SH2 Binding Site Comparison: A New Application of the SURFCOMP Method

Christian Hofbauer and András Aszódi*

In Silico Sciences Unit, Informatics and Knowledge Management, Novartis Institutes for BioMedical Research Vienna, Brunnerstrasse 59, A-1235 Vienna, Austria

Received October 1, 2004

To avoid side effects, it is often desirable to increase the specificity of a drug candidate when targeting one member of a family of related proteins, whereby one exploits small differences between the structures of the binding sites. Identification of such differences can be carried out by analyzing the distributions of physicochemical properties mapped onto molecular surfaces. Here we demonstrate that SURFCOMP, our local surface similarity detection method, is able to detect differences between the binding sites of two closely related proteins. We analyzed the SH2 domains of Sap and Eat-2, two highly similar signal transduction molecules involved in inflammatory processes and found differences between their binding sites that can possibly lead to a better understanding of the different specificities of the two proteins.

INTRODUCTION

A main effort in modern structural biology is the detection of functional similarities and dissimilarities between proteins. Such relations are usually established by means of the amino acid sequence or the three-dimensional structure of the proteins. Although the sequence information is easily available for most proteins, the function of a protein is in general determined by its three-dimensional structure. Structures are more conserved through evolution than sequences, and it is therefore not surprising that functional and structural similarity can often be detected between proteins that do not have any significant sequence identity. Similarly, proteins with very similar amino acid sequences can show considerable differences in their 3D structures as will be demonstrated on Sap and Eat-2, two promising drug targets of the SH2 family.

The Src homology 2 (SH2) domain, which is an approximately 100 amino acid long conserved domain found in a large number of proteins, is a key element in tyrosine kinase regulation of cellular processes. SH2 domains play an important role in phosphorylation-based signal transduction mechanisms, where the trigger event is the binding of the SH2 domain to peptide sequences that contain phosphorylated tyrosine residues (pTyr). Blocking the protein−protein interactions of SH2 domains is a promising strategy for fighting a variety of different diseases reaching from cancer to inflammatory diseases.[2] But to avoid side effects, it is absolutely crucial to target only one member of the SH2 family by an inhibitor in a highly selective manner.

Sap (**s**ignalling lymphocyte activation molecule (Slam)-**a**ssociated **p**rotein) is a protein composed of a single SH2 domain that inhibits signal transduction events initiated by a series of receptors on the surface of T lymphocytes and natural killer cells. Sap interacts with the consensus motif in the cytoplasmic tail of Slam (CD150) in the phosphorylated and—unlike other SH2 domains—also in the dephosphorylated form, thereby blocking the recruitment of the Shp-2 phosphatase to that position in the receptor. Recently two groups independently discovered that this interaction is mediated via the kinase FynT.[3,4] Furthermore, a mutation in the gene encoding Sap (*SH2DIA*) is involved in the X-linked lymphoproliferative disease (XLP), a rare immune disorder that renders the immune system unable to respond effectively to the Epstein−Barr virus.[5]

Eat-2 (**e**ws/fli1 **a**ctivated **t**ranscript **2**), expressed in macrophages and b-lymphocytes,[6] is a very similar SH2 domain to Sap (see sequence and structure alignments in Figure 1). Eat-2 too can be associated to Slam and acts as a Shp-2 blocker, but no interactions with the SH3 domain of FynT are reported. Analogously to Sap, it binds to the phosphorylated cytoplasmic tail of Slam, but unlike Sap it does not bind to the dephosphorylated receptor. Therefore, in contrast to Sap, the binding of Eat-2 to Slam is significantly more dependent on tyrosine phosphorylation. This selectivity toward pTyr and the different cellular localizations of Sap and Eat-2 make the system an interesting target for a selective blocking of the SH2/Slam interactions. Earlier studies revealed that the consensus sequence motif T/S-x-pY/Y-x-x-V/I is responsible for the Slam recognition of Sap,[7,8] where x represents any amino acid and pY stands for phosphotyrosine. The three conserved residues of this motif are located at three characteristic cavities on the surface of Sap, and corresponding sites can be found in Eat-2. It was now of particular interest for us to investigate these cavities and to detect any differences in the molecular structure around those regions. If such differences are found, they may highlight positions where an inhibitor could selectively bind to Sap but not to Eat-2.

Several methods have been published that are able to detect common structural motifs on proteins.[9−13] Considering the lock and key principle as the typical mechanism of biomolecular recognition and neglecting effects such as induced fit, shape, and the distribution of physicochemical properties on the reactive interface seem to be critical elements in the interaction patterns of proteins with ligands. These properties can be mapped onto solvent-accessible[14] or solvent-excluded

* Corresponding author phone: ++43 (1) 86634-452; fax: ++43 (1) 86634-727; e-mail: andras.aszodi@pharma.novartis.com.

10.1021/ci0497049 CCC: $30.25
Published on Web 03/02/2005

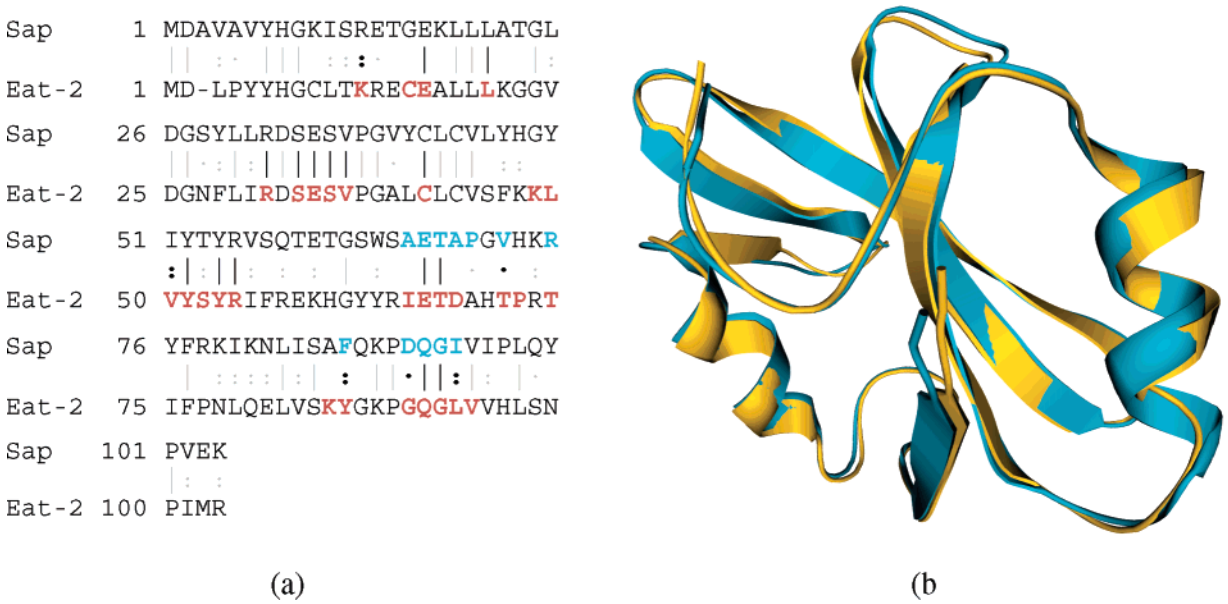

Figure 1. (a) Sequence alignment between Sap and Eat-2. The residues that are in close contact (6.0 Å) to the ligand peptide are displayed in blue (Sap) and red (Eat-2). A | means residue identity; : and · mean strong and weak evolutionary similarity. The alignment was performed by ClustalW with the PAM350 matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.1. (b) Structure alignment based on the backbone and side chain atoms of Sap (yellow) and Eat-2 (blue). Both the sequence and the structural superposition underline the strong similarity between the two proteins.

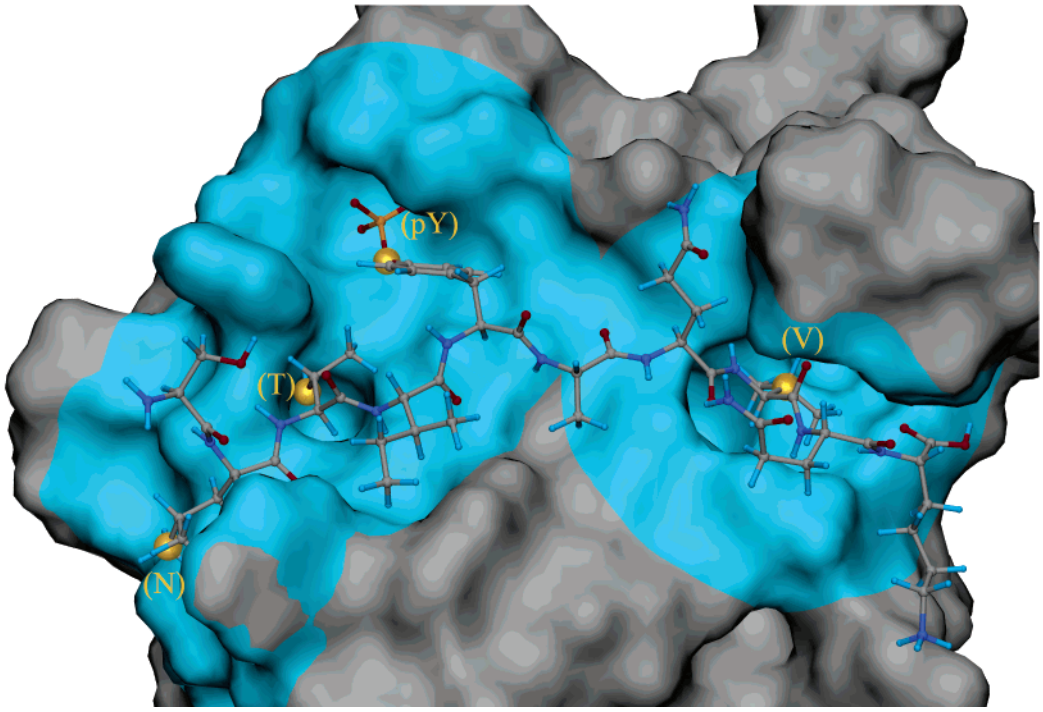

Figure 2. To detect differences in the molecular surface beneath the ligand peptide pSlam, sections of the molecular surfaces of two SH2 proteins Sap (picture) and Eat-2 were selected and analyzed for dissimilarities. The selection process was performed by marking all points on the surfaces (blue) within a cut off distance of 8 Å adjacent to selected anchor residues (yellow) of the pSlam peptide. The list of anchor residues includes the N-terminal residue (N), the threonine 279 (T), the phosphotyrosine 281 (pY), and the valine 284 (V) of pSlam.

molecular surfaces.[15] A comparison between such property-mapped surfaces of different proteins can then reveal common surface motifs that are functionally important. An excellent review about different surface comparison methods for proteins has been published by Via et al.,[16] who also discuss the relationship between sequence, structure, and surface. One application of molecular surface comparison is protein/protein and protein/ligand docking, and many algorithms have been published that use surface or shape complementarity as key heuristics.[17−21] Besides docking, several methods are now available that are able to detect global[22−26] or local surface similarities[27−29] including our own program SURFCOMP, which uses a set of chemical and shape filters to extract all possible local matches between molecular surfaces.[1] But for most of these programs, especially the more powerful local search methods, to the best of our knowledge, no applications to protein surface comparison have been reported so far.

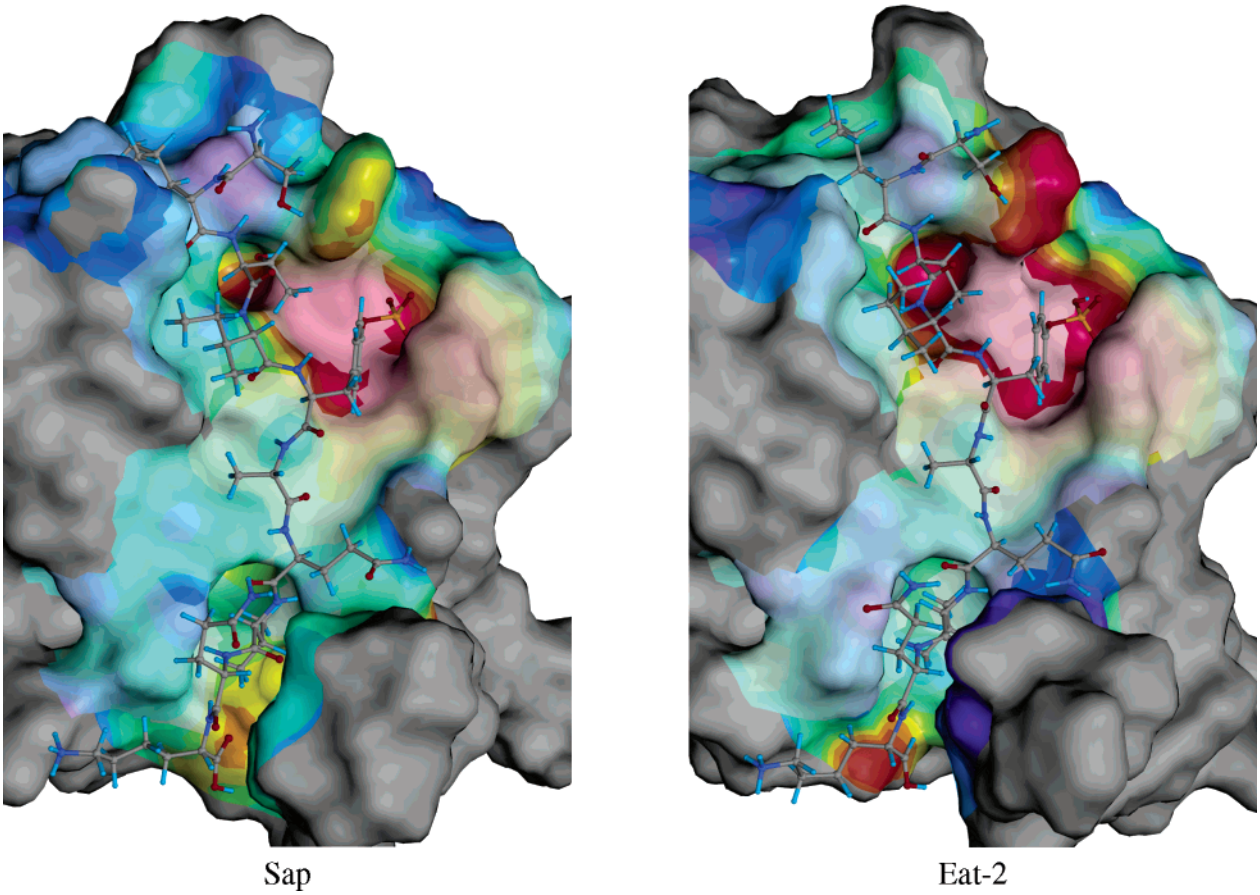

Figure 3. Differences in the surface areas of Sap (left) and Eat-2 (right) that are involved in the pSlam binding are shown in strong colors. Similar parts of the surfaces are represented as less intensive colors while the surface areas that were not compared are displayed in gray. The colors are encoding the electrostatic potential (ESP) of the surfaces, where blue indicates negative and red indicates positive areas. The surfaces are superimposed with the structure of the pSlam peptide.

The main obstacle with comparing macromolecular surfaces is that even at low resolution the number of points that form the surface is approximately 1 order of magnitude higher than for small compounds. Because most of the algorithms for surface comparison scale quadratically with the number of surface points, this implies a massive increase in computational time, memory usage, and number of candidate alignments that have to be evaluated. Fortunately, protein functional sites, where the interactions, with the ligands take place, usually cover only a small fraction of the total protein surface. Therefore, the comparison of two proteins can be reduced in many cases to the comparison of their functional sites. This will allow the investigation of the relevant parts of the proteins' surfaces at the same resolution as low molecular weight compounds.

In this paper we show that only minor modifications to SURFCOMP enable this algorithm to perform comparisons of protein functional sites. In combination with a comparative scoring methodology, the program was able to detect significant differences on the surfaces of Sap and Eat-2 that are in close contact to the phosphorylated signaling peptide Slam. These differences could possibly be the structural basis for a rational design of specific Sap inhibitors.

METHODS AND COMPUTATIONAL PROCEDURES

The dissimilarities between the molecular surfaces of Sap and Eat-2 have been analyzed by our program SURFCOMP. Since a detailed description of the algorithm has been published recently,[1] we provide here an overview of the method and explain only the details of the modifications that were necessary to apply SURFCOMP to the comparison of protein surfaces.

The methodology is based on an 3D structure search algorithm where maximal common subgraph isomorphism is used to detect local similarities.[30] To apply this principle to molecular surface comparison, the graph nodes were defined as critical points corresponding to "hills" and "valleys" on the surface and are augmented by surrounding surface patches. Since the number of these points can be quite large, it is necessary to reduce the complexity of the problem by a set of filters that implement various geometric and physicochemical heuristics. Among all possibly matching point pairs, those are selected first that show sufficient chemical similarity, defined by a fuzzy dissimilarity index F,[29] between physicochemical properties mapped onto the surface points. Then the curvature patterns around all remaining point pairs are compared by harmonic shape image matching[31] to discard points that are not embedded in a similar shape (determined by the correlation coefficient R between the patches attached to the points). Finally the distances and the bearings (relative patch orientations) between combinations of similar pairs are checked to be within certain boundaries (t and ϕ_{min}, respectively) to form an association graph that is then analyzed by clique detection.

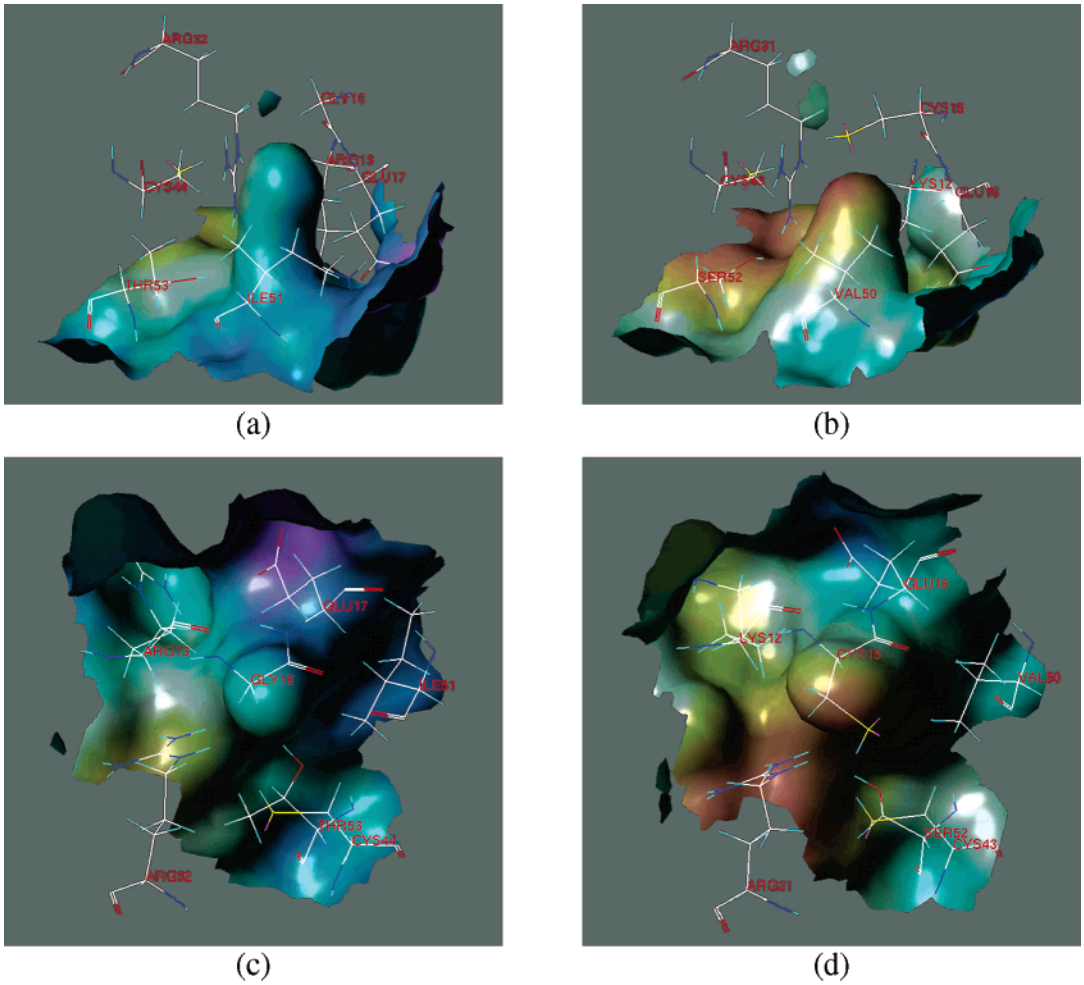

Figure 4. Structural conformation of the threonine cavity in Sap and Eat-2. All four images are presenting the inside of the cavities' surfaces of Sap (left) and Eat-2 (right). In panels a and b, the different depth of the pockets is illustrated; in panels c and d, the effect of the cysteine side chain is shown. From these, one can figure out easily how the mercapto-methyl group is limiting the extension of the cavity. The color code of the surface patches represents the electrostatic potential with red and orange as positively charged regions and blue and green as negatively charged regions.

The cliques represent the local surface similarities, and an alignment between the two molecular surfaces can be calculated based on the corresponding points. The alignments are clustered to reveal a picture of the total surface similarity between the two molecules.

The SURFCOMP program has been designed to retrieve all the detected similarities between two molecular surfaces. In the case of large molecules such as proteins the comparison produces a huge number of alternative solutions, which makes it difficult to determine the best matches by visual inspection of the alignments. It is therefore necessary to find a suitable ranking that automatically identifies the promising matches. To this end, we implemented a consensus scoring algorithm based on the rank-by-rank scheme,[32] which is used, for example, in molecular docking.[33] For each detected similarity, the program calculates the average rank determined by the root-mean-square deviation (RMSD) of the corresponding alignment, the number of corresponding surface points (N_{points}) that build the similarity, and the chemical correlation of these points (R_{chem}):

$$\text{consensus rank} = \frac{1}{3}[\text{rank(RMSD)} + \text{rank}(N_{points}) + \text{rank}(R_{chem})] \quad (1)$$

Thereafter the consensus ranks are sorted in ascending order to place the most promising matches at the top of the list. An evaluation of this scoring method is given in the Appendix.

To detect differences between Sap and Eat-2 concerning the mode of binding of Slam, the investigations focused on those sections of the molecular surface that were in close contact with the phosphorylated Slam (pSlam) peptide. We used the MOLCAD[34] module of the Sybyl[35] modeling package to generate solvent-excluded surfaces from the crystal structures of the Sap/pSlam (PDB entry 1D4W) and Eat-2/pSlam (PDB entry 1I3Z) complexes. Close contact was defined by selecting only those points of the surfaces that were located within 8.0 Å of the following atoms on the pSlam peptide: (1) the carbon atom of the closer methyl group in the side chain of leucine 278 (CD1), (2) the oxygen of the hydroxyl group of threonine 279 (OG1), (3) the oxygen connecting the phosphate group with the side chain of *p*-tyrosine 281 (OH), and (4) the β carbon in the side chain of valine 284 (CB).

The first center represents the N-terminal part of the ligand peptide, and the last three atoms are placed within the three binding cavities of the proteins that bind the conserved residues of the consensus sequence motif. Figure 2 shows

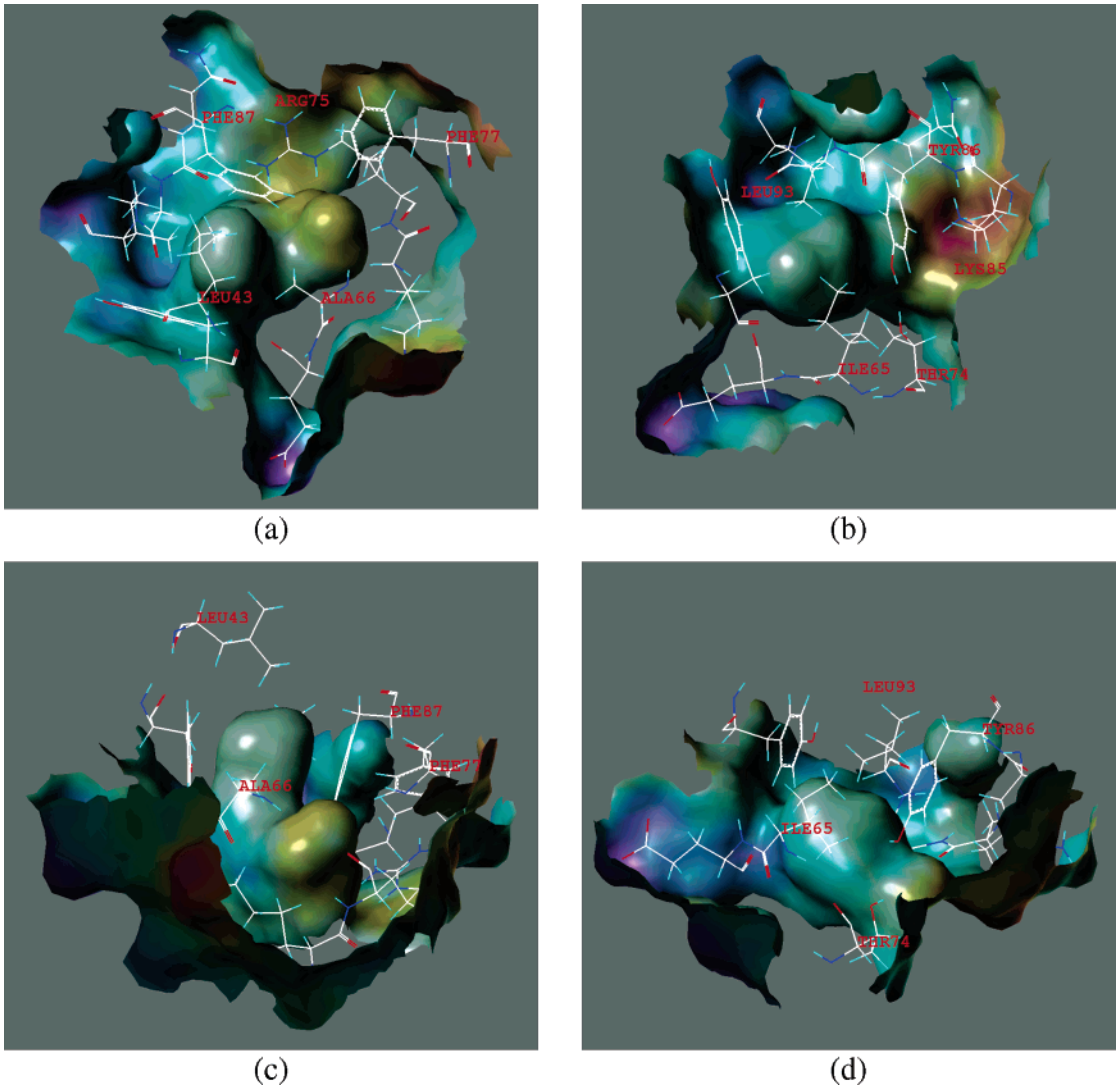

Figure 5. Structural conformation of the valine cavity in Sap and Eat-2. All four images are presenting the inside of the cavities' surfaces of Sap (left) and Eat-2 (right). The top row (a and b) shows which residues of both molecules define the borders of the cavities. The bottom row shows how Ile-65 and Leu-93 prevent the further extension of the pocket into the inner parts of Eat-2 (d) while the same hole reaches to Leu-43 in Sap (c). The color code of the surface patches represents the electrostatic potential with red and orange as positively charged regions and blue and green as negatively charged regions.

the molecular surface of Sap with the considered regions highlighted.

A surface comparison was performed for each of the four corresponding centers on Sap and Eat-2. To find all the possible differences, the SURFCOMP parameters were tuned in a way to retrieve only the most significant surface similarities: curvature cutoff range $c_{CR} = 2.0$ Å, neighborhood $n_{CP} = 2.0$ Å, fuzzy threshold $F = 0.3$, shape threshold $R = 0.6$, distance tolerance $t = 1.0$ Å, minimum distance $\delta_{min} = 0.5$ Å, and angular tolerance $\phi_{tol} = 15.0°$. For the physicochemical property used in the fuzzy filtering, we selected the electrostatic potential of the protein, which was calculated using the atomic point charges of the corresponding atom types of the Amber force field.[36] To simplify the ESP calculations, we neglected the atomic charges of the pSlam ligand, which was acceptable because we were interested in differences only and not in absolute values. Initially the results of each comparison highlighted only the differences in one region. To get the overall surface alignment of the complete binding area the best clusters of all four computations were combined.

To identify the residues that give rise to the dissimilarities in the binding surfaces, the alignments were visualized with SYBYL 6.9[35] together with the corresponding structural and surface data. In the molecular viewer, the residues that are responsible for the differences in that area could be identified easily, and the surface was regenerated only for those amino acids to focus the attention of the observer on the relevant parts.

RESULTS AND DISCUSSION

Figure 3 shows that differences between the binding surfaces are located at the N-terminal part, at the threonine binding pocket, around the pTyr-281 location and inside the valine-284 cavity. The central pTyr-284 binding pocket seems to be different on its right side and upper rim as seen in the orientation of Figure 3 where it flanks the threonine cavity. The first one is caused by different conformations of the side chains of the glutamic acids 34 (Eat-2) and 35 (Sap), and the latter one is due to surface features that correspond to residues (Lys-12 in Eat-2 and Arg-13 in Sap) that do not show any strong interactions with the ligand. This

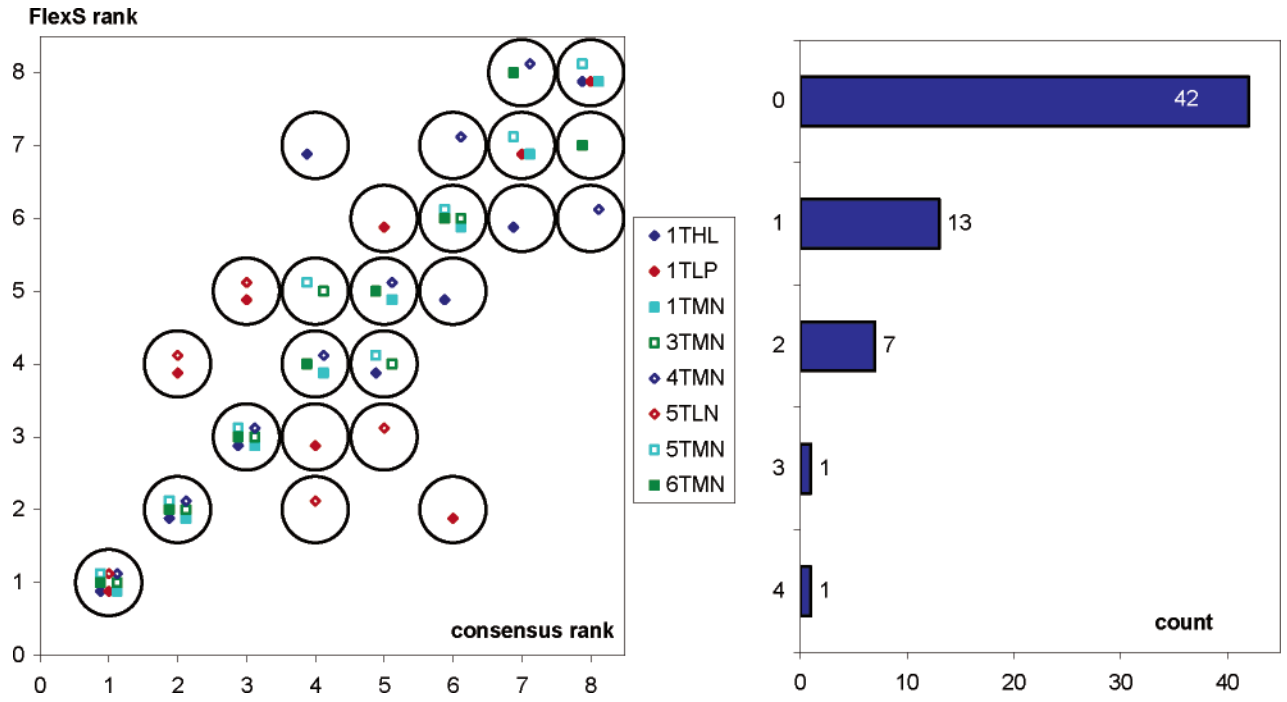

Figure 6. Results of the comparative ranking. (left) Mapping between the SURFCOMP consensus ranking and the FlexS ranks of all comparative ranking experiments. Each circle represents a distinct mapping between the two rankings that occurs at least once in the calculations. All correct matches appear in the diagonal of the graph. (right) A histogram of the mismatches (0 indicates a correct match).

leaves us with the differences in the threonine binding pocket and the valine cavity, which cover two of the three structural motifs that seem to be responsible for the recognition of the ligand.

The threonine cavity in Eat-2 is wider than but not as deep as the corresponding section on the Sap surface. Furthermore, the entrance to the cavity from below (in Figure 3) is steeper in Sap than in Eat-2.

The residues that form this cavity in Sap and Eat-2 are very similar, and the relative conformations of the residues in each pocket are also highly conserved (Figure 4). But the surfaces are nevertheless dissimilar at several points that are related to the differences in the amino acid structure. As mentioned above, a significant dissimilarity is caused by the patches that are placed around Arg-13 in Sap and Lys-12 in Eat-2. The most important difference, however, is located right at the center of the cavities where a glycine residue in Sap (Gly-16) is exchanged by a cysteine residue in Eat-2 (Cys-15). The missing side chain causes the pocket of Sap to extend deeper into the protein than in Eat-2, where the side chain of the cysteine is blocking the way. In the crystal structure of Sap, the larger cavity is occupied by two water molecules that seem to be tightly bound to the protein as judged by their low B-factors of 15.25 $Å^2$ for the inner and 17.89 $Å^2$ for the outer water, respectively. In Eat-2 the corresponding pocket holds only one molecule of water that is much more mobile (B-factor of 39.07 $Å^2$).

The situation around the second major difference, the valine pocket, is even more interesting because the differences there are larger and more complex. The finger that encloses the cavity from the right side is much more negatively charged in Eat-2 than in Sap, and the shape of that region is also quite divergent (see Figure 3 and Figure 5). The most important differences are found at the bottom of the pocket. There Sap has two little extra cavities that are separated by a small ridge. On Eat-2 the bottom of the valine pocket is rather flat and has no pronounced hole or ridge.

Similarly to the situation of the threonine cavity, some of the residues that line the pocket binding the Val-284 residue of the pSlam ligand are conserved between Sap and Eat-2 (see Figure 5). In contrast to the threonine cavity, the shapes of these valine cavities differ not only at the center but also at the peripheral areas. However, the most interesting part is again the central pocket. In the middle of the valine cavity Eat-2 has only a single shallow hole that is enclosed by a leucine (Leu-93) and an isoleucine (Ile-65) residue. Sap has two deeper but smaller cavities at the same position that share a common entrance similar to the entrance of the single Eat-2 hole. These two cavities are surrounded by two phenylalanine residues (Phe-77 and Phe-87), one alanine (Ala-66), and one leucine (Leu-43). In contrast to the threonine binding site, the valine pockets in Sap and Eat-2 do not contain bound water molecules in the crystal, which is probably due to the hydrophobic character of the residues involved. To illustrate how different the depth of the two pockets actually is, consider that the bottom of the cavity in Sap is formed by Leu-43. This residue corresponds to Leu-42 in Sap which is buried deep inside the protein and does not have any contact to solvent molecules.

It is noteworthy that the corresponding surface patches of the proteins, which are in contact with variable parts of the consensus sequence motif T/S-x-pY/Y-x-x-V/I, are highly conserved. Neither the region beneath the Ile-280 nor the patch close to Ala-282 and Glu-282 show any significant differences, although they do not have a lot of significant surface features. These findings support the consensus motif from the perspective of the surfaces because a flat and featureless region does not provide many anchor points, which are necessary for discrimination.

CONCLUSION

Molecular surfaces describe the interface of a molecule between its atoms and the environment. By analyzing the patterns of physicochemical properties mapped onto molecular surfaces one can gain insight into the details of molecular interactions, which in the case of biological macromolecules complements the information obtainable from sequence or 3D structure alignments. Although the comparison of protein surfaces can be very CPU-intensive in general, we have shown that the detection of surface similarities between proteins can be performed as efficiently as for small molecules if the search is restricted to the functional sites of the proteins. Such a selection could be implemented in our surface similarity search program SURFCOMP.

The comparison of Sap and Eat-2 showed that it can be rewarding to look for dissimilarities between the surfaces of active sites with similar functions in order to find ways to selectively influence one target molecule over the other, which is often a very important problem in rational drug design. With a sequence or structural alignment only the differences in the amino acid sequences or the atomic positions can be detected. Molecular surface comparison can make the influence of these variations on the interface between the receptor and the ligand visible. One can then focus on those dissimilarities in the sequences that are responsible for the significant differences detected between the binding site surfaces.

The present implementation of SURFCOMP can only compare static surfaces that were generated from fixed molecular conformations. Conformational flexibility could possibly be taken into account by comparing surfaces derived from a set of relevant low-energy conformers. However, for the comparisons discussed here, this approach was not necessary because the presence of the bound ligand strongly limits the conformational flexibility of the binding site.

Although we have analyzed only one pair of related proteins, it is conceivable that many such systems exist where an investigation of molecular surfaces can help to find promising drug candidates. In principle, it should even be possible to carry out a large-scale comparison of all protein structures in a non-redundant subset of the PDB, thus enabling the discovery of new, unexpected functional similarities and the identification of novel drug targets.

Availability of Programs. Readers wishing to use the SURFCOMP package can obtain the source code and installation packages for several Linux distributions at http://teachme.tuwien.ac.at/surfcomp.

APPENDIX: EVALUATION OF THE COMPARATIVE SCORING ALGORITHM

To test our scoring approach, we compared the results with the ranking produced by the program FlexS[37] for a flexible superposition of eight thermolysin inhibitors. We used FlexS as a reference method because it incorporates a volumetric technique to generate the flexible alignments, which is comparable to surface or shape matching. For each structure in the thermolysin data set, a flexible alignment with all the other structures in the set was generated, and the conformations that produced the best alignment with the current template structure were taken to form the data for the surface similarity searches. Solvent-excluded surfaces were generated for all structures in that set and compared to the surface of the template molecule with SURFCOMP. The resulting tables of alternative alignments were combined into one table for each template molecule and ranked by the consensus scoring approach. From this scoring a ranking of the molecules of the data set was assembled based on the first occurrence of the best alignment of each molecule.

In Figure 6, the results of all 8 comparative ranking experiments are summarized. Overall the agreement between the ranking based on FlexS' total score[38] and the consensus scoring of the SURFCOMP program is very good. More than 65% of the structures were assigned the same rank by both methods and another 20% showed only a ranking difference of 1. Furthermore, many of the mismatches are still in a correct relative order. The larger differences were mainly caused by the comparative scoring experiments against 1TLP and 5TLN. 5TLN does not have any significant surface similarities with any of the other molecules,[1] which makes a reasonable ranking based on that criterion most unlikely.

ACKNOWLEDGMENT

We thank Dr. Tamás Schweighoffer for his support in the preparation of the project and Dr. Siegfried Höfinger for revising the manuscript.

REFERENCES AND NOTES

(1) Hofbauer, C.; Lohninger, H.; Aszódi, A. SURFCOMP: a novel graph-based approach to molecular surface comparison. *J. Chem. Inf. Comput. Sci* **2004**, *44*, 837−847.
(2) Botfield, M. C.; Green, J. SH2 and SH3 domains: choreographers of multiple signaling pathways. *Annu. Rep. Med. Chem.* **1995**, *30*, 227−237.
(3) Latour, S.; Roncagalli, R.; Chen, R.; Bakinowski, M.; Shi, X.; Schwartzberg, P. L.; Davidson, D.; Veillette, A. Binding of SAP SH2 domain to FynT SH3 domain reveals a novel mechanism of receptor signalling in immune regulation. *Nat. Cell Biol.* **2003**, *5*, 149−154.
(4) Chan, B.; Lanyi, A.; Song, H. K.; Griesbach, J.; Simarro-Grande, M.; Poy, F.; Howie, D.; Sumegi, J.; Terhorst, C.; Eck, M. J. SAP couples Fyn to SLAM immune receptors. *Nat. Cell Biol.* **2003**, *5*, 155−160.
(5) Morra, M.; Howie, D.; Grande, M. S.; Sayos, J.; Wang, N.; Wu, C.; Engel, P.; Terhorst, C. X-linked lymphoproliferative disease: a progressive immunodeficiency. *Annu. Rev. Immunol.* **2001**, *19*, 657−682.
(6) Morra, M. Structural basis for the interaction of the free SH2 domain EAT-2 with SLAM receptors in hematopoietic cells. *EMBO J.* **2001**, *20*, 5840.
(7) Li, C.; Iosef, C.; Jia, C. Y.; Han, V. K.; Li, S. S. Dual functional roles for the X-linked lymphoproliferative syndrome gene product SAP/SH2D1A in signaling through the signaling lymphocyte activation molecule (SLAM) family of immune receptors. *J. Biol. Chem.* **2003**, *278*, 3852−3859.
(8) Li, S. C.; Gish, G.; Yang, D.; Coffey, A. J.; Forman-Kay, J. D.; Ernberg, I.; Kay, L. E.; Pawson, T. Novel mode of ligand binding by the SH2 domain of the human XLP disease gene product SAP/SH2D1A. *Curr. Biol.* **1999**, *9*, 1355−1362.
(9) Kleywegt, J. Recognition of spatial motifs in protein structures. *J. Mol. Biol.* **1999**, *285*, 1887.
(10) Kauvar, M. Deciphering cryptic similarities in protein binding sites. *Curr. Opin. Biotechnol.* **1998**, *9*, 390.
(11) Jambon, M. A new bioinformatic approach to detect common 3D sites in protein structures. *Proteins* **2003**, *52*, 137.
(12) Bostick, D. A new topological method to measure protein structure similarity. *Biochem. Biophys. Res. Commun.* **2003**, *304*, 320.
(13) Maggiora, M. Comparing protein structures: a Gaussian-based approach to the three-dimensional structural similarity of proteins. *J. Mol. Graphics* **2001**, *19*, 168.
(14) Lee, B.; Richards, F. M. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379−400.
(15) Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **1983**, *16*, 548−558.
(16) Via, A.; Ferrè, F.; Brannetti, B.; Helmer-Citterich, M. Protein surface similarities: a survey of methods to describe and compare protein surfaces. *Cell. Mol. Life Sci.* **2000**, *57*, 1970−1977.

(17) Lin, S. L.; Nussinov, R.; Fischer, D.; Wolfson, H. J. Molecular surface representations by sparse critical points. *Proteins: Struct. Funct. Genet.* **1994**, *18*, 94−101.
(18) Fischer, D.; Lin, S. L.; Wolfson, H. L.; Nussinov, R. A geometry-based suite of molecular docking processes. *J. Mol. Biol.* **1995**, *248*, 459−477.
(19) Goldman, B. B.; Wipke, W. T. QSD quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock). *Proteins: Struct. Funct. Genet.* **2000**, *38*, 79−94.
(20) Wang, H. Grid-search molecular accessible surface algorithm for solving the protein docking problem. *J. Comput. Chem.* **1991**, *12*, 746−750.
(21) Connolly, M. L. Shape distributions of protein topography. *Biopolymers* **1992**, *32*, 1215−1236.
(22) Chau, P. L.; Dean, P. M. Molecular recoginition: 3D surface structure comparison by gnomic projection. *J. Mol. Graphics* **1987**, *5*, 97−100.
(23) Bladon, P. A rapid method for comparing and matching the spherical parameter surfaces of molecules and other irregular objects. *J. Mol. Graphics* **1989**, *7*, 130−137.
(24) Blaney, F.; Flinn, P.; Phippen, R.; Wyatt, R. Molecular surface comparison: application to drug design. *J. Mol. Graphics* **1993**, *11*, 98−105.
(25) Leicester, S. E.; Finney, J. L.; Bywater, R. P. Description of molecular surface shape using Fourier descriptors. *J. Mol. Graphics* **1988**, *6*, 104−8, 100.
(26) Ritchie, D. W. A. Protein docking using spherical polar Fourier correlations. *Proteins*: *Struct. Funct. Genet.* **2000**, *39*, 178−194.
(27) Cosgrove, D.; Bayada, D.; Johnson, A. A novel method of aligning molecules by local surface shape similarity. *J. Comput.- Aided Mol. Des.* **2000**, *14*, 573−591.
(28) Goldman, B. B.; Wipke, W. T. Quadratic shape descriptors. 1. Rapid superposition of dissimilar molecules using geometrically invariant surface descriptors. *J. Chem. Inf. Comput. Sci* **2000**, *40*, 644−658.
(29) Exner, T. E.; Keil, M.; Brickmann, J. Pattern recognition strategies for molecular surfaces. I. Pattern generation using fuzzy set theory. *J. Comput. Chem.* **2002**, *23*, 1176−1187.
(30) Martin, Y. C.; Bures, M. G.; Danaher, E. A.; DeLazzer, J.; Lico, I.; Pavlik, P. A. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 83−102.
(31) Zhang, D.; Herbert, M. Harmonic maps and their applications in surface matching. Presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '99) 1999,
(32) Wang, R.; Wang, S. How does consensus scoring work for virtual library screening? An idealized computer experiment. *J. Chem. Inf. Comput. Sci* **2001**, *41*, 1422.
(33) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, *42*, 5100−5109.
(34) Brickmann, J.; Goetze, T.;, Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C. D. Interactive visualization of molecular scenarios with MOLCAD/SYBYL. *Data Visualization Mol. Sci.* **1995**.
(35) *SYBYL 6.9*; Tripos Inc.: St. Louis, MO, 2003.
(36) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Mertz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(37) Lemmen, C.; Lengauer, T. Time-efficient flexible superposition of medium-sized molecules. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 357−368.
(38) Lemmen, C.; Lengauer, T.; Klebe, G. FLEXS: a method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502.

CI0497049