Workshop on Structure-based ligand design: *Lecture 1: Introduction (GRID and 'Flu)*

Rebecca C. Wade

European Media Laboratory
Heidelberg

rebecca.wade@eml.villa-bosch.de
http://www.eml.org/english/Research/MCM

EML
MCM
EMBL

Workshop schedule

- **Lecture 1:** Introduction to Structure-based Drug Design, (GRID and 'flu)
- **Practical 1:** GRID
- **Lecture 2:** COMBINE Analysis overview
- **Lecture 3:** Molecular modeling for COMBINE
- **Practical 2:** COMBINE Analysis- molecular modeling
- **Lecture 4:** Chemometric analysis for COMBINE
- **Practical 3:** COMBINE Analysis- chemometrics
- **Lecture 5/Demo/Discussion/Practical 4**

Acknowledgements

- Ingrid Pettersson
- Anders Karlen
- Wesley Schaal

- **Ting Wang (EML)**

- Angel Ortiz (Mt Sinai School Medicine, New York)
- Sanja Tomic (Rudjer Boskovic Inst, Zagreb)
- Jan Kmunicek, Jiri Damborsky (Brno, Czech Republic)
- Federico Gago (Univ. Alcala de Henares, Madrid)

- Gabriele Cruciani, Manolo Pastor (Univ. Perugia)
- Peter Goodford (Univ. Oxford)

Lecture 1: Overview

- Rationally designed drugs in the clinic
- How? Sometimes via:
  - Structure-based drug design
    - Dynamic-structure-based-specific-drug-design
  - GRID: Anti-influenza agents
- Beyond the first compound
  - ADMET
  - Resistance
  - Patents
  - New targets
- QSARs
  - COMBINE Analysis

Drug design strategies

- Natural products:
  - Observation (often accidental) of their biological effects
- Screening organic compounds:
  - Serendipity very important
- Template design based on natural effector molecules (neurotransmitters, hormones) or enzyme substrates:
  - Design new receptor agonists or antagonists
  - Design enzyme inhibitors as transition-state analogues
- **Structure-based design:**
  - **Compounds rationally designed to bind to 3D structure of macromolecular receptor**
- Combinatorial library design:
  - Screen designed library of compounds
  - Library can be designed using receptor structure

The Structure-based Approach to Drug Design?

- **Drug design and development is multidisciplinary**
- **Different computational methods should be integrated**
- **Structure-based and combinatorial library approaches are complementary, synergetic**
- **Drugs are not simply ligands, they must be bioavailable and have suitable ADMET** (Absorption, Distribution, Metabolism, Excretion, Toxicity) **properties**

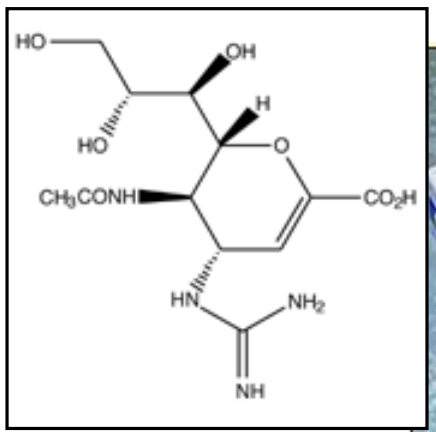
Clinical drugs designed by structure-based approaches

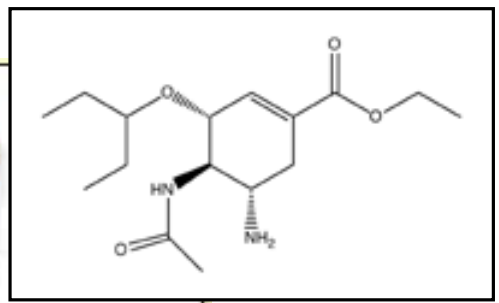
Some successes:

- **1970s: Antihypertensive angiotensin-converting enzyme (ACE) inhibitor: Captopril** *(Capoten, Lopirin, BMS)*
- **1995: Antiglaucoma carboanhydrase inhibitor: dorzolamide** *(Trusopt, Merck)*
- **1995-7:HIV protease inhibitors: saquinavir** *(Invirase, Roche)*, **indinavir** *(Crixivan, Merck)*, **ritonavir** *(Norvir, Abbott)*, **nelfinavir** *(Viracept, Agouron)*
- **1999: Anti-influenza neuraminidase inhibitors: zanamivir** *(Relenza, Biota/Glaxo Wellcome)*, **Oseltamivir** *(Tamiflu, Gilead/Hoffmann-La Roche)*

Drugs against influenza

- Two drugs acting specifically against influenza came into the clinic in the last three years that were designed by structure-based drug design
  - GG167/Relenza/Zanamivir
  - GS4104/Tamiflu/Oseltamivir

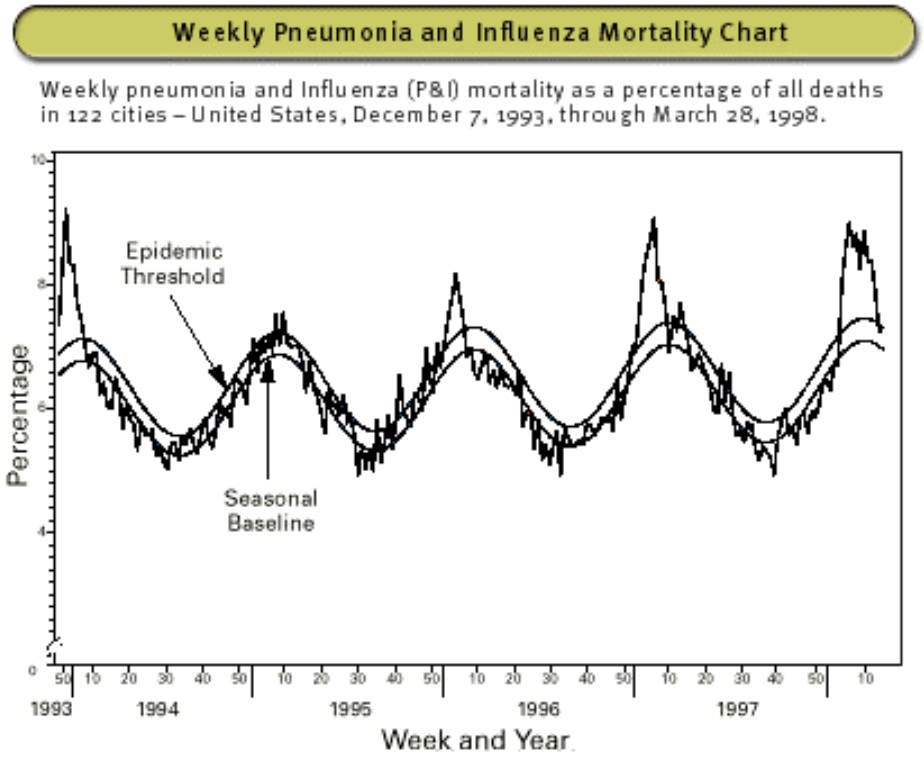




Influenza

‘Flu affects large #s of people

- Economic toll
- Secondary infections can be serious and lethal
- **Pandemics** are possible
  - ***1918***: Spanish ‘flu led to more than 20 million deaths (> 2x # of deaths due to WWI)
  - ***1997:*** lethal ‘flu variant in Hong Kong → 6 deaths. Pandemic was contained by identifying and destroying the source – infected chickens, ducks and geese.



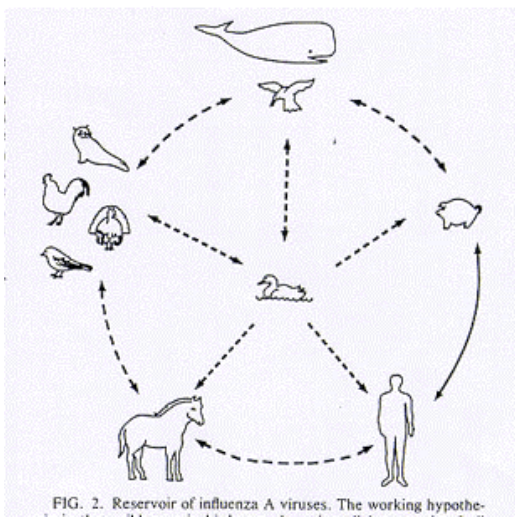
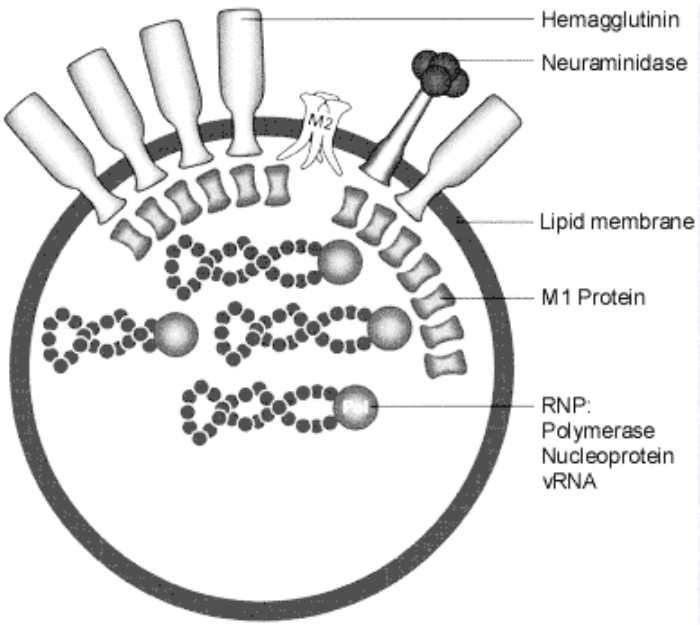


FIG. 2. Reservoir of influenza A viruses. The working hypothe-

Influenza is always evolving



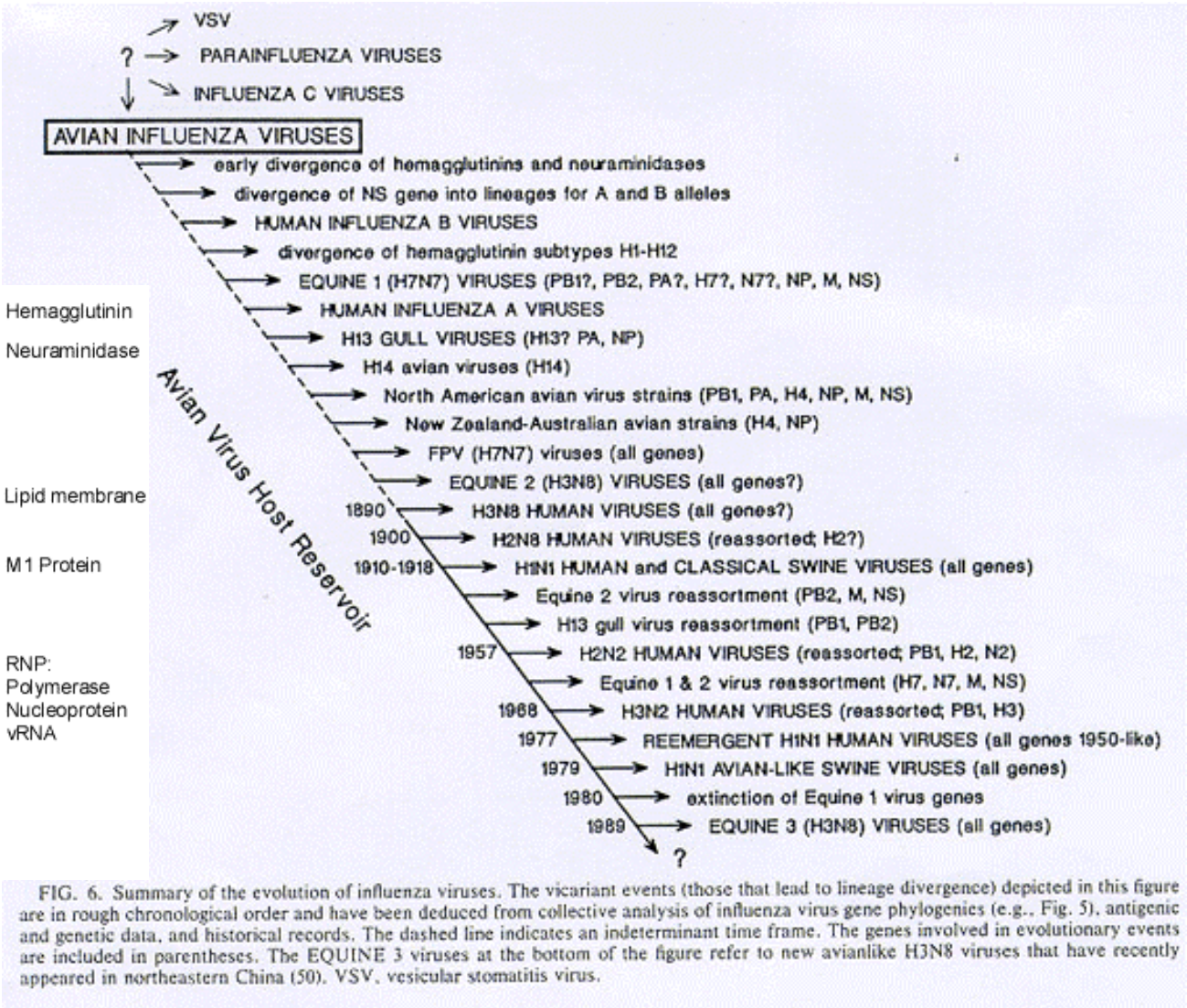
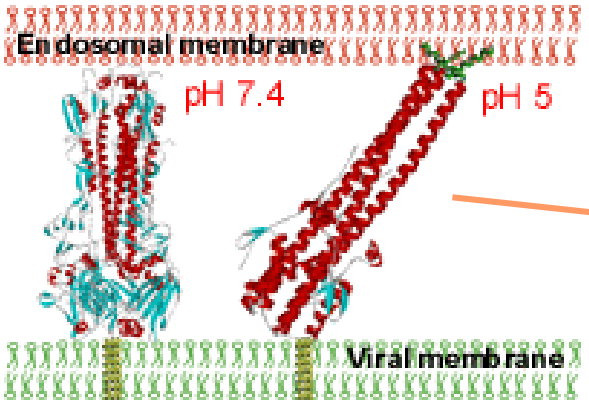
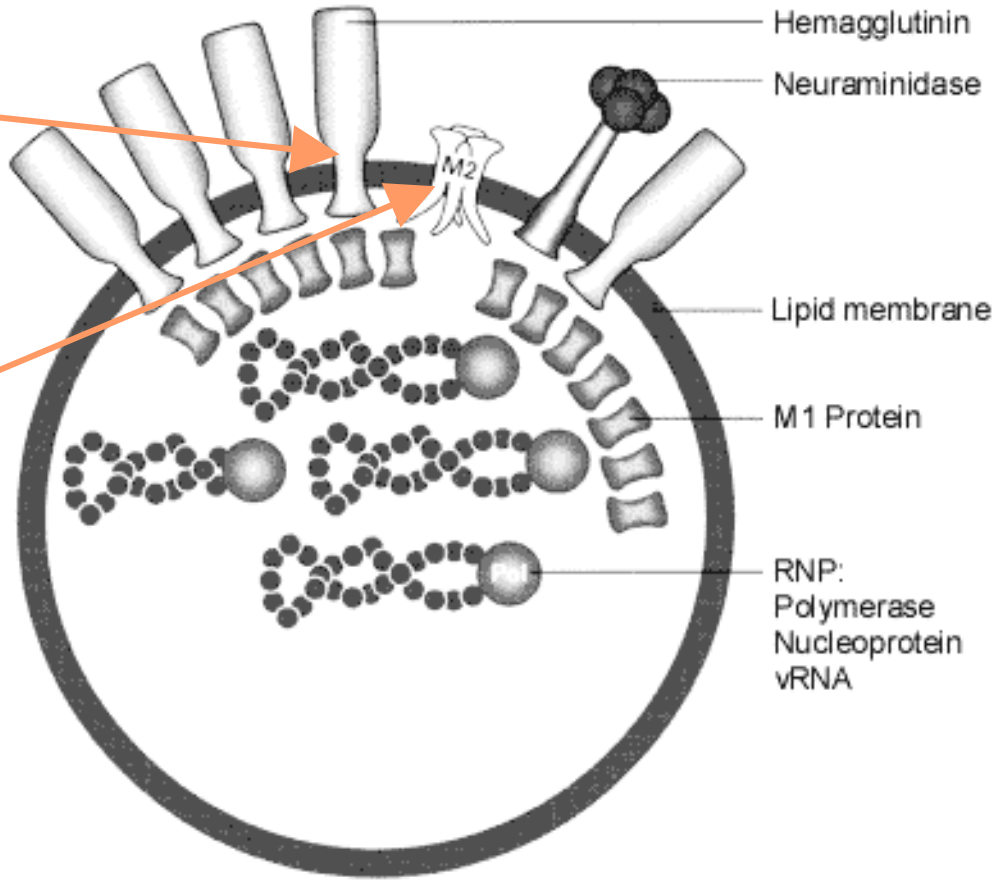


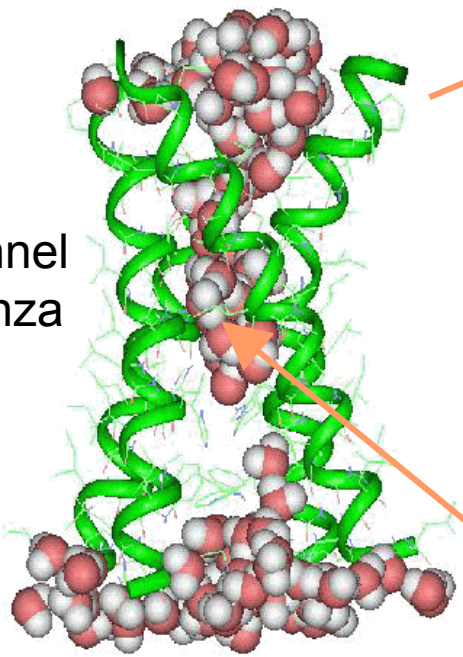
FIG. 6. Summary of the evolution of influenza viruses. The vicariant events (those that lead to lineage divergence) depicted in this figure are in rough chronological order and have been deduced from collective analysis of influenza virus gene phylogenies (e.g., Fig. 5), antigenic and genetic data, and historical records. The dashed line indicates an indeterminant time frame. The genes involved in evolutionary events are included in parentheses. The EQUINE 3 viruses at the bottom of the figure refer to new avianlike H3N8 viruses that have recently appeared in northeastern China (50). VSV, vesicular stomatitis virus.

Influenza virus design targets



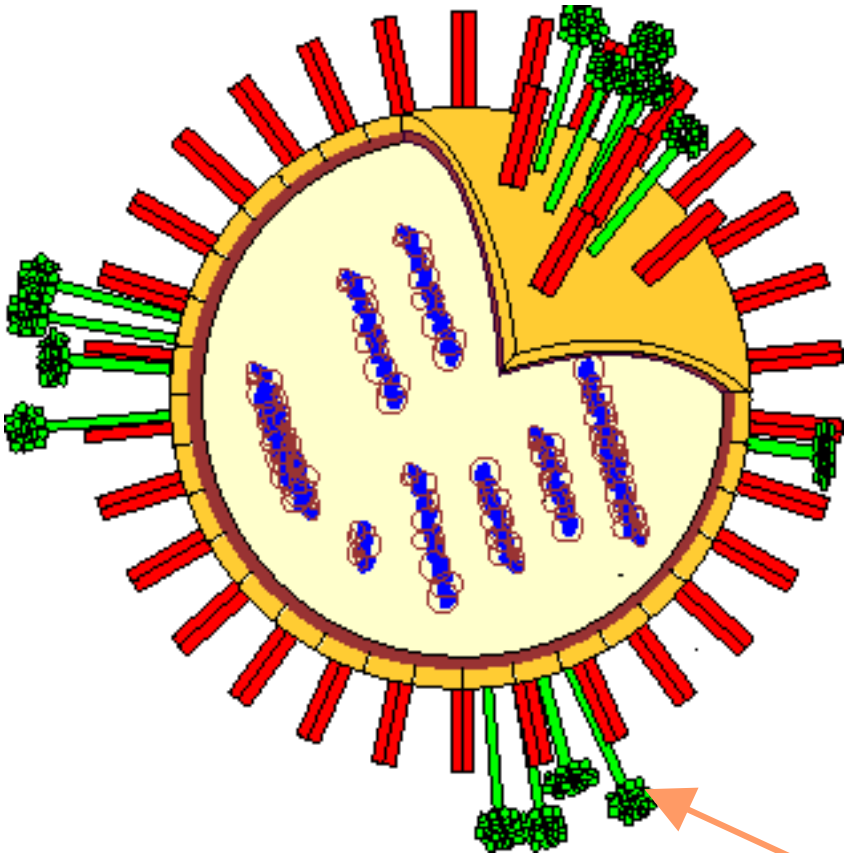


M2 channel in influenza A



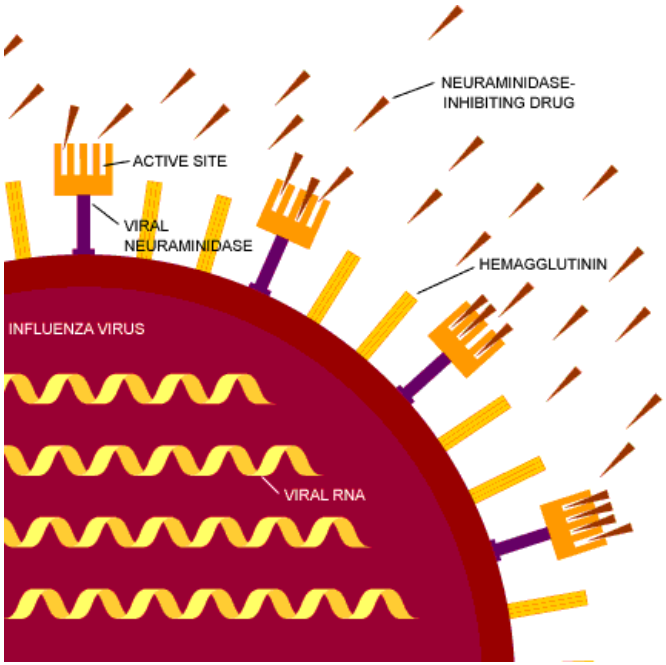
Amantidine & rimantadine block the M2 channel, prevents acidification of viral interior, inhibits viral uncoating

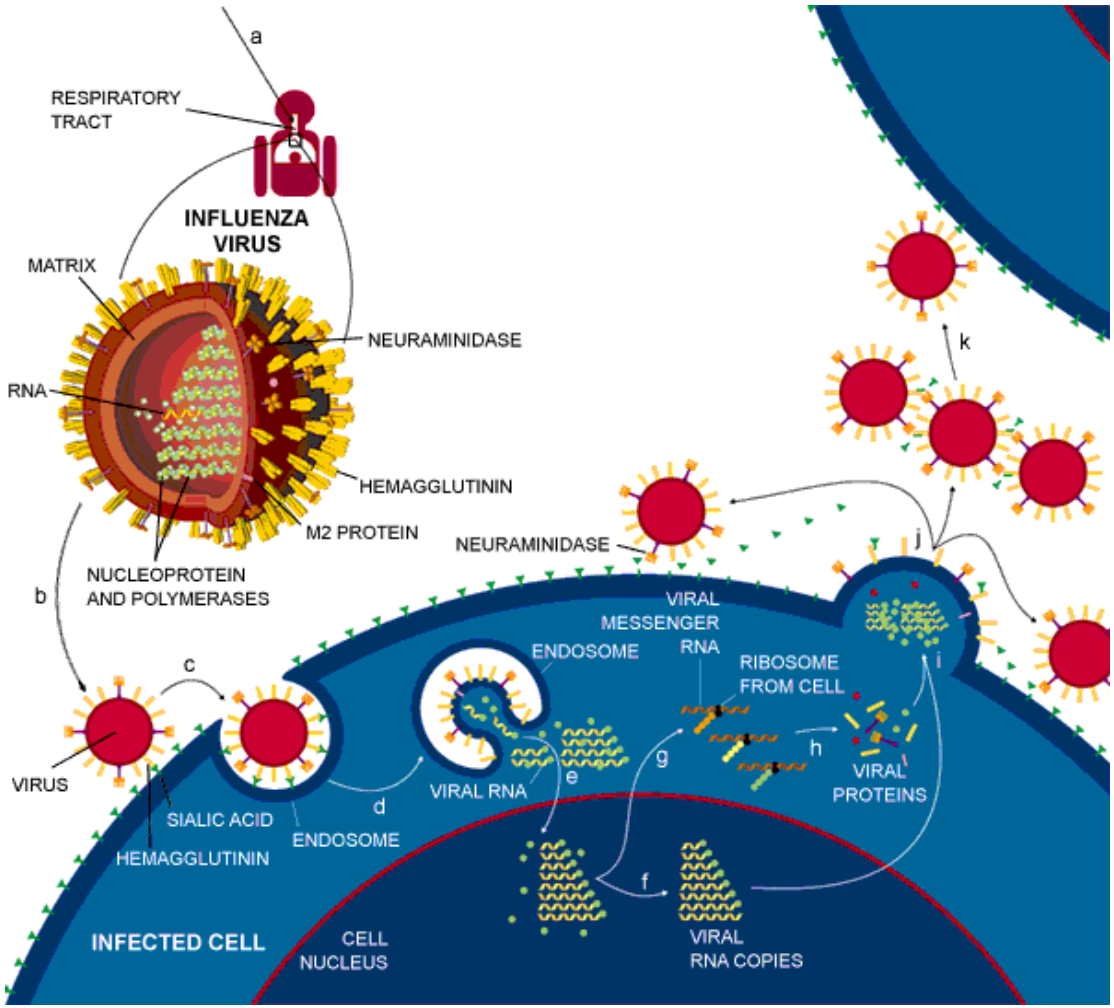
Influenza virus: drug target



Neuraminidase

Influenza neuraminidase inhibition





Laver, G., Sci. Am.

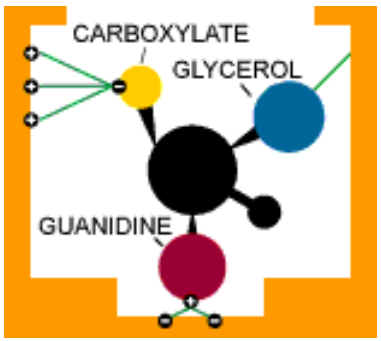
Highly abridged & selective chronology of structure-based drug design

1957	1st protein structure determined – Myoglobin *(Kendrew)*
1973	Structure-based design of haemoglobin ligand – allosteric effector *(Beddell & Goodford)*
1981	Crystal structure of influenza haemagglutinin *(Wiley)*
1983	Crystal structure of influenza neuraminidase *(Colman)*
1985	GRID method for predicting binding sites *(Goodford)*
1986	DOCK method for docking ligands *(Kuntz)*
1993	SBDD of influenza sialidase inhibitors *(von Itzstein)*
1999	1st anti-influenza drug in the clinic

Abridged chronology of neuraminidase inhibitor design

- 1966 **Random screening** $k_i \sim 10^{-4}$ M
- 1969 Sialic acid binding $k_i \sim 10^{-3}$ M
- 1974 **Mechanism-based design**
  Transition-state analogues $k_i \sim 10^{-6}$ M
- 1983 Crystal structure
  Structure-based drug design
- 1993 4-guanidino-Neu5Ac2en $k_i \sim 10^{-10}$-10^{-11} M
  (GG167/Zanamivir/Relenza)
- 1997 Optimization/Core redesign
  (GS1404/Oseltamivir/Tamiflu)
- 1999 1st Drug in the clinic

Chronology of influenza neuraminidase inhibitor design



Laver, G., Sci. Am.

1966 **Random screening**
Compound databases → *N*-phenyl oxamic acid + others
$k_i \sim 10^{-4}$ M

1969 **Mechanism-based design**
Sialic acid/ Neu5Ac
$k_i \sim 10^{-3}$ M
Transition-state analogue
Neu5Ac2en
$k_i \sim 10^{-6}$ M
C5-substitution
FANA
$k_i \sim 10^{-6}$ M

1974

1983 **Structure-based design**
Crystal structure of NA
C4-substitution of Neu4Ac2en
4-guanidino-Neu5Ac2en (GG167)
$k_i \sim 10^{-10}$ -10^{-11} M

1993

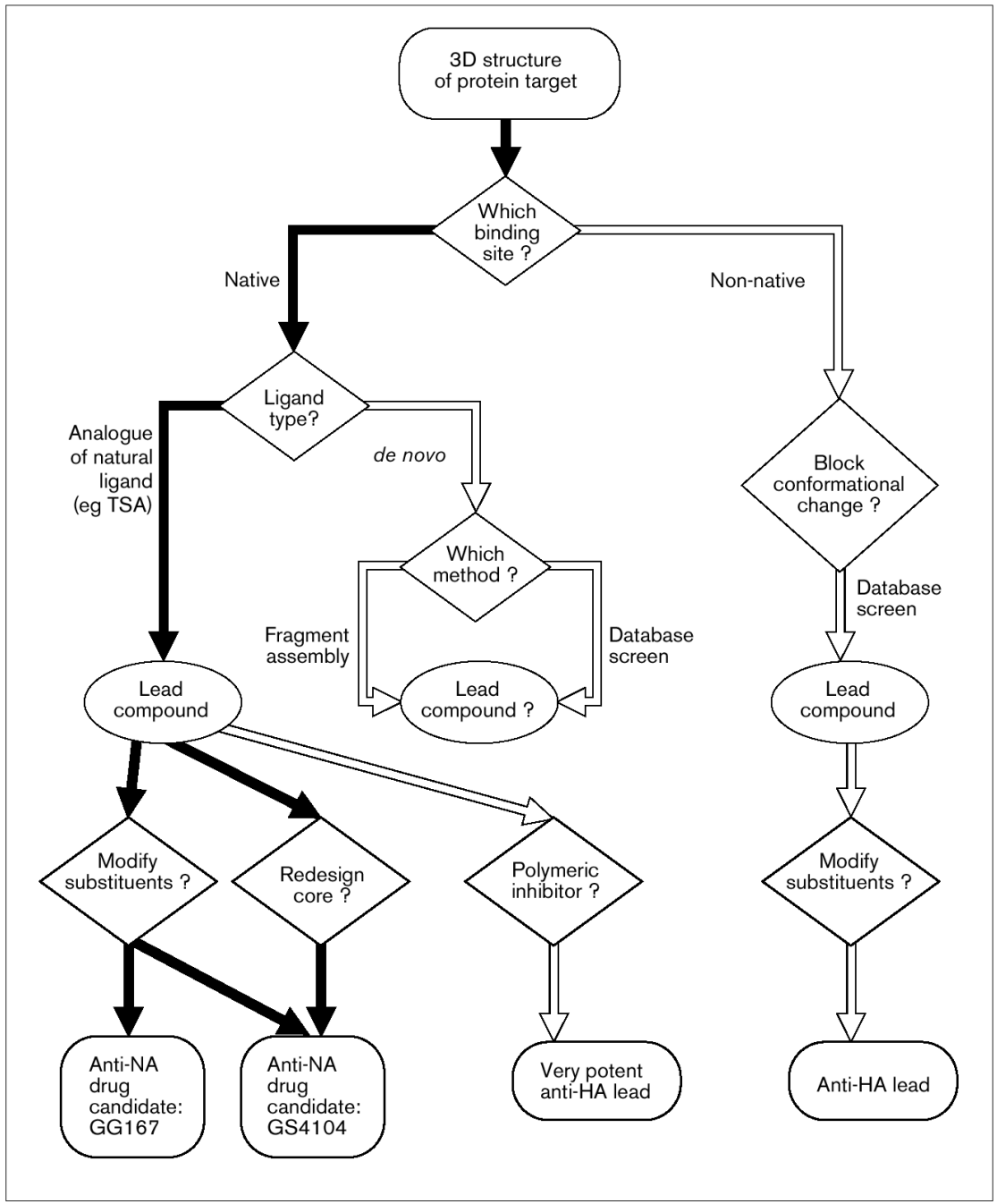
Benzoic acid analogues
BANA113

Carbocyclic analogues with lipophilic substituent
GS4104

1997

Wade,RC *Structure* (1997) **5**, 1139–1145

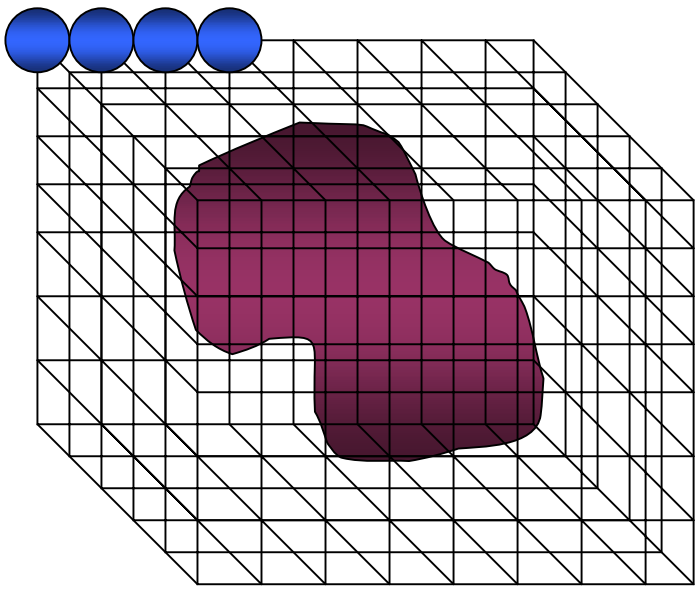
SBDD strategies



Wade,RC *Structure* (1997) **5**, 1139–1145

GRID

- Detect energetically favorable binding sites on molecules of known structure



$$\Delta E = \sum_i E_{LJ} + \sum_i E_{EL} + \sum_i E_{HB} + S$$

Goodford, PJ *J. Med. Chem.* (1985) 28, 849-857.

GRID: Some uses

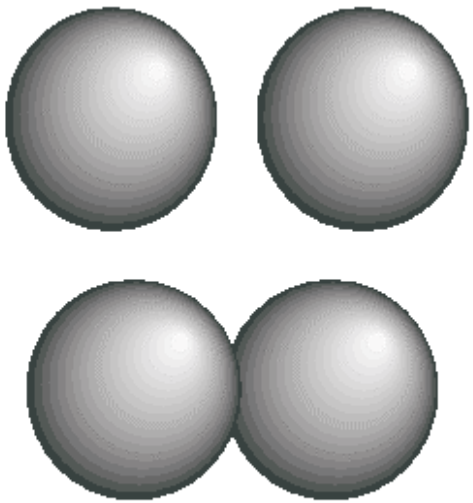
- Structure-based drug design
- Locate ordered water molecule network, e.g. prior to energy minimization/molecular dynamics
- Designing molecules to accommodate/displace interfacial water molecules
- Designing selective molecules (comparing GRID maps)
- Compute molecular interaction fields (MIFs) for generating structure-activity relationships (SARs) for sets of molecules
  - For small molecules (GRID/GOLPE, cf CoMFA)
  - **For macromolecules (PIPSA)**
- Docking molecules (with GROUP)

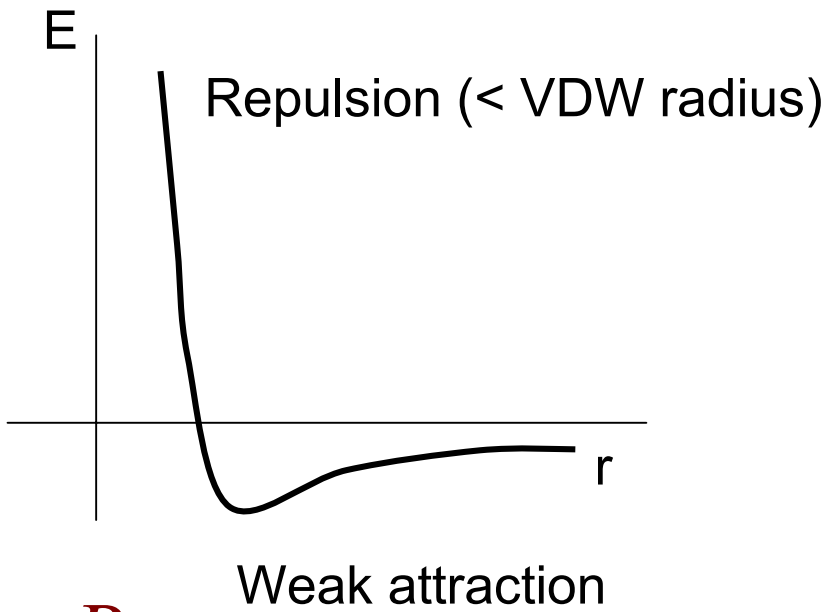
GRID Energy Function

- Empirical, based on observed geometries in small molecule and protein crystals

$$\Delta E = \sum_i E_{LJ} + \sum_i E_{EL} + \sum_i E_{HB} + S$$

Non-bonded interactions: Lennard-Jones

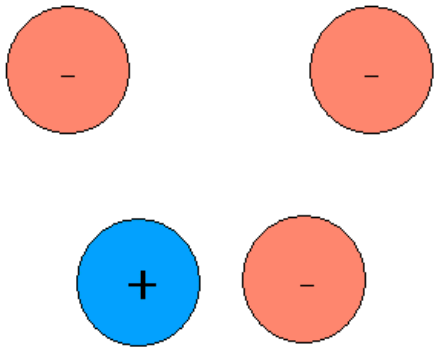




$$E_{LJ} = \frac{A}{r^{12}} - \frac{B}{r^{6}}$$

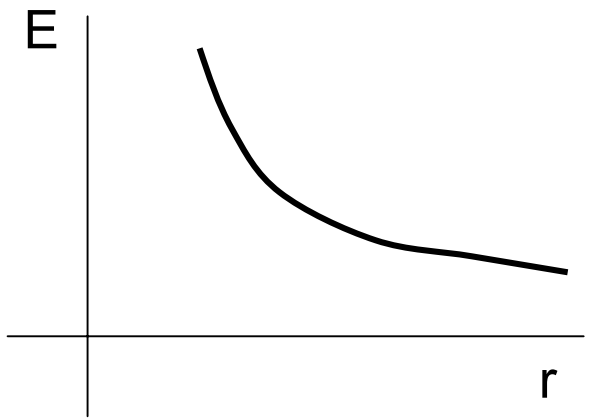
- van der Waals excluded volume repulsion
- London dispersive induced-dipole induced-dipole attraction
- Short-range (8-10 A cutoff)

Non-bonded interactions: Electrostatic



Repulsion

Attraction



$$E_{EL} = \frac{q_i q_j}{K\zeta}\left[\frac{1}{r} + \frac{(\zeta-\varepsilon)/(\zeta+\varepsilon)}{\sqrt{r^2 + 4s_{q_i}s_{q_j}}}\right] \qquad \varepsilon = 80 \qquad \zeta = 4$$

- Each atom has partial atomic charge
- Coulombic forces
- Long range (~1/r) (No cutoff)
- Screened by dielectric medium (~1/ε)
  - Method of images – assuming planar dielectric boundary

GRID Energy Function

- Empirical, based on observed geometries in small molecule and protein crystals

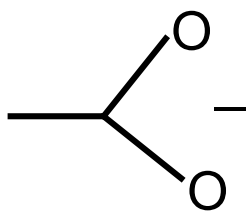
$$\Delta E = \sum_i E_{LJ} + \sum_i E_{EL} + \sum_i E_{HB} + S$$

$$E_{LJ} = \frac{A}{r^{12}} - \frac{B}{r^{6}} \qquad E_{HB} = \left[\frac{C}{r^{8}} - \frac{D}{r^{6}}\right] f(\Theta, \Phi...) f'(Q)$$

$$E_{EL} = \frac{q_i q_j}{K\zeta}\left[\frac{1}{r} + \frac{(\zeta - \varepsilon)/(\zeta + \varepsilon)}{\sqrt{r^2 + 4 s_{q_i} s_{q_j}}}\right] \qquad \varepsilon = 80 \qquad \zeta = 4$$

GRID probes

- Most are appropriately parameterized single-point spheres
  - E.g. carbonyl oxygen, hydroxyl, amino, water, dry, sulfate
- Some are multiatom
  - E.g. carboxylate, amide



- GROUP:
  - Docking (rigid body) of whole molecules using 3+ GRID probe maps

GRID water and DRY probes

- Water
  - Sphere (LJ)
  - H-bonds: 4, orientational dependence

$$\Delta E = \sum_i E_{LJ} + \sum_i E_{HB}$$

- Dry
  - Sphere (LJ)
  - minus H-bonds: as water
  - Water entropy
    - Displacement of an ordered water molecule to bulk is assumed to result in reduction in its # of H-bonds from 4 to 3, permitting 4 H-bond permutations (123, 234, 341, 412): Went = RTln(4)= -0.848 kcal/mole

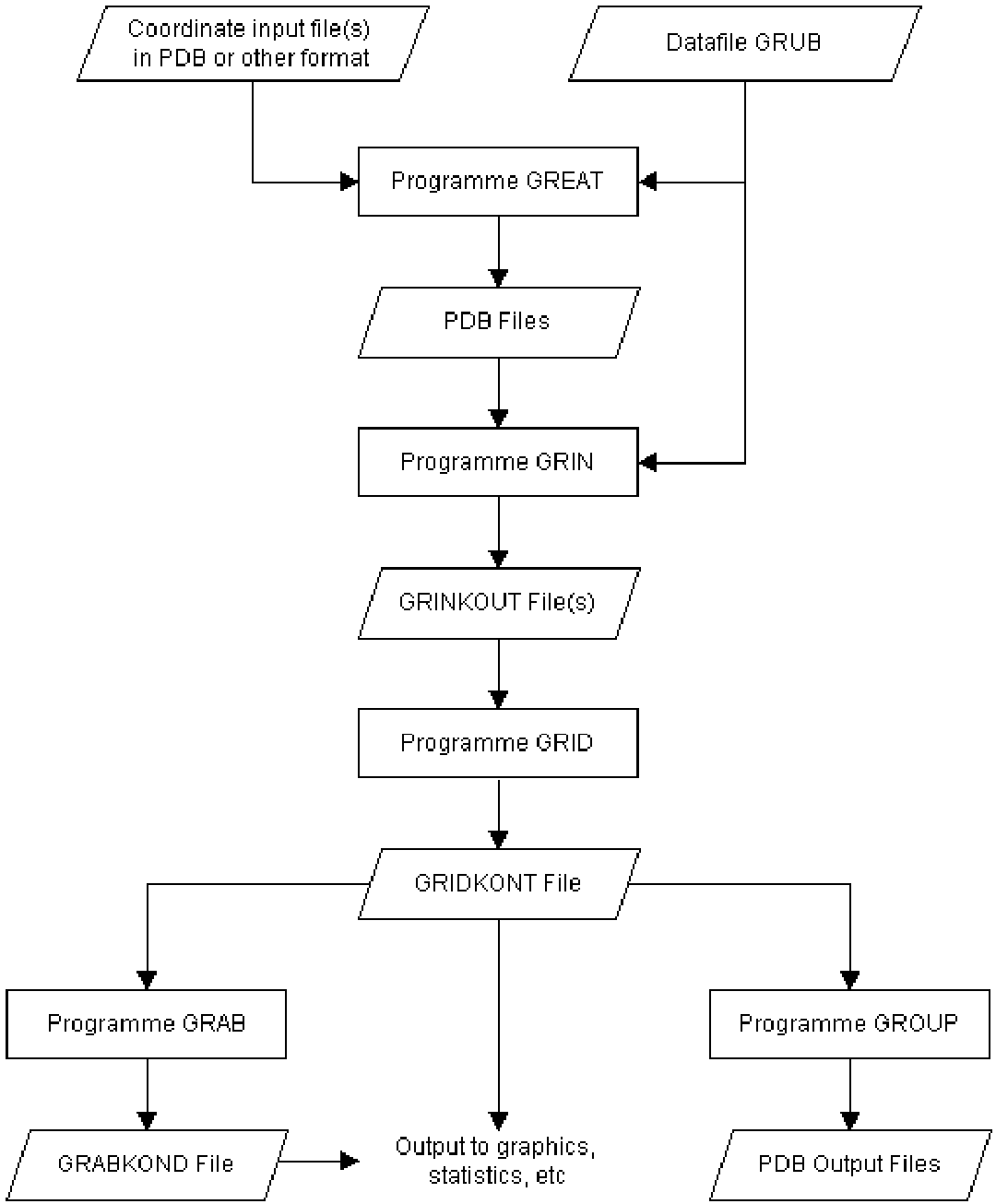
$$\Delta E = W_{ENT} + \sum_i E_{LJ} - \sum_i E_{HB}$$

Receptor flexibility in GRID

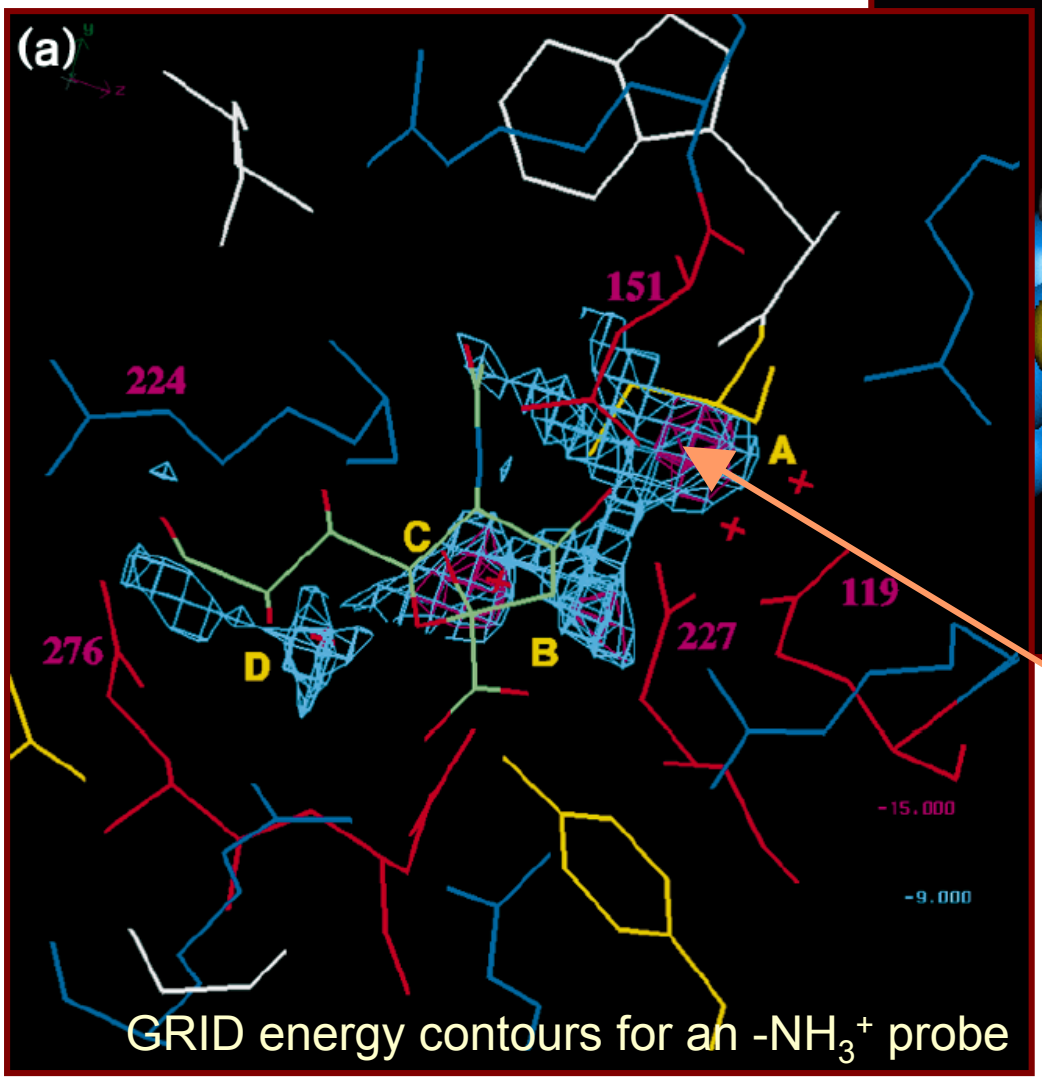
- Hydrogens
  - Polar Hs added to protein by GRID
  - Can adapt positions according to probe location
  - Histidine can change tautomer according to probe location, if desired
- Water molecules (LEAU)
  - Can be positioned between probe and target to investigate whether that would improve binding
- Side-chains (MOVE>0)
  - Can adopt position according to probe location and type
- Counterions (MOVE = -1, >0)

Running GRID

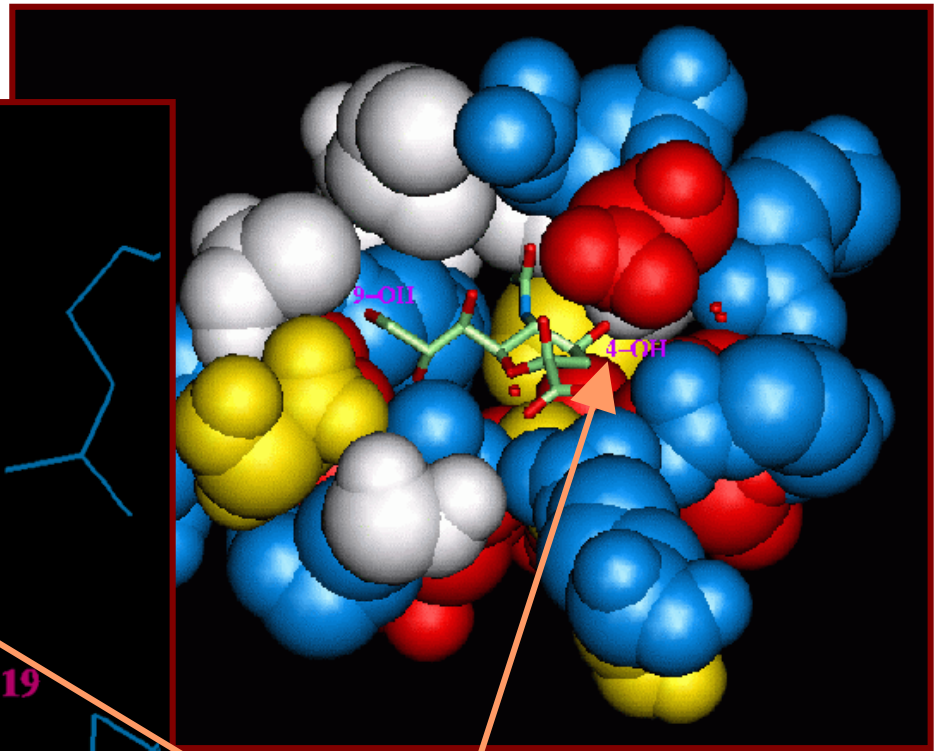
New in GRID20:
A GUI called
Greater



Neuraminidase: finding the right plug

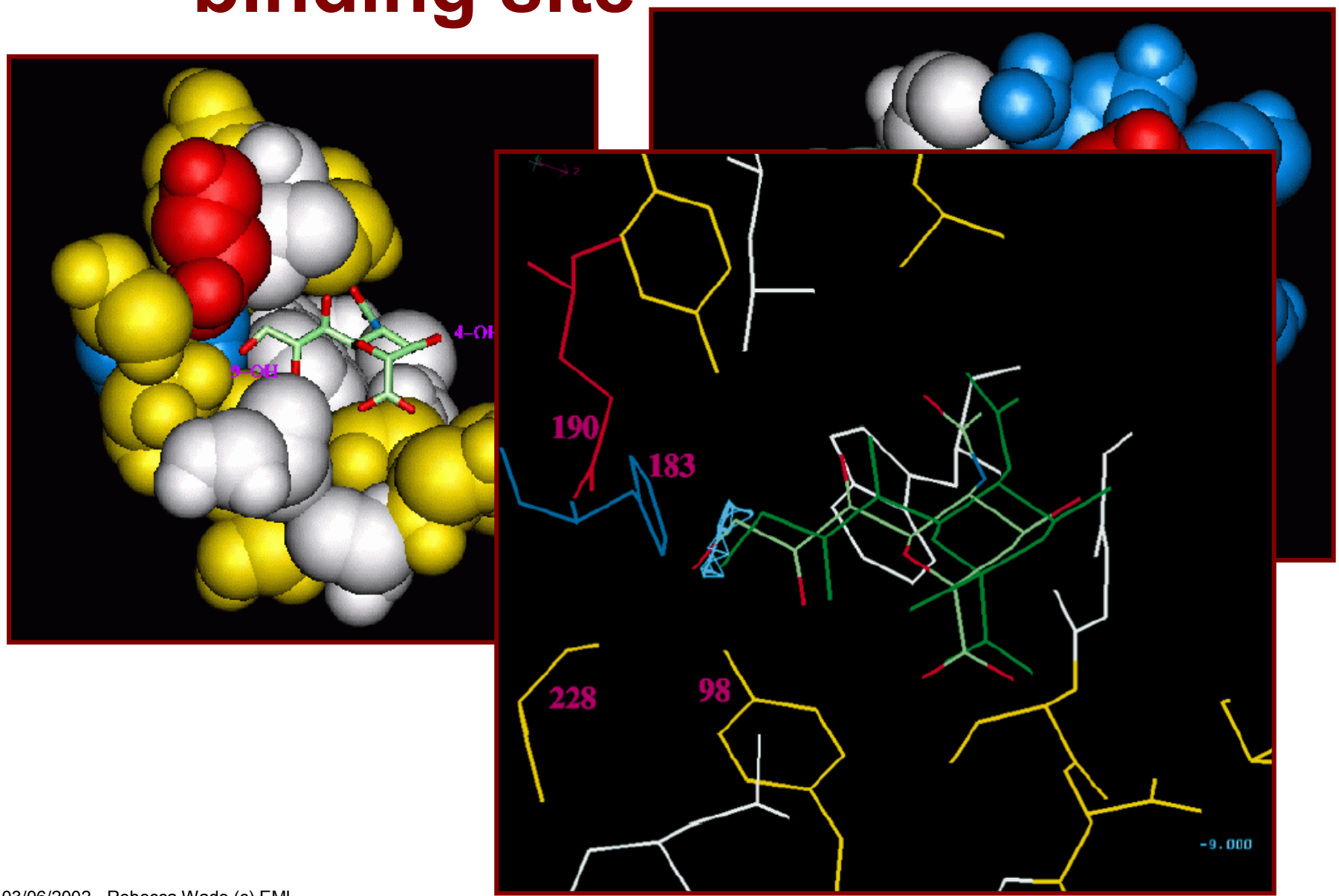


GRID energy contours for an -NH_3^+ probe

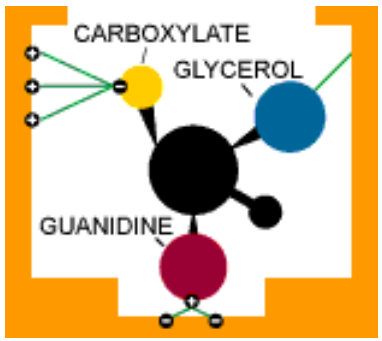


Amino (GS1404)/ guanidino (GG167) group binds here

Haemagglutinin: sialic acid binding site



Chronology of influenza neuraminidase inhibitor design



1966 **Random screening**

Compound databases → *N*-phenyl oxamic acid + others

$k_i \sim 10^{-4}$ M

1969 **Mechanism-based design**

Sialic acid/ Neu5Ac

$k_i \sim 10^{-3}$ M

Transition-state analogue

Neu5Ac2en

$k_i \sim 10^{-6}$ M

C5-substitution

1974

FANA

$k_i \sim 10^{-6}$ M

1983 **Structure-based design**

Crystal structure of NA

C4-substitution of Neu4Ac2en

4-guanidino-Neu5Ac2en (GG167)

$k_i \sim 10^{-10}$ -10^{-11} M

1993

Benzoic acid analogues

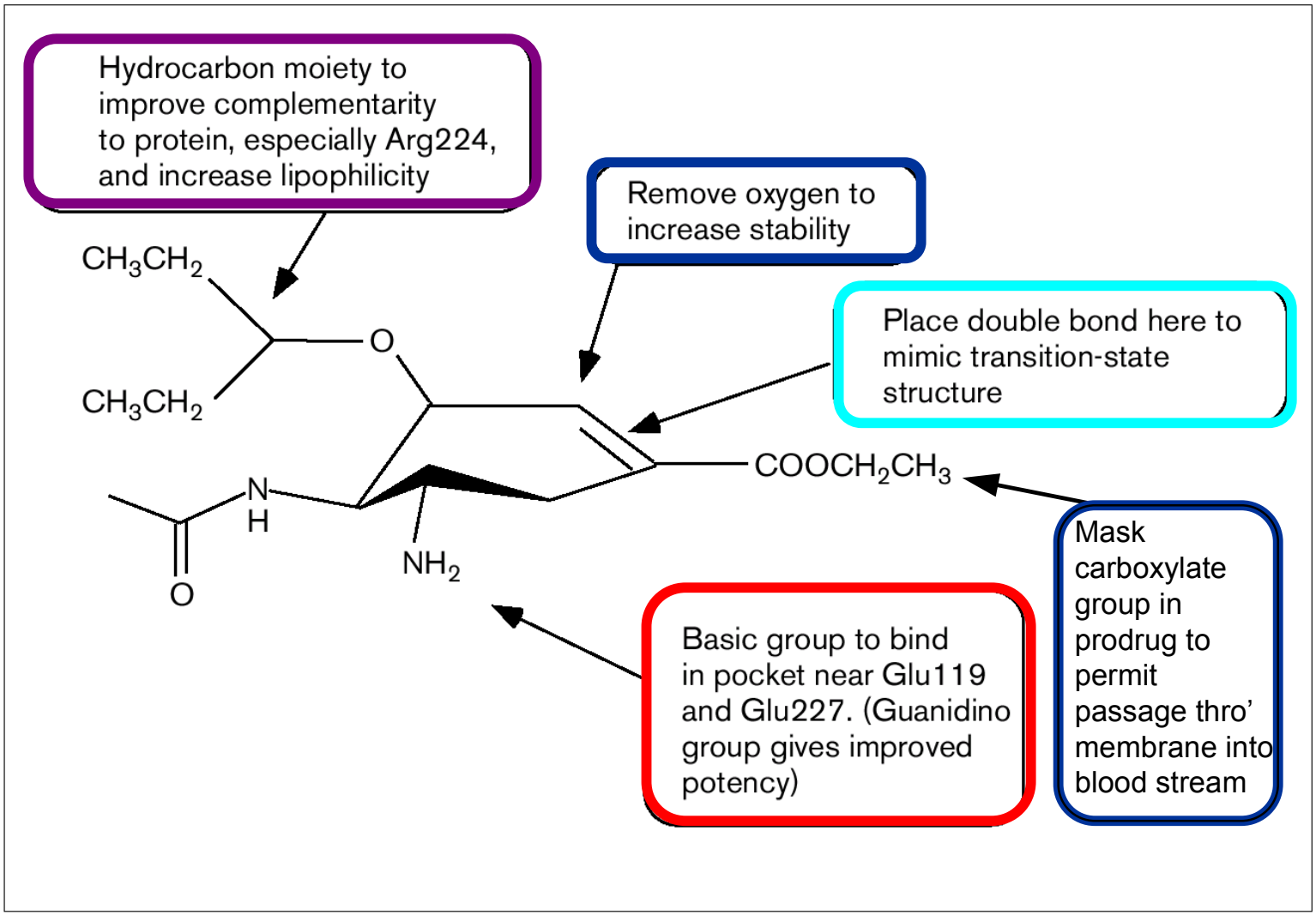
Carbocyclic analogues with lipophilic substituent

1997

BANA113

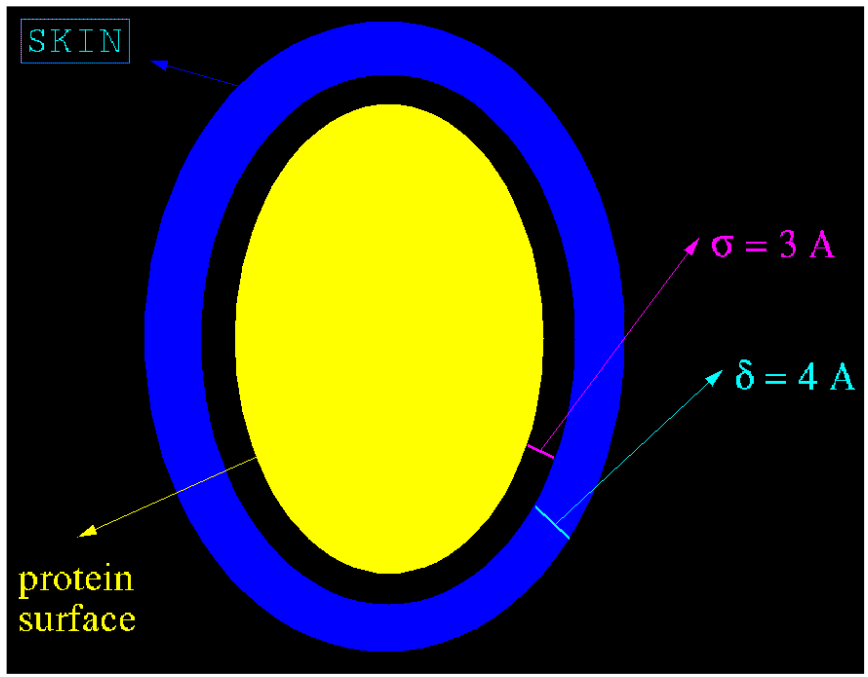
GS4104

Design rationale for drug against influenza (GS4104)



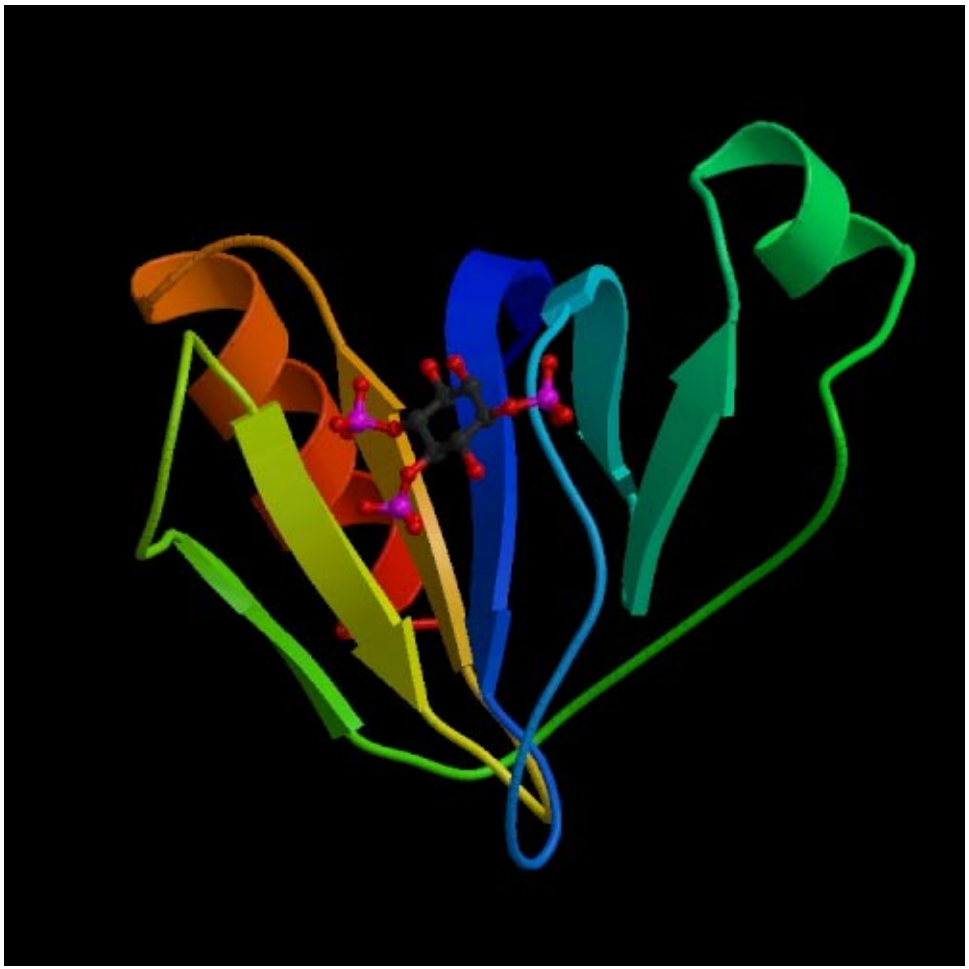
Protein Interaction Property Similarity Analysis (PIPSA)

- Pairwise comparison of molecular interaction fields of proteins *a* and *b*
  - ϕ_a (i,j,k); ϕ_b (i,j,k) at (i,j,k) grid points
- Hodgkin similarity index
  - $SI(a,b) = 2M_aM_b/(M_a^2+M_b^2)$
  - $M_aM_b = \Sigma_{i,j,k}\phi_a(i,j,k)\phi_b(i,j,k)$
- Skin (region where MIFs are compared)
  - $\sigma = 3$Å from the vdw surface
  - $\delta = 4$Å thickness
  - **complete** skin
  - **restricted** regions: binding or ET



Blomberg et al. Proteins 1999;
De Rienzo et al. Protein Sci. 2000
Wade et al. IJQC 2001

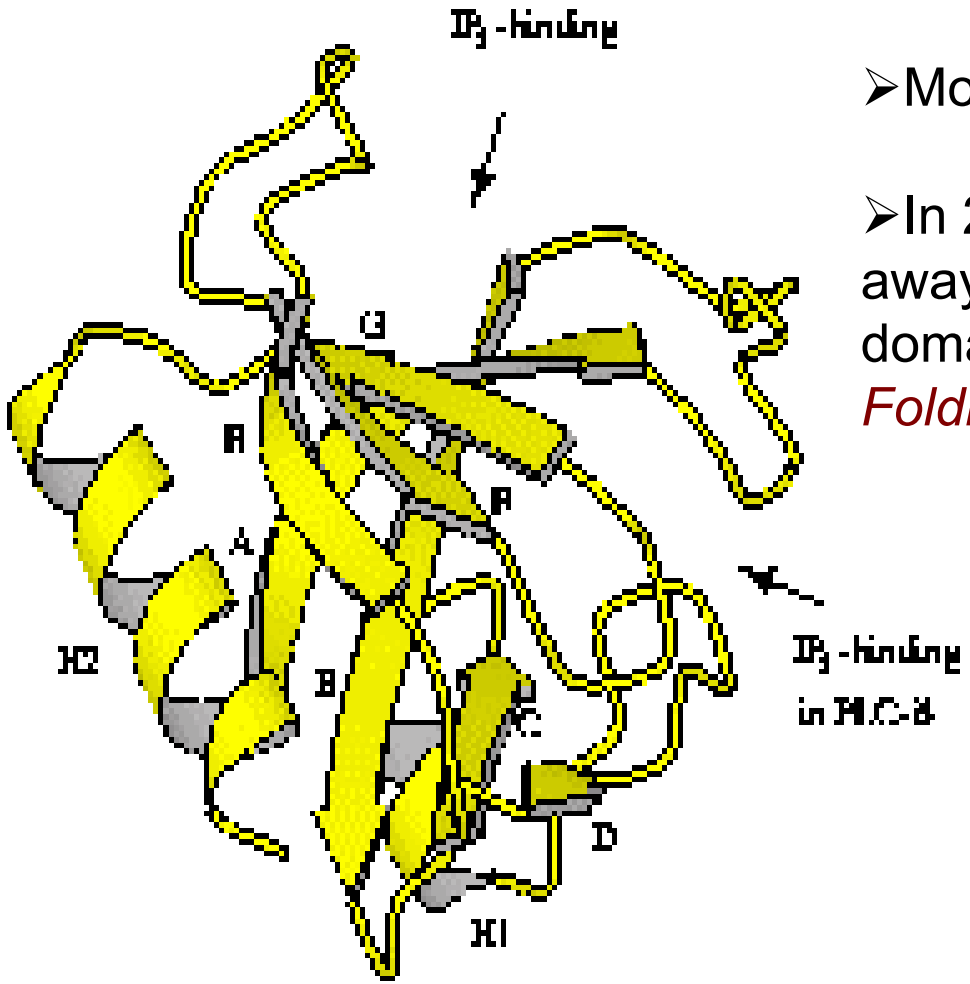
Pleckstrin homology (PH) Domains



- Found in proteins involved in signal transduction or cytoskeletal organization
- Mostly phospholipid membrane binding
- Also protein-protein interactions

Spectrin PH domain with inositol-(1,4,5)-triphosphate (IP_3) bound

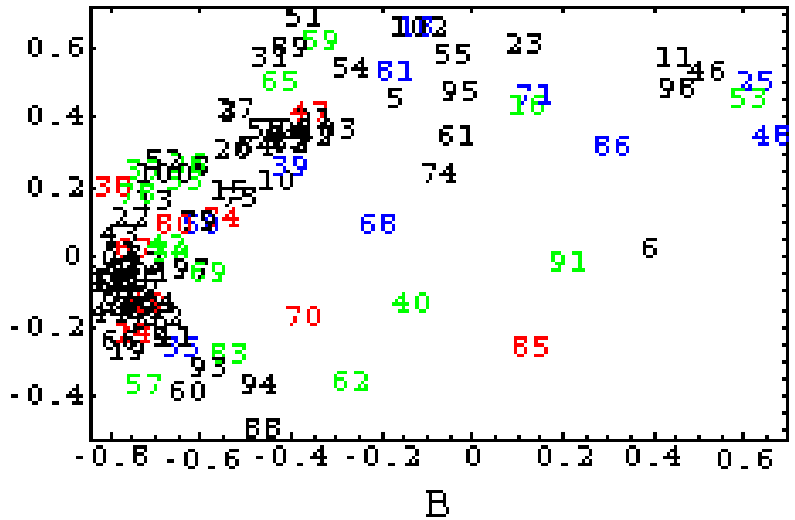
104 PH domains: classification on electrostatic interaction properties



- More than 1 phospholipid binding site
- In 25% of PH domains, dipole points away from sites in spectrin and PLC-δ domains *Blomberg & Nilges (1997) Folding & Design 2, 343-355*

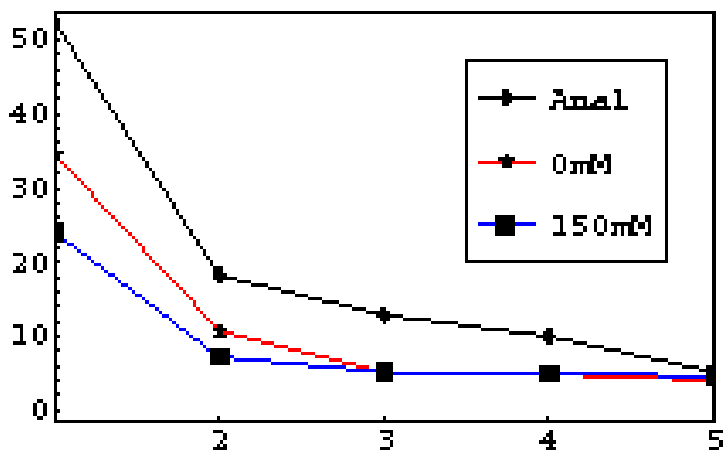
Blomberg, Gabdoulline, Nilges & Wade (1999) Proteins 37:379-87

PH domains - distribution in electrostatic potential similarity space



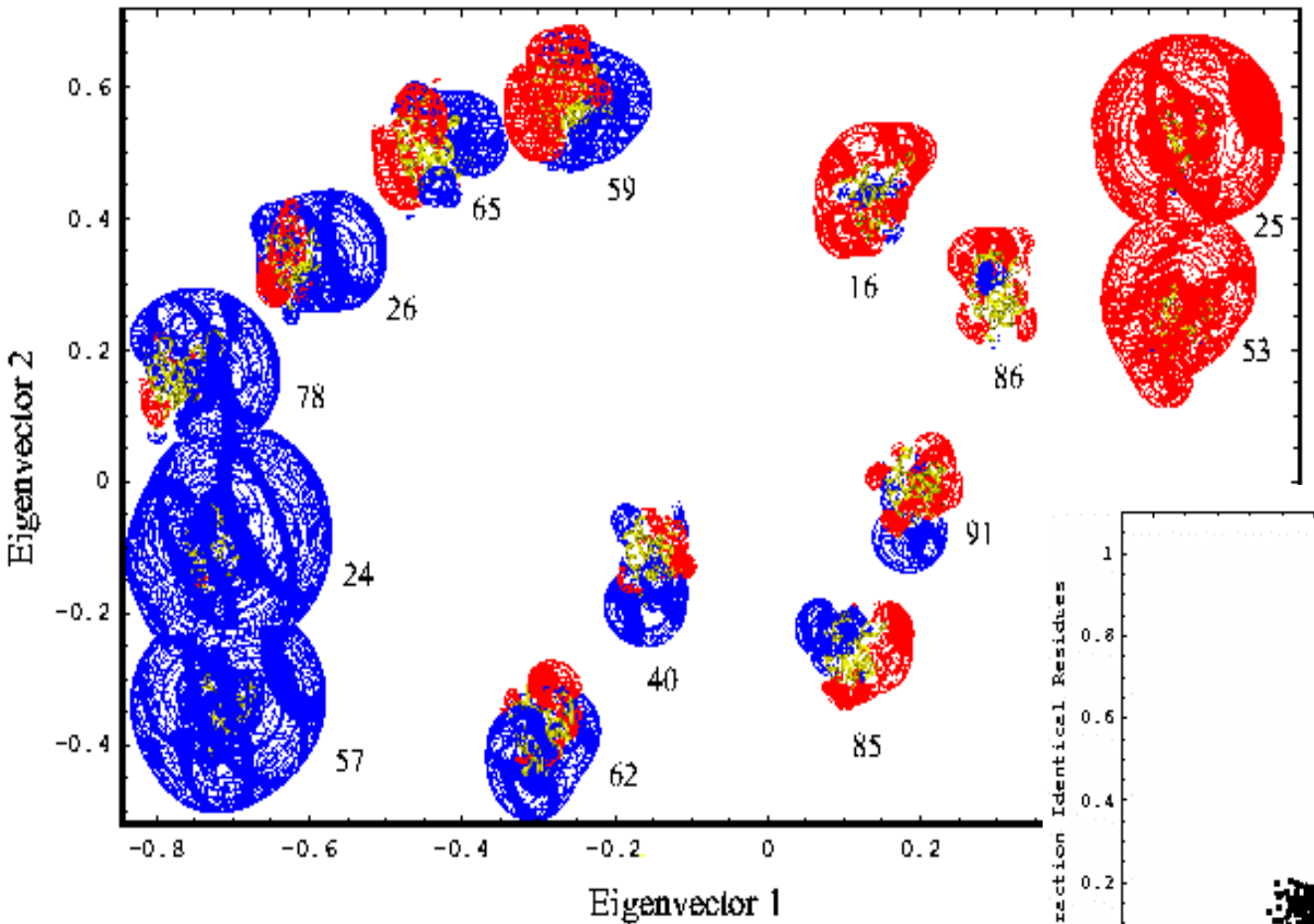
Majority:
conserved +ve

15%, incl.
DH-linked
PH repeats
(Nter, Cter)

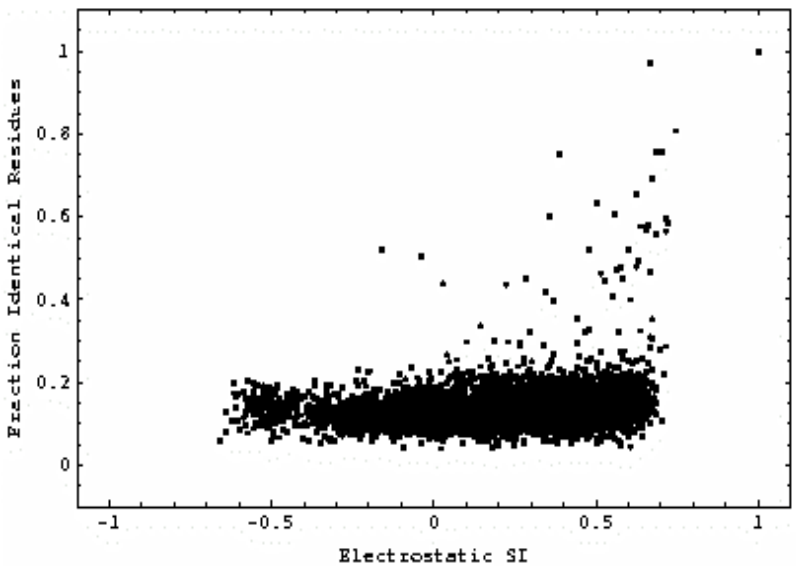


Blomberg, Gabdoulline, Nilges & Wade (1999) Proteins 37:379-87

PH domains - distribution in electrostatic potential similarity space



Distribution for DH-linked and internal PH repeat domains



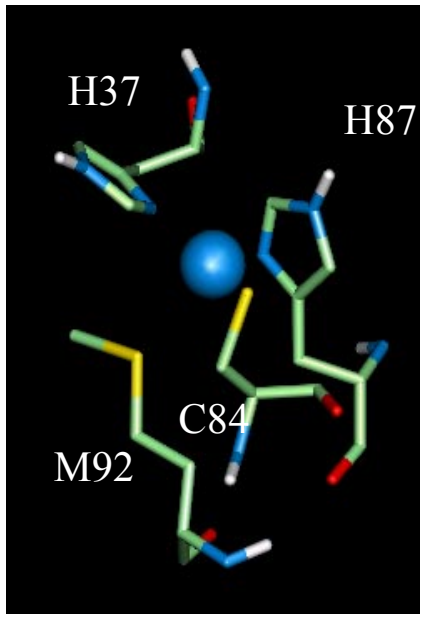
Cupredoxins: Blue Copper Proteins

- **Shuttle electrons from donor to acceptor protein**

$$A_{ox} + B_{red} \overset{k2}{\leftrightarrow} A_{red} + B_{ox}$$

- **Structures (27):**
  - plastocyanins (12), azurins (6), pseudoazurins (4), amicyanins (2), rusticyanin (1), cucumber basic protein (1) & stellacyanin (1)
  - Complexes of cupredoxins bound to their redox partners (2)

- Greek key β-barrel fold
- 10-14 kDa

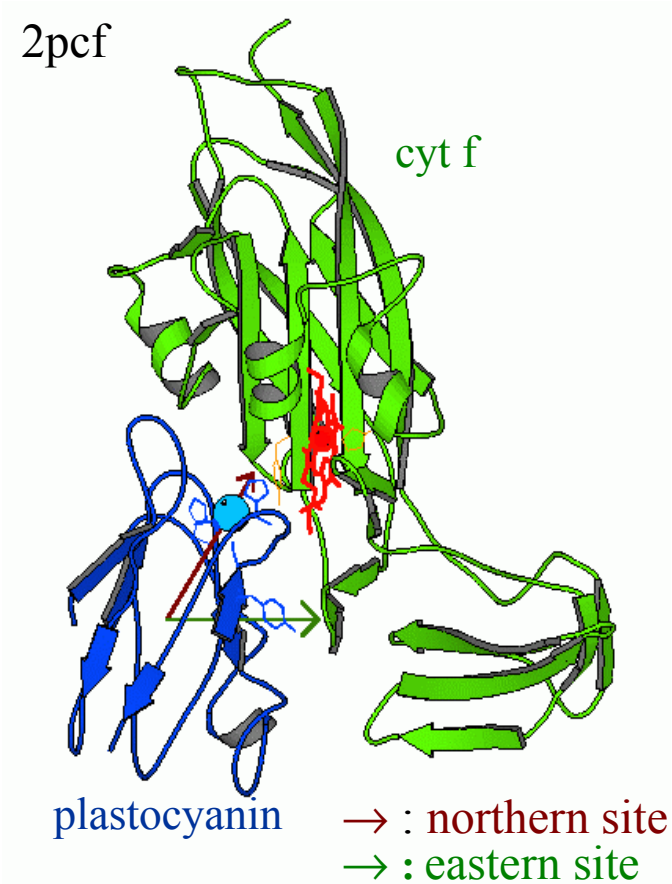


- Type I Cu active site

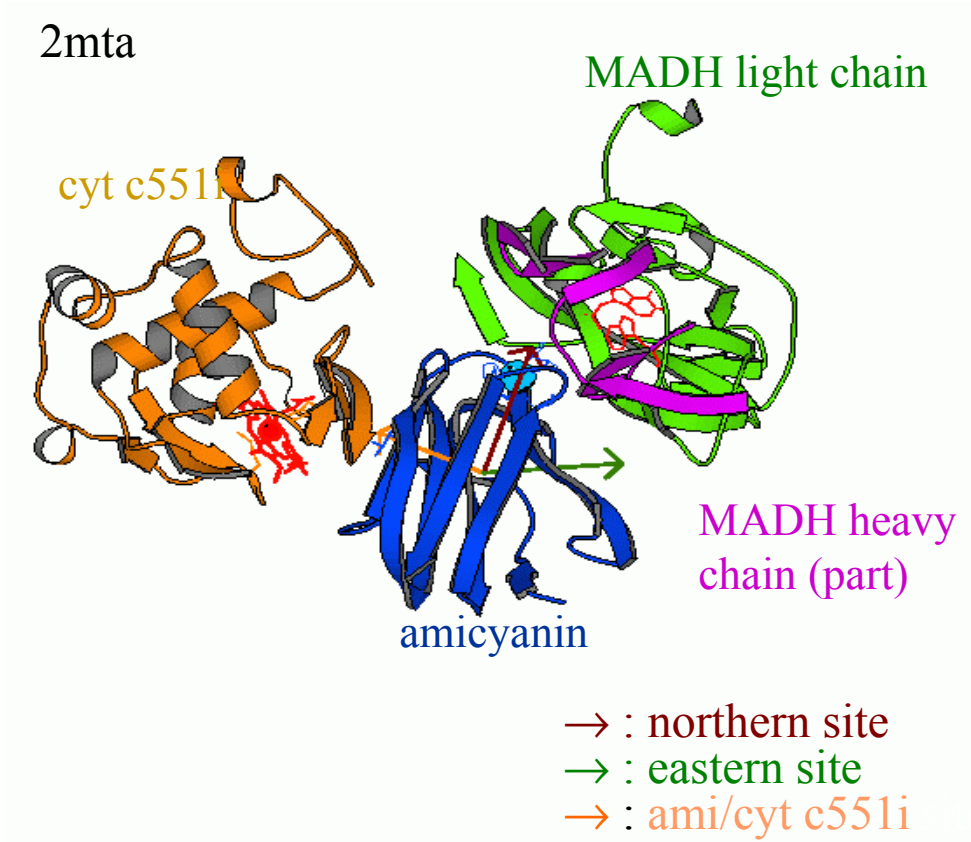
De Rienzo, Gabdoulline, Menziani & Wade Protein Sci. 2000

Cupredoxin Complexes:

plastocyanin/cytochrome f



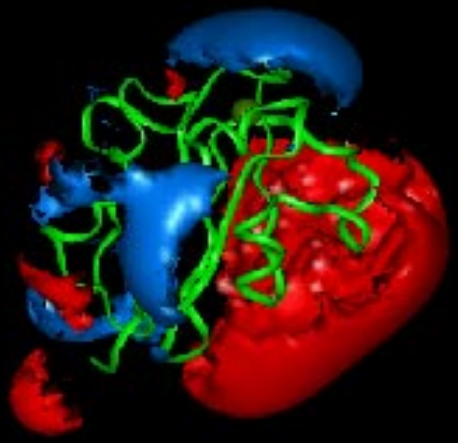
cytochrome c551i/amicyanin/MADH



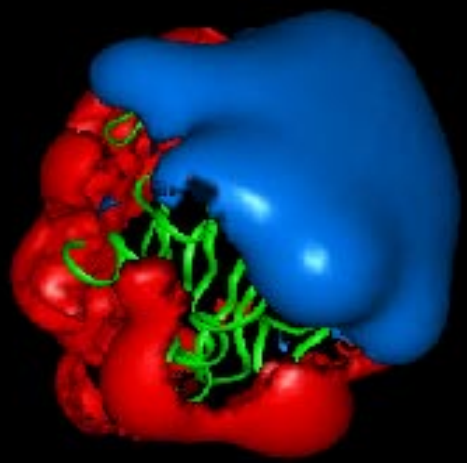
De Rienzo, Gabdoulline, Menziani & Wade Protein Sci. 2000

MEPs - iso-potential contours at +0.5 and -0.5 kcal/mol/e

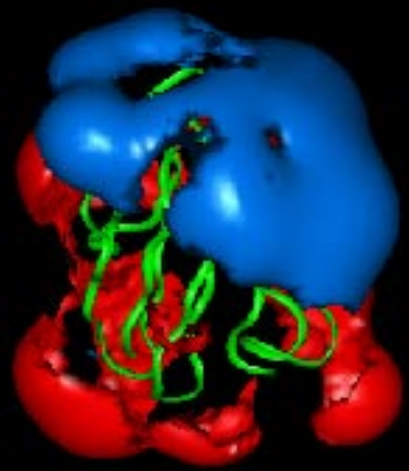
EUK. PLASTOCYANIN



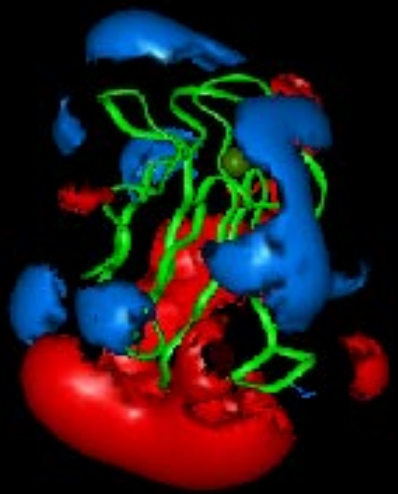
AZURIN



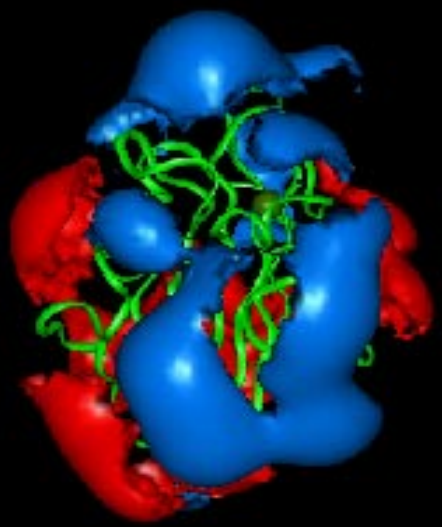
PSEUDOAZURIN



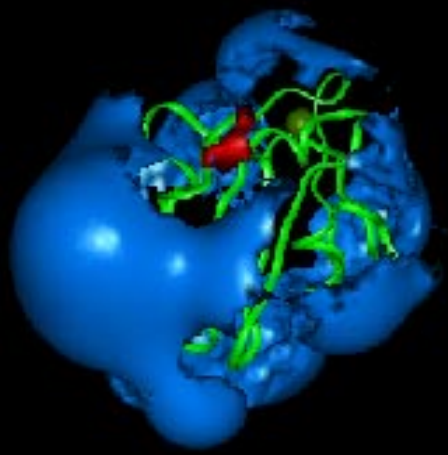
AMICYANIN



BACT. PLASTOCYANIN



RUSTICYANIN

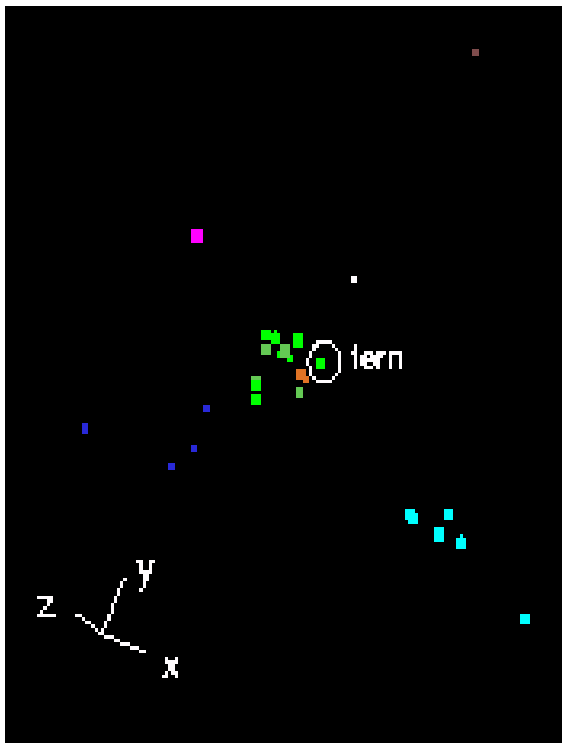


CBP

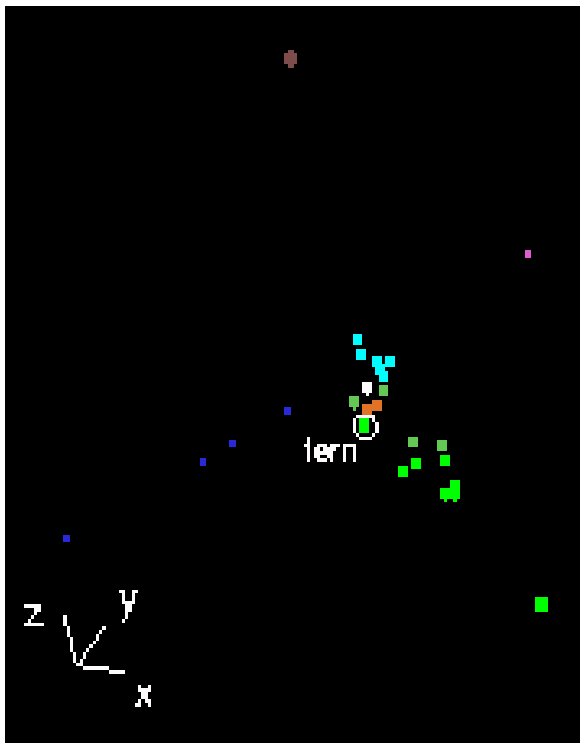
STELLACYANIN

Clustering of cupredoxins by PIPSA

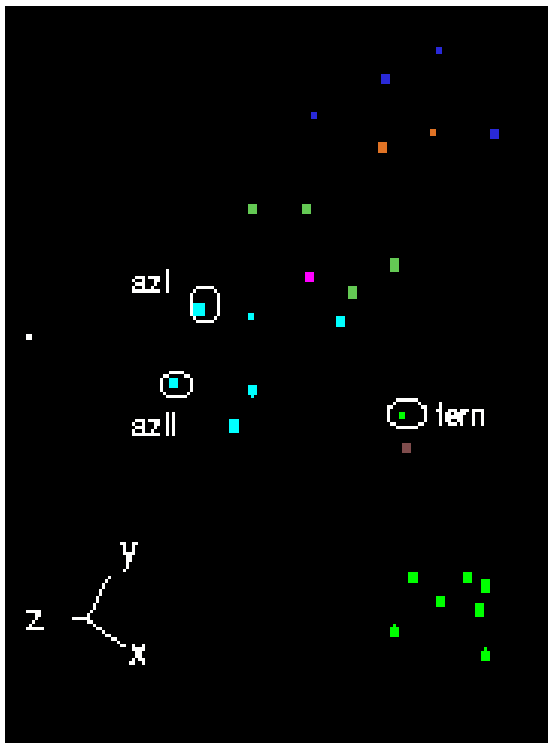
De Rienzo, Gabdoulline, Menziani & Wade Protein Sci. 2000



Pairwise Sequence identity



Hydrophobic potential



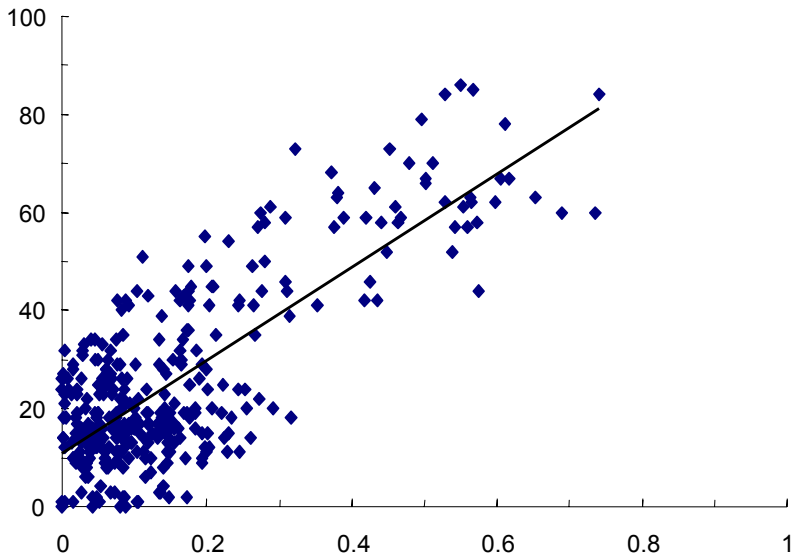
Electrostatic potential (150mM)

light green - cyanobacterial plastocyanins; dark green - eukaryotic plastocyanins; cyan - azurins; violet - pseudoazurins; orange - amicyanins; magenta - rusticyanin; black - CBP; brown - stellacyanin.

Correlation:

Pairwise sequence identity: similarity index for 3D interaction properties

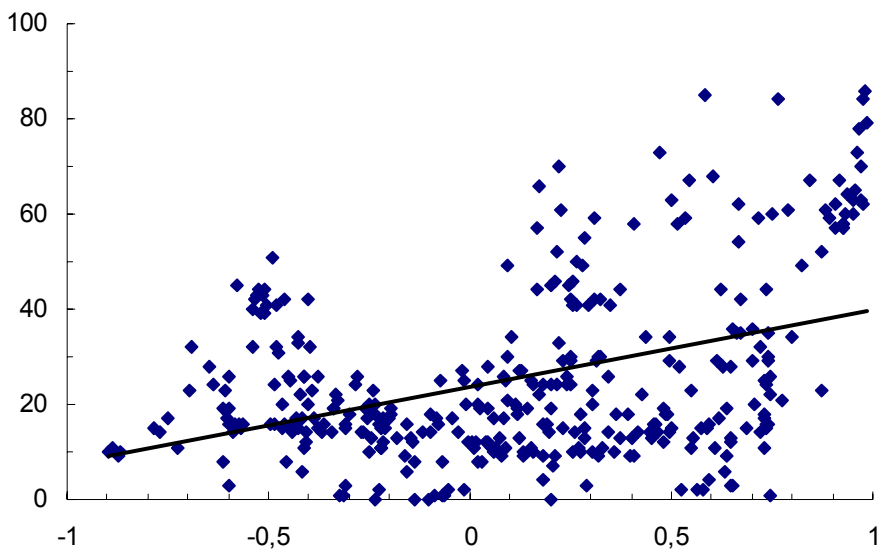
id%



hydrophobicity SI

$id\% = 95.05si_{hyd} + 10.84$

$r^2 = 0.62$; $r = 0.78$

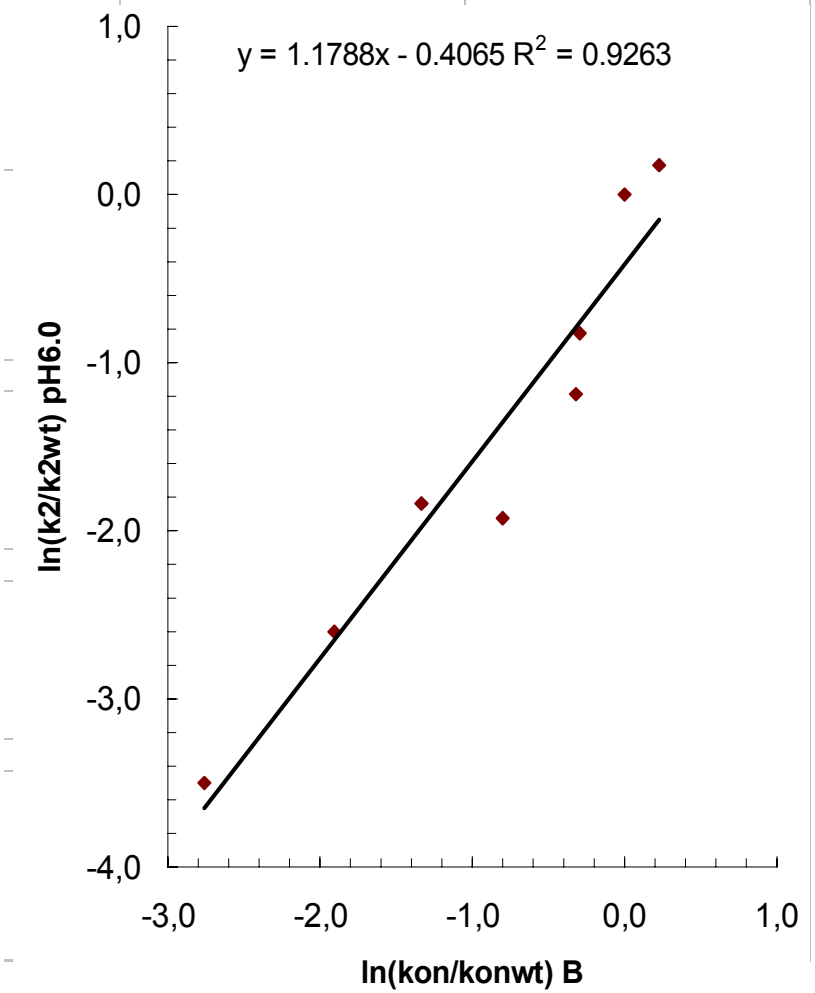


electrostatic potential SI

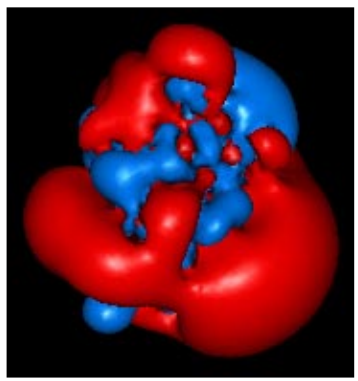
$id\% = 16.23si_{mep} + 23.54$

$r^2 = 0.17$; $r = 0.41$

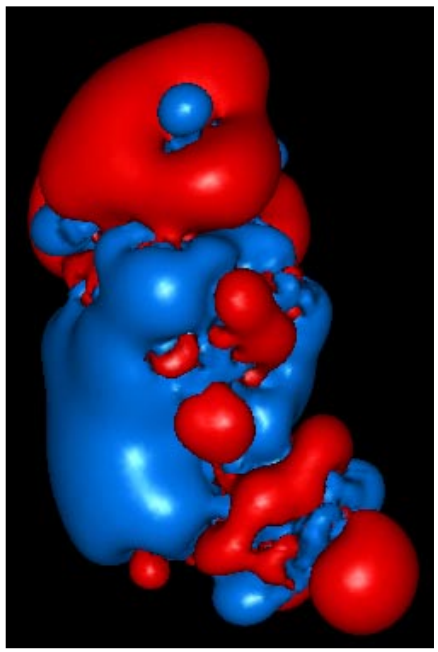
Association (ET) rates for plastocyanin mutants to cytochrome f



Brownian Dynamics



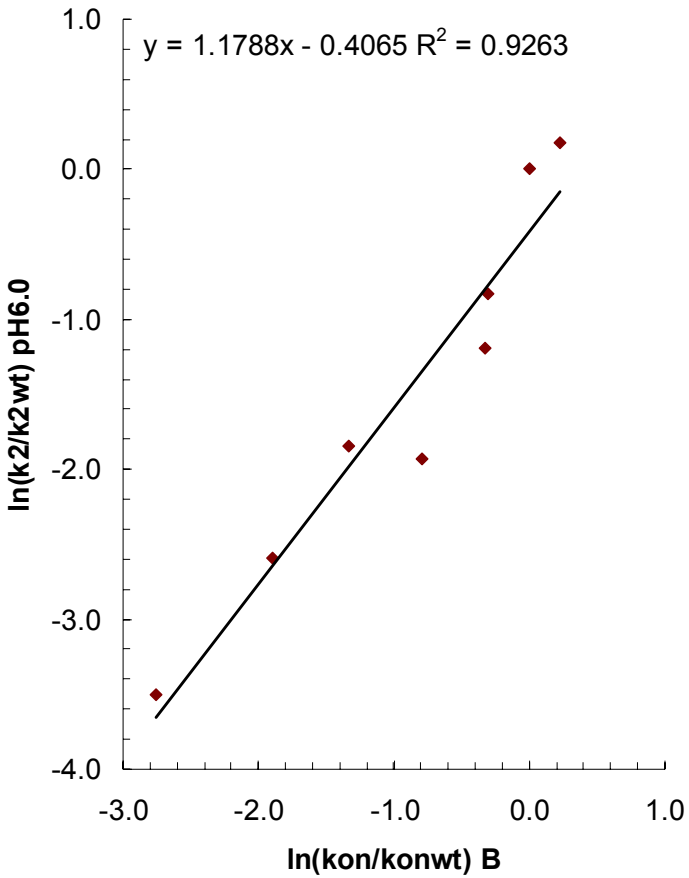
plastocyanin:



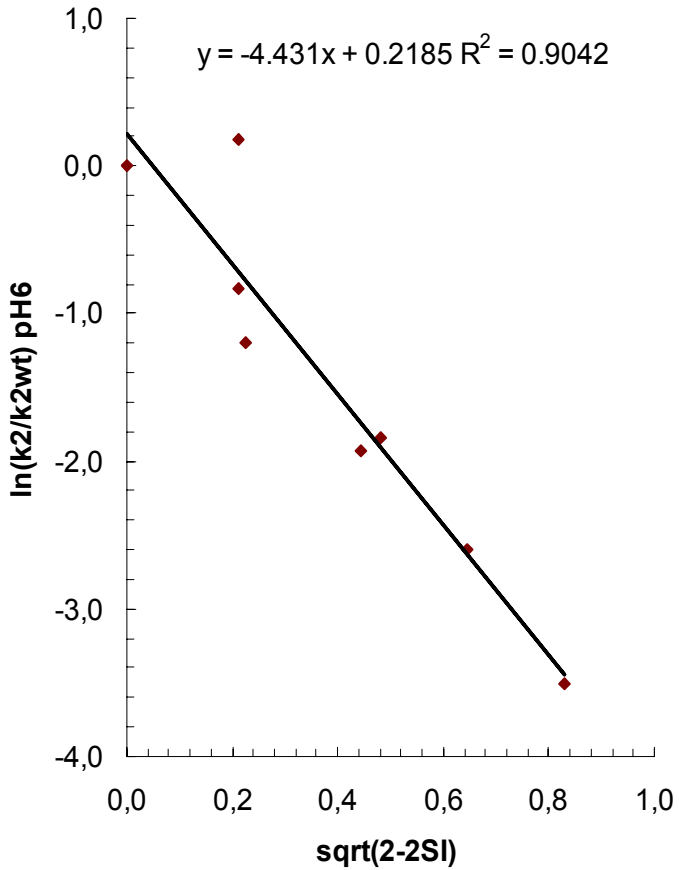
cytochrome f

De Rienzo, Gabdoulline, Menziani & Wade Biophys J. 2001

Association (ET) rates for plastocyanin mutants to cytochrome f



Brownian Dynamics



PIPSA: electrostatic potential