Computing and Displaying Intermolecular Negative Volume for Docking

Chang Ha Lee and Amitabh Varshney

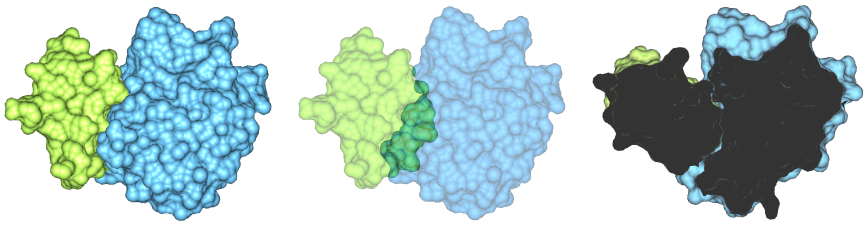
Department of Computer Science and UMIACS, University of Maryland, College Park, MD 20742, USA, {chlee, varshney}@cs.umd.edu

Summary. Protein docking is a Grand Challenge problem that is crucial to our understanding of biochemical processes. Several protein docking algorithms use shape complementarity as the primary criterion for evaluating the docking candidates. The intermolecular volume and area between docked molecules is useful as a measure of the shape complementarity. In this paper we discuss an algorithm for interactively computing intermolecular negative volume and the area of docking site using graphics hardware. We also present the design considerations for building an interactive 3D visualization tool for visualizing intermolecular negative volumes.

1 Introduction

Several drug development processes have so far begun with large-scale random screening of candidate inhibitors. These initial discoveries are improved through well-defined approaches to find new drugs. As molecular structure determination techniques and computational methods progress, protein docking methods using structure-based molecular complementarity have become an important substitute for random screening in the drug design process [Kun92].

Among many factors involved in protein-protein interactions such as electrostatics, hydrophobicity, and hydrogen bonding, shape complementarity is of major importance for protein docking. Purely geometric approach can restrict the time-consuming calculations of interaction energy to be performed only for those cases that have a good geometric fit. Geometric methods can also be used as foundations for more complete approaches considering chemical and energetic characteristics [Con86]. A complete search of all possible geometric fits of two flexible molecules takes too much time because of the extremely large degrees of freedom. Therefore, molecules have been often assumed as rigid bodies. Even with the rigid body assumption, finding accurate shape complementarities remains a challenging problem. Most existing methods provide a list of candidates sorted by complementarity criteria and the final decision by human is needed. Therefore, an interactive tool for visualizing the shape complementarity would be useful.



(a) Solid Surfaces (b) Translucent Surfaces (c) Cross Sections

Fig. 1. Traditional Complementarity Visualization Methods

There are many methods for visualizing the steric fit between molecules. These include visualization using solid solvent-accessible smooth molecular surfaces, translucent molecular surfaces, and cross-sections of molecular surfaces (See Figure 1). The solid molecular surface representation is unsuitable for complementarity visualization because the interface between molecules is difficult to observe due to occlusions from the solid surfaces. Translucent molecular surfaces allow the visualization of the interface between molecules. However, visual interference from other parts of the molecules prevents a clear visualization of the intermolecular interface. The cross-section method, also called the Z-clip method in graphics, visualizes the molecular interface by displaying cross sections of the molecules at varying depths from the viewer. Although the interface can be visualized clearly using two-dimensional cross sections, it is difficult to construct a mental model of the three-dimensional spatial structure of the interface. Therefore, in addition to the visualization of molecular surfaces, we need new and more informative methods for the visualization of the interface between molecules.

In this paper we present a method that computes *the negative volume* between molecules to visualize their interface. Our method leverages the recent advances in the 3D graphics hardware to achieve interactive rates of performance. Using this method scientists can interactively study various possible docking conformations and visualize the quality of the steric fit.

The remaining paper is organized as follows. In Section 2, we give an overview of the previous work. The concept of intermolecular negative volume and the description of the algorithm for computing intermolecular negative volume are given in Sections 3, 4, and 5. The algorithm for computing the area and volume of the docking site is described in Section 6. In Section 7, we describe our interactive 3D visual tool to assist protein docking. We conclude this paper and discuss future work in Section 8.

2 Previous Work

The early research on drug design focused on geometric shape complementarity. Connolly [Con86] has proposed a protein docking algorithm based on geometric shape complementarity. He defines a molecule's *shape function* parameterized by scale R, at a surface point p as the volume of the molecule that lies inside a sphere of radius R centered at p. He defines the *knobs* as the local minima in the shape function and *holes* as the local maxima. His method finds the transform to dock two proteins by finding matches between quartets of knobs and holes on the two proteins. Katchalski-Katzir *et al.* [KKSE+92] have proposed a Fourier-transform-based geometric recognition algorithm for molecular surface complementarity.

Edelsbrunner *et al.* [EFL98, ELW98] have defined a *pocket* as a region in the complement if it can be reached only via narrow pathways. They have proposed an algorithm to compute pockets in a protein. They have also proposed a method to measure properties of surface pockets such as volume and area. They have applied their method to discover the binding sites between molecules.

Word *et al.* [WLL+99] have proposed a method to measure the goodness-of-fit for molecular interfaces. They have described small-probe contact dots for measuring and visualizing the atomic contacts inside or between molecules. Their algorithm is similar to the Connolly's algorithm [Con83] for computing solvent-accessible molecular surfaces, in that a probe sphere is rolled over the spherical model of a molecule. The difference is that they leave a dot when the probe touches atoms of two molecules. Quantitative measure for goodness-of-fit is defined by the volume measured by the length between dots.

Wintner and Moallemi [WM00] have proposed the concept of *Quantized Surface Complementarity Diversity*, QSCD, for measuring complementarity between molecules. Diversity is defined as the measure of the difference, or similarity, between small molecules. They have defined a set of theoretical target surfaces that approximate all possible binding pockets with a volume limited by a predefined threshold. Each target surface is formed by cubic units carved out of the surface. These cubic units represent negative space that a potential ligand could occupy. To measure the complementarity of a molecule, the molecule is also quantized into a set of cubic units, and the quantized cubes are compared with the target surfaces. In this paper we discuss a shape complementarity definition based on the ratio of the negative volume to area of the interface between two molecules.

Several researchers have worked on analyzing and classifying molecular interfaces and interactions. Kuntz [Kun92] has proposed strategies for drug design based on the structure of molecules. He finds possible docking sites by locating the grooves on the surface and creating their negative images by using spheres. Then he matches the ligand and the receptor by placing the ligand into the site using the isomorphic subgraph matching algorithms. Jones and Thornton [JT96] have analyzed protein complexes for better understanding of

the principles of protein–protein interactions. They have defined the protein–protein interface based on the changes in the solvent-accessible surfaces when going from a monomeric to a dimeric state. They have examined structural properties of protein–protein interfaces as well as the biochemical properties. Specifically, they have measured the complementarity between surfaces using the *gap index*. The gap index is defined as the gap volume between molecules divided by the interface area. They calculate the gap volume using a method by Laskowski [Las95]. Laskowski defines the *gap sphere* as the largest sphere which can be placed between two atoms from each molecule without penetrating either of the molecular surfaces. He computes the gap volume by adding the volumes of all gap spheres. This is a good estimate for the intermolecular volume, however: 1) gap spheres might overlap resulting in possible overestimation of the gap volume, 2) the gap spheres might not cover the entire intermolecular volume resulting in possible underestimation of the gap volume, 3) isotropic spheres might not be adequate to measure an anisotropic intermolecular region, and 4) this method could take quadratic time if it considers all possible pairs of atoms. Our approach computes the intermolecular volume accurately to any desired level of precision and runs in linear time.

Nadassy *et al.* [NTOJ+01] have measured how compactly atoms are packed in molecular interfaces compared to the internal spaces by computing the atomic volumes in double-stranded DNA and in protein–DNA interfaces. They have also used two measures for assessing packing in interfaces of macromolecular complexes: (a) the gap volume index as defined above, and (b) the *shape correlation index* by Lawrence and Colman [LC93]. The shape correlation index is derived from the distance between two points on the surfaces of interacting molecules and the angle between normal vectors of these points. This is an intuitive estimate for the shape complementarity and could be implemented to run in linear time with a good data structure for identifying nearest points. However, their metric is insensitive to the area of the intermolecular interface.

Varshney *et al.* [VJR+95] have proposed an analytic approach for computing and visualizing molecular interfaces in linear time. However, their primary goal is to visualize the interface surfaces, not the intermolecular negative volume. Also, they do not use the graphics hardware to accelerate the computation of the interface surfaces. To the best of our knowledge, no previous work has been done in interactively computing and visualizing intermolecular negative volume using linearly scalable algorithms with user-specifiable accuracy.

Domik and Fels [DF96] have developed a visual tool for studying molecular docking. Their system enables users to visually determine prospective binding sites by visualizing collision detections. Recently Olson *et al.* [Ols03] have developed an augmented reality tool to study molecules. As the user rotates and translates a 3D printed replica of a molecule, their system tracks the molecule's movements and mimics them in conjunction with a virtual molecule on the screen. This provides the users a compelling sense of the shape of a molecule

and how it relates to other molecules. This can be valuably used in shape complementarity studies. Another powerful tool for protein visualization has been recently developed by Kreylos *et al.* [KMH+03]. Their tool allows the users to design proteins ab-initio using primary, secondary, and tertiary structures. Their tool also allows inverse kinematics and interactive visualization of proteins using a variety of motif visualizations as well as Ramachandran plots and intra-molecular collisions.

Interactivity is crucial for task-completion in 3D visualization applications. This has been proven by several researchers including Smets and Overbeeke [SO95] and Hawkes *et al.* [HRS95]. To achieve interactivity while studying shape complementarity for a pair of molecules, we have developed an algorithm for computing and visualizing the intermolecular negative volume and the area of docking site using graphics hardware. Our work complements previous work on protein visualization in that it provides a new way to interactively visualize molecular interfaces.

3 Defining the Intermolecular Negative Volume

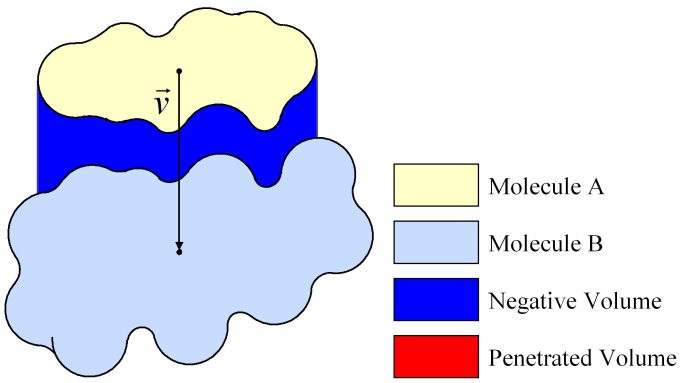
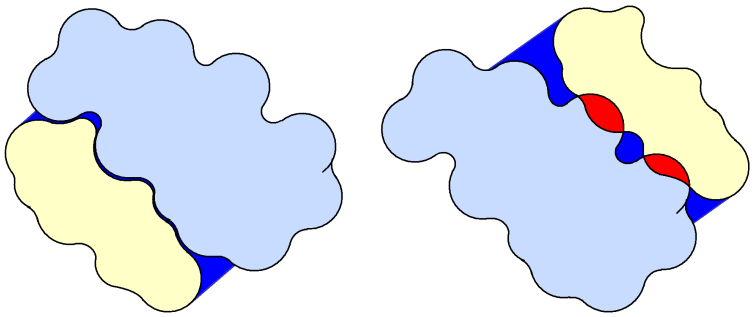


Fig. 2. Computing Intermolecular Negative Volume

The smooth molecular surface, first proposed by Richards [Ric77], is defined as the surface which an external probe sphere touches as it is rolled over the spherical atoms of a molecule. This representation is useful for studying interaction between molecules as it provides a smooth surface approximation to a molecule while retaining its most important shape features. Specifically, this surface representation is useful for studying shape complementarity since the two molecular surfaces are approximately coincident in the interfacial region [Con86]. We use the smooth solvent-accessible molecular surface to represent a molecule.



(a) Intermolecular Volume of Docked Molecules

(b) Intermolecular Volume of Penetrating Molecules

Fig. 3. Intermolecular Negative Volume

Given two molecules A and B, let the set of points on the molecular surfaces be represented by A_o and B_o respectively. Here the subscript o denotes that the points in these are defined in the object (world) coordinate system. Further, let the cardinalities of A_o and B_o be given by n and m, respectively. The centers of A_o and B_o are defined as the average of their respective surface points: $\mathbf{c}_A = \frac{1}{n}\sum_{j=0}^{n} \mathbf{x}_{oj}$ and $\mathbf{c}_B = \frac{1}{m}\sum_{k=0}^{m} \mathbf{x}_{ok}$, where $\mathbf{x}_{oj} \in A_o, \mathbf{x}_{ok} \in B_o$.

We define *the aligning direction* $\mathbf{d}$ between molecules A and B as $(\mathbf{c}_B - \mathbf{c}_A)/|\mathbf{c}_B - \mathbf{c}_A|$. The aligning direction is the unit vector from the center of A to the center of B. We define the aligning direction this way only as a heuristic. Our system can accept user-defined aligning directions as well. We define two mutually orthogonal vectors $\mathbf{u}$, $\mathbf{v}$, perpendicular to the vector (d) to construct the *interface coordinate system*. In this coordinate system x axis is considered to be along $\mathbf{u}$, y along $\mathbf{v}$, and z along (d). Now consider an axis-aligned bounding box in the interface coordinate system that contains both molecules. We assume the origin lies at the center of that face of the bounding box which is parallel to the x-y plane and below the molecule A as shown in Figure 4. This assumption makes all z values in the interface coordinate system positive. We shall use the subscript i to denote the interface coordinate system. Let the matrix to transform a point from the object coordinate system to the interface coordinate system be given by M_i. The point sets in the interface coordinate system A_i and B_i can be defined as $A_i = \{\mathbf{x}_i | \mathbf{x}_i = M_i\mathbf{x}_o,\ \mathbf{x}_o \in A_o\}$, and $B_i = \{\mathbf{x}_i | \mathbf{x}_i = M_i\mathbf{x}_o,\ \mathbf{x}_o \in B_o\}$.

Let $\mathcal{P}(A_i, \mathbf{d})$ and $\mathcal{P}(B_i, \mathbf{d})$ be the parallel projections of A_i and B_i onto the x-y plane of the interface coordinate system. Then, we define the *intersection of projected regions*, $P(A_i, B_i, \mathbf{d}) = \mathcal{P}(A_i, \mathbf{d}) \cap \mathcal{P}(B_i, \mathbf{d})$. This region is shown in Figure 4 in dark green color between the two molecules.

The docking region of a molecule is the region of the molecular surface which borders the intermolecular negative volume. We define the *docking region of* A as $R_A = \{(x_i, y_i, z_i) | (x_i, y_i) \in P(A_i, B_i, \mathbf{d}) \ \wedge \ z_i = \max\{z|$

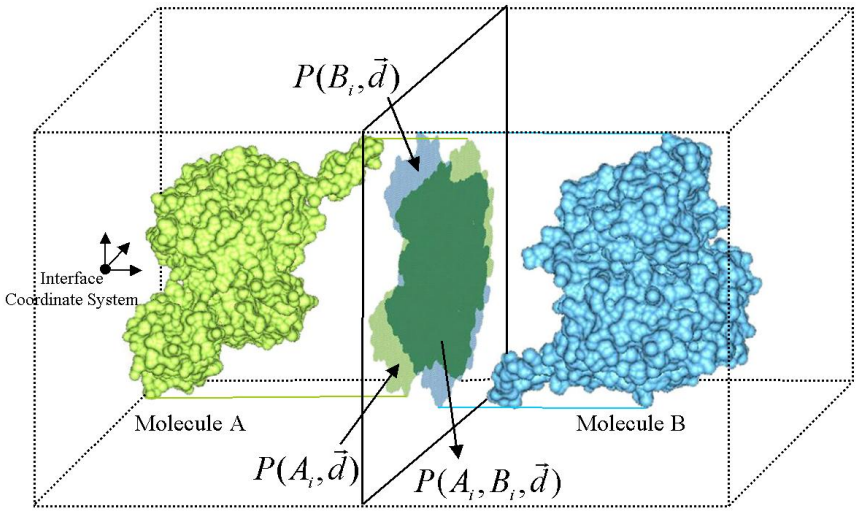


Fig. 4. The Interface Coordinate System and the Molecular Projection Regions

$(x_i, y_i, z) \in A_i\}$. Similarly, the *docking region of* B is defined as $R_B = \{(x_i, y_i, z_i) | (x_i, y_i) \in P(A_i, B_i, \mathbf{d}) \wedge z_i = \min\{z | (x_i, y_i, z) \in B_i\}$. The intermolecular negative volume between R_A and R_B is defined in the interface coordinate system as $V_i = \{(x_i, y_i, z_i) | (x_i, y_i) \in P(A_i, B_i, \mathbf{d}) \wedge z_i \in [R_A(x_i, y_i), R_B(x_i, y_i)]\}$, where $R_A(x_i, y_i) = \{z_i |\ (x_i, y_i, z_i) \in R_A\}$ and $R_B(x_i, y_i) = \{z_i |\ (x_i, y_i, z_i) \in R_B\}$.

The *intermolecular negative volume* between molecules A and B is transformed back to the object coordinate system as $V(A, B) = \{\mathbf{x_o} | \mathbf{x_o} = M_i^{-1}\mathbf{x}_i,\ \mathbf{x}_i \in V_i\}$. Figure 2 shows the intermolecular negative volume in blue.

4 Computing the Intermolecular Negative Volume

We compute the intermolecular negative volume using the graphics hardware. The graphics hardware depth-testing functionality can compute the distance from the viewing plane to the nearest (or the farthest) surface for each pixel in the viewing plane. For instance, in OpenGL the depth-testing option to select the surface closest to the viewing plane is *GL_LESS*, and the option to select the farthest surface is *GL_GREATER*. We can also change the level of detail of the interface by simply changing the resolution of the viewing plane.

The algorithm for computing intermolecular negative volume using depth buffer is as follows. First, we set the viewing direction as the aligning direction $\mathbf{d}$ and draw molecule A with the depth-testing option to select the farthest depth coordinate. The depth buffer now contains the distance from the viewing plane to the farthest surface of A (the surface that defines one side of the

intermolecular volume). The depth buffer is read-back and saved as D_A and the buffer is reset. Second, with the viewing direction still $\mathbf{d}$ we draw molecule B with the depth-testing option to select the nearest depth coordinate. The depth buffer now contains the distance from the viewing plane to the nearest surface of B (which is the surface that defines the second side of the intermolecular volume). The depth buffer is again read back and saved this time as D_B.

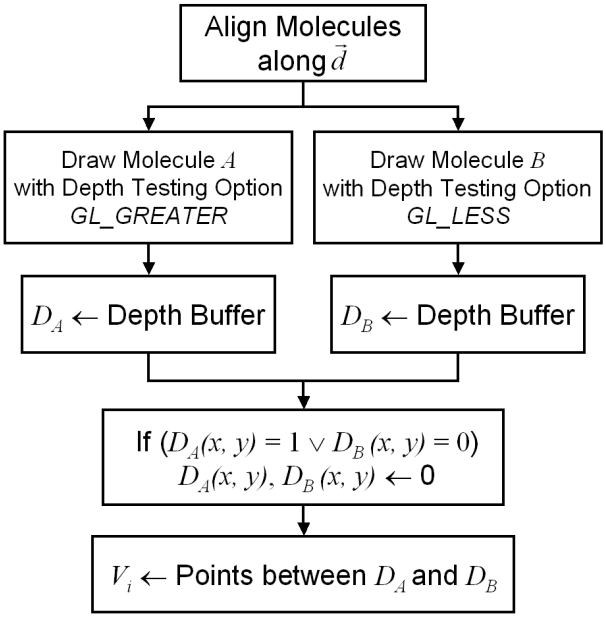


Fig. 5. Algorithm for Computing Intermolecular Negative Volume

In D_A, the region that lies outside the molecule has the largest possible value (1 in OpenGL), and in D_B, the region outside the molecule has the smallest possible value (0 in OpenGL). For each (x, y) value, if $D_A(x, y)$ is equal to 1 or $D_B(x, y)$ is equal to 0, we set both $D_A(x, y)$ and $D_B(x, y)$ to 0. This gives us the intersection of projected regions $P(A_i, B_i, \mathbf{d})$ and sets the remaining region to 0 in D_A and D_B. Now, D_A and D_B store R_A and R_B, respectively, which are the docking regions of A and B as defined in the previous section. Therefore, the volume between D_A and D_B is the intermolecular negative volume between A and B. The overview of our algorithm is shown in Figure 5. To visualize the intermolecular negative volume, we build a triangle mesh using D_A and D_B. The method for triangulating the intermolecular negative volume is described in the next section.

5 Visualizing the Intermolecular Negative Volume

The intermolecular negative volume is the set of voxels between two 2D pixelated depth buffers, D_A and D_B. An obvious choice to extract an isosurface from volume data is the Marching Cubes algorithm [LC87]. However, we have additional information in this case that we can use. Since the non-zero regions of D_A and D_B are the same, we can reduce the marching cubes algorithm to its two-dimensional analog, the *marching squares* algorithm. For each vertex in the triangle mesh, we produce x and y coordinates using the marching squares algorithm and get the z coordinate from D_A and D_B.

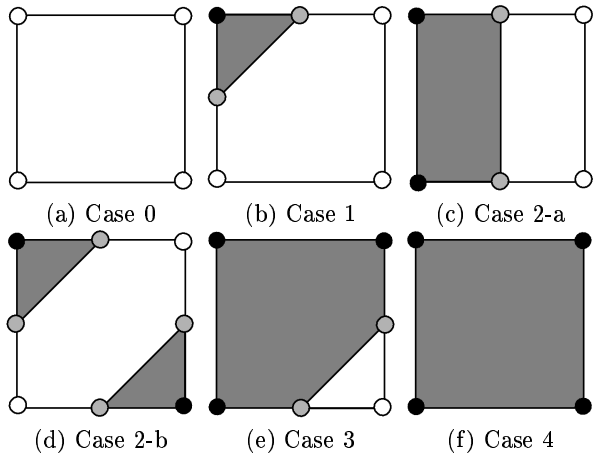
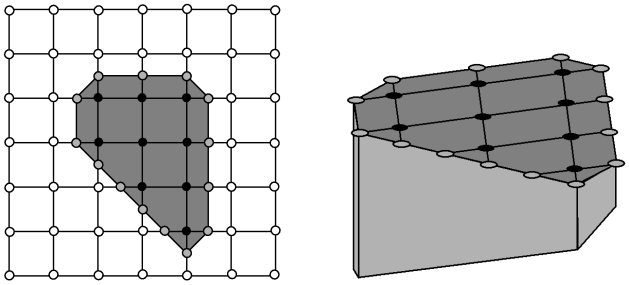


Fig. 6. Cases for building Triangle Meshes : White dots are zero points, black dots are non-zero points, and gray dots are middle points added by the algorithm. Shaded regions show the polygonal mesh.

There are six cases for the squares according to the values of four corner points. In Figure 6, the white dots are points with zero values, and the black dots are points with non-zero values. For each square (pixel) of D_A, we repeat the following process. We create a gray point in the middle of an edge that connects a zero (white) point to a non-zero (black) point. The black and gray points are added to the vertex list of the resulting triangle mesh. The z coordinate of a black point is the value of D_A at the point, and the z coordinate of the gray point is same as that of the adjacent black point. We then generate triangles connecting black and gray dots, which are the shaded regions in Figure 6. These triangles are part of the docking region of A. Similarly, we can produce triangles with same x and y coordinates and the z values from D_B, and these triangles form the docking region of B.

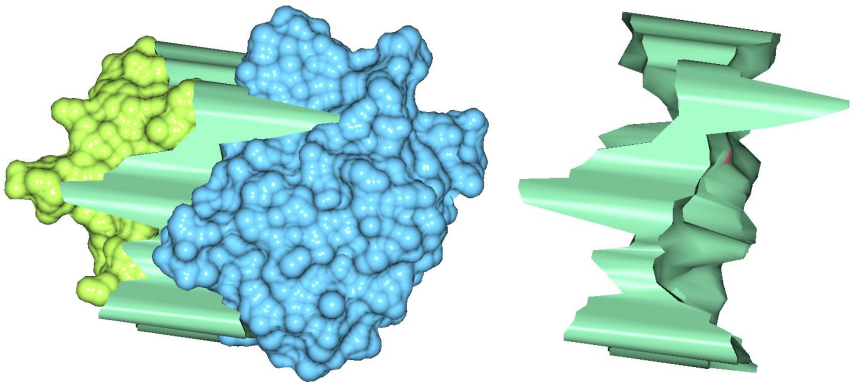
Next we connect the boundaries of the two surfaces representing docking regions of A and B. We define a *gray edge* as the edge formed by two adja-



(a) Marching Squares Algorithm (b) 3D Intermolecular Volume

Fig. 7. A Mesh Construction Using Marching Squares Algorithm

cent gray points. The boundary of the docking region consists of gray edges. Therefore, we create two triangles to connect two gray edges with same x and y coordinates and different z coordinates from D_A and D_B. Figure 7 shows the construction of the mesh. Figure 8 shows the mesh of the intermolecular negative volume between the Proteinase and its Inhibitor in the Protein Data Bank complex 4SGB.



(a) Visualization with Solid Molecules (b) Intermolecular Volume Only

Fig. 8. Interactive Visualization of the Intermolecular Negative Volume between Proteinase and Inhibitor (4SGB)

6 Modifications for More Accurate Computation

6.1 Intermolecular Negative Volume with a Threshold

The algorithm for computing intermolecular negative volume described in Section 4 might not produce desirable results when the docking site is relatively small. The actual docking site might be smaller than the intersection of projected region as shown in Figure 9. This problem causes the thick borders in the intermolecular negative volume that you can see in Figure 8.

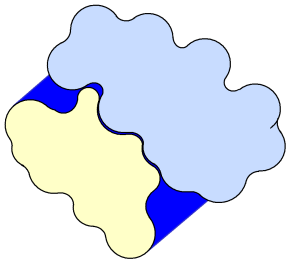
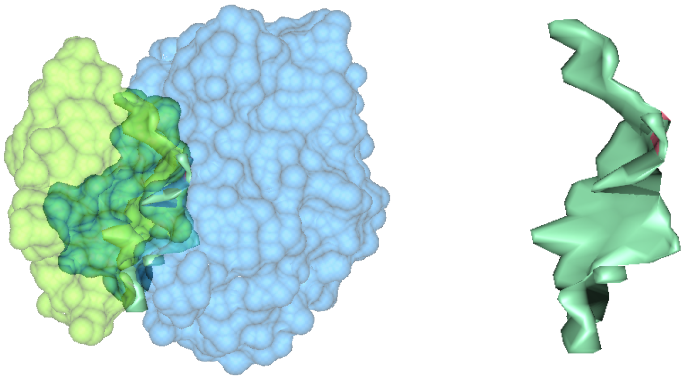


Fig. 9. Intermolecular Negative Volume with a Small Docking Site

We have extended our algorithm to trim the thick borders that do not, in general, define the intermolecular volume. We exclude the regions where the distance between surfaces is greater than a certain threshold ϵ when we compute the intersection of projected regions. Specifically, we compute the difference of z values $Dist(x, y) = D_A(x, y) - D_B(x, y)$ for each (x, y) in the depth buffers D_A and D_B. Only if the absolute value of the difference is less than a threshold ϵ, $|Dist(x, y)| < \epsilon$, we add (x, y) to the docking site. The rest of algorithm is same as described in Section 4. Figure 10 shows the visualization of the intermolecular negative volume modified from Figure 8. We have currently set the ϵ to be the diameter of a water molecule, 2.8 Å to make the intermolecular volume solvent inaccessible.

6.2 Computing the Area of the Docking Site

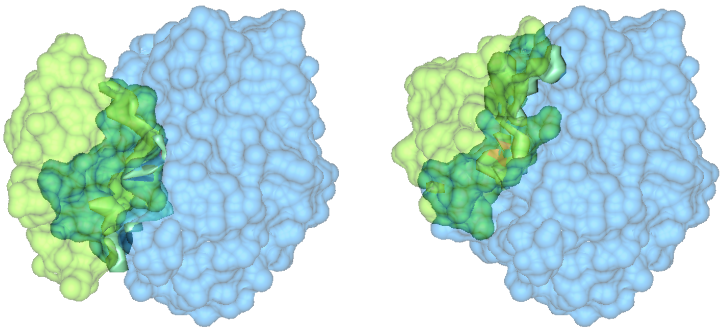
The intermolecular negative volume is often not enough to characterize the docking site. For instance, in Figure 11, the correct fit (a) has larger area-volume ratio even though the incorrect fit (b) has smaller volume. We observe that the ratio of the area of the interface to the intermolecular volume is a much better heuristic than just using the intermolecular volume. We compute the area of the molecular interface by simply adding up the areas of the mesh triangles in the intermolecular volume that are defined by the surfaces of one of the two molecules A and B. The fit between molecules is considered good



(a) Visualization with Molecular Surface (b) Intermolecular Volume Only

Fig. 10. The Intermolecular Negative Volume between Proteinase and Inhibitor (4SGB) with a threshold $\epsilon = 2.8$ Å

when the volume between them is small *and* the docking site area is large. We propose the ratio of the area of the interface and intermolecular volume as a criterion for characterizing the goodness-of-fit for protein docking as well as rational drug design. The larger this ratio, the better the fit between molecules.



(a) Correct Fit:
Volume = 195.54
Docking Site Area = 283.20
Area/Volume = 1.45

(b) Incorrect Fit:
Volume = 179.73
Docking Site Area = 148.46
Area/Volume = 0.83

Fig. 11. Intermolecular Negative Volume between Proteinase and Inhibitor (4SGB)

Table 1. Intermolecular Negative Volume and Interface Area

Complex Names	Volume ($Å^3$)	Interface Area ($Å^2$)	Ratio (Area/Volume)
Protease/Inhibitor	538.60	443.46 / 510.64	0.89
alpha-chymotrypsinogen/Trypsin inhibitor	539.75	602.76 /651.60	1.16
beta-trypsinogen/Trypsin inhibitor	521.61	479.96 / 493.38	0.93
Subtilisin Novo/Chymotrypsin inhibitor 2	666.67	605.26 / 679.02	0.96
Subtilisin BPN/Subtilisin inhibitor	553.15	536.00 / 609.72	1.04
Barnase/Barstar	1224.65	1229.19 / 1171.00	0.98
Acetylcholinesterase/Inhibitor	955.03	837.64 / 810.42	0.86
Ribonuclease inhibitor/Ribonuclease A	536.34	406.89 / 401.39	0.75
IgG1 D44.1 Fab Fragment/Lysozyme	603.93	353.73 / 363.46	0.59
IgG1 E8 Fab Fragment/Cytochrome C	892.38	637.04 / 648.74	0.72
Antibody Hulysll Fv/Lysozyme	1674.65	1174.88 / 1140.83	0.69
CDK2 cyclin-dependant kinase 2/Cyclin	1534.82	1512.47 / 1705.81	1.05
Methylamine dehydrogenase/Amicyanin	2017.78	1517.69 / 1503.73	0.75

7 Interactive Visualization

We have developed a 3D visualization tool for visualizing the intermolecular negative volume interactively, which can be used as a complementary tool to existing drug-design systems. Users can manipulate the molecules together or separately while the intermolecular negative volume is computed and rendered for every frame at interactive rates. Our system also provides various visualization options. The intermolecular negative volume can be visualized alone, or together with molecules as the images in this paper show. The molecules can be visualized in solid surfaces or translucent surfaces.

Table 1 shows the intermolecular volumes, the interface area contributions from the two molecules, and the ratio of the average interface area to the intermolecular volume for a number of naturally occurring molecular complexes. Table 2 shows the times for computing intermolecular negative volumes for the same molecular complexes. This includes the time to generate and render the triangle mesh for visualizing the intermolecular volume. We have computed the intermolecular volume at various resolutions: 64^3, 128^3, and 256^3. The computing time of our algorithm is linearly related to the number of atoms and runs at interactive rates as shown in Figure 12. We have obtained these results on a Dell Precision Workstation with 1.5 GHz Pentium 4, 1 GB RAM, and a nVidia GeForce FX 5900 graphics card.

In the study of intermolecular negative volumes, the variation of the distance between the two molecular surfaces is also important in addition to aggregate information such as area and volume. We have encoded the distance information into the color of the intermolecular negative volume as shown in Figure 13. Deep blue shows larger interface distance and light blue shows a smaller interface distance.

Table 2. Timing Information for Various Complexes

Complex Names	Number of Atoms	Time for 64^3 (msec)	Time for 128^3 (msec)	Time for 256^3 (msec)
Protease/Inhibitor	1310/380	20	50	170
alpha-chymotrypsinogen/Trypsin inhibitor	1799/440	30	50	170
beta-trypsinogen/Trypsin inhibitor	1629/454	30	50	170
Subtilisin Novo/Chymotrypsin inhibitor 2	1938/513	30	60	170
Subtilisin BPN/Subtilisin inhibitor	1938/764	30	70	210
Barnase/Barstar	2581/2059	40	70	190
Acetylcholinesterase/Inhibitor	4116/460	40	70	190
Ribonuclease inhibitor/Ribonuclease A	951/3411	40	70	190
IgG1 D44.1 Fab Fragment/Lysozyme	4291/4291	60	90	211
IgG1 E8 Fab Fragment/Cytochrome C	3340/823	40	70	210
Antibody Hulysll Fv/Lysozyme	3488/2002	50	70	191
CDK2 cyclin-dependant kinase 2/Cyclin	4796/4202	71	110	241
Methylamine dehydrogenase/Amicyanin	4280/4281	81	120	261

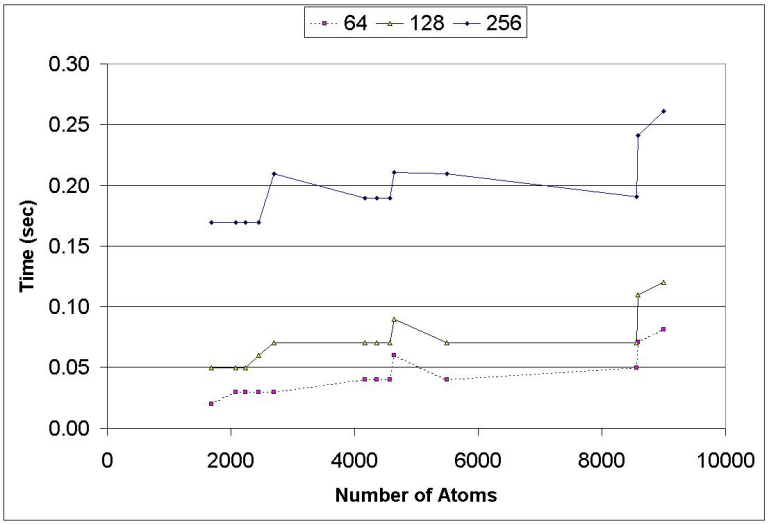
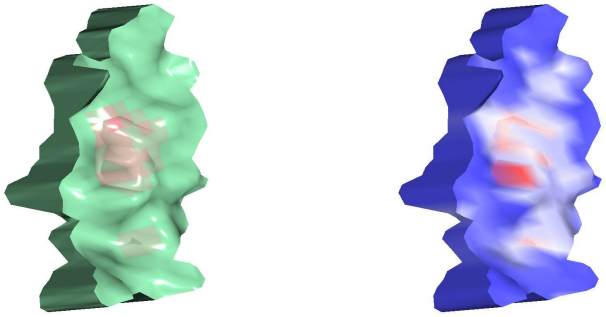


Fig. 12. Time for Computing Intermolecular Negative Volumes

8 Conclusions and Future Work

Our algorithm for computing intermolecular negative volume and the area of the molecular interfaces can be used for computing criteria for shape complementarity. Our 3D visualization tool for visualizing intermolecular negative volume interactively can be used as a complementary tool for existing protein docking and rational drug design systems. The visualization tool can be used to select the best fit among the candidates, and to improve the result by moving the molecules interactively to produce a better fit. Currently, we have only considered the geometric shape complementarity as a way of characterizing the goodness-of-fit between two molecules. Shape complementarity is an



(a) Intermolecular Negative Volume (b) Intermolecular Negative Volume with Color Coding

Fig. 13. Color Coding of Intermolecular Negative Volume between Proteinase and Inhibitor (4SGB). The red color shows overlap between the two molecules of the 4SGB complex.

important criterion, but not the only one. It would be very helpful to include other criteria such as electrostatics, hydrophobicity, hydrogen bonding in developing a comprehensive metric for characterizing good molecular interfaces that can be used in protein docking and rational drug design applications.

Acknowledgements

We would like to thank Sergei Sukharev at the Department of Biology at the University of Maryland, Ron Unger at the Bar-Ilan University in Israel, and John Moult at the Center for Advanced Research in Biotechnology at Rockville, Maryland for valuable discussions and for providing the data sets for testing. This work has been supported in part by the NSF grants: IIS 00-81847, CCF 04-29753, and CNS 04-03313.

References

[Con83] M. L. Connolly. Solvent-accessible surfaces of proteins and nucleic acids. *Science*, 221:709–713, 1983.

[Con86] M. L. Connolly. Shape complementarity at the hemoglobin a1b1 subunit interface. *Biopolymers*, 25:1229–1247, 1986.

[DF96] G. Domik and G. Fels. HotDock: An interactive approach to molecular docking, 1996. http://www.uni-paderborn.de/ lst/HotDock/.

[EFL98] H. Edelsbrunner, M. A. Facello, and J. Liang. On the definition and the construction of pockets in macromolecules. *Discrete Applied Mathematics*, 88(4):83–102, 1998.

[ELW98] H. Edelsbrunner, J. Liang, and C. Woodward. Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. *Protein Science*, 7(9):1884–1897, 1998.

[HRS95] R. Hawkes, S. Rushton, and M. Smyth. Update rates and fidelity in virtual environments. *Virtual Reality: Research, Applications and Design*, 1(2):46–51, 1995.

[JT96] S. Jones and J. M. Thornton. Principles of protein–protein interactions. In *Proc. Natl. Acad. Sci. USA*, volume 93, pages 13–20, 1996.

[KKSE+92] E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A. A. Friesem, C. Aflalo, and I. A. Vakser. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. In *Proceedings of the National Academy of Sciences of United States of America*, volume 89, pages 2195–2199, March 1992.

[KMH+03] O. Kreylos, N. L. Max, B. Hamann, S. N. Crivelli, and E. W. Behel. Interactive protein manipulation. In *Proceedings of the IEEE Visualization*, pages 581–588, Seattle, Washington, October 2003.

[Kun92] I. D. Kuntz. Structure-based strategies for drug design and discovery. *Science*, 257:1078–1082, 1992.

[Las95] R. A. Laskowski. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *Journal of Molecular Graphics*, 13(5):323–330, 1995.

[LC87] W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3d surface construction algorithm. In *Proceedings of SIGGRAPH*, pages 163–169, July 1987.

[LC93] M. C. Lawrence and P. M. Colman. Shape complementarity at protein/protein interfaces. *Journal of Molecular Biology*, 234(4):946–950, 1993.

[NTOJ+01] K. Nadassy, I. Tomas-Oliveira, I. Alberts J., Janin, and S. J. Wodak. Standard atomic volumes in double-stranded DNA and packing in protein–DNA interfaces. *Nucleic Acids Research*, 29(16):3362–3376, 2001.

[Ols03] A. Olson. Tangible interfaces for molecular biology. In *Demos at the IEEE Visualization*, page D12, Seattle, Washington, October 2003.

[Ric77] F. M. Richards. Areas, volumes, packing and protein structures. In *Annual Review of Biophysics and Bioengineering*, volume 6, pages 151–176, 1977.

[SO95] G. J. F. Smets and K. J. Overbeeke. Trade-off between resolution and interactivity in spatial task performance. *IEEE Computer Graphics and Applications*, 15(5):46–51, 1995.

[VJR+95] A. Varshney, F. P. Brooks Jr., D. C. Richardson, W. V. Wright, and D. Manocha. Defining, computing, and visualizing molecular interfaces. In *IEEE Visualization*, pages 36–43, October 1995.

[WLL+99] J. M. Word, S. C. Lovell, T. H. LaBean, H. C. Taylor, M. E. Zalis, B. K. Presley, J. S. Richardson, and D. C. Richardson. Visualizing and quantifying molecular goodness-of-fit: Small-probe contact dots with explicit hydrogen atoms. *Journal of Molecular Biology*, 285(4):1711–1733, 1999.

[WM00] E. A. Wintner and C. C. Moallemi. Quantized surface complementarity diversity (QSCD): A model based on small molecule-target complementarity. *Journal of Medicinal Chemistry*, 43:1993–2006, 2000.

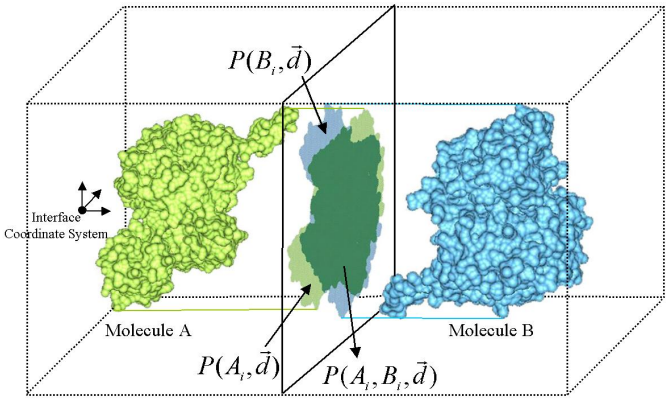
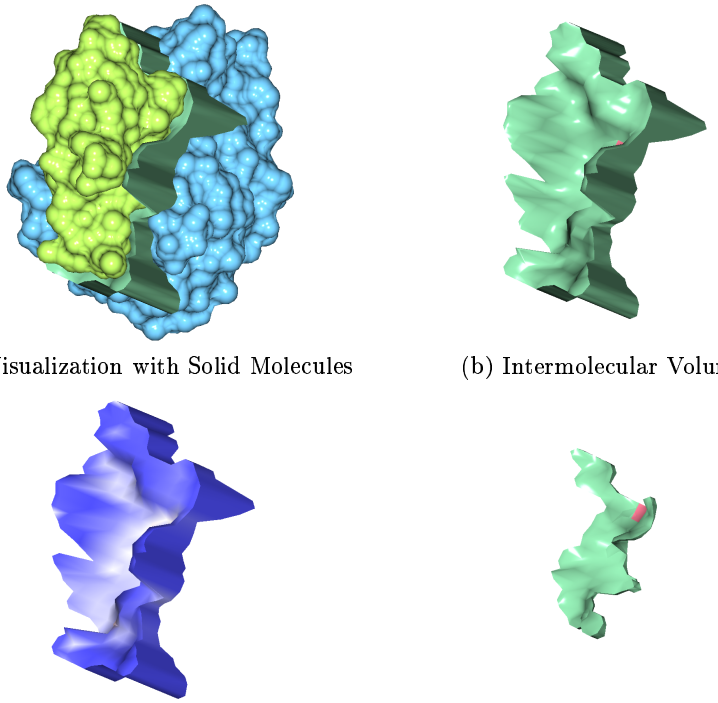


Fig. 4. The Interface Coordinate System and the Molecular Projection Regions



(a) Visualization with Solid Molecules

(b) Intermolecular Volume

(c) Color-coded Negative Volume

(d) Thresholded Negative Volume ($\epsilon = 2.8\mathring{A}$)

Fig. 14. Interactive Visualization of the Intermolecular Negative Volume between Proteinase and Inhibitor (4SGB)