Journal of Computer-Aided Molecular Design, **12:** 563–572, 1998.
KLUWER/ESCOM

Do active site conformations of small ligands correspond to low free-energy solution structures?

Michal Vieth*, Jonathan D. Hirst & Charles L. Brooks III**
*Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, TPC-6, La Jolla, CA 92037, U.S.A.; *Current address: Department of Computational Chemistry and Macromolecular Structure Research, Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, IN 46032, U.S.A.*

Received 6 January 1998; Accepted 13 April 1998

Key words: anchor points, conformational search, ligand binding

Summary

We compare the low free energy structures of ten small, polar ligands in solution to their conformations in their respective receptor active sites. The solution conformations are generated by a systematic search and the free energies of representative structures are computed with a continuum solvation model. Based on the values of torsion angles, we find little similarity between low energy solution structures of small ligands and their active site conformations. However, in nine out of ten cases, the positions of 'anchor points' (key atoms responsible for tight binding) in the lowest energy solution structures are very similar to the positions of these atoms in the active site conformations. A metric that more closely captures the essentials of binding supports the basic premise underlying pharmacophore mapping, namely that active site conformations of small flexible ligands correspond to their low energy structures in solution. This work supports the efforts of building pharmacophore models based on the information present in solution structures of small isolated ligands.

Introduction

One important activity in rational drug design is the generation of pharmacophore models. A pharmacophore model is a spatial arrangement of atoms or functional groups believed to be responsible for biological activity [1]. Pharmacophore elucidation provides one with a means of searching databases of three dimensional structures for putative bioactive molecules [2]. Pharmacophore searching techniques have been quite effective in discovery of novel ligands [3, 4]. When the structure of a receptor is not known, pharmacophore models are often constructed from information about solution structures, which may come from NMR, computer calculations or crystal structures of isolated ligands [5]. In these cases, pharmacophore mapping techniques are used to elucidate a pharmacophore from a series of active and inactive compounds [6–11]. Underlying these approaches are the assumptions that there are commonalities between different ligands binding to the same receptor [9], and that these common features are represented in the low energy solution structures of the ligands.

The degree of similarity between active site conformations and low energy solution structures of the same ligands will provide an indication of the physics of binding. For example, to what extent does the receptor induce structural changes in the ligand when bound to the active site? Does it substantially deform the solution structure of the ligand with binding involving 'key reshaping' of the ligand as proposed for peptides [12], or is the change in ligand structure negligible, with the lock and key mechanism [13] holding? How 'rusty' is the lock and key model [14]? It is interesting to note that antibodies have been shown to evolve from an induced fit scenario favoring broad specificity of binding to a lock and key type of binding, resulting in higher affinity [15].

In recent years, the relationship between the conformation of small molecules in isolated crystals and

**To whom correspondence should be addressed.

in receptor complexes has been examined a number of times [14, 16, 17]. There are examples that support the view of similarity for nonpolar ligands [16, 17]. For example, the crystal structure of isolated retinol is very similar to its conformation in the active site of retinol binding protein [18]. However, for polar molecules, a number of examples show differences between conformations in crystals of isolated molecules and in their protein-complexed form [17]. The conclusion reached by a number of researchers is that, in most cases, there is no relationship between the structure of an isolated ligand in its molecular crystal and the structure of the ligand in a protein active site [16, 17]. It was concluded by Nicklaus that flexible ligands undergo substantial deformations upon binding to receptors [17].

In this paper our goal it to examine a principal assumption that underlies the basis of pharmacophore construction. We try to find a similarity between low energy solution structures of small ligands and their conformation in receptor complexes. We stress that this paper does not provide a general means of pharmacophore identification from solution structures. In addition, we will not predict the ability of a given ligand to bind to the receptor. We will use information from crystal structures of complexes to establish properties of ligands in the active sites of complexes and compare them to low energy solution structures. A similarity between the bound and unbound forms of the ligand may correspond to a necessary condition for a given ligand to bind to its receptor, and this may be used as means of focusing design efforts on a subset of ligands and eliminating a large number of others. However, this condition is not sufficient for design, as binding properties are an interplay between thermodynamics of the free ligand and protein in solution and the ligand/receptor complex and as such cannot, in principle, be inferred from properties of the free ligand alone.

Before we examine the definition of similarity in detail, we should note that ligands often show some flexibility in the active site. This flexibility may be observed and modeled through computer simulation of ligands in their receptor sites. Such flexibility may not be inconsistent with a single crystal conformation as there is some imprecision in ligand atomic position assignment [19]. Ligand flexibility in the active site prompts us to define similarity in broad terms. In this paper, we consider two metrics of similarity: one based on torsion angles and the other based on the relative position of a subset of atoms. First we can compare low energy solution structures to a family of conformations observed in active sites and explore the hypothesis: 'There is a similarity between the family of active site conformations and the low energy structures adopted in solution.' Another observation, evident from simulations of ligands in their receptor active sites is that some atoms, 'anchor points', are much less mobile than others. These atoms may be assumed to be those most responsible for tight binding, since we find that mobility is well correlated with ligand/receptor interaction energy. This observation prompts us to formulate a second hypothesis for exploration: 'There is a similarity between the spatial orientation of anchor points observed in active site ligand conformations and the position of those anchor points in low energy solution structures.' In this paper, we test these two hypotheses.

Methodology

In order to investigate the similarity between low energy solution structures of small ligands and their active site conformations, we have selected ten protein-ligand complexes from the Brookhaven Protein Data Bank (PDB) [20]. The ligands are diverse enough to draw general conclusions and sufficiently small to permit the use of an exhaustive search strategy to explore the solution energy hypersurface. The ligands are depicted in Figure 1. They include dipeptides, peptide-like molecules, small and medium-sized organic molecules. The number of rotatable bonds in these ligands varies from two to five. In this counting and in the subsequent calculations, we only consider rotatable bonds that involve rotation of at least one more independent rotatable bond. For example, for the isoleucine side chain, the torsion angle χ_1 is considered, whereas χ_2 is not. This ensures that we focus our computational effort on the most important part of conformational space (i.e. that part giving rise to the largest changes in size and shape of the substrate).

Our identification of low energy structures of uncomplexed ligands in solution is based on a systematic search [8] over the molecular torsions. The generated conformations are then minimized and clustered. Subsequently, cluster centers are subjected to free energy evaluations using a continuum model for solvent and compared to the population of ligands in the active sites generated by molecular dynamics (MD) simulations. In the following section we introduce the force field, the techniques used to select anchor points,

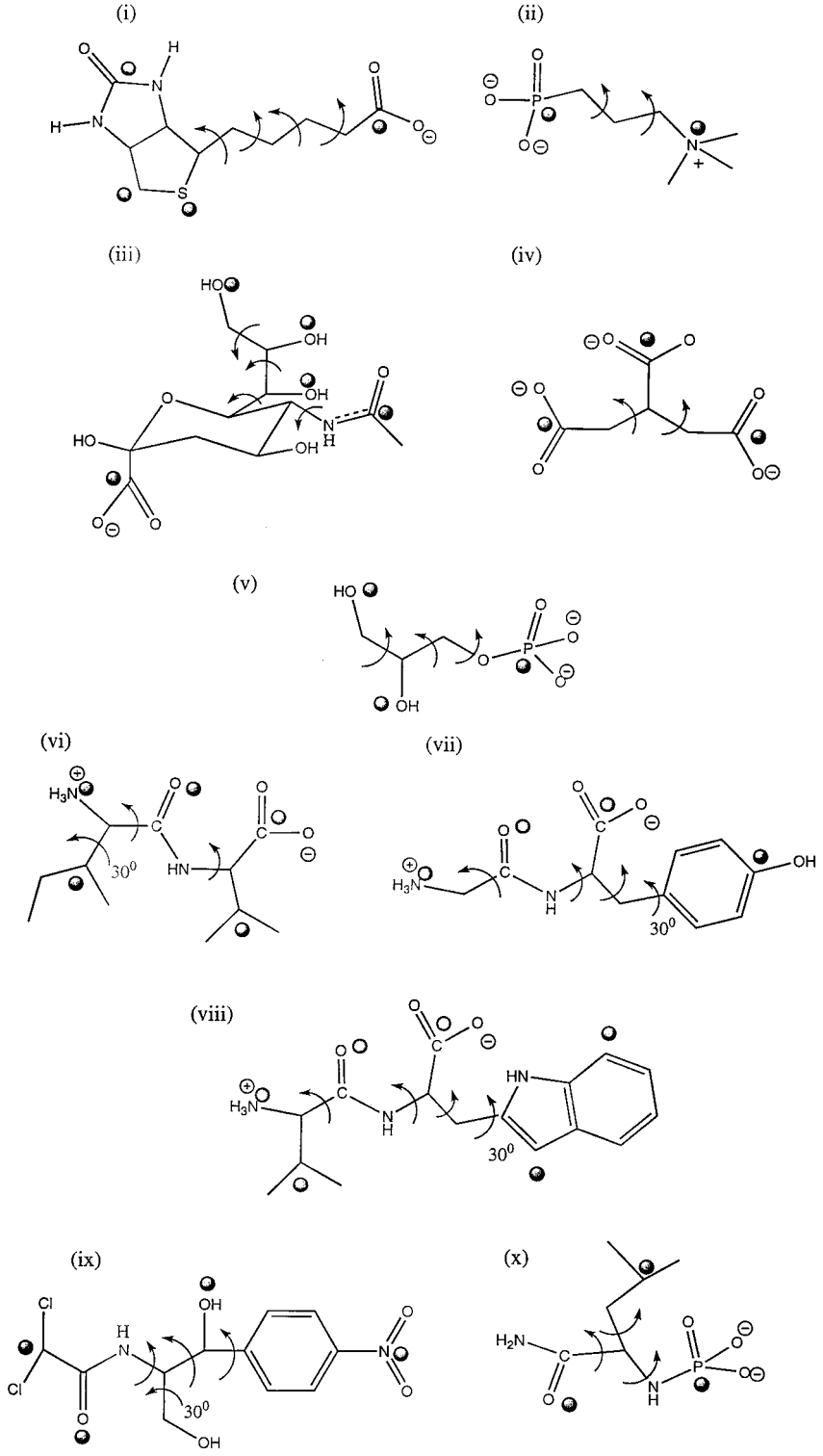

Figure 1. Ligands used in this study: (i) biotin (streptavidin) [32], (ii) phosphocholine (Fab McPC-603) [33], (iii) sialic acid (hemagglutinin) [34], (iv) tricarballylic acid (aconitase) [35], (v) glycerol 3-phosphate (Triose Phosphate Isomerase) [36], (vi) Ile-Val (Pancreatic Trypsin Inhibitor) [37], (vii) Gly-Tyr (Carboxypeptidase A) [38], (viii) Val-Trp (Thermolysin) [39], (ix) Chloramphenicol (Chloramphenicol Acetyltransferase) [40], (x) N-phosphoryl-L-leucinamid (Thermolysin) [41]. Anchor points are marked with gray circles, the rotatable bonds considered are indicated by arrows, rotation increments (if different from 15°) are shown.

our methods used to organize the data, and a fitness function (related to the ligand free energy in solution) for the identification of highly populated solution structures.

Force field

For all peptidic ligands, the polar hydrogen CHARMM toph19/param19 topology and parameter sets [21] were used. The force field parameters for nonpeptidic ligands were obtained from Quanta [22]. The charges were generated by a template method. In the cases where no templates were present in the Quanta database, the charges were fitted to an HF/6-31G electrostatic potential [23]. As discussed later, there is some evidence that our conclusions are not particularly sensitive to the protocol used for generating charges. Files containing all non-standard charges and parameters are available through anonymous ftp and are also mirrored on our web page [24].

Structure generation

To examine the structural diversity of a ligand in its active site, we analyzed 100 trials (100 ligand replicas using the multiple copy simultaneous search method [25]) of constant temperature MD simulation. The ligands were free to move, but the receptors were held rigid. All crystallographic water molecules were removed, except those that bridge interactions between ligand and protein (as occurs in the following complexes: tricarballylic acid with aconitase, chloramphenicol with chloramphenicol acetyltransferase, Ile-Val dipeptide with pancreatic trypsin inhibitor). Metal ions that mediate interactions between the ligand and protein were retained. A distance dependent dielectric (of the form of $\epsilon = 2r$) was used in the simulations. A nonbonded interaction truncation of 8 Å was applied. A switching function for the electrostatic and van der Waals interactions was used with switching distance of 6 Å. The nonbonded list consisted of all atom pairs. The simulations were carried out at a constant temperature of 300 K for 75 ps and the final structures were quenched to 50 K in 1 ps.

For the solution structures, conformational flexibility was explored by a systematic search [8]. Each rotatable bond was sampled in 15° increments and the bonds responsible for ring rotations were sampled in 30° increments. The magnitudes of increments for each bond rotation are shown in Figure 1. The structures resulting from the systematic search were minimized using 1000 steps of conjugate gradient minimization with an energy tolerance of 0.0001 kcal/mol. This procedure ensures that the generated structures fall into the nearest local minimum. The number of individual structures produced by the systematic search ranged from 1036 for phosphocholine to 316 954 for biotin. Clustering was then used to identify the distinct structures.

Selection of anchor points

There have been a number of methods proposed to elucidate pharmacophores from crystal structures of complexes. In Clark's scheme [1], anchor points for pharmacophore building include heteroatoms, hydrogen-bond donors and acceptors, aromatic and lipophilic groups. The purpose of our work is to examine the basic premise underlying pharmacophore generation from solution structures of isolated ligands. For this reason, we exploit all available information about the structure of the ligand-receptor complex to develop an appropriate similarity metric. In this work, the anchor points were selected based on MD simulations of the ligands in their receptor active sites. The average interaction energy with the receptor was computed on a per atom basis. Then, based on the mobility of atoms during the course of active site simulations, atoms of lowest mobility were identified in each of the functional groups of the ligand. Anchor points were selected based on the largest contribution to the interaction energy with the receptor and on the lowest mobility. The above selection corresponds to choosing those ligand atoms most responsible for tight binding with the receptor. A detailed description of the anchor point selection is shown in Figure 2. The use of more than one conformation complementary to the active site is chosen to eliminate small differences in less important regions of structures and to emphasize the similarities based on anchor points. Furthermore this strategy substitutes some bound conformations of high-energy (based on our force field) with low-energy structures, thus providing consistency between our models of conformations in solution and in the active site.

It is worth noting that some functional groups do not provide any anchor points (e.g., one of the hydroxyl groups in chloramphenicol). Typically, in such a case, the functional group may be highly mobile with a low interaction energy with the receptor, as seen in groups that are exposed to solvent. This highlights some of the difficulties associated with generating a pharmacophore model in the absence of detailed struc-

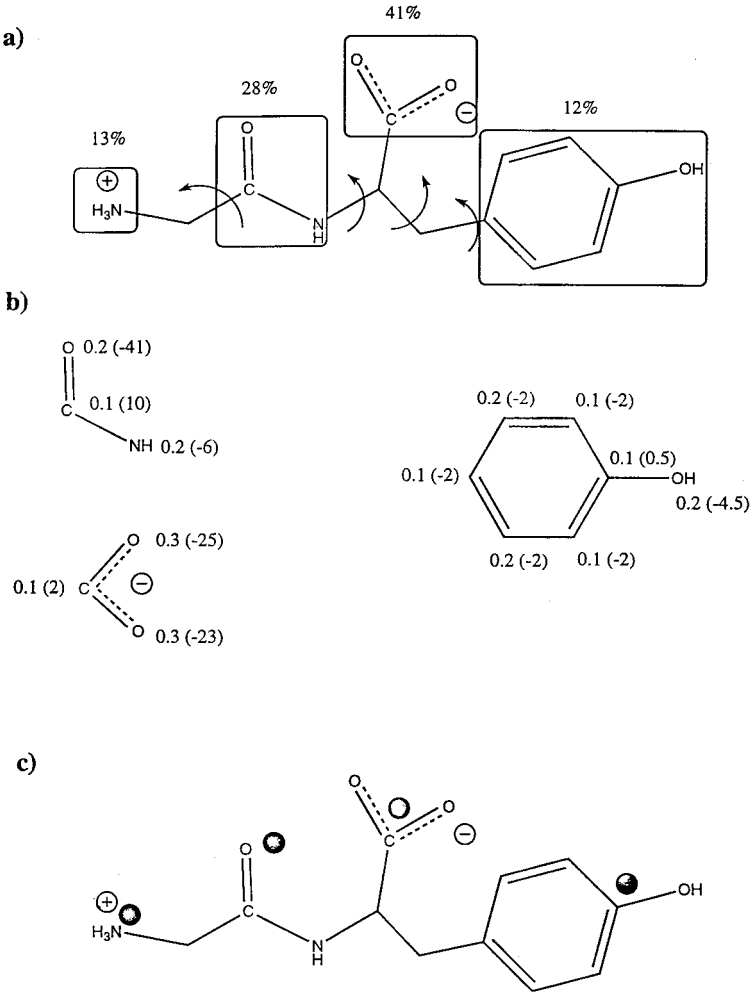

Figure 2. Scheme for the selection of anchor points, illustrated for Gly-Tyr. The molecule is divided into functional groups (donors, acceptors, aromatic, aliphatic). The aliphatic linkers are not considered to be functional groups. Within each functional group, atoms with the highest contribution to the interaction energy and the lowest mobility (as computed from MD simulations) with the receptor are selected. (a) The division of the molecule into functional groups, and the fraction of the total interaction energy contributed by each functional group. If this fraction is less than 50% of the average fraction of interaction energy per group (12.5% for four functional groups) and there are no atoms with mobility below the mean (0.3 in this case) no anchor points will be assigned for this group. (b) For each heavy atom in a functional group, we compute its mobility (first number in Å) and interaction energy (second number in parentheses in kcal/mol) as indicated. (c) In general, anchor points are heavy atoms with a mobility lower than the mean over all heavy atoms (0.3 in this case) and which contribute substantially to the interaction energy of the functional group. Substantial is defined as greater than average for the heavy atoms in the functional group. For the case of symmetric atoms (i.e., the carboxylic group) the atom closest to the average position of other possible anchor atoms (i.e., two oxygens) is selected. For peptide bonds, carbonyl oxygens are selected. All amino acid side chains are considered to be potential functional groups. For all single rings, one anchor point is selected. For fused rings, two (e.g., side chain of tryptophan) or three points (e.g., biotin) are selected.

tural data. Fortunately, we need not dwell on this issue, but we note in passing there are software packages that address this issue [26, 27] and that general ideas from our method for identifying anchor points could be used in the case of unknown receptor structure (see Figure 2). However, it would still be difficult to identify functional groups that do not interact strongly with the receptor, unless this group is common in a series of active compounds.

Organization of the structures

In order to group the structures into families and to select a small number of representative structures, clustering techniques [28] were used. The ligand conformations in the active site were clustered in an

hierarchical agglomerative fashion. The clustering was carried out until the cluster radius was less than 15° for the torsion angle metric. For the anchor point distance metric, the stopping criterion was a cluster radius of 0.5 Å. The number of clusters for the active site conformations was between three and ten for the torsion metric and was one or two for the distance metric. There were too many solution structures to carry out hierarchical agglomerative clustering, thus the data were partitioned in 1–5° increments (0.1–0.2 Å increments for the distance metric). In this partitioning, a partition center was defined as the structure closest to the average structure for that partition. The partition centers were then used in hierarchical agglomeration. For the test case of phosphocholine, conformational partitioning followed by clustering on the partition centers gives an overall partitioning of structures into clusters which are roughly 80% identical to full hierarchical agglomerative clustering. The number of clusters of solution structures varied from 14 (four for the distance metric) to 225 (105 for the distance metric). After agglomeration, clustering representatives from each cluster were selected based on the energy (computed using a distance dependent dielectric of 2r) to be used to estimate the free energy of each cluster. In other words, the partition centers (referred later to as cluster representatives) were defined as the lowest energy structures from each cluster. The number of structures in each stage of the procedure is given in Table 1.

Free energy of solution structures

In order to estimate the Boltzmann probability to observe a given cluster in solution, the free energy was calculated based on the population of the cluster (entropic contribution), the solvation energy of the cluster representative and the internal energy of a cluster representative:

$$A = E_{\text{solv,PB}} + E_{\text{int}} - k_B T \, ln \, N, \qquad (1)$$

where A is the free energy of a cluster, $E_{\text{solv,PB}}$ is the solvation free energy of a cluster representative calculated using the finite difference Poisson Boltzmann (PB) equation [29] (with a grid size 0.4 Å and an interior dielectric of 2), E_{int} is the internal free energy of a ligand, $k_B T = 0.6$ kcal/mol and N is the population of a cluster. In this approach, all structures within one cluster are treated as isoenergetic. Clustering reduces the amount of computer time required for the free energy calculations by up to 1000-fold. The clusters were rank ordered based on their free energies and their Boltzmann probabilities were computed. For each ligand, cluster representatives from solution are compared to cluster representatives from the active site simulations based on the torsion and distance metrics. The low free energy solution structures that are similar to the active site conformations are identified as the lowest free energy cluster representatives that are close (i.e., within the clustering radius) to any of the active site cluster representatives. That is, we compare one representative from each cluster in solution to one representative from each active site cluster. For the anchor point clusters, there is only one cluster in the active site (and it contains the crystal structure conformation). For the torsion angle metric, there are more active site clusters (one of which contains the crystal structure conformation). We note in passing that some crystal structures minimized in the CHARMM force-field are not identical to the unminimized crystal structures, based on the relatively tight criterion of 15° RMSD that we use in this study.

The precision in our free energy estimates is about ±3 kcal/mol because the free energy assignment for each cluster depends slightly on the choice of the cluster representative. Thus, all structures/clusters with free energies within 3 kcal/mol of the minimum free energy cluster should be treated as indistinguishable. The 3 kcal/mol imprecision arises in part from differences between energy calculations using an effective distance dependent dielectric solvent model and calculations including PB solvation energy [30].

Results

Torsion angle metric

In this section we address our first hypothesis ‘Is there similarity in important torsions between the active site conformations and low energy solution structures?’ Table 2 shows the lowest free-energy solution structures that are similar to the active site conformations. Representative active site conformations were chosen based on the clustering of 100 structures generated by MD simulations as described above. For four out of ten cases, the lowest free energy solution structures (within 3 kcal/mol from the lowest free energy structure) have the same torsion angles as do the corresponding ligands in their respective active sites. Two other ligands have solution structures with the same torsion angles as ligand conformations that are within

Table 1. Results of clustering

Ligand (PDB code of complex)	Number of active site clusters		Number of solution structures		
	Torsion metric	Anchor point distance metric	Total number from systematic search	After clustering with torsion metric	After clustering with anchor point distance metric
Chloramphenicol (3cla)	15	2	77 532	118	20
N-Phosphoryl-L-leucinamid (2tmn)	7	1	82 012	132	25
Phosphocholine (2mcp)	9	1	1 036	14	4
Tricarballylic acid (6acn)	1	1	1 296	70	5
Biotin (1stp)	24	1	316 954	225	105
Glycerol 3-phosphate (6tim)	9	2	13 785	60	19
Ile-Val (2tpi)	10	1	16 137	36	11
Gly-Tyr (3cpa)	9	1	82 547	92	31
Val-Trp (3tmn)	5	1	6 912	110	47
Sialic acid (4hmg)	11	1	72 588	156	61

5 kcal/mol of the lowest free energy solution structure. For sialic acid and the Val-Trp dipeptide, solution structures that correspond to the ligand conformation in the active site are of high free energy (9–10 kcal/mol higher than the lowest free energy solution structure). Thus, in agreement with previously published data [16, 17], based on a torsion angle metric, we find little similarity between low-energy solution structures of small, polar ligands and their active site conformation. For some ligands there is some similarity, but this is only evident if one considers a family of conformations in the active site which has a broader distribution of torsion angles than the single crystal conformation. In general, in order to recover the active site ligand structure from conformational searches of the free ligand, a very generous 10 kcal/mol energy cutoff has to be applied. Thus, solution search strategies aiming to provide exact values of torsion angles and to relate those to active site structures will probably not be useful for more than five rotatable bonds due to the exponential growth of conformational space [17].

Anchor point distance metric

In this section the second hypothesis is examined. Table 3 compares low free energy solution structures having the same relative disposition of anchor points as the active site ligand conformations. It is apparent from Table 3, that for nine out of ten complexes the position of the anchor points in the lowest energy solution structure corresponds to the position in the active site ligand conformation. This observation for small, polar ligands should also hold for larger ligands, but this remains to be verified. The similarity is not apparent only in the case of sialic acid, probably because the number of anchor points is largest for this case. Based on the results from Table 3 we support the hypothesis that ‘there is a similarity in the position of anchor points (atoms responsible for tight binding with the receptor) between low energy solution structures and

Table 2. Results for the torsion angle metric

Ligand	Rank[a]	TARMSD[b]	ΔA[c]	Population[d]
Chloramphenicol	1	12.7	0.0	0.45
N-phosphoryl-L-leucinamid	5	5.5	3.0	0.01
Phosphocholine	2	3.1	0.3	0.24
Tricarballylic acid	5	10.4	5.9	0.00
Biotin	27	6.3	4.5	0.00
Glycerol 3-phosphate	2	11.5	2.8	0.01
Ile-Val	6	15.6	4.4	0.00
Gly-Tyr	20	6.2	7.7	0.00
Val-Trp	37	11.6	9.0	0.00
Sialic acid	8	7.4	5.7	0.00

[a] Rank of the lowest solution free energy cluster, whose center is within 15° Torsion Angle Root Mean Square Deviation (TARMSD) from any of the active site cluster centers.
[b] TARMSD of the lowest solution free energy cluster center from the closest active site cluster center.
[c] Free energy difference between the lowest free energy solution cluster and the lowest free energy solution cluster that is similar to any of the active site clusters in kcal/mol.
[d] Population (based on Boltzmann probability at 300 K) of the lowest free energy solution cluster similar to any active site cluster.

Table 3. Results for the anchor point distance metric

Ligand name	Rank[a]	APDRMSD	ΔA	Population
Chloramphenicol	1	0.35	0.0	0.72
N-phosphoryl-L-leucinamid	2	0.20	1.1	0.13
Phosphocholine	1	0.26	0.0	0.99
Tricarballylic acid	1	0.27	0.0	0.99
Biotin	8	0.5	2.7	0.01
Glycerol 3-phosphate	1	0.33	0.0	0.96
Ile-Val	1	0.27	0.0	0.99
Gly-Tyr	1	0.29	0.0	0.75
Val-Trp	2	0.20	0.6	0.20
Sialic acid	10	0.30	6.5	0.00

[a] Rank of the lowest solution free energy cluster, whose center is within 0.5 Å Anchor Point Distance RMS Deviation (APDRMSD) from any of the active site cluster centers. Other columns are analogous to those in Table 2.

Table 4. The lowest free energy solution structures similar to crystallographic structures of isolated ligands

Ligand (CSD code) [a]	Metric	Rank	RMSD	ΔA	Population
Chloramphenicol	Torsion angle	1	12.7	0.0	0.72
(CLM01)	Anchor point distance	1	0.35	0.0	0.45
Biotin	Torsion angle	32	3.3	5.2	0.00
(BIOT01)	Anchor point distance	14	0.27	4.5	0.00

[a] See footnote to Table 2 for explanation of column headings.

the active site conformations.' Thus, receptors may change the detailed solution structure of ligands that bind to their active sites [17], however we postulate that the majority of changes take place in the regions that contribute little to binding. The relative position of atoms responsible for tight binding seems to be very similar in low energy solution structures and in the active site of the receptor. Figure 3 shows a comparison of representatives of the lowest free energy solution cluster to the crystal structures of the ligand in the ligand-receptor complexes for two prototypic ligands. It is interesting that for Gly-Tyr peptide both the anchor point metric and torsion angle metric provide cluster representatives visually similar to the active site conformations. However, based on the torsion angle metric the lowest free energy solution structure is different from the active site conformation.

Comparison of solution and crystal conformations of isolated ligands

In addition to the similarity between solution structure and an active site ligand conformation, it is interesting to compare the former to the crystal structure of an isolated ligand. The comparison is made for two ligands, biotin and chloramphenicol, whose crystal structures are available from the Cambridge Structural Database (CSD) [31]. The results are presented in Table 4. For chloramphenicol the similarity is apparent and independent of the metric, whereas for biotin the similar structure is of relatively higher free energy (about 5 kcal/mol from the lowest). Biotin solution structures with anchor point positions similar to active site con-

Table 5. Results for two different charge sets for Gly-Tyr dipeptide

Metric	Charges	Rank[a]	RMSD	ΔA	Population	Number of clusters
Torsion	Param19	20	6.2	7.7	0.00	92
	HF/6-31G	14	6.4	6.4	0.00	122
Anchor point distance	Param19	1	0.29	0.0	0.75	31
	HF/6-31G	1	0.39	0.0	0.99	42

[a] See footnotes to Table 2.

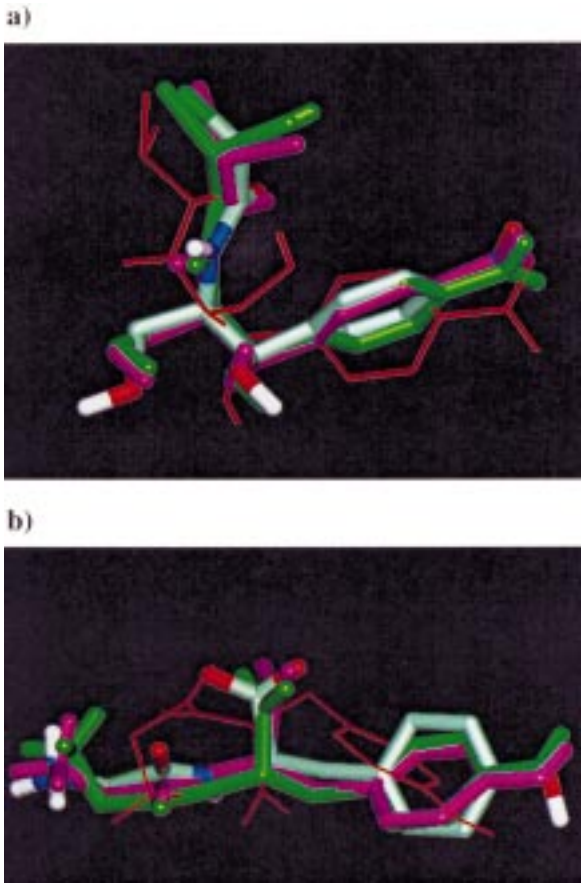

Figure 3. Superposition of the lowest free energy solution structure, that is closest to the family of the active site conformation, onto the crystal structure of the ligand in the ligand-receptor complex. The crystal structure is colored based on atom type, the best structure as judged by the torsion metric is shown in magenta, the best structure as judged by the anchor point distance metric is shown in green. To contrast low free energy solution structures with others, the highest free energy solution structure is shown in red. (a) Chloramphenicol. (b) Gly-Tyr dipeptide.

formations are of slightly lower free energy than those similar to crystal structures of the isolated ligand. This counterintuitive observation might be due to the choice of anchor points based on the behavior of biotin in the environment the receptor provides, rather than in the crystal environment of other biotin molecules. It is possible that selection of anchor points from the crystal environment of other biotin molecules could be slightly different.

Dependence on the charge parameters

We have investigated the dependence of our results on the charge parameters using the Gly-Tyr dipeptide as a test case. For this ligand we have used standard toph19/param19 [21] charges and compared the results to a charge set generated by charge fitting to the HF/6-31G *ab initio* electrostatic potential [23]. The *ab initio* charges were calculated for the active site dipeptide geometry with no geometry optimization because the purpose was merely to generate another reasonable charge set. The average root mean square (RMS) difference of charge per atom is about 0.3 electrons. The results, shown in Table 5, indicate that, although a different charge set changes the number of clusters and the minimum energy structure, the general conclusions remain invariant. For both charge sets the active site conformation is similar (based on torsion angles) to a relatively high energy solution structure. On the other hand, the positions of the anchor points in the lowest energy solution structures generated with the two different charge sets are both identical to the positions of the anchor points in the protein-ligand complex.

Conclusions

We have examined the similarity between low free energy solution structures of small ligands and their

conformations in ligand-receptor complexes. We have found little evidence of detailed similarity, as measured in terms of values of torsion angles. However, similarity in terms of the position of so-called anchor points in low-energy solution conformations of small ligands and the corresponding protein-ligand complexes has emerged. Our results are relatively insensitive to the charge parameters used for the tested example and they appear to be applicable to a wide range of ligands. We observe that receptors often deform a ligand from its low energy solution structure. This deformation, however, takes place in the regions less important for tight binding. The evidence for the similarity of low energy solution structures and active site ligand conformations strongly supports current efforts to build pharmacophore models based on solution structures and identification of commonalities in position of key atoms in a series of ligands known to bind to a given receptor [6–8]. This work cannot, however, judge whether ligands with the appropriate anchor points will bind to the receptor tightly, because the binding is a balance of many factors including shape complementarity, desolvation, electrostatic and dispersion interactions and as such cannot be predicted from solution shape information alone.

Acknowledgements

Helpful discussions with Drs A. Kolinski, A. Godzik, T. Cleveland and A. Ortiz are acknowledged. This work was supported by NIH GM37554 and RR12255.

References

1. Clark, D.E., Westhead, D.R., Sykes, R.A. and Murray, C.W., J. Comput.-Aided Mol. Design, 10 (1996) 397.
2. Martin, Y.C., J. Med. Chem., 35 (1992) 2145.
3. Wang, S., Zaharewitz, D.W., Sharma, R., Marquez, V.E., Lewin, N.E., Du, L., Blumberg, P.M. and Milne, G.W., J. Med. Chem., 37 (1994) 4479.
4. Kiyama, R., Honma, T., Hayashi, K., Ogawa, M., Hara, M., Fujimoto, M. and Fujishita, T.J., J. Med. Chem., 38 (1995) 2728.
5. Klebe, G., J. Mol. Biol., 237 (1994) 212.
6. Hodgkin, E.E., Miller, A. and Whittaker, M., J. Comput.-Aided Mol. Design, 7 (1993) 515.
7. Dammkoehler, R.A., Karasek, S.F., Shands, E.F. and Marshall, G.R., J. Comput.-Aided. Mol. Design, 3 (1989) 3.
8. Dammkoehler, R.A., Karasek, S.F., Shands, E.F. and Marshall, G.R., J. Comput.-Aided Mol. Design, 9 (1995) 491.
9. Jones, G., Willet, P. and Glen, R.C., J. Comput.-Aided Mol. Design, 9 (1995) 532.
10. Mayer, D., Naylor, C.B., Motoc, I. and Marshall, G.R., J. Comput.-Aided Mol. Design, 1 (1987) 3.
11. Sheridan, R.P., Nilakantan, R., Dixon, J.S. and Venkataraghavan, R., J. Med. Chem., 29 (1986) 899.
12. Schwyzer, R., Biopolymers, 37 (1995) 5.
13. Fisher, E. and Thierfelder, H., Berl. Dtsh, Chem. Ges., 27 (1894) 2985.
14. Jorgensen, W.L., Science, 254 (1991) 954.
15. Patten, P.A., Gray, N.S., Yang, P.L., Marks, C.B., Wedemayer, G.J., Boniface, J.J., Stevens, R.C. and Schultz, P.G., Science, 271 (1996) 1086.
16. Böhm, H.-J. and Klebe, G., Angew. Chem. Int. Ed. Engl., 35 (1996) 2588.
17. Nicklaus, M.C., Shaomeng, W., Driscoll, J.S. and Milne, W.A., Bioorg. Med. Chem., 3 (1995) 411.
18. Klebe, G. In Burgi, H.B. and Dunitz, J.D. (Eds.) Structure Correlation, VCH, Weinheim, 1994, p. 543.
19. Allen, F.H. In Domenicano, A. and Hargittai, I. (Eds.) Accurate Molecular Structures, Oxford University Press, Oxford, 1992, p. 355.
20. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
21. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M., J. Comput. Chem., 4 (1983) 187.
22. Quanta, Molecular Simulations Inc., San Diego, CA, 1997.
23. Gaussian 94, Gaussian, Inc., Pittsburgh, PA, 1995.
24. Vieth, M., Parameter files for ligands are available through anonymous ftp from ftp.scripps.edu/ficio/parameters/ or from our web site at www.scripps.edu/brooks/ficio/parameters/param.html, 1997.
25. Miranker, A. and Karplus, M., Proteins, 11 (1991) 29.
26. Cerius2, Molecular Simulations Inc., San Diego, CA, 1997.
27. Catalyst, Molecular Simulations Inc., San Diego, CA, 1997.
28. Jain, A.K. and Dubes, R.C., Algorithms for clustering data, Prentice Hall, Englewood Cliffs, NJ, 1988.
29. Warwicker, J. and Watson, H.C., J. Mol. Biol., 157 (1982) 671.
30. Edinger, S.R., Cortis, C., Shenkin, P.S. and Friesner, R.A., J. Phys. Chem., B101 (1997) 1190.
31. Allen, F.H., Bellard, S., Brice, M.D., Cartwright, B.A., Doubleday, A., Higgs, A., Hummelink-Peters, T., Kennard, O., Motherwell, W.D.S., Rodgers, J.R. and Watson, D.G., Acta Crystallogr., 35 (1979) 2331.
32. Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J. and Salemme, F.R., Science, 243 (1989) 85.
33. Segal, D.M., Padlan, E.A., Cohen, G.H., Rudikoff, S., Potter, M. and Davies, D.R., Proc. Natl. Acad. Sci. USA, 71 (1974) 4298.
34. Weis, W.I., Brünger, A.T., Skehel, J.J. and Wiley, D.C., J. Mol. Biol., 212 (1990) 737.
35. Robbins, A.H. and Stout, C.D., Proc. Natl. Acad. Sci. USA, 86 (1989) 3639.
36. Noble, M.E.M., Wierenga, R.K., Lambeir, A.M., Opperdoes, F.R., Thunnissen, A.M.W.H., Kalk, K.H., Groendijk, H. and Hol, W.G.J., Proteins, 10 (1991) 50.
37. Walter, J., Steigemann, W., Singh, T.P., Bartunik, H., Bode, W. and Huber, R., Acta Crystallogr. B, 38 (1982) 1462.
38. Christianson, D.W. and Lipscomb, W.N., Proc. Natl. Acad. Sci. USA, 83 (1986) 7568.
39. Holden, H.M. and Matthews, B.W., J. Biol. Chem., 263 (1988) 3256.
40. Leslie, A.G.W., J. Mol. Biol., 213 (1990) 167.
41. Tronrud, D.E., Monzingo, A.F. and Matthews, B.W., Eur. J. Biochem., 157 (1986) 261.