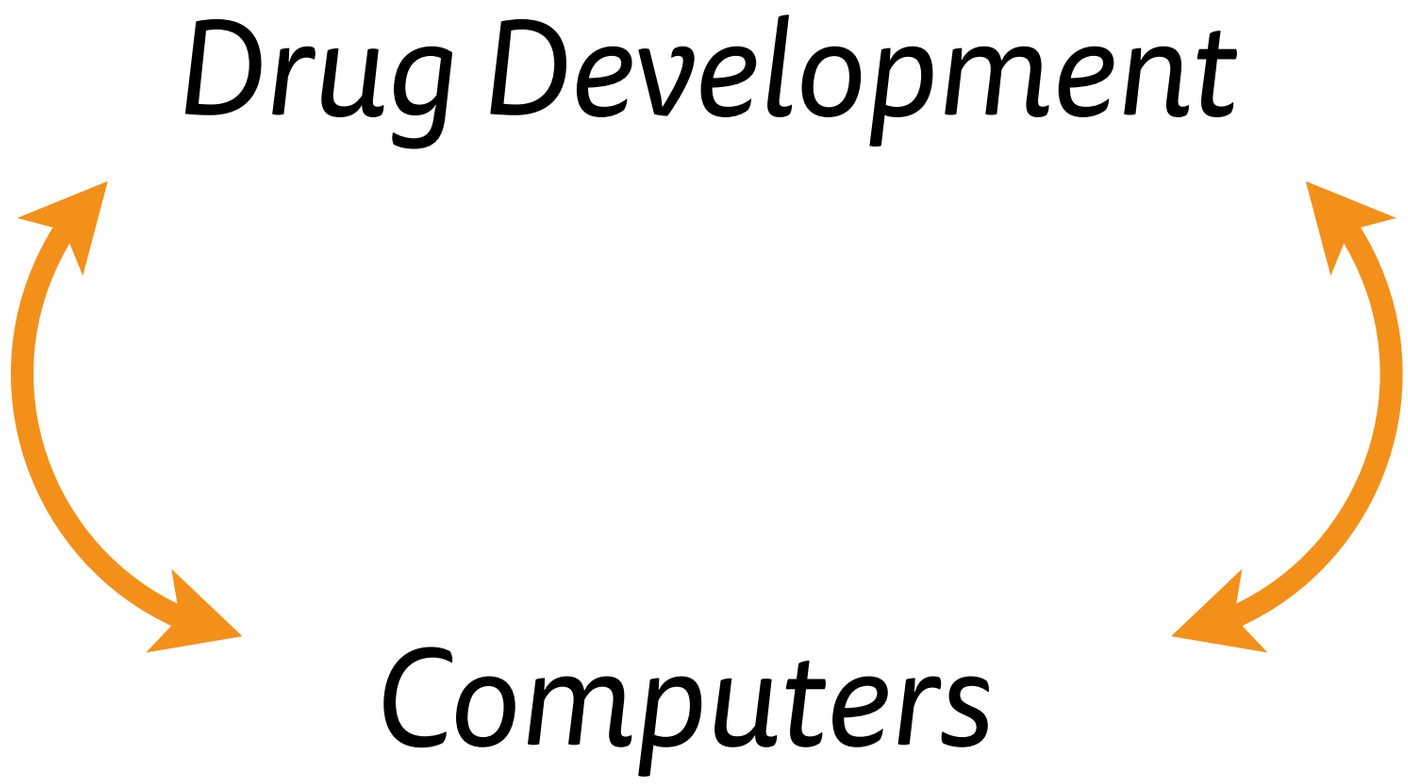
Do enzyme-inhibiting drugs show increased reliance on certain chemical properties for binding to their respective enzymes?

Chemical properties that affect binding of enzyme-inhibiting drugs to enzymes

Research proposal by Dan Nacu

Why is this significant?

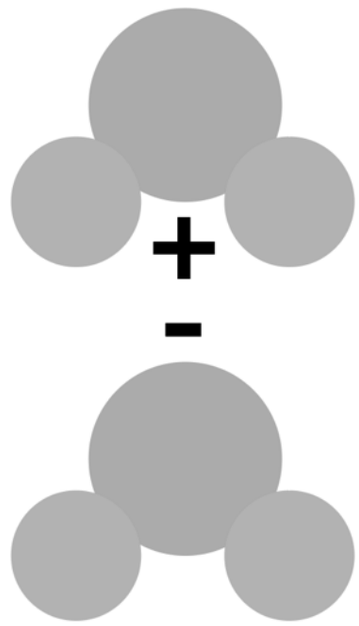


How can it be done?

Simulation Models

Shape Complementarity

Chemical Properties



How can it be done?

Chemical Properties

Solvent Accessible Surface Area
Hydrophobicity
Electrostatics
Van Der Waal's Forces
Residue Pair potential
Desolvation Energies
Atomic Contact Energies
Complementary Determining Regions
etc...

A lot of options...

Its been done before...in a different way.

Li et al, 2007

ELSEVIER

Biophysical Chemistry

Biophysical Chemistry 129 (2007) 1 – 10

http://www.elsevier.com/locate/biophyschem

Complex-type-dependent scoring functions in protein–protein docking

Chun Hua Li [a,1], Xiao Hui Ma [a,b,1], Long Zhu Shen [a], Shan Chang [a], Wei Zu Chen [a], Cun Xin Wang [a,*]

[a] College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100022, People's Republic of China
[b] Department of Pharmaceutical and Biomedical Sciences, University of George, Athens, GA 30602, USA

Received 14 February 2007; received in revised form 24 April 2007; accepted 25 April 2007
Available online 8 May 2007

Abstract

A major challenge in the field of protein–protein docking is to discriminate between the many wrong and few near-native conformations, i.e. scoring. Here, we introduce combinatorial complex-type-dependent scoring functions for different types of protein–protein complexes, protease/inhibitor, antibody/antigen, enzyme/inhibitor and others. The scoring functions incorporate both physical and knowledge-based potentials, i.e. atomic contact energy (ACE), the residue pair potential (RP), electrostatic and van der Waals' interactions. For different type complexes, the weights of the scoring functions were optimized by the multiple linear regression method, in which only top 300 structures with ligand root mean square deviation (L_RMSD) less than 20 Å from the bound (co-crystallized) docking of 57 complexes were used to construct a training set. We employed the bound docking studies to examine the quality of the scoring function, and also extend to the unbound (separately crystallized) docking studies and extra 8 protein–protein complexes. In bound docking of the 57 cases, the first hits of protease/inhibitor cases are all ranked in the top 5. For the cases of antibody/antigen, enzyme/inhibitor and others, there are 17/19, 5/6 and 13/15 cases with the first hits ranked in the top 10, respectively. In unbound docking studies, the first hits of 9/17 protease/inhibitor, 6/19 antibody/antigen, 1/6 enzyme/inhibitor and 6/15 others' complexes are ranked in the top 10. Additionally, for the extra 8 cases, the first hits of the two protease/inhibitor cases are ranked in the top for the bound and unbound test. For the two enzyme/inhibitor cases, the first hits are ranked 1st for bound test, and the 119th and 17th for the unbound test. For the others, the ranks of the first hits are the 1st for the bound test and the 12th for the 1WQ1 unbound test. To some extent, the results validated our divide-and-conquer strategy in the docking study, which might hopefully shed light on the prediction of protein–protein interactions.

Keywords: Binding affinity; Scoring function; Protein–protein docking

1. Introduction

Protein–protein interaction is the basis of many biological regulations. Knowledge of 3-dimensional (3D) protein–protein structures is important for an adequate description of protein–protein interactions. However, large macromolecular assemblies are a major challenge for structural biology. The amount of experimental structures of protein–protein complexes is relatively quite small and the cost is very expensive. Thus, a combination of protein modeling and experimental structure determination increases knowledge of structure-based analysis of the protein–protein interaction network [1–4]. As a part of molecular modeling, docking algorithms are designed to model protein–protein complexes based on the component structures.

Docking algorithms have progressed in recent years, which can dock unbound (separately crystallized) proteins to obtain the structure of the complex with small structural changes accompanying complexation [5–19]. The accuracy and reliability of docking algorithms still need to be assessed if they are to become widely used. This depends on docking algorithms with an efficient procedure to generate potential structures and a good scoring function to distinguish the near-native structures from a large number of non-native ones. The known scoring functions include surface complementarity (SC) [5,6], surface complementarity together with an electrostatic filter [20,21], knowledge-based statistical potential such as atomic contact energy (ACE) [22], the residue pair potential (RP) [23] and DFIRE [24]. Some combinatorial functions are used in docking

* Corresponding author.
E-mail address: cxwang@bjut.edu.cn (C.X. Wang).
[1] Both are first authors.

0301-4622/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2007.04.014

Their Equation

$$\text{Score} = w_1 E_{\text{RP}} + w_2 E_{\text{ACE}} + w_3 E_{\text{vdw}}^{\text{attr}} + w_4 E_{\text{vdw}}^{\text{rep}} + w_5 E_{\text{ele}}^{\text{sa}} + w_6 E_{\text{ele}}^{\text{sr}} + w_7 E_{\text{ele}}^{\text{la}} + w_8 E_{\text{ele}}^{\text{lr}} ,$$

Their Results

Name	Success Ratio
Protease/Inhibitor	16/17
Enzyme/Inhibitor	6/6
Antibody/Antigen	18/19
Other	11/15

How will this be different?

Introducing HINT

***H**ydropathic **INT**eractions*

$$b_{ij} = a_i \, a_j \, S_i \, S_j \, T_{ij} \, R_{ij} + r_{ij}$$

The HINT Equation

What can be done?

By weighing each variable in HINT, the most important chemical property for enzyme/ inhibitor complexes can be found.

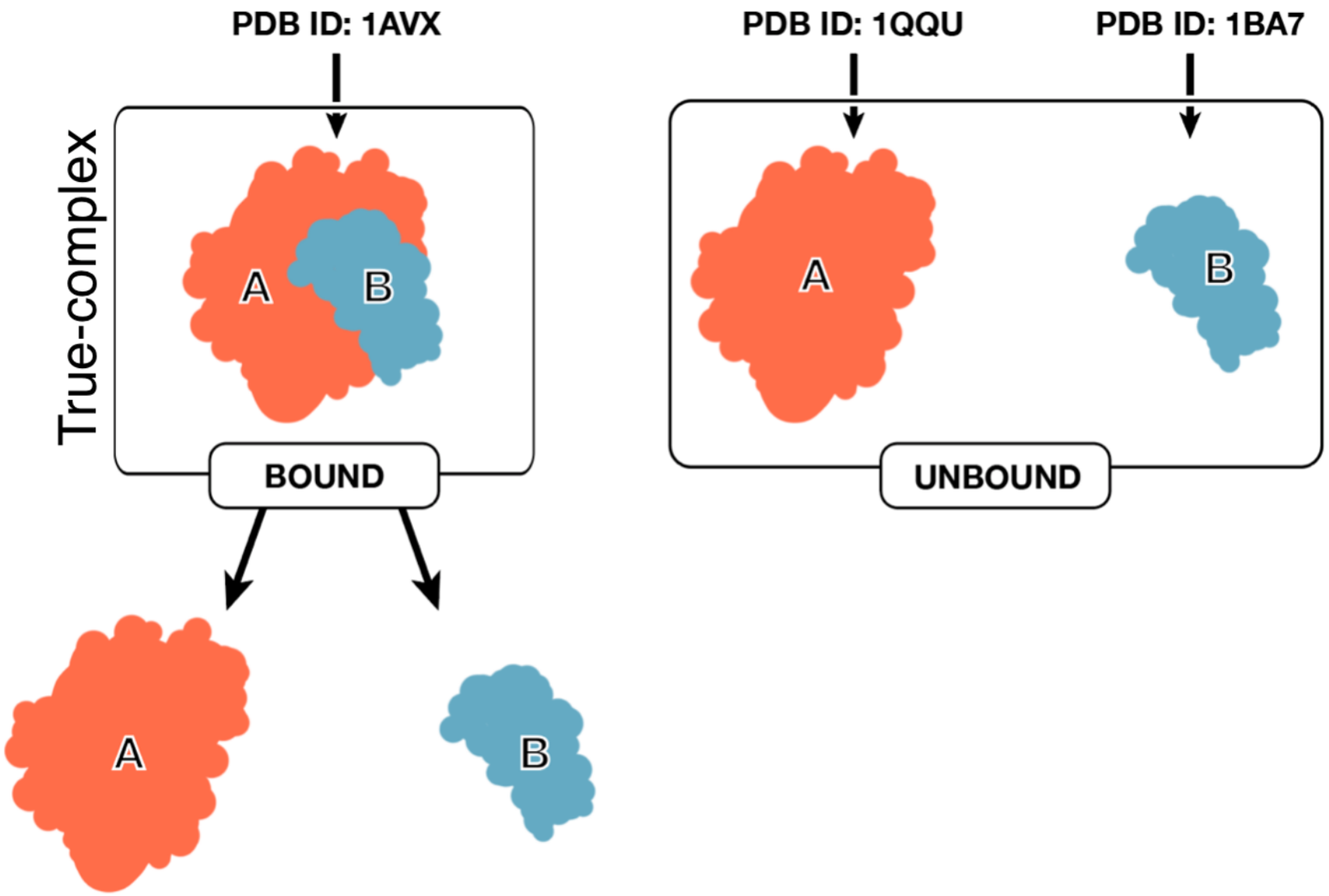
$$b_{ij} = a_i \, a_j \, S_i \, S_j \, T_{ij} \, R_{ij} + r_{ij}$$

Remember The Question

Do enzyme-inhibiting drugs show increased reliance on certain chemical properties for binding to their respective enzymes?

What's the experiment?

Start with 46 enzyme inhibitor complexes from ***the Benchmark 5.***



What's the experiment?

Bound

Unbound

FTDock

Huge list of possible complexes

What's the experiment?

Huge list of possible complexes

L_RMSD Testing

Ligand_Root-Mean-Square-Deviation

Top 20 Structures

What's the experiment?

Top 20 Structures

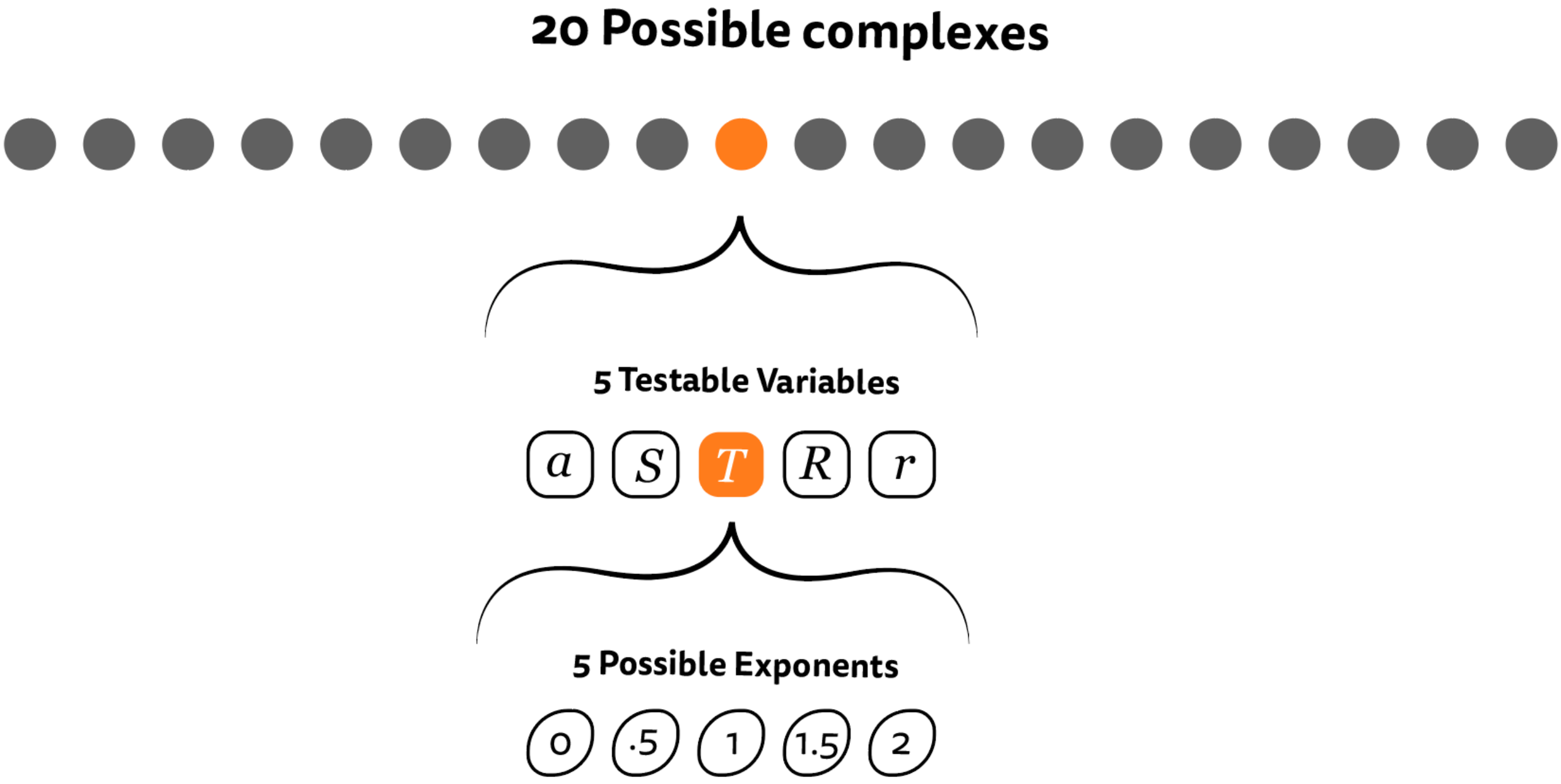
*For **46** complexes =*

***920** simulated structures.*

*For both bound and unbound, **1,840 total***

Lets look at just one

What's the experiment?



What's the experiment?

In the end...

23,000 HINT Scores for Bound

23,000 HINT Scores for Unbound

46,000 scores

46•20•5•5 = 23,000

What's the experiment?

Highest 50 HINT Scores for each complex

L_RMSD Testing

Ligand_Root-Mean-Square-Deviation

Find best match for each complex

Results!

Possible Results

Complex	Final L_RMSD Score	Weighing Used	Significant Chemical Property
#1 Bound	4 Å	$a_i a_j (S_i S_j)^{1.5} T_{ij} R_{ij} + r_{ij}$	Solvent Accessible Surface Area
#1 Unbound	6 Å	$a_i a_j S_i S_j (T_{ij})^2 R_{ij} + r_{ij}$	**Electrostatics**
#2 Bound	2 Å	$a_i a_j S_i S_j T_{ij} (R_{ij})^{0.5} + r_{ij}$	Atomic Distance
#2 Unbound	4 Å	$a_i a_j S_i S_j (T_{ij})^{1.5} R_{ij} + r_{ij}$	**Electrostatics**
#3 Bound	3 Å	$a_i a_j (S_i S_j)^{1.5} T_{ij} R_{ij} + r_{ij}$	Solvent Accessible Surface Area
#3 Unbound	5 Å	$a_i a_j S_i S_j (T_{ij})^{1.5} R_{ij} + r_{ij}$	**Electrostatics**
…	…	…	…
#46 Bound	2 Å	$a_i a_j S_i S_j (T_{ij})^2 R_{ij} + r_{ij}$	**Electrostatics**
#46 Unbound	6 Å	$(a_i a_j)^{0.5} S_i S_j T_{ij} R_{ij} + r_{ij}$	Hydrophobic Atom Constant

Possible Results

Complex	Final L_RMSD Score	Weighing Used	Significant Chemical Property
#1 Bound	4 Å	$a_i a_j (S_i S_j)^1 T_{ij} R_{ij} + r_{ij}$	Solvent Accessible Surface Area
#1 Unbound	6 Å	$a_i a_j S_i S_j (T_{ij})^2 R_{ij} + r_{ij}$	Electrostatics
#2 Bound	2 Å	$a_i a_j S_i S_j T_{ij} (R_{ij})^{0.5} + r_{ij}$	Atomic Distance
#2 Unbound	4 Å	$a_i a_j S_i S_j (T_{ij})^{1.5} R_{ij} + r_{ij}$	Electrostatics
#3 Bound	3 Å	$a_i a_j (S_i S_j)^{1.5} T_{ij} R_{ij} + r_{ij}$	Solvent Accessible Surface Area
#3 Unbound	5 Å	$a_i a_j S_i S_j (T_{ij})^{0.5} R_{ij} + r_{ij}$	Electrostatics
…	…	…	…
#46 Bound	2 Å	$a_i a_j (S_i S_j)^0 T_{ij} R_{ij} + r_{ij}$	Solvent Accessible Surface Area
#46 Unbound	6 Å	$(a_i a_j)^{0.5} S_i S_j T_{ij} R_{ij} + r_{ij}$	Hydrophobic Atom Constant

In the future…

Different models (besides HINT)

Different complexes (besides enzyme/inhibitor)

Questions?