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?

Drug Development

Computers
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...
It's been done before...in a different way.

Li et al, 2007

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

Chun Hua Li, Xiao Hui Ma, Long Zhu Shen, Shan Chang, Wei Zu Chen, Cun Xin Wang

Abstract

A major challenge in the field of protein–protein docking is to discriminate between the many wrong and few superative conformations, i.e., scoring. Here, we introduce combinatorial complex-type-dependent scoring functions for different types of protein–protein complexes, protein/inhibitor, antibody/antigen, enzyme/inhibitor and others. The scoring functions incorporate both physical and knowledge-based potentials, i.e., atomic contact energy (ACE), the residual pair potential (RPP), 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 deviations (RMSD) less than 2.0 Å from the bound (co-crystallized) docking of 37 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 even the protein–protein complexes. In bound docking of the 37 years, the first hits of protein/inhibitor cases are all ranked in the top 5. For the cases of antibody/antigen, enzyme/inhibitor and others, there are 15/19, 56 and 55 cases with the first hits ranked in the top 10, respectively. In unbound docking studies, the first hits of 417 protein/inhibitor, 67/9 antibody/antigen, 6 enzyme/inhibitor and 6 3 others' complexes are ranked in the top 10. Additionally, for the 8 cm–8 cases, the first hits of the two protein/inhibitor cases are ranked in the top 10 for the bound and unbound test. For the two antibody/antigen cases, the first hits are ranked 1st for bound test, and the 19th and 57th 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 VDO 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.

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 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 in-house (separately crystallized) proteins to obtain the structure of the complex with small structural changes accompanying complexation [5–9]. 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 residual pair potential (RPP) [23] and DFIRE [24]. Some combinatorial functions are used in docking

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

Their Equation

<table>
<thead>
<tr>
<th>Name</th>
<th>Success Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protease/Inhibitor</td>
<td>16/17</td>
</tr>
<tr>
<td>Enzyme/Inhibitor</td>
<td>6/6</td>
</tr>
<tr>
<td>Antibody/Antigen</td>
<td>18/19</td>
</tr>
<tr>
<td>Other</td>
<td>11/15</td>
</tr>
</tbody>
</table>

Their Results
How will this be different?

Introducing HINT

Hydropathic INTeractions

\[ b_{ij} = a_i \ a_j \ S_i \ S_j \ T_{ij} \ R_{ij} + r_{ij} \]

The HINT Equation
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

FTDock

Unbound

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

Let's look at just one
What's the experiment?

20 Possible complexes

5 Testable Variables

\[ a \quad S \quad T \quad R \quad r \]

5 Possible Exponents

\[ 0 \quad .5 \quad 1 \quad 1.5 \quad 2 \]
What's the experiment?

In the end...

23,000 HINT Scores for Bound
23,000 HINT Scores for Unbound

46,000 scores

46 \times 20 \times 5 \times 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

<table>
<thead>
<tr>
<th>Complex</th>
<th>Final L_RMSD Score</th>
<th>Weighing Used</th>
<th>Significant Chemical Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 Bound</td>
<td>4 Å</td>
<td>(a_ia_j \left( S_iS_j \right)^{1.5} T_{ij} R_{ij} + r_{ij})</td>
<td>Solvent Accessible Surface Area</td>
</tr>
<tr>
<td>#1 Unbound</td>
<td>6 Å</td>
<td>(a_ia_j S_iS_j (T_{ij})^2 R_{ij} + r_{ij})</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>#2 Bound</td>
<td>2 Å</td>
<td>(a_ia_j S_iS_j T_{ij} (R_{ij})^{0.5} + r_{ij})</td>
<td>Atomic Distance</td>
</tr>
<tr>
<td>#2 Unbound</td>
<td>4 Å</td>
<td>(a_ia_j S_iS_j (T_{ij})^{1.5} R_{ij} + r_{ij})</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>#3 Bound</td>
<td>3 Å</td>
<td>(a_ia_j (S_iS_j)^{1.5} T_{ij} R_{ij} + r_{ij})</td>
<td>Solvent Accessible Surface Area</td>
</tr>
<tr>
<td>#3 Unbound</td>
<td>5 Å</td>
<td>(a_ia_j S_iS_j (T_{ij})^{1.5} R_{ij} + r_{ij})</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>#46 Bound</td>
<td>2 Å</td>
<td>(a_ia_j S_iS_j (T_{ij})^2 R_{ij} + r_{ij})</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>#46 Unbound</td>
<td>6 Å</td>
<td>((a_ia_j)^{0.5} S_iS_j T_{ij} R_{ij} + r_{ij})</td>
<td>Hydrophobic Atom Constant</td>
</tr>
</tbody>
</table>
# Possible Results

<table>
<thead>
<tr>
<th>Complex</th>
<th>Final L_RMSD Score</th>
<th>Weighing Used</th>
<th>Significant Chemical Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 Bound</td>
<td>4 Å</td>
<td>$a_i a_j (S_i S_j) T_{ij} R_{ij} + r_{ij}$</td>
<td>Solvent Accessible Surface Area</td>
</tr>
<tr>
<td>#1 Unbound</td>
<td>6 Å</td>
<td>$a_i a_j S_i S_j (T_{ij})^2 R_{ij} + r_{ij}$</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>#2 Bound</td>
<td>2 Å</td>
<td>$a_i a_j S_i S_j T_{ij} (R_{ij})^{0.5} + r_{ij}$</td>
<td>Atomic Distance</td>
</tr>
<tr>
<td>#2 Unbound</td>
<td>4 Å</td>
<td>$a_i a_j S_i S_j (T_{ij})^{1.5} R_{ij} + r_{ij}$</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>#3 Bound</td>
<td>3 Å</td>
<td>$a_i a_j (S_i S_j)^{1.5} T_{ij} R_{ij} + r_{ij}$</td>
<td>Solvent Accessible Surface Area</td>
</tr>
<tr>
<td>#3 Unbound</td>
<td>5 Å</td>
<td>$a_i a_j S_i S_j (T_{ij})^{0.5} R_{ij} + r_{ij}$</td>
<td>Electrostatics</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>#46 Bound</td>
<td>2 Å</td>
<td>$a_i a_j (S_i S_j)^9 T_{ij} R_{ij} + r_{ij}$</td>
<td>Solvent Accessible Surface Area</td>
</tr>
<tr>
<td>#46 Unbound</td>
<td>6 Å</td>
<td>$(a_i a_j)^{0.5} S_i S_j T_{ij} R_{ij} + r_{ij}$</td>
<td>Hydrophobic Atom Constant</td>
</tr>
</tbody>
</table>
In the future...

Different models (besides HINT)
Different complexes (besides enzyme/inhibitor)
Questions?