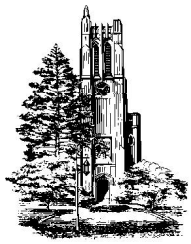
Modeling of Hydrophobic Interactions in Computational Ligand Screening: A Template Matching Approach

Paul C. Sanschagrin[1], Maria I. Zavodszky[1], and Leslie A. Kuhn[1,2]

[1]Department of Biochemistry and Molecular Biology and [2]Center for Biological Modeling
Michigan State University, East Lansing, MI USA 48824
http://www.bch.msu.edu/labs/kuhn/; sanschag@sol.bch.msu.edu



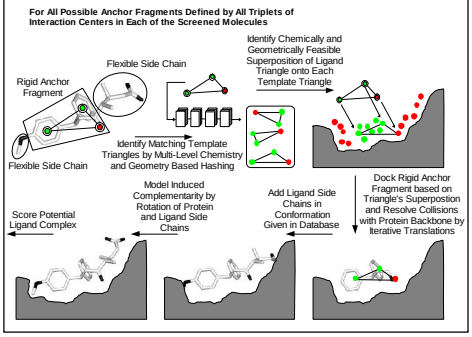
Abstract

Computational ligand screening is an approach to identify potential protein ligands efficiently using computational rather than experimental techniques. One approach is to test for the ability to identify and dock true ligands from a large number of molecular structures. Using this approach as implemented in our software SLIDE(1), screenings of 100,000-plus molecules can be achieved in a few hours to a day, with each molecule taking less than one second to screen and dock. SLIDE does this screening while including both ligand and protein side-chain flexibility and water-mediated interactions, with a template-matching approach based points of complementary chemistry between the protein and each ligand candidate. The description of the protein template, representing the binding site, involves points where favorable hydrophobic or hydrogen bonding interactions can be made. Here, we present an improved method of modeling the hydrophobic interaction component of the SLIDE screening method. Protein template points are identified by representing the major patches of hydrophobic surface on the protein with a uniform density of points, while not increasing the number of points in the protein binding site. Hydrophobic interaction points on the ligand are assigned using rules which assign points at a similar density and spacing to protein template points, improving the quality of ligands dockings, especially for predominantly hydrophobic ligands. We present use of this new method on several test cases, showing docking of known ligands closer to the orientation found in the X-ray crystallographic structure and faster screening times on a test database. Combined with the new hydrogen-bond template design developed by Maria Zavodszky (see related poster), the new knowledge-based hydrophobic and hydrogen-bond template method allows unbiased screening of ligands while considering the shape and chemistry of the protein target.

Introduction

- When screening a database of 100,000+ molecules for potential ligands, only 1 second can be used to achieve and score a single docking in order to complete the screening in one to a few days.
- Induced complementarity has previously been implemented in molecular docking algorithms (2-7), but has been too computationally time consuming to screen a large database; SLIDE overcomes this by using a geometric hash scheme to quickly eliminate infeasible dockings and then a mean-field optimization (adapted from 9,10) to model ligand and protein side-chain flexibility upon docking.
- SLIDE describes the protein binding site as a collection of points of favorable hydrogen bonding or hydrophobic interaction and, similarly, describes each molecule in the database as a set of potential interaction points. SLIDE then selects potential ligands from the database by matching these points and checking for steric complementarity.
- Here, we present an improved method for modeling the hydrophobic interaction

SLIDE uses a multistep algorithm for screening large molecule databases with induced ligand and protein flexibility.



SLIDE uses a geometric hashing scheme to identify infeasible dockings for each triplet of interaction points in each compound in the database relative to triplets of interaction points in the protein template.

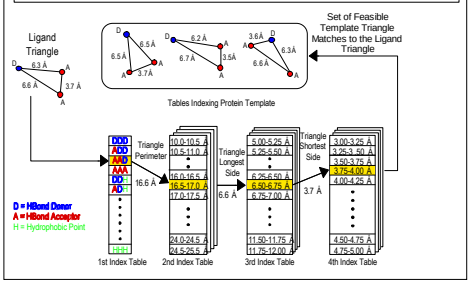


Figure 1. Templates for the Estrogen Receptor using the 3ERD PDB structure with the original and new methods for hydrophobic template point generation. The DES ligand is shown in tubes colored by atom type (carbon, green; oxygen, red). The template points are shown as spheres, colored as follows:

Hydrophobic Point - Green
Hydrogen Bond Donor - Blue
Hydrogen Bond Acceptor - Red
Hydrogen Bond Donor or Acceptor - White

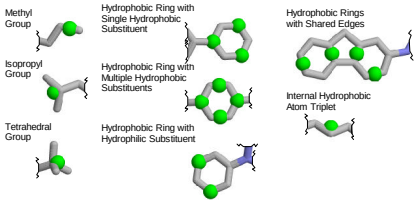
SLIDE represents the major hydrophobic surface patches of the protein in the template.

- Examine protein environment of each point
- Determine a hydrophobic enhancement score for the point

Enhancement Score =
Number of Hydrophobic Protein Atoms within 5.2 Å
- Number of Hydrophilic Protein Atoms within 5.2 Å

- If score > cutoff (3), assign as H

SLIDE assigns ligand interaction points using a rule based approach that provides even point density in hydrophobic regions of ligands.



The new hydrophobic interaction method was tested against a set of three proteins with varying ligands.

Thrombin

PDB Code	Ligand Name
1VR1	**Uniganded**
1A2C	Aeruginosin298-A
1A3B	Borolog1
1A3E	Borolog2
1A46	b-strand Mimetic Inhibitor
1A4W	Thiazole Containing Inhibitor
1A5G	Synthetic Peptide Mimetic
1A61	b-strand Mimetic Inhibitor
1ABJ	PPACK
1AD8	MDL 103752
1AE8	L-Ethoxycarbonyl-D-Phe-Pro-2(4-aminobutyl)hydrazine
1AFE	n-Carbobenzoxy-Pro-a-azalus p-nitrophenyl ester
1AHT	p-Amidino-phenyl-pyruvate
1AI8	Morpholino-diphenylAla-methoxypropylboronic acid
1AIX	Phenylmethylenecarboxy-(methyleneamino-formyl-diphenylmethyl)methy-Pro-boroVal
1AWF	GR133487
1AWH	GR133686
1AY6	Macrocyclic Tripeptide
1B5G	Peptide Mimetic Inhibitor
1BA8	Rigid Tripeptidyl Adehyde
1BB0	Rigid Tripeptidyl Adehyde
1BCU	Proflavin
1BHX	SDZ 229-357
1BMM	BMS-186282
1BMN	BMS-189090
1CA8	Rigid Tripeptidyl Adehyde
1DWB	Bezamidine
1DWC	Argatroban (MD-805)
1DWD	NAPAP
1DWE	PPACK
1FPC	DAPA
1HAI	PPACK
1HAP	PPACK
1HDT	BMS-183507
1HLT	PPACK
1LHC	Ac-(D)Phe-Pro-boroArg-OH
1LHD	Ac-(D)Phe-Pro-boroLys-OH
1LHE	Ac-(D)Phe-Pro-boro-N-butyl-amidino-Gly-OH
1LHG	Ac-(D)Phe-Pro-boroornithine-OH
1NRR	PPACK
1NRS	Receptor Based Peptide
1PPB	PPACK
1TBZ	(D)-Phe-Pro-Arg-benzothiazole
1TMB	Cyclotheonamide A
1TMT	(D)-Phe-Pro-Arg
1TMU	PPACK
1TOM	Methyl-Phe-Pro-amino-cyclohexylGly
1UMA	"(N,N)-Dimethylcarbamoyl-a-azaLys"
1UVS	BM51.1011
1UVT	BM14.1248
1UVU	BM12.1700
2HPQ	PPCK
3HAT	Fibrinopeptide A Mimic
7KME	SEL2711
8KME	SEL2770

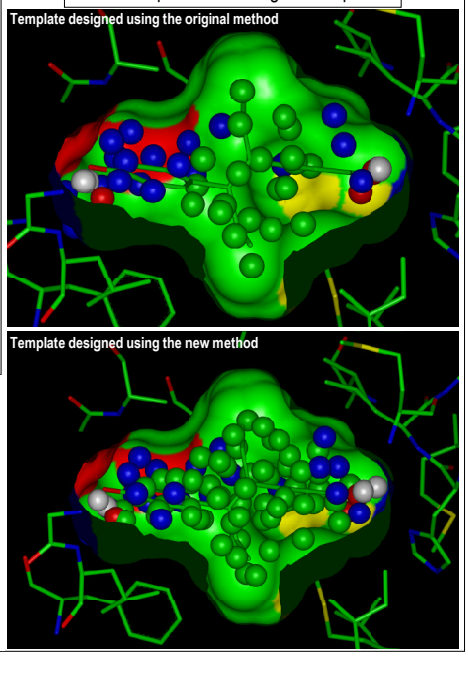
PolyADP Ribose Polymerase

PDB Code	Ligand Name
2paw	**Uniganded**
1pax	3,4-Dihydro-5-methyl-isoqinolinone
2pax	4-Amino-1,8-Naphthalimide
3pax	3-Methylbenzamide
4pax	8-Hydroxy-2-methyl-hydro-quinazolin-4-one

Estrogen Receptor

PDB Code	Ligand Name
1a52	Estradiol
1ere	Estradiol
1err	Raloxifene
1qkt	Estradiol
1qku	Estradiol
3erd	Dethylstilbestrol (DES)
3ert	4-Hydroxytamoxifen

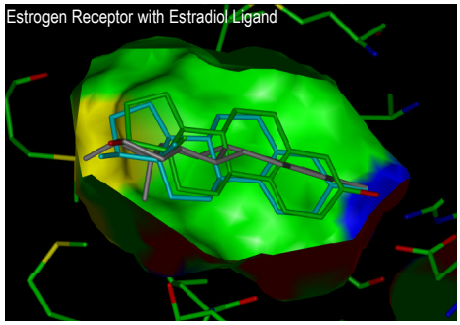
The new hydrophobic template better represents the hydrophobicity of the ligand and the protein -- Estrogen Receptor

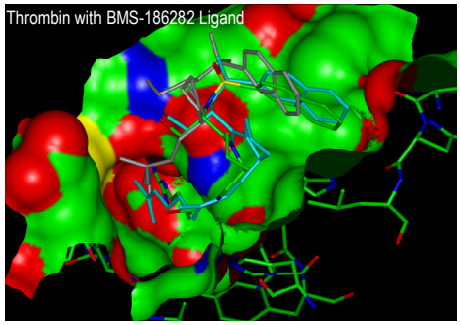


Using the new method for finding hydrophobic interaction points, *SLIDE* achieves dockings which are closer to the

Protein	Number of Ligands In Database	Number of Ligands Docked Original Method	Number of Ligands Docked New Method	RMSD Improvement*	Drugscore Improvement*
Estrogen	7	5	5	-1.33 5/0/0	-65587 5/0/0
Thrombin	56	34	38	+0.16 13/13/5	-23742 14/7/10
Poly ADP Ribose Polymerase	4	3	4	-0.34 2/0/1	+912 1/0/2

*Improvements are scaled such that <0 is better and >0 is worse; below is # Better/# Same/# Worse





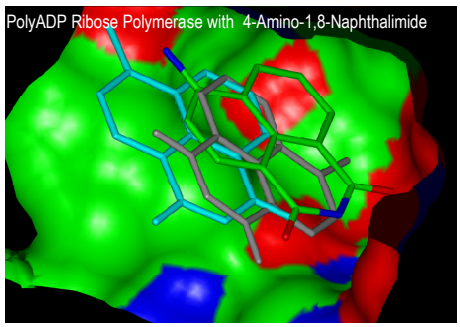


Figure 2. Best dockings for example ligands for the three test cases. In all figures, the real ligand (docked into the free structure by superposition with the ligand bound structure for thrombin and polyADP ribose polymerase) is shown colored by atom(crabon, green; oxygen, red; nitrogen, blue), the best docked ligand based on RMSD using the original interaction point method is in grey and using the new method in blue. The new method does better, especially on ligands with many hydrophobic rings, such as the estradiol.

Conclusions

- The new method for hydrophobic template design yields better representation of the hydrophobic character of the binding site than the original method.
- The new hydrophobic ligand interaction point assignment yields more accurate dockings than the original method.

References

1. V. Schnecke, L. A. Kuhn. *Perspectives in Drug Discovery and Design* **20**, 171, (2000)
2. R. M. A. Knegtel, et. al., *J. Comput. Aided. Mol. Des.*, **13**, 167, (1999)
3. B. Sandak, et. al., In D. Hirschberg and G. Myers, eds. *Combinatorial Pattern Matching* **vol 1075** of *LNCS*, 271, (1996)
4. B. Sandak, et. al., *Proteins: Struct. Funct. Genet*, **32**, 159 (1998)
5. M. Totrov, R. Abagyan, *Proteins: Struct. Funct. Genet*, **Supplement 1, 215** (1997)
6. Z. R. Wasserman, C. N. Hodge, *Proteins: Struct. Funct. Genet*, **24**, 227, (1996)
7. M. Rarey, et. al., *J. Mol. Biol.*, **261**, 470, (1996)
8. V. Schnecke, et. al, *Proteins: Struct. Funct. Genet*, **33**, 74, (1998)
9. R. M. Jackson, et. al., *J. Mol. Biol.*, **276**, 265, (1998)
10. P. Koehl, M. Delarue, *J. Mol. Biol.*, **239**, 249, (1994)