Specificity of Ligand Binding in a Buried Nonpolar Cavity of T4 Lysozyme: Linkage of Dynamics and Structural Plasticity[†,‡]

Andrew Morton and Brian W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, and Departments of Chemistry and Physics, University of Oregon, Eugene, Oregon 97403

Received February 9, 1995; Revised Manuscript Received April 21, 1995[⊗]

ABSTRACT: To better understand the role of shape complementarity in ligand binding and protein core interactions, the structures have been determined of a set of ligands bound within a cavity-containing mutant of T4 lysozyme. The interior cavity is seen to consist of two parts that respond very differently to the binding of ligands. First, there is a relatively rigid region that does not relax significantly upon binding any ligand. Second, there is a more flexible region that moves to various extents in response to binding the different ligands. The part of the binding site that remains rigid is characterized by low temperature factors and strong protection from hydrogen exchange. This part of the site appears to be primarily responsible for discriminating between ligands of different shape (i.e., for determining specificity). The more flexible region, characterized by relatively high temperature factors and weak protection from hydrogen exchange, allows some promiscuity in binding by undergoing variable amounts of deformation at essentially the same energetic cost. This linkage between the dynamic information represented by crystallographic temperature factors and hydrogen-exchange behavior on the one hand, and structural plasticity in response to ligand binding on the other hand, suggests a way to improve our understanding of steric interactions in protein cores and protein−ligand binding sites. Ligand design and packing algorithms might take advantage of this information, requiring complementary interactions where the protein is rigid and allowing some overlap in regions where the protein is flexible.

Interactions between residues in a protein core and between a ligand and its binding site are strongly influenced both by the hydrophobic effect and by shape-dependent packing effects. The preceding paper describes a quantitative analysis of the relative contributions of these effects, using a system involving binding of simple nonpolar ligands to a buried nonpolar cavity in T4 lysozyme (Morton et al., 1995). It was shown that while the relative importance of packing and hydrophobicity depends on the structural class of the ligand, in general the specificity of the system is most sensitive to differences in packing energetics. In the prior analysis (Morton et al., 1995) the free energy of packing was defined in two different ways, $\Delta G_{PAK,O}$ relative to an octanol reference state and $\Delta G_{PAK,V}$ relative to the vapor phase. In the present paper, the term "packing energy" will refer exclusively to $\Delta G_{PAK,O}$. This corresponds to the work required to reorganize the binding site to accommodate the ligand (ΔG_{REORG}) plus the interaction between the ligand and the reorganized site ($\Delta G_{P:L}$) relative to the interaction between the ligand and octanol ($\Delta G_{O:L}$) [see Figure 6 of Morton et al. (1995)].

Current understanding of shape-dependent interactions is reflected in computer algorithms which attempt to predict core packing arrangements or protein−ligand interactions, such as PROPAK (Ponder & Richards, 1987) and DOCK (Kuntz et al., 1994). Both of these algorithms rely heavily on shape criteria to predict allowable combinations of functional groups. In the case of core interactions, recent work has shown that protein main-chain flexibility acts to relax the constraints of shape on packing interactions (Baldwin et al., 1993; Lim et al., 1994), but the details of where and how this flexibility arises are still uncertain. In the case of ligand design, impressive results from shape-based algorithms such as DOCK have been obtained, even in cases where the shape of the docking site is not accurately known (Ring et al., 1993). In some cases of structure-based design where the structure of the predicted complex has been determined, changes in protein structure are observed (Shoichet et al., 1993; Mattos et al., 1994). On the basis of these results, it would seem that any method to simplify the prediction of where and how protein elements will display structural plasticity would be useful.

In this light, we present X-ray crystal structures of complexes of the Leu 99 → Ala mutant of T4 lysozyme (L99A) with nine different ligands whose packing energetics were quantitatively analyzed in the accompanying paper. We observe that ligand binding causes various amounts of reorganization of the binding site relative to its conformation in the apoprotein, as well as large increases in its flexibility. The possible linkage of binding site flexibility and structural reorganization is discussed.

EXPERIMENTAL PROCEDURES

The L99A mutant protein (also referred to as the apoprotein) was purified and crystallized essentially as described (Eriksson et al., 1993). The crystals were maintained in a mother liquor consisting of 2.3 M phosphate and 0.23 M sodium chloride, pH 6.7. Crystals were equilibrated with ligands by vapor diffusion for 3−10 days prior to data collection (Eriksson et al., 1992a). X-ray diffraction data were collected using a Xuong-Hamlin area detector and

[†] A.M. was a Howard Hughes Predoctoral Fellow. This work was supported in part by NIH Grant GM21967 to B.W.M.

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (access codes T6278−T6285).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1995.

0006-2960/95/0434-8576$09.00/0

structures refined using TNT (Tronrud et al., 1987; Tronrud, 1992). For each complex, the binding of the ligand was first confirmed from inspection of the $F_{PL} - F_P$ electron density maps, where F_P and F_{PL} are the observed structure amplitudes for the L99A protein and the protein−ligand complex. The protein was then refined against the structure amplitudes of the complex using ~15 rounds of conjugate direction minimization of coordinates and temperature factors. $F_{PL} - F_C$ and $2F_{PL} - F_C$ electron density maps were examined, where F_C is the calculated structure amplitude of the protein. (For the ethyltoluenes, low occupancy of the binding site resulted in unclear density. These ligands also refined with very high temperature factors (60−100 Å^2) and are not discussed here.) The protein−ligand complex was then refined to convergence. The expected bond lengths and angles of each ligand for use in crystallographic refinement were determined using energy minimization of bond lengths and bond angles with the MM2 force field contained in the program ALCHEMY (Tripos Associates). Comparisons of the refined X-ray models were made after overlaying each coordinate set onto the model for the apoprotein based on the C^α atoms within the carboxy-terminal domain (residues 80−160). These and other coordinate manipulations were performed using the program EDPDB (Zhang & Matthews, 1995). Cavity size was estimated with the program MS (Connolly, 1983) using a probe radius of 1.2 Å.

RESULTS

Crystallographic Analysis of Complexes. X-ray data and refinement statistics for nine L99A−ligand complexes are collected in Table 1. Figure 2 shows representative difference Fourier electron density maps. Each of these nine ligands produced strong density upon binding, indicating that the occupany is generally high. This is also indicated by the thermal factors of the ligands, which are only slightly above those of the adjacent atoms in the protein core (Table 1). In most cases the difference density maps (Figure 2) show negative density in the vicinity of helix F (residues 109−113), indicating increased disorder in this region. In some cases, accompanying positive features indicate positional shifts. The final refined coordinates of the complexes are compared to those of the apoprotein in Figures 3 and 4. The thermal factors are compared in Figure 5. Comparative illustrations of the ligand conformations and of the protein structural changes are presented in Figures 6 and 7, respectively. Coordinates were deposited in the Brookhaven Protein Data Bank.

The Cavity. To facilitate direct comparisons, the crystal structures of the L99A apoprotein and the L99A−benzene complex were rerefined in parallel with the complexes reported below. They are essentially identical to those reported by Eriksson et al. (1992a,b). The cavity has a volume of 153 Å^3 and contains no unpaired hydrogen bond donors or acceptors, with the exception of S^δ of Met 102, which is not hydrogen bonded in either the wild-type or L99A apoprotein structure (Figure 1).

The cavity has the general shape of a benzene molecule plus an additional "bulge" (cf. Figure 1) sufficient to accommodate a benzene molecule with a single small substituent. Model building based on the benzene complex suggested that an ethyl group would fit, while a propyl group would require a highly strained dihedral angle to avoid bad contacts. Accommodation of groups larger than propyl would require some accommodation by the protein. Any multiply-substituted benzene derivative would require either significant reorganization of the protein, an altered binding configuration relative to the observed benzene configuration, or both.

(1) Benzene. The structure of the L99A−benzene complex is essentially identical to that previously described (Eriksson et al., 1992a). The only significant movement relative to the apoprotein is a slight adjustment (0.2−0.3 Å) of helix F (Figure 4a). This is accompanied, however, by a significant increase in thermal factors over the same region (Figure 5a).

(2) Ethylbenzene. The difference map (Figure 2g) and refined ethylbenzene complex (Figure 3g) show that ethylbenzene binds in a position similar to that of benzene. The ethyl group occupies the extra bulge in the cavity, as anticipated (see also Figure 6c). Binding is accompanied by slightly larger shifts in the protein than in the benzene case, again centered on helix F (Figure 4a,g). The thermal factors in helix F are also higher in this complex than in the benzene complex (Figure 5a,g). There is some evidence for disorder of the ligand in the binding site: thermal factors of the atoms within the phenyl ring are roughly equal (20−24 Å^2) and are comparable to those of the side chains lining the cavity (~15−30 Å^2), but those of the atoms in the ethyl substituent are significantly higher (45−50 Å^2). This disorder is consistent with the difference map (Figure 2g) which shows only poor density for the ethyl side chain.

(3) n-Butylbenzene and Isobutylbenzene. *n*-Butylbenzene is the least soluble of all the ligands studied (~0.1 mM). This may limit the saturation of the binding sites in the crystals. The ligand has thermal factors significantly higher than all the others (Table 1) and presumably binds with less than 100% occupancy. The X-ray data are not of sufficient resolution to refine occupancy and thermal factors independently, nor did we attempt simultaneous refinement of bound and unbound protein conformations.

n-Butylbenzene binds in a configuration similar to that adopted by ethylbenzene. In both cases the phenyl ring binds in approximately the same location as benzene alone (Figure 6c). Movements of the protein are modest, occurring mainly in helix F (Figure 4h). This movement appears sufficient to accommodate the relatively long butyl group. Although close contacts remain in the refined structure (e.g., 3.2 Å to Val 111), they involve atoms whose thermal factors are very high. As in the case of ethylbenzene, the temperature factors within helix F increase substantially (Figure 5h).

In the case of isobutylbenzene, however, the isobutyl group apparently cannot be accommodated by small movements in helix F. The difference map (Figure 2i) and refined coordinates (Figures 3i, 4) indicate movements of the protein backbone in this region by as much as 2 Å. The thermal factors of helix F are very similar for isobutyl- and *n*-butylbenzene. Thus, these two ligands, although having similar hydrophobicities and similar binding energies, appear to cause different amounts of reorganization in the protein. It should be noted, however, that the high thermal factors of helix F indicate substantial mobility and corresponding coordinate uncertainty in both complexes. $F_{PL} - F_C$ difference maps calculated from the final refined coordinates show positive difference features at positions other than the atoms of helix F, suggesting that the protein in this region may adopt multiple conformations.

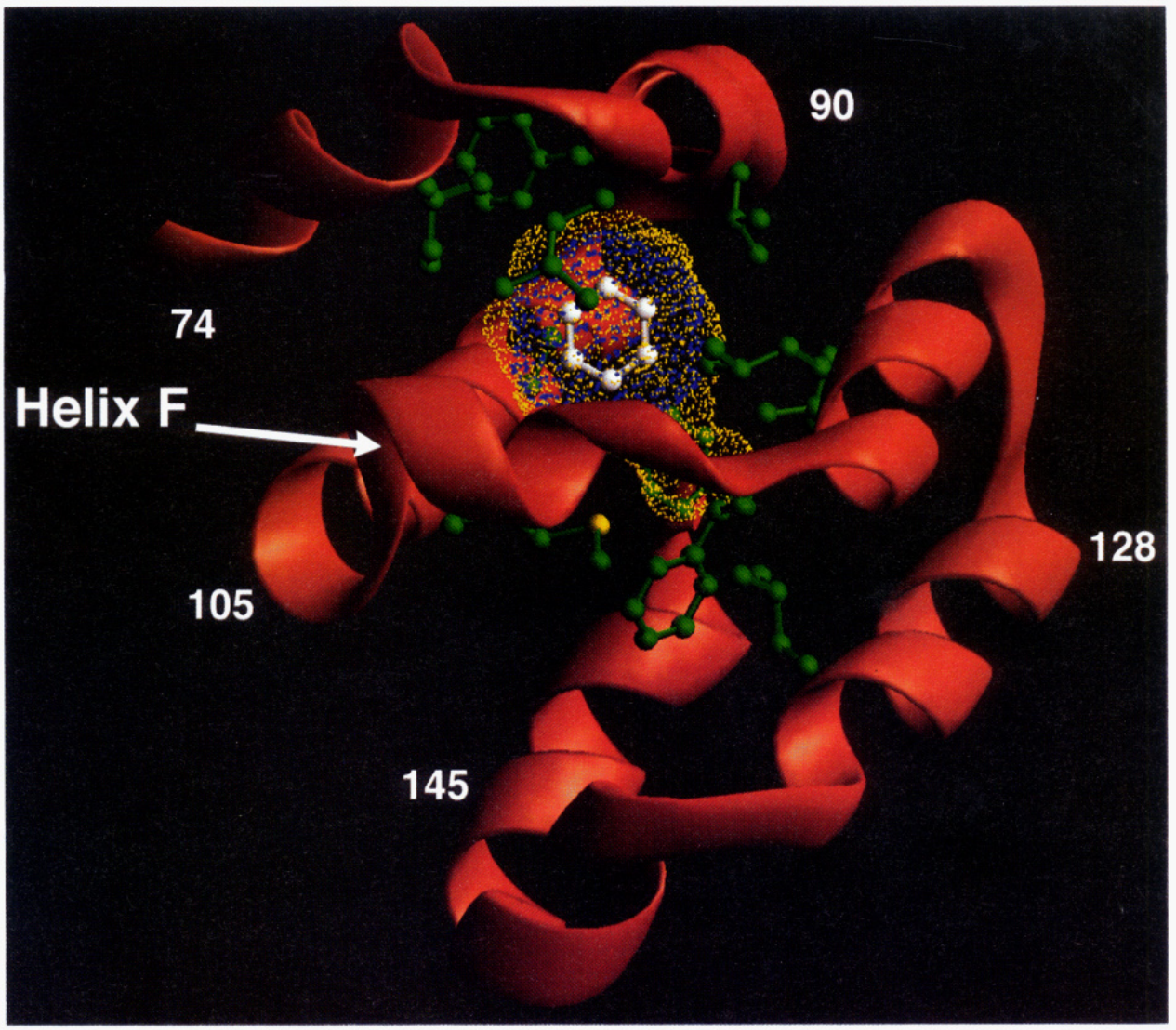


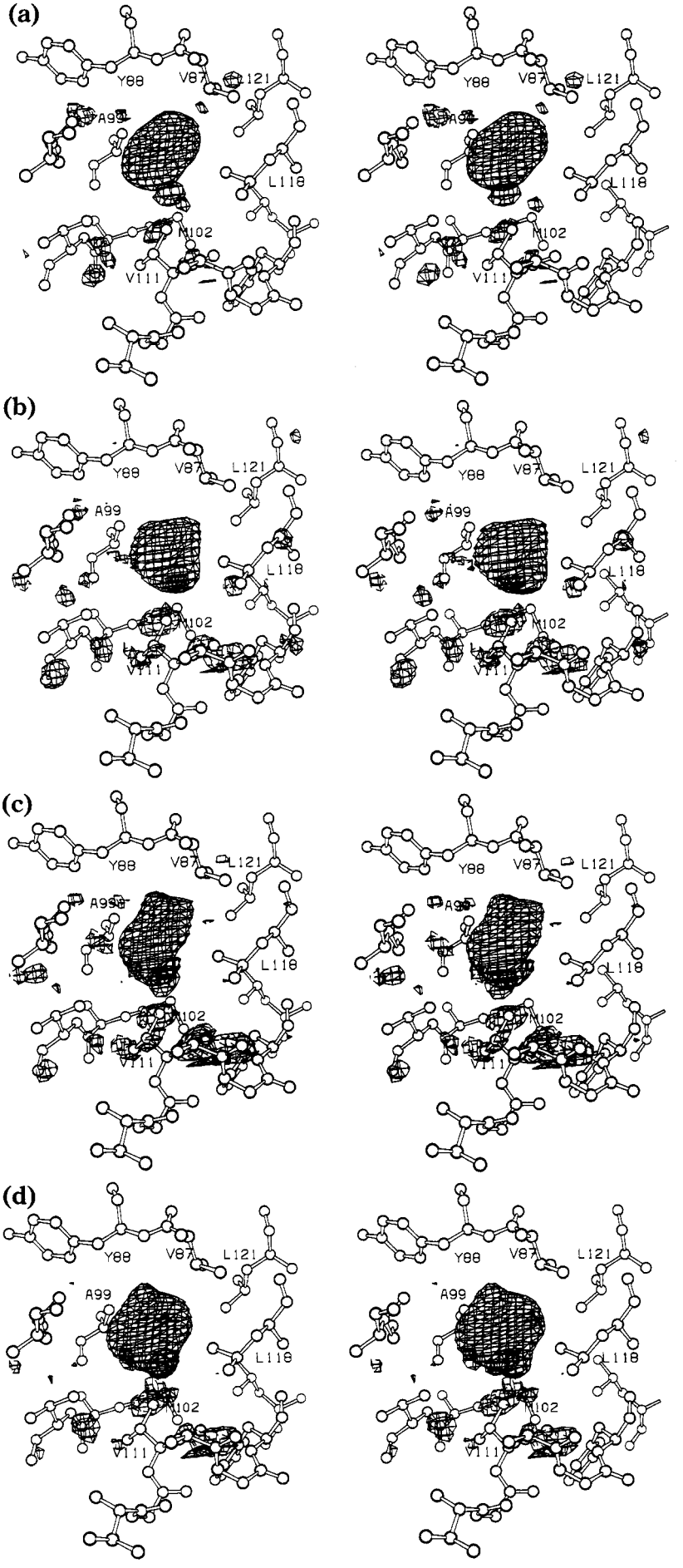
FIGURE 1: Structure of the complex of benzene with the L99A mutant of T4 lysozyme. Only the carboxy-terminal domain of the protein is shown. The molecular surface of the cavity (gold) and the van der Waals surface of the benzene (blue) are represented as dot surfaces. The bound benzene is shown in white. The side chains which line the cavity are represented as color-coded ball-and-stick models, where nitrogen is blue, oxygen is red, sulfur is yellow, and carbon is green. Helix F runs horizontally through the middle of the figure.

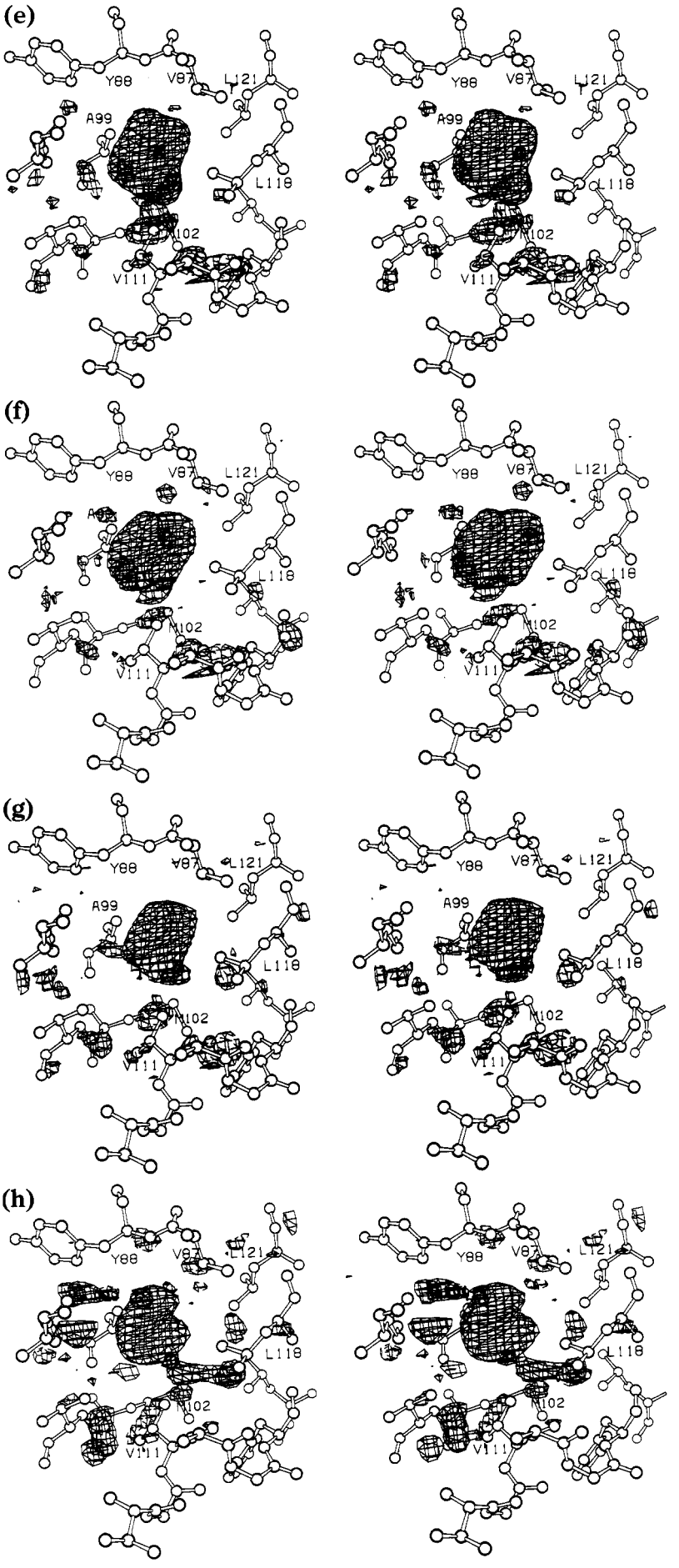
(4) p-Xylene and o-Xylene. Initial electron density maps ($F_{PL} - F_P$; see Experimental Procedures) did not unambiguously indicate the position of these ligands. After initial refinement with no ligand in the model, these ligands were positioned on the basis of $F_{PL} - F_C$ maps, as well as difference maps between the data sets for these ligands and the benzene complex data set.

p-Xylene binds with movement of the main chain of the protein up to ~0.6 Å (Figure 4b), and the thermal factors of helix F increase markedly (Figure 5b). Also, the side chain of Val 111 rotates approximately 140°, to a χ_1 value of −35°, to avoid a close contact with the ligand. The ligand itself is also shifted ~1 Å relative to the position occupied by benzene and the monosubstituted benzenes. In the case of *o*-xylene, the position occupied by the ligand corresponds more closely to that of benzene (Figure 6a), but in this case binding causes a larger reorganization of helix F, with backbone shifts up to 2 Å (Figure 4c), very similar to those of isobutylbenzene.

(5) Benzofuran, Indene, and Indole. These three isosteric ligands bind in significantly different configurations in their respective complexes, corresponding to rotations of up to 41°. The ligand configurations are coplanar to within 10° (Figure 6b). In the indole complex the protein movements are less than ~0.7 Å (Figure 4f), with slight increases in the thermal values of helix F, very similar to those seen in the benzene complex. The nitrogen of the indole ring is 3.6 Å from S^{δ} of Met 102, comparable to the distances of 3.3−3.6 Å observed for NH···S hydrogen bonds (Saenger, 1984). Indene binds in a different orientation (compare panels f and e of Figure 3; see also Figure 6b) and causes significant movement in helix F (Figure 4e), similar to the isobutylbenzene and *o*-xylene structures. Benzofuran binds in a conformation intermediate between the other two (Figures 3d, 6b).

Some question exists as to the exact orientation of these molecules, because of their approximate 2-fold symmetry. The disposition of the nitrogen in the indole molecule was confirmed by examination of a difference map (not shown) for 5-fluoroindole. No such fluorinated derivative was available for benzofuran, however. On the basis of the distribution of thermal factors of benzofuran after refinement in four possible orientations, we have placed the oxygen so that it is directed away from S^{δ} of Met 102. Regardless of


(a)
Y88
V87
L121
A99
L118
M102
V111
(b)
Y88
V87
L121
A99
L118
M102
V111
(c)
Y88
V87
L121
A99
L118
M102
V111
(d)
Y88
V87
L121
A99
L118
M102
V111


(e)
Y88
V87
L121
A99
L118
N102
V111
(f)
Y88
V87
L121
L118
N102
V111
(g)
Y88
V87
L121
A99
L118
N102
V111
(h)
Y88
V87
L121
L118
N102
V111

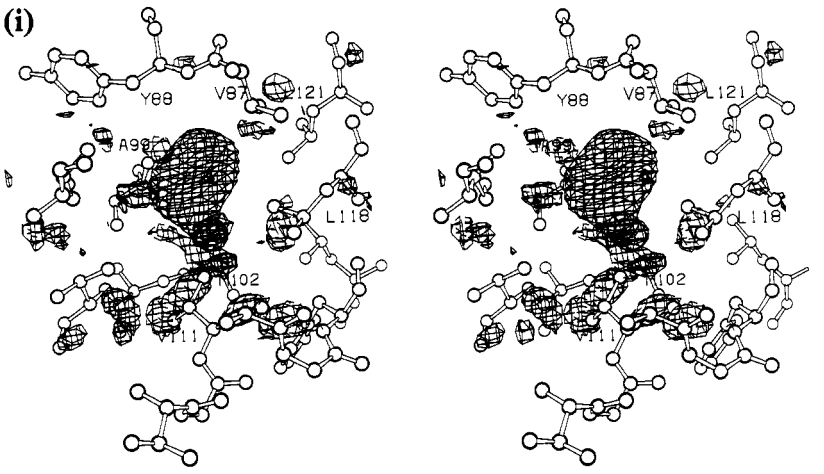


FIGURE 2: Electron density maps calculated with amplitudes ($F_{PL} - F_P$), where F_{PL} and F_P are respectively the observed amplitudes for the ligand−L99A complex and for the apoprotein (i.e., L99A). Phases are from the refined structure of L99A. The map for the ligand complex is superimposed on the refined L99A apoprotein coordinates. Maps are calculated at 2.1 Å resolution and contoured at +3.5 (solid lines) and −3.5 (broken lines) standard deviations. Panels: (a) benzene; (b) *p*-xylene; (c) *o*-xylene; (d) benzofuran; (e) indene; (f) indole; (g) ethylbenzene; (h) *n*-butylbenzene; (i) isobutylbenzene.

which orientation is correct, the oxygen of the bound benzofuran cannot make a hydrogen bond to a neighboring atom of the protein.

DISCUSSION

The X-ray crystal structures of the L99A−ligand complexes described here reveal a relation between the structural and dynamic consequences of ligand binding which explains the specificities among the various classes of ligands (Morton et al., 1995). The most striking features common to all the protein−ligand complexes are the changes in the dynamic and structural character of helix F (residues 109−113). These changes play a central role in determining the interactions between the ligands and the binding site. For every protein−ligand complex, the temperature factors in helix F are significantly higher than the corresponding values for the apoprotein, indicating that increased disorder in this region of the protein is associated with binding regardless of the specific interactions made between each individual ligand and the binding site (Figure 5). Helix F is also the site of the largest changes in atomic positions caused by binding, although the magnitude of these changes ranges widely for the various complexes (Figure 4). The extent of positional shift is not correlated with the structural class of the ligand. Rather, each of the three classes contains at least one ligand that causes large positional shifts, as well as other ligands that cause relatively modest shifts.

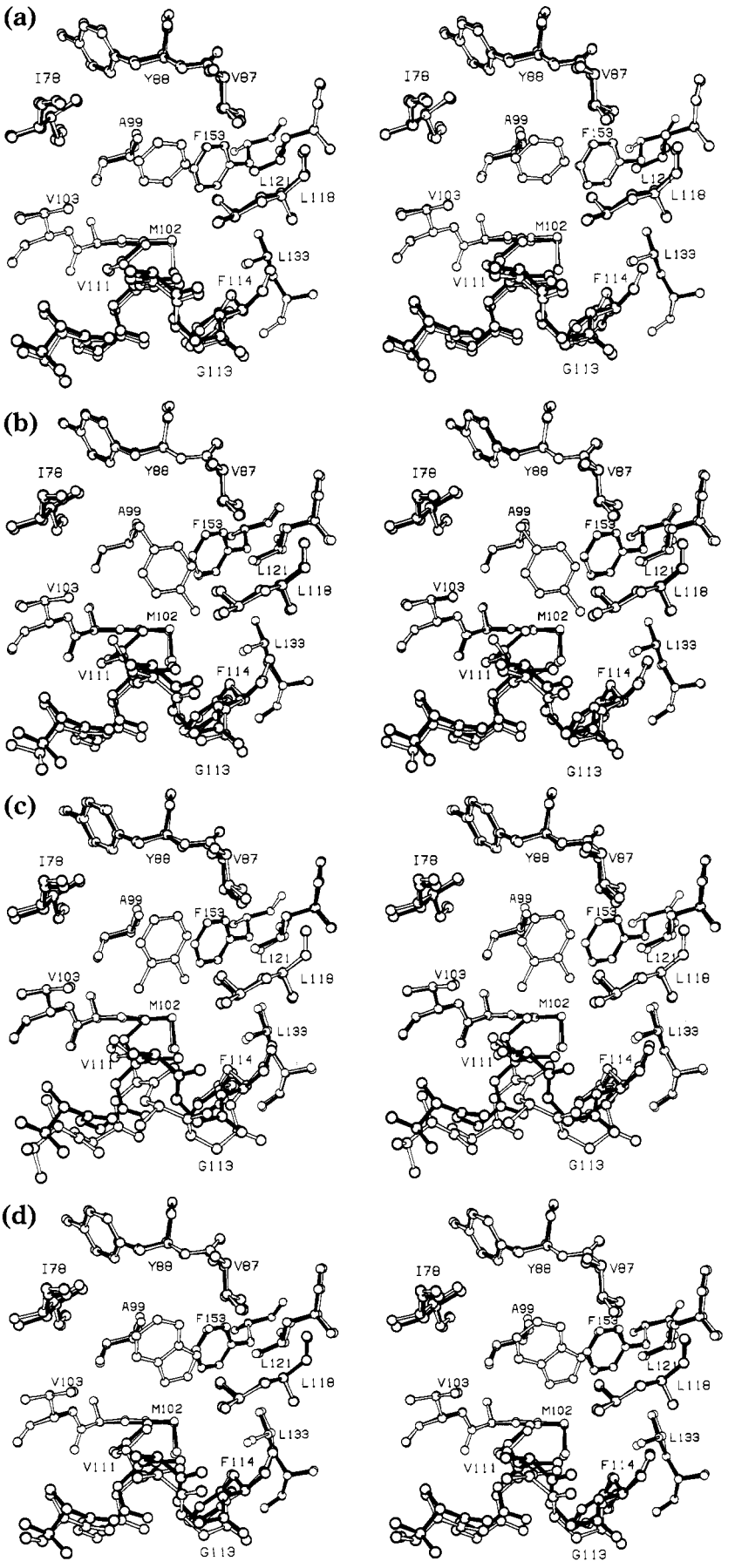
The behavior of helix F may be explained by a two-part mechanism, termed *induced flexibility*. Binding of any ligand at this site leads to a conformational destabilization of helix F, as reflected in its increased temperature factors in all complexes. As a result of this destabilization, the helix displays increased mobility and is apparently able to undergo large positional shifts (~1.5−2.5 Å) with relatively low cost in free energy (Morton et al., 1995). The differences in the affinities of the various ligands are thus determined largely by their interactions with the remaining, relatively rigid, portion of the binding site.

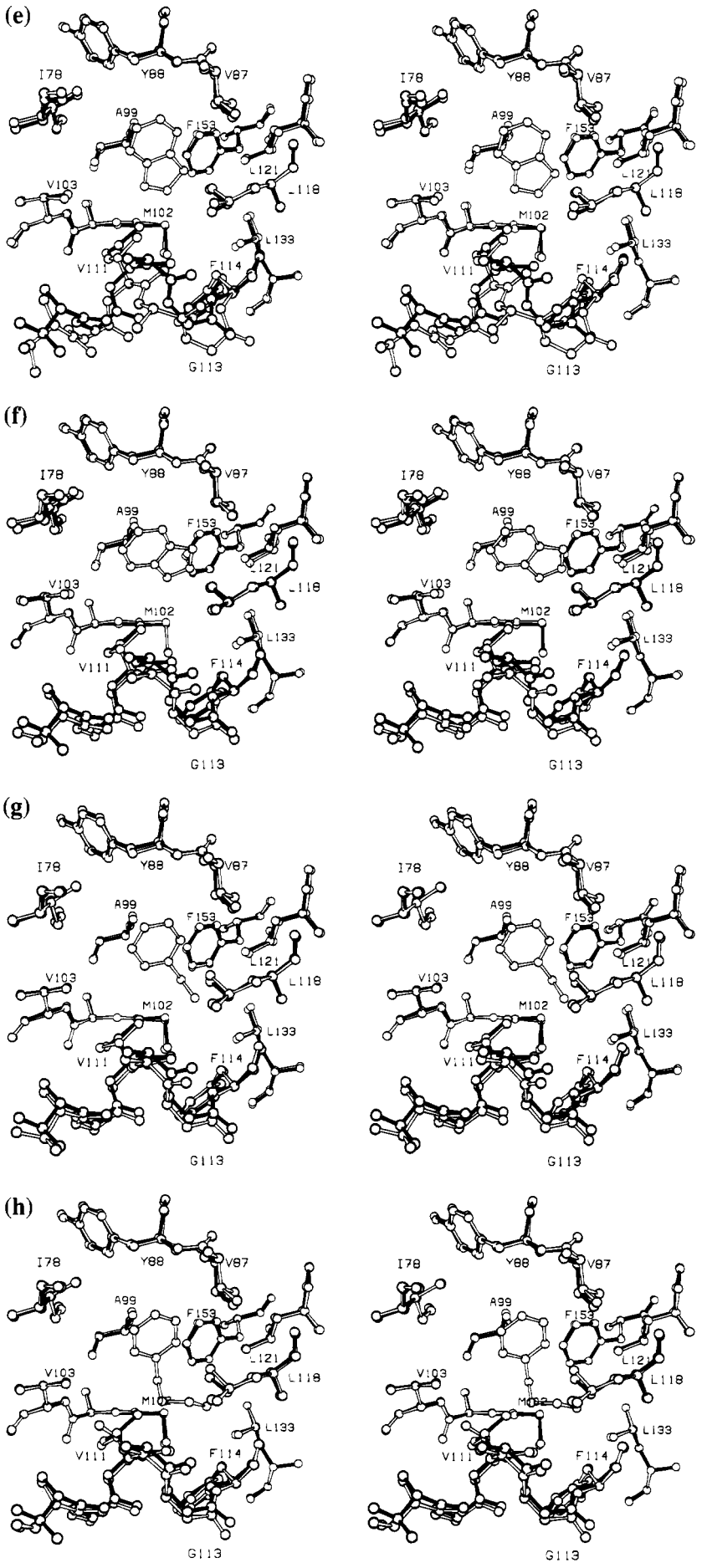
If such linkages between structural and dynamic effects are common, they may provide an important clue to understanding the well-documented role played by structural plasticity in core repacking [e.g., Baldwin et al. (1993)] and protein−ligand interactions [e.g., Bone et al. (1989) and Varadarajan and Richards (1992)]. We consider the roles played by helix F and the remainder of the binding site for each class of ligand in turn.

Helix F. Helix F rests on the protein surface, where it connects two antiparallel α-helices (E and G) which form part of the four helix bundle of the domain architecture. The helix is quite short, consisting of only seven residues (~1.5 turns), two of which (110 and 113) are glycine. Within the helix are five intrahelical hydrogen bonds, of which two are 3_{10} and three are α. In the apoprotein the main-chain atoms within helix F have a mean thermal factor of 31 Å^2, whereas the other helices involved in forming the binding site, helices D, E, and G, have mean main-chain thermal factors of 18, 16, and 20 Å^2, respectively. Hydrogen-exchange experiments also indicate that helix F is among the most dynamic secondary structure elements in the protein (F. W. Dahlquist, personal communication). Arnold and Ornstein (1992) performed a molecular dynamics simulation and found that helix F experienced the greatest fluctuations of any helix in the C-terminal domain of T4 lysozyme. These features suggest that helix F is intrinsically less rigid than a typical helix.

The Isophobes (Class I). The binding analysis (Morton et al., 1995) suggests that the differences in binding free energies among the class I isophobic ligands arise primarily from differences in the packing of the ligands in the binding site, including the cost of reorganizing the binding site to accommodate the ligands, rather than from differences in the hydrophobicity of the ligands. In particular, the packing interactions made by *o*-xylene and *p*-xylene are less favorable by 1.8 and 2.2 kcal/mol, respectively, than for ethylbenzene (Morton et al., 1995). [As noted in the introduction, "packing interactions" refer to $\Delta G_{PAK,O}$ and correspond to the energy of deformation of the protein plus the ligand−site interaction relative to ligand−octanol interaction; see eq 8 of Morton et al. (1995).]

The observed crystal structures reveal that the changes in structure are least for ethylbenzene, slightly greater for *p*-xylene, and substantially larger for *o*-xylene (Figure 4). All three complexes show significant increases in temperature factors within helix F (Figure 5). The simplest interpretation


(a)
I78
Y88
V87
A99
F153
L121
L118
V103
M102
L133
V111
F114
G113
(b)
(c)
(d)


(e)
I78
Y88
V87
A99
F153
L121
L118
V103
M102
L133
V111
F114
G113
(f)
I78
Y88
V87
A99
F153
L121
L118
V103
M102
L133
V111
F114
G113
(g)
I78
Y88
V87
A99
F153
L121
L118
V103
M102
L133
V111
F114
G113
(h)
I78
Y88
V87
A99
F153
L121
L118
V103
L133
V111
F114
G113

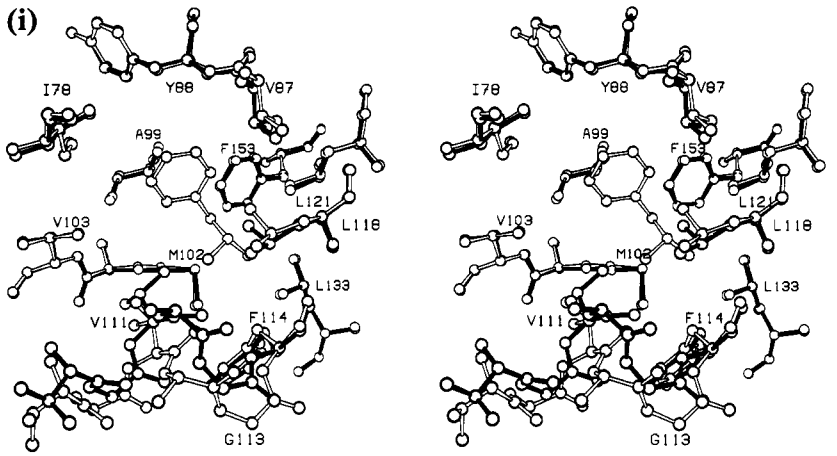


FIGURE 3: Comparison of refined coordinates of the ligand−L99A complexes (open bonds) with the apoprotein (solid bonds), after superposition of each complex onto the apoprotein. The superposition is based on root-mean-square minimization of the discrepancies between the α-carbon atoms in the C-terminal domains (i.e., the domain that includes the L99A cavity). Panels: (a) benzene; (b) *p*-xylene; (c) *o*-xylene; (d) benzofuran; (e) indene; (f) indole; (g) ethylbenzene; (h) *n*-butylbenzene; (i) isobutylbenzene.

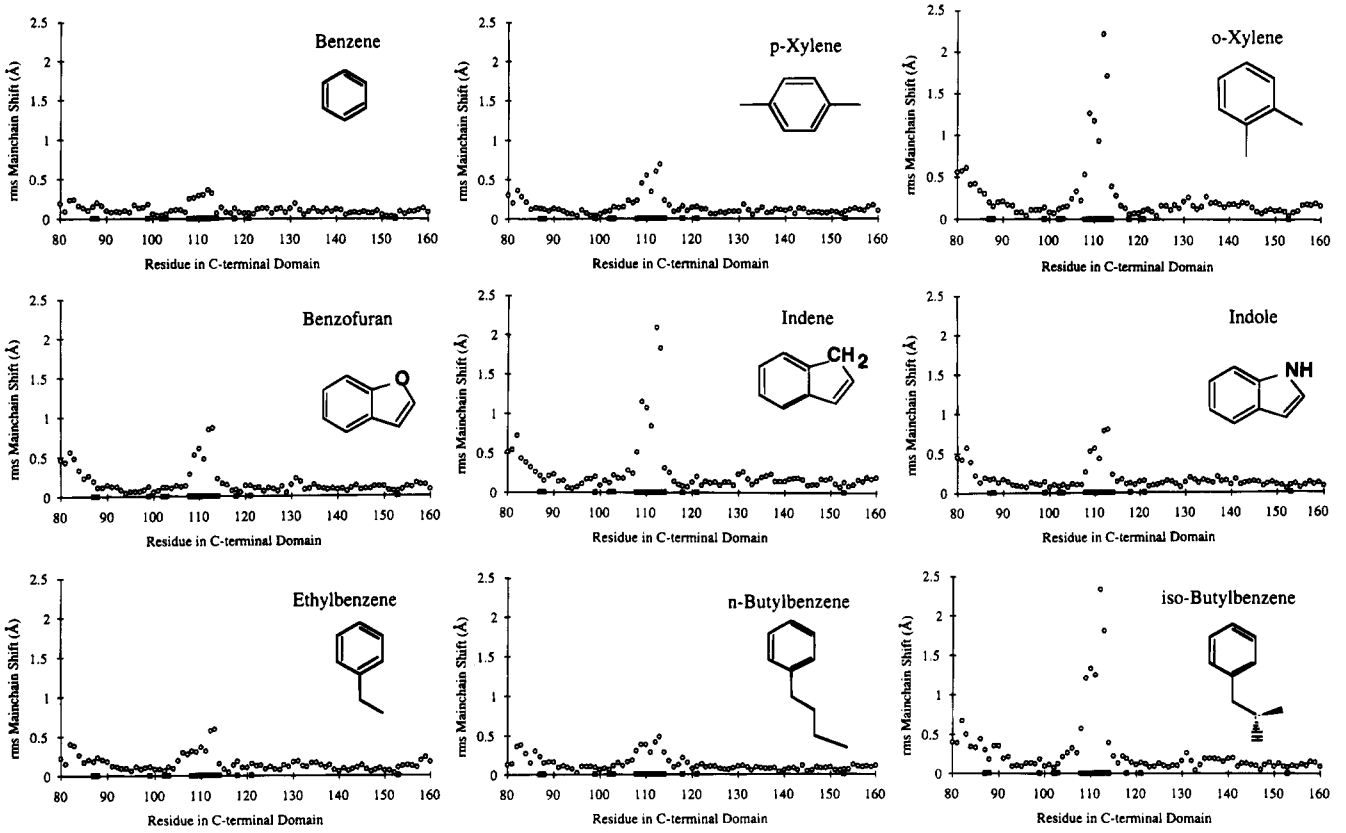


FIGURE 4: Main-chain positional shifts in the complexes relative to the apoprotein. The value plotted for each residue is the root-mean-square average of the shift of the backbone atoms. The complexes were superimposed on the basis of the carboxy-terminal domain as in Figure 3. (The amino-terminal domains are not included in the comparison and are not included in the figure). Residues that comprise the cavity are indicated by the solid rectangles at the bottom of the figure. Panels (from left to right): (a) benzene; (b) *p*-xylene; (c) *o*-xylene; (d) benzofuran; (e) indene; (f) indole; (g) ethylbenzene; (h) *n*-butylbenzene; (i) isobutylbenzene.

of these results is that the relatively large differences between the packing energies of ethylbenzene and the two xylenes arise from the way in which these molecules interact with the main part of the cavity, which is benzene-shaped and relatively rigid. This rigid part of the cavity can accommodate well the singly substituted phenyl ring of ethylbenzene but not so well the doubly substituted rings of the xylenes. While binding of these ligands is permitted by their translation toward the flexible region of the site (Figure 6a), this still results in poor packing interactions with the rigid part of the site. The large shift of helix F in the *o*-xylene complex relative to the other complexes corresponds to only a very small change in packing energy. Thus, in the present case, differences in packing energy are determined more by the interactions between the ligands and the site than by the reorganization of the site in response to ligand binding. This situation arises because of the induced flexibility of that part of the protein which undergoes reorganization, helix F.

The Isosteres (Class II). The isosteric ligands of class II were chosen to directly measure the effect of differences in

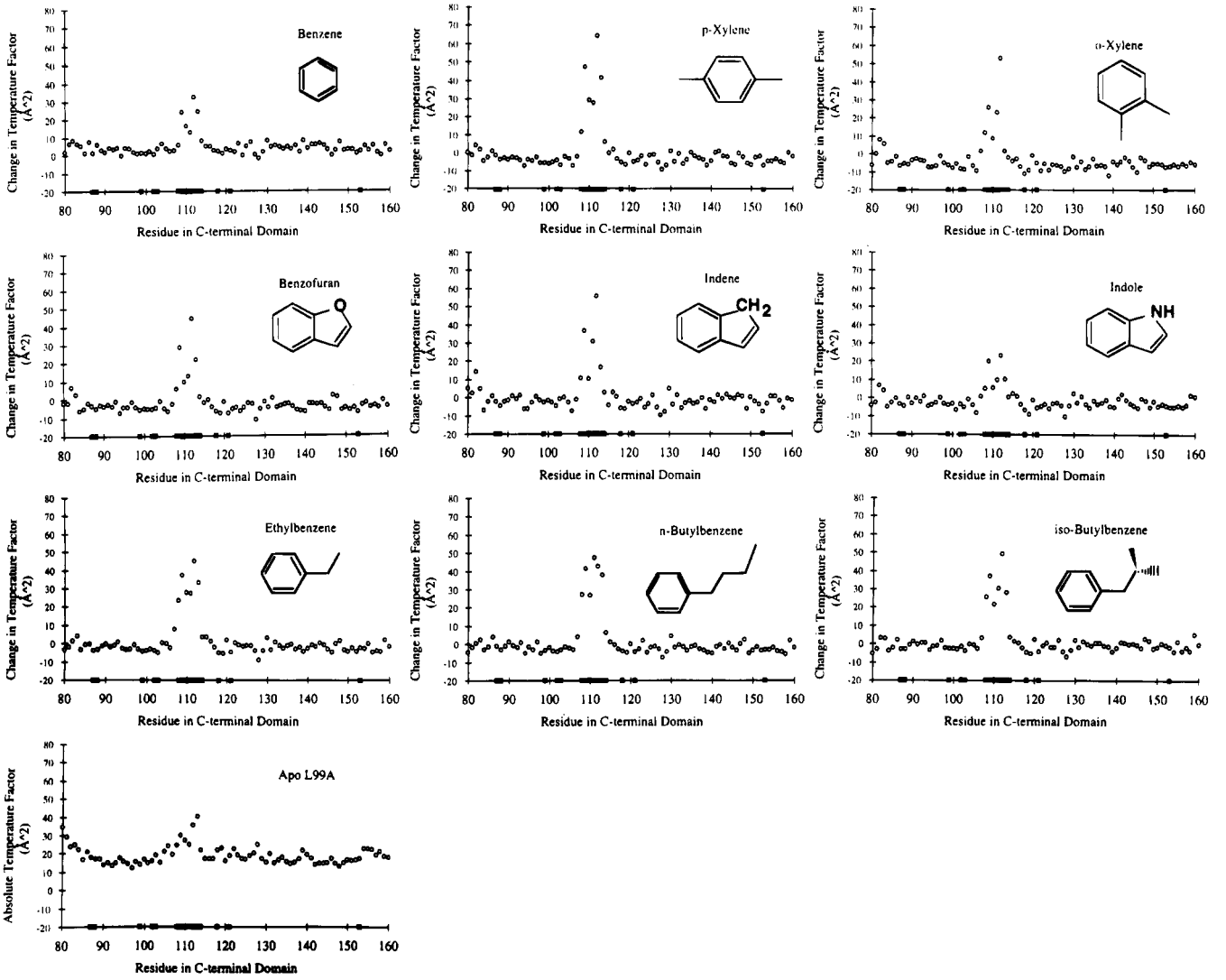


FIGURE 5: Change in the backbone thermal factors of the complex structures relative to those of the apoprotein. The value plotted for each residue is the root-mean-square average of the change in crystallographic thermal factor of the main-chain atoms. Residues that comprise the cavity are indicated by the solid rectangles at the bottom of the figure. Panels (from left to right): (a) benzene; (b) *p*-xylene; (c) *o*-xylene; (d) benzofuran; (e) indene; (f) indole; (g) ethylbenzene; (h) *n*-butylbenzene; (i) isobutylbenzene; (j) apo L99A.

Table 1: Crystallographic Data Collection and Refinement Statistics[a]

ligand	benzene	indene	benzofuran	indole	*o*-xylene	*p*-xylene	ethylbenzene	*n*-butylbenzene	isobutylbenzene
cell									
a, *b* (Å)	60.9	60.7	60.9	60.7	60.8	60.7	60.8	60.9	60.9
c (Å)	97.0	97.2	97.4	97.6	96.7	97.1	97.5	97.0	97.3
reflections	14844	13459	17566	16208	14000	17279	17896	18632	17474
% completeness	74	65	85	79	68	84	87	90	84
resolution (Å)	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
R-factor (%)	15.8	17.7	16.8	16.9	15.6	17.0	16.5	16.4	16.2
Δ_{bond} (Å)	0.014	0.017	0.017	0.015	0.016	0.016	0.015	0.015	0.015
Δ_{angle} (deg)	2.0	2.3	2.1	2.0	2.1	2.0	2.1	2.0	2.1
B(ligand) (Å^2)	24	19	21	17	29	36	28	46	27
B(cavity wall) (Å^2)	22	17	15	16	15	14	18	26	22

[a] The space group is $P3_221$. The unique reflections typically result from 30 000 to 60 000 independent observations. The *R*-factor is the residual between the crystallographically refined model and the observed data. *B*(ligand) is the average thermal factor for atoms in the ligand; *B*(cavity wall) is the average thermal factor for the surrounding atoms in the wall of the cavity.

hydrophobicity on binding free energy in the protein core. Unexpectedly, however, the binding affinities of these ligands are not dominated by their hydrophobicity. Rather, differences in protein−ligand packing interactions appear to be nearly as important (Morton et al., 1995). These results are consistent with the structural analyses in that the binding configurations of indole, benzofuran, and indene are significantly different despite their similar shapes (Figure 6b). Also, the structural changes in the protein range from modest (~0.7 Å for indole) to large (~2 Å for indene) (Figure 4). These differences in ligand configuration and protein conformation may explain the different packing energetics of the "isosteric" ligands. They underscore the danger in assuming that small differences in ligand size and shape will give only small

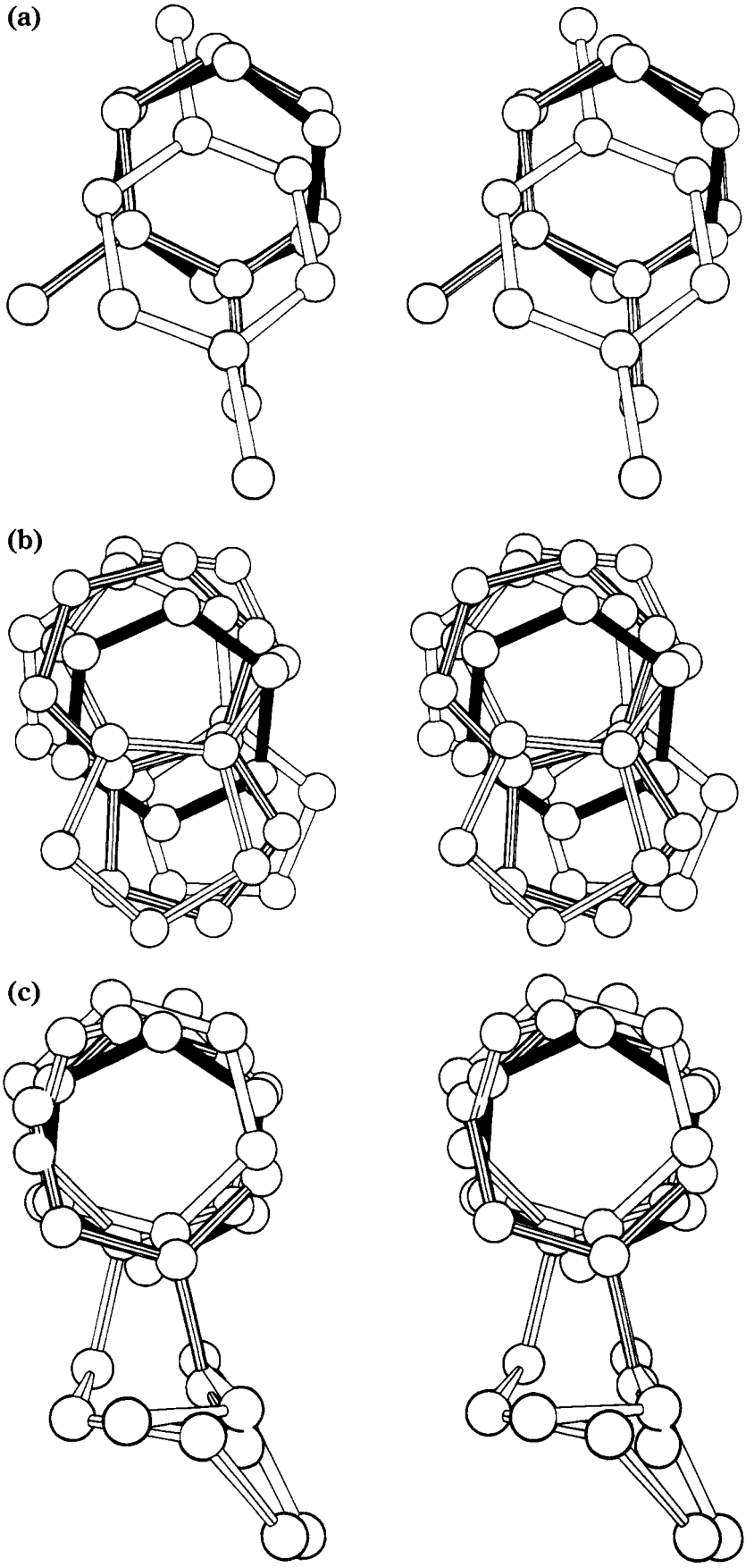


FIGURE 6: Superposition of different classes of ligands to compare their respective modes of binding in the L99A cavity. The ligands are shown on the basis of superposition of the carboxy-terminal domains of the respective protein−ligand complex. Stereo pairs: (a) superposition of benzene, *o*-xylene, and *p*-xylene (darkest to lightest); (b) superposition of benzene, benzofuran, indene, and indole (darkest to lightest); (c) superposition of benzene, ethylbenzene, *n*-butylbenzene, and isobutylbenzene (darkest to lightest).

differences in packing interactions and highlight the need for structural analysis, even when comparing closely related ligands.

Among the class II ligands only indole (a hydrogen bond donor) forms a hydrogen bond in the bound state (to S^{δ} of Met 102). Neither benzofuran nor thianaphthene (potential hydrogen bond acceptors) have partners available in the cavity. Hydrogen bonds made by furans or thiofurans in solution should, however, be much weaker than the hydrogen bond made by the indole. Thus, the penalty caused by losing hydrogen bonds to solvent may be fairly low in the case of benzofuran and thianaphthene and offset by the hydrogen bond to Met 102 in the case of indole. This may explain why hydrogen bonding does not appear to dominate the relative binding energies of these molecules.

In general, crystallographic analysis has shown that the formation of enzyme−inhibitor complexes leads to local *reduction* of mobility (Petsko & Ringe, 1984; Ringe & Petsko, 1985). In the present case, however, binding leads to *increased* mobility, suggesting that the present cavity differs from a typical active site. In this regard, however, we note that indole, the only ligand to make a hydrogen bond in the complex, causes the smallest increase in the mobility of helix F (Figure 5).

The Phenylalkanes (Class III). For benzene, ethylbenzene, and even *n*-butylbenzene the structural changes in the protein upon binding are relatively modest (Figure 4), and the packing energies of these ligands are fairly constant. The presence of the bulky isobutyl group results in much larger structural perturbations and a modest increase in the packing free energy (0.6 kcal/mol relative to *n*-butylbenzene; Morton et al., 1995). The temperature factors in helix F increase markedly in all of these complexes. The relative constancy of the packing free energies is surprising in light of the different sizes of these ligands. For example, the volume of benzene in 80 Å^3 while the volume of *n*-butylbenzene is 160 Å^3. Apparently, either both the reorganization and interaction parts of the packing free energy are relatively constant for these ligands or they vary in a mutually compensating fashion.

The induced flexibility model suggests that the differential cost of reorganization among these ligands is small. All the ligands destabilize the helical conformation of the flexible region upon binding (Figure 5). The comparison of isobutyl- and *n*-butylbenzene, however, suggests that nearly 2 Å of additional movement over several residues can cost as little as 0.6 kcal/mol. This is likely to be an upper estimate for the differential cost of reorganization, since the isobutyl group is not expected to fit into the cavity (as it exists in the respective complexes) significantly better than the *n*-butyl group. Thus, as in the case of the class I (isophobic) ligands, the packing energetics among most of the class III ligands are dominated by the interaction terms rather than the reorganization terms. The difference between the two classes is that in the case of the class III ligands, which all have a singly substituted phenyl ring that complements the rigid portion of the cavity, most of the ligands have similar interaction energies (defined relative to their interaction with octanol solution) and therefore similar packing free energies. The importance of the complementarity between their phenyl rings and the rigid portion of the cavity is demonstrated by the fact that all the class III ligands, including isobutylbenzene, bind in remarkably similar configurations (Figure 6c).

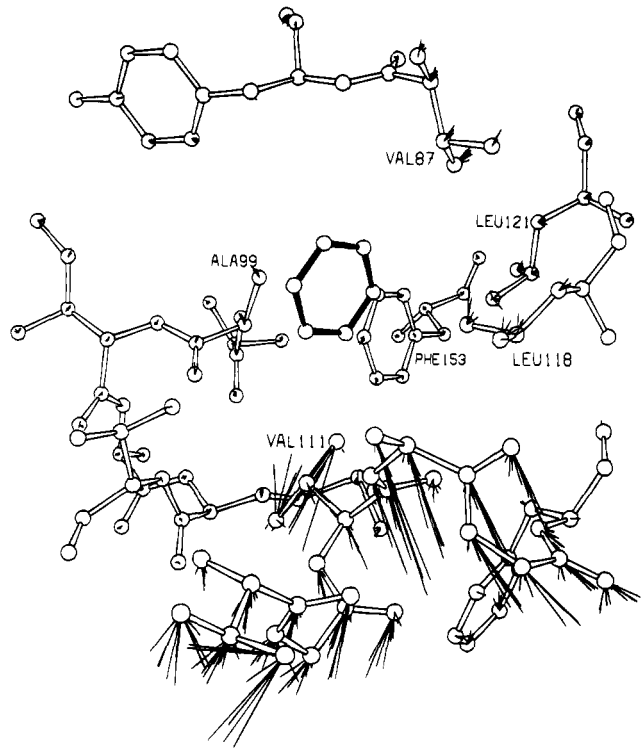


FIGURE 7: Illustration of the location of the structural changes in the L99A binding site upon binding each of nine ligands. Open bonds correspond to the position of atoms in the apoprotein. From each atom nine vectors are drawn, leading to the position of that atom in each of the nine complexes. In some cases, these vectors are shorter than the atomic radii used in the figure. A benzene molecule has been included according to its position in the L99A−benzene complex to indicate the position of the interior cavity.

Implications for Structure-Based Ligand Design. In the absence of specific polar interactions, the binding strength and specificity for most of the ligands studied here are determined by hydrophobicity, entropic, and packing effects (Morton et al., 1995). In principle, the results from this system should allow refinement of the manner in which ligand design algorithms treat hydrophobicity and packing. The hydrophobic effect can be modeled by using the water−octanol tranfer free energies of the different ligands. The consequences of packing forces, however, can be quite subtle. For example, a series of isosteres bind with distinct orientations and cause different degrees of perturbation of the protein structure (Figures 4, 6).

The energetic response of the protein is strongly influenced by its ability to deform in some directions much more easily than others. In the present case, the major mode of relaxation arises from the deformation of helix F. Variations of this deformation are seen in complexes from each of the three ligand classes (Figures 3, 4, 7). The latent mobility of helix F is indicated by the high thermal factors and low hydrogen-exchange protection factors of this region, as well as by its glycine-rich sequence and its structural role in the architecture of the C-terminal domain.

Woodward (1993) has noted a correlation between regions of a protein which appear highly dynamic in hydrogen-exchange measurements and those that respond to mutational perturbation with relatively large protein reorganization. Our results are consistent with this observation. The specific types of deformation that are possible, however, only become apparent in the presence of a variety of perturbations (in this

case different ligands). The advantage of this additional information is demonstrated by the observation that successful structure-based design often utilizes information from one or more protein−ligand complex structures, rather than the structure of the apoprotein alone [e.g., Shoichet et al. (1993)].

The present results hold the hope that ligand design might take advantage of the relation, described above, between the dynamic and structural responses of a binding site, perhaps without the need for predicting the fine details of complex geometry or calculating precise interaction energies. As a first approximation, information from crystallographic thermal factors, hydrogen-exchange protection factors, secondary structure, and amino acid sequence can be used to determine which regions of a binding site are inherently more flexible than others. Ligands with shapes that require protein reorganization only in flexible regions while remaining complementary to more rigid regions of a binding site can be expected to bind better than anticipated on the basis of rigid shape criteria alone. Ligand design algorithms could incorporate this information, requiring complementary ligand−site interactions where the binding site is rigid and allowing some overlap between ligand and binding site in regions where the site is flexible.

ACKNOWLEDGMENT

The authors thank Brian Shoichet, Doug Barrick, and members of the Matthews Laboratory for advice and encouragement, Rebekka Wachter for help with X-ray data collection, and Joan Wozniak for technical assistance.

REFERENCES

Arnold, G. E., & Ornstein, R. L. (1992) *Protein Eng. 5*, 703−714.
Baldwin, E. P., Hajiseyedjavadi, O., Baase, W. A., & Matthews, B. W. (1993) *Science 262*, 1715−1718.
Bone, R., Silen, J. L., & Agard, D. A. (1989) *Nature 339*, 191−195.
Connolly, M. L. (1983) *J. Appl. Crystallogr. 16*, 548−558.
Eriksson, A. E., Baase, W. A., Wozniak, J. A., & Matthews, B. W. (1992a) *Nature 355*, 371−373.
Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992b) *Science 255*, 178−183.
Eriksson, A. E., Baase, W. A., & Matthews, B. W. (1993) *J. Mol. Biol. 229*, 747−769.
Kuntz, I. D., Meng, E. C., & Shoichet, B. K. (1994) *Acc. Chem. Res. 27*, 117−123.
Lim, W. A., Hodel, A., Sauer, R. T., & Richards, F. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 423−427.
Mattos, C., Rasmussen, B., Ding, X., Petsko, G. A., & Ringe, D. (1994) *Struct. Biol. 1*, 55−57.
Morton, A., Baase, W. A., & Matthews, B. W. (1995) *Biochemistry 34*, 8564−8575.
Petsko, G. A., & Ringe, D. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 331−371.
Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
Ring, C. S., Sun, E., McKerrow, J. H., Lee, G. K., Rosenthal, P. J., Kuntz, I. D., & Cohen, F. E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3583−3587.
Ringe, D., & Petsko, G. A. (1985) *Prog. Biophys. Mol. Biol. 45*, 197−235.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Shoichet, B. K., Stroud, R. M., Santi, D. V., Kuntz, I. D., & Perry, K. M. (1993) *Science 259*, 1445−1450.
Tronrud, D. E. (1992) *Acta Crystallogr. A48*, 912−916.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−503.
Varadarajan, R., & Richards, F. M. (1992) *Biochemistry 31*, 12315−12327.
Woodward, C. K. (1993) *Trends Biochem. Sci. 18*, 359−360.
Zhang, X.-J., & Matthews, B. W. (1995) *J. Appl. Crystallogr.* (in press).

BI950298J