Molecular Docking of Carbohydrate Ligands to Antibodies: Structural Validation against Crystal Structures

Mark Agostino,[†,‡] Cassandra Jene,[†,‡] Tristan Boyle,[†] Paul A. Ramsland,*[,§,||,⊥] and Elizabeth Yuriev*[,†]

Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia, Centre for Immunology, Burnet Institute, Studley Road, Heidelberg, VIC 3084, Australia, Department of Surgery Austin Health, University of Melbourne, Studley Road, Heidelberg, VIC 3084, Australia, and Department of Immunology, Monash University, Alfred Medical Research and Education Precinct, Commercial Road, Melbourne, VIC 3004, Australia

Received October 3, 2009

Cell surface glycoproteins play vital roles in cellular homeostasis and disease. Antibody recognition of glycosylation on different cells and pathogens is critically important for immune surveillance. Conversely, adverse immune reactions resulting from antibody−carbohydrate interactions have been implicated in the development of autoimmune diseases and impact areas such as xenotransplantation and cancer treatment. Understanding the nature of antibody−carbohydrate interactions and the method by which saccharides fit into antibody binding sites is important in understanding the recognition process. In silico techniques offer attractive alternatives to experimental methods (X-ray crystallography and NMR) for the study of antibody−carbohydrate complexes. In particular, molecular docking provides information about protein−ligand interactions in systems that are difficult to study with experimental techniques. Before molecular docking can be used to investigate antibody−carbohydrate complexes, validation of an appropriate docking method is required. In this study, four popular docking programs, Glide, AutoDock, GOLD, and FlexX, were assessed for their ability to accurately dock carbohydrates to antibodies. Comparison of top ranking poses with crystal structures highlighted the strengths and weaknesses of these programs. Rigid docking, in which the protein conformation remains static, and flexible docking, where both the protein and ligand are treated as flexible, were compared. This study has revealed that generally molecular docking of carbohydrates to antibodies has been performed best by Glide.

INTRODUCTION

Cell surface glycoproteins play vital roles in cellular homeostasis and disease. The ability of antibodies (Ab) to recognize glycosylation on a variety of cellular structures is critically important for immune surveillance.[1] Carbohydrate epitopes on bacterial, viral, and tumor cells are important disease markers and targets for therapeutic[2−4] and diagnostic[4] antibodies and for vaccine development.[5,6] Conversely, certain Ab−carbohydrate interactions stimulate adverse immune reactions that may cause the development of autoimmune diseases[7] and detrimentally impact such areas as transplantation (including xenotransplantation[8−10]) and cancer treatment by therapeutic antibodies.[11,12] Despite its clinical importance, the knowledge of the structural basis of antibody recognition of carbohydrates is limited. Structures of Ab−carbohydrate complexes can be investigated experimentally by X-ray crystallography or NMR. However, the flexibility of carbohydrates makes the crystallization of these complexes difficult, and NMR is more suited to determining the conformation of the bound carbohydrate, rather than the nature of the interactions occurring with a protein. In silico techniques offer attractive alternatives for the study of Ab−carbohydrate interactions. Specifically, molecular docking can provide information about protein−ligand interactions in systems that are difficult to study experimentally. Molecular docking identifies binding modes of a ligand by flexing and orienting it in a protein active site and scoring them. Available docking methods differ mainly in ligand placement in the active site, exploration of conformational space, and scoring or binding affinity estimation.

Evaluation of the predictive accuracy of docking programs routinely attracts considerable attention.[13,14] Evaluation studies are performed for individual programs (e.g., for GOLD[15]) or comparatively (i.e., for a range of programs). Typically, they assess the accuracy of ligand binding compared to crystal structures (pose prediction) and/or the ability to identify known binders from a diverse set of binding/nonbinding compounds. Docking evaluations are commonly done using diverse test sets of protein−ligand complexes (studies[14,16−18] and references therein). Given the known target-dependence of scoring functions[19] employed in docking, evaluations are often focused on a given type of protein target.[14,18] An evaluation designed for a particular ligand type/chemotype would significantly benefit a study aimed at delineating fundamental interaction phenomena, such as Ab−carbohydrate recognition.

* Corresponding authors. E-mail: elizabeth.yuriev@pharm.monash.edu.au (E.Y.) and pramsland@burnet.edu.au (P.A.R.).
[†] Medicinal Chemistry and Drug Action, Monash University.
[‡] The first two authors contributed equally to this work.
[§] Burnet Institute.
[||] University of Melbourne.
[⊥] Department of Immunology, Monash University.

10.1021/ci900388a
Published on Web 12/08/2009

Table 1. Antibody−carbohydrate Crystal Structures Used as Test Cases

PDB code	carbohydrate ligand[a]	resolution (Å)
1CLY[2]	Fucα(1−2)Galβ(1−4) [Fucα(1−3)]GlcNAcβ-Non[b]	2.50
1M7D[29]	Rhaα(1−3)2-deoxy-Rhaα(1−3)GlcNAc	2.30
1MFA[30]	Galα(1−2)[Abeα(1−3)]Manα(1-OMe)	1.70
1MFC[31]	Galα(1−2)[Abeα(1−3)]Manα(1−4)Rha	2.10
1MFD[32]	Galα(1−2)[Abeα(1−3)]Manα(1-OMe)	2.10
1MFE[33]	Galα(1−2)[Abeα(1−3)]Man	2.00
1OP3[6]	Manα(1−2)Man	1.75
1S3K[3]	Fucα(1−2)Galβ(1−4)[Fucα(1−3)]GlcNAc	1.90
1UZ8[4]	Galβ(1−4)[Fucα(1−3)]GlcNAc	1.80
1ZLS[5]	Manα(1−2)Manα(1−2)Manα(1−2)Man	2.00
1ZLU[5]	Manα(1−2)Man	2.75

[a] All monosaccharides are D, except Fuc and Rha. [b] Non is the PDB 3 letter code for nonoate methyl ester.

There are some peculiarities to protein−carbohydrate interactions, compared to general protein−ligand and even protein−peptide interactions, that make carbohydrate docking particularly challenging. First, many carbohydrates are extremely flexible and carry a large number of hydroxyl groups. This leads to the formation of often extensive hydrogen-bonding networks, which significantly contribute to the overall binding free energy. Second, a distinctive feature of protein−carbohydrate interaction is the formation of crucial CH-π contacts between aromatic side chains of the protein and the C−H bonds of the carbohydrates (on their hydrophobic faces).[20] Therefore, widely used docking programs and scoring functions, which account differently for these types of interactions, may not preform as well for protein−carbohydrate complexes.

Thus, before docking can be confidently used for Ab−carbohydrate complexes, a comparison and validation of docking programs for their specific use in this area of research are needed.[18] Recently, the performance of AutoDock, DOCK, and Glide was evaluated for lectin−carbohydrate systems using seven crystal structures.[21] It was found that Glide marginally outperformed the other programs. Glide and GOLD have been also compared for docking of glucosyltriazolylacetamide to glycogen phosphorylase b.[22] Both programs performed well in predicting the experimentally determined binding poses for the ligand, with Glide, in particular, producing "impressive" results. To the best of our knowledge, these are the only docking comparisons executed for carbohydrate ligands. To address the need for more knowledge and information in this field, we have set out to evaluate several docking programs (Glide, AutoDock, GOLD, and FlexX) by cognate and cross-docking Ab−carbohydrate simulations and to evaluate the performance of rigid vs flexible protein docking. Our results demonstrated that generally Ab−carbohydrate docking has been performed well by Glide. GOLD and AutoDock had several problems and FlexX performed poorly.

METHODS

The structures of the Ab−carbohydrate complexes (Table 1) were reduced to just the fragment variable (Fv) portion containing the antibody binding site. Water molecules were removed, hydrogens added, and bond orders corrected using the Protein Preparation Wizard workflow implemented in Maestro 8.0 (Schrödinger, Inc.). Prime Refinement was used to predict side chains for antibodies containing residues with missing atoms. Ligands were extracted from the crystal structures within Maestro. Programs Glide 4.5, AutoDock 4.0 (with AutoDock Tools 1.4.5), FlexX 1.2 (implemented in Sybyl 7.3), and GOLD 3.1.1 were examined. All programs were used in their latest versions available at the time of this study. Grid boxes were centered on ligand centroids and default options were used unless stated otherwise. Input file preparation was carried out separately as prescribed by the manuals for each program (details below).

Rigid Receptor Docking. In this method, the ligand is treated as flexible while the protein is held rigid.

Glide. The Ab−carbohydrate complexes were employed to build energy grids using the default value of protein atom scaling (1.0) within a cubic box with a side of 34 Å. The ligand and protein were parametrized with the OPLS force field. Docking calculations were performed in Standard Precision mode, and the option to generate ring conformations was switched off in order to maintain the input conformations. Generated ligand poses were clustered within a root-mean-square deviation (rmsd) of 2.0 Å and scored by GlideScore.

GOLD. Receptor and ligand input files were prepared in Sybyl 7.3. Gasteiger−Marsilli charges were added to both the carbohydrate and antibody. The grid box was defined by the default radius. The docking was terminated when the top three solutions found were within an rmsd of 2.0 Å. The poses were scored by GoldScore.

AutoDock. Receptor and ligand input files were generated in AutoDock Tools (the user interface for AutoDock 4.0). Gasteiger−Marsilli charges were added to the ligand, and Kollman charges and solvation parameters were added to the protein. Search parameters were set to those previously established for AutoDock docking of carbohydrates:[23] 245 runs for each complex, 800 000 evaluations per run, and a population size of 200. The mutation and crossover rates were set to 0.80 and 0.02, respectively, and the maximal number of generations was 27 000. Elitism was set to 1 and the local search frequency to 0.06. The grid box was defined using the default radius.

FlexX. Receptor and ligand input files were prepared and the binding site defined as described for GOLD above. All parameters for the various stages of the incremental construction process and for the input file set up were set to defaults. The poses were ranked by ChemScore.

Flexible Receptor Docking. In this method, both the protein and ligand are treated as flexible.

Glide/Prime Flexible Receptor Docking. The flexible receptor docking is implemented in the Schrödinger Suite via the Glide/Prime Induced Fit Docking (IFD) workflow. However, the standard IFD workflow has limitations in the settings used in the Glide portion. To circumvent this, the ensemble protocol was adapted to manually perform the flexible docking task. After an initial round of rigid receptor docking using Glide (as described above for rigid docking), top 20 ligand poses were placed back into the antibody crystal structure, and protein residues within 5 Å of the carbohydrate were minimized using Prime. The ligand was redocked rigidly into each of the newly minimized protein binding site conformations as described above, and 20 poses

Table 2. Binding Site Topographies and Degrees of Carbohydrate Protrusion

PDB code	binding site topography	number of carbohydrate residues	number of residues protruding from the binding site
1CLY	valley	4	1[a]
1M7D	canyon	3	0
1MFA	crater	3	2
1MFC	crater	4	2
1MFD	crater	3	1
1MFE	crater	3	1
1OP3	canyon	2	0
1S3K	valley	4	0
1UZ8	canyon	3	0
1ZLS	crater	4	2
1ZLU	crater	2	0

[a] Noncarbohydrate residue.

per protein conformation were kept. The 400 generated poses were rescored based on IFDScore.[24]

Flexible Receptor Docking with AutoDock. The ligand and protein were prepared as described above for AutoDock rigid docking. The selection sphere, which determines protein residues to be treated as flexible, was defined by the centroid of the ligand with a radius of 5.0−6.0 Å, depending on the number of combined ligand and protein rotatable bonds.

Analysis of Docking Results. To determine antibody binding site topographies, molecular surfaces were generated for each antibody using Maestro. Binding sites were inspected visually and classified according to the scheme by Lee et al.[25] Docked poses were compared to the respective experimental structures by calculating heavy atom rmsd values in Maestro using the Superposition command and tabulated for both the top ranked and "best" (i.e., lowest rmsd) poses. An rmsd lower than or equal to 2.0 Å, compared to the corresponding crystal structure, was regarded as a successful prediction of the experimental binding mode.[13,16,18,26,27] Ab−carbohydrate intermolecular interactions were analyzed using LIGPLOT.[28] The maximum interatomic distances were 3.9 Å for van der Waals (vdW) contacts and 3.35 Å for hydrogen bonds. For all other settings the default was kept.

RESULTS

Test Set. *Choice of Ab−Carbohydrate Complexes.* The Protein Data Bank (PDB) contains crystal structures of a number of antibodies complexed with oligosaccharides associated with cancerous,[2−4] viral,[5,6] parasitic,[4] and bacterial[29−33] antigens. All high-resolution Ab−carbohydrate complexes selected for this study (Table 1) were chosen to have short two to four residue carbohydrates with a similar level of flexibility. Henceforth, the complexes are referred to by their PDB codes.

Binding Site Topography. In this study, all antibody binding sites belong to a cavity class (according to Webster et al.[34]). We have adopted the Lee et al.[25] algorithm to further define the binding sites based on familiar geological features. A range of binding site types (valleys, craters, and canyons) were chosen (Table 2). Antibodies in the 1CLY and 1S3K complexes contain the valley binding site type (Figure 1a), which cuts across the Fab region of the antibody with a large shallow opening being exposed to the solvent. The crater antibody binding sites in the 1MFA, 1MFC, 1MFD, 1MFE, 1ZLS, and 1ZLU complexes (Figure 1b) contain shallow bowl-shaped, approximately circular openings with sloping sides. The canyon antibody binding site type of the 1M7D, 1OP3, and 1UZ8 complexes (Figure 1c) is a steep cut across the Fab with end(s) of the canyon open and exposed to the solvent.

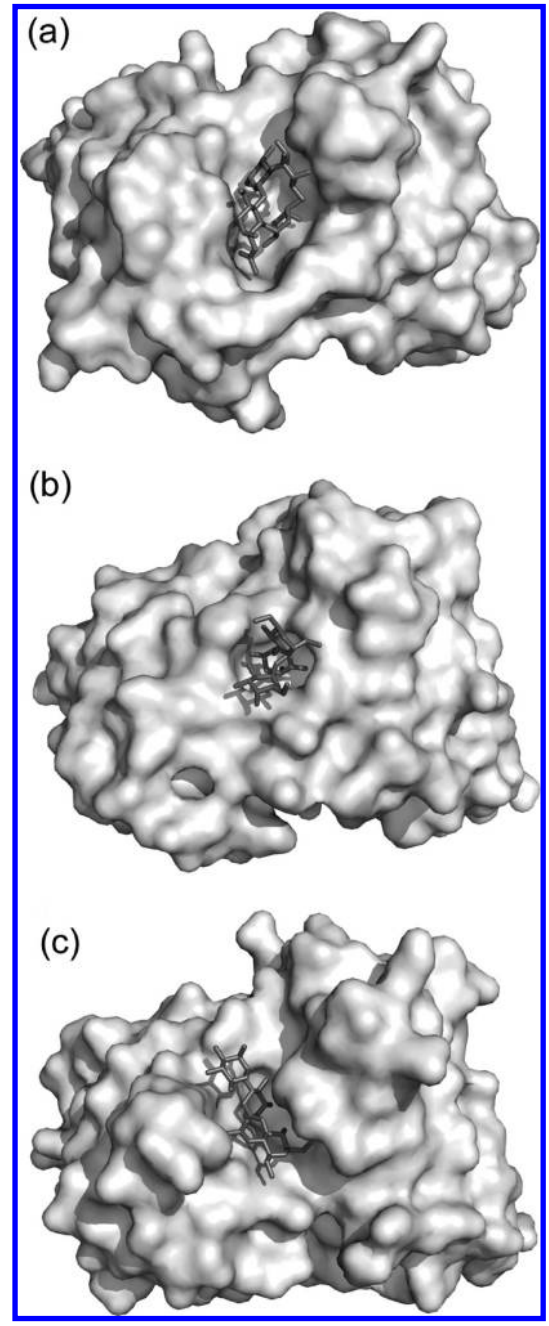

Figure 1. Antibody binding site topographies: (a) valley (example, antibody hu3S193, PDB code 1S3K), (b) crater (example, antibody Se155-4, PDB code 1MFD), and (c) canyon (example, antibody SYA/J6, PBD code 1M7D).

Rigid Receptor Docking Performance of Four Programs. *Glide.* The Glide (Grid-Based Ligand Docking with Energetics) docking program (Schrödinger, Inc.)[35] uses a hierarchical series of filters to perform a systematic search for possible locations of the ligand in the protein active site. Ligand conformational flexibility is handled by a conformational search. GlideScore is an empirical scoring function that considers an energetic contribution as well as the effects

Table 3. Docking Results: Rigid Receptor Docking

PDB code	rmsd values (Å)							
	Glide		GOLD		AutoDock		FlexX	
	top[a]	best[b]	top	best	top	best	top	best
1CLY[c]	4.4	0.9 (3)	9.76	4.23 (6)	7.20	4.3 (13)	2.71	2.71 (1)
1CLY[d]	0.7	0.7 (1)	7.13	6.1 (9)				
1M7D	0.6	0.6 (1)	1.02	1.0 (1)	1.47	1.47 (1)	6.3	5.0 (10)
1MFA	0.8	0.8 (1)	1.61	0.79 (4)	1.4	1.4 (1)	1.38	1.29 (3)
1MFC	2.05	1.8 (3)	0.89	0.67 (2)	1.21	1.21 (1)	6.50	2.51 (9)
1MFD	1.9	1.9 (1)	1.05	0.86 (2)	5.67	0.91 (2)	1.96	1.56 (3)
1MFE	0.80	0.80 (1)	6.21	0.96 (6)	1.17	1.17 (1)	1.79	1.79 (1)
1OP3	0.57	0.57 (1)	0.57	0.53 (1)	0.78	0.78 (1)	1.59	1.34 (2)
1S3K	0.45	0.45 (1)	0.85	0.85 (1)	1.14	1.14 (1)	−[e]	0.99 (12)
1UZ8	0.37	0.37 (1)	0.8	0.8 (1)	2.54	2.54 (1)	5.38	1.44 (7)
1ZLS	2.94	2.92 (4)	3.44	3.44 (1)	3.37	2.71 (4)	2.77	2.66 (2)
1ZLU	0.9	0.9 (1)	0.5	0.5 (1)	3.08	2.83 (2)	−[e]	1.43 (10)
average	1.43		2.42		2.63		3.37	

[a] Highest ranked pose is referred to as the top pose. [b] The pose achieving the lowest rmsd value is referred to as the best pose. The value in parentheses corresponds to pose rank. [c] When the whole ligand was docked. [d] When only carbohydrate portion was docked. [e] Top pose docked in incorrect location.

of hydrophobic, hydrogen bonding, and metal−ligand interactions and penalizes steric clashes.

Table 3 lists the rmsd values of docked poses generated by Glide for each of the 11 Ab−carbohydrate complexes in the test set. In 64% of cases the top pose deviated by less than 1.0 Å from the crystal structure. This data is consistent with the results reported in the original Glide publication from Friesner et al.,[35] where nearly 50% of the compounds from the test set were docked with rmsd values below 1.0 Å. In the same study, about three-quarters of the compounds were docked with an rmsd below 2.0 Å. This correlates well to our results: 73% of top poses deviating by less than 2.0 Å. In the majority of cases (73%), the top ranked pose was also the pose with the lowest rmsd to the crystal structure (i.e., the best pose). These results indicate that Glide is successful at predicting Ab−carbohydrate complexes. To strengthen this conclusion, we have carried out detailed analyses of the test cases, where the docked outcomes deviated most from the experimental structures: 1CLY, 1ZLS, and 1MFC.

The worst rmsd result from Glide rigid docking of the Ab−carbohydrate test set was that for 1CLY. In this case, a tetrasaccharide ligand with a large flexible noncarbohydrate tail is complexed with an antibody containing a valley-shaped binding site. Figure 2a shows the overlay of the crystal structure, top pose, and best docked poses. The top pose gave an rmsd of 4.4 Å, the binding mode closest to the experiment was achieved in pose 3 (0.9 Å). The main source of deviation for this ligand in both poses is the flexible nonoate methyl ester tail (Non 5). Located at the solvent exposed end of the binding site, this noncarbohydrate tail has very little (if any) conformational restriction compared to the rest of the ligand. All four hexose units of the top pose are placed accurately in the binding pocket, with rmsd for these residues of just 0.8 Å. When the ligand was docked without the tail, the top pose was the best and had the rmsd of 0.7 Å (Table 3). Shown in Figure 2b,c are the protein−ligand interactions in the binding site. Three hydrogen bonds (Gal 2-Tyr 35H, GlcNAc 3-Tyr 33H, and GlcNAc 3-Gln 52AH) are seen in both the crystal structure and the top docking pose. The alternative conformation of the flexible tail in the docked

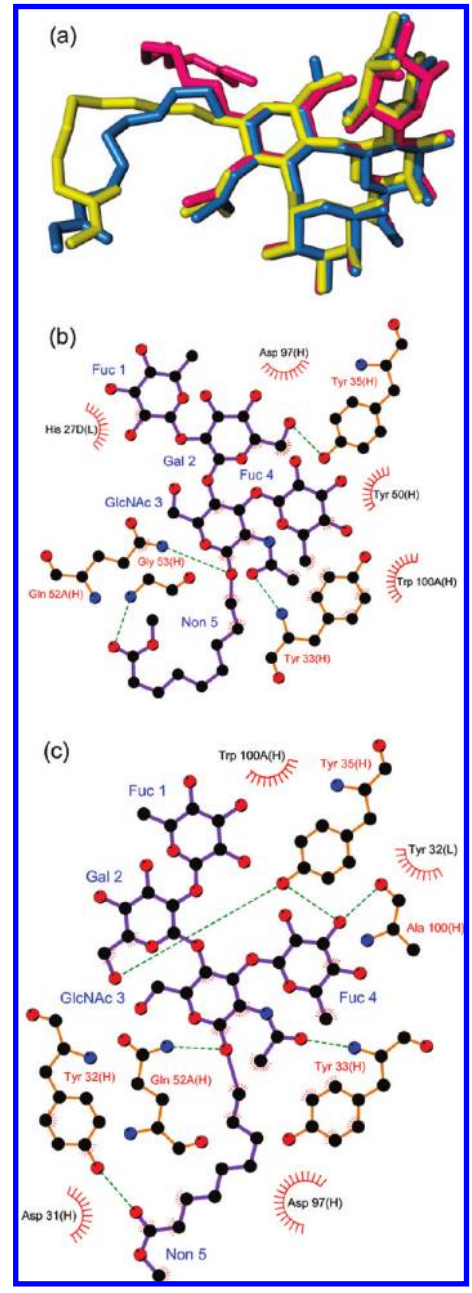

Figure 2. Glide rigid docking results for 1CLY. (a) Superimposition of the crystal structure (blue), top pose (magenta), and best pose (yellow). LIGPLOT diagrams show intermolecular interactions in the top pose (b) and crystal structure (c). In LIGPLOT diagrams: hydrogen bonds are shown as dotted green lines between the atoms involved; hydrophobic contacts are represented by an arc with red spokes radiating toward the ligand atoms they contact, the contacted atoms are shown with spokes radiating back. Hydrophobic vdW intermolecular contacts are defined in LIGPLOT as only those occurring between carbon/sulfur atoms and other carbon/sulfur atoms. Ligands are shown in purple. Protein bonds are shown in orange. Protein residues are designated by their number and chain identifier, as per the PDB files.

pose prevents interaction with Tyr 32H (crystal structure), which is compensated by a hydrogen bond with Gly 53H in the docked complex.

In the case of 1ZLS, a tetrasaccharide ligand bound to an antibody with a crater-shaped binding site, the top pose gave an rmsd of 2.94 Å. The best pose, ranked fourth, gave an rmsd of 2.92 Å (Figure 3a). The Manα(1−2)Man terminus of this sugar penetrates the binding site to the greatest extent and is more conformationally restricted than the remaining

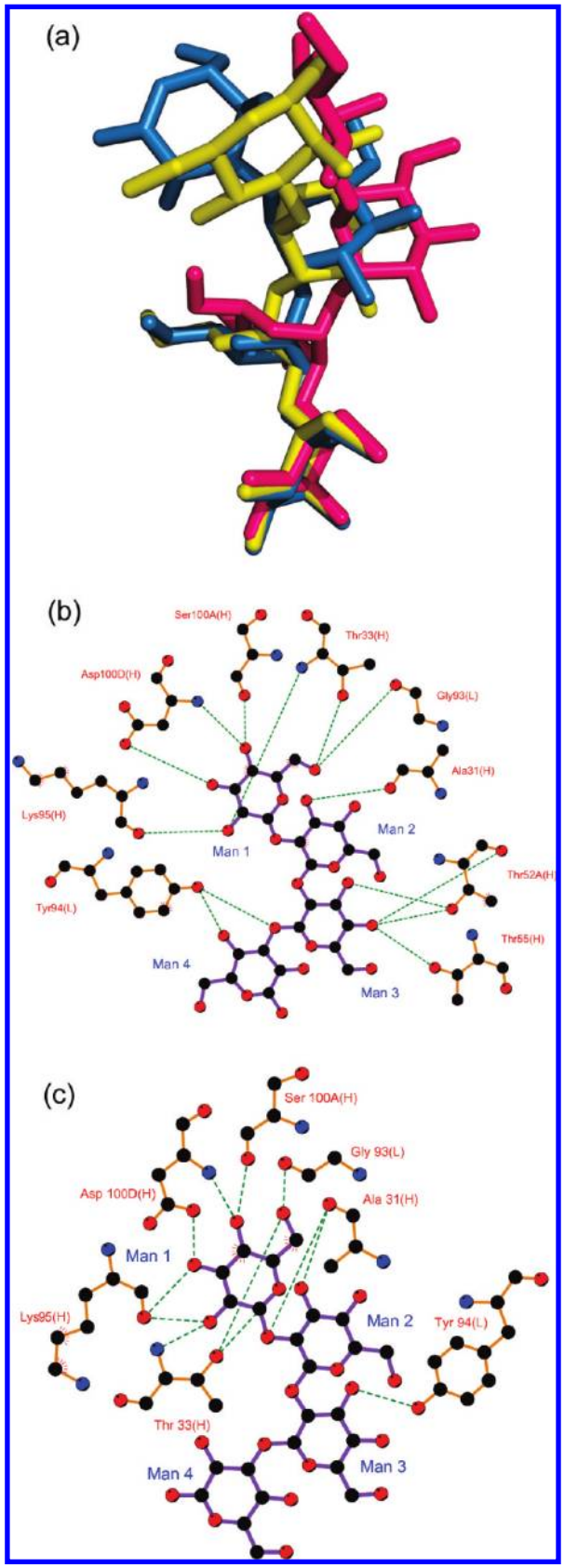

Figure 3. Glide rigid docking results for 1ZLS. For color coding and LIGPLOT conventions, please refer to Figure 2.

mannose residues. As a result, the terminus is docked accurately, deviating from the crystal structure by 1.0 Å in the top pose. An extensive network of hydrogen bonds between this terminus and the protein, present in the crystal structure, can be seen in the top pose (Figure 3b,c). Residues Man 3 and Man 4 demonstrate a greater deviation from the crystal structure. A single hydrogen bond between Tyr 94L and OH(3) of Man 3 is present in the crystal structure. A compensatory hydrogen bond to Tyr 94L is seen in the top pose, with the glycosidic oxygen between Man 3 and Man 4. It must be noted that the conformation of the Man 4 ring is distorted in the crystal structure.[5] As a result, it could be very safely suggested that the rmsd for the whole docked ligand is artificially inflated.

In 1MFC, a tetrasaccharide ligand bound to an antibody with a crater-shaped binding site, the Glide top pose gave an rmsd of 2.05 Å. The correct binding mode was ranked third (1.8 Å) (Figure 4a). In this structure, Abe 3 penetrates the binding site to the greatest extent and is conformationally restricted by the surrounding protein residues. In the top pose, the terminal Abe 3 and Man 2 show the least deviation from the crystal structure, with a combined rmsd for these two rings of 0.5 Å. Also, all experimentally detected hydrogen bonds and two out of three vdW contacts are successfully replicated (Figure 4b,c).

GOLD. GOLD (genetic optimization for ligand docking, Cambridge Crystallographic Data Centre)[36] utilizes a genetic algorithm (GA), which mimics the process of evolution by representing ligand descriptors on a chromosome. Ligand state is encoded by a chromosome, describing its location, orientation, and conformation. Energy-based scoring function GoldScore includes hydrogen bonding, vdW interactions, and ligand internal energy.

Table 3 lists the results of docking the test set with GOLD. In 8 of the 11 cases (73%), the top pose achieved an rmsd of less than 2.0 Å. For 5 of the 11 cases, the top pose was also the best. In three cases, the program failed to predict the experimental binding mode correctly: 1CLY, 1MFE, and 1ZLS. The 1ZLS test case and its challenge to pose prediction is described above. 1CLY and 1MFE will be discussed below in more detail.

For 1CLY the top ranking pose gave a very poor rmsd of 9.76 Å (Figure 5a). GOLD was unsuccessful in generating the experimental ligand binding mode in this case as the best pose (ranked sixth) gave an rmsd of 4.23 Å (Figure 5b). Unlike the case with Glide (see above), the four hexose rings of the tetrasaccharide in the top pose are incorrectly placed in the binding site. No single residue is accurately overlaid with the respective unit of the crystal structure. The non-carbohydrate tail (Non 5) is also located in an alternative area, compared to the crystal structure. As a result of this wrong pose prediction, the interactions are also not reproduced well (data not shown). Also unlike the situation for 1CLY docking with Glide, when the tail was removed and ligand redocked, the results improved only marginally for the top pose and actually worsened for the best pose (Table 3). One possible explanation for the decreased accuracy of GOLD for 1CLY, compared to that of Glide, could be more efficient exploration of conformational space in Glide compared to the GA methods employed by GOLD.

It is interesting to consider GOLD's poor result for 1CLY in comparison to its good result for 1S3K. We have previously observed[3,37] that the most notable difference between these two complexes occurs at the terminal Le^Y-specific fucose residue, which participates in two hydrogen bonds with Asn 28L in 1S3K but not in 1CLY. It is likely that, in the absence of these additional "recognition handles"

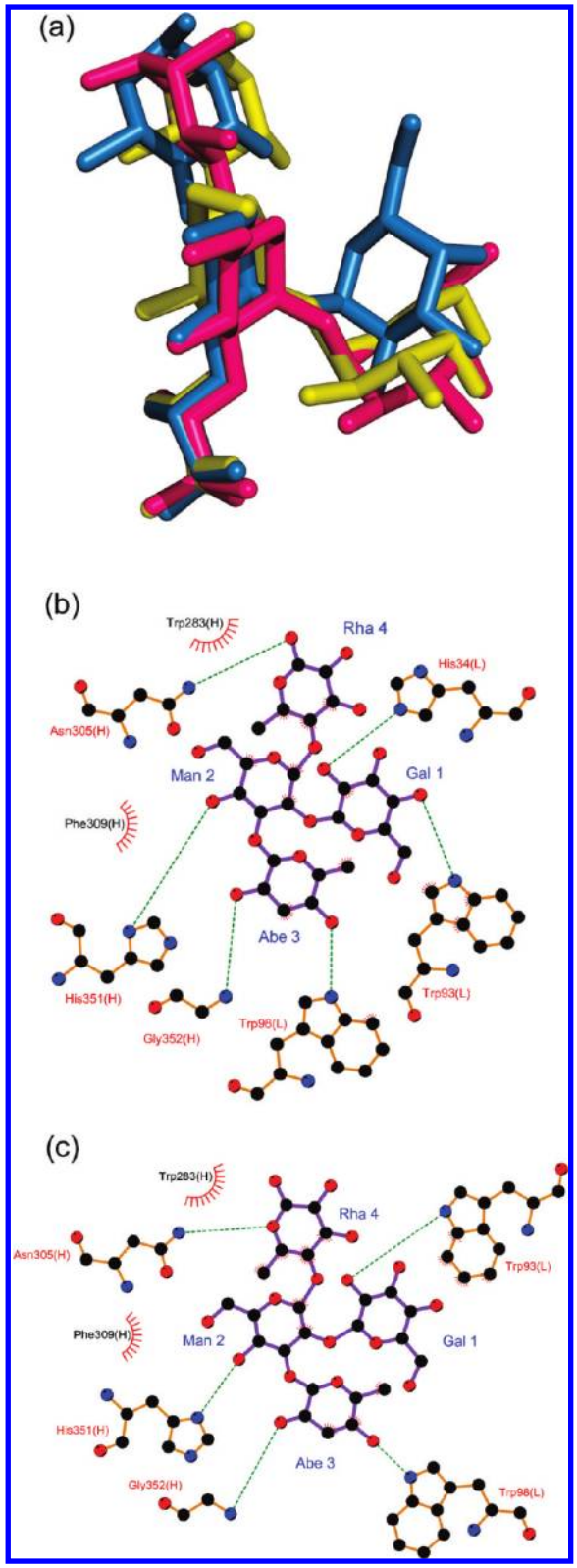

Figure 4. Glide rigid docking results for 1MFC. For color coding and LIGPLOT conventions, please refer to Figure 2.

in 1CLY, GOLD fails to place this residue correctly, which leads to incorrect placement of the whole ligand.

For 1MFE, a trisaccharide ligand bound to a crater-shaped antibody binding site, the top pose gave an rmsd of 6.21 Å. The best pose (ranked sixth) deviated by 1.0 Å (Figure 6a). Man 2 is fairly accurately placed in the top pose, but the placement of the other two residues (Gal 1 and Abe 3) is switched relative to that in the crystal structure, indicating

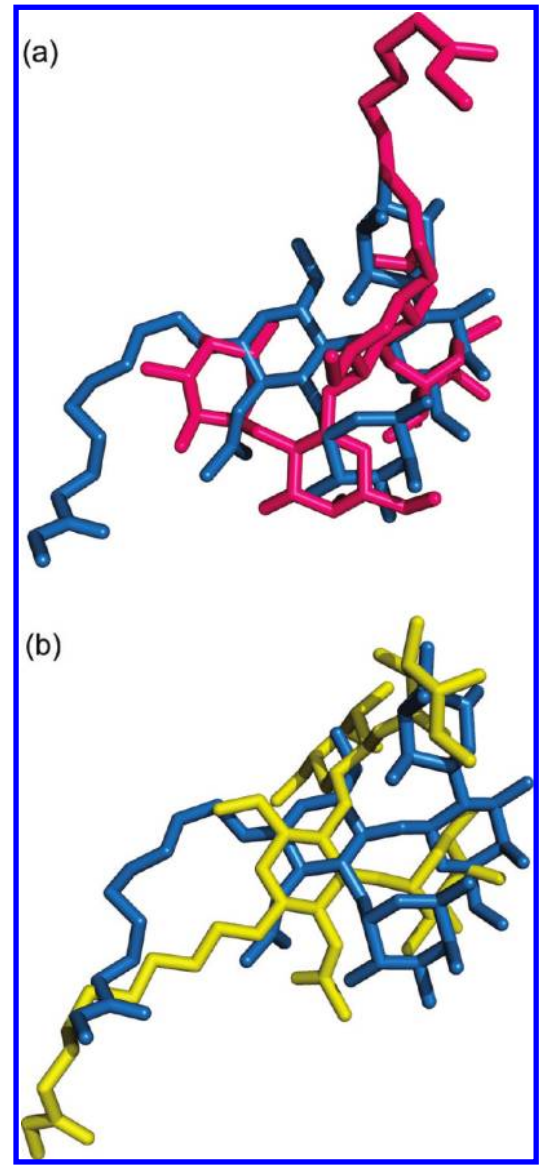

Figure 5. GOLD rigid docking results for 1CLY. (a) Superimposition of the crystal structure (blue) and the top pose (magenta). (b) Superimposition of the crystal structure (blue) and the best pose (yellow).

that GOLD has docked the ligand into the binding site "back to front". The accurate placement of Man 2 in the top pose allows for a hydrogen-bond interaction to His 351H also seen in the crystal structure (Figure 6b,c). In the crystal structure,[33] Abe 3 is totally buried and provides the majority of intermolecular interactions, particularly vdW contacts. However, as hydrogen bonding is factored into the GoldScore algorithm, the high degree of hydrogen bonding in the top docked pose (six contacts compared to four in the crystal structure (Figure 6b,c), may have caused the unreasonably high scoring of this pose compared to the best pose that had fewer of these interactions (data not shown).

AutoDock. Similar to GOLD, AutoDock (Scripps Research Institute)[38] is based on a genetic algorithm. The AutoDock Lamarckian GA is a hybrid of a genetic algorithm with a local adaptive search method. An energy-based scoring function is used to assess the generated solutions.

Only 55% of the top docked poses generated by AutoDock for the Ab−carbohydrate test set had rmsd lower than 2.0 Å (Table 3). Not surprisingly, the worst result was obtained

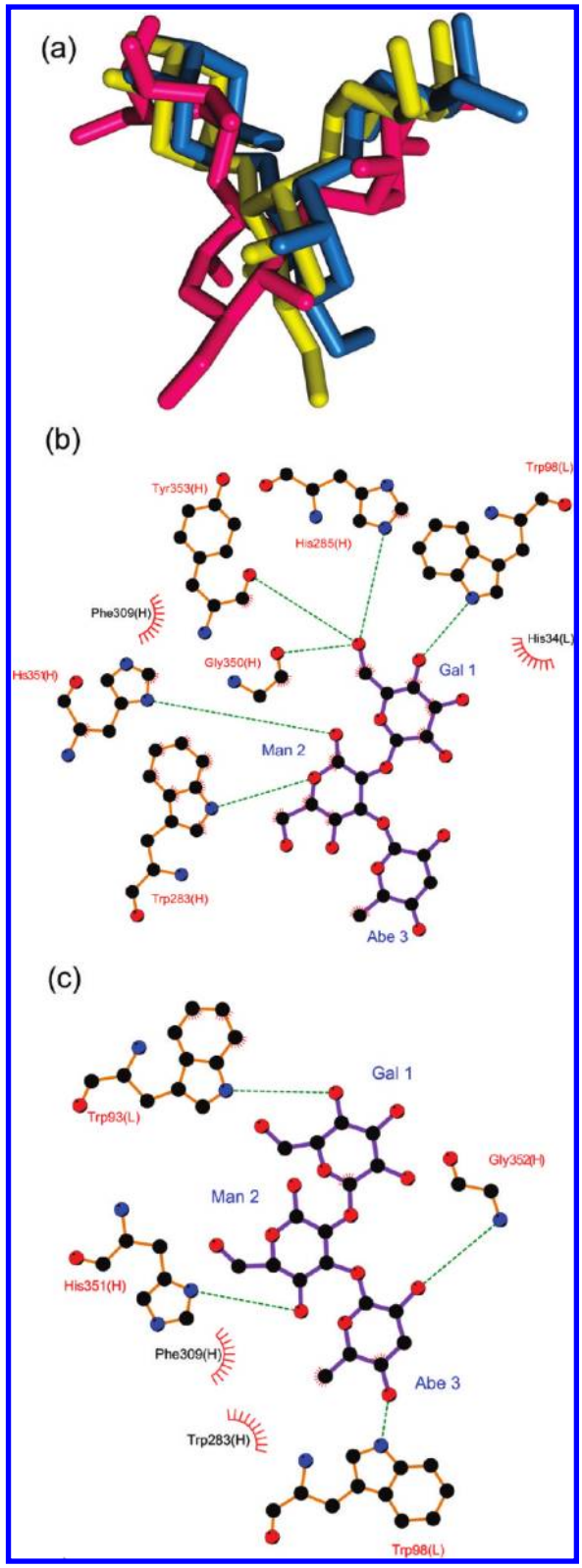

Figure 6. GOLD rigid docking results for 1MFE. For color coding and LIGPLOT conventions, please refer to Figure 2.

for 1CLY (7.20 and 4.30 Å for top and best pose, respectively). Also similar to Glide and GOLD, a suboptimal outcome was obtained for 1ZLS (3.37 and 2.71 Å for top and best pose, respectively). In the 1MFD test case, the top docking pose was very poor (5.67 Å), but the best pose was excellent (0.91 Å) and very highly ranked (second). In the two remaining "incorrect" cases (1UZ8 and 1ZLU), the carbohydrate residues are placed nearly correctly (Figure 7),

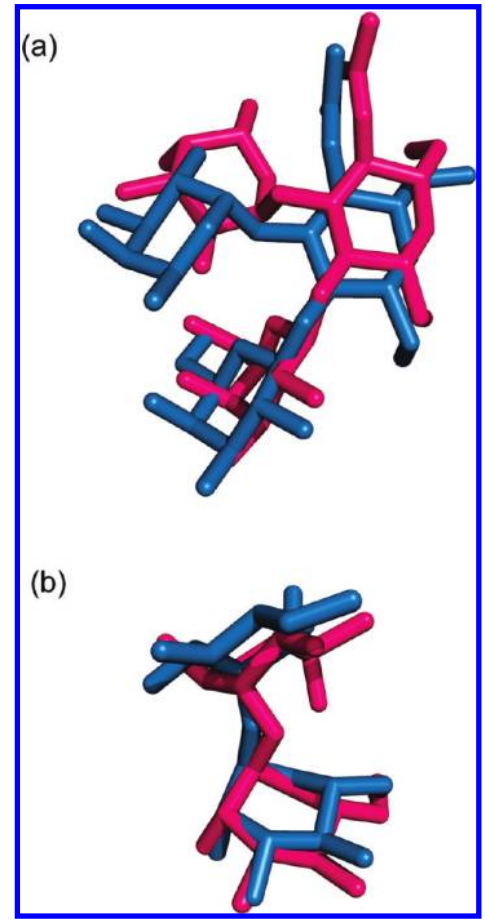

Figure 7. AutoDock rigid docking results for 1UZ8 (a) and 1ZLU (b). For color coding, please refer to Figure 2.

resulting in borderline rmsd (2.54 and 3.08 Å for their top poses, respectively).

FlexX. FlexX (BioSolveIT)[39] features an incremental construction algorithm. Ligands are docked starting with a base fragment and the best solution is selected according to ChemScore, which takes into account hydrogen bonding, ionic, aromatic, and lipophilic interactions, and entropy.

Table 3 lists the docking results for FlexX. In 55% of cases the top pose deviated from the crystal structure by more than 2.0 Å. In the cases of 1S3K and 1ZLU all obtained poses were docked to alternative sites on the antibody Fv surface and thus rmsd values were not recorded. Similar to Glide and GOLD, FlexX produced poor results for 1CLY and 1ZLS. Test cases 1M7D, 1MFC, and 1UZ8 challenged FlexX to reproduce the experimentally observed poses. The failure of FlexX to correctly predict the experimental binding modes in these cases is somewhat surprising. On the basis of our experience with the other three programs, it was reasonable to expect that at least a terminal residue(s) should be placed correctly in the binding site. One possible explanation for the poor docking outcome is that FlexX fails to establish crucial intermolecular interactions. In the test cases, many of the specific hydrogen-bonding interactions in the crystal structure were facilitated by water molecules. Particularly in 1UZ8 and 1MFC, buried water molecules have been credited with increasing the complementarity of complexation.[4,31] In the docking procedures adopted here for all four programs, crystallographic waters were removed prior to docking. It is possible that this produced an insurmountable challenge for FlexX.

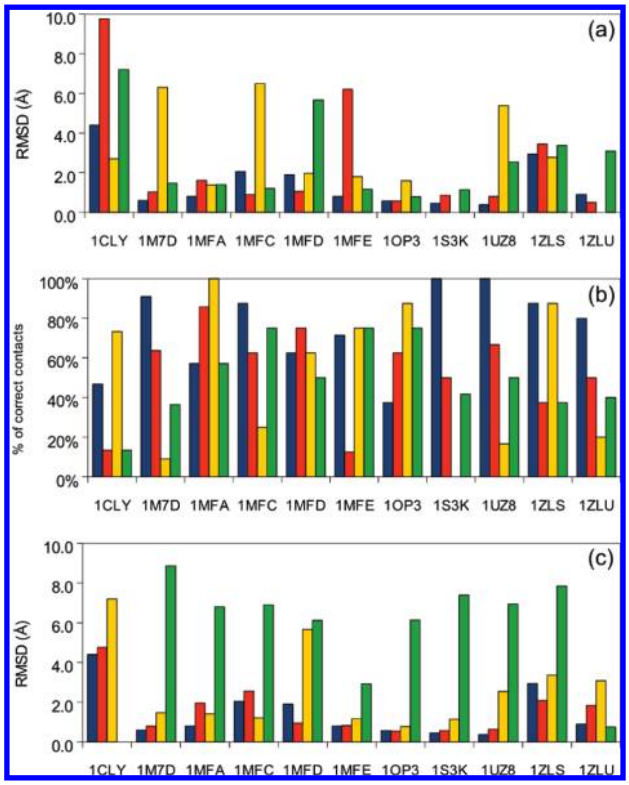

Figure 8. Comparative results for four programs. (a) Rigid receptor docking. The column height corresponds to the rmsd of the top pose. Color coding: Glide, blue; GOLD, red; FlexX, yellow; AutoDock, green. (b) Rigid receptor docking. The column height corresponds to the percentage of the intermolecular interactions from crystal structure reproduced in the top pose. For color coding, see part a. (c) Flexible receptor docking. The column height corresponds to the rmsd of the top pose. Color coding: Glide rigid, blue; Glide flexible, red; AutoDock rigid, yellow; AutoDock flexible, green.

Comparison of Experience with Four Programs. Figure 8a reports the rmsd for top ranked poses for each of the complexes in the test set. The results attest to the accuracy of Glide in docking of carbohydrates to antibodies, with an average rmsd of 1.43 Å for top ranked poses. GOLD also performed well for all but a few cases, averaging a 2.42 Å deviation for top ranked poses. In comparison, AutoDock and FlexX produced less accurate results (average rmsd values of 2.63 and 3.37 Å, respectively).

Figure 8b shows the percentage of interactions displayed in the crystal structure, replicated in the top docked complexes. It is clear that experimentally observed contacts are most consistently reproduced in the Glide docked complexes. These findings correlate directly with the accuracy of the docking results (Figure 8a). In most of the cases, where the top pose deviated from the crystal structure by less than 2.0 Å, more than 80% of intermolecular interactions were replicated.

Flexible Receptor Docking Performance of Two Programs. Flexible receptor docking was performed on the Ab−carbohydrate test set with the Glide/Prime IFD protocol and AutoDock 4.0. Table 4 lists the top and best pose rmsd values and Figure 8c shows the top ranked pose rmsd for each of these two methods. The AutoDock flexible docking could not be applied to 1CLY. In this case, the number of combined rotatable bonds in the ligand and protein residue side chains of the binding site exceeded the limit of 32 set in AutoDock 4.0.

Table 4. Docking Results: Flexible Receptor Docking

	rmsd values (Å)			
	Glide/Prime		AutoDock	
PDB code	top	best	top	best
1CLY	4.76	3.27 (20)	—[a]	—[a]
1M7D	0.79	0.28 (13)	8.86	2.69 (104)
1MFA	2.22	0.38 (8)	6.80	3.41 (143)
1MFC	2.56	1.51 (2)	6.90	1.4 (6)
1MFD	1.04	0.95 (7)	6.12	3.53 (16)
1MFE	0.82	0.64 (3)	2.92	2.68 (58)
1OP3	0.54	0.4 (12)	6.14	1.86 (3)
1S3K	0.56	0.56 (1)	7.39	2.35 (14)
1UZ8	0.64	0.47 (15)	6.94	3.08 (8)
1ZLS	2.09	1.89 (7)	7.84	2.79 (75)
1ZLU	1.84	0.77 (14)	0.75	0.75 (1)
average	1.62		6.07	

[a] The system contained too many rotatable bonds for this procedure.

Glide/Prime showed a high level of prediction accuracy, averaging an rmsd of 1.62 Å for top poses, comparing well to the results obtained from rigid docking with Glide. The accuracy of flexible docking with AutoDock was quite poor, with an average rmsd for the top pose of 6.07 Å.

Unlike Glide rigid docking, the Glide/Prime protocol produced the top pose with the lowest rmsd (best pose) only in one case (1S3K). However, 63% of cases gave a top pose with an rmsd below 2.0 Å. The results for the best poses show that, while in some cases the rmsd is slightly lower than that achieved by rigid docking, the ranking of these best poses is worse. Similarly to rigid docking with Glide, a poor outcome was received for the 1CLY test case. For three test cases (1MFA, 1MFC, and 1ZLS), the rmsd of the top pose was above 2.0 Å, but only marginally. Two of these cases and their docking challenges have been discussed previously.

For flexible receptor docking with AutoDock, the top poses deviated from the crystal structures in most cases by more than 4.0 Å. In 30% of cases the best pose was under 2.0 Å; however, these poses were very poorly ranked. The only case where AutoDock produced a good result is 1ZLU, where a disaccharide ligand is bound to a crater-shaped antibody binding site.

Cross-Docking. Complexes 1MFC, 1MFD, and 1MFE were used for cross-docking with Glide as each contains the same antibody, Se155-4. In each case the antibody was cocrystallized with carbohydrate ligands consisting of the same anchoring residues (Gal 1 and Abe 2) occupying the binding site and either one or two additional hexose rings extending into the solvent-exposed region at the mouth of the binding site. The ligand from each complex was docked into the antibody binding sites of the other two complexes (i.e., six cross-dockings were performed).

Table 5 shows the rmsd values of top docked poses in both cognate rigid docking into the original complex (shown in bold) and cross-docking flexible simulations. In all cases, flexible receptor docking with the Glide/Prime protocol achieved very accurate rmsd values for top ranked poses, relative to the corresponding crystal structures. The intermolecular interactions were the same as obtained with the rigid docking results (data not shown).

Table 5. Glide/Prime Flexible Cross-Docking Results

	rmsd values for top ranked poses (Å)		
	1MFC ligand[a]	1MFD ligand[b]	1MFE ligand[b]
1MFC antibody	**2.05**[c]	1.05	1.02
1MFD antibody	2.7	**1.89**	1.7
1MFE antibody	1.75	1.64	**0.8**

[a] Tetrasaccharide. [b] Trisaccharide. [c] The rmsd values of top docked poses in rigid docking into the original complex are shown in bold.

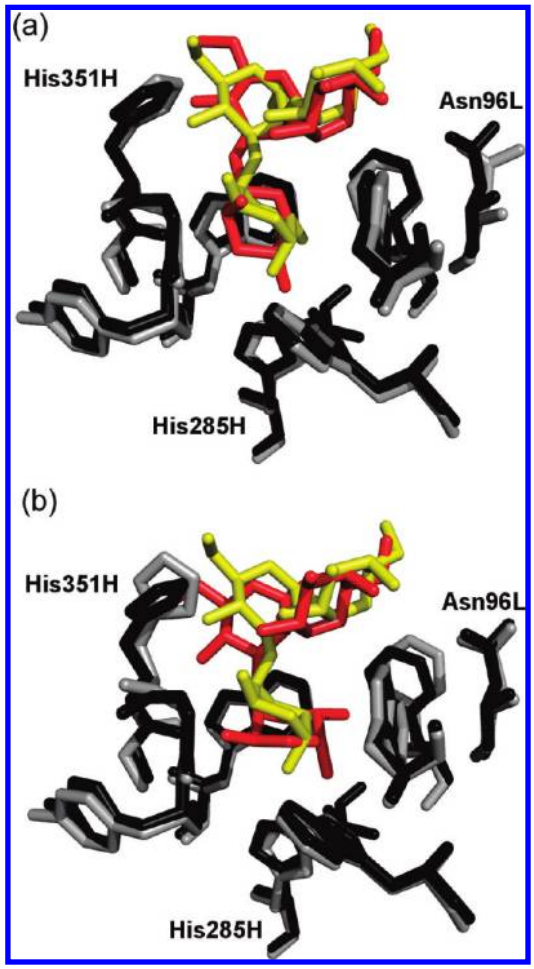

Figure 9. Cross-docking results for 1MFD. (a) Superimposition of 1MFD crystal structure pose (yellow) and binding site protein residues (black) with docked top pose (red) and binding site protein residues (gray) from docking into the 1MFC antibody. (b) Superimposition of 1MFD crystal structure pose (yellow) and binding site protein residues (black) with docked top pose (red) and binding site protein residues (gray) from docking into the 1MFE antibody.

Results from cross-docking of the 1MFD trisaccharide ligand are used as an example to demonstrate the effects of this flexible docking method (Figure 9). Cross-docking the 1MFD ligand into the 1MFC binding site resulted in a structure with only slight shifting of protein positions and a top ligand pose with an rmsd of 1.05 Å (Figure 9a). In comparison, cross-docking of the same ligand into the 1MFE binding site shows a large shifting of the protein residues His 351H, His 285H, and Asn 96L and a top ligand pose with an rmsd of 1.64 Å (Figure 9b). Thus, while the rmsd values are close, the movements of the protein residues in each antibody binding site are quite different. This suggests a greater level of ligand-induced conformational change in the binding site of the antibody optimized for the 1MFE ligand when cross-docked with the 1MFD ligand.

DISCUSSION

Comparative Analysis of Rigid Receptor Docking Outcomes. *Pose Prediction.* The first aim of this study was to identify which program out of four could most reliably dock carbohydrates to antibodies for application to situations where a crystal structure is not available. The results for rigid docking of the Ab−carbohydrate complexes used in this study indicate that, while different programs succeeded and/or failed in different test cases, Glide was more successful than the other three programs overall. GOLD produced results that were only slightly worse than Glide overall, but it displayed two specific points of weakness. Namely, those in the GA (1CLY test case) and in the potential overestimation of hydrogen bonds by GoldScore (1MFE test case). Both points may not be as obvious when docking small organic molecules but become more pronounced for highly flexible molecules with a multitude of hydrogen bond donor/acceptor atoms, such as carbohydrates.

Effect of Binding Site Topography and Carbohydrate Length on Docking. The greatest influence on the accuracy of the top pose, as defined by rmsd and intermolecular interactions, appears to be the binding site topography combined with the degree of ligand protrusion from the binding site (Table 2). The latter is, in itself, a consequence of the binding site topography and ligand size. Since Glide performed best in this comparison, we have limited the discussion of these factors to Glide docking outcomes.

Canyon-shaped binding sites (1M7D, 1OP3, and 1UZ8) delivered the best results due to the degree by which the ligand is surrounded by protein residues in this topography class. The carbohydrates mostly bind parallel to the binding site floor (Figure 10a), each of the ligand rings tightly surrounded by protein residues contained in the floor and walls of the site, causing a high degree of conformational restriction. Thus, it could be cautiously suggested that canyon-shaped sites lend themselves best to more reliable docking. With respect to carbohydrate ligands, shorter ligands are more likely to reproduce experimental structures when docked.

Valley-shaped binding sites are represented by 1CLY and 1S3K in our test set. The 1CLY antibody (Figure 10b) binds a tetrasaccharide with a large flexible noncarbohydrate tail residue. Located at the "open" end of the binding site, this flexible tail deviates most from the crystal structure in docking simulations, compared to the four carbohydrate residues (Figure 2a). In contrast, the 1S3K antibody, which also binds to a tetrasaccharide ligand, displays a docking outcome of 0.45 Å. Thus, valley-shaped binding sites warrant further investigations (once relevant crystal structures become available) with respect to their effect on docking accuracy.

The crater-shaped antibody binding sites cause the carbohydrate to bind almost perpendicular to the cavity floor (Figure 10c), with the residues anchoring the ligand in the cavity completely buried and surrounded by protein residues. Those at the opposite end of the ligand protrude into the bulk solvent at the opening of the binding site and have a

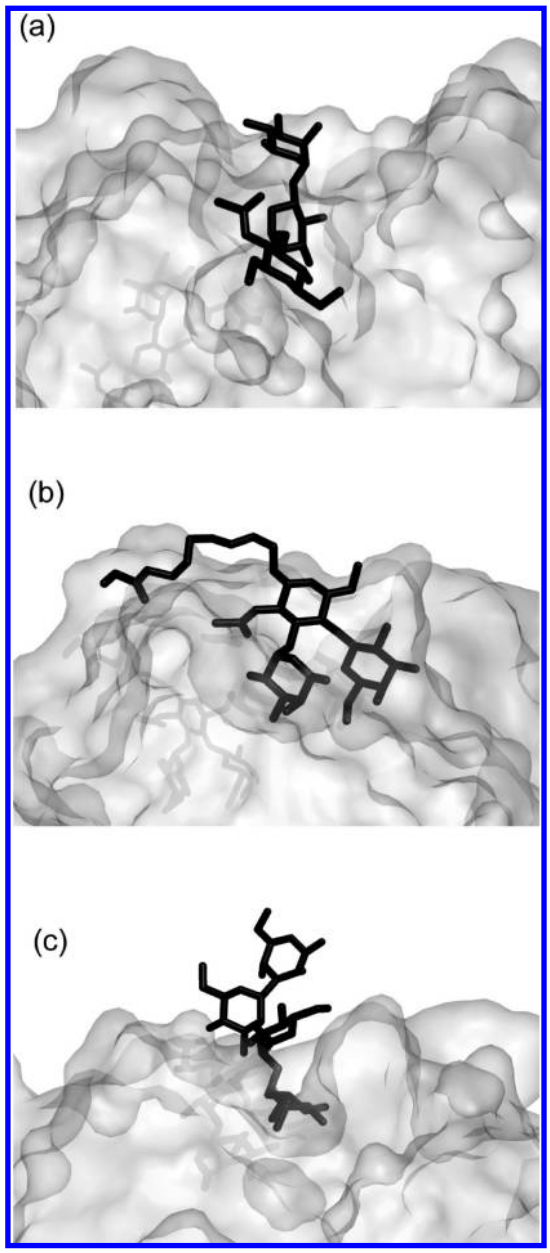

Figure 10. Ligand accommodation tactics in binding sites of different topography: (a) Canyon-shaped binding sites (e.g., 1M7D). (b) Valley-shaped binding sites (e.g., 1CLY). (c) Crater-shaped binding sites (e.g., 1ZLS).

much greater degree of conformational freedom. In all cases of crater-shaped binding sites, docking pose deviation from crystal structures is largely due to these protruding terminal residues. In the cases of 1ZLS and 1MFC, which contained large carbohydrate ligands with a two residue protrusion from the binding site, Glide produced top poses deviating by more than 2.0 Å. When the ligand was small and did not protrude from the binding site at all (1ZLU), the rmsd was just 0.9 Å. Thus, for most crater-shaped binding sites, the smaller the carbohydrate and, consequently, the degree of protrusion, the better the docking accuracy.

Intermolecular Interactions. Interestingly, in all cases where docking results (expressed in terms of rmsd) were detrimentally affected by ligand protrusion from the binding site (e.g., 1CLY, 1ZLS, and 1MFC), top docking poses reproduced all or almost all hydrogen-bonding interactions and the majority of vdW contacts observed experimentally. These cases demonstrate particularly well that the de facto measure of pose prediction — rmsd — does not do justice to the intermolecular interactions in protein−ligand complexes. As our results show, these two measures correlate well in many, but not all, cases. Despite the ease of calculation and, therefore, the practical appeal of rmsd, using it as the only measure of docking accuracy is a flawed and inconsistent approach. Kroemer et al.[40] were the first to investigate the merits of interactions as pose prediction measure and have developed the interactions-based accuracy classification (IBAC) to group docked poses into "correct, nearly correct, and incorrect" based on the intermolecular interactions observed in docked poses compared to crystal structures. Vaque et al.[13] pointed out that generally intermolecular interactions are weak, directional, and hard to predict. However, for carbohydrate ligands the IBAC or IBAC-like approach would be appropriate due exactly to the directional preference of their hydrogen bonding (compared to small organic molecules) and to the CH-π stacking. The drawback of IBAC is the need for visual inspection and lack of automation.[27,41] Recently, we have developed a computational procedure that makes use of intermolecular interactions in multiple docked poses and generates a binding site map.[42] We have validated this procedure against the same set of Ab−carbohydrate crystal structures utilized here and demonstrated its utility both as an automated measure of pose assessment and as a predictive tool of likely antibody residues required for carbohydrate recognition.

Summary of Flexible Docking. In many popular docking methods, the ligand is treated as flexible, but the protein conformation is kept rigid. This relies on the lock-and-key hypothesis for protein ligand binding.[43,44] However, it is now widely accepted that ligand binding is not a static event but a dynamic process, in which both the ligand and protein may undergo conformational change(s).[45,46] Antibody−antigen systems form an important class of protein complexes exhibiting induced-fit (conformational changes upon binding).[47] Structural and kinetic evidence suggests that antibodies can exhibit conformational diversity by adopting multiple free structures capable of binding unrelated antigens.[48] In docking, incorporating protein flexibility exponentially expands the potential search space and quickly becomes impractical. However, algorithms that allow for partial treatment of protein flexibility have now been implemented in several docking programs. In this study, we have evaluated two of such algorithms: Glide/Prime Induced Fit Docking protocol and AutoDock 4.0.

The test cases used for cognate docking are the biased protein complexes, in which the active site is already optimized for the ligand being docked. During flexible docking, protein conformations are altered from those in the crystal structures, to which the comparisons are being made. Therefore, in pose prediction exercises, the cognate rigid receptor docking is likely to outperform the flexible approach. Thus, we have carried out the flexible receptor docking to establish whether it could be relied upon to produce accurate results. As expected, the results of flexible docking were less accurate than those from rigid docking. Overall, Glide/Prime IFD has performed significantly better than AutoDock 4.0 in the flexible receptor docking exercise. The poor performance of AutoDock was rather puzzling. Previously, AutoDock has been widely used for carbohydrate docking and

structurally and energetically validated for a range of systems.[23,49−57] However, most of the previous studies and all of the reported carbohydrate docking validations have been performed with AutoDock 3.x, which does not allow protein flexibility. AutoDock 4.0 is a fairly new development. To the best of our knowledge, simulations described here present a first attempt at flexible receptor carbohydrate docking with AutoDock 4.0 and its first structural evaluation for this ligand system.

While crystal structures were used as validation test cases in this study, we had to consider an application of Ab−carbohydrate docking to a real scenario, where a complex crystal structure is not available. Then, to produce a complex one needs to dock a ligand into the native protein (crystal structure or an in silico model) or into the protein complexed to another ligand. A large body of literature indicates that flexible treatment is preferable in such cases. However, since the input binding site conformation may be far removed from that in the complex being modeled, an appropriate validation of such approach is cross-docking, where a ligand is docked into a protein active site, which is free or bound to a different ligand. Thus, the final aim of this study was to evaluate docking for the ability to accurately reproduce experimental binding modes by docking carbohydrates into antibodies optimized for different carbohydrates. Only the Glide/Prime protocol was used for this evaluation, since it performed best in both rigid and flexible cognate trials. The results of cross-docking with the Glide/Prime IFD were promising. They indicated that the flexible receptor docking implemented in this protocol is capable of producing correct binding modes, using the binding site not experimentally optimized for the ligand in question. As the number of Ab−carbohydrate complexes in the PDB increases, cross-docking using antibodies with a more diverse set of ligands could expand this area of study.

CONCLUSIONS

The reliability of four different programs in performing molecular docking of carbohydrates to antibodies has been evaluated and compared. Glide stood out as the most accurate program for rigidly docking carbohydrates to antibodies. GOLD and AutoDock had several problems, and FlexX performed poorly. The Glide/Prime flexible receptor docking protocol produced results with a level of accuracy comparable to that of the cognate rigid receptor docking. The accuracy of the Glide/Prime protocol in both cognate and cross-docking offers a promising method for the study of the dynamic nature of the Ab−carbohydrate recognition process. Our findings are in agreement with the experience of others. While a perfect docking program does not currently exist, Glide has ranked consistently high in docking studies,[21,58,59] with some listing its treatment of flexible ligands as an important factor.[13]

The diverse antibody binding sites and carbohydrate ligand lengths presented challenges for accurate docking. Large flexible noncarbohydrate residues and crater-shaped binding sites, accommodating large carbohydrate ligands, were found to be the most limiting features. The best results were produced for antibodies with canyon-shaped binding sites, which offered good results for both small and large carbohydrate ligands.

Our results illustrated the importance of intermolecular interactions as a measure of pose prediction. We have demonstrated that the ability of docked poses to display critical intermolecular interactions is a significant criterion and should always be used in conjunction with rmsd.

ACKNOWLEDGMENT

M.A. is a recipient of an Australian Postgraduate Award (APA). P.A.R. is a recipient of an R. Douglas Wright Career Development Award (ID365209) from the NHMRC.

REFERENCES AND NOTES

(1) Nores, G. A.; Lardone, R. D.; Comin, R.; Alaniz, M. E.; Moyano, A. L.; Irazoqui, F. J. Anti-GM1 antibodies as a model of the immune response to self-glycans. *Biochim. Biophys. Acta* **2008**, *1780*, 538–545.
(2) Jeffrey, P. D.; Bajorath, J.; Chang, C. Y.; Yelton, D.; Hellstrom, I.; Hellstrom, K. E.; Sheriff, S. The X-ray structure of an anti-tumour antibody in complex with antigen. *Nat. Struct. Biol.* **1995**, *2*, 466–471.
(3) Ramsland, P. A.; Farrugia, W.; Bradford, T. M.; Mark Hogarth, P.; Scott, A. M. Structural convergence of antibody binding of carbohydrate determinants in Lewis Y tumor antigens. *J. Mol. Biol.* **2004**, *340*, 809–818.
(4) van Roon, A. M.; Pannu, N. S.; de Vrind, J. P.; van der Marel, G. A.; van Boom, J. H.; Hokke, C. H.; Deelder, A. M.; Abrahams, J. P. Structure of an anti-Lewis X Fab fragment in complex with its Lewis X antigen. *Structure* **2004**, *12*, 1227–1236.
(5) Calarese, D. A.; Lee, H. K.; Huang, C. Y.; Best, M. D.; Astronomo, R. D.; Stanfield, R. L.; Katinger, H.; Burton, D. R.; Wong, C. H.; Wilson, I. A. Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13372–13377.
(6) Calarese, D. A.; Scanlan, C. N.; Zwick, M. B.; Deechongkit, S.; Mimura, Y.; Kunert, R.; Zhu, P.; Wormald, M. R.; Stanfield, R. L.; Roux, K. H.; Kelly, J. W.; Rudd, P. M.; Dwek, R. A.; Katinger, H.; Burton, D. R.; Wilson, I. A. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* **2003**, *300*, 2065–2071.
(7) Albert, H.; Collin, M.; Dudziak, D.; Ravetch, J. V.; Nimmerjahn, F. In vivo enzymatic modulation of IgG glycosylation inhibits autoimmune disease in an IgG subclass-dependent manner. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 15005–15009.
(8) Galili, U. The alpha-gal epitope and the anti-Gal antibody in xenotransplantation and in cancer immunotherapy. *Immunol. Cell Biol.* **2005**, *83*, 674–686.
(9) Sandrin, M. S.; Vaughan, H. A.; Dabkowski, P. L.; McKenzie, I. F. Anti-pig IgM antibodies in human serum react predominantly with Gal(alpha 1−3)Gal epitopes. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11391–11395.
(10) Sandrin, M. S.; McKenzie, I. F. Gal alpha (1,3)Gal, the major xenoantigen(s) recognised in pigs by human natural antibodies. *Immunol. Rev.* **1994**, *141*, 169–190.
(11) Chung, C. H.; Mirakhur, B.; Chan, E.; Le, Q. T.; Berlin, J.; Morse, M.; Murphy, B. A.; Satinover, S. M.; Hosen, J.; Mauro, D.; Slebos, R. J.; Zhou, Q.; Gold, D.; Hatley, T.; Hicklin, D. J.; Platts-Mills, T. A. Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N. Engl. J. Med* **2008**, *358*, 1109–1117.
(12) Arnold, D. F.; Misbah, S. A. Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N. Engl. J. Med.* **2008**, *358*, 2735 (author reply 2735−2736).
(13) Vaque, M.; Ardrevol, A.; Blade, C.; Salvado, M. J.; Blay, M.; Fernandez-Larrea, J.; Arola, L.; Pujadas, G. Protein−ligand docking: A review of recent advances and future perspectives. *Curr. Pharm. Anal.* **2008**, *4*, 1–19.
(14) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
(15) Paula, S.; Monson, N.; Ball, W. J., Jr. Molecular modeling of cardiac glycoside binding by the human sequence monoclonal antibody 1B3. *Proteins* **2005**, *60*, 382–391.
(16) Corbeil, C. R.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 3. Impact of input ligand conformation, protein flexibility, and water molecules on the accuracy of docking programs. *J. Chem. Inf. Model.* **2009**, *49*, 997–1009.
(17) Cheng, T.; Li, X.; Li, Y.; Liu, Z. C.; Wang, R. Comparative assessment of scoring functions on a diverse test set. *J. Chem. Inf. Model.* **2009**, *49*, 1079–1093.

(18) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of several molecular docking programs: Pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2009**, *49*, 1455–1474.
(19) Mooij, W. T.; Verdonk, M. L. General and targeted statistical potentials for protein−ligand interactions. *Proteins* **2005**, *61*, 272–287.
(20) Laughrey, Z. R.; Kiehna, S. E.; Riemen, A. J.; Waters, M. L. Carbohydrate−π interactions: What are they worth. *J. Am. Chem. Soc.* **2008**, *130*, 14625–14633.
(21) Nurisso, A.; Kozmon, S.; Imberty, A. Comparison of docking methods for carbohydrate binding in calcium-dependent lectins and prediction of the carbohydrate binding mode to sea cucumber lectin CEL-III. *Mol. Simul.* **2008**, *34*, 469–479.
(22) Alexacou, K. M.; Hayes, J. M.; Tiraidis, C.; Zographos, S. E.; Leonidas, D. D.; Chrysina, E. D.; Archontis, G.; Oikonomakos, N. G.; Paul, J. V.; Varghese, B.; Loganathan, D. Crystallographic and computational studies on 4-phenyl-*N*-(beta-D-glucopyranosyl)-1*H*-1,2,3-triazole-1-acetamide, an inhibitor of glycogen phosphorylase: Comparison with alpha-D-glucose, *N*-acetyl-beta-D-glucopyranosylamine and *N*-benzoyl-*N*′-beta-D-glucopyranosyl urea binding. *Proteins* **2008**, *71*, 1307–1323.
(23) Rockey, W. M.; Laederach, A.; Reilly, P. J. Automated docking of alpha-(1→4)- and alpha-(1→6)-linked glucosyl trisaccharides and maltopentaose into the soybean beta-amylase active site. *Proteins* **2000**, *40*, 299–309.
(24) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534–553.
(25) Lee, M.; Lloyd, P.; Zhang, X.; Schallhorn, J. M.; Sugimoto, K.; Leach, A. G.; Sapiro, G.; Houk, K. N. Shapes of antibody binding sites: Qualitative and quantitative analyses based on a geomorphic classification scheme. *J. Org. Chem.* **2006**, *71*, 5082–5092.
(26) Zhao, Y.; Sanner, M. F. Protein-ligand docking with multiple flexible side chains. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 673–679.
(27) Hawkins, P. C. D.; Warren, G. L.; Skillman, A. G.; Nicholls, A. How to do an evaluation: Pitfalls and traps. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 179–190.
(28) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein−ligand interactions. *Protein Eng.* **1995**, *8*, 127–134.
(29) Vyas, N. K.; Vyas, M. N.; Chervenak, M. C.; Johnson, M. A.; Pinto, B. M.; Bundle, D. R.; Quiocho, F. A. Molecular recognition of oligosaccharide epitopes by a monoclonal Fab specific for *Shigella flexneri* Y lipopolysaccharide: X-ray structures and thermodynamics. *Biochemistry* **2002**, *41*, 13575–13586.
(30) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. Structure of a single-chain antibody variable domain (Fv) fragment complexed with a carbohydrate antigen at 1.7-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6423–6427.
(31) Cygler, M.; Wu, S.; Zdanov, A.; Bundle, D. R.; Rose, D. R. Recognition of a carbohydrate antigenic determinant of *Salmonella* by an antibody. *Biochem. Soc. Trans.* **1993**, *21*, 437–441.
(32) Bundle, D. R.; Baumann, H.; Brisson, J. R.; Gagne, S. M.; Zdanov, A.; Cygler, M. Solution structure of a trisaccharide−antibody complex: Comparison of NMR measurements with a crystal structure. *Biochemistry* **1994**, *33*, 5183–5192.
(33) Cygler, M.; Rose, D. R.; Bundle, D. R. Recognition of a cell-surface oligosaccharide of pathogenic *Salmonella* by an antibody Fab fragment. *Science* **1991**, *253*, 442–445.
(34) Webster, D. M.; Rees, A. R. Molecular modeling of antibody-combining sites. *Methods Mol. Biol.* **1995**, *51*, 17–49.
(35) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(36) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(37) Yuriev, E.; Farrugia, W.; Scott, A. M.; Ramsland, P. A. Three-dimensional structures of carbohydrate determinants of Lewis system antigens: Implications for effective antibody targeting of cancer. *Immunol. Cell Biol.* **2005**, *83*, 709–717.
(38) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(39) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
(40) Kroemer, R. T.; Vulpetti, A.; McDonald, J. J.; Rohrer, D. C.; Trosset, J. Y.; Giordanetto, F.; Cotesta, S.; McMartin, C.; Kihlen, M.; Stouten, P. F. Assessment of docking poses: Interactions-based accuracy classification (IBAC) versus crystal structure deviations. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 871–881.
(41) Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection−What can we learn from earlier mistakes. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 213–228.
(42) Agostino, M.; Sandrin, M. S.; Thompson, P. E.; Yuriev, E.; Ramsland, P. A. In silico analysis of antibody−carbohydrate interactions and its application to xenoreactive antibodies. *Mol. Immunol.* **2009**, *47*, 233−246.
(43) Fischer, E. Einfluss der konfiguration auf die wirkung der enzyme. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993.
(44) Ehrlich, P., Partial cell functions, Nobel Lecture, December 11, 1908. In *Nobel Lectures, Physiology or Medicine 1901−1921*; Elsevier: Amsterdam, 1967; pp 304−320.
(45) Koshland, D. E., Jr. The key−lock theory and the induced fit theory. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 2375–2378.
(46) Koshland, D. E., Jr. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.
(47) Herron, J. N.; He, X. M.; Ballard, D. W.; Blier, P. R.; Pace, P. E.; Bothwell, A. L.; Voss, E. W., Jr.; Edmundson, A. B. An autoantibody to single-stranded DNA: Comparison of the three-dimensional structures of the unliganded Fab and a deoxynucleotide−Fab complex. *Proteins* **1991**, *11*, 159–175.
(48) James, L. C.; Roversi, P.; Tawfik, D. S. Antibody multispecificity mediated by conformational diversity. *Science* **2003**, *299*, 1362–1367.
(49) Allen, M. J.; Laederach, A.; Reilly, P. J.; Mason, R. J.; Voelker, D. R. Arg343 in human surfactant protein D governs discrimination between glucose and *N*-acetylglucosamine ligands. *Glycobiology* **2004**, *14*, 693–700.
(50) Cantu, D.; Nerinckx, W.; Reilly, P. J. Theory and computation show that Asp463 is the catalytic proton donor in human endoplasmic reticulum alpha-(1→2)-mannosidase I. *Carbohydr. Res.* **2008**, *343*, 2235–2242.
(51) Mertz, B.; Hill, A. D.; Mulakala, C.; Reilly, P. J. Automated docking to explore subsite binding by glycoside hydrolase family 6 cellobiohydrolases and endoglucanases. *Biopolymers* **2007**, *87*, 249–260.
(52) Mulakala, C.; Nerinckx, W.; Reilly, P. J. Docking studies on glycoside hydrolase family 47 endoplasmic reticulum alpha-(1→2)-mannosidase I to elucidate the pathway to the substrate transition state. *Carbohydr. Res.* **2006**, *341*, 2233–2245.
(53) Mulakala, C.; Reilly, P. J. Understanding protein structure−function relationships in family 47 alpha-1,2-mannosidases through computational docking of ligands. *Proteins* **2002**, *49*, 125–134.
(54) Allen, M. J.; Laederach, A.; Reilly, P. J.; Mason, R. J. Polysaccharide recognition by surfactant protein D: Novel interactions of a C-type lectin with nonterminal glucosyl residues. *Biochemistry* **2001**, *40*, 7789–7798.
(55) Mulakala, C.; Reilly, P. J. Force calculations in automated docking: Enzyme−substrate interactions in *Fusarium oxysporum* Cel7B. *Proteins* **2005**, *61*, 590–596.
(56) Laederach, A.; Reilly, P. J. Specific empirical free energy function for automated docking of carbohydrates to proteins. *J. Comput. Chem.* **2003**, *24*, 1748–1757.
(57) Hill, A. D.; Reilly, P. J. A Gibbs free energy correlation for automated docking of carbohydrates. *J. Comput. Chem.* **2008**, *29*, 1131–1141.
(58) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225–242.
(59) Englebienne, P.; Fiaux, H.; Kuntz, D. A.; Corbeil, C. R.; Gerber-Lemaire, S.; Rose, D. R.; Moitessier, N. Evaluation of docking programs for predicting binding of Golgi alpha-mannosidase II inhibitors: A comparison with crystallography. *Proteins* **2007**, *69*, 160–176.

CI900388A