RNA2: Last week's take home lessons

- Clustering by gene and/or condition
- Distance and similarity measures
- Clustering & classification
- Applications
- DNA & RNA motif discovery & search

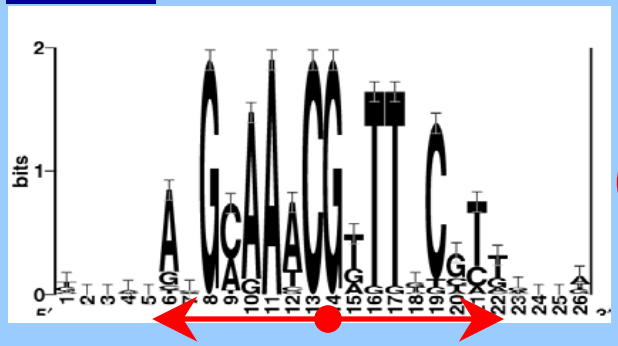
Protein1: Today's story & goals

- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

Palindromicity

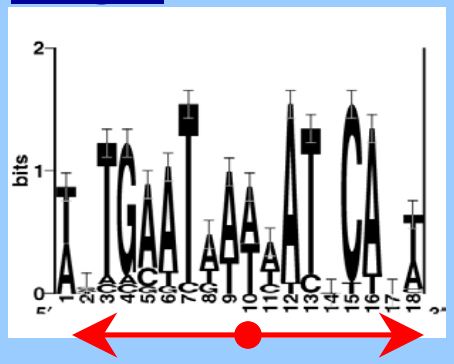
- CompareACE score of a motif versus its reverse complement
- Palindromes: CompareACE > 0.7
- Selected palindromicity values:

PurR



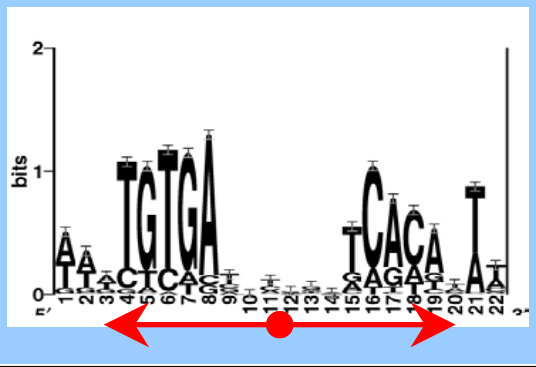
0.97

ArgR



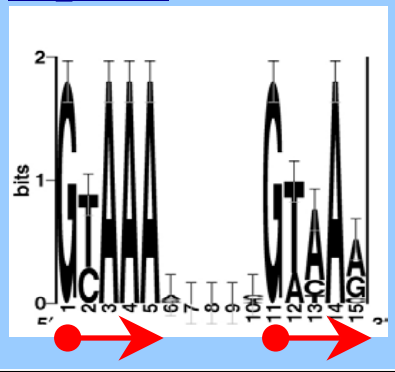
0.92

Crp



0.92

CpxR



0.39

Is there a code for protein interactions with DNA or RNA?

ABCs of Protein Structure

Fig **(See http://ntri.tamuk.edu/hplc/protein.html)**

Interactions of Adjacent Basepairs in EGR1 Zinc Finger DNA Recognition

See Isalan et al., *Biochemistry* (‘98) 37:12026-12033

Wildtype
RSDHLTT

Motifs: weight all 64 K_a^{app}

TGG **2.8 nM**

RGPDLAR
REDVLIR

GCG **16 nM**

2.5 nM

LRHNLET

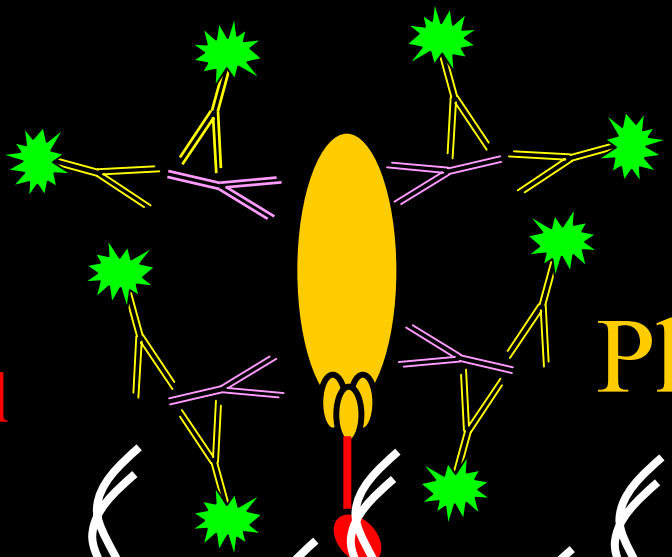
TAT **5.7 nM**

KASNLVS

AAA,AAT,ACT,AGA,
AGC,AGT,CAT,CCT,
CGA,CTT,TTC,TTT

AAT **240 nM**

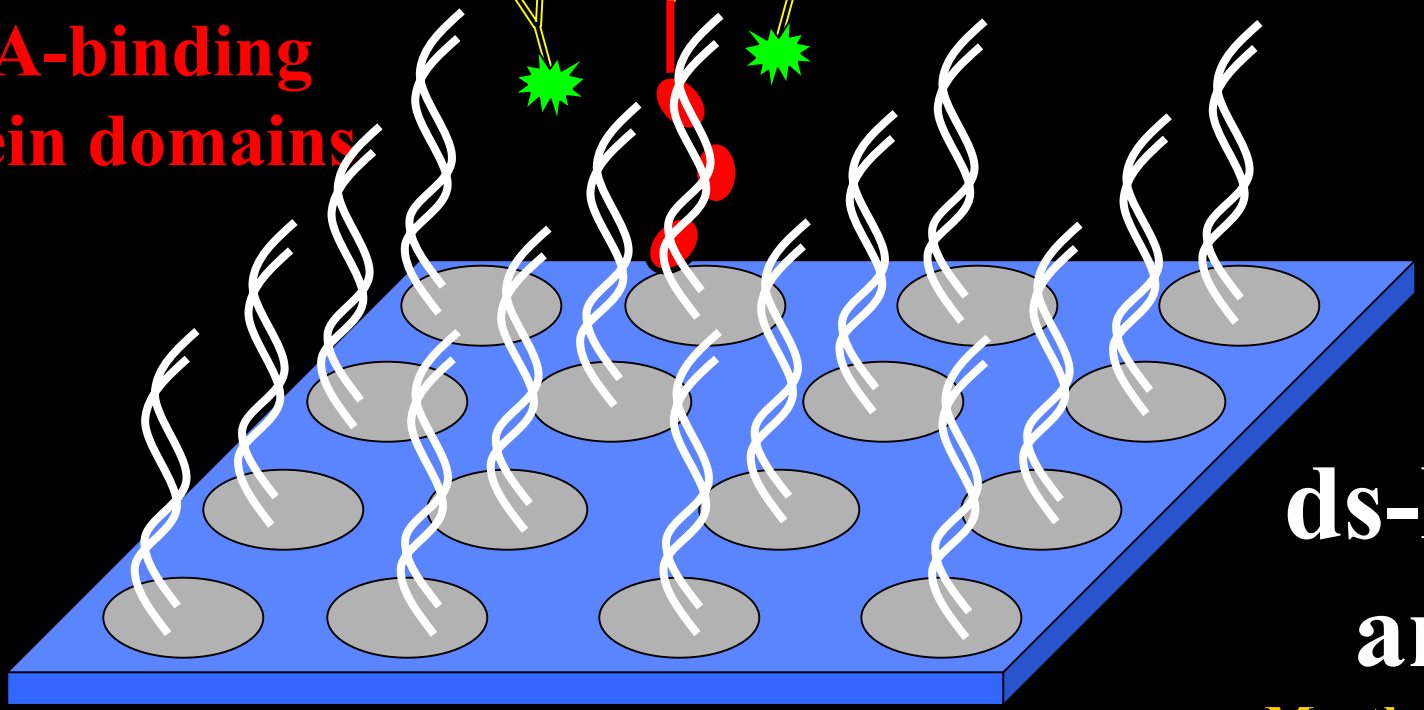
Combinatorial arrays for binding constants



Phycoerythrin
- 2º IgG

Phage

Combinatorial
DNA-binding
protein domains



ds-DNA
array

Martha Bulyk et al

Ka apparent (association constant)

$$P + D \xleftrightarrow{Ka} P \cdot D$$

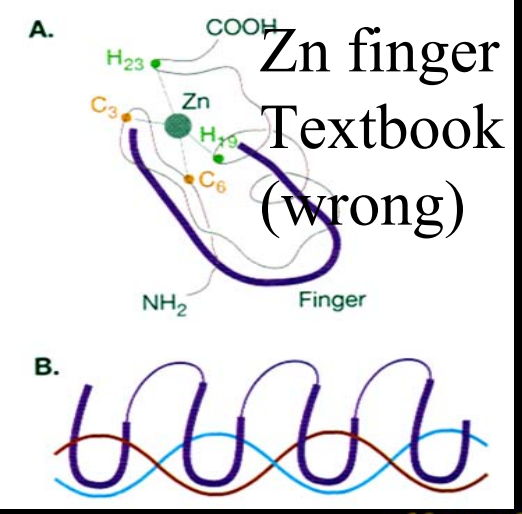
$$Ka = \frac{[P \cdot D]}{[P][D]}$$

The fraction of DNA molecules with protein bound is:

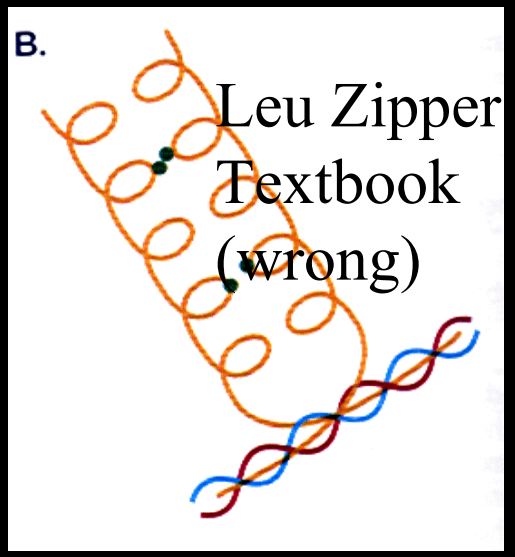
$$= \frac{[P \cdot D]}{[D] + [P \cdot D]} = \frac{Ka[P][D]}{[D] + Ka[P][D]} = \frac{[P]}{1/Ka + [P]}$$

relative signal intensity is expected to be directly proportional

DNA binding



Zn finger Textbook (wrong)



Leu Zipper Textbook (wrong)

zif268/DNA (zinc finger)

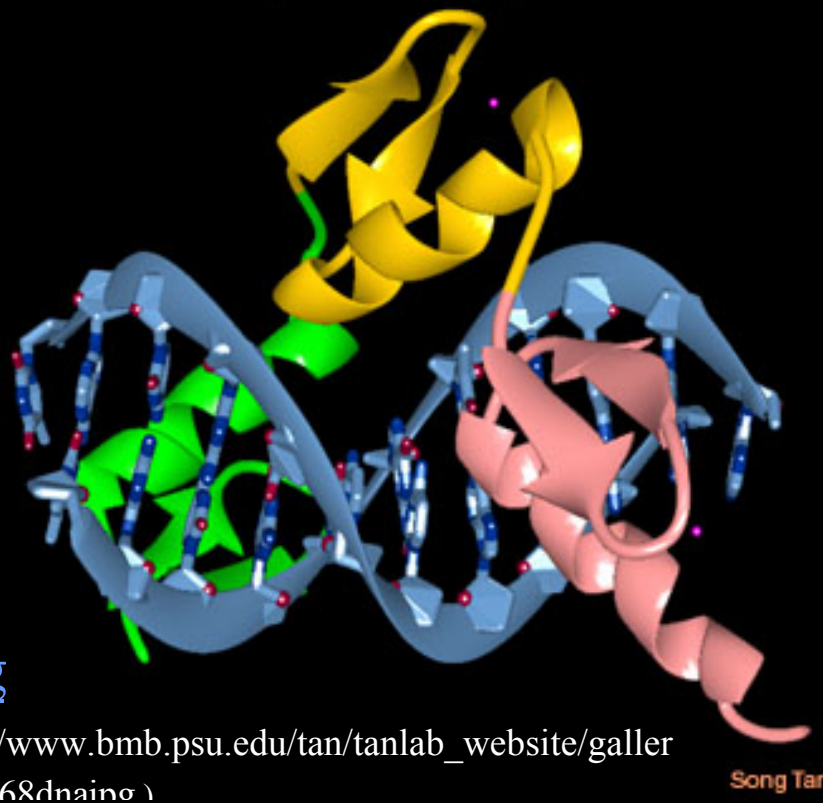


Fig (http://www.bmb.psu.edu/tan/tanlab_website/gallery/zif268dnaipg)

GCN4 COMPLEX WITH AP-1 DNA (1YSA)

GCN4 fig (http://www.rtc.riken.go.jp/jouhou/image/dna-protein/all/small_N1ysa.gif

A code for protein interactions with RNAs?

See Wang et al. (2001) Expanding the genetic code of Escherichia coli. Science 292:498-500

Protein1: Today's story & goals

- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

Real world programming (3D + time)

Perl exercises & central dogma:
Bit I/O, syntax, memory, conditionals,
loops, operators, functions, documentation.

For real world interfaces add:
Sensors & actuators
Issues of feedback, synchrony,
analog to digital to analog

Scary proteins

Anthrax

Protectve Antigen (transport)
Edema Factor
Lethal Factor
(Nature Biotech 19:958)
(http://arep.med.harvard.edu/pdf/Mourez01.pdf)

HIV-1 Polymerase
ApoE4 Atherosclerosis & Alzheimer's
Staph hemolysin (Net2)

Protein programming time scales

f- to nsec	atomic motion
μ- to msec	enzyme turnover
sec	drug cell diffusion
min	transcription
hr-day	cell-cycle
day	circadian
17 years	cicada
100 years	aging

What good are 3D protein structures?

Depends on accuracy.

See Baker & Sali (2001) Science 294/5540/93/F1

Structure Based Drug Design

Stout TJ, et al. Structure-based design of inhibitors specific for bacterial thymidylate synthase. Biochemistry. 1999 Feb 2;38(5):1607-17.

Frecer V, Miertus S, Tossi A, Romeo D Drug Des Discov 1998 Oct;15(4):211-31. Rational design of inhibitors for drug-resistant HIV-1 aspartic protease mutants.

Kirkpatrick DL, Watson S, Ulhaq S Comb Chem High Throughput Screen 1999 2:211-21. (Pub) Structure-based drug design: combinatorial chemistry and molecular modeling.
(http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?uid=10469881&form=6&db=m&Dopt=b)

Guo et al. Science 2000 288:2042-5. Designing small-molecule switches for protein-protein interactions. (Pub)
(http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10856217&dopt=Abstract)

Lee et al. PNAS 1998 95:939-44. Analysis of the S3 and S3' subsite specificities of feline immunodeficiency virus (FIV) protease: development of a broad-based protease inhibitor efficacious against FIV, SIV, & HIV in vitro & ex vivo. (Pub)
(http://www.pnas.org/cgi/content/full/95/3/939)

Covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance

See Huang et al. Science 1998 282:1669-75.. (Pub)
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=9831551&dopt=Abstract)

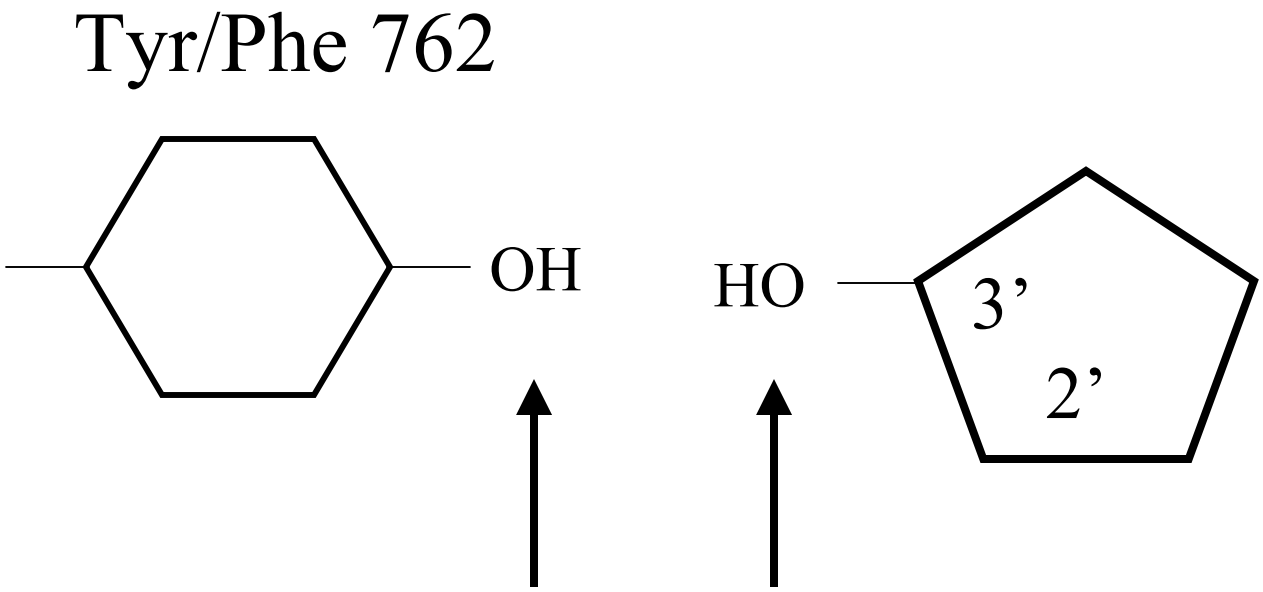
3D structure & chemical genetics

Tabor & Richardson PNAS 1995 92:6339-43 A single residue in DNA polymerases of the Escherichia coli DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. (Pub)
F to Y (one atom) gives up to a 8000-fold specificity effect, hence dye-terminators feasible (and uniform).
(http://www.ncbi.nlm.nih.gov/entrez/utils/fref.fcgi?http://www.pnas.org/cgi/pmidlookup?view=reprint&pmid=7603992)

Louvion et al. Gene 1993 131:129-34. Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast. (Pub)

Shakespeare et al. PNAS 2000 97:9373-8. Structure-based design of an osteoclast-selective, nonpeptide src homology 2 inhibitor with in vivo antiresorptive activity. (Pub)

Compensating steric hinderance in DNA polymerases



Real world programming with proteins

Transgenics: Overproduction or restoration
Homologous recombination: Null mutants
Point Mutants: Conditional mutants, SNPs

Chemical genetics & drugs:
Combinatorial synthesis
Structure-based design
Mining biodiversity compound collections
Quantitative Structure-Activity Relationships QSAR
(http://mmlin1.pha.unc.edu/~jin/QSAR/)

Protein1: Today's story & goals

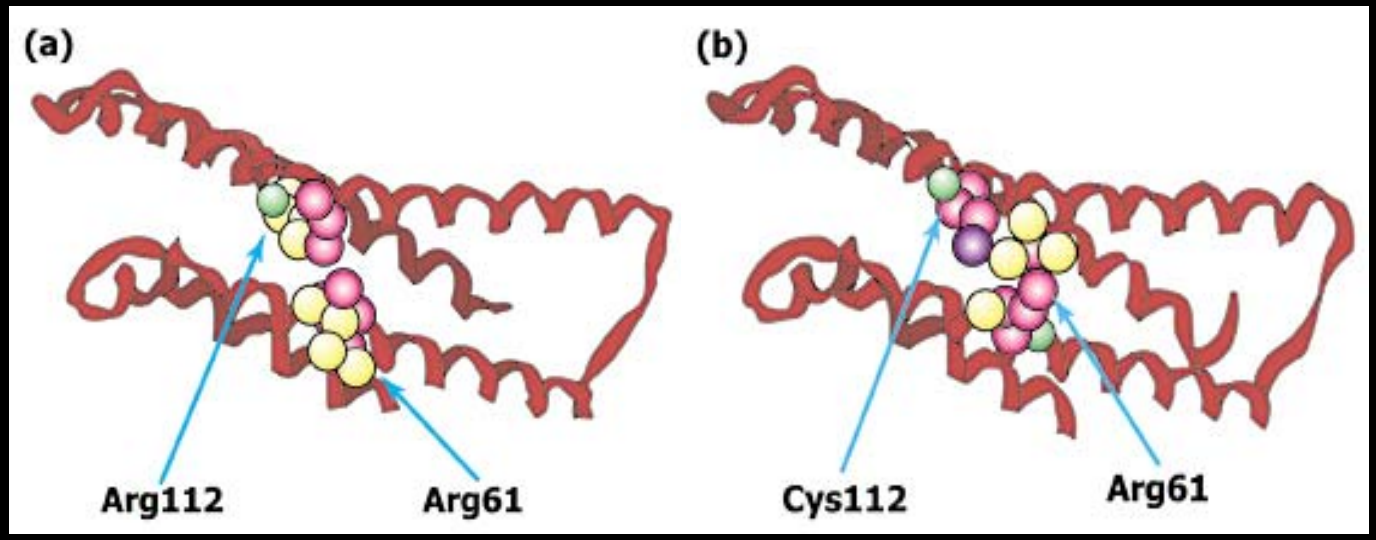
- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

Altered specificity mutants (continued)

Genetic strategy for analyzing specificity of dimer formation: Escherichia coli cyclic AMP receptor protein mutant altered in dimerization
Immunoglobulin V region variants in hybridoma cells. I. Isolation of a variant with altered idiotypic and antigen binding specificity.
In vitro selection for altered divalent metal specificity in the RNase P RNA.
In vitro selection of zinc fingers with altered DNA-binding specificity.
In vivo selection of basic region-leucine zipper proteins with altered DNA-binding specificities.
Isolation and properties of Escherichia coli ATPase mutants with altered divalent metal specificity for ATP hydrolysis.
Isolation of altered specificity mutants of the single-chain 434 repressor that recognize asymmetric DNA sequences containing TTAA
Mechanisms of spontaneous mutagenesis: clues from altered mutational specificity in DNA repair-defective strains.
Molecular basis of altered enzyme specificities in a family of mutant amidases from Pseudomonas aeruginosa.
Mutants in position 69 of the Trp repressor of Escherichia coli K12 with altered DNA-binding specificity.
Mutants of eukaryotic initiation factor eIF-4E with altered mRNA cap binding specificity reprogram mRNA selection by ribosomes in
Mutational analysis of the CitA citrate transporter from Salmonella typhimurium: altered substrate specificity.
Na+-coupled transport of melibiose in Escherichia coli: analysis of mutants with altered cation specificity.
Nuclease activities of Moloney murine leukemia virus reverse transcriptase. Mutants with altered substrate specificities.
Probing the altered specificity and catalytic properties of mutant subtilisin chemically modified at position S156C and S166C in the S1
Products of alternatively spliced transcripts of the Wilms' tumor suppressor gene, wt1, have altered DNA binding specificity and regulate
Proline transport in Salmonella typhimurium: putP permease mutants with altered substrate specificity.
Random mutagenesis of the substrate-binding site of a serine protease can generate enzymes with increased activities and altered
Redesign of soluble fatty acid desaturases from plants for altered substrate specificity and double bond position.
Selection and characterization of amino acid substitutions at residues 237-240 of TEM-1 beta-lactamase with altered substrate specificity
Selection strategy for site-directed mutagenesis based on altered beta-lactamase specificity.
Site-directed mutagenesis of yeast eEF1A. Viable mutants with altered nucleotide specificity.
Structure and dynamics of the glucocorticoid receptor DNA-binding domain: comparison of wild type and a mutant with altered specificity.
Structure-function analysis of SH3 domains: SH3 binding specificity altered by single amino acid substitutions.
Sugar-binding and crystallographic studies of an arabinose-binding protein mutant (Met108Leu) that exhibits enhanced affinity & altered
T7 RNA polymerase mutants with altered promoter specificities.
The specificity of carboxypeptidase Y may be altered by changing the hydrophobicity of the S'1 binding pocket.
The structural basis for the altered substrate specificity of the R292D active site mutant of aspartate aminotransferase from E. coli.
Thymidine kinase with altered substrate specificity of acyclovir resistant varicella-zoster virus.
U1 small nuclear RNAs with altered specificity can be stably expressed in mammalian cells and promote permanent changes in
Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex.
Use of Chinese hamster ovary cells with altered glycosylation patterns to define the carbohydrate specificity of Entamoeba histolytica
Using altered specificity Oct-1 and Oct-2 mutants to analyze the regulation of immunoglobulin gene transcription.
Variants of subtilisin BPN' with altered specificity profiles.
Yeast and human TFIID with altered DNA-binding specificity for TATA elements.

SNPs & Covariance in proteins

ApoE-e4 (20%) **e3**



Ancestral = Arg 112 Thr 61

Prediction of deleterious human alleles

1) Binding site,
2) buried charge or hydrophobic change
3) Disulfide loss
4) Solubility
5) Proline in helix
6) Incompatible with multisequence profile

Hum Molec Gen 10:591-7.

Protein1: Today's story & goals

- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

Biochemical diversity

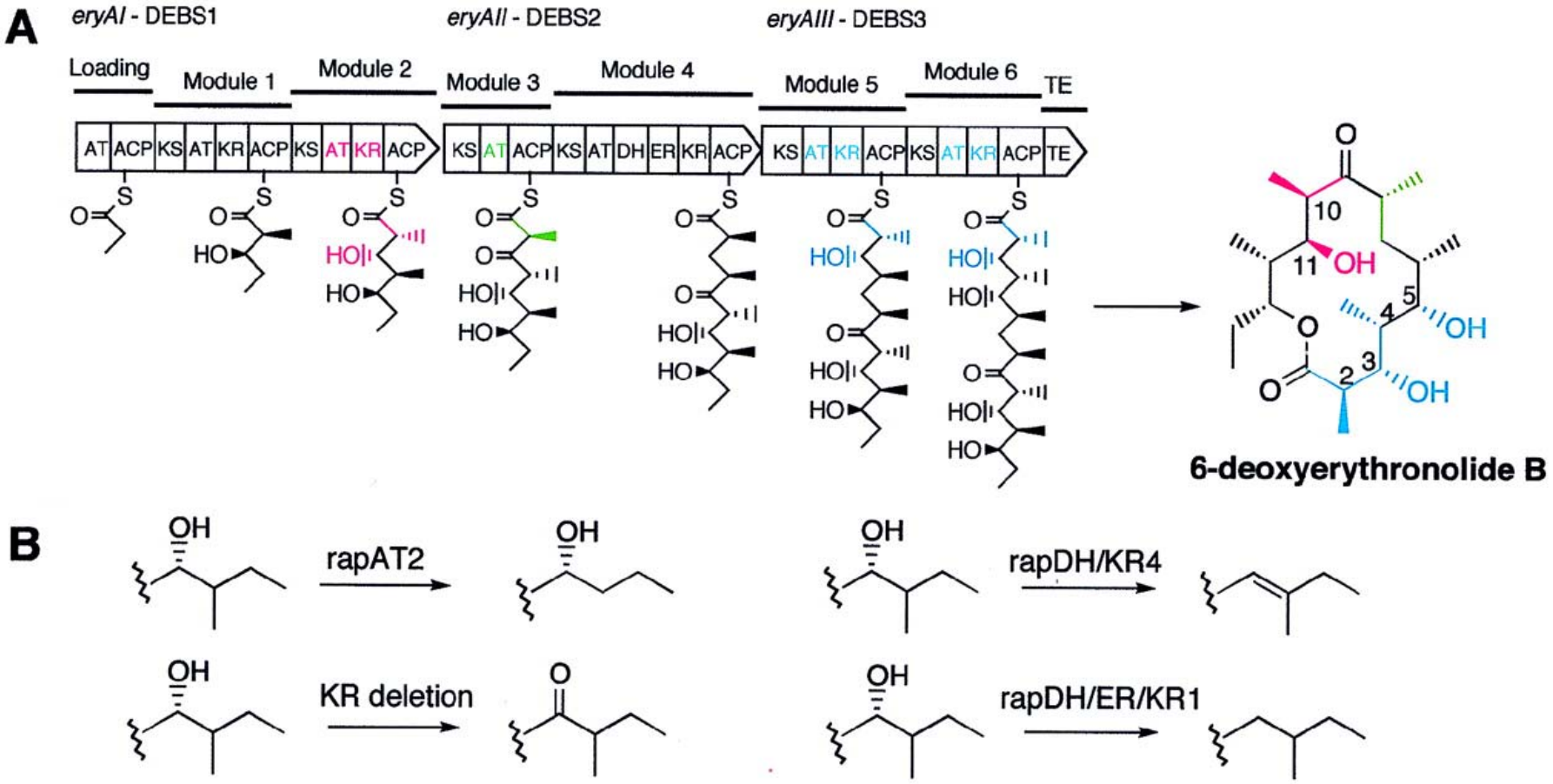
See Xue Q, et al. 1999 PNAS 96:11740-5 A multiplasmid approach to preparing large libraries of polyketides.

and

Olivera BM, et al. 1999 Speciation of cone snails and interspecific hyperdivergence of their venom peptides. Ann NY Acad Sci. 870:223-37.

Immune receptor
diversity

Polyketide engineering



Protein interaction assays

Harvard ICCB (http://iccb.med.harvard.edu/)

Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors.

See Maly et al. PNAS 2000 97:2419-24 (Pub)

Protein1: Today's story & goals

- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

Computational protein target selection

Homologous: for example to successful drug targets

Conserved: Arigoni et al. Nat Biotechnol 1998 16: 851-6
A genome-based approach for the identification
of essential bacterial genes. (Pub)
(http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?uid=9743119&form=6&db=m&Dopt=b)

Surface accessible: antibodies or cell excluded drugs
(e.g. from membrane topology prediction)

Disease associated: differential gene expression clusters

Given many genome sequences
(of accuracy 99.99%)

Sequence to exon 80% [Laub 98]
Exons to gene (without cDNA or homolog) ~30% [Laub 98]
Gene to regulation ~10% [Hughes 00]
Regulated gene to protein sequence 98% [Gesteland]
Sequence to secondary-structure (α,β,c) 77% [CASP5 Dec'02]
Secondary-structure to 3D structure 25% [CASP]
3D structure to ligand specificity ~10% [Johnson 99]

Expected accuracy overall $\sim =$ 0.8*.3*.1*.98*.77*.25*.1 = .0005 ?

http://cubic.bioc.columbia.edu/papers/2002_rev_dekker/paper.html
http://depts.washington.edu/bakerpg/
CASP = Computational Assessment of Structure Prediction

Measuring 3D protein family relationships

3D to 3D comparsions:
CATH Class, Architecture, Topology & Homology (UCI)
CE Combinatorial Extension of the optimal path (RCSB)
FSSP Fold class by Structure-Structure alignment of Proteins (EBI)
SCOP Structural Classification Of Proteins (MRC)
VAST Vector Alignment Search Tool (NCBI)

3D to sequence: "Threading"

Ref
(http://www.rcsb.org/pdb/cgi/explore.cgi?job=neighbors&pdbId=2NLL&page=&pid=12722972873212)

Structural genomics projects

Goals:
1) Assign function to proteins with only cellular or phenotypic function
2) Assign functional differences within a sequence family
3) Interpret disease associated single nucleotide polymorphisms (SNPs).

Selection criteria 35% identity clusters:
Large Families with a predefined limit on sequence length
Families in all 3 main domains of life (prokaryotes, archaea, eukaryotes)
Families with a human member
Families without a member of known structure
Non-transmembrane families

www.nih.gov/nigms/news/meetings/structural_genomics_targets.html
Current estimated cost: **$200K**/structure
Target cost: 10,000 per 5 years = **$8K/**structure.

Programming cells via membrane proteins

Number of types of ligands larger
Number of potential side-reactions smaller
Basic cell properties:
Adhesion, motility, immune recognition

Membrane protein 3D structures

Soluble fragments of fibrous & membrane proteins

Myosin, flu hemagglutinin, histocompatibility antigens, T-cell receptor, etc.

Integral membrane proteins

Prostaglandin H2 synthase, Cyclooxygenase, Squalene-hopene cyclase, **Bacteriorhodopsin,** Photosynthetic Reaction Centers, Light Harvesting Complexes, Photosystem I, Multi-,monomeric beta-barrel pores, Toxins, Ion Channels, Fumarate Reductase, Cytochrome C Oxidases, Cytochrome bc1 Complexes, Ca ATPase Water & Glycerol channels, GPCR-Rhodopsin, F1-ATPase

blanco.biomol.uci.edu/Membrane_Proteins_xtal.html

(http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html)

Transmembrane prediction

J Mol Biol 2001 Oct 5;312(5):927-34 Energetics, stability, and prediction of transmembrane helices. Jayasinghe et al.

Backbone constraint, identifies TM helices of membrane proteins with an accuracy greater than 99 %. (& energetics of salt-bridge formation. Falsely predicts 17 to 43 % of a set of soluble proteins to be MPs, depending upon the hydropathy scale used

"function from structure"

Surface electrostatics, as displayed, (e.g., GRASP, Nicholls, et al.) can identify DNA & RNA binding sites, occasionally, other features.

Thornton et al: small ligand binding sites are almost always associated with the largest depressions in the surface of a protein... visually

Conserved motifs in a family (on the surface of a structure) as a method of finding functional features, particularly protein-protein interaction sites.

3D catalytic motifs can be catalogued & used to identify the catalytic function of new structures.

Methods developed in drug design to identify potential lead compounds are expected to be applicable to deducing ligand-binding specificity.

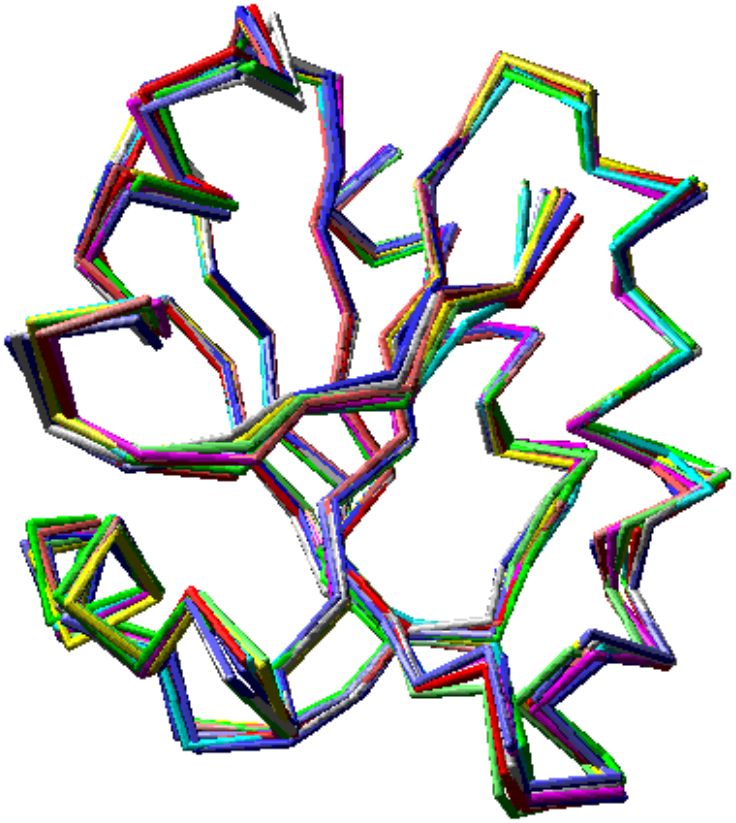
http://www.nih.gov/nigms/news/meetings/structural_genomics_targets.html
http://bioinfo.mbb.yale.edu/genome/foldfunc/

Where do 3D structures come from?

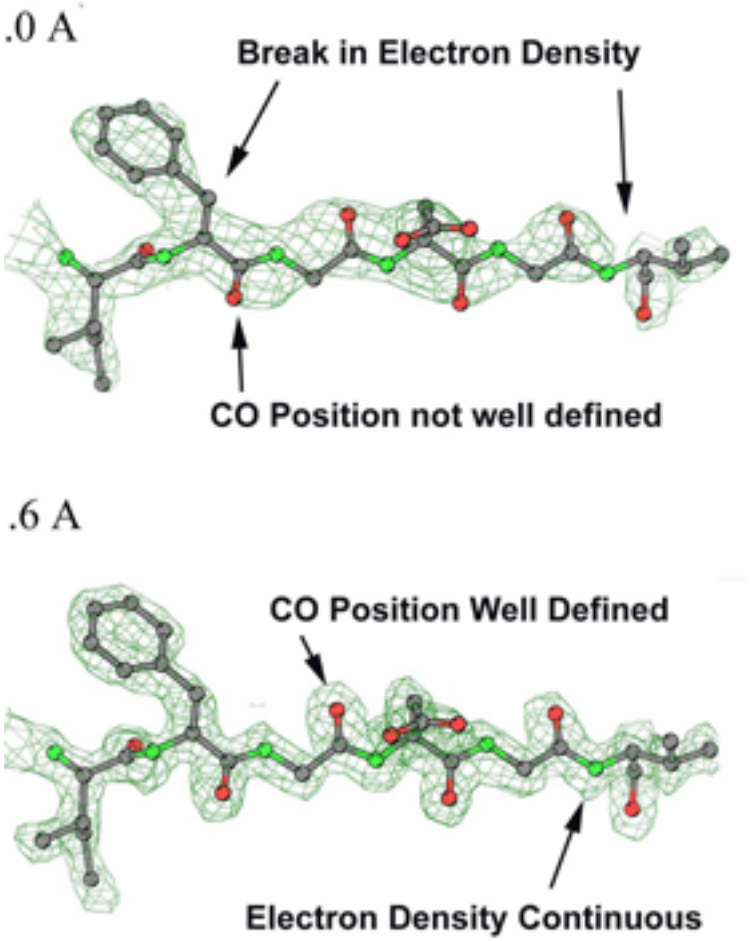
Research Collaboratory for Structural Bioinformatics
Protein Data Bank (RCSB PDB)
(http://www.rcsb.org/pdb/cgi/queryForm.cgi)

```
HEADER    COMPLEX (TRANSCRIPTION REGULATION/DNA)  23-NOV-93   1HCQ       1HCQ   2
COMPND   2 MOLECULE: HUMAN/CHICKEN ESTROGEN RECEPTOR;                     1HCQ   4
REMARK   2 RESOLUTION. 2.4  ANGSTROMS                                     1HCQ  39
REMARK   3   PROGRAM 1                  X-PLOR                            1HCQ  42
REMARK   3   R VALUE                    0.204                             1HCQ  46
SEQRES   1 A   84  MET LYS GLU THR ARG TYR CYS ALA VAL CYS ASN ASP TYR     1HCQ  60
SEQRES   1 C   18    C   C   A   G   G   T   C   A   C   A   G   T   G     1HCQ  74
FORMUL   9   ZN    8(ZN1 2+)                                              1HCQ 107
FORMUL  10  HOH   *158(H2 O1)                                             1HCQ 108
HELIX    1   1 GLU A   25  ILE A   35  1                                  1HCQ 109
ATOM      1  N   MET A   1      50.465  24.781  79.460  1.00 60.88       1HCQ 133
ATOM      2  CA  MET A   1      50.332  26.116  80.055  1.00 61.13       1HCQ 134
CONECT 2983 2747 2789                                                     1HCQ4038
MASTER       22    3    8    9    8    0    0    6 3864    8   34   36  1HCQ4039
END                                                                       1HCQ4040
```

NMR distance-constrained ensembles
Crystallographic phases & electron density



Cα trace



Ref1, 2 (http://www.usm.maine.edu/~rhodes/ModQual/index.html#Electron-density map)

Crystallographic refinement

Fourier transform relates scattered X-rays, F, to electron density, ρ.
Δk is the scattering vector.

$$\mathbf{F}(\Delta\mathbf{k}) = V \int_{x=0}^{x=1} \int_{y=0}^{y=1} \int_{z=0}^{z=1} \rho(\mathbf{x},\mathbf{y},\mathbf{z})\, e^{i\Delta\mathbf{k}\cdot(\mathbf{x}\mathbf{a}+\mathbf{y}\mathbf{b}+\mathbf{z}\mathbf{c})}\, dx dy dz$$

Minimize Fo-Fc.
Linearize with a first order Taylor expansion; parameters p (e.g. = x,y,z)

$$\Delta(\mathbf{p} + \xi) = \Delta(\mathbf{p}) - \sum_{i=1}^{n} \xi_i \frac{\partial |F_{calc}|}{\partial p_j}$$

(ref)
(http://www.ysbl.york.ac.uk/~mgwt/thesis-tth/chapter2.html)

Crystallography & NMR System(CNS) X-plor

Heavy atom searching, experimental phasing (MAD & MIR), density modification, crystallographic refinement with maximum likelihood targets.

NMR structure calculation using NOEs, J-coupling, chemical shift, & dipolar coupling data.

http://cns.csb.yale.edu/v1.0/

Measure Structure Quality

R factor $= \Sigma \, ||Fo|-|Fc|| \, / \, \Sigma \, |Fo|$ < 0.25 good > 0.4 crude

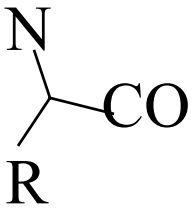
Correlation Coefficient > 0.7

RMSD (root mean square deviation) $= \text{sqrt}\left[\Sigma \, (X_{i1} - X_{i2})^2\right]$
compare models 1 & 2 $i = 1$ to n (#atoms)
canonical peptide geometry

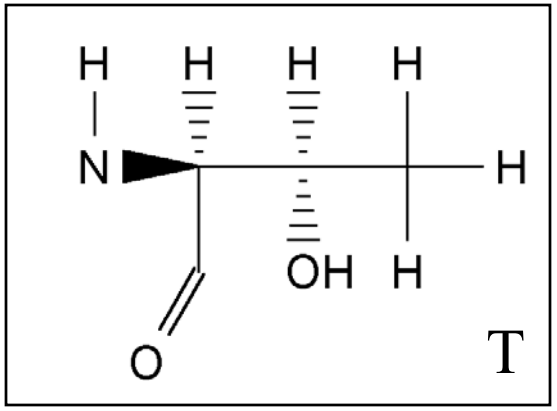
Protein1: Today's story & goals

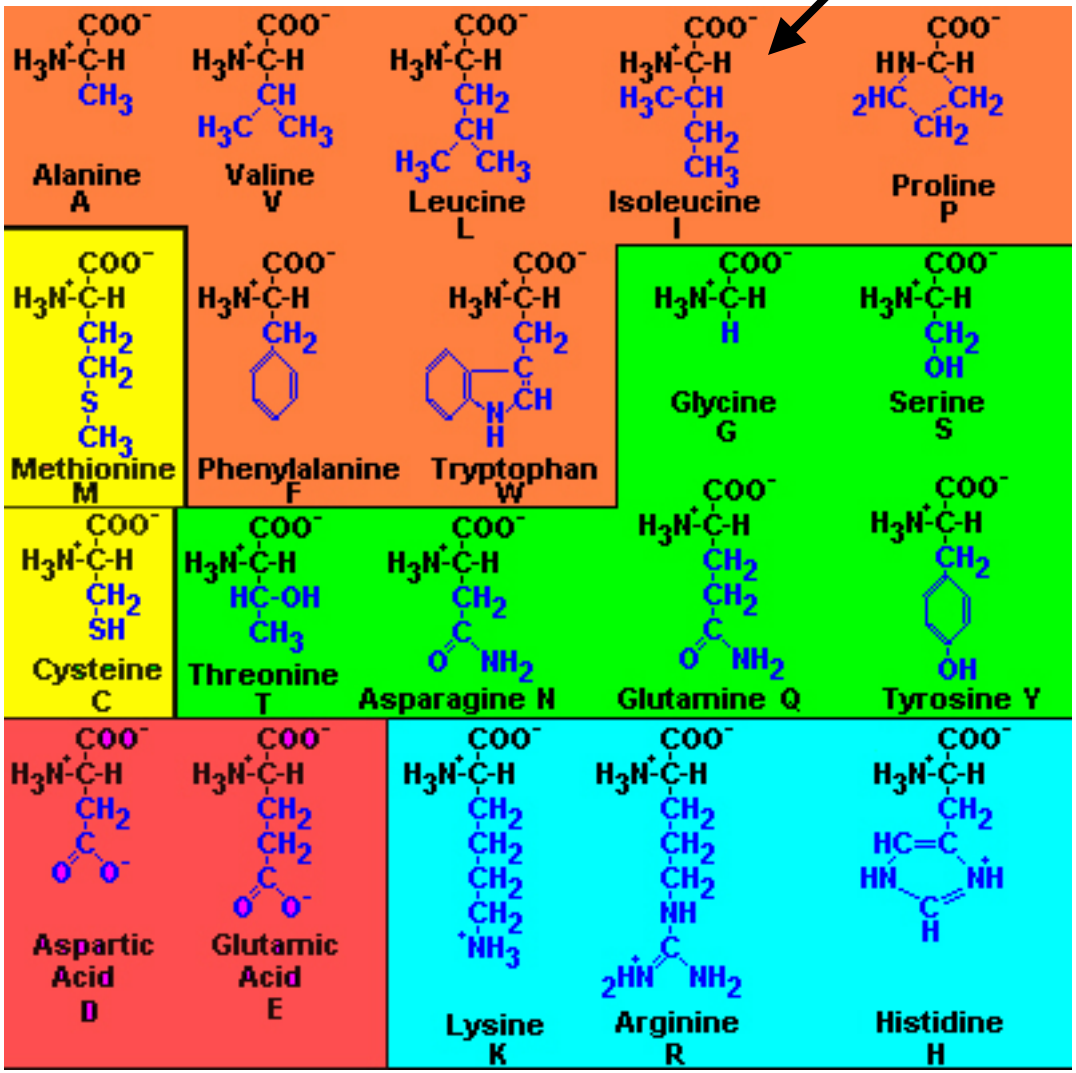
- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

20 Amino acids of 280



19 L-amino acids:
H toward you; CO R N clockwise.





www.people.virginia.edu/~rjh9u/aminacid.html
www-nbrf.georgetown.edu/pirwww/search/textresid.html

Favored peptide conformations

Fig
(http://iona.cryst.bbk.ac.uk/course/section3/rama.html)

Molecular dynamics (Energy minimization, trajectories, approximations)

Quantum Electrodynamics (QED) Schwinger
Born-Oppenheimer Approximation

Quantum Engines Molecular Orbital Methods
Semiempirical Hartree-Fock methods
Modified Intermediate Neglect of Differential Overlap (MINDO)
Modified Neglect of Diatomic Overlap (MNDO) - AMPAC, MOPAC
SemiChem Austin Model 1 (SAM1) - Explicitly treats d-orbitals.

ab initio Hartree-Fock programs:
GAMESS, Gaussian

Semiempirical Engines (Molecular Mechanics) from above & spectroscopy
AMBER, Discover, SYBYL, CHARMM, MM2, MM3, ECEPP.
(**C**hemistry at **HAR**vard **M**olecular **M**echanics),

http://cmm.info.nih.gov/modeling/guide_documents/tocs/computation_software.html
http://www.foresight.org/Nanosystems/toc.html

Molecular mechanics

$$F = m\,a$$

$$-dE/dr_i = F_i = m_i\, d^2r_i/dt^2$$ r = position (radius)

$dt \sim= 1$ fs (1e-15 sec)

$$v_i(t+dt/2) = v_i(t-dt/2) + a_i(t)\, dt$$ update velocity & r

$$r_i(t+dt) = r_i(t) = v(t+dt/2)dt$$

$$E = E_b + E_\theta + E_\omega + E_{vdw} + E_{electrostatic}$$

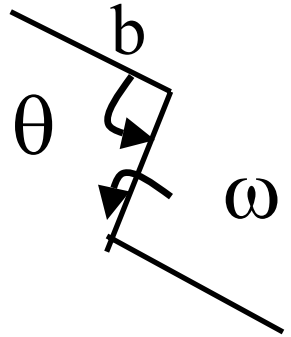
$$E_b = 0.5\, k_b(r-r_0)^2$$

$$E_\theta = 0.5\, k_\theta(\theta - \theta_0)^2$$

$$E_\omega = k_\omega\, [\, 1 + \cos(\, n\, \omega - 1)]$$

$$E_{vdw} = A(r/r_{v0})^{-12} - B(r/r_{v0})^{-6}$$

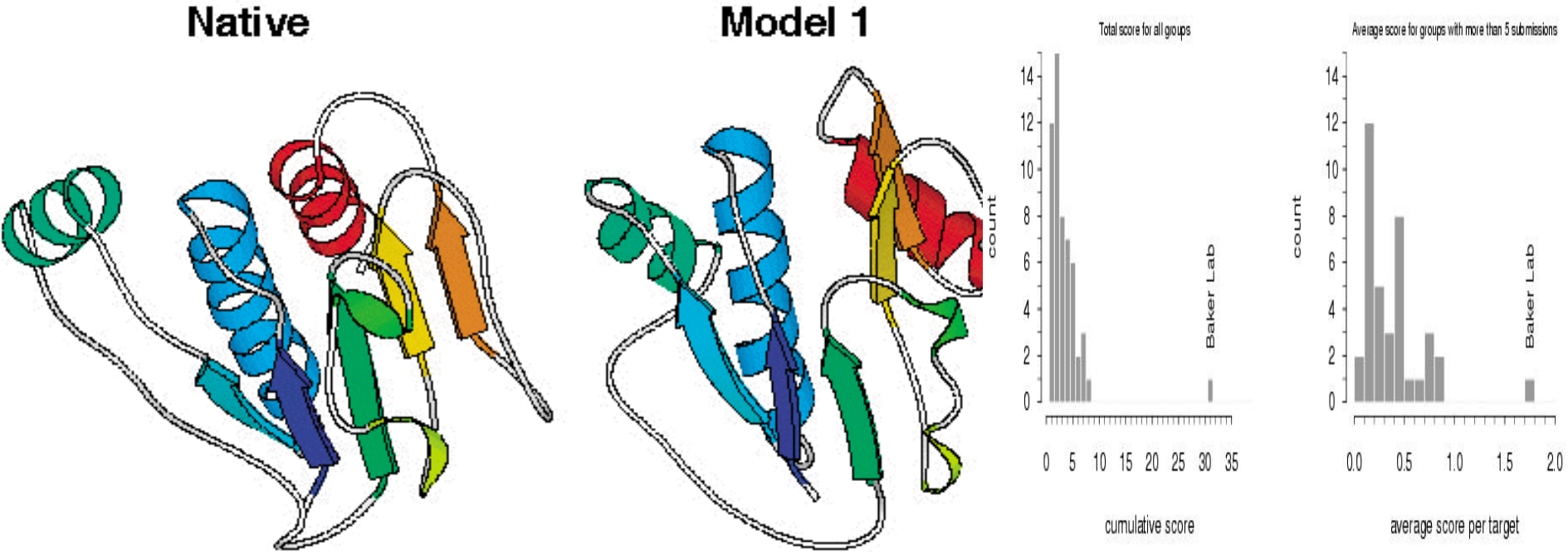
$$E_{electrostatic} = qi\, qj\, /\, e\, r$$



(Ref) (http://www.tau.ac.il/~becker/course/energy.html)

Rosetta (for Ab Initio Structure Prediction CASP4)

T087 - PPase (Domain 1: 2-192)

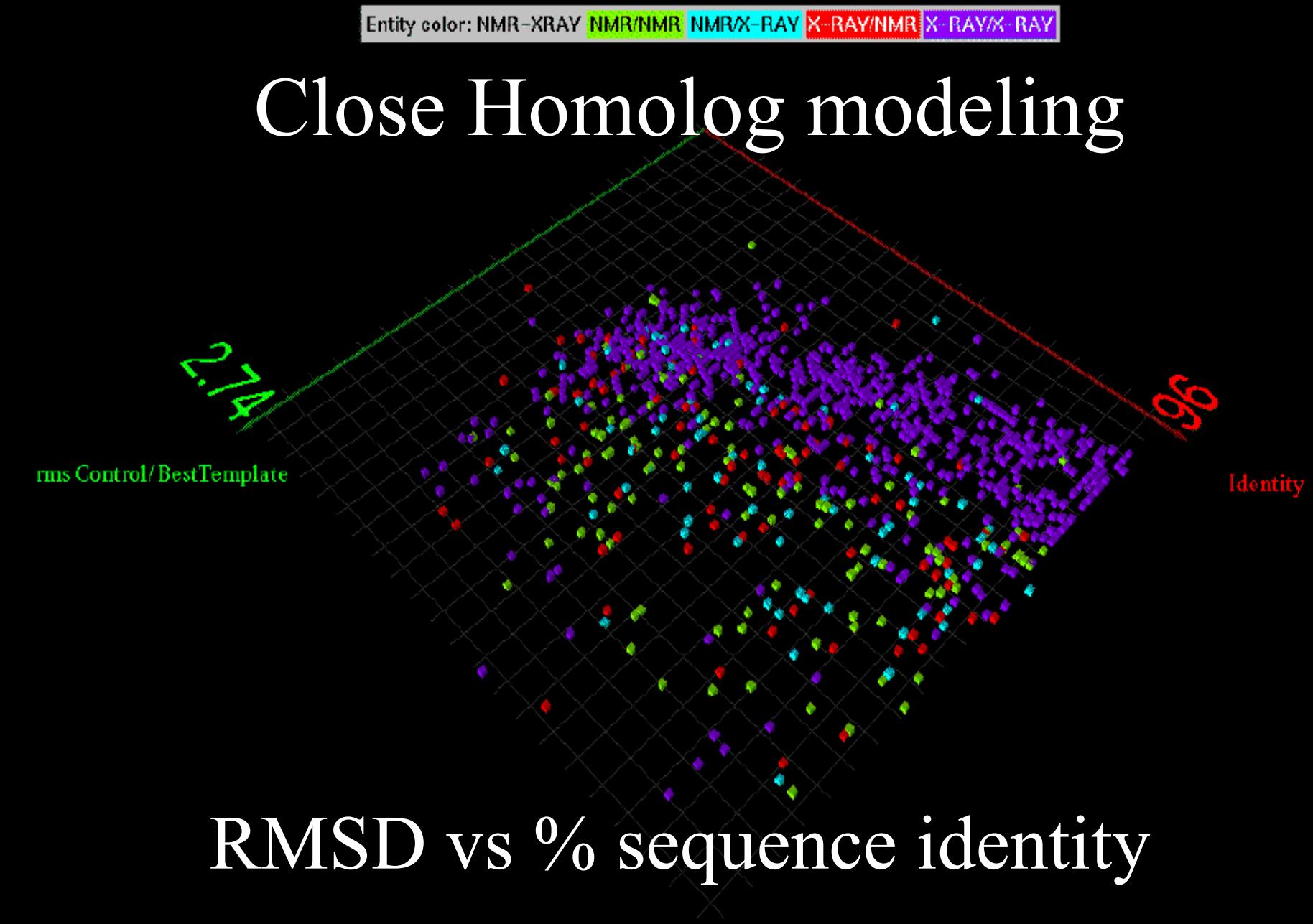


RMSD = 6.5 (128 Cα)
RMSD = 4.8 (95 Cα)

(2 pt for largely correct prediction, 1 point for a somewhat)

http://depts.washington.edu/bakerpg/

Close Homolog modeling



RMSD vs % sequence identity

Small protein molecular dynamics (only water as ligand)

See IBM Blue Gene $100M **(http://www.research.ibm.com/news/detail/bluegene.html)**

and

Duan Y, Kollman PA Science 1998 282:740-4 Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. (36 aa)

and

Daura X, van Gunsteren WF, Mark AE Proteins 1999 Feb 15;34(3):269-80 Folding-unfolding thermodynamics of a beta-heptapeptide from equilibrium simulations.

Docking

Knegtel et al J Comput Aided Mol Des 1999 13:167-83 Comparison of two implementations of the incremental construction algorithm in flexible docking of thrombin inhibitors.

A set of 32 known thrombin inhibitors representing different chemical classes has been used to evaluate the performance of two implementations of incremental construction algorithms for flexible molecular docking: DOCK 4.0 and FlexX 1.5. Both docking tools are able to dock **10-35%** of our test set within **2 A** of their known positions.

Liu M, Wang S J Comput Aided Mol Des 1999 Sep;13(5):435-51 MCDOCK: a Monte Carlo simulation approach to the molecular docking problem. The root-mean-square (rms) of atoms of the ligand between the predicted and experimental binding modes ranges from **0.25 to 1.84** A for the 19 test cases.

Protein1: Today's story & goals

- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk

Top 10 drugs (20-42 M units/yr of 1.6 G units)

Premarin	Estrone, estradiol, estriol replacement
Synthroid	Synthetic thyroid hormone
Lipitor	LDL cholesterol uptake
Prilosec	Ulcers: proton pump inhibitor
Norvasc	Blood Pressure: calcium channel blocker
Prozac	Depression: serotonin uptake
Claritin	Allergy: histamine receptor antagonist
Zithromax	Antibiotic: Erythromycin-like (ribosome)
Zoloft	Depression: serotonin uptake
Glucophage	Diabetes: Insulin signal transduction?

www.cyberpharmacy.co.kr/topic/brand2.html
drwhitaker.com/wit_drug_land.php

Estrogen Receptor DNA binding domain

Gewirth & Sigler Nature Struct Biol 1995 2:386-94. The basis for half-site specificity explored through a non-cognate steroid receptor-DNA complex. Ref (http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?form=6&db=m&Dopt=b&uid=7664096)

Figure (http://www.acsu.buffalo.edu/~jbarnard/Estrogen.html)

rcsb

Estrogen binding domain

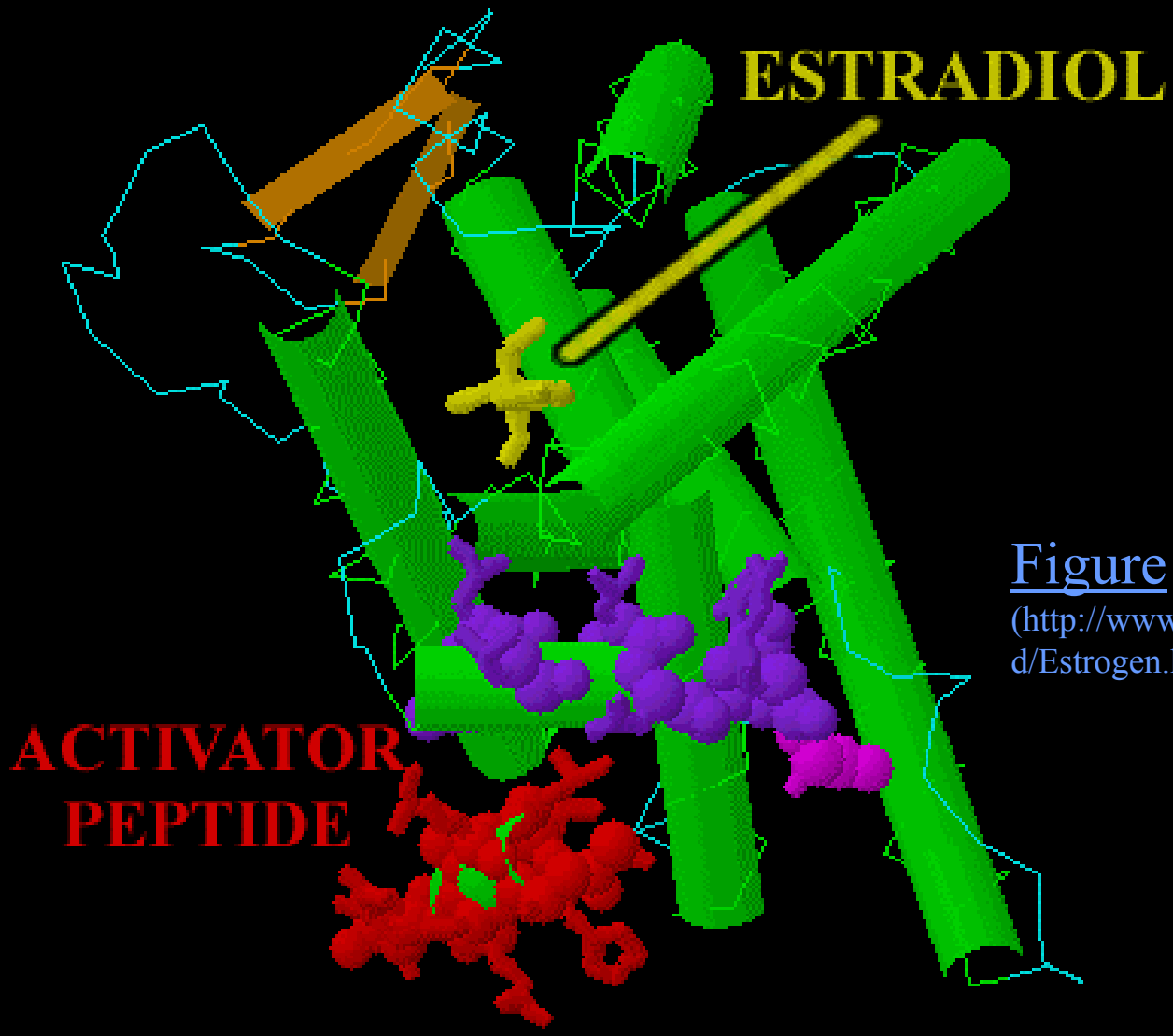


Figure
(http://www.acsu.buffalo.edu/~jbarnard/Estrogen.html)

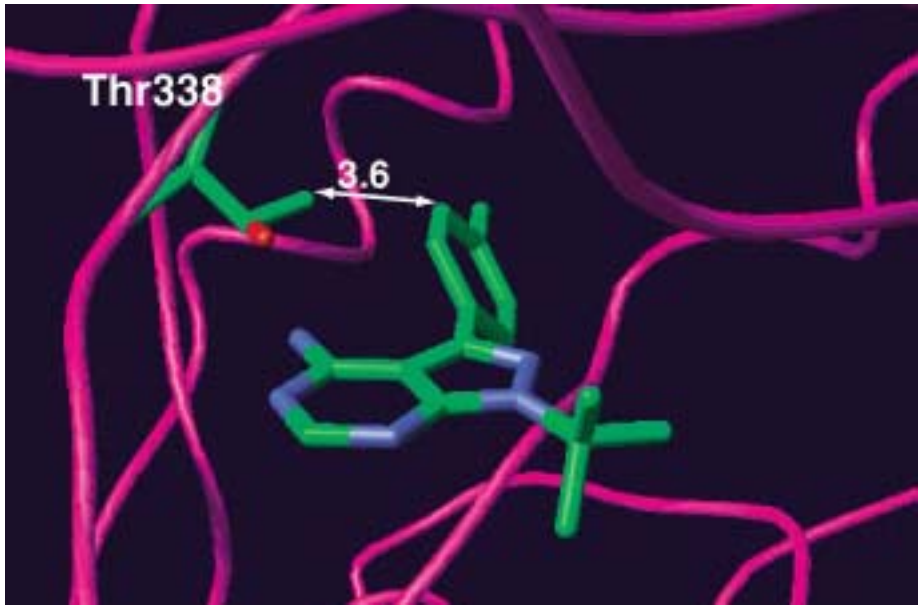
Avoiding receptor cross-talk

Ligands: steroids, retinoids, vitaminD, thyroid hormone
Transduction specificity: Steroid response elements

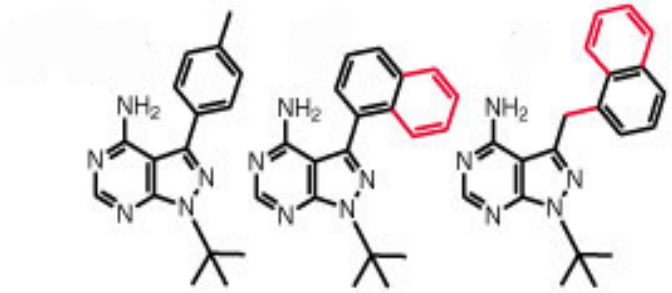
```
 AGGTCA Nn AGGTCA

Half site: AGGTCA or rGkTCr or TAAGGTCA (GR: AGAACA)
DR3          VDR   Vitamin D3
DR2,IR0      RAR   9-cis-retinoate
DR5,DR15     RXR   trans-Retinoate
DR4          T3R   thyroid
IR3,DR15     ER    estrogen
```

Targeting one member of a protein family

A chemical switch for inhibitor-sensitive alleles of any protein kinase.



IC_{50} in μM



Bishop et al. Nature 2000 407: 395-401 (Pub)
(http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=11014197&dopt=Abstract)

T/F338G mutations:

CDK2	22	18	29
CAMK II	17	22	24
CDK2-as1		0.015	0.0050
CAMKII-as1		0.097	0.0080

Protein1: Today's story & goals

- Protein interaction codes(s)?
- Real world programming
- Pharmacogenomics : SNPs
- Chemical diversity : Nature/Chem/Design
- Target proteins : structural genomics
- Folding, molecular mechanics & docking
- Toxicity animal/clinical : cross-talk