical melting point. *P.endeavori*, *T.litoralis*, *T.profundus* and *A.fulgidus* GDH, interestingly, all use exclusively NADP as cofactor, in contrast to the hyperthermostable GDHs mentioned before. *P.islandicum*, however, is only able to use NAD. All enzymes of which the amino acid sequence is known, cluster with the other hyperthermophilic sequences in family II (Fig. 1).

Hybrid glutamate dehydrogenase

The construction of hybrid enzymes by the exchange of large parts of the polypeptide chain between homologous proteins, has been an established approach to study the contribution of large substructures in determining stability and activity [28]. This approach was taken in order to investigate the role of the substrate and the cofactor domains of GDH [15]. Hybrid enzymes were constructed between the GDHs from the hyperthermophilic archaeon *P.furiosus* and the mesophilic bacterium *C. difficile*. Although the two microorganisms are phylogenetically very distant, their GDHs share 52% amino acid identity and both belong to family II (Fig. 1). On the other hand, there are considerable differences in thermostability and kinetic properties between the enzymes. Two hybrid genes were constructed, containing gene fragments coding for either the substrate or the cofactor-binding domain of *P.furiosus* or *C.difficile* GDH (Fig. 3). The gene fusion was located in between two conserved glycine residues in the loop connecting the two domains, in order to disturb as little as possible the folding of the resulting hybrids in comparison with the parental enzymes. Expression of the genes in *E.coli* resulted in formation of hybrid proteins as shown by SDS-PAGE and immunoblotting using antibodies raised against pyrococcal GDH. Activity of the hybrid containing the substrate binding domain of *C.difficile* and the cofactor binding domain of *P.furiosus* GDH could not be demonstrated, indicating that correct folding and/or assembly into an active hexameric structure is impaired in this mutant. In contrast, the complementary hybrid GDH with the substrate binding

domain from *P.furiosus* and the cofactor domain from *C.difficile* GDH was produced in a hexameric, active conformation. This *P.furiosus-C.difficile* hybrid maintained efficient cofactor binding as indicated by similar affinities for NAD and NADH as found for parental *C.difficile* GDH and the inability to use the phosphorylated cofactor. However, its overall catalytic activity was low, probably caused by less efficient substrate binding. This might be due to changed interdomain interactions that are involved in positioning the two domains in the right orientation with respect to each other. The temperature optimum of the *P.furiosus-C.difficile* hybrid is, apparently, determined by the weakest or more flexible part of the protein, because the optimum of the hybrid is only a few degrees higher than that of the *C.difficile* parental enzyme (65°C instead of 60°C). It may be envisaged that the mesophilic domain reaches its maximum flexibility at the temperature optimum and starts unfolding, while the thermostable domain is still too rigid to be able to bind and release substrates efficiently, explaining the low catalytic efficiency of the hybrid. In contrast, the presence of the pyrococcal substrate binding domain is increasing the transition point in guanidinium chloride induced denaturation by as much as two-fold from 1.5 M GdmCl for *C.difficile* GDH to 3.0 M for the hybrid; *P.furiosus* GDH has a transition point at 6.1 M. However, the effect of domain exchange on thermostability is the most dramatic: Instead of obtaining a hybrid displaying a thermostability in between that of the parental enzymes, as found in the case of chemical stability, the *P.furiosus-C.difficile* hybrid is less thermostable than *C.difficile* GDH. These results indicate that, while in hyperthermostable enzymes properties like thermoactivity, thermostability and stability towards denaturants are very well optimised, they are not necessarily linked.

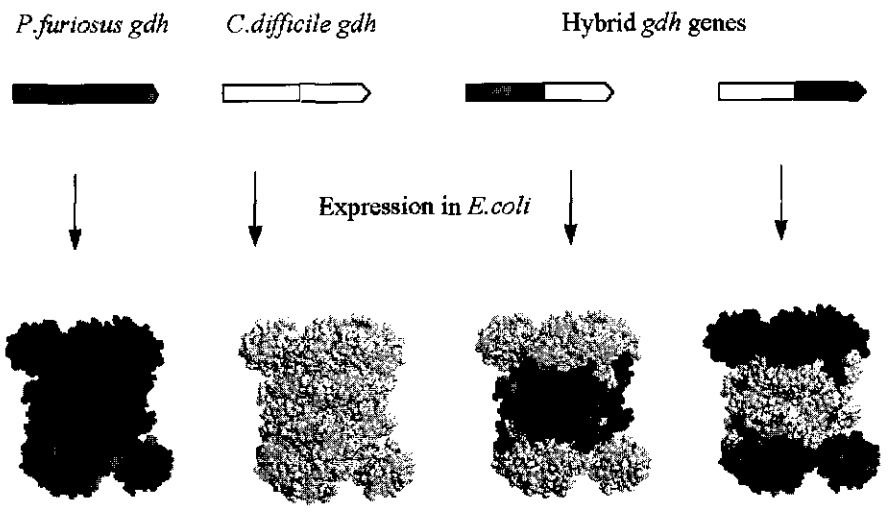


Figure 3: Schematic drawing of wild-type and hybrid *gdh* genes and the expression in *E.coli* of their corresponding gene products.

Ion-pairs and ion-pair networks in glutamate dehydrogenase

The recent elucidation of a large number of three-dimensional structures from enzymes from hyperthermophilic Bacteria and Archaea, comprise GDH, aldehyde ferredoxin oxidoreductase, ornithine carbamoyl transferase and citrate synthase from *P.furiosus* [10,29,30,31], Fe-superoxide dismutase from *Aquifex pyrophilus* [32], β-glycosidase and indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus* [33,34] and GDH, D-glyceraldehyde-3-phosphate dehydrogenase, phosphoribosyl anthranilate isomerase and signal transduction protein CheY from *T.maritima* [11,35,36,37]. In the majority of these structures, a high number of ion-pairs and large ion-pair networks were identified that are only partially or not at all present in homologous enzymes from mesophilic organisms. This finding strongly suggests an important role of these electrostatic interactions in determining enzyme hyperthermostability and/or thermoactivity.

The role of ion-pair interactions in stabilisation of proteins has been a subject of debate for years. Some twenty years ago, Perutz and coworkers already showed that ion-pairs may play a role in stabilisation of proteins [38,39]. Site-directed mutagenesis involving single, surface-located salt bridges in T4 lysozyme showed that their contribution to the stability of the protein is only marginal (0.1-0.25 kcal/mol) [40,41]. This is caused by the fact that the gain in free energy of folding is about equal to the entropic cost of dehydration and the reduction of the conformational freedom. In contrast, a buried salt-bridge in the same enzyme was shown to stabilize the native state by 3-5 kcal/mol as compared to the unfolded state [42]. However, a buried salt-bridge triad in the Arc repressor from bacterial phage P22 could be replaced by more stable hydrophobic interactions [43].

Theoretical considerations indicate that extensive networks may play an important role in maintaining enzyme stability or function at extreme temperatures. For each additional ion-pair that is added to a network, only a single residue needs to be desolvated and fixated. Indeed, multiple salt bridges in barnase show cooperativity; each single ion-pair in a three-residue network was found to enhance the strength of the other interaction by 0.8 kcal/mol [41]. In addition, part of the entropic cost of fixation is in many cases already provided during the folding of the protein since networks are often located in cavities and at interfaces. Furthermore, at high temperatures electrostatic interactions may become more important because hydration effects play a minor role and the dielectric constant decreases with temperature, resulting in an increased electrostatic energy upon formation of an ion-pair.

Ion-pair networks have been characterized most extensively in GDH. GDH from *P.furiosus* was found to contain more ion-pairs and larger ion-pair networks than GDHs from the mesophilic homologues from *E.coli*, *Neurospora crassa* and *C.symbiosum* (Table 3) [10,44]. The number of ion-pairs in *P.furiosus* GDH has been doubled compared to *C.symbiosum* and *E.coli* GDH. In the hyperthermostable enzyme, the residues involved in these ionic interactions, preferably pair with more than one partner. While in the *C.symbiosum* hexamer only about 10% of these residues form two and three ionic interactions, in the *P.furiosus* hexamer more than half of the residues form multiple ion-pair interactions. Consequently, extensive interacting networks are present in the

P.furiosus GDH. In *C.symbiosum* and *E.coli* GDH, approximately 25% of the ion-pairs are arranged in networks of four or more residues, respectively; in *P.furiosus* GDH this number amounts to 62%. While these four-residue clusters are the largest to be identified in the mesophilic enzymes, *P.furiosus* GDH contains networks comprising five, six and up to 18 residues.

The elucidation of the crystal structure of GDH from the hyperthermophilic bacterium *T.maritima* confirmed the ion-pair network hypothesis [11]. In line with the stability of this enzyme being intermediate between that of *C.symbiosum* and *P.furiosus* GDH, the total number of ion-pair interactions per monomer is in between that of the other enzymes (Table 3). In comparison to the *C.symbiosum* hexamer, the number of three- and four-residue networks in *T.maritima* GDH increased and 6 networks of seven residues are present. However, these extra ion-pairs are mainly located within each subunit, while in *P.furiosus* GDH many large networks are found at the subunit interface. In contrast, the subunit interface in *T.maritima* GDH seems to be optimized by hydrophobic interactions.

In order to obtain more support for the ion-pair network hypothesis, a structure-based homology modelling study was recently carried out using sequences from glutamate dehydrogenases from ten different sources, spanning a large temperature spectrum [45]. Specific attention was given to the 18-residue network that is three times present at the subunit interface of *P.furiosus* GDH, and the three pairs of 6-residue networks, also present at the subunit interface in between the 18-residue networks (Fig. 2). For all GDHs that were studied, the networks decreased in size, became fragmented and finally completely disappeared with decreasing thermostability. The formation of large, intricate ion-pair networks seems to be specifically correlated with enzymatic operation close to or above 100°C.

The first experimental approach addressing the ion-pair network

Table 3: Analysis of ion-pairs and ion-pair networks in GDH from *E.coli*, *C.symbiosum* and *P.furiosus* [44], and *T.maritima* [11].

	E.coli	*C.symbiosum*	*T.maritima*	*P.furiosus*
no. of ion-pairs per subunit	26	28	37	45
no. of residues per hexamer forming 1/2/3 ion-pairs	168/36/6	144/36/6	132/78/18	108/102/30
no. of residues forming 2/3/4 membered networks	84/24/12	72/24/12	66/33/18	54/24/12
no. of residues forming 5/6/7/18 membered networks	0/0/0/0	0/0/0/0	0/0/6/0	12/6/0/3
% of ion-pairs in ≥ 4 residue networks	21	23	40	62
no. of intersubunit ion-pairs	nd	36	34	54

hypothesis in a hyperthermophile, is the acid-induced denaturation of *P.furiosus* GDH and biochemical characterisation of the unfolded enzyme [46]. Lowering the pH into the acidic region results in protonation of negatively charged groups in the enzyme, and the shielding of dipoles by the added ions resulting in the disruption of electrostatic interactions. Therefore any residual structure at low pH values should be a result of hydrophobic interactions. GdmCl-induced denaturation was found to be pH dependent, indicating that electrostatic interactions play indeed a major role in determining the stability of the enzyme. However, at pH 1 a monomeric protein was formed that could be further unfolded by GdmCl with a higher transition midpoint than found for other thermophilic and mesophilic glutamate dehydrogenases [46]. This finding indicates that, at least within one subunit, electrostatic interactions cannot be the predominant force in *P.furiosus* GDH.

Preliminary results on mutagenesis of *T.litoralis* GDH indicate that the introduction of a charged residue at the subunit interface (T138E) in first instance results in a lower thermostability of the mutant *T.litoralis* GDH [47]. The introduction of a second, nearby mutation (D157T) that does not take part in the presumed ionic network, elevates the thermostability of the double mutant above that of the wild-type enzyme. These results confirm that, assuming that the presumed ionic interactions are indeed formed in the mutant enzyme, ion-pair networks might indeed be a stabilizing feature in hyperthermostable GDH.

Introduction of ion-pair networks into *T.maritima* GDH

In order to experimentally verify the presumed stabilizing contribution of ion-pair networks to the thermostability of enzymes from hyperthermophiles, several strategies can be followed. One is to remove the interaction from the enzyme by replacement of the charged amino acid for an uncharged residue using site-directed mutagenesis. This approach was used to demonstrate the stabilizing effect of small ion-pair networks in glyceraldehyde-3-phosphate dehydrogenase from *T.maritima* [48,49]. A drawback of this approach, however, is the fact that always more than only the ionic interaction is affected. The hydrogen bonding potential will be changed and cavities or steric hindrance may be introduced. An alternative approach is to introduce charged residues that are present in a thermostable enzyme, into a less stable homologue. We employed this strategy to study the role of *P.furiosus* GDH ion-pair networks by introducing these into the less stable GDH from the hyperthermophilic bacterium *T.maritima* [50]. The GDHs from *P.furiosus* and *T.maritima* share 55% amino acid identity and are very homologous in secondary and tertiary structure as well [11]. However, the *P.furiosus* GDH has a higher melting point than the *T.maritima* GDH (113°C versus 93°C) [19,11] and contains significantly larger ion-pair networks. The trimer interface in *P.furiosus* is almost completely charged because of the presence of three 18-residue networks which are separated by three pairs of 6-residue networks (Fig. 2) [10]. The *P.furiosus* GDH 18-residue network is fragmented into two small ones of only four residues in the *T.maritima* GDH and from the pyrococcal six-residue network only one residue is conserved. The major difference between the hinge regions connecting the two domains, is the presence of a five-residue ion-pair network in the *P.furiosus* enzyme that is

absent in the less stable *T.maritima* GDH. This five-residue ion-pair network links secondary structure elements from both domains, is surface exposed, and located opposite to the substrate binding site at a distance of more than 12 Å (Fig. 2). The five participating residues in *P.furiosus* GDH are positioned in such a way that four salt bridges are formed. *T.maritima* GDH contains, instead of aspartate and lysine, asparagine and glycine at positions 97 and 376, respectively, resulting in the loss of three of the four ionic interactions. An additional lysine in *T.maritima* GDH (which is an asparagine in the pyrococcal enzyme and does not interact at all with the network) brings the number of ion-pairs in this region to two [50].

Two single mutant *T.maritima* GDHs were generated and characterized, containing the substitutions asparagine to aspartate at position 97 and glycine to lysine at position 376, as well as double mutant N97D/G376K [50]. The three-dimensional structure of the double mutant was solved at 3.0 Å resolution and revealed that a six-residue ion-pair network is present in this mutant, this network being even larger than the one in the *P.furiosus* enzyme. Apparent melting temperatures of 91, 92 and 91°C were determined for N97D, G376K and the double mutant, respectively, not differing significantly from that of the wild-type GDH (93°C). In addition to this, identical transition midpoints in guanidinium chloride induced unfolding were found; 3.5 M for wild-type GDH and the single mutants, 3.6 M for the double mutant. In contrast to this, thermal inactivation at 85°C occurred more than two-fold faster for all mutant enzymes than for the wild-type GDH. The effect of the two single mutations was not additive, suggesting a stabilizing feature between these two residues in the double mutant which might be either or both the ionic interaction or the extra hydrogen bond.

In addition to the effect of the mutations on the stability of the enzyme, changes in the thermoactivity and kinetic parameters of the mutants in comparison with the wild-type GDH were studied [50]. At temperatures of 65°C and higher, the wild-type and the three mutant enzymes showed identical specific activities. However, at 58°C the specific activity of N97D/G376K and G376K was found to be significantly higher than that of the wild-type and N97D GDHs. Enzyme inactivation started at 58°C for G376K and N97D/G376K, at 62°C for the wild-type GDH and at 66°C for N97D. For the wild-type GDH, K_m values for α-ketoglutarate and NADH were found to decrease, and that for ammonia to increase, with increasing temperature. K_m values for all substrates changed differently for each mutant GDH and with temperature, indicating that the mutations have pronounced effects on catalysis and enzyme efficiency at different temperatures. The most pronounced effect was found for mutant N97D that has a ten-fold lower K_m value for NADH at 58°C than at 25°C and is at this temperature more efficient than the wild-type GDH on this substrate.

In conclusion, these results indicate that the engineered ion-pair interactions in the hinge region do not affect the stability towards temperature- or guanidinium chloride-induced denaturation but rather affect the specific activity and substrate affinity of the enzyme and the temperature at which it functions optimally.

In addition to studying the role of ion-pair networks in the flexible hinge region of GDH, we studied the contribution of the large intersubunit network

that comprises 18 residues in *P.furiosus* GDH and is fragmented into two small networks of only four residues in *T.maritima* GDH (Fig.2) (Lebbink *et al.*, manuscript in preparation). In order to reconstruct the network into the less stable enzyme, two positive and two negative charges were introduced into *T. maritima* GDH as single substitutions, several double and triple combinations and the quadruple mutant. Preliminary analysis revealed different effects of each substitution on thermal inactivation and activity of the enzyme and also pointed out the need for balanced charges (Lebbink *et al.*, manuscript in preparation).

Conclusion

A large amount of structural and biochemical data of wild-type and mutant GDHs has been generated during recent years, contributing greatly to our knowledge of the molecular adaptations that govern enzyme hyperthermostability and activity. This wealth of information and the fact that GDH is a large, multimeric enzyme containing a high number of complex ion-pair networks, makes it an excellent model system to study the role of these features in determining protein hyperthermostability, which is undoubtedly one of the main topics in protein chemistry at the present time. We have shown that the introduction of a six-residue ion-pair network in the flexible hinge region of *T.maritima* GDH is not affecting the melting temperature of the enzyme, leads to somewhat faster thermal inactivation, but has pronounced effects on thermoactivity and kinetic parameters [50]. The finding that the presence of the ion-pair network does not seem to increase the stability of the enzyme, may indicate that in this region of the enzyme, ion-pairs indeed do not contribute to stability. Alternatively, these results may reflect the considerable differences between the archaeal and the bacterial enzyme (amino acid identity 55%). Although the direct environment of the hinge network is highly conserved, small changes that are not visible in the three-dimensional structures, or unfavourable interactions with macroscopic parameters like the existing electrostatic field, may lead to non-optimal interactions in the mutant enzyme, resulting in a faster thermal inactivation. This consideration is supported by the fact that in *T.litoralis* GDH two mutations (one introducing a charge, the second one being nearby but removing a charge) apparently are needed for stabilisation of this enzyme [47]. The degree of homology between *P.furiosus* and *T.litoralis* GDH is much higher (87%) than that between P.furiosus and *T.maritima* (55%) and the results suggest that already in a highly homologous system, second-site mutations are required to obtain the desired stabilisation. For a further interpretation of these preliminary data, it is essential to determine the three-dimensional structures of wild-type and mutant *T.litoralis* GDH, and investigate whether the additional charge indeed participates in an ion-pair network and what is the basis for the stability-rescuing effect of the second-site mutation.

The fact that GDH is a large, multimeric enzyme is not only an advantage. Due to this complexity, GDH does not unfold reversibly during thermal incubation or guanidinium chloride induced denaturation and therefore no thermodynamic analysis of wild-type and mutant GDHs can be performed. This means that only data about the thermoactivity and the kinetic stability of GDH can be obtained. To circumvent this problem, smaller model systems like monomeric proteins or

isolated domains could be used. However, only few monomeric thermostable enzymes have been reported to date and the enzymatic activity of isolated domains is difficult to assess. Furthermore, much smaller networks are to be expected in these systems, and intersubunit networks - the most prominent feature in the most stable GDH - cannot be studied in this way. Therefore, GDH will remain one of the challenging model systems for analysing structure-stability relations in the near future.

Acknowledgements

We sincerely thank Stefan Knapp and Rudolf Ladenstein (Huddinge, Sweden), Valerio Consalvi, Roberta Chiaraluce and Roberto Scandurra (Roma, Italy), David Rice (Sheffield, UK) and Ans Geerling and Rik Eggen (Wageningen, NL) for their extensive and valuable contributions to the work described in this overview. This work was partially supported by the EU contracts BIOT-CT93-0274, BIO 4-CT96-0488 and BRPR-CT97-0484.

References

[1] R. Ladenstein and G. Antranikian, in: Th. Scheper (Ed.), *Advances in Biochemical Engineering/Biotechnology*, **Vol 61**. Springer-Verlag Berlin Heidelberg, 1998, pp. 37-85.
[2] E.L. Smith, B.M. Austen, K. M. Blumenthal and J.F. Nyc, in P.D. Boyer (Ed.), *The Enzymes*, **Vol 11**, 3rd edition, Academic Press, New York (1975) pp. 293-367.
[3] R.I.L. Eggen, A.C.M. Geerling, W.G.B. Voorhorst, R. Kort and W.M. de Vos, *Biocatalysis*, **11**(1994) 131-141.
[4] R.I.L. Eggen, A.C.M. Geerling, K. Waldkötter, G. Antranikian and W.M. de Vos, *Gene*, **132** (1993) 143-148.
[5] R. Kort, W. Liebl, B. Labedan, P. Forterre, R.I.L. Eggen and W.M. de Vos, *Extremophiles*, 1 (1997) 52-60.
[6] D.M. Lyerly, L.A. Barosso and T.D.Wilkins, *J. Clin. Micr.*, **29** (1991) 2639-2642.
[7] N. Benachenhou-Lahfa, P. Forterre and B. Labedan, *Mol. Evol.*, **36** (1993) 335-346.
[8] P.J. Baker, K.L. Britton, P.C. Engel, G.W. Farrants, K.S. Lilley, D. Rice, and T.J. Stillman, Proteins: *Struct. Funct. Genet.*, **12** (1992) 75-86.
[9] T.J. Stillman, P.J. Baker, K.L. Britton and D.W. Rice, *J. Mol. Biol.*, **234** (1993) 1131-1139.
[10] K.S.P. Yip, T.J. Stillman, K.L. Britton, P.J. Artymiuk, P. J., Baker, S.E. Sedelnikova, P.C. Engel, A. Pasquo, R. Chiaraluce, V. Consalvi, R. Scandurra and D.W. Rice, *Structure*, **3** (1995) 1147-1158.
[11] S. Knapp, W.M. de Vos, D. Rice and R. Ladenstein, *J. Mol. Biol.*, **267** (1997) 916-932.
[12] T.J. Stillman, P.J. Baker, K.L. Britton, D.W. Rice and F.H. Rodgers, *J.Mol. Biol.*, **224** (1992) 1181-1184.
[13] A. Pasquo, K.L. Britton, T.J. Stillman, D.W. Rice, H. Cölfen, S.E. Harding, R. Scandurra and P.C. Engel, *Biochim. Biophys. Acta*, **1297** (1996) 149-158.
[14] J.L.E. Dean, H. Cölfen, S.E. Harding, D.W. Rice and P.C. Engel., *Eur. Biophys. J.*, **25** (1997) 417-422.
[15] J.H.G. Lebbink, R.I.L. Eggen, A.C.M. Geerling, V.Consalvi, R. Chiaraluce, R. Scandurra, and W.M. de Vos, *Prot. Eng.*, **8** (1995) 1287-1294.
[16] V. Consalvi, R. Chiaraluce, L. Politi, R. Vaccaro, M. de Rosa and R. Scandurra, *Eur. J. Biochem.*, **202** (1991) 1189-1196.
[17] V. Consalvi, R. Chiaraluce, L.Politi, A. Pasquo, M.De Rosa, R. Scandurra, *Biochimica Biophysica Acta*, **1202** (1993) 207-215.
[18] V. Consalvi, R. Chiaraluce, L. Politi, A. Gambacorta, M.de Rosa and R. Scandurra, *Eur. J. Biochem.*, **196** (1991) 459-467.
[19] H. Klump, J. DiRuggiero, M. Kessel, J-B Park, M.W.W. Adams and F.T. Robb, *J. Biol. Chem.*, **267** (1992) 22681-22685.
[20] J. Diruggiero and F.T. Robb, *Appl. Environ. Microbiol.*, **61** (1995) 159-164.
[21] V. Consalvi, R. Chiaraluce, S. Millevoi, A. Pasquo, P. Vecchini, E. Chiancone and R. Scandurra, *Eur. J. Biochem.*, **239** (1996) 679-685.
[22] J. Diruggiero, F.T. Robb, R. Jagus, H.H. Klump, K.M. Borges, M. Kessel, X. Mai and M.W.W. Adams, *J. Biol. Chem.*, **268** (1993) 17767-17774.
[23] C. Kujo and T. Ohshima, *Appl. Env. Microbiol.*, **64** (1998), 2152-2157.
[24] R.N.Z.A. Rahman, S. Fujiwara, M. Takagi and T.Imanaka, *Mol. Gen. Genet.*, **257** (1998) 338-347.
[25] K. Ma, F.T. Robb, and M.W.W. Adams, *Appl. Env. Micr.*, **60** (1994) 562-568.
[26] T. Kobayashi, S. Higuchi, K. Kimura, T. Kudo and K. Horikoshi, *J. Biochem.*, **118** (1995) 587-592.

[27] N. Aalén, I.H. Steen, N-K. Birkeland, T. Lien., *Arch. Microbiol.* **168** (1997) 536-539.
[28] A.E. Nixon, M. Ostermeier and S.J. Benkovic, *Trends in Biotech.*, **16** (1998) 258-264.
[29] M.K. Chan, S. Mukund, A. Kletzin, M.W.W. Adams and D.C. Rees, *Science*, **267** (1995) 1463-1469.
[30] V. Villeret, B. Clantin, C. Tricot, C. Legrain, M. Roovers. V. Stalon, N. Glansdorff and J. van Beeumen, *Proc. Natl. Acad. Sci.*, **95** (1998) 2801-2806.
[31] R.J.M. Russell, J.M.C. Ferguson, D.W. Hough, M.J. Danson and G.L. Taylor, *Biochemistry*, **36** (1997) 9983-9994.
[32] J-H. Lim, Y.G. Yu, Y.S. Han, S-j. Cho, B-Y. Ahn, S.H. Kim and Y. Cho, *J. Mol. Biol.*, **270** (1997) 259-274.
[33] C.F. Aguilar, I. Sanderson, M. Moracci, M. Ciaramella, R. Nucci, M. Rossi, M. and L.H. Pearl, *J. Mol. Biol.*, **271** (1997) 789-802.
[34] M. Hennig, B. Darimont, R. Sterner, K. Kirschner and J.N. Jansonius, *Structure*, **3** (1995) 1295-1306.
[35] I. Korndörfer, B. Steipe, R. Huber, A. Tomschy and R. Jaenicke, *J. Mol. Biol.*, **246** (1995) 511-521.
[36] M. Hennig, R. Sterner, K. Kirschner and J.N. Jansonius, *Biochemistry*, **36** (1997) 6009-6016.
[37] K.C. Usher, A.F.A. de la Cruz, F.W. Dahlquist, R.V. Swanson, M.I. Simon and S.J. Remington, *Prot. Science*, **7** (1998) 403-412.
[38] M.F. Perutz and H. Raidt, *Nature*, **255** (1975) 256-259.
[39] M.F. Perutz, *Science*, **201** (1978) 1187-1191.
[40] S. Dao-pin, U. Sauer, H. Nicholson and B.W. Matthews, *Biochemistry*, **30** (1991) 7142-7153.
[41] A. Horovitz, L. Serrano, B. Avron, M. Bycroft and A.R. Fersht, *J. Mol. Biol.*, **216** (1990) 1031-1044.
[42] D.E. Anderson, W.J. Becktel and F.W. Dahlquist, *Biochemistry*, **29** (1990) 2403-2408.
[43] C.D. Waldburger, J.F. Schildbach and R.T. Sauer, *Structural Biology*, **2** (1995) 122-128.
[44] D.W. Rice, K.S.P. Yip, T.J. Stillman, K.L. Britton, A. Fuentes, I. Connerton, A. Pasquo, R. Scandurra and P.C. Engel, *FEMS Micr. Rev.*, **18** (1996) 105-117.
[45] K.S.P. Yip, K.L. Britton, T.J. Stillman, J.H.G. Lebbink, W. M. de Vos, F.T. Robb, C. Vetriani, D. Maeder and D.W. Rice, *Eur. J. Biochem.*, **255** (1998) 336-346.
[46] R. Chiaraluce, R.M. Schwerdtfeger, R. Scandurra, G. Antranikian and V. Consalvi, *Eur. J. Biochem.*, **247** (1997) 224-230.
[47] C. Vetriani, D.L. Maeder, N. Tolliday, K.S. Yip, T.J. Stillman, K.L. Britton, D. Rice, H.H. Klump and F.T. Robb. *Abstracts from the International Congress on Extremophiles '98*, pp. 252, Yokohama, 18-22 January 1998, Japan.
[48] A. Tomschy, G. Böhm and R. Jaenicke, *Prot. Eng.*, **7** (1994) 1471-1478.
[49] G. Pappenberger, H. Schurig and R. Jaenicke, *J. Mol. Biol.*, **274** (1997) 676-683.
[50] J.H.G. Lebbink, S. Knapp, J. van der Oost, D. Rice, R. Ladenstein and W.M. de Vos, *J. Mol. Biol.*, **280** (1998) 287-296.
[51] M. Molitor, C.Dahl, I. Molitor, U. Schafer, N. Speich, R. Huber, R. Deutzmann, H.G. Truper., *Microbiology*, **144** (1998) 529-541.
[52] M.Morikawa, Y. Izawa, N. Rashid, T. Hoaki, T. Imanaka, *Appl. Env. Microbiol.*, **60** (1994) 4559-4566.
[53] K.O. Stetter, *Syst. Appl. Env. Microbiol.*, **10** (1988) 172-173.
[54] T. Kobayashi, Y.S. Kwak, T. Akiba, T. Kudo and K. Horikoshi, *Syst. Appl. Microbiol.*, **17** (1994) 232-236.
[55] B. Maras, V. Consalvi, R. Chiaraluce, L.Politi, M. de Rosa and R. Scandurra, *Eur. J. Biochem.*, **203** (1992) 81-87.
[56] S.E.H.. Syed, *PhD Thesis* (1987) University of Sheffield, UK.
[57] J.K. Teller, R.J. Smith, M.J. McPherson, P.C. Engel and J.R. Guest, *Eur. J. Biochem.*, **206** (1992) 151-159.
[58] R. Sayle, RasMol V2.5, Molecular Visualisation Program, (1992) Glaxo Research and Development, Greenford, Middlesex, U.K.
[59] H.P. Klenk *et al.*, *Nature*, **390** (1997) 364-70.

Chapter 3

Exchange of domains of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* and the mesophilic bacterium *Clostridium difficile*: effects on catalysis, thermoactivity and stability.

Joyce H.G. Lebbink, Rik I.L. Eggen, Ans C.M. Geerling, Valerio Consalvi, Roberta Chiaraluce, Roberto Scandurra and Willem M. de Vos.

Reprinted with permission from
Protein Engineering **8**, 1287-1294, 1995.

Exchange of domains of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* and the mesophilic bacterium *Clostridium difficile*: effects on catalysis, thermoactivity and stability

Joyce H.G.Lebbink[1], Rik I.L.Eggen[1,2], Ans C.M.Geerling[1], Valerio Consalvi[3], Roberta Chiaraluce[3], Roberto Scandurra[3] and Willem M.de Vos[1,4]

[1]Department of Microbiology, Wageningen Agricultural University, Hesselink van Suchtelenweg 4, NL-6703 CT Wageningen, The Netherlands and [3]Dipartimento di Scienze Biochimiche 'A Rossi Fanelli', Università 'La Sapienza', Piazzale Aldo Moro 5, I-00185, Roma, Italy

[2]Present address: Swiss Federal Institute for Environmental Science and Technology (EAWAG), Ueberlandstrasse 133, CH-8600 Duebendorf, Switzerland

[4]To whom correspondence should be addressed

The glutamate dehydrogenase gene from the hyperthermophilic archaeon *Pyrococcus furiosus* has been functionally expressed in *Escherichia coli* under the control of the λ P_L promoter. The *P.furiosus* glutamate dehydrogenase amounted to 20% of the total *E.coli* cell protein, and the vast majority consisted of hexamers. Following activation by heat treatment, an enzyme could be purified from *E.coli* that was indistinguishable from the glutamate dehydrogenase purified from *P.furiosus*. Hybrid genes, that consisted of the coding regions for the homologous glutamate dehydrogenases from *P.furiosus* and the mesophilic bacterium *Clostridium difficile*, were constructed and successfully expressed in *E.coli*. One of the resulting hybrid proteins, containing the glutamate binding domain of the *C.difficile* enzyme and the cofactor binding domain of the *P.furiosus* enzyme, did not show a detectable activity. In contrast, the complementary hybrid containing the *P.furiosus* glutamate and the *C.difficile* cofactor binding domain was a catalytically active hexamer that showed a reduced substrate affinity but maintained efficient cofactor binding with the specificity found in the *Clostridium symbiosum* enzyme. Compared with the *C.difficile* glutamate dehydrogenase, the archaeal–bacterial hybrid is slightly more thermoactive, less thermostable but much more stable towards guanidinium chloride-induced inactivation and denaturation.

Keywords: glutamate dehydrogenase/heterologous expression/hybrid enzymes/*Pyrococcus furiosus*/thermostability

Introduction

The discovery of microbes that have the capacity to grow optimally at or around the boiling temperature of water (Stetter *et al.*, 1990) has greatly stimulated the analysis of the molecular mechanisms of adaptation of these hyperthermophilic organisms to their extreme environment. Considerable attention has been focused on protein thermostability that is generally accompanied by a stability towards chemical denaturants and proteolysis (Fontana, 1991). Theoretical considerations in combination with protein sequence comparisons have indicated that protein thermostability is a consequence of small but multiple changes distributed all over the protein (Jaenicke, 1991). Site-directed and random mutagenesis studies have been used to investigate the contribution of specific interactions to protein stability and have confirmed that only a few residues play a critical role in achieving thermal and chemical stability (Goldenberg, 1992). The recent elucidation of the 3-D structures of several proteins from thermophilic microbes and their comparison with proteins from mesophilic sources have advanced the understanding of the thermal stability of proteins. Although various different mechanisms have been identified by which the thermal stabilization of proteins may occur, an analysis of the structures of rubredoxin and aldehyde ferredoxin oxidoreductase from the hyperthermophilic *Pyrococcus furiosus* has indicated that they both share an increased number of electrostatic interactions in comparison with other proteins (Day *et al.*, 1992; Chan *et al.*, 1995). A large number of extra ion pairs was also observed in the structure of the glyceraldehyde-3-phosphate dehydrogenase from the thermophilic *Thermotoga maritima* (Korndörfer *et al.*, 1995). However, site-directed mutagenesis studies have so far failed to identify the specific ion pairs that contribute to the high thermostability of the latter enzyme (Tomschy *et al.*, 1994).

Complementary to studying the role of specific residues, hybrid proteins in which large parts of the polypeptide chain are exchanged between homologous proteins offer the possibility of investigating the role of larger substructures in conferring functionality and stability to a protein. This approach has been applied to bacterial 3-isopropylmalate dehydrogenase from *Thermus thermophilus* and the archaeal glyceraldehyde-3-phosphate dehydrogenase from *Methanothermus fervidus*, in which parts of mesophilic counterparts were introduced (Biro *et al.*, 1990; Onodera *et al.*, 1994). In addition, hybrid enzymes have been made between the bacterial *Thermoactinomyces intermedius* phenylalanine dehydrogenase and *Bacillus stearothermophilus* leucine dehydrogenase (Kataoka *et al.*, 1994). In all cases, novel dehydrogenases with characteristics from both parental enzymes were generated that allowed the localization of regions involved in the thermostability and binding of substrate and cofactor. We have extended this approach of producing hybrid proteins by exchanging domains in glutamate dehydrogenase (GDH). GDH is a well studied enzyme catalyzing the oxidative deamination of L-glutamate to α-ketoglutarate and ammonia, accompanied by the reduction of a single molecule of NAD^+ or $NADP^+$ (Smith *et al.*, 1975). Primary sequences of GDH are available from a wide variety of organisms and have been used to arrange GDHs into two different families (Benachenhou-Lahfa *et al.*, 1993). In general, GDH is a hexameric enzyme that consists of identical subunits. Recently, the 3-D structure of the hexameric GDH from *Clostridium symbiosum*, belonging to family I, has been solved (Baker *et al.*, 1992). Each subunit can be divided into two domains: domain I forms the core of the enzyme and binds the substrate, while domain II points outwards and binds the cofactor.

The most thermostable GDH is found in the hyperthermophilic archaeon *P.furiosus* with a half-life of ~12 h at 100°C

Table I. Universal and gene-specific primers used for construction and DNA sequencing of the hybrid *gdh* genes

Primer	Sequence	Position		
BG26	5'-GGATAACAATTTCACACAGG-3'	pUC19	354–373	S
BG27	5'-GTTTTCCCAGTCACGAC-3'	pUC19	492–476	A
BG18	5'-TGTCTTGCCCTTGAGGG-3'	*P.furiosus gdh*	629–645	A
BG56	5'-TATACCTGCATCCACCAATGCTAAGAGGC-3'	*P.furiosus gdh*	550–531	A
BG57	5'-TATACCTGCAATTTGGTGGTTCTTTAGG-3'	*C.difficile gdh*	539–557	S
BG58	5'-TATACCTGCTGGTGGATCACTTGGAAGAA-3'	*P.furiosus gdh*	543–563	S
BG59	5'-TATACCTGCAAAGATCCACCAAATTCAACTG-3'	*C.difficile gdh*	547–533	A
BG65	5'-CCACAAATAATGGCATG-3'	*P.furiosus gdh*	448–464	S
BG66	5'-AGCTATAGAAGAAGTG-3'	*P.furiosus gdh*	888–903	S

The position of the primers is indicated in the sequence of pUC19 (Sambrook *et al.*, 1989) or that of the different *gdh* genes numbered from first base of the initiation start codon (Lyerly *et al.*, 1991; Eggen *et al.*, 1993). S, sense primer; A, antisense primer.

(Consalvi *et al.*, 1991). Remarkably, GDH may comprise up to 20% of the total cell protein in *P.furiosus* and plays an important role in the removal of reduction equivalents (Kengen and Stams, 1994). We have cloned and characterized the *gdh* gene from *P.furiosus* and found it to encode a GDH belonging to family II (Eggen *et al.*, 1993). Recently, we described the overexpression of the *P.furiosus gdh* gene in *Escherichia coli* and the formation of an active GDH in this host (Eggen *et al.*, 1994). Here we present the characterization of the *P.furiosus* GDH produced by *E.coli* and a comparison with the enzyme isolated from *P.furiosus*. To study the role of the different domains in catalytic activity and stability, hybrid proteins were constructed between GDH from *P.furiosus* and the phylogenetically very distant mesophilic bacterium *Clostridium difficile* that share 52% amino acid sequence identity and both belong to the family II GDHs. Two hybrids were generated that contained either the substrate or the cofactor binding domain of *P.furiosus* or *C.difficile* GDH. The functionality, catalytic activity and both the chemical and thermal stabilities of these hybrids were compared with those of the parental enzymes produced in *E.coli*. An active archaeal–bacterial hybrid GDH was obtained with new properties.

Materials and methods

Bacterial strains and media

Escherichia coli TG1 [*supE hsd* Δ5 *thi* Δ(*lac-proAB*) F' (*tra*D 35 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15); Gibson, 1984] was cultivated in LB-based medium at 37°C. *Escherichia coli* K12ΔH1Δ*trp* [M72 Sm^R*lacZ*am Δ*bio-uvrB* Δ*trpEA2*(λ*N*am7-*N*am53-*c*I857ΔH1); Bernard *et al.*, 1979] was used in combination with expression plasmids containing the λ P_L promoter and was cultivated in the same medium.

Enzymes, chemicals and primers

Restriction enzymes and DNA modifying enzymes were purchased from Life Technologies Inc. and Pharmacia LKB Biotechnologies. Biochemicals of the highest purity were obtained from Boehringer and Fluka. Oligonucleotides were purchased from Pharmacia LKB Biotechnologies and are listed in Table I.

DNA isolation and manipulation

Small-scale plasmid DNA isolations, ligations, restrictions, DNA purification, agarose gel electrophoresis and transformation of *E.coli* were all performed as described by Sambrook *et al.* (1989). The plasmids pUC18 and pUC19 and the bacteriophages M13mp18 and M13mp19 (Yannisch-Perron *et al.*, 1985) were used as cloning vectors. The expression vector pPLc2833 is a derivative of pPLc28 (Remaut *et al.*, 1981) containing additional *Pst*I–*Sal*I–*Xba*I–*Sal*I–*Xba*I–*Bam*HI restriction sites in between the unique *Bam*HI and *Hin*dIII sites immediately downstream of the P_L promoter in pPLc2833. The hybrid genes were sequenced according to the dideoxy chain termination method (Sanger *et al.*, 1977) using the T7 sequencing kit (Pharmacia). Universal and gene-specific oligonucleotides were used as primers (Table I).

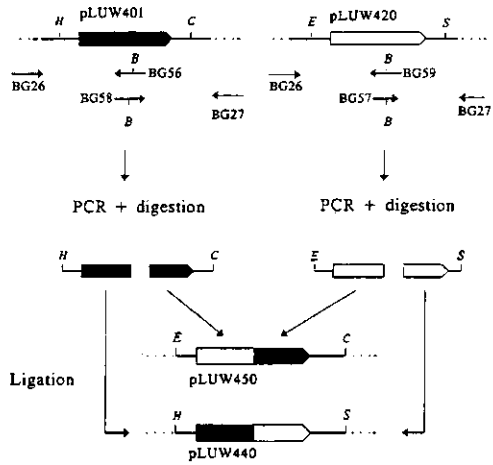


Fig. 1. Construction of hybrid genes. The physical maps of the parental plasmids pLUW401, containing the 1.6 kb *Hin*dIII–*Hin*cII fragment with the *P.furiosus gdh* gene (black), and pLUW420, containing the 1.6 kb *Eco*RI–*Sph*I fragment with the *C.difficile gdh* gene (white). Relevant restriction sites are indicated: H, *Hin*dIII; C, *Hin*cII; E, *Eco*RI; S, *Sph*I; B, *Bsp*MI.

Construction of the expression plasmids for wild-type and hybrid gdh genes

A 1.6 kb *Hin*dIII–*Hin*cII fragment containing the *P.furiosus gdh* gene was isolated from pLUW400 (Eggen *et al.*, 1994). Sticky ends were filled in using the Klenow fragment of *E.coli* DNA polymerase. The fragment was cloned into the blunted *Eco*RI restriction site of pPLc2833. The orientation of the *P.furiosus gdh* gene under the control of the P_L promoter was confirmed by restriction analysis, and the construct was denoted pLUW410. A 1.6 kb *Eco*RI–*Sph*I chromosomal fragment containing the *C.difficile gdh* gene (Lyerly *et al.*, 1991; kindly provided by Dr Lisa Barosso, Department of Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA) was cloned into an *Eco*RI–*Sph*I linearized pUC18 vector to position the *gdh* gene under the control of the *lac*UV5 promoter, resulting in pLUW420. Hybrid

genes were constructed according to a method described by Tomic *et al.* (1990) (Figure 1). Parts of the *P.furiosus* and *C.difficile gdh* genes encoding either domain I or II of GDH were amplified by PCR using 100 ng pLUW401 (pUC19 containing the *P.furiosus gdh* gene on a 1.6 kb *Hin*dIII–*Hin*cII fragment) or pLUW420 as template and, for each reaction, 100 ng universal vector primer and 200 ng of a gene-specific primer with a *Bsp*MI site (Table I). Restriction endonuclease *Bsp*MI digests DNA at a short distance from its recognition site, allowing the formation of complementary overhanging ends on the PCR products that may be ligated. In addition to the template and primers, the 100 µl reaction mixtures contained 0.2 mM dNTPs, *Taq* DNA polymerase buffer, 5 mM $MgCl_2$, 5 U *Taq* DNA polymerase and were subjected to 30 cycles of amplification (1.5 min at 92°C, 2.5 min at 50°C and 3.5 min at 72°C) on a programmable Ori-Block (New Brunswick Scientific). As determined by agarose gel electrophoresis, the amplification of the *P.furiosus* glutamate and cofactor binding domains resulted in the expected 0.6 and 1.0 kb products, respectively. Amplification of the *C.difficile* glutamate and cofactor binding domains also resulted in the expected 0.6 and 1.0 kb products, respectively. Each product was digested with *Bsp*MI and *Hin*dIII, *Hin*cII, *Eco*RI or *Sph*I that cut in the polylinker sequence (Figure 1). A three-point ligation of the 5′ *P.furiosus gdh* gene fragment and the 3′ *C.difficile gdh* gene fragment into *Hin*dIII–*Sph*I linearized pUC19 resulted in pLUW440. A three-point ligation of the 5′ *C.difficile gdh* gene fragment and the 3′ *P.furiosus gdh* gene fragment into *Eco*RI–*Hin*cII linearized pUC18 resulted in pLUW450. All constructions were performed in *E.coli* TG1.

Purification of GDH

GDH was purified from *P.furiosus* as described previously (Consalvi *et al.*, 1991). To isolate *P.furiosus* GDH from *E.coli*, strain K12ΔH1Δ*trp* harboring pLUW410 was cultured at 28°C until an OD_{600} of 0.5 was reached and incubated for another 2 h at 42°C. Subsequently, the culture was centrifuged for 10 min at 5000 r.p.m. and 4°C, medium was discarded and cells were resuspended in 20 mM potassium phosphate buffer, pH 6.5, 1 mM EDTA (buffer A) containing 1 µg/ml DNase. The suspension was sonified using a Branson sonifier and cell debris was removed by centrifugation at 13 000 r.p.m. for 10 min. The resulting supernatant was incubated for 10 min at 90°C and precipitated proteins were removed by centrifugation for another 10 min at 13 000 r.p.m. The sample was dialyzed against buffer A and applied to a 60 ml Matrex gel Red A column (Amicon) equilibrated with buffer A. After extensive washing with buffer A, GDH was eluted from the column using a step gradient of 120 ml of 0.10, 0.25, 0.35 and 1.00 M NaCl in buffer A. The fractions 0.25–1.00 M were pooled, concentrated and dialyzed against 0.1 M sodium phosphate buffer, pH 7.0, and contained pure GDH as judged by SDS–PAGE (12% polyacrylamide).

To purify *C.difficile* and the *P.furiosus–C.difficile* hybrid GDH, *E.coli* TG1 cultures harboring either pLUW420 or pLUW440 were centrifuged for 10 min at 5000 r.p.m. and 4°C. The medium was discarded and the cells were resuspended in 10 mM potassium phosphate buffer, pH 7.2, 1 mM EDTA (buffer B) containing 1 µg/ml DNase. The cells were broken in a French press under 1000 atmospheres pressure and cell debris was removed by centrifugation at 5000 r.p.m. for 20 min. Crude extracts were dialyzed against buffer B and applied to a Q Sepharose column (Pharmacia) equilibrated with buffer B. The column was washed until a stable baseline was recorded. The elution of GDH was performed by applying a gradient of 0.0–1.0 M NaCl in buffer B. Active fractions were pooled and dialyzed against buffer B. Ammonium sulfate was added up to 40% saturation and the precipitated proteins were removed by centrifugation for 10 min at 13 000 r.p.m. The sample was loaded onto a phenyl Sepharose column (Pharmacia) equilibrated with 40% ammonium sulfate saturated buffer B. After extensive washing using equilibration buffer, elution was performed by a 40–0% ammonium sulfate gradient in buffer B. Active fractions were pooled, dialyzed against 1.0 mM potassium phosphate buffer, pH 6.8 (buffer C), and applied to a hydroxyl apatite column equilibrated with buffer C. After extensive washing, elution was performed by applying a 1.0 mM–1.0 M gradient of potassium phosphate, pH 6.8. Active fractions were pooled and dialyzed against 0.1 M sodium phosphate buffer, pH 7.0. To obtain pure *P.furiosus* GDH produced by *E.coli* without heat incubation, the protocol described for the *C.difficile* GDH purification was followed. The purity of the samples was analyzed by SDS–PAGE.

GDH activity assay

GDH activity was assayed using a double-beam Kontron Uvikon 860 spectrophotometer at 340 nm and 25°C. The 1 ml standard assay reaction mixtures (reverse reaction, unless mentioned otherwise) contained 100 mM potassium phosphate buffer, pH 7.6, and 0.75 mM α-ketoglutarate, 75 mM ammonia and 100 µM NADPH for *P.furiosus* GDH; 4 mM α-ketoglutarate, 600 mM ammonia and 0.2 mM NADH for *C.difficile* GDH; and 6 mM α-ketoglutarate, 900 mM ammonia and 0.2 mM NADH for the hybrid GDH. Fixed substrate concentrations for glutamate oxidation (forward reaction) were 6 mM L-glutamate and 0.25 mM NADP for *P.furiosus* GDH; 50 mM L-glutamate and 4 mM NAD for *C.difficile* GDH; and 60 mM L-glutamate and 6 mM NAD for the hybrid GDH. The amount of enzyme was varied to obtain a rate of ~0.03 Δ absorbance/min. We define 1 U of enzyme activity as the amount of enzyme which produces 1 µmol of product per minute under standard conditions. Kinetic parameters were calculated by a nonlinear regression analysis of the data using ENZFITTER (Leatherbarrow, 1987). For each calculation, at least eight velocity/substrate data pairs were used. Substrate specificity was studied by measuring GDH activity in the presence of 20 and 100 mM D-glutamate, glycine, L-leucine, L-threonine, L-asparagine, L-valine, L-proline, L-aspartate, L-glutamine or L-norvaline and 5 and 25 mM oxaloacetate, glyoxylate, 2-oxobutyrate, 2-oxo-L-methyl-pentanoate, pyruvate or 2-oxovalerate.

Protein analysis

Protein concentrations were determined using Bradford reagent (Bio-Rad, Veenendaal, The Netherlands) with bovine serum albumin as a standard. Protein concentrations of pure samples were determined using the following extinction coefficients at 280 nm, calculated as described by Gill and von Hippel (1989): ε_m = 82 500 $M^{-1}cm^{-1}$ for *P.furiosus* GDH, ε_m = 66 000 $M^{-1}cm^{-1}$ for *C.difficile* GDH and ε_m = 86 910 $M^{-1}cm^{-1}$ for the hybrid GDH. Protein bands on SDS–PAGE gels were scanned and quantified using an Image Analysis System (Visionary, Fotodyne) equipped with ImageQuant (Molecular Dynamics) software. UV absorbance spectra were recorded at 20°C on a Perkin Elmer Lambda 16 spectrophotometer. The electrophoretic separation of proteins under native conditions was performed on 8% running gels. Apparent molecular masses were determined by an analytical gel filtration using a Superose 12

column (Pharmacia) as described previously (Consalvi *et al.*, 1991). Immunodetection was performed using antibodies raised against purified pyrococcal GDH (Eggen *et al.*, 1993). Protein microsequencing was carried out by cyclic Edman degradation on an Applied Biosystems amino acid sequenator model 477A (Eurosequence, Groningen, The Netherlands) using protein samples dialyzed against H_2O.

Effect of guanidinium chloride (GdmCl) on spectral properties

The enzymes (80 μg/ml) were incubated at 20°C for 24 h in 20 mM phosphate, pH 7.0, in the presence of increasing concentrations of GdmCl ranging from 0.0 to 7.2 M. UV absorbance spectra were monitored at 20°C in a Perkin Elmer Lambda 16 spectrophotometer. Fluorescence properties were studied at 20°C in a Perkin Elmer LS50B fluorimeter. Excitation was performed at 295 nm and emission was monitored at 300–400 nm. Far-UV CD spectra (210–250 nm) were recorded at 20°C in a Jasco J500A spectropolarimeter using a 0.2 cm cuvette. Fluorescence emission at 342 nm and observed ellipticities at 222 nm were normalized according to Pace (1986).

Thermostability

Enzyme thermostability was studied at 63, 65, 67 and 69°C for *C.difficile* GDH and hybrid GDH (5–50 μg/ml) and at 90 and 105°C for *P.furiosus* GDH (50 μg/ml). The enzymes were incubated in 20 mM sodium phosphate, pH 7.0, in sealed glass tubes for 0–24 h. Residual enzyme activity (forward reaction in the case of *C.difficile* GDH and hybrid GDH) was assayed at 25°C with respect to a control kept at room temperature. Slopes of first-order inactivation plots yielded rate constants for inactivation. These rate constants were used to calculate the free energy of the transition state for thermal inactivation ($\Delta G^{\ddagger}$) according to Segel (1975).

Results

Expression of the P.furiosus gdh gene in E.coli

High level overproduction of the *P.furiosus* GDH was obtained by cloning the *P.furiosus gdh* gene under the control of the λ P_L promoter on the expression vector pPLc2833 and introducing the resulting plasmid pLUW410 into *E.coli* ΔH1Δ*trp* that carries a temperature-sensitive *c*I857 repressor. Incubation at 42°C inactivates the *c*I857 repressor and the *P.furiosus gdh* gene is expressed. The *P.furiosus* GDH produced in this system amounted to up to 20% of the total cell protein in *E.coli* and showed a subunit molecular mass of 46 kDa, similar to GDH isolated from *P.furiosus* cells (Figure 2, lanes 2 and 3). Heat incubation of the cell-free extract at 90°C for 10 min and the subsequent removal of precipitated *E.coli* proteins resulted in highly purified *P.furiosus* GDH (Figure 2, lane 4). The residual traces of *E.coli* proteins were removed by bioaffinity chromatography on a Matrex gel Red A column (Figure 2, lane 5). The enzyme showed a strong cross-reaction with antibodies raised against purified pyrococcal GDH, and native PAGE and subsequent immunoblotting revealed that at least 90% of the GDH was in the hexameric conformation (results not shown). N-terminal sequencing (results not shown) indicated that >98% of the *P.furiosus* GDH produced in *E.coli* contained its N-terminal methionine that is also present in the majority of GDH purified from *P.furiosus* (Maras *et al.*, 1994).

Total GDH activity of the incubated cell-free extract samples increased ~3-fold during the heat treatment, indicating an activation of *P.furiosus* GDH or the inactivation of inhibitors

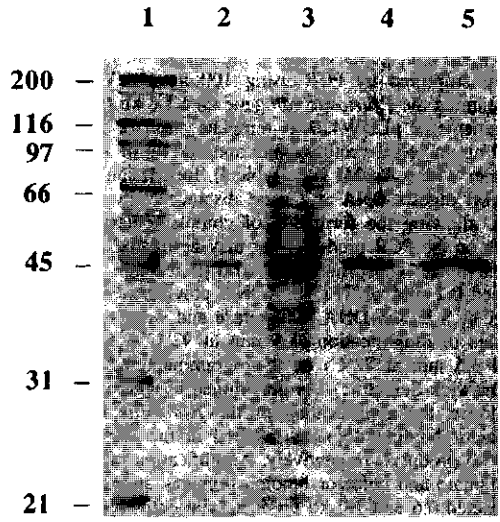


Fig. 2. SDS–PAGE of *P.furiosus*, *C.difficile* and hybrid GDH. Lane 1, broad range molecular weight marker (Bio-Rad); lane 2, GDH isolated from *P.furiosus*; lanes 3–5, *E.coli* harboring pLUW410 cell-free extract (lane 3), supernatant after heat incubation (lane 4) and Matrex gel Red A fraction (lane 5).

present in the cell-free extract. The latter possibility was excluded because the addition of cell-free extract from an *E.coli* culture containing pPLc2833 to the reaction mixture did not affect the specific activity of purified *P.furiosus* GDH. *Pyrococcus furiosus* GDH purified from *E.coli* without heat treatment showed an apparent molecular mass of 260 kDa on a gel permeation column. Heat treatment of this 260 kDa fraction increased its activity ~10-fold, indicating that *P.furiosus* GDH is present as partially active hexamers in the *E.coli* cell-free extract which attained a fully active conformation upon incubation at 90°C. No differences in secondary and tertiary structure between the partially and fully active conformations could be detected in UV absorbance difference spectra, fluorescence emission and far-UV CD spectra (results not shown). These results indicate that *P.furiosus* GDH produced in *E.coli* is identical to that produced in *P.furiosus*, allowing for its use in protein engineering studies.

Construction and expression of hybrid gdh genes

Hybrid *gdh* genes were constructed containing either the substrate binding domain I or the cofactor binding domain II of *P.furiosus* GDH and the complementary domain of *C.difficile* GDH. The latter GDH was chosen because the family II *C.difficile* and *P.furiosus* GDHs share 52% identical amino acid residues (Lyerly *et al.*, 1991; Eggen *et al.*, 1993). In addition, there are considerable differences in the thermal stability and kinetic properties of the two enzymes (see below). To maximize the possibility of obtaining hybrid enzymes with catalytic activity, the fusion was located in between two fully conserved glycine residues, that in the known GDH structure of *C.symbiosum* are located in the loop between β-sheet f and α-helix 9 connecting the two domains (corresponding to residues Gly200 and Gly201 in *C.symbiosum*; Baker *et al.*, 1992).

Based on two pUC19 derivatives, pLUW401 and pLUW420, carrying the *P.furiosus* and *C.difficile gdh* genes, respectively,

Table II. Kinetic parameters for *P.furiosus* GDH isolated from *P.furiosus* and *E.coli*, *C.difficile* GDH and hybrid GDH obtained at 25°C by varying the concentration of the respective substrate while keeping all other parameters fixed

Substrate	*P.furiosus* GDH		*P.furiosus* GDH (*E.coli*)		*C.difficile* GDH		Hybrid GDH	
	K_m (mM)	Catalytic efficiency[a] ($min^{-1}mg^{-1}$)	K_m (mM)	Catalytic efficiency[a] ($min^{-1}mg^{-1}$)	K_m (mM)	Catalytic efficiency[a] ($min^{-1}mg^{-1}$)	K_m (mM)	Catalytic efficiency[a] ($min^{-1}mg^{-1}$)
L-Glutamate	1.0	5.5	–	–	2.7	28	18.1	0.69
α-Ketoglutarate	0.039	405	0.029	600	0.16	5906	0.74	225
Ammonia	4.6[b]	3.4	5.1[b]	1.9	121[d]	7.9	228[d]	0.74
	24.0[c]	1.1	30.0[c]	0.6				
NAD	–	–	–	–	0.65	145	1.08	12
NADH	–	–	–	–	0.043	27 621	0.047	3932
NADPH	0.0030	6533	0.0050	4600	ND	ND	ND	ND

ND, not detectable; –, not determined.
[a]V_{max}/K_m.
[b]1.25–20 mM ammonia.
[c]30–150 mM ammonia.
[d]8–900 mM ammonia.

two reciprocal hybrid genes were constructed following a procedure described previously (Tomic *et al.*, 1990). The essential reactions included PCR amplification with appropriate primers that created a *Bsp*MI site (Table I) followed by digestion with *Bsp*MI and ligation into linearized pUC19 (Figure 1). The complete DNA sequence of each hybrid gene was determined and found to be identical to that expected from the *P.furiosus* and *C.difficile gdh* gene sequences (Lyerly *et al.*, 1991; Eggen *et al.*, 1993). The resulting plasmids pLUW440 and pLUW450 (Figure 1) were each introduced into *E.coli* strain TG1. Cell-free extracts were assayed for GDH activity and compared with those obtained with the parental plasmids carrying the wild-type genes.

The activity of the hybrid GDH specified by pLUW440, containing the substrate binding domain I of *P.furiosus* and the cofactor binding domain II of *C.difficile*, could be detected using NAD(H) as cofactor. No activity could be detected with NADP(H). This *P.furiosus–C.difficile* hybrid enzyme amounted to up to 1% of total cell protein according to SDS–PAGE, possessed a subunit molecular mass of 46 kDa like the *P.furiosus* GDH, and showed a weak cross-reaction with antibodies raised against pyrococcal GDH (data not shown).

The complementary hybrid GDH specified by pLUW450, containing domain I of *C.difficile* GDH and domain II of *P.furiosus* GDH, also amounted to up to 1% of total cell protein according to SDS–PAGE, had a molecular mass of 45 kDa like *C.difficile* GDH, and showed a strong cross-reaction with the pyrococcal GDH antibodies (results not shown). However, despite its efficient production, this hybrid GDH showed no activity above the *E.coli* background GDH activity. In contrast, active *C.difficile* GDH was highly over-produced in *E.coli* TG1 harboring pLUW420, the parental plasmid that only differs from pLUW450 by the absence of *P.furiosus* DNA encoding domain II (results not shown; Lyerly *et al.*, 1991).

The active *P.furiosus–C.difficile* hybrid GDH specified by pLUW440 and the *C.difficile* GDH were subsequently purified to homogeneity, as judged by SDS–PAGE, and used for further characterization.

Kinetic parameters and substrate specificity

Pyrococcus furiosus GDH was purified from either *P.furiosus* or *E.coli* and used to determine the kinetic parameters as well as the substrate and cofactor specificities that were compared with each other and those obtained with purified *C.difficile* or *P.furiosus–C.difficile* hybrid GDH (Table II).

Similar K_m values for α-ketoglutarate, ammonia and NADPH were obtained with *P.furiosus* GDH isolated from *P.furiosus* and *E.coli*. Even the Lineweaver–Burk plots that were obtained with the *P.furiosus* GDH from either source showed similar curvature for ammonia, indicating that at low concentrations the affinity for ammonia was higher than at high concentrations. The affinity for ammonia of the *P.furiosus* GDH from both sources at 25°C was the same as that reported previously at 60°C (Consalvi *et al.*, 1991). However, the K_m values of the *P.furiosus* GDH for NADPH and notably α-ketoglutarate appeared to be temperature dependent because those found at 25°C were four and nine times lower, respectively, than those reported previously at 60°C (Consalvi *et al.*, 1991). The calculated V_{max}/K_m values of the *P.furiosus* GDH isolated from either *P.furiosus* or *E.coli* were similar, indicating that this GDH showed a host-independent specific activity and efficiency.

The purified *C.difficile* GDH followed Michaelis–Menten kinetics for NAD, NADH and ammonia. No activity was found with phosphorylated cofactors. For L-glutamate and α-ketoglutarate, substrate inhibition was observed above 50 and 4 mM, respectively. *Clostridium difficile* GDH was unable to convert any of the tested substrate analogs.

The hybrid *P.furiosus–C.difficile* GDH also followed Michaelis–Menten kinetics only for NAD and NADH. For L-glutamate and α-ketoglutarate, substrate inhibition was observed for this enzyme above 60 and 6 mM, respectively. No saturation was observed with the hybrid GDH for ammonia up to 900 mM. V_{max}/K_m values indicated that the hybrid GDH was much less efficient than both parental enzymes. Substrate binding appeared to be affected in the hybrid because K_m values for glutamate, α-ketoglutarate and ammonia were much higher than for both parental enzymes (Table II). In addition, the hybrid was found to be more specific than GDH isolated from *P.furiosus* because it was not able to convert L-norvaline or 2-oxovalerate, while the *P.furiosus* GDH was able to utilize any of these substrate analogs (Consalvi *et al.*, 1991). Comparable K_m values for NAD and NADH were found with the hybrid and *C.difficile* GDH, while neither showed activity with the phosphorylated cofactor. The identical cofactor affinity and specificity of the *P.furiosus–C.difficile* hybrid and the *C.difficile* GDH indicate that these properties are solely determined by the cofactor binding domain.

Optimum temperature

Pyrococcus furiosus GDH isolated from *E.coli* showed an identical temperature optimum of 100°C as the enzyme isolated from *P.furiosus* (Consalvi *et al.*, 1991; results not shown). This temperature optimum was much higher than that observed with the *P.furiosus–C.difficile* hybrid GDH that showed a similar optimum to the *C.difficile* GDH for glutamate formation (56–63°C). For glutamate breakdown, the optimum temperature of the hybrid GDH was slightly higher than that of *C.difficile* GDH (61–64 and 57–61°C, respectively; Figure 3). This indicates that exchange of the *P.furiosus* cofactor binding domain for that from *C.difficile* lowers the optimum temperature of the resulting hybrid enzyme to the level obtained with the GDH from the mesophilic parent.

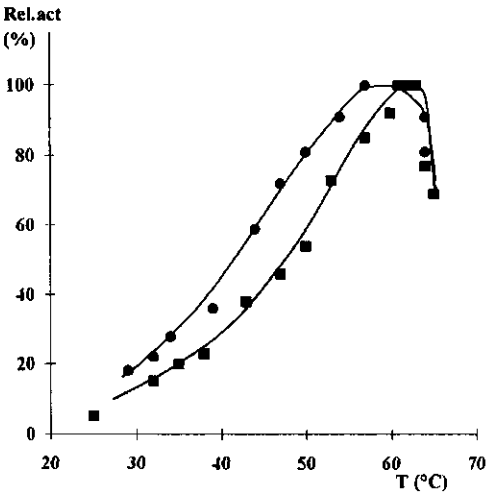


Fig. 3. Temperature dependence of *C.difficile* GDH (●) and hybrid GDH (■) for glutamate breakdown.

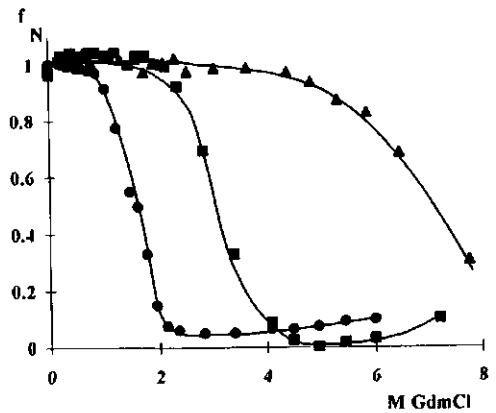


Fig. 4. GdmCl-induced denaturation of *C.difficile* GDH (●), hybrid GDH (■) and *P.furiosus* GDH (▲) followed by fluorescence emission at 342 nm. The fraction of GDH present in the native conformation (f_N) is plotted against the GdmCl concentration.

Influence of GdmCl on spectral properties

The effect of GdmCl on fluorescence emission intensity and ellipticity of the *P.furiosus*, *C.difficile* and hybrid *P.furiosus–C.difficile* GDHs was studied. Identical denaturation curves were obtained with both spectroscopic techniques. Transition midpoints were found at 6.6, 3.2 and 1.5 M GdmCl for *P.furiosus* GDH, the hybrid and *C.difficile* GDH, respectively (Figure 4). The loss of secondary structure was accompanied by a loss of ability to regain catalytic activity after dilution of the denaturant because transition midpoints found by fluorescence and CD were similar to those found in reactivation experiments in which GdmCl was removed after 24 h of incubation by an at least 50-fold dilution (data not shown).

Thermostability

The thermostability of *P.furiosus* GDH purified from *P.furiosus* or *E.coli* was studied at 90 and 105°C (Table III). The 260 kDa gel permeation fraction containing *P.furiosus* GDH that was purified from *E.coli* without heat treatment lost its activity almost completely within several minutes of incubation at 105°C (data not shown). However, *Pyrococcus furiosus* GDH that was purified from *E.coli* using heat incubation showed an initial fast inactivation (loss of >50% activity within several minutes) and a second much slower inactivation at 105°C. The half-life of this second inactivation process was similar to that of GDH purified from *P.furiosus* (~1 h; Table III). At 90°C, there was only a small initial loss of activity of the heat-treated GDH purified from *E.coli*. *Pyrococcus furiosus* GDH isolated from *E.coli* showed a similar half-life at this temperature to that isolated from *P.furiosus* (~6 h; Table III).

The thermostability of *C.difficile* GDH and the *P.furiosus–C.difficile* hybrid GDH was determined at 63, 65, 67 and 69°C. At all these incubation temperatures the hybrid *P.furiosus–C.difficile* GDH showed a 2-fold shorter half-life than the *C.difficile* GDH, irrespective of the protein concentration (data not shown). At 65°C, this difference in stability amounted to up to 1.89 kJ mol^{-1} (Table III). Both *P.furiosus* and *C.difficile* GDH remained soluble after thermal inactivation, whereas centrifugation of the thermally inactivated hybrid resulted in the

Table III. Half-life and $\Delta G^{\ddagger}$ values for the thermal inactivation of *P.furiosus* GDH isolated from *P.furiosus* and *E.coli* at 90 and 105°C, and of *C.difficile* GDH and hybrid GDH at 65°C

	65°C		90°C		105°C	
	$t_{1/2}$ (h)	$\Delta G^{\ddagger}$ (kJ mol^{-1})	$t_{1/2}$ (h)	$\Delta G^{\ddagger}$ (kJ mol^{-1})	$t_{1/2}$ (h)	$\Delta G^{\ddagger}$ (kJ mol^{-1})
C.difficile GDH	5.1	111.70	–	–	–	–
Hybrid GDH	2.6	109.81	–	–	–	–
P.furiosus GDH	–	–	5.5	120.41	1.20	120.72
P.furiosus GDH (*E.coli*)	–	–	6.4	120.87	0.86	119.68

$\Delta G^{\ddagger}$ values were calculated according to Segel (1975). –, not determined.

complete precipitation of the aggregated enzyme presumably because of the exposure of buried hydrophobic residues.

Discussion

We have reported previously the functional expression of the *P.furiosus gdh* gene in *E.coli* under the control of the λ P_L promoter (Eggen *et al.*, 1994). After further optimization described here, up to 20% of the total *E.coli* cell protein could be identified as *P.furiosus* GDH. After heat treatment, the enzyme was fully active, suggesting that it needs a temperature closer to the *in vivo* growth optimum of *P.furiosus* to be able to adopt its correct conformation. *Pyrococcus furiosus* GDHs isolated from *P.furiosus* and *E.coli* have similar apparent molecular weights, N-terminal sequences, subunit compositions, specific activities, thermoactivities, kinetic parameters and both chemical and thermal stabilities. This confirms that the overproduced enzyme can be used to study molecular adaptations to high temperatures.

Pyrococcus furiosus GDH produced by *E.coli* appears to be partially active in untreated cell-free extract. Native PAGE and subsequent immunodetection showed that at least 90% of the GDH formed in *E.coli* is present in a hexameric conformation. Furthermore, the hexameric GDH fraction obtained by gel filtration showed a clear increase in specific activity after heat treatment. These results strongly suggest the presence of partially active, hexameric GDH in the *E.coli* cell-free extract, which upon heat treatment arranges into a fully active structure. Furthermore, the enzyme isolated from *E.coli* without heat treatment is not fully thermostable because it showed an almost complete loss of activity upon incubation at 105°C, while the heat-incubated enzyme showed a considerably longer half-life (almost 1 h) at this temperature. This implies that either the *P.furiosus* GDH is present in *E.coli* in two conformations, one fully thermostable and one less thermostable, or that the enzyme always assumes a partially active conformation which is lost during its purification from *P.furiosus* but maintained in the purification method applied to *E.coli*. At present, we cannot discriminate between these two possibilities. It was reported recently that the expression of the *P.furiosus gdh* gene as a transcriptional and translational fusion using the T7 system in *E.coli* resulted in ~15% GDH of total cell protein. However, this overproduced *P.furiosus* GDH consisted of 50% inactive monomeric enzyme that assembled into hexamers upon heat incubation (Diruggiero and Robb, 1995). The correct hexameric subunit assembly in *E.coli*, which is reported here when the *P.furiosus gdh* gene is expressed under the control of the λ P_L promoter, might be caused by the activity of *E.coli* heat-shock proteins that are induced upon incubation at 42°C. However, more trivial explanations may also account for the previously observed incorrect folding in *E.coli* because the *P.furiosus gdh* gene used in those studies was PCR amplified and may contain unforeseen mutations (Diruggiero and Robb, 1995).

Exchanging parts of the polypeptide chain between homologous enzymes with different characteristics can be very useful in studying the role and localization of residues involved in achieving thermostability. The cloning and successful expression of the *P.furiosus gdh* gene allowed us to exchange domains between this GDH and GDH from any other organism to study the effects on its functionality and stability. The phylogenetically very distant mesophilic bacterium *C.difficile* was chosen to be the partner in the domain exchange. The gene fusion was located between two conserved glycine residues in the loop connecting the two domains, so as to disturb as little as possible the folding of the resulting hybrids in comparison with the parental enzymes.

The activity of the *C.difficile–P.furiosus* hybrid GDH, containing the substrate binding domain I of *C.difficile* GDH and the cofactor binding domain II of *P.furiosus* GDH, could not be detected above the *E.coli* background, although it was clearly present in *E.coli* cell-free extract as judged by SDS–PAGE and immunoblotting. Apparently, the correct folding and/or assembly into an active hexameric structure was impaired in this mutant. In contrast, the complementary *P.furiosus–C.difficile* hybrid, with the substrate binding domain I from *P.furiosus* GDH and the cofactor binding domain II of *C.difficile* GDH, showed significant GDH activity when produced in *E.coli*. According to gel permeation chromatography, it had the same apparent molecular weight as *P.furiosus* GDH and *C.difficile* GDH, indicating a hexameric assembly similar to that of the parental enzymes. The *P.furiosus–C.difficile* hybrid enzyme is less efficient than both parental GDHs, as indicated by its kinetic parameters. Similar affinities for NAD and NADH for the hybrid and *C.difficile* GDH, and the inability to use the phosphorylated cofactor, indicate that binding of the cofactor is not altered in the hybrid that contains domain II from *C.difficile* GDH. This result agrees with the structures of the binary and tertiary complexes of *C.symbiosum* GDH, which show that the cofactor is almost exclusively bound by domain II (Baker *et al.*, 1992; Stillman *et al.*, 1993). Moreover, this result indicates that the cofactor binding domain II of GDH is structurally independent of domain I with respect to the binding of the cofactor. A similar independent activity was observed with the cofactor binding domain of the glyceraldehyde-3-phosphate dehydrogenase from *T.maritima* which, after overproduction in *E.coli*, retained its capacity to bind NAD and NADH (Jecht *et al.*, 1994). Both dehydrogenases show a similar domain arrangement, with the main difference being the orientation of the cofactor binding domain that is C-terminal in GDH and N-terminal in glyceraldehyde-3-phosphate dehydrogenase.

The low catalytic efficiency of the *P.furiosus–C.difficile* hybrid enzyme is very likely caused by less efficient substrate binding, as indicated by the higher K_m values for L-glutamate and α-ketoglutarate as well as ammonia compared with both parental enzymes. Interactions in the pocket on the surface of domain I, where the substrate side chain is bound (Stillman *et al.*, 1993), are not expected to be changed significantly because they are solely made by residues from domain I and the hybrid is even more specific towards substrate analogs than *P.furiosus* GDH. It is possible that the interdomain interactions involved in positioning the two domains in the correct orientation with respect to each other (and thus the positioning of substrate and cofactor) are less efficient in the hybrid, although preliminary modeling studies indicated that these interactions are considerably conserved in *C.difficile* and *P.furiosus* GDH.

The *C.difficile–P.furiosus* hybrid GDH shows a thermoactivity that is only slightly higher than that of the mesophilic parental enzyme. Apparently, the presence of a domain from a hyperthermostable enzyme does not significantly enhance the temperature at which the enzyme becomes susceptible towards thermal inactivation during catalysis, and thermoactivity might be determined by the weakest or more flexible part of the protein. It may be envisaged that the mesophilic domain reaches its maximum flexibility at the temperature optimum and starts unfolding, while the thermostable domain

is still too rigid to be able to bind and release substrates efficiently, explaining the low catalytic efficiency of the hybrid. Although the thermoactivity of the hybrid is relatively low, it is more stable towards chemical denaturation than its mesophilic parental enzyme, most likely because of the presence of the *P.furiosus* glutamate binding domain. The apparent two-state cooperative transition profile, obtained by fluorescence emission and ellipticity measurements of GdmCl-denatured hybrid GDH, is the same as for *C.difficile* GDH. Although kinetic parameters suggest that both domains are structurally independent and may be able to fold independently, the hybrid unfolds without any detectable preliminary loss of secondary structure of the *C.difficile* cofactor binding domain.

The effect of domain exchange on thermostability is dramatic. Instead of obtaining a hybrid displaying a thermostability in between those of the parental enzymes, as found in the case of chemical stability, the hybrid is even less thermostable than *C.difficile* GDH. Enzymes from (hyper)-thermophilic organisms are usually characterized by a high stability towards physical and chemical denaturation (Fontana, 1991). In the case of the hybrid, the increased stability towards GdmCl with respect to *C.difficile* GDH is not accompanied by a corresponding increase in thermostability. It should be noted that domain I of the hybrid does not consist completely of *P.furiosus* residues; the C-terminal α-helix 17 (numbering as in *C.symbiosum* GDH; Baker *et al.*, 1992) crosses over from domain II and is composed of *C.difficile* residues. Preliminary modeling studies indicate that the efficient packing of this α-helix against the *P.furiosus* domain is impaired because the possibility of forming hydrogen bridges is reduced and several residues will cause steric hindrance. The importance of a C-terminal α-helix for thermostability has been shown previously with the glyceraldehyde-3-phosphate dehydrogenase because its removal led to a dramatic decrease in the thermostability of the *M.fervidus* enzyme (Biro *et al.*, 1990). Furthermore, the packing of this C-terminal helix in GDH may also have effects on thermoactivity and catalytic efficiency because it is involved in orientating the two domains with respect to each other. The construction of hybrid GDHs with an additional exchange of this α-helix is in progress to allow us to study the possibility of obtaining more efficient and thermostable enzymes.

In conclusion, we have demonstrated that it is possible to successfully exchange parts of homologous proteins from micro-organisms that are phylogenetically very distant and belong to the archaea and bacteria. Furthermore, it has been shown that thermoactivity and thermostability are not necessarily linked to chemical stability and may be engineered independently, offering the possibility of constructing enzymes that are optimally active at low temperatures while retaining an extreme chemical stability.

Acknowledgements

We are grateful to Professor David Rice, Linda K.Britton and Timothy J.Stillman for their help with the protein modeling. This work was partially supported by contract BIOT-CT93-0274 of the European Union.

References

Baker,P.J., Britton,K.L., Engel,P.C., Farrants,G.W., Lilley,K.S., Rice,D.W. and Stillman,T.J. (1992) *Proteins: Struct. Funct. Genet.*, **12**, 75–86.

Benachenhou-Lahfa,N., Forterre,P. and Labedan,B. (1993) *J. Mol. Evol.*, **36**, 335–346.

Bernard,H.U., Remaut,E., Hershfield,M.V., Das,H.K., Helmshi,D.D., Yanofsky,C. and Franklin,N. (1979) *Gene*, **5**, 59–76.

Biro,J., Fabry,S., Dietmaier,W., Bodegain,C. and Hensel,R. (1990) *FEBS Lett.*, **275**, 130–134.

Chan,M.K., Mukund,S., Kletzin,A., Adams,M.W.W. and Rees,D.C. (1995) *Science*, **267**, 1463–1469.

Consalvi,V., Chiaraluce,R., Politi,L., Vaccaro,R., de Rosa,M. and Scandurra,R. (1991) *Eur. J. Biochem.*, **202**, 1189–1196.

Day,M.W., Hsu,B.T., Joshua-Tor,L., Park,J., Zhou,Z.H., Adams,M.W.W. and Rees,D.C. (1992) *Protein Sci.*, **1**, 1494–1507.

Diruggiero,J. and Robb,F.T. (1995). *Appl. Environ. Microbiol.*, **61**, 159–164.

Eggen,R.I.L., Geerling,A.C.M., Waldkötter,K., Antranikian,G. and de Vos,W.M. (1993) *Gene*, **132**, 143–148.

Eggen,R.I.L., Geerling,A.C.M., Voorhorst,W.G.B., Kort,R. and de Vos,W.M. (1994) *Biocatalysis*, **11**, 131–141.

Fontana,A. (1991) In de Presco,G. (ed.), *Life under Extreme Conditions: Biochemical Adaptation*. Springer-Verlag, Berlin, Germany, pp. 89–113.

Gibson,T.J. (1984) PhD thesis, Cambridge University, Cambridge, UK.

Gill,S.C. and von Hippel,P.H. (1989) *Anal. Biochem.*, **182**, 319–326.

Goldenberg,D.D. (1992) In Creighton,T.E. (ed.), *Protein Folding*. W.H. Freeman and Co., NY, pp. 353–403.

Jaenicke,R. (1991) *Eur. J. Biochem.*, **202**, 715–728.

Jecht,M., Tomschy,A., Kirscher,K. and Jaenicke,R. (1994) *Protein Sci.*, **3**, 411–418.

Kataoka,K., Takada,H., Tanizawa,K., Yoshimura,T., Esaki,N., Ohsima,T. and Soda,K. (1994) *J. Biochem.*, **116**, 931–936.

Kengen,S.W.M. and Stams,A.J.M. (1994) *Arch. Microbiol.*, **161**, 168–175.

Korndörfer,I., Steipe,B., Huber,R., Tomschy,A. and Jaenicke,R. (1995) *J. Mol. Biol.*, **246**, 511–521.

Leatherbarrow,R.J. (1987) *ENZFITTER*. Elsevier-Biosoft, Cambridge, UK.

Lyerly,D.M., Barosso,L.A. and Wilkins,T.D. (1991) *J. Clin. Microbiol.*, **29**, 2639–2642.

Maras,B., Valiante,S., Chiaraluce,R., Consalvi,V., Politi,L., de Rosa,M., Bossa,F., Scandurra,R. and Barra,D. (1994) *J. Protein Chem.*, **13**, 253–259.

Onodera,K., Sakurai,M., Moriyama,H., Tanaka,N., Numata,K., Ohshima,T., Sato,M. and Katsube,Y. (1994) *Protein Engng*, **7**, 453–459.

Pace,C.N. (1986) *Methods Enzymol.*, **131**, 266–280.

Remaut,E., Stanssens,P. and Fiers,W. (1981) *Gene*, **15**, 81–93.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Segel,I.H. (1975) *Enzyme Kinetics*. Wiley, New York, pp. 844–944.

Smith,E.L., Austen,B.M., Blumenthal,K.M. and Nyc,J.F. (1975) In Boyer, P.D. (ed.), *The Enzymes*. Academic Press, New York, Vol. XI, pp. 293–367.

Stetter,K.O., Fiala,G., Huber,G., Huber,R. and Segerer,A. (1990) *FEMS Microbiol. Rev.*, **75**, 117–124.

Stillman,T.J., Baker,P.J., Britton,K.L. and Rice,D.W. (1993) *J. Mol. Biol.*, **234**, 1131–1139.

Tomic,M., Sunjevaric,I., Savtchenko,E.S. and Blumenberg,M. (1990) *Nucleic Acids Res.*, **18**, 1656.

Tomschy,A., Böhm,G. and Jaenicke, R. (1994) *Protein Engng*, **7**, 1471–1478.

Yannisch-Perron,C., Viera,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received July 13, 1995; revised September 29, 1995; accepted October 6, 1995

Chapter 4

Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the Glutamate Dehydrogenase family.

Kitty S. Yip, K. Linda Britton, Timothy J. Stillman, Joyce H. G. Lebbink, Willem M. de Vos, Frank T. Robb, Carlos Vetriani, Dennis Maeder, and David W.Rice.

Reprinted with permission from
European Journal of Biochemistry **255**, 336-346, 1998

Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the Glutamate Dehydrogenase family

Kitty S. P. YIP[1], K. Linda BRITTON[1], Timothy J. STILLMAN[1], Joyce LEBBINK[2], Willem M. DE VOS[2], Frank T. ROBB[3], Constantino VETRIANI[3], Dennis MAEDER[3] and David W. RICE[1]

[1] The Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK
[2] Wageningen Agricultural University, Department of Microbiology, Wageningen, The Netherlands
[3] Centre of Marine Biotechnology, University of Maryland, Baltimore, USA

(Received 22 January/9 April 1998) – EJB 98 0097/3

The recent structure determination of glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus* and the comparison of this structure with its counterparts from the mesophiles *Clostridium symbiosum* and *Escherichia coli* has highlighted the formation of extended networks of ion-pairs as a possible explanation for the superior thermal stability of the hyperthermostable enzyme. In the light of this, we have carried out a homology-based modelling study using sequences of a range of glutamate dehydrogenases drawn from species which span a wide spectrum of optimal growth temperatures. We have attempted to analyse the extent of the formation of ion-pair networks in these different enzymes and tried to correlate this with the observed thermal stability. The results of this analysis indicate that the ion-pair networks become more fragmented as the temperature stability of the enzyme decreases and are consistent with a role for the involvement of such networks in the adaptation of enzymes to extreme temperatures.

Keywords: glutamate dehydrogenase; thermal stability; homology modeling; ion-pair network; Archaea.

A proper understanding of the molecular basis of thermal stability in proteins could have important consequences for their application in a range of biotechnological processes. For example, the availability of enzymes with the appropriate specificity and capable of surviving for long periods at extreme temperatures could lead to the creation of novel applications of enzyme-based technology in industries such as those involved in the processing of paper, pulp or fibres [1–2]. The opportunities afforded by such applications have generated significant interest in the field and a large number of studies have been undertaken to unravel the molecular mechanisms involved in generating thermostable enzymes. Recently, increasing attention has been focused on proteins from Archaea, a phylogenetically distinct evolutionary kingdom that includes many extremophilic microorganisms and whose study has expanded our horizons on the limits of biodiversity [3]. This genus includes hyperthermophiles that thrive at temperatures above the boiling point of water and studies on proteins from such organisms, which are necessarily more thermostable, are expected to provide clues to understanding the molecular basis of thermal stability. The crystallographic analysis of a number of proteins from hyperthermophiles to mesophiles has permitted direct structural comparisons, homology-based modelling studies and site-directed mutagenesis to be undertaken, combining sequence data on extremophilic enzymes with structural data from their mesophilic counterparts. However, to date, these analyses have not produced a consistent picture on the origins of thermal stability and suggestions put forward to explain enhanced stability properties of proteins have included decreased flexibility of the protein [4–5] arising from changes in interactions such as the increase in the number of ion-pairs [6–9], an increase in packing density [10], a decrease in the degree of cavity formation [11–12], a decrease in the sizes of loops linking secondary-structure elements [11], an increase in alanine residues located at the termini of surface α helices [13], a reduction in exposed surface area [14], an increase in proline content [15–16], an increase in the number of disulphide bonds [17] and an increase in hydrophobic interactions at subunit interfaces [18–20].

Equally, a number of lines of evidence have been used to discount the importance of some of these factors. For example, the replacement by mutagenesis of three charged residues which form part of a small ion-pair network in the Arc repressor by three hydrophobic residues resulted in an increase in stability [21]. Furthermore, the comparative structural analysis of the neutral protease from *Bacillus cereus* and the thermostable enzyme thermolysin failed to reveal a dramatic increase in the number of ion-pairs, suggesting that this type of interaction is of little importance in generating stability [22]. One possible explanation for discrepancies in the relative importance of individual factors to stability in different proteins could be that thermal stability arises from a combination of different molecular interactions that are differentially weighted on a case by case basis. However, since the context in which a particular type of interaction occurs is of profound importance in determining the

Correspondence to D. W. Rice, The Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, S10 2TN, UK
Phone: +44 114 2824242.
Fax: +44 114 2728697.
E-mail: d.rice@sheffield.ac.uk

Abbreviations. GluDH, glutamate dehydrogenase; Cs, *Clostridium symbiosum*; Pf, *Pyrococcus furiosus*; Pe, *Pyrococcus endeavori*; Tl, *Thermococcus litoralis*; Ss, *Sulfolobus solfataricus*; Tm, *Thermotoga maritima*; Cd, *Clostridium difficile*; Pa, *Peptostreptococcus asaccharolyticus*; Nc, *Neurospora crassa*; Ec, *Escherichia coli*.

Enzyme. Glutamate dehydrogenase (EC 1.4.1.2–4).

contribution it might make to stability, it is often hard to assess the subtle influence of small structural changes [23].

Glutamate dehydrogenase (GluDH) is one of the most abundant proteins in many Archaea and the extensive biochemical literature on mesophilic GluDHs [24] has led to its adoption as a model system for the analysis of thermal stability and to the sequencing of this enzyme from a range of species which span a wide spectrum of thermal stability [25]. Analysis of the sequences of these enzymes shows that they are all closely related (Table 1). The GluDH isolated from *Pyrococcus furiosus* (Pf), an organism that grows at temperatures up to 103°C [26] lies at one extreme of temperature stability with a half-life at 100°C of 12 h [27]. In contrast, the GluDH from *Clostridium symbiosum* (Cs) is much less stable losing 56% of its activity following incubation for 23 min at 52°C [28]. The analysis of the structure of Pf GluDH has shown that the hyperthermostable enzyme contains a striking series of ion-pair networks on the surface of the protein subunits and partially buried at inter-subunit and inter-domain interfaces, not found in the equivalent mesophilic enzymes [29–30]. In a attempt to examine the variation in the extent of the ion-pair networks across the family of GluDHs we have carried out an homology-based modelling study using the available structures of both the mesophilic (Cs and Ec) and hyperthermostable (Pf) GluDHs as a guide. Here we report the correlation of the extent of ion-pair formation across the GluDH family with the enzyme thermal stability and discuss the implications of this for a general understanding of this phenomenon.

MATERIALS AND METHODS

Sequence alignment and stability ranking. Ten hexameric GluDHs from species which span the temperature spectrum from mesophiles to hyperthermophiles were approximately ranked in terms of their thermal stability using a combination of data drawn from studies on the melting temperatures, the residual activity after incubation of the enzyme for a fixed period at different temperatures and the maximal growth temperature of the organism [*Clostridium symbiosum* (Cs) GluDH [28], *Escherichia coli* (Ec) GluDH [31], *Neurospora crassa* (Nc) [32], *Clostridium difficile* (Cd) [33, 34], *Peptostreptococcus asaccharolyticus* (Pa); Pasquo, A., Consalvi, V. and Scandurra, R., unpublished results, *Thermotoga maritima* (Tm) GluDH [35], *Sulfolobus solfataricus* (Ss) GluDH [36], *Thermococcus litoralis* (Tl) GluDH [37], *Pyrococcus endeavori* (Pe also known as ES4) GluDH [38] and *Pyrococcus furiosus* (Pf) GluDH [39]]. Their amino acid sequences were aligned against one another by visual inspection (Fig. 1), maximising the high level of sequence similarity seen across the family of enzymes (Table 1) and using the three-dimensional structures of the *C. symbiosum* (Cs) [40], *E. coli* (Ec) [30–31], *N. crassa* (Nc) (Stillman, T. J., unpublished results) and *P. furiosus* (Pf) [29] enzymes as a further guide. Throughout this report the Pf GluDH sequence numbering is used to identify equivalent residues in the other GluDH sequences unless otherwise stated.

Homology modelling. The fold of the structures of Cs and Pf GluDH, enzymes which lie towards the extreme ends of the temperature-stability spectrum considered here, are very similar. In each subunit the polypeptide chain is folded into two domains separated by a deep cleft. The N-terminal portion of the chain (Fig. 2) folds to form domain I which contains most of the residues responsible for forming interactions across the subunit interfaces of the hexamer. Domain II contains the binding site for the $NAD(P)^+$ moiety and is formed from residues at the C-terminal end of the polypeptide chain (Fig. 2). There are two significant regions of difference in both sequence and structure between these two enzymes. The first of these occurs at the N-terminus, where the poor sequence similarity and difference in length between the enzymes corresponds to structural differences in a small helical subdomain which forms part of domain I. The second difference lies in domain II and arises from structural changes associated with a region close to the adenine-ribose-binding site (residues 251–290 in the Pf enzyme). The structural differences in the latter region are thought to be involved in the different nucleotide specificity exhibited by members of this family [29] which range from the strict NAD^+- or $NADP^+$-dependence of some of the bacterial GluDHs (for example, the NAD^+-dependent GluDH from Cs [41], or the $NADP^+$-dependent GluDH from Ec [42] to dual specificity of some of the vertebrate or archaeal enzymes (for example, the dual specificity GluDHs from Pf [27] and bovine liver [43]). With the exception of the above regions, the structural equivalence of the residues in the four structures is good and, using one or other of the structures the sequences of the proteins which possess intermediate temperature stability, can be modelled with confidence.

RESULTS AND DISCUSSION

Analysis of residue composition in regions of high charge density. The extensive nature of the ion-pair networks in Pf GluDH arise from interactions in the assembled enzyme of a patch on the subunit surface that contains a high proportion of charged residues (Fig. 3a). In particular, the most extensive networks arise from association of residues from the enzyme's N-terminal domains. This interaction gives rise to the formation of three unique ion-pair networks including the two largest ones [cluster 1 (18 residues) and cluster 2 (6 residues)] and also a smaller cluster of three residues (cluster 1A) that lies close to cluster 1. This extensive patch involves 25 residues, 17 of which are charged (Arg35, Glu44, Asp46, Arg117, Glu120, Arg121, Arg124, Arg128, Asp132, Glu138, Asp157, Glu158, Glu160, Arg164, Arg165, Lys166 and Arg414), and which arise (primarily) from segments of the polypeptide chain including $\alpha 4$, $\alpha 5$, βa, βb, $\alpha 14$, βe and the loop between $\alpha 5$ and βf (Fig. 2). Overall, the solvent-accessible area of the monomer covered by the 25 residues in these two patches of high charge density is approximately 2100 Å^2.

Cluster 1 is the most extensive ion-pair network in Pf GluDH and involves the interaction of 18 charged residues. It is formed by the interaction of residues from four monomers across the twofold axis relating dimers in the hexameric enzyme with 16 of the 18 charged residues belonging to one or other of the

Fig. 1. Alignment of representative sequences for GluDH from species that cover a broad spectrum of growth temperatures. The enzymes chosen cover a number of mesophilic, thermophilic and hyperthermophilic organisms. The residue numbering quoted at the top and bottom of the sequence alignment is for the Cs and Pf GluDHs, respectively. Positions of insertion and deletion are indicated as a dash. The 24 residues which form the patch of high charge density in the Pf enzyme and which can be equivalenced across the entire enzyme family are highlighted in black boxes. References for the GluDH sequences given are: *Clostridium symbiosum* (Cs) [45]; *Escherichia coli* (Ec) [31]; *Neurospora crassa* (Nc) [46]; *Clostridium difficile* (Cd) [47]; *Peptostreptococcus asaccharolyticus* (Pa) [48]; *Thermotoga maritima* (Tm) [35]; *Sulfolobus solfataricus* (Ss) [36]; *Thermococcus litoralis* (Tl) [37], *Pyrococcus endeavori* (Pe) [38]; *Pyrococcus furiosus* (Pf) [49].

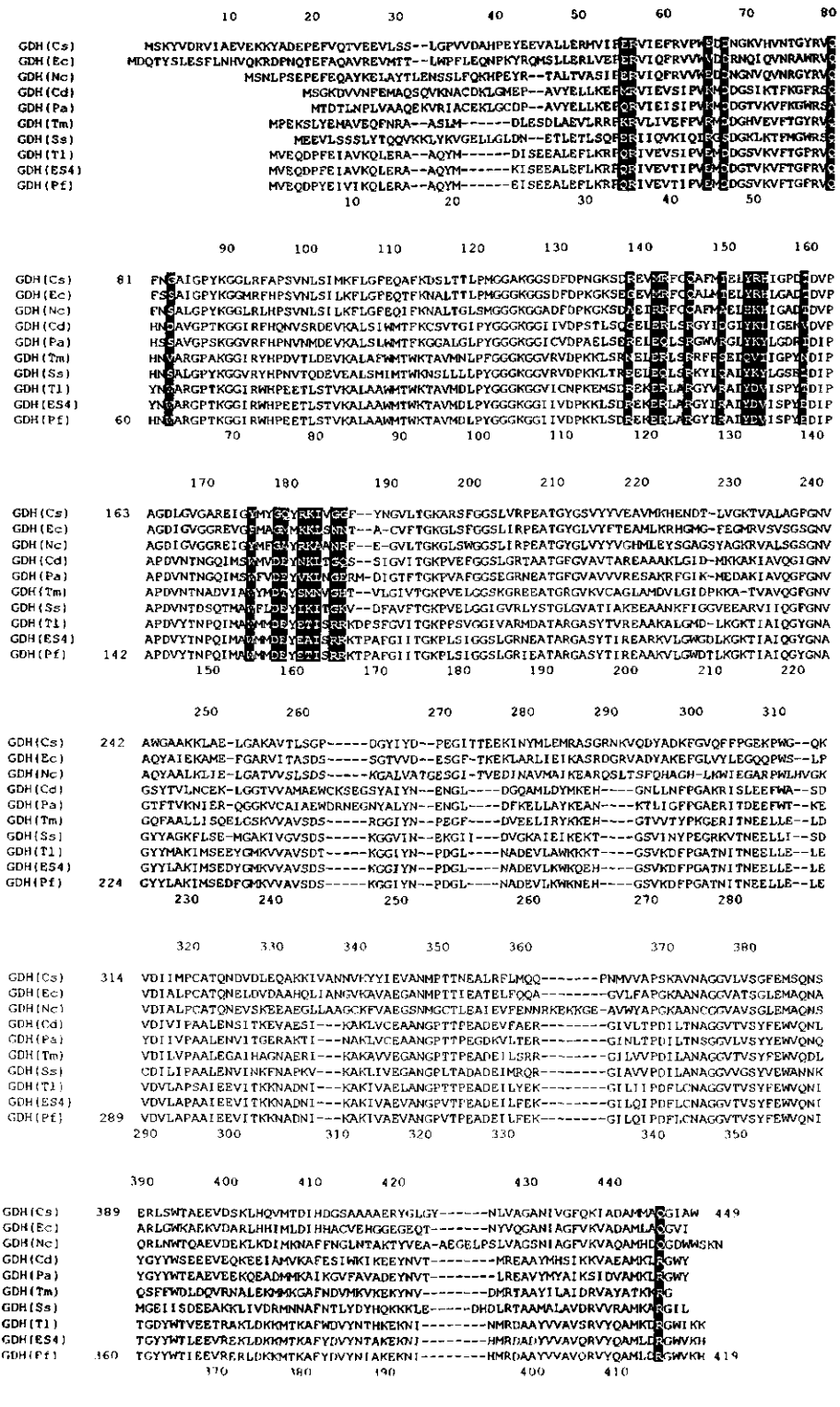

GDH(Cs)
GDH(Ec)
GDH(Nc)
GDH(Cd)
GDH(Pa)
GDH(Tm)
GDH(Ss)
GDH(Tl)
GDH(ES4)
GDH(Pf)

Table 1. The percentage sequence identity of 9 GluDH species when compared with the Pf enzyme. Calculated in a pair-wise manner between the equivalent regions of the Pf GluDH sequence and that of the other enzymes.

GluDH species	Sequence identity	Melting temperature	Half-life	
	%	°C		
P. furiosus	100	114.5	4.8 h	104 °C
P. endeavori	97	113	3.5 h	105 °C
T. litoralis	87	109	0.3 h	104 °C
S. solfataricus	44	n.d.	15 h	80 °C
T. maritima	55	93	105 min	85 °C
P. asaccharolyticus	47	n.d.	50 min	78 °C
C. difficile	52	n.d.	5.1 h	65 °C
N. crassa	33	n.d.	170 min	59 °C
E. coli	30	n.d.	n.d.	
C. symbiosum	33	55	20 min	52 °C

patches in the different subunits (Fig. 3c). The interactions between charged residues from this patch on three different subunits forms a further six-residue cluster (cluster 2) and two of these clusters lie close to cluster 1. Additionally a smaller cluster of three residues (cluster 1A) which lie close to cluster 1 is also formed by residues from this patch. Overall, the interaction of these patches contributes to a region of the structure which is dominated by the presence of interacting charged residues and which involves 52 residues of which 16 are uncharged and 36 are charged and which are responsible for the formation of 32 ion-pairs. Since this interaction occurs across a twofold axis there are necessarily 26 residues that are unique to a single subunit. In total, of the 36 charged residues in these clusters (one of cluster 1, two cluster 1A and two cluster 2), 34 arise from the patch on the subunit surface that is rich in charged residues.

Viewed down the twofold axis, the ion-pair network involving the three-residue cluster (cluster 1A), lies close to and above the 18-residue network, whilst the six-residue cluster lies to one side. Two six-residue clusters lie close to each other, across the twofold axis, between dimers, and three of these pairs almost connect the three 18-residues clusters which occur around the threefold axis (Fig. 3d). This arrangement results in the formation of an almost seamless belt of ion-pair interactions embracing all six subunits (Fig. 3b).

A comparison of the structures of Cs and Pf GluDH over the region of high charge density reveals that the two enzymes are very closely related. The only exception is the small change in the fold of the loop between α5 and βf which is subtly modified due to an insertion in the Pf enzyme at residue 167. Hence due to the effect of this minor structural change, Lys166 which forms part of cluster 2, has been omitted in the analysis. Overall therefore, 24 unique residues which form the patch that gives rise to the ion-pair networks can be unambiguously equivalenced between the Cs and Pf enzymes and the α carbon atoms of these residues can be superimposed with an rms error of 0.84 Å, highlighting their structural similarity. Examination of the sequence alignment presented in Fig. 1 shows that the sequences at these 24 positions are such that the structures of the Cs and Pf enzymes can be regarded as representative for the entire enzyme family. Thus, it is straightforward to analyse the variation in the residue distribution in this area across those enzymes with intermediate thermal stability whose structures are not yet available.

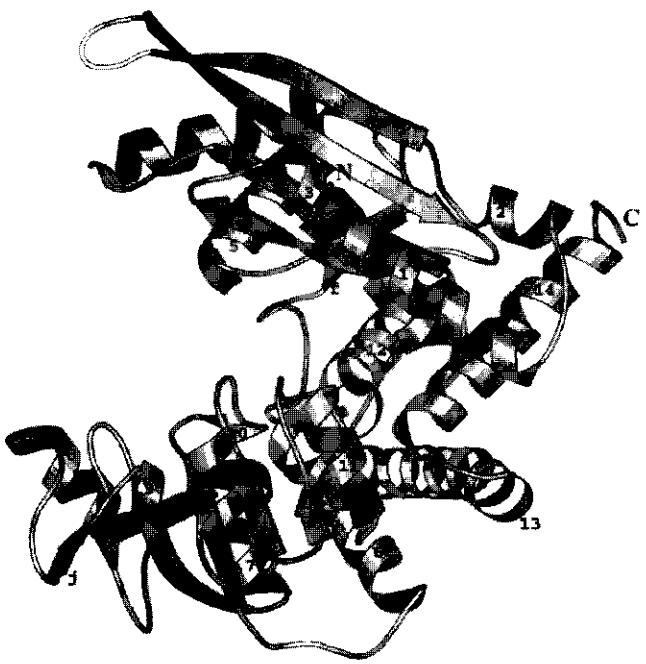


Fig. 2. Schematic representation of the secondary structure of the monomer of Pf GluDH with strands (a–m) and helices (1–14) marked. Domain I which is formed largely from residues from the N-terminal portion of the polypeptide chain and is uppermost in this figure whilst domain II is at the bottom. This figure was prepared using the MOLSCRIPT program [50].

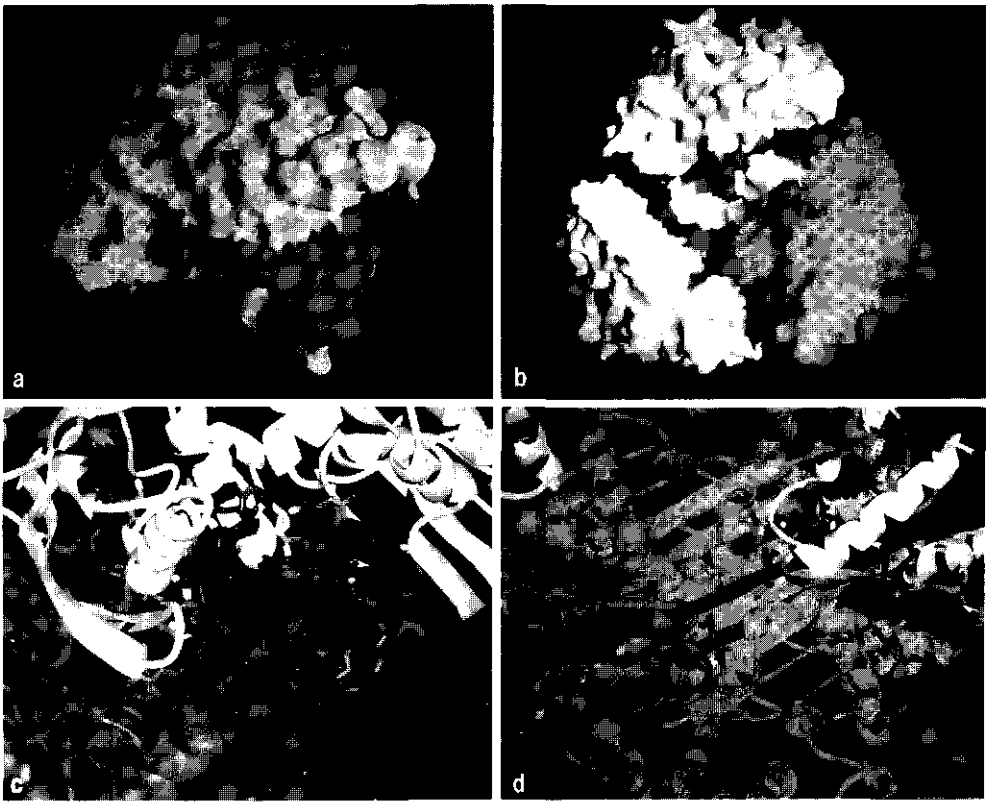


Fig. 3. Schematic views of the molecular surfaces involved in ion-pair interactions of Pf GluDH. (a) A molecular surface diagram of a Pf GluDH monomer drawn in white (GRASP [51]) which highlights the patch of high charge density (orange) which in the assembled hexamer gives rise to the extensive ion-pair networks. (b) A molecular surface diagram of a Pf GluDH trimer viewed down the threefold axis showing the interface that is involved in trimer-trimer interaction in the hexamer. The surfaces associated with the charged residues in clusters 1 and 1A are shown in red and those involved in cluster 2, in pink (drawn using the GRASP program [51]). (c) A schematic diagram showing four subunits of Pf GluDH in a ribbon representation viewed down the twofold axis between dimers. The subunits are shown in different colours and the side chains of the residues which contribute to the ion-pair networks cluster 1 and 1A are indicated in all atom representations in red. Surrounding cluster 1 are four sets of cluster 2 with their charged side chains shown in pink (MIDAS program [52]). (d) A similar representation of the Pf GluDH viewed down the twofold axis relating two monomers drawn in the style of (c). Again two cluster 2 networks are shown in pink, and at either side, two cluster 1 and cluster 1A networks are shown in red.

In Table 2 we summarise the characteristics of the residues which are found in the patch of high charge density across the various GluDH species which span a wide spectrum of growth temperatures. Specifically, since ion-pairs necessarily involve the interaction of residues of opposite charge, we have analysed the number of balanced charges, rather than the total number of charged residues across these GluDHs that comprise these regions against the difference in enzyme stability. Whilst in the absence of structures for each of these enzymes, we cannot be certain that the change in the number of balanced charges necessarily equates to the change in the number of ion pairs, we assume that the analysis provides at least a guideline as to how many ion-pairs might be formed. The number of balanced charges in this area is plotted against the rank-ordered temperature stability in Fig. 4. Examination of this figure clearly shows that the hyperthermostable enzymes contain significantly more balanced charges than their mesophilic equivalents with the enzymes which possess an intermediate stability occupying positions between the two extremes. If the assumption that the number of balanced charges is in some way related to the formation of ion-pairs is realistic, then this analysis reveals a possible correlation between ion-pairs and enzyme stability. The decrease in the number of balanced charges observed in passing from the hyperthermostable enzymes to the GluDHs from the mesophilic organisms is accompanied by a corresponding increase in both hydrophobic and polar residues. Furthermore, we also note that the number of glycines in the patch increases in the mesophilic enzymes. This is consistent with earlier analyses of preferential substitutions in thermophilic proteins involving the replacement of glycine [7, 44].

Detailed comparison of the major ion-pair clusters across the enzyme family. *Analysis of clusters 1 and 1A.* A schematic diagram which illustrates the spatial arrangement of the complex ion-pair interactions in cluster 1 [and its associated smaller cluster of three residues in cluster 1A (Asp46, Arg121 (subunit A#), Glu44 (subunit A)] is presented in Fig. 5a [Note, subunits identified with the same suffix are related by threefold symmetry

Table 2. Variation of residue composition with species for the patch of high charge density in Pf GluDH. The distribution of residue types in the GluDH from each of the ten different species that comprise the 24 unique residues which are responsible for the formation of the patch of high charge density in the Pf enzyme and have counterparts across all the GluDHs. The actual numbers of ion-pairs formed between these residues in the three-dimensional structures of Cs, Ec, Nc and Pf GluDHs have been determined following X-ray analyses and are included in the penultimate column of the table. The analysis of the patch across the members of the GluDH family excludes residue K166 due to small structural differences between the enzymes. Data are arranged in decreasing order of temperature stability with the most stable enzymes at the top.

GluDH species	No. of (+) residues	No. of (−) residues	No. of polar residues	No. of hydrophobic residues	No. of glycines	No. of balanced changes	No. of unique ion-pairs within defined region	Total no. of ion-pairs/ subunit
Pf	8	8	5	3	–	8	14	45
Pe	8	8	6	2	–	8	–	–
Tl	8	7	7	2	–	7	–	–
Ss	8	5	8	3	–	5	–	–
Tm	6	3	11	4	–	3	–	–
Pa	8	6	6	4	–	6	–	–
Cd	7	6	6	4	1	6	–	–
Nc	8	3	8	4	1	3	3	–
Ec	5	2	10	5	2	2	1	28
Cs	6	3	7	4	4	3	4	26

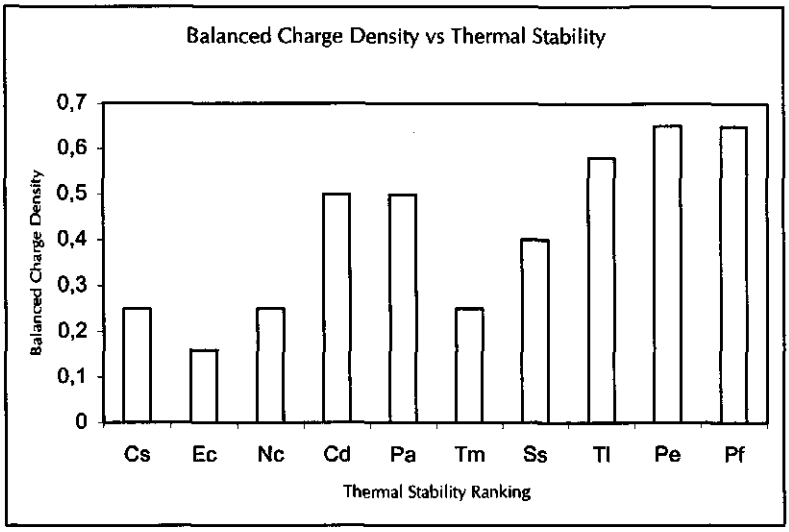
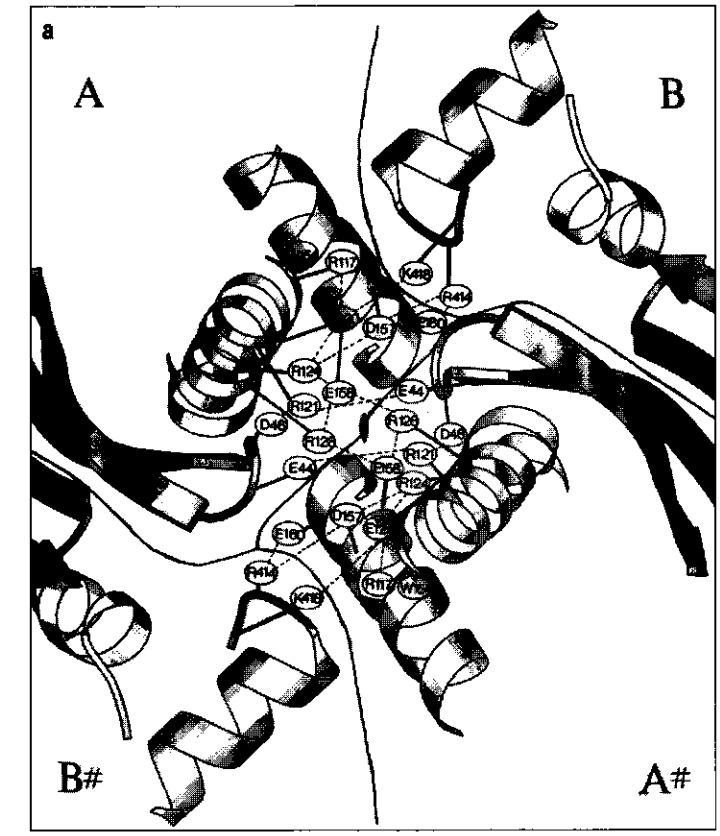


Fig. 4. A box chart to represent the correlation between the balanced charge density against increasing thermal stability ranking of the GluDHs from different species. Balanced Charge Density = $(n^+ + n^-) - |n^+ - n^-|/N_{total}$ where n^- and n^+ are the numbers of positively and negatively charged residues respectively in the two patches of high charge density and N_{tot} is the overall number of unique residues in the patches.

(e.g. A, B and C or A#, B# and C#) whereas those bearing a different suffix (e.g. A and A#) are related by twofold symmetry]. Cluster 1 involves the formation of ion pairs between the side chains of Glu160, Asp157, Arg124, Arg117, Glu120, Glu158, Arg128 (subunit A), Lys418 and Arg414 (subunit B) and their twofold related partners. In the global analysis conducted above we have considered all three positions of cluster 1A and sequence changes at 17 of the 18 positions within cluster 1 (Lys418 was omitted since it lies beyond the boundary of the patch). Amongst the most thermostable GluDHs from *Thermococcus litoralis* (Tl) [37] and *Pyrococcus endeavori* (Pe) (Baross, J. A., personal communication), all the residues which



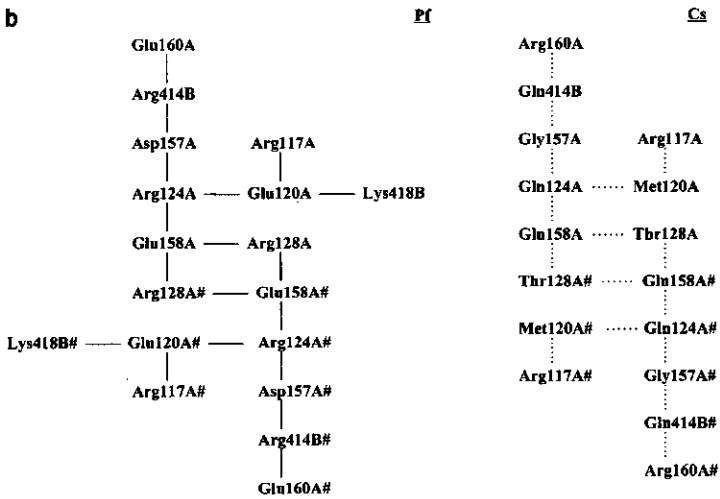


Fig. 5. Residues involved in the 18 residue ion-pair network in Pf GluDH and equivalents in the Cs enzyme. (a) A schematic representation showing the ion-pair interactions which contribute to the 18-residue ion-pair cluster in the Pf GluDH (cluster 1 and 1A). The twofold axis relating the dimer AB# to the dimer BA# is indicated by a diad (◆). Ion-pair partners are shown by dashed lines. The backbone of the polypeptide chain is shown as a grey ribbon and the approximate side-chain positions are indicated in black. The uncharged residues in this region are also marked. The diagram was prepared using MOLSCRIPT [50]. (b) An illustration of the ion-pair interactions in cluster 1 in Pf GluDH (left) and of the equivalent residues in Cs GluDH (right). The sequence changes between the GluDH from these two species result in the total breakdown of this network in the mesophilic enzyme. All ionic interactions are indicated as solid lines and those that are lost are depicted as dashed lines. In the Pf enzyme cluster 1 also includes K418 which has been omitted from this analysis as it lies beyond the edge of the patch.

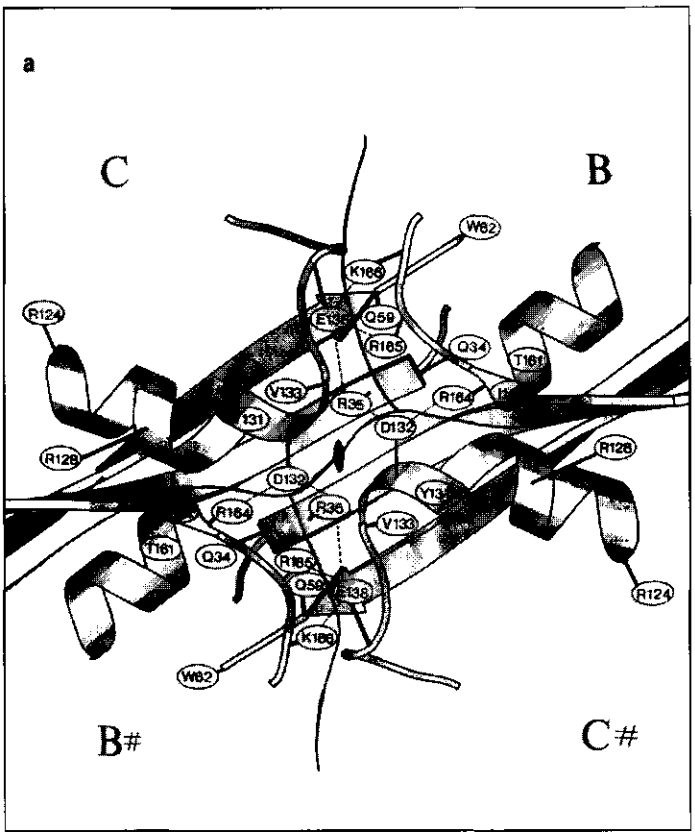


Fig. 6. The six residue ion-pair network in Pf GluDH and equivalents in the less thermostable counterparts. (a) A schematic representation drawn in the same style as Fig. 5a to show the ion-pair network cluster 2 which comprises six residues. The twofold axis relating a CC# dimer is indicated by a diad. Potential ion-pair interactions are indicated by dashed lines. Again uncharged residues in this region are displayed, together with R124, and R128 from cluster 1 to illustrate the relative position of cluster 2 to cluster 1 at the periphery. (b) An illustration of the ion-pair interactions in cluster 2 in Pf GluDH and the equivalent residues in Tl, Ss, Nc, and Cs GluDHs. The partial and complete fragmentation of this cluster as the temperature stability of the enzyme falls can be seen. K166 which forms part of cluster 2 in the Pf enzyme has been omitted from the analysis due to minor structural differences in some of the GluDHs.

participate in both cluster 1 and cluster 1A in the Pf enzyme are conserved and the formation of an identical set of ion-pair interactions can confidently be predicted.

However, moving down the temperature spectrum towards the less thermostable enzymes, it is clear that this network collapses. For example, in *Sulfolobus solfataricus* (Ss) GluDH, sequence changes include the replacement of Arg128 by glutamine. This residue is situated at the centre of cluster 1 forming an ion-pair handshake with both symmetry-related carboxyl side chains of Glu158 across the twofold axis and therefore the sub-

Table 3. Closest approach distance between atoms bearing formal charge and involved in ion-pairs in the structures of Pf and Cs GluDH.

Closest approach distance	Number of such ion-pairs in GluDH from	
	Cs	Pf
Å		
≤2.5	2	4
>2.5 ≤3.0	21	25
>3.0 ≤3.5	3	9
>3.5 ≤4.0	0	7

stitution leads to the loss of multiple ion-pair interactions. Further sequence changes are also found at the periphery of cluster 1 with the replacement of Arg117 and Glu160 and the deletion of Lys418. As a result of these sequence changes, of the six triple ion pairs formed by residues in cluster 1 in the Pf enzyme, only two (involving the symmetry-related interactions of Arg124) appear to be conserved and thus, of this once extensive 18-residue network, only two, five-residue networks are likely to be found in Ss GluDH. Furthermore, if we move lower down the temperature spectrum to the less thermostable GluDHs then the complete collapse of the network of ion-pairs in cluster 1 is apparent. For example structural data on both Cs (Fig. 5b) [40], Ec GluDHs (Stillman, T. J., unpublished results), and Nc GluDH (Stillman et al., unpublished results) confirm the absence of ion-pair networks in this region (Table 3). The smaller ion-pair cluster (1A) involving Asp46 and Arg121 from subunit A and Glu44 from a twofold related subunit appears to be more strongly (although not completely) retained across the enzyme family. The possible functional significance of this is not currently understood.

Analysis of cluster 2. This is the second largest ion-pair network in the Pf enzyme and involves six residues with the formation of ion pairs between the side chains of Arg164, Arg165 and Lys166 from subunit B, Arg35 and Glu138 from subunit C and Asp132 from subunit C# (Fig. 6a).

In the modelling study we were only able to consider sequence changes at five of the six positions within cluster 2 (Lys166 was omitted on account of the small structural differences close to this position found in some members of the enzyme family). However, amongst the most stable GluDHs (Pf, Pe, Tl), the sequence alignment is such that changes at all six positions can be analysed. This analysis showed that all six positions are conserved in the two more stable Pf and Pe GluDHs. However, in Tl GluDH, the least stable enzyme of the three (though still an enzyme of considerable thermal stability with a half-life of 2 h at 98°C [37]), Glu138, which lies at the centre of this cluster forming a triple ion-pair to Arg35, Arg165 and Lys166, is replaced by threonine. This substitution would remove this potentially key glutamate and reduce the network to one of only three residues. Further analysis of the remaining five positions of cluster 2 for the enzymes of both intermediate and low thermal stability would appear to indicate that the entire network is lost. This is illustrated in Fig. 6b where the complete absence of complementary charged side chains at equivalent positions can be seen in the structure of the mesophilic Cs and Nc enzymes and in the moderately thermostable GluDH from Ss.

Ion-pairing in hyperthermophiles. The dramatic nature of the extensive ion-pair networks in Pf GluDH was an unexpected feature of the structure and the absence of such networks in the comparable mesophilic enzymes led to the suggestion that they play a key role in the superior thermal properties of the hyperthermostable enzyme [29]. At present, three-dimensional structures are only available for enzymes which represent the extremes of stability and thus it is not possible to compare directly the degree of ion-pair formation across the family and correlate it with thermal stability. However, the similarity in the structures of these enzymes that can confidently be predicted from the sequence similarity has allowed us to carry out a modelling exercise to try to predict how these networks vary across the enzyme family. One problem associated with such an approach is that the side chains of the charged residues, particularly arginine and lysine, are inherently very flexible. Hence, it is difficult to model with confidence the exact position of the charged portion of the side chain and therefore the determination of the exact number of ion-pairs is problematic. Nevertheless, in the case of the Pf enzyme, the formation of extensive ion-pair networks is a natural consequence of the high concentration of basic and acidic residues on adjacent regions of the enzyme to such an extent that the formation of networks is almost unavoidable. Therefore, in the first instance, we have tried to form an assessment of the number of ion pairs formed across the family by examining the variation in the fraction of charged residues which comprise the two patches which give rise to the ion-pair clusters.

The study appears to provide supporting evidence of possible correlations between the extent and character of ion-pair clusters in the members of the GluDH family with their relative thermal stabilities. Overall, not only does the net number of ion-pairs appear to decrease between the hyperthermostable enzymes and the enzymes from mesophilic organisms, but the larger clusters seem to disappear rapidly as enzyme thermal stability falls below 100°C with the networks becoming less extensive and more fragmented. The apparent absence of the larger, more intricate networks of ion pairs in enzymes operating at temperatures below 100°C might well indicate that the multiple ion-pair interactions are critical only in the proteins of those organisms that function at the highest temperatures. However, if the formation of such networks is responsible for high temperature stability of proteins, it still remains to be established whether this is the only mechanism by which such stability may be generated or only one of a number of distinct molecular solutions that Nature has utilised to solve the problem of surviving such harsh conditions.

Clearly, further three-dimensional structural information on the family of GluDHs will permit a more detailed analysis of the changes to the ion-pair clusters to be made and enhance our understanding of the molecular basis of thermal stability. However, in the meantime, the analysis already highlights a key area of the structure of this enzyme that may be important in stability and pinpoints a number of sequence changes whose roles in generating stability at high temperatures can be tested by site-directed mutagenesis. In the future such work will allow this proposal to be tested in detail.

This work was supported by grants from the UK Biotechnology and Biological Sciences Research Council, the EU Biotechnology programs: BIO2-CT930274 and BIO4-CT960488; the New Energy and Industrial Technology Development Organisation and the National Science Foundation grants BCS-9320069 and BES-9410687. The Krebs Institute is a designated Biotechnology and Biological Sciences Research Council Biomolecular Sciences Centre.

REFERENCES

1. Flam, F. (1994) The chemistry of life at the margins, *Science 265*, 471–472.
2. Adams, M. W. W. & Kelly, R. M. (1995) Enzymes from microorganisms in extreme environments, *Chem. Eng. News 73*, December 18, 32–42.

3. Woese, C. R., Kandler, O. & Wheelis, M. L. (1990) Towards a natural system of organisms: Proposal for the domains; Archaea, Bacteria and Eucarya, *Proc. Natl Acad. Sci. USA 87*, 4576–4579.
4. Vihinen, M. (1987) Relationship of protein flexibility to thermostability, *Protein Eng. 1*, 481–485.
5. Mathews, B. W., Nicholson, H. & Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, *Proc. Natl Acad. Sci. USA 84*, 6663–6667.
6. Day, M. W., Hsu, B. T., Joshua-tor, L., Park, J.-B., Zhou, Z. H., Adams, M. W. W. & Rees, D. (1992) X-Ray crystal structures of the oxidised and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*, *Protein Sci. 1*, 1494–1507.
7. Kelly, C. A., Nishiyama, M., Ohnishi, Y., Beppu, T. & Birktoft, J. J. (1993) Determinants of protein thermostability observed in the 1.9 Å crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*, *Biochemistry 32*, 3913–3922.
8. Perutz, M. F. & Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2, *Nature 255*, 256–259.
9. Perutz, M. F. (1978) Electrostatic effects in proteins, *Science 201*, 1187–1191.
10. Britton, K. L., Baker, P. J., Borges, K. M. M., Engel, P. C., Pasquo, A., Rice, D. W., Robb, F., Scandurra, R., Stillman, T. J. & Yip, K. S. P. (1995) Insights into thermal stability from a comparison of the glutamate dehydrogenase from *Pyrococcus furiosus and Thermococcus litoralis*, *Eur. J. Biochem. 229*, 688–695.
11. Russell, R. J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. (1994) The crystal structure of citrate synthase from the thermophilic archaeon *Thermoplasma acidophilum*, *Structure 2*, 1157–1167.
12. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect, *Science 255*, 178–183.
13. Völkl, P., Markiewicz, P., Stetter, K. O. & Miller, J. (1994) The sequence of a subtilisin-type protease (aerolysin) from the hyperthermophilic archaeum *Pyrobaculum aerophilum* reveals sites important to thermostability, *Protein Sci. 3*, 1329–1340.
14. Chan, M. K., Mukland, S., Kletzin, A., Adams, M. W. W. & Rees, D. C. (1995) Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase, *Science 267*, 1463–1469.
15. Tesfay, H. S., Amelunxen, R. E. & Goldberg, I. D. (1989) Nucleotide-sequences of genes encoding heat-stable and heat-labile glyceraldehyde-3-phosphate dehydrogenases – Amino acid sequence and protein stability, *Gene (Amst.) 82*, 237–248.
16. Suzuki, Y., Hatagaki, K. & Oda, H. (1991) A hyperthermostable pullulanase produced by an extreme thermophile, *Bacillus flavocaldarius* KP 1228, and evidence for the proline theory of increasing protein thermostability, *Appl. Microbiol. Biotechnol. 34*, 707–714.
17. Matsumura, M., Signor, G. & Matthews, B. W. (1989) Substantial increase of protein stability by multiple disulphide bonds, *Nature 342*, 291–293.
18. Imada, K., Sato, M., Tanaka, N., Katsube, Y., Matsura, Y. & Oshima, T. (1991) Three dimensional structure of a highly thermostable enzyme, 3-Isopropylmalate dehydrogenase of *Thermus thermophilus* at 2.2 Å resolution, *J. Mol. Biol. 222*, 725–738.
19. Hurley, J. H., Baase, W. A. & Matthews, B. W. (1992) Design and structural analysis of alternative core packing arrangements in Bacteriophage T4 Lysozyme, *J. Mol. Biol. 224*, 1144–1159.
20. Anderson, D. E., Hurley, J. H., Nicholson, H., Baase, W. A. & Matthews, B. W. (1993) Hydrophobic core repacking and aromatic-aromatic interaction in the thermostable mutant of T4 lysozyme Ser117→Phe, *Protein Sci. 2*, 1285–1290.
21. Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995) Are buried salt bridges important for protein stability and conformational specificity? *Struct. Biol. 2*, 122–128.
22. Pauptit, R. A., Karlsson, R., Picot, D., Jenkins, J. A., Niklaus-Reimer, A.-S. & Jansonius, J. N. (1988) Crystal structure of neutral protease from *Bacillus cereus* refined at 3.0 resolution and comparison with the homologous but more thermostable enzyme thermolysin, *J. Mol. Biol. 199*, 525–537.
23. Creighton, T. E. (1985) The problem of how and why proteins adopt folded conformations, *J. Phys. Chem. 89*, 2452–2459.
24. Boyer, P. D. (1975) *The enzymes II*, 3rd edn, Academic Press, New York.
25. Hudson, R. C. & Daniel, R. M. (1993) L-Glutamate dehydrogenases: distribution, properties and mechanism, *Comp. Biochem. Physiol. B 106*, 767–792.
26. Fiala, G. & Stetter, K. O. (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C, *Arch. Microbiol. 145*, 56–61.
27. Consalvi, V., Chiaraluce, R., Politi, L., Vaccaro, R., De Rosa, M. & Scandurra, R. (1991) Extremely thermostable glutamate dehydrogenase from the hyperthermophilic archaebacterium *Pyrococcus furiosus*, *Eur. J. Biochem. 202*, 1189–1196.
28. Syed, S. E. H. (1987) *Glutamate dehydrogenase from* Clostridium symbiosum, PhD Thesis, University of Sheffield, UK.
29. Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R. & Rice, D. W. (1995) The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion pairs networks in maintaining enzyme stability at extreme temperatures, *Structure 3*, 1147–1158.
30. Rice, D. W., Yip, K. S. P., Stillman, T. J., Britton, K. L., Fuentes, A., Connerton, I., Pasquo, A., Scandurra, R. & Engel, P. C. (1996) Insights into the molecular basis of thermal stability from the structure determination of *Pyrococcus furiosus* glutamate dehydrogenase, *FEMS Microbiol. Rev. 18*, 105–117.
31. Abeysinghe, I. S. B. (1993) Structural studies of the $NADP^+$-linked glutamate dehydrogenase from *E. coli*, PhD Thesis, University of Sheffield, UK.
32. Fincham, J. R. S. & Garner, H. R. (1967) Effects of pH and temperature on the wild type and a mutant form of *Neurospora* glutamate dehydrogenase, *J. Biochem. (Tokyo) 103*, 705–708.
33. Anderson, B. M., Anderson, C. D., Van Tassell, R. L., Lyerly, D. M. & Wilkins, T. D. (1993) Purification and characterisation of *Clostridium difficile*, *Arch. Biochem. Biophys. 300*, 483–488.
34. Lebbink, J. H. G., Eggen, R. I. L., Geering, A. C. M., Consalvi, V., Chiaraluce, R., Scandurra, R. & de Vos, W. M. (1995) Exchange of domains of GluDH from the hyperthermophilic archaeon *Pyrococcus furiosus* and the mesophilic bacterium *Clostridium difficile* effects on catalysis, thermoactivity and stability, *Protein Eng. 8*, 1287–1294.
35. Kort, R., Liebl, W., Labedan, B., Forterre, P., Eggen, R. I. L. & de Vos, W. M. (1997) *Extremophiles*, in the press.
36. Maras, B., Consalvi, V., Chiaraluce, R., Politi, L., De Rosa, M., Bossa, F., Scandurra, R. & Barra, D. (1992) The protein sequence of glutamate dehydrogenase from *Sulfolobus solfataricus*, a thermoacidophilic archaebacterium. Is the presence of *N*-ε-methyllysine related to thermostability? *Eur. J. Biochem. 203*, 81–87.
37. Ma, K., Robb, F. T. & Adams, M. W. W. (1994) Purification and characterisation of NADP-specific alcohol dehydrogenase and glutamate dehydrogenase from the hyperthermophilic archaeon *Thermococcus litoralis*, *Appl. Environ. Microbiol.* February, 562–568.
38. DiRuggiero, J., Robb, F. T., Jagus, R., Klump, H. H., Borges, K. M., Kessel, M., Mai, X. & Adams, M. W. W. (1993) Characterisation, cloning and *in vitro* expression of the extremely thermostable glutamate dehydrogenase from the hyperthermophilic archaeon, *ES4*, *J. Biol. Chem. 268*, 17767–17774.
39. Klump, H., DiRuggiero, J., Kessel, M., Park, J.-B., Adams, M. W. W. & Robb, F. T. (1992) Glutamate dehydrogenase from the hyperthermophilic *Pyrococcus furiosus*: Thermal denaturation and activation, *J. Biol. Chem. 267*, 22681–22685.
40. Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W. & Stillman, T. J. (1992) Subunit assembly and active site location in the structure of glutamate dehydrogenase, *Proteins 12*, 75–86.
41. Rice, D. W., Hornby, D. P. & Engel, P. C. (1985) Crystallisation of an NAD^+ dependent glutamate dehydrogenase from *Clostridium symbiosum*, *J. Mol. Biol. 181*, 147–149.

42. McPherson, M. J. & Wootton, J. C. (1983) Complete nucleotide sequence of the *Escherichia coli* gdhA gene, *Nucleic Acids Res. 11*, 5257–5266.
43. Koberstein, R. & Sund, H. (1973) Studies of glutamate dehydrogenase. The influence of ADP, GTP, L-glutamate on the binding of the reduced coenzyme to beef liver GluDH, *Eur. J. Biochem. 36*, 545–552.
44. Menéndez-Arias, L. & Argos, P. (1989) Engineering protein thermal stability, *J. Mol. Biol. 206*, 397–406.
45. Teller, J. K., Smith, R. M., McPherson, M. J., Engel, P. C. & Guest, J. R. (1992) The glutamate dehydrogenase gene of *Clostridium symbiosum*. Cloning by polymerase chain reaction sequence analysis and over expression in *Escherichia coli*, *Eur. J. Biochem. 206*, 151–159.
46. Kinnaird, J. H. & Fincham, J. R. S. (1983) The complete nucleotide sequence of the *Neurospora crassa* am (NADP-specific glutamate dehydrogenase) gene, *Gene (Amst.) 26*, 253–260.
47. Lyerly, D. M., Barroso, L. A. & Wilkins, T. D. (1991) Identification of the latex test-reactive protein of *Clostridium difficile* as glutamate dehydrogenase, *J. Chem. Microbiol. 29*, 2639–2642.
48. Snedecor, B., Chu, H. & Chen, E. (1991) Selection, expression, and nucleotide sequencing of the glutamate dehydrogenase gene of *Peptostreptococcus asaccharolyticus*, *J. Bacteriol. 173*, 6162–6167.
49. Eggen, R. I. L., Geerling, A. C. M., Waldkötter, K., Antranikan, G. & de Vos, W. M. (1993) The glutamate dehydrogenase encoding gene of the hyperthermophilic archaeon *Pyrococcus furiosus*: sequence, transcription and analysis of the deduced amino acid sequence, *Gene (Amst.) 132*, 143–148.
50. Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. 24*, 946–950.
51. Nicholls, A., Sharp, K. A. & Honig, B. (1991) Protein folding and association – Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins Struct. Funct. Genet. 11*, 281–296.
52. Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988) The MIDAS display system, *J. Mol. Graphics 6*, 13–27.

Chapter 5

Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. I. Introduction of a six-residue ion-pair network in the hinge region.

Joyce H.G. Lebbink, Stefan Knapp, John van der Oost, David W. Rice, Rudolf Ladenstein, and Willem M. de Vos.

Reprinted with permission from
Journal of Molecular Biology **280**, 287-296, 1998.

JMB

Engineering activity and stability of *Thermotoga maritima* Glutamate Dehydrogenase. I. Introduction of a Six-residue Ion-pair Network in the Hinge Region

Joyce H. G. Lebbink[1], Stefan Knapp[2], John van der Oost[1], David Rice[3] Rudolf Ladenstein[2] and Willem M. de Vos[1]*

[1]*Laboratory for Microbiology Wageningen Agricultural University, Hesselink van Suchtelenweg 4, NL-6703 CT Wageningen, The Netherlands*

[2]*Center for Structural Biochemistry, NOVUM Karolinska Institute S-14157 Huddinge, Sweden*

[3]*The Krebs Institute for Biomolecular Research Department of Molecular Biology and Biotechnology University of Sheffield PO box 594, Sheffield S10 2UH UK*

**Corresponding author*

Comparison of the recently determined three-dimensional structures of several glutamate dehydrogenases allowed for the identification of a five-residue ion-pair network in the hinge region of *Pyrococcus furiosus* glutamate dehydrogenase (melting temperature 113°C), that is not present in the homologous glutamate dehydrogenase from *Thermotoga maritima* (melting temperature 93°C). In order to study the role of this ion-pair network, we introduced it into the *T. maritima* enzyme using a site-directed mutagenesis approach. The resulting *T. maritima* glutamate dehydrogenases N97D, G376 K and N97D/G376 K as well as the wild-type enzyme were overproduced in *Escherichia coli* and subsequently purified. Elucidation of the three-dimensional structure of the double mutant N97D/G376 K at 3.0 Å, showed that the designed ion-pair interactions were indeed formed. Moreover, because of interactions with an additional charged residue, a six-residue network is present in this double mutant. Melting temperatures of the mutant enzymes N97D, G376 K and N97D/G376 K, as determined by differential scanning calorimetry, did not differ significantly from that of the wild-type enzyme. Identical transition midpoints in guanidinium chloride-induced denaturation experiments were found for the wild-type and all mutant enzymes. Thermal inactivation at 85 C occured more than twofold faster for all mutant enzymes than for the wild-type glutamate dehydrogenase. At temperatures of 65°C and higher, the wild-type and the three mutant enzymes showed identical specific activities. However, at 58°C the specific activity of N97D/G376 K and G376 K was found to be significantly higher than that of the wild-type and N97D enzymes. These results suggest that the engineered ion-pair interactions in the hinge region do not affect the stability towards temperature or guanidinium chloride-induced denaturation but rather affect the specific activity of the enzyme and the temperature at which it functions optimally.

Keywords: glutamate dehydrogenase; thermostability; hinge ion-pair networks; *Pyrococcus furiosus*; *Thermotoga maritima*

Introduction

During recent years an increasing amount of genetic, biochemical and structural data have become available of enzymes isolated from hyperthermophilic Bacteria and Archaea. This has allowed for an extensive comparison with data obtained from moderate thermophiles, mesophiles and psychrophiles, in order to determine the molecular basis for adaptations to high temperatures. In many cases the involvement of charged residues has been proposed, including optimized electrostatic interactions (Spassov & Atanasov, 1994; Spassov *et al.*, 1995), increased numbers of ion-pairs (Korndörfer *et al.*, 1995; Hennig *et al.*, 1997), and extensive ion-pair networks (Yip *et al.*, 1995; Hennig *et al.*, 1995; Knapp *et al.*, 1997; Russell *et al.*, 1997). Several studies involving single, surface-located salt bridges in T4 lysozyme have shown

Abbreviations used: GDH, glutamate dehydrogenase.

that their contribution to the stability of the protein is only marginal (0.1 to 0.25 kcal/mol) because the gain in free energy of folding is about equal to the entropic cost of dehydration and the reduction of the conformational freedom (Dao-pin *et al.*, 1991; Horovitz *et al.*, 1990). However, theoretical considerations indicate that extensive networks may play an important role in maintaining enzyme stability or function at extreme temperatures. First, each extra ion-pair added to the network requires the desolvation and localization of only a single residue. In addition, networks are often located in cavities and at interfaces where their conformational freedom is already restricted. As a consequence, part of the entropic cost has already been provided during the folding of the protein. Finally, hydration effects play a minor role at high temperatures and the dielectric constant decreases with temperature, resulting in an increased electrostatic energy upon formation of a salt bridge.

Studies on the functionality of electrostatic interactions in hyperthermostable enzymes are limited to the deletion of either isolated ion-pairs or the disruption of small ion-pair networks in *Thermotoga maritima* glyceraldehyde-3-phosphate dehydrogenase (Tomschy *et al.*, 1994; Pappenberger *et al.*, 1997). Removal of the side-chain from the central arginine in a four-residue network resulted in faster thermal denaturation of the mutant, suggesting that ion-pair interactions do indeed contribute to the stability of the enzyme (Pappenberger *et al.*, 1997). The question remains, however, if stability in systems like this would be affected if only the ionic interaction was removed, without introducing cavities. Here we have taken the complementary approach by introducing an ion-pair network that is present in a thermostable model enzyme, into a less thermostable counterpart. Apart from studying the effects on stability, the effects on catalytic efficiency and temperature optimum were investigated as well. These characteristics all have to be adapted to high temperatures, but especially the latter two have not yet received much attention. The well-studied enzyme glutamate dehydrogenase (GDH) is, for several reasons, an excellent model. Three-dimensional structures have been solved for GDHs from the mesophilic bacterium *Clostridium symbiosum* (Baker *et al.*, 1992), the hyperthermophilic bacterium *T. maritima* (Knapp *et al.*, 1997) and the hyperthermophilic archaeon *Pyrococcus furiosus* (Yip *et al.*, 1995). All these GDHs consist of six identical subunits, each of which can be divided into a substrate- and a cofactor-binding domain (Baker *et al.*, 1992). During catalysis, the cofactor binding domain rotates with respect to the substrate binding domain, thereby opening and closing the cleft in between the two domains where the active site is located (Stillman *et al.*, 1993). Residues and interactions in the so-called hinge region connecting the two domains, are thought to mediate this movement and therefore are assumed to play an important role in determining activity and stability of the enzyme (Yip *et al.*, 1995; Knapp *et al.*, 1997). Comparison of those structures and a homology modelling study involving GDHs from ten different sources (Yip *et al.*, 1998), showed a strong correlation of the number of ion-pairs and the extent of ion-pair networks with the thermostability of the enzyme. We previously cloned, sequenced and overexpressed in *Escherichia coli* the *gdh* genes from the hyperthermophilic archaeal and bacterial representatives *P. furiosus* and *T. maritima*, respectively (Eggen *et al.*, 1993; Lebbink *et al.*, 1995; Kort *et al.*, 1997). *P. furiosus* GDH overproduced by *E. coli* has successfully been used in protein engineering experiments to study the contribution of the different domains to activity and stability of GDH (Lebbink *et al.*, 1995).

The glutamate dehydrogenases from *P. furiosus* and *T. maritima* share 55% amino acid identity and are very homologous in secondary and tertiary structure as well (root-mean-square deviation of 3.5 Å for the C^{α} positions of the cofactor binding domain and 2.2 Å for the C^{α} positions of the substrate binding domain). However, the *P. furiosus* GDH has a higher melting point ($t_m = 113°C$; Klump *et al.*, 1992) and contains significantly larger ion-pair networks than the *T. maritima* GDH ($t_m = 93°C$; Knapp *et al.*, 1997). We studied the role of a five-residue ion-pair network that is located in the hinge region of the GDH from *P. furiosus* by introducing the involved charged amino acids into the GDH from *T. maritima*. Subsequently, the biochemical and structural properties of the mutant GDHs were compared to those of the wild-type *T. maritima* GDH.

Results

Design, production and structure of mutant and wild-type *T.maritima* GDHs

The major difference between the hinge regions of the GDHs from *P. furiosus* and *T. maritima*, is the presence of a five-residue ion-pair network in the *P. furiosus* enzyme that is absent in the less stable GDH. This five-residue ion-pair network links secondary structure elements from both domains, is surface exposed, and located oppositely to the substrate binding site at a distance of more than 12 Å (Figure 1). The five participating residues are positioned in such a way that four salt bridges are formed, i.e. between R64-D97, D97-K379, K379-D383 and D383-R406. Two of those salt bridges cross the domain interface: D97-K379 linking the loop between α-helix 3 and β-sheet d (substrate binding domain) with α-helix 13 (cofactor binding domain), and D383-R406 linking α-helix 13 (cofactor binding domain) with α-helix 14 (substrate binding domain; Figure 2(a)). In the less stable *T. maritima* GDH the corresponding residues in a structure-based alignment are R64, N97, G376, D380 and R403; consequently only the salt bridge between D380 and R403 is present (Figure 2(b)).

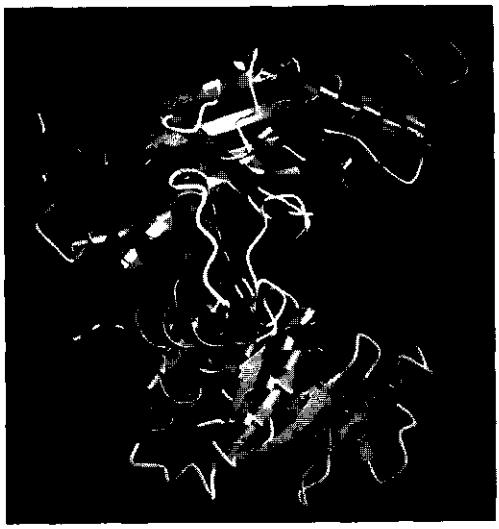


Figure 1. Schematic drawing of a wild-type *T. maritima* GDH subunit. The upper (N-terminal) domain binds the substrate in the cleft in between the two domains, the lower domain (C-terminal) binds the cofactor along its surface that is facing the upper domain (Baker *et al.*, 1992; Stillman *et al.*, 1993). The secondary structure elements linked by the ion-pair network in the double mutant and *P. furiosus* GDH are highlighted in blue. The Figure has been generated with the programme MOLSCRIPT (Kraulis, 1991) together with the rendering programme Raster3D (Bacon & Anderson, 1988; Merrit & Murphy, 1994).

In order to introduce the five-residue network into the hinge of *T. maritima* GDH, residues N97 and G376 of *T. maritima* GDH were mutated to the *P. furiosus* residues D97 and K379 using a PCR-based method. Production of the wild-type *T. maritima* GDH and the N97D, G376 K and N97D/G376 K mutant enzymes in *E. coli* in all instances exceeded 5% of total cell protein. The wild-type and mutant *T. maritima* GDHs were purified to homogeneity as judged by SDS-PAGE, and used for further structural and biochemical analysis.

Double mutant N97D/G376 K has been crystallized under conditions similar to those described for the wild-type *T. maritima* GDH (Knapp *et al.*, 1997) and in the same spacegroup ($P3_1$ 2 1) with identical cell dimensions (Table 1). The side-chain of the lysine residue introduced at position 376 could easily be built into a 6-fold averaged $(2F_O^{Mu} - F_C^{WT})\exp(i\alpha_C^{WT})$ map (not shown). In addition, the density of the lysine side-chain was clearly visible in a $(F_O^{Mu} - F_C^{WT})\exp(i\alpha_C^{WT})$ difference map (Figure 3(a)). A representative density of the averaged electron density map $(2F_O^{MU} - F_C^{MU})\exp(i\alpha_C^{MU})$ is shown in Figure 3(b). The three-dimensional structure of this double mutant GDH was subsequently refined at 3.0 Å. Its overall structure is unchanged; the C^α positions of the subunits differ with an r.m.s.d. of only 0.13 Å when compared with the wild-type *T. maritima* GDH structure. The absence of density for residue 97 in the difference Fourier map (Figure 3(a)) indicates that the aspartate in the double mutant has adopted a similar conformation to the asparagine which is present in the wild-type GDH. The introduced lysine side chain adopts a similar conformation to the *P. furiosus* GDH structure (Table 2). Its environment is conserved and no other interactions than with D380 and D97 are formed. Comparison of the structure of the double mutant with that from the *P. furiosus* GDH revealed that the five-residue network is even extended to a larger network in the double mutant GDH (Figure 2(c)). In addition to the five *P. furiosus* residues, the network involves also amino acids E177 in subunit C which is in close proximity to R64 (in subunit A), and K383 which forms now a salt bridge to D380. These two ion-pairs are also present in the wild-type *T. maritima* GDH structure. The conformation of D380 is identical in *P. furiosus* wild-type and *T. maritima* wild-type and double mutant GDHs, indicating that the extra ion-pair interaction with K383 does not change the geometry of the remaining part of the salt-bridge network. Residue E177 is disordered in the wild-type GDH (Knapp *et al.*, 1997) as well as in the double mutant. Therefore, almost no density has been observed beyond the C^β atom and the side-chain has been modelled by a database rotamer. Hence it is not expected that E177 interacts strongly with R64 but will be rather flexible, and this residue is therefore not considered as contributing to the network. Comparing the environment of the ion-pair network in *P. furiosus* GDH to the situation in the double mutant, we observed, apart from the two extra ion-pair interactions, only one additional difference within 4.0 Å distance of any of the participating residues. This concerns valine 402 in *P. furiosus* GDH which is isoleucine 399 in *T. maritima* GDH at 3 Å distance from the C^β of D380. This difference is compensated for by a nearby change from isoleucine 387 in *P. furiosus* GDH to valine 384 in *T. maritima* GDH.

Thermal inactivation and conformational stability

To study enzyme inactivation at high temperatures, the wild-type GDH and the three mutants N97D, G376 K and N97D/G376 K GDHs were incubated at 85°C and the remaining activity was determined (Table 3). The wild-type GDH is significantly more resistant against thermal inactivation than any of the three mutant GDHs, that showed two to threefold reduced half-lifes of inactivation.

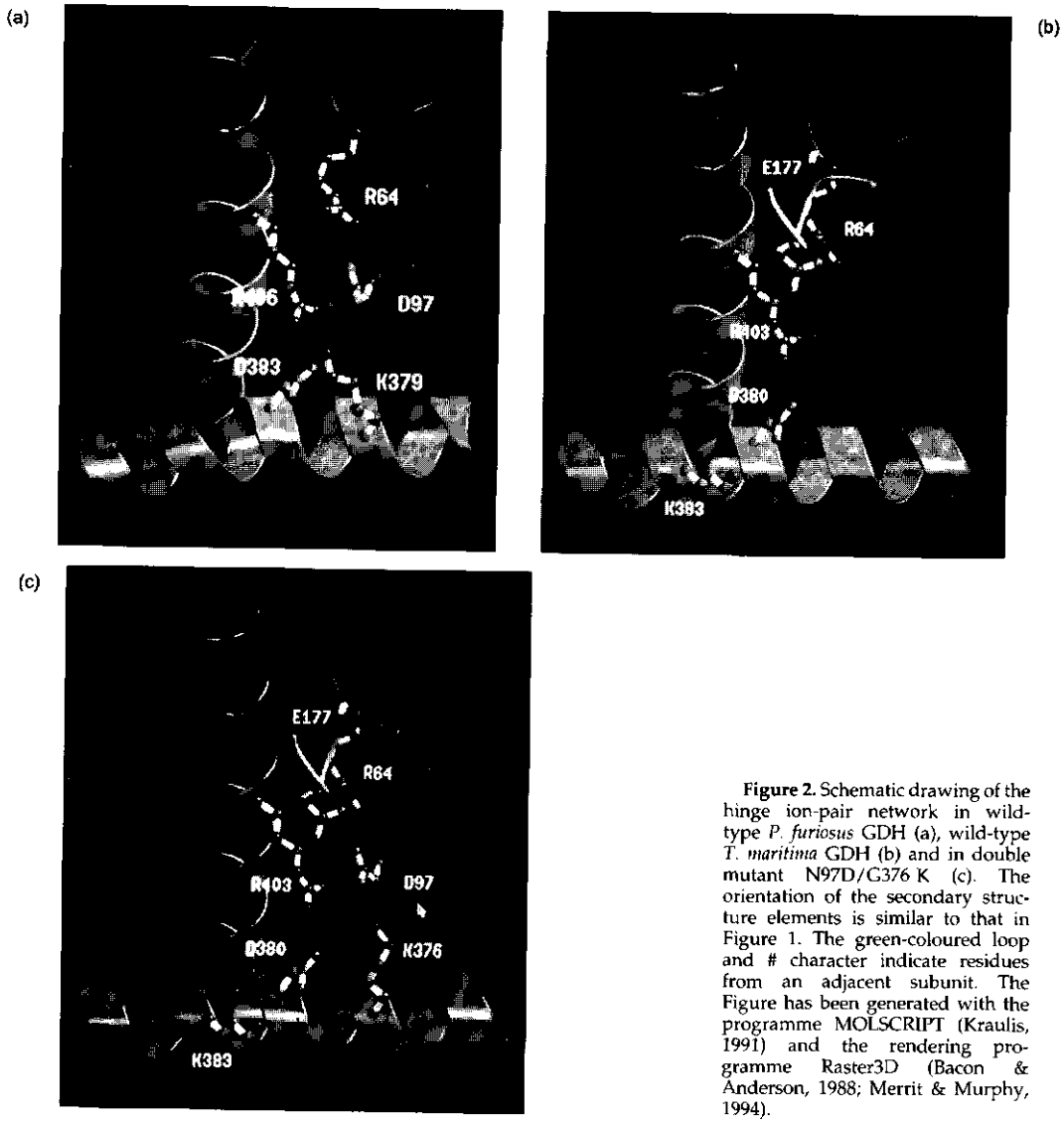


Figure 2. Schematic drawing of the hinge ion-pair network in wild-type *P. furiosus* GDH (a), wild-type *T. maritima* GDH (b) and in double mutant N97D/G376 K (c). The orientation of the secondary structure elements is similar to that in Figure 1. The green-coloured loop and # character indicate residues from an adjacent subunit. The Figure has been generated with the programme MOLSCRIPT (Kraulis, 1991) and the rendering programme Raster3D (Bacon & Anderson, 1988; Merrit & Murphy, 1994).

Melting temperatures of wild-type and the three mutant GDHs have been determined using differential scanning calorimetry (Table 3). Since the thermal unfolding transition was irreversible for all enzymes, unfolding enthalpies were not determined. No significant differences in melting temperatures were found between the wild-type GDH and any of the three mutant GDHs.

To obtain further information on the conformational stability against denaturation, unfolding of wild-type and mutant GDHs by increasing amounts of guanidinium chloride was assayed using circular dichroism spectroscopy. Identical transition midpoints were found for the wild-type and the three mutant GDHs (Table 3).

Optimum temperature and kinetic parameters

Optimum temperatures were determined by measuring enzyme activity at increasing temperatures (Figure 4). The activities of wild-type and N97D GDH showed comparable temperature dependencies, with optima around 65 C. Remarkably, the increase in activity which was found between 20 and 40 C, levelled off at higher temperatures, resulting in a "bulge" in the curves around 40 C. This behaviour was not found with

Table 1. Data collection and structural refinement statistics for mutant N97D/G376 K

Data collection:	
Space group	$P3_121$
Cell dimensions (Å)	$a = b = 147.4$; $c = 273.5$
Cell angles (deg.)	$\alpha = \beta = 90$; $\gamma = 120$
Max. resolution (Å)	3.0
Completeness to 3.0 Å (%)	94
Last shell (3.14 to 3.0 Å) (%)	60
R_{merge}[a] (%)	9.6
R_{sym}[b] (%)	5
No. of reflections	151,200 (53,329 unique)
Structural refinement:	
R factor	22.5 (R_{free} 29.8%)
R.m.s.d. of bond length (Å)	0.006
R.m.s.d. of bond angle	1.03
R.m.s.d. of dihedrals	21.1
R.m.s.d. of improper	1.06
Ramachandran plot	
- most favourable regions (%)	88
- additionally allowed regions (%)	12
- disallowed regions	None
Total number of amino acid residues	412
Total number of subunits	6
Total number of non-hydrogen atoms	19,004
R.m.s.d. C^{α} of subunit A (Å)	0.18
R.m.s.d. between the domains of the subunits (Å):	
Domain I	0.1
Domain II	0.2

[a] R_{merge} compares multiple measured reflections according to:

$$R_{merge} = \frac{\sum_{hkl} \sum_{j=1}^{N} |I_{hkl} - I_{hkl}(j)|}{\sum_{hkl} N_x I_{hkl}}$$

[b] R_{sym} compares intensities (I) of symmetry related reflections:

$$R_{sym} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

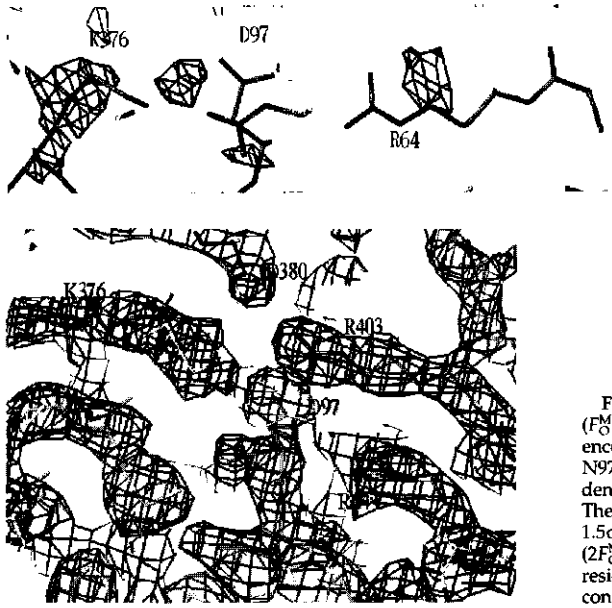


Figure 3. (a) 6-fold averaged $(F_O^{Mu} - F_C^{WT})$ $\exp(i\alpha_C^{WT})$map (difference map) of double mutant N97D/G376 K. Shown are electron densities at residues 376, 97 and 64. The map has been contoured at 1.5σ. (b) Averaged density map $(2F_O^{MU} - F_C^{MU})$ $\exp(i\alpha_C^{MU})$ of five residues involved in the network, contoured at 1.5 σ.

Table 2. Distances between the charged groups of the residues forming the hinge ion-pair network in subunit A of wild-type *P. furiosus* GDH (Yip *et al.*, 1995; structure determined at 2.2 Å resolution), wild-type *T. maritima* GDH (Knapp *et al.*, 1997; 3.0 Å) and in double mutant N97D/G376 K (this study; 3.0 Å)

Wild-type *P. furiosus* GDH	Distance (Å)	Wild-type *T. maritima* GDH	Distance (Å)	N97D/G376K *T. maritima* GDH	Distance (Å)
R64 - D97	3.1	R64 - N97	3.7	**R64 - D97**	3.4
D97 - K379	3.6	N97/G376	n.i.	**D97- K376**	3.4
K379 - D383	3.7	G376/D380	n.i.	**K376 - D380**	3.1
D383 - N386	n.i.	**D380 - K383**	4.0	**D380 - K383**	4.2
D383 - R406	3.0	**D380 - R403**	3.1	**D380 - R403**	2.9

For glutamate, aspartate and arginine, residues with two atoms that may be charged, the interaction with the smallest distance is listed. Residues forming ion-pairs are indicated in bold. N.i., no interaction.

the other two mutant GDHs. Above 65°C, G376 K and N97D/G376 K showed similar specific activities to the wild-type and N97D GDH. However, below this temperature both these mutants were much more active. At 58°C, the difference between the double mutant and N97D is almost twofold (89 units/mg *versus* 51 units/mg, respectively). Independently cultured and purified batches of wild-type and double mutant GDHs in all cases showed this higher specific activity for the double mutant. Enzyme inactivation started at 58°C for G376 K and N97D/G376 K, at 62°C for the wild-type GDH and at 66°C for N97D.

Kinetic parameters were determined in order to study whether the differences in specific activities are caused by differences in apparent substrate affinities (Table 4). For α-ketoglutarate, similar K_m values were found for all the enzymes at 25°C. With increasing temperature this affinity increased for the wild-type and N97D but remained unchanged for G376 K and N97D/G376 K. K_m values for ammonia increased at higher temperature. Interestingly, G376 K and N97D/G376 K GDH appeared to have much lower K_m values for ammonia at 25°C than the wild-type and N97D, but higher K_m values at their temperature optimum. At 25°C the wild-type GDH had the lowest K_m for NADH, followed by N97D, while G376 K and the double mutant showed the highest K_m. At 58°C, K_m values were lower than at 25°C, which contrasts with the results found for ammonia. Remarkably, for N97D GDH the decrease is almost ten-fold.

Table 3. Half-lifes of thermal inactivation, apparent melting temperatures (t_m) determined by differential scanning microcalorimetry and transition points for guanidinium chloride induced denaturation (t_s) determined for wild-type *T. maritima* GDH and mutants N97D, G376 K and N97D/G376 K

	Half-life at 85°C (minutes)	t_m (°C)	t_s (M GdmCl)
Wild-type GDH	105	93	3.5
N97D	36	91	3.5
G376K	35	92	3.5
N97D/G376K	41	91	3.6

Discussion

In order to study the role of a five-residue ion-pair network in the hinge region of *P. furiosus* GDH, the participating charged residues were introduced into the less stable GDH from *T. maritima*. Elucidation of the three-dimensional structure of the double mutant N97D/G376 K showed that the designed ion-pair interactions are indeed formed and, moreover, because of an additional charged residue present in *T. maritima* GDH, a six-residue ion-pair network was found to be formed. Detailed analysis showed that the configuration of the participating residues is similar to those in the *P. furiosus* structure and that, apart from the extra ion-pair, no additional interactions are formed. The environment of the network is very well conserved in both enzymes and there is only one conservative valine to isoleucine substitution in the immediate vicinity of the participating residues. Due to the resolution limit of 3.0 Å in the *T. maritima* GDH structures, interactions with surrounding water molecules could not be considered. Using differential scanning calorimetry and guanidinium chloride-induced unfolding, we did not observe an increase in melting temperature or a shift of the

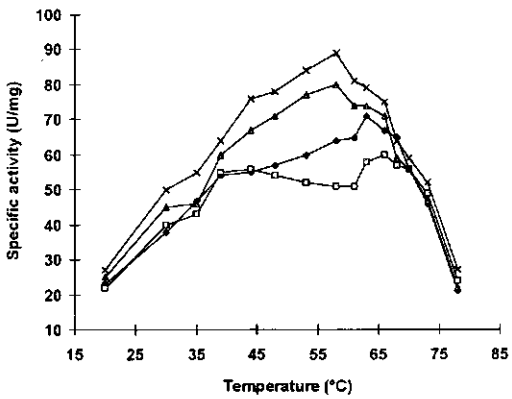


Figure 4. Specific activities as a function of temperature for wild-type *T. maritima* GDH (◆) and mutants N97D (■), G376 K(▲) and double mutant N97D/G376 K (×).

Table 4. K_m values, specific activities and catalytic efficiencies for NADH (concentration range 4 to 314 mM at 25°C and 19 to 376 mM at 58°C), α-ketoglutarate and ammonia for wild-type *T. maritima* GDH and mutants N97D, G376 K and N97D/G376 K, determined at room temperature and 58°C

	25°C			58°C		
	K_m (mM)	V_{max} (units/mg)	Catalytic efficiency[a] ($min^{-1}mg^{-1}ml$)	K_m (mM)	V_{max} (units/mg)	Catalytic efficiency[a] ($min^{-1}mg^{-1}ml$)
NADH						
Wild-type	0.395	103	261	0.101	95	941
N97D	0.692	160	231	0.073	80	1096
G376K	0.914	182	199	0.256	201	785
N97D/G376K	0.810	171	211	0.260	216	830
α-Ketoglutarate						
Wild-type	1.95	48.0	24.6	1.52	75.3	49.5
N97D	2.08	48.4	23.3	1.28	63.4	49.5
G376K	2.42	46.7	19.3	2.69	125	46.5
N97D/G376K	2.50	50.0	20.0	2.58	124	48.1
ammonia						
Wild-type	108	49.6	0.46	147	81.1	0.55
N97D	91.8	47.4	0.52	150	70.4	0.47
G376K	63.3	40.2	0.64	209	123	0.59
N97D/G376K	55.1	41.2	0.75	186	132	0.71

Reproducibility was in all cases within 10%.
[a] V_{max}/K_m.

denaturation midpoint towards higher guanidinium chloride concentrations of the mutant GDHs. Thermal inactivation of the enzymes was affected; all enzymes withstand incubation at 85°C for a considerable period of time but the wild-type shows at least a twofold longer half-life. The faster inactivation caused by the two single mutations is not additive since the half-life of the double mutant is similar to that of either of the single mutants. This might be due to a favourable ionic and/or H-bond interaction between the two introduced side-chains. For mutant N97D the faster inactivation cannot be explained by the loss of a hydrogen bridge due to the removal of the amine group; rather, in addition to the introduction of ionic interactions, a hydrogen bridge was gained in the mutant because of the presence of two hydrogen acceptors in the aspartate side-chain that can now interact with the neighbouring amine groups from lysine 376 and arginine 64. In all mutants all participating residues are surface-exposed and new residues are not introduced into a hydrophobic environment. Less obvious reasons, for example unfavourable interactions with the existing electrostatic potential, must play a role. In conclusion, no significant stabilizing effect of the introduced ion-pair network could be detected.

It is interesting to note that the K_m values of the *T. maritima* GDH for NADH and α-ketoglutarate decreased with increasing temperature, suggesting that substrate/cofactor binding has been adapted to the high growth temperature of the organism (80 C). However, for ammonia the opposite effect was observed, to the extent that there is not even an increase in catalytic efficiency at the higher temperature. In comparison with the wild-type GDH, the different mutant enzymes showed lower affinity for NADH and higher affinity for ammonia at 25°C but lower affinity for the latter substrate at 58°C. The changes of these values with temperature were different for each enzyme.

Enzymes need to compromise between a certain stability (or rigidity) in order to maintain the native conformation, and conformational flexibility in order to function. It is anticipated that the modulation of flexibility is substantially affected by interactions between regions of the enzyme where major conformational changes occur during catalysis, such as the hinge region in GDH. Sequence variations in these areas may account for the optimization of enzyme catalysis at the growth temperature of a certain organism. This idea is enforced by the observation that many of the sequence variations between the 90% identical GDHs from the hyperthermophilic archaeon *Thermococcus litoralis* (optimum growth $t = 88$°C; Neuner *et al.*, 1990) and *P. furiosus* (optimum growth $t = 100$°C; Fiala & Stetter, 1986) are located in the hinge region (Britton *et al.*, 1995). Mutants G376 K and N97D/G376 K have significantly higher specific activities at moderate temperatures (between 40 and 60°C) than the wild-type GDH and N97D, with an almost twofold difference between N97D and the double mutant at 58°C. Substitution G376 K seems to be responsible for this effect. Glycine is the amino acid which imposes the least conformational strain on a peptide chain. It has been a well-established idea that enzymes from hyperthermophiles are less flexible at room temperature than their less stable homologues, and gain comparable flexibility and maximum catalytic activity around the optimum growth temperature of the organisms in which they reside. The glycine to lysine substitution might result in a GDH that is less flexible in the hinge region and might have its optimum activity

shifted to a higher temperature. However, the two GDHs containing the G376 K substitution display higher specific activities at lower temperatures than the wild-type GDH. Above 65°C the specific activities of the mutant enzymes do not differ significantly from that of the wild-type enzyme.

In conclusion, our results indicate that large ion-pair networks which have been reported for several hyperthermostable enzymes may play an important role in the adaptation of their specific activity and temperature optima. A role in conferring stability could not be demonstrated, and considering the complexity of these multi-subunit enzymes, higher amino acid identity and structural homology might be required in order to obtain confirmation of these results. This is the first experimental report on the effect of large ion-pair networks on protein stability and activity and may contribute to a further understanding of these characteristic features of hyperthermostable enzymes.

Materials and Methods

Construction of expression plasmids for wild-type and mutated *gdh* genes

Wild-type *T. maritima* GDH was expressed in *E. coli* TG1 (Gibson, 1984) from a 1.7 kb fragment, containing the *T. maritima gdh* gene (Kort *et al.*, 1997), cloned into a pTZ19R-derivative (Mead *et al.*, 1986). Due to a spontaneous mutation during subcloning, the gene product differs from the enzyme described by Kort *et al.* (1997) at position 152 at which a methionine instead of an isoleucine is present. This methionine at position 152 is partially exposed, located at the bottom of the substrate binding domain where no important interactions are formed and is present in all investigated GDHs, including the wild-type enzyme. Because the substitution does not lead to any conformational change, it is not taken into consideration. To facilitate further manipulations, a *Bam*HI restriction site was introduced at position 542 of the coding region by PCR without changing the amino acid sequence. The single mutations leading to amino acid substitutions N97D and G376 K were introduced using cassette mutagenesis and the megaprime method (Landt *et al.*, 1990; Kuipers *et al.*, 1991) and were combined by the exchange of restriction fragments. DNA sequencing was performed to verify the sequence of the amplified DNA and to confirm the presence of the desired mutations.

Purification and crystallization of GDH

Wild-type and mutant GDH was purified from overnight cultures as described before (Knapp *et al.*, 1997). Protein concentrations were determined using an extinction coefficient at 280 nm of $\varepsilon_m = 46{,}490\ M^{-1}\ cm^{-1}$ calculated as described by Gill & von Hippel (1989). Crystals of mutant N97D/G376 K were grown as described for the wild-type GDH (Knapp *et al.*, 1997) using slightly higher concentrations of polyethylene glycol (4 to 6%, w/v).

Data collection, analysis and refinement

Data were collected as described for the wild-type GDH (Knapp *et al.*, 1997). Since the double mutant was identical to the wild-type GDH in cell parameters and spacegroup, its structure was directly refined with the program X-PLOR (Brünger, 1992) using the phases determined for the wild-type structure (Knapp *et al.*, 1997) and parameters developed by Engh & Huber (1991). 10% of all reflections were excluded from the refinement to calculate a free *R*-factor. All temperature factors were initially set to 20 Å^2. Initially a rigid body refinement was carried out using (1) the hexamer, (2) the subunits and (3) the domains as rigid bodies. After the rigid body refinement the *R*-value dropped from 34% (R_{free} = 35%) to 31% (R_{free} 32%). In the calculated electron density $((2F_O^{Mu} - F_C^{WT}) \exp i\alpha_C^{WT})$ as well as in the difference map $((F_O^{Mu} - F_C^{WT}) \exp i\alpha_C^{WT})$ the introduced lysine residue at position 376 was clearly visible. The density was improved by 6-fold averaging using the program MAIN (Turk, 1992). No density in the difference map but a clear density in the $(2F_O^{Mu} - F_C^{WT}) \exp i\alpha_C^{WT}$ map was visible at position 97 where an asparagine residue was replaced by an aspartate. The new side-chains were built into the density and the mutant structure was subsequently refined using strict non-crystallographic symmetry and grouped temperature factors. Here the conventional *R*-value dropped to 27% (R_{free} 31%). Finally the positions of the atoms were restrained with an effective energy term of 300 kcal/(mol Å^2) and in the last refinement cycles the restraints were weakened for the amino acid side-chains as well as for the crystal-contact regions. The final model was refined to a similar *R*-value as the wild-type structure (*R*-value = 22.5%, R_{free} 29.8%). The coordinates will be submitted to the Brookhaven Protein Data Bank, and are available from the authors on request until processed and released.

Characterization of wild-type and mutant GDHs

GDH activity was assayed using a Cary 4E or Beckman DU7500 spectrophotometer at 340 nm and 25°C. The 1.0 ml standard assay reaction mixtures (reductive amination) contained 100 mM potassium phosphate buffer (pH 7.6) and 10 mM α-ketoglutarate, 500 mM ammonia and 200 μM NADH. Corrections were made for the increasing auto-oxidation of NADH at elevated temperatures. The amount of enzyme was adjusted to obtain a rate of about 0.03 $\Delta A_{340\ nm}$/minute. One unit of enzyme activity is defined as the amount of enzyme which converts 1 μmol NADH per minute under standard conditions. At 25°C NADH was not saturating at concentrations as high as 310 μM. The data could, however, be accurately fitted according to Michaelis-Menten kinetics and the curve showed clear levelling off at higher concentrations. At 58°C, saturation was reached at 280 μM NADH in the case of the wild-type enzyme and N97D. For G376 K and the double mutant, saturation was almost complete at this temperature. Using NADPH instead of NADH resulted in immediate decrease of initial activity in the assay under all conditions used, and suggested product inhibition by $NADP^+$. This finding was confirmed by addition of a pulse of 5 mM $NADP^+$ to a running NADH reaction which immediately abolished all activity. Thermal inactivation was studied at an enzyme concentration of 50 μg/ml and 85°C. Residual enzyme activity was assayed at 25°C as described above. Guanidinium chloride-induced denaturation was studied as described before (Lebbink *et al.*,

1995) using an AVIV 62 DS circular dichroism spectrometer. Differential scanning microcalorimetry was carried out on a MicroCal MCS calorimeter controlled by the MCS OBSERVER program. Protein samples (1 mg/ml) were dialysed extensively against 30 mM sodium phosphate (pH 7.0) and degassed for five minutes. Samples were heated at a rate of 1 deg.C/minute.

Acknowledgements

We thank Kitty Yip, Linda Britton and Timothy Stillman (Krebs Institute, Sheffield, UK) for their contribution to designing the mutants, Walter van Dongen (Wageningen Agricultural University, NL) for helpful discussion and we are very grateful to Valerio Consalvi and Roberta Chiaraluce (Universita "La Sapienza", Roma, Italy) for critical comments on the manuscript. This work has been partially supported by contracts BIOT-CT93-0274 and BIO 4-CT96-0488 of the European Union.

References

Bacon, D. J. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219–220.

Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W. & Stillman, T. J. (1992). Subunit assembly and active site location in the structure of glutamate dehydrogenase. *Proteins: Struct. Funct. Genet.* **12**, 75–86.

Britton, K. L., Baker, P. J., Borges, K. M. M., Engel, P. C., Pasquo, A., Rice, D. W., Robb, F. T., Scandurra, R., Stillman, T. J. & Yip, K. S. P. (1995). Insights into thermal stability from a comparison of the glutamate dehydrogenase from *Pyrococcus furiosus* and *Thermococcus litoralis*. *Eur. J. Biochem.* **229**, 688–695.

Brünger, A. T. (1992). *X-PLOR (Version 3.1) Manual*, Yale University Press, New Haven and London.

Dao-pin, S., Sauer, U., Nicholson, H. & Matthews, B. W. (1991). Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. *Biochemistry*, **30**, 7142–7153.

Eggen, R. I. L., Geerling, A. C. M., Waldkötter, K., Antranikian, G. & de Vos, W. M. (1993). The glutamate dehydrogenase-encoding gene of the hyperthermophilic archaeon *Pyrococcus furiosus*: sequence, transcription and analysis of the deduced amino acid sequence. *Gene*, **132**, 143–148.

Engh, R. H. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect A*, **47**, 392–400.

Fiala, G. & Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* **145**, 56–61.

Gibson, T. J. (1984). PhD thesis, Cambridge University, Cambridge.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.

Hennig, M., Darimont, B., Sterner, R., Kirschner, K. & Jansonius, J. N. (1995). 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure*, **3**, 1295–1306.

Hennig, M., Sterner, R., Kirschner, K. & Jansonius, J. N. (1997). Crystal structure at 2.0 Å resolution of phosphoribosyl anthranilate isomerase from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability. *Biochemistry*, **36**, 6009–6016.

Horovitz, A., Serrano, L., Avron, B., Bycroft, M. & Fersht, A. R. (1990). Strength and cooperativity of contributions of surface salt bridges to protein stability. *J. Mol. Biol.* **216**, 1031–1044.

Klump, H., Di Ruggiero, J., Kessel, M., Park, J-B., Adams, M. W. W. & Robb, F. T. (1992). Glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*. Thermal denaturation and activation. *J. Biol. Chem.* **267**, 22681–22685.

Knapp, S., de Vos, W. M., Rice, D. W. & Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3.0 Å resolution. *J. Mol. Biol.* **267**, 916–932.

Korndörfer, I., Steipe, B., Huber, R., Tomschy, A. & Jaenicke, R. (1995). The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga* maritima at 2.5 Å resolution: structural basis for thermostability. *J. Mol. Biol.* **246**, 511–521.

Kort, R., Liebl, W., Labedan, B., Forterre, P., Eggen, R. I. L. & de Vos, W. M. (1997). Glutamate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: molecular characterization and phylogenetic implications. *Extremophiles*, **1**, 52–60.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

Kuipers, O. P., Boot, H. J. & de Vos, W. M. (1991). Improved site-directed mutagenesis method using PCR. *Nucl. Acids Res.* **19**, 4558.

Landt, O., Grunert, H. & Hahn, U. (1990). A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene*, **96**, 125–128.

Lebbink, J. H. G., Eggen, R. I. L., Geerling, A. C. M., Consalvi, V., Chiaraluce, R., Scandurra, R. & de Vos, W. M. (1995). Exchange of domains of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* and the mesophilic bacterium *Clostridium difficile*: effects on catalysis, thermoactivity and stability. *Protein Eng.* **8**, 1287–1294.

Mead, D. A., Szczesna-Skorupa, E. & Kemper, B. (1986). Single-stranded DNA "blue" T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Protein Eng.* **1**, 67–74.

Merrit, E. A. & Murphy, M. E. P. (1994). Raster 3D version 2.0: a program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869–873.

Neuner, A., Yannasch, H., Bellun, S. & Stetter, K. O. (1990). *Thermococcus litoralis* sp. nov.: a new species of extremely thermophilic marine archaebacteria. *Arch. Microbiol.* **153**, 205–207.

Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676–683.

Russell, R. J. M., Ferguson, J. M. C., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry*, **36**, 9983–9994.

Spassov, V. Z. & Atanasov, B. P. (1994). Spatial optimization of electrostatic interactions between the ionized groups in globular proteins. *Protein Struct. Funct. Genet.* **19**, 222–229.

Spassov, V. Z., Karshikoff, A. D. & Ladenstein, R. (1995). The optimization of protein-solvent interactions: thermostability and the role of hydrophobic and electrostatic interactions. *Protein Sci.* **4**, 1516–1527.

Stillman, T. J., Baker, P. J., Britton, K. L. & Rice, D. W. (1993). Conformational flexibility in glutamate dehydrogenase. The role of water in substrate recognition and catalysis. *J. Mol. Biol.* **234**, 1131–1139.

Tomschy, A., Böhm, G. & Jaenicke, R. (1994). The effect of ion-pairs on the thermal stability of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *Protein Eng.* **7**, 1471–1478.

Turk, D. (1992). PhD thesis, Technical University, München.

Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, R., Scandurra, R. & Rice, D. W. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure*, **3**, 1147–1158.

Yip, K. S. P., Britton, K. L., Stillman, T. J., Lebbink, J. H. G., de Vos, W. M., Robb, F. T. & Rice, D. W. (1998). Insights into the molecular basis of thermostability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* In the press.

Edited by R. Huber

(Received 30 October 1997; received in revised form 9 April 1998; accepted 9 April 1998)

Chapter 6

Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. II. Construction of a 16-residue ion-pair network at the subunit interface.

Joyce H.G. Lebbink, Stefan Knapp, John van der Oost, David W. Rice, Rudolf Ladenstein, and Willem M. de Vos.

Reprinted with permission from
Journal of Molecular Biology **289**, 357-369, 1999

JMB

Engineering Activity and Stability of *Thermotoga maritima* Glutamate Dehydrogenase. II†: Construction of a 16-Residue Ion-pair Network at the Subunit Interface

Joyce H. G. Lebbink[1], Stefan Knapp[2], John van der Oost[1], David Rice[3], Rudolf Ladenstein[2] and Willem M. de Vos[1]*

[1]*Laboratory of Microbiology Wageningen Agricultural University, Hesselink van Suchtelenweg 4, NL-6703 CT Wageningen The Netherlands*

[2]*Center for Structural Biochemistry, NOVUM Karolinska Institute S-14157 Huddinge, Sweden*

[3]*The Krebs Institute for Biomolecular Research Department of Molecular Biology and Biotechnology University of Sheffield, PO Box 594 Sheffield, S10 2UH, UK*

**Corresponding author*

The role of an 18-residue ion-pair network, that is present in the glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*, in conferring stability to other, less stable homologous enzymes, has been studied by introducing four new charged amino acid residues into the subunit interface of glutamate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. These two GDHs are 55% identical in amino acid sequence, differ greatly in thermo-activity and stability and derive from microbes with different phylogenetic positions. Amino acid substitutions were introduced as single mutations as well as in several combinations. Elucidation of the crystal structure of the quadruple mutant S128R/T158E/N117R/S160E *T. maritima* glutamate dehydrogenase showed that all anticipated ion-pairs are formed and that a 16-residue ion-pair network is present. Enlargement of existing networks by single amino acid substitutions unexpectedly resulted in a decrease in resistance towards thermal inactivation and thermal denaturation. However, combination of destabilizing single mutations in most cases restored stability, indicating the need for balanced charges at subunit interfaces and high cooperativity between the different members of the network. Combination of the three destabilizing mutations in triple mutant S128R/T158E/N117R resulted in an enzyme with a 30 minutes longer half-life of inactivation at 85 °C, a 3 °C higher temperature optimum for catalysis, and a 0.5 °C higher apparent melting temperature than that of wild-type glutamate dehydrogenase. These findings confirm the hypothesis that large ion-pair networks do indeed stabilize enzymes from hyperthermophilic organisms.

Keywords: glutamate dehydrogenase; thermostability; subunit interface ion-pair network; *Pyrococcus furiosus*; *Thermotoga maritima*

Introduction

The recent elucidation of a large number of three-dimensional structures from enzymes isolated from hyperthermophilic bacteria and archaea, has revealed a strong correlation between the number of ion-pairs and enzyme thermostability (Aguilar *et al.*, 1997; Hennig *et al.*, 1995, 1997; Knapp *et al.*, 1997; Körndorfer *et al.*, 1995; Russell *et al.*, 1997; Yip *et al.*, 1995). Furthermore, with increasing temperature, these ion-pairs tend to be more and more organized in large networks. These findings strongly suggest an important role for ion-pair networks in the thermal adaptation of proteins. Until recently, experimental evidence testing this hypothesis has been scarce. Studies on ion-pairs in mesophilic systems gave rise to controversial results (Dao-pin *et al.*, 1991; Anderson *et al.*, 1990; Waldburger *et al.*, 1995). It has been argued that the gain in electrostatic energy upon formation of salt-bridges, is largely offset by the unfavorable entropy of side-chain fixation as well as desolva-

Abbreviations used: GDH, glutamate dehydrogenase.

†Paper I in this series is Lebbink *et al.* (1998).

E-mail address of the corresponding author: willem.devos@algemeen.micr.wau.nl

tion energies (Dao-pin *et al.*, 1991; Hendsch & Tidor, 1994). However, arrangement of charged residues into networks at high temperatures may play a key role in the acquisition of thermostability of proteins, since (i) the dielectric constant decreases significantly with temperature, (ii) each new member in a network of ion-pairs requires the desolvation of just one residue, (iii) hydration free energies are reduced at elevated temperatures and because these reductions are considerably larger for charged residues than for hydrophobic isosteres, ion-pairs may become more stabilizing than hydrophobic interactions at elevated temperatures (Elcock & McCammon, 1997; Elcock, 1998). As a result the argument that in comparison with hydrophobic isosteres, surface salt bridges destabilize proteins, may no longer be true for the optimal temperatures of hyperthermostable proteins.

The first experimental results on ion-pair networks in hyperthermostable proteins are currently appearing and shed more light on the mechanisms involved. The pH-induced denaturation of glutamate dehydrogenase (Chiaruluce *et al.*, 1997), rubredoxin (Cavagnero *et al.*, 1995) and methionine aminopeptidase (Ogasahara *et al.*, 1998) from the hyperthermophilic archaeon *Pyrococcus furiosus*, emphasized the role of electrostatic interactions in the acquisition of enzyme thermostability. Comparison of the unfolding kinetics of pyrococcal rubredoxin with that of the mesophile *Clostridium pasteurianum* showed an enhanced kinetic barrier towards thermal unfolding (Cavagnero *et al.*, 1998). Likewise, an increase in the free energy of activation for thermal unfolding was described for salt-bridges in a comparison of wild-type and mutant glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* in which ionic interactions were disrupted (Pappenberger *et al.*, 1997). Removal of a central charged residue that, according to homology with *P. furiosus* GDH, belongs to an 18-residue ion-pair network of glutamate dehydrogenase from *P. kodakaraensis* resulted in faster thermal inactivation and inability to attain native-like tertiary structure (Rahman *et al.*, 1998). Moreover, the introduction of a central charged residue, presumably restoring a six-residue ion-pair network as present in the very homologous GDH from *P. furiosus*, resulted in an increased temperature optimum and half-life of inactivation at 100 °C (Rahman *et al.*, 1998). The same substitution in *Thermococcus litoralis* GDH, was unexpectedly found to be destabilizing. A second-site mutation, not being part of the network, was required to raise the stability of the mutant GDH over that of the wild-type enzyme (Vetriani *et al.*, 1998).

T. maritima GDH is a thermostable enzyme with a melting temperature of 93 °C and its structure is stabilized by ion-pairs (intrasubunit) and hydrophobic interactions (intersubunit) (Knapp *et al.*, 1997; Kort *et al.*, 1997). It shares considerable amino acid identity (55 %) with the GDH from *P. furiosus* (Eggen *et al.*, 1993) and high structural homology (Knapp *et al.*, 1997; Yip *et al.*, 1995). However, *P. furiosus* GDH has a melting temperature of 113 °C (Klump *et al.*, 1992). In addition, it contains significantly more inter- and intrasubunit ion-pairs and larger ion-pair networks than the *T. maritima* enzyme. Recently, we described the role of a five-residue intrasubunit ion-pair network in the hinge region of *P.furiosus* GDH, of which the missing charged amino acids were successfully introduced into *T. maritima* GDH (Lebbink *et al.*, 1998). Thermal inactivation and thermal and guanidinium chloride-induced denaturation showed that this ion-pair network did not increase the stability of *T. maritima* GDH, but did result in significant differences in specific activities and substrate affinities.

We continued our strategy of introducing ion-pair networks into *T. maritima* GDH here and changed focus from intrasubunit interactions towards ionic interactions crossing subunit interfaces. In *P. furiosus* GDH the subunit interfaces are highly charged because of the presence of multiple extensive ion-pair networks (Yip *et al.*, 1995). The largest network comprises 18 residues, cross-links secondary structure elements from four different subunits, and is three times present in each hexamer at the trimer interface. In *T. maritima* GDH this network is largely reduced in size and fragmented. We introduced multiple substitutions in this less stable GDH in order to restore the *P. furiosus* GDH network, and compared structural and biochemical properties of the mutant GDHs to those of the wild-type. Here we describe the results of this study that provided insight into the role of extensive ion-pair interactions at subunit interfaces, indicated the need for balanced charges, and showed that a stabilization strategy employed by an archaeal enzyme can be succesfully introduced into a bacterial homolog that has a different mode of thermostabilization.

Results

Design of *T. maritima* GDH subunit interface mutants

In general, GDH is a homohexamer with its subunits arranged into two trimers that are stacked upside down on top of each other (Figure 1(a); Baker *et al.*, 1992). At this resulting trimer interface the largest ion-pair networks were identified in the GDH from *P. furiosus* (Figure 1(b); Yip *et al.*, 1995). Three 18-residue networks are present, each connecting secondary structure elements belonging to four different subunits. Each network is made up from two symmetry-related parts that are connected by multiple interactions crossing the trimer interface. Ion-pairs are formed between the side-chains of Glu160, Asp157, Arg117, Glu120, Arg124, Glu158 and Arg128 (subunit C), Arg414 and Lys418 (subunit A) and their 2-fold related partners (Figure 2(a) and Table 1). The majority of these interactions are salt-bridges, meaning that apart from the ionic interaction, also hydrogen

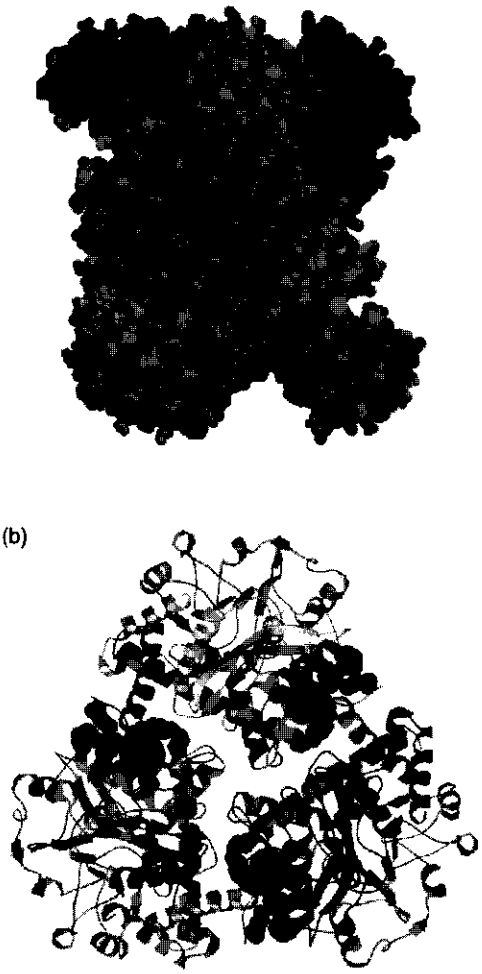


Figure 1. (a) Space-filling model of *P. furiosus* GDH showing the arrangement of six identical subunits into two trimers that are stacked upside down on top of each other. Furthermore the composition of each subunit in two domains is visible, with the substrate binding domains forming the core of the hexamer and being responsible for the majority of intersubunit contacts, and the cofactor-binding domains projected outwards. (b) View of one trimer along the 3-fold axis, looking down upon the trimer interface. The secondary structure elements of three subunits are shown in different colors. Amino acid side-chains belonging to the 18-residue ion-pair network are depicted in space-filling presentation. Only one half of each network is visible, the symmetry-related half belongs to the second trimer. The Figures have been generated with the programme MOLSCRIPT (Kraulis, 1991) together with the rendering programme Raster3D (Bacon & Anderson, 1988; Merrit & Murphy, 1994).

bridges are formed. Out of 18 ion-pair interactions, eight are crossing subunit interfaces, and multiple residues form ion-pairs with three oppositely charged partners (Glu120, Arg124 and Glu158).

T. maritima GDH is 55% identical on amino acid level with the pyrococcal enzyme and has a 20°C lower apparent melting temperature. The intersubunit network is drastically reduced in size and fragmented into two parts (Figure 2(b); Table 1). The corresponding residue for the pyrococcal Lys418 in *T. maritima* GDH is Pro415, therefore no ion-pair is formed by this residue. Further interactions between the network and this residue cannot be analyzed due to the absence of electron density for the three C-terminal residues 413-415 in the wild-type *T. maritima* GDH structure. Two residues at the periphery of the ion-pair network, Arg117 and Glu160, are exchanged for asparagine and serine, respectively, in *T. maritima*. Apart from the removal of the ionic interaction also the hydrogen-bonding capacity is reduced, because the side-chains of asparagine and serine are shorter than those of arginine and glutamate. The central arginine 128 and glutamate 158 of *P. furiosus* GDH correspond to serine and threonine in *T. maritima* GDH. This results in a complete loss of all ionic and hydrogen-bond interactions between the two symmetry-related parts of the network. The net effect is the presence of two small four-residue ion-pair networks in *T. maritima* GDH, each containing only three ion-pair interactions, that no longer cross the trimer interface.

In order to study the role of the 18-residue network in conferring thermo-activity and stability to GDH, we introduced the missing charged residues into the GDH from *T. maritima*. The C-terminal Pro415 was ignored because the absence of density in the wild-type structure would not allow structural interpretation of any substitution at this position. Four single mutants were constructed; containing the peripheral Asn117 to Arg (this mutant GDH will be referred to as N117R) and Ser160 to Glu substitutions (S160E) and the central Ser128 to Arg (S128R) and Thr158 to Glu substitutions (T158E). The central double mutant S128R/T158E and a second double mutant T158E/S160E were constructed by combination of the respective single mutations. Subsequently two triple mutants were constructed by addition of either the peripheral N117R or S160E to central double mutant S128R/T158E. In the quadruple mutant S128R/T158E/N117R/S160E all single substitutions were combined.

gdh gene expression in *Escherichia coli* and purification of wild-type and mutant enzymes

Expression in *E.coli* of wild-type and mutated *T. maritima gdh* genes in all cases resulted in the overproduction of a new 260 kDa protein, exceeding 5% of total cell protein as judged by SDS-PAGE (not shown). No differences in expression levels could be detected. In contrast, large

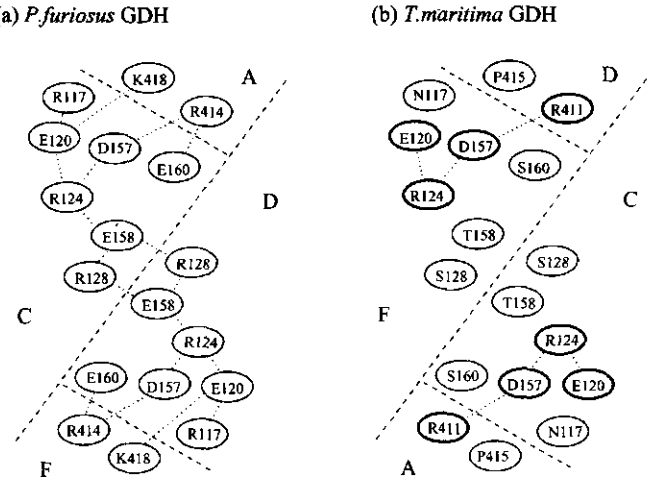


Figure 2. Schematic presentation of the 18-residue ion-pair network in *P. furiosus* GDH (a) and the homologous residues in *T. maritima* GDH (b). Residues forming ion-pairs are connected by dotted lines, broken lines indicate subunit interfaces.

differences in specific GDH activity in cell-free extracts were measured. Wild-type GDH has a specific activity of 2.5 units/mg at 20 °C, while mutants T158E, S128R/T158E and S128R/T158E/N117R showed activities as low as 0.02 unit/mg. Substitution T158E seems to be responsible for this. During heat-treatment at 70 °C, the GDH activity of the wild-type cell-free extract remained constant, whereas that of the aforementioned three mutants showed a progressive increase in total activity. This increase is more pronounced for the double and the triple mutant than for single mutant T158E (data not shown). However, wild-type activity levels were not reached, because simultaneous inactivation of the mutants occurred. Not only heat-treatment but also the presence of mutation S160E resulted in "active-state" enzyme for the mutants containing the T158E substitution, since all mutants (and most convincingly, double mutant T158E/S160E) containing this combination were fully active. Native PAGE of heat-incubated extracts revealed that all enzymes were hexamers, with sharp bands for all mutant enzymes except for T158E that migrated as a diffuse band (results not shown).

Table 1. Interatomic distances between charged groups of residues forming the 18-residue ion-pair network

Wild-type *P. furiosus* GDH	Distance (Å)	Wild-type *T. maritima* GDH	Distance (Å)	Quadruple mutant *T. maritima* GDH	Distance (Å)
Glu160 C-Arg414 A	**2.6**	Ser160 F/Arg411 D	n.i.	**Glu160 F-Arg411 D**	**2.8**
Glu120 C-Lys418 A	**2.8**	Glu120 F/Pro415 D	n.i.	Glu120 F/Pro415 D	n.i.
Glu120 C-Arg117 C	**2.8**	Glu120 F/Asn117 F	n.i.	**Glu120 F-Arg117 F**	**3.8**
Glu120 C-Arg124 C	**3.0**	**Glu120 F-Arg124 F**	**3.1**	**Glu120 F-Arg124 F**	**2.7**
Asp157 C-Arg124 C	**2.7**	**Asp157 F-Arg124 F**	**3.0**	**Asp157 F-Arg124 F**	**3.2**
Asp157 C-Arg414 A	**3.1**	**Asp157 F-Arg411 D**	**2.7**	**Asp157 F-Arg411 D**	**3.2**
Glu158 C-Arg124 C	**2.9**	Thr158 F/Arg124 F	2.9	**Glu158 F-Arg124 F**	**3.3**
Glu158 C-Arg128 C	**3.0**	Thr158 F/Ser128 F	n.i.	**Glu158 F-Arg128 F**	**3.4**
Glu158 C-Arg128 D	**2.8**	Thr158 F/Ser158 C	n.i.	**Glu158 F-Arg128 C**	**3.2**
Glu158 D-Arg128 C	**2.8**	Thr158 C/Ser128 F	n.i.	**Glu158 C-Arg128 F**	**3.4**
Glu158 D-Arg128 D	**3.0**	Thr158 C/Ser128 C	n.i.	**Glu158 C-Arg128 C**	**3.1**
Glu158 D-Arg124 D	**2.9**	Thr158 C/Arg124 C	3.1	**Glu158 C-Arg124 C**	**3.5**
Asp157 D-Arg124 D	**2.7**	**Asp157 C-Arg124 C**	**2.9**	**Asp157 C-Arg124 C**	**2.9**
Asp157 D-Arg414 F	**3.1**	**Asp157 C-Arg411 A**	**3.2**	**Asp157 C-Arg411 A**	**2.9**
Glu120 D-Arg124 D	**3.0**	**Glu120 C-Arg124 C**	**3.2**	**Glu120 C-Arg124 C**	**2.8**
Glu120 D-Lys418 F	**2.8**	Glu120 C/Pro415 A	n.i.	Glu120 C/Pro415 A	n.i.
Glu120 D-Arg117 D	**2.8**	Glu120 C/Asn117 C	n.i.	**Glu120 C-Arg117 C**	**3.7**
Glu160 D-Arg414 F	**2.6**	Ser160 C/Arg411 A	n.i.	**Glu160 C-Arg411 A**	**3.0**

Interatomic distances between the charged groups of the residues forming the 18-residue ion-pair network at the subunit interface of wild-type *P. furiosus* GDH (Yip *et al.*, 1995; structure determined at 2.2 Å resolution), wild-type *T. maritima* GDH (Knapp *et al.*, 1997; 3.0 Å) and in quadruple mutant *T. maritima* GDH (this study; 2.9 Å). For glutamate, aspartate and arginine, residues with multiple atoms that may be charged, the interaction with the smallest distance is listed. Residues forming ion-pairs are indicated in bold. n.i., no interaction.

Wild-type and single mutants N117R, S128R, T158E, double mutant S128R/T158E, both triple mutants and the quadruple mutant GDHs were purified to homogeneity as judged by SDS-PAGE, and used for further structural and biochemical analysis. Based on thermal inactivation of heat-stable cell-free extracts, the effect of S160E is similar in all mutant GDHs containing this substitution. A detailed biochemical characterization of the effect of S160E was therefore restricted to triple mutant S128R/T158E/S160E and the quadruple mutant. During the purification procedure, the double and triple mutant GDHs containing the T158E substitution, reached full activity either because of prolonged handling at room-temperature, or because of favorable interactions with for example the cofactor analog on the affinity column, whereas single mutant T158E never reached the same specific activity as wild-type GDH (see below).

Structure of quadruple mutant GDH

The quadruple mutant GDH was crystallized under similar conditions and in the same space-group as the wild-type *T. maritima* GDH and the hinge double mutant N97D/G376 K (Knapp *et al.*, 1997; Lebbink *et al.*, 1998; Table 2). The electron density was significantly improved by averaging using the non-crystallographic symmetry. In the wild-type structure the first three and last three amino acids were not included because no density for these residues has been observed. In the quadruple mutant, however, weak density for the last three amino acids was observed in a 6-fold averaged $(2F_o^{MU} - F_c^{WT})$ exp $i\alpha^{WT}$ map probably due to the improved quality of the diffraction data. The three C-terminal residues were built, and their positions and grouped and restrained *B*-values were refined. However, all three residues had *B*-values of 90 Å^2 and higher and were therefore still not included into the final model of the enzyme. New densities at the sites of the amino acid substitutions in the network were clearly visible in the $(2F_o^{MU} - F_c^{WT})$ exp $i\alpha^{WT}$ map (Figure 3(a), (c), (e)). The side-chains of the new arginine and glutamate residues were built into the model and the structure was subsequently refined at 2.9 Å, a slightly better resolution than for the wild-type GDH and hinge mutant N97D/G376 K (Figure 3(b), (d), (f); Knapp *et al.*, 1997; Lebbink *et al.*, 1998; Table 2). The overall structure of the quadruple mutant was not different from that of the wild-type GDH. The C^α positions of the hexamer differ with an r.m.s.d. of 0.57 Å and those of the subunits vary between 0.32 and 0.42 Å. These differences are mainly due to variations in hinge opening angles between the two domains of each subunit, which are likely caused by the small differences in crystallization conditions.

Analysis of all the residues involved in the large intersubunit ion-pair networks revealed

Table 2. Data collection and structural refinement statistics for wild-type *T. maritima* GDH and quadruple mutant S128R/T158E/N117R/S160E

	Wild-type GDH	Quadruple mutant
A. *Data collection*		
Spacegroup	$P3_121$	$P3_121$
Cell dimensions (Å)	a=b=147.26; c=273.82	a=b=145.1; c=272.5
Cell angles (deg.)	α=β=90; γ=120	α=β=90; γ=120
Max. resolution (Å)	3.0	2.9
Completeness to 3 Å (%)	94	98
last shell	75 (3.14-3 Å)	97.1 (2.9-3 Å)
R_{merge} (%)	9.1	7.0
R_{sym} (%)	5.0	4.5
No. of reflections (unique)	220,980 (57,307)	153,233 (72,518)
B. *Structural refinement*		
R factor (%)	22.5 (R_{free}: 29.5)	21.1 (R_{free}: 27.4)
Overall temperature factors (Å^2)		
Main-chain	33.5	47
Side-chain	47.1	56
All atoms	40.2	51
Deviations from standard bond length and angles		
r.m.s.d. bond length (Å)	0.006	0.005
r.m.s.d. bond angles (deg.)	1.032	0.906
r.m.s.d. bond dihedrals (deg.)	21.141	25.7
r.m.s.d. bond improper (deg.)	1.058	0.537
Ramachandran plot		
Most favored regions (%)	89.8	87.9
Additional allowed (%)	9.4	11.5
Generously allowed (%)	0.8	0.6
Disallowed (%)	0	0
r.m.s.d. deviation of C^α positions from wild-type		
Hexamer	-	0.57
Subunit A	-	0.32

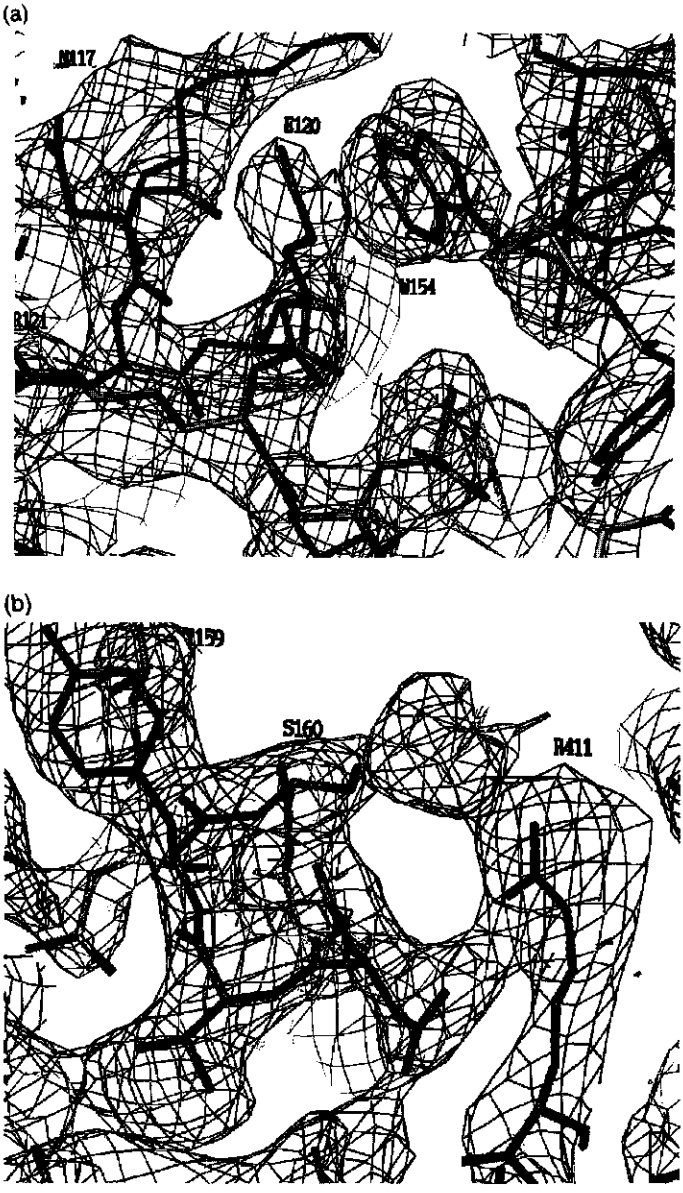


Figure 3. *(Legend shown on page 364)*

that the conserved *T. maritima* GDH residues have retained their conformation in comparison with the wild-type enzyme and furthermore, that all new side-chains adopt conformations as designed, i.e. similar to that of their counterparts in *P. furiosus* GDH. Table 1 lists the interatomic distances between the involved charged groups in one of the three networks. All distances are well within the cut-off value for ion-pairs of 4.0 Å (Barlow & Thornton, 1983). This means that in the quadruple mutant a 16-residue ion-pair network is present, which is almost as large as the network in *P. furiosus* GDH (Figure 4).

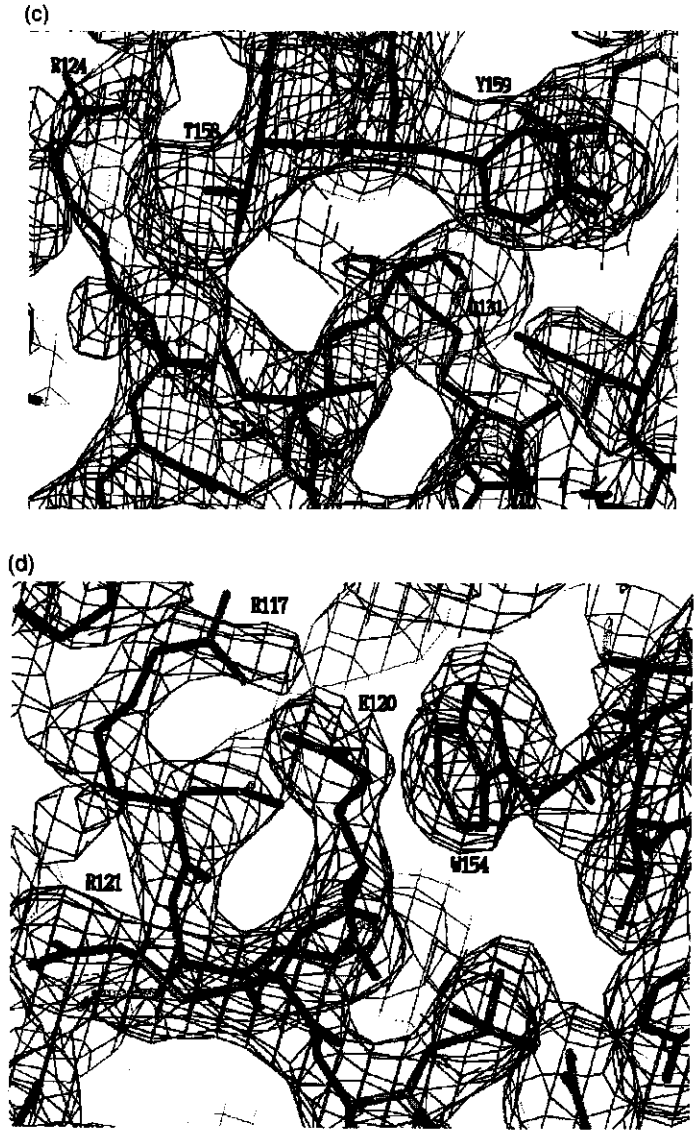


Figure 3. (*Legend shown on page 364*)

Thermoactivity and resistance towards thermal inactivation and denaturation

Wild-type *T. maritima* GDH showed optimal activity of 72 units/mg at 65°C, identical to the values determined earlier (Lebbink *et al.*, 1998). Specific activities at the respective temperature optima of the mutant GDHs showed that most of these have similar activity to the wild-type GDH, with N117R somewhat lower and S128R/T158E slightly higher values (Table 3). Only single mutant T158E shows a large change in comparison with the wild-type GDH, with 40 units/mg its activity is almost two times lower.

The effect of the introduced ion-pairs was addressed by comparing thermo-activity and resistance towards thermal inactivation and thermal denaturation of the wild-type and the mutant

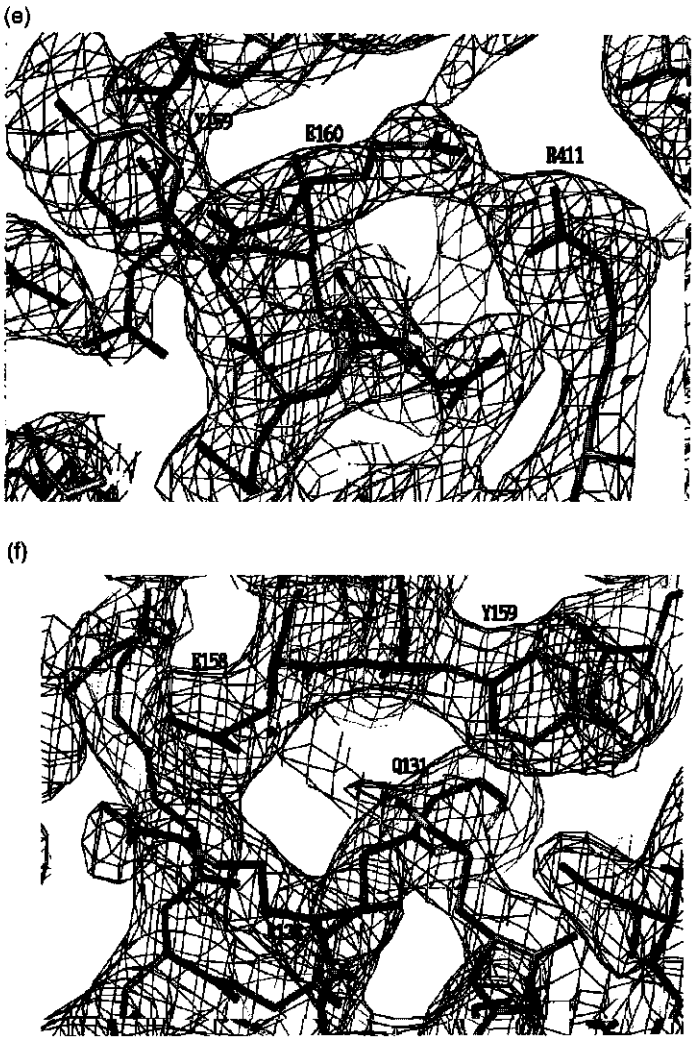


Figure 3. 6-fold averaged electron density maps ($(2F_o^{MU} - F_c^{WT})$ exp $i\alpha^{WT}$) of quadruple mutant GDH showing new density at positions of substition N117R (a, d), S160E (b, e) and S128R and T158E (c, f) before (a, b, c) and after (d, e, f) refinement.

GDHs (Table 3). Similar trends were found with all three techniques. Single mutants S128R and especially T158E have lower temperature optima, lower apparent melting temperatures and inactivate much faster than the wild-type GDH. Also single mutant N117R is less stable than the wild-type although differences are only marginal. The combination of the two destabilizing central mutations into double mutant S128R/T158E results in a remarkable restoration of stability compared to the wild-type GDH. The apparent melting temperature of this mutant is decreased by only 0.4 °C and its half-life at 85 °C is only 20 % lower and their temperature optima are identical. A remarkable effect was observed by the addition of arginine 117, which on its own slightly destabilizes, but added to the central double mutant shows a significant increase of the half-life by a factor of 1.5 and an increase in apparent melting temperature of 0.9 °C. Moreover, this triple mutant S128R/T158E/N117R has a higher melting temperature and temperature optimum than the wild-type GDH and is more resistant to thermal inactivation. In all cases substitution S160E is deleterious, as indicated by the low half-lives of the second triple and the quadruple mutant at 85 °C. Again, however, the triple mutant has been stabilized by the additional N117R substitution.

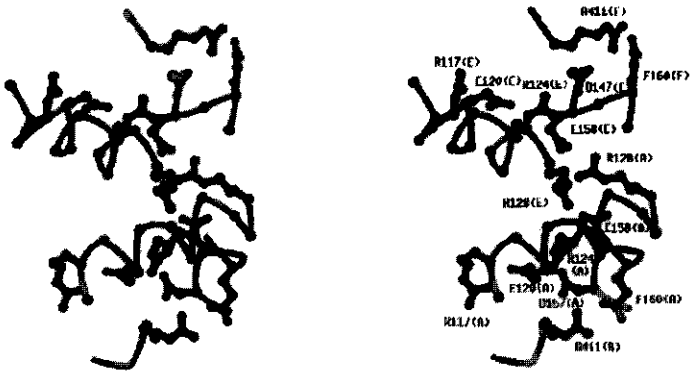


Figure 4. Stereoview of the 16-residue network in *T. maritima* quadruple mutant GDH.

Table 3. Biochemical characterization of wild-type and mutant *T. maritima* GDHs

	t_{opt} (°C)	Specific activity at t_{opt} (units/mg)	$t_{1/2}$ at 85 °C (min)	t_m (°C)
Wild-type GDH	65	72	209 ± 13	89.1
S128R	58	77	16 ± 0.2	87.0
T158E	55	40	<0.1 -	83.1
N117R	63	58	192 ± 24	88.7
S128R/T158E	65	92	162 ± 6.6	88.7
S128R/T158E/N117R	68	74	240 ± 12	89.6
S128R/T158E/S160E	60	79	1 ± 0.3	88.0
Quadruple mutant	63	76	5 ± 1	88.3

The first two columns list the specific activity of each enzyme at its respective optimum temperature for catalysis. Thermal inactivation of wild-type and mutant *T. maritima* GDH at 85 °C is tabulated as half-lives of inactivation in minutes with standard errors. The last column lists the apparent melting temperatures as determined by differential scanning microcalorimetry. The estimated error in t_m is 0.2 °C.

Discussion

The largest ion-pair network in any hyperthermostable protein reported to date is the 18-residue network that is three times present at the trimer interface of the GDH from *P. furiosus* (Yip *et al.*, 1995). In order to study the role of this network in conferring thermostability, we decided to introduce it into the less thermostable GDH from *T. maritima*, analogous to our study of the hinge ion-pair network that was described before (Lebbink *et al.*, 1998). GDH mutants containing single or multiple amino acid substitutions were purified from *E. coli* and biochemically characterized with respect to optimal activity and their resistance towards thermal inactivation and denaturation. The three-dimensional structure of the quadruple mutant was elucidated, and together with the structure of the double mutant in the hinge region (Lebbink *et al.*, 1998), has been crucial in the interpretation of the biochemical data because they show that the predicted networks are indeed formed and that unexpected results cannot be explained by the absence of expected interactions. Furthermore, the structures show that although especially glutamate, arginine and lysine side-chains are long and flexible, they do in all cases adopt the conformation favorable for formation of the ion-pair interactions. This suggests that oppositely charged groups in close proximity are not only a prerequisite for, but also unavoidably lead to the formation of ion-pair interactions. Because most of the interactions within the network in *P. furiosus* GDH are salt-bridges, the effect of the network on the stability of the hexamer will be conferred by both ionic interactions and hydrogen bridges. Hydrogen bond formation in the *T. maritima* structures, however, is difficult to assess because of the medium resolution of the models (2.9 and 3.0 Å , while that of *P. furiosus* GDH is 2.2 Å). Furthermore, crystallographic *B*-factors for several of the charged side-chains are considerably high. Grouped *B*-factor values for Arg117 and Arg128 , for example, are 80 and 95 Å^2, respectively, indicating that these two residues are very flexible. However, the position of the side-chains was clearly visible in the $(2F_o^{MU} - F_c^{WT})$ exp $i\alpha^{WT}$ map. Even if both H-bond distance and H-bond geometry criteria would be satisfied, the high mobility of the side-chains indicates that the actual formation of these interactions will only occur for very short periods of time. In contrast, the assignment of ion-pairs does not

suffer from such considerations; only a distance criterium has to be fulfilled and even if large movements temporarily result in charged groups being further than 4.0 Å apart, the interaction may easily be restored because of the long-range character of electrostatic interactions. These high *B*-factors may reflect just why ion-pair networks may be favorable stabilizing mechanisms at high temperatures; the long-range nature of electrostatics allows considerable movement of side-chains without interactions being broken and at the same time lowering the entropic cost of side-chain fixation. Similar conclusions were drawn from the crystal structure of a thermostable β-glycosidase that contains not only many ion-pair interactions, but in addition contains buried solvent networks, that may stabilize the enzyme in a similar manner (Aguilar *et al.*, 1997).

The slight destabilization of *T. maritima* GDH by the introduction of an extra intersubunit ion-pair between Glu120 and Arg117, may be explained by the repulsion between the nitrogen at the epsilon-position of the introduced arginine side-chain and the 3.9 Å distant N^{ε} atom of arginine 44. In *P. furiosus* GDH this interaction is not present because the corresponding glutamate 44 does not interact with arginine 117, but rather forms an ion-pair with the arginine 121. The destabilization of single mutants S128R and T158E and the subsequent restoration of enzyme stability upon their combination, indicates that the presence of unbalanced, similar charges in close proximity is highly deleterious and shows the need for balanced charges. Electrostatic interactions are generally optimized in proteins in a way that leads to the rejection of conformers with repulsive charge-charge interactions (Spassov *et al.*, 1994). Considering not only the destabilization but also the inactivation by the T158E substitution, the effect of this mutation may well be more than local. In native PAGE this enzyme migrates as a diffuse band which might indicate the presence of multiple (hexameric) conformations. This hypothesis is supported by the fact that heat-treatment is able to increase the activity of mutants containing T158E, reminiscent of heat-incubation being a trigger for a subtle conformational change in *P. furiosus* GDH that is needed for the enzyme to reach full activity (Consalvi *et al.*, 1996; Lebbink *et al.*, 1995). Furthermore, the addition of substitution S160E, which is located only two residues apart in the α-helix, appears to be able to restore full activity to all mutants containing the T158E mutation. This suggests that if the T158E mutants are indeed trapped in a local energy minimum, the new ion-pair interaction between Glu160 and Arg411 is sufficient to pull the enzyme into its fully active conformation. It is therefore surprising that, although a strong ion-pair interaction seems to be formed in the mutants containing the S160E substitution, all these mutants are considerably destabilized. The importance of correct interactions formed by central residue Glu158 is also shown by the observation that its removal from the network in GDH from *P. kodakaraensis* results in an enzyme that is no longer capable of attaining its native conformation (Rahman *et al.*, 1998).

The combined effects of the central charged residues have clear evolutionary implications, and are in accordance with the results from a comparative study on the breakdown of the network going from the most thermostable GDHs down along the temperature spectrum to mesophilic GDHs (Yip *et al.*, 1998). In all hyperthermostable enzymes the central charge cluster connects the symmetry-related parts of the ion-pair network across the subunit interface, while in all less stable enzymes both charges have been substituted. Most probably *in vivo* only one substitution leads to an enzyme not able to function properly, as was shown *in vitro* here and by Rahman *et al.* (1998). Although there is a clear stabilizing interaction between the central residues Arg128 and Glu158, it is not strong enough to compensate for the loss of stability conferred by the serine and threonine side-chains to the wild-type enzyme. This means either that ion-pair interactions may not contribute more to enzyme stability than polar interactions, or alternatively, regarding the only moderate homology between the wild-type archaeal and bacterial enzymes, it could indicate that the bacterial enzyme needs further optimization in order for the electrostatics to exert their full effect on the enzyme's stability. This is supported by the observation that the subunit interfaces in *T. maritima* GDH are much more hydrophobically optimized than in the pyrococcal enzyme (Knapp *et al.*, 1997) and by the work by Vetriani *et al.* (1998). They showed that restoration of an ion-pair network in *T. litoralis* GDH (this GDH is 87% identical on amino acid level to *P. furiosus* GDH, while identity between *T. maritima* and *P. furiosus* GDH is only 55%) only resulted in stabilization when accompanied by a second-site mutation (Vetriani *et al.*, 1998).

The effect of the combination of mutations in stabilized S128R/T158E/N117R GDH is non-additive. This finding indicates cooperativity between the central charge cluster Arg128-Glu158, symmetry-related Glu158*-Arg128* and the peripheral ion-pair Glu120-Arg117. A direct interaction between these structural elements is not possible, because the C^{α} atom of Arg117 is located 14 and 17 Å from the C^{α} atoms of residues 158 and 128, respectively. Similar distances separate the charged groups of the side-chains of these residues. Therefore the cooperativity is exerted either *via* backbone communication (residues 117, 120 and 128 are all part of the same α-helix) or *via* the electrostatic fields generated by the different mutations. Differences between backbone conformations of the wild-type and the quadruple mutant are either not present or too subtle to be visualized in the 3 Å structures. The significant differences in stability

between GDH mutants containing small and large networks show that large ion-pair networks may be profitable for this enzyme and furthermore indicate that cooperativity of electrostatic interactions may extend over large distances. The increase in the half-life of inactivation due to substitution N117R indicates that ion-pair formation stabilizes the enzyme by increasing the kinetic barrier towards thermal inactivation. These data are in line with the studies on thermal unfolding of *P. furiosus* rubredoxin and *T. maritima* glyceraldehyde-3-phosphate dehydrogenase (Cavagnero *et al.*, 1998; Pappenberger *et al.*, 1997). In a study of the stability of salt bridges at high temperatures (Elcock, 1998) simple model calculations have suggested that a sizeable energy barrier exists for breaking a salt bridge and that the height of this barrier increases with temperature. A similar barrier is not seen with hydrophobic isosteres. The presence of this energy barrier suggests an apparent role of salt bridges in increasing the kinetic barrier towards thermal inactivation or unfolding.

In conclusion, we have succeeded in introducing three 16-residue ion-pair networks into the trimer interface of a bacterial GDH which originally has been hydrophobically optimized. Successive introduction of the amino acids that form the network, does not result in incremental stabilization of the enzyme. Several combinations of individual destabilizing mutations, however, result in restored activity and stability. On the one hand, these findings indicate that exchange of structural features between archaeal and bacterial enzymes that share only moderate amino acid identity (but high structural homology) and have developed different modes of stabilization, is not straightforward. On the other hand, if the right conditions are met stabilization is possible, as shown in the combinatorial mutant S128R/T158E/N117R. The high cooperativity detected between the different members of the network may be the ultimate stabilizing feature and may explain why large ion-pair networks are found in many hyperthermostable enzymes.

Materials and Methods

Construction of mutated *gdh* genes

The expression plasmid containing the wild-type *T. maritima gdh* gene that was used as template for the production of mutant *gdh* genes, was described by Lebbink *et al.* (1998). Single mutations were introduced using two PCR-based methods and appropriate wild-type or mutant template DNA; one method is based on megaprime and casette mutagenesis (Kuipers *et al.*, 1991; Landt *et al.*, 1990), the other on the overlap extension method (Ho *et al.*, 1989). Combinations of mutations were constructed using either one of the aforementioned techniques or by exchange of restriction fragments. Mutated gene fragments were sequenced to verify the sequence of the amplified DNA and to confirm the presence of the desired mutations.

Purification and crystallization of GDH

Wild-type and mutant *T. maritima* GDH were purified as described (Lebbink *et al.*, 1998), except for the decrease in heat-incubation temperature from 80 to 70 °C because of the thermolability of some of the mutants. Because of this an additional chromatography step was needed in order to remove the last impurities. The samples were applied onto an hydroxylapatite column (BioRad, Veenendaal, The Netherlands). GDH was eluted using a linear phosphate gradient (from 20 to 300 mM). Purity of the samples was analyzed by SDS-PAGE. Pure GDH samples were pooled and dialyzed against 20 mM sodium phosphate buffer (pH 7.2). Protein concentrations were determined using an extinction coefficient at 280 nm of $\varepsilon_{zn} = 46{,}490\ M^{-1}\ cm^{-1}$ calculated as described by Gill & von Hippel (1989). Crystals of mutant S128R/T158E/N117R/S160E were grown as described for the wild-type GDH (Knapp *et al.*, 1997) but using slightly higher concentrations of polyethylene glycol (4 to 6%, w/v).

Data collection, analysis and refinement

Data were collected at the synchrotron facility in Hamburg (beamline X11). Images were recorded by the rotation technique using a 300 mm image plate, a wavelength of 0.9096 Å and a detector to crystal distance of 400 mm. The raw data were indexed and evaluated with Denzo and scaled using Scalepack (Otwinowski & Minor, 1997). Intensity measurements of one crystal were used for the final dataset. Since the mutant was identical to the wild-type GDH in cell parameters and spacegroup, its structure was directly refined with the program X-PLOR (Brünger, 1992a) using the phases determined for the wild-type structure (Knapp *et al.*, 1997) and parameters developed by Engh & Huber (1991). Each model was refined (X-PLOR) using the rigid body refinement option as well as conventional conjugate gradient minimization procedures on positional parameters. Temperature factors were initially set to 20 Å^2 and were refined (grouped and restrained) after the *R*-value of the model dropped below 30%. The improvement of the model was monitored with the help of the free *R*-value calculated from 10% of randomly selected reflections which were excluded from the refinement procedure (Brünger, 1992b). After the *R*-value dropped below 25% (*R*-free 29%) the positions of the atoms in the two domains were restrained with an effective energy term of 200 kcal/(mol Å^2). In the final refinement cycles the restraints were weakened (100 kcal/(mol Å^2)). In the calculated electron density ($(2F_O^{Mu} - F_C^{WT})$ exp $i\alpha_C^{WT}$) the side-chains of the introduced residues were clearly visible. The quality of the electron density map was significantly improved by 6-fold averaging of the electron density using the program RAVE (Kleywegt & Jones, 1994) and a mask calculated at 3 Å distance around the initial model with the help of the program MAMA (Kleywegt & Jones, 1993). The rotation/translation matrices were refined with IMP (Kleywegt & Jones, 1994). The new side-chains were built into the density and the mutant structure was subsequently refined as described above. The final model was refined to a similar *R*-value as the wild-type structure (*R*-value 21.0%, R_{free} 27.4%). The geometry of the final model was analyzed using the program PROCHECK (Laskowski *et al.*, 1993).

Characterization of wild-type and mutant GDHs

GDH activity was assayed on a Hitachi U2010 spectrophotometer as described before (Lebbink *et al.*, 1998), except for the change of the final concentration of the potassium phosphate reaction buffer from 100 mM to 20 mM. Thermal inactivation was studied at an enzyme concentration of 50 μg/ml and 85 °C. Residual enzyme activity was assayed at 25 °C. Slopes of first-order inactivation plots yielded rate constants for inactivation from which half-lives were calculated. Differential scanning microcalorimetry was carried out as described before (Lebbink *et al.*, 1998) using 20 mM sodium phosphate (pH 7.4). Temperature profiles with sharp transitions were obtained, that allowed for determination of apparent t_m values with 0.2 °C accuracy. Since the thermal unfolding transition was irreversible, thermodynamic parameters were not determined.

Atomic coordinates

The atomic coordinates have been deposited in the Brookhaven protein database under accession code 2tmg.

Acknowledgments

We thank Kitty Yip, Linda Britton and Timothy Stillman (Krebs Institute, Sheffield, UK) for their contribution to designing the mutants and the staff of the Synchrotron facility at EMBL, Hamburg, for their assistance with data collection. This work has been partially supported by contracts BIOT-CT93-0274 and BIO 4-CT96-0488 of the European Union.

References

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the β-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. (1990). pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry*, **29**, 2403-2408.

Bacon, D. J. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219-220.

Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W. & Stillman, T. J. (1992). Subunit assembly and active site location in the structure of glutamate dehydrogenase. *Proteins: Struct. Funct. Genet.* **12**, 75-86.

Barlow, D. J. & Thornton, J. M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867-885.

Brünger, A. T. (1992a). *X-PLOR (Version 3.1) Manual*, Yale University Press, New Haven and London.

Brünger, A. T. (1992b). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **255**, 472-475.

Cavagnero, S., Zhou, Z. H., Adams, M. W. & Chan, S. I. (1995). Response of rubredoxin from *Pyrococcus furiosus* to environmental changes: implications for the origin of hyperthermostability. *Biochemistry*, **34**, 9865-9873.

Cavagnero, S., Debe, D. A., Zhou, Z. H., Adams, M. W. W. & Chan, S. L. (1998). Kinetic role of electrostatic interactions in the unfolding of hyperthermophilic and mesophilic rubredoxins. *Biochemistry*, **37**, 3369-3376.

Chiaraluce, R., Schwerdtfeger, R. M., Scandurra, R., Antranikian, G. & Consalvi, V. (1997). Acid-induced disassembly of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* occurs below pH 2.0. *Eur. J. Biochem.* **247**, 224-230.

Consalvi, V., Chiaraluce, R., Millevoi, S., Pasquo, A., Vecchini, P., Chiancone, E. & Scandurra, R. (1996). Refolding pathway and association intermediates of glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*. *Eur. J. Biochem.* **239**, 679-685.

Dao-pin, S., Sauer, U., Nicholson, H. & Matthews, B. W. (1991). Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. *Biochemistry*, **30**, 7142-7153.

Eggen, R. I., Geerling, A. C., Waldkotter, K., Antranikian, G. & de Vos, W. M. (1993). The glutamate dehydrogenase-encoding gene of the hyperthermophilic archaeon *Pyrococcus furiosus*: sequence, transcription and analysis of the deduced amino acid sequence. *Gene*, **132**, 143-148.

Elcock, A. H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489-502.

Elcock, A. H. & McCammon, J. A. (1997). Continuum solvation model for studying protein hydration thermodynamics at high temperatures. *J. Phys. Chem.* **101**, 9624-9634.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A*, **47**, 392-400.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.

Hendsch, Z. S. & Tidor, B. (1994). Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* **3**, 211-226.

Hennig, M., Darimont, B., Sterner, R., Kirschner, K. & Jansonius, J. N. (1995). 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure*, **3**, 1295-1306.

Hennig, M., Sterner, R., Kirschner, K. & Jansonius, J. N. (1997). Crystal structure at 2.0 Å resolution of phosphoribosyl anthranilate isomerase from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability. *Biochemistry*, **36**, 6009-6016.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51-59.

Kleywegt, G. J. & Jones, T. A. (1993). Masks made easy. In *ESF/CCP4 Newsletter*, 28, pp. 56-59 .

Kleywegt, G. J. & Jones, T. A. (1994). Halloween, masks and bones. In *From First Map to Final Model*, (CCP4) pp. 59-66.

Klump, H., Di Ruggiero, J., Kessel, M., Park, J. B., Adams, M. W. & Robb, F. T. (1992). Glutamate dehydrogenase from the hyperthermophile *Pyroco-*

ccus furiosus. Thermal denaturation and activation. *J. Biol. Chem.* **267**, 22681-22685.

Knapp, S., de Vos, W. M., Rice, D. & Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3.0 Å resolution. *J. Mol. Biol.* **267**, 916-932.

Körndorfer, I., Steipe, B., Huber, R., Tomschy, A. & Jaenicke, R. (1995). The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J. Mol. Biol.* **246**, 511-521.

Kort, R., Liebl, W., Labedan, B., Forterre, P., Eggen, R. I. & de Vos, W. M. (1997). Glutamate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: molecular characterization and phylogenetic implications. *Extremophiles*, **1**, 52-60.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Kuipers, O. P., Boot, H. J. & de Vos, W. M. (1991). Improved site-directed mutagenesis method using PCR. *Nucl. Acids Res.* **19**, 4558.

Landt, O., Grunert, H. P. & Hahn, U. (1990). A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene*, **96**, 125-128.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

Lebbink, J. H. G., Eggen, R. I. L., Geerling, A. C. M., Consalvi, V., Chiaraluce, R., Scandurra, R. & Devos, W. M. (1995). Exchange of domains of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* and the mesophilic bacterium *Clostridium difficile*: effects on catalysis, thermoactivity and stability. *Protein Eng.* **8**, 1287-1294.

Lebbink, J. H., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. & de Vos, W. M. (1998). Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. I. Introduction of a six-residue ion-pair network in the hinge region. *J. Mol. Biol.* **280**, 287-296.

Merrit, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0 - a program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869-873.

Ogasahara, K., Lapshina, E. A., Sakai, M., Izu, Y., Tsunasawa, S., Kato, I. & Yutani, K. (1998). Electrostatic stabilization in methionine aminopeptidase from hyperthermophile *Pyrococcus furiosus*. *Biochemistry*, **37**, 5939-5946.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676-683.

Rahman, R., Fujiwara, S., Nakamura, H., Takagi, M. & Imanaka, T. (1998). Ion-pairs involved in maintaining a thermostable structure of glutamate dehydrogenase from a hyperthermophilic archaeon. *Biochem. Biophys. Res. Commun.* **248**, 920-926.

Russell, R. J. M., Ferguson, J. M. C., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry*, **36**, 9983-9994.

Spassov, V. Z., Karshikoff, A. D. & Ladenstein, R. (1994). Optimization of the electrostatic interactions in proteins of different functional and folding type. *Protein Sci.* **3**, 1556-1569.

Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S. P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H. & Robb, F. T. (1998). Protein thermostability above 100 C: a key role for ionic interactions. *Proc. Natl Acad. Sci. USA*, **95**, 12300-12305.

Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995). Are buried salt bridges important for protein stability and conformational specificity? *Nature Struct. Biol.* **2**, 122-128.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R. & Rice, D. W. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure*, **3**, 1147-1158.

Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D. & Rice, D. W. (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336-346.

Edited by R. Huber

(Received 12 January 1999; received in revised form 1 April 1999; accepted 5 April 1999)

Chapter 7

The β-glucosidase CelB from *Pyrococcus furiosus*: Production by *Escherichia coli*, purification and *in vitro* evolution.

Joyce H.G. Lebbink, Thijs Kaper, Servé W.M. Kengen, John van der Oost and Willem M. de Vos.

Parts of this chapter will be published in
Methods in Enzymology,
Volume '**Hyperthermophilic enzymes**'.

Introduction

One of the key enzymes of the hyperthermophilic archaeon *Pyrococcus furiosus* involved in growth on β-linked sugars, is the inducible β-glucosidase (CelB). This enzyme is very well suited to function as a model in studying the molecular mechanisms that are employed by hyperthermophilic organisms to optimise enzyme stability and catalysis. In this chapter we describe (i) an overview of the state of the art on this hyperthermostable enzyme, (ii) protocols for improved and scaled-up heterologous production and enzyme purification, and (iii) the development and application of a directed evolution procedure, that resulted in the isolation and characterisation of an active site mutant.

Characterisation and application potential of CelB

The cytoplasmic CelB was purified from *P.furiosus*, showed a molecular mass of 230 kD, and is composed of four identical subunits (Kengen *et al.*, 1993). The enzyme is one of the most thermostable glycosyl hydrolases described to date, and is optimally active at 102-105°C and pH 5.0 (Kengen *et al.*, 1993). It shows high activity on aryl-glucosides and aryl-galactosides, as well as on the β-(1,4)-linked disaccharides cellobiose and lactose and the β-(1,3)-linked disaccharide laminaribiose (Kengen *et al.*, 1993; Bauer *et al.*, 1999). It furthermore has low β-mannosidase and β-xylosidase activity (Kengen *et al.*, 1993). High catalytic efficiency (4200 s^{-1} M^{-1}) is reported for hydrolysis of *p*-nitrophenyl-β-D-glucopyranoside (pNp-glu), with a V_{max} of 700 U/mg and a K_m for this substrate of 0.15 mM at 90°C (Kengen *et al.*, 1993). The enzyme is not dependent on bivalent cations, and thiol-groups are not essential for activity. Finally, the enzyme is competitively inhibited by several ground-state analogues like glucose (K_i = 40 mM), and transition-state analogues like gluconolactone (K_i = 0.080 mM), as well as tris(hydroxymethyl)aminomethane (Tris) (K_i = 1.3 mM) (Bauer & Kelly, 1998).

The *celB* gene was cloned and sequenced, and the deduced amino acid sequence showed high homology with β-glycosidases that belong to glycosyl hydrolase family I (Voorhorst *et al.*, 1995). Overexpression in *Escherichia coli*

Abbreviations

pNp	*p*-nitrophenol
pNp-glu	*p*-nitrophenol-β-D-glucopyranoside
pNp-gal	*p*-nitrophenol-β-D-galactopyranoside
pNp-man	*p*-nitrophenol-β-D-mannopyranoside
X-glu	5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
CelB	β-glucosidase from *Pyrococcus furiosus*
LacG	6-phospho-β-galactosidase from *Lactococcus lactis*
BglA	β-glucosidase from *Bacillus polymyxa*
dNTP	deoxy-nucleotide-tri-phosphate
IPTG	isopropyl-β-D-thiogalactopyranoside
Tris	tris(hydroxymethyl)aminomethane

resulted in high level production of a thermostable β-glucosidase that was purified and found to have similar kinetic and stability properties as CelB purified from *P.furiosus*. This indicated that the *E.coli*-produced enzyme could replace the native enzyme in studying features of hyperthermostability and -activity. Furthermore, the heterologous expression system allowed for the first protein engineering studies of a hyperthermostable enzyme that was realised by changing glutamate 372, which is conserved among family I enzymes. Substitution of this residue by an aspartate or a glutamine resulted in 200-fold and 1000-fold reduction in specific activity, respectively. This confirmed that this conserved glutamate residue is the active site nucleophile involved in catalysis above 100°C (Voorhorst *et al.*, 1995). A comparison of kinetic properties of CelB with those of the β-glucosidase BglA from *Agrobacterium faecalis*, indicated that these homologous enzymes share a common catalytic mechanism (Bauer & Kelly, 1998). Substrate hydrolysis occurs via a double displacement mechanism, involving a covalent glucosyl-enzyme intermediate and results in retention of the configuration at the anomeric carbon atom (Kempton & Withers, 1992; Wang, 1995). A three-dimensional CelB model, based on crystals diffracting to 3.3 Å, shows the complete conservation of the active site architecture compared to other family I enzymes (Chapter 8 of this thesis). Evidently, during evolution of family I enzymes from organisms growing optimally at different temperatures, the overall protein structure has been adapted to the changing environmental conditions while the integrity of the active site and thereby substrate binding and enzyme turnover, has been maintained.

It has been reported that CelB production is highest when *P.furiosus* is grown on β-linked sugars and may reach 5% of total cell protein in this organism (Kengen *et al.*, 1993). This production is controlled at the transcriptional level and gene expression is induced within ten minutes after addition of cellobiose (Voorhorst *et al.*, 1999). Although CelB is highly active on the β-(1,4)-linked disaccharide cellobiose, it is not able to liberate glucose from β-(1,4)-linked polysaccharides like cellulose. It does, however, show low activity on the β-(1,3)-linked polymer laminarin (Kengen *et al.*, 1993). In fact, a strong synergy between CelB and the secreted endo-β-(1,3)-glucanase (LamA) from *P.furiosus* on laminarin has been described (Gueguen *et al.*, 1997). This indicates that these two enzymes cooperate to enable *P.furiosus* to grow efficiently on β(1,3-1,4)-glucans, that may be ubiquitously available in its marine environment as constituents of the cell wall of eukaryotic algae (laminarin) or methanogenic archaea (pseudopeptidoglycan) (Gueguen *et al.*, 1997; König *et al.*, 1983). Recently, a second endoglucanase (EglA) has been characterised, with specificity for β-(1,4)-linkages. LamA and EglA may work in concert to hydrolyse β(1,3-1,4) β-glucan, delivering short oligosaccharides to the intracellular CelB (Bauer *et al.*, 1999).

CelB is the most thermostable member of the family I of glycosyl hydrolases. Half-life values for thermal inactivation of 85 hours and 13 hours at 100 and 110°C were reported, respectively, and an apparent melting temperature determined by differential scanning microcalorimetry of 108°C (Kengen *et al.*, 1993; Bauer & Kelly, 1998). Half-life values for thermal

inactivation were lowered considerably when the incubation buffer was changed from 140 mM Tris (pH 8.5) to 150 mM sodium citrate (pH 5.0). This stabilising effect of Tris became apparent during SDS-PAGE analysis that shows that CelB can only be completely denatured when boiled in non-conventional sample buffer, in which Tris has been replaced with citrate as buffer component. The stabilising effect of Tris was attributed to its structural analogy with several compatible solutes that are present in hyperthermophiles and have been shown to act as thermostabilisers (Kengen *et al.*, 1993; Kengen & Stams, 1994). Alternatively, the observed stabilisation may be related to the fact that Tris is a competitive inhibitor of CelB (Bauer & Kelly, 1998). Interaction between this organic compound and active site-residues may stabilise the enzyme against thermal inactivation. Incubation in high concentrations of denaturants like 4 M urea, 1M guanidine hydrochloride or 8.5 M ethanol and subsequent assaying for activity in the absence of these compounds, revealed no loss of activity, and indicates that CelB is also very resistant against chemical treatments (Kengen & Stams, 1994).

The nearest family I relative of CelB is the β-glycosidase (LacS) from the hyperthermophilic archaeon *Sulfolobus solfataricus* (Pisani *et al.*, 1990). This enzyme has a half-life of inactivation of 3 hours at 85°C and shares with CelB a similar broad substrate specificity (Pisani *et al.*, 1990; Nucci *et al.*, 1993). Stability of LacS towards high temperature, SDS and alkaline pH has been extensively studied (Nucci *et al.*, 1995; d'Auria *et al.*, 1998; d'Auria *et al.*, 1996). The elucidation of the three-dimensional structure of LacS revealed that thermostability is achieved by a relatively high number of large ion-pair networks and solvent-filled, hydrophilic cavities that were proposed to confer resilience to the enzyme, thereby protecting against denaturation upon large conformational fluctuations at high temperature (Aguilar *et al.*, 1997). In spite of their high degree of amino acid homology and three-dimensional structure similarity, the temperatures at which CelB and LacS inactivate for 50% within one hour, is at least 20°C higher for CelB. The molecular basis for this difference in thermostability between CelB and LacS is currently under investigation (Pouwels et *al.*, 1999).

CelB is an enzyme with an interesting potential for industrial or diagnostic applications. A series of transposons carrying the *celB* gene as genetic marker has been developed for gram-negative bacteria in order to study plant-bacterium interactions (Sessitsch *et al.*, 1996). In plants or plant-root ecosystems, endogenous background β-glucosidase and β-galactosidase activity are easily heat-inactivated and the combination of *celB* and the *gusA* markers allows for simultaneous detection of multiple-strain occupancy of plants and individual nodules by Rhizobia. This application has been developed into the CelB Gene Marking Kit that is marketed by FAO/IAEA (Sessitsch *et al.*, 1996). CelB is furthermore a very suitable enzyme to perform transglycosylation and glucoconjugation with retention of configuration at the anomeric carbon atom. Due to its ability to accept a wide variety of aglycones like primary and tertiary alcohols, CelB seems a promising biocatalyst for regio- and stereoselective sugar derivative synthesis (Fischer *et al.*, 1996). In addition, efficient production of β-galacto-oligosaccharides from lactose has been described (Boon *et al.*, 1999). At present, its performance in high-temperature lactose hydrolysis, oligosaccharide

synthesis, and application in biosensors is tested (Kaper and Kengen, personal communication).

Small and large-scale heterologous production of CelB in *Escherichia coli*

High level overexpression of the *celB* gene in *Escherichia coli* MC1061 was initially achieved by cloning a 1.9-kb *Ssp*I-*Sma*I DNA fragment into the *Sma*I site of vector pTTQ19 under control of the vector-located *tac* promoter (Voorhorst *et al.*, 1995). Upon induction with β-D-thiogalactopyranoside (IPTG), CelB levels in *E.coli* MC1061 harbouring this plasmid pLUW510, amounted up to 20% of total cell protein (Voorhorst *et al.*, 1995). However, high selection pressure had to be applied in order to achieve these levels of overexpression (500 μg/ml ampicilline) and even then in medium-scale fermentations, low expression levels, probably due to structural instability of the plasmid, were often observed. For this reason we decided to develop a new and stable expression system, that (i) would preferably confer resistance to an antibiotic that is metabolised *inside* the cells, thereby exerting stronger selection pressure than ampicilline which is detoxified *outside* the cell, (ii) would greatly reduce the amount of antibiotic to be added, and (iii) would result in high-level overexpression and would allow large-scale fermentations without the problems of plasmid instability.

The coding region of the *celB* gene was PCR-amplified using primers BG238 and BG239 which overlap the start- and stop codons of the gene, respectively, and have the following sequence (with the start and stop codons in italics and restriction sites underlined): BG238 = 5'-GCGCG*CCATGG*CAAAGTTCCCAAAAAACTTCATGTTTG; BG239 = 5'-CGCGC*GGATCC*CTACTTTCTTGTAACAAATTTGAGG. BG238 is homologous to the coding strand and introduces a *Nco*I restriction site (*CCATGG*) which overlaps the startcodon. In order to be able to introduce this restriction site, three extra bases were introduced (GCA) that encode for an alanine residue that is now inserted in between the N-terminal methionine and lysine at position 2 in the enzyme. This alanine will be referred to as residue number 1a and its introduction resulted in an enzyme with comparable kinetic parameters and resistance against thermal inactivation as the enzyme isolated from *P.furiosus* (Voorhorst *et al.*, 1995; Chapters 9 and 10 of this thesis). BG239 is homologous to the non-coding strand and introduces a *Bam*HI restriction site immediately downstream of the stop-codon. The introduction of the *Nco*I and *Bam*HI restriction sites allowed for a translational fusion of the *celB* gene to the ϕ10 translation initiation and termination signals on expression plasmid pET9d (Rosenberg *et al.*, 1987). The resulting plasmid pLUW511 carries *celB* under control of the bacteriophage T7 promoter and could be stable maintained in *E.coli* BL21(DE3) (*hsdS gal* (λcI*ts*857 *ind*1Sam7 *nin*5 *lac*UV5-T7 gene *1*)), which contains a T7 RNA polymerase gene in its chromosome under control of the *lac*UV5 promoter (Studier *et al.*, 1986). pLUW511 confers kanamycine resistance to BL21(DE3) and relatively low amounts of this antibiotic (30-50 μg/ml) are required for stable plasmid maintenance. Induction of a log-phase culture ($OD^{600\ nm}$ = 0.5) with IPTG resulted in high level overexpression, with CelB amounting to at least 20% of total cell protein. The specific activity of CelB

produced from this plasmid amounted to 260 U/mg in cell free extract, compared to 220 U/mg for the enzyme produced from pLUW510. Overnight growth without induction resulted in a lower expression level but a higher total CelB yield because of much higher cell densities. This indicates that the *lac*UV5 promoter is not completely blocked by the repressor, and that under these conditions, T7 DNA polymerase gene expression is sufficient for high-level overproduction of CelB. Because of practical considerations and the high cell densities that can be reached during overnight growth without induction, this procedure was employed as standard protocol. The new construct pLUW511 was found to be completely stable during selective and non-selective growth for at least 50 generations, and is, therefore, very well suited for large-scale fermentations. In an 8-liter fermentor a 20 h culture of BL21(DE3) harbouring pLUW511 yielded 140 grams wet weight *E.coli*, from which 720 mg pure CelB was obtained (see below).

Purification of recombinant CelB

Purification of CelB from *E.coli* has been reported as a simple two-step method including denaturation of most *E.coli* proteins by heat-incubation, and further purification using anion exchange chromatography (Voorhorst *et al.*, 1995). A standard procedure included in this method is the incubation of the cell extract with DNase and RNase. However, in spite of this treatment, oligonucleotides were sometimes found to co-elute with the β-glucosidase from the anion exchange column. The comparison of kinetic and stability parameters of mutated variants of CelB with the wild-type enzyme, that is routinely performed in our laboratory, requires identical enzyme preparations which are completely free of oligonucleotides and other contaminants. We therefore adapted the purification procedure, which is described below.

Purification of CelB from 1 liter batch cultures

One litre TY medium (1% trypton, 0.5% yeast extract, 0.5% NaCl) containing 30 μg/ml kanamycine was inoculated with 10 ml overnight culture of *E.coli* BL21(DE3) harbouring expression plasmid pLUW511 and cultured overnight in 2 litre baffled Erlenmeyer flasks at 37°C in a rotary shaker. Cells (typically 6 g wet weight/litre) were collected at 5400 g, resuspended in 20 ml 20 mM sodium citrate (pH 4.8) and disrupted using a French Press at 110 MPa (Aminco, Silver Spring, Maryland, USA). The resulting cell lysate was incubated for 20 minutes at 80°C and cell debris and denatured *E.coli* proteins were removed by centrifugation at 48,000 g for 30 minutes. The use of 20 mM sodium citrate (pH 4.8) instead of 50 mM sodium phosphate pH 7.5 (Voorhorst *et al.*, 1995) or 20 mM Tris-HCl (pH 8.0), resulted in a more efficient removal of *E.coli* proteins in this step. The resulting heat-stable cell free extract was dialysed overnight against 20 mM Tris-HCl (pH 8.0). A 15 ml Q-Sepharose anion exchange column (Pharmacia, Uppsala, Sweden), operated by the HiLoad System with Pump and Gradient Kit P50 (Pharmacia), was equilibrated with the same buffer. The dialysed heat-stable cell-free extract was loaded onto this column and after extensive washing with 20 mM Tris-HCl (pH 8.0), a linear 0.0-1.0 M sodium chloride gradient in the same buffer was applied. CelB eluted at

0.5 M NaCl under these conditions. Active fractions were pooled and concentrated to less than 1 ml total volume using Centricon-30 devices (Amicon, Beverly, USA). The resulting sample was applied onto a 300 ml Superdex 75 gelfiltration column (Pharmacia) equilibrated with 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, in order to remove last traces of *E.coli* contaminating proteins and oligonucleotides. Purity of the fractions was analysed by SDS-PAGE. Pure fractions typically showed the presence of a protein band with the expected molecular weight for the monomer of 58 kDa, as well as a higher molecular weight form presumably corresponding to the tetrameric form, as described before (Kengen *et al.*, 1993; Kengen & Stams, 1994). Absence of oligonucleotides was spectroscopically verified using a Hitachi U-2010 UV-VIS spectrophotometer, by the location of the UV absorption maximum at 280 nm and the absence of a shoulder in the UV spectrum at 260 nm. Pure fractions were pooled and dialysed against an appropriate buffer. The exchange of Tris-HCl to another buffer is required because Tris is a competitive inhibitor of CelB and other glycosyl hydrolases with an inhibition constant around 1 mM (Bauer & Kelly, 1998; Kempton & Withers, 1992). Sodium citrate (pH 4.8) buffer should be avoided because at higher protein concentrations (starting at approximately 1 mg/ml) CelB reversibly precipitates in this buffer. CelB has a pI of 4.40 (Kengen *et al.*, 1993) and apparently the charge distribution on the enzyme surface at pH 4.8 is such that aggregation occurs at high protein concentrations. A preferred buffer is sodium phosphate at pH 6 or 7. Protein concentrations were determined at 280 nm using an extinction coefficient for one subunit of $\varepsilon_m^{280nm} = 1.28*10^5\ M^{-1}cm^{-1}$, calculated according to Gill & von Hippel (1989). Typically 50-80 mg of pure CelB was obtained from a 1 litre batch culture using this protocol.

Purification of CelB from 8-liter fermentations

A fermentor containing 8 litre TY medium with 50 μg/ml kanamycin and 55 mM glucose was inoculated with 50 ml of an overnight culture of BL21(DE3) harbouring pLUW511. A continuous airflow of 18 l/h and a stirring speed of 400 rpm were employed, and to prevent excessive foaming, 600 μl antifoam was added (Antifoam 289, Sigma). The pH was kept constant at 7.0. After growth at 37°C for approximately 20 h, an $OD_{660\ nm}$ of 8 was reached, and cells were collected using a continuous centrifuge (Heraeus Biofuge28RS). Cells (140 g wet weight) were resuspended in 420 ml of a 20 mM sodium citrate buffer (pH 4.8) and broken using a French Press at 110 MPa. The resulting cell lysate was incubated for 60 minutes at 80°C and cell debris and denatured *E.coli* proteins were removed by centrifugation at 48,000 g for 30 minutes. After dialysis against 20 mM Tris-HCl (pH 8.0), the heat-stable cell-free extract was loaded onto a 50-ml Q-sepharose column (Pharmacia) equilibrated with the same buffer. After extensive washing with the equilibration buffer, CelB was eluted using a 0.0-1.0 M NaCl gradient in the same buffer. CelB eluted at 0.45 M NaCl. The large volume of the pooled samples (330 ml containing 810 mg protein) ruled out gel filtration as the subsequent step. The pooled fractions were saturated with 1 M ammonium sulphate and loaded onto a Phenyl Sepharose column (Pharmacia), after which a linear decreasing ammonium sulphate gradient (1.0-0.0 M) in 20 mM Tris (pH 8.0) was applied. CelB eluted at the end of the gradient, separated

from oligonucleotides that do not bind to the column. Last traces of a contaminating *E.coli* protein were removed by a 30 min heat-treatment at 100°C. Final yield amounted up to 60% and 720 mg pure CelB was obtained.

Construction of a random *CelB* library

Directed evolution by random mutagenesis and in vitro recombination is a powerful approach for studying many characteristics of a model system and to evolve a desired new property (Stemmer, 1994a; Stemmer, 1994b; Crameri *et al.*, 1998; Kuchner & Arnold, 1997; Arnold, 1998). This approach has so far been restricted to enzymes from mesophiles and thermophiles and has not yet been applied to hyperthermostable enzymes. CelB is a very suitable candidate to be used as model in the development of such a procedure for hyperthermostable enzymes. The enzyme is efficiently overexpressed in *E.coli* and therefore easily accessible for the introduction of random mutations by genetic techniques. The enzyme is extremely thermostable and resistant to inactivation by chemical denaturants (Kengen & Stams, 1994). Furthermore, because of its broad substrate specificity, it is able to hydrolyse a variety of chromogenic substrate analogues. Therefore, high-throughput screening and selection methods can be applied that are based on the production and analysis of large libraries consisting of random CelB variants.

Description of the model system

As starting material for a random CelB library, we used the expression plasmids pET9d and its derivative pLUW511 in *E.coli* strain JM109(DE3). The genotype of JM109(DE3) is *recA1 supE44endA1 hsdR17 gyrA96 relA1 thi* Δ(*lac-proAB*) F' [*traD36 proAB*$^+$ *lacI*^{q} *lacZ*ΔM15] (λ*clts857 ind1Sam7 nin5 lacUV5-T7* gene *1*) and it is therefore deficient in β-galactosidase (LacZ) activity, in contrast to BL21(DE3) (see above). In combination with vector pET9d, which does not code for the α-peptide of the *E.coli* β-galactosidase, this strain will result in white colony formation on agar plates containing the chromogenic substrates 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (X-glu) or 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), and will not hydrolyse the artificial substrates pNp-glu and pNp-gal. When a gene coding for a functional CelB is cloned into pET9d, colonies will develop a blue colour due to hydrolysis of either the X-glu or X-gal substrates. Alpha-complementation between the truncated LacZ produced by *E.coli* JM109(DE3), and parts of, or the complete CelB enzyme has not been observed and is thought to be highly unlikely; LacZ and CelB belong to different families of glycosyl hydrolases, which do not share amino acid or structural homology. While in family I glycosyl hydrolases the active site is buried at the inside of an $(\alpha\beta)_8$ barrel within each individual monomer, the active site in LacZ involves residues from different subunits (Wiesmann *et al.*, 1997; Aguilar *et al.*, 1997; Jacobson *et al.*, 1994). Even in the absence of an inducer, it is evident that the *lacUV5* promoter in JM109(DE3) is not fully repressed and CelB is produced, as described above for BL21(DE3). Although the activity of the thermostable CelB at 37°C is only 2% of its activity at 90°C, this is sufficiently high to hydrolyse the chromogenic substrates during overnight growth (Figure 1). Blue colour formation in general correlates with the activity of

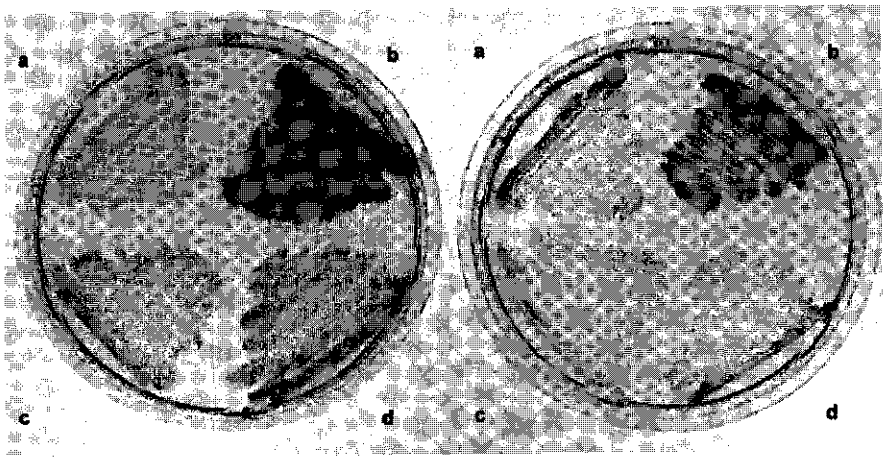


Figure 1: Agar plates containing 16 μg/ml of the chromogenic substrates X-glu (left) and X-gal (right) with negative controls (a), and clones producing wild-type CelB (b), and active site-mutants E417S (c) and (d) N206S.

the CelB variant that is expressed; wild-type CelB-producing colonies develop a darker blue colour than, for example, colonies harbouring plasmid pLUW513, coding for mutant E417S CelB that has a 10-fold reduced specific activity on pNp-glucose (Chapter 8 of this thesis). A major advantage of using the pET9d expression plasmid, is the possibility of constructing translational fusions, using restriction enzyme recognition sites that are overlapping the start and stop codons of the gene of interest. This means that solely the coding region of the *celB* gene is mutagenised, and that the library will not contain random clones with mutations in promoter or terminator sequences or mutations affecting the plasmid replication, stability or copy number. These kind of mutations may result in fluctuating expression levels of the gene and thereby total enzyme activity, and were shown previously to be responsible for part of the increase in activity found for an *in vitro* evolved fucosidase (Zhang et al., 1997).

Random mutagenesis of the celB gene

Random mutagenesis and DNA shuffling of the *celB* gene was essentially performed according to the protocol of Stemmer (Stemmer, 1994a) with the optimisation described by Lorimer and Pastan (Lorimer & Pastan, 1995). Individual steps in the procedure are visualised in Figure 2. For the preparation of template for DNA shuffling, the *celB* gene was PCR-amplified using *Taq* DNA polymerase, thereby exploiting the lack of proofreading activity of this polymerase to introduce random mutations. Primers were developed that just overlap the start and stop codons of the *celB* gene. This allowed mutagenesis of

the complete coding region except for the start and stop codons themselves without contaminating mutations in the flanking regions. The *celB* gene was PCR amplified using pLUW511 as template and primers BG238 and BG309 (5'-GTTAGCAGCCGGATCCCTA; with the *Bam*HI site underlined and the stop codon in italics). While BG238 is the same primer as used for the construction of expression clone pLUW511, BG309 resembles BG239 but allows amplification (and mutagenesis) of all codons up to, but not including, the stop codon. In later procedures BG238 has been replaced by primer BG417 (5'-CTTTAAGAAGGAGATATACCATG), which allows mutagenesis of all codons immediately downstream of the startcodon. The PCR was performed in the following reaction mixture (PCR-MIX): 10 ng pLUW511, 100 ng of the primers BG238 and BG309, 0.2 mM dNTP's, *Taq* DNA polymerase buffer and 5 units *Taq* DNA polymerase (Pharmacia). 30 cycles of 30'' 94°C, 30'' 45°C, 90'' 72°C on a Perkin Elmer GeneAmp 2400 were performed, preceded by 5' denaturation at 94°C and followed by 7' extension at 72°C. PCR products obtained in 24 independent amplification reactions were pooled and purified using the QIAquick PCR purification kit (Qiagen, Westburg, Leusden, The Netherlands). PCR products (200 ng/µl) were incubated with 1 ng/µl DNaseI (Sigma, Zwijndrecht, The Netherlands) in 50 mM Tris-HCl (pH 7.4) containing 10 mM $MnCl_2$ at 20°C. The replacement of magnesium with manganese in this step (i) results in the introduction of double-stranded nicks into the DNA instead of single-stranded nicks, which means that also after denaturation the single-stranded fragments will have the length as visualised before denaturation on the agarose gel, and (ii) lowers the affinity of DNaseI for fragments smaller than 50 bp, which makes the timing of the DNaseI treatment less critical (Lorimer & Pastan, 1995). The reaction was stopped after 10 minutes by addition of 0.1 volume of 0.5 M EDTA pH 8.0. This treatment typically resulted in the formation of 50-200 bp fragments (Figure 2). These fragments were purified using the Qiaquick Nucleotide Removal kit (Qiagen, Westburg) according to the supplied protocol except for the centrifugation at both washing steps that was performed at 6000 rpm. Alternatively, DNA fragments of 50-150 bp were excised from an agarose gel and purified using the QIAEX II Gel Extraction kit. The concentration of DNA fragments after purification was increased by removal of excess water in a vacuum excicator. The resulting gene fragments were assembled in a PCR reaction without primers (Figure 2). Incubation mixtures containing Taq DNA polymerase buffer, 0.2 mM dNTP's, 20 ng/µl *celB* fragments and 5 units Taq DNA polymerase (Pharmacia) were subjected to 40 cycles of 30'' 94°C, 30'' 55°C, 90'' 72°C, preceded by 5' 94°C and followed by 7' 72°C. Correct fragments were enriched in a subsequent PCR reaction including primers BG238 and BG309 as follows; the mixtures were diluted 10-fold in PCR-MIX lacking pLUW511, and subjected to 30 cycles of amplification as described for the first PCR-reaction (see above). Critical in the whole procedure was to avoid contamination with intact *celB* during reassembly steps. A total of 100 cycles of PCR amplification using *Taq* DNA polymerase was performed in this way. The resulting products, a smear of DNA fragments equal to and larger than 1.5 kb, were digested with *Nco*I and *Bam*HI (Figure 2). Fragments of approximately 1.5 kb length were isolated from an agarose gel using the QIAEX II Gel Extraction kit and ligated into *Nco*I-*Bam*HI linearised pET9d using a vector:insert molar ratio of

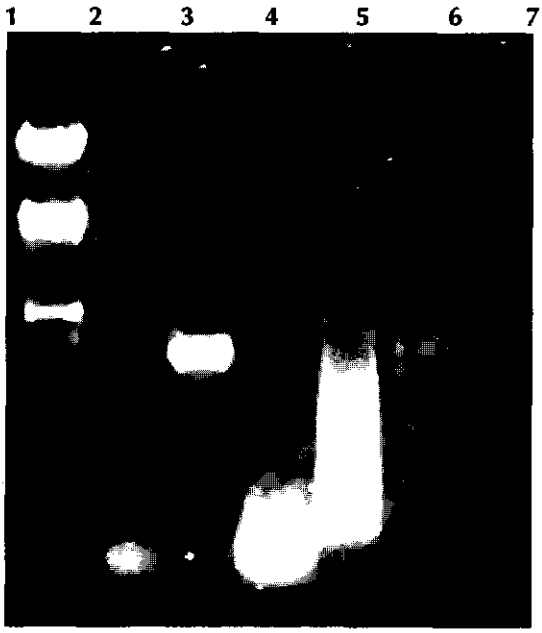


Figure 2: 2% agarose gel showing DNA shuffling of *P.furiosus celB* gene. Lane 1 and 2: Markers, phage λ DNA digested with *Hin*dIII and *Eco*RI and a 36 bp oligonucleotide, respectively. Lane 3: Amplified *celB* gene, Lane 4: DNaseI treated *celB* gene, Lane 5: Reassembled *celB* gene, Lane 6: Enrichment of *celB* gene by conventional PCR on assembly product, lane 7: Reassembled *celB* gene after digestion with restriction enzymes *Nco*I and *Bam*HI and gel purification.

1:3. *E.coli* strain JM109(DE3) was transformed with the ligation mixture and plated onto TY agar plates containing 30 μg/ml kanamycine and 16 μg/ml X-glu (Biosynth, Switzerland). Transformants containing functional *P.furiosus* CelB were identified by blue colour formation. Approximately 5000 transformants per μg of transformed linearised vector DNA were obtained, from which ¼ appeared to be blue and contained a complete *celB* gene, in contrast to the white colonies that contained with high frequency *celB* inserts smaller than 50-100 bp.

Mutation frequency

Nine functional CelB clones were randomly picked, and plasmid DNA was isolated and sequenced in order to determine the mutation frequency. The number of mutations ranged from 0 to 5 and showed the following distribution: no mutations (2), one mutation (1), two mutations (2), three mutations (2), four mutations (1) and five mutations (1). These mutations were randomly distributed along the entire *celB* gene. The average mutation frequency is 2.3 base changes per *celB* gene, which results in one to two amino acid changes per CelB enzyme. This is an acceptable compromise between having a minimum of wild-type sequences in the library and a maximum of amino acid substitutions per enzyme while preventing the occurrence of compensatory mutations starting to interfere with desired phenotypic changes.

Construction and screening of the random celB library

Single blue colonies were picked and transferred to microtiter plates containing 200 µl of TY medium with 30 µg/ml kanamycine and 10% glycerol. Each microtiter plate contained negative and positive controls, namely JM109(DE3) harbouring pET9d, pLUW511 or pLUW513 coding for CelB mutant E417S (Chapter 8 of this thesis). Seventy microtiter plates were prepared (containing approximately 6200 random clones), incubated overnight at 37°C and stored at –80°C. For screening on chromogenic substrates, the library was replica-plated using a 96-pins replicator stamp onto TY agar plates containing 30 µg/ml kanamycine and 16 µg/ml X-glu or X-gal. Plates were incubated overnight at 37°C and intensity of blue colour formation was assayed by visual inspection. This resulted in blue colour formation of similar intensity on both substrates for the wild-type CelB, while clones with empty vector pET9d remained completely white, and the E417S mutant developed only a faint blue colour (Figure 1). Multiple clones were identified that showed unchanged activity on X-glu but largely reduced or completely abolished colour formation on X-gal.

Characterisation of CelB mutant N206S

One of the clones displaying a higher ratio in colour formation on X-glu over X-gal was selected and a single colony was picked (Figure 1). DNA sequencing revealed a single base change of A to G at position 617 of the *celB* coding region. This mutation results in an amino acid substitution of asparagine 206 to serine in the CelB enzyme. Investigation of this amino acid position in a multiple amino acid sequence alignment of family I glycosyl hydrolases, revealed that this asparagine is a conserved residue and is located next to active site glutamate 207, which is proposed to act as the acid/base catalyst during substrate hydrolysis.

The putative role of asparagine 206 has been deduced from an investigation of the 3D-structures from a mutant form of the 6-phospho-β-galactosidase LacG from *Lactococcus lactis* and BglA from *Bacillus polymyxa* (Wiesmann *et al.*, 1997; Sanz-Aparicio *et al.*, 1998). These two family I glycosyl hydrolases were co-crystallised with the substrate 6-phospho-galactose and the inhibitor gluconate in their active site, respectively. Analysis of the three-dimensional structures revealed that the amide side chain of the corresponding asparagine of the CelB Asn206 is in direct contact with the substrate, with a hydrogen bond presumably formed between the side chain nitrogen and the hydroxyl group at the C2 position of the sugar moiety of the substrate (Figure 3). This interaction would be formed in the case of glucose- and galactose-moieties at this position in the substrate, because both sugars have the C2-hydroxyl group in an equatorial position. In the case of mannose, however, this hydroxyl group is axial, and visual inspection of the models suggests that in this case no hydrogen bond can be formed between the C2-OH and the amide side chain of Asn 206. Rather, the mannose-C2 hydroxyl group may be stabilised by an interaction with Glu 207. In mutant N206S the asparagine side chain is shortened to a serine, which is still a polar side chain but because of the increased distance, a favourable interaction between the C2-hydroxyl of the substrate and this side chain is considered highly unlikely.

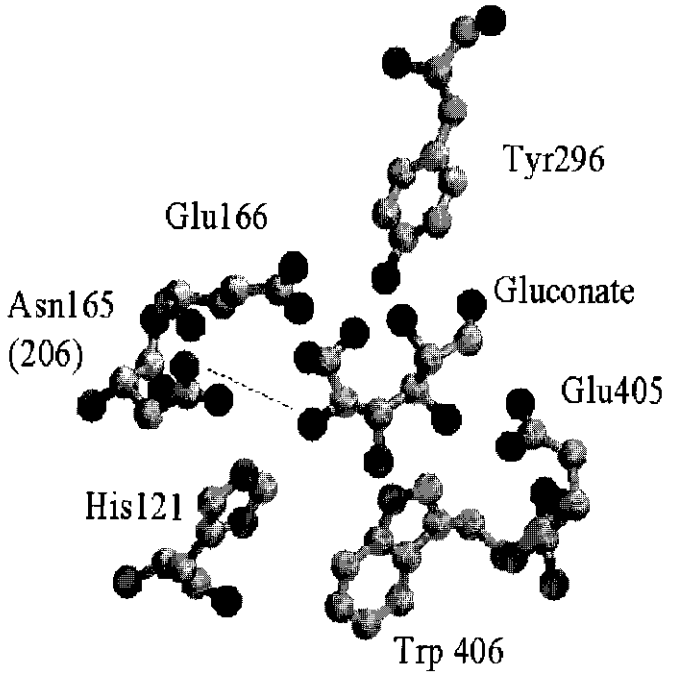


Figure 3: Model of the active site of *B.polymyxa* β-glucosidase complexed with the inhibitor gluconate. Residue numbering is according to the *B.polymyxa* enzyme with CelB numbering in between brackets. The amide group of the asparagine 206 is hydrogen bonding the hydroxyl group at the C2 position of substrates containing glucose and galactose, but not mannose. In the random CelB mutant N206S, this hydrogen bond cannot be formed. The affinity for glucose and galactose substrates in this mutant is much reduced compared to wild-type CelB, while the affinity for mannose is virtually unchanged. The figure has been prepared using the program RASMOL.

In order to investigate the validity of the analysis described above, we purified wild-type CelB and mutant N206S and determined kinetic parameters on the artificial substrates pNp-glu, pNp-gal and pNp-mannose. Wild-type and mutant CelB were purified from 1 litre cultures according to the protocol described above. Kinetic parameters at 90°C were determined on a Hitachi U-2010 spectrophotometer equipped with SPR-10 temperature controller (Hitachi, Tokyo, Japan). A detailed description of the CelB activity assay is described elsewhere (Kaper *et al.*, in preparation). Initial free pNp-formation was continuously followed at 405 nm in 150 mM sodium citrate (pH 4.8; set at room temperature) containing increasing concentrations of the aryl-glucosides pNp-glucose (0.05-40 mM), pNp-galactose (0.2-40 mM) or pNp-mannose (0.4-26 mM). An extinction coefficient of 0.561 M^{-1} cm^{-1} was used for pNp in the assay buffer. V_{max} and K_m values were derived from fitting the data according to the Michaelis-Menten equation corrected, if appropriate, for substrate inhibition, using the non-linear regression program Tablecurve (Tablecurve 2D for Windows, version 2.03 Jandel Scientific). Data are listed in Table I. A more detailed description of the kinetic data for the wild-type enzyme is described elsewhere (Chapter 9 of this thesis). On all three substrates, the maximum

Table 1: Kinetic parameters for wild-type and mutant N206S CelB determined at 90°C on aryl-glucosides pNp-glucose, pNp-galactose and pNp-mannose. *Maximum substrate concentrations employed for pNp-gal are 40 mM, and the calculated K_m value solely serves to indicate that N206S no longer has noteworthy affinity for this substrate.

	pNp-Glu		pNp-Gal		pNp-Man	
	V_{max} (U/mg)	K_m (mM)	V_{max} (U/mg)	K_m (mM)	V_{max} (U/mg)	K_m (mM)
wild-type CelB	1800	0.42	2600	5.0	78	3.5
N206S	340	12	1200*	200*	17	4.7

activity of the mutant dropped considerably compared to wild-type CelB, suggesting that Asn206 is not only involved in substrate binding, but also plays a role in catalysis. Indeed, it has been described for several β-glycosidases from mesophilic organisms that interactions at the C2 position of the substrate strongly stabilise the transition states both for the glycosylation and deglycosylation of the enzyme (Namchuk & Withers, 1995; McCarter *et al.*, 1992; Mega & Matsushima, 1983). However, its role in substrate binding appears to be more important since the raise in K_m for pNp-glu and pNp-gal is much more dramatic than its effect on V_{max} values. Because the maximum pNp-galactose concentration used (40 mM) was far below the calculated K_m value for this substrate for mutant N206S, the calculated kinetic parameters for this mutant/substrate combination only serve to illustrate that affinity for this substrate has been lost in the mutant. This is in agreement with the observation that *E.coli* JM109(DE3) producing mutant N206S remained almost white after overnight growth on agar plates containing X-gal (Figure 1). The effects on K_m values can be directly correlated to the removal of the hydrogen bond acceptor in the mutant N206S. This is supported by the fact that the K_m for pNp-mannose in the mutant is virtually unchanged with respect to the wild-type CelB, which indicates that the asparagine is not a ligand for the axial C2-OH of this substrate.

Conclusions

CelB from *P.furiosus* serves as a very suitable model glycosyl hydrolase to study substrate specificity as well as adaptations of stability and activity to extreme temperatures. Stable production in *E.coli* and an efficient purification protocol allow for simple and rapid preparation of pure wild-type and mutant CelB enzymes. These routine procedures have been applied in studying substrate recognition, thermostability and low- and high temperature dependent activity (Chapter 8 and 9 of this thesis). The *in vitro* evolution procedure that has been developed, allows for easy screening of the existing mutant CelB library for altered characteristics and enables further rounds of directed evolution. The described results on CelB mutant N206S demonstrates (i) the successful construction of the random CelB library, (ii) the feasibility to select mutants with

changed catalytic properties and/or substrate specificity, and (iii) the applicability of LacG and BglA crystal structures in developing models for interactions between CelB and its substrates.

Acknowledgements

This work has been partially supported by contracts BIOT-CT93-0274, BIO 4-CT96-0488, FAIR CT96-1048 and BRPR-CT97-0484 of the European Union. The authors are very grateful to Peter Bron and Melike Balk for their contribution to the work described in this paper.

References

Aguilar, C.F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the beta-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Arnold, F. H. (1998). Design by directed evolution. *Acc. Chem. Res.* **31**, 125-131.

Bauer, M. W., Driskill, L. E., Callen, W., Snead, M. A., Mathur, E. J. & Kelly, R. M. (1999). An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyses beta-1,4 bonds in mixed-linkage (1−>3),(1−>4)-beta-D-glucans and cellulose. *J. Bacteriol.* **181**, 284-90.

Bauer, M. W. & Kelly, R. M. (1998). The family 1 beta-glucosidases from *Pyrococcus furiosus* and *Agrobacterium faecalis* share a common catalytic mechanism. *Biochemistry* **37**, 17170-8.

Boon, M. A., Van der Oost, J., De Vos, W.M., Janssen, A.E.M., Van `t Riet, K. (1998). Synthesis of oligosaccharides catalyzed by thermostable β-glucosidase from *Pyrococcus furiosus*. *Appl. Biochem. Biotechnol.* **75**, 269-278.

Crameri, A., Raillard, S. A., Bermudez, E. & Stemmer, W. P. C. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288-291.

Fischer, L., Bromann, R., Kengen, S. W. M., de Vos, W. M. & Wagner, F. (1996). Catalytical potency of beta-glucosidase from the extremophile *Pyrococcus furiosus* in glucoconjugate synthesis. *Biotechnology* **14**, 88-91.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-26.

Gueguen, Y., Voorhorst, W. G. B., van der Oost, J. & de Vos, W. M. (1997). Molecular and biochemical characterisation of an endo-beta-1,3-glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **272**, 31258-31264.

Jacobson, R. H., Zhang, X. J., DuBose, R. F. & Matthews, B. W. (1994). Three-dimensional structure of beta-galactosidase from *E. coli*. *Nature* **369**, 761-6.

Kempton, J. B. & Withers, S. G. (1992). Mechanism of *Agrobacterium* beta-glucosidase: kinetic studies. *Biochemistry* **31**, 9961-9.

Kengen, S. W., Luesink, E. J., Stams, A. J. & Zehnder, A. J. (1993). Purification and characterisation of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**, 305-12.

Kengen, S. W. M., Stams, A.J.M. (1994). An extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*; a comparison with other glycosidases. *Biocatalysis* **11**, 79-88.

Konig, H., Kandler, M., Jensen, M., Rietschel, E.T. (1983). *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 627-636.

Kuchner, O. & Arnold, F. H. (1997). Directed evolution of enzyme catalysts. *Trends Biotechnol.* **15**, 523-530.

Lorimer, I. A. & Pastan, I. (1995). Random recombination of antibody single chain Fv sequences after fragmentation with DNaseI in the presence of Mn^{2+}. *Nucl. Acids Res.* **23**, 3067-8.

McCarter, J. D., Adam, M.J., Withers, S.G. (1992). Binding energy and catalysis. Fluorinated and deoxygenated glycosides as mechanistic probes of *Escherichia coli* (lacZ) beta-galactosidase. *Biochem. J.* **286**(721-7).

Mega, T., Matsushima, Y. (1983). Energy of binding of *Aspergillus oryzae* beta-glucosidase with the substrate, and the mechanism of its enzymic action. *J. Biochem.* **94**, 1637-47.

Namchuk, M. N., Withers, S.G. (1995). Mechanism of *Agrobacterium* beta-glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry* **34**, 16194-202.

Nucci, R., Moracci, M., Vaccaro, C., Vespa, N. & Rossi, M. (1993). Exo-glucosidase activity and substrate specificity of the beta-glycosidase isolated from the extreme thermophile *Sulfolobus solfataricus*. *Biotechnol. Appl. Biochem.* **17**, 239-50.

Nucci, R., S, D. A., Febbraio, F., Vaccaro, C., Morana, A., De Rosa, M. & Rossi, M. (1995). A thermostable beta-glycosidase from *Sulfolobus solfataricus*: Temperature and SDS effects on its functional and structural properties. *Biotech. Appl. Biochem.* **21**, 265-274.

Pisani, F. M., Rella, R., Raia, C. A., Rozzo, C., Nucci, R., Gambacorta, A., De Rosa, M. & Rossi, M. (1990). Thermostable beta-galactosidase from the archaebacterium *Sulfolobus solfataricus*. Purification and properties. *Eur. J. Biochem.* **187**, 321-8.

Pouwels, J., Moracci, M., Cobucci-Ponzano, B., Perugino, G., van der Oost, J., Kaper, T., Lebbink, J., de Vos, W.M., Ciaramella, M. and Rossi, M. (1999). Activity and stability of hyperthermophilic enzymes: a comparative study on two archaeal β-glycosidases. *Submitted for publication.*

Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J. & Studier, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125-35.

d'Auria, S., Moracci, M., Febbraio, F., Tanfani, F., Nucci, R. & Rossi, M. (1998). Structure-function studies on beta-glycosidase from *Sulfolobus solfataricus*. Molecular bases of thermostability. *Biochimie* **80**, 949-957.

d'Auria, S., Morana, A., Febbraio, F., Vaccaro, C., De Rosa, M. & Nucci, R. (1996). Functional and structural properties of the homogeneous beta-glycosidase from the extreme thermoacidophilic archaeon *Sulfolobus solfataricus* expressed in *Saccharomyces cerevisiae*. *Prot. Expr. Purif.* 7, 299-308.

SanzAparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). Crystal structure of beta-glucosidase A from *Bacillus polymyxa*: Insights into the catalytic activity in family 1 glycosyl hydrolases. *J. Mol. Biol.* **275**, 491-502.

Sessitsch, A., Wilson, K. J., Akkermans, A. D. & de Vos, W. M. (1996). Simultaneous detection of different Rhizobium strains marked with either the *Escherichia coli gusA* gene or the *Pyrococcus furiosus celB* gene. *Appl. Environ. Microbiol.* **62**, 4191-4.

Stemmer, W. P. (1994a). DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10747-51.

Stemmer, W. P. (1994b). Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-91.

Studier, F.W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113-30.

Voorhorst, W. G., Eggen, R. I., Luesink, E. J. & de Vos, W. M. (1995). Characterization of the *celB* gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *J. Bacteriol.* **177**, 7105-11.

Voorhorst, W. G. B., Gueguen, Y., Geerling, A.C.M., Schut, G., Dahlke, I., Thomm, M., Van der Oost, J., de Vos, W.M. (1999). Transcriptional regulation in the hyperthermophilic archaeon *Pyrococcus furiosus*: Coordinated expression of divergently oriented genes in response to β-linked glucose polymers. *J. Bacteriol.* **181**, 3777-3783.

Wang, Q., Trimbur, D., Graham, R., Warren, R.A., Withers, S.G. (1995). Identification of the acid/base catalyst in *Agrobacterium faecalis* beta-glucosidase by kinetic analysis of mutants. *Biochemistry* **34**, 14554-62.

Wiesmann, C., Hengstenberg, W. & Schulz, G. E. (1997). Crystal structures and mechanism of 6-phospho-beta-galactosidase from *Lactococcus lactis*. *J.Mol. Biol.* **269**, 851-860.

Zhang, J. H., Dawes, G. & Stemmer, W. P. C. (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4504-4509.

Chapter 8

Engineering 6-phospho-β-galactosidase activity into the hyperthermostable *Pyrococcus furiosus* β-glucosidase CelB.

Thijs Kaper, Joyce H.G. Lebbink, Jeroen Pouwels, Jürgen Kopp, Georg E. Schulz, John van der Oost, Willem M. de Vos.

Manuscript in preparation.

Abstract

The hyperthermostable *Pyrococcus furiosus* β-glucosidase, belonging to family 1 of glycosyl hydrolases, has been purified from an overproducing *Escherichia coli* strain, crystallised and a structural model was established at 3.3 Å resolution using the molecular replacement method. The pyrococcal enzyme is a homotetramer, with subunits that form a common $(\beta\alpha)_8$-barrel, as previously found in other family 1 glycosidases. The active site of the *P. furiosus* β-glucosidase appeared very similar to that of *Lactococcus lactis* 6-phospho-β-galactosidase, although their substrate specificities differ considerably. Three substitutions were designed to accommodate the *P. furiosus* β-glucosidase for the hydrolysis of 6-phospho-glycosides. The mutations E417S, M424K and F426Y were introduced by site-directed mutagenesis and the three single mutants, the double mutant M424K/F426Y, the triple mutant and the wild-type β-glucosidase, designated CelB, were overproduced in *E. coli*, purified and characterised by determining the pH optima and kinetic parameters. The k_{cat}/K_m ratios of single mutant E417S and the triple mutant for the hydrolysis of *ortho*-nitrophenol-β-D-galactopyranoside were 4.9 and 2.7 times higher as compared to the wild-type CelB. Both mutants had 30 to 300-fold lowered k_{cat}/K_m ratios for the hydrolysis of non- phosphorylated sugars compared to wild-type CelB. These results support the structural model of CelB and indicate that 6-phospho-galactosidase activity can be engineered in a thermoactive β-glucosidase.

Introduction

Glycosyl hydrolases are widely spread among living organisms. Members of this enzyme superfamily are involved in a broad range of physiological processes, such as food supply and storage, cell wall synthesis and degradation, defence systems and signalling events (Davies & Henrissat, 1995). A large number of glycosyl hydrolases has been isolated, characterised and the corresponding amino acid sequences have been determined. Comparison of the amino acid sequences has led to the identification of over 70 families of glycosyl hydrolases (Henrissat, 1999). Family 1 consists of β-glycosidases, 6-phospho-β-glycosidases, lactase-phlorizin hydrolases and myrosinases originating from all three domains of life: Bacteria, Archaea and Eucarya. These enzymes hydrolyse their substrate with retention of configuration at the anomeric carbon atom, which occurs via a double displacement mechanism, in which a covalent intermediate is believed to be formed (Kempton & Withers, 1992; Sinnott, 1990; Wang *et al.*, 1995). Two glutamate residues serve as a general acid/base and a nucleophile in this reaction. Although family 1 glycohydrolases in general display a broad substrate specificity, the glycosidases have been optimised for the hydrolysis of non-phosphorylated glycosides and have hardly any activity on phosphorylated sugars, while the opposite is true for the family 1 6-P-glycosidases.

In the last years, the 3D structures have been solved of family 1 glycosyl hydrolases from Eucarya (Barrett *et al.*, 1995; Burmeister *et al.*, 1997), Bacteria

(Wiesmann *et al.*, 1995; Sanz-Aparicio *et al.*, 1998) and Archaea (Aguilar *et al.*, 1997). The enzymes display a common TIM barrel motif with the two catalytic glutamates located at the C-terminal end of β-strand 4 and 7. The involvement of the two glutamates in catalysis has been widely shown by mutagenesis and inhibitor studies (Street *et al.*, 1992; Trimbur *et al.*, 1992; Voorhorst *et al.*, 1995; Wiesmann *et al.*, 1995; Moracci *et al.*, 1996; Febbraio *et al.*, 1997). Recently, the 3D structure of an inactive mutant of the *Lactococcus lactis* 6-P-galactosidase LacG has been elucidated with a galactose-6-P bound in the active site (Wiesmann *et al.*, 1997).

Previously, a β-glucosidase has been purified from the hyperthermophilic archaeon *Pyrococcus furiosus* (Kengen et al., 1993). The enzyme was active as a homotetramer with 58 kD subunits and displayed a high thermostability, with a half life of thermal inactivation of 85 h at 100°C. The corresponding *celB* gene has been cloned, functionally expressed in *E. coli* and its characterisation designated the β-glucosidase to family 1 of glycosyl hydrolases, in which E372 was identified as the nucleophilic catalytic residue by site directed mutagenesis (Voorhorst *et al.*, 1995). A translational fusion of the *celB* gene to the T7 promoter of the pET9d vector, in which an alanine residue was inserted after the N-terminal methionine, allowed for stable overproduction in *E. coli* of the *P. furiosus* β-glucosidase (Chapter 7 of this thesis). The activity, stability and kinetic parameters of the resulting enzyme, referred to as CelB, were comparable to those of the β-glucosidase isolated from *P. furiosus* (Chapters 7, 9 & 10 of this thesis). Operating at 100°C, CelB serves as a model system to study the substrate specificity of family 1 glycosyl hydrolases, since it has been well characterised, can be produced in *E. coli* in wild-type and mutant forms, and shows extreme stability, broad specificity and high activity.

Here, we report the investigation and mutagenesis of the active site of CelB, based on a structural model of the *P. furiosus* β-glucosidase, that was built using medium resolution X-ray diffraction data and elucidated at 3.3 Å resolution by molecular replacement, using the structures of *L. lactis* 6-P-β-galactosidase LacG (Wiesmann *et al.*, 1995) and *Sulfolobus solfataricus* β-glycosidase LacS (Aguilar *et al.*, 1997) as search models. Based on the structural model we designed mutations to adapt the enzyme to the hydrolysis of phosphorylated galactose. Five mutant CelB enzymes have been purified, characterised and compared to the wild-type enzyme. The results confirmed the model of CelB and indicate that the single substitution E417S caused a significant increase of the catalytic activity on phosphorylated galactose.

Materials and methods

Strains, vectors and chemicals

Escherichia coli BL21(DE3) (*hsdS gal* (*λcIts* 857 *ind*1 *Sam*7 *nin*5 *lac*UV5-T7 gene 1) was used for heterologous expression, *E. coli* TG1 (*supE hsdΔ*5 *thi* Δ(*lac-proAB*) F'[*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]) was used in the construction of derivatives of pET9d (T7 promoter, kan^R) (Novagen). All chemicals used were of analytical grade. Ortho-nitrophenyl-β-D-glucose (oNp-glu), ortho-nitrophenyl-β-

D-galactose (oNp-gal) and ortho-nitrophenyl-β-D-galactose-6-phosphate (oNp-gal-6P) were purchased from Sigma (Zwijndrecht, The Netherlands). *Pfu* DNA polymerase was obtained from Stratagene.

Construction of mutants

Previously, the *celB* gene (GenBank accession AF013169) has been translationally fused to the bacteriophage T7 ϕ10 protein promoter located on the pET9d vector, resulting in pLUW511 (Chapter 7). In combination with an *E. coli* strain that carries the bacteriophage T7 polymerase integrated on the chromosome, CelB could be produced up to 259 U/mg in total cell extracts (Chapter 7). Mutations were introduced in *celB* using *Pfu* polymerase in the PCR-based overlap extension method (Higuchi *et al.*, 1988). For each mutation a sense/antisense primer pair was designed. BG266 (5'-AACAGACAACTACTCCTG -3') / BG267 (5'-GGGCCCAGGAGTAGTTGTCTGTT-3') introduced mutation E417S (introduced mutations in bold face, introduced restrictions sites underlined and restriction enzymes bracketed). BG268 (5'-CAAGGGTTCAGGAA AAGATACGGATTGGTTT-3') / BG269 (5'-AAACCAATCCGTATCTTTCCTGAAC CCTTG-3') introduced M424K/F426Y. BG333 (5'-GGGTTCAGGAAAAGATTCG GATTGGTT-3') / BG334 (5'-AACCAATCCGAATCTTTCCTGAACCC-3') (*Xmn*I) introduced M424K. F426Y was introduced by BG335 (5'-GGTTCAGGATGAGG TACGGATTGGTT-3') / BG336 (5'-AACCAATCCGTACCTCATCCTGAACC-3') (*Rsa*I). In each case the sense primer BG238 (5'-GCGCGCCATGGCAAAGTTCCC AAAAAACTTCATGTTTG-3', *celB* start codon italic face (*Nco*I)) and antisense primer BG239 (5'-CGCGCGGATCCCTACTTTCTTGTAACAAATTTGAGG-3', *celB* stop codon italic face (*Bam*HI)) were used as flanking primers for the amplification of the *celB* gene. The resulting PCR products were digested with *Nco*I and *Bam*HI and ligated in *Nco*I/*Bam*HI-digested pET9d vector. In Table I the resulting constructs have been listed. Plasmid pLUW511 was used as a template in PCR reactions for the construction of pLUW513, 515 and 522. Plasmid pLUW513 was used as a template in the construction of pLUW519.

Table I. List of pET9d constructs with *celB* derivatives

Construct	Gene product
pLUW511	CelB wild-type
pLUW513	CelB E417S
pLUW515	CelB M424K
pLUW518	CelB M424K/ F426Y[1]
pLUW519	CelB E417S/ M424K/F426Y[2]
pLUW522	CelB F426Y

[1] referred to as CelB double mutant

[2] referred to as CelB triple mutant

Plasmid pLUW518 was constructed by replacing a pLUW519 *Pst*I-*Afl*II fragment with a pLUW511 *Pst*I-*Afl*II fragment. Mutations were verified by DNA sequence analysis using the Thermosequenase cycle sequencing kit with infrared dye labelled primers. Reactions were analysed on a Licor 4000L automated sequencer (Licor, Lincoln, USA) (data not shown).

Enzyme production and purification

CelB was routinely produced in *E. coli* BL21 (DE3) harbouring pET9d derivatives (Chapter 7). From a -80°C glycerol stock *E. coli* BL21(DE3) cells harbouring pLUW511 or a pLUW511 derivative were streaked out on a TYA plate (1% trypton, 0.5% yeast extract, 0.5% NaCl, 1.5% granulated agar) with 50 μg/ml kanamycine and cultered overnight at 37°C. One colony was used to inoculate a 10 ml TYK (1% trypton, 0.5% yeast extract, 0.5% NaCl, 30 μg/ml kanamycine) preculture. After 8 h growth in a 37°C shaker this culture was used to inoculate 1 litre TYK in a 2 litre baffled Erlenmeyer. After overnight growth at 37°C while shaking at 110 rpm the cells were harvested by centrifugation for 10 min 5400 g at 4°C. Yields were between 4-8 g wet cells per litre. The cells were resuspended in 25 ml 20 mM Na-citrate (pH 4.8) and the cells were lysed by passage through a French press (110 Mpa) three times. The resulting cell free extract was incubated at 80°C for 45 min. The denatured *E. coli* proteins were precipitated by centrifugation in an ultracentrifuge (Centrikon) for 30 min 20000 rpm at 20°C. The heat incubated soluble fraction was dialysed against 20 mM Tris-HCl (pH 8.0) (buffer A) and applied on a Q-sepharose anion exchange column (Amersham-Pharmacia, Uppsala, Sweden) equilibrated with buffer A. Protein was eluted by a linear increasing NaCl gradient to 1.0 M. A dominant peak with β-glucosidase activity eluted around 0.5 M NaCl. The active fractions were pooled and concentrated in Centricon 30 concentrators (Amicon, Inc, Beverly, MA, USA) followed by separation over a Hiload 26/60 Superdex 75 prepgrade column using a Hiload FPLC (Amersham-Pharmacia, Uppsala, Sweden) and buffer A with 0.10 M NaCl as elution buffer. Fractions with activity were checked for DNA contamination using a spectrophotometer (Beckman, Fullerton, CA, USA). Fractions without DNA were pooled, dialysed against buffer A and concentrated in a Filtron 30 concentrator. The resulting solution contained pure CelB, as was confirmed by SDS-PAGE electrophoresis, and was stored at 4°C with 0.02% Na-azide. Protein concentrations were determined spectrophotometrically at 280 nm with a molar extinction coefficient ε_{280} of 128280 $M^{-1}cm^{-1}$, calculated according to Gill & Von Hippel (1989).

Enzyme preparation, crystallisation and data collection

P. furiosus ß-glucosidase was purified from *E. coli* MC1061 cells harbouring pLUW510 as described previously (Voorhorst *et al.*, 1995) except for the heat denaturation of the cell free extract, which, in this case, was 15 min at 60°C. Initial crystallisation conditions were determined after screening according to Jancarik & Kim (1991). After fine-tuning, the enzyme was found to crystallise optimally by vapour diffusion using the hanging drop method at 20°C. Drops of 5 or 10 μl contained 10 mg/ml protein, 8% PEG-400, 150 mM $CaCl_2$ and 10 mM Tris pH 8.0 and the reservoir contained 20% PEG-400. Within a few days crystals grew to an average size of 600 x 600 x 600 μm^3. The crystals diffracted

to a maximum resolution somewhat less than 3 Å on a rotating X-ray generator. A native data set at 3.3 Å was collected at RT with a multiwire area detector (X1000, Nicolet/Siemens) on a rotating anode with exposure times up to 240 sec pro frame. The structures of 6-phospho-β-galactosidase LacG of *L. lactis* (Wiesmann *et al.*, 1995) and β-glycosidase LacS of *S. solfataricus* (Aguilar *et al.*, 1997) were used as search models for molecular replacement and subsequent construction of the structural model of *P. furiosus* ß-glucosidase.

Kinetic analyses

pH optima. The pH optima of wild-type CelB and CelB mutants were determined at 90°C in combined citrate/phosphate buffer (80mM/80mM) in a pH range of 3.5 - 8.0 (set at 25°C). As substrates oNp-glu, oNp-gal and oNp-gal-6-P were used. In the case of oNp-gal-6P, reaction vials (1.5 ml) containing 49 µl 40 mM oNp-gal-6-P in buffer were preincubated in a water bath at 90°C for 2 min. The reaction was started by the addition of 1 µl enzyme solution to the vials. After 5 min the reaction was stopped by the addition of 100 µl ice-cold 0.5 M Na_2CO_3, which terminated the reaction and augmented the molar extinction coefficient of the released nitrophenol. 125 µl samples were transferred to a 96-wells microtiter plate and the extinction of the reaction mixtures was determined at 405 nm in a microtiter plate reader (SLT Lab Instruments 340 ATTC). Maximum extinction was set at 100%. When oNp-glu and oNp-gal were used as substrates, reaction vials containing 195 µl 10 mM substrate in buffer were treated as for oNp-gal-6-P, only the reactions were initiated by the addition of 5 µl enzyme, stopped with 0.4 ml carbonate and 150 µl samples were transferred to a microtiter plate for analysis.

K_m and V_{max} determinations. The K_m and V_{max} of wild-type CelB and CelB mutants were measured for the hydrolysis of the three substrates at 90°C in 100 mM citrate/phosphate buffer. We followed the release of nitrophenol at 405 nm (pH5.0) or 470 nm (pH6.0) in quartz cuvettes using a spectrophotometer equipped with a temperature controller (Hitachi, San Jose, CA, USA). For each substrate the activities were measured at different concentrations (oNp-glu: 0.05-50 mM; oNp-gal: 0.10-50 mM; oNp-gal-6-P: 0.4-24 mM). Molar extinction coefficients of ortho-nitrophenol at 90°C in assay mixtures were determined ($\varepsilon_{405;pH5.0}$: 0.6243 $cm^{-1}mM^{-1}$; $\varepsilon_{470;pH6.0}$: 0.6339 $cm^{-1}mM^{-1}$, $\varepsilon_{405,pH4.5}$: 0,5614 $cm^{-1}mM^{-1}$, $\varepsilon_{405,pH4.0}$: 0,5398 $cm^{-1}mM^{-1}$). K_m and V_{max}-values were calculated by fitting the activities according to Michaelis-Menten kinetics using the non-linear regression program TableCurve (Jandel Scientific, AISN Software)

Results

CelB structure

The crystals obtained with the *P. furiosus* β-glucosidase allowed data collection to 3 Å resolution, which is too low for a very detailed structural analysis. The tetrameric structure of the enzyme was established in the molecular replacement procedure (Fig. 1). The search models in this procedure were the 3D structures of *L. lactis* 6-phospho-β-galactosidase LacG and the *S.*

solfataricus β-glycosidase LacS, which show 16% and 53% identical amino acid residues to the pyrococcal enzyme, respectively. Upon analysis of the crystal packing, residue K132 was found to participate in the largest crystal contact interaction and was subjected to mutagenesis in order to obtain better crystals. For this purpose, the *P. furiosus celB* gene from the pET9d construct, coding for CelB, with an alanine inserted after the N-terminal methionine was mutagenised, expressed, and CelB mutants K132G, K132L, K132I, K132Q, K132Y, K132H, H129D and H129D/K132Y were purified and crystallised. Crystals were obtained, but they failed to diffract to higher resolution than those with the *E. coli* produced β-glucosidase (data not shown). Therefore, the initial structural model was used in the analysis of the CelB tertiary structure, the active site and for the design of the proposed substitutions.

The CelB homotetramer has point symmetry 222, with one subunit in the crystallographically asymmetric unit. The tetramer consists of two barrels, placed next to each other and two dimers placed on top of each other with the barrel openings facing each other. The overall size of the tetramer is 110 Å x 90 Å x 55 Å and it appearance is that of a slightly twisted square. The axes of the $(\beta\alpha)_8$-barrels are almost parallel to one of the tetramer axes (Fig. 1). Each subunit of CelB consists of a single domain of 472 amino acids, with 18 helices and 16 β-strands. The centre of the monomer is formed by a $(\beta\alpha)_8$-barrel, with the catalytic residues E207 and E372 at the C-terminal end of the 4th and 7th β-strand of the barrel, respectively.

Figure 1. Ribbon model of *P. furiosus* β-glucosidase, viewed along one of the two-fold axis of the tetramer.

Architecture of the CelB active site

The positions of the active site residues in LacG and CelB agree with each other, as can be seen in a superposition (Fig. 2). E207 and E372 in CelB are the equivalents of the catalytic glutamate residues in LacG and N17, R77, H150, N206, Y307 and W410 are conserved residues and likely to be involved in substrate binding. However, the residues that interact with the phosphate group of a galactose-6-P molecule bound in the active site of LacG, serine, lysine and tyrosine, differ at corresponding positions in CelB (E417, M424 and F426) (Fig 3). These residues have been conserved in all known β-glycosidases that hydrolyse 6-P-glycosides.

E207 and E372 have been identified as catalytic residues by comparison with other β-glycosidases and the importance of E372 has been shown by site-directed mutagenesis (Voorhorst *et al.*, 1995). The glutamates are essential in the catalysis and serve as a general acid/base and nucleophile. The average distance between the oxygen atoms in the carboxylate residues in CelB is 4.3 Å. The active site is more or less located in the centre of mass of a subunit and can be reached by a 20 Å-long channel. The channel runs between the fifth and sixth βα-motive and is opened by a 50°-bending of the helix of the fifth βα-unit. The unusual amino acid sequence W287-W288-F289-F290 at this position in *S. solfataricus* LacS is not present in CelB.

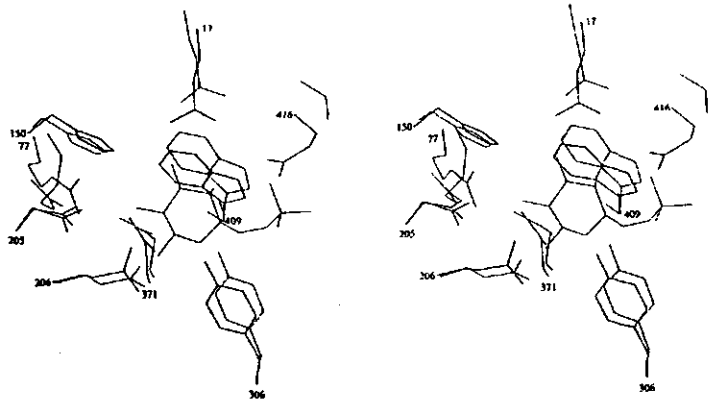


Figure 2. Superposition of *P. furiosus* CelB (thicker line) with *L. lactis* LacG (thinner line) that has galactose-6-phosphate bound in the active site. The residues that are involved in the glucose or galactose binding in both enzymes have been represented. Due to the limited resolution, the accuracy of the CelB model can be expected to be in the 1 Å range.

```
Pfu_CelB    NAEELP--MIITMMADAADR----------YRPHMLVSMLKA : 393
Sso_LacS    R-MHLY--MYVTMMIADDADY----------QRPYMLVSMVYQ : 408

Lla_LacG    D-MPNYKKIYITMMLGYKDEFVDN--TVYDDGRIDMVKQMLEV : 405
Sau_LacG    D-MPNYHKIYITMMLGYKDEFIESEKTVHDDARIDMVRQMLNV : 407
Eco_BglB    R-MQ--KPLFIVMMLGAKDS-VEADGSIQDDYRIAMLNDMLVQ : 391
                        #

Pfu_CelB    VYNAMKEMADVRMMLHMSLTMNYMWAQG-FRMMMLVYMMFETK : 436
Sso_LacS    VHRAINSMADVRMMLHMSLAMNYMWASG-FSMMMLLKMMYNTK : 451

Lla_LacG    LSDAIADMANVKMMFIMSLMMVFMWSNG-YEMMMLFYMMFDTQ : 448
Sau_LacG    IADAIIDMANVKMMFIMSLMMVFMWSNG-YEMMMLFYMMFETQ : 450
Eco_BglB    VNEAIADMVDIMMTSMGPIMLVMASHSQMSMMMFIYMMRDDN : 435
                                  *       * *

Pfu_CelB    K-----RYLRPMALVFREIATQKEIPEELAHLADLKFVT--RK. : 472
Sso_LacS    R-----LYWRPMALVYREIATNGAITDEIEHLNSVPPVKPLRH. : 489

Lla_LacG    E-----RYPKKMAHWYKKLAETQVIE.                  : 468
Sau_LacG    E-----RYPKKMAYWYKELAETKEIK.                  : 470
Eco_BglB    GEGSLTRTRKKMFGWYAEVIKTRGLS-----LKKIT----IKAP.: 470
```

Figure 3. Alignment of partial amino acid sequences. *Pfu*_CelB, *P. furiosus* β-glucosidase CelB (Swiss Prot Q51763); *Sso*_LacS, *S. solfataricus* β-glycosidase LacS (P22498); *Lla*_LacG, *L. lactis* 6-phospho-β-galactosidase LacG (P11546); *Sau*_LacG, *Staphylococcus aureus* 6-phospho-β-galactosidase LacG (P11175); *Eco*_BglB, *Escherichia coli* 6-phospho-β-glucosidase BglB (P11988). Conserved residues in all proteins have been shaded light gray, residues that interact with the phosphate group of galactose-6-phosphate in *L. lactis* LacG have been indicated (*) and shaded black in the 6-phospho-β-glycosidases and dark gray at the corresponding positions in the β-glycosidases. The conserved catalytic nucleophile in all enzymes has been indicated (#).

Design and production of wild-type and mutant *P. furiosus* CelB

In 6-P-β-galactosidase LacG of *Lactococcus lactis* S428, K435 and Y437 form a binding site for the phosphate group of galactose-6-phosphate (Fig. 2). These three residues are highly conserved among family 1 6-P-β-galactosidases (Fig 3). In *P. furiosus* CelB the corresponding residues at these positions are E417, M424 and F426 residues which, in turn, have been highly conserved among family 1 β-glycosidases that hydrolyse non-phosphorylated substrates. The residues that create the phosphate binding-site in LacG were introduced in CelB by a PCR-based method. This resulted in three CelB single mutants, E417S, M424K and F426Y, one CelB double mutant M424K/F426Y, and one CelB triple mutant E417S/M424K/F426Y. The wild-type and mutant *celB* genes were overexpressed in *E. coli* to more than 20% of the total cell protein and subsequently purified to apparent homogeneity as judged by SDS-PAGE and spectrophotometry (not shown) and finally used for further characterisation.

pH optima and kinetic parameters of mutant and wild-type CelB

Wild-type CelB hydrolyses the chromogenic substrate pNp-β-glu optimally at pH 5.0 (Kengen *et al.*, 1993). The glycosidic bond is cleaved by a common double displacement mechanism (Bauer & Kelly, 1998). The pH optimum for this mechanism is found to be determined by the pK_a values of the

Table II. pH optima of wild-type CelB and CelB mutants at 90°C in 0.1 M Na-citrate/NaP_i

	oNp-glu	oNp-gal	oNp-gal-6-P
CelB wt	5.0	5.0	4.0
CelB E417S	4.5	4.5	5.0
CelB M424K	6.0	5.75	4.5
CelB F426Y	5.0	5.0	4.0
CelB double mutant	6.0	6.25	4.5
CelB triple mutant	5.0	5.0	6.0

two catalytic residues (McIntosh *et al.*, 1996). Since the substitutions E417S and M424K introduced charge differences in the active site, these could influence the pK_a values of these catalytic residues and, as such, change the pH-optimum for hydrolysis. Therefore, the pH optima for the hydrolysis of ortho-nitrophenol-β-D-glucopyranoside (oNp-glu), ortho-nitrophenol-β-D-galactopyranoside (oNp-gal) and ortho-nitrophenol-β-D-galactopyranoside-6-phosphate (oNp-gal-6-P) were determined for different mutants (Table II).

The pH optima for the hydrolysis of oNp-glu and oNp-gal were essentially the same; this was found for wild-type CelB as well as the CelB mutants (Table II). Large differences were observed in the shape of the pH-curves. Three groups could be distinguished: i) wild-type CelB, CelB M424K, and CelB F426Y showed an optimum activity at a broad pH range for oNp-glu and oNp-gal, but a narrow curve for oNp-gal-6-P with the optimum at a lower pH than with oNp-glu and oNp-gal as substrates (Fig. 4A and 4B), ii) CelB E417S and the CelB triple showed a pH dependency of activity for the three substrates that was opposite of that of wild type CelB (Fig. 4C and 4D), iii) the CelB double mutant hydrolysed both oNp-glu and oNp-gal as well as oNp-gal-6-P optimally at a relatively wide pH range., the optimal pH for hydrolysis with oNp-gal-6-P being lower than with oNp-glu and oNp-gal as substrates (Fig. 4E and 4F). The mutation F426Y did not have an effect on the pH optimum of hydrolysis, compared to the wild-type. However, CelB M424K hydrolysed oNp-glu and oNp-gal optimally at pH 6.0, one unit higher than the wild-type. This mutant also hydrolysed oNp-gal-6-P optimally at a higher pH than the wild-type.

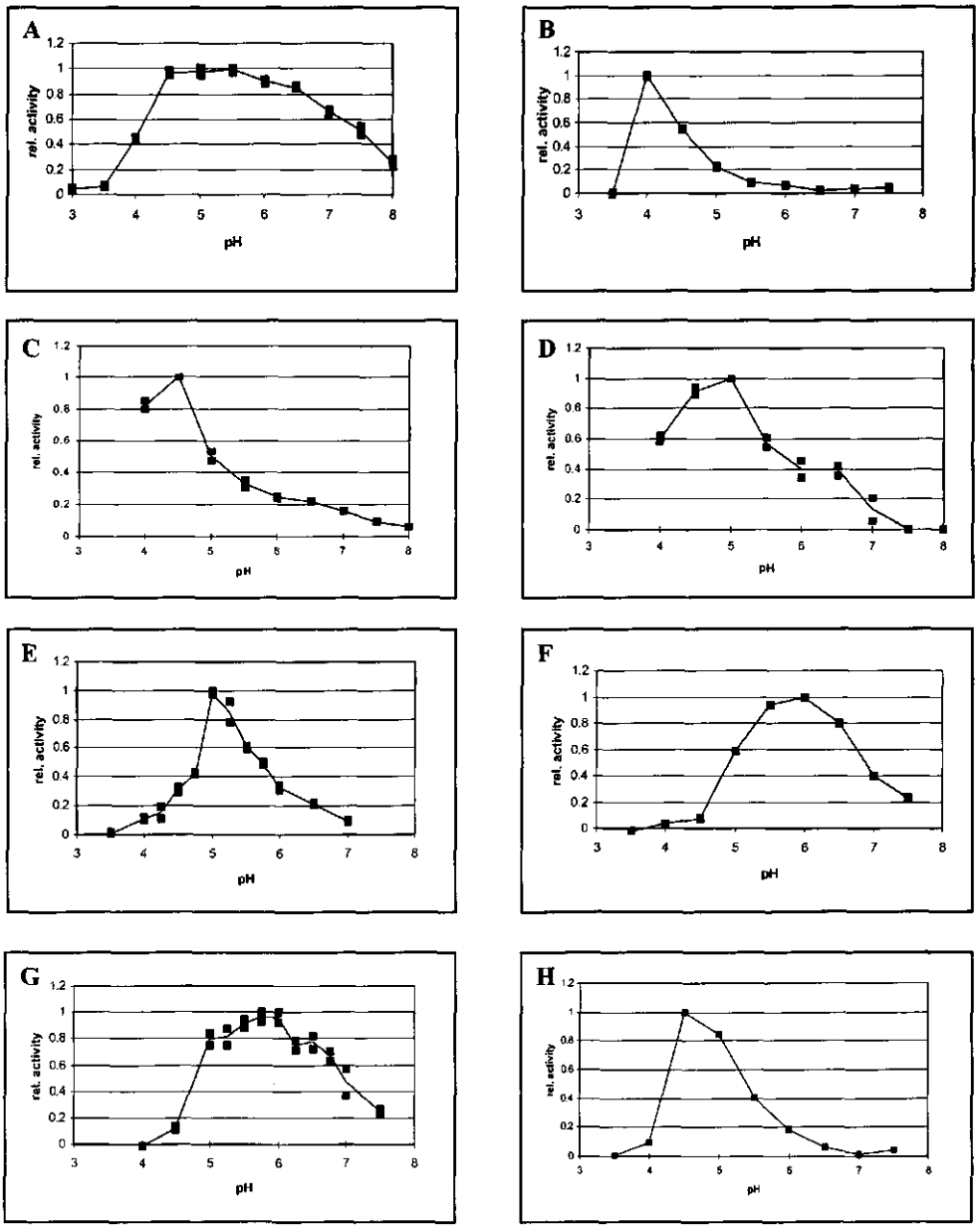


Figure 4. Typical effects of pH on enzyme activity at 90°C in combined citrate-phosphate buffer. Assays were performed with 8.3 ng to 3.3 µg enzyme. Maximum activity in each case was set at 1. A: wild-type CelB with oNp-glu (max. 920 U/mg), B: wild-type CelB with oNp-gal-6P (max. 30.7 U/mg), C: CelB E417S with oNp-glu (max. 28.7 U/mg), D: CelB E417S with oNp-gal-6-P (max. 88.6 U/mg), E: CelB triple with oNp-glu (max. 124 U/mg), F: CelB triple with oNp-gal-6-P (max. 54.3 U/mg), G: CelB double mutant with oNp-glu (max. 1241 U/mg), H: CelB double mutant with oNp-gal-6-P (max. 57.9 U/mg).

Table III. Kinetic constants of wild-type and mutant CelB with oNp-glu and oNp-gal as substrates at 90°C in 0.1 M Na-citrate/NaP_i at determined pH-optimum

	oNp-glu			oNp-gal		
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
CelB wt	0.276	1093	4048	2.28	2050	899
CekB E417S	29.3	3304	113	41.9	1058	25.3
CelB M424K	0.511	772	1510	4.76	1883	396
CelB F426Y	0.566	840	1484	0.891	892	1001
CelB double	3.66	1546	422	2.98	969	325
CelB triple	59.3	777	13.1	40.1	175	4.36

Activity assays were performed at the pH optima for the substrates and the kinetic constants for the hydrolysis of oNp-glu and oNp-gal were determined (Table III), as well as those for the hydrolysis of oNp-gal-6-P (Table IV). For wild-type CelB we determined K_m-values of 0.276 mM for oNp-glu and 2.28 mM for oNp-gal. All introduced mutations increased the K_m for oNp-glu compared to the wild-type. Moreover, the turnover number decreased for most mutants and this resulted in lower efficiencies for the hydrolysis of oNp-glu. The K_m for oNp-gal increased in CelB E417S, CelB M424K, and the CelB triple mutant, was unaffected in the CelB double mutant, but was significantly lowered in CelB F426Y. This made CelB F426Y slightly more efficient on oNp-gal than the wild-type. The K_m-values for oNp-glu and oNp-gal in the mutants that contain the mutation E417S (CelB E417S and CelB triple) increased 20-fold to 200-fold.

Wild-type CelB is able to hydrolyse oNp-gal-6P inefficiently with a K_m of 30.9 mM and a k_{cat} of 49.4 s^{-1}(Table IV). CelB E417S and the CelB triple mutant have lower K_m-values and higher turnover numbers. CelB F426Y has a slightly higher K_m-value and a reduced k_{cat}-value with respect to wild-type CelB. With the analysis of the observed K_m-values it has to be taken into account that some of them were too high to allow assays at saturating substrate concentrations (CelB E417S and CelB triple with oNp-glu and oNp-gal; CelB wild-type and CelB F426Y with oNp-gal-6-P). The data obtained for CelB M424K and the CelB double mutant could not be fitted with the Michaelis-Menten equation, because the measured activities showed a linear relationship within used the substrate concentration range. CelB E417S and CelB triple are 4.9-fold and 2.7-fold more efficient in hydrolysing oNp-gal-6P than wild-type CelB.

Table IV. Kinetic constants of wild-type and mutant CelB at 90°C with oNp-gal-6-P as substrate in 0.1 M Na-citrate/NaP_i at pH-optimum

	oNp-gal-6-P		
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
CelB wt	30.9	49.4	1.57
CelB E417S	22.3	171	7.67
CelB M424K	-	(11.3)[1]	-
CelB F426Y	38.4	25.6	0.66
CelB double mutant	-	(57.9)[1]	-
CelB triple mutant	16.2	68.8	4.24
LacG[2]	0.4	83	208

[1] data could not be fitted according to Michaelis-Menten kinetics, highest measured k_{cat} shown
[2] LacG activity measured at 25°C, data obtained from Hengstenberg *et al.*, 1995

Discussion

CelB structural model and active site

A structural model of *P. furiosus* ß-glucosidase with a resolution of 3.3 Å has been built using the molecular replacement method with the 3D structures of *L. lactis* LacG and *S. solfataricus* LacS. The resulting model has been analysed with respect to the active site.

The central fold of CelB is a $(\beta\alpha)_8$-barrel (Fig. 1), first found in triose-phosphate isomerase (Banner *et al.*, 1975). This fold has been predicted for all family 1 glycosyl hydrolases and has been observed for all established structures of family 1 glycosyl hydrolases (Barrett *et al.*, 1995; Wiesmann *et al.*, 1995, 1997; Aguilar *et al.*, 1997; Burmeister *et al.*, 1997; Sanz-Aparicio *et al.*, 1998). The positions of the active site residues in CelB, LacG and LacS have been highly conserved. The distance between the carboxylic oxygen atoms of the catalytic residues E206 and E372 of 4.3 Å is clearly shorter than the average 5.3 Å from 10 known retaining β-glucosidases (McCarter & Withers, 1994). The residues are located at the C-terminal ends of the fourth and seventh β-strand of the barrel and identify CelB as a member of the 4/7-super family ((Jenkins *et al.*, 1995).

The family 1 6-P-β-glycosidases have evolved a binding site for the phosphate group of galactose-6-phosphate. In *L. lactis* LacG S428, K435 and Y437 were found to interact with the phosphate group (Wiesmann et *al.*, 1997). The corresponding residues in CelB, E417, M424 and F426, are highly

conserved in the β-glycosidases (Fig 3). The negatively charged E417 would have a repulsive action on the phosphate group of an incoming 6-phospho-glycoside. The mutations E417S, M424K and F426Y were therefore proposed to create a phosphate binding site in CelB, in order to optimise it for the hydrolysis of 6-phospho-glycosides. In *S. solfataricus* LacS E432, the corresponding residue of E417 in CelB, was postulated to interfere with the carboxylate group of substrates such as galuronic acid or glucuronic acid (Aguilar *et al.*, 1997). From the crystal structure of the family 1 *B. polymyxa* β-glucosidase BglA with a gluconate molecule bound in the active site (Sanz-Aparicio et *al.*, 1998), it can be deduced that only E417 is likely to interact with a non-phosphorylated substrate in CelB. The proposed mutation E417S, therefore, can be expected to have a significant effect on the binding of non-phosphorylated sugars.

Characterisation of mutants

CelB hydrolyses its substrates by a retaining mechanism in which a covalent intermediate is formed (Bauer & Kelly, 1998). In the catalytic cycle, the C1-atom of an incoming glycoside is attacked by the nucleophilic glutamate. The glycosidic bond is broken and a covalent enzyme-substrate intermediate is formed. An incoming water molecule hydrolyses the covalent bond and the enzyme returns to the initial state (White & Rose, 1997). The hydrolysis of the glycosyl-enzyme intermediate is the rate-limiting step in the reaction with oNp-glu, and the rate is dependent on the ionisation state of the general acid/base catalytic residue that abstracts a proton of the incoming water which hydrolyses the covalent bond between the enzyme and the sugar moiety (Kempton & Withers, 1992).

The wild-type CelB and CelB E417S, CelB M424K, CelB F426Y, CelB double and triple were tested for their ability to hydrolyse oNp-glu, oNp-gal and oNp-gal-6-P. Wild-type CelB has little activity with oNp-gal-6-P as substrate and the pH optimum for hydrolysis is rather low, pH 4.0. This pH optimum is rather narrow, which indicates that wild-type CelB is not adapted to this substrate (see Fig 4). In CelB M424K, the optimal pH of hydrolysis of oNp-glu and oNp-gal was pH 6.0, 1 unit higher than wild-type CelB. Probably, the introduced positive charge at this position has an effect on the pK_a-value of the general acid/base. The distance between E207 and K424, ca. 15 Å, is too large for a direct interaction, but there could be an influence of the charge. Besides, mutation of the general acid/base residue in *S. solfataricus* β-glycosidase LacS and *Agrobacterium faecalis* β-glucosidase Abg did not abolish the activity completely, which means that other residues are capable of taking over the role of the glutamate. K424 could have an effect on these residues. The lysine contributes to the position of the pH optimum for the hydrolysis of phosphorylated galactose, as can be concluded from the difference in the pH optima of CelB E417S and the CelB triple mutant with oNp-gal-6-P as substrate. The lysine at this position is also present in wild-type β-glucosidase from *Thermotoga maritima* and *Caldocellum saccharolyticum*, and, remarkably, the pH optimum of the β-glucosidase of *T. maritima* was the same as that found for CelB M424K (Gabelsberger et *al.*, 1993),

The mutation E417S has the largest effect on the hydrolysis of oNp-gal-6-P. CelB E417S and CelB triple are the mutants that have an improved efficiency

for the hydrolysis of phosphorylated galactose. The combination of the three mutations in CelB triple results in the lowest K_m, which is 40-fold higher at 90°C than the K_m of LacG for oNp-gal-6P at 25°C, but the k_{cat} is in the same range. The effect of temperature on the K_m is not predictable, as has been found in other studies (Chapters 3, 5 and 9 of this thesis). The corresponding glutamate residue of E417 in *Bacillus polymyxa* β-glucosidase BglA was found to interact with the hydroxyl group of C4 and C6 of a gluconate molecule bound in the active site. Interactions with the hydroxyl group at the C4 position are of importance for the binding of glucose and galactose, which are identical molecules except for the position of the hydroxyl group on the C4 position, equatorial in glucose and axial in galactose. It is believed that the position of this glutamate residue is partly responsible for the difference in substrate specificity between a glucosidase and a galactosidase (Sanz-Aparicio *et al.*, 1998). Furthermore, E417 is important for the stabilisation of the glycosyl-enzyme intermediate in the hydrolysis cycle (Namchuk & Withers, 1995). The effects of the removed interaction with the C4 hydroxyl group become clear from the kinetic parameters on glucose and galactose in CelB E417S. The K_m-values for oNp-glu and oNp-gal were found to increase 100-fold and 20-fold.

Based on a comparison of the 3D structure of *L. lactis* 6-phospho-β-galactosidase LacG and the model of the *P. furiosus* β-glucosidase, three substitutions were introduced in CelB, which increased its activity towards phosphorylated galactose. However, the catalytic efficiency of *L. lactis* LacG for oNp-gal-6-P is higher than the CelB mutants, mainly because the K_m-values of the CelB mutants are relatively high. Probably, subtle changes in the backbone will be needed to further optimise CelB for phosphorylated galactose, besides the introduction of residues in the active site.

The comparison of related structures can provide a basis to introduce substrate specificity characteristics of one enzyme into another. This study shows that this may result in an enzyme that is more efficient for a desired reaction than the wild-type. This is the first example of a β-glucosidase that has been optimised for the hydrolysis of oNp-gal-6-P. The introduction of a functional phosphate binding site in CelB indicates that the molecular adaptations of 6-P-glycosidases from mesophilic organisms for their substrates are also applicable for glycosidases that are active at extreme temperatures.

Acknowledgements

We would like to thank Wilfried Voorhorst and Jurre Koning for their contributions in early stages of this work, Willem van Berkel for valuable discussions, and Lawrence Pearl and Mose Rossi for sharing the *S. solfataricus* LacS coordinates prior to publication. This work has been partly supported by the EU in contracts FAIR CT96-1048 and BIO 4-CT96-0488.

References

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the β-glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D. & Waley, S. G. (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. *Nature* **255**, 609-14.

Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. & Hughes, M. A. (1995). The crystal structure of a cyanogenic β-glucosidase from white clover, a family 1 glycosyl hydrolase. *Structure* **3**, 951-60.

Bauer, M. W. & Kelly, R. M. (1998). The family 1 β-glucosidases from *Pyrococcus furiosus* and *Agrobacterium faecalis* share a common catalytic mechanism. *Biochemistry* **37**, 17170-8.

Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. (1997). The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure* **5**, 663-75.

Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-9.

Febbraio, F., Barone, R., S, D. A., Rossi, M., Nucci, R., Piccialli, G., De Napoli, L., Orru, S. & Pucci, P. (1997). Identification of the active site nucleophile in the thermostable β-glycosidase from the archaeon *Sulfolobus solfataricus* expressed in *Escherichia coli*. *Biochemistry* **36**, 3068-75.

Gabelsberger, J., Liebl, W. & Schleifer, K. (1993). Purification and properties of recombinant β-glucosidase of the hyperthermophilic bacterium *Thermotoga maritima*. *Appl. Microbiol. Biotechnol.***40**, 44-52.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-26.

Henrissat, B. (1999) http://afmb.cnrs-mrs.fr/~pedro/CAZY/ghf.html.

Hengstenberg, W., Kohlbrecher, D., Witt. E, Kruse, R., Christansen, I., Peters, D., Von Strandman, R.P., Städtler, P., Koch, B. & Kalbitzer, H.R. (1993) Structure and function of proteins of the phosphotransferase system and of 6-phospho-β-galactosidases in Gram-positive bacteria. *FEMS Microbiol. Rev.***12**, 149-64

Higuchi, R., Krummel, B. & Saiki, R. K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucl. Ac. Res.* **16**, 7351-67.

Jancarik, J., Scott, W. G., Milligan, D. L., Koshland, D. E., Jr. & Kim, S. H. (1991). Crystallisation and preliminary X-ray diffraction study of the ligand-binding domain of the bacterial chemotaxis-mediating aspartate receptor of *Salmonella typhimurium*. *J. Mol. Biol.* **221**, 31-4.

Jenkins, J., Lo Leggio, L., Harris, G. & Pickersgill, R. (1995). β-glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β-strands four and seven. *FEBS Lett.* **362**, 281-5.

Kempton, J. B. & Withers, S. G. (1992). Mechanism of *Agrobacterium* β-glucosidase: kinetic studies. *Biochemistry* **31**, 9961-9.

Kengen, S. W., Luesink, E. J., Stams, A. J. & Zehnder, A. J. (1993). Purification and characterisation of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**, 305-12.

McCarter, J. D. & Withers, S. G. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **4**, 885-92.

McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W. & Withers, S. G. (1996). The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a 13C-NMR study of *Bacillus circulans* xylanase. *Biochemistry* **35**, 9958-66.

Moracci, M., Capalbo, L., Ciaramella, M. & Rossi, M. (1996). Identification of two glutamic acid residues essential for catalysis in the β-glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Prot. Engng.* **9**, 1191-5.

Namchuk, M. N. & Withers, S. G. (1995). Mechanism of Agrobacterium β-glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry* **34**, 16194-202.

Sanz-Aparicio, J., Hermoso, J. A., Martinez Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). Crystal structure of β-glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. *J. Mol. Biol.* **275**, 491-502.

Sinnot, M.L. (1990). Catalytic mechanisms of enzymatic glycosyl transfer. *Chem. Rev.* **90**, 1171-1202

Street, I. P., Kempton, J. B. & Withers, S. G. (1992). Inactivation of a β-glucosidase through the accumulation of a stable 2-deoxy-2-fluoro-alpha-D-glucopyranosyl-enzyme intermediate: a detailed investigation. *Biochemistry* **31**, 9970-8.

Trimbur, D. E., Warren, R. A. & Withers, S. G. (1992). Region-directed mutagenesis of residues surrounding the active site nucleophile in β-glucosidase from *Agrobacterium faecalis*. *J. Biol. Chem.* **267**, 10248-51.

Voorhorst, W. G. B., Eggen, R. I. L., Luesink, E. J. & De Vos, W. M. (1995). Characterisation of the *celB* gene coding for β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutagenesis in *Escherichia coli*. *J. Bacteriol.* **177**, 7105-11.

Wang, Q., Trimbur, D., Graham, R., Warren, R. A. & Withers, S. G. (1995). Identification of the acid/base catalyst in *Agrobacterium faecalis* β-glucosidase by kinetic analysis of mutants. *Biochemistry* **34**, 14554-62.

White, A. & Rose, D. R. (1997). Mechanism of catalysis by retaining β-glycosyl hydrolases. *Curr. Opin. Struct. Biol.* **7**, 645-51.

Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). The three-dimensional structure of 6-phospho-β-galactosidase from *Lactococcus lactis*. *Structure* **3**, 961-8.

Wiesmann, C., Hengstenberg, W. & Schulz, G. E. (1997). Crystal structures and mechanism of 6-phospho-β-galactosidase from *Lactococcus lactis*. *J. Mol. Biol.* **269**, 851-60.

Chapter 9

Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* β-glucosidase CelB by directed evolution.

Joyce H.G. Lebbink, Thijs Kaper, Peter Bron, John van der Oost and Willem M. de Vos.

Submitted for publication.

Abstract

The β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* (CelB) is the most thermostable and thermoactive family I glycosyl hydrolase described to date. To obtain more insight in the molecular determinants of adaptations to high temperatures and study the possibility of optimising low-temperature activity of a hyperthermostable enzyme, we generated a library of random CelB mutants in *Escherichia coli*. This library was screened for increased activity on *p*-nitrophenol-β-D-glucopyranoside at room temperature. Multiple CelB variants were identified with up to three-fold increased rates of hydrolysis of this aryl-glucoside, and ten of them were characterised in detail. Amino acid substitutions were identified in the active site region, at subunit interfaces, at the enzyme surface, and buried in the interior of the monomers. Characterisation of the mutants revealed that the increase in low-temperature activity was achieved in different ways, including altered substrate specificity and increased flexibility by an apparent overall destabilisation of the enzyme. Kinetic characterisation of the active site mutants showed that in all cases the catalytic efficiency at 20°C on *p*-nitrophenol-β-D-glucose as well as on the disaccharide cellobiose, was up to two-fold increased. In most cases, this was achieved at the expense of β-galactosidase activity at 20°C and total catalytic efficiency at 90°C. Substrate specificity was found to be affected by many of the observed amino acid substitutions, of which only some are located in the vicinity of the active site. The largest effect on substrate specificity was observed with the CelB variant N415S that showed a 7.5-fold increase in the ratio of *p*-nitrophenol-β-D-glucopyranoside/*p*-nitrophenol-β-D-galactopyranoside hydrolysis. This asparagine at position 415 is predicted to interact with active site residues that stabilise the hydroxyl group at the C4 position of the substrate, the conformation of which is equatorial in glucose-containing, and axial in galactose-containing substrates.

Introduction

Carbohydrate polymers can be degraded by a wide variety of glycosyl hydrolases that have been classified into more than 60 different families based on amino acid sequence comparisons (Davies & Henrissat, 1995; Henrissat & Davies, 1997). In recent years, considerable progress has been made in the determination and characterisation of primary and three-dimensional structures from family I glycosyl hydrolases. From among these enzymes isolated from mesophiles, the crystal structures have been solved of the 6-phospho-β-galactosidase (LacG) from *Lactococcus lactis* and the β-glucosidase from *Bacillus*

Abbreviations

pNp	*p*-nitrophenol
pNp-glu	*p*-nitrophenol-β-D-glucopyranoside
pNp-gal	*p*-nitrophenol-β-D-galactopyranoside
X-glu	5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside

polymyxa (Wiesmann et al., 1995; Wiesmann et al., 1997; Sanz-Aparicio et al., 1998). Both of these enzymes have been cocrystallised with either the substrate or an inhibitor in the active site and therefore allow analysis of the interactions between the enzyme and its substrate. Biochemical characterisation and primary structure determination has been reported for several family I glycosyl hydrolases from hyperthermophilic organisms, that optimally grow at and around the normal boiling point of water. These include the β-glucosidase from *Thermotoga maritima* (Gabelsberger et al., 1993), the β-glycosidase LacS from *Sulfolobus solfataricus* (Cubellis et al., 1990; Pisani et al., 1990) and the β-glucosidase CelB and β-mannosidase BmnA from *Pyrococcus furiosus* (Voorhorst et al., 1995; Kengen et al., 1993; Bauer et al., 1996). The crystal structure of LacS and a three-dimensional model for CelB, based on 3.25 Å X-ray diffraction data, have recently been reported, and their comparison with the homologous structures from the mesophiles, allows for studying of the molecular adaptations of family I enzymes to high temperatures (Aguilar et al., 1997; Chapter 8).

We have focused on the most thermostable member of the family I glycosyl hydrolases, CelB from the archaeon *P.furiosus*. CelB was purified from *P.furiosus*, characterised, and found to be a tetrameric enzyme of 230 kD with identical 58 kD subunits (Kengen et al., 1993). The enzyme is very well adapted to the high growth optimum of the organism (100°C), with a half-life for thermal inactivation of 85 hours at 100°C and an optimum temperature for activity of 102-105°C. CelB shows high activity on the aryl-glucosides *p*-nitrophenol-β-1,4-D-glucopyranoside (pNp-glu) and pNp-β-1,4-D-galactose (pNp-gal), as well as on β-(1,4)-linked disaccharides cellobiose and lactose and on the β-(1,3)-linked disaccharide laminaribiose (Kengen et al., 1993; Driskill et al., 1999). Low activity was detected on pNp-mannose and pNp-xylose. The *celB* gene was cloned, sequenced and functionally overexpressed in *Escherichia coli* (Voorhorst et al., 1995). Using site-directed mutagenesis, Glu 372 was identified as the active site nucleophile, being the equivalent of the nucleophile in mesophilic homologues according to a multiple sequence alignment (Voorhorst et al., 1995). Moreover, a recent study indicated that the enzymes from *P.furiosus* and the mesophilic bacterium *A.faecalis* share a common catalytic mechanism (Bauer & Kelly, 1998). The 3-dimensional model for CelB shows that the conformation of the active site is identical to that of other family I enzymes (Chapter 8).

An alternative approach to study structure-function relations in enzymes, is that of directed evolution by combining random mutagenesis, *in vitro* recombination and rapid screening procedures (Stemmer, 1994; Kuchner & Arnold, 1997; Arnold, 1998). Subsequent biochemical and structural analysis of selected mutant enzymes has been shown to contribute to understanding the molecular basis of the observed phenotypic changes and as such to our knowledge of enzyme catalysis and stability. Directed evolution approaches have hitherto been restricted to enzymes with moderate thermostability and have not yet been applied to enzymes from hyperthermophiles. CelB is extremely suitable to be used as hyperthermostable model enzyme in the development of such a procedure, since it is efficiently produced in *E.coli*, is not post-translationally modified, and is capable of hydrolysing chromogenic

substrates, both at low and high temperatures. In this study we describe the construction of a random CelB library and the screening of this library for increased activity on pNp-glucose at room temperature. Multiple random mutants were selected and characterised, with the aim to gain insight in substrate recognition and catalysis both at low and at high temperatures, and into the relation of these features with thermoactivity and thermostability.

Materials and methods

Construction of a random mutant CelB library

Random mutations were introduced into the *celB* gene by PCR amplification using primers BG238 (5'-GCGCGCCATGGCAAAGTTCCCAAAAA ACTTCATGTTTG; *Nco*I restriction site underlined) and BG309 (5'-GTTAGCAGCCGGATCCCTA; *Bam*HI restriction site underlined) and proofreading deficient *Taq* DNA polymerase (Pharmacia, Uppsala, Sweden) (Chapter 7). DNA shuffling of the *celB* gene was essentially performed according to the protocol of Stemmer (1994) with the optimisation described by Lorimer & Pastan (1995). The resulting shuffled DNA fragments were digested with *Nco*I and *Bam*HI, the restriction sites of which overlapped the *celB* start- and stop-codon, respectively. The fragments were subsequently cloned into expression vector pET9d as translational fusion with phage ϕ10 translation initiation and termination signals (Rosenberg et al., 1987). The resulting plasmids were used to transform *E.coli* JM109(DE3) (Studier et al., 1990; Yanisch-Perron & Messing, 1985). Transformation mixtures were plated onto selective TY agar plates (1% trypton, 0.5% yeast extract, 0.5% NaCl; 1.5% granulated agar and 30 µg/ml kanamycine) supplemented with 16 µg/ml of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (X-glu, Biosynth, Switzerland). Transformants harbouring only pET9d or non-functional *celB* variants will remain white after overnight growth at 37°C, while transformants harbouring functional *celB* genes will develop a blue colour due to hydrolysis of the X-glu. Single blue colonies were transferred to microtiter plates containing 200 µl/well of TY broth supplemented with 30 µg/ml kanamycine.

Screening for increased activity on pNp-glu at room temperature

A replica of the random CelB library was prepared in microtiter plates and bacterial growth was quantified by determination of the optical density at 560 nm using a Thermomax microplate reader (Molecular Devices). Data was exported into a spreadsheet (Excel, Microsoft, USA) and corrected for blank absorption. *E.coli* cells were lysed by a combination of freezing/thawing and addition of 10 µl chloroform and 5 µl of 0.1% SDS. No negative effect of these additives on CelB activity was detected in control experiments. Viscosity was reduced by adding 10 µg/ml DNaseI and 10 mM $MgCl_2$. CelB activity at 20°C was assayed by transfer and mixing of 5 µl of lysed cells to a second set of microtiter plates containing 100 µl/well of 150 mM sodium citrate (pH 4.8) and 1 mM pNp-glu. After 30 minutes reactions were stopped by addition of 200 µl 0.5 M Na_2CO_3. Free pNp formation was quantified at 405 nm using the Thermomax microplate reader. Potential high-performance mutants were

selected on the basis of an increased ratio of $OD^{405\ nm}/OD^{560\ nm}$ relative to wild-type controls. These mutants were regrown and rescreened according to the same procedure. Plasmid DNA of confirmed high-performance mutants was subjected to a second round of mutation and selection.

Characterisation of high-performance mutants

High performance mutants were grown overnight in 10 ml selective TY medium at 200 rpm. Cells were spun down, resuspended in 1 ml of a 20 mM sodium citrate buffer (pH 4.8) and lysed by sonication (Branson sonifier). Cell debris was spun down and cell free extract was incubated for 10 min at 80°C. Denatured *E.coli* proteins were precipitated by centrifugation. Protein concentration of the heat-stable cell free extract was determined according to Bradford (Bradford, 1976) using the Bio-Rad protein assay (Biorad, Veenendaal, The Netherlands). SDS-PAGE analysis revealed that CelB was at least 90% pure in heat-stable cell-free extract (not shown). CelB activity at 20°C and 90°C was determined relative to wild-type CelB in 0.5 ml 150 mM sodium citrate (pH 4.8), containing 3 mM pNp-glu or 40 mM pNp-gal. Reactions were stopped by addition of 1 ml 0.5 M Na_2CO_3 and the free pNp concentration was quantified at 405 nm on a Hitachi U-1100 spectrophotometer. Temperature optima were determined in 150 mM sodium citrate buffer (pH 4.8) containing 3 mM pNp-Glu at desired temperatures up to 98°C in a water bath. Thermostability was analysed in an oil bath at 106°C. Heat-stable cell free extract was diluted to 50 µg/ml protein in a 20 mM sodium citrate buffer pH 4.8. 100 µl samples were incubated in glass vials closed by screw caps with teflon inlay and incubated at 106°C. Residual CelB activity after 1 h incubation was determined at 20°C as described above and compared to activity of a sample kept at room temperature.

DNA sequencing and three-dimensional structure analysis

Plasmid DNA was isolated from 3 ml cultures using the Qiaprep spin plasmid kit (Qiagen, Westburg, NL) according to the included protocol. Sequencing was performed using infrared-labelled oligos (MWG, Germany) complementary to pET9d sequences immediately flanking the cloning site on the plasmid, and the Thermo Sequenase kit (Amersham Life Science) on the Li-Cor 4000L automated sequencer. Amino acid substitutions were deduced from identified mutations in the DNA sequence of the *celB* gene and their position and possible interactions in the three-dimensional structure of CelB was analysed using the 3D-structures of BglA from *B.polymyxa* (Sanz-Aparicio et al., 1998; pdb accession code 1bgg.pdb), LacS from *S.solfataricus* (Aguilar et al., 1997; pdb accession code 1gow.pdb) and the three-dimensional model for CelB (Chapter 8), with the use of the molecular visualisation software Rasmol (Raswin Molecular Graphics V 2.6; Sayle, Glaxo Research and Development, U.K.) on a personal computer, and InsightII (V 97.0, Molecular Simulations Inc.) on a Silicon Graphics Indy workstation.

Purification of CelB and determination of biochemical parameters

One litre selective TY medium in 2 litre baffled Erlenmeyer flasks, was inoculated with *E.coli* harbouring expression plasmids containing either wild-type or mutant *celB* genes and cultured overnight at 37°C while vigorously

shaking. Cell lysate was prepared and CelB was purified using heat-precipitation and anion exchange chromatography essentially as described by Voorhorst et al., 1995) with an additional gelfiltration in 20 mM Tris-HCl buffer (pH 8.0) on a 300 ml Superdex 75 column (Pharmacia, Sweden). Pure CelB fractions were pooled and dialysed against a 20 mM sodium citrate buffer (pH 4.8). Protein concentrations were determined at 280 nm using an extinction coefficient for one subunit of $\varepsilon_m^{280nm} = 1.28*10^5\ M^{-1}cm^{-1}$ according to Gill & von Hippel (1989). Initial CelB activity of pure fractions was monitored continuously in a total volume of 1.0 ml 150 mM sodium citrate buffer (pH 4.8, set at 20°C) at 20 and 90°C on a Hitachi U-2010 spectrophotometer equipped with SPR-10 temperature controller (Hitachi, Tokyo, Japan). Free pNp formation was followed at 405 nm. Specific activities were calculated using determined extinction coefficients for pNp in the assay buffer of $\varepsilon_m^{405\ nm}$ = 0.178 mM^{-1} cm^{-1} at 20°C and $\varepsilon_m^{405\ nm}$ = 0.561 mM^{-1} cm^{-1} at 90°C. Kinetic parameters were determined by measuring initial velocity at varying substrate concentrations, ranging from 0 to 8 mM pNp-glu (up to 40 mM for N415S at 90°C) and 0.2-40 mM pNp-gal. Hydrolysis of disaccharides was performed in 150 mM citrate buffer (pH 4.8) with 0-150 mM cellobiose or 0-300 mM lactose at 20 and 90°C. Glucose formation was quantified using the Peridochrom Glucose kit (Boehringer Mannheim, Germany) according to the included protocol. The signals of the reaction substrates cellobiose and lactose and the product galactose, were found to be negligible. For each K_m and V_{max} determination, at least 14 substrate/velocity data pairs were determined. Kinetic data were fitted according to functions describing Michaelis-Menten kinetics, and corrected, when necessary, according to a function describing substrate inhibition, using the non-linear regression program Tablecurve (Tablecurve 2D for Windows, version 2.03, Jandel Scientific). k_{cat} values were calculated by normalising V_{max} according to the subunit molecular weight M_r = 54,580 Da. The formation of reaction products after 30 minutes incubation at 20 and 90°C of 20 µg wild-type CelB with 20 and 150 mM cellobiose was analysed by HPLC using a Polyspher OAHY column (Merck, Darmstadt, Germany). 20 mM solutions of glucose, cellobiose and a mixture of cellulodextrines were used as standards. Thermal inactivation of pure enzyme samples was studied by incubation of 50 µg/ml samples of wild-type and mutant CelB at 106°C in 20 mM sodium citrate buffer (pH 4.8). Remaining activity as a function of time was assayed relative to a standard kept at room temperature. First-order plots of inactivation yielded half-life values for thermal inactivation.

Results and discussion

Construction of a random mutant CelB library

The hyperthermostable β-glucosidase CelB from *P.furiosus* is optimally active at 102-105°C, and its activity at room temperature does not exceed 1% of its optimal activity. In order to study the possibility of increasing this low activity at room temperature and in order to gain more insight in the factors determining this temperature dependance of activity, we subjected the *celB* gene to random mutagenesis and a low-temperature activity screening. Random mutations were introduced into the *celB* gene using error-prone PCR and *in vitro* recombination

by DNA shuffling according to Stemmer (1994). Upon cloning the mutated *celB* genes in *E.coli*, transformants harbouring functional CelB enzymes were selected by their ability to hydrolyse the chromogenic substrate analogue X-glu at low temperature. An ordered library consisting of 6160 random *E.coli* mutants was constructed. Complete DNA sequence analysis of 9 randomly picked mutants revealed an average mutation frequency of 2.3 basepair per *celB* gene, indicating that on average 1 or 2 amino acids will be changed in each CelB enzyme (data not shown).

Screening for increased activity on pNp-Glu at room temperature

The mutant CelB library was screened for increased activity on pNp-glu at room temperature. This resulted in the identification of about 400 mutants with significantly higher activity than the wild-type controls. These mutants were transferred to new microtiter plates and rescreened. Heat-stable cell-free extract of the 42 most active mutants was further analysed for low-temperature β-glucosidase activity. Eventually, the nine most active mutants (harbouring plasmids pLUW838-pLUW846) were selected, analysed in detail and found to contain heat-stable β-glucosidase activities that were up to three-fold higher at room temperature than that of *E.coli* producing wild-type CelB (Table 1). Variation in enzyme production levels was ruled out as an explanation for this increased activity since protein analysis using SDS-PAGE, showed that CelB expression levels were similar in all mutants (data not shown). Plasmid inserts were sequenced and amino acid substitutions were deduced from the changes in the DNA sequence (Table 1).

Amino acid substitutions in high performance mutants

The 3-dimensional model of CelB (Chapter 8), and the crystal structures of LacS (Aguilar et al., 1997) and BglA from *Bacillus polymyxa* complexed with gluconate in the active site (Sanz-Aparicio et al., 1998), were used to analyse the position and possible interactions of each of the amino acid substitutions found in the high-performance mutants. Their position is marked in the CelB sequence in a multiple amino acid sequence alignment, which furthermore highlights the active site residues and residues involved in intersubunit contacts (Figure 1). Moreover, analysis of the position of the substitutions in the CelB model revealed a more or less uniform distribution over the enzyme (Figure 2). Based on the position of the substitutions in the enzyme structure, three classes of CelB mutants were distinguished (Table 1, Figure 2). The first class (Class I) involves substitutions of residues in the immediate vicinity of the active site, that often directly interact with active site residues. The second class (Class II) involves substitutions that are located at or near subunit interfaces. The remainder of the mutations was assigned to Class III, and involves substitutions at the protein surface or buried in the interior of the monomers. Five of the high-performance mutants contained single amino acid substitutions. While in these mutants the genotype-phenotype relation is straightforward, this is less clear in the remaining mutants containing multiple amino acid substitutions.

Table 1: Characterisation of high performance CelB mutants. For each mutant the amino acid substitutions are listed, as well as their classification. Class I involves amino acid substitutions in the active site region, Class II at or very near to subunit interface, and Class III at surface of the protein or buried in interior and not in one of the other categories. * pLUW847 was selected from a second generation library. Relative activities on saturating amounts of pNp-Glu (3 mM) and pNp-Gal (40 mM) at 20 and 90°C are given. Temperature optima were determined up to 98°C and thermostability relative to the wild-type was determined at 106°C.

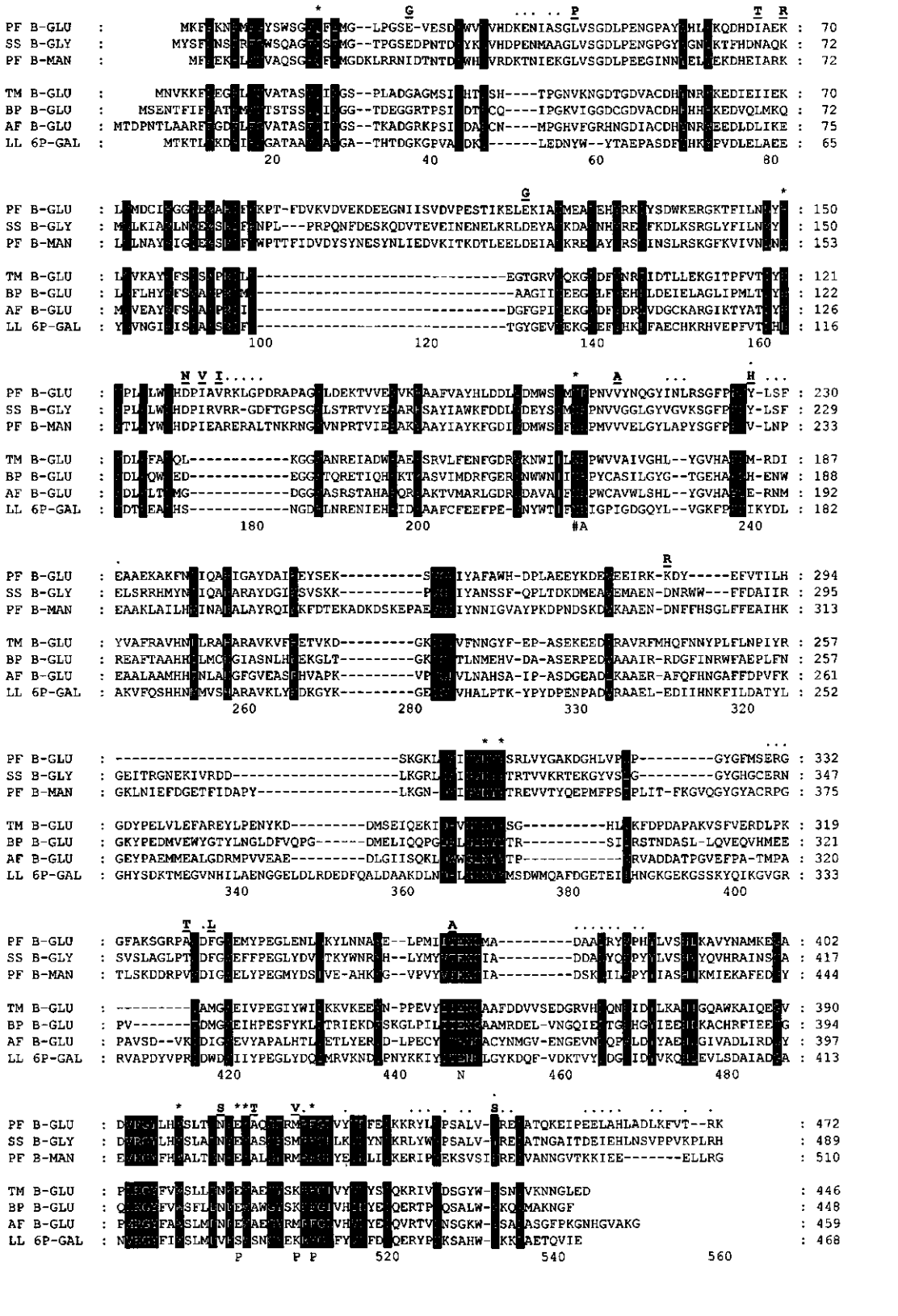
Plasmid	Amino acid substitution(s)	Class	% Act. on pNp-Glu at 20°C	% Act.on pNp-Gal at 20°C	% Act. on pNp-Glu at 90°C	% Act. on pNp-Gal at 90°C	Temperature optimum (°C)	Thermostability (% remaining activity after 1 h at 106°C)
pLUW511	wild-type		100	100	100	100	≥ 98	100
pLUW839	N415S	I	259	64	35	35	≥ 98	100
pLUW842	M424V	I	199	140	53	49	85	24
pLUW843	T371A	I	148	139	101	91	85-90	14
pLUW846	A419T	I	159	109	64	98	85	53
pLUW844	K285R	III	155	103	55	81	85-90	58
pLUW841	V211A	I	260	151	141	120	85-90	< 1
	V163I	II						
	F447S	II						
pLUW847 *	N415S	I	308	27	24	63	85	< 1
	V211A	I						
	Y227H	II						
	E26G	III						
pLUW845	I161V	II	110	37	86	125	98	84
	E119G	III						
pLUW840	K70R	II	175	147	42	< 1	70-80	< 1
	L45P	II						
	F344L	III						
pLUW838	I67T	III	110	97	70	70	≥ 98	100
	D159N	III						
	A341T	III						

The majority of the amino acid substitutions in the mutants containing only a single substitution belong to Class I (Figure 3). These mutations predominantly affect the C-terminal part of the enzyme. Plasmid pLUW839 encodes amino acid substitution N415S, which involves an asparagine residue that is hydrogen bonded to active site residues glutamine 17 and glutamate 417. These two residues directly interact with the hydroxyl group at the C4-position of the substrate. The serine side chain in the mutant is much shorter than the original asparagine. Therefore, interactions will be significantly weakened or destroyed and a cavity may be introduced. In homologous family I glycosyl hydrolases this residue is conserved except for *L.lactis* LacG (and other 6-phospho-β-galactosidases), which contains a valine (or a leucine or glutamine) at this position (Figure 1) (Witt et al., 1993). Mutant N415S displays the highest increase in β-glucosidase activity and is as thermostable as the wild-type enzyme. DNA of pLUW839, coding for this mutant, was therefore used to start a second generation evolution, resulting in a derivative (harbouring pLUW847) with considerably higher activity on pNp-glu than the N415S mutant CelB (Table 1).

Plasmid pLUW842 encodes the single amino acid substitution M424V, which involves a methionine that, like the asparagine at position 415, contacts active site glutamate 417. The valine in the resulting mutant CelB introduces a shorter but more bulky side chain, which may influence the conformation of this active site glutamate. Residues at position 424 in homologous enzymes are either methionine or lysine, which both have long, unbranched side chains (Figure 1). In LacG this lysine, together with a serine instead of E417, and a third conservative mutation (a tyrosine at position 426 instead of phenylalanine), serves to stabilise binding of the negatively charged phosphate group on the substrate (Wiesmann et al., 1997; Chapter 8).

Plasmid pLUW843 encodes T371A CelB, with a substitution that results in the removal of the polar side chain of a threonine residue that is located adjacent to the catalytic nucleophile E372. Although this threonine does not seem to interact directly with the glutamate side chain, the fact that this residue and several adjacent residues are completely conserved in all family I glycosyl

Figure 1 (next page): Alignment of family I glycosyl hydrolases. Sequences are obtained from GenBank and Swiss Prot and include: Pf β-glu = β-glucosidase (CelB) from *P.furiosus* (AF013169), Ss β-gal = β-galactosidase (LacS) from *Sulfolobus solfataricus* (P22498), Pf β-man = β-mannosidase from *P.furiosus* (U60214), Tm β-glu = β-glucosidase from *T.maritima* (Q08638), Bp β-glu = β-glucosidase (BglA) from *Bacillus polymyxa* (P22505), Af β-glu = β-glucosidase (Abg) from *Agrobacterium faecalis* (P12614), Ll 6P-gal = 6-phospho-β-galactosidase (LacG) from *L.lactis* (P11546). The active site nucleophile and acid-base catalyst are marked with N and A below the sequences, respectively; other active site residues are marked with an asterisk. Residues involved in intersubunit interactions are indicated with dots. Identified amino-acid substitutions in high-performance mutants are indicated in bold and underscored. Active site mutation N206S, a random mutant obtained from the same library with a different screening method (Chapter 7) is marked #; residues involved in stabilisation of the negatively charged phosphate group on the substrate in LacG are marked P below the sequences (Wiesmann et al., 1997; Chapter 8).


PF B-GLU
SS B-GLY
PF B-MAN
TM B-GLU
BP B-GLU
AF B-GLU
LL 6P-GAL
20
40
60
80
100
120
140
160
180
200
#A
240
260
280
330
320
340
360
380
400
420
440
N
460
480
P
P P
520
540
560

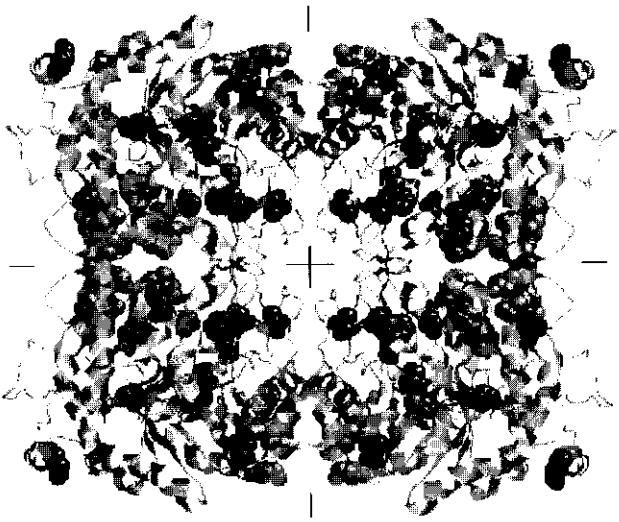


Figure 2: View of the CelB tetramer (Chapter 8). Subunits are depicted in light-grey ribbon presentation. Side chains of the amino acids that have been substituted in the high-performance CelB mutants, are shown in space-filling representation, with Class I substitutions in dark-grey, Class II substitutions in light grey and Class III substitutions in black. The figure has been generated with RASMOL.

hydrolases, strongly suggests that interactions in this region are critical for correct catalysis (Figure 1).

Plasmid pLUW846 encodes A419T CelB, which involves a buried alanine at the top of the active site barrel next to W418 that interacts with the hydroxyl group at the C3 position of the substrate. Because the alanine is buried, the larger threonine side chain will probably change the side chain packing in this area. Remarkably, the same substitution into threonine or into the similar polar residue serine has also occured in 6-phospho-β-galactosidases (Figure 1) (Witt et al., 1993).

Another plasmid that contains only a single base change in the *celB* gene, is pLUW844 coding for the amino acid substitution K285R. This mutation is located at the kink in the α-helix that creates a tunnel in the side of the $(\alpha\beta)_8$-barrel, which probably forms the substrate entrance to the active site (Aguilar et al., 1997). K285R is, however, a very conservative mutation and the residue is not conserved at all in homologous enzymes. The model of CelB predicts an ion-pair interaction between K285 and E281, which may still be formed in the mutant K285R CelB.

The fifth substitution in the active site region, although not as close to the active site itself as the other substitutions, is V211A that involves a completely buried valine in the same α-helix as active site residues N206 and the acid/base catalyst E207. Residues at this position are always hydrophobic and the mutation will introduce a cavity. This substitution is found in combination with two class II substitutions, V163I and F447S, at the subunit interface in the CelB mutant encoded by pLUW841. The V163I substitution is located in a loop that is only present in the tetrameric family I enzymes. The phenylalanine at position 447 is

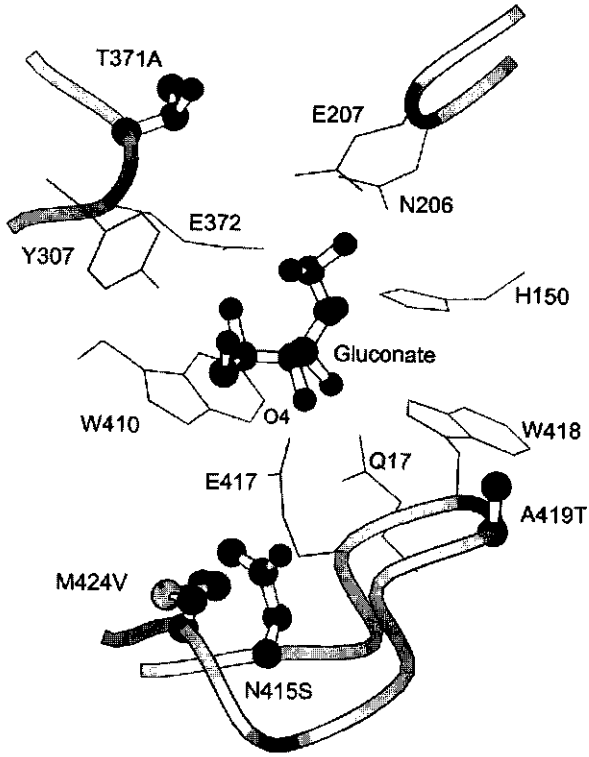


Figure 3: Close-up of the active site of *B.polymyxa* BglA with inhibitor gluconate (Sanz-Aparicio *et al.*, 1998). Gluconate is depicted in ball-and-stick representation. Active site residues are shown in stick representation. Corresponding residues that are changed in the high performance CelB mutants (T371A, M424V, N415S and A419T) are shown in ball-and-stick representation. Numbering of the residues is according to the homologous CelB residues in order to allow comparison with the text. The figure has been generated using Molscript (Kraulis, 1991).

completely buried close to the subunit interface and its substitution to serine will therefore result in a cavity. The aromatic ring at this position in the enzyme seems to be important, since in all family I enzymes either a phenylalanine or a tyrosine residue is present.

Two of the above mentioned active site substitutions, N415S and V211A, are also present in the enzyme encoded by plasmid pLUW847 that was selected from a second-generation library. The N415S substitution is derived from the template plasmid in the DNA shuffling procedure; the V211A substitution is introduced independently in this second round of evolution. Two additional mutations are present in pLUW847, resulting in E26G in which a surface-exposed charge is removed, and Y227H, in which intersubunit interactions are changed. This second-generation CelB variant is significantly more active at room temperature than N415S CelB and exceeds the wild-type CelB activity more than three-fold (Table 1).

Two mutants contain multiple substitutions belonging to Class II and III. pLUW845 codes for substitution I161V that is also located in the subunit interface loop that is unique in tetrameric family I enzymes. Furthermore, the

surface-exposed E119 has been exchanged for a glycine. This may result in the loss of electrostatic interactions and may have significant effects on the local flexibility of the polypeptide because of the large conformational freedom that is introduced with the glycine residue. pLUW840 codes for the three substitutions K70R, L45P and F344L. This triple mutant is highly destabilised, possibly mainly by the L45P substitution in which the conserved leucine is changed to a proline, which may introduce severe strain into the polypeptide chain.

Plasmid pLUW838 codes for three class III substitutions, viz. I67T, D159N and A341T. This CelB mutant displays an unchanged temperature optimum curve, is as thermostable as the wild-type enzyme, and shows relatively minor effects on activity (Table 1). These properties indicate that even a highly optimised, hyperthermostable enzyme is able to incorporate changes without loosing stability and thermoactivity. None of these residues is completely conserved among close relatives, and moreover, LacS also contains a threonine at position 341 (Figure 1).

Characterisation of high-performance CelB variants

The increase in rate of β-glucosidase activity on pNp-glu as substrate that is found in the heat-stable extract of the first generation mutants differs from 1.1- to 2.6- fold (Table 1). A further increase to more than 3-fold higher activity was obtained in the second-generation mutant. However, β-galactosidase activity is increased in only four of the mutant CelB enzymes, is unchanged in three other mutants, and has decreased in the remaining three mutant CelB enzymes (Table 1). The largest increase in the ratio of pNp-glu/pNp-gal hydrolysis is observed in the mutants containing the active site-associated mutation N415S (pLUW839 and pLUW847) as well as in the enzyme encoded by pLUW845. This latter mutant CelB is remarkable because its amino acid substitutions are at the enzyme surface (E119G) and on the subunit interface (I161V), indicating that substrate specificity is not solely determined by interactions in the active site region. Large reductions in both β-glucosidase and β-galactosidase activity at 90°C are found in the mutant CelB mutants, except for the enzyme encoded by plasmid pLUW841 that has increased activity on both substrates, the active site mutant T371A that has unchanged activity, and, again remarkably, the CelB mutant encoded by pLUW845 that shows increased β-galactosidase activity, although this was severely reduced at room temperature.

To study whether the low temperature adaptation of the mutant enzymes affected the activities at other temperatures, their temperature optima were determined. Changes that were detected in the optimum curves, could be divided in three categories (Fig 4); (i) flattening of the curve, as is found for mutant N415S (Fig 4a) that shows below 50°C higher activity than the wild-type CelB, while the reverse was observed above 60°C, (ii) shifting of the curve to lower temperature values as observed for mutant M424V (Fig 4b), and (iii) a much earlier inactivation as found for the mutant encoded by pLUW840 (Fig 4c). All mutants with lower temperature optima for activity also have a lower resistance towards thermal inactivation (Table 1).

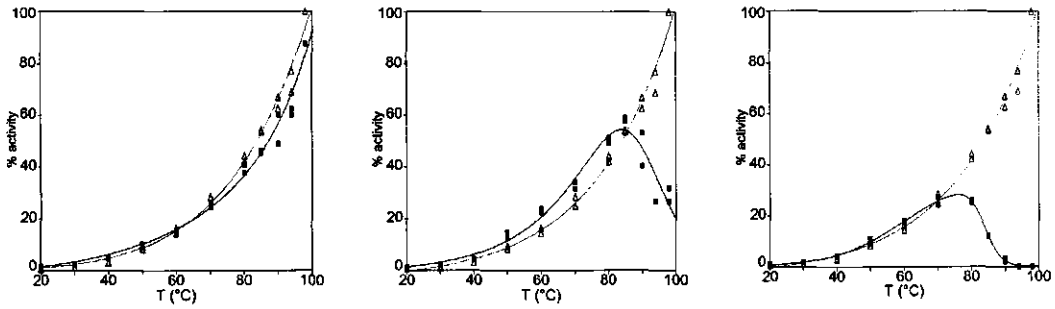


Figure 4: Optimum temperature curve for wild-type CelB (Δ) compared to a) mutant CelB N415S (▪), b) mutant CelB M424V (▪), and c) the enzyme encoded by plasmid pLUW840 (▪).

Biochemical characterisation of wild-type and active site mutant CelB

In order to gain more insight in the effects realised by the amino acid changes in the vicinity of the active site, we decided to completely purify and characterise mutant CelB enzymes containing a single substitution in this region, namely N415S, M424V, T371A and A419T. In order to obtain an accurate measure of the persistence of catalytic function at high temperature, thermal inactivation of purified wild-type and CelB mutants was followed in time (Table 2). Mutant N415S CelB is as thermostable as the wild-type enzyme, while the other mutants are considerably destabilised; compared to all other family I glycosyl hydrolases they are, however, still extremely thermostable enzymes because they all retain a considerable amount of activity after incubation at 106°C. Kinetic parameters for all these mutants were determined and compared to those of the wild-type CelB at 20°C on the aryl-glucosides pNp-glucose and pNp-galactose, as well as on β-(1,4)-linked disaccharides cellobiose and lactose (Table 2 and see below). Kinetic parameters at 90°C were determined for wild-type and mutant N415S CelB (Table 3 and see below).

Kinetic characterisation of wild-type CelB at 20 and 90°C. The hydrolysis of pNp-glu by wild-type CelB revealed an apparent inhibition in a hyperbolic Michaelis-Menten plot, starting at substrate concentrations above 1 mM at 20°C (Figure 5) and above 2-4 mM pNp-glu at 90°C (not shown). A similar phenomenon was observed during cellobiose hydrolysis (Figure 6). Plotting the data according to a Lineweaver-Burk equation resulted in a curve typically found when substrate inhibition occurs (inset in Figure 6). This process has not been reported in the previous kinetic characterisation of CelB, presumably because employed maximum substrate concentrations did not exceed 1.5 mM pNp-glu at 90°C (Bauer & Kelly, 1998). Apparent substrate inhibition was, however, also observed with the β-glucosidase from *Thermotoga maritima* (Gabelsberger et al., 1993). Fitting of the data according to Michaelis-Menten kinetics, corrected for substrate inhibition, resulted in a V_{max} of 15 U/mg and a K_m of 0.53 mM for pNp-glucose at 20 °C (Table 2). The enzyme is two-

Table 2: Kinetic parameters of wild-type and mutant CelB enzymes at 20°C on pNp-Glu, pNp-Gal and on β-(1,4)-linked disaccharides cellobiose and lactose. In the last column the half-life of thermal inactivation at 106 °C is listed.

CelB variant	pNp-Glu			Cellobiose			pNp-Gal			Lactose			Thermostability
	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})	V_{max} (U/mg)	K_m (mM)	Cat. Eff.[1] (mM^{-1} s^{-1})	$t_{1/2}$ at 106°C (h)
wild-type	15	0.53	25	18	23	0.69	9.9	0.70	13	14	52	0.25	3.3
N415S	51	1.1	43	23	20	1.0	4.5	1.2	3.5	23	95	0.22	3.2
M424V	43	0.92	43	19	16	1.1	13	3.4	3.4	23	97	0.22	1.1
T371A	28	0.57	45	17	11	1.4	12	1.3	8.2	19	52	0.33	0.65
A419T	43	1.3	31	21	25	0.75	15	4.5	3.1	16	83	0.18	0.82

[1]Cat. eff. = k_{cat}/K_m

Table 3: Kinetic parameters of wild-type CelB and mutant N415S at 90°C on pNp-Glu, pNp-Gal and on β-(1,4)-linked disaccharides cellobiose and lactose.

CelB variant	pNp-Glu			cellobiose			pNp-gal			lactose		
	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})	V_{max} (U/mg)	K_m (mM)	Cat.eff.[1] (mM^{-1} s^{-1})
wild-type	1800	0.42	3900	720	14	48	2600	5.0	480	1500	120	11
N415S	1600	1.4	1000	270	38	6.4	1100	8.5	120	590	220	2.4

[1]Cat. eff. = k_{cat}/K_m

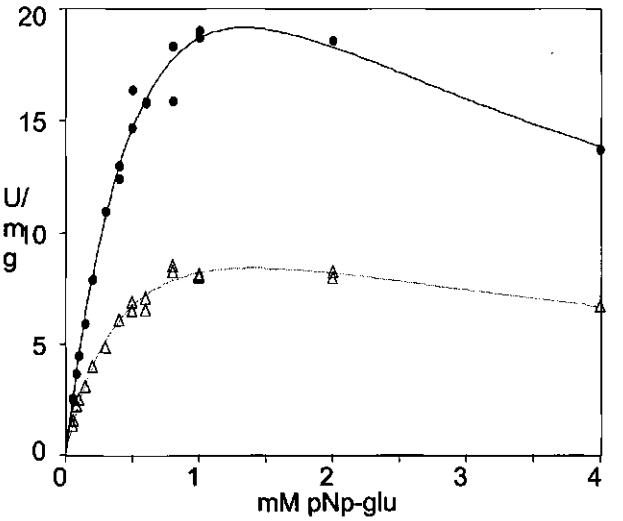


Figure 5: Comparison of rate of pNp-glu hydrolysis by wild-type (Δ) and mutant CelB N415S (•) as a function of substrate concentration at 20°C.

fold less efficient on pNp-gal than on pNp-glu, due to a lower hydrolysis rate and a lower affinity for the galactosyl substrate. On cellobiose and lactose, similar V_{max} values are found, but affinity for the β-(1,4)-linked disaccharides is relatively low. V_{max}-values for the aryl-glycosides at 90°C are 1.5- and 4-fold higher than reported earlier for the enzyme purified from *E.coli* by Voorhorst and coworkers (1995) and Bauer & Kelly (1998), respectively. This difference may be related to the observed substrate inhibition or may reflect the variability between independent enzyme preparations that have been purified by different methods. Calculated K_m values at 90°C are more or less similar to those already reported. Interestingly, while K_m values of wild-type CelB for pNp-glu and cellobiose at 90°C *decrease* compared to those at 20°C, these values *increase* for pNp-gal and lactose. Substrate affinity is, apparently, highly influenced by the operation temperature. It should be noted that, although V_{max} values for pNp-gal and lactose are considerably higher than for pNp-Glu and cellobiose, the enzyme shows, based on catalytic efficiency, a higher β-glucosidase than β-galactosidase activity.

Transglycosylation. Transglycosylation is common among family 1 glycosyl hydrolases. It has been described before for *P.furiosus* CelB and *A.faecalis* Abg in the case of pNp-xylose and pNp-arabinose substrates and for CelB as well using cellobiose and lactose (Bauer & Kelly, 1998; Kempton & Withers, 1992; Fischer et al., 1996; Boon et al., 1999). Cellotriose formation from cellobiose has been reported for BglA from *B.polymyxa* and for the β-

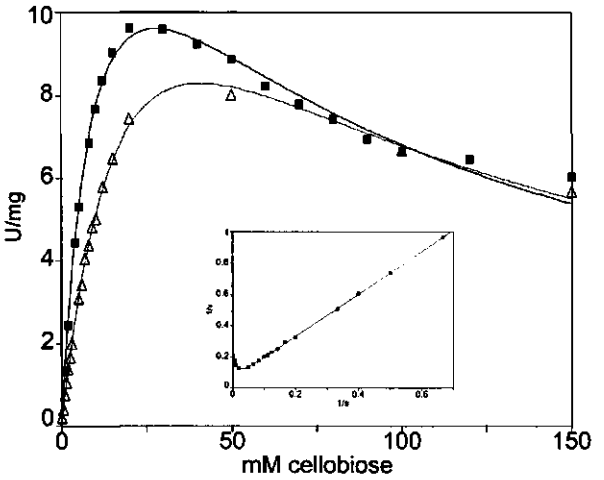


Figure 6: Comparison of the rate of cellobiose hydrolysis by wild-type (Δ) and mutant T371A CelB (■) as a function of substrate concentration at 20°C. Inset is Lineweaver-Burk plot for wild type CelB.

glucosidase from *T.maritima* (Painbeni et al., 1992; Gabelsberger et al., 1993). Analysis of reaction products of CelB incubated with 150 mM cellobiose by HPLC revealed, apart from glucose and cellobiose, the presence of an additional product with the same retention time as cellotriose, strongly suggesting that transglycosylation is indeed occurring. Kinetic characterisation of this process was not possible because its effect on product formation was completely masked by the large observed substrate inhibition (see above).

Kinetic characterisation of active site mutant CelB enzymes at 20°C. The catalytic efficiencies of mutated CelB on pNp-glu and cellobiose are increased up to two-fold in comparison with wild-type CelB (Table 2). This has either been achieved by large increases in V_{max} values, as in the case of mutant N415S CelB on pNp-glu (Figure 5), or by a decrease in K_m values as observed for T371A CelB on cellobiose (Figure 6). Increase in efficiency of pNp-glu hydrolysis is at the expense of catalytic efficiency on pNp-gal. However, a similar effect is not observed on β-(1,4)-linked disaccharides, moreover, a significant increase in efficiency is reported for T371A CelB on lactose. Evidently, the effect of the leaving group (either pNp or glucose for the aryl substrates and disaccharides, respectively) on catalytic efficiency is not comparable for glucosyl and galactosyl substrates. This is most clearly illustrated for mutant N415S CelB, for which a large reduction in maximum activity on pNp-galactose is found, that is now more than ten-fold lower than maximum activity on pNp-glucose. In contrast, cellobiose is hydrolysed with the same maximum velocity as lactose.

Kinetic characterisation of active site mutant CelB N415S at 90°C. Because N415S CelB is as thermoactive and thermostable as the wild-type CelB, we decided to analyse its kinetic parameters at 90°C in order to see the relation between these characteristics at low and high temperatures (Table 3). Maximum activity of mutant N415S CelB at 90°C was lower than for the wild-type enzyme for pNp-glu and especially pNp-gal. Because also K_m values increase, N415S CelB is less efficient than wild-type CelB on pNp-glu and pNp-gal. The same is found for the disaccharides. Higher performance on pNp-glu at room temperature is therefore not only at the expense of galactosidase activity at room temperature, but also at the expense of high-temperature catalysis. Because of the apparent unchanged temperature optimum and thermostability of this mutant, this effect is not caused by an overall destabilisation of the enzyme, but must be a direct result of the effects of the single mutation on substrate binding and turnover rates.

Conclusions

Screening of a random CelB library resulted in the selection of CelB variants with significantly higher activity on pNp-glu at room temperature than the wild-type enzyme. Amino acid substitutions were located in the active site region (Class I), at or close to subunit interfaces (Class II), or either at the protein surface or buried in the subunit interior (Class III). All mutants containing subunit interface substitutions were less thermostable and had lower temperature optima than the wild-type CelB, suggesting that subunit interfaces play an important role in thermoadaptation. Hyperthermostable enzymes are most probably highly optimised with respect to packing efficiency (Jaenicke & Bohm, 1998; Scandurra et al., 1998) However, a recent database survey showed that proteins from mesophiles and thermophiles essentially do not differ in packing (Karshikoff & Ladenstein, 1998). The present study indicates that CelB is able to accommodate amino acids in its interior with larger or smaller side chains and different properties, without affecting its thermostability or temperature optimum for catalysis.

Remarkably, substrate specificity of CelB seems to be a characteristic that is determined by the enzyme as a whole, since changes in the ratio on pNp-glu and pNp-gal are not restricted to substitutions in the active site region. However, the most drastic effect on the pNp-glu/pNp-gal ratio (a 7.5-fold increase) is caused by the disruption of interactions between asparagine at position 415 and two active site residues that directly contact the hydroxyl group on the substrate at the C4 position. This is exactly the (only) position at which glucose and galactose differ in orientation of this hydroxyl group. The loss of hydrogen bonds between the substituted asparagine and active site residues Q17 and E417 may influence the position and the flexibility of these side chains, thereby changing their interactions with the substrate and the ratio between binding and turnover on substrates with either glucose- or galactose moieties at this position. This is indeed found for mutant N415S on aryl-glycosides and, to a lesser extent, on β-(1,4)-linked disaccharides. In fact, a key role in determining glucose/galactose specificity for residues interacting with the active site Q17 and E417 was proposed recently on the basis of the interactions between the inhibitor

gluconate and active site residues of *B.polymyxa* BglA (Sanz-Aparicio et al., 1998). Further indications that slight changes in conformational freedom of active site residues may have large effects on catalysis and substrate recognition, is provided by the large and opposing effect of operational temperature on affinities for different substrates. While wild-type CelB has an 8-fold higher efficiency on pNp-glucose than on pNp-galactose at 90°C, this ratio is reduced to 2-fold at 20°C. Until recently it was assumed, based on lactate dehydrogenase from *T.maritima* and *T.neapolitana* xylose isomerase, that an increased temperature causes the decrease of one enzyme's affinity for its substrate (Hecht et al., 1989; Vieille et al., 1995; Vieille & Zeikus, 1996). However, the results presented here indicate that this assumption does not hold. Also for glutamate dehydrogenase from *P.furiosus* and *T.maritima* we have reported different changes in substrate affinity with increasing temperature, depending on the substrate (Chapters 3 and 5 of this thesis).

Our results show that the quality of the random CelB library is good and that the sensitivity of the screening procedure is sufficient to isolate mutants with only 10% increase in activity. Mutants displaying evolved properties, like N415S CelB with two-fold increased catalytic efficiency on pNp-glu but unchanged stability, and mutant T371A CelB with two-fold higher affinity for cellobiose, could be interesting candidates for industrial or diagnostic applications.

In conclusion, this random screening approach using a hyperthermostable enzyme resulted in the identification of residues that are critical in determining thermostability, low-temperature activity, and substrate recognition. It was shown that low-temperature activity can be engineered into a hyperthermostable enzyme without affecting its extreme stability. Not only is directed evolution a powerful approach to introduce a desired property into an industrially relevant biocatalyst, but, as we have shown here, it may facilitate our understanding of the mechanisms determining enzyme catalysis, stability, and substrate recognition at physiological or extreme conditions.

Acknowledgements

The authors are very grateful to Servé Kengen for helpful discussion, Ans Geerling and Ineke Heikamp-de Jong for technical assistance, and Leo de Graaff for kindly providing access to the microtiter plate reader. This work has been partially supported by contracts BIOT-CT93-0274 and BIO 4-CT96-0488 of the European Union.

References

Aguilar, C.F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., & Pearl, L. H. (1997) *J. Mol. Biol.* **271**, 789-802.

Arnold, F. H. (1998) *Acc. Chem. Res.* **31**, 125-131.

Bauer, M. W., Bylina, E. J., Swanson, R. V., & Kelly, R. M. (1996) *J. Biol. Chem.* **271**, 23749-55.

Bauer, M. W., & Kelly, R. M. (1998) *Biochemistry* **37**, 17170-8.

Boon, M. A., Van der Oost, J., de Vos, W.M., Janssen, A.E.M., Van 't Riet, K. (1998) *Appl. Biochem. Biotechnol.* **75**, 269-278.

Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-54.

Cubellis, M. V., Rozzo, C., Montecucchi, P., & Rossi, M. (1990) *Gene* **94**, 89-94.

Davies, G., & Henrissat, B. (1995) *Structure* **3**, 853-9.

Driskill L.E., Kusy, K., Bauer, M.W., Kelly, R. M. (1999) *Appl. Environ. Microbiol.* **65**, 893-7.

Fischer, L., Bromann, R., Kengen, S. W. M., de Vos, W. M., & Wagner, F. (1996) *Biotechnology* **14**, 88-91.

Gabelsberger, J., Liebl, W., Schleifer, K.-H. (1993) *Appl. Microb. Biotechnol.* **40**, 44-52.

Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem.* **182,** 319-26.
Hecht, K., Wrba, A., & Jaenicke, R. (1989) *Eur. J. Biochem.* **183,** 69-74.
Henrissat, B., & Davies, G. (1997) *Curr. Opin. Struct. Biol.* **7,** 637-644.
Jaenicke, R., & Bohm, G. (1998) *Curr. Opin. Struct. Biol.* **8,** 738-748.
Karshikoff, A. & Ladenstein, R. (1998) *Prot. Engng.* **11,** 867-872.
Kempton, J. B., & Withers, S. G. (1992) *Biochemistry* **31,** 9961-9969.
Kengen, S. W. M., Luesink, E. J., Stams, A. J., & Zehnder, A. J. (1993) *Eur. J. Biochem.* **213,** 305-312.
Kraulis, P.J. (1991) *J. Appl. Crystall.* **24,** 946-950.
Kuchner, O., & Arnold, F. H. (1997) *Trends Biotechnol.* **15,** 523-530.
Lorimer, I. A., & Pastan, I. (1995) *Nucl. Acids Res.* **23,** 3067-8.
Painbeni, E., Valles, S., Polaina, J., & Flors, A. (1992) *J. Bacteriol.* **174,** 3087-3091.
Pisani, F. M., Rella, R., Raia, C. A., Rozzo, C., Nucci, R., Gambacorta, A., De Rosa, M., & Rossi, M. (1990) *Eur. J. Biochem.* **187,** 321-328.
Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J., & Studier, F. W. (1987) *Gene* **56,** 125-135.
Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L., & Polaina, J. (1998) *J. Mol. Biol.* **275,** 491-502.
Scandurra, R., Consalvi, V., Chiaraluce, R., Politi, L., & Engel, P. C. (1998) *Biochimie* **80,** 933-941.
Stemmer, W. P. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 10747-10751.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60-89.
Vieille, C., Hess, J. M., Kelly, R. M., & Zeikus, J. G. (1995) *Appl. Environ. Microbiol.* **61,** 1867-1875.
Vieille, C., & Zeikus, J. G. (1996) *Trends Biotechnol.* **14,** 183-190.
Voorhorst, W. G., Eggen, R. I., Luesink, E. J., & de Vos, W. M. (1995) *J. Bacteriol.* **177,** 7105-7111.
Wiesmann, C., Beste, G., Hengstenberg, W., & Schulz, G. E. (1995) *Structure* **3,** 961-968.
Wiesmann, C., Hengstenberg, W., & Schulz, G. E. (1997) *J. Mol. Biol.* **269,** 851-860.
Witt, E., Frank, R., & Hengstenberg, W. (1993) *Protein Engng.* **6,** 913-920.
Yanisch-Perron C, V. J., Messing J. (1985) *Gene* **33,** 103-119.

Chapter 10

Thermostability of the β-glucosidase CelB from the hyperthermophilic archaeon *Pyrococcus furiosus*

Joyce H.G. Lebbink, Thijs Kaper, Sander Smits, Mark van Passel, John van der Oost and Willem M. de Vos.

Summary

The β-glucosidase CelB from the hyperthermophilic archaeon *Pyrococcus furiosus* is a highly thermostable and thermoactive enzyme with an optimal temperature for activity of 102-105°C, a half life for thermal inactivation of 85 hours at 100°C, and an apparent melting temperature of 108°C. The molecular basis for this extreme stability has been addressed by comparative analysis of the three-dimensional model of CelB and the crystal structure of the homologous β-glycosidase LacS from *Sulfolobus solfataricus*, as well as by biochemical characterisation of wild-type CelB and mutants thereof, containing substitutions or deletions of charged residues at the C-terminus, which may be involved in the formation of ion-pair networks. All mutant CelB enzymes show decreased optimal temperatures for activity and increased rates of thermal inactivation, indicating that ion-pairs at the C-terminus of CelB play a role in determining the extreme thermostability of this enzyme. Furthermore, increased thermal inactivation of CelB in the presence of sodium chloride and at extreme pHs, supports the role of electrostatic interactions in this enzyme. The CelB structure indicates that part of CelB stability is derived from structural adaptations that allow oligomerisation and the presence of solvent-filled, hydrophilic cavities, which have earlier been identified in LacS. However, structural comparison of several large ion-pair networks that are present in LacS, reveals that a five-residue and a six-residue membered network are not present in CelB and that a large 16-membered network at the subunit interface, extensively cross-linking the subunits, has been fragmented. CelB must therefore be stabilised to a large extent by other factors, such as efficient hydrophobic packing, optimised hydrogen bonding patterns or reduction of exposed solvent area.

Introduction

The recent elucidation of a variety of three-dimensional structures of enzymes from hyperthermophilic Bacteria and Archaea has revealed a large variety of stabilising mechanisms. Among these are reduction of the number of thermolabile residues, decreased surface-to-volume ratio, improved hydrophobic packing, improved hydrogen bonding and electrostatic interactions, and oligomerisation (recently reviewed by Ladenstein & Antranikian, 1998; Jaenicke & Bohm, 1998). Many of these strategies are theoretically supported by model calculations and data base surveys (Spassov *et al.*, 1995; Elcock, 1998). One of the strategies that seems to be employed by many enzymes from hyperthermophiles, is an increase in the number of ion-pairs and the arrangement of these ion-pairs into large networks (Yip *et al.*, 1995; Aguilar *et al.*, 1997). Several protein engineering studies have corroborated the stabilising effect of large ion-pair networks at subunit interfaces in oligomeric enzymes (Pappenberger *et al.*, 1997; Vetriani *et al.*, 1998; *Chapter 6 of this thesis*). However, for networks at the protein surface and at domain interfaces of the same oligomeric enzymes, no stabilising effect could be detected (Tomschy *et al.*, 1994; Chapter 5 of this thesis).

Glycosyl hydrolases are enzymes that are able to hydrolyse oligosaccharides, and they are classified in more than 60 sequence-based

families (Davies & Henrissat, 1995; Henrissat & Davies, 1997). Glycosyl hydrolases belonging to family I include β-glucosidases, β-galactosidases, and 6-phospho-β-glycosidases. The most thermostable representative of this family is β-glucosidase (CelB) from the hyperthermophilic archaeon *Pyrococcus furiosus*, encoded by the *celB* gene (Kengen *et al.*, 1993; Voorhorst *et al.*, 1995). This enzyme is involved in the intracellular cleavage of oligosaccharides during growth of this organism on polysaccharides at temperatures up to 103°C. CelB has been purified from *P.furiosus* and overproducing *E.coli* strains, and found to be a tetrameric enzyme with 58 kD subunits that retained their kinetic and stability properties irrespective of its source (Kengen *et al.*, 1993; Voorhorst *et al.*, 1995). This allowed for the first protein engineering studies of a hyperthermostable enzyme and showed that the active site nucleophile at 100°C is the same as that of other family I glycosyl hydrolases at ambient temperatures (Voorhorst *et al.*, 1995). CelB is very well adapted to high temperatures, with an optimal temperature for activity at 102-105°C, a half life of thermal inactivation of 85 h at 100°C, and an apparent melting temperature of 108°C (Kengen *et al.*, 1993; Bauer & Kelly, 1998).

Several three-dimensional structures of family I glycosyl hydrolases are available, among which that of the homologous β-glycosidase (LacS) from the hyperthermophilic archaeon *Sulfolobus solfataricus* (Aguilar *et al.*, 1997). LacS is a thermostable, tetrameric enzyme with a half-life of thermal inactivation of 3 h at 85°C and a temperature optimum for activity of 95°C (Pisani *et al.*, 1990). The active site nucleophile and acid-base catalyst (Glu 387 and Glu 206, respectively) were identified by a site-directed mutagenesis approach (Moracci *et al.*, 1996). The crystal structure of LacS revealed two possible mechanisms, that may be responsible for the high thermostability of LacS, i.e. the presence of large, solvent-filled, hydrophilic cavities on the one hand, and the formation of large ion-pair networks on the other hand (Aguilar *et al.*, 1997). C-terminal deletions and single amino acid substitutions, that are predicted to disrupt an extensive ion-pair network which tightly crosslinks the subunits, significantly reduced the resistance of LacS towards thermal inactivation at high temperatures (Moracci *et al.*, 1998). Recently, a three-dimensional model of CelB has been reported, based on 3.25 Å X-ray diffraction data and LacS as search model in the molecular replacement procedure (Chapter 8). LacS and CelB share 55% amino acid identity, a similar overall three-dimensional fold as well as many catalytic properties, but differ considerably in thermostability.

In this chapter we describe a survey of the unique characteristics of CelB by comparative analysis of the deduced primary, and determined tertiary and quaternary structure of CelB with that of homologous family I enzymes. Specific attention has been focused on the oligomerisation state of the mesophilic and thermostable members and to the detailed comparison of CelB with LacS as representatives of hyperthermophilic archaea. The contribution of ionic interactions to the thermostability of CelB was determined by the substitution and deletion of charged residues at the C-termini of CelB, and comparison of biochemical parameters of these mutants to the wild-type enzyme.

Materials and Methods

Construction of mutant *celB* genes

Wild-type CelB was produced in *E.coli* BL21(DE3) from the expression plasmid pLUW511, which contains the *celB* coding region fused to the start and stop codons of the ϕ10 gene on plasmid pET9d under control of the T7 promoter (Chapter 7). The translational fusion required the introduction of an additional triplet coding for an alanine immediately downstream of the start-codon. This alanine will be referred to as residue number 1a. Mutations were introduced by PCR amplification of the *celB* gene using *Pfu* polymerase (Stratagene). One universal primer (BG238 = 5'- GCGCGCCATGGCAAAGTTCCCAAAAAACTTC ATGTTTG), overlapping the startcodon of the *celB* gene (italic) and introducing an *Nco*I restriction site (underlined) was used for the construction of all mutated gene fragments. A second, mutagenic primer was used for each mutant, overlapping the stop-codon (italic) of the *celB* gene and introducing a *Bam*HI restriction site (underlined): BG298 for amino acid substitution R471A = 5'-CGCGCGGATCCCTACTTGGCCGTAACAAATTTGAGGTCTGCG, BG442 for K472A = 5'-GCGCGCGGATCCCTAGGCTCTTGTAACAAATTTGAGGTC, BG300 for K472Δ = 5'-CGCGCGGATCCCTATCTTGTAACAAATTTGAGG, BG443 for R471A/K472A = 5'-GCGCGCGGATCCCTAGGCGGCTGTAACAAAT TTGAGGTC, BG406 for R471Δ/K472Δ = 5'-GGCGGGGATCCCTATGTAACAA ATTTTAAATCTGCGAGGTG. Amplified gene fragments were digested with *Nco*I and *Bam*HI and subsequently ligated into the expression vector pET9d. BL21(DE3) was transformed with the ligation mixtures, plated onto selective medium and single colonies were picked. DNA was isolated using the Qiaprep spin plasmid kit (Qiagen, Westburg, The Netherlands) and the DNA sequence of the complete gene of each mutant was confirmed by DNA sequencing using infrared-labelled oligos (MWG, Germany) complementary to pET9d sequences immediately flanking the cloning site on the plasmid, and the Thermo Sequenase kit (Amersham Life Science) on a Li-Cor 4000L automated sequencer.

Enzyme purification

One litre selective TY medium in 2 litre baffled Erlenmeyer flasks, was inoculated with *E.coli* BL21(DE3) harbouring an expression plasmid containing either wild-type or mutated *celB* genes and cultured overnight at 37°C while vigorously shaking. Cell lysate was prepared and CelB was purified using heat-precipitation, anion exchange chromatography and gel filtration as described in Chapter 7 of this thesis. Pure CelB fractions were pooled and dialysed against 20 mM sodium phosphate buffer (pH 6.0). Protein concentrations were determined at 280 nm using an extinction coefficient for one subunit of $\varepsilon_m^{280nm} = 1.28*10^5\ M^{-1}cm^{-1}$ according to Gill & von Hippel (1989).

β-glucosidase activity assay

Standard activity measurements were performed in 0.5 ml 150 mM sodium citrate buffer (pH 4.8) containing 3 mM of the artificial substrate *p*-nitrophenol-β-D-glucopyranoside (pNp-Glu). Assay mixtures were preheated when necessary. Typically an aliquot (10 μl) of an appropriate enzyme dilution was added and the mixture was incubated for 10 min. Under these conditions

pNp-liberation was linearly related to enzyme concentration and no substrate depletion occured. After 10 min the reaction was stopped by addition of 1 ml 0.5 M sodium carbonate, raising the pH of the mixture to approximately 10 and increasing the $\varepsilon_m^{405\ nm}$ for the free pNp-molecule to 18,500 M^{-1} cm^{-1} at room temperature. The amount of free pNp was quantified at 405 nm on a Hitachi U-1100 spectrophotometer or, in the case of 100 µl volumes, in microtiter plates using a Thermomax microplate reader (Molecular Devices) at the same wavelength.

Thermal inactivation

The effect of the additional alanine, which was introduced in the enzyme for cloning purposes immediately downstream of the N-terminal methionine, was studied in heat-treated cell free extract of *E.coli* cultures harbouring either pLUW510 (coding for CelB without the alanine) or pLUW511 (CelB including the alanine) (Chapter 7). Heat-treated cell-free extract and purified wild-type and mutant CelB enzymes were diluted to 50 µg/ml in 20 mM sodium citrate buffer pH 4.8 (pH set at room temperature) and 100 µl samples were incubated at desired temperatures in an oil bath in glass vials closed with a screw cap with teflon inlay. Following incubation, samples were cooled down and remaining activity was determined relative to a sample kept at room temperature. Data were fitted according to first-order inactivation and half-life values were obtained from these fits. Dependance of the rate of inactivation on protein concentration was studied at 5, 25, 50, 100 and 500 µg/ml. Samples of 200 µg/ml were used for analysis on SDS and native PAGE. Lyotropic salts were added at 1M *I* ionic strength. Buffers for pH dependent stability were 20 mM sodium acetate from pH 3.74 to 5.8, 20 mM sodium phosphate from pH 6.2 to 8.2 and 20 mM sodium carbonate from 8.7 to 10.5. pH values are set at room temperature. Theoretical changes in the pH values of the phosphate and acetate buffers at 90°C do not exceed 0.1 unit, while for the carbonate buffers changes may amount up to 0.4 unit (Bates, 1970; Stoll, 1990).

Guanidinium chloride induced inactivation

Wild-type and mutant CelB enzymes were diluted to 50 µg/ml in 20 mM sodium citrate buffer (pH 4.8) containing increasing concentrations of guanidinium chloride (GdmCl) ranging from 0 to 2.5 M. After 24 hour incubation at room temperature, enzymes were 50-fold diluted into the assay mixture and remaining activity was determined. From plots of activity versus concentration denaturant, the concentration GdmCl was determined at which 50% of the enzyme was irreversibly inactivated. Samples of 200 µg/ml were incubated with different concentrations of GdmCl and used after 24 h for analysis on native PAGE.

Results

Stability of wild-type CelB

The routine expression plasmid for production of CelB (pLUW511) contains, in comparison with the originally reported gene sequence (Voorhorst *et al.*, 1995), three additional bases immediately downstream of the N-terminal methionine codon. These bases were introduced for cloning purposes and result in the presence of an additional alanine residue in the produced CelB enzyme (Chapter 7). In order to address the effect of this additional residue, the thermal inactivation of CelB expressed from pLUW510 (coding for CelB without the alanine; Voorhorst et *al.*, 1995) and pLUW511 was studied. Comparable rates for thermal inactivation of both enzymes at 106°C were found, with the half-life for inactivation of the CelB containing the alanine, being approximately 25% lower than that of the enzyme produced from pLUW510. This indicates a local, slightly destabilising effect of the alanine without the occurrence of large conformational changes and indicates that the CelB enzyme produced from pLUW511 (referred to as wild-type CelB) may be used to replace the enzyme produced from pLUW510 in studies of structure-function and structure-stability relations in CelB.

The half-life value of wild-type CelB was found to be 2.3 hours at 106°C, 11 hours at 101°C and 16.5 hours at 96°C. This stability was independent of protein concentration over the range 5-500 μg/ml. When the buffer was changed from 20 mM sodium citrate (pH 4.8) to 20 mM Tris (pH 8.0), a remarkable stabilisation was achieved, with a 7-fold increase of the half-life at 96°C. This stabilising effect of Tris was reported earlier and may reflect the *in vivo* stabilisation of CelB by organic compounds in the cytoplasm containing primary amine groups (Kengen, 1994). Alternatively, it may be related to the fact that Tris is a competitive inhibitor of CelB (Bauer & Kelly, 1998). Interaction between this organic compound and active site-residues may serve to stabilise

Table 1: Effect of the addition of 1 M *I* lyotropic salts on the half-life for thermal inactivation of wild-type CelB at 101°C.

lyotropic salt	$t_{1/2}$ (hours)
-	11
sulphate	9.6
acetate	23
chloride	0.1
nitrate	< 0.01

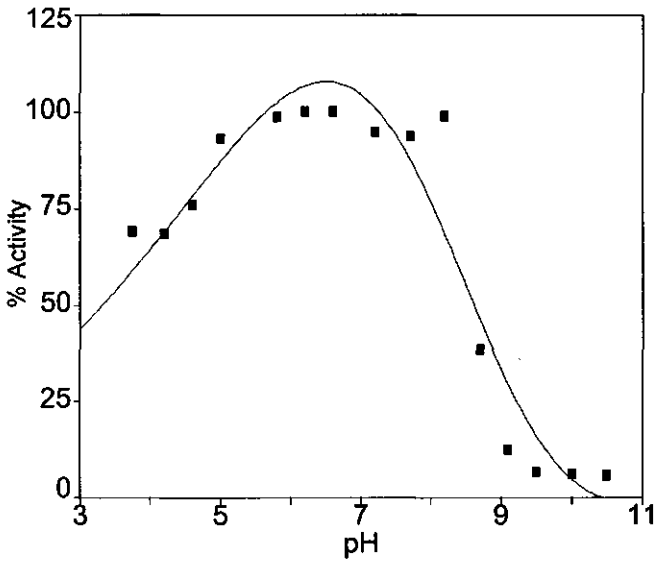


Figure 1: pH dependent thermostability of wild-type CelB. Residual activity after 1 hour incubation at 90°C was determined and plotted relative to the remaining activity at pH 6.

the enzyme against thermal inactivation. The addition of low concentrations of sodium chloride was also found to be stabilising. The enzyme was stabilised more than three-fold at 150 mM NaCl at 101°C. No effect of 0.2 M GdmCl on CelB stability was observed, while addition of 1 and 3 M GdmCl resulted in immediate and complete inactivation at 106°C.

The effect on CelB stability by addition of lyotropic salts more or less followed their position in the Hofmeister series of anions (Table 1). The addition of 1.0 M *I* (ionic strength equivalent) sodium sulphate or sodium acetate slightly increased or doubled, respectively, the half-life of CelB inactivation at 101°C. The addition of 1.0 M *I* sodium chloride decimated, and that of 1 M *I* sodium nitrate completely abolished activity after incubation of only a few minutes at 101°C.

Wild-type CelB was diluted to 50 µg/ml in 20 mM sodium acetate, sodium phosphate or sodium carbonate buffer with a pH between 3.7 and 10.5. At room temperature, CelB is completely stable at all pH values. After 1 h incubation at 90°C, remaining activity was determined (Figure 1). At the high pH- end of the stability curve, wild-type CelB inactivates around pH 8.5. At the low pH-end of the stability curves, activity diminishes below pH 5.

Three-dimensional structure analysis

A three-dimensional model of CelB has recently been reported, based on 3.3 Å X-ray diffraction data and using the crystal structure of LacS from *S.solfataricus* as search model in the molecular replacement procedure (Chapter 8). Because of the low resolution of the diffraction data, and the anticipated model bias towards LacS, the CelB model can not be used for detailed analysis of hydrogen bonds, ion-pairs and residue packing efficiency. These analyses are crucial in accessing the molecular basis for thermostability. However, the overall folding and the role of certain residues within larger structural features may be considered, as is detailed below.

A schematic alignment of CelB with several homologous family I glycosyl hydrolases revealed a clustering in two groups (Figure 2). CelB clusters with LacS as well as with the thermostable β-mannosidase from *P.furiosus* (Bauer *et al.*, 1996). The mannosidase is a tetrameric enzyme like CelB and LacS, and its half-life for thermal inactivation is in between that of the other two glycosidases (60 hours at 90°C). The second group in the alignment is formed by the β-glucosidases from the hyperthermophilic bacterium *Thermotoga maritima* and the mesophile *Bacillus polymyxa*, and the 6-P-β-galactosidase (LacG) from *Lactococcus lactis*. The enzyme from *T.maritima* is thermostable (half-life for thermal inactivation is approximately 6 hours at 95°C) and reported to be a dimer, while LacG is a monomer and the enzyme from *Bacillus* is an octamer (Gabelsberger *et al.*, 1993; Wiesmann *et al.*, 1995; Sanz-Aparicio *et al.*, 1998).

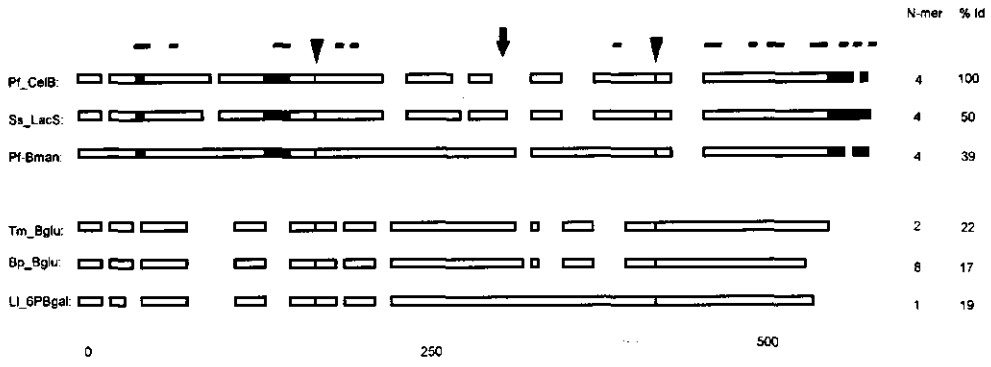


Figure 2: Schematic alignment of *P.furiosus* CelB with several other members of family 1 of glycosyl hydrolases, namely LacS from *S.solfataricus*, the β-mannosidase from *P.furiosus*, β-glucosidases from *T.maritima* and *B.polymyxa* and the 6-P-β-galactosidase from the mesophile *Lactococcus lactis*. The conserved glutamate residues (▼) that have been identified as the acid-base catalyst and the nucleophile are indicated. Regions that are involved in subunit interactions in CelB and LacS are overlined, with the three insertions indicated as filled parts of the sequences. The loop at the site of the barrel at the substrate entrance, which has been shortened in LacS and deleted in CelB, is indicated by an arrow. The state of oligomerisation and percentage amino acid identity with CelB is indicated.

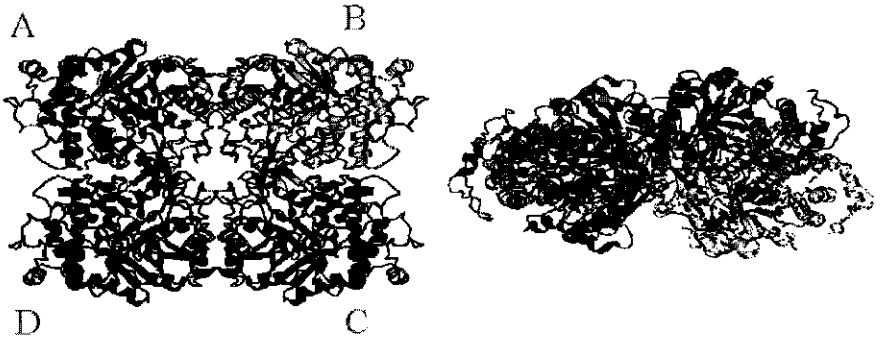


Figure 3: Overview of the tetrameric composition of CelB.

The quaternary conformation of CelB was found to be very similar to that from LacS and shows an arrangement of four identical subunits, termed A, B, C and D, in a slightly twisted square (Figure 3). No interactions between subunits A and C are present. The intersubunit surface between A and B is twice as large (1377 Å) as the one between A and D (814 Å) that forms a roof over the active site barrel (Kopp, 1998). Three of the regions that are involved in intersubunit contacts in LacS and CelB, are part of insertions which are not present in the less stable homologous enzymes (Fig 2). Upon formation of the tetramer, a little over 11% of the total solvent accessibility of each subunit is buried (Kopp, 1998).

Each CelB monomer is predicted to adopt a classical $(\beta\alpha)_8$ barrel fold with elaborate extensions at the top of the barrel (Fig 4a). The overall fold of the CelB monomer is identical to that of LacS with the exception of the deletion of one loop on the outside (see below) (Figure 4b and c). The active site is located inside the barrel and is connected to the outside via a radial channel. Access to the channel is generated by a kink in one of the helices lining the barrel.

CelB is also predicted to contain the large, solvent-filled, hydrophilic cavities that were reported for LacS and may play a major role in thermostabilisation by conferring resilience to the enzyme (Aguilar et *al.*, 1997). The residues that are located within 4 Å distance of the solvent molecules in the largest cavity, are completely conserved, with the only exception of serine at position 172 being an alanine in CelB. This is, however, one of the residues that contacts the solvent molecules with its backbone carbonyl group. The conservation of the involved residues is highest among the thermostable members of the family I glycosyl hydrolases and progressively diminishes in enzymes from less thermophilic and mesophilic species.

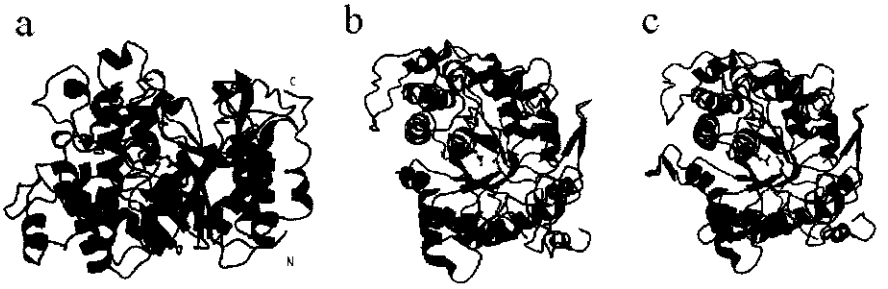


Figure 4: View of a CelB monomer from the side (a) and from the top (b), and view of a LacS monomer from the top (c). Active site glutamates are depicted in ball-and-stick representation. The orientation of the monomer in panel a is similar to that of subunit D in figure 3. The subunit in panel b is turned 90° across the x and z axis compared to panel a.

Entry to the active site is via a radial channel in the side of the $(\alpha\beta)_8$-barrel. Opening of this channel is generated by a kink in the helix of the fifth $\beta\alpha$-unit and by drastic shortening of a loop that is folded over this side of the barrel in the homologous enzymes that have their active site entrance at the top of the barrel. In LacS this loop is shortened to a small beta-strand, composed of two small beta-sheets, by residues 297-310 and is attached to the barrel by extensive ion-pair networks. One ion-pair network comprises Arg313-Asp248-Arg245-Asp309-Arg307-Asp104, the other network links Arg378-Asp291-Arg295-Glu297-Lys304. In CelB, the loop is completely deleted and consequently the ion-pair networks will not be present (Fig 2 and 4).

The C-termini in LacS join in the central part of the enzyme and have been crosslinked to their own and two neighbouring subunits by one large ion-pair network involving 12 residues: The C-terminal negatively charged carboxyl group of His489 interacts with the positively charged side chain of the same residue, and as well ion-pairs with Arg488, which in turn interacts with Glu345 from its own and an opposite subunit (Aguilar *et al.*, 1997). In CelB the C-terminus has been shortened by two residues, although the forelast residue is still an arginine and the last residue is still positively charged (lysine). The large ion-pair network that is crosslinking four subunits, however, seems not to be present. In contrast, the CelB structure indicates that the side chain of arginine 471 (488 in LacS) is pointing the other way and is now ion-pairing with Asp358 and Asp392 from an adjacent subunit (Figure 5). In this way, five-residue ion-pair networks crossing one subunit interface, seem to be formed.

A detailed analysis of ion-pairs and networks in the CelB structure is not possible due to the low resolution of the model. However, taken into consideration the described absence of several of the largest LacS networks, it is unlikely that CelB will contain a significantly higher number of large ion-pair networks than LacS. In contrast, it seems more likely that the extent of ion-pair formation in the most thermostable glycosidase is *less* than in its less stable counterpart. A further indication for a reduced number of large networks in CelB

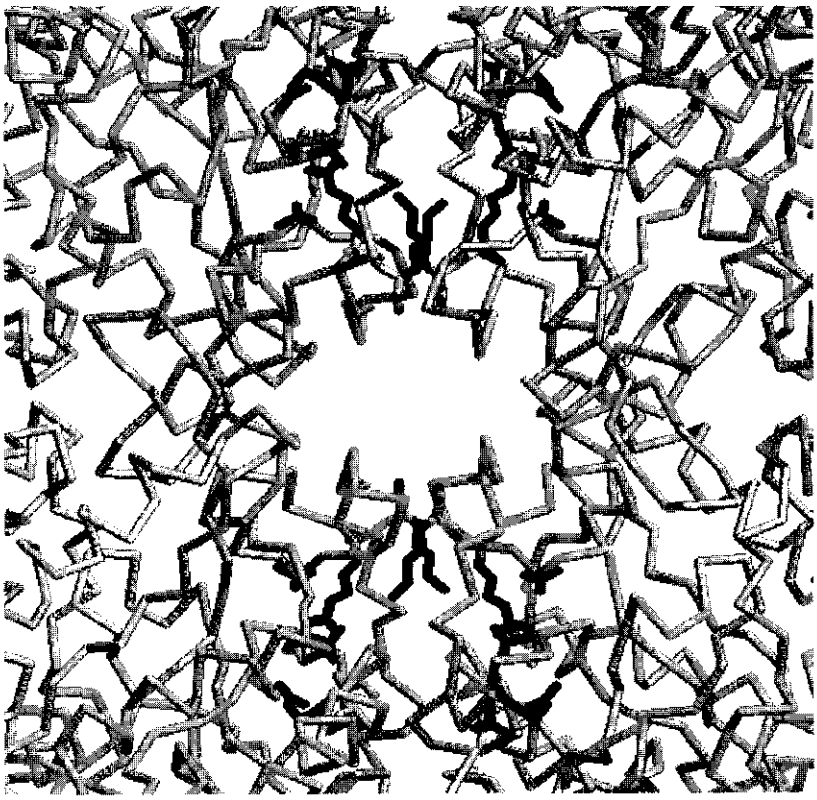


Figure 5: Proposed ionic interactions at the C-termini of the CelB monomers based on the three-dimensional CelB model (Chapter 8). Residues from two opposite monomers interact and in total four 5-residue ion-pair networks are formed.

is the reversal of the Arg/Lys ratio between this enzyme and LacS. Arginines are often found in large ion-pair networks because of their ability to form bidentate interactions. The Arg/Lys ratio in LacS is 1.4, while in CelB it is only 0.6, an observation that could reflect the higher number of large ion-pair networks in LacS.

Design and analysis of C-terminal CelB mutants

In order to study the role of the proposed C-terminal ion-pair interactions in CelB, we constructed a series of mutants with progressive deletion of parts of or the complete amino acid residues at the C-terminus; amino acid substitutions R471A, K472A and R471A/K472A remove the positive side chains, while substitutions K472Δ and R471Δ/K472Δ remove the complete amino acid residues at the indicated positions. Since the C-terminal backbone carboxyl group is negatively charged, deletions will also influence electrostatic interactions between this group and other residues in the enzyme, if present.

Expression of the mutated genes in E.coli resulted in high level production of a protein at the same position on an SDS-PAGE gel as the wild-

Table 2: Half-life values for thermal inactivation at 106°C for wild-type CelB and C-terminal mutants with standard error in minutes, and optimal temperature for catalysis. nd = not determined.

	Half-life values at 106°C (minutes)		optimum T for catalysis (°C)
wild-type	139	± 3.5	100-105
R471A	91	± 14	98
K472A	43	± 9.4	90
K472Δ	58	± 4.8	94
R471A/K472A	37	± 7.4	85
R471Δ/K472Δ	137 at 90°C	± nd	80

type CelB (not shown). Wild-type and mutant CelB enzymes were purified from *E.coli* and biochemically characterised. Half-life values for thermal inactivation at 106°C were determined and revealed a progressive decrease of thermostability with the severeness of the respective mutations (Table 2). The effect of the complete deletion of two C-terminal residues even forced us to lower the incubation temperature with 15°C to be able to determine an accurate half-life value. Furthermore, the optimal temperature for catalytic activity was determined, which followed the same trend as found for half-life of thermal inactivation (Table 2). Reactivation curves after prolonged incubation of wild type CelB and mutant enzymes R471A and R471Δ/K472Δ in guanidinium chloride were determined (Figure 6). Mutants R471A and R471Δ/K472Δ reactivate less efficient than the wild-type CelB. Concentrations at which 50% of the enzymes is irreversibly inactivated are 1.6, 1.5 and 1.3 M GdmCl for wild-type CelB, R471A and R471Δ/K472Δ, respectively.

Correlation between enzyme activity and presence of tetrameric conformation

Wild-type CelB and mutants R471A and R471Δ/K472Δ were incubated at 200 μg/ml at 106°C for different periods of time. Mutant R471Δ/K472Δ was furthermore incubated at 90°C. Remaining activity was determined relative to a sample kept at room temperature. Aliquots of the controls and the incubated samples were run on a native polyacrylamide gel (Figure 7). Progressive incubation of wild-type and mutant R471A CelB enzymes leads to inactivation of the enzyme and to a lower intensity of the tetramer band on native PAGE. Mutant R471Δ/K472Δ is not stable at 106°C and tetrameric enzyme can only be detected in the sample kept at room temperature. Loss of tetramer of this mutant can be followed at 90° (Figure 7). Also for guanidinium dependent reactivation,

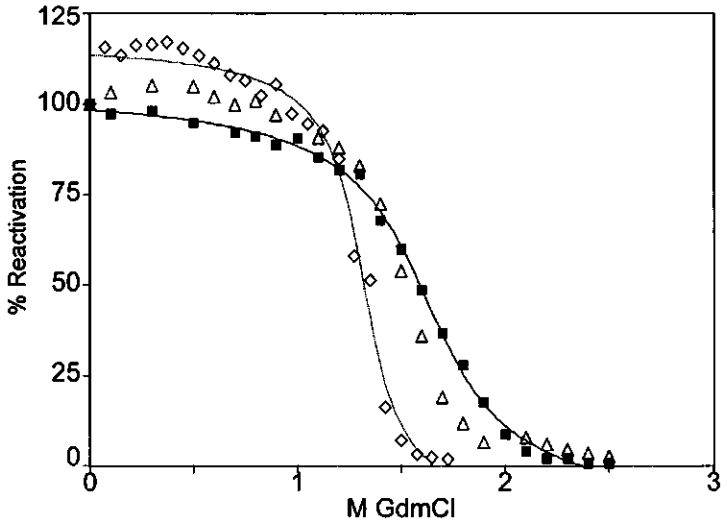


Figure 6: Reversibility of inactivation after 24 hours incubation in increasing concentrations of guanidinium chloride. ■ = wild-type CelB, Δ = R471A, ◊ = R471Δ/K472Δ

a correlation between percentage of reactivation and the presence of CelB tetramers on native PAGE was found (results not shown).

Discussion and conclusions

In order to obtain more insight into the molecular basis for the extreme thermostability of CelB from *P.furiosus*, we analysed its three-dimensional model and biochemically characterised the wild-type enzyme as well as several CelB mutants containing deletions of positively charged amino acids at the C-terminal end that are likely involved in ion-pair interactions.

Inactivation of CelB at 106°C seems not to be caused by aggregation. Aggregation is a multimolecular event that would not result in the observed first-order kinetics and the independence of the kinetics of inactivation on protein concentration. Furthermore, no stabilisation was achieved by addition of GdmCl to the incubation mixtures. Glutamate dehydrogenase from *Sulfolobus solfataricus* is a thermostable enzyme which inactivates due to aggregation, and this enzyme can be stabilised by the addition of GdmCl in pre-transition region concentrations (Consalvi *et al.*, 1993). The mechanism of this stabilisation is suggested to be the shielding of hydrophobic patches of reversible denatured

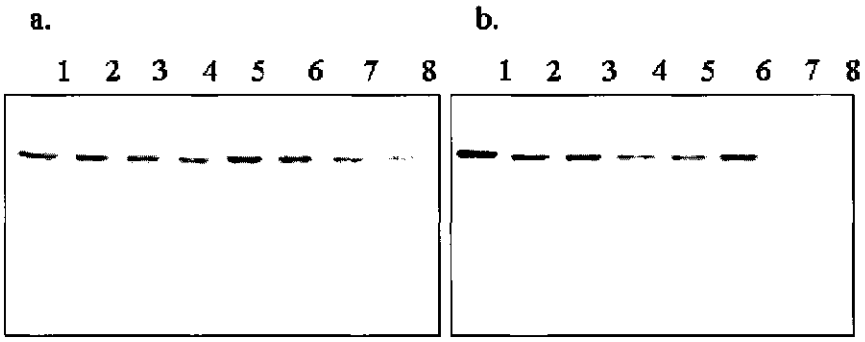


Figure 7: Native polyacrylamide gels with 15 microliter samples of 200 microgram/ml wild-type CelB and mutants that have been incubated for different periods of time at 106 and 90°C. Panel a, lane 1-4 and 5-8: wild-type CelB and mutant R471A incubated at 106°C for 0, 0.8, 2 and 4 h, respectively. Panel b, lane 1: wild-type CelB kept at room temperature, lane 2-5: R471Δ/K472Δ incubated at 90°C for 0, 1, 2 and 3 hours, lane 6-8, R471Δ/K472Δ incubated at 106°C for 0, 0.8 and 1.3 hours.

enzyme by the denaturant, which would otherwise interact and result in irreversible aggregation.

A correlation was found between β-glucosidase activity and the presence of tetrameric CelB on native PAGE. Loss of quaternary structure is probably followed by immediate unfolding of the individual subunits since there is no indication for completely folded monomers on the native gels. For the mutant CelB enzymes a correlation between remaining activity and tetramer presence on native gels was also observed, indicating that removal of ion-pair interactions at the C-terminus results in accelerated subunit dissociation. These results suggest that loss of quaternary structure may be responsible for thermal inactivation of CelB. However, an apparent correlation may also be the result of subunit dissociation, occurring immediately after an unidentified inactivating event. In this case the loss of tetrameric conformation is only indicative of but not the reason for, the observed inactivation of CelB.

Lyotropic salts have different effects on the stability of enzymes. Sulphate and acetate ions are cosmotropic salts and stabilise hydrophobic interactions by preferentially being excluded from the native protein surface (Arakawa, 1990). Nitrate is a chaotropic anion that decreases the strength of hydrophobic interactions. Chloride ions have neither of these effects (Hatefi, 1969; von Hippel, 1969). All ions at high concentrations have the ability to interfere with ion-pair interactions, decreasing the strength of these interactions by charge shielding. The net effect of lyotropic salts on enzyme stability will therefore be the sum of the effect on hydrophobic interactions and ion-pair interactions. The presented results show that wild-type CelB is strongly stabilised by acetate ions and that sulphate ions result in a more or less unchanged half-life for inactivation. Sodium chloride at 1 M *I* is strongly destabilising, indicating that CelB is indeed stabilised by ion-pair interactions. A comparative study towards

the stability of CelB and LacS confirmed this accelerated inactivation in the presence of high salt concentrations and showed that LacS is more sensitive than CelB (Pouwels *et al.*, 1999). Nitrate combines the effects of charge shielding and weakening of hydrophobic interactions and inactivates the enzyme at 101°C completely. Remarkable is the effect of sodium chloride at low concentrations where charge shielding is not or only partially taking place. Up to 150-200 mM NaCl, CelB is stabilised by this compound. Moreover, CelB activity is stimulated by these amounts of NaCl (Kengen *et al.*, 1993).

Further evidence for the stabilisation of CelB by ionic interactions is the pH dependent resistance towards thermal inactivation. Upon increasing pH, protons are abstracted from positively charged groups, resulting in the breakage of ion-pairs and salt-bridges and unfolding of the enzyme. Decreasing pH results in protonation of negatively charged groups and concomitant loss of electrostatic interactions, and hence implies less resistance to unfolding. The increased loss of CelB activity above pH 8 and below pH 5 suggests a role for charged groups in maintaining an active enzyme structure at high temperatures. The lower sensitivity of CelB to pH induced and sodium chloride induced inactivation in comparison with LacS suggest different stabilisation strategies in both archaeal enzymes. While LacS seems to be mainly stabilised by ion-pair networks, these features may play a less prominent role in CelB. Because CelB is a much more thermostable enzyme than LacS, it must have been stabilised by other interactions like improved hydrophobic interactions, hydrogen bond formation or optimised packing efficiency.

There seems to be no correlation between family I glycosyl hydrolase thermostability and their state of oligomerisation (Aguilar *et al.*, 1997). It is, however, interesting to note that the variability in oligomerisation of the two enzymes from mesophiles and the one from *T.maritima* (1-mer, 8-mer, 2-mer) is achieved without large deletions and insertions between the amino acid sequences (Fig. 2). Amino acid substitutions seem to have been sufficient to induce oligo-(or mono-) merisation. In fact, also some tetra- and octameric species of the *T.maritima* β-glucosidase were detected using size-exclusion chromatography, indicating that there is equilibrium between the different oligomerisation states (Gabelsberger *et al.*, 1993). In contrast, many of the subunit interfaces in LacS and CelB are formed between residues of which some or all are part of insertions (Fig. 1). This means that in these enzymes, tetramerisation has been achieved after extensive structural adaptations, in contrast to the enzymes from mesophilic sources. While there is no correlation between oligomerisation and thermal stability, the mechanism by which oligomerisation has been achieved in the most thermostable enzymes differs greatly from that in mesophiles and, hence, may be an important adaptation towards high temperatures.

All mutant CelB enzymes inactivate faster at 106°C than the wild-type enzyme, have lower temperature optima for activity, and do not reactivate as efficient as the wild-type CelB after prolonged incubation in guanidinium chloride. The loss of activity is paralleled by a loss of tetrameric conformation as visualised by native PAGE for some of the mutants. This indicates that both the C-terminal arginine and lysine residues are involved in interactions that contribute to the thermal and other stability of CelB. These interactions involve

the side chains of these residues, because both mutants R471A and K472A CelB are less stable than the wild-type CelB. However, not only the side chains, but also the backbone atoms of these residues are involved in stabilising interactions. Complete deletion of these residues, as in R471Δ/K472Δ CelB, results in an even faster thermal inactivation. The fact that R471Δ/K472Δ CelB already inactivates at 90°C, while wild-type CelB and the other mutants survive significantly at 106°C, indicates that the effect of this C-terminal deletion may be more than local.

In conclusion, we suggest that CelB is stabilised in relation to homologues from mesophilic sources by oligomerisation and the formation of ion-pair networks. The fact that CelB is more thermostable than LacS but probably does not contain a higher number of large ion-pair networks, most likely means that CelB must be stabilised to a high extent by hydrophobic interactions or more efficient hydrogen bonding. However, confirmation of this hypothesis awaits the availability of a high-resolution structure for CelB and a more detailed biochemical and structural analysis of wild-type and mutant CelB en LacS enzymes.

Acknowledgements

We are very grateful to Prof. Georg Schulz and Jürgen Kopp for the collaboration on the determination of the three-dimensional model for CelB.

References

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the beta-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Arakawa, T., Bhat, R., Timasheff, S.N. (1990). Why preferential hydration does not always stabilise the native structure of globular proteins. *Biochemistry* **29**, 1924-1931.

Bates, R. G., Paabo, M. (1970). Measurement of pH. *In: Handbook of Biochemistry, Selected data for Molecular Biology,* (H.A.Sober, Ed.) 2nd edition, The Chemical Rubber Company, Cleveland, Ohio.

Bauer, M. W., Bylina, E. J., Swanson, R. V. & Kelly, R. M. (1996). Comparison of a beta-glucosidase and a beta-mannosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Purification, characterization, gene cloning, and sequence analysis. *J. Biol. Chem.* **271**, 23749-23755.

Bauer, M. W. & Kelly, R. M. (1998). The family 1 beta-glucosidases from *Pyrococcus furiosus* and *Agrobacterium faecalis* share a common catalytic mechanism. *Biochemistry* **37**, 17170-17178.

Consalvi, V., Chiaraluce, R., Politi, L., Pasquo, A., De Rosa, M. & Scandurra, R. (1993). Glutamate dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*: studies on thermal and guanidine-dependent inactivation. *Biochim. Biophys. Acta.* **1202**, 207-215.

Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859.

Elcock, A. H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489-502.

Gabelsberger, J., Liebl, W., Schleifer, K.-H. (1993). Purification and properties of recombinant beta-glucosidase of the hyperthermophilic bacterium *Thermotoga maritima*. *Appl. Microbiol. Biotechnol.* **40**, 44-52.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.

Hatefi, Y., Hanstein, W.G. (1969). Solubilisation of particulate proteins and nonelectrolytes by chaotropic agents. *Proc. Natl. Acad. Sci. U.S.A.* **62**, 1129-36.

Henrissat, B. & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637-644.

Jaenicke, R. & Bohm, G. (1998). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738-748.

Kengen, S. W., Luesink, E. J., Stams, A. J. & Zehnder, A. J. (1993). Purification and characterisation of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**, 305-312.

Kengen, S. W. M., Stams, A.J.M. (1994). An extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*; a comparison with other glycosidases. *Biocatalysis* **11**, 79-88.

Kopp, J. (1998). PhD Thesis University Freiburg, Germany.

Ladenstein, R., Antranikian, G. (1998). Proteins from hyperthermophiles: Stability and enzymatic catalysis close to the boiling point of water. *In: Biotechnology of Extremophiles, Adv. Biochem. Engng/Biotechnol*, Ed G. Antranikian, Springer Verlag, Heidelberg, Germany, Vol **61**, p37-86.

Moracci, M., Ciaramella, M., Cobucci-Ponzano, B., Perugino, G., Rossi, M. (1998). The β-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *International Conference on Thermophiles 1998*, Brest, France.

Moracci, M., Capalbo, L., Ciaramella, M. & Rossi, M. (1996). Identification of two glutamic acid residues essential for catalysis in the beta-glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Prot. Engng.* **9**, 1191-1195.

Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676-683.

Pisani, F. M., Rella, R., Raia, C. A., Rozzo, C., Nucci, R., Gambacorta, A., De Rosa, M. & Rossi, M. (1990). Thermostable beta-galactosidase from the archaebacterium *Sulfolobus solfataricus*. Purification and properties. *Eur. J. Biochem.* **187**, 321-328.

Pouwels, J., Moracci, M., Cobucci-Ponzano, B., Perugino, G., van der Oost, J., Kaper, T., Lebbink, J., de Vos, W.m., Ciaramella, M. and Rossi, M. (1999). Activity and stability of hyperthermophilic enzymes: a comparative study on two archaeal β-glycosidases. *Submitted for publication.*

Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). Crystal structure of beta-glucosidase A from *Bacillus polymyxa*: Insights into the catalytic activity in family 1 glycosyl hydrolases. *J. Mol. Biol.* **275**, 491-502.

Spassov, V. Z., Karshikoff, A. D. & Ladenstein, R. (1995). The optimisation of protein-solvent interactions: thermostability and the role of hydrophobic and electrostatic interactions. *Prot. Sci.* **4**, 1516-1527.

Stoll, V. S., Blanchard, J.S. (1990). *Meth. Enzymol.* **182**, 24-30.

Tomschy, A., Bohm, G. & Jaenicke, R. (1994). The effect of ion pairs on the thermal stability of D-glyceraldehyde 3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *Prot. Engng.* 7, 1471-1478.

Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K. S. P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. (1998). Protein thermostability above 100°C: A key role for ionic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12300-12305.

von Hippel, P. H., Schleich, T. (1969). *in: Structure and stability of biological macromolecules* (Timasheff S.N. & Fasman, G.D. Eds.) Marcel Dekker New York., 417-574.

Voorhorst, W. G., Eggen, R. I., Luesink, E. J. & de Vos, W. M. (1995). Characterisation of the celB gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *J. Bacteriol.* **177**, 7105-7111.

Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). The three-dimensional structure of 6-phospho-beta-galactosidase from *Lactococcus lactis*. *Structure* 3, 961-968.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V. & et al. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3**, 1147-1158.

Chapter 11

Summary and concluding remarks

Summary and concluding remarks

Hyperthermophilic organisms are able to survive and reproduce optimally between 80°C and 113°C. Most of them belong to the domain of the Archaea, although several hyperthermophilic Bacteria have been described. One of the major questions regarding hyperthermophiles concerns the molecular mechanisms that determine the extreme stability of their macromolecules. In particular, enzymes on the one hand should be flexible in order to efficiently perform catalysis, while on the other hand they should be sufficiently rigid in order to prevent thermal unfolding and inactivation. In other words, they have to be thermostable as well as thermoactive. Therefore, the research described in this thesis has been focused on analysing the molecular determinants of enzyme thermostability and thermoactivity.

In the first chapter of this thesis, hyperthermophilic microorganisms are introduced, and the origin of the hyperthermophilic phenotype is discussed in the light of established and emerging phylogenetic, biochemical and structural biological considerations. Furthermore, thermodynamic and kinetic protein stability is explained, and an overview is given of the strategies that have been suggested to confer thermostability to enzymes from hyperthermophiles, by analysing amino acid sequences, homology-based structural models, and three-dimensional structures. One of the proposed general strategies for achieving hyperthermostability relates to the number of ion-pairs and their organisation in extensive networks in hyperthermostable enzymes. While biochemical studies have shown that isolated ion-pairs at room temperature do not contribute significantly to protein stability, several theoretical considerations indicate that arrangements of ion-pairs into networks at high temperatures may play an important role in protein stabilisation. In contrast to the need for maximum stability at high temperatures, it is the requirement for flexibility that enables an enzyme to efficiently perform its catalytic function. The mechanisms that determine hyperthermostability on the one hand, hyperthermoactivity on the other hand, and the interplay between these two characteristics, have been investigated in the research described in this thesis, using three different approaches of protein engineering, namely (i) domain swapping, (ii) site-directed mutagenesis, and (iii) random mutagenesis and directed evolution.

In order to study the molecular determinants of enzyme thermostability and thermoactivity, we decided to include representatives of both hyperthermophilic Archaea and Bacteria in the analysis. *Pyrococcus furiosus* was selected as hyperthermophilic Archaeon, since it optimally grows at 100°C (Fiala & Stetter, 1986). Moreover, *P.furiosus* is very well studied with regard to its physiology, biochemistry and genomics. As representative of the hyperthermophilic Bacteria we choose *Thermotoga maritima* that grows optimally at 80°C (Huber *et al.*, 1986). Both organisms are able to ferment mixtures of polypeptides and a variety of oligosaccharides. One of the enzymes that was chosen as model, is β-glucosidase, a glycosyl hydrolase involved in the hydrolysis of disaccharides and oligosaccharides into the corresponding monomeric sugars that are further metabolised. The second model enzyme is glutamate dehydrogenase (GDH), a key enzyme in a pathway that is involved in the disposal of reducing equivalents that are generated during sugar

fermentation. In both *P.furiosus* and *T.maritima* these two key enzymes are expected to perform similar functions.

GDH is extensively introduced in Chapter 2 and is a useful model enzyme since it is well studied, catalyses an important reaction, and has been isolated and characterised from many organisms in all three kingdoms of life, ranging from psychrophiles to hyperthermophiles. Multiple amino acid sequences are available, as well as high resolution three-dimensional structures from GDH of various sources including the hyperthermophiles *P.furiosus* and *T.maritima* (Yip *et al.*, 1995; Knapp *et al.*, 1997). Furthermore, the *gdh* genes from these hyperthermophiles can be easily overexpressed in *Escherichia coli*, allowing efficient enzyme purification and mutagenesis (Chapter 3; Kort *et al.*, 1997). GDH is in general a hexameric enzyme, composed of six identical subunits that are arranged in two trimers that are stacked upside down on top of each other (Baker *et al.*, 1992). Each subunit is composed of a substrate and a nicotinamide cofactor binding domain, linked by a flexible hinge region and separated by a deep cleft in which the active site is located. During catalysis, the hinge region mediates the opening and closing of the cleft by a rotation of the cofactor-domain with respect to the substrate-binding domain. In this way an adequate hydrophobic environment for hydride transfer during catalysis is created (Stillman *et al.*, 1993).

Chapter 3 describes the functional expression of the *gdh* gene from *P.furiosus* in *E.coli*. The *P.furiosus* GDH amounted to 20 % of total *E.coli* cell protein, with the majority of the expressed protein in the hexameric conformation. Following activation by a heat-treatment, the GDH that could be purified from *E.coli* was indistinguishable from that purified from *P.furiosus*. The role of the GDH substrate and cofactor binding domains in conferring thermoactivity and thermostability were studied by exchanging them between the GDHs of the hyperthermophilic archaeon *P.furiosus* and the mesophilic bacterium *C.difficile*. Hybrid genes were constructed and successfully expressed in *E.coli*. One of the resulting hybrid proteins, containing the glutamate-binding domain of the *C.difficile* and the cofactor-binding domain of the *P.furiosus* enzyme, did not show detectable activity. In contrast, the complementary hybrid, containing the *P.furiosus* glutamate and the *C.difficile* cofactor binding domain was found to be a catalytically active hexamer that showed a reduced substrate affinity but maintained efficient cofactor binding with the specificity found in the *C.difficile* enzyme. Compared to the *C.difficile* GDH, the latter archaeal-bacterial hybrid is slightly more thermoactive, less thermostable but much more stable towards guanidinium chloride-induced inactivation and denaturation. These results indicate (i) that the cofactor-binding domain is structurally independent from the substrate-binding domain with respect to affinity for the coenzyme, and (ii) that thermal and chemical stability may be uncoupled.

The elucidation of the three-dimensional structure of GDH from *P.furiosus* and its comparison with the homologous GDH from the mesophilic bacterium *C.symbiosum*, has highlighted the formation of extensive ion-pair networks at domain and subunit interfaces as a possible explanation for the superior thermostability of the archaeal enzyme (Yip et al., 1995). In order to assess whether this is a more general stabilising strategy which is not only employed in GDH from *P.furiosus*, but also in enzymes from other

hyperthermophilic Archaea and Bacteria, an homology-based modelling study was carried out using GDH amino acid sequences derived from species spanning a wide spectrum of optimal growth temperatures. Chapter 4 describes the observed correlation between the amount and extent of ion-pair networks and the thermal stability of the studied enzymes, which is consistent with a role for the involvement of such networks in the adaptation of enzymes to extreme temperatures.

In order to analyse the role of ion-pair networks in more detail, we decided to rationally engineer the networks that are found in the GDH from *P.furiosus*, into the GDH from *T.maritima*, in which most of the networks are largely reduced or absent and which is much less thermostable than the pyrococcal GDH (Knapp et al., 1997). In the study that is described in Chapter 5, we focused on the flexible hinge region that is connecting the two domains. While in GDH from *P.furiosus* a five-residue ion-pair network is present, this is reduced to two single ion-pairs in *T.maritima* GDH. Using a site-directed mutagenesis approach, the missing charged residues were introduced into *T.maritima* GDH. The resulting mutant GDHs as well as the wild-type enzyme were overproduced in *Escherichia coli* and subsequently purified. Elucidation of the three-dimensional structure of the double mutant N97D/G376K at 3.0 Å, showed that the designed ion-pair interactions were indeed formed. Moreover, because of interactions with an additional charged residue, a six-residue network is present in this double mutant. Melting temperatures of the mutant enzymes N97D, G376K and N97D/G376K, as determined by differential scanning calorimetry, did not differ significantly from that of the wild-type enzyme. Identical transition midpoints in guanidinium chloride-induced denaturation experiments were found for the wild-type and all mutant enzymes. Thermal inactivation at 85°C occured more than two-fold faster for all mutant enzymes than for the wild-type GDH. At temperatures of 65°C and higher, the wild-type and the three mutant enzymes showed identical specific activities. However, at 58°C the specific activity of N97D/G376K and G376K was found to be significantly higher than that of the wild-type and N97D enzymes. These results suggest that the engineered ion-pair interactions in the hinge region do not affect the stability towards temperature- or guanidinium chloride-induced denaturation but rather affect the specific activity of the enzyme and the temperature at which it functions optimally.

Chapter 6 focuses on the role of an 18-residue ion-pair network, which is present at the subunit interface of the GDH from *P.furiosus*. This network has been studied by introducing four new charged amino acid residues into the subunit interface of GDH from *T.maritima*. Amino acid substitutions were introduced as single mutations as well as in several combinations. Elucidation of the crystal structure of the quadruple mutant S128R/T158E/N117R/S160E *T.maritima* GDH showed that all anticipated ion-pairs were formed and that a 16-residue ion-pair network was present. Enlargement of existing networks by single amino acid substitutions unexpectedly resulted in a decrease in resistance towards thermal inactivation and thermal denaturation. However, combination of destabilising single mutations in most cases restored stability, indicating the need for balanced charges at subunit interfaces and high cooperativity between the different members of the network. Combination of the three destabilising

mutations in triple mutant S128R/T158E/N117R resulted in an enzyme with a 30 minutes longer half-life of inactivation at 85°C, a 3°C higher temperature optimum for catalysis, and a 0.5°C higher apparent melting temperature than that of wild type GDH. These findings confirm the hypothesis that large ion-pair networks do indeed stabilise enzymes from hyperthermophilic organisms.

Regarding subunit interfaces in GDH, a more or less complete picture of the role of ion-pair networks has emerged by now. The modelling study that involved GDHs derived from different sources spanning the complete temperature spectrum, showed a progressive increase in formation of networks at the subunit interfaces with increasing temperature. The removal of central ion-pairs from the large 18-residue ion-pair network in *P. kodakaraensis* GDH has been reported to reduce the half-life for thermal inactivation (Rahman *et al.*, 1998). Furthermore, there are now multiple examples of reconstituted large networks in less thermostable GDHs that significantly increase half-life values for thermal inactivation and apparent melting temperatures (Rahman *et al.*, 1998; Vetriani *et al.*, 1998; Chapter 6). Ion-pair networks therefore play an important role in the thermostabilisation and maintenance of a correct structure of this enzyme. However, except for the studies described in this thesis (Chapter 5 & 6), three-dimensional structures of mutated GDHs have not been determined. This implies that only for the charged residues introduced into the subunit interface and hinge region of *T.maritima* GDH it has actually been proven that the predicted networks are indeed formed. The role of ion-pairs in the flexible hinge region of GDH is less well resolved. These networks are apparently involved in determining thermoactivity and kinetic parameters, which is consistent with the hinge region being involved in positioning of the domains with regard to each other during catalysis. However, the introduction of a 6-residue ion-pair network in the hinge of *T.maritima* GDH did not increase the thermostability of this enzyme. This result may reflect the considerable differences between the archaeal and the bacterial enzyme, which share 55% amino acid identity. This consideration is supported by the fact that (i) in *Thermococcus litoralis* GDH, which has a much higher homology (87% identity) with *P.furiosus* GDH, a restored ion-pair network was found to be stabilising only in combination with a second site mutation (Vetriani *et al.*, 1998), and (ii) stabilisation of the *T.maritima* GDH subunit interface is only observed after the combination of multiple, destabilising single mutations (Chapter 6). Apparently, different strategies for stabilisation have evolved within the hyperthermostable members of this protein family. While the subunit interfaces of the most thermostable, archaeal members are highly charged, the subunit interface of the *T.maritima* GDH is optimised by hydrophobic interactions (Knapp et al., 1997). In spite of this, we have shown that it is possible to successfully introduce an archaeal feature into a bacterial enzyme that is stabilised by a different mechanism. The electrostatic optimisation of the subunit interfaces in the archaeal GDHs, on the one hand, and the hydrophobic optimisation of the bacterial GDH subunit interface on the other hand, is not a universal division. Remarkably, exactly the opposite situation has been observed in the superoxide dismutase family, which shows electrostatic optimisation of the subunit interfaces in the enzyme from the bacterium *Aquifex pyrophilus*, and hydrophobic optimisation in the enzyme of the archaeon *Sulfolobus acidocaldarius* (Lim *et al.*, 1997; Knapp *et al.*, 1999).

The thesis continues with an overview of the current knowledge on the well-studied second model enzyme, the β-glucosidase CelB from *P.furiosus* (Chapter 7). CelB is involved in the hydrolysis of β-(1,4)-linked disaccharides during growth on sugars, and is highly thermostable and thermoactive, with a half-life of thermal inactivation of 85 hours at 100°C and an optimum temperature for catalysis of 102-105°C (Kengen *et al.*, 1993). Heterologous expression and enzyme purification procedures have been developed that allow for routine site-directed mutagenesis and *in vitro* applications (Voorhorst *et al.*, 1995 and Chapter 7).

Directed evolution is a potent approach to evolve a desired property into a biocatalyst, as well as a promising method to study structure-function and structure-stability relationships. So far, this approach was restricted to model enzymes from mesophilic and thermophilic sources. Chapter 7 describes the development of a directed evolutionary approach for the thermostable CelB by random mutagenesis, *in vitro* recombination and a rapid screening procedure. The constructed library contains over 6000 random CelB clones with an average of 1 or 2 amino acid substitutions per enzyme. The ability of CelB to hydrolyse chromogenic glucose- and galactose derivatives was used to isolate mutants with a changed ratio in the hydrolysis of these enantiomeric sugars. One of the isolated mutants was characterised in detail and found to contain a single amino acid substitution of asparagine to serine at position 206 in the active site. In the wild-type enzyme this asparagine forms a hydrogen bond with the hydroxyl group of the substrate at the C2 position. Biochemical analysis of the mutant using different substrates provided insight in the active-site architecture of CelB, and indicated that the interactions between CelB and its substrates are similar to those observed in three-dimensional structures of mesophilic homologues complexed with substrate analogues.

Chapter 8 describes a three-dimensional model of CelB and the validation of this model by engineering substrate specificity. The model is based on 3.3 Å X-ray diffraction data and on the tetrameric β-glycosidase LacS from *Sulfolobus solfataricus* and the monomeric 6-phospho-β-galactosidase from *Lactococcus lactis* (LacG) as search models (Aguilar *et al.*, 1997; Wiesmann *et al.*, 1995). CelB shows high structural homology with LacS and contains subunits that adopt the common $(\alpha\beta)_8$ barrel motif and are arranged in a slightly twisted tetramer. The difference in substrate specificity between β-glycosidases and 6-phospho-β-glycosidases seems to be determined by three residues in the active site that are conserved within, but not between, both types of glycosidases. This hypothesis was tested by replacing these residues in CelB with the corresponding substitutes, as present in LacG. Mutant CelB containing all three substitutions showed a reduced catalytic activity on non-phosphorylated sugars and increased efficiency on 6-phospho-β-galactose. These results indicate that the phosphate-binding site is functional at elevated temperatures.

Low-temperature activity and substrate specificity of CelB was studied in Chapter 9 using the random CelB library described in Chapter 7. Screening of this library at room temperature on the artificial substrate pNp-glucose resulted in the isolation and characterisation of mutants with up to three-fold increased rates of hydrolysis of this aryl-glucoside. Amino acid substitutions were identified in the active site region, at subunit interfaces, at the enzyme surface, and buried

in the interior of the monomers. Characterisation of the mutants revealed that the increase in low-temperature activity was achieved in different ways, including altered substrate specificity and increased flexibility by an apparent overall destabilisation of the enzyme. Kinetic characterisation of active site mutants showed that in all cases the catalytic efficiency at 20°C on *p*-nitrophenol-β-D-glucose as well as on the disaccharide cellobiose, was up to two-fold increased. In most cases, this was achieved at the expense of β-galactosidase activity at 20°C, and of total catalytic efficiency at 90°C. Substrate specificity is affected by many of the observed amino acid substitutions, only some of which are located in the vicinity of the active site. The largest change in the ratio of *p*-nitrophenol-β-D-glucopyranoside/*p*-nitrophenol-β-D-galactopyranoside hydrolysis is a 7.5-fold increase, which was observed in N415S CelB. This asparagine at position 415 interacts with active site residues that stabilise the hydroxyl group at the C4 position of the substrate, the conformation of which is equatorial in glucose-containing and axial in galactose-containing substrates. In conclusion, this random screening approach using a hyperthermostable enzyme resulted in the identification of residues that are critical in determining thermostability, low-temperature activity, and substrate recognition. It was shown that low-temperature activity may be engineered into a hyperthermostable enzyme without affecting its extreme stability. Several mutants display evolved properties that could enable them to perform more efficiently in industrial or biotechnological applications than the wild-type CelB. It is feasible that N415S CelB with its increased catalytic efficiency on cellobiose but unchanged stability, or T371A CelB with its much better affinity for cellobiose, could be applied instead of wild-type CelB in biosensors for the detection of cellobiose at low temperatures. Not only is directed evolution a potent approach to evolve a desired property into a biocatalyst of potential industrial interest, but, as is shown here, it may enhance our understanding of the mechanisms determining enzyme catalysis, stability, and substrate recognition at physiological or extreme conditions.

In Chapter 10 the molecular basis for the extreme thermostability of CelB has been addressed by comparative analysis of the three-dimensional model of CelB and the crystal structure of the homologous β-glycosidase LacS from *S.solfataricus*. Furthermore, this chapter describes the biochemical characterisation of wild-type CelB and mutants thereof, containing substitutions or deletions of charged residues at the C-terminus, which may be involved in the formation of ion-pair networks. All mutant CelB enzymes show decreased optimal temperatures for activity and increased rates of thermal inactivation, indicating that ion-pairs at the C-terminus of CelB play a role in determining the extreme thermostability of this enzyme. Moreover, increased thermal inactivation of CelB in the presence of sodium chloride and at extreme pHs, supports the role of electrostatic interactions in this enzyme. The CelB structure indicates that part of CelB stability is derived from structural adaptations that allow oligomerisation and from the presence of solvent-filled, hydrophilic cavities, which have earlier been identified in LacS. However, structural comparison of several large ion-pair networks that are present in LacS, reveals that a five-residue and a six-residue membered network is not present in CelB

and a large 16-membered network at the subunit interface, extensively cross-linking the subunits, has been fragmented.

The picture emerging from the structural and biochemical characterisation of wild-type, site-directed, and random mutants of CelB, is that, compared to homologous enzymes from mesophilic sources, active site architecture, substrate recognition, acid-base catalysis and transition state stabilisation have been preserved, while thermoactivity and thermostability have been adapted to the high-temperature conditions in which the enzyme has to function optimally.

An evaluation of the obtained results and conclusions against the original aims as set in the introduction of this thesis reveals the following. With the GDH model we have established and validated a strategy for enzyme stabilisation that seems to be employed by many different hyperthermophiles. Furthermore, we have shown that these features can be successfully engineered into less thermostable enzymes, in order to raise their thermostability. In addition, we have increased our knowledge of enzyme catalysis and substrate recognition at high temperatures and in relation with low-temperature catalysis, using site-directed mutagenesis and directed evolution of the hyperthermostable β-glucosidase. The directed evolutionary approach, in particular, has been shown to be a potent technique which will in the near future rapidly increase our knowledge of enzyme characteristics and may be employed to increase the biotechnological potential of the enzymes under study.

References

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the beta-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W. & Stillman, T. J. (1992). Subunit assembly and active site location in the structure of glutamate dehydrogenase. *Proteins* **12**, 75-86.

Fiala, G. & Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* **145**, 56-61.

Huber, R., Langworthy, T. A., König, H., Thomm, M., Woese, C. W., Sleytr, U. B. & Stetter, K. O. (1986). *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch. Microbiol.* **144**, 324-333.

Kengen, S. W., Luesink, E. J., Stams, A. J. & Zehnder, A. J. (1993). Purification and characterisation of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**, 305-312.

Knapp, S., de Vos, W. M., Rice, D. & Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3.0 Å resolution. *J. Mol. Biol.* **267**, 916-932.

Knapp, S., Kardinahl, S., Hellgren, N., Tibbelin, G., Schafer, G. & Ladenstein, R. (1999). Refined crystal structure of a superoxide dismutase from the hyperthermophilic archaeon *Sulfolobus acidocaldarius* at 2.2 angstrom resolution. *J. Mol. Biol.* **285**, 689-702.

Kort, R., Liebl, W., Labedan, B., Forterre, P., Eggen, R. I. L. & de Vos, W. M. (1997). Glutamate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: molecular characterisation and phylogenetic implications. *Extremophiles* **1**, 52-60.

Lim, J. H., Yu, Y. G., Han, Y. S., Cho, S., Ahn, B. Y., Kim, S. H. & Cho, Y. (1997). The crystal structure of an Fe-superoxide dismutase from the hyperthermophile *Aquifex pyrophilus* at 1.9 Å resolution: structural basis for thermostability. *J. Mol. Biol.* **270**, 259-274.

Rahman, R., Fujiwara, S., Nakamura, H., Takagi, M. & Imanaka, T. (1998). Ion pairs involved in maintaining a thermostable structure of glutamate dehydrogenase from a hyperthermophilic archaeon. *Biochem. Biophys. Res. Comm.* **248**, 920-926.

Stillman, T. J., Baker, P. J., Britton, K. L. & Rice, D. W. (1993). Conformational flexibility in glutamate dehydrogenase. Role of water in substrate recognition and catalysis. *J. Mol. Biol.* **234**, 1131-1139.

Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K. S. P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. (1998). *Protein thermostability above 100°C: A key role for ionic* interactions. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12300-12305.

Voorhorst, W. G., Eggen, R. I., Luesink, E. J. & de Vos, W. M. (1995). Characterisation of the celB gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *J. Bacteriol.* **177**, 7105-7111.

Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). The three-dimensional structure of 6-phospho-beta-galactosidase from *Lactococcus lactis*. *Structure* **3**, 961-968.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V. & et al. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3**, 1147-1158.

Samenvatting

Hyperthermofiele micro-organismen zijn staat te overleven en optimaal te groeien tussen 80 en 113°C. De meeste hyperthermofielen zijn Archaea, hoewel eveneens verscheidene hyperthermofielen tot het domein van de Bacteria behoren. Een belangrijke vraag op het gebied van hyperthermofielen, is welke moleculaire mechanismen de extreme stabiliteit van hun macromoleculen bepalen. Hun enzymen moeten aan de ene kant voldoende flexibel zijn om hun catalytische functie te kunnen uitoefenen, terwijl ze aan de andere kant voldoende rigide moeten zijn om inactivatie en ontvouwing te kunnen weerstaan. Met andere woorden, enzymen van hyperthermofielen moeten zowel thermostabiel als thermoactief zijn. Het onderzoek dat beschreven wordt in dit proefschrift, is gericht op de analyse van de moleculaire mechanismen van enzym thermostabiliteit en -activiteit.

In het eerste hoofdstuk van dit proefschrift worden de hyperthermofiele micro-organismen geïntroduceerd en wordt de oorsprong van het hyperthermofiele fenotype bediscussieerd in het licht van bestaande en nieuwe fylogenetische, biochemische en structuurbiologische argumenten. Daarnaast worden de principes van thermodynamische en kinetische eiwitstabiliteit uitgelegd, alsmede de mechanismen die mogelijk verantwoordelijk zijn voor de hoge stabiliteit van enzymen uit hyperthermofiele organismen, en die afgeleid zijn uit analyse van aminozuursequenties, homologie-gebaseerde structuurmodellen en drie-dimensionale kristalstructuren. Eén van de mogelijk algemene strategieën voor het verkrijgen van thermostabiliteit, betreft het aantal ionparen en hun organisatie in grote netwerken in hyperthermostabiele enzymen. Terwijl biochemische studies hebben aangetoond dat geïsoleerde ionparen bij kamertemperatuur niet significant bijdragen aan eiwitstabiliteit, suggereren verscheidene theoretische argumenten dat de organisatie van ionparen in netwerken bij hoge temperatuur een belangrijke rol zouden kunnen spelen in eiwitstabilisatie. Recht tegenover de vraag naar stabiliteit, of rigiditeit, bij hoge temperatuur, staat de eis van flexibiliteit waardoor het mogelijk is voor een enzym om catalytisch actief te zijn. De mechanismen die stabiliteit aan de ene kant en activiteit aan de andere kant bepalen, alsmede de wisselwerking tussen deze twee eigenschappen, zijn onderzocht in dit proefschrift door middel van drie verschillende benaderingen van 'protein engineering' (eiwitaanpassing), namelijk (i) domeinuitwisseling, (ii) plaatsgerichte mutagenese, en (iii) random mutagenese en geforceerde evolutie.

Om de moleculaire mechanismen van thermostabiliteit en thermoactiviteit in enzymen te kunnen bestuderen, zijn zowel hyperthermofiele Bacteria als Archaea onderzocht. *Pyrococcus furiosus* is geselecteerd als hyperthermofiel archaeon, omdat dit organisme optimaal groeit bij 100°C (Fiala & Stetter, 1986). Bovendien zijn de fysiologie, biochemie en genetica van *P.furiosus* zeer goed bestudeerd. Als representant voor de Bacteria is *Thermotoga maritima* gekozen, een bacterie die optimaal groeit bij 80°C (Huber *et al.*, 1986). Beide organismen zijn in staat mengsels van eiwitten en suikers te fermenteren. Eén van de enzymen die als model zijn gekozen in dit proefschrift is β-glucosidase, een glycosyl hydrolase dat disacchariden en oligosacchariden hydrolyseert tot de overeenkomstige monomere suikers die verder

gemetaboliseerd worden. Het tweede modelenzym is glutamaat dehydrogenase (GDH), dat een sleutelenzym is in een metabole route, betrokken bij de verwijdering van overtollige reductie-equivalenten die gegenereerd zijn gedurende de suikerfermentatie. In zowel *P.furiosus* als *T.maritima* wordt aan beide enzymen dezelfde functie toebedacht.

GDH wordt uitgebreid geïntroduceerd in hoofdstuk 2 van dit proefschrift en is een geschikt modelenzym aangezien het goed bestudeerd is, een belangrijke metabole reactie katalyseert, en is geïsoleerd en gekarakteriseerd in veel organismen in alle drie domeinen van het leven, variërend van psychrofielen tot hyperthermofielen. Veel aminozuursequenties zijn beschikbaar, alsmede verschillende hoog-resolutie structuren van GDH waaronder de hyperthermofielen *P.furiosus en T.maritima* (Yip *et al.*, 1995; Knapp *et al.*, 1997). Verder komt het *gdh* gen gemakkelijk tot expressie in *Escherichia coli*, wat een efficiente eiwitzuivering en mutagenese mogelijk maakt (Kort *et al.*, 1997; hoofdstuk 3). GDH is in het algemeen een hexameer enzym, bestaande uit zes identieke subeenheden die twee trimeren vormen (Baker *et al.*, 1992). Elke monomeer bestaat uit een substraatbindend en een cofactorbindend domein, verbonden door middel van een flexibele scharnier (hinge) en gescheiden door een diepe groef waarin het actieve centrum zich bevindt. Gedurende de katalyse reguleert de hinge regio het openen en sluiten van de groef door middel van een rotatie van het cofactorbindend domein ten opzichte van het substraatbindend domein. Op deze manier wordt een adequate hydrofobe omgeving gecreëerd voor de protonoverdracht gedurende de katalyse (Stillman et *al.*, 1993).

Hoofdstuk 3 beschrijft de functionele expressie van het *gdh* gen van *P.furiosus* in *E.coli*. Het *P.furiosus* GDH omvatte 20% van het totale *E.coli* eiwitgehalte, met de meerderheid daarvan in de hexamere conformatie. Na hitte-incubatie werd een GDH gezuiverd dat niet te onderscheiden was van het GDH dat uit *P.furiosus* is geïsoleerd. De rol van de substraat- en cofactorbindende domeinen in het bepalen van thermoactiviteit en thermostabiliteit werd bestudeerd door het uitwisselen hiervan tussen de GDHs van de hyperthermofiele archaeon *P.furiosus* en de mesofiele bacterie *C.difficile*. Hybride genen werden geconstrueerd en tot expressie gebracht in *E.coli*. Een van de resulterende hybride eiwitten, met het glutamaatbindend domein van *C.difficile* GDH en het cofactorbindend domein van *P.furiosus* GDH, had geen detecteerbare katalytische activiteit. Het complementaire hybride enzym, met het *P.furiosus* glutamaatbindend domein en het *C.difficile* cofactorbindend domein, was een katalytisch actieve hexameer met gereduceerde substraataffiniteit waarin een efficiente cofactorbinding behouden was met een specificiteit zoals die van het *C.difficile* enzym. Vergeleken met het *C.difficile* GDH is de laatstgenoemde archaële-bacteriële hybride thermoactiever, minder thermostabiel maar veel resistenter tegen guanidiniumchloride geïnduceerde inactivatie en denaturatie. Deze resultaten geven aan dat (i) het cofactorbindend domein structureel onafhankelijk is van het substraatbindend domein met betrekking tot affiniteit voor de cofactor, en dat (ii) thermostabiliteit en chemische stabiliteit ontkoppeld kunnen worden.

De opheldering van de driedimensionale structuur van GDH van *P.furiosus* en de vergelijking hiervan met de homologe structuur van de

mesofiele *C.symbiosum* heeft de aanwezigheid van grote netwerken van ionparen op domein- en subeenheid contactvlakken (subunit interfaces) als mogelijke verklaring aangewezen voor de superieure thermostabiliteit van het archaële enzym (Yip et al., 1995). Om te bepalen of dit een meer algemene strategie is, die niet alleen wordt aangewend door *P.furiosus* maar ook in enzymen van andere hyperthermofiele Archaea en Bacteria, is een op homologie gebaseerde modelleringsstudie uitgevoerd met GDH aminozuursequenties uit verschillende soorten afkomstig uit een wijd spectrum van optimale groeitemperaturen. Hoofdstuk 4 beschrijft de gevonden correlatie tussen de hoeveelheid en de vorming van netwerken van ionparen en de thermostabiliteit van de bestudeerde enzymen, die consistent is met een rol van deze netwerken in de adaptatie van enzymen aan extreme temperaturen.

Om de rol van ion-paren in meer detail te bestuderen is gepoogd de netwerken, zoals ze aanwezig zijn in het GDH van *P.furiosus*, te introduceren in het GDH van *T.maritima*, waarin de meeste netwerken veel kleiner of zelfs afwezig zijn en dat veel minder thermostabiel is dan het enzym uit *P.furiosus* (Knapp et al., 1997). In de studie die beschreven wordt in hoofdstuk 5, is de flexibele hinge regio die beide domeinen verbindt, bestudeerd. Terwijl in *P.furiosus* GDH een 5-residuen ionpaar netwerk aanwezig is, is dit gereduceerd tot twee geïsoleerde ionparen in *T.maritima* GDH. Gebruik makend van plaatsgerichte mutagenese zijn de ontbrekende residuen geïntroduceerd in het *T.maritima* GDH. Zowel de resulterende mutante GDHs als het wild-type enzym werden overgeproduceerd in *E.coli* en vervolgens gezuiverd. Opheldering van de kristalstructuur van de dubbelmutant N97D/G376K, liet zien dat de ontworpen ionpaar-interacties inderdaad gevormd zijn. Vanwege interacties met een extra geladen residue is zelfs een 6-residuen netwerk gevormd in deze dubbelmutant. Smelttemperaturen van de mutante enzymen N97D, G376K en N97D/G376K, bepaald met behulp van differentiële scanning microcalorimetrie, verschilden niet significant van die van het wildtype enzym. Identieke transitiemiddelpunten in guanidiniumchloride geïnduceerde ontvouwingsexperimenten werden gevonden voor het wild-type en alle mutante enzymen. Thermische inactivatie bij 85°C vond twee keer sneller plaats voor de mutanten dan in het wild-type GDH. Bij temperaturen van 65°C en hoger, hadden het wild-type en de mutanten identieke specifieke activiteiten. Echter, bij 58°C is de specifieke activiteit van N97D/G376K en G376K significant hoger dan dat van N97D en het wild-type GDH. Deze resultaten suggereren dat de ontworpen ionpaar interacties in de hinge regio geen invloed hebben op de stabiliteit versus temperatuur- of guanidinium-geïnduceerde ontvouwing maar eerder effect hebben op de specifieke activiteit van het enzym en de temperatuur waarbij het optimaal functioneert.

Hoofdstuk 6 behandelt de rol van een 18-residuen ionpaar netwerk dat aanwezig is op de subunit interface van het GDH van *P.furiosus*. Dit netwerk is bestudeerd door vier ontbrekende geladen aminozuren te introduceren in de subunit interface van het GDH van *T.maritima*. Aminozuursubstituties werden geintroduceerd als enkelvoudige mutaties, alsmede in verschillende combinaties. Opheldering van de kristalstructuur van de viervoudige mutant S128R/T158E/N117R/S160E *T.maritima* GDH, liet zien dat alle geanticipeerde ionparen inderdaad gevormd zijn en dat een 16-residuen ionpaar netwerk

aanwezig is. Uitbreiding van bestaande kleine netwerken door toevoeging van enkelvoudige mutaties resulteerde in een verlaging van de resistentie tegen thermische inactivatie en denaturatie. Combinatie van destabiliserende enkelvoudige mutaties herstelde in de meeste gevallen de stabiliteit, wat aangeeft dat gebalanceerde ladingen belangrijk zijn en dat er grote coöperativiteit is tussen verschillende delen van het netwerk. Combinatie van destabiliserende mutaties in de drievoudige mutant S128R/T158E/N117R resulteerde in een enzym met een 30 minuten langere halfwaardetijd voor inactivatie bij 85°C, een 3°C hoger temperatuuroptimum voor katalyse, en een 0.5°C hogere smelttemperatuur dan die van het wildtype GDH. Deze resultaten bevestigen de hypothese dat grote ionpaar netwerken inderdaad een stabiliserende invloed hebben op enzymen uit hyperthermofiele organismen.

Met betrekking tot subunit interfaces in GDH is nu een min of meer compleet beeld ontstaan. De modelleringsstudie met de GDHs uit soorten afkomstig uit het complete temperatuurspectrum, liet een progressieve trend in de vorming van netwerken van ionparen met stijgende temperatuur zien. Het verwijderen van centrale ionparen uit het grote 18-residue netwerk in *P.kodakaraensis* GDH reduceerde de halfwaardetijd voor thermische inactivatie (Rahman *et al.*, 1998). Verder zijn er meerdere voorbeelden van gereconstitueerde grote netwerken in minder thermostabiele GDHs met significant hogere halfwaardetijden voor thermische inactivatie en schijnbare smelttemperaturen (Rahman et al., 1998; Vetriani *et al.*, 1998; hoofdstuk 6). Netwerken van ionparen leveren dus een belangrijke bijdrage aan de stabilisatie en het behoud van een correcte structuur van dit enzym. Echter, met uitzondering van de studies beschreven in dit proefschrift (hoofdstuk 5& 6), zijn er geen kristalstructuren bepaald voor de gemuteerde enzymen, wat betekent dat alleen voor de geladen residuen die geïntroduceerd zijn in de subunit interface en de hinge regio van *T.maritima* GDH werkelijk is aangetoond dat de voorspelde netwerken inderdaad gevormd zijn. De rol van ionparen in de flexibele hinge regio van GDH is minder duidelijk. Deze ionparen zijn blijkbaar betrokken bij het bepalen van de thermoactiviteit en kinetische parameters, wat overeenkomt met de rol van de hinge regio in de positionering van de domeinen ten opzichte van elkaar gedurende katalyse. Echter, de introductie van een 6-residuen netwerk van ionparen in de hinge regio van *T.maritima* GDH verhoogde niet de thermostabiliteit van dit enzym. Dit resultaat kan een afspiegeling zijn van aanzienlijke verschillen tussen het bacteriële en het archaële enzym, die 55% identiek zijn op aminozuurniveau. Deze overweging wordt versterkt door het feit dat (i) in *Thermococcus litoralis* GDH, dat veel homologer (87% identiteit) is met het *P.furiosus* GDH, slechts in combinatie met een tweede mutatie een hersteld netwerk stabiliserend was (Vetriani et al., 1998), en (ii) dat stabilisatie van de *T.maritima* subunit interface alleen bereikt werd door combinatie van destabiliserende enkelvoudige mutaties (hoofdstuk 6). Blijkbaar zijn verschillende strategieën voor eiwitstabilisatie geëvolueerd binnen de GDH familie. Terwijl de subunit interfaces van de meest stabiele, archaële leden geladen zijn, is de subunit interface van het *T.maritima* GDH geoptimaliseerd door middel van hydrofobe interacties (Knapp et al., 1997). Desondanks is aangetoond dat het mogelijk is een archaële eigenschap te introduceren in een bacteriëel enzym dat met behulp van een andere strategie

gestabiliseerd is. De electrostatische optimalisatie van de subunit interfaces in de archaële GDHs aan de ene kant, en de hydrofobe optimalisatie van het bacteriële GDH aan de andere kant, is geen algemeen onderscheid. Integendeel, precies de tegenovergestelde situatie is waargenomen in de superoxide dismutase familie, dat electrostatische optimalisatie laat zien van de subunit interfaces van het enzym uit de bacterie *Aquifex pyrophilus* en hydrofobe optimalisatie in het enzym van het archaeon *Sulfolobus acidocaldarius* (Lim et *al.*, 1997; Knapp *et al.*, 1999).

Het proefschrift vervolgt met een overzicht van de bestaande kennis over het goed bestudeerde tweede modelenzym, het β-glucosidase CelB van *P.furiosus* (Hoofdstuk 7). CelB hydrolyseert β-(1,4)-gekoppelde disacchariden gedurende groei van *P.furiosus* op suikers, en is buitengewoon thermostabiel en thermoactief, met een halfwaardetijd van 85 uur bij 100°C en een optimum temperatuur voor katalyse van 102-105°C (Kengen et *al.*, 1993). Verbeterde heterologe expressie en enzymzuiveringsprotocollen zijn beschreven, wat routinematige plaatsgerichte mutagenese en *in vitro* toepassingen mogelijk maakt (Voorhorst *et al.*, 1995; hoofdstuk 7).

Geforceerde evolutie is een krachtige benadering om een gewenste eigenschap te introduceren in een industrieel relevante biokatalysator, alsmede een veelbelovende methode om structuur-functie en structuur-stabiliteitsrelaties te bestuderen. Tot zover is deze benadering beperkt gebleven tot modelenzymen uit mesofiele en thermofiele organismen. Hoofdstuk 7 beschrijft de ontwikkeling van een geforceerde evolutiebenadering voor het thermostabiele CelB door middel van random mutagenese, *in vitro* recombinatie en snelle screeningsmethoden. De geconstrueerde bank bevat meer dan 6000 random CelB varianten met een gemiddelde van 1 tot 2 aminozuurveranderingen per enzym. De mogelijkheid van CelB om chromogene glucose- en galactose substraten te splitsen, is gebruikt om random mutanten te isoleren met een veranderde verhouding in hydrolyse van deze enantiomere suikers. Eén van de geisoleerde mutanten is in detail gekarakteriseerd en bevatte een enkele aminozuurverandering van asparagine naar serine op positie 206 in het actieve centrum van het enzym. In het wild-type CelB vormt deze asparagine een waterstofbrug met de hydroxylgroep van het substraat op de C2 positie. Biochemische analyse van de mutant met verschillende substraten verleende inzicht in de architectuur van het actieve centrum van CelB, en toonde aan dat de interacties tussen CelB en zijn substraten hetzelfde zijn als geobserveerd in de kristalstructuren van mesofiele homologe glycosidases gecomplexeerd met substraat analogen.

Hoofdstuk 8 beschrijft een driedimensionaal model van het *P.furiosus* CelB en de validering hiervan door middel van het veranderen van de substraatspecificiteit. Het model is gebaseerd op 3.3 Å X-ray diffractie data en het tetramere β-glycosidase LacS uit *Sulfolobus solfataricus* en het monomere 6-fosfo-β-galactosidase van *Lactococcus lactis* (LacG) als zoekmodellen (Aguilar *et al.*, 1997; Wiesmann et *al.*, 1995). CelB vertoont hoge structurele homologie met LacS en bevat subeenheden die het bekende $(\alpha\beta)_8$ barrel motief vormen. Het verschil in substraatspecificiteit tussen β-glycosidases en 6-fosfo-β-glycosidases lijkt te worden bepaald door drie residuen in het actieve centrum die geconserveerd zijn binnen, maar niet tussen, de beide typen glycosidasen. Deze

hypothese is getest door de drie residuen in CelB te vervangen voor de corresponderende LacG residuen. Mutant CelB dat alle drie aminozuursubstituties bevat, had een lagere katalytische activiteit op niet-gefosforyleerde substraten dan wildtype CelB en hogere efficientie op 6-fosfo-β-galactose. Deze resultaten suggereren dat de geïntroduceerde fosfaatbindingsplaats functioneel is bij hoge temperatuur.

Activiteit bij lage temperatuur en substraatspecificiteit van CelB is bestudeerd in hoofdstuk 9, door gebruik te maken van de random CelB bank die beschreven is in hoofdstuk 7. Screening van deze bank bij kamertemperatuur met het kunstmatige substraat pNp-glucose, resulteerde in de isolatie en karakterisatie van mutanten met tot drie keer hogere hydrolysesnelheden op dit arylglucoside. Aminozuurveranderingen werden geïdentificeerd in of vlakbij het actieve centrum, op subunit contactvlakken, aan het enzymoppervlak en binnen in de monomeren. Karakterisatie van de mutanten liet zien dat de verhoging in lage-temperatuursactiviteit bereikt is op verschillende manieren, waaronder veranderende substraatspecificiteit en verhoogde flexibiliteit door middel van een schijnbare destabilisatie van het eiwit. Kinetische karakterisatie van actief centrum mutanten liet zien dat in alle gevallen de catalytische efficientie bij 20°C zowel op *p*-nitro-β-D-glucose als op de disaccharide cellobiose, tot twee keer toe toegenomen is. In de meeste gevallen is dit bereikt ten koste van β-galactosidase activiteit bij 20°C en totale activiteit bij 90°C. Substraatspecificiteit wordt beïnvloed door veel van de geobserveerde aminozuursubstituties, waarvan slechts enkelen in de buurt van het actieve centrum gelokaliseerd zijn. De grootste verandering in de verhouding van *p*-nitrofenol-β-D-glucopyranoside/*p*-nitrofenol-β-D-galactopyranoside hydrolyse is een 7,5-voudige verhoging, die veroorzaakt werd door aminozuursubstitutie asparagine 415 naar serine (N415S). De asparagine op positie 415 gaat een interactie aan met de actieve-centrumresiduen die de hydroxylgroep op de C4-positie van het substraat stabiliseren, waarvan de conformatie equatoriaal is in glucose-bevattende substraten en axiaal in galactose-bevattende substraten. Concluderend heeft deze random-screening benaderingsmethode geresulteerd in de identificatie van residuen die van belang zijn voor de bepaling van de thermostabiliteit, lage temperatuursactiviteit en substraatherkenning. Lage-temperatuursactiviteit kan verhoogd worden in een hyperthermostabiel enzym zonder de thermostabiliteit aan te tasten. Verscheidene mutanten laten geëvolueerde eigenschappen zien die interessant zouden kunnen zijn voor industriële of biotechnologische toepassingen. Mutant N415S CelB met een hogere efficientie op cellobiose en onveranderde stabiliteit, of mutant T371A CelB met een hogere affiniteit voor cellobiose, zouden in plaats van het wild-type CelB toegepast kunnen worden in biosensoren die gebruikt kunnen worden om cellobiose aan te tonen bij lage temperatuur. Geforceerde evolutie is niet alleen een krachtige benadering om een gewenste eigenschap in een biokatalysator van potentieel industrieel belang te evolueren, maar kan, zoals hier is laten zien, ons begrip van de mechanismen die enzymcatalyse, -stabiliteit en substraatherkenning bepalen, vergroten.

In hoofdstuk 10 wordt de moleculaire basis voor de thermostabiliteit van CelB onderzocht door middel van vergelijkbare analyse van het driedimensionale model van CelB en de kristalstructuur van het homologe β-glycosidase van *S.solfataricus*. Verder beschrijft dit hoofdstuk de biochemische

karakterisatie van wildtype CelB en mutanten hiervan met substituties of deleties van geladen residuen aan de C-terminus, die ionparen kunnen vormen. Alle CelB mutanten hebben lagere optimumtemperaturen voor katalyse en inactiveren sneller dan het wildtype CelB, wat aantoont dat ionparen aan de C-terminus van CelB een rol spelen in het bepalen van de extreme thermostabiliteit van dit eiwit. Het CelB model geeft aan dat een deel van de stabiliteit afgeleid wordt van structurele adaptaties die oligomerisatie mogelijk maken en door de aanwezigheid van watergevulde, hydrofiele gaten, die eerder geïdentificeerd waren in LacS. Echter, vergelijking van verscheidenen grote netwerken die aanwezig zijn in LacS, laat zien dat een 5 en een 6 residuen network niet aanwezig zijn in CelB en een groot 16-residuen netwerk op de subunit interface, dat subunits verbindt, gefragmenteerd is.

Het beeld dat de structurele en biochemische karakterisatie van wildtype, plaatsgerichte en random mutanten CelB oplevert, is dat in vergelijking met mesofiele homologen, actief-centrum architectuur, substraatherkenning, zuur-base catalyse en overgangstoestand-stabilisatie behouden zijn, terwijl thermoactiviteit en thermostabiliteit aangepast zijn aan de hoge-temperatuurscondities waarin het enzym optimaal moet functioneren.

Een evaluatie van de verkregen resultaten en conclusies ten opzichte van de originele doelen zoals beschreven is in de inleiding van dit proefschrift, laat het volgende zien. Met het GDH model is een algemene strategie voor enzymstabilisering bij hoge temperatuur uitgewerkt en gevalideerd. Eveneens is gevonden dat de stabiliserende netwerken van ionparen succesvol in minder stabiele enzymen geïntroduceerd kunnen worden om de thermostabiliteit te verhogen. Tenslotte is onze kennis van enzymcatalyse en substraatherkenning bij hoge temperatuur in relatie met laag-temperatuur catalyse vergroot, gebruik makend van plaatsgerichte en geforceerde evolutie van het hyperthermostabiele β-glucosidase. Met name de geforceerde evolutiebenadering is een krachtige techniek die in hoog tempo de kennis over enzymeigenschappen zal uitbreiden en eveneens gebruikt zal worden om de biotechnologische potentieel van bestudeerde enzyme uit te breiden.

Referenties

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the beta-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.

Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W. & Stillman, T. J. (1992). Subunit assembly and active site location in the structure of glutamate dehydrogenase. *Proteins* **12**, 75-86.

Fiala, G. & Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* **145**, 56-61.

Huber, R., Langworthy, T. A., König, H., Thomm, M., Woese, C. W., Sleytr, U. B. & Stetter, K. O. (1986). *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch. Microbiol.* **144**, 324-333.

Kengen, S. W., Luesink, E. J., Stams, A. J. & Zehnder, A. J. (1993). Purification and characterisation of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**, 305-312.

Knapp, S., de Vos, W. M., Rice, D. & Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3.0 Å resolution. *J. Mol. Biol.* **267**, 916-932.

Knapp, S., Kardinahl, S., Hellgren, N., Tibbelin, G., Schafer, G. & Ladenstein, R. (1999). Refined crystal structure of a superoxide dismutase from the hyperthermophilic archaeon *Sulfolobus acidocaldarius* at 2.2 angstrom resolution. *J. Mol. Biol.* **285**, 689-702.

Kort, R., Liebl, W., Labedan, B., Forterre, P., Eggen, R. I. L. & de Vos, W. M. (1997). Glutamate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: molecular characterisation and phylogenetic implications. *Extremophiles* **1**, 52-60.

Lim, J. H., Yu, Y. G., Han, Y. S., Cho, S., Ahn, B. Y., Kim, S. H. & Cho, Y. (1997). The crystal structure of an Fe-superoxide dismutase from the hyperthermophile *Aquifex pyrophilus* at 1.9 Å resolution: structural basis for thermostability. *J. Mol. Biol.* **270**, 259-274.

Rahman, R., Fujiwara, S., Nakamura, H., Takagi, M. & Imanaka, T. (1998). Ion pairs involved in maintaining a thermostable structure of glutamate dehydrogenase from a hyperthermophilic archaeon. *Biochem. Biophys. Res. Comm.* **248**, 920-926.

Stillman, T. J., Baker, P. J., Britton, K. L. & Rice, D. W. (1993). Conformational flexibility in glutamate dehydrogenase. Role of water in substrate recognition and catalysis. *J. Mol. Biol.* **234**, 1131-1139.

Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K. S. P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. (1998). Protein thermostability above 100°C: A key role for ionic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12300-12305.

Voorhorst, W. G., Eggen, R. I., Luesink, E. J. & de Vos, W. M. (1995). Characterisation of the celB gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *J. Bacteriol.* **177**, 7105-7111.

Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). The three-dimensional structure of 6-phospho-beta-galactosidase from *Lactococcus lactis*. *Structure* **3**, 961-968.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V. & et al. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3**, 1147-1158.

Curriculum vitae

Joyce Helen Geraldine Lebbink werd geboren op 6 oktober 1969 te Steenderen. Na een verhuizing richting het westen van het land, doorkroop en –liep zij de peuterklas, kleuterschool en basisschool in Hoogmade (1972-1982). Vervolgens koos zij exact en slaagde in 1988 voor het VWO examen op het St. Bonaventuracollege te Leiden. Het fundamentele en multidisciplinaire karakter van de studie Moleculaire Wetenschappen aan de Universiteit Wageningen sprak haar zeer aan en vol goede moed stortte Joyce zich in het roerige studentenleven. De propedeuse werd binnen een jaar afgesloten met een positief resultaat maar eveneens met de nodige vraagtekens over het verdere verloop van haar carriere. Joyce besloot haar wetenschappelijke vorming enige tijd in te vriezen en zich te gaan bemoeien met het reilen en zeilen van Jongerenvereniging Unitas. Na een hectisch jaar vervolgde zij met frisse moed haar studie in de biochemische specialisatie. Haar eerste afstudeervak Moleculaire Genetica (Prof. dr. C. Heytink) werd gevolgd door een afstudeervak Bacteriële Genetica aan de vakgroep Microbiologie, onder begeleiding van dr. Rik Eggen en Prof. dr. Willem M. de Vos. Het geïnitieerde onderzoek werd succesvol voortgezet met een verblijf van een half jaar in het Biochemisch Laboratorium van Universiteit 'La Sapienza' in Rome, Italie (Prof. R. Scandurra). Vervolgens studeerde Joyce in augustus 1994 af en werd aangesteld als toegevoegd onderzoeker bij de vakgroep Microbiologie. In december 1994 werd deze positie omgezet naar Assistant in Opleiding. Tot april 1999 heeft Joyce binnen de werkgroep Bacteriële Genetica gewerkt aan de moleculaire karakterisatie van de thermostabiliteit en katalytische eigenschappen van enzymen uit hyperthermofielen. Gedurende deze periode deed zij een grote fascinatie op voor de structuurbiologie, wat zich mede uit in dit proefschrift en heeft geresulteerd in de aanstelling (per mei 1999) als Postdoctoral Fellow aan het Center for Structural Biochemistry van het Karolinska Institute in Zweden, onder leiding van Prof. Rudolf Ladenstein.

List of publications:

Lebbink, J.H.G., Eggen, R. I. L., Geerling, A. C. M., Consalvi, V., Chiaraluce, Scandurra, R. and de Vos, W. M. Exchange of domains of glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* and the mesophilic bacterium *Clostridium difficile*: effects on catalysis, thermoactivity and stability. *Prot. Engng.* **8**, 1287-1294, 1995.

Ramos, A., Raven, N. D. H., Sharp, R. J., Bartolucci, S., Rossi, M., Cannio, R, Lebbink, J.H.G., van der Oost, J., de Vos, W. M. & Santos, H. Stabilisation of enzymes against thermal stress and freeze-drying by mannosylglycerate. *Appl. Environ. Microbiol.* **63**, 4020-4025, 1997.

Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D. & Rice, D. W. Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336-346, 1998.

Lebbink, J. H.G., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. & de Vos, W. M. Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. I. Introduction of a six-residue ion-pair network in the hinge region. *J. Mol. Biol.* **280**, 287-296, 1998.

Lebbink, J. H.G., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. & de Vos, W. M. Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. II. Construction of a 16-residue ion-pair network at the subunit interface. *J. Mol. Biol.* **289**, 357-369, 1999.

Lebbink, J.H.G., Kengen, S.W.M., van der Oost, J. & de Vos, W.M. Glutamate dehydrogenase from hyperthermophilic Bacteria and Archaea: determinants of thermostability and catalysis at extremely high temperatures. *J. Mol. Catalysis B: Enzym.* **7**, 133-145, 1999.

Lebbink, J.H.G., Kaper, T., Kengen, S. W. M., van der Oost, J., de Vos, W.M. The β-glucosidase CelB from *Pyrococcus furiosus*: Production in *Escherichia coli*, purification and *in vitro* evolution. To be published in *Methods in Enzymology*, Volume 'Hyperthermophilic enzymes'.

Kaper, T., Verhees, C.H., Lebbink, J.H.G., van Lieshout, J.F.T., Kluskens, L.D., Ward, D.E., Kengen, S.W.M., de Vos, W.M., van der Oost, J. Characterization of β-glycosyl hydrolases from *Pyrococcus furiosus*. To be published in *Methods in Enzymology*, Volume 'Hyperthermophilic enzymes'.

Lebbink, J. H. G., Kaper, T., Bron, P., van der Oost, J., de Vos, W.M. Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* β-glucosidase CelB by directed evolution. *Submitted for publication.*

Brinkman, A., Dahlke, I., Tuininga, J., Lammers, T., Dumay, V., de Heus, E., Lebbink, J., Thomm, M., de Vos, W. M., van der Oost, J. The transcriptional regulator LrpA from the hyperthermophilic archaeon *Pyrococcus furiosus* is negatively autoregulated. *Submitted for publication.*

Pouwels, J., Moracci, M., Cobucci-Ponzano, B., Perugino, G., van der Oost, J., Kaper, T., Lebbink, J.H.G., de Vos, W.M., Ciaramella, M., Rossi. M. Activity and stability of hyperthermophilic enzymes: a comparative study on two archaeal β-glycosidases. *Submitted for publication.*

Kaper, P., Lebbink, J.H.G., Pouwels, J., Kopp, J., Schulz, G. E., van der Oost, J., de Vos, W.M. Engineering 6-phospho-β-galactosidase activity into the hyperthermostable *Pyrococcus furiosus* β-glucosidase CelB. *In preparation.*