J Mol Model (2015) 21: 164
DOI 10.1007/s00894-015-2713-2

ORIGINAL PAPER

Comparative evaluation of several docking tools for docking small molecule ligands to DC-SIGN

Gregor Jug[1] · **Marko Anderluh**[1] · **Tihomir Tomašič**[1]

Received: 6 February 2015 / Accepted: 19 May 2015 / Published online: 4 June 2015

Abstract Five docking tools, namely AutoDock, FRED, CDOCKER, FlexX and GOLD, have been critically examined, with the aim of selecting those most appropriate for use as docking tools for docking molecules to the lectin dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN). This lectin has been selected for its rather non-druggable binding site, which enables complex interactions that guide the binding of the core monosaccharide. Since optimal orientation is crucial for forming coordination bonds, it was important to assess whether the selected docking tools could reproduce the optimal binding conformation for several oligosaccharides that are known to bind DC-SIGN. Our results show that even widely used docking programs have certain limitations when faced with a rather shallow and featureless binding site, as is the case of DC-SIGN. The FRED docking software (OpenEye Scientific Software, Inc.) was found to score as the best tool for docking ligands to DC-SIGN. The performance of FRED was further assessed on another lectin, Langerin. We have demonstrated that this validated docking protocol could be used for docking to other lectins similar to DC-SIGN.

Keywords Antagonists · Anti-infectives · DC-SIGN · Docking · Docking validation · HIV-1

Electronic supplementary material The online version of this article (doi:10.1007/s00894-015-2713-2) contains supplementary material, which is available to authorized users.

✉ Tihomir Tomašič
tihomir.tomasic@ffa.uni-lj.si

[1] Faculty of Pharmacy, University of Ljubljana, Aškerčeva 7, 1000 Ljubljana, Slovenia

Introduction

Developments of computer science over the last few decades have enabled the use of computational methods for drug discovery and drug design to become one of the most important and, now, routinely used tools in drug discovery. Of these, protein-ligand docking is one of the most widely used methods in research on protein-ligand interactions [1]. A prerequisite for the success of this method in predicting proper protein-ligand interactions, detecting and ranking best ligand poses, assessing binding free energies, etc. is the appropriate quality of the input data. The minimum requirement for the latter includes accurate structural data for the ligand(s) and the protein (i.e., target) of interest. The number of protein-ligand structures solved using X-ray and NMR techniques is increasing and is available at the RCSB Protein Data Bank (PDB), thus providing the input for docking studies and consequently leading to their development [2, 3].

Docking programs have been developing over more than 30 years [4]. There are now more than 50 different docking software packages available to researchers [5]. They differ mainly in the algorithm and scoring functions used for predicting and ranking ligands' docking poses and, consequently, not every docking program can be used for any particular docking study [6]. All docking programs require 3D structures of the target protein and of the ligand. Since some of the latest docking tools are capable of generating various 3D conformations of the ligand, the initial 3D structure of the ligand must be provided although, in some cases, SMILES string notation can be sufficient for the docking program to generate the 3D structure of the ligand. If conformation generation is not enabled by the docking program itself, a set of different ligand conformations have to be generated, using appropriate software [5].

Alongside the advantages that docking tools possess, their deficiencies also have to be kept in mind. In this paper we consider two of these that we believe are important and should be focused on in future improvement of docking tools: (i) the protein is (in most cases) kept rigid during the docking study and (ii) the contribution of desolvation is not considered explicitly or is ignored, which is expected to be of paramount importance in docking studies of highly solvated ligands like (oligo)saccharides [6, 7].

Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) is a type II transmembrane C-type lectin located specifically on dendritic cells (DCs). It acts as an adhesion molecule and enables certain functions of DCs, including migration, pathogen recognition, internalization, processing and binding to T cells. HIV-1 has been reported to enter DCs by being bound to DC-SIGN, escaping the normal lytic pathway in the endosomes of DCs and thus avoiding the immunity defense system. DC-SIGN is thus a receptor of interest in the design of anti-infectives that would inhibit DC-SIGN-pathogen interaction by blocking the very first step in pathogen infection. The main scope of our research is dedicated to the design of new DC-SIGN carbohydrate recognition domain (CRD) antagonists that could act as potential anti-infectives [8–10]. Docking programs can be a great help in the design of new DC-SIGN antagonists, not only helping to evaluate the binding of the newly designed molecules to the target, but also enable to select, and hence synthesize, molecules for which the predicted binding energy (i.e., K_d in nM range) is estimated to be favorable.

DC-SIGN is a C-type lectin that binds glycosylated proteins and oligosaccharides [10]. The main determinant of its binding site is the hexacoordinated Ca^{2+} ion (Fig. 1). In the unbound state, it forms two complex bonds with water molecules, and similar interactions are formed with the hydroxyls of the bound monosaccharide at positions 3 and 4. Since all the DC-SIGN ligands found in crystal structures bind tightly to DC-SIGN by forming complex bonds with the Ca^{2+} ion, we sought a docking protocol that would enable the correct prediction of these bonds.

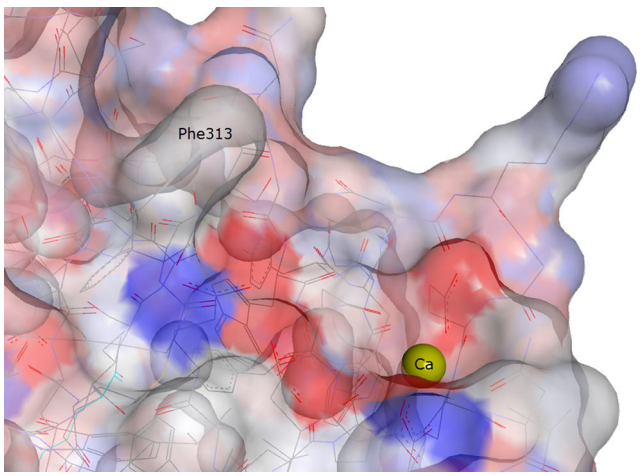

Fig. 1 The DC-SIGN binding site from a crystal structure (PDB code 2IT6). The protein is represented as an atom charge surface. The Ca^{2+} ion is represented as a yellow sphere. Phe313 is also exposed. It acts as a steric hindrance, affecting the binding of linear oligosaccharides and constitutes the hydrophobic groove of the receptor, which could be exploited to improve affinity of the novel DC-SIGN antagonists [8, 11]

Docking programs in general differ in a number of important features [7], among which prediction of complex binding with metal ions is the most important feature for the study of C-type lectins like DC-SIGN, so a validation study of five different docking tools was carried out, namely FlexX as available in LeadIT 2.1.2 [12], CDOCKER as available in Accelrys Discovery Studio 3.0 [13], FRED 2.2.5 (OpenEye Scientific Software, Inc.) [14–16], AutoDock 4.2 [17–20], and GOLD 5.2 [21]. These programs were selected on the basis of the intrinsic differences in the algorithms they use to dock the ligand to the receptor's binding site. GOLD and AutoDock use a genetic algorithm [7], FRED uses a docking algorithm based on Gaussian function [16], CDOCKER uses a molecular dynamics simulated annealing-based docking algorithm (CHARMm) [22], and FlexX uses an incremental construction algorithm [7, 23, 24].

The main goal was to test these five docking tools and select the ones that are most appropriate to be used for docking studies on DC-SIGN and, potentially, to extrapolate our findings to other C-type lectins with similar poorly defined and non-druggable binding sites. The goal was also to determine whether the docking algorithm can predict the coordination bond between the core mannose and DC-SIGN and whether the scoring functions are able to evaluate the coordination bonds accurately. Furthermore, since DC-SIGN binds oligosaccharides that undergo massive desolvation upon binding, it was important to test whether different docking algorithms could correctly evaluate their binding, even if desolvation effects are not considered explicitly.

Methods

Crystal structures

Structures of protein-ligand complexes were downloaded from the RCSB Protein Data Bank (PDB) [25]. Several predefined criteria were selected to help select the most appropriate protein-ligand complexes that would later be used for comparative evaluation of docking tools: (i) the structure of the ligand should contain the core mannose entity to enable the coordination bonds with calcium ion to be studied, (ii) ligand structures should be variable, (iii) the molecular weights of the selected ligands should cover different ranges and (iv) the resolution of the crystal structure should be less than 2.00 Å. To meet these requirements, we selected the crystal structures of the following protein-ligand complexes:

(i) 2IT6—the complex of DC-SIGN CRD with dimannoside (Fig. 1) [26], (ii) 2XR5—the complex of the DC-SIGN CRD with pseudo-dimannoside mimic [27], and (iii) 2XR6—the complex of DC-SIGN CRD with pseudo-trimannoside mimic [28]. In addition, a protein-ligand complex of Langerin with dimannoside (PDB code: 3P5F) was used [29] to demonstrate the possible usefulness of the validated docking protocol for docking to other lectins similar to DC-SIGN.

The establishment of interactions between DC-SIGN and ligand's core mannose is, as discussed later, of paramount importance to ensure high binding affinity. Regardless of the size and flexibility of the ligands used in this study, all of them predominantly interact with the protein through the core mannose. Thepaut et al. [27] and Sutkeviciute et al. [28] stressed that, in the case of binding of pseudo-dimannoside [27] and pseudo-trimannoside mimics [28], most of the ligand that is attached to the core mannose does not form any "significant" interactions with the protein and remains flexible within the complex with the protein. Since pseudo-dimannoside mimic [27] and pseudo-trimannoside mimic [28] have multiple rotatable bonds and, in complex with the protein, tend to jut out toward the solvent and away from the protein (Fig. 2), we expected some difficulties during validation of docking, especially when compared to the docking validation of dimannoside Manα1-2Man [26]. Another reason for this is that the dimannoside in a crystal structure with DC-SIGN (PDB code: 2IT6) also forms hydrophobic interactions with Phe313 [26]. Besides forming interactions with the ligands, Phe313 also provides steric hindrance (Fig. 2), weakening the binding of linear oligosaccharides and defines a hydrophobic groove of the receptor, which could be exploited to improve affinity of the novel DC-SIGN antagonists [8, 11]. Thepaut et al. [27] suggested that the important amino acid residues for establishing the interactions between DC-SIGN and core mannose hydroxyl groups (especially 3-OH and 4-OH) are Glu347, Asn349, Glu354, and Asn365. This information was considered when defining pharmacophore constraints during the setup of docking protocols.

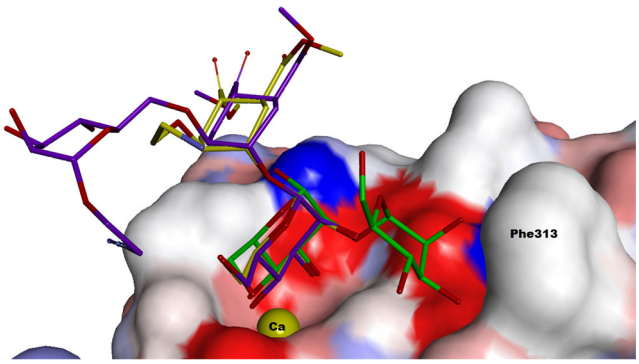

Fig. 2 Crystal structure of DC-SIGN CRD, showing the binding site with the Ca^{2+} ion and Phe313. The protein is presented as an atom charge surface and the Ca^{2+} ion is presented as yellow sphere. Ligands shown on this figure were taken from the above stated crystal structures and are presented as follows: ligand from DC-SIGN CRD-2IT6 (dimannoside Manα1-2Man [26]) is presented as green sticks, ligand from DC-SIGN CRD-2XR5 (pseudo-dimannoside mimic [27]) is presented as yellow sticks and ligand from DC-SIGN CRD-2XR6 (pseudo-trimannoside mimic [28]) is presented as violet sticks. Oxygen atoms are colored red, nitrogen atoms are colored blue (valid for all ligand structures)

The two dimensional structures of ligands are shown in Supplementary material S1 (Fig. S1–S4). The figures also contain information about interactions between the proteins and the ligands, as suggested by the program LeView [30].

Docking

General

The three dimensional structures of all ligands used for docking validation were extracted from selected crystal structures (PDB codes: 2IT6, 2XR5, and 2XR6; [2, 3]). The hydrogen atoms were added using Accelrys Discovery Studio 3.0 [13]. Addition of hydrogen atoms as well as the ligand structures were manually checked for possible irregularities. Charges were assigned to the protein during the protein preparation phase which is done prior to docking. The charges at the physiological pH=7.4 were considered. The geometries and charges of the molecules (except for FRED, as explained later) were optimized and assigned using the MMFF94 [31] force field and partial atomic charges in ChemBio3D Ultra 13.0. The energy was minimized until the gradient value was smaller than 0.001 kcal (mol Å)$^{-1}$. The optimized structure was further refined with GAMESS interface in ChemBio3D Ultra 13.0, using the semi-empirical PM3 method, QA optimization algorithm and Gasteiger Hückel charges for all atoms for 100 steps [32]. For the final structure refinement we selected one of the advanced semi-empirical methods (i.e., PM3) as available in ChemBio3D Ultra 13.0 [32].

All crystallographic water molecules were removed during preparation of the proteins. After superimposition of all three crystal structures we thoroughly analyzed the presence of water molecules around the binding site to identify possible recurrence of a specific water molecule in all three crystal structures. This would suggest that this water molecule has a crucial role in establishing binding interactions between different ligands and the protein. Including such a water molecule would possibly improve our docking validation. However, if such a recurring water molecule is not present, we considered that including specific crystallographic water molecules, important for binding of a particular ligand, would not be beneficial for predicting the binding of designed DC-SIGN antagonists with unknown affinity. Only one water molecule was identified that was in the proximity of the binding site and recurred in all three selected crystal structures. However it was still not close enough to affect the binding of ligands and thus the results of our docking experiment (see

Supplementary material S2, Fig. S5). We believe that water molecules affect the ligand binding in vivo, but for our in silico experiment the presence of crystallographic water (currently) adds no value.

A binding site was predefined in such a way that (i) all amino acids relevant for establishing the interactions with the ligand and (ii) the Ca^{2+} ion, which is the most important for binding of the core mannose to the binding site, were selected as part of the sphere/grid box. During the detailed examination of the binding site [10, 27] some amino acid residues that are involved in forming hydrogen bonds with either core mannose or with other functional groups of the ligand were recognized. Those amino acids were also included in the binding site definition (see Supplementary material S3, Fig. S6).

The size of the binding site was defined by a sphere/grid box, centered on the co-crystallized ligand. The size was set on the basis of the ligand pose in the crystal structure. We did our utmost to define the binding sites of comparable sizes. Since different programs have a different geometrical body as a "marker" for the chosen active site (sphere, cube or rectangle) the exact size for all selected programs could not be chosen. Details on the volume/size of selected active sites for each selected program and for each examined protein are described below in more detail. The reason for the variation of the sphere/grid box volume is the difference in the sizes of the ligands used in our experiment. With setting the proper sphere/grid box volume, free rotation and placement of ligand were allowed and situations were avoided in which these two parameters would be constrained by the sphere/grid box volume and, in such a way as to (negatively) impact the docking results [33].

The protein was kept rigid during all docking iterations within all docking programs, whereas the ligands were treated in accordance with the predefined docking settings. To decide whether we should also perform flexible docking, we performed a superposition of all three selected crystal structures (see Supplementary material, Fig. S7) at the very beginning of our study. We found that there are no major conformational changes of the protein, even after different ligands are bound to it. Especially we put the emphasis to the Ca^{2+} binding site of the DC-SIGN. Except for the slight shift of Phe313 no significant shifts of other, for the ligand binding crucial amino acid residues (Glu347, Asn349, Glu354 and Asn365, as described above) were spotted. Thus, we decided to progress only with the rigid docking.

Ten to 25 docking runs using each of the above stated docking programs were made using each of the three protein-ligand complexes. Settings of the docking program were kept as set by default or, more commonly, modified in a logical fashion before each iteration. Individual changes of selected parameters were made with the purpose of optimizing the default settings. The number of docking runs differs between the selected programs. If the results for a ligand, after a certain number of iterations, were still way (e.g., RMSD> 4.00 Å) from the desired value of RMSD (≤2.00 Å), the program was marked as "not suitable" for docking of ligands to DC-SIGN. It is important that, even for those programs, we calculated the RMSD of the core mannose pose compared to the core mannose pose from the crystal structure. The core mannose pose is of paramount importance [27] for establishing bonds between the ligand and the protein. Thus, its position is presumed to be even more important than the overall position of the ligand.

The aim was to set up the docking protocols within all numbered programs in such a way that the top pose of docking will also have an RMSD less than 2.00 Å or, in the best possible situation, that the top pose will also have the lowest RMSD value. The second aim was to check, within the top ranked docking results that are either suitable or not (in regard to RMSD), how well the selected docking programs are able to reproduce the pose of the core mannose in comparison with its pose in the crystal structure. Detailed information about the docking protocols for all the selected programs is given in Supplementary material S5 (Tables S1 and S2).

FlexX

Protein preparation The protein was prepared using FlexX (Receptor Definition Wizard) [12]. Water molecules were deleted from the crystal structures of DC-SIGN-ligand complexes (PDB codes: 2IT6, 2XR5, and 2XR6). The amino acid residues within a radius from 6 to 8 Å around the ligand were defined as the ligand-binding site. Due to its importance, the Ca^{2+} ion was preserved at the binding site. Ligands were also preserved to be used later as the reference ligands, which allows direct RMSD calculation immediately after the docking run is finished. Instead of pharmacophore constraint definition [8], spatial constraints for the core mannose of ligands extracted from certain DC-SIGN-ligand complexes (PDB codes: 2XR5 and 2XR6) were defined.

Ligand docking Ligands were docked in 25 independent runs. The certain docking parameters were changed between the runs with the purpose to optimize the docking efficiency. Amongst those the scoring parameters, chemical parameters, and docking details were altered in a controlled fashion. On the other hand the ligand binding was, after the initial runs, where each of the three possible approaches were tested, always driven by enthalpy and entropy (hybrid approach). Only the top 50 docked poses were preserved after each docking run. All were inspected visually and the best ranked FlexX-calculated conformation was used for analysis and representation. Top ten docking poses were re-scored with HYDE scoring function, as available in LeadIT 2.1.2, which is based

on a consistent description of hydrogen bond and desolvation energies in protein-ligand complexes [34].

CDOCKER

Protein preparation The protein was prepared using Accelrys Discovery Studio 3.0 [13] with the built in functionality "Prepare Protein". Water molecules were deleted from the crystal structures of DC-SIGN-ligand complexes (PDB codes: 2IT6, 2XR5, and 2XR6). The amino acid residues within a radius from 10 to 12 Å around the center of the ligand were defined as the ligand-binding site. We also defined a key amino acid residues (pharmacophore constraints), that are, besides the Ca^{2+} ion, involved in protein-ligand interactions. Three amino acid residues were thus selected as important for protein-ligand interactions—Asn349, Glu354, and Asn365.

Ligand docking Accelrys Discovery Studio 3.0 has been used to integrate several docking algorithms (CDOCKER, LibDock, LigandFit, and MCSS) [13]. Amongst these the one that has the most suitable docking algorithm for our docking study, i.e., CDOCKER [22], was selected, being a grid-based molecular docking method that employs CHARMm. As was done for experiments with other docking tools, the receptor was kept rigid during each iteration. Different ligand conformations were generated from the initial ligand structure through high temperature molecular dynamics, followed by random rotations. These conformations were then refined by grid-based, simulated annealing molecular dynamics with a final grid-based or full force field minimization. The CHARMm energy (interaction energy plus ligand strain) and the interaction energy alone were calculated for each docked pose. The latter were then sorted by CHARMm energy and just the top scoring poses retained [13]. The default settings of CDOCKER protocol were aligned slightly to comply with our needs. The number of random conformations to be calculated and refined by the program was increased. This allows a wider conformational field of the investigated ligand to be explored. Ligands were docked in 25 independent runs. All docking solutions were inspected visually and the best ranked CDOCKER-calculated conformation was used for analysis and representation.

FRED

Ligand and protein preparation An initial docking study was performed using FRED software (Release 2.2.5, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com/). The protein was prepared using FRED Receptor (Release 2.2.5, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com/) [14–16]. We determined the grid box size; inner and outer contours of the grid box were calculated automatically. The outer contour size ranged from 3190 to 4257 $Å^3$ (depending on the size of the ligand), while the grid box sizes varied from 4627 to 5751 $Å^3$. Pharmacophore constraints were then defined that are important for interactions between ligands (especially the core mannose) and the protein. Amino acid residues Asn365, Glu347, and Glu354 and Ca^{2+} ion were identified as being important. FRED requires the preparation of input conformers of each ligand prior to docking. These conformers were generated using another program from OpenEye Scientific Software, i.e., OMEGA 2.4.3 [35].

In an additional docking study the docking performances of two different releases of FRED software were compared. In particular, docking results obtained from FRED 2.2.5 were compared with those obtained from a newer release of FRED 3.0.1 [14–16], available in OEDocking software (Release 3.0.1, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com/) [36]. The protein was prepared using MAKE Receptor application (Release 3.0.1, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com/) [37]. The protein was prepared in nearly the same way as with FRED Receptor 2.2.5, as described above. The grid box size and contour sizes were set to have approximately the same values as for proteins prepared by FRED Receptor 2.2.5. The outer contour size ranged from 3232 to 4526 $Å^3$, and the grid box sizes from 5906 to 6608 $Å^3$. The setup of contours was set as "balanced". The same pharmacophore constraints were selected as for the initial docking study. No new input conformers of the ligands were generated with OMEGA—the same ones as those for initial docking study were also used with FRED 3.0.1.

Ligand docking Ligands were docked using FRED (version 2.2.5. OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com/) [14–16] in 20 independent runs. Exhaustive scoring was performed using Chemgauss3 scoring function. Further optimization was done using OEChemscore scoring function. Scoring and consensus pose selection were performed using Chemgauss3, OEChemscore, and CGT scoring functions. Other settings were set as default. All docking solutions were inspected visually and the best ranked FRED-calculated conformation was used for analysis and representation.

Two further independent runs were performed (named "variant 1" and "variant 2") for each of the ligands, using the newer release of FRED docking software (ver. 3.0.1). For "variant 1" we used the prepared proteins and ligand conformations already used for the initial docking using FRED 2.2.5. However, the modified docking protocol had to be applied, as a consequence of the changes made to the FRED due to the upgrade to version 3.0.1 (see release notes on the OpenEye home page for more details: http://docs.eyesopen.com/oedocking/releasenotes.html). For "variant 2" the re-prepared proteins were used (described above) with the

same ligand conformations and the same modified docking protocol as for "variant 1".

AutoDock 4.2

Ligand and protein preparation Ligand and protein preparation as well as docking with AutoDock 4.2 [17–20] were performed using the AutoDockTools (ADT) 1.5.4 [18]. In the first step ADT was used to prepare the protein and ligands: i.e., to determine protonation, to calculate charges, and specify rotatable bonds of the ligands and the protein and to define the position of the grid box and its dimensions. The grid box size varied from 16×16×16 Å to 18×18×18 Å, depending on the size of the ligand. The protein and ligands were stored as the .pdbqt format, which is a prerequisite for docking using AutoDock [17–20].

Ligand docking Initially the default settings suggested by the AutoDock were used. Eventually independent changes were made to some parameters, e.g., the translational step (Å), torsion step (degrees), cluster tolerance (Å), etc. Twenty individual GA runs were performed with different combinations of parameter settings or customized docking settings. Lamarckian GA [20] was used throughout all runs.

GOLD 5.2

Protein preparation Molecular docking calculations were performed using GOLD program version 5.2 [21]. Ligands and water molecules were deleted from the crystal structures of DC-SIGN-ligand complexes (PDB codes: 2IT6, 2XR5, and 2XR6), and hydrogen atoms were added to the protein using GOLD. The amino acid residues within a radius of 6 Å around the ligand were defined as the ligand-binding site.

Ligand docking Ligands were docked in ten independent GA runs. The GA parameters were set as suggested by GOLD 5.2. ASP [38]. CHEMPLP [39], GOLDScore [40] or ChemScore [41, 42] was used as scoring function. All docking solutions were inspected visually and the best ranked GOLD-calculated conformation was used for analysis and representation.

Visualisation of docking results and RMSD calculation

Docking solutions were visualized directly in the docking program, if applicable (e.g., FlexX). In cases where this functionality is not built in the docking program or has too many limitations, either PyMol 1.3.1 [43] or Accelrys Discovery Studio 3.0 [13] were used.

The root-mean-square deviation, RMSD was selected as the representative parameter for evaluating docking success [44]. As a reference for calculation of RMSD values the ligand poses from the above mentioned crystal structures were used. After every successfully completed docking run we performed a visual check, i.e., comparison of crystal structure ligand pose and the docked poses. After the visual check was done and the results were evaluated as promising we used crystal structure ligand poses to obtain the RMSD values for the top ten ranked poses. FRED, which gave the best RMSD values for the top ranked poses, also performed the best when considering overall RMSD values, i.e., taking into account RMSD values of the top ten ranked poses for all three ligands. Supplementary material S6 (Table S3) shows how many of the top ten ranked poses for each docking tool correspond to the acceptance criteria of RMSD ≤2.00 Å.

The heavy atom RMSD values were retrieved in two ways. If the docking program did not support direct calculation of RMSD (e.g., FRED), Accelrys Discovery Studio 3.0 [13] was used for the calculation. Otherwise, some of the docking programs were used to calculate the RMSD values instantly just after the end of the docking experiment (e.g., FlexX). Only the position of heavy atoms of the ligand was taken into account for RMSD calculation.

Results and discussion

The main goal of our experiment was to find the most appropriate docking program and protocol for use in the structure-based design of novel DC-SIGN ligands. The following criteria were used:

a. The docking program should be able to satisfactorily re-dock the ligand from the crystal structure of the protein-ligand complex. The upper RMSD limit which still fulfills our criteria of acceptance was set to 2.00 Å, this being the most common limit in docking studies within various groups [45, 46].
b. The docking program should be able to re-dock the core mannose of the ligand contained in the crystal structure of the protein-ligand complex satisfactory well. The core mannose is of paramount importance for establishing interactions between mannose containing ligands and DC-SIGN CRD [27]. Its pose in the binding site is strongly related to establishment of the coordination bond with Ca^{2+} and the hydrogen bonds with the amino acid residues in the active site [10].

Docking programs able to re-dock ligands and core mannose within an RMSD≤2.00 Å were selected and evaluated as most appropriate of the already selected docking programs for use for further docking studies of various newly designed ligands.

Table 1 RMSD values obtained for re-docked ligands of different DC-SIGN CRD-ligand complexes

	FRED	AutoDock	FlexX	HYDE	GOLD	CDOCKER
RMSD (Å)	DC-SIGN 2IT6					
	0.81	3.50	0.61	0.79	1.68	0.97
	DC-SIGN 2XR5					
	1.32	4.21	7.26	2.01	4.31	1.22
	DC-SIGN 2XR6					
	1.68	5.48	7.44	6.49	6.82	2.34

Re-docking of ligands

Docking results obtained after the appropriate docking parameters for each of the selected programs were established were evaluated. The RMSD was calculated as described above. The calculated RMSD values for the top ranked ligand pose for all three crystal structures and all stated docking programs are listed in Table 1). The information on how many of the top ten ranked ligand poses met the selected criteria of RMSD≤2.00 Å is presented in Supplementary material S6 (Table S3).

Using FRED, all three ligands were re-docked within the RMSD cut-off value of 2.00 Å. Promising results were also achieved using CDOCKER, which showed a minor deviation from the desired result only for the most complex ligand (DC-SIGN 2XR6). FlexX showed exceptional performance with the first ligand (DC-SIGN 2IT6), but failed later with the other two ligands, which were far away from the desired value. When FlexX docking poses were re-scored by HYDE, significant improvements of RMSD values were observed in the case of more complex ligands. GOLD and AutoDock did not perform satisfactorily, even at the first stage, i.e., re-docking of the least complex disaccharide ligand (DC-SIGN 2IT6). Re-docking of all the ligands, using all five docking programs, showed that the RMSD increases with greater complexity of the ligand. The docking results obtained by FRED and CDOCKER are shown in graphical form (Figs. 3, 4, and 5).

Re-docking of the core mannose

Docking results were also evaluated considering only the pose of the core mannose. The RMSD was calculated using the Accelrys Discovery Studio 3.0 [13]. Only the positions of heavy atoms were considered in the calculations. The RMSD values for core mannose for all of the three crystal structures are presented (Table 2).

Once again FRED performed well, with an average RMSD for the core mannose of 0.79 Å. CDOCKER outperformed FRED with an average RMSD for the core mannose of 0.62 Å. Of the three other docking programs only FlexX and GOLD showed promising results, while AutoDock was not satisfactory. HYDE re-scoring improved the RMSD in the case of the least complex ligand (DC-SIGN 2IT6), while in the other two cases RMSD slightly increased. The graphic presentations of the docking results obtained by FRED and CDOCKER are shown in Figs. 6 and 7.

When RMSD values for core mannose alone were considered, most programs successfully predicted its pose, but not when RMSD values for the whole ligands were considered.

The best results altogether were obtained using FRED although CDOCKER predicted the correct pose of the core mannose even slightly better than FRED. Since the positioning of core mannose is of paramount importance [10] and is the key decision-making parameter, the fair performance of FlexX should be noted. Quite promising values were obtained when focusing only on the positioning of core mannose. The satisfactory performance for docking of mannose-based ligands to DC-SIGN CRD using FlexX was also tested and confirmed [8].

Gowthaman et al. showed that docking of small molecules to a flatter binding site is highly likely to fail, since in this case docking programs rarely predict the right poses of small molecules [47], i.e., there is a strong chance that a docking algorithm will dock different (either correct, wrong or both) ligand

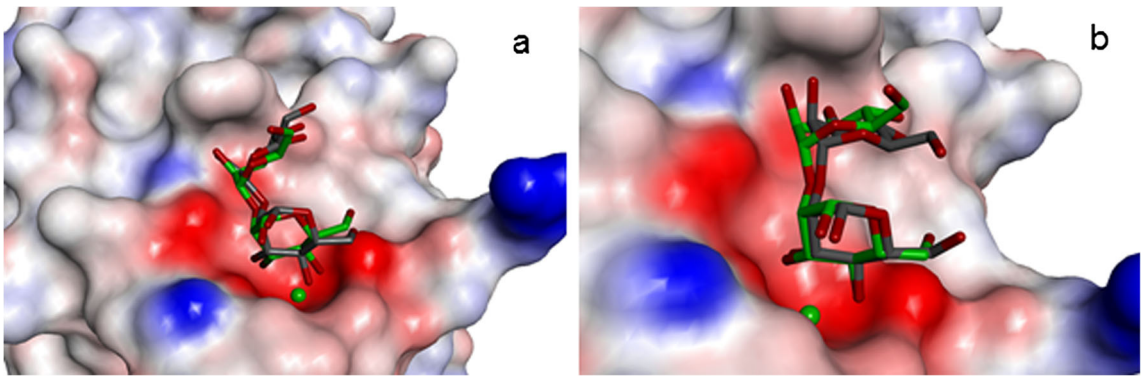

Fig. 3 Re-docked dimannoside using **a** FRED (RMSD=0.81 Å) and **b** CDOCKER (RMSD=0.97 Å) to DC-SIGN CRD, 2IT6. The position of the ligand from X-ray determination is *colored green*; the docked structure is *colored gray*; oxygen atoms are *colored red* and nitrogen atoms are *colored blue* (valid for both structures)

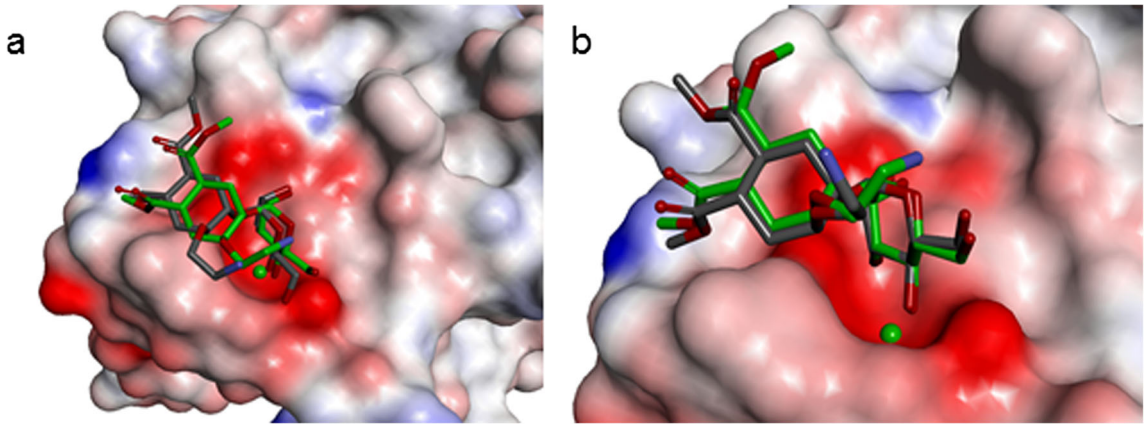

Fig. 4 Re-docked pseudo-1,2-mannobioside using **a** FRED (RMSD= 1.32 Å) and **b** CDOCKER (RMSD=1.22 Å) to DC-SIGN CRD, 2XR5; the position of the ligand from X-ray determination is *colored green*; the docked structure is *colored gray*; oxygen atoms are *colored red* and nitrogen atoms are *colored blue* (valid for both structures)

poses to the binding site. The values of assigned binding energy could be of similar values because scoring functions are not able to accurately estimate all of the possible interactions and/or desolvation of the target protein and ligand pose. To epitomize: there is a high risk for an error to be made in this kind of protein-ligand docking studies, thus only the scarce docking programs (or more exactly—docking algorithms) are appropriate for use.

To better understand the differences in performance between the docking programs, we tried relating the docking results to the algorithms used by each of the docking programs. From the results we can conclude that the genetic algorithm (used by GOLD and AutoDock [7]) is inappropriate for the docking of various ligands to binding site of DC-SIGN. The results also suggest that the incremental construction algorithm (used by FlexX) [7, 23, 24] should not be considered as the first choice for docking studies involving DC-SIGN. As was shown by Tomašić et al. [8], FlexX can more or less successfully reproduce only the ligands' residues that interact with the protein, but does not function well once the ligand is reaching out to the solvent rather than interacting with the protein. Considering this, FlexX is conditionally useful as a docking program when dealing with DC-SIGN. Thus the two most appropriate docking algorithms for use with DC-SIGN are CDOCKER's molecular dynamics simulated-annealing-based docking algorithm (CHARMm) [22] and FRED's docking algorithm that is based on Gaussian function [16].

Asensio et al. [48] discussed the importance of carbohydrate recognition by aromatic rings of receptors. As briefly mentioned above, this phenomenon was also noted within the binding of dimannoside to DC-SIGN (PDB code 2IT6), where a rather hydrophilic mannose forms hydrophobic interactions with Phe313 [26]. The prediction of hydrophobic interactions by docking programs is very difficult due to the entropic effect of desolvation of interacting groups. On the basis of our results we can conclude, that FRED and CDOCKER are both capable of reproducing quite well not only complex interactions but also hydrophobic interactions between the ligand and the protein.

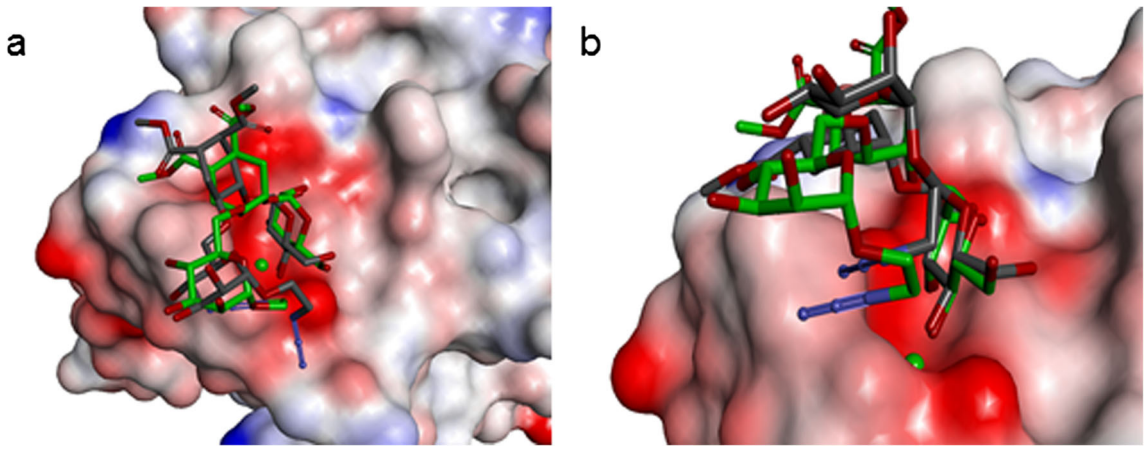

Fig. 5 Re-docked pseudo-trimannoside mimic using **a** FRED (RMSD= 1.68 Å) and **b** CDOCKER (RMSD=2.34 Å) – DC-SIGN CRD, 2XR6; position of the ligand from X-ray is *colored green*, docked structure is *colored gray*; oxygen atoms are *colored red* and nitrogen atoms are *colored blue* (valid for both structures)

Table 2 RMSD values obtained for re-docked core mannose

	FRED	AutoDock	FlexX	HYDE	GOLD	CDOCKER
RMSD (Å)	DC-SIGN 2IT6					
	0.65	3.40	0.79	0.55	1.04	0.73
	DC-SIGN 2XR5					
	0.97	1.36	0.79	0.92	0.80	0.34
	DC-SIGN 2XR6					
	0.76	2.35	2.96	3.11	8.71	0.78

Re-docking of ligands using two different releases of FRED

In addition to successful validation of FRED 2.2.5 we aimed to compare and validate the newer release of FRED (ver. 3.0.1). There are several differences between the two versions (see release notes on the OpenEye home page for more details: http://docs.eyesopen.com/oedocking/releasenotes.html). We concluded that the changes made to scoring functions (the majority of the latter were omitted, Chemgauss scoring function was improved), could have the biggest impact on the validation of the docking program. Thus, two independent runs were performed, "variant 1" and "variant 2", as described in Methods, using FRED 3.0.1. The results, compared to ones from initial docking validation with FRED 2.2.5 are shown in Table 3.

It is evident that the newer version of FRED did not deliver any promising results. The RMSD values are above the set limit of 2.00 Å for all three ligands.

As we have shown earlier in this paper, there is a possibility within some of the docking programs that, even if the RMSD for the whole ligand is high, the RMSD of core mannose could be suitable. Considering this lesson learned, we have also calculated the RMSD solely for the core mannose of the re-docked ligands. The results are presented in Table 4.

The RMSD values for core mannose are somehow promising when compared to the RMSD values of the whole ligands. Nevertheless, the RMSD values for core mannose (especially the value from the more complex ligands) are well above the desired value of 2.00 Å.

We conclude, therefore, that the newer release of FRED is in our case not predictive as the previous release FRED 2.2.5. We believe that omission of most of the scoring functions is the most probable cause for the differences in RMSD values. As already shown earlier, the consensus scoring of scoring functions Chemgauss3, OEChemscore, and CGT was used for the validation of FRED 2.2.5. In contrast, FRED 3.0.1 does not support consensus scoring, but rather applies only one scoring function, i.e., Chemgauss4.

We conclude that, for the docking studies on DC-SIGN, the older version of FRED that supports consensus scoring of several scoring functions is more appropriate and is a preferable choice.

Modeling of complex bonds by selected docking tools

Modeling of complex bonds still represents a challenge in the development of docking tools, especially due to parameterization and accurate modeling of interaction geometry (a geometry of complex bond is very hard to reproduce) [49, 50].

To the best of our knowledge only FlexX allows advanced modeling of complex bonds that is done by defining a pharmacophore constraint, which forces a specific (complex) interaction between the ligand and the protein. The rest of selected docking programs unfortunately have limitations in regard to complex bonds modeling [49].

Since we got the best docking results using FRED and CDOCKER we can summarize that the modeling of complex bonds only was not sufficient for the successful docking. Beside complex bonds, we also have to take into account hydrogen bonds, van der Waals, and hydrophobic interactions. Thus, we incorporated the amino acid residues (Glu347, Asn349, Glu354, and Asn365) that were also recognized as important for ligand binding and protein-ligand interaction establishment [27]. If the interaction of core mannose with above mentioned amino acid residues is reached we can presume that the complex bonds would be established in vivo due to proximity of 3-OH and 4-OH groups of core mannose to Ca^{2+} ion.

Performance of FRED 2.2.5 docking protocol in a virtual screening-like experiment

In order to establish whether FRED 2.2.5 docking protocol can discriminate between active and inactive compounds, we performed a virtual screening-like experiment using our

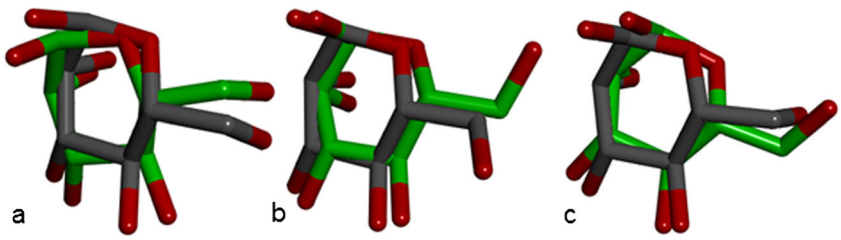

Fig. 6 The core mannose (**a**: 2IT6, **b**: 2XR5, and **c**: 2XR6; docking results obtained by FRED) — the pose of the core mannose from X-ray *colored green*, docked structure *colored gray*

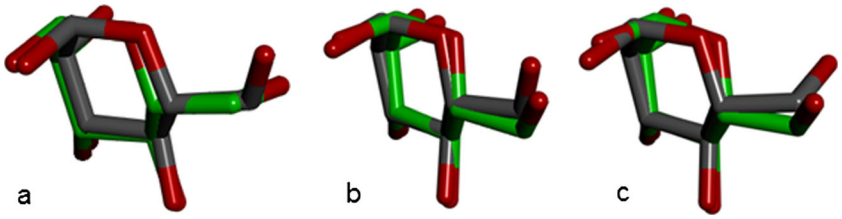

Fig. 7 The core mannose (**a**: 2IT6, **b**: 2XR5, and **c**: 2XR6; docking results obtained by CDOCKER) — pose of the core mannose from X-ray *colored green*, docked structure *colored gray*

recently reported DC-SIGN antagonists [8, 51] with IC_{50} values between 0.040 mM and 14.94 mM as known actives (see Supplementary material S7, Table S4). Since some of these compounds are mixtures of two diastereomers, both isomers were considered for docking, which resulted in a set of 32 compounds. A set of 50 decoys per ligand having similar physico-chemical properties but dissimilar 2D topology were generated using the Database of Useful Decoys: Enhanced (DUD-E) server [52]. The obtained ligand library thus consisted of 1632 compounds, for which an average of 181 conformers per ligand were generated using OMEGA 2.4.3 [35]. The above described docking protocol with Chemgauss3 exhaustive scoring, OEChemscore optimization and final consensus scoring was used for virtual screening-like experiment using crystal structure of DC-SIGN in complex with dimannoside (PDB code 2IT6). Analysis of the results showed that the consensus scoring did not provide the sufficiently good results meaning that it did not successfully retrieve known active compounds from the ligand library. However, when the hitlist was ranked using only OEChemscore scoring function, all 32 compounds with known activity were identified up to rank 91, which is amongst the top 5.6 % of the docked poses. Although OEChemscore was not able to rank the actives according to their experimental affinity, the two most active compounds 2d and 6 were ranked as fourth and fifth, which altogether highlights the FRED 2.2.5 docking protocol as useful for virtual screening for novel DC-SIGN antagonists. Experimental IC_{50} values, OEChemscore scores, and hitlist ranks are shown in Table S4 in the Supplementary material S7.

Applicability of FRED docking protocol to other, similar, lectins

To prove that the validated FRED 2.2.5 docking protocol is useful, not just for DC-SIGN, but can be applied to other lectins with a poorly defined binding site, we performed a further docking study on Langerin (PDB code 3P5F). Langerin is known to bind several ligands that also bind to DC-SIGN, including HIV-1 gp120 [53, 54].

The ligand, which was taken from the crystal structure, and protein were prepared and the docking performed in the same way as described for DC-SIGN. The aim was to re-dock the dimannoside, the grid box size, and outer contour size were set to values as similar as possible to those for DC-SIGN (PDB code: 2IT6). The size of the outer contour was 3174 $Å^3$ and the grid box size was 4264 $Å^3$. Since a different protein with a different crystal structure is the subject of this docking run, we had to choose different pharmacophore constraints. Feinberg et al. [29] suggested some amino acid residues that are important for interactions between the ligand's core mannose and Langerin. We selected the following ones as pharmacophore constraints for hydrogen bond formation: Asn287, Asn307, Glu285, and Glu293. In addition Asp308 was selected as a pharmacophore constraint, as it could be important for establishing the hydrogen bond with the ligand.

The best ranked pose gave an RMSD value of 0.94 Å when compared to the ligand pose from the crystal structure (PDB code: 3P5F). Six out of ten top ranked poses had RMSD values≤2.00 Å. It can be concluded therefore that the FRED

Table 3 RMSD values obtained for re-docked ligands of different DC-SIGN CRD-ligand complexes using two different releases of FRED

	FRED 2.2.5	FRED 3.0.1 (»Variant 1«)	FRED 3.0.1 (»Variant 2«)
RMSD (Å)	DC-SIGN 2IT6		
	0.81	2.51	2.52
	DC-SIGN 2XR5		
	1.32	5.26	5.32
	DC-SIGN 2XR6		
	1.68	9.03	9.46

Table 4 RMSD values obtained for re-docked core mannose using two different releases of FRED

	FRED 2.2.5	FRED 3.0.1 (Variant 1)	FRED 3.0.1 (Variant 2)
RMSD (Å)	DC-SIGN 2IT6		
	0.65	0.85	0.80
	DC-SIGN 2XR5		
	0.97	2.88	3.06
	DC-SIGN 2XR6		
	0.76	6.33	3.05

2.2.5 docking protocol, validated for use with DC-SIGN, can also be of value for docking studies on other lectins having similar binding site properties and, potentially, on other proteins with shallow and poorly defined binding sites.

Conclusion

Five of the currently available docking tools have been validated for docking to a protein with a shallow binding site. Besides locating the best program for docking studies on DC-SIGN, two different versions of the same program have been tested for appropriateness for the docking to DC-SIGN. The validated docking protocol has also been shown to be applicable to other C-type lectins with similar binding site properties.

Of the docking programs selected only FRED 2.2.5 and CDOCKER (Accelrys Discovery Studio 3.0) were found to be appropriate for docking studies on DC-SIGN. This is based on the calculated RMSD values for each of the re-docked ligands and on the calculated RMSD values for the core mannose. Both programs outperformed other docking programs when considering both criteria. Performance of FRED 2.2.5 docking protocol was also tested in a virtual screening-like experiment during which we demonstrated that validated docking protocol can discriminate fairly well between active and inactive compounds.

We also compared the appropriateness of a different version of the same program, the newly released FRED 3.0.1 (as available in OEDocking 3.0.1), for docking studies on DC-SIGN. Repeating the docking and calculating the RMSD values for whole ligands and core mannose with FRED 3.0.1, we showed that the older version of FRED functions better in docking to DC-SIGN. We presume this is due to the possibility of consensus scoring with FRED 2.2.5 that is not available in the newer release of FRED.

The possible use of the FRED 2.2.5 docking protocol for docking studies on other C-type lectins that are similar to DC-SIGN (i.e., possess shallow and poorly defined binding sites) was examined. The successful re-docking of the dimannoside ligand to another C-type lectin, Langerin (PDB code: 3P5F) with RMSD≤2.00 Å suggests the usefulness of the FRED 2.2.5 docking protocol for other, similar C-type lectins.

Acknowledgments The authors thank OpenEye Scientific Software, Inc. for free academic licenses for their software. The authors thank Professor Roger Pain for proofreading the manuscript.

Compliance with ethical guideline

Funding This work was supported by the Slovenian Research Agency (Grant No. P1-0208).

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Brooijmans N, Kuntz ID (2003) Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct 32:335–373. doi:10.1146/annurev.biophys.32.110601.142532
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28(1):235–242. doi:10.1093/nar/28.1.235
3. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1978) The protein data bank: a computer-based archival file for macromolecular structures. Arch Biochem Biophys 185(2):584–591. doi:10.1016/0003-9861(78)90204-7
4. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) A geometric approach to macromolecule-ligand interactions. J Mol Biol 161(2):269–288. doi:10.1016/0022-2836(82)90153-X
5. Bharatham N, Bharatham K, Shelat AA, Bashford D (2014) Ligand binding mode prediction by docking: mdm2/mdmx inhibitors as a case study. J Chem Inf Model 54(2):648–659. doi:10.1021/ci4004656
6. Kroemer RT (2007) Structure-based drug design: docking and scoring. Curr Protein Pept Sci 8(4):312–328. doi:10.2174/138920307781369382
7. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3(11):935–949. doi:10.1038/nrd1549
8. Tomašić T, Hajšek D, Švajger U, Luzar J, Obermajer N, Petit-Haertlein I, Fieschi F, Anderluh M (2014) Monovalent mannose-based DC-SIGN antagonists: targeting the hydrophobic groove of the receptor. Eur J Med Chem 75:308–326. doi:10.1016/j.ejmech.2014.01.047
9. Obermajer N, Sattin S, Colombo C, Bruno M, Švajger U, Anderluh M, Bernardi A (2011) Design, synthesis and activity evaluation of mannose-based DC-SIGN antagonists. Mol Divers 15(2):347–360. doi:10.1007/s11030-010-9285-y
10. Anderluh M, Jug G, Švajger U, Obermajer N (2012) DC-SIGN antagonists, a potential new class of anti-infectives. Curr Med Chem 19(7):992–1007. doi:10.2174/092986712799320664
11. Feinberg H, Mitchell DA, Drickamer K, Weis WI (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. Science 294(5549):2163–2166. doi:10.1126/science.1066371
12. LeadIT version 2.1.3 is available from BioSolveIT (GmbH), http://www.biosolveit.de/
13. Accelrys Discovery Studio 3.0 is available from Accelrys Inc, San Diego http://accelrys.com/
14. McGaughey GB, Sheridan RP, Bayly CI, Culberson JC, Kreatsoulas C, Lindsley S, Maiorov V, Truchon JF, Cornell WD (2007) Comparison of topological, shape, and docking methods in virtual screening. J Chem Inf Model 47(4):1504–1519. doi:10.1021/ci700052x
15. McGann M (2011) FRED pose prediction and virtual screening accuracy. J Chem Inf Model 51(3):578–596. doi:10.1021/ci100436p
16. McGann MR, Almond HR, Nicholls A, Grant JA, Brown FK (2003) Gaussian docking functions. Biopolymers 68(1):76–90. doi:10.1002/bip.10207
17. Morris GM, Goodsell DS, Huey R, Olson AJ (1996) Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J Comput Aid Mol Des 10(4):293–304. doi:10.1007/BF00124499
18. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30(16):2785–2791. doi:10.1002/jcc.21256

19. Goodsell DS, Olson AJ (1990) Automated docking of substrates to proteins by simulated annealing. Proteins 8(3):195–202. doi:10.1002/prot.340080302
20. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19(14):1639–1662. doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AIDJCC10>3.0.CO;2B
21. GOLD version 5.2 is available from The Cambridge Crystallographic Data Centre, Cambridge, www.ccdc.cam.ac.uk/
22. Wu G, Robertson DH, Brooks CL 3rd, Vieth M (2003) Detailed analysis of grid-based molecular docking: a case study of CDOCKER-A CHARMm-based MD docking algorithm. J Comput Chem 24(13):1549–1562. doi:10.1002/jcc.10306
23. Bursulaya BD, Totrov M, Abagyan R, Brooks CL 3rd (2003) Comparative study of several algorithms for flexible ligand docking. J Comput Aid Mol Des 17(11):755–763. doi:10.1023/B:JCAM.0000017496.76572.6f
24. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. J Mol Biol 261(3):470–489. doi:10.1006/jmbi.1996.0477
25. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) The protein data bank. Acta Crystallogr D Biol Crystallogr 58(Pt 6 No 1):899–907. doi:10.1107/S0907444902003451
26. Feinberg H, Castelli R, Drickamer K, Seeberger PH, Weis WI (2007) Multiple modes of binding enhance the affinity of DC-SIGN for high mannose N-linked glycans found on viral glycoproteins. J Biol Chem 282(6):4202–4209. doi:10.1074/jbc.M609689200
27. Thepaut M, Guzzi C, Sutkeviciute I, Sattin S, Ribeiro-Viana R, Varga N, Chabrol E, Rojo J, Bernardi A, Angulo J, Nieto PM, Fieschi F (2013) Structure of a glycomimetic ligand in the carbohydrate recognition domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design. J Am Chem Soc 135(7):2518–2529. doi:10.1021/ja3053305
28. Sutkeviciute I, Thepaut M, Sattin S, Berzi A, McGeagh J, Grudinin S, Weiser J, Le Roy A, Reina JJ, Rojo J, Clerici M, Bernardi A, Ebel C, Fieschi F (2014) Unique DC-SIGN clustering activity of a small glycomimetic: a lesson for ligand design. ACS Chem Biol 9(6):1377–1385. doi:10.1021/cb500054h
29. Feinberg H, Taylor ME, Razi N, McBride R, Knirel YA, Graham SA, Drickamer K, Weis WI (2011) Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site. J Mol Biol 405(4):1027–1039. doi:10.1016/j.jmb.2010.11.039
30. Caboche S (2013) LeView: automatic and interactive generation of 2D diagrams for biomacromolecule/ligand interactions. J Cheminf 5(1):40. doi:10.1186/1758-2946-5-40
31. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 17(5–6):490–519. doi:10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P
32. GAMESS interface, ChemBio3D Ultra 13.0, CambridgeSoft
33. Wilantho A, Tongsima S, Jenwitheesuk E (2008) Pre-docking filter for protein and ligand 3D structures. Bioinformation 3(5):189–193. doi:10.6026/97320630003189
34. Schneider N, Lange G, Hindle S, Klein R, Rarey M (2013) A consistent description of hydrogen bond and dehydration energies in protein-ligand complexes: methods behind the HYDE scoring function. J Comput Aided Mol Des 27(1):15–29. doi:10.1007/s10822-012-9626-2
35. OMEGA 2.4.3: OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com/. Hawkins PCD, Skillman AG, Warren GL, Ellingson BA, Stahl MT (2010) Conformer generation with OMEGA: Algorithm and validation using high quality structures from the Protein Data Bank and Cambridge Structural Database. J Chem Inf Model 50:572–584. doi: 10.1021/ci100031x
36. OEDocking 3.0.1: OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com/
37. MAKE Receptor 3.0.1: OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com/
38. Mooij WTM, Verdonk ML (2005) General and targeted statistical potentials for protein—ligand interactions. Proteins 61(2):272–287. doi:10.1002/prot.20588
39. Korb O, Stützle T, Exner TE (2009) Empirical scoring functions for advanced protein—ligand docking with plants. J Chem Inf Model 49(1):84–96. doi:10.1021/ci800298z
40. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267(3):727–748. doi:10.1006/jmbi.1996.0897
41. Eldridge M, Murray C, Auton T, Paolini G, Mee R (1997) Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J Comput Aid Mol Des 11(5):425–445. doi:10.1023/A:1007996124545
42. Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD (1998) Flexible docking using tabu search and an empirical estimate of binding affinity. Proteins 33(3):367–382. doi:10.1002/(SICI)1097-0134(19981115)33:3<367::AID-PROT6>3.0.CO;2-W
43. Pymol is available from Delano Scientific LLC, San Francisco, CA, http://pymol.sourceforge.net/
44. Jain A (2008) Bias, reporting, and sharing: computational evaluations of docking methods. J Comput Aid Mol Des 22(3–4):201–212. doi:10.1007/s10822-007-9151-x
45. Cross JB, Thompson DC, Rai BK, Baber JC, Fan KY, Hu Y, Humblet C (2009) Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. J Chem Inf Model 49(6):1455–1474. doi:10.1021/ci900056c
46. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS (2006) A critical assessment of docking programs and scoring functions. J Med Chem 49(20):5912–5931. doi:10.1021/jm050362n
47. Gowthaman R, Deeds EJ, Karanicolas J (2013) Structural properties of non-traditional drug targets present new challenges for virtual screening. J Chem Inf Model 53(8):2073–2081. doi:10.1021/ci4002316
48. Asensio JL, Arda A, Canada FJ, Jimenez-Barbero J (2013) Carbohydrate-aromatic interactions. Accounts Chem Res 46(4):946–954. doi:10.1021/ar300024d
49. Seebeck B, Reulecke I, Kämper A, Rarey M (2008) Modelling of metal interaction geometries for protein-ligand docking. Proteins 71(3):1237–1254. doi:10.1002/prot.21818
50. Hu X, Balaz S, Shelver WH (2004) A practical approach to docking of zinc metalloproteinase inhibitors. J Mol Graph Model 22(4):293–307. doi:10.1016/j.jmgm.2003.11.002
51. Tomašić T, Rabbani S, Gobec M, Mlinarič Raščan I, Podlipnik Č, Ernst B, Anderluh M (2014) Branched α-D-mannopyranosides: a new class of potent FimH antagonists. Med Chem Commun 5(8):1247–1253. doi:10.1039/C4MD00093E
52. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK (2012) Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J Med Chem 55(14):6582–6594. doi:10.1021/jm300687e
53. Holla A, Skerra A (2011) Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. Protein Eng Des Sel 24(9):659–669. doi:10.1093/protein/gzr016
54. de Witte L, Nabatov A, Geijtenbeek TB (2008) Distinct roles for DC-SIGN + −dendritic cells and Langerhans cells in HIV-1 transmission. Trends Mol Med 14(1):12–19. doi:10.1016/j.molmed.2007.11.001