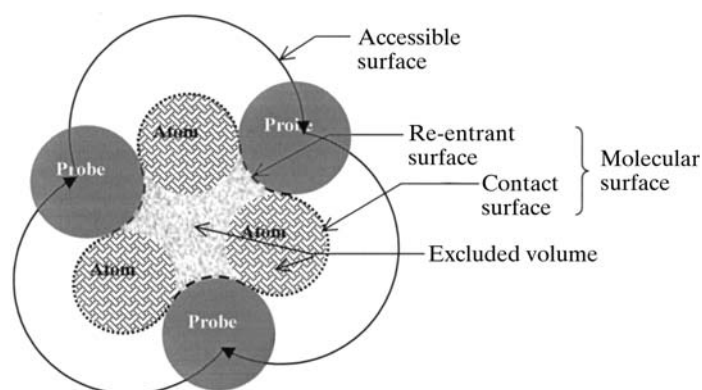22. MOLECULAR GEOMETRY AND FEATURES

Fig. 22.1.2.1. Surfaces in a plane cut through a hypothetical molecule. The molecular surface consists of the sum of the atomic surfaces that can be contacted by solvent molecules and the surface of the space between atoms from which solvent molecules are excluded. The solvent-accessible surface is the surface formed by the set of the centres of spheres that are in closest contact with the molecular surface.

surface area, also known as the solvent-excluding surface, is the outer surface of the volume from which solvent molecules are excluded. Strictly, this would depend on the orientation of non-spherically symmetric solvents such as water. However, since hydrogen atoms are smaller than oxygen atoms, for current purposes it is sufficient to consider water as a sphere with a radius of 1.4 to 1.7 Å, approximating the 'average' distance from the centre of the oxygen atom to the van der Waals surface of water. The practical definition of the molecular surface is, then, the area of the volume excluded to a spherical probe of 1.4 to 1.7 Å radius.

As an aside, it is important to note that surface-area calculations depend on inexact parameterization. For example, there is no radius of any hard-sphere model that can give a realistic representation of the solvent. Furthermore, the choice of van der Waals radii can depend on whether the distance of zero or minimum potential energy is estimated and the potential-energy function or experimental data used. (Tables of common values are given by Gerstein & Richards in Section 22.1.1.) Thus, calculations of molecular and accessible surfaces are approximate. However, when the errors are averaged over large areas of a macromolecule, the numbers can be precise enough to give important insights into function.

Fig. 22.1.2.1 shows that the molecular surface consists of two components. The contact surface is part of the van der Waals surface. The re-entrant surface encloses the interstitial volume and has components that are the exterior surfaces of atoms (contact surface) and parts of the surfaces of probes placed in positions where they are in contact with van der Waals surfaces of two or more atoms (re-entrant surface).

The occluded molecular surface is an approximate complement to the solvent-accessible surface. It is the part of the surface that would be inaccessible to solvent because of steric conflict with neighbouring macromolecular atoms. It is an approximation in that current calculations use van der Waals surfaces, ignoring the differences between atomic and re-entrant surfaces (see below), and the volume of the probe is not fully accounted for (Pattabiraman *et al.*, 1995). Occluded area is defined as the atomic area whose normals cannot be extended 2.8 Å (the presumptive diameter of a water molecule) without intersecting the van der Waals volume of another atom. This crude approximation to the surface that is inaccessible to water not only increases the speed of calculation, but enables surface areas to be partitioned between the atoms. It is used primarily to evaluate model protein structures by comparing the fraction of each amino acid's surface area that is occluded with that

calculated for the same residue types in a database of accurate structures.

22.1.2.1.3. *Hydration surface*

Whether graphically displaying a molecule or examining potential docking interactions, it is usually the molecular surface or solvent-accessible surface that is used. However, macromolecules also interact through the small (solvent) molecules that are more or less tightly bound (Gerstein & Lynden-Bell, 1993*c*). There is a gradation of how tightly solvent molecules are bound and how many are bound around different side chains. With dynamics simulations, Gerstein & Lynden-Bell (1993*c*) showed that the second hydration shell was a reasonable, practical 'average' limit to which water atoms should be considered significantly perturbed by the protein. They defined a hydration surface as the surface of this second shell and presented evidence that it approximates the boundary between bound and bulk solvent. They presented calculations that showed that molecules interact significantly when their hydration surfaces interact, and not just when they are close enough for their molecular surfaces to form contacts. It may be computationally impractical to perform the simulations required to calculate the hydration surfaces of many proteins, but this work reminds us that energetically significant interactions occur over a wider area than the commonly computed contact molecular-surface area.

22.1.2.1.4. *Hydrophobicity*

The hydrophobic effect (Kauzmann, 1959; Tanford, 1997) has its origins in unfavourable entropic terms for water molecules immediately surrounding a hydrophobic group. In the bulk solvent, each water molecule can be oriented in a variety of ways with favourable hydrogen bonding. At the interface with a hydrophobic group, hydrogen bonds are possible only in some directions, with some configurations of the water molecules. When a hydrophobic group is embedded in water, the surrounding solvent molecules have a more restricted set of hydrogen-bonding configurations, resulting in an unfavourable entropic term. The magnitude of the entropic term should be proportional to the number of solvent molecules immediately surrounding the hydrophobic group. This integer number can be considered very approximately proportional to the area of the surface made by the centres of the set of possible solvent probes contacting the solute, *i.e.* the *solvent-accessible surface area* (Fig. 22.1.2.1). When large areas are considered, summed over many hydrophobic atoms, the errors of this non-integer approximation are insignificant. It is now common practice to estimate the hydrophobic effect free-energy contribution by multiplying the change in macromolecular surface area by an energy per unit area [~ 80 J mol^{-1} Å^{-2} (Richards, 1985), but see also below].

22.1.2.2. *Calculation of surface area and energies of interaction*

22.1.2.2.1. *Introduction*

The first method to be discussed allows the calculation of an accessible surface. The first method for calculating molecular surface involved raining water down on a model of a macromolecule and constructing a surface by making a net under the spheres in their landing positions (Greer & Bush, 1978). This ignored overhangs and was replaced by the dot surface method. More recently, methods were developed to make polyhedral surfaces of triangles by contouring between lattice points or by delimiting with arcs the spherical and toroidal surfaces and then subdividing the piece-wise quartic molecular surface. The surface is

540