PROTEINS: Structure, Function, and Genetics 41:173–191 (2000)

DARWIN: A Program for Docking Flexible Molecules

Jeffrey S. Taylor and Roger M. Burnett*
The Wistar Institute, Philadelphia, Pennsylvania

ABSTRACT **A new program named "DARWIN" has been developed to perform docking calculations with proteins and other biological molecules. The program uses the Genetic Algorithm to optimize the molecule's conformation and orientation under the selective pressure of minimizing the potential energy of the complex. A unique feature of DARWIN is that it communicates with the molecular mechanics program CHARMM to make the energy calculations. A second important feature is its parallel interface, which allows simultaneous use of multiple standalone copies of CHARMM to rapidly evaluate large numbers of potential solutions. This permits an "accuracy first" approach to docking, which avoids many of the common assumptions and shortcuts often made to reduce computation time. The method was applied to three protein-carbohydrate complexes: the crystallographically determined structures of Concanavalin A and Fab Se155-4; and a model structure for Fab ME36.1. Conformations close to the crystal structures were obtained with this approach, but some "false positive" solutions were also selected. Many of these could be eliminated by introducing different methods for simulating solvent effects. An effective screening method for docking a database of compounds to a single target enzyme using DARWIN is also presented. Proteins 2000;41:173–191.** © 2000 Wiley-Liss, Inc.

Key words: genetic algorithm; protein-ligand interaction; carbohydrate binding; antibody-antigen recognition; drug design

INTRODUCTION

In recent years, considerable effort has been devoted to finding an effective automated procedure for molecular docking. A major driving force behind this effort is the need for cost-effective ways to identify lead compounds for new drugs. Additional momentum has been gained with the exponential growth in the number of high-resolution structures that have been determined for enzymes in critical metabolic pathways. With these structures, computer-based methods can be used to identify or design ligands that have good structural and chemical complementarity to the enzymatic active sites to serve as candidates for drug development. A common strategy for such structure-based drug design is to perform broad screening of many compounds using rapid approximations of the complex's potential energy. This approach is typified by the programs DOCK[1] and Hammerhead.[2]

A different approach, which is possible when only one ligand is to be studied, is the use of molecular dynamics techniques similar to those used in the refinement of molecular models against crystallographic or nuclear magnetic resonance experimental data. Several sophisticated molecular mechanics programs, including CHARMM[3] and AMBER,[4] have been developed to calculate the potential energy of proteins and their associated ligands, and to perform molecular dynamics. These "accuracy first" programs are computationally expensive, and therefore slow, but they usually yield the best results when evaluating possible docked solutions.[5–7] The slow speed of this approach has limited its application to cases where small ligands are to be docked into well-described binding sites. Here, the extent of the necessary search, also called the "solution space," is small.[8]

The work described in this article arose from a desire to refine a molecular model of the adenovirus capsid that had been derived from positioning the crystallographic model of hexon[9] within an image reconstruction of the virion determined by electron microscopy.[10] Initially, this was not possible due to the size of hexon (327 KDa) and the large number of copies in the icosahedral capsid (240). Currently available docking programs either could not handle large molecules, were too slow to be useful, or did not provide the required accuracy. A new docking program, "DARWIN," was developed to overcome these problems. The overall aim was to expand the scope of the molecular dynamics docking approach, while retaining its accurate energy calculations. DARWIN uses a standard genetic algorithm (GA)[11,12] to search large solution spaces, and a parallel interface to the program CHARMM to perform the energy calculations (note that while CHARMM is used here, any program can be linked to DARWIN for evaluating potential solutions). DARWIN drives the search for the correct placement of the ligand by using CHARMM to evaluate each potential solution. This unique approach

Grant sponsor: National Institutes of Health, Structural Biology Training Program; Grant number: T32-GM 08275; Grant sponsor: National Institutes of Health; Grant Number: AI 17270, CA 10815; Grant sponsor: National Science Foundation; Grant number: MCB 95-07102; Grant sponsor: Wistar Cancer Center.

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/index.html

Jeffrey S. Taylor's present address is Department of Biochemistry, Duke University Medical Center, Box 3711, Nanaline Duke Building, Durham, NC 27710.

*Correspondence to: Roger M. Burnett, The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104. E-mail: Burnett@wistar.upenn.edu

Received 23 November 1999; Accepted 11 May 2000

uses established well-tested CHARMM protocols to evaluate potential solutions, whereas most other docking programs require the user to use the program's own evaluation function that may not meet the requirements of the problem at hand. The parallel interface allows up to 128 copies of CHARMM to simultaneously evaluate these solutions. The easy scalability of DARWIN allowed its testing on a cluster of workstations in the laboratory, and then a facile move to a supercomputer for production runs. DARWIN is written in the C computer language and its source code is available from the authors.

DARWIN was first tested on carbohydrate docking problems as these were somewhat simpler than docking viral capsid proteins, although still challenging, and were relevant to other work in the laboratory. Oligosaccharides present difficulties for docking programs as they are flexible and have relatively weak binding constants, often in the range 10^{-3} to 10^{-6} M, which make it difficult to evaluate potential solutions. Nevertheless, previous studies of carbohydrates and carbohydrate-protein complexes have demonstrated CHARMM's utility in this area.[13–15] A primary biological goal of the study was to evaluate a model of the complex between the GD2 ganglioside and its monoclonal antibody ME36.1. Gangliosides are a particular class of cell surface glycolipids, and GD2 is an important marker for the progression of melanoma. Normal melanocytes, which are pigment-forming skin cells, express ganglioside GM3 most prominently. In melanoma, as the transformation to a cancerous cell begins, there is a change in the distribution and type of gangliosides on the cell surface. Thus, the progression of the disease is marked by the appearance of certain altered gangliosides (GD2, 9-O acetylated GD3, GD3 and GM2).[16] Monoclonal antibodies, such as ME36.1, can be used in immunotherapy against melanoma to identify and eliminate cancerous cells, and thus prevent its further spread.[17] To characterize the interaction of ME36.1 with GD2, the crystal structure of its Fab fragment was determined in this laboratory.[16] As GD2 could not be identified in the electron density for crystals of a putative complex, the Fab/GD2 complex was modeled manually.[16] It was thus of interest to use DARWIN to evaluate this model analytically.

METHODS

DARWIN

DARWIN combines a genetic algorithm with a gradient minimization search strategy. The GA was chosen for DARWIN as there were indications from other fields, such as aircraft design[18] and telecommunications network routing,[19] that GAs performed well in multi-variable searches within large solution spaces. At the outset of the project, GA approaches to docking had not been reported. Since then, a GA has been added to the popular program DOCK, and GAs have become the basis for at least two new docking programs.[20–22]

DARWIN uses a "simple" GA as described by Judson et al.,[22] but with a parallelized evaluation loop to provide the computational speed required to tackle large docking problems. GAs are loosely modeled on biological evolution, and are characterized by the unique manner in which the variables determining possible solutions are represented and altered. Each solution is represented by a binary "chromosome" containing the variables. A population of solutions is randomly generated and the individual chromosomes are evaluated and given a "fitness" ranking. The most fit chromosomes "survive" and continue into the next round or "generation," and the least fit chromosomes are deleted. Superior chromosomes become "parents" of the next generation. New chromosomes are created from the parent solutions by operations analogous to genetic mutations and crossovers. The new generation is evaluated and the cycle repeats until a specified number of generations is completed (Fig. 1).

At the start of a calculation, DARWIN reads a file containing the parameters that control how the GA will function. These can be customized to meet the needs of different docking problems. Two important parameters are the population's size and the number of generations to run. Their product determines the total number of evaluations to be performed. As the evaluation function is the rate-limiting step, this number is important as it determines how much computation time is needed for the calculation. As the population should be as large as practical, a single evaluation is first timed and used to decide the population size and the number of generations to run. Experience with DARWIN, as well as other studies,[20,23] show that population sizes from 100 to 1,000 chromosomes are good starting points for docking. The number of generations needed for convergence was usually between 5 and 10% of the population size.

The following default values were found useful for DARWIN's breeding parameters: mutation rate (20%), crossover rate (40%), and double-crossover rate (40%). The breeding parameters control the relative frequency that each genetic operator is used when creating new chromosomes. The defaults provided a good mix of small movements through solution space by mutations, and larger movements by crossovers. For very small solution spaces, the mutation rate can be increased to 40% and the crossover rates reduced to 30%. This permits more small mutational alterations to promising solutions before a crossover causes a jump to a new region of solution space. The "survival" rate, which defines the number of chromosome copied directly to the next generation, is DARWIN's method for assuring that good solutions are not lost in the new generation. Early trials showed that increasing the survival rate reduced the convergence rate, and allowed DARWIN to make a more complete search. However, survival rates over 30% often required many more generations for convergence because fewer new chromosomes were created in each generation. In contrast, expanding the population size had little effect when the survival rate was below ~10%, as large numbers of good solutions were lost at each new generation. The default rate of 25% was suitable for most searches. The "death" rate is the threshold below which solutions are deleted and not used to create new chromosomes; the default of 5% worked well.

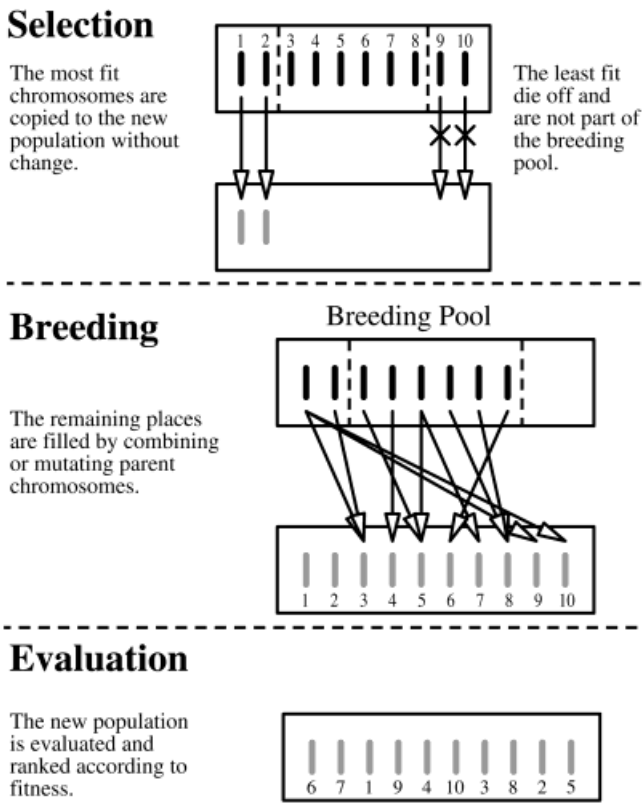


Fig. 1. DARWIN's genetic algorithm. A population of ten possible solutions is randomly generated, evaluated, and sorted by fitness (1–10). In the selection process, the most fit solutions (1, 2) survive and are copied to the next generation, the least fit (9, 10) die and are deleted. The remaining chromosomes are used as a breeding pool to create eight new solutions, using crossovers and mutations. The new population is then evaluated and sorted by fitness so that the cycle can be repeated for the next generation.

CHARMM

The well-known molecular mechanics program CHARMM is used to provide DARWIN's evaluation function. The advantage of this approach is that CHARMM is driven by a command language, and so the evaluation function can be customized readily to meet the requirements of a particular docking target. Additional benefits of using CHARMM are that it has been widely used for modeling a variety of proteins and small molecules, and it is being continuously developed. The modularity of DARWIN ensures that the updates and improvements made available through the CHARMM user group can immediately be incorporated into the program. CHARMM provides DARWIN with various coordinate manipulation, energy calculation, and analysis functions including: manipulation of ligand coordinates; calculation of root-mean-square-deviations (RMSD) for comparing sets of atomic coordinates; potential energy calculations; surface area calculations; coordinate adjustment by gradient energy minimization; and molecular dynamics. Using CHARMM syntax, the user can program DARWIN to use any property that can be calculated by CHARMM to evaluate the molecular system described by a particular chromosome.

The standard amino acids are well studied, and the CHARMM "all atom" amino acid topology and parameter files[24] were used without modification. The CHARMM start-up command sequences, and the commands to calculate energy for each chromosome, for each of the three molecules studied are available as supplementary material. As parameters for the carbohydrates were not readily available, parameterization was completed using information from three sources: the Cambridge Structure Database[25] (for bond lengths and angles); the Harvard CHARMM parameters for glucose[26] (for angle and torsion constants); and ab initio calculations were performed using GAMESS[27] quantum chemistry software (for ligand atom partial charge information). The GAMESS input files used for the partial charge calculations, and the CHARMM parameters defining the sugars, are also available as supplementary material. For energy calculations, the non-bonded interaction list was kept to 12 Å (using the CHARMM keyword CUTNB) and updated every 50 steps. Non-bonded interactions were truncated at 10 Å (CTOFNB). Electrostatic forces were calculated, using a constant dielectric (CDIE) of 1.0, with grouped force switching[28] over the range of 8 Å (CTONNB) to 10 Å (CTOFNB).

In testing DARWIN, the system's potential energy, used to rank each generation's chromosomes, was calculated using several different methods. Portions of the energy calculations were varied in the different tests, but the evaluation always began with a gradient energy minimization of the ligand's atomic coordinates. This aspect of the method is important and has significant consequences, as local minimization within the overall GA search significantly reduces the solution space that the GA must explore. Thus, a DARWIN chromosome does not describe the final position of the ligand, for which the energy is returned by CHARMM. Instead, the chromosome codes a starting point, from which the ligand coordinates are locally optimized using the gradient energy minimization. Local minimization corrects small overlaps of van der Waals radii, rotates bonds to bring groups into good hydrogen bonding geometry, and can reposition the ligand slightly to improve other interactions. This approach relieves DARWIN from the task of making many small adjustments to a promising position using the genetic algorithm. For simplicity, the docking experiments described here held the protein fixed and minimized only the ligand. While this assisted the experiments by speeding the calculations, it is not an intrinsic limitation of DARWIN and the method is easily expanded to include minimization of the protein during an evaluation.

Solvent Modeling

Initial experiments showed that docking a target and ligand in a vacuum, i.e., without including water molecules, did not always yield the correct solution. A likely explanation is that without water, the electrostatic portion of the evaluation was inaccurate and so gave misleading results. A possible approach using CHARMM was to use its optional hydrophobic energy term, which uses atomic solvation parameters (ASPs) to estimate the free energy

changes due to desolvation.[29] This approach was not used because ASPs for the three ligands under study were not available; moreover, there are disagreements between the ASP sets available for proteins.[30] Instead, the system's electrostatic potential energy, which includes the hydrophobic and charged effects of bulk solvent, was modeled with the program "Qnifft" (Dr. Kim Sharp, personal communication). Qnifft calculates a system's electrostatic potential energy from finite difference solutions to the Poisson-Boltzmann equation, and is an outgrowth of the "DelPhi" program[31–33] that has been utilized in other docking protocols.[32,34] When given atomic coordinates and charges, Qnifft returns the electrostatic potential energy; a sum of Coulombic and solvation terms. Because Qnifft does not calculate bonded or van der Waals energy terms, and has no means to perform local energy minimization, it was used only with CHARMM and not alone. To find the ligand/target interaction energy, three Qnifft energy calculations are made. First, the energies of the target alone and the ligand alone are calculated, as these calculations are only made once. Then, the energy for each complex is calculated, and the energies from the target and ligand are subtracted out, leaving the interaction energy. For all the Qnifft calculations, a 65^3 grid was used with a minimum solvent border of 15 Å, which provided scales of 0.85, 0.62, and 0.73 grid/Å for ConA, Se155-4, and ME36.1, respectively. The atoms were charged using the same partial charge values as used for the CHARMM calculations. The boundary conditions at the edge of the box were computed using the dipolar option, with a solvent dielectric of 80 and an interior dielectric of 4. Salt effects were calculated for an ionic strength of 0.145 M, using an ion-exclusion radius of 2 Å. The solution to the non-linear Poisson-Boltzmann equation was assumed to have reached convergence when the residual dropped below 1×10^{-3} (the default value).

Parallelization

Genetic algorithms are inherently parallel. The evaluation of one chromosome does not affect the evaluation of the next chromosome, and thus two or more evaluations can be carried out simultaneously. By implementing DARWIN in a parallel architecture, multiple central processing units (CPUs) can work together to speed up the chromosome evaluation loop and so reduce the real time spent on a single DARWIN run. DARWIN was developed on a local network of laboratory workstations using the freely available "Parallel Virtual Machine" (PVM) software.[35] PVM provided the routines needed for message passing, data conversion, task spawning, and other required elements for parallel execution. PVM is supported at most supercomputing centers, and its use allowed a transparent transition of DARWIN from the laboratory to the massively parallel SP/2 computer at the Cornell Theory Center, where as many as 128 CPUs were utilized.

DARWIN solves molecular docking problems by using CHARMM as an evaluation function. Conventionally, this would be accomplished by making CHARMM a program subroutine, but this would have been cumbersome and required considerable modifications to the CHARMM source code. A superior method was devised that relies on the UNIX operating system's ability to pass data between programs (Fig. 2A). DARWIN was designed to consist of a master process containing the GA, and slave processes directing the chromosome evaluations. When DARWIN is run, the master process creates multiple slaves that each start a CHARMM process with the UNIX "pipe" system call, which is a UNIX feature for passing data from one program to another. Each slave then reads the commands needed to initialize CHARMM from a user file and sends them "down the pipe" to CHARMM. Once initialized, the slaves enter a loop where they receive a chromosome from the master, evaluate it, and return the evaluation to the master. The slaves evaluate chromosomes by sending the appropriate commands to CHARMM, which write the results in a file for retrieval by the slave. This method allowed an efficient scheme of dynamic load balancing to be used, in which fast and slow computers can be used together without wasting processing time (Fig. 2B). With DARWIN and CHARMM linked in this manner, DARWIN can optimize any function that CHARMM can evaluate.

Docking With DARWIN

The basic strategy for docking with DARWIN is straightforward and can be broken into three steps:

1. Molecular Definition. Define the two molecules to be docked (a target protein and carbohydrate ligand in the test cases) and generate the required coordinate, structure, and parameter files for CHARMM. This is a standard procedure for any investigation using CHARMM and general studies of proteins using this program are well documented.[3]
2. DARWIN Setup. Choose GA parameters and a CHARMM evaluation function for DARWIN to optimize.
3. Docking Run and Evaluation of Possible Solutions. Run DARWIN and examine the resulting minimum energy structures for clusters indicating possible solutions. These may be further evaluated by their agreement with any available experimental data.

The task of evaluating the correctness of the potential solutions is straightforward for known structures, as in the two test cases. Here, performance of the docking program can be measured by a "discrepancy plot" in which energy is plotted against the RMSD of the ligand atomic positions from those in the known structure. For unknown cases, it is harder to evaluate progress, and two other plots were devised. "Molecular cluster" plots (similar to the cloud of possible solutions derived from nuclear magnetic resonance spectroscopy) were very helpful in identifying regions of solution space containing the top scoring solutions. These highlight local minima determined by DARWIN and so can be used to narrow future searches. It is important to consider regions where good solutions frequently appear as indicative of potential binding sites rather than concentrating solely on the single best scoring structure. The molecular cluster plots are supplemented

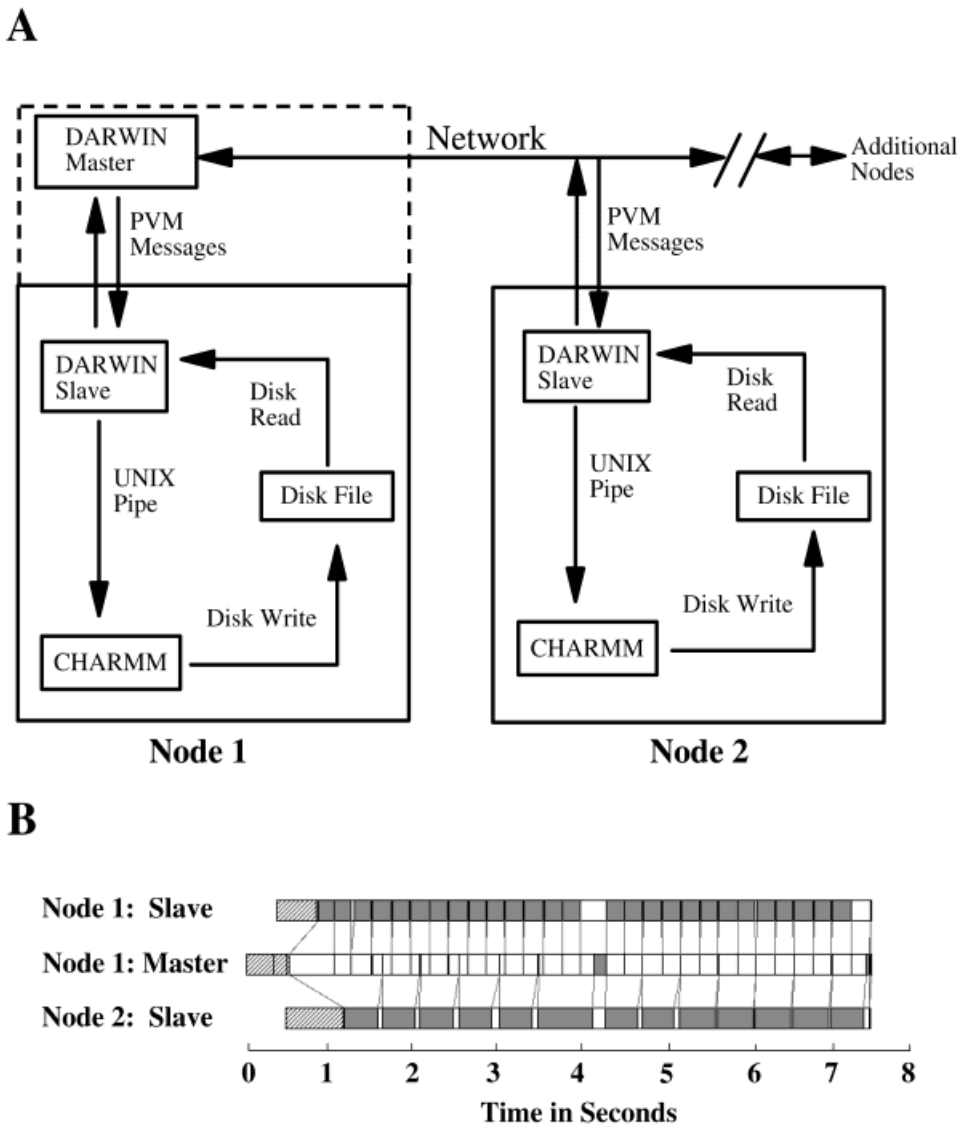


Fig. 2. Parallel docking. **A:** Diagram of DARWIN's data flow. The diagram illustrates how the GA is linked to CHARMM and depicts two nodes of a virtual machine connected by a network. The DARWIN master generates chromosomes with the GA and sends them to the slaves as PVM messages for evaluation. The slaves pipe commands to CHARMM, retrieve the evaluations from a disk file, and send them back to the master using PVM. **B:** Dynamic load balancing. Multiple slaves cannot be counted on to evaluate their chromosomes in equal times, as nodes may have different speeds or loads. Therefore, load balancing is required to avoid idle time while the master waits for a slow evaluation. Load balancing is visualized in this time line for a docking run on two nodes. Each horizontal bar represents a process (master or slave). The bars are shaded to depict the activity of the node. The crosshatched regions represent overhead time for setting up the process, the shaded regions computing time, and the white regions idle time while waiting for an incoming message. Lines between processes represent messages. In this run, 2 generations of 20 chromosomes were evaluated. Of the 40 total evaluations, Node 1 performed 27 and the slower Node 2 performed 13.

by simple "energy plots" of the individual variables on the chromosome to determine whether each variable has converged. Again, this information is useful in narrowing future searches. These two plots are described further in Results.

In general, it was found best to begin a search with as few restrictions as possible. After a feel for the problem has been obtained, searches can be narrowed by several methods. Often, some experimental data is available on the location of the binding site, or mutational data has implicated a particular residue as critical for binding the substrate. Clusters of solutions that do not agree with available experimental data can be removed by excluding particular ligand orientations, or some portions of the target's surface. By narrowing the search area, DARWIN can concentrate on optimizing binding at a particular location without being distracted by competing binding sites. This approach assists in reducing the possible solutions to a few probable sites or orientations.

RESULTS

To test DARWIN and gain experience in docking carbohydrates before tackling the unknown ME36.1 structure, two simpler protein-carbohydrate complexes with known crystallographic structures were investigated. The reproduction of a known configuration is a logical first step towards predicting the structures of unknown complexes, because it is verifiable. The test cases chosen were Concanavalin A[36] and Fab Se155-4.[37] ConA was chosen for its well-defined structure (the complex with the carbohydrate was solved at 2.0 Å resolution) and its relatively simple ligand containing only 27 atoms (Fig. 3A). The structure of antibody fragment Fab Se155-4 is also known at high resolution (2.0 Å resolution), but the ligand is a trisaccharide. The three pyranose rings, with 64 total atoms, and the flexibility from the glycosidic ring linkages between rings (Fig. 3B), provided a greater docking challenge than the ConA test.

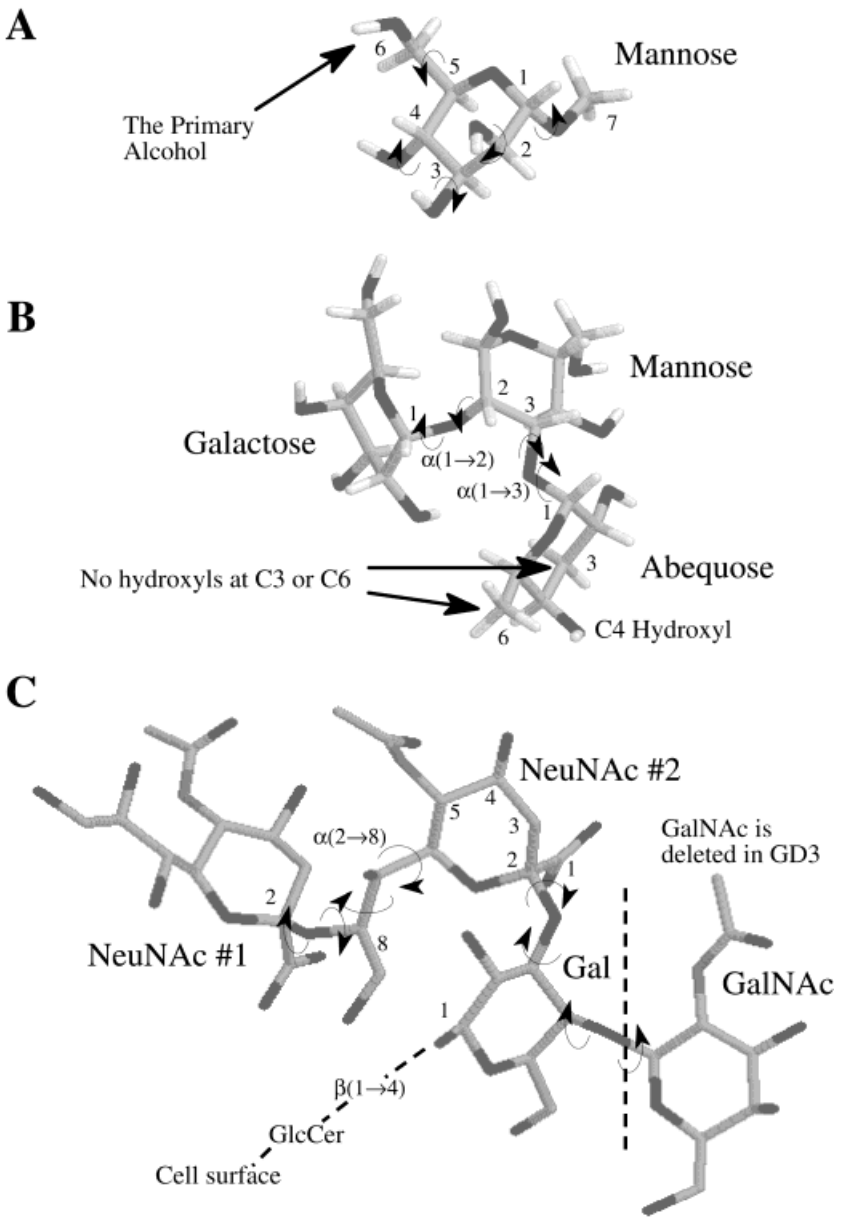


Fig. 3. The carbohydrate ligands. Carbon atoms are numbered and rotatable bonds included in the search are indicated (arrows). **A:** The ConA ligand, methyl α-D-mannopyranoside. **B:** The Se155-4 antigen, trisaccharide α-D-Galactose-[α-D-Abequose]-α-D-Mannose. The abequose residue is particularly hydrophobic as it is missing the usual hydroxyl groups at the C3 and C6 positions. **C:** The ME36.1 antigen, ganglioside GD2, has an unusual α(2→8) linkage between two sialic acid residues (NeuNAc), which are linked by galactose (Gal) to N-acetylgalactosamine (GalNAc). The GD3 ganglioside is identical, except that the GalNAc residue is deleted. The four-sugar chain is linked to the cell surface by a β(1→4) linkage from the galactose residue to a glucose-ceramide (GlcCer) residue in the cell membrane. To maintain this link, the rotation of the Gal residue was restricted in the docking search so that Gal C1 remained solvent accessible.

Concanavalin A

Atomic coordinates for the complex of methyl α-D-mannopyranoside and ConA were obtained from the Protein Data Bank (PDB no. 5CNA). ConA is a dimer or tetramer in solution, depending on pH, and crystallizes as a tetramer.[36] There is one carbohydrate binding site per subunit, located at the ends of the molecule far from the subunit interfaces. Since the sites behave independently, and binding is not cooperative,[38] only one of the ConA monomers was considered in the docking experiments. This reduced the number of atoms in the target molecule and the calculation time. The complex structure also contains two bound metal ions (calcium and magnesium) per monomer, which were included in the target. In addition, there are ~130 ordered water molecules bound to each ConA monomer. To investigate the effect of water on docking, the tests were performed both with and without these water molecules. As hydrogen atoms are not defined by the crystal structure, they were added using CHARMM's HBUILD command.[39] Then, while all other atoms were kept fixed in position, the hydrogen positions were energy minimized to relax atomic overlaps and create the model of a ConA subunit.

Both global and local searches were attempted for ConA. In solution, a ligand is presented with the entire surface of a protein and somehow finds its way to the correct binding site. As this is the most general docking problem, it is more difficult than a search restricted to the volume around a known binding site and is likely to reveal any deficiencies in the evaluation function used to evaluate potential solutions.

TABLE I. ConA Docking Results

Complex	Solvent	Search space	Evaluation function	Lowest CHARMM energy (Kcal/mol)
Crystal structure	Water		Bonded + VDW + Electrostatics	24.6
Crystal structure	None		Bonded + VDW + Electrostatics	34.1
Crystal structure	None		Bonded + VDW	−8.9
Mannopyranoside	Water	Global	Bonded + VDW + Electrostatics	19.5
Mannopyranoside	None	Global	Bonded + VDW + Electrostatics	0.6
Mannopyranoside	None	Binding site	Bonded + VDW + Electrostatics	21.8
Mannopyranoside	None	Global	Bonded + VDW	−16.9
Mannopyranoside	Qnifft			
All 8 pyranosides	Water	Binding site	Bonded + VDW + Electrostatics	19.2[a]

[a]α-D-mannopyranoside.

The requirements were less stringent in most reported ligand docking experiments, as these concentrated on searching the local area of a binding site. This is a practical approach for drug design applications because the binding site is nearly always known, and the question is then "how do other ligand analogs bind in the receptor site?"

For docking, DARWIN uses the GA to generate a chromosome describing the starting orientation and position of the ligand, and the disposition of the rotatable bonds. The atomic positions, including those for hydrogen, are then varied within a local minimization step in which CHARMM minimizes the energy of the system. First, the ligand's atomic positions were optimized using adopted basis Newton-Raphson (ABNR) energy minimization with a time step of 50 fs. Minimization proceeded until the energy change was less than 0.002 Kcal, or a maximum of 320 steps. During the minimization, the target's atoms were held fixed and all of the ligand's atoms were free to move. The resultant potential energy then was returned.

For ConA, the global search space encompassed a 50 × 50 × 40 Å box centered around the subunit, making its entire surface available to the ligand. The local search space encompassed a 15 Å cube centered on the known binding site. Each search space was divided into a grid with 0.5 Å steps. Thus, the global box contained 8×10^5 discrete points. It should be noted that a 0.5 Å grid size is coarse when compared with the values of 0.016–0.08 Å used in previous studies.[23] DARWIN compensates for the loss of accuracy arising from such a coarse grid size by the local minimization step in its evaluation function, which can move atoms in steps as small as 0.001 Å. This method was economical, as it reduced the search space, while it maintained high accuracy by allowing local minimization to explore the region between grid points. The three angles describing the orientation of the ligand were represented on the GA chromosomes with integer values from 0 to 31 (11.25° steps), and each of the five rotatable bonds with integers from 0 to 5 (60° steps). As before, the rotational step size was coarse when compared to other GA docking experiments (steps as small as 0.35° were reported by Judson et al., 1995[22]). But with DARWIN, the local minimization makes the fine adjustments needed to give a final precision of ~0.02°.

Evaluating the Results for ConA

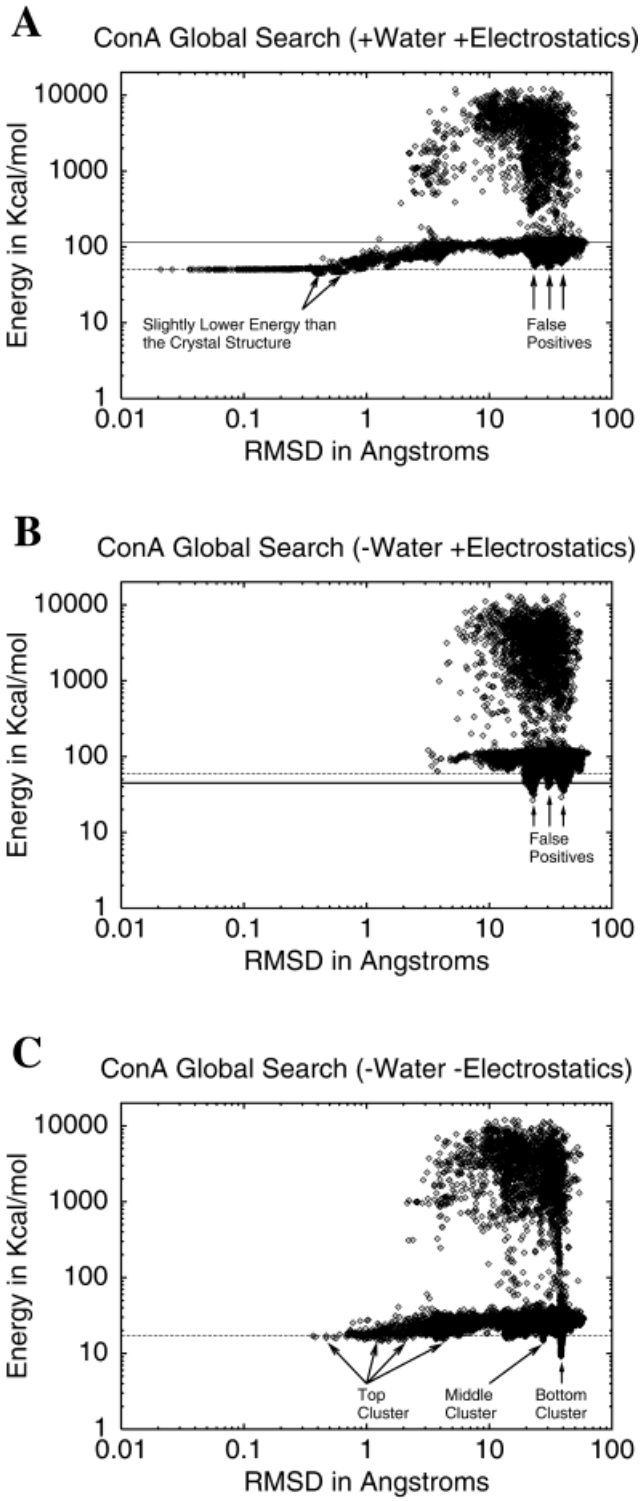
The results of the first experiment (Table I) were evaluated from discrepancy plots (Fig. 4). These face the minor difficulty that calculated potential energies have a wide range and include negative values. Thus, the origin is always shifted to make all energies positive and so allow use of a logarithmic scale. Here, the internal energy for the ligand alone is 116 Kcal/mol, after an origin shift of +26 Kcal/mol. This provides a base line for interpreting the first discrepancy plot (Fig. 4A), which shows two major groups of structures: those that fall above the line, representing structures where the ligand has an unfavorable interaction with the protein; and those falling below the line, representing structures with favorable interactions. Obviously, the latter group contains the structures of interest.

Discrepancy plots must be examined with the following three questions in mind:

1. *Are solutions found at or very near the corresponding crystal structure?* For ConA, DARWIN found solutions very near the correct structure, several with less than 0.1 Å RMSD. If the search had not found this region, it would have indicated either a problem with the search itself, such as premature convergence, or that the CHARMM evaluation incorrectly assigned poor scores to good structures ("false negatives") or good scores to poor structures ("false positives"). Where the correct solution is unknown, premature convergence is difficult to identify. Possible indications are nonsense solutions, or docking runs converging on different solutions. If this occurs, the population size or survival rate should be increased.
2. *Are solutions found that are close to the correct structure but have high energies?* Such false negatives indicate problems with the evaluation function and show that DARWIN cannot converge on the correct solution. False negatives are very common when local minimization is not performed as part of the evaluation because a small misplacement of the ligand can cause van der Waals collisions, with a large rise in energy. False negatives of this sort are not a significant problem for Monte Carlo

type docking programs that walk away from the current position in small steps, but they can be devastating in the GA approach. Without local minimization, a nearly perfect chromosome could be lost if the evaluation function returns a high energy due to overlap, causing the GA to omit that region of space in subsequent generations. A warning sign would be a few solutions with high energy near the correct structure. For ConA, false negatives were not found.

3. *Are there solutions far from the correct structure, but with lower energies?* False positives also indicate problems with the evaluation function. In the crystal, the ligand is observed at only one site per subunit. Although it is possible that binding could occur at low occupancy elsewhere on the protein, the observed binding site clearly has the highest specific binding affinity, and the resultant complex should have the lowest energy. In the first ConA experiment, which included the water molecules, the energy of the correct complex structure was 24.6 Kcal/mol (Table I, Fig. 4A). Although DARWIN found a few complexes with slightly lower energy (19.5 Kcal/mol), these were all very close to the crystal structure (between 0.4 and 0.6 Å RMSD). When displayed on a graphics workstation, it appeared that the orientations of several hydroxyl groups on the ligand had altered to create additional hydrogen bonds. Because these solutions were so nearly correct, they were not considered misleading.



The discrepancy plot (Fig. 4A) shows that all solutions between 2 and 20 Å RMSD have an energy that is at least 20 Kcal/mol above that for the correct complex structure. This indicates that the ligand lies in an energy well in the crystal structure, as would be expected for a binding site. However, several groups of structures that are even further from the crystal structure (20 and 40 Å RMSD) have energies that are only 8 Kcal/mol higher. These solutions might be false positives, or could accurately represent other potential binding sites on the target's surface. They are explored in greater detail in the following section.

Effect of Water on Concanavalin A Docking

In the ConA crystal structure, several water molecules in and around the binding pocket assist in coordinating the ligand through hydrogen bonds. As the locations of water molecules are not usually available to the modeler, it would be extremely helpful if docking problems could be solved without this knowledge. To test DARWIN's performance in a more realistic setting than the first experiment, a docking run was performed in which the water molecules were removed from the target. Disconcertingly, the results

Fig. 4. Discrepancy plots of the ConA global search results. For each potential solution returned by DARWIN, the RMSD of the solution's atomic positions from those in the crystal structure are plotted against potential energy. The energy values (Table I) have been origin shifted by +26 Kcal/mol. **A:** The target included experimentally observed water molecules. The internal energy of the ligand alone (116 Kcal/mol) is shown by the solid line, and the energy for the correct complex by the dashed line. **B:** The target omitted experimentally observed water molecules. The energy for the correct complex is shown by a dashed line. Structures below the heavy solid line are drawn in a molecular cluster plot (Fig. 5A). **C:** The target omitted experimentally observed water molecules, and the electrostatic term was omitted from the evaluation function. The dashed line marks the energy for the correct complex. When solutions below this line are drawn in a molecular cluster plot (Fig. 5B), they form three main clusters (arrows).

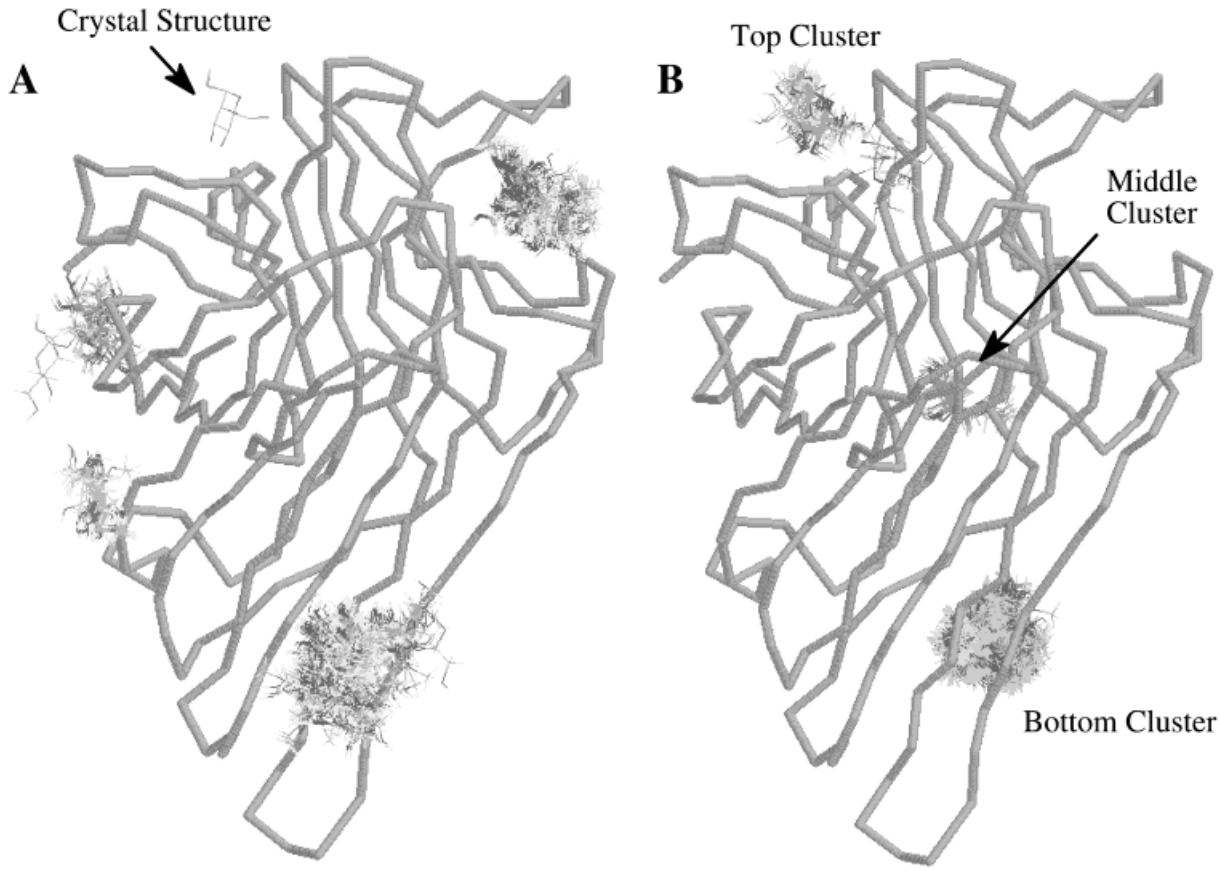


Fig. 5. Molecular cluster plots for ConA. These demonstrate the differing results from inclusion or omission of the electrostatics term from the evaluation function. **A:** All the ligand structures with energy less than 20 Kcal/mol that were found in the global search omitting water from the target (Fig. 4B, below the heavy solid line) are displayed on the ConA backbone. The position of the ligand in the crystal structure is indicated, although it was not sampled in this search. The lowest energy structure lies in the cluster at the upper right. **B:** Ligand structures resulting from a search conducted as in A, except that the electrostatics term was omitted. The 420 solutions at, or below, the energy of the correct complex (Fig. 4C, below the dashed line) are plotted on the ConA backbone. All of the clusters in A, including the polar site at the upper right, have disappeared and are replaced by three new clusters. The crystal structure lies in the top cluster, and the lowest energy structure lies in the new cluster at the bottom. As this site lies on the dimer interface, it would be blocked with a dimer target.

were quite different (Table I). Many false positives were found (Fig. 4B), and the search failed to find structures within 3 Å RMSD of the correct solution. The false positives corresponded to groups at the same three locations (~20, 30, 40 Å) found in the previous experiment (Fig. 4A).

The failure to find solutions closer than 3 Å RMSD to the correct ConA structure could be due to several reasons. The search may have converged prematurely without finding the true global minimum. On the other hand, a problematical energy function could have assigned artificially good (low) values to structures far from the correct solution, so that the true energy minimum was not explored. To determine which was the case, DARWIN was run again with the search space reduced from a global search of the entire target surface to a local search around the binding site. This forced DARWIN to search the area of the correct binding site more thoroughly by testing many variations of the ligand structure, to find any low-energy structures that were missed in the global search. This local search area did not include any of the target's surface regions corresponding to the locations of the false positives in Figure 4B. The best energy found in the local search was 21.8 Kcal/mol (Table I). This is similar to the energy of the correct complex, but well above the energy of the best structures found in the global search. As low-energy structures at the correct site could not be found, it could be concluded that the problem lay with the evaluation function rather than with premature convergence.

Molecular cluster plots (Fig. 5), in which the structures of the best scoring ligands are superimposed on the α-carbon backbone of the target molecule, were very helpful in visualizing the various possible solutions identified from the discrepancy plots (Fig. 4). A problem with the discrepancy plots is that solutions with similar deviations do not necessarily lie at the same location on the target's surface. The molecular cluster plot shows where the solutions fall on the protein surface and so can reveal sticky spots that may be alternative binding sites. Figure 5A shows the ligand positions for those ConA complexes that had an energy less than 20 Kcal/mol in the global search without water (Fig. 4B). The correct solution is also

shown, although it was not sampled during the search. The ligand positions with lowest energy cluster at a highly polar region of ConA (upper right in Fig. 5A). In the crystal structure, this region contains several ordered water molecules that would presumably block the ligand from binding at this location.

Docking by Shape Complementarity Alone

It was hypothesized that the difficulties with water could be overcome by reducing the contribution of electrostatic energy to the evaluation. This would place greater weight on the complementarity of the ligand's shape to that of its binding site. In the next ConA test, the electrostatic term in the CHARMM energy function was deleted to assess whether this omission would compensate for omitting water in the complex. This forced DARWIN to search using only the bonded and van der Waals energies in the evaluation function, thus emphasizing the similarity of the interface surfaces. It was tempting also to exclude the electrostatic term from the local minimization, to save its high computational cost. This was not done as electrostatic forces are very important for correctly aligning a ligand's hydrogen bond donors and acceptors with those in the binding site. As these act at comparatively long range, they play a greater role than the other energy terms in moving a ligand into contact with the target's surface during minimization.

The global search was performed again, omitting water, and with the electrostatic portion of the energy omitted from the evaluation (Table I, Fig. 4C). The results improved dramatically and, in contrast to the earlier search, many low-energy structures lay within 1 Å of the correct solution. However, although the false positives found previously were all eliminated, a new set of false positives was located more than 10 Å from the correct structure. A molecular cluster plot (Fig. 5B) of the 420 structures found at, or below, the energy of the correct complex (Fig. 4C) shows three distinct local minima: A cluster at the correct site (top); a small cluster ~20 Å from the correct site (middle); and a larger cluster ~40 Å from the correct site (bottom). Although the latter contained the structure with the lowest energy, its location in the dimer interface would not be available if the search were performed with the complete molecule. It represented, perhaps, not a false positive so much as an artifact caused by considering only the ConA monomer as the target. If this site was ignored, the most favored site was at the correct location.

Modeling Solvent Effects

The ConA experiments showed that the evaluation function could be misleading when solvent was not included in the target. To investigate if this problem could be overcome by including a model for bulk solvent, without explicitly including water molecules in the system, the program Qnifft was used. Qnifft calculates the electrostatic potential energy for a system from its atomic coordinates and charges. The 420 top scoring structures from the global search without waters and the electrostatic term (Fig. 5B) were evaluated by Qnifft (Fig. 6). The results

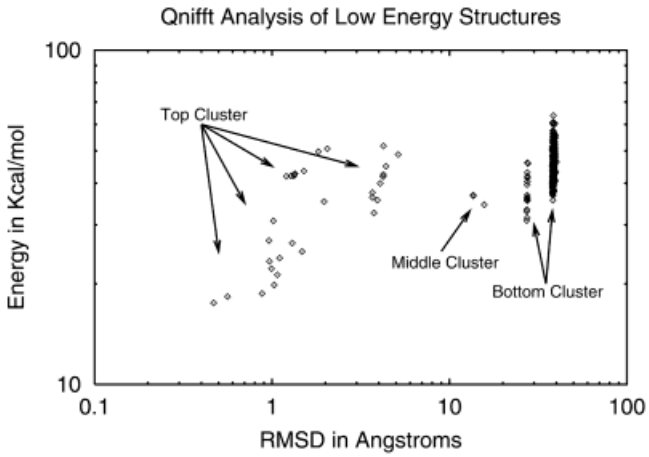


Fig. 6. Qnifft results for ConA. Qnifft was used to calculate ConA/ligand interaction energies for the 420 structures in the three clusters shown in Figure 5B. The energy values have been origin shifted. The solutions nearest to the crystal structure (Top Cluster) have the lowest energy, and those in the dimer interface (Bottom Cluster) have poor energies.

were encouraging as the ten best scoring structures were now all <2 Å RMSD from the crystal structure, and the closest (RMSD <0.5 Å) had the lowest energy. Additionally, the solutions that fell in the dimer interface had some of the highest energies, indicating that it is not a reasonable binding site.

Docking With Multiple Ligands

An experiment was designed to test whether DARWIN could select between multiple ligands for the target, an ability that could be used for designing drugs. The binding site was searched with the eight possible α-D-pyranosides (ido-, gulo-, talo-, galacto-, altro-, allo-, manno-, and glucopyranoside). DARWIN's flexibility facilitated this task as the new experiment merely required adding three variables to the chromosome. These describe the hydroxyl groups at carbons 2, 3, and 4 as either axial or equatorial, and so define all eight sugars. The variables directed CHARMM to modify the ligand via the "PATCH" command before it performed the local energy minimization. With this modification, mutations and crossovers not only reoriented the ligand, but also changed the ligand's isomeric state to that of a different sugar. Thus, the selection process driving the DARWIN search to the correct ligand binding conformation, could simultaneously drive the chemical structure of the ligand to fit the available binding sites. Of the eight possible sugars, the mannopyranoside ligand had the best energy score (Fig. 7). Altropyranoside and glucopyranoside were the next best, and galactopyranoside was the worst.

Antibody Fragment Se155-4

The monoclonal antibody Se155-4, which recognizes a complex carbohydrate, has been used by several groups as a model system. Its advantages include a well-refined crystal structure (2.0 Å resolution),[40] and a detailed

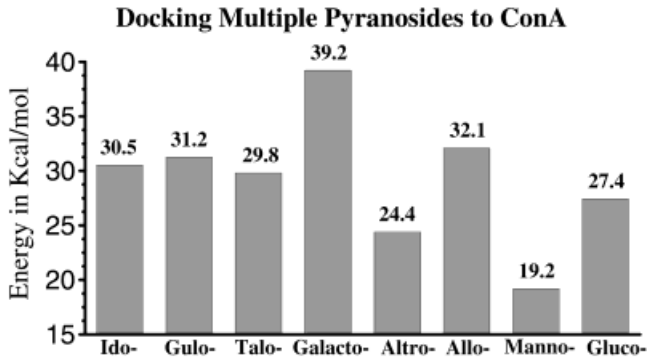


Fig. 7. Selection of structurally-related potential ligands by DARWIN. Eight sugars were simultaneously docked to ConA. For each ligand, the lowest energy found by DARWIN is shown.

description of its binding thermodynamics.[41] The crystal structure of the Se155-4 Fab complexed with its dodecasaccharide antigen was solved at 2.0 Å resolution.[40] In this structure, the three terminal sugar residues galactose, mannose, and abequose (3,6-dideoxy-mannose) lie in well-defined electron density at the complementarity-determining region (CDR) of Se155-4. However, the remaining nine residues were poorly defined. A structure at 1.7 Å resolution[37] was obtained by engineering a single-chain variant of the Fab bound to a shorter trisaccharide, α-D-Gal(1→2)-[α-D-Abe(1→3)]-α-D-Man (Fig. 3B), containing the three sugar residues clearly identified in the initial structure. The antigen positions in the two crystal structures differ only slightly, with an RMSD of 0.46 Å (the individual galactose, mannose, and abequose residues have deviations of 0.51, 0.51, and 0.26 Å, respectively). The trisaccharide was used as the ligand in the Se155-4 docking experiments. Its oxygen linkages provided a ligand with greater flexibility than that used in the ConA studies, but less than the ligand that was the ultimate goal, GD2.

Se155-4 has a pocket-shaped binding site, rather than the grooved type seen in other sugar binding Fabs. The bound trisaccharide is bent into an "L" shape with its abequose ring completely buried in the pocket, and its mannose and galactose rings on the surface of the protein partially exposed to solvent. Abequose is an unusually hydrophobic sugar because it lacks hydroxyl groups at the C3 and C6 positions (Fig. 3B) and is buried in a binding pocket with mostly hydrophobic amino acids, including four tryptophans and a phenylalanine. The pocket also contains a water molecule, which assists in binding the abequose residue through hydrogen bonds with the C4 hydroxyl and His 285.

DARWIN and CHARMM were set up as in the ConA experiments. The coordinates from the original 2.0 Å structure of the complex[40] (PDB no. 1MFE) were used, as this structure has greater biological relevance. In any case, the CDR regions are nearly identical (0.34 Å RMSD). Hydrogen atoms were added by CHARMM and, while holding non-hydrogen atoms fixed, the coordinates were energy minimized to relax bad contacts. For the ligand, CHARMM parameters were already available for galactose and mannose[26] and only minor modifications were needed to define abequose. The chromosomes were coded as in the ConA experiments with two changes: variables describing the ϕ and ψ torsion angles of the sugar ring linkages were added; and the variables for the hydroxyl rotations were deleted, as the ConA experiments had indicated that local minimization was sufficient to correctly orient the hydroxyls. As the portion of a Fab molecule that interacts with the antigen is well known, the search area was restricted to a 15 Å cube including the entire CDR.

In the Se155-4 crystal structure, four water molecules are within hydrogen bonding distance of the antigen. Water 601 bonds to the abequose, water 602 bonds to the mannose, and waters 757 and 758 bond to the galactose. The Se155-4 target was searched with, and without, the 184 water molecules. In contrast to the ConA experiments, the presence or absence of water had little effect on the docking results (Table II). When the electrostatic term was omitted in the chromosome evaluations, as had been found most effective for the ConA target without water, DARWIN identified ten low-energy structures close to the experimentally observed binding mode (Fig. 8). The best scoring structure had an RMSD of 2.4 Å. Although this was high, most of the deviation was due to the galactose ring (3.8 Å), while the other two sugar residues were much closer to the correct structure (0.6 and 1.3 Å). Including the electrostatic term in the evaluation function improved the position of the galactose residue, but slightly worsened the positions of the abequose and mannose residues (Table II). For both evaluation functions, water had little effect.

In the crystal structure, the abequose residue is completely buried in the protein and is not accessible to solvent. Molecular cluster plots (Fig. 9) revealed that the best DARWIN solutions all placed the abequose ring in the correct position and orientation. The abequose cluster is very tight, with less than 0.6 Å deviation in the averaged atomic positions. The mannose residue is also in essentially the same position in all the low energy structures. However, in each case a rotation of ~120° about the C5 to C6 bond (Fig. 3B) changed the position of the primary alcohol, when compared to the crystal structure, and this displaced the mannose ring by approximately 0.5 Å. The galactose residue has the greatest positional variation, and shows the largest deviation from the correct structure. The lowest energy structure has its galactose ϕ and ψ linkages twisted by 6° and 43°, respectively, which moves the galactose ring by approximately 3.8 Å from the location in the crystal structure.

Qnifft was used to analyze the cluster of ten best solutions determined by DARWIN (Fig. 8). The Qnifft interaction energies (data not shown) for all the structures were very similar (52.3 ± 2.8 Kcal/mol) and the only outlier (92.1 Kcal/mol) corresponded to a ligand with its galactose residue farther away (4.3 Å RMSD) from its correct placement than those in the other solutions (~3.4 Å). This agreed with the behavior of galactose in the earlier experiments. The Qnifft interaction energy for the crystal structure was 53.8 Kcal/mol.

TABLE II. Se155-4 Docking Results

Complex	Solvent	Evaluation function	Lowest CHARMM energy (Kcal/mol)	Abequose RMSD (Å)	Mannose RMSD (Å)	Galactose RMSD (Å)	Total RMSD (Å)
Crystal structure	None	Bonded + VDW	−29.0				
Crystal structure	Water	Bonded + VDW	−29.7				
Crystal structure	None	Bonded + VDW + Electrostatics	25.5				
Crystal structure	Water	Bonded + VDW + Electrostatics	17.9				
Gal→Man←Abe	None	Bonded + VDW	−26.7	0.6	1.3	3.8	2.4
Gal→Man←Abe	Water	Bonded + VDW	−27.6	0.7	1.2	3.9	2.4
Gal→Man←Abe	None	Bonded + VDW + Electrostatics	17.9	0.8	1.8	1.4	1.4
Gal→Man←Abe	Water	Bonded + VDW + Electrostatics	11.7	0.7	1.8	1.4	1.4
Gal→Man←Abe	Qnifft			1.3	1.8	4.3	2.9

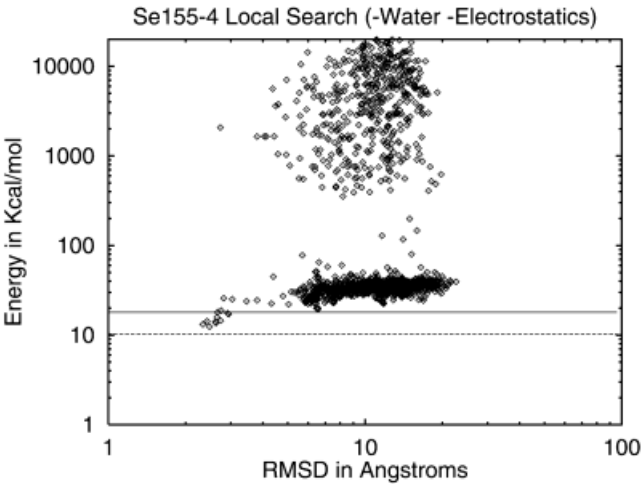


Fig. 8. DARWIN results for target Se155-4. Discrepancy plot for the local search of the CDR region, omitting water from the target and the electrostatics term from the evaluation function. The energy values have been origin shifted by +40 Kcal/mol. The dashed line represents the energy of the minimized PDB coordinates. Ten solutions with low energy (below the solid line) were found with energy near to that of the crystal structure.

Antibody Fragment ME36.1

In the "real world" docking experiment, the carbohydrate portion of the GD2 ganglioside was docked to the Fab fragment of the ME36.1 antibody. The GD2 ganglioside contains a complex oligo-saccharide with five pyranoside residues (Fig. 3C). As the glucose-ceramide residue that anchors GD2 to the cell's outer membrane is unlikely to contact the Fab, it was not included in the modeled structure of the complex (PDB no. 2PSK),[16] and it was omitted from the docking experiment for the same reason. Coordinates for the Fab ME36.1 structure[16] were obtained from the PDB (PDB no. 1PSK) and hydrogen atoms were added by CHARMM. The two waters in the antigen binding site were removed because they would presumably be displaced upon antigen binding. GD2 has an unusual α2→8 linkage between two sialic acid residues (NeuNAc), and a total of eight torsion angles in the linkages between the four rings. Thus, the GA search included the standard six positional variables, plus eight variables for the interring linkage torsion angles. The search space was restricted to a 20 Å cube that included the entire CDR. The search was not confined to match the docking criteria defined by Pichla et al.,[16] except that the orientation of the galactose residue was restricted so that it could not be buried against the protein. This restriction was accomplished by reducing the range of possible rotation angles for the galactose residue to include only those that left the residue exposed. This had the added advantage of reducing the search space by half.

DARWIN and CHARMM were set up as in the previous experiments, and the ME36.1 docking runs were first made with the same evaluation function used for the ConA and Se155-4 targets. The initial searches found only structures that had very few antigen/Fab contacts. This problem arose because GD2 is highly flexible and tended to fold back on itself and form structures with many favorable *intra*-molecular contacts. The result was that docking runs converged to local minima at which the antigen made a tight ball and had few interactions with the Fab. To avoid this problem, the evaluation function was changed to include only *inter*-molecular interaction energy. By removing the *intra*-molecular non-bonded interactions from the fitness score, complexes were searched in which the antigen could only satisfy its bonding requirements with protein contacts. To accomplish this, the ABNR minimization was performed as before, but the intra-molecular energy term was omitted from the total calculated energy (CHARMM's "INTERACTION" command was useful for this purpose). This strategy prevented convergence at structures with many intra-molecular contacts and few antigen/Fab contacts.

The correct structure of the bound antigen was not known in the real ME36.1 case, and, therefore, progress could not be measured by the usual discrepancy plots. Instead, the six positional variables on the chromosome determining the position of the galactose were plotted separately as an indicator of convergence (Fig. 10). This type of plot can be helpful in narrowing future searches by revealing the range of values for a particular variable that creates favorable interaction energies. DARWIN's 50 best antigen structures had a common conformation and position with an average RMSD of 1.3 Å (Table III). A

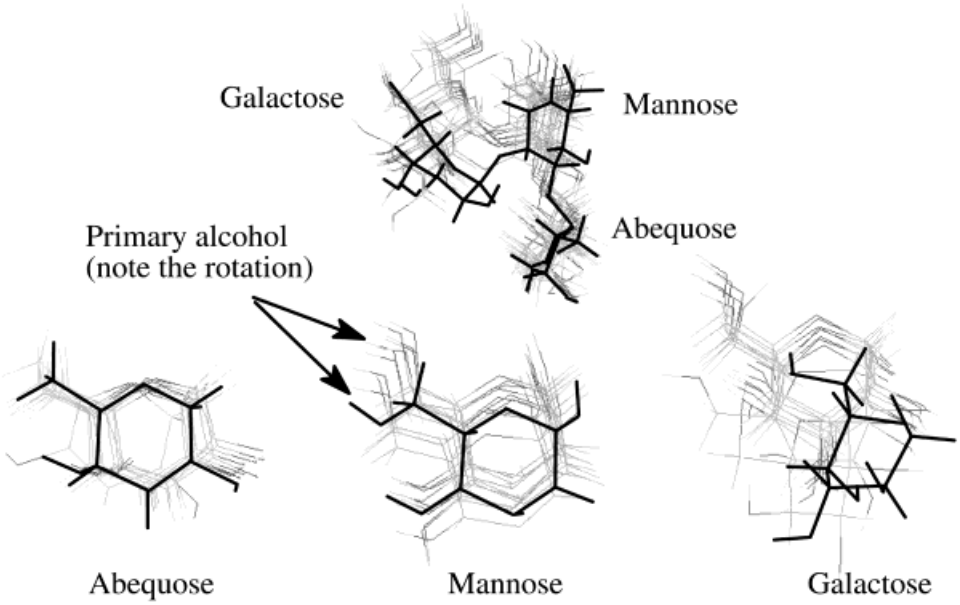


Fig. 9. Se155-4 molecular cluster plot. The ten best scoring structures from Figure 8 are shown (**top**), superimposed on the crystal structure (heavy lines). For clarity, each ring is drawn separately below in an appropriate orientation.

molecular cluster plot (Fig. 11) showed that all 50 structures were very similar except for the location of the GalNAc residue, which had two distinct orientations. The results suggested that there were two variants of one basic solution for GD2.

In all 50 structures for GD2, the two sialic acids were the most tightly defined residues, with all solutions within 1 Å RMSD from the average. The first sialic acid residue was in a similar position and orientation as in the model structure,[16] with several conserved hydrogen bonds. The position of the second sialic acid residue was very different. In the model, the carboxyl group on C2 is oriented away from the protein, and the acetyl group on C5 is oriented toward the heavy chain. The DARWIN solution flipped the ring over, so that the C2 carboxyl group was buried against the protein and the C5 acetyl group was oriented towards the light chain (Fig. 12A,B). The galactose residue had no significant contacts with the protein in either the model structure or the DARWIN solution, and in both cases the mandatory linkage with the cell surface was maintained. In the model, the GalNAc residue partially occupies the location where DARWIN placed the second sialic acid residue. DARWIN moved the GalNAc residue to two alternate positions with approximately equal energy, a "horizontal" position and a "vertical" position (Fig. 12B,C).

To investigate whether solvent would play a role in selecting between the two possible configurations, the antigen/Fab interaction energies for the top 50 solutions were calculated using Qnifft as in the earlier experiments (data not shown). The average interaction energy was 35 ± 7 Kcal/mol compared with an energy for the model structure of 41.4 Kcal/mol. There were four structures with energy more than two deviations lower than average. As these contained both vertical and horizontal structures with equal probability, the Qnifft results suggested that the bimodal distribution was a real effect.

DISCUSSION

DARWIN was developed for large docking problems; those with big solution spaces, or lengthy evaluation schemes, or both. We assumed that substantial computer time would be available, and that accuracy was much more important than speed. CHARMM is highly customizable and many of the evaluation schemes in other docking programs could be replicated by using the appropriate CHARMM parameters. The design of DARWIN is such that improvements made to CHARMM simulation methods are immediately applicable. An example is the planned incorporation of a Poisson-Boltzmann based energy term into CHARMM. This will immediately facilitate additional study of the Qnifft approach to solvent modeling that was helpful in the ConA experiments.

Direct comparisons of DARWIN's docking speed with other GA-based docking protocols on identical test cases were not performed. However, for small ligands and search spaces, it is likely that DARWIN would lag behind other programs, primarily due to the slow local minimization procedure in the evaluation function. DARWIN completed the ConA global search tests in ~19 hours on a single workstation, which is not prohibitively slow. The advantages conferred by DARWIN's parallel processing format were shown by the reduction of the run time for this problem to as little as 40 minutes when using 32 CPUs (90MHz R6000) on the Cornell Theory Center's SP2 computer. Looking ahead, it may be possible to speed up DARWIN's search method by including a feedback loop in the evaluation function to update the chromosome with the ligand's post-minimization position and conformation. Closing the loop in this fashion will create a new GA/gradient-minimizer hybrid search algorithm that should be significantly faster.

The oligosaccharide studies avoided potential issues with induced fit binding modes, in which there are confor-

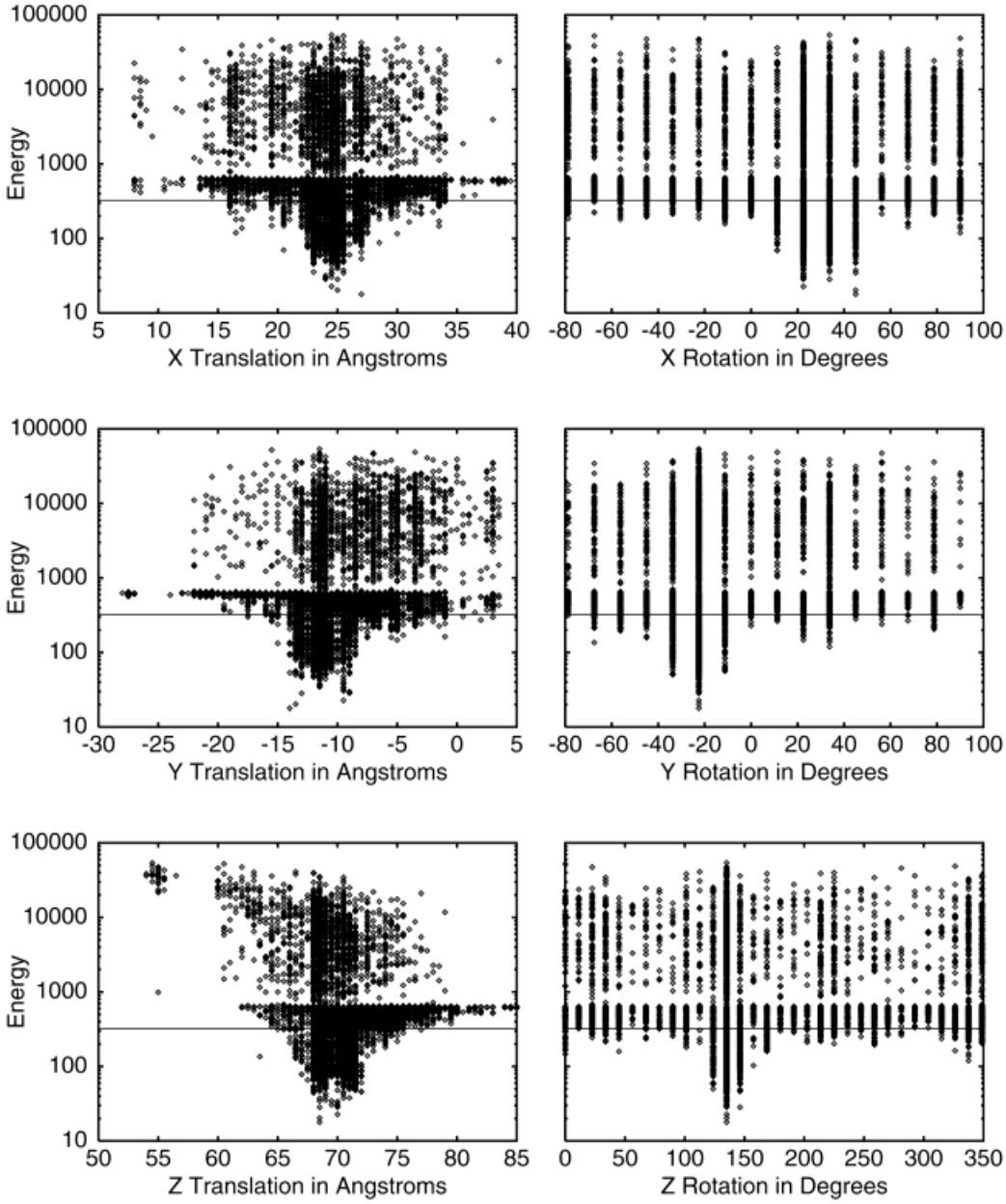


Fig. 10. DARWIN search results for target ME36.1. The six positional variables in the search are shown as a function of energy (other search variables not shown). All energy values have been origin shifted by 560 Kcal/mol. The energy of the model[16] is marked by a solid line.

mational changes in the protein upon ligand binding. A more realistic, and more difficult, target would be the *apo* crystal structures for such proteins. To tackle these targets, it would be necessary to introduce flexibility into the protein molecule. While DARWIN could accomplish these tasks, *apo* structures were not used as it seemed best to begin with relatively simple tests to gain experience with the program. Limited target flexibility could easily be introduced during the local minimization phase of the chromosome evaluation. While computationally expensive, somewhat larger target motions might be successfully modeled in a similar way, with short molecular dynamics integrations.

TABLE III. ME36.1 Docking Results

	RMSD from the average of the top 50 structures (Å)				
	NeuNAc 1	NeuNAc 2	Gal	GalNAc	Total
Top DARWIN structure	0.7	1.3	2.9	4.2	2.1
Model structure	2.2	5.3	4.6	7.1	4.6
All 50 DARWIN structures[a]	0.7	1.0	1.3	2.4	1.3

[a]Averaged values.

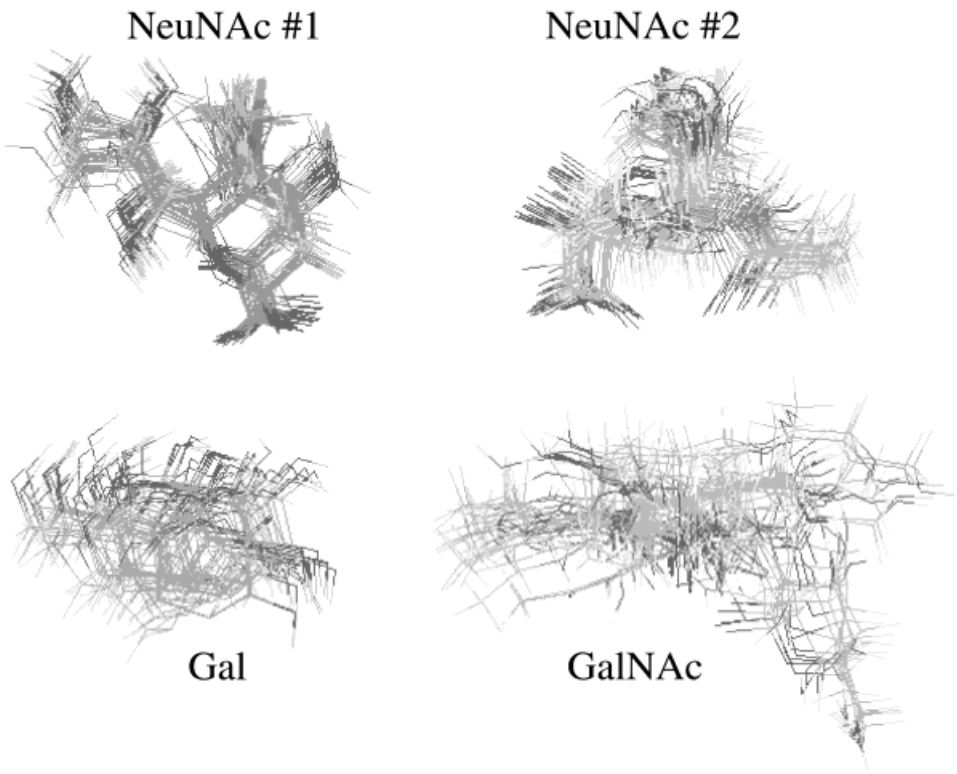


Fig. 11. Molecular cluster plot for GD2. The sugars in the 50 lowest energy structures are shown individually. The two sialic acid (NeuNAc) residues show very little deviation in their positions. The galactose residues show some translational spread, but generally maintain a common orientation. The GalNAc residue has the greatest variation, but has a bimodal distribution, with the pyranose ring oriented either horizontally or vertically.

Carbohydrates often have weak binding constants, 10^{-3} to 10^{-6} M, which contribute to the difficulties in obtaining crystals of protein-carbohydrate complexes. Even when the complex can be crystallized, the carbohydrate is often only partially visible in the electron density due to partial occupancy or static disorder. Most docking studies are based on the search for enzyme inhibitors with much stronger binding constants, in the range of 10^{-7} to 10^{-9} M, and molecules with binding constants similar to carbohydrates would not normally be considered as potential drug candidates. The results of the trials with DARWIN, in which three cases of increasing complexity were tackled with good results, show that the prediction of structures for carbohydrate complexes is not impossible, although difficult.

It is encouraging that the DARWIN docking results are well correlated with experimentally known binding affinities for the different ligands. An example is Se155-4, where the crystal structure temperature factors of the antigen are in accord with the different deviations for each residue in the DARWIN solution. Temperature factors reflect mobility and the average values for each of the sugar residues are 15.9, 31.5, and 44.3 $Å^2$ for abequose, mannose, and galactose, respectively. This indicates that the abequose residue, which is the most deeply buried in the binding pocket, is more tightly anchored to the protein than the mannose and galactose residues, which are on the surface. The DARWIN results reflected this situation as solutions were found in which the deviation of each residue from the correct solution correlated well with that residue's temperature factor (abequose 0.7 Å RMSD, mannose 1.2 Å RMSD, galactose 3.9 Å RMSD). This result suggests that the scatter among top-scoring solutions from DARWIN can be taken as a good measure of the binding affinities for different parts of flexible ligands.

An unexpected problem that arose with docking was the generation of intra-molecular antigen interactions resulting in nonsense solutions for the initial ME36.1/GD2 searches. Similar problems did not arise in the ConA and Se155-4 experiments for simple reasons. The ConA ligand was not flexible enough to make intra-molecular contacts. The Se155-4 antigen could possibly make a few intra-molecular contacts but, when so folded, it could not fit into the deep pocket at the binding site. As this pocket can completely surround a sugar ring, it provided a much larger number of inter-molecular contacts than could be made intra-molecularly. The problem for ME36.1 could be fixed by removing intra-molecular contacts from the evaluation function, but not from the local minimization. This adaptability illustrates one of the strengths of the two-part DARWIN procedure. While the first two docking experiments were not repeated with the updated evaluation function, the intra-molecular non-bonded energies for the ConA and Se155-4 ligands are very small. Omitting this term solely from the fitness score would be unlikely to alter the distribution of possible binding sites, and would have no effect on the final conformation within each site.

The ConA experiments highlighted bound water as an important and difficult factor in docking, as recognized by others in the field.[22] No practical method for including explicit solvent in the docking target has been developed so far. Our experiments with the ConA system demonstrated

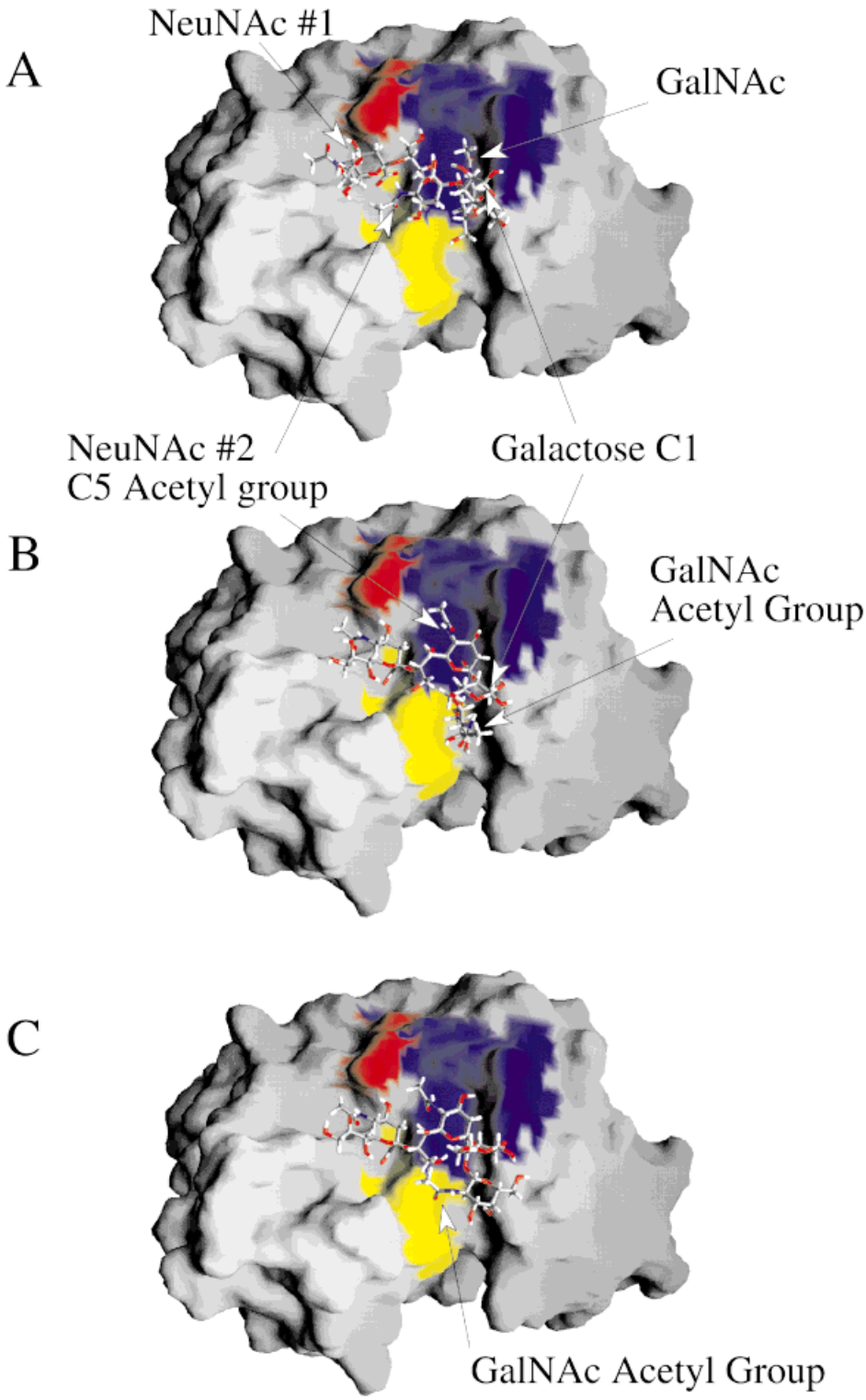


Fig. 12. Comparison of the model and DARWIN structures for the ME36.1/GD2 complex. The surface of the ME36.1 CDR is colored according to the underlying residues. The light chain L1 and L3 loops are blue, the heavy chain H3 loop is yellow, and a tyrosine on the L3 loop is red. **A:** The model structure.[16] The C5 acetyl group of NeuNAc #2 contacts the heavy chain and GalNAc contacts the light chain (arrows). **B:** The DARWIN "vertical" solution. NeuNAc #2 is flipped over so that the C5 acetyl group contacts the light chain (arrow). GalNAc contacts the heavy chain with the ring positioned vertically so that the acetyl group points straight up (arrow). In both the model and the DARWIN solution, NeuNAc #1 is in approximately the same position and Galactose C1, which makes the link to the cell surface, is exposed. **C:** The DARWIN "horizontal" solution. GalNAc is horizontal with the acetyl group to the left. Illustration was made with the program GRASP.[46]

how the lack of solvent made the CHARMM electrostatic energy term a liability in the search, even though it was important in identifying the final solution. When explicit solvent was not included, the calculated electrostatic forces were artificially high and led to improper placement of the ligand. Removing the electrostatics term from the energy improved the ConA search, but left an incomplete function that led the search to stumble into a false binding site in the dimer interface. Although this site would have been avoided if the dimer of ConA had been used for the target, it is clear that docking by shape complementarity alone is not sufficiently accurate. It is therefore encouraging that trials using Qnifft, which calculates a Poisson-Boltzmann based energy term, to model the effect of solvent gave good results. Practical difficulties in using two separate programs (CHARMM and Qnifft) for the evaluation function prevented more thorough testing of the Qnifft energy term. However, the Poisson-Boltzmann energy term that is being developed for the next version of CHARMM will permit its incorporation into DARWIN and facilitate further study in this area.

The DARWIN results for docking GD2 to ME36.1 highlight the difficulty of manually docking flexible molecules. The four main assumptions as to where GD2 should be placed in the Fab's antigen binding site,[16] appear fairly restrictive. (1) As the galactose residue provides the connecting point to the cell's surface, it must be accessible when the antigen is bound. (2) As iodination of the Fab blocks GD2 binding, it is likely that at least one tyrosine is involved in antigen binding. (3) Carbohydrate binding proteins often have two distinct globular domains that together form a cleft in which the carbohydrate can bind.[42] This arrangement is analogous to that in Fabs, where antigens should bind in a groove between the variable light and heavy domains. (4) The H3 and L3 loops on the heavy and light chains of Fabs are particularly important for specificity.[43] As the sequence in the L1 loop of ME36.1 is identical to the L1 loop from Fab J539, which also binds a carbohydrate antigen,[44] it was assumed that this loop is likely to be important.

Despite these constraints, DARWIN found a solution, with two variations, that fit the same criteria as the manual model and yet differed in the details of nearly every Fab/antigen interaction. Overall, the DARWIN solution appears reasonable. That it fits the four criteria is remarkable, as the search space was only restricted to satisfy the first of these. It validates the original model as the antigen is in the same position, with the same general orientation, but has improved antigen/Fab contacts and lower energy. The DARWIN solutions for ME36.1 accurately reflect weak binding for a portion of GD2 in a manner similar to that seen in the Se155-4 experiment. The position of the GalNAc was not well defined in the DARWIN solutions. DARWIN could not choose between the horizontal and vertical variations of the GalNAc residue's position, and these alternative conformations could not distinguished by Qnifft. It has been shown experimentally that ME36.1 can bind GD3, which lacks the GalNAc residue altogether.[45] This implies that the two sialic acid residues account for most of the binding energy, which is reasonable as these are the two sugars with net charge. The conclusion is that ME36.1 binds the GalNAc residue loosely and so it can adopt alternative conformations. The improved model for the ME36.1/GD2 complex will be valuable for ongoing studies on the interactions of ME36.1 with anti-idiotypic antibodies (Dr. Dorothee Herlyn, personal communication), which are part of a wider program on the induction of T-cell immunity.

For many drug design applications, docking is complicated by the fact that while the target enzyme is defined, the ligand to be docked is not. The experiment in which all eight pyranosides were simultaneously docked to ConA showed that DARWIN is likely to be effective and useful in selecting drug candidates from a pool of related ligands. The results for ConA fit well with experimental data showing that mannopyranoside binds more tightly than glucopyranoside, and that galactopyranoside has no measurable binding.[38] The remaining pyranosides have not been experimentally tested. As DARWIN can easily use CHARMM to make more drastic changes to a ligand than simple isomerization, it could use the GA to select functional groups on a template ligand. In this approach to "virtual combinatorial chemistry," different functional groups could be selected by DARWIN for addition to a defined ligand backbone. For each chromosome evaluation, CHARMM would add the selected functional groups to the ligand and calculate the complex's energy. In this manner, libraries of compounds could be searched for leads that then could be experimentally tested for their binding affinity to the drug target.

CONCLUSIONS

The docking problem has two parts, a search and an evaluation. The main conclusion of the work performed with DARWIN is that the search problem is tractable, and that the main difficulty is to evaluate potential solutions. The genetic algorithm was capable of optimizing the energy landscape in the problems tested, which had up to 14 degrees of freedom. Others have had success using a GA for problems with as many as 22 degrees of freedom.[21] The search algorithm was robust as DARWIN found solutions with fitness at least as good as the correct structure in each of the tests. In the cases where DARWIN failed to identify the correct solution, it was because structures with even lower energy were discovered (false positives).

The evaluation part of the docking problem is more difficult, but DARWIN provides a framework to continue work towards a robust general procedure. It would seem that DARWIN could be used advantageously for docking flexible ligands, and evaluating structurally related potential ligands. Although an evaluation function was not found that worked in all cases, successful methods could be found with some fine tuning. This is particularly encouraging for studies in which a protein-ligand structure is known and the goal is to test new ligands. In these cases, the experimentally determined structure

provides a control around which an evaluation scheme can be developed for other ligands. The use of DARWIN to refine a molecular model of the adenovirus capsid will be described in a forthcoming publication.

In summary, a general recipe for docking with DARWIN can be provided. It is best to begin with as few assumptions as possible as to the location and orientation of the ligand. An evaluation function should be selected, and a population size chosen based on how much computation time is available. Molecular cluster plots can identify possible binding sites and preferred ligand orientations. This information can be used to progressively narrow the search without overlooking possible solutions. Even if the binding site is known, other areas of the protein's surface should not be excluded immediately, as the detection of the correct binding site can be a good internal control for assessing the evaluation function. Electrostatics energy should be used with caution, and it is wise to perform docking runs with and without this term in the evaluation function. Likewise, where the structure is known, docking runs should be made with and without ordered water molecules in the target. If possible, test the evaluation function on a similar complex with a known structure. Remember that molecules are dynamic and that crystal structures are averaged positions, temperature factors can indicate mobility, and multiple docking solutions may be a reasonable reflection of reality. Take advantage of DARWIN's parallelism and, above all, do not let laboratory computers sit idle but install PVM and sign them up as evaluation slaves!

ACKNOWLEDGMENTS

The authors thank Dr. Kim Sharp for his assistance with Qnifft, and for his critical reading of the manuscript. We are grateful to Dr. Susan L. Pichla for providing assistance with the GD2 model, and Drs. Mitchell Lewis and G. Patrick Brady for helpful discussions and assistance in coding DARWIN's genetic algorithm. Computer time for this project was made available by the Cornell Theory Center. This work was performed at the Wistar Institute while J.S.T. was in the doctoral program of the Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine. He was supported partially by the Structural Biology Training Program (T32-GM 08275). The research was supported by grants to R.M.B. from the National Institutes of Health and the National Science Foundation, and by the Wistar Cancer Center.

REFERENCES

1. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. J Mol Biol 1982;161:269–288.
2. Welch W, Ruppert J, Jain AN. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. Chem Biol 1996;3:449–462.
3. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J Comp Chem 1983;4: 187–217.
4. Weiner PK, Kollman PA. AMBER: assisted model building with energy refinement. A general program for modeling molecules and their interactions. J Comp Chem 1981;2:287–303.
5. Leach AR. Ligand docking to proteins with discrete side-chain flexibility. J Mol Biol 1994;235:345–356.
6. Cummings MD, Hart TN, Read RJ. Monte Carlo docking with ubiquitin. Protein Sci 1995;4:885–899.
7. Wasserman ZR, Hodge CN. Fitting an inhibitor into the active site of thermolysin: a molecular dynamics case study. Proteins 1996;24: 227–237.
8. Rosenfeld R, Vajda S, DeLisi C. Flexible docking and design. Annu Rev Biophys Biomol Struct 1995;24:677–700.
9. Athappilly FK, Murali R, Rux JJ, Cai Z, Burnett RM. The refined crystal structure of hexon, the major coat protein of adenovirus type 2, at 2.9 Å resolution. J Mol Biol 1994;242:430–455.
10. Stewart PL, Burnett RM, Cyrklaff M, Fuller SD. Image reconstruction reveals the complex molecular organization of adenovirus. Cell 1991;67:145–154.
11. Holland JH. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. Cambridge: MIT Press; 1992. 211 p.
12. Forrest S. Genetic algorithms: principles of natural selection applied to computation. Science 1993;261:872–878.
13. Ha SN, Giammona A, Field M, Brady JW. A revised potential-energy surface for molecular mechanics studies of carbohydrates. Carbohydr Res 1988;180:207–221.
14. Taylor JS, Teo B, Wilson DB, Brady JW. Conformational modeling of substrate binding to endocellulase E2 from *Thermomonospora fusca*. Protein Eng 1995;8:1145–1152.
15. Naidoo KJ, Denysyk D, Brady JW. Molecular dynamics simulations of the N-linked oligosaccharide of the lectin from *Erythrina corallodendron*. Protein Eng 1997;10:1249–1261.
16. Pichla SL, Murali R, Burnett RM. The crystal structure of a Fab fragment to the melanoma-associated GD2 ganglioside. J Struct Biol 1997;119:6–16.
17. Herlyn M, Koprowski H. Melanoma antigens: immunological and biological characterization and clinical significance. Annu Rev Immunol 1988;6:283–308.
18. Bramlette MF, Bouchard EE. Genetic algorithms in parametric design of aircraft. In: Davis L, editor. Handbook of genetic algorithms. Boston: International Thomson Computer Press; 1996. p 109–123.
19. Cox LA Jr, Davis L, Qiu Y. Dynamic anticipatory routing in circuit-switched telecommunications networks. In: Davis L, editor. Handbook of genetic algorithms. Boston: International Thomson Press; 1996. p 124–143.
20. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 1995;245:43–53.
21. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol 1997;267:727–748.
22. Judson RS, Tan YT, Mori E, Melius C, Jaeger EP, Treasurywala AM, Mathiowetz A. Docking flexible molecules: a case study of three proteins. J Comp Chem 1995;16:1405–1419.
23. Oshiro CM, Kuntz ID, Dixon JS. Flexible ligand docking using a genetic algorithm. J Comput Aided Mol Des 1995;9:113–130.
24. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Nguyen DT, Ngo T, Prodhom B, Roux B, Schlenkrich M, Smith J, Stote R, Straub J, Wiorkiewicz-Kuczera J, Karplus M. Self-consistent parameterization of biomolecules for molecular modeling and condensed phase simulations. FASEB J 1992;6:A143.
25. Allen FH, Davies JE, Galloy JJ, Johnson O, Kennard O, Macrae CF, Mitchell EM, Mitchell GF, Smith JM, Watson DG. The development of versions 3 and 4 of the Cambridge Structure Database System. J Chem Inf Comput Sci 1991;31:187–204.
26. Liang G, Brady JW. CHARMM parameters for alpha-D-glucopyranose, Version 24b2. Cambridge: Harvard University; 1994.
27. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA Jr. General atomic and molecular electronic structure system. J Comp Chem 1993;14:1347–1363.
28. Steinbach PJ, Brooks BR. New spherical-cutoff methods for

long-range forces in macromolecular simulation. J Comp Chem 1994;15:667–683.
29. Wesson L, Eisenberg D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. Protein Sci 1992;1:227–235.
30. Vajda S, Sippl M, Novotny J. Empirical potentials and functions for protein folding and binding. Curr Opin Struct Biol 1997;7:222–228.
31. Gilson MK, Sharp KA, Honig BH. Calculating the electrostatic potential of molecules in solution: method and error assessment. J Comp Chem 1987;9:327–335.
32. Morris GM, Goodsell DS, Huey R, Olson AJ. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J Comput Aided Mol Des 1996;10:293–304.
33. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free energies using macroscopic solvent molecules. J Phys Chem 1994;98:1978–1988.
34. Shoichet BK, Kuntz ID. Protein docking and complementarity. J Mol Biol 1991;221:327–346.
35. Geist A, Beguelin A, Dongarra J, Jiang W, Manchek R, Sunderam V. PVM: parallel virtual machine. A user's guide and tutorial for networked parallel computing. Cambridge: MIT Press; 1994. 299 p.
36. Naismith JH, Emmerich C, Habash J, Harrop SJ, Helliwell JR, Hunter WN, Raftery J, Kalb (Gilboa) AJ, Yariv J. Refined structure of Concanavalin A complexed with methyl α-D-mannopyranoside at 2.0 Å resolution and comparison with the saccharide-free structure. Acta Crystallogr D Biol Crystallogr 1994;50:847–858.
37. Zdanov A, Li Y, Bundle DR, Deng S-J, MacKenzie CR, Narang SA, Young NM, 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 USA 1994;91:6423–6427.
38. Dani M, Manca F, Rialdi G. Calorimetric study of concanavalin A binding to saccharides. Biochim Biophys Acta 1981;667:108–117.
39. Brünger AT, Karplus M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. Proteins 1988;4:148–156.
40. Cygler M, Rose DR, Bundle DR. Recognition of a cell-surface oligosaccharide of pathogenic *Salmonella* by an antibody Fab fragment. Science 1991;253:442–445.
41. Sigurskjold BW, Bundle DR. Thermodynamics of oligosaccharide binding to a monoclonal antibody specific for a *Salmonella* O-antigen point to hydrophobic interactions in the binding site. J Biol Chem 1992;267:8371–8376.
42. Quiocho FA. Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. Annu Rev Biochem 1986;55:287–315.
43. Davies DR, Padlan EA, Sheriff S. Antibody-antigen complexes. Annu Rev Biochem 1990;59:439–473.
44. Ekborg G, Ittah Y, Glaudemans CPJ. Monoclonal IgA J539 binds galactopyranosyl antigens on its surface. Mol Immunol 1983;20:235–238.
45. Thurin J, Thurin M, Kimoto Y, Herlyn M, Lubeck MD, Elder DE, Smereczynska M, Karlsson K-A, Clark WM Jr, Steplewski Z, Koprowski H. Monoclonal antibody-defined correlations in melanoma between levels of G_{D2} and G_{D3} antigens and antibody-mediated cytotoxicity. Cancer Res 1987;47:1229–1233.
46. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.