J. Mol. Biol. (1991) **220**, 133–149

Analysis of Protein Loop Closure

Two Types of Hinges Produce One Motion in Lactate Dehydrogenase

Mark Gerstein and Cyrus Chothia

MRC Laboratory of Molecular Biology
Hills Road, Cambridge CB2 2QH, U.K.

(Received 23 October 1990; accepted 29 January 1991)

As shown in previous crystallographic investigations, upon binding lactate and NAD, lactate dehydrogenase undergoes a large conformational change that results in a surface loop moving roughly 10 Å to cover the active site. In addition, there are appreciable movements (≈2 Å) of five helices and three other loops.

We demonstrate by a new fitting procedure that the loop moves on two hinges separated by a relatively rigid type II turn. The first hinge has few steric constraints on it, and its motion can be well accounted for by large changes in two torsion angles, i.e. as in a classic hinge motion. In contrast, the second hinge, which is part of a helix connected to the end of the loop, has many more constraints on it and distributes its deformation over more torsion angles. This novel motion involves the helix stretching and splitting into α-helical and 3_{10}-helical components and substantial side-chain repacking in the sense of "cogs hopping between grooves" at its interface with the end of a neighboring helix.

The loop is stabilized by five transverse (across loop) hydrogen bonds. These are preserved, through the conformational change and through 17 lactate dehydrogenase sequences, more than the longitudinal hydrogen bonds down the sides of the loop.

Through a network of contacts, many of them conserved hydrophobic residues, the motion of the loop is propagated outward to structures that have no direct contact with the ligands. These moving structures are on the surface of the protein, and the whole protein can be subdivided into concentric shells of increasing mobility.

Keywords: lactate dehydrogenase; loop closure; conformational change; surface loops; helix movements

1. Introduction

The activity of many proteins requires conformational change. The molecular mechanisms behind some of these changes have been analyzed and found to occur on a variety of different structural scales. At one end of the spectrum are movements of whole subunits, e.g. for the transmission of allosteric effects in hemoglobin and other proteins (Perutz, 1989). Next, there are hinge movements of domains, e.g. in the immunoglobulins (Bennet & Huber, 1984; Lesk & Chothia, 1988), lysozyme (Faber & Matthews, 1990) and alcohol dehydrogenase (Eklund *et al.*, 1981). On a secondary structure scale, there are rigid-body motions of helices that occur in insulin and that facilitate domain closure in citrate synthase (Chothia *et al.*, 1983; Lesk & Chothia, 1984).

Other conformational changes occur on a smaller, sub-secondary structure scale. An important example of these changes is the closure of surface loops over substrates bound in active sites. These changes are an essential step in the catalytic mechanisms of certain enzymes. They have been observed in lactate dehydrogenase (Rossmann *et al.*, 1971; White *et al.*, 1976), triose phosphate isomerase (Phillips *et al.*, 1977) and HIV-1 protease (Miller *et al.*, 1989). Here, we analyze the mechanism that underlies one of these cases of loop closure: that which occurs in lactate dehydrogenase (LDH†).

We first use a variety of fitting techniques to categorize objectively the loop closure motion. Then to explain it we investigate the loop structure in depth: first, its main-chain torsions and then the packing, hydrogen bonding and sequence similarity of its side-chains. Finally, we show how the motion of the loop is integrated with the lesser motions of other parts of the protein.

† Abbreviations used: LDH, lactate dehydrogenase; VDW, van der Waals; r.m.s., root-mean-square; k_{cat}, catalytic rate constant; TIM, triose phosphate isomerase.

0022-2836/91/130133-17 $03.00/0

2. An Overview of LDH Structure and Loop Closure

LDH catalyzes the interconversion of lactate and pyruvate in the Cori cycle (for a review, see Holbrook *et al.*, 1975). It uses NAD as a coenzyme. In the early 1970s, Rossmann and his colleagues (Adams *et al.*, 1970, 1973) determined the structure of the apo enzyme. The loop closure was first noticed in 1971 (Rossmann *et al.*, 1971) and, subsequently, LDH ternary complexes (LDH–NAD–pyruvate, LDH–NAD–malate and LDH–NAD–oxamate) were solved to 3 Å (1 Å = 0·1 nm) (White *et al.*, 1976). Recently, the apo form and one ternary complex have been refined to high resolution and the co-ordinates deposited in the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). The apo form has a resolution of 2·0 Å and a residual of 0·202 (Abad-Zapatero *et al.*, 1987; Brookhaven file 6LDH), and the abortive ternary complex with substrate analogue oxamate and NADH has a resolution of 2·1 Å and a residual of 0·173 (unpublished refinement by J. P. Griffith & M. G. Rossmann; file 1LDM). These two structures will be considered here.

A standard labeled schematic drawing of LDH is shown in Figure 1(a), and its secondary structure assignments are given in Table 1. The enzyme is a tetramer. Each subunit consists of two domains: an NAD-binding domain and a catalytic domain. The structure of the NAD-binding domain found in LDH is very similar to that found in other dehydrogenases (Rossmann *et al.*, 1975). It consists of a large sheet of six strands (βA, βB, βC, βD, βE and βF) with three helices packed onto one face (αB, αC and α3G) and three packed onto the other (αD, αE and α1F). The catalytic domain consists of four helices (α2F, α1G, α2G and αH) packed on the outside of two sheets, which each have three strands (βG, βH and βJ, and βK, βL and βM). These two sheets have an extremely twisted conformation. There is one additional helix in the protein (αA) which is part of an extended piece of chain at the N terminus. In the ternary complex a flexible loop between βD and αD (Ala96 (98) to Phe119 (122)†) closes over the active site, covering the ligands (Fig. 1(b) and in closeup in Fig. 4). Relative to their position in the apo structure, atoms in the loop move up to 15 Å. To distinguish it from other loops in the protein, the Loop will henceforth be capitalized.

The detailed energetics of the Loop closure will not be discussed in depth here. A number of experimental investigations have suggested that electrostatics is the principal factor (e.g. see Wilks *et al.*, 1988). The Loop is held open by repulsion between positively charged residues in it (Arg99 (101) and Arg106 (109)) and in the active-site (H193 (195) and R169 (171)). Upon binding, the carboxylate on the lactate counterbalances these positive charges and so allows the loop to close.

3. Fitting Methodology

We used three fitting procedures to probe rigorously and objectively the structural differences between the two conformations of LDH on successively finer scales: sieve-fit, fit-refit and fit-all. The fit-all procedure is a novel procedure devised especially for its application here. We describe first how these fits are done and then their results‡. It is useful at the outset to make one definition: Δ = average (root-mean-square (r.m.s.)) deviation in main-chain atom positions in Å/atom after doing a least-squares fit on the main-chain atoms of the same residues in the apo and holo forms.

(a) *Sieve-fit*

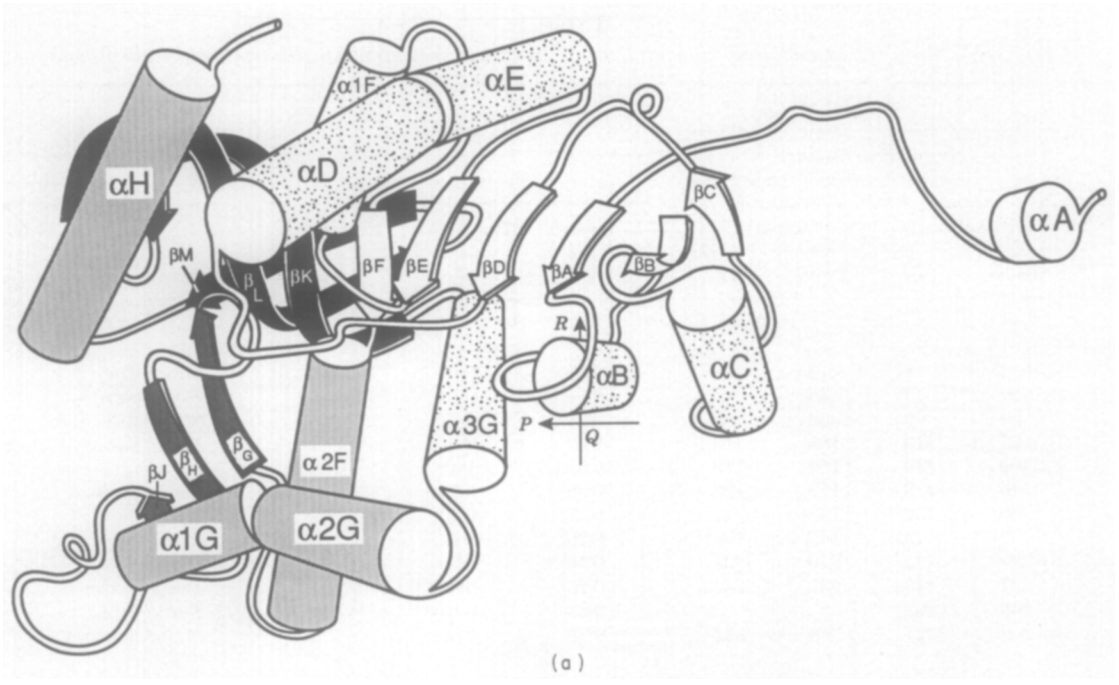
The "sieve-fit" procedure (Lesk, 1991) was used to characterize the conformational change at the grossest level. The apo form was fitted onto the ternary complex. Those residues that had the worst fit (i.e. had the highest deviation) were then excluded, and the two structures were refitted again with this smaller set of residues. This fitting and excluding was repeated a number of times until the deviation Δ for the regions being fitted dropped below a certain threshold.

Two threshold values were used with this procedure. First, the experimental uncertainty in atomic positions in the apo structure is estimated to be about 0·25 Å/atom (Abad-Zapatero *et al.*, 1987), and the ternary complex will have a similar value.

† The residue numbering convention used here will be the strictly sequential numbering of the Protein Data Bank file. An alternative X-ray numbering convention exists. Equivalences between these numbering systems are summarized below and given in brackets after selected residue names.

Sequential	X-ray
1–20	1–20
21–81	22–81
81–101	83–103
102–128	105–131
129–130	132A, 132B
131–206	133–208
207–209	209A, 210A, 210B
210–298	211–299
299–327	301–329
328–329	330A, 331

‡ All the fits (except fit-all), the counting of non-bonded contacts, the measurements of torsion angles and the accessible area calculations were carried out using programs written by A. M. Lesk (see Lesk, 1986, and references therein). For the purposes of these calculations a van der Waals (VDW) contact occurs when the centers of 2 atoms are less than the sum of their VDW radii plus 0·6 Å. These VDW radii are chosen to include hydrogen atoms, so, for example, C^{β} on Ala has the radius of a methyl group. The hydrogen bonding criteria is that donor and acceptor heavy (non-hydrogen) atoms be separated by less than 3·5 Å and that the donor-acceptor-next atom angle is less than 110°.



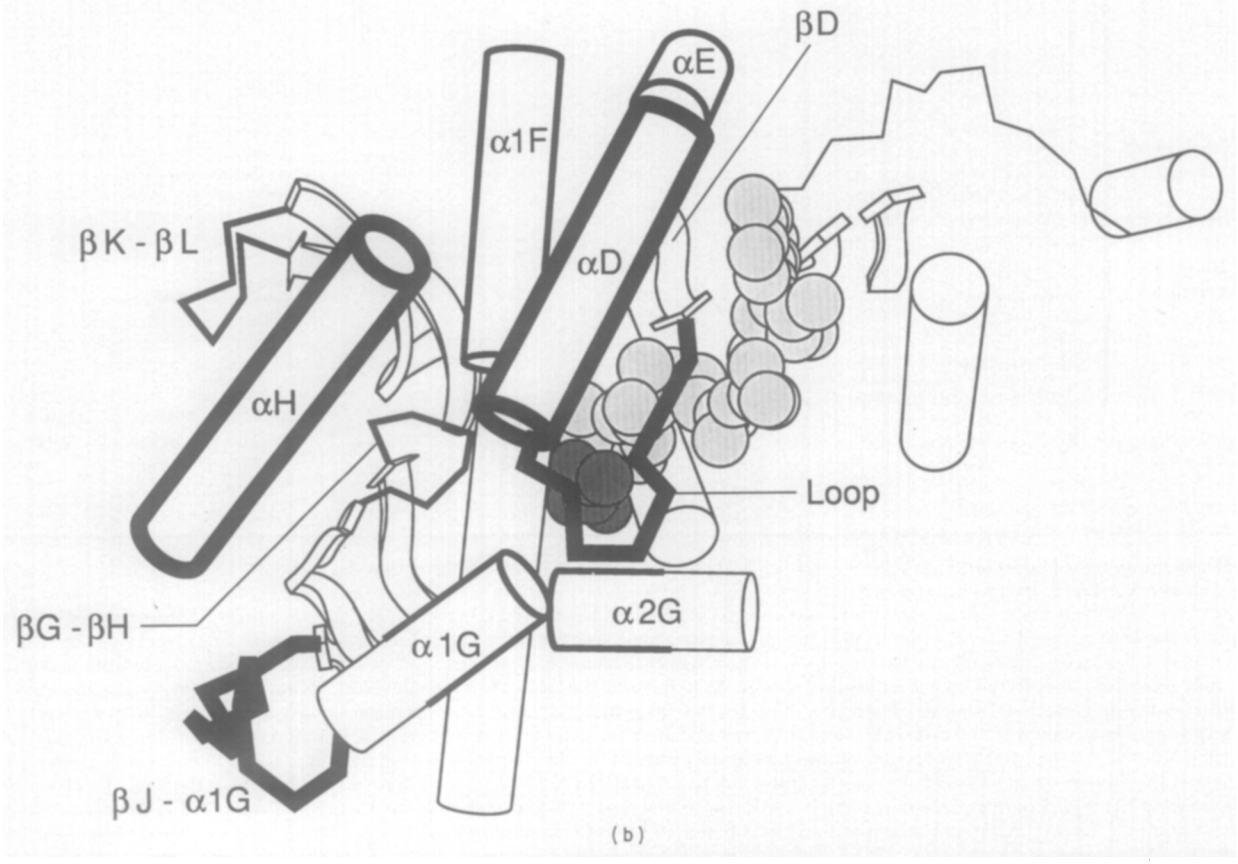


Figure 1. Schematics showing spatial arrangement of LDH secondary structure. (a) Figure of the apo form adapted from Abad-Zapatero *et al.* (1987) delineates labeling conventions for secondary structure. Sheet 1 is shown in white, sheet 2 in dark grey and sheet 3 in black. Helices packing onto sheet 1 are filled with dots, while those packing onto sheets 2 and 3 are shown in light grey. Only 1 subunit of tetramer is shown here. Other subunits are formed by rotating this one by 180° around the molecular *P*, *Q* and *R* axes. Views of the whole tetramer can be seen in Figs 7 and 9. (b) Figure of the ternary complex highlights the effects of loop closure. The NAD is represented by light grey spheres and the oxamate by dark grey ones. Parts of the protein changing conformation most from the apo form are labeled. Those moving most, the major movers, are shown with a heavy line, and those moving less, the minor movers, are shown with a medium-weight line. For clarity, some parts of the protein that are stationary have been omitted from this diagram. Figs 1(b), 4 and 6 are derived from computer programs written by A. M. Lesk (Lesk & Hardman, 1982).

Table 1

LDH secondary structures and comparison of their arrangement in the apo and holo forms

	(a) Secondary structure				(b) Fit-refit results			(c) Sieve-fit	
	Normally considered		Truncated for fits						
	From	To	From	To	Plastic deform Δ, (Å/atom)	Rigid-body translate T, (Å)	Rigid-body rotate, Ω (deg.)	Movement classification	White *et al.* (1976) comparison
Helices									
Overall					0·39				
A	2	8	3	8	0·26	0·13	2·7	Static core	0
B	29	43	30	43	0·29	0·14	1·6	Static core	0
C	54	69	56	69	0·24	0·07	1·2	Static core	0
D	105	119	105	119	1·39	2·81	22·6	Major mover	9000
cut D	105	119	110	119	0·34	1·38	13·8		
E	120	128	122	128	0·35	0·70	3·8	Minor mover	2000
1F	139	152	139	147	0·27	0·20	7·6	Minor mover	7000
2F	163	179	165	178	0·28	0·16	2·0	Static core	0
1G	224	233	226	233	0·36	0·39	8·4	Minor mover	0
2G	234	243	237	242	0·25	0·61	1·9	Minor mover	0
3G	246	263			0·24	0·15	1·9	Static core	0
H	306	324	308	322	0·39	2·56	7·2	Major mover	7800
Sheets									
Sheet 1									
Overall					0·31	0·12	0·5	Static core	
A	22	27							
B	47	51							
C	76	80	76	79					
D	90	95							
E	132	136							
F	157	161							
Sheet 2									
Overall					0·26	0·08	1·7	Static core	
G	186	192	186	190					
H	197	207	197	206					
J	209	211							
Sheet 3									
Overall					0·27	0·10	0·7	Static core	
K	265	274	265	272					
L	284	293	288	293					
M	296	301							
Loops									
Loop	96	104			1·18	5·59	35·2	Major mover	11,100
βG-βH	191	196			0·76	0·53	5·9	Minor mover	3000
βJ-α1G	212	223			1·75	0·82	8·6	Major mover	0
α1G-α2G	234	236			0·29	1·34	16·6	Minor mover	0
βK-βL	273	287			0·38	0·76	3·9	Minor mover	3000

Three groupings of information are presented in this Table. (a) This part of the Table defines the standard LDH secondary structures as well as other LDH substructures referred to in the text. The secondary-structure assignments are based on those in the Protein Data Bank files. However, the secondary-structure assignments in the data base differed markedly for the apo and holo conformations. These differences were resolved to make the final scheme delineated here. Occasionally, slightly truncating secondary structures was felt to give more meaningful values for the various fits and these truncations are indicated in the Table as well. (b) The results from doing the fit-refit procedure on selected substructures are shown. As defined in the text, the r.m.s. deviation Δ in Å/atom provides a measure of substructure deformation, whereas T and Ω refer to the magnitude of rigid-body translation (in Å) and rotation (in deg.). In interpreting these values it is worthwhile to keep in mind that Δ for fitting the whole subunit is 1·79 Å/atom and for fitting just the static core 0·32 Å/atom and that the uncertainty in atom positions in the apo form is 0·25 Å/atom. (c) An indication is given as to whether the substructure was included in the static core or considered as a major or minor mover based on the sieve-fit procedure discussed in the text. For comparison, results derived from White *et al.* (1976) on the unrefined structures are listed in the next column. These numbers are the maximum magnitude of the differences between apo and holo forms for each structural element as measured from Fig. 4(b) ($n = 4$) in White *et al.* (1976). The difference is expressed as a product of the fit of the holo structure into the electron density of the apo and *vice versa*.

Deviations Δ obtained from fits of individual secondary structures in both conformations had roughly similar values, 0·25 to 0·4 Å (see below and Table 1). So 0·4 Å was taken as a threshold to separate what did not move from what moved. Second, for fitting the whole protein, the mean deviation Δ is 1·8 Å, so 2 Å was taken to divide what moved a little from what moved a lot. These two thresholds were, in turn, used to partition the residues in protein into three categories:

(1) The static core: residues that were both part of well-defined secondary structure and that remained after sieve-fitting with the lower threshold.

(2) Major movers: contiguous residues that were excluded by the higher threshold.

(3) Minor movers: contiguous residues that were excluded by the lower but included by the higher threshold.

(b) *Fit-refit*

A "fit-refit" procedure, as described by Lesk & Chothia (1984), was used to characterize the motion of individual elements of secondary structure. In this procedure the apo and holo forms were first fit and superimposed on the basis of residues in the static core. Then a particular element of secondary structure was refitted and resuperimposed. The second fit results in a deviation Δ, which gives a measure of the deformation of the secondary structure in going between apo and holo forms, and the resuperimposition, a vector that measures its rigid-body translation and rotation. Occasionally, one or two residues at the termini of helices that unwound to some degree were not included in these fits.

(c) *Fit-all*

A "fit-all" procedure was devised to locate hinge regions in loops. It consists of four steps:

(1) Consider the Loop and its surroundings, residues 90 to 135, to be a single region. Calculate the deviations Δ for fitting all possible contiguous subregions. That is, for all residues r_1 and r_2 in the single region, fit r_1 to r_2 in the apo form with r_1 to r_2 in the ternary complex. Possible subregions include 96 to 106 or just 109 but not 96 to 106 together with 109, which are two non-contiguous subregions.

(2) The resulting deviations can naturally be arranged into a surface plot having the following form: $\Delta = f(r_1, r_2)$. This surface has a number of relevant characteristics. Since Δ does not depend on which direction the fits are done within, $f(r_1, r_2) = f(r_2, r_1)$, and it is only necessary to look at the subsurface where $r_2 > r_1$. Usually, but not always, it will be the case that the deviation Δ from fitting a subregion is less than that from fitting the whole region. (This would not be the case, however, if, say, residues 95 to 110 fit very poorly but residues 111 to 130 fit very well. A fit from 95 to 130 would have more overall deviation but less average deviation Δ than one done from just 95 to 110.) Because of this poorer fitting of larger regions, the surface will tend to slope up from the fits of small regions (say, 92 to 94 or just 94) to the fits of larger regions (say, 92 to 110). Eventually, it will climax approximately at the fit of the whole region, 90 to 135.

(3) Against this steady background increase there may be a large increase when one adds a residue that changes conformation significantly to a fit of a rigid region. Suppose that residues 90 to 95 move rigidly but there is a deformable hinge point at residue 96. One would expect that all possible fits from residues 90 to 95 would give relatively small deviations Δ of, say, 0·3 Å. After residue 96 is added to the fit, one would expect a sudden jump to a deviation Δ of, say, 0·8 Å.

(4) Consequently, maxima in the slope of the surface (i.e. points of inflection, where $\nabla|\nabla\Delta| = 0$) correspond in the protein to deformable boundaries between rigid regions.

4. Conformational Differences between the Apo and Holo Forms

(a) *Large-scale differences*

The regions with the same structure in the apo and ternary forms were determined by the sieve-fit procedure as summarized in Table 1 and Figure 1(b). Two-thirds of moving residues are in the catalytic domain and one-third in the NAD binding domain. The static core includes five helices (αA, αB, αC, α2F and α3G) and all three sheets. A fit of the residues in the static core resulted in a deviation Δ of 0·33 Å. The major movers include the Loop and αD, αH and the C terminus, and the random coil connecting βJ and α1G (βJ-α1G). The minor movers include four helices (α1F, αE, and the end of α1G through the beginning of α2G (α1G-α2G)) and two loops (the one connecting βH and βG (βG-βH) and one connecting βK and βL (βK-βL)). A graphical representation of the final iteration of sieve-fit can be seen in Figure 2, which shows the deviation in C^{α} position *versus* residue number after the structures were superimposed based on a fit of the static core. Altogether the major and minor movers form just over one-third of the structure (Table 1 and Fig. 2).

As summarized in Table 1, the magnitude and character of the large-scale structural changes was determined by the fit-refit procedure. With the exception of αD, the deviations Δ of all the helices (between 0·25 and 0·45 Å) are comparable to the accuracy of the co-ordinates, so the helices do not deform appreciably. Furthermore, except for αD, all the moving helices move as rigid bodies within the accuracy of the co-ordinates; for three of the six moving helices (αE, α2G and αH) the magnitude of rigid-body translation is 2 to 6·5 times Δ, and two of the remaining helices, α1F and α1G, rotate by roughly 8°. However, this rigid-body motion is clearly not the case for the mobile loop regions, which have large deviations Δ. As the whole of αD appears to deform as much as these mobile loops and as it is directly connected to the Loop, it is best to consider it as an integral part of the Loop region and not a separate, rigid helix.

(b) *Detailed characterization of the differences in Loop conformation*

The motion of the Loop region (i.e. Ala96 (98) to Phe119 (122)) is clearly the largest movement in LDH. The structural changes that occur within it were determined by the fit-all procedure, and a surface plot resulting from this procedure is shown in Figure 3. Inspection of this graph for points of inflection reveals two areas of very steep slope, which indicate that most of the motion in the Loop can be localized in two deformable regions or hinge

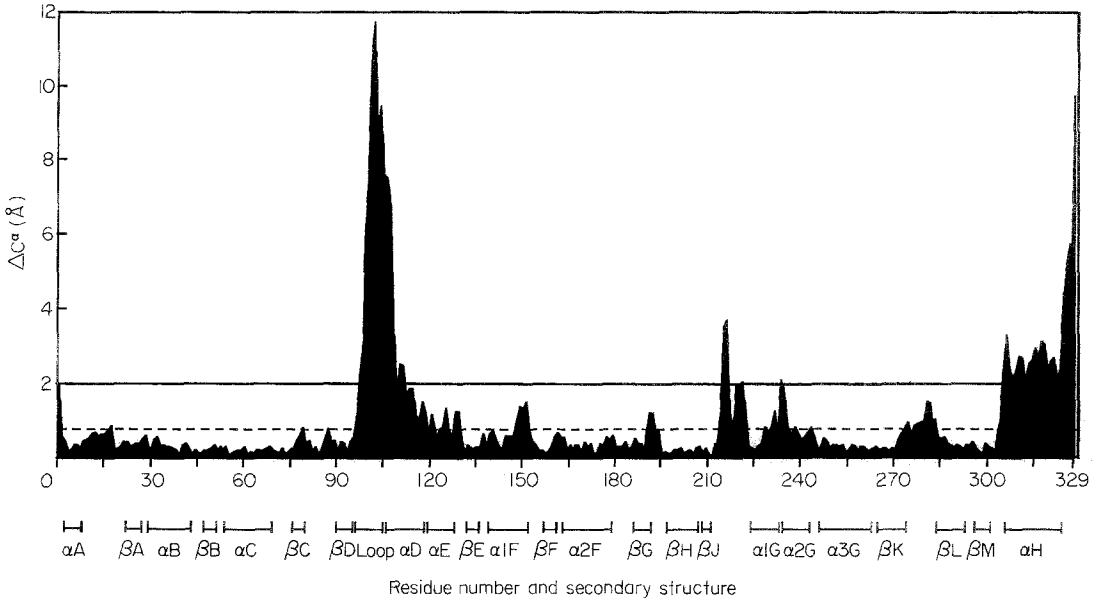


Figure 2. Graph showing the difference in C^{α} position between the 2 conformations after superposing the structures based on a least-squares fit of the main-chain atoms in the static core. The broken line at 0·4 Å/atom separates the static core from the minor movers, and the continuous line at 2 Å/atom, the major movers from the minor ones. Secondary structure is shown below.

joints. The first hinge is from Ala96 to Gln100 (102) at the beginning of the Loop, and the second hinge is from Ser105 (108) to Val110 (113) at the beginning of αD. The type II turn from Gln101 (103) to Gly104 (107), which connects these hinges, is carried along by their motion as a relatively rigid body. This double-hinged system is anchored on its N terminal side by βD and on its C terminal side by the non-deforming part of the αD that runs from Gln111 (114) to Phe119 (see Table 1). There is also a third region of moderately steep slope on the graph. This corresponds to a slight kinking or deformation at the junction of αD and αE, i.e. around Phe119. Figure 4 shows the Loop backbone in both conformations, and the hinge joints are visible on careful inspection.

(c) *Comparison with previous results on LDH loop closure*

The results of White *et al.* (1976) are qualitatively similar to those described here. However, because of the medium resolution and unrefined nature of their structures, it was not possible then to describe accurately the exact extent of the movement, its rigid-body nature, or its double-hinged character. Some of the differences from our results are summarized in Table 1. In particular, White *et al.* (1976) found a much larger movement for αIF and a smaller movement for αIG-α2G than we do.

A hinge-jointed mechanism for the Loop has been alluded to previously but never discussed in depth (e.g. Abad-Zapatero *et al.*, 1987; Wilks *et al.*, 1988; Birktoft *et al.*, 1982). In particular, contrary to our findings, Gly103 (106) has been suggested as the location of a hinge joint. A mutagenesis experiment described by Waldman *et al.* (1988) supports our conclusion that Gly103 is not a hinge joint. This residue was changed to tryptophan. The rate constant (k_{cat}) was reduced by only 40%, and the Michaelis constant was unaffected. Loop closure is the rate-limiting step in normal LDH catalysis, occurring at a maximal rate of about 250 times per second (Holbook & Gutfreund, 1973; Clarke *et al.*, 1986). Consequently, the fact that k_{cat} was essentially unchanged indicates that the mutation did not affect the dynamics of loop closure. Such a mutation clearly puts many steric constraints on the flexibility of residue 103 and would probably have a large effect on k_{cat} if Gly103 actually did deform in hinge-like fashion.

Two additional LDH crystal structures solved by Rossmann and his colleagues provide further evidence for the double hinged-structure proposed here. These structures are not to as high a resolution or degree of refinement as the dogfish (*Squalus ascanthius*) apo and holo structures, so they were not consulted for the fine points of the analysis. However, they have noticeable features that clearly point out the flexibility of the second hinge. In the crystal structure of the apo mouse testicular C_4 LDH, Musick & Rossmann (1979) found residues Leu107 (110) to Leu109 (112) at the beginning of αD to be in an extended chain with an intermediate conformation between that of the open and closed forms. And in the structure of pig heart LDH with an S-Lac-NAD substrate analog, Grau *et al.* (1981) found that the beginning of αD differed the most in conformation from the dogfish holo structure considered here.

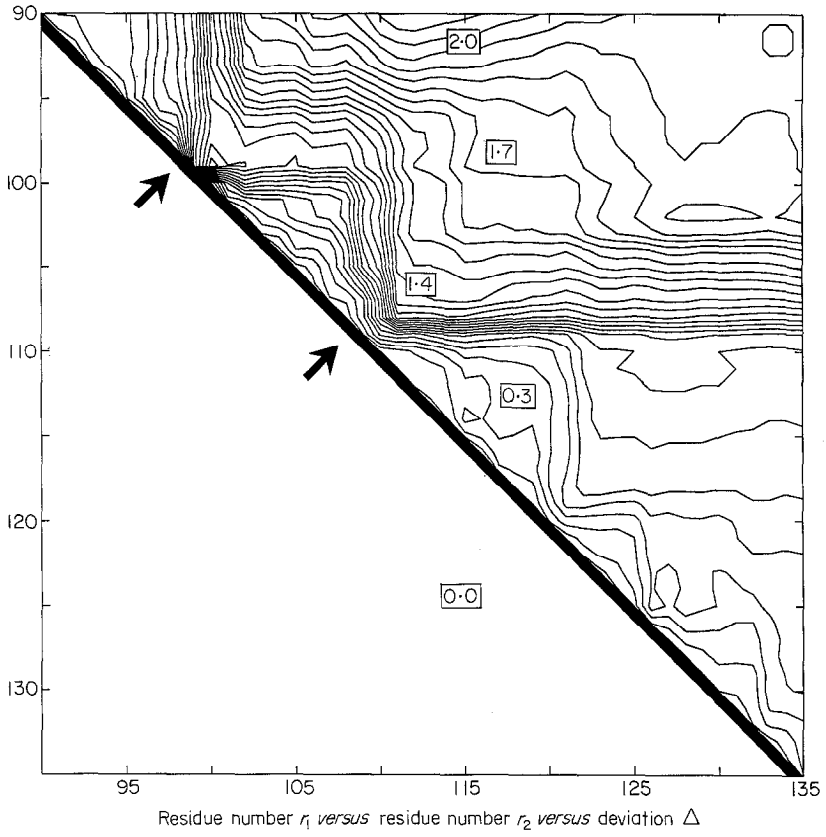


Figure 3. Contour plot showing results of doing fit-all procedure on the loop region of LDH. Least-square backbone fits between apo and holo forms were done from residue number r_1 to residue number r_2. The r.m.s. deviation Δ of atoms in Å/atom is shown as contours of increasing deviation. Forty contours between 0 and 2·25 Å/atom are shown. Rough values of Δ for various parts of the surface are shown in boxes. The 2 regions of maximal slope in this graph, indicated by arrows, correspond to hinges in the Loop.

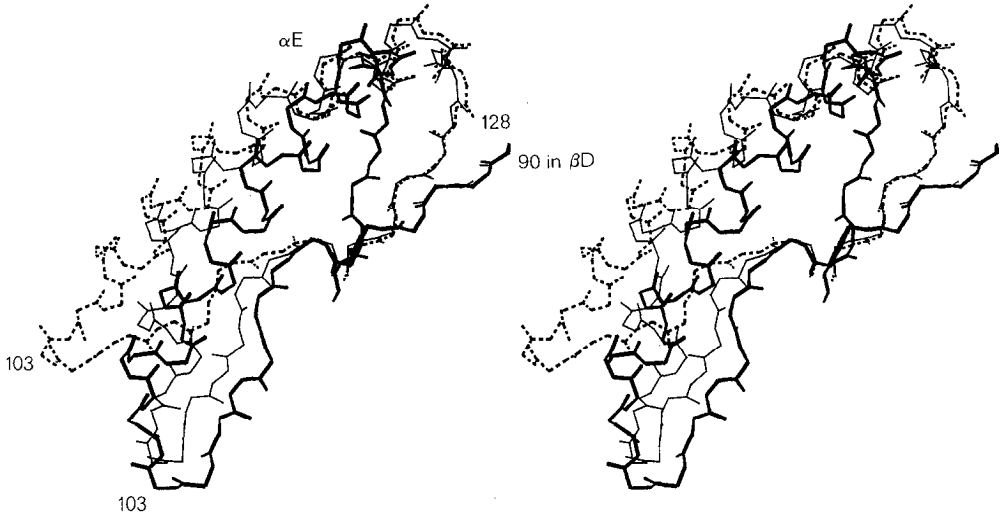


Figure 4. Superposition of 3 conformations of the Loop shows how loop closure can be roughly approximated by changing 3 torsion angles. The light continuous line shows holo conformation main-chain atoms from Lys90 to Val135. The broken line shows apo conformation of loop superposed on the holo form based on a least-square fit of just the backbone atoms in βD and βE, Lys90 to Thr95 and Ile132 to Val135. The bold continuous line shows this apo conformation with 3 key torsion angles changed to their values in the holo form: ϕ_{96}, which changes for −105° to −69°; ϕ_{97}, −174° to 146°; and ϕ_{105}, −94° to −75°. The apo loop with 3 torsions changed fits the holo loop with deviation Δ of 1·08 Å/atom and a translation vector magnitude of 2·6 Å, while an identical fit of the unchanged apo loop gives 1·20 and 5·7.

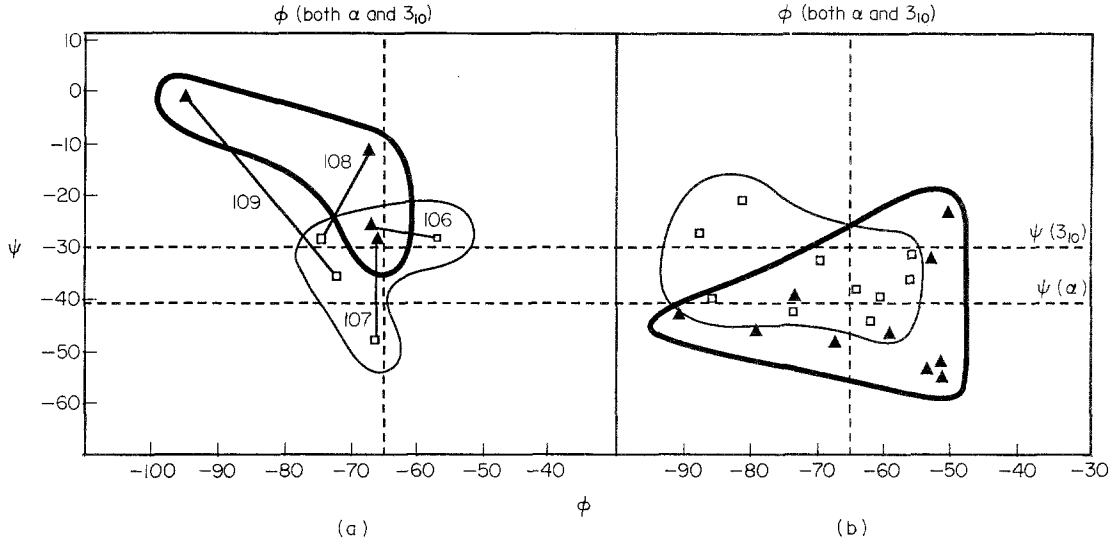


Figure 5. Ramachandran diagrams for residues in αD. (a) For the 2nd hinge (Arg106 to Leu109) and (b) for the rest of αD (Val110 to Phe119). Only the lower left-hand corner of the diagram is shown as ψ, ϕ values for all residues are clustered here. (▲——) Values for the holo conformation; (□ ——) values for the apo form. For comparison, values of ideal α-helical and 3_{10}-helical conformations are marked by the intersection of the broken lines. For the 4 residues in the 2nd hinge, the residue number of the ψ, ϕ value is indicated next to a line connecting the apo to holo values. Note both the 2nd hinge and the rest of αD roughly occupy the same region in torsion-angle space in the apo conformation. However, in the ternary complex, the 2nd hinge stretches into a markedly 3_{10}-helical conformation, while the rest of αD moves slightly in the opposite direction towards a more α-helical conformation.

5. Loop Main-chain Torsion Angle Analysis

Analysis of structural features of the Loop support and corroborate a main result of the fitting procedures, that the Loop has a double-hinged structure. It is best to consider first the main-chain and move later to the side-chains. As shown in Figure 4, if one starts with the apo form and changes three main-chain torsion angles to their values in the ternary complex it is possible to reproduce much of the effect of loop closure. The importance of these three torsions is easily understandable. ϕ_{96} and ϕ_{97} are the first torsion angles to change appreciably between conformations in the first hinge, and ϕ_{105} is the first torsion to change appreciably in the second hinge. The large impact of changing these torsions simply results from the fact that in any joint the changes near the beginning will have greater effect than those later on.

Torsion angle changes in other loop residues do not have the large, concerted effect of the changes in these three key torsions. A case in point are the changes in ψ_{99} (162°) and ϕ_{100} (148°), which simply result in the flipping of the Arg99-Gln100 (101–102) peptide plane. Consequently, simply plotting the magnitude of torsion angle changes *versus* residue number does not furnish useful information about loop motion. Plotting the pseudo-torsion angle α, (the torsion angle for virtual bonds connecting C^{α} atoms), *versus* residue number averages out some of these "local" fluctuations and gives a roughly bimodal distribution that corroborates the proposed double-hinged structure. However, the peaks on this distribution were not nearly as pronounced as those coming from the fit-all procedure.

From a detailed inspection of views such as those shown in Figure 4, it is clear that the motion of the first hinge can be better accounted for by changing a few torsions than that of the second hinge. That is, after changing the three key torsions, the conformation of Lys90 (92) to Glu104, which reflects only the effects of the first hinge, matches the conformation of the holo form much better than the conformation of Ala98 (100) to Val135 (137), which reflects only the effects of the second hinge. This degree of "matching" can be quantified by comparing the fit of the apo form, with torsions first unchanged and then changed, to the holo form and judging the relative improvement. It turns out that changing ϕ_{96} and ϕ_{97} improves the fit of Lys90 to Gln104 by 22%, but changing ϕ_{105} only improves the fit of Ala98 to Val135 by 2%.

It is possible to understand better the flexibility of the second hinge by plotting its torsions in a Ramachandran diagram, as shown in Figure 5. In the apo form residues Arg106 (109) to Phe119 (all of αD) occupy one relatively large region of $\phi\psi$-space. On average they have a configuration ($\langle\psi\rangle = -36°$) midway between that of a 3_{10}-helix ($\psi = -30°$) and an α-helix ($\psi = -41°$). However, in the ternary complex, αD splits into two distinct regions of $\phi\psi$-space. Arg106 to Leu109 ($\langle\psi\rangle = -17°$) occupy a region closer to an ideal 3_{10}-helix, while Asn110 to Phe119 ($\langle\psi\rangle = -44°$) move towards a region with a more α-helical configuration. This splitting of αD in $\phi\psi$-space closely matches its segmentation with the fit-all procedure into a deformable hinge region at the beginning and a relatively rigid helix at the end. Its real-space manifestation is that the second hinge

Table 2
Fits of all and parts of αD to ideal 3_{10} and α helices

	r.m.s. deviation Δ in fits to ideal helices	
	α	3_{10}
All of helix D (106–118)		
Apo	0·75	1·86
Holo	1·51	2·14
Second hinge (106–109)		
Apo	0·25	0·56
Holo	0·39	0·28
Rest of helix (110–119)		
Apo	0·52	1·12
Holo	0·55	1·43

Δ is as defined in the text, the deviation in Å/atom after doing a least-squares fit on the main-chain atoms.

stretches considerably: the C^{α} atoms on Ser105 and Gln111 move apart by 3·3 Å.

The splitting of αD is also evident using other main-chain conformation analysis techniques. As shown in Table 2, fits of parts of αD against ideal α- and 3_{10}-helices show the bifurcation clearly. In the conformational change, the main-chain hydrogen bonding in the second hinge shifts from predominantly 1,4 to 1,3. However, clear analysis of the hydrogen bonding is not as straightforward because in either conformation many of amide and carbonyl groups have an intermediate geometry that can accommodate both 1,3 and 1,4 bonding.

6. Packing and Sequence Conservation of Loop Side-chains

At this point, it is convenient to switch focus from describing the main-chain deformations that occur upon loop closure to analyzing how the specific side-chains present in the Loop make this motion possible. It will be worth while to look at the Loop residues for patterns of sequence similarity, VDW contacts and hydrogen bonding.

(a) *Sequence alignment*

To facilitate sequence comparison 17 LDH sequences were aligned. Parts of this alignment corresponding to the major and some minor movers are shown in Table 3. The sequences used were from 79% to 33% identical with dogfish muscle LDH. They were grouped into ten eukaryotic sequences with an average identity of 73% to dogfish LDH and seven prokaryotic sequences with an average identity of 35%. In the overall alignment 12% of the residues (41 of 329) were completely conserved.

In terms of overall criteria for sequence similarity, the Loop was one of the most highly conserved substructures in LDH. This contrasts with the *a priori* expectation that surface loops are the least well conserved parts of protein structure. Of the 23 residues between Ala96 and Phe119, six were totally, and 11 highly, conserved. On average, at a given residue position, the dogfish muscle sequence matched 12·5 of the 16 other sequences, *versus* an average of nine matches overall and roughly ten for residues in the static core.

(b) *Hydrogen-bonding patterns*

The hydrogen bonds internal to the Loop and αD can be divided into three classes.

(1) "Transverse" hydrogen bonds join one side of the loop region to the other.

(2) "Longitudinal" hydrogen bonds link residues along one side of the Loop *via* side-chain atoms.

(3) "Helical" hydrogen bonds link the main-chain atoms of αD in a conventional 1,4 or 1,3 pattern. These were discussed in the previous section on main-chain conformation.

As shown in Figure 6, in both the apo and holo conformation, the same five transverse hydrogen bonds are present. Three of these transverse hydrogen bonds involve just residues in the type II turn. They undoubtedly stabilize this substructure and perhaps account for its rigidity compared to the hinges surrounding it. In the apo form there are five longitudinal hydrogen bonds. Two of these are involved in an Arg112-Glu104 (115–107) salt bridge that spans the second hinge. When the Loop closes, these five bonds are broken and replaced with three different ones. However, the salt bridge is still preserved to a degree. It stretches, the C^{α}-to-C^{α} distance between Arg112 and Glu104 increasing from 10·4 to 13·4 Å, and a new hydrogen bond ($N^{\eta 2}$–$O^{\varepsilon 2}$) takes the place of the two previous ones (N^{ε}–$O^{\varepsilon 2}$ and $N^{\eta 2}$–$O^{\varepsilon 1}$).

Comparison of the list of hydrogen bonds attached to Figure 6 with sequence alignment in Table 3 shows that the side-chains involved in transverse hydrogen bonds, i.e. Asn113 (116) and Glu104 (107), are better conserved than those involved in longitudinal hydrogen bonds. So transverse hydrogen bonds are also probably preserved to a greater extent than the longitudinal ones across LDH sequences. Moreover, VDW contacts across the Loop region are probably preserved through many LDH sequences. In particular, there are three highly conserved, hydrophobic residues along the inside of αD (Leu109 (112), Ile116 (119) and Phe117 (120)) that with Asn113 pack against the opposite side of the Loop and give it a small hydrophobic "core".

The retention of essentially the same loop conformation in the open and closed forms is facilitated by the conservation of transverse hydrogen bonds. This is in accord with the findings of Tramantano *et al.* (1989) that structures of medium sized loops are mainly stabilized by hydrogen bonds to their inward pointing polar groups.

(c) *Structural constraints determine how the hinges deform*

The steric environment of each hinge is very different. Residues of the first hinge (Ala96-Gly-Ala-

Table 3

Selected parts of the alignment of 17 LDH sequences

Key	Description	Reference	096	097	098	099	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
DM	M chain – dogfish	Taylor (1977)	A	G	A	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	I	F	K	F
HM	M chain – human	Tsujibo *et al.* (1985)	A	G	A	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	I	F	K	F
PM	M chain – pig	Klitz *et al.* (1977)	A	G	A	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	I	F	K	F
RM	M chain – rat	Matrisian *et al.* (1985)	A	G	A	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	I	F	K	F
CM	M chain – chicken	Torff *et al.* (1977)	A	G	A	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	I	F	K	F
HUB	B chain – human	Takeno & Li (1989)	A	G	V	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	V	F	K	F
PH	H chain – pig	Klitz *et al.* (1977)	A	G	V	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	V	F	K	F
CH	H chain – chicken	Torff *et al.* (1977)	A	G	V	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	V	F	K	F
MUC	C chain – mouse	Sakai *et al.* (1987)	A	G	A	R	M	V	S	G	E	T	R	L	D	L	L	Q	R	N	V	A	I	M	K	A
RX	X chain – rat	Pan *et al.* (1983)	A	G	A	R	M	V	S	G	Q	S	R	L	A	L	L	Q	R	N	V	T	I	M	K	A
BSN	*Bacillus stearothermophilus*	Barstow *et al.* (1986)	A	G	A	N	Q	K	P	G	E	T	R	L	D	L	V	D	K	N	I	A	I	F	R	S
BM	*Bacillus megaterium*	Widerkehr (1982)	A	G	A	N	Q	A	P	G	E	T	R	L	D	L	V	E	K	N	V	X	I	F	E	X
LC	*Lactobacillus casei*	Hensel *et al.* (1983)	A	G	A	P	K	Q	P	G	E	T	R	L	D	L	V	N	K	N	L	K	I	L	K	S
BSU	*Bacillus subtilis*	Hediger *et al.* (1986)	A	G	A	N	Q	K	P	G	E	T	R	L	E	L	V	E	K	N	L	K	I	F	K	G
BCA	*Bacillus caldotenax*	Barstow *et al.* (1987)	A	G	A	N	Q	K	P	G	E	T	R	L	D	L	V	D	K	N	I	A	I	F	R	S
TAQ	*Thermus aquaticus*	Kunai *et al.* (1986)	A	G	V	A	Q	R	P	G	E	T	R	L	Q	L	L	D	R	N	A	Q	V	F	A	Q
BL	*Bifidobacterium longum*	Minowa *et al.* (1989)	A	G	P	R	Q	K	P	G	Q	S	R	L	E	L	V	G	A	T	V	N	I	L	K	A

α D

Loop

Key	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242
DM	H	P	L	D	G	T	N	K	D	K	Q	D	W	K	K	L	H	K	D	V	V	D	S	A	Y	E	V	I	K	L	K	G
HM	H	P	D	L	G	T	D	K	D	K	E	Q	W	K	E	V	H	K	Q	V	V	E	S	A	Y	E	V	I	K	L	K	G
PM	H	P	E	L	G	T	D	A	D	K	E	H	W	K	A	V	H	K	E	V	V	D	S	A	Y	E	V	I	K	L	K	G
RM	N	P	Q	L	G	T	D	A	D	K	E	Q	W	K	D	V	H	K	Q	V	V	D	S	A	Y	E	V	I	K	L	K	G
CM	H	P	D	M	G	T	D	A	D	K	E	H	W	K	E	V	H	K	Q	V	V	D	S	A	Y	E	V	I	K	L	K	G
HUB	N	P	E	M	G	T	D	N	D	S	E	N	W	K	E	V	H	K	M	V	V	E	S	A	Y	E	V	I	K	L	K	G
PH	D	P	A	M	G	T	D	K	D	S	E	N	W	K	E	V	H	K	Q	V	V	E	S	A	Y	E	V	I	R	L	K	G
CH	N	P	E	M	G	T	D	N	D	S	E	N	W	K	E	V	H	K	M	V	V	E	S	A	Y	E	V	I	K	L	K	G
MUC	N	P	A	I	G	T	D	S	D	K	E	H	W	K	N	V	H	K	Q	V	V	E	G	G	Y	E	V	L	N	M	K	G
RX	N	P	A	I	G	S	D	S	N	K	Q	E	W	K	T	V	H	K	Q	V	V	D	G	G	Y	E	V	L	D	L	K	G
BSN	V	E	S	K	G	E	E	A	Q	K	D	L	—	E	R	I	F	V	N	V	R	D	A	A	Y	Q	I	I	E	K	K	G
BM	I	N	T	A	A	I	E	K	E	P	D	M	—	Q	H	L	F	S	Q	X	R	D	A	A	Y	T	I	I	N	X	K	X
LC	V	K	A	H	P	E	I	K	E	D	K	L	—	V	K	M	F	E	D	V	R	D	A	A	Y	E	I	I	K	L	K	G
BSU	V	E	K	N	D	A	Y	K	Q	E	E	L	—	D	Q	I	V	D	D	V	K	N	A	A	Y	T	I	I	E	K	K	G
BCA	V	E	S	K	G	E	E	A	Q	K	E	L	—	E	R	I	F	V	N	V	R	D	A	A	Y	Q	I	I	E	K	K	G
TAQ	A	E	A	R	G	R	P	L	S	P	E	D	R	A	R	I	D	E	G	V	R	R	A	A	Y	R	I	I	E	G	K	G
BL	T	P	L	P	G	H	D	P	L	D	A	D	K	E	E	I	H	Q	E	V	K	N	A	A	Y	K	I	I	N	G	K	G

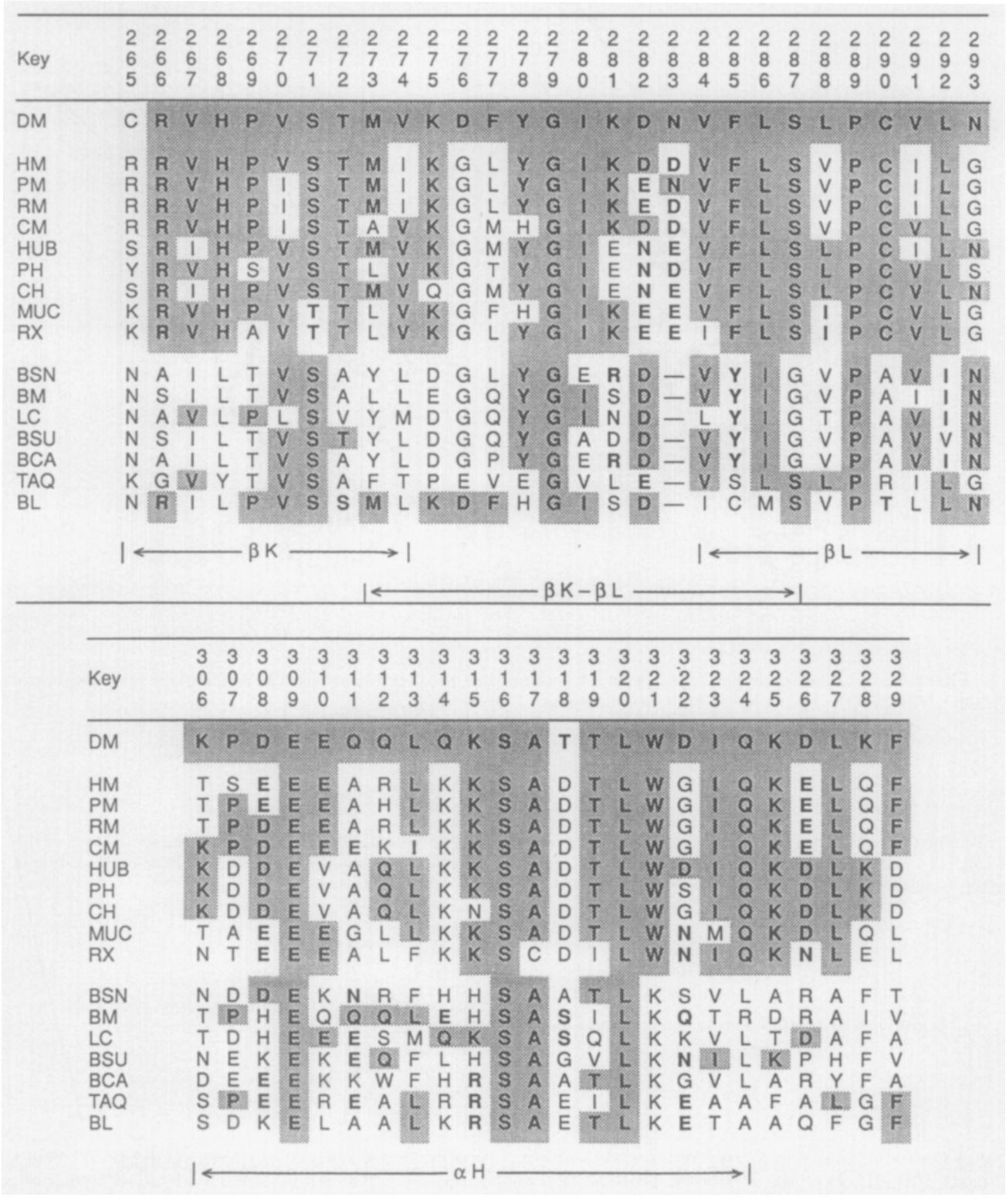
α 1G

α 2G

β J – α 1G

α 1G – α 2G

Table 3 *(continued)*



As indicated by the secondary structure assignment at the bottom of the Table, only the residues corresponding to the Loop, αH, βK-βL, α1G-α2G and βJ-α1G are shown. The numbering is sequential and secondary structure assignment is for dogfish muscle LDH. References for all sequences are listed.

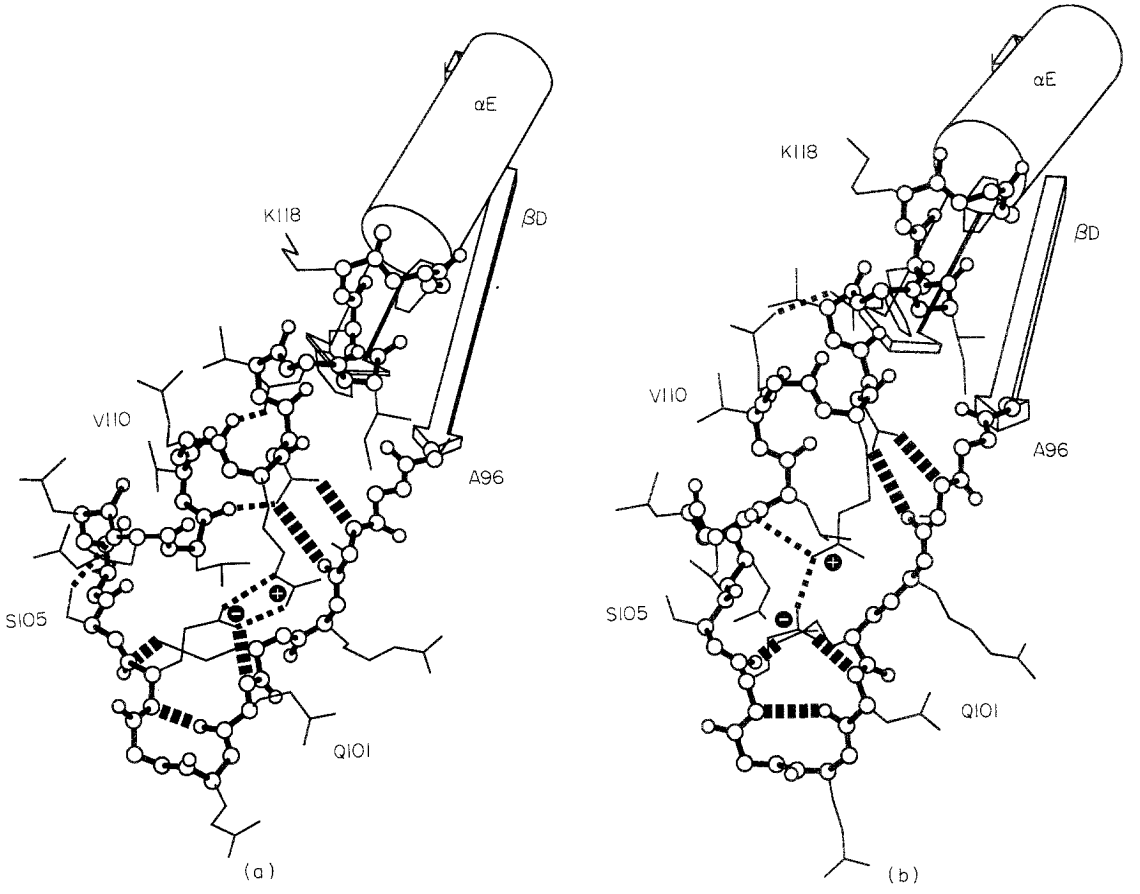


Figure 6. Hydrogen bonding in Loop region for both apo and holo conformations. The 5 transverse hydrogen bonds (thick broken lines) are present in both conformations.

Ala98	N	to Asn113	$O^{\delta 1}$
Gln101	N	to Asn113	$O^{\delta 1}$
Gln101	N	to Glu104	$O^{\delta 1}$
Glu104	N	to Gln101	O
Asn113	$N^{\delta 2}$	to Ala98	O (only marginal geometry)

The longitudinal hydrogen bonding (thinner broken lines) changes completely between conformations. However, the Arg112-Glu104 salt bridge (indicated by + and − signs) is maintained to some degree. No helical hydrogen bonding is shown. (a) Longitudinal hydrogen bonds in the apo form are:

Asn108	N	to Ser105	O^{δ}
Arg112	N^{ϵ}	to Glu104	$O^{\epsilon 1}$
Arg112	$N^{\eta 2}$	to Glu104	$O^{\epsilon 2}$
Asn113	$N^{\delta 2}$	to Leu109	O
Asn115	$N^{\delta 2}$	to Gln111	O

(b) Longitudinal hydrogen bonds in the holo form are:

Gln111	$N^{\epsilon 2}$	to Asn115	$O^{\delta 1}$
Arg112	$N^{\eta 2}$	to Glu104	$O^{\epsilon 2}$
Arg112	$N^{\eta 2}$	to Asn108	O

Arg-Gln100) make few VDW contacts with other parts of the Loop or the rest of the protein. This is particularly true of Arg99 (101), which points directly out into solution and makes no contacts, and of the completely conserved, small residues, Ala96 (98) and Gly97 (99). Partially, as a result of the small side-chains and few VDW contacts in the first hinge, much of its main chain is exposed to solvent: on average each residue of the hinge has 23 Å^2 of main chain solvent-accessible surface area (Lee & Richards, 1971). In contrast, residues of the second hinge (Arg106-Leu-Asn-Leu-Val110) have larger side-chains that are involved in numerous packing interactions with αH and the rest of the Loop. The main chain of the second hinge is completely buried by its tightly packed side-chains and only exposes 3 Å^2 of surface area per residue.

The clearly different steric environments of the hinges manifest themselves in the constraints that they put on possible deformations, and these constraints become very clear when the interaction of the Loop and the two ligands is considered. The

first hinge primarily interacts with the coenzyme, which make 21 VDW contacts and two hydrogen bonds with it *versus* 11 contacts and no hydrogen bonds with the rest of the Loop. Seventeen of these 21 contacts are to main-chain atoms which, because of their few structural constraints, are free to "wrap" around the NAD. Moreover, the absence of steric hindrance allows the main chain to twist greatly on one torsion, i.e. to kink, rather than having to spread its deformation over many torsions. Thus, the loose structural constraints on the first hinge explain why its motion can so well be accounted for by just changing two torsion angles, ϕ_{96} and ϕ_{97}.

Structural constraints can also help explain the motion of the second hinge, though in a slightly different fashion. It interacts primarily with the oxamate, which makes six VDW contacts and three hydrogen bonds with Arg106 (109). The oxamate makes only one other contact, a hydrogen bond, with a residue in the rest of the Loop, Gln100 (102), and this lone hydrogen bond would not even be present for lactate or pyruvate, the true substrates, since it involves the amino group of oxamate. Because of the numerous contacts it makes with the substrate, it is not surprising that Arg106 is conserved in all known LDH sequences and that Clarke *et al.* (1986) found that changing Arg106 to glutamine reduced k_{cat} by a factor of roughly 400. In contrast to the ligand interaction in the first hinge, the oxamate does not make any contact with the second hinge main-chain atoms. It has to affect their conformation indirectly through a flexible arginine side-chain, which stretches from 5·5 to 7·1 Å (C^{α} to C^{ε}) in the conformational change. Moreover, because of the many packing constraints imposed on it, the main-chain is not as free to deform as that in the first hinge and so must spread its deformation over many torsion angles. Thus, tight structural constraints help to explain why the deformation in the second hinge cannot be accounted for by changing one or two torsions and rather involves a bending motion spread over much of the helix. These constraints also help to rationalize the highly irregular conformation that αD assumes in the holo form. In the conformational change, the standard deviations of both its torsion angles increase considerably, more than doubling from 7° to 17° for $\langle \Delta\psi \rangle$.

7. Propagation of Conformational Change

With the exception of the Loop and βG-βH, neither ligand has significant, direct interaction with any of the other major or minor movers, i.e. no contacts with αH and βK-βL and only marginal contacts to α1G-α2G, αE and α1F.

How do the ligands, when they bind, induce conformational changes in these structures without making contact with them? The Loop is the key. As White *et al.* (1976) suggested, all the major displacements can be directly associated with the Loop movement. Of the 101 VDW contacts that the Loop and αD make in the apo form with the rest of the protein, 92 are to the major and minor movers. The Loop and αD make contacts to α1F, αE and αH in both conformations, and loop closure leads to interactions with α2G and βG-βH in the holo form. As shown in Figure 7, the effects of Loop closure are propagated through this network of contacts.

(a) *Change in packing of the second hinge and αH*

Most of the Loop contacts, i.e. 51 VDW contacts and two hydrogen bonds, are to αH and the C terminus, and this part of the protein undergoes the second largest conformational change (after the Loop). Figure 8 shows a schematic view of the packing, where the packing position occupied by each side-chain is represented by a sphere. Most side-chains accommodate the conformational change by just slightly rocking. That is, they change their side-chain torsion angles by about 10°. They do not repack and change their immediate neighbors, and they still have the same packing position. Some of these rocking side-chains are well conserved, especially the hydrophobic ones such as Val110 (113), Val114 (117) and Ile323 (325).

However, as shown in Figure 8, some side-chains completely repack and move to new positions. In particular, the protrusions in the second hinge formed by the side-chains of Arg106 and Leu107 completely repack with respect to the groove in αH formed by Thr319 (321), Asp322 (324) and Ile323 (325). In a motion similar to a cog hopping between grooves, Arg106 leaves its packing position in the αH groove and moves down to make new contacts with the substrate. Leu107, in turn, moves into the groove and leaves its previous packing position unoccupied. This "groove-hopping" motion clearly manifests how tight the steric constraints are on the second hinge. Leu107 just barely fits into the αH groove and makes two very close contacts with Ile323. As perhaps is to be expected because of these tight constraints, Leu107 and Arg106 are completely conserved in all LDH sequences (though as discussed before, there are also other reasons why Arg106 would be conserved), and Ile323 is also highly conserved (Table 3). The C terminus executes a similar form of musical chairs. In the conformational change, six side-chains rotate through four packing positions. Leaving its apo packing position unoccupied, Lys328 (330A) takes the place of Phe329 (331); Phe329, in turn, takes the place of Leu327 (329); Leu327 takes the place of Asp326 (328); and Asp326 takes the place of Lys325 (327), which moves to a previously unoccupied position. The C terminus has very high temperature factors and perhaps is somewhat disordered. However, this repacking is a gross change that is still meaningful despite the poorer quality of the co-ordinates in this region.

(b) *Contacts to α1G-α2G and βK-βL*

When the Loop closes it makes contact with α1G and α2G. The formation of 18 new VDW contacts

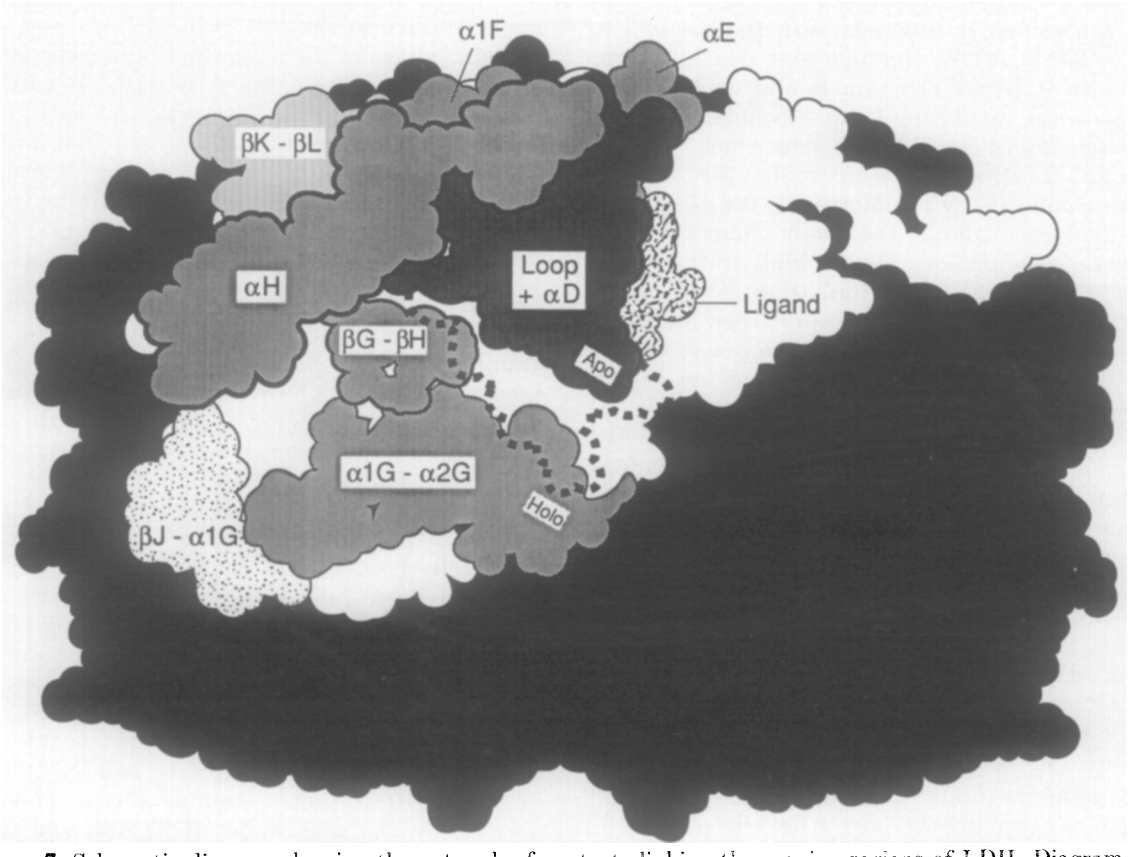
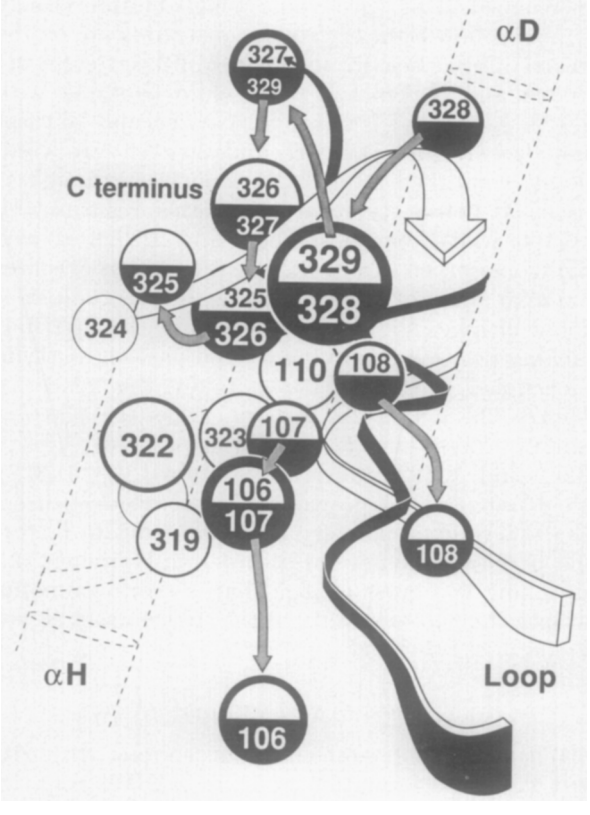


Figure 7. Schematic diagram showing the network of contacts linking the moving regions of LDH. Diagram shows outlines of VDW surfaces (as seen in projection from solvent) of the major and minor movers in the apo form. The outline of the Loop in holo form is indicated by the thick dotted line. The outline of the part of the ligand not covered by the Loop is also shown. Note how the Loop makes contact with βG-βH and α1G-α2G only in the holo form and how the minor mover βK-βL is connected to the ligand only through αH and the Loop.



between the Loop plus αD and Ala235 (236), Tyr236 (237), Ile239 (240) and Lys240 (241) on α2G provides a driving force for loop closure. As shown in Table 3, two of these contact residues, Tyr236 and Ile239, are conserved in all LDH sequences, and Parker *et al.* (1982) found that nitration of Tyr236 in pig heart LDH reduces enzymatic activity, presumably by interfering with the contacts it forms with Glu102 on the Loop. Wilks *et al.* (1990) made three sets of mutations that reduced the size of the residues at the Loop–α2G interface, and in each case k_{cat} decreased commensurately with the magnitude of the change. In bacterial LDH, the mutations were Gln100-Lys101-Pro102 to Met-Val-Ser, k_{cat} down by 3·8; Ala235-Ala236, k_{cat} down by 1·5; and a

Figure 8. Schema showing the packing of αD onto αH and the C terminus. Circles represent packing positions for side-chains, and the numbers inside circles indicate residue numbers of side-chains. Ribbons show main-chain conformation. Circles split into 2 halves indicate that a packing position is occupied by different side-chains in apo and holo conformation. Black filled half-circles and ribbons indicate holo conformation, while white half-circles show apo. Grey arrows show the movement of side-chains through packing positions. Note the "cog-hopping-between-grooves" movement of Arg106 and Leu107 and the elaborate rearrangement of the C terminus.

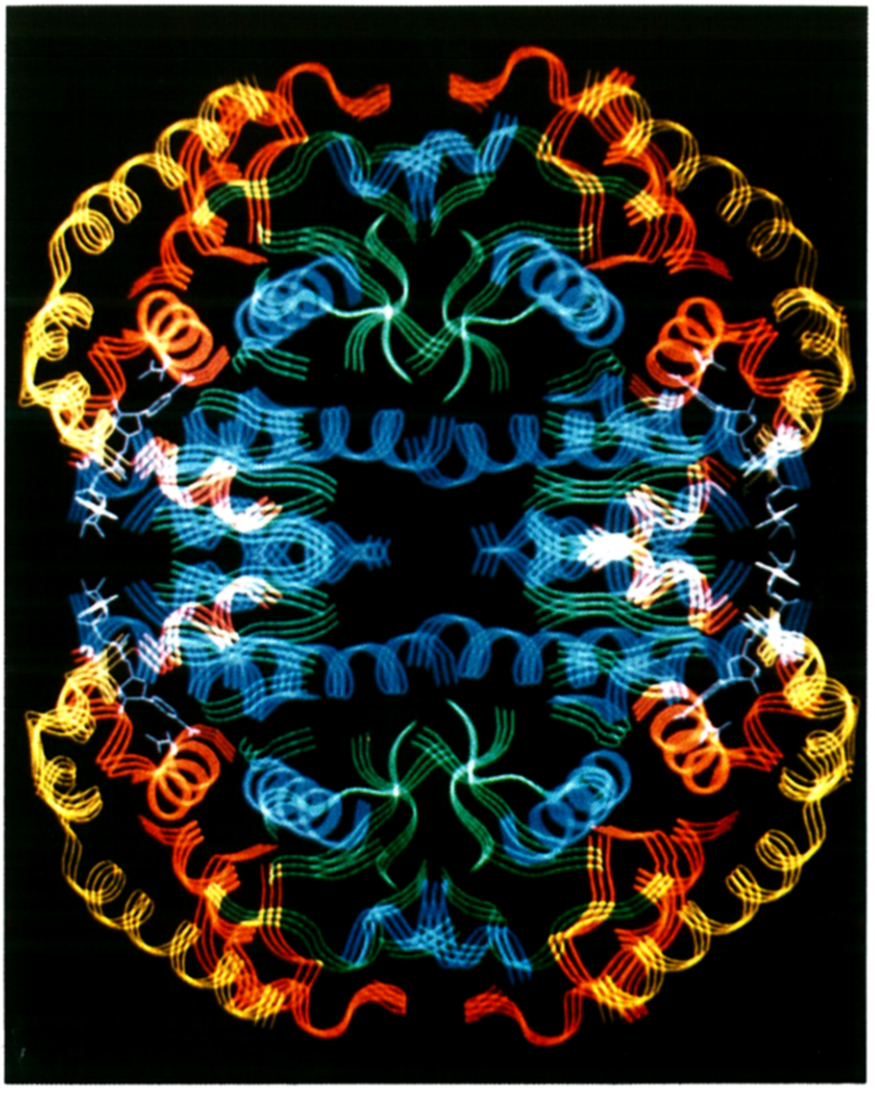


Figure 9. Division of tetramer into 4 concentric shells based on the degree-of-movement classification produced by the sieve-fit procedure. The holo conformation is shown. This graphic was generated with FRODO (Jones *et al.*, 1982). Blue, helices in the static core; green, sheets in the static core packed against the helices in the blue shell; red, helices and loops classified as minor movers packed against the sheets; yellow, the major movers; white, ligands.

set of mutations combining both the previous sets, k_{cat} was down by 7·8. Furthermore, the first set of mutations halved the rate of loop closure, while the others did not affect it. This relatively small reduction in the rate of loop closure is reasonable considering that one of the mutated residues was part of the first hinge and the other two were in the relatively rigid turn.

The motion of the minor mover βK-βL adds an additional degree of complexity to this account of propagation of conformational change. It does not make any contacts to the Loop in either the apo or holo forms, but it makes many VDW contacts and hydrogen bonds to the major mover αH as well as the minor movers α1F and βG-βH in both conformations. Consequently, as shown in Figure 7, ligand binding affects it twice indirectly, i.e. through the Loop and then (mostly) through αH. As shown in Table 3, many of the side-chains making contacts in this linkage are conserved, i.e. Ser316 (318), Ala317 (319), Leu320 (322) and Trp321 (323) on αH and Gly279 (280), Ile280 (281), Val284 (285), Phe285 (286) and Ile286 (287) on βK-βL. Note the many hydrophobic side-chains.

The only moving structure not connected to this network of contacts is βJ-α1G. It does not make contact with the Loop or the ligands and is connected to the other movers only tenuously. It has very low sequence conservation, and its average temperature factor, 32 Å^2 per atom in the apo form, is by far the highest in the protein, roughly four times that of the static core. As pointed out by Abad-Zapatero *et al.* (1987), it is probably poorly defined in the crystal structure and its apparent motion may be an artifact.

(c) *Concentric shells of increasing mobility*

A logical (though not necessarily temporal or causal) progression emerges in the picture of confor-

mational change. The coenzyme binds and the first hinge is free to deform with large twists of a few torsions. The oxamate binds, and the second hinge deforms by splitting and stretching. The Loop closes and, through its network of contacts, induces αH to move a lot and αE, α1F, and α1G-α2G to move to a smaller degree. βK-βL moves because of its contacts with αH, and the small loop βG-βH moves in response to its contacts with the ligands, the Loop and the other movers.

What about the parts of the protein that do not move? As shown in Figure 9, based on categorization of the sieve-fit procedure, the tetramer can be divided into four concentric shells. The central shell consists of helices in the static core. The next shell consists of the sheets in the static core that pack onto these helices. The third shell consists of the helices and loops classified as minor movers, which also pack on top of the sheets. And the fourth and outermost shell consists of the major movers. Almost all the quaternary structure contacts are made by residues in the innermost two shells.

The major and minor movers form an autonomous region on the surface of the tetramer (Figs 7 and 9). This autonomous region moves over a fixed central core, which is constrained by quaternary structure contacts. As discussed above, with the exception of αD, the moving helices act as rigid bodies, and at the interface between the autonomous region and the fixed core, their motion is accommodated by side-chains just slightly rocking (without large-scale repacking as in the Loop–αH interface). Their motion, consequently, can be described by the helix interface shear mechanism found for the helices in insulin and citrate synthnase (Chothia *et al.*, 1983; Lesk & Chothia, 1984).

9. Conclusion: a Comparison with TIM Loop Closure

Joseph *et al.* (1990) investigated the mechanism of loop closure in triose phosphate isomerase (TIM). With a view towards illuminating the essential features of loop closure, it is worth summarizing our results and comparing them to their findings on TIM.

Both the LDH and TIM loops have a double-hinged structure with a relatively rigid central section bracketed by two flexible hinges. The motion of each hinge in TIM can be approximated by changing two pseudo torsion angles (i.e. α angles) and allowing the rest of the molecule to move rigidly. In LDH, the first hinge motion can also be simply represented by changing two torsion angles. However, the second hinge has a more complex and unusual motion that involves the stretching and splitting of a helix into α-helical and 3_{10}-helical components. The difference between the two LDH hinge motions can be understood in terms of the different number of steric constraints on each hinge, many for the second hinge and few for the first. Probably because of the importance of constraints on LDH Loop motion, we found that the hinges in LDH were highly conserved, while Joseph *et al.* (1990) found the opposite for the hinges in TIM.

The sections of LDH and TIM loops that are not part of the hinges retain their conformation through closure, and these rigid sections are stabilized by internal hydrogen bonding. In TIM, a fourth hydrogen bond is added to the three present when the loop closes. In LDH, the hydrogen bonds can be divided into transverse ones, which are conserved during closure and also across LDH sequences, and longitudinal and helical ones, which rearrange during the conformational change.

The LDH Loop motion is integrated with many other smaller motions, while TIM loop motion is more localized. All the motion in LDH is confined to the tetramer surface and mediated by a network of contacts. In this network, the interface between the Loop and the helix packed against it, αH, is particularly significant, and it undergoes extensive side-chain repacking in the sense of "cogs hopping between grooves".

We thank A. M. Lesk for helpful conversations and use of computer programs, M. G. Rossmann for providing the refined LDH co-ordinates before they were available from the Protein Data Bank and useful comments on the manuscript, and A. Lenton for assistance with the figures. M.G. would like to acknowledge support from a Herchel-Smith Havard Fellowship.

References

Abad-Zapatero, C., Griffith, J. P., Sussman, J. L. & Rossmann, M. G. (1987). *J. Mol. Biol.* **198**, 445–467.

Adams, M. J., Ford, G. C., Koekoek, R., Lentz, P. J., Jr, McPherson, A., Jr, Rossmann, M. G., Smiley, I. E., Schevitz, I. E. & Wonacott, A. J. (1970). *Nature (London)*, **227**, 1098–1103.

Adams, M. J., Liljas, A. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 519–531.

Barstow, D. A., Clarke, A. R., Chia, W. N., Wigley, D., Sharman, A. F., Holbrook, J. J., Atkinson, T. & Minton, N. P. (1986). *Gene*, **46**, 47–55.

Barstow, D. A., Murphy, J. P., Sharman, A. F., Clarke, A. R., Holbrook, J. J. & Atkinson, T. (1987). *Eur. J. Biochem.* **165**, 581–586.

Bennett, W. S. & Huber, R. (1984). *Crit. Rev. Biochem.* **15**, 291–384.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Birktoft, J. J., Fernley, R. T., Bradshaw, R. A. & Banaszak, L. J. (1982). *Proc. Nat. Acad. Sci., U.S.A.* **79**, 6166–6170.

Chothia, C., Lesk, A. M., Dodson, G. G. & Hodgkin, D. C. (1983). *Nature (London)*, **302**, 500–505.

Clarke, A. R., Wigley, D. B., Chia, W. N., Barstow, D., Atkinson, T. & Holbrook, J. J. (1986). *Nature (London)*, **324**, 699–702.

Eklund, H., Samaha, J. P., Wallen, L., Branden, C. I., Akeson, A. & Jones, T. A. (1981). *J. Mol. Biol.* **146**, 561–587.

Faber, H. R. & Matthews, B. W. (1990). *Nature (London)*, **438**, 263–266.

Grau, U. M., Trommer, W. E. & Rossmann, M. G. (1981). *J. Mol. Biol.* **151**, 289–307.

Hediger, M. A., Frank, G. & Zuber, H. (1986). *Hoppe-Seyler's Biol. Chem.* **367**, 891–903.
Hensel, R., Mayr, U. & Yang, C. (1983). *Eur. J. Biochem.* **134**, 503–511.
Holbrook, J. J. & Gutfreund, H. (1973). *FEBS Letters*, **31**, 157–169.
Holbrook, J. J., Liljas, A., Steindel, S. J. & Rossmann, M. G. (1975). In *The Enzymes* (Boyer, P. D., ed.), 3rd edit., vol. 11, pp. 190–292, Academic Press, New York.
Jones, T. A. (1982). In *Computational Crystallography* (Sayre, D., ed.), pp. 303–317, Clarendon Press, Oxford.
Joseph, D., Petsko, G. A. & Karplus, M. (1990). *Science*, **249**, 1425–1428.
Kiltz, H. H., Keil, W., Griesbach, M., Petry, K. & Meyer, H. (1977). *Hoppe-Seyler's Z. Physiol. Chem.* **358**, 123–127.
Kunai, K., Machida, M., Matsuzawa, H. & Ohta, T. (1986). *Eur. J. Biochem.* **160**, 433–440.
Lee, B. K. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.
Lesk, A. M. (1986). In *Computer Applications in the Biosciences* (Saccone, C., ed.), pp. 23–28, EEC, Brussels.
Lesk, A. M. (1991). *Protein Architecture: a Practical Guide*, IRL Press, Oxford.
Lesk, A. M. & Chothia, C. (1984). *J. Mol. Biol.* **174**, 175–191.
Lesk, A. M. & Chothia, C. H. (1988). *Nature (London)*, **335**, 188–190.
Lesk, A. M. & Hardman, K. D. (1982). *Science*, **216**, 539–540.
Matrisian, L. M., Rautmann, G., Magun, B. E. & Breathnach, R. (1985). *Nucl. Acids Res.* **13**, 711–726.
Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H., Wlodawer, A. (1989). *Science*, **246**, 1149–1152.
Minowa, T., Iwata, S., Sakai, H., Masaki, H. & Ohta, T. (1989). *Gene*, **85**, 161–168.
Musick, W. D. L. & Rossmann, M. G. (1979). *J. Biol. Chem.* **254**, 7611–7620.
Pan, Y. C. E., Sharief, F. S., Okabe, M., Huang, S. & Li, S. S. L. (1983). *J. Biol. Chem.* **258**, 7005–7016.
Parker, D. M., Jeckel, D. & Holbrook, J. J. (1982). *Biochem. J.* **201**, 465–471.
Perutz, M. (1989). *Quart. Rev. Biophys.* **22**, 139–236.
Phillips, D. C., Rivers, P. S., Sternberg, M. J. E., Thornton, J. M. & Wilson, I. A. (1977). *Trans. Biochem. Soc.* **5**, 642–647.
Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Lentz, P. J., Jr, McPherson, A., Jr, Schevitz, R. W. & Smiley, I. E. (1971). *Cold Spring Harbor Symp. Quant. Biol.* **36**, 179–191.
Rossmann, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. (1975). In *The Enzymes* (Boyer, P. D., ed.), 3rd edit., vol. 11, pp. 61–102, Academic Press, New York.
Sakai, I., Sharief, F. S. & Li, S. S. L. (1987). *Biochem. J.* **242**, 619–622.
Takeno, T. & Li, S. S. L. (1989). *Biochem. J.* **257**, 921–924.
Taylor, S. S. (1977). *J. Biol. Chem.* **252**, 1799–1806.
Torff, H. J., Becker, D. & Schwarzwalder, J. (1977). In *Pyridine Nucleotide Dependent Dehydrogenases* (Sund, H., ed.), pp. 31–42, Walter de Gruyter, Berlin.
Tramantano, A., Chothia, C. H. & Lesk, A. M. (1989). *Proteins*, **6**, 382–394.
Tsujibo, H., Tiano, H. F. & Li, S. S. L. (1985). *Eur. J. Biochem.* **147**, 9–15.
Waldman, A. D. B., Hart, K. W., Clarke, A. R., Wigley, D. B., Barstow, D. A., Atkinson, T., Chia, W. N. & Holbrook, J. J. (1988). *Biochem. Biophys. Res. Commun.* **150**, 752–759.
White, J., Hackert, M. L., Buehner, M., Adams, M. J., Ford, G. C., Lentz, P. J., Jr, Smiley, I. E., Steindel, S. J. & Rossmann, M. G. (1976). *J. Mol. Biol.* **102**, 759–779.
Widerkehr, F. (1982). Ph.D. thesis, ETH Zurich.
Wilks, H. M., Hart, K. W., Feeney, R., Dune, C. R., Muirhead, H., Chia, W. N., Barstow, D., Atkinson, T., Clarke, A. R. & Holbrook, J. J. (1988). *Science*, **242**, 1541–1544.
Wilks, H. M., Halsall, D. J., Atkinson, T., Chia, W. N., Clarke, A. R. & Holbrook, J. J. (1990). *Biochemistry*, **29**, 8587–8591.

Edited by B. W. Matthews