Revisiting the description of Protein-Protein interfaces. Part I: Algorithms

Frédéric Cazals, Flavien Proust

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Revisiting the description of Protein-Protein interfaces. Part I: Algorithms

Frédéric Cazals — Flavien Proust

N° 5346 – version 2

version initiale Octobre 2004 – version révisée Février 2005

Thème SYM

ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

Revisiting the description of Protein-Protein interfaces. Part I: Algorithms

Frédéric Cazals , Flavien Proust

Thème SYM — Systèmes symboliques
Projet Geometrica

Rapport de recherche n° 5346 – version 2 *— version initiale Octobre 2004 — version révisée Février 2005 29 pages

Abstract: The description of non-covalent contacts in (macro-)molecular assemblies is of fundamental interest for understanding the formation and the stability of such complexes. This description focuses on shape complementarity and packing properties at the interface, connectivity and chemical relationships across the interface, as well as global measures such as the Solvent Accessible Surface Area (SASA) lost upon complex formation, or the interface planarity.

This paper develops a notion of interface which (i)unifies most (if not all) the statistics used so far to describe interfaces (ii)provides additional insights from a chemical, geometric and topological standpoints, from the atomic scale to the interface scale.

Consider a complex made of two proteins. First, we define from the Voronoi diagram of the atomic balls an interface between the proteins, which is coherent with the atoms loosing accessibility to the solvent upon formation of the complex. Second, we show how to study the topology and the geometry of the interface. Third, we show how to weight the interface facets so as to recover the SASA lost upon formation of the complex. Fourth, we revisit the previous contributions to accommodate crystallographic water molecules at the interface. Our interface model is implemented using the CGAL library, and the corresponding module integrated to a VMD plug-in. A detailed experimental study on the usual database of protein-protein complexes is provided in a companion paper.

Key-words: Structural Biology, Molecular Modeling, Molecular Interfaces, Voronoï diagrams.

* Revision of [CP04] triggered by [CP05]. The initial title of [CP04], *On the topology and the geometry of (Voronoi) molecular interfaces. Part I: algorithms* was changed to comply with that of [CP05]. Main changes between the text bodies are: section on previous works, appendix I.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Re-examen des interfaces de complexes Protéine - Protéine. Première partie: Algorithmes

Résumé : La description de contacts non covalents au sein d'assemblages (macro-)moléculaires est d'un intérêt capital pour la compréhension de la stabilité et des mécanismes de formation de tels complexes. La description de tels contacts est généralement faite en considérant la complémentarité stérique et chimique des atomes à l'interface, les propriétés de packing à l'interface, ainsi également que des mesures globales telles que la Surface Accessible au Solvant Perdue (SASL) lors de la formation du complexe, ou la planarité de l'interface.

Ce travail a pour objet le développement d'un modèle d'interface qui (i)unifie les statistiques utilisées précédemment (ii)permet de réaliser une analyse plus fine des interfaces d'un point de vue chimique, géométrique, et topologique, à l'échelle des atomes mais aussi à celle de l'interface.

Considérons un complexe formé de deux protéines. Premièrement, en utilisant le diagramme de Voronoï des sphères atomiques, nous définissons une interface entre les protéines, interface cohérente avec les atomes perdant de l'accessibilité au solvant lors de la formation du complexe. Deuxièmement, nous montrons comment étudier la topologie et la géométrie de l'interface. Troisièmement, nous présentons une construction permettant de pondérer les facettes de l'interface par la Surface Accessible au Solvant Perdue, et montrons comment ce mécanisme de pondération peut également impliquer des atomes situés à proximité de l'interface. Enfin, nous montrons comment faire intervenir les molécules d'eau cristallographiques dans les constructions précédentes. L'algorithme de calcul d'interface utilise la librairie CGAL et peut être appelé depuis VMD. Une étude expérimentale exhaustive sur les bases de données usuelles de complexes protéine-protéine est présentée dans un papier associé.

Mots-clés : Biologie Structurale, Modélisation Moléculaire, Interfaces, Protéines, Diagrammes de Voronoï.

1 Introduction

1.1 Molecular interfaces

Non covalent interactions in-between proteins are central in many structural biology research avenues, among which the understanding of flexibility issues and induced fit phenomena in docking, the characterization of stability issues in complex formation, or the characterization of the specificity of interfaces. Understanding the mechanisms underlying these interactions is therefore central for the investigations of biological function. The description of these non covalent interactions involves listing the atoms in contact together with the corresponding properties. Following [CJ99], the corresponding aspects of structure involved cover (i)the size and the chemical character of the protein surfaces which are buried at the interface (ii)the connectivity and packing properties of atoms making contacts across the interface (iii)the number of hydrogen bonds and interface water molecules involved in polar interactions. These statistics are ubiquitous in the definition and the calibration of scoring functions used by docking or folding algorithms.

Interface atoms and interfaces. In the previous discussion, we implicitly assumed one can identify interface atoms and pairs of atoms across an interface. The most usual constructions to do so are the following —see also section 2. Consider a complex made of two (macro-)molecules A and B, and given a water probe of fixed radius, denote SAS_A / SAS_B / $SAS_{A\cup B}$ the Solvent Accessible Surface (SAS) of molecule A / molecule B / the complex $A \cup B$. Also denote SASA the SAS Area of a given SAS.

On one hand, the interface is defined as the collection of atoms loosing accessibility to the solvent upon formation of the complex $A \cup B$. This definition does not properly define an *interface* i.e. a geometric entity separating the atoms of the two molecules, but rather focuses on the SASA lost upon formation of the complex, that is on $SASAL = SASA_A + SASA_B - SASA_{A\cup B}$. (Notice that this calculation does not involve steric change upon formation of the complex since $SASA_A$ and $SASA_B$ are computed from the coordinates of the molecules in the complex.) On the other hand, given a threshold distance —typically 8, the interface is defined as the collection of pairs of atoms whose distance is less than the threshold. This second definition identifies pairs of atoms across the interface, but does not define an *interface* neither.

While the former construction is mainly used for the classification of interfaces (in particular protein-protein interfaces), the later is ubiquitous in the derivation of statistical potentials. More generally, the definition and the study of interfaces usually involves numerous statistics computed from a variety of algorithms requiring user-defined parameters. Our goal is to develop a notion of interface unifying all these notions.

1.2 Contributions

Motivated by the previous discussion, this paper aims at providing a more accurate description of interfaces, so as to distinguish families of protein-protein complexes, assess the role of water molecules, ease the identification of functional epitopes on proteins, etc. We make the following contributions. First, we define a notion of interface which accommodates crystallographic water molecules and is coherent with atoms loosing accessibility in a complex. Second, we study the topology (connected components, holes) and the geometry of interfaces. Third, we show how to weight the interface facets so as to recover the SASA lost upon formation of the complex. A detailed experimental study on the complexes of [CJ99, BCRJ04] can be found in the companion paper [CP05]. These contributions exploit the relationship between the boundary of a union of balls —the atomic balls expanded by the water probe— and the so-called α-complex and α-shape of the balls, as well as properties of the arrangement of balls.

Our construction unifies most (if not all) the classical statistics and algorithms used to describe complexes, and also allows a more local and detailed analysis, from the topological, geometric, and chemical perspectives. Moreover, it is based on the Delaunay and does not involve any user-defined parameter.

1.3 Model and notations

To study interfaces within a complex, we shall use the following terminology. We assume the complex features four molecular species, A, B, W, U. Typically, A and B are two proteins — or A is a protein and B a substrate or a drug, while W are water molecules and U additional molecules —for example example co-factors. The focus will be on interfaces involving two species —called bicolor interfaces, and on the interfaces involving the three species ABW —called the tricolor interface.

Consequently, when raising properties of a bicolor interface regardless of the species involved, we shall replace the labels A and B by the generic labels X and Y. Since pairs of atoms will be examined within triple or quadruples, the generic labels accompanying X, Y will be Z and T. On the other hand, when discussing specific properties of the tricolor interface, we shall use the labels A, B, W. When considering pairs of atoms with different types, we shall also says these atoms have different colors. Accordingly, bicolor, tricolor and quadricolor simplices refer to simplices featuring two, three or four types —out of A, B, W, U or X, Y, Z, T.

A ubiquitous notion used in this paper is the Solvent Accessible Surface and its area, denoted SAS and SASA. When restricted to a particular atom identified by its index i, the corresponding notations are SAS_i and $SASA_i$. Finally, the SAS lost by atom i in the formation of a complex is denoted $SASL_i$, and the corresponding area $SASAL_i$. Finally, when displaying bicolor arrangements of balls, we shall use the following conventions.

- Balls: a ball represents an atom expanded by a water probe.

- Circle arcs: solid arcs represent the boundary of the complex, dashed arcs represent the SAS lost upon formation of the complex, dotted arcs represent the SAS of the individual atoms lost upon formation of the spare molecules. See Fig. 14 for an illustration.
- Edges in Delaunay and Voronoi diagrams: an edge is colored purple when connecting two balls from different molecules. The edge (and its dual Voronoi facet) are depicted in solid when contributing to the interface, and dashed otherwise.

All the constructions presented thereafter are related to the Voronoi digram of balls (also called the power diagram). Readers not familiar with Voronoi diagrams, α-complexes and α-shapes, as well as the relationship between these constructions and molecular surfaces are referred to section 9.

1.4 Paper overview

Previous work is discussed in section, 2. Bicolor interfaces are defined in section 3 and section 4 is devoted to the study of their topology and geometry. The strategy used to weight the interface facets with the SAS lost is explained in section 5. Tricolor interfaces accommodating water molecules are described in section 6. Concluding remarks are reported in section 8. Finally, a primer on Voronoi diagrams and related constructions is provided in section 9, while complementary material and proofs of the claims are provided in section 10.

For illustrations of the constructions on real examples, the reader is referred to the companion paper [CP05].

2 Review of previous work

The characterization of interfaces has been subject to intense research. Based on atoms loosing accessibility in a complex, [CJ99, CJ02, BCRJ04] study parameters of interfaces at the atomic and the interface scales. At the atomic level, the classification of interface atoms as buried or exposed, the description of neighbors, the packing properties using Voronoi volumes, the dissection of chemical properties of interface atoms have been studied. At a more global level, one encounters the solvent accessible surface area lost upon formation of the complex (SASL), the number of connected components of the interface (based on a clustering of interface atoms), or the characterization of the flatness of interfaces (based on the atomic deviation wrt a least-square plane through the interface atoms). In a docking perspective, [PGF98] investigates molecular surface patches at the interface —atoms in contact are defined by a distance in-between 0.5 and 2.8 . In order to discriminate homodimeric and monomeric proteins from the crystalline state, statistical potentials mixing distances and surface areas are developed in [PHT00]. In a similar vein but with applications to docking, statistical potentials based on triples of atoms in contact are investigated in [AJBL03] —triples being read from the α-complex of the atomic balls.

Aside these applied contributions, more conceptual interface models have been developed. In [GW96], an interface is defined as the points in-between two molecules where two potentials emitted by the molecules take the same value. In [BER04], an interface based on a *retraction* process using the Voronoi diagram of the atomic balls is defined. Speaking of Voronoi diagram, the packing properties of atoms at interfaces has been characterized using Voronoi cells volumes [Ric77]. Such contributions are also related to the question of shape complementarity. Shape complementarity at interfaces has been measured by the integral of a soft potential between the molecular surfaces [aMS92], by a correlation index based on the the integral of a function involving the normals of the solvent accessible surfaces [LC93], and by the volume of the envelope of spheres filling inter-facial cavities [Las95].

As outlined by this review, interfaces have been investigated using a variety of statistics, algorithms —and actually cut-offs. The goal of this paper is to reconcile these contributions by developing a coherent treatment of interfaces from the chemical, geometric and topological standpoints, from the atomic scale to the interface scale.

3 Bicolor Voronoi interfaces

In this section, we define bicolor Voronoi interfaces. We use the terminology of bicolor AB interface —the interface between the two proteins A and B, although the presentation is identical for any bicolor XY interface.

3.1 Bicolor interface and interface neighbors

Following the definitions recalled in introduction, an atom contributes to the interface if it is loses exposition to the solvent, which is equivalent to saying that its SAS penetrates the SAS of an atom of the second molecule. Notice that loosing accessibility either means loosing it completely or retaining a contact with the solvent. In terms of pairwise contacts, an atom loosing accessibility but retaining contact contributes to the boundary of the union of balls through edges which are either singular or regular in the corresponding α-complex. On the other hand, an atom loosing accessibility and getting buried makes contacts to atoms of the second molecules through interior edges of the corresponding α-complex.

Let us model a molecule (and thus a complex) as a collection $\{B_i\}_{i=1,\ldots n}$ of Van der Waals balls. More precisely, denote $B_i(a_i, r_i)$ the atomic ball of center a_i and radius r_i, and let $S_i(a_i, r_i)$ be the corresponding sphere. Also denote r_w the radius of the water probe used to define the SAS surface. Summarizing the previous discussion discussion, we define:

Definition. 1 *An AB interface edge is an edge of type AB in the α-complex of the balls $B_i(a_i, r_i + r_w)$, with $\alpha = 0$. The interface neighbors of a sphere S_i are the atoms of the second molecule sphere S_i is connected to through an interface edge. The interface neighbors of S_i are denoted $N(i)$.*

The AB interface is defined as the collection Voronoi facets dual of the AB interface edges. A Voronoi edge bounding an interface Voronoi facet is called an interface Voronoi edge.

Interface neighbors and intersecting balls. To begin with, we prove that the notion of interface neighbors is sufficient to account for the loss of accessibility:

Observation. 1 *Interface neighbors exactly identify the atoms loosing accessibility.*

From an applied perspective, definition 1 deserves a discussion wrt the following classification of interface atoms, introduced in [CJ99]. Let the distance between two atoms be the distance between their Van der Waals spheres. Call an interface atom a contact atom if it has a party at distance less than 0.5 , and call it buried if it looses contact to the solvent. (The threshold of 0.5 corresponds to Van der Waals contacts across the interface.) Based on this terminology, interface atoms are classified as A (retaining accessibility, no contact), C (retaining accessibility, contact), B (buried, contact). Note that this classification does not qualify buried atoms having neighbors on the second molecule at distance within the range 0.5 to 2.8 .

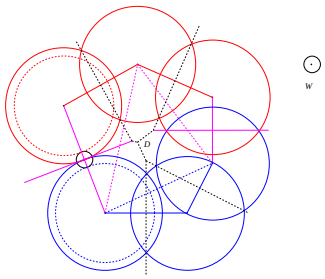

Figure 1: A dome of free space for water molecules

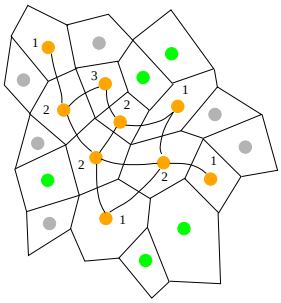

Figure 2: Facets are at depth one to three in the hydrophobic core

3.2 Solvent accessibility, the core of a bicolor interface, and hot spots

The SAS separates the expanded atoms from the free volume accessible by water molecules. Consider an interface edge. If the intersection circle of the two atoms associated to the edge contributes to the SAS of the complex, there then is a packing defect since the atoms contribute to the SAS of the complex. This motivates the following:

Definition. 2 *A* core *interface edge is an interface edge which which is interior in the α-complex. The core interface consists of the core interface edges.*

From a chemical standpoint, it is well known that all the amino acids present at an interface do not play an equivalent role, and those whose energetic contribution is prominent are termed as *hot* [CW95, AUVW97, BT98]. It is usually agreed that these amino acids are always buried —a necessary but not sufficient condition for being hot— and face one another. Therefore, the core of the bicolor interface is likely to yield important informations on the hot residues. Notice that the property of being buried for an interface edge does not imply the same for the atoms —which may contribute to the SAS.

With a slight anticipation, we discuss an algorithm to investigate the hot residues. Assume we have computed the core of the interface and isolated its connected components. Then, an appealing strategy consists of seeking large connected components featuring the same chemical (typically hydrophobic) properties. See Fig. 2 for an illustration.

Finally, while discussing the relationship of bicolor interfaces with the SAS, we observe that boundary edges of bicolor interfaces always witness contact with the solvent:

Observation. 2 *Consider a Voronoi edge bounding a single interface Voronoi facet f —of any bicolor type. Then, the intersection circle associated to the dual edge of f contributes to the SAS of the complex.*

Is should be noticed however, that the condition is sufficient but not necessary. A triangle AAB may contribute two Voronoi facets dual of the AB edges, yet the two intersection circles may contribute to the SAS. See Fig. 1 for an illustration.

4 Topology and geometry of bicolor interfaces

In this section, we investigate the topology and the geometry of a bicolor XY interface.

4.1 Topology of bicolor Voronoi interfaces

Consider a bicolor interface of type XY. Studying its topology amounts to describing the way the Voronoi facets dual of interface edges patch with one together. A Voronoi facet belongs to exactly two Voronoi regions and separates two balls of types X and Y, so that the Voronoi interface is a surface with boundary. To describe the topology of the interface, one has to examine the configurations an XY interface edge is involved in. Since such an edge belongs to a tetrahedron featuring at least the two molecular species XY, we define:

Definition. 3 *The type of a d-simplex is defined as the $(d+1)$-tuple enumerating the multiplicities of the molecular species represented at the vertices of the simplex.*

Notice that the previous definition does not impose any ordering on the vertices. We have:

Observation. 3 *A bicolor tetrahedron has type $(1,3)$ or $(2,2)$. A tricolor tetrahedron has type $(1,1,2)$, while a quadricolor tetrahedron has type $(1,1,1,1)$.*

Examining these configurations yields the following:

Observation. 4 *The bicolor interface has the following properties:*

- *The interface is a cell complex with boundary.*
- *A Voronoi edge bounding an XY Voronoi facet has one or two incident XY Voronoi facets. Such a Voronoi edge is on the interface boundary iff only one edge of its dual triangle is an interface edge of type XY.*
- *The neighborhood of every Voronoi vertex is either a topological disk, a half-topological disk, or two half-topological disks pinched together at the Voronoi vertex. This later case correspond to a tetrahedron of type $(2,2)$, where two bicolor edges are interface edges, and these two edges span the four vertices of the tetrahedron.*

An interface can be connected through a pinched Voronoi vertex of a $(2,2)$ tetrahedron. For such a tetrahedron, the situation is that depicted on Fig. 3(b), namely out of the 4 potential bicolor intersections, only two actually occur. Phrased differently, we have two pairs of independent atomic balls. Since the intersection of the corresponding Voronoi facets does not encode a joint property of the two pairs of balls, we define:

Definition. 4 *Two Voronoi facets are called edge-connected if they share a Voronoi edge. An edge-connected component of the interface is a collection of edge-connected Voronoi facets.*

The definition is illustrated on Fig. 4, which features two edge-connected components. Although the light grey interface is locally pinched around vertex v_1, the two corresponding Voronoi facets belong to the same edge connected component. Since there is no ambiguity, we shall just talk of connected components in the sequel.

It should be noticed that the edge connectivity just defined could be tightened by requiring the three balls to intersect —a stronger condition than the pairwise intersection along the two bicolor edges.

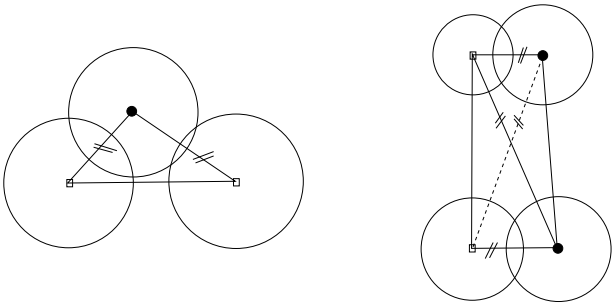

Figure 3: Interface connectivity (a)Edge connected Voronoi facets —through the dual of the triangle (b)Pinched Voronoi vertex for a $(2,2)$ tetrahedron

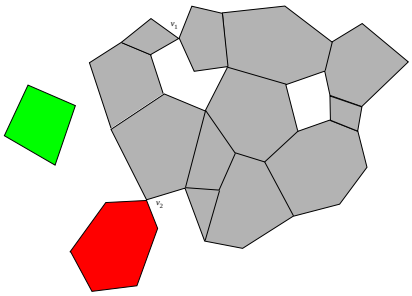

Figure 4: An interface with three connected components and four boundary loops

4.2 Computing bicolor interfaces and their boundaries

Retrieving connected components. Given an initial Voronoi facet f_0 dual of an edge e_0, the exploration of the corresponding connected component requires running a breadth-first-search (or depth-first-search) like algorithm anchored at f_0. To run such an algorithm, the only information required is the list of Voronoi facets edge connected to a given facet f. But two edge connected Voronoi facets are dual of edges of the same triangle. Therefore, to report the facets edge connected to a facet f dual of an edge e, we just need to report the Delaunay triangles (i)incident to e (ii)featuring a second edge of type $tag(e)$ —with $tag(e) \in \{AB, AW, BW\}$ the type edge e. This is easily done by rotating around edge e in the Delaunay triangulation.

Retrieving the cycles bounding a connected component. Consider an edge connected interface of type AB, AW or BW, and call the closed curves bounding this component the *loops*.

Starting from an initial boundary Voronoi edge e, a given loop can be computed by iteratively following the successor of e on the boundary of the connected component. Assume e is oriented and denote $s(e), t(e)$ the corresponding source and target Voronoi vertices, and let $T(v)$ be the tetrahedron associated with Voronoi vertex v. The successor of edge e is one of the four Voronoi edges dual of the facets of $T(t(e))$. Following observation 4, if the neighborhood of $t(e)$ is a half-topological disk, tetrahedron $T(t(e))$ has only two facets whose dual Voronoi edges are boundary edges. Since e is one of them, extending the loop requires retrieving and following the second one. On the other hand, if $t(e)$ is a pinched Voronoi vertex and if the four edges belong to the boundary of the connected component processed, there are three potential outgoing edge, but only one bounding the Voronoi facet which has e on its boundary. Again, extending the loop requires finding and following this edge.

Equipped with this extension operation, computing the loop of a given edge e_0 requires picking an arbitrary orientation for e_0, and following the boundary until the Voronoi vertex $T(s(e_0))$ is encountered again. To retrieve all the loops, we just have to iterate over the remaining boundary edges which are not already part of a loop.

The previous description actually eludes a difficulty, namely the way infinite tetrahedra are handled. (Recall these are the tetrahedra featuring a triangle of the convex hull of the Delaunay triangulation). When such a tetrahedron is encountered during the extension of a loop, we do not report its Voronoi center but the weighted circumcenter of the finite facet —that is the center of the smallest sphere orthogonal to the three spheres associated with the vertices of the facet.

A final comment is in order. A pinched vertex may not disconnect an edge connected interface, but in some cases, it might be interesting to report all the loops regardless of non-local connexions. For example on Fig. 5(a), one may wish to report two loops and not one. The previous algorithm remains valid, provided a straightforward modification of the extension rule at Voronoi vertices of $(2, 2)$ tetrahedra.

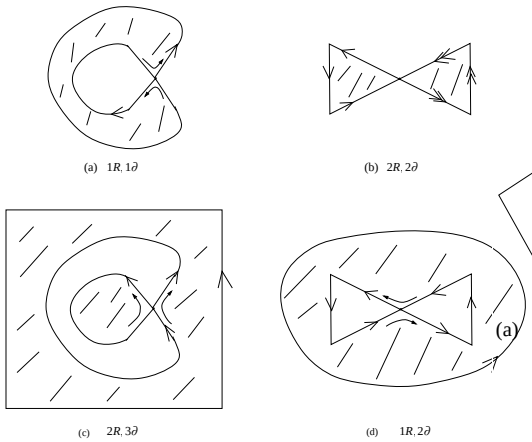

Figure 5: Edge connected regions, pinched vertices and boundaries: R=Edge Connected Region, ∂=boundary curve

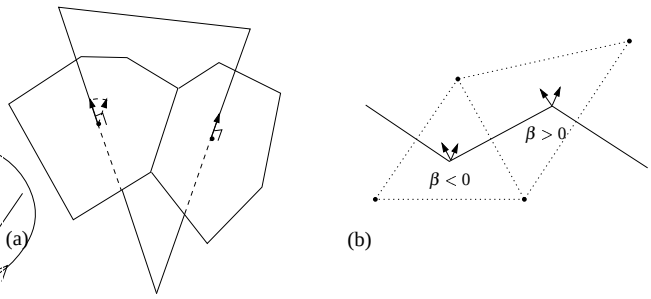

Figure 6: Mean curvature of the Voronoi interface (a)Convex angle between two facets (b)Side view

4.3 Geometry of connected components

The most straightforward statistic for interfaces is the surface area and the boundary length. But one may also check their curvature. Since an interface is characterized by a collection of interface atoms, it is most natural to study its curvature using some discrete notion of curvature. Since a bicolor interface is an orientable surface, the natural way to characterize its *extrinsic* curvature consists of using the mean curvature. (Which is more suited than the Gauss curvature which is *intrinsic*.) Recall that the mean curvature of a polyhedral surface is carried out along edges, the amount of mean curvature attached to an edge being defined by $h(e) = \beta(e)l(e)$, with $\beta(e)$ the angle between the normals to the facets incident to e —this angle being counted positively (negatively) is e is convex (concave), and $l(e)$ the edge length. (See e.g. [San79, CSM03].) The calculation of this angle is especially easy in our setting, since the angle between two incident Voronoi facets is the angle between the dual Delaunay edges.

The mean curvature just introduced actually has two drawbacks. First, in a chemical setting, one expects dihedral angles to alternate. Second, although each connected component is a (possibly pinched) orientable surface, in order for the sum of the mean curvature of edges of different components to make sense, one needs to orientate these components coherently. To get around these two difficulties, we use $\mid h(e) \mid$ rather $h(e)$.

4.4 On the geometry of interface facets

Defining the interface from the α-complex provides a filtering mechanism for long bicolor edges located either inside a cavity of the complex or near the boundary of the convex hull. Indeed, long edges belong to the Delaunay triangulation but not the the α-complex. This filtering mechanism avoids using explicit solvent molecules, a strategy often resorted to in applications deriving statistical potentials from the Delaunay triangulation. However, this filtering mechanism does not provide any control on the geometry of the interface Voronoi facets, and in particular, large facets are expected near the convex hull of the atoms centers. In other terms, interface edges encode the topology of the interface but no its geometry.

To retrieve a relevant geometric information, we propose two alternatives. The first one consists of weighting the interface facets so as to encode exactly the SAS lost upon formation of the complex —this is the weighting scheme presented in section 5. The second one builds upon the observation that boundary atoms do not play a major role from an energetic standpoint [BT98], so that one may discard selected boundary edges these atoms are involved in. One way to discard large Voronoi facets is the following. Recall than any simplex in the α-complex comes with an $\overline{\mu}$ value which gives the weight of its largest orthogonal ball. For an interface edge e, denoting w_e the weight of its smallest ball, one can therefore discard the edge if $\overline{\mu}/w_e \geq M$, with M a positive number. Since weights of balls are equal to their square radii, the conditions amounts to saying that the radii of the balls are within a factor $\sqrt{M}$.

Finally, we provide a relationship between the core of interfaces —def. 2 and the previous filtering mechanism. Since w_e is always positive, discarding all the edges such that $\overline{\mu} \geq 0$ means keeping those such that $\overline{\mu} < 0$, which are interior in the α-complex with $\alpha = 0$. Therefore:

Observation. 5 *Filtering the interface edges with the threshold $M = 0$ amounts to selecting the interface core.*

4.5 About Voronoi interfaces

Our interface definition deserves s special comment wrt the construction performed in [BER04], which consists of using the filtration encoded in the Delaunay triangulation so as to define the interface as a subset of the bicolor Voronoi facets. But there is a fundamental differences between the two Voronoi interfaces.

As recalled by observation 8 —section 9, using the Delaunay filtration amounts to enlarging the atomic balls by an amount which depends of the radii of the atoms. Instead, using $\alpha = 0$ has the advantage of enlarging all the atomic balls by the same amount. Consequently, the atoms we select are exactly those loosing accessibility, a property not enjoyed by the construction of [BER04].

5 Weighting the bicolor interface in terms of SAS lost

5.1 Balls, Voronoi diagram of balls, and SAS lost

Bicolor Voronoi interfaces are interesting since they define an interface and exactly identify the atoms loosing accessibility, but they do not give a precise account of the SASA lost upon formation of the complex. An appealing strategy would be to weight the Voronoi interface facets in such a way that the sum of the weights equals SASAL.

To weight a Voronoi facet, the natural strategy would be to charge the surface area lost by the corresponding intersecting spheres to the facet. But in doing so, whenever a spherical cap is contained in more than one sphere, the corresponding surface area is reported as many times. For example on Fig. 7, $SASL_1$ consists of three circle arcs contained in one or two balls of the second molecule. An appealing alternative would be to use the partition of each ball by the Voronoi diagram, from which one could charge a spherical cap of one atom not contained in its Voronoi region to the sphere of the Voronoi region it belongs to. For example on Fig. 7, one could charge the circle arc af of S_2 to S_1 since this circle arc belongs to the Voronoi region of S_1. But in doing so, one does not precisely account for the SAS lost —which consists of the circle-arc $ag \subset af$ on this example.

Summarizing this discussion, weighting the interface facets so as to recover SASL requires considering the spherical caps of balls of one color contained in balls of other colors, and weighting them according to the number of containing balls. We now formalize this construction.

5.2 Weighting the bicolor facets

Given a collection of balls $\mathcal{S} = \mathcal{A} \cup \mathcal{B}$ carrying the labels of the two molecules A and B, consider its arrangement, that is the decomposition of spheres in spherical caps, circle arcs and vertices. Denote $\mu(R)$ the surface area of any spherical region R. For a sphere S_i, consider the partition of $SASL_i$ by its interface neighbors. In the arrangement of all balls, these are exactly the spherical caps of S_i which are contained in balls of the second molecule. We assume each such region is identified by the multi-index I featuring the indices of the atoms of the second molecule, and we denote $\pi(i)$ the collection of all such multi-indices. Finally, given a collection of balls I, denote $C_{i;\ I}$ the portion of S_i contained in the intersection of the balls in I, that is

$$C_{i;\ I} = S_i \cap \cap_{j \in I} B_j. \tag{1}$$

With the previous notations, since by observation 1 the interface neighbors cover $SASL_i$, we have:

$$SASAL_i = \mu(SASL_i) = \sum_{I \in \pi(i)} \mu(C_{i;\ I}), \tag{2}$$

$$SASAL = \sum_{i \in A} SASAL_i + \sum_{j \in B} SASAL_j. \tag{3}$$

Recall that the size $\mid I \mid$ of a multi-index I is the number of its arguments. We define:

Definition. 5 *Consider two intersecting balls B_i and B_j corresponding to an interface edge ij. The* inclusion coefficient *of sphere i wrt ball j is defined by:*

$$\mu^{\prime}_{i\rightarrow j} = \sum_{I\in\pi(i), j\in I} \frac{\mu(C_{i;\ I})}{\mid I \mid}. \tag{4}$$

Also define the strength F_{ij} *of the interface facet as $F_{ij} = \mu^{\prime}_{i\rightarrow j} + \mu^{\prime}_{j\rightarrow i}$.*

Recalling that the interface neighbors of a sphere S_i are denoted $\mathrm{N}(i)$, we have:

Observation. 6 *The inclusion coefficients satisfy $\sum_{j\in N(i)} \mu^{\prime}_{i\rightarrow j} = SASAL_i$. Moreover, denoting IE the set of interface edges, one has:*

$$\sum_{ij\in IE} F_{ij} = SASAL. \tag{5}$$

Equipped with this weighting strategy, the Voronoi interface describes the topology of the interface, and also encodes the Surface Area lost upon formation of the complex. Notice also that from an algorithmic standpoint, the definition of inclusion coefficients just involves the neighbors of an atom in the α-complex.

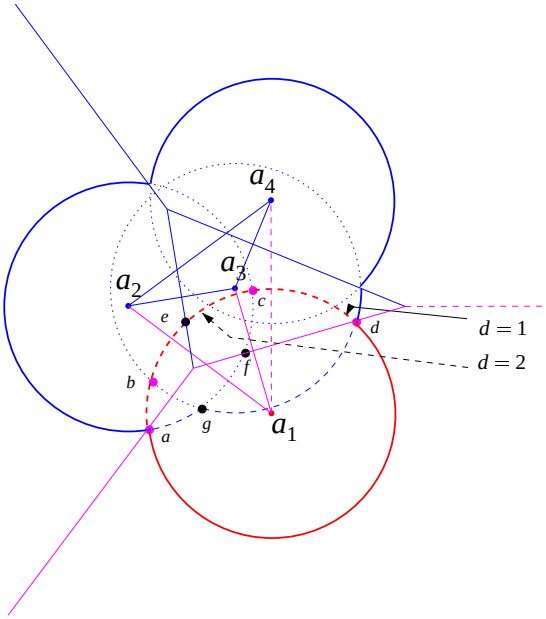

Figure 7: Caps of $SASL_1$ are at depth one or two in the arrangement of balls. (Bottom ball is red, top ones are blue.)

5.3 Application to statistical potentials

Statistical potentials are often based upon pairwise frequencies of distances, yet variants based upon SAS have been proposed. While the focus often consists of using SASL as a whole [PHT00], attempts have been made [AT97] at charging surface related quantities to pairs of residues. But whenever the SAS of a residue is occluded by several other residues, the pairwise values do not account for the influence of the environment. To account precisely for SAS lost and give an account of multiple contacts, we propose to use the weighting scheme just defined.

In the previous construction, we partitioned $SASL_i$ by interface neighbors which form a subset of all balls intersecting S_i. Replacing the former by the later yields the following:

Observation. 7 *For a ball S_i, let $\pi(i)$ be the collection of multi-indices corresponding to the partition of $SASL_i$ by all balls of the second molecule, and define the inclusion coefficients accordingly —by Eq. (4). Then, the sum of inclusion coefficients over all intersecting pairs of balls satisfies:*

$$\sum_{i\in A}\sum_{j\in B}\mu'_{i\to j}+\sum_{j\in B}\sum_{i\in A}\mu'_{j\to i}=SASAL. \tag{6}$$

5.4 Algorithms

From an algorithmic perspective, the calculation of the coefficients introduced in the previous two sections can be done as follows. First, recall that due to packing properties of atoms, an atomic ball is intersected by a constant number of atomic balls, so that arrangements of atomic balls have linear complexity [HO94]. Also recall that if one is given a collection of balls $\{B_j\}$ intersecting a given ball B_i, the arrangement on B_i of the intersection circles $B_i \cap B_j$ can be computed in constant time since the number of neighbors is bounded by a constant [HS98].

Consider the problem of computing the inclusion coefficients based on interface neighbors. First, one computes the Delaunay triangulation and the corresponding α-complex, and retrieves all neighbors —of any color. Second, for each ball B_i, one computes the arrangements of intersecting circles as explained in [HS98]. The relevant portions of S_i are those covered by balls of the second color only.

Consider now the coefficients based on all balls intersecting a given ball. Starting from the neighbors of a ball B_i in the Voronoi diagram, one first collects all the balls interesting B_i by recursively visiting the neighbors of S_i. (Care has to be taken if there are hidden balls, but if no ball contains the center of any other ball —which is usually the case for atoms, each ball has a Voronoi region and the difficulty does not occur.) Then one computes the arrangement of intersecting circles, and focuses only on the caps of S_i contained in balls of the second color.

In any case, for a total of n atoms, the costly operation is the construction of the Delaunay triangulation, which has $O(n^2)$ worst-case complexity. Practically for atoms, the triangulation has linear size and can be computed in expected $O(n \log n)$ time [BY98].

6 Tricolor interfaces and water molecules

6.1 The $AW - BW$ interface

When considering an interface, an interesting question is the role played by crystallographic water molecules. If one has three molecular species A, B, W, one can define three types of bicolor interfaces. But since we primarily care for the AB interface, contact of type AW and BW are of interest only when located near the AB interface. We therefore define —see Fig. 8:

Definition. 6 *An interface water molecule is a ball of type W which is the vertex of at least one edge of type AW and at least one edge of type BW, both edges belonging to the α-complex of the balls $B_i(a_i, r_i + r_w)$ —with $\alpha = 0$. An AW (or BW) interface edge is an edge of type of type AW (or BW), with W an interface water molecule. The AW (BW) interface is defined as the collection Voronoi facets dual of the AW (BW) interface edges.*

A further refinement of constrained bicolor interface consists of aggregating Voronoi facets of type AW and BW:

Definition. 7 *The $AW - BW$ interface is the collection of Voronoi facets dual of edges of type AW or BW. A connected component of $AW - BW$ interface is a collection of edge-connected Voronoi facets dual of interface edges of type AW or BW.*

Edge connected Voronoi facets of types AW and BW are found in bicolor, tricolor or quadricolor tetrahedra. Let us analyze the last two cases, depicted on Fig. 21. Since in such tetrahedra we disregard AB edges, the configurations found are those of bicolor tetrahedra. More precisely, a tetrahedron of type $AABW$ where AB edges are omitted is equivalent to a bicolor $(3,1)$ tetrahedron. A tricolor tetrahedron of type $ABWW$ is similar to a bicolor $(2,2)$ tetrahedron. Finally, a $ABWX$ tetrahedron is equivalent to a $(2,1,1)$.

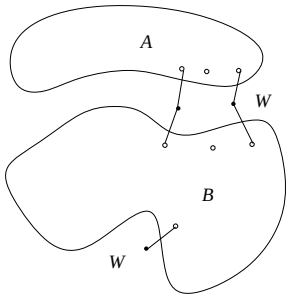

Figure 8: The top two contacts qualify as AW and BW. The bottom one does not

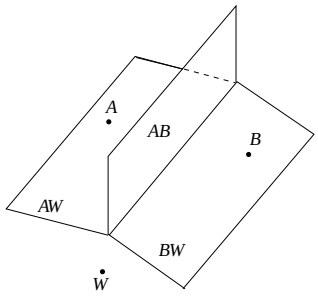

Figure 9: The boundary of the union of the AB and $AW - BW$ interfaces may not be a one-manifold

6.2 The ABW interface

As just observed, the topological properties of the tricolor interface are the same as those of any bicolor interface. In particular, both are surfaces with possibly pitched vertices. From an application perspective and to assess the role of water molecules, it makes sense to study the union of the AB and the $AW - BW$ interfaces, and we define:

Definition. 8 *The ABW interface is defined as the union of the AB and $AW - BW$ interfaces.*

But the topology of the union is more involved than that of the singletons. First, non-manifold Voronoi edges may appear —if the three facets dual of the corresponding triangle are present in the union of the two interfaces. Second, the boundary of the union may not be a one-manifold, and we call it a curve network or net for short —Fig. 9. To deal with these difficulties, it is actually sufficient to compute the interfaces separately, run a Union-Find algorithm to maintain the connected components of the edge-connected components, and another union-find algorithm to maintain the connected components of the boundary loops of the connected components of the union. Finding the Voronoi edges along which connexions occur can be done while computing the interfaces, while running m Union-Find operations on an n-element set takes $O(m\alpha(m,n)$ with $\alpha(m,n)$ the inverse of Ackerman's function [Tar83].

6.3 $SASL$ and water molecules

So far, we have defined the SAS lost as $SASAL = SASA_A + SASA_B - SASA_{A\cup B}$. But whenever molecules of other types (W or U) are present, some of them may account for a loss of accessibility, and one can define $SASAL = SASA_A + SASA_B - SASA_{A\cup B \text{wrt} A\cup B\cup W\cup U}$. The algorithms described in section 5 remain the same since one just needs to compute the spherical caps of atoms of A and B contained in balls on any other color.

7 Implementation

The algorithms described in sections 3 and 6 have been implemented using the CGAL library. The corresponding module, *intervor* can either be called from the command-line on a pdb file containing a complex, or from a VMD plugin. Both *intervor* and the VMD plugin will be made public. A detailed experimental study on the complexes [CJ99, BCRJ04] can be found in the companion paper [CP05].

8 Conclusion and future work

This paper introduces a notion of interface which provides a compact representation of molecules and complexes, and makes the following contributions. First, we introduce a notion of molecular interface coherent with atoms loosing accessibility in a complex, and

which also accommodates crystallographic water molecules. Second, we show how to study the topology and the geometry of interfaces. Third, we show how to weight interface facets so as to recover the SASA lost upon formation of the complex, and extend the construction to define statistical potentials based on SASL. A detailed experimental study on the complexes of [CJ99, BCRJ04] can be found in the companion paper [CP05]. Apart from these contributions, several research directions may benefit from the ideas and algorithms developed in this paper.

Contacts within a protein and applications. Our focus has been on interfaces between two (macro-)molecules, but one could extend this notion to contacts between atoms or amino-acids within a protein. Such a development might find applications in folding, for the study of allosteric pathways, or for the classification of proteins.

Relationship to Knowledge Based Scoring Functions. We have shown how to charge SASL to all pairs of intersecting balls. This construction deserves further investigation in the context of statistical potentials. Notice also that the notion of contact across the interface could be used to define a notion of chemical environment and accompanying statistical potentials.

Connexions with higher order Voronoi diagrams. Our definition of interface is based upon interface edges. Since intersecting Van der Waals balls may not yield interface edges, it might be interesting to explore generalizations of our constructions with higher-order Voronoi diagrams so as to account for these contacts.

The Kinetic setting. Our focus has been on static interfaces defined from pdb files. An appealing extension would be to integrate the calculation to simulation techniques (e.g. molecular dynamics) to see whether important properties can be inferred.

Acknowledgments. We thank Mariette Yvinec for interesting discussions and for polishing the CGAL *Alpha_Shape_3* package.

References

[AJBL03] L. Adamian, R. Jackups, T.A. Binkowski, and J. Liang. Higher-order interhelical spatial interactions in membrane proteins. *J Mol Biol.*, 327(1), 2003.

[aMS92] P. Walls ans M. Sternberg. New algorithm to model protein - protein recognition based on surface complementarity. *J. Mol. Bio.*, 228, 1992.

[AT97] R. Abagyan and M. Totrov. Contact area difference (cad): A robust measure to evaluate accuracy of protein models. *J. Mol. Biol.*, 268, 1997.

[AUVW97] S. Atwell, M. Ultsch, A.M. De Vos, and J.A. Wells. Structural plasticity in a remodeled protein-protein interface. *Science*, 1997.

[BCRJ04] R.P. Bahadur, P. Chakrabarti, F. Rodier, and J. Janin. A dissection of specific and non-specific protein-protein interfaces. *J. Mol. Bio.*, 336, 2004.

[BER04] Y.-E. A. Ban, , H. Edelsbrunner, and J. Rudolph. Interface surfaces for protein-protein complexes. In *RECOMB*, 2004.

[BT98] A.A. Bogan and K.S. Thorn. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.*, 280, 1998.

[BY98] Jean-Daniel Boissonnat and Mariette Yvinec. *Algorithmic Geometry*. Cambridge University Press, UK, 1998. Translated by Hervé Brönnimann.

[CJ99] L. Lo Conte and J. Janin. The atomic structure of protein-protein recognition sites. *J. Mol. Bio.*, 185, 1999.

[CJ02] P. Chakrabarti and J. Janin. Dissecting protein-protein recognition sites. *Proteins*, 47, 2002.

[CP04] F. Cazals and F. Proust. On the topology and the geometry of (voronoi) molecular interfaces. part i: algorithms. Technical Report 5346, INRIA, October 2004.

[CP05] F. Cazals and F. Proust. Revisiting the description of protein-protein interfaces. part ii: experimental study. Technical Report 5501, INRIA, February 2005.

[CSM03] David Cohen-Steiner and Jean-Marie Morvan. Restricted delaunay triangulations and normal cycle. In *Proc. 19th Annu. ACM Sympos. Comput. Geom.*, pages 237–246, 2003.

[CW95] T. Clackson and J. Wells. A hot spot of binding energy in a hormone-receptor interface. *Science*, 267, 1995.

[Ede92] H. Edelsbrunner. Weighted alpha shapes. Technical Report UIUCDCS-R-92-1760, Dept. Comput. Sci., Univ. Illinois, Urbana, IL, 1992.

[EFL98] H. Edelsbrunner, M. Facello, and J. Liang. On the definition and the construction of pockets in macromolecules. *Discrete Appl. Math.*, 88:83–102, 1998.

[GW96] R.R. Gabdoulline and R.C. Wade. Analytically defined surfaces to analyze molecular interactions proeprties. *J. Molecular Graphics*, 14, 1996.

[HO94] D. Halperin and M. H. Overmars. Spheres, molecules, and hidden surface removal. In *Proc. 10th Annu. ACM Sympos. Comput. Geom.*, pages 113–122, 1994.

[HS98] Dan Halperin and Christian R. Shelton. A perturbation scheme for spherical arrangements with application to molecular modeling. *Comput. Geom. Theory Appl.*, 10:273–287, 1998.

[Las95] R. Laskowski. Surfnet: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.*, 13(5), 1995.

[LC93] M. Lawrence and P. Colman. Shape complementarity at protein-protein interfaces. *J. Mol. Bio.*, 234, 1993.

[PGF98] R. Prei"ner, A. Goede, and C. Fr"ommel. Dictionary of interfaces in proteins (dip). data bank of complementary molecular surface patches. *JMB*, 280, 1998.

[PHT00] H. Ponstingl, K. Henrick, and J.M. Thornton. Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins Struct. Funct. Genet.*, 41, 2000.

[Ric77] F. M. Richards. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.*, 6:151–176, 1977.

[San79] L. Santaló. *Integral Probability and Geometric Probability*, volume 1 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1979.

[Tar83] R. E. Tarjan. *Data Structures and Network Algorithms*, volume 44 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1983.

9 Appendix I: A primer on weighted α-shapes

In this section, we provide a quick overview of weighted α-shapes and recall the connexion with molecular surfaces. See [Ede92, EFL98] for more details.

Collections of balls. From a geometric standpoint, we model a molecule by the collection of its Van der Waals balls, and we shall indistinctly use the terms ball or atom. A given ball is denoted $B_i(a_i, r_i)$ and the corresponding sphere $S_i(a_i, r_i)$. For convenience, we shall denote w_i the square radius of a ball. For example, we shall speak of the sphere $S_i(a_i, w_i)$. Finally, we denote r_w the radius of the water probe used to define the SAS surface. For example, an atomic ball enlarged by the water probe corresponds to the ball $B_i(a_i, r_i + r_w)$.

We considering a molecule or a complex, the collection of all Van der Waals balls is denoted $\mathcal{S}$. We assume each atom is identified by its index in the collection, and a subset of atoms is therefore represented by the multi-index $I = \{i_1, \ldots, i_k\}$ corresponding to these atoms. The number of atoms in such a subset, or equivalently the size of the multi-index is denoted $\mid I \mid$.

(Weighted) Voronoi and Delaunay diagram. Let $S_i(a_i, w_i = r_i^2)$ and $S_j(a_j, w_j = r_j^2)$ be two spheres. Their *power distance* is defined by $\pi(S_i, S_i) = a_i a_j^2 - w_i - w_j$. The two spheres are called orthogonal if $\pi(S_i, S_j) = 0$, and conflict free if $\pi(S_i, S_j) > 0$. The *power* of a point p wrt a sphere S_i is defined by $\pi(p, S_i) = a_i p^2 - w_i$. The sets of points having equal power wrt two spheres is called the radical flat of the spheres. Notice that the power of a point wrt a sphere is just the equation of the sphere, so that point p is inside / on / outside sphere S precisely when $\pi(p, S) < 0, \pi(p, S) = 0, \pi(p, S) > 0$.

The Weighted Voronoi diagram $V(P)$ —also known as the power diagram— of $\mathcal{S}$ is a cell decomposition of $\mathbb{R}^3$ in convex polyhedra. Every *Voronoi cell* corresponds to exactly one ball and contains all points of $\mathbb{R}^3$ that do not have a smaller power distance to any other ball, i.e. the Voronoi cell corresponding to a ball $S_i \in \mathcal{S}$ is defined by $V_i = \{p \in \mathbb{R}^3 : \forall j \in \mathcal{S} \quad \pi(p, S_i) \leq \pi(p, S_j)\}$. Closed facets shared by two Voronoi cells are called *Voronoi facets*, closed edges shared by three Voronoi cells are called *Voronoi edges*, and the points shared by four Voronoi cells are called *Voronoi vertices*.

The Delaunay triangulation $D(P)$ of P is the dual of the Voronoi diagram, in the following sense. Whenever a collection $I = \{i_1, \ldots, i_k\}$ of Voronoi cells corresponding to balls $S_{i_1}, \ldots, S_{i_k}$ have a non-empty intersection, the simplex whose vertices are $a_1, \ldots, a_k$ belongs to the Delaunay triangulation. Under the hypothesis of general position, it is a simplicial complex that decomposes the convex hull of the centers of the balls. This simplicial complex contains vertices (the centers of the balls), edges, triangles and tetrahedra. See Figure 12 for a two-dimensional example of a weighted Voronoi and Delaunay diagrams.

The following observations are straightforward:

1. If all the balls have the same radius, the radical flat defined as the set of points p such that $\pi(p, S_i) = \pi(p, S_j)$ is the hyper-plane bisecting the line-segment $a_i a_j$, so

that the weighted diagrams just introduced reduce to the usual Euclidean Voronoi and Delaunay diagrams. Although atoms of different types have different radii, from now on, we shall therefore skip the adjective weighted.

2. A Voronoi cell (of any dimension, but considered as relatively open) consists of centers of spheres which are orthogonal to the balls centered at the vertices of its dual simplex, and are conflict free wrt the other spheres.

3. If the square radii of all the spheres are shifted by the same amount, the radical flats remain the same, and so do the Voronoi and Delaunay diagrams.

4. If a portion of a ball extends into the Voronoi region of another ball, it is contained in the ball associated to that region.

α-complex and α-shapes. As observed above, the Voronoi and Delaunay diagrams remain the same if the radii of the balls are shifted by the same amount. Let us consider the collection of balls $\{S_i(a_i, w_i+\alpha\}$ for α growing from $-\infty$ to ∞. As soon as $w_i+\alpha \geq 0$, define $R_i = B_i \cap V_i$, that is R_i is the portion of the growing ball contained within its Voronoi region. Consider now a collection of balls I. Similarly to the Delaunay diagram, the α-complex is the simplicial complex containing a simplex spanning the vertices of the balls in I as soon as $\cap_{i\in I} R_i \neq \emptyset$. The geometric domain spanned by the simplices of the α-complex is called the α-shape —see Fig. 13 for an illustration.

Relationship to molecular surfaces It is well known [Ede92, EFL98] that the boundary of a union of balls is encoded by the boundary of the corresponding α-shape for $\alpha = 0$. For example, the Van der Waals surface just reads from the α-shape of the balls $S_i(a_i, r_i)$, with $\alpha = 0$. Similarly, the SAS surface is described by the boundary of the union of balls $S_i(a_i, r_i + r_w)$ —the Van der Waals balls expanded by the water probe.

Finally, we make the following straightforward observation, which will be used when discussing our definition of interfaces:

Observation. 8 *Filtering the α-complex of the balls $S_i(a_i, r_i + r_w)$ with $\alpha = 0$ consists of selecting orthogonal spheres whose centers are at distance $r_i + r_w$ of the atoms' centers. Filtering the α-complex of the balls $S_i(a_i, r_i)$ with $\alpha = r_w^2$ consists of selecting orthogonal spheres whose centers are at distance $\sqrt{r_i^2 + r_w^2}$ of the atoms' centers.*

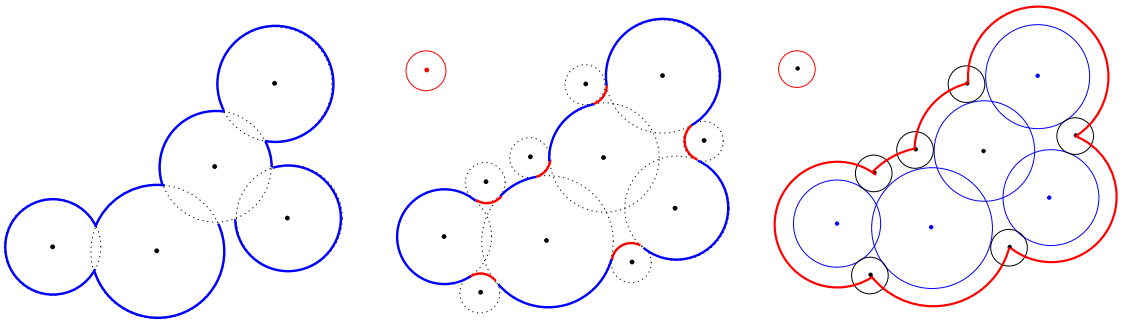

Figure 10: Molecular surfaces: (a)Van der Waals surface (b)Molecular (Connolly) surface (c)Solvent Accessible Surface

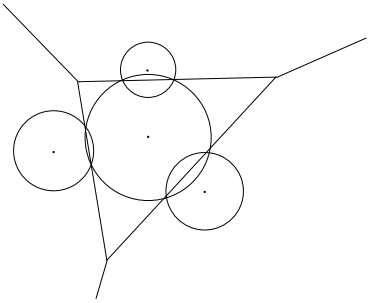

Figure 11: (a)Radical axis (b)(Weighted Voronoi and Delaunay diagrams)

Figure 12: Weighted Voronoi and Delaunay diagrams

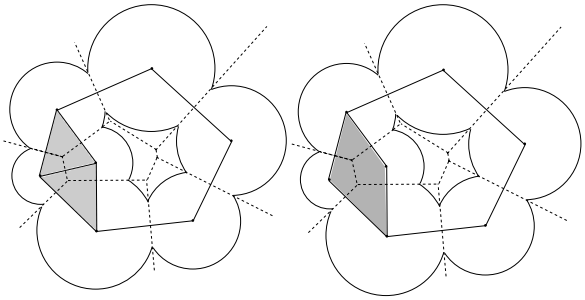

Figure 13: (a)α-complex (b)α-shape

10 Appendix II: complementary material

10.1 Proof of Observation 1

The following proves that interface neighbors exactly identify the atoms loosing accessibility:

Proof.[Observation 1.]
An interface edge witnesses two intersecting spheres of different colors, which implies the corresponding atoms loose contact to the solvent.

To prove that an atom S_i loosing accessibility has at least the ball of one interface neighbor, let $p \in SASL_i$. We shall prove that p belongs to at least one interface neighbor of S_i. Denote a_j the center of a sphere S_j, and denote $F_{j,k}$ the Voronoi facet of two incident Voronoi regions V_j and V_k. Assume no sphere center lies on the radical flat of any other two balls —a generic condition that can be met using symbolic perturbations. Since $p \in SASL_i$, point p is contained in at least one ball of the second color. Denote $S_{i_k}, S_{i_{k-1}}, \ldots, S_{i_1} = S_i$ the balls attached to the Voronoi cells traversed when moving from p to a_i along the line-segment pa_i. Since —by the previous generic assumption— the line-segment pa_i intersects each Voronoi facet transversally, this sequence is well defined. For $j = k, \ldots, 2$, the intersection point between pa_i and $F_{i_j,i_{j-1}}$ witnesses the fact that point p belongs to the half-space of points whose power wrt S_{i_j} is less than wrt $S_{i_{j-1}}$. Therefore

$$\pi(p, S_{i_k}) < \pi(p, S_{i_{k-1}}) < \cdots < \pi(p, S_1) = 0, \tag{7}$$

which also means that p belongs to every ball $S_{i_j}, j = k, \ldots, 2$, and in particular to S_{i_2} which is a neighbor of S_i in the Voronoi diagram. □

10.2 Interface neighbors and intersecting balls

To get more insights into the relationships between a ball S_i and its neighbors, call a ball *principal* if removing it from the complex results in SAS recovery for S_i, and call it *redundant* otherwise. One can observe that:

- intersecting balls might not correspond to an edge in the Delaunay triangulation —S_1 and S_3 on Fig. 14.
- an intersecting pair having an edge in the Delaunay triangulation might not correspond to an edge in the α-complex —S_1 and S_4 on Fig. 15. However, it can be checked that the corresponding pair corresponds to a cell in a higher-order Voronoi diagram. To see which, write down the power of an intersection point between these two balls wrt all the other balls.
- a principal ball always corresponds to an interface neighbor, but a given sphere may not have any principal neighbor —Figs. 16and 17.

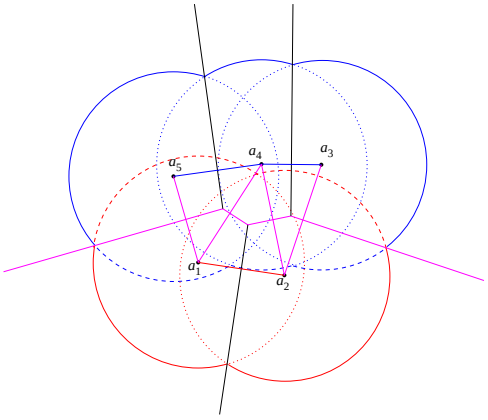

Figure 14: Intersecting balls not connected in Delaunay: S_1 and S_3

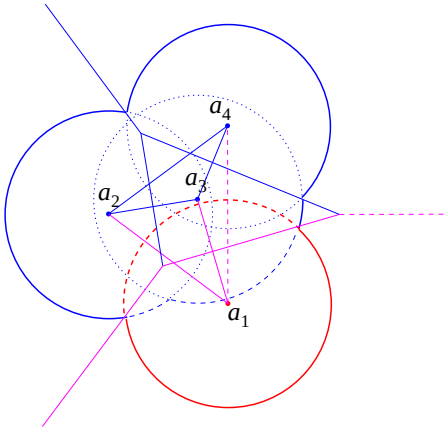

Figure 15: Intersecting balls. Delaunay edge, no interface edge: S_1 and S_4

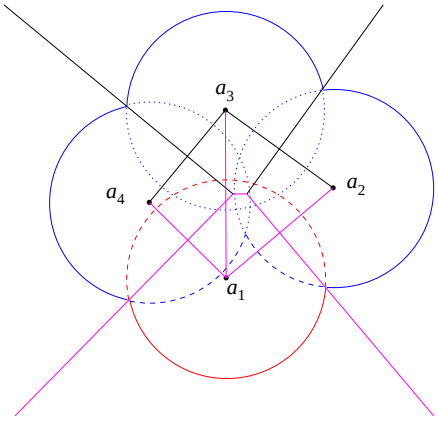

Figure 16: An interface neighbor which is not principal: S_3 for S_1

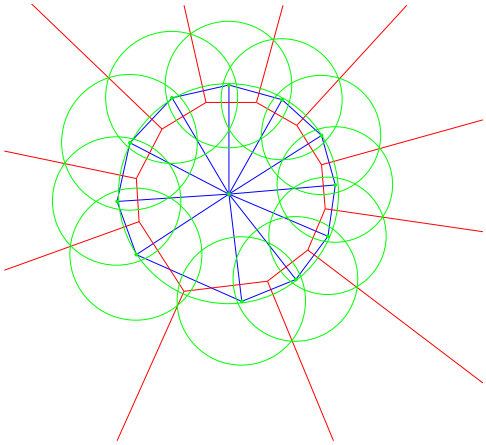

Figure 17: The central ball does not have any principal neighbor

10.3 Configurations of tetrahedra for bicolor and tricolor interfaces

In this section, we list the configurations of tetrahedra contributing to bicolor and tricolor interfaces. When studying the interface XY (or ABW), the relevant edges are marked with

the sign // on the figures. Notice that such an edge does not contribute to the interface if not present in the α-complex of the balls $B_i(a_i, r_i + r_w)$, with $\alpha = 0$.

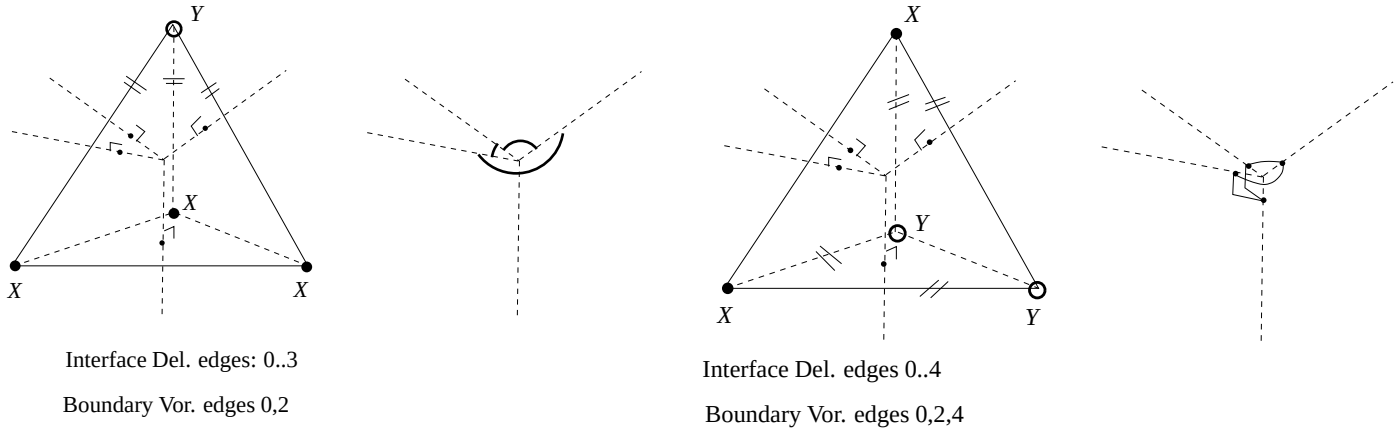

Figure 18: Bicolor interface (a)Bicolor tetrahedron 1-3 (b)Bicolor tetrahedron 2-2

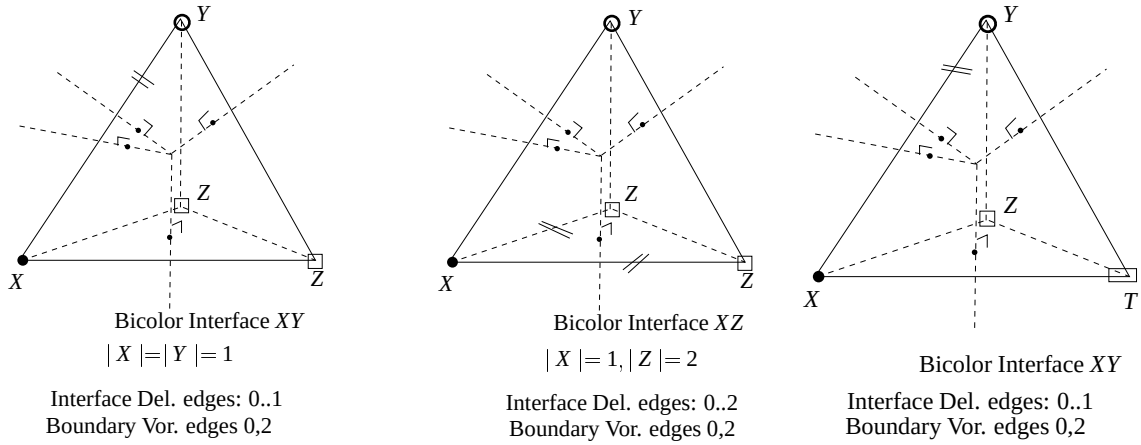

Figure 19: Bicolor interface (a)tricolor tetrahedron $(1, 1, 2)$ (b)quadricolor tetrahedron $(1, 1, 1, 1)$

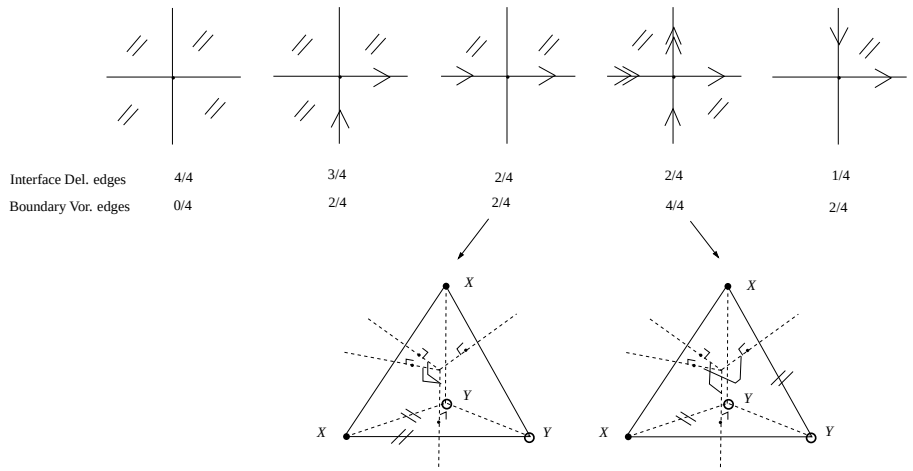

Figure 20: Neighborhood of the Voronoi vertex of a bicolor 2-2 tetrahedron as a function of the number of bipolar edges selected. Arrows materialize the boundaries

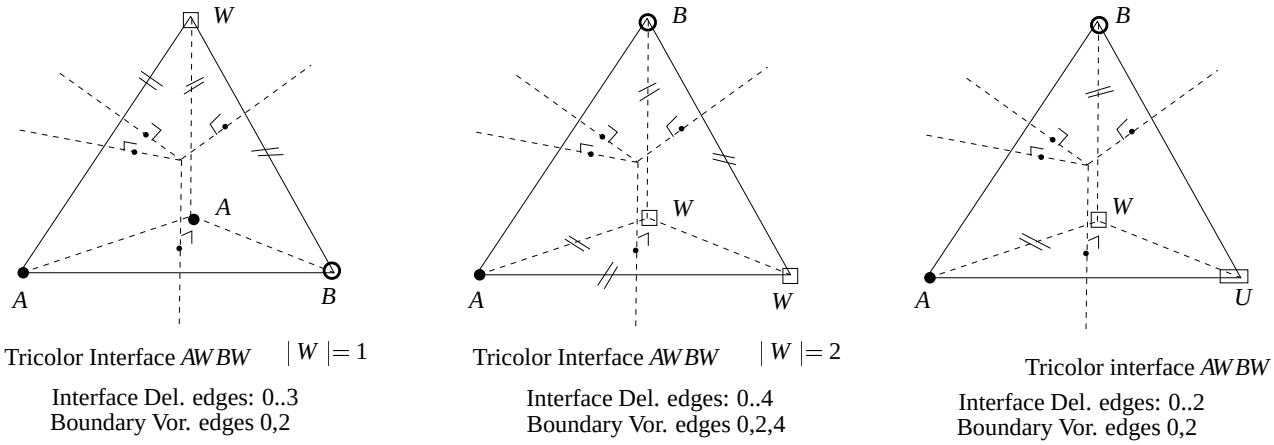

Figure 21: Tricolor interface (a)tricolor tetrahedron $(1,1,2)$ —two cases (b)quadricolor tetrahedron $(1,1,1,1)$

10.4 Proof of observation 4

The following establishes the topological properties of the bicolor interface:

Proof.[Observation 4.]
The first claim stems from the fact that the interface is a subset of the Voronoi diagram. For the second one, consider a Voronoi edge bounding an XY Voronoi facet. Such an edge is the dual of a bicolor triangle which has at most two edges of type XY. Since the three edges of the triangle cannot have the same type XY, the Voronoi edge cannot be non-manifold.
The proof of the third claim follows from an easy analysis of the configurations a given interface Delaunay edge of type XY is involved with. For convenience, we detail the case of bicolor tetrahedra and refer the reader to section 10.3 for the other cases.
Bicolor tetrahedra —types $(1,3)$ and $(2,2)$. To analyze the neighborhood of a Voronoi vertex, there are two cases to be analyzed since a bicolor tetrahedron can either be $(1,3)$ or $(2,2)$. The first case is trivial and yields a topological disk or a half-topological disk. Consider a $(2,2)$ tetrahedron —see Fig. 20. The cases where 4,3 or 1 bicolor edges belong to the α-complex are trivial. But for 2 edges only do so, we need to distinguish between two cases: the two Delaunay edges belong to the same triangle, or span the four vertices of the tetrahedron. In the first case, the contribution of the tetrahedron to the interface is a region with boundary —with two boundary Voronoi edges. In the second one, the contribution consists of two regions meeting a the Voronoi vertex —which we may called a pinched Voronoi vertex.
Tricolor and Quadricolor tetrahedra. See section 10.3. □

10.5 Proof of Observation 6.

Proof. By Observation 1, the interface neighbors account for $SASL_i$. Moreover, exchanging the sigmas from the previous definitions yields:

$$\sum_{j\in \mathrm{N}(i)} \mu^{\flat}_{i\to j} = \sum_{j\in \mathrm{N}(i)} \sum_{I\in\pi(i), j\in I} \frac{\mu(C_{i;\ I})}{\mid I \mid} = \sum_{I\in\pi(i)} \sum_{j\in \mathrm{N}(i), j\in I} \frac{\mu(C_{i;\ I})}{\mid I \mid} = \sum_{I\in\pi(i)} \mu(C_{i;\ I}) = SASAL_i. \tag{8}$$

The second claim follows directly from the first one. □

Contents

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399