DOI: 10.1002/cmdc.200700010

From Modeling to Medicinal Chemistry: Automatic Generation of Two-Dimensional Complex Diagrams

Katrin Stierand and Matthias Rarey*[a]

As a result of the increasing application of structure-based drug design, the visualization of protein–ligand complexes has become an important feature in medicinal chemistry. The large number of experimentally resolved complex structures and the further development of computer-aided methods like docking or de novo design establishes new possibilities in this field. During lead finding and optimization, a manual investigation of many complexes and their interaction patterns is typically performed. We present an algorithm that automatically generates 2D-protein–ligand diagrams as a possible solution for a transparent visualization of the contact partners in a complex and as a support for scientists in the evaluation of structure-based design results. Running the software on representative test data sets, it generates collision free layouts for ~76% of the cases in the range of tenths of a second per complex. The success rate for complexes with ligands which have a molecular weight < 500 Da is 87%.

Introduction

The interaction pattern between proteins and their ligands is an important deciding criterion during structure-based lead identification and optimization. The manual evaluation of protein–ligand complexes plays an important role for the modeler in drug discovery. Besides other attributes the interaction pattern determines the further usage of the ligand in the lead generation process and gives important hints for further structure modifications. The method of choice for evaluating molecular complexes is a graphical 3D representation requiring a time consuming analysis to acquire all interaction information. Especially when a large number of complexes are available for analysis, the two-dimensional illustration of protein–ligand complexes including visualization of both hydrophobic contacts and hydrogen bonds is a helpful medium for the modeler to communicate results from structure-based design efforts to the medicinal chemist. Although not all information is presented, the 2D sketch of the complex is useful as it focuses on the most important aspects of molecular interactions.

In this paper we present an algorithm implemented in the software tool POSEVIEW[1] that automatically arranges a ligand, hydrogen bonds, metal interaction, and the corresponding amino acids and co-factors as well as hydrophobic contacting amino acids in a complex diagram. The layout is calculated independently from the 3D coordinates based only on the connection tables of the molecules and the interaction data of the complex. Whereas hydrogen bonds between protein and ligand are drawn as dashed lines and the appropriate residues and the ligand are visualized as structure diagrams, the hydrophobic contacts are represented more indirectly by means of spline sections around the ligand and the label of the contacting amino acid. A spline section highlights the hydrophobic part of the ligand that contains most of the atoms participating in the contact to the depicted amino acid. The aim is to generate a complex diagram without collisions between the different diagram elements—the structure diagrams, the interaction lines, the spline segments, and the amino acid labels—maintaining a clear and easily ascertainable arrangement.

A formerly developed program called Ligplot[2] automatically plots protein–ligand interactions. Besides POSEVIEW it is the only tool known from the literature that deals with the problem of automatic generation of 2D depictions of complexes. However, in contrast to POSEVIEW, the 2D layout is directly derived from the 3D coordinates. All molecules included in the complex plot are connected to the ligand by either one or more hydrogen bonds or, in the case of hydrophobic contacts, a virtual bond between the contact atoms and the residue. The resulting 3D graph is flattened on the 2D plane by unrolling it. A subsequent cleanup method is applied to the plot to remove as many collisions between structures and interactions as possible. Additionally, the software tool MOE offers the possibility of 2D depiction of ligand-interaction diagrams.[3] In this case the interacting residues of the active site are represented as colored dots only, such that the interacting atoms on the protein side cannot be identified.

This paper gives a methodical overview of the POSEVIEW algorithm. The new method is applied to three different representative subsets from the Brookhaven Protein Data Bank (PDB).[4] We can show that for 76.3% of all complexes a collision-free layout can be generated, whereas for 16.4% collisions

[a] K. Stierand, Prof. Dr. M. Rarey
Center for Bioinformatics
University of Hamburg
Bundesstrasse 43
20146 Hamburg (Germany)
Fax: (+49) 40-428387352
E-mail: rarey@zbh.uni-hamburg.de

occur although a collision-free layout is possible in principle, and for 3.4% of the cases a collision-free drawing is impossible because of the reduction from three to two dimensions. The remaining 3.9% could not be drawn for different reasons. Further on, different complex examples of the proteins HIV-1 protease (1hvr), flavodoxin (1akr), estrogen receptor alpha (3erd, 3ert), and chymotrypsin (1ghb) will be discussed in detail. Finally an application scenario for the protein cyclin dependent kinase 2 with different ligands will be presented and the use of POSEVIEW in drug discovery projects will be discussed.

Methods

The workflow of POSEVIEW is subdivided into several steps (see Figure 1). Details concerning the arrangement of hydrogen bonds and the corresponding amino acids can be found in a former publication.[1]

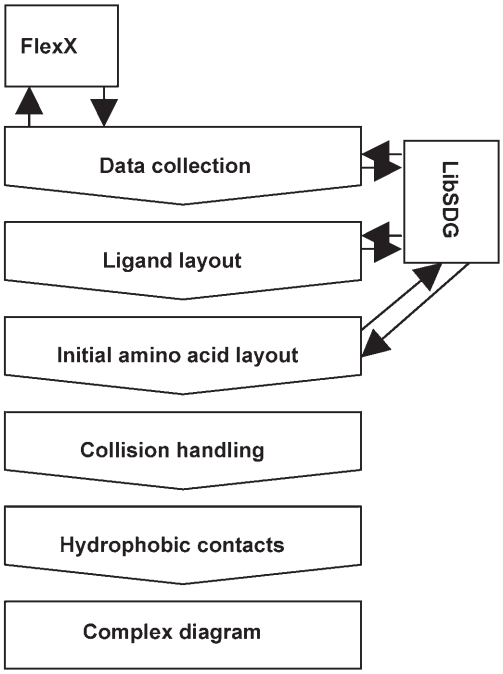

Figure 1. Workflow of POSEVIEW. The call of the external programs FlexX and LIBSDG by the different modules of POSEVIEW are marked by arrows.

For data preparation, FlexX [5] is used to read the protein and the ligand file and estimate the interaction energy between receptor and ligand or compute a docking result. Based on these calculations, the selection of complex diagram elements is performed. Amino acids are selected for drawing if their related hydrogen bond and metal interactions offer a score contribution of at least one third of the maximal hydrogen bond score ($-1.55\ kJ\,mol^{-1}$) calculated by Böhm's scoring function[6] with minor changes as it is used in FlexX; all other interactions are ignored. The inclusion of residues with hydrophobic contacts depends on their distance to the ligand and the number, accessibility, and type of the contact atoms. In this context, the following elements are treated as hydrophobic: carbon, sulfur (depending on its binding mode), chlorine, bromine, and iodine. If the number and geometry of outgoing bonds of a hydrophobic atom provides surface accessibility, for example sp2 hybridization and three outgoing ring bonds for carbon atoms, then this atom is selected as a contact atom. The maximum distance between two contact partners is set to the sum of both van-der-Waals radii plus a tolerance of 0.8 Å. An amino acid is included in the complex diagram if at least three contacts to different atoms of the ligand are found independent of its known tendency to make hydrophobic interactions. Co-factors and water molecules are selected for drawing, if any interactions with the required strength to the ligand are found. An example for the inclusion of a co-factor is shown in Figure 2: FlexX calculated a metal interaction with maximal

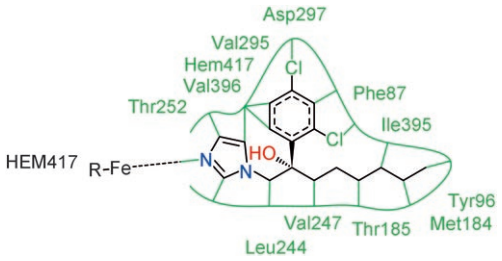

Figure 2. Some base points of the spline are joined or deleted to avoid loops. The straight green lines show the connection between the base points and the corresponding atoms. Three base points were joined and two base points were deleted: one point corresponding to the hydroxyl group and one corresponding to a carbon in the aliphatic chain on the right-hand side of the molecule.

score ($-4.7\ kJ\,mol^{-1}$) between the Hem417 and an imidazole nitrogen of the ligand. Additionally, hydrophobic contacts between the porphyrin and the surrounding ligand atoms are found. Note that interactions between co-factors or water molecules and protein residues are not shown.

The selected amino acids with hydrogen bonds and the ligand structure diagrams are generated by LibSDG.[7] The structure diagrams are generated independently from the 3D coordinates only based on the atom connectivity and the following layout arrangement is calculated on the basis of the new 2D coordinates. Important information about the configuration of the molecules, such as *cis*/*trans*-isomerism or stereocenters, are maintained and correctly displayed in the structure diagrams.

A complex diagram without collisions between the structure diagrams of the molecules and the interaction lines requires a ligand layout that allows an intersection-free arrangement of hydrogen bonds. Thus, all possible ligand layout modifications are enumerated until the neighborhood of interaction atoms with hydrogen bonds to the same residue is not interrupted by other interaction atoms. Modification operations are rotations of bonds and exchanges of two bonds which are connected to the same atom; all modifications applied to the ligand maintain the stereo information.

Subsequent to the ligand layout calculation, the selected residues with hydrogen bonds are initially placed on the basis

of the convex hull[8] of the ligand atoms. The directions of interaction lines determine the positions of the amino acid structure diagrams: they are first calculated and then cut at a certain distance from the ligand. Finally the amino acid structure diagrams are placed with the interaction atoms at the end of the interaction. If the ligand interaction atom is a vertex of the convex hull, the interaction direction extends the resulting direction of all bonds leading to this atom. For all other cases, the direction is set perpendicular to an appropriate edge of the convex hull.

As a consequence of the amino acid positioning, collisions between the diagram elements may occur. Hence a collision handling is performed by moving the amino acids on a grid such that each diagram can be placed in nine different positions including the initial one. Employing a branch and bound method[9] the best relative arrangement of amino acids is computed.

Hydrophobic contacts between the ligand and the amino acids are represented in a more implicit way than hydrogen bonds in the complex diagram. Whereas the amino acids are drawn as labels, the hydrophobic regions of the ligand are highlighted by spline segments. The arrangement of hydrophobic contacts is done subsequent to the described hydrophilic layout calculation without a rearrangement of the former placed diagram elements. It is subdivided into two steps: at first the spline segments around the ligand are calculated and afterwards the labels are placed based on the coordinates of the selected spline segments.

The cubic spline around the ligand is calculated on the basis of a polygon whose vertices are a selection of ligand atoms. Two different atom types are selected as vertices: Terminal atoms and atoms with two outgoing bonds if the exterior angle of the two bonds is oriented outwards with respect to the convex hull of the ligand. For the selection only explicitly drawn atoms are considered, implicit hydrogens do not influence the degree of an atom; the carbon of a terminal methyl group, for example, has the bond degree one. Outgoing from this polygon the base points for the cubic spline are calculated by shifting the vertices along the bisecting line of the polygon angles 1.5 standard bond lengths outwards.

In a post-processing step, base points which cause loops are removed. The reasons for loops can either be two base points which are located a small distance from each other or an interruption of the circular order of the base points, which can be generated by shifting the base points outwards or by ligand structure diagrams with a concave bond arrangement (see Figure 2).

After the spline generation the hydrophobic areas of the ligand are identified and the corresponding segments are cut from the spline; all other spline parts are deleted. Hydrophobic areas are defined as a continuous sequence of polygon vertex atoms which fulfill the criteria of a hydrophobic atom as described at the beginning of the method section. The spline segments are selected for drawing according to the list of contact atoms of the hydrophobic contacting amino acids. If one amino acid has contacts to more than one spline segment, the segment that contains most of the contact atoms is chosen. Spline segments which have no corresponding residue are deleted.

In case of collisions between spline segments and other complex diagram elements, like hydrogen bonds or amino acid structure diagrams, the affected spline segment parts are deleted. If the segment is subdivided in two sections by a deletion, the shorter end is also deleted such that each hydrophobic region is represented by one continuous line.

In analogy to the hydrogen bond directions, a virtual direction for each hydrophobic contact is computed. Initially, the label is placed on the intersection of the virtual interaction and the spline segment that is assigned to the considered residue. In some cases, the interaction direction does not intersect the spline segment, as the bonds which influence the interaction direction do not point towards the segment. Then the label of the residue is moved to the nearer end of the selected segment.

Following the initial placement, a collision handling is performed. Compared to the other diagram elements, the labels of the hydrophobic contacting residues are relatively small. Thus, placing the labels is done very flexibly and new collisions can be avoided. The collision handling is done by means of a grid that covers the smallest enclosing box of the complete diagram. Each grid point within a convex polygon of the amino acid atoms, the cubic spline around the ligand, or on the labels of the amino acid structure diagrams and the interaction lines is labeled as occupied. Then all hydrophobic labels are iteratively moved to the next free space on the grid. Moving away from the initial position the grid is searched for free spaces with two different strategies: first, along the spline segment and second, in a spiral around the starting point.

Finally an output is generated either in a browser or in a png, pdf, or xfig file.

Results and Discussion

To estimate the performance of POSEVIEW, the algorithm was applied to three different published data sets (PDBbind,[10] CCDC Astex,[11] and FlexX200;[12] the data sets are not disjunctive) containing a representative selection of complexes from the PDB database. All ligands in the chosen complexes have a molecular weight between 50 and 1000 Da. On average, the computing time for a complex layout generation on a standard PC (Intel Xeon2 EMT 64, 3.0 GHz, 2 GB main memory running under SuSE Linux 9.2) amounts to 0.64 s and the data preparation by FlexX is done in ~2 s per complex. Whereas the collision-free layouts could be generated in an average of 0.24 s, the diagrams of lower quality show higher computing times. This difference is due to the collision handling strategy: the number of iterations increases with the number of collisions and complex diagram elements.

The resulting diagrams are subdivided into three quality groups. The first group contains complex diagrams with collision-free layouts whereas the second and the third groups consist of diagrams with collisions. The distinctive feature between the second and the third group is the reason for the collisions. Whereas in the second group the ligand layout pro-

vides a collision-free arrangement of diagram elements (improvable layouts), the interactions in the third group cannot be arranged intersection-free because of the reduction from 3D to 2D coordinates (unsolvable layouts). As described in the method section, the arrangement of the diagram elements visualizing the hydrophobic contacts is done very flexibly. Thus, no new collisions are caused by the addition of hydrophobic contacts. Detailed information about computing times and result quality can be found in Table 1 and Table 2.

POSEVIEW performs on the FlexX200 and the CCDC Astex data set in computing time and number of good results better than on the PDBbind data set. This relates to the different average molecular weights and numbers of amino acids and interactions (see Table 3).

Figure 3 shows the percentage of the different layout qualities subdivided in bins of 100 in the range of 1 to 1000 Da. For all three data sets POSEVIEW performs well on complexes with ligands which have a molecular weight <600 Da. According to Lipinski's rule of five,[13] drug-like molecules frequently have a weight <500 Da, such that a good applicability of POSEVIEW in drug discovery can be expected. In the case of the PDB bind data set the results for complexes with a molecular weight >600 Da are of more composite quality than for the other data sets because many of these complexes have just a small number of hydrogen bonds. For the other two data sets, the number of interactions and amino acids increases with the molecular weight.

Table 1. POSEVIEW applied on different data sets.

	FlexX200		PDBbind		CCDC Astex	
	Number	%	Number	%	Number	%
Number of complexes in data set	200	100.0	793	100.0	305	100.0
Collision-free layouts	166	83.0	601	72.5	244	80.0
Improvable layouts	25	12.5	140	18.6	38	12.5
Unsolvable layouts	6	3.0	24	4.0	6	2.0
No interactions/contacts	0	0.0	6	1.0	2	0.6
Not drawn[a]	3	1.5	28	3.9	15	4.9

[a] Some complex diagrams could not be drawn for the following technical reasons: complex ring systems in the ligand molecule or difficulties during the data preparation. For complexes with more than 17 hydrogen bonds or 15 amino acids no layouts are computed because the computing times for more elements lie in the range of minutes to hours.

Table 2. Computing times of POSEVIEW for different data sets.

	FlexX200	PDBbind	CCDC Astex
Computing time/complex	0.25 s	0.84 s	0.82 s
Collision-free layouts	0.11 s	0.53 s	0.09 s
Improvable layouts	0.85 s	0.80 s	5.77 s
Unsolvable layouts	1.70 s	7.62 s	0.24 s

Table 3. Averaged attributes of protein–ligand complexes from the three different test data sets.

	FlexX200	PDBbind	CCDC Astex
Molecular weight of the ligand [Da]	332.8	408.9	349.5
Number of amino acids forming hydrogen bonds	3.8	4.2	3.4
Number of hydrogen bonds	4.8	5.3	4.2

The results for the different data sets show a strong dependency between the number of interactions and amino acids and the computing time of the layout generation. Figure 4 shows the computing times depending on the number of interactions for all three data sets in a plot. The variation in computing times increases with the number of interactions. This is caused by the distribution of the amino acid structure diagrams around the ligand: if more than one interaction with different amino acids starts at the same ligand atom, then the collision handling algorithm is needed to find an overlap-free layout. Often small hydrophilic ligands have many interactions and their layout provides an intersection-free arrangement, however, it is sterically not possible to arrange all amino acids overlap-free with the given length of interaction lines. Another reason for high computing times can be large topological distances between two ligand atoms which interact with the same amino acid. In this case it is sometimes not possible to optimize the ligand layout such that the interactions do not cross any structure diagram part.

Selected examples

In the following, some selected complex diagrams will be presented pointing out the different challenges during the layout generation. Again, the focus lies on the hydrophobic part of the complexes.

A large contribution arises to the energy of binding between ligand and receptor by hydrophobic contacts for the HIV-1 protease and its inhibitors. Figure 5 shows a complex diagram containing a ligand that was developed by structure-based drug design. The aim was to replace a water molecule in the active site that forms two hydrogen bonds to the native ligand and one to the isoleucines 50 of both protein chains. The complex diagram contains three amino acids which form hydrogen bonds: the two catalytic aspartates and the isoleucine 50 A. The isoleucine 50 B was omitted, because the contribution of the interaction falls below the specified threshold. The two hydrophobic parts of the ligand are flanked by long spline segments. In the post-processing step of the spline generation, for both spline segments a base point in the cavity between the two covered substituents was deleted and two other base points were joined to eliminate loops from the spline. Whereas the upper spline segment has got a steady run, for the lower spline segment a deletion of one more base point would be nec-

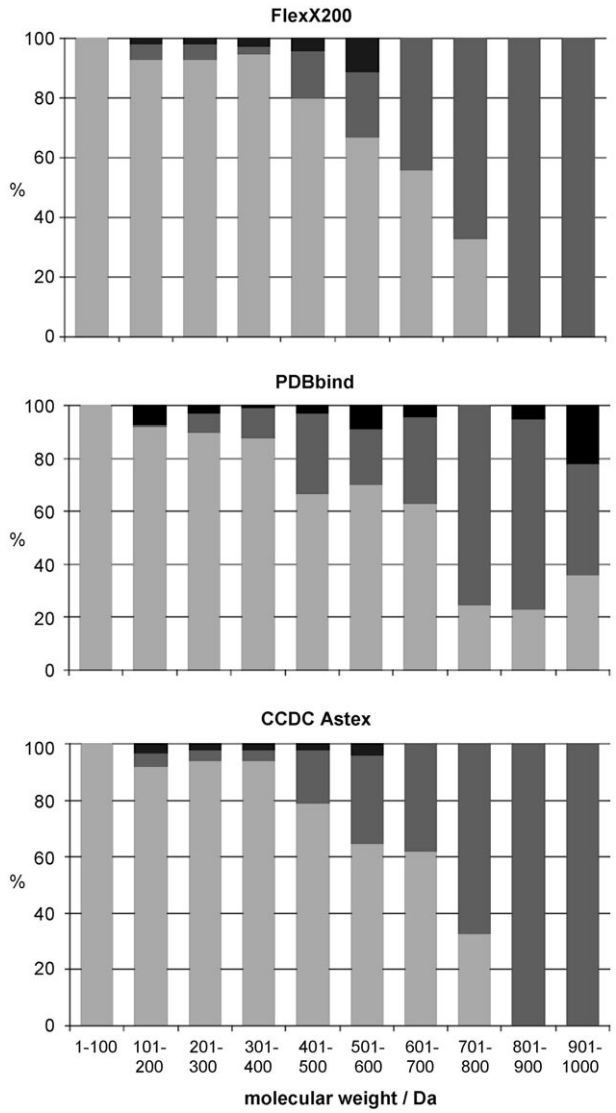

Figure 3. Comparison of percentage of different layout qualities for different molecular weights for the three test data sets (light gray: collision-free layouts, dark gray: improvable layouts, black: unsolvable layouts).

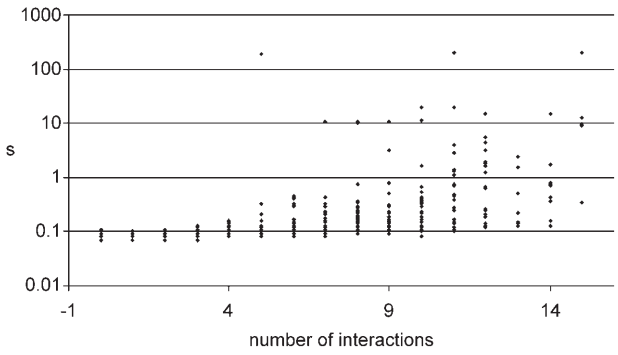

Figure 4. Increase in computing time with the number of interactions.

essary to reach the same result. The right end of the upper segment was cut to avoid intersections with the hydrogen bond between the carbonyl group of the ligand and the isoleucine 50 A.

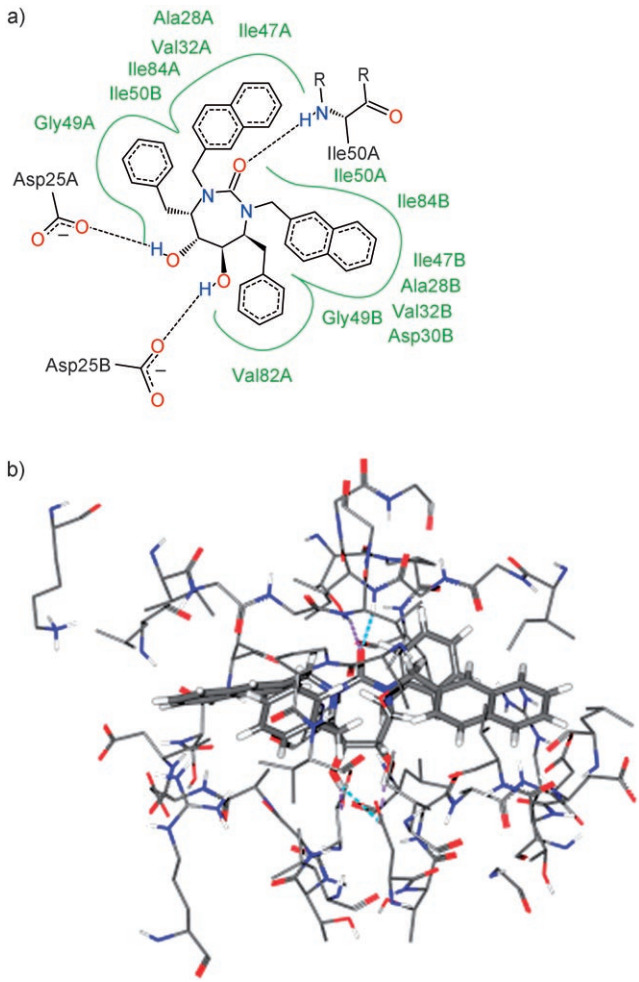

Figure 5. HIV-1 protease complexed with a substituted 1,3-Diazepin-2-one (PDB code: 1hvr)[14] drawn as a) complex diagram and b) in 3D.

The active site of the HIV-1 protease is a symmetrically arranged homodimer. As a consequence the three-dimensional representation of the ligand shows a symmetrical arrangement of the ring substituents, but this symmetry is not represented in the diagram because it is not considered in the drawing methods. As other symmetric ligands are arranged asymmetrically in the active site to form an optimal hydrogen bonding pattern we decided not to check for further active site symmetry which would be possible in principle.

Flavodoxins are electron-transfer proteins that function in various electron-transport systems. They are not found in mammalian, but in pathogens like Helicobacter pylori.[15] In contrast to the HIV-1 protease complex, the complex of flavodoxin and flavin mononucleotide is characterized by a large number of hydrogen bonds. The hydrophobic contact area of the ligand is reduced to a part of the flavin ring system (see Figure 6), and only two amino acids fulfilling the hydrophobic contact criteria were found. The overlap-free layout requires a collision handling for the amino acid structure diagrams because the initial interaction directions for the pairs of amino acids which are interacting with the same ligand atom are identical. Furthermore, a layout optimization for the amino

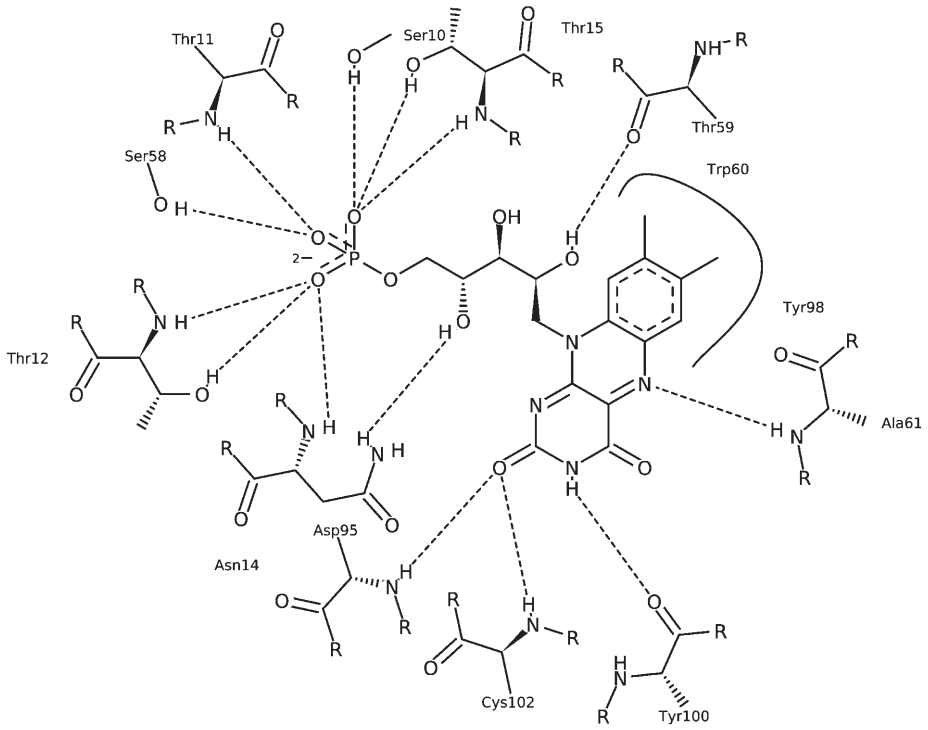

Figure 6. Complex of flavodoxin and flavin mononucleotide (PDB code:1akr).[16]

areas and hydrogen bonds whereas in the other case the ligand shows parts which are not covered by the active site. According to the symmetric ligand layout in Figure 7, the drawn spline segments are also symmetric. The clustered distribution of the amino acid labels has its origin in the positions of the contact atoms.

Chymotrypsin is a digestive protein and cleaves peptides at the carboxyl side of tyrosine, tryptophan, and phenylalanine. Figure 9 shows the complex of chymotrypsin and *N*-acetyl-D-tryptophan. For this example a nonoptimal layout was computed. This is caused by the initial arrangement of the hydrophobic amino acid labels because the virtual interaction direction for all included residues is very similar. Moreover, the interaction line between the ligand and the serine 190E is situated near the spline section such that no label can be placed between these two diagram elements. Although no collision is caused by this placement, a method that computes a distribution of labels more bound to the spline segment would be preferable.

acids with two interactions to the ligand (threonine 12, threonine 15, and asparagine 14) was performed to approximate the interaction starting points.

Estrogen receptors are nuclear receptors implicated in different types of cancer. Figure 7 shows the complex of an estrogen receptor with the agonist diehtylstilbesterol and Figure 8 shows the complex with the antagonist 4-hydroxytamoxifen. Tamoxifen is a drug used in the therapy of breast cancer. In the 3D pose the diehtylstilbesterol is completely buried, the tamoxifen points partly out of the active site of the receptor. This is also represented in the 2D depictions, in the one case the ligand is completely surrounded by hydrophobic contact

Figure 10 shows two different ligands from different inhibitor classes[19] and their interactions with the active site of the protein cyclin dependant kinase 2 (CDK2) (PDB codes: 1e1x,[20] 1h1s).[21] Both ligands form three hydrogen bonds to the hinge region that consists of the residues glutatmate 81 and leucine 83. The orientation of the ligands with respect to the hinge region is different for the complex diagrams; in Fig-

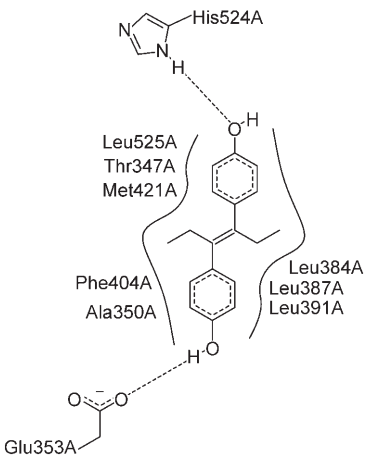

Figure 7. Complex of estrogen receptor alpha and diehtylstilbesterol (PDB code: 3erd).[17]

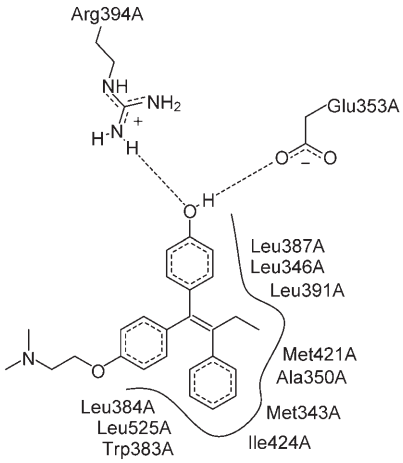

Figure 8. Complex of estrogen receptor alpha and 4-hydroxytamoxifen (PDB code: 3ert).[17]

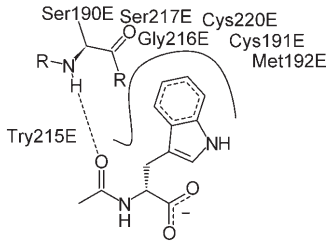

Figure 9. Complex of gamma-chymotrypsin and N-acetyl-D-tryptophan (PDB code: 1ghb).[18]

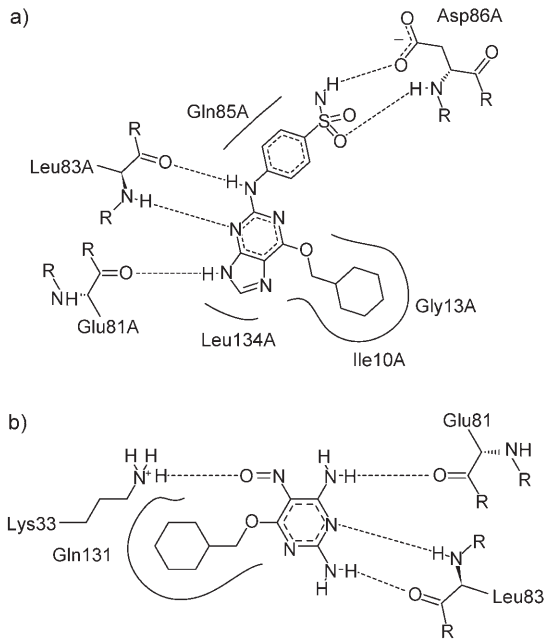

Figure 10. Two different inhibitors in complex with cyclin dependent kinase 2 (PDB codes: a) 1e1x,[20] b) 1h1s).[21]

ure 10a the hinge region is situated to the right of the ligand and in Figure 10b to the left of the ligand. For a quick comparison of the interaction patterns, a similar orientation would be helpful. In this case the common parts of the active sites would be ascertainable instantaneously for the viewer. Also, a presentation that emphasizes the similarities between the ligands could be useful for an easily readable representation of the complexes.

As presented in this paper, POSEVIEW has a wide range of application in structure-based drug design. It can be a helpful medium in the communication between modeler and medicinal chemist for the discussion about single complexes and complex series. Especially the illustration of series of related complexes, for example from a docking experiment with a special target, offers potential for further methodical improvement as was discussed with the example of Figure 10.

Apart from the structure-based drug design, it is further conceivable to find other application fields for POSEVIEW. Although the 3D structures of some protein families are not known today, they represent an interesting target for pharmaceutical intervention. Based on the amino acid sequence, for some of these proteins the binding mode between ligand and receptor could be determined experimentally, for example by mutations. As POSEVIEW is independent from 3D coordinates, it would be able to generate complex diagrams of the same quality based on the interaction pattern and the contact atoms only.

Availability

POSEVIEW is available as a web service at http://www.zbh.uni-hamburg.de/poseview. Complex diagrams can be generated from a PDB file alone or from a PDB file for the protein and a separate mol2 file for the ligand. The png output can be viewed in the web interface and a pdf-file of the plot can be downloaded. The stand alone tool with full functionality will be available shortly from BioSolveIT GmbH (http://www.biosolveit.de).

Conclusions

POSEVIEW is a program based on a multiphase algorithm for the automatic generation of complex diagrams which contain the ligand, the interacting amino acids of the active site, and the hydrogen bonding pattern. The diagram generation is done in the range of tenths of a second, which makes POSEVIEW applicable for large numbers of complexes whereas the quality is in most cases comparable to hand-generated plots.

POSEVIEW generates complex diagrams, which show a clear arrangement of all their elements such that the contained information is easily ascertainable for the user. Redundant information like noninteracting parts of amino acid structure diagrams are removed from the plot to keep it as simple as possible. Beyond the depicted details no further information about the complex is available. For many applications it would be conceivable to know details such as the length or strength of a hydrogen bond or the contact ligand atoms of a hydrophobic amino acid. The mentioned additional features could be added with little effort, but this leads to a conflict between the attempt at simplicity and the need for details. A possible solution would be to provide these features as optional and adjustable by the user.

Acknowledgements

The authors thank Patrick C. Maaß for providing the LibSDG and for many fruitful discussions during the development phases of POSEVIEW.

Keywords: 2D visualization · molecular interactions · protein–ligand complex · structure diagrams

[1] K. Stierand, P. Maaß, M. Rarey, *Bioinformatics* **2006**, *22*, 1710.
[2] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng. Des. Sel.* **1995**, *8*, 127.
[3] A. M. Clark, Chemical Computing Group Inc., **2006**.
[4] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235.

[5] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, *J. Mol. Biol.* **1996**, *261*, 470.
[6] H. J. Böhm, *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243.
[7] P. C. Fricker, M. Gastreich, M. Rarey, *J. Chem. Inf. Comput. Sci* **2004**, *44*, 1065.
[8] R. L. Graham, *Information Processing Lett.* **1972**, *1*, 132.
[9] E. L. Lawler, D. W. Wood, *Oper. Res.* **1966**, *14*, 699.
[10] R. Wang, X. Fang, Y. Lu, S. Wang, *J. Med. Chem.* **2004**, *47*, 2977.
[11] J. W. Nissink, C. Murray, M. Hartshorn, M. L. Verdonk, J. C. Cole, R. Taylor, *Proteins* **2002**, *49*, 457.
[12] B. Kramer, M. Rarey, T. Lengauer, *Proteins* **1999**, *37*, 228.
[13] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **2001**, *46*, 3.
[14] P. Y. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong et al., *Science* **1994**, *263*, 380.
[15] J. Freigang, K. Diederichs, K. P. Schäfer, W. Welte, R. Paul, *Protein Sci.* **2002**, *11*, 253.
[16] P. A. O'Farrell, M. A. Walsh, A. A. McCarthy, T. M. Higgins, G. Voordouw, S. G. Mayhew, *Biochemistry* **1998**, *37*, 8405.
[17] A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene, *Cell* **1998**, *95*, 927.
[18] H. P. Yennawar, N. H. Yennawar, G. K. Farber, *J. Am. Chem. Soc.* **1995**, *117*, 577.
[19] R. Buijsman in *Chemogenomics in Drug Discovery, Vol. 22* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, p. 191.
[20] C. E. Arris, F. T. Boyle, A. H. Calvert, N. J. Curtin, J. A. Endicott, E. F. Garman, A. E. Gibson, B. T. Golding, S. Grant, R. J. Griffin, P. Jewsbury, L. N. Johnson, A. M. Lawrie, D. R. Newell, M. E. Noble, E. A. Sausville, R. Schultz, W. Yu, *J. Med. Chem.* **2000**, *43*, 2797.
[21] T. G. Davies, J. Bentley, C. E. Arris, F. T. Boyle, N. J. Curtin, J. A. Endicott, A. E. Gibson, B. T. Golding, R. J. Griffin, I. R. Hardcastle, P. Jewsbury, L. N. Johnson, V. Mesguiche, D. R. Newell, M. E. M. Noble, J. A. Tucker, L. Wang, H. J. Whitfield, *Nat. Struct. Mol. Biol.* **2002**, *9*, 745.

Received: January 15, 2007
Revised: February 9, 2007
Published online on April 13, 2007

