STRUCTURAL STUDIES ON DIHYDROLIPOYL TRANSACETYLASE

the core component of the pyruvate dehydrogenase complex of *Azotobacter vinelandii*

Promotor: dr. C. Veeger, hoogleraar in de Biochemie

Co-promotor: dr. A. de Kok

J. R. O. Hanemaaijer

STRUCTURAL STUDIES ON DIHYDROLIPOYL TRANSACETYLASE

the core component of the pyruvate dehydrogenase complex of *Azotobacter vinelandii*

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN
DOCTOR IN DE LANDBOUWWETENSCHAPPEN
OP GEZAG VAN DE RECTOR MAGNIFICUS,
DR. H. C. VAN DER PLAS,
IN HET OPENBAAR TE VERDEDIGEN
OP WOENSDAG 5 OKTOBER 1988
DES NAMIDDAGS TE VIER UUR IN DE AULA
VAN DE LANDBOUWUNIVERSITEIT TE WAGENINGEN

ISN 4347325

Aan Marja
en mijn ouders

Stellingen

1.
De konklusie dat de E_1 - en de E_3 komponenten van het pyruvaat dehydrogenase complex uit E.coli op verschillende plaatsen op de E_2-keten binden, is niet af te leiden uit de beschreven experimenten.

L.C. Packman & R.N. Perham (1986), FEBS Lett. 206, 193-198

2.
De veronderstelling dat de drie lipoyldomeinen in E_2 van het pyruvaat dehydrogenase complex uit E.coli onafhankelijke mobiliteit bezitten, lijkt tegengesproken te worden door resultaten uit fluorescentiemetingen.

Dit proefschrift

3.
De concentratie van een eiwit met een afwijkend gehalte aan tryptofaan, tyrosine of arginine, of van een eiwit waar de aminozuursamenstelling niet van bekend is, kan noch met de methode van Lowry, noch met de Coomassie brilliant blue methode nauwkeurig bepaald worden.

O.H. Lowry, N.J. Rosebrough, A.L. Farr & R.J. Randall (1951) J.Biol.Chem. 193, 265-275
S.J. Compton & C.G. Jones (1985), Anal.Biochem. 151, 369-374

4.
Het toenemend gebruik in wetenschappelijke terminologie van het voorvoegsel bio-, waar pseudo- bedoeld wordt, kan ten koste gaan van de waardering van het vakgebied biochemie.

5.
De konklusie van Makinen & Makinen, dat modifikatie van een tyrosyl residue ten grondslag ligt aan de inaktivatie van collagenase door o-cloranil, is voorbarig en niet uit de beschreven resultaten af te leiden.

P-L Makinen & K.K. Makinen (1988)
Biochem.Biophys.Res.Commun. 153, 74-80

6.
De konstatering van Yang et al. dat STEM-metingen een uit 24 subunits bestaande PDC-core ondersteunen, is onjuist.

H. Yang, J.F. Hainfeld, J.S. Wall & P.A. Frey (1985) J.Biol.Chem. 260, 16049-16052

7.
De kwaliteit van de beoordeling van wetenschappelijke publikaties is niet gediend met een uitbreiding van het forum met goochelaars, journalisten en Robin Hood's.

E.Davenas et al. (1988), Nature 333, 816-818
J.Maddox J. Randi & W.W.Stewart (1988), Nature 334, 287-290
J. Benveniste (1988), Nature 334, 291

8.
Het verwijzen van ethische vraagstukken, bijvoorbeeld die betreffende de prenatale diagnostiek, naar de aktuele fatsoensopvattingen van de maatschappij biedt weinig garantie voor de zwakken in de samenleving.
["De polis van de partikuliere verzekeraar is net zo fatsoenlijk als de samenleving zelf wenst te zijn", Dr. R Scheerens, direkteur KLOZ, Volkskrant 14-11-1987]

R.J. Lifton (1986), The Nazi Doctors, Basic Books, Inc., Publishers, New York

9.
Een titel stelt geen jota voor

10.
Passende kleding zit niet lekker.

Roeland Hanemaaijer

"Structural studies on dihydrolipoyl transacetylase"

Wageningen, 5 oktober 1988

Voorwoord

In tegenstelling tot wat de voorkant van dit proefschrift doet vermoeden is het werk dat in dit boekje beschreven staat niet van een persoon afkomstig, maar een samenvoeging van ideeen en bijdragen van meerdere mensen.
In de eerste plaats wil ik Aart de Kok noemen en bedanken. Aart, je direkte betrokkenheid en stimulerende invloed zijn op veel plaatsen in dit proefschrift terug te vinden.
Adrie Westphal heeft mij wegwijs gemaakt in het PDC gebeuren en in het kloneringswerk. Hij heeft in de beginfase en in de eindfase aan het onderzoek meegewerkt en tussen deze periodes in is hij er steeds direkt bij betrokken gebleven. Adrie, bedankt!
Ook mijn promotor, prof. Cees Veeger ben ik dankbaar voor de waardevolle suggesties en de kritische op- en aanmerkingen.
Jacques Vervoort, Carlo van Mierlo, Ton Visser en Walter van Dongen zijn direkt betrokken geweest bij een aantal experimenten.
Adrie, Jac, Egbert, Jack, Walter, Willy en Anita zorgden voor een prima werksfeer, waarbij je soms blij was dat je op het lab aan 't werk was, maar je het af en toe ook heerlijk vond om op je kamer te kunnen werken.
Hans Wassink wil ik bedanken voor het in meer of mindere mate spontane uitlenen van biochemicalien in financieel krappe tijden.
Hilda Toussaint, Tini van der Heiden, Felix Kormelink, Paul Corstjens, Remco Masurel, Anja Janssen en Axel Berg hebben als doctoraalstudent een niet gering aandeel in het onderzoek gehad.
Jenny Toppenberg-Fang en Ivonne Soekhram hebben ale verbeteringen en veranderingen in de tekst verwerkt. Martin Bouwmans heeft al het tekenwerk verzorgd. De mensen van de fotolokatie van het Biotechnion hebben alle foto's verzorgd.
Henk Bak van de Rijksuniversiteit Groningen en vooral Wim Roelofsen van de vakgroep Microbiologie hebben tal van aminozuuranalyses uitgevoerd. Jan van Breemen van de Rijksuniversiteit Groningen heeft de elektronenmicroscopische opnames gemaakt en prof. de Kruijff van de Rijksuniversiteit Utrecht stelde ons in de mogelijkheid om de lichtverstrooings-experimenten uit te voeren.
I like to thank Jacqueline Jolles of the University of Paris V for the determination of the N-terminal sequences of the catalytic domain and of E2.
Niet alle bijdragen waren van wetenschappelijke aard. Ik wil mijn vrienden en kennisen bedanken voor het kompenseren van het fundamenteel wetenschappelijke met zoveel wat daarbuiten is.

Hoewel als laatste genoemd, zeker niet in de laatste plaats wil ik mijn ouders en Marja bedanken, resp. voor het stimuleren en het relativeren.

The research reported in this thesis was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid from the Netherlands Organization for Scientific Research

Contents

List of abbreviations

ATCC	american type of culture collection
bp	basepairs
BCOADC	branched-chain 2-oxoacid dehydrogenase complex
CoA	coenzyme A
E_1	pyruvate dehydrogenase
E_2	dihydrolipoyl transacetylase
E_3	lipoamide dehydrogenase
EDTA	ethylenediamine tetraacetate
FAD	flavine adenine dinucleotide, oxidized form
GdnHCl	guanidine hydrochloride
IPTG	isopropoyl-β-D-thiogalactoside
kbp	kilo basepairs
LGT	low-gelling temperature
M_r	relative molecular mass
NAD^+	nicotinamide adenine dinucleotide, oxidized form
NADH	nicotinamide adenine dinucleotide, reduced form
OGDC	2-oxoglutarate dehydrogenase complex
PAGE	polyacrylamide gelelectrophoresis
PDC	pyruvate dehydrogenase complex
PEG	poly(ethyleneglycol)
$PhMeSO_2F$	phenylmethylsulphonyl fluoride
$s_{20,w}$	sedimentation coefficient in water at 293 K
SDS	sodium dodecylsulphate
TNBS	trinitrobenzene sulfonic acid
TPP	thiamine diphosphate
Tris	tris(hydroxymethyl)aminomethane
U	unit of enzyme activity (1 U = 16.6 nkatal)
X-gal	5-bromo-4chloro-3-indolyl-β-D-galactopyranoside

List of enzymes

EC number	Systematic name	Trivial name
1.2.4.1	pyruvate:lipoate oxidoreductase	pyruvate dehydrogenase
	oxidoreductase	
2.3.1.21	acetylCoA:dihydrolipoamide	dihydrolipoyl transacetylase
	S-acetyltransferase	
1.8.1.4	NADH:lipoamide oxidoreductase	lipoamide dehydrogenase

CHAPTER 1

INTRODUCTION

INTRODUCTION

Basic structure and mechanism

The pyruvate dehydrogenase complex is a multi-enzyme complex which catalyzes the oxidative decarboxylation of pyruvate to acetylCoA. It is built from multiple copies of three enzymes: pyruvate dehydrogenase (E_1), dihydrolipoyl transacetylase (E_2) and lipoamide dehydrogenase (E_3) [1]. The mechanism of reaction is shown schematically in Fig. 1. E_1, which requires TPP as an essential cofactor, catalyzes the decarboxylation of pyruvate with formation of an hydroxyethyl intermediate. E_1 then catalyzes the reductive acetylation of the lipoic acid cofactor of E_2. This lipoyl group is covalently bound to the E_2 core through an amide linkage with a lysine ϵ-aminogroup [2]. The acetylgroup is transferred by E_2 to CoA. The reduced lipoic acid cofactor is reoxidized by E_3, which via FAD and a redox-active S-S bridge transfers the reduction equivalents to NAD^+.
PDC is closely related to two other 2-oxoacid dehydrogenase multi-enzyme complexes: the 2-oxoglutarate dehydrogenase complex (OGDC) of the tricarboxylic acid cycle, and the branched-chain 2-oxoacid dehydrogenase complex (BCOADC) of the common pathway for the catabolism of the essential branched-chain amino acids

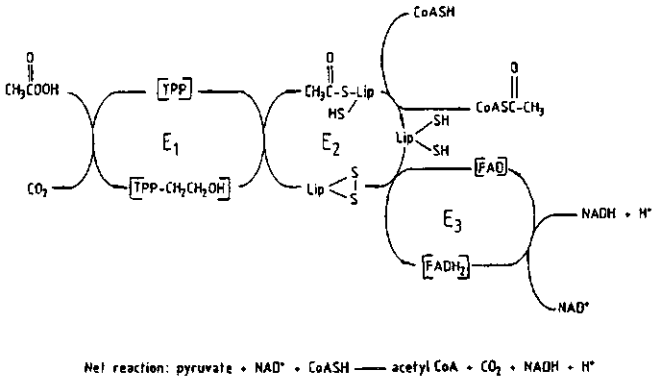


Figure 1. Reaction mechanism of the pyruvate dehydrogenase complex.

[3](Fig. 2). These complexes are built from similar components, 2-oxoacid dehydrogenase (E1), acyltransferase (E_2) and lipoamide dehydrogenase (E_3). E_3 is apparently common to all three complexes [4], whereas E_1 and E_2 are unique to each complex. The basis of overall structure of the three complexes is similar, with E_2 forming a structural, symmetrical core to which the E_1 and E_3 components are tightly but non-covalently bound [5]. OGDC is isolated from Gram-negative bacteria [6,7] and mammals [8]. It is composed of a core of 24 E_2 subunits, arranged in a cubic structure with 432 symmetry. The same structure is found for BCOADC from mammals [9]. The quaternary structure is most variable in

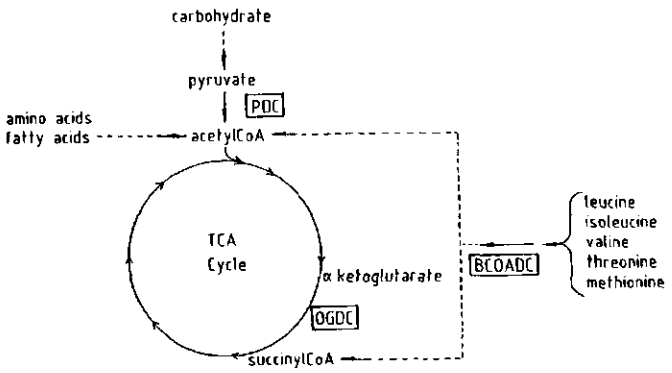


Figure 2. Outline of the metabolic role of the 2-oxoacid dehydrogenase complexes. PDC = pyruvate dehydrogenase complex; OGDC = 2-oxoglutarate dehydrogenase complex; BCOADC = branched-chain 2-oxoacid dehydrogenase complex; TCA cycle = tricarboxylic acid cycle. The broken lines indicate that several steps are involved.

PDC. The E_2 core of PDC from the Gram-negative bacterium E.coli is also composed of 24 subunits [6,10], but the E_2 core of PDC from mammalian sources [5], yeast [11] or Gram-positive bacteria [12] consists of 60 subunits, showing icosahedral symmetry. In the gram negative A.vinelandii PDC is based on a tetrameric E_2 core [13]. This core, upon removal of the periphal components, associates to an E.coli-like structure, as observed on electronmicrographs and in the analytical ultracentrifuge [14].

In all organisms studied E_1 and E_3 are bound as dimers to the E_2 core [5,11,12]. In Gram-negative bacteria E_1 is identified as a

homodimer [5], in Gram-positive bacteria and mammals as a dimer of non-identical chains [9,12]. E_3 is always identified as a dimer, the same is shown for E_1 from OGDC [15]. Transmission electron microscopy of partly assembled PDC and OGDC from E.coli reveals E_3 as binding on the faces of the cube with E_1 bound along the edges [16]. This has been confirmed by computer modelling [17] and by scanning transmission electron microscopy [18,19].

Regulation

Bacterial PDC's are regulated by feed-back inhibition (NADH, acetylCoA) and by cooperative binding of pyruvate and CoA [20-22]. The sites of allosteric action are located on the E_1-component. It was shown for PDC from Pseudomonas aerigunosa that also NAD^+, which will act on the E_3 component, shows cooperativity towards the enzyme complex [23]. Mammalian PDC has a more sophisticated regulation mechanism. To the E_2 component a kinase is tightly bound and a phosphatase is loosely bound [24,25]. Inactivation of the complex occurs after phosphorylation by kinase of the α-subunit of E_1 [26]. Multiple sites are phosphorylated, although inactivation occurs with phosphorylation of one specific site [27,28]. The role of the additional sites are unclear yet. The inactivation is reversible. After action of the phosphatase, activity is restored for 100%. The kinase and phosphatase activities are regulated by the present amount of substrates and products, cations (Mg^{2+}, Ca^{2+}) and hormones [29-31]. Recently an additional component of mammalian PDC, protein X, has been identified. It possesses a lipoic acid cofactor which can be acetylated [32]. The role of protein X is unclear yet, but a possible involvement in anchoring the kinase to the PDC core has been suggested [33].

Stoichiometry and size

The best studied PDC is that of E.coli. However, subunit structure and molecular weight have been subject to considerable controversy. Reconstitution experiments by the group of Reed [10] showed a catalytic optimum stoichiometry of 24:24:12 (E_1:E_2:E_3). By the group of Perham [34] in reconstitution experiments a

stoichiometry of 38:24:19 is reported, which agrees well with direct chain stoichiometry determinations of 28-36:24:14-20 [35,36]. These ratios have been recalculated [37] from the original values using the amino acid compositions and molecular masses obtained from the DNA-sequences of the structural genes for E_1, E_2 and E_3 [38-40]. Recent reconstitution experiments by De Kok et al. [41] confirmed the catalytic optimum stoichiometry of 24:24:12. In reconstitution experiments using analytical ultracentrifugation saturation at an E_1/E_2 ratio of 2.0 and an E_3/E_2 ratio of 2.5 was observed [41]. Obviously, more E_1 and E_3 can bind without effecting the activity. Therefore it is concluded that the complex is not an exact, homogeneous structure. This is in agreement with the determination of the molecular mass of the whole complex, ranging from 3.8-6.1 MDa [42-45]. PDC from A.vinelandii shows a catalytic optimum at a stoichiometry of 6:4:2 [13]. It is the smallest complex isolated, Mr ± 860 kDa [14]. Because of their large structural cores complexes from other sources are large enzymes with M_r 5-10x10^6. On electron micrographs they can easily be seen as particles having a diameter of 20-45 nm [12,16]. This means that they are larger than ribosomes. There are reports [46-48] that the isolated E.coli PDC is not homogeneous, but that in addition to the main component with $s_{20,w}$ = 60S it contains a smaller subcomplex, $s_{20,w}$ = 17S, which is enzymatically active. Also in pigeon breast muscle in the presence of Triton a 1000 kDa active PDC has been determined [49], and in mammalian PDC in the presence of Triton an equilibrium between a small enzymatically active particle of 1000-3000 kDa and the regular PDC of 8000 kDa [50] has been observed [51]. After immunolabelling of innermembrane vesicles no large PDC was found. The same result was obtained after immunolabelling of permeabilized mitochondria. These studies suggest that in situ and probably in vivo, PDC is not present as the large molecule which is found after isolation of the complex.

The E_2 component of PDC

Primary structure and the lipoic acid cofactor

In A.vinelandii 2-2.5 lipoyl groups per E_2 chain were determined by ^{14}C-acetyl incorporation [52], NEM-labelling [53] and by modification with TNBS [14]. E.coli E_2 possesses two lipoyl groups per chain after labelling of the complex with substrates [54-56] and by direct measurements by means of mass spectrometry [57], although the possibility of three lipoyl groups per E_2 chain, found after growing E.coli on (^{35}S)-sulphate, could not be excluded [57]. Since only one lipoyl-lysine residue containing sequence could be identified, the presence of repeating sequences in the primary structure of E_2 was suggested [58]. Such an organization of the E_2 chain became evident from the primary structure, inferred from the determination of the DNA-sequence of the structural gene encoding E_2 [39]; three highly homologous regions of sequence, comprising about 100 amino acid residues each, are repeated in the N-terminal half of the protein (Fig. 3). A potential lipoylation site is present in each repeat. All three sites in E_2 are at least partly lipoylated [37] and each lipoyl group can be reductively acetylated by pyruvate.

Three lipoyl groups per E_2 chain of PDC is specific for Gram-negative bacteria. E_2 chains from PDC from mitochondria or Gram-positive bacteria contain only one lipoyl group [51,59-61]. Also OGDC's possess only one lipoyl group per E_2 chain [57,62]. The three repeats in E_2 of PDC from E.coli can be obtained by limited proteolysis (see below) as functional entities [37], indicating an independent folding. Why the E_2 chains from E.coli and A.vinelandii possess a three-fold repeat of lipoyl domains is obscure. It is shown that not all lipoyl domains in E.coli are required for activity. A part of them can be chemically modified or excised enzymatically without a corresponding loss of complex activity [63-66]. With genetic engineering experiments, in which deletions were made of one or two of the repeating units in the gene encoding E.coli E_2 [67,68], a fully active complex was obtained, suggesting that the extra two lipoyl domains can be regarded as surplus.

Domain structure

The tertiary structure of E_2 contains specific features to locate the lipoyl groups, the transacetylase active site, the binding sites for the E_1 and E_3 components and the E_2 intersubunit binding sites. The E_2 components of all 2-oxoacid dehydrogenase complexes are exceptionally sensitive to proteolytic cleavage under non-denaturing conditions [59,62,63,69-71]. After limited proteolysis usually a domain bearing the lipoyl group(s) is readily cleaved off, suggesting that this domain distinguishes itself physically from a structural core. In E.coli the lipoyl domain can be proteolysed into three stable domains, resembling the single stable domain obtained after proteolysis of E_2 from Gram-positive bacteria or mitochondria [37]. In the structural core domain the quaternary structure of intact E_2 and the transacetylase active site are retained [62,63]. The binding sites for the E_1 and E_3 components are retained [63,72] or lost [70,73-75] during proteolysis, depending on the source of the E_2 and on the conditions of proteolysis. After limited proteolysis of E.coli E_2 a core domain is found which is still able to bind E_1 and E_3 [63,76]. Under harsher conditions a ± 6 kDa smaller domain is found (called catalytic domain), which no longer contains the E_3 binding site [74]. The E_1 binding site has been claimed to be retained on the catalytic domain, although limited proteolysis experiments with PDC and OGDC from other organisms suggest that the E_1 and E_3 binding sites are located in the same regions [70,73].

The region between the lipoyl domain and the catalytic domain is easily proteolysed. It contains many charged residues and is therefore indicated as a hinge region. From limited proteolysis experiments and the primary structure a model for the structure of the E_2 chain can be derived as shown in Fig. 3. On electron micrographs the structural core shows a similar, but somewhat clearer cubic appearance as intact E_2 [63]. Obviously, in electron micrographs of E_2 the lipoyl domain, accounting for about half of the E_2 chain, is only visible as a diffuse appearance around the cubic structure.

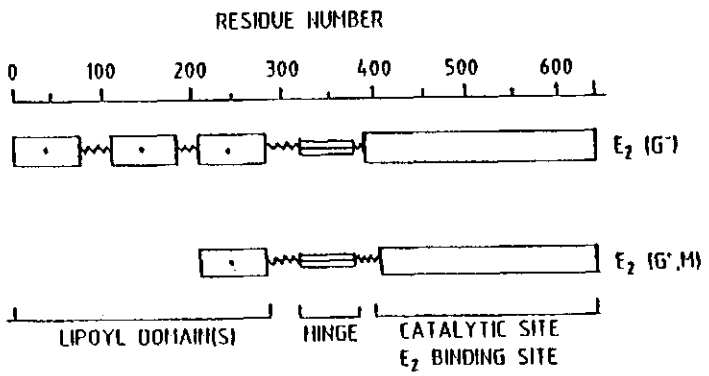


Figure 3. A model for the structure of the E_2 chain from PDC. E_2(G-)= E_2 chain from Gram-negative bacteria: E_2(G+,M)= E_2 chain from Gram-positive bacteria and mitochondria. It also represents the E_2-chain from OGDC. The site of lipoylation is indicated by o. represents a part of the E_2 chain rich in alanine, proline and charged amino acids. The residue numbering is based on data from E.coli.

Mobility and active-site coupling

The extent of reductive acetylation of the reconstructed complex with one lipoyl domain is 40-50% below the expected value, indicating that only 50% of the sites are lipoylated [64-66], while still full activity was present. In Bacillus staerothermophilus PDC, containing only one lipoyl domain, about half of the lipoyl domains could be removed from E_2 by chymotryptic digestion, without loss of activity [60]. These results can be explained in the light of active site coupling: the lipoyl groups of different E_2 chains can interact with each other via an extensive network of intramolecular coupling reactions that permits the transfer of acetyl groups between E_2 subunits within the enzyme core [55,56,77]. In addition, the active site of an E_1 subunit can be visited by more than one lipoyl group [65,66,78]. The lipoyl group, shuttling the substrates between the different active sites in the complex, is thought to rotate as a 'swinging arm' between the active sites [79]. After labelling the lipoyl group with spinlabels indeed mobility was found [80,81]. However, in fluorescence experiments no energy transfer between the labelled active sites of the different components was found, from

which it was concluded that the active sites of the different components are at least 4 nm apart in the complex [82-84], too far away for a single lipoyl-lysine group. Both the large distance and the active site coupling involve movement of the region of the E_2 chain around the lipoyl group. Strong evidence in favor of the existence of conformational mobility in the E_2 chain has been obtained from ^{1}H-NMR spectroscopy [85,86]. The ^{1}H-NMR-spectrum of the intact complex features some sharp signals characteristic of mobile polypeptide chain [85]. A major source of the unusually sharp signal is thought to be the three long alanine+proline-rich regions which link the lipoyl domains to each other and to the E_3 binding domain [37,87-89]. In addition a second sharp signal is found, probably arising from a smaller alanine+proline-rich region, located at the N-terminus of the catalytic domain [87,89]. Because the E_3 binding site is located in the region between the lipoyl domain and the catalytic (core) domain [74,75], it is suggested [89,90] that, if this alanine-proline segment is conformationally mobile, this could explain the unexpected mobility of E_3, bound to the E_2 core [91]. Because of the limited solubility of E.coli E_2 no ^{1}H-NMR spectrum of isolated E_2 is known.

A.vinelandii PDC

The A.vinelandii PDC is the smallest PDC which is isolated. Its sedimentation coefficient is 17-19S [14,92], and its molecular mass is about 850 kDa. Previously it was isolated as a four-component complex [92,93]. The function of the fourth component (63 kDa on SDS-gel) was obscure. It has been shown that it could act as an additional transacetylase [94], may be analogous to the protein X, which is found as a fourth component in mammals [32,33]. The fourth component of A.vinelandii PDC can be resolved from the complex without loss in activity (6U/mg)[93], and via a modification in the isolation procedure a three component complex is isolated with a high specific activity (15-19U/mg)[52].
The complex can be resolved into its components by thiol-Sepharose chromatography [95]. After reconstitution the overall-activity is restored for 100% [96]. After resolution the tetrameric E_2 core associates to a multimeric structure,

analogous to the E.coli structure [14]. Based on the molecular mass of the large structure (1.9-2.1 MDa), the cubic appearance on electronmicrographs, and the E.coli model of a cube composed of eight morphological subunits, a 32-meric structure was concluded [14], built from eight tetramers. Within the tetrameric E_2 core active-site coupling is observed, but, in contrast to E.coli, no coupling is observed within the large aggregated core [52]. Upon addition of the peripheral components the 32-mer dissociates again into tetramers. Besides aggregation of the isolated E_2 core also aggregation of the whole complex is found. In the presence of polyethylene glycol and $MgCl_2$ the A.vinelandii complex associates to a well defined, E.coli-like, polymer with a sedimentation coefficient of 56-60S [97]. From pressure-dependent light-scattering studies [42] an association from tetramer into 32-mer is proposed, although from the results an association into 24-mer cannot be excluded. In reconstitution experiments [96] active hybrid complexes can be obtained from either an E_1E_2 subcomplex from A.vinelandii and the E_3 component from E.coli or from an E_2E_3 subcomplex from E.coli and the E_1 component from A.vinelandii, showing that the binding sites on E_2 for the peripheral components are largely conserved.
Obviously, many similarities exist between the A.vinelandii and the E.coli complexes, although differences are present like molecular mass, stoichiometry and active site coupling. Since the similarities and differences are mainly concerned within the E_2 component, we focussed our attention to this component of the pyruvate dehydrogenase complex from A.vinelandii.
In this thesis the domain structure of A.vinelandii E_2 is analysed by limited proteolysis, the primary strucure is determined by cloning and sequencing the gene encoding E_2. A production clone is obtained of E_2 and of the two main domains, yielding a 15-fold higher expression of E_2 with a two-fold higher specific activity as found in A.vinelandii itself. The mobility of the lipoyl domain is analysed by time-resolved fluorescence anisotropy decay. Because of the high solubility of E_2 and the catalytic domain we were able to obtain 600 MHz ^{1}H-NMR spectra of OGDC, PDC, E_2 and the catalytic domain. From hydrodynamic studies on E2 and the catalytic domain it is shown that isolated

A.vinelandii E_2 consists of 24 subunits. A model is presented in which the A.vinelandii complex can be regarded as a morphological subunit of the cubic pyruvate dehydrogenase complex from E.coli. This model may be universal for all the related 24-meric 2-oxoacid dehydrogenase complexes.

References

1. Reed, L.J. & Cox, D.J. (1966) Annual Reviews Biochem. 35, 57.
2. Nawa, H., Brady, W.T., Koike, M. & Reed, L.J. (1960) J.Am.Chem.Soc. 82, 896.
3. Martin, R.R., Marshall, V.P., Sokatch, J.R. & Unger, L. (1973) J.Bacteriol. 115, 198.
4. Reed, L.J., Pettit, F.H., Yeaman, S.J. & Teague, W.M. (1980) in "Enzyme Regulation and Mechanism of Action" (Mildner, P. & Ries, B. eds.) pp. 47-56, Pergamon Press Ltd., Oxford.
5. Reed, L.J. (1974) Acc.Chem.Res. 7, 40.
6. Reed, L.J. & Oliver, R.M. (1968) Brookhaven Symp.Biol. 21, 397.
7. Pettit, F.H., Hamilton, L., Munk, P., Namihira, G., Eley, M.H., Wilms, C.R. & Reed, L.J. (1973) J.Biol.Chem. 248, 5282.
8. Koike, M., Hamada, M., Koike, K., Hiraoka, T. & Nakaula, Y. (1976) In "Thiamine" (Gubler, C.J., Fujiwara, M & Dreyfus, P.M., eds.) pp. 5-18. John Wiley, New York.
9. Reed, L.J., Damuni, Z. & Merryfield, M.L. (1985) Curr.Top.Cell.Regul. 27, 41.
10. Eley, M.H., Namihira, G., Hamilton, L., Munk, P. & Reed, L.J. (1972) Arch.Biochem.Biophys. 152, 655.
11. Keha, E.E., Ronft, H. & Kresze, G.-B. (1982) FEBS Lett. 145, 289.
12. Henderson, C.E., Perham, R.N. & Finch, J.T. (1979) Cell 17, 85.
13. Bosma, H.J., Westphal, A.H., de Kok, A. & Veeger, C. (1984) Eur.J.Biochem. 142, 541.
14. Bosma, H.J., de Kok, A., van Markwijk, B.W. & Veeger, C. (1984) Eur.J.Biochem. 140, 273.
15. Stepp, L.R., Pettit, F.H., Yeaman, S.J. & Reed, L.J. (1983) J.Biol.Chem. 258, 9454.

16. Oliver, R.M. & Reed, L.J. (1982) in"Electron Microscopy of Proteins" (Harris, R., ed.), vol. 1, pp. 1-48, Academic Press, London.
17. Hackert, M.L., Oliver, R.M. & Reed, L.J. (1983) Proc.Natl.Acad.Sci.USA 80, 2907.
18. Yang, H., Hainfield, J.F., Wall, J.S. & Frey, P.A. (1985) J.Biol.Chem. 260, 16049.
19. Wagenknecht, T., Francis, N., DeRosier, D.J., Hainfield, J.F. & Wall, J.S. (1987) J.Chem.Biol. 262, 877.
20. Schwartz, E.R. & Reed, L.J. (1970) Biochemistry 9, 1434.
21. Bisswanger, H. & Henning, U. (1971) Eur.J.Biochem. 24, 376.
22. Bisswanger. H. (1974) Eur.J.Biochem. 48, 377.
23. Ghosh, R., Guest, J.R. & Jeyaseelan, K. (1981) Biochim.Biophys.Acta 658, 232.
24. Linn, T.C., Pelley, J.W., Hucho, F. & Reed, L.J. (1972) Arch.Biochem.Biophys. 148, 327.
25. Barrera, C.R., Namihira, G., Hamilton, R., Munk, P., Eley, M.H., Linn, T.C. & Reed, L.J. (1972) Arch.Biochem.Biophys. 148, 343.
26. Roche, T.E. & Reed, L.J. (1972) Biochem.Biophys.Res.Commun. 48, 840.
27. Cook, K.G., Lawson, R. & Yeaman, S.J. (1983) FEBS Lett. 157, 59.
28. Denton, R.M. & McCormack, J.G. (1985) Am.J.Physiol. 249, E543.
29. Denton, R.M., Randle, P.J. & Martin, B.R. (1972) Biochem.J. 128, 161.
30. Randle, P.J. (1978) Trends Biochem.Sci. 3, 217.
31. Block, K.P., Heywood, B.W., Buse, M.G. & Harper, A.E. (1985) Biochem.J. 232, 593.
32. Hodgson, J.A., De Marcucci, O.G. & Lindsay, J.G. (1986) Eur.J.Biochem. 158, 595.
33. Rahmatullah, M., Jilka, J.M., Radke, G.A. & Roche, T.E. (1986) J.Biol.Chem. 261, 6515.
34. Bates, D.L., Danson, M.J., Hale, G., Hooper, E.A. & Perham, R.N. (1977) Nature 268, 313.
35. Bates, D.L., Harrison, R.A. & Perham, R.N. (1975) FEBS Lett. 60, 427.
36. Hale, G. & Perham, R.N. (1979) Biochem.J. 177, 129.

37. Packman, L.C., Hale, G. & Perham, R.N. (1984) EMBO J. 3, 1315.
38. Stephens, P.E., Darlison, M.G., Lewis, H.M. & Guest, J.R. (1983) Eur.J.Biochem. 133, 35.
39. Stephens, P.E., Darlison, M.G., Lewis, H.M. & Guest, J.R. (1983) Eur.J.Biochem. 133, 481.
40. Stephens, P.E., Lewis, H.M., Darlison, M.G. & Guest, J.R. (1983) Eur.J.Biochem. 135, 519.
41. De Kok, A., Bosma, H.J., Westphal, A.H. & Veeger, C. (1988) in "Thiamine Pyrophate Biochemistry", Vol. 2, (Schowen, R. & Schellenberger, A., eds.), pp. 19-36, CRC Press, Boca Raton.
42. Danson, M.J., Hale, G., Johnson, P., Perham, R.N., Smith, J. & Spragg, P. (1979) J.Mol.Biol. 129, 603.
43. Anglides, K.J., Akiyama, S.K. & Hammes, G.G. (1979) Proc.Natl.Acad.Sci.USA 75, 4877.
44. Durchschlagg, H. (1975) Mechanism 1, 153.
45. Vogel, O., Hoehn, B. & Henning, U. (1972) Eur.J.Biochem. 30, 354.
46. Schmitt, B. (1975) Biochimie 58, 1405.
47. Schmitt, B. & Cohen, R. (1980) Biochem.Biophys.Res.Commun. 135, 802.
48. Gilbert,G.A. & Gilbert, L.M. (1980) J.Mol.Biol. 144, 405.
49. Forster, M., Schumacher, J., Brouwer, M. & Staib, W. (1985) Hoppe-Seylers Z.Physiol.Chem. 366, 787.
50. Hayakawa, T., Kanazaki, T., Kitamura, T., Fukoyoshi, Y., Sakurai, Y., Koike, K., Suematsu, T. & Koike, M.(1969) J.Biol.Chem. 244, 3660.
51. Sumegi, B., Liposits, Z., Inman, L., Paull, W.K. & Srere, P.A. (1987) Eur.J.Biochem. 169, 223.
52. Bosma, H.J., de Graaf-Hess, A.C., de Kok, A., Veeger, C., Visser, A.J.W.G. & Voordouw, G. (1982) Ann.N.Y.Acad.Sci. 378, 265.
53. De Kok, A., Visser, A.J.W.G. & de Graaf-Hess, A.C. (1982) In "Flavins and Flavoproteins" (Massey, V. & Williams, C.H., eds.) pp. 61-67, Elsevier, Amsterdam.
54. Danson, M.J. & Perham, R.N. (1976) Biochem.J. 159, 677.
55. Bates, D.L., Danson, M.J., Hale, G., Hooper, E.A. & Perham, R.N. (1977) Nature, Lond. 268, 313.
56. Collins, J.H. & Reed, L.J. (1977) Proc.Natl.Acad.Sci.USA 76,

4385.

57. White, R.H., Bleile, D.M. & Reed, L.J. (1980) Biochem.Biophys.Res.Commun. 94, 78.
58. Hale, G. & Perham, R.N. (1980) Biochem.J. 187, 905.
59. Bleile, D.M., Hackert, M.L., Pettit, F.H. & Reed, L.J. (1981) J.Biol.Chem. 256, 514.
60. Duckworth, H.W., Jaenicke, R., Perham, R.N., Wilkie, A.G., Finch, J.T. & Roberts, G.C.K. (1982) J.Biol.Chem. 260, 13779.
61. Hamada, M., Hiraoka, T., Koike, K., Ogasahara, K., Kanzaki, T. & Koike, M. (1979) J.Biochem. 79, 1273.
62. Perham, R.N. & Roberts, G.C.K. (1981) Biochem.J. 199, 733.
63. Bleile, D.M., Munk, P., Oliver, R.M. & Reed, L.J. (1979) Proc.Natl.Acad.Sci.USA 76, 4385.
64. Ambrose-Griffin, M.C., Danson, M.J., Griffin, W.G., Hale, G. & Perham, R.N. (1980) Biochem.J. 187, 393.
65. Stepp, L.R., Bleile, D.M., McRorie, D.K., Pettit, F.H. & Reed, L.J. (1981) Biochemistry 20, 4555.
66. Berman, J.N., Chen, G.-X., Hale, G. & Perham, R.N. (1981) Biochem.J. 199, 513.
67. Guest, J.R., Lewis, H.M., Graham, L.D., Packman, L.C. & Perham, R.N. (1985) J.Mol.Biol. 185, 743.
68. Graham, L.D., Guest, J.R., Miles, J.S., Packman, L.C., Perham, R.N. & Radford, S.E. (1986) Phil.Trans.R.Soc.Lond.A317, 391.
69. Packman, L.C. & Perham, R.N. (1984) Biochem.J. 242, 531.
70. Perham, R.N. & Wilkie, A.O.M. (1980) Biochem.Int. 1, 470.
71. Chuang, D.T., Hu, C.-W.C., Ku, L.S., Markovitz, P.J. & Cox, R.P. (1985) J.Biol.Chem. 260, 13779.
72. Packman, L.C., Perham, R.N. & Roberts, G.C.K. (1984) Biochem.J. 217, 219.
73. Kresze, G.B., Ronft, H. & Dietl, B. (1980) Eur.J.Biochem. 105, 371.
74. Packman, L.C. & Perham, R.N. (1986) FEBS Lett. 306, 193.
75. Packman, L.C. & Perham, R.N. (1987) Biochem.J. 242, 531.
76. Hale, G. & Perham, R.N. (1979) FEBS Lett. 105, 263.
77. Packman, L.C., Stanley, C.J. & Perham, R.N. (1983) Biochem.J. 213, 331.
78. Hackert, M.L., Oliver, R.M. & Reed, L.J. (1983) Proc.Natl.Acad.Sci.USA 80, 2907.

79. Koike, M., Reed, L.J. & Caroll, W.R. (1963) J.Biol.Chem. 235, 1924.
80. Grande, H.J., Bresters, T.W., de Abreu, R.A., de Kok, A. & Veeger, C. (1975) Eur.J.Biochem. 59, 355.
81. Ambrose, M.C. & Perham, R.N. (1976) Biochem.J. 155, 429.
82. Moe, O.A.Jr., Lerner, D.A. & Hammes, G.G. (1974) Biochemistry 13, 2552.
83. Shepherd, G.B. & Hammes, G.G. (1977) Biochemisty 16, 5234.
84. Scouten, W.H., de Graaf-Hess, A.C., de Kok, A., Visser, A.J.W.G. & Veeger, C. (1978) Eur.J.Biochem. 84, 17.
85. Perham, R.N., Duckworth, H.W. & Roberts, G.C.K. (1981) Nature, Lond. 292, 474.
86. Roberts, G.C.K., Duckworth, H.W., Packman, L.C. & Perham, R.N. (1983) Ciba Found.Symp. 93, 47.
87. Spencer, M.E., Darlison, M.G., Stephens, P.E., Duckenfield, I.K. & Guest, J.R. (1984) Eur.J.Biochem. 141, 361.
88. Radford, S.E., Laue, E.D. & Perham, R.N. (1986) Biochem.Soc.Trans. 14, 1231.
89. Radford, S.E., Laue, E.D., Perham, R.N., Miles, J.S. & Guest, J.R. (1987) Biochem.J. 247, 641.
90. Perham, R.N., Packman, L.C. & Radford, S.E. (1988) Biochem.Soc.Symp. 54, 67.
91. Grande, H.J., Visser, A.J.W.G. & Veeger, C. (1980) Eur.J.Biochem. 106, 361.
92. Bresters, T.W., De Abreu, R.A., de Kok, A., Visser, J. & Veeger, C. (1975) Eur.J.Biochem. 59, 335.
93. De Abreu, R.A., de Kok, A. & Veeger, C. (1977) FEBS Lett. 82, 99.
94. De Abreu, R.A., de Kok, A., de Graaf-Hess, A.C. & Veeger, C. (1977) Eur.J.Biochem. 81, 357.
95. De Graaf-Hess, A.C. & de Kok, A. (1982) FEBS Lett. 143, 261.
96. De Kok, A. & Westphal, A.H. (1985) Eur.J.Biochem. 152, 35.
97. Bosma, H.J., Voordouw, G., de Kok, A. & Veeger, C. (1980) FEBS Lett. 120, 179.

CHAPTER 2

THE DOMAIN STRUCTURE OF THE DIHYDROLIPOYL TRANSACETYLASE COMPONENT OF THE PYRUVATE DEHYDROGENASE COMPLEX FROM AZOTOBACTER VINELANDII

Roeland Hanemaaijer, Arie de Kok, Jacqueline Jollès and Cees Veeger

Eur. J. Biochem. 169, 245–252 (1987)

The domain structure of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*

Roeland HANEMAAIJER[1], Arie de KOK[1], Jacqueline JOLLES[2] and Cees VEEGER[1]

[1] Department of Biochemistry, Agricultural University, Wageningen
[2] Laboratoire des Protéines, Université de Paris V, Paris

(Received April 2/July 22, 1987) – EJB 87 0309

Limited proteolysis with trypsin has been used to study the domain structure of the dihydrolipoyltransacetylase (E_2) component of the pyruvate dehydrogenase complex of *Azotobacter vinelandii*. Two stable end products were obtained and identified as the N-terminal lipoyl domain and the C-terminal catalytic domain. By performing proteolysis of E_2, which was covalently attached via its lipoyl groups to an activated thiol-Sepharose matrix, a separation was obtained between the catalytic domain and the covalently attached lipoyl domain. The latter was removed from the column after reduction of the S-S bond and purified by ultrafiltration. The lipoyl domain is monomeric with a mass of 32.6 kDa. It is an elongated structure with $f/f_o = 1.62$. Circulair dichroic studies indicates little secondary structure. The catalytic domain is polymeric with $s_{20,w} = 17$ S and mass = 530 kDa. It is a compact structure with $f/f_o = 1.24$ and shows 40% of the secondary structure of E_2. The cubic structure of the native E_2 is retained by this fragment as observed by electron microscopy. Ultracentrifugation in 6 M guanidine hydrochloride in the presence of 2 mM dithiothreitol yields a mass of 15.8 kDa. An N-terminal sequence of 36 amino acids is homologous with residues 370–406 of *Escherichia coli* E_2. The catalytic domain possesses the catalytic site, but in contrast to the *E. coli* subunit binding domain the pyruvate dehydrogenase (E_1) and lipoamide dehydrogenase (E_3) binding sites are lost during proteolysis. From comparison with the *E. coli* E_2 sequence a model is presented in which the several functions, such as lipoyl domain, the E_3 binding site, the catalytic site, the E_2/E_2 interaction sites, and the E_1 binding site, are indicated.

The pyruvate dehydrogenase complex (PDC) catalyzes the oxidative decarboxylation of pyruvate resulting in the production of acetyl-coenzyme A and NADH. In gram-negative bacteria three enzymes take part in this reaction: pyruvate dehydrogenase (E_1), lipoylacetyltransferase (E_2) and lipoamide dehydrogenase (E_3). E_1 and E_3 are non-covalently bound as dimers to a core of E_2 [1]. The structure of PDC from *Azotobacter vinelandii* is based on a tetrameric core of E_2, to which three E_1 dimers and one E_3 dimer are bound [2]. This quaternary structure differs considerably from that of the *Escherichia coli* PDC which is based on a core of 24 E_2 chains arranged in a cubic structure with 432 symmetry [3]. To this core twelve E_1 dimers and six E_3 dimers are bound in the catalytically optimum stoichiometry [4]. Upon removal of the peripheral components, the *A. vinelandii* E_2 associates to a multimer. Although its appearance in the electron microscope is similar to that of *E. coli* E_2, hydrodynamic studies indicate differences in quaternary structure [5, 6]. Experiments with hybrid *E. coli*/*A. vinelandii* PDC complexes showed that the differences in quaternary structure reside in the E_2 component [7].

Limited proteolysis has proved to be useful tool in studying the domain structure of proteins. In the case of multifunctional proteins such as the fatty acid synthetase complex [8] and the enzyme complex for the synthesis of aromatic amino acids [9] structural domains are obtained which catalyze partial reactions. The E_2 chain of PDC comprises many functions: the E_1, E_2 and E_3 intersubunit binding sites, the covalently bound lipoyl groups which are involved in the transport of activated substrates between the three different active site of the complex, and the site for acetyltransferase activity. Bleile et al. [10] and Packman et al. [11] have studied the domain structure of the *E. coli* E_2. Bleile et al. [10] obtained two stable fragments by limited proteolysis with trypsin: an N-terminal lipoyl domain containing the potential lipoyl binding sites, and a C-terminal domain which contains the active site, the E_2 intersubunit binding sites and the binding site for E_1 and E_3. Sequence determination [12] showed that the lipoyl domain consists of three homologous sequences which contain alanyl prolyl-rich regions at their C-termini. Packman et al. [11] showed by digestion with *Staphylococcus aureus* V8 proteinase that these homologous sequences can be considered as separate domains, each containing a potential lipoyl binding site. We will use here the term lipoyl domain to indicate the intact N-terminal fragment.

The C-terminal domain (residues 317–629) is usually called the subunit binding domain [10, 12–14], although other names such as inner core fragment [11], binding and catalytic domain [12] or cat [13, 14] have been used as well.

Correspondence to A. de Kok, Laboratorium voor Biochemie, Landbouwuniversiteit, De Dreijen 11, NL-6703 BC Wageningen, The Netherlands

Abbreviations. Pyruvate dehydrogenase complex, PDC; phenylmethylsulfonyl fluoride, $PhMeSO_2F$.

Enzymes. Pyruvate dehydrogenase (E_1), pyruvate:lipoate oxidoreductase (EC 1.2.4.1); lipoate acetyltransferase (E_2), acetylCoA:dihydrolipoamide *S*-acetyltransferase (EC 2.3.1.21); lipoamide dehydrogenase (E_3) NADH:lipoamide oxidoreductase (EC 1.8.1.4); trypsin (EC 3.4.21.4).

In this paper we will refer to this fragment as the subunit binding domain.

In order to analyse the structural basis for the differences between the E_2 components of *A. vinelandii* and *E. coli* we performed limited proteolysis with trypsin of the *A. vinelandii* E_2 component. Again two stable fragments were isolated, a lipoyl domain and a domain which contained the catalytic activity and the E_2 intersubunit binding sites, which we call the catalytic domain. In contrast to limited proteolysis of *E. coli* E_2, both the E_1 and E_3 binding sites are lost. A comparison with the sequence data of the *E. coli* E_2 was made which resulted in the assignment of the function in the primary structure.

MATERIALS AND METHODS

Materials

Trypsin (treated with tosylphenylalanylchloromethane) was obtained from Worthington Biochemical Corp. Activated thiol-Sepharose 4B was from Pharmacia.

Dihydrolipoamide was synthesized by the method of Reed [15]. Eosin 5-maleimide was obtained from Molecular Probes Inc. (Oregon). All other chemicals used were of analytical grade.

Isolation of the complex and resolution into its components

The pyruvate dehydrogenase complex was isolated from *Azotobacter vinelandii* as described by Bosma et al. [16]. After ethanolamine – Sepharose chromatography the enzyme was concentrated by precipitation with 10% poly(ethyleneglycol) 6000 instead of ultracentrifugation. The complex was resolved into its components by covalent chromatography on thiol-Sepharose 4B as described before [17] with the following modifications. Dissociation of E_1 was performed at pH 9.4 instead of pH 8.8. The E_3 was eluted overnight with 3 M KBr in ethanolamine buffer pH 9.0 to remove small amounts of E_1 still present. The E_2 and E_3 components were dialysed against standard buffer (50 mM potassium phosphate pH 7.0 containing 0.5 mM EDTA and 0.05 mM $PhMeSO_2F$) and used immediately or stored under liquid nitrogen. The E_1 component was dialysed against standard buffer containing 0.1 mM thiamin diphosphate and 1 mM $MgCl_2$, and stored in 20% (v/v) glycerol at −20 °C.

Limited proteolysis

The isolated E_2 component (0.8 – 1.0 mg/ml) was incubated with trypsin (10 μg/ml) in standard buffer without $PhMeSO_2F$ at 0 °C. Samples were withdrawn at timed intervals and added to an equal volume of sample buffer (80 °C) for SDS/polyacrylamide gel electrophoresis.

Isolation of the main fragments obtained by limited proteolysis of E_2

Isolation of the main proteolytic fragments was performed by proteolysis of E_2 which was covalently bound on thiol-Sepharose 4B through covalent linkage of the activated thiol groups with the lipoyl SH groups, which were generated by incubation of the complex with pyruvate, thiamin diphosphate and $MgCl_2$. The E_1 and E_3 components were eluted as described before [17]. After washing with standard buffer without $PhMeSO_2F$ the E_2 component, still bound on the matrix, was digested by incubating with standard buffer containing 10 μg/ml trypsin at 4 °C for 1 h. After washing with standard buffer containing 0.2 mM $PhMeSO_2F$ the catalytic domain fragment was eluted with standard buffer containing 3 M KBr. The lipoyl domain fragment, still covalently attached to the column matrix was eluted like undigested E_2 by standard buffer containing 0.5 M hydroxylamine or 50 mM Tris/HCl pH 8.5 containing 20 mM dithiothreitol. This fraction was separated from large proteins such as the E_3 dimer and undigested E_2 by ultrafiltration using an Amicon YM-100 membrane. The lipoyl domain, which passed the membrane, was concentrated on an Amicon YM-10 membrane. Both fragments were dialysed immediately against standard buffer and stored under liquid nitrogen.

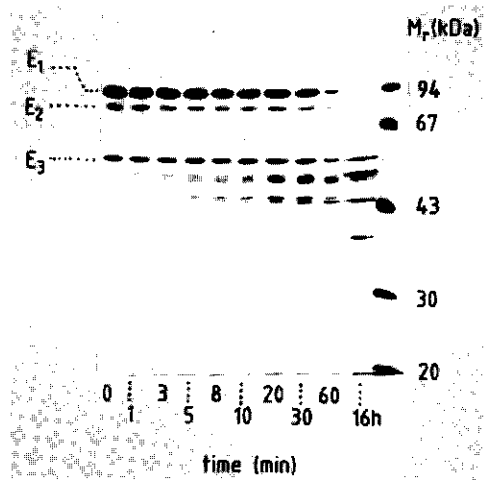


Fig. 1. *Limited proteolysis of pyruvate dehydrogenase complex.* PDC (2 mg/ml) was incubated with trypsin (10 μg/ml) in standard buffer without $PhMeSO_2F$ at 0 °C. Aliquots of 5 μl were taken at the indicated timed intervals, added to an equal volume of sample buffer (80 °C) and analysed by SDS gel electrophoresis

Other techniques

Sedimentation velocity and equilibrium experiments were performed in standard buffer or standard buffer containing 6 M guanidine hydrochloride in an MSE Centriscan 75, equipped with an ultraviolet/visible monochromator. Corrections for the solvent density and viscosity were made according to standard procedures [18]. For the E_2 component, the catalytic and the lipoyl domain, partial specific volumes of 0.740, 0.747 and 0.736 ml/g respectively were calculated from the amino acid composition, according to the method of Cohn and Edsall [19]. For the determination of the molecular mass of the catalytic domain monomer, gel chromatography in 6 M guanidine hydrochloride was carried out, using FPLC equipped with a Superose 12 HR 10/30 column (Pharmacia). The peak fraction was directly used for sedimentation equilibrium experiments.

For amino acid analysis, samples were hydrolysed under nitrogen in sealed tubes in 3 M mercaptoethanesulfonic acid for 72 h at 110 °C [20]. Analyses were performed on a Kontron Liquimat III analyser. Corrections were made for threonine, serine and tyrosine. For protein sequencing the samples were

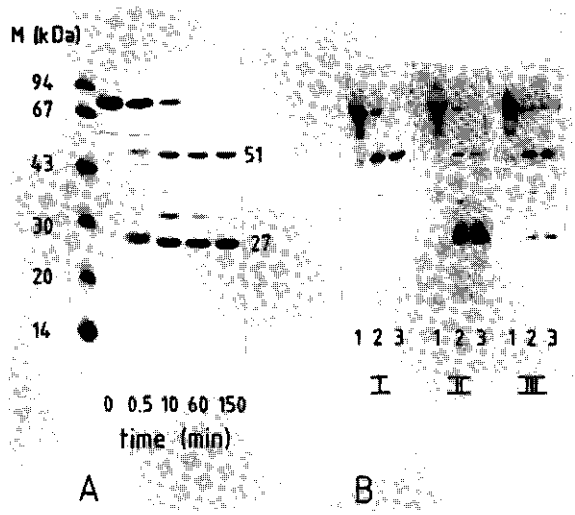


Fig. 2. *Limited proteolysis of dihydrolipoyltransacetylase (E_2).* (A) SDS gel electrophoresis of E_2 and its tryptic digest at timed intervals. Lane 1, markers (mass is indicated in kDa); lanes 2–6, E_2 (1 mg/ml) incubated with trypsin for 0 min (lane 2), 0.5 min (lane 3), 10 min (lane 4), 60 min (lane 5) and 150 min (lane 6). (B) Western blotting of E_2 and its tryptic digest after incubation for 0 min (lanes 1), 10 min (lanes 2) and 90 min (lanes 3). The blots were incubated with antibodies raised against the 51-kDa fragment (I), the 27-kDa fragment (II) and intact E_2 (III)

purified as follows. Gel chromatography was carried out in 6 M guanidine hydrochloride, after which the protein fraction was *S*-carboxymethylated. After dialysis against ammonium carbonate the samples were concentrated by freeze-drying, and gel filtration in 50% acetic acid was carried out.

Protein sequencing was carried out by automated Edman degradation in an Applied Biosystems 470 A gas-phase protein sequencer. The phenylthiohydantoin derivatives were automatically identified with an Applied Biosystems 120 A PTH analyser used on-line with the sequencer.

CD spectra were measured with a Jobin Yvon mark V dichrograph using quartz cells with a 0.1–0.5-mm pathlength in the wavelength region 195–255 nm. The cell holder was thermostatted at 20 °C and the cell compartment purged with nitrogen.

Electron microscopy of E_2 was carried out after cross-linking of the protein by reaction with 2% (w/v) glutaraldehyde in standard buffer, at room temperature for 10 min. The sample was stained with 1% (w/v) uranyl acetate and examined in a Jeol electron microscope operated at 80 kV. The catalytic domain was examined without cross-linking beforehand. Enzyme activities were assayed at 25 °C as described previously [21]. Protein concentrations were calculated from measurements of absorbance at 280 nm. The absorption coefficients $A_{280}^{0.1\%}$ for E_2 and the catalytic domain were 0.45 and 0.93, respectively, as determined by amino acid analysis. Because of the low coefficient for the lipoyl domain ($A_{280}^{0.1\%} < 0.05$) this protein concentration was measured according to the method of Lowry [22]. Gel electrophoresis in the presence of sodium dodecyl sulfate was carried out essentially as described by Laemmli [23].

For the production of antisera, digested and undigested E_2 were subjected to preparative SDS gel electrophoresis. After lightly staining, the protein bands were cut out and protein was extracted from the gel. The protein was dissolved in phosphate-buffered saline pH 7.2 and mixed with Freund's incomplete adjuvants. For immunisation male New Zealand White rabbits were used. After five weeks a booster injection was given and ten days later the antisera were collected. For the immunological detection of proteins the western blotting method was used as described by Voordouw et al. [24], whereby visualization of the bands was performed with alkaline-phosphatase-linked immunodetection (Protoblot Immunoblotting System, Promega Biotec).

RESULTS

Limited proteolysis of PDC and the E_2 core

Incubation of the pyruvate dehdrogenase complex with trypsin at pH 7 and 4 °C resulted in digestion of the E_1 and E_2 components (Fig. 1). The overall activity was lost as well as the E_1 activity, but no loss was observed of dihydrolipoyltransacetylase (E_2) activity, i.e. the acetylation of exogenous dihydrolipoamide, or of lipoamidehydrogenase (E_3) activity.

It is well known that E_2 and its fragments behave anomalously on SDS gels [6, 10]. Therefore the chain masses given below are an indication of relative position rather than presenting true molecular mass data. When trypsin-digested E_2 (mass 83 kDa on SDS gels) was subjected to SDS gel electrophoresis two main products were found: a 51-kDa fragment and a 27-kDa fragment (Fig. 2A). With antibodies raised against the 51-kDa fragment, a 40-kDa minor fragment was identified as a proteolytic degradation product of the 51-kDa fragment. A 34–37-kDa minor fragment (mass varying between isolations) showed cross-reactivity with antibodies raised against the 27-kDa fragment (Fig. 2B). Some cross-reactivity is observed with the 51-kDa fragment. This could be due to contamination of the 27-kDa fragment with degra-

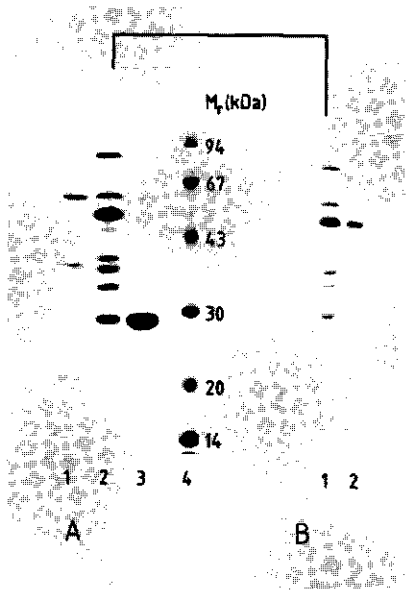


Fig. 3. *Isolation of the proteolytic fragments of E_2.* (A) After incubation of E_2, covalently bound to thiol-Sepharose, with trypsin, fractions were eluted with standard buffer containing 0.2 mM $PhMeSO_2F$ (lane 1), standard buffer containing 3 M KBr (lane 3) or standard buffer containing 0.5 M hydroxylamine (lane 2). (B) Purification of the 51-kDa fragment by ultrafiltration. The fraction containing the 51-kDa fragment (lane 1) was eluted over an YM 100 membrane (lane 2)

dation products originating from the 51-kDa fragment. Alternatively the presence of homologous epitopes (e.g. the Ala – Pro-rich sequence in the N-terminal sequence of the 27-kDa fragment, see Fig. 4B) may be the cause of this weak cross-reactivity. The cross-reactivity of antibodies raised against SDS-denatured E_2 was larger with the 51-kDa fragment than with the 27-kDa fragment. This was confirmed with the ELISA technique. With antibodies raised against native E_2 this difference was even larger. No fragments below 27 kDa could be detected on SDS gels up to 20% acrylamide or by using antibodies raised against intact E_2, the 51-kDa or the 27-kDa fragment.

Proteolysis of E_2 with trypsin after amidination of the lysyl residues (cf. [25]) yielded essentially the same results, which indicates that proteolysis takes place on arginyl residues, or on closely spaced lysyl and arginyl residues. These experiments indicate that the presence of E_3 or E_1 does not protect potential cleavage sites on E_2.

Purification of the proteolytic fragments

The purification of the tryptic fragments was carried out as described in Methods. After elution with buffer containing 3 M KBr, pure 27-kDa fragment was obtained (Fig. 3A, lane 3). After elution with buffer containing hydroxylamine or dithiothreitol a fraction containing the 51-kDa fragment was obtained (Fig. 3A, lane 2, and Fig. 3B, lane a). This fragment was purified by removing undigested E_2 and the 27-kDa fragment, which are polymeric (>0.5 MDa), by ultrafiltration which resulted in a fraction containing the 51-kDa fragment and a trace of its 40-kDa degradation product (Fig. 3B, lane 2).

Characterization of the fragments

The 51-kDa fragment was eluted in the same way as undigested E_2 and should for this reason contain the lipoyl moieties by which E_2 is covalently bound to the matrix. This was confirmed by labelling E_2 with an SH-binding fluorescent probe, eosin 5-maleimide. After trypsinolysis of the labelled E_2 a fluorescent 51-kDa band was observed on the SDS gel. This fragment, which we will call the lipoyl domain in analogy with *E. coli* E_2 [10], has a molecular mass of 32.6 kDa as determined by sedimentation equilibrium analysis. Its sedimentation coefficient $s_{20,w} = 2.2$ S, from which a frictional ratio can be calculated of 1.62. By using analytical ultracentrifugation no binding of E_3 with the lipoyl domain was observed.

The transacetylase activity, which is retained after E_2 digestion, is located on the 27-kDa fragment. When this fragment is titrated with either the E_1 or the E_3 component no significant change in sedimentation coefficient was found. During E_3 titration, when scanned at 456 nm where E_3 is detected, a sedimentation coefficient of 5.9 S was found, belonging to the E_3 dimer. These results show that the 27-kDa fragment possesses no detectable E_1 or E_3 binding sites. In this respect it differs from the subunit-binding domain obtained by limited proteolysis with trypsin of *E. coli* E_2 [10], which has a larger peptide chain (29.6 kDa instead of 15.8 kDa, see below) and still possesses the E_1 and E_3 binding sites. To distinguish the *A. vinelandii* 27-kDa fragment from the *E. coli* subunit-binding domain it will be called the catalytic domain. Also in contrast to the *E. coli* domain, the catalytic domain remains soluble in 50 mM potassium phosphate buffer and is suitable for further characterization such as sedimentation analysis, circular dichroism and electron microscopy. It has a sedimentation coefficient of 16.8 S; for intact E_2 a sedimentation coefficient of 21.8 S was found. By sedimentation equilibrium analysis of the catalytic domain a molecular mass of 530 ± 17 kDa ($n = 3$) was determined (intact E_2 mass = 1.9 MDa); apparently it still possesses the E_2 intersubunit binding sites, resulting in the same quaternary structure as E_2 (see electron microscopy below). From the molecular mass and sedimentation coefficient a frictional ratio of 1.24 was calculated, indicating a compact, globular structure. The mass of the catalytic domain as determined by sedimentation equilibrium analysis in the presence of 6 M guanidine hydrochloride and 2 mM dithiothreitol was found to be 15.8 ± 0.5 kDa ($n = 6$), which agrees with a 32-mer structure as proposed by Bosma et al. [2, 6].

Because the E_1 and E_3 binding sites are present neither on the lipoyl domain nor on the catalytic domain they are lost during proteolysis. This loss is also observed in an experiment where an E_2-E_3 subcomplex, covalently bound on a thiol-Sepharose matrix, is digested with trypsin. After washing with standard buffer more than 75% of the bound E_3 was lost. Apparently bound E_3 does not protect its binding site. In the eluted E_3 fraction no protein fragments could be observed by using SDS gel electrophoresis, probably because they are too small for detection. The results indicate that upon proteolysis of E_2 (mass 63 kDa by sedimentation equilibrium analysis) two large fragments, the lipoyl domain (32.6 kDa) and the catalytic domain (15.8 kDa) are formed, but that in addition

Table 1. *Amino acid compositions of the* A. vinelandii E_2 *component and its proteolytic fragments*
For comparison the amino acid composition of *E. coli* E_2 and its lipoyl domain are given. The *E. coli* amino acid compositions are from Stephens et al. [12]. The tryptophan content of the *A. vinelandii* polypeptides was determined by direct amino acid analysis and by the method of Pajot [27]

Amino acid	E_2		Lipoyl domain		Catalytic domain
	A. vin	*E. coli* 1–629	*A. vin*	*E.coli* 1–316	*A. vin*
	residues/molecule				
Asx	39	50	19	23	10
Thr	20	27	17	11	6
Ser	35	29	24	13	8
Glx	68	73	46	44	14
Pro	53	37	26	22	9
Gly	55	51	26	25	11
Ala	132	96	68	67	23
Cys	4	1	0	0	1
Val	48	68	35	42	10
Met	10	16	4	7	3
Ile	28	45	17	19	8
Leu	53	33	24	9	18
Tyr	6	3	1	0	3
Phe	11	19	2	5	5
Lys	34	53	10	29	9
His	7	5	1	0	3
Arg	23	20	9	0	6
Trp	3	3	0	0	2
Total	629	629	328	316	149
Mass	kDa				
	63	66	32.6	31.6	15.8

small fragments which constitute the E_1 and E_3 binding sites, are lost (total molecular mass 14 kDa).

Amino acid composition of E_2 and its domains

The amino acid compositions of E_2, the lipoyl domain and the catalytic domain are listed in Table 1. The compositions of the *E. coli* E_2 and lipoyl domain as determined from the DNA sequence by Stephens et al. [12] are shown for comparison. Much similarity exists in the compositions of the lipoyl domains of both organisms. Both contain large amounts of prolyl and alanyl residues which indicates a common domain structure. One striking difference is the presence of a single histidyl, a tyrosyl and nine argininyl residues in the *A. vinelandii* lipoyl domain. In the *E. coli* sequence a tyrosyl and a histidyl residue are found 9–11 amino acids C-terminal of the trypsin cleavage site at residue 316 [11]. This result, together with a slight difference in molecular mass of the *A. vinelandii* lipoyl domain (32.6 kDa) and the *E. coli* lipoyl domain (31.6 kDa [10]) indicates that the *A. vinelandii* lipoyl domain extends about 10–25 residues beyond the *E. coli* cleavage site. The small fragmented material of proteolysed *A. vinelandii* E2, compared to the corresponding *E. coli* sequences (residues 329–369 and 519–629) contains many prolyl, alanyl and lysyl residues. In *E. coli* E_2 a hinge region is indicated between the lipoyl domain and the subunit binding domain [10, 12]. This region is rich in prolyl, alanyl and lysyl residues. The difference in arginyl residue content explains why the *E. coli* lipoyl domain, lacking arginyl residues, can hardly be visualized on SDS gels by the Coomassie staining [10, 25], whereas the *A. vinelandii* domain gives a well-coloured band [26].

Table 2. *Cross-reactivity between* A. vinelandii E_2, *lipoyl domain and catalytic domain and* E. coli E_2 *of antibodies raised against* A. vinelandii E_2, *its lipoyl domain and its catalytic domain*

Antibodies to *A. vinelandii*	*A. vinelandii*			*E. coli*
	lipoyl domain	catalytic domain	E_2	
Anti-E_2	++	++	++	++
Anti-(lipoyl domain)	++	–	++	+
Anti-(catalytic domain)	–	++	+	+

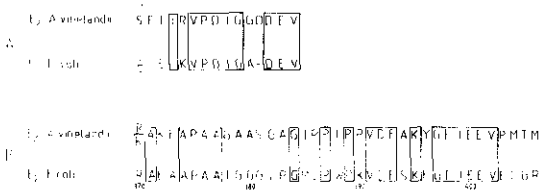


Fig. 4. *Alignment for amino acid sequences of the N-terminal sequence of E_2, the catalytic domain and* E. coli E_2. The sequences, shown in single-letter amino acid code, were aligned for maximum homology. Identical residues are enclosed by solid boxes. For *E. coli* E_2 only the related amino acid sequence is shown. The *E. coli* amino acid sequence is from Stephens et al. [12]

The largest differences in amino acid composition between the three PDC components of *A. vinelandii* and *E. coli* are found in E_2 [6]. The differences in number of prolyl and small-chain residues were considered to produce differences in secondary structure. Nevertheless, antibodies raised against *A. vinelandii* E_2, its lipoyl domain and its catalytic domain show extensive cross-reactivity between the *E. coli* and *A. vinelandii* E_2 (Table 2), whereas with antibodies raised against *A. vinelandii* E_3 no cross-reactivity between the E_3 components of the two organisms as found.

N-terminal amino acid sequences of intact E_2 and the catalytic domain

The N-terminal amino acid sequences of the intact E_2 chain and the catalytic domain were determined by automated Edman degradation. In both cases 1 nmol of the carboxymethylated sample was used. The N-terminal sequence of the intact E_2 chain is shown in Fig. 4A. The homology with the *E. coli* E_2 sequence is obvious. It shows that the lipoyl domain is, like that of *E. coli* E_2, located at the N-terminal part of the E_2 chain.

The catalytic domain fragment shows two N-termini in the ratio 3:2. They show considerable homology with residues 371–406 and 373–406 of the *E. coli* E_2 amino acid sequence as reported by Stephens et al. [12]. The 36-amino-acid-long N-terminal sequence is shown in Fig. 4B. Apparently the two N-termini are the result of random cleavage at two basic residues separated by one amino acid. If the residue at position

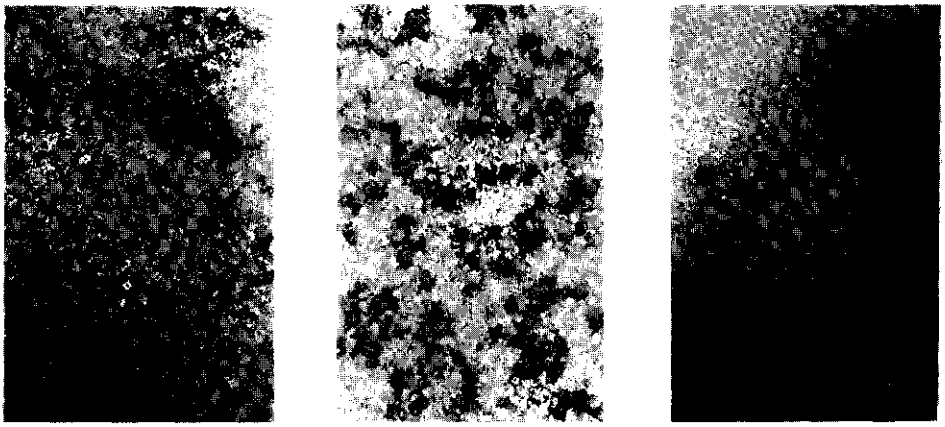


Fig. 5. Electron micrographs of the isolated E_2 (A) and the catalytic domain (B, C). (A) E_2 (0.01 mg/ml) was cross-linked by reaction with 2% glutaraldehyde at room temperature for 10 min and applied on a grid. Uranyl acetate was used as a negative stain. (B) Isolated catalytic domain (0.01 mg/ml), negatively stained with uranyl acetate. (C) Isolated catalytic domain (0.01 mg7ml) applied on a negatively loaded grid. Uranyl acetate was used as a negative stain

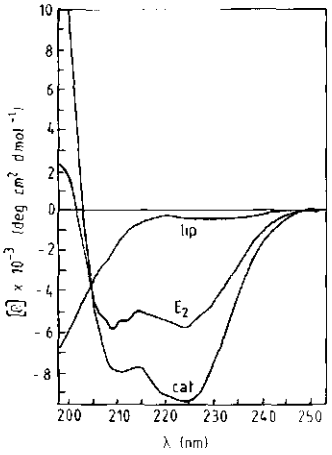


Fig. 6. *Circular dichroic spectra of isolated E_2, lipoyl domain and catalytic domain.* The peptides were dissolved in standard buffer. The cell holder was thermostatted at 20 C. The response time was 10 s

370 is an arginine it may explain why amidination does not effect the proteolysis pattern.

Electron microscopy of intact E_2 and the catalytic domain

We have shown previously [f6] that intact E_2, when freed of E_1 and E_3, aggregates from a tetramer to a large particle of similar sedimentation coefficient but larger mass than that of *E. coli* E_2. For the *A. vinelandii* E_2 rather blurred cubic structures with edges of 13–15 nm are observed by electron microscopy (Fig. 5A). The catalytic domain gives much sharper structures with the same dimensions. Apparently the

Table 3. *Percentages of α-helix, β-sheet and random coil in E_2, the lipoyl domain and the catalytic domain of* A. vinelandii *PDC*

Protein	α-Helix	β-Sheet	Random coil
	%		
E_2	11	37	52
Lipoyl domain	6	24	70
Catalytic domain	13	50	37

information for the quaternary structure is retained in the catalytic domain (Fig. 5B). The lipoyl domain and the E_1 and E_3 binding sites are not observed although they constitute three quarters of the intact E_2 chain. They may be the cause of the vagueness of the cubes, probably because they form open, extended structures, which are not well stained. When a negative loaded grid was used the catalytic domain shows the tendency to form semi-crystalline aggregates (Fig. 5C), which may be used for three-dimensional reconstruction (work in progress).

Circular dichroism

In order to determine the secondary structure of E_2 and its proteolytic fragments CD studies in the peptide region were carried out. The spectra are shown in Fig. 6. From the spectra the percentages of secondary structure were calculated according to the method of Saxena and Wetlaufer [28] (Table 3). The lipoyl domain has relatively little secondary structure. The catalytic domain and the part of the E_2 chain which is proteolysed into small fragments contribute the same amount to the E_2 secondary structure, as calculated from the difference in molar ellipticity between intact E_2 and the isolated fragments.

DISCUSSION

The data presented in this paper show that the *A. vinelandii* transacetylase consists of at least two distinct folding domains: the lipoyl domain and the catalytic domain. Together these domains represent about 80% of the native E_2 chain. The E_1 and E_3 binding sites are lost during proteolysis.

The N-terminal lipoyl domain contains two to three lipoyl moieties [2] and makes up half of the E_2 chain. It has many properties in common with the lipoyl domain of *E. coli* . The high Ala + Pro content is indicative of an analogous structure with three distinct domains [11]. The lipoyl domain constitutes the flexible protuberant structures [30] which gives E_2 its swollen and extended character. This can be concluded from a comparison of the frictional ratios of 2.41 for the polymeric E_2, 1.62 for the monomeric lipoyl domain and 1.24 for the polymeric catalytic domain. The experiments with antibodies against native E_2 are in agreement with this model. Westhof et al. [29] have shown that a good correlation exists between chain flexibility and antigenicity. This may explain the higher cross-reactivity with the lipoyl domain compared to the catalytic domain. It is more difficult to understand why this difference in response, though less pronounced, is also found with antibodies against SDS-denatured E_2. The lipoyl domain contributes little to the secondary structure of E_2. Nevertheless it is well protected against limited proteolysis, despite its many lysyl and arginyl residues. Probably these residues are confined to regions which do possess secondary structure, whereas other parts devoid of these residues, but rich in alanyl and prolyl residues, form coiled regions. Only at longer reaction times or higher trypsin concentration does further degradation takes place.

The catalytic domain, which also possesses the E_2/E_2 interaction sites, forms a compact, globular core with apparent cubic symmetry. In contrast with the *E. coli* subunit binding domain [10] no E_1 and E_3 binding sites are located on this fragment. In this respect it is more comparable to the inner core domain obtained by limited digestion of the *Bacillus stearothermophilus* pyruvate dehydrogenase complex [31] and the branched-chain 2-oxo-acid dehydrogenase complex from bovine liver [32]. It is readily soluble in standard buffer, whereas the *E. coli* subunit binding domain is not. This makes the *A. vinelandii* catalytic domain suitable for accurate characterization.

Size determinations of E_2 and its domains give large deviations depending on the method used. By SDS gel electrophoresis E_2 has an apparent mass of 83 kDa (*E. coli* 82 kDa). By sedimentation equilibrium analysis in 6 M guanidine hydrochloride a value of 63 kDa (*E. coli* 63 kDa [6]) is found. The lipoyl domain has on gel an apparent mass of 51 kDa (in our gel system *E. coli* 56 kDa) and by sedimentation equilibrium analysis 32.6 kDa (*E. coli* 31.6 kDa [10]). It is thought that the large deviations are caused by the acidic nature of the lipoyl domain [10], or by the alanine + proline-rich regions [13], which are thought to be responsible for anomalous electrophoretic mobility in other proteins, such as the bovine β-crystallin B_1 chains [33]. But also the catalytic domain gives a substantial difference in mass determinations: an apparent mass of 27 kDa on gel and of 15.8 kDa by sedimentation equilibrium analysis. For the *E. coli* subunit-binding domain values of 36 – 39 kDa and 29.6 kDa respectively, have been reported [10].

Although the amino acid sequence of *A. vinelandii* E_2 is not yet known, we can compare our data with the known sequence of *E. coli* E_2 [12] (Fig. 7). From the size of the lipoyl domain and the presence of an extra tyrosyl and histidyl residue, which are not present in the *E. coli* lipoyl domain but are present at position 325 and 327, we locate the C-terminus of the *A. vinelandii* lipoyl domain between residues 325 – 340. The N-terminus of the catalytic domain is homologous with the sequence 370 – 406 of *E. coli* E_2 and therefore this domain will have its C-terminal residue near position 520. Recently Packman and Perham [34, 35] have shown that the E_3 binding domain in the *E. coli* transacetylase and transsuccinylase is located between the lipoyl domain and the catalytic domain. This binding domain is probably located between two 'hinge' regions which allow a high mobility of E_3 when bound to the complex. Fluorescence anisotropy measurements of the FAD in the free or bound E_3 indicate almost no change on rotational correlation coefficient, both in the *A. vinelandii* PDC [16] and in the much larger *E. coli* PDC [36]. Because a large fragment at the C-terminus of E_2 is lost during proteolysis it may be argued that this fragment may represent the E_1 binding site. On the other hand E_1 and E_3 may share the same binding site. Binding experiments have shown competition between E_1 and E_3 and an equivalent binding site was proposed [7]. However, such a binding site could be composed from sequences quite far apart in the primary structure but closely related in the tertiary structure. Further studies will be devoted to elucidating the primary structure of E_2 and the exact location of the E_1 and E_3 binding sites.

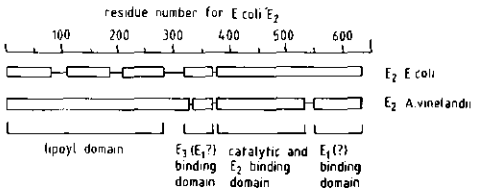


Fig. 7. *A model for the structure of* E. coli E_2 *and* A. vinelandii E_2. The *E. coli* structure is according to Packman and Perham [34]

We wish to thank Mr H. J. Bak for the amino acid analysis and Dr J. F. L. van Breemen (State University Groningen, The Netherlands) for his work on the electron microscope, Mr M. M. Bouwmans for drawing the figures and Miss Y. T. Soekhram for preparation of the manuscript. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid from the Netherlands Organisation for the advancement of Pure Research (ZWO).

REFERENCES

1. Koike, M., Reed, L. J. & Caroll, W. R. (1963) *J. Biol. Chem. 238*, 30 – 39.
2. Bosma, H. J., de Kok, A., Westphal, A. H. & Veeger, C. (1984) *Eur. J. Biochem. 142*, 541 – 549.
3. Reed, L. J. & Oliver, R. M. (1968) *Brookhaven Symp. Biol. 21*, 397 – 411.
4. Reed, L. J., Pettit, F. H., Eley, M. H., Hamilton, L., Collins, J. W. & Oliver, R. M. (1975) *Proc. Natl Acad. Sci. USA 72*, 3068 – 3072.
5. Bosma, H. J., Voordouw, G., de Kok, A. & Veeger, C. (1980) *FEBS Lett. 120*, 179 – 182.
6. Bosma, H. J., de Kok, A., van Markwijk, B. W. & Veeger, C. (1984) *Eur. J. Biochem. 140*, 273 – 280.
7. De Kok, A. & Westphal, A. H. (1985) *Eur. J. Biochem. 152*, 35 – 41.
8. Hardie, D. G. & McCarthy, A. D. (1985) *Biochem. Soc. Trans. 13*, 297 – 299.

9. Coggins, J. R., Boocock, M. R., Campbell, M. S., Chaudhuri, S., Lambert, J. M., Lewendon, A., Mousdale, D. M. & Smith, D. D. S. (1985) *Biochem. Soc. Trans. 13*, 299–303.
10. Bleile, D. M., Munk, P., Oliver, R. M. & Reed, L. J. (1979) *Proc. Natl Acad. Sci. USA 76*, 4385–4389.
11. Packman, L. C., Hale, G. & Perham, R. N. (1984) *EMBO J. 3*, 1315–1319.
12. Stephens, P. E., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983) *Eur. J. Biochem. 133*, 481–489.
13. Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C. & Perham, R. N. (1985) *J. Mol. Biol. 185*, 743–754.
14. Graham, L. D., Guest, J. R., Lewis, H. M., Miles, J. S., Packman, L. C., Perham, R. N., Radford, F. R. S. & Radford, S. E. (1986) *Phil. Trans. R. Soc. Lond. 317*, 391–404.
15. Reed, L. J., Leach, F. R. & Koike, M. (1958) *J. Biol. Chem. 232*, 132–142.
16. Bosma, H. J., de Graaf-Hess, A. C., de Kok, A., Veeger, C., Visser, A. W. J. G. & Voordouw, G. (1982) *Ann. N. Y. Acad. Sci. 378*, 265–285.
17. De Graaf-Hess, A. C. & de Kok, A. (1982) *FEBS Lett. 143*, 261–264.
18. Schachman, K. (1959) *Ultracentrifugation in biochemistry*, Academic Press, New York, London.
19. Cohn, E. J. & Edsall, J. T. (1943) *Proteins, amino acids and peptides*, pp. 370–381, Rheinhold, New York.
20. Penke, B., Ferenczi, R. & Kovacs, K. (1974) *Anal. Biochem. 60*, 45–50.
21. Bresters, T. W., de Abreu, R. A., de Kok, A., Visser, J. & Veeger, C. (1975) *Eur. J. Biochem. 59*, 335–345.
22. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.
23. Laemmli, U. K. (1970) *Nature (Lond.) 227*, 680–685.
24. Voordouw, G., Walker, J. E. & Brenner, S. (1985) *Eur. J. Biochem. 148*, 509–514.
25. Hale, G. & Perham, R. N. (1979) *Eur. J. Biochem. 94*, 119–126.
26. Compton, S. J. & Jones, C. G. (1985) *Anal. Biochem. 151*, 369–374.
27. Pajot, P. (1976) *Eur. J. Biochem. 63*, 263–269.
28. Saxena, V. P. & Wetlaufer, D. B. (1971) *Proc. Natl Acad. Sci. USA 68*, 969–972.
29. Westhof, E., Altschuh, D., Moras, D., Bloomer, A. C., Mondragon, A., Klug, A. & van Regenmortel, M. H. V. (1984) *Nature (Lond.) 311*, 123–126.
30. Stepp, L. R., Bleile, D. M., McRorie, D. K., Pettit, F. H. & Reed, L. J. (1981) *Biochemistry 20*, 4555–4560.
31. Duckworth, H. W., Jaenicke, R., Perham, R. N., Wilkie, A. G. M., Finch, J. T. & Roberts, G. C. K. (1982) *Eur. J. Biochem. 124*, 63–69.
32. Chuang, D. T., Hu, C.-W. C., Ku, L. S., Markovitz, P. J. & Cox, R. P. (1985) *J. Biol. Chem. 260*, 13779–13786.
33. Berbers, G. A. M., Hoekman, W. A., Bloemendal, H., de Jong, W. W., Kleinschmidt, T. & Braunitzer, G. (1983) *FEBS Lett. 161*, 225–229.
34. Packman, L. C. & Perham, R. N. (1986) *FEBS Lett. 206*, 193–198.
35. Packman, L. C. & Perham, R. N. (1987) *Biochem. J. 242*, 531–538.
36. Grande, H. J., Visser, A. J. W. G. & Veeger, C. (1980) *Eur. J. Biochem. 106*, 361–369.

CHAPTER 3

THE DIHYDROLIPOYL TRANSACETYLASE COMPONENT OF THE PYRUVATE DEHYDROGENASE COMPLEX FROM AZOTOBACTER VINELANDII: MOLECULAR CLONING AND SEQUENCE ANALYSIS

Roeland Hanemaaijer, Anja Janssen, Arie de Kok and Cees Veeger

Eur. J. Biochem. *174*, 593 – 599 (1988)

The dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*

Molecular cloning and sequence analysis

Roeland HANEMAAIJER, Anja JANSSEN, Arie de KOK and Cees VEEGER

Department of Biochemistry, Agricultural University, Wageningen

(Received January 26/April 5, 1988) – EJB 88 0100

The gene encoding the dihydrolipoyltransacetylase component (E_2) of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* has been cloned in *Escherichia coli*. A plasmid containing a 2.8-kbp insert of *A. vinelandii* chromosomal DNA was obtained and its nucleotide sequence determined.

The gene comprises 1911 base pairs, 637 codons excluding the initiation codon GUG and stop codon UGA. It is preceded by the gene encoding the pyruvate dehydrogenase component (E_1) of pyruvate dehydrogenase complex and by an intercistronic region of 11 base pairs containing a good ribosome binding site. The gene is followed downstream by a strong terminating sequence. The relative molecular mass (64913), amino acid composition and N-terminal sequence are in good agreement with information obtained from studies on the purified enzyme. Approximately the first half of the gene codes for the lipoyl domain. Three very homologous sequences are present, which are translated in three almost identical units, alternated with non-homologous regions which are very rich in alanyl and prolyl residues. The N-terminus of the catalytic domain is sited at residue 381. Between the lipoyl domain and the catalytic domain, a region of about 50 residues is found containing many charged amino acid residues. This region is characterized as a hinge region and is involved in the binding of the pyruvate dehydrogenase and lipoamide dehydrogenase components. The homology with the dihydrolipoyltransacetylase from *E. coli* is high: 50% amino acid residues are identical.

Dihydrolipoyltransacetylase (E_2) is the core component of the pyruvate dehydrogenase complex. This complex catalyzes the oxidative decarboxylation of pyruvate to acetyl-CoA and NADH [1]:

$$\text{Pyruvate} + \text{NAD}^+ + \text{CoA} \rightarrow \text{AcetylCoA} + \text{CO}_2 + \text{NADH} + \text{H}^+.$$

The E_2 component comprises many functions. In the *Azotobacter vinelandii* enzyme three pyruvate dehydrogenase (E_1) dimers and one lipoamide dehydrogenase (E_3) dimer are bound to a core of four E_2 chains [2]. The E_2 chain contains covalently bound lipoyl moieties which transport the substrates between the different active sites of the complex.

Limited proteolysis studies with trypsin have shown that E_2 consists of at least two domains: an N-terminal lipoyl domain which contains the lipoyl moieties and the C-terminal catalytic domain which possesses the catalytic site and the E_2-E_2 intersubunit binding sites [3]. The binding sites for the E_1 and E_3 components were not found on these domains.

Correspondence to A. de Kok, Laboratorium voor Biochemie, Landbouw-Hogeschool, De Dreijen 11, NL-6703 BC Wageningen, The Netherlands

Abbreviations. kbp, 10^3 base pairs; LGT-agarose, low-gelling-temperature agarose.

Enzymes. Pyruvate dehydrogenase (E_1), pyruvate:lipoate oxidoreductase (EC 1.2.4.1); dihydrolipoyltransacetylase (E_2), acetylCoA:dihydrolipoamide *S*-acetyltransferase (EC 2.3.1.21); lipoamide dehydrogenase (E_3), NADH:lipoamide oxidoreductase (EC 1.8.1.4); DNA polymerase I (Klenow fragment) (EC 2.7.7.7); calf intestinal alkaline phosphatase (EC 3.1.3.1); restriction endonucleases *Bam*HI, *Eco*RI, *Hin*dIII, *Sau*3A and *Sma*I (EC 3.1.24.4).

The E_2 component of the pyruvate dehydrogenase complex from *Escherichia coli* has a comparable domain structure: an N-terminal lipoyl domain and a C-terminal domain, which is called the catalytic and subunit-binding domain. In contrast to the *A. vinelandii* E_2, the C-terminal domain of *E. coli* possesses binding sites for the E_1 and E_3 components [4, 5]. This domain contains an N-terminal extension of about 50 residues and this region is indicated as a binding site for the E_3 component [6, 7]. The quaternary structure of the *E. coli* complex differs from that of *A. vinelandii*, being based on an E_2 core of 24 subunits [8], whereas the *A. vinelandii* one is based on a tetrameric E_2 core [2]. Upon removal of the peripheral components from the *A. vinelandii* complex, the E_2 tetramers associate to an *E. coli*-like multimer, as shown by sedimentation velocity experiments and electron microscopy [3, 9].

The gene for *E. coli* dihydrolipoyltransacetylase has been cloned and sequenced [10]. It is preceded by the gene encoding the pyruvate dehydrogenase component (E_1). Both genes are transcribed in one transcription unit, from a promoter upstream of the gene encoding E_1 to a terminating sequence downstream of the gene encoding E_2. The gene encoding the E_3 component lies downstream of this terminating sequence [11]. The gene organization in *A. vinelandii* is different: in this organism the gene encoding the E_3 component is located downstream of the gene encoding the E_2 component of the 2-oxoglutarate dehydrogenase complex [12]. This paper reports on the cloning and nucleotide sequence analysis of the gene encoding the dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *A. vinelandii*. Infor-

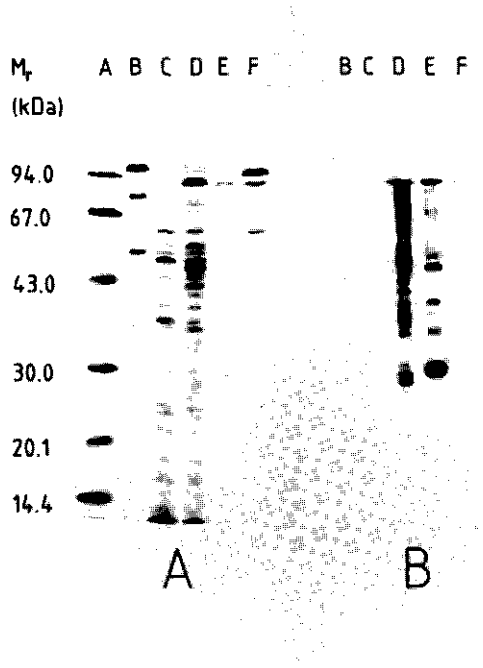


Fig. 1. *Expression of* A. vinelandii *dihydrolipoyltransacetylase (E_2) in* E. coli *TG_2 (pRA 282).* (A) SDS gel electrophoresis of cell-free extracts, purified E_2 and the pyruvate dehydrogenase complex stained with Coomassie brilliant blue R. (B) Western blot of cell-free extracts, purified E_2 and pyruvate dehydrogenase complex. Lane A, marker proteins (molecular mass in kDa on left); lane B, pyruvate dehydrogenase complex from *E. coli*; lane C, cell-free extract of *E. coli* TG_2; lane D, cell-free extract of *E. coli* TG_2 (pRA 282); lane E, purified E_2 (with degradation products due to storage); lane F, pyruvate dehydrogenase complex from *A. vinelandii*

mation is given about the size and the structural organization, and a comparison with the *E. coli* gene is made.

MATERIALS AND METHODS

Materials

Restriction endonucleases were purchased from Bethesda Research Laboratories (BRL) and Boehringer. DNA polymerase (Klenow fragment) and T4-DNA ligase were obtained from BRL. Calf intestinal phosphatase, universal M13 sequencing primer and 7-deaza-dGTP were purchased from Boehringer. [α-^{32}P]dATP (3000 Ci/mmol) was obtained fron New England Nuclear. All other chemicals used were of analytical grade.

Bacterial strains and vectors

A. vinelandii (ATCC478) was grown as described [13] and used for DNA isolation. *E. coli* strain TG_2, a *recA*$^-$ version of TG_1 [*Δ(lac-pro), thi, supE, hsdM*$^-$, *hsdR*$^-$, F′(*traD36 proA*$^+$ *B*$^+$ *lacZ lacI*^{q} *ΔM15*)] was used throughout [14]. Plasmid pUC9 was used as a cloning vector [15] and phage M13mp18 and M13mp19 were used for sequencing [16].

DNA isolation and cloning

Standard DNA operations were performed as described [17]. Chromosomal DNA was isolated from *A. vinelandii* cells [12] and partially digested with *Sau*3A. Fragments of 9–23 kb were isolated after LGT-agarose gel electrophoresis. These fragments were ligated into the *Bam*HI site of pUC9. *E. coli* TG_2 cells, transformed with recombinant plasmids, were grown on two-gridded nitrocellulose filters, placed on TY-agar plates, lysed according to [12], and screened for production of *A. vinelandii* dihydrolipoyltransacetylase (E_2) using antibodies raised against the catalytic domain of E_2. A 14-kbp plasmid of a positive clone [TG_2(pRA177)] was isolated as described in [18], with modifications according to [12]. After digestion with *Bam*HI the insert of the plasmid was isolated, partially digested with *Sau*3A and fragments of 2–3 kbp and 3–4 kbp were isolated from electrophoresis in LGT-agarose. These fragments were ligated into the *Bam*HI site of pUC9 and, after transformation in *E. coli* TG_2 cells, colonies were screened as described above.

DNA sequence determinations and analysis

The insert from a positive recombinant was isolated after digestion with *Bam*HI. The double-stranded DNA was sonicated for three 30-s periods with 30-s intervals at 0 C. Fragments of 300–600 bp were isolated, end-repaired as described [19] and isolated after LGT-agarose gel electrophoresis. The fragments were cloned into the *Sma*I site of the replicative form of M13mp19. The single-stranded form of the recombinant phage DNA was prepared by standard techniques [20] and sequence data were obtained according the dideoxy-chain-termination method of Sanger [21], using the

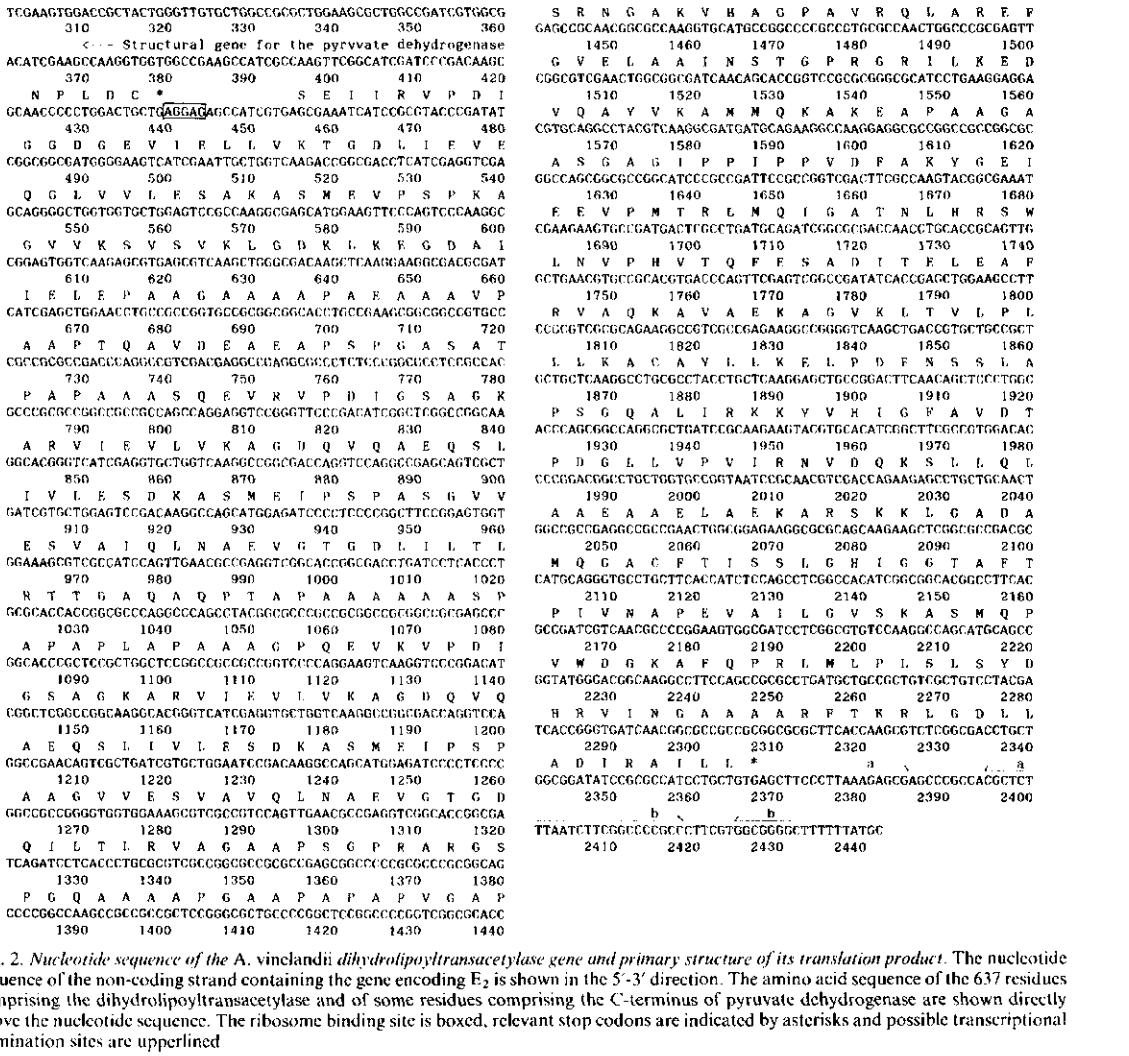


Fig. 2. *Nucleotide sequence of the* A. vinclandii *dihydrolipoyltransacetylase gene and primary structure of its translation product.* The nucleotide sequence of the non-coding strand containing the gene encoding E_2 is shown in the 5′-3′ direction. The amino acid sequence of the 637 residues comprising the dihydrolipoyltransacetylase and of some residues comprising the C-terminus of pyruvate dehydrogenase are shown directly above the nucleotide sequence. The ribosome binding site is boxed, relevant stop codons are indicated by asterisks and possible transcriptional termination sites are upperlined

universal M13 primer or two synthetic oligonucleotides, representing two regions in the lipoyl domain.

The sequencing reaction was carried out at 50°C. In some cases instead of dGTP, 7-deaza-dGTP was used [22]. Gels were run at 60°C to avoid secondary structure in the G+C-rich regions [23].

The data were compiled in a VAX computer using the programs of Staden [24, 25].

Other techniques

Enzyme activities were assayed at 25°C as described previously [12]. Antisera were collected as described in [3]. Because of cross-reactivity with *E. coli* E_2, antisera were saturated with *E. coli* TG_2 cell-free extract before use. Gel electrophoresis in the presence of sodium dodecyl sulphate was carried out essentially as described by Laemmli [26].

RESULTS AND DISCUSSION

Cloning strategy

From partially *Sau*3A-digested *A. vinelandii* chromosomal DNA fragments of 9–23 kbp were isolated and ligated into the *Bam*HI site of pUC9. *E. coli* TG_2 cells transformed with the recombinant plasmids were screened using antiserum aginst the catalytic domain of E_2. This antiserum shows less cross-reactivity with *E. coli* enzymes in cell-free extract than the antisera raised against intact E_2 or against the lipoyl domain. Of 600 colonies screened two positives were found as shown by SDS gel electrophoresis of cell-free extracts followed by Western blotting. From one of the positive colonies a 14-kbp recombinant plasmid was isolated. After partial digestion of the insert with *Sau*3A, fragments of 2–3 kbp and of 3–4 kbp were isolated. After ligation in the *Bam*HI site of pUC9 and transformation to *E. coli* TG_2 cells, screening for

Table 1. *Codon usage in the gene encoding dihydrolipoyltransacetylase from* A. vinclandii
The GUG initiation codon is not included

F	UUU	0	S	UCU	1	Y	UAU	0	C	UGU	0
F	UUC	10	S	UUC	11	Y	UAC	5	C	UGC	2
L	UUA	0	S	UCA	0	*	UAA	0	*	UGA	1
L	UUG	3	S	UCG	6	*	UAG	0	W	UGG	2
L	CUU	0	P	CCU	3	H	CAU	1	R	CGU	1
L	CUC	10	P	CCC	17	H	CAC	5	R	CGC	19
L	CUA	0	P	CCA	0	Q	CAA	3	R	CGA	0
L	CUG	44	P	CCG	31	Q	CAG	25	R	CGG	4
I	AUU	1	T	ACU	1	N	AAU	0	S	AGU	2
I	AUC	34	T	ACC	16	N	AAC	10	S	AGC	21
I	AUA	0	T	ACA	0	K	AAA	0	R	AGA	0
M	AUG	10	T	ACG	4	K	AAG	34	R	AAG	0
V	GUU	2	A	GCU	6	D	GAU	6	G	GGU	4
V	GUC	30	A	GCC	78	D	GAC	20	G	GGC	42
V	GUA	3	A	GCA	6	E	GAA	19	G	GGA	2
V	GUG	26	A	GCG	28	E	GAG	24	G	GGG	5

E_2 production was carried out. Of 400 colonies with a 3–4 kbp insert screened, 19 were positive, 18 producing intact E_2 as judged by SDS gel electrophoresis and Western blotting. Of 340 colonies with a 2–3 kbp insert screened 13 were positive, of which 12 produced intact enzyme as is shown for colony RA282 in Fig. 1. The plasmid of this colony contained an insert of 2.8 kbp and was used for sequencing analysis.

The transacetylase activity in the cell-free extract of this colony was 15-fold higher than the activity of the *E. coli* enzyme measured in the cell-free extract of TG_2 cells containing pUC9 without insert. No increased E_2 activity was found when cells were grown without addition of isopropyl β-D-thiogalactoside to the growth medium, indicating that transcription is dependent on the vector-encoded *lacZ* promoter.

Sequencing strategy

The 'shot-gun' cloning method was used to sequence the 2.8-kbp fragment. The complete nucleotide sequence of the region containing the dihydrolipoyltransacetylase gene is shown in Fig. 2. The sequence was 91% overlapping (100% in the coding region) and 79% (94% in the coding region) was derived from both strands of DNA (residues 1345–1355, 2009–2019 and 2209–2310 are derived from one strand of DNA). Because it was difficult to sequence a very G+C-rich part of DNA in the region encoding the lipoyl domain (residues 1030–1120), 7-deaza-dGTP was used in the sequencing reaction and gels were run at 60 °C. Also two synthetic oligonucleotides were used (identical to residues 996–1011 and 1146–1131) reading this sequence from both strands.

Coding regions

A 1911-bp open reading frame was found, which was identified as the gene encoding the dihydrolipoyltransacetylase, its initiation codon being GUG. This codon usually codes for the amino acid valine, but as a start codon it codes for formylmethionine [26]. *In vitro* binding experiments with fMet-tRNA do not show strong discrimination against GUG [26]. About 3% of the known genes possesses the GUG start codon [27].

Table 2. *Amino acid composition of dihydrolipoyltransacetylase from* A. vinelandii
The amino acid composition deduced from the nucleotide sequence is compared with the composition reported from the amino acid analysis on the purified enzyme [3]. The methionine initiation residue is not included

Amino acid	Number of residues from DNA sequence	Composition from	
		DNA sequence	amino acid analysis
		mol/100 mol	
Asp	26	4.1	6.2
Asn	10	1.6	
Thr	21	3.3	3.2
Ser	41	6.4	5.5
Glu	43	6.8	10.8
Gln	28	4.4	
Pro	51	8.0	8.4
Gly	53	8.3	8.8
Ala	118	18.5	21.0
Cys	2	0.3	0.6
Val	61	9.6	7.7
Met	10	1.6	1.6
Ile	35	5.5	4.5
Leu	57	8.9	8.4
Tyr	5	0.8	1.0
Phe	10	1.6	1.7
Lys	34	5.3	5.4
His	6	0.9	1.1
Arg	24	3.8	3.7
Trp	2	0.3	0.4
Total	637		
M_r	64913		

The molecular mass of the protein deduced from the open reading frame corresponds to that determined by sedimentation equilibrium analysis on the purified enzyme in 6 M guanidine hydrochloride [3]. One other open reading frame was identified preceding the gene encoding E_2. The deduced amino acid sequence (145 residues) showed a high degree of homology (49%) with the C-terminal part of the *E. coli* pyruvate dehydrogenase (E_1, residues 774–886).

The intercistronic region

Between the stop codon of the first open reading frame (E_1) and the start codon of the second open reading frame (E_2) an intercistronic region of 11 bp is found. It contains a good ribosome binding site (position 440–444, Fig. 2). No independent promotor sequence could be identified, which indicates that, as in *E. coli* [11], the genes encoding E_1 and E_2 are transcribed from the same promotor, suggesting that the difference in subunit ratio of the two components in the complex ($E_1 : E_2 = 1.5 : 1.0$) is regulated at the translational level.

The nucleotide sequence downstream of the gene encoding E_2

Downstream of the termination codon of the gene encoding E_2 two regions of dyad symmetry are found, positions 2376–2404 (a) and 2411–2432 (b). In these regions the mRNA transcripts can form very stable stem and loop structures with a free energy of −56.1 kJ (a) and −126.5 kJ (b), respectively. The presence of six thymidine residues fol-

Fig. 3. *Comparison of the amino acid sequence of dihydrolipoyltransacetylase from* A. vinelandii *and* E. coli *[10].* Residues are given in the single-letter amino acid code. Homologous residues are indicated by asterisks

lowing the very G + C-rich dyad symmetry, characteristic for *rho*-independent terminators [29], makes it likely that the transcription termination takes place in this region.

In *E. coli* the *lpd* gene encoding E_3 was found downstream of the gene encoding E_2 [11]. In the region downstream of the gene encoding *A. vinelandii* E_2 (500 bp) no reading frame could be found which codes for E_3. This is accordance with the results of Westphal and de Kok [12] who showed that the gene encoding E_3 was located downstream of the gene encoding the dihydrolipoyltranssuccinylase component of the 2-oxoglutarate dehydrogenase complex. The E_3 component is shared by both complexes [30].

Codon usage

The codon usage is shown in Table 1. It is non-random with a typical high overall G + C content of 71%, even higher than the G + C content of other *A. vinelandii* genes [12, 31]. The G + C content in the third position of the codons is 90%.

Primary structure and composition of the E_2 component

From the nucleotide sequence the primary structure was derived as shown in Fig. 2. It is assumed that the initiating

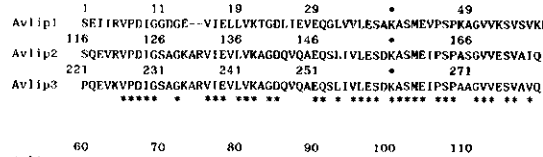


Fig. 4. *Repeating units in the lipoyl domain.* The three homologous repeating units are aligned for maximum homology. Residues that are identical are indicated by asterisks. Residues are given in the single-letter amino acid code. Potential lipoyl-binding sites are postulated from homology with other organisms [10, 30, 31] and are indicated by a closed circle (●) above the sequence

formylmethionine residue is removed post-translationally, so that the serine residue corresponds to the first residue of the N-terminal amino acid sequence that has been determined by automated Edman degradation of the purified protein as isolated from *A. vinelandii* [3]. An internal sequence of 36

residues, obtained by Edman degradation of the N-terminus of the catalytic domain [3], corresponds exactly with the sequence 381 – 416. Only the last residue proved to be an arginyl residue instead of the reported methionyl residue.

The gene encodes a polypeptide of 637 amino acid residues with a total mass of 64.9 kDa. This is in excellent agreement with the mass of 63 kDa as determined by sedimentation equilibrium analysis [3]. The chain mass is significant lower than that based on SDS/polyacrylamide gel electrophoresis (82 kDa). The predicted amino acid composition agrees well with that determined for the purified enzyme [3] (Table 2). The sequence indicates the presence of two cysteinyl residues, located in the catalytic domain. SH group determination with 5,5'-dithiobis(2-nitrobenzoic acid) in the presence of 6 M guanidine · HCl without prior reduction indicated 1 – 2 SH groups, whereas no reaction was observed with the native enzyme, in agreement with previous labeling experiments of the whole complex with *N*-ethyl[2,3-^{14}C]maleimide [32]. This indicates that both SH groups are buried.

Domain structure and comparison with the E. coli *sequence*

The alignment of the amino acid sequences of E_2 from *E. coli* and *A. vinelandii* is shown in Fig. 3. The overall homology is very high: 50% of the residues are identical or 69% if closely related amino acids are included. In the N-terminal part of the E_2 chain, concerning the lipoyl domain [10], the primary structure shows three repeating units, alternated by regions which are very rich in alanyl and prolyl residues (Fig. 4). In purified E_2 2 – 3 lipoyl groups per subunit were estimated [33] so that potentially each unit will possess a lipoyl binding site. In *E. coli* the region around the lipoyl binding site was found to be very conserved in all three repeating structures [10, 34]. In *A. vinelandii* the same region is found in all three repeating units; therefore, we conclude that the lipoyl groups are attached to Lys39, Lys156 and Lys261. Analogous regions have been found in the E_2 chain of *E. coli* [35] and *A. vinelandii* (unpublished results) 2-oxoglutarate dehydrogenase complexes and in the bovine heart pyruvate dehydrogenase complex [36]. An homologous region is present in the biotin carboxyl carrier protein (EC 2.1.3.1) from *Propionibacterium shermanii* [37]. The biotinyl prosthetic group is bound to a lysyl residue in an equivalent position as the lipoyl carrying lysyl residue.

The C-terminus of the lipoyl domain has not been determined. From the molecular mass of 32.8 kDa of the lipoyl domain, as determined by analytical ultracentrifugation, cleavage products with other proteases and homology with the *E. coli* sequence, the most probable location of the C-terminus is in the region between residue 331 and 347.

The N-terminal sequence of the *A. vinelandii* catalytic domain starts at residue 381, as concluded from a comparison with the N-terminal sequence of the catalytic domain, obtained by limited proteolysis [3]. Packman et al. [5] have reported that limited proteolysis of *E. coli* E_2 by trypsin resulted in cleavage at position 316 of the *E. coli* sequence. A fragment called 'inner core domain' was isolated which still contained the binding sites for E_1 and E_3. More recently [7] a smaller fragment was obtained after prolonged digestion with trypsin. This fragment, starting at position 370, equivalent with that of the *A. vinelandii* catalytic domain (Fig. 3), has lost the E_3 binding site but seems to retain the E_1 binding site. In recent experiments (unpublished) we have determined the C-terminus of the catalytic domain using carboxypeptidase Y. It appeared to be identical to the C-terminus of intact E_2. Therefore proteolytic removal of the C-terminus, a conclusion inferred from molecular mass determination of the catalytic domain, seems not to be justified. This will be more fully discussed in a subsequent paper.

From the indicated domain structure it is concluded that the lipoyl domain and the catalytic domain are separated by about 50 amino acid residues. This region is very sensitive to proteolysis by trypsin [3], chymotrypsin and *Staphylococcus aureus* proteinase (unpublished results) and is therefore indicated as a hinge region.

In previous experiments we have shown that, after limited proteolysis of E_2 with trypsin, the E_1 and E_3 binding sites are lost [3]. This indicates that the hinge region is involved in the binding of both components, although from these experiments we cannot exclude that other regions on E_2 might contribute to the binding.

We have shown previously that hybrid active complexes could be formed betweeen E_3 from *E. coli* and the E_1E_2 subcomplex from *A. vinelandii* and between E_1 from *A. vinelandii* and the E_2E_3 subcomplex from *E. coli* [30]. Furthermore the E_3 component of each organism binds to the E_2 components of both pyruvate and 2-oxoglutarate dehydrogenase complexes. Therefore sequence homology between the E_3 binding sites is expected. From an inspection of the sequences (Fig. 3) it is clear that it is mainly the charged amino acid residues in the hinge region which are conserved. This observation extends to the E_2 chain of 2-oxoglutarate dehydrogenase complex from *A. vinelandii* (unpublished results). A further indication that charged residues are involved in the binding of the peripheral components is given by the methods used for resolution of the components from the complex: either high pH or high salt concentration. Further research is being directed to elucidating the E_1 and E_3 binding sites.

We wish to thank A. H. Westphal for experimental advice and for isolation of *A. vinelandii* chromosomal DNA. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid from the Netherlands Organisation for the advancement of Pure Research (ZWO).

REFERENCES

1. Koike, M., Reed, L. J. & Caroll, W. R. (1963) *J. Biol. Chem. 238*, 30 – 39.
2. Bosma, H. J., de Kok, A., Westphal, A. H. & Veeger, C. (1984) *Eur. J. Biochem. 142*, 541 – 549.
3. Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987) *Eur. J. Biochem. 169*, 245 – 252.
4. Bleile, D. M., Munk, P., Oliver, R. M. & Reed, L. J. (1979) *Proc. Natl Acad. Sci. USA 76*, 4385 – 4389.
5. Packman, L. C., Hale, G. & Perham, R. N. (1984) *EMBO J. 3*, 1315 – 1319.
6. Packman, L. C. & Perham, R. N. (1986) *FEBS Lett. 206*, 193 – 198.
7. Packman, L. C. & Perham, R. N. (1987) *Biochem. J. 242*, 531 – 538.
8. Reed, L. J. & Oliver, R. M. (1968) *Brookhaven Symp. Biol. 21*, 397 – 411.
9. Bosma, H. J., de Graaf-Hess, A. C., de Kok, A., Veeger, C., Visser, A. W. J. G. & Voordouw, G. (1982) *Ann. NY Acad. Sci. 378* 265 – 285.
10. Stephens, P. E., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983) *Eur. J. Biochem. 133*, 481 – 489.
11. Spencer, M. E. & Guest, J. R. (1985) *Mol. Gen. Genet. 200*, 145 – 154.
12. Westphal, A. H. & de Kok, A. (1988) *Eur. J. Biochem. 172*, 299 – 305.

13. Bresters, T. W., de Abreu, R. A., de Kok, A., Visser, J. & Veeger, C. (1975) *Eur. J. Biochem. 59*, 335–345.
14. Gibson, T. J. (1984) Ph.D. Thesis, University of Cambridge.
15. Vieira, J. & Messing, J. (1982) *Gene 19*, 259–268.
16. Norrander, J., Kempe, T. & Messing, J. (1983) *Gene 26*, 101–106.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York.
18. Voordouw, G., Walker, J. E. & Brenner, S. (1985) *Eur. J. Biochem. 148*, 509–514.
19. Bankier, A. T. & Barrel, B. G. (1983) in *Techniques in nucleic acid biochemistry* (Flavell, R. A., ed.) pp. 1–34, Elsevier Scientific Publishers, Ireland.
20. Sanger, F., Coulson, A. R., Barrel, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161–178.
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl Acad. Sci. USA 74*, 5463–5467.
22. Mizusawa, S., Nishimura, S. & Seela, F. (1986) *Nucleic Acids Res. 14*, 1319–1324.
23. Martin, R. (1985) *Focus 7*, 1–6.
24. Staden, R. (1982) *Nucleic Acids Res. 10*, 4731–4751.
25. Staden, R. (1984) *Nucleic Acids Res. 12*, 499–503.
26. Marcker, K. A., Clark, B. F. C. & Anderson, J. F. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 279–285.
27. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981) *Annu. Rev. Microbiol. 35*, 365–403.
28. Tinoco, I. Jr, Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nat. New Biol. 246*, 40–41.
29. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet. 13*, 319–353.
30. De Kok, A. & Westphal, A. H. (1985) *Eur. J. Biochem. 152*, 35–41.
31. Brigle, K. E., Newton, W. E. & Dean, D. R. (1985) *Gene 37*, 37–44.
32. De Abreu, R. A., De Kok, A., De Graaf-Hess, A. E. & Veeger, C. (1977) *Eur. J. Biochem. 81*, 357–364.
33. De Kok, A., Visser, A. J. W. G. & De Graaf-Hess, A. C. (1982) in *Flavins and flavoproteins* (Massey, V. & Williams, C. H., eds) pp. 61–65, Elsevier, Amsterdam.
34. Hale, G. & Perham, R. N. (1979) *Eur. J. Biochem. 94*, 119–126.
35. Spencer, M. E., Darlison, M. G., Stephens, P. E., Duckenfield, I. K. & Guest, J. R. (1984) *Eur. J. Biochem. 141*, 361–374.
36. Bradford, A. P., Howell, S., Aitken, A., James, L. A. & Yeaman, S. J. (1987) *Btochem. J. 245*, 919–922.
37. Maloy, W. L., Bowien, B. U., Zwolinski, G. K., Kumar., K. G., Wood, H. G., Ericsson, L. H. & Walsh, K. A. (1979) *J. Biol. Chem. 254*, 11615–11622.

CHAPTER 4

THE GENE ENCODING DIHYDROLIPOYL TRANSACETYLASE FROM *AZOTOBACTER VINELANDII*: EXPRESSION IN *ESCHERICHIA COLI* AND ACTIVATION AND ISOLATION OF THE PROTEIN

Roeland Hanemaaijer, Adrie H. Westphal, Axel Berg, Arie de Kok and Cees Veeger

Summary

The gene encoding the dihydrolipoyl transacetylase (E_2) component from Azotobacter vinelandii has been cloned in Escherichia coli. High expression of the gene was found when the cells were grown for more than 14 hours. The produced E_2 was partially inactive, varying between 10-90% in different experiments. By limited proteolysis of the protein it was shown that the catalytic domain was incorrectly folded, caused by formation of inter- or intramolecular S-S-bridges. The enzyme was fully activated after unfolding in 2.5 M guanidine hydrochloride containing 2 mM dithiothreitol, followed by refolding by dialysis. Active E_2 was isolated in a simple three-step isolation procedure. It possessed a specific activity more than twice as high as found after isolation of E_2 from purified pyruvate dehydrogenase complex from A.vinelandii. Active E_2 comprises about 7% of the total soluble cellular protein in the E.coli clone.
By genetic manipulation deletion mutants of E_2 were created, one encoding the lipoyl domain and the N-terminal half of the E_1- and E_3 binding domain, the other encoding the catalytic domain and the C-terminal half of the E_1-and E_3 binding domain. In E.coli expression of both mutants was observed.

Introduction

Dihydrolipoyl transacetylase (E_2) is the core component of the pyruvate dehydrogenase complex (PDC).It comprises many functions in which several sites are involved: binding sites for the pyruvate dehydrogenase (E_1) and lipoamide dehydrogenase (E_3) components, the E_2 intersubunit binding sites, the transacetylase active site and the covalently bound lipoyl moieties, which transport the substrates between the different active sites [1]. Limited proteolysis studies [2] have shown that E_2 consists of at least two stable domains: an N-terminal lipoyl domain, containing the lipoyl moieties, and a C-terminal catalytic domain, possessing the transacetylase active site and the E_2 intersubunit binding sites. The E_1 and E_3 binding sites are lost during proteolysis.

The gene encoding E_2 of PDC from Azotobacter vinelandii has been cloned and sequenced [3]. It consists of 637 codons. It is preceded by the gene encoding E_1 and by an intercistronic region of 11 basepairs, containing a ribosome binding site, which has a high degree of homology with the E.coli consensus ribosome binding site. The lipoyl domain is located between residues 1 and 331/347 [2,4]. The catalytic domain is located between residue 381 and 637. The region between these two domains is thought to possess the E_1 and E_3 binding sites [2]. It possesses many charged residues and is easily digested by various proteases [2,4]. The homology of A.vinelandii E_2 with E.coli E_2 is high: 50% of the amino acid residues is identical.
E.coli cells, transformed with plasmid pRA282, that contain the gene encoding A.vinelandii E_2 on a 2.8 kbp insert, show a high expression of this gene. A 15-fold higher activity was found compared to the background of E.coli E_2 [2].
In this paper some characteristics of the A.vinelandii protein, synthesized in the E.coli host, are shown. A high E_2 production is found when the cells are grown for more than 14 hours. E_2 can be isolated in a simple three-step isolation procedure. A variable amount of the protein was inactive. An activation procedure is described by which fully active E_2 is obtained. Two plasmids are constructed encoding the lipoyl domain fused with six and the catalytic domain fused with ten vector-encoded residues. These separate domains are synthesized as stable proteins in the E.coli host.

Materials and methods

Bacterial strains and plasmids

E.coli TG2 [Δ(lac-pro) thi recA⁻ supE hsdR hsdM F'(traD36 proA⁺B⁺ lacIq lacZΔM15)] [5] or E.coli MC1061 [araD139 Δ(ara leu7697)ΔlacY74 galU galK hsdM⁺ hsdR⁻ strA] [6] were used as host for the expression of plasmid-encoded E_2. E.coli Y1083 [thi Δ(lacIPOZYA)U169 araD139 Δlon phx rpsL hflA150[chr::Tn10]], derived from Y1089 [7], was used as host for the expression of the plasmid-encoded lipoyl domain and catalytic domain.

Materials

Standard restriction endonucleases, T4 DNA ligase, DNA polymerase I (Klenow fragment) were obtained from Bethesda Research Laboratories (BRL). Restriction endonuclease RsrII was obtained from Janssen Biochimica. Ampicillin, ethidium bromide and IPTG were from Sigma. Tosylphenylalanylchloromethane-treated trypsin was obtained from Worthington Biochemical Corp. DEAE-Sephacel and Sephacryl S300 were from Pharmacia Fine Chemicals.

General methods

Plasmid DNA was isolated from cleared lysates by equilibrium centrifugation in cesium chloride containing ethidium bromide [8]. Small-scale preparations were prepared according to [9]. Methods of digestion with restriction endonucleases, DNA ligation, plasmid transformation and agarose gel electrophoresis of DNA have been described before [2,10] or were carried out according to Maniatis et al. [11]. Western-blotting was carried out according to [12]; visualization of bound antibodies was with secondary antibodies (goat-anti-rabbit) conjugated to alkaline-phosphatase (Protoblot, Promega Biotec). E_2 transacetylase activity was assayed spectrophotometrically at 240 nm and 25°C as described previously [13]. The lipoyl content was determined using the 5,5'-dithio-bis(2-nitrobenzoic acid) cleavage reaction [14].

Construction of plasmids

Plasmid pRA282 (Fig.1), derived from the vector pUC9 [15] and encoding E_2, has been described previously [2]. Plasmid pRC1 was constructed from pRA282 by deleting the sequences encoding the lipoyl domain of E_2 (Fig.1). Therefore, pRA282 was cut with restriction endonucleases RsrII and HindIII; the ends were filled-up with Klenow polymerase and the blunt-ended plasmid was closed with T4 DNA ligase and introduced into E.coli. Plasmid pRL1 was also derived from pRA282 by removing the sequences encoding the catalytic domain. (Fig.1). To this end pRA282 was cut with

restriction endonucleases RsrII and EcoRI. After end-repair with Klenow the plasmid was closed and introduced into E.coli. For pRC1 and pRL1 E.coli TG2 and E.coli Y1083 were used as host.

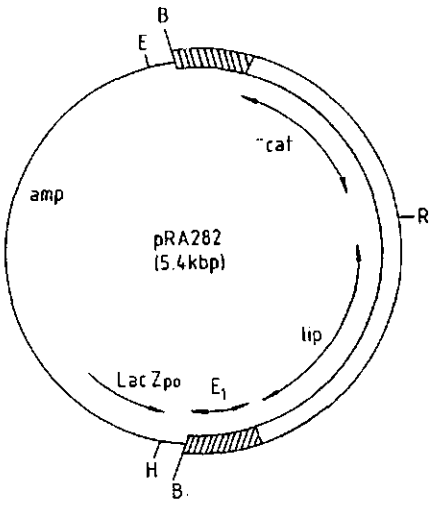


Figure 1. Physical map of plasmid pRA 282. Shown are the positions of the gene encoding E_2 (white box), with the regions encoding the lipoyl (lip) and the catalytic (cat) domain and part of the gene for E_1 (hatched) and a non-coding region (also hatched). The single line represents the vector, pUC9, with the gene encoding β-lactamase (amp) and the lacZ promoter/operator (lacZpo), whose direction is indicated by the arrow. Restriction sites are shown for BamHI (B), EcoRI (E), HindIII (H) and RsrII (R).

Preparation and analysis of cell-free extracts

E.coli TG2, E.coli MC1061 or E.coli Y1083, harbouring recombinant plasmids, were grown at 37°C in TY medium [16], containing ampicillin (75 µg/ml) and IPTG (20 µg/ml), for at least 17 h, unless described otherwise. A 10 ml cell suspension was concentrated 20-fold and cells were disrupted by sonication. After centrifugation for 10 min. at 13,000xg the supernatant was assayed for E2-activity and subjected to SDS-PAGE.

Isolation of A.vinelandii E_2 expressed in E.coli

A single colony of E.coli TG2(pRA282) was picked from a TY-plate containing 50 µg/ml ampicillin and grown overnight at 37°C in 10 ml TY containing 75 µg/ml ampicillin and 20 µg/ml IPTG. 0.5 ml of this culture was used to inoculate 50 ml of the same medium and

this culture was grown at 37°C for 8 hours. The culture was stored overnight at 4°C and the next day used for inoculation of 20 litres of cold (4°C) growth-medium. The large culture was grown for 40 hours at 37°C. Using these growth-conditions less than 10% of E_2 was in the inactive form. Cells (80-100 grams) were harvested and used for isolation of E_2. Cells were suspended in 50 mM potassium phosphate pH 7.0, containing 3 mM EDTA and 0.1 mM $PhMeSO_2F$ (standard buffer), and disrupted using a Manton-Gaulin laboratory homogenizer at 9000 psi. After centrifugation for 30 minutes at 14000xg a PEG-$MgCl_2$-precipitation was carried out in two steps: after addition of 10% (w/v) PEG-6000 and 0.75 mM MgCl2 (final concentrations) a large amount of protein was precipitated, while E_2 remained in solution. At 12 mM $MgCl_2$ E_2 was precipitated and after centrifugation for 30 minutes at 20000xg the pellet was suspended in standard buffer containing 12 mM EDTA. The clear solution was applied to a DEAE-Sephacel column and eluted using a 0-500 mM KCl-gradient in standard buffer. The peak-fraction was concentrated by ultrafiltration (Amicon YM100) and applied to a Sephacryl S300 column. The fractions of this column were analysed for activity and on SDS-gel.

Activation-procedure for E_2

Cell-free extract from E.coli MC1061(pRA282), which had grown for 23 h, was diluted 6-fold in 50 mM potassium phosphate, pH 7.0, 0.1 mM EDTA and 0.1 mM $PhMeSO_2F$. After addition of 0.5 volume of a 7.5 M GdnHCl stock solution (final concentration 2.5 M GdnHCl),the samples were incubated for at least 15 minutes at 0°C. After incubation CoA and/or lipoic acid were added from a stock solution to a final concentration of 0.5 mg/ml CoA and/or 1 mM lipoic acid. In the samples without CoA or lipoic acid a same volume of standard buffer was added.
Renaturation and analysis of the activity was carried out in three ways: 1. dilution of the sample directly in the assay mix (the final concentration of GdnHCl is 25 mM; this concentration does not inhibit E_2-activity); 2. a 5-fold dilution in standard buffer, and after incubation for at least 1 h measurement of the E_2-activity and 3. dialysis (o/n) against standard buffer, after which E_2-activity is measured.

Results and discussion

Expression of A.vinelandii E_2 in E.coli

Plasmid pRA282 (Fig.1) was used for expression of A.vinelandii E_2 in E.coli TG2. The gene encoding E_2 is preceded by a region of about 500 nucleotides, encoding the C-terminal part of the pyruvate dehydrogenase (E_1) component of the pyruvate dehydrogenase complex [2]. No terminator for transcription was found in the intercistronic region of 11 bp. The gene encoding E_2 is under regulation of the vector-encoded lacZ-promoter.

The E_2-activity of the E.coli clone that expresses A.vinelandii E_2 is shown in Fig. 2 at different stages of growth. Expression of E_2-activity is low in the exponential growth-phase (less than eight hours of growth), but steadily increases during the end-log and stationary phases of growth.

After approximately 18-20 hours, when the cells are well within the stationary phase, a peculiar event takes place. In most experiments the E_2-activity, expressed by the E.coli clone, remained fairly constant after this time (as represented by the

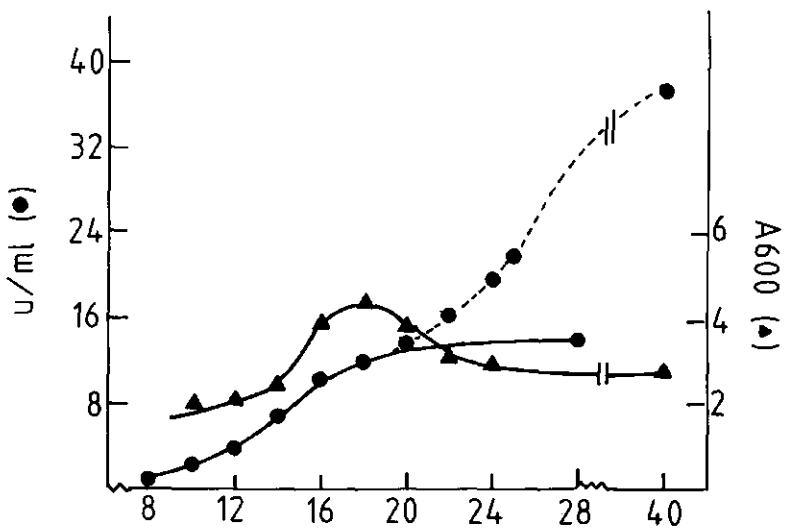


Figure 2. Expression of A.vinelandii E_2-activity in E.coli TG2(pRA282) during cell-growth. A colony was picked up from a TY-plate and grown for 8 hours in TY containing 75 µg/ml ampicillin and 20 µg/ml IPTG. 10 ml cultures were inoculated with 25 ul of the pre-culture and grown for variable times. The OD600 was determined (▲) and after preparation of cell-free extract E_2-activity was measured (●). The solid line (●—●) shows an experiment in which E_2-activity did not increase after 18-20 hours of growth, and is typical for most experiments performed; the dashed line (●--●) represents an experiment, in which activity increased after this time.

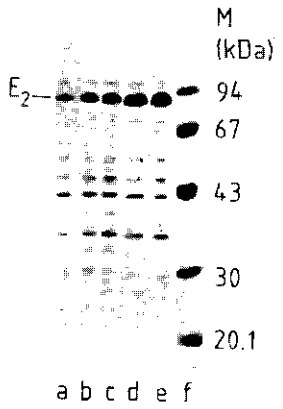


Figure 3. Expression of A.vinelandii E_2 protein in E.coli TG2(pRA282) during cell-growth. Cells were grown as described in the legends of Fig. 2; a culture was used that did not show an increase in E_2-activity after 18-20 hours of growth. Cell-free extracts were subjected to SDS-PAGE. The cell-free extracts were prepared from cells that had grown for 12 (lane a), 16 (lane b), 20 (lane c), 24 (lane d) and 40.5 hours (lane e). In lane f molecular weight markers are shown.

solid line in Fig. 2). In some experiments, however, the activity further increased more than twofold over the next 20 hours, as shown by the dashed line in Fig. 2. In earlier experiments, this dramatic increase in the expression of E_2-activity was never noticed, because E.coli strains with pRA282 were grown less than 16 hours.

We also studied whether the amounts of E_2 protein, produced by the E.coli clone, paralleled the changes in E_2-activity expressed by the clone. Therefore, samples of the respective E.coli clone were collected after different growth-times, and cell-free extracts of these clones were analysed by SDS-PAGE. Fig. 3 shows the electrophoretic profiles of cell-free extracts of these samples drawn from a growth-experiment, in which no increase in E_2-activity was found after 18-20 hours of growth (corresponding to the solid line in Fig. 2). Surprisingly, the amount of E_2 protein still increased after 18-20 hours of growth, as can be noticed from the increase in intensity of the E_2 protein band on gel, although the activity remained fairly constant. Apparently, after approximately 18 hours of growth, E_2 was still produced, but was not enzymatically active in this experiment. After we

succeeded in activating this inactive E_2 (see next section), from the ratio between the activities before and after activation we were able to estimate the relative amount of inactive E_2 that was produced. This amount varied between 10% and 90% in different experiments. Usually, the longer the recombinant bacteria were cultured, the more inactive protein was observed. Several reasons might exist for the accumulation of inactive E_2: a lack of an (unknown) cofactor, aggregation caused by the high protein concentration, or a wrong folding. Inactive E_2 was not readily pelletted after centrifugation of sonicated extracts, suggesting that it is not present as large aggregates, that are often found after production of high amounts of eukaryotic [17,18] or prokaryotic [19] proteins expressed from recombinant plasmids, and that might form separate entities, the so-called inclusion-bodies. It is not known, why some cultures accumulate much more inactive E_2 than others, although the same growth conditions were applied.

The inactive protein was tested by limited proteolysis for correct folding of the domains. Limited proteolysis with trypsin of E_2, as isolated from A.vinelandii, yields two stable domains: the lipoyl domain and the catalytic domain [3]. When cell-free extract of an E.coli TG2(pRA282) culture that contained a large

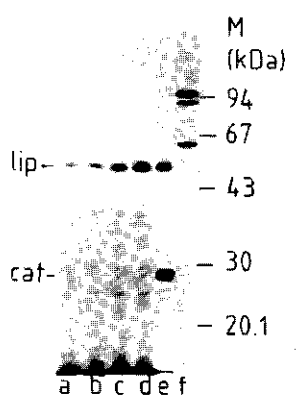


Figure 4. Limited proteolysis of A.vinelandii E_2 expressed in E.coli TG2(pRA282). Proteolysis with trypsin was carried out on cell-free extracts, prepared from cells that produced predominantly inactive E_2. Cells were grown as described in the legends of Fig. 2. Cell-free extracts were incubated with trypsin (0.2 µg/ml cell-free extract) at 0°C for 90 minutes, and analysed by means of SDS-PAGE. Lane a-d: digests from extracts from cells after 16,20,24 and 40.5 hours of growth, respectively. Lane e, E_2 isolated from A.vinelandii PDC, digested with trypsin, lane f, isolated A.vinelandii PDC.

amount of inactive E_2 was subjected to limited proteolysis with trypsin a usual amount of lipoyl domain was found (Fig.4), but hardly any catalytic domain could be observed by SDS-PAGE. This indicates that the lipoyl domain is folded correctly. The catalytic domain, on the other hand, is degradaded into small fragments upon proteolysis with trypsin. Obviously, this C-terminal part of the E_2-chain is unfolded or folded in an incorrect way, causing that lysyl- and arginyl residues are available for trypsin digestion.

Finally, a comment should be made on the late expression of E_2 protein and activity that was noticed in these experiments. One of the factors that might be responsible for this, could be the dependence of the E_2 gene from the vector-encoded lacZ-promoter; this promotor is known to be stimulated by guanosine-5'-diphosphate-3'diphosphate (ppGpp)[20]. The concentration of ppGpp is low at high growth rates and is highly raised in a process called "stringent respons" [21] in the stationary phase of the growth-cycle or when specific nutrients, like amino acids, are depleted.

Activation of inactive E_2 produced in E.coli

From the experiments described in the previous part, it appeared that in the inactive protein the catalytic domain was incorrectly folded. We undertook a series of experiments aimed at restoration of the native folding. Therefore, E_2 was first incubated in 2.5 M GdnHCl in order to obtain denaturation, followed by renaturation by removal of the GdnHCl under different conditions. Experiments were performed in cell-free extract because we were not able to purify the inactive protein. Coenzymes are known to have a positive influence on the rate and yield of reactivation and renaturation [22]. Therefore the lipoic acid cofactor and the substrate CoA, which both visit the active site in the catalytic domain, were added in the activation assay. The results of the activation experiments are shown in Table 1. After incubation with 2.5 M GdnHCl for 15 minutes at 0°C, transition of E2 to the unfolded state occurs, as has been observed by circular dichroism and kinetic measurements [4]. When a cell-free extract,

Table 1. Activation of inactive A.vinelandii E_2 expressed in E.coli TG2(pRA282).

method/sample	CFE	+L	+D	+G	+G+A	+G+L	+G+A+L	+G+D	+G+D+A
direct	100	100	ND	10	10	10	10	ND	ND
diluted	60	100	ND	20	30	20	40	ND	ND
dialysed	100	110	100	160	170	120	140	1030	1030

Cell-free extracts (=CFE) of E.coli TG2(pRA282) containing partially inactive A.vinelandii E_2 were incubated with 2.5 M GdnHCl (15 minutes, 0°C;=G), followed by addition of acetylCoA (0.5 mg.ml;=A), lipoic acid (1 mM;=L) and/or dithiothreitol (2 mM;=D). After incubation with GdnHCl, activation was carried out by different methods and activity was measured. Activities are shown as percentage of the blank, to which no GdnHCl is added. The terms direct, diluted and dialysed are explained in the text. ND= not done.

containing more than 90% inactive E_2, was denatured with 2.5 M GdnHCl, followed by dialysis to remove GdnHCl, the E_2-activity increased 1.6-fold (Table 1). Only a small amount of activity was found when the GdnHCl-denatured extract was brought to low GdnHCl concentration by dilution. The addition of CoA or lipoic acid did not result in a significant higher activity.

The addition of dithiothreitol in the activation assay resulted in a significant activation: a 10-fold higher activity was found. Upon addition of dithiothreitol to cell-free extract, which had not been incubated with GdnHCl, no increase in activity was observed. Obviously, inactive, incorrectly folded E_2 results from the formation of intermolecular or intramolecular S-S bridges. E_2 contains two (reduced) cysteine residues in the catalytic domain, that might be involved in the formation of erroneous S-S-bridges in the inactive protein.

Correct folding of the activated protein was checked by limited proteolysis. After limited proteolysis of the E_2-containing cell-free extract and analysis of the digestion products by SDS-PAGE, the catalytic domain could be detected, suggesting that it is correctly folded (Fig. 5). In the same sample an about 2 kDa

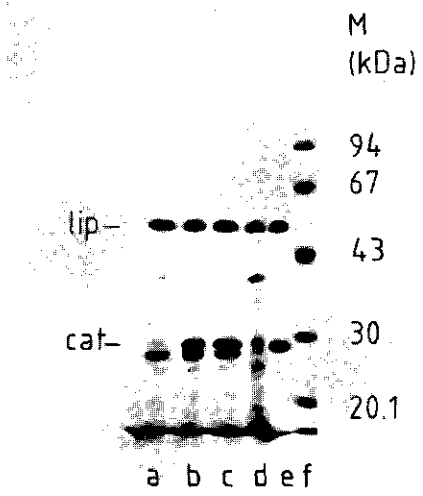


Figure 5. Limited proteolysis of A.vinelandii E_2 expressed in E.coli TG2(pRA282), after application of different activation procedures. Digestion products were subjected to SDS-PAGE. Cells were grown for 24 hours. Preparation of cell-free extracts (CFE), the activation procedures and the meaning of the abbreviations are described in the method-section or in the legends of Table 1. Digestion with trypsin was carried out for 90 minutes at 0°C. Lane a, CFE+G; lane b, CFE+G+D; lane c, CFE+G+D+A; lane d, CFE; lane e, E_2, isolated from A.vinelandii PDC and digested with trypsin, lane f, molecular weight markers.

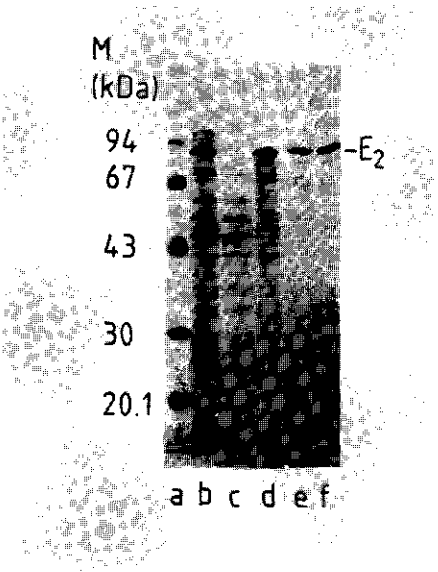


Figure 6. Purification of A.vinelandii E_2 expressed in E.coli TG2(pRA282). Samples from different stages of the isolation procedure of A.vinelandii E_2 expressed in E.coli TG2(pRA282) were subjected to SDS-gelelectrophoresis. Lane a, cell-free extract; lane b, supernatant after PEG/12 mM $MgCl_2$-precipitation; lane c, pellet after PEG/12 mM $MgCl_2$-precipitation; lane d, pool DEAE-Sephacel; lane e, pool Sephacryl-S300; lane f, molecular weight markers.

smaller protein-band was observed (Fig.5,lane 2) which also sometimes is observed after limited proteolysis of E_2 which has been stored for a long time, or after very long incubation with trypsin (days).

Isolation of A.vinelandii E_2 expressed in E.coli

E.coli TG2(pRA282) cells were grown as described in the methods, using conditions resulting in production of E_2 of which less than 10% is inactive. E_2 was isolated as described in the methods from 80-100 gram cells, in a three-step procedure. The results of a

Table 2. Purification of A.vinelandii E_2 from E.coli TG2(pRA282).

sample	volume	protein	specific activity	total units	yield
	(ml)	(mg)	(U/mg)	(U)	(%)
crude extract	325	9653	4.4	42150	100
PEG-precipitation	40	1044	38.9	40640	96
DEAE-Sephacel	275	591	66.1	39050	92
Sephacryl-S300	81	406	79.4	32240	76

The culture of E.coli TG2(pRA282), used for isolation, produced E_2 that was for more than 90% active.

purification are shown in Table 2 and Fig. 6. E_2 possessed a specific activity of 79 U/mg. This is much higher than the specific activity which was found before in E_2 resolved from isolated A.vinelandii PDC. A value of 12.6 U/mg has been reported [23], and in subsequent experiments a value of 30 U/mg was obtained (unpublished result). In the isolation described in this paper a more than twofold higher specific activity was found. The harsh conditions during resolution of the components from PDC (pH 9.4, 3M KBr) might be responsible for the lower activity obtained before, although previously no difference was found between activity of the isolated E_2 or of E_2 in the complex.
We obtained 406 mg of pure E_2 from 100 gram cells (Table 2). During the isolation of PDC from A.vinelandii, 280 mg PDC, corresponding to 70 mg E_2, was obtained from 250 gram cells [24]. Consequently, a 15-fold higher amount of E_2 can be obtained (per gram of cell wet-weight) from the E.coli clone compared to A.vinelandii. The E_2 is very soluble in standard buffer and concentrations of 1 mM are easily obtained. This makes A.vinelandii E_2 very suitable for NMR-experiments.
The lipoyl groups are attached to specific sites in the lipoyl domain [25]. In E_2 of PDC from A.vinelandii three potential lipoyl binding sites were indicated in the lipoyl domain [2]. In E_2 isolated from A.vinelandii 2-2.5 lipoyl groups were determined per E_2-chain [26,27], but about 50% could be excised without

affecting the activity, as was shown for E.coli E_2 by enzymatic removal of the lipoyl groups [28] and for mutant E.coli PDC, in which one or two of the lipoyl domains were deleted by genetic engineering [29,30]. Determination of the lipoyl groups in A.vinelandii E_2 produced in E.coli showed that only about one fifth of the potential lipoyl binding sites are actually occupied by lipoyl groups. Obviously, E.coli is able to incorporate lipoyl groups in A.vinelandii E_2, but the efficiency is very low. It is not exactly known how lipoyl groups are attached to the binding sites, and by what factors this binding is influenced. It cannot be excluded that incorporation of these moieties requires a specific mechanism, that is not very efficient towards the heterologous A.vinelandii E_2. Also the high synthesis rates of E_2 or the high amounts of E_2 produced might influence the effectiveness of cofactor incorporation, if the amount of cofactor produced cannot cope with the amount of E_2 produced. This latter factor might certainly be relevant: also in an E.coli clone that produced high amounts of E.coli lipoyl domain from a recombinant plasmid, both lipoylated and non-lipoylated domains were observed [31]. Further research will be carried out in which the incorporation of lipoyl groups is determined at different stages of the growth.
Despite the small amount of incorparated lipoyl groups the lipoyl domain is correctly folded, as judged by the relative resistance to tryptic digestion. Obviously, this cofactor is not essential for correct folding.

Separate expression of the E_2 catalytic and lipoyl domains from plasmids pRC1 and pRL1

Previous experiments have demonstrated that the two functional domains of E_2 can be considered as two separate and stable structural domains. These two domains are relatively resistant to trypsin digestion and are connected by a short region that is sensitive to trypsin digestion, and is thought to possess the E_1- and E_3 binding sites. Here we describe experiments aimed at the separate expression of these domains from recombinant plasmids

that only contain the genetic information for either the lipoyl or the catalytic domain of E_2. Therefore pRA282, which has a 2.8 kbp insert, was digested with endonuclease RsrII, that has a unique restriction site around residue 1534 (Fig.1), in the middle of the region (of the gene) that is proposed to contain the information for the E_1- and E_3 binding sites. Using the HindIII-site of the multiple cloning site of the vector, plasmid pRC1 is constructed by a deletion of the RsrII-HindIII fragment. After deletion of the ±1540 residues, the resulting plasmid still has the genetic information for the catalytic domain and part of the proposed E_1- and E_3 binding region, but has lost the information of the lipoyl domain and the start codon, GTG. Translation might start, however, from the in-frame AUG start codon for the vector-encoded-β-galactosidase, resulting in the expression of the catalytic domain that is extended at the N-terminus with the six N-terminal residues of β-galactosidase and 20 residues of the proposed E_1- and E_3 binding region (=E_2cat+). Using the EcoR1-site of the multiple cloning site of the vector, plasmid pRL1 is constructed by a deletion of the RsrII-EcoRI fragment. The resulting plasmid pRL1, after deletion of the ±1260 bp, still has the genetic information for the lipoyl domain and part of the proposed E_1- and E_3 binding region, but sequences for the catalytic domain and the stop codon are lost. Translation might terminate at an in-frame stop codon, found 30 nucleotides downstream of the truncated E_2-gene in the vector, resulting in the expression of the lipoyl domain that has a 10 residues long C-terminal extension of vector-encoded residues and 14-30 residues of the proposed E_1- and E_3 binding region (=E_2lip+). E_2cat+ and E_2lip+ are produced in low amounts in E.coli TG2(pRC1) or E.coli TG2(pRL1), respectively. They can hardly be seen on SDS-gel (Fig. 7A) but after Western blotting the proteins could be observed after reaction with antibodies specific for the homologous domains of E_2 (Fig. 7B,C). E_2cat+ had, as expected, an apparent molecular mass of 30 kDa. Also a ± 2 kDa smaller protein band was observed, having the same molecular mass as the catalytic domain obtained after digestion of E_2 with trypsin. Probably the region around the original claevage site is still easily digested. E_2lip+ had the same apparent molecular mass (52 kDa) as the lipoyl domain obtained after digestion of E_2 with

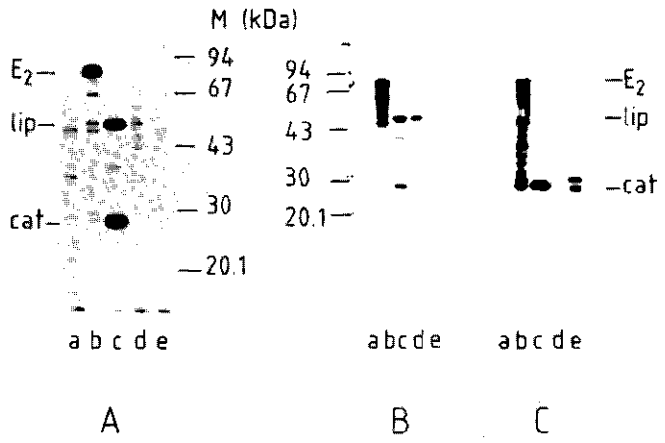


Figure 7. Expression of A.vinelandii lipoyl domain and catalytic comain (E_2lip+ and E_2cat+) in E.coli TG2(pRL1) and TG2(pRC1). A. SDS-PAGE, stained with Coomassie brilliant blue R. B. Western-blot, incubated with antiserum against the tryptic lipoyl domain. C. Western-blot, incubated with antiserum against the tryptic catalytic domain. Lane a, cell-free extract of E.coli TG2(pUC9); lane b, E_2, isolated from A.vinelandii PDC, partially degraded during storage); lane c, E_2, isolated from A.vinelandii PDC, and digested by trypsin at 0°C for 90 minutes; lane d, cell-free extract of E.coli TG2(pRL1); lane e, cell-free extract of TG2(pRC1). In this experiment E_2cat+ and E_2lip+ were expressed in E.coli TG2; later experimets showed a much higher level of expression in the Δlon E.coli strain Y1083.

trypsin, although a 2-3 kDa higher molecular mass was expected. Obviously, a part of the protein is cleaved off by cellular proteases, probably around the original tryptic cleavage site. To diminish the possibility of cellular cleavages, expression of the recombinant plasmids was also studied in E.coli MC1061, a Δlon strain that is deficient in the lon-protease, thereby reducing the degradation of expressed proteins [10]. A high production of both E_2cat+ and E_2lip+ were observed: clearly visible bands were observed when cell-free extracts of the clones were subjected to SDS-PAGE (data not shown). The E_2cat+ had an apparent molecular mass of 30 kDa, the E2lip+ one of 52 kDA, the same as in E.coli strain TG2. The much higher production suggests that the proteins are not very stable in E.coli TG2 . Obviously, the domains of E2 are not as stable as the whole enzym. E_2cat+ has been isolated using the same method as used for isolation of E_2 (data not shown). A corresponding high specific activity was obtained: 130 U/mg (in accordance with 55 U/mg E_2), showing that transacetylase activity is fully retained in the domain.

Acknowledgements

We wish to thank Mr. M.M. Bouwmans for drawing the figures and Dr. W. van Dongen for advice and the critical reading of the manuscript. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid of the Netherlands Organisation for the advancement of Scientific Research (NWO).

References

1. Koike, M., Reed, L.J. & Caroll, W.R. (1963)
   J.Biol.Chem. 238, 30-39
2. Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988)
   Eur.J.Biochem., in press
3. Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987)
   Eur.J.Biochem. 169, 245-252
4. Hanemaaijer, R., Westphal, A.H., van der Heiden, T., de Kok, A. & Veeger, C. (1988)
   Eur.J.Biochem., in press
5. Gibson, T.J. (1984)
   Ph.D.Thesis, Cambridge
6. Casadaban, M. & Cohen, S. (1980)
   J.Mol.Biol. 138, 179
7. Young, R.A. & Davis, R.W. (1983)
   Proc.Natl.Acad.Sci.USA 80, 1194
8. Radloff, R., Bauer, W. & Vinograd, J. (1967)
   Proc.Natl.Acad.Sci.USA 57, 1514-1521
9. Birnboim, H.C. & Doly, J. (1979)
   Nucleic Acids Res. 7, 1513-1524
10. Westphal, A.H. & de Kok, A. (1988)
    Eur.J.Biochem. 172, 299-305
11. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982)
    "Molecular Cloning: A Laboratory Manual" Cold Spring Harbor Lab., Cold Spring Harbor, New York
12. Voordouw, G., Walker, J.E. & Brenner, S. (1985)
    Eur.J.Biochem. 148, 509-514
13. Bresters, T.W. de Abreu, R.A., de Kok, A., Visser, J. & Veeger, C. (1977)

Eur.J.Biochem. 59, 335-345

14. De Kok, A., Kornfeld, S., Benziman, M. & Milner, Y. (1980)
Eur.J.Biochem. 106, 49-58

15. Vieira, J. & Messing, J. (1982)
Gene 19, 259-268

16. Miller, J.H. (1977)
"Experiments in Molecular Genetics" Cold Spring Harbor Lab.,
Cold Spring Harbor, New York

17. Marston, F.A.O., Lowe, P.A., Doel, M.T., Schoemaker, J.M.,
White, S. & Angal, S. (1984)
Bio/Technology 2, 800-804

18. Sekine, S., Mizukami, T., Nishi, T., Kuwana, Y., Saito, A.,
Sato, M., Itoh, S. & Kawauchi, H. (1985)
Proc.Natl.Acad.Sci.USA 82, 4306-4310

19. Chang, T.W., Kato, I., McKinney, S., Chanda, P., Barone,
A.D., Wong-Staal, F., Callo, R. & Chang, N.T. (1985)
Bio/Technology 3, 905-909

20. Primakoff, P. (1981)
J.Bacteriol. 145, 410-416

21. Ryals, J., Little, R. & Bremer, H. (1982)
J.Bacteriol. 151, 1261-1268

22. Krebs, H., Rudolph, R. & Jaenicke, R. (1979)
Eur.J.Biochem. 100, 359-364

23. De Graaf-Hess, A.C. & de Kok, A. (1982)
FEBS Lett. 143, 261-264

24. Bosma, H.J. (1984)
Ph.D. Thesis, Wageningen

25. Hale, G. & Perham, R.N. (1980)
Biochem.J. 187, 905-908

26. Bosma, H.J., de Graaf-Hess, A.C., de Kok, A., Veeger, C.,
Visser, A.J.W.G., & Voordouw, G. (1982)
Ann.N.Y.Acad.Sci. 378, 265-285

27. De Kok, A., Visser, A.J.W.G. & de Graaf-Hess, A.C. (1982)
In "Flavins and Flavoproteins" (Massey, V. & Williams, C.H.,
eds.) pp. 61-67, Elsevier, Amsterdam

28. Bleile, D.M., Munk, P., Oliver, R.M. & Reed, L.J. (1979)
Proc.Natl.Acad.Sci.USA 76, 4385

29. Guest, J.R., Lewis, H.M., Graham, L.D., Packman, L.C. &

Perham, R.N. (1985)
J.Mol.Biol. 185, 743-754

30. Graham, L.D., Guest, J.R., Lewis, H.M., Miles, J.S., Packman, L.C., Perham, R.N. & Radford, S.E. (1986)
Philos.Trans.R.Soc.London A 317, 391-404

31. Miles, J.S. & Guest, J.R. (1987)
Biochem.J. 245, 869-874

CHAPTER 5

TIME-RESOLVED FLUORESCENCE STUDIES ON THE DIHYDROLIPOYL TRANSACETYLASE (E_2) COMPONENT OF THE PYRUVATE DEHYDROGENASE COMPLEX FROM AZOTOBACTER VINELANDII

Roeland Hanemaaijer, Remco Masurel, Antonie J.W.G. Visser, Arie de Kok and Cees Veeger

SUMMARY

The dihydrolipoyl transacetylase (E_2) component of A.vinelandii PDC and its lipoyl domain shows similar dynamic properties as revealed with fluorescnece anisotropy decay of lipoyl-bound IAANS. The lipoyl domain (32.6 kDa), containing three almost identical subdomains shows a mode of rotation characteristic for a protein of about 30 kDa. A similar rotation is found in E_2, indicating an independent rotational mobility of the whole domain in the multimeric E_2-core (M_r = 1.6 MDa). No independent rotation of a single lipoyl subdomain (M_r = 10 kDa) is observed. The E_1 component, in contrast to the E_3 component, shows interaction with the lipoyl domain.

1. INTRODUCTION

The pyruvate dehydrogenase complex catalyzes the oxidative decarboxylation of pyruvate to acetylCoA. The structural core of the A.vinelandii complex is composed of four dihydrolipoyl transacetylase (E_2) chains, to which three dimers of pyruvate dehydrogenase (E_1) and one dimer of lipoamide dehydrogenase (E_3) are bound [1]. After removal of these peripheral components the E_2-core aggregates to a cubic 24-meric structure [2,3]. The substrates are transferred between the different components by lipoyl-lysine residues, which act as swinging arms between the different active sites [4]. In E.coli and in A.vinelandii E_2 the N-terminal part of the E_2-chain consists of three homologous repeating sequences [5,6]. Each repeat of about 80 amino acid residues contains a lysyl residue that is a potential site for lipoylation. Each repeat is separated from its neighbour by a region of about 20 residues, very rich in alanyl- and prolyl residues. After limited proteolysis of E_2 with trypsin this N-terminal part, called lipoyl domain (32.6 kDa), is separated from a part called the catalytic domain, which forms the multimeric core [7]. In E.coli it is shown that after limited proteolysis with Staphylococcus aureus V8 proteinase the three repeats can be isolated separately as folded subdomains [8]. In other organisms such as gram-positive bacteria and eukaryotes, and also in the closely related 2-oxoglutarate dehydrogenase complex, only one lipoyl subdomain is found [9-11]. Spin label experiments with the PDC from E.coli have demonstrated that the dithiolane ring of the lipoyl group can rotate freely in the complex as is indicated by a rotational correlation time of 0.2-1.0 ns [12]. Also a correlation time of more than 50 ns was found, which was attributed to the rotation of the whole complex (expected corre-

lation time 2 μs). When using the triplet probe eosin-maleimide only mobility of the whole complex was found, showing that the label is folded back to the protein surface [13]. From a lack in energy transfer between the lipoyl domain and the FAD-group of E_3 it is suggested that the lipoyl-lysine residue is too short to serve all catalyic centres [14] and therefore at least a part of the protein chain around the lipoyl residue is thought to be flexible. Direct evidence for conformational flexibility in E_2-chains was obtained from ^{1}H-NMR spectroscopy [15]. In these experiments the regions, which are very rich in alanyl-and prolyl residues, are shown to possess conformational flexibility. Thus, from these experiments it has been suggested that all three subdomains in the lipoyl domain can rotate independently. In this paper it is shown with fluorescence anisotropy decay experiments that, although the alanine-proline rich region may possess internal flexibility, the lipoyl domain moves as a single entity within the large multimeric E_2-core. This movement is restricted by the binding of the peripheral components.

2. MATERIALS AND METHODS

2-(4,-iodoacetamidoanilino)-naphtalene-6-sulfonic acid (IAANS) was obtained from Molecular Probes. Dihydrolipoyl transacetylase (E_2) was isolated from the pyruvate dehydrogenase complex by covalent chromatography on thiolsepharose 4B as described previously [16], with modifications according to [7]. The lipoyl domain was obtained after limited proteolysis with trypsin of the E_2 component, which was covalently bound on thiolsepharose 4B and purified as described before [7]. Labelling was carried out as follows. E_2 and the lipoyl domain were incubated for 30 minutes at 4°C in 20 mM Tricine pH 8.5 containing 20 mM dithiotreitol, and anaerobically dialysed against standard buffer (50 mM potassium phosphate pH 7.0 containing 0.5 mM EDTA and 0.1 mM phenylmethylsulfonylfluoride). From a 15.9 mM IAANS stock-solution in H_2O/ethanol (1:1 v/v) a fifteen times molar excess to sulfhydryl-groups of the lipoyl-moieties was added, and after 30 minutes the sample was dialysed against standard buffer. From optical density measurements it was observed that over 80% of the lipoyl groups were (doubly) labelled. Since it has been shown [2] that only 2 - 2.5 of the three potential sites per E_2-chain contain a lipoyl group, it is clear that all present lipoyl groups were (doubly) labelled. No bound label

was detected without prior reduction. After labelling the complex activity had totally disappeared, but the E_2 (transacetylase) activity was fully retained. Fluorescence spectra were recorded on an Aminco SPF-500 fluorimeter. Time-resolved fluorescence decay was measured with a system consisting of a frequency-doubled synchronously pumped dye laser for excitation and time-correlated single photon counting in detection. The excitation wavelength was at 310 nm and the emission was monitored via a 450 nm band-pass filter (Balzers K45). Details of the experimental set-up have been described in [17]. All experiments were carried out at 20°C. Data analysis was performed as described in [18]. The order parameter S_1 is defined as $[(\beta_2+\beta_3)/(\beta_1+\beta_2+\beta_3)]^{\frac{1}{2}}$ [19]. The parameter S_2, representing the degree of order only connected with the very slow rotation is defined as $[(\beta_3/\beta_2+\beta_3)]^{\frac{1}{2}}$. The related cone angle θ_c for the order parameter S_1 is obtained from the relation $(S_1)^2 = \frac{1}{2}\cos\theta_c(1+\cos\theta_c)$.

3. RESULTS AND DISCUSSION

The initial fluorescence anisotropy decay curves are shown in Figs. 1 and 2. Fig. 1 shows the experimental and the fitted curve of the fluorescence anisotropy decay of E_2, in Fig. 2 only the fitted curves of the lipoyl domain, E_2, and E_2 with bound E_1- and E_3 components are shown. For the lipoyl domain the anisotropy decays as a double exponential function with a short (0.6 ns) and a longer (11.4 ns) characteristic time constant. The short component can be ascribed to the motion of the lipoylgroup rotating freely around the linkage with the protein chain. This time constant correlates well with the values found previously with spin labels [12]. The longer component can be ascribed to the rotation of the whole domain. The correlation time can be calculated on the basis of an empirical formula relating the correlation time Φ with M_r of a hydrated, spherical polypeptide at 20°C: $\Phi(ns) = 3.84 \times 10^{-4} M_r$ [20]. From this formula a correlation time of about 12.5 ns is calculated for the lipoyl domain having M_r = 32.6 kDa [7]. Despite the fact that the lipoyl domain is thought to have a swollen or extended structure, f/f_0 = 1.7 [7], the agreement is rather good. For E_2 the anisotropy decay curve can be described as a triple exponential function with a short (0.4 ns), a longer (10.7 ns) and a very long (600 ns) time constant. The latter is fixed in the fitting procedure and represents the rotation of the whole protein (1.6 MDa). The short component is in the same order as found in the lipoyl

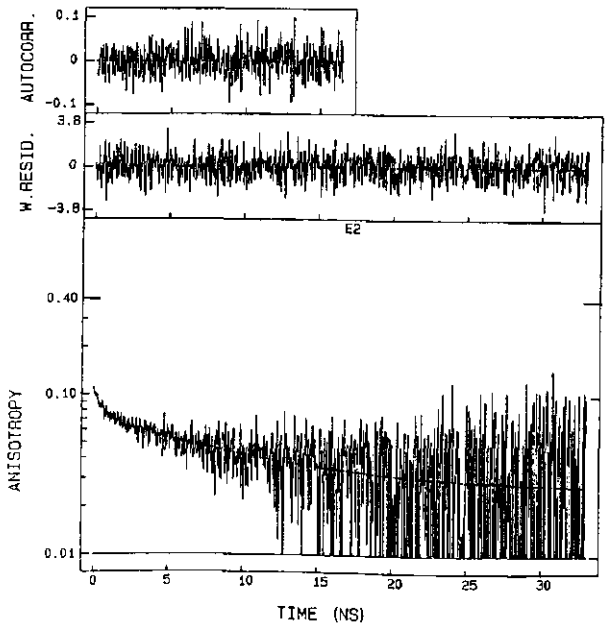


Figure 1. Fluorescence anisotropy decay of E_2. Shown are two curves: the experimental (noisy curve) and the calculated (smooth curve) fluorescence anisotropy decay. The parameters of the triple exponential decay are listed in Table 1. The quality of the fit is indicated by the weighted residuals and the autocorrelation function, shown in the upper panels. The statistical parameters are χ^2 = 1.06, and the Durbin-Watson parameter = 2.03.

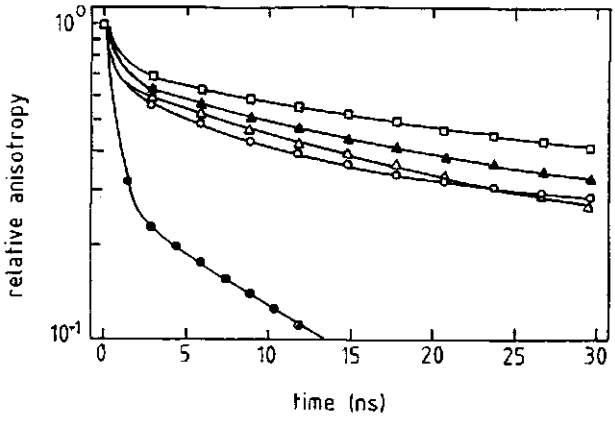


Figure 2. Fluorescence anisotropy decay of the lipoyl domain, E_2, and E_2 complexed with E_1 or E_3. Shown are the fitted curves of a) the lipoyl domain (●), b) E_2 (○), c) E_2 + E_3 (▲) in a 4:2 molar concentration, when the E_3 binding site is saturated by E_3, d) E_2 + E_3 (△) in a 4:8 molar concentration, when both the E_3- and E_1 binding sites are saturated by E_3, and e) E_2 + E_1 (□) in a 4:6 molar concentration, when the E_1 binding sites are saturated by E_1. The parameters af the double (a) and the triple (b-e) exponential decays and the standard errors derived from the fits are listed in Table 1.

domain and accounts for the free rotation of the lipoyl group. The time constant of 10.7 ns will represent the rotation of the lipoyl domain, which is thought to rotate independently of the large E_2-core. From the order parameter S_1 and the related cone angle θ_c that can be derived from a wobbling-in-cone model [19] it is shown that in E_2, where the lipoyl domain is bound to the core, the order increases and the motion of the label is more restricted. This indicates that the motion of the lipoyl domain is not totally independent of the core, but a certain degree of interaction exists. Upon binding of E_1- or E_3 components to the E_2-core no significant difference in correlation times is found (table 1,exps. 3-5). From β_3, S_1 and S_2 it is clear that, despite the dissociation of E_2 which takes place upon addition of E_1 (or excess E_3), a significant difference is found in order and in motional restriction upon binding of the E_1 component (exp. 5). Upon addition of excess E_3 (exp. 4), which is thought to bind on or near the E_1 binding site [1] no significant difference is observed. Obviously some interaction exists between E_1 and the lipoyl domain which is not present between E_3 and the lipoyl domain. This could be related to the observation of Packman et al. [8] who showed that E_3 is able to use lipoamide as a substrate, whereas E_1 needs the intact lipoyl subdomain as substrate. Mobility of lipoyl subdomains has been suggested from ^{1}H-NMR experiments, based on observed conformational mobility of the alanine and proline rich regions. In these fluorescence anisotropy experiments we present a direct indication that the lipoyl domain as a whole contains mobility more or less independent of the large E_2-core. The hinge of this mobility should be located between the lipoyl domain and the catalytic domain (residue 331-381 [6]). In this region also the binding sites for the E_1- and the E_3 components are located [7]. After binding of these components to the E_2-core no dramatic limitation of mobility is observed, suggesting that the hinge of the mobility of the lipoyl domain is located N-terminal of the E_1- and E_3 binding sites. Previous fluorescence anisotropy experiments of the FAD in free and in bound E_3 have shown that E_3, bound to the E_2-core, still possesses a high mobility [21]. This indicates a second hinge C-terminal of the E_3 binding site. This region may correspond with a high mobile region observed in ^{1}H-NMR experiments near the N-terminus of the catalytic domain [22].

Table 1. Fluorescence decay parameters of the lipoyl domain, E_2 and E_2 complexed with E_3 or E_1.

Exp.	Protein[a]	M_r (kDa)	β_1	ψ_1 (ns)	β_2	ψ_2 (ns)	β_3	ψ_3[b] (ns)	S_1[c]	θ_c[d]	S_2[e]
1	Lipoyl domain	33	0.160±0.011	0.62±0.19	0.039±0.002	11.4±1.5	-	-	0.44±0.03	56±2	-
2	E_2	1560	0.069±0.002	0.38±0.07	0.048±0.002	10.7±1.4	0.024±0.003	600	0.74±0.04	36±3	0.57±0.03
3	E_2+E_3 (4:2)	2160	0.042±0.004	0.91±0.13	0.051±0.006	18.1±3.7	0.025±0.006	830	0.80±0.11	31±8	0.57±0.11
4	E_2+E_3 (4:8)	660	0.055±0.008	0.56±0.08	0.060±0.005	16.9±2.5	0.026±0.005	253	0.78±0.10	32±8	0.55±0.08
5	E_2+E_3 (4:6)	860	0.041±0.004	1.00±0.14	0.047±0.006	20.0±4.7	0.055±0.007	330	0.84±0.10	27±9	0.73±0.09

a) See legends fig. 2 for details.

b) Fixed in the analysis.

c) From $(S_1)^2 = \dfrac{\beta_2 + \beta_3}{\beta_1 + \beta_2 + \beta_3}$.

d) From $(S_1)^2 = \frac{1}{4}\cos\theta_c(1+\cos\theta_c)$.

e) From $(S_2)^2 = \dfrac{\beta_3}{\beta_2 + \beta_3}$.

ACKNOWLEDGEMENTS

We like to thank Ing. A. van Hoek for assistance in the fluorescence decay experiments and Mrs. J.C. Toppenberg-Fang for the preparation of the manuscript. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid from the Netherlands Organisation for the advancement of Pure Research (NWO).

REFERENCES

1. Bosma, H.J., de Kok, A., Westphal, A.H. and Veeger, C. (1984) Eur.J.Biochem. 142, 541-549
2. Bosma, H.J., de Kok, A., van Markwijk, B.W. and Veeger, C. (1984) Eur.J.Biochem. 140, 273-280
3. Hanemaaijer, R., Westphal, A.H., van der Heiden, T., de Kok, A. and Veeger, C. Submitted to Eur.J.Biochem.
4. Koike, M., Reed, L.J. and Caroll, W.R. (1963) J.Biol.Chem. 238, 30-39
5. Stephens, P.E., Darlison, M.G., Lewis, H.M. and Guest, J.R. (1983) Eur.J.Biochem. 133, 481-489
6. Hanemaaijer, R., Janssen, A., de Kok, A. and Veeger, C. (1988) Eur.J.Biochem., in press
7. Hanemaaijer, R., de Kok, A., Jolles, J. and Veeger, C. (1987) Eur.J.Biochem. 169, 245-252
8. Packman, L.C., Hale, G. and Perham, R.N. (1984) EMBO J. 3, 1315-1319
9. Duckworth, H.W., Jaenicke, R., Perham, R.N., Wilkie, A.O.M., Finch, J.T. and Roberts, G.C.K. (1982) Eur.J.Biochem. 124, 63-69
10. White, R.H., Bleile, D.M. and Reed, L.J. (1980) Biochem.Biophys.Res.Commun. 94, 78-84
11. Spencer, M.E., Darlison, M.G., Stephens, P.E., Duckenfield, I.K. and Guest, J.R. (1984) Eur.J.Biochem. 141, 361-374
12. Ambrose-Griffin, M.C. and Griffin, W.G. (1984) Biochim.Biophys.Acta 789, 87-97
13. Visser, A.J.W.G., Scouten, W.H. and Lavalette, D. (1981) Eur.J.Biochem. 121, 233-235
14. Oliver, R.M. and Reed, L.J. (1982) in Electron Microscopy of Proteins (Harris, R., ed), Vol I, pp. 1-48, Academic Press, London
15. Texter, F.L., Radford, S.E., Laue, E.D., Perham, R.N., Miles, J.S. and Guest, J.R. (1988) Biochemistry 27, 289-296

16. De Graaf-Hess, A.C. and de Kok, A. (1982) FEBS Lett. 143, 261-264
17. Van Hoek, A. and Visser, A.J.W.G. (1985) Anal.Instrum. 14, 359-378
18. Vos, K., van Hoek, A. and Visser, A.J.W.G. (1987) Eur.J.Biochem. 165, 55-63
19. Lipari,G. and Szabo, A. (1980) Biophys.J. 30, 489-506
20. Visser, A.J.W.G., Penners, N.H.G. and Muller, F. (1983) in Mobility and Recognition in Cell Biology (Sund, H. & Veeger, C. eds) pp.137-152, Walter de Gruyter, Berlin
21. Grande, H.J., Visser, A.J.W.G. and Veeger, C. (1980) Eur.J.Biochem. 106, 361-369
22. Radford, S.E., Laue, E.D., Perham, R.N., Miles, J.S. and Guest, J.R. (1987) Biochem.J. 247, 641-649

CHAPTER 6

MOBILE SEQUENCES IN THE PYRUVATE DEHYDROGENASE COMPLEX, THE E_2 COMPONENT, THE CATALYTIC DOMAIN AND THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX OF AZOTOBACTER VINELANDII, AS DETECTED BY 600 MHz ^{1}H-NMR SPECTROSCOPY

Roeland Hanemaaijer, Jacques Vervoort, Adrie Westphal, Arie de Kok and Cees Veeger

Summary

600 MHz ^{1}H-NMR spectroscopy demonstrates that the pyruvate dehydrogenase complex of Azotobacter vinelandii contains regions of the polypeptide chain with intramolecular mobility. This mobility is located in the E_2 component and can probably be ascribed to alanine-proline rich regions that link the lipoyl subdomains to each other as well as to the E_1 and E_3 binding domain. In the catalytic domain of E_2, which was thought to form a compact, rigid core, also conformational flexibility is observed. It is conceivable that the N-terminal region of the catalytic domain, containing many alanine residues, is responsible for the observed mobility. In the low-field region of the ^{1}H-NMR spectrum of E_2 specific resonances are found, which are not present in the corresponding part of the ^{1}H-NMR spectrum of the catalytic domain. These resonances can be ascribed to phenylalanine, histidine and/or tyrosine residues. These residues are not present in the lipoyl domain, and therefore the resonances must arise from mobile residues located in the E_1 and E3 binding domain that links the lipoyl domain to the catalytic domain. In the ^{1}H-NMR spectrum of the intact complex these resonances are not clearly visible anymore, indicating a decreased mobility of the E_1 and E_3 binding domain.

Introduction

The pyruvate dehydrogenase complex catalyzes the following reaction:

Pyruvate + CoA + NAD^+ ----> AcetylCoA + CO_2 + NADH + H^+

It is composed of multiple copies of three enzymes: pyruvate dehydrogenase (E_1), dihydrolipoyl transacetylase (E_2) and lipoamide dehydrogenase (E_3). The E_2 component forms a structural core to which E_1 and E_3 components are bound non-covalently. During catalysis the substrate is carried in a thioester linkage by lipoyl-lysine residues that are part of the E_2 component and act as "swinging arms" while transferring the substrates,acetyl groups and reduction equivalents, between the active sites of the different components [1].

The domain structure of E_2 has been examined by limited

proteolysis [2]. After digestion of E_2 with trypsin a lipoyl domain and a catalytic domain are obtained. The lipoyl domain carries the lipoyl groups. The catalytic domain possesses the transacetylase active site and the E_2 intersubunit binding sites, showing the same 24-meric structure as intact E_2 [3]. The E_1 and E_3 binding sites are lost during proteolysis. They are thought to be located on a ±6 kDa region that links the catalytic domain to the lipoyl domain.

The gene encoding E_2 of the pyruvate dehydrogenase complex of Azotobacter vinelandii has been cloned and sequenced [4]. From the amino acid sequence derived from the DNA sequence it is shown that the lipoyl domain accounts for the N-terminal half of the polypeptide chain and comprises three highly homologous structures, each carrying a potential lipoyl binding site. In Escherichia coli it is shown that these repeats exist as three independent folded entities that retain their function as substrate for reductively acetylation by the E_1 component [5]. They are linked to each other and to the E_1 and E_3 binding domain by means of long regions (±20-30 residues) that are very rich in alanine and proline residues [6].

The active sites of the different components are at least 4 nm apart in the complex, too far away for a single lipoyl-lysine group [7,8]. Therefore it has been suggested that the lipoyl domain might be sufficiently flexible to facilitate movement of the lipoyl group between the different active sites. Strong evidence in favor of the existence of conformationally flexible regions in 2-oxoacid dehydrogenase complexes from E.coli [9,10], Bacillus staerothermophilus [11,12] and ox heart [13] was obtained by using ^{1}H-NMR spectroscopy. Unusual sharp resonances in the 270 and 400 MHz ^{1}H-NMR spectra of the intact complexes are thought to originate from alanine-proline rich regions in the E_2 chains. In E.coli it has been suggested that the mobility of these regions is responsible for the independent mobility of the three repeats in the lipoyl domain [14,15]. Mobility of the lipoyl domain has also been detected by time-resolved fluorescence spectroscopy [16]. After labelling of the lipoyl groups of E_2 from A.vinelandii with a fluorophore mobility is observed of the labelled lipoyl group, of the whole complex and

of the lipoyl domain. No independent mobility of the three repeating units has been observed.
After cloning in E.coli a high expression of the gene, encoding E_2 was obtained [17]. Both E_2 and the catalytic domain, obtained by limited proteolysis, showed to be soluble at high protein concentrations and we were able to carry out 600 MHz ^{1}H-NMR experiments.

Materials and methods

Materials

Tosylphenylananylchloromethane-treated trypsin was obtained from Worthington and phenylmethylsulphonyl fluoride was from Sigma. Deuterium oxide (99.8%) was purchased from Ega-chemie.

Isolation of the proteins

The pyruvate dehydrogenase complex was isolated from Azotobacter vinelandii as described by Bosma et al. [18]. The 2-oxoglutarate dehydrogenase complex was isolated from A.vinelandii as described in [19]. The E_2 component of the A.vinelandii pyruvate dehydrogenase complex was isolated from Escherichia coli, in which the gene encoding A.vinelandii E_2 was cloned [17]. The catalytic domain, obtained by limited proteolysis of A.vinelandii E_2 with trypsin, was isolated by gelchromatography, using FPLC equipped with a Superose 6 K26/70 column (Pharmacia), and concentrated by ultrafiltration using an Amicon YM-100 membrane.

^{1}H-NMR spectroscopy

Protein samples for ^{1}H-NMR spectroscopy (in standard buffer: 50 mM potassium phosphate, pH 7.0 containing 0.1 mM EDTA and 0.1 mM phenylmethylsulphonyl fluoride) were concentrated to about 0.5-1.0 mM. ^2H_2O was added to a final concentration of 10%(v/v). ^{1}H-NMR spectra were obtained at 600 MHz with a Bruker AM-600 spectrometer using a 16.7 kHz spectral width, 30° pulses and a repetition time of 1 second. 8K data points were used. Quadrature detection and quadrature phase cycling were used. Chemical shifts are relative to trimethylsilylpropionate (TSP). Wilmed 5mm NMR-tubes were used. The solvent resonance irridation (18 dB

attenuation from 0.2 Watt) was applied at all times, except during data aquisition. The transmitter frequence was placed exactly at the waterresonance. The sample temperature was 23°C. The NMR-data were transferred to a MicroVAX II and evaluated with (modified) software, kindly given to us by dr. R. Boelens. The FID's were zerofilled once before Fourier transformation (without any window function). The resulting frequency specrum was manipulated by a fourth order polynome for baseline correction.

Other techniques

Protein concentrations were measured according the method of Lowry [20]. A stock solution of 1M phenylmethylsulphonyl fluoride was prepared in $C^2H_3O^2H$ and diluted 100-fold in standard buffer.

Results

High-field resonances

In ^{1}H-NMR spectra of large protein complexes (MDa) in general no distinct resonances can be observed; the calculated linewidth of a methylene proton resonance is in the range of kHz. However, in the 600 MHz spectrum of the pyruvate dehydrogenase complex (Fig. 1A), superimposed on the expected broad component in the spectrum, many sharp resonances are observed (linewidth in the range of 10-100 Hz). These resonances must arise from a region or regions of the polypeptide chain having substantial mobility. The major sharp resonance occurs at 1.39 ppm which, on basis of the chemical shift, can be ascribed to the methyl side chains of alanine residues. In the intact pyruvate dehydrogenase complex of E.coli a similar sharp resonance has been found [10]. It was ascribed to the alanine-proline rich regions of the E_2 component, located at the links between the repeating units in the lipoyl domain and between the lipoyl domain and the E_1 and E_3 binding domain. Indeed, in the spectrum of E_2 (Fig. 1B) the same sharp resonance is found. Also many sharp resonances are found between 1.5 and 2.5 ppm, which partly can be ascribed to mobile proline residues present in the same regions. In the high-field spectrum of OGDC (Fig. 1D) a similar pattern of sharp resonances is observed, suggesting a similar mobile region. In the E_2 sequence of A.vinelandii OGDC also a alanine-proline rich region has been

found between its single lipoyl domain and the proposed E_1 and E_3 binding domain (unpublished results).
Two additional resonances of ethanol can be observed, which are due to a contamination. The concentration of ethanol is estimated to be around 5 - 10 mM

Low-field resonances in E_2

In the aromatic region of the spectrum of E_2 (Fig.1B) two specific sharp resonances (7.1 and 7.3 ppm) are observed which are not present in the spectrum of PDC (Fig.1A) or in that of the catalytic domain (Fig.1C). The intensity of these residues is about 3% of the alanine methyl resonance, which indicates that they can be assigned to either one or two aromatic residues, i.e. to phenylalanine, histidine or tyrosine. These amino acid residues are not present in the lipoyl domain of E_2 (Fig.2) [4].

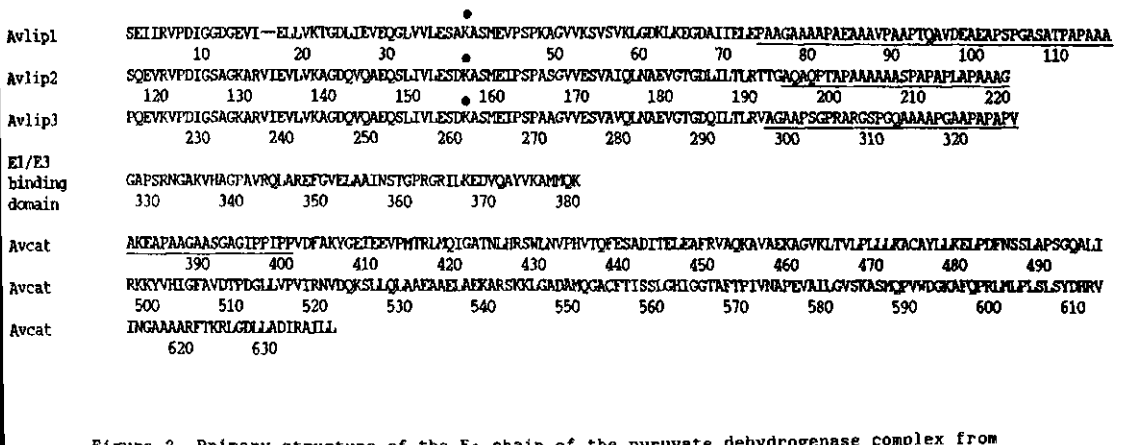
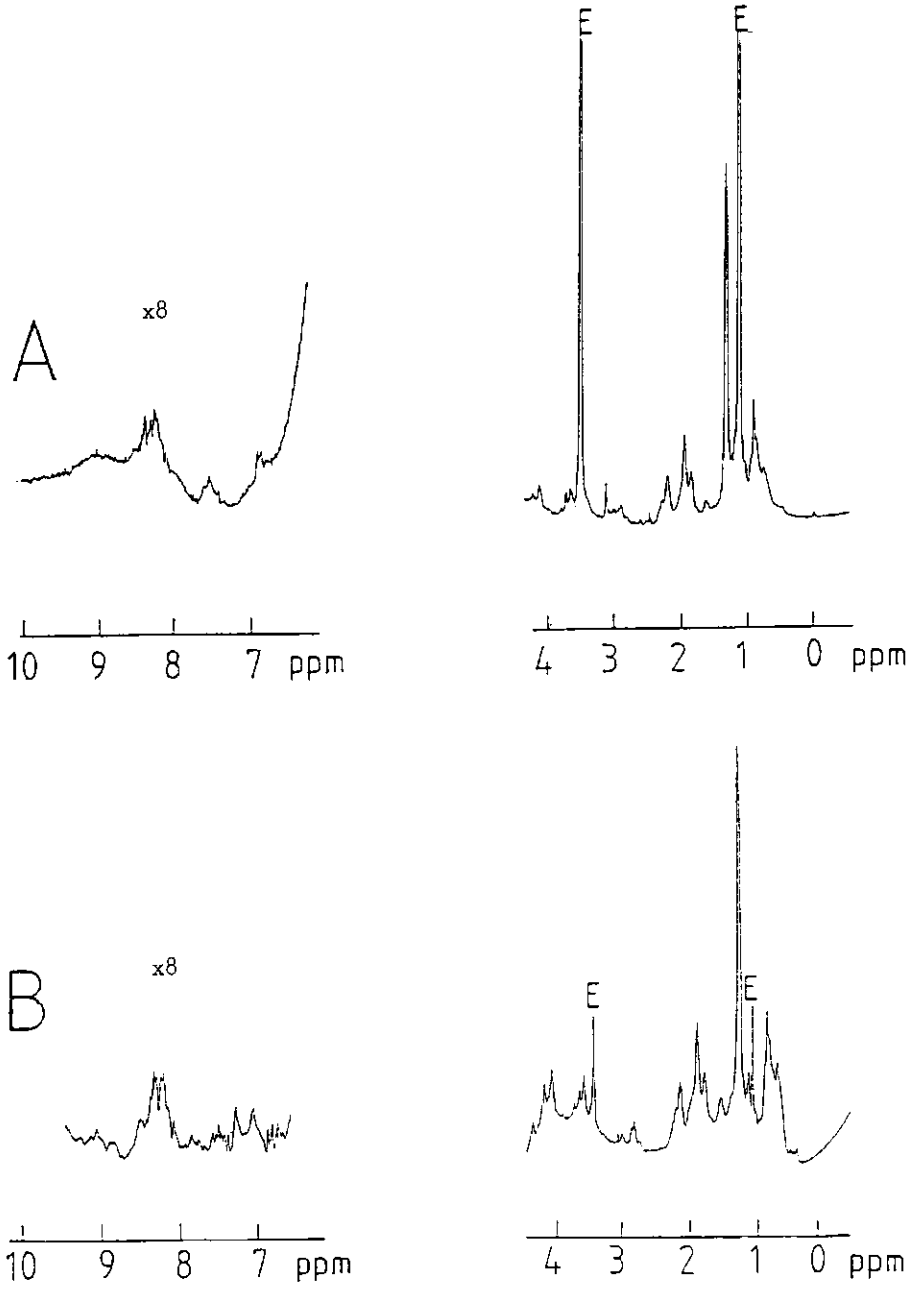


Figure 2. Primary structure of the E_2 chain of the pyruvate dehydrogenase complex from A.vinelandii. The proposed lipoyl binding sites are indicated (●). The alanine-proline rich regions are underlined. lip = lipoyl domain and cat = catalytic domain.

Most of the aromatic amino acid residues are located in the catalytic domain but obviously not in a mobile region since in the spectrum of the catalytic domain domain no sharp resonances are observed which can be ascribed to these residues, although sharp resonances are present at low field (Fig. 1C). However, both a histidine, a phenylalanine and a tyrosine residue are


A
x8
10 9 8 7 ppm
E
E
4 3 2 1 0 ppm
B
x8
10 9 8 7 ppm
E
E
4 3 2 1 0 ppm

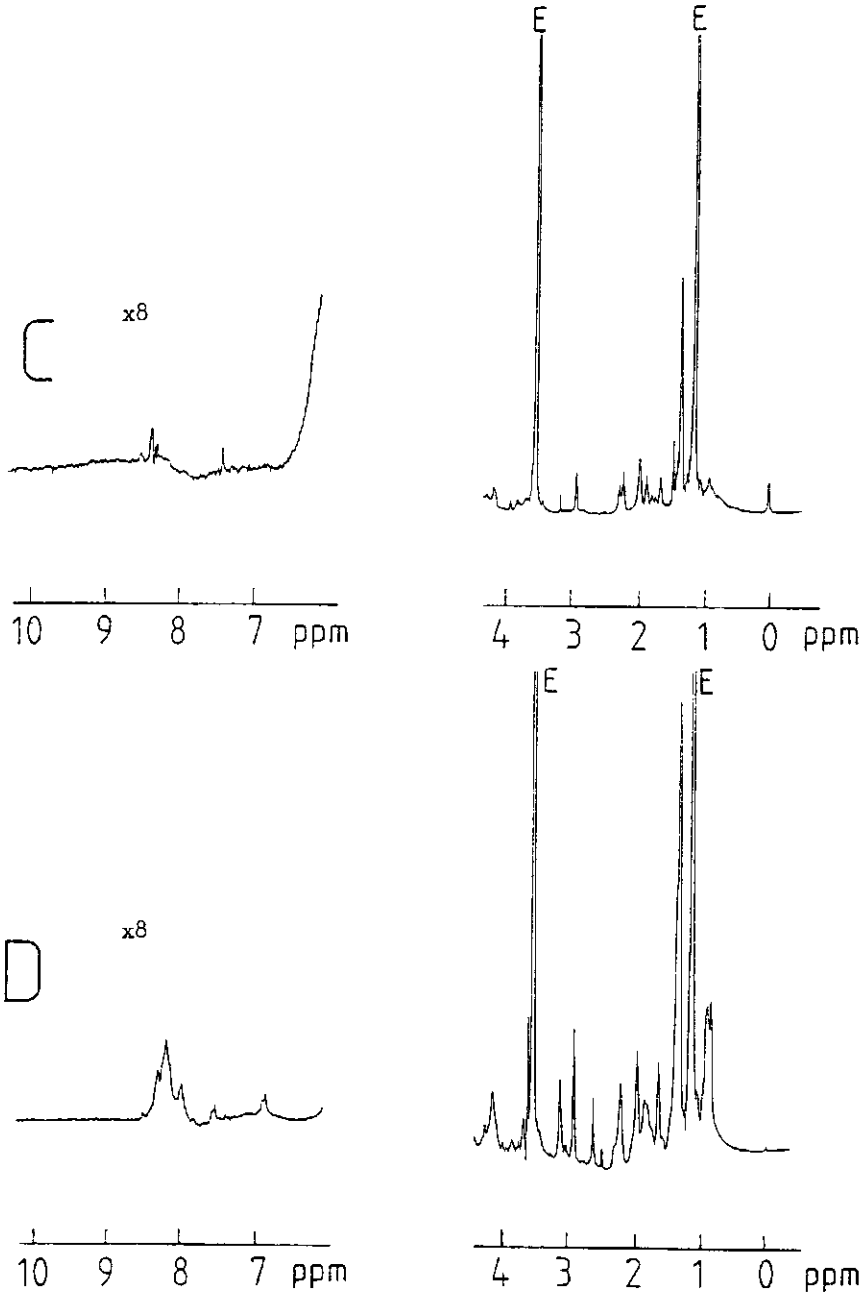


Figure 1. 600 MHz ^{1}H-NMR spectra of the pyruvate dehydrogenase complex, the E_2 component, the catalytic domain and the 2-oxoglutarate dehydrogenase complex from A.vinelandii. A, pyruvate dehydrogenase complex, 34 mg/ml (=159 μM E_2); B, E_2 component of the pyruvate dehydrogenase complex, 43.6 mg/ml (=770 μM E_2); C, catalytic domain of E_2, 23.7 mg/ml (=862 μM); D, 2-oxoglutarate dehydrogenase complex, 66.7 mg/ml (=550 μM E_2). The sharp resonances marked E are from ethanol.

present in the region possessing the E_1 and E_3 binding sites that is located between the lipoyl and the catalytic domain. The sharp resonances of histidine, phenylalanine and/or tyrosine residues are not observed in the whole complex. In E.coli NMR-spectra mobility of residues in this region has never been observed since E_2 of E.coli PDC aggregates at the high protein concentrations needed for NMR experiments [9].

Mobility in the catalytic domain

The catalytic domain is thought to form a compact rigid core (0.5-0.6 kDa) [2]. However, also in the spectrum of this domain sharp resonances are observed (Fig. 1C). The most prominent sharp resonance is at 1.39 ppm; the methyl resonance of alanine residues. This linewidth of this peak is less than the one observed in the spectra of PDC, E_2 and OGDC. Its intensity, however, is much less than in the other spectra. A second sharp resonance is found very close to it at 1.50 ppm. This resonance cannot be observed in PDC and OGDC. Probably both resonances can be attributed to a region rich in alanine residues located at the N-terminus of the catalytic domain. A mobile region in the catalytic domain has also been observed in E.coli PDC. A sharp resonance at 1.52 ppm was observed which was not present in the whole complex. It was ascribed to a very small alanine rich region at the N-terminus of the catalytic domain. In the spectrum of the catalytic domain of OGDC E_2 from Bacillus staerothermophilus [12] a small resonance at 1.36 ppm was observed, but was ascribed (may be by the lack of the amino acid sequence) to a non-protein contaminant. Recently a part of the amino acid sequence of this protein has been published and a small alanine rich region is shown at the N-terminus of the catalytic domain [21].

The spectrum of the catalytic domain of E_2 from A.vinelandii shows sharper resonances than those found in the spectra of PDC and E_2. This can probably be explained by the smaller region that is mobile. Although a limited amount of sharp resonances is present in the spectrum of the catalytic domain, these resonances cannot exclusively be ascribed to alanine residues. In addition to the 1.39 ppm and the 1.50 ppm resonances in the high-field region (0-4.5 ppm) some resonances can be observed which can pro-

bably be ascribed to amino acid residues present in the N-terminal region of the catalytic domain (e.g. prolines, lysine and isoleucines). Also in the low-field region (7.5 - 8.5 ppm) a few resonances are observed which can probably be ascribed to resonances from amide protons that are protected from water, suggesting that part of the mobile region has a folded structure [22]. The resonance at 7.45 ppm can be ascribed to phenylmethylsulfonyl fluoride, which is added in high amounts during isolation of the catalytic domain.

Discussion

The experiments described above clearly show sharp resonances in the ^{1}H-NMR spectra of OGDC, PDC, E_2 and the catalytic domain, indicating conformational mobility. Like in *E.coli* [9,10], in *A.vinelandii* PDC sharp resonances are observed which in part can be ascribed to alanine-proline rich regions. In contrast to the experiments with the *E.coli* complex [9] we were able to isolate E_2 and concentrate it up to 1 mM, without aggregation of the protein. The 600 MHz NMR spectrum of E_2 shows similar sharp resonances as the spectrum of the whole complex, which is a direct evidence that the mobile regions are located on the E_2 chain. Three long alanine-proline rich regions (20-30 residues) are present in E_2, located between the repeats in the lipoyl domain and between the lipoyl domain and the E_1 and E_3 binding domain. Our results are in agreement with the proposal that the conformational mobility can be ascribed to these regions, which is confirmed by experiments with genetic engineered *E.coli* complexes, in which mutations or deletions have been introduced in the alanine-proline rich regions [14,15]. A further confirmation came from experiments using a synthetic alanine-proline rich polypeptide, representing one of these regions in *E.coli* E_2 [23].

From the NMR experiments on the *E.coli* complex it has been suggested that the presence of the highly mobile alanine-proline sequences gives an independent conformational mobility to the three repeating subdomains in the lipoyl domain [14,15]. However, no sharp resonances of the subdomains (±8 kDa) are observed in

the spectra of the whole complex. Similar results have been obtained from NMR spectra of OGDC of Bacillus staerothermophilus [12]. This complex contains only one lipoyl domain per E_2 chain. This lipoyl domain contains one tryptophan, one tyrosine and four phenylalanine residues. Since no aromatic proton resonances of the lipoyl domain were visible in the spectrum of the intact complex, it has been suggested that the lipoyl domain is not free to move rapidly, but must spend a major part of the time bound to one of the other components of the complex. Thus, a mobile alanine-proline region can be observed without the corresponding mobility of the lipoyl domain. In the spectra of A.vinelandii PDC, OGDC or E_2 no indications for a small, mobile lipoyl-subdomain are found. This is in agreement with results obtained from fluorescence anisotropy experiments in which mobility of the single lipoyl-label group, the whole complex and the whole lipoyl domain (32 kDa) was found, but no mobility of a lipoyl-subdomain (9 kDa) could be observed. Obviously, using ^{1}H-NMR spectroscopy, no conclusions can be drawn about mobility of the lipoyl-subdomains, when only mobility of the alanine-proline rich region is observed.

In the low-field part of the spectrum of E_2 two specific sharp resonances are observed, which are ascribed to phenylalanine, histidine and/or tyrosine residues. Since no similar resonances are observed in the spectrum of the catalytic domain and no such residues are present in the lipoyl domain. Thus these resonances must be ascribed to the phenylalanine, histidine and/or tyrosine residues that are present in the region between the lipoyl and the catalytic domain. In this region the binding sites for the E_1 and E_3 components are located [2]. Recently, evidence has been obtained for a E_1 and E_3 binding domain in this region of the E_2 chain of PDC from Bacillus staerothermophilus [21]. Obviously, this region possesses conformational mobility. This is in agreement with previous results from fluorescence anisotropy experiments on FAD in free and bound E_3. After binding of E_3 to the E_2 core almost no change in rotational correlation coefficient was observed, from which was suggested that the binding region should possess mobility [24].

No sharp resonances of residues from this region have been observed in the spectrum of the whole complex. Obviously after binding of the large E_1 dimer (200 kDa) or E_3 dimer (100 kDa) the mobility of this region in E_2 diminishes to an extent which cannot be observed as a sharp resonance by ^{1}H-NMR spectroscopy.

The spectrum of the catalytic domain of E_2 from A.vinelandii (M=0.5-0.6 kDa), in contrast to the spectrum of the same domain from E.coli [9], shows a certain degree of conformational mobility. Probably the sharp resonances in the spectrum can partly be ascribed to an alanine-proline rich region located at the N-terminus of the catalytic domain. Besides the sharp resonance at 1.39 ppm a smaller resonance at 1.50 ppm is observed. In the spectrum of the whole complex this resonance is not observed; in the spectrum of E_2 it is present, partly masked by the large resonance at 1.39 ppm. Probably the alanine residues, located at the N-terminus of the catalytic domain become more mobile when the N-terminal half of the E_2 chain is excised after proteolysis, and partly change to a somewhat different environment. In E_2 the environment can differ from the one in the whole complex, since no large components are bound and the binding domain of E_1 and E_3 possesses mobility.
It has been shown that ^{1}H-NMR spectroscopy can be used to obtain information about conformational mobility in proteins, but clear results can only be obtained when residues, showing specific resonances in the ^{1}H-NMR spectrum, are present on well-defined sites in the protein. Therefore, the use of genetic engineered complexes in which mutations have been introduced, together with ^{1}H-NMR spectroscopy is a useful combination to study conformational mobility in specific regions of the polypeptide chain. With the use of both the intact complex, isolated E_2 and domains of E_2 new possibilities are created in studying the structure and mechanism of 2-oxoacid dehydrogenase complexes.

Acknowledgements

We would like to thank Mr. C. van Mierlo for recording some NMR spectra. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid from the Netherlands Organisation for the advancement of Scientific Research (ZWO).

References

1. Koike M. Reed, L.J. & Caroll W.R. (1963)
   J.Biol.Chem. 238, 30-39
2. Hanemaaijer, R., de Kok A. Jolles, J. & Veeger C. (1987)
   Eur.J.Biochem. 169, 245-252
3. Hanemaaijer R. Westphal A.H. van def Heiden T. de Kok, A. & Veeger C.
   Eur.J.Biochem., submitted
4. Hanemaaijer, R., Janssen A. de Kok A & Veeger C. (1988)
   Eur.J.Biochem., in press
5. Packman L.C., Hale, G. & Perham, R.N. (1984)
   EMBO J. 3, 1315-1319
6. Stephens, P.E., Darlison, M.G., Lewis H.M. & Guest J.R. (1983)
   Eur.J.Biochem. 133, 481-489
7. Shepherd, G.B. & Hammes, G.G. (1976)
   Biochemistry 16, 5234-5241
8. Scouten, W.H., de Graaf-Hess, A.C. de Kok, A., Visser, A.J.W.G. & Veeger C. (1978)
   Eur.J.Biochem. 84, 17-
9. Perham R.N., Duckworth, H.W. & Roberts G.C.K. (1981)
   Nature (London) 292, 474-477
10. Roberts, G.C.K., Duckworth H.W., Packman, L.C. & Perham, R.N. (1983)
    Ciba Symp. 93, 47-62
11. 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
12. Packman, L.C., Perham, R.N. & Roberts, G.C.K. (1984)
    Biochem.J. 217, 219-227

13. Wawrzynczak E.J., Perham R.N. & Roberts, G.C.K. (1981)
FEBS Lett. 131, 151-154
14. Radford, S.E., Laue, E.D., Perham, R.N., Miles, J.S. % Guest, J.R. (1987)
Biochem.J. 247, 641-649
15. Texter, F.L., Radford, S.E., Laue, D.E., Perham, R,N., Miles, J.S. & Guest, J.R. (1988)
Biochemistry 27, 289-296
16. Hanemaaijer, R., Masurel, R. Visser, A.J.W.G., de Kok, A. & Veeger, C. (1988)
FEBS Lett. in press
17. Hanemaaijer, R., Westphal, A.H., Berg A., de Kok A. & Veeger, C.
Eur.J.Biochem. submitted
18. Bosma, H.J., de Graaf-Hess, A.C., de Kok, A., Veeger, C., Visser, A.J.W.G. & Voordouw, G. (19829
Ann.N.Y.Acad.Sci. 378, 265-285
19. Bosma, H.J. (1984)
Ph.D. Thesis, Wageningen
20. Lowry, O.H., Rosebrough, N.J., Farr, A.L., & Randall, R.J. (1951)
J.Biol.Chem. 193, 265-275
21. Packman, L.C., Borges, A. & Perham, R.N. (1988)
Biochem.J. 252, 79-86
22. Wüthrich, K. (1986)
In "NMR of Proteins and Nucleic Acids", pp. 24, John Wiley & Son, Inc., New York
23. Radford, S.E., Laue, E.D. & Perham, R.N. (1986)
Biochem.Soc.Trans. 14, 1231-1232
24. Grande, H.J., Visser, A.J.W.G. & Veeger, C. (1980)
Eur.J.Biochem. 106, 361-369

CHAPTER 7

THE QUATERNARY STRUCTURE OF THE DIHYDROLIPOYL TRANSACETYLASE COMPONENT OF THE PYRUVATE DEHYDROGENASE COMPLEX FROM AZOTOBACTER VINELANDII: A RECONSIDERATION

Roeland Hanemaaijer, Adrie H. Westphal, Tini van der Heiden, Arie de Kok and Cees Veeger

Summary

After limited proteolysis of the dihydrolipoyl transacetylase component (E_2) of A.vinelandii pyruvate dehydrogenase complex (PDC) a C-terminal domain was obtained which retained the transacetylase active site and the quaternary structure of E_2 but had lost the lipoyl containing N-terminal part of the chain and the binding sites for the peripheral components pyruvate dehydrogenase and lipoamide dehydrogenase.

The C-terminus of this domain was determined by treatment with carboxypeptidase Y and shown to be identical with the C-terminus of E_2. Together with the previously determined N-terminus and the known amino acid sequence of E_2, as determined from the DNA sequence, a molecular mass of 27.5 kDa was calculated.

Ultracentrifugation in 6 M guanidine hydrochloride of the tryptic catalytic domain resulted in large systematic deviations. Similar deviations were obtained with a fragment of 30.7 kDa, obtained by genetic engineering. These deviations are probably due to incomplete unfolding. From the molecular mass of the native catalytic domain, 530 kDa, and the symmetry of the cubic structures observed on electron micrographs a 24-meric structure is concluded instead of the 32-meric structure proposed previously.

From the effect of guanidine hydrochloride on the lightscattering of intact E_2 it was concluded that dissociation occurs in a two-step reaction resulting in particles with an average mass 1/6 of the original mass before the N → D transition takes place. The first step is correlated with the transition of one or both tryptophan residues from a buried environment to an intermediate exposed one.

Crosslinking experiments with the catalytic domain indicated that the multimeric E_2 is built from tetramers and that the tetramers are arranged as a dimer of dimers.

A model for the quaternary structure of E_2 is given, in which it is assumed that the tetrameric E_2 core of PDC is formed from each of the six morphological subunits, located at the lateral faces of the cube. Binding of peripheral components to a site that interferes with the cubic assembly causes dissociation resulting in the unique small PDC of A.vinelandii.

Introduction

The pyruvate dehydrogenase complex (PDC) catalyzes the following reaction:

$$Pyruvate + NAD^+ + CoA \longrightarrow AcetylCoA + CO_2 + NADH + H^+.$$

It is composed of multiple copies of three enzymes: pyruvate dehydrogenase (E_1), dihydrolipoyl transacetylase (E_2) and lipoamide dehydrogenase (E_3)[1]. The E_1- and E_3 components are non-covalently bound as dimers to a core of E_2-chains [2]. PDC from gram-positive bacteria [3] or from mammalian sources [4] is based on a core of 60 E_2-chains with an icosahedral symmetry, whereas the E_2-core of PDC from Escherichia coli is composed of 24 subunits, arranged in a cubic structure with 432 symmetry [1,5]. The PDC from Azotobacter vinelandii is the smallest complex known [6]. It is based on a tetrameric E_2-core [7,8]. The sedimentation coefficient of this complex is 17-19S [7]; the complex from E.coli is 53-60S [3,9,10]. However, a 17-20S form of E.coli PDC has also been observed [5,9]. This small form is enzymatically active and is present in small amounts in E.coli PDC preparations. It is indicated as having a trimeric core, being the morphological subunit of both the cubic and the icosahedral complexes [11,12]. On the other hand, upon removal of the peripheral components of the A.vinelandii tetrameric core, E_2 associates to a multimer. In electron micrographs a similar cubic appearance as the E.coli E_2 core is observed [1,13]. On basis of its molecular mass of about 2 MDa [7,13] and its association-dissociation behaviour [13] it was thought to be composed of tetrameric instead of trimeric morphological subunits, arranged at the vertices of the E_2-cube. This would result in an E_2-core composed of 32 subunits [7].

The E_2 components of PDC's from different organisms are exquisitely sensitive to proteolytic cleavage under non-denaturing conditions [14-17]. After proteolysis usually a domain comprising the lipoyl moieties and a structural domain are obtained. In the latter the quaternary structure of intact E_2 and the transacetylase active site are retained. The binding sites for the E_1- and E_3 components are retained [14,15] or lost [16-

20] after proteolysis, depending on the source and on the conditions of proteolysis [19]. After limited proteolysis with trypsin of the E_2 component of A.vinelandii PDC a stable multimeric fragment was found which had lost the E_1- and E_3 binding sites but still possessed the transacetylase active site and the E_2 intersubunit binding sites [16]. Electron micrographs of this domain showed the same cubic appearance as intact E_2. The N-terminal amino acid sequence has been determined by automated Edman degradation and is located at residue 381 of the amino acid sequence derived from DNA sequence analysis of the gene encoding A.vinelandii E_2 [16,21]. The molecular mass determination by sedimentation equilibrium experiments in the presence of 6M GdnHCl and 2 mM dithiotreitol gave a value of 15.8 kDa [16]. The molecular mass of the native catalytic domain core was found to be 530 kDa [16], which supported the model of a core of 32 subunits. However, no definite data on the chain mass of the catalytic domain, such as obtained from the amino acid sequence, were available. In this paper we report the determination of the C-terminal sequence of the catalytic domain, obtained by reaction with carboxypeptidase Y. It showed to be identical to the C-terminal sequence of the intact E_2-chain. This means that the catalytic domain has a molecular mass of 27 kDa. Together with results obtained from light-scattering experiments and crosslinking experiments a model for the quaternary structure of E_2 is proposed, showing an E_2-core with a cubic structure, being comprised of 24 subunits, with tetrameric morphological subunits forming the lateral faces of the cube.

Materials and methods

Materials

Carboxypeptidase Y was obtained from Boehringer. Dimethylsuberimidate was from Sigma, all other crosslinkers were synthesized as described in [22]. Guanidine hydrochloride was purchased from Merck and re-crystallized in ethanol, resulting in less than 0.1 absorption unit at 240 nm for a 1 M solution. All other chemicals were analytical grade.

Protein purifications

The E_2 component was isolated from the pyruvate dehydrogenase complex by covalent chromatography on thiolsepharose 4B as described before [23], with modifications according to [16]. The catalytic domain was obtained after limited proteolysis with trypsin of the E_2 component, which was covalently bound on thiolsepharose 4B, as described previously [16]. When used in sedimentation equilibrium experiments the catalytic domain was purified by FPLC, equipped with a Superose 12 column, in standard buffer (50 mM potassium phosphate pH 7.0, 0.5 mM EDTA and 0.05 mM phenylmethylsulfonylfluoride) containing 6M GdnHCl. An extended catalytic domain was isolated as a fusion protein from a production clone which was obtained by transformation of the C-terminal part of the gene encoding E_2 (residue 1522-2365 [21]) cloned into the HindIII/BamHI sites of the multiple cloning site of the pUC9 vector (details to be published). Its N-terminus contains the 6 N-terminal residues of the vector encoded β-galactosidase.

C-terminus determination

The catalytic domain of the E_2 component (2.2 mg) was dialyzed against 25 mM ammoniumhydrogencarbonate, freeze-dried and dissolved in 50% acetic acid. The sample was eluted on a Sephadex G25 column to remove small material which could result from the digestion with trypsin. After elution the void fraction was again freeze-dried and citraconylated according the method described by Perham [24]. The sample was eluted on a Sephadex G25 column in 100 mM sodiumacetate pH 6.0. 40 nmol protein was incubated with 15 µg carboxypeptidase Y (dissolved in the same buffer) and samples of 10 nmol protein were withdrawn after 0, 1, 10 and 90 minutes, respectively. The reaction was stopped by addition of 2 mg sulfosalicylic acid, followed by incubation at 0°C for at least 30 minutes. The precipitated protein was removed by centrifugation and the supernatant, containing the free amino acid residues, was directly used for amino acid analysis.

Crosslinking

Crosslinking experiments were essentially performed as described in [22]. Dimethylsuberimidate or other crosslinking reagents were dissolved in 100 mM triethanolamine-HCl pH 8.5 and immediately mixed with the protein solution (1 mg/ml) to final concentrations of 10-100 mM. After incubation for 3.5 h at room temperature the reaction was stopped by addition of glycine at a 100-500 mM concentration. Samples were analysed by SDS-gelelectrophoresis.

Light-scattering

Light-scattering measurements were performed at 436 nm with a Fica 50 photometer. Measurements were carried out at 25°C in Danliker cuvettes. Samples were filtered before use to remove dust particles. Variable GdnHCl concentrations were obtained by addition from a 7.5 M stock solution to the protein sample. After mixing the solution it was incubated for 15 minutes at 25°C and measured at different angles between 30° and 135°. Results were corrected for the signal obtained from a blank without protein. Light scattering data were analysed according [25]. The ratio of molecular mass in the presence of GdnHCl (M_i) to that in the absence of GdnHCl (M_o) was determined from Eqn. 1

$$\frac{M_i}{M_o} = \frac{(n_o)^2 \, (dn/dc_o)_\mu^2 \; (c_o)(I_{i,x})}{(n_i)^2 \, (dn/dc_i)_\mu^2 \; (c_i)(I_{o,x})} \qquad (1)$$

where n represents the refractive index, $(dn/dc)_\mu$ the refractive index increment at constant chemical potential, c the enzym concentration, I the scattered light intensity at x degrees, and the subscripts i and o refer to the quantities in the presence and absence of GdnHCl, respectively. Eqn. 1 was derived assuming that the second virial coefficient is negligibly small. The refractive indices of the solutions with different concentrations GdnHCl were obtained from [26], the refractive index increments were taken from data obtained for bovine serum albumine [27],

assuming that for all globular proteins under the same conditions the refractive index increments are similar. In the analysis of M_1/M_0 the mean intensity of the signals obtained at 45°,75° and 90° was used, because at these angles the lowest background signal from the buffer was observed.

Other methods

E_2 transacetylase activity was assayed spectrophotometrically at 240 nm and 25°C, measuring the formation of S-acetyldihydrolipoamide as described previously [6]. Amino acid analysis was performed on a Biotronik LC 6000E analyser, equipped with a Durrum DC6A ion-exchanger (physiological run). Sedimentation equilibrium experiments were performed in a MSE Centriscan 75, equipped with an ultraviolet/visible monochromator. Solutions contained 50 mM potassium phosphate pH 7.0, 0.5 mM EDTA, 0.05 mM phenylmethylsulfonylfluoride, 6M GdnHCl and 2 mM dithiotreitol. Corrections for the solvent density were made according to standard procedures [28]. Circular dichroic spectra were measured as described previously [16]. Fluorescence spectra were measured on an Aminco SPF-500 spectrofluorimeter, thermostatted at 25°C.

Results and discussion

C-terminus determination

Protein samples of the catalytic domain fragment, obtained by trypsinolysis as described in methods, were incubated with carboxypeptidase Y. Already after a reaction time of one minute a release of amino acids is observed (see table 1). Hardly any arginyl- or lysyl residues were released which was expected if the C-terminus of the catalytic domain was formed by an internal cleavage of trypsin in the E_2-chain. Instead, from the start leucyl- and isoleucyl residues were found in a 2:1 ratio. After longer incubation times (10 minutes) also alanyl-, arginyl- and aspartyl residues are observed. This finding excludes the possibility of an internal cleavage. It is exactly to be expected from the sequence -DIRAILL at the C-terminus of intact E_2 [21].

No similar region is found elsewhere in the E_2-sequence. As expected, with intact E_2 the same pattern is observed. Because the N-terminus has been determined previously at residue 381 [16,21] this result indicates that the catalytic domain represents peptide fragment 381-637.

Table 1. Number of amino acid residues (nmol) released from the E_2 C-terminus after incubation with carboxypeptidase Y during 0, 1, 10 and 90 minutes, respectively. Residues which are not shown were present at less than 2.5 nmol after 90 minutes incubation.

Residue	Time (minutes)			
	t=0	t=1	t=10	t=90
Leu	0.0	2.8	10.5	16.3
Ile	0.0	1.1	5.8	8.3
Ala	0.8	1.1	3.3	8.0
Arg	0.0	0.9	2.3	4.3
Asp	0.0	0.4	1.0	4.8

Previously, we have also shown that neither the catalytic domain nor the lipoyl domain (fragment 1-about 330) binds E_1 or E_3. Therefore these results also indicate that the region 330-380 has an important role in the binding of E_1 and E_3, although we cannot exclude the involvement of either domain in the binding. Especially the observation that E_1-binding (or E_3 in excess) leads to dissociation of the E_2-multimer [8] indicates an interference with E_2-E_2 interaction sites, which are located at the catalytic domain. Such interference does not necessarily require binding on the catalytic domain but may be due to steric factors. In ultracentrifugation experiments no effect of E_1 or E_3 has been observed on the aggregation state of the catalytic domain.

Sedimentation equilibrium experiments

Previously, the molecular mass of the catalytic domain was estimated at 15.8 ±0.5 kDa [16] by sedimentation equilibrium experiments of the protein in 6M GdnHCl in the presence of 2 mM dithiotreitol or after reduction and carboxymethylation. In contrast on SDS-gels a mass of 27 kDa was observed. It is known

however for the intact protein that large deviations on SDS-gels are observed: 82 kDa instead of 65 kDa. Ultracentrifugation of the intact E_2 in 6M GdnHCl gave a value of 61-63 kDa [13,16]. Therefore the value of 15.8 kDa for the catalytic domain was considered a good estimate. The large deviation from the true molecular mass of 27 kDa as obtained from sequence information must therefore arise from a systematic error such as the presence of considerable amounts of small material, incomplete unfolding or an internal cleavage by trypsin. To avoid small material, which could have arisen from the proteolytic cleavage, the catalytic domain was purified extensively under denaturing conditions, as desribed in the methods. From the carboxypeptidase treatment and the determination of the N-terminus it was clear that neither small proteolysed material was present nor internal cleavage had taken place in the samples. To check whether the prolonged incubation during ultracentrifugation did not result in further fragmentation the protein was dialysed after ultracentrifugation and resubjected to gelelectrophoresis. Again a value of 27 kDa was found.

Incomplete unfolding seems unlikely because the N -> D transition takes place at about 2.5 M GdnHCl (see below). However, it is known that even in 6M GdnHCl not all secondary structure is lost [29]. When the ultracentrifugation experiment was carried out at pH 5.0 (or pH 1.0) a somewhat higher molecular mass was observed, 19.6 kDa instead of 15.8 kDa. By genetic engineering we obtained a protein fragment corresponding to residue 353-637 of E_2, which contains 6 N-terminal amino acid residues of the E.coli β-galactosidase. It shows normal catalytic activity and properties comparable to the tryptic catalytic domain (to be published), having a calculated molecular mass of 30.7 kDa. When the molecular mass of this protein was analysed by sedimentation equilibrium experiments in 6M GdnHCl and 2 mM dithiotreitol at pH 5.0, a value of 23.4 kDa is obtained, showing the same absolute deviation as the tryptic catalytic domain.

In conclusion, we cannot exclude that this relatively large deviation is due to incomplete unfolding, but also other unknown effects may contribute to this anomaly.

The subunit mass of the E_2 component is of importance in the

determination of the quaternary structure of the E_2-core. Hydrodynamic data of the intact, native E_2-core indicated a mass of 1.8-2.0 MDa [13], significantly larger than that of the E.coli E_2 [30]. Combined with the monomer weight, the chain stoichiometry of the components of PDC, and the cubic structure of E_2, which in E.coli is thought to be composed of eight morphological subunits [11], a model was proposed based on 32 subunits [8], in which the tetrameric E_2-core of PDC constitutes the corners of the isolated E_2-cube. With structures of this size with a known tendency to either dissociate or associate to dimers [13], large absolute deviations could result from heterodispersity of the system. The catalytic domain structure on the other hand is more stable (e.g. no crosslinking was required in preparations for electronmicroscopy) and no dimerization was observed in sedimentation velocity experiments. Sedimentation equilibrium experiments gave a molecular mass of 530 kDa [16] resulting in a multimer of 19-20 subunits. This indicates that probably the A.vinelandii cubic structure, in analogy with E.coli [11] and with the related 2-oxoglutarate dehydrogenase complex [31] is composed of 24 subunits.

Dissociation and unfolding

Further support for a 24-meric structure was obtained by light-scattering experiments of E_2 in GdnHCl. The ratio of the molecular mass (M_i) of E_2 in different concentrations GdnHCl to the molecular mass (M_o) of the enzyme in standard buffer (see methods) is shown in fig. 1. Since the subunits of E_2 are identical, the molecular mass ratio is a direct measure of the extent of dissociation of the multimer.

From fig. 1 it is clear that two transitions in the dissociation are observed, at 1.0 M and 1.8 M, respectively. At 1.0 M GdnHCl the multimer dissociated to one third of its original mass, and at 1.8 M GdnHCl a particle of about one sixth of the original mass resulted. At still higher concentrations a further small decrease in light-scattering is observed. Up to 2.0 M GdnHCl E_2-activity is still retained (fig. 2) and hardly any secondary structure is lost (fig. 3). At slightly higher concentrations GdnHCl a very clear transition is shown and both

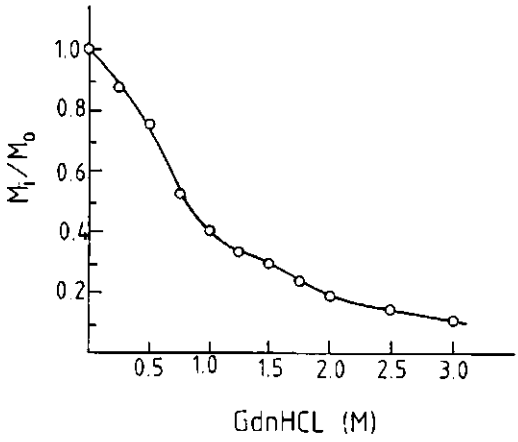


Figure 1. Variation of the molecular mass ratio of E_2 with varying GdnHCl-concentrations. The mean light scattering intensity at 45°, 75° and 90° of a solution of E_2 (1.8 mg/ml) in 50 mM potassium phosphate, 0.05 mM EDTA and 0.05 mM phenylmethylsulfonylfluoride was measured at 20°C. Concentrated GdnHCl was then added to yield the molarity indicated, and the resulting scattering was measured. The molecular mass ratio was calculated according Eqn. 1.

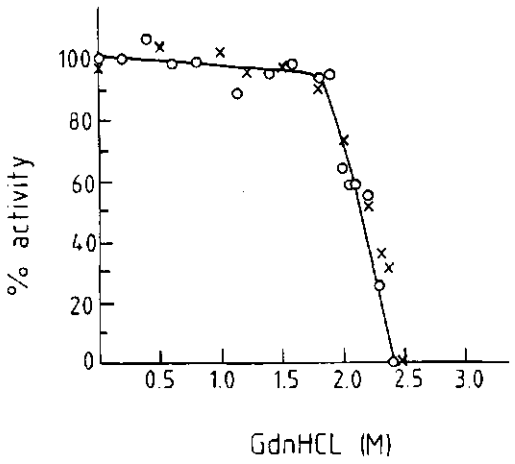


Figure 2. Variation of the percentatage activity of E_2 and of catalytic domain with varying GdnHCl-concentrations. The transacetylase activity of E_2 (o) and of the catalytic domain (x) was measured at 25°C.

activity and secondary structure are lost (figs. 2,3). The catalytic domain shows a comparable behaviour with respect to catalytic activity and molar ellipticity. These results show that one sixth of the multimeric E_2-core forms an enzymatically active morphological subunit.

The two tryptophan residues of the E_2 component (Trp429 and Trp591) are located in the catalytic domain. Fluorescence is a powerful technique to indicate the environment and changes in the environment of tryptophan residues. Fluorescence emission spectra (not shown) indicate that both tryptophan residues are buried, belonging to type I (notation by Burnstein, ref. 32). Upon addition of GdnHCl up to 1 M a decrease in the emission at 331 nm is observed (fig. 4) and a small red shift occurs to give a spectrum of the intermediate type II. Above 1.8 M GdnHCl a further decrease at 331 nm is observed which is due to a shift of the emission to 359 nm , characteristic of a completely exposed tryptophan residue. This coincides with the N-->D transition

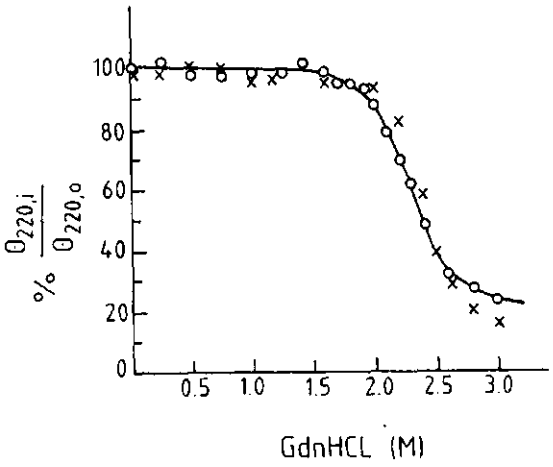


Figure 3. Variation of the percentage of the mean residue ellipticity at 220 nm of E_2 and the catalytic domain with varying GdnHCl-concentrations. Circular dichroic spectra of E_2 (o) and of the catalytic domain (x) were measured at 25°C at protein concentrations of 0.2-0.4 mg/ml.

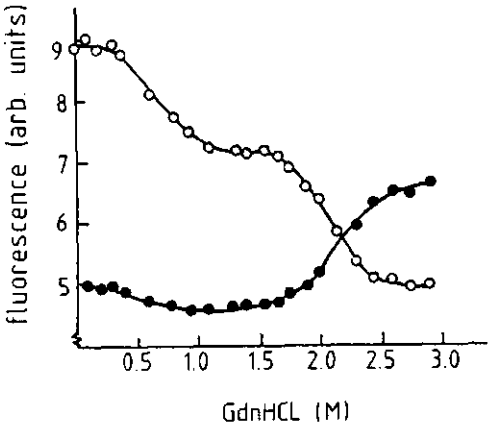


Figure 4. Variation of the fluorescence emission of E_2 with varying GdnHCl-concentrations. Fluorescence measurements at λem=330 nm (o) and λem=359 nm (•) were carried out at 25°C at 0.15 mg/ml and λexc=295 nm. Corrections for the influence of GdnHCl on fluorescence emission were made with samples containing tryptophan instead of protein.

discussed above. In comparison with fig. 1 it is concluded that the changes that occur up to 1.0 M GdnHCl reflect the dissociation into octamers. This causes a change in the environment of one or both tryptophan residues from a completely buried type to an intermediate exposed one. From 1.0 - 1.8 M GdnHCl, when the octamers dissociate to tetramers, almost no difference in tryptophan fluorescence is observed. Obviously the interactions between the tetramers are different from those between the octamers. At still higher GdnHCl concentrations the protein unfolds and both tryptophan residues become exposed.

Crosslinking

The catalytic domain forms the same multimeric structure as intact E_2 but has a much lower molecular mass. This makes it very suitable for cross-linking experiments, followed by analysis with SDS-gelelectrophoresis. The catalytic domain was incubated with the bifunctional crosslinking reagent dimethylsuberimidate (DMS,

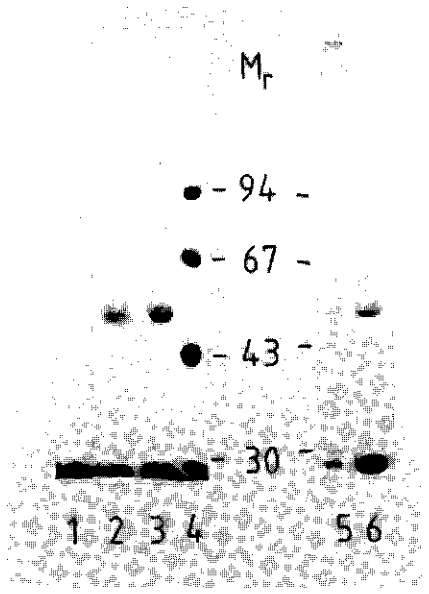


Figure 5. SDS-gelelectrophoresis of the catalytic domain of E_2 crosslinked with various crosslinking reagents. Crosslinking was carried out in 100 mM trietnanolamine-HCl pH 8.5, for 3.5 h at room temperature. Lane 1-3, with 0 mM, 25 mM and 75 mM dimethylsuberimidate, respectively, lane 4, marker proteins, lane 5, with 1.6 mM dimethyldodecylimidate, lane 6, with 14 mM dimethyladipylimidate. Molecular masses are indicated in kDa.

1.1 nm) at pH 8.5 and 20°C for 3.5 h. The protein concentration was kept low to avoid intermolecular crosslinking [33]. Analysis by SDS-gelelectrophoresis (fig. 5) showed that crosslinking took place up to a tetramer. When a shorter crosslinker was used like dimetyladipimidate (DMA, 0.9 nm) only monomer and dimer were found (fig. 5, lane 6) and using dimethyldodecylimidate (DMD, 1.5 nm) only large material was found, not able to enter the gel (fig.5, lane 5). These results confirm the proposed model of tetramers being the morphological subunits of the large, 24-meric core. Within the tetramer the arrangement of subunits is such that a preference for dimerization exists.

A model for the quaternary structure of E_2

From the data presented above we conclude that the isolated E_2 from A.vinelandii is a 24-meric structure, analogous to that of E.coli. Unlike the E_2-core of E.coli, upon addition of the

peripheral components the A.vinelandii E_2-core dissociates into tetramers. This gives a clue to the structure of the morphological subunits from which the E_2-cube is built. A model for this structure should explain 1) the stoichiometry and symmetry of component binding and 2) the dissociation and vice versa the assembly of the E_2-cube. From the crystal structure (Schierbeek, personal communication) it is clear that the subunits in E_3 are oriented around a twofold symmetry axis, as is the case with glutathion reductase [35]. Binding of a single E_3 dimer to E_2 does

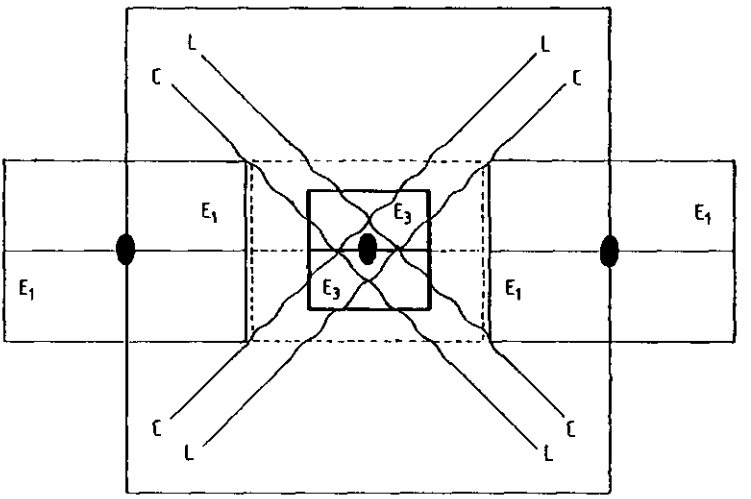


Figure 6. A schematic model of the tetrameric A.vinelandii PDC. The tetrameric E_2, which upon removal of the peripheral components aggregates to a 24-meric cubic structure, represents the lateral face of a cube. The E_1 and E_3 dimers possess a two-fold symmetry axis (●). The E_2-chain is indicated by L~C, in which L represents the lipoyl domain, C the catalytic domain and the connecting region 330-380 is represented by a wavy line.
The mobile lipoyl domain is not necessarily present at the indicated location. The dashed box represents the additional E_1 binding site at the rear face of the tetramer.

not lead to dissociation [8]. Therefore a unique E_3 binding site is indicated in the centre of the tetramer, analogous to the structure proposed for the E.coli PDC [11]. The arrangement of E_2 subunits should obey this symmetry and therefore an arrangement is proposed as indicated in fig. 6. In this arrangement a dimer of dimers is assumed as indicated by the crosslinking experiments. The binding sites of E_1 and E_3 are located in the connecting region (indicated in fig. 6 by a wavy line) between the

lipoyl and catalytic domains.

In the optimum catalytic stoichiometry three E_1 dimers and one E_3 dimer are bound to the E_2 tetramer. Because competition between binding of E_3 and E_1 is observed [8] it is assumed that the subunit orientation in the E_1 dimer is identical to the E_3 dimer. Compared to other 2-oxoacid dehydrogenase complexes the A.vinelandii PDC contains one additional binding site for E_1 per E_2 tetramer. We assume that this extra binding site becomes available during dissociation. This assumption is based on the observation that the dissociation is driven by component binding [8,13]. The additional binding site, indicated in fig. 6 by a dashed line, is located at the interface of the E_2 tetramer that is involved in cubic assembly. The three potential E_1 binding sites must have a comparable affinity for E_1 because the extent of dissociation is proportional to the extent of binding.

The model could also apply to the substructure of the E.coli PDC. Dissociation to an active 17S species of E.coli, in equilibrium with the usual 53-60S species, is observed [5,9]. The difference with the A.vinelandii E_2 is probably due to the absence of the additional binding site for E_1, as no evidence is presented that the association-dissociation equilibrium of E.coli PDC is shifted by binding of components.

The model does not explain the 2-step dissociation observed in the light-scattering experiments. According to the model the tetramers are considered to have equivalent positions in the cube. The fluorescence experiments indicate that the environmental changes which occur upon dissociation, as reported by tryptophan fluorescence, are completed at the dimer stage. Apparently the tetramers associate in a different way to octamers than do the octamers to form the cube. Whether this 2-step process is due to the presence of Gdn-HCl or indicates the in vivo pathway of assembly remains an open question.

Acknowledgements

We wish to thank Mr W. Roelofsen for carrying out the amino acid analysis and Prof. de Kruijff (State University Utrecht, The Netherlands) for the hospitality for the use of the light-

scattering instrumentation. We are also grateful to Mr M.M. Bouwmans for the drawing of the figures and to Mrs J.C. Toppenberg-Fang for the preparation of the manuscript. This work was supported by the Netherlands Foundation for Chemical Research (SON) with the financial aid from the Netherlands Organisation for the Advancement of Pure Research (NWO).

References

1. Reed, L.J. & Oliver, R.M. (1968).
Brookhaven Symp. Biol. 21, 397-411.
2. Koike, M., Reed, L.J. & Caroll, W.R. (1963).
J.Biol.Chem. 238, 30-39.
3. Henderson, C.E., Perham, R.N. & Finck, J.T. (1979).
Cell 17, 85-93.
4. Reed, L.J., Damuni, Z. & Merryfield, M.L. (1985).
Curr.Top.Cell.Regul. 27, 41-49.
5. Danson, M.J., Hale, G., Johnson, P., Perham, R.N., Smith, J. & Spragg, P. (1979).
J.Mol.Biol. 129, 603-617.
6. Bresters, T.W., de Abreu, R.A., de Kok, A., Visser, J. & Veeger, C. (1975).
Eur.J.Biochem. 59, 335-345.
7. De Abreu, R.A., de Kok, A., de Graaf-Hess, A.C. & Veeger, C. (1977).
Eur.J.Biochem. 81, 357-364.
8. Bosma, H.J., de Kok, A., Westphal, A.H. & Veeger, C. (1984).
Eur.J.Biochem. 142, 541-549.
9. Schmitt, B. (1976).
Biochimie (Paris) 58, 1405-1407.
10. Dennert, G. & Höglund, S. (1970).
Eur.J.Biochem. 12, 502-507.
11. Oliver, R.M. & Reed, L.J. (1982).
in Electron Microscopy of Proteins, Vol. 2. (Harris, R. ed.)
Academic Press, London, pp. 1-48.
12. Bleile, D.M., Hackert, M.L., Pettit, F.H. & Reed, L.J.(1981).
J.Biol.Chem. 256, 514-519.
13. Bosma, H.J., de Kok, A., van Markwijk, B.W. & Veeger, C.

(1984).
Eur.J.Biochem. 140, 273-280.

14. Bleile, D.M., Munk, P., Oliver, R.M. & Reed, L.J. (1979).
Proc.Natl.Acad.Sci. USA 76, 4385-4389.

15. Packman, L.C., Perham, R.N. & Roberts, G.C.K. (1984).
Biochem.J. 217, 219-227.

16. Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987).
Eur.J.Biochem. 169, 245-252.

17. Perham, R.N. & Wilkie, A.O.M. (1980).
Biochem.Int. 1, 470-477.

18. Kresze, G.B., Ronft, H., & Dietl, B. (1980).
Eur.J.Biochem. 105, 371-379.

19. Packman, L.C. & Perham, R.N. (1986).
FEBS Lett. 139, 155-158.

20. Packman, L.C. & Perham, R.N. (1987).
Biochem.J. 242,531-538.

21. Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988).
Eur.J.Biochem., in press.

22. Davies, G.A. & Stark, G.R. (1970).
Proc.Natl.Acad.Sci.USA 66, 651-656.

23. De Graaf-Hess, A.C. & de Kok, A. (1982).
FEBS Lett. 143, 261-264.

24. Perham, R.N. (1978).
in Techniques in Protein and Enzyme Biochemistry, Vol. B110,
(Kornberg, H.L., Metcalfe, J.C., Northcote, D.H., Pogson,
C.I. & Tipton, K.F. eds.), Elsevier/North-Holland, Amsterdam,
pp.1-39.

25. Parr, G.R. & Hammes, G.G. (1975).
Biochemistry 14, 1600-1605.

26. Nozaki, Y. (1972).
Meth.Enzymol. 26, 43-50.

27. Noelken, M.E. & Timasheff, S.N. (1967).
J.Biol.Chem. 242, 5080-5085.

28. Schachman, K. (1959).
Ultracentrifugation in biochemistry, Academic Press, N-Y,
London.

29. Tanford, C. (1968).
Advan.Protein Chem. 23, 202.

30. Eley, M.H., Namihira, G., Hamilton, L., Munk, P. & Reed, L.J. (1972).
Arch.Biochem.Biophys. 152, 655-669.

31. Koike, M. & Koike, K. (1976).
Adv.Biophys. 9, 187-227.

32. Burnstein, E.A., Vedenkia, N.S. & Ivkova, M.N. (1973).
Photochem. Photobiol. 18, 263-279.

33. Hajdu, J., Bartha, F. & Friedrich, P. (1976).
Eur.J.Biochem. 68, 373-383.

34. Packman, L.C., Hale, G. & Perham, R.N. (1984).
EMBO J. 3, 1315-1319.

35. Thieme, R., Pal, E.F., Schirmer, R.H. & Schulz, G.E. (1981).
J.Mol.Biol. 152, 763-782.

Summary

The studies described in this thesis deal with the structure of the Azotobacter vinelandii dihydrolipoyl transacetylase, the core component (E_2) of the pyruvate dehydrogenase complex.
In all organisms the pyruvate dehydrogenase complex is closely related to the 2-oxoglutarate dehydrogenase complex and, if present, the branched-chain 2-oxoacid dehydrogenase complex. These enzyme complexes are large multimeric structures. The smallest known is the pyruvate dehydrogenase complex from A.vinelandii. Upon resolution of the other components, the tetrameric core component of this complex aggregates to a well-defined multimeric structure, resembling the structure from the large complexes from other organisms.. Therefore, it seems likely that the A.vinelandii complex could represent the model for the building unit of the large complexes from other organisms. Since the core component (E_2) carries all the information concerning the quaternary structure of the complex, we focussed our attention on this intriguing enzyme.
The domain structure of E_2 has been examined by limited proteolysis of E_2, as described in chapter 2. After limited proteolysis with trypsin two stable domains were obtained. The lipoyl domain carries the lipoyl groups which are concerned with the transport of the substrates between the active sites of the different components. The catalytic domain possesses the transacetylase active site and the E_2-intersubunit binding sites, responsible for the quaternary structure of E_2. The binding sites for the E_1 and E_3 components are lost during proteolysis.
The cloning and sequencing of the gene encoding dihydrolipoyl transacetylase have been described in chapter 3. The gene, located downstream of the gene encoding the PDC E_1 component, does not possess an own promoter, but is probably regulated by the E_1-promoter. The gene possesses a strong terminating sequence. Downstream the gene encoding E_2 no open reading frame, that codes for the E_3 component, has been identified, as has been found in E.coli. The primary structure of E_2, derived from the DNA sequence, is homologous to that of E_2 from E.coli. The lipoyl domain, located at the N-terminus, is built from three repeating

sequences, separated by regions which are very rich in alanine and proline residues. The catalytic domain, located at the C-terminus, comprises the transacetylase active site and the E_2 intersubunit binding sites. The region, located between the lipoyl and the catalytic domain contains many charged amino acid residues and is thought to possess the E_1 and E_3 binding sites. The expression of the gene encoding E_2, located on plasmid pRA282 and cloned in E.coli, has been described in chapter 4. A high production of E_2 was obtained. The production raised dramatically when the cells were in the stationery phase of the growth-cycle. The percentage active E_2 varied strongly per culture. The inactivation was found to be caused by formation of intramolecular or intermolecular S-S-bridges, resulting in incorrect folding of the catalytic domain. An activation and an isolation procedure have been described.

Mobility of the repeating units within the lipoyl domain has been studied using time-resolved fluorescence, as described in chapter 5. It has been shown that the repeats show no independent rotational mobility, but rotate as one unit, serving the active sites of the different components.

Internal mobility within the lipoyl domain has been observed by ^{1}H-NMR experiments, as described in chapter 6. Probably this internal mobility, that is ascribed to the alanine-proline rich region, does not result into an independent mobility of the three repeats. The catalytic domain, despite its compact structure, still possesses a certain amount of internal mobility. This can partly be ascribed to alanine and proline residues, probably the N-terminal region of the domain, which is rich in these residues. In the spectrum of E_2 sharp resonances have been observed that can be ascribed to mobility of the E_1 and E_3 binding domain. Such mobility has not been found after binding of E_1 and E_3 components, in the whole complex.

The molecular mass of the native catalytic domain and of the single polypeptide chain have been determined, and from this and light-scattering and crosslinking experiments it has been concluded that the large multimeric structure of the isolated catalytic domain (and of E_2) is built from 24 subunits in contrast to a 32-meric structure as proposed previously. A model has been pre-

sented for the quaternary structure of E_2, in which it is assumed that the multimeric E_2-core is built from six tetrameric morphological subunits, forming the lateral faces of the cubic 24-mer. These tetrameric subunits represent the E_2-core of the intact complex. Compared to other 2-oxoacid dehydrogenase complexes, the A.vinelandii PDC contains one additional binding site for E_1 per E_2 tetramer. It is assumed that this extra binding site becomes available during dissociation, resulting in the unique small PDC of A.vinelandii.

Samenvatting

Suikers, die we door onze voeding opnemen, worden in het lichaam afgebroken tot energierijke verbindingen. Een belangrijke stap in dit proces is de omzetting van pyruvaat in acetylCoA dat vervolgens in de citroenzuurcyclus verder wordt omgezet tot energierijke verbindingen, zoals ATP, nodig om cellen en organen optimaal te doen funktioneren. De omzetting van pyruvaat in acetylCoA vindt plaats o.i.v. het pyruvaat dehydrogenase complex (PDC), een complex van drie verschillende enzymen. Door als complex aanwezig te zijn kunnen deze enzymen efficienter funktioneren dan in het geval dat ze elk apart zouden voorkomen. Het produkt van het eerste enzym (E_1) is direct op de goede plaats om te reageren met enzym twee (E_2), dat vervolgens zijn produkt weer direkt kan afgeven aan enzym drie (E_3).
Het complex is opgebouwd uit een vast aantal E_2 componenten, waar de E_1 en E_3 componenten aan kunnen binden. Het complex is een van de grootste enzymcomplexen die in de cel voorkomen, van ongeveer dezelfde grootte als ribosomen. De grootte van het PDC verschilt per organisme en wordt gedikteerd door de quaternaire structuur van de E_2 component. In zoogdieren bestaat het bv. uit 60 E_2 ketens, terwijl het in gram-positieve bacterien en in de gram-negatieve darmbacterie E.coli uit 24 E_2 ketens is opgebouwd.
Het onderzoek, dat in dit proefschrift beschreven is, handelt over PDC uit de gram-negatieve bacterie Azotobacter vinelandii. Het unieke van dit PDC is dat het is opgebouwd uit slechts vier E_2 ketens. De vraag die we ons dan ook stellen is of de grootte ook een funktionele betekenis heeft. Ook is het van belang te weten hoe de reakties binnen zo'n complex verlopen. Om dit te bestuderen is onderzoek naar de structuur van het eiwit noodzakelijk. Omdat veel van de karakteristieke eigenschappen van het complex op E_2 gelegen zijn, is de aandacht op dit enzym gericht.
In de inleiding (hoofdstuk 1) is een overzicht gegeven wat er over PDC (en twee verwante enzymcomplexen, OGDC en BCOADC) bekend is. De vraag wordt gesteld of het kleine PDC uit A.vinelandii de bouwsteen kan zijn van de grotere enzymcomplexen. Omdat ook in het onderzoek geldt: "hoe kleiner, hoe eenvoudiger", neemt PDC uit A.vinelandii een centrale plaats in in het totale onderzoek aan

deze enzymcomplexen.
De domeinstructuur van E_2, de opbouw van de eiwitketen, is onderzocht m.b.v. beperkte proteolyse, hetgeen beschreven is in hoofdstuk 2. Na digestie van natief E_2 met trypsine worden twee stabiele domeinen verkregen: het lipoyl domein, dat de lipoylgroepen bezit die het transport van de substraten tussen de verschillende componenten verzorgen, en het katalytisch domein, dat de actieve plaats bezit en de bindingsplaatsen voor de E_2 subeenheden, die de quaternaire structuur van het complex bepalen. De bindingsplaatsen voor de E_1- en E_3 component zijn door proteolyse verloren gegaan.
In hoofdstuk 3 zijn de clonering en sequentiebepaling van het gen dat codeert voor E_2 beschreven. Het gen dat codeert voor E_2 ligt op de DNA keten stroomafwaarts van het gen dat codeert voor E_1. Het bezit geen eigen promoter, maar wel een sterke terminatie sequentie. De aminozuurvolgorde van E_2, is uit de DNA sequentie afgeleid en komt sterk overeen met de aminozuurvolgorde van E_2 uit E.coli. Het lipoyl domein, dat N-terminaal gelegen is, bestaat uit drie repeterende sequenties, met daartussen gebieden van 20-30 aminozuren die zeer veel alanine- en proline residuen bevatten. Het katalytisch domein ligt C-terminaal. Het gebied tussen het lipoyl en katalytisch domein, waarvan gedacht wordt dat het de bindingsplaatsen voor de E_1- en E_3 component bezit, bevat veel geladen aminozuren.
Het gen, dat codeert voor A.vinelandii E_2, is in een plasmide gezet en gecloneerd in E.coli, zoals beschreven is in hoofdstuk 4. In E.coli werd een hoge produktie aan A.vinelandii E_2 gevonden, ongeveer 15 maal zo hoog als in A.vinelandii zelf. De produktie van E_2 nam sterk toe als de cellen in de stationaire fase van de groeicyclus waren. Een deel van het E_2 (variërend per celkweek) bleek inaktief te zijn. De oorzaak hiervan bleek te liggen in de vorming van (normaal niet voorkomende) intermoleculaire- of intramoleculaire S-S-bruggen. Een methode om het enzym weer aktief te krijgen en een methode om het uit E.coli te isoleren zijn in hoofdstuk 4 beschreven.
De mobiliteit van het lipoyl domein is bestudeerd zoals beschreven in hoofdstuk 5. Door het lipoyl domein specifiek te labelen kon m.b.v. tijdsafhankelijk fluorescentie anisotropie verval aange-

toond worden dat het lipoyl domein als geheel erg beweeglijk is, maar dat geen onafhankelijke beweeglijkheid van de drie repeterende structuren in het lipoyl domein aanwezig is. Ook is de mobiliteit onderzocht m.b.v. proton NMR, zoals beschreven in hoofdstuk 6. De drie alanine-proline rijke gebieden, die tussen de drie repeterende strukturen in het lipoyl domein liggen, bleken wel mobiel te zijn, maar blijkbaar resulteert deze mobiliteit niet in een afzonderlijke beweeglijkheid van de repeterende strukturen. In E_2 bleek het E_1- en E_3 bindings domein ook zeer mobiel, terwijl na binding van de componenten (in PDC) geen mobiliteit meer werd gevonden. Omdat in PDC het lipoyldomein nog steeds beweeglijk is, moet zich tussen het lipoyl domein en het E_1- en E_3 bindingsdomein een scharnierpunt bevinden. Het katalytisch domein bleek ook een mobiel gedeelte te bezitten, waarschijnlijk een deel van de N-terminus van dit domein. Hier zou een tweede scharnierpunt kunnen liggen, dat verantwoordelijk is voor de beweeglijkheid van het E_1-en E_3 bindingsdomein.

Het E_2 aggregeert, na verwijdering van de gebonden E_1-en E_3 componenten tot een grote E.coli-achtige structuur. Het katalytisch domein heeft eenzelfde struktuur. Uit vroegere metingen werd een struktuur verondersteld, bestaande uit 32 ketens. In hoofdstuk 7 is beschreven, dat uit bepalingen van het molecuulgewicht van het natieve katalytisch domein en dat van de enkele keten, alsmede uit lichtverstrooiings- en crosslinking experimenten kan worden gekonkludeerd dat het katalytisch domein (en E_2) bestaat uit 24 ketens. Dit betekent dat A.vinelandii PDC meer op E.coli lijkt dan altijd gedacht werd. Een model is voorgesteld, waarin zes A.vinelandii PDC's kunnen aggregeren tot een E.coli-achtig PDC. Het verschil tussen de twee complexen kan een extra bindingsplaats voor de E_1 component op A.vinelandii E_2 zijn. Na dissociatie van de grote structuur komt deze bindingsplaats vrij, wat resulteert in de unieke structuur van A.vinelandii PDC.

Het (relatief) kleine A.vinelandii pyruvaat dehydrogenase complex lijkt dus model te staan voor de grotere pyruvaat dehydrogenase (en verwante) complexen uit andere organismen.

Curriculum vitae

Roeland Hanemaaijer werd op 21 februari 1959 te Bunschoten geboren. In 1977 behaalde hij het diploma Atheneum-B aan het Farel College te Amersfoort en begon met de studie Moleculaire Wetenschappen aan de Landbouwuniversiteit te Wageningen. In 1980 werd het kandidaatsexamen afgelegd en in november 1984 werd de studie afgesloten, met als hoofdvakken Moleculaire biologie (Prof.dr. A. van Kammen), Experimentele diermorfologie en celbiologie (Prof.dr. W. van Muiswinkel) en Toxicologie (Prof.dr. J.H. Koeman) en als bijvak Biochemie (Prof.dr. F. Müller). Tevens werd de onderwijsbevoegdheid scheikunde verkregen. De praktijkperiode van 6 maanden werd uitgevoerd aan het Rijksinstituut voor volksgezondheid en milieuhygiene (RIVM) te Bilthoven. Op 1 februari 1985 trad hij in dienst van de Nederlandse organisatie voor wetenschappelijk onderzoek (NWO) aan de afdeling Biochemie van de Landbouwuniversiteit te Wageningen.