

## Proteins with Similar Architecture Exhibit Similar Large-Scale Dynamic Behavior

O. Keskin,\* R. L. Jernigan,<sup>†</sup> and I. Bahar\*<sup>†</sup>

\*Chemical Engineering Department and Polymer Research Center, Bogazici University, and TUBITAK Advanced Polymeric Materials Research Center, Bebek 80815, Istanbul, Turkey, and <sup>†</sup>Molecular Structure Section, Laboratory of Experimental and Computational Biology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892-5677 USA

**ABSTRACT** We have investigated the similarities and differences in the computed dynamic fluctuations exhibited by six members of a protein fold family with a coarse-grained Gaussian network model. Specifically, we consider the cofactor binding fragment of CysB; the lysine/arginine/ornithine-binding protein (LAO); the enzyme porphobilinogen deaminase (PBGD); the ribose-binding protein (RBP); the N-terminal lobe of ovotransferrin in apo-form (apo-OVOT); and the leucine/isoleucine/valine-binding protein (LIVBP). All have domains that resemble a Rossmann fold, but there are also some significant differences. Results indicate that similar global dynamic behavior is preserved for the members of a fold family, and that differences usually occur in regions only where specific function is localized. The present work is a computational demonstration that the scaffold of a protein fold may be utilized for diverse purposes. LAO requires a bound ligand before it conforms to the large-scale fluctuation behavior of the three other members of the family, CysB, PBGD, and RBP, all of which contain a substrate (cofactor) at the active site cleft. The dynamics of the ligand-free enzymes LIVBP and apo-OVOT, on the other hand, concur with that of unliganded LAO. The present results suggest that it is possible to construct structure alignments based on dynamic fluctuation behavior.

### INTRODUCTION

Understanding the relationship between molecular structure and biological function is of utmost importance for protein design and engineering. Despite the rapidly increasing number of x-ray-elucidated three-dimensional structures and advances in techniques for probing or controlling the function and dynamics of proteins, relatively few connections between function and structure have been established, except on the local scale, e.g., enzyme active sites. Based upon the small number of known protein structural families, it appears that the large number of functions required is being achieved by decorating a given structure in different ways with residues requisite for its specific function. Thus structural determinations reveal the existence of common (or closely similar) folds in proteins that are involved in different biological activities. This diversity in function is generally accompanied by changes in sequence, whereas the structure can be preserved on either a local (structural motif) or a global (domain) scale. Like the common binding/recognition properties of structural motifs, common mechanisms of action for globally similar proteins might be expected. A challenging issue is to explore the similarities and differences in the global dynamics manifested by structurally similar but sequentially and functionally distinct proteins. This is the subject of the present work.

Despite the lack of similarity in primary sequence, a number of substrate-binding proteins (SBPs) exhibit similar three-dimensional structures. The examples considered here are 1) the cofactor binding fragment of CysB (Tyrrell et al., 1997), a member of the LysR family of transcriptional regulatory proteins; 2) the substrate-specific receptor of permeases referred to as lysine/arginine/ornithine-binding protein (LAO) (Kang et al., 1991; Oh et al., 1993), a member of the periplasmic SBPs (PBP) family; 3) the enzyme porphobilinogen deaminase (PBGD) (Louie et al., 1992), a member of the Lac repressor (LacR) family; 4) the ribose-binding protein (RBP), an L-arabinose-binding protein; 5) the apo-form of the N-terminal lobe of ovotransferrin (an iron-binding protein), (apo-OVOT); and 6) the leucine/isoleucine/valine-binding protein (LIVBP), another PBP. All have 1) two globular domains (at least) with a substrate-binding site located in the cleft between the two domains, 2) two or three peptide segments connecting the two lobes, and 3) a common  $\alpha/\beta$  architecture of the domains, each comprising a  $\beta$ -sheet core, flanked on both sides by  $\alpha$ -helices, similar to a Rossmann fold.

These proteins undergo a conformational change from the “open” to the “closed” form upon substrate binding. They have essentially two functions: binding to substrate and another subsequent specific biological activity. These functions could be assayed separately, by mutation experiments in different parts of the protein.

The dynamic characteristics of these proteins will be analyzed here, using the so-called Gaussian network model (GNM) of proteins (Bahar et al., 1997; Haliloglu et al., 1997). The GNM has proved itself in numerous applications to be a simple yet useful tool for the efficient investigation of the collective dynamics of proteins and of biomolecular

---

Received for publication 29 June 1999 and in final form 16 December 2000.

Address reprint requests to Dr. Robert Jernigan, Laboratory of Experimental and Computational Biology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, MSC 5677, Rm. B-116, Bldg. 12B, Bethesda, MD 20892-5677. Tel.: 301-496-4783; Fax: 301-402-4724; E-mail: robert\_jernigan@nih.gov.

© 2000 by the Biophysical Society

0006-3495/00/04/2093/14 \$2.00

complexes (Bahar et al., 1998a,b, 1999; Bahar and Jernigan, 1998, 1999; Demirel et al., 1998; Jernigan et al., 1999). The intricate coupling of all structural elements is considered therein at the level of individual residues, with the aid of a Kirchhoff matrix (Flory, 1976) of interresidue contacts. The decomposition of the total fluctuations serves to identify the different individual modes of motion, ranging from the most cooperative ones involving the overall structure, to the most localized ones, manifested as high-frequency fluctuations of individual residues.

## PROTEINS

Table 1 lists the family, the Brookhaven Protein Databank (PDB) codes and resolution of the examined crystal structures, the residue intervals forming the different domains, and the cross-over regions between the domains. Further details about these proteins follow.

CysB has four identical subunits. It controls the expression of genes associated with the biosynthesis of cysteine in bacteria. It is both an activator and repressor of transcription. It binds to a number of *cys* promoters, and in the presence of *N*-acetylserine, it activates transcription, while it acts as a repressor for transcription of its own gene. In the crystal structure examined here—a fragment of CysB in a biologically active dimeric form (Tyrrell et al., 1997)—each monomer consists of the C-terminal residues 88–324, hence the notation Cys(88–324). The monomers are folded each as two similar  $\alpha/\beta$  domains, I and II, connected by two cross-over regions (Table 1). There is a cavity between the two domains, which contains a sulfate anion. There is also an extended region (residues 291–324) appended to domain I, comprising three helices and one  $\beta$ -sheet. The ribbon diagram of CysB(88–324) is shown in Fig. 1. In Fig. 1 *a* the dimeric form is displayed, and in Fig. 1 *b* the monomeric form. We note the large surface buried at the interface between the monomers; the dimerization geometry differs from the common subunit rearrangement of the structurally similar cofactor-binding domains of LacR family proteins.

LAO is a periplasmic substrate-binding protein that transports a wide variety of substrates such as sugars, vitamins, and inorganic ions (Kang et al., 1991) (Fig. 1 *c*). It captures its ligands by a large relative movement between the two domains, leading to sequestration into the buried cavity. Each domain consists of one  $\beta$ -sheet and four  $\alpha$ -helices. Both the liganded and unliganded crystal structures of LAO have been determined by x-ray crystallography, and the motions of both will be analyzed here. In the former case, a lysine is bound in the cleft between the two domains, and the molecule assumes a closed form (Kang et al., 1991), similar to the structure of CysB(88–324). In the unliganded form, on the other hand, an open conformation is taken (Oh et al., 1993). The passage between open and closed forms is attributed to the hinge role of residues at the peptides connecting the two domains.

PBGD is a key enzyme in the biosynthetic pathway of tetrapyrroles. It is folded into three domains of roughly equal size (Fig. 1 *d*). The crystal structure of PBGD (Louie et al., 1992) contains a cofactor dipyrromethane (DPM), covalently bonded to the conserved Cys<sup>242</sup> in domain III. The cofactor lies in the deep cleft between domains I and II, thus bridging the three domains. The opening of the cleft—which is also the catalytic site of the enzyme—has been suggested (Louie et al., 1992) to involve a hinge-bending motion of domains I and II, similar to that of other SBPs, accompanied by the movement of domain III away from the interdomain interface, so as to reposition the growing polypyrrole chain (product of polymerization) and accommodate the incoming porphobilinogen substrate (reactant).

RBP is a periplasmic binding protein of Gram-negative bacteria. It is the receptor for both chemotaxis and transport of sugars and small peptides. The protein consists of two highly similar structural domains, each of which is composed of a core of a six-stranded  $\beta$ -sheet flanked on both sides by  $\alpha$ -helices (Fig. 1 *e*) (Mowbray and Cole, 1992; Binnie et al., 1992). Two short stretches of amino acids connect the two domains (Table 1). The two domains are related to each other by an almost perfect twofold axis of

**TABLE 1** Proteins analyzed in the present study

Protein	CysB(88-324)	LAO	PBGD	RBP	OVOT	LIVBP
Family	LysR	PBP	LacR	L-Arabinose BP	Iron BP	PBP
PDB codes	1al3	1l1t, 2lao*	1pda	2dri	1tfa	2liv
Resolution (Å)	1.8	2.7	1.9	1.6	1.9	2.4
Domain I residues	88–162, 270–324 <sup>†</sup>	1–89, 193–238	4–99, 200–217	1–103, 236– 264	4–82, 251–316	1–118, 253–326
Domain II residues	167–265	92–186	105–193	104–235, 265–271	83–250, 317–332	124–247, 332–344
Domain III residues	—	—	221–307	—	—	—
Crossover regions	163–166, 266–269	90–91, 187–192	100–104, 194–199, 218–220	—	—	119–123, 248–252, 327–331

\* 1l1t for the closed (liganded) form, 2lao for the open (unliganded) form.

<sup>†</sup> Residues 291–324 form an extended region appended to domain I.

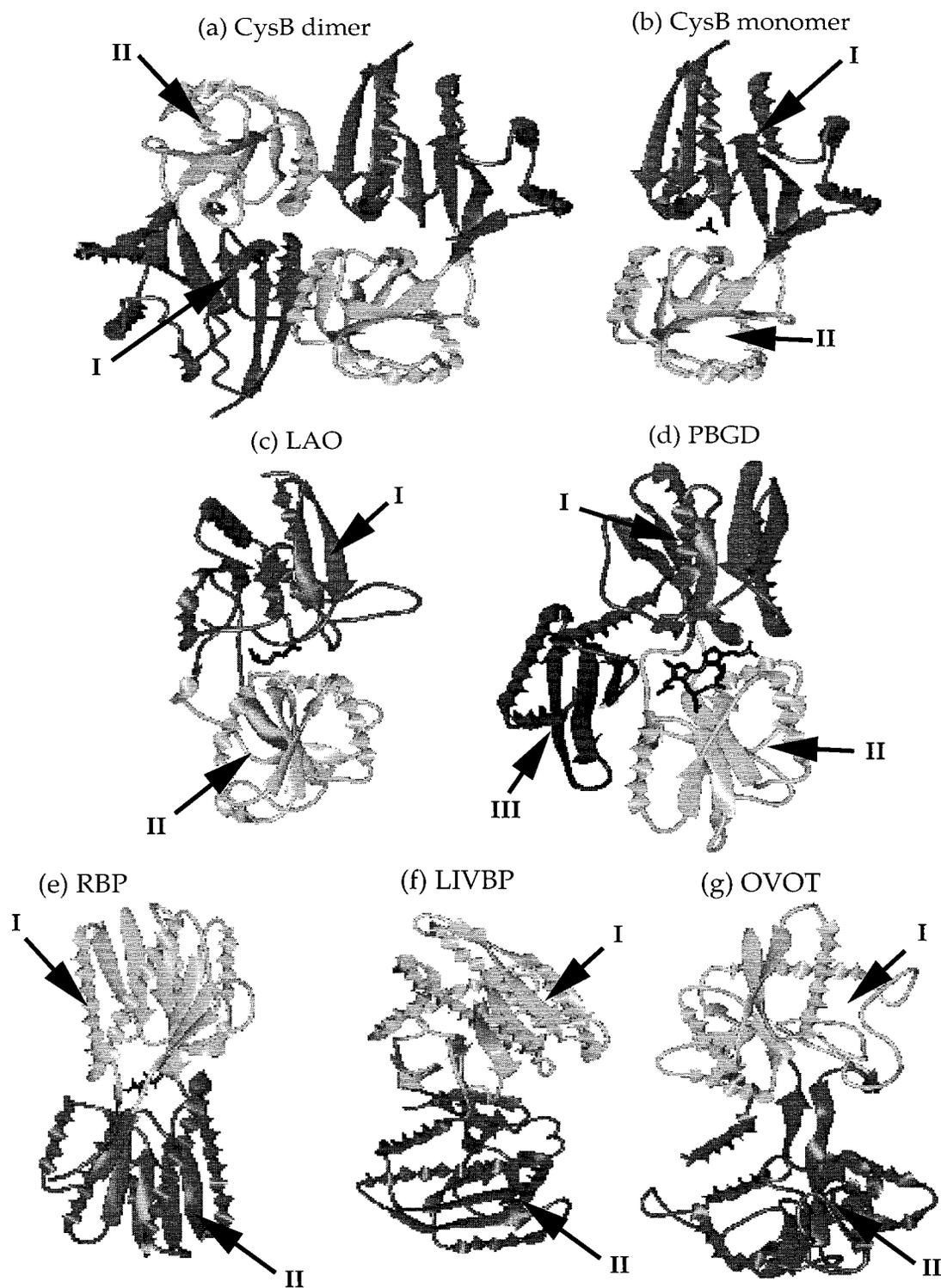


FIGURE 1 Ribbon diagrams of (a) CysB dimer, (b) CysB monomer, (c) LAO (closed), (d) PBGD, (e) RBP, (f) LIVBP, and (g) OVOT (see Table 1 for PDB names). All proteins consist of two domains, I and II, except PBGD, which has three domains. All domains are Rossmann-like folds with a  $\beta$ -sheet core and  $\alpha$ -helices outside. The crystallographic structure for CysB contains a sulfate ion at the cleft between the two domains, as shown in *b*. LAO contains a lysine bound to the cleft, as seen in *c*. PBGD contains the cofactor DPM, covalently bonded to the conserved Cys<sup>242</sup> residues of domain III, and RBP has the ligand  $\beta$ -D-ribose in *e*.

rotation. The binding of a small ligand,  $\beta$ -D-ribose, between domains I and II causes the activation of the protein, allowing it to be recognized by membrane components (Ordal et al., 1985; Stewart and Dahlquist, 1988; Mowbray and Cole, 1992).

Translocation of amino acids, sugars, oxyanions and other nutrients into the cytoplasm of Gram-negative bacteria is accomplished by a series of steps involving a number of different proteins, one of which is LIVBP (Fig. 1 *f*). LIVBP has two domains with Rossmann-like folds similar to the domains of the other SBPs. Between the domains there is a cleft that is easily accessible to ligands. Each domain has the same supersecondary structure, a central  $\beta$ -pleated sheet flanked on either side by helices. Here the dynamics of the unliganded (open) structure of LIVBP (Sack et al., 1989a,b) will be examined.

OVOT belongs to the family of iron-binding proteins, along with other transferrins such as serum transferrin and lactoferrin. These proteins serve to control the iron level in the body fluid of vertebrates by their ability to bind very tightly two  $\text{Fe}^{3+}$  ions. Here the N-terminal lobe of ovotransferrin in the apo-form is analyzed. The structure consists of two similarly sized domains having Rossmann-like folds (Fig. 1 *g*). The two iron-binding sites are located within the interdomain cleft of each domain, and the domains are connected by two  $\beta$ -strands (Mizutani et al., 1999).

Comparison of the dynamics of apo-OVOT, unliganded LAO, and unliganded LIVB, with the four liganded structures, CysB(88–324), liganded LAO, PBGD, and RBP, will shed light on the differences in the motions of the closed and open forms of SBPs.

Table 2 compares the six proteins studied. The amino acid sequence identities obtained from the FSSP comparisons (Holm and Sander, 1994) range between 4% and 15%. The values in parentheses represent the root mean square deviations (RMSDs) based on FSSP calculations, the most similar proteins being LAO and OVOT (2.8 Å) and the most dissimilar ones LIVBP and PBGD (8.3 Å).

## GAUSSIAN NETWORK MODEL

A summary of the GNM is given here. The reader is referred to our previous studies for details (Bahar et al., 1997, 1998a; Haliloglu et al., 1997; Bahar and Jernigan, 1998). The GNM

essentially relies on the construction of a Kirchhoff matrix of contacts,  $\Gamma$ . For a protein of  $N$  residues,  $\Gamma$  is an  $N \times N$  symmetrical matrix whose  $ij$ th off-diagonal element is  $-1$  if  $\alpha$ -carbons  $i$  and  $j$  are separated by a distance lower than a cutoff value,  $r_c$ , for effective interaction, and zero otherwise; and the  $i$ th diagonal element is equal to the negative sum of the off-diagonal elements in the same row (or column). In the GNM, all contacting residues are assumed to be coupled by a single parameter ( $\gamma$ ) harmonic potential, giving rise to Gaussian fluctuations about their mean (native) positions. Cross-correlations between residue fluctuations are found from

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = (3kT/\gamma)[\Gamma^{-1}]_{ij} \quad (1)$$

where  $\Delta \mathbf{R}_i$  is the fluctuation in the position of residue  $i$ ,  $k$  is the Boltzmann constant,  $T$  is the absolute temperature,  $[\Gamma^{-1}]_{ij}$  designates the  $ij$ th element of the inverse of  $\Gamma$ , and the angular brackets refer to the average over all modes of motion. Mean square (ms) fluctuations  $\langle (\Delta R_i)^2 \rangle$  are found from Eq. 1, using  $i = j$ . It should be noted that the determinant of  $\Gamma$  is zero. Therefore  $\Gamma$  cannot be inverted.  $\Gamma^{-1}$  is an approximation calculated from the  $N - 1$  nonzero eigenvalues  $\lambda_k$  and corresponding eigenvectors  $\mathbf{u}_k$  of  $\Gamma$  using

$$\Gamma^{-1} = \sum_k \lambda_k^{-1} [\mathbf{u}_k \mathbf{u}_k^T] \quad (2)$$

The above summation is carried out over the range  $2 \leq k \leq N$ , omitting the zero eigenvalue  $\lambda_1$ . The superscript T designates the transpose.  $[\mathbf{u}_k \mathbf{u}_k^T]$  is an  $N \times N$  matrix, representing the contribution of the  $k$ th eigenvector to  $\Gamma$ . Equation 2 provides a simple means of decomposing the dynamics into a series of modes. Combining Eqs. 1 and 2, the ms fluctuations of the  $i$ th residue are expressed as

$$\langle (\Delta R_i)^2 \rangle = (3kT/\gamma) \sum_k 1/\lambda_k [\mathbf{u}_k \mathbf{u}_k^T]_{ii} \quad (3)$$

where  $2 \leq k \leq N$ , and the subscript  $ii$  designates the  $i$ th diagonal element of the matrix enclosed in brackets. These elements, when plotted against the index  $i$ , represent the  $k$ th mode shape, i.e., the normalized distribution of ms fluctuations of the  $N$  residues as driven by the  $k$ th mode of motion. The eigenvalue  $\lambda_k$  is a measure of the frequency of the  $k$ th mode. In particular, the lowest nonzero eigenvalue ( $\lambda_2$ )

**TABLE 2** Comparison of the proteins from FSSP calculations: sequence percentage identities and RMSD (Å)

	CysB (88–324)	LAO (closed)	PBGD	RBP	LIVBP	OVOT
CysB(88-324)	100 (0.0)	9 (4.0)	12 (4.0)	7 (4.5)	11 (7.8)	12 (6.7)
LAO	9 (4.0)	100 (0.0)	12 (4.7)	11 (3.5)	7 (6.4)	11 (2.8)
PBGD	12 (4.0)	12 (4.7)	100 (0.0)	4 (3.6)	8 (8.3)	15 (5.6)
RBP	7 (4.5)	11 (3.5)	4 (3.6)	100 (0.0)	12 (7.0)	9 (3.7)
LIVBP	11 (7.8)	7 (6.4)	8 (8.3)	12 (7.0)	100 (0.0)	7 (4.7)
OVOT	12 (6.7)	11 (2.8)	15 (5.6)	9 (3.7)	7 (4.7)	100 (0.0)

See Holm and Sander (1994).

refers to the frequency of the most cooperative (slowest) mode of motion, and the corresponding eigenvector  $\mathbf{u}_2$  reflects the shape of this so-called global, or dominant, collective mode. The latter mode is usually implicated in biological function.

We note that the GNM modes are equivalent to the normal modes of the molecule under the two simplifying assumptions of 1) identical force constants,  $\gamma$ , for all interacting pairs of residues, and 2) isotropic fluctuations. This reduces the number of distinct modes from  $3N - 6$  internal modes in conventional normal mode analysis to only  $N - 1$ .

In the present analysis,  $\langle(\Delta R_i)^2\rangle$  values will be calculated first. These will be shown to be in excellent agreement with those indicated by x-ray crystallographic temperature factors  $B_i = (8\pi^2/3) \langle(\Delta R_i)^2\rangle$ . Second, the most cooperative, dominant modes of motion will be identified for each protein, which will illustrate the close similarity in the dynamics of the investigated systems, accompanying their structural similarities. Thus a common mechanism of motion will be identified for the different proteins, directly imparted by their common overall architecture. The effects of dimerization and ligand binding on this mechanism of motion will be explored. This analysis will reveal the identity of the structural motifs, or subdomains, involved in comparable modes of action (hinge-bending, or large-amplitude swinging, or propagation of movements), thus opening the possibility for an alignment of residues based on their dynamic characteristics. Structural superposition based on such an alignment will, on the other hand, disclose the few elements distinguished by their unique conformations, essentially those involved in the specific function of the particular proteins.

## RESULTS AND DISCUSSION

### Temperature factors

Fig. 2 displays the fluctuations of CysB, LAO, and PBGD residues. Solid and dashed curves represent the calculated and experimental results, respectively. The experimental data are the x-ray crystallographic B factors (also called Debye-Waller factors) of the individual  $\alpha$ -carbons, reported in the PDB files of the respective structures (Table 1). Theoretical results are found from the GNM. The agreement between theory and experiments is excellent. The respective force constants,  $\gamma$  in the GNM, chosen for each protein so as to scale overall the calculated curves to best fit the experimental data, are 4.72, 4.05, and 4.19 kcal/Å<sup>2</sup> for CysB, LAO, and PBGD, respectively. The cutoff distance for interactions is taken as 10 Å. The results for CysB (Fig. 2 *a*) are calculated for the dimeric form but are displayed for one monomer only. We note that coordinates were not reported for the segment of PBGD between residues 49 and 58, and the corresponding sections of the curves in Fig. 2 *c* have simply been connected by a straight line. Calculations per-

formed for the other proteins, RBP, LIVB and OVOT, similarly yielded results (not shown) in good agreement with experimental data.

### Dominant modes of motion

Fig. 3 displays the shape of the global, most cooperative mode of motion for 1) CysB(88–324), 2) LAO (closed form), and 3) PBGD. These three structures (and RBP) are all liganded and assume the so-called closed conformations. They will be shown below to have closely superimposable dynamic characteristics, whereas the three unliganded structures (open form of LAO, apo-OVOT, and unliganded LIVB) will exhibit a substantial departure from the first set, illustrating the loss in the generic behavior of SBPs in the absence of their ligands.

The dashed curve in Fig. 3 *a* displays the global mode shape of CysB(88–324) in the dimeric form. It is found from the weighted contribution of the slowest two modes of motion  $2 \leq k \leq 3$ , following Eq. 3, while LAO and PBGD curves in Fig. 3, *b* and *c*, refer to the slowest mode ( $k = 2$ ) only. The total number of modes in the dimer is twice as many as for the monomer—hence the equivalence of the slowest two modes of CysB dimer to the single slowest mode of LAO or PBGD.

Our previous GNM analyses (Bahar et al., 1998a, 1999; Bahar and Jernigan, 1998, 1999; Demirel et al., 1998; Jernigan et al., 1999) indicate that the minima in the global mode shapes generally coincide with the residues acting as hinges; the same regions are also usually observed to be correlated with (or juxtaposed to) biologically active sites, such as catalytic sites in enzymes. Maxima, on the other hand, correspond to segments distinguished by their enhanced mobilities, often implicated in substrate recognition. In the present calculations, the cross-over regions between the domains (Table 1) emerge as minima (Fig. 3), which is consistent with their hinge-bending role. Other minima, indicative of sites critically important for monitoring the global motions, are observed as well. These will be examined separately for each protein.

CysB residues T100, T102, Q103, T149, W166, and T270 are pointed out to be critical for the stabilization of protein, as they form hydrogen bonds with the sulfate ligand (Tyrrell et al., 1997). It is interesting to observe that these residues are all located at minima (highly constrained regions) in the global mode (Fig. 3 *a*). The cleft between domains I and II is the substrate-binding site of CysB. The cleft is lined with the polar side chains of H101, Q128, G129, P131, E150, Y197, T202, and T225, besides the ones mentioned above. Residues around the cleft participate in the active site. Reported experiments show indeed that mutations around T100–T102, T149, and T202 affect cofactor responsiveness (Tyrrell et al., 1997). These residues are all located at the minima in the global mode shape, in parallel with previous GNM results.

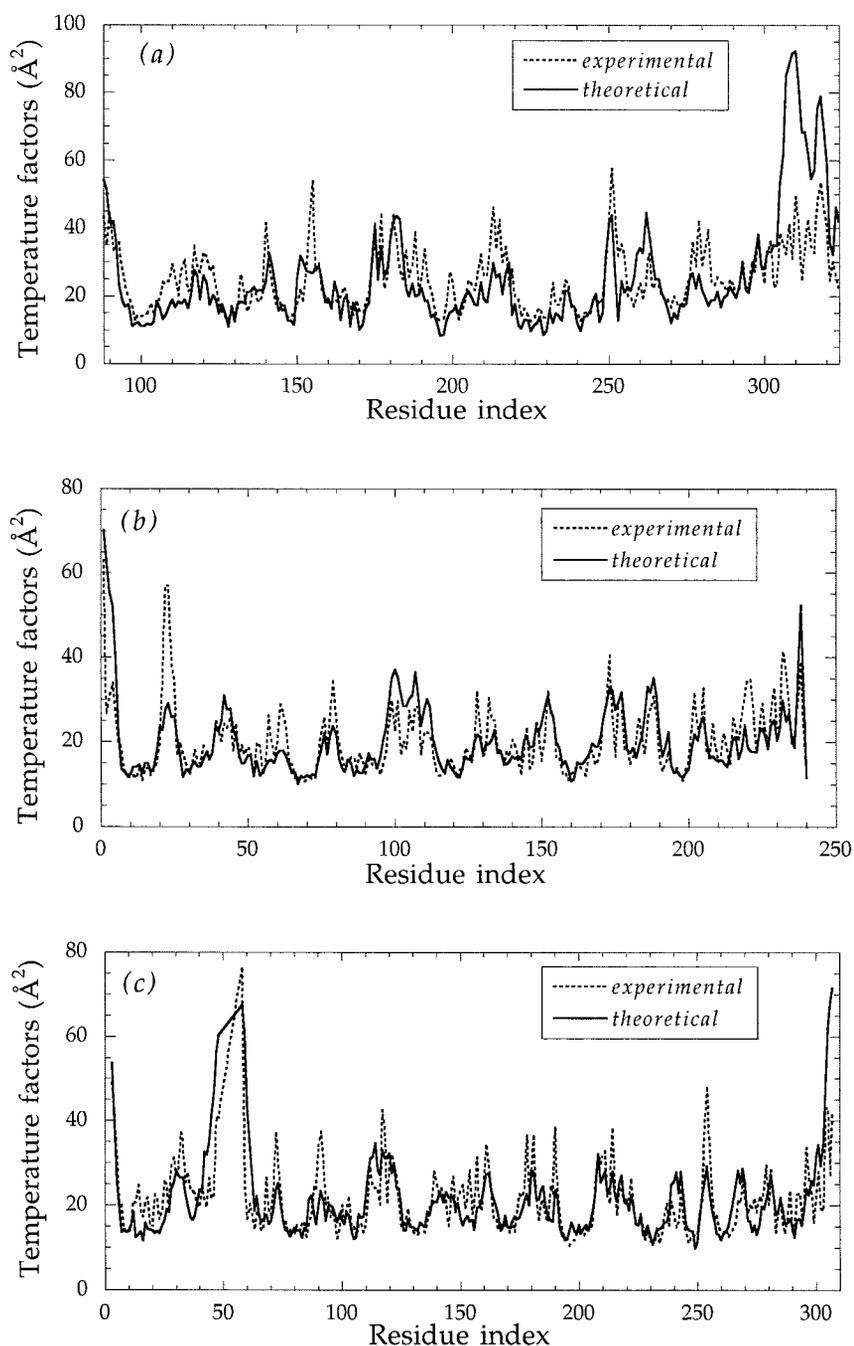


FIGURE 2 Crystallographic temperature factors for (a) CysB(88–324), (b) LAO (closed), and (c) PBGD. ---, Experimental temperature factors; —, the GNM results. The PDB structure of the CysB dimer and the LAO and PBGD monomers are used in calculations.

A second group of mutations, clustered around residue V248, was observed to affect the cofactor response and/or DNA binding affinity of CysB (Tyrrell et al., 1997). These residues, situated on a loop that lines a groove across the dimer surface, were suggested to interfere with DNA binding or multimerization. A minimum is again observed in this region, although it is not as pronounced as the preceding ones, supporting the possible involvement (presumably relevant to function) of this site in controlling the global motions of the molecule.

Finally, the dimerization interface is well conserved among all CysB proteins (Tyrrell et al., 1997) whose se-

quences differ substantially. These regions are also observed to be constrained in our analysis. Maxima, on the other hand, indicate the most mobile regions in the global modes. They correspond to loop regions in CysB (88–324), except for the two helical segments 175–190 and 310–320. These regions are likely to act as recognition sites; however, there are no relevant experimental data.

LAO residues S18, F52, S120, Y190, and D193 are known to be ligand-binding sites (Kang et al., 1991). These are all located at minima in the curve displayed in Fig. 3 b. The ligand-binding role of the latter two is worth noting in that they are located next to the hinge residues F191, G192,

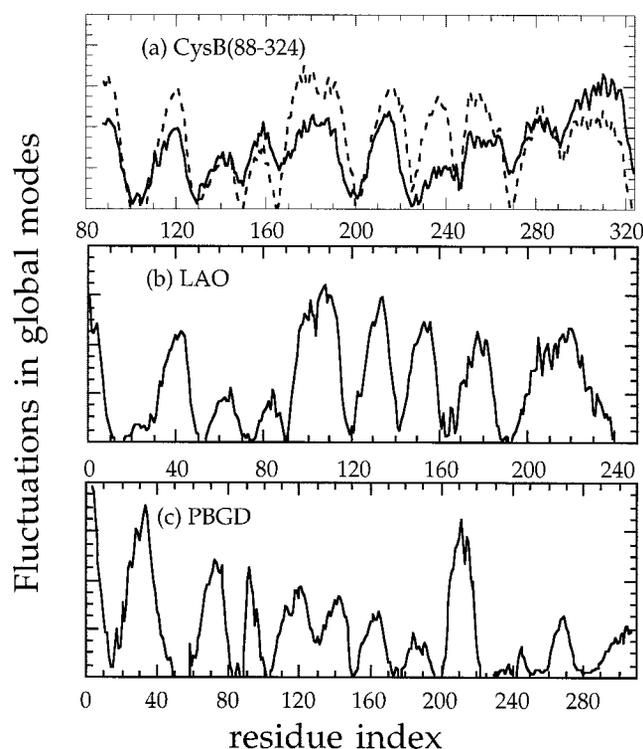


FIGURE 3 Comparison of the slowest mode shapes of (a) CysB(88–324), (b) LAO (closed), and (c) PBGD. The solid curves are calculated for the monomers, alone, and the dashed curve in *a* is for the monomer within the dimer. The list of residues at the minima are listed in the top of Table 3. These assume a hinge-bending role in the global movements of the molecules.

and G194, which are also noted below as the key residues, distinguished by their high-frequency/small-amplitude fluctuations. Apparently, the segment of residues Y190–G194 is actively communicating interdomain signals induced by ligand binding. Other minima, S70–L71, A90, E167, and A141–N142, correspond to the residues that line the cleft.

The maxima around residues Q42, K62, S84, S100–K110, T132, A152–G153, E177, and R200–D220, on the other hand, correspond to amino acid or membrane recognition regions. In particular, D149 and R154 were pointed out in mutation experiments (Kang et al., 1991) to be important for membrane interactions.

PBGD crystal structure contains a cofactor DPM, covalently bonded to C242. The global mode curve of PBGD (Fig. 3 *c*) yields the active site C242 as a minimum. Likewise, F62 and D84, two residues that were reported to be important for cofactor binding, and residues R101 and R232, whose point mutations inhibit enzyme activity (Louie et al., 1992), are all located at minima. Finally, the residues lining the substrate-binding cavity (R132, R11, R149, A170, A172, G173, and R176) are all distinguished by severely suppressed fluctuations in the global mode.

Maxima all belong to surface-exposed loop regions in the first and second domains, suggesting that domain III is not

directly involved in substrate recognition. Instead, the fact that a large portion of it is severely constrained in the global mode signals its possible involvement in assisting in enzymatic activity.

Similar analysis performed for RBP (not shown) yielded minima near residues N13, D89, R90, R141, N190, D215, and Q235, all of which are known to interact with the ligands (Mowbray and Cole, 1992). The active site of RBP is reported to comprise residues 13, 15–16, 89–90, 141, 164, 190, 215, and 235 in PDBSUM (Laskowski et al., 1997), in accord with our results. In particular, the putative hinge of the RBP is composed of residues 103–104, 235–236, and 263–265 (Mowbray and Cole, 1992), and these are confirmed in the present analysis to be located at the minima of the global mode shape curves.

A summary of the residues acting as hinges is presented in the first eight rows of Table 3, for the four liganded structures described above. The hinge sites are all minima in the slowest mode shapes (Fig. 3). Residues in the same row of the table occupy the same sequential position in the structurally aligned forms of the three structures, as determined from the FSSP program (Holm and Sander, 1994, 1996).

### Comparison of the global dynamics of the proteins

The fluctuations and the hinges in the slowest mode are illustrated in Fig. 4 by the colored ribbon diagrams for CysB(88–324) dimer (Fig. 4 *a*), CysB(88–324) monomer (Fig. 4 *b*), LAO (Fig. 4 *c*), PBGD (domains I and II) (Fig. 4 *d*), RBP (Fig. 4 *e*), LIVBP (Fig. 4 *f*), and OVOT (Fig. 4 *g*). Six colors are used to represent different levels of

TABLE 3 Hinge sites and other key residues\*

CysB(88-324)	LAO	PBGD	RBP†
Minima (slowest modes)			
T100-Q103	T12-F17	R11-A16	T10-F16
G129-S130	S50-D53	G58-F62	S39-N42
T149-A150	S70-S72	K83-D84	T66-D67
H165-W166	A89-A90	R101-P104	S103-D104
T202	S120	T127-S128	G134-S136
T225-D226	A141-N142	G150-N151	D163-F164
A247-P250	A163-F168	G173-K175	D215-P218
H267-T270	Y190-G192	P194-G199	Q235-P237
Maxima (fastest modes)			
Y96	R7	R7	L6
I134	L55	L66	
D144-I147	D65-I68	D76-V79	P65
R168, V171	S92, I95	D106, V109	G108
G238-V240	L155-A157	Y164-A166	G169, G213-P214
H267-T270	Y190-G192	P194-G199	T232

\* Residues in a given row play an analogous role in the different structures.

† RBP has additional minima (slow modes) D89–R90 and N190–D192, and maxima (fast modes) at A188–A194 and G244–V245.

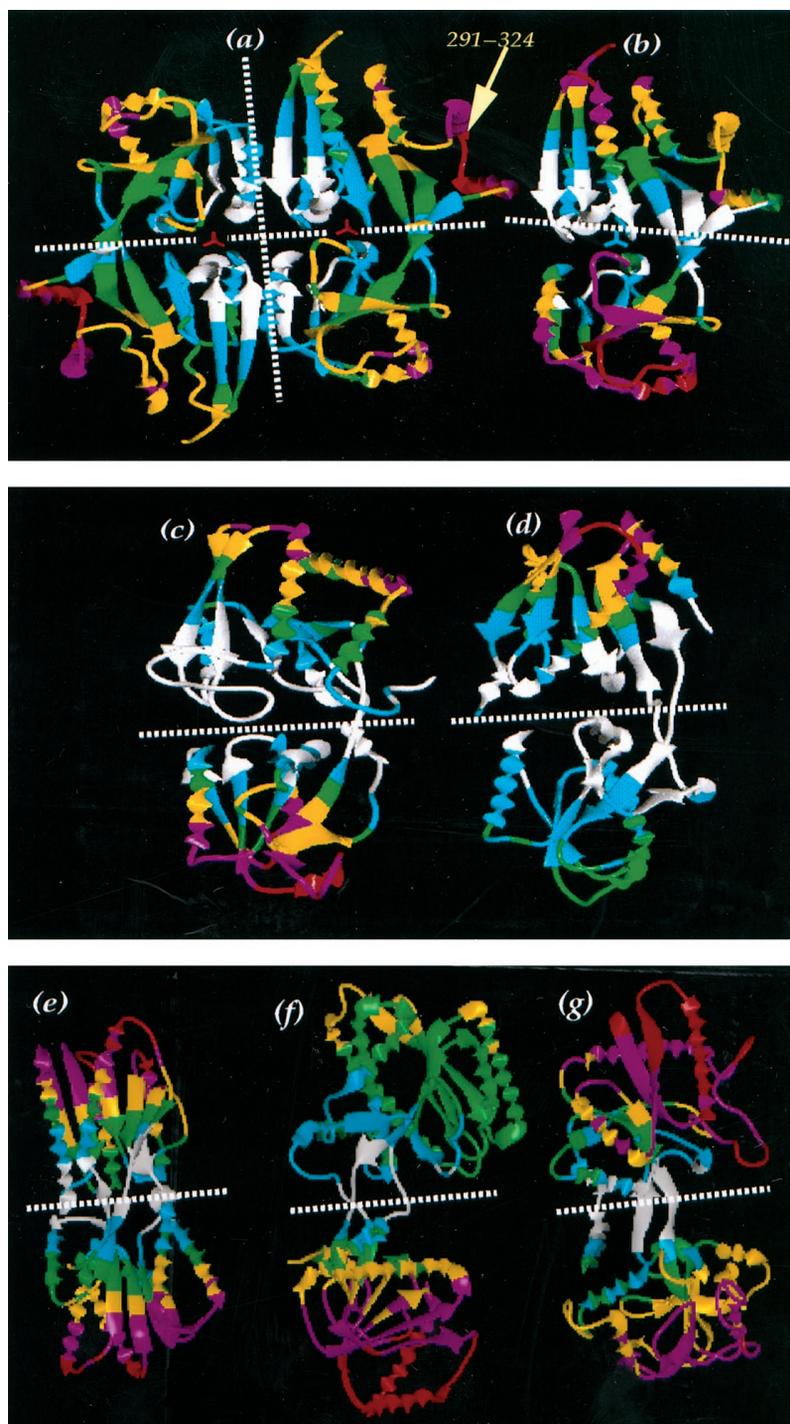


FIGURE 4 Ribbon diagrams illustrating the amplitudes of motion for the different structural regions of (a) CysB dimer, (b) CysB(88–324) monomer, (c) LAO (closed), (d) PBGD, (e) RBP, (f) LIVBP, and (g) OVOT in their most cooperative, slowest modes. Six different colors are used to represent the different levels of flexibilities: white, cyan, green, yellow, magenta, and red, where the smallest displacement level is white and the highest is red. The dashed white lines are the hinge rotation axes about which the concerted movements of the domains occur. There are two axes in the case of the CysB (88–324) dimer, suggestive of a saddle-forming motion.

flexibility: white, cyan, green, yellow, magenta, and red. White regions are almost fixed in the global mode; red regions are the most mobile. In general, the helices and loops are more flexible, while the  $\beta$ -strands are more severely constrained. The lower mobility of domain II in PBGD is due to the juxtaposition of domain III.

The dashed lines indicate the loci of residues that are severely constrained in the global mode. These may alter-

natively be viewed as the axes of symmetry about which domains I and II are engaged in their cooperative, opposite direction fluctuations driven by the global mode. Residues at the interdomain interfaces, or near the cleft, generally exhibit severely suppressed fluctuations. The peptides linking the domains are almost fixed in space, consistent with their hinge-bending role. Interestingly, the axes are located precisely at the interface between domains I and II in the

case of LAO and PBGD but are shifted toward domain I in the CysB monomer. In the latter structure, domain II enjoys a higher conformational freedom than domain I. The disparity between the mobilities of the two domains is observed to disappear upon dimerization of the molecule (see Fig. 4 *a*).

The fact that the two domains move in opposite directions is verified by calculating the cross-correlations,  $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle$ , between residue fluctuations. The results are illustrated for LAO in the correlation map displayed in Fig. 5 over all modes. Similar results (not shown) were found for the other proteins. As seen in the map, residue pairs within a given domain are positively correlated, i.e., they move in the same direction, whereas those belonging to the two different domains are negatively correlated, i.e., they undergo opposite direction fluctuations. We note that the C-terminal segment of  $\sim 20$  residues, which is generally considered to be part of domain I, is instead coupled with domain II. Similarly, the extension 291–324 appended to domain I in CysB(88–324) was found to undergo concerted motions with domain II, rather than domain I.

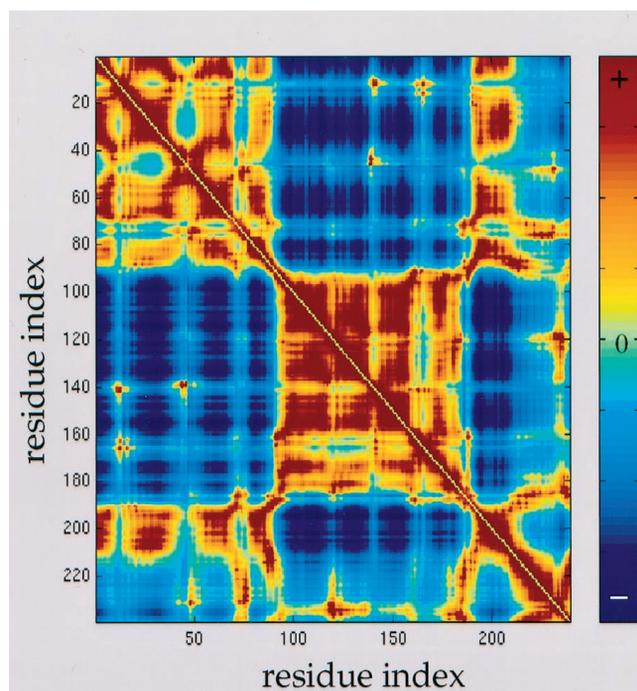


FIGURE 5 Cross-correlations  $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle$  for residue fluctuations in LAO for all modes of motion. The two axes refer to residue indices. Blue and red regions refer to anticorrelated and positively correlated regions, as indicated by the vertical color scale, on the right. Residue pairs in a given domain are positively correlated (i.e., they undergo the same direction fluctuations), whereas those in two different domains are anticorrelated (i.e., they fluctuate in opposite directions) in general. The C-terminal segment, which is considered to be a part of domain I, appears to be dynamically coupled to domain II rather than to domain I.

## Alignment based on global dynamics

The global mode shapes of the examined proteins (Fig. 3) exhibit some similarities. However, as the lengths of the secondary structures and loops differ in each protein, the peaks and minima do not coincide. To visualize the extent to which these structures exhibit a common dynamic behavior on a global scale, and which particular regions are responsible for departures among them, alignments of the proteins based on their global mode shapes were undertaken here.

Fig. 3 reveals that domain II in each protein exhibits a relatively more persistent behavior, the qualitative shape of the corresponding parts of the three curves being more closely concordant. Based on this observation, the second domains of the proteins were superposed first. Domain II residues 92–186 of LAO (see Table 1) were taken as the basis. The ranges of residues of CysB and PBGD (to be superimposed) were determined by requiring the global mode shape of their domain II to be maximally correlated with that of LAO, without gaps. Segments of 94 contiguous residues were thus considered in CysB and PBGD, shifting one residue at a time. The highest correlation was achieved when CysB(88–324) and PBGD were shifted by  $-79$  and  $-10$ , respectively, i.e., the domain II residues 92–186 of LAO (see Table 1) were aligned with residues 171–265 of CysB and residues 102–196 of PBGD. After optimal superposition of the domains II, an iterative procedure was performed for domain I residues. Gaps of various sizes were included in this case, and we inspected their FSSP alignments. No attempt to develop an automated algorithm was made here, but the prospect of accomplishing a meaningful alignment based on dynamic characteristics was explored. The result, shown in Fig. 6, demonstrates that the examined SBPs indeed obey a closely superimposable mechanism of

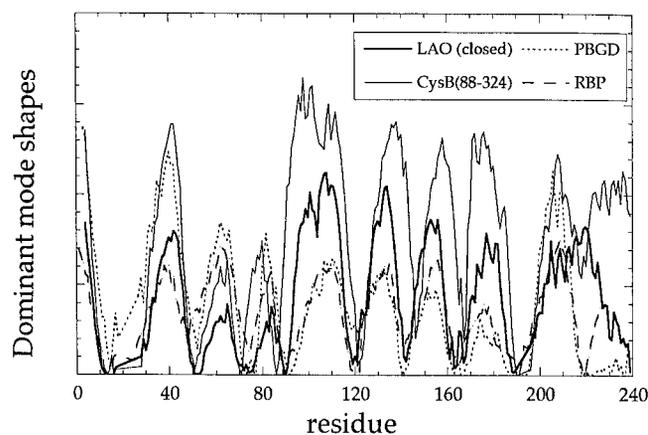


FIGURE 6 Superposition of the global mode shapes of the indicated four proteins, demonstrating that a common mechanism of cooperative motion is operative for all four SBPs. The abscissa refers to the residue indices of LAO. The global mode shapes are matched after the short segments shown in Fig. 7 are deleted.

motion on a global scale upon suitable alignment of their residues. Details are presented below. Interestingly, the regions deleted for achieving optimal alignment are those usually involved in the particular functions of the individual proteins, revealing that a common scaffold and global mechanism can underlie a diversity of functions.

In the case of the LAO-CysB(88–324) pair, an adequate superposition of the global mode shapes requires the deletion of the residues 18–27 and 191–195 in domain I of LAO. The C-terminal segments are the only regions exhibiting different behavior. We note that the extended region appended to domain I in CysB, comprising the C-terminal residues 291–324, is distinguished by its unique structural and dynamic characteristics. The fact that this region cannot be matched invites attention to its unique behavior, which presumably is relevant to its specific function and dynamics. Fig. 7 *a* illustrates the structural alignment of CysB(88–324) and LAO deduced from this analysis. The rms deviation between the  $\alpha$ -carbon coordinates, calculated for the residue {4–17, 28–190, 196–216} for LAO and {93–290} for CysB, deduced from the present superposition of the global mode shapes, is 7.15 Å. The structural elements, distinguished by their unique dynamic features, are indicated therein.

A similar analysis for the LAO-PBGD pair (Fig. 6) indicates the equivalence of the stretches of the respective residues {3–19, 20–45, 63–196, 197–220} and LAO residues {3–19, 27–52, 53–186, 192–215}, approximately. Therefore, residues 46–62 are to be omitted in PBGD, along with the two stretches of residues 20–26 and 187–191 in LAO. The corresponding structural alignment is shown in

Fig. 7 *b*. The rms deviation is 7.02 Å. Therein, the third domain of PBGD is shown in yellow. The substrates, lysine and DPM, are shown in red and magenta, respectively. The regions distinguished by their peculiar behavior are indicated by boxes or arrows, in parallel with Fig. 7 *a*. We note that a large portion (49–57) of the loop 46–62 of PBGD was not visible in x-ray. The segments with known coordinates, 46–48 and 58–61, are displayed in orange, and the LAO residues 187–191 are in green. For clarity, the side chain of I48, which is located precisely at the active site cleft entrance, is also displayed. Finally, LAO and RBP can be aligned by deleting residues {18–24, 182–188} from LAO and {78–98, 137–144, 177–206} from RBP, using both their global mode shapes and their FSSP alignments.

Fig. 6 demonstrates that the global dynamics of the four structures are quite similar. Thus a common mechanism of motion is operative in all four cases on the global scale. Certain regions, however, can be differentiated. It is natural to ascribe the specific function of these structurally similar proteins to these particular regions. These are residues 18–25 and 191–195 and the C-terminal segment 217–238 of LAO, the extended portion of residues 291–324 appended to domain I in CysB, and residues 46–62 in PBGD. Among these, LAO 18–27 and CysB 291–324 are surface-exposed, suggesting a specific recognition role; whereas the remainder are located near the active site cleft, which could be related to their involvement in specific enzymatic functions. In PBGD, it is additionally clear that a number of domain III residues (near 220–245) also participate in the enzymatic function, as implied by their severe confinement in the global modes (see Figs. 3 *c* and 6).

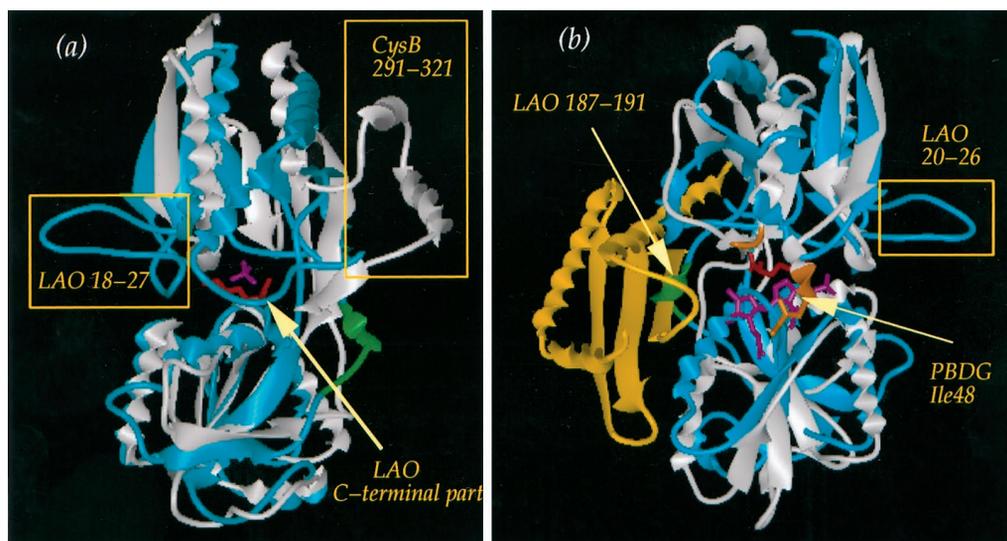


FIGURE 7 Ribbon diagrams of the pair of structures (a) CysB(88–324)–LAO and (b) PBGD–LAO, superimposed following the dynamic alignment described in Fig. 6. The rms deviations between the pairs of structures are 7.15 and 7.02 Å, respectively. Elements discriminated by their unique behavior and likely to be involved in the distinct, specific function of the proteins are specially indicated.

## Key residues

The above analysis provides an estimation of the structural regions that assume an equivalent role in the global dynamic behavior of the three proteins. However, both the global mode shape and the superposition of the modes carried out above give insights regarding “structural blocks” in general. Whereas the loci of hinge sites are sufficiently precise (Table 3), as indicated by the sharp minima in the global mode shapes, the key elements coordinating the large-scale concerted movements of the domains or subdomains are not distinguishable. The latter are probably located in the core sites of structural elements and likely participate in highly constrained, finely tuned tertiary interactions. For identification of such residues we also examine the fastest mode shapes of the proteins. In contrast to the slow modes that yield relatively smooth distributions in fluctuations, the fastest/smallest amplitude modes yield sharp peaks at individual sites, which permit us to unambiguously identify the residues most tightly packed and most severely constrained in a particular tertiary structure. Previous studies showed that the latter, referred to as *kinetically hot residues*, are rather conserved and often take part in folding nuclei (Bahar et al., 1998; Demirel et al., 1998).

Calculations yield the set of residues listed in the lower part of Table 3. These appear as peaks in the fastest 10 modes of motions. Residues in a given row are counterparts of one another in the different proteins. The rms deviation values for the aligned structures of the respective pairs CysB(88–324)-LAO, LAO-PBGD, CysB(88–324)-PBGD, and LAO-RBP are 4.0, 4.7, 4.0, and 4.5 Å, respectively.

Interestingly, the kinetically hot residues are located almost invariably adjacent to the hinge-bending sites, suggesting that their tight packing and close communication are a requirement for efficient propagation of the motion.

Fig. 8 illustrates the position of these residues in the structure. Results are displayed for PBGD only, because in all cases, counterparts in the other proteins lie at the same positions in the three-dimensional structures. It is interesting to observe that these residues form a network of interactions, as indicated by the dashed yellow lines, propagated from the interface toward the inner regions of the two domains. Domain I appears to be stabilized through inter-strand interactions, whereas the domain II core is located mostly on one central strand, which interacts at its terminal part with a second one. Among these interactions, the most pronounced one, distinguished by the highest peaks in the mode shapes of all proteins, is located between two strands of domain I, labeled here by indices 7 and 77.

## Effect of ligand binding on the dynamic characteristics of LAO

The structure of LAO examined above is the closed form complexed with lysine. An unliganded, open form of the same protein has also been determined by x-ray crystallography (Oh et al., 1993). Maiorov and Abagyan (1997) stated that multidomain proteins may undergo substantial domain displacements, while their interdomain structure remains essentially unchanged. They studied the stabilizing role of ligand in LAO and could predict the domain linkers cor-

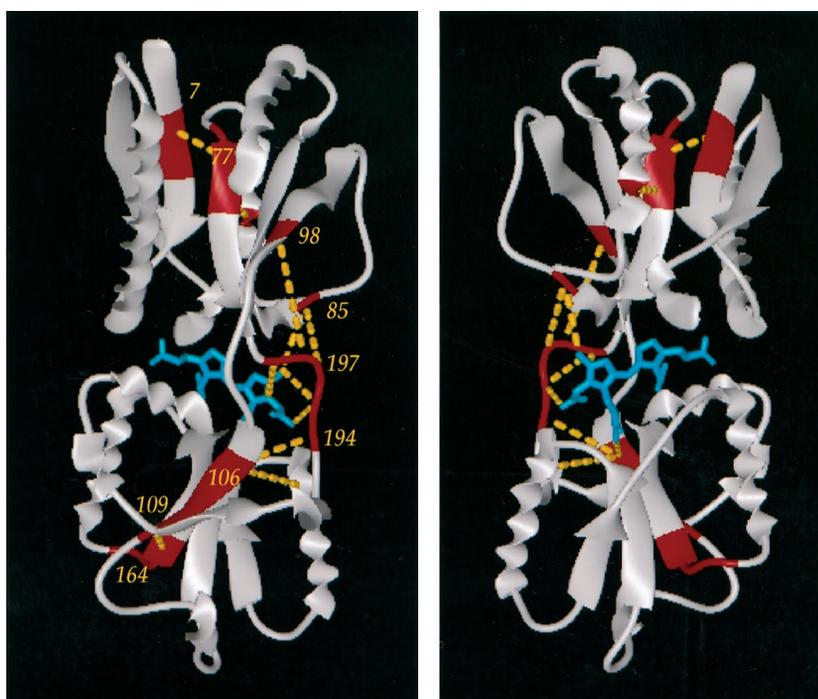


FIGURE 8 Kinetically hot residues in the examined structures, illustrated for PBGD in front and back views. The regions colored in red are the structural elements that are most efficiently communicating the cooperative movement across the molecule. See Table 3 for the complete list of these residues and their counterparts.

rectly. These were pointed out to act as hinges that accommodate closed and open conformations (Maiorov and Abagyan, 1997). To see the role of the lysine ligand in the dynamic characteristics of the protein, we repeated the GNM analysis for the open structure. Experimental and theoretical temperature factors (not shown) again are found to be in good agreement.

Fig. 9 *a* presents a comparison of the slowest mode shapes of the open and closed conformations of LAO. The open conformation is represented by the solid curve, whereas the closed conformation is given by the dotted curve. The two curves disclose the differences between the global modes of the two forms, as also illustrated in Fig. 10. Ligand binding induces both qualitative and quantitative changes in the flexibility of certain structural elements. The arrows in Fig. 9 *a* indicate the regions whose mobilities were of considerable amplitude before binding but almost completely suppressed upon substrate binding. This depression of flexibility is indicative of the involvement of these particular regions in the binding of substrate. Interestingly, domain I residues around the loop 20–25, which were distinguished above by their demarcating behavior, lie among the group of residues strongly affected by substrate binding. This confirms their role in the specific binding of LAO. The other structural element remarked upon was the

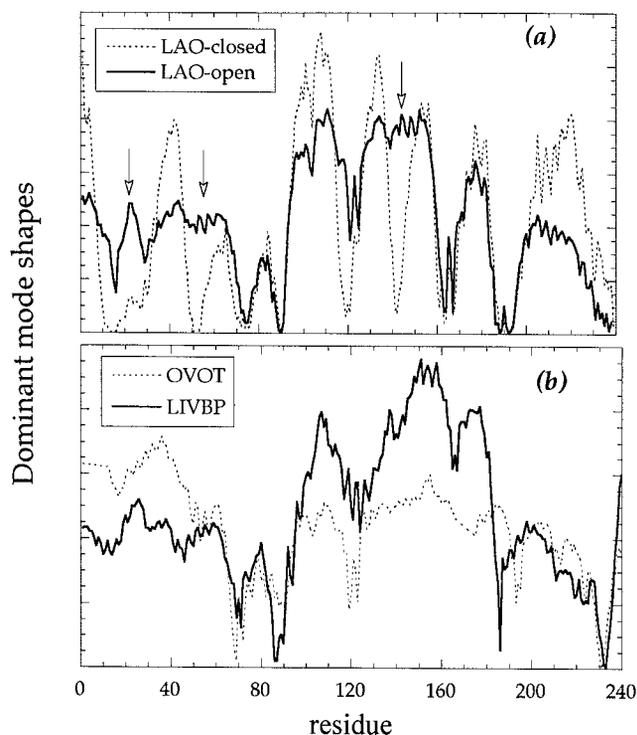


FIGURE 9 (a) Comparison of the global mode shapes of LAO in the closed (---) and open (—) forms of the molecule. The arrows indicate the regions in which flexibility becomes severely suppressed upon ligand binding. (b) Global mode shapes of LIVBP and OVOT, to be compared to that of LAO (open).

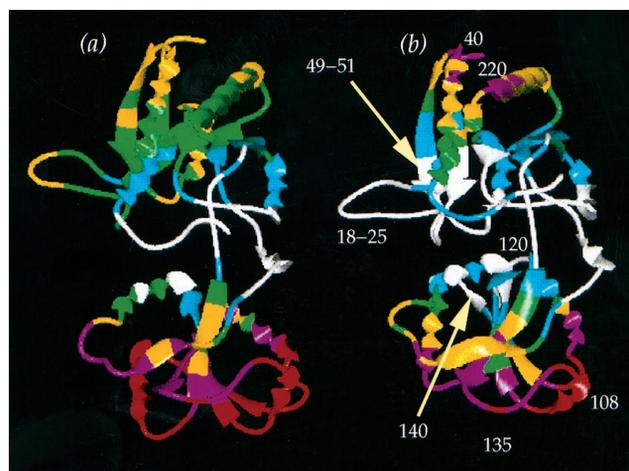


FIGURE 10 Ribbon diagrams illustrating the relative flexibilities of different structural segments in the (a) open and (b) closed forms of LAO. The color scheme is the same as in Fig. 4. Residues whose flexibilities are significantly affected are labeled in *b*.

C-terminal segment, which is again significantly affected, but in the opposite direction—an enhancement in mobility is observed. This C-terminal region includes an HTH motif. This element, along with two others centered about residues 40 and 105, might be involved in a signaling process activated by ligand binding.

Fig. 9 *b* displays the slowest mode shape for the two other unliganded structures, LIVBP and OVOT. These two structures were aligned with respect to LAO (open form) by deleting residues {1–10, 40–50, 60–70, 200–230, and 250–280} in LIVBP and {81–90, 124–144} in OVOT. The global mode shapes of the three proteins exhibit similar patterns, departing from the generic behavior of the liganded structures (Fig. 6).

## CONCLUSION

The four structurally similar but functionally different proteins, representative of four distinct classes, CysB (88–324), LAO, PBGD, and RBP, are found to obey a unique global mode shape (displayed in Fig. 6) indicative of a general mechanism of action dictated solely by their common architecture. These proteins do indeed have a common function of binding substrates to the cleft between their two domains, while their surface-exposed segments interact with nucleotides or membrane complexes. The Rossmann fold organization has indeed been pointed out to represent a robust scaffold upon which to develop a repertoire of versatile ligand-binding loops (Tyrrell et al., 1997). The present analysis supports this view. The Rossmann-like fold imparts a well-defined character, not only in structure, but also in dynamics; and the local peculiarities of the different proteins do not significantly affect the observed generic mode shape. The common hinge-bending role of particular

sites, listed in Table 3, are subsumed by different types of residues that are the counterparts of one another in the different aligned structures.

The determination of a common mode shape provides, on the other hand, a means of identifying the structural elements whose dynamic characteristics do not conform to the general consensus. An alignment of the residues of these proteins from different families was performed, based on observed global mode shapes, which divulged the approximate stretches of residues having unique distinctive dynamic features. The global modes are adequately superimposable by deleting only a few segments, mainly residues 18–27 and 185–195 in LAO, 46–62 and 220–245 in PBGD, and residues 78–98, 137–144, and 177–206 in RBP, apart from the C-terminal segments of ~20 residues in each case. These are implicated by the present analysis to be involved in the specific functions of recognition, binding, catalytic activity, or signal transduction. That the loop 18–27 is involved in substrate binding in LAO is verified by comparing the closed (ligand-bound) and open (unliganded) forms of the protein. Likewise, the variability in the size of the connection, such as residues 191–195 in LAO, may arise from the need to accommodate different size substrates in the cleft. These were interestingly observed to assume a hinge-bending role, along with their kinetically hot character.

The alignment of structurally similar but functionally distinct proteins based on their global dynamics does not aim at constructing optimal structural alignments. Well-established structural alignment methods do exist, which can be readily resorted to. Our purpose here was to demonstrate that it is alternatively possible to make an alignment based on dynamic characteristics, rather than structural similarities, and more importantly, such an analysis might stipulate, within similar structures, the regions responsible for functional diversity. We have not yet performed a systematic analysis, nor did we propose an automated method for alignments based on dynamics. Yet it is worth noting the possibility of making such comparisons—taking advantage of efficient approaches such as the GNM—while organizing or clustering structures in data bases and assigning functions to particular structural regions.

Interestingly, the dimerization of CysB(88–324) is observed to affect the relative amplitudes of motion, establishing a balance in the mobilities of the two domains and inducing an increased flexibility at the extended region 291–324 appended to domain I. This is a surface-exposed region, comprising two helices, reminiscent of nucleotide-binding HTH motifs. Its counterpart in the LacR family of proteins ensures dimerization by forming a four-helix bundle (Tyrrell et al., 1997). The role of this element in CysB remains to be established.

Finally, it is worth noting that the open and closed forms of LAO exhibit significant departures in their global mode shapes (Fig. 9 a). This is remarkable, because the secondary structural elements are the same, and the two structures

deviate in their  $\alpha$ -carbon coordinates by 6.4 Å only. Yet the implication of ligand binding for the global dynamics of the molecule is larger than its effect on the structure. Likewise, the behaviors of the two other proteins with similar folds, unliganded LIVBP and apo-OVOT, conform to that of unliganded LAO. It is known that in many ligand-receptor systems, the ligand plays a signaling role and appears not to have any other function. The ligand is not metabolized to useful products, is not an intermediate in any cellular activity, and has no enzymatic properties. The only function of these ligands appears to be to change the structural properties of the protein it binds to, whether dynamic or static, which is then activated or deactivated. The present analysis demonstrates that only after ligand binding, with the molecule assuming its closed conformation, are the hinge-bending role of specific elements and the enhanced mobility of others imparted. These particular changes lead to relative flexibilities conforming precisely to the generic mode shape (Fig. 6) of the presently investigated SBPs.

In summary, the findings of this paper are as follows: 1) Members of the same fold family exhibit common dynamic behavior on a global scale, such that their lowest frequency (most cooperative) mode shapes are closely superimposable, except for a few segments exhibiting distinct behavior. 2) These segments, distinguished by their unique dynamic behavior, are precisely those conferring the specific functional properties of the particular protein. 3) Ligand binding, dimerization (in CysB), or a third domain juxtaposition (in PBGD) helps to confer the generic mechanism of global motion. Two corollaries important for protein design and engineering are as follows: 1) It is possible to utilize the same scaffold (3-D fold) for diverse functions, and 2) it is possible to identify/recognize functionally important regions of a structure as those segments whose behaviors depart from that of the generic global mode shape for a given fold.

The finding—that a given protein architecture dictates the global dynamic behavior of all members of a protein family—is general. This invites one to consider more broadly the consensual global dynamic behaviors of all protein fold families, such as those defined in the SCOP (Murzin et al., 1995) or CATH (Orengo et al., 1997) data bases.

## REFERENCES

- Bahar, I., A. R. Atilgan, M. C. Demirel, and B. Erman. 1998a. Vibrational dynamics of folded proteins: significance of slow and fast modes in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.
- Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential. *Folding Des.* 2:173–181.
- Bahar, I., B. Erman, R. L. Jernigan, A. R. Atilgan, and D. G. Covell. 1999. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. *J. Mol. Biol.* 285:1023–1037.
- Bahar, I., and R. L. Jernigan. 1998. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase bound forms. *J. Mol. Biol.* 281:871–884.

- Bahar, I., and R. L. Jernigan. 1999. Subunit communication, processing and cooperative fluctuations in tryptophan synthase. *Biochemistry*. 38: 3478–3490.
- Bahar, I., A. Wallqvist, D. G. Covell, and R. L. Jernigan. 1998b. Correlation between native-state hydrogen exchange and cooperative residue fluctuations from a simple model. *Biochemistry*. 37:1067–1075.
- Binnie, R. A., H. Zhang, S. Mowbray, M. A. Hermodson. 1992. Functional mapping of the surface of *Escherichia coli* ribose-binding protein: mutations which affect chemotaxis and transport. *Protein Sci*. 1:1642–1651.
- Demirel, M. C., A. R. Atilgan, R. L. Jernigan, B. Erman, and I. Bahar. 1998. Identification of kinetically hot residues in proteins. *Protein Sci*. 7:2522–2532.
- Flory, P. J. 1976. Statistical thermodynamics of random networks. *Proc. R. Soc. Lond. A*. 351:351–380.
- Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett*. 79:3090–3093.
- Holm, L., and C. Sander. 1994. The FSSP database of structurally aligned protein fold families. *Nucleic Acids Res*. 22:3600–3609.
- Holm, L., and C. Sander. 1996. Mapping of the protein universe. *Science*. 273:595–560.
- Jernigan, R. L., M. C. Demirel, and I. Bahar. 1999. Relating structure to function through the dominant modes of motion of DNA topoisomerase II. *Int. J. Quant. Chem*. 75:301–312.
- Kang, C. H., W. C. Shin, Y. Yamagata, S. Gokcen, F. L. Ames, and S. H. Kim. 1991. Crystal structure of the lysine-, arginine-, ornithine-binding protein (LAO) from *Salmonella typhimurium* at 2.7 Å resolution. *J. Biol. Chem*. 15:23893–23899.
- Laskowski, R. A., E. G. Hutchinson, A. D. Michie, A. C. Wallace, M. L. Jones, and J. M. Thornton. 1997. PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci*. 22:488–490.
- Louie, G. V., P. D. Brownlie, R. Lambert, J. B. Cooper, T. L. Blundell, S. P. Wood, M. J. Warren, S. C. Woodcock, and P. M. Jordan. 1992. Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. *Nature*. 359:33–39.
- Maiorov, V., and R. Abagyan. 1997. A new method for modelling large-scale rearrangements of protein domains. *Proteins*. 27:410–424.
- Mizutani, K., H. Yamashita, H. Kurokawa, B. Mikami, and M. Hirose. 1999. Alternative structural state of transferrin—the crystallographic analysis of iron-loaded but domain-opened ovotransferrin N-lobe. *J. Biol. Chem*. 274:10190–10194.
- Mowbray, S. L., and L. B. Cole. 1992. 1.7 Å x-ray structure of the periplasmic ribose receptor from *Escherichia coli*. *J. Mol. Biol*. 225: 155–175.
- Murzin, A. G., S. E. Brenner, T. Hubbard, and C. Chothia. 1995. Scop: a structural classification of protein database for the investigation of sequences and structures. *J. Mol. Biol*. 247:536–540.
- Oh, B.-H., J. Pandit, C.-H. Kang, K. Nikaïdo, S. Gokcen, G. F.-L. Ames, and S.-H. Kim. 1993. Three-dimensional structures of the periplasmic lysine-, arginine-, ornithine-binding protein with and without a ligand. *J. Biol. Chem*. 268:11348–11355.
- Ordal, G. W. 1985. Bacterial chemotaxis: biochemistry of behavior in a single cell. *CRC Crit. Rev. Microbiol*. 12:95–130.
- Orengo, C. A., A. D. Michie, S. Jones, D. T. Jones, M. W. Swindells, and J. M. Thornton. 1997. CATH—a hierarchical classification of protein domain structures. *Structure*. 5:1093–1108.
- Sack, J. S., M. A. Saper, and F. A. Quioco. 1989a. Periplasmic binding protein structure and function. Refined x-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *J. Mol. Biol*. 206:171–191.
- Sack, J. S., S. D. Trakhanov, I. H. Tsigannik, and F. A. Quioco. 1989b. Structure of the L-leucine-binding protein refined at 2.4 Å resolution and comparison with the Leu/Ile/Val-binding protein structure. *J. Mol. Biol*. 206:193–207.
- Stewart, R. C., and F. W. Dahlquist. 1988. N-terminal half of CheB is involved in methylesterase response to negative chemotactic stimuli in *Escherichia coli*. *J. Bacteriol*. 170:5728–5738.
- Tyrrell, R., K. H. G. Verschuere, E. J. Dodson, G. N. Murshudov, C. Addy, and A. J. Wilkinson. 1997. The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure*. 5:1017–1032.