PROTEINS: Structure, Function, and Genetics 39:261–268 (2000)

Side-Chain Flexibility in Proteins Upon Ligand Binding

Rafael Najmanovich,* Josef Kuttner, Vladimir Sobolev, and Marvin Edelman
Plant Sciences Department, Weizmann Institute of Science, Rehovot, Israel

ABSTRACT **Ligand binding may involve a wide range of structural changes in the receptor protein, from hinge movement of entire domains to small side-chain rearrangements in the binding pocket residues. The analysis of side chain flexibility gives insights valuable to improve docking algorithms and can provide an index of amino-acid side-chain flexibility potentially useful in molecular biology and protein engineering studies. In this study we analyzed side-chain rearrangements upon ligand binding. We constructed two non-redundant databases (980 and 353 entries) of "paired" protein structures in complexed (holo-protein) and uncomplexed (apo-protein) forms from the PDB macromolecular structural database. The number and identity of binding pocket residues that undergo side-chain conformational changes were determined. We show that, in general, only a small number of residues in the pocket undergo such changes (e.g., ~85% of cases show changes in three residues or less). The flexibility scale has the following order: Lys > Arg, Gln, Met > Glu, Ile, Leu > Asn, Thr, Val, Tyr, Ser, His, Asp > Cys, Trp, Phe; thus, Lys side chains in binding pockets flex 25 times more often then do the Phe side chains. Normalizing for the number of flexible dihedral bonds in each amino acid attenuates the scale somewhat, however, the clear trend of large, polar amino acids being more flexible in the pocket than aromatic ones remains. We found no correlation between backbone movement of a residue upon ligand binding and the flexibility of its side chain. These results are relevant to 1. Reduction of search space in docking algorithms by inclusion of side-chain flexibility for a limited number of binding pocket residues; and 2. Utilization of the amino acid flexibility scale in protein engineering studies to alter the flexibility of binding pockets. Proteins 2000;39:261–268.** © 2000 Wiley-Liss, Inc.

Key words: database analysis; docking predictions; side-chain conformation; binding pocket; holo-protein; apo-protein

INTRODUCTION

Knowledge of the structure of a ligand-protein complex assists in understanding enzymatic mechanisms and provides a rational approach to drug design and engineering of binding pockets. Such structural information is valuable to solve problems in fields as diverse as medicine, agronomy and industry. A significant number of protein structures have been determined by X-ray crystallography in both complexed and uncomplexed forms and are available from the macromolecular protein databank.[1] Comparison of these structures is valuable for revealing general features of ligand binding.

Ligand binding may induce rearrangements of the protein structure, in particular, side-chain movements. Knowledge of the extent with which side-chain rearrangements occur is therefore important to developing improved docking prediction algorithms. Predicting the structure of the complex of a small ligand with a protein (molecular docking) is still a complex task, the two major problems being the definition of an appropriate scoring function of the candidate structures and the size of search space. Assuming that one knows the correct scoring function (as discussed in Petrella et al.[2]), a successful searching procedure should consider the three factors that give rise to the size of search space: the relative positions of ligand and receptor, the different ligand conformations and protein flexibility.

Initial docking procedures dealt with both ligand and receptor as rigid bodies.[3–5] Some approaches included ligand flexibility.[6–9] Presently, several approaches include, to some extent, also protein flexibility.[10–18] Ligand binding may induce large structural changes in the receptor protein, such as the movement of loops[19,20] or even large domains.[21,22] Nevertheless, in most cases changes in backbone structure are negligible and only side-chain reorientation (if any) occur upon ligand binding.[23,24] Consequently, combinatorial approaches making use of side-chain rotamer libraries are considered very efficient. Accounting for side-chain reorientation during docking procedures is a similar task to that of predicting side-chain conformations in homology modelling.[25–27]

A fast method for finding the global minimum (or maximum) of the scoring function in side-chain rotamer space uses the dead-end elimination theorem.[28–31] Later, it was shown that this method could miss the global minimum during the searching procedure. Its modification, based on the fuzzy-end elimination theorem,[32,33] corrects this problem but the size of the search space becomes huge. Thus, additional information that allows the restriction of search space can be very valuable.

In this paper we analyze to what extent amino acid side-chains belonging to the binding pocket undergo conformational changes upon binding of a small ligand. We constructed a complete database (3,827 entries) of protein

*Correspondence to: Rafael Najmanovich, Plant Sciences Department, Weizmann Institute of Science, Rehovot 76100, Israel. E-mail: rafael.najmanovich@weizmann.ac.il

Received 3 September 1999; Accepted 20 January 2000

structures in complexed (holo-protein) and uncomplexed (apo-protein) forms from the PDB macromolecular structural databank of March 1999 from which two different non-redundant databases (980 and 353 entries) were defined. The number and type of binding pocket undergoing side-chain conformational changes in holo- and apo-proteins were then analyzed.

METHODS

Database Creation

The first step in creating our database was the selection of ligands. In the PDB, a ligand is described by its three-letter code name and listed as HETATM or ATOM in the coordinates section of an entry. Our analysis is restricted to ligands listed as HETATM (thus excluding nucleic acids and peptides that are listed as ATOM). Furthermore, we excluded from our analysis ligands that are covalently bound to protein atoms as well as PO_4^- and SO_4^- molecules. If different parts of a single ligand appear with different codes they are considered as separate ligands. No distinction was made between ligands diffused into the protein crystal or co-crystallized.

The binding pocket is defined as consisting of those amino acids in contact with ligand atoms.[34] "Interatomic contact" refers to the contact surface between atoms.[35] Two binding pockets are considered different if the list of residues in contact with the ligand differs by one or more residues.

Our goal is to use the present analysis to improve docking algorithms. At the atomic level, numerous locations can be found at the protein surface having surface complementarity for very small ligands. Similarly, non-specific locations can be found for ligands that form only a small number of contacts. We therefore imposed a minimum number of five heavy atoms (viz. non-hydrogen) to the ligand as well as a minimum number of five contact residues between the ligand and the protein. The minimum number of atoms required for the ligand excludes atomic ions and water molecules.

A PDB entry that contains a ligand (as defined above) is termed holo-protein with respect to the ligand in consideration. To be considered as an apo-protein (of a given holo-protein), a PDB entry must have an identical amino acid sequence as that of the holo-protein, and none of its binding pocket residues can be in contact with another ligand that is not also present in the given holo-protein. The last requirement ensures that the only difference between the two proteins in the region of the binding pocket is due to the ligand being considered. Only PDB entries determined by X-ray crystallography to a resolution equal or better than 2.5 Å were used in the present study.

Definition of an entry in our database

An entry in our database is composed of the holo-protein PDB entry identifier, the associated three-letter ligand identifier and an apo-protein PDB entry identifier (also see legend to Fig. 1, below).

Side Chain Rearrangement

For the purpose of the present analysis, we compared the value of side-chain dihedral angles for binding pockets residues in both holo- and apo-protein entries. We define a conformational change to have occurred if a difference larger than a threshold value exists for at least one dihedral angle.

Several studies stress that in known protein structures a particular dihedral angle could differ from the angle determined by the torsional energy minima or from the statistical average by more than $\pm 40°$ and still belong to the same rotamer.[2,30] It then follows that occasionally a difference of $\sim 80°$ can be found for a particular angle with the side-chains still belonging to the same rotamer. However, in the great majority of instances a difference of $\sim 80°$ would indicate structures belonging to different rotamers. For completeness, our analyses were performed at three different threshold values: 45°, 60°, and 75°. The trends for all were very similar. As expected, the higher the threshold value the smaller the number of binding pocket residues which undergo side-chain conformational change. However, the differences are not pronounced (e.g., the percentage of binding pockets with up to three flexible residues was about 80% at a threshold of 45°, 85% at 60°, and 90% at 75°). Moreover, the probability for a specific amino acid to change side-chain conformation upon ligand binding is insensitive to variation in the threshold. Thus, for clarity, we present only the results for the threshold value of 60°. A similar threshold value has been used in recent analysis of amino acid conformational changes in protein association.[36]

RESULTS AND DISCUSSION

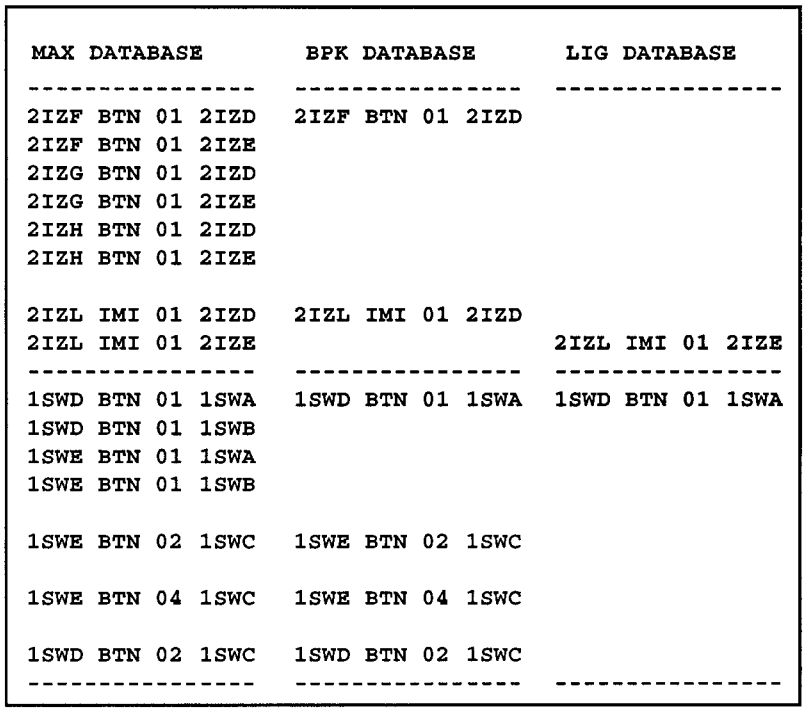
Databases

Maximal database

Using the definitions and rules discussed in Methods, we built a database with 3,827 entries. This maximal (MAX) database is the largest one that meets our criteria. It contains 221 different protein sequences, defining a total of 980 binding pockets and 353 different ligands. The MAX database is composed of blocks of entries, with all PDB files in a given block having the same amino acid sequence (i.e., all deriving from the same protein). Every block is further divided into sub-blocks, with all PDB files in a given sub-block being grouped together because they have the same ligand and list of binding-pocket residues. Blocks and sub-blocks may be composed of single entries. A small section of the MAX database is shown in Figure 1 to help visualize how the apo- and holo-structures are paired for analysis in the derived, non-redundant databases.

The MAX database is large due to the combinatorial nature of the process used to create it: 1. A given protein (apo- and/or holo-form) might have been crystallized several times with the same ligand(s) but under different conditions, giving rise to different PDB entries; or 2. The same PDB entry might be considered as a holo-protein for a given ligand but as apo-protein for another ligand, thus increasing the total number of entries in the database.

The MAX database is clearly redundant yet useful as a



MAX DATABASE	BPK DATABASE	LIG DATABASE
2IZF BTN 01 2IZD	2IZF BTN 01 2IZD	
2IZF BTN 01 2IZE		
2IZG BTN 01 2IZD		
2IZG BTN 01 2IZE		
2IZH BTN 01 2IZD		
2IZH BTN 01 2IZE		
2IZL IMI 01 2IZD	2IZL IMI 01 2IZD	
2IZL IMI 01 2IZE		2IZL IMI 01 2IZE
1SWD BTN 01 1SWA	1SWD BTN 01 1SWA	1SWD BTN 01 1SWA
1SWD BTN 01 1SWB		
1SWE BTN 01 1SWA		
1SWE BTN 01 1SWB		
1SWE BTN 02 1SWC	1SWE BTN 02 1SWC	
1SWE BTN 04 1SWC	1SWE BTN 04 1SWC	
1SWD BTN 02 1SWC	1SWD BTN 02 1SWC	

Fig. 1. Database structure. An entry in each of the three databases is composed of a holo-protein PDB ID (e.g., 2IZF), followed by a modified ligand ID (e.g., BTN 01) and, finally, a apo-protein PDB ID (e.g., 2IZD). The number in the modified ligand ID numbers the ligand in the order of its appearance in the PDB file. All MAX database entries between two consecutive dashed lines form a block which represents the same protein sequence, while empty lines are used to separate entries, forming sub-blocks, that differ either in the ligand or the set of residues in contact with the ligand. The BPK database was created by choosing one entry from each sub-block (in the example shown, the first entry in each case). The LIG database was created by choosing one entry for each three-letter ligand ID.

starting point for studies related to ligand binding that require structural comparisons between apo- and holo-forms. In this study, we have derived two subsets from the MAX database, each with different minimal redundancy criteria.

Binding pocket database

The binding pocket database (BPK) consists of one entry for each different binding pocket found in the MAX database. Any two entries in this database must differ in at least one of three factors: the protein sequence, the ligand and the binding pocket. In terms of the structure of the database, applying the above criteria is equivalent to choosing one entry from each sub-block in the MAX database (see Fig. 1). In the particular analysis presented, we used the first entry appearing in each sub block. However, the results were essentially the same when we used different choices (data not shown). The BPK database consists of 980 entries.

Ligand database

The ligand database (LIG) was created by randomly selecting a single entry for each different ligand present in the MAX database. Having only a single entry for each ligand, stringently avoids the possibility that a set of binding-pocket contacts will be counted more than once. Either because the ligand defining the binding pocket does not duplicate all the contacts in two different entries, or two protein sequences differ in a small number of distal amino acids which do not influence the structure of the binding pocket. On the other hand, in this way we may lose cases in which a given ligand binds genuinely to different binding pockets, as well as cases where the same ligand was crystallized with decidedly different proteins. The LIG database consists of 353 entries.

TABLE I. Database Characteristics

Database	Entries[a]	PDB files[b]	Unique protein sequences
MAX	3827	998	221
BPK	980	729	221
LIG	353	473	154

[a]Number of holo- and apo-protein pairs present in each database. The difference between the number of entries and the number of unique protein sequences is mainly due to the frequent existence of several independent PDB entries for the same apo- (or holo-) form of a protein (see example in Fig. 1).
[b]Number of different PDB files used to build each database irrespective of their role (apo or holo) in each database.

A potential source of redundancy that was not excluded in either of our sub-databases is that of structural similarity of ligands. Two ligands are considered different if their three-letter PDB codes are not the same. However, there might be several such ligands that are structurally or functionally very similar. It is very difficult to judge in advance the effect on binding of even a single atom difference between ligands.

We present further information with respect to the composition of the three databases in Table I. The three databases mentioned in this study can be obtained at the following URL: *http://sgedg.weizmann.ac.il/ferafael/ligdb.html.*

Analyses

Two questions were addressed: 1. How many binding pocket residues of paired protein structures undergo side

TABLE II. Distribution of Flexible Side Chains

Amino acid[a]	Rotatable bonds	BPK database Total	BPK database Flexible side chains[c]	LIG database Total	LIG database Flexible side chains[c]
Arg[b]	4	1199	285	418	113
Asn	2	686	78	248	24
Asp	2	893	51	327	19
Cys	1	204	4	108	3
Gln	3	474	107	205	51
Glu	3	795	109	308	48
His	2	862	45	370	27
Ile	2	708	89	273	43
Leu	2	932	118	381	52
Lys	4	653	247	274	111
Met	3	196	48	92	15
Phe	2	616	9	311	5
Ser	1	751	46	366	25
Thr	1	875	63	338	30
Trp	2	674	14	205	4
Tyr	2	955	71	337	23
Val	1	731	61	349	27

[a]Side chain conformational changes in proline were not considered since they are invariably accompanied by changes in backbone conformation. Glycine and alanine do not have any rotatable bonds.
[b]Rotation of the NE-CZ bond in Arg was not considered because the CD, NE, CZ, NH1, NH2 atoms form a structure close to planar that practically does not change in shape under various conditions (atom names follow PDB nomenclature).
[c]The probability to undergo side-chain conformational change for each amino acid is obtained by dividing the number of flexible side chains by the total number of observed side chains. We assume that this probability is a quantitative measure of flexibility.

chain conformational changes? 2. Which amino acids are more likely to undergo such changes (i.e., which amino acids are more flexible upon ligand binding)?

Binding pocket flexibility

We determined the frequency of side-chain conformational changes occurring in binding pocket residues upon ligand binding. Only dihedral angles of rotatable side-chain bonds were considered, thus excluding alanine and glycine from the present analysis. Side-chain conformational changes in proline were not considered since they are invariably accompanied by changes in backbone conformation. The number of rotatable bonds for each amino acid as well as the total number of residues studied is indicated in Table II.

The distributions of the number of amino acid residues that undergo side-chain conformational change for the BPK and LIG databases are shown in Figure 2. The probability of a pocket to have N residues undergoing conformational change decreases asymptotically with N in such a way that changes in up to three residues account for ~85% of the cases (inset, Fig. 2). This result supports restriction of side-chain flexibility to a small number of residues in docking predictions. In addition, 94.4% of χ_1 angles and 95.7% of χ_2 angles (irrespective of the amino acid) do not undergo conformational change. These results are somewhat higher than those presented by Betts and Sternberg[36] (83.1% and 87.9% respectively for χ_1 and χ_2 for surface exposed residues) in the case of protein-protein association. This difference suggests that side chains in binding pockets are more rigid than those in protein interfaces (perhaps due to functional constraints in ligand recognition). Although the averages for χ_1 and χ_2 are similar, detailed analysis for each residue show larger differences. In most cases χ_1 is more flexible than χ_2 however, we see the opposite behavior in Asn (96.6% and 92.0%) and Ile (95.9% and 91.5%).

Amino acid flexibility

We calculated the probability with which each amino acid undergoes side-chain conformational changes as follows:

$$p_i = \frac{N_C^i}{N_T^i} \pm \frac{\sqrt{N_C^i}}{N_T^i} \quad (1)$$

where N_C^i is the total number of amino acids of type i undergoing conformational changes and N_T^i is the total number of amino acids of type i present in all binding pockets, the second term is the error estimation involved in the measurement.

Our purpose is to estimate the probability of an amino acid already present in a binding pocket to undergo conformational changes, therefore, we did not normalize p_i by the probability of occurrence of different amino acids in binding pockets.

The data summarized in Table II is sufficient for statistical analysis of flexibility for individual amino acids, including the least frequently occurring one (Met) with 196 cases

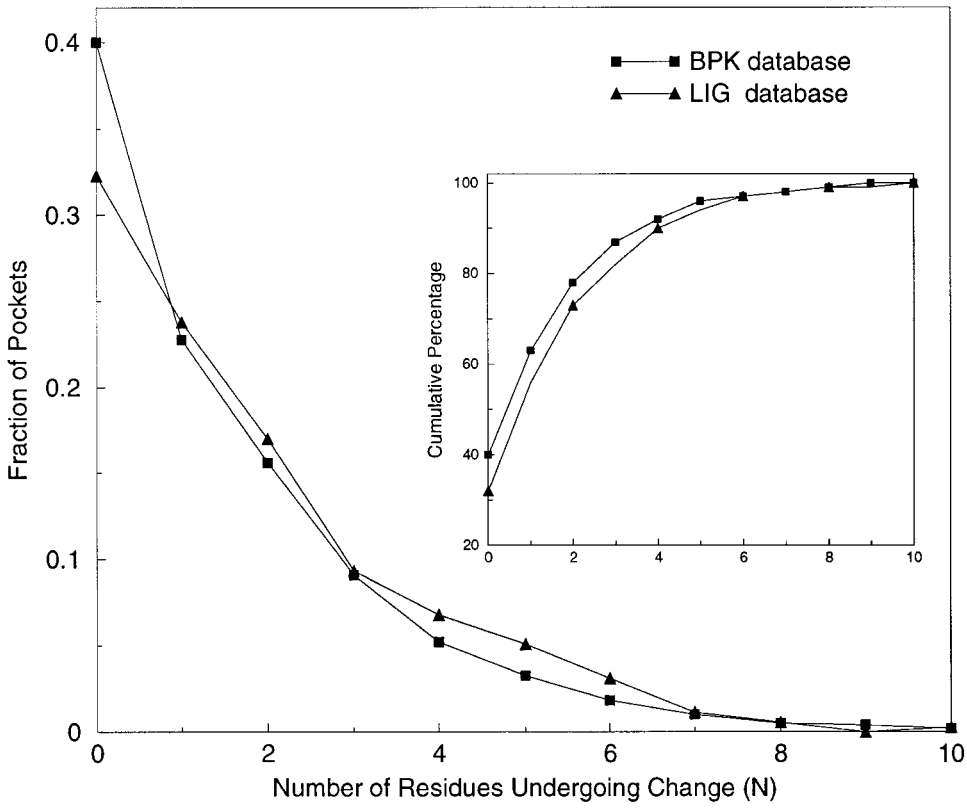


Fig. 2. Flexibility of binding pockets. Changes in side-chain conformations were analyzed in 980 pockets for the BPK database and 353 pockets for the LIG database. The fraction of pockets was plotted versus the number of pocket residues undergoing changes in their side-chain conformation upon ligand binding. The inset shows the cumulative percentage of pockets in which not more than the indicated number of residues (co-ordinate axis) undergo conformational change.

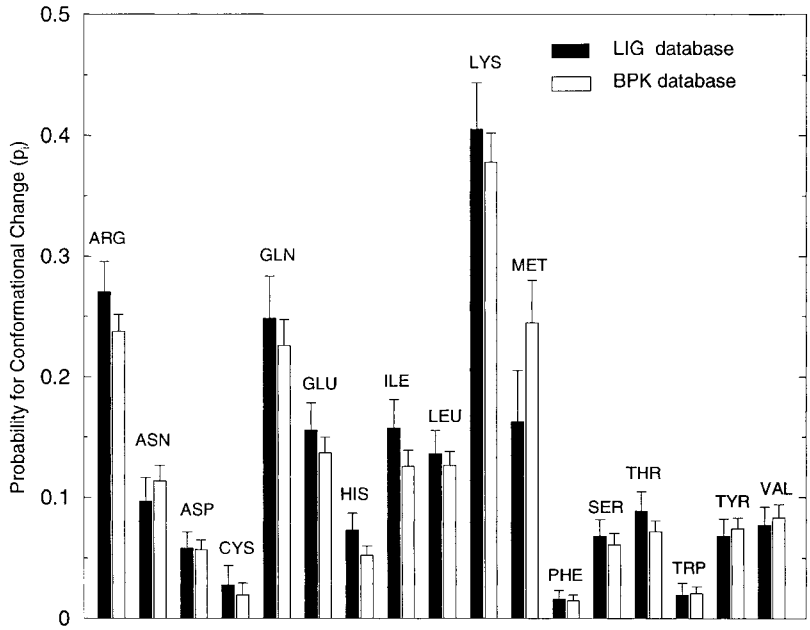


Fig. 3. Flexibility residue side chains. The probability for side-chain conformational change for the different amino acids (p_i) is shown for the BPK and LIG databases. Bars represent error estimation according to Eq. 1.

in the BPK database and 92 cases in the LIG database. Side-chain flexibility upon ligand binding, p_i (Eq. 1), for the amino acids listed in Table II is presented in Figure 3. The results indicate the following order: Lys > Arg, Gln, Met > Glu, Ile, Leu > Asn, Thr, Val, Tyr, Ser, His, Asp > Cys, Trp, Phe; with a 25-fold difference in the probability to undergo side-chain conformational changes between Lys and Phe.

The low flexibility of Cys can only be partially explained by its participation in disulfide bonds since, about 50% of the cysteines are involved in disulfide bonds irrespective of whether their side-chains undergo conformational change.

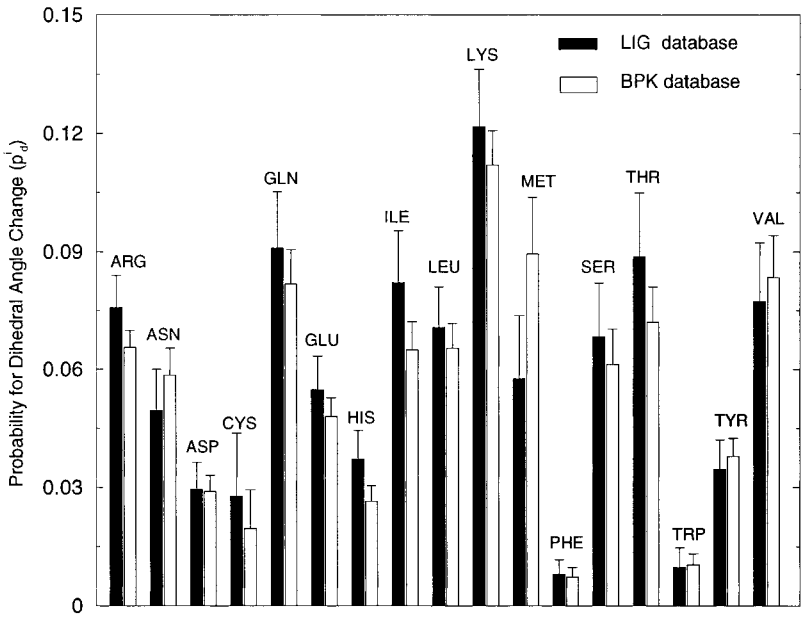


Fig. 4. Flexibility of dihedral angles. The probability for side-chain conformational change for the different amino acids from Figure 3 were used to estimate the probability of a single dihedral angle to undergo change according to Eq. 2.

It is interesting to know whether flexibility measured as rotations of multiple bonds differs from flexibility measured as the rotation of at least one bond. We calculated this and found that flexibility of multiple bonds shows a similar trend to that for at least one bond (data not shown).

Correlation with structural and chemical features

We noticed a tendency in the results of Figure 3 for some amino acids with three or four side-chain rotatable bonds (such as Arg, Gln, Lys, and Met) to be more flexible, while several amino acids with one or two such bonds (such as Asp, Cys, Phe, and Trp) were more rigid. As a first approximation, one can suppose that the different dihedral angles can rotate independently. We denote p_d^i as the probability of a single dihedral angle to undergo conformational changes in amino acid i. In order to estimate p_d^i, we denote the probability that a given bond does *not* undergo conformational change as $1 - p_i$. For an amino acid side chain with n_d flexible dihedral angles, p_d^i is then given by:

$$p_d^i = 1 - \sqrt[n_d]{1 - p_i} \tag{2}$$

The root comes from the fact that there are n_d independent dihedral angles in a given side chain.

Analyzing the calculated bond probabilities (Eq. 2), we still find significant differences in flexibility among the amino acids. Thus, while there is a correlation between the number of flexible dihedral angles and the probability for a side chain to undergo conformational change, in general, differences among amino acids present in Figure 3 are still apparent in Figure 4, although somewhat attenuated. No correlation was observed between side-chain flexibility and number of atoms rotated.

Recently, Kawashima and Kanehisa[37] created a database of amino acid indexes containing 437 different sets of values that reflect structural propensities as well as physicochemical properties of the different amino acids. We calculated the correlation between the flexibility scale (Fig. 3 and Table II) and every entry in their database. The highest correlation obtained (0.74) is to an index of average accessible surface area.[38] Analysis of other high ranking indexes (correlation larger than 0.70) showed clustering in two categories; viz., indexes related to surface accessibility and α-helix-stability.

Backbone versus side-chain flexibility

The parameter we use to estimate the extent of backbone movements of the binding pocket residues is the maximal displacement of C_α atoms ($\Delta d_{\max}$). From the list of residues in contact one calculates all possible pairwise distances, $d_{i,j}$, between C_α atoms in the apo- and holoprotein entries. The value of $\Delta d_{\max}$ is given by:

$$\Delta d_{\max} = \max_{\langle i,j \rangle} \mid d_{i,j}^{APO} - d_{i,j}^{HOLO} \mid \tag{3}$$

where $\langle i,j \rangle$ denotes all pairwise combinations of C_α atoms from residues in contact with the ligand.

Ligand binding can cause large backbone displacements. If side-chain flexibility were of minor importance compared to backbone movements, it would be unrealistic to study the former in a database where the latter effect was major. Analysis of the BPK database shows that only 12% of the cases have backbone displacements ($\Delta d_{\max}$) larger than 2 Å; indeed, 75% of the cases show $\Delta d_{\max}$ of less than 1 Å (see Fig. 5). Thus, backbone displacements in binding pockets are, on the average, of minor importance when compared to side-chain flexibility. Similar conclusions were recently arrived at for protein-protein association.[36]

The inset in Figure 5 is an enlargement of the plot of backbone versus side-chain movements in the region of 1 to 5 Å of $\Delta d_{\max}$. One can readily see that the data is scattered, suggesting that there is little correlation between backbone displacement and side-chain flexibility.

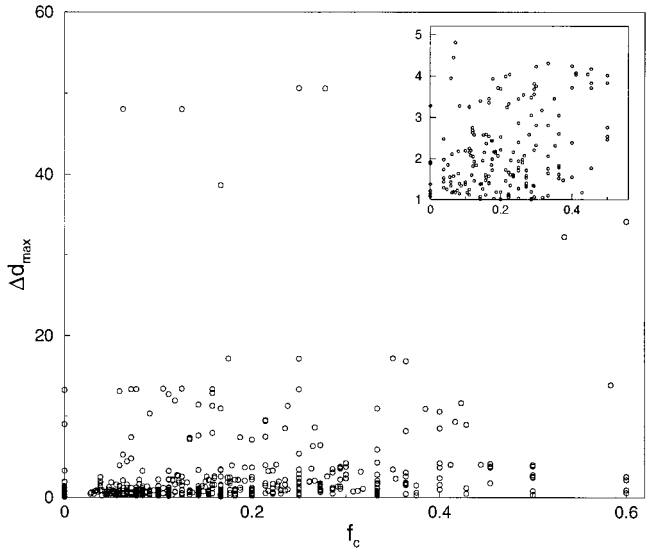


Fig. 5. Correlation between backbone movements and side-chain flexibility. Δd_{max} is the maximal displacement of C_α atoms of the binding pocket (Eq. 4) while f_c is the fraction of residues undergoing side-chain conformational changes. f_c is calculated as the ratio of the number of amino acids undergoing conformational change to the total number of residues in that pocket. Each point represents a single binding pocket. The data shown is for the BPK database. The inset presents a magnified view for the 1–5 Å range of Δd_{max}.

We note that in the few cases of very large backbone displacement (i.e., >18 Å) in our database, the fraction of residues undergoing side-chain conformational changes is not larger than average. Thus, it is likely that the flexibility of side-chains in pockets subject to very large motions does not differ from that of side-chains involved in more usual, smaller backbone displacements.

Side-chain conformational changes in apo protein entries

We analyzed pairs of apo protein entries belonging to the same sub block on the MAX database to determine to what extent side-chain conformational changes are due to ligand binding or to variations among independently determined structures of the same protein. We observed that changes of more than 60° are rarer among apo-apo pairs than among holo-apo ones. However, the flexibility scale for the two cases showed the same trend, suggesting that this is an intrinsic property of the amino acids.

APPLICATIONS

We are interested in learning how to diminish the complexity of searching the side-chain conformational space. This would be useful for improving the efficiency of docking predictions. Our results show that only a small number of amino-acid side chains belonging to binding pocket residues need to be considered flexible at any one time in a search algorithm (e.g., amino-acid side chains of three or fewer residues). In addition, the flexibility scale is useful in decreasing the size of the combinatorial problem of choosing a set of residues likely to contribute to side-chain flexibility (e.g., the triplet [Arg, Lys, Met] would be chosen for calculation while [Cys, Phe, Trp] would be eliminated).

The side-chain flexibility scale should also prove useful in planning protein engineering experiments by suggesting those mutations most likely to affect the plasticity of a binding pocket. Intuitively, it seems likely that by creating a more rigid pocket one might narrow the range of compounds which could otherwise bind to it.

ACKNOWLEDGEMENTS

We thank Drs. E. Domany, A. Finkelstein, S. Pietrokovsky, and E. Trifonov for their valuable comments. V.S. was partially supported by the State of Israel, Ministry of Absorption and the Center for Absorption of Scientists.

REFERENCES

1. Bernstein F, Koetzle T, Williams G, et al. The Protein data bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
2. Petrella R, Lazaridis T, Karplus M. Protein side chain conformer prediction: a test of the energy function. Fold Des 1998;3:353–377.
3. Kuntz I, Blaney J, Oatley S, Langridge R, Ferrin T. A geometric approach to macromolecule-ligand interactions. J Mol Biol 1982; 161:269–288.
4. Boehm H-J. The computer program LUDU: a new method for the de novo design of enzyme inhibitors. J Comput Aided Mol Des 1992;6:61–78.
5. Sobolev V, Wade R, Vriend G, Edelman M. Molecular docking using surface complementarity. Proteins 1996;25:120–129.
6. Oshiro C, Kuntz I, Dixon J. Flexible ligand docking using a genetic algorithm. J Comput Aided Mol Des 1995;9:113–130.
7. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–489.
8. Sobolev V, Moallem T, Wade R, Vriend G, Edelman M. CASP2 Molecular docking predictions with the LIGIN software. Proteins Suppl. 1997;1:210–214.
9. Makino S, Kuntz I. Elect++: Faster conformational search method for docking flexible molecules using molecular similarity. J Comput Chem 1998;19:1834–1852.
10. Leach A. Ligand docking to proteins with discrete side-chain flexibility. J Mol Biol 1995;235:345–356.
11. Luty B, Wasserman Z, Stouten P, Hodge N, Zacharias M, McCammon J. A molecular mechanics/grid method for the evaluation of ligand-receptor interactions. J Comput Chem 1995;16:454–464.
12. Jones G, Willett P, Glen R. Molecular recognition of receptor-sites using a genetic algorithm with a description of desolvation. J Mol Biol 1995;245:43–53.
13. Knegtel R, Kuntz I, Oshiro C. Molecular docking to ensembles of protein structures. J Mol Biol 1997;266:424–440.
14. Apostolakis J, Pluckthun A, Calfish A. Docking small ligands in flexible binding sites. J Comput Chem 1998;19:21–37.
15. Sandak B, Wolfson H, Nussinov R. Flexible docking allowing induced fit in proteins: insights from an open to closed conformational isomers. Proteins 1998;32:159–174.
16. Mangoni M, Roccatano D, Di Nola A. Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulations. Proteins 1999;35:153–162.
17. Trosset J-Y, Scheraga H. Flexible docking simulations: scaled collective variable monte carlo minimization approach using bezier splines, and comparison with a standard Monte Carlo algorithm. J Comput Chem 1999;20:244–252.
18. Zacharias M, Sklenar H. Harmonic modes as variables to approximately account for receptor flexibility in ligand-receptor docking simulations: application to DNA minor groove ligand complex. J Comput Chem 1999;20:287–300.
19. Fraser M, Strynadka N, Bartlett P, Hanson J, James M. Crystallographic analysis of transition-state mimics bound to penicillopepsin: phosphorus-containing peptide analogues. Biochemistry 1992; 31:5201–5214.
20. Hecht H, Szardenings M, Collins J, Schomburg D. Three-dimensional structure of a recombinant variant of human pancre-

atic secretory trypsin inhibitor (Kazal type). J Mol Biol 1992;225: 1095–1103.
21. Lesk A, Chothia C. Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint. Nature 1988;335:188–190.
22. Ikura M, Clore G, Gronenborn A, Zhu G, Klee C, Bax A. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. Science 1992;256:632–638.
23. Katzin B, Collins E, Robertus J. Structure of ricin A-chain at 2.5 Å. Proteins 1991;10:251–259.
24. Xu Z, Bernlohr D, Banaszak L. Crystal structure of recombinant murine adipocyte lipid-binding protein. Biochemistry 1992;31: 3484–3492.
25. Bower M, Cohen F, Dunbrack RJ. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modelling tool. J Mol Biol 1997;267:1268–1282.
26. Huang E, Koehl P, Levitt M, Pappu R, Ponder J. Accuracy of side-chain prediction upon near-native protein backbones generated by ab initio folding methods. Proteins 1998;33:204–217.
27. Samudrala R, Moult J. Determinants of side chain conformational preferences in protein structures. Protein Eng 1998;11:991–997.
28. Desmet J, De Maeyer M, Hazes B, Lasters I. The dead-end elimination theorem and its use in protein side-chain positioning. Nature 1992;356:539–542.
29. Goldstein R. Efficient rotamer elimination applied to protein side-chains and related spin glasses. Biophys J 1994;66:1335–1340.
30. De Maeyer M, Desmet J, Lasters I. All in one: a highly detailed rotamer library improves both accuracy and speed in the modelling of side chains by dead-end elimination. Fold Des 1997;2:53–66.
31. Lasters I, De Maeyer M, Desmet J. Enhanced dead-end elimination in the search for the global minimum energy conformation of a collection of protein side chains. Protein Eng 1995;8:815–822.
32. Lasters I, Desmet J. The fuzzy-end elimination theorem: correctly implementing the side chain placement algorithm based on the dead-end elimination theorem. Protein Eng 1993;6:717–722.
33. Keller D, Shibata M, Markus E, Ornstein R, Rein R. Finding the global minimum: a fuzzy end elimination implementation. Protein Eng 1995;8:893–904.
34. Sobolev V, Sorokine S, Prilusky J, Abola E, Edelman M. Automated analysis of interatomic contacts in proteins. Bioinformatics 1999;15:327–332.
35. Sobolev V, Edelman M. Modelling the quinone-B binding site of the photosystem-II reaction center using notions of complementarity and contact-surface between atoms. Proteins 1995;21:214–225.
36. Betts M, Sternberg M. An analysis of conformational changes on protein-protein association: implications for predictive docking. Protein Eng 1999;12:271–283.
37. Kawashima S, Kanehisa M. AAindex: amino acid index database. Nucleic Acids Res 2000;28:374.
38. Janin J, Wodak S, Levitt M, Maigret B. Conformation of amino acid side chains in proteins. J Mol Biol 1978;125:357–386.