

Welcome to the SDSU Structural Biology Program!

Wednesday, November 19, 2008

7:00 - 8:00 p.m.

GMCS 305

<http://sci.sdsu.edu/sbp/>

Protein structure determination by X-ray crystallography

- Derive a source of material for study
 - Cloning, expression, purification
- Grow single crystals of subject
 - Crystallization
- Collect x-ray diffraction data
 - X-ray source, experimental design
- X-ray diffraction data processing
 - Convert spots to numbers

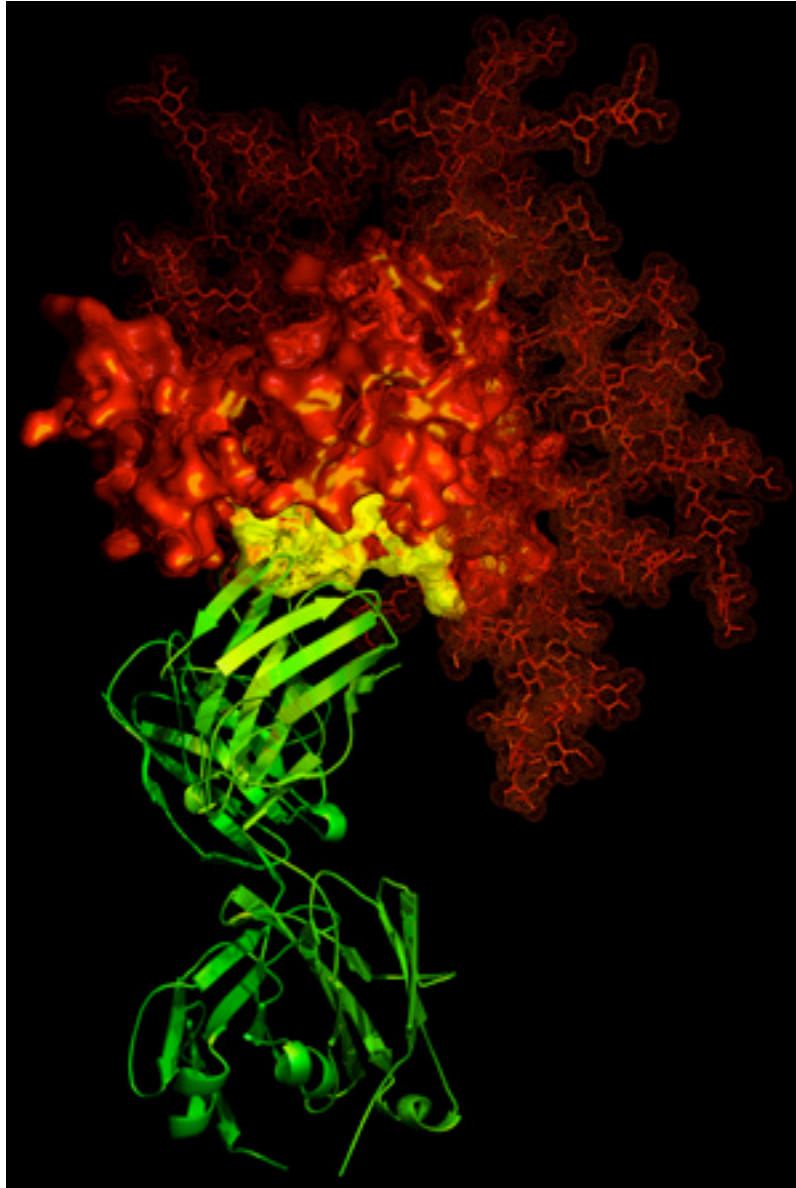
Protein structure determination by X-ray crystallography

- Solve the “phase problem”
 - MIR
 - MR
 - MAD/SAD
- Build and refine model
 - Computer graphics workstation
- Analyze structure
 - Propose and then test structure-based functional hypotheses

Protein structure determination by X-ray crystallography

- Analyze structure
 - Propose and then test structure-based functional hypotheses

Everyone's favorite--Flashy Figures!



**b12 antibody Fab fragment
bound to the CD4 interaction
site of the HIV-1 gp12
surface glycoprotein**

Final output-the atomic coordinate file

	Name of crystallized molecule	pdb code
HEADER	IMMUNE SYSTEM	20-JUL-05 2ADG
TITLE	CRYSTAL STRUCTURE OF MONOCLONAL ANTI-CD4 ANTIBODY Q425	
COMPND	MOL_ID: 1;	
COMPND	2 MOLECULE: Q425 FAB LIGHT CHAIN;	
COMPND	3 CHAIN: A;	Annotation:molecule(s),
COMPND	4 MOL_ID: 2;	species, author(s), reference
COMPND	5 MOLECULE: Q425 FAB HEAVY CHAIN;	citation(s),
COMPND	6 CHAIN: B	
SOURCE	MOL_ID: 1;	
SOURCE	2 ORGANISM_Scientific: MUS MUSCULUS;	
SOURCE	3 ORGANISM_COMMON: MOUSE;	
SOURCE	4 OTHER_DETAILS: HYBRIDOMA;	
SOURCE	5 MOL_ID: 2;	
SOURCE	6 ORGANISM_Scientific: MUS MUSCULUS;	
SOURCE	7 ORGANISM_COMMON: MOUSE;	
SOURCE	8 OTHER_DETAILS: HYBRIDOMA	
KEYWDS	ANTI-CD4, INTERFACIAL METAL, ANTIBODY RECOGNITION	
EXPDTA	X-RAY DIFFRACTION	
AUTHOR	T.ZHOU,D.H.HAMER,W.A.HENDRICKSON,Q.J.SATTENTAU,P.D.KWONG	
REVDAT	2 18-OCT-05 2ADG 1 JRNL	
REVDAT	1 20-SEP-05 2ADG 0	
JRNL	AUTH T.ZHOU,D.H.HAMER,W.A.HENDRICKSON,Q.J.SATTENTAU,	
JRNL	AUTH 2 P.D.KWONG	
JRNL	TITL INTERFACIAL METAL AND ANTIBODY RECOGNITION.	
JRNL	REF PROC.NATL.ACAD.SCI.USA V. 102 14575 2005	
JRNL	REFN ASTM PNASA6 US ISSN 1091-6490	
REMARK	1	
REMARK	2	
REMARK	2 RESOLUTION. 2.50 ANGSTROMS.	

Final output-the atomic coordinate file

Refinement statistics

```
REMARK 3 REFINEMENT.  
REMARK 3 PROGRAM : CNS 1.1  
REMARK 3 AUTHORS : BRUNGER, ADAMS, CLORE, DELANO, GROS, GROSSE-  
REMARK 3 : KUNSTLEVE, JIANG, KUSZEWSKI, NILGES, PANNU,  
REMARK 3 : READ, RICE, SIMONSON, WARREN  
REMARK 3  
REMARK 3 REFINEMENT TARGET : ENGH & HUBER  
REMARK 3  
REMARK 3 DATA USED IN REFINEMENT.  
REMARK 3 RESOLUTION RANGE HIGH (ANGSTROMS) : 2.50  
REMARK 3 RESOLUTION RANGE LOW (ANGSTROMS) : 10.00  
REMARK 3 DATA CUTOFF (SIGMA(F)) : 0.000  
REMARK 3 DATA CUTOFF HIGH (ABS(F)) : 62726.540  
REMARK 3 DATA CUTOFF LOW (ABS(F)) : 0.0000  
REMARK 3 COMPLETENESS (WORKING+TEST) (%) : 96.2  
REMARK 3 NUMBER OF REFLECTIONS : 18319  
REMARK 3  
REMARK 3 FIT TO DATA USED IN REFINEMENT.  
REMARK 3 CROSS-VALIDATION METHOD : THROUGHOUT  
REMARK 3 FREE R VALUE TEST SET SELECTION : RANDOM  
REMARK 3 R VALUE (WORKING SET) : 0.203  
REMARK 3 FREE R VALUE : 0.244  
REMARK 3 FREE R VALUE TEST SET SIZE (%) : 4.900  
REMARK 3 FREE R VALUE TEST SET COUNT : 894  
REMARK 3 ESTIMATED ERROR OF FREE R VALUE : 0.008
```

Resolution range and
data completeness

R-work and R-free

Final output-the atomic coordinate file

REMARK 3 FIT IN THE HIGHEST RESOLUTION BIN.
REMARK 3 TOTAL NUMBER OF BINS USED : 6
REMARK 3 BIN RESOLUTION RANGE HIGH (A) : 2.50
REMARK 3 BIN RESOLUTION RANGE LOW (A) : 2.66
REMARK 3 BIN COMPLETENESS (WORKING+TEST) (%) : 84.60
REMARK 3 REFLECTIONS IN BIN (WORKING SET) : 2562
REMARK 3 BIN R VALUE (WORKING SET) : 0.3370
REMARK 3 BIN FREE R VALUE : 0.3350
REMARK 3 BIN FREE R VALUE TEST SET SIZE (%) : 4.70
REMARK 3 BIN FREE R VALUE TEST SET COUNT : 127
REMARK 3 ESTIMATED ERROR OF BIN FREE R VALUE : 0.030

Refinement statistics for
highest resolution data

REMARK 3
REMARK 3 NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK 3 PROTEIN ATOMS : 3264
REMARK 3 NUCLEIC ACID ATOMS : 0
REMARK 3 HETEROGEN ATOMS : 0
REMARK 3 SOLVENT ATOMS : 67

Number of atoms in final model
*Remember this will affect the
observations/parameters ratio

REMARK 3 B VALUES.
REMARK 3 FROM WILSON PLOT (A**2) : 15.10
REMARK 3 MEAN B VALUE (OVERALL, A**2) : 31.00

REMARK 3
REMARK 3 ESTIMATED COORDINATE ERROR.
REMARK 3 ESD FROM LUZZATI PLOT (A) : 0.32
REMARK 3 ESD FROM SIGMAA (A) : 0.53

Note that these numbers
are much less than 2.50 Å

Final output-the atomic coordinate file

Atom #	AA #	AA type	x,y,z coordinates	B-factor	Element	
ATOM	1	N	GLU A 1	17.329 64.570 43.981	1.00 47.17	N
ATOM	2	CA	GLU A 1	18.786 64.289 44.167	1.00 46.98	C
ATOM	3	C	GLU A 1	19.536 64.291 42.829	1.00 44.67	C
ATOM	4	O	GLU A 1	19.234 65.098 41.944	1.00 46.12	O
ATOM	5	CB	GLU A 1	18.965 62.943 44.880	1.00 49.84	C
ATOM	6	CG	GLU A 1	17.692 62.092 44.974	1.00 52.37	C
ATOM	7	CD	GLU A 1	17.157 61.664 43.613	1.00 53.80	C
ATOM	8	OE1	GLU A 1	16.221 60.834 43.571	1.00 54.59	O
ATOM	9	OE2	GLU A 1	17.667 62.156 42.585	1.00 54.49	O
ATOM	10	N	THR A 2	20.515 63.404 42.675	1.00 40.45	N
ATOM	11	CA	THR A 2	21.263 63.357 41.424	1.00 36.65	C
ATOM	12	C	THR A 2	20.721 62.304 40.467	1.00 35.48	C
ATOM	13	O	THR A 2	20.607 61.127 40.813	1.00 35.96	O
ATOM	14	CB	THR A 2	22.752 63.069 41.654	1.00 34.59	C
ATOM	15	OG1	THR A 2	23.304 64.062 42.525	1.00 33.14	O
ATOM	16	CG2	THR A 2	23.497 63.101 40.335	1.00 32.39	C
ATOM	17	N	THR A 3	20.386 62.741 39.260	1.00 33.43	N
ATOM	18	CA	THR A 3	19.868 61.848 38.240	1.00 30.95	C
ATOM	19	C	THR A 3	20.956 61.553 37.226	1.00 29.87	C
ATOM	20	O	THR A 3	21.722 62.441 36.853	1.00 30.95	O

Chain #

Occupancy

Atom type

Where do we find coordinate files?

The Protein Data Bank

<http://www.rcsb.org>

The screenshot shows the RCSB PDB website homepage. At the top left is the RCSB PDB logo. To the right, it says "A MEMBER OF THE CPDB" and "An Information Portal to Biological Macromolecular Structures". Below this, it states "As of Tuesday Nov 18, 2008 there are 54298 Structures" and "PDB Statistics". A search bar is located below the header, with options for "PDB ID or keyword" and "Author". Navigation links include "CONTACT US", "FEEDBACK", "HELP", and "PRINT".

The main content area features a yellow banner with the message: "Are you missing data updates? The PDB archive has moved to <ftp://ftp.wwpdb.org>. For more information click [here](#)." Below this is a "Welcome to the RCSB PDB" section. It states: "The RCSB PDB provides a variety of tools and resources for studying the structures of biological macromolecules and their relationships to sequence, function, and disease." It also mentions that the RCSB is a member of the wwPDB and offers tools for browsing, searching, and reporting.

On the left side, there is a navigation menu with links such as "Home", "Getting Started", "Download Files", "Deposit and Validate", "Structural Genomics", "Dictionaries & File Formats", "Software Tools", "General Education", "Site Tutorials", "BioSync", "General Information", "Acknowledgments", and "Frequently Asked Questions".

On the right side, there is a "News" section with links for "Complete News", "Newsletter", "Discussion Forum", and "Job Listings". Below this, there is a "View Domain Annotations in 3D" link and a 3D protein structure visualization labeled "SCOP".

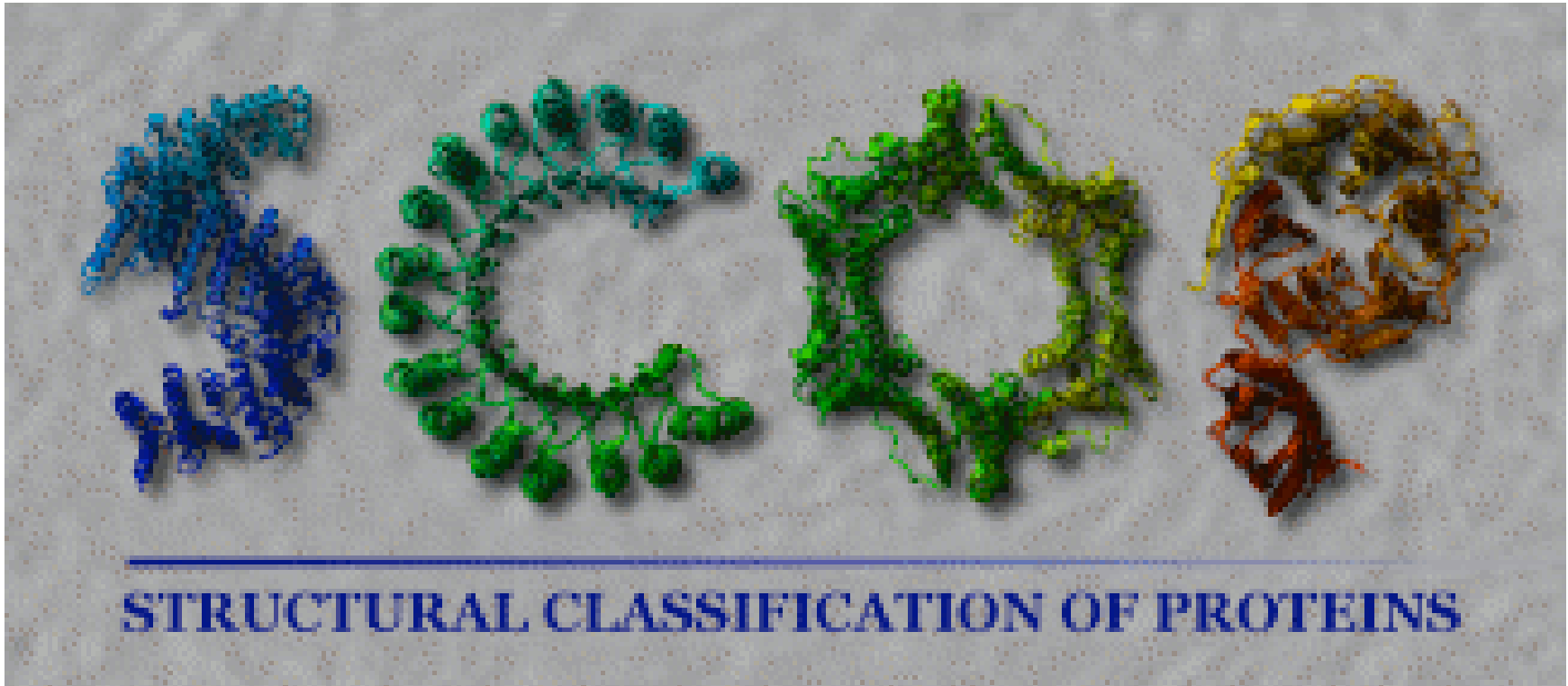
At the bottom, there is a "Quick Tips" box that says: "Do you use www.PDB.org to teach a class? Look at the [Molecule of the Month](#) here." Below this, there is a "Molecule of the Month: Mechanosensitive Channels" section. It includes a 3D protein structure visualization and a text description: "We are remarkably resistant to changes in our surrounding environment. Our bulky bodies allow us to weather extremes of heat and cold, and our skin protects us if we go for a swim in fresh water or salty water. If things get too uncomfortable, we can always get up and walk away, finding a warmer or cooler or drier place. Bacteria don't have as many options. They are tiny and they are immersed in water, so changes in the environment can pose life-threatening challenges. For instance, if it rains they may be suddenly surrounded by fresh water. This is dangerous because the water seeps into the cell through osmosis and increases the pressure inside. At other times, the bacterium may be shifted suddenly to salty conditions, which pulls water out and dehydrates the cell. Bacteria have methods for resisting these changes, so they can keep a steady, comfortable osmotic pressure inside." Below the text are links for "More ..." and "Previous Features".

Some common first observations

- Define the protein fold (SCOP)
- Calculate the root-mean-squared (rms) deviation from a related protein
- Calculate the protein solvent exposed surface area
- Calculate the electrostatic surface potential
- Calculate the hydrophobic surfaces, channel volume, etc.

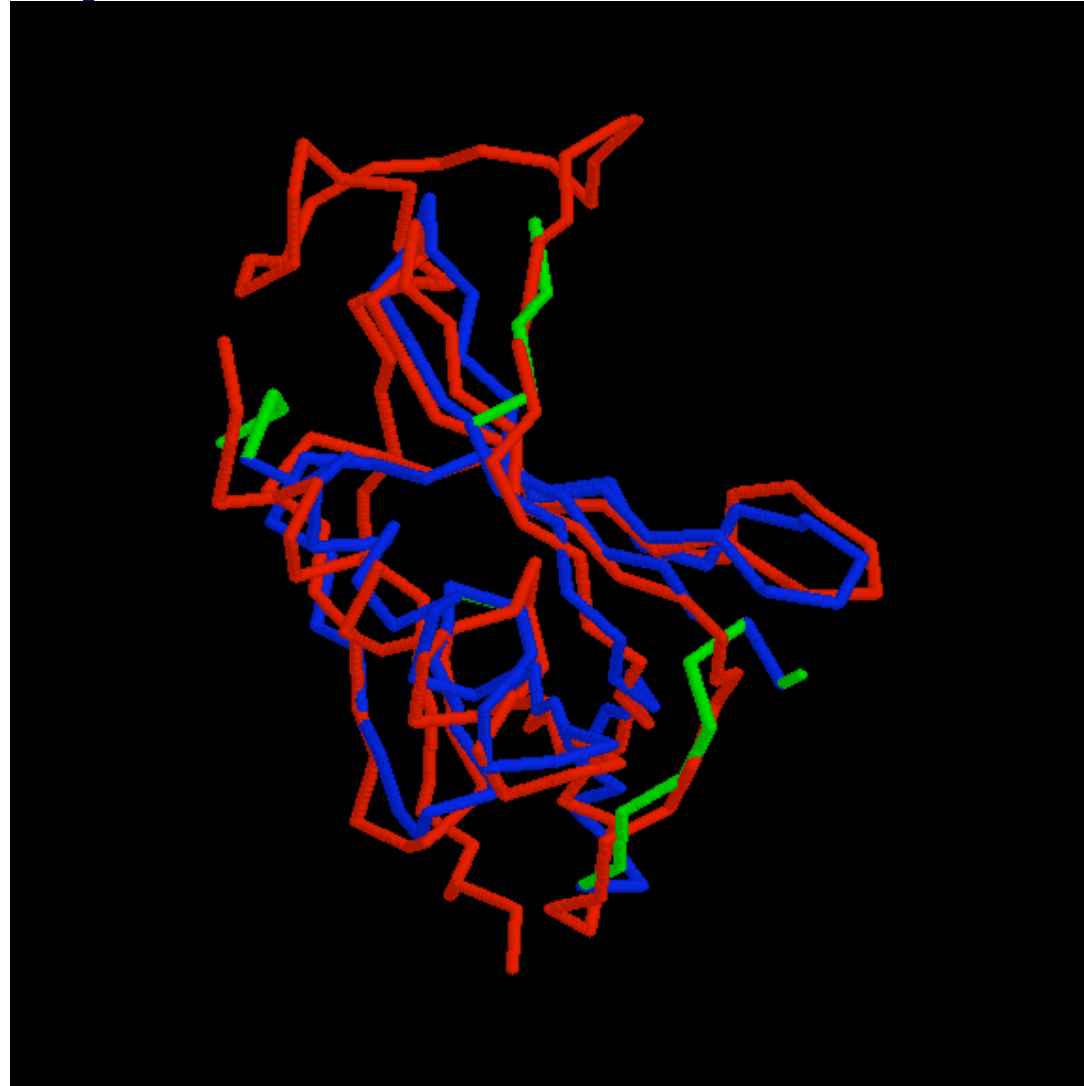
Structural Classification of Proteins

<http://scop.mrc-lmb.cam.ac.uk/scop/>



Identifies tertiary structure, domains, fold classes, supersecondary structure motifs, structural families, homologous structures, etc.

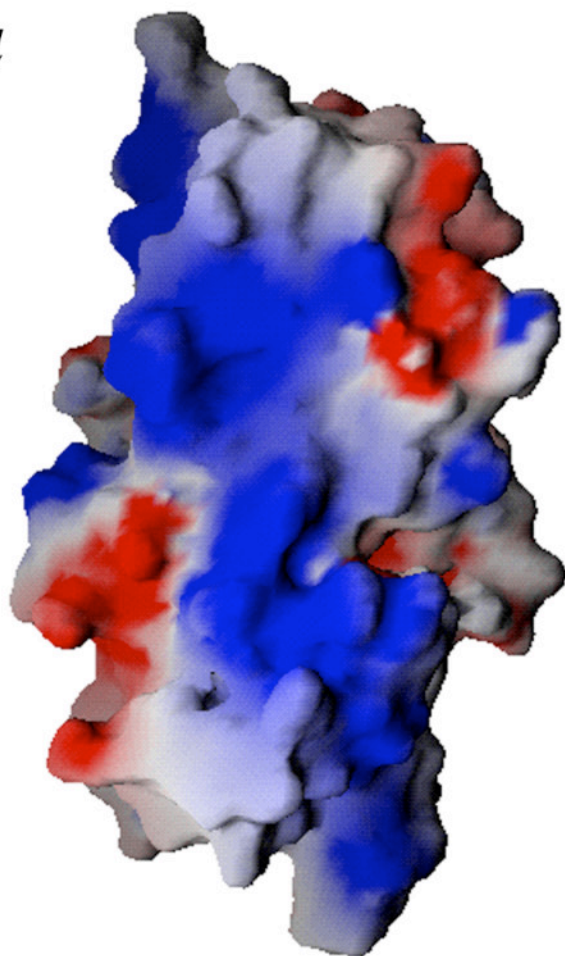
Superposition and rmsd calculation



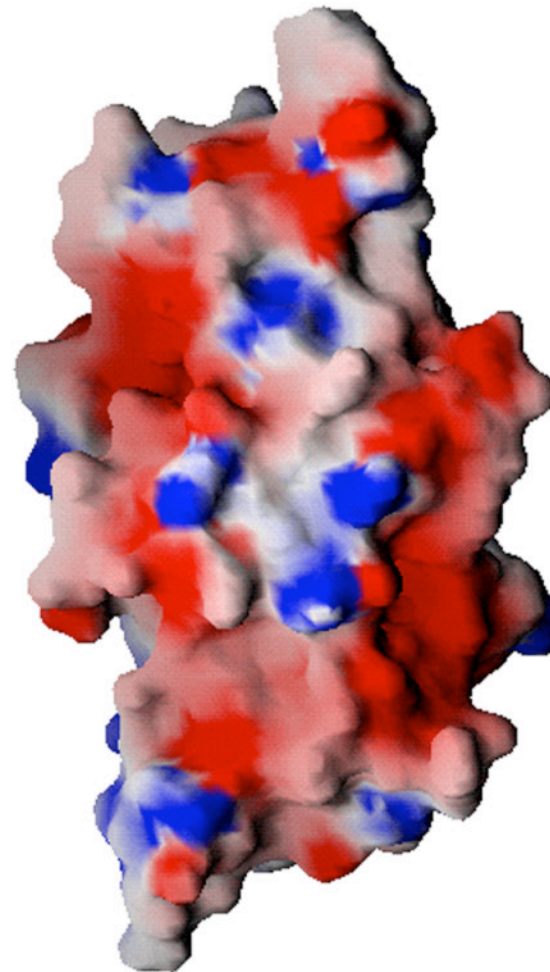
The function is usually carried out on C α atoms of amino acids

Electrostatic surface potential

a



b



Acidic (negatively charged) surfaces are modeled red and basic (positively charged) surfaces are depicted in blue.

Molecular modeling

- Use computer-based algorithms that work on coordinate files to carry out discovery-level investigations
- Dynamics-attempt to model macromolecular motion and simulate biological chemistry in real-time
- Homology modeling-use known structures as templates to predict the structure of related molecules

Structure-based biochemistry

- Use the information contained in a newly determined x-ray crystal structure to suggest testable hypotheses about how the molecule functions *in vivo*
- Get into the lab and test your hypothesis!
 - Mutagenesis
 - Deletions
 - Chimeras

Assignment

- Have a great Thanksgiving holiday!
- Give up on the Detroit Lions in the first half and complete Gale Rhodes' online tutorial on Swiss Pdb Viewer

<http://spdbv.vital-it.ch/TheMolecularLevel/SPVTut/index.html>