BEILSTEIN INSTITUT

Discovering the Subtleties of Sugars
June 10th – 14th, 2013, Potsdam, Germany

Prediction of Binding Poses and Binding Affinities for Glycans and their Binding Proteins using a Robust Scoring Function for General Protein-Ligand Interactions

Nan-Lan Huang[1] **and Jung-Hsin Lin**[1,2,3,*]

[1]Research Center for Applied Sciences and [2]Institute of Biomedical Sciences, Academia Sinica, 128 Academia Rd., Sec. 2, Nankang, Taipei 115, Taiwan;

[3]School of Pharmacy, National Taiwan University, 1 Jen-Ai Rd., Sec. 2, Taipei 10051, Taiwan

E-Mail: *jlin@ntu.edu.tw, jhlin@gate.sinica.edu.tw

Received: 30th September 2013 / Published: 22nd December 2014

Abstract

The binding of glycans to proteins represents the major way in which the information contained in glycan structures is recognised, deciphered and put into biological action. The physiological and pathological significance of glycan–protein interactions are drawing increasing attention in the field of structure-based drug design. We have implemented a quantum chemical charge model, the Austin-model 1-bond charge correction (AM1-BCC) method, into a robust scoring function for general protein ligand interactions, called, AutoDock^{RAP}. Here we report its capability to predict the binding poses and binding affinities of glycans to glycan-binding proteins. Our benchmark indicates that this generally applicable scoring function can be adopted in virtual screening of drug candidates and in prediction of ligand binding modes, given the structures of the well-defined recognition domains of glycan-binding proteins.

This article is part of the Proceedings of the Beilstein Glyco-Bioinformatics Symposium 2013.
www.proceedings.beilstein-symposia.org

INTRODUCTION

Glycans, the freestanding or covalently linked monosaccharide or oligosaccharide entities in the cells, mediate a wide variety of biological functions of both prokaryotic and eukaryotic cells. A majority of these biological processes involve the recognition of glycans by glycan-binding proteins (GBP). Particularly, glycans located on cell surfaces or secreted biomolecules play a crucial role in cell–cell interaction, including the interaction between host cell and pathogens [1]. For example, understanding of the recognition of host glycoprotein receptors by viral neuraminidase led to the development of high-affinity inhibitors in use to reduce the prevalence of influenza. It is of high priority to understand the molecular basis of the interaction between GBP and glycans involved in the various physiological or pathological events.

Evaluation of the binding affinities of drug-like molecules with the target proteins is crucial for discriminating drug candidates from weak-binding or even nonbinding small molecules. Rigorous statistical mechanical approaches for evaluation of binding free energies are theoretically most satisfactory [2, 3], but such approaches are computationally too demanding for virtual screening. Due to practical consideration, most, if not all, computational docking methods rely greatly on empirical or semi-empirical scoring functions to evaluate protein–ligand interactions. Semi-empirical models based on molecular mechanics have the advantages of easier rational interpretation of binding modes, and they are more sensitive to protein conformational changes. Frequently used energetic terms include van der Waals energy, electrostatic energy, hydrogen-bond energy, desolvation energy, hydrophobic interaction, torsional entropy, and so on [4, 5]. Among these terms, the atomic partial charges of biomolecules are considered of central importance, because they are essential for evaluation of the long-ranged electrostatic interaction, which is known to be a key factor for biomolecular association. Due to the extremely low computational cost, current molecular docking programs often use regression models with distance-dependent molecular descriptors or energy terms to predict the possible binding poses and to evaluate the binding affinity of a small molecule. Such descriptors are also used for large-scale virtual chemical library screening to rapidly narrowing down the chemical space and for subsequent identification of potential drugs.

The affinity of most single glycan–protein interactions is generally low, with K_d values of mM to μM levels [1]. In nature, many GBPs are oligomeric or membrane-associated proteins, which allow aggregation of the GBP in the plane of the membrane. Many of the glycan ligands for GBPs are also multivalent. The interaction of multiple subunits with a multivalent display of glycans raises the affinity of the interaction by several orders of magnitude under the physiological conditions. However, most of the currently used scoring functions may not have comparable performance for the individual ‘‘weak binder’’ as for

small molecules with submicromolar to picomolar affinities [6]. There is a thirst for a general scoring function that has equivalent performance on the weak interactions between glycans and GBP.

Robust scoring functions for protein–ligand interactions with quantum chemical charge models

In a previous study, we have employed two well-established quantum chemical approaches, namely the restrained electrostatic potential (RESP) and the Austin-model 1-bond charge correction (AM1-BCC) methods, to obtain atomic partial charges [7] for deriving new scoring functions for the automated molecular docking software package, AutoDock4 [8], which has been widely adopted in virtual screening of drug candidates and prediction of ligand binding poses in protein pockets.

The AutoDock4 scoring function comprises five energetic terms: the van der Waals interactions, the hydrogen bonding interactions, the electrostatic interactions, the desolvation energy, and the torsional entropy. The AutoDock4 scoring function predicts the binding free energy with the following formula:

$$\begin{aligned}\Delta G_{bind} = {} & W_{vdw} \times \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ & + W_{H-bond} \times \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ & + W_{estat} \times \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \\ & + W_{desol} \times \sum_{i,j} \left(S_i V_j + S_j V_i \right) e^{\left(-r_{ij}^2 / 2\sigma^2 \right)} \\ & + W_{tor} \times N_{tors}\end{aligned}$$

The atomic charges used to evaluate the electrostatics energy term of the original 2007 AutoDock4 scoring function were calculated using the Gasteiger charge model [9], whose primary advantages lie in its simplicity and speed. However, such charge calculations can generate atomic charges that are less accurate than those determined by quantum chemical methods.

We implemented two variants of AutoDock4 scoring functions using two well-established charge models for ligands, namely, RESP [9] and AM1-BCC [11 – 12], that have been used widely in molecular dynamics simulations with the AMBER force field. RESP is a two-stage restrained electrostatic fit charge model, while AM1-BCC is a quick and efficient semi-empirical atomic charge model that aims to achieve the accuracy of RESP. The atomic

charges of proteins were retrieved from the AMBER parm99SB force field parameters, which were mainly derived by the RESP methodology [13 – 15]. The abbreviations ''AP'' for AM1-BCC (ligand)/Amber PARM99SB (protein) and ''RP'' for RESP (ligand)/Amber PARM99SB (protein) will be used in the following sections.

In combination with robust regression analysis and outlier exclusion, our protein–ligand free energy regression with the robust AP (RAP) charge model achieves lowest root-mean-squared error of 1.637 kcal/mol for the training set of 147 complexes and 2.176 kcal/mol for the external test set of 1.427 complexes. The assessment for binding pose prediction with the 100 external decoy sets indicates very high success rate of 87% with the criteria of predicted root-mean-squared deviation of less than 2 Å (Table 1 and Figure 1). The success rates and statistical performance of our robust scoring functions are only weakly dependent on the type of protein–ligand interactions (Table 2).

Table 1. Success rates of binding site prediction by different scoring functions[a] [7]

	success rate (%) for different RMSD criteria				
scoring function	≦1Å	≦1.5Å	≦2Å	≦2.5Å	≦3Å
DrugScore^{CSD}	83	85	87		
AutoDock4^{RAP}	83	85	87	87	87
AutoDock4^{RGG}	80	82	86	86	86
AutoDock4^{RRP}	79	81	84	85	85
original AutoDock4^{GG}	74	76	79	79	79
Cerius2/PLP	63	69	76	79	80
SYBYL/F-Score	56	66	74	77	77
Cerius2/LigScore	64	68	74	75	76
DrugScore	63	68	72	74	74
Cerius2/LUDI	43	55	67	67	67
X-Score	37	54	66	72	74
AutoDock3	34	52	62	68	72
Cerius2/PMF	40	46	52	54	57
SYBYL/G-Score	24	32	42	49	56
SYBYL/ChemScore	12	26	35	37	40
SYBYL/D-Score	8	16	26	30	41

[a] Except for the results of the AutoDock4 scoring functions, the results of DrugScore^{CSD} and other scoring functions were taken from Velec *et al.*[26][18] and Wang *et al.* [19], respectively.
[b] Scoring functions are sorted by the number of cases under 2Å.

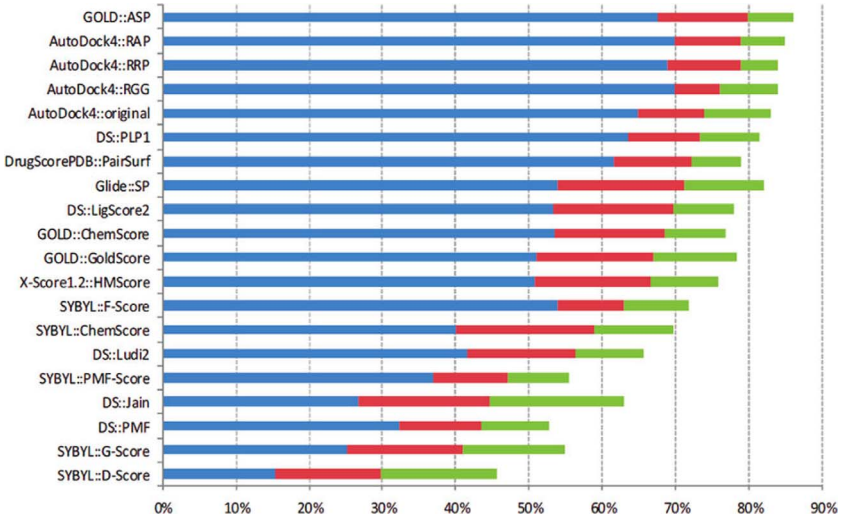

Figure 1. Comparison of the success rates of AutoDock4 scoring functions and 16 scoring functions provided by Cheng *et al* [20]. The cutoffs are rmsd < 1.0 Å (blue bars), < 2.0 Å (red bars), and < 3.0 Å (green bars), respectively. The native binding poses of ligands were included in the decoy sets. Scoring functions are sorted by the number of cases under 2 Å [7].

Table 2. Success rates of binding pose prediction of various scoring functions[a] on three classes of complexes [7]

	success rate (%; RMSD ≤2Å)			
	Overall	hydrophilic	mixed	hydrophobic
scoring function	(100)	(44)	(32)	(24)
AutoDock4^{RAP}	87	89	91	79
AutoDock4^{RGG}	86	86	91	79
AutoDock4^{RRP}	84	84	91	75
original AutoDock4^{GG}	79	77	81	79
Cerius2/PLP	76	77	78	71
SYBYL/F-Score	74	75	75	71
Cerius2/LigScore	74	77	75	67
DrugScore^{PDB}	72	73	81	58
Cerius2/LUDI	67	75	66	54
X-Score	66	82	59	46
AutoDock3	62	73	53	54
Cerius2/PMF	52	68	44	33
SYBYL/G-Score	42	55	34	29
SYBYL/ChemScore	35	32	34	42
SYBYL/D-Score	26	23	28	29

[a] Data were adopted from Wang *et al.*[19] except for AutoDock4 scoring functions.
[b] Scoring functions are sorted according to the overall success rates.

Recognition of glycan by proteins is a key to the specificity in glycobiology

Binding of glycans to proteins represents the major way in which the information contained in glycan structures is recognised, deciphered, and put into biological action [1]. The structures of hundreds of glycan–protein complexes have been determined by X-ray crystallography and NMR spectroscopy. In most cases, the glycan-binding sites typically accommodate one to four sugar residues. Unveiling the three-dimensional structure of a glycan–protein complex can reveal much about the specificity of binding, changes in conformation that take place on binding, and the contribution of specific amino acids to the interaction.

Hydrophobic interactions are very common in glycan–protein complexes and can involve aromatic residues as well as alkyl side chains of amino acids in the binding pocket [1]. Since the forces involved in the binding of a glycan to a protein are the same as for the binding of a ligand to its receptor (hydrogen bonding, electrostatic or charge interactions, van der Waals interactions, and dipole attraction), it is tempting to try to calculate their contribution to overall binding energy. Unfortunately, calculating the free energy of association is difficult for several reasons, including problems in defining the conformation of the unbound versus the bound glycan, changes in bound water within the glycan and the binding site, and conformational changes in the GBP upon binding. To take the first step to tackle these problems, we tested the capability of our established AutoDock4^{RAP} scoring function to predict the binding affinities of glycans to GBP.

Performance of AutoDock4^{RAP} on predicting binding affinities of glycans to GBP

GBP can be broadly classified into two major groups: glycosaminoglycan-binding proteins and lectins. Because glycosaminoglycan-binding proteins do not have shared structural features, we applied the AutoDock4^{RAP} scoring function to the crystal structures of glycan–lectin complexes for which the binding affinities have been determined experimentally [16].

Lectins tend to recognise specific terminal aspects of glycan chains by fitting them into shallow but relatively well-defined binding pockets, namely, ‘‘carbohydrate-recognition domains’’ (CRD) that often retain specific features of primary amino acid sequence or three-dimensional structure [1]. The binding affinities to a single CRD in many lectins appear to be low (with K_d values in the micromolar range).

During the initial preparation work on the 23 complex structures for subsequent docking, we did not include the crystal structure with PDB code 1EN2 because the frequently appeared missing residues in the protein coordinates led to abrupt termination of the process. Among the 22 crystal structures used in the current validation study (Table 3), four complexes have glycosylated residues (1AXO, 1AX1, 1AX2, and 1AXZ). The covalently linked oligosaccharides are excluded from the analyses since they do not serve as ligands for the proteins.

Table 3. Validation of AutoDock4^{RAP} on glycan–lectin complexes.
The AutoDock4^{RAP} scoring function was applied to the crystal structures of glycan–lectin complexes for which the binding affinities have been determined experimentally [15]. The refined ligand binding modes of the complexes used in the study have root-mean-square deviations (RMSD) no more than 1.21 Å in reference to the corresponding crystal binding modes, and the refined free energy of binding for the 22 glycan–lectin complexes has a root-mean-squared error of 1.606 kcal/mol in reference to the experimental values. [a] Complex with crystal packing effect at the binding site.

PDB ID	Protein name	ΔG_{exp}	Rescore	Refine		Docking ^{rank1}		Docking ^{rank2}		Docking ^{rank3}		Ligand in crystal
		(kcal/mol)	ΔG	ΔG	RMSD	ΔG	RMSD	ΔG	RMSD	ΔG	RMSD	
1J4U[a]	**Artocarpin**	**-4.36**	-2.95	**-3.69**	0.73	**-7.41**	26.95	**-6.65**	1.35	**-6.48**	26.76	O1-Methyl-Mannose
5CNA[a]	**Concanavalin A**	**-5.31**	-4.35	**-4.76**	0.35	**-6.53**	0.98	**-5.84**	16.85	**n.a.**	n.a.	O1-Methyl-Mannose
1GIC[a]	**Concanavalin A**	**-4.61**	-4.57	**-4.82**	0.36	**-6.37**	0.77	**-6.13**	1.75	**-5.46**	17.51	Methyl-α-D-Glucopyranoside
1QDO[a]	**Concanavalin A**	**-6.81**	-3.35	**-4.70**	0.69	**-7.26**	3.12	**-6.92**	2.05	**-6.59**	2.94	(α-D-Man)–(O1-Methyl-Man)
1QDC[a]	**Concanavalin A**	**-5.31**	-4.20	**-4.52**	0.51	**-8.18**	16.57	**-6.55**	2.48	**-6.26**	1.24	(α-D-Mannose)–(O1-Methyl-Mannose)
1ONA[a]	**Concanavalin A**	**-7.41**	-1.54	**-5.50**	0.50	**-6.83**	2.32	**-6.82**	1.74	**-5.57**	3.75	(α-D-Mannose)–(O1-Methyl-Mannose)–(α-D-Mannose)
1DGL[a]	**Lectin**	**-8.21**	-4.41	**-5.24**	0.37	**-6.77**	1.96	**-6.47**	5.25	**-6.26**	2.36	(α-D-Mannose)–(O1-Methyl-Mannose)–(α-D-Mannose)
1AXZ[a]	**Lectin**	**-4.35**	-2.33	**-3.46**	0.81	**-6.27**	17.32	**-5.54**	0.99	**-5.18**	10.11	α-D-Galactose; β-D-Galactose
1AX0[a]	**Lectin**	**-4.28**	-3.36	**-4.86**	0.62	**-7.22**	16.10	**-6.72**	1.11	**-5.93**	16.49	N-Acetyl-2-Deoxy-2-Amino-Galactose
1AX1[a]	**Lectin**	**-4.50**	-1.75	**-2.43**	0.69	**-7.32**	11.97	**-5.86**	3.11	**-5.08**	2.66	(β-D-Glucose)–(β-D-Galactose)
1AX2[a]	**Lectin**	**-5.43**	-1.90	**-2.50**	0.90	**-5.97**	18.72	**-5.93**	20.84	**-5.38**	3.26	[2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose]–[β-D-Galactose]
2BQP	**Lectin**	**-3.35**	-3.09	**-3.84**	0.42	**-6.35**	0.91	**-5.69**	19.18	**-5.19**	19.39	α-D-Glucose
1BQP	**Lectin**	**-3.97**	-3.74	**-4.14**	0.28	**-7.23**	8.30	**-6.50**	19.04	**-6.47**	9.76	$(\alpha\text{-D-Mannose})_2$
1QF3	**Agglutinin**	**-4.06**	-1.53	**-2.85**	0.79	**-6.77**	11.82	**-5.14**	9.12	**-5.00**	1.88	Methyl-β-Galactose
2PEL[a]	**Agglutinin**	**-4.25**	-1.16	**-2.66**	0.63	**-6.51**	2.76	**-5.05**	3.04	**-4.65**	9.12	α-Lactose (LBT) *3 +β-Lactose (LAT) *1
1EHH[a]	**Agglutinin isolectin VI**	**-5.11**	-0.97	**-2.81**	1.21	**-6.34**	13.14	**-5.09**	12.71	**-4.54**	15.03	$(\text{N-Acetyl-D-Glucosamine})_3$
1K7U	**Agglutinin isolectin III**	**-5.11**	-0.28	**-3.58**	1.01	**-6.59**	16.61	**-5.35**	17.34	**-5.33**	13.93	$(\text{N-Acetyl-D-Glucosamine})_2$
1KUJ[a]	**Agglutinin**	**-4.14**	-4.27	**-5.40**	0.45	**-6.95**	11.14	**-6.07**	10.41	**-5.83**	12.84	O1-Methyl-Mannose
1GZC	**Lectin**	**-4.76**	-1.60	**-2.42**	0.73	**-7.54**	11.53	**-6.45**	18.26	**-5.81**	19.97	β-Lactose
1HKD[a]	**Lectin**	**-3.83**	-3.45	**-3.89**	0.25	**-6.21**	19.16	**-5.03**	7.26	**-5.02**	1.14	Methyl-α-D-Glucopyranoside
4GAL[a]	**Human galectin-7**	**-4.62**	+16.23	**-2.95**	1.06	**-7.06**	11.17	**-5.86**	10.9	**-4.65**	9.55	(β-D-Galactose)–(β-D-Glucose)
5GAL[a]	**Human galectin-7**	**-4.40**	+1.23	**-2.76**	0.95	**-6.60**	12.88	**-6.46**	9.52	**-6.32**	7.58	(N-Acetyl-D-Glucosamine)–(β-D-Galactose)

In the analyses of free energy of binding, we carried out three stages of measures. The first one was to ‘‘rescore’’ the original binding mode in the crystal coordinates without moving the ligand. Next, we allowed the ligand to move in a restricted space using local search parameters, without torsional or rotational modification; thereby ‘‘refining’’ the ligand to a potential position with lower binding free energy in the crystal binding site. The refined ligand binding modes of the complexes used in the study have root-mean-square deviations

(RMSD) no more than 1.21 Å in reference to the corresponding crystal binding modes (Table 3). The refined free energy of binding for the 22 glycan–lectin complexes has a root-mean-squared error of 1.606 kcal/mol in reference to the experimental values.

We also carried out a comprehensive search, rendering the ligand to have translational and rotational alterations, "docking" the ligand to a larger space in the binding pocket of the protein. Because the lectin structures used in the study all have shallow CRD with relatively large area, we enclosed the entire CRD for the docking analysis of each complex.

As we inspected the structures of these glycan–lectin complexes, we found out that 17 out of the 22 complexes used in the study have crystal packing effects at the binding sites, that is, the ligand (glycan) bound to the protein (lectin) at the interface of different symmetry mates when we generated them using the crystallographic symmetry information (Table 3 and Figure 2). The crystal packing effect could be indicative of an artefact in the crystal binding mode for certain complex structures. For the glycan–lectin complexes, however, the crystal packing effects at the binding sites do not necessarily reflect the dockability of the glycan ligands. Using the AutoDock4^{RAP} scoring function, we can still reproduce the crystal binding modes (with RMSD less than 2 Å) after comprehensive docking analyses on many of the complexes with such crystal packing effect (Table 3 and Figures 3 – 5). This could be due to the spatial arrangement in the recognition of glycan by the shallow CRD of lectin, which is quite different from that of the proteins with deep ligand binding pockets.

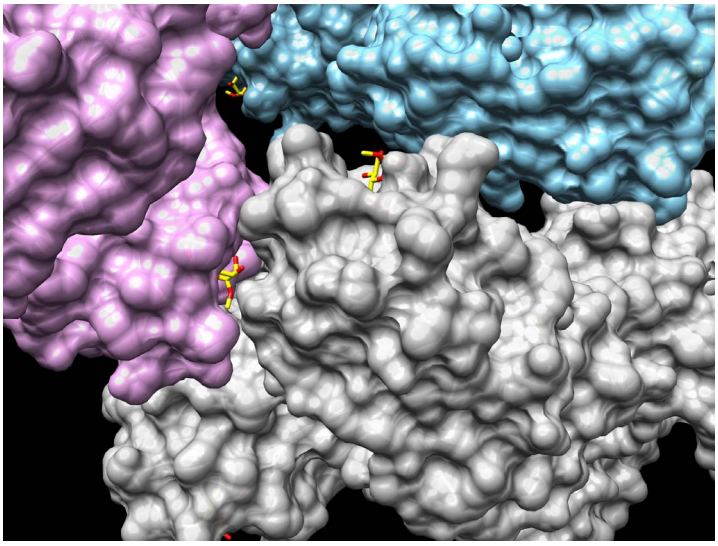

Figure 2. Representative image of crystal packing effects at the binding site. The symmetry mates of artocarpin (PDB code: 1J4U) were generated using the crystal symmetry information and are shown in different colours. The O1-methyl-mannose bound to the interface of the violet and grey molecules indicates crystal packing effect at this ligand-binding site.

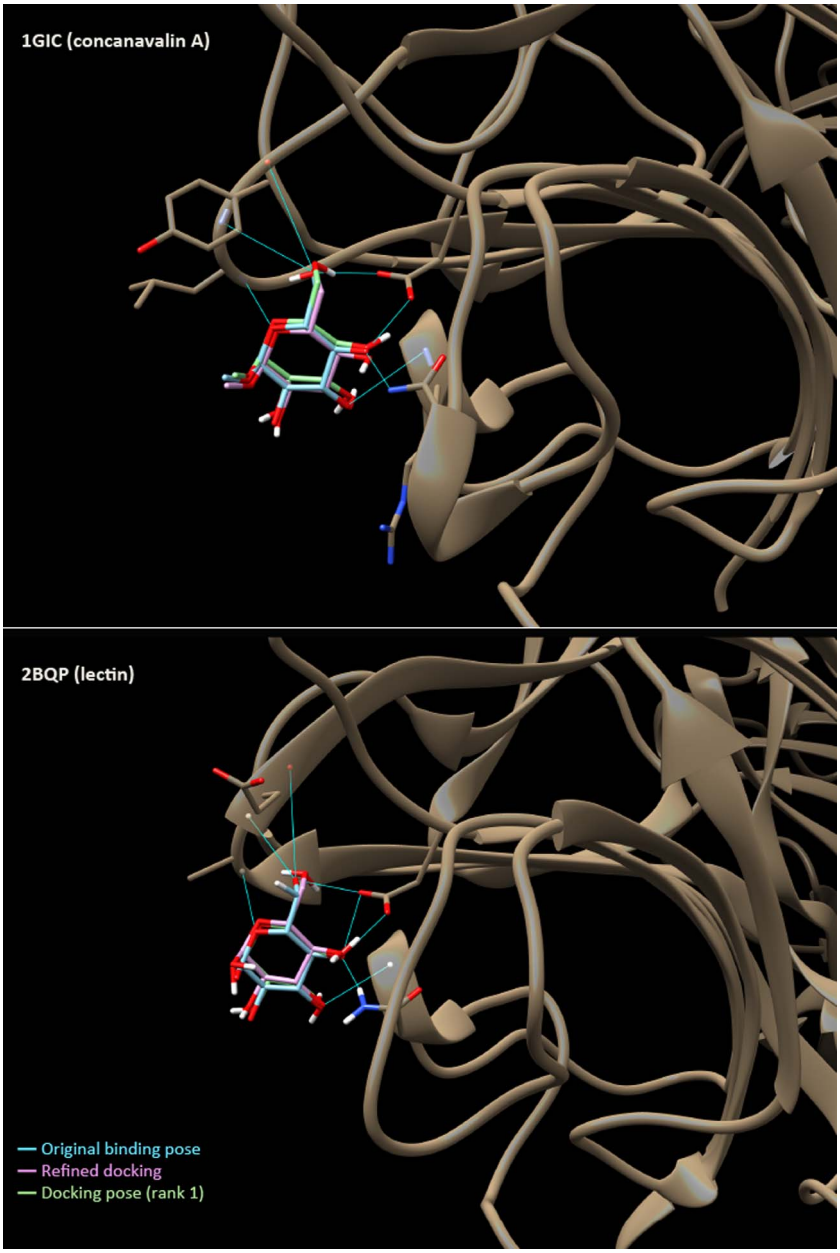

Figure 3. Crystal and predicted binding modes of the glycan ligands to concanavalin A (1GIC) and lectin (2BQP). Cyan lines indicate the hydrogen bonds formed between the glycan ligand and the protein.

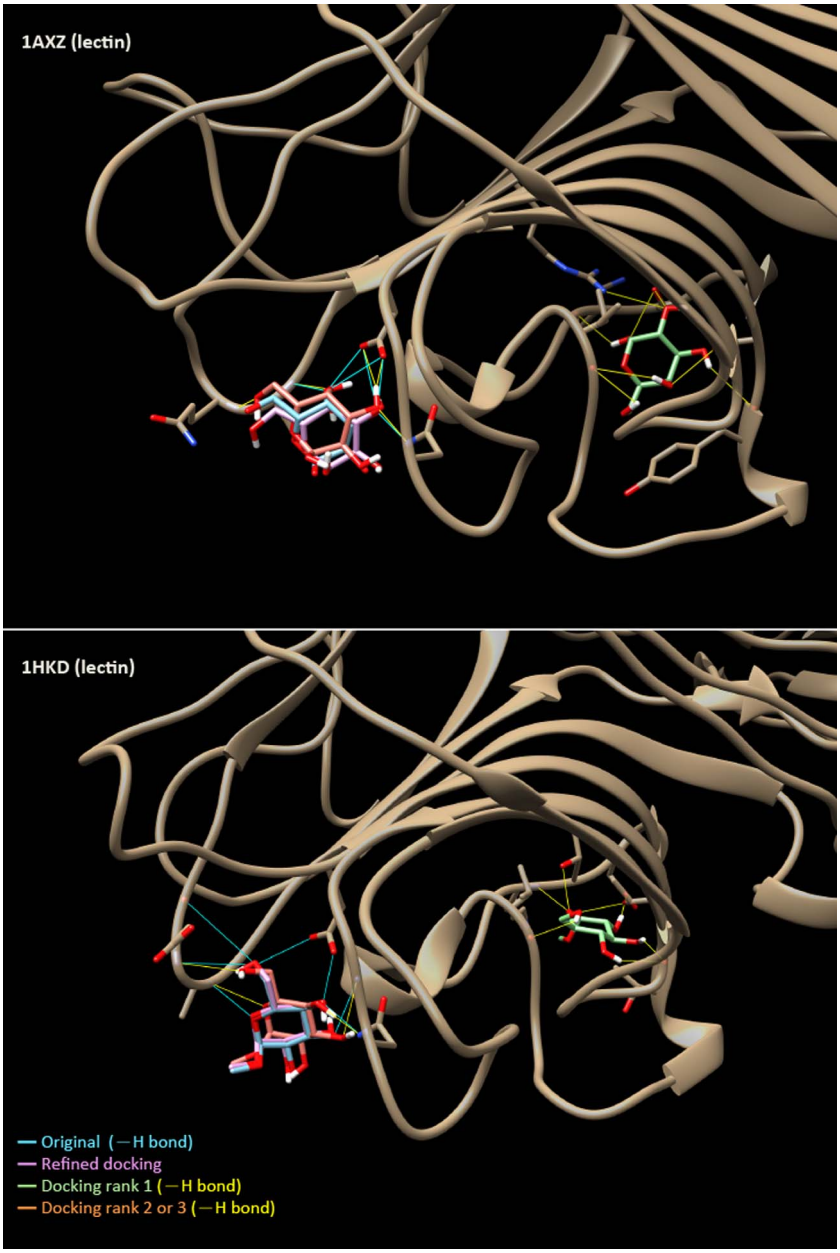

Figure 4. Crystal and predicted binding modes of the glycan ligands to lectins (1AXZ and 1HKD). Cyan lines indicate the hydrogen bonds formed between the protein and the ligand in the crystal binding mode, while yellow lines indicate the hydrogen bonds formed between the ligand in the predicted binding modes.

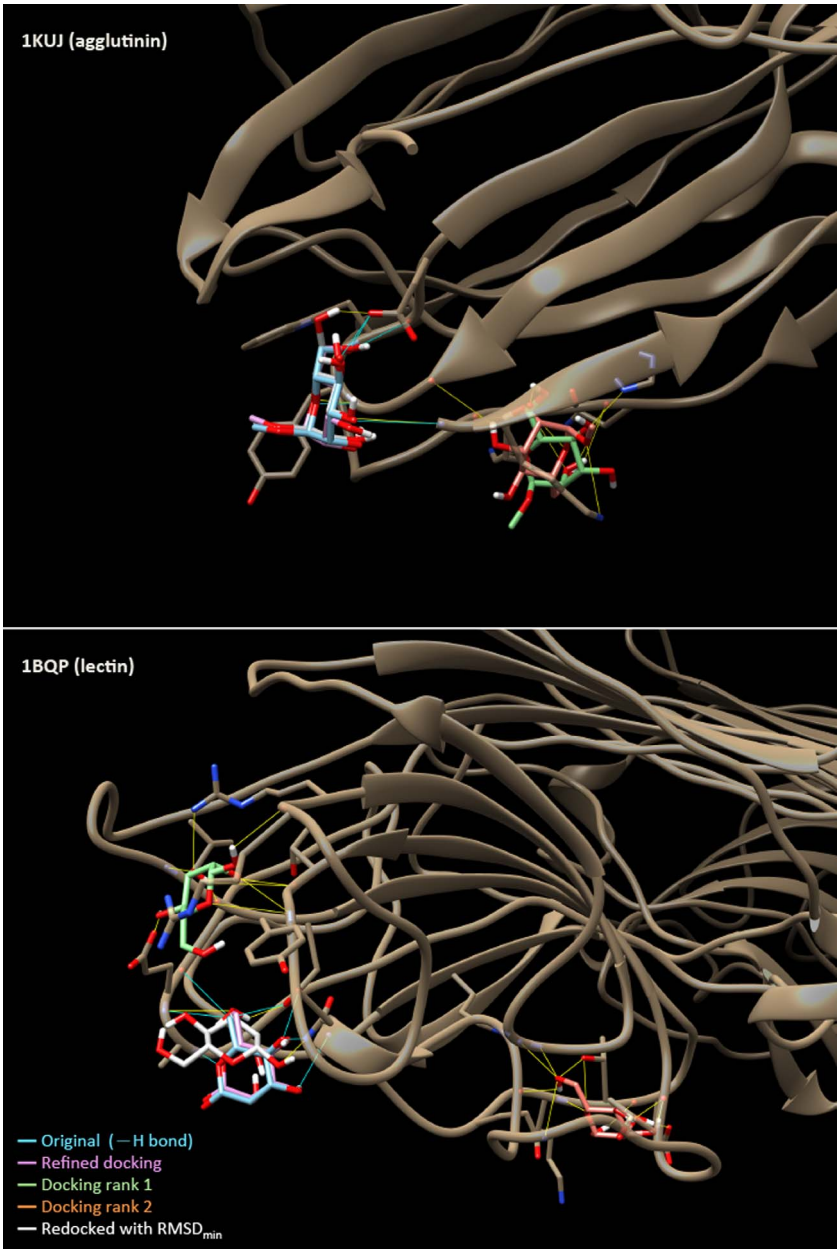

Figure 5. Crystal and predicted binding modes of the glycan ligands to lectins (1KUJ and 1BQP). Cyan lines indicate the hydrogen bonds formed between the protein and the ligand in the crystal binding mode, while yellow lines indicate the hydrogen bonds formed between the ligand in the predicted binding modes.

Potential application of AutoDock4^{RAP} on the glycan–GBP systems

During the initial development of the molecular biology, studies of glycans lagged far behind those of other major classes of molecules [1]. This was in large part due to their inherent structural complexity, the great difficulty in determining their sequences, and the fact that their biosynthesis could not be directly predicted from a DNA template. This delayed development in experimental methodology could reflect the performance of computational work on glycobiology. Kerzmann *et al*. presented a method specifically designed for the docking of carbohydrate-like compounds [15]. In contrast, although the AutoDock4^{RAP} scoring function was not tailored for glycan–GBP complexes, the current validation study revealed the capability of AutoDock4^{RAP} to predict the binding affinities for GBP. The use of a general, robust scoring function can facilitate the virtual screening on compounds with more diverse chemical scaffolds, the essential work in the very beginning of rational drug design, for the various GBP.

One of the critical issues in calculating the free energy of binding lies in the conformational changes in the protein (GBP) upon ligand (glycan) binding. This can be addressed using the relaxed complex scheme [17 – 19] since the atomic charge models used in the current scoring function are those have been widely used in molecular dynamics simulations with the AMBER force field.

In the complexes used in the study, there are a considerable number of hydrogen bonds between the glycans and lectins, as demonstrated in both the crystal and predicted binding modes (Figures 3 – 5). The contribution of hydrogen bonding to the binding affinities of glycans to GBP can be further examined with energy decomposition analysis when the relaxed complex scheme is used. Nevertheless, the amino acid residues with potentials to form hydrogen bonds with glycans may also serve as target residues in the design of lead compounds to inhibit glycan–GBP interactions.

Summary

The physiological and pathological significances of glycan–binding proteins are drawing more and more attention, both in basic and applied sciences. In the current study, we have demonstrated the capability of a general, robust scoring function, AutoDock4^{RAP}, to predict the binding affinities for glycan–binding proteins, without any calibration to this specific class of protein-ligand interactions. The free energy of binding for the 22 glycan–lectin complexes has a root-mean-squared error of 1.606 kcal/mol in reference to the experimental values. The use of AutoDock4^{RAP} can therefore facilitate the virtual screening on compounds with more diverse chemical scaffolds, as well as further rigorous studies, such as those with use of relaxed complex scheme and energy decomposition analysis.

References

[1] Varix A., Cummings R.D., Esko J.D., Freeze, H.H., Stanley, P., Bertozzi, C. R., Hart, G.W., Etzler, M.E., Eds. (2009) *Essentials of Glycobiology*. 2nd edition. Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).

[2] Gilson, M.K., Zhou, H.-X. (2007) Calculation of Protein–Ligand Binding Affinities. *Ann. Rev. Biophys. Biomol. Struc.* **36**:21 – 42. doi: http://dx.doi.org/10.1146/annurev.biophys.36.040306.132550.

[3] Gilson, M.K., Given, J.A., Bush, B.L., McCammon, J.A. (1997) The statistical–thermodynamic basis for computation of binding affinities: A critical review. *Biophysical J.* **72**:1047 – 1069. doi: http://dx.doi.org/10.1016/S0006-3495(97)78756-3.

[4] Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* **19**:1639 – 1662. doi: http://dx.doi.org/10.1002/(sici)1096-987x(19981115)19:14<1639::aid-jcc10>3.0.co;2-b.

[5] Huey, R., Morris, G.M., Olson, A.J., Goodsell, D.S. (2007) A semi-empirical free energy force field with charge-based desolvation. *J. Comp. Chem.* **28**:1145 – 1152. doi: http://dx.doi.org/10.1002/jcc.20634.

[6] Wang, J.C., Lin, J.H. (2013) Scoring functions for prediction of protein-ligand interactions. *Curr. Pharm. Des.* **19**:2174 – 2182. doi: http://dx.doi.org/10.2174/1381612811319120005.

[7] Wang, J.-C., Lin, J.-H., Chen, C.-M., Perryman, A.L. Olson, A.J. (2011) Robust scoring functions for protein-ligand interactions with quantum chemical charge models. *J. Chem. Inf. Model.* **51**:2528 – 2537. doi: http://dx.doi.org/10.1021/ci200220v.

[8] Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., and Olson, A.J. (2009) AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **30**(16):2765 – 2791. doi: http://dx.doi.org/10.1002/jcc.21256.

[9] Gasteiger, J., Marsili, M. (1980) Iterative partial equalization of orbital electronegativity – a rapid access to atomic charges. *Tetrahedron* **36**:3219 – 3228. doi: http://dx.doi.org/10.1016/0040-4020(80)80168-2.

[10] Bayly, C.I., Cieplak, P., Cornell, W., Kollman, P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the resp model. *J. Phys. Chem.* **97**:10269 – 10280. doi: http://dx.doi.org/10.1021/j100142a004.

[11] Jakalian, A., Bush, B.L., Jack, D.B., Bayly, C.I. (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comp. Chem.* **21**:132 – 146.
doi: http://dx.doi.org/10.1002/(SICI)1096-987X(20000130)21:2<132::AID-JCC5>3.0.CO;2-P

[12] Jakalian, A., Jack, D.B., Bayly, C.I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comp. Chem.* **23**:1623 – 1641.
doi: http://dx.doi.org/10.1002/jcc.10128.

[13] Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. (1995) A second generation force-field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**:5179 – 5197.
doi: http://dx.doi.org/10.1021/ja00124a002.

[14] Ponder, J.W., Case, D.A. (2003) Force fields for protein simulations. *Adv. Prot. Chem.* **66**:27 – 85.
doi: http://dx.doi.org/10.1016/S0065-3233(03)66002-X.

[15] Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J.M., Kollman, P.A. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comp. Chem.* **24**:1999 – 2012.
doi: http://dx.doi.org/10.1002/jcc.10349.

[16] Kerzmann, A., Fuhrmann, J., Kohlbacher, O., Neumann, D. (2008) BALLDock/SLICK: A new method for protein-carbohydrate docking. *J. Chem. Inf. Model.* **48**:1616 – 1625.
doi: http://dx.doi.org/10.1021/ci800103u.

[17] Lin J.H., Perryman A.L., Schames J.R., McCammon J.A. (2002) Computational drug design accommodation receptor flexibility: The relaxed complex method. *J. Am. Chem. Soc.* **68**:47 – 62.

[18] Lin J.H., Perryman A.L., Schames J.R., McCammon J.A. (2003) The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers* **68**:47 – 62.
doi: http://dx.doi.org/10.1021/ja0260162.

[19] Lin J.H. (2011) Accommodating protein flexibility for structure-based drug design. *Curr. Top. Med. Chem.* **11**:171 – 178.
doi: http://dx.doi.org/10.2174/156802611794863580

[20] Cheng, T., Li, X., Li, Y., Liu, Z., Wang, R. (2009). Comparative Assessment of Scoring Functions on a Diverse Test Set. *J. Chem. Inf. Model.* **49**:1079 – 1093. doi: http://dx.doi.org/10.1021/ci9000053.