Pyruvate dehydrogenase

Its structure, function and interactions within the pyruvate dehydrogenase multienzyme complex.

Promotoren:	Prof Dr C. Veeger
	Hoogleraar in de biochemie
Co-promotor:	Dr. A de Kok
	Universitair hoofddocent
	Laboratorium voor Biochemie
Promotie commissie	Dr. V. Bunnik, Birk Medical Research Institute, New York
	Prof. Dr. W.G.J. Hol, Washington University, Seattle
	Prof. Dr. J.F. Koster, Erasmus Universiteit, Rotterdam
	Prof. Dr. A.J.W.G. Visser, Wageningen Universiteit
	Prof. Dr. W.M. de Vos, Wageningen Universiteit
	Prof. Dr. S.C. de Vries, Wageningen Universiteit

Pyruvate dehydrogenase

Its structure, function and interactions within the pyruvate dehydrogenase multienzyme complex.

Annechien Frederique Hengeveld

Proefschrift

ter verkrijging van de graad van doctor

op gezag van de rector magnificus

van Wageningen Universiteit

Prof. Dr. Ir. L. Speelman

in het openbaar te verdedigen

op woensdag 24 april 2002

des namiddags om half twee in de Aula

ISBN: 90-5808-620-8

Stellingen

1. Op basis van de hoge homologie tussen prokaryote pyruvaat dehydrogenases (E1p) uit pyruvaat dehydrogenase multienzym complex kan geëxtrapoleerd worden dat prokaryoot E1p bindt aan de centrale component (E2p) via een N-terminaal bindingsdomein.
   Dit proefschrift

2. Op basis van de voorgestelde wijze van binding voor heterotetrameer E1b zijn er zeer grote verschillen tussen de wijze van binding van E1 aan kubisch E2 (homodimeren aan een 24-meer) en aan dodecaedrisch E2 (heterotetrameren aan een 60-meer), de voorgestelde wijze van binding van heterotetrameer E1 behoeft echter meer experimentele ondersteuning.
   Dit proefschrift & Ævarsson, A., Seger, K., Turley, S., Sokatch, J.R. and Hol, W.G.J.: Crystal structure of 2-oxoisovalerate and dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes (1999) Nat. Struct. Biol. 6, 785-792

3. De overgang van enzymologie naar farmacologie is vooral moeilijk omdat er in verschillende "talen" over dezelfde zaken wordt gediscussieerd.

4. Eiwit homologie modelling zonder experimentele validatie is niet meer dan speculeren over de structuur van een eiwit. Om deze reden zou er altijd een nauwe samenwerking moeten zijn tussen theoretici en experimentele biochemici / biofysici.

5. Alhoewel het ophelderen van de structuur van rhodopsine een enorme bijdrage heeft geleverd aan inzicht in het mechanisme van G-eiwit gekoppelde receptoren moet voorzichtigheid in acht worden genomen bij het extrapoleren van deze structuur naar andere G-eiwit gekoppelde receptoren.
   Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., Miyano, M.: Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor (2000) Science 289, 739

6. Het beter om in een goed milieu te leven dan om uit een goed milieu te komen.

7. Omdat het krijgen van een kind altijd slecht uitkomt binnen een wetenschappelijke carrière, is dit geen reden om het krijgen van kinderen uit te stellen.

8. Elke plant is een waterplant.

Stellingen behorende bij het proefschrift
'Pyruvate dehydrogenase. Its structure, function and interactions within the pyruvate dehydrogenase multienzyme complex.'

Annechien F. Hengeveld
Wageningen, 24 april 2002

Contents

Abbreviations and symbols

AD	Alzheimer's disease
ADHC	Acetoin dehydrogenase complex
BCAA	Branched-chain amino acid
BCDC	Branched-chain oxoacid dehydrogenase complex
BCKA	Branched- chain α-ketoacid
CAT	Chloramphenicol acetyltransferase
Cl_2Ind	2,6-dichlorophenol-indophenol
E1	α-ketoacid dehydrogenase component
E1b	Branched-chain α-ketoacid dehydrogenase (EC 1.2.1.25)
E1o	2-oxoglutarate dehydrogenase (EC 1.2.4.2)
E1p	Pyruvate dehydrogenase (EC 1.2.4.1)
E2	Dihydrolipoyl acyltransferase component
E2p	Dihydrolipoyl transacetylase (EC 2.3.1.12)
E3	Dihydrolipoamide dehydrogenase component (EC 1.8.1.4)
E3BP	E3 binding protein (protein X)
GDHC	Glycine dehydrogenase complex
LS	Leigh's syndrome
MP-8	Microperoxidase 8
MSUD	Maple Syrup Urine Disease
NOE	Nuclear Overhauser effect
Nterm-E1p	Peptide representing amino acid 4 to 45 of E1p from *Azotobacter vinelandii*
OGDC	Oxoglutarate dehydrogenase complex
PDC	Pyruvate decarboxylase
PDHC	Pyruvate dehydrogenase multienzyme complex
PFOR	Pyruvate ferredoxin oxidoreductase
POX	Pyruvate dehydrogenase
ThDP	Thiamin diphosphate
TK	transketolase
TPPI	time proportional phase incrementation

1 Introduction

Pyruvate dehydrogenase multi-enzyme complex (PDHC) is member of a family of multi-enzyme complexes that catalyse the irreversible decarboxylation of various 2-oxoacid substrates to their corresponding acyl-CoA derivatives, NADH and CO_2. PDHC, like the other members of this family, consists of an assembly of non-covalently attached proteins that catalyse successive steps in a reaction sequence. The combined catalytic activity of the enzymes in the assembly is greatly enhanced through the coupling of the successive reactions of the complex. Covalent attachment of reaction intermediates forces them to complete the reaction sequence and prevents "scavenging" by competing enzymes. Additionally, active site coupling and substrate channelling limit the diffusion of the substrates and degradation of unstable reaction intermediates. The coupling of the successive reaction steps finally results in a very effective regulation of the overall reaction sequence that can occur on all participating enzymes [Reed, 1974; Hammes, 1981].

2-oxoacid dehydrogenase complexes hold key points in primary energy metabolism (figure 1): PDHC at the branch-point of glycolysis and citric acid cycle, the oxoglutarate dehydrogenase complex (OGDC) in the citric acid cycle, the branched-chain oxoacid dehydrogenase complex (BCDC) in amino acid catabolism, the glycine dehydrogenase complex (GDHC) in the glycine cleavage system and in the photosynthetic carbon cycle and acetoin dehydrogenase complex (ADHC) in acetoin metabolism. Malfunctioning of any of these complexes in man leads to a broad variety of clinical manifestations [Wynn et al., 1996; Indo & Matsuda, 1996; Kerr et al., 1996; Danner & McKean, 1996; Chuang & Shih 1995; Robinson, 1995; Kerr & Zinn 1995; Dahl, 1995; Patel & Harris 1995; Roche & Patel, 1989; Kish, 1997; Mizuno et al., 1994]. Consequently, there is a broad interest in the functioning of these complexes on a molecular level. This thesis describes the studies on the pyruvate dehydrogenase (E1p) component of the PDHC from *Azotobacter vinelandii*, a gram-negative bacterium. PDHC from *Azotobacter vinelandii* is of special interest as it is the smallest and simplest know 2-oxoacid dehydrogenase. Consequently, in this introduction I will mainly focus on the structure and

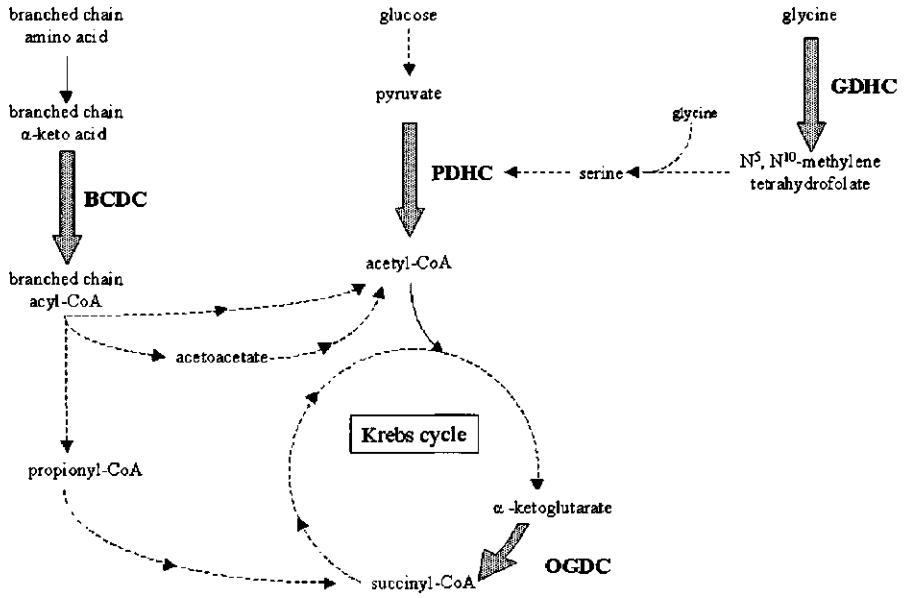


Figure 1. Metabolic location of 2-oxoacid dehydrogenase complexes [Olson, 1989].

function of PDHC from gram-negative bacteria. The functioning and dysfunctioning of human 2-oxoacid dehydrogenase multi-enzyme complexes and their involvement in metabolic disease is reviewed in chapter 7.

Reaction mechanism and architecture

Five 2-oxoacid dehydrogenase complexes are found in gram-negative bacteria: PDHC, OGDC, BCDC, GDHC and ADHC. All, except GDHC [Douce et al., 2001], share a common architecture. The complexes consist of multiple copies of three enzymes: 2-oxoacid dehydrogenase or E1-component, dihydrolipoyl acyltransferase (E2) and dihydrolipoamide dehydrogenase (E3). Mammalian and yeast PDHC contain a so-called E3-binding protein (E3BP, formerly called protein X) [de Marcucci & Lindsay, 1985; Lawson et al., 1991; Neagle & Lindsay, 1991] that is involved in the binding of E3 to the complex. Eukaryotic PDHC and BCDC in addition, contain specific regulatory proteins (E1-kinase and E1 phosphatase) [Linn et al., 1969; Patel & Roche, 1990]. E2 forms the central core of the complex with either octahedral (24 subunits) or icosahedral (60 subunits) symmetry. Multiple

copies of the peripheral components E1 and E3 are non-covalently attached to this core. The resulting multi-enzyme complexes are of such an enormous size (~5-14 MDa) that complexes

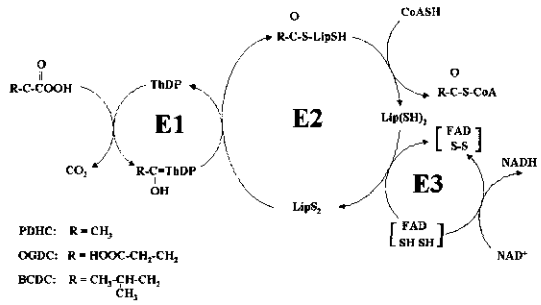


Figure 2. Reaction mechanism of 2-oxoacid dehydrogenase complexes. A 2-oxo acid dehydrogenase (E1p, E1b or E1o) uses thiamin diphosphate (ThDP) to carry out the oxidative decarboxylation of the substrate, forming an acylated lipoic acid, covalently attached to the acyltransferase component (E2p, E2b or E2o) and CO2. The E2-component catalyses the transfer of the acyl group to coenzyme A (CoA). The resulting dihydrolipoyl group is then reoxidised by the dihydrolipoamide dehydrogenase component (E3) using NAD+, forming NADH.

of octahedral and of icosahedral symmetry can be distinguished by electron microscopy [Oliver & reed, 1982]. As an exception, the E2p core (the subscript "p" denotes the origin of the E2 component: "p" for PDHC, "o" for OGDC and "b" for BCDC) of *Azotobacter vinelandii* dissociates into trimers upon binding of the peripheral E1p or E3 component, resulting in a complex of only 700 kDa [Schulze et al., 1991]. The same phenomenon occurs as a result of the expression of the *Escherichia coli* E2o catalytic domain with a C-terminal $(his)_6$ tag [Knapp et al., 2000]. The C-terminus of the trimeric form is rotated 37° compared to the cubic form, disrupting the two-fold interface. No effects however were observed in the active site.

PDHC irreversibly catalyses the oxidative decarboxylation of pyruvate, generating acetyl-CoA, NADH and CO_2 (figure 2). In this reaction three prosthetic groups (thiamin diphosphate, lipoic acid and flavin adenine dinucleotide) and two cofactors (CoA and NAD^+) are involved. E1p utilises thiamin diphosphate (ThDP) to catalyse the oxidative decarboxylation of pyruvate and to subsequently reductively acetylate the lipoic acid that is attached to E2p. This lipoic acid, which is covalently attached to a lysine on E2p forming the

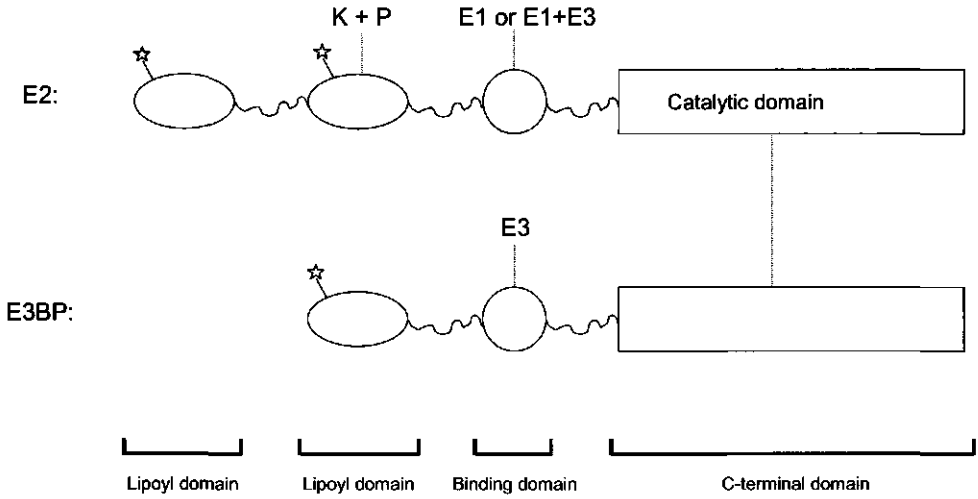


Figure 3. The domain structure of E2. The separately folding domains are connected by flexible linker sequences shown in wavy lines. An asterisk indicates the lipoic acid that is covalently attached to the lipoyl domain.

lipoyl-group, is involved in all three steps of the reaction mechanism and visits the active sites of all three enzymes. E2p catalyses the transfer of the acetyl group from the lipoyl group to CoA, forming and releasing acetyl CoA. Finally, the reduced lipoyl group is reoxidised by E3 utilising flavin adenine dinucleotide (FAD) and transferring the reduction equivalents to NAD^+, forming NADH. Unlike eukaryotic PDHC gram-negative PDHC does not contain additional regulatory enzymes, but is regulated via allosteric mechanisms.

The acyltransferase component (E2) and dihydrolipoamide dehydrogenase (E3)

This introduction mainly focuses on the E1 component of 2-oxoacid dehydrogenase complexes. For a better understanding however, the most important features of the other components of the complex will be discussed in this section [for reviews see Perham, 1991; Mattevi et al., 1992a; Berg & de Kok, 1997; de Kok et al., 1998; Perham, 2001].

Limited proteolysis studies revealed that E2 has a multi-domain structure that contains three independent folding units that are separated by linker sequences of about 20-50 amino acids (Figure 3) [Bleile et al., 1979; Bleile et al., 1981; Packman et al., 1984a; Packman et al.,

1984b; Packman & Perham, 1987; Chuang, 1985; Hanemaaijer et al., 1987]. From N- to C-terminus it contains two or three lipoyl domains (dependent on the type and source of the complex), an E1/E3 binding domain and a catalytic domain. NMR experiments indicate that the linker sequences are extended and highly flexible [Perham et al., 1981; Duckworth et al., 1982; Hanemaaijer et al., 1988; Texter et al., 1988]. This flexibility is essential for the mechanism of active site coupling, by which the substrate, covalently attached to the lipoyl domain, is transferred between the active sites of E1, E2 and E3 (figure 2) [Miles et al., 1988; Turner et al., 1993; Schulze et al., 1993].

The role of the lipoyl domain and its role in active site coupling was extensively reviewed recently by [Berg & de Kok, 1997] and [Perham, 2000]. 3-dimensional structures of lipoyl domains of several sources have been solved [Dardel et al., 1993; Green et al., 1995; Berg et al., 1996; Berg et al., 1997; Ricaud et al., 1996]. All show a very similar overall fold: two four-stranded β-sheets packed in a sandwich like manner. A conserved lysine residue is exposed in a β-turn at the far end of the domain. A lipoic acid prosthetic group is attached to this residue. An exposed loop close to the lipoylated lysine residue was shown to be involved in the species specific recognition of E1, it has become clear however, that recognition of the lipoyl domain by its partner E1 is a complex process and can not be attributed to a single determinant on the lipoyl domain [Jones et al., 2000; Jones et al., 2001; Howard et al., 2000; Gong et al., 2000].

The peripheral E1/E3 binding domain (~35 amino acid residues) is one of the smallest proteins that has a stable globular fold in the absence of disulphide bridges or prosthetic groups [Brocklehurst et al., 1994]. The structure of the binding domain of *E. coli* E2o and *Bacillus stearothermophilus* E2p were solved by NMR [Robien et al., 1992; Kalia et al., 1993] and additionally, the structure of the binding domain of *B. stearothermophilus* E2p in complex with E3 was solved by X-ray crystallography [Mande et al., 1996]. In accord with the substantial sequence homology for the binding domains of different complex type and of different sources, the solved structures of the binding domain were very similar. The global fold comprises two parallel a-helices, separated by a short loop, a short helix and a longer disordered loop.

The C-terminal catalytic domain of E2 harbours the acyltransferase active site and in addition the sites responsible for the formation of the octahedral or icosahedral core. Crystal structures are now at hand for both the 24-meric cubic structure and the 60-meric dodecahedral structure [Mattevi et al., 1992; 1993a; 1993b; Knapp et al., 1998; Izard et al., 1999]. Both have a

hollow interior and contain large pores at their faces. Trimeric E2-units form the building blocks of these aggregates, with extensive and strong interactions. The inter-trimer interactions are much weaker. The active site of E2 is located at each interface of 2 E2-subunits in a trimer; a 3 nm long channel is formed, with two entrances. Lipoamide enters this channel from the outside, whereas CoA enters from the inside of the hollow core. A conserved histidine and serine/threonine residue are involved in the reaction mechanism, as was shown by extensive mutagenesis/X-ray crystallography studies [Mattevi et al., 1993c; Hendle et al., 1995]. More extensive reviews on E2 can be read in [de Kok et al., 1998; Berg & de Kok, 1997; Perham 1991].

E3 is an FAD-linked homodimer that reoxidises the reduced lipoamide, forming NADH [Williams, 1992]. Several 3D structures have been solved by X-ray crystallography [Mattevi et al., 1991; 1992b; 1993b; de la Sierra et al., 1997]. The two identical subunits of E3 are composed of four distinct domains: an FAD binding domain, an NAD binding domain, a central domain and an interface domain. The catalytic side of the enzyme is located at the dimer interface. The catalytic mechanism of E3 has been extensively investigated: the final production of NADH involves the transfer of reducing equivalents from the E2-linked lipoamide via the FAD cofactor and a redox-active cysteine disulphide pair [Hopkins & Williams, 1995].

The 2-oxoacid dehydrogenase component (E1)

History

The E1 component catalyses the rate limiting step in the overall complex reaction; the reductive acylation of the lipoyl groups [Cate et al., 1980]. Until quite recently it was by far the least characterised component of the complex. This lack of (structural) information was largely due to the lack of a stable form of the enzyme: when resolved from the complex, E1 is unstable. At the start of this project only the recombinant expression of two functional heterotetrameric E1 [Wynn et al., 1992b; Lessard & Perham, 1994] had been reported, whereas no stable homodimeric E1 was available.

In this chapter I will review the structural and functional data that are now at hand for E1. To put this thesis in its right perspective, the reader should keep in mind however, that most of this information has only come available very recently.

The E1-component exists either as a homodimer (α_2) or heterotetramer ($\alpha_2\beta_2$), depending on the type and source of the complex [Perham, 1991]. 2-oxoglutarate dehydrogenase (E1o) and pyruvate dehydrogenase (E1p) from gram-negative sources are homodimeric enzymes with subunits of ~100 kDa. Branched-chain 2-oxoacid dehydrogenase (E1b) and E1p from other sources exist as heterotetramers with subunits of ~36 kDa and ~ 41kDa. Exceptional E1 components were found in *Z. mobilis* (a tetrameric E1p that contains an N-terminal domain in its β-subunit) [Neveling et al., 1999] and *Alcaligenes eutrophus* that possesses two E1p components {Hein & Steinbüchel, 1996]. There is strikingly little sequence similarity between homodimeric and heterotetrameric E1 or even between E1o and homodimeric E1p. An exception to this is the so-called "Thiamin diphosphate-binding motif", that has so far been found in all known thiamin diphosphate (ThDP) dependent enzymes [Hawkins et al., 1989].

ThDP-dependent enzymes

E1 is a member of the family of ThDP dependent enzymes. In this paragraph, I will briefly describe some of the other members of this family, as much of the mechanistic information on E1 is extrapolated from these enzymes.

In all ThDP dependent enzymes the diphosphorylated state of thiamin (vitamin B1) (Figure 4)

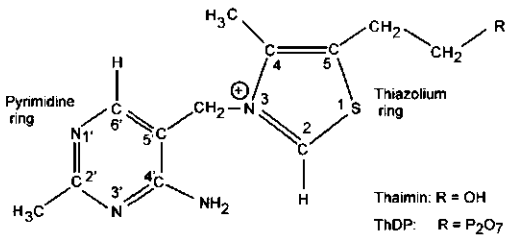


Figure 4 Thiamin diphosphate.

functions as a cofactor, along with a metal ion such as Mg^{2+} or Ca^{2+}, in many reactions where a C-C bond adjacent to a carbonyl group is cleaved. The chemical steps that occur during non-enzymatic ThDP catalysis are well known and have been reviewed in detail [Kluger, 1987; Schellenberger, 1998]. With the structural information at an atomic level available for a number of ThDP dependent enzymes (pyruvate decarboxylase (PDC) [Dyda et al., 1993], pyruvate oxidase (POX) [Muller & Schulz, 1993], transketolase (TK) [Lindqvist et al., 1992] and pyruvate ferredoxin oxidoreductase (PFOR) [Chabrière et al., 1999] significant advances have been made in the understanding of the enzyme-catalysed mechanism.

TK catalyses the transfer of a dihydroxyethyl group between a ketose and an aldose sugar phosphate (figure 5) [Sundström et al., 1993]. TK is found in the non-oxidative branch of the pentose phosphate pathway where it, together with transaldose, creates a link to glycolysis.

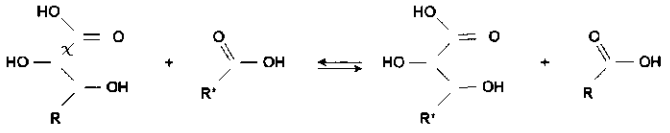


Figure 5. The general reaction mechanism transketolase.

TK is a homodimer, with subunits of 74 kDa. The determination of the 3D structure of yeast TK by X-ray crystallography revealed for the first time the fold of a ThDP dependent enzyme and provided the first insights into the structural basis for ThDP binding and catalysis [Lindqvist & Schneider, 1993]. TK has three domains: i) a pyrophosphate (pp) binding domain, ii) the pyrimidine (pyr) binding domain and iii) the C-terminal domain. ThDP is bound at the interface between the pp-domain of one subunit and the pyr-domain of the other subunit. Of all ThDP dependent enzymes TK is the most homologous to the homodimeric E1p (see also chapter 6 of this thesis).

POX catalyses the oxidative decarboxylation of pyruvate in the presence of phosphate and O2 during the formation of acetylphosphate and hydrogen peroxide [Sedewitz et al., 1984a; 1984b]. The enzyme uses FAD as a second cofactor. The X-ray structure of tetrameric POX from *Lactobacillus plantarum* was solved [Muller & Schulz, 1993]. The enzyme consists of three domains: i) the pyrimidine (pyr) binding domain, ii) an FAD binding domain and iii) a pyrophosphate (pp) binding domain. Analogous to TK, ThDP is bound at the interface of the pyr-domain and the pp-domain originating from different subunits.

PDC is a key enzyme in alcoholic fermentation that catalyses the conversion of pyruvate to acetaldehyde:

$$CH_3COCOO^- + H_20 \rightarrow CH_3CHO + OH^- + CO_2$$

The catalytic mechanisms of PDC from yeast [Lobell & Crout, 1996] and *Zymomonas mobilis* [Sun et al., 1995] have been investigated in detail. Yeast PDC is allosterically regulated via substrate activation. 3D high-resolution structures were solved for PDC from yeast [Dyda et al., 1993] and from *Z. mobilis* [Dobritzsch et al., 1998]. Both exist as homotetramers in the native state with subunits of ~ 60 kDa. The overall domain structure of PDC is very similar to that of POX, despite the fact that the overall sequence identity is less than 20% [Lindqvist & Schneider, 1993].

PFOR is utilised by anaerobic bacteria and archaea and can be regarded as the precursor of PDHC [Witzmann & Bisswanger, 1998]. It catalyses the decarboxylation of pyruvate, forming acetyl-CoA. The surplus reducing equivalents are transferred via ferredoxin to cytochrome C [Charon et al., 1999]. Unlike PDHC, PFOR can also perform the reverse reaction; that is, the formation of pyruvate from CO2 and acetyl-CoA. The structure of PFOR from *Desulfovibrio africanus* was solved by X-ray crystallography [Chabrière et al., 1999]. It is a homodimer of 266 kDa, each subunit contains one ThDP, two ferredoxin type [4Fe-4S] clusters and one [4Fe-4S] cluster coordinated by an atypical cysteine-containing motif. In contrast with other ThDP dependent enzymes the two domains that bind ThDP (the pp- and the pyr-domain) originate from one subunit.
In addition to the above-described enzymes, the structure of benzoyl formate decarboxylase was solved [Hasson et al., 1998], that also shows the same general ThDP-binding fold.

Reaction mechanism of E1

ThDP catalyses numerous reactions that all are initiated by the deprotonation of the catalytic centre of the coenzyme (the C_2-atom of the thiazolium moiety) (figure 6, step 1). The dissociation rate of the C_2-H is enhanced about four orders of magnitude through the action of a conserved glutamine, via the N_1' and N_4' groups of the pyrimidine moiety of ThDP [Kern et al., 1997]. In all known structures of ThDP dependent enzymes, ThDP has a highly constrained, protein induced V-shaped conformation [Muller et al., 1993; Shin et al., 1977], which prevents the simultaneous protonation of the thiazole C_2 and the pyrimidine N_4' atom. The formed C_2-carbanion carries out a nucleophilic attack on the substrate (figure 6, step 2). This activation mechanism has been confirmed by ^{13}C-NMR for PDC, TK [Kern et al., 1997],

addition of an electrophile, such as a proton, via a general acid catalyst located on E1 to one of the sulphur atoms of the lipoic acid, thereby shifting the equilibrium to the products [Jordan, 1999].

For further detail on ThDP-chemistry and ThDP enzymes the reader is referred to excellent reviews published in a special issue of Biochimica et Biophysica Acta [Schellenberger, 1998a] and [Charon et al., 1999].

Regulation of E1

E1 catalyses the rate limiting step of the overall complex reaction and is therefore an ideal target for complex regulation. In prokaryotic complexes many effectors influence the activity of PDHC in a positive or negative manner. Metabolic inhibitors of PDHC include NADH, acetyl-CoA and GTP [Bremer, 1969; Schwartz et al., 1968; Schwartz & Reed, 1970;

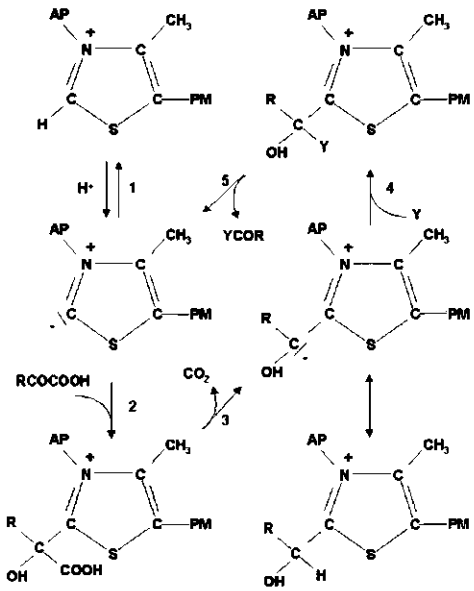


Figure 6 The catalytic mechanism of ThDP enzymes. A description of the mechanism is given in the section "reaction mechanism of E1".

Bisswanger & Henning, 1971; Bosma, 1984], additionally, regulation of E1p occurs by its substrate pyruvate. Mammalian E1p is regulated in a more complex fashion by phosphorylation/dephosphorylation of the α-subunit [Patel & Roche, 1990; Patel et al., 1996;

Kolobova et al., 2001; Korotchkina & Patel, 2001]. E1 is the target of (synthetic) substrate analogues and ThDP-analogues [Lowe et al., 1983; Gutowski & Lienhard, 1976]. Fluoropyruvate [Leung & Frey, 1978; Bisswanger, 1980; 1981; Flourney & Frey, 1989] and phosphonate [Kluger & Pike 1977] and phosphonate analogues of pyruvate [Schonbrunn-Hanebeck et al., 1990] act as competitive inhibitors, whereas bromopyruvate [Maldonado et al., 1972; Apfel et al., 1984; Lowe & Perham, 1984] and 2-oxo-3-alkynoic acids [Brown et al., 1997] act as suicide inhibitors. In addition, PDHC activity is very sensitive for ionic strength and pH [Pawelczyk et al., 1992; Rodriguez-Zavala et al., 2000].

The mechanism of regulation of *A. vinelandii* E1o by its substrate and by structural substrate analogs can be represented by a scheme that considers the formation of a pre-catalytic complex "SE" between the substrate and the enzyme, before the catalytic complex "ES" is formed. In the catalytic complex the ThDP bound enamine/$C_2\alpha$-carbanion intermediate is formed. The incorrect orientation of the substrate in the pre-catalytic complex, or the occupation of this site by a substrate analog, causes the substrate or substrate analog inhibition [Bunik et al., 2000]. Evidence of a precatalytic complex was found in the X-ray structure of PFOR [Chabrière et al., 1999], where one of the carboxyl oxygen atoms of pyruvate, instead of the carbonyl oxygen, interacts with the N_4' of ThDP. In *P. putida* E1b [Ævarsson, 1999] the substrate analog a-chloroisocaproate binds in a similar fashion. Therefore, the precatalytic binding of 2-oxoacid substrates could be a general step in ThDP catalysis.

Hennig and co-workers [Hennig et al., 1997] investigated the mechanism of the regulation of *E. coli* E1p by pyruvate by testing the influence of pyruvate on the interaction of E1 with several ThDP-analogs. Pyruvate enhances the binding of ThDP and coenzymatically active ThDP-analogs, which point to an increased binding affinity of the reaction intermediate-ThDP complex to E1. Lowering of the pyruvate concentration has an opposite effect: it enhances the dissociation of ThDP. At physiological ThDP concentrations ThDP binding and dissociation are slow compared to the catalytic activity, and therefore E1 can efficiently be regulated through this pyruvate-mediated effect.

3D structure of E1

The crystal structures of human and *P. putida* E1b (an $\alpha_2\beta_2$ tetramer) [Ævarsson et al., 1999; Ævarsson et al., 2000] reveal a tightly packed arrangement of the four subunits, with the β_2-

Figure 7. Overall structure of human heterotetrameric E1b [Ævarsson et al., 2000].

dimer held between the jaws of a 'vise' formed by the α_2-dimer (figure 7). The α-subunit is composed of three domains: an extended N-terminal tail, a large domain and a small C-terminal domain. The β-subunit consists of two domains. Each subunit in the tetramers interacts with all three other subunits through extensive surface contacts (even though the interactions are mainly between the domains of the α-subunit and the N-terminal domain of the β-subunit). The structures of human and *P. putida* E1b are very similar: both have a topology that is again very similar to that of TK from yeast: residues from both the α- and β-subunit form a core that is also found in PDC, POX, TK, PFOR and BFD [Muller et al., 1993; Lindqvist & Schneider, 1993]. The α-subunit of E1b corresponds to the N-terminal half of

TK, the β-subunit to the C-terminal half (Chapter 7 in this thesis compares heterotetrameric E1 and TK in more detail). Most likely, homodimeric E1 has the same principle structure as heterodimeric E1 (see chapter 6 for a more detailed comparison of homodimeric and heterotetrameric E1).

A 20Å long hydrophobic channel suitable to accommodate the E2-lipoyl-lysine arm leads to the active site of E1b. Two ThDP-binding pockets are formed at the interface between the α- and β'-subunit and visa versa. The α-subunit binds the phosphate end of ThDP and the β'-subunit binds the pyrimidine end of ThDP (figure 8). In a nutshell the following residues are

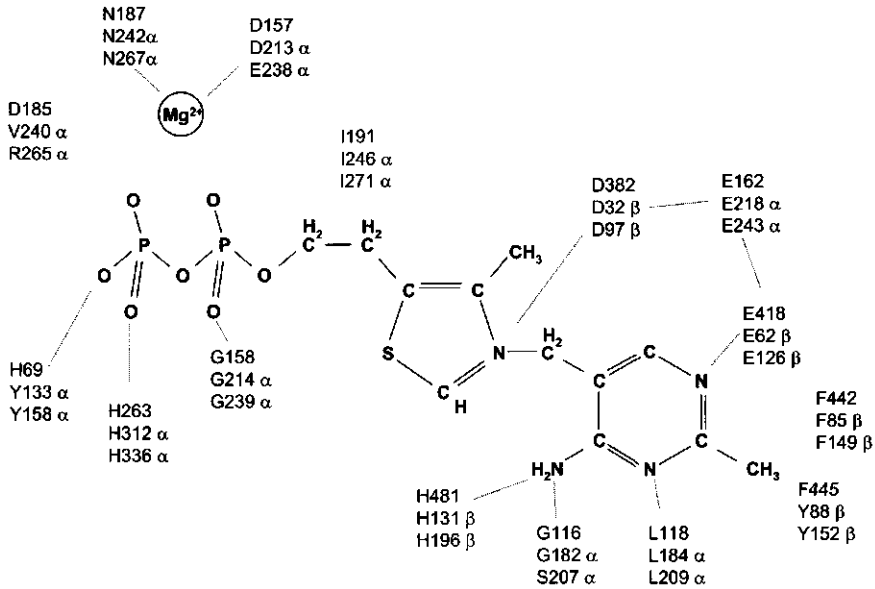


Figure 8: Conserved thiamin diphosphate (ThDP) protein interactions in the active site of from top to bottom, transketolase (TK), P. putida E1b and human E1b.

of importance in the active site of E1b: residues from the so-called "ThDP-binding motif" (GDG-$X_{(25\text{-}30)}$-N) [Hawkins et al., 1989] interact with the phosphate moiety of ThDP, whereas the characteristic V-conformation of the cofactor is formed through the action of hydrophobic residues in the β-subunit. A glutamate of the β-subunit (E62β in *P. putida* E1b) has an important role in cofactor activation, as was confirmed by mutagenesis [Wynn et al. 2001]. The protonation potential of this glutamate is increased by a second glutamate from the

α-subunit (E218α in *P. putida*). Two histidines are oriented towards the reactive C_2 of ThDP and are most likely involved in substrate recognition and binding. The hydrophobic channel positions the disulphide exactly between these two residues. Possibly one of these histidine residues acts as a proton donor in the reduction of the disulphide bond. The importance of the corresponding active site residues in *E. coli* E1p was confirmed by site directed mutagenesis [Russell et al., 1992; Yi et al., 1996]. For more detail on the structural properties of E1 and its role in the dysfunctioning of the complexes the reader is referred to chapter 7 of this thesis.

Assembly

As was already described in an earlier section, the structural core of the 2-oxoacid dehydrogenase complexes is formed by aggregates of E2. E2p from gram-negative bacteria, E2o and most E2b's form a cubic assembly of 24-subunits [Perham, 1991]. All other E2's form a core of 60 subunits with icosahedral symmetry. The non-covalent binding of the peripheral components E1 and E3 results in huge complexes of 5-14 MDa [Oliver & Reed, 1982]. *E. coli* E2p is able to bind up to 24 E1p dimers or 24 E3 dimers [Reed et al., 1975], but in the isolated complex the stochiometry varies [Hale & Perham, 1979]. Optimal stochiometry for *E. coli* PDHC is 12 E1p dimers and 6 E3 dimers bound to the 24-meric cubic core (E1p to the edges and E3 to the faces of the cube) [Reed et al., 1975]. PDHC from *A. vinelandii* forms the smallest known complex: only 700 kDa [Schulze 1992]. It exists of a functional E2p trimer, two E1p dimers and one E3 dimer. In vitro, 2-oxoacid dehydrogenase complexes are self-assembling: they can be reconstituted from their individual components [Bates et al., 1977; Bosma, 1984; Domingo et al., 1999]. As an exception, human BCDC requires chaperones for in vitro reassociation [Wynn et al., 2000; Chuang et al., 1999].

The binding of E1 and E3 to E2 is competitive: an excess of either component can displace the other component [Bosma, 1984;Reed et al., 1975; Lessard et al., 1996]. For *B. stearothermophilus* E2p it was shown that E1p and E3 couldn't bind simultaneously to the same binding domain [Lessard & Perham, 1995]. Mutagenesis experiments of *A. vinelandii* E2p identified separate binding regions on the binding domain for E1p and E3 [Schulze et al., 1993]. The elucidation of the X-ray structure of the isolated binding domain of *B. stearothermophilus* in complex with E3 revealed that an E3 dimer binds to the N-terminal helix of the binding domain of E2p, mainly via electrostatic interactions [Mande et al., 1996]. The subunit interface of E3 is involved in the binding as was shown in the E3 - binding

domain structure and as was shown by the stabilisation of the subunit interactions upon binding to E2p [Westphal et al., 1995]. Based on the stochiometry of binding and on the fact that the dissociation of E1 requires monomerisation (high pH), it seemed likely that E1p also binds at its subunit interface. In contrast, mutagenesis experiments of *A. vinelandii* E2p show that E1p requires binding sites on both the binding- and catalytic-domain of E2p [Schulze et al., 1991c; 1992; 1993], suggesting two binding sites on E1p. In this thesis, we identified the extended N-terminal domain of *A. vinelandii* E1p as the domain that is necessary for the binding of E1p to E2p [Hengeveld 1997, 1999] (chapters 2-4). For efficient binding two N-terminal binding domains are required (see chapter 3) and it seems likely that one N-terminal domain of E1p binds to the binding domain of E2p whereas the second N-terminal domain of E1p binds to the catalytic domain of E2p. Therefore, Cryo electron microscopy of frozen-hydrated complexes [Wagenknecht et al., 1990; 1992] shows that *E. coli* E1p and E3 are separated from the E2p core by 3-5 nm. This suggests a flexible mode of attachment of E1p and E3 to the E2p core, by the flexible linker sequences that connect the binding domain to the catalytic domain. Alternatively, for E1p, the distance might be bridged by an extended N-terminal domain (see chapters 2, 3 and 4 of this thesis). The advantage of flexible binding of E1 is that it does not impose constraints that could prevent allosteric transitions in E1p. Recently, a size variation of 40 Å was demonstrated for *B. stearothermophilus* PDHC, emphasising the dynamic character of the complexes [Zhou et al., 2001].

Outline of this thesis

To obtain a better insight into the structure and functioning of the E1-component of the pyruvate dehydrogenase multi-enzyme complex from *A. vinelandii* a PhD project was started in 1995. The results of this project are presented in this thesis.

Much was known about the multienzyme complexes at the start of this project, but very little of this knowledge concerned E1. The enzyme had not been cloned or sequenced and the wild type E1p was unstable when separated from the complex. Chapter 2 describes for the first time the over-expression and purification of a stable homodimeric E1p. In addition, limited proteolysis indicates that the N-terminal region of the enzyme is involved in the binding of E1p to E2p. Chapter 2 and 3 illustrate the importance of this N-terminal region of E1p. N-terminal deletion mutants were constructed to investigate the role of the N-terminal region of E1p in more detail (chapter 3). In Chapter 4 we describe in detail the properties of a synthetic

peptide representing this region using advanced spectroscopic techniques. In chapter 5 point mutations reveal the specific binding residues in this region.

Detailed structural information on a heterotetrameric E1 has come available in the final stages of this project. Chapter 6 highlights the similarities between the homodimeric and heterotetrameric E1's and we propose a model for the homodimeric E1p. Finally, in chapter 7 the structural basis of the dysfunctioning of the 2-oxo acid dehydrogenase complexes is reviewed.

2 Expression and characterisation of the homodimeric E1-component of the *Azotobacter vinelandii* pyruvate dehydrogenase complex.

This chapter is a modified version of Hengeveld, A.F.; Westphal, A.H. & de Kok, A. (1997): "Expression and characterisation of the homodimeric E1 component of the *Azotobacter vinelandii* pyruvate dehydrogenase complex" Eur. J. Biochem. **250** , 260-268.

Abstract

We have cloned and sequenced the gene encoding the homodimeric pyruvate dehydrogenase component (E1p) of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* and expressed and purified the E1p component in *Escherichia coli*.

The cloned E1p can be used to fully reconstitute complex activity. The enzyme is stable in high ionic strength buffers, but is irreversibly inactivated when incubated at high pH, which presumably is caused by its inability to redimerize correctly. This explains the previously found low stability of the wild-type E1p component after resolution from the complex at high pH.

The cloned E1p shows a kinetic behaviour exactly as the wild-type complex-bound enzyme with respect to its substrate (pyruvate), its allosteric properties and its effectors. These experiments show that acetyl coenzyme A acts as a feedback inhibitor by binding to the E1p component.

Limited proteolysis experiments show that the N-terminal region of E1p is easily removed. The resulting protein fragment is still active with artificial electron acceptors but has lost its ability to bind to the core component (E2p) and thus reconstitute complex activity. E1p is protected against proteolysis by E2p. The allosteric effector pyruvate changes E1p into a conformation which is more resistant to proteolysis.

Introduction

The pyruvate dehydrogenase complex (PDHC) from *Azotobacter vinelandii* catalyses the oxidative decarboxylation of pyruvate and the subsequent acetylation of coenzyme A (CoA) to acetyl-CoA (for recent reviews see (Perham,1991; Mattevi, De Kok, & Perham, 1992a; Berg, & De Kok, 1997). The complex is composed of multiple copies of three different components: pyruvate dehydrogenase (E1p), dihydrolipoyl acyltransferase (E2p), and lipoamide dehydrogenase (E3). The E2p component plays a central role in the functioning of the complex. It has a multidomain structure: the N-terminal part contains three lipoyl domains each carrying a lipoamide moiety covalently attached to a specific lysyl residue, following

there is an E1p/E3 binding domain, and at the C-terminus there is a catalytic domain containing the acyltransferase catalytic site, which also forms the structural core of the complex. The different E2p domains are linked by proline-alanine-rich spacer sequences that are highly flexible and which are essential for the PDHC activity (Perham, Duckworth, & Roberts, 1981).

Because the crystallisation of the entire complex is most likely prohibited by the flexibility of the lipoyl domains caused by its flexible linker regions, the structures of the several different components of the *A. vinelandii* PDHC complex were solved. Todate crystal structures of the 24-meric octahedral catalytic domain of *A. vinelandii* E2p (Mattevi et al., 1992c; Mattevi et al., 1993a) and of E3 (Mattevi et al., 1991) were solved. The solution structure of the E2p lipoyl domain (Berg, Vervoort & De Kok, 1997) was solved by NMR spectroscopy. The binding mode of E3 to the binding domain has been solved by X-ray crystallography for the PDHC from *B. stearothermophilus* (Mande et al., 1996). As yet no structural information on the E1 component is available.

The E1 component can be a homodimer ($\alpha 2$) or a heterotetramer ($\alpha 2\beta 2$). Homodimeric E1 with a molecular mass of approximately 100 kDa is found in all 2-oxoglutarate dehydrogenases (E1o) and in pyruvate dehydrogenases (E1p) from gram-negative bacteria. The heterotetramer with subunits of approximately 41 and 36 kDa is found in branched-chain 2-oxoacid dehydrogenases (E1b) and in E1p from eukaryotes and gram-positive bacteria. It is striking how little sequence similarity is observed between the homodimeric and the heterotetrameric E1 components or even between the homodimeric E1p and E1o (Matuda et al., 1991). One common motif was identified in all E1 components: a putative thiamin diphosphate-binding motif (Hawkins et al., 1989). This motif was found in all known thiamin diphosphate-dependent enzymes and is predicted to be involved in the thiamin diphosphate binding (ThDP). Recent three-dimensional structure determinations of several ThDP-dependent enzymes (Lindqvist & Schneider, 1993) - transketolase (TK) (Lindqvist et al., 1992), pyruvate oxidase (POX) (Muller & Schulz, 1993) and pyruvate decarboxylase (PDC) (Dyda et al., 1993) - have shown that the motif is indeed involved in binding the metal ion and the diphosphate group of ThDP. Mutagenesis experiments have also proven that these residues are involved in the binding of the metal ion and ThDP (Russell et al., 1992; Diefenbach et al., 1992).

Several heterotetrameric E1's were previously cloned and expressed in *E. coli.* Genes expressing the E1α and the E1β subunit of the mammalian branched-chain 2-oxoacid

dehydrogenase complex were overexpressed in *E. coli* (Wynn et al., 1992a; Wynn et al., 1992b; Davie et al., 1992). Coexpression of the E1α subunit fused to a maltose-binding protein (MBP) and the E1β subunit resulted in an active E1b (α2β2), but mixing *in vitro* did not (Davie et al., 1992). In contrast, the E1α and E1β subunits of the E1p component of the PDHC from *B. stearothermophilus* were produced separately in *E. coli* (Lessard & Perham, 1994) and a functional E1p (α2β2) was produced by *in vitro* mixing of the individual subunits with one subunit in excess. In addition, the genes encoding the E1α and E1β subunits of the mammalian PDHC were co-expressed, producing a functional E1p-component (Jeng et al., 1994). Likewise, the active E1b component of *P. putida* was overexpressed in both *P. putida* and in *E. coli* (Hester et al., 1995) and a functional E1b of *Streptomyces avermitilis* was produced in *E. coli* (Skinner et al., 1995).

The homodimeric E1p component has, until now, resisted most attempts to characterisation. No procedure for the expression of E1p has been described sofar. The main problem in characterising the homodimeric E1p has been its instability when resolved from the complex. The *A. vinelandii* E1p is dissociated from the complex by binding the complex on a Thiol-Sepharose matrix and eluting the E1 component at pH 9.4, a procedure which incubates the components at alkaline pH for a short time only (De Graaf-Hess & De Kok, 1982). The dissociation of the complex is thought to be caused by the monomerisation of the E1 component if incubated at alkaline pH (Reed & Oliver, 1968b).

This paper describes the cloning and expression of an active and stable homodimeric E1p and its characterisation.

Experimental procedures

Materials

Restriction endonucleases and T4-DNA-ligase were purchased from Bethesda Research Laboratory (BRL). The Wizard Miniprep DNA purification system and the Wizard DNA clean-up system were purchased from Promega corporation BNL. Pefabloc SC (4-(2-aminoethyl)-benzene-sulfonylfluorid hydrochlorid) was purchased from Merck. Activated Thiol-Sepharose, Fastflow Q-Sepharose and HiLoad Q-Sepharose (preparative grade), Sephacryl S400 (preparative grade), Superose-6 (analytical grade) and Superdex-200 (analytical grade) were purchased from Pharmacia Fine Chemicals. The sequencing primers

were from Pharmacia Biotech. Chymotrypsin and trypsin were from Sigma. Endoproteinase Glu-C was from Boehringer. Molecular weight markers were from Pharmacia Fine Chemicals. All other chemicals used were of analytical grade.

The *E. coli* strain TG2, a *recA*$^-$ version of TG1 F' *tra D36 lacI*^{q} *Δ(lacZ) M15 proA*$^+$*B*$^+$*/supE Δ(hsdM-m*$_{cr}$*B) 5* (r_K^- m_K^- $M_{cr}B^-$) *thi Δ(lac-proAB)*, was used (Gibson, 1984). The plasmids pUC9 and pUC18 were used as cloning vectors (Vieira & Messing, 1982).

DNA cloning

Standard DNA operations were performed as described (Ausubel, 1987). For construction of an overexpression system of *A. vinelandii* E1p the plasmid pRA177 (Hanemaaijer et al., 1988) was partially digested with *Sau*3A. Fragments of 2-3 and 3-4 kb were isolated and ligated in the *Bam*HI site of pUC9. A clone containing the gene encoding E1p was found by screening with *A. vinelandii* E1p antiserum. This clone (p205) was completely sequenced. Digestion of p205 with *Dde*I rendered a fragment of 2688 bp encoding for E1p with only 30 bp before the start codon and none after the stop codon. The ends of this fragment were filled up with Klenow polymerase, ligated in the *Bam*HI site of pUC18 (pAFH001) and transformed into *E. coli* TG2. Colonies were screened for E1p expression using SDS-PAGE and Western blotting using E1p antibodies.

DNA sequence determination and analysis

The DNA-sequence of the double-stranded plasmid p205 was determined using oligonucleotide primers specifically annealing to the plasmid. The sequencing primers used are shown in figure 1. The DNA-sequence was determined using an Applied Biosystems 373 DNA sequencer-stretch, wrt 48 cm. The sequencing kit used was the ABI PRISM dye terminator sequencing ready reaction kit with amplitaq DNA polymerase.

Expression

E. coli TG2 harboring the recombinant plasmid pAFH001 were grown at 35 °C or 37 °C in TY medium, containing 75 μg/ml ampicillin and 20 μg/ml IPTG, for 12 to 20 h. For analysis 5.0 ml samples were taken from the cultures. The cells were harvested by centrifugation and

resuspended in 400 µl 50 mM potassium phosphate buffer, pH 7.0, containing 0.4 mM thiamin diphosphate, 0.4 mM $MgCl_2$, 25 µM EDTA, 50 µM Pefabloc and 0.02% NaN_3 (standard buffer). The cells were disrupted by sonification, after which samples of the supernatant and pellet were separately subjected to SDS-PAGE, or cells were directly boiled in SDS-PAGE-sample buffer and subjected to SDS-PAGE.

Isolation

A single colony from *E. coli* TG2 (pAFH001) was picked from a TY-plate containing 50 µg/ml ampicillin and resuspended in 3 ml TY. 0.2 ml of this suspension was used to inoculate each of ten 3-liter flasks containing 500 ml TY, ampicillin (75 µg/ml) and IPTG (20 µg/ml). This culture was grown for 14 hours at 35 °C. Using these conditions 80-90% of the expressed E1p was soluble. Cells were harvested and used for isolation of E1p. All steps described below were carried out at 4 °C unless mentioned otherwise.

The harvested cells were resuspended in ice-cold standard buffer and disrupted using a French press. The presence of E1p was monitored using SDS-PAGE. The disrupted cells were centrifuged at 33000 x g for 30 minutes, the resulting supernatant was fractionated by protamine sulphate precipitation (2 mg/ml final concentration) and again centrifuged at 33000 x g for 30 minutes. The resulting supernatant was subjected to a two step ammonium sulphate fractionation. The fraction between 20% and 50% saturation was collected by centrifugation, dissolved in standard buffer (20 mM potassium phosphate) and directly loaded onto a Fastflow Q-Sepharose column. The column was washed with standard buffer (20 mM potassium phosphate) and proteins were eluted with a linear gradient of 0-0.6 M KCl in the washing buffer. E1p eluted from the column at 0.40-0.42 M KCl. Fractions containing E1p were pooled and concentrated by ultrafiltration (Filtron Ω100) and loaded onto a Sephacryl S400 column. Proteins were eluted in standard buffer, containing 150 mM KCl. Fractions containing E1p were again pooled and loaded onto a HiLoad Q-Sepharose column. The column was washed and eluted as described for the Fastflow Q-Sepharose. Fractions containing E1p were pooled, concentrated to a 10 mg/ml solution and stored at 4 °C. All solutions including protein solutions were kept sterile by filtration using a sterile 0.22 µm filter.

Activity assays

For assaying the E1p component two methods were used. Firstly the oxidative decarboxylation of pyruvate was measured using 2,6-dichlorophenol-indophenol (Cl_2Ind) as an artificial electron acceptor (Khailova et al., 1977). The activity was measured in 200 µl 50 mM Tris·HCl, pH 7.5, 0.1 mM $MgCl_2$, 0.1 mM thiamin diphosphate, 20 µM Cl_2Ind and 5.0 mM pyruvate. To this mixture 5 -15 µl enzyme was added. The absorbance decrease at 600 nm due to the oxidation of Cl_2Ind was followed in time ($\varepsilon = 21.7\ M^{-1}cm^{-1}$) on a Pharmacia Ultrospec III spectrophotometer. One unit of activity is defined as the amount of enzyme required for the oxidation of 1.0 µmole of Cl_2Ind per minute. Secondly, the E1 activity was measured by reconstitution of the overall PDHC-activity. The PDHC-activity was assayed spectrophotometrically at 340 nm and 25°C (Schwartz & Reed, 1970). For this purpose PDHC was isolated as described in (Bosma et al., 1984) and the E2p and E3 components were separated and isolated after immobilization on Thiol-Sepharose (De Graaf-Hess & De Kok, 1982).

The irreversible inactivation of E1p by incubation at different pH values was measured after incubating 1 ml E1p for 15 min in 9 ml 50 mM potassium phosphate, 50 mM ethanolamine, 0.1 mM thiamin diphosphate, 0.1 mM $MgCl_2$, 50 mM Pefabloc, 25 mM EDTA, pH 7.4 to pH 10 at room temperature. After 15 min either PDHC activity was reconstituted by combining the E2p and E3 component with the E1p component in a cuvette or Cl_2Ind activity was measured as described above.

Limited proteolysis

The isolated *A. vinelandii* E1p component (1.0 mg/ml) was incubated with trypsin, chymotrypsin or endoproteinase Glu-C (10 µg/ml) in standard buffer without Pefabloc at 0°C, either nothing, 5 mM pyruvate or equimolar amounts of E2p were added to this mixture. Samples were withdrawn at timed intervals and the proteolysis was stopped by adding either an equal volume of sample buffer (80°C) for SDS-PAGE, 100 µM Pefabloc (trypsin or chymotrypsin) or 330 mM L-1-chloro-3-(4-tosylamido)-7-amino-2-heptanone-HCl (endoproteinase Glu-C) for analytical gelfiltration or activity measurements.

N-terminal sequence determination

The samples to be analysed were applied to SDS-PAGE according to Laemmli (Laemmli, 1970). Separated protein bands were blotted onto polyvinylidene difluoride membranes. The protein sequencing was carried out by Edman-degradation on an automated sequenator (Model 477A, Applied Biosystems).

Analytical size-exclusion chromatography

The interaction of E1p with E2p was studied using two different methods. Firstly, the dissociation of the 24-meric E2p core into trimers upon binding of E1p was monitored on Superose-6 and Superdex 200 gelfiltration columns using an ÄKTA explorer system (Pharmacia Biotech). The proteins were eluted with standard buffer, containing 150 mM KCl (Schulze et al., 1991c). Secondly, the PDHC activity, reconstituted from wild type E2p, wild type E3 and overexpressed E1p was assayed as described above.

Samples obtained from limited proteolysis experiments and from incubations at various pH were also analysed on Superose 6 and Superdex 200 columns. Both columns were calibrated using the following proteins: cytochrome c (12.4 kDa), myoglobin (17.8 kDa), chymotrypsinogen (25 kDa), bovine albumin (67 kDa), alcohol dehydrogenase (144 kDa), catalase (232 kDa), pyruvate kinase (237 kDa), vanilyl alcohol oxidase (500 kDa), citrate lyase (550 kDa) and dextran blue (2000 kDa).

Others

SDS-PAGE was performed as described by Schägger and von Jagow (Schägger & Von Jagow, 1987). For the purpose of Western blotting SDS-PAGE was performed as described by Laemmli (Laemmli, 1970). Protein concentrations were estimated by the microbiuret method (Goa, 1953), Bovine serum albumin was used as a standard.

For the production of antiserum E1p was extracted from SDS gels as described in (Hanemaaijer et al., 1987) and mixed with Freund's incomplete adjuvant. Male New Zealand White rabbits were used for immunisation. After five weeks a booster injection was given and ten days later the antisera were collected. The monoclonal antibodies 18A9 against *E. coli* E1p were kindly provided by Prof. F. Jordan (McNally et al., 1995b; McNally et al., 1995a).

Results and discussion

Cloning strategy

From partially digested pRA177 fragments of 3-4 kbp were isolated and ligated into the *Bam*HI site of pUC9. *E. coli* TG2 cells transformed with these plasmids were screened using antiserum raised against *A. vinelandii* E1p. Of 400 colonies screened 10 positives were found, 2 of which produced the intact E1p as judged by SDS-PAGE and western blotting. The plasmid containing an insert of 3301 bp (p205) was used for sequence analysis. Expression of E1p in *E. coli* TG2 harboring the recombinant plasmid p205 could be detected by Western blotting, but no clear protein band could be seen on SDS-PAGE or activity could be measured through reconstitution with E2p and E3. For the purpose of overexpression of E1p a new construct was made. To have E1p expressed under the pUC18 promoter region, the p205 insert was truncated by digestion with *Dde*I and ligated into pUC18. This gave a fragment with only 30 bp before the start codon and 3 bp after the stop codon. The 30 bp before the start codon did include the -10 ribosome binding site region but not the -35 region (figure 1). The cell-free extract of a colony of this construct (pAFH001) showed a thick band on SDS-PAGE at about 100 kDa, which cross-reacted with *A. vinelandii* E1p antiserum. Also an increase in PDHC activity could be measured when *A. vinelandii* E2p and E3 were added to the cell-free extract. This increase in PDHC activity can only be due to the presence of *A. vinelandii* E1p, since *E. coli* E1p is not able to reconstitute PDHC activity if combined with *A. vinelandii* E2 and E3 components.

Sequence

The 3301-bp insert in p205 was completely sequenced in both directions using oligo-nucleotide primers designed to specifically anneal to the double stranded p205 (figure 1). A 2655-bp open reading frame was found just downstream of the gene encoding E2p and identified as the gene encoding the pyruvate dehydrogenase component of PDHC. The deduced amino acid sequence showed a high degree of identity with other known E1p sequences from gram-negative bacteria, e.g. 58 % identity with *E. coli* E1p. None or very low similarity could be found with heterotetrameric E1p and E1b or with homodimeric E1o. The ThDP binding motif (Hawkins et al., 1989), was also found in the *A. vinelandii* E1p, starting at residue 227. Even though this motif is shared by all ThDP-dependent enzymes its location

in the amino acid sequence is different for the various enzymes (Hawkins et al., 1989; Berg et al., 1996).

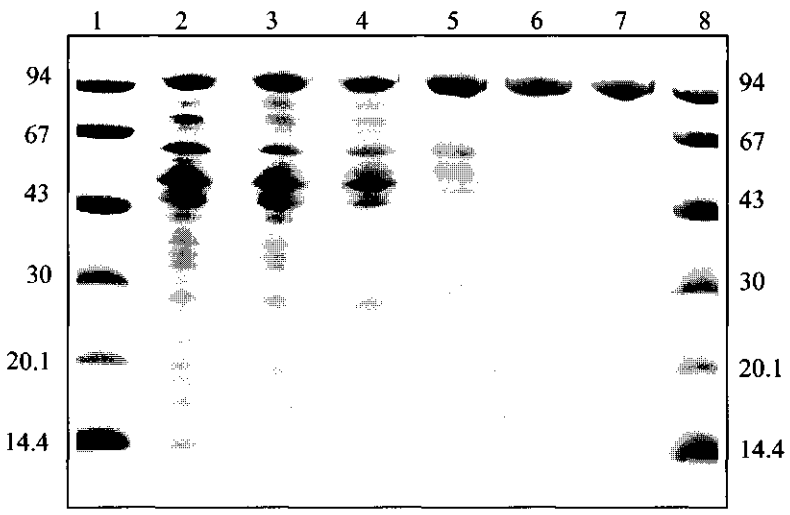


Figure 2. Purification of pyruvate dehydrogenase from *A. vinelandii*, expressed in *E. coli* TG2. Fractions were analysed using SDS-PAGE. Lane 1 and 8, marker proteins; lane 2, cell extract; lane 3, supernatant after protamine sulphate treatment; lane 4, resuspended fraction after ammonium sulphate fractionation; lane 5, fractions pooled from FastFlow Q-Sepharose chromatography; lane 6, fractions pooled from Sephacryl S400 chromatography; lane 7, fractions pooled from HiLoad Q-Sepharose chromatography.

Expression

Both complete cells, cell pellet and cell-free extract from cultures grown at 35 and 37 °C for 14 to 20 hours were applied to SDS-PAGE to determine optimal growth conditions. The amount of expressed E1p is equal under all conditions, but the amount of soluble E1p is highest (80-90%) when the cells are grown at 35 °C for 14 hours. If grown at 37 °C or at 35 °C for 20 hours the expression level of E1p is equally high, but 60-90% of the enzyme is found in the pellet and cannot be dissolved, indicating E1p to be present in inclusion bodies. Thus, the expression of soluble E1p is highly sensitive to small changes in the growth conditions.

Purification

A. vinelandii E1p was expressed and purified as described in Materials and Methods. 30-g of *E. coli* TG2 cells harboring the plasmid pAFH001 yielded 63 mg of purified E1p. It is very important that the disruption of the cells and the precipitation steps are carried out very fast and at 0 °C, because the expressed E1p is highly sensitive to protease degradation. Only the water soluble Pefabloc was used as a protease inhibitor. Protease inhibitor mixtures, e.g. Complete Protease inhibitor mix (Boehringer), inactivated E1p.

Figure 2 shows the results of the purification of E1p as monitored using SDS-PAGE. Due to the high NADH oxidase activity and the *E. coli* PDHC activity in the crude preparations it is not possible to accurately measure PDHC-activity in the first steps of the purification. The measurement of the E1 activity in the crude preparations is not possible because of the low specific activity of E1p in this assay. For these reasons only the specific activities of the purified enzymes have been determined. The specific activity of purified E1p using Cl_2Ind as artificial electron acceptor is 0.086 U/mg, which is comparable to wild-type E1p activity (0.082 U/mg). Reconstitution of the overall PDHC activity results in a specific activity of 28.9 U/mg E2p, at an optimum molar ratio of 1.9 dimers E1p per trimer E2p and dimer E3. The wild type activity at comparable conditions is 30 U/mg E2p, with an optimum ratio of 2 dimers E1p per trimer E2p and dimer E3. From these results it can be concluded that the cloned E1p is fully active.

Identification

The determined N-terminal sequence, using the Edman-degradation method, of the cloned and expressed E1p is DMQDLDP. The determined sequence of wild type E1p is MQDMQDLDPI. The size of the cloned E1p as determined by electronspray mass determination is 99863 ± 127 Da, but the calculated molecular mass of E1p is 100051 Da. This also indicates that the cloned E1p lacks one or two N-terminal residues. It is not clear what effect causes this truncation at the N-terminus, but clearly it has no influence on the activity of E1p nor on other properties (see below). From the alignment of the N-terminal sequences of all know homodimeric E1p it can be seen that these residues are not conserved. The observed N-terminal sequence is not found in the *E. coli* E1p amino acid sequence proving the isolated E1p being *A. vinelandii* E1p.

The interaction of the *A. vinelandii E2p* with the cloned E1p was studied by two different methods. First complex activity could be reconstituted as described in the previous section. Secondly the dissociation of the 24-subunit E2p core into trimers upon binding of E1p was monitored by analytical gelfiltration (Schulze et al., 1991a). The *A. vinelandii* E2p core consists of 24 subunits that form a cubic structure, yet it behaves in a unique way: upon binding of E1p and/or E3 the core dissociates into trimers (Mattevi et al., 1992c; Bosma, et al., 1984). Figure 3, panel A shows that the cloned E1p is, like the wild type E1p, able to dissociate the 24 meric E2p into trimers. The binding properties of the cloned E1p are thus unchanged compared to wild type E1p. These results again confirm our previous conclusion that the cloned E1p is fully functional.

Stability

When dissociated from the complex E1p from *A. vinelandii* is very unstable and inactivates in a few days. We have noticed that the expressed E1p is much more stable and keeps its activity after storage for 6 months at 4 ^{o}C in saturated ammonium sulphate, pH 7.0. Because buffers with a high pH, e.g. ethanolamine pH 9.4, are used in the dissociation procedure (De Graaf-Hess & De Kok, 1982), the effect of incubation in buffers of different pH was studied as described in Experimental procedures. Only 60% of the PDHC activity remains after incubation of E1p for 15 min. at pH 9.5. Thus at this pH E1p partially loses its ability to reconstitute PDHC activity. On the other hand the Cl_2Ind activity remains high. At pH 9.5 E1p eluted as a single peak with a retention volume corresponding to a molecular weight of approximately 120 kDa, indicating monomerization. Monomerization has also been observed with the *E. coli* E1p at high pH (Reed & Oliver, 1968b) but no evidence of irreversible loss of activity was given. Presumably only the dimer interacts with E2p, comparable to the binding of E3 to E2p (Mande et al., 1996; Schulze et al., 1991a). We conclude that the monomer undergoes time and pH-dependent structural changes and upon lowering the pH, part of E1p does not dimerize correctly and is unable to interact with E2p.

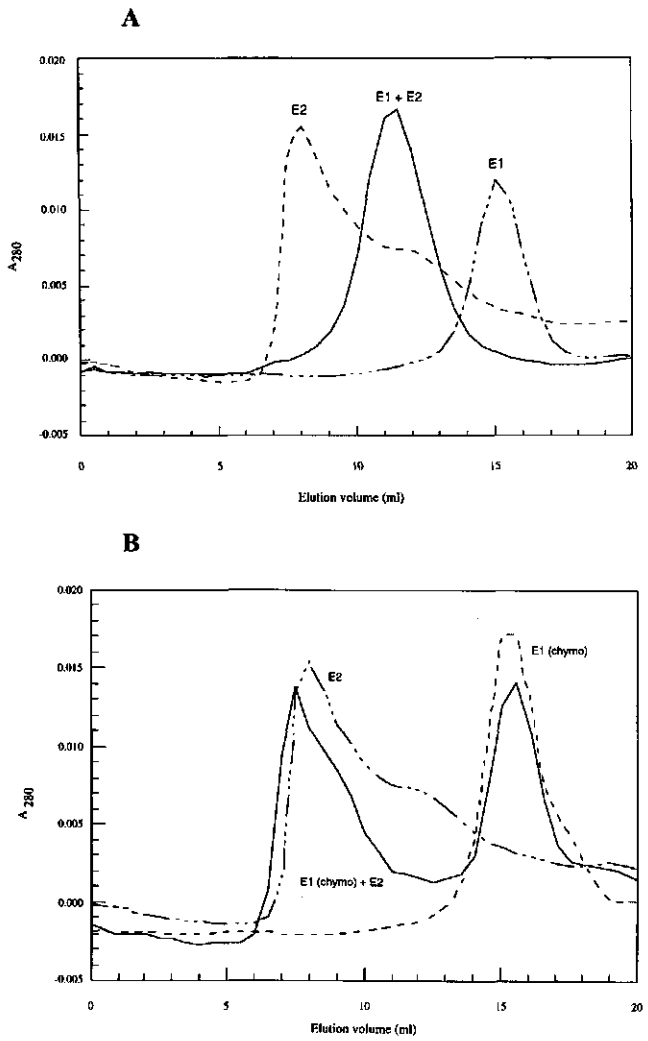


Figure 3. Binding properties of the expressed E1p component of the pyruvate dehydrogenase complex (Panel A) and of the E1p proteolysed by chymotrypsin (Panel B). E2p was added in molar ratios to the E1p samples. 50 ml samples were loaded on the Superose 6 column and eluted with 50 mM potassium phosphate buffer pH 7.0 containing 150 mM KCl and 50 mM Pefabloc.

In general, the enzyme is unstable in low ionic strength buffers and in the presence of organic solvents. Even the addition of 1-2 % ethanol results in rapid inactivation.

Kinetics

The kinetic behaviour of complex-bound E1p has been investigated previously (Bosma 1984; Bresters, De Kok & Veeger, 1975). For measuring the oxidative decarboxylation of pyruvate ferricyanide was used as an artificial electron acceptor. A Hill-coefficient of 1.4-1.6 (Bosma,

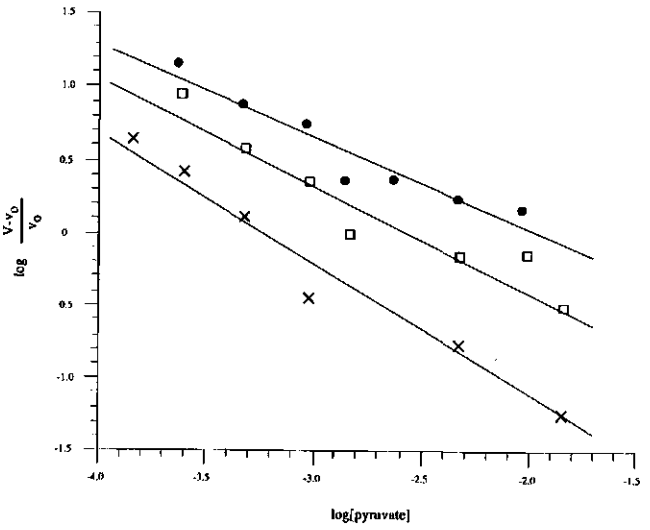


Figure 4. Hill-plots of the dependence of the rate of the Cl_2Ind reaction on the pyruvate concentration in the presence and absence of effectors. The reaction mixture is as described in the experimental procedures, pyruvate is added as indicated. (□) no effector present; (•) 45 mM AcCoA; (×) 1 mM AMP.

1984) or 1.1-1.5 (Bresters, De Kok, & Veeger, 1975) was found for the complex bound E1p. The Hill-coefficient did not change significantly in the presence of positive (AMP) or negative (acetyl CoA) effectors. Although these effects were found with the E1p reaction it could not be excluded that some of these effects were exerted through interaction with the other components of the complex. Therefore we have repeated these experiments with the expressed (non-complex-bound) E1p. In figure 4 a Hill-plot of the expressed E1p is shown. A Hill-coefficient of 1.2 - 1.6 was calculated from this plot. This demonstrates that non-complex-bound E1p shows the same cooperativity as found for complex bound E1p. The Hill-coefficients again do not seem to change significantly in the presence of negative (acetyl coenzyme A) or positive (AMP) effectors, thus the effect exerted by pyruvate is apparently not influenced by these effectors. It must be mentioned that the Cl_2Ind activity assay is, like the ferricyanide assay, quite insensitive and much more difficult to handle than the PDHC-reaction, but since all results found sofar are consistent we conclude that the observed deviation from h=1.0 is significant. The allosteric inhibition of E1p by acetyl CoA is remarkable and sofar only shown for *E. coli* PDHC (Schrenk, & Bisswanger, 1984). Acetyl CoA shows strong product inhibition (Ki 8 μM) in the PDHC reaction (Bresters De Kok, & Veeger, 1975). The results shown here indicate that, apart from interaction with the active site on E2p, acetyl CoA interacts directly with the E1p component. Due to the insensitivity of the

Cl_2Ind assay an accurate $S_{0.5}$ value could not be determined however it is estimated at about 10 µM.

Limited proteolysis

Limited proteolysis experiments were performed to obtain more information about the domain structure of the homodimeric E1p. Both proteolysis with trypsin and chymotrypsin yields two main fragments (T1, T2 and C1, C2) (figure 5). The sizes of these fragments as determined by SDS-PAGE are 40 kDa and 54 kDa for the tryptic fragments and 59 kDa and 38 kDa for the chymotryptic fragments. Size determination by MALDI was not possible because the fragments were very unstable in buffer with a KCl concentration below 150 mM. The N-terminal amino acid sequence of the four fragments was determined using Edman-degradation (figure 1).

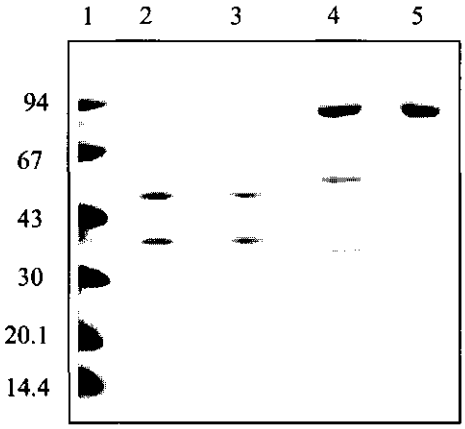


Figure 5. Limited proteolysis of cloned E1p. Lane 1, marker proteins; lane 2, incubation with trypsin; lane 3, incubation with trypsin and 5 mM pyruvate; lane 4, incubation with chymotrypsin; lane 5, incubation with chymotrypsin and 5 mM pyruvate.

Analytical gelfiltration of the samples obtained after proteolysis shows a single peak with a retention volume corresponding to a molecular weight of 140 kDa or 146 kDa respectively for trypsin and chymotrypsin. SDS-PAGE shows that the E1p in this peak is completely proteolysed by either trypsin or chymotrypsin and only the two main fragments (T1, T2 or C1, C2) are present. Intact E1p has a single peak with a retention volume corresponding to a molecular weight of 174 ± 9 kDa. From these results it can be concluded that proteolysed E1p is smaller than the intact E1p, but that it must still be present as a dimeric molecule. Thus, the proteolysed residues are not involved in the monomer-dimer interaction.

For both the trypsin and the chymotrypsin fragments it was not possible to reconstitute PDHC activity after only 30 s digestion time, while the Cl_2Ind-activity was still 100%. At this time a very small increase in migration was observed, without the appearance of other degradation products. This indicates that the active site is still present after proteolysis, but that E1p has lost its ability to interact with E2p. This is confirmed by the loss of their ability to dissociate the 24-subunit E2p core into trimers (figure 3 B).

Figure 1 shows that fragments T1 and C1 should contain the ThDP-binding motif. Western blot analysis using a monoclonal antibody (mAb) raised against *E. coli* E1p (mAb 18A9) (McNally, Motter & Jordan, 1995b; McNally, Mattsson & Jordan, 1995a) showed the highest cross reactivity against these fragments. mAb 18A9 inhibits *E. coli* PDHC for more than 98% and is most likely directed against the active site of E1p. Thus these results confirm that active site residues are most likely present in fragment C1 and T1. Moreover, E1p is partially protected against proteolysis by binding to E2p (results not shown), which confirms the conclusion that E1p is digested in a region responsible for the binding to E2p.

From the N-terminal sequences of the fragments and a comparison of fragment size as determined by SDS-PAGE and calculated from the sequence it is concluded that T1 is fragment 44-400 and fragment T2 is fragment 418-885. Likewise fragment C1 is 38-552 and C2 is 553-885. The start of fragment C1 is puzzling because digestion at R37 is not expected for chymotrypsin. Rather the sequence YLL just prior to R37 seems a more likely target. The C-termini of the fragments T1 and C1 are not well defined as there are several potential cleavage sites in this region.

Since E1p digested by both trypsin and chymotrypsin behaves in a similar way, e.g. a loss of PDHC-activity combined with a small decrease in molecular weight, it is most likely that the 38 - 40 N-terminal amino acids are involved in the binding of E1p to the E2p-component and the internal "loops" available to the proteases are not.

Experiments with endoproteinase Glu-C confirm this conclusion. This protease only cuts off a small fragment, comparable to the initial cleavage product by trypsin or chymotrypsin, leaving the remaining protein intact. The digestion causes a time dependent loss of complex activity without affecting the E1 activity, like the other proteases. The N-terminal sequence of the large fragment obtained by limited proteolysis with endoprotease Glu-C was found to be LATRTGT, starting at position 41 in the original sequence (figure1).

Previously it was shown that a point mutation in the N-terminal region of the catalytic domain of *A. vinelandii* E2p resulted in a very weak E1p binding; thus, E1p binds not only to the binding domain but also to the E2p catalytic domain (Schulze et al., 1991c). Cryoelectron microscopy of frozen-hydrated pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase complexes from *E. coli* shows that the E1 and E3 subunits seem to be separated from the surface of the E2 core by 3 - 5 nm, and thin bridges of density are seen in the gap between the E2 core and the bound subunits (Wagenknecht, Grassucci & Schaak, 1990). It is therefore assumed that E1 itself provides a linker sequence to the core. The experiments here indicate that the N-terminal part of E1p is involved in this interaction. Furthermore, incubation of E1p in potassium phosphate buffer, pH 7.0, containing $MgCl_2$ and thiamin diphosphate, with 5 mM pyruvate partially protects against proteolysis (figure 5). This indicates that E1p changes its conformation upon binding of pyruvate as expected from the cooperative behaviour of this enzyme.

Summarising, the limited proteolysis studies show that the N-terminal region of E1p is easily split off and thus it can be concluded that it has an extended conformation or forms a separate domain. Furthermore, this N-terminal region seems to be involved in the binding to E2p. The digested enzyme is still present as a dimer.

Acknowledgements: We thank Dr. R. Amons and Dr. D. van Wassenaar for determining the N-terminal amino acid sequences. We thank Dr. T. Muisers for determining the molecular mass by electron spray. We thank Dr. T. van Kampen for performing the DNA sequencing. This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

```
   1 GATCGCGCAC CCGTCTCGAG GACGAACCCG CTGTTCGCGC AGACCCGTGC GAAGCCGCGC GCCGCCGGCG

  71 GCCAGATCGG CAAAGGGCGT CGCCGGACTT GTCGCCGAGC GTGCCGGCGG GCGTCCCCGA AGCGTTGCAG

 141 GGCCTGATTG CCCAGGGTCT CCAGCGCGGG GGGGAAGTTC GATCAGCGTA GGTAGGCTCA TGCTCGTGTC

 211 CTGATATCGG CCAGACTTTC CAAAGCGGGT TGAGCCCGAC CCGTGGGGTA ATTTTACAAA AGTCCTCGGA

 281 GGGGGGGCTG AAAACCCAAT TGTTTTGTAG TAAAACTACA ACGACAGTCG GATCACCCCG GCCCGAACCA

 351 ACTAACAACA ATCCCGCAGC CGCCCATAAG GCCGGCTGTG TATCCGGGCT CCCAACTCAG GGATGCCTTC

                     M  Q  D  M  Q   D  L  D   P  I  E   T  Q  E  W   L  D  S   18
 421 CGCTATGGAG CAAGACATGC AAGACATGCA AGACCTTGAC CCCATCGAGA CCCAGGAGTG GCTGGACTCG

     L  E  S   V  L  D  H   E  G  E   E  R  A   H  Y  L  L   T  R  M   G  E  L   41
 491 CTCGAATCAG TCCTCGACCA CGAAGGCGAA GAGCGCGCCC ATTATCTGCT GACCCGGATG GGCGAACTGG

     A  T  R  T   G  T  Q   L  P  Y   A  I  T  T   P  Y  R   N  T  I   P  V  T  H   65
 561 CCACTCGCAC GGGCACCCAA CTGCCCTATG CGATCACCAC CCCCTATCGC AACACCATTC CGGTGACCCA

     E  A  H   M  P  G   D  L  F  M   E  R  R   I  R  S   L  V  R  W   N  A  L   88
 631 CGAAGCGCAC ATGCCCGGCG ACCTGTTCAT GGAACGCCGC ATCCGCTCCC TGGTGCGCTG GAATGCGCTG

     A  T  V   M  R  A  N   K  K  D   P  D  L   G  G  H  I   S  T  F   A  S  S   111
 701 GCCACGGTCA TGCGGGCGAA CAAGAAGGAC CCGGACCTGG GTGGACACAT CTCCACCTTC GCCTCCAGCG

     A  T  L  Y   D  I  G   F  N  Y   F  F  Q  A   P  T  A   E  H  G   G  D  L  I  135
 771 CCACGCTCTA CGATATCGGC TTCAACTATT TCTTCCAGGC GCCCACGGCA GAACACGGCG GCGACCTGAT

     Y  F  Q   G  Q  A   A  P  G  V   Y  A  R   A  F  L   E  G  R  I   S  E  A   158
 841 CTACTTCCAG GGCCAGGCCG CTCCCGGCGT CTACGCCCGC GCCTTCCTCG AAGGCCGCAT CAGCGAAGCG

     Q  L  Q   Q  F  R  Q   E  V  D   G  D  G   L  S  S  Y   P  H  P   H  L  M   181
 911 CAACTGCAGC AGTTCCGGCA GGAAGTCGAC GGCGACGGCC TGTCCTCCTA TCCGCACCCG CACCTGATGC

     P  D  F  W   Q  F  P   T  V  S   M  G  L  G   P  I  Q   A  I  Y   Q  A  R  F  205
 981 CGGACTTCTG GCAGTTCCCC ACCGTTTCCA TGGGGCTGGG CCCGATCCAG GCGATCTACC AGGCGCGCTT

     M  K  Y   L  E  H   R  G  F  I   P  A  G   K  Q  K   V  W  C  F   M  G  D   228
1051 CATGAAGTAC CTGGAGCACC GCGGCTTCAT TCCCGCCGGC AAGCAGAAGG TCTGGTGCTT CATGGGCGAC

     G  E  C   D  E  P  E   S  L  G   A  I  S   L  A  G  R   E  K  L   D  N  L   251
1121 GGCGAATGCG ACGAGCCGGA ATCCCTCGGC GCCATCTCCC TGGCCGGCCG CGAGAAGCTG GACAACCTGA

     I  F  V  I   N  C  N   L  Q  R   L  D  G  P   V  R  G   N  G  K   I  I  Q  E  275
1191 TCTTCGTCAT CAACTGCAAC CTGCAGCGCC TCGACGGCCC GGTCCGCGGT AACGGCAAGA TCATCCAGGA

     L  E  G   V  F  R   G  A  Q  W   N  V  N   K  V  V   W  G  R  F   W  D  P   298
1261 ACTCGAAGGG GTGTTCCGCG GCGCCCAGTG GAACGTCAAC AAGGTGGTCT GGGGCCGCTT CTGGGACCCG

     L  F  A   K  D  H  A   G  L  L   Q  Q  R   M  D  E  V   V  D  G   D  Y  Q   321
1331 CTGTTCGCCA AGGACCATGC CGGCCTGCTG CAGCAGCGCA TGGACGAGGT GGTCGACGGC GACTACCAGA

     N  Y  K  A   K  D  G   A  F  V   R  K  H  F   F  G  A   R  P  E   L  L  E  L  345
1401 ACTACAAGGC CAAGGACGGA GCCTTCGTAC GCAAGCATTT CTTCGGCGCG CGCCCCGAGC TGCTGGAGTT

     V  K  D   M  S  D   E  D  I  W   K  L  N   R  G  G   H  D  P  Y   K  V  Y   368
1471 GGTCAAGGAC ATGTCCGACG AGGACATCTG GAAGCTCAAC CGTGGCGGCC ACGACCCCTA CAAGGTCTAC

     A  A  Y   H  Q  A  V   N  H  Q   G  Q  P   S  V  I  L   A  K  T   I  K  G   391
1541 GCAGCCTATC ACCAGGCGGT CAACCACCAG GGCCAGCCGA GCGTGATCCT CGCCAAGACC ATCAAGGGCT

     Y  G  T  G   A  G  E   A  K  N   I  A  H  N   V  K  K   V  D  V   E  S  L  K  415
1611 ACGGCACCGG CGCCGGCGAG GCGAAGAACA TCGCCCACAA CGTGAAGAAG GTCGATGTCG AGAGCCTCAA

     L  F  R   D  K  F   D  V  P  L   K  D  E   E  L  E   D  L  P  F   Y  R  P   438
1681 GCTGTTCCGC GACAAGTTCG ACGTCCCGCT GAAGGACGAG GAGCTGGAAG ACCTGCCGTT CTACCGGCCG

     D  E  N   S  P  E  M   K  Y  L   R  S  R   R  E  A  L   G  G  F   V  P  Q   461
1751 GACGAGAACA GCCCGGAGAT GAAGTACCTG CGCAGCCGCC GCGAGGCCCT GGGCGGCTTC GTGCCGCAGC

     R  R  R  K   S  I  S   I  P  T   P  P  L  D   S  L  K   A  I  L   D  G  T  G  485
1821 GCCGCAGGAA GAGCATCAGC ATCCCGACCC CGCCGCTGGA CAGCCTCAAG GCGATCCTCG ACGGCACCGG
```

```
      D  R  E  I  S  T  T  M  A  F  V  R  I  L  A  Q  L  V  K  D  K  E  L  508
1891 CGACCGCGAA ATCTCCACCA CCATGGCCTT CGTGCGAATC CTCGCGCAAC TGGTCAAGGA CAAGGAGCTG

      G  S  R  I  V  P  I  I  P  D  E  A  R  T  F  G  M  E  G  M  F  R  Q  531
1961 GGCTCGCGCA TCGTGCCGAT CATCCCCGAC GAGGCGCGCA CCTTCGGCAT GGAAGGCATG TTCCGCCAGC

     L  G  I  Y  S  S  V  G  Q  L  Y  E  P  V  D  K  D  Q  V  M  F  Y  R  E 555
2031 TCGGCATCTA CTCCTCGGTC GGCCAGCTCT ACGAGCCGGT GGACAAGGAC CAGGTGATGT TCTATCGCGA

      D  K  K  G  Q  I  L  E  E  G  I  N  E  A  G  A  M  S  S  W  I  S  A  578
2101 GGACAAGAAG GGCCAGATCC TCGAGGAAGG CATCAACGAA GCCGGCGCCA TGTCCAGTTG GATCTCCGCG

      A  T  A  Y  S  N  H  N  Q  P  M  L  P  F  Y  V  F  Y  S  M  F  G  F  601
2171 GCGACCGCGT ACAGCAACCA CAACCAGCCG ATGCTGCCGT TCTACGTGTT CTATTCGATG TTCGGCTTCC

     Q  L  I  G  D  L  A  W  T  A  G  D  S  Q  A  R  G  F  L  I  G  G  T  A 625
2241 AACTCATCGG CGACCTCGCC TGGACGGCCG GCGACAGCCA GGCGCGCGGC TTCCTGATCG GCGGTACCGC

      G  R  T  T  L  N  G  E  G  L  Q  H  E  D  G  H  S  H  I  L  A  S  T  648
2311 CGGCCGCACC ACGCTCAACG GCGAAGGCCT GCAGCACGAG GACGGCCACA GCCACATCCT CGCCTCGACC

      I  P  N  C  R  T  Y  D  P  T  Y  A  Y  E  M  A  V  I  I  R  E  G  I  671
2381 ATCCCGAACT GTCGCACCTA CGATCCGACC TATGCCTATG AGATGGCGGT GATCATCCGC GAAGGCATCC

      R  Q  M  T  E  E  Q  Q  N  V  F  Y  Y  I  T  A  M  N  E  A  Y  T  Q  P 695
2451 GCCAGATGAC CGAGGAGCAG CAGAACGTCT TCTACTACAT CACCGCGATG AACGAGGCCT ACACCCAGCC

      A  M  P  E  G  A  E  A  G  I  V  K  G  M  Y  L  L  E  E  D  K  R  D  718
2521 GGCCATGCCG GAAGGGGCCG AGGCAGGCAT CGTCAAGGGC ATGTACCTGC TCGAGGAAGA CAAGCGCGAC

      A  A  H  H  V  Q  L  L  G  S  G  T  I  L  R  E  V  R  E  A  A  K  I  741
2591 GCCGCGCACC ATGTCCAGTT GCTCGGCAGC GGCACCATCC TGCGCGAAGT GCGCGAAGCG GCGAAGATCC

     L  R  E  D  Y  N  V  A  A  D  V  W  S  V  T  S  F  N  E  L  R  R  N  G 765
2661 TCCGCGAGGA CTACAACGTC GCCGCCGACG TGTGGAGCGT GACCAGCTTC AACGAACTGC GCCGCAACGG

      L  A  V  E  R  R  N  R  L  H  P  E  Q  K  P  E  Q  S  Y  V  E  Q  C  788
2731 CCTGGCCGTG GAACGCCGCA ACCGCCTGCA CCCGGAACAG AAGCCGGAGC AGAGCTACGT CGAGCAGTGC

      L  N  G  R  K  G  P  V  V  A  S  T  D  Y  M  K  L  F  A  D  Q  I  R  811
2801 CTGAACGGCC GCAAGGGCCC GGTGGTGGCC TCCACAGACT ACATGAAGCT GTTCGCCGAC CAGATCCGCC

     Q  W  V  P  S  R  E  Y  K  V  L  G  T  D  G  F  G  R  S  D  T  R  K  K 835
2871 AGTGGGTCCC GAGCCGCGAG TACAAGGTGC TCGGCACCGA TGGCTTCGGC CGCAGCGATA CCCGCAAGAA

      L  R  H  F  F  E  V  D  R  Y  W  V  V  L  A  A  L  E  A  L  A  D  R  858
2941 GCTGCGTCAT TTCTTCGAAG TGGACCGCTA CTGGGTTGTG CTGGCCGCGC TGGAAGCGCT GGCCGATCGT

      G  D  I  E  A  K  V  V  P  K  P  F  A  K  F  G  I  D  P  D  K  R  N  881
3011 GGCGACATCG AAGCCAAGGT GGTGCCGAAG CCATTCGCCA AGTTCGGCAT CGATCCCGAC AAGCGCAACC

     P  L  D  C  *                 S  E  I  I  R  V  P  D  I  G  G  D  G  E
3081 CCCTGGACTG CTGAGGAGAG CCATCGTGAG CGAAATCATC CGCGTACCCG ATATCGGCGG CGATGGGGAA

      V  I  E  L  L  V  K  T  G  D  L  I  E  V  E  Q  G  L  V  V  L  E  S
3151 GTCATCGAAT TGCTGGTCAA GACCGGCGAC CTCATCGAGG TCGAGCAGGG GCTGGTGGTG CTGGAGTCCG

     A  K  A  S  M  E  V  P  S  P  K  A  G  V  V  K  S  V  S  V  K  L  G  D
3221 CCAAGGCGAG CATGGAAGTT CCCAGTCCCA AGGCCGGAGT GGTCAAGAGC GTGAGCGTCA AGCTGGGCGA

      K  L  K  E  G  D  A  I
3291 CAAGCTCAAG GAAGGCGACG CGATC
```

Figure 1. Nucleotide sequence of the *A. vinelandii* pyruvate dehydrogenase gene (E1p) and surrounding regions and the deduced amino acid sequence. The deduced N-terminal amino acid residues of the dihydrolipoyl transacetylase component (E2p) are also shown. The ribosome binding site is shown in bold italics. The specific oligo nucleotide primers used for sequencing the complete gene are underlined. The restriction sites for *Dde*I (CTGAG) are marked with bold letters. The N-terminal sequences of the limited proteolyses fragments T1 (TGTQLPY), T2 (DKFDVPL), C1 (MGELATR), C2 (REDKKGQ) and G1 (LATRTGT) are marked with bold letters.

3 Pyruvate dehydrogenase from *Azotobacter vinelandii*. Properties of the N-terminally truncated enzyme.

This chapter is a modified version of Hengeveld, A.F.; Schoustra, S.E.; Westphal, A.H. & de Kok, A (1999): "Pyruvate dehydrogenase from *Azotobacter vinelandii*: Properties of the N-terminally truncated enzyme' Eur. J. Biochem. **265**, 1098-1107.

Abstract

The pyruvate dehydrogenase multienzyme complex (PDHC) catalyses the oxidative decarboxylation of pyruvate and the subsequent acetylation of coenzyme A to acetyl-CoA. Previously, limited proteolysis experiments indicated that the N-terminal region of the homodimeric pyruvate dehydrogenase (E1p) from *Azotobacter vinelandii* could be involved in the binding of E1p to the core protein (E2p) (Hengeveld, A. F., Westphal, A. H. and De Kok, A. (1997) *Eur J. Biochem.* **250**, 260-268.). To further investigate this hypothesis N-terminal deletion mutants of the E1p component of *Azotobacter vinelandii* pyruvate dehydrogenase complex were constructed and characterised.

Up to 9 N-terminal amino acids could be removed from E1p without effecting the properties of the enzyme. Truncation of up to 48 amino acids did not effect the expression or folding abilities of the enzyme, but the truncated enzymes could no longer interact with E2p. The 48 amino acids deletion mutant (E1pΔ48) is catalytically fully functional: it has a V_{max} identical to that of wild type E1p, it can reductively acetylate the lipoamide group attached to the lipoyl domain of the core enzyme (E2p) and it forms a dimeric molecule. On the other hand, the $S_{0.5}$ for pyruvate is decreased.

A heterodimer was constructed containing one subunit of wild type E1p and one subunit of E1pΔ48. From the observation that the heterodimer was not able to bind to E2p it is concluded that both N-terminal domains are needed for the binding of E1p to E2p. The interactions are thought to be mainly of electrostatic nature involving negatively charged residues on the N-terminal domains of E1p and previously identified positively charged residues on the binding and catalytic domain of E2p.

Introduction

The pyruvate dehydrogenase multienzyme complex (PDHC) from gram-negative bacteria consists of multiple copies of three different enzyme components: pyruvate dehydrogenase (E1p), dihydrolipoyl acyltransferase (E2p) and lipoamide dehydrogenase (E3). The complex catalyses the oxidative decarboxylation of pyruvate and the subsequent acetylation of

coenzyme A to acetyl-CoA (for recent reviews: Perham 1991; Mattevi et al., 1992a; Berg & De Kok, 1997; De Kok et al., 1998). The substrate specific thiamin diphosphate dependent E1p catalyses the decarboxylation of pyruvate and subsequently the reductive acetylation of the lipoamide groups attached to E2. The E2 component then transfers the acyl group to CoA. Finally, the reduced lipoyl group is reoxidised by the FAD dependent E3 component, which transfers the reduction equivalents to NAD^+.

The E2 component plays a central role in of the complex, both catalytically and structurally. This component consists of three to five domains interconnected by flexible linkers of 20 to 40 amino acids. This architecture allows a high degree of flexibility of the individual domains, required for catalysis. One to three N-terminal lipoyl domains, each carrying a lipoyl-lysine group, are attached to the E1/E3 binding domain. The C-terminal catalytic domain forms the structural core of the complex. Three-dimensional structures of the different domains of E2 have been solved either by X-ray crystallography (Mattevi et al., 1992c; Mattevi et al., 1993a) or by NMR (Berg et al., 1997; Berg et al., 1996b; Ricaud et al., 1996; Green et al., 1995; Dardel et al., 1993). Structures of E3 from several sources have been solved by X-ray crystallography (Mattevi et al., 1991; Mattevi et al., 1993b). The well-conserved E2p binding domain consists of two parallel helices connected by a short strand, a turn and a disordered loop (Kalia et al., 1993; Robien et al., 1992). The binding mode of *Bacillus stearothermophilus* E3 to the binding domain was recently solved by X-ray crystallography (Mande et al., 1996). The interactions between E3 and the binding domain are dominated by an electrostatic zipper formed by basic residues of the binding domain and acidic residues of one of the subunits of E3. Because E3 is homodimeric the binding domain can interact with both E3 subunits, but the binding site is located so close to the twofold symmetry axis that it is impossible for two molecules of the binding domain to bind simultaneously to one E3 dimer. The E2p binding domain behaves like a Janus-face protein: While E3 interacts with the N-terminal helix, mutagenesis experiments indicate that E1p from *Azotobacter vinelandii* interacts with the C-terminal part (Schulze et al., 1991a). The point mutation R416D in the catalytic domain from *A. vinelandii* also resulted in highly impaired E1p binding. Deletions in the flexible linker region between the binding and the catalytic domain had no effect on the affinity of E1p for E2p. Thus it can be concluded that the binding site for E1p consists of two regions, one located on the binding domain and one on the catalytic domain.

The E1 component exists both as a homodimer ($\alpha 2$) or a heterotetramer ($\alpha 2\beta 2$) depending on the source and type of complex. Homodimeric E1 with a molecular mass of approximately

100 kDa is found in all 2-oxoglutarate dehydrogenases (E1o) and in pyruvate dehydrogenases from gram-negative bacteria. Heterotetrameric E1 is found in branched-chain-2-oxoacid dehydrogenases (E1b) and in pyruvate dehydrogenases from eukaryotes and gram-positive bacteria. Little sequence similarity is found between the different classes of E1 (Matuda et al., 1991), with exception of the thiamin diphosphate binding motif, present in all thiamin diphosphate dependent enzymes (Hawkins et al., 1989). Until now no structural information is available for either homodimeric or heterotetrameric E1 and any information on the interaction between E1 and E2 is lacking. Previously, cryoelectron microscopy experiments showed that E1o, E1p and E3 are separated from the surface of the E2 core by 4 - 6 nm, and sometimes thin bridges of density are visible in the gap between the core and the bound subunits. The bridging density between E1 and E2 seems to represent more mass than between E2 and E3 (Wagenknecht et al., 1990; Wagenknecht et al., 1992). Studies by scanning transmission electron microscopy show that E1 is bound along the edges of the cubic core, while E3 is present on the faces of the core (Yang et al., 1985). When E1 is bound to the complex, the lipoyl domains are located mostly at the E2-E1 interfaces, while removal of E1 from the complex gives a much larger conformational freedom to the lipoyl domains (Yang et al., 1994).

E1p from *A. vinelandii* has recently been cloned and expressed in *Escherichia coli*, which makes it possible to study the interaction between E1 and E2 in more detail. Limited proteolysis experiments of *A. vinelandii* E1p showed that its N-terminal region can easily be cleaved off (Hengeveld et al., 1997). The remaining fragment is still active, but unable to bind to E2p. This suggests that E1p has an N-terminal domain that is involved in the binding to E2p, but not in catalysis. To further investigate the role of this domain in the binding to E2p N-terminal deletion mutants were made using endonuclease *Bal*-31. The characterisation of these mutants is described in this paper.

Experimental Procedures

Materials

Restriction endonucleases, Endonuclease *Bal*-31, T4-DNA-ligase and Taq polymerase were obtained from Bethesda Research Laboratory (BRL). Klenow Polymerase was obtained from Biolabs. Alkaline phosphatase (calf intestinal), Goat-anti-rabbit IgG-alkaline phosphatase and Endoproteinase GluC were obtained from Boehringer Mannheim. The QIAprep Spin

Miniprep kit and the QIAquick Nucleotide Removal kit were obtained from Qiagen. The GenecleanII kit was obtained from Westburg. α-32P-ATP was obtained from Amersham. Sodium (2-^{14}C)-pyruvate was obtained from New England Nuclear research products. Pefabloc SC (4-(2-aminoethyl)-benzene-sulfonylfluorid hydrochloride) was obtained from Merck. Activated Thiol-Sepharose, Fastflow Q-Sepharose and HiLoad Q-Sepharose (preparative grade), Sephacryl S400 (preparative grade), Superose-6 (analytical grade), Superdex-200 (analytical grade), MonoQ (prepacked, 1 ml) and molecular weight markers were obtained from Pharmacia Fine Chemicals. The sequencing primers were from Pharmacia Biotech. Chymotrypsin and trypsin were from Sigma. All other chemicals used were of analytical grade.

The *E. coli* strain TG2, a *recA-* version of TG1 [Δ*(lac⁻proAB), thi, supE, Res⁻ Mod⁻ (k), F' (traD36 proA⁺B⁺ ,lacI^q lacZ* ΔM15] was used (Gibson 1984). The plasmids pUC9 and pUC18 were used as cloning vectors (Vieira & Messing, 1982).

Construction of N-terminal deletion mutants.

Standard DNA operations were performed as described (Ausubel, 1987).

pAFH001 (Hengeveld, Westphal. & De Kok, 1997), containing the complete E1p gene, was used as the starting material for the construction of the endonuclease *Bal*-31 deletion mutants. The general approach for preparing *Bal*-31 deletion mutants was derived from (Sambrook, Fritsch, & Maniatis, 1989). The region encoding *A. vinelandii* E1p was excised from pAFH001 using *Eco*RI and *Hind*III and cloned into pUC18 digested with *Hinc*II. The resulting construct pSES01 was used for the deletion experiments. A 1829 basepair λ-DNA fragment (*Eco*RI-*Xma*I) was inserted upstream from the E1p encoding region to prevent removal of bases from the region encoding the promoter and ribosome binding region of the vector (pSES02) (figure 1). pSES02 (5 mg) was linearised with *Sma*I and incubated with *Bal*-31 nuclease (0.015 U) at 37 °C yielding a digestion speed of 12 basepairs per min. Aliquots were taken at timed intervals up to 25 minutes and the reaction was stopped by addition of EGTA, pH 9.0 up to a final concentration of 70 mM. The samples were digested by *Eco*RI to remove the remaining λ-DNA fragment, made blunt-ended, ligated and transformed in *E. coli* TG2 cells. The transformed *E. coli* TG2 cells were screened for E1p expression by Western blotting using antiserum raised against *A. vinelandii* E1p (Hanemaaijer, 1987). Positive colonies were subsequently analysed by SDS-PAGE. The plasmids were isolated and the

DNA-sequences of the N-terminal coding regions were determined using Taq polymerase according to (Sambrook, Fritsch, & Maniatis, 1989).

Enzyme expression and purification

E. coli TG2 harbouring either the recombinant plasmid pAFH001, expressing wild type E1p, or a deletion mutant were grown and purified as described in (Hengeveld et al., 1997). However, E1pΔ48 elutes at 0.2 M KCl from the Q-Sepharose column while wild type E1p elutes at 0.4 M KCl. *A. vinelandii* wild type E2p and E3 were expressed and purified from *E. coli* TG2 as described in (Hanemaaijer et al., 1989; and Westphal & De Kok, 1988).

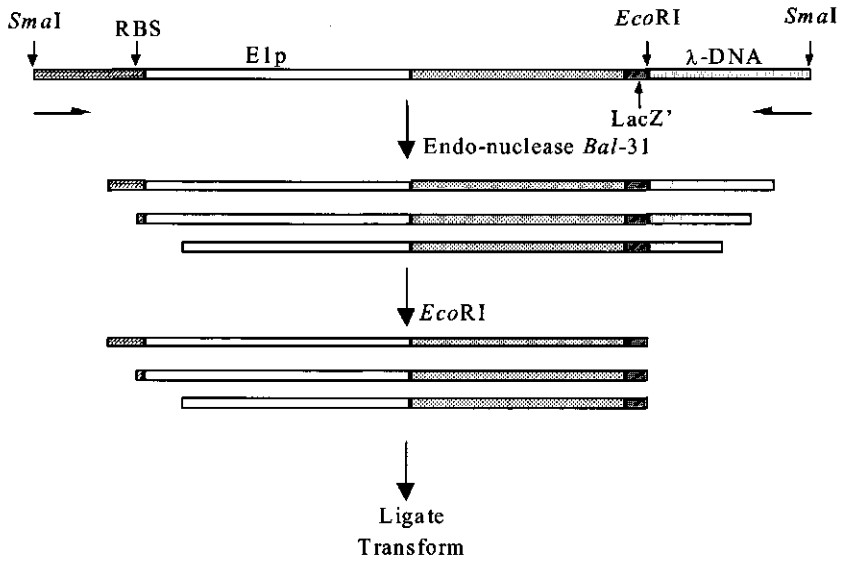


Figure 1 Scheme of the construction of the *Bal-31* deletion mutants. The top bar represents linearised pSES02. pSES02 (5 mg) was incubated with *Bal-31* nuclease (0.015 U) at 37 °C yielding a digestion speed of 12 basepairs per min.. Aliquots were taken at timed intervals up to 25 minutes. The samples were digested by *Eco*RI to remove the remaining l-DNA fragment, made blunt-ended, ligated and transformed in *E. coli* TG2 cells. RBS is the *Azotobacter vinelandii* E1p ribosome binding site. LacZ' is the promoter and ribosome binding site region of the vector and the region encoding the N-terminal part of β-galactosidase.

Activity assays

For the assay of the E1p component two methods were used. The oxidative decarboxylation of pyruvate was measured at 600 nm using 2,6-dichlorophenol-indophenol (Cl_2Ind) (ε = 21.7×10^3 $M^{-1}cm^{-1}$) as an artificial electron acceptor (Khailova et al., 1977). One unit of activity is defined as the amount of enzyme required for the oxidation of 1.0 µmole of Cl_2Ind per minute. The ability of E1p to reconstitute overall PDHC-activity was measured spectrophotometrically at 340 nm as described in (Schwartz & Reed, 1970).

The pH-inactivation of E1p and E1pΔ48 was measured by incubating the enzyme for 30 minutes at a pH ranging from 5.5 to 9.5. Subsequently the E1-activity was measured as described above.

Reductive acetylation assay

Reductive acetylation of the lipoyl domains by wild type E1p or a deletion mutant in the presence of ^{14}C-labelled pyruvate was assayed as described in (Packman et al., 1984a; Berg et al., 1998) with the following modifications. Lipoyl domain (6 µM to 60 µM) was incubated with 2.5 mg E1p in 100 ml 50 mM potassium phosphate pH 7.0, containing 0.1 mM thiamin diphosphate and 1.0 mM $MgCl_2$. The reaction was started by addition of sodium (2-^{14}C) pyruvate to a final concentration of 0.25 mM. The reaction was stopped after 30 seconds.

Interaction of E1p and E1pΔ48 with E2p.

The interaction of E1p with E2p was studied by two methods. Firstly, the dissociation of the 24-meric E2p core into trimers upon binding of E1p was monitored on Superose-6 and Superdex 200 gelfiltration columns using a ÄKTA explorer system (Pharmacia Biotech). The proteins were eluted with 50 mM potassium phosphate buffer, pH 7.0, containing 0.1 mM thiamin diphosphate (ThDP), 0.1 mM $MgCl_2$, 25 µM EDTA, 50 µM Pefabloc and 0.02% NaN_3 (standard buffer), containing 150 mM KCl (Schulze et al., 1991a; Hengeveld et al., 1997). Both columns were calibrated using the following proteins: cytochrome c (12.4 kDa), myoglobin (17.8 kDa), chymotrypsinogen (25 kDa), bovine albumin (67 kDa), alcohol dehydrogenase (144 kDa), catalase (232 kDa), pyruvate kinase (237 kDa), vanilyl alcohol oxidase (500 kDa), citrate lyase (550 kDa) and dextran blue (2000 kDa).

Secondly, the PDHC activity reconstituted from wild type E2p, wild type E3 and wild type or mutant E1p was assayed as described above.

Preparation of heterodimers

For the preparation of heterodimers containing one subunit of the wild type E1p and one subunit of a deletion mutant, wild type E1p and the deletion mutant were mixed in an equimolar ratio in standard buffer (0.4 mg/ml). The pH of this mixture was raised to 9.5 by addition of 7.5 µl ethanolamine (1% w/v) to 100 µl mixture. The mixture was incubated on ice for 15 min. Subsequently the pH was lowered to 7 by addition of potassium phosphate (1 M, pH 6.9) to a final concentration of 75 mM and incubated on ice for 1 hour. After the addition of potassium phosphate in some cases fresh ThDP was added to a final concentration of 0.1 mM. The resulting mixture of homodimers of wild type E1p, deletion mutant and heterodimers was separated on a MonoQ column (1 ml) in standard buffer with a gradient of 0 - 0.6 M KCl using an ÄKTA explorer system (Pharmacia Biotech) for monitoring. The heterodimers were used immediately for further analysis.

Others

SDS-PAGE was performed as described in (Schägger & Von Jagow, 1987). For the purpose of Western blotting SDS-PAGE was performed as described in (Laemmli, 1970). Protein concentrations were estimated using the microbiuret method (Goa, J., 1953), bovine serum albumin was used as a standard. Antiserum against *A. vinelandii* E1p was produced as described in (Hanemaaijer et al., 1987).

Limited proteolysis experiments using trypsin, chymotrypsin and endoproteinase GluC were performed as described in (Hengeveld et al., 1997).

Results and discussion

Construction of N-terminal deletion mutants

When studying the structure/function relationships of E1p it was found that the N-terminal region can easily be removed by limited proteolysis without effecting catalysis. This truncation caused the loss of its ability to bind to the core component of the complex (E2p).

Therefore, we decided to study the role of the N-terminal region in greater detail by constructing N-terminal deletion mutants of *A. vinelandii* E1p.

Incubation of pSES02 with endonuclease *Bal*-31 starting 60 basepairs upstream of the start codon of E1p (figure 1) resulted in 412 mutants, divided into 4 groups, based on the incubation time (Table 1). Because the deletion mutants lack the *A. vinelandii* ribosome binding site and the start codon, they were expressed as a protein fused to the first 6 amino acids (MTMITN) of β-galactosidase. The number of deletion mutants that result in an expressed fusion-product (Table 1) corresponds well to the expected number. The sequencing results in Table 2 show that the largest deletion expressing E1p in group III is 196 basepairs even though the group contains deletions up to 300 basepairs. This indicates that deletions beyond 196 basepairs do not result in expression of a folded protein. As seen from Table 2 a preference is seen for some truncations, mostly situated in the region where the proteolytic sites for trypsin (R44), chymotrypsin (R37) and endoproteinase GluC (E40) are found. Proteolysis sites are usually in flexible regions on the surface of a protein. Truncated proteins starting in these regions are therefore more likely to fold correctly. This could explain why many of deletions are found in a region highly susceptible to proteolysis. Since truncations of 26 and 29 amino acids are also prominent, this suggests a more flexible or loop region.

Table 1 The results of the construction of N-terminal deletion mutants using *Bal*-31, divided into groups based on the incubation time (12 basepairs sec^{-1}). The mutants were constructed as described in the experimental procedures.

Group	Number of Colonies	Deleted basepairs	Antibody screening [a]	Overexpression [b]	Sequenced [c]
I	154	0-24	129 (84%)	n.a.	0
II	128	48-96	60 (47%)	18	13
III	64	144-192	20 (31%)	13	13
IV	66	240-300	5 (8%)	4	4

[a] The number of colonies that showed E1p expression using antibodies raised against E1p, and the percentage of colonies showing E1p-expression to the total number in this group.

[b] The number of colonies that showed over-expression of E1p in SDS-PAGE. Only colonies that were found positive by antibody screening were used. n.a. means not analysed

[c] Number of mutants that were sequenced. Only mutants that showed E1p over-expression on SDS-PAGE were used.

A number of mutants were tested for their ability to reconstitute complex activity (Table 2). Up to 13 amino acids could be fused to the N-terminus without impairing the ability to reconstitute complex activity (6 from β-galactosidase + 7 upstream from the start codon of E1p). Up to 9 amino acids could be deleted without effecting this ability. Truncation from 9 up to 48 amino acids resulted in expression of folded enzyme, but in all cases the ability to reconstitute complex activity was completely lost. Truncations larger than 48 amino acids did not result in a folded enzyme. This demonstrates that almost the entire N-terminal region is necessary for the folding of a functional N-terminal domain. If only part of the N-terminal domain is removed and part is present as a misfolded or unfolded structure, this might interfere with the interactions between E1p and E2p. For this reason the mutant E1pΔ48 (-196 basepairs) was selected for further characterisation.

Expression and purification of deletion mutant E1pΔ48

The expression of the mutant E1pΔ48 differed considerably from that of the wild type enzyme (Hengeveld et al., 1997). The mutant shows a much higher expression level, while it is not found in inclusion bodies. This can be explained by the fact that wild type E1p from *A. vinelandii* binds to *E. coli* E2p, but this does not result in a completely active complex (60 %) (De Kok & Westphal, 1985; Schulze et al., 1992). E1pΔ48 lacks its N-terminal domain and thus does not bind to and inhibit the *E. coli* PDHC.

The purification procedure of E1pΔ48 was identical to that described for wild type E1p. E1pΔ48 shows an increased stability compared to wild type E1p, indicating that the instability of E1p is (partly) caused by its N-terminal domain. This could also explain the increased stability of E1p when bound to E2p.

Limited proteolysis

Proper folding of E1pΔ48 was checked by limited proteolysis studies. The proteolysis patterns were compared to those found for E1p as described previously (Hengeveld ,et al., 1997). Endoproteinase GluC does not cleave E1pΔ48. In wild type E1p it cleaves at a single site (E40) in the N-terminal region and since no extra cleavage sites are found in E1pΔ48, E1pΔ48 clearly folds in a similar way. Both trypsin and chymotrypsin cleave wild type E1p at two sites, one of them in the N-terminal region (R44 + R418 and R37 + Y553 respectively).

Because E1pΔ48 is truncated very close to the N-terminal proteolysis sites the large fragments obtained after proteolysis are expected to be almost identical (only 4 and 11 amino acids difference respectively) to those obtained for E1p. The digestion patterns obtained for E1pΔ48 with trypsin and chymotrypsin are indeed very similar to those of E1p, again showing that E1pΔ48 is correctly folded.

Interactions of E1pΔ48 with E2p

The binding of E1p and E1pΔ48 to E2p was studied by two different methods. Firstly the E2p-E3 subcomplex was titrated with either wild type E1p or the deletion mutant. For the wild type enzyme reconstitution of the overall PDHC activity resulted in a specific activity of 28.9 U/mg E2p at an optimum molar ratio of 1.9 E1p dimers per E2p trimer and E3 dimer (Hengeveld et al., 1997). Using up to 40 E1pΔ48 dimers per E2p trimer and E3 dimer we were not able to reconstitute any PDHC-activity.

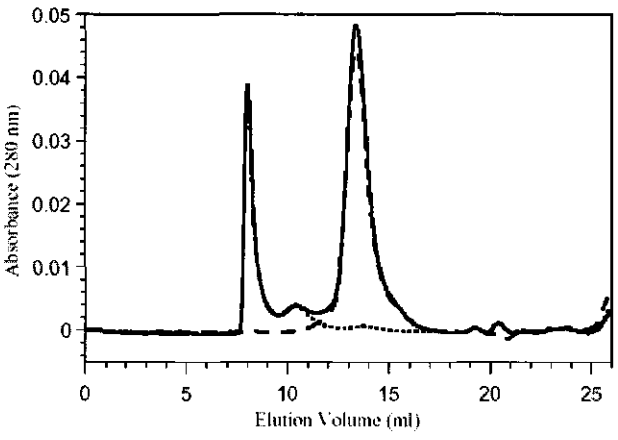


Figure 2. Binding properties of the N-terminal deletion mutant E1pΔ48. *A. vinelandii* E2p was added in equimolar ratios to the E1pΔ48 sample. 50 ml samples were loaded on a Superose 6 column and eluted with 50 mM potassium phosphate, pH 7.0, containing 150 mM KCl, 0.1 mM $MgCl_2$, 0.1 mM ThDP and 50 mM Pefabloc. — — — is E1pΔ48, ---- is E2p and ——— is E1pΔ48 + E2p.

Table 2 Ability of a number of *Bal*-31 deletion mutants to reconstitute PDHC-activity. PDHC-activity was measured by recombining cell extract of the deletion mutant with E2p and E3 from *Azotobacter vinelandii*, as described in the experimental procedures. Mutants were N-terminally fused to the amino acids "MTMITN".

[a] The 6 amino acids from β-galactosidase fused to the N-terminus are not included in this number.

Group	Number of deleted basepairs	Number of deleted/added amino acids [a]	PDHC-activity	Number of mutants
II	24-48	wild type	+	7
II	25	+9	n.a.	1
II	31	+7	+	1
II	40	+4	+	1
II	52	0	+	1
II	79	-9	+	1
II	97	-15	-	1
IV	130	-26	-	2
III	139	-29	-	3
III	151	-33	-	1
III	160	-36	-	1
III	163	-37	-	3
III	166	-38	-	1
IV	172	-40	-	1
IV	181	-43	-	4
IV	196	-48	-	1

Secondly, the binding of E1pΔ48 to E2p was monitored by analytical gelfiltration experiments. The *A. vinelandii* E2p core consists, like that of other gram-negative bacteria, of 24 subunits, forming a cubic structure. However, upon binding of E1p and/or E3 the 24-subunit core dissociates into trimers (Mattevi et al., 1992c). This property can easily be monitored by analytical gelfiltration (Schulze et al., 1991b). Figure 2 shows that E1pΔ48 has no influence on the 24-subunit core and elutes as a separate fraction. The size of E1pΔ48 was estimated from these experiments and it was evident that E1pΔ48 is present as a dimer and is only slightly smaller in size than the wild type enzyme.

The above-mentioned measurements all confirm our hypothesis that the N-terminal region of E1p is required for the binding of E1p to E2p. The fact that E1pΔ48 lacks this region, but is still correctly folded and present as a dimer, points to the presence of a separate folding domain. It is clear that the N-terminal domain is necessary for the binding of E1p to E2p, but

it is unclear if one or both N-terminal domains in a dimer are involved in the binding. For this reason a heterodimeric molecule was prepared containing one subunit Elp and one subunit ElpΔ48.

Heterodimer Preparation and Isolation

Elp can be monomerised by raising the pH (Reed. & Oliver, 1968a). If the pH is lowered, Elp dimerises, although after this procedure the stability of the enzyme is somewhat decreased as shown before (Hengeveld et al., 1997). A mixture of Elp and ElpΔ48 was monomerised by

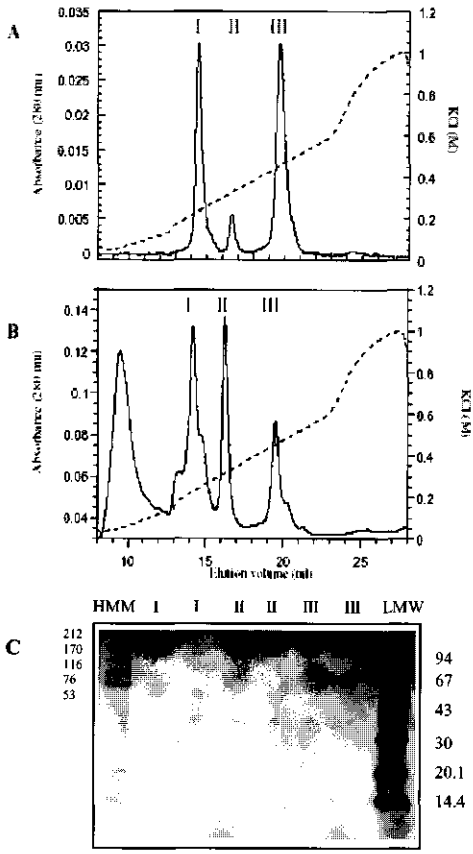


Figure 3. Preparation and isolation of heterodimeric Elp-ElpΔ48.
Heterodimeric Elp-ElpΔ48 was prepared as described in detail in the experimental procedures. The dimers were separated on a MonoQ column (1 ml) in standard buffer using a gradient of 0 - 0.6 M KCl. Panel A, 50 ml sample. Panel B, 1 ml in which, after the addition of potassium phosphate ThDP was added to a final concentration of 0.1 mM. Part of the protein was not bound to the column and washed out at 0 M KCl. Peak I contains ElpΔ48, peak II contains the heterodimer Elp-ElpΔ48 and peak III contains Elp. Panel C, SDS-PAGE of fractions of peak I, II and III from panel B. The gel was stained using silver stain. HMW is high molecular weight marker, LMW is low molecular weight marker. I are two fractions of peak I, II are two fractions of peak II and III are two fractions of peak III.

incubation at pH 9.5 for 15 minutes and subsequently dimerised by lowering the pH to 7. Theoretically the mixture should contain three dimers in a ratio 1:2:1 (Elp-Elp : Elp-ElpΔ48 : ElpΔ48-ElpΔ48).

As was already seen during the purification of ElpΔ48 the deletion mutant binds weaker to an anion exchange column than wild type enzyme. The observed difference in binding can be due

to the differences in pI as calculated from the primary sequences (5.8 (E1p) and 6.3 (E1pΔ48) respectively). The pI value of the heterodimer is expected to be in between these two values Figure 3A and 3B show elution patterns of dimer mixtures to which, after lowering the pH to 7, 0.1 mM ThDP was (fig 3B) or was not added (fig 3A). Fractions corresponding to the three peaks shown in both figures 3A and 3B were subjected to SDS-PAGE (fig 3C). Peak I shows a single protein band corresponding to E1pΔ48, peak III shows a single band corresponding to E1p. Peak II shows two bands corresponding to both E1p and E1pΔ48, present in equal amounts. Analytical gelfiltration using a Superose 6 column showed that the molecules in all three peaks were present as dimers. Fractions of all three peaks were tested for their ability to reconstitute PDHC-activity. Complex activity could only be detected using peak III. Using up to a 40 times excess of enzyme from peaks I and II did not result in the reconstitution of any complex activity. From these experiments it is clear that peak II contains heterodimeric molecules consisting of one subunit of E1p and one subunit of E1pΔ48.

Comparing figure 3A and 3B shows that in the absence of ThDP (fig 3A) during dimerisation there is a strong preference for the formation of homodimeric molecules. This preference disappears if 0.1 mM ThDP is added during dimerisation (fig 3B) and the ratio of the formed dimers is much closer to the theoretically expected 1:2:1 ratio. ThDP is rapidly hydrolysed at high pH (Sober & Harte, 1970) and therefore it might be degraded during monomerisation. Therefore, in the absence of ThDP most likely apo-enzyme is formed, while in the presence of ThDP holo-enzyme is formed. The high sequence identity (44 %) between the ThDP binding site of E1p and yeast transketolase suggest a similar binding mode of ThDP in both enzymes. In transketolase the binding of ThDP occurs on the dimer interface and involves residues of both subunits of the dimer (Lindqvist et al., 1992). In E1p ThDP influences the dimerisation, confirming the binding of ThDP at the dimer interface. The heterodimer is a very unstable molecule; concentrating it by several different methods resulted in inactivation. Due to these difficulties only small amounts of heterodimer could be obtained and used in activity measurements (see below) or binding studies. The concentration of the heterodimer is in the range where the 24-meric E2p partly dissociates into trimers. Therefore, the patterns obtained in analytical gelfiltration were not clear-cut. Using specific antibodies raised against both E1p and E2p it was shown however, that no binding occurs between the heterodimer and E2p. This demonstrates that for the binding of E1p to E2p both N-terminal domains are required.

Kinetic characterisation

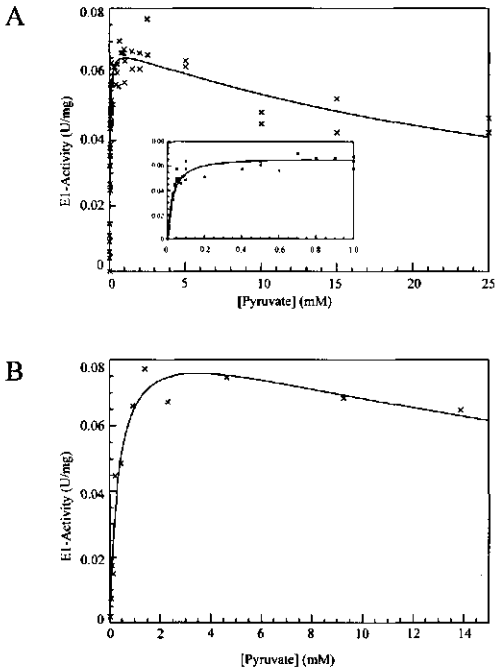


Figure 4. Substrate titration curve of E1pΔ48 and E1p in 50 mM potassium phosphate, pH 7. The reaction mixture is as described in the experimental procedures. Panel A shows the titration of E1pΔ48. The inset shows a detail of the titration of 0 - 1 mM pyruvate. Panel B shows the titration of E1p.

To analyse whether the N-terminal domain solely functions as a binding domain (analogous to the binding domain of E2p), or if it has more functions, we investigated the kinetic properties of E1pΔ48. The reaction mechanism of E1p consists of several steps, as described in the introduction. Using an artificial electron acceptor (Cl_2Ind), the oxidative decarboxylation of pyruvate can be measured. In figure 4 v/s curves of E1p and E1pΔ48 are shown. The resulting $S_{0.5}$ values, V_{max} values and Hill coefficients for E1pΔ48 at different conditions were calculated (Table 3). The kinetic behaviour of wild type E1p, both bound in the complex and free has been investigated previously (Hengeveld et al., 1997; Bosma, 1984; Bresters et al., 1975). These data are now compared to those of the deletion mutant E1pΔ48 (Table 3). From figure 4 one can see that both E1p and E1pΔ48 are inhibited by its substrate at high concentrations. The K_i values were calculated to be 31 mM (E1p) and 37 mM (E1pΔ48) respectively.

Wild type E1p and E1pΔ48 both have a similar V_{max}. The Hill-coefficients of E1pΔ48, indicative of the allosteric nature of the enzyme, remain the same within the experimental error. Besides this, the pH-inactivation curves of E1p and E1pΔ48 are very similar (results not

shown) with an optimum activity if incubated at pH 7. These experiments indicate that the deletion mutant is not only correctly folded but that is also has the same catalytic properties as wild type E1p.

Surprisingly however, the $S_{0.5}$ values of E1pΔ48 all decreased an order of magnitude compared to those of wild type E1p. Both AMP and potassium phosphate have a strong decreasing effect on the $S_{0.5}$ values of both E1p and E1pΔ48. The inhibitor acetyl CoA decreases the $S_{0.5}$ values for E1p while for E1pΔ48 the $S_{0.5}$ value increases. The $S_{0.5}$ value of E1pΔ48 in the absence of effectors (0.24 mM) is very similar to that of E1p in the presence of an activator (0.2 mM - 0.4 mM). It seems therefore, that the removal of the N-terminal domain results in activation of E1p. E1pΔ48 in this "activated conformation" has retained most of the kinetic properties of wild type enzyme; effectors show similar effects on the enzyme and its allosteric properties do not seem to have changed significantly. Moreover, the inhibition of E1p by its substrate does not change significantly due to truncations of the N-terminal region. Slight alterations in the Hill-coefficients could possibly be due to the insensitivity and inaccuracy of the Cl_2Ind-assay. On the other hand, the proximity of the highly negatively charged N-terminal domain (P_i 4.0) to the active site in wild type E1p might decrease the effective concentration of pyruvate resulting in an apparent higher $S_{0.5}$

Table 3 Summary of the kinetic parameters for the E1-reaction by *Azotobacter vinelandii* E1p and E1pΔ48. $S_{0.5}$ and V_{max} values are calculated from the v/s curves in the presents of several different effectors. The Hill coefficients are calculated from the log(s) / log(V_{max}-v/v) curve. The reactions were performed as described under the experimental procedures.

	$S_{0.5}$ (mM)		Hill coefficient		V_{max} (U/mg)	
	E1p [1)]	E1pΔ48	E1p [1)]	E1pΔ48	E1p [1)]	E1pΔ48
Tris, pH 7.4	2.0	0.2	1.3 ± 0.2	1.0 ± 0.2	0.07 ± 0.01	0.06 ± 0.01
KPi, pH 6.9	0.2	0.03	1.2 ± 0.2	1.4 ± 0.2	0.08 ± 0.01	0.07 ± 0.01
KPi, pH 7.4	n.a.	0.05	n.a.	1.4 ± 0.2	n.a.	0.07 ± 0.01
Tris&KPi, pH 7.4	n.a.	0.01	n.a.	1.5 ± 0.2	n.a.	0.06 ± 0.01
Tris&acetylCoA, pH 7.4	1.5	0.5	1.3 ± 0.2	1.2 ± 0.2	0.09 ± 0.01	0.07 ± 0.01
Tris&AMP, pH 7.4	0.4	0.02	0.8 ± 0.1	0.6 ± 0.1	0.1 ± 0.02	0.04 ± 0.01

1) Kinetic constants for E1p were partially obtained from (24). n.a. is not analysed

The kinetic characterisation of the heterodimer shows that it is catalytically functional, although quite unstable. The V_{max} (0.06 U/mg) is very similar to that of both E1p and E1pΔ48. The low concentration and the instability of the heterodimer, combined with the insensitivity of the Cl_2Ind-assay resulted in rather inaccurate data in the v/s curve, so that no conclusions on the $S_{0.5}$ could be drawn.

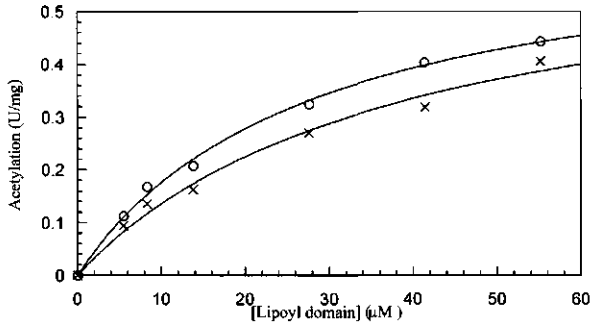


Figure 5. Reductive acetylation of the lipoyl domain of *A. vinelandii*, catalysed by E1p or E1pΔ48. The reaction mixture was as described in the experimental procedures. (x) E1p and (o) E1pΔ48.

The second part of the E1-reaction, the transfer of the acyl group from hydroxy-ethyl-ThDP to the free lipoylated lipoyl domain was measured using radioactive labelled pyruvate at different lipoyl domain concentrations (figure 5). K_m and k_{cat} values for free lipoyl domain found for E1pΔ48 (24 μM and 1.1 s^{-1}) are very similar to those found for E1p (27 μM and 0.9 s^{-1}). The removal of the N-terminal domain clearly has no influence on the interactions of E1p with the lipoyl domain and on catalysis in the acyl-transfer. The fact that the K_m and k_{cat} values for complex-bound E1p (K_m 43 μM and a k_{cat} 0.8 s^{-1} (Berg et al., 1998)) and non-complex-bound are highly comparable also confirm this: the binding of E1p to E2p has no influence on the interactions of E1p with free lipoyl domain.

Summarising, the deletion mutant E1pΔ48 is catalytically fully functional. It can both decarboxylate pyruvate and transfer the acyl moiety to the lipoyl group. On the other hand, removal of the N-terminal domain decreases the $S_{0.5}$ values for pyruvate, indicating the proximity of the N-terminal domain to the active site.

Model of the interactions between E1p and E2p The N-terminal domain is highly conserved among all E1p's from gram-negative bacteria (figure 6). Previously, it was stated that no significant homology could be found between different ThDP-dependent enzymes, except in the ThDP-binding motif (Matuda et al., 1991; Hawkins et al., 1989). An overall identity of 24% was found however between *A. vinelandii* E1p and yeast transketolase, even though E1p is 884 amino acids long and transketolase only 680 amino acids. Amino acids 1-70 from E1p

are completely absent in transketolase, indicating that these residues have a function that is absent in transketolase. The proposed binding function of the N-terminal region of E1p is in good agreement with this. Surprisingly the N-terminal sequence from E1p is completely absent in heterotetrameric E1, indicative of a different binding mode in these enzymes.

The secondary structure prediction program PHDsec (Rost & Sander, 1993) predicts a structure as shown in figure 6. Amino acids 1-9, which could be removed from E1p without effecting its binding properties, are predicted to form a loop-like structure. Amino acids 13-25 and 28-44 are predicted to form two helices. It is likely that this region forms the core of the N-terminal domain; two helices connected by a small loop. From table 2 it can be seen that from amino acid 15- 26 no truncated enzymes were found, while truncated enzymes starting at residues 26 and 29 are very prominent, indicating an exposed region around residues 26 and 29. Finally, a second loop region is predicted, separating the N-terminal domain from the

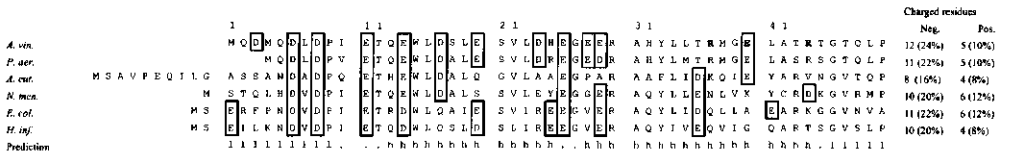


Figure 6. Aligned N-termini of E1p from several Gram-negative bacteria. *A. vin. Azotobacter vinelandii* E1p (Hengeveld et al., 1997); *P. aer. Pseudomonas aeruginosa* E1p (Rae et al., 1997); *A. eut. Alcaligenus eutrophus* E1p (Hein, & Steinbüchel, 1994); *N. men. Neisseria meningitidis* E1p (Ala' Aldeen et al., 1996); *E Col. Escherichia coli* E1p (Stephens et al., 1983); *H. inf. Haemophilus influenzae* E1p (Roe, Clifton, Dyer, Genococcal Genome Sequencing Project). Acidic residues are shown in boxes. The proteolytic sites for trypsin (R44), chymotrypsin (R37) and endoproteinase GluC (E40) are shown in bold. The bottom line shows the secondary structure as given by the secondary structure prediction program PHDsec (Rost & Sander, 1993). "l" is loop, "h" is helix, "• " means no prediction is made for this residue. On the right the amount and percentage (between brackets) of positively and negatively charged residues in the N-terminal region is shown.

remaining part of the enzyme. Analysis of the structure of a synthetic peptide representing the N-terminal sequence of E1p is in progress.

From previous experiments (see introduction) it was concluded that the binding site for E1p on E2p consists of two regions, one on the binding domain and one on the catalytic domain. In both regions, as shown by mutagenesis (Schulze et al., 1991a), positively charged residues are involved in the interactions (figure 6). The N-terminal region of E1p is highly negatively

charged. *A. vinelandii* E1p contains 24 % acidic residues in the first 50 amino acids compared to an average of 14 % in the complete enzyme. The amount of basic residues is relatively low (10% in the first 50 amino acids, compared to 14 % in the complete enzyme).

Residues E14, D17, E20, D24 and E28 would, according to the secondary structure prediction, all be present in a helix and moreover all on the same side of this helix. This indicates that these residues might be involved in the interactions between E1p and E2p. Site directed mutagenesis studies are in progress to confirm this hypothesis.

Binding studies using the heterodimeric E1p-E1pΔ48 showed that both N-terminal domains are essential for the binding of E1p to E2p. Because binding occurs to both the binding domain and the catalytic domain of E2p (Schulze et al., 1991a), we propose that one N-terminal domain binds to the binding domain, while the other one interacts with the catalytic domain. Most likely the acidic residues on the N-terminal domain and basic residues (Schulze et al., 1991a) on E2p are involved these interactions. The observed stochiometry in *A. vinelandii* PDHC (three E1p dimers per E2 trimer) (Bosma , 1984) seems to exclude a symmetric binding mode in which one dimer of E1p interacts with two E2p binding domains. Apparently an asymmetric binding mode is preferred, as is the case with E3 (Mande et al., 1996).

Summarising we can conclude that E1p has an N-terminal domain functioning as the binding domain for its binding to E2p. The E1p dimer binds to E2p through interactions of this domain with both the binding domain of E2p and the catalytic domain of E2p. Besides its binding function the N-terminal domain also influences the kinetic behaviour of E1p, indicating that the N-terminal domain is in proximity of the active site.

Acknowledgements: This investigation was supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

4 Functional and structural characterisation of the N-terminal domain of prokaryotic pyruvate dehydrogenase.

A modified version of this chapter has been submitted for publication to Biochemistry:
Annechien F. Hengeveld , Carlo P. M. van Mierlo, Henno W. van den Hooven , Antonie J. W. G. Visser, Aart de Kok. Functional and structural characterisation of the N-terminal domain of prokaryotic pyruvate dehydrogenase.

Abstract

We used a synthetic peptide (Nterm-E1p) to characterise the structure and function of the N-terminal region (amino acid residues 4-45) of the pyruvate dehydrogenase component (E1p) from the pyruvate dehydrogenase multienzyme complex (PDHC) from *Azotobacter vinelandii*. Activity and binding studies established that Nterm-E1p specifically competes with E1p for binding to the dihydrolipoyl acyltransferase component (E2p) of PDHC. Moreover, the experiments show that the N-terminal region of E1p forms an independent folding domain that functions as a binding domain. CD measurements, 2D-^{1}H NMR analysis and secondary structure prediction all indicate that Nterm-E1p has a high a-helical content. Here we propose a structural model of the N-terminal domain.

We demonstrate that the peptide is present in two conformations, the population of which depends on the sample conditions. The conformations are designated "unfolded" at a pH above 6 and "folded" at a pH below 5. The 2D-^{1}H TOCSY spectrum of a mixture of folded and unfolded Nterm-E1p shows exchange cross peaks that "link" the folded and unfolded state of Nterm-E1p. We suggest that the exchange rate between the two species is in the range of 0.5 to 5 seconds^{-1}. Sharp resonances in the NMR-spectra of wild type E1p demonstrate that this 200 kDa enzyme contains highly flexible regions. The observed dynamic character of E1p and of Nterm-E1p is likely required for the binding of the E1p dimer to the two different binding sites on E2p. Moreover, the flexibility might be essential to sustain the allosteric properties of the enzyme bound in the complex.

Introduction

The pyruvate dehydrogenase multienzyme complex (PDHC) from gram-negative bacteria consists of multiple copies of three different enzyme components: pyruvate dehydrogenase (E1p), dihydrolipoyl acyltransferase (E2p) and lipoamide dehydrogenase (E3). The complex catalyses the oxidative decarboxylation of pyruvate and the subsequent acetylation of coenzyme A to acetyl-CoA (for reviews see Perham, 1991; Mattevi et al., 1992a; Berg and De Kok, 1997; De Kok et al., 1998; Perham, 2000). The substrate specific, thiamin diphosphate

dependent E1p catalyses the decarboxylation of pyruvate and subsequently the reductive acetylation of the lipoamide groups attached to E2p. The E2p component then transfers the acyl group to CoA. Finally, the reduced lipoyl group is re-oxidised by the FAD dependent E3 component, which transfers the reduction equivalents to NAD^+.

E2p plays a central role in the complex, both catalytically and structurally. It consists of four to five domains interconnected by flexible alanine/proline rich linkers of 20 to 40 amino acids. This architecture allows a high degree of flexibility of the individual domains, required for catalysis. From N- to C-terminus, E2p consists of two or three lipoyl domains, each carrying a lipoyl-lysine group, an E1/E3 binding domain and a catalytic domain that also forms the structural core of the complex. Three-dimensional structures of the different domains of E2p have been solved either by X-ray crystallography (Mattevi et al., 1992c; Mattevi et al., 1993a; Knapp et al., 1998; Knapp et al., 2000; Izard et al., 1999) or by NMR spectroscopy (Dardel et al., 1993; Green et al. 1995; Ricaud et al., 1996; Berg et al., 1997).

The E1/E3 binding domain of E2p behaves like a Janus-face protein: E3 interacts solely with the N-terminal helix of the E1/E3 binding domain, while E1p interacts with the C-terminal part of the E1/E3 binding domain. The binding mode of *B. stearothermophilus* E3 to the binding domain was solved by X-ray crystallography (Mande et al., 1996). An electrostatic zipper dominates the interactions between the E3 dimer and the well-conserved E1/E3 binding domain, with acidic residues on E3 and basic residues on the binding domain of E2. Since point mutations in the E1/E3 binding domain and in the catalytic domain from *A. vinelandii* E2p both result in highly impaired E1p binding (Schulze et al., 1991a), it can be concluded that the binding site for E1p, unlike the binding site for E3, consists of two regions, one located on the E1/E3 binding domain of E2p and one on the catalytic domain of E2p. This was confirmed by the construction of *A. vinelandii – Escherichia coli* chimeric E2p (Schulze et al., 1992): E1p only interacts strongly with a chimeric E2p when both the E1/E3 binding domain and the catalytic domain are of the same origin as E1p.

The E1p component exists both as a homodimer ($\alpha 2$) or a heterotetramer ($\alpha 2\beta 2$) depending on the source and type of complex. Recently the 3-dimensional structures of the heterotetrameric branched-chain 2-oxoacid dehydrogenase (E1b) from *Pseudomonas putida* and from human were solved (Ævarsson et al., 1999; Ævarsson et al., 2000). No structural information is available though for the homodimeric E1. However, recent studies (Hengeveld et al., 1997) have given insight into the binding mode of homodimeric E1p to E2p. Limited proteolysis experiments of *Azotobacter vinelandii* E1p showed that its N-terminal region (amino acids 1-

50) could easily be cleaved off. The remaining fragment is still active, but unable to bind to E2p. This suggests that the N-terminal region is involved in the binding to E2p, but not in catalysis. The role of the N-terminal sequence of *A. vinelandii* E1p in the binding to E2p was confirmed by the construction of a series of N-terminal deletion mutants (Hengeveld et al., 1999). Construction of a heterodimeric E1p, containing only one N-terminal region, demonstrated that not one but both N-termini of the homodimer are necessary for good binding.

From the previous studies it is not evident whether the N-terminal region of E1p forms an independent folding domain, and if so, how this domain functions. To attempt to answer these questions, we used a synthetic peptide called Nterm-E1p that comprises amino acids 4-45 of the N-terminal region of E1p from *A. vinelandii* in both functional and detailed spectroscopic studies. By answering these questions we expect to gain more insight in the binding process of E1p to its two different binding sites on E2p.

Experimental Procedures

Peptide synthesis

A peptide consisting of 42 amino acid residues (designated Nterm-E1p) (MQDLDPIETQEWLDSLESVLDHEGEERAHYLLTRMGELATRT) that mimics the N-terminal region of E1p was synthesised by Genosys Biotechnologies. The peptide was supplied lyophilised and purified (84%) by reversed phase high performance liquid chromatography (HPLC) using an acetonitrile/water solvent gradient containing 0.1% trifluoroacetic acid (TFA). The peptide was further purified to 98% by HPLC using the same solvents.

Enzyme isolation

The *E. coli* strain TG2, a *recA*- version of TG1 [Δ*(lac⁻proAB), thi, supE, Res⁻ Mod⁻ (k), F' (traD36 proA⁺B⁺, lacI ᵠ lacZ* ΔM15)] (Gibson, 1984) harbouring the recombinant plasmid pAFH001, expressing wild type E1p was grown and purified as described in (Hengeveld et al., 1997). *A. vinelandii* wild type E2p and E3 were expressed and purified from *E. coli* TG2 as described in (Hanemaaijer et al., 1989) and (Westphal and De Kok, 1988). *A. vinelandii* E3-Y16F was expressed and purified as described in (Benen et al., 1992).

Activity assays

The oxidative decarboxylation of pyruvate by E1p was measured at 600 nm using 2,6-dichlorophenol-indophenol (Cl_2Ind) ($\varepsilon = 21.7\times10^3$ $M^{-1}cm^{-1}$) as an artificial electron acceptor (Khailova et al., 1977). The E2p component was assayed as described in (Bresters et al., 1977). The E3 and E3-Y16F component were assayed as described in (Van den Broek, 1971). The overall PDHC-activity was reconstituted by incubating E2p with E1p and E3 or E3-Y16F and was measured spectrophotometrically at 340 nm as described in (Schwartz and Reed, 1970).

Binding studies

The binding of Nterm-E1p to E2p was analysed by analytical size-exclusion chromatography using a Superose 6 column (Pharmacia Biotech) as described in (Hengeveld et al., 1997). The collected fractions were analysed by SDS-PAGE (Schägger and Von Jagow, 1987). The E2p concentration was 8.3 μM; the Nterm-E1p concentration was 40.6 μM.

Protein concentrations were estimated using the microbiuret method (Goa, 1953), Bovine serum albumin was used as a standard.

Limited proteolysis and MALDI-TOF MS

Truncated *A. vinelandii* E1p for NMR spectroscopy was obtained by limited proteolysis of E1p by endoproteinase GluC as described in (Hengeveld et al., 1999). A 1 mg/ml concentration of E1p was used and the reaction was stopped by addition of TLCK (260 μM final concentration). After proteolysis, the small fragments and the protease were removed from the sample by ultrafiltration, using a 100 kDa cut-off filter. This step was repeated three times. The sample was concentrated to the original starting volume of the undigested E1p NMR sample, resulting in a concentration of approximately 0.5 mM of digested E1p.

For MALDI-TOF MS truncated *A. vinelandii* E1p was obtained by limited proteolysis of the enzyme by endoproteinase GluC as described in (Hengeveld et al., 1999). The small fragments were separated from the large (core) fragment on a Superdex peptide column. The collected fractions were freeze-dried and subsequently dissolved in 20-100 μl water. Remaining salt was removed by use of cation exchange chromatography (Biorad AG50W-X8 column). Finally the samples, which were changed by hydrophobic interaction chromatography (Millex-FH, Millipore Corporation) into 50% AcNi + 0.1% TFA / H_2O, were analysed by MALDI-TOF MS on an applied Biosystems voyager DE RP using x-cyano 4 hydroxy cynamic acid as

a matrix solution. TFA was used to decrease the pH of the samples. Microperoxidase (MP-8) (Kraehenbuhl et al., 1974) was used for calibration.

Spectroscopic measurements

Circular dichroism spectroscopy

CD spectra were recorded at 20 °C on a Jasco spectropolarimeter, model J-715. Spectra were recorded in cuvettes with a 0.1 cm path length. The peptide concentration was 10.1 μM (50 μg/ml) in 10 mM potassium phosphate (pH 4.8). The pH was adjusted by addition of aliquots of 250 mM NaOH or of 250 mM HCl to values ranging from pH 3 - 7. The ionic strength was increased by the addition of aliquots of 3M KCl.

Fluorescence spectroscopy

The fluorescence emission spectra were measured on a Fluorolog 3 (SPEX) fluorimeter at several temperatures using excitation wavelengths of 280 or 295 nm, with 1 nm slits. The sample conditions were identical to those described for the CD measurements. Spectra were recorded in 1 ml fluorescence cuvettes with a 1 cm path length.

Time-resolved fluorescence spectroscopy

Time-resolved fluorescence measurements were carried out using mode-locked continuous wave lasers and time-correlated counting as detection technique. A mode-locked CW Yttrium Lithium Fluoride (YFL) laser was used for the synchronously pumping of a cavity-dumped Rhodamine 6G dye laser. The samples were 1 ml with a light path of 1 cm. Fluorescence was detected at an angle of 90° with respect to the excitation light beam. An interference filter (Schott 348.8 nm) was used for detection. Measurements consisted of repeated sequences of measuring during 10 s parallel and 10 s perpendicular polarised emission. Background emission of the buffer solution was measured and used for background subtraction. The fast and single exponential fluorescence decay of para-terphenyl in 50/50% cyclohexane/CCl_4 was measured for obtaining a dynamic instrumental response. Data analysis was performed using a home-built computer program (Novikov et al., 1999). The peptide concentration was 10.1 or 20.2 μM in 10 mM potassium phosphate (pH 4.8). The pH was adjusted by the addition of aliquots of 250 mM NaOH or of 250 mM HCl. The E2p concentration was 3.5 or 7.0 μM.

NMR spectroscopy

Samples of the peptide Nterm-E1p contained 4.05 mM protein in 90% H_2O / 10% ^2H_2O (1D NMR measurements) or 8.1 mM protein in 90% H_2O / 10% ^2H_2O (2D NMR measurements). The E1p samples contained 0.5 mM protein in 90% H_2O / 10% ^2H_2O. The pH was adjusted by the addition of a small amount of a solution of 250 mM NaOH or of 250 mM HCl. ThDP was added to the E1p and to the truncated E1p sample to a concentration where the NMR signals originating from free ThDP were just visible in the recorded NMR spectra. Truncated E1p was obtained by limited proteolysis of E1p by endoproteinase GluC as described above. The E1p NMR-sample was diluted to a 1 mg/ml solution. NOESY (τ_m=50 ms) (Kumar et al., 1980; Macura & Ernst, 1980) and TOCSY (τ_m =50 ms) (Braunschweiler and Ernst, 1983) spectra were acquired on a Bruker AMX500 spectrometer. The used carrier frequency coincided with the water resonance. The sample temperature was maintained at 284 or 298 K as calibrated with tetramethylammonium nitrate (TMA). The NOESY spectra were acquired with pre-saturation during the relaxation delay and mixing time. The TOCSY spectra were acquired using homonuclear Hartman-Hahn transfer using the clean-MLEV17 sequence for mixing, using two power levels for excitation and spinlock (Bax et al., 1988). Time proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983) was used in all experiments. The TOCSY experiments were recorded with 120 t_1 experiments and 2048 data points. The TOCSY experiment of E1p was recorded with 64 t_1 experiments and 2048 data points. The NOESY experiments were recorded with 176 t_1 experiments and 2048 data points. The spectral width was 8406 Hz. The chemical shift is reported relative to an external standard (2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS)). The obtained data were processed using XWINNMR software (Bruker). The data were digitally filtered using a sine-bell shifted by 45° in the t_2 dimension and a squared cosine-bell in the t_1 dimension. After double Fourier transformation and interactive phase correction a baseline correction with a polynomial of maximum order 5 was performed to all spectra. The 2D spectra are presented as contour plots with 14 levels increasing by a factor of 1.3.

A 1D version of the 3-pulse NOESY sequence and phase cycle was used to collect the 1D NMR data. No window functions were applied to the 1D NMR data and no zero filling was applied.

Results and discussion

Functionality of the peptide representing the N-terminal region of E1p

To establish the function of the N-terminal region of E1p from *A. vinelandii*, we used a synthetic peptide representing amino acid residues 4-45 of E1p. Initially we tried to obtain the N-terminal region by limited proteolysis of E1p using endoproteinase GluC. The small fragments obtained after partial proteolyses were analysed by MALDI-TOF MS. A wide range of fragment sizes was found ranging from 392 to 1507 Da. Based on the sequence, fragments ranging from 285-4759 Da (3-40 amino acids) were predicted. The expected molecular mass of the entire N-terminal domain is about 4700 Da, indicating that the N-terminal region itself is sensitive to the protease and that it is digested into small (3-13 amino acids) fragments.

Since we were not able to isolate the N-terminal region itself by the above-described method, we used the synthetic peptide "Nterm-E1p" in our studies. For reasons described in the discussion section we have chosen a region of the N-terminus of E1p that starts at amino acid 4 and that ends at amino acid 45. To determine whether this peptide forms a functional binding domain, we performed several functional studies.

First, complex activity was measured: E2p was incubated with different concentrations of Nterm-E1p, subsequently E1p and E3 were added and finally complex-activity was measured

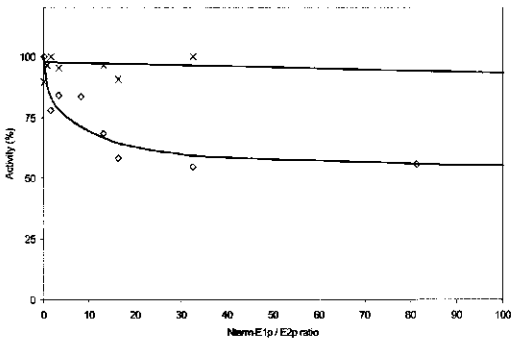


Figure 1 Reconstitution of *A. vinelandii* PDHC in the presence of a varying concentration of Nterm-E1p. E2p (1.6 μM) was incubated with varying concentrations of Nterm-E1p. Subsequently it was recombined with E1p (1.6 μM) and mutant E3 (Y16F) (1.0 μM) (Benen et al., 1992), and finally, after 15 minutes incubation, complex-activity (◊) and E3 activity (×) were measured.

(figure 1). To establish whether Nterm-E1p competes for binding to E2p with E1p, E3 or both, we used the mutant E3-Y16F (Benen et al., 1992) instead of wild type E3. This dimer-interface mutant is impaired in its dimerisation (E3-E3) and virtually inactive. Upon binding to E2p however, it forms a stable and active dimer. This enables us to distinguish between the

bound and unbound form of E3. Figure 1 shows the results of a titration of E2p with Nterm-E1p and the subsequent recombination with E1p and E3-Y16F. The results show that Nterm-E1p inhibits complex activity, whereas it has no influence on the activity of E3-Y16F. Maximum inhibition is about 45% at a 20 : 1 Nterm-E1p to E2p ratio. This partial inhibition can be caused by several factors. First of all, it is possible that the apparent maximum inhibition is not a real plateau but a very shallow increasing slope that will reach 100 % inhibition at infinite peptide concentration. On the other hand, it could also be that the peptide can only interact with one of the two E1p binding sites on E2p and therefore can only partially inhibit the E1p binding to E2p. The activities of the different components of the complex in the presence of a 15-fold excess of Nterm-E1p were measured and no significant effect was found. Consequently, the inhibitory effect on complex activity must be due to competition of Nterm-E1p with either E1p or E3 or both for binding to E2p. As no effect of Nterm-E1p on the Y16F-E3 activity was measured (figure 1) Nterm-E1p must be competing with E1p for specific interactions with one or both of the E1p binding site(s) on E2p. It is very difficult to accurately calculate a binding constant for Nterm-E1p. The measured reaction is very complex: it involves three enzyme components and an inhibitor. Moreover, E1p and E3 influence each other's binding to E2p (Bosma et al., 1984).

Secondly, the binding of Nterm-E1p to E2p was monitored by analytical gelfiltration (data not shown): free E2p mainly elutes in a "24-mer" peak, but upon addition (and binding) of E1p or E3 it dissociates into trimers. An analytical gelfiltration column can separate these fractions very well (Hengeveld et al., 1997; Hengeveld et al., 1999; Mande et al., 1996; Schulze et al., 1991a). When E2p is pre-incubated with an excess amount of Nterm-E1p (1:5 ratio) one extra elution peak appears at an elution volume expected for Nterm-E1p. Whereas the integral of the E2p elution peak increases its elution volume does not change compared to the elution volume of the free E2p sample. This demonstrates that although Nterm-E1p binds to E2p it does not (unlike wild type E1p) influence the trimer-trimer interactions of E2p.

Nterm-E1p has a pH dependent conformation

CD measurements

The conformational behaviour of Nterm-E1p was investigated using CD-, fluorescence- and NMR-spectroscopy. The CD-results are shown in figure 2. At pH 4.8, a spectrum with two minima at 208 nm and at 222 nm and a crossover point at 200 nm is observed, which is

characteristic of a highly α-helical conformation. At pH 6.8, the CD-spectrum exhibits only a minimum at 203 nm and the crossover point is shifted to 193 nm, which is more typical for an unfolded protein. To evaluate the secondary structure of Nterm-E1p we used the program CONTIN (Provencher and Glockner, 1981) (table 1). Indeed, at pH 4.8 the peptide is mostly α-helical (68%), while at pH 6.8 it is mostly unordered (54%). For brevity we will, from here on, refer to the conformation of the peptide at pH 4.8 as "folded" and to the conformation of the peptide at pH 6.8 as "unfolded". We are aware that this does not fully describe the conformation of the peptide under these conditions and have conducted more detailed studies using NMR- and fluorescence-spectroscopy to determine the conformation of Nterm-E1p at various conditions in more detail (see below).

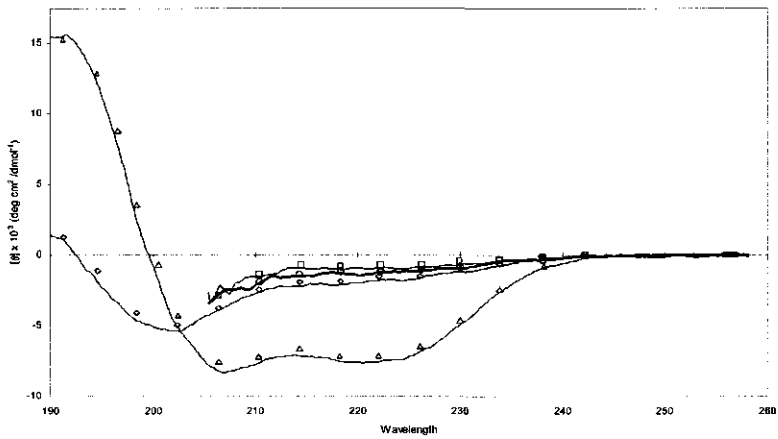


Figure 2 Far-UV CD spectrum of Nterm-E1p at varying pH values and in the presence of different concentration of GuHCl. The samples were prepared as described in the methods section. Nterm-E1p at pH 4.8 (Δ), Nterm-E1p at pH 6.8 (◊), Nterm-E1p at pH 6.8 with 0.6 M GuHCl (O), Nterm-E1p at pH 6.8 with 1.2 M GuHCl (□).

Fluorescence measurements

Nterm-E1p contains a single tryptophan and therefore fluorescence spectroscopy is a very suitable technique for studying the microenvironment of this residue. All experiments were recorded with excitation at 295 nm, where only tryptophan absorbs light. At pH 4.8 the tryptophan has a fluorescence maximum at 335 nm (figure 3A), while between pH 5.5 and pH

6.8 the maximum is at 350 nm. The emission maximum at 350 nm is typical for a tryptophan in a polar environment (with few protein contacts). The blue-shifted maximum at 335 nm at pH 4.8 results from a less polar environment: a tryptophan buried in a protein (i.e. in a folded structure). Thus, this local structure is in agreement with the CD observations that the peptide can exist in two conformations: folded and unfolded. The fluorescence of the peptide at pH 6.8 is quenched by 40 % compared to the fluorescence at pH 4.8. This is caused by the exposure of the tryptophan to the solvent, as a tryptophan's emission pattern is highly sensitive to the polarity of the solution. Since the peptide, besides the tryptophan, also contains one tyrosine, emission spectra with excitation at 280 nm, where both tryptophan and tyrosine absorb light, were recorded as well. Difference spectra were obtained by normalising the spectra and subsequently subtracting the spectrum with excitation at 295 nm from the

Table 1. The secondary structure content of Nterm-E1p at different pH values as calculated from CD measurements by the program CONTIN (Provencher and Glockner, 1981). The predicted secondary structure content of the peptide as determined using the neural network algorithms from the European Molecular Biology Laboratory (EMBL) (Rost and Sander, 1993; Rost and Sander, 1994) is also shown. The samples were prepared as described in the methods section.

	α-helix (%)	β-sheet (%)	β-turn (%)	Unfolded (%)	Unknown (%)
pH 4.8	67	0	14	19	-
pH 6.8	12	24	10	54	-
Prediction	73	0	17	-	10

spectrum with excitation at 280 nm. Both difference spectra (figure 3D) (pH 4.8 and pH 6.8) resemble the spectrum expected for a tyrosine, but the relative contribution of the tyrosine has increased at pH 6.8. This increase in tyrosine fluorescence probably occurs because of a larger distance between the tryptophan and the tyrosine due to unfolding of the peptide, which would lead to a decreased energy transfer between the two aromatic residues.

Besides steady state fluorescence, we also measured time dependent fluorescence decay and fluorescence anisotropy of the peptide at 348.8 nm using time resolved fluorescence measurements (table 2). The fluorescence decay was in both cases (pH 4.8 and pH 7.0) best fitted using a 4 component exponential decay. In the time resolved fluorescence measurements the average fluorescence lifetime of the peptide at pH 7.0 is slightly shorter than at pH 4.8

(table 2), which is in agreement with the decreased fluorescence intensity in the steady state measurements (figure 3A). A fluorescence lifetime of the tryptophan close to 2 ns was found for the peptide at both pH 4.8 and pH 7.0, indicating quenching of the tryptophan fluorescence at both conditions (the average lifetime of a tryptophan is about 3 ns).

Table 2. The fluorescence decay and fluorescence anisotropy parameters of Nterm-E1p at pH 4.8 and pH 7.0 (duplicates). Samples were prepared as described in the methods section.

	Fluorescence decay		Fluorescence anisotropy				
	χ^2 [1]	τ [2]	χ^2	ϕ1 [3] (ns) [CI][4]	β1 [5]	ϕ2 (ns) [CI]	β2
pH 4.8	1.05	2.05	1.05	0.28 [0.14-0.52]	0.11 [0.073-0.14]	2.05 [1.62-2.85]	0.13 [0.094-0.17]
pH 4.8	1.08	2.08	1.05	0.31 [0.17-0.49]	0.12 [0.09-0.15]	2.32 [1.85-3.20]	0.12 [0.088-0.15]
pH 7.0	1.02	1.97	1.07	0.19 [0.032-ND]	0.066 [0.025-ND]	1.32 [1.10-1.92]	0.17 [0.087-0.20]
pH 7.0	0.99	2.00	1.03	0.18 [0.051-ND]	0.071 [0.033-0.12]	1.31 [1.11-1.75]	0.17 [0.11-0.20]

[1] Quality of fit criterion
[2] Average fluorescence lifetime
[3] Rotational correlation times
[4] 67 % confidence interval
[5] Relative contribution of f to the total anisotropy
ND is not defined

This quenching can be caused by a nearby tyrosine, glutamine, glutamic acid, aspartic acid or histidine (Lakowicz et al., 1998).

The anisotropy decays of the peptide at pH 4.8 and at pH 7.0 contain two components (table 2): a short sub-nanosecond component and a component of 1-2 ns. The short component probably originates from internal flexibility of the tryptophan. The smaller contribution of this component to the total fluorescence anisotropy at pH 7.0 compared to the contribution of the sub-nanosecond component at pH 4.8 probably results from a smaller amplitude of the movement of the tryptophan in the pH 7.0 sample.

The second anisotropy component of the peptide at pH 4.8 correlates excellently with the expected rotational correlation time of a globular protein of 4.9 kDa (1.9 ns) (Visser and Lee, 1980). The shorter rotational correlation time (1.3 ns) of the peptide at pH 7.0 supports a changed, less compact, conformation under these conditions, i.e. the peptide being "unfolded".

NMR measurements

1D-^{1}H-NMR spectra of Nterm-E1p were recorded at pH values ranging from pH 5.0 to pH 6.5 (figure 4). Due to overlap of most of the NMR resonances, assignment of the resonances to their corresponding protons is not possible based on these 1D-NMR spectra.

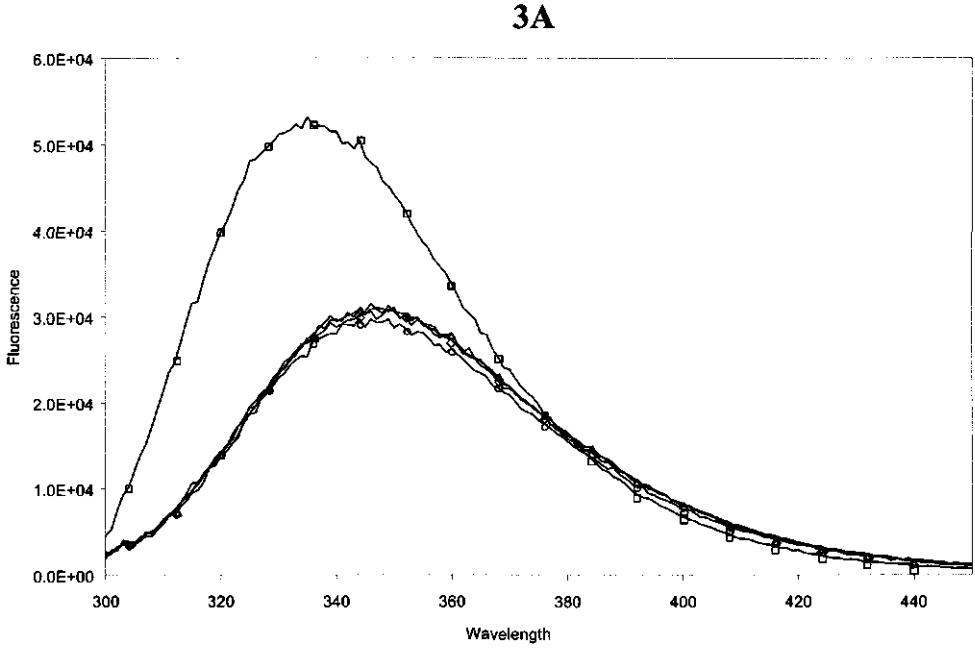
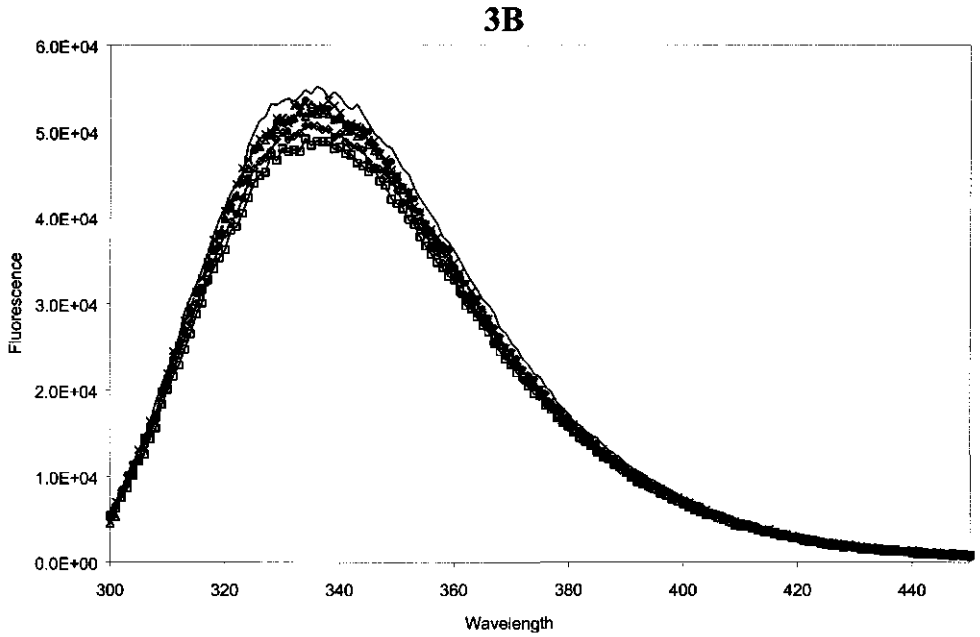


Figure 3 Fluorescence spectra obtained with excitation at 295 nm of Nterm-E1p at varying pH values (A), at varying ionic strengths (B) and in the presence of various concentrations of the denaturant GuHCl (C). Normalised fluorescence spectra obtained with excitation at 280 nm of Nterm-E1p at pH 4.8 and pH 6.8 are shown in (D). Difference spectra were obtained by normalising the fluorescence spectra and subtracting the spectrum obtained with excitation at 295 nm from the spectrum obtained with excitation at 280 nm. (A) Nterm-E1p at pH 4.8 (□), at pH 5.5 (•), at pH 6.2 (Δ) and at pH 6.8 (◊);(B) Nterm-E1p at pH 4.8 (Δ), Nterm-E1p at pH 5 with 50 mM KCl (□), Nterm-E1p at pH 5 with 100 mM KCl (◊), Nterm-E1p at pH 5 with 150 mM KCl (×), Nterm-E1p at pH 5.2 with 250 mM KCl (O) and Nterm-E1p at pH 5.2 with 400 mM KCl (line); (C) Nterm-E1p at pH 4.8 (thin dotted line), Nterm-E1p at pH 6.8 (thick dotted line), Nterm-E1p at pH 6.8 with 0.6 M GuHCl (thin line) and Nterm-E1p at pH 6.8 with 1.2 M GuHCl (thick line); (D) Nterm-E1p at pH 4.8 with excitation at 280 nm (□), Nterm-E1p at pH 4.8 with excitation at 295 nm (----), Nterm-E1p at pH 6.8 with excitation at 280 nm (O), Nterm-E1p at pH 6.8 with excitation at 295 nm (thick dotted line), the difference spectrum of Nterm-E1p at pH 4.8 (◊), the difference spectrum of Nterm-E1p at pH 6.8 (Δ).



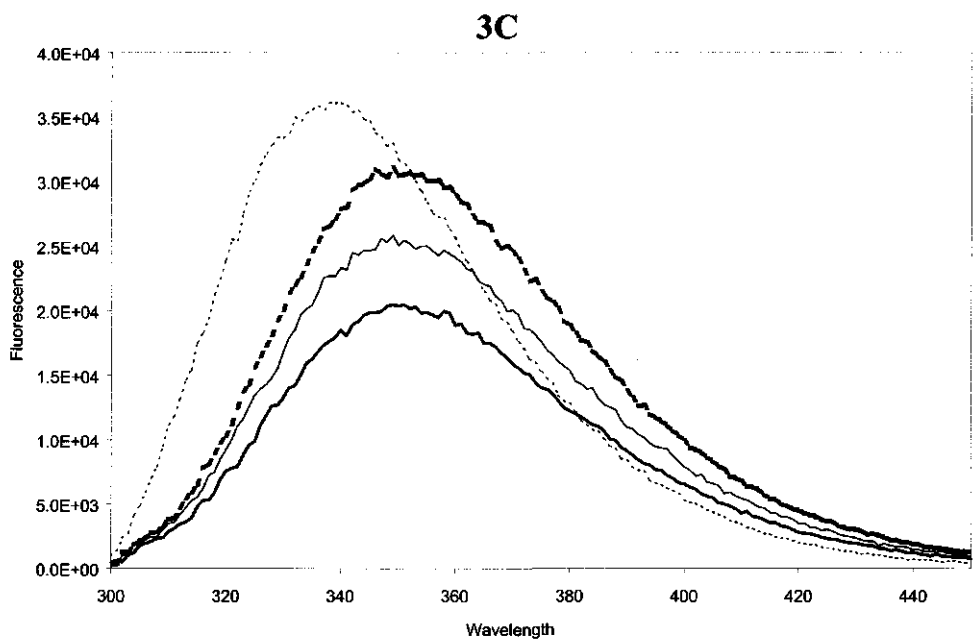


Figure 3, continued

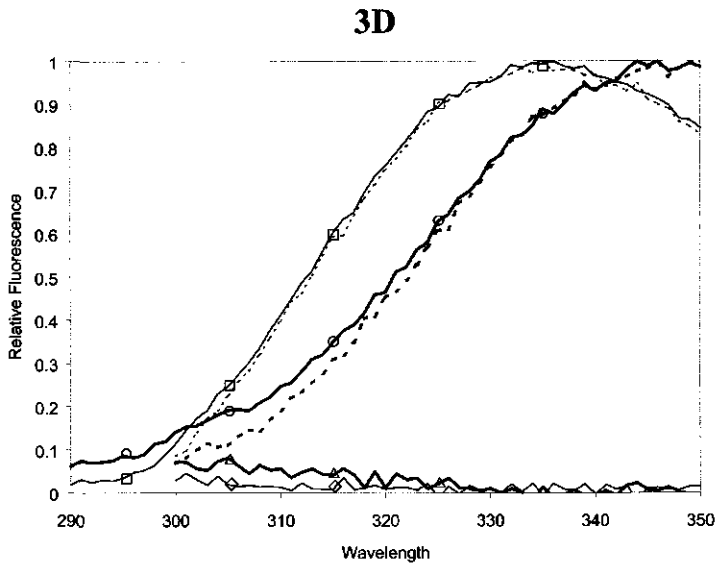


Figure 3, continued

An exception to the latter forms the resonance with a chemical shift of 10.12 ppm (in the spectrum recorded at pH 6.5). This resonance is assigned to the aromatic side chain NH of the single tryptophan in Nterm-E1p (Bundi, and Wüthrich, 1979). At low pH two resonances instead of one are found in the low-field region of the NMR spectrum. The new resonance emerges at a chemical shift of 10.43 ppm, and also originates from the aromatic side chain NH of the tryptophan residue (see following sections). The resonance at 10.43 ppm gives rise to many NOE cross peaks in the 2D NOESY NMR spectrum (figure 5), whereas the resonance at 10.12 ppm does not. Consequently, the resonance at 10.43 ppm corresponds to the aromatic NH of a tryptophan in a folded environment. The tryptophan aromatic NH resonance at 10.12 ppm has virtually no NOE contacts with the remainder of the peptide and its chemical shift is typical for that of an exposed tryptophan, i.e. a tryptophan in an unfolded protein.

The equilibrium between the folded and the unfolded conformation of Nterm-E1p

The experiments described above show that the conformation of Nterm-E1p is pH dependent. At pH 6.5 it seems to be unfolded, while at pH 5.0 a mixture of folded and unfolded peptides exists. To test the influence of a change in sample condition on this equilibrium, 1D-^{1}H-NMR

spectra were recorded at different ionic strengths (figure 4). Additionally, fluorescence spectra were recorded at different ionic strengths (figure 3B), at different temperatures and in the presence of different concentrations of GuHCl (figure 3C). Finally, CD-spectra were recorded in the presence of different concentrations of GuHCl (figure 2).

Table 3 shows the ratios of the integrals of the tryptophan resonances at 10.43 and 10.12 ppm at different pH values and at different ionic strengths. From figure 4 and table 3 it is clear that at pH 6.5 the folding equilibrium is completely shifted towards the unfolded state of the peptide. Under these conditions, only one resonance is observed for a tryptophan aromatic NH proton and the ^{1}H-NMR spectrum has limited chemical shift dispersion, typical for that of a largely unfolded protein.

The other sample conditions tested always resulted in a mixture of folded and unfolded Nterm-E1p. The folding equilibrium is shifted towards the folded state upon lowering the pH, but this effect is limited by the solubility of the peptide: at pH 3.5-4.3 the peptide precipitates due to protonation of the carboxyl groups of the peptide (pKa ~ 4). For this reason the lowest pH value used in our experiments is pH 4.8. Raising the ionic strength of the solution by the addition of KCl results in a further shift in the folding equilibrium towards the folded state: a ratio of folded : unfolded of about 2:1 was the maximum reached under the conditions used by us. However, under the latter circumstances, significant line broadening of the NMR resonances is observed. Consequently, we did not use these sample conditions for the 2D NMR measurements.

Raising the ionic strength of the solution at pH 5 by the addition of KCl had no noteworthy effect on the fluorescence maximum of the tryptophan (figure 3B). Also, addition of a denaturant (GuHCl) to an Nterm-E1p sample at pH 6.8 (unfolded state) had no effect on the tryptophan fluorescence maximum (figure 3C). It did however slightly affect the CD spectrum (figure 2): addition of GuHCl results in a small reduction of the far-UV CD signal suggesting some slight further unfolding of Nterm-E1p. In conclusion, both fluorescence and NMR measurements suggest that the peptide is largely unfolded under these conditions.

2D-NMR: analysis of the secondary structure of Nterm-E1p

To obtain more detailed information about the conformation of Nterm-E1p in the folded (pH 5.0) and unfolded (pH 6.5) state, 2D ^{1}H-TOCSY and 2D ^{1}H-NOESY NMR spectra were recorded (figure 6 and figure 5). We did not attempt to obtain a complete sequential resonance

assignment of the proton resonances of Nterm-E1p at pH 5 as the 2D-NMR spectra are very complex: two species are populated to a significant extent at pH 5, considerable resonance overlap is observed, and several signals are broadened due to exchange phenomena. In addition, as the synthetic peptide is isotropically unenriched, heteronuclear multidimensional NMR spectroscopy could not be employed to our benefit. For these reasons, only the main features of the 2D ^{1}H NMR spectra are discussed below.

Figure 6 shows the fingerprint and the NH cross-peak region of a TOCSY spectrum at pH 5 (figure 6A, 6C) and at pH 6.5 (figure 6B, 6D). The 2D NMR spectrum of Nterm-E1p at pH 5, where a mixture of folded and unfolded peptide exists, shows much more NH resonances and cross peaks than the 2D NMR spectrum of Nterm-E1p at pH 6.5. The extra resonances observed fall outside the chemical shift ranges expected for an unfolded peptide. A about 10-15 d_{NN} NOE's are observed for the folded species. This is consistent with the CD results (table 1) that indicate a significant α-helicity of the folded state. The 2D ^{1}H-TOCSY spectrum recorded at pH 6.5 is typical for that of a random coil protein. Cross peaks corresponding to amide groups of the same amino acid residue types in the unfolded protein have minimal chemical shift differences.

The backbone amide proton resonances resonate between 7.8 ppm and 8.6 ppm. The resonances with chemical shifts between 6.6 ppm and 7.6 ppm can be assigned to the aromatic side chain protons. Consistently, the 2D-^{1}H-NMR NOESY spectrum recorded at pH 6.5 is typical for that of a flexible peptide.

Most of the NH-resonances and cross peaks in the TOCSY spectrum of unfolded Nterm-E1p are also present in the TOCSY spectrum acquired at pH 5. However, some of these resonances could not be observed in the TOCSY spectrum of the mixture of folded and unfolded Nterm-E1p. This can be due to several causes: firstly, the intensities of the cross peaks of the peptide in the unfolded state are reduced in the TOCSY spectrum of the mixture of both folding states as only about half of the molecules are unfolded. Secondly, some signals in the 2D ^{1}H NMR spectra are severely broadened (compare figure 6A and 6B) as a result of exchange at an intermediate exchange rate on the chemical shift timescale, presumably between partially folded and fully unfolded molecules. Thirdly, the pH change can cause the shift of several cross peaks.

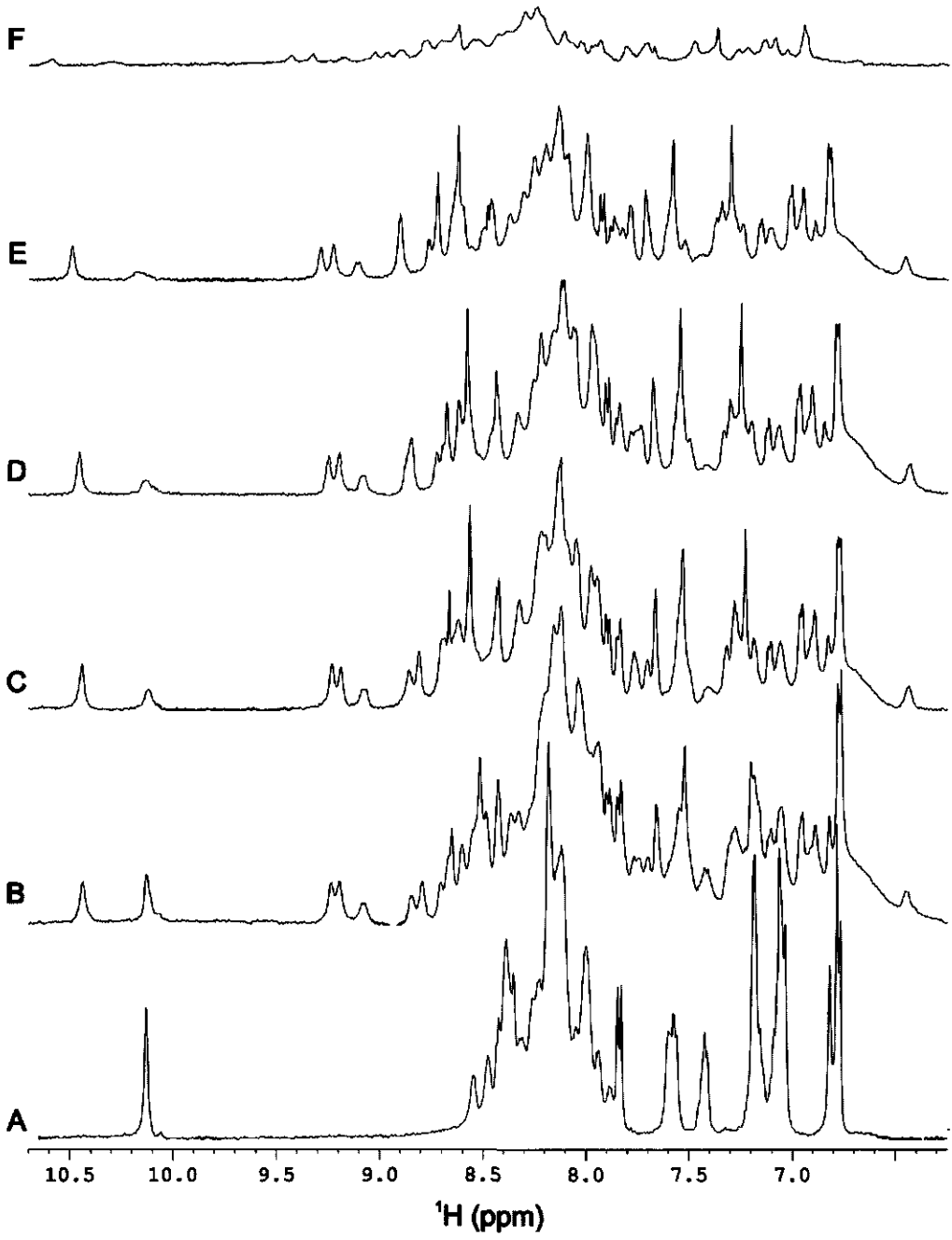


Figure 4 500 MHz 1D-1H NMR spectra of Nterm-E1p in 90% H2O/10% $^{2}H_2O$ recorded at various pH values and various ionic strengths, at a temperature of 298 K. The samples were prepared as described in the methods section. (A) Nterm-E1p at pH 6.5, (B) Nterm-E1p at pH 5.8, (C) Nterm-E1p at pH 5.0, (D) Nterm-E1p at pH 5.0 with 138 mM KCl, (E) Nterm-E1p at pH 5.3 with 415 mM KCl, (F) Nterm-E1p at pH 5.5 with 1 M KCl.

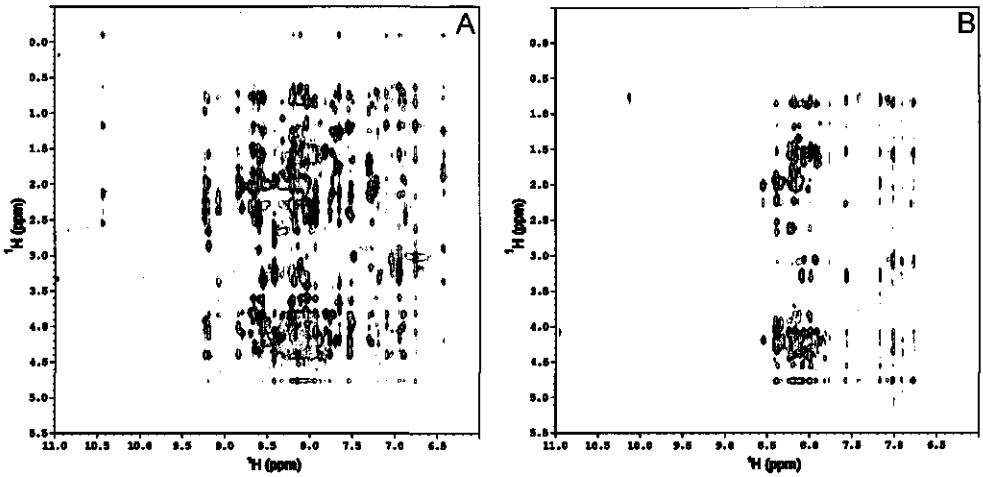


Figure 5:500 MHz 2D-^{1}H NOESY spectra of Nterm-E1p in 90% H_2O/10% ^2H_2O recorded at pH 5.0 and at pH 6.5 at a temperature of 298 K. The mixing time was 50 ms. (A) Fingerprint region of the NOESY spectrum of Nterm-E1p at pH 5.0, (B) Fingerprint region of the NOESY spectrum of Nterm-E1p at pH 6.5.

pH	KCl (mM)	Ratio 10.43 ppm : 10.12 ppm
6.5	0	0 : 1
5.8	0	0.82 : 1
5.0	0	1.49 : 1
5.0	138	1.67 : 1
5.3	415	1.85 : 1
5.5	1000	1.20 : 1

Table 3. The ratios of the integrals of the tryptophan resonances at 10.43 and 10.12 ppm as a function of pH and ionic strengths. The tryptophan resonance at 10.43 ppm represents the "folded" state of Nterm-E1p, whereas the tryptophan resonance at 10.12 ppm represents the "unfolded" state of the peptide. The samples were prepared as described in the methods section.

The rate of exchange between folded and unfolded Nterm-E1p

Thus far we described the peptide's function and conformation and studied the conditions that influence the equilibrium between the two conformations of the peptide. Here we focus on an interesting phenomenon displayed by the 2D ^{1}H-TOCSY spectrum recorded at pH 5.0 and

298 K. This spectrum not only shows the two aromatic side chain NH resonances of the tryptophan in its folded and unfolded state, but it also shows a cross peak that "links" these two resonances. Several similar cross peaks are observed in the low-field region of the TOCSY spectrum (figure 6C). These cross peaks must be due to chemical exchange that is fast on the NMR T_1 timescale, but not on the chemical shift timescale ($\Delta\delta > k_{exch} > T_1^{-1}$).

For the majority of these exchange cross peaks, one of the frequencies could be assigned to the folded state of Nterm-E1p and the other frequency component resulted from the corresponding amide of unfolded Nterm-E1p. The exchange cross peaks disappear once the folded state of Nterm-E1p is fully populated. Similar exchange cross peaks "linking" folded and unfolded species that are in equilibrium have been observed for folding intermediates of bovine pancreatic trypsin inhibitor (BPTI) (van Mierlo et al., 1991; van Mierlo, et al., 1993). Based on estimates of the likely T_1 values for Nterm-E1p, we suggest that the exchange rate between the folded and the unfolded Nterm-E1p molecules at 298 K is in the range 0.5 to 5 $seconds^{-1}$.

The TOCSY experiment was repeated at 284 K, which yielded the same folded : unfolded ratio for the Nterm-E1p species as was found at 298 K. However, now no exchange cross peaks between the folded and unfolded tryptophan NH resonance are observed. Additionally, no exchange cross peaks in the low-field region of the TOCSY spectrum are observed under these conditions. Thus, temperature does seriously affect the exchange rate, but it does not affect the folding equilibrium of the Nterm-E1p species.

NMR characterisation of the 200 kDa E1p homodimer

The observed dynamic exchange between the folded and unfolded state of Nterm-E1p suggests that the N-terminal domain of E1p is highly mobile. In the past, flexibility of regions of E1p was suggested on basis of the sensitivity of E1p to proteases (Hengeveld et al., 1997) and on its resistance to crystallisation (personal communication, W. Hol). Highly mobile parts of a protein should be visible in an NMR experiment, even if the entire protein is very large (the E1p dimer has a molecular weight of 2 × 100 kDa). For this reason we recorded 1D-^{1}H (figure 7A) and 2D-^{1}H-TOCSY (figure 8A) NMR spectra of the intact E1p at pH 5.5. The 1D and 2D NMR spectra show a number of sharp resonances normally not expected for a structured enzyme with the size of E1p. ThDP (E1p's cofactor) was present in the sample in

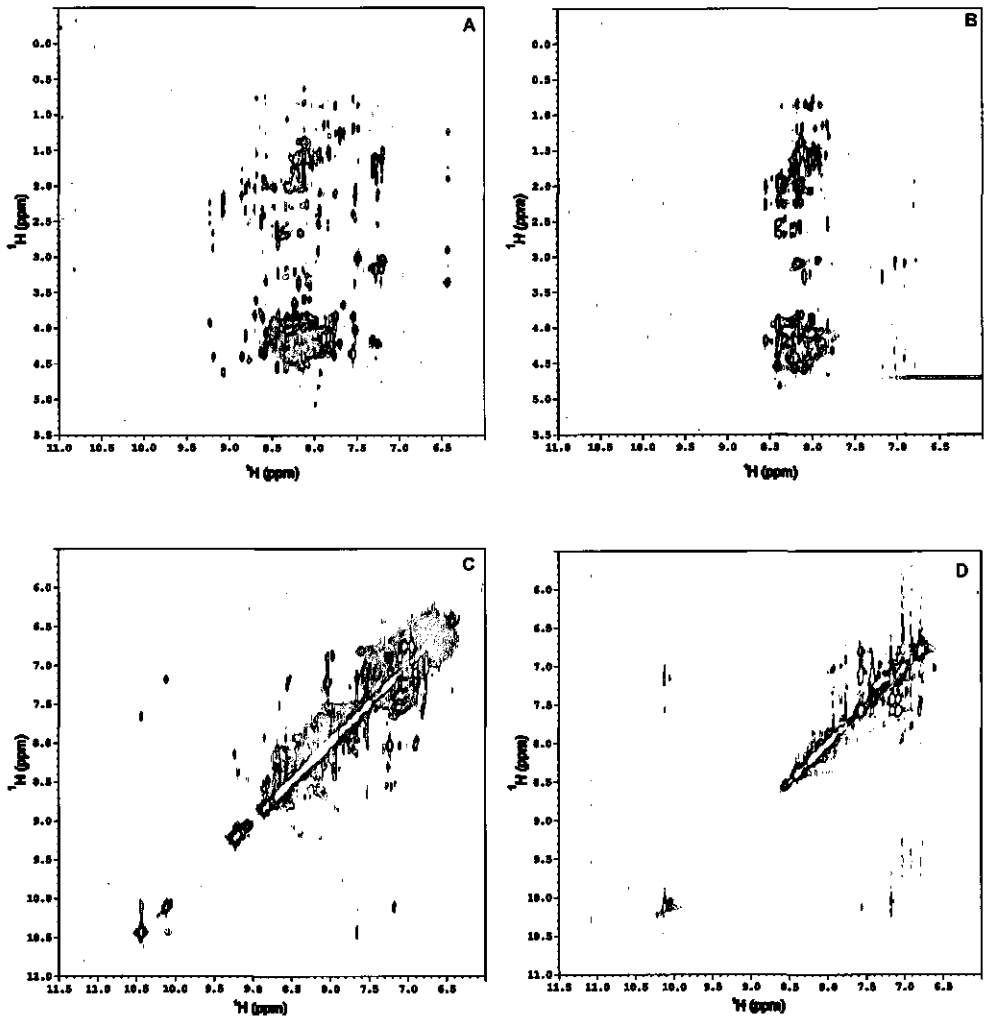


Figure 6:500 MHz 2D-^{1}H TOCSY spectra of Nterm-E1p in 90% H_2O/10% ^2H_2O recorded at pH 5.0 and at pH 6.5 at a temperature of 298 K. The mixing time was 50 ms. (A) Fingerprint region of the TOCSY spectrum of Nterm-E1p at pH 5.0, (B) Fingerprint region of the TOCSY spectrum of Nterm-E1p at pH 6.5, (C) NH cross-peak region of the TOCSY spectrum of Nterm-E1p at pH 5.0, (D) NH cross-peak region of the TOCSY spectrum of Nterm-E1p at pH 6.5.

very small excess. To identify the resonances originating from free ThDP we recorded the spectrum of ThDP (figure 7C).

Clearly, besides these resonances, the spectra contain additional sharp resonances (figure 7A and 8A), indeed confirming the presence of highly mobile regions in E1p. Most of these resonances are found in the 7.8-8.6 ppm region, suggesting unstructured parts of the protein. Upon comparison of the E1p NMR spectra with the NMR spectra recorded of Nterm-E1p, we were unable to identify distinct cross peak patterns at comparable chemical shifts in both

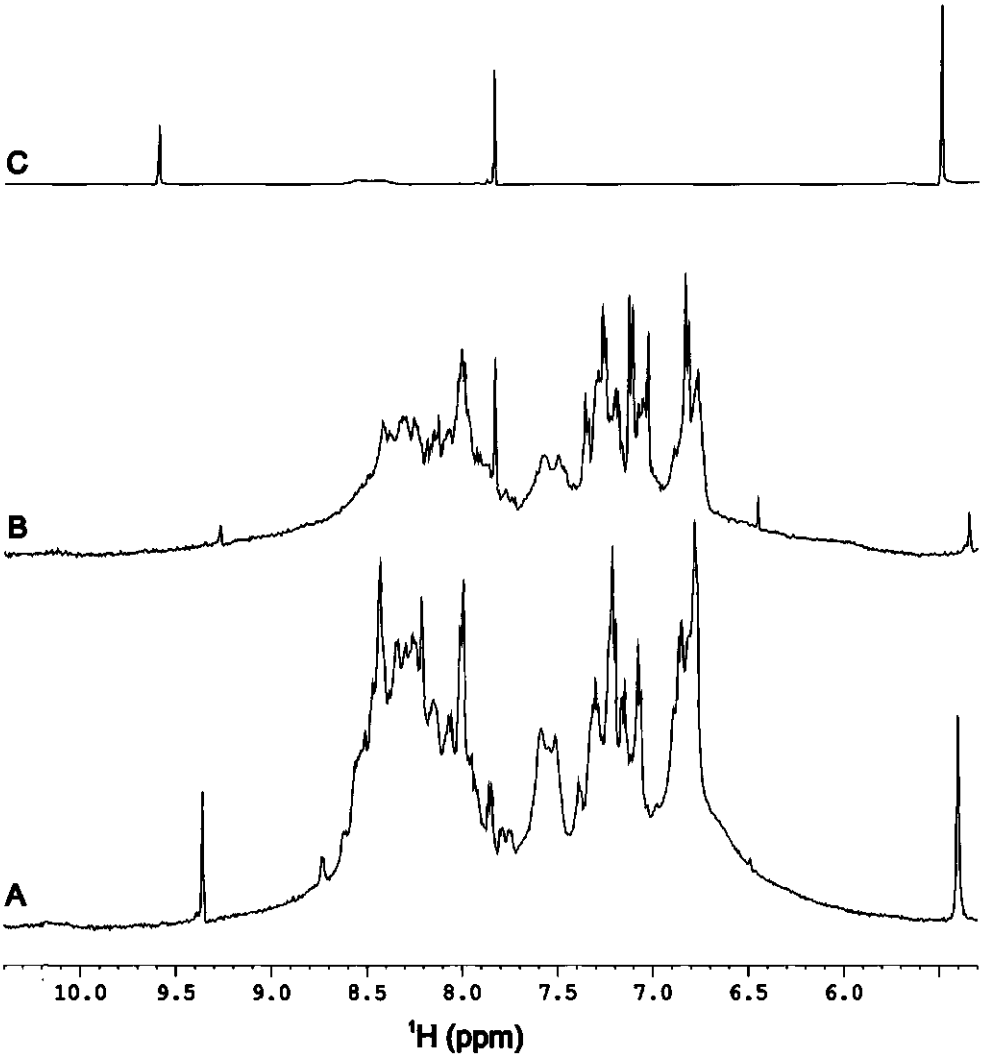


Figure 7: 500 MHz 1D-^{1}H NMR spectra of E1p and truncated E1p in 90% H_2O/10% $^{2}H_2O$ recorded at pH 5.5 at a temperature of 298 K. (A) E1p, (B) E1p truncated by endoproteinase GluC, (C) ThDP. Samples were prepared as described in the methods section.

spectra. Nevertheless, the TOCSY spectrum of E1p shows some similarity to the spectrum of the unfolded peptide. Thus, although we demonstrate that the entire E1p contains highly mobile/flexible regions, we are, based on these experiments, unable to unambiguously proof that these peaks originate from residues in its N-terminal domain.

To determine whether the observed sharp resonances originate from the N-terminal region of E1p, or from other flexible regions in the enzyme, we partially digested the sample with endoproteinase GluC. This protease cuts E1p at amino acid residue E40 and thus digestion results in removal of amino acid residues 1-40. The subsequently recorded 1D-^{1}H and 2D-^{1}H-TOCSY NMR spectra are shown in figure 7B and 8B. Comparison of the spectra of digested and undigested E1p shows that quite a number of sharp resonances in the spectrum of the complete E1p have disappeared in the spectrum of the truncated E1p. Thus, it seems that at

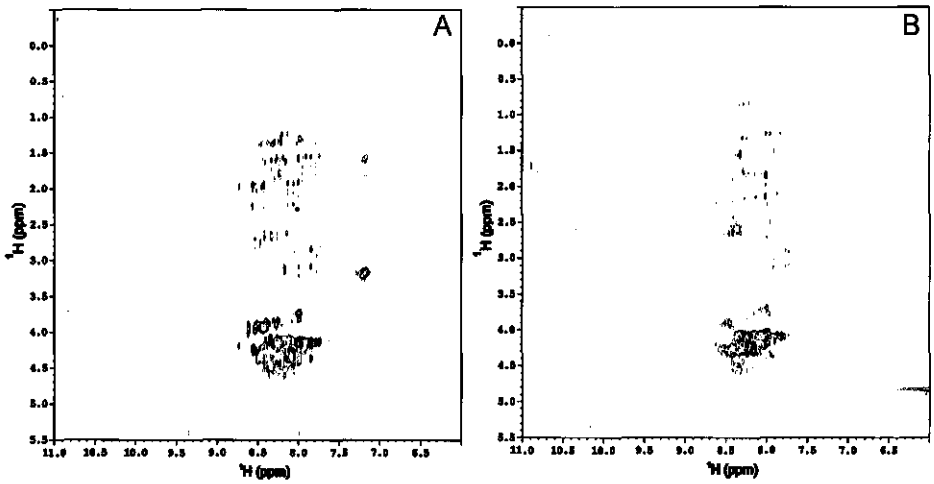


Figure 8: 500 MHz 2D-^{1}H TOCSY spectra of E1p and truncated E1p in 90% H_2O/10% ^2H_2O recorded at pH 6.5 at a temperature of 298 K recorded at 500 MHz. The mixing time was 50 ms. (A) Fingerprint region of the TOCSY spectrum of E1p, (B) Fingerprint region of the TOCSY spectrum of E1p truncated using endoproteinase GluC.

least part of the sharp resonances originate from the N-terminal residues 1-40 of E1p. Some sharp resonances however are not affected by truncating E1p, suggesting that these resonances either originate from another mobile region of E1p, or from impurities in the sample that are not sensitive to the protease. Previously we demonstrated that the core of E1p starts at amino acid 49, therefore truncation at amino acid 40 likely results in a small N-terminal "tail". This

could explain that the 2D-^{1}H TOCSY spectrum of the truncated E1p still contains sharp resonances. Moreover, E1p is cleaved by chymotrypsin and trypsin at respectively residue R418 and residue R554 (Hengeveld et al., 1997).

It seems likely that exposed loops or turns are present at these sites. Possibly these loops are flexible causing sharp NMR resonances. Nevertheless, these results strongly indicate that the N-terminal region of E1p is flexible, analogous to Nterm-E1p.

Discussion

In the present study we describe the structural and functional characterisation of a synthetic peptide (Nterm-E1p) that represents the N-terminal region of E1p.

Previous experiments showed that the N-terminal region of E1p is necessary for the binding of E1p to E2p (Hengeveld et al., 1999). The design of the peptide representing the N-terminal region is based on the following considerations. 1st, deletion experiments show that the structural core of E1p starts from amino acid 49 (Hengeveld et al., 1999). 2nd, amino acids R37, E40 and R44 are all highly sensitive to proteases indicating a more flexible/extended region (Hengeveld et al., 1997). 3rd, deletion experiments show that amino acid residues 1-8 are not necessary for the binding of E1p to E2p (Hengeveld et al., 1999). 4th, due to the decrease of the accuracy of the peptide synthesis upon the increase of the peptide size, the size of the peptide had to be minimised. Consequently, we decided to construct a peptide starting at Met4 and ending at Thr45.

Using activity and binding essays we demonstrate that the peptide Nterm-E1p binds to E2p and that in doing so it specifically competes with E1p. Furthermore, we show that the peptide has no influence on the binding of E3 to E2p. Thus, here we unambiguously proof that the N-terminal region of E1p forms a separate domain involved in the binding to E2p. Moreover, these experiments justify the use of the synthetic peptide for the characterisation of the structure and function of the N-terminal domain of E1p.

Both the CD experiment (table 1) and the 2D NMR analysis of Nterm-E1p at pH 4.8 indicate that the peptide has a high α-helical content. The calculated secondary structure content of Nterm-E1p at pH 4.8 based on the CD data is remarkably similar to the predicted secondary structure content of the peptide based on an analysis by the neural network algorithms from the European Molecular Biology Laboratory (EMBL) (Rost and Sander, 1993; Rost and

Sander 1994). From the secondary structure prediction and our own data we infer the following model of the N-terminal domain of E1p. The domain likely consists of two α-helices connected by a loop or turn. Because residues 1-8 could be removed without affecting E1p, these residues probably do not belong to the first helix. Since the E1p "core" starts at amino acid 49 and residues R37, E40 and R44 are highly sensitive to proteases, it seems very likely that this region forms an extended "linker", connecting the N-terminal domain to the E1p core. Analysis of the sequences of the putative α-helices in the N-terminal domain shows that mainly the first helix (approximately amino acids 13 – 25) has a high content of acidic residues. A helical wheel of this sequence shows that D17, E20 and D24 are all on the same side of the amphipatic helix, separated by a single turn. The opposite side of the helix has the same pattern, but now with hydrophobic residues. The second helix (for which it is less eminent where is starts and ends) also has one side that has a high content of hydrophobic residues. We suggest therefore that the N-terminal domain is formed by these two helices, the hydrophobic patches of which interact, forming a hydrophobic "core". The acidic residues of helix 1 form the binding region to E2p. Mutagenesis experiments are in progress to prove this hypothesis.

Using CD, fluorescence and NMR spectroscopy, we demonstrate that the peptide is present in two conformations, the population of which depends on the sample conditions. The conformations are designated "unfolded" at high pH (above pH 6) and "folded" at low pH (below pH 5). The secondary structure estimation based on the CD data suggests that the unfolded state is not entirely unfolded. Addition of GuHCl however, results in only a slight further unfolding of the peptide (figure 2). Addition of GuHCl has no influence on the fluorescence maximum of the tryptophan; the immediate environment of this residue is thus largely unfolded at pH 6.8. Additionally, the TOCSY and NOESY NMR spectra of Nterm-E1p at pH 6.5 suggest a largely unfolded protein. We therefore believe that the peptide is largely unfolded at high pH.

The equilibrium between the folded and unfolded state of the peptide is not affected only by the pH but also by the ionic strength as NMR spectra show. The 2D ^{1}H TOCSY spectrum of the mixture of folded and unfolded Nterm-E1p at pH 5 and 298 K shows a cross peak linking the two aromatic side chain NH resonances of the single tryptophan, corresponding to the folded and unfolded state of Nterm-E1p. The absence of the exchange peaks at 284 K compared to 298 K in the 2D TOCSY NMR spectrum demonstrates that the exchange rate is dependent on the temperature. The exchange rate between the folded and unfolded Nterm-E1p

molecules is in the range 0.5 to 5 seconds^{-1}. Similar exchange rates were found for the exchange between the folded and unfolded state of BPTI folding intermediate (Van Mierlo et al., 1991; Van Mierlo, et al., 1993). Just as in case of the BPTI folding intermediates, the folded conformation of Nterm-E1p is not very stable.

As mutagenesis experiments identified basic residues in the E1/E3 binding domain and catalytic domain as the binding residues for E1p (Schulze et al., 1991a) it is likely that the acidic residues in the N-terminal domain are involved in ionic interactions between E1p and E2p. Shielding of these charges might mimic the bound state of the N-terminal domain, resulting in a shift in the equilibrium towards the folded state.

It is surprising that at high pH the peptide is unfolded as at this pH E1p binds very well to E2p. However, it is possible that interaction with E2p induces a folded conformation (analogous to the charge shielding effect). Alternatively, removal of folded N-terminal domain by binding to E2p will force the equilibrium in the folded direction, since the folded and unfolded conformations are in slow exchange. Thirdly, we cannot exclude that interactions of the N-terminal domain of E1p with the E1p core stabilise the folded conformation of the N-terminal domain.

Sharp resonances in the NMR-spectra of the complete E1p (figures 7 and 8) unambiguously demonstrate that the 200 kDa enzyme contains highly flexible regions. Because several of these resonances are absent in the spectrum of the truncated E1p it is very likely that the N-terminal region of E1p (amino acid 1-48) is flexible in the native enzyme. Efficient binding of E1p to E2p requires residues on both the binding and the catalytic domain of E2p (Schulze et al., 1991a). Only an E1p dimer (containing two N-terminal domains), and not a monomer or heterodimer (containing one N-terminal domain) binds strongly to E2p (Hengeveld et al., 1999). The requirement for two N-terminal domains is most likely due to the distance between the binding residue R416 on the catalytic domain of E2p and the C-terminal helix of the binding domain of E2p, which is too large to be spanned by one N-terminal domain of E1p. Since E1p interacts with two different binding sites on E2p, it is likely that the flexible regions are necessary for the enzyme to be able to adapt to the two required "binding conformations". The observed dynamic character of Nterm-E1p is possibly required for it's binding to E2p. After binding of the first N-terminal domain of the E1p dimer to E2p a high degree of flexibility of E1p will be required to interact with the second binding site on E2p. Recent 3D electron microscopy of *Saccharomyces cerevisiae* E2p (Zhou et al., 2001) revealed striking dynamics of the E2p core; a 40 Å difference in diameter was observed. Moreover, the intact

complex exhibits similar flexibility to that of the E2p core and the E1p on the outside of the core mimics the expansion/contraction of the underlying E2p core. It is therefore quite likely that the flexible regions that we observed in E1p and Nterm-E1p are involved in this dynamic process, assuming that this core flexibility is also present in the cubic E2p cores.

Finally, the flexibility of E1p, and possibly the flexibility of the N-terminal domain of E1p can also be of importance for the catalytic properties of the enzyme. E1p is an allosteric enzyme, therefore requiring conformational freedom of the active site of the enzyme in the multienzyme complex.

In summary, we demonstrated that E1p has an independent N-terminal folding domain that is necessary for binding of E1p to E2p. The domain (in the isolated peptide) samples both a folded and an unfolded conformation, the population of which is determined by pH and ionic strength. The observed exchange between the two conformations of the N-terminal domain and the flexible regions that we identified in the entire E1p are very likely necessary for the binding of E1p to E2p.

Acknowledgements: We thank Arie van Hoek for his contribution to the time resolved fluorescence measurements and Sjef Boeren his contribution to the MALDI measurement.

5 Identification of the E2-binding residues in the N-terminal domain of prokaryotic pyruvate dehydrogenase.

A modified version of this chapter has been submitted to FEBS Letter: Annechien F. Hengeveld and Aart de Kok. Identification of the binding residues in the N-terminal domain of pyruvate dehydrogenase.

Abstract

Pyruvate dehydrogenase (E1p) is one of the components of the pyruvate dehydrogenase multienzyme complex (PDHC). Previously, it was shown that the N-terminal domain of E1p is involved in its binding to the core component (E2p) of PDHC.

We constructed point mutations in this domain (D17Q, D17R, E20Q, E20R, D24Q and D24R) to identify the specific residues involved in these interactions. Kinetic and binding studies show that D17 is essential for the binding of E1p to E2p. D24 is involved in the binding, but not essential, whereas E20 is not involved. None of the mutations affects the folding or dimerisation of E1p

Based on these observations we have refined the model of the binding mode of E1p and E2p In the proposed "helix-turn-helix" motif residues D17, E20 and D24 are on one side of the helix, whereas on the opposite side W15, L19 and V22 form a hydrophobic "patch". Helix II also has a hydrophobic side, which can form a hydrophobic "pocket" with the residues of helix I. None of the mutations result in complete abolishment of the binding capacity of E1p to E2p; neither do point mutations in E2p. Apparently the binding interface is composed of more than one region both on E2p (binding domain and catalytic domain) and on E1p.

The binding of E1p to E2p is highly species specific: *E. coli* E1p binds very poorly to *A. vinelandii* E2p. The fact that the mutations D17Q and D24R both have such a strong effect on the binding abilities of E1p correlates well with this observation.

Introduction

The pyruvate dehydrogenase multienzyme complex (PDHC) from gram-negative bacteria consists of multiple copies of three different enzyme components: pyruvate dehydrogenase (E1p), dihydrolipoyl acyltransferase (E2p) and lipoamide dehydrogenase (E3). It catalyses the oxidative decarboxylation of pyruvate, resulting in the acetylation of coenzyme A and the transfer of reducing equivalents to NAD^+ (for reviews see Perham, 1991: Mattevi et al., 1992a; Berg and De Kok 1997; De Kok et al., 1998; Perham, 2000).

From N- to C-terminus, E2p consists of two or three lipoyl domains, each carrying a lipoyl-lysine group, an E1/E3 binding domain and a catalytic domain, which also forms the structural

core of the complex. Flexible alanine/proline rich linkers of 20 to 40 amino acids interconnect these domains.

The E1 component exists both as a homodimer (α_2) or a heterotetramer ($\alpha_2\beta_2$) depending on the source and type of complex. Recently the 3-dimensional structure of the heterotetrameric branched-chain oxoacid dehydrogenase (E1b) from *Pseudomonas putida* and human were solved (Ævarsson et al., 1999; Ævarsson et al., 2000). For the homodimeric E1 however, no detailed structural information is available. Recent studies have given insight into the binding mode of homodimeric E1p from *Azotobacter vinelandii*; limited proteolysis experiments (Hengeveld et al., 1997) and mutagenesis studies (Hengeveld et al., 1999) showed that the N-terminal region of E1p (amino acid 1-48) is essential for its binding to E2p. Additionally, mutagenesis experiments of E2p identified binding sites for E1p on both the E1/E3 binding domain of E2p and on the catalytic domain of E2p (Schulze et al., 1991a). The construction of chimeric E2ps confirmed the importance of both these domains in the interactions between E1p and E2p (Schulze et al., 1992). For tight binding the homodimeric E1p requires two N-terminal binding regions (Hengeveld et al., 1999), therefore it seems likely that one N-terminus of E1p interacts with the E1/E3 binding domain of E2p, whereas the other N-terminal region of E1p binds to the catalytic domain of E2p. Structure and function of the N-terminal region of E1p were studied in more detail using a synthetic peptide representing this region and it was shown that this region of E1p forms an independent folding domain that specifically competes with E1p for the binding to E2p. CD and NMR measurements indicated that the N-terminal domain is mostly α-helical. (unpublished observations) These data combined with a predicted secondary structure suggest that the domain consists of a "helix-turn-helix" motif, attached to the core of E1 by an extended linker or loop. Using an amino acid sequence alignment of the N-terminal region of E1p from several gram-negative bacteria, combined with the above, we have identified the residues D17, E20 and D24 in the N-terminal domain to be likely candidates for the binding of E1p to E2p. Additionally, these residues might be involved in the species-specific binding of E1p to E2p. Here we describe a mutational analysis of these putative binding residues of E1p.

Experimental procedures

Construction of E1p mutants

The E1p mutants D17Q, D17R, E20Q, E20R, D24Q and D24R were constructed using standard PCR mutagenesis techniques. pAFH001 (Hengeveld et al., 1997) was used as the wild type template plasmid DNA.

Enzyme expression and purification

The *E. coli* strain TG2, a *recA*- version of TG1 [*Δ(lac proAB), thi, supE, Res⁻ Mod⁻ (k), F'* (*traD36 proA⁺B⁺, lacI^q lacZ* ΔM15)] (Gibson, 1984) harbouring the recombinant plasmid pAFH001, expressing wild type E1p was grown and purified as described in (Hengeveld et al., 1997). E1p mutants were expressed and purified from *E. coli* as wild type E1p to approximately 80% purity, omitting the gelfiltration and the second ion-exchange chromatography step. The purity of the mutant E1ps was estimated from their relative intensity on SDS-PAGE (Schägger and Von Jagow 1987) coloured with Coomassie stain. *A. vinelandii* wild type E2p and E3 were expressed and purified from *E. coli* TG2 as described in (Hanemaaijer et al., 1989; Westphal & De Kok 1988).

Enzyme activity assays

The oxidative decarboxylation of pyruvate by E1p was measured at 600 nm using 2,6-dichlorophenol-indophenol (Cl_2Ind) ($\varepsilon = 21.7\times10^3\ M^{-1}cm^{-1}$) as an artificial electron acceptor (Khailova et al., 1977). The overall PDHC-activity was reconstituted by incubating E2p with E1p or mutant E1p and E3 and was measured spectrophotometrically at 340 nm as described in (Schwartz & Reed 1970). Protein concentrations were estimated using the microbiuret method (Goa 1953), Bovine serum albumin was used as a standard.

Interaction of E1p and E1p mutants with E2p

The binding of E1p or mutant E1p to E2p was analysed by analytical size-exclusion chromatography using a Superose 6 column (Pharmacia Biotech) as described in (Hengeveld et al., 1997).

Results and Discussion

Based on a secondary structure prediction (figure 1) and on experimental data showing that the N-terminal domain has a high α-helical content, we previously proposed a "helix-turn-helix" motif for this domain (Hengeveld et al., 2001). Mainly the N-terminal α-helix has a very high content of acidic amino acid residues. The interactions between E1p and E2p are highly species specific, and it is therefore not likely that conserved residues are solely involved in these interactions. Residues D17, E20 and D24 of *A. vinelandii* E1p are not conserved and moreover, the construction of a helical wheel of amino acid residues 9-25 shows that they are all on the same side of the helix. It therefore seems very likely that (some of) these residues are involved in the interactions between E1p and E2p. To test this hypothesis, we constructed the following mutants of E1p: E1p-D17Q, E1p-D17R, E1p-E20Q, E1p-E20R, E1p-D24Q and E1p-D24R.

```
A. vin              MQDMQDLDPI ETQEWLDSLE SVLDHEGEER AHYLLTRMGE LATRTGTQLP
P. aer                MQDLDPV ETQEWLDALE SVLDREGEDR AHYLMTRMGE LASRSGTQLP
A. eut MSAVPEQILGASSANDADPQ ETHEWLDALQ GVLAAEGPAR AAFLIDKQIE YARNNGVTQP
N. men            MSTQLHDVDPI ETQEWLDALS SVLEYEGGER AQYLLENLVK YCRDKGVRMP
E. col          MSERFPNDVDPI ETRDWLQAIE SVIREEGVER AQYLIDQLLA EARKGGVNVA
H. inf          MSEILKNDVDPI ETQDWLQSLD SLIREEGVER AQYIVEQVIG QARTSGVSLP
```

Figure 1 Sequence alignment of the N-terminal sequence of E1p from several Gram negative bacteria. A. vin, *A. vinelandii* E1p; P. aer, *Pseudomonas aeruginosa* E1p; A. eut, *Alcaligenes eutrophus* E1p; N. men, *Neisseria meningitidis* E1p; E. coli, *E. coli* E1p; H. infl, *Haemophilus influenzae* E1p.

Table 1 shows the kinetic parameters of these mutants and of wild-type E1p. Clearly, the mutations D17R, E20Q, E20R and D24Q have no significant effects on the kinetic properties of the enzyme. The mutations D17Q and D24R on the other hand, do change the kinetics of the enzyme. Especially E1p-D24R is strongly impaired in its E1-activity, whereas the effect on E1-D17Q is intermediate. Removal of amino acids 1-48 affects the affinity of E1p for pyruvate (Hengeveld et al., 1999) and it is therefore not entirely surprising that mutations in this region affect the kinetic parameters of E1p. From the structural and functional information available at this time it is not possible to give an explanation of this effect. Results shown below prove however that global misfolding of these mutants does not cause the effect on the activity of E1p-D24R and E1p-D17Q and therefore it is likely that the effects

on the catalytic activity of these mutants is not directly correlated with their ability to interact with E2p.

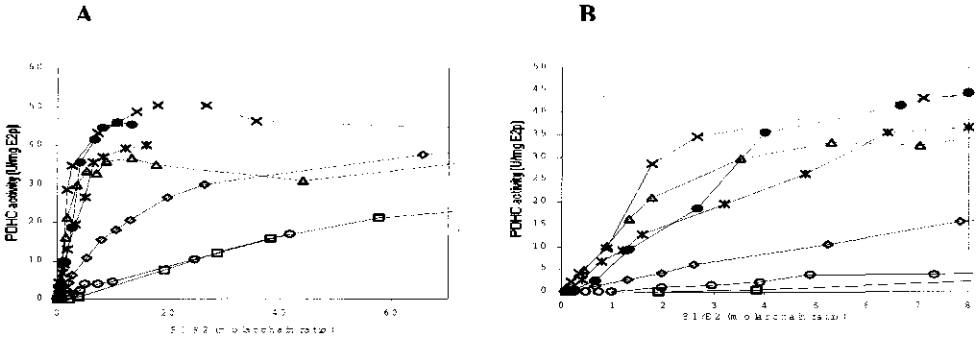


Figure 2 A Reconstitution of *A. vinelandii* PDHC using wild type E2p (3.9 μM) and E3 (2.8 μM) and either wild type or mutant E1p. E2p is preincubated with E3 at a molar ratio of 1:0.7 (E2p:E3). Subsequently varying amounts of E1p are added to restore complex activity. E1p-D17Q (◊), E1p-D17R (□), E1p-E20Q (Δ), E1p-E20R (×), E1p-D24Q (*), E1p-D24R (O) and wild type E1p (•). **B,** Idem at low E1/E2 chain ratio's.

Binding properties of wild-type and mutant E1p

Two different methods were used to study the binding of wild type and mutant E1p to E2p. Firstly, PDHC was reconstituted by the addition of wild type or mutant E1p to the E2p/E3 sub-complex and PDHC-activity was measured (figure 2). Secondly, the binding properties of wild type and mutant E1p were studied by analytical gelfiltration.

Binding studies by PDHC reconstitution.

The first method shows that at an E1p:E2p ratio of 5:1 wild type E1p virtually saturates E2p (based on the PDHC-activity). E1p-E20Q and E1p-E20R also saturate E2p at this ratio and consequently the residue E20 does not seem to be involved in the interactions between E1p and E2p. Addition of E1p-D17R, E1p-D17Q or E1p-D24R on the other hand, results in only 1%, 28% or 10% of maximum PDHC-activity at a 5:1 E1p:E2p ratio. The mutation D24Q has an intermediate effect (65%) on the PDHC-activity at this ratio. At a 60-fold excess of E1p the PDHC-activity of the severely affected E1p mutants (E1p-D17R and E1p-D24R) reaches about 50% of maximum PDHC-activity. E1p-D17Q almost saturates E2p at this ratio, whereas the mutant E1p-D24Q attains saturation of E2p at 16-fold excess. Since at a 60-fold

excess of E1p the PDHC-activity curves of the severely affected E1p mutants do not reach saturation, but are still increasing, it seems likely that at infinite mutant E1p concentrations saturation may occur.

Comparing the effects of the mutations D17Q and D17R, one sees that even though E1p-D17Q is quite impaired in its E1-activity the mutation D17R (which has no effect on the E1-activity) has a stronger effect on the complex activity. Consequently, the impaired E1-activity of E1p-D17Q cannot (fully) cause the decrease in complex activity. Similar considerations are valid for E1p-D24R and E1p-D24Q; here it is also likely that the effect of E1p-D24Q on the complex-activity is (at least partly) caused by effects on the binding properties of this mutant.

Table 1 Summary of the kinetic parameters for the E1-reaction by *Azotobacter vinelandii* E1p and *A. vinelandii* E1p mutants and the binding properties of wild-type and mutant E1p as determined by analytical gelfiltration. $S_{0.5}$ and V_{max} values are calculated from V vs. [S] curves. The Hill coefficients are calculated from log([S]) vs. log((V_{max}-v)/v) curves. The reactions were performed as described in the experimental procedures. ND, not determined.

	$S_{0.5}$ (mM)	V_{max} (U mg^{-1})	Hill coefficient	Dissociation of E2p
E1p (wt)	0.03 ± 0.007	0.06 ± 0.003	1.1 ± 0.1	yes
D17Q	0.02 ± 0.008	0.02 ± 0.002	0.5 ± 0.2	no
D17R	0.02 ± 0.004	0.04 ± 0.001	1.0 ± 0.1	no
E20Q	0.008 ± 0.0009	0.06 ±0 .001	1.2 ± 0.1	yes
E20R	0.02 ± 0.003	0.09 ± 0.003	0.9 ± 0.1	yes
D24Q	0.02 ± 0.005	0.06 ± 0.003	1.0 ± 0.1	yes
D24R	0.01 ± 0.003	0.01 ± 0.0004	ND	no

Binding studies by analytical gelfiltration.

Free *A. vinelandii* E2p forms a 24-meric cubic core structure of about 1.5 MDa. Upon binding of wild type E1p or E3 the core dissociates into trimers with E1p or E3 bound (~400-800 kDa). These two forms of E2p can be separated by gelfiltration chromatography (table 2). The results obtained by this method correlate very well with the results that we described above.

The absence of dissociation of E2p by E1p-D17Q, E1p-D17R or E1p-D24R is analogous to that observed in figure 2 at "low" (2:1) E1p:E2p ratio; no, or very low PDHC activity is measured at this ratio using these mutants. These experiments confirm that the low PDHC-activities of E2p/E3 sub-complex recombined with either E1p-D17Q or E1p-D24R are not caused by their decreased E1-activity, but (at least partially) by a decreased binding affinity for E2p. Even though binding of E1p-D24Q to E2p results in dissociation of E2p, a considerable fraction of E1p-D24Q remains unbound (data not shown), which correlates well with the intermediate effect that this mutant has on the complex-activity as was measured above.

The elution volumes of all six mutant E1ps are identical to the elution volumes of wild type E1p. This demonstrates that none of the mutations affects the folding or dimerisation of E1p (as the molecular sizes of the mutants are not affected). This important observation shows that the observed effects on the activity and on the binding properties of these mutants are not caused by misfolding of E1p. Dimerisation of E1p is crucial for the binding of E1p to E2p and these results rule out that effects on the binding are caused by an indirect effect via an impaired dimerisation. Summarising, the above-described measurements clearly establish that D17 is essential for the binding of E1p to E2p. D24 on the other hand seems to be involved in the interactions, but does not seem to be essential. Mutation of the middle amino acid residue (E20) of the acidic "patch" has no influence on the interactions of E1p to E2p. This last observation indicates that the effects on D17 and D24 are specific and not due to local misfolding.

Model of the binding mode of E1p to E2p

Using the above-described results we have refined the model of the binding mode of E1p and E2p that was postulated before. A helical wheel of helix I of the proposed "helix-turn-helix" motif (amino acids 9-25) shows that the residues D17, E20 and D24 are on one side of the helix, whereas on the opposite side W15, L19 and V22 form a hydrophobic "patch". Helix II, of which it is less evident where it starts and finishes, also has a hydrophobic side, which can form a hydrophobic "pocket" with the residues of helix I. As shown above, we have established that E20 is not involved in interactions between E1p and E2p. D24 likely is (indirectly) involved in the interactions between E1p and E2p, whereas D17 is an essential binding residue. None of the mutations result in complete abolishment of the binding capacity of E1p to E2p; neither do point mutations in E2p (Schulze et al., 1991a). Apparently the

binding interface is composed of more than one region both on E2p (binding domain and catalytic domain) and on E1p.

Species specificity

Not only from their proposed location in the first helix of the N-terminal domain, but also based on species differences, it seems likely that D17, E20 or D24 are involved in the binding of E1p to E2p. Previously, it was shown that the binding of E1p to E2p is highly species specific (Schulze et al., 1992). Based on this, it seems likely that the binding residues in the N-terminal binding domain of E1p are not entirely conserved i.e. difference between species should be observed. Whereas the surrounding acidic residues (E11, E14 and E26) are highly conserved, D17, E20 and D24 of *A. vinelandii* E1p exhibit differences with some of the other species (figure 1). In *E. coli* E1p these residues are Q17, E20 and R24. Previously, it was shown, using chimeras of *E. coli* and *A. vinelandii* E2p that *E. coli* E1p solely binds to *E. coli* E2p and not at all to *A. vinelandii* E2p. Binding, on the other hand, of, visa versa, *A. vinelandii* E1p to *E. coli* E2p still results in 60% residual PDHC activity. Testing the reverse effect (whether the mutations D17Q and D24R result in an increased binding to *E. coli* E2p) is extremely difficult, as it would require *E. coli* E2p that contains absolutely no *E. coli* E1p. The fact that the mutations D17Q and D24R both have such a strong effect on the binding abilities of E1p nevertheless correlates well with the fact that *E. coli* E1p binds very poorly to *A. vinelandii* E2p.

Acknowledgements: This work was supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organisation for Scientific Research (NWO). We like to thank Adrie Westphal for his contributions to this work.

6 A structural alignment of the homodimeric and heterotetrameric 2-oxoacid dehydrogenases and transketolase.

Abstract

Several structures of thiamin diphosphate-dependent enzymes are known at this time. In striking contradiction with their structural similarity is their apparent lack of sequence similarity. Combining primary and secondary structural information enabled us to obtain a sequence alignment of homodimeric and heterotetrameric 2-oxoacid dehydrogenases and transketolase. This alignment consists of two parts, corresponding to the α- and β-strand of heterotetrameric 2-oxoacid dehydrogenase. Aligning the sequences of homodimeric pyruvate dehydrogenase (E1p) (*Azotobacter vinelandii* and *Escherichia coli*), transketolase (TK) (*Saccharomyces cerevisiae*), branched-chain oxoacid dehydrogenase (E1b) (*Pseudomonas putida* and human) and heterotetrameric E1p (*Bacillus stearothermophilus* and human), 36 residues are identical while 32 residues are highly conserved. These residues are evenly spread throughout the sequence, although multiple gaps and/or insertions had to be introduced. The most likely candidates for ThDP binding and active site residues were identified. Furthermore, a common function was suggested for regions shared by TK and the 2-oxoacid dehydrogenases. A few conserved regions remain that no definite function could be assigned to.

This new alignment opens possibilities to construct model structures for E1. Especially for the homodimeric E1p, that up to now has resisted crystallisation attempts, this will be of great value.

Introduction

With the recent report on the high-resolution 3-D structure of branched chain 2-oxoacid dehydrogenase (E1b) from *Pseudomonas putida* (Ævarsson et al., 1999) and *Homo sapiens* (Ævarsson et al., 2000) it has become possible to compare the structures of five different thiamin diphosphate (ThDP)-dependent enzymes; transketolase (TK) (Lindqvist et al., 1992, Nikkola et al., 1994), pyruvate oxidase (POX) (Muller & Schulz, 1993), pyruvate decarboxylase (PDC) (Dyda et al., 1993), pyruvate ferredoxin oxidoreductase PFOR (Chabrière et al., 1999) and E1b. These structures all contain the consensus sequence of the previously identified ThDP-binding site motif "GDG $X^{26/27}$ NN" (Hawkins et al., 1989) that is involved in the binding of the pyrophosphate end of the ThDP molecule using a Mg^{2+} or Ca^{2+} ion. Moreover, a common fold for ThDP-dependent enzymes was identified by

comparison of TK, POX and PDC (Lindqvist & Schneider, 1993, Muller et al., 1993). The recently solved E1b shows a similar fold (Ævarsson et al., 1999, Ævarsson et al., 2000).

In striking contradiction to this structural similarity is the apparent lack of sequence homology between different members of the ThDP-dependent enzyme family (Hawkins et al., 1989, Wexler et al., 1991, Burns et al., 1988, Matuda et al., 1991, Green, 1989). Especially the reported lack of sequence similarity between the homodimeric and the heterotetrameric form of pyruvate dehydrogenase (E1p) from the pyruvate dehydrogenase complexes (PDHC) is puzzling.

To solve this puzzle we have combined primary and secondary structural information to obtain a sequence alignment between the different ThDP-dependent enzymes.

Materials and methods

We used the amino acid sequence of *Saccharomyces cerevisiae* TK as reported in the 3-D structure (Lindqvist et al., 1992) because it differed considerably from the amino acid sequence as deduced from the nucleic acid sequence. We compared this sequence with the amino acid sequence from *Pseudomonas putida* E1b (E1bα and E1bβ) (Burns et al., 1988), Human E1b (E1bα and E1bβ) (Zhang et al., 1988, Chuang et al., 1990), Human E1p (E1pα and E1pβ) (Koike et al., 1988), *B. stearothermophilus* E1p (Hawkins et al., 1990), *Azotobacter vinelandii* E1p (Hengeveld et al., 1997) and *E. coli* E1p (Stephens et al., 1983). Attempts to obtain computerised alignments by several different algorithms all failed to yield results that were compliant with know structural and functional information. These initial homology searches all resulted in mis-alignments in the know functional motifs. For this reason first the ThDP-binding motifs of TK and *A. vinelandii* E1p were manually aligned. This partial alignment was used as starting point for an alignment using a standard alignment program (CLUSTAL W) (Thompson et al., 1994). The resulting alignment was of bad quality, but it showed that the 680 amino acids sequence of TK could be aligned with the entire 885 amino acid sequence of *A. vinelandii* E1p. This alignment was optimised by visual inspection. Secondly the primary sequence of TK and *P. putida* E1b were aligned on basis of their secondary structure. This alignment was also optimised by visual inspection. Heterotetrameric E1ps, heterotetrameric E1bs and homodimeric E1ps were aligned by using the standard alignment program. Heterotetrameric E1p vs. E1b was also aligned in this manner. The above-mentioned alignments (TK vs. homodimeric E1p, TK vs. heterotetrameric E1b and

heterotetrameric E1b vs. heterotetrameric E1p) were combined, on basis of both primary and secondary structural information, to the alignment shown in figure 1.

Results and Discussion

Multiple research groups (Hawkins et al., 1989, Wexler et al., 1991, Burns et al., 1988, Matuda et al., 1991, Green, 1989) previously concluded that homodimeric and heterotetrameric E1s show very low homology, restricted to the ThDP-binding motif. Robinson and Chun (Robinson & Chun, 1993) previously aligned different ThDP-dependent enzymes. Although they did state that all enzymes could be aligned, most of the functional groups were mis-aligned due to the lack of structural data.

Using secondary structure information now available for *P. putida* E1b (Ævarsson et al., 1999), human E1b (Ævarsson et al., 2000) and TK (Lindqvist et al., 1992, Nikkola et al., 1994) we were able to align amino acid sequences of homodimeric and heterotetrameric E1s and TK (figure 1). PDC and POX are organised in a different manner. An alignment between the two groups can only be obtained by dividing the PDC or POX sequence into two parts and reversing its order (Robinson & Chun, 1993). We have not attempted to make a detailed alignment is this way. The homodimeric oxoglutarate dehydrogenase (E1o) cannot be aligned to any other known ThDP-dependent enzyme (apart from the ThDP-binding motif) and is therefore not included in our alignment. The sequence alignment in figure 1 can be divided into two parts, corresponding to the α- and β-strand of heterotetrameric E1. 36 residues are identical in all four types of enzymes while 32 residues are highly conserved. These residues were evenly spread throughout the sequence. In 63 cases 3 out of 4 types of enzymes in the alignment had identical residues and in 63 cases 3 out of 4 were highly conserved. More important, we could identify most known functional groups in all four types of enzymes. We will discuss this in more detail below.

The α-strand shows three regions of high homology between the four types of enzymes, corresponding to amino acids 108-132, 153-191 and 208-252 of TK respectively. Two of these regions are involved in the binding of ThDP.

The first region contains the residues involved in binding the N_4' and N_3' of the amino pyrimidine-moiety of ThDP. The specific residues were identified as G_{116} and L_{118} for TK (Lindqvist et al., 1992, Nikkola et al., 1994, Schneider & Lindqvist, 1998), G_{182} and L_{184} for *P. putida* E1b (Ævarsson et al. 1999) and S_{207} and L_{209} from human E1b (Ævarsson et al.

2000). Homologous residues in homodimeric E1p are G_{195} and I_{197} for *A. vinelandii* E1p and G_{197} and I_{200} for *E. coli* E1p. In heterotetrameric E1p the residues were identified as and G_{165} and V_{167} (human) and I_{142} and I_{144} (*B. stearothermophilus*). The second region is the well-recognised ThDP-binding motif (Hawkins et al., 1989). It contains amino acids that bind the phosphate moiety of ThDP via a divalent metal ion (Ca^{2+} or Mg^{2+}). The residues were, among others, identified as D_{157} and N_{187} for TK (Wang et al., 1997), D_{228} and N_{258} for *A. vinelandii* E1p, D_{213} and N_{242} for *P. putida* E1b and D_{196} and N_{225} for human E1p. Moreover, it contains a conserved glutamate involved in increasing the protonation potential of the glutamate necessary for cofactor activation (E_{218} in *P. putida* E1b, E_{243} in human E1b and E_{162} in TK). The counterpart in homodimeric E1p is E_{233} (*A. vinelandii*) or E_{235} (*E. coli*). In heterotetrameric E1p however, no likely candidate could be identified. The third conserved region in the α-strand does not contain active site residues or any other residue that has a previously identified function. This region was absent in PDC and POX, indicating a function only present in TK and the 2-oxoacid dehydrogenases. In both TK and E1b this third conserved region is found on the exterior of the protein, close to the N-terminal residues. Wexler et al. (Wexler et al., 1991) previously identified this region as being highly conserved among heterotetrameric E1ps and E1bs but not in homodimeric E1ps. They suggested it to be involved in the αb-subunit interaction. We can now conclude from both our alignment (that includes homodimeric E1p) and the 3-D structures of TK and E1b that this is not the case. Since this region is found in both the 2-oxoacid dehydrogenases and in TK a function in the binding of E1 to another component of the multienzyme complex seems unlikely.

Besides these three conserved regions we identified conserved amino acids involved in the co-ordination of the pyrophosphate moiety of ThDP or catalysis. H_{69} (TK) (Wikner et al., 1997) interacts with the phosphate moiety of ThDP. Analogous residues in *P. putida* E1b (Y_{133}) and human E1b (Y_{158}) also form a hydrogen bond with a phosphate oxygen of ThDP (Ævarsson et al., 1999, Ævarsson et al., 2000). In heterotetrameric E1p this residue is conserved as a tyrosine while in homodimeric E1p a conserved histidine was found. *A. vinelandii* E1p is an exception to this: the histidine is replaced by a glutamine. H_{263} (TK) (Wikner et al., 1997) (H_{312} in *P. putida* E1b and H_{336} in human E1b) is oriented towards the reactive C_2 of ThDP and is likely involved in catalysis. This histidine is conserved throughout all 2-oxoacid dehydrogenases and TK.

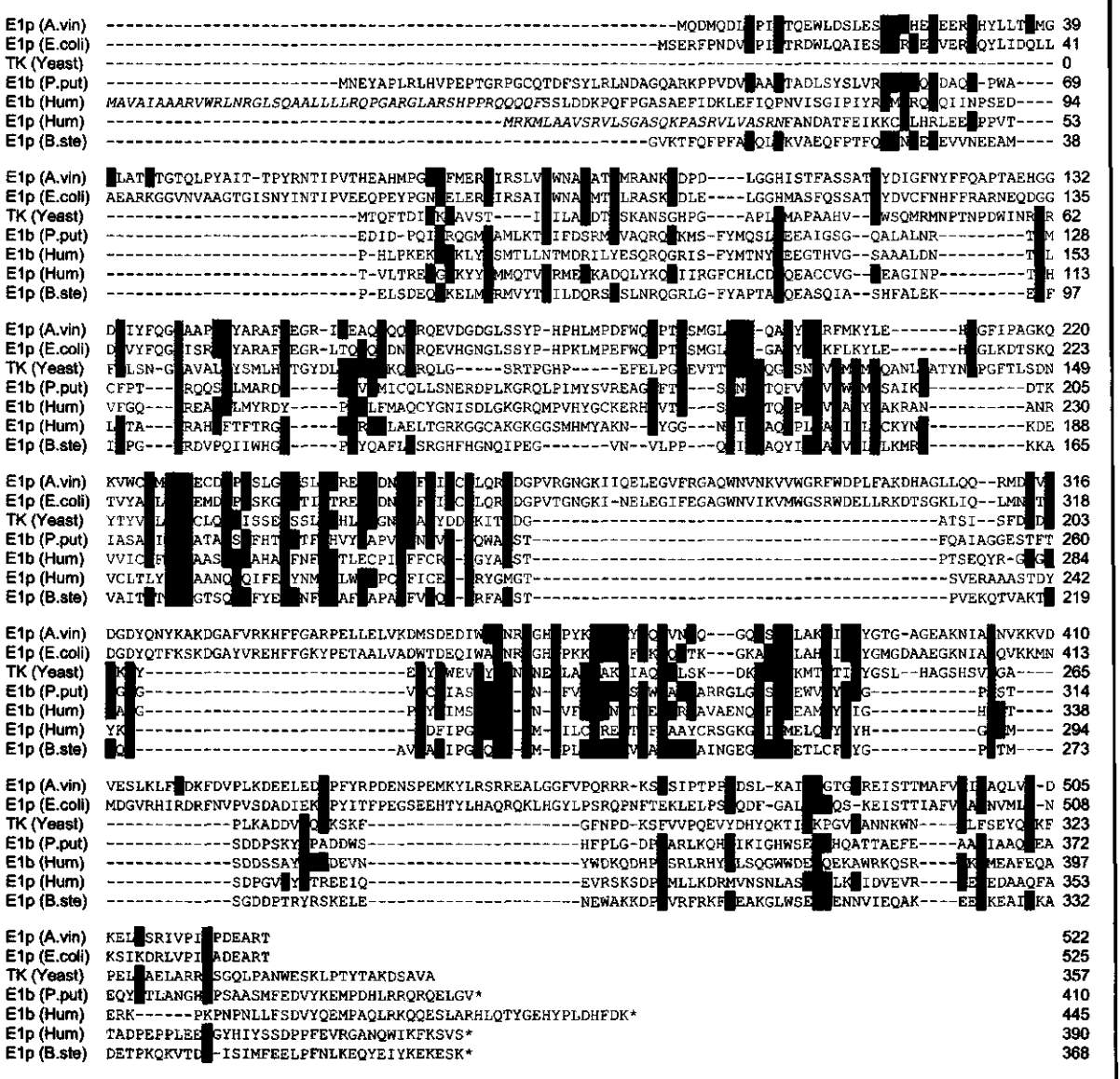
The phosphorylation sites in human E1 were previously identified as S_{293} and S_{300} for E1p (Korotchkina & Patel, 1995) and S_{337} and S_{347} in E1b (Zhao et al., 1994). As expected these amino acids are not found in the other enzymes since regulation by phosphorylation/dephosphorylation is unique to eukaryotic 2-oxoacid dehydrogenases. One S (S_{313}) is present in *P. putida* E1b, but does not function as a phosphorylation site. Interestingly, the first phosphorylation site is very near the catalytic active site histidine (H_{336} in human E1b).

The β-strand shows high homology in the region corresponding to amino acids 417-448 of TK. This region contains the main ThDP binding residues that are present in the β-strand. E_{418} (TK) (Wikner et al., 1994, Hawes et al., 1995), E_{89} for human E1p (Fang et al., 1998), E_{127} (human E1b) and E_{62} (*P. putida* E1b) co-ordinate the N_1' of the amino pyrimidine moiety of ThDP. This glutamate is important in cofactor activation. It is completely conserved in all 2-oxoacid dehydrogenases and TK. Secondly, hydrophobic residues in this region form a hydrophobic binding pocket for the pyrimidine moiety. In TK these residues are F_{442}, F_{445} and Y_{448} forming an FxxFxxY motif (Lindqvist et al., 1992, Nikkola et al., 1994). The analogous motif in E1b is FxxYxxP in which only the F and Y are found in the binding pocket. The residues FxxFxxX form this motif in heterotetrameric E1p. For homodimeric E1p the motif is slightly unclear and can either be FxxFxxM or YxxY/FxxF.

Besides the residues in the above-described region we identified another conserved histidine. H_{481} (TK) (Wikner et al., 1997) is near the N_4' of ThDP, oriented towards the reactive C_2 and likely involved in catalysis. This histidine is conserved in all 2-oxoacid dehydrogenases and TK.

The remaining part of the β-strand is very well conserved, although multiple gaps and/or insertions had to be introduced into the sequences. The C-terminal β-domain is not conserved throughout the entire ThDP-family, but it is found in all E1s, pyruvate:ferredoxin oxidoreductase (PFOR) (Chabrière et al., 1999) and TK. Its function in TK is unknown but for E1b a protein-protein (E1-E2) interaction function has been proposed (Ævarsson et al., 1999., Ævarsson et al., 2000). In particular it has been suggested that the C-terminal domain of the β-subunit binds to the small E1/E3-binding domain of the E2 core protein. In homodimeric E1p, on the other hand, limited proteolysis (Hengeveld et al., 1997) and mutagenesis experiments (Hengeveld et al., 1999) show that an extended N-terminal domain is responsible for the

Alpha chain



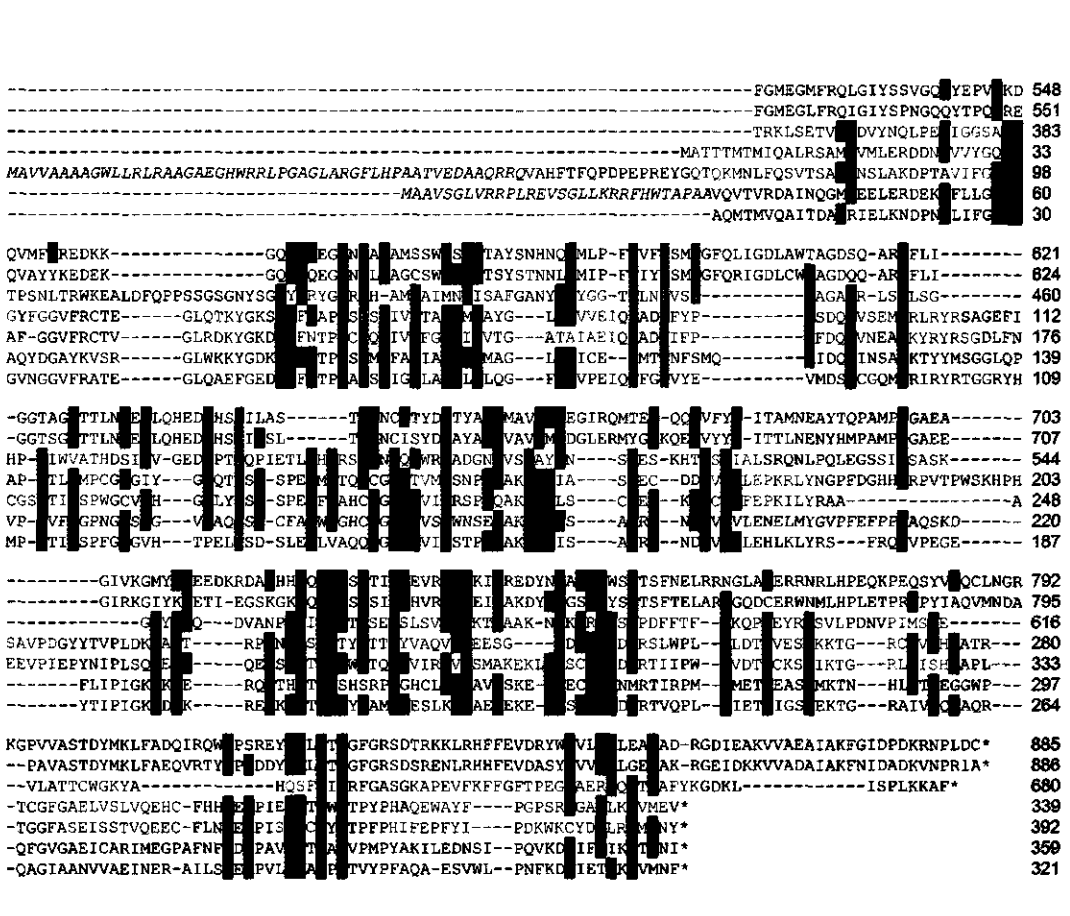


Figure 1 Sequence alignment of ThDP-dependent enzymes. The alignment was constructed as described in the methods and contains the following sequences: E1p (A.vin), *A. vinelandii* pyruvate dehydrogenase, accession no. Y15124; E1p (*E.coli*), *E. coli* pyruvate dehydrogenase, accession no. P06958 and P78049; TK (Yeast), *S. cerevisiae* transketolase, accession no. P23254; E1b (P.put), *P. putida* branched chain α-keto acid dehydrogenase (α and β chain), accession no. P09060 and P09061; E1b (Hum) branched chain α-keto acid dehydrogenase (α and β chain), accession no. P12694 and P21953; E1p (B.ste) *B. stearothermophilus* pyruvate dehydrogenase (α and β chain), accession no. P21873 and P21874; E1p (Hum) Human pyruvate dehydrogenase (α and β chain), accession no. P08559 and P11177. Residues of a known secondary structure are blue and green for α-helices and β-sheets respectively. Conserved residues are highlighted in red. Active site residues are highlighted in yellow. Proteolysis sites in *A. vinelandii* E1p are highlighted in green. The mitochondrion import sequences of human E1p and E1b are shown in italics.

binding to E2. Even more puzzling is the fact that the C-terminal domain of the β-subunit of heterotetrameric E1 is also found in homodimeric E1p, TK and PFOR (Chabrière et al., 1999) indicating a common function in all these enzymes. Clearly more experiments are necessary to elucidate the binding mode of both kinds of E1 to E2 and to reveal the function of the C-terminal domain.

Up to this point we tried to identify the function of identical or homologous residues and regions in this alignment. Besides the similarities between the four types of enzymes there are also some regions of large differences. First of all, in the N-terminal region the 2-oxoacid dehydrogenases have about 70 additional amino acids compared to TK. In *A. vinelandii* E1p this region forms an N-terminal domain necessary for binding to the E2 component of the complex (Hengeveld et al., 1999). This domain is extending from E1p by a flexible loop or linker that contains three proteolysis sites (Hengeveld et al., 1997) (figure 1). Some homology to this domain is found in *P. putida* E1b and *B. stearothermophilus* E1p, but, contrasting, the X-ray structure of *P. putida* E1b does not show an extended N-terminal domain. Parts of the N-terminal sequence of human E1p and E1b form a leader peptide necessary for import into the mitochondrion (Takakubo et al., 1995). The remaining part lacks any homology to the other two 2-oxoacid dehydrogenases.

All other large insertions in the alignment are found in the homodimeric E1p sequence. For three out of four regions we can suggest a putative function. The first region (264-311 in *A. vinelandii*) overlaps with a region that was identified as a flexible loop in apo-TK, involved in ThDP binding (Sundström et al., 1992) and could have a similar function in E1p. The second region (321-348 in *A. vinelandii*) was identified as a putative DNA-binding region (R339-L357) in *A. vinelandii* E1p (Regnström et al., 1999). The third region (405-424 in *A. vinelandii*) contains a limited proteolysis site (R418 in *A. vinelandii*) and therefore is likely to form an extended loop (Hengeveld et al., 1997). At this moment we cannot suggest any function to the fourth insertion (437-460 in *A. vinelandii*).

The above analysis shows that ThDP-dependent enzymes can be divided into two groups: the first group containing POX and PDC, the second group containing 2-oxoacid dehydrogenases and TK. We have analysed the second group in detail and shown that the members of this group have conserved regions throughout the entire sequence. We identified the most likely candidates for ThDP binding and active site residues. Furthermore, this group of ThDP-dependent enzymes contains conserved regions that we cannot assign definite functions to. The fact that both transketolase and 2-oxoacid dehydrogenases share these regions suggests a

common function. More research will have to be done to elucidate the function of these regions.

The availability of 3-D structures for *P. putida* E1b, human E1b and TK and this new alignment open possibilities to construct model structures for both homodimeric and heterotetrameric E1p. In combination with site-directed mutagenesis this will yield new insight into the structure-function relationships of E1p. Especially for the homodimeric E1p, that up to now, has resisted all crystallisation attempts this will be of great value.

Acknowledgements: This work was supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

7 Structural basis of the dysfunctioning of human 2-oxoacid dehydrogenase complexes.

A modified version of this chapter has been accepted for publication in Current Medicinal Chemistry A.F. Hengeveld and A. de Kok. Structural basis of the dysfunctioning of human 2-oxo acid dehydrogenase complexes.

Abstract

2-oxoacid dehydrogenase complexes are a ubiquitous family of multienzyme systems that catalyse the oxidative decarboxylation of various 2-oxoacid substrates. They play a key role in the primary energy metabolism: in glycolysis (pyruvate dehydrogenase complex), the citric acid cycle (2-oxoglutarate dehydrogenase complex) and in amino acid catabolism (branched-chain 2-oxoacid dehydrogenase complex). Malfunctioning of any of these complexes leads to a broad variety of clinical manifestations. Deficiency of the pyruvate dehydrogenase complex predominantly leads to lactic acidosis combined with impairment of neurological function and/or delayed growth and development. Maple urine disease is an inborn metabolic error caused by dysfunction of the branched-chain 2-oxoacid dehydrogenase complex. An association between both Alzheimer disease and Parkinson's disease and the 2-oxoglutarate dehydrogenase gene has been reported.

Currently a wealth of both genetic and structural information is available. Three-dimensional structures of three components of the complex are presently available: of the pyruvate dehydrogenase component (E1), of the dihydrolipoyl acyltransferase component (E2) and of the lipoamide dehydrogenase component (E3). Moreover, detailed information on the reaction mechanism, regulation and the interactions between the different components of the complex is now at hand. Although only one of the structures is of human origin (E1b), model building by homology modelling allows us to investigate the causes of dysfunction.

In this review we have combined this knowledge to gain more insight into the structural basis of the dysfunctioning of the 2-oxoacid dehydrogenase complexes.

Introduction

2-oxoacid dehydrogenase multienzyme complexes constitute a family of enzymes that catalyse the irreversible oxidative decarboxylation of various 2-oxoacids, generating the corresponding acyl-CoA derivative, NADH and CO_2 (Perham, 1991; Mattevi et al., 1992a; De Kok et al., 1998; Patel & Roche, 1990). In man the members of this family comprise the pyruvate dehydrogenase complex (PDHC), the oxoglutarate dehydrogenase complex (OGDC)

and the branched-chain 2-oxoacid dehydrogenase complex (BCDC). All three of these complexes are located in the mitochondrial matrix compartment, associated with the inner membrane. Unique mitochondrial protein kinases and protein phosphatases are dedicated to the regulation of the complexes and closely interact with the multienzyme complexes (Roche, et al., 1996; Reed et al., 1996; Harris & Popov, 1996; Randle & Priestman, 1996; Abe, et al., 1996). Each of the three members functions at a strategic point in primary energy metabolism: in glycolysis (PDHC), the citric acid cycle (OGDC) or amino acid catabolism (BCDC). PDHC uses pyruvate to produce acetyl CoA, which is used in the citric acid cycle. OGDC has part in transforming 2-oxoglutarate into succinyl CoA, while BCDC catalyses the oxidative decarboxylation of branched-chain α-keto acids (BCKA) produced by the transamination of the branched-chain amino acids (BCAA) valine, leucine and isoleucine. As glucose is of critical importance to some cells in the central nervous system conservation of it is of vital importance during carbohydrate depletion. Since the PDHC-reaction controls the oxidative utilisation of glucose, at times of depletion this reaction is reduced and fatty acids or ketone bodies are used as a source for acetyl-CoA in the central nervous system (Randle, 1986). In the brain, on the other hand, acetyl-CoA is nearly exclusively synthesised from pyruvate. Acetylcholine, the excitatory neurotransmitter is synthesised from acetyl-CoA. PDHC-activity is also reduced under diabetic conditions or starvation (Roche et al., 1996). The regulation of these mechanisms is primarily controlled by the phosphorylation / dephosphorylation of E1 (Roche et al., 1996).

PDHC, OGDC and BCDC convert their respective substrates (figure 1) to the corresponding acyl-CoA derivative by the combined action of three enzymes: an α-keto acid dehydrogenase component (E1p, E1o or E1b), an acyltransferase component (E2p, E2o or E2b) and a dihydrolipoamide dehydrogenase component (E3). The last component (E3) is shared by all three enzymes. In PDHC an additional component is involved in the catalytic mechanism: the E3-binding component (E3BP, previously known as protein X).

Malfunctioning of any of the complexes leads to a broad variety of clinical manifestations. Because of their important functions complete inactivation of any of the complexes is not expected and not found. Deficiency of PDHC leads to lactic acidosis, primary leading to impairment of neurological function, growth and development. Maple syrup urine disease (MSUD) is an inborn error caused by dysfunctioning of BCDC. Both deficiencies have been reviewed several times in the past years (Wynn et al., 1996; Indo & Matsuda, 1996; Kerr, et

al., 1996; Danner & McKean, 1996; Chuang & Shih, 1995; Robinson, 1995; Kerr & Zinn,

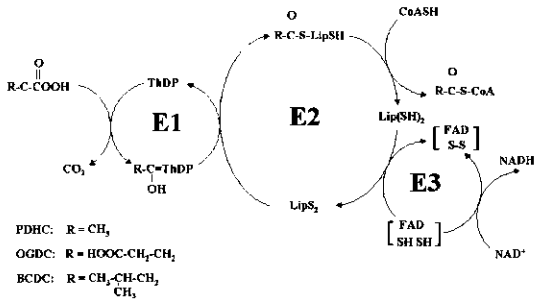


Figure 1 Reaction mechanism of 2-oxoacid dehydrogenase complexes. A 2-oxoacid dehydrogenase (E1p, E1b or E1o) uses thiamin diphosphate (ThDP) to carry out the oxidative decarboxylation of the substrate, forming an acylated lipoic acid, covalently attached to the acyltransferase component (E2p, E2b or E2o) and CO_2. The E2-component catalyses the transfer of the acyl group to coenzyme A (CoA). The resulting dihydrolipoyl group is then reoxidised by the dihydrolipoamide dehydrogenase component (E3) using NAD^+, forming NADH.

1995; Dahl, 1995; Patel & Harris, 1995; Brown, et al., 1989a).

An association of both Alzheimer disease and Parkinson's disease with OGDC has been reported (Kish, 1997; Mizuno et al., 1994).

These multienzyme systems form large and complicated complexes with molecular weights ranging up to 10 million Dalton. They share a common building design in which many peripheral components bind to a central core with either octahedral (24-subunits) or icosahedral (60-subunits) symmetry.

Currently three-dimensional structures of all three components of the complex are available: E1b from *Pseudomonas putida* (Ævarsson et al., 1999) and man (Ævarsson et al., 2000), E2p (Mattevi et al., 1992c; Knapp et al., 1998; Izard et al., 1999 and Dardel et al., 1993; Green et al., 1995; Berg et al., 1996b; Berg et al., 1997; Ricaud et al., 1996 and Kalia et al., 1993; Robien et al., 1992; Mande et al., 1996) and E3 (Toyoda et al., 1998; De la Sierra et al., 1997; Mattevi et al., 1991; Mattevi et al., 1992b; Mattevi et al., 1993b) from several organisms. Although most of these structures are not of human origin, model building by homology modelling allows us to investigate the causes of dysfunctioning of human complexes. In this

review we will combine the current wealth of both genetic and structural information to gain more insight into the structural basis of the dysfunctioning of the 2-oxoacid dehydrogenase complexes. We will concentrate on those effects of dysfunctioning that have a molecular explanation. In order to understand the effect of different mutations in the complexes, first we will address the structure and functioning of the individual components of the complex and then their organisation within the complexes. Following, we will extrapolate this information to the human complexes.

Dysfunctioning of 2-oxoacid dehydrogenase complexes

As all 2-oxoacid dehydrogenase complexes hold crucial roles in energy metabolism deficiency of any of the complexes seems to be entirely incompatible with survival. We will try to shed some light onto this paradox by attempting to explain these deficiencies on a molecular and structural level. Previously, mutations in the chromosomal DNA of patients were identified, inventoried and structure-function relationships were postulated. Several review articles on both PDHC and BCDC-deficiencies have been published recently (Wynn et al., 1996; Indo & Matsuda, 1996; Kerr et al., 1996; Danner & McKean, 1996; Chuang & Shih, 1995; Robinson, 1995; Kerr & Zinn, 1995; Dahl, 1995; Patel & Harris, R.A., 1995; Brown et al., 1989a). In case of E1 this has always been with the lack of an actual structure. Very recently the X-ray structure of E1b from *P. putida* and human origin were solved (Ævarsson et al., 1999; Ævarsson et al., 2000). This makes it possible to access the structure-function relationships of identified mutations in the availability of a structure. Mutations in E1b that cause MSUD were evaluated in this manner by Ævarsson et al. (Ævarsson et al., 2000). For this reason our focus in this review will be mainly on E1p. In the next section we will re-evaluate the structure-function relationship of identified mutations of E1. Brief sections will be devoted to E2, E3 and E3BP. Finally we will discuss the assembly of the complex and mutations that influence the complex activity on this level of organisation.

PDHC-deficiency is extremely heterogeneous in its clinical manifestations (Robinson, 1995; Dahl, 1995; Brown & Brown, 1994). All forms include lactic acidosis, ranging from very severe, mostly resulting in death at birth, to more moderate clinical manifestations. Other

manifestations of PDHC-deficiency are neural disorders resulting in mental retardation or only mild developmental delay (Robinson, 1995). A poor correlation has been found between the severity of the symptoms and the residual activity of PDHC measured in different tissues. Factors that could attribute to this bad correlation are differences in phosphorylation and the amount of enzyme found in different tissues. The E1pα gene is located on the X-chromosome (Brown et al., 1989), which explains the relatively high prevalence and large heterogeneity of its mutations (Dahl, 1995; Brown et al., 1989). Random inactivation of the major part of one X-chromosome occurs in somatic cells of females during early embryogenesis; therefore, women heterozygous for a polymorphic X-linked gene have a mixture of cells expressing one or the other allele in their tissues (Lyon, 1962). Thus, a variable pattern of X-chromosome inactivation in females can also explain the poor correlation between PDHC activity and the severity of the symptoms. Up to now no null mutations of E1pα were identified in males, confirming that the total absence of PDHC would be lethal. In PHDC at least 65 mutations in E1pα were reported, while no mutations were reported to effect the E1pβ subunit or the E2p or the regulatory enzymes E1-kinase and phosphatase. Only a few mutations were reported in E3 (Randle & Priestman, 1996) and E3BP (Patel & Roche, 1990).

Leigh syndrome (LS) is a progressive neurodegenerative disorder with onset usually in infancy or early childhood (Leigh, 1951; Rahman et al., 1996). LS has been associated with deficiencies of PDHC and of respiratory chain complexes. Moreover, mitochondrial DNA point mutations were found in LS patients. The portion of LS patients that were identified to have PDHC-deficiencies all had mutations in the E1pα gene. We discuss these mutations together with the other PDHC-deficiency causing mutations.

MSUD is an autosomal recessive inborn error caused be mutations in any of the components of BCDC. All forms of MSUD result in keto-acidosis due to the accumulation of BCAA in plasma resulting in increased urinary excretion of the corresponding BCKA. The BCAA exert a direct neurotoxic effect, although its mechanism is yet unclear (Schadewaldt & Wendel, 1997). This results in clinical effects ranging from mild symptoms to mental retardation and even death. The distinctive odour of the urine of the patient is caused by sotolone, a flavour component also found in maple syrup (Podebrad et al., 1999). The wide range of residual activity of BCDC has resulted in a classification of patients with different clinical features into five phenotypes: classic, intermediate, intermittent, thiamin-responsive and E3-deficient (Chuang & Shih, 1995). Patients with an E3-deficiency are deficient in all three 2-oxoacid

complexes and therefore usually show very severe clinical symptom, usually resulting in early (infant) death. There seems to be a correlation between the severity of the symptoms of the disease and the measured activity of BCDC. Most patients are now identified through new born screening programs. The incidence of MSUD varies from 1/150.000-1/200.000 in a general population (Danner & McKean, 1996) to 1/176 within an isolated Mennonite population (Peinemann & Danner, 1994). When an identified patient maintains a diet restricted in BCAA, sometimes supplemented with thiamin, it usually shows normal development (Hilliges et al., 1993; Treacy et al., 1992).

Even though no molecular evidence was found up to now to link the OGDC with inborn errors several reports have been published that show a correlation between the activity of this complex and for example Alzheimer's disease (AD) (Kish, 1997; Cruts et al., 1995), Parkinson's disease (Mizuno et al., 1994), Wernicke-Korsakoff (Butterworth et al., 1993) or Freidrich's ataxia (Mastrogiacomo et al., 1996). As our scope in this review is to find relations between the molecular basis of the disease and the structure-function relationship of the corresponding enzyme we will not discuss this complex in any detail.

Lissens et al. have recently publicised a mutation update on all the mutations that are found for the E1p gene (Lissens et al., 2000). They have evaluated the occurrence of these X-linked mutations in males and females and the diagnostic relevance of this evaluation. The total number of female and male patients is almost the same, although a difference in the distribution of the type of mutations is seen between the sexes. Ævarsson et al. list and discuss all mutations that cause MSUD found in E1b (Ævarsson et al., 2000). We have summarised this information again, only listing mutations that result in point mutations, small deletions or insertions into the amino acid sequences of the different components of the complexes in Table 1. Table 1 also includes mutations found in E3.

Alignment analysis: extrapolation to human structures

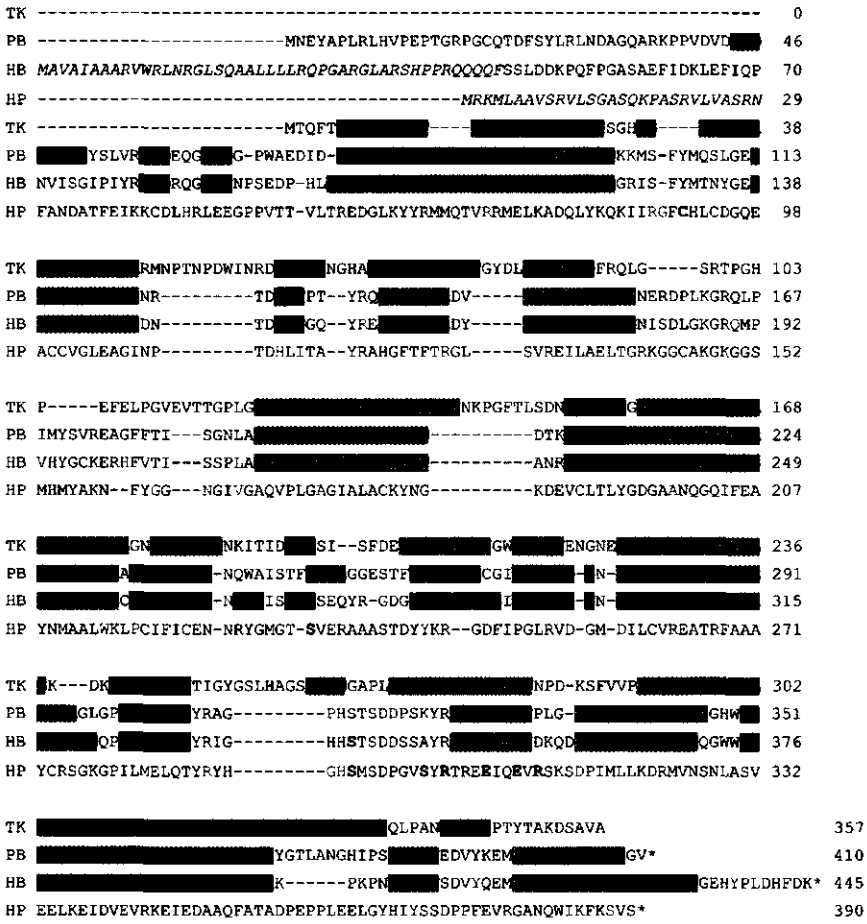
Most structural information available on the catalytic components of the multienzyme complexes is from sources other than human or even mammalian. The availability of a large number of amino acid sequences of the different components of the complex makes it possible

to extrapolate known structural information to the human complexes. For both E2 and E3 multiple structure based sequence alignments have been published before (Mattevi et al., 1992c; Jentoft et al., 1992; Russell & Guest, 1991a) that show a high degree of similarity between enzymes from the different sources. No specific consensus sequences were found that distinguish E2-cores based on 24 subunits as found in OGDC and BCDC from those based on 60 subunits as found in eukaryotic PDHC (Izard et al., 1999). All E2 sequences contain from N- to C-terminus 1 to 3 lipoyl domains, an E1/E3 binding domain and a catalytic domain, separated by alanine/proline rich linker sequences. The main difference in E2 sequences is the number of lipoyl domains present. Mammalian OGDC forms an exception to this rule as it lacks the binding domain. The binding of the E1 and E3 component to E2 will be discussed further along this review.

For E1 the alignments were up to now limited to E1's from the same kind of complex and even from the same class: either homodimeric or heterotetrameric (Wexler et al., 1991; Burns et al., 1988; Matuda et al., 1991). All OGDC and PDHC from gram-negative bacteria have homodimeric E1, BCDC and PDHC from eukaryotes contain heterotetrameric E1. Determination of the X-ray structure of E1b from *P. putida* (Ævarsson et al., 1999) and human (Ævarsson et al., 2000) has shown that the heterotetrameric E1b has a large structural homology to transketolase (TK). As TK in its turn shows significant (24%) sequence similarity to homodimeric E1p we are now able to combine the information available for both homodimeric and heterotetrameric E1. Figure 2 shows the alignment of human E1p and E1b with E1b from *P. putida* and TK from Yeast (TK was recently reviewed by (Schneider & Lindqvist, 1998)). This alignment was obtained based on both primary and secondary structural information using standard alignment programs and optimised by visual inspection.

The structure of the homodimeric E1o remains a puzzle as, besides the well recognised ThDP-binding motif (Hawkins et al., 1989), no significant homology to any known ThDP-dependent enzyme could be found. Based on the large structural similarity found between different ThDP-dependent enzymes it is very likely though that E1o has a similar structure as other ThDP-dependent enzymes and E1 in particular.

Alpha strand



Beta strand

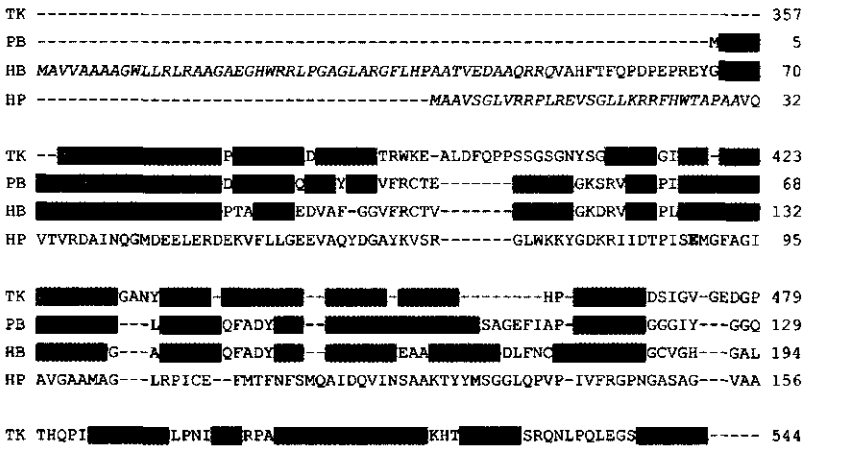


Figure 2: Sequence alignment of E1 amino acid sequences from several sources and transketolase based on structural homology. TK transketolase from Yeast, PB *Pseudomonas putida* E1b, HB human E1b, HP human E1p. Known secondary structure is highlighted in green and yellow: α-helices are highlighted in yellow, β-sheets are highlighted in green. Mutations in human sequences that cause dysfunctioning of the compexes are shown in red (see also Table 1). Site directed mutant with known function (Table 2) are shown in bold and blue (in HB and HP). The putative lipoyl domain binding residues are shown in blue (in PB). The leader peptide sequences are shown in italics.

A structural analysis of the dysfunctioning of E1

Replacement of active site residues

The E1 component catalyses the reductive acylation of the E2 bound lipoyl groups (figure 1). It uses thiamin diphosphate (ThDP) to carry out the oxidative decarboxylation of the substrate (figure 3). Excellent reviews on ThDP-chemistry and ThDP-enzymes have been published recently in a special issue of Biochimica et Biophysica Acta (Schellenberger 1998). The cofactor-binding site is highly conserved between the different members of the ThDP-dependent enzymes (Muller et al., 1993; Lindqvist & Schneider, 1993). Both by site directed mutagenesis and chemical modification several catalytically important amino acid residues have been identified in eukaryotic E1 (Table 2).

The recent determination of the 3-D structures of human and *P. putida* E1b (Ævarsson et al., 1999; Ævarsson et al., 2000) has added a great deal of knowledge with respect to active site

residues of E1. Figure 4 shows the residues involved in ThDP-binding in both human and *P. putida* E1b and in TK from yeast (Lindqvist et al., 1992; Nikkola et al., 1994). Essentially, most residues involved in ThDP-binding in E1b are conserved in all E1's and TK (figure 2 and 4). All ThDP-enzymes require bivalent metal ions (Mg^{2+} or Ca^{2+}) for coenzyme binding (Wang et al., 1997).

Table 2. Catalytically important amino acid residues in eukaryotic E1 as identified by site directed mutagenesis or chemical modification.

Subunit	Residue [1)]	Function	Source	Reference
E1pα	C91	Catalysis	Bovine, human	Ali et al., 1993 Korotchkina et al., 1999
E1pα	S232	Phosphorylation	Human	Korotchkina et al., 1995
E1pα	S293	Phosphorylation	Human	Korotchkina et al., 1995
E1pα	S300	Phosphorylation	Human	Korotchkina et al., 1995
E1pα	R302	Activation by proteolysis	Porcine	Koike et al., 1992
E1pα	E306	Activation by proteolysis	Porcine	Koike et al., 1992
E1pα	E309	Activation by proteolysis	Porcine	Koike et al., 1992
E1pα	R311	Activation by proteolysis	Porcine	Koike et al., 1992
E1p	R??	Catalysis	Pigeon	Nemerya et al., 1996
E1pβ	E89	Catalysis	Human	Fang et al., 1998
E1pβ	W165	ThDP-binding	Human	Korotchkina et al., 1999; Ali et al., 1995
E1pβ	R269	Pyruvate binding	Bovine	Eswaran et al., 1995
E1bα	S337	Phosphorylation	Human	Zhao et al., 1994
E1bα	S347	Phosphorylation	Human	Zhao et al., 1994

[1)] The amino acid is numbered according to the human homolog.

The diphosphate-metal-anchor is the main linker group to the protein, the metal ion functions in complexing the pyrophosphate group to a conserved group of amino acid side chains in the

apo-enzyme. These amino acids are found in the ThDP-binding motif, a conserved sequence found in all ThDP-dependent enzymes (GDG-$(X)_{25\text{-}30}$-N) (Hawkins et al., 1989).

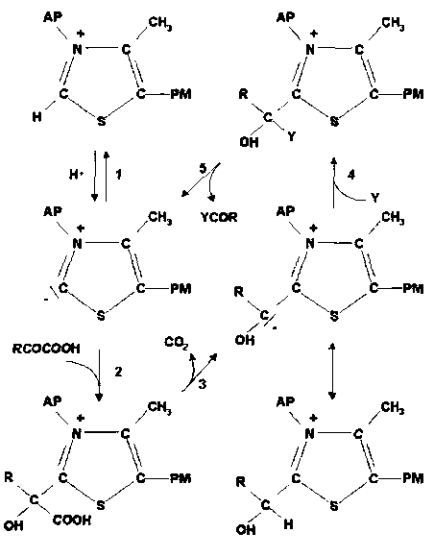


Figure 3: The catalytic mechanism of ThDP enzymes is described as a cyanide-analogous charge catalysis (Schellenberger, 1998). The deprotonation of the catalytic centre of the coenzyme (the C_2-atom of the thiazolium moiety) is the key step in the reaction (1). The C_2-H dissociation rate is enhanced about 4 orders of a magnitude through the action of a conserved glutamine via the N_1'and the 4'-NH_2 group of the amino-pyrimidine moiety of ThDP (Kern, 1997). The so-called V-conformation of ThDP (Shin, 1977) that is found for all enzyme-bound ThDP's brings the 4'-NH_2 group and the C_2-H in direct contact. The C_2-carbanion forms a nucleophilic adduct with the substrate (2) which is rapidly decarboxylated (3) producing a ThDP-bound enamine. The enamine then reacts as a carbanion with lipoamide to form a tetrahedral intermediate (4). Rearrangement results in the formation of the acylated dihydrolipoamide and ThDP (5). The reaction mechanism of the second substrate (lipoamide) is far from understood.

A tyrosine is stacked against the amino pyrimidine ring; hydrophobic sandwich the amino pyrimidine and thiazolium ring resulting in the characteristic V-conformation of enzyme bound ThDP. An important residue in cofactor activation is a glutamate from the β-subunit (E62β in *P. putida*), conserved in all E1's and TK. The protonation potential of this glutamate is increased by a second conserved glutamate (E218α in *P. putida*) (Dyda et al., 1993; Arjunan et al., 1996; Meshalkina et al., 1997). A conserved glycine (G182α in *P. putida*) forms a second H-bond with the 4'-N atom of ThDP. Side chains of two conserved histidines (His312α and His 131β in *P. putida*) are oriented toward the reactive C_2. Both histidines are conserved in all E1's and TK and are most likely involved in catalysis. In E1b H131β is covalently bound to the substrate analogue inhibitor a-chloroisocaproate (Ævarsson et al.,

1999). H131β is equivalent to H481 in TK that is proposed to be involved in substrate recognition and binding. H312α is analogous to H263 in TK and both are proposed to be involved in polarisation of the substrate and catalysis (Wikner et al., 1995; Wikner et al., 1997).

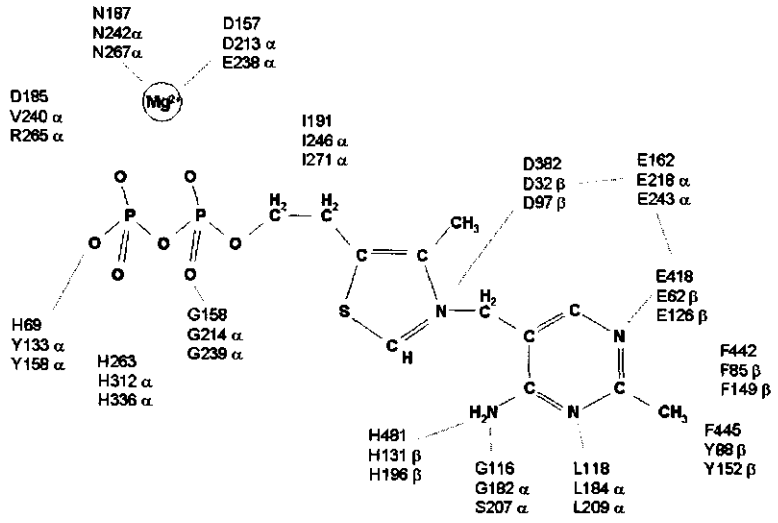


Figure 4: Thiamin diphosphate (ThDP)-protein interactions in the active site of, from top to bottom: transketolase (TK), human E1p and human E1b. Only conserved interactions are shown.

The second substrate in the reaction mechanism, lipoamide, is covalently attached via an aliphatic lipoyl-lysine arm, specifically attached to a lipoyl domain. A channel lined with hydrophobic residues was identified in E1b positioning the disulphide of the lipoamide exactly in between the two previously mentioned active site histidines. Both are, from this model, very likely involved in catalysis, possibly as proton donor in the reduction of the disulphide bond.

Free lipoamide is a very poor substrate for E1 and the binding of the lipoyl domain to E1 is very specific for the parent complex. Surface regions that are possibly involved in the recognition and binding of the lipoyl domain were proposed for E1b (figure 5). These proposed regions form the rims of the lipoamide channel opening.

Active site mutants of E1pα

Of all mutations found in E1pα the mutations that have a very clear structure-function relationship are those found in its active site, involved in the binding of ThDP. The residues involved in the binding of ThDP are, as mentioned above, highly conserved throughout all ThDP dependent-enzymes and thus extrapolation does not give any problems for these residues.

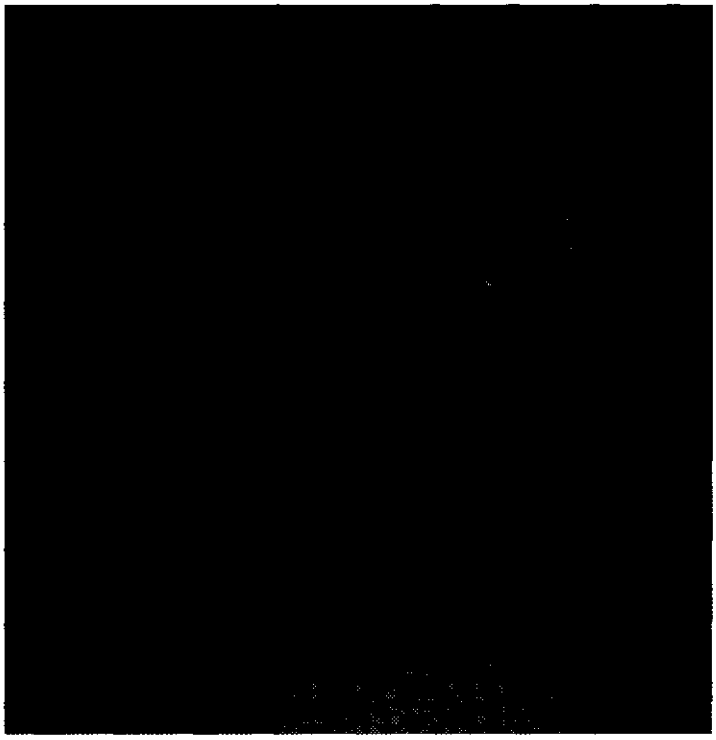


Figure 5: Overall structure of human E1b (Ævarsson et al., 2000).

The active site mutation that is probably the easiest to explain is V167M. V167α is conserved as a large hydrophobic amino acid chain (V, L or I) in all ThDP-dependent enzymes (L184α in *P. putida* E1b). It interacts with the N3' of the amino-pyrimidine moiety of ThDP, supporting the V-conformation of ThDP. Site directed mutagenesis was performed on this residue in pyruvate decarboxylase (PDC) from *Saccharomyces cerevisiae* (I415) showing that decreasing

the size of this residue leads to diminished specific activity (Guo et al., 1998; Jordan et al., 1996).

The importance of this side chain is also shown by the effect of mutations on nearby residues as T231α. Changing T231α to an alanine will affect the position of V167, having a similar effect as reducing the size of V167.

The residues G162α and N164α are not directly involved in interactions with ThDP, but both are very near the glycine residue that interacts with the N4' of the amino pyrimidine moiety. Alteration of these two residues (G162R and N164S) will therefore have an effect on the positioning of the residues that interact with the amino pyrimidine moiety.

Only one mutation was found in the region where the phosphate moiety of ThDP interacts: A199T. At first glance it is surprising that this mutation has such a drastic effect because human A199α is equivalent to a threonine in *P. putida*. Looking at the surrounding amino acids in both human E1p (DGAAN) and P. putida E1b (DGATA) or human E1b (EGAAS) suggest that the mutation A199T might result in a small change in the positioning of the helix that positions the phosphate moiety binding residues.

Two other mutations in the active site are G291R and H292L. H292 is the highly conserved histidine (H312α in *P. putida* E1b) that is proposed to be involved in polarisation of the substrate and catalysis. As expected the effect of this mutation is very severe, resulting in only 2% PDHC activity in fibroblasts. The mutation was found in a female patient and the (short) survival of this patient is probably due to the variable X-chromosome inactivation pattern. In TK this histidine (H263 in TK) was mutated resulting in severely impaired specific activity (Wikner et al., 1997). Changing its neighbouring residue from a glycine to an arginine will drastically affect the position and pKa of the histidine and thus affect the catalytic activity of E1.

Human P217α is conserved in heterotetrameric E1. It forms the turn between an α-helix and a β-sheet that both contain several residues involved in ThDP binding. Mutating the proline to a leucine will drastically effect the positioning of this helix and sheet. Moreover, this β-sheet forms part of the hydrophobic core of the enzyme. Site-directed mutagenesis of this residue shows that indeed the thermostability of mutants is strongly affected (and thus the structural integrity of the enzyme) (Tripatara et al., 1999).

5 of the E1pa mutations listed in Table 1 are responsive to ThDP-treatment (H44R, R88S, G89S, R263G and V389fs). Of these 5 residues only R88 and G89 are clearly near the active

site. Even though both are not conserved active site residues, they are close to a conserved phenylalanine (F105 in *P. putida* E1b or F90 in human E1p) that is postulated to be involved in substrate binding. Thus, alteration of neighbouring residues will affect the catalytic activity of the enzyme. Moreover, C91 in human E1p has been identified as an active site residue by site directed mutagenesis (Table 2) (Ali et al., 1993; Korotchkina et al., 1999).

Why the other three mutations are ThDP-responsive is not clear, all are found on the surface on the enzyme and not in any direct contact with the active site. As they are more likely involved in tetramerisation or interactions with E2, they will be discussed further in this section.

The last mutation that clearly effects the active site is R288fs. This mutation results in a 283 amino acid truncated protein, of which part of the active site has been deleted. Most likely this mutant will not result in any stable folded protein and the residual activity that was measured (88%) in the female patient is probably due to the X-chromosome inactivation pattern.

Regulatory sites in E1

The activity of both PDHC and BCDC are primarily controlled by a phosphorylation-dephosphorylation cycle. Specific regulatory enzymes carry out this cycle: E1-kinase and E1-phosphatase. Pyruvate and ADP inhibit E1-kinase. As a result E1 is activated at high pyruvate concentrations and at high ATP demand. Dichloroacetate (DCA) also inhibits E1-kinase and is therefore used in the treatment of PDHC deficient patients. In human E1p and E1b 3 and 2 serine residues respectively have been identified that can be phosphorylated (E1p: S232α, S293α, S300α and E1b: S337α, S347α) (Table 2) (Zhao et al., 1994; Korotchkina et al., 1995). Phosphorylation of the primary sites (S337 for E1b and S293 for E1p) results in inactivation of the complexes (Zhao et al., 1994; Hawes et al., 1995). These primary sites are located very close to the active site, right next to a conserved histidine that is involved in catalysis (H312α in *P. putida*). A few mutations are found in the vicinity of the third phosphorylation site (S300α): R302C, R302H, R302L, and the insertion mutants E306ins7 and I306ins11. R302 is a conserved surface residue very close to S300α. It is also in proximity of the conserved histidine and the primary phosphorylation site. Changing this large charged residue into a hydrophobic residue would quite likely affect the surrounding residues or even the stability of the entire enzyme. Both insertion mutants will affect the level of phosphorylation and

therefore the catalytic activity of the enzyme. Since the insertions are found at the surface of E1p, they will probably not be as destructive as if they would have been found in the interior of the enzyme. In the last case this would probably lead to complete disruption of the structural core.

The mutation T231A that has already been mentioned above will probably, besides its effects on the active site, also influence phosphorylation, because of its position right next to S232α. It can be expected though that from both effects the disruption of the active site will have the largest effect on catalysis.

Protein-protein interactions in E1

Human E1b and E1p are heterotetramers formed by two α and two β-subunits (figure 5). In *P. putida* E1b the α-subunit forms a large domain, an extended N-terminal tail and a small C-terminal domain. The β-subunit is composed of two domains. The interactions between the α and β-subunit are mainly between the α-domains and the N-terminal β-domain. The C-terminal β-domain is not conserved throughout the entire ThDP-family, but it is found in all E1's, pyruvate:ferredoxin oxidoreductase (PFOR) (Chabrière et al., 1999) and transketolase (TK). Its function in TK is unknown but for *P. putida* E1b a protein-protein (E1-E2) interaction function has been proposed (Ævarsson et al., 1999).

Besides the region responsible for the specific interactions with the lipoyl domain, a second region in E1 is important for protein-protein interactions: the region that binds E1 to E2. In heterotetrameric E1's results have been obtained that indicate that the β-subunit is responsible for these interactions (Ævarsson et al., 2000; Perham, 1996; Wynn, 1992a; Lessard et al., 1996). In particular it has been suggested that the C-terminal domain of the β-subunit binds to the small E1/E3-binding domain. In homodimeric E1, on the other hand, limited proteolysis (Hengeveld et al., 1997) and mutagenesis experiments (Hengeveld et al., 1999) show that an extended N-terminal domain is responsible for the binding to E2. Even more puzzling is the fact that the C-terminal domain of the β-subunit of heterotetrameric E1 is also found in TK and PFOR (Ævarsson et al., 1999). Alignment of homodimeric E1p with TK also reveals the presence of this C-terminal domain, indicating a common function in all these enzymes. Clearly more experiments are necessary to elucidate the binding mode of both kinds of E1 to E2 and to reveal the function of the C-terminal domain.

The fact that ThDP binds on the αβ-interface does not make it surprising that many mutations are found that affect these interactions. Disturbing the subunit interactions will very likely affect ThDP binding and therefore catalysis. Three surface residues in E1pα were found to be responsive to thiamin treatment (Table 1). The H44R mutation is particularly difficult to explain, as the N-terminal part of the different E1's have a low homology. Using the alignment in figure 2, H44 would be analogous to Q61 in *P. putida* E1b. This residue is an exposed surface residue analogous to residues comprising the N-terminal domain of *A. vinelandii* E1p. This domain is involved in the binding of E1p to E2p (Hengeveld et al., 1997). Removal of this domain also had effect on the catalytic properties of E1p (Hengeveld et al., 1999). Therefore, indirect effects of surface mutations on catalysis have to be taken into account, and the effect of the mutation H44R may be analogous. Residue R263 is analogous to Y283 in *P. putida* E1b, again a surface residue, now closer to the N-terminal tail of the other α-subunit. The mutation R263G is thus possibly involved in disrupting the α-α subunit interactions. Since ThDP binds on the α-β interface it is very likely that any disruption of subunit interactions can be (partially) compensated by extreme elevation of the ThDP concentration and thus it is very likely that ThDP has some effect on all mutations that destabilise α-α or α-β interactions. The mutation V389fs results in the elongation of the C-terminus of the α-subunit, disrupting the α-β interactions. Such a mutation will also be sensitive to ThDP-treatment.

Other mutations that are found on either the α-α or α-β interface are F205L, M210V, W214R, Y243N, Y369Q, R378C, R378H and S388X. All of these mutations involve quite drastic changes and will for this reason interfere with the tetramerisation of E1p.

Finally a few mutations (R72C, R127W, D315N, P316L and L329R) are found on the surface of the α-subunit of E1p, but not involved in subunit interactions. These exposed residues are clearly important for the functioning of E1p considering their effect on catalytic activity or stability. They could possibly be involved in binding to the other components of the complex (binding to E2 or interaction with the lipoyl domain). On the other hand, the mutations could simply be so drastic that the structural integrity of the enzyme is affected. The latter is most likely the case for the mutation R127W, where only trace amounts of E1pα were found using immuno-staining of fibroblasts (Table 1). The mutation D315N, on the other hand, resulted in normal amounts of E1pα indicating its importance in catalysis.

Mutations in E1b causing MSUD

The crystal structure of human E1b has revealed that it is very alike *P. putida* E1b. The main difference between the two structures is the presence of two structural potassium (K^+) ions in human E1b that were not found in *P. putida* E1b. The first K^+-ion assist a loop close to ThDP to adopt its proper conformation, while the second K^+-ion is found in the β-subunit at the α-β interface and is probably necessary for proper tetramerisation. Ævarsson et al. (Ævarsson et al., 2000) have discussed recently in detail all known mutations in E1b that cause MSUD. The mutations can be divided into three groups based on the effect that they have on the enzyme (Table 1). The first group contains mutations that affect ThDP-binding, mutations that were discussed above as active site mutants. The second group comprises mutations affecting the hydrophobic core or structural integrity of the enzyme. The third (largest) group contains mutants that affect the subunit association or protein-protein interactions. In the case of MSUD, unlike what has been found for mutations in E1p, the effect of the mutations on structure and catalysis correlates very well with the severity of the disease.

Structure and dysfunctioning of E2

The E2-component has a multidomain structure in which the independent folding units of lipoyl domains, an E1 and /or E3 binding domain and a catalytic domain are separated by flexible linker sequences of about 20-50 amino acids (figure 6). The interdomain fragments appear to be extended and highly flexible, as judged from limited proteolysis and NMR experiments (Perham & Roberts, 1981; Texter et al., 1988). The high flexibility is essential for the mechanism of active site coupling, by which the substrate is transferred between the active sites of E1, E2 and E3 (figure 1). Within this common architecture the number of lipoyl domains differs between the different complexes. Human PDHC has two lipoyl domains, human BCDC and OGDC both have only one lipoyl domain. Several structures of lipoyl domains have been solved by NMR (Dardel et al., 1993; Green, et al., 1995; Berg et al., 1996b; Berg et al., 1997; Ricaud et al., 1996), all showing a very similar overall fold. Residues in an exposed loop close to the lipoylated lysine residue might be involved in the recognition of E1. Mutagenesis and NMR studies have shown however that also other parts of the lipoyl domain are involved in the interactions with E1. The role of the lipoyl domains in

multienzyme catalysis was extensively reviewed recently in (Berg & De Kok, 1997). Besides its role in catalysis, the inner lipoyl domain of PDHC is also involved in binding the regulatory components of the complex, the E1-kinase and E1-phosphatase.

The E1 and/or E3 binding domain is a very small domain of only about 35 amino acids, consisting of two almost parallel α-helices. Structures of bacterial PDHC and OGDC E1/E3 binding domain have been solved by NMR (Kalia et al., 1993; Robien et al., 1992; Mande et al., 1996). The structure of a complex between a lipoyl/binding di-domain and the E3-component from *Bacillus stearothermophilus* has been solved by X-ray crystallography (Cruts et al., 1995).

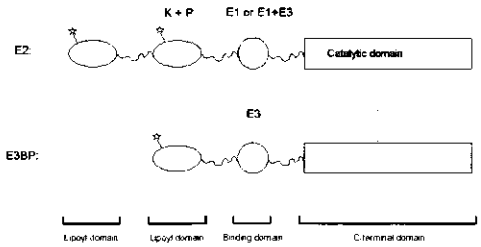


Figure 6. Domain structure of E2 and E3BP and the interactions of the domains with the other components of the mammalian 2-oxoacid dehydrogenase complexes. The separately folding domains are connected by flexible linker sequences shown in wavy lines. Dotted lines indicate interactions between different components. K means E1-kinase, P is E1-phosphatase. The lipoic acid covalently attached to the lipyl domain is indicated by an asterisk. The C-terminal domain of E2 and E3BP do not show sequence identity, therefore the C-terminal domain of E2 is indicated as the catalytic domain and the domain of E3BP as the C-terminal domain. E2 contains either one or two lipoyl domains depending on the source (E2p 2 lipoyl domains, E2o and E2b one lipoyl domain).

The site of interaction between E3 and the binding domain is so close to the E3 twofold symmetry axis that it is impossible for two molecules of the binding domain to bind simultaneously to one E3 dimer. The interactions between E3 and the binding domain are dominated by an electrostatic zipper formed by basic residues of the binding domain and acidic residues of one of the subunits of E3. The binding domain behaves like a Janus-face protein: while E3 interacts with the N-terminal helix, mutagenesis experiments in *Azotobacter vinelandii* E2p indicate that E1 interacts with the C-terminal part of the binding domain (Schulze et al., 1991a). Considerable homology between binding domain sequences of many organisms indicates a similar structure and binding mode in complexes from other sources.

Surprisingly E2o of higher organisms seems to lack a binding domain (Rice et al., 1992). In mammalian OGDC sequences related to the E3-BP are integrated into the N-terminal region of the E1-component. This region is responsible for its interaction with the E2 core and for directing E3 binding to the complex. Thus uniquely to the mammalian OGDC, E3 interacts via the E1 component and directly via the binding domain in E2 (Rice et al., 1992).

The structural core of all 2-oxoacid dehydrogenase complexes is formed by aggregation of the C-terminal catalytic domain of E2. Trimeric E2 units form a cubic or dodecahedral core. Crystal structures are now available for both the 24-meric cubic structure and the 60-meric dodecahedral structure (Mattevi et al., 1992c; Knapp et al., 1998; Izard et al., 1999). Both the cube and the dodecahedron have a hollow interior and contain large pores at their faces. While the intratrimer interactions are extensive and strong, the intertrimer interactions are much weaker.

The active site of E2 is located at the interface of two E2-subunits in a trimer, forming a 3 nm long channel with two entrances. Lipoamide enters this channel from the outside while CoA enters from the inside of the core. A high sequence and structural similarity between trimers of E2 and chloramphenicol acetyltransferase (CAT) was revealed by the determination of the crystal structure of E2. From sequence alignments of 13 known E2-sequences (Russell & Guest, 1991b) and the similarity to CAT (Guest, 1987) two highly conserved residues, a histidine and an aspartic acid, were predicted to be involved in the catalytic reaction of E2. Mutagenesis studies and the crystal structures of E2 have confirmed the proposed mechanism (Mattevi, A. et al., 1993c). The involvement of a third conserved residue, a serine or threonine, was revealed by the crystal structure of *A. vinelandii* E2p and has been confirmed by mutagenesis studies on E2 from several organisms (Mattevi et al., 1993c; Hendle et al., 1995). The histidine acts as a general base in the proposed mechanism, while the serine acts as a hydrogen bond donor to the putative negatively charged transition state. The acyl transfer is completed by rearrangements resulting in the formation of acylated CoA and dihydrolipoamide, which then dissociate from the active site.

Even though the amino acid sequences of various E2's are highly homologous, PDHC OGDC and BCDC each have their own substrate specificity. Mutagenesis studies by (Meng & Chuang, 1994) suggest that substrate specificity is sensitive to alterations in the size of residue 348 of human BCDC. In BCDC this residue is either an alanine or serine, in PDHC a phenylalanine and in OGDC it is either a glycine or serine. The inability to convert BCDC

specificity to PDHC specificity indicates however, that also other elements are important for creating the optimal environment within the active site of E2.

Defects of the E2 component of the 2-oxoacid dehydrogenase complexes have been reported for BCDC and OGDC. Up to now no defects were found in the E2 component of PDHC.

Even though no molecular evidence was found up to now to link the OGDC with inborn errors several reports have been published that show a correlation between the activity of this complex and for example Alzheimer's disease (AD) (Kish, 1997; Cruts et al., 1995), Parkinson's disease (Mizuno et al., 1994), Wernicke-Korsakoff (Butterworth et al., 1993) or

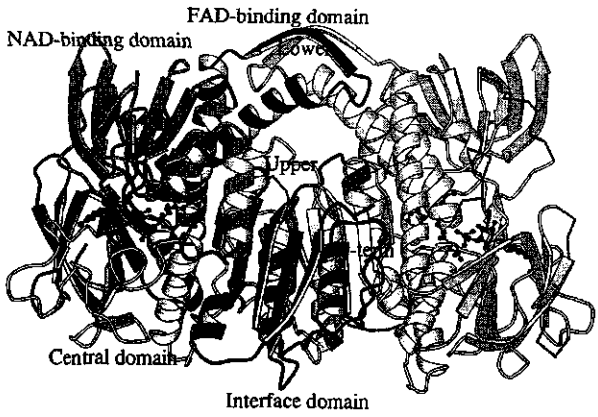


Figure 7 Ribbon representation of the structure of *P. fluorescens* lipoamide dehydrogenase (E3). Subunit one is indicated in dark grey shading, the other in light grey. The NAD-binding domain, the FAD-binding domain and the central and interface domain are shown. FAD is drawn in ball-and-stick representation.

Freidrich's ataxia (Mastrogiacomo et al., 1996). There is extensive evidence of mitochondrial abnormalities in AD, specifically of the activity and levels of OGDC (Gibson et al., 1998). Moreover, polymorphisms of the OGDC gene have been associated with AD (Sheu et al., 1999). On the other hand, the extracellular deposition in the brain of amyloid beta peptide (a characteristic of AD) has recently been suggested to have a direct toxic effect via activation of transcription factors. In particular a strong upregulation of gadd45 was found, indicating DNA damage. (Santiard-Baron et al., 1999). The polymorphisms in the OGDC gene might therefore be a result of this neurotoxic effect of amyloid beta peptide and thus be a secondary effect.

Several mutations in the E2b gene of MSUD patients have been described including insertions, substitutions and deletions in both intronic and exonic sequences. Most mutations described were found in introns, resulting in splicing problems. A number of these mutations involve rare deletions of intronic segments leading to secondary insertions/deletions in the

transcript through utilisation of cryptic or new splice sites (Chuang et al., 1997). Other mutations include small deletions in exon 2 and 3 (Chuang et al., 1991) or intronic mutations resulting in the entire skipping of exons (Fisher et al., 1991; Mitsubuchi et al., 1991; Herring et al., 1992). Point mutations were up to now found in exon 6,7, and 8. The first point mutation resulted in the substitution of E163 for a stop codon, yielding a truncated protein that only includes the lipoyl and binding domains (Fisher et al., 1993). The other two mutations are F215C and R230G (Chuang et al., 1997). F215 is conserved in all E2 as a hydrophobic residue, in contact with the active site histidine. Replacing it for a cysteine will certainly affect the catalytic activity of E2. R230 is completely conserved in all E2 sequences. It is also present in the active site, in contact with CoA. Replacing the positively charged arginine for the small and uncharged glycine will affect the binding of CoA and thus the activity of E2.

Structure and dysfunctioning of E3

Lipoamide dehydrogenase (E3) belongs to a family of FAD-dependent disulphide oxidoreductases. 3-D structures of E3 from multiple sources have been solved by X-ray crystallography (figure 7) (Toyoda et al., 1998; De la Sierra et al., 1997; Mattevi et al., 1991; Mattevi et al., 1992b; Mattevi, 1993b.). Jentoft et al. (Jentoft et al., 1992) have constructed a model structure for human E3 Since recent review articles describe the enzyme in depth (Williams, 1992; De Kok & Van Berkel, 1996) we will only give a short description of the enzyme. E3 is usually shared by all of the 2-oxoacid dehydrogenase complexes, although exceptions were found where different E3's were found for the different complexes (Palmer et al., 1991; Claiborn et al., 1994). E3 is a dimer of identical subunits, each comprising of three domains: the FAD-binding domain, the NAD-binding domain and the interface domain. The catalytic site is located at the subunit interface in an 11 Å deep cleft (figure (8). The enzyme catalyses the oxidation of the dihydrolipoamide group attached to one of the lipoyl domains of E2. In the first step of this ping-pong mechanism the electrons are transferred from the reduced lipoyl group to a reactive disulphide group. Secondly, the electrons are transferred via the FAD to NAD^+, the final electron acceptor. All important active site residues, as indicated by structural comparison with other disulphide reductases and mutagenesis are conserved within all E3's (De Kok et al., 1996).

E3-deficiency presents itself in two clinical forms, either in infancy with severe mental retardation and lactic acidemia or later in life with episodes of liver failure and myoglobinuria.

Seven mutations were identified to have the first phenotype (Table 1) and recently one new mutation was associated with the second phenotype (G229C). As all of these mutations have been discussed extensively we will only give a summary of the proposed functions. The mutations K72E and P488L are always found combined in one patient. K72 is an exposed residue that likely has negligible effect. P488 is in the cis conformation and positions the backbone containing the his-glu diad of the active site. G136D and E375K are also found together, G136 is part of an α-helix in the FAD domain and deletion of this residue affects FAD binding. E375 is part of the central domain of E3. Mutation of this residue results in drastic reduction of the enzyme activity. Mutating either G229 to a C or D479 to a V affects protein stability and possibly the dimerisation of E3. R495 is conserved in all E3 as a bulky hydrophilic residue, mutating it to a glycine results in disruption of dimer formation. Finally one mutation of E3 results in the replacement of the last amino acid of the leader peptide (Y35) into a stop codon and a frameshift. This mutation results in complete loss of E3.

Function and dysfunctioning of E3BP

So far we have considered the three major catalytic components of the complexes. A fourth component has been identified in mammalian PDHC's: the E3BP (De Marcucci & Lindsay, 1985). The E3BP has a similar domain structure to that of E2 (figure 6): a lipoyl domain, an E3 binding domain and a C-terminal domain. The C-terminal domain has no acetyltransferase activity or sequence similarity to any other known protein (Reed & Hackert, 1990). A few copies of E3BP associate with the catalytic core of E2p, facilitating the proper binding of the E3-component (Sanderson et al., 1996). In Yeast PDHC 12 subunits of the E3BP bind to E2, probably on the inside of the dodecahedral E2-core structure (Stoops et al., 1997). It has been shown that the lipoyl domain of the E3BP can participate in the reaction mechanism (Yang et al., 1997).

Five cases of E3BP deficiency have been reported so far (Aral et al., 1997; De Meirleir et al., 1998a; Geoffroy et al., 1996; Ling et al., 1998; Marsac et al., 1993). All cases show neonatal lactic acidemia. The deficiencies were demonstrated by Western blot analysis, in which the E3BP protein was either altered or completely absent. Analysis on a molecular level showed that only in one case a small deletion was causing the deficiency. A four bp deletion was found in the leader peptide sequence resulting in a premature stop codon (Ling et al., 1998). In two other cases large deletions lead to truncated E3BP of only 25 or 321 amino acids, instead

of the 501 amino acid long wild type protein (Aral et al., 1997). It is clear from these results that the absence of E3BP leads to heavily impaired PDHC.

Assembly of the complexes

So far we have discussed the different components of the 2-oxoacid dehydrogenase complexes and the effect of mutations on these components. The assembly of the components into the functional multi-enzyme complexes is of vital importance for the complexes. All components of the mitochondrial complexes are encoded by nuclear genes. The cytosolic precursors of the components have to be actively imported into the mitochondrion. For BCDC the relative rates of entry were shown to be E1bα > E2 > E1bβ. This makes the import rate of E1bβ the limiting factor in the amount of complex formed (Sittler et al., 1998). Point mutations in the leader sequence (or mitochondrial targeting sequence) have been found in E1pα and E3, both strongly affecting the activity of the activity of the entire complex. This shows not only the importance of these sequences but also the effect the partial or total absence of one component of the complex has on the total activity.

The mature form of E2 is formed in the mitochondrial matrix through lipoylation of the lipoyl domain by lipoyltransferases (Motokawa, 1996). The two lipoyltransferases found in bovine liver were able to lipoylate lipoyl domains from several origins (rat E2p and E2o and bovine E2b). In the mitochondrial lipoylation system a third enzyme is required for lipoic acid activation.

Chuang et al. (Chuang et al., 1999) have shown in *in vitro* experiments that E1b requires molecular chaperones (GroEL/GroEs) for the reconstitution of the α2β2 tetramers. The reconstitution proceeds via the formation of an αβ intermediate.

The assembly of the different 2-oxoacid dehydrogenase complexes can result in aggregates ranging in size from only 700 kDa (the smallest PDHC as found in *A. vinelandii*) up to 10 MDa in mammalian PDHC. Mammalian complexes all contain the three catalytic components (E1, E2 and E3), both PDHC and BCDC contain two regulatory enzymes (E1 kinase and E1 phosphatase) and PDHC also contains the E3BP. The association of the 6 components of PDHC, with the emphasis on the association of the regulatory enzymes to E2 has been reviewed by Roche et al. (Roche al., 1996) (figure 6). Typical bovine PDHC contains about 20-30 E1 tetramers bound to the dodecahedral assembly of E2, and only about one kinase molecule. Both E1-kinase and E1-phosphatase bind to the inner lipoyl domain of E2. Hand

over hand movement, from one lipoyl domain to another, of the kinase or phosphatase is necessary to reach all bound E1's.

From the number of mutations found on the E1 and E3 components that affect the subunit interactions it is clear that disturbance of the assembly of the components of the complex has a drastic effect on the functioning of the complex. It would therefore not be surprising if some of the mutations found on the surface of E1p, to which at this moment no definite function could be assigned, are involved in the association of this component to the other components of the complex. Since the assembly of the complex and the interactions between the different

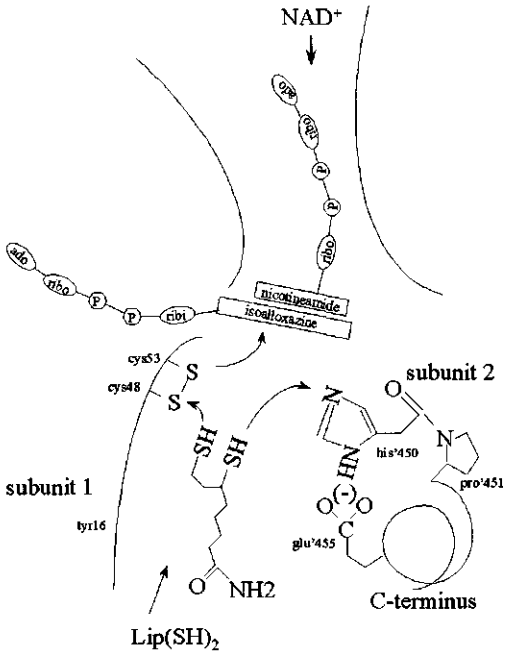


Figure 8 Schematic drawing of the active site and reaction mechanism of E3. For an *explanation see text.*

components of the complex is still quite unclear, it would be very useful to investigate the function of the residues in more detail in *in vitro* experiments.

Concluding remarks

Dysfunctioning of any of the components of the 2-oxoacid dehydrogenase complexes leads to severe disorders in humans. In this paper we have tried to relate mutations in these components, that were found by molecular analysis of patients, to a more definite function of the affected residues. Especially in the case of PDHC deficiency, the need for uniform kinetic data has become obvious from our inventarisation. Because of, for example, the variable expression pattern of the X-chromosome linked E1pα kinetic data from tissues result in large variations in the measured activities. The use of a uniform model system using site directed mutants for analysis of mutations identified in patients would certainly improve the understanding of the structure-function relationship of these complicated systems. Such an analysis can help to improve our understanding of the mechanism of deficiency and might lead to new or more effective treatments.

Table 1. Mutations in human E1, E3 and E3BP that cause the dysfunctioning of the 2-oxoacid dehydrogenase complexes.

Mutat. [1]	Sex [2]	Act. [3]	Blot [4]	Km ThDP	Km pyr	Location	Comments	Reference
Point mutations in human E1pα								
R10P	♂	28	R α			Mitochondrial targeting peptide	Defective import E1α	Takakubo et al., 1995
H44R	♂	0-17	N	↑		Protein interaction?	Defective activation by phosphorylation, ThDP responsive	Naito et al., 1994
R72C	4♂	20-30				Protein interaction		Lissens et al., 2000; Dahl et al., 1992; Lissens et al., 1996
R88S	♂	18	R αβ			Active site	ThDP responsive	Lissens et al., 1997
G89S	♀	100		↑		Active site	DCA + ThDP responsive	Matsuda et al., 1994; Naito et al., 1999
H113D	♀	12				?		Lissens et al., 1996
R127W	♀,3♂	59,7-15	T			Protein interaction		Lissens et al., 2000; Fujii et al., 1994
R141Q	♂	32				Protein interaction		Lissens et al., 200; Morten et al., 1999
G162R	♀	40				Active site		Lissens et al., 1996
N164S	3♂	26				Active site		Lissens et al., 2000

Mutat. [1]	Sex [2]	Act. [3]	Blot [4]	Km ThDP	Km pyr	Location	Comments	Reference
V167M	2♀	1-7				Active site		Chun et al., 1993
A175P	♀	51	N			Subunit interaction		Takakubo et al., 1993b
A199T	♂	17				Active site		Chun et al., 1993
F205L	2♂	25-36	R			Subunit interaction		Chun et al., 1995; Dahl et al., 1994
M210V	♂	6		↑	↑	Subunit interaction	Kcat 38%	Tripatara et al., 1999; Triptara et al., 1996
W214R	♂	47				Subunit interaction		Lissens et al., 2000
P217L	♂	12		=	=	Active site, structural integrity	Kcat 24%, Structural integrity	Triptara et al., 1999; Triptara et al., 1996
T231A	♂	2				Active site, phosphorylation		Chun et al., 1993
Y243N	♂	32	R α			Subunit interaction	Unstable E1	Matthews et al., 1994
D258A	♂	10	R			Protein interaction		Matthews et al., 1993b
R263G	♀,12♂	16-168	R αβ			Subunit interaction	ThDP responsive, unstable E1	Lissens et al., 2000; Chun et al., 1995; Lissens et al., 1996; Marsac et al., 1997; Chun et al., 1993; Wexler et al., 1992; Wexler et al., 1997; Naito et al., 1997; Briones et al., 1996
R263Q	2♂	38	AL			Subunit interaction		Awata et al., 1994
M282L	♀	115				?		Matthews et al., 1994
R288H	♀	25				Protein interaction		Lissens et al., 2000
G291R	♀	100				Active site		Matsuda et al., 1994
H292L	♀	2				Active site		Chun et al., 1993
R302L	♀					Phosphorylation		Lissens et al., 2000
R302H	♀	40				Phosphorylation		Otero et al., 1998
R302C	8♀	5-89	N/Rαβ			Phosphorylation		Lissens et al., 2000; Dahl et al., 1992; Lissens et al., 1996; Morten et al., 1999; Otero et al., 1998; Morten et al., 1998; Fujii et al., 1996; De Meirleir et al., 1993

Mutat. [1]	Sex [2]	Act. [3]	Blot [4]	Km ThDP	Km pyr	Location	Comments	Reference
D315N	♂	26-40	N			Protein interaction		Matthews et al., 1994
P316L	♀	53				Protein interaction		Takakubo et al., 1995
L329R	♂	50				Protein interaction		Lissens et al., 2000
Y369Q	♂	30				Subunit interaction		Lissens et al., 2000
R378C	♀		Rαβ			Subunit interaction		Lissens et al., 2000
R378H	♀, 7♂	13-50	N/Rαβ			Subunit interaction	PDHC-act. 1.2%	Lissens et al., 2000; ripatara et al., 1999; Chun et al., 1995; Morten et al., 1999; Matthews et al., 1994; Wexler et al., 1997; Hansen et al., 1991; Wexler et al., 1995
S388X	♂	15				Subunit interaction		Linssen et al., 2000
Insertions and fusions in human E1pα								
V171Δ31	♀, 2♂	4/40	A					Dahl et al., 1992; Chun et al., 1995
Y243Δ	♀					Subunit interaction		Lissens et al., 2000
R288fs	♀	88					Truncated at 283 aa	Matthews et al., 1994
S300fs	♀	16	A				Truncated at 276 aa, unstable E1	Matthews et al., 1993a
R302fs	♀	5	A				Truncated at 297 aa	Hansen et al., 1993
E306ins7	♂	10	N					De Meirleir et al., 1992
I307ins11	♀	15						Hansen et al., 1994
R311Δ	2♂	11/0	N			Protein interaction		Chun et al., 1995,Triptara et al., 1996
S312fs	7♀	8-31	N/Rαβ				Truncated at 293 aa	Lissens et al., 2000; Green et al., 1995; Chun et al., 1993; Fujii et al., 1996; Dahl et al., 1990; Lissens et al., 1995
K313Δ	♀	40	AL			Protein interaction		Hansen et al., 1991
P316fs	♀	1	SR				Truncated at 308 aa	Chun et al., 1993
321ins5	♂	15						Lissens et al., 1996
N326ins5	♂	18	Rαβ					Fujii et al., 1996
S327fs	♀	61	SR				Truncated at 313 aa	Chun et al., 1995

Mutat. [1]	Sex [2]	Act. [3]	Blot [4]	Km ThDP	Km pyr	Location	Comments	Reference
L335ins12	♂							De Meirleir et al., 1998
E358fs	♀	26	S				Truncated at 339 aa	Chun et al., 1991
P360fs	♀	100	S				Truncated at 332 aa	Chun et al., 1995
E362+6	♀	36						Ito et al., 1995
G365+3	♂		Rαβ					Fujii et al., 1996
E376fs	♀	20					Elongated to 394 aa	Chun et al., 1993
Q382fs	6♀	6-46	Rα				Truncated at 358 aa	Lissens et al., 2000; Chun et al., 1993; Endo et al., 1991; Ito et al., 1992; Takakubo et al., 1993a
K387fs	♂	41	R				Elongated to 362 aa	Tripatara et al., 1996
K387fs	♂	25	T				Elongated to 400 aa	Hansen et al., 1991
K387fs	♂		Tα,Rβ				Elongated to 431 aa	Morten et al., 1999
S388fs	♀, 4♂	8-14	Rαβ			Structural integrity	Truncated at 358 aa	Lissens et al., 2000; Naito et al., 1994; Chun et al., 1995
S388ins2	♂		Rαβ					Marsac et al., 1997; Fujii et al., 1996
V389fs	♂	25	R				ThDP responsive, elongated to 392 aa	Endo et al., 1989
Point mutations in human E1bα								
M109T		6				Hydrophobic core	Intermediate	Ævarssonet al., 2000
R159W						ThDP-binding	Classic	Nobukino et al., 1993
Q190K		1.4				Subunit association	Intermediate	Nobukino et al., 1993
T211M		0				ThDP-binding	Classic	Ævarsson et al., 2000
G249S		0				Hydrophobic core	Classic	Wynn et al., 1998
A253T		0				Hydrophobic core	Classic	Nobukino et al., 1993
A254D		27				Subunit association	Intermediate	Ævarsson et al., 2000
R265W		0				ThDP-binding	Classic	Wynn et al., 1998
N267S		1.3				ThDP-binding	Classic	Chuang et al., 1995
A285P		0				Subunit association	Classic	Wynn et al., 1998
G290R		4				Subunit association	Intermediate	Chuang et al., 1990
R297H		27				Subunit association	Intermediate	Ævarsson et al., 2000
T310R		0				Hydrophobic core	Classic	Wynn et al., 1998

Mutat. [1]	Sex [2]	Act. [3]	Blot [4]	Km ThDP	Km pyr	Location	Comments	Reference
I326T		1.4				Hydrophobic core	Intermediate	Nobukino et al., 1993
F409C		0				Subunit association	Classic	Chuang et al., 1995
Y413C		0				Subunit association	Intermediate	Chuang et al., 1990
Y438N		0				Subunit association	Classic	Matsuda et al., 1990; Fisher et al., 1991
Point mutations in human E1bβ								
N176Y		<1				Subunit association	Classic	Ævarsson et al., 2000
H206R		0				Subunit association	Classic	Nobukino et al.,
Point mutations in human E3								
Y35X	♂	8-20	20%			Leader peptide		Hong et al., 1996; Elpeleg et al., 1997; Shaag et al., 1999
K72E		6	R			Exposed, solvent		Liu et al., 1993; Liu et al., 1995; Lanterman et al., 1996
G136Δ	♂	3	N			FAD binding		Hong et al., 1997
G229C	♀, ♂	9-21	25-60%			Protein stability, dimerisation		Shaag et al., 1999
E375K	♂	3	N			?		Hong et al., 1997
D479V	♀	15	8%			Stability, dimer interface		Shany et al., 1999
P488L		6	R			Positioning active site		Liu et al., 1993; Liu et al., 1995; Lanterman et al., 1996
R495G	♀	14	53%			Dimer formation		Hong et al., 1996

[1] Mutat. means mutation. Amino acid numberings include leader peptide sequences.

[2] ♀ is female, ♂ is male

[3] Activity was measured in fibroblast.

[4] Blot means Western blot. R is reduced, N is normal, T is trace, AL is altered, A is absent, SR is slightly reduced, S is smaller molecular weight.

Acknowledgements: This work was supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

8 Summary and concluding remarks

Pyruvate dehydrogenase multi-enzyme complex (PDHC) is member of a family of multi-enzyme complexes that catalyse the irreversible decarboxylation of various 2-oxoacid substrates to their corresponding acyl-CoA derivatives, NADH and CO_2. 2-oxoacid dehydrogenase complexes hold key points in primary energy metabolism: PDHC in glycolysis, the oxoglutarate dehydrogenase complex (OGDC) in the citric acid cycle, the branched-chain oxoacid dehydrogenase complex (BCDC) in amino acid catabolism, the glycine dehydrogenase complex (GDHC) in the glycine cleavage system and acetoin dehydrogenase complex (ADHC) in acetoin metabolism. Malfunctioning of a 2-oxo acid dehydrogenase complex in man leads to a broad variety of clinical manifestations, e.g. lactic acidosis (PDHC-deficiency) or maple syrup urine disease (BCDC-deficiency).

PDHC consists of multiple copies of three enzymes: 2-oxoacid dehydrogenase or E1p-component, dihydrolipoyl acyltransferase (E2p) and dihydrolipoamide dehydrogenase (E3). The E2 component plays a central role in the complex, both catalytically and structurally. It consists of multiple domains (from N- to C-terminus, 1-3 lipoyl domains, an E1/E3-binding domain and a catalytic domain) connected by flexible linkers, allowing a high degree of flexibility of the individual domains. In gram-negative bacteria, the C-terminal catalytic domain forms a cubic core of 24 subunits. Multiple copies of the peripheral components E1p and E3 are non-covalently attached to this core. Three-dimensional structures of the different components and domains of PDHC have been solved by X-ray crystallography or NMR.

E1 catalyses the rate limiting step in the overall complex reaction; the reductive acylation of the lipoyl groups of E2 utilising thiamin diphosphate (ThDP). The E1-component exists either as a homodimer (α_2) or heterotetramer ($\alpha_2\beta_2$), depending on the type and source of the complex. There is strikingly little sequence similarity between homodimeric and heterotetrameric E1 or even between homodimeric E1o and homodimeric E1p. An exception to this is the so-called "ThDP-binding motif", that has so far been found in all known ThDP dependent enzymes. Until quite recently E1 was by far the least characterised component of the complex. This lack of (structural) information was largely due to the lack of a stable form

of the enzyme: when resolved from the complex, E1 is unstable. To obtain a better insight into the structure and functioning of the E1-component of the pyruvate dehydrogenase multi-enzyme complex from *A. vinelandii* and its interactions within the complex, a PhD project was started in 1995. The results of this project are presented in this thesis.

Chapter 2 describes the cloning and sequencing of the gene encoding E1p from *Azotobacter vinelandii* and, for the first time, the expression and purification of a homodimeric E1. The cloned E1p shows a kinetic behaviour exactly as the wild-type complex-bound enzyme with respect to its substrate (pyruvate), its allosteric properties and its effectors. Moreover, it can be used to fully reconstitute complex activity. The enzyme is stable in high ionic strength buffers, but is irreversibly inactivated when incubated at high pH, which presumably is caused by its inability to redimerise correctly. This explains the previously found low stability of the wild-type E1p component after resolution from the complex at high pH. Kinetic experiments show that acetyl coenzyme A acts as a feedback inhibitor by binding to the E1p component.
Limited proteolysis experiments show that the N-terminal region of E1p is easily removed. The resulting protein fragment is still active with artificial electron acceptors but has lost its ability to bind to the core component (E2p) and thus reconstitute complex activity, indicating that the N-terminal region of E1p is involved in the binding of E1p to E2p.
To further investigate this hypothesis N-terminal deletion mutants of E1p were constructed and characterised (Chapter 3).
Up to 9 N-terminal amino acids could be removed from E1p without affecting the properties of the enzyme. Truncation of up to 48 amino acids did not effect the expression or folding abilities of the enzyme, but these truncated enzymes could no longer interact with E2p, confirming the involvement of the N-terminal region in the binding of E1p to E2p. The 48 amino acids deletion mutant (E1pΔ48) was characterised in detail: it is catalytically fully functional, it has a V_{max} identical to that of wild type E1p, it can reductively acetylate the lipoamide group attached to the lipoyl domain of the core enzyme (E2p) and it forms a dimeric molecule. On the other hand, the $S_{0.5}$ for pyruvate is significantly decreased.
A heterodimer was constructed containing one subunit of wild type E1p and one subunit of E1pΔ48. As the heterodimer is not able to bind to E2p therefore both N-terminal domains are needed for the binding of E1p to E2p.
Based on sequence alignments and mutagenesis experiments on E2p the interactions between the N-terminal region of E1p and E2p are thought to be mainly of electrostatic nature

involving negatively charged residues on the N-terminal domains of E1p and previously identified positively charged residues on both the binding and catalytic domain of E2p. These mutagenesis experiments show that E1p, unlike E3, interacts with both the binding- and the catalytic-domain of E2p.

In chapter 4 we use a synthetic peptide (Nterm-E1p) to characterise the structure and function of the N-terminal region (amino acid residues 4-45) of E1p. Activity and binding studies establish that Nterm-E1p specifically competes with E1p for binding to the E2p component. Moreover, the experiments show that the N-terminal region of E1p forms an independent folding domain that functions as a binding domain. CD measurements, 2D-^{1}H NMR analysis and secondary structure prediction all indicate that Nterm-E1p has a high α-helical content.

Chapter 5 describes the construction of point mutations in the N-terminal domain (D17Q, D17R, E20Q, E20R, D24Q and D24R) to identify the specific residues involved in the interactions between E1p and E2p. Kinetic and binding studies show that D17 is essential for the binding of E1p to E2p. D24 is involved in the binding, but not essential, whereas E20 is not involved. None of the mutations affects the folding or dimerisation of E1p

We have refined the model of the binding mode of E1p and E2p proposed in chapter 4: in the "helix-turn-helix" motif residues D17, E20 and D24 are on one side of the helix, whereas on the opposite side W15, L19 and V22 form a hydrophobic "patch". Helix II also has a hydrophobic side, which can form a hydrophobic "pocket" with the residues of helix I. None of the mutations result in complete abolishment of the binding capacity of E1p to E2p; neither do point mutations in E2p. Apparently the binding interface is composed of more than one region both on E2p (binding domain and catalytic domain) and on E1p.

Finally, it was shown before that the binding of E1p to E2p is highly species specific: *E. coli* E1p binds very poorly to *A. vinelandii* E2p. The fact that the mutations D17Q and D24R both have such a strong effect on the binding abilities of E1p correlates well with this observation.

Spectroscopic measurements in chapter 4 show that the peptide is present in two conformations, the population of which depends on the sample conditions. The two conformations are designated "unfolded" at a pH above 6 and "folded" at a pH below 5. The 2D-^{1}H TOCSY spectrum of a mixture of folded and unfolded Nterm-E1p shows exchange cross peaks that "link" the folded and unfolded state of Nterm-E1p. We suggest that the exchange rate between the two species is in the range of 0.5 to 5 seconds^{-1}. Sharp resonances in the NMR-spectra of wild type E1p demonstrate that this 200 kDa enzyme contains highly flexible regions that partly originate from its N-terminal region. The observed dynamic

character of E1p and of Nterm-E1p is likely required for the binding of the E1p dimer to the two different binding sites on E2p. Moreover, the flexibility might be essential to sustain the allosteric properties of the enzyme when bound in the complex.
In summary, chapter 2-5 describe in detail the elucidation of the binding mode of E1p to E2p.

Only in the final stages of this PhD project detailed structural information came available for E1. Therefore, in chapter 6 we combined primary and secondary structural information to obtain a sequence alignment of homodimeric and heterotetrameric 2-oxoacid dehydrogenases and transketolase. In striking contradiction with their structural similarity is the apparent lack of sequence similarity between these enzymes. The alignment consists of two parts, corresponding to the α- and β-strand of heterotetrameric 2-oxoacid dehydrogenase. Aligning the sequences of homodimeric pyruvate dehydrogenase (E1p) (*Azotobacter vinelandii* and *Escherichia coli*), transketolase (TK) (*Saccharomyces cerevisiae*), branched-chain oxoacid dehydrogenase (E1b) (*Pseudomonas putida* and human) and heterotetrameric E1p (*Bacillus stearothermophilus* and human), 36 residues are identical while 32 residues are highly conserved. These residues are evenly spread throughout the sequence, although multiple gaps and/or insertions had to be introduced. The most likely candidates for ThDP binding and active site residues were identified. Furthermore, a common function was suggested for regions shared by TK and the 2-oxoacid dehydrogenases. A few conserved regions remain that no definite function could be assigned to. It would be very valuable to study the function of these regions by for example mutagenesis experiments. In addition, puzzling differences between the binding mode of homodimeric E1 and heterotetrameric E1 are found and clearly need more detailed study. This new alignment opens possibilities to construct model structures for E1. Especially for the homodimeric E1p, that up to now has resisted crystallisation attempts, this will be of great value.
Finally, chapter 7 reviews the structural basis of the dysfunctioning of human 2-oxoacid dehydrogenase complexes. As was mentioned before, malfunctioning of any of these complexes leads to a broad variety of clinical manifestations. Deficiency of the PDHC predominantly leads to lactic acidosis combined with impairment of neurological function and/or delayed growth and development. Maple urine disease is an inborn metabolic error caused by dysfunction of the BCDC. An association between both Alzheimer disease and Parkinson's disease and the OGDC gene has been reported. In this chapter the wealth of both genetic and structural information available is combined to gain more insight into the

structural basis of the dysfunctioning of the 2-oxoacid dehydrogenase complexes. Known mutations in the components of human complexes were explained based on their putative function in the enzyme.

Even though this thesis sheds light onto the structure and binding mode of homodimeric E1p, many questions remain. For example, the differences between homodimeric and heterotetrameric E1; why does heterotetrameric E1 lack an N-terminal domain? Putative active site residues were identified based on sequence alignments and homology models, but their function has to be experimentally verified; what residues form the lipoamide binding channel; which residues bind ThDP or what residues bind pyruvate, either in a pre-catalytic complex or in a catalytic complex. Finally, our experiments show the presence of several highly flexible regions in E1p. Removal of these regions could be an important step in obtaining suitable crystals that will ultimately lead to a high-resolution structure of homodimeric E1.

Butterworth, R.F., Kril, J.J. & Harper, C.G. (1993), *Clin. Exp. Res. 17*, 1084-1088.
Cate, R.L., Roche, T.E. & Davis, L.C. (1980), *J. Biol. Chem. 255*, 7556-7562.
Chabrière, E., Charon, M.H., Volbeda, A., Pieulle, L., Hatchikian, E.C. & Fontecilla-Camps, J.C. (1999), *Nature Struct. Biol. 6*, 182-190.
Charon, M.M., Volbeda, A., Chabrière, E., Pieulle, L. & Fontecilla-Camps, J.C. (1999), *Curr. Op. struct. Biol. 9*, 663-669.
Chen, Y. & Barkley, M.D. (1998), *Biochemistry 37*, 9976-9982.
Chuang, D.T., Hu, C.C., Ku, L.S., Markovitz, P.J. & Cox, R.P. (1985), *J. Biol. Chem. 260*, 13779-13786.
Chuang, D.T., Fisher, C.W., Laue, K.S., Griffin, T.A., Wynn, R.M. & Cox, R.P. (1991), *Mol. Biol. Med. 8*, 49-63.
Chuang, D.T. & Shih, V.E. (1995), In: *The metabolic and molecular bases of inherited disease, 7th ed.* Scriver, S.R., Beaudet, A.L., Sly, W.S., Valle, D., Eds., McGraw-Hill Inc., pp. 1239-1277.
Chuang, J.L., Cox, R.P. & Chuang, D.T. (1990), *FEBS Lett. 262*, 305-309.
Chuang, J.L., Fisher, C.R., Cox, R.P. & Chuang, D.T. (1994), *Am. J. Hum. Genet, 55*, 297-304.
Chuang, J.L., Davie, J.R., Chinsky, J.M., Wynn, R.M., Cox, R.P. & Chuang, D.T. (1995), *J. Clin. Invest. 95*, 954-963.
Chuang, J.L., Cox, R.P. & Chuang, D.T. (1997), *J. Clin. Invest. 100*, 736-744.
Chuang, J.L., Wynn, R.M., Song, J.L. & Chuang, D.T. (1999), *J. Biol. Chem, 274*, 10395-10404.
Chun, K., MacKay, N., Petrova-Benedict, R. & Robinson, B.H. (1991), *Am. J. Hum. Genet, 49*, 414-420.
Chun, K., MacKay, N., Petrova-Benedict, R. & Robinson, B.H. (1993), *Hum. Mol. Genet, 2*, 449-454.
Chun, K., MacKay, N., Petrova-Benedict, R., Federico, A., Fois, A., Cole, D.E.C., Robertson, E. & Robinson, B.H. (1995), *Am. J. Hum. Genet. 56*, 558-569.
Claiborn, A., Ross, R.P., Ward, D., Parsonage, D. & Crane III Jr., E.J. (1994), In: *Flavins and flavoproteins*, Yagi, K., Ed., Walter de Gruyter, Berlin, pp. 587-596.
Cruts, M., Backhovens, H., Van Gassen, G., Theuns, J., Wang, S.Y., Wehnert, A., Van Duijn, C.M., Karlsson, T., Hofman, A., Adolfsson, R., Martin, J.J. & Van Broeckhoven, C. (1995), *Neurosci. Lett. 199*, 73-77.
Dahl, H.H.M., Maragos, C., Brown, R.M., Hansen, L.L. & Brown, G.K. (1990), *Am. J. Hum. Genet. 47*, 286-293.
Dahl, H.H.M., Brown, G.K., Brown, R.M., Hansen, L.L., Kerr, D.S., Wexler, I.D., Patel, M.S., De Meirleir, L., Lissens, W., Chun, K., MacKay, N. & Robinson, B.H. (1992), *Hum. Mutat. 1*, 97-102.
Dahl, H.H.M., Brown, G.K. (1994), *Hum. Mutat. 3*, 152-155.
Dahl, H.H.M. (1995), Am. J. *Hum. Genet. 56*, 553-557.
Danner, D.J. & McKean, M.C. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 271-293.
Dardel, F., Davis, A. L., Laue, E.D. & Perham, R.N. (1993), *J. Mol. Biol. 229*, 1037-48.
Davie, J.R., Wynn, R.M., Cox, R.P. & Chuang, D.T. (1992), *J. Biol. Chem. 267*, 16601-16606.
De Graaf-Hess, A.C. & De Kok, A. (1982), *FEBS Lett. 143*, 261-264.
De Kok, A. & Westphal, A.H. (1985), *Eur. J. Biochem. 152*, 35-42.
De Kok, A., Hengeveld, A.F., Martin, A. & Westphal, A.H. (1998), *Biochim. Biophys. Acta. 1385*, 353-66.
De Kok, A. & Van Berkel, (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 53-70.
De la Sierra, I.L., Pernot, L., Prangé, T., Saludjian, P., Schiltz, M., Fourme, R. & Padron, G.J. (1997), *Mol. Biol. 269*, 129-141.
De Marcucci, O. & Lindsay, J.G. (1985), *Eur. J. Biochem. 149*, 641-648.
De Meirleir, L., Lissens, W., Vamos, E. & Liebaers, I. (1992), *Hum. Genet. 88*, 649-652.
De Meirleir, L., Lissens, W., Denis, R., Wayenberg, J.L., Michotte, A., Brucher, J.M., Vamos, E., Gerlo, E. & Liebaers, (1993), *Pediat. Neurol. 9*, 216-220.
De Meirleir, L., Lissens, W., Benelli, C., Marsac, C., De Klerk, J., Scholte, J., Van Diggelen, O., Kleijer, W., Seneca, S. & Liebaers, I. J. (1998a), *Inher. Metab. Dis. 21*, 9-16.
De Meirleir, L., Specola, N., Seneca, S. & Lissens, W. (1998b), *J. Inher. Metab. Dis. 21*, 224-226.
Diefenbach, R.J., Candy, J.M., Mattick, J.S. & Duggleby, R.G. (1992), *FEBS Lett. 296*, 95-98.
Dobritzsch, D., Konig, S., Schneider, G. & Lu, G. (1998), *J. Biol. Chem. 273*, 20196-20204.
Domingo, G.J., Chauhan, H.J., Lessard, I.A.D., Fuller, C. & Perham, R.N. (1999), *Eur. J. Biochem. 266*, 1136-1146.
Douce, R., Bourguignon, J., Neuburger, M. & Beill, F.R. (2001), *TIPS 6*, 167-176.
Duckworth, H.W., Jaenicke, R., Perham, R.N., Wilkie, A.O.M., Finch, J.T. & Roberts, G.C.K. (1982), *Eur. J. Biochem. 124*, 63-69.
Dyda, F., Furey, W., Subramanyam, S., Sax, M., Farrenkopf, B. & Jordan, F. (1993), *Biochemistry 32*, 6165-6170.

Elpeleg, O.N., Shaag, A., Glustein, J.Z., Anikster, Y., Joseph, A. & Saada. A. (1997), *Hum. Mutat. 10*, 256-257.
Endo, H., Hasegawa, K., Narisawa, K., Tada, K., Kawaga, Y. & Ohta, S. (1989), *Am. J. Hum. Genet. 44*, 358-364.
Endo, H., Miyabayashi, S., Tada, K. & Narisawa, K. (1991), *J. Inherit. Metab. Dis. 14*, 793-799.
Eswaran, D., Ali, M.S., Shenoy, B.C., Korotchkina, L.G., Roche, T.E. & Patel, M.S. (1995), *Biochim. Biophys. Acta 1252*, 203-208.
Fang, R., Nixon, P.F. & Duggleby, R.G. (1998), *FEBS Lett. 437*, 273-277.
Fisher, C.R., Fisher, C.W., Chuang, D.T. & Cox, R.P (1991a), *Am. J. hum. Genet. 49*, 429-434.
Fisher, C.W., Laue, K.S., Fisher, C.R., Wynn, R.M., Cox, R.P. & Chuang, D.T. (1991b), *Biochem. Biophys. Res. Comm. 174*, 804-809.
Fisher, C.W., Fisher, C.R., Chuang, J.L., Laue, K.S., Chuang, D.T. & Cox, R.D. (1993), *Am. J. Hum. Genet. 52*, 414-424.
Flourney, D.S. & Frey, P.A. (1989), *Biochemistry 28*, 9594-9602.
Fujii, T., Van Coster, R.N., Old, S.E., Medori, R., Winter, S., Gubits, R.M., Matthews, P.M., Brown, R.M., Brown, G.K., Dahl, H.H.M. & De Vivo, D.C. (1994), *Ann. Neurol. 36*, 83-89.
Fujii, T., Garcia-Alvarez, M.B., Sheu, K.F., Kranz-Eble, P.J. & De Vivo, D.C. (1996), *Pediat. Neurol. 14*, 328-334.
Geoffroy, V., Fouque, F., Benelli, C., Poggi, F., Saudubray, J.M., Lissens, W., De Meirleir, L.D., Marsac, C., Lindsay, J.G. & Sanderson, S.J. (1996), *Pediatrics 97*, 267-272.
Gibson, T. J. (1984) *Ph. D. Thesis, University of Cambridge.*
Gibson, G.E., Sheu, K.F.R. & Blass, J.P. (1998), *J. Neorol. Transm. 105*, 855-870.
Goa, J. (1953), *Scand. J. Clin. Lab. Invest. 5*, 218-222.
Gong, X., Peng, T., Yakhnin, A., Zolkiewski, M., Quinn, J., Yeaman, S.J. & Roche, T.E. (2000), *J. Biol. Chem. 275*, 13645-13653.
Green J.B.A.. (1989), *FEBS Lett. 246*, 1-5.
Green, J.D., Laue, E.D., Perham, R.N., Ali, S.T. & Guest, J.R. (1995), *J. Mol. Biol. 248*, 328-43.
Griesinger C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988), *J. Am. Chem. Soc. 110*, 7870.
Guest, J.R. (1987), *FEMS Microbiol. Lett. 44*, 417-422.
Guo, F., Zhang, D., Kahyaoglu, A., Farid, R.S. & Jordan, F. (1998), *Biochemistry 37*, 13379-13391.
Gutowski, J.A. & lienhard, G.E. (1976), *J. Biol. Chem. 251*, 2863-2866.
Hale, G. & Perham, R.N. (1979), *Biochem. J. 177*, 129-137.
Hammes, G.G. (1981), *Biochem. Soc. Symp. 46*, 73-90.
Hanemaaijer, R., De Kok, A., Jolles, J. & Veeger, C. (1987), *Eur. J. Biochem. 169*, 245-252.
Hanemaaijer, R., Janssen, A., De Kok, A. & Veeger, C. (1988a), *Eur. J. Biochem. 174*, 593-599.
Hanemaaijer, R., Vervoort, J., Westphal, A.H., de Kok, A. & Veeger, C. (1988b), *FEBS Lett. 240*, 205-210.
Hanemaaijer, R., Westphal, A.H., Berg, A., Van Dongen, W., De Kok, A. & Veeger, C. (1989), *Eur. J. Biochem. 181*, 47-53.
Hansen, L.L., Brown, G.K., Kirby, D.M. & Dahl, H.H.M. (1991), *J. Inherit. Metab. Dis. 14*, 140-151.
Hansen, L.L., Brown, G.K., Brown, R.M. & Dahl, H.H.M. (1993), *Hum. Mol. Genet. 2*, 805-807.
Hansen, L.L., Horn, N., Dahl, H.H.M. & Kruse, T.A. (1994), *Hum. Mol. Genet. 3*, 1021-1022.
Harris, R.A. & Popov, K.M. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 139-150.
Hasson, M.S., Muscate, A., McLeish, M.J., Polonikova, L.S., Gerlt, J.A., Kenyon, G.L., Petsko, G.A. & Ringe, D. (1998), *Biochemistry 37*, 9918-9930.
Hawes, J.W., Schnepf, R.J., Jenkins, A.E., Shimomura, Y., Popov, K.M. & Harris R.A. (1995), *J. Biol. Chem. 270*, 31071-31076.
Hawkins, C.F., Borges, A. & Perham, R.N. (1989), *FEBS Lett. 255*, 77-82.
Hawkins, C.F., Borges A. & Perham R.N. (1990), *Eur. J. Biochem. 191*, 337-346.
Hein, S. & Steinbüchel, A. (1994), *J. Bacteriol. 176*, 4394-4408.
Hemalatha, S.G., Kerr, D.S., Wexler, I.D., Lusk, M.M., Kaung, M., Du, Y., Kolli, M., Schelper, L. & Patel, M.S. (1995), *Hum. Mol. Genet. 4*, 315-318.
Hendle, J., Mattevi, A., Westphal, A.H., Spee, J., De Kok, A., Teplyakov, A. & Hol, W.G.J. (1995), *Biochemistry, 34*, 4287-4298.
Hengeveld, A.F., Westphal, A.H. & De Kok, A. (1997), *Eur. J. Biochem. 250*, 260-268.
Hengeveld, A.F., Schoustra, S.E., Westphal, A.H. & De Kok, A. (1999), *Eur. J. Biochem. 265*, 1098-1107.
Hennig, J., Kern, G., Neef, H., Spinka, M., Bisswanger, H. & Hubner, G. (1997), *Biochemistry 36*, 15772-15779.
Herring, W.J., McKean, M., Dracopdi, N. & Danner, D.J. (1992), *Biochim. Biophys. Acta 1138*, 236-242.
Hester, K., Luo, J., Burns, G., Braswell, E.H. & Sokatch J.R. (1995), *Eur. J. Biochem. 233*, 828-836.
Hilliges, C., Awiszus, D. & Wendel, U. (1993), *Eur. J. Pediatr. 152*, 144-147.

Hong, Y.S., Kerr, D.S., Craigen, W.J., Tan, J., Pan, Y., Lusk, M. & Patel, M.S. (1996), *Hum. Mol. Genet. 5*, 1925-1930.
Hong, Y.S., Kerr, D.S., Liu, T.C., Lusk, M., Powell, B.R. & Patel, M.S. (1997), *Biochim. Biophys. Acta 1362*, 160-168.
Hopkins, N. & Williams, C.H. (1995), *Biochemistry 34*, 11757-11765.
Howard, M.J., Chaunan, H.J., Domingo, G.J., Fuller, C. & Perham, R.N. (2000), *J. Mol. Biol. 295*, 1023-1037.
Hübner, G., Tittman, K., Killenberg-Jabs, M., Schaffner, J., Spinka, M., Neef, H., Kern, D., Kern, G., Schneider, G., Wikner, C. & Ghisla, S. (1998), *Biochim. Biophys. Acta 1385*, 221-228.
Indo, Y. & Matsuda, J. (1996), I. In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 227-247.
Ito, M., Huq, A.H.M.M., Naito, E., Saijo, T., Takeda, E. & Kuroda, Y. (1992), *J. Inherit. Metab. Dis. 15*, 848-856.
Ito, M., Naito, E., Yokota, I., Takeda, E., Matsuda, J., Hirose, M., Sejima, H., Aiba, H., Hojo, H. & Kuroda, Y. (1995), *J. Inherit. Metab. Dis. 18*, 547-557.
Izard, T., Ævarsson, A., Allen, M.D., Westphal, A.H., Perham, R.N., de Kok, A. & Hol, W.G.J. (1999), *Proc. Natl. Acad. Sci. U.S.A. 96*, 1240-1245.
Jeng, J., Huh, T-L. & Song, B.J. (1994), *Biochem. Biophys. Res. Comm. 203*, 225-230.
Jentoft, J.E., Shoham, M. & Hurst, D. (1992), *Proteins 14*, 88-101.
Jones, D.D., Horne, H.J., Reche, P.A. & Perham, R.N. (2000), *J. Mol. Biol. 295*, 289-306.
Jones, D.D., Stott, K.M., Reche, P.A. & Perham, R.N. (2001), *J. Mol. Biol. 305*, 49-60.
Jordan, F., Baburina, I., Gao, Y., Guo, F., Kahyaoglu, A., Nemeria, N., Volkov, A., Yi, J., Zhang, D., Machado, R., Guest, J.R., Furey, W. & Hohmann, S. (1996), In: *Biochemistry and physiology of thiamin diphosphate enzymes*. Bisswanger, H., Schellenberger, A. Eds., Intemann wissenschaftlicher verlag, Prien, pp. 53-69.
Jordan, F. (1999), *FEBS Lett. 457*, 298-301.
Kalia, Y.N., Brocklehurst, S.M., Hipps, D.S., Appella, E., Sakaguchi, K. & Perham, R.N. (1993), *J. Mol. Biol. 230*, 323-41.
Kern, D., Kern, G., Neef, H., Tittmann, K., Killenberg-Jabs, M., Wikner, C., Schneider, G. & Hübner, G. (1997), *Science 275*, 67-70.
Kerr, D.S., Zinn, A.B. (1995) In: *Inherited Metabol. Dis. Diagnosis and treatment.* Fernandes, J., Saudubray, J.M., van Berghe, G. Eds., Springer Verlag, Berlin, pp. 109-119.
Kerr, D.S., Wexler, I.D., Tripatara, A. & Patel, M.S. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 249-269.
Khailova, L.S., Bernhardt, R. & Huebner, G. (1977) *Biokhimya 42*, 113-117.
Kish, S.J. (1997), *Ann. N.Y. Acad. Sci 826*, 218-228.
Kluger, R. & Pike, D.C. (1977), *J. Am. Chem. Soc. 99*, 4504-4506.
Kluger, R. (1987), *Chem. Rev. 87*, 863-876.
Knapp, J.E., Mitchell, D.T., Yazdi, M.A., Ernst, S.R., Reed, L.J. & Hackert, M.L. (1998), *J. Mol. Biol. 280*, 655-68.
Knapp, J.E., Carroll, D., Lawson, J.E., Ernst, S.R., Reed, L.J. & Hackert, M.L. (2000), *Protein Sci. 9*, 37-48.
Koike, K., Ohta, S., Urata, Y., Kagawa, Y. & Koike, M. (1988), *Proc. Natl. Acad. Sci. U.S.A. 85*, 41-45.
Koike, K., Urata, Y., Goto, S., Biochim. (1992), *Biophys. Acta 1118*, 223-230.
Kolobova, E., Tuganova, A., Boulatnikov, I. & Popov, K.M. (2001), *Biochem. J. 358*, 69-77.
Korotchkina, L.G. & Patel, M.S. (1995), *J. Biol. Chem. 269*, 14297-14303.
Korotchkina, L.G., Ali, M.S. & Patel, M.S. (1999), *Archiv. Biochem. Biophys. 369*, 277-287.
Korotchkina, L.G. & Patel, M.S. (2001), *J. Biol. Chem. 276*, 5731-5738.
Kraehenbuhl, J.P., Galardy, R.E. & Jamieson, J.D. (1974), *J. Exp. Med. 139*, 208-223.
Kumar, A., Ernst, R.R. & Wüthrich, K. (1980), *Biochem. Biophys. Res. Commun. 95*, 1-6.
Laemmli, U.K. (1970), *Nature (London) 227*, 680-685.
Lakowicz, J.R. (1999), in *Principles of fluorescence spectroscopy, 2nd Ed.* Kluwer Academic/Plenum Publishers, New York.
Lanterman, M.M., Dickinson, J.R. & Danner, D.J. (1996), *Hum. Mol. Genet. 5*, 1643-1648.
Lawson, J.E., Behal, R.H. & Reed, L.J. (1991), *Biochemistry 30*, 2834-2839.
Leigh, D. J. (1951), *Neurol. Neurosurg. Psychiatry 14*, 216-221.
Lessard, I.A.D. & Perham, R.N. (1994), *J. Biol. Chem. 269*, 10378-10383.
Lessard, I.A.D. & Perham, R.N. (1995), *Biochem. J. 306*, 726-733.
Lessard, I.A.D., Fuller, C. & Perham, R.N. (1996), *Biochemistry 35*, 16863-16870.
Leung, L.S. & Frey, P.A. (1978), *Biochem. Biophys. Res. Comm. 81*, 274-279.
Lindqvist, Y., Schneider, G., Ermler, U. & Sundström, M. (1992), *EMBO J. 11*, 2373-2379.
Lindqvist, Y. & Schneider, G. (1993), *Curr. Opin. Struct. Biol. 3*, 896-901.

Ling, M., McEachern, G., Seyda, A., MacKay, N., Scherer, S.W., Bratinova, S., Beatty, B., Giovannucci-Uzielli, M.L. & Robinson, B.H. (1998), *Hum. Mol. Genet. 7*, 501-505.
Linn, T.C., Pettit, F.H. & Reed, L.J. (1969), *Proc. Natl. Acad. Sci. USA 62*, 234-241.
Lissens, W., Desguerre, I., Benelli, C., Marsac, C., Fouque, F., Haenggeli, C., Ponsot, G., Seneca, S., Liebaers, I. & De Meirleir, L. (1995), *Hum. Mol. Genet. 4*, 307-308.
Lissens, W., De Meirleir, L., Seneca, S., Benelli, C., Marsac, C., ThePoll, B.T., Briones, P., Ruitenbeek, W., Van Diggelen, O., Chaigne, D., Ramaekers, V. & Liebaers, I. (1996), *Hum. Mutat. 7*, 46-51.
Lissens, W., De Meirleir, L., Seneca, S., Liebaers, J., Brown, G.K., Brown, R.M., Ito, M., Naito, E., Kuroda, Y., Kerr, D.S., Wexler, J.D., Patel, M.S., Robinson & B.H., Seyda, A. (2000), *Hum. Mutat. 15*, 209-219.
Liu, T.C., Kim, H., Arizmendi, C., Kitano, A. & Patel, M.S. (1993), *Proc. Natl. acad. Sci. USA 90*, 5186-5190.
Liu, T.C., Korotchkina, L.G., Hyatt, S.L., Vettakkorumakankav, N.N. & Patel, M.S. (1995), *J. Biol. Chem. 270*, 15545-15550.
Lobell, M. & Crout, D.H.G. (1996), *J. Am. Chem. Soc. 118*, 1867-1873.
Lowe, P.N., Leeper, F.J. & Perham, R.N. (1983), *Biochemistry 22*, 150-157.
Lowe, P.N. & Perham, R.N. (1984), *Biochemistry 23*, 91-97.
Lyon, M.F. (1962), *Am. J. Hum. Genet. 14*, 135-148.
Macura, S. & Ernst, R.R. (1980), *Mol. Phys. 41*, 95-117.
Maldonado, M.E., Oh, K.J. & Frey, P.A. (1972), *J. Biol. Chem. 247*, 2711-2716.
Mande, S.S., Sarfaty, S., Allen, M.D., Perham, R.N. & Hol, W.G.J. (1996), *Structure 4*, 277-286.
Marion, D. & Wüthrich, K. (1983), *Biochem. Biophys. Res. Commun. 113*, 967-974.
Marsac, C., Stansbie, D., Bonne, G., Cousin, J., Jehenson, P., Benelli, C., Leroux, J.P. & Lindsay, G. J. (1993), *Pediatr. 123*, 915-920.
Marsac, C., Benelli, C., Desguerre, I., Diry, M., Fouque, F., De Meirleir, L., Ponsot, G., Seneca, S., Poggi, F., Saudubray, J.M., Zabot, M.T., Fontan, D. & Lissens, W. (1997), *Hum. Genet. 99*, 785-792.
Mastrogiacomo, F., LaMarche, J., Dozic, S., Lindsay, G., Bettendorff, L., Robitaille, Y., Schut, L. & Kish, S.J. (1996), *Neurodegen. 5*, 27-33.
Matsuda, I, Nobukuni, Y., Mitsubuchi, H., Indo, Y., Endo, F., Asaka, J. & Harada, A. (1990), *Biochem. Biophys. Res. Comm. 172*, 451-464.
Matsuda, J., Ito, M., Naito, E., Yokota, I. & Kuroda, Y. (1994), *Proc. VIth Int. Cong. Inborn Errors Metab. Milan*, pp. 65.
Mattevi, A., Schierbeek, A.J. & Hol W.G.J. (1991), *J. Mol. Biol. 220*, 975-994.
Mattevi, A., De Kok, A. & Perham, R.N. (1992a), *Curr. Opin. Struct. Biol. 2*, 877-887.
Mattevi, A., Obmolova, G., Kalk, K.H., Sokatch, J., Betzel, C.H. & Hol, W.G.J. (1992b), *Proteins 13*, 336-351.
Mattevi, A., Obmolova, G., Schulze E., Kalk, K.H., Westphal, A.H., De Kok, A. & Hol, W.G.J. (1992c), *Science 255*, 1544-1550.
Mattevi, A., Obmolova, G., Kalk, K.H., Westphal, A.H., De Kok, A. & Hol, W.G.J. (1993a), *J. Mol. Biol. 230*, 1183-1199.
Mattevi, A., Obmolova, G., Kalk, K.H., van Berkel, W.J. & Hol, W.G.J.. (1993b), *J. Mol. Biol. 230*, 1200-15.
Mattevi, A., Oblomova, G., Kalk, K.H., Teplyakov, A. & Hol, W.G.J. (1993c), *Biochemistry 32*, 3887-3901.
Matthews, P.M., Brown, R.M., Otero, L., Marchington, D.R., Leonard, J.V. & Brown, G.K. (1993a), *Neurology 43*, 2025-2030.
Matthews, P.M., Marchington, D.R., Squier, M., Land, J., Brown, R.M. & Brown, G.K. (1993b), *Ann. Neurol. 33*, 652-655.
Matthews, P.M., Brown, R.M., Otero, L.J., Marchington, D.R., LeGris, M., Howes, R., Meadows, L.S., Shevell, M., Scriver, C.R. & Brown, G.K. (1994), *Brain 117*, 435-443.
Matuda S, Nakano K, Ohta S, Saheki T, Kawanishi Y & Miyata T. (1991), *Biochim. Biophys. Acta 1089*, 1-7.
McNally, A.J., Mattsson L. & Jordan F. (1995a), *J. Biol. Chem. 270*, 19744-19751.
McNally, A.J., Motter K. & Jordan F. (1995b), *J. Biol. Chem. 270*, 19736-19743.
Meng, M. & Chuang, D.T. (1994), *Biochemistry 33*, 12879-12885.
Menon, S. & Ragsdale, S.W. (1997), *Biochemistry 36*, 8484-8494.
Meshalkina, L., Nilsson, U., Wikner, C., Kostikova, T. & Schneider, G. (1997), *Eur. J. Biochem. 244*, 646-652.
Miles, J.S., Guest, J.R., Radford, S.F. & Perham, R.N. (1988), *J. Mol. Biol. 202*, 97-106.
Mitsubuchi, H., Nobukuni, Y., Akaboshi, I., Indo, Y., Endo, F. & Matsuda, I. (1991), *J. Clin. Invest. 87*, 1207-1211.
Mizuno, Y., Matsuda, S., Yoshino, H., Mori, H., Hattori, N. &, Ikebe, S. (1994), *Ann. Neurol. 35*, 204-210.
Morten, K.J., Caky, M. & Matthews, P.M. (1998), *Neurology 51*, 1324-1330.
Morten, K.J., Beattie, P., Brown, G.K. & Matthews, P.M. (1999), *Neurol. 53*, 612-616.
Motokawa (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 119-130.

Muller, Y.A., Lindqvist, Y., Furey, W., Schulz, G.E., Jordan, F. & Schneider, G. (1993), *Structure 1*, 95-103.
Muller, Y.A. & Schulz, G.E. (1993), *Science 259*, 965-967.
Naito, E., Ito, M., Takeda, E., Yokota, I., Yoshijima, S. & Kuroda, Y. (1994), *Pediat. Res. 36*, 340-346.
Naito, E., Ito, M., Yokota, I., Saijo, T., Matsuda, J., Osaka, H., Kimura, S. & Kuroda, Y. J. (1997), *Inherit. Metab. Dis. 20*, 539-548.
Naito, E., Ito, M., Yokota, I., Saijo, T., Chen, S., Maehara, M. & Kuroda, Y. (1999), *J. Neurol. Sci. 171*, 56-9.
Neagle, J.C. & Lindsay, J.G. (1991), *Biochem. J. 278*, 423-427.
Nemerya, N.S., Zemskova, M.A., Nyukhalkina, I.A. & Khailova, L.S. (1996), *FEBS Lett. 394*, 96-98.
Neveling, U., Bringer-Meyer, S. & Sahm, H. (1999), *FEMS Microb. Lett. 177*, 117-121.
Nikkola, M., Lindqvist, Y. & Schneider, G. (1994), *J. Mol. Biol. 238*, 387-404.
Nobukino, Y., Mitsubuchi, H., Hayashida, Y., Ohta, K., Indo, Y., Ichiba, Y., Endo, F. & Matsuda, I. (1993), *Biochim. Biophys. Acta 1225*, 64-70.
Novikov, E.G., van Hoek, A., Visser, A.J.W.G. & Hofstraat, J.W. (1999), *Opt. Commun. 166*, 189-198.
Oliver, R.M. & Reed, L.J. (1982), *In: Electron microscopy of proteins* (Harris, J.R. eds.), Acadamic Press, London, pp. 1-48.
Otero, L.J., Brown, R.M. & Brown, G.K. (1998), *Hum. Mutat. 12*, 114-121.
Packman, L.C., Perham, R.N. & Roberts, G.C. (1984a), *Biochem. J. 217*, 219-227.
Packman, L.C., Hale, G. & Perham, R.N. (1984b), *EMBO J. 3*, 1315-1319.
Packman, L.C. & Perham, R.N. (1987), *Biochem. J. 242*, 531-538.
Palmer, J.A., Madhusud, K.T., Hatter, K. & Sokatch, J.R. (1991), *Eur. J. Biochem. 202*, 231-240.
Pan, K. & Jordan, F. (1998), *Biochemistry 37*, 1357-1364.
Patel, M.S. & Roche, T.E. (1990), *FASEB J. 4*, 3224-3233.
Patel, M.S. & Harris, R.A. (1995), *FASEB J. 9*, 1164-1172.
Patel, M.S., Roche, T.E. & Harris, R.A. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Birkhäuser Verlag, Basel, pp. 1-317.
Pawelczyk, T., Easom, R.A. & Olson, M.S. (1992), *Arch. Biochem. Biophys. 294*, 44-49.
Peinemann, F. & Danner, D.J. J. (1994), *Inherit. Metab. Dis. 17*, 3-15.
Perham, R.N., Duckworth, H.W. & Roberts, G.C.K. (1981), *Nature 292*, 474-477.
Perham, R.N. & Roberts, G.C.K. (1981), *Biochem. J.* 199, 733-740.
Perham, R.N. (1991), *Biochemistry 30*, 8501-8512.
Perham, R.N. (1996), In: *Alpha-keto acid dehydrogenase complexes*. Patel, M.S., Roche, T.E., Harris, R.A., Eds., Birkhäuser verlag, Basel, pp. 1-15.
Perham, R.N. (2000), *Ann. Rev. Biochem. 69*, 961-1004.
Podebrad, F., Heil, M., Reichert, S., Mosandl, A., Sewell, A.C. & Bohles, H. J. (1999), *Inher. Metab. Dis. 22*, 107-114.
Provencher, S.W. & Glockner, J. (1981), *Biochemistry* 20, 33-37.
Rae, J.L., Cutfield, J.F., Lamont, I.L. (1997), *J. Bacteriol. 179*, 3561-3571.
Rahman, S., Blok, R.B., Dahl, H.H.M., Danks, D.M., Kirby, D.M., Chow, C.W., Christodoulou, L. & Thorburn, D.R. (1996), *Ann. Neurol. 39*, 343-351.
Randle, P.J. (1986), *Biochem. Soc. Trans. 14*, 1799-1806.
Randle, P.J., Priestman, D.A. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 151-162.
Reed, L.J. & Oliver, R.M. (1968a), Brookhaven symposia in biology, Structure, Function and Evolution in *Protein 21* 397-411.
Reed, L.J. & Oliver, R.M. (1968b), Structure, Funct. Evol. *Protein 21*, 397-411.
Reed, L.J. (1974), *Acc. Chem. Res. 7*, 40- 46.
Reed, L.J., Pettit, F.H., Eley, M.H., Hamilton, L., Collins, J.H. & Oliver, R.M. (1975), *Proc. Natl. Acad. Sci. USA 72*, 3068-3072.
Reed, L.J. & Hackert, M.L. (1990), *J. Biol. Chem. 265*, 8971-8974.
Reed, L.J., Lawson, J.E., Niu, X.D. & Yan, J. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 131-138.
Regnström, K., Sauge-Merle, S., Chen, K. & Burgess, B.K. (1999), *Proc. Natl. Acad. Sci. U.S.A. 96*, 12389-12393.
Ricaud, P.M., Howard, M.J., Roberts, E.L., Broadhurst, R.W. & Perham R.N. (1996), *J. Mol. Biol. 264*, 179-90.
Rice, J.E., Dunbar, B. & Lindsay, J.G. (1992), *EMBO. J. 11*, 3229-3235.
Robien, M.A., Clore, G.M., Omichinski, J.G., Perham, R.N., Appella, E., Sakaguchi, K. & Gronenborn, A.M. (1992), *Biochemistry 31*, 3463-3471.
Robinson, B.H. & Chun, K. (1993), *FEBS Lett. 328*, 99-102.

Robinson, B.H. (1995) In: *The metabolic and molecular bases of inherited disease*, 7th ed. Scriver, S.R., Beaudet, A.L., Sly, W.S., Valle, D., Eds., McGraw-Hill Inc., pp. 1463-1471.
Roche, T.E. & Patel, M.S. (1989) *Ann. NY Acad. Sci. 573*, 1-474.
Roche, T.E., Liu, S., Ravindran, S., Baker, J.C. & Wang, L. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 33-52.
Rodriguez-Zavala, J.S., Pardo, J.P. & Moreno-Sanchez, R. (2000), *Arch. Biochem. Biophys. 379*, 78-84.
Rost, B. & Sander, C. (1993), *J. Mol. Biol. 232*, 584-599.
Rost, B. & Sander, C. (1994), *Proteins 19*, 55-72.
Russell, G.C. & Guest, J.R. (1991a), *Biochim. Biophys. Acta 1076*, 225-232.
Russell, G.C. & Guest, J.R. (1991b), *Proc. Roy. Soc. Lond. (B) 243*, 155-160.
Russell, G.C., Machado, R.S. & Guest, J.R. (1992), *Biochem. J. 287*, 611-619.
Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989), *Molecular Cloning. A Laboratory Manual, 2nd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanderson, S.J., Khan, S.S., McCarthey, R.G., Miller, C. & Lindsay, J.G. (1996), *Biochem. J. 319*, 109-116.
Santiard-Baron, D., Gosset, P., Nicole, A., Sinet, P.M., Christen, Y. & Ceballos-Picot, I. (1999), *Exp. Neurol. 158*, 206-213.
Schadewaldt, P. & Wendel, U. (1997), *Eur. J. Pediatr. 156*, S62-S66.
Schägger, H. & Von Jagow, G. (1987), *Anal. Biochem. 166*, 368-379.
Schellenberger, A. (1998a), *Biochim. Biophys Acta 1385*, 173-428.
Schellenberger, A. (1998b), *Biochim. Biophys. Acta, 1385*, 177-186.
Schneider, G. & Lindqvist, Y. (1998), *Biochim. Biophys. Acta 1385*, 387-398.
Schonbrunn-Hanebeck, E., Laber, B. & Amrhein, N. (1990), *Biochemistry 29*, 4880-4885.
Schrenk, D.F. & Bisswanger, H. (1984), *Eur. J. Biochem. 143*, 561-566.
Schulze, E., Benen, J.A.E., Westphal, A.H. & De Kok, A. (1991a), *Eur. J. Biochem. 200*, 29-34.
Schulze E., Westphal, A.H., Obmolova, G., Mattevi, A., Hol, W.J.G. & De Kok, A. (1991b), *Eur. J. Biochem. 201*, 561-568.
Schulze E., Westphal, A.H., Boumans, H. & De Kok, A. (1991c), *Eur. J. Biochem. 202*, 841-848.
Schulze E., Westphal, A.H., Veeger, C. & De Kok, A. (1992), *Eur. J. Biochem. 206*, 427-435.
Schulze, E., Westphal, A.H., Hanemaaijer, R. & De Kok, A. (1993), *Eur. J. Biochem. 211*, 591-599.
Schwartz, E.R., Old, L.O. & Reed, L.J. (1968), *Biochem. Biophys. Res. Comm. 31*, 495-500.
Schwartz, E.R. & Reed, L.J. (1970), *Biochemistry 9*, 1434-1439.
Sedewitz, B., Schleifer, K.H. & Götz, F. (1984a), J. Bacteriol. 160, 273-278.
Sedewitz, B., Schleifer, K.H. & Götz, F. (1984b), J. Bacteriol. 160, 462-465.
Shaag, A., Saada, A., Berger, I., Mandel, H., Joseph, A., Feigenbaum, A. & Elpeleg, O.N. (1999), *Am. J. Med. Genet. 82*, 177-182.
Shany, E., Saada, A., Landau, D., Shaag, A., Herskovitz, E. & Elpeleg, O.N. (1999), *Biochem. Biophys. Res. Comm. 262*, 163-166.
Sheu, K.F.R., Brown, A.M., Haroutunian, V., Kristal, B.S., Thaler, H., Lesser, M., Kalaria, R.N., Relkin, N.R., Mohs, R.S., Lilius, L., Lannfelt, L. & Blass, J.P. (1999), *Ann. Neurol. 45*, 48-53.
Shin, W., Pletcher, J., Blamk, G. & Sax, M. (1977), *J. Am. Chem. Soc. 99*, 3491-3499.
Sittler, T.L., McKean, M.C., Peinemann, F., Jackson, E. & Danner, D.J. (1998), *Biochim. Biophys. Acta 1404*, 385-392.
Skinner, D.D., Morgenstern, M.R., Fedechko, R.W. & Denoya, C.D. (1995), *J. Bacteriol. 177*, 183-190.
Sober, H.A. & Harte, R.A. (1970), *Handbook of Biochemistry, 2nd ed.* The Chemical rubber Co., Cleveland, Ohio.
Stephens, P.E., Darlison, M.G., Lewis, H.M. & Guest J.R. (1983), *Eur. J. Biochem. 133*, 155-162.
Stevenson, K.J., Hale, G. & Perham, R.N. (1978), *Biochemistry 17*, 2189-2192.
Stoops, J.K., Cheng, R.H., Yazdi, M.A., Maeng, C.Y., Schroeter, J.P., Klueppelberg, U., Kolodziej, S.J., Baker, T.S. & Reed, J.L. (1997), *J. Biol. Chem. 272*, 5757-5764.
Sun, S., Duggleby, R. & Schowen, R. (1995), *J. Am. Chem. Soc. 117*, 7317-7322.
Sundström, M., Lindqvist, Y. & Schneider, G. (1992), *FEBS Lett. 313*, 229-231.
Sundström, M., Lindqvist, Y., Schneider, G., Hellman, U. & Ronne, H. (1993), *J. Biol. Chem. 268*, 24346-24352.
Takakubo, F., Thorburn, D.R. & Dahl, H.H.M. (1993a), *Hum. Mol. Genet. 2*, 473-474.
Takakubo, F., Thorburn, D.R. & Dahl, H.H.M. (1993b), *Hum. Mol. Genet. 2*, 1961-1962.
Takakubo, F., Thorburn, D.R., Brown, R.M., Brown, G.K. & Dahl, H.H.M. (1995), *Hum. Mutat. 6*, 274-275.
Texter, F.L., Radford, S.E., Laue, E.D., Perham, R.N., Miles, J.S. & Guest, J.R. (1988), *Biochemistry 27*, 289-296.
Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994), *Nucleic Acids Res. 22*, 4673-4680.

Toyoda, T., Suzuki, K., Sekiguchi, T., Reed, L.J. & Takenaka, A. J. (1998), *Biochem. 123*, 668-674.
Treacy, E., Clow, C.L., Reade, T.R., Chitayat, D., Manner O.A. & Scriver, C.R. (1992), *J. Inherit. Metab. Dis. 15*, 121-135.
Tripatara, A., Kerr, D.S., Lusk, M.M., Kolli, M., Tan, J.& Patel, M.S. (1996), *Hum. Mutat. 8*, 180-182.
Tripatara, A., Korotchkina, L.G. & Patel, M.S. (1999), *Arch. Biochem. Biophys. 367*, 39-50.
Turner, S.L., Russell, G.C., Williamson, M.P. & Guest, J.R. (1993), *Protein Eng. 6*, 101-108.
Van den Broek (1971), *PhD Thesis, Wageningen University*, The Netherlands.
Van Mierlo, C.P.M., Darby, N.J., Neuhaus, D. & Creighton, T.E. (1991), *J. Mol. Biol. 222*, 373-390.
Van Mierlo, C.P.M., Darby, N.J., Keeler, J., Neuhaus, D. & Creighton, T.E. (1993), *J. Mol. Biol. 229*, 1125-1146.
Vieira, J. & Messing J. (1982), *Gene 19*, 259-268.
Visser, A.J.W.G. & Lee, J. (1980), *Biochemistry 19*, 4366-4372.
Wagenknecht, T., Grassucci, R. & Schaak, D. (1990), *J. Biol. Chem. 265*, 22402-22408.
Wagenknecht, T., Grassucci, R., Berkowitz, J. & Forneris, C. (1992), *J. Struct. Biol. 109*, 70-77.
Wang, J.J.L., Martin, P.R. & Singleton, C.K. (1997), *Biochem. Biophys. Acta 1341*, 165-172.
Westphal, A.H. & De Kok, A. (1988), *Eur. J. Biochem. 172*, 299-305.
Westphal, A.H., Fabisz-Kijowska, A., Kester, H., Obels, P.P. & De Kok, A. (1995), *Eur. J. Biochem. 234*, 861-870.
Wexler, I.D., Hemalatha, S.G. & Patel, M.S. (1991), *FEBS Lett. 282*, 209-213.
Wexler, I.D., Hemalatha, S.G., Liu, T.C., Berry, S.A., Kerr, D.S. & Patel, M.S. (1992), *Pediat. Res. 32*, 169-174.
Wexler, I.D., Hemalatha, S.G., Dahl, H.H.M., Buist, N.R., Berry, S.A., Cederbaum, S.D., Du, Y., Kuang, M., McConnell, J., Patel & M.S., Kerr, D.S. (1995), *Proc. Soc. Inher. Metab. Disorders, Perdido Beach, FL.*
Wexler, I.D., Hemalatha, S.G., McConnell, J., Buist, N.R., Dahl, H.H.M., Berry, S.A., Cederbaum, S.D., Patel, M.S. & Kerr, D.S. (1997), *Neurology 49*, 1655-1661.
Wikner, C., Meshalkina, L., Nilsson, U., Nikkola, M., Lindqvist, Y., Sundström, M. & Schneider, G. (1994), *J. Biol. Chem. 269*, 32144-32150.
Wikner, C., Meshalkina, L., Nilsson, U., Bäckström, S., Lindqvist, Y. & Schneider, G. (1995), *Eur. J. Biochem. 233*, 750-755.
Wikner, C., Nilsson, U., Meshalkina, L., Udekwu, C., Lindqvist, Y. & Schneider, G. (1997), *Biochemistry 36*, 11643-11649.
Williams Jr., C.H. (1992), In: *Chemistry and biochemistry of flavoenzymes*, Müller, F., Ed., CRC Press: Boca Raton, 3, pp.121-211.
Witzmann, S. & Bisswanger, H. (1998), *Biochim. Biophys. Acta 1385*, 341-352.
Wynn, R.M., Chuang, J.L., Davie, J.R., Fisher, C.W., Hale, M.A., Cox, R.P. & Chuang D.T. (1992a), *J. Biol. Chem. 267*, 1881-1887.
Wynn, R.M., Davie, J.R., Cox, R.P. & Chuang D.T. (1992b), *J. Biol. Chem. 267*, 12400-12403.
Wynn, R.M., Davie, J.R., Meng, M. & Chuang, D.T. (1996), In: *Alpha-keto acid dehydrogenase complexes*, Patel, M.S., Roche, T.E., Harris, R.A., Ed., Birkhäuser Verlag: Basel, pp. 101-107.
Wynn, R.M., Davie, J.R., Chuang, J.L., Cote, C.D. & Chuang, D.T. (1998), *J. Biol. Chem. 273*, 13110-13118.
Wynn, R.M., Song, J.L. & Chuang, D.T. (2000), *J. Biol. Chem. 275*, 2786-2794.
Wynn, R.M., Ho., R., Chuang, J.L. & Chuang, D.T. (2001), *J. Biol. Chem. 276*, 4168-4174.
Yang, D., Song, J., Wagenknecht, T. & Roche, T.E. (1997), *J. Biol. Chem. 272*, 6361-6369.
Yang, H.C., Hainfeld, J.F., Wall, J.S. & Frey, P.A. (1985), *J. Biol. Chem. 260*, 16049-16051.
Yang, Y.S., Datta, A., Hainfeld, R., Wall, J.S. & Frey, P.A. (1994), *Biochemistry 33*, 9428-9437.
Yi, J., Nemeria, N., McNally, A., Jordan, F., Machado, R.S. & Guest, J.R. (1996), *J. Biol. Chem. 271*, 33192-33200.
Zhang, B., Crabb, D.W. & Harris, R.A. (1988), *Gene 69*, 159-164.
Zhao, Y., Hawes, J., Popov, K.M., Jaskiewicz, J., Shimomura, Y., Crabb, D.W. & Harris, R.A. (1994). *J. Biol. Chem. 269*, 18583-18587.
Zhou, Z.H., Liao, W., Cheng, R.H., Lawson, J.E., McCarthy, D.B., Reed, L.J. & Stoops, J.K. (2001), *J. Biol. Chem. 276*, 21704-21713.

Samenvatting

Het pyruvaat dehydrogenase multienzym complex (PDHC) behoort tot een familie van multienzym complexen die de irreversibele oxidatieve decarboxylatie katalyseert van verschillende 2-oxozuur substraten naar hun corresponderende acyl-CoA derivaten, NADH and CO_2. 2-oxozuur dehydrogenase complexen bezetten strategische punten in het primaire energie metabolisme: PDHC in de glycolyse, het oxoglutaraat dehydrogenase complex (OGDC) in de citroenzuur cyclus, het branched-chain oxozuur dehydrogenase complex (BCDC) in aminozuur katabolisme, het glycine dehydrogenase complex (GDHC) in het glycine splitsingsysteem en het acetoine dehydrogenase complex (ADHC) in acetoine metabolisme. Het slecht, of niet functioneren, van een 2-oxozuur dehydrogenase complex in de mens resulteert in een grote verscheidenheid aan klinische verschijnselen, zoals bijvoorbeeld melkzuur acidose (PDHC-deficientie) of de 'maple syrup urine' ziekte (BCDC-deficientie).

Het PDHC bestaat uit meerdere kopieën van drie enzymen: het 2-oxozuur dehydrogenase of de E1p component, het dihydrolipoyl acyltransferase (de E2p component) en het dihydrolipoamide dehydrogenase (de E3 component). In de afkorting E1p verwijst de "p" naar het bijbehorende complex; PDHC, "o" verwijst naar OGDC en "b" verwijst naar BCDC. De E2 component speelt zowel katalytisch als structureel een centrale rol in het complex. Het bestaat uit meerdere domeinen: van N- naar C-terminus, 1 tot 3 lipoyl domeinen, een E1/E3-bindingsdomein en een katalytisch domein. De domeinen zijn met flexibele 'linkers' aan elkaar verbonden. Deze architectuur zorgt voor de hoge mate van flexibiliteit van de individuele domeinen die essentieel is voor het voor het complex zo belangrijke 'active site coupling' mechanisme. In Gram-negatieve bacteriën vormt het C-terminale katalytisch domein een kern die bestaat uit 24 subeenheden; een kubus gevormd uit 8 trimeren (op elk hoekpunt een trimeer). Meerdere kopieën van E1p en E3 binden, niet covalent, aan de buitenzijde van deze kern. Van al deze verschillende enzymen of domeinen van enzymen van PDHC zijn inmiddels drie dimensionale structuren opgelost met behulp van Röntgen kristallografie of NMR.

Om een beter inzicht te krijgen in de structuur en in het functioneren van de E1p component van het PDHC uit *A. vinelandii* en van zijn interacties binnen het complex, is in 1992 een SON/NWO OIO project gestart. De resultaten van dit onderzoek zijn beschreven in dit

proefschrift. E1 katalyseert de snelheidslimiterende stap van het reactiemechanisme; de reductieve acylering van de lipoyl groepen van E2, waarbij het gebruik maakt van thiamine difosfaat (ThDP). E1 bestaat of als een homodimeer (α_2), of als een heterotetrameer ($\alpha_2\beta_2$), afhankelijk van het soort en de oorsprong van het complex. Er zijn verrassend weinig overeenkomsten tussen de sequenties van homodimeer en heterotetrameer E1, of zelfs maar tussen homodimeer E1o en homodimeer E1p. Een uitzondering hierop is het zogenaamde 'ThDP-bindingsmotief'. Dit motief is gevonden in alle tot nu toe bekende ThDP afhankelijke enzymen. Tot zeer recent was E1 verreweg de slechtst gekarakteriseerde component van het complex. Dit gebrek aan (structurele) informatie werd grotendeels veroorzaakt door het gebrek aan een stabiele vorm van het eiwit: wanneer E1 niet gebonden is aan het complex is het instabiel.

Hoofdstuk 2 beschrijft de klonering en het sequensen van het gen dat codeert voor E1p uit *Azotobacter vinelandii* en, voor de eerste maal, de expressie en zuivering van een homodimeer E1. Het beschikbaar komen van een stabiel E1 maakte het mogelijk om dit enzym in meer detail te bestuderen. Het katalytisch mechanisme van het gekloneerde E1p is exact hetzelfde als dat van het wildtype complexgebonden enzym; wat betreft de substraatbinding (pyruvaat) en het allostere gedrag werden geen verschillen waargenomen. Bovendien resulteert het combineren van E2 en E3 met het gekloneerde E1 in volledige complex activiteit. E1 is stabiel in buffers met een hoge zoutsterkte, maar het wordt irreversibel geïnactiveerd bij een hoge pH. Waarschijnlijk wordt dit veroorzaakt door het onvermogen van E1 om na monomeriseren correct dimeren te vormen. Dit verklaart de eerder gevonden instabiliteit van het wildtype E1p nadat het van het complex verwijderd is door incubatie bij hoge pH. Experimenten laten zien dat acetyl coenzym A zich gedraagt als een 'feedback inhibitor' via binding aan de E1p component.

Gelimiteerde proteolyse experimenten laten zien dat het N-terminale gebied van E1p makkelijk kan worden verwijderd. De ontstane eiwitfragmenten zijn actief wanneer een artificiële elektronenacceptor wordt gebruikt, maar hebben hun vermogen verloren om aan de centrale component (E2p) te binden en op die manier complex activiteit te herstellen. Dit wijst er op dat het N-terminale gebied van E1p betrokken is bij de binding van E1p aan E2p.

Om deze hypothese verder te onderzoeken zijn er N-terminale deletiemutanten van E1p gemaakt en gekarakteriseerd. Deze experimenten worden beschreven in hoofdstuk 3.

Maximaal 9 N-terminale aminozuren kunnen verwijderd worden zonder effect te hebben op de eigenschappen van het enzym. Het inkorten van E1p tot en met 48 aminozuren heeft geen

effect op de expressie of de vouwing van het enzym, maar deze ingekorte enzymen zijn niet langer instaat om aan E2p te binden. Dit bevestigt de betrokkenheid van het N-terminale gebied bij de binding van E1p aan E2p. De mutant waarvan 48 aminozuren zijn verwijderd (E1pΔ48) is in detail gekarakteriseerd: Het is katalytisch volledig functioneel, het heeft een V_{max} welke gelijk is aan die van het wildtype E1p, het kan de lipoamide groep gebonden aan het lipoyl domein reductief acetyleren en het vormt dimeren. Aan de andere kant is de $S_{0.5}$ voor pyruvaat significant afgenomen.

Om te bestuderen of één of twee N-termini van E1p betrokken zijn bij de binding aan E2p is er een heterodimeer gemaakt met één subunit van het wildtype E1p en één subunit van E1pΔ48. Uit het feit dat de heterodimeer niet in staat is om aan E2p te binden volgt dat beide N-terminale domeinen in een E1p dimeer nodig zijn voor de binding van E1p aan E2p.

Verschillende experimenten (sequentievergelijkingen en mutagenese) aan E2p wijzen er op dat de interacties tussen E1p en E2p hoofdzakelijk elektrostatisch van karakter zullen zijn. De interacties vinden plaats tussen negatief geladen residuen in het N-terminale domein van E1p en positief geladen residuen in zowel het bindingsdomein als het katalytisch domein van E2p. De bovengenoemde mutagenese experimenten demonstreren dat E1p, anders dan E3, bindt aan zowel het bindingsdomein als het katalytisch domein van E2p.

Hoofdstuk 4 beschrijft in detail de karakterisering van de structuur en van de functie van een synthetisch peptide (Nterm-E1p) dat het N-terminale gebied (aminozuur 4-45) van E1p vertegenwoordigd. Activiteitsmetingen en bindingsstudies bevestigen dat Nterm-E1p specifiek met E1p concurreert voor de binding aan E2p. Bovendien tonen deze experimenten aan dat het N-terminale gebied van E1p een zelfstandig vouwend domein vormt dat functioneert als een bindingsdomein. Zowel CD metingen, 2D-^{1}H NMR analyse en secondaire structuurvoorspellingen geven aan dat Nterm-E1p voornamelijk uit α-helix bestaat.

Hoofdstuk 5 beschrijft de constructie van puntmutaties in het N-terminale domein (D17Q, D17R, E20Q, E20R, D24Q en D24R) met als doel de specifieke residuen te identificeren welke betrokken zijn bij de binding van E1p aan E2p. Analyse van het katalytisch gedrag en bindingsstudies tonen aan dat aminozuur D17 essentieel is voor de binding van E1p aan E2p. Aminozuur D24 is betrokken bij de binding, maar lijkt niet essentieel. Aminozuur E20 is niet betrokken bij de binding van E1p aan E2p. Geen van de mutaties hebben invloed op de vouwing of op de dimeer vorming van E1p, wat aantoont dat het effect van de mutaties op de binding niet veroorzaakt wordt door slechte vouwing van het eiwit.

We hebben het model van de interacties tussen E1p en E2p zoals dat is voorgesteld in hoofdstuk 4 in hoofdstuk 5 verder verfijnt: in het 'helix-turn-helix' motief van het N-terminale domein zitten de residuen D17, E20 en D24 aan één kant van de helix. Aan de tegenoverliggende zijde van deze helix bevinden zich W15, L19 en V22 die samen een hydrofoob gebiedje vormen. Helix II bevat ook een hydrofobe kant, welke een hydrofobe 'pocket' kan vormen met de residuen uit helix I. Geen van de mutaties resulteert in totale vernietiging van de bindingscapaciteit van E1p; hetzelfde geldt voor de puntmutaties gemaakt in E2p. Het raakvlak tussen E1p en E2p is schijnbaar opgebouwd uit meer dan één regio in zowel E2p (het bindingsdomein en het katalytisch domein) alsmede in E1p.
Al eerder werd aangetoond dat de binding tussen E1p en E2p soort specifiek is: *E. coli* E1p bind vrijwel niet aan *A. vinelandii* E2p. Het feit dat de mutaties D17Q en D24R (*A. vinelandii* aminozuren naar *E. coli* aminozuren) beiden zo'n sterk effect hebben op het bindingsvermogen van E1p correleert goed met deze observatie.
De spectroscopische metingen in hoofdstuk 4 tonen aan dat het peptide Nterm-E1p aanwezig is in twee toestanden, de populatie van de twee vormen hangt af van de condities van het monster. De twee gevonden structuren zijn respectievelijk "ontvouwen" bij een pH boven de 6 en "gevouwen" bij een pH beneden de 5 genoemd. Het 2D-^{1}H TOCSY spectrum van "gevouwen" Nterm-E1p laat zien dat er hier eigenlijk sprake is van een mengsel van gevouwen en ontvouwen Nterm-E1p. Bovendien zijn er uitwisselingskruispieken tussen de twee toestanden te zien. De uitwisselingssnelheid tussen de twee populaties valt in het gebied van 0.5 tot 5 seconden^{-1}. De geobserveerde scherpe resonanties in de NMR spectra van wildtype E1p tonen aan dat dit 200 kDa grote enzym zeer flexibele gebieden bevat. Deels vinden deze scherpe resonanties hun oorsprong in de N-terminus van E1p. Het geobserveerde dynamische karakter van E1p en van Nterm-E1p is waarschijnlijk nodig voor de binding van de E1p dimeer aan de twee verschillende bindinggebieden in E2p. De flexibiliteit kan ook nodig zijn voor het behoud van het allostere karakter van het enzym wanneer het gebonden is aan het complex.
Samenvattend beschrijven de hoofdstukken 2-5 tot in detail de opheldering van de binding tussen E1p en E2p.

Pas aan het einde van dit project is gedetailleerde structurele informatie beschikbaar gekomen over E1. Om deze reden hebben we in hoofdstuk 6 primaire en secondaire structurele informatie gecombineerd met als doel het verkrijgen van een sequentievergelijking van

homodimeer en heterotetrameer 2-oxozuur dehydrogenase en van transketolase. Het is opvallend dat in tegenstelling tot hun structurele overeenkomsten, er een totaal gebrek aan sequentie identiteit lijkt te bestaan tussen deze eiwitten. De gemaakte sequentievergelijking bestaat uit twee delen, overeenkomstig met de α- en de β- keten van het heterotetramere 2-oxozuur dehydrogenase. Vergelijken van de sequenties van homodimeer pyruvaat dehydrogenase (E1p) (*Azotobacter vinelandii* en *Escherichia coli*), transketolase (TK) (*Saccharomyces cerevisiae*), branched-chain oxozuur dehydrogenase (E1b) (*Pseudomonas putida* en de mens) en heterotetrameer E1p (*Bacillus stearothermophilus* en de mens) laat zien dat 36 aminozuren identiek zijn terwijl 32 residuen zeer sterk geconserveerd zijn. Deze residuen zijn gelijkmatig verspreid over de sequentie, alhoewel meerdere tussenruimtes ('deletions') en/of invoegingen ('insertions') geïntroduceerd moesten worden. De meest aannemelijke kandidaten voor ThDP binding en voor de 'active site' aminozuren zijn in dit hoofdstuk geïdentificeerd. Daarnaast is er een gemeenschappelijke functie voorgesteld voor gebieden die aanwezig zijn in zowel TK als in beide soorten 2-oxozuur dehydrogenase. Enkele geconserveerde gebieden blijven over, waarvoor geen duidelijke functie kon worden toegekend op basis van de gemaakte vergelijking tussen de verschillende sequenties. Het zou zeer waardevol zijn om de functie van deze gebieden te bestuderen bijvoorbeeld met behulp van mutagenese experimenten. Daarnaast zijn er verschillen gevonden tussen de manier van binden van homodimeer E1 en van heterotetrameer E1 aan E2p welke oproepen tot vragen en dus nader dienen te worden bestudeerd. Deze nieuwe sequentievergelijking opent, samen met de nu beschikbare drie dimensionale informatie, de mogelijkheid tot het maken van zeer bruikbare model structuren van E1. Speciaal voor homodimeer E1p, welke tot nu toe niet gekristalliseerd kon worden, zal dit van zeer grote waarde zijn.

Als laatste, wordt in hoofdstuk 7 een overzicht gegeven van het structurele grondvlak van de verstoorde of verdwenen functie van humane 2-oxozuur dehydrogenase complexen. Zoals al eerder vermeld, resulteert het slecht of niet functioneren van elk van deze complexen in een brede variatie aan klinische verschijnselen. Slecht functioneren of gebrek aan PDHC leidt met name tot melkzuur acidose gecombineerd met verminderde neurologische functie en/of een vertraagde groei en ontwikkeling. Maple syrup urine ziekte is een aangeboren stofwisselingsstoornis veroorzaakt door het slecht functioneren van het BCDC. Daarnaast zijn er verbanden gevonden tussen de ziekte van Alzheimer en de ziekte van Parkinson en het gen dat OGDC codeert. In dit laatste hoofdstuk wordt de overvloed aan zowel genetische alsook aan structurele informatie die nu beschikbaar is gecombineerd om op deze manier een beter

inzicht te verkrijgen in het structurele grondvlak van de verstoorde functie van de 2-oxozuur dehydrogenase complexen als gevolg van mutaties. De effecten van mutaties in de verschillende componenten van de humane complexen worden verklaard gebaseerd op hun vermoedelijke functie.

Alhoewel dit proefschrift licht werpt op de bindingswijze van homodimeer E1p in het complex en op de structuur van dit enzym, blijven er vele vragen open met betrekking tot dit enzym. Bijvoorbeeld, de verschillen tussen homodimeer en heterotetrameer E1; waarom heeft heterotetrameer E1 geen N-terminaal domein? We hebben, gebaseerd op sequentie vergelijkingen en homologie modellen, de hypothetische aminozuren in de 'active site' geïdentificeerd, maar hun functie moet experimenteel geverifieerd worden; welke aminozuren vormen het lipoamide bindingkanaal; welke residuen binden ThDP en welke aminozuren binden pyruvaat, hetzij in een pre-katalytisch complex, hetzij in een katalytisch complex. Tenslotte hebben onze experimenten het bestaan van verschillende zeer flexibele gebieden in E1p aangetoond. Het verwijderen van deze gebieden zou een belangrijke stap kunnen zijn in de richting van het verkrijgen van geschikte eiwitkristallen. Uiteindelijk zou dit kunnen leiden tot het oplossen van de hoge resolutie structuur van een homodimeer E1, wat weer hele nieuwe wegen zal openen in het onderzoek naar 2-oxozuur dehydrogenase complexen.

Nawoord

Eindelijk: Het is af!!

Eigenlijk zegt dit genoeg, toch wil ik op deze plek iedereen bedanken die heeft bijgedragen aan het tot stand komen van dit proefschrift.

Bedankt allemaal !!

Een paar mensen wil ik speciaal noemen: Aart bedankt voor je interesse, begeleiding en steun in de afgelopen jaren. Adrie, kamer en labgenoot die altijd klaar stond om te helpen: ontzettend bedankt!!
Prof. Veeger en Prof. Laane voor het mogelijk maken van mijn promotie.
En dan natuurlijk alle labgenoten die door de jaren heen biochemie bevolkt hebben; het was altijd enorm gezellig.
Petra, kom op meid nu jij nog!!

Bart, je hebt van al die tekst een boekje gemaakt, dat hebben we mooi voor mekaar!
Mam, zonder al het oppassen was het niet gelukt.

Polleke en Loes, nu gaan we weer wat leuks doen....

Curriculum vitae

Annechien Hengeveld is geboren op 16 september 1969 te Leiden. In 1988 behaalde zij het "highschool" diploma te Burbank, California, USA. In 1989 behaalde zij het VWO diploma aan de Rijksscholen gemeenschap "de Drie Waarden" te Schoonhoven. Aansluitend werd de studie Moleculaire Wetenschappen aangevangen aan de Landbouwuniversiteit te Wageningen. Tijdens de doctoraalfase werden afstudeervakken gevolgd bij de vakgroep Biochemie (Dr. A. de Kok en Prof. C. Veeger: "Studies naar de beweeglijkheid van het lipoyl domein van het dihydrolipoyl acyltransferase van het pyruvaat dehydrogenase multi-enzym complex uit *Azotobacter vinelandii* met behulp van tijdopgeloste fluorescentie"), bij de vakgroep microbiologie (Dr. R. Eggen, Prof. Dr. W. de Vos: "Expressie and mutagenese van het β-glucosidase gen uit *Pyrococcus furiosus* in *E. coli*") en bij het Department of Structural Biology, School of Medicine, University of Washington, Seattle USA (Prof. Dr. W. Hol: "Energetic analysis of CoA bound in two different conformations to the catalytic domain of dihydrolipoyl acyltransferase from *A. vinelandii*" en "Crystallization experiments on different components of the pyruvate dehydrogenase multienzyme complex from *A. vinelandii*"). In maart 1995 werd het doctoraalexamen behaald.

Van juli 1995 tot november 1999 was zij als onderzoeker in opleiding (OIO) verbonden aan de vakgroep Biochemie bij de Universiteit Wageningen. Hier is het onderzoek, zoals beschreven in dit proefschrift, uitgevoerd onder supervisie van Dr. Aart de Kok en Prof. C. Veeger. Prof. C. Laane was tijdens een deel van deze periode hoogleraar bij de vakgroep. Het onderzoek werd gefinancierd door de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO) onder toezicht van de Nederlandse stichting voor Chemische Wetenschappen (CW). In mei 2000 kreeg zij een dochter. Sinds augustus 2001 is zij als post-doctoral fellow werkzaam bij de vakgroep farmacochemie van het Leiden/Amsterdam Centre for Drug Research (LACDR) aan de Vrije Universiteit van Amsterdam (Prof. Dr. R. Leurs).

Publications

Berg, A., Bosma, H., Westphal, A.H., **Hengeveld, A.F.**, Vervoort, J. & De Kok A. (1996) The pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Structure and function of the lipoyl domain, in: Biochemistry and physiology of thiamin diphosphate enzymes (Bisswanger, H. & Schellenberger, A., eds.) pp 278-291, A.u. C.Intemann, Wissenschaftlicher Verlag, Prien, Germany.

Hengeveld, A.F., Westphal, A.H. & De Kok, A. (1997) Expression and characterisation of the homodimeric E1 component of the *Azotobacter vinelandii* pyruvate dehydrogenase complex, Eur. J. Biochem. 250, 260-268.

De Kok, A., **Hengeveld, A.F.**, Martin, A. & Westphal, A.H. (1998) The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria, Biochim. Biophys. Acta. 1385, 353-66.

Hengeveld A.F., Schoustra S.E., Westphal A.H., De Kok A. 1999. Pyruvate dehydrogenase from *Azotobacter vinelandii*. Properties of the N-terminally truncated enzyme. Eur. J. Biochem. 265:1098-1107.

Hengeveld, A.F. & De Kok A. 2002. The structural basis of the dysfunctioning of human 2-oxo acid dehydrogenase complexes, Curr. Med. Chem. In press.

Hengeveld, A.F., Hooven van den, H.W., Mierlo van, C.P.M., Visser, A.J.W.G. & De Kok, A. Functional structural characterisation of the N-terminal domain of pyruvate dehydrogenase, submitted to Biochemistry.

Hengeveld, A.F. & De Kok, A. Identification of the E2-binding residues in the N-terminal domain of prokaryotic pyruvate dehydrogenase. Submitted

Hengeveld, Annechien Frederique
Pyruvate dehydrogenase: Its structure, function and interactions within the pyruvate dehydrogenase multienzyme comple
Thesis Wageningen. – With ref. – With summary in Dutch
ISBN: 90-5808-620-8

The research described in this thesis was carried out at the Department of Biochemistry of Wageningen University, The Netherlands.
The investigations were supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organisation for Scientific Research (NWO).