Volume 326, number 1,2,3, 4–10 FEBS 12673 July 1993

Minireview

Movable lobes and flexible loops in proteins

Structural deformations that control biochemical activity

E.S. Kempner

Laboratory of Physical Biology National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Received 19 April 1993; revised version received 19 May 1993

Two classes of protein whose structure is modified by small ligands are reviewed. Proteins of one group contain two massive domains joined by a flexible link; in response to small molecules, the two lobes approach and enclose the ligand. In the other, a short segment of amino acids moves as a flexible loop over the ligand which often is trapped in a non-aqueous environment. Biochemical reaction rates are altered dramatically by these movements.

Enzyme; Reaction rate; Structure–function; Domain

1. INTRODUCTION

The conformation of a protein is defined by the three-dimensional folding of a polypeptide chain and, in oligomers, the orientation of one chain with respect to another. There has been considerable speculation about the effect of changes in conformation on the biochemical activity expressed by these molecules.

Detailed three-dimensional structures of many proteins have been established by X-ray diffraction of crystals, and for small proteins in solution by ≥ 2D NMR. In many of these studies, the structure was also determined for the protein-ligand complex (e.g. [1–6]). The results range all the way from proteins which are not appreciably altered by the binding of ligands, to those which show many complex position shifts – multiple structural changes unique to each protein.

In many proteins, binding of ligands at one site alters the structure of other regions of a polypeptide. Two general patterns have emerged. Members of one group show large-scale mass movement: two large domains on a single polypeptide move in relation to each other about a central flexible region. In the other class, binding (usually of a substrate) at the active site results in movement of a relatively small loop elsewhere on the polypeptide: this movement encloses the ligand in a cage-like structure. Some examples of both types are described below.

Correspondence address: E.S. Kempner, Bldg. 6, Room 120, NIH, Bethesda, MD 20892, USA. Fax: (1) (301) 402 0009.

2. MOVABLE LOBES

This group of proteins is composed of molecules with two major domains, as in calmodulin. The crystal structure of the calcium-calmodulin complex has been described, showing a central not-too-rigid helix. The two globular regions at the amino- and carboxy-termini of calmodulin can move and each can bind to myosin light chain kinase (MLCK). This further distorts the quasi-helical hinge region, but the hinge amino acid residues themselves do not interact with MLCK [1]. This process does not seem to be a directed movement in response to MLCK binding. It is envisaged that MLCK first binds to one lobe; subsequent movement of the other lobe brings it near to MLCK, permitting additional interactions which stabilize the structure.

Similar conformational alterations occur in several proteins (Table I). All involve movement of two massive domains, commonly called 'lobes', about a flexible connecting region. In tomato bush stunt virus two large domains are connected by a flexible hinge which permits rotation of approximately 20° [7]. In enzymes, the binding of a small molecule such as a substrate can cause movement of the two major lobes towards each other, enclosing the substrate-binding region. This movement occurs about a 'hinge region' of the polypeptide. Some examples include hexokinase [8], alanine racemase [9], bacterial neutral proteases [10], phosphoglycerate kinase [11] and cAMP-dependent protein kinase [12]. In another member of this group, maltodextrin binding protein, the two lobes shifted position by 35° on binding

 Published by Elsevier Science Publishers B.V.

of ligand [2] as shown in Fig. 1. The general pattern of this 'lobed' protein class is that substrate binding to one region of a polypeptide results in movement of two massive protein domains about a 'hinge' region (which is sometimes rich in glycine); the hinge amino acids do not interact with the substrate molecule. These proteins are clear examples of Koshland's 'induced fit' model [13].

3. FLEXIBLE LOOPS

Another group of proteins shows limited changes in polypeptide folding, involving only small stretches of amino acid residues. These restricted conformational changes were associated with dramatic changes in biochemical activity. There is a wide variety of biochemical systems in which similar subtle structural effects have been observed (Table I). The clearest example of this phenomenon is found in the X-ray diffraction studies of streptavidin [3].

Streptavidin is composed of four identical 15 kDa polypeptide chains. The X-ray diffraction patterns from streptavidin crystals were compared to those from crystals of streptavidin–biotin complexes. The overall structure of the streptavidin tetramer was similar with and without biotin, except that two disordered loops became ordered. One of these polypeptide loops changed posi-

Table I

Structural modifiers

A. Lobes:	Effector	Conformational change
Calmodulin	Target peptide	Hinge region (9 residues); 2 large lobes move towards each other
Hexokinase	Substrate (glucose)	Hinge region moves two large lobes towards each other
Phosphoglycerate kinase	MgATP and 3-phosphoglycerate	Hinge region moves two large lobes together
Alanine racemase	Substrate (alanine)	Hinge region (5 residues) folds two domains into a compact structure
Tomato bushy stunt virus	–	Flexible hinge
Maltodextrin binding protein (MBP)	Maltose	Rigid-body 'hinge-bending'; 2 large lobes move by 35°
cAMP-dependent protein kinase	Inhibitor peptide	Large and small lobes move 39° around a hinge

B. Loops:	Effector	Conformational change	Flexible region
Streptavidin	Biotin	Flexible loop	6 residues
Glutathione synthase	Either ATP or γ-Glu-cys	Flexible loop	16 residues
Triosephosphate isomerase	Substrate (dihydroxy-acetone phosphate) or inhibitors	Hinged lid Flexible loop	10 residues
Tryptophan synthase	Substrate (Glyceraldehyde 3-phosphate)	Flexible loop	10 residues
Lipases	Oil–water interface or *N*-hexylchloro-phosphonate ethyl ester	Helical lid	3(hinge)-7(rigid lid)-4(hinge)
Glucose permease		Mobile loop	11 residues
Phosphoglycerate mutase	Phosphoglycerate	Flexible carboxyl terminus	10 residues
Adenylate kinase	ATP	Flexible loop	7 residues
Lactate dehydrogenase	Substrate (pyruvate)	Loop	13 residues
Dihydrofolate reductase	Substrate	Loop	15 residues
Rubisco	Substrate?	Loop	8 residues
HIV-1 aspartic protease	Substrate-derived inhibitor	2 loops	15 residues

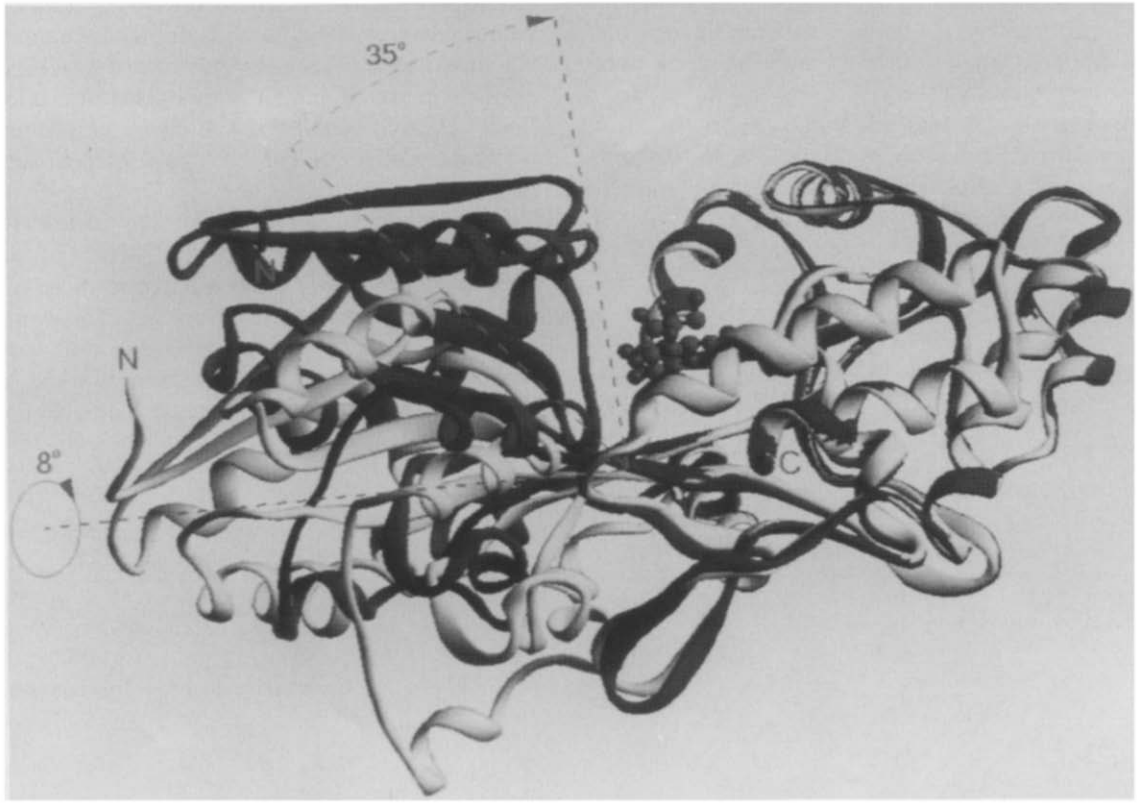

Fig. 1. Changes in the structure of maltodextrin binding protein on binding of maltose [7]. The two movable lobes shift their positions by 35°. Reprinted with permission from Biochemistry 31, 10657–10663; copyright 1992 American Chemical Society.

tion (Fig. 2): the movement of this 6-residue loop (Table I) brings it to a position near the biotin to which it forms a hydrogen bond. The polypeptide loop effectively shields almost all of the biotin molecule (save for a carboxyl group) from the aqueous environment. The biotin moiety is effectively confined in the modified avidin structure and is not able to move out into the bulk aqueous phase. Similarly, extraneous biotin molecules are not able to enter this structure. In addition to the noncovalent binding, these detailed molecular structures revealed a new steric hindrance which prevents removal of the bound biotin.

A surprisingly similar model was reported for the enzyme glutathione synthetase [4]. A sixteen amino-acid region of the polypeptide is mobile compared to the rest of the protein. When the substrates (ATP and γ-glutamylcysteine) bind, this flexible loop closes the active site. Unstable intermediates formed during the enzymatic reactions are protected from decomposition by solvent water.

A disordered 11-residue stretch of the triosephosphate isomerase polypeptide extends into the solvent. On binding of substrate, this loop moves as a rigid body on two hinges and closes over the active site. The lid makes several hydrogen bonds with other residues of the polypeptide, but only one with the substrate [5]. The presence of this loop protects the phosphorylated intermediate from decomposition [14].

In the α-subunit of tryptophan synthase there is a ten-residue stretch which is disordered in the crystal. This loop effects the binding of the substrate (indole glycerol phosphate) and in addition communicates 'the effects of ligand binding from the α subunit to the β subunit in the $\alpha_2\beta_2$ complex' [15].

In fungal lipase there is a similar phenomenon: a rigid lid made of a helical 7-residue stretch flanked by 3-residue and 4-residue hinge regions. Movement of the lid 'exposes catalytic groups and creates a nonpolar surface which stabilizes contact between the enzyme and the lipid surface' [6]. In the human pancreatic lipase [16] the hydrolytic site is covered by a surface loop and is inaccessible to the solvent. Activation of the enzyme involves a reorientation of this flap. The surface loop lies between disulfide-bridged residues 237 and 261, covering the active site with a short one-turn α-helix [16]. The human LPL similarly contains a loop (residues 217–238) bounded by two conserved cysteine residues; it can be replaced with a similar loop from hepatic triglyceride lipase and still shows activity. Although these two loops differ in primary sequence, they both show a similar predicted (by the Chou–Fasman algorithm) alternation of α-helical and β-sheet regions.

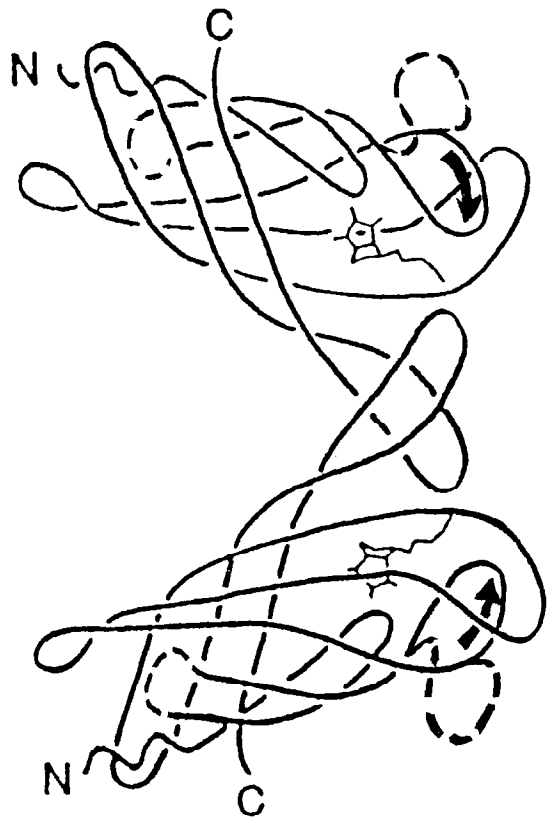

Fig. 2. The structure of streptavidin (modified from [9]). Two 15 kDa subunits are shown; two small (dashed) portions of the polypeptide appear disordered. On binding of biotin, one of these (residues 45 to 50) changes position. It entraps the biotin molecule and excludes it from contact with the aqueous solvent. Reprinted with permission from 'Structural Origins of High-Affinity Biotin Binding to Streptavidin' Science 243, 85–88, Copyright 1989 by AAAS.

However, replacement of the human LPL loop with the loop from pancreatic lipase yields an inactive human LPL [17].

The binding of substrate (pyruvate) to lactate dehydrogenase causes a conformational change due to movement of a 13-residue polypeptide folding over the active center pocket. Smaller conformational changes are 'associated with the large movement of the loop and the different position and conformation of the coenzyme' [18]. Genetic engineering techniques were used to change the length of the loop. Mutant enzymes with three-residue-shorter loops still showed tight binding of substrate, but catalytic activity was greatly reduced since the substrate was no longer isolated from bulk water. Longer loops (by 4 residues) decreased the activity over a million-fold. Interestingly, activity with phenylpyruvate as a substrate was unaffected [19].

There are conformational changes in rabbit muscle adenylate kinase on binding of metal-ATP substrate. The largest displacements involve a 7-residue glycine-rich loop which approaches the effector molecule, but makes little or no contact with it [20].

A hairpin turn in the dihydrofolate reductase polypeptide closes down over both $NADP^+$ and folate in the ternary complex. In addition, this molecule also displays some movable lobe properties associated with the adenosine-binding domain [21].

Activation of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) involves both small domain movements and a large movement of the active site loop on the large subunit [22].

Glucose permease contains two mobile loop regions: a 5-residue loop and also a 11-residue loop which is close to the active site. In addition, there is a very mobile 13-residue segment at the amino terminus [23].

Phosphoglycerate mutase, a tetramer, also shows a movable region but at the carboxy terminus of each polypeptide. This 10-residue stretch forms a flexible tail which has no well-defined conformation in crystals of the dephosphorylated enzyme [24]. Models indicate that this tail 'can adopt a conformation which will hinder access to the active site' which lies at the bottom of a deep hollow formed entirely by the residues of one subunit. They also suggest that this tail could then exclude water from the active site, 'ensuring that the phosphoryl group on the enzyme is transferred to the substrate rather than to a water molecule'.

Movable 'flaps' are reported in the aspartic proteinases. The retroviral HIV-1 PR is a homodimer; a substrate-derived inhibitor is bound in a groove across the dimer interface. Identical polypeptide loops from each subunit of the dimer move by more than 7 Å and completely wrap around the inhibitor. This embrace desolvates the central region of the inhibitor [25,26]. In the eukaryotic aspartic proteinase, the dimeric structure differs and only a single flap (of unrelated amino acid sequence) covers the inhibitor [26].

Most of these loops are small; only a few amino acid residues are involved. Only a small amount of energy would be required to shift their position. Yet their effect on reaction rates is considerable. The wide range of proteins displaying these flexible loops is mirrored in the diversity of amino acid sequences in these loops (Table II). No pattern is evident among these sequences.

Throughout these various examples, a small stretch of polypeptide changes its position relative to the rest of the polymer. Frequently, the X-ray or 2D NMR data indicate that in the absence of the small molecule effectors the loop region is not sharply defined, i.e. the lid position is variable. In general, the small molecule effectors 'lock' the loop in one position. This can result in sterically isolating a small molecule (such as a biochemical intermediate or a ligand) from the bulk phase, and this has been shown to alter reaction rates and pathways. In addition, as a loop closes down over a small molecule, it may also expose another polypeptide region which is potentially the active site for other reactions. The ill-defined position of many of these loops suggests a somewhat restricted access to the hypothetical second

Table II

Amino acid sequence of hinges and loops

Amino acid sequence of hinge region between Lobes:

Calmodulin

74	75	76	77	78	79	80	81	82
Arg	Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu
R	K	M	K	D	T	D	S	E

Alanine racemase

253	254	255	256	257
Tyr	Gly	Gly	Gly	Tyr
Y	G	G	G	Y

Tomato bushy stunt virus

269	270	271	272	273
Thr	Asn	Thr	Leu	Leu
T	N	T	L	L

Amino acid sequence of flexible Loops:

Streptavidin

45	46	47	48	49	50
Ser	Ala	Val	Gly	Asn	Ala
S	A	V	G	N	A

Glutathione Synthetase

226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242
Ile	Pro	Gln	Gly	Gly	Glu	Thr	Arg	Gly	Asn	Leu	Ala	Ala	Gly	Gly	Arg	Gly
I	P	Q	G	G	E	T	R	G	N	L	A	A	G	G	R	G

Triosephosphate isomerase

-------HINGE------- **********LID**********---------HINGE----------

166	167	168	169	170	171	172	173	174	175	176	177
Pro	Val	Trp	Ala	Ile	Gly	Thr	Gly	Lys	Thr	Ala	Thr
P	V	Y	A	I	G	T	G	K	T	A	T

Tryptophan synthetase (in α subunit)

178	179	180	181	182	183	184	185	186	187	188	189	190	191	192
Ser	Arg	Ser	Gly	Val	Thr	Gly	Ala	Glu	Asn	Arg	Gly	Ala	Leu	Pro
S	R	S	G	V	T	G	A	E	N	R	G	A	L	P

Adenylate kinase (rabbit muscle)

15	16	17	18	19	20	21
Gly	Gly	Pro	Gly	Ser	Gly	Lys
G	G	P	G	S	G	K

Phosphoglycerate mutase carboxy terminal flexible tail

232	233	234	235	236	237	238	239	240	241
Gly	Ala	Ala	Val	Ala	Asn	Gln	Lys	Lys	Gly
G	A	A	V	A	N	Q	K	K	G

Lactate dehydrogenase (there is no residue # 104)

97	98	99	100	101	102	103	105	106	107	108	109	110
Cys	Ala	Gly	Ala	Asn	Gln	Lys	Pro	Gly	Glu	Thr	Arg	Leu
C	A	G	A	N	Q	K	P	G	E	T	R	L

Fungal lipase

--------Hinge-------- **********Helix********** ----------Hinge----------

82	83	84	85	86	87	88	89	90	91	92	93	94	95
Ser	Ser	Ser	Leu	Arg	Asn	Trp	Ile	Ala	Asp	Leu	Thre	Phe	Val
S	S	S	L	R	N	W	I	A	D	L	T	F	V

Table II *(continued)*

Amino acid sequence of hinges and loops

Amino acid sequence of hinge region between Lobes:																				
Human pancreatic lipase																				
237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257
Lys	Lys	Asn	Ile	Leu	Ser	Gln	Ile	Val	Asp	Ile	Asp	Gly	Ile	Trp	Glu	Gly	Thr	Arg	Asp	Phe
K	K	N	I	L	S	Q	I	V	D	I	D	G	I	W	E	G	T	R	D	F
258	259	260	261																	
Asp	Asp	Cys	Asn																	
A	A	C	N																	
Dihydrofolate reductase																				
9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24					
Ala	Val	Asp	Arg	Val	Ile	Gly	Met	Glu	Asn	Ala	Met	Pro	Trp	Asn	Leu					
A	V	D	R	V	I	G	M	E	N	A	M	P	W	N	L					
Rubisco																				
331	332	333	334	335	336	337	338													
Val	Val	Gly	Lys	Leu	Glu	Gly	Glu													
V	V	G	K	L	E	G	E													
HIV-1 Aspartic proteinase																				
44	45	46	47	48	49	50	51	52	53	54	55	56	57	58						
Pro	Lys	Met	Ile	Gly	Gly	Ile	Gly	Gly	Phe	Ile	Lys	Val	Arg	Gln						
P	K	M	I	G	G	I	G	G	F	I	K	V	R	Q						

active sites; in the case of an enzyme, this would imply a limited reaction rate (i.e. in the 'basal' state). The ligand of the primary reaction which causes a shift in position of the flexible loop would then appear to be an 'activator' for the putative other reaction. In this model, the activator has no chemical relationship to the active site or the substrate of the second reaction.

4. GENERAL PRINCIPLES

Lobes and loops in proteins often move in response to small molecules, covering small surface areas of the protein; ligands are trapped and solvent molecules are excluded. Subunit association, just like 'flexible loops' and 'movable lobes' can cover (or uncover) reactive surfaces. Conformational changes in a loop of the α-subunit of tryptophan synthase also affect the β-subunits (see above). Subunits with identical amino acid sequences in enzymatically-different proteins may reflect the same general principles. Uracil DNA glycosylase has the same sequence as glyceraldehyde 3-phosphate dehydrogenase (GAPDH). GAPDH has no glycosylase activity, but on denaturation of the tetrameric dehydrogenase to monomers, glycosylase activity appears [28]. Another identical amino acid sequence is found both in *S*-adenosyl-L-methionine:mRNA (nucleoside-$O^{2'}$-)-methyltransferase and in VP39, a subunit of poly(A) polymerase [29]. The polymerase is a heterodimer also containing a catalytic subunit, VP55; VP39 accelerates the reaction catalyzed by VP55. In this case, the polymerase dimer retains the methyltransferase activity seen in monomeric VP39.

Meyer-Siegler et al. [28] speculated that oligomerization might be a deliberate mechanism for altering biochemical activities. The appearance of new activities after oligomerization might be due to the proximity of complementary regions on two subunits, but could just as well be due to the uncovering of new regions by movement of a loop or arm away from the original monomer and towards its new molecular mate. Cyclic nucleotide phosphodiesterase from brain is an inactive dimer which is converted by calmodulin to two functional monomers [30]. In view of the calmodulin-MLCK studies described above, the association of calmodulin with phosphodiesterase may open the enzyme dimer.

Movable lobes and flexible loops are found in many proteins. These deformations can have profound biochemical consequences, such as permitting reactions which would otherwise be prevented; oligomerization may be a related mechanism to attain the same end. Since the movement of lobes and loops is often initiated by binding of small molecules, the general concept of allosterism underlies all of these structural modifications. The structural deformations reviewed here reveal two alternative patterns of conformational changes which could explain allosteric effects in molecular detail.

REFERENCES

[1] Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. and Bax, A. (1992) Science 256, 632–638.
[2] Sharff, A.J., Rodseth, L.E., Spurlino, J.C. and Quiocho, F.A. (1992) Biochemistry 31, 10657–10663.

[3] Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J. and Salemme, F.R. (1989) Science 243, 85–88
[4] Tanaka, T., Kato, H., Nishioka, T. and Oda, J. (1992) Biochemistry 31, 2259–2265
[5] Joseph, D., Petsko, G.A. and Karplus, M. (1990) Science 249, 1425–1428.
[6] Brzozowski, A.M., Derewenda, U., Derewenda, Z.S., Dodson, G.G., Lawson, D.M., Turkenburg, J.P., Bjorkling, F., Huge-Jensen, B., Patkar, S.A. and Thim, L. (1991) Nature 351, 491–494.
[7] Winkler, F.K., Schutt, C.E., Harrison, S.C. and Bricogne, G. (1977) Nature 265, 509–513.
[8] Bennett, W.S., Jr. and Steitz, T.A. (1978) Proc. Natl. Acad. Sci. USA 75, 4848–4852.
[9] Galaktos, N.G. and Walsh, C.T. (1989) Biochemistry 28, 8167–8174.
[10] Holland, D.R., Tronrud, D.E., Pley, H.W., Flaherty, K. M., Stark, W., Jansonius, J.N., McKay, D.B. and Matthews, B.W. (1992) Biochemistry 31, 11310–11316.
[11] Pickover, C.A., McKay, D.B., Engelman, D.M. and Steitz, T.A. (1979) J. Biol. Chem. 254, 11323–11329.
[12] Olah, G.A., Mitchell, R.D., Sosnick, T.R., Walsh, D.A. and Trewhella, J. (1993) Biochemistry 32, 3649–3657.
[13] Koshland, D.E., Jr. (1958) Proc. Natl. Acad. Sci. USA 44, 98–104.
[14] Pompliano, D.L., Peyman, A. and Knowles, J.R. (1990) Biochemistry 29, 3186–3194.
[15] Miles, E.W. (1991) J. Biol. Chem. 266, 10715–10718.
[16] Winkler, F.K., D'Arcy, A. and Hunziker, W. (1990) Nature 343, 771 774.
[17] Faustinella, F., Smith, L.C. and Chan, L. (1992) Biochemistry 31, 7219–7223 .
[18] Adams, M.J., Buehner, M., Chandrasekhar, K., Ford, G.C., Hackert, M.L., Liljas, A., Rossmann, G., Smiley, I.E., Allison, W.S., Everse, J., Kaplan, N.O. and Taylor, S.S. (1973) Proc. Natl. Acad. Sci. USA 70, 1968–1972.
[19] Wilks, H.M., Moreton, K.M., Halsall, D.J., Hart, K.W., Sessions, R.D., Clarke, A.R. and Holbrook, J.J. (1992) Biochemistry 31, 7802–7806.
[20] Fry, D.C., Kuby, S.A. and Mildvan, A.S. (1986) Proc. Natl. Acad. Sci. USA 83, 907–911.
[21] Bystroff, C. and Kraut, J. (1991) Biochemistry 30, 2227–2239.
[22] Schreuder, H.A., Knight, S., Curmi, P.M.G., Andersson, I., Cascio, D., Branden, C.-I. and Eisenberg, D. (1993) Proc. Natl. Acad. Sci. USA, in press.
[23] Stone, M.J., Fairbrother, W.J., Palmer, A.G. III, Reizer, J., Saier, M.H. Jr. and Wright, P.E. (1992) Biochemistry 31, 4394–4406.
[24] Winn, S.I., Watson, H.C., Harkins, R.N. and Fothergill, L.A. (1981) Phil. Trans. Roy. Soc. Lond. B 293, 121–130.
[25] Jaskolski, M., Tomasselli, A.G., Sawyer, T.K., Staples, D.G., Heinrikson, R.L., Schneider, J., Kent, S.B.H. and Wlodawer, A. (1991) Biochemistry 30, 1600–1609.
[26] Rao, J.K.M., Erickson, J.W. and Wlodawer, A. (1991) Biochemistry 30, 4663–4671.
[27] Hopper, P., Harrison, S.C. and Sauer, R.T. (1984) J. Mol. Biol. 177, 701–713.
[28] Meyer-Siegler, K., Mauro, D.J., Seal, G., Wurzer, J., de Riel, J.K. and Sirover, M.A. (1991) Proc. Natl. Acad. Sci. USA 88, 8460–8464.
[29] Schnierle, B.S., Gershon, P.D. and Moss, B. (1992) Proc. Natl. Acad. Sci. USA 89, 2897–2901.
[30] Kincaid, R.L., Kempner, E.S., Manganiello, V.C., Osborne, J.C. Jr. and Vaughn, M. (1981) J. Biol. Chem. 256, 11351–11355.