Molecular Biophysics & Biochemistry 447b3 / 747b3

Bioinformatics
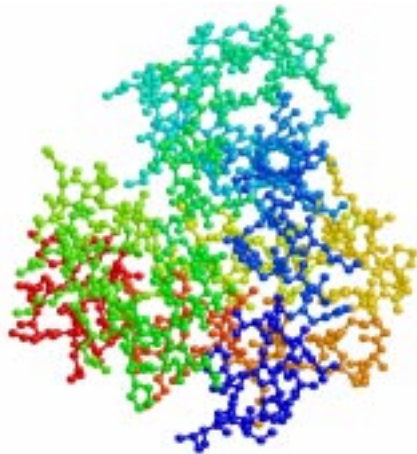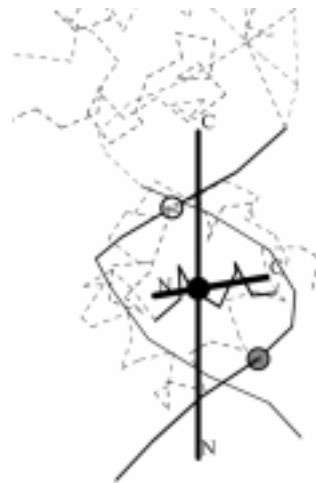
Surfaces & Volumes

Mark Gerstein

Class 6b/7a, 1/28/98
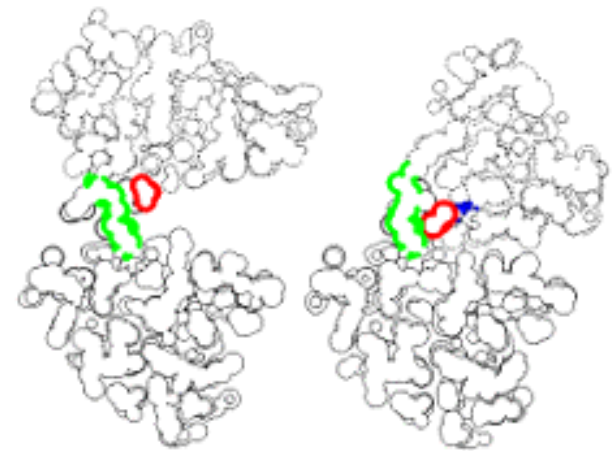
Yale University

Other Aspects of Structure, Besides just Comparing Atom Positions
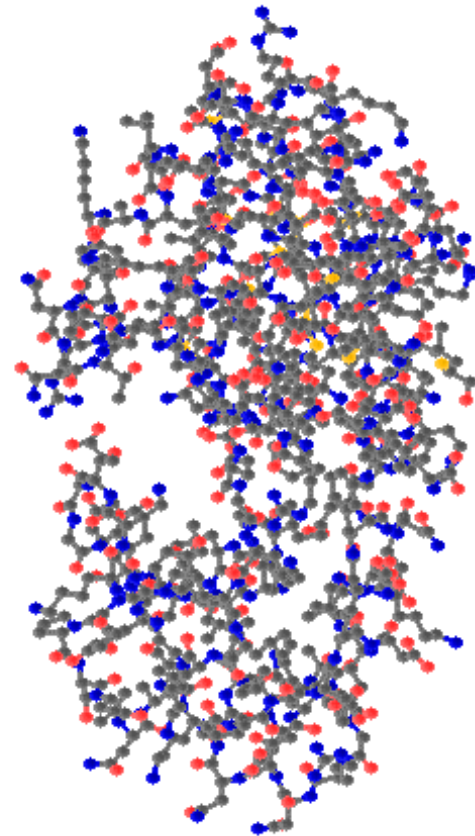
Atom Position,
XYZ triplets

Lines, Axes,
Angles

Surfaces, Volumes

What is Protein Geometry?

- Coordinates (X, Y, Z's)
- Derivative Concepts
  - Distance, Surface Area, Volume, Cavity, Groove, Axes, Angle, &c
- Relation to
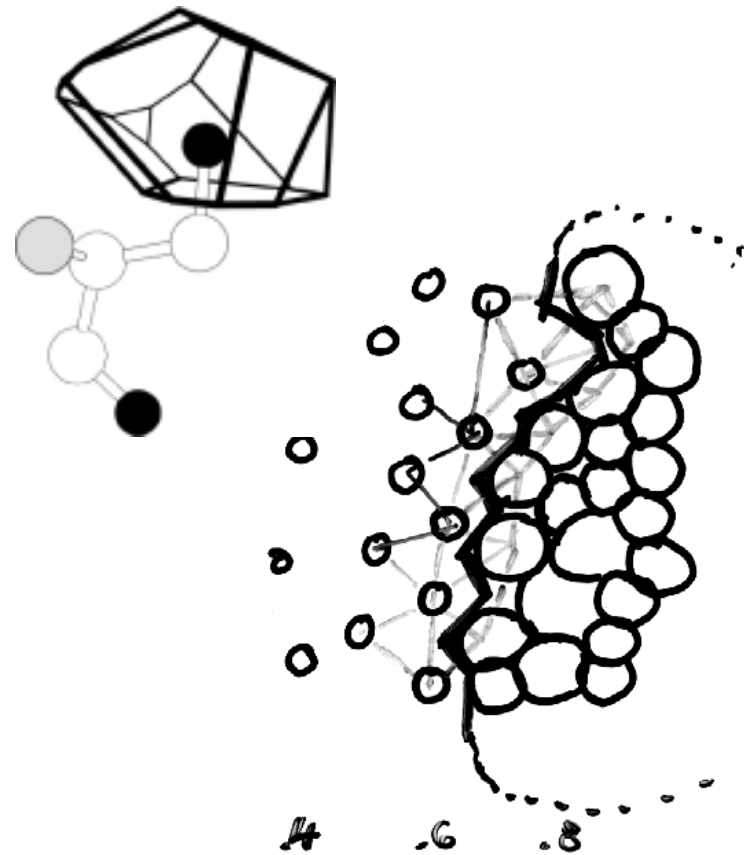 - Function, Energies (E(x)), Dynamics (dx/dt)

3

Surfaces
(slides 1-10 from website)

Packing at Interfaces

- Voronoi volumes (and D. triangulation) to measure packing

- Tight core packing v. Loose surface packing

- Grooves & ridges: close-packing v. H-bonding

- How packing defines a surface (hydration surface)

- Implications for Motions

Classic Papers

- Lee, B. & Richards, F. M. (1971). "The Interpretation of Protein Structures: Estimation of Static Accessibility," *J. Mol. Biol.* **55**, 379-400.

- Richards, F. M. (1974). "The Interpretation of Protein Structures: Total Volume, Group Volume Distributions and Packing Density," *J. Mol. Biol.* **82**, 1-14.

- Richards, F. M. (1977). "Areas, Volumes, Packing, and Protein Structure," Ann. Rev. Biophys. Bioeng. 6, 151-76.

Packing ~ VDW force

- Longer-range isotropic attractive tail provides general cohesion

- Shorter-ranged repulsion determines detailed geometry of interaction

- Billiard Ball model, WCA Theory
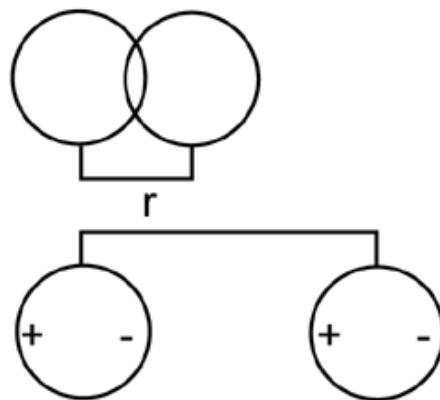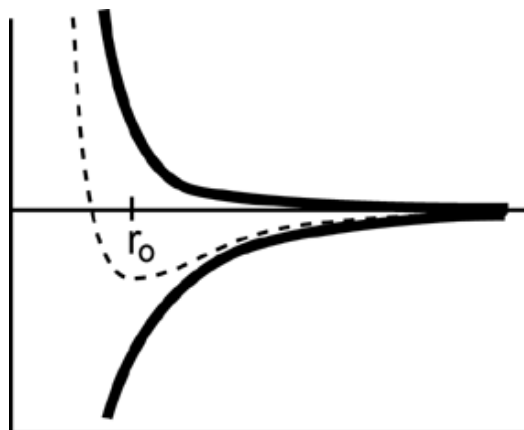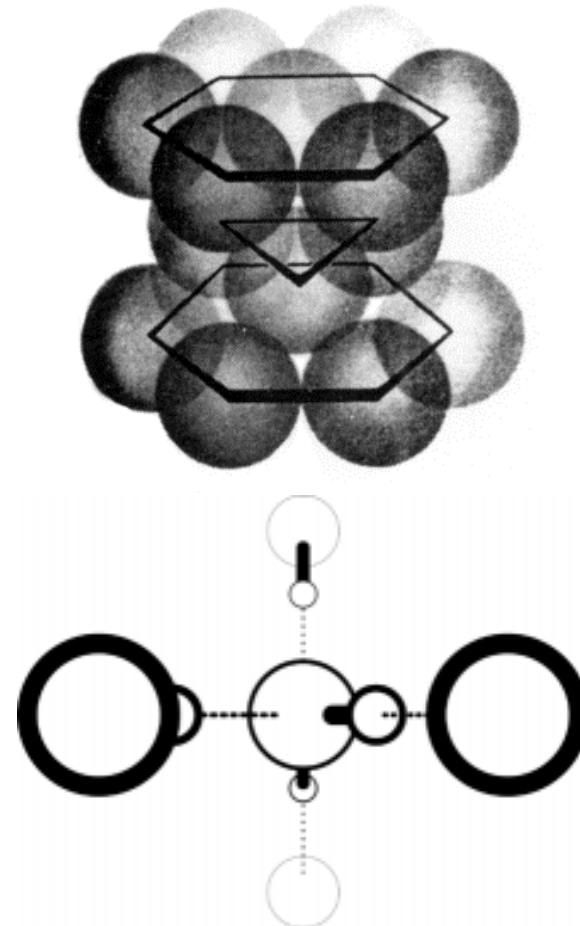

Electron Overlap Replusion

$$U = \varepsilon \left(\frac{r_0}{r} \right)^{12}$$

Dispersion Attraction

$$U = -4\varepsilon \left(\frac{r_0}{r} \right)^{6}$$

Close-packing is Default

- No tight packing when highly directional interactions (such as H-bonds) need to be satisfied

- Packing spheres (.74), hexagonal

- Water (~.35), "Open" tetrahedral, H-bonds

Water v. Argon

9
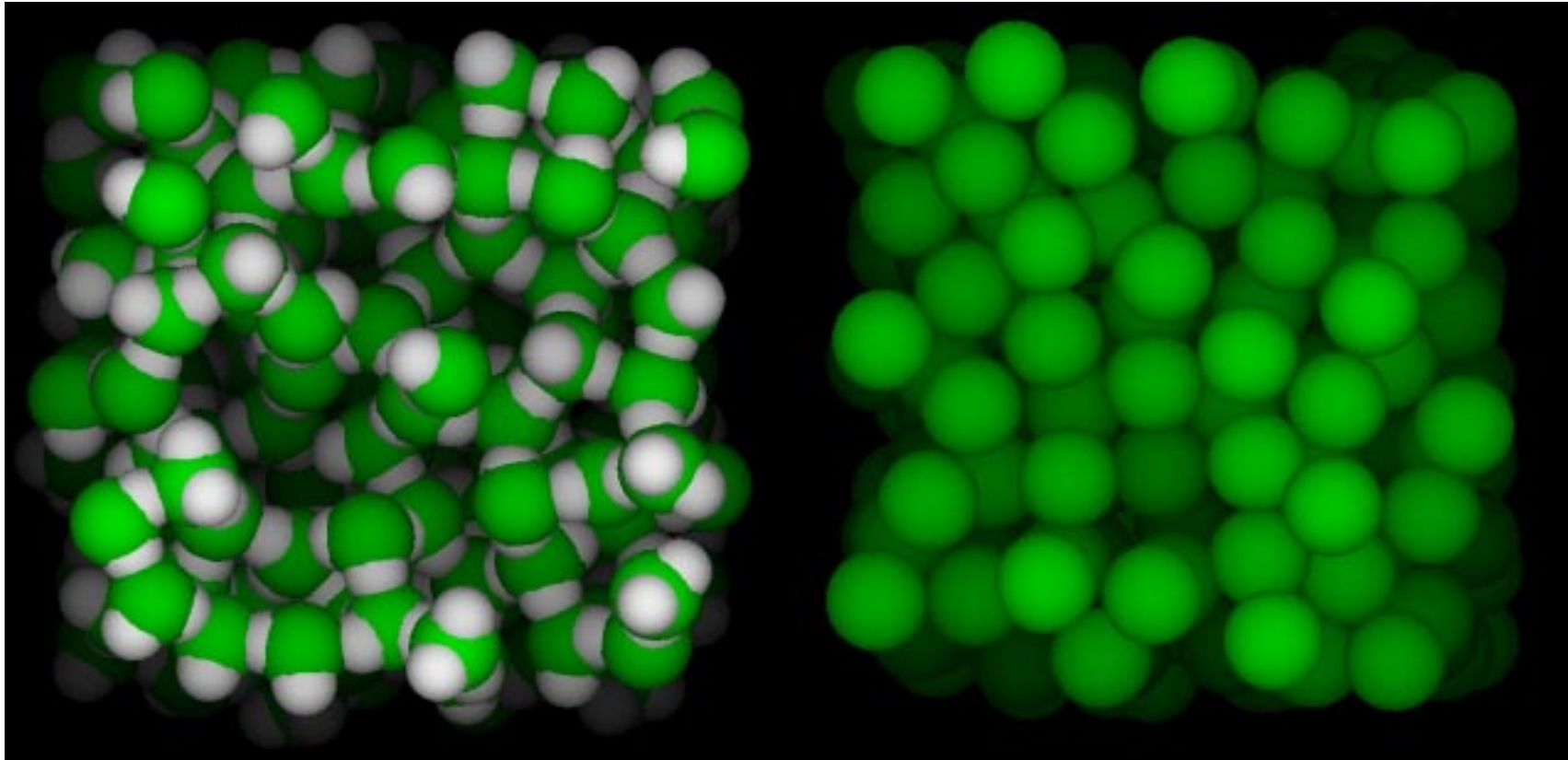
Voronoi Volumes

- Each atom surrounded by a single convex polyhedron and allocated space within it

  - Allocation of all space (large V implies cavities)

- 2 methods of determination

  - Find planes separating atoms, intersection of these is polyhedron

  - Locate vertices, which are equidistant from 4 atoms

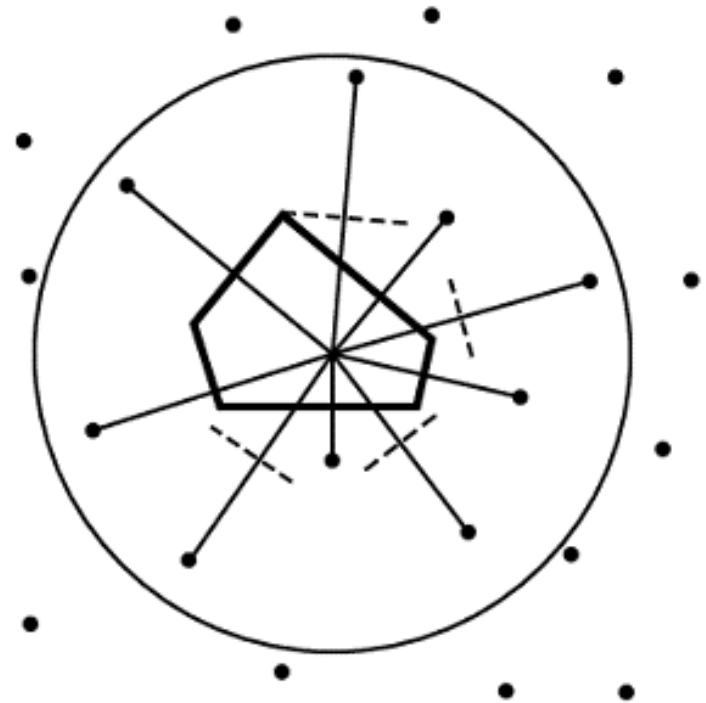# Calclating Volumes with Voronoi polyhedra

- In 1908 Voronoi found a way of partitioning all space amongst a collection of points using specially constructed polyhedra. Here we refer to a collection of "atom centers" rather than "points."

- In 3D, each atom is surrounded by a unique limiting polyhedron such that all points within an atom's polyhedron are closer to this atom than all other atoms.

- Likewise, points equidistant from 2 atoms form planes (lines in 2D). Those equidistant from 3 atoms form lines, and those equidistant form 4 centers form vertices.

Voronoi Volumes, the Natural Way to Measure Packing

Packing Efficiency

    `= Volume-of-Object`

    `------------------`

    `Space-it-occupies`

    = V(VDW) / V(Voronoi)

- Absolute v relative eff.

  V1 / V2

- Other methods

  – Measure Cavity Volume (grids, constructions, &c)
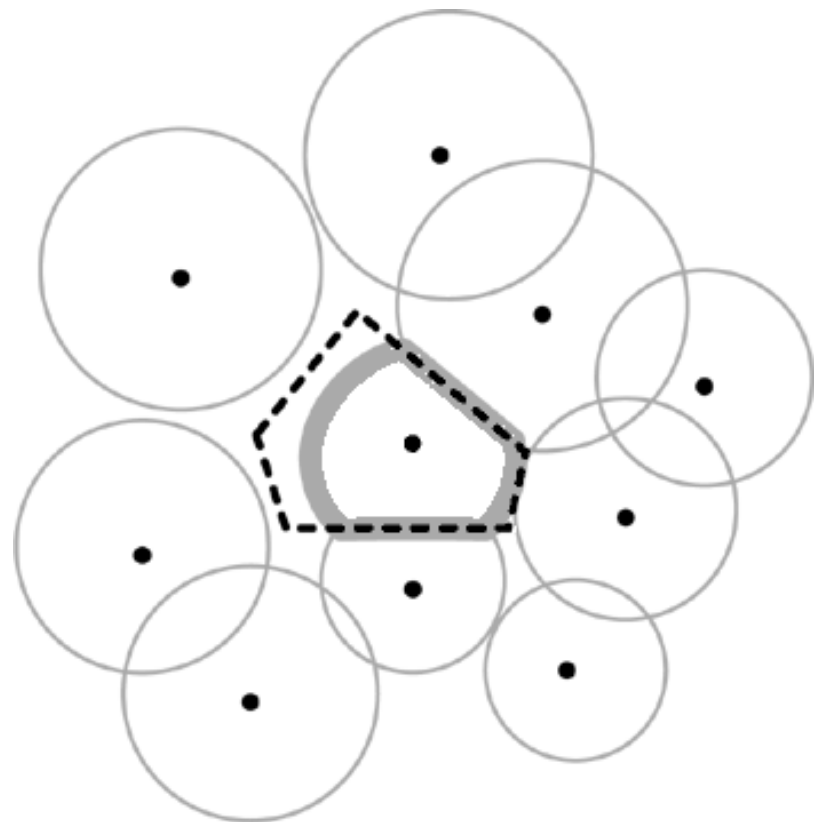
Delauney Triangulation, the Natural Way to Define Packing Neighbors

- Related to Voronoi polyhedra (dual)
- What "coordination number" does an atom have? Doesn't depend on distance
- alpha shape
- threading

Properties of Voronoi Polyhedra
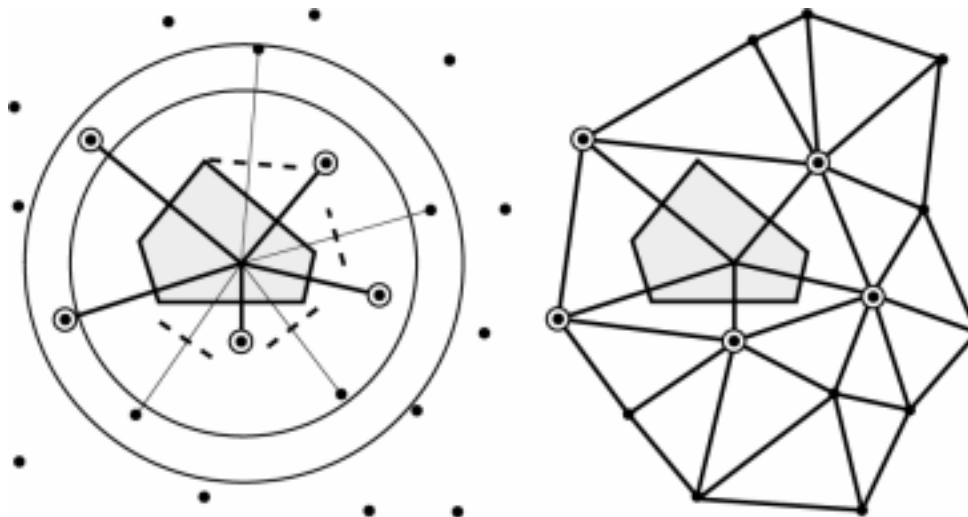
- If Voronoi polyhedra are constructed around atoms in a periodic system, such as in a crystal, all the volume in the unit cell will be apportioned to the atoms. There will be no gaps or cavities as there would be if one, for instance, simply drew spheres around the atoms.

- Voronoi volume of an atom is a weighted average of distances to all its neighbors, where the weighting factor is the contact area with the neighbor.

Voronoi diagrams are generally useful, beyond proteins

- Border of D.T. is Convex Hull

- D.T. produces "fatest" possible triangles which makes it convenient for things such as finite element analysis.

- Nearest neighbor problems. The nearest neighbor of a query point in center of the Voronoi diagram in which it resides

- Largest empty circle in a collection of points has center at a Voronoi vertex

- Voronoi volume of "something" often is a useful weighting factor. This fact can be used, for instance, to weight sequences in alignment to correct for over or under-representation

Volume Calc Methods
(slides 20-40 from website)

Atoms have different sizes

- Difficulty with Voronoi Meth. Not all atoms created equal

- Solutions
  - Bisection -- plane midway between atoms
  - Method B (Richards) Positions the dividing plane according to ratio
  - Radical Plane

- VDW Radii Set
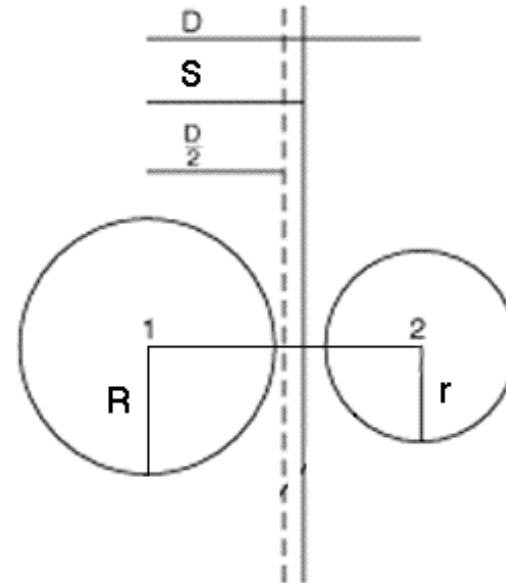
Set of VDW Radii

- Great differences in a sensitive parameter (Radii for carbon 1.87 vs 2.00)

- Complex calculation: minimizing SD, iterative procedure, from protein structures

- Look for common distances in CCD

- <u>Preliminary</u> Solution

Atom	Bondi	New
C4__	1.87	1.88
C3H1	1.76	1.76
C3H0	1.76	1.61
O1HO	1.40	1.42
O2H1	1.40	1.46
N___	1.65	1.64
S___	1.85	1.77

Standard Residue Volumes

- Database of many hi-res structures (~100, 2 Å)

- Volumes statistics for buried residues
  (various selections, resample, &c)

- Standard atomic volumes harder…
  parameter set development...

G	64	c	105	T	120	V	139	H	159	M	168	R	194
A	90	C	113	P	124	E	140	L	165	K	170	Y	198
S	94	D	117	N	128	N	150	I	165	F	193	W	233

Standard Core Volumes (Prelim.)

Atom Types		Num.	Volume (Å^3)	Error (%)
Mainchain Atoms				
carbonyl carbon (except G)	C	8361	9.2	.08
alpha carbon (except G)	CA	7686	13.4	.09
nitrogen (except P)	N	9042	13.9	.09
carbonyl oxygen	O	7831	15.8	.10
Gly C		811	10.2	.27
Gly CA		522	23.5	.39
Pro N		334	8.6	.39
Sidechain atoms				
trigonal or aromatic carbon	>C=	3026	10.3	.13
aromatic CH (H,F,W,Y)	-CH=	4333	21.1	.14
aliphatic CH	>CH-	3411	14.6	.14
methylene group	-CH2-	5427	23.7	.12
methyl group (A,V,L,I)	-CH3	5273	36.7	.11
hydroxyl oxygen (S,T)	-OH	851	17.2	.36
carbonyl oxygen (N,Q)	=O	272	16.8	.76
carboxyl oxygen (D,E)	-O	517	16.0	.53
2° amine (R,H,W)	-NH-	530	15.6	.53
1° amine or amide (R,N,Q)	-NH2	355	23.4	.52
tetrahedral nitrogen (K)	-NH3	31	20.0	1.40
thioether or disulfide (C,M)	-S-	1242	19.3	1.22

27.8   1.33   20

Small Packing Changes Significant

- Exponential dependence
- Bounded within a range of 0.5 (.8 and .3)
- Many observations in standard volumes gives small error about the mean (SD/sqrt(N))
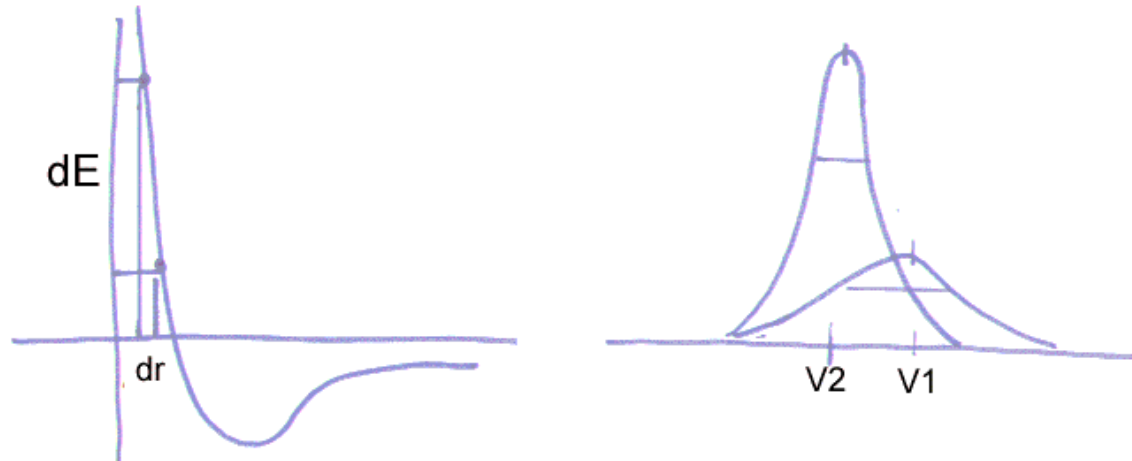
# Packing defines the "Correct Definition" of the Protein Surface

- Voronoi polyhedra are the *Natural* way to study packing!

- How reasonable is a geometric definition of the surface in light of what we know about packing

- The relationship between
  - accessible surface
  - molecular surface
  - Delauney Triangulation (Convex Hull)
  - polyhedra faces
  - hydration surface

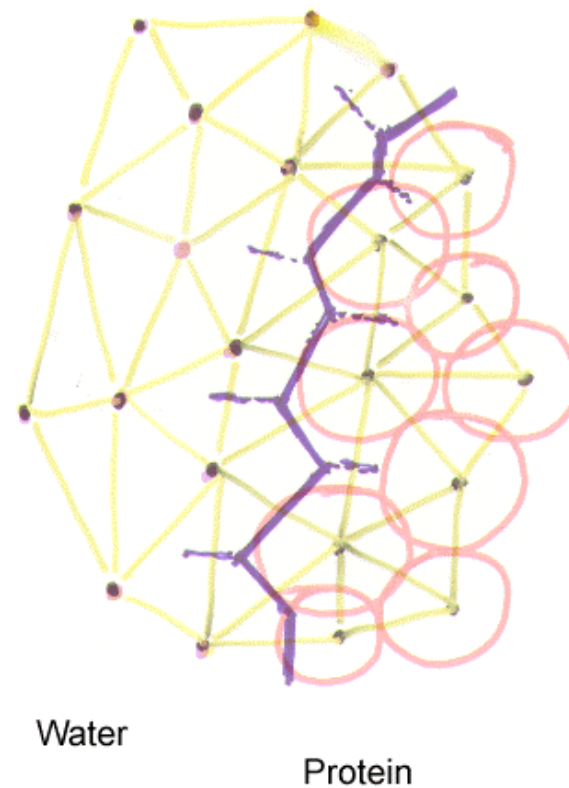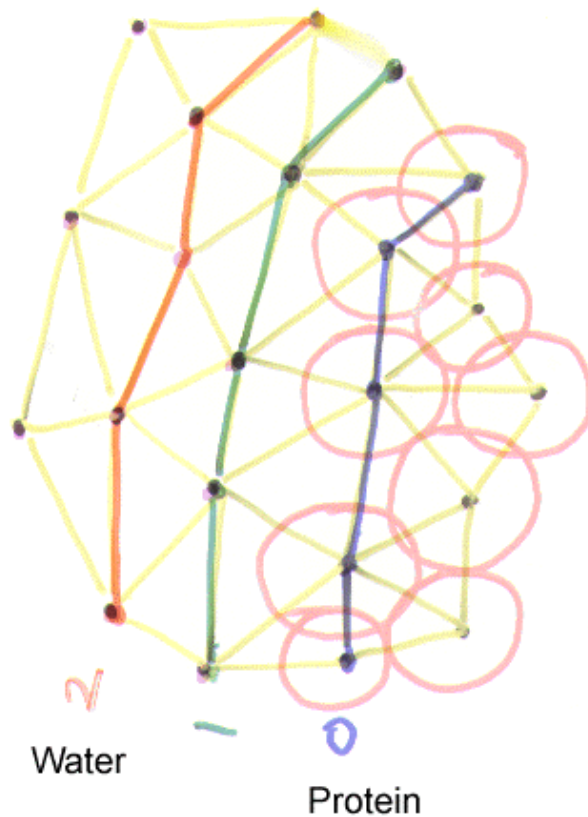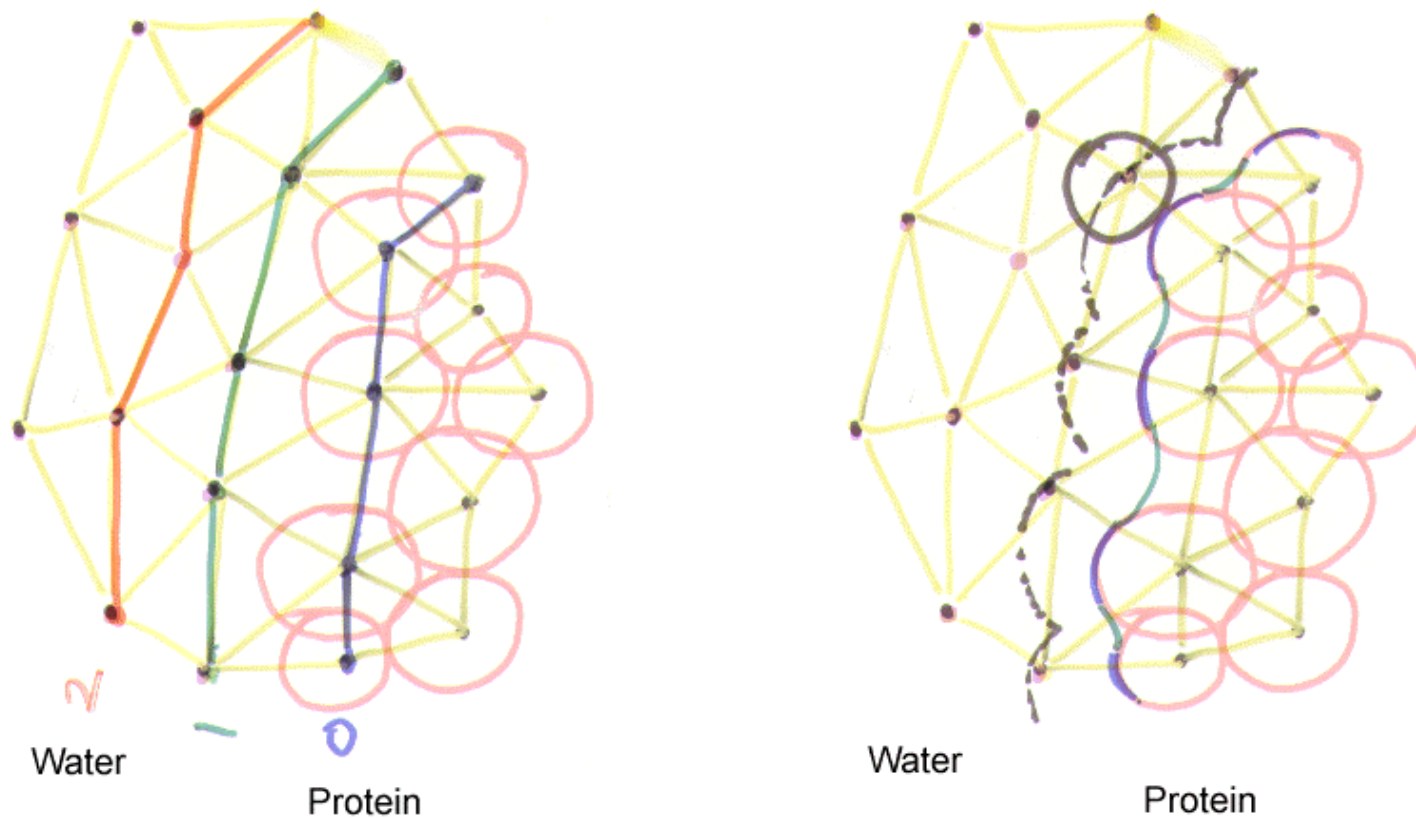# Defining Surfaces from Packing: Convex Hull and Layers of Waters
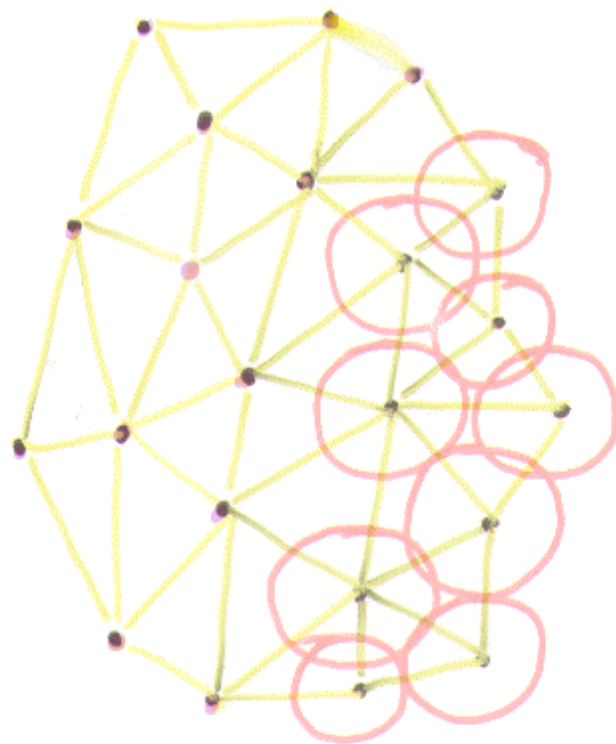
Water

Protein
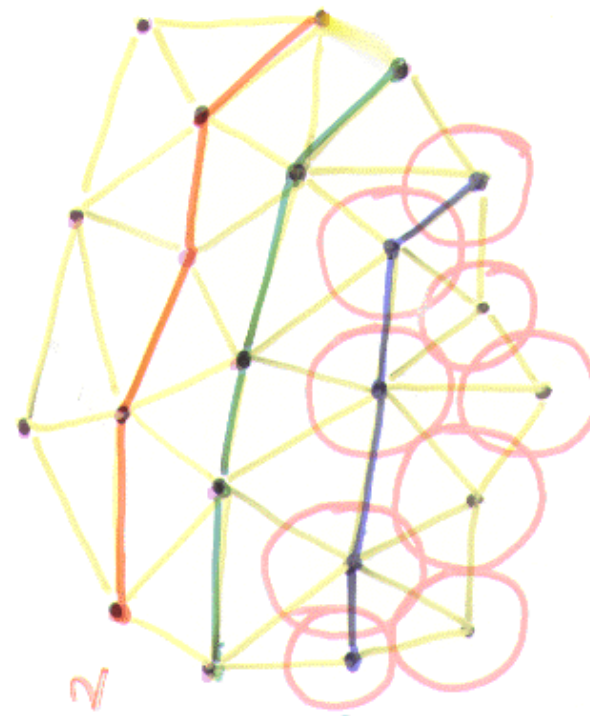
Water

Protein

Defining a Surface from the Faces of Voronoi Polyhedra

Water

Protein

Water

Protein

Accessible Surface
as a Time-averaged Water Layer
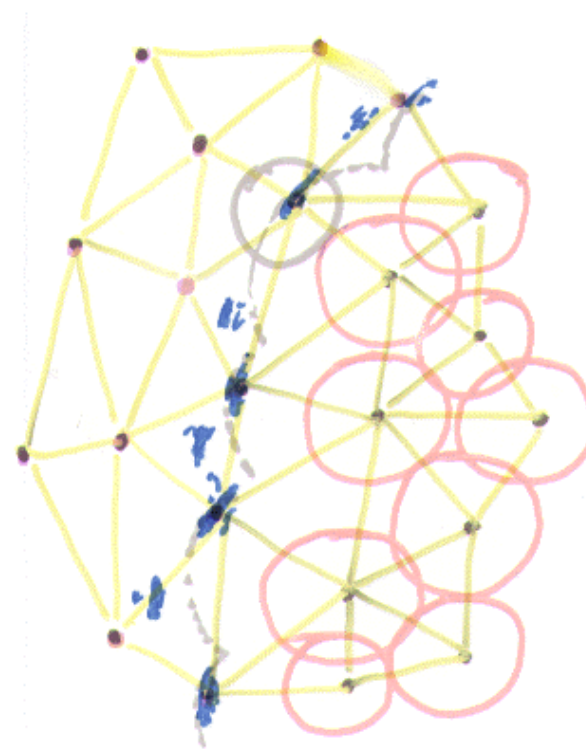

Water

Protein

Water

Protein

The Hydration Surface:
Trying to Model Real Water

Water

Protein

Water

Protein