Biochemistry

 Volume 30, Number 35 September 3, 1991

Perspectives in Biochemistry

Domains, Motifs, and Linkers in 2-Oxo Acid Dehydrogenase Multienzyme Complexes: A Paradigm in the Design of a Multifunctional Protein†

Richard N. Perham

Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, England

Received May 28, 1991

Multienzyme complexes are noncovalent aggregates of protein subunits that catalyze successive steps of a multistep chemical reaction. It is almost self-evident that such a system can have properties that go beyond the sum of those of its parts and these have been reviewed before (Reed, 1974; Perham, 1975; Hammes, 1981) under three main headings: (1) enhancement of catalytic activity, (2) substrate channeling, and (3) the "hot potato" hypothesis. The first two of these take account of the limited diffusion required of a common intermediate if the active sites of successive enzymes are suitably juxtaposed. This can extend to the directed diffusion of the intermediate, even to its covalent transfer, which may be of particular importance in selecting between competing pathways at branch points of metabolism. The hot potato hypothesis (Perham, 1975) recognizes that intramolecular transfer in a protein milieu can protect an intermediate that is insufficiently stable to survive the normal diffusion process in aqueous solution at neutral pH and, say, 37 °C. The scope of active-site chemistry that might be employed is thus widened.

In this paper I discuss several lines of work on the 2-oxo acid dehydrogenase multienzyme complexes that have uncovered some unexpected features of the enzyme proteins. Despite the formidable size of these protein complexes, the combined techniques of limited proteolysis, NMR spectroscopy, and protein engineering have proved remarkably effective in tackling the problem of their structure and function. Apart from providing a new structural basis for understanding the mechanism of these enzyme complexes in particular, the results illuminate the properties of multienzyme complexes in general. In addition they put flesh on another principle, that of parsimony in the choice of protein subunits in the construction of macromolecular assemblies (Perham, 1975; Perham et al., 1978), while disclosing an apparent disregard of this principle in certain aspects of the domain structure of the enzyme complexes that remains a puzzle. There are useful lessons to be learned here about the design of multidomain proteins, which carry important implications for the analysis of protein–protein interactions in a wide range of biological systems.

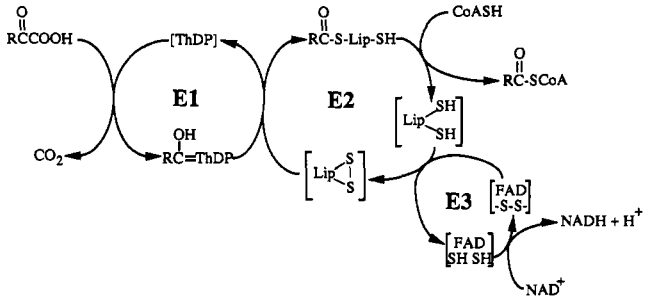
The interested reader who wishes to learn more about these complexes, their genetics, their part in metabolism, and their importance in clinical medicine, is referred to the proceedings of a recent conference devoted exclusively to them in all their different aspects (Roche & Patel, 1989; Patel & Roche, 1990).

Symmetry, Subunit Structure, and Basic Mechanism

2-Oxo acid dehydrogenase complexes (Reed, 1974) consist of multiple copies of three different enzymes that catalyze the oxidative decarboxylation of 2-oxo acids, as shown schematically in Figure 1. The three constituent enzymes of the pyruvate dehydrogenase (PDH) complex are pyruvate decarboxylase [pyruvate dehydrogenase (lipoamide), E1p, EC 1.2.4.1], dihydrolipoamide acetyltransferase (E2p, EC 2.3.1.12) and dihydrolipoamide dehydrogenase (E3, EC 1.8.1.4). For the 2-oxoglutarate dehydrogenase (2OGDH) complex, the corresponding enzymes are 2-oxoglutarate decarboxylase (E1o, EC 1.2.4.2), dihydrolipoamide succinyltransferase (E2o, EC 2.3.1.61) and dihydrolipoamide dehydrogenase (E3, EC 1.8.1.4). The three short branched-chain 2-oxo acids produced by transamination of the amino acids leucine, isoleucine, and valine are oxidatively decarboxylated by a single complex (BCODH) of comparable structure (Reed & Yeaman, 1987; Randle et al., 1987). During catalysis, the substrate is attached by thioester linkage to lipoyllysine residues of the E2 chains (Figure 1). Spin-label experiments with the PDH complex of *Escherichia coli* have demonstrated that the dithiolane ring of the lipoyl group has a correlation time of ca. 10^{-9} s, which indicates that it has essentially total freedom to rotate in the complex (Ambrose & Perham, 1976; Grande et al., 1976) and to act as the classic "swinging arm", conveying substrate between the three successive active sites (Reed, 1974). In all the complexes, substrate specificity resides

† Work performed in the author's laboratory was supported by grants from the Science and Engineering Research Council and The Wellcome Trust.

0006-2960/91/0430-8501$02.50/0



Net Reaction

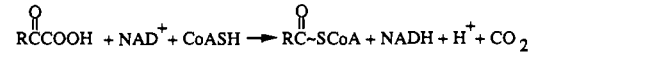


FIGURE 1: Schematic diagram of the reaction mechanism of 2-oxo acid dehydrogenase complexes [after Reed (1974)]. ThDP, thiamin diphosphate; Lip, lipoic acid.

in the E1 and E2 components, whereas E3 fulfills a common service function to reoxidize the enzyme-bound dihydrolipoamide (Figure 1). Thus, in keeping with the principle of parsimony in the assembly of biomolecular structures, the E3 component is normally identical in the various complexes elaborated by a given organism. The only known exception is that of *Pseudomonas putida*, which has three different E3 genes, of which two are known to encode dihydrolipoamide dehydrogenases specific to the branched-chain 2-oxo acid and to the pyruvate and 2-oxoglutarate dehydrogenase complexes, respectively (Burns et al., 1989). The reason for this is obscure.

The 2-oxo acid dehydrogenase complexes have a high M_r [$(5-10) \times 10^6$, depending on the underlying symmetry] and are seen as particles of 30–40 nm in the electron microscope (Reed, 1974; Henderson et al., 1979; Oliver & Reed, 1982; Hackert et al., 1989). They are thus substantially larger than a ribosome. The E2 component forms a structural core to which the E1 and E3 components, both dimers, are bound tightly but noncovalently. In the PDH and 2OGDH complexes of *E. coli* (Reed, 1974; Danson et al., 1979) and *Azotobacter vinelandii* (Hanemaaijer et al., 1989) and the 2OGDH (Reed, 1974) and BCODH (Griffin et al., 1988; Hackert et al., 1989) complexes of mammals, this core consists of 24 copies of the E2 chain arranged with octahedral symmetry, whereas in the PDH complexes of mammals (Reed, 1974), *Bacillus stearothermophilus* (Henderson & Perham, 1980), *Bacillus subtilis* (Lowe et al., 1983), *Streptococcus faecalis* (Reed, 1974; A. Allen, J. N. Berman, L. C. Packman, and R. N. Perham, unpublished work), and *Saccharomyces cerevisiae* (Keha et al., 1982), the core comprises 60 E2 chains arranged with icosahedral symmetry. In the octahedral PDH complexes the E1p polypeptide chains are unsplit, whereas in the icosahedral PDH complexes the E1p component is composed of E1α and E1β subunits, i.e., E1α_2E1β_2. Thus, the PDH complexes of eukaryotes, which are located in the mitochondria, and those of Gram-positive *Bacillus* spp. and *S. faecalis*, form a class that can be distinguished in subunit composition and E2 symmetry from the corresponding complexes of the Gram-negative bacteria, *E. coli* and *A. vinelandii*. This may be of interest in considering the evolutionary origins of mitochondria (Henderson et al., 1979; Keha et al., 1982).

The Gram-positive bacterial and eukaryotic PDH complexes themselves differ in that the latter are importantly subject to control in vivo by a system of phosphorylation–dephosphorylation of the E1α subunits, as are the eukaryotic BCODH complexes. This metabolic control will not be considered further here, as detailed reviews can be found elsewhere (Reed & Yeaman, 1987; Randle et al., 1987; Yeaman, 1989; Patel & Roche, 1990). The structure of the E2p core of the ox heart and kidney (and presumably other mammalian) PDH complexes is additionally complicated by the presence of a few (about 6) copies of an extra lipoic acid containing polypeptide chain, protein X (Jilka et al., 1986; Hodgson et al., 1986). A similar protein is present in the PDH complex from yeast (Behal et al., 1989). It may be concerned (see below) with binding the E3 component and facilitating its participation in electron transfer (Gopalakrishnan et al., 1989; Neagle et al., 1991; Lawson et al., 1991). There is no evidence as yet for the existence of a protein X in the E2p core of the PDH complex of prokaryotes, nor in 2OGDH or BCODH complexes, but this cannot be ruled out.

Despite the participation of three enzymes in the overall reaction (Figure 1), the active sites are not represented in equal numbers in the enzyme particle. For example, each E2p chain of the PDH complex of *E. coli* is able to bind up to two E1p chains (one E1p dimer) (Bates et al., 1975; Reed et al., 1975) and, although the E1p:E2p:E3 chain ratios in the purified complex have been reported widely (Reed, 1974; Hammes, 1981; Yang et al., 1985) as 1.0:1.0:0.5, there is other evidence that these chain ratios are not always strictly adhered to (Bates et al., 1975; de Kok & Westphal, 1985). Competition between E1p and E3 subunits for space on the surface of the E2p core during the assembly process (Reed et al., 1975) might be the basis, at least in part, for any variations in subunit stoichiometry. The E1p dimers appear to bind along the 12 edges of the octahedral core, with the E3 dimers integrated into the 6 faces (Oliver & Reed, 1982; Yang et al., 1985; Reed & Hackert, 1990). Similar considerations apply to the 2-oxo acid dehydrogenase complexes from other sources, especially the PDH complexes of icosahedral symmetry whose E3:E2p ratios are significantly less than 0.5 (Reed, 1974; Henderson & Perham, 1980; Wu & Reed, 1984). A structural heterogeneity in the *E. coli* PDH complex consistent with these observations has been inferred from two different methods of ultracentrifugation analysis (Schmitt & Cohen, 1980; Gilbert & Gilbert, 1980) although no such heterogeneity could be detected by STEM measurements (Yang et al., 1985). Moreover, the 2OGDH complex of *E. coli*, with E1o:E2o:E3 chain ratios of 0.5:1.0:0.5 (Reed, 1974), has been postulated to exhibit more than 100 000 different quaternary structures as a result of different permutations in the packing of E1o and E3 subunits round the E2o core during assembly (Wagenknecht et al., 1986, 1987). In any event, these enzyme complexes plainly do not conform entirely to the simple principles of exact assembly expected from a study of less complicated oligomeric systems (Perham, 1975), and we must evidently seek a molecular model that is not dependent on a simple direct transfer of substrate between three active sites organized with 1:1:1 stoichiometry in a strict geometrical arrangement.

Domains and Linkers in the E2 Polypeptide Chains

The structural and mechanistic cores of the 2-oxo acid dehydrogenase complexes are provided by the E2 polypeptide chains. Given their location as core enzymes, around which the E1 and E3 components are assembled, the E2 components are remarkably sensitive to limited proteolysis, readily shedding fragments of the E2 chains carrying the lipoyl groups. These fragments, or lipoyl domains, are therefore thought to protrude physically from an inner part of the E2 core, interdigitating between the E1 and E3 subunits. The residual parts of the E2 chains retain the dihydrolipoamide acyltransferase activity and remain assembled as the inner core, still exhibiting their octahedral or icosahedral symmetry [see Bleile et al. (1981)

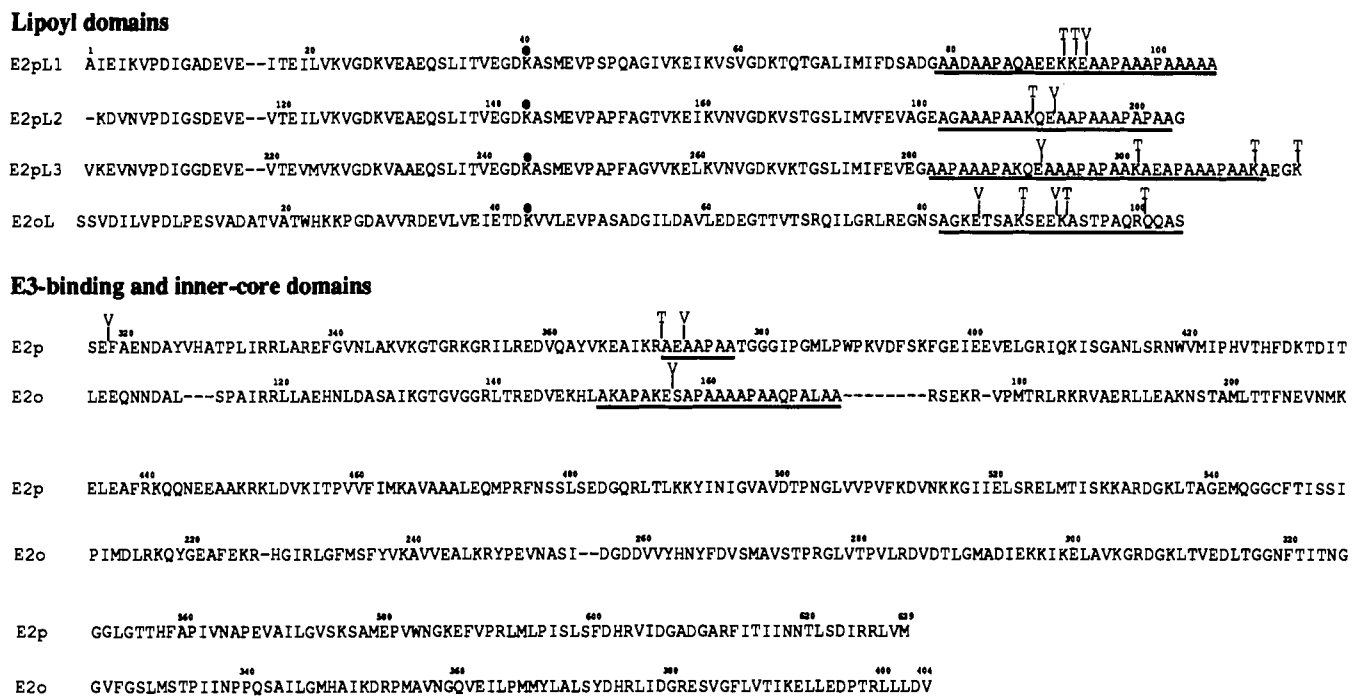


FIGURE 2: Amino acid sequences of the E2p and E2o chains of the PDH and 2OGDH complexes of *E. coli*. Segments of polypeptide chain rich in alanine, proline, and charged amino acids, which link folded domains, are underlined. Cleavage sites for trypsin (T) and *S. aureus* V8 proteinase (V) are indicated. ●, lipoyllysine residue.

and Duckworth et al. (1982) and references therein]. The lipoyl fragments, unlike the inner part of the E2 core, are difficult to visualize in negatively stained electron micrographs, perhaps because of physical flexibility (Bleile et al., 1979; Perham & Wilkie, 1980; Oliver & Reed, 1982), and they behave hydrodynamically as unusually swollen or extended structures (Bleile et al., 1979, 1981).

It has become apparent that the highly segmented structure of the E2 chains holds the key to a fuller understanding of the structure and mechanism of the complexes. The unexpected variety and almost gothic complexity of the domain and linker structure of the E2 chains of 2-oxo acid dehydrogenase complexes has emerged from a study of their patterns of limited proteolysis, combined with a knowledge of the DNA sequences of the structural genes that encode them (Perham & Packman, 1989; Guest et al., 1989; Reed & Hackert, 1990). The E2p and E2o chains of the PDH and 2OGDH complexes of *E. coli* have figured prominently in this approach. The amino acid sequences of the E2p (Stephens et al., 1983) and E2o (Spencer et al., 1984; Packman & Perham, 1987) chains, together with the proteolytic cleavage sites for trypsin and *Staphylococcus aureus* V8 proteinase (Packman & Perham 1986, 1987), are shown in Figure 2. The clustering of the proteolytic cleavage sites in the interdomain segments of the polypeptide chains defines the approximate limits of the domains and permits their purification (Packman et al., 1984a; Packman & Perham, 1986, 1987). Thus, the N-terminal half of the E2p chain (66 kDa) consists of three highly homologous lipoyl domains, each of about 80 amino acid residues, linked together in tandem array by long peptide sequences (about 25–30 amino acid residues) rich in alanine, proline, and charged amino acids. These are followed by a domain, about 50 amino acids in length, whose primary purpose at least appears to be that of binding the E3 subunits to the E2 core (Packman & Perham, 1986; Radford et al., 1987) and which is linked in turn to a large C-terminal region (29 kDa) by another (but less conspicuous) region rich in alanine, proline, and charged amino acids. It is this C-terminal (29 kDa) fragment that remains assembled as the octahedral inner core and that retains the binding site for the E1p subunits and the active site for the acetyltransferase step of the complex reaction (Radford et al., 1987). The E2o chain of *E. coli* is homologous in sequence and is organized similarly (Figure 2) except that it has only one lipoyl domain and the compositions of the interdomain regions differ somewhat from those of their counterparts in the E2p chain (see below).

This segmented structure for the E2 chains turns out to be general and widespread. The E2p chain from the PDH complex of *B. stearothermophilus* can be cleaved successively with chymotrypsin and trypsin to produce the single lipoyl domain, the E3-binding domain, and the residual (icosahedral) inner core domain, all of which show significant sequence homology with the corresponding domains from the E2p and E2o chains of *E. coli* (Packman et al., 1988; Borges et al., 1990). An important difference is that the E3-binding domain (in this as in other icosahedral complexes) plays a part also in binding the E1p component, for its removal by proteolysis causes the complex to disassemble to a naked inner core (Perham & Wilkie, 1980; Packman et al., 1988). Similar segmented structures for other E2 chains have been inferred from the DNA sequences of genes or cDNAs encoding them, notably the E2p chains from PDH complexes of *A. vinelandii* (Hanemaaijer et al., 1988), human placenta (Coppel et al., 1988), human liver (Thekkumkara et al., 1988), yeast (Niu et al., 1988), *B. subtilis* (Hemilä et al., 1990), and *S. faecalis* (Allen & Perham, 1991); the E2o chain from the 2OGDH complexes of *B. subtilis* (Carlsson & Hederstedt, 1989) and *A. vinelandii* (Westphal & de Kok, 1990); and the E2b chains from the BCODH complexes of *Ps. putida* (Burns et al., 1988), ox liver (Griffin et al., 1988), human liver (Danner et al., 1989), and human placenta (Nobukuni et al., 1989). These sequences have recently been listed and their extensive similarities compared (Russell & Guest, 1991a).

As pointed out in detail elsewhere (Perham & Packman, 1989), there is no simple correlation between the number of lipoyl domains per E2 chain, the symmetry of the E2 core, and

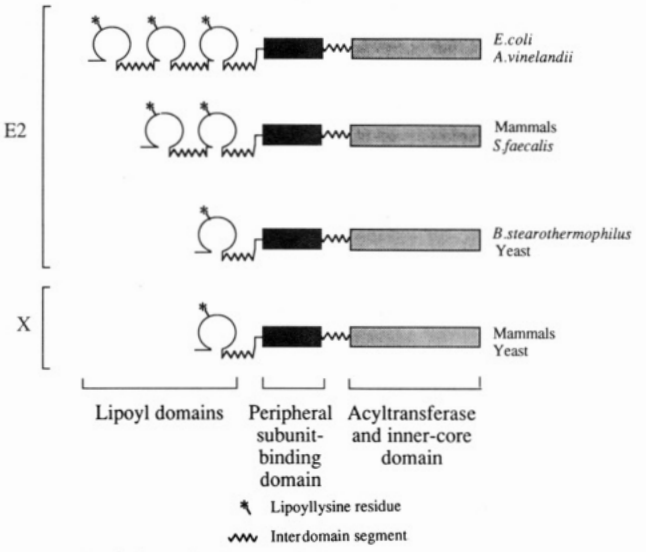


FIGURE 3: Schematic layout of domains in the E2 and protein X subunits of 2-oxo acid dehydrogenase complexes.

the phylogeny of the source of the E2 core. Thus, the octahedral PDH complexes of *E. coli* and *A. vinelandii* (prokaryotes) have three, the icosahedral *S. faecalis* (prokaryote) and mammalian (eukaryotic) PDH complexes have two, and the icosahedral *Bacillus* (prokaryote) and yeast (eukaryote) enzymes have one (Figure 3). No 2OGDH or BCODH complex has yet been found with more than one lipoyl domain per E2 chain, and all of these, with one exception, are of octahedral symmetry. The exception is the BCODH complex of *B. subtilis* and *B. stearothermophilus*, which appears to be identical with the icosahedral PDH complex, an example of multiple substrate specificity, related perhaps to the requirement for branched chain fatty acids as essential components of the cell membranes of Gram-positive organisms (Lowe et al., 1983).

Protein X, the extra component found thus far only in PDH complexes of eukaryotic origin (see above), has a similarly segmented structure. Limited proteolysis of the ox heart PDH complex (Rahmatullah et al., 1989; Neagle et al., 1989) and sequence analysis of the gene encoding the yeast protein X (Behal et al., 1989) indicate that protein X from both sources has only a single lipoyl domain and a peripheral subunit-binding domain, which are homologous with their counterparts in the yeast E2 chain (Figure 3). Most surprisingly, perhaps, the sequence homology is limited to this N-terminal half of the molecule; the C-terminal half has little or no similarity in sequence to the E2 protein and, in particular, lacks a C-terminal segment of about 80 amino acid residues (Behal et al., 1989). This may indicate that it is unable to catalyze the acyltransferase reaction associated with E2 (see below).

Substrate Channeling and the Role of the Lipoyl Domains

In all 2-oxo acid dehydrogenase complexes studied thus far, the lipoyl groups are attached to small protein domains, similar in sequence and presumably in three-dimensional structure. Why is this? The swinging arm mechanism does not require it, except to limit diffusion of the dithiolane ring (see above), and it is not demanded by the chemical mechanism (Figure 1). Indeed, free lipoic acid or lipoamide can act as substrates for the E2p and E3 components of the *E. coli* PDH complex (Reed et al., 1958), yet direct reductive acetylation of these compounds by the E1p component could not be detected to any appreciable extent (Reed et al., 1958; Reed, 1966). On the other hand, lipoyl domains excised proteolytically from the PDH complexes of ox heart (Bleile et al., 1981), *E. coli* (Packman et al., 1984a), and *B. stearothermophilus* (Packman et al., 1984b) are reductively acetylated by the relevant E1p component in the presence of ThDP, Mg^{2+}, and pyruvate.

Only the *R*-enantiomers of lipoate and dihydrolipoate function with the PDH and 2OGDH complexes of *E. coli* (Yang & Frey, 1989). Taking this stereospecificity into account, free lipoamide is actually a very poor substrate for *E. coli* E1p, with a K_m in excess of 2 mM and a k_{cat}/K_m of 3.0 M^{-1} s^{-1} (Graham et al., 1989). A lipoylated decapeptide with an amino acid sequence identical with that surrounding each of the three lipoyllysine residues in the *E. coli* E2p chain is equally ineffective as a substrate, whereas the free lipoyl domains are reductively acetylated under identical conditions with a K_m of approximately 26 μM and a k_{cat}/K_m of approximately 3.0×10^4 M^{-1} s^{-1}. Moreover, the lipoyl domains from the E2p and E2o chains of *E. coli* function well as substrates only with the E1p and E1o components of their respective parent complexes (Graham et al., 1989).

As judged by these values for k_{cat}/K_m, the ability of lipoic acid to act as a substrate for *E. coli* E1p is greatly enhanced (k_{cat}/K_m raised by a factor of 10000) by virtue of its attachment to a protein domain. The lipoyl domain acts further to confer specificity on the reductive acylation of the dithiolane ring by the cognate E1 component, and the feebleness of the lipoylated decapeptide as a substrate indicates that a folded lipoyl domain is necessary to the reaction (Graham et al., 1989), though it is not simply a question of enhanced binding of the lipoyl domain (Graham & Perham, 1990). These results have yet to be extended to other complexes, but it is known that free lipoyl domains from the PDH complexes of ox heart (Bleile et al., 1981) and *B. stearothermophilus* (Packman et al., 1984b) are reductively acetylated by the relevant E1p component, with K_m values around 20 μM. Therefore, it is more than likely that, in all 2-oxo acid dehydrogenase complexes, the lipoyl group must be attached to a specific lipoyl domain to promote reductive acylation of the dithiolane ring. This is all the more remarkable for the fact that the dithiolane ring is at the end of a 1.4-nm swinging arm and is free to rotate with respect to the bulk of the protein (see above). It is reasonable to infer that some process of molecular recognition and interaction between the lipoyl domain and its cognate E1 component forms an essential part of this first step in the reaction catalyzed by the enzyme complex. The specificity conferred by the lipoyl domain provides a beautiful mechanism for substrate channeling, the lipoyl group that undergoes reductive acylation being confined to the lipoic acid residue covalently attached to the intended E2 component.

Although the underlying requirement for lipoyl domains in 2-oxo acid dehydrogenase complexes has thus been laid bare, another and initially more obvious question to ask is why the *E. coli* and *A. vinelandii* PDH complexes have three lipoyl domains per E2 chain and the mammalian and *S. faecalis* PDH complexes have two, whereas other 2-oxo acid dehydrogenase complexes function well with just one lipoyl domain per E2 chain. Protein engineering experiments have thrown much light on this problem: deletion of one or even two of the lipoyl domains in the E2 chain of the *E. coli* PDH complex (Figure 4) had no adverse effect on its assembly or its catalytic activity (Guest et al., 1985), and the conversion of the lipoyllysine residue to a glutamine residue of course rendered the lipoyl domain nonfunctional and the complex inactive, but the creation of E2 chains containing various combinations and permutations (+/−, −/+, −/−/+, +/+/−)

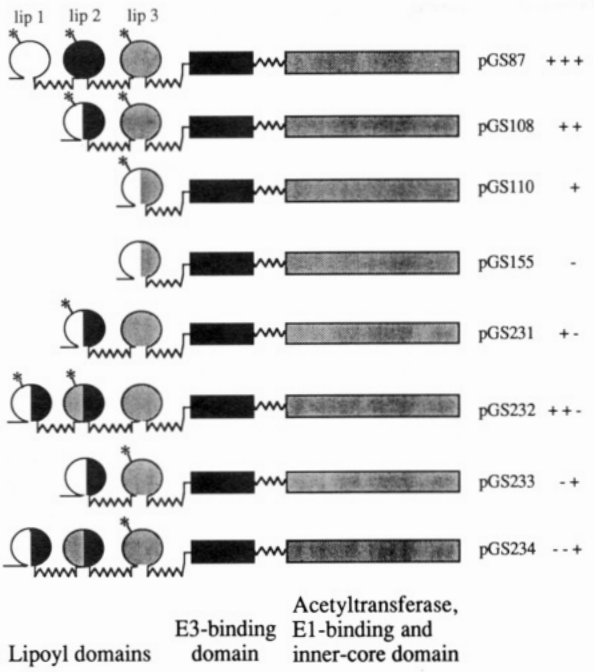


FIGURE 4: Schematic layout of domains in wild-type and genetically engineered E2 chains of the PDH complex of *E. coli*. +, a functional lipoyl domain; −, a lipoyl domain in which the lipoyllysine residue has been replaced with glutamine.

of functional (+) and nonfunctional (−) lipoyl domains (Figure 4) restored the catalytic activity to essentially wild-type levels, with all the functional domains being capable of participating in the reaction (Allen et al., 1989).

Several important conclusions can be deduced from these results. First, the lipoyl domains are clearly independent entities in the folding process, and the folding of the other domains in the E2 chains, together with the ensuing assembly of the complex, is independent of the lipoyl domains. Second, one lipoyl domain is sufficient for complex activity. Nature appears to have been unnecessarily prodigal in furnishing some 2-oxo acid dehydrogenase complexes with a superfluity of lipoyl domains, in contravention of the normal principle of parsimony in biomolecular assembly. The reason for this remains obscure. Third, any catalytic mechanism that requires serial transfer of the acetyl group between outer and inner lipoyl domains (or vice versa) in a given E2 chain can be ruled out. Fourth, each lipoyl domain can function independently of its neighbors, heightening our appreciation of the need for conformational flexibility in the E2 chain to permit the lipoyl domains to participate in active-site coupling (see below).

Structure and Lipoylation of the Lipoyl Domain

Despite much effort in numerous laboratories over many years, the 2-oxo acid dehydrogenase complexes have resisted all efforts to crystallize them, except as assemblies with what are now known to be proteolyzed E2 cores (DeRosier et al., 1971; Fuller et al., 1979). The E3 components of the yeast (Takenaka et al., 1988) and *A. vinelandii* (Schierbeek et al., 1989) PDH complexes have been crystallized and high-resolution structures are awaited. The domain-and-linker structure of the E2 chain, however, opens up another approach: systematic determination of the structures of the individual domains and of the linkers that join them. The structure of the lipoyl domain is of unusual interest, given the part it plays in promoting the reductive acylation of its pendant lipoyl group (see above).

A subgene encoding the lipoyl domain (residues 1–85) of the *B. stearothermophilus* PDH complex has been created and overexpressed in *E. coli* (Dardel et al., 1990). The product was found to exist in three forms: an unlipoylated domain (80%), a domain correctly lipoylated on Lys-42 (16%), and a domain aberrantly acylated on Lys-42 with an octanoyl group (4%). It appears that the overexpression of the subgene swamps the ability of the *E. coli* cell to lipoylate the domain and, moreover, induces the mismodification with an octanoyl group, a likely biosynthetic precursor of lipoic acid (Dardel et al., 1990). Octanoylation of an *E. coli* lipoyl domain overexpressed in a lipoic acid deficient mutant of *E. coli* has also been observed (Ali et al., 1990), but no mismodification of the lipoyl domains of the native *E. coli* PDH complex could be detected (Packman et al., 1991). These results are the first report of an aberration in the posttranslational modification of a recombinant protein and serve as a timely reminder of the need to check as carefully as possible the product of a foreign gene overexpressed in *E. coli* (or another host).

It is worth noting that the relevant enzymes of the *E. coli* cell can evidently recognize the lipoyl domain from the *B. stearothermophilus* PDH complex and correctly lipoylate it (Dardel et al., 1990). On the other hand, a gene encoding the E2 chain of the ox liver BCODH complex can be expressed in *E. coli* but the product is not lipoylated, although it can be by subsequent exposure to ox liver mitochondrial extracts (Griffin et al., 1990). This indicates that the lipoylating enzymes can distinguish between lipoyl domains, an interesting problem of molecular recognition in the field of posttranslational modification that should repay further study. In addition, all three lipoyl domains of the native *E. coli* E2p chain are fully lipoylated (Packman et al., 1991). This observation is in marked contrast with reports from many laboratories that a maximum of approximately two lipoyl groups can be reductively acetylated per E2p chain in the *E. coli* PDH complex [reviewed by Packman et al. (1984a)] and suggests that previous results have suffered from systematic errors in estimating the extent of reductive acetylation.

The availability of large amounts of the *B. stearothermophilus* lipoyl domain as the product of an overexpressed gene in *E. coli* has made it possible to tackle the structure of the domain by NMR spectroscopy. The 400-MHz ^{1}H NMR spectra of the lipoylated and unlipoylated forms of the domain are essentially identical, apart from resonances attributable to the protons of the lipoyl group (Dardel et al., 1990). Thus, there is no little or no conformational change in the protein as the result of the posttranslational modification. The key to the facilitation of the reductive acylation of the dithiolane ring (see above) must therefore lie in the structure of the lipoyl domain and its interaction with the E1 component. The secondary structure of the lipoyl domain from the *B. stearothermophilus* PDH complex (Dardel et al., 1991) is shown in Figure 5. The domain is formed chiefly from two large regions of antiparallel β-sheet, with the N- and C-terminal ends of the folded chain close in space in one of them. The lipoyllysine residue is prominently displayed in the other, on the only sharp turn in the structure. Thus it is easy to see, at least in principle, how the lipoyl domain might present the exposed lipoyllysine residue to the active site of the E1 component. Full and detailed structures of this domain and of a similar lipoyl domain from the *E. coli* PDH complex (F. Dardel, J. D. F. Green, E. D. Laue, and R. N. Perham, unpublished work) are nearing completion, which will allow this aspect of the complex mechanism to be explored directly.

Flexible Linkers in the E2 Polypeptide Chains

With the observation that the E2 chains in these enzyme complexes are highly susceptible to proteolysis and the rec-

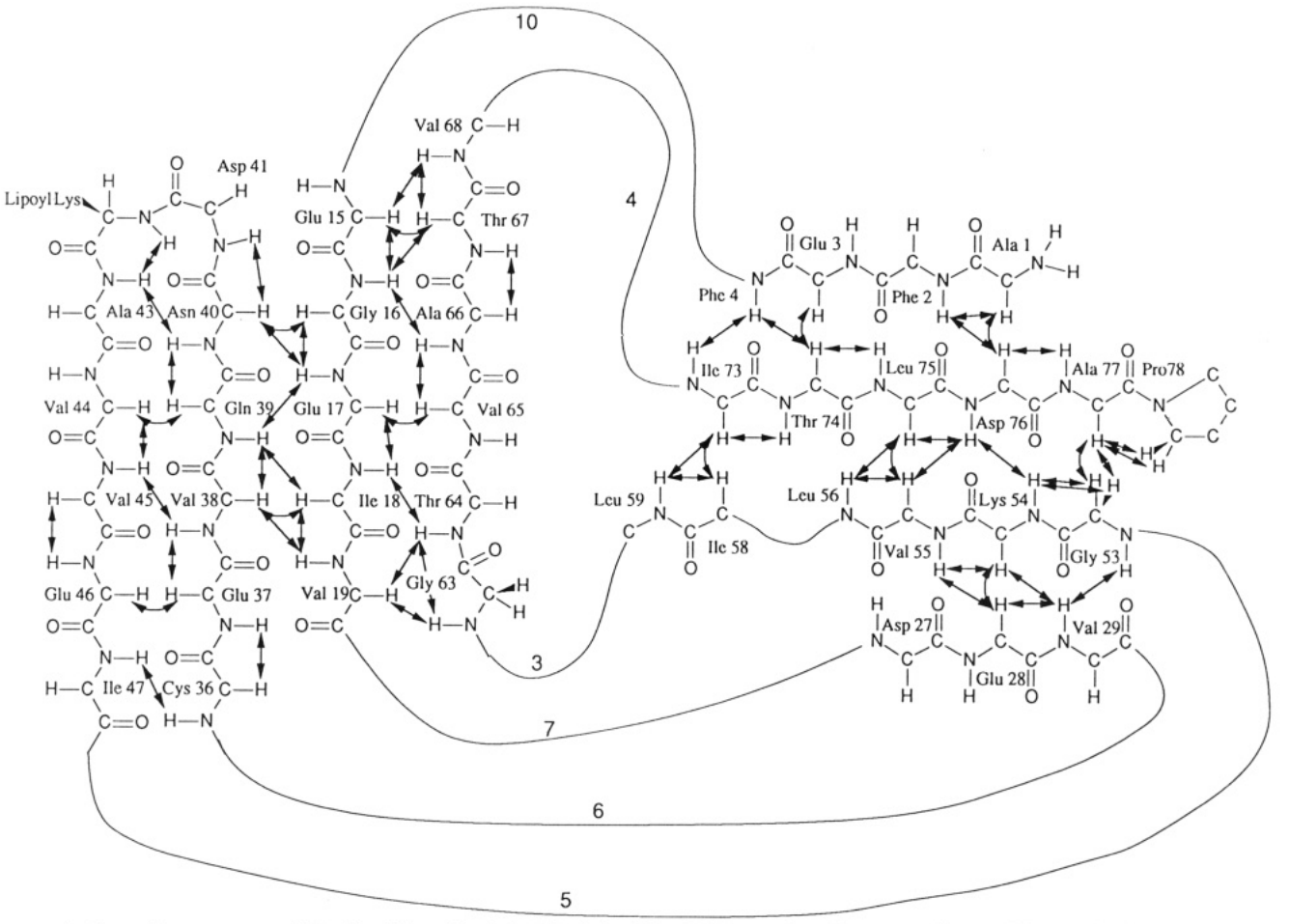


FIGURE 5: Secondary structure of the lipoyl domain of the E2 chain of the PDH complex of *B. stearothermophilus*. NOESY connectivities used to establish the positions of β-strands are indicated by double-headed arrows. The numbers on the lines connecting the β-sheets indicate the number of amino acids in the connecting segments of polypeptide chain.

ognition that proteolytic cleavage sites are generally located in exposed and disordered segments of polypeptide came the suggestion that the E2 chains embody flexible regions that link the lipoyl domains to the rest of the enzyme complex (Bleile et al., 1979; Hale & Perham, 1979). It was envisaged further that this might contribute to the mechanism of active-site coupling. The best and most direct evidence in favor of this has come from NMR spectroscopy.

For a protein complex with an M_r of (5–10) × 10^6, the calculated line width of a methylene proton resonance is about 8 kHz. However, in the ^{1}H NMR spectra of all 2-oxo acid dehydrogenase complexes studied thus far, there exists a family of surprisingly sharp resonances (line width about 30–50 Hz), implying that some regions of the polypeptide chains enjoy substantial conformational flexibility with respect to the bulk of the structure. A combination of limited proteolysis, electron microscopy, and NMR spectroscopy pointed to these regions being part of the E2 chains and likely to be the interdomain segments of polypeptide chain in particular (Perham et al., 1981; Perham & Roberts, 1981; Duckworth et al., 1982; Packman et al., 1984a,b; Spencer et al., 1984). Unequivocal identification has been provided by the alliance of NMR spectroscopy with protein engineering. Up to two lipoyl domains have been deleted from the E2 chain of the *E. coli* PDH complex (Figure 4) and a nested set of deletions has been constructed (Miles et al., 1988) in the long 32-residue (alanine + proline)-rich region that joins the residual lipoyl domain to the E3-binding domain (Figure 6). In the 400-MHz ^{1}H NMR spectra of the pGS110- and pGS156-encoded PDH complexes, there were appropriate falls in the intensities of

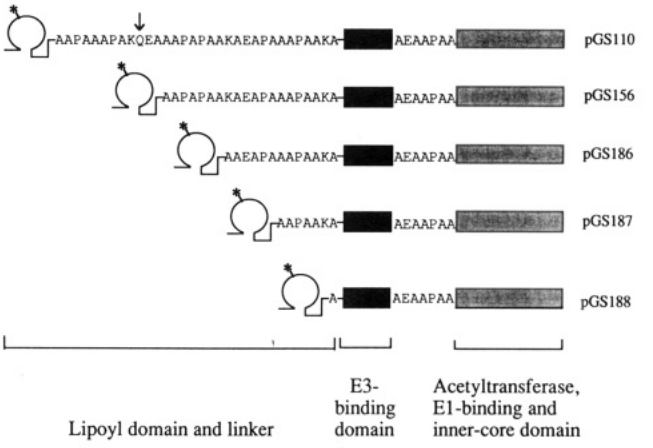


FIGURE 6: Nested deletions in an interdomain linker. The starting protein is the *E. coli* E2 chain with only one lipoyl domain (see Figure 4). *, lipoyllysine residue; ↓, site of the Q291H mutation.

the sharp resonances (Radford et al., 1987).

More direct still, a histidine residue was introduced in place of Gln-291 [original numbering (Stephens et al., 1983)] in the E2p chain of the pGS110-encoded complex (Figure 6) and two sharp signals were observed in the aromatic region of the ^{1}H NMR spectrum of the mutated complex (Texter et al., 1988). The interdomain segments of the native *E. coli* E2p chain lack aromatic amino acid residues and hence contribute no sharp resonances to that part of the ^{1}H NMR spectrum. The chemical shifts of the two new signals are characteristic of the C2 and C4 protons of a histidine residue and were lost when

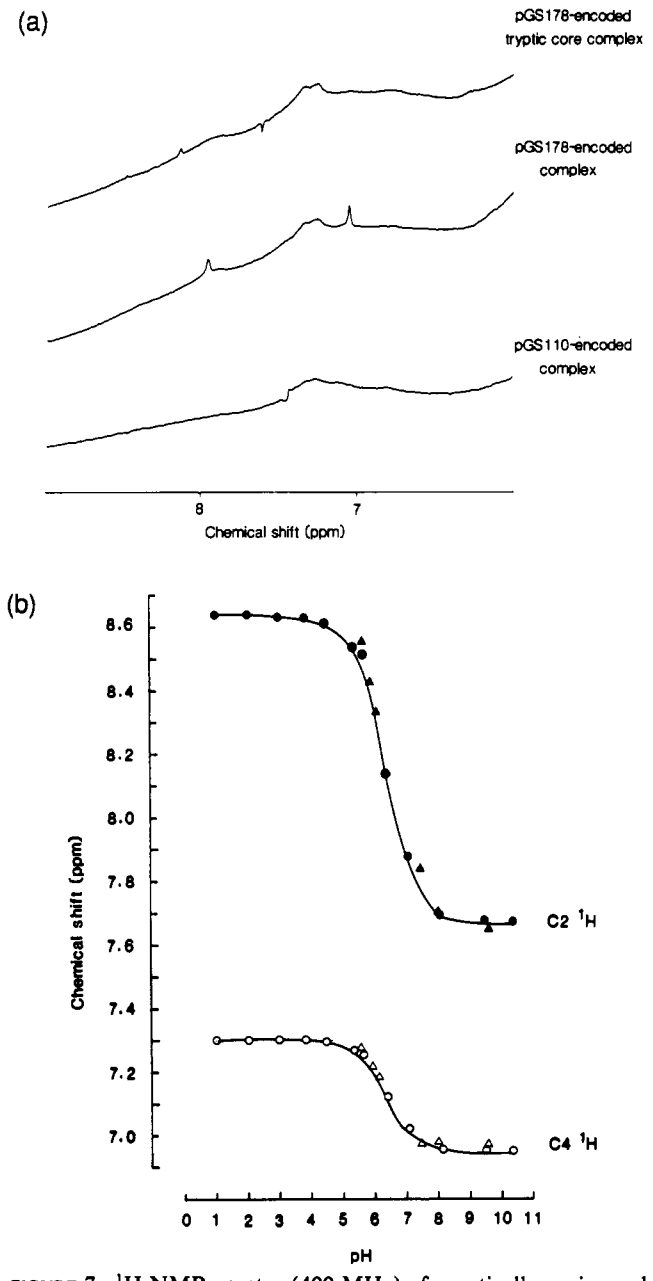
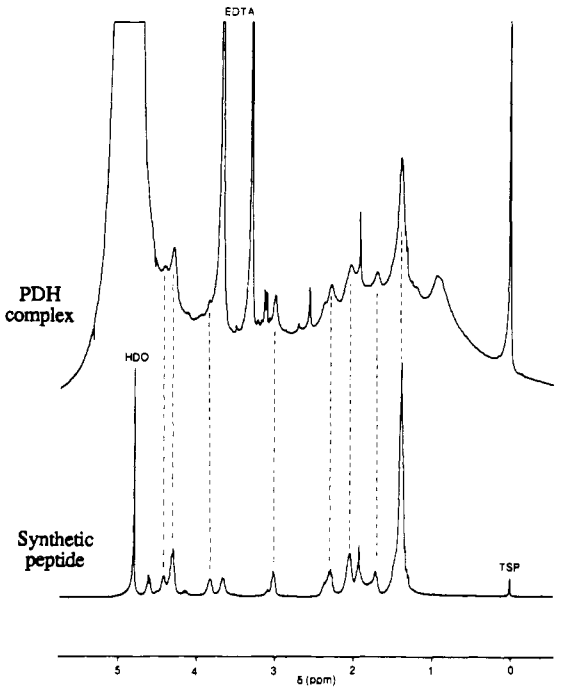


FIGURE 7: ^{1}H NMR spectra (400 MHz) of genetically engineered *E. coli* PDH complexes. (a) Aromatic region. pGS110-encoded, one lipoyl domain per E2 chain (see Figure 6); pGS178-encoded, as pGS110 but carrying the Q291H mutation (see Figure 6). Tryptic cleavage occurs at Lys-316 (Figure 2); E1p and E3 remain bound. (b) Titration curves of His-291 in the E2p chain of the pGS178-encoded complex and of a histidine residue in a 32-residue synthetic peptide whose sequence was identical with that of the interdomain segment. ▲ and △, C2 and C4 protons, respectively, of His-291 in the E2p chain of pGS178-encoded complex; ● and ○, C2 and C4 protons, respectively, of the histidine residue in the synthetic peptide.

the enzyme complex was subjected to limited proteolysis to remove the lipoyl domain and linker (Figure 7). Moreover, the pK_a (6.4) of the introduced histidine residue (Texter et al., 1988) is identical with that of a histidine residue at the same position in a 32-residue synthetic peptide designed to mimic the interdomain segment (Radford et al., 1989a), as judged by the pH dependence of the C2 and C4 proton resonances in their respective NMR spectra. It is a tribute to the sensitivity and resolving power of modern NMR spectrometers, and unequivocal proof of the conformational flexibility of the interdomain segment of polypeptide chain, that a single histidine residue can be detected in a protomer of 200 kDa in a complex of M_r approximately 4×10^6. It opens up new approaches to the study of conformational flexibility in proteins.

Limited proteolysis of the 2OGDH complex of *E. coli* with trypsin cleaves the E2o chain at Arg-100 (Packman & Perham, 1987), but loss of the lipoyl domain together with an appreciable part of the linker region that attaches it to the E3-binding domain (Figure 2) is not accompanied by a corresponding fall in the sharp resonances in the ^{1}H NMR spectrum (Perham & Roberts, 1981). These resonances are dominated by signals whose chemical shift (1.39 ppm) suggests that they derive from the methyl protons of alanine (or threonine) residues, as with the PDH complex. Since the interdomain segment that links the E3-binding domain to the inner-core (acyltransferase) domain in the E2o chain is rich in alanine and proline residues (Figure 2), it is conceivable that this segment of polypeptide chain enjoys substantial conformational flexibility in the 2OGDH complex. On the other hand, insertion of a histidine residue into the corresponding region of the E2p chain of the PDH complex, as described above for the lipoyl domain linker, has failed to disclose any accompanying sharp signals in the aromatic region of the ^{1}H NMR spectrum (N. Allison, J. D. F. Green, J. R. Guest, and R. N. Perham, unpublished work). However, this segment of the E2p chain is less conspicuous in its content of alanine and proline residues (Figure 2) and is not easily cleaved with proteinases (Packman & Perham, 1987; Radford et al., 1987). Thus it may be less exposed and more constrained than its counterpart in the *E. coli* E2o chain. Likewise, the results of limited proteolysis (Packman et al., 1988) and NMR spectroscopy (Duckworth et al., 1982; Packman et al., 1984b) are consistent with, but do not prove, conformational flexibility in the region of polypeptide chain that links the peripheral subunit-binding domain to the inner-core domain in the E2 chain of the PDH complex of *B. stearothermophilus*. Further work needs to be done here; there may be significant differences between different 2-oxo acid dehydrogenase complexes in this regard.

More information about conformational flexibility in the interdomain segments of polypeptide chain has come from a study of synthetic peptides (Radford et al., 1989a). The 400-MHz ^{1}H NMR spectra of the *E. coli* PDH complex and of a 32-residue synthetic peptide with an amino acid sequence identical with that of the segment of E2p chain that links the inner lipoyl domain to the E3-binding domain (Figures 2 and 4) are remarkably similar (Figure 8). Given the huge disparity in the size (M_rs of ca 4.5×10^6 and 2800, respectively), this not only helps confirm the identity of the conformationally flexible regions in the PDH complex but also emphasizes the considerable freedom to move that they must enjoy with respect to the bulk structure of the complex. The CD spectrum of the peptide indicates that it is essentially unfolded in aqueous solution but a more detailed analysis of the ^{1}H and ^{13}C NMR spectra reveals that significant structural constraints are operating; in particular, all six Ala–Pro bonds in the peptide exist (>95%) in the trans form, compared with the 80% or so expected of a random coil. The trans configuration is retained in the interdomain segment in the assembled PDH complex (Radford et al., 1989a). This limits the flexibility around the Ala–Pro bond, stiffening and elongating the structure of the peptide. Similar results have been obtained for synthetic peptides representing the two other interlipoyl domain segments in the *E. coli* PDH complex (J. D. F. Green, E. D. Laue, R. N. Perham, S. R. Martin, and E. Appella, unpublished work). These structural constraints may restrict the movement of the lipoyl domains and it is not inconceivable that they are



Peptide sequence: AAPAAPAKQEAAAPAPAAKAEPAAAPAAKA

FIGURE 8: ^{1}H NMR spectra (400 MHz) of wild-type *E. coli* PDH complex and a 32-residue synthetic peptide. The amino acid sequence of the peptide was identical with that linking the innermost lipoyl domain to the E3-binding domain in the E2p chain (see Figures 2 and 6). TSP, 3-(trimethylsilyl)-2,2,3,3-tetradeuteriopropionic acid.

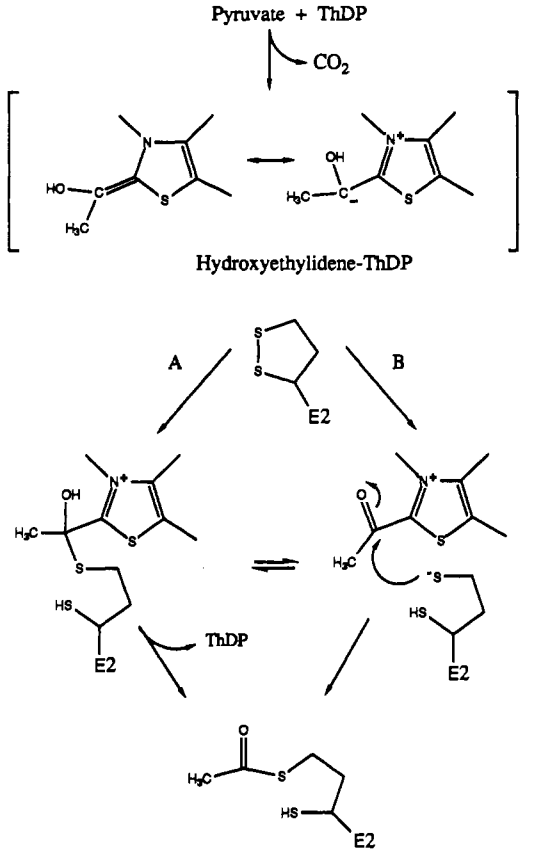
the result of evolutionary pressures to optimize interactions between the three active sites in the enzyme complex.

The interdomain segments from other 2-oxo acid dehydrogenase multienzyme complexes all share a family resemblance, but they exhibit a wide spectrum of sequences, ranging from those dominated by alanine and proline residues to others more obviously hydrophilic and charged, as detailed by Perham and Packman (1989). Some, like those of *B. stearothermophilus* E2p, are combinations of the two types (Packman et al., 1988). Aromatic residues are conspicuously rare. Amino acid sequences of this kind are increasingly being found in a wide range of proteins, where they appear to function as interdomain segments or pendant extensions to more tightly folded domains [Radford et al. (1989a) and references therein]. Conformational calculations on the (Ala-Pro) repeat at the N-terminus of the rabbit myosin light chains are consistent with it adopting a semiextended structure as part of a flexible antenna region (Abillon et al., 1990). In all cases such segments can be envisaged as conformationally flexible structures that have the potentially beneficial property of keeping domains apart as well as allowing them to move in pursuit of essential interactions with other proteins, membranes, or nucleic acids (Radford et al., 1989a).

Active Sites of the E1, E2, and E3 Components

It is not the purpose of this review to discuss the detailed chemical mechanism of each component enzyme. Indeed, such discussion is hampered still by the lack of three-dimensional structures for the E1 and E2 components, but information is beginning to emerge about domains and motifs that house the active sites and this warrants inclusion. However, the E3 component is a member of the growing band of flavoprotein disulfide oxidoreductases, about which much is known (Williams, 1976; Carothers et al., 1989; Roche & Patel, 1989;

Scheme I[a]



[a] After Gruys et al. (1989).

Curti et al., 1991), including X-ray crystallographic structures for the yeast (Takenaka et al., 1988) and *A. vinelandii* (Schierbeek et al., 1989) dihydrolipoamide dehydrogenases. All these enzymes are homodimers (subunit M_r approximately 50 000), with a mechanism that involves the concerted participation of the enzyme-bound FAD and an intrachain disulfide bridge that undergoes alternate reduction and oxidation.

As outlined above, a single E3 gene (*lpd*) is generally able to fulfill the need for an E3 component in different 2-oxo acid dehydrogenase complexes, but in *Ps. putida* three *lpd* genes have been identified (Burns et al., 1989). A second dihydrolipoamide dehydrogenase has also been reported for *E. coli*, where it is suggested that it may be involved with lipoic acid in the binding protein dependent transport of ribose, galactose, and maltose (Richarme, 1989). To add to this, dihydrolipoamide dehydrogenase has been detected in halophilic and thermophilic archaebacteria and in the bloodstream form of the eukaryote *Trypanosoma brucei*, none of which possesses a 2-oxo acid dehydrogenase complex (Danson, 1988). In *T. brucei* at least, the enzyme is associated with the plasma membrane (Jackman et al., 1990). This promises to be an interesting line of inquiry for the future.

Little is known about the active site of the E1 component, though much is known in general about the chemistry of ThDP-dependent reactions (Sable & Gubler, 1982). 2-Acetyl-ThDP has been detected in the active site of the E1 component of the *E. coli* PDH complex, but it is not known whether reductive acetylation of the lipoyl group compulsorily follows pathway B in Scheme I (Gruys et al., 1989). Apart from exercising stereospecificity for the *R*-enantiomer of lipoic acid (Yang & Frey, 1989), the *E. coli* E1p is thought to produce 8-*S*-acetyldihydrolipoamide as its initial product; this can be followed by intramolecular acetyl transfer to produce an equilibrium mixture of S^8- and S^6-isomers, though this

```
Zmpdc    APERRNILMVGDGS.FQLTAQEVAQMVR.LKL.PVIIFLINNY.GYTIEVM
Scpdc1   DPKKRVILFIGDGSL.QLTVQEISTMIRWGLK.PY.LFVLNND.GYTIEKL
Ntsura   RPDEVVVDIDGDGS.FIMNVQELATIKVENL..PVKIMLLNNQ.HLGMVVQ
Scilv2   KPESLVIDIDGDAS.FNMTLTELSS.AVQAGT.PVKILILNNEE.QGMVTQ
Ecilvi   LPEETVVCVTGDGSI.QMNIQELST.ALQYEL.PVLVVNLNNR.YLGMVKQ
Ecpoxb   EPERQVVAMCGDGG.FSMLMGDFLS.VVQMKL.PVKIVVFNNS.VLGFVAM
HuE1αp   GKDEVCLTLYGDGAANQGQIFEAYNMAALWKL.PCIFICENNY.GMGTSVE
BsE1α    GKKAVAITYTGDGGTSQGDFYEGINFAGAFKA.PAIFVVQNNR..FAISTP
EcaceE   TSKQTVYAFLGDGEMDEPESKGAITIATREKL.DNLVFVINCNLQR.LDGP
Sckgd1   LLHGDA.AFAGQGVV.YETM.GFLTL.PEYSTGGTIHVITNNQIGFT.TDP
EcsucA   TIHGDA.AVTGQGVV.QETL.NMS.KARGYEVGGTVRIVINNQVGFTTSNP
HuE1αb   NANRVVICYFGEGAASEGDAHDGFNFAATLEC.PIIFFCRNN..GYAISTP
OxE1αb   NANRVVICYFGEGAASEGDAHAGFNFAATLEC.PIIFFCRNN..GYAISTP
RaE1αb   NANQIVICYFGEGAASEGDAHAGFNFAATLEC.PIIFFCRNN..GYAISTP
PpbkdA1  GDTKIASAWIGDGATAESDFHTALTFAHVYRA.PVILNVVNNQ..WAISTF
Ysmdas   IITNKVYCMVGDACLQEGPALESISLAGHMGLDNLIVLYDNNQVCCDGSVD

Rcrcfp   QPVGDTIAIIGDGSITAGMAYEALNHAGHLK..SRMFVILNDND.MSIAPP
          ----β----t----------α----------t--β---t----α----
```

FIGURE 9: Putative ThDP-binding motif in ThDP-utilizing enzymes. The enzymes from which the sequences are taken are listed in Hawkins et al. (1989). Rcrcfp, an unidentified open reading frame adjacent to the genes encoding the photosynthetic reaction center proteins in *Rhodobacter capsulata*. The secondary structure prediction is shown beneath the sequences.

cannot be part of the normal catalytic process since the isomerization rate constant (0.045 s^{-1} at pH 6.5) is so much smaller than the turnover number (1.6×10^2 s^{-1}) of the enzyme (Yang & Frey, 1986).

An important clue to the location of the active site of E1 has emerged from a comparison of the primary structures of ThDP-binding enzymes (Hawkins et al., 1989). A surprising feature of the amino acid sequences of the E1p and E1o chains of *E. coli* was their lack of similarity to each other or to the E1α or E1β chains of the BCODH or eukaryotic PDH complexes (Guest et al., 1989; Patel & Roche, 1990). However, there is a conserved sequence motif of about 30 amino acid residues, beginning with the sequence -GDG- and ending with the sequence -NN- (Figure 9), in all ThDP-utilizing enzymes irrespective of their lack of sequence homology otherwise, which secondary structure predictions suggest will adopt a $\beta\alpha\beta$ fold (Hawkins et al., 1989). It is tempting to suppose that this may be a common motif related to binding ThDP. In keeping with this, the motif occurs in the E1α subunits of 2-oxo acid dehydrogenase complexes, the subunit associated on other grounds (Stepp & Reed, 1985) with the decarboxylation step, whereas the E1β subunit may be more closely implicated in the subsequent reductive acylation (Scheme I) of the incoming dithiolane ring of the lipoyl group. Further alignments of E1α and E1β primary structures have indicated additional conserved sequences, one in E1α that may be involved in interaction with E1β, one in E1β that may be involved in the binding to E2, and others with no obvious function as yet (Wexler et al., 1991).

Another common sequence motif, -HXXXDG-, this time in the E2 chains, has led to speculation that it may form part of the acyltransferase active site. It is known to be critically involved in the active site of chloramphenicol acetyltransferase, an enzyme to which the acyltransferase domains of E2 chains appear to show remote but significant similarity, suggesting that the histidine side chain acts to promote the nucleophilic attack of coenzyme A on 8-*S*-acyldihydrolipoamide (Guest, 1987). Such speculation is supported by the loss of acyltransferase activity that accompanies the replacement of this histidine residue in *E. coli* E2p (Russell & Guest, 1990) or the E2 of the ox BCODH complex (Griffin & Chuang, 1990), although conflicting results have been reported for the corresponding mutagenesis of yeast E2p (Niu et al., 1990). However, the case for a resemblance between the mechanisms of E2 and chloramphenicol acetyltransferase has been strengthened by the observation that mutagenesis of a serine residue corresponding to one that stabilizes a tetrahedral transition state in the latter enzyme leads to loss of activity by *E. coli* E2p (Russell & Guest, 1991b). The -HXXXDG- motif reposes near the C-terminus of all E2 chains but is lacking from the putative acetyltransferase domain (Figure 3) of the yeast protein X (Behal et al., 1989). This would suggest that protein X, at least in yeast, is unable to catalyze the acetyltransferase reaction.

Mechanism of Active-Site Coupling

The lipoyllysine swinging arm (Reed, 1974) has stood the test of time as a basic component of the mechanism of active-site coupling in the 2-oxo acid dehydrogenase multienzyme complexes. Similar mechanisms have been invoked for other multienzyme complexes that utilize biotin or pantethenic acid as prosthetic group (Hammes, 1981). But it is clear that the 2-oxo acid dehydrogenase complexes have requirements over and above the simple swinging arm. For example, the distance between the active sites in the *E. coli* PDH complex has been estimated from fluorescence energy transfer experiments to be some 4.5 nm (Shepherd & Hammes, 1977; Hammes, 1981), substantially greater than the span ($\leq$3 nm) of a single lipoyllysine side chain. Moreover, as described above, the lipoyl group must be attached to a specific protein domain if it is to function as a substrate.

When the PDH complex of *E. coli* (Stepp et al., 1981) or mammals (Rahmatullah et al., 1990) is subjected to limited proteolysis, about half of the lipoyl domains can be removed without significant loss of overall catalytic activity. Similarly, the lipoyl groups of the *E. coli* complex can be chemically modified (Berman et al., 1981) or enzymically removed (Stepp et al., 1981; Berman et al., 1981) more rapidly than complex activity is lost. This is borne out by the protein engineering experiments that demonstrated deletion or inactivation of lipoyl domains without loss of complex activity (Figure 4). Thus, what would conventionally be regarded as catalytically essential intermediates can be excised or modified without effect. This paradox has been resolved by noting that E1 catalyzes the rate-determining step in the mechanism (Danson et al., 1978; Cate et al., 1980) and then postulating that any function of a lipoyl group that has been modified or excised can be taken over by one that remains. This can be achieved physically by several different lipoyl groups, i.e., domains, interacting with a given E1 active site (Perham et al., 1981; Berman et al., 1981; Stepp et al., 1981), to which can be added a system of rapid intramolecular transacylation reactions that permits the acyl group to be transferred between neighboring lipoyl domains (Bates et al., 1977; Collins & Reed, 1977). In this way too, reducing equivalents can be funneled into the E3 component, which (see above) can be present in relatively few copies per enzyme complex particle. A computer simulation of the kinetics supports this unusual mechanism, termed multiple random coupling (Hackert et al., 1983).

Such a mechanism clearly rests on the interdigitation of the lipoyl domains between the peripheral E1 and E3 subunits and the conformational flexibility in the interdomain segments of the E2 polypeptide chains that permits these domains to link the different and widely separated active sites. This has been emphasized by the demonstration that F_{ab} fragments of antibodies raised against the interdomain segment that joins the innermost lipoyl domain to the E3-binding domain in the *E. coli* E2p chain inhibit the overall PDH complex activity (and the formation of the acetyl-enzyme intermediate) without effect on the individual part-reactions catalyzed by the E1, E2, or E3 components (Radford et al., 1989b). This is most simply explained by the F_{ab} fragments acting, like wedges, to prevent the movement of the lipoyl domains necessary for active-site coupling.

The presence of protein X in the icosahedral eukaryotic PDH complexes may also be important in this regard. Limited proteolysis of protein X in mammalian PDH complexes leads to a lower affinity of E3 for the core assembly (Gopalakrishnan et al., 1989; Neagle et al., 1991), but removal of only the lipoyl domain of protein X (Figure 3) can be achieved without loss of catalytic activity (Neagle et al., 1991). Mutagenesis of the yeast gene for protein X has indicated that it is not required for the formation of the E2 core, that it is essential for the binding of E3, and that the lipoyl domain can be deleted without much effect, but that deletion of the putative peripheral subunit-binding domain leads to lowered affinity for E3 and concomitant loss of overall PDH complex activity (Lawson et al., 1991). Given the few copies (about 6) of protein X and of the E3 dimer in these complexes, it is likely that protein X plays some important part in binding and positioning E3 at specific sites on the E2 core (Lawson et al., 1991). Its lipoyl domain, though capable of becoming reductively acetylated (Jilka et al., 1986; Hodgson et al., 1986), is apparently superfluous, as are many of the lipoyl domains in the E2 chains themselves.

Conclusion

The 2-oxo acid dehydrogenase complexes can be seen to provide an elegant solution to the problem of catalyzing an integrated set of biochemical reactions. The enhancement of catalytic activity that accompanies attachment of the lipoyl group to a protein domain and the mechanism of substrate channeling conferred by the need to recognize that domain are important manifestations of properties long inferred for multienzyme complexes. Whether in addition there is any component of the hot potato hypothesis involved remains to be determined. The system of active-site coupling is uniquely a property of a molecular assembly; the larger the number of subunits in the E2 core, as with cubic point group symmetries, the more opportunities there are for such effects (Danson et al., 1978). These would be denied an identical number of E2 polypeptide chains used to form separate E1E2E3 trimers. Moreover, the complex is relieved of the necessity to assemble with a strict stoichiometry of active sites in perfect geometrical arrangement with respect to one another.

The key to all this obviously lies in the domain-and-linker design of the E2 polypeptide chain, which enables it both to assemble as the structural and catalytic core of the complex and to have its lipoyl domains interdigitating flexibly between the peripheral E1 and E3 subunits. Why there is a superfluity of lipoyl domains remains a puzzle, and there is much still to be learned about the structures of the individual domains and of the peripheral subunits before a full understanding of the complex and its mechanism of active-site coupling can be claimed. But the 2-oxo acid dehydrogenase complexes continue to offer lessons about the design of multidomain, multisubunit systems with important applications in wider studies of biomolecular assembly.

Registry No. 2-Oxo acid dehydrogenase, 9067-96-3.

References

Abillon, E., Bremier, L., & Cardinaud, R. (1990) *Biochim. Biophys. Acta 1037*, 394–400.

Ali, S. T., Moir, A. J. G., Ashton, P. R., Engel, P. C., & Guest, J. R. (1990) *Mol. Microbiol. 4*, 943–950.

Allen, A. G., & Perham, R. N. (1991) *FEBS Lett.* (in press).

Allen, A. G., Perham, R. N., Allison, N., Miles, J. S., & Guest, J. R. (1989) *J. Mol. Biol. 208*, 623–633.

Ambrose, M. C., & Perham, R. N. (1976) *Biochem. J. 155*, 429–432.

Bates, D. L., Harrison, R. A., & Perham, R. N. (1975) *FEBS Lett. 60*, 427–430.

Bates, D. L., Danson, M. J., Hale, G., Hopper, E. A., & Perham, R. N. (1977) *Nature (London) 268*, 313–316.

Behal, R. H., Browning, K. S., Hall, T. B., & Reed, L. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8732–8736.

Berman, J. N., Chen, G.-X., Hale, G., & Perham, R. N. (1981) *Biochem. J. 199*, 513–520.

Bleile, D. M., Munk, P., Oliver, R. M., & Reed, L. J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4385–4389.

Bleile, D. M., Hackert, M. L., Pettit, F. H., & Reed, L. J. (1981) *J. Biol. Chem. 256*, 514–519.

Borges, A., Hawkins, C. F., Packman, L. C., & Perham, R. N. (1990) *Eur. J. Biochem. 194*, 95–102.

Burns, G., Brown, G., Hatter, K., & Sokatch, J. R. (1988) *Eur. J. Biochem. 176*, 165–169.

Burns, G., Sykes, P. J., Hatter, K., & Sokatch, J. R. (1989) *J. Bacteriol. 171*, 665–668.

Carlsson, P., & Hederstedt, L. (1989) *J. Bacteriol. 171*, 2110–2115.

Carothers, D. J., Pons, G., & Patel, M. S. (1989) *Arch. Biochem. Biophys. 268*, 409–425.

Cate, R. L., Roche, T. E., & Davis, L. C. (1980) *J. Biol. Chem. 225*, 7556–7562.

Collins, J. H., & Reed, L. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4223–4227.

Coppel, R. L., McNeilage, L. J., Surh, C. D., Van de Water, J., Spithill, T. W., Whittingham, S., & Gershwin, M. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7317–7321.

Curti, B., Ronchi, S., & Zanetti, G., Eds. (1991) *Flavins and Flavoproteins 1990*, Walter de Gruyter, Berlin.

Danner, D. J., Litwer, S., Herring, W. J., & Pruckler, J. (1989) *J. Biol. Chem. 264*, 7742–7746.

Danson, M. J. (1988) *Biochem. Soc. Trans. 16*, 87–89.

Danson, M. J., Fersht, A. R., & Perham, R. N. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5386–5390.

Danson, M. J., Hale, G., Johnson, P., Perham, R. N., Smith, J., & Spragg, P. (1979) *J. Mol. Biol. 129*, 603–617.

Dardel, F., Packman, L. C., & Perham, R. N. (1990) *FEBS Lett. 264*, 206–210.

Dardel, F., Laue, E. D., & Perham, R. N. (1991) *Eur. J. Biochem.* (in press).

de Kok, A., & Westphal, A. H. (1985) *Eur. J. Biochem. 152*, 35–41.

DeRosier, D. J., Oliver, R. M., & Reed, L. J. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1135–1137.

Duckworth, H. W., Jaenicke, R., Perham, R. N., Wilkie, A. O. M., Finch, J., & Roberts, G. C. K. (1982) *Eur. J. Biochem. 124*, 63–69.

Fuller, C. C., Reed, L. J., Oliver, R. M., & Hackert, M. L. (1979) *Biochem. Biophys. Res. Commun. 90*, 431–438.

Gershwin, M. E., Mackay, I. R., Sturgess, A., & Coppel, R. L. (1987) *J. Immunol. 138*, 3525–3531.

Gilbert, G. A., & Gilbert, L. M. (1980) *J. Mol. Biol. 144*, 405–408.

Gopalakrishnan, S., Rahmatullah, M., Radke, G. A., Powers-Greenwood, S., & Roche, T. E. (1989) *Biochem. Biophys. Res. Commun. 160*, 715–721.

Graham, L. D., & Perham, R. N. (1990) *FEBS Lett. 262*, 241–244.

Graham, L. D., Packman, L. C., & Perham, R. N. (1989) *Biochemistry 28*, 1574–1581.

Grande, H. J., Van Telgen, H. J., & Veeger, C. (1976) *Eur. J. Biochem. 71*, 509–518.
Griffin, T. A., & Chuang, D. T. (1990) *J. Biol. Chem. 265*, 13174–13180.
Griffin, T. A., Lau, K. S., & Chuang, D. T. (1988) *J. Biol. Chem. 263*, 14008–14014.
Griffin, T. A., Wynn, R. M., & Chuang, D. T. (1990) *J. Biol. Chem. 265*, 12104–12110.
Gruys, K. J., Datta, A., & Frey, P. A. (1989) *Biochemistry 28*, 9071–9080.
Guest, J. R. (1987) *FEMS Microbiol. Lett. 44*, 417–422.
Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C., & Perham, R. N. (1985) *J. Mol. Biol. 185*, 743–754.
Guest, J. R., Angier, S. J., & Russell, G. C. (1989) *Ann. N.Y. Acad. Sci. 573*, 76–99.
Hackert, M. L., Oliver, R. M., & Reed, L. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2907–2911.
Hackert, M. L., Zu, W.-X., Oliver, R. M., Wall, J. S., Hainfeld, J. F., Mullinax, T. R., & Reed, L. J. (1989) *Biochemistry 28*, 6816–6821.
Hale, G., & Perham, R. N. (1979) *FEBS Lett. 105*, 263–266.
Hammes, G. G. (1981) *Biochem. Soc. Symp. 46*, 73–90.
Hanemaaijer, R., Janssen, A., de Kok, A., & Veeger, C. (1988) *Eur. J. Biochem. 174*, 593–599.
Hawkins, C. F., Borges, A., & Perham, R. N. (1989) *FEBS Lett. 255*, 77–82.
Hemilä, H., Palva, A., Paulin, L., Arvidson, S., & Palva, I. (1990) *J. Bacteriol. 172*, 5052–5063.
Henderson, C. E., & Perham, R. N. (1980) *Biochem. J. 189*, 161–172.
Henderson, C. E., Perham, R. N., & Finch, J. T. (1979) *Cell 17*, 85–93.
Hodgson, J. A., De Marcucci, O., & Lindsay, J. G. (1986) *Eur. J. Biochem. 158*, 595–600.
Jackman, S. A., Hough, D. W., Danson, M. J., Stevenson, K. J., & Opperdoes, F. R. (1990) *Eur. J. Biochem. 193*, 91–95.
Jilka, J. M., Rahmatullah, M., Kazemi, M., & Roche, T. E. (1986) *J. Biol. Chem. 263*, 1858–1867.
Keha, E. E., Ronft, H., & Kresze, G.-B. (1982) *FEBS Lett. 145*, 289–292.
Lawson, J. E., Behal, R. H., & Reed, L. J. (1991) *Biochemistry 30*, 2834–2839.
Lowe, P. N., Hodgson, J. A., & Perham, R. N. (1983) *Biochem. J. 215*, 133–140.
Miles, J. S., Guest, J. R., Radford, S. E., & Perham, R. N. (1988) *J. Mol. Biol. 202*, 97–106.
Neagle, J. C., De Marcucci, O., Dunbar, B., & Lindsay, J. G. (1989) *FEBS Lett. 253*, 11–15.
Neagle, J. C, Stansbie, D., & Lindsay, J. G. (1991) in *Biochemistry and Physiology of Thiamin Diphosphate Enzymes* (Bisswanger, H., & Ullrich, J., Eds.) pp 157–163, VCH Verlagsgesellschaft, Weinheim.
Niu, X.-D., Browning, K. S., Behal, R. H., & Reed, L. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7546–7550.
Niu, X.-D., Stoops, J. K., & Reed, L. J. (1990) *Biochemistry 29*, 8614–8619.
Nobukuni, Y., Mitsubuchi, H., Endo, F., & Matsuda, I. (1989) *Biochem. Biophys. Res. Commun. 161*, 1035–1041.
Oliver, R. M., & Reed, L. J. (1982) In *Electron Microscopy of Proteins* (Harris, R., Ed.) Vol. 1, pp 1–48, Academic Press, London.
Packman, L. C., & Perham, R. N. (1986) *FEBS Lett. 206*, 193–198.
Packman, L. C., & Perham, R. N. (1987) *Biochem. J. 242*, 531–538.
Packman, L. C., Hale, G., & Perham, R. N. (1984a) *EMBO J. 3*, 1315–1319.
Packman, L. C., Perham, R. N., & Roberts, G. C. K. (1984b) *Biochem. J. 217*, 219–227.
Packman, L. C., Borges, A., & Perham, R. N. (1988) *Biochem. J. 252*, 79–86.
Packman, L. C., Green, B., & Perham, R. N. (1991) *Biochem. J. 277*, 153–158.
Patel, M. S., & Roche, T. E. (1990) *FASEB J. 4*, 3224–3233.
Perham, R. N. (1975) *Philos. Trans. R. Soc. London, Ser. B 272*, 123–136.
Perham, R. N., & Wilkie, A. O. M. (1980) *Biochem. Int. 1*, 470–477.
Perham, R. N., & Roberts, G. C. K. (1981) *Biochem. J. 199*, 733–740.
Perham, R. N., & Packman, L. C. (1989) *Ann. N.Y. Acad. Sci. 573*, 1–20.
Perham, R. N., Harrison, R. A., & Brown, J. P. (1978) *Biochem. Soc. Trans. 6*, 47–50.
Perham, R. N., Duckworth, H. W., & Roberts, G. C. K. (1981) *Nature (London) 292*, 474–477.
Radford, S. E., Laue, E. D., Perham, R. N., Miles, J. S., & Guest, J. R. (1987) *Biochem. J. 247*, 641–649.
Radford, S. E., Laue, E. D., Perham, R. N., Martin, S. R., & Appella, E. (1989a) *J. Biol. Chem. 264*, 767–775.
Radford, S. E., Perham, R. N., Ullrich, S. J., & Appella, E. (1989b) *FEBS Lett. 250*, 336–340.
Rahmatullah, M., Gopalakrishnan, S., Radke, G. A., & Roche, T. E. (1989) *J. Biol. Chem. 264*, 1245–1251.
Rahmatullah, M., Radke, G. A., Powers-Greenwood, S. L., Andrews, P. C., & Roche, T. E. (1990) *J. Biol. Chem. 265*, 14512–14517.
Randle, P. J., Patson, P. A., & Espinal, J. (1987) in *The Enzymes* (Boyer, P. D., & Krebs, E. G., Eds.) Vol. 18, pp 97–121, Academic Press, Orlando, FL.
Reed, L. J. (1966) in *Comprehensive Biochemistry* (Florkin, M., & Stotz, E. H., Eds.) Vol. 14, pp 99–125, Elsevier, London.
Reed, L. J. (1974) *Acc. Chem. Res. 7*, 40–46.
Reed, L. J., & Yeaman, S. J. (1987) in *The Enzymes* (Boyer, P. D., & Krebs, E. G., Eds.) Vol. 18, pp 77–96, Academic Press, Orlando, FL.
Reed, L. J., & Hackert, M. L. (1990) *J. Biol. Chem. 265*, 8971–8974.
Reed, L. J., Koike, M., Levitch, M. E., & Leach, F. R. (1958) *J. Biol. Chem. 232*, 143–158.
Reed, L. J., Pettit, F. H., Eley, M. H., Hamilton, L., Collins, J. H., & Oliver, R. M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3068–3072.
Richarme, G. (1989) *J. Bacteriol. 171*, 6580–6585.
Roche, T. E. & Patel, M. S., Eds. (1989) Alpha-Keto Acid Dehydrogenase Complexes: Organization, Regulation, and Biomedical Ramifications. *Ann. N.Y. Acad. Sci. 573*, 1–462.
Russell, G. C, & Guest, J. R. (1990) *Biochem. J. 269*, 443–450.
Russell, G. C. & Guest, J. R. (1991a) *Biochim. Biophys. Acta 1076*, 225–232.
Russell, G. C., & Guest, J. R. (1991b) *Proc. R. Soc. London, B 243*, 155–160.
Sable, H. Z., & Gubler, C. J., Eds. (1982) Thiamin: Twenty Years of Progress. *Ann. N.Y. Acad. Sci. 378*, 1–470.

Schierbeek, A. J., Swarte, M. B. A., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G. J., & Drenth, J. (1989) *J. Mol. Biol. 206*, 365-380.
Schmitt, B., & Cohen, R. (1980) *Biochem. Biophys. Res. Commun. 93*, 709-712.
Shepherd, G., & Hammes, G. G. (1977) *Biochemistry 16*, 5234-5241.
Spencer, M. E., Darlison, M. G., Lewis, H. M., & Guest, J. R. (1984) *Eur. J. Biochem. 141*, 361-374.
Stephens, P. E., Darlison, M. G., Lewis, H. M., & Guest, J. R. (1983) *Eur. J. Biochem. 133*, 481-489.
Stepp, L. R., & Reed, L. J. (1985) *Biochemistry 24*, 7187-7191.
Stepp, L. R., Bleile, D. M., McRorie, D. K., Pettit, F. H., & Reed, L. J. (1981) *Biochemistry 20*, 4555-4560.
Takenaka, A., Kizawa, K., Hata, T., Sato, S., Misaka, E.-J., Tamura, C., & Sasada, Y. (1988) *J. Biochem. (Tokyo) 103*, 463-469.
Texter, F. L., Radford, S. E., Laue, E. D., Perham, R. N., Miles, J. S., & Guest, J. R. (1988) *Biochemistry 27*, 289-296.
Thekkumkara, T. J., Ho, L., Wexler, I. D., Pons, G., Liu, T.-C., & Patel, M. S. (1988) *FEBS Lett. 240*, 45-48.
Wagenknecht, T., Francis, N., & DeRosier, D. J. (1986) *Biochem. Biophys. Res. Commun. 135*, 802-807.
Wagenknecht, T., Francis, N., DeRosier, D. J., Hainfeld, J. F., & Wall, J. S. (1987) *J. Biol. Chem. 262*, 877-882.
Westphal, A. H., & de Kok, A. (1990) *Eur. J. Biochem. 187*, 235-239.
Wexler, I. D., Hemalatha, S. G., & Patel, M. S. (1991) *FEBS Lett. 282*, 209-213.
Williams, C. H., Jr. (1976) in *The Enzymes* (Boyer, P. D., Ed.) 3rd ed., Vol. 13, pp 89-173, Academic Press, New York.
Wu, T.-L., & Reed, L. J. (1984) *Biochemistry 23*, 221-226.
Yeaman, S. J. (1989) *Biochem. J. 257*, 625-632.
Yang, H., Hainfeld, J. F., Wall, J. S., & Frey, P. A. (1985) *J. Biol. Chem. 260*, 16049-16051.
Yang, Y.-S., & Frey, P. A. (1986) *Biochemistry 25*, 8173-8178.
Yang, Y.-S., & Frey, P. A. (1989) *Arch. Biochem. Biophys. 268*, 465-474.

Accelerated Publications

Exon Organization of the Human FKBP-12 Gene: Correlation with Structural and Functional Protein Domains‡

Anthony G. DiLella* and Robert J. Craig

Department of Molecular Genetics, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

Received June 14, 1991

ABSTRACT: FKBP-12, the major T-cell binding protein for the immunosuppressive agents FK506 and rapamycin, catalyzes the interconversion of the cis and trans rotamers of the peptidyl–prolyl amide bond of peptide and protein substrates. The function of rotamase activity in cells and the role of FKBP-12 in immunoregulation is uncertain. In this paper we report the cloning and characterization of the human chromosomal FKBP-12 gene and four processed FKBP-12 pseudogenes. The FKBP-12 gene is 24 kilobases in length and contains five exons. The protein-coding region of the gene is divided into four exon modules that correlate with the structural and functional domains of the protein. The novel structure of FKBP-12 resulting from the topology of the antiparallel β-sheet is the topological crossing of two loops that are encoded by separate exons. Separate exons also encode the antiparallel β-sheet and α-helical region that define the drug-binding pocket and enzyme activity site of FKBP-12. The exon organization of the FKBP-12 gene also provided insight into the genetic evolution of the immunophilin family. Knowledge of the FKBP-12 gene structure will enable inactivation of this gene by homologous recombination in cells to provide a model to study the role of FKBP-12 in immunoregulation and normal cellular processes.

Early events in T-lymphocyte activation result from the stimulation of the T-cell receptor signal transmission pathway by a specific antigen (Crabtree, 1989). Immunosuppressive agents that specifically inhibit this pathway have been used to prevent graft rejection of organ and bone marrow transplants (Schreiber, 1991). FK506 and CsA are structurally unrelated immunosuppressive drugs that block the T-cell receptor signal transmission pathway in T-cell activation by inhibiting expression of the same set of lymphokine genes, such as IL-2 (Tocci et al., 1989; Mattila et al., 1990; Emmel et al., 1989). Although FK506 and CsA each inhibit T-cell activation by a similar mechanism, they bind to distinct and abundant T-cell cytoplasmic receptors, termed immunophilins (Schreiber, 1991). The CsA-binding protein is cyclophilin (Handschumacher, 1984), while the major FK506-binding protein in T cells is the 12-kDa cytosolic protein FKBP-12 (Harding et al., 1989; Siekierka et al., 1989; Fretz et al., 1991). Rapamycin, an immunosuppressive agent structurally related to FK506 (Schreiber, 1991), also specifically binds to FKBP-12 (Fretz et al., 1991) but inhibits a later stage of T-cell activation by blocking the lymphokine receptor signal trans-

‡The nucleotide sequence in this paper has been submitted to GenBank under Accession Number J05340.

0006-2960/91/0430-8512$02.50/0