

Solution structures of lipoyl domains of the 2-oxo acid dehydrogenase complexes from *Azotobacter vinelandii*

Implications for molecular recognition

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

Co-promotor: dr. A. De Kok
universitair hoofddocent, vakgroep Biochemie

1102201, 2219

**Solution structures of lipoyl domains of the 2-oxo acid
dehydrogenase complexes from *Azotobacter vinelandii***

Implications for molecular recognition

Axel Berg

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van de Landbouwniversiteit Wageningen,
dr. C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 7 februari 1997
des namiddags te vier uur in de Aula

Van de rector



The research described in this thesis was carried out at the Department of Biochemistry, Agricultural University, Wageningen, The Netherlands.

The investigations were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

ISBN 90-5485-626-2

BIBLIOTHEEK
LANDBOUWUNIVERSITEIT
WAGENINGEN

Stellingen

1. De suggestie dat een conformationele overgang tussen een type I- en een type I' turn plaatsvindt, en dat deze een rol speelt bij het bewegen van de biotinylgroep van het biotine carboxyl carrier eiwit (BCCP) tussen de verschillende actieve centra van het acetyl-CoA carboxylase, is niet gebaseerd op enige experimentele waarneming of analogie.

Athappilly, F.K. & Hendrickson, W.A. (1995) *Structure* 3, 1407-1419

2. De hogere mobiliteit, tijdens natieve gel-electroforese, van gelipoyeerde lipoyl-domeinen ten opzichte van de niet-gelipoyeerden, is het gevolg van het verdwijnen van een positieve lading, en wordt niet veroorzaakt door een conformatie verandering die leidt tot een compactere eiwitstructuur.

Liu, S., Baker, J.C., Andrews, P.C. & Roche, T.E. (1995) *Arch. Biochem. Biophys.* 316, 926-940.

3. Men mag wel vaker, zoals Havel, van geluk spreken in wetenschappelijke publicaties.

Havel, T.F. (1991) *Prog. Biophys. Molec. Biol.* 56, 43-78.

4. Berekende ¹H-NMR chemische verschuivingen kunnen nog niet worden gebruikt om de kwaliteit van eiwitstructuren te meten.

Williamson, M.P., Kikuchi, J. & Asakura, T. (1995) *J. Mol. Biol.* 247, 541-546.

5. De mathematische beschrijving, door Forman-Kay *et al.*, van de t_1 -evolutie van de magnetisatie in een HMQC-*J* experiment voor een multiplet gecentreerd rond de nulrequentie, is onjuist.

Forman-Kay, J.D., Gronenborn, A.M., Kay, L.E., Wingfield, P.T. & Clore, G.M. (1990) *Biochemistry* 29, 1566-1572.

6. De meeste stellingen zijn geen stellingen, maar waarheden als koeien.
7. Het afschaffen van de dienstplicht zal het verzet tegen militarisme, oorlog en geweld doen afnemen.
8. Als men zegt geen tijd te hebben wordt veelal bedoeld dat men geen prioriteit heeft. Prioriteit zou daarom volgens de nieuwe spellingsregels als prioritijd geschreven moeten worden.
9. Een gepopulariseerde Nederlandse samenvatting voor niet-ingewijden in een proefschrift roept bij hen in het algemeen meer vragen op dan worden beantwoord.
10. De maatregel van NWO dat een OIO pas dan wordt opgevolgd als voor zijn/haar voorganger geen wachtgeld meer hoeft te worden betaald, kost op termijn meer geld dan zij oplevert.
11. Het opzuiveren van eiwitten leidt tot niets.

Stellingen behorende bij het proefschrift:

'Solution structures of lipoyl domains of the 2-oxo acid dehydrogenase complexes from *Azotobacter vinelandii*. Implications for molecular recognition'

Axel Berg

Wageningen, 7 februari 1997

CONTENTS

	Page
List of abbreviations	7
Chapter 1 General introduction: 2-oxo acid dehydrogenase complexes	9
Chapter 2 Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from <i>Azotobacter vinelandii</i>	43
Chapter 3 Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from <i>Azotobacter vinelandii</i> . Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex	71
Chapter 4 Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from <i>Azotobacter vinelandii</i>	95
Chapter 5 Three-dimensional structure in solution of the N-terminal lipoyl domain of the pyruvate dehydrogenase complex from <i>Azotobacter vinelandii</i>	115
Chapter 6 Reductive acylation of lipoyl domains of 2-oxo acid dehydrogenase complexes from <i>Azotobacter vinelandii</i>	135
Chapter 7 Summary and concluding remarks	147
Nederlandse samenvatting	153
Curriculum vitae	157
List of publications	159
Nawoord	160

LIST OF ABBREVIATIONS

Å	0.1 nm
BCDHC	branched-chain 2-oxo acid dehydrogenase complex
CAT	chloramphenicol acetyltransferase
$\Delta\delta$	secondary shift
E1	2-oxo acid dehydrogenase
E1b	decarboxylase component of BCDHC
E1o	2-oxoglutarate dehydrogenase
E1p	pyruvate dehydrogenase
E2	dihydrolipoyl acyltransferase
E2b	acyltransferase component of BCDHC
E2o	dihydrolipoyl succinyltransferase
E2p	dihydrolipoyl acetyltransferase
E3	lipoamide dehydrogenase
FAD	flavin adenine dinucleotide
IPTG	isopropyl β -D-thiogalactopyranoside
M_r	relative molecular mass
OGDHC	2-oxoglutarate dehydrogenase complex
PDHC	pyruvate dehydrogenase complex
STEM	scanning transmission electron microscopy
ThDP	thiamin diphosphate
NMR	nuclear magnetic resonance
1D, 2D, 3D	one-, two-, three-dimensional
DG-SA	distance geometry/simulated annealing
DQF-COSY	double-quantum filtered correlation spectroscopy
E.COSY	exclusive correlation spectroscopy
HMQC	heteronuclear multiple-quantum coherence
HSQC	heteronuclear single-quantum coherence
NOESY	nuclear Overhauser enhancement spectroscopy
SA	simulated annealing
SCUBA	stimulated cross-peaks under bleached alphas
TOCSY	total correlation spectroscopy
TPPI	time-proportional phase incrementation
TQF-COSY	triple-quantum filtered correlation spectroscopy
r.m.s.	root mean square
r.m.s.d.	root mean square deviation
VDW	van der Waals

CHAPTER 1

General introduction: 2-oxo acid dehydrogenase complexes

Introduction

The 2-oxo acid dehydrogenase complexes are commonly regarded as classic examples of multienzyme complexes. Ever since the pioneering work of the group of Lester J. Reed (Reed, 1974), these complexes have expansively disclosed many of their secrets, providing currently a wealth of information on macromolecular structure, assembly and symmetry, active-site coupling, conformational mobility, substrate specificity and metabolic regulation (Roche & Patel, 1989; Patel *et al.*, 1996). In particular, by effectively combining protein biochemistry with modern biophysical and genetic techniques, which are currently undeniable indispensable in contemporary biochemistry, the knowledge of structural and mechanistic properties of these complexes of gigantic size (0.7 - ~14 MDa) has expanded largely during the last decade. Major contributions to this field, i.e. on bacterial complexes, have been made by the laboratories of Richard N. Perham (Perham *et al.*, 1988; Perham & Packman, 1989; Perham, 1991), John R. Guest (Miles & Guest, 1987a; Guest *et al.*, 1989), Aart de Kok and Wim G.J. Hol (Mattevi *et al.*, 1992a; De Kok, 1996).

Multienzyme complexes are defined as noncovalent aggregates of enzymes that catalyse two or more consecutive steps in a metabolic sequence (Reed, 1974). The purpose of multienzyme complexes appear to be manifold. Reaction intermediates can be forced to complete the intended reaction sequence instead of escaping to conversion by enzymes that compete for the same reaction intermediate (substrate channelling) (Reed, 1974; Hammes, 1981). This is particular efficient if intermediates are covalently bound to the complex, like in the cases of the 2-oxo acid dehydrogenase complexes and the fatty acid synthases. Besides, by sequestering reactive intermediates, their conversion by undesired chemical reactions is also prevented (Perham, 1975). Another advantage of multienzyme complexes is that the catalytic activity can be enhanced because the local substrate concentrations are increased significantly (Reed, 1974). Since the enzyme activities in the complex are coupled, also better and more efficient regulation of the overall reaction is possible.

The large attention for the family of 2-oxo acid dehydrogenase multienzyme complexes originates certainly, at least in part, from the key positions they occupy in energy metabolism. The pyruvate dehydrogenase complex (PDHC) catalyses the irreversible oxidative decarboxylation of pyruvate to acetyl-CoA, linking the glycolysis with the tricarboxylic acid cycle. The 2-oxoglutarate dehydrogenase complex (OGDHC) converts 2-oxoglutarate into succinyl-CoA as part of the tricarboxylic acid cycle itself, and the branched-chain 2-oxo acid dehydrogenase complex (BCDHC) catalyses an irreversible step in the catabolism of the branched-chain amino acids by converting the 2-oxo acids derived from valine, leucine and isoleucine.

In this introduction I will review mainly developments concerning the structural and mechanistic features of the 2-oxo acid dehydrogenase complexes, with special emphasis on the structure and role of the lipoyl domains in the complex. Developments on genetic defects and regulation of the eukaryotic complexes have been reviewed elsewhere (Yeaman, 1989; Patel & Roche, 1990; Chuang *et al.*, 1991; Patel *et al.*, 1992; Behal *et al.*, 1993; Patel & Harris, 1995).

Structure and mechanism

2-Oxo acid dehydrogenase complexes are composed of multiple copies of at least three different enzymes, which are the major catalytic components: a substrate specific 2-oxo acid dehydrogenase (E1), a dihydrolipoyl acyltransferase (E2), and a common lipoamide dehydrogenase (E3). The complex-specific components are commonly abbreviated according to the origin of their complex, where E1p, E1o and E1b, and E2p, E2o and E2b, indicate the E1 and E2 components of PDHC, OGDHC, and BCDHC, respectively. The three components catalyse the set of sequential reactions as shown in Figure 1, involving three prosthetic groups (thiamin diphosphate (ThDP), lipoic acid and flavin adenine dinucleotide (FAD)), and two cofactors (NAD⁺ and CoA). The E1 component catalyses the oxidative decarboxylation of the 2-oxo acid and the subsequent reductive acylation of lipoic acid. The lipoic acid is covalently bound in an amide linkage to the N^ε group of a specific lysine residue of the E2 component (Nawa *et al.*, 1960), forming the so-called lipoyl group. The acyl group is then transferred to CoA, catalysed by the E2 component. Finally, the reduced lipoyl group is reoxidised by the E3 component with the concomitant reduction of NAD⁺.

In addition to the three major catalytic components, mammalian and yeast PDHCs also contain a component called protein X (De Marcucci & Lindsay, 1985), which is involved in the binding of the E3 component to E2 (Neaige & Lindsay, 1991; Lawson *et*

al., 1991a). Furthermore, eukaryotic PDHC and BCDHC contain an E1-specific kinase and phosphatase, which are involved in regulation of the complex activity by a phosphorylation/dephosphorylation mechanism (Linn *et al.*, 1969; Patel & Roche, 1990).

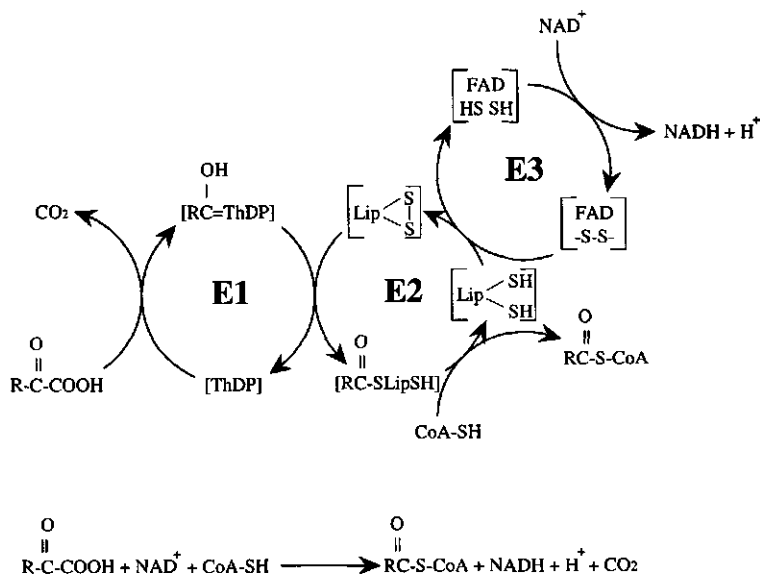


Figure 1. Reaction scheme for the oxidative decarboxylation of 2-oxo acids by the 2-oxo acid dehydrogenase complexes (Reed, 1974). R = CH₃ for the pyruvate dehydrogenase complex (PDHC), R = CH₂CH₂COOH for the 2-oxoglutarate dehydrogenase complex (OGDHC), and R = CH(CH₃)₂, CH₂CH(CH₃)₂ or CH(C₂H₅)(CH₃) for the branched-chain 2-oxo acid dehydrogenase complex (BCDHC). ThDP, thiamin diphosphate; Lip, lipoic acid.

The structural core of all 2-oxo acid dehydrogenase complexes is formed by an aggregate of the E2 component. The E2 components of PDHC from Gram-negative bacteria, and the OGDHC and BCDHC from all sources, except BCDHC from *Bacilli* (Perham & Packman, 1989), are assemblies of 24 identical subunits arranged with octahedral symmetry (Perham, 1991). The PDHC of mammals, yeast and Gram-positive bacteria form a core of 60 E2 subunits with icosahedral symmetry. Multiple copies of the peripheral components E1 and E3 bind tightly but noncovalently to the E2 core. Together this results in multienzyme complexes of enormous size (~5-14 MDa), which can easily be observed as large particles in electron micrographs (Oliver & Reed, 1982). A well-known exception is the PDHC core from *Azotobacter vinelandii*, which dissociates into functional

trimers upon binding of the peripheral components to the E2 component (Bosma *et al.*, 1984).

The E1 component (2-oxo acid dehydrogenase) catalyses the rate-limiting step in the overall complex reaction, the reductive acylation of lipoyl groups (Cate *et al.*, 1980). It is the least-characterised enzymatic component of the complex, and no structural information of any E1 component at atomic resolution is yet available. The E1 component exists in two forms, dependent on the type of complex and its symmetry (Perham, 1991). A homodimeric form (α_2) is found in PDHC and OGDHC with octahedral symmetry, and a heterodimeric ($\alpha_2\beta_2$) form is found in all BCDHCs, and in PDHCs with icosahedral cores. Despite a remarkable absence of sequence similarity among the different E1 subunits (e.g. E1p and E1o from *E. coli* (Stephens *et al.*, 1983a; Darlison *et al.*, 1984)), a common structural motif for a ThDP-binding site was proposed in all sequences of E1 components and other ThDP-dependent enzymes (Hawkins *et al.*, 1989). From the recent determination of the three-dimensional structures of a number of different ThDP-dependent enzymes, i.e. transketolase (Lindqvist *et al.*, 1992), pyruvate oxidase (Muller & Schulz, 1993) and pyruvate decarboxylase (Dyda *et al.*, 1993), this motif was shown to be involved in binding the metal ion and the diphosphate group (Lindqvist & Schneider, 1993; Muller *et al.*, 1993). The lack of structural information on E1 components seems partly due to the limited availability of a stable form of the enzyme. There are only a small number of reports on the recombinant expression of functional heterodimeric E1 (Wynn *et al.*, 1992; Lessard & Perham, 1994; Hester *et al.*, 1995), and only the first report this year on the expression of a functional homodimeric E1 (Berg *et al.*, 1996a). However, the work mentioned above seems promising enough to expect new structural information shortly.

In marked contrast with the E1 component, many structural and mechanistic details for the E3 component (lipoamide dehydrogenase) are available. The E3 component is usually common to all 2-oxo acid dehydrogenase complexes from the same source. Exceptions occur e.g. in *Pseudomonas putida*, where three different E3 enzymes have been demonstrated (Palmer *et al.*, 1991), and in *Enterococcus faecalis* where two *lpd* genes have been found, one related to the PDHC (Allen & Perham, 1991) and one of unknown function but probably related to BCDHC (Claiborn *et al.*, 1994). On the other hand, it has been shown that in pea leaf mitochondria the PDHC and the glycine decarboxylase complex share the same lipoamide dehydrogenase (Bourguignon *et al.*, 1996). Lipoamide dehydrogenase belongs to the family of flavin-dependent disulphide oxidoreductases and is a homodimeric enzyme. Four three-dimensional structures of E3 components have been solved by means of X-ray crystallography: E3 from *A. vinelandii* (Schierbeek *et al.*, 1989; Mattevi *et al.*, 1991), *Pseudomonas fluorescens* (Mattevi *et al.*, 1993a), *Bacillus*

stearothermophilus in complex with the PDHC binding domain (Mande *et al.*, 1996), and LPD-*val* from *P. putida* in complex with NAD⁺ (Mattevi *et al.*, 1992b). The enzyme catalyses the regeneration (oxidation) of the dihydrolipoyl group of the E2 component, using a ping-pong mechanism (Massey, 1960). In the first step the electrons are transferred from the reduced lipoyl group to a reactive disulphide group of the enzyme. In the second step the electrons are transferred via the FAD group to the final electron acceptor NAD⁺. Further details regarding structural and mechanistic properties of this enzyme are beyond the scope of this introduction and can be obtained elsewhere (Massey, 1960; Williams, 1992; De Kok & van Berkel, 1996).

The acyltransferase component

The acyltransferase (E2) components of all 2-oxo acid dehydrogenase complexes are multidomain proteins sharing a common but extraordinary design (Figure 2). Three different types of functional and separately folded domains have been disclosed from limited proteolysis studies (Bleile *et al.*, 1979; Bleile *et al.*, 1981; Packman *et al.*, 1984a; Packman *et al.*, 1984b; Chuang, 1985; Hanemaaijer *et al.*, 1987; Packman & Perham, 1987), amino acid sequence comparisons (Russell & Guest, 1991) and functional expression of separate domains by means of genetic engineering (Miles & Guest, 1987b; Dardel *et al.*, 1990; Schulze *et al.*, 1991a; Higgs & Perham, 1992; Berg *et al.*, 1994; Meng & Chuang, 1994).

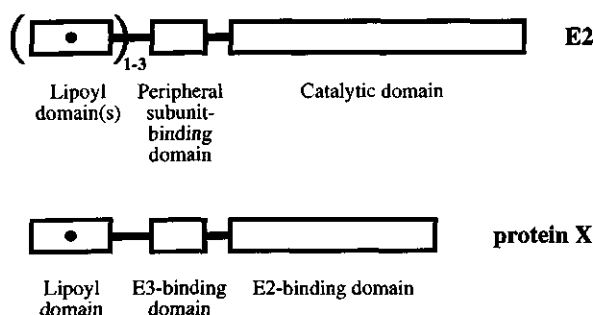


Figure 2. Schematic representation of the structural domains of the E2 components and protein X. The domains are connected by flexible linker segments (thick line). The approximate position of the lipoylation site in the lipoyl domains is indicated by a dot. PDHCs contain one, two or three lipoyl domains depending on the source, OGDHCs and BCDHCs contain only one lipoyl domain per E2 chain.

The E2 polypeptide chain contains at the N-terminus one to three lipoyl domains (~ 80 amino acid residues), each containing a covalently bound lipoyl group, followed by a peripheral subunit-binding domain (~ 35 amino acid residues) involved in the binding of the E3 and/or E1 components to the E2 core. The C-terminal catalytic domain (~ 29 kDa) accommodates the acyltransferase active site and the intersubunit binding sites responsible for the formation of the multimeric (24 or 60 subunit) core of the complex. The domains are linked by long (15 to 40 amino acid residues) flexible linker segments rich in alanine, proline and usually charged residues.

All OGDHCs and BCDHCs known so far contain a single lipoyl domain per E2 chain. In PDHCs, E2 components with one, two or three lipoyl domains are found. For example, *E. coli* (Stephens *et al.*, 1983b) and *A. vinelandii* E2p (Hanemaaijer *et al.*, 1988a) have three lipoyl domains, *Enterococcus faecalis* (Allen & Perham, 1991), *Haemophilus influenzae* (Fleischmann *et al.*, 1995), *Alcaligenes eutrophus* (Hein & Steinbüchel, 1994), *Acholeplasma laidlawii* (Wallbrandt *et al.*, 1992), *Neisseria meningitidis* (Ala'Aldeen *et al.*, 1995) and most mammalian E2ps (Thekkumkara *et al.*, 1988) possess two lipoyl domains, and *Bacillus subtilis* (Hämälä *et al.*, 1990), *B. stearothermophilus* (Borges *et al.*, 1990) and yeast E2p (Niu *et al.*, 1988) contain only one lipoyl domain. The lipoyl domains of the same E2p chain show a very high amino acid sequence identity. It appears that there is no obvious correlation between the number of lipoyl domains per E2 chain and the source or symmetry of the E2 core (Perham, 1991). The structure, function, expression and lipoylation of lipoyl domains are discussed in other paragraphs below.

The isolated peripheral subunit-binding domain (~35 amino acid residues) is one of the smallest proteins having a stable globular fold without the help of disulphide bridges or prosthetic groups (Brocklehurst *et al.*, 1994). The solution structures of the chemically synthesised binding domains of *E. coli* E2o (Robien *et al.*, 1992) and *B. stearothermophilus* E2p (Kalia *et al.*, 1993), and the crystal structure of the latter binding domain in complex with the E3 component (Mande *et al.*, 1996), have been solved. The global fold of the domain comprises two almost parallel α -helices, a short helix and a short and a long more disordered loop (Figure 3). The binding domain of *A. vinelandii* and *B. stearothermophilus* E2p seems to be involved in binding of both the E3 and E1 components (Hanemaaijer *et al.*, 1987; Packman *et al.*, 1988), as is the case for all BCDHCs (Wynn *et al.*, 1992). In other complexes with octahedral cores, like OGDHC from *E. coli* (Packman & Perham, 1986), the binding domain is responsible for binding only the E3 component. In eukaryotic PDHCs, the binding domain is involved in binding of the E1 component (Rahmatullah *et al.*, 1989a; Lawson *et al.*, 1991b), whereas the E3

component is bound to the complex by protein X (Rahmatullah *et al.*, 1989a; Neagle & Lindsay, 1991; Lawson *et al.*, 1991a).

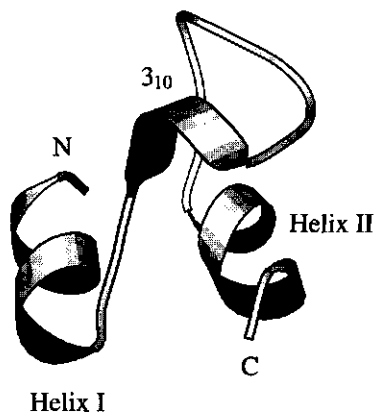


Figure 3. Schematic drawing (Kraulis, 1991) of the peripheral subunit-binding domain of *B. stearothermophilus* E2p (Kalia *et al.*, 1993), indicating its overall fold and the elements of secondary structure.

The structural core of all 2-oxo acid dehydrogenase complexes is formed by aggregation of the C-terminal catalytic domain of the E2 polypeptide chain. From early electron microscopy studies (Reed, 1974) and crystallographic symmetries observed in X-ray diffraction analysis (DeRosier *et al.*, 1971; Fuller *et al.*, 1979), a cubic or dodecahedral core formed by trimeric E2 units was already proposed (Reed & Hackert, 1990). Although for some time tetrameric E2 building blocks were suggested for the *A. vinelandii* E2p (Bosma *et al.*, 1984; Hanemaaijer *et al.*, 1989), the determination of the crystal structure of the catalytic domain of *A. vinelandii* E2p (Mattevi *et al.*, 1992c, 1993b) clearly showed a hollow 24-meric truncated cube with an edge of 125 Å, with trimers at its vertices (Figure 4). Based on the sequence similarity among E2p amino acid sequences, the same trimeric building blocks are also assumed for complexes with 60-meric dodecahedral cores (Mattevi *et al.*, 1992a, 1992c), for which no crystal structure is yet available.

The determination of the crystal structure of the catalytic domain revealed a high structural similarity between the trimers and chloramphenicol acetyltransferase (CAT), as predicted earlier on the basis of sequence homology between CAT and *E. coli* E2p (Guest, 1987). The active site is located at each interface of two E2 subunits in a trimer, forming a channel where lipoamide enters from the outside and CoA arrives from the inside of the cube. Guest (1987) also suggested for the E2p acetyltransferase reaction a similar reaction mechanism to CAT, with an active-site histidine residue acting as a general base. Since

then, a number of site-directed mutagenesis experiments on different E2 components of various sources have confirmed and refined the proposed reaction mechanism (Griffin & Chuang, 1990; Russell & Guest, 1990; Russell *et al.*, 1992; Meng & Chuang, 1994; Hendle *et al.*, 1995), with the single exception of yeast E2p, where substitution of the proposed active-site histidine by alanine or asparagine did not have a significant effect on the activity (Niu *et al.*, 1990).

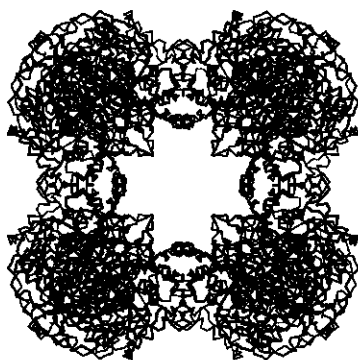


Figure 4. C α -trace (Kraulis, 1991) of the cubic catalytic core domain of *A. vinelandii* E2p (Mattevi *et al.*, 1992c).

The involvement of the histidine in the reaction mechanism has been confirmed by the crystal structures of binary and ternary complexes of the *A. vinelandii* catalytic domain with the substrates lipoamide and CoA (Mattevi *et al.*, 1993c).

The separate domains in the acyltransferase chain are connected to each other by flexible polypeptide segments of unusual composition with a large majority of alanine, proline, and usually charged amino acid residues. Their structure, mobility and role in active-site coupling will be discussed in the paragraph 'active-site coupling and role of linkers'.

Related complexes

In this intervening paragraph, a brief description of several multienzyme complexes and multicomponent enzymes that show some structural or functional relationship to the 2-oxo acid dehydrogenase complexes is given, aimed at a better understanding of several comparisons that are made with these complexes elsewhere in this introduction. The

glycine decarboxylase complex, the acetoin dehydrogenase enzyme system, the acetyl-coenzyme A carboxylase and the fatty acid synthase are considered.

The glycine decarboxylase complex, also known as the glycine cleavage system, catalyses the reversible oxidative decarboxylation and deamination of glycine yielding carbon dioxide, ammonia, NADH, and 5,10-methylene-5,6,7,8-tetrahydropteroyl-glutamic acid ($\text{CH}_2\text{H}_4\text{PteGlu}_n$) (Douce *et al.*, 1994). The latter component is recycled by serine hydroxymethyltransferase, which is linked to the complex. The glycine decarboxylase complex consists of four protein components, named P-, H-, T- and L-protein. The P-protein, a homodimer containing pyridoxal phosphate, catalyses the decarboxylation of glycine and transfer of the remaining methylamine to the lipoyl group of the H-protein. The T-protein catalyses the transfer of the methylene carbon from the H-protein to $\text{H}_4\text{PteGlu}_n$ forming $\text{CH}_2\text{H}_4\text{PteGlu}_n$ with the release of ammonia. Finally, the reduced lipoyl group of the H-protein is reoxidised by the L-protein, a lipoamide dehydrogenase. The H-protein couples the activities of the multienzyme complex enzymes with its covalently bound lipoyl group, analogous to the lipoyl domains of 2-oxo acid dehydrogenase complexes. The structure of the H-protein has shown to be similar to the lipoyl domain structures (see below).

The acetoin dehydrogenase enzyme system catalyses the cleavage of acetoin to acetyl-CoA and acetaldehyde, comparable to the oxidative decarboxylation of 2-oxo acids by the 2-oxo acid dehydrogenase complexes (Opperman *et al.*, 1989). This multienzyme complex also comprises multiple copies of three enzymes, a ThDP-dependent acetoin dehydrogenase (E1), a dihydrolipoamide acetyltransferase (E2), and lipoamide dehydrogenase (E3). The nucleotide-derived amino acid sequences of the acetoin dehydrogenase complex components of the Gram-negative bacteria *Pelobacter carbinolicus* (Opperman & Steinbüchel, 1994) and *Alcaligenes eutrophus* (Pfiefert *et al.*, 1991), and the Gram-positive bacterium *Clostridium magnum* (Krüger *et al.*, 1994) show a high similarity with each other and with the components of the 2-oxo acid dehydrogenase complexes. The E1 components of acetoin dehydrogenase enzyme systems are heterodimers, the E2 components have one or two lipoyl domains, and the E3 component was found to have an N-terminal lipoyl domain in *C. magnum*. No structural information on any of these components is yet available.

Acetyl-CoA carboxylase catalyses the first committed step in long-chain fatty acid biosynthesis, converting acetyl-CoA to malonyl-CoA. It belongs to the family of biotin-dependent carboxylases, which include also propionyl-CoA carboxylase, oxaloacetate decarboxylase, pyruvate decarboxylase and transcarboxylase (Toh *et al.*, 1993). They are composed of three subunits: a biotin carboxylase, a biotin carboxyl carrier protein, and a

carboxyltransferase. The biotin carboxylase catalyses the ATP-dependent carboxylation of biotin using bicarbonate. Biotin is linked to the N^ε group of a specific lysine residue of the biotin carboxyl carrier protein. It acts as a swinging arm to transfer the carboxyl group to the carboxyltransferase, much like the lipoyl domains of the 2-oxo acid dehydrogenase complexes. In acetyl-CoA carboxylase, the carboxyl group is finally transferred to acetyl-CoA to form malonyl-CoA, catalysed by the carboxyltransferase subunit. The three-dimensional structures of the *E. coli* biotin carboxylase and the biotin carboxyl carrier protein of acetyl-CoA carboxylase have been determined by X-ray crystallography (Waldrop *et al.*, 1994; Athappilly & Hendrickson, 1995). It shows that the structures of the biotin carboxyl carrier protein and lipoyl domains are strikingly similar (see below).

The fatty acid synthase system is considered here, because a swinging arm is involved in the catalytic mechanism (Hammes, 1981). The enzymes of fatty acid synthesis constitute a multienzyme complex, in which all reaction intermediates are bound to an acyl carrier protein. The acyl carrier protein contains a phosphopantetheine moiety covalently linked to the hydroxyl group of a serine residue. During catalysis, the acyl groups are bound via a thioester linkage to the -SH group of this prosthetic group, which serves as a swinging arm to deliver the acyl group at the different active sites. With the presence of a swinging arm the similarity with 2-oxo acid dehydrogenase complexes ends. There is no structural or sequential homology between the acyl carrier protein and lipoyl domains.

Subunit assembly and quaternary structure

The quaternary structure and the modes of binding of the individual components of 2-oxo acid dehydrogenase complexes would be easiest assessed by crystal structures of the complete multienzyme complexes, but the elucidation of these will probably be for a long time, if not forever, just a wish. Naturally, large amounts of labour have been put into obtaining well-ordered crystals of complete complexes or of isolated E2 components, but so far without any success. Likely, the high mobility of the linker segments and lipoyl domains prevented the growth of crystals diffracting at atomic resolution (Mattevi *et al.*, 1992c). However, the approach of structure determination of the individual components, domains and linkers of the complexes (Perham, 1991; Mattevi *et al.*, 1992a) has resulted recently in three-dimensional structures of various E3 components (see above) and of the three different domains of the E2 component (see above and below). Despite the fact that these tertiary structures have provided a wealth of new information and insights in many aspects of 2-oxo acid dehydrogenase complexes, a quaternary structure of the complex is of course not obtained by simply adding up these tertiary structures. But combining these

data with other biochemical data and electron microscopy studies could lead at least to an increased insight, although limited, into subunit assembly of these complexes.

The quaternary structure of the E2 component has already been discussed, and has shown to being assembled from 24 or 60 subunits in octahedral and icosahedral complexes, respectively. Multiple copies of the E1 and E3 components bind to the E2 cores, but not in stoichiometric amounts. In *E. coli* PDHC, 12 E1 dimers bind to the edges of the octahedral core, whilst 6 E3 dimers bind in the 6 faces of the cube (Koike *et al.*, 1963). The optimal E1o:E2o:E3 chain-stoichiometry for *E. coli* and *A. vinelandii* OGDHC have been estimated at 12:24:12 (Pettit *et al.*, 1973; Bosma, 1984). An unusual exception to the organisation of PDHCs from Gram-negative bacteria is the PDHC from *A. vinelandii*. As isolated, it consists of a trimeric E2 core to which two E1 dimers and a single E3 dimer are bound (Schulze *et al.*, 1992). With that it is the smallest 2-oxo acid dehydrogenase complex known ($M_r \sim 700$ kDa). Upon removal of the peripheral components the E2 component aggregates to the common 24-meric cubic core with 432 symmetry. However, in the presence of CoA or acetyl-CoA, the cubic core does not dissociate upon addition of the peripheral components, which suggests that *in vivo* this PDHC is also likely based on a 24-meric E2p core (Schulze *et al.*, 1993).

The icosahedral cores of mammalian, Gram-positive bacterial and yeast PDHC bind about 30 E1 tetramers ($\alpha_2\beta_2$) at the edges and 6 E3 dimers in the faces of the pentagonal dodecahedral E2 core (Henderson *et al.*, 1979; Wu & Reed, 1984). The E1p components are bound to the peripheral subunit-binding domain of the E2p component via their β -subunits (Wynn *et al.*, 1992; Lessard & Perham, 1995). In addition, mammalian and yeast PDHC contain 6 or 12 protein X subunits (Jilka *et al.*, 1986; Maeng *et al.*, 1994; Sanderson *et al.*, 1996), and one to three copies of a specific kinase and phosphatase (Reed, 1974; Reed & Hackert, 1990). Protein X has a similar domain structure as the E2 components (Figure 2), with one N-terminal lipoyl domain, a small E3-binding domain, and a C-terminal domain that binds protein X to the E2 core (Behal *et al.*, 1989; Rahmatullah *et al.*, 1989b; Lawson *et al.*, 1991a). The main function of protein X seems binding of the E3 dimers to the complex (Neagle & Lindsay, 1991; Lawson *et al.*, 1991a; Maeng *et al.*, 1994), although it has been suggested that the lipoyl domain of protein X can also play a role in the catalytic mechanism (Rahmatullah *et al.*, 1990; Lawson *et al.*, 1991b; Sanderson *et al.*, 1996). The E1 kinases were shown to associate to the mammalian PDHC via binding to, preferentially, the inner lipoyl domain of the E2 component (Radke *et al.*, 1993; Liu *et al.*, 1995a), which also appears to play a role in kinase stimulation (Ono *et al.*, 1993; Ravindran *et al.*, 1996). Finally, mammalian OGDHC lacks protein X, but the E2o component also lacks a sequence motif of the putative peripheral-subunit binding

domain (Nakano *et al.*, 1991). However, in the N-terminal region of the E1 α component a sequence similar to protein X is found to being involved in E3 binding (Rice *et al.*, 1992).

Many 2-oxo acid dehydrogenase complexes have shown to be self-assembling (Koike *et al.*, 1963; Reed *et al.*, 1975; Bates *et al.*, 1977; Bosma *et al.*, 1984), at least *in vitro*, which means that they can be functionally reconstituted from their individual components. Interestingly, reconstitution experiments have revealed that in cases where the binding domain is involved in binding of both peripheral components, there seems competition for binding sites during assembly, possibly caused by steric hindrance (Reed *et al.*, 1975). For example, the prokaryotic E2p components from *A. vinelandii* and *E. coli* can bind one peripheral component in a chain ratio of 1:2 for E2p:E1 or E2p:E3 in the absence of the other peripheral component (Reed *et al.*, 1975; Bosma *et al.*, 1984). Addition of the other component causes displacement of the bound component. For the *B. stearotherophilus* E2p it has been shown that the E1p and E3 components cannot bind simultaneously to the same isolated binding domain (Lessard & Perham, 1995). This suggests that one peripheral subunit-binding domain is involved in the binding of a dimer of E1(β) or E3. Other studies on the binding of E3 to the isolated di-domain (lipoyl domain plus binding domain) (Hippis *et al.*, 1994) or the complete E2p or E2 α (Westphal *et al.*, 1995) confirmed this suggestion. Furthermore, it has been shown that a dimeric E3 is essential for binding to the E2 component (Schulze *et al.*, 1991b).

An exciting observation by Westphal and co-workers was that binding to the *A. vinelandii* E2p stabilised the E3 by tightening the intersubunit interaction, making this component less sensitive against over-reduction (Westphal *et al.*, 1995). From this study it was also concluded that the E3 component must bind with its subunit interface near the dyad axis to the E2 component, thereby preventing sterically the binding of a second E3 dimer. This mode of interaction between E3 and the binding domain has recently been justified by the elucidation of the X-ray crystal structure of the *B. stearotherophilus* E3 component with the binding domain of E2p (Mande *et al.*, 1996). The E3 component was shown to bind mainly to the N-terminal part of the binding domain, via predominantly electrostatic interactions with both E3 subunits.

The interaction between the E1 and the E2 component has not been studied in such large detail. Reconstitution experiments of PDHCs based on chimeric E2p components from *A. vinelandii* and *E. coli* have shown that E1p interacts with both the binding domain and the catalytic domain of E2p (Schulze *et al.*, 1992). Site-directed mutagenesis experiments of *A. vinelandii* E2p already had indicated that the binding sites for E1p are located on the binding domain and catalytic domain (Schulze *et al.*, 1991c).

The spatial distribution of the complex components has been probed mainly by different electron microscopy techniques. From electron microscopy studies on negatively stained *E. coli* complexes (Oliver & Reed, 1982), and confirmed by cryoelectron microscopy of the frozen-hydrated complexes (Wagenknecht *et al.*, 1990), it has been shown that the E1 and E3 subunits are separated from the E2 core by a gap of 3-5 nm. These results have been interpreted by a flexible mode of attachment of E1 and E3 to the E2 core, conferred by the linker sequence connecting the peripheral subunit-binding domain with the core-forming catalytic domain. Similar results have been obtained for the mammalian PDHC, where the E1p and E3 components also appear not being bound directly to the E2p-X core (Wagenknecht *et al.*, 1991). Whether the suggested mobility of the peripheral components plays a role in active-site coupling remains to be determined. It could well be that a function as a spacer rather than being flexible is more important for this linker segment. It should be noted that scanning transmission electron microscopy (STEM) studies of cross-linked *E. coli* PDHC suggest more distinct E1 and E3 binding sites on the E2 core (CaJacob *et al.*, 1985; Yang *et al.*, 1986). Finally, various electron microscopy studies of *E. coli* PDHC, e.g. of negatively stained E2p cores (Bleile *et al.*, 1979), using cryoelectron microscopy (Wagenknecht *et al.*, 1990, 1992), or using STEM of gold cluster labelled lipoyl groups (Yang *et al.*, 1994), all indicate that the lipoyl domains extend from the surface of the E2p core. This has been confirmed by cross-linking studies of *E. coli* PDHC with avidin, which tightly binds lipoyl groups (Hale *et al.*, 1992).

Structure and role of lipoyl domains

The lipoyl groups, which are covalently attached to lipoyl domains, are essential for the coupling of the activities of the separate multienzyme components, by acting as reaction intermediate carriers. As such, they are substrates for the three different active sites in the multienzyme complex, and are indispensable for the efficient functioning of the complex. Lipoyl domains are independently folded and functioning protein units, as has been shown by reductive acylation of lipoyl domains obtained by limited proteolysis (Bleile *et al.*, 1981; Packman *et al.*, 1984a, 1984b) or by expression of sub-genes encoding them (Ali & Guest, 1990; Dardel *et al.*, 1990; Quinn *et al.*, 1993; Berg *et al.*, 1994, 1995; Liu *et al.*, 1995b). In combination with amino acid sequence comparisons of many acyltransferases (Russell & Guest, 1991; Matuda *et al.*, 1992; Dardel *et al.*, 1993) (see also Figure 5), it was concluded that lipoyl domains comprise approximately 80 residues,

A. vin (p1)	1	10	20	30	40	50	60	70	80
A. vin (p2)	-	SELIRVPDVG	--GDGEVIE	LLVKTGDLIE	VEQGLWLES	AKASMEVPSP	KAGVKSVSV	KLGDK-LKEG	DAIIIELEPAA
A. vin (p3)	-	SQEVKVPDVG	-SAGKARVIE	VLVKGAGDQVQ	AEQSLIVLES	DKASMEIPLSP	AGSVVESVAI	QLNAE-VGTG	DLIIILTRTTG
A. vin (o)	-	AIDIKAFTEP	ESIADGTVAT	WHKKPGEAVK	RDELIVDIET	DKVMEVLAE	ADGVIAEIVK	NEGDT-VLSG	ELLGKLTGEG
E. coli (p1)	-	AIEIKVDPDG	--ADEVEITE	ILVKVGDKVE	AEQSLITVEG	DKASMEVPSP	QAGIVKEIKV	SVGDK-TQTG	ALIMIFDSAD
E. coli (p2)	-	AKDENVDPDG	--SDEVEVTE	ILVKVGDKVE	AEQSLITVEG	DKASMEVPAP	FAGTVKEIKV	NVGDK-VSTG	SLIMVFEVAG
E. coli (p3)	-	VKEENVDPDG	-GDEVEVTE	VMKVGDKVA	AEQSLITVEG	DKASMEVPAP	FAGVVKELKV	NVGDK-VKTG	SLIMIFEVEG
E. coli (o)	S	SKDILVPLDP	ESVADATVAT	WHKKPDAVV	RDEVLVEIET	DKVLEVPAS	ADGLTDAVLE	DEGTT-VTSR	QILGRLREGN
H. inf (p1)	-	SQDIQIPDVG	--SDEVVNTE	VMNVGDITIS	VDQSLINVEG	DKASMEVPAP	EAGVVKELLV	KVGDK-VSTG	TPMLVLEAAG
H. inf (p2)	-	IVEENVDPDG	-GDEVVNTE	IMVAVGDITIS	AEQSLITVEG	DKASMEVPAP	FQGVVKEILV	KSGDK-VSTG	SLIMRFEVLG
H. inf (o)	-	AIEILVPLDP	ESVADATVAT	WHKKLGDTVK	RDEVIVEIET	DKVLEVPAL	SDGVLAENVQ	AEGET-VVSK	QLLGKLISTAQ
B. sub (p)	-	AFEFKLPDVG	EGHEGEIVK	WFVKNPDEV	EDVLAEVQN	DKAVVEIPLSP	VKGVLELKV	EEGTV-ATVG	QIIITFDAPG
B. sub (o)	-	AIEIKVPELA	ESISEGTTAQ	WLKQPGDYVE	EGVLELELET	DKAVNVEITAE	ESGVLEVLK	DSGDT-VQVG	EIIIGTISEGA
B. sub (b)	A	IEQMTMPQJG	ESVTEGTISK	WLVAPEGKVN	KYDPIAEVMT	DKVNAEVPSS	FTGTTITELVG	EEGQT-LQVG	EMICKIETEG
B. ste (p)	-	AFEFKLPDVG	EGHEGEIVK	WFVKNPDEVN	EDVLAEVQN	DKAVVEIPLSP	VKGVLELKV	PEGTV-ATVG	QTLITLDDAPG
S. cer (o)	-	HTIIGMPALS	PTMTQGNLAA	WTKKEGDQLS	PGEVIAEIEIET	DKAQMDEFQ	EDGYLAKILV	PEGTKDIPVN	KPLAVVVEDEK
Human (p1)	-	HQKVLPLSLS	PTMQAGTIAR	WKKKEGDKIN	EGDLLAEVET	DKATVGFESL	EECYMAKILV	AEGRDVPDG	AIICITVGEK
Human (p2)	-	HMQLVLPALS	PTMTMGTVQR	WEKKVGEKLS	EGDLLAEIET	DKATIGFEVQ	EEGLAKILV	PEGTRDVPDG	TPLCIIIVEKE
Human (o)	-	LVTVKTPAFA	ESVTEGDV-R	WEKAVGDTVA	EDEVVCEIET	DKTSVQVPS	ANGVIEALLV	PDGTV-VEGG	TPLFTLRKTG
Human (b)	-	VVQFKLSDIG	EGIREVTVKE	WVKEGDTVS	QFDSICEVQS	DKASVTTISR	YDGVIKKLYY	NLDDI-AVVG	KPLVDIETEA
Rat (p2)	-	HMQIVLPALS	PTMTMGTVQR	WEKKVGEKLS	EGDLLAEIET	DKATIGFEVQ	EEGLAKILV	PEGTRDVPDG	TPLCIIIVEKQ
Rat (o)	-	VITVQTPAFA	ESVTEGDV-R	WEKAVGDAVA	EDEVVCEIET	DKTSVQVPS	ANGVIEALLV	PDGTV-VEGG	TPLFTLRKTG
Bovine (p1)					VET	DKATVGF			
Bovine (o)					IET	DKTSVQVPS	ANG		
Bovine (b)	-	IVQFKLSDIG	EGIREVTVKE	WVKEGDTVS	QFDSICEVQS	DKASVTTISR	YDGVIKKLYY	NLDDT-AVVG	KPLVDIETEA

Figure 5. Amino acid sequence alignment of lipoyl domains of various sources. Only those lipoyl domains, for which the amino acid sequence of more than one type of complex from one source is known, are listed. The amino acid sequence of the lipoyl domain of *B. stearrowthermophilus* E2p is shown because a three-dimensional structure is available. The boldfaced lysine residue is the lipoylation site. A. vin(p), *A. vinelandii* E2p (Hanemaaijer *et al.*, 1988a); A. vin(o), *A. vinelandii* E2o (Westphal & de Kok, 1990); E. coli(p), *E. coli* E2p (Spencer *et al.*, 1984); H. inf(p), *Haemophilus influenzae* E2p (Fleischmann *et al.*, 1995); H. inf(o), *H. influenzae* E2o (Fleischmann *et al.*, 1995); B. sub(p), *Bacillus subtilis* E2p (Hämälä *et al.*, 1990); B. sub(o), *B. subtilis* E2o (Carlsson & Hederstedt, 1989); B. sub(b), *B. subtilis* E2b (Wang *et al.*, 1993); B. ste(p), *B. stearrowthermophilus* E2p (Borges *et al.*, 1990); S. cer(p), *Saccharomyces cerevisiae* E2p (Niu *et al.*, 1988); S. cer(o), *S. cerevisiae* E2o (Repetto & Tzagoloff, 1990); Human(p), human liver E2p (Thekkumkara *et al.*, 1988); Human(o), human E2o (Nakano *et al.*, 1994); Human(b), human E2h (Danner *et al.*, 1989); Rat(p), rat E2p (Matuda *et al.*, 1992); Rat(o), rat E2o (Nakano *et al.*, 1991); Bovine(p), bovine E2p (Bradford *et al.*, 1987a); Bovine(o), bovine E2o (Bradford *et al.*, 1987b); Bovine(b), bovine E2b (Griffin *et al.*, 1988).

each containing one fully conserved lysine residue as potential lipoylation site.

The structure of the lipoyl domain is not only required for the specific attachment of the lipoic acid prosthetic group (Wallis & Perham, 1994), but also increases the efficiency of reductive acylation of its lipoyl group dramatically. While free lipoamide or lipoic acid are good substrates for the E2 and E3 components, lipoamide is an extremely poor substrate for the E1 component (Reed *et al.*, 1958a). Likewise, reductive acetylation by E1p of a lipoylated decapeptide, with an amino acid sequence matching that surrounding the lipoylation site of *E. coli* E2p, is only barely detectable (Graham *et al.*, 1989). However, lipoyl groups, when bound to the lipoyl domains, are readily and efficiently reductively acylated by their appropriate E1 components (Bleile *et al.*, 1981; Packman *et al.*, 1984a, 1984b; Berg *et al.*, 1994, 1995). It has been suggested, on the basis of the large difference between the K_m (~ 33 μM) and the K_s (> 0.3 mM), that the enlarged efficiency of reductive acylation of the lipoyl group by a folded lipoyl domain is not directly a matter of enhanced binding to the E1 component (Graham & Perham, 1990).

The lipoyl domain is also responsible, at least in part, for the specificity of the reductive acylation reaction. Lipoyl domains are only efficiently reductively acylated by the E1 component of their parent complex, as has been shown for the *E. coli* (Graham *et al.*, 1989) complexes, and subsequently for the *A. vinelandii* complexes (chapter 6, this thesis). Reduced overall activity of reconstituted *E. coli* PDHC containing *A. vinelandii* E1p (De Kok & Westphal, 1985), and of *E. coli* and *A. vinelandii* PDHCs containing each others engineered lipoyl domains (Schulze *et al.*, 1992), is also ascribed to reduced efficiency of reductive acylation. Together this indicates that molecular recognition occurs between lipoyl domains and E1 components. It is obvious, however, that a complete picture of the specific molecular interactions involved in recognition of lipoyl domains is impaired by the lack of structural information of the E1 component at atomic resolution.

Another very intriguing question regarding lipoyl domains is why a number of PDHCs have more than one (two or three) lipoyl domain per E2 chain? It has been shown for the *E. coli* PDHC, containing three lipoyl domains per E2p chain, that nearly half of the lipoyl domains can be removed by limited proteolysis without significant loss in overall complex activity (Berman *et al.*, 1981; Stepp *et al.*, 1981). Likewise, the rate of chemical modification of enzymatic excision of lipoyl groups was shown to be faster than the rate at which complex activity decreased (Ambrose-Griffin *et al.*, 1980; Berman *et al.*, 1981; Danson *et al.*, 1981; Stepp *et al.*, 1981). Furthermore, by genetic engineering two of the three *E. coli* PDHC lipoyl domains can be removed with no apparent effect on overall complex activity or active-site coupling (Guest *et al.*, 1985; Graham *et al.*, 1986). This is explained by an active-site coupling mechanism in which the rate-limiting E1 component

can serve many lipoyl domains, and in which acyl groups can rapidly transfer among different lipoyl domains (Bates *et al.*, 1977; Collins & Reed, 1977; Danson *et al.*, 1978a, 1978b). In this way the function of the removed or inactivated lipoyl domains can be taken over by the remainders. However, this mechanism does still not explain the apparent excess of lipoyl domains in a number of PDHCs. Only recently, an alternative approach of comparing isogenic strains of *E. coli* containing PDHCs with one, two or three lipoyl domains per E2p chain, showed that the maximum growth rates of these strains in minimal medium are directly correlated with the number of lipoyl domains (Dave *et al.*, 1995). These results show at least the advantage for *E. coli* having PDHC with three lipoyl domains per E2p chain for efficient balanced growth on carbon sources that need this complex in their metabolic route.

Lipoyl domains are not exclusively found at the N-terminal part of the E2 chains of 2-oxo acid dehydrogenase complexes and acetoin dehydrogenase enzyme systems. As mentioned earlier, the protein X component of eukaryotic PDHCs also contains an N-terminal lipoyl domain, and it has been shown that this lipoyl domain is able to function in the overall complex reaction (Rahmatullah *et al.*, 1990; Lawson *et al.*, 1991b). The lipoyl domain of protein X of *S. cerevisiae* shows about 50% amino acid sequence identity to the lipoyl domains of its E2p (Behal *et al.*, 1989). Only recently, the PDHCs from *Alcaligenes eutrophus* (Hein & Steinbüchel, 1994) and *Neisseria meningitidis* (Ala'Aldeen *et al.*, 1996), and the acetoin dehydrogenase enzyme system from *Clostridium magnum* (Krüger *et al.*, 1994), were found to have an E3 component containing a lipoyl domain connected to its N-terminus by a linker segment. Although the role of these lipoyl domains in the multienzyme complex has not yet been established, their high amino acid sequence identity with the lipoyl domains of the E2 component suggests that they could take part in the overall reaction.

An additional role for the lipoyl domains of mammalian PDHCs has found to be involvement in binding of the pyruvate dehydrogenase kinase. It was shown that the kinase selectively binds to the inner lipoyl domain of the two lipoyl domains of the mammalian E2 component (Liu *et al.*, 1995a), and that this association involves the hydrophobic inner portion of the lipoyl group (Radke *et al.*, 1993). The kinase activity is regulated through the redox state of the inner lipoyl domain, showing an increased activity upon reduction or acetylation of the lipoyl group (Ravindran *et al.*, 1996). To explain the rapid phosphorylation of many E1 components by a limited number of kinase molecules, a mechanism has been proposed in which the bound kinase directly moves between the different inner lipoyl domains without dissociating from the complex (Ono *et al.*, 1993; Liu *et al.*, 1995a).

Recently, several three-dimensional structures of lipoyl domains have been solved, all by means of NMR spectroscopy. These are the single lipoyl domain of *B. stearrowthermophilus* PDHC (Dardel *et al.*, 1993), a non-native hybrid lipoyl domain of *E. coli* PDHC (Green *et al.*, 1995a) and the N-terminal lipoyl domains of *A. vinelandii* PDHC and OGDHC (Berg *et al.*, 1996b), which are described in this thesis. All lipoyl domain structures show a very similar overall fold, which is now considered a new class of all- β folds called β -barrel-sandwich hybrids (Chothia & Murzin, 1993), or flattened β -barrels (Green *et al.*, 1995a). The structure of the lipoyl domain is formed by two very similar four-stranded antiparallel β -sheets, which are packed around a core of hydrophobic residues in a sandwich-like manner (Figure 6).

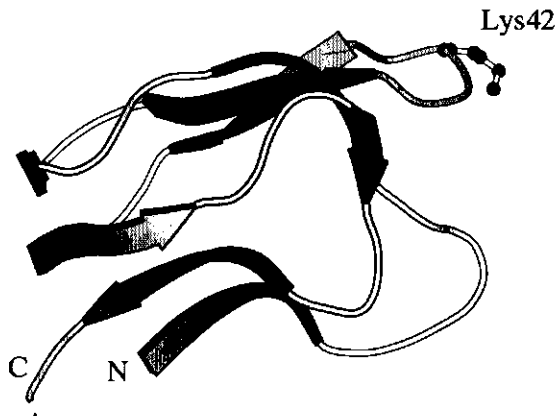


Figure 6. Schematic drawing (Kraulis, 1991) of the lipoyl domain of *B. stearrowthermophilus* PDHC (Dardel *et al.*, 1993). Lys42 is the lipoylation site.

The lipoylation site is exposed in a β -turn at the far end of one of the sheets, while the N-terminus and C-terminus meet at the opposite side of the domain, in two adjacent β -strands in the other β -sheet. The lipoyl domain displays a remarkable internal symmetry, relating the two halves of the molecule by a two-fold rotational axis. Regarding the lipoyl domain structures determined so far, in combination with the conservation of key-residues in all lipoyl domain amino acid sequences, it is concluded that all lipoyl domains will have highly similar folds (Dardel *et al.*, 1993; Green *et al.*, 1995a; Berg *et al.*, 1996b). Further structural details and comparisons of lipoyl domains can be found in the chapters 4 and 5. A final noticeable feature of lipoyl domains is that their structure is not altered by lipoylation. In the NMR spectra only very small differences in chemical shifts of residues close to the lipoylation site are observed between the unlipoylated and lipoylated forms of the lipoyl domain (Dardel *et al.*, 1991; Berg *et al.*, 1994).

Lipoyl domains derived from 2-oxo acid dehydrogenase complexes show high structural homology to the H-protein of the glycine decarboxylase system and the biotinyl domain of acetyl-CoA carboxylase, as had been predicted on the basis of (low) sequence similarity to lipoyl domains (Brocklehurst & Perham, 1993; Toh *et al.*, 1993). The structure of the H-protein has been solved by X-ray crystallography (Pares *et al.*, 1994; Cohen-Addad *et al.*, 1995), and consists of a β -barrel-sandwich structure similar to the lipoyl domain (Figure 7). The lipoyl-lysine is analogously presented in a β -hairpin turn and is rather flexible, as concluded from the relatively high B-factors. The N-terminal exposed loop of the lipoyl domain is replaced by a helix in the H-protein, which is also in proximity of the lipoylation site. The H-protein (~ 130 amino acid residues) is larger than a lipoyl domain (~ 80 amino acid residues) and contains two additional β -strands at the N-terminal end, and a short and a long C-terminal helix. Very interestingly, the X-ray crystal structure of the methylamine loaded form of the H-protein shows that the lipoyl-methylamine group interacts with several specific conserved residues, located in a cleft formed by the β -sandwich and the N-terminal helix (Cohen-Addad *et al.*, 1995). The strong interactions between the protein and the methylamine group explain why the methylamine-loaded form of the H-protein is stable (Neuberger *et al.*, 1991), and show that in this form the lipoyl group is not free to rotate.

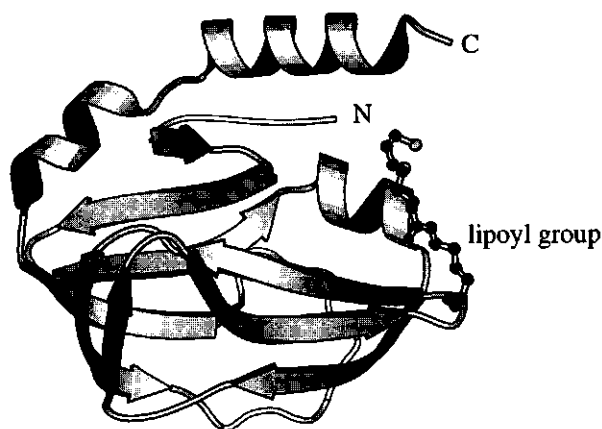


Figure 7. Schematic drawing (Kraulis, 1991) of the H-protein of the glycine decarboxylase complex from pea leaves (Pares *et al.*, 1994). The lipoyl group is represented in ball-and-stick.

The X-ray crystal structure of the biotinyl domain of acetyl-CoA carboxylase from *E. coli* is very similar to the structure of lipoyl domains (Athappilly & Hendrickson, 1995). The only difference in overall fold is the absence of the exposed N-terminal loop of the lipoyl domain in the biotinyl domain, which in turn possesses a large loop between the second and third β -strand (Figure 8). The biotinylated lysine residue resides in a β -hairpin turn, a structural feature which seems conserved in all proteins containing lipoic acid or biotin. The biotinyl group is well defined in the electron-density map, and interacts with residues of the large loop that is absent in lipoyl domains. This indicates that the biotinyl group is not completely free to swing, at least in its noncarboxylated form, but is partially buried in the surface of the domain.

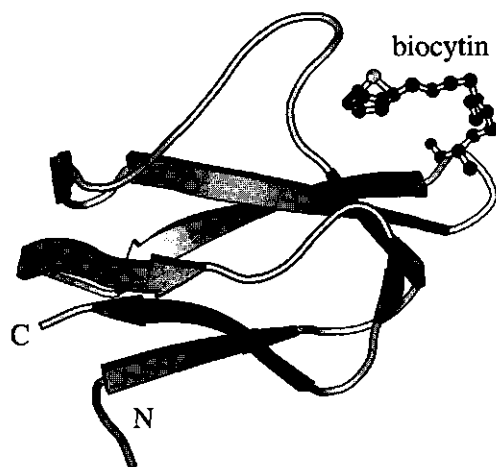


Figure 8. Schematic drawing (Kraulis, 1991) of the biotinyl domain of acetyl-CoA carboxylase from *E. coli* (Athappilly & Hendrickson, 1995). The biocytin is shown in ball-and-stick representation.

Lipoylation

Lipoic acid (6,8-thioctic acid or 1,2-dithiolane-3-pentanoic acid) is the prosthetic group of the lipoyl domains of 2-oxo acid dehydrogenase complexes and acetoin dehydrogenase enzyme systems, and of the H-protein of the glycine decarboxylase complex. The carboxyl group of lipoic acid is bound in an amide linkage to the N^ε group of a specific lysine residue in a posttranslational modification process called lipoylation. Until recently, very little was known about the enzyme(s) and mechanism(s) involved in lipoylation of apo-lipoyl domains, and about the biosynthesis of lipoic acid, which will not

be discussed here. Early work by the group of Lester J. Reed on the lipoylating systems of *E. coli* and *S. faecalis* showed that incorporation of lipoic acid into apo-PDHC required lipoic acid, ATP, inorganic phosphate and a divalent metal ion, and that lipoic acid and ATP could be replaced by lipoyladenylate (Reed *et al.*, 1958a, 1958b). Only after the observation of unlipoylated and mismodified (octanoylated) lipoyl domains after overexpression of their subgenes in *E. coli* under certain conditions, the attention to the lipoylation process seemed renewed.

Unlipoylated and octanoylated lipoyl domains have been observed in a number of cases. Expression of an *E. coli* PDHC lipoyl domain in *E. coli* produced a mixture of the lipoylated and unlipoylated form of the domain (Miles & Guest, 1987b; Ali & Guest, 1990). Likewise, lipoylated and unlipoylated forms were detected when the human PDHC inner lipoyl domain (Quinn *et al.*, 1993), the lipoyl domains of *A. vinelandii* PDHC and OGDHC (Berg *et al.*, 1994, 1995), and the H-protein from pea (Macherel *et al.*, 1996) were overproduced in *E. coli*. Addition of exogenous lipoic acid to the growth medium commonly resulted in increased amounts of the lipoylated form, indicating that the overexpression exceeded the cell's capacity for lipoylation. In addition to the lipoylated and unlipoylated forms, small amounts of an octanoylated form were observed when the lipoyl domains of *B. stearothersophilus* and human PDHC, and the H-protein of bovine glycine decarboxylase complex were expressed in *E. coli* (Dardel *et al.*, 1990; Fujiwara *et al.*, 1992; Hipps & Perham, 1992; Liu *et al.*, 1995b), and when *E. coli* PDHC lipoyl domains were expressed in a lipoate-deficient *E. coli* strain (Ali *et al.*, 1990). The identification of the different forms of lipoyl domains and H-proteins was greatly facilitated by the application of new and advanced mass-spectrometry techniques, e.g. electrospray mass spectrometry.

It is interesting that lipoyl domains and H-proteins from different sources are being lipoylated by the lipoylating system of *E. coli*, with the exception of the lipoyl domain of bovine BCDHC which is not lipoylated (Griffin *et al.*, 1990). The lipoyl domain and its particular lipoyl-lysine residue are selected specifically for lipoylation, and the question arises how this site is recognised by the lipoylating enzyme(s)? This question was addressed by site-directed mutagenesis experiments of residues about the lipoyl-lysine residue of the *B. stearothersophilus* PDHC lipoyl domain (Wallis & Perham, 1994). They showed that the position of the lipoyl-lysine residue in the β -turn where it is found, is essential for lipoylation, rather than the residues directly surrounding it.

Lipoate protein ligases, responsible for the lipoylation reaction, have been isolated from *E. coli* (Brookfield *et al.*, 1991; Green *et al.*, 1995b) and from bovine liver mitochondria (Fujiwara *et al.*, 1994). In *E. coli* two distinct genes have been cloned

(Morris *et al.*, 1994, 1995), the products of which are involved in lipoylation. The *lplA* gene product is responsible for the incorporation of exogenous lipoic acid, via a mechanism using a lipoyl-AMP intermediate, which is consistent with the early observations by Reed (Reed *et al.*, 1958a, 1958b), and analogous to the biotin protein ligase (Cronan, 1989). The other lipoate protein ligase, the *lipB* gene product, utilises lipoyl groups generated via endogenous lipoic acid biosynthesis. This indicates that two redundant pathways with two different lipoate protein ligases for lipoylation exist in *E. coli*. Two isoforms of lipoyltransferase were purified from bovine liver mitochondria that could use lipoyl-AMP but not lipoic acid plus MgATP for lipoylation (Fujiwara *et al.*, 1994), suggesting that two enzymes are involved in the complete lipoylation reaction.

From the amino acid sequences of E2 components the number of potential lipoylation sites can be determined. The extent of lipoylation of the potential sites has been a subject of controversy for a long time. Various studies using different methods of determination, mainly on the E2p component of *E. coli*, resulted in numbers ranging from 1.7 to 2.0 lipoyl groups per E2p chain (Packman *et al.*, 1984a). However, re-assessment of the number of functional lipoyl groups per E2p and E2o chain by a combination of protein-chemical and modern mass-spectrometric techniques clearly showed that all potential lipoylation sites contain functional lipoyl groups (Packman *et al.*, 1991).

Active-site coupling and role of linkers

The coupling of the activities of the three enzymatic components of the 2-oxo acid dehydrogenase complexes is brought about by their lipoyl groups. A swinging arm mechanism, in which the long and flexible lipoyl groups rotate among the different active sites, had been proposed responsible for active-site coupling (Koike *et al.*, 1963). However, by fluorescence energy transfer measurements it was shown that the distances between the active sites in PDHC were at least 40 Å (Shepherd & Hammes, 1977), a gap that cannot be narrowed by a rotating lipoyl group (~ 28 Å). Thus, more than just a swinging arm is required for active-site coupling, and the participation of two or more lipoyl groups in a catalytic cycle, or additional movement of protein parts, was suggested to be involved in the mechanism.

The mechanism of active-site coupling is more complicated than the simple direct transfer of an intermediate from an active site via the lipoyl group to the subsequent active site. The observation that only a few E1 dimers can reductively acylate many lipoyl groups in *E. coli* PDHC and OGDHC (so-called servicing experiments) (Bates *et al.*, 1977; Collins & Reed, 1977; Danson *et al.*, 1978a), and at a rate comparable with the overall complex

activity (Danson *et al.*, 1978b), suggests that acyl-transfer reactions among lipoyl groups need to be accounted for in the active-site coupling mechanism. The involvement of intramolecular transacylation reactions can be used to explain the observations that the removal of lipoyl domains by limited proteolysis (Berman *et al.*, 1981; Stepp *et al.*, 1981) or genetic engineering (Guest *et al.*, 1985), or enzymatic release or chemical inactivation of the lipoyl groups (Ambrose-Griffin *et al.*, 1980; Berman *et al.*, 1981; Danson *et al.*, 1981; Stepp *et al.*, 1981) proceed faster than the accompanying loss of overall complex activity. The kinetics of inactivation have been simulated by a computer model, using a multiple random coupling mechanism for active-site coupling (Hackert *et al.*, 1983a, 1983b). This mechanism was supported by site-directed mutagenesis experiments of *E. coli* PDHC, in which was shown that various permutations of functional and non-functional (by lipoyl-lysine to glutamine mutations) lipoyl domains did not effect active-site coupling (Allen *et al.*, 1989). These elegant experiments strongly suggest that the three lipoyl domains function completely independent and that the reductive acetylation of them is random order.

The multiple random coupling mechanism suggests a rapid transfer of acyl groups among lipoyl groups. From the X-ray crystal structures of the E2 and E3 components, this acyl transfer is expected to be a chemical reaction and not an enzymatically catalysed reaction, which means that both the 6-S-acyldihydrolipoyl group and the 8-S-acyldihydrolipoyl group will be formed. Since the 6-S-acyldihydrolipoyl group is likely not converted by the E2 component, a rapid decrease of the complex activity during servicing experiments could be expected, but is, however, not observed. Chemical intramolecular isomerisation of the inactive enzyme-bound 6-S-acyldihydrolipoyl groups to 8-S-acyldihydrolipoyl groups is too slow during normal catalytic turnover (Yang & Frey, 1986) to compensate for this. The question why rapid inactivation of the complexes during servicing experiments, due to formation of the inactive 6-S-acyldihydrolipoyl groups, is not observed, has not been addressed until now and remains to be solved.

The observation of unexpectedly sharp resonances in the $^1\text{H-NMR}$ spectrum of the *E. coli* PDHC ($M_r \sim 5$ MDa) indicated the presence of conformational mobile regions in E2p (Perham *et al.*, 1981). Later on, similar signals have also been detected in other complexes, like *E. coli* OGDHC (Perham & Roberts, 1981), *B. stearothermophilus* PDHC (Duckworth *et al.*, 1982), and *A. vinelandii* PDHC, OGDHC and E2p (Hanemaaijer *et al.*, 1988b). These signals were ascribed to the alanine and proline rich linker segments in the E2 chain, which could be important for facilitating movement of lipoyl domains for active-site coupling. This was supported by the similarity of $^1\text{H-NMR}$ spectra of a 32-residue synthetic peptide representing the amino acid sequence of a linker (Radford *et al.*, 1986),

and by the reduction of sharp signals in the spectrum of *E. coli* PDHC with two deleted lipoyl domains and linkers (Radford *et al.*, 1987). Direct evidence for mobility of the linker segments came from the observation of sharp ¹H-NMR signals assigned to a histidine residue that was introduced in the interdomain segment of a mutant *E. coli* PDHC with only one lipoyl domain (Texter *et al.*, 1988).

A more detailed analysis, by means of NMR spectroscopy and circular dichroism, of several synthetic peptides with amino acid sequences representing *E. coli* PDHC linkers, showed that these peptides are very flexible in solution (Radford *et al.*, 1989; Green *et al.*, 1992). Moreover, it was shown that their structures were disordered but not random coil, and in particular that all Ala-Pro peptide bonds were in the *trans* conformation. This suggests a certain stiffening of the flexible linkers, which would facilitate protrusion from the core of the complex.

The importance of the flexibility of the linkers for active-site coupling was shown by making deletions in the 32-residue linker segment of an *E. coli* PDHC mutant containing one lipoyl domain (Miles *et al.*, 1988). Reduction of the linker to less than 13 residues caused significant loss in active-site coupling. Studies in which different linkers, varying in length and composition, were engineered in this complex confirmed the importance of the linker size (Turner *et al.*, 1993). They also showed that the amino acid composition of the linker is of less importance, although highly charged linkers were not allowed. Finally, the flexibility and importance of the short linker segment connecting the peripheral subunit-binding domain with the catalytic domain for active-site coupling needs to be established unequivocally (Schulze *et al.*, 1993), and may vary between different 2-oxo acid dehydrogenase complexes (Perham, 1991).

Outline of this thesis

In this thesis the determination of the three-dimensional structures of the N-terminal lipoyl domains of *A. vinelandii* PDHC and OGDHC by means of NMR spectroscopy is described. The structure of the lipoyl domain is of particular interest and importance for several reasons. The attachment of the lipoyl group to a lipoyl domain dramatically enhances the efficiency of reductive acylation of the reactive dithiolane ring by the 2-oxo acid dehydrogenase (E1) component. Moreover, the lipoyl domain causes its lipoyl group to be only an effective substrate for the E1 of its parent multienzyme complex, and thus effectuates the reductive acylation reaction to being specific. It is therefore clear that a process of molecular recognition occurs between E1 components and lipoyl domains. Besides, the selection of a specific lysine residue of the lipoyl domain by the

enzymes responsible for its lipoylation is also determined by the structure of the lipoyl domain.

The present study aimed at gaining a more detailed insight in the structural basis for the specific interaction between lipoyl domains and E1 components, and in particular what residues or structural features of the lipoyl domain are involved in this process. The determination of the three-dimensional structure of lipoyl domains is naturally a prerequisite for such studies. Their size and apparent stability make the lipoyl domains suitable for a structure determination using NMR spectroscopy.

To obtain sufficient amounts of the N-terminal lipoyl domains of *A. vinelandii* PDHC and OGDHC for structural investigations, sub-genes encoding these lipoyl domains (79 amino acid residues) were expressed in *E. coli* (chapters 2 and 3). The lipoyl domains were isotopically enriched with ^{15}N to facilitate the NMR spectral assignments. The sequential ^1H and ^{15}N assignments of the N-terminal lipoyl domain of PDHC, as well as the derivation of its secondary structure in solution, are described in chapter 2. Furthermore, the ^1H -NMR spectra of the separately isolated lipoylated and nonlipoylated forms of this lipoyl domain are compared, and several ^1H resonances of the lipoyl group were assigned. The sequential ^1H and ^{15}N assignments and secondary structure of the single lipoyl domain of OGDHC were also obtained (chapter 3). From a comparison of a variety of NMR-derived parameters of the PDHC and OGDHC lipoyl domains, it is suggested that their structures in solution are very similar.

The determination of the three-dimensional solution structures of the PDHC and OGDHC lipoyl domains is described in the chapters 4 and 5, respectively. These structures are compared with each other, and with the structures of the lipoyl domains of *B. stearothermophilus* PDHC and *E. coli* PDHC, which became available during the course of the structure determination of the *A. vinelandii* lipoyl domains. On the basis of a comparison of these structures and many lipoyl domain amino acid sequences, several residues are proposed, including those of a solvent-exposed loop close in space to the lipoylation site, that could be responsible for the specific recognition of lipoyl domains by E1. Cross-acylation experiments of *A. vinelandii* PDHC and OGDHC lipoyl domains catalysed by *A. vinelandii* and *E. coli* complexes, and site-directed mutagenesis experiments of the exposed loop of the OGDHC lipoyl domain were performed, to study the role of this loop in molecular recognition (chapter 6).

References

- Ala'Aldeen, D. A. A., Weston, V., Baldwin, T. J. & Borriello, S. P. (1995) The gene cluster encoding components of the pyruvate dehydrogenase complex of *Neisseria meningitidis*: detection and sequence analysis. *J. Med. Microbiol.* **42**, 148-159.
- Ala'Aldeen, D. A. A., Westphal, A. H., Weston, V., Atta, M., Baldwin, T. J., Bartley, J., de Kok, A. & Borriello, S. P. (1996) Cloning, sequencing, characterisation and implications on vaccine design of the novel dihydrolipoyl acetyltransferase of *Neisseria meningitidis*. *J. Med. Microbiol.* in the press.
- Ali, S. T. & Guest, J. R. (1990) Isolation and characterization of lipoylated and unlipoylated domains of the E2p subunit of the pyruvate dehydrogenase complex of *Escherichia coli*. *Biochem. J.* **271**, 139-145.
- Ali, S. T., Moir, A. J. G., Ashton, P. R., Engel, P. C. & Guest, J. R. (1990) Octanoylation of the lipoyl domains of the pyruvate dehydrogenase complex in a lipoyl-deficient strain of *Escherichia coli*. *Mol. Microbiol.* **4**, 943-950.
- Allen, A. G., Perham, R. N., Allison, N., Miles, J. S. & Guest, J. R. (1989) Reductive acetylation of tandemly repeated lipoyl domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli* is random order. *J. Mol. Biol.* **208**, 623-633.
- Allen, A. G. & Perham, R. N. (1991) Two lipoyl domains in the dihydrolipoamide acetyltransferase chain of the pyruvate dehydrogenase multienzyme complex of *Streptococcus faecalis*. *FEBS Lett.* **287**, 206-210.
- Ambrose-Griffin, M. C., Danson, M. J., Griffin, W. G., Hale, G. & Perham, R. N. (1980) Kinetic analysis of the role of lipoic acid residues in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochem. J.* **187**, 393-401.
- Athappilly, F. K. & Hendrickson, W. A. (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**, 1407-1419.
- Bates, D. L., Danson, M. J., Hale, G., Hooper, E. A. & Perham, R. N. (1977) Self-assembly and catalytic activity of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Nature* **268**, 313-316.
- Behal, R. H., Browning, K. S., Hall, T. B. & Reed, L. J. (1989) Cloning and nucleotide sequence of the gene for protein X from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **86**, 8732-8736.
- Behal, R. H., Buxton, D. B., Robertson, J. G. & Olson, M. S. (1993) Regulation of the pyruvate dehydrogenase multienzyme complex. *Annu. Rev. Nutr.* **13**, 497-520.
- Berg, A., de Kok, A. & Vervoort, J. (1994) Sequential ¹H and ¹⁵N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **221**, 87-100.
- Berg, A., Smits, O., de Kok, A. & Vervoort, J. (1995) Sequential ¹H and ¹⁵N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex. *Eur. J. Biochem.* **234**, 148-159.
- Berg, A., Bosma, H., Westphal, A. H., Hengeveld, J., Vervoort, J. & de Kok, A. (1996a) The pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Structure and function of the lipoyl domain. In: *Biochemistry and physiology of thiamin diphosphate enzymes* (Bisswanger, H. & Schellenberger, A., eds.), pp. 278-291, A.u. C. Intemann, Wissenschaftlicher Verlag, Prien, Germany.
- Berg, A., Vervoort, J. & de Kok, A. (1996b) Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. *J. Mol. Biol.* **261**, 432-442.
- Berman, J. N., Chen, G.-X., Hale, G. & Perham, R. N. (1981) Lipoic acid residues in a take-over mechanism for the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochem. J.* **199**, 513-520.
- Bleile, D. M., Munk, P., Oliver, R. M. & Reed, L. J. (1979) Subunit structure of dihydrolipoyl transacetylase component of pyruvate dehydrogenase complex from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **76**, 4385-4389.
- Bleile, D. M., Hackert, M. L., Pettit, F. H. & Reed, L. J. (1981) Subunit structure of dihydrolipoyl transacetylase component of pyruvate dehydrogenase complex from bovine heart. *J. Biol. Chem.* **256**, 514-519.
- Borges, A., Hawkins, C. F., Packman, L. C. & Perham, R. N. (1990) Cloning and sequence analysis of the genes encoding the dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase components of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Eur. J. Biochem.* **194**, 95-102.

- Bosma, H. J. (1984) *Studies on 2-oxoacid dehydrogenase multienzyme complexes of Azotobacter vinelandii*. Ph.D. Thesis, Agricultural University Wageningen, The Netherlands.
- Bosma, H. J., de Kok, A., Westphal, A. H. & Veeger, C. (1984) The composition of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Does a unifying model exist for the complexes from gram-negative bacteria? *Eur. J. Biochem.* **142**, 541-549.
- Bourguignon, J., Merand, V., Rawsthorne, S., Forest, E. & Douce, R. (1996) Glycine decarboxylase and pyruvate dehydrogenase complexes share the same dihydrolipoamide dehydrogenase in pea leaf mitochondria: evidence from mass spectrometry and primary-structure analysis. *Biochem. J.* **313**, 229-234.
- Bradford, A. P., Howell, S., Aitken, A., James, L. A. & Yeaman, S. J. (1987a) Primary structure around the lipate attachment site on the E2 component of bovine heart pyruvate dehydrogenase complex. *Biochem. J.* **245**, 919-922.
- Bradford, A. P., Aitken, A., Beg, F., Cook, K. G. & Yeaman, S. J. (1987b) Amino acid sequence surrounding the lipoic acid cofactor of bovine kidney 2-oxoglutarate dehydrogenase complex. *FEBS Lett.* **222**, 211-214.
- Brocklehurst, S. M. & Perham, R. N. (1993) Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: A new automated method for the prediction of protein tertiary structure. *Protein Sci.* **2**, 626-639.
- Brocklehurst, S. M., Kalia, Y. N. & Perham, R. N. (1994) Protein-protein recognition mediated by a mini-protein domain: possible evolutionary significance. *TIBS* **19**, 360-361.
- Brookfield, D. E., Green, J., Ali, S. T., Machado, R. S. & Guest, J. R. (1991) Evidence for two protein-lipoylation activities in *Escherichia coli*. *FEBS Lett.* **295**, 13-16.
- CaJacob, C. A., Frey, P. A., Hainfeld, J. F., Wall, J. S. & Yang, H. (1985) *Escherichia coli* pyruvate dehydrogenase complex: particle masses of the complex and component enzymes measured by scanning transmission electron microscopy. *Biochemistry* **24**, 2425-2431.
- Carlsson, P. & Hederstedt, L. (1989) Genetic characterization of *Bacillus subtilis* *odhA* and *odhB*, encoding 2-oxoglutarate dehydrogenase and dihydrolipoamide transsuccinylase, respectively. *J. Bacteriol.* **171**, 3667-3672.
- Cate, R. L., Roche, T. E. & Davis, L. C. (1980) Rapid intersite transfer of acetyl groups and movement of pyruvate dehydrogenase component in the kidney pyruvate dehydrogenase complex. *J. Biol. Chem.* **255**, 7556-7562.
- Chothia, C. & Murzin, A. G. (1993) New folds for all- β proteins. *Structure* **1**, 217-222.
- Chuang, D. T., Hu, C. C., Ku, L. S., Markovitz, P. J. & Cox, R. P. (1985) Subunit structure of the dihydrolipoyl transacylase component of branched-chain α -keto acid dehydrogenase complex from bovine liver. Characterization of the inner transacylase core. *J. Biol. Chem.* **260**, 13779-13786.
- Chuang, D. T., Fisher, C. W., Lau, K. S., Griffin, T. A., Wynn, R. M. & Cox, R. P. (1991) Maple syrup urine disease: domain structure, mutations and exon skipping in the dihydrolipoyl transacylase (E2) component of the branched-chain α -keto acid dehydrogenase complex. *Mol. Biol. Med.* **8**, 49-63.
- Claiborn, A., Ross, R. P., Ward, D., Parsonage, D. & Crane III, E. J. (1994) Flavoprotein peroxide and disulfide reductases and their roles in Streptococcal oxidative metabolism. In: *Flavins and flavoproteins 1993* (Yagi, K., eds.), pp. 587-596, Walter de Gruyter, Berlin.
- Cohen-Addad, C., Pares, S., Sieker, L., Neuburger, M. & Douce, R. (1995) The lipoamide arm in the glycine decarboxylase complex is not freely swinging. *Nat. Struct. Biol.* **2**, 63-68.
- Collins, J. H. & Reed, L. J. (1977) Acyl group and electron pair relay system: a network of interacting lipoyl moieties in the pyruvate and α -ketoglutarate dehydrogenase complexes from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **74**, 4223-4227.
- Cronan, J. E. (1989) The *E. coli* *bio* operon: transcriptional repression by an essential protein modification enzyme. *Cell* **58**, 427-429.
- Danner, D. J., Litwer, S., Herring, W. J. & Pruckler, J. (1989) Construction and nucleotide sequence of a cDNA encoding the full-length preprotein for human branched-chain acyltransferase. *J. Biol. Chem.* **264**, 7742-7746.
- Danson, M. J., Hooper, E. A. & Perham, R. N. (1978a) Intramolecular coupling of active sites in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochem. J.* **175**, 193-198.
- Danson, M. J., Fersht, A. R. & Perham, R. N. (1978b) Rapid intramolecular coupling of active sites in the pyruvate dehydrogenase complex of *Escherichia coli*: Mechanism for rate enhancement in a multimeric structure. *Proc. Natl. Acad. Sci. USA* **75**, 5386-5390.
- Danson, M. J., Hale, G. & Perham, R. N. (1981) The role of lipoic acid residues in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochem. J.* **199**, 505-511.

- Dardel, F., Packman, L. C. & Perham, R. N. (1990) Expression in *Escherichia coli* of a sub-gene encoding the lipoyl domain of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *FEBS Lett.* **264**, 206-210.
- Dardel, F., Laue, E. D. & Perham, R. N. (1991) Sequence-specific ¹H-NMR assignments and secondary structure of the lipoyl domain of the *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *Eur. J. Biochem.* **201**, 203-209.
- Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- Darlison, M. G., Spencer, M. E. & Guest, J. R. (1984) Nucleotide sequence of the *sucA* gene encoding 2-oxoglutarate dehydrogenase of *Escherichia coli* K12. *Eur. J. Biochem.* **141**, 351-359.
- Dave, E., Guest, J. R. & Attwood, M. M. (1995) Metabolic engineering in *Escherichia coli*: lowering the lipoyl domain content of the pyruvate dehydrogenase complex adversely affects the growth rate and yield. *Microbiology* **141**, 1839-1849.
- De Kok, A. & Westphal, A. H. (1985) Hybrid pyruvate dehydrogenase complexes reconstituted from components of the complexes from *Escherichia coli* and *Azotobacter vinelandii*. *Eur. J. Biochem.* **152**, 35-41.
- De Kok, A. (1996) Dihydrolipoyl acetyltransferase, the core protein of the pyruvate dehydrogenase multienzyme complex from *Azotobacter vinelandii*. In: *Perspectives on protein engineering 5* (Geisow, M., eds.), CD-ROM, Biodigm Ltd., Nottingham, UK.
- De Kok, A. & van Berkel, W. J. H. (1996) Lipoamide dehydrogenase. In: *Alpha-keto acid dehydrogenase complexes* (Patel, M. S., et al., eds.), pp. 53-70, Birkhäuser Verlag, Basel.
- De Marcucci, O. & Lindsay, J. G. (1985) Component X. An immunologically distinct polypeptide associated with mammalian pyruvate dehydrogenase multi-enzyme complex. *Eur. J. Biochem.* **149**, 641-648.
- DeRosier, D. J., Oliver, R. M. & Reed, L. J. (1971) Crystallization and preliminary structural analysis of dihydrolipoyl transsuccinylase, the core of the 2-oxoglutarate dehydrogenase complex. *Proc. Natl. Acad. Sci. USA* **68**, 1135-1137.
- Douce, R., Bourguignon, J., Macherel, D. & Neuburger, M. (1994) The glycine decarboxylase system in higher plant mitochondria: structure, function and biogenesis. *Biochem. Soc. Trans.* **22**, 184-188.
- Duckworth, H. W., Jaenicke, R., Perham, R. N., Wilkie, A. O. M., Finch, J. T. & Roberts, G. C. K. (1982) Limited proteolysis and proton NMR spectroscopy of *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *Eur. J. Biochem.* **124**, 63-69.
- Dyda, F., Furey, W., Subramanyam, S., Sax, M., Farrenkopf, B. & Jordan, F. (1993) Catalytic centers in the thiamin diphosphate dependent enzyme pyruvate decarboxylase at 2.4-Å resolution. *Biochemistry* **32**, 6165-6170.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. & Venter, J. C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.
- Fujiwara, K., Okamura-Ikeda, K. & Motokawa, Y. (1992) Expression of mature bovine H-protein of the glycine cleavage system in *Escherichia coli* and *in vitro* lipoylation of the apoforn. *J. Biol. Chem.* **267**, 20011-20016.
- Fujiwara, K., Okamura-Ikeda, K. & Motokawa, Y. (1994) Purification and characterization of lipoyl-AMP:N₅-lysine lipoyltransferase from bovine liver mitochondria. *J. Biol. Chem.* **269**, 16605-16609.
- Fuller, C. C., Reed, L. J., Oliver, R. M. & Hackert, M. L. (1979) Crystallization of a dihydrolipoyl transacetylase-dihydrolipoyl dehydrogenase subcomplex and its implications regarding the subunit structure of the pyruvate dehydrogenase complex from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **90**, 431-438.
- 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) The pyruvate dehydrogenase multi-enzyme complex of *Escherichia coli*: genetic reconstruction and functional analysis of the lipoyl domains. *Phil. Trans. R. Soc. Lond. A* **317**, 391-404.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.

- Graham, L. D. & Perham, R. N. (1990) Interactions of lipoyl domains with the E1p subunits of the pyruvate dehydrogenase multienzyme complex from *Escherichia coli*. *FEBS Lett.* **262**, 241-244.
- Green, J. D. F., Perham, R. N., Ullrich, S. J. & Appella, E. (1992) Conformational studies of the interdomain linker peptides in the dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Biol. Chem.* **267**, 23484-23488.
- Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995a) Three-dimensional structure of a lipoyl domain from dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.
- Green, D. E., Morris, T. W., Green, J., Cronan, J. E. & Guest, J. R. (1995b) Purification and properties of the lipoate protein ligase of *Escherichia coli*. *Biochem. J.* **309**, 853-862.
- Griffin, T. A., Lau, K. S. & Chuang, D. T. (1988) Characterization and conservation of the inner E2 core domain structure of branched-chain α -keto acid dehydrogenase complex from bovine liver. Construction of a cDNA encoding the entire transacylase (E2b) precursor. *J. Biol. Chem.* **263**, 14008-14014.
- Griffin, T. A. & Chuang, D. T. (1990) Genetic reconstruction and characterization of the recombinant transacylase (E2b) component of bovine branched-chain α -keto acid dehydrogenase complex. Implication of histidine 391 as an active site residue. *J. Biol. Chem.* **265**, 13174-13180.
- Griffin, T. A., Wynn, R. M. & Chuang, D. T. (1990) Expression and assembly of mature apotransacylase (E2b) of bovine branched-chain α -keto acid dehydrogenase complex in *Escherichia coli*. *J. Biol. Chem.* **265**, 12104-12110.
- Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C. & Perham, R. N. (1985) Genetic reconstruction and functional analysis of the repeating lipoyl domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **185**, 743-754.
- Guest, J. R. (1987) Functional implications of structural homologies between chloramphenicol acetyltransferase and dihydrolipoamide acetyltransferase. *FEMS Microbiol. Lett.* **44**, 417-422.
- Guest, J. R., Angier, S. J. & Russel, G. C. (1989) Structure, expression, and protein engineering of the pyruvate dehydrogenase complex of *Escherichia coli*. *Ann. N.Y. Acad. Sci.* **573**, 76-99.
- Hackert, M. L., Oliver, R. M. & Reed, L. J. (1983a) Evidence for a multiple random coupling mechanism in the α -ketoglutarate dehydrogenase multienzyme complex of *Escherichia coli*: a computer model analysis. *Proc. Natl. Acad. Sci. USA* **80**, 2226-2230.
- Hackert, M. L., Oliver, R. M. & Reed, L. J. (1983b) A computer model analysis of the active-site coupling mechanism in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **80**, 2907-2911.
- Hale, G., Wallis, N. G. & Perham, R. N. (1992) Interaction of avidin with the lipoyl domains in the pyruvate dehydrogenase multienzyme complex: three-dimensional location and similarity to biotinyl domains in carboxylases. *Proc. R. Soc. Lond. B* **248**, 247-253.
- Hämälä, H., Palva, A., Paulin, L., Arvidson, S. & Palva, I. (1990) Secretory S complex of *Bacillus subtilis* and identity to pyruvate dehydrogenase. *J. Bacteriol.* **172**, 5052-5063.
- Hammes, G. G. (1981) Processing of intermediates in multienzyme complexes. *Biochem. Soc. Symp.* **46**, 73-90.
- Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987) The domain structure of the dihydrolipoyl transacylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **169**, 245-252.
- Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988a) The dihydrolipoyltransacylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Molecular cloning and sequence analysis. *Eur. J. Biochem.* **172**, 593-599.
- Hanemaaijer, R., Vervoort, J., Westphal, A. H., de Kok, A. & Veeger, C. (1988b) Mobile sequences in the pyruvate dehydrogenase complex, the E2 component, the catalytic domain and the 2-oxoglutarate dehydrogenase complex of *Azotobacter vinelandii*, as detected by 600 MHz proton NMR spectroscopy. *FEBS Lett.* **240**, 205-210.
- Hanemaaijer, R., Westphal, A. H., Van der Heiden, T., De Kok, A. & Veeger, C. (1989) The quaternary structure of the dihydrolipoyl transacylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. A reconsideration. *Eur. J. Biochem.* **179**, 287-292.
- Hawkins, C. F., Borges, A. & Perham, R. N. (1989) A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS Lett.* **255**, 77-82.
- Hein, S. & Steinbüchel, A. (1994) Biochemical and molecular characterization of the *Alcaligenes eutrophus* pyruvate dehydrogenase complex and identification of a new type of dihydrolipoamide dehydrogenase. *J. Bacteriol.* **176**, 4394-4408.

- Henderson, C. E., Perham, R. N. & Finch, J. T. (1979) Structure and symmetry of *B. stearothermophilus* pyruvate dehydrogenase multienzyme complex and implications for eukaryote evolution. *Cell* **17**, 85-93.
- Hendle, J., Mattevi, A., Westphal, A. H., Spee, J., de Kok, A., Teplyakov, A. & Hol, W. G. J. (1995) Crystallographic and enzymatic investigations on the role of Ser558, His610, and Asn614 in the catalytic mechanism of *Azotobacter vinelandii* dihydrolipoamide acetyltransferase (E2p). *Biochemistry* **34**, 4287-4298.
- Hester, K., Luo, J., Burns, G., Braswell, E. H. & Sokatch, J. R. (1995) Purification of active E1 $\alpha(2)\beta(2)$ of *Pseudomonas putida* branched-chain 2-oxo acid dehydrogenase complex. *Eur. J. Biochem.* **233**, 828-836.
- Hipps, D. S. & Perham, R. N. (1992) Expression in *Escherichia coli* of a sub-gene encoding the lipoyl and peripheral subunit-binding domains of the dihydrolipoamide acetyltransferase component of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *Biochem. J.* **283**, 665-671.
- Hipps, D. S., Packman, L. C., Allen, M. D., Fuller, C., Sakaguchi, K., Appella, E. & Perham, R. N. (1994) The peripheral subunit-binding domain of the dihydrolipoamide acetyltransferase component of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*: preparation and characterization of its binding to the dihydrolipoamide dehydrogenase component. *Biochem. J.* **297**, 137-143.
- Jilka, J. M., Rahmatullah, M., Kazemi, M. & Roche, T. E. (1986) Properties of a newly characterised protein of the bovine kidney pyruvate dehydrogenase complex. *J. Biol. Chem.* **261**, 1858-1867.
- Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K. & Perham, R. N. (1993) The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *J. Mol. Biol.* **230**, 323-341.
- Koike, M., Reed, L. J. & Carroll, W. R. (1963) α -Keto acid dehydrogenase complexes. IV. Resolution and reconstitution of the *Escherichia coli* pyruvate dehydrogenase complex. *J. Biol. Chem.* **238**, 30-39.
- Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
- Krüger, N., Opperman, F. B., Lorenzl, H. & Steinbüchel, A. (1994) Biochemical and molecular characterization of the *Clostridium magnum* acetoin dehydrogenase enzyme system. *J. Bacteriol.* **176**, 3614-3630.
- Lawson, J. E., Behal, R. H. & Reed, L. J. (1991a) Disruption and mutagenesis of the *Saccharomyces cerevisiae* PDX1 gene encoding the protein X component of the pyruvate dehydrogenase complex. *Biochemistry* **30**, 2834-2839.
- Lawson, J. E., Niu, X. & Reed, L. J. (1991b) Functional analysis of the domains of dihydrolipoamide acetyltransferase from *Saccharomyces cerevisiae*. *Biochemistry* **30**, 11249-11254.
- Lessard, I. A. D. & Perham, R. N. (1994) Expression in *Escherichia coli* of genes encoding the E1 α and E1 β subunits of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus* and assembly of a functional E1 component ($\alpha(2)\beta(2)$) *in vitro*. *J. Biol. Chem.* **269**, 10378-10383.
- Lessard, I. A. D. & Perham, R. N. (1995) Interaction of component enzymes with the peripheral subunit-binding domain of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*: stoichiometry and specificity in self-assembly. *Biochem. J.* **306**, 727-733.
- Lindqvist, Y., Schneider, G., Ermiler, U. & Sundström, M. (1992) Three-dimensional structure of transketolase, a thiamin diphosphate dependent enzyme, at 2.5 Å resolution. *EMBO J.* **7**, 2373-2379.
- Lindqvist, Y. & Schneider, G. (1993) Thiamin diphosphate dependent enzymes: transketolase, pyruvate oxidase and pyruvate decarboxylase. *Curr. Opin. Struct. Biol.* **3**, 896-901.
- Linn, T. C., Pettit, F. H. & Reed, L. J. (1969) α -keto acid dehydrogenase complexes: X. regulation of the activity of pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation. *Proc. Natl. Acad. Sci. USA* **62**, 234-241.
- Liu, S., Baker, J. C. & Roche, T. E. (1995a) Binding of the pyruvate dehydrogenase kinase to recombinant constructs containing the inner lipoyl domain of the dihydrolipoamide acetyltransferase component. *J. Biol. Chem.* **270**, 793-800.
- Liu, S., Baker, J. C., Andrews, P. C. & Roche, T. E. (1995b) Recombinant expression and evaluation of the lipoyl domains of the dihydrolipoamide acetyltransferase component of the human pyruvate dehydrogenase complex. *Arch. Biochem. Biophys.* **316**, 926-940.
- Macherel, D., Bourguignon, J., Forest, E., Faure, M., Cohen-Addad, C. & Douce, R. (1996) Expression, lipoylation and structure determination of recombinant pea H-protein in *Escherichia coli*. *Eur. J. Biochem.* **236**, 27-33.

- Maeng, C.-Y., Yazdi, M. A., Niu, X.-D., Lee, H. Y. & Reed, L. J. (1994) Expression, purification, and characterization of the dihydrolipoamide dehydrogenase-binding protein of the pyruvate dehydrogenase complex from *Saccharomyces cerevisiae*. *Biochemistry* **33**, 13801-13807.
- Mande, S. S., Sarfaty, S., Allen, M. D., Perham, R. N. & Hol, W. G. J. (1996) Protein-protein interactions in the pyruvate dehydrogenase multienzyme complex: dihydrolipoamide dehydrogenase complexed with the binding domain of dihydrolipoamide acetyltransferase. *Structure* **4**, 277-286.
- Massey, V., Gibson, Q.H. & Veeger, C. (1960) Intermediates in the catalytic action of lipoyl dehydrogenase. *Biochem. J.* **77**, 341-351.
- Mattevi, A., Schierbeek, A. J. & Hol, W. G. J. (1991) Refined crystal structure of lipoyl dehydrogenase from *Azotobacter vinelandii* at 2.2 Å resolution. A comparison with the structure of glutathione reductase. *J. Mol. Biol.* **220**, 975-994.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992a) The pyruvate dehydrogenase multienzyme complex. *Curr. Opin. Struct. Biol.* **2**, 877-887.
- Mattevi, A., Obmolova, G., Kalk, K. H., Sokatch, J., Betzel, C. H. & Hol, W. G. J. (1992b) The refined crystal structure of *Pseudomonas putida* lipoyl dehydrogenase complexed with NAD⁺ at 2.45 Å resolution. *Proteins* **13**, 336-351.
- Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1992c) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
- Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. H. & Hol, W. G. J. (1993a) 3-Dimensional structure of lipoyl dehydrogenase from *Pseudomonas fluorescens* at 2.8 Å resolution - analysis of redox and thermostability properties. *J. Mol. Biol.* **230**, 1200-1215.
- Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1993b) Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Å resolution. *J. Mol. Biol.* **230**, 1183-1199.
- Mattevi, A., Obmolova, G., Kalk, K. H., Teplyakov, A. & Hol, W. G. J. (1993c) Crystallographic analysis of substrate binding and catalysis in dihydrolipoyl transacetylase (E2p). *Biochemistry* **32**, 3887-3901.
- Matuda, S., Nakano, K., Ohta, S., Shimura, M., Yamanaka, T., Nakagawa, S., Titani, K. & Miyata, T. (1992) Molecular cloning of dihydrolipoamide acetyltransferase of the rat pyruvate dehydrogenase complex: sequence comparison and evolutionary relationship to other dihydrolipoamide acyltransferases. *Biochim. Biophys. Acta* **1131**, 114-118.
- Meng, M. & Chuang, D. T. (1994) Site-directed mutagenesis and functional analysis of the active-site residues of the E2 component of bovine branched-chain α -keto acid dehydrogenase complex. *Biochemistry* **33**, 12879-12885.
- Miles, J. S. & Guest, J. R. (1987a) Molecular genetic aspects of the citric acid cycle of *Escherichia coli*. *Biochem. Soc. Symp.* **54**, 45-65.
- Miles, J. S. & Guest, J. R. (1987b) Subgenes expressing single lipoyl domains of the pyruvate dehydrogenase complex of *Escherichia coli*. *Biochem. J.* **245**, 869-874.
- Miles, J. S., Guest, J. R., Radford, S. E. & Perham, R. N. (1988) Investigation of the mechanism of active site coupling in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli* by protein engineering. *J. Mol. Biol.* **202**, 97-106.
- Morris, T. W., Reed, K. E. & Cronan, J. E. (1994) Identification of the gene encoding lipoyl-protein ligase A of *Escherichia coli*. Molecular cloning and characterization of the *lplA* gene and gene product. *J. Biol. Chem.* **269**, 16091-16100.
- Morris, T. W., Reed, K. E. & Cronan, J. E. (1995) Lipoyl acid metabolism in *Escherichia coli*: the *lplA* and *lipB* genes define redundant pathways for ligation of lipoyl groups to apoprotein. *J. Bacteriol.* **177**, 1-10.
- Muller, Y. A. & Schulz, G. E. (1993) Structure of the thiamine- and flavin-dependent enzyme pyruvate oxidase. *Science* **259**, 965-967.
- Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F. & Schneider, G. (1993) A thiamin diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase. *Structure* **1**, 95-103.
- Nakano, K., Matuda, S., Yamanaka, T., Tsubouchi, H., Nakagawa, S., Titani, K., Ohta, S. & Miyata, T. (1991) Purification and molecular cloning of succinyltransferase of the rat α -ketoglutarate dehydrogenase complex. Absence of a sequence motif of the putative E3 and/or E1 binding site. *J. Biol. Chem.* **266**, 19013-19017.
- Nakano, K., Takase, C., Sakamoto, T., Nakagawa, S., Inazawa, J., Ohta, S. & Matuda, S. (1994) Isolation, characterization and structural organization of the gene and pseudogene for the dihydrolipoamide

- succinyltransferase component of the human 2-oxoglutarate dehydrogenase complex. *Eur. J. Biochem.* **224**, 179-189.
- Nawa, H., Brady, W. T., Koike, M. & Reed, L. J. (1960) Studies on the nature of protein-bound lipoic acid. *J. Am. Chem. Soc.* **82**, 896-903.
- Neagle, J. C. & Lindsay, J. G. (1991) Selective proteolysis of the protein X subunit of the bovine heart pyruvate dehydrogenase complex. Effects on dihydrolipoamide dehydrogenase (E3) affinity and enzymic properties of the complex. *Biochem. J.* **278**, 423-427.
- Neuberger, M., Jourdain, A. & Douce, R. (1991) Isolation of H-protein loaded with methylamine as a transient species in glycine decarboxylase reactions. *Biochem. J.* **278**, 765-769.
- Niu, X., Browning, K. S., Behal, R. H. & Reed, L. J. (1988) Cloning and nucleotide sequence of the gene for dihydrolipoamide acetyltransferase from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **85**, 7546-7550.
- Niu, X., Stoops, J. K. & Reed, L. J. (1990) Overexpression and mutagenesis of the catalytic domain of dihydrolipoamide acetyltransferase from *Saccharomyces cerevisiae*. *Biochemistry* **29**, 8614-8619.
- Oliver, R. M. & Reed, L. J. (1982) Multienzyme complexes. In: *Electron microscopy of proteins* (Harris, J. R., eds.), pp. 1-48, Academic Press, London.
- Ono, K., Radke, G. A., Roche, T. E. & Rahmatullah, M. (1993) Partial activation of the pyruvate dehydrogenase kinase by the lipoyl domain region of E2 and interchange of the kinase between lipoyl domain regions. *J. Biol. Chem.* **268**, 26135-26143.
- Opperman, F. B., Steinbüchel, A. & Schlegel, H. G. (1989) Evidence for oxidative thiolytic cleavage of acetoin in *Pelobacter carbinolicus* analogous to aerobic oxidative decarboxylation of pyruvate. *FEMS Microbiol. Lett.* **60**, 113-118.
- Opperman, F. B. & Steinbüchel, A. (1994) Identification and molecular characterization of the *aco* genes encoding the *Pelobacter carbinolicus* acetoin dehydrogenase enzyme system. *J. Bacteriol.* **176**, 469-485.
- Packman, L. C., Hale, G. & Perham, R. N. (1984a) Repeating functional domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *EMBO J.* **3**, 1315-1319.
- Packman, L. C., Perham, R. N. & Roberts, G. C. K. (1984b) Domain structure and ¹H-n.m.r. spectroscopy of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *Biochem. J.* **217**, 219-227.
- Packman, L. C. & Perham, R. N. (1986) Chain folding in the dihydrolipoamide acetyltransferase components of the 2-oxo-acid dehydrogenase complexes from *Escherichia coli*. Identification of a segment involved in binding the E3 subunit. *FEBS Lett.* **206**, 193-198.
- Packman, L. C. & Perham, R. N. (1987) Limited proteolysis and sequence analysis of the 2-oxo acid dehydrogenase complexes from *Escherichia coli*. Cleavage sites and domains in the dihydrolipoamide acetyltransferase components. *Biochem. J.* **242**, 531-538.
- Packman, L. C., Borges, A. & Perham, R. N. (1988) Amino acid sequence analysis of the lipoyl and peripheral subunit-binding domains in the lipoate acetyltransferase component of the pyruvate dehydrogenase complex from *Bacillus stearothermophilus*. *Biochem. J.* **252**, 79-86.
- Packman, L. C., Green, B. & Perham, R. N. (1991) Lipoylation of the E2 components of the 2-oxo acid dehydrogenase multienzyme complexes of *Escherichia coli*. *Biochem. J.* **277**, 153-158.
- Palmer, J. A., Madhusudhan, K. T., Hatter, K. & Sokatch, J. R. (1991) Cloning, sequence and transcriptional analysis of the structural gene for LPD-3, the third lipoamide dehydrogenase of *Pseudomonas putida*. *Eur. J. Biochem.* **202**, 231-240.
- Pares, S., Cohen-Addad, C., Sieker, L., Neuberger, M. & Douce, R. (1994) X-ray structure determination at 2.6-Å resolution of a lipoate-containing protein: The H-protein of the glycine decarboxylase complex from pea leaves. *Proc. Natl. Acad. Sci. USA* **91**, 4850-4853.
- Patel, M. S. & Roche, T. E. (1990) Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *FASEB J.* **4**, 3224-3233.
- Patel, M. S., Kerr, D. S. & Wexler, I. D. (1992) Biochemical and molecular aspects of pyruvate dehydrogenase complex deficiency. *Int. Pediatr.* **7**, 16-22.
- Patel, M. S. & Harris, R. A. (1995) Mammalian α -keto acid dehydrogenase complexes: gene regulation and genetic defects. *FASEB J.* **9**, 1164-1172.
- Patel, M. S., Roche, T. E. & Harris, R. A., eds. (1996) Alpha-keto acid dehydrogenase complexes. *Molecular and Cell Biology Updates* pp. 1-321, Birkhäuser Verlag, Basel.
- Perham, R. N. (1975) Self-assembly of biological macromolecules. *Phil. Trans. R. Soc. Lond., Ser. B* **272**, 123-136.
- Perham, R. N. & Roberts, G. C. K. (1981) Limited proteolysis and proton n.m.r. spectroscopy of the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochem. J.* **199**, 733-740.

- Perham, R. N., Duckworth, H. W. & Roberts, G. C. K. (1981) Mobility of polypeptide chain in the pyruvate dehydrogenase complex revealed by proton NMR. *Nature* **292**, 474-477.
- Perham, R. N., Packman, L. C. & Radford, S. E. (1988) 2-Oxo acid dehydrogenase multi-enzyme complexes: in the beginning and halfway there. *Biochem. Soc. Symp.* **54**, 67-81.
- Perham, R. N. & Packman, L. C. (1989) 2-Oxo acid dehydrogenase multienzyme complexes: domains, dynamics and design. *Ann. N.Y. Acad. Sci.* **573**, 1-20.
- Perham, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.
- Pettit, F. H., Hamilton, L., Munk, P., Namihira, G., Eley, M. H., Willms, C. R. & Reed, L. J. (1973) α -Keto acid dehydrogenase complexes: XIX. subunit structure of the *Escherichia coli* α -ketoglutarate dehydrogenase complex. *J. Biol. Chem.* **248**, 5282-5290.
- Pfiefert, H., Hein, S., Krüger, N., Zeh, K., Schmidt, B. & Steinbüchel, A. (1991) Identification and molecular characterization of the *Alcaligenes eutrophus* H16 *aco* operon genes involved in acetoin metabolism. *J. Bacteriol.* **173**, 4056-4071.
- Quinn, J., Diamond, A. G., Masters, A. K., Brookfield, D. E., Wallis, N. G. & Yeaman, S. J. (1993) Expression and lipoylation in *Escherichia coli* of the inner lipoyl domain of the E2-component of the human pyruvate dehydrogenase complex. *Biochem. J.* **289**, 81-85.
- Radford, S. E., Laue, E. D. & Perham, R. N. (1986) Nuclear-magnetic-resonance spectroscopy studies of the flexible linkages between lipoyl domains in the pyruvate dehydrogenase multienzyme complex from *Escherichia coli*. *Biochem. Soc. Trans.* **14**, 1231-1232.
- Radford, S. E., Laue, E. D., Perham, R. N., Miles, J. S. & Guest, J. R. (1987) Segmental structure and protein domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. Genetic reconstruction *in vitro* and ^1H -n.m.r. spectroscopy. *Biochem. J.* **247**, 641-649.
- Radford, S. E., Laue, E. D., Perham, R. N., Martin, S. R. & Appella, E. (1989) Conformational flexibility and folding of synthetic peptides representing an interdomain segment of polypeptide chain in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Biol. Chem.* **264**, 767-775.
- Radke, G. A., Ono, K., Ravindran, S. & Roche, T. E. (1993) Critical role of a lipoyl cofactor of the dihydrolipoyl acetyltransferase in the binding and enhanced function of the pyruvate dehydrogenase kinase. *Biochem. Biophys. Res. Commun.* **190**, 982-991.
- Rahmatullah, M., Gopalakrishnan, S., Andrews, P. C., Chang, C. L., Radke, G. A. & Roche, T. E. (1989a) Subunit associations in the mammalian pyruvate dehydrogenase complex. Structure and role of protein X and the pyruvate dehydrogenase component binding domain of the dihydrolipoyl transacetylase component. *J. Biol. Chem.* **264**, 2221-2227.
- Rahmatullah, M., Gopalakrishnan, S., Radke, G. A. & Roche, T. E. (1989b) Domain structure of the dihydrolipoyl transacetylase and the protein X components of mammalian pyruvate dehydrogenase complex. Selective cleavage by protease Arg C. *J. Biol. Chem.* **264**, 1245-1251.
- Rahmatullah, M., Radke, G. A., Andrews, P. C. & Roche, T. E. (1990) Changes in the core of the mammalian pyruvate dehydrogenase complex upon selective removal of the lipoyl domain from the transacetylase component but not from the protein X component. *J. Biol. Chem.* **265**, 14512-14517.
- Ravindran, S., Radke, G. A., Guest, J. R. & Roche, T. E. (1996) Lipoyl domain-based mechanism for the integrated feedback control of the pyruvate dehydrogenase complex by enhancement of pyruvate dehydrogenase kinase activity. *J. Biol. Chem.* **271**, 653-662.
- Reed, L. J., Koike, M., Levitch, M. E. & Leach, F. R. (1958a) Studies on the nature and reactions of protein-bound lipoic acid. *J. Biol. Chem.* **232**, 143-158.
- Reed, L. J., Leach, F. R. & Koike, M. (1958b) Studies on a lipoic acid-activating system. *J. Biol. Chem.* **232**, 123-142.
- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.
- Reed, L. J., Pettit, F. H., Eley, M. H., Hamilton, L., Collins, J. H. & Oliver, R. M. (1975) Reconstitution of the *Escherichia coli* pyruvate dehydrogenase complex. *Proc. Natl. Acad. Sci. USA* **72**, 3086-3072.
- Reed, L. J. & Hackert, M. L. (1990) Structure-function relationships in dihydrolipoamide acyltransferases. *J. Biol. Chem.* **265**, 8971-8974.
- Repetto, B. & Tzagoloff, A. (1990) Structure and regulation of *KGD2*, the structural gene for yeast dihydrolipoyl transsuccinylase. *Mol. Cell. Biol.* **10**, 4221-4232.
- Rice, J. E., Dunbar, B. & Lindsay, J. G. (1992) Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *EMBO J.* **11**, 3229-3235.
- Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K. & Gronenborn, A. M. (1992) Three-dimensional solution structure of the E3-binding domain of the dihydrolipoamide

- succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **31**, 3463-3471.
- Roche, T. E. & Patel, M. S., eds. (1989) Alpha-keto acid dehydrogenase complexes: organization, regulation, and biomedical ramifications. *Ann. N.Y. Acad. Sci.* **573**, pp. 1-474.
- Russell, G. C. & Guest, J. R. (1990) Overexpression of restructured pyruvate dehydrogenase complexes and site-directed mutagenesis of a potential active-site histidine residue. *Biochem. J.* **269**, 443-450.
- Russell, G. C. & Guest, J. R. (1991) Sequence similarities within the family of dihydrolipoamide acyltransferases and discovery of a previously unidentified fungal enzyme. *Biochim. Biophys. Acta* **1076**, 225-232.
- Russell, G. C., Machado, R. S. & Guest, J. R. (1992) Overproduction of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli* and site-directed substitutions in the E1p and E2p subunits. *Biochem. J.* **287**, 611-619.
- Sanderson, S. J., Miller, C. & Lindsay, J. G. (1996) Stoichiometry, organisation and catalytic function of protein X of the pyruvate dehydrogenase complex from bovine heart. *Eur. J. Biochem.* **236**, 67-77.
- Schierbeek, A. J., Swarte, M. B. A., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G. J., Drenth, J. & Betzel, C. (1989) X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *J. Mol. Biol.* **206**, 365-379.
- Schulze, E., Westphal, A. H., Obmolova, G., Mattevi, A., Hol, W. G. J. & de Kok, A. (1991a) The catalytic domain of the dihydrolipoamide acetyltransferase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* and *Escherichia coli*. Expression, purification, properties and preliminary X-ray analysis. *Eur. J. Biochem.* **201**, 561-568.
- Schulze, E., Benen, J. A. E., Westphal, A. H. & de Kok, A. (1991b) Interaction of lipoamide dehydrogenase with the dihydrolipoamide acetyltransferase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **200**, 29-34.
- Schulze, E., Westphal, A. H., Boumans, H. & de Kok, A. (1991c) Site-directed mutagenesis of the dihydrolipoamide acetyltransferase component (E2p) of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Binding of the peripheral components E1p and E3. *Eur. J. Biochem.* **202**, 841-848.
- Schulze, E., Westphal, A. H., Veeger, C. & de Kok, A. (1992) Reconstitution of pyruvate dehydrogenase multienzyme complexes based on chimeric core structures from *Azotobacter vinelandii* and *Escherichia coli*. *Eur. J. Biochem.* **206**, 427-435.
- Schulze, E., Westphal, A. H., Hanemaaijer, R. & de Kok, A. (1993) Structure/function relationships in the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Role of the linker region between the binding and catalytic domain of the dihydrolipoamide acetyltransferase component. *Eur. J. Biochem.* **211**, 591-599.
- Shepherd, G. B. & Hammes, G. G. (1977) Fluorescence energy transfer measurements in the pyruvate dehydrogenase multienzyme complex from *Escherichia coli* with chemically modified lipoic acid. *Biochemistry* **16**, 5234-5241.
- Spencer, M. E., Darlison, M. G., Stephens, P. E., Duckenfield, I. K. & Guest, J. R. (1984) Nucleotide sequence of the *sucB* gene encoding the dihydrolipoamide succinyltransferase of *Escherichia coli* K12 and homology with the corresponding acetyltransferase. *Eur. J. Biochem.* **141**, 361-374.
- Stephens, P. E., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983a) The pyruvate dehydrogenase complex of *Escherichia coli* K12. Nucleotide sequence of the pyruvate dehydrogenase component. *Eur. J. Biochem.* **133**, 155-162.
- Stephens, P. E., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983b) The pyruvate dehydrogenase complex of *Escherichia coli* K12. Nucleotide sequence encoding the dihydrolipoamide acetyltransferase component. *Eur. J. Biochem.* **133**, 481-489.
- Stepp, L. R., Bleile, D. M., McRorie, D. K., Pettit, F. H. & Reed, L. J. (1981) Use of trypsin and lipoamidase to study the role of lipoic acid moieties in the pyruvate and α -ketoglutarate dehydrogenase complexes of *Escherichia coli*. *Biochemistry* **20**, 4555-4560.
- Texter, F. L., Radford, S. E., Laue, E. D., Perham, R. N., Miles, J. S. & Guest, J. R. (1988) Site-directed mutagenesis and ^1H NMR spectroscopy of an interdomain segment in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **27**, 289-296.
- Thekkumkara, T. J., Ho, L., Wexler, I. D., Pons, G., Liu, T. C. & Patel, M. S. (1988) Nucleotide sequence of a cDNA for dihydrolipoamide acetyltransferase component of human pyruvate dehydrogenase complex. *FEBS Lett.* **240**, 45-48.
- Toh, H., Kondo, H. & Tanabe, T. (1993) Molecular evolution of biotin-dependent carboxylases. *Eur. J. Biochem.* **215**, 687-696.

- Turner, S. L., Russell, G. C., Williamson, M. P. & Guest, J. R. (1993) Restructuring an interdomain linker in the dihydrolipoamide acetyltransferase component of the pyruvate dehydrogenase complex of *Escherichia coli*. *Protein Eng.* **6**, 101-108.
- Wagenknecht, T., Grassucci, R. & Schaak, D. (1990) Cryoelectron microscopy of frozen-hydrated α -ketoacid dehydrogenase complexes from *Escherichia coli*. *J. Biol. Chem.* **265**, 22402-22408.
- Wagenknecht, T., Grassucci, R., Radke, G. A. & Roche, T. E. (1991) Cryoelectron microscopy of mammalian pyruvate dehydrogenase complex. *J. Biol. Chem.* **266**, 24650-24656.
- Wagenknecht, T., Grassucci, R., Berkowitz, J. & Forneris, C. (1992) Configuration of interdomain linkers in pyruvate dehydrogenase complex of *Escherichia coli* as determined by cryoelectron microscopy. *J. Struct. Biol.* **109**, 70-77.
- Waldrop, G. L., Rayment, I. & Holden, H. M. (1994) Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry* **33**, 10249-10256.
- Wallbrandt, P., Tegman, V., Jonsson, B. H. & Wieslander, A. (1992) Identification and analysis of the genes coding for the putative pyruvate dehydrogenase enzyme complex in *Acholeplasma laidlawii*. *J. Bacteriol.* **174**, 1388-1396.
- Wallis, N. G. & Perham, R. N. (1994) Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**, 209-216.
- Wang, G. F., Kuriki, T., Roy, K. L. & Kaneda, T. (1993) The primary structure of branched-chain α -oxo acid dehydrogenase from *Bacillus subtilis* and its similarity to other α -oxo acid dehydrogenases. *Eur. J. Biochem.* **213**, 1091-1099.
- Westphal, A. H. & de Kok, A. (1990) The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur. J. Biochem.* **187**, 235-239.
- Westphal, A. H., Fabisz-Kijowska, A., Kester, H., Obels, P. P. & de Kok, A. (1995) The interaction between lipoamide dehydrogenase and the peripheral-component-binding domain from the *Azotobacter vinelandii* pyruvate dehydrogenase complex. *Eur. J. Biochem.* **234**, 861-870.
- Williams, C. H. J. (1992) Lipoamide dehydrogenase, glutathione reductase, thioredoxin reductase and mercuric reductase-family of flavoenzyme transhydrogenases. In: *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., eds.), pp. 121-211, CRC Press, Boca Raton.
- Wu, T. L. & Reed, L. J. (1984) Subunit binding in the pyruvate dehydrogenase complex from bovine kidney and heart. *Biochemistry* **23**, 221-226.
- Wynn, R. M., Chuang, J. L., Davie, J. R., Fisher, C. W., Hale, M. A., Cox, R. P. & Chuang, D. T. (1992) Cloning and expression in *Escherichia coli* of mature E1 β subunit of bovine mitochondrial branched-chain α -keto acid dehydrogenase complex. Mapping of the E1 β -binding region on E2. *J. Biol. Chem.* **267**, 1881-1887.
- Yang, Y.-S. & Frey, P. A. (1986) Dihydrolipoyl transacetylase of *Escherichia coli*. Formation of 8-S-acetyldihydrolipoamide. *Biochemistry* **25**, 8173-8178.
- Yang, H., Frey, P. A., Hainfeld, J. F. & Wall, J. S. (1986) Pyruvate dehydrogenase complex of *Escherichia coli*. Radial mass analysis of subcomplexes by scanning transmission electron microscopy. *Biophys. J.* **49**, 56-58.
- Yang, Y.-S., Datta, A., Hainfeld, J. F., Furuya, F. R., Wall, J. S. & Frey, P. A. (1994) Mapping the lipoyl groups of the pyruvate dehydrogenase complex by use of gold cluster labels and scanning transmission electron microscopy. *Biochemistry* **33**, 9428-9437.
- Yeaman, S. J. (1989) The 2-oxo acid dehydrogenase complexes: recent advances. *Biochem. J.* **257**, 625-632.

CHAPTER 2

Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*

ABSTRACT

The N-terminal lipoyl domain (79 residues) of the transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* has been sub-cloned and produced in *Escherichia coli*. Over-expression exceeds the capacity of *E. coli* cells to lipoylate all expressed lipoyl domain, but addition of lipoic acid to the growth medium results in expression of fully lipoylated domain. A two-dimensional homo- and heteronuclear NMR study of the lipoyl domain has resulted in sequential ^1H and ^{15}N resonance assignments of the unlipoylated form of the protein. Small differences in chemical shift values for protons of residues in the vicinity of the lipoyl-lysine residue are observed for the lipoylated form of the domain, suggesting that the conformation of the lipoyl domain is not altered significantly by the coupled prosthetic group. From nuclear Overhauser effects, backbone coupling constants, and slowly exchanging amide protons, two antiparallel β -sheets, each containing four strands, were identified. The lipoyl-lysine residue is exposed to the solvent and located in a type-I turn between two strands. The N- and C-terminal residues of the folded chain are close together in the other sheet. Preliminary data on the relative three-dimensional orientation of the two β -sheets are presented. Comparison with the solution structure of the lipoyl domain of the *Bacillus stearothermophilus* pyruvate dehydrogenase complex shows resemblance to a large extent, despite the sequence identity of 31%.

INTRODUCTION

The pyruvate dehydrogenase complex (PDHC) from *Azotobacter vinelandii* catalyses the oxidative decarboxylation of pyruvate to acetyl-CoA [for recent reviews on PDHC, see Perham (1991) and Mattevi *et al.* (1992a)]. The complex is composed of multiple copies of three enzymes: pyruvate dehydrogenase (E1p), dihydrolipoyl transacetylase or acetyltransferase (E2p) and dihydrolipoyl dehydrogenase or lipoamide dehydrogenase (E3). The central core of the complex is formed by a trimer of E2p (Mattevi *et al.*, 1992b), to which two dimers of E1p and one dimer of E3 are tightly but non-covalently bound (Schulze *et al.*, 1992). The E2p monomer is a highly segmented protein in which five separate and independently folded domains can be recognised, connected by mobile linker sequences rich in alanine and proline residues (Hanemaaijer *et al.*, 1988). The N-terminal part contains three highly homologous lipoyl domains, each bearing a lipoyl group covalently bound to a specific lysine residue. Between the lipoyl domains and the C-terminal catalytic domain, which contains the acetyltransferase catalytic site and aggregates to form the core structure of the complex (Hanemaaijer *et al.*, 1987; Mattevi *et al.*, 1992b), the peripheral subunit-binding domain is situated. This domain is responsible for binding of the E3 and the E1p component to the acetyltransferase core. The E1p component also interacts with the catalytic domain (Schulze *et al.*, 1991a, 1992).

The lipoyl-lysine residues of lipoyl domains are central to the activity of the complex, providing swinging arms that are highly mobile and responsible for substrate channelling among the three successive active sites (Reed, 1974). A folded structure of the lipoyl domain attached to the lipoic acid prosthetic group has proven to be essential for recognition by the E1 component of the parent 2-oxo acid dehydrogenase complex (Graham *et al.*, 1989), in contrast to the E2 and E3 component which can use free lipoamide as a substrate (Reed *et al.*, 1958a). In addition, a specific lysine residue of each lipoyl domain is recognised and modified by the lipoylating enzymes responsible for covalent attachment of the lipoic acid to the domain (Reed *et al.*, 1958b).

The high mobility of the lipoyl domains caused by the linker sequences has been the most likely cause of failure of crystallising any 2-oxo acid dehydrogenase complex or any of its acyltransferase components (Mattevi *et al.*, 1992b). Structure determination of the individual components and the domains and linkers of the acyltransferase component is the way to overcome the problems of crystallisation. This approach has resulted in the crystal structure of the catalytic domain of E2p from *A. vinelandii* (Schulze *et al.*, 1991b; Mattevi *et al.*, 1992b, 1993a) and the solution structures of the E2p lipoyl domain from *B. stearothermophilus* (Dardel *et al.*, 1991, 1993), the E3-binding domain from *E. coli*

dihydrolipoyl succinyltransferase (E2o) (Robien *et al.*, 1992) and the E1p/E3-binding domain from *B. stearothermophilus* E2p (Kalia *et al.*, 1993). Crystal structures of the lipoamide dehydrogenase of *A. vinelandii* (Schierbeek *et al.*, 1989; Mattevi *et al.*, 1991), *Pseudomonas putida* (Mattevi *et al.*, 1992c) and *P. fluorescens* (Mattevi *et al.*, 1993b) have been resolved. So far no structural information on the E1 component or on the integration of the individual components in a functional complex is available.

The objective of this study is the determination of the solution structure of the N-terminal lipoyl domain of E2p of *A. vinelandii* PDHC. NMR spectroscopy has the advantage of obtaining structural as well as dynamic information of a protein. Besides structural information, it is important to know the dynamic properties of the lipoyl domain since they are essential for efficient multienzyme catalysis. Moreover, the interactions of the lipoyl domain with the three different complex components can be studied under various conditions by NMR.

Sufficient amounts of lipoyl domain for structural NMR studies were obtained from a sub-clone of the acetyltransferase gene in *E. coli* expressing exclusively the N-terminal lipoyl domain of *A. vinelandii* E2p. Nearly complete ^1H and ^{15}N resonance assignments for the unlipoylated lipoyl domain were obtained using two-dimensional homo- and heteronuclear NMR experiments. By two-dimensional ^1H NMR the unlipoylated and the lipoylated form of the domain were compared. The secondary structure of the lipoyl domain, based on characteristic medium-range and long-range NOE connectivities, amide hydrogen exchange patterns and vicinal coupling constants, is presented. This structure is compared with the solution structure of the lipoyl domain of the *B. stearothermophilus* PDHC (Dardel *et al.*, 1991, 1993).

RESULTS

Over-expression and lipoylation

The vector pAR1 expressing the N-terminal lipoyl domain (residues 1 to 79) of the dihydrolipoyl transacetylase component of the PDHC from *A. vinelandii* was constructed as described under Experimental Procedures, for the following reasons. By sub-cloning the N-terminal lipoyl domain the translation start of the complete transacetylase is not modified, thereby retaining the good *lacZ* promoter activity for the lipoyl domain. The C-terminal residue of the expressed lipoyl domain is determined by the position of the available *SsrII* restriction site in the DNA encoding the complete transacetylase component from which the sub-clone has been derived. Plasmid pAR1 encodes the complete N-

terminal lipoyl domain and a few amino acids of the adjacent Ala+Pro-rich linker sequence. The exact structural C-terminal residue of the N-terminal lipoyl domain is not known but is expected to be situated around Pro74 from sequence comparison with other lipoyl domains. The advantage of expressing a few amino acid residues beyond the expected lipoyl domain is, apart from the fact that one is sure to express at least the complete structural domain, that it might be possible by NMR to observe the expected mobility and low degree of secondary structure of the linker residues. The C-terminal arginine residue is not present in the original sequence but results from the fusion of DNA fragments after DNA manipulation.

Induction with IPTG of *E. coli* TG2 cells carrying plasmid pAR1 resulted in high expression of a soluble polypeptide of approximately 8 kDa (Figure 1A, lane 2), as expected for a single lipoyl domain. This peptide was not detected in cells containing the control plasmid pUC9 (Figure 1A, lane 1).

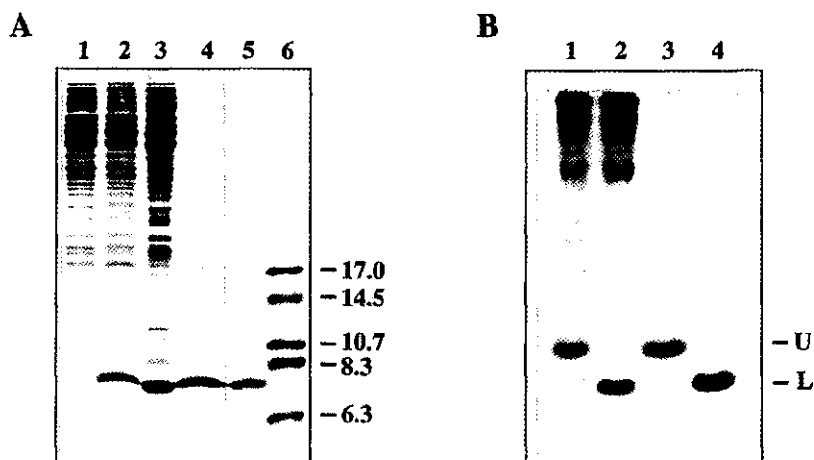


Figure 1. Expression, purification, and characterisation of the lipoyl domain. (A) SDS/PAGE of the expression and purification of the N-terminal lipoyl domain of *A. vinelandii* PDHC in *E. coli*. Lane 1, cell-free extract of *E. coli* TG2 transformed with plasmid pUC9; lane 2, cell-free extract of *E. coli* TG2 transformed with plasmid pAR1; lane 3, cell-free extract after 3 h at 140000 x g; lane 4, filtrate after Amicon YM 30; lane 5, purified lipoyl domain after Q-Sepharose chromatography; lane 6, molecular mass markers (values given in kDa). (B) Non-denaturing PAGE of the unlipoylated and lipoylated forms of the lipoyl domain. Lane 1, cell-free extract of *E. coli* TG2 transformed with plasmid pAR1 expressing the *A. vinelandii* PDHC lipoyl domain, and grown on TY medium without added lipoic acid; lane 2, identical to lane 1 but with lipoic acid added to the TY growth medium; lane 3, isolated unlipoylated lipoyl domain (U); lane 4, isolated lipoylated lipoyl domain (L).

When *E. coli* TG2(pAR1) cells were grown on TY medium without additional lipoic acid, only a minor part (less than 10%) of expressed lipoyl domain was modified, as

judged by native PAGE of these *E. coli* cell-free extracts (Figure 1B, lane 1). Apparently over-expression exceeds the ability of the *E. coli* cells to correctly modify all the expressed *A. vinelandii* lipoyl domain. We were not able to lipoylate the lipoyl domain completely *in vitro* with *E. coli* cell-free extract following the procedure described by Brookfield *et al.* (1991). Only a minor portion (less than 5%) of the unmodified domain could be modified in this manner. However, addition of 10 $\mu\text{g}/\text{mL}$ lipoic acid to the TY growth medium resulted in complete modification of the expressed lipoyl domain (Figure 1B, lane 2). Thus the *E. coli* cells are very well capable of lipoylating the *A. vinelandii* lipoyl domain, but the rate of synthesis of lipoic acid or one of its precursors is too slow compared to the production rate of the recombinant lipoyl domain. Presumably uptake of sufficient amounts of lipoic acid from the growth medium complements the lacking production of lipoic acid in the cell.

Purification and characterisation

The expressed N-terminal lipoyl domain of *A. vinelandii* PDHC was purified from *E. coli* as described under Experimental Procedures. The results of the purification are shown in Figure 1. The Amicon YM30 membrane extremely efficiently separated the lipoyl domain from most other proteins in the cell-free extract (Figure 1A, lane 4), although not all lipoyl domain could be recovered from the protein concentrate. The lipoyl domain was further purified to homogeneity by Q-Sepharose column chromatography, which also separated the unmodified from the modified form of the lipoyl domain. The unmodified form eluted with 0.35 M KCl, while the modified form eluted with 0.6 M KCl. The modified and unmodified form of the lipoyl domain were isolated from 6-L cultures with and without added lipoic acid to the growth medium with yields of approximately 40 mg. No significant difference in expression level of the lipoyl domain between the two cultures could be observed. The small amount of modified lipoyl domain in the culture without additional lipoic acid was not routinely purified.

Both forms of the lipoyl domain appeared homogeneous on SDS/PAGE. When samples of unmodified and modified lipoyl domain were subjected to IEF, two very small additional protein bands were observed for the two forms of the lipoyl domain, both with identical pI difference to the main protein band on the gel. The measured isoelectric points of the two primary forms were 4.5 and 4.4 respectively. The reductive acetylation assay of the modified lipoyl domain showed that approximately 0.8 mol $^{14}\text{C}/\text{mol}$ lipoyl domain could be incorporated. This strongly indicates that the modification and folding of the modified lipoyl domain is correct, since commonly the reductive acetylation assay slightly

underestimates the amount of incorporated label (Packman *et al.*, 1991). No reductive acetylation could be detected with the unmodified lipoyl domain.

Electrospray mass spectrometry of the unmodified form of the domain gave a single component with an M_r of 8166 ± 0.8 , which corresponds to the M_r of the unlipoylated domain form (calculated M_r 8165). The modified form of the protein gave significantly a more noisy spectrum, but the determined M_r of 8352 ± 1.7 was in excellent agreement with the M_r of a lipoylated domain (calculated M_r 8352). Partial mismodification with octanoic acid of expressed lipoyl domains in *E. coli* has been reported earlier for the *B. stearothermophilus* E2p lipoyl domain (Dardel *et al.*, 1990), the human E2p inner lipoyl domain (Quinn *et al.*, 1993), and the *E. coli* E2p lipoyl domain only when expressed in a lipoic-acid-deficient *E. coli* strain (Ali *et al.*, 1990). However, no octanoyl-modified lipoyl domain could be detected in pools of *A. vinelandii* E2p lipoyl domain isolated from *E. coli* TG2(pAR1) cultures grown in the presence or absence of additional lipoic acid in the growth medium.

NMR spectroscopy of the lipoyl domain

The two-dimensional $^1\text{H-NMR}$ spectra of the lipoyl domain show greater overlap than expected for a protein of this size, due to several reasons. First, the lipoyl domain does not contain any aromatic residues, which are known to cause large conformation-dependent shifts (Wüthrich, 1976) as a result of ring current field effects. Secondly, only 14 types of amino acids occur in the lipoyl domain, and even these are unequally distributed. The lipoyl domain consists of 11 Val, 10 Glu, 9 Gly, 8 Leu, 7 Ala, 7 Lys, 7 Ile, 6 Ser, 5 Asp, 4 Pro, 2 Arg, and 1 Thr, Gln and Met. Chemical shift differences of protons of identical residue type are caused by differences in local environment and conformation. The lipoyl domain consists of approximately 60% β -sheet and lacks any helical conformation; hence it is clear that the probability of overlap of resonance lines in the lipoyl domain is considerable. It therefore appeared to be an essential factor in the entire backbone assignment procedure that spectra were recorded at a relatively broad range of temperatures.

Initially, assignments were obtained for the unlipoylated lipoyl domain, since the lipoylated protein has a higher tendency to aggregate, probably due to hydrophobic interactions. On the basis of these assignments the lipoylated protein was assigned and compared with the unlipoylated protein.

During the final stage of the assignment procedure two-dimensional heteronuclear NMR experiments were performed with the ^{15}N -labelled lipoyl domain to solve a few

ambiguities and to confirm the sequential assignments made on the basis of the homonuclear ^1H -NMR experiments.

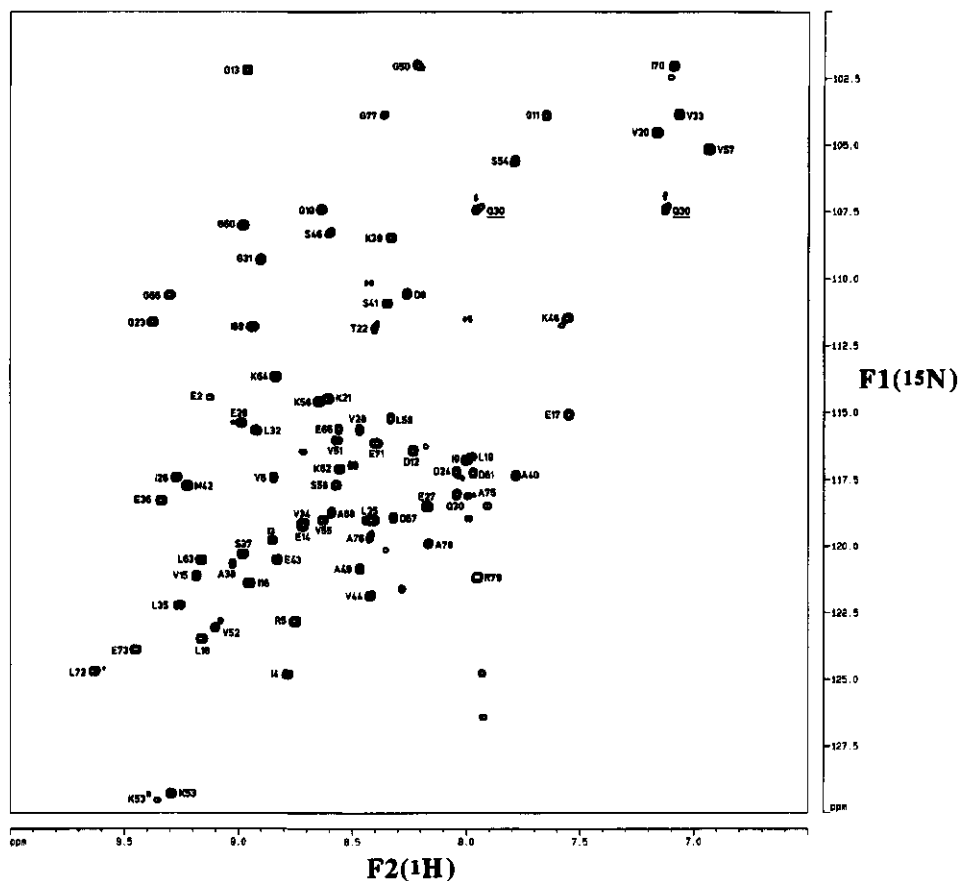


Figure 2. 2D ^{15}N - ^1H HSQC spectrum of the N-terminal lipoyl domain of *A. vinelandii* PDHC, recorded at 500 MHz, 37°C and pH 5.5. The side-chain amide resonances of Gln30 are indicated by an underlined label. The 'satellite' peak of Lys53 is indicated by an asterisk.

Spin system identification

Amino acid spin systems were identified based on their characteristic scalar connectivities, starting with inspection of 2D ^1H DQF-COSY and TOCSY spectra. Since the lipoyl domain lacks any aromatic residues, identification of spin systems could also be based more reliably on the C^βH proton chemical shifts.

Because of resonance overlap in the spectral region where α/β scalar connectivities can be observed, a TQF-COSY experiment was recorded to discriminate between spin systems with one or two $C^\beta H$ protons. A cross peak for two coupled spins can only be observed in the TQF-COSY if these spins are mutually coupled to at least one other spin (Braunschweiler *et al.*, 1983). In this way especially isoleucine $C^\beta H$ protons can be distinguished from $C^\beta H$ protons of lysine, leucine, and arginine, all having their $C^\beta H$ protons resonating at approximately the same frequency (Wishart *et al.*, 1991). The TQF-COSY experiment was also useful to detect glycine residues as they exclusively give rise to cross peaks in the NH- $C^\alpha H$ cross peak region. Seven of nine glycine residues could be identified in this manner.

All seven lysine residues were identified by the weak connectivity of their amide proton to the $C^\epsilon H$ proton in the TOCSY experiment. One of two arginines was identified by the COSY connectivity between the $C^\delta H$ protons and the $N^\epsilon H$ proton. Both the $N^\epsilon H$ proton as well as the backbone amide proton of this arginine residue showed cross peaks toward the $C^\delta H$ protons in the TOCSY experiment. The single glutamine was found by the COSY cross peak between the two $N^\epsilon H$ protons. Through the NOESY connectivity between the $N^\epsilon H$ protons and the $C^\gamma H$ protons, the resonances of the glutamine (Gln30) could be completely identified. One serine was unambiguously identified by the TOCSY cross peak between the amide proton and the hydroxyl proton, and the COSY connectivities between the hydroxyl proton and the two $C^\beta H$ protons.

Together only 54 spin systems could be identified in the initial stage of the assignment procedure. The remaining spin systems, including three of the four proline residues, were identified in the course of the sequential assignments. In the final stage of the assignment procedure several spin systems were completed after inspection of the HMQC-TOCSY spectra.

Sequential assignments

Sequential resonance assignments for the lipoyl domain were obtained according to standard methods (Wüthrich, 1986; Gronenborn *et al.*, 1989), and were based primarily on the 2D 1H -NOESY spectra in H_2O . Most residues in the antiparallel β -sheet regions were readily identified and connected by their strong sequential $d_{\alpha N}$ cross peaks in the NOESY spectra. Most sequential assignments in the β -sheet regions were confirmed by long-range NOEs to the neighbouring strand. Outside these regions residues were connected by significant but weaker sequential $d_{\alpha N}$ NOEs. When $C^\alpha H$ protons kept overlapping with the water resonance even at different temperatures, as occurred for Leu19 and Leu32, sequential connections were solved by $d_{\beta N}$ NOEs.

Table 1. ^1H and ^{15}N chemical shifts (ppm) for the unlipoylated form of the N-terminal lipoyl domain of the pyruvate dehydrogenase complex from *A. vinelandii* at pH 5.5 and 37°C, in 50 mM potassium phosphate containing 150 mM potassium chloride.

Residue	Amide		C^αH	C^βH	Other
	^{15}N	^1H			
Ser1			4.42	4.07, 3.95	
Glu2	118.5	9.14	4.89	2.11, 2.00	C^γH 2.25
Ile3	123.8	8.87	4.71	1.93	1.64, 1.30, 0.88
Ile4	128.9	8.79	4.50	2.04	1.33, 0.99, 0.83
Arg5	126.9	8.76	5.32	1.63	C^γH 1.37, C^δH 3.48, 3.21, $\text{N}^\epsilon\text{H}$ 7.46, N^ϵ 81.3
Val6	121.5	8.85	3.98	2.39	$\text{C}^\gamma\text{H}_3$ 1.13
Pro7			4.97	2.35	C^γH 2.26, C^δH 3.98, 3.50
Asp8	114.7	8.27	4.53	2.88, 2.72	
Ile9	120.9	8.01	4.52	2.20	1.45, 1.15, 0.87, 0.75
Gly10	111.5	8.64	4.14, 3.91		
Gly11	108.0	7.66	4.32, 3.99		
Asp12	120.5	8.23	5.26	2.69	
Gly13	106.4	8.96	3.34		
Glu14	123.4	8.73	5.36	2.05	C^γH 2.16
Val15	125.2	9.20	4.14	2.52	$\text{C}^\gamma\text{H}_3$ 1.09, 0.85
Ile16	125.6	8.95	4.73	2.11	1.04
Glu17	119.2	7.56	4.50	1.98, 1.91	C^γH 2.33, 2.22
Leu18	127.5	9.17	4.77	1.87, 1.75	C^γH 1.70, $\text{C}^\delta\text{H}_3$ 1.02, 0.89
Leu19	120.8	7.99	4.76	1.86, 1.71	C^γH 1.59, $\text{C}^\delta\text{H}_3$ 0.85, 0.68
Val20	108.6	7.17	4.98	2.33	$\text{C}^\gamma\text{H}_3$ 1.18, 0.95
Lys21	118.6	8.62	4.76	1.93, 1.87	1.67, 1.34, ($\text{C}^\epsilon\text{H}$ 2.86)
Thr22	115.9	8.41	3.55	4.07	$\text{C}^\gamma\text{H}_3$ 1.30
Gly23	115.7	9.38	4.51, 3.58		
Asp24	121.3	8.05	4.60	2.86, 2.62	
Leu25	123.1	8.41	4.93	1.85, 1.66	$\text{C}^\delta\text{H}_3$ 0.99
Ile26	121.4	9.28	5.06	1.80	1.31, 1.23, 0.91, 0.79
Glu27	122.7	8.18	4.78	2.28, 1.95	C^γH 2.54, 2.43
Val28	119.8	8.47	3.29	2.00	$\text{C}^\gamma\text{H}_3$ 1.10
Glu29	119.5	8.99	3.81	2.44	
Gln30	122.2	8.05	4.18	2.06	C^γH 2.34, 2.56, $\text{N}^\epsilon\text{H}$ 7.13, 7.91, N^ϵ 107.4
Gly31	113.4	8.91	3.90, 3.58		
Leu32	119.8	8.93	4.76	2.06, 1.97	C^γH 1.30, $\text{C}^\delta\text{H}_3$ 0.79, 0.66
Val33	108.0	7.08	4.98	2.29	$\text{C}^\gamma\text{H}_3$ 1.03, 0.94

Table 1. (Continued)

Residue	Amide		C $^{\alpha}$ H	C $^{\beta}$ H	Other
	15 N	1 H			
Val34	123.3	8.72	4.91	1.97	C $^{\gamma}$ H $_3$ 1.02, 0.99
Leu35	126.3	9.27	5.51	1.81, 1.11	C $^{\gamma}$ H 1.67, C $^{\delta}$ H $_3$ 0.82
Glu36	122.4	9.35	5.14	2.16, 2.04	C $^{\gamma}$ H 2.32
Ser37	124.3	8.99	5.12	4.20, 3.91	
Ala38	124.8	9.02	4.23	1.58	
Lys39	112.5	8.33	4.36	1.92	1.51, C $^{\epsilon}$ H 3.12
Ala40	121.4	7.80	4.68	1.51	
Ser41	115.1	8.35	5.47	3.83	
Met42	121.9	9.23	4.88	2.09	C $^{\gamma}$ H 2.52
Glu43	124.6	8.85	5.10	2.04	C $^{\gamma}$ H 2.18
Val44	125.9	8.42	4.42	2.06	C $^{\gamma}$ H $_3$ 0.96
Pro45			4.92	2.16	C $^{\gamma}$ H 2.06, C $^{\delta}$ H 4.22
Ser46	112.4	8.60	4.76	4.20, 3.89	OH 5.44
Pro47			4.00	2.37	C $^{\gamma}$ H 2.37, 2.00, C $^{\delta}$ H 4.11, 4.00
Lys48	115.6	7.57	4.70	(1.75, 1.60)	1.36, C $^{\epsilon}$ H 3.04
Ala49	124.9	8.47	4.41	1.42	
Gly50	106.1	8.23	4.60, 3.93		
Val51	120.1	8.58	5.04	1.95	C $^{\gamma}$ H $_3$ 1.03, 0.90
Val52	127.1	9.11	3.94	2.49	C $^{\gamma}$ H $_3$ 1.06, 0.88
Lys53	133.4	9.30	4.49	1.78, 1.66	1.50, (C $^{\epsilon}$ H 3.05)
Ser54	109.7	7.80	4.68	3.92, 3.86	
Val55	123.1	8.63	4.59	2.24	C $^{\gamma}$ H $_3$ 1.30, 0.98
Ser56	121.8	8.57	4.61	3.98, 3.47	
Val57	109.3	6.95	4.90	2.41	C $^{\gamma}$ H $_3$ 1.26, 0.93
Lys58	118.7	8.66	4.70	1.88	1.77, 1.62, 1.38, C $^{\epsilon}$ H 3.15
Leu59	119.3	8.34	3.82	1.77, 1.74	C $^{\gamma}$ H 1.50, C $^{\delta}$ H $_3$ 1.05, 0.89
Gly60	112.1	8.99	4.49, 3.91		
Asp61	121.4	7.98	4.63	2.91, 2.67	
Lys62	121.1	8.55	5.03	1.89, 1.85	1.58, 1.46, C $^{\epsilon}$ H 3.10
Leu63	124.6	9.16	5.13	1.72	1.02, 0.89
Lys64	117.8	8.85	4.80	1.78	1.42, C $^{\epsilon}$ H 3.14
Glu65	119.7	8.57	3.87	1.94	C $^{\gamma}$ H 2.18
Gly66	114.7	9.31	4.60, 3.69		
Asp67	123.0	8.33	4.77	2.92, 2.81	
Ala68	122.8	8.60	4.04	1.39	
Ile69	115.9	8.94	4.41	1.58	0.95, 0.79, 0.74

Table 1. (Continued)

Residue	Amide		C^αH	C^βH	Other
	^{15}N	^1H			
Ile70	106.2	7.10	5.18	2.28	1.16, 0.97
Glu71	120.3	8.42	5.11	2.01	C^γH 2.24
Leu72	128.4	9.63	5.15	1.76, 1.30	$\text{C}^\delta\text{H}_3$ 0.63
Glu73	128.0	9.46	5.20	2.23, 1.96	C^γH 2.36
Pro74					
Ala75	122.2	7.99	4.42	1.54	
Ala76	123.7	8.43	4.44	1.51	
Gly77	108.0	8.36	4.05		
Ala78	124.0	8.18	4.47	1.50	
Arg79	125.2	7.96	4.28	1.97, 1.83	C^γH 1.72, C^δH 3.31

^1H chemical shifts (± 0.02 ppm) are referenced to internal trimethylsilyl propionate; ^{15}N chemical shifts (± 0.2 ppm) are referenced to external liquid ammonia. Parentheses indicate that the resonance position of the proton is determined exclusively via TOCSY spectra, and should therefore be regarded as tentative. Resonances in the final column without a label belong to the spin system but their position in the side chain could not be determined.

In regions connecting the different β -strands several residues were also connected through their strong sequential d_{NN} NOEs.

Sequential assignments past proline residues could not be obtained except for Pro47 to which sequential $d_{\alpha\delta}$ NOEs were observed. No sequential contacts between Val6 and Pro7 could be distinguished, since the C^αH proton of Val6 overlaps with one of the C^βH protons of Pro7. A sequential $d_{\alpha\alpha}$ NOE, characteristic for a *cis*-proline could not be observed either. Sequential contacts between Val44 and Pro45 could not be observed, probably due to overlap.

It appeared difficult to assign the residues of the presumably flexible C-terminus of the expressed lipoyl domain (from Pro74 to Arg79). The resonances from these residues were shown to be significantly more narrow than resonances from other residues, suggesting a higher mobility. These residues could only be assigned by combining the two-dimensional homonuclear NMR experiments with the HMQC-NOESY and HMQC-TOCSY experiments recorded at 14°C . The starting point for the assignment of these residues was Gly77, which was readily identified in the HMQC-TOCSY spectra by its high field amide ^{15}N chemical shift. No assignments could be made for Pro74, probably due to flexibility combined with the possibility of *cis-trans* isomerisation.

The ^{15}N resonances were readily assigned according to the amide proton assignments using the HSQC and the HMQC-TOCSY experiments. All sequential

resonance assignments obtained from the 2D ^1H -NOESY experiments were unambiguously confirmed by the HMQC-NOESY experiments. The remaining ambiguities caused by overlapping amide protons were solved by the chemical shift dispersion in the ^{15}N dimension, as is shown in the HSQC spectrum (Figure 2). After completion of the sequential connections, several incomplete spin systems were assigned through intraresidue NOEs. Most of the isoleucine and lysine residues could not be completely assigned due to overlap in the upfield regions of the spectra. Complete ^{15}N and ^1H resonance assignments were obtained for 57 residues, and partial assignments for 21 residues. A survey of the assigned ^{15}N and ^1H chemical shifts in the lipoyl domain is given in Table 1.

After assignment of all amide protons and observable side-chain protons of the lipoyl domain in the HSQC spectrum, several peaks with relative low intensity remained unassigned (Figure 2). Close examination of these peaks revealed that they might belong to peaks in their direct vicinity in the spectrum, therefore designated as 'satellite peaks' here. One of the most intense satellite peaks, in close vicinity of Lys53, showed identical NOE cross peaks as Lys53 itself in the HMQC-NOESY spectrum (Figure 3). Moreover, tentatively assigned satellite peaks of slowly exchanging amide protons showed similar exchange rates as their 'parent peaks'. From these observations the existence of multiple conformations is suggested. Although it is not possible to prove which satellite peak belongs to which residue, on the basis of vicinity and hydrogen exchange rate for Lys48, Val52, Lys53, Ile70, and Leu72 a satellite peak was tentatively assigned. The remaining 12 peaks with low intensity in the HSQC spectrum cannot yet be assigned. Due to overlap, the satellite peaks were not recognised in the homonuclear NMR spectra, except for Leu72 (Figure 4). The satellite peak for Leu72 is also present in homonuclear 1D and 2D NMR spectra of the lipoylated form of the lipoyl domain.

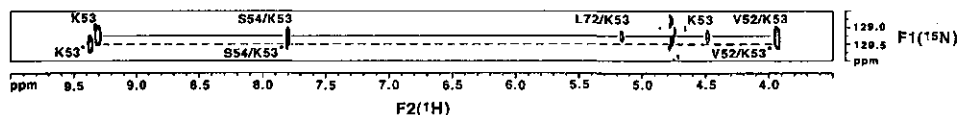


Figure 3. Selected region of a 2D ^{15}N - ^1H HMQC-NOESY spectrum of the lipoyl domain. The sequential $d_{\alpha\text{N}}(52,53)$ and $d_{\text{NN}}(53,54)$ connectivities for Lys53 (solid line) and its 'satellite' partner Lys53* (broken line) are labelled. The labelled intraresidue $d_{\alpha\text{N}}$ connectivity and the long-range $d_{\alpha\text{N}}(72,53)$ connectivity are only observed for Lys53.

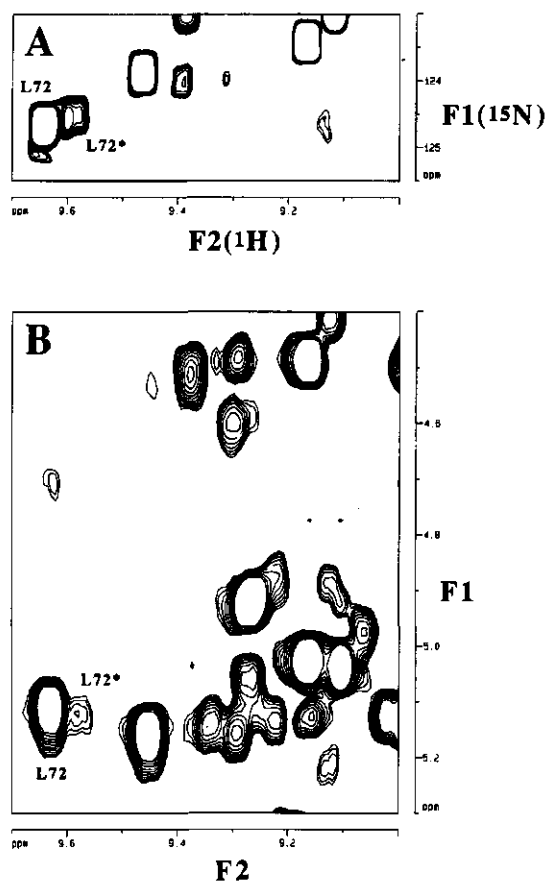


Figure 4. Selected regions of a 2D ^{15}N - ^1H HMQC-NOESY spectrum (A) and a 2D ^1H -NOESY spectrum (B) of the lipoyl domain. The intraresidue $d_{\alpha\text{N}}$ connectivities for Leu72 and its satellite peak Leu72* are labelled in both spectra.

Residues with satellite peaks appear to be situated around Lys53 in β -strand S6 or in the neighbouring strand S8, with the exception of Lys48. This suggests that a population of the lipoyl domain possesses a slightly different local conformation with altered resonance positions for residues around Lys53. Integration of all peaks in the HSQC spectrum revealed that the satellite peaks show about 2-10% of the intensity of the majority of the peaks. The intensity of the satellite peaks is not changed in a HSQC spectrum of the lipoyl domain recorded at an increased pH value of 7. Thus, only a small population of the lipoyl domain exists in a locally slightly different conformation. The structural change is

not situated in the β -sheet containing the lipoyl-lysine residue (Lys39), and no satellite peaks are observed in the vicinity of this residue. The exchange rate between the different forms of the lipoyl domain is slow compared to the NMR time scale, since separate narrow resonances for the protons in different local conformations can be observed at 500 MHz. The chemical shift difference between a peak and its satellite partner is at least 25 Hz, suggesting that the exchange rate constant for the different conformations is much smaller than about 5 s^{-1} .

Sequential resonance assignments for the lipoylated form of the domain were obtained using the assignments for the unlipoylated domain. No ^{15}N assignments for the lipoylated domain were obtained. The 2D ^1H -NOESY and COSY spectra of the lipoylated and the unlipoylated forms of the lipoyl domain recorded in H_2O are very similar. Only protons of Lys39, residues in the direct neighbourhood of the lipoyl-Lys39 residue, and residues around Gly10 show differences in chemical shift values. Residues for which chemical shift values of protons of the lipoylated domain differed more than 0.05 ppm from protons of the unlipoylated domain are listed in Table 2 together with their chemical shift difference. Most amide protons are shifted upfield while most C^αH protons show a downfield shift.

Table 2. Differences in chemical shift value of protons of the lipoylated form of the lipoyl domain at pH 5.5 and 30°C , relative to the chemical shift values of the unlipoylated domain. Only chemical shift differences larger than 0.05 ppm are reported. (+) downfield shift, (-) upfield shift.

Residue	Proton(s)	Chemical shift difference (ppm)
Ile9	NH	-0.07
Gly10	NH	-0.06
Gly11	NH	-0.09
Asp12	C^αH	+0.07
Glu14	NH	+0.08
Glu36	C^γH	+0.08
Ser37	$\text{C}^\beta\text{H}_1, \text{C}^\beta\text{H}_2$	+0.20, +0.14
Ala38	NH	-0.14
Lys39	NH	-0.27
	C^αH	+0.07
Ala40	NH	-0.17
	C^αH	-0.13
Ser41	NH	-0.12
	C^αH	+0.06

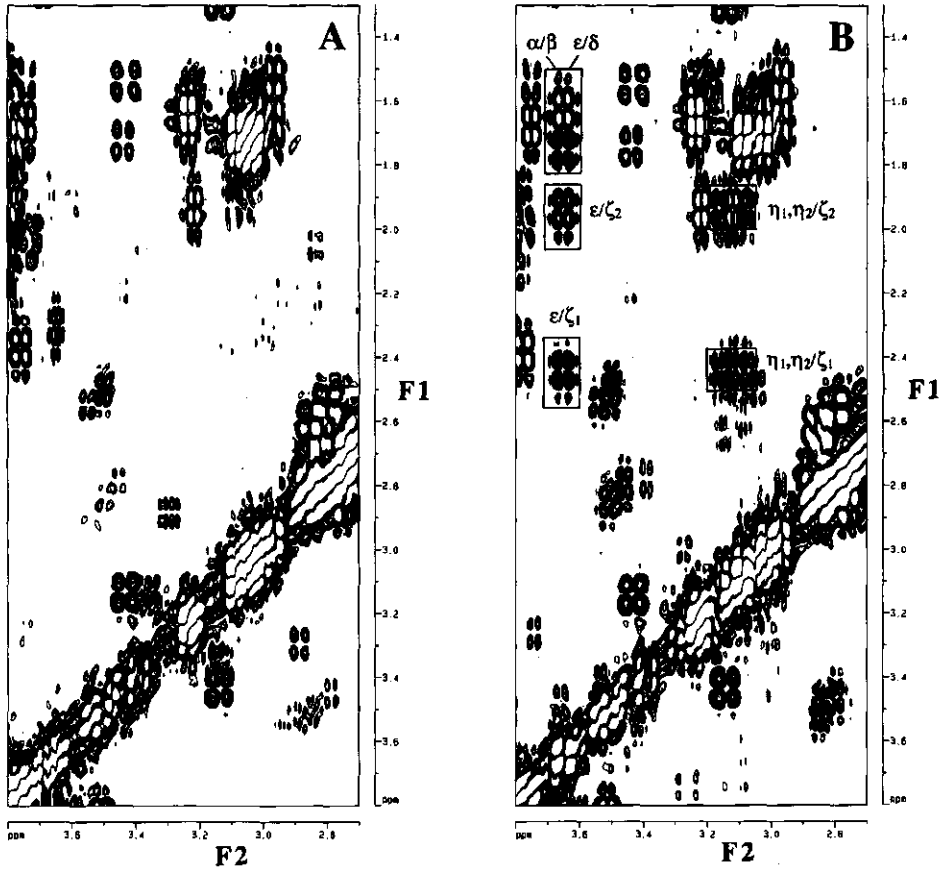


Figure 5. Comparison of a selected region of the 2D DQF-COSY spectrum of the unlipoylated form (A) and the lipoylated form (B) of the lipoyl domain, recorded at 30°C and pH 5.5 in 90% $\text{H}_2\text{O}/10\%$ D_2O . Both positive and negative levels are drawn without distinction. The cross peaks belonging to the lipoyl group are boxed and the assignments are indicated [proton(s) in F_2 / proton(s) in F_1].

Comparing the DQF-COSY spectra of the unlipoylated and lipoylated forms of the protein, several scalar connectivities belonging to the liponic acid moiety can be recognised (Figure 5). On the basis of their COSY coupling network and chemical shift position these resonances are tentatively assigned.

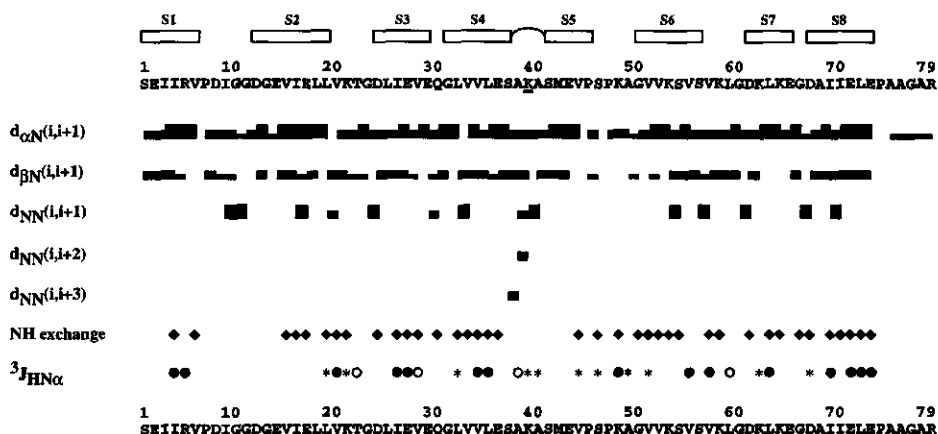


Figure 6. Summary of sequential and medium-range NOE connectivities, slowly exchanging amide protons, and $^3J_{HN\alpha}$ coupling constants observed for the lipoyl domain. In the lines showing the amino acid sequence, the lipoyl-lysine residue is underlined. NOE intensities were derived from NOESY spectra with a mixing time of 150 ms. The thickness of the bars reflects a qualitative measure of the strength of the NOEs, classified as strong, medium or weak. Amide protons that could still be observed 14 h after dissolving the protein in D_2O , at pH 5.5 and $30^\circ C$, are marked with (♦). The $^3J_{HN\alpha}$ coupling constants are classified as (○) < 6 Hz, and (●) > 8 Hz. An asterisk indicates that the determination of the coupling constant was impaired by overlap. In the upper line the secondary structure elements, as derived from NMR data, are denoted.

Secondary structure

Secondary structure elements of the lipoyl domain were identified from characteristic NOEs, NH-exchange rates, and $^3J_{HN\alpha}$ values (Figure 6). On the basis of strong sequential $d_{\alpha N}$ NOEs, large $^3J_{HN\alpha}$ values (> 8 Hz), and interstrand d_{NN} , $d_{\alpha N}$, and $d_{\alpha\alpha}$ connectivities, eight extended peptide segments could be deduced, interrupted by loop regions. These segments could then be assembled into two antiparallel β -sheets of four strands each (Figure 7). One consists of S1(Ser1-Val6), S3(Asp24-Glu29), S6(Gly50-Ser56), and S8(Asp67-Glu73), and the other of S2(Asp12-Leu19), S4(Gly31-Ser37), S5(Ser41-Pro45), and S7(Asp61-Glu65). The neighbouring β -strands S4 and S5 are connected by a well-defined type I turn (Richardson, 1981; Wüthrich *et al.*, 1984) from Ser37 to Ala40, characterised by the strong $d_{NN}(39,40)$ and medium $d_{NN}(38,39)$, and the small $^3J_{HN\alpha}$ value (< 6 Hz) of Ala38. This turn is located at the corner of one of the sheets, and contains Lys39 that is modified with lipoic acid forming the lipoyl-lysine residue. The majority of the residues bordering the β -strands show strong sequential d_{NN} contacts, indicating a strong bending of the backbone.

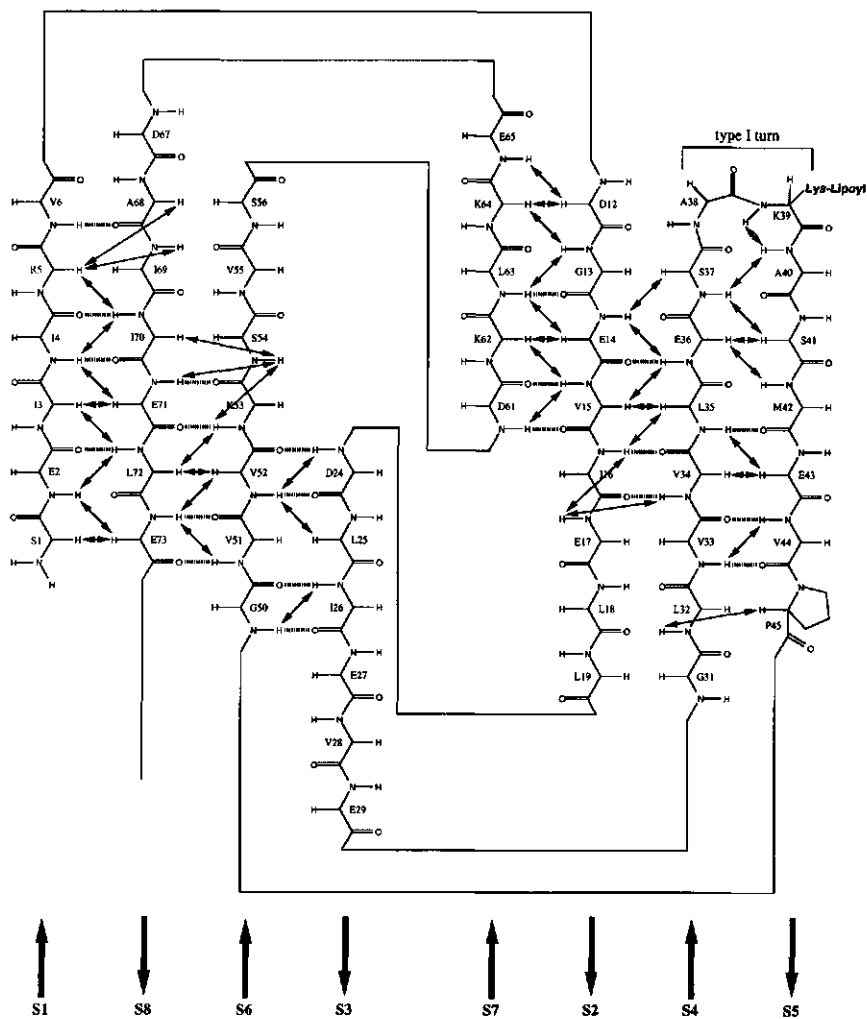


Figure 7. Schematic drawing of the secondary structure of the lipoyl domain in solution. Arrows denote observed medium- and long-range medium and strong NOE connectivities between backbone protons. Hydrogen bonds are indicated by hatched lines, and are drawn on the basis of the observation of slowly exchanging amide protons.

A sharp hairpin-like turn in the protein backbone is located around Gly10, which shows strong d_{NN} contacts with its neighbouring residues Ile9 and Gly11. This turn is further characterised by NOEs between the amide proton of Gly11 and the side-chain protons of Ile9. It can however not be classified as belonging to type I or type II tight turns.

To detect slowly exchanging amide protons, HSQC spectra of a freshly prepared sample in D_2O were recorded at different time intervals. Many of these protons are located

in the β -strands and, on the basis of the postulated interstrand hydrogen bonds, the NH-exchange experiments confirm the arrangement of these strands into two antiparallel β -sheets. 2D ^1H -NOESY spectra of an identically prepared sample of unlabelled lipoyl domain verified the identity of the hydrogen-bonded amide protons. In addition, these spectra revealed the presence of a slowly exchanging (several hours) hydroxyl proton of Ser46. This hydroxyl proton shows a very strong NOE connectivity to the amide proton of Lys48, suggesting that this proton is hydrogen-bonded to the backbone nitrogen atom of Lys48.

The slow exchange rate of the amide protons of Val20, Lys21, Gly23, Val57, and Lys58, which are situated in the peptide segments connecting the two β -sheets, suggests that at least some of these segments possesses secondary structure that is not random coil. The amide protons of Ser37 and Ala40 show a relatively fast exchange rate. These residues are situated in the type I turn at the corner of one of the two sheets which contains the lipoyl-lysine residue.

Several irregularities in the β -sheets are observed. A bulge is located in strand S8 around Ala68, which shows rapid NH-exchange and no d_{NN} contacts with the opposite strand. Instead, a very strong $d_{\alpha\alpha}$ contact between Arg5 and Ala68 is observed together with a $d_{\alpha\text{N}}$ contact between Arg5 and Ile69. Other irregularities are situated around Ser54 in strand S6 which shows strong unexpected d_{NN} contacts with Glu71 and Lys53, and around Glu17 in strand S2 showing strong unexpected d_{NN} contacts with Val34 and Ile16.

The five C-terminal residues of the lipoyl domain show only short-range NOEs and high amide proton exchange rates. The sequential NOEs for these residues became visible only in the NOESY spectra recorded at lower temperatures. Together with the observation of significant narrower linewidths of resonances from these residues, it is concluded that the C-terminus has a disordered structure and is not likely to be involved in the global fold of the lipoyl domain. These flexible residues are thought to belong to the first mobile Ala+Pro-rich linker sequence separating the N-terminal lipoyl domain from the second lipoyl domain.

The two-dimensional ^1H -NOESY spectra of the unlipoylated and lipoylated forms of the lipoyl domain are almost identical, except for several additional resonances belonging to the lipoic acid moiety, and chemical shift differences of protons of residues in the neighbourhood of the lipoyl-lysine residue. No changes in NOE intensities connecting these residues nor addition or loss of short-, medium- or long-range NOEs could be observed however, suggesting that the structure of the lipoyl domain is not altered much if any by its covalently coupled lipoic acid prosthetic group.

Global fold

The secondary structure of the lipoyl domain consists of two antiparallel β -sheets of four strands each, connected by peptide segments that do not form regular secondary structure. Both β -sheets show some kind of amphiphilic character, with an asymmetric distribution of polar and nonpolar side chains, resulting in a predominantly hydrophobic and hydrophilic side of each sheet. It is likely that the two hydrophobic sides of the sheets will face each other to form a hydrophobic interior, while the two hydrophilic sides will form the surface of the protein. This relative three-dimensional orientation of the two β -sheets is supported by a number of long-range NOEs observed between residues in the two different sheets. For example, the C^βH proton of Leu72 shows a NOE cross peak to the C^αH proton of Ile16, and the C^αH proton of Glu65 shows a NOE cross peak to a $\text{C}\gamma\text{H}$ proton of Val6.

DISCUSSION

Toward the solution structure of the N-terminal lipoyl domain of dihydrolipoyl transacetylase of the *A. vinelandii* PDHC, nearly complete sequential ^1H and ^{15}N resonance assignments were obtained using two-dimensional homo- and heteronuclear NMR. The secondary structure of the lipoyl domain, as revealed from the reported NMR data, is well defined in solution. The domain contains two four-stranded antiparallel β -sheets, and a short disordered C-terminus from Pro74 to Arg79. In one of the sheets, the lipoyl-lysine residue is situated in a well-defined type I turn which connects strand S4 with the peripheral strand S5. Consequently, the lipoyl group is exposed to the solvent where it can act as a substrate anchor connecting the cascade of three different enzyme activities in the multienzyme complex. In the other sheet, the N-terminus and C-terminus of the folded domain are close together in the neighbouring strands S1 and S8. Together with the observation of several NOE connectivities between the hydrophobic sites of both β -sheets and the topology of the eight β -strands, it is suggested that the lipoyl domain folds like a jelly roll barrel around a predominantly hydrophobic core. By NMR it is observed that the folded lipoyl domain ends around Pro74. The five C-terminal residues prove to be flexible, as expected, and thus belong to the mobile linker sequence rich in Ala and Pro residues. The mobility of the linkers, separating the lipoyl domains from each other and from the peripheral subunit-binding domain, is essential for the lipoyl domains to reach the three different active sites in the multienzyme complex.

The NMR spectra of the lipoyl domain show a number of extra cross peaks in the fingerprint region with low intensity which remain unassigned. The excess number of cross peaks are most obviously detected in the HSQC spectra. These peaks seem to belong to peaks in their direct vicinity in the spectra and were therefore designated as satellite peaks. The satellite peaks are proposed to arise from conformational inhomogeneity of the lipoyl domain, and most residues with satellite peaks appear to be situated around Lys53. A tentative explanation for the small population of lipoyl domain with a local deviating structure could be found in the nature of the residues surrounding Lys53. In close vicinity Glu71 and Glu73 are situated, which in principle could both be candidates to form a salt bridge with Lys53. Instability of this salt bridge could then lead to a population of lipoyl domain with a slightly different structure. Biochemical support for this hypothesis comes from the observation of minor but identical inhomogeneity for both unlipoylated and lipoylated domain when subjected to IEF. This means that two extra protein bands are observed for the two forms of the lipoyl domain, both with identical pI difference to the main protein band on the gel. The presence or absence of an additional salt bridge could lead to a different pI of a protein, and therefore could be an explanation for the identical inhomogeneity found on the IEF gel.

The two-dimensional proton NMR spectra of the unlipoylated and lipoylated forms of the lipoyl domain are almost identical. This suggests that the structure of the lipoyl domain is not altered significantly by the covalently coupled and solvent-exposed lipoic acid prosthetic group, as expected. However, chemical shift differences of protons of residues surrounding the lipoyl-lysine residue and the lipoyl-lysine residue itself are observed (Table 2). Most shifted amide protons in this region show an upfield shift indicating an increased shielding. This could possibly be attributed to a slight loosening of the already weak hydrogen bonds involved in this tight turn, but will not result in significant structural changes.

The changes of chemical shifts, due to the coupled prosthetic group, of protons belonging to residues around Gly10 however adds more to the understanding of the possible function of the loop comprising this residue. These changes strongly indicate that this loop is situated in close vicinity of the lipoyl group, which is possible on the basis of the topology of the β -strands (Figure 7). Based on the observation that this region shows the highest level of local similarity among all lipoyl domain sequences, Dardel *et al.* (1993) proposed that this loop might be an important determinant of the lipoyl domain for the process of molecular recognition by the E1 component in the multienzyme complex. However, a close examination of the sequence alignment of lipoyl domains (Dardel *et al.*, 1993) reveals that in this region significant sequential differences can be observed between

different acyltransferases from the same source. For example, for the lipoyl domains of E2p from *A. vinelandii* and *E. coli* the sequence toward Gly10 is -PDIG-, whereas the aligned sequences for the lipoyl domains of E2o (Westphal & de Kok, 1990) are -PTFP- and -PDLP- respectively. When this sequence is to serve as a determinant of the lipoyl domain for the specific recognition by the E1 component, it needs to be different among different 2-oxo acid dehydrogenase multienzyme complexes from the same source. Furthermore, this determinant should be in close proximity of the lipoyl group, as is proven by our findings for the hairpin-like loop around Gly10. It is therefore proposed that this loop is the major determinant of the lipoyl domain for the specific protein-protein interactions within the enzyme complex.

The tentative assignment of protons of the lipoic-acid moiety from the DQF-COSY spectra of the lipoylated form of the domain (Figure 5) contradicts the assignment of Dardel *et al.* (1991) for this protein-bound prosthetic group. Our observation suggests that the C^αH protons of the lipoic-acid moiety resonate at nearly but not identical frequencies, based on the exceptionally large cross-peak width, and cross-peak pattern. Both C^αH protons show then connectivities to the separate C^βH protons. On the basis of the cross-peak connectivity pattern, it was not possible to discriminate between the resonance positions of the C^βH protons and the C^δH protons, since the C^αH protons and the $\text{C}^\epsilon\text{H}$ protons resonate at approximately the same frequency.

Comparison of the secondary structure of the N-terminal lipoyl domain of the *A. vinelandii* PDHC with the solution structure of the lipoyl domain of the *B. stearothermophilus* PDHC (Dardel *et al.*, 1991, 1993) shows resemblance to a large extent. First, the arrangement of the eight β -strands into two antiparallel β -sheets is identical, as is the position of lipoyl-lysine residue in the type-I turn at the corner of one of the sheets. Secondly, the loop region around Gly10 is present in both proteins, although this loop might be somewhat smaller in the *A. vinelandii* PDHC lipoyl domain. After Gly11 a gap of three residues occurs in the *A. vinelandii* lipoyl domain sequence when compared to the sequence of the *B. stearothermophilus* PDHC lipoyl domain. Furthermore, all strong sequential d_{NN} contacts, indicating strong bends of the backbone of the folded protein, occur at identical positions in the aligned sequences. The exceptions are the strong sequential d_{NN} contacts in the flexible C-terminus that are not found for the *A. vinelandii* PDHC lipoyl domain. In addition, the strongly hydrogen-bonded hydroxyl proton of Ser46, surrounded by Pro45 and Pro47, was not observed for the *B. stearothermophilus* PDHC lipoyl domain. The sequence -PSP- is conserved in both lipoyl domains, but the proposed hydrogen-bond acceptor Lys48 is replaced by a valine residue in the *B. stearothermophilus* PDHC lipoyl domain sequence. This however cannot explain why this seryl hydroxyl

proton is not observed by Dardel *et al.* (1993) for the *B. stearrowophilus* PDHC lipoyl domain, as we propose hydrogen bonding from the hydroxyl group toward the amide backbone function.

It is clear that the three-dimensional solution structures of the N-terminal lipoyl domain of *A. vinelandii* PDHC and the lipoyl domain of *B. stearrowophilus* PDHC are expected to be strikingly similar, although only a secondary structure of the *A. vinelandii* PDHC lipoyl domain is available as yet. Our results strongly confirm the suggestion of Dardel *et al.* (1993) that all lipoyl domains fold in a similar way. Hence it is evident that the second and third lipoyl domain of the *A. vinelandii* PDHC will have almost identical tertiary structures as its N-terminal lipoyl domain, since their sequences are highly similar and they should act as equal substrates for their parent complex.

The tertiary structure of the lipoyl domain of *A. vinelandii* PDHC is now being determined by distance geometry calculations and is a further step to the completion of the three-dimensional structure of the entire multienzyme complex. Moreover it opens a unique way to study the interactions between the lipoyl domain and the other complex components, since the crystal structures of both the acetyltransferase core and the lipoamide dehydrogenase component of the *A. vinelandii* PDHC are available. The interactions of the lipoyl domain with other complex components will be studied by NMR and crystallography, and can be combined with docking experiments to gain a better understanding of the different processes of interaction and recognition in this multienzyme complex.

EXPERIMENTAL PROCEDURES

Gene cloning and expression

For the construction of the plasmid, expressing the N-terminal lipoyl domain of E2p of *A. vinelandii* PDHC, the plasmid pRA282 (Hanemaaijer *et al.*, 1988) containing the complete E2p gene in pUC9, was used as starting material. In plasmid pRA282 two *Cla*I sites are present at positions 407 and 2828 in the original sequence, but the *Cla*I site at position 407 is sensitive to *E. coli* *dam* methylation and could not be cleaved. Plasmid pRA282 was digested with *Cla*I and *Sst*II, and a 3.4-kb fragment comprising the N-terminal lipoyl domain and the pUC9 part of the vector was isolated. The *Cla*I and *Sst*II sites were made blunt using T4 DNA polymerase. Plasmid pAR1 was obtained after ligation of the resulting fragment, thereby introducing a stop codon (UGA) at position 692.

Plasmid pAR1 was characterised by digestion analysis and by sequencing the coding region of its insert including the introduced stop codon.

This plasmid encodes the complete N-terminal lipoyl domain and the first few adjacent amino acids of the Ala+Pro rich sequence. The codon encoding the C-terminal amino acid (Arg79) results from the fusion after ligation of the blunt ends of the DNA fragment.

E. coli strain TG2 (Gibson, 1984), a *recA*⁻ version of TG1 [$\Delta(lac-pro)$, *thi*, *supE*, *Res*⁻ *Mod*⁻ (*k*), *F'* (*traD36 proA*⁺*B*⁺, *lacIq lacZ* Δ M15)] was used for the expression of the lipoyl domain. *E. coli* TG2 cells transformed with the plasmid pAR1, were grown at 37°C in TY medium containing 75 μ g/mL ampicillin and 20 μ g/mL IPTG. When expression of lipoylated protein was required, 10 μ g/mL lipoic acid was added. After incubation for 24 h, 1.5 mL cell culture was concentrated fourfold and cells were disrupted by sonication. After centrifugation for 10 min at 13000 \times g, a fraction of the supernatant (cell-free extract) was subjected to polyacrylamide gel electrophoresis for analysis of expression.

Isolation of the expressed lipoyl domain

The N-terminal lipoyl domain of *A. vinelandii* E2p was isolated from *E. coli* TG2(pAR1). 30 g cells from a 6-L culture in TY medium supplemented with 10 mg/L lipoic acid (optional) was suspended in 100 mL 20 mM potassium phosphate pH 6.0 containing 0.5 mM EDTA and 0.02% NaN₃ (buffer A). Throughout the isolation the presence of the lipoyl domain was monitored by SDS/PAGE. The cells were disrupted using a French press. After centrifugation for 3 h at 140000 \times g, the supernatant was concentrated to 10 mL by ultrafiltration (Amicon YM 30). The ultrafiltration was repeated several times after a fourfold dilution of the concentrate. The filtrate, containing the lipoyl domain, was applied to a Pharmacia HiLoad Q-Sepharose column (55 mL bed volume, 5 mL/min flow rate) equilibrated with buffer A and eluted with a gradient of 0-1.0 M KCl in buffer A. Fractions containing the lipoyl domain were analysed by native polyacrylamide gel electrophoresis, pooled, concentrated by ultrafiltration (Amicon YM 2), and dialysed against the appropriate buffer.

Isotopic labelling

Uniformly ¹⁵N-labelled lipoyl domain was obtained by growing bacteria on medium with ¹⁵NH₄Cl as the sole nitrogen source. The medium contained 0.5 g/L ¹⁵NH₄Cl, 4 g/L glucose, 6 g/L Na₂HPO₄, 3 g/L KH₂PO₄, 0.5 g/L NaCl, 2 mM MgSO₄, 0.1 mM CaCl₂, 10 μ M FeCl₃, and 5 mg/L thiamin. The growth of bacteria, the isolation, and

the purification of the ^{15}N -labelled lipoyl domain followed the protocol described above, except that the bacteria were grown for 40 h.

Characterisation of the isolated lipoyl domain

Polyacrylamide gel electrophoresis was performed according to Schägger & von Jagow (1987) in the presence or absence of SDS with resolving gel 16.5% T, 3% C, and stacking gel 4% T, 3% C.

Samples of the isolated lipoylated and unlipoylated forms of the domains were run on Phastgel IEF 3-9 gels as recommended by Pharmacia-LKB.

Reductive acetylation of active lipoyl domain by the *A. vinelandii* PDHC in the presence of $[2\text{-}^{14}\text{C}]$ pyruvate was assayed in a similar manner as described by Packman *et al.* (1984). Purified lipoyl domain (12 nmol) was incubated at 4°C for 5 min in the presence of *A. vinelandii* PDHC (0.014 nmol) in 100 μL 50 mM potassium phosphate pH 7.0, containing 0.2 mM thiamin diphosphate, 2 mM MgCl_2 and 3 mM NAD^+ . After addition of 5 μL 5 mM sodium $[2\text{-}^{14}\text{C}]$ pyruvate (31.7 Ci/mol), the mixture was left for 25 min at 4°C , followed by addition of 100 μg of bovine serum albumin as carrier protein. The protein was precipitated by immediate addition of 2 mL 10% (mass/vol.) ice-cold trichloroacetic acid. This suspension was kept on ice for 5 min, and then the precipitate was collected on two stacked Whatman GF/C filters. The filters were washed with 10 mL ice-cold 10% (mass/vol.) trichloroacetic acid followed by 3 mL of ice-cold acetone, dried under vacuum and the radioactivity measured.

Electrospray mass spectrometry of the lipoyl domain in 80% (by vol.) aqueous methanol containing 1% (by vol.) acetic acid was performed with a Finnigan MAT 900 double focused mass spectrometer equipped with a Finnigan MAT electrospray interface.

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

NMR spectroscopy

The lipoyl domain was dialysed against 50 mM potassium phosphate pH 5.5 containing 150 mM potassium chloride and 0.02% NaN_3 , and 10% (by vol.) D_2O was added for the lock signal. The D_2O sample was prepared by twice lyophilising the dialysed protein sample and dissolving it in D_2O . The lipoylated form of the domain was treated likewise except that the potassium chloride was omitted from the dialysis buffer to prevent rapid aggregation. The final protein concentration was 7 mM. The concentration of the ^{15}N -labelled unlipoylated lipoyl domain was 5 mM.

NMR spectra were routinely recorded on a Bruker AMX500 spectrometer operating at a proton frequency of 500.13 MHz and a ^{15}N frequency of 50.68 MHz, at 14, 30 and 37°C. Several homonuclear ^1H -NMR spectra were recorded at 41°C on a Bruker AM600 spectrometer.

Two-dimensional ^1H -NMR spectra were collected with the carrier frequency coinciding with the water resonance. The following homonuclear 2D experiments were performed: double-quantum filtered correlation spectroscopy (DQF-COSY) (Rance *et al.*, 1983), triple-quantum filtered correlation spectroscopy (TQF-COSY) (Piantini *et al.*, 1982; Shaka & Freeman, 1983), total correlation spectroscopy (TOCSY) using a clean-MLEV17 sequence for spin-locking (Griesinger *et al.*, 1988) with a mixing time of 70 ms, NOESY (Jeener *et al.*, 1979; Kumar *et al.*, 1980) with mixing times of 25, 50, 100 and 150 ms. Water protons were irradiated for 1.5 s during the relaxation delay. In most experiments the stimulated cross peaks under bleached alphas (SCUBA) sequence (Brown *et al.*, 1988) was incorporated using a SCUBA delay of 70 ms to recover saturated resonances under the solvent peak. Time-proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983) was used in all experiments. All homonuclear 2D spectra were recorded with 512 t_1 experiments with 2048 data points, and a spectral width of 7042 Hz at 500 MHz.

^1H - ^{15}N heteronuclear single-quantum coherence (HSQC) (Bodenhausen & Ruben, 1980) spectra were recorded using 256 increments of t_1 , 2048 real data points, and spectral widths of 10 kHz in ω_2 and 1723 Hz in ω_1 at 500 MHz. ^{15}N decoupling during acquisition was achieved using the GARP (Shaka *et al.*, 1985) decoupling sequence. The residual water peak was suppressed by low-power presaturation. The ^{15}N carrier frequency was placed in the centre of the backbone amide ^{15}N spectrum at 120.9 ppm.

^1H - ^{15}N heteronuclear multiple-quantum coherence (HMQC) (Müller, 1979; Bax *et al.*, 1983) experiments combined with NOESY (mixing time 150 ms) and clean-TOCSY (mixing time 80 ms) with an inserted SCUBA delay, were recorded using identical spectral widths and decoupling as used in the HSQC experiments.

NMR data were processed on a Bruker X32 data station using the UXNMR software. The free induction decays were zero-filled and multiplied by the appropriately matched sine-bell functions prior to Fourier transformation, followed by interactive phase correction and baseline correction.

The chemical shifts are reported in ppm relative to internal trimethylsilyl propionate for ^1H and to external liquid NH_3 for ^{15}N (Live *et al.*, 1984).

Backbone $^3J_{\text{HN}\alpha}$ coupling constants were measured from a DQF-COSY spectrum in H_2O zero-filled to 1k x 4k data points (digital resolution in ω_2 1.7 Hz/point).

Acknowledgements

We thank Dr. Wilfried Niessen (Gorlaeus Laboratories, Leiden University, The Netherlands) for performing the electrospray mass spectrometry, and Luuk van Langen for assistance with the isolation of the ^{15}N -labelled lipoyl domain and the hydrogen exchange experiments. Several NMR spectra were recorded at the Dutch National Hf-NMR Facility (Nijmegen, The Netherlands), which is supported by the Netherlands Foundation for Chemical Research (SON). This work was financially supported by the Netherlands Organisation for Scientific Research (NWO) under the auspices of the Netherlands Foundation for Chemical Research (SON).

Note: a value of 4.1 ppm needs to be added to the ^{15}N chemical shifts in Figures 2, 3 and 4.

REFERENCES

- Ali, S. T., Moir, A. J. G., Ashton, P. R., Engel, P. C. & Guest, J. R. (1990) Octanoylation of the lipoyl domains of the pyruvate dehydrogenase complex in a lipoyl-deficient strain of *Escherichia coli*. *Mol. Microbiol.* **4**, 943-950.
- Bax, A., Griffey, R. H. & Hawkins, B. L. (1983) Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Reson.* **55**, 301-315.
- Bodenhausen, G. & Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185-189.
- Braunschweiler, L., Bodenhausen, G. & Ernst, R. R. (1983) Analysis of networks of coupled spins by multiple quantum N.M.R. *Mol. Phys.* **48**, 535-560.
- Brookfield, D. E., Green, J., Ali, S. T., Machado, R. S. & Guest, J. R. (1991) Evidence for two protein-lipoylation activities in *Escherichia coli*. *FEBS Lett.* **295**, 13-16.
- Brown, S. C., Weber, P. L. & Müller, L. (1988) Toward complete ^1H NMR spectra in proteins. *J. Magn. Reson.* **77**, 166-169.
- Dardel, F., Packman, L. C. & Perham, R. N. (1990) Expression in *Escherichia coli* of a sub-gene encoding the lipoyl domain of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *FEBS Lett.* **264**, 206-210.
- Dardel, F., Laue, E. D. & Perham, R. N. (1991) Sequence-specific ^1H -NMR assignments and secondary structure of the lipoyl domain of the *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *Eur. J. Biochem.* **201**, 203-209.
- Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- Gibson, T. J. (1984) *Studies on the Epstein-Bar virus genome*. Ph.D. Thesis, University of Cambridge, UK.
- Goa, J. (1953) A micro biuret method for protein determination: determination of total protein in cerebrospinal fluid. *Scand. J. Clin. Lab. Invest.* **5**, 218-222.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.
- Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988) Clean TOCSY for ^1H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870-7872.
- Gronenborn, A. M., Bax, A., Wingfield, P. T. & Clore, G. M. (1989) A powerful method of sequential proton resonance assignment in proteins using relayed ^{15}N - ^1H multiple quantum coherence spectroscopy. *FEBS Lett.* **243**, 93-98.

- Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987) The domain structure of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **169**, 245-252.
- Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988) The dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Molecular cloning and sequence analysis. *Eur. J. Biochem.* **172**, 593-599.
- Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546-4553.
- Kalia, Y. N., Brocklehurst, S. M., Hippos, D. S., Appella, E., Sakaguchi, K. & Perham, R. N. (1993) The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *J. Mol. Biol.* **230**, 323-341.
- Kumar, A., Ernst, R. R. & Wüthrich, K. (1980) A two-dimensional nuclear overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1-6.
- Live, D. H., Davis, D. G., Agosta, W. C. & Cowburn, D. (1984) Long range hydrogen bond mediated effects in peptides: ^{15}N NMR study of gramicidin S in water and organic solvents. *J. Am. Chem. Soc.* **106**, 1939-1941.
- Marion, D. & Wüthrich, K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurement of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967-974.
- Mattevi, A., Schierbeck, A. J. & Hol, W. G. J. (1991) Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 Å resolution. A comparison with the structure of glutathione reductase. *J. Mol. Biol.* **220**, 975-994.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992a) The pyruvate dehydrogenase multienzyme complex. *Curr. Opin. Struct. Biol.* **2**, 877-887.
- Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1992b) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
- Mattevi, A., Obmolova, G., Kalk, K. H., Sokatch, J., Betzel, C. H. & Hol, W. G. J. (1992c) The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD^+ at 2.45 Å resolution. *Proteins* **13**, 336-351.
- Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1993a) Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Å resolution. *J. Mol. Biol.* **230**, 1183-1199.
- Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. H. & Hol, W. G. J. (1993b) 3-Dimensional structure of lipoamide dehydrogenase from *Pseudomonas fluorescens* at 2.8 Å resolution - analysis of redox and thermostability properties. *J. Mol. Biol.* **230**, 1200-1215.
- Müller, L. (1979) Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J. Am. Chem. Soc.* **101**, 4481-4484.
- Packman, L. C., Perham, R. N. & Roberts, G. C. K. (1984) Domain structure and ^1H -n.m.r. spectroscopy of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *Biochem. J.* **217**, 219-227.
- Packman, L. C., Green, B. & Perham, R. N. (1991) Lipoylation of the E2 components of the 2-oxo acid dehydrogenase multienzyme complexes of *Escherichia coli*. *Biochem. J.* **277**, 153-158.
- Perham, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.
- Piantini, U., Sørensen, O. W. & Ernst, R. R. (1982) Multiple quantum filters for elucidating NMR coupling networks. *J. Am. Chem. Soc.* **104**, 6800-6801.
- Quinn, J., Diamond, A. G., Masters, A. K., Brookfield, D. E., Wallis, N. G. & Yeaman, S. J. (1993) Expression and lipoylation in *Escherichia coli* of the inner lipoyl domain of the E2-component of the human pyruvate dehydrogenase complex. *Biochem. J.* **289**, 81-85.
- Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983) Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479-485.
- Reed, L. J., Koike, M., Levitch, M. E. & Leach, F. R. (1958a) Studies on the nature and reactions of protein-bound lipoic acid. *J. Biol. Chem.* **232**, 143-158.
- Reed, L. J., Leach, F. R. & Koike, M. (1958b) Studies on a lipoic acid-activating system. *J. Biol. Chem.* **232**, 123-142.
- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.

- Richardson, J. S. (1981) The anatomy and taxonomy of protein structures. *Adv. Protein. Chem.* **34**, 167-339.
- Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K. & Gronenborn, A. M. (1992) Three-dimensional solution structure of the E3-binding domain of the dihydrolipoamide succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **31**, 3463-3471.
- Schägger, H. & von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379.
- Schierbeek, A. J., Swarte, M. B. A., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G. J., Drenth, J. & Betzel, C. (1989) X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *J. Mol. Biol.* **206**, 365-379.
- Schulze, E., Westphal, A. H., Boumans, H. & de Kok, A. (1991a) Site-directed mutagenesis of the dihydrolipoyl transacetylase component (E2p) of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Binding of the peripheral components E1p and E3. *Eur. J. Biochem.* **202**, 841-848.
- Schulze, E., Westphal, A. H., Obmolova, G., Mattevi, A., Hol, W. G. J. & de Kok, A. (1991b) The catalytic domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* and *Escherichia coli*. Expression, purification, properties and preliminary X-ray analysis. *Eur. J. Biochem.* **201**, 561-568.
- Schulze, E., Westphal, A. H., Veeger, C. & de Kok, A. (1992) Reconstitution of pyruvate dehydrogenase multienzyme complexes based on chimeric core structures from *Azotobacter vinelandii* and *Escherichia coli*. *Eur. J. Biochem.* **206**, 427-435.
- Shaka, A. J. & Freeman, R. (1983) Simplification of NMR spectra by filtration through multiple-quantum coherence. *J. Magn. Reson.* **51**, 169-173.
- Shaka, A. J., Barker, P. B. & Freeman, R. (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.* **64**, 547-552.
- Westphal, A. H. & de Kok, A. (1990) The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur. J. Biochem.* **187**, 235-239.
- Wishart, D. S., Sykes, B. D. & Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein structure. *J. Mol. Biol.* **222**, 311-333.
- Wüthrich, K. (1976) *NMR in biological research: peptides and proteins*. North-Holland Publishing Company, Amsterdam.
- Wüthrich, K., Billeter, M. & Braun, W. (1984) Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J. Mol. Biol.* **180**, 715-740.
- Wüthrich, K. (1986) *NMR of proteins and nucleic acids*. Wiley, New York.

CHAPTER 3

Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex

ABSTRACT

A 79-amino-acid polypeptide, corresponding to the lipoyl domain of the succinyltransferase component of the 2-oxoglutarate dehydrogenase multienzyme complex from *Azotobacter vinelandii*, has been subcloned and produced in *Escherichia coli*. Complete sequential ^1H and ^{15}N resonance assignments for the lipoyl domain have been obtained by using homo- and hetero-nuclear NMR spectroscopy. Two antiparallel β -sheets of four strands each were identified from characteristic NOE connectivities and $^3J_{\text{HN}\alpha}$ values. The lipoyl-lysine residue is found in a type-I turn connecting two β -strands. The secondary structure of the lipoyl domain very much resembles the secondary solution structure of the N-terminal lipoyl domain of the *A. vinelandii* pyruvate dehydrogenase complex, despite the sequence identity of 25%. A detailed comparison of the NMR-derived parameters of both lipoyl domains, i.e. chemical shifts, NH-exchange rates, NOEs, and $^3J_{\text{HN}\alpha}$ values suggests a high structural similarity in solution between the two lipoyl domains. Preliminary tertiary-structure calculations confirm that these lipoyl domains have very similar overall folds. The observed specificity of the 2-oxo acid dehydrogenase components of both complexes for these lipoyl domains is discussed in this respect.

INTRODUCTION

The 2-oxo acid dehydrogenase multienzyme complexes catalyse the irreversible oxidative decarboxylation of 2-oxo acids to acyl-CoA [for recent reviews see Perham (1991) and Mattevi *et al.* (1992a)]. In *Azotobacter vinelandii* two members of this family of multienzyme complexes are present, the pyruvate dehydrogenase complex (PDHC) and the 2-oxoglutarate dehydrogenase complex (OGDHC). These complexes have a very similar design and share many structural and catalytic properties. The complexes are composed of multiple copies of three enzymes: a substrate-specific 2-oxo acid dehydrogenase [pyruvate dehydrogenase (E1p) or 2-oxoglutarate dehydrogenase (E1o)], an acyltransferase [acetyltransferase (E2p) or succinyltransferase (E2o)], and a common lipoamide dehydrogenase (E3). The E2 component forms the oligomeric cubic core of the complex to which the peripheral subunits E1 and E3 are bound as dimers. In the OGDHC 12 dimers of E1o and 6 dimers of E3 are bound to a core of E2o consisting of 24 subunits. The PDHC from *A. vinelandii* consists of a trimeric core of E2p (Mattevi *et al.*, 1992b) to which two dimers of E1p and one dimer of E3 are bound (Schulze *et al.*, 1992). Upon removal of the peripheral E1p and E3 components the trimer aggregates to the 24-subunit core structure. This feature is unique for the PDHC from *A. vinelandii*.

The acyltransferase monomers are highly segmented and consist of three types of separate and independently folded domains (Hanemaaijer *et al.*, 1988). They contain at the N-terminal part one (OGDHC) or three (PDHC) lipoyl domains of about 80 residues each, all carrying a liponic acid prosthetic group bound to a specific lysine residue. This lipoyl group visits the three active sites in the multienzyme complex and is responsible for the transfer of substrate and reduction equivalents among them (Reed, 1974). The lipoyl domain is followed by a peripheral subunit-binding domain of about 35 residues, which is responsible for binding of the E3 and the E1 components to the structural core. The C-terminal catalytic domain (29 kDa) catalyses the acyltransferase reaction and aggregates to form the core of the complex (Hanemaaijer *et al.*, 1987; Mattevi *et al.*, 1992b). For PDHC from *A. vinelandii* it was shown that the catalytic domain also interacts with the E1p component (Schulze *et al.*, 1991a, 1992). The domains are linked by conformationally flexible sequences of 20 to 40 amino acid residues with an unusual high content of Ala and Pro residues.

Recently our knowledge and understanding of the 2-oxo acid dehydrogenase complexes expanded largely since several structures of enzymes and domains of these multienzyme complexes became available. By X-ray diffraction the crystal structures of the catalytic core domain of E2p from *A. vinelandii* (Schulze *et al.*, 1991b; Mattevi *et al.*,

1992b, 1993a), and the lipoamide dehydrogenases from *A. vinelandii* (Schierbeek *et al.*, 1989; Mattevi *et al.*, 1991), *Pseudomonas putida* (Mattevi *et al.*, 1992c) and *P. fluorescens* (Mattevi *et al.*, 1993b) have been determined. The solution structures of the lipoyl domain of *Bacillus stearothermophilus* PDHC (Dardel *et al.*, 1993), a hybrid lipoyl domain of *E. coli* PDHC (Green *et al.*, 1995), the E3-binding domain of *E. coli* E2o (Robien *et al.*, 1992), the E1p/E3-binding domain of *B. stearothermophilus* E2p (Kalia *et al.*, 1993), and the secondary structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC (Berg *et al.*, 1994) have been resolved by NMR. No structural information on the E1 component is available as yet.

The specificity of the 2-oxo acid dehydrogenase complexes is determined by the E1 and E2 components. E1 decarboxylates only its specific 2-oxo acid substrate and the E2 component converts CoA exclusively to its specific acyl-CoA derivative. For efficient multienzyme catalysis the specificity at the level of the active sites of the E1 and E2 components should be sufficient. However, free lipoamide is an extremely poor substrate for the E1 component, and a folded structure of the lipoyl domain attached to lipoamide has proven to be essential for reaction with E1 (Graham *et al.*, 1989), in contrast to the E2 and E3 components which can use free lipoamide as a substrate (Reed *et al.*, 1958). Moreover, lipoyl domains can only be fully reductively acylated by the E1 component of their parent complex (Graham *et al.*, 1989; chapter 6, this thesis), and are therefore involved in specificity as well. Thus, the E1 components are not only specific for their 2-oxo acid substrate but also for their lipoyl domains. Specificity of E1p for lipoyl domains was also demonstrated by reconstitution experiments of pyruvate dehydrogenase complexes based on chimeric core structures from *A. vinelandii* and *E. coli* (Schulze *et al.*, 1992). It was shown that the reaction of lipoyl domains with E1p of different origin caused a decrease in overall complex activity.

In this paper we present complete sequence-specific ^1H and ^{15}N resonance assignments and secondary structure of the *A. vinelandii* OGDHC lipoyl domain, purified from a subclone of the succinyltransferase gene in *E. coli*. The secondary structure of this lipoyl domain, characterised via NOE connectivities, slowly exchanging amide protons, and vicinal coupling constants, provides the first structural information of a lipoyl domain of a 2-oxoglutarate dehydrogenase complex. The secondary structure obtained has been compared in detail with the solution secondary structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC. Preliminary tertiary-structure models of the *A. vinelandii* PDHC and OGDHC lipoyl domains, based on a limited number of constraints, are presented.

By determining the solution structure of the lipoyl domain of E2o of *A. vinelandii* OGDHC we expect to gain more insight in the process of molecular recognition of lipoyl

domains by E1, and thereby unravel what is the determinant of the lipoyl domain in this process. In particular, comparison of structures of lipoyl domains from different acyltransferases from the same source, in our case *A. vinelandii*, is a unique way to tackle this interesting problem. In addition, the interactions of the different lipoyl domains with the different 2-oxo acid dehydrogenase complexes from *A. vinelandii* can be studied by NMR and other techniques.

RESULTS AND DISCUSSION

Expression and isolation

Sequencing of the complete insert of plasmid pAB1, encoding the lipoyl domain (residues 1-79) of the succinyltransferase component of *A. vinelandii* OGDHC, revealed an error in the previously published DNA sequence of the succinyltransferase encoding gene (Westphal & de Kok, 1990). This error was confirmed by another examination of the original sequence data of the succinyltransferase, and was not a result of DNA manipulation. The change in the DNA sequence has consequences for the protein sequence, and in the amended amino acid sequence of succinyltransferase of *A. vinelandii* OGDHC, Pro28 is replaced by Ala28.

Induction of *E. coli* TG2(pAB1) cells with IPTG results in a high expression of lipoyl domain. The expression level is comparable to that of the N-terminal lipoyl domain of the acetyltransferase component of *A. vinelandii* PDHC in *E. coli* (Berg *et al.*, 1994). Addition of lipoic acid to the TY growth medium results in expression of lipoylated lipoyl domain, as judged by native PAGE of cell-free extracts of the induced *E. coli* cells. The unlipoylated and lipoylated forms of the lipoyl domain were purified with yields of about 7 mg lipoyl domain/L bacterial culture. Approximately 20 mg ¹⁵N-labelled unlipoylated lipoyl domain was isolated from a 6-L culture grown on minimal medium. The lipoylated lipoyl domain could be reductively succinylated by catalytic amounts of *A. vinelandii* OGDHC in the presence of [U-¹⁴C]2-oxoglutarate. This proves that the expressed lipoyl domain is correctly modified with lipoamide and that its folding is correct. In most experiments this lipoyl domain showed very similar properties as the N-terminal lipoyl domain of *A. vinelandii* PDHC.

NMR assignments

Sequential resonance assignments were obtained using standard methods (Wüthrich, 1986; Gronenborn *et al.*, 1989). Amino acid spin systems were identified based

on their characteristic through-bond connectivities using DQF-COSY and TOCSY experiments recorded in H_2O and D_2O , in combination with ^1H - ^{15}N HMQC-TOCSY experiments. Together 65 spin systems could be distinguished at this stage of the assignment procedure. Many spin systems were then tentatively assigned to individual residue types on the basis of side-chain proton chemical shifts. The AMX spin system and the spin system of the aromatic ring protons belonging to a single aromatic residue were connected through intra-residue NOEs between one or two ring protons and the C^αH and C^βH protons. The remaining spin systems were identified in the course of the sequential assignments.

Sequence-specific resonance assignments were readily obtained using 2D ^1H NOESY in combination with 2D ^1H - ^{15}N HMQC-NOESY. In this manner most residues with amide-proton chemical shift degeneracy could easily be resolved, as can be judged by the spreaded dispersion of peaks in the HSQC spectrum (Figure 1). Only residues Ala6 and Ile35 have nearly degenerate ^1H and ^{15}N chemical shifts. By using NMR spectra recorded at two different temperatures, nearly all overlapping C^αH protons could be resolved. If C^αH protons kept partially overlapping with the solvent resonance or other C^αH protons, sequential connections were certain by $d_{\beta\text{N}}$ or d_{NN} contacts. Many sequential assignments in the β -sheet regions were confirmed by interstrand long-range NOEs. Sequential assignments past the three proline residues were obtained using $d_{\alpha\delta}$ NOEs. Several methylene C^βH protons and valine methyl protons were stereospecifically assigned on basis of $^3J_{\alpha\beta}$ values and relative intra-residue NOE intensities (Zuiderweg *et al.*, 1985; Wagner *et al.*, 1987). NOE intensities were measured from the 70-ms NOESY spectra and $^3J_{\alpha\beta}$ coupling constants from the E.COSY spectrum. Most of the $^3J_{\alpha\beta}$ values obtained from the E.COSY spectrum were confirmed qualitatively from analysis of peak shapes in the clean-TOCSY spectra, as described by Driscoll *et al.* (1989). By making effective use of the combination of two-dimensional homo- and hetero-nuclear NMR experiments, and due to the relative low amount of chemical shift degeneracy of backbone protons, we did not need to turn to 3D NMR spectroscopy at this stage of the structure determination of the lipoyl domain. The ^{15}N resonances were determined according to the amide proton assignments.

In Table 1 the ^{15}N and ^1H chemical shifts of the unlipoylated lipoyl domain are summarised. Complete assignments were obtained for 67 residues, and 12 residues were partly assigned. For some Leu, Ile and Lys residues, side-chain protons past the C^βH proton(s) were observed, but these could not be assigned to a specific position in the side chain.

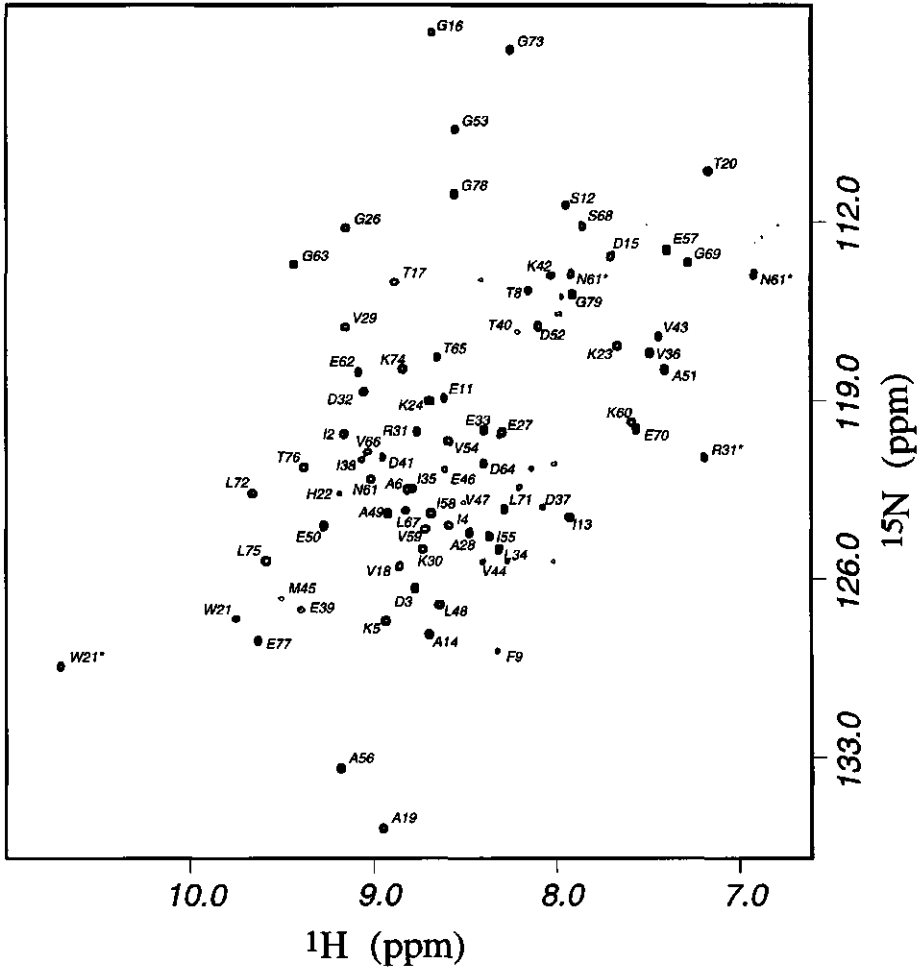


Figure 1. 2D ^1H - ^{15}N HSQC spectrum of the lipoyl domain of *A. vinelandii* OGDHC, recorded at 30°C and pH 5.5. The observable side-chain ^1H - ^{15}N cross peaks of Trp21, Arg31, and Asn61 are marked with an asterisk.

Table 1. ^1H and ^{15}N chemical shifts (ppm) for the unlipoylated form of the lipoyl domain at pH 5.5 and 30°C, in 50 mM potassium phosphate containing 100 mM potassium chloride.

Residue	Amide		C^αH	C^βH	Other
	^{15}N	^1H			
Ala1			4.50	1.47	
Ile2	120.4	9.16	4.44	1.82	$\text{C}^\gamma\text{H}_2$ 1.14, $\text{C}^\gamma\text{H}_3$ 0.92, $\text{C}^\delta\text{H}_3$ 0.73
Asp3	126.4	8.77	4.75	<u>2.32, 2.58</u>	
Ile4	124.0	8.58	4.15	1.55	0.98, 0.84, 0.72
Lys5	127.8	8.93	5.01	1.65, 1.54	C^γH 1.26, $\text{C}^\epsilon\text{H}$ 2.90
Ala6	122.6	8.81	4.17	1.36	
Pro7			4.48	2.44	C^γH 2.03, 1.86, C^δH 3.57, 3.27
Thr8	114.8	8.14	4.07	3.95	$\text{C}^\gamma\text{H}_3$ 1.13
Phe9	128.9	8.32	4.95	3.33, 3.00	C^δH 7.04, $\text{C}^\epsilon\text{H}$ 6.70, C^ζH 6.70
Pro10			4.57	2.44	C^γH 2.11, 1.98, C^δH 3.95, 3.49
Glu11	119.0	8.61	4.08	2.08	C^γH 2.36
Ser12	111.4	7.93	4.24	4.09, 3.91	
Ile13	123.7	7.92	4.27	1.92	$\text{C}^\gamma\text{H}_2$ 1.13, $\text{C}^\gamma\text{H}_3$ 0.65, $\text{C}^\delta\text{H}_3$ 0.78
Ala14	128.3	8.69	4.39	1.46	
Asp15	113.4	7.69	4.93	2.98, 2.76	
Gly16	104.7	8.66	4.71, 3.25		
Thr17	114.4	8.87	5.23	3.79	$\text{C}^\gamma\text{H}_3$ 1.07
Val18	125.6	8.85	3.65	2.42	$\text{C}^\gamma\text{H}_3$ <u>0.84, 0.70</u>
Ala19	135.9	8.95	4.62	1.38	
Thr20	110.1	7.17	4.50	3.63	$\text{C}^\gamma\text{H}_3$ 0.52
Trp21	127.6	9.75	4.66	<u>3.09, 3.52</u>	C^δH 7.14, $\text{C}^\epsilon\text{H}$ 7.92, C^ηH 6.68, C^ζH 7.63, C^ξH 6.69, $\text{N}^{\epsilon\text{H}}$ 10.70, $\text{N}^{\epsilon\text{I}}$ 129.5
His22	122.8	9.16	4.37	<u>2.89, 3.43</u>	C^δH 6.46, $\text{C}^{\epsilon\text{H}}$ 8.51
Lys23	117.0	7.66	4.91	1.74	C^γH 1.50, $\text{C}^\epsilon\text{H}$ 2.90
Lys24	119.1	8.69	4.64	1.50	1.87
Pro25			3.90	1.98, 1.79	C^γH 2.14, C^δH 3.50, 3.42
Gly26	112.3	9.15	4.36, 3.53		
Glu27	120.4	8.29	4.47	2.09	C^γH 2.40, 2.25
Ala28	124.3	8.46	4.92	1.47	
Val29	116.2	9.15	4.71	1.82	$\text{C}^\gamma\text{H}_3$ 0.84
Lys30	124.9	8.73	4.97	1.80, 1.58	C^γH 1.44, $\text{C}^\epsilon\text{H}$ 3.00

Table 1. (Continued)

Residue	Amide		C α H	C β H	Other
	^{15}N	^1H			
Arg31	120.3	8.75	3.24	1.64, 1.49	C γ H 1.49, 1.33, C δ H 3.17, N ϵ H 7.18; N ϵ 84.6
Asp32	118.8	9.04	4.22	<u>2.91, 3.23</u>	
Glu33	120.2	8.39	4.24	2.17, 2.03	C γ H 2.39, 2.29
Leu34	124.9	8.30	4.06	1.54, 1.22	C γ H 0.76, C δ H $_3$ 0.35
Ile35	122.6	8.78	4.37	1.66	C γ H 1.07, 0.85
Val36	117.2	7.48	5.04	1.10	C γ H $_3$ <u>-0.24, 0.12</u>
Asp37	123.2	8.08	5.44	<u>2.63, 2.51</u>	
Ile38	121.4	9.04	4.63	1.83	C γ H 1.16, 0.72
Glu39	127.3	9.38	4.76	2.03, 1.92	C γ H 2.15
Thr40	116.4	8.20	4.80	4.14	C γ H $_3$ 0.72
Asp41	121.3	8.95	4.31	2.85, 2.62	
Lys42	114.2	8.02	4.32	1.80, 1.93	C γ H 1.42, C δ H 1.66, C ϵ H 2.99
Val43	116.6	7.43	4.46	2.11	C γ H $_3$ <u>0.90, 0.81</u>
Val44	125.4	8.40	4.64	1.94	C γ H $_3$ 0.94, 0.80
Met45	126.8	9.50	4.64	<u>1.82, 2.07</u>	C γ H 2.44
Glu46	121.8	8.60	4.49	1.92, 1.78	C γ H 2.28
Val47	123.1	8.50	4.17	2.10	C γ H $_3$ 0.74
Leu48	127.1	8.63	4.99	1.69, 1.07	0.88, 0.68
Ala49	123.5	8.92	4.21	1.54	
Glu50	124.0	9.27	4.23	2.13, 1.98	C γ H 2.42, 2.28
Ala51	117.9	7.40	4.32	1.28	
Asp52	116.1	8.09	4.88	<u>2.62, 2.80</u>	
Gly53	108.5	8.55	4.38, 4.13		
Val54	120.7	8.58	4.89	1.73	C γ H $_3$ <u>0.83, 0.72</u>
Ile55	124.4	8.36	3.95	2.28	C γ H $_2$ 1.56, 1.38, C γ H $_3$ 0.58, C δ H $_3$ 0.79
Ala56	133.5	9.17	4.51	1.38	
Glu57	113.2	7.39	4.37	1.82	C γ H 1.98
Ile58	123.5	8.68	3.91	1.19	C γ H $_2$ -0.82, -0.13, C γ H $_3$ 0.13, C δ H $_3$ 0.62
Val59	124.1	8.71	4.29	1.86	C γ H $_3$ <u>0.73, 0.89</u>
Lys60	120.0	7.58	4.94	2.11, 1.68	
Asn61	122.2	9.01	4.68	<u>2.61, 2.87</u>	N δ H 6.91, 7.91, N δ 114.1
Glu62	118.0	9.08	3.79	2.04	C γ H 2.72, 2.40
Gly63	113.8	9.43	4.50, 3.61		

Table 1. (Continued)

Residue	Amide		C^αH	C^βH	Other
	^{15}N	^1H			
Asp64	121.6	8.39	4.81	<u>2.73</u> , <u>3.05</u>	
Thr65	117.4	8.64	4.97	4.07	$\text{C}^\gamma\text{H}_3$ 1.31
Val66	121.1	9.03	4.79	1.95	$\text{C}^\gamma\text{H}_3$ 0.88, 0.86
Leu67	123.4	8.83	4.77	1.61	C^γH 1.79, $\text{C}^\delta\text{H}_3$ 1.04, 0.95
Ser68	112.3	7.85	2.87	3.66, 3.57	
Gly69	113.7	7.27	3.80		
Glu70	120.2	7.56	4.01	<u>1.84</u> , <u>2.16</u>	C^γH 2.42, 2.36
Leu71	123.4	8.27	4.19	1.88, 1.59	C^γH 1.47, $\text{C}^\delta\text{H}_3$ 0.91
Leu72	122.7	9.66	4.71	<u>1.69</u> , <u>1.86</u>	C^γH 2.03, $\text{C}^\delta\text{H}_3$ 0.83, 0.63
Gly73	105.3	8.24	4.56, 4.07		
Lys74	117.9	8.83	5.40	1.71	C^γH 1.47, 1.35, C^δH 1.73, $\text{C}^\epsilon\text{H}$ 2.86, 2.80
Leu75	125.4	9.57	5.02	1.73, 1.45	$\text{C}^\delta\text{H}_3$ 0.71
Thr76	121.7	9.37	4.46	4.15	$\text{C}^\gamma\text{H}_3$ 1.23
Glu77	128.5	9.64	4.38	<u>2.00</u> , <u>2.10</u>	C^γH 2.30, 2.27
Gly78	111.0	8.55	3.97		
Gly79	114.9	7.90	3.77		

^1H chemical shifts (± 0.02 ppm) are referenced to internal trimethylsilyl propionate; ^{15}N chemical shifts (± 0.2 ppm) are referenced to external liquid ammonia. Underlined chemical shifts indicate stereospecific assignments for β -methylene protons (β^2 and β^3 respectively) and methyl groups of valines (γ^1 and γ^2 respectively). Resonances in the final column without a label belong to the spin system but their position in the side chain could not be determined.

Secondary structure

The NMR data that were used to identify elements of regular secondary structure are summarised in Figure 2. Two four-stranded antiparallel β -sheets were identified from long-range interstrand d_{NN} , $d_{\alpha\alpha}$, and $d_{\text{N}\alpha}$ NOEs, together with data on slowly exchanging amide protons and $^3J_{\text{HN}\alpha}$ values. The strands are defined as follows: S1(Ala1-Ala6), S2(Gly16-Trp21), S3(Glu27-Lys30), S4(Ile35-Thr40), S5(Val44-Leu48), S6(Gly53-Val59), S7(Asp64-Leu67), S8(Leu71-Thr76). The secondary structure of the lipoyl domain is depicted in Figure 3. The observed interstrand and some intrastrand NOEs are shown for residues in the β -sheet regions. It should be noted that the absence of some contacts could be caused by overlap. We were unable to observe some interstrand $d_{\alpha\alpha}$ NOEs because of C^αH chemical shift degeneracy resulting in NOEs too close to the diagonal. The proposed main-chain hydrogen bonds based on slowly exchanging amide protons are also indicated

in Figure 3. The pairing of the hydrogen bond donors and acceptors is of course tentative. For the majority of the residues in the β -sheet regions the value of the coupling constant ${}^3J_{\text{HN}\alpha}$ is high. A type-I turn is suggested at the corner of the neighbouring strands S4 and S5, from Thr40 to Val43. The type-I turn is characterised by the strong $d_{\text{NN}}(42,43)$ connectivity, the medium $d_{\text{NN}}(41,42)$ connectivity and the small ${}^3J_{\text{HN}\alpha}$ value of Asp41. In this turn the lipoyl-lysine residue Lys42 is situated. The N-terminus of the lipoyl domain is close to its C-terminus in the other sheet. A large loop region comprising nine residues connects β -strand S1 with strand S2. A β -bulge is found in strand S8 around Leu71 and Leu72, which are opposed by only a single amino acid (Lys5) in the neighbouring strand S1. Other irregularities in β -strands are observed around Glu57 in strand S6, Thr20 in strand S2, and Ile35 in strand S4. These residues show all unexpected d_{NN} contacts to the opposite strand.

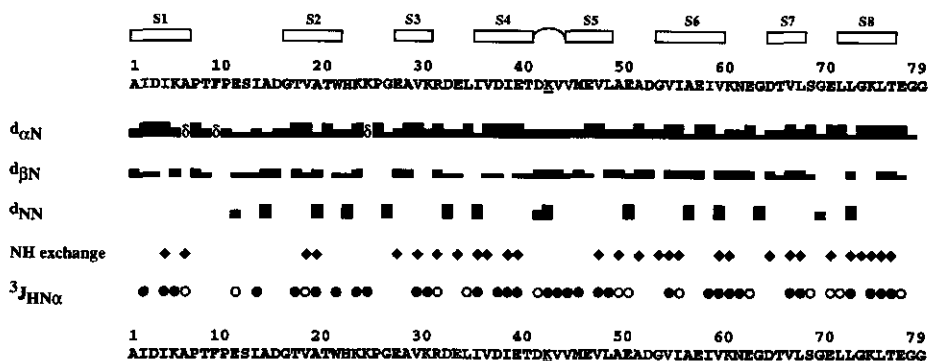


Figure 2. Summary of sequential and medium-range NOE connectivities, slowly exchanging amide protons, and ${}^3J_{\text{HN}\alpha}$ coupling constants observed for the lipoyl domain. In the lines showing the amino acid sequence, the lipoyl-lysine residue is underlined. NOE intensities were derived from NOESY spectra with a mixing time of 70 ms. The thickness of the bars reflects a qualitative measure of the strength of the NOEs, classified as strong, medium or weak. Sequential $d_{\alpha\delta}$ connections involving a proline residue at position $i + 1$ are indicated by δ . Amide protons that could still be observed 16 h after dissolving the protein in D_2O , at pH 5.5 and 30°C , are marked with (\blacklozenge). The ${}^3J_{\text{HN}\alpha}$ coupling constants are classified as (O) < 6 Hz, and (\bullet) > 8 Hz. In the upper line the regular secondary structure elements, as derived from NMR data, are denoted.

The defined structure of the lipoyl domain is likely to end near Gly78. The backbone proton resonances from Gly79 are significantly more narrow than the signals from other residues, and Gly78 and Gly79 show no medium or long-range NOEs. It is therefore concluded that at least Gly79 is not involved in the global fold of the lipoyl

domain, and is a residue of the linker sequence connecting the lipoyl domain with the binding domain.

The assignment of the lipoyl domain was facilitated by the presence of two aromatic residues in the sequence. In particular, Trp21 shows many long-range NOEs to various residues with hydrophobic side chains, including Val18, Val36, Ile58, and Leu72.

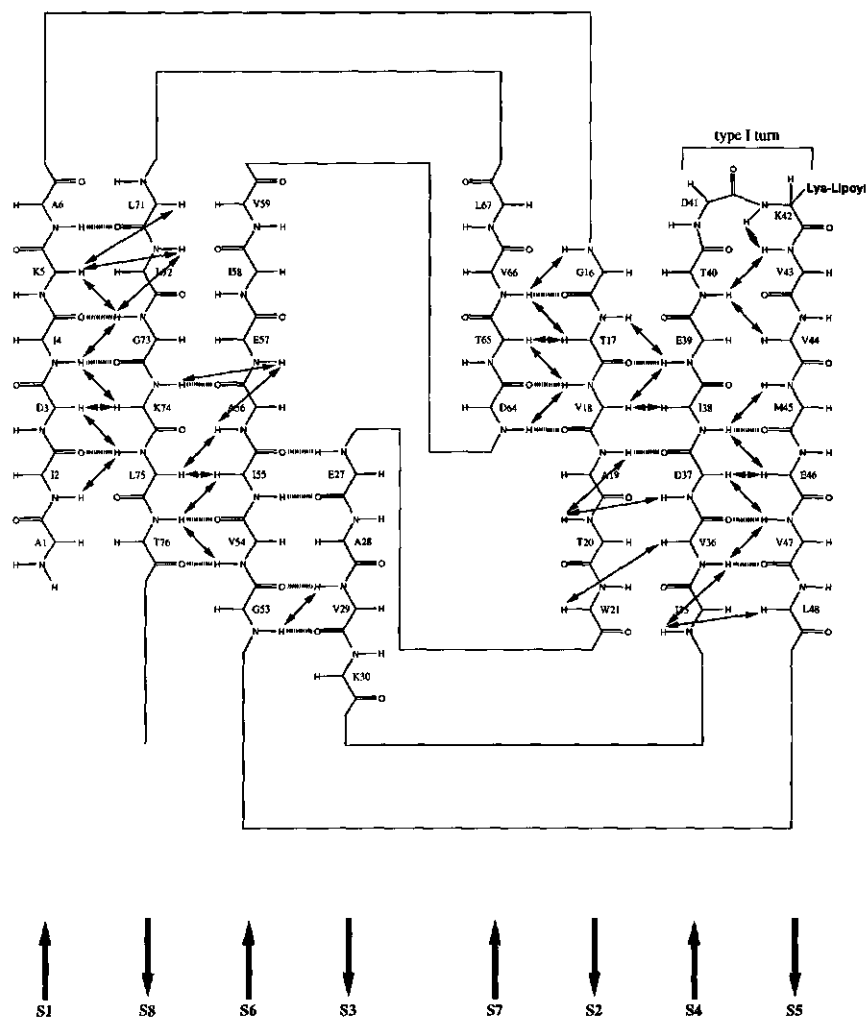


Figure 3. Schematic drawing of the secondary structure of the lipoyl domain in solution. Arrows denote observed medium- and long-range medium and strong NOE connectivities between backbone protons. Hydrogen bonds are indicated by hatched lines, and are drawn on the basis of the observation of slowly exchanging amide protons.

Two of these residues, i.e. Ile58 and Leu72, are situated in the opposite β -sheet as where Trp21 is positioned. It is therefore suggested that the two β -sheets are formed around a hydrophobic core with a central position of Trp21. The suggested overall fold of the lipoyl domain is in agreement with a calculated preliminary tertiary-structure model, based on a limited number of NOE-distance and torsion-angle constraints (see below).

Comparison of *A. vinelandii* OGDHC and PDHC lipoyl domain

The NMR-derived parameters and secondary structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC have been determined earlier in our laboratory (Berg *et al.*, 1994). The NMR experiments with both lipoyl domains have been performed under nearly identical conditions with respect to pH, salt concentration, and temperature. This allows a detailed comparison with the properties of the *A. vinelandii* OGDHC lipoyl domain described in this paper. Despite the sequence identity of only 25% (Figure 4), the secondary structure and overall fold of both lipoyl domains, as will be discussed below, prove to be very similar. Comparison of the observed NMR-derived parameters, i.e. NOEs, NH-exchange rates, vicinal coupling constants, and chemical shifts will provide information about local differences in the solution structures of the two lipoyl domains.

Matching both lipoyl domains reveals that the elements of the regular secondary structure and their positions in the amino acid sequence are very similar. They consist of two four-stranded antiparallel β -sheets, with an identical pairing of the strands and approximately the same lengths of the extended strands. The lipoyl-lysine residue is located at an identical position in the single type-I turn connecting strands S4 and S5. This is an important feature since, although exposed to solvent, the exact position of the lipoyl-lysine residue in this structure element is very likely required for lipoylation and reaction with E1, as has been demonstrated for the *B. stearrowthermophilus* PDHC lipoyl domain (Wallis & Perham, 1994). Furthermore, all the observed irregularities in the β -sheets occur at identical positions, i.e. around Thr20, Ile35, Glu57, and Leu72 for the OGDHC lipoyl domain, and at the corresponding positions in the PDHC lipoyl domain. The major difference in secondary structure between both lipoyl domains is a gap of three residues in the PDHC lipoyl domain sequence at position 11 to 14 in the OGDHC lipoyl domain sequence, resulting in a shortened loop between strand S1 and S2.

The pattern of long-range NOEs connecting the different β -strands for the OGDHC lipoyl domain, as shown in Figure 3, is very similar to that found previously for the PDHC lipoyl domain. Observed differences in this pattern may not reflect true differences in the NOE pattern because of chemical shift degeneracy. For example, the C $^{\alpha}$ H protons of Ala1 and Thr76, and of Glu39 and Val44 of the OGDHC lipoyl domain would be expected to

give an interstrand $d_{\alpha\alpha}$ NOE as is the case for the PDHC lipoyl domain, but they do not as their chemical shifts are nearly identical. For the same reason several interstrand long-range $d_{\text{N}\alpha}$ contacts are observed only for the OGDHC or for the PDHC lipoyl domain. There are, however, certain true differences in long-range NOEs between the two domains. The $d_{\alpha\alpha}$ contact between Leu67 and Asp15, and the $d_{\text{N}\alpha}$ between Ser68 and Asp15 are not observed for the OGDHC lipoyl domain, whereas these contacts are present at the equivalent positions in the PDHC lipoyl domain structure. Likewise, $d_{\text{N}\alpha}$ NOEs between Glu57 and Gly73, and between Thr17 and Thr40 are not observed for the OGDHC lipoyl domain. The $d_{\alpha\alpha}$ NOE between Lys64 and Asp12, and also the $d_{\text{N}\alpha}$ NOE between Glu65 and Asp12 in the PDHC lipoyl domain are however only weak.

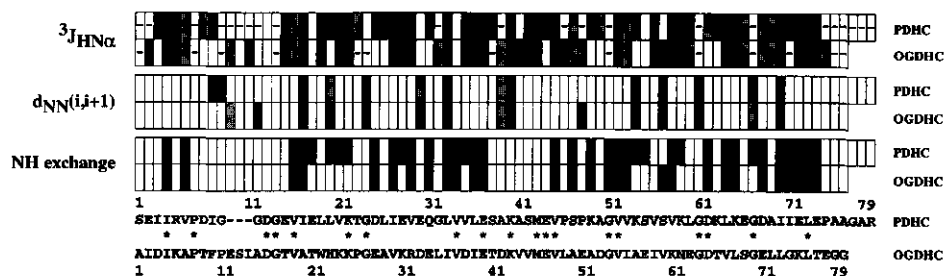


Figure 4. Comparison of NMR-derived parameters between the lipoyl domains of PDHC and OGDHC from *A. vinelandii*, including $^3J_{\text{HN}\alpha}$ coupling constants, strong and medium $d_{\text{NN}(i,i+1)}$ NOEs, and slowly exchanging amide protons. In the upper bar, large $^3J_{\text{HN}\alpha}$ values (> 8 Hz) are indicated by a black filled box, small $^3J_{\text{HN}\alpha}$ values (< 6 Hz) by a grey filled box, and $^3J_{\text{HN}\alpha}$ values between 6 and 8 Hz by a blank box. Residues for which the cross-peak intensity was too weak to measure the coupling are indicated by (-), as are Pro and Gly residues. In the second bar, strong sequential d_{NN} contacts are indicated by a black filled box and medium sequential d_{NN} contacts by a grey filled box. In the lower bar, slowly exchanging amide protons are indicated by a black filled box. The upper half of each bar indicates the parameters for the N-terminal lipoyl domain of PDHC, the lower half of each bar the parameters for the lipoyl domain of OGDHC. At the bottom the sequences of both lipoyl domains are compared. Identical residues are indicated by asterisks.

Strong sequential d_{NN} NOEs, when not occurring successively as in α -helices, are indicative of strong bending of the backbone. A comparison of the occurrence of strong and medium sequential d_{NN} connectivities for the two lipoyl domains shows a very similar pattern (Figure 4). The majority of these contacts are observed at identical positions in the backbone of the two domains, and only in a few cases a relatively small difference in NOE intensity (from strong to medium) is observed. A distinct difference in sequential d_{NN} contacts is observed however in the loop region connecting strand S1 with strand S2, suggesting a rather large difference in backbone conformation in this loop between the two

lipoyl domains. As will be discussed below, this observation supports the hypothesis that differences in this loop are responsible for the specificity of the 2-oxo acid dehydrogenase component for their lipoyl domains. The only other difference between the two domains with respect to the occurrence of sequential d_{NN} contacts is the connectivity observed for Glu50 of the OGDHC lipoyl domain. In the PDHC lipoyl domain a proline is present at this position, which explains the absence of this contact here.

The hydrogen exchange rates of the lipoyl domains were measured from ^1H - ^{15}N HSQC spectra of the ^{15}N -labelled domains in D_2O . An amide hydrogen was designated to exchange slowly with the solvent if it could still be detected after 16 h at 30°C , pH 5.5. The slowly exchanging amide protons of the two domains are compared in Figure 4. The pattern of hydrogen exchange is also similar. In the PDHC lipoyl domain several amide protons exchange somewhat slower with the solvent than in the OGDHC lipoyl domain. These differences however are rather small. In most cases the differences in exchange rate imply a change from a slowly exchanging proton in the PDHC lipoyl domain to a medium exchanging (8.5 h) amide proton in the OGDHC lipoyl domain (data not shown). The exceptions are the fast exchanging amide protons of His22 to Lys24 of the OGDHC lipoyl domain compared to the slowly exchanging amide protons of Leu19 to Lys21 of the PDHC lipoyl domain, and the reversed situation for Val59 and Ser56 of the OGDHC and PDHC lipoyl domain respectively. The amino acid residues that show the large differences in hydrogen exchange rate occur at the edges of regular β -sheet structure or in regions connecting different β -strands. A large difference in hydrogen exchange rate may reflect a significant difference in local backbone conformation, caused by a difference in backbone-backbone hydrogen bonding. It can not be ruled out however, that hydrogen bonding between the backbone amide proton and a certain side chain cause these differences, and that the effect on the backbone conformation is less drastic.

Vicinal coupling constants ($^3J_{\text{HN}\alpha}$) can be translated into ϕ torsion angles by means of the Karplus relationship (Karplus, 1963), and provide therefore very useful information about local backbone conformation. The $^3J_{\text{HN}\alpha}$ values of the two lipoyl domains were determined from fitting t_1 vectors of the HMQC- J spectra with respect to the coupling constant and line width. This yields reasonably accurate $^3J_{\text{HN}\alpha}$ values with a precision of about 0.5 Hz for large couplings, and about 1-1.5 Hz for small couplings. The determined values of the vicinal coupling constants are compared in Figure 4 in a qualitative manner, i.e. they are divided into three categories: larger than 8 Hz, between 6 and 8 Hz, and smaller than 6 Hz. Again, the pattern of small and large $^3J_{\text{HN}\alpha}$ values is very similar, indicating a similar backbone conformation of the two lipoyl domains. Only relative small

differences in vicinal coupling constants are observed, and no changes to an opposite category can be found (i.e. from large to small $^3J_{\text{HN}\alpha}$ values and *visa versa*).

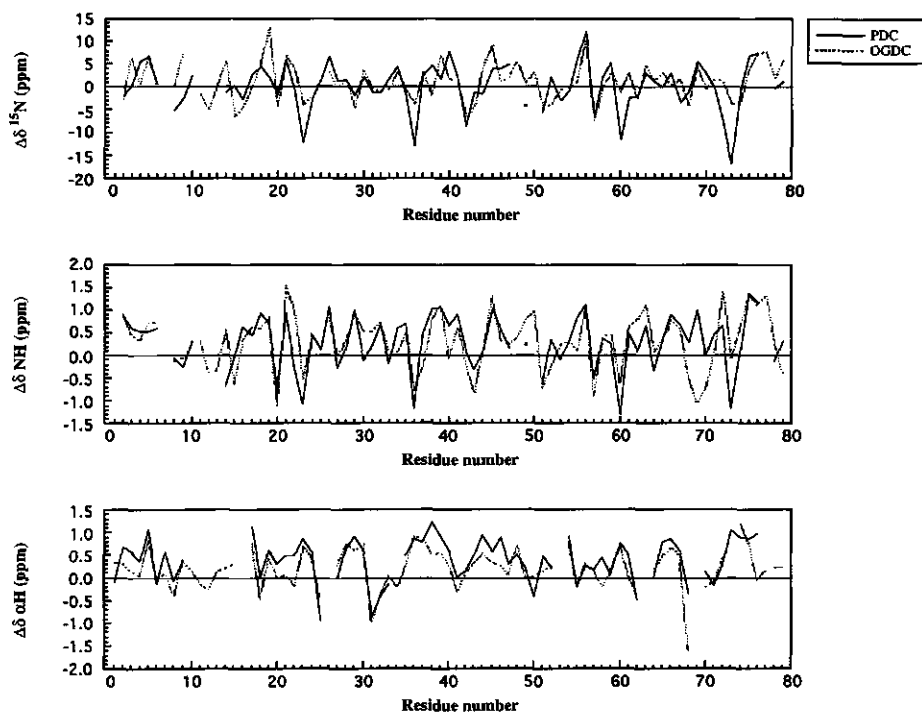


Figure 5. Comparison of the secondary shifts for the ^{15}N atoms, the NH and $\text{C}^{\alpha}\text{H}$ protons between the PDHC and OGDHC lipoyl domains. The secondary shifts were calculated by subtracting random coil contributions (values from Wishart *et al.* (1991)) from the experimental chemical shifts for each residue. The straight line in each plot represents the connected secondary shifts of the N-terminal lipoyl domain of PDHC and the dotted line the connected secondary shifts of the lipoyl domain of OGDHC. The secondary shifts are plotted versus the residue numbers of the OGDHC lipoyl domain, with the exception of NH proton and ^{15}N atom chemical shifts at positions of prolines in either lipoyl domain, and with the exception of $\text{C}^{\alpha}\text{H}$ chemical shifts at positions of Gly residues in either lipoyl domain. The data points of the PDHC lipoyl domain were aligned as in (Figure 4.) The residue numbers given on the x-axis are therefore approximate for the PDHC lipoyl domain. To all experimental ^{15}N chemical shifts of the PDHC lipoyl domain a correction of 4.1 ppm was added as compared to Berg *et al.* (1994) (see correction (1994) *Eur. J. Biochem.* 223, 1079).

The chemical shift is the most accessible NMR-derived parameter since it can be measured very easily and with a relatively high accuracy. Furthermore, clear relationships have been found between chemical shifts of backbone protons and secondary structure elements of proteins (Wishart *et al.*, 1991). We measured the chemical shifts of the lipoyl domains of PDHC and OGDHC at 37 °C and 30 °C, respectively. By comparing the NMR spectra of the PDHC lipoyl domain recorded at 30 °C with those recorded at 37 °C we

observe however, that the changes of the chemical shifts upon this change of temperature are relatively small ($\Delta\delta < 0.02$ ppm). This legitimises a comparison between the backbone chemical shifts obtained for the two lipoyl domains. A comparison of the backbone ^{15}N , NH, and C^αH chemical shifts between the PDHC and OGDHC lipoyl domains is shown in a plot versus the amino acid sequence (Figure 5). The chemical shifts are corrected for random coil contributions by subtraction of the random coil chemical shifts obtained by Wishart *et al.* (1991) from NMR-assigned proteins with known three-dimensional structures. The corrected chemical shifts are called secondary shifts or conformation-dependent shifts. A positive value of the secondary shift means a downfield shift with respect to the random coil chemical shift, a negative value means an upfield shift. The gaps in the different chemical shift profiles of the ^{15}N and NH atoms and the C^αH protons are caused respectively by the various Pro and Gly residues, and the three-residue gap in the amino acid sequence of the PDHC lipoyl domain.

The secondary shift profiles of the NH and C^αH protons and the ^{15}N backbone atoms in Figure 5 show a very high similarity between the two lipoyl domains, despite the primary sequence identity of only 25%. The mean difference in secondary shift between the PDHC and OGDHC lipoyl domain for the NH protons, the C^αH protons and the ^{15}N nuclei is -0.01 ppm (r.m.s.d. 0.51 ppm), 0.17 ppm (r.m.s.d. 0.37 ppm), and -0.8 ppm (rmsd 4.56 ppm) respectively. Large differences (> 1.0 ppm) in secondary shifts between the two domains are only found for four amide protons and two C^αH protons. These are the NH protons of residues Ala14, His22, Gly69, and Gly73, and the C^αH protons of residues Ser68 and Thr76 of the OGDHC lipoyl domain. We will try to explain large differences in secondary shifts between the two lipoyl domains in terms of differences in other NMR-derived parameters, although the chemical shift is determined by a number of different effects, and it is not yet possible to relate thoroughly chemical shift differences with structural differences in proteins.

In all cases where large differences in secondary shift between the two lipoyl domains are observed, a difference in one or more other NMR-derived parameter is noticed. Since the detailed three-dimensional solution structures of the two lipoyl domains are not yet available, only a few of these cases are self-evident. For example, a large upfield secondary shift of 2.03 ppm of the NH proton of Gly69 of the OGDHC lipoyl domain with respect to the comparable Gly66 of the PDHC lipoyl domain is observed. This suggests that the hydrogen bonding of the NH proton of Gly69 is somewhat weaker than for Gly66 of the PDHC lipoyl domain, which is confirmed by the significant faster NH-exchange rate for Gly69 (Figure 4). In addition, a structural difference is also suggested near this proton since a stronger sequential d_{NN} contact is observed for Gly66 in

the PDHC lipoyl domain (Figure 4). As a second example, the large upfield secondary shift of 1.29 ppm of the C^αH proton of Ser68 of the OGDHC lipoyl domain with respect to the comparable Glu65 of the PDHC lipoyl domain coincides with the absence of the long-range $d_{\text{N}\alpha}$ (68,15) contact in the OGDHC lipoyl domain. In this case however, the large difference in secondary shift can not yet be explained in detail, but can only be related with a change in the local structure.

It is clear from the high similarity in secondary shifts of backbone protons between the two lipoyl domains, that their structures in solution are very similar. The high structural similarity between the two lipoyl domains is confirmed by the high similarity of the other NMR-derived parameters, e.g. vicinal coupling constants, NOEs, and NH-exchange rates, and by preliminary tertiary-structure calculations. The overall folds of both the OGDHC lipoyl domain and the PDHC lipoyl domain were computed using distance geometry methods with only a limited number of input constraints. The minimised average low-resolution structures of the two lipoyl domains are schematically drawn and compared in Figure 6, and prove to be very similar. The overall fold, two antiparallel β -sheets formed around a hydrophobic core, is the same for both lipoyl domains. The suggested global fold of the lipoyl domain is in agreement with the suggested global fold for the PDHC lipoyl domain from *A. vinelandii* (Berg *et al.*, 1994) and the solution structures of the lipoyl domain of *B. stearothermophilus* PDHC (Dardel *et al.*, 1993) and a hybrid lipoyl domain of *E. coli* PDHC (Green *et al.*, 1995).

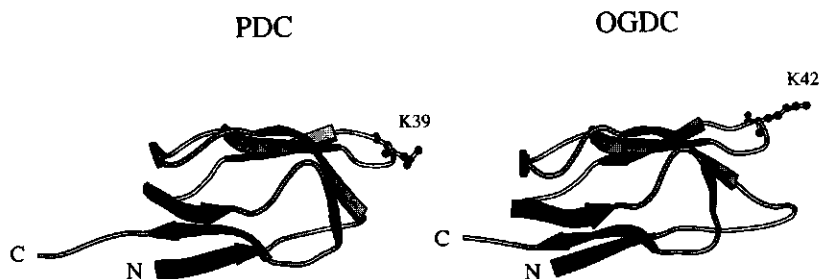


Figure 6. Comparison of the overall folds of the PDHC and OGDHC lipoyl domains from *A. vinelandii*. The tertiary-structure models were obtained by distance geometry calculations using only a limited number of constraints. The figure was generated using the program MOLSCRIPT (Kraulis, 1991).

Despite the proposed high structural similarity between the PDHC and OGDHC lipoyl domains from *A. vinelandii*, these lipoyl domains react only with the 2-oxo acid dehydrogenase (E1) of their parent complex. A difference in overall fold between the lipoyl domains from different acyltransferases is thus not the cause of the specificity of the

E1 component for lipoyl domains. The sequence around the lipoylated lysine is highly conserved in lipoyl domains from different acyltransferases. Moreover, in the *A. vinelandii* PDHC and OGDHC lipoyl domains this lipoyl-lysine residue is found in a distinct type-I turn. This suggests that not only the sequence but also the structure around the lipoyl-lysine residue is conserved in lipoyl domains from different acyltransferases from the same source. It is therefore evident that the sequence as well as the local structure around the lipoyl-lysine residue are not responsible for the specificity of E1 for its parent lipoyl domain(s).

On the basis of sequence alignment of lipoyl domains, and with the supposition that the part of the lipoyl domain that determines specificity towards E1 is likely to be in proximity of the lipoyl group, Berg *et al.* (1994) proposed that the loop connecting strand S1 with strand S2 could be a determinant of the lipoyl domain for specific interaction with E1. The sequence alignment of lipoyl domains shows that in this loop significant sequential differences occur between different acyltransferases from the same source. Our results clearly indicate that the solution structures of the *A. vinelandii* PDHC and OGDHC lipoyl domains are very similar. Large structural differences between the two lipoyl domains are suggested only for residues in this loop, as can be judged by the large difference in sequential d_{NN} contacts in this loop. This supports our hypothesis that this loop is important for specificity of E1 for lipoyl domains. Mutagenesis experiments on the lipoyl domains to test this hypothesis are in progress.

The assignment of the lipoyl domain of *A. vinelandii* OGDHC is an important step in the determination of its three-dimensional solution structure. By determining the detailed tertiary solution structures of the lipoyl domains of both *A. vinelandii* PDHC and *A. vinelandii* OGDHC, which is currently in progress in our laboratory, a comparison can be made between structures of lipoyl domains from different acyltransferases from the same source. These structures should provide more detailed information about the differences between the structures of these lipoyl domains. It is clear however, that distinct conclusions about the interactions, and the specificity of those interactions, between the lipoyl domains and the E1 components are impaired by the lack of structural information on the E1 component. Studies on the interactions of the two lipoyl domains with different E1 components of the *A. vinelandii* 2-oxo acid dehydrogenase complexes by NMR, which are now possible, can provide further information about the part of the lipoyl domain that is involved in the interaction with E1.

EXPERIMENTAL PROCEDURES

Construction of expression vector

By using the plasmid pAE2 (Westphal & de Kok, 1990) encoding the complete E2o gene as a template, the region encoding the lipoyl domain was selectively amplified by PCR. The T7/T3 α reversed sequencing primer, and a mutagenic primer that introduces a stop codon after Gly79 and contains an engineered *EcoRI* site, were used as PCR primers. The generated 375-bp PCR product was digested with *EcoRI* and *HindIII*, and the resulting 319-bp fragment ligated into pUC9. The recombinant plasmid pAB1 encodes the first 79 amino acids of the succinyltransferase component of *A. vinelandii* OGDHC. The C-terminal amino acid of the lipoyl domain was determined by sequence alignment.

Expression and isolation

E. coli strain TG2 (Gibson, 1984), a *recA*⁻ version of TG1 [Δ (*lac-pro*), *thi*, *supE*, *Res*⁻ *Mod*⁻ (*k*), *F'* (*traD36 proA*⁺*B*⁺, *lacI9 lacZ* Δ M15)] was used for the expression of the lipoyl domain. *E. coli* TG2 cells transformed with the plasmid pAB1 were grown at 37°C in TY medium containing 75 $\mu\text{g}/\text{mL}$ ampicillin and 20 $\mu\text{g}/\text{mL}$ isopropyl β -D-thiogalactopyranoside (IPTG). When expression of lipoylated protein was required, 10 $\mu\text{g}/\text{mL}$ lipoic acid was added.

The lipoyl domain of *A. vinelandii* E2o was isolated from a 6-L culture of *E. coli* TG2(pAB1) in TY medium essentially as described for the N-terminal lipoyl domain of *A. vinelandii* E2p (Berg *et al.*, 1994), with the following minor modifications. Throughout the isolation of the lipoyl domain 20 mM piperazine pH 6.0 containing 0.5 mM EDTA and 0.02% NaN₃ was used instead of potassium phosphate. The ultrafiltration (Amicon YM 30) of the cell-free extract was performed at pH 7.0 instead of pH 6.0, to prevent aggregation of proteins and subsequent clogging of the ultrafiltration membrane. In later purifications of the lipoyl domain the ultrafiltration step (Amicon YM 30) was replaced by Sephadex G-75 gel-filtration chromatography. This modification resulted in a doubling of the yield of lipoyl domain.

Uniformly ^{15}N -labelled lipoyl domain was obtained by growing bacteria on medium with $^{15}\text{NH}_4\text{Cl}$ as the sole nitrogen source, following the protocol as described by Berg *et al.* (1994). The ^{15}N -labelled lipoyl domain was isolated using the same purification method as for unlabelled lipoyl domain.

Reductive succinylation of the lipoyl domain by the *A. vinelandii* OGDH complex in the presence of [^{14}C]2-oxoglutarate was assayed in a similar manner as described by Berg *et al.* (1994).

NMR spectroscopy

The lipoyl domain was dialysed against 50 mM potassium phosphate pH 5.5 containing 100 mM potassium chloride and 0.02% NaN₃, and 10% (by vol.) D₂O was added for the lock signal. The D₂O sample was prepared by twice lyophilising the dialysed protein sample and dissolving it in D₂O. The final protein concentration was 6-8 mM. The concentration of the ¹⁵N-labelled unlipoylated lipoyl domain was 4 mM.

NMR data were acquired with a Bruker AMX500 spectrometer at 14 °C and 30 °C. All two-dimensional spectra were recorded in the phase-sensitive mode using the time-proportional phase incrementation method (TPPI) (Marion & Wüthrich, 1983). The following regular two-dimensional ¹H-NMR experiments were performed: double-quantum filtered correlation spectroscopy (DQF-COSY) (Rance *et al.*, 1983), total correlation spectroscopy (TOCSY) using the clean-MLEV17 mixing sequence (Griesinger *et al.*, 1988) and a mixing time of 80 ms, and nuclear Overhauser enhancement spectroscopy (NOESY) (Jeener *et al.*, 1979; Kumar *et al.*, 1980) using mixing times of 70, 100 and 150 ms. Solvent suppression was achieved using selective low-power irradiation for 1.5 s during the relaxation delay. For spectra recorded in H₂O, presaturation was employed together with the SCUBA (stimulated cross peaks under bleached alphas) sequence (Brown *et al.*, 1988) using a SCUBA delay of 70 ms to recover saturated resonances under the solvent peak. All homonuclear 2D spectra were acquired with 512 *t*₁ experiments (64 or 80 transients/*t*₁ increment) with 2048 data points, and a spectral width of 7042 Hz at 500 MHz in both dimensions. For measuring ³J_{αβ} coupling constants an exclusive correlation spectroscopy (E.COSY) (Griesinger *et al.*, 1985; Griesinger *et al.*, 1987) spectrum was acquired in D₂O at 30 °C, with 920 *t*₁ increments of 4k data points.

¹H-¹⁵N heteronuclear single-quantum coherence (HSQC) spectra (Bodenhausen & Ruben, 1980) were recorded using 256 increments of *t*₁, 2048 data points, and spectral widths of 10 kHz in ω_2 and 1824 Hz in ω_1 . The (1/4J_{NH}) delay was set to 2.7 ms. ¹⁵N decoupling during acquisition was accomplished using the GARP decoupling sequence (Shaka *et al.*, 1985). The residual water peak was suppressed by low-power presaturation. The ¹⁵N carrier frequency was placed in the centre of the backbone amide ¹⁵N spectrum at 120.9 ppm.

¹H-¹⁵N heteronuclear multiple-quantum coherence (HMQC) spectra (Müller, 1979; Bax *et al.*, 1983) combined with NOESY (mixing time 150 ms) and clean-TOCSY (mixing time 80 ms) with an inserted SCUBA delay, were acquired using identical spectral widths and decoupling as used in the HSQC experiments. The (1/2J_{NH}) delay in the HMQC experiments was set to 5.3 ms.

To obtain reasonable accurate $^3J_{\text{HN}\alpha}$ values, ^1H - ^{15}N HMQC- J spectra (Kay & Bax, 1990) were recorded for the *A. vinelandii* OGDHC and the PDHC lipoyl domain at 30 °C according to Forman-Kay *et al.* (1990). The HMQC- J experiments were recorded with 1248 t_1 increments of 2k data points per increment. Different sine and sine-squared apodization functions were applied in ω_1 that was zero-filled to 4k data points (digital resolution 0.44 Hz/point). Coupling constants were derived by non-linear least-squares fitting of t_1 cross sections with respect to the coupling constant and line width (Redfield & Dobson, 1990). For peaks for which no satisfactory fit was obtained, due to a poor signal/noise ratio or partial overlap with another peak, the coupling constant was estimated from comparison with spectral simulations.

NMR data were processed using the program FELIX version 2.3 from Biosym Technologies Inc.(San Diego), running on a Silicon Graphics Indigo² workstation. For most spectra, a phase-shifted (35-90°) sine-squared apodization function was applied prior to Fourier transformation. Most homonuclear spectra were zero-filled to 2048 x 2048 data points. For the extraction of $^3J_{\alpha\beta}$ couplings, the E.COSY spectrum was zero-filled to 4096 x 4096 data points to increase the digital resolution in ω_2 to 1.72 Hz/point. Baseline corrections were applied using the FLATT routine (Güntert & Wüthrich, 1992) of the FELIX program.

The chemical shifts are reported in ppm relative to internal trimethylsilyl propionate for ^1H and to external liquid NH_3 for ^{15}N . The ^{15}N chemical shifts were calculated relative to the ^1H standard according to Live *et al.* (1984).

Slowly exchanging amide protons were identified in HSQC experiments started 30 min after dissolving the lyophilised protein in D_2O . Subsequently, every 90 min a HSQC spectrum was recorded over a period of 24 h.

Structure calculations

Preliminary tertiary structures of the lipoyl domains were computed using the program DG-II (Havel, 1991) in the NMR refine module of the InsightII package (Biosym Technologies Inc., San Diego), according to the standard procedure as described by Havel (1991). A limited number of experimental constraints was used in the distance geometry calculations: 451 NOE distance and 47 ϕ torsion angle constraints for the OGDHC lipoyl domain, and 387 NOE distance and 53 ϕ torsion angle constraints for the PDHC lipoyl domain. A total of 20 structures was calculated for each lipoyl domain, and structures possessing no distance violations larger than 0.05 nm were averaged and minimised to obtain a tentative tertiary-structure model for each lipoyl domain.

Acknowledgements

We wish to thank dr. Jack Stokkermans for assistance with PCR and Adrie Westphal for help with the purification of the lipoyl domain. We are very grateful to dr. Ulrich Hommel and dr. Paul C. Driscoll for generously providing us with their program for fitting HMQC-J traces, and to Frank Vergeldt for assistance with the spectral simulations. This work was financially supported by the Netherlands Organisation for Scientific Research (NWO) under the auspices of the Netherlands Foundation for Chemical Research (SON).

REFERENCES

- Bax, A., Griffey, R. H. & Hawkins, B. L. (1983) Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Reson.* **55**, 301-315.
- Berg, A., de Kok, A. & Vervoort, J. (1994) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **221**, 87-100.
- Bodenhausen, G. & Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185-189.
- Brown, S. C., Weber, P. L. & Müller, L. (1988) Toward complete ^1H NMR spectra in proteins. *J. Magn. Reson.* **77**, 166-169.
- Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- Driscoll, P. C., Clore, G. M., Beress, L. & Gronenborn, A. M. (1989) A proton nuclear magnetic resonance study of the anti-hypertensive and anti-viral protein BDS-I from the sea anemone *Anemonia sulcata*. *Biochemistry* **28**, 2178-2187.
- Forman-Kay, J. D., Gronenborn, A. M., Kay, L. E., Wingfield, P. T. & Clore, G. M. (1990) Studies on the solution conformation of human thioredoxin using heteronuclear ^{15}N - ^1H nuclear magnetic resonance spectroscopy. *Biochemistry* **29**, 1566-1572.
- Gibson, T. J. (1984) *Studies on the Epstein-Bar virus genome*. Ph.D. Thesis, University of Cambridge, UK.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.
- Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995) Three-dimensional structure of a lipoyl domain from dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.
- Griesinger, C., Sørensen, O. W. & Ernst, R. R. (1985) Two-dimensional correlation of connected transitions. *J. Am. Chem. Soc.* **107**, 6394-6396.
- Griesinger, C., Sørensen, O. W. & Ernst, R. R. (1987) Practical aspects of the E.COSY technique. Measurement of scalar spin-spin coupling constants in peptides. *J. Magn. Reson.* **75**, 474-492.
- Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988) Clean TOCSY for ^1H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870-7872.
- Gronenborn, A. M., Bax, A., Wingfield, P. T. & Clore, G. M. (1989) A powerful method of sequential proton resonance assignment in proteins using relayed ^{15}N - ^1H multiple quantum coherence spectroscopy. *FEBS Lett.* **243**, 93-98.
- Güntert, P. & Wüthrich, K. (1992) FLATT-A new procedure for high-quality baseline correction of multidimensional NMR spectra. *J. Magn. Reson.* **96**, 403-407.

- Hancmaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987) The domain structure of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **169**, 245-252.
- Hancmaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988) The dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Molecular cloning and sequence analysis. *Eur. J. Biochem.* **172**, 593-599.
- Havel, T. F. (1991) An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Prog. Biophys. Molec. Biol.* **56**, 43-78.
- Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546-4553.
- Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K. & Perham, R. N. (1993) The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *J. Mol. Biol.* **230**, 323-341.
- Karplus, M. (1963) Vicinal proton coupling in nuclear magnetic resonance. *J. Am. Chem. Soc.* **85**, 2870-2871.
- Kay, L. E. & Bax, A. (1990) New methods for the measurement of NH-CoH coupling constants in ^{15}N -labelled proteins. *J. Magn. Reson.* **86**, 110-126.
- Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
- Kumar, A., Ernst, R. R. & Wüthrich, K. (1980) A two-dimensional nuclear overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1-6.
- Live, D. H., Davis, D. G., Agosta, W. C. & Cowburn, D. (1984) Long range hydrogen bond mediated effects in peptides: ^{15}N NMR study of gramicidin S in water and organic solvents. *J. Am. Chem. Soc.* **106**, 1939-1941.
- Marion, D. & Wüthrich, K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurement of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967-974.
- Mattevi, A., Schierbeek, A. J. & Hol, W. G. J. (1991) Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 Å resolution. A comparison with the structure of glutathione reductase. *J. Mol. Biol.* **220**, 975-994.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992a) The pyruvate dehydrogenase multienzyme complex. *Curr. Opin. Struct. Biol.* **2**, 877-887.
- Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1992b) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
- Mattevi, A., Obmolova, G., Kalk, K. H., Sokatch, J., Betzel, C. H. & Hol, W. G. J. (1992c) The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD^+ at 2.45 Å resolution. *Proteins* **13**, 336-351.
- Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1993a) Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Ångstrom resolution. *J. Mol. Biol.* **230**, 1183-1199.
- Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. H. & Hol, W. G. J. (1993b) 3-Dimensional structure of lipoamide dehydrogenase from *Pseudomonas fluorescens* at 2.8 Ångstrom resolution - analysis of redox and thermostability properties. *J. Mol. Biol.* **230**, 1200-1215.
- Müller, L. (1979) Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J. Am. Chem. Soc.* **101**, 4481-4484.
- Perham, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.
- Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983) Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479-485.
- Redfield, C. & Dobson, C. M. (1990) ^1H NMR studies of human lysozyme: spectral assignment and comparison with hen lysozyme. *Biochemistry* **29**, 7201-7214.
- Reed, L. J., Koike, M., Levitch, M. E. & Leach, F. R. (1958) Studies on the nature and reactions of protein-bound lipoic acid. *J. Biol. Chem.* **232**, 143-158.

- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.
- Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K. & Gronenborn, A. M. (1992) Three-dimensional solution structure of the E3-binding domain of the dihydrolipoamide succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **31**, 3463-3471.
- Schierbeek, A. J., Swarte, M. B. A., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G. J., Drenth, J. & Betzel, C. (1989) X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *J. Mol. Biol.* **206**, 365-379.
- Schulze, E., Westphal, A. H., Boumans, H. & de Kok, A. (1991a) Site-directed mutagenesis of the dihydrolipoyl transacetylase component (E2p) of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Binding of the peripheral components E1p and E3. *Eur. J. Biochem.* **202**, 841-848.
- Schulze, E., Westphal, A. H., Obmolova, G., Mattevi, A., Hol, W. G. J. & de Kok, A. (1991b) The catalytic domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* and *Escherichia coli*. Expression, purification, properties and preliminary X-ray analysis. *Eur. J. Biochem.* **201**, 561-568.
- Schulze, E., Westphal, A. H., Veeger, C. & de Kok, A. (1992) Reconstitution of pyruvate dehydrogenase multienzyme complexes based on chimeric core structures from *Azotobacter vinelandii* and *Escherichia coli*. *Eur. J. Biochem.* **206**, 427-435.
- Shaka, A. J., Barker, P. B. & Freeman, R. (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.* **64**, 547-552.
- Wagner, G., Braun, W., Havel, T. H., Schaumann, T., Go, N. & Wüthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* **196**, 611-639.
- Wallis, N. G. & Perham, R. N. (1994) Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**, 209-216.
- Westphal, A. H. & de Kok, A. (1990) The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur. J. Biochem.* **187**, 235-239.
- Wishart, D. S., Sykes, B. D. & Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein structure. *J. Mol. Biol.* **222**, 311-333.
- Wüthrich, K. (1986) *NMR of proteins and nucleic acids*. Wiley, New York.
- Zuiderweg, E. R. P., Boelens, R. & Kaptein, R. (1985) Stereospecific assignments of ¹H-NMR methyl lines and conformation of valyl residues in the lac repressor headpiece. *Biopolymers* **24**, 601-611.

CHAPTER 4

Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*

ABSTRACT

The three-dimensional solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii* has been determined from nuclear magnetic resonance data by using distance geometry and dynamical simulated annealing refinement. The structure determination is based on a total of 580 experimentally derived distance constraints and 65 dihedral angle constraints. The solution structure is represented by an ensemble of 25 structures with an average root-mean-square deviation between the individual structures of the ensemble and the mean coordinates of 0.71 Å for backbone atoms and 1.08 Å for all heavy atoms. The overall fold of the lipoyl domain is that of a β -barrel-sandwich hybrid. It consists of two almost parallel four-stranded anti-parallel β -sheets formed around a well-defined hydrophobic core, with a central position of the single tryptophan 21. The lipoylation site, lysine 42, is found in a β -turn at the far end of one of the sheets, and is close in space to a solvent-exposed loop comprising residues 7 to 15. The lipoyl domain displays a remarkable internal symmetry that projects one β -sheet onto the other β -sheet after rotation of approximately 180° about a 2-fold rotational symmetry axis. There is close structural similarity between the structure of this 2-oxoglutarate dehydrogenase complex lipoyl domain and the structures of the lipoyl domains of pyruvate dehydrogenase complexes from *Bacillus stearothermophilus* and *Escherichia coli*, and conformational differences occur primarily in a solvent-exposed loop close in space to the lipoylation site. The lipoyl domain structure is discussed in relation to the process of molecular recognition of lipoyl domains by their parent 2-oxo acid dehydrogenase.

INTRODUCTION

Lipoyl domains are the small (8 kDa) substrate-carrying domains of 2-oxo acid dehydrogenase multienzyme complexes. This family of multienzyme complexes consists of pyruvate, 2-oxoglutarate, and branched-chain 2-oxo acid dehydrogenase complexes, and catalyse the irreversible oxidative decarboxylation of their respective 2-oxo acids to the corresponding acyl-CoA derivatives [reviewed by Perham (1991) and Mattevi *et al.* (1992a)]. These complexes have a very similar design and share many structural and catalytic properties. The 2-oxoglutarate dehydrogenase complex (OGDHC) from *Azotobacter vinelandii*, which is part of the tricarboxylic acid cycle, consists of multiple copies of three protein components. The structural core of the complex is formed by 24 subunits of dihydrolipoyl succinyltransferase (E2o) arranged with octahedral symmetry, to which 12 dimers of 2-oxoglutarate dehydrogenase (E1o) and six dimers of lipoamide dehydrogenase (E3) are bound in a non-covalent manner. The E2o monomer is composed of three separate and independently folded domains (Westphal & de Kok, 1990): (1) an N-terminal lipoyl domain (about 80 residues) containing the covalently bound prosthetic group lipoic acid, (2) a peripheral subunit-binding domain (about 35 residues) involved in the binding of the E1 and E3 dimers to the structural core, and (3) a C-terminal catalytic domain (29 kDa) that accommodates the succinyltransferase activity and aggregates to form the core of the multienzyme complex. The domains are connected by long flexible linker segments unusually rich in alanine and proline residues.

The structures of several individual domains of acyltransferase components of various 2-oxo acid dehydrogenase multienzyme complexes have recently been determined. By X-ray diffraction the crystal structure of the catalytic core domain of the E2 component of the pyruvate dehydrogenase complex (PDHC) from *A. vinelandii* has been determined (Mattevi *et al.*, 1992b, 1993). The structures of the lipoyl domain of *Bacillus stearothermophilus* PDHC (Dardel *et al.*, 1993) and a non-native hybrid lipoyl domain of *Escherichia coli* PDHC (Green *et al.*, 1995) have been determined by means of NMR spectroscopy. The structures of the peripheral subunit-binding domains of E2o of *E. coli* OGDHC (Robien *et al.*, 1992) and of dihydrolipoyl acetyltransferase (E2p) of *B. stearothermophilus* PDHC (Kalia *et al.*, 1993) have also been solved by means of NMR spectroscopy.

The lipoyl domains play a crucial role in coupling the activities of the three enzyme components in 2-oxo acid dehydrogenase complexes. A specific lysine residue side-chain of each lipoyl domain is modified with lipoic acid to form a so-called lipoyl group, which visits the three successive active sites (Reed, 1974). With that, it transports acyl groups

from E1 to E2, and reduction equivalents from E2 to E3. A folded structure of the lipoyl domain attached to the lipoyl group has been shown to be required for the reaction with the E1 enzyme (Graham *et al.*, 1989), in contrast to the E2 and E3 enzymes which can use free lipoamide as a substrate (Reed *et al.*, 1958). Moreover, lipoyl domains can only be efficiently reductively acylated by the E1 enzyme of their parent complex (Graham *et al.*, 1989; chapter 6, this thesis). The E1 enzymes are thus not only specific for their 2-oxo acid substrate but also for their lipoyl domains.

One of the important objectives of our current research on lipoyl domains is to understand in more detail the structural basis for the molecular recognition of lipoyl domains by their parent E1. By a comparison of structures of different lipoyl domains we expect to advance our insight in this process. It must be noted that no structural information on any E1 component is yet available. Comparison of the solution structures of the lipoyl domain of *B. stearothermophilus* PDHC and the hybrid lipoyl domain of *E. coli* PDHC showed that all lipoyl domains will likely have similar overall folds (Green *et al.*, 1995). A comparison of the solution structures of lipoyl domains from different acyltransferases from the same source (in our case E2p and E2o from *A. vinelandii*) with known specificities for each other's E1 components, should lead to a more detailed picture of the process of molecular recognition. We have recently reported on the subcloning, over-expression, and ¹H and ¹⁵N nuclear magnetic resonance assignments of these lipoyl domains (Berg *et al.*, 1994, 1995), which provide the necessary background data for their structure determination. Here we present the three-dimensional solution structure of one of them, the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *A. vinelandii*.

RESULTS

Structure calculations

From the NOESY spectra 668 distance constraints were collected, but 137 were found to be redundant and were not included in the structure calculations. The number of NOE distance constraints that were non-trivial (531) increased slightly to 550 due to the DIANA treatment of stereospecifically unassigned diastereotopic hydrogen atoms (see Materials and Methods). An additional 30 distance constraints derived from the identification of 15 hydrogen bonds were only included in the final round of structure calculations. A total of 65 dihedral angle constraints was obtained, 47 for ϕ angles (34 for $^3J_{\text{HN}\alpha}$ values ≥ 8 Hz), and 18 for χ_1 angles. The distribution of the constraints (NOEs plus hydrogen bonds) as a function of the residue number is shown in Figure 1(a).

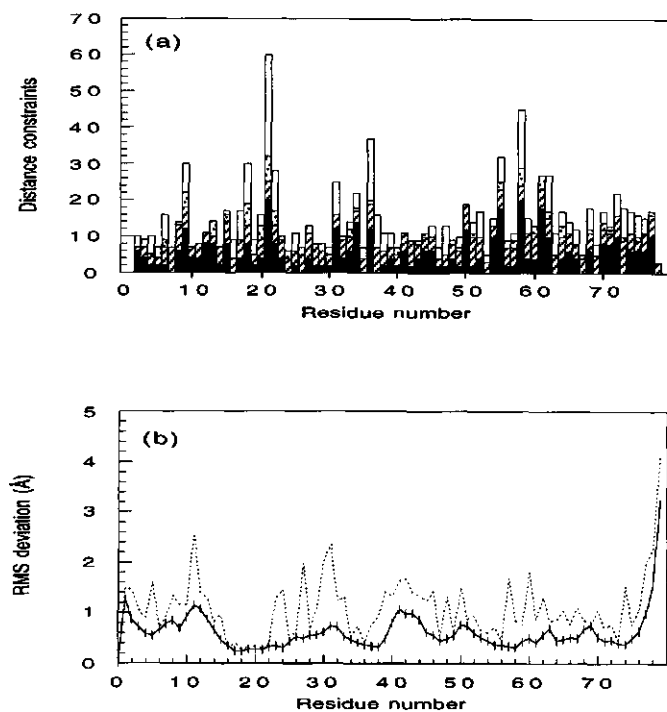


Figure 1. (a) Plot of the number of experimentally derived distance constraints (NOEs plus hydrogen bonds) as a function of the residue number. The bars representing the different distance constraint categories are as follows: intra-residue (filled), sequential (hatched), medium range (stippled), and long range (open). Each inter-residue distance constraint is counted twice, once for each residue involved. (b) Plot of the residue-based r.m.s. deviation of the individual DG-SA structures from the mean-structure $(DG-SA)_r$ of backbone heavy atoms (continuous line) and all heavy atoms (broken line).

With the final set of 645 experimental constraints 50 structures were calculated, and solely on the basis of constraint satisfaction, 25 structures were selected to represent the solution structure of the lipoyl domain. These structures show no distance constraint violations greater than 0.2 Å (1 Å = 0.1 nm) and no dihedral angle constraint violations greater than 1°, which reveals that good constraint satisfaction can be achieved. The structures also have small deviations from idealised geometry and reasonable non-bonded contacts (Table 1), as indicated by the low value of the repulsive van der Waals energy term. A good fit of the final structures to the experimental constraints could only be achieved using relatively long simulation times in the simulated annealing calculations. However, a comparable constraint satisfaction could be achieved with several successive

rounds of simulated annealing refinement (Brünger, 1992) following the initial distance geometry/simulated annealing (DG-SA) calculation which used shorter periods of dynamics.

Table 1. Structural statistics^a

	< DG-SA >	(DG-SA) _r
R.m.s. deviations from experimental		
distance constraints (Å) ^b :		
All NOE constraints (550)	0.015(±0.001)	0.012
Intra-residue (200)	0.013(±0.002)	0.011
Sequential (li-jl) = 1 (178)	0.014(±0.003)	0.011
Short range (li-jl) ≤ 4 (39)	0.0006(±0.0004)	0.0002
Long range (li-jl) > 4 (133)	0.0013(±0.0009)	0.0009
Hydrogen bond (30)	0.0006(±0.0004)	0.0008
R.m.s. deviations from experimental		
dihedral constraints (°) (65)	0.12(±0.01)	0.16
R.m.s. deviations from idealised covalent		
geometry:		
Bonds (Å)	0.0019(±0.0001)	0.0017
Angles (°)	0.479(±0.001)	0.470
Impropers (°)	0.284(±0.002)	0.273
Energy (kcal/mol) ^c :		
F _{NOE}	8.468(±1.078)	5.484
F _{cdih}	0.061(±0.011)	0.107
F _{repel}	3.832(±0.296)	3.204

^a < DG-SA > refers to the 25 final structures obtained by the hybrid distance geometry/simulated annealing protocol. (DG-SA)_r is the restrained minimised mean structure obtained by averaging the atomic coordinates of the final structures best fitted to each other over the backbone atoms.

^b The number of each type of constraint used in the structure calculations is given in parentheses.

^c The force constants used for these calculations were 50 kcal/mol per Å², and 200 kcal/mol per rad², for F_{NOE} and F_{cdih}, respectively. F_{repel} was calculated using the final value of 4 kcal/mol per Å⁴ with the van der Waals hard sphere radii set to 0.75 times the standard values used in the CHARMM empirical energy function (Brooks *et al.*, 1983). The paralldg.pro parameter set supplied with X-PLOR version 3.1 (Brünger, 1992) was used for the calculations.

Description of the structures

The ensemble of 25 structures representing the solution structure of the lipoyl domain shows reasonably good convergence. A superposition of the 25 DG-SA structures is shown in Figure 2. These structures display a root-mean-square (r.m.s.) deviation from the mean structure of 0.71 Å for backbone atoms and 1.08 Å for all non-hydrogen atoms (Table 2). If the less well-defined N-terminal residue and two C-terminal residues are excluded, these values drop to 0.60 Å and 1.01 Å for backbone and heavy atoms, respectively. The r.m.s. deviation values indicate that the structure of the lipoyl domain is well defined in solution. In Figure 1(b) the r.m.s. deviation of the 25 structures from the mean structure for backbone atoms and all heavy atoms as a function of the residue number is depicted.

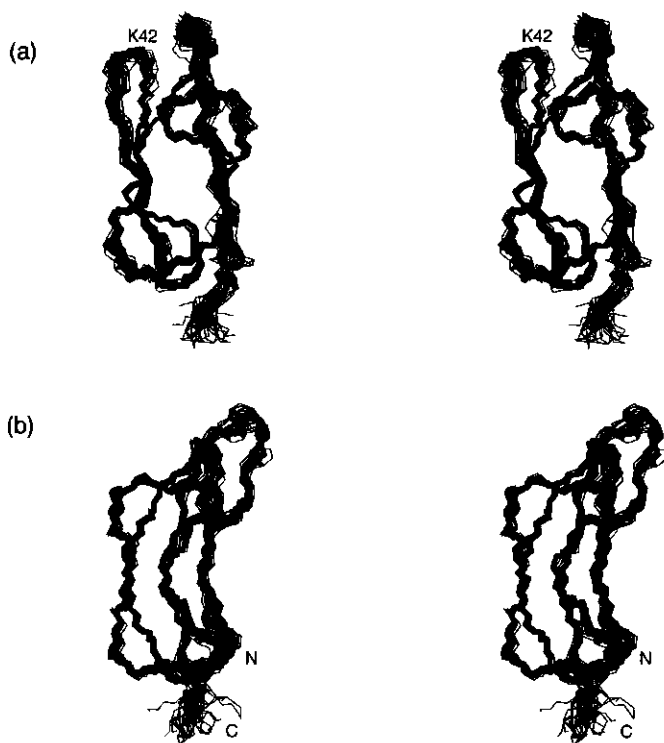


Figure 2. (a) Stereoview of the 25 selected structures of the lipoyl domain, superposed over the backbone (N, C α , C) atoms for residues 2 to 77. Only the backbone heavy atoms are shown. (b) Side view of the superposition shown in (a), obtained by a 90° rotation about the vertical axis.

Table 2. Atomic r.m.s. differences for the 25 selected structures of the lipoyl domain ^a

	Residues 1 to 79	Residues 2 to 77 ^b	β -Sheet ^c
To mean			
Backbone ^d (Å)	0.71(\pm 0.16)	0.59(\pm 0.13)	0.48(\pm 0.10)
Heavy (Å)	1.08(\pm 0.15)	1.03(\pm 0.15)	0.85(\pm 0.08)
Pairwise			
Backbone (Å)	1.03(\pm 0.21)	0.84(\pm 0.18)	0.62(\pm 0.15)
Heavy (Å)	1.56(\pm 0.19)	1.47(\pm 0.19)	1.19(\pm 0.13)

^a The r.m.s. deviations from the mathematical mean structure, obtained by a least-squares fit of the 25 selected structures over the backbone heavy atoms (N, C $^{\alpha}$, C) of residues 2 to 77

^b The r.m.s. deviation of residues 2 to 77, excluding the less well-defined N-terminal and C-terminal residues

^c The r.m.s. deviation of regions adopting regular β -sheet structure; residues 2 to 6, 16 to 21, 27 to 29, 34 to 39, 44 to 48, 53 to 58, 64 to 66 and 71 to 76

^d These include the backbone heavy atoms N, C $^{\alpha}$, and C only

Four regions in the protein show a backbone r.m.s. deviation to the mean structure greater than 1.0 Å. These regions coincide approximately with regions containing residues for which no long-range NOEs were obtained. Two of these are the N-terminal residue and the two C-terminal residues, which are in proximity in the structure. In particular the C-terminal Gly79 shows a significantly smaller linewidth of the backbone amide proton resonance (Berg *et al.*, 1995), suggesting a higher mobility of this residue as compared to the rest of the molecule. The other two parts of the lipoyl domain that show greater backbone r.m.s. deviation values are a solvent-exposed loop region near Glu11 and the β -turn region around Lys42, the active site lipoyl-lysine residue. These regions are also close together in three-dimensional space, at the opposite side of the N- and C-terminal ends of the molecule.

Analysis of the angular order parameters (Hyberts *et al.*, 1992) for both ϕ and ψ angles revealed that most residues have well-defined ($S^{\phi,\psi} > 0.9$) backbone dihedral angles (data not shown). All of the non-glycine residues show sterically allowed ϕ,ψ angle combinations, with 66% of the residues lying in the most-favoured regions of the ϕ,ψ conformational space (Morris *et al.*, 1992). It should be noted however that, as pointed out by Morris *et al.* (1992), the ϕ,ψ distribution cannot be used as an absolute measure of the accuracy of structures calculated by an energy-based method such as X-PLOR. The only non-glycine residue with a positive ϕ angle is Asp32. This residue has well-defined

backbone torsion angles ($S\phi,\psi = 0.99$) of $70.8(\pm 4)^\circ$ (ϕ) and $44.0(\pm 9)^\circ$ (ψ) and lies in the α_L region of ϕ,ψ conformational space. For Asp32 a $^3J_{HN\alpha}$ coupling constant between 6 and 8 Hz was measured, together with a very strong intra-residue NH-C $^\alpha$ H contact. Both these observations are consistent with residues showing positive ϕ angles (Ludvigsen & Poulsen, 1992).

A schematic ribbon drawing of the energy-minimised average structure of the lipoyl domain is illustrated in Figure 3. The β -strands are numbered sequentially. The overall fold is that of a β -barrel-sandwich hybrid (Chothia & Murzin, 1993) which is now recognised as being typical for lipoyl domains. The molecule is approximately 30 Å long (backbone-to-backbone distance), and 15 Å wide and deep, with the long side being parallel to the β -sheets from the lipoyl-lysine turn to the N and C termini. The overall topology of the β -strands (Berg *et al.*, 1995) is somewhat different from a classic Greek key, and should be regarded as a separate class of β -sandwich protein structures as discussed by Green *et al.* (1995). The lipoyl domain consists of six major (five to six residues) and two minor (three residues) antiparallel β -strands that form two very similar β -sheets of four strands each, which run approximately parallel. Sheet A is formed by the strands S1, S3, S6 and S8, and sheet B is formed by the strands S2, S4, S5 and S7. The two sheets are connected by loops and turns adopting no regular repetitive secondary structure, and share an interface of hydrophobic residues forming the core of the domain.

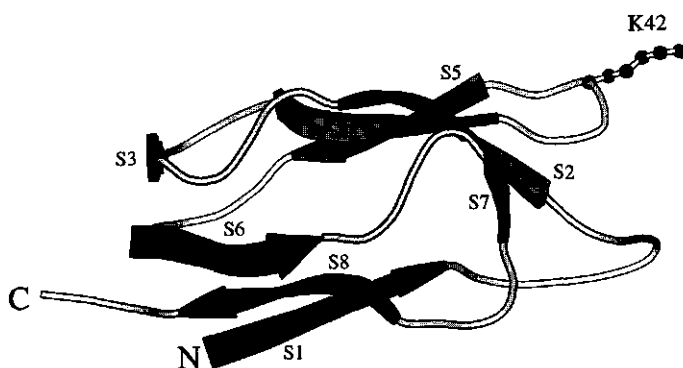


Figure 3. Schematic representation of the minimised average structure of the lipoyl domain, generated using the program MOLSCRIPT (Kraulis, 1991). The β -strands are labelled with their strand numbers. The N and C termini are close together in β -sheet A (dark grey), the lipoylation site Lys42 is situated in β -sheet B (light grey), at the opposite side of the domain.

The structure of the lipoyl domain is remarkably symmetrical and contains an internal 2-fold symmetry axis through the centre of the hydrophobic core and parallel to the plane formed by either one of the β -sheets. Rotation of approximately 180° about this axis aligns the two halves of the molecule (Gly16 to Glu39 and Gly53 to Thr76) with an r.m.s. deviation of only 0.83 Å for backbone atoms (Figure 4). The rotation projects one β -sheet onto the original position of the other sheet, and the β -turn containing the lipoyl-lysine residue onto the N- and C-terminal residues, and *vice versa*. Other structural features, like four type II turns and four β -bulges (Figure 5), are all strikingly found, in pairs, at exact symmetrical positions. The internal symmetry is also logically reflected in a repeating pattern of observed NMR parameters, e.g. coupling constants, chemical shifts and NOEs. The 2-fold quasi-symmetry seems a highly conserved feature of lipoyl domains.

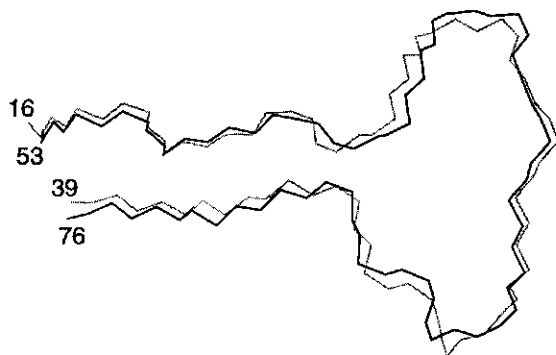


Figure 4. Superposition of two halves of the lipoyl domain, illustrating the internal 2-fold symmetry. The backbone heavy atoms of residues 16 to 39 (grey line) were aligned with those of residues 53 to 76 (black line).

The N-terminal and C-terminal ends of the lipoyl domain are close together and meet in the adjacent β -strands S1 and S8 in sheet A. At the exact opposite side of the molecule, at the far end of sheet B, the active site lipoyl-lysine residue (Lys42) is presented to the solvent at position 3 of a β -turn that connects the successive strands S4 and S5. Strand S1 is connected to strand S2 by an exposed loop comprising residues Pro7 to Asp15. Residues in this loop all show fast amide-proton exchange rates (Berg *et al.*, 1995). The minor strand S3 is connected to the major strands S2 and S4 by loop regions which both contain a type II β -turn. Similarly, as a rational consequence of the internal symmetry of the lipoyl domain, the minor strand S7 is linked to the strands S8 and S6 in the opposing sheet by loop regions each containing a type II turn as well. The β -strands S5 and S6 are

connected by a loop comprising residues Ala49 to Asp52. This loop is shorter and adopts a different conformation than its symmetry-related exposed loop (connecting strand S1 with S2), and with that it disturbs the symmetry created by the rest of the domain.

	1	5	10	15	20	25	30	35																															
Av_O	A	I	D	I	K	A	P	T	F	P	E	S	I	A	D	G	T	V	A	T	W	H	K	K	P	G	E	A	V	K	R	D	E	L	I	V	D	I	
Bs_P	A	F	E	F	K	L	P	D	I	G	E	G	I	H	E	G	E	I	V	K	W	F	V	K	P	G	D	E	V	N	E	D	D	V	L	C	E	V	
Ec_P1,3	A	I	E	I	K	V	P	D	I	G	--	A	D	E	V	E	I	T	E	I	L	V	K	V	G	D	K	V	E	A	E	Q	S	L	I	T	V		
Av_P1	S	E	I	I	R	V	P	D	I	G	---	G	D	G	E	V	I	E	L	L	V	K	T	G	D	L	I	E	V	E	Q	G	L	V	L				
sheet		<u>a</u>	<u>a</u>	<u>a</u>	<u>a</u>	<u>a</u>						<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>				<u>a</u>	<u>a</u>	<u>a</u>			<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>							
turn																																							
consvd		*	*	*								*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
	40	*	45	50	55	60	65	70	75																														
Av_O	E	T	D	K	V	V	M	E	V	L	A	E	A	D	G	V	I	A	E	I	V	K	N	E	G	D	T	V	L	S	G	E	L	Z	G	K	L	T	E
Bs_P	Q	N	D	K	A	V	V	E	I	P	S	P	V	K	G	K	V	L	E	I	L	V	P	E	G	T	V	A	T	V	G	Q	T	L	I	T	L	D	A
Ec_P1,3	E	G	D	Q	A	S	M	E	V	P	A	P	F	A	G	V	V	K	E	L	K	V	N	V	G	D	K	V	K	T	G	S	L	I	M	I	F	E	V
Av_P1	E	S	A	K	A	S	M	E	V	P	S	P	K	A	G	V	V	K	S	V	S	V	K	L	G	D	K	L	K	E	G	D	A	I	I	E	L	E	P
sheet	<u>b</u>			<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	<u>b</u>	
turn				1	1																																		
consvd	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Figure 5. Structure-based sequence alignment of the lipoyl domains of *A. vinelandii* OGDHC (Av_O), *B. stearothermophilus* PDHC (Bs_P) (Dardel *et al.*, 1993), a hybrid lipoyl domain of *E. coli* PDHC (Ec_P1,3) (Green *et al.*, 1995), and the N-terminal lipoyl domain of *A. vinelandii* PDHC (Av_P1) (Berg *et al.*, 1994, 1995). Residue numbers above the sequence refer to the Av_O sequence. Residues belonging to the hydrophobic core of the lipoyl domain are indicated in the sequence by bold italic type, except for Av_P1, for which only a low resolution three-dimensional structure is yet available. The lipoyl-lysine residue is underlined. The locations of the β -strands in the Av_O structure are indicated in the line labelled "sheet", where a represents β -sheet A, and b represents β -sheet B. In the line labelled "turn" the positions of type I (1) turns, type II (2) turns and β -bulges (β) in the structure of Av_O are indicated. The bottom-line, consvd, indicates amino acid residues conserved or semi-conserved in all lipoyl domains.

The hydrophobic core of the lipoyl domain consists of a number of residues all having a well-defined side-chain conformation with an r.m.s. deviation smaller than 1.0 Å, with the exception of Val29 (r.m.s. deviation 1.09 Å). The side-chains of the following residues form the hydrophobic core; Ile4, Val18, Trp21, Val29, Ile35, Val36, Ile38, Val47, Ala49, Ile55, Ile58, Val66, Leu72 and Leu75. The central position in the core is occupied by the single tryptophan residue 21. This residue shows many long-range NOEs to other members of the core, and consequently has a very well-defined side-chain conformation (r.m.s. deviation 0.29 Å). The top of the core, as shown in Figure 2(a), is covered by this

Trp21 and by Ile58, while the bottom is closed by residues Ile4, Ile38 and Val47. The hydrophobic core is flanked by the residues Ile38 and Val66 on one side, and Ala49 and Val29 on the other side. The hydrophobic interactions between the side-chains of the residues forming the hydrophobic core undoubtedly keep the two β -sheets together and are essential for the stability of the lipoyl domain.

DISCUSSION

The structure of the lipoyl domain provides a specific attachment site for the lipoyl group. The lipoylation site, Lys42, is located at position 3 in a β -turn at the periphery of one of the β -sheets where it is solvent exposed. By using site-directed mutagenesis Wallis & Perham (1994) clearly showed that a structural rather than a sequence motif around the lipoylation site is responsible for the recognition of this site by the lipoylating enzymes, and that the lipoyl-lysine residue should be in position 3 of this β -hairpin turn in order to become lipoylated. It is worth noting here that lipoylation of the specific lysine residue only induces small changes in the chemical shift of residues close to the lipoylation site, suggesting that the conformation of the lipoyl domain is not altered by the covalently coupled lipoyl group (Dardel *et al.*, 1991; Berg *et al.*, 1994). This also indicates that the lipoyl group, although quite hydrophobic in nature, protrudes into the solvent and does not bind back to a hydrophobic surface on the lipoyl domain, as for example in the case of the methylamine-loaded H-protein of the glycine decarboxylase complex from pea leaf mitochondria (Cohen-Addad *et al.*, 1995).

A noticeable characteristic of the lipoyl domain is its highly negative charge, with a measured isoelectric point of approximately 4. The high number of charged residues, with the majority being negatively charged, are found exclusively on the surface of the domain. The distribution of charges on the surface of the lipoyl domain seems to be uniform and no obvious clustering of charges is observed. Despite the lack of experimental evidence, we believe that the highly charged surface of the lipoyl domain prevents interaction with the hydrophobic linker sequence, other lipoyl domains in the complex, and other aspecific protein-protein interactions. These interactions would be highly disadvantageous for the functioning of the lipoyl domain moving rapidly among the different active sites in the complex, and would undo the advantages of being part of an efficient multienzyme complex.

Comparison with other lipoyl domain structures

We describe here the first three-dimensional structure of a lipoyl domain of a 2-oxoglutarate dehydrogenase complex, from *A. vinelandii*. Two other three-dimensional solution structures of lipoyl domains have recently been described, both from PDH complexes. These are the lipoyl domain from *B. stearrowthermophilus* PDHC (Dardel *et al.*, 1993; Protein Data Bank accession code 1LAC) and a non-native lipoyl domain from *E. coli* PDHC (Green *et al.*, 1995), which is a hybrid between the N- and C-terminal halves of the first and third lipoyl domain, respectively, and in which the lipoyl-lysine residue is replaced by a glutamine residue. In addition, the X-ray crystal structures of the structurally and functionally related lipoylated H-protein of the glycine decarboxylase system from pea leaves (Pares *et al.*, 1994), and the biotinyl domain of acetyl-coenzyme A carboxylase from *E. coli* (Athappilly & Hendrickson, 1995) have been determined. We will concentrate here only on the comparison of structures of lipoyl domains derived from 2-oxo acid dehydrogenase complexes.

The three lipoyl domain solution structures have all been determined to approximately the same precision, using a comparable number of constraints. The minimised average structures overlay with an r.m.s. deviation of 2.5 Å for the backbone atoms of the *A. vinelandii* OGDHC lipoyl domain and the *B. stearrowthermophilus* PDHC lipoyl domain (Figure 6), and with an r.m.s. deviation of 2.5 Å for the backbone atoms of the *B. stearrowthermophilus* PDHC lipoyl domain and the hybrid *E. coli* PDHC lipoyl domain (Brocklehurst & Perham, 1993). The lipoyl domains thus have essentially the same overall fold and are very similar. We have shown previously (Berg *et al.*, 1995), based on a high similarity of NMR-derived parameters, e.g. chemical shifts, and by the calculation of a preliminary tertiary-structure model, that the structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC is also likely to have a very similar global fold. This suggests again that all lipoyl domains will have similar structures in solution. The determination of the structure of an OGDHC lipoyl domain (this study) now implies that this will be irrespective of the type of complex.

Alignment of the amino acid residue sequences of the lipoyl domains reveals that many conserved or partially conserved residues contribute to the hydrophobic core of the domain (Figure 5). Among the most conserved residues that are not part of the hydrophobic core are the lipoyl-lysine residue and its preceding residue (DK), which reside in the type I β -hairpin turn, and a number of conserved glycine residues that occur at position 3 of a β -turn, which reflects their structural importance. The reason for the high conservation of the exposed Lys24 seems not directly apparent.

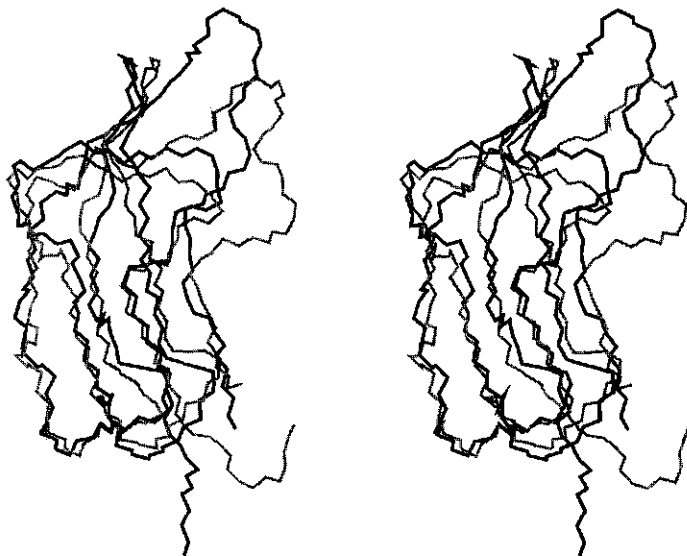


Figure 6. Stereoview of the superposition of the minimised average solution structure of the lipoyl domain of *A. vinelandii* OGDHC (black line) and the lipoyl domain of *B. stearothermophilus* PDHC (grey line). Only the backbone heavy atoms are drawn.

Despite their high structural similarity, there are significant conformational differences among the lipoyl domains, which occur primarily in an exposed loop close in space to the lipoylation site. This loop connects β -strand S1 with S2 (Figure 3) and is one of the least-defined regions in all three lipoyl domain structures. The r.m.s. deviation values of backbone atoms between the different average lipoyl domain structures drop by at least 0.7 Å if this loop is excluded from the alignment. It should be noted, however, that alignment of the isolated loops (residues 7 to 15) of the *A. vinelandii* OGDHC and *B. stearothermophilus* PDHC lipoyl domains reveals that their average structures are less different than expected from alignment of the complete structures, and their backbone atoms align with an r.m.s. deviation of 1.9 Å. This suggests that the loops show some structural similarity but are connected to the rest of the domain in a different way, with the hinge points being residues 7 and 15. The corresponding loop region in the *E. coli* PDHC lipoyl domain comprises two residues less, and its structure seems more different compared to the structures of the other two lipoyl domains.

The exposed loop of the *A. vinelandii* OGDHC lipoyl domain is formed around the side-chain of Phe9, which is almost completely buried, and which makes hydrophobic contacts with the other hydrophobic residues in the loop, i.e. Pro7, Pro10 and Ile13. This

small cluster of hydrophobic side-chains could provide some rigidity to this exposed loop, possibly necessary for its function. In the loops of the *B. stearrowthermophilus* PDHC and the *E. coli* PDHC lipoyl domains also a number of these hydrophobic contacts are found (Brocklehurst & Perham, 1993).

Implications for molecular recognition

The three-dimensional structure of the lipoyl domain is required for an efficient reaction of its lipoyl group with the E1 component (Graham *et al.*, 1989). Free lipoamide is an extremely poor substrate for E1. This is remarkable since the reactive dithiolane ring is at the end of the 14 Å long lipoyl group and protrudes into the solvent. Furthermore, the structure of the lipoyl domain is also responsible, at least in part, for the specificity of the reductive acylation reaction with the E1 component. It has been shown that the 2-oxo acid dehydrogenases (E1p and E1o) from *E. coli* are specific for their lipoyl domains (Graham *et al.*, 1989). We have confirmed this specificity for the 2-oxo acid dehydrogenase complexes from *A. vinelandii* (chapter 6, this thesis). This suggests that some kind of molecular recognition occurs between the E1 and the lipoyl domain.

It is not yet understood why the specificity of reductive acylation of lipoyl domains by only the E1 component of the parent 2-oxo acid dehydrogenase complex is required. Since the lipoyl groups are able to reach to within at least 10 to 20 Å of the outer surface of the multienzyme complex (Hale *et al.*, 1992), there could be a chance of cross-reaction between two different multienzyme complexes if they can approach one another closely enough. However, this cross-reaction will only occur very rarely, if at all, since the rate of inter-core transacetylation is already very slow (Bosma, 1984). We believe that the only reasonable explanation for the necessity of the observed specificity is the possible occurrence *in vivo* of hybrid complexes with respect to their E1 component. The 2-oxoglutarate dehydrogenase complex isolated from *E. coli* was found to be a hybrid complex containing a small amount (~10%) of pyruvate dehydrogenase (Steginsky *et al.*, 1985). The reaction of this E1p with pyruvate and the OGDHC lipoyl domains would rapidly acetylate these lipoyl domains, resulting in an inactive complex. The exclusive reaction of lipoyl domains with only their parent E1 component would prevent such rapid inactivation.

This leaves us with the question of what is the determinant of the lipoyl domain in this process of molecular recognition? Since no structural information is available on any E1 component, at present only a comparison of sequences and structures of lipoyl domains could provide some clues to a possible answer. The comparison is impaired by the lack of information about the differences in specificity among the three lipoyl domain structures

determined so far. It is evident, however, that the local structure around the lipoyl-lysine residue as well as its preceding residue are not involved in the specific recognition process, since these both seem highly conserved in lipoyl domains. Furthermore, it is also clear that the molecular recognition is brought about by small differences in structure and/or charge of the lipoyl domains, since the lipoyl domain structures determined so far are all very similar. The largest structural differences observed among the lipoyl domains occur in the loop region connecting the β -strand S1 with S2 (Figure 3). Alignment of many lipoyl domain amino acid residue sequences reveals that in the first four residues of this loop (residues 7 to 10) the sequence PDIG is frequently found, in particular in lipoyl domains of PDH complexes. However, significant differences in the amino acid residue sequence in this part of the loop are always observed among lipoyl domains of different acyltransferases from the same source. In addition, in this loop occasionally a deletion of one or two residues is found, resulting in loops of different lengths. Furthermore, this loop is at the same side of the lipoyl domain as its lipoyl-lysine residue, and is supposedly somewhat more flexible (low number of NOE distance constraints) than the body of the domain, not uncommon for interaction sites. Together this could suggest that this loop may play a role in the process of molecular recognition of lipoyl domains by only their parent E1 component. Other possible candidates that could be involved in the specific recognition process are the exposed residues at positions 43 and 44 (Val,Val in Av_O, Ala,Val in Bs_P, Ala,Ser in Ec_P1,3 and Av_P1) in the Av_O sequence (Figure 5). As is true for the loop discussed above, these exposed residues are close to the lipoylation site and are nearly always different among lipoyl domains of different acyltransferases from the same source.

Since the sequence similarity among lipoyl domains is not very high, and, as stated earlier, the specificity of E1 components for the three lipoyl domains with resolved structure is not known, it proves difficult as yet to be conclusive about residues or structural motifs of the lipoyl domain responsible for molecular recognition by their E1. The determination of the three-dimensional structure of the *A. vinelandii* OGDHC lipoyl domain provides significant structural information necessary for further studies on the interaction between lipoyl domains and E1, e.g. by NMR spectroscopy and by site-directed mutagenesis. Such studies, which are currently in progress, should aid elucidating which residues of the lipoyl domain are involved in the specific interaction with E1.

MATERIALS AND METHODS

Sample preparation

Unlabelled and uniformly ^{15}N -isotopically labelled lipoyl domain was expressed and purified as reported previously (Berg *et al.*, 1995). The unlipoylated form of the lipoyl domain was used in all experiments. NMR samples contained 8 mM of unlabelled lipoyl domain or 4 mM of ^{15}N -labelled lipoyl domain in 50 mM potassium phosphate (pH 5.5) containing 100 mM potassium chloride in either 90% $\text{H}_2\text{O}/10\%$ D_2O or 99.95% D_2O .

NMR spectroscopy

Details of the NMR experiments, resonance assignments and analysis of the secondary structure of the lipoyl domain have been described elsewhere (Berg *et al.*, 1995). All NMR experiments were performed on a Bruker AMX500 spectrometer, at 14°C and 30°C, and processed using the program FELIX version 2.3 from Biosym Technologies Inc., San Diego.

Derivation of structural constraints

NOE cross-peaks were assigned in an iterative manner, in several rounds of structure calculations. An initial ensemble of structures was calculated using a set of only unambiguously identified NOEs. In each round of calculations a number of ambiguous NOEs were resolved based on analysis of the family of calculated structures, and these were included in the next round of calculations. In the last stage of structure refinement, the repeating pattern of NOEs resulting from the 2-fold internal symmetry of the lipoyl domain also helped to resolve some ambiguous NOEs. The majority of the NOE distance constraints were derived from two-dimensional homonuclear NOESY spectra acquired with mixing times of 70, 100 and 150 ms. In addition, a small number of NOE distance constraints were derived from a 150 ms two-dimensional HMQC-NOESY spectrum. NOE volumes were measured using the FELIX program and calibrated for conversion to approximate interproton distances with the volumes of sequential $d_{\alpha\text{N}}$ cross-peaks in regular antiparallel β -sheet, using the standard distance of 2.2 Å for this interaction (Wüthrich, 1986). NOEs were grouped into strong, medium and weak categories, corresponding to upper-bound interproton distance constraints of 2.7, 3.3 and 5.0 Å, respectively. An additional 0.5 Å was added to the upper limits for distances involving methyl groups (Clore *et al.*, 1987; Wagner *et al.*, 1987). In all cases, the lower-bound distance constraint was set to 1.8 Å, the approximate sum of the van der Waals radii. NOEs which partially overlapped in the spectra were conservatively assigned an upper-bound

constraint of 5.0 Å if their intensity grouping was ambiguous. The standard pseudoatom corrections (Wüthrich *et al.*, 1983) were applied to methyl groups, degenerate diastereotopic hydrogen atoms, and diastereotopic hydrogen groups for which only one NOE was observed. For stereospecifically unassigned diastereotopic substituents for which NOEs to both hydrogen atoms were observed, the DIANA treatment of diastereotopic hydrogen atoms was used (Güntert *et al.*, 1991). In cases where both NOEs had an equal upper distance limit, no pseudoatom correction was applied. In cases where both NOEs had different upper distance limits, the upper distance limit of both NOEs was set equal to the higher limit of the two, and an extra DIANA pseudoatom distance constraint was calculated and added to the constraint list.

Backbone ϕ dihedral angle constraints were derived from $^3J_{\text{HN}\alpha}$ coupling constants measured in a ^1H - ^{15}N HMQC- J spectrum, as described by Berg *et al.* (1995). The ϕ dihedral angles were constrained to $-120(\pm 30)^\circ$ when $^3J_{\text{HN}\alpha} \geq 9$ Hz, and to $-120(\pm 40)^\circ$ when $^3J_{\text{HN}\alpha}$ was between 8 and 9 Hz. For $^3J_{\text{HN}\alpha}$ values smaller than 6 Hz, ϕ angles were constrained to $-55(\pm 30)^\circ$. Stereospecific resonance assignments of methylene C^βH protons and valine residue methyl groups were obtained from $^3J_{\alpha\beta}$ values (from DQF-COSY and E.COSY spectra) and relative intra-residue NOE intensities (Zuiderweg *et al.*, 1985; Wagner *et al.*, 1987). The derived χ_1 angles were constrained to one of the three rotamers -60 , 180 , or 60° , with an allowed range of $\pm 30^\circ$.

Hydrogen-bond constraints were included only in the final round of structure calculations, in a conservative way. A hydrogen-bond constraint was defined only if (1) the amide proton had been shown to exchange slowly in ^1H - ^{15}N HSQC experiments of freshly prepared lipoyl domain in D_2O , (2) the hydrogen bond was present in the majority ($> 90\%$) of the calculated structures, and (3) the hydrogen bond was present in regular antiparallel β -sheet and between backbone atoms only. Two constraints were used for each hydrogen bond, 1.5 to 2.3 Å for the NH-O distance, and 2.5 to 3.3 Å for the N-O distance.

Structure calculations

The first family of three-dimensional structures of the lipoyl domain was calculated using the program DG-II (Havel, 1991) in the NMR refine module of the InsightII package (Biosym Technologies Inc., San Diego). All subsequent structure calculations were performed with the program X-PLOR version 3.1 (Brünger, 1992) using a protocol based on the hybrid distance geometry-dynamical simulated annealing calculation strategy developed by Nilges *et al.* (1988). An improved protocol as described in the X-PLOR version 3.1 manual was used to produce the structures. In the first stage of the protocol a family of embedded substructures was produced with distance geometry using X-PLOR/dg

(Kuszewski *et al.*, 1992). These substructures only included C, C α , C α H, N, NH, C β and C γ atoms. After fitting the remaining atoms to the substructures, a simulated annealing protocol was used to regularise and refine the structures. The first stage of the simulated annealing protocol consisted of 45 ps of dynamics at high temperature (2000 K), thereby gradually introducing the covalent and NOE force constants and lowering the repulsive van der Waals energy term. In the second stage, the repulsive force constant was increased to its final value (see Table 1) during slow cooling of the system to 100 K in 75 ps. Finally, the structures were subjected to 1000 steps of restrained Powell minimisation. Structures were accepted in the final family if showing no distance constraint violations greater than 0.2 Å, no dihedral angle constraint violations greater than 1°, and no bond-length and angle violations greater than 0.05 Å and 5°, respectively. To ensure the uniqueness of the final structures, two other methods, full-structure distance geometry (Kuszewski *et al.*, 1992) and *ab initio* simulated annealing (Nilges *et al.*, 1991) were also applied. All methods produced very similar sets of structures.

A mean structure was obtained by fitting all backbone atoms (residues 2 to 77) of the final set of accepted structures to each other and averaging their atomic coordinates. The mathematical average structure was then minimised by 2000 steps of restrained Powell minimisation.

Molecular graphics analysis of the structures was performed using the InsightII program (Biosym Technologies Inc., San Diego). Analysis of the quality of the structures was performed by tools provided with X-PLOR and with PROCHECK (Laskowski *et al.*, 1993).

The coordinates of both the energy-minimised average structure and the ensemble of 25 structures representing the solution structure of the lipoyl domain, together with the NMR constraints used for their determination, have been deposited with the Protein Data Bank, Brookhaven National Laboratory, Upton, NY, USA. The entry codes are 1GHJ and 1GHK, respectively, for the energy-minimised average structure and the ensemble of 25 structures.

Acknowledgements

We thank Frank Vergeldt for assistance with computing. This work was financially supported by the Netherlands Organisation for Scientific Research (NWO) under the auspices of the Netherlands Foundation for Chemical Research (SON).

REFERENCES

- Athappilly, F. K. & Hendrickson, W. A. (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**, 1407-1419.
- Berg, A., de Kok, A. & Vervoort, J. (1994) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **221**, 87-100.
- Berg, A., Smits, O., de Kok, A. & Vervoort, J. (1995) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex. *Eur. J. Biochem.* **234**, 148-159.
- Bosma, H. J. (1984) *Studies on 2-oxoacid dehydrogenase multienzyme complexes of Azotobacter vinelandii*. Ph.D. Thesis, Agricultural University Wageningen, The Netherlands.
- Brocklehurst, S. M. & Perham, R. N. (1993) Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: a new automated method for the prediction of protein tertiary structure. *Protein Sci.* **2**, 626-639.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187-217.
- Brünger, A. T. (1992) *X-PLOR (Version 3.1): a system for x-ray crystallography and NMR*. Yale University Press, New Haven.
- Chothia, C. & Murzin, A. G. (1993) New folds for all- β proteins. *Structure* **1**, 217-222.
- Clore, G. M., Gronenborn, A. M., Nilges, M. & Ryan, C. A. (1987) Three-dimensional structure of potato carboxypeptidase inhibitor in solution. A study using nuclear magnetic resonance, distance geometry, and restrained molecular dynamics. *Biochemistry* **26**, 8012-8023.
- Cohen-Addad, C., Pares, S., Sieker, L., Neuburger, M. & Douce, R. (1995) The lipamide arm in the glycine decarboxylase complex is not freely swinging. *Nat. Struct. Biol.* **2**, 63-68.
- Dardel, F., Laue, E. D. & Perham, R. N. (1991) Sequence-specific ^1H -NMR assignments and secondary structure of the lipoyl domain of the *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *Eur. J. Biochem.* **201**, 203-209.
- Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.
- Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995) Three-dimensional structure of a lipoyl domain from dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.
- Güntert, P., Braun, W. & Wüthrich, K. (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517-530.
- Hale, G., Wallis, N. G. & Perham, R. N. (1992) Interaction of avidin with the lipoyl domains in the pyruvate dehydrogenase multienzyme complex: three-dimensional location and similarity to biotinyl domains in carboxylases. *Proc. R. Soc. Lond. B* **248**, 247-253.
- Havel, T. F. (1991) An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Prog. Biophys. Molec. Biol.* **56**, 43-78.
- Hyberts, S. G., Goldberg, M. S., Havel, T. F. & Wagner, G. (1992) The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci.* **1**, 736-751.
- Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K. & Perham, R. N. (1993) The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *J. Mol. Biol.* **230**, 323-341.

- Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
- Kuszewski, J., Nilges, M. & Brünger, A. T. (1992) Sampling and efficiency of metric matrix distance geometry: a novel partial metrization algorithm. *J. Biomol. NMR* **2**, 33-56.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
- Ludvigsen, S. & Poulsen, F. M. (1992) Positive ϕ -angles in proteins by nuclear magnetic resonance spectroscopy. *J. Biomol. NMR* **2**, 227-233.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992a) The pyruvate dehydrogenase multienzyme complex. *Curr. Opin. Struct. Biol.* **2**, 877-887.
- Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1992b) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
- Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1993) Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Ångstrom resolution. *J. Mol. Biol.* **230**, 1183-1199.
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins: Struct. Funct. Genet.* **12**, 345-364.
- Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988) Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* **229**, 317-324.
- Nilges, M., Kuszewski, J. & Brünger, A. T. (1991) Sampling properties of simulated annealing and distance geometry. In: *Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy* (Hoch, J. C., et al., eds.), pp. 451-455, Plenum Press, New York.
- Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. & Douce, R. (1994) X-ray structure determination at 2.6-Å resolution of a lipooate-containing protein: the H-protein of the glycine decarboxylase complex from pea leaves. *Proc. Natl. Acad. Sci. USA* **91**, 4850-4853.
- Perham, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.
- Reed, L. J., Koike, M., Levitch, M. E. & Leach, F. R. (1958) Studies on the nature and reactions of protein-bound lipoic acid. *J. Biol. Chem.* **232**, 143-158.
- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.
- Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K. & Gronenborn, A. M. (1992) Three-dimensional solution structure of the E3-binding domain of the dihydrolipoamide succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **31**, 3463-3471.
- Steginsky, C. A., Gruys, K. J. & Frey, P. A. (1985) α -Ketoglutarate dehydrogenase complex of *Escherichia coli*. A hybrid complex containing pyruvate dehydrogenase subunits from pyruvate dehydrogenase complex. *J. Biol. Chem.* **260**, 13690-13693.
- Wagner, G., Braun, W., Havel, T. H., Schaumann, T., Go, N. & Wüthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* **196**, 611-639.
- Wallis, N. G. & Perham, R. N. (1994) Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**, 209-216.
- Westphal, A. H. & de Kok, A. (1990) The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur. J. Biochem.* **187**, 235-239.
- Wüthrich, K., Billeter, M. & Braun, W. (1983) Pseudostructures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949-961.
- Wüthrich, K. (1986) *NMR of proteins and nucleic acids*. Wiley, New York.
- Zuiderweg, E. R. P., Boelens, R. & Kaptein, R. (1985) Stereospecific assignments of $^1\text{H-NMR}$ methyl lines and conformation of valyl residues in the lac repressor headpiece. *Biopolymers* **24**, 601-611.

CHAPTER 5

Three-dimensional structure in solution of the N-terminal lipoyl domain of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*

ABSTRACT

The three-dimensional structure of the N-terminal lipoyl domain of the acetyltransferase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii* has been determined using heteronuclear multidimensional NMR spectroscopy and dynamical simulated annealing. The structure is compared with the solution structure of the lipoyl domain of the *A. vinelandii* 2-oxoglutarate dehydrogenase complex. The overall fold of the two structures, described as a β -barrel-sandwich hybrid, is very similar. This agrees well with the high similarity of NMR-derived parameters, e.g. chemical shifts, between the two lipoyl domains. The main structural differences between the two lipoyl domains occur in a solvent-exposed loop close in space to the lipoylation site. Despite their high structural similarity, these lipoyl domains show a high preference for being reductively acylated by their parent 2-oxo acid dehydrogenase. Potential residues of the lipoyl domain involved in this process of molecular recognition are discussed.

INTRODUCTION

The pyruvate dehydrogenase complex (PDHC) catalyses the irreversible overall conversion of pyruvate to acetyl-CoA [for a recent review see Mattevi *et al.*, (1992a)]. This multienzyme complex plays an important regulating role in the aerobic catabolism of carbohydrates, where it links the glycolysis with the citric acid cycle. Pyruvate dehydrogenase complexes from Gram-negative bacteria are composed of multiple copies of three enzymatic components: pyruvate dehydrogenase (E1p), dihydrolipoyl acetyltransferase (E2p), and lipoamide dehydrogenase (E3). The structural core of these complexes consists of a cubic assemblage of 24 E2p chains arranged with octahedral symmetry. The E1p and E3 components are tightly but non-covalently bound to the E2p core as dimers. The 2-oxoglutarate dehydrogenase complex (OGDHC), which occurs in the citric acid cycle, belongs to the same family of 2-oxo acid dehydrogenase complexes as the PDHC, as do the branched-chain 2-oxo acid dehydrogenase complexes. These complexes all have similar structural and catalytic properties.

The chain of the acetyltransferase component (E2p) of PDHC from *A. vinelandii* is, like other E2 components, highly segmented. It consists of five independently folded domains that are separated by long stretches (25 to 40 amino acids) of polypeptide chain which are flexible and unusually rich in alanine and proline residues (Hanemaaijer *et al.*, 1988). From the N-terminus the E2p chain consists of three highly homologous lipoyl domains (about 80 residues) that each contain one covalently bound lipoic acid prosthetic group, a peripheral subunit-binding domain (about 35 residues), and a catalytic domain (29 kDa) which accommodates the acetyltransferase active site and is responsible for the formation of the core of the complex (Hanemaaijer *et al.*, 1987).

The lipoyl domains attached to the flexible linkers fulfil an indispensable role in coupling the three separate enzyme activities in the multienzyme complex (Reed, 1974). The single oxidised lipoyl group of each lipoyl domain, which consists of lipoic acid bound in an amide linkage to the side chain of a specific lysine residue, becomes reductively acetylated by pyruvate catalysed by E1p. The acetyl group is delivered at the active site of E2p to form acetyl-CoA, after which the lipoyl domain visits the active site of E3 where it becomes re-oxidised. The three-dimensional structure of the lipoyl domain is required for the efficient reductive acetylation of its lipoyl group by E1p (Graham *et al.*, 1989). Moreover, lipoyl domains can only be efficiently reductively acylated by an E1 of their parent complex (Graham *et al.*, 1989), which indicates that molecular recognition occurs between E1 components and lipoyl domains.

So far any 2-oxo acid dehydrogenase complex or any of its acyltransferase components refused to crystallise to an appreciable resolution, most likely caused by the high flexibility of their lipoyl domains and/or linkers. Therefore we have undertaken, among others, the approach of structure determination of the individual components and the separate domains of the acyltransferase component of these large multienzyme complexes. For the *A. vinelandii* 2-oxo acid dehydrogenase complexes this has resulted in the unique crystal structure of the catalytic cubic core domain of E2p (Mattevi *et al.*, 1992b, 1993), the crystal structure of lipoamide dehydrogenase (Schierbeek *et al.*, 1989; Mattevi *et al.*, 1991), and the solution structure of the single lipoyl domain of OGDHC (Berg *et al.*, 1996). The solution structures of the lipoyl domain of *Bacillus stearothermophilus* PDHC (Dardel *et al.*, 1993), an inactive hybrid lipoyl domain from *Escherichia coli* (Green *et al.*, 1995), and the peripheral subunit-binding domains of dihydrolipoyl succinyltransferase (E2o) of *E. coli* OGDHC (Robien *et al.*, 1992) and of E2p of *B. stearothermophilus* PDHC (Kalia *et al.*, 1993) have also been solved. No structural information on any E1 component is yet available.

In this paper we describe the determination of the three-dimensional solution structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC. We have recently reported on the subcloning, overexpression, ¹H and ¹⁵N NMR assignments and secondary structure of this lipoyl domain (Berg *et al.*, 1994), as well as a comparison of NMR-derived parameters with those of the lipoyl domain of *A. vinelandii* OGDHC (Berg *et al.*, 1995). From this study it became clear that the overall fold of the PDHC and OGDHC lipoyl domains must be very similar. This implies that the observation that lipoyl domains are only efficiently reductively acylated by the E1 component of the parent complex is a result of subtle differences between the lipoyl domains. The aim of this work is to gain a better understanding at the atomic level of this process of molecular recognition by comparing the three-dimensional structures of both lipoyl domains. Furthermore, elucidation of the structure of the PDHC lipoyl domain provides the basic structural information for the design of significant mutants to investigate the specific interaction of this domain with E1, and provides naturally a further step in the completion of the three-dimensional structure of the entire pyruvate dehydrogenase multienzyme complex from *A. vinelandii*.

RESULTS

Structure determination and analysis

The calculation of the final set of structures was based on a total of 474 non-trivial NOE distance constraints, 26 hydrogen-bond distance constraints (derived from 13 hydrogen bonds), and 66 dihedral angle constraints (53 ϕ and 13 χ_1 angles). The NOE distance constraints comprised 156 intra-residue, 161 sequential, 38 short-range ($1 < |i-j| < 5$), and 119 long-range constraints ($|i-j| \geq 5$). The distribution of distance constraints (NOEs plus hydrogen bonds) along the amino acid sequence is shown in Figure 1(a). The absence of any aromatic residues in the lipoyl domain results in a relatively poor chemical shift dispersion in the NMR spectra, in particular in the aliphatic regions of the spectra. Therefore, a number of NOEs remained ambiguous, even after several rounds of structure calculations. For example, a number of resonances could be assigned belonging to a certain amino acid residue, but their exact position in the side chain could not be determined unambiguously (Berg *et al.*, 1994). Despite thorough inspection of the NMR spectra, including a 3D NOESY-HMQC spectrum, the number of unambiguous long-range NOEs that could be identified was relatively low.

A total of 50 structures was generated using *ab initio* simulated annealing in X-PLOR, and of these, 29 structures were selected based on experimental constraint satisfaction criteria only. These structures show no distance constraint violations $> 0.2 \text{ \AA}$ ($1 \text{ \AA} = 0.1 \text{ nm}$) and no dihedral angle violations $> 1^\circ$. A summary of structural statistics for the ensemble of 29 structures is given in Table 1. A stereoview of a best fit superposition (using backbone atoms of residues 1 to 73) of the 29 SA structures is shown in Figure 2. In combination with Figure 1(b), showing the r.m.s. deviation of the 29 structures from the mathematical average structure for backbone and all heavy atoms as a function of the residue number, it can be seen that the structures show reasonable convergence. The last six residues at the C-terminus are poorly defined, and exclusion of these residues results in r.m.s. deviation values of the ensemble from the mean structure of 0.87 \AA for backbone heavy atoms and 1.27 \AA for all heavy atoms (Table 1). For the disordered C-terminal residues no non-trivial NOEs could be observed. It is noteworthy that the linewidth of the NH resonances from the C-terminal residues is considerably narrower than those from residues of the rest of the lipoyl domain (Berg *et al.*, 1994), implying a higher mobility for these residues.

Pyruvate dehydrogenase complex lipoyl domain structure

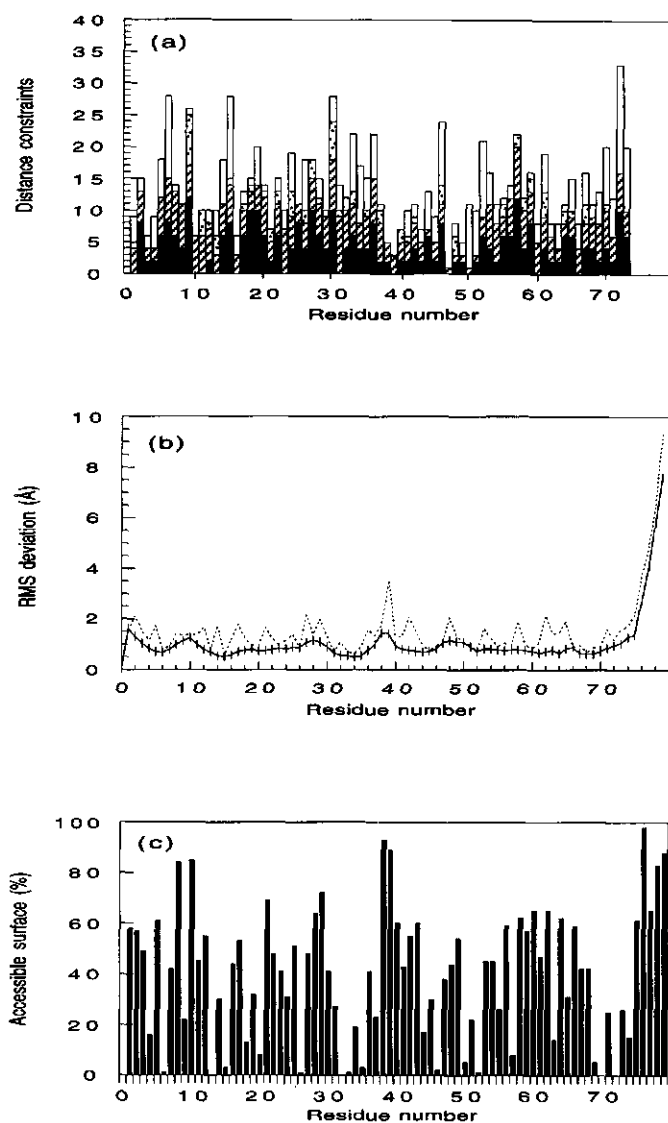


Figure 1. (a) Plot of the number of experimentally derived distance constraints (NOEs plus hydrogen bonds) as a function of the residue number. The bars representing the different distance constraint categories are as follows: intra-residue (filled), sequential (hatched), medium range (stippled), and long range (open). Each inter-residue distance constraint is counted twice, once for each residue involved. (b) Plot of the residue-based r.m.s. deviation of the 29 individual simulated annealing structures from the mean-structure $(SA)_T$ of backbone heavy atoms (continuous line) and all heavy atoms (broken line). (c) Plot of the percentage of residue surface exposed to solvent (all atoms) of the minimised average structure.

Table 1. Structural statistics for the 29 converged structures of the lipoyl domain^a

	< SA >	(SA) _r
R.m.s. deviations from experimental		
distance constraints (Å) ^b :		
All NOE constraints (474)	0.0068(±0.002)	0.0045
Intra-residue (156)	0.0092(±0.003)	0.0067
Sequential (li-jl) = 1 (161)	0.0053(±0.002)	0.0038
Short range (li-jl) < 5 (38)	0.00042(±0.0010)	0.0049
Long range (li-jl) ≥ 5 (119)	0.0035(±0.001)	0.0012
Hydrogen bond (26)	0.0007(±0.0007)	0.0001
R.m.s. deviations from experimental		
dihedral constraints (°) (66)	0.057(±0.03)	0.034
R.m.s. deviations from idealised covalent		
geometry:		
Bonds (Å)	0.0019(±0.0001)	0.0018
Angles (°)	0.544(±0.001)	0.538
Impropers (°)	0.317(±0.003)	0.317
Energy (kcal/mol) ^c :		
F _{NOE}	1.381(±0.724)	0.565
F _{cdih}	0.017(±0.018)	4.745
F _{repel}	1.865(±0.363)	1.384
Atomic r.m.s. differences (Å)		
	backbone ^d	heavy
Residues 1 to 79, to mean (SA) _r	1.45(±0.36)	1.91(±0.37)
Residues 1 to 73, to mean (SA) _r	0.87(±0.18)	1.27(±0.15)
Residues 1 to 79, pairwise	2.07(±0.57)	2.74(±0.59)
Residues 1 to 73, pairwise	0.97(±0.19)	1.62(±0.17)

^a < SA > refers to the 29 final structures obtained by the simulated annealing protocol, (SA)_r is the restrained minimised mean structure obtained by averaging the atomic coordinates of the final structures best fitted to each other over the backbone atoms (residues 1 to 73).

^b The number of each type of constraint used in the structure calculations is given in parentheses.

^c The force constants used for these calculations were 50 kcal/mol per Å², and 200 kcal/mol per rad², for F_{NOE} and F_{cdih}, respectively. F_{repel} was calculated using the final value of 4 kcal/mol per Å⁴ with the van der Waals hard sphere radii set to 0.75 times the standard values used in the CHARMM empirical energy function (Brooks *et al.*, 1983). The paralhhdg.pro parameter set supplied with X-PLOR version 3.1 (Brünger, 1992) was used for the calculations.

^d These include the backbone heavy atoms N, C^α, and C only.

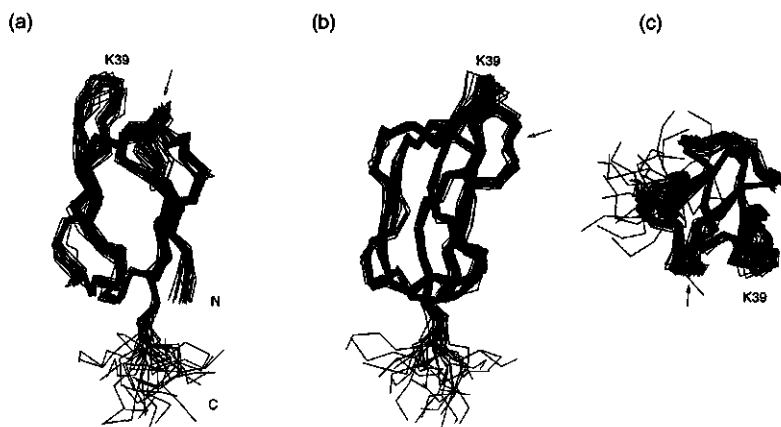


Figure 2. (a) View of the 29 selected structures of the lipoyl domain, superposed over the backbone (N, C α , C) atoms for residues 1 to 73. Only the C α atoms are shown. The N-terminus and C-terminus, the lipoylation site (K39) and the solvent-exposed loop close in space to the lipoylation site (arrow), are indicated (b) Side view of the superposition shown in (a), obtained by a 90° rotation about the vertical axis. (c) Front view at the lipoylation site (K39) and the solvent-exposed loop.

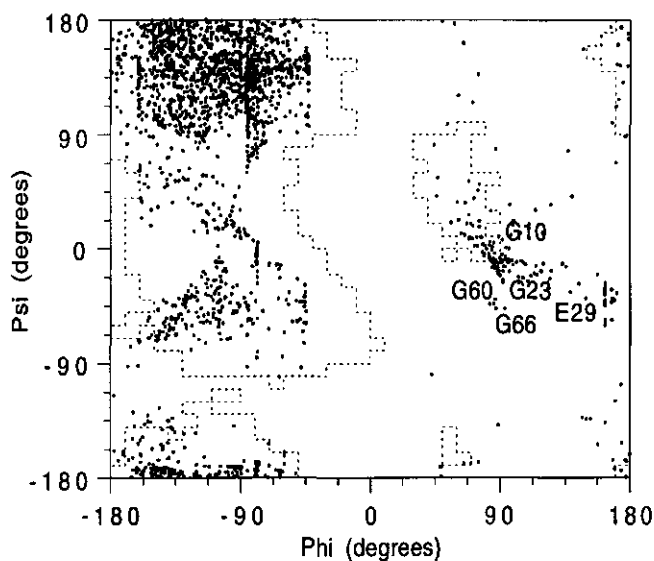


Figure 3. Ramachandran plot for the non-terminal residues in the 29 final structures of the lipoyl domain. Residues with positive ϕ angles are labelled.

A Ramachandran plot for all 29 structures in the ensemble is depicted in Figure 3. Most non-glycine residues lie within the sterically allowed regions (Morris *et al.*, 1992). The six C-terminal residues show backbone dihedral angle (ϕ and ψ) combinations in all regions of the Ramachandran plot, and have consequently very poorly defined angles (average $S^{\phi,\psi} < 0.5$) (Hyberts *et al.*, 1992). The only non-glycine residue showing a positive ϕ angle is Glu29. This residue lies in the generously allowed region, and has reasonably high angular order parameters ($S^{\phi} = 0.86$, $S^{\psi} = 0.96$). For Glu29 no dihedral angle constraint was applied since a ${}^3J_{\text{HN}\alpha}$ coupling constant between 6 and 8 Hz was measured. Furthermore a very strong intra-residue NH-C α H NOE was observed for this residue. Both these observations are consistent with residues showing positive ϕ angles (Ludvigsen & Poulsen, 1992).

Description of the structure

The ensemble of 29 structures representing the solution structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC is shown in Figure 2. A schematic representation of the minimised average structure is illustrated in Figure 4. The overall fold of the lipoyl domain can be described as a β -barrel-sandwich-hybrid (Chothia & Murzin, 1993). The domain is an all- β -sheet protein, and consists of two very similar four-stranded antiparallel β -sheets with three major and one minor strand each, that are formed around a core of hydrophobic residues. One β -sheet (sheet A) is formed by the β -strands S1, S3, S6 and S8, and the other β -sheet (sheet B) is formed by the strands S2, S4, S5 and S7, with β -strands S3 and S7 being the minor strands. The β -strands are alternated and connected by loop regions and turns adopting no regular repetitive secondary structure. In one of the sheets the successive strands S4 and S5 are connected by a type I β -turn comprising residues Ser37 to Ala40. The turn holds the lipoyl-lysine residue Lys39 at position 3, which has been proved to be the crucial position for this residue to become lipoylated (Wallis & Perham, 1994). The strands S1 and S2 are connected by a short five-residue loop (Pro7 to Gly11) which is solvent exposed and lies in the vicinity of the lipoylation site. This loop contains a non-classical tight turn of type $\alpha\gamma$ (Wilmot & Thornton, 1990). The two minor strands (S3 and S7) are both surrounded by small loop regions that each contain a type II β -turn, and that link them to the larger strands. The four turns all show a small (< 6 Hz) ${}^3J_{\text{HN}\alpha}$ coupling constant for residues at position 2 and slow amide-proton exchange for residues at position 4 (Berg *et al.*, 1994, 1995).

The lipoyl domain displays a remarkable internal symmetry with a 2-fold rotational axis. The two halves of the molecule (residues 13 to 36 and residues 50 to 73) align with

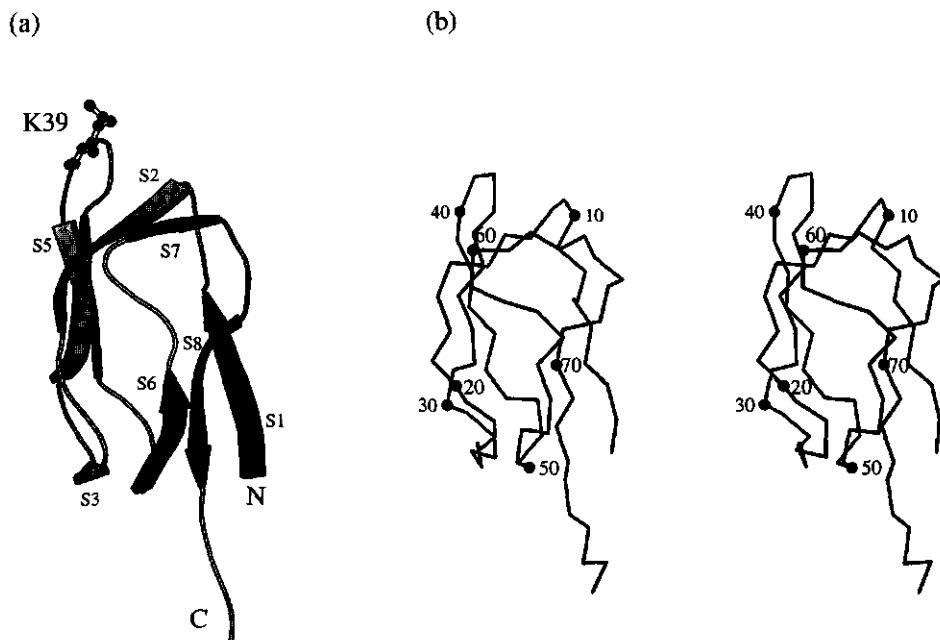


Figure 4. (a) MOLSCRIPT representation (Kraulis, 1991) of the minimised average structure of the lipoyl domain. The β -strands are labelled with their strand numbers. The N and C termini are close together in β -sheet A (dark grey), the lipoylation site Lys39 is situated in the other β -sheet B (light grey), at the opposite side of the domain. (b) Stereoview of the C α trace of the lipoyl domain, with every tenth residue numbered.

an r.m.s. deviation of only 0.72 Å for backbone atoms after rotation of approximately 180° about this axis. This symmetry seems conserved in all lipoyl domains.

The core of the lipoyl domain is formed mainly by hydrophobic side chains of residues of the two β -sheets. Sheet A contributes eight residues (Ile4, Val6, Ile26, Val52, Val55, Ile69, Ile70 and Leu72), and the sheet B contributes seven residues (Val15, Leu18, Leu32, Val33, Leu35, Val44 and Leu63) to the hydrophobic core. In addition, the side chains of three residues (Val20, Pro47, Val57), which reside in regions connecting the β -strands, protrude into the core, and these residues show hydrophobic contacts to core residues in both sheets. The absence of aromatic residues in the hydrophobic core is remarkable and unusual, but seems to have no effectual consequence for the stability of the domain. Most residues in the core have a well-defined side-chain conformation (Figure 1a) and a low side-chain solvent accessible surface (Figure 1c). The six C-terminal residues of the subcloned fragment are disordered and are not involved in the global fold of the domain. These residues belong to the first mobile Ala+Pro-rich linker sequence separating the N-terminal lipoyl domain from the second lipoyl domain in the E2p chain.

DISCUSSION

Lipoyl domains are essential for the optimal functioning of the 2-oxo acid dehydrogenase complexes. By acting as substrate carrying domains they couple the activities of the three enzyme components in the multienzyme complex. Being coupled to the flexible linkers, the lipoyl domains are capable of moving rapidly from active site to active site while keeping substrate diffusion limited and thus increasing the local substrate concentration significantly. The lipoyl domains with their covalently coupled lipoyl group are as such a substrate for all three enzyme components in the complex. With the lipoyl group in the oxidised form the lipoyl domain is a substrate for the E1 component, in the acylated form a substrate for the E2 component, and in the reduced form a substrate for the E3 component.

The structure of the lipoyl domain not only provides a specific point of attachment for the lipoyl group, but interestingly enhances the reaction efficiency with the E1 component dramatically (10.000 times) (Graham *et al.*, 1989). Free lipoamide, as well as a lipoylated decapeptide with an amino acid sequence identical to that found surrounding the lipoylation site of *E. coli* PDHC lipoyl domains, are extremely poor substrates for E1. Although free lipoamide proves a good substrate for the E2 and E3 components, it cannot be ruled out that attachment of the protein domain to the lipoyl group also increases the reaction rate with these components. Recently, an indication of such an enhancement for the E3 reaction has been reported in a paper by Ravindran *et al.* (1996). Furthermore, the structure of the lipoyl domain is also responsible, at least in part, for the specificity of the reductive acylation reaction with the E1 component. Lipoyl domains are only being efficiently reductively acylated by the E1 component of their parent complex. This means, for the *A. vinelandii* and *E. coli* complexes, that PDHC lipoyl domains are a good substrate for E1p but not for E1o, and that OGDHC lipoyl domains are only a good substrate for E1o and not for E1p. Specificity of E2 components for their lipoyl domains has not been reported but seems not likely, since lipoyl domains can be selected by the E2 component on the basis of the acyl-group that is bound to the lipoyl group. Considering that PDHC and OGDHC share a common E3 component, specificity is not involved in the reaction of lipoyl domains catalysed by E3. Thus, the three-dimensional structure of the lipoyl domain is very important because it is responsible for both an efficient and a specific reaction of the bound lipoic acid prosthetic group with the E1 component. This is remarkable since the reactive dithiolane ring of this prosthetic group is at the far end of the long lipoyl group and protrudes into the solvent.

Several solution structures of lipoyl domains have recently been described. These are the PDHC lipoyl domains from *B. stearothermophilus* (Dardel *et al.*, 1993) and *E. coli* (Green *et al.*, 1995), and the OGDHC lipoyl domain from *A. vinelandii* (Berg *et al.*, 1996). In addition, the X-ray crystal structures of the sequentially and functionally related lipoylated H-protein of the glycine decarboxylase system from pea leaves (Pares *et al.*, 1994), and the biotinyl domain of acetyl-coenzyme A carboxylase from *E. coli* (Athappilly & Hendrickson, 1995) have been determined. The general fold of all these domains/proteins is similar and has been described as a new class of all β -sheet proteins called β -barrel-sandwich hybrids, and is thought to be typical for proteins containing a lipoyl or biotinyl group. The solution structure of the lipoyl domain of *A. vinelandii* PDHC, described in the present study, is no exception and adopts a similar overall fold. Many of the common structural features of lipoyl domains and related proteins, like the topology and arrangement of the β -strands, the hairpin turn containing the lipoylated or biotinylated lysine residue, and the internal 2-fold symmetry, have been reviewed earlier in different papers describing these structures. We will concentrate here mainly on the comparison of the structures of the *A. vinelandii* PDHC and OGDHC lipoyl domains, for reasons outlined below.

A comparison of the secondary chemical shifts and other NMR-derived parameters (like $^3J_{\text{HN}\alpha}$ coupling constants and NH-exchange rates) between the PDHC and OGDHC lipoyl domain presented earlier (Berg *et al.*, 1995), indicates that there are no major structural differences in solution between the two domains. A comparison of the PDHC and the OGDHC lipoyl domain structures, together with their superposition, is shown in Figure 5. It is immediately clear that the two structures are strikingly similar. All secondary structure elements, as well as the far majority of residues forming the hydrophobic core, are highly conserved (Figure 6). The positions of the β -strands in both structures are very similar, including the occurrence of four classic β -bulges within them. The positions of the four type II turns surrounding the two small β -strands, and the position of the type I turn that holds the lipoyl-lysine residue, are also equivalent. The minimised average PDHC and OGDHC lipoyl domain structures overlay with an r.m.s. deviation of only 1.5 Å for C α atoms (residues 1 to 6, 12 to 73 for PDHC, residues 1 to 6, 15 to 76 for OGDHC), despite the relatively low sequence identity of 25%.

It is clear from the structure-based sequence alignment in Figure 6 that the second and the third lipoyl domain of PDHC (Av_p2,3) will have highly similar structures to the N-terminal PDHC lipoyl domain, which is of course expected on the basis of their high sequence conservation and their identical function in the same multienzyme complex. In all lipoyl domains known so far, a strictly conserved aspartic acid residue precedes the

lipoyl-lysine residue, with the single exception of the N-terminal lipoyl domain of *A. vinelandii* PDHC (this study), where an alanine is found in this position. The second and third lipoyl domain of *A. vinelandii* PDHC thus also contain an aspartic acid at this position.

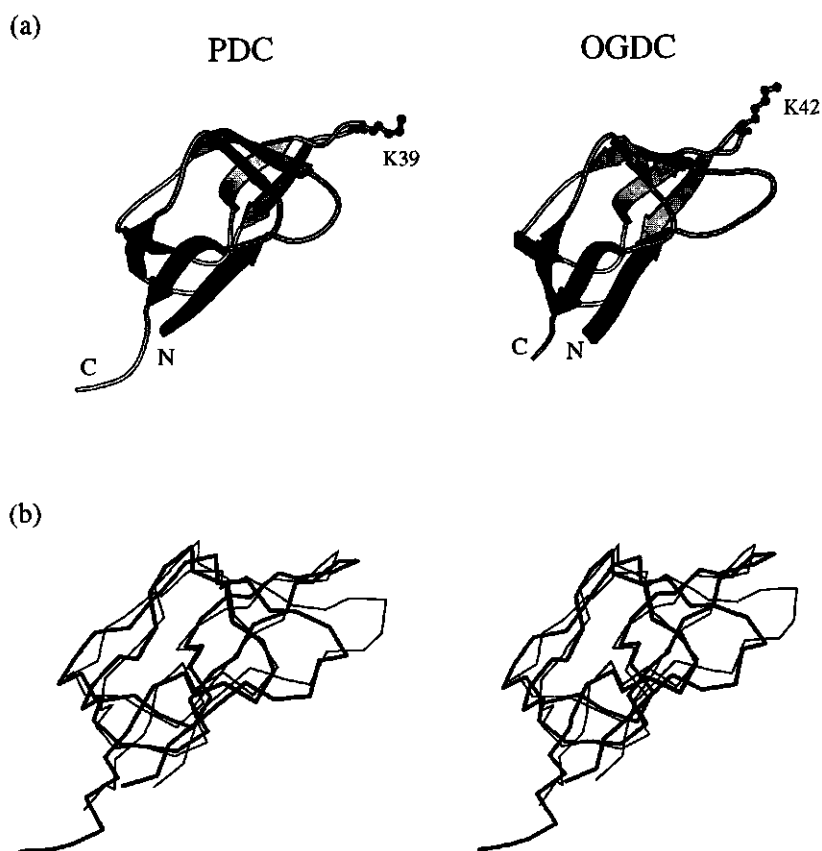


Figure 5. (a) Comparison of the schematic drawings of the N-terminal lipoyl domain of PDHC and the lipoyl domain of OGDHC (Berg *et al.*, 1996), from *A. vinelandii*. (b) Stereoview of the superposition of the minimised averaged solution structures of the *A. vinelandii* PDHC lipoyl domain (thick line) and the OGDHC lipoyl domain (thin line).

A comparison of the structure of the N-terminal lipoyl domain of *A. vinelandii* PDHC with other lipoyl domain structures shows that this substitution does not alter the conformation of the turn (type I) which holds the lipoyl-lysine residue, and in which this residue takes position 2. This lipoyl domain, when separately isolated from the complex, can still be

efficiently reductively acetylated by E1p (Berg *et al.*, 1994). It cannot be ruled out, however, that the aspartic acid to alanine residue substitution has an effect on the rate of this reaction. Site-directed mutagenesis experiments on the single lipoyl domain of *B. stearothermophilus* PDHC (Wallis & Perham, 1994) showed that this substitution resulted in a decrease of the initial rate of reductive acetylation to 37% of that of the wild-type domain.

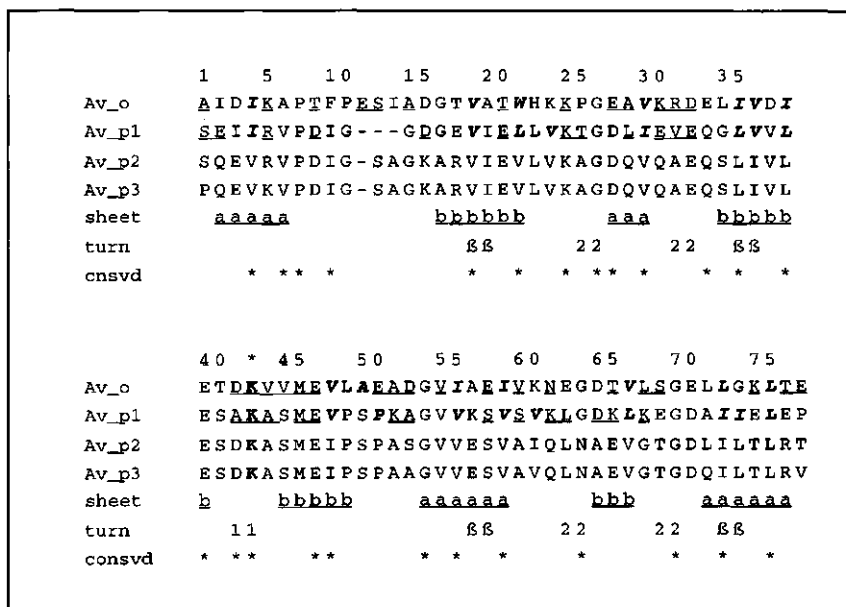


Figure 6. Structure-based sequence alignment of the lipoyl domains of PDHC (Av_p1-3) (Hanemaaijer *et al.*, 1988) and OGDHC (Av_o) (Westphal & de Kok, 1990) from *A. vinelandii*. Residues belonging to the hydrophobic core of the lipoyl domain are in bold italic in the sequence. Exposed residues (more than 50% of side-chain surface accessible to solvent) are underlined. The lipoyl-lysine residue is in bold. The locations of the β -strands are indicated in the line labelled "sheet", where a represents β -sheet A, and b represents β -sheet B. In the line labelled "turn" the conserved positions of type I (1) and type II (2) turns and β -bulges (β) in the sequence are indicated. The bottom-line "consvd" indicates conserved or semi-conserved amino acid residues in all lipoyl domains.

It is interesting to compare lipoyl domain structures and amino acid sequences in search for possible residues or structural motifs involved in the process of molecular recognition of lipoyl domains by E1 components. This is however not as easy as it seems at first sight. First, the specificity of reductive acylation of lipoyl domains by the E1 components of their parent complex has been shown only for PDHC and OGDHC from *E. coli* (Graham *et al.*, 1989) and *A. vinelandii* (chapter 6, this thesis). We assume that this

specificity extends to 2-oxo acid dehydrogenase complexes from other sources. Although several structures of lipoyl domains have recently been solved, we do not know the specificity of these lipoyl domains for the E1 components of the various complexes, except for the two lipoyl domains of the complexes from *A. vinelandii*. This limits the structural comparison of the *B. stearothermophilus* PDHC lipoyl domain and the *E. coli* PDHC lipoyl domain with each other and with the *A. vinelandii* lipoyl domains. In the search for the determinant(s) of the lipoyl domain involved in molecular recognition we will therefore focus mainly on the structural comparison of the two *A. vinelandii* lipoyl domains and on the comparison of a large number of amino acid sequences of lipoyl domains that are available. In particular, amino acid residues of lipoyl domains that are always different among different complexes from the same source, but are identical among the lipoyl domains (if more than one) within one complex of a certain source, will be considered. Furthermore, we assume that amino acid residues of the lipoyl domain that are involved in the specific reaction with E1 are exposed and likely reside on the same side (half) of the domain where the lipoyl-lysine residue is found. The high structural similarity between the *A. vinelandii* PDHC and OGDHC lipoyl domains, and among these and other lipoyl domains, implies that molecular recognition of lipoyl domains by their parent E1's is due to relative small differences in structure and/or charge or hydrophobicity among the lipoyl domains.

Keeping all the above mentioned considerations in mind, a number of candidate amino acid residues involved in molecular recognition can be proposed. As has been suggested earlier (Brocklehurst & Perham, 1993; Dardel *et al.*, 1993; Berg *et al.*, 1994, 1995, 1996), one of regions that could be involved is a solvent-exposed loop that connects β -strand S1 with S2 (Figure 4), and which is close in space to the lipoylation site. The largest structural differences that are found among the different lipoyl domains occur in this loop. Although three amino acid residues in the first part of this loop (residues 8 to 10) seem conserved among lipoyl domains (very frequently the amino acid sequence DIG is found in PDHC lipoyl domains), this sequence is usually different in lipoyl domains of different acyltransferases from the same source. This is also true for the second part of this loop (residues 11 to 15, Figure 6). However, in the case of the N-terminal lipoyl domain of *A. vinelandii* PDHC (this study), the second part of the loop comprises two residues less than the two other lipoyl domains of the same E2p chain. This implies that the second part of the exposed loop would be less likely to participate in the specific interaction with E1 than would the first part. The question whether the structure and/or certain loop residues could be important determinants of the lipoyl domain in the molecular recognition process can however not be answered unequivocally at this stage.

There are several other exposed amino acid residues that could very well be suggested to be involved in the specific recognition process, based on the conditions stated earlier. These are the two amino acid residues succeeding to the lipoyl-lysine residue (Ala40, Ser41 in Av_p1 and Val43, Val44 in Av_o, see Figure 6) (Berg *et al.*, 1996), and the residue at position 3 of the type II turn between β -strands S3 and S4 (Glu29 in Av_p1 and Asp32 in Av_o). Of these, the last amino acid residue may be less likely than the two residues in the immediate vicinity of the lipoylation site, since this residue is at a distance of at least 15 Å from the lipoylation site. Site-directed mutagenesis of these amino acid residues including residues in the exposed loop (see above) should help elucidate if these residues are responsible for the specific interaction of lipoyl domains with the E1 component of their parent complex.

The question of what determines the specificity of the reductive acylation reaction cannot be uncoupled from the question of what is important for the dramatic enhancement of efficiency of reductive acylation of lipoamide when coupled to the lipoyl domain. Although it has already been suggested that the lipoyl domain does not promote reductive acylation of lipoamide just by enhanced binding to the E1 component (Graham & Perham, 1990), the lack of any structural information on E1 severely impedes detailed studies on this subject at the moment. Nonetheless, the determination of the three-dimensional structure of the lipoyl domain is undeniably a prerequisite for all studies concerning the (specific) interaction of lipoyl domains with E1 components.

MATERIALS AND METHODS

NMR experiments

The unlipoylated form of the lipoyl domain was used in all NMR experiments. The preparation of the NMR samples and most of the NMR experiments have been described previously (Berg *et al.*, 1994, 1995). NMR samples contained typically 7 mM unlabelled lipoyl domain or 5 mM uniformly ^{15}N -labelled lipoyl domain. A three-dimensional ^{15}N NOESY-HMQC spectrum (Marion *et al.*, 1989) was recorded at 40°C on a Bruker AM600 spectrometer, with 256(t_1) x 96(t_2 , ^{15}N) x 1024(t_3) points and a NOESY mixing time of 100 ms. NMR spectra were processed using the program FELIX version 2.3 from Biosym Technologies Inc., San diego.

Derivation of structural constraints

The majority of the NOE distance constraints was derived from two-dimensional homonuclear NOESY spectra acquired with mixing times of 25, 50, 100 and 150 ms. Approximately 60 NOE distance constraints were derived from a 100 ms three-dimensional NOESY-HMQC spectrum. NOE volumes were calibrated for conversion to approximate interproton distances with the volumes of sequential $d_{\alpha N}$ cross-peaks in regular antiparallel β -sheet. NOEs were classified as strong, medium or weak, and converted into upper-bound interproton distance constraints of 2.7, 3.3 and 5.0 Å, respectively. For distances involving methyl groups an additional 0.5 Å was added to the upper limits (Clore *et al.*, 1987; Wagner *et al.*, 1987). The lower-bound distance constraint was set to 0.0 Å in all cases (Hommel *et al.*, 1992). Partially overlapping NOEs were conservatively given an upper-bound constraint of 5.0 Å if their intensity grouping was ambiguous. The standard pseudoatom corrections (Wüthrich *et al.*, 1983) were applied to methyl groups, degenerate diastereotopic hydrogen atoms, and diastereotopic hydrogen groups for which only one NOE was observed. For stereospecifically unassigned diastereotopic substituents for which NOEs to both hydrogen atoms were observed, the DIANA treatment of diastereotopic hydrogen atoms was used (Güntert *et al.*, 1991). In cases where both NOEs had an equal upper distance limit, no pseudoatom correction was applied. In cases where both NOEs had different upper distance limits, the upper distance limit of both NOEs was set equal to the higher limit of the two, and an extra DIANA pseudoatom distance constraint was calculated and added to the constraint list.

Backbone ϕ dihedral angle constraints were derived from $^3J_{HN\alpha}$ coupling constants measured in a 1H - ^{15}N HMQC- J spectrum, as described in Berg *et al.* (1995). The ϕ dihedral angles were constrained to $-120(\pm 30)^\circ$ for $^3J_{HN\alpha} \geq 9$ Hz, to $-120(\pm 40)^\circ$ for $^3J_{HN\alpha}$ between 8 and 9 Hz, and to $-55(\pm 30)^\circ$ for $^3J_{HN\alpha}$ values smaller than 6 Hz. Stereospecific resonance assignments of methylene $C^\beta H$ protons and valine residue methyl groups were obtained from $^3J_{\alpha\beta}$ values (from DQF-COSY and E.COSY spectra) and relative intra-residue NOE intensities (Zuiderweg *et al.*, 1985; Wagner *et al.*, 1987). The derived χ_1 angles were constrained to one of the three rotamers -60 , 180 , or 60° , with a range of $\pm 30^\circ$.

Hydrogen-bond constraints were included only in the later rounds of structure calculations, in a conservative way. A hydrogen-bond constraint was defined only if (1) the amide proton had been shown to exchange slowly in 1H - ^{15}N HSQC experiments of a freshly prepared lipoyl domain sample in D_2O , (2) the hydrogen bond was present in the majority ($> 90\%$) of the calculated structures, and (3) the hydrogen bond was present in regular antiparallel β -sheet and between backbone atoms only. Two constraints were used

for each hydrogen-bond, 1.5 to 2.3 Å for the NH-O distance, and 2.5 to 3.3 Å for the N-O distance.

Structure calculations

Structures were calculated with the program X-PLOR version 3.1 (Brünger, 1992), using an *ab initio* simulated annealing protocol (Nilges *et al.*, 1991). The starting structure had randomised ϕ and ψ angles, and χ_1 angles set to 180°. The protocol started with 90 ps of restrained dynamics at 1000 K during which the weighting of the quartic van der Waals term was kept very low to allow atoms to pass through each other. The structures were then gradually cooled (50 K steps) to 100 K in 75 ps while increasing the weight of the repulsive van der Waals energy term to its final value (see Table 1). Finally, 1000 steps of restrained Powell minimisation were applied.

A mean structure (<SA>) was obtained by fitting all backbone atoms (residues 1 to 73) of the final set of selected structures to each other and averaging their atomic coordinates. The mathematical average structure was then minimised by 2000 steps of restrained Powell minimisation.

Structures were visualised using the InsightII program (Biosym Technologies Inc., San Diego), and analysed by tools provided with X-PLOR, PROCHECK (Laskowski *et al.*, 1993) and NAOMI (Brocklehurst & Perham, 1993).

The coordinates of both the energy-minimised average structure and the ensemble of 29 structures representing the solution structure of the lipoyl domain, together with the NMR constraints used for their determination, have been deposited with the Protein Data Bank, Brookhaven National Laboratory, Upton, NY, USA. The entry codes are I1YU and I1YV, respectively, for the energy-minimised average structure and the ensemble of 29 structures.

Acknowledgements

Several NMR experiments were performed at the SON National Hf-NMR Facility (Nijmegen, The Netherlands). This work was financially supported by the Netherlands Organisation for Scientific Research (NWO) under the auspices of the Netherlands Foundation for Chemical Research (SON).

REFERENCES

- Athappilly, F. K. & Hendrickson, W. A. (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**, 1407-1419.
- Berg, A., de Kok, A. & Vervoort, J. (1994) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **221**, 87-100.
- Berg, A., Smits, O., de Kok, A. & Vervoort, J. (1995) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex. *Eur. J. Biochem.* **234**, 148-159.
- Berg, A., Vervoort, J. & de Kok, A. (1996) Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. *J. Mol. Biol.* **261**, 432-442.
- Brocklehurst, S. M. & Perham, R. N. (1993) Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: a new automated method for the prediction of protein tertiary structure. *Protein Sci.* **2**, 626-639.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187-217.
- Brünger, A. T. (1992) *X-PLOR (Version 3.1): a system for x-ray crystallography and NMR*. Yale University Press, New Haven.
- Chothia, C. & Murzin, A. G. (1993) New folds for all- β proteins. *Structure* **1**, 217-222.
- Clore, G. M., Gronenborn, A. M., Nilges, M. & Ryan, C. A. (1987) Three-dimensional structure of potato carboxypeptidase inhibitor in solution. A study using nuclear magnetic resonance, distance geometry, and restrained molecular dynamics. *Biochemistry* **26**, 8012-8023.
- Dardel, F., Davis, A. L., Lauc, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.
- Graham, L. D. & Perham, R. N. (1990) Interactions of lipoyl domains with the E1p subunits of the pyruvate dehydrogenase multienzyme complex from *Escherichia coli*. *FEBS Lett.* **262**, 241-244.
- Green, J. D. F., Lauc, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995) Three-dimensional structure of a lipoyl domain from dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.
- Güntert, P., Braun, W. & Wüthrich, K. (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517-530.
- Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987) The domain structure of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **169**, 245-252.
- Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988) The dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Molecular cloning and sequence analysis. *Eur. J. Biochem.* **172**, 593-599.
- Hommel, U., Harvey, T. S., Driscoll, P. C. & Campbell, I. D. (1992) Human epidermal growth factor: high resolution solution structure and comparison with human transforming growth factor α . *J. Mol. Biol.* **227**, 271-282.
- Hyberts, S. G., Goldberg, M. S., Havel, T. F. & Wagner, G. (1992) The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci.* **1**, 736-751.
- Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K. & Perham, R. N. (1993) The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *J. Mol. Biol.* **230**, 323-341.

- Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
- Ludvigsen, S. & Poulsen, F. M. (1992) Positive ϕ -angles in proteins by nuclear magnetic resonance spectroscopy. *J. Biomol. NMR* **2**, 227-233.
- Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M. & Clore, G. M. (1989) Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1β . *Biochemistry* **28**, 6150-6156.
- Mattevi, A., Schierbeek, A. J. & Hol, W. G. J. (1991) Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 Å resolution. A comparison with the structure of glutathione reductase. *J. Mol. Biol.* **220**, 975-994.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992a) The pyruvate dehydrogenase multienzyme complex. *Curr. Opin. Struct. Biol.* **2**, 877-887.
- Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1992b) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
- Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1993) Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Å resolution. *J. Mol. Biol.* **230**, 1183-1199.
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins: Struct. Funct. Genet.* **12**, 345-364.
- Nilges, M., Kuszewski, J. & Brünger, A. T. (1991) Sampling properties of simulated annealing and distance geometry. In: *Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy* (Hoch, J. C., et al., eds.), pp. 451-455, Plenum Press, New York.
- Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. & Douce, R. (1994) X-ray structure determination at 2.6-Å resolution of a lipoate-containing protein: the H-protein of the glycine decarboxylase complex from pea leaves. *Proc. Natl. Acad. Sci. USA* **91**, 4850-4853.
- Ravindran, S., Radke, G. A., Guest, J. R. & Roche, T. E. (1996) Lipoyl domain-based mechanism for the integrated feedback control of the pyruvate dehydrogenase complex by enhancement of pyruvate dehydrogenase kinase activity. *J. Biol. Chem.* **271**, 653-662.
- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.
- Robien, M. A., Clore, G. M., Omichinski, J. G., Perham, R. N., Appella, E., Sakaguchi, K. & Gronenborn, A. M. (1992) Three-dimensional solution structure of the E3-binding domain of the dihydrolipoamide succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **31**, 3463-3471.
- Schierbeek, A. J., Swarte, M. B. A., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G. J., Drenth, J. & Betzel, C. (1989) X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *J. Mol. Biol.* **206**, 365-379.
- Wagner, G., Braun, W., Havel, T. H., Schaumann, T., Go, N. & Wüthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* **196**, 611-639.
- Wallis, N. G. & Perham, R. N. (1994) Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**, 209-216.
- Westphal, A. H. & de Kok, A. (1990) The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur. J. Biochem.* **187**, 235-239.
- Wilmot, C. M. & Thornton, J. M. (1990) β -Turns and their distortions: a proposed new nomenclature. *Protein Eng.* **3**, 479-493.
- Wüthrich, K., Billeter, M. & Braun, W. (1983) Pseudostructures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949-961.
- Zuiderweg, E. R. P., Boelens, R. & Kaptein, R. (1985) Stereospecific assignments of ^1H -NMR methyl lines and conformation of valyl residues in the lac repressor headpiece. *Biopolymers* **24**, 601-611.

CHAPTER 6

Reductive acylation of lipoyl domains of 2-oxo acid dehydrogenase complexes from *Azotobacter vinelandii*

ABSTRACT

The rate and specificity of reductive acylation of lipoyl domains derived from *Azotobacter vinelandii* 2-oxo acid dehydrogenase complexes, catalysed by *A. vinelandii* and *Escherichia coli* complexes, have been investigated. The turnover rate of reductive acetylation of complex-bound lipoyl domains by pyruvate dehydrogenase (E1p) is more than 50 times higher than of free lipoyl domains under comparable conditions. This gain in catalytic rate indicates a large limitation of diffusion of lipoyl domains when attached via the flexible linker segments to the complex, and illustrates the efficiency of substrate channeling in the multienzyme complex. The 2-oxo acid dehydrogenases exhibit specificity for lipoyl domains in the reductive acylation reaction. The *A. vinelandii* pyruvate dehydrogenase complex (PDHC) derived lipoyl domain is a good substrate for *A. vinelandii* pyruvate dehydrogenase, but not for *A. vinelandii* 2-oxoglutarate dehydrogenase (E1o), and *vice versa*. The *A. vinelandii* PDHC lipoyl domain is also, although at a lower rate, reductively acetylated by *E. coli* E1p and reductively succinylated by *E. coli* E1o. Likewise, the *A. vinelandii* 2-oxoglutarate dehydrogenase complex (OGDHC) derived lipoyl domain is recognised by *E. coli* E1o, but not by *E. coli* E1p. This suggests that common determinants of the lipoyl domains exist that are responsible for recognition by the E1 components. On the basis of the observed specificity and lipoyl domain sequences and structures, an exposed loop of the *A. vinelandii* OGDHC lipoyl domain was subjected to mutagenesis. Although the reductive acylation experiments of mutants of the lipoyl domain indicate the importance of this loop for recognition, it is probably not the single determinant for specificity.

INTRODUCTION

The 2-oxo acid dehydrogenase complexes catalyse the oxidative decarboxylation of 2-oxo acids to the corresponding acyl-CoA derivatives, accompanied with the reduction of NAD⁺ [for a recent review see Mattevi *et al.* (1992a)]. In the Gram-negative bacteria *Azotobacter vinelandii* and *Escherichia coli* two of these multienzyme complexes, which are very similar, are present. The pyruvate dehydrogenase complex (PDHC) produces acetyl-CoA from pyruvate, and the 2-oxoglutarate dehydrogenase complex (OGDHC) converts 2-oxoglutarate into succinyl-CoA. The complexes effectuate these reactions by combining the activities of three enzymes: a substrate-specific 2-oxo acid dehydrogenase [pyruvate dehydrogenase (E1p) or 2-oxoglutarate dehydrogenase (E1o)], an acyltransferase [acetyltransferase (E2p) or succinyltransferase (E2o)], and a common lipoamide dehydrogenase (E3). The oligomeric cubic core of the complex is provided by the E2 component (Mattevi *et al.*, 1992b), to which the peripheral components E1 and E3 are non-covalently bound as dimers. The E2 component consists of three types of domains (Stephens *et al.*, 1983; Packman *et al.*, 1984a; Hanemaaijer *et al.*, 1987, 1988; Westphal & de Kok, 1990): one (E2o) or three (E2p) homologous N-terminal lipoyl domains, a peripheral subunit-binding domain, and a C-terminal catalytic domain. The domains are separated by long (20 to 40 amino acids) flexible linker segments rich in alanine, proline and frequently charged residues.

The lipoyl domains play a vital role in coupling of the activities of the three enzyme components in the complex. They each contain the prosthetic group lipoic acid, which is covalently bound to a specific lysine residue forming the so-called lipoyl group (Reed, 1974). In the multienzyme reaction sequence, these lipoyl groups are subsequently reductively acylated by the E1 component, de-acylated by the E2 component, and re-oxidised by the E3 component. The flexibility of the linker segments has proven to be essential for this active-site coupling (Miles *et al.*, 1988; Radford *et al.*, 1989). The importance for lipoyl domains to being covalently linked to the complex for efficient multienzyme catalysis is also apparent (this chapter; Graham *et al.*, 1989). Furthermore, the three-dimensional structure of the lipoyl domain is required for promoting reductive acylation of its pendant lipoyl group by the E1 component (Graham *et al.*, 1989), although this does not seem to happen just simply by enhancement of the binding to E1 (Graham & Perham, 1990). Free lipoamide has shown to be a very poor substrate for E1.

An interesting observation by Graham *et al.* (1989) was that isolated lipoyl domains of the 2-oxo acid dehydrogenase complexes of *E. coli* were only efficiently reductively acylated by the E1 components of their parent complex. The specificity of E1

components for their lipoyl domains has been generalised in the following years, but has however never been proven to extend to complexes from other sources. We will show now that this specificity is also present in the complexes from *A. vinelandii*. The requirement for specificity towards lipoyl domains is not directly evident but, as discussed by Berg *et al.* (1996), the possible occurrence of hybrid complexes with respect to their E1 components (Steginsky *et al.*, 1985) could be a reasonable explanation for the observed specificity.

During the last years a number of different lipoyl domain structures of various complexes have become available (Dardel *et al.*, 1993; Green *et al.*, 1995; Berg *et al.*, 1996; chapter 5, this thesis). No three-dimensional structure of an E1 component is yet available. In the light of the specific recognition of lipoyl domains by E1, lipoyl domain structures and sequences were compared in search for part(s) of the lipoyl domain that could be important for molecular recognition between lipoyl domains and E1 components (Berg *et al.*, 1996; chapter 5, this thesis). Potential residues of the lipoyl domain that could be involved in recognition include residues of a solvent-exposed loop close in space to the lipoylation site. Here we describe cross-acylation experiments of *A. vinelandii* PDHC and OGDHC lipoyl domains catalysed by *E. coli* complexes, and site-directed mutagenesis experiments of the exposed loop of the OGDHC lipoyl domain, to investigate the role of this loop in molecular recognition.

RESULTS AND DISCUSSION

Kinetics of reductive acylation

The kinetics of reductive acetylation of the free PDHC-derived N-terminal lipoyl domain, catalysed by *A. vinelandii* PDHC as the source of E1p, were studied. The rate of radioactivity incorporation from [2-¹⁴C]pyruvate was measured as a function of the lipoyl domain concentration (Figure 1). The apparent kinetic constants from the double reciprocal plot are a K_m value of $\sim 43 \mu\text{M}$ and a k_{cat} of $\sim 0.8 \text{ s}^{-1}$, with a k_{cat}/K_m of $\sim 1.9 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$. The kinetic constants for the reductive succinylation of the free OGDHC-derived lipoyl domain catalysed by *A. vinelandii* OGDHC were of the same order of magnitude (data not shown). These data are highly comparable with kinetic data obtained for *E. coli* PDHC ($K_m = 26 \mu\text{M}$, $k_{cat} = 0.8 \text{ s}^{-1}$) (Graham *et al.*, 1989). A comparison of the k_{cat}/K_m values for lipoyl domains with the k_{cat}/K_m value for free lipoamide ($K_m > 4 \text{ mM}$, $k_{cat}/K_m \sim 1.5 \text{ M}^{-1}\text{s}^{-1}$) as a substrate for *E. coli* E1p shows that the lipoyl domain greatly promotes the reductive acetylation of its lipoyl group (Perham, 1991).

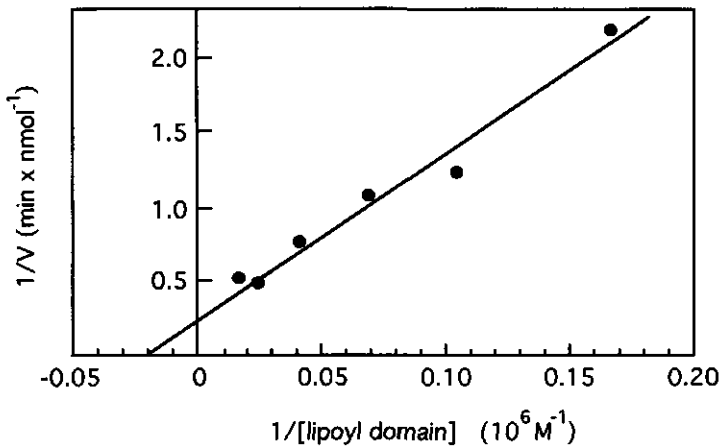


Figure 1. Kinetics of reductive acetylation of the N-terminal lipoyl domain of *A. vinelandii* PDHC, catalysed by PDHC as source of E1p. All reactions were performed at 25 °C. Reaction mixtures contained 0.1 U PDHC, 0.4 mM thiamin diphosphate, 2 mM MgCl_2 , 3 mM NAD^+ , 0.25 mM $[2\text{-}^{14}\text{C}]$ pyruvate and lipoyl domain (6–60 μM) in 50 μl of 50 mM potassium phosphate pH 7.0. A double reciprocal plot is shown relating the amount of radioactivity incorporation with the lipoyl domain concentration.

It is also interesting to compare the k_{cat} value for the reductive acetylation of free lipoyl domains ($\sim 0.8 \text{ s}^{-1}$) with that of complex-bound lipoyl domains. Since the reductive acetylation of lipoyl domains by E1p is known to be the rate limiting step in the overall complex reaction (Danson *et al.*, 1978; Bosma, 1984), a k_{cat} of 46 s^{-1} for E1p can be calculated from the specific activity of 16 $\mu\text{mol NAD}^+ \text{ min}^{-1} \text{ mg}^{-1}$ of the *A. vinelandii* PDHC. It appears that, under comparable conditions, acetylation of complex-bound lipoyl domains is nearly 60 times faster than acetylation of free lipoyl domains. This can be ascribed to diffusion limitation, and very nicely illustrates the efficiency of substrate channeling inside the multienzyme complex. It is not inconceivable that this rate enhancement occurs in each of the steps of the complex reaction cycle involving the lipoyl domain. This could explain why no PDHC activity was detected in cell-free extracts of a strain of *E. coli* deleted for the PDHC genes, in which two independent plasmids encoding, respectively, a PDHC lacking lipoyl domains and a discrete lipoyl domain, were expressed (Russell *et al.*, 1989).

Specificity of reductive acylation

To determine the specificity of the reductive acylation reaction, the incorporation of acyl groups in the *A. vinelandii* PDHC and OGDHC lipoyl domains, catalysed by various

complexes as sources of the E1 component was assayed. The results of these cross-acylation experiments are shown in Table 1. If we first consider the incubations of lipoyl domains with complexes from *A. vinelandii*, it is apparent that the *A. vinelandii* lipoyl domains are only efficiently reductively acylated by the E1 components of their parent complexes. Reductive acetylation of the OGDHC lipoyl domain by *A. vinelandii* E1p could hardly be detected, and the PDHC lipoyl domain is only a poor substrate for *A. vinelandii* E1o. These results are in agreement with the results obtained for the *E. coli* complexes (Graham *et al.*, 1989), including the observation that the PDHC lipoyl domain is a somewhat better substrate for E1o than is the OGDHC lipoyl domain for E1p.

Table 1. Reductive acylation of lipoyl domains of *A. vinelandii* PDHC and OGDHC catalysed by PDHC and OGDHC from *A. vinelandii* and *E. coli* as the sources of the respective E1 components. The lipoyl domains are designated as follows: Av PDHC is the N-terminal lipoyl domain of the acetyltransferase component of *A. vinelandii* PDHC, Av OGDHC is the lipoyl domain of the succinyltransferase component of *A. vinelandii* OGDHC, and Av OGDHC_m2 and Av OGDHC_m3 are Av OGDHC lipoyl domain loop mutants as described in Figure 3.

Lipoyl domain	Complex	% acylation (1 min.)	% acylation (30 min.)
Av PDHC	Av PDHC	100	112
	Av OGDHC	0	22
	Ec PDHC	13	106
	Ec OGDHC	28	82
Av OGDHC	Av PDHC	2	2
	Av OGDHC	100	51
	Ec PDHC	0	0
	Ec OGDHC	93	17
Av OGDHC_m2	Av PDHC	1	8
	Av OGDHC	13	48
Av OGDHC_m3	Av PDHC	3	18
	Av OGDHC	27	106

To obtain more information about the specificity of the reductive acylation reaction, we extended the experiments with incubations of the *A. vinelandii* lipoyl domains with PDHC and OGDHC from *E. coli*, which are very homologous to the *A. vinelandii*

complexes. The experiments show that the *A. vinelandii* PDHC lipoyl domain is recognised by both E1p and E1o from *E. coli*, but that the rate of acyl incorporation is much slower than with *A. vinelandii* E1p (Table 1). The *A. vinelandii* OGDHC lipoyl domain is acylated by *E. coli* OGDHC to an amount comparable with the incubation with *A. vinelandii* OGDHC for 1 min. Finally, the OGDHC lipoyl domain is not reductively acetylated by *E. coli* PDHC.

The observation that the amount of incorporated radioactivity into the OGDHC lipoyl domain is decreased after prolonged incubation (30 min.) with both the *A. vinelandii* and the *E. coli* OGDHC could be due, at least in part, to the instability of the protein-bound succinyl groups under the applied conditions (Collins & Reed, 1977). This situation does, however, not seem to apply to the incubations of the PDHC lipoyl domain and the OGDHC lipoyl domain mutants (discussed below) with OGDHCs. Although we have at the moment no good explanation for the apparent greater stability of the protein-bound succinyl groups in those cases, we could speculate that here always a PDHC lipoyl domain exposed loop is present, that could play a role in stabilising the succinyl groups.

Recognition of *A. vinelandii* PDHC lipoyl domains by *E. coli* E1p has been observed earlier in an experiment with a reconstituted *E. coli* PDHC based on a chimeric E2p core, in which the lipoyl domains were replaced by those of *A. vinelandii* E2p (Schulze *et al.*, 1992). This complex showed only 44% overall complex activity as compared with the same complex reconstituted with E1p from *A. vinelandii* instead of E1p from *E. coli*. The reaction rate in the complex can however not be compared directly with the rate of reductive acetylation of free lipoyl domains, since the local lipoyl domain concentration in the complex is much higher (estimated in the mM range) than in the reductive acetylation assay (60 μ M). An example of recognition of *E. coli* PDHC lipoyl domains by *A. vinelandii* E1p is given by a hybrid *E. coli* PDHC reconstituted with *A. vinelandii* E1p, which showed around 50% of the wild-type complex activity (De Kok & Westphal, 1985; Schulze *et al.*, 1992).

OGDHC lipoyl domain loop mutants

The results obtained from the cross-acylation reactions suggest that there are common determinants of the lipoyl domains which are responsible for the specific interactions of lipoyl domains with the E1 components. A comparison of lipoyl domain structures and many lipoyl domain sequences revealed that an exposed loop of the lipoyl domain, which is close in space to the lipoylation site, could be involved in the process of molecular recognition (Berg *et al.*, 1996). In order to test if this solvent-exposed loop is a determinant for recognition, we performed loop-directed mutagenesis experiments on the

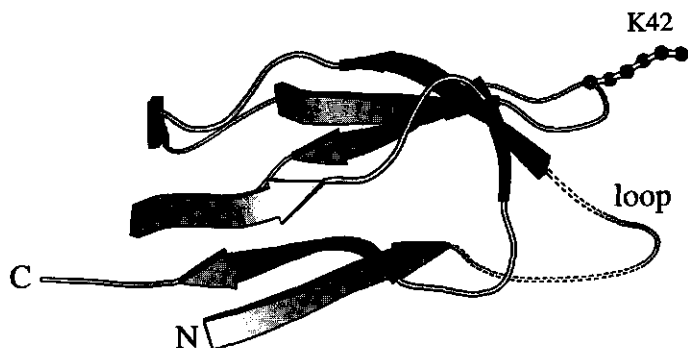


Figure 2. Schematic representation (Kraulis, 1991) of the three-dimensional solution structure of the *A. vinelandii* OGDHC lipoyl domain (Berg *et al.*, 1996), showing the lipoylation site (K42) and the exposed loop which has been mutated.

	1	5	10	15
Av PDHC_1	S	E I I R V P D I G	---	G D G E V
Av PDHC_2	S	Q E V R V P D I G	-	S A G K A R V
Av PDHC_3	P	Q E V K V P D I G	-	S A G K A R V
Av OGDHC	A	I D I K A P T F P E S I A D G T V		
Ec PDHC_1	A	I E I K V P D I G	---	A D E V E I
Ec PDHC_2	A	K D V N V P D I G	---	S D E V E V
Ec PDHC_3	V	K E V N V P D I G	---	G D E V E V
Ec OGDHC	S	V D I L V P D L P E S V A D A T V		
Av OGDHC_m1	A	I D I K A P D I G	---	A D G T V
Av OGDHC_m2	A	I D I K A P D I G	-	S I A D G T V
Av OGDHC_m3	A	I D I K A P D I G	-	S S A D G T V
sec. struct.	β	β β β β β	<u>l l l l l l l l l l</u>	β β β

Figure 3. Sequence alignment of the N-terminal residues of the lipoyl domains of *A. vinelandii* PDHC (Av PDHC_1-3) (Hanemaaijer *et al.*, 1988) and OGDHC (Av OGDHC) (Westphal & de Kok, 1990), *E. coli* PDHC (Ec PDHC_1-3) (Stephens *et al.*, 1983) and OGDHC (Ec OGDHC) (Spencer *et al.*, 1984), and three *A. vinelandii* OGDHC lipoyl domain mutants (Av OGDHC_m1-3). The locations of the exposed loop (l) and β-strands (β) are indicated in the bottom line.

OGDHC lipoyl domain (Figure 2), changing the sequence of the loop into different sequence variants of the loop of the lipoyl domain of PDHC. Three different OGDHC lipoyl domain loop mutants were created (Figure 3). One of the mutants (Av OGDHC_m1), in which the OGDHC lipoyl domain sequence TFPESI was replaced by the PDHC sequence DIG, did not express well in *E. coli*, showed a poor solubility and could not be purified and tested. The two other mutants designated Av OGDHC_m2 and Av OGDHC_m3, in which the sequence TFPESI has been replaced by DIGSI and DIGSS respectively, expressed well and could be purified. Both mutants were lipoylated in *E. coli* when exogenous lipoic acid was added to the growth medium (Berg *et al.*, 1994), as judged by comparison with their unlipoylated form after electrophoresis on a native polyacrylamide gel.

The purified lipoyl domain mutants were both tested in the reductive acylation assays with the two *A. vinelandii* complexes as the source of the E1 component (Table 1). These experiments show that both mutants are still being reductively succinylated by OGDHC, but at a lower rate as compared with the native OGDHC lipoyl domain. However, in particular the Av OGDHC_m3 mutant is now reductively acetylated by PDHC to a higher level than the native OGDHC lipoyl domain after 30 min. incubation. It is clear from these experiments that substitution of the sequence of the exposed loop of the OGDHC lipoyl domain by a sequence of the PDHC lipoyl domain does not alter the specificity of reductive acylation considerably. Only a small change in specificity towards PDHC is observed. It is therefore concluded that residues in this loop, and possibly the conformation of this loop, may very well be involved in the specific interaction between lipoyl domains and 2-oxo acid dehydrogenases, but that it is certainly not the only determinant for specificity. Other residues of the lipoyl domain, some of which have been suggested recently (Berg *et al.*, 1996), must also play a role in this process of molecular recognition. This will be the subject of further research.

MATERIALS AND METHODS

Materials

Sodium [2-¹⁴C]pyruvate (23.0 Ci/mol) and sodium [U-¹⁴C]2-oxoglutarate (293 Ci/mol) were obtained from NEN Research Products. All other chemicals used were of analytical grade.

Construction of lipoyl domain mutants

Standard methods were used for DNA manipulations (Ausubel *et al.*, 1987). Site-directed mutagenesis was performed by the mega-primer method (Landt *et al.*, 1990) using as a template the plasmid pAB1 (Berg *et al.*, 1995) encoding the first 79 amino acids of the succinyltransferase component of *A. vinelandii* OGDHC. The complete sub-gene encoding the lipoyl domain was sequenced after the mutagenesis reactions to ensure that besides the desired mutations no other mutations were present.

Protein isolation

The lipoylated lipoyl domain (residues 1-79) of OGDHC from *A. vinelandii* was isolated from a sub-gene overexpressed in *E. coli*, as described by Berg *et al.* (1995). The same purification procedure was used for the purification of the lipoylated N-terminal lipoyl domain (residues 1-79) of PDHC from *A. vinelandii* (Berg *et al.*, 1994) and the mutant OGDHC lipoyl domains. The lipoyl domains were dialysed against 50 mM potassium phosphate pH 7.0. The purity and lipoylation (see also Berg *et al.*, 1994) of the lipoyl domains was analysed by denaturing and non-denaturing polyacrylamide gel electrophoresis (Schägger & von Jagow, 1987), with resolving gel 16.5% T, 3% C and stacking gel 4% T, 3% C. Lipoyl domain concentrations were estimated using the microbiuret method (Goa, 1953).

The pyruvate dehydrogenase complex and the 2-oxoglutarate dehydrogenase complex from *A. vinelandii* were isolated as described by Bosma *et al.* (1984) and by Bosma (1984), respectively. The pyruvate dehydrogenase complex from *E. coli* was purified according to De Kok & Westphal (1985). The 2-oxoglutarate dehydrogenase complex from *E. coli* was isolated in a similar manner as described for the complex from *A. vinelandii*. PDHC overall activity was assayed as described by Schwartz & Reed (1970). In the assay for OGDHC the pyruvate was replaced by 2-oxoglutarate.

Reductive acylation assay

Reductive acylation of lipoyl domains by 2-oxo acid dehydrogenase in the presence of the corresponding ¹⁴C-labelled 2-oxo acid was assayed in a similar manner as described by Packman *et al.* (1984b). The general assay conditions were the following: lipoyl domain was incubated at 25°C for 5 min. in the presence of a catalytic amount (1 U/ml, 1 U = 1 μmol NADH min.⁻¹) of 2-oxo acid dehydrogenase complex (PDHC or OGDHC) in 50-300 μl 50 mM potassium phosphate pH 7.0, containing 0.4 mM thiamin diphosphate, 2 mM MgCl₂, and 3 mM NAD⁺. The reaction was started with the addition of sodium [¹⁴C]pyruvate or sodium [U-¹⁴C]2-oxoglutarate to a final concentration of 0.25 mM, and

was left at 25 °C during the reaction. After the addition of bovine serum albumine as a carrier protein (final concentration 1 mg/ml) the reaction was stopped by addition and immediate mixing of 1-2 ml of ice-cold 10% (mass/vol.) trichloroacetic acid. This suspension was kept on ice for 5 min. and then the protein precipitate was collected on two stacked Whatman GC/F filters. The filters were washed with 10 ml ice-cold 10% (mass/vol.) trichloroacetic acid, followed by 3 ml of ice-cold acetone. The dried filters were counted for radioactivity. In the control experiments the lipoyl domain was omitted from the incubation mixture.

For the determination of the specificity of reductive acylation, typically 60 μM of isolated *A. vinelandii* N-terminal PDHC or OGDHC lipoyl domain was incubated in 100 μl reaction mixture for 1 min. and 30 min. with PDHC or OGDHC from *A. vinelandii* or *E. coli* and the corresponding ^{14}C -labelled 2-oxo acid. For the determination of the kinetic constants of the *A. vinelandii* 2-oxo acid dehydrogenases (E1p and E1o) for their respective lipoyl domains as substrates, various concentrations (6-60 μM) of lipoyl domain were incubated for 1 min. at 25 °C with the corresponding complex in a reaction volume of 50 μl . The catalytic amount of complex used for these reactions was chosen such that the reductive acylation reactions were linear up to at least 1.5 min. for the highest concentration of lipoyl domain that was used.

Acknowledgements

We wish to thank Anton Veenstra for help with the preparation of one of the lipoyl domain mutants. We thank Prof. C. Veeger for critically reading the manuscript. This work was financially supported by the Netherlands Organisation for Scientific Research (NWO) under the auspices of the Netherlands Foundation for Chemical Research (SON).

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) *Current protocols in molecular biology*. John Wiley & Sons, New York.
- Berg, A., de Kok, A. & Vervoort, J. (1994) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **221**, 87-100.
- Berg, A., Smits, O., de Kok, A. & Vervoort, J. (1995) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex. *Eur. J. Biochem.* **234**, 148-159.

- Berg, A., Vervoort, J. & de Kok, A. (1996) Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. *J. Mol. Biol.* **261**, 432-442.
- Bosma, H. J. (1984) *Studies on 2-oxoacid dehydrogenase multienzyme complexes of Azotobacter vinelandii*. Ph.D. Thesis, Agricultural University Wageningen, The Netherlands.
- Bosma, H. J., de Kok, A., Westphal, A. H. & Veeger, C. (1984) The composition of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Does a unifying model exist for the complexes from gram-negative bacteria? *Eur. J. Biochem.* **142**, 541-549.
- Collins, J. H. & Reed, L. J. (1977) Acyl group and electron pair relay system: a network of interacting lipoyl moieties in the pyruvate and α -ketoglutarate dehydrogenase complexes from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **74**, 4223-4227.
- Danson, M. J., Fersht, A. R. & Perham, R. N. (1978) Rapid intramolecular coupling of active sites in the pyruvate dehydrogenase complex of *Escherichia coli*: mechanism for rate enhancement in a multimeric structure. *Proc. Natl. Acad. Sci. USA* **75**, 5386-5390.
- Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- De Kok, A. & Westphal, A. H. (1985) Hybrid pyruvate dehydrogenase complexes reconstituted from components of the complexes from *Escherichia coli* and *Azotobacter vinelandii*. *Eur. J. Biochem.* **152**, 35-41.
- Goa, J. (1953) A micro biuret method for protein determination: determination of total protein in cerebrospinal fluid. *Scand. J. Clin. Lab. Invest.* **5**, 218-222.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.
- Graham, L. D. & Perham, R. N. (1990) Interactions of lipoyl domains with the E1p subunits of the pyruvate dehydrogenase multienzyme complex from *Escherichia coli*. *FEBS Lett.* **262**, 241-244.
- Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995) Three-dimensional structure of a lipoyl domain from dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.
- Hanemaaijer, R., de Kok, A., Jolles, J. & Veeger, C. (1987) The domain structure of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. *Eur. J. Biochem.* **169**, 245-252.
- Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988) The dihydrolipoyltransacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Molecular cloning and sequence analysis. *Eur. J. Biochem.* **172**, 593-599.
- Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
- Landt, O., Grunert, H. P. & Hahn, U. (1990) A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**, 125-128.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992a) The pyruvate dehydrogenase multienzyme complex. *Curr. Opin. Struct. Biol.* **2**, 877-887.
- Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A. & Hol, W. G. J. (1992b) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
- Miles, J. S., Guest, J. R., Radford, S. E. & Perham, R. N. (1988) Investigation of the mechanism of active site coupling in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli* by protein engineering. *J. Mol. Biol.* **202**, 97-106.
- Packman, L. C., Hale, G. & Perham, R. N. (1984a) Repeating functional domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *EMBO J.* **3**, 1315-1319.
- Packman, L. C., Perham, R. N. & Roberts, G. C. K. (1984b) Domain structure and ^1H -n.m.r. spectroscopy of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *Biochem. J.* **217**, 219-227.
- Perham, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.
- Radford, S. E., Perham, R. N., Ullrich, S. J. & Appella, E. (1989) Antibodies against an inter-domain segment of polypeptide chain inhibit active-site coupling in the pyruvate dehydrogenase multienzyme complex. *FEBS Lett.* **250**, 336-340.
- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.

- Russell, G. C., Williamson, R. A. & Guest, J. R. (1989) Partial complementation of pyruvate dehydrogenase deficiency by independently expressed lipoyl and catalytic domains of the dihydrolipoamide acetyltransferase component. *FEMS Microbiol. Lett.* **60**, 267-272.
- Schägger, H. & von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379.
- Schulze, E., Westphal, A. H., Veeger, C. & de Kok, A. (1992) Reconstitution of pyruvate dehydrogenase multienzyme complexes based on chimeric core structures from *Azotobacter vinelandii* and *Escherichia coli*. *Eur. J. Biochem.* **206**, 427-435.
- Schwartz, E. R. & Reed, L. J. (1970) Regulation of the activity of the pyruvate dehydrogenase complex of *Escherichia coli*. *Biochemistry* **9**, 1434-1439.
- Spencer, M. E., Darlison, M. G., Stephens, P. E., Duckenfield, I. K. & Guest, J. R. (1984) Nucleotide sequence of the *sucB* gene encoding the dihydrolipoamide succinyltransferase of *Escherichia coli* K12 and homology with the corresponding acetyltransferase. *Eur. J. Biochem.* **141**, 361-374.
- Steginsky, C. A., Gruys, K. J. & Frey, P. A. (1985) α -Ketoglutarate dehydrogenase complex of *Escherichia coli*. A hybrid complex containing pyruvate dehydrogenase subunits from pyruvate dehydrogenase complex. *J. Biol. Chem.* **260**, 13690-13693.
- Stephens, P. E., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983) The pyruvate dehydrogenase complex of *Escherichia coli* K12. Nucleotide sequence encoding the dihydrolipoamide acetyltransferase component. *Eur. J. Biochem.* **133**, 481-489.
- Westphal, A. H. & de Kok, A. (1990) The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. 2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component. *Eur. J. Biochem.* **187**, 235-239.

CHAPTER 7

Summary and concluding remarks

The 2-oxo acid dehydrogenase complexes are large multienzyme complexes that catalyse the irreversible oxidative decarboxylation of a specific 2-oxo acid to the corresponding acyl-CoA derivative. The pyruvate dehydrogenase complex (PDHC) converts the product of the glycolysis, pyruvate, to acetyl-CoA, which enters the tricarboxylic acid cycle. The 2-oxoglutarate dehydrogenase complex (OGDHC) functions in the tricarboxylic acid cycle itself by converting 2-oxoglutarate to succinyl-CoA. The branched-chain 2-oxo acid dehydrogenase complex (BCDHC) is involved in the catabolism of branched-chain amino acids. Since these complexes play vital roles in metabolism, impairment in their functioning by genetic defects naturally causes severe diseases in humans, e.g. lactic acidosis (PDHC deficiency) and maple syrup urine disease (BCDHC deficiency) (Patel & Harris, 1995).

The 2-oxo acid dehydrogenase complexes have a very similar design and share many structural and catalytic properties. They convert their substrate by the combined activity of multiple copies of three enzymes: a substrate-specific 2-oxo acid dehydrogenase (E1), an acyltransferase (E2), and a lipoamide dehydrogenase (E3). The E2 component forms the central oligomeric core of the complex to which the peripheral subunits E1 and E3 are noncovalently bound. The acyltransferase (E2) component is a highly segmented and multifunctional protein in which three different independently folded domains can be recognised, connected by mobile linker sequences. The N-terminal part consists of one to three lipoyl domains (~ 80 amino acid residues each) containing the covalently bound prosthetic group lipoic acid. Between the lipoyl domain(s) and the C-terminal catalytic domain (~ 29 kDa), which bears the acyltransferase active site and which aggregates to form the oligomeric core of the complex, the peripheral subunit-binding domain (~ 35 amino acid residues) is found.

The lipoyl domains play a crucial role in coupling the activities of the three multienzyme components by providing swinging arms that are mobile and responsible for substrate channelling among the three successive active sites. A specific lysine side chain

of each lipoyl domain is modified with lipoic acid to form a lipoyl group, which transports acyl groups from E1 to E2, and reduction equivalents from E2 to E3. The structure of the lipoyl domain is required for the efficient reaction of its lipoyl group with the E1 component. Furthermore, for *E. coli* PDHC and OGDHC it was shown that lipoyl domains can only be reductively acylated by the E1 enzyme of their parent complex, which indicates that molecular recognition occurs between E1 components and lipoyl domains (Graham *et al.*, 1989).

The research presented in this thesis aimed at gaining insight in the interaction between lipoyl domains and E1 components, and in particular what part of the lipoyl domain determines the specificity of the reductive acylation reaction. By the determination and comparison of the three-dimensional structures of the lipoyl domains of PDHC and OGDHC from *A. vinelandii*, for which their specificities in the reductive acylation reaction were determined, we expected to shed already some light on this process of molecular recognition. It also provides the structural basis for further studies on the specific interaction between E1 and the lipoyl domain, some of which are described in this thesis. The results of this work are summarised hereafter.

The N-terminal lipoyl domain (residues 1-79) of the E2p component of PDHC and the single lipoyl domain of the E2o component of OGDHC have both been sub-cloned and expressed in *E. coli* (chapters 2 and 3). The expression exceeded the capacity of the *E. coli* cells to lipoylate all the produced lipoyl domain, and only 5-10% of the lipoyl domain was found to be modified with lipoic acid. The unlipoylated and lipoylated forms of the lipoyl domain could be separated by anion-exchange chromatography. Addition of a supplementary amount of lipoic acid to the growth medium resulted in full lipoylation of the expressed lipoyl domain. The ability of the purified lipoylated lipoyl domains to become reductively acylated by the E1 components of their parent complex proved that their folding and modification were correct. The correct modification of the E2p lipoyl domain has been confirmed by electrospray mass spectrometry.

Two-dimensional homo- and heteronuclear NMR studies of the *A. vinelandii* lipoyl domains have resulted in sequential ^1H and ^{15}N resonance assignments and the secondary structure of both domains (chapters 2 and 3). The 2D ^1H -NOESY spectra of the unlipoylated and lipoylated forms of the E2p lipoyl domain are almost superimposable, except for several additional resonances that could be assigned to the lipoic acid moiety, and small differences in chemical shift of protons of residues in the direct vicinity of the lipoyl-lysine residue (chapter 2). No changes in NOE intensities connecting these residues nor addition or loss of NOEs could be observed however, suggesting that the structure of the lipoyl domain is not altered much if any upon lipoylation. A detailed comparison of the

NMR-derived parameters of both lipoyl domains, i.e. chemical shifts, NH-exchange rates, NOEs, and $^3J_{\text{HN}\alpha}$ coupling constants suggests a high structural similarity in solution between the two lipoyl domains, despite their amino acid sequence identity of only 25% (chapter 3).

The three-dimensional solution structures of the two lipoyl domains have been determined using distance geometry and/or dynamical simulated annealing calculations (chapters 4 and 5). The overall fold of both lipoyl domains is very similar and can be described as a β -barrel-sandwich hybrid, which is now known to be typical for lipoyl domains. The domain is formed by two very similar and almost parallel four-stranded antiparallel β -sheets connected by loops and turns. The β -sheets each consist of three major and one minor strand, and are formed around a well-defined core of hydrophobic residues. At the far end of one of the sheets the lipoyl-lysine residue is presented to the solvent in a β -turn connecting two successive strands. The N-terminal and C-terminal ends of the folded domain meet at the exact opposite of the domain in two adjacent β -strands of the other sheet. The lipoyl domains display a remarkable internal symmetry that projects one β -sheet onto the other β -sheet after rotation of approximately 180° about a 2-fold rotational symmetry axis. The last six and two C-terminal residues of the cloned fragments of the PDHC and OGDHC lipoyl domains, respectively, are poorly defined and belong to the flexible linker sequences connecting the lipoyl domains to the remainder of the acyltransferase chain. These residues also show a significant narrower linewidth of their amide protons in the NMR spectra, which is an indication of increased mobility.

The number of long-range NOE-distance constraints that have been obtained for the two lipoyl domains is not very large, but is comparable to the number obtained for other lipoyl domains, i.e. from *B. stearothermophilus* PDHC and *E. coli* PDHC (Dardel *et al.*, 1993; Green *et al.*, 1995). This seems inherited with the particular type of protein structure, in combination with the absence (*A. vinelandii* PDHC lipoyl domain) or low amount (only one Trp residue in the core of the *A. vinelandii* OGDHC lipoyl domain) of aromatic residues. Many long-range contacts between hydrophobic residues in the core of the domains are side-chain side-chain contacts. The unambiguous assignment of many of these contacts is impaired by overlap, and these contacts will only be accessible after ^{13}C -labelling of the lipoyl domains. Although the three-dimensional structures of the lipoyl domains that have been determined are not of high resolution, they provide good and suitable structural models for comparisons and the design of significant mutants to investigate the specific interactions between lipoyl domains and other complex components.

A comparison of the structures of the *A. vinelandii* PDHC and OGDHC lipoyl domains with those of the PDHC lipoyl domains of *B. stearrowthermophilus* and *E. coli* shows that their overall fold is strikingly similar. In particular, that fact that the two *A. vinelandii* lipoyl domains, for which their specificity in the reductive acylation reactions has been demonstrated now (chapter 6), have similar structures, indicates that molecular recognition of lipoyl domains by E1 is a result of only delicate differences among lipoyl domains. On the basis of a careful comparison of lipoyl domain structures and sequences, potential residues of the lipoyl domain that could be important for molecular recognition are proposed. These include residues of an exposed loop connecting the first two β -strands in the sequence, and which lies close in space to the lipoylation site. In this loop the largest structural differences among lipoyl domains are found. Other potential candidates are the two amino acid residues immediately succeeding the lipoyl-lysine residue.

Site-directed mutagenesis experiments of the exposed loop of the *A. vinelandii* OGDHC lipoyl domain, and cross-acylation experiments of *A. vinelandii* PDHC and OGDHC lipoyl domains catalysed by *E. coli* complexes, were performed to investigate the role of this loop in molecular recognition (chapter 6). These experiments indicate that this loop is very likely involved in the interaction with the E1 component, but that it is probably not the single determinant conferring specificity to the reductive acylation reaction. Additional site-directed mutagenesis experiments on this loop and the residues following the lipoyl-lysine residue are required to further investigate their role in molecular recognition.

All studies on the interaction between lipoyl domains and E1 components are impaired by the lack of a three-dimensional structure of any E1 component. Furthermore, although the interaction between the lipoyl domain and E1 is specific, it is supposedly weak (Graham & Perham, 1990). This has been substantiated by initial NMR experiments in which, upon addition of *A. vinelandii* E1p to the ^{15}N -labelled lipoylated E2p lipoyl domain (1:2 ratio), no broadening or shift of the lipoyl domain resonances could be observed (Berg *et al.*, 1996) (not described in this thesis). From addition of pyruvate to this NMR sample it was suggested that also the acetylated form of the lipoyl domain does not bind to E1p, at least not under the applied conditions. Together this implies that the lipoyl domain does not interact (strongly) with the E1 component, except when the hydroxyethyl-ThDP moiety is present in E1. Such a ping-pong type of mechanism may be advantageous for the lipoyl domain that needs to interdigitate rapidly among the different active sites in the complex, but is unfortunately less advantageous for the observer.

A final remark is dedicated to the rotational flexibility of the lipoyl group of the lipoyl domain, that is thought to be required to act as a swinging arm in the complex

(Reed, 1974). Recently, the X-ray crystal structures of the H-protein of the glycine decarboxylase system with its lipoyl group loaded with methylamine (Cohen-Addad *et al.*, 1995), and of the biotinyl domain of acetyl-CoA carboxylase (Athappilly & Hendrickson, 1995), showed that their lipoyl/biotinyl group binds back to the protein surface. This may even not be very surprising considering the hydrophobic nature of the lipoyl group, and the question if this also could happen with lipoyl groups of 2-oxo acid dehydrogenase complexes is therefore relevant. In the case of the H-protein, only the methylamine-loaded form of the lipoyl group binds to the protein, while the oxidised lipoyl group does not. Furthermore, the methylamine-loaded lipoyl group binds to residues that are conserved in H-proteins, several of which are located in a N-terminal helix that is absent in lipoyl domains. In the case of the biotinyl domain, the biotinyl group is partly buried in a thumb-like protruding loop that is not found in lipoyl domains. For the lipoyl domains of 2-oxo acid dehydrogenase complexes no X-ray crystal structures, or NMR structures of their lipoylated form, are available. As mentioned earlier, only small differences in chemical shifts of residues close in space to the lipoylation site (average difference ~ 0.15 ppm), including residues of the exposed loop (average difference ~ 0.07 ppm), are observed between the lipoylated and unlipoylated forms of the lipoyl domain (chapter 2). These differences are considered too small to suggest binding of the lipoyl group to e.g. the exposed loop. Together with the fact that in the H-protein and the biotinyl domain the lipoyl/biotinyl group binds to parts of the protein that are absent in lipoyl domains, this indicates that the lipoyl groups of 2-oxo acid dehydrogenase complexes are likely to swing freely.

References

- Athappilly, F. K. & Hendrickson, W. A. (1995) Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**, 1407-1419.
- Berg, A., Bosma, H., Westphal, A. H., Hengeveld, J., Vervoort, J. & de Kok, A. (1996) The pyruvate dehydrogenase complex from *Azotobacter vinelandii*. Structure and function of the lipoyl domain. In: *Biochemistry and physiology of thiamin diphosphate enzymes* (Bisswanger, H. & Schellenberger, A., eds.), pp. 278-291, A.u. C. Intemann, Wissenschaftlicher verlag, Prien, Germany.
- Cohen-Addad, C., Pares, S., Sieker, L., Neuburger, M. & Douce, R. (1995) The lipoamide arm in the glycine decarboxylase complex is not freely swinging. *Nat. Struct. Biol.* **2**, 63-68.
- Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. (1993) Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048.
- Graham, L. D., Packman, L. C. & Perham, R. N. (1989) Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry* **28**, 1574-1581.
- Graham, L. D. & Perham, R. N. (1990) Interactions of lipoyl domains with the E1p subunits of the pyruvate dehydrogenase multienzyme complex from *Escherichia coli*. *FEBS Lett.* **262**, 241-244.

Chapter 7

- Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T. & Guest, J. R. (1995) Three-dimensional structure of a lipoyl domain from dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Mol. Biol.* **248**, 328-343.
- Patel, M. S. & Harris, R. A. (1995) Mammalian α -keto acid dehydrogenase complexes: gene regulation and genetic defects. *FASEB J.* **9**, 1164-1172.
- Reed, L. J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46.

NEDERLANDSE SAMENVATTING

De 2-oxozuurdehydrogenase complexen zijn zeer grote multi-enzymcomplexen die de oxidatieve decarboxylering van een specifiek 2-oxozuur tot het corresponderende acyl-CoA derivaat katalyseren. Het pyruvaatdehydrogenase complex (PDHC) zet het eindproduct van de glycolyse, pyruvaat, om in acetyl-CoA. Acetyl-CoA wordt vervolgens in de citroenzuurcyclus verder verbrand. Het α -ketoglutaraatdehydrogenase complex (OGDHC) maakt deel uit van de citroenzuurcyclus en zet α -ketoglutaraat om in succinyl-CoA. Verder is er nog een derde oxozuurdehydrogenase complex (BCDHC) dat is betrokken bij de omzetting van oxozuren afgeleid van aminozuren met vertakte ketens (valine, leucine, isoleucine). Al deze multi-enzymcomplexen spelen een essentiële rol in het energiemetabolisme. Het is dan ook niet verwonderlijk dat defecten aan deze belangrijke multi-enzymcomplexen door erfelijke afwijkingen ernstige ziekten veroorzaken.

De drie 2-oxozuurdehydrogenase complexen hebben veel gemeenschappelijke structurele en katalytische eigenschappen. Ze zetten allemaal het betreffende substraat stapsgewijs om door middel van deelreacties, die worden gekatalyseerd door drie verschillende enzymen in het complex: een substraatspecifiek 2-oxozuurdehydrogenase (E1), een acyltransferase (E2) en een lipoamide-dehydrogenase (E3). Deze enzymen komen allemaal in meerdere kopieën voor in het complex. De E2-component vormt de structurele kern van het complex door te aggregeren tot een 24-meer of een 60-meer. Aan de E2-kern zijn meerdere kopieën van de E1- en de E3-component niet-covalent gebonden. De E2-component is een multifunctioneel eiwit bestaande uit drie verschillende soorten domeinen, die onafhankelijk van elkaar vouwen en functioneren. Het N-terminale gedeelte van de E2-eiwitketen bestaat uit één tot drie lipoyldomeinen (~ 80 aminozuren elk) waaraan covalent de prosthetische groep lipoaat is gebonden. Tussen de lipoyldomeinen en het C-terminale katalytische domein (~ 29 kDa), dat het acyltransferase actieve centrum bevat en de kern van het complex vormt, bevindt zich een bindingsdomein. Dit kleine domein (~ 35 aminozuren) is betrokken bij de binding van de E1- en/of E3-componenten aan de kern van het complex. De domeinen zijn met elkaar verbonden door flexibele stukken eiwitketen (zgn. linkers), die essentieel zijn voor het goed functioneren van het complex.

De lipoyldomeinen spelen een cruciale rol in het multi-enzymcomplex. Zij zorgen voor het transport van de reactie-intermediären tussen de verschillende enzymen in het

complex, waardoor alle actieve centra in het complex als het ware worden gekoppeld (active-site coupling). Elk lipoyldomein bevat een lipoylgroep, die wordt gevormd door de modificatie van een specifiek lysine-residu met lipoaat. De lipoylgroep transporteert de acylgroepen van E1 naar E2, en vervolgens de reductie-equivalenten van E2 naar E3. De geoxideerde lipoylgroep is dan weer beschikbaar om te worden geacyleerd door de E1-component. De structuur van het lipoyldomein is belangrijk en essentieel voor een efficiënte reactie van zijn lipoylgroep met E1. Vrij lipoamide is een zeer slecht substraat voor E1. Voor *E. coli* PDHC en OGDHC is aangetoond dat de lipoyldomeinen alleen efficiënt worden geacyleerd door de E1-component van hun eigen complex. Dit duidt erop dat de lipoyldomeinen specifiek worden herkend door de E1-component.

Het doel van het onderzoek, dat in dit proefschrift wordt beschreven, is om meer inzicht te krijgen in de interacties tussen lipoyldomeinen en E1-componenten, en in het bijzonder welk deel van het lipoyldomein nu zorgt voor de specifieke herkenning door de E1-component. Door het bepalen en vergelijken van de drie-dimensionale structuren van de lipoyldomeinen van het PDHC en het OGDHC uit *A. vinelandii*, waarvoor de specificiteit in de reductieve acyleringsreactie is bepaald, wordt verwacht al wat meer te weten te komen over deze moleculaire herkenning. De bepaling van de structuur van de lipoyldomeinen vormt in ieder geval een goede structurele basis voor verdere studies op dit gebied, waarvan er een aantal ook in dit proefschrift wordt beschreven. De resultaten van het bovengenoemde werk zullen nu worden samengevat.

Het N-terminale lipoyldomein (aminozuurresiduen 1-79) van de acetyltransferase-component van het PDHC, en het lipoyldomein van de succinyltransferase-component van het OGDHC, zijn beiden gesubcloneerd en tot expressie gebracht in *E. coli* (hoofdstukken 2 en 3). Het vermogen van de *E. coli* cellen tot het lipoyleren van het lipoyldomein is echter niet toereikend, en slechts 5-10% van het geproduceerde lipoyldomein wordt voorzien van een lipoylgroep. De gelipoyleerde en niet-gelipoyleerde vormen van het lipoyldomein kunnen eenvoudig worden gescheiden met behulp van een anionenwisselaar. Het toevoegen van extra lipoaat aan het groeimedium resulteert echter in het volledig lipoyleren van al het geproduceerde lipoyldomein. De vouwing en de modificatie van de recombinante lipoyldomeinen is goed, want ze kunnen allebei worden geacyleerd door de betreffende E1-component. De correcte modificatie van het PDHC lipoyldomein is ook bevestigd door electrospray massaspectrometrie.

Twee-dimensionale homonucleaire en heteronucleaire NMR (kernspinresonantie) metingen hebben geresulteerd in de toekenning van de sequentiële ^1H en ^{15}N resonanties en de secundaire structuur van beide lipoyldomeinen (hoofdstukken 2 en 3). De ^1H -NOESY spectra van de gelipoyleerde en niet-gelipoyleerde vormen van het PDHC

lipoyldomein kunnen vrijwel exact over elkaar worden gelegd, met uitzondering van enkele extra resonanties die kunnen worden toegekend aan de lipoylgroep (hoofdstuk 2). Ook zijn er kleine verschuivingen waar te nemen van resonanties afkomstig van residuen die zich in de directe omgeving van het lipoyl-lysine-residu bevinden. Er zijn echter geen veranderingen waarneembaar in de intensiteit van NOEs die deze residuen verbinden, noch is er een verlies of het verschijnen van extra NOEs geconstateerd. Dit duidt erop dat de structuur van het lipoyldomein niet, of niet noemenswaardig, verandert door de covalentgebonden lipoylgroep. Een uitgebreide vergelijking van de chemische verschuivingen (chemical shifts), NH-uitwisselingsnelheden, NOEs en $^3J_{\text{HN}\alpha}$ koppelingsconstanten wijst op een zeer groot structureel verwantschap tussen beide lipoyldomeinen, alhoewel hun aminozuurvolgorden slechts voor 25% identiek zijn (hoofdstuk 3).

De drie-dimensionale structuren in oplossing van de PDHC en OGDHC lipoyldomeinen zijn bepaald met behulp van distance geometry en/of simulated annealing berekeningen (hoofdstukken 4 en 5). De algemene vouwing van beide lipoyldomeinen is hetzelfde, en kan worden beschreven als een β -barrel-sandwich hybrid (afgeplatte cylinder gevormd door β -platen), een vouwing die nu als kenmerkend wordt beschouwd voor lipoyldomeinen. Het domein bestaat uit twee vrijwel identieke antiparallelle β -sheets die nagenoeg parallel ten opzichte van elkaar staan, en die met elkaar worden verbonden door loops (lussen) en turns (bochten). Elke β -sheet wordt gevormd door drie grote en één kleine β -strand, die rond een goed gedefinieerde kern van hydrofobe aminozuurresiduen liggen. Aan het uiteinde van één van de β -sheets, in een β -turn die twee opeenvolgende β -strands met elkaar verbindt, bevindt zich het lipoyl-lysine-residu, dat zo direct de oplossing insteekt. Precies daartegenover, aan de andere kant van het domein, komen de N- en C-termini samen in de andere β -sheet. Het lipoyldomein is opvallend symmetrisch, en een tweevoudige rotatiesymmetrie-as kan worden gedefinieerd dwars door het centrum van het domein heen. Daarmee kunnen de twee β -sheets op elkaar worden gedraaid. De laatste zes en twee C-terminale aminozuurresiduen van respectievelijk de recombinante PDHC en OGDHC lipoyldomeinen, zijn slecht gedefinieerd in de structuren. Ze behoren tot de flexibele linkers die de lipoyldomeinen met de rest van de acyltransferase polypeptideketen verbinden. De NH-protonen van deze residuen hebben een smallere lijnbreedte in het NMR-spectrum dan de NH-protonen van de rest van het lipoyldomein, hetgeen erop duidt dat deze residuen mobieler zijn.

De algemene vouwing van de tot nu toe bekende lipoyldomeinen, nl. de *A. vinelandii* PDHC en OGDHC lipoyldomeinen en de PDHC lipoyldomeinen van *B. stearothermophilus* en *E. coli*, is nagenoeg hetzelfde. Vooral het feit dat de twee *A.*

vinelandii lipoyldomeinen, waarvan hun specificiteit in de reductieve acyleringsreactie is vastgesteld (hoofdstuk 6), zeer overeenkomstige structuren hebben, suggereert dat de moleculaire herkenning van lipoyldomeinen door de E1-component berust op subtiel verschillen tussen de lipoyldomeinen. Op basis van het nauwkeurig vergelijken van lipoyldomeinstructuren en -aminozuurvolgorden, kunnen een aantal aminozuurresiduen worden voorgesteld, die mogelijk zijn betrokken bij de moleculaire herkenning. Daaronder bevinden zich onder andere aminozuurresiduen die deel uitmaken van een exposed (aan de oplossing blootgestelde) loop, die de eerste twee β -strands met elkaar verbindt, en die in de buurt ligt van het lipoyl-lysine-residu. De grootste structurele verschillen tussen de diverse lipoyldomeinen worden juist in deze loop waargenomen. Ook de twee aminozuurresiduen, die in de aminozuurvolgorde direct na het lipoyl-lysine-residu komen, zouden mogelijk betrokken kunnen zijn bij de specifieke herkenning door E1.

Om de mogelijke betrokkenheid van de exposed loop bij de moleculaire herkenning te onderzoeken, zijn plaatsgerichte mutagenese-experimenten aan de loop van het OGDHC lipoyldomein uitgevoerd (hoofdstuk 6). Ook zijn daartoe cross-acyleringsreacties van de *A. vinelandii* PDHC en OGDHC lipoyldomeinen, gekatalyseerd door *E. coli* complexen, gedaan. Deze experimenten tonen aan dat de exposed loop waarschijnlijk is betrokken bij de specifieke herkenning van lipoyldomeinen door de E1-component, maar dat niet alleen deze loop de specificiteit van de reductieve acyleringsreactie bepaalt. Aanvullende plaatsgerichte mutagenese-experimenten aan deze loop, en aan de andere voorgestelde aminozuurresiduen, zullen nodig zijn om hun rol in de moleculaire herkenning definitief vast te stellen.

CURRICULUM VITAE

Axel Berg is geboren op 28 augustus 1965 te Jutphaas. In 1983 behaalde hij het diploma ongedeeld VWO aan het Cals College te Nieuwegein. Aansluitend werd de studie Moleculaire Wetenschappen aangevangen aan de Landbouwniversiteit (voorheen Landbouwhogeschool) in Wageningen. Tijdens de doctoraalfase werden afstudeervakken gedaan bij de vakgroep Organische chemie (dr. M.C.R. Franssen, prof. H.C. van der Plas: de oxidatie van indolen door chloorperoxidase) en bij de vakgroep Biochemie (dr. A. de Kok, prof. C. Veeger: de expressie in *E. coli* van het gen coderend voor dihydrolipoyl transacetylase uit *Azotobacter vinelandii*). Tevens werden er stages uitgevoerd bij het Preclinical Research Department van Sandoz A.G., Basel, Zwitserland (dr. H. Kocher: electroforetische eigenschappen van chimere antilichamen) en bij het Chemistry Department van Baylor University, Waco, Texas, USA (prof. W.H. Scouten: ontwikkeling van een snelle screeningsmethode voor dye-ligand affiniteitschromotografie). Het doctoraalexamen werd behaald in augustus 1989.

In de periode van oktober 1989 tot mei 1995 is hij als onderzoeker in opleiding (OIO) verbonden geweest aan de vakgroep Biochemie van de Landbouwniversiteit te Wageningen. Hier werd het onderzoek, zoals beschreven in dit proefschrift, uitgevoerd onder supervisie van dr. A. de Kok, dr. J. Vervoort en prof. C. Veeger. Het onderzoek werd gefinancierd door de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO) onder toezicht van Scheikundig Onderzoek Nederland (SON).

Van juni 1995 tot juni 1996 is hij als vervangend dienstplichtige werkzaam geweest binnen de Metallo-eiwit en Protein-engineering groep (prof. G.W. Canters: NMR aan paramagnetische eiwitten) van het Leids Instituut voor Chemieonderzoek aan de Rijksuniversiteit Leiden.

LIST OF PUBLICATIONS

Hanemaaijer, R., Westphal, A.H., Berg, A., van Dongen, W., de Kok, A. & Veeger, C. (1989) The gene encoding dihydrolipoyl transacetylase from *Azotobacter vinelandii*. Expression in *Escherichia coli* and activation and isolation of the protein, *Eur. J. Biochem.* 181, 47-53.

Berg, A. & Scouten, W.H. (1990) Dye-ligand centrifugal affinity chromatography, *Bioseparation 1*, 23-31.

Schulze, E., Westphal, A.H., Berg, A. & de Kok, A. (1990) Time-resolved fluorescence studies on mutants of the dihydrolipoyl transacetylase (E2) component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*, *FEBS Lett.* 273, 46-50.

Berg, A., de Kok, A. & Vervoort, J. (1994) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the N-terminal lipoyl domain of the dihydrolipoyl transacetylase component of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*, *Eur. J. Biochem.* 221, 87-100.

De Kok, A., Berg, A., Van Berkel, W., Fabisz-Kijowska, A., Westphal, A.H., van den Akker, F., Mattevi, A. & Hol, W.G.J. (1994) The pyruvate dehydrogenase complex from *Azotobacter vinelandii*, in *International symposium on flavins and flavoproteins* (Yagi, K., ed.) pp. 535-544, Walter de Gruyter, New York, Berlin.

Berg, A., Smits, O., de Kok, A. & Vervoort, J. (1995) Sequential ^1H and ^{15}N nuclear magnetic resonance assignments and secondary structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*. Evidence for high structural similarity with the lipoyl domain of the pyruvate dehydrogenase complex, *Eur. J. Biochem.* 234, 148-159.

Berg, A., Vervoort, J. & de Kok, A. (1996) Solution structure of the lipoyl domain of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*, *J. Mol. Biol.* 261, 432-442.

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

Berg, A., Vervoort, J. & de Kok, A. (1997) Three-dimensional structure in solution of the N-terminal lipoyl domain of the pyruvate dehydrogenase complex from *Azotobacter vinelandii*, *Eur. J. Biochem.*, in the press.

Berg, A., Westphal, A.H., van Montfort, B.A. & de Kok, A. (1996) Specificity of reductive acylation of the lipoyl domains of the 2-oxo acid dehydrogenase multienzyme complexes from *Azotobacter vinelandii*, to be submitted.

Dennison, C., Berg, A., de Vries, S. & Canters, G.W. (1996) ^1H NMR studies of the paramagnetic Cu_A centre of cytochrome oxidase, *FEBS Lett.* 394, 340-344.

Dennison, C., Berg, A. & Canters, G.W. (1996) A ^1H NMR study of the paramagnetic active site of the Cu_A variant of amicyanin, *Biochemistry*, in the press.

Ubbink, M., Pfuhl, M., van der Oost, J., Berg, A. & Canters, G.W. (1996) NMR assignments and relaxation studies of *Thiobacillus versutus* cytochrome *c-550* indicate the presence of a highly mobile 13-residues long C-terminal tail, *Protein Science*, in the press.

Vervoort, J., Berg, A., Vossen, P., Vogelsang, R., Joosten, M.H.A.J. & de Wit, P.J.G.M. (1996) Sequence specific ^1H -NMR assignments, secondary structure and global fold of the race-specific elicitor AVR9 of the tomato pathogen *Cladosporium fulvum*, submitted to *Eur. J. Biochem.*

NAWOORD

MIJN SPECIALE DANK AAN: PROFESSOR C. VEEGER EN AART DE KOK, MIJN PROMOTOR EN MIJN CO-PROMOTOR, VOOR DE SUPERVISIE EN DE BEGELEIDING, JACQUES VERVOORT, MAÎTRE JACQUES, VOOR JE BIJZONDERE HULPVAARDIGHEID EN HET SAMEN PIELEN ACHTER (NEE, VÓÓR) DE COMPUTER, ADRIE WESTPHAL, VOOR ALLE HULP BIJ VEELERLEI EXPERIMENTEN EN HET VELE DAT JE MIJ HEBT GELEERD OVER CLONEREN, ZUIVEREN, COMPLEXEN EN DE REST, FRANK VERGELDT, VOOR AL JE ONDERSTEUNING EN MEER MET COMPUTERS EN PROGRAMMA'S, RENÉ WIENTJENS EN MARTINA DUYVIS, MIJN PARANIMFEN EN OUD-KAMERGENOTEN, VOOR JULLIE VRIENDSCHAP EN GEZELLIGHEID, LAURA, MARTIN EN BERY, VOOR DE ADMINISTRATIE EN HULP, ADRIE DE JAGER EN ARIE VAN HOEK, VOOR JULLIE ASSISTENTIE EN DE GOEDE CONDITIE VAN JULLIE FANTASTISCHE APPARATUUR, DE STUDENTEN ARTHUR JAARTSVELD, HANS MASSELINK, LUUC VAN LANGEN, OTTO SMITS, ANNECHIEN HENGEVELD, ANTON VEENSTRA EN BART VAN MONTFORT, VOOR JULLIE BIJDRAGE AAN HET ONDERZOEK, DE BOLLE VAN BREDA, STOKKIE, BEENIE BOY, KAANTJE, PHIL & PAP, PIERIKSMA, PROFESSOREN VERHAGEN EN WESTPHAL, RIK TROM, WILD-TYPE WILLY, EPPIE, DUKVIS, GJOEP, ANDERE LAB EN/OF LOTGENOTEN ZONDER BIJNAAM, EN OVERIGE (EX-)LEDEN VAN DE VAKGROEP BIOCHEMIE, VOOR DE FANTASTISCHE SFEER EN JULLIE COLLEGIALITEIT, ELLES STEENSMA, MIJN KAMERGENOTE, VOOR DE GEZELLIGHEID EN DE GEZELLIGHEID, MARC, FRANK, ANTONIO EN MARTINA, VOOR HET HEERLIJKE ETEN EN DE BIJBEHORENDE INSPIRERENDE DISCUSSIES, DE GETROUWE LOBURG-GANGERS, VOOR DE KONINCKLIJKE UURTJES, DE 'MOLECULEN CLUB', VOOR DE ONVERGETELIJKE WEEKENDS, PROFESSOR CANTERS EN DE METALLO-EIWIT GROEP UIT LEIDEN, VOOR JULLIE GASTVRIJHEID, BELANGSTELLING EN DE LEUKE TIJD, EN TOT SLOT MAUREEN, MIJN OUDERS, ZUSJE, FAMILIE, VRIENDEN EN BEKENDEN, VOOR JULLIE ONVOORWAARDELIJKE STEUN, BETROKKENHEID EN LIEFDE.