FOR THE RECORD

Computational models explain the oligosaccharide specificity of cyanovirin-N

YUKIJI K. FUJIMOTO,[1] RYAN N. TERBUSH,[2] VADIM PATSALO,[3] AND DAVID F. GREEN[3,4]

[1]Department of Chemistry, Stony Brook University, Stony Brook, New York 11794-3400, USA
[2]Smithtown West High School, Smithtown, New York 11787-1630, USA
[3]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, New York 11794-3600, USA
[4]Graduate Program in Biochemistry and Structural Biology, Stony Brook University, Stony Brook, New York 11794-3600 USA

(RECEIVED January 19, 2008; FINAL REVISION July 30, 2008; ACCEPTED July 31, 2008)

Abstract

The prokaryotic lectin cyanovirin-N (CV-N) is a potent inhibitor of HIV envelope-mediated cell entry, and thus is a leading candidate among a new class of potential anti-HIV microbicides. The activity of CV-N is a result of interactions with the D1 arm of high-mannose oligosaccharides on the viral glycoprotein gp120. Here, we present computationally refined models of CV-N recognition of the di- and trisaccharides that represent the terminal three sugars of the D1 arm by each CV-N binding site. These models complement existing structural data, both from NMR spectroscopy and X-ray crystallography. When used with a molecular dynamics/continuum electrostatic (MD/PBSA) approach to compute binding free energies, these models explain the relative affinity of each site for the two saccharides. This work presents the first validation of the application of continuum electrostatic models to carbohydrate–protein association. Taken as a whole, the results both provide models of CV-N sugar recognition and demonstrate the utility of these computational methods for the study of carbohydrate-binding proteins.

Keywords: virucidal lectins; continuum electrostatics; molecular recognition; carbohydrate-binding protein; binding affinity

Supplemental material: see www.proteinscience.org

Since the discovery of the antiviral activity of legume lectins such as Concanavalin A more than 20 years ago (Lifson et al. 1986), it has been known that one mechanism of blocking HIV-envelope-mediated cell entry is the association of a carbohydrate-binding protein with the oligosaccharides on the surface of the viral envelope. In particular, the outer envelope glycoprotein, gp120, is heavily glycosylated by *N*-linked carbohydrates, both of the high-mannose and complex types (Leonard et al. 1990). As this glycosylation plays a key role in viral avoidance of natural immune responses, inhibitors that work through specific interactions with these carbohydrates provide a unique opportunity to interfere with cellular infection in a manner that makes it difficult for viral resistance to evolve (Scanlan et al. 2007).

To date, numerous lectins from diverse sources have been identified as having virucidal activity against HIV, but among the best characterized is cyanovirin-N (CV-N), originally isolated from the cyanobacterium *Nostoc ellipsosporum* (Boyd et al. 1997). Both inhibition and calorimetric studies have determined that CV-N specifically targets the α(1–2)-linked mannoses found on the D1 arm of high-mannose oligosaccharides; CV-N contains two pseudo-symmetric binding domains which have differing

Reprint requests to: David F. Green, Department of Applied Mathematics and Statistics, Stony Brook University, Math Tower Room 1-117, Stony Brook, NY 11794-3600, USA; e-mail: david.green@stonybrook.edu; fax: (631) 632-8490.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.034637.108.

affinities and specificities for various oligosaccharides (Bewley and Otero-Quintero 2001; Bewley et al. 2002). Several structures of CV-N have also been solved, both in solution (Bewley et al. 1998; Bewley 2001) and crystal phases (Yang et al. 1999; Botos et al. 2002). Both its remarkable stability to denaturation and its potent antiviral activity have made CV-N one of the leading candidates for use as a biopharmaceutical. Preclinical trials in primates have shown promise for the use of CV-N as a topical agent to prevent sexual transmission of HIV (Tsai et al. 2003, 2004).

Despite the apparent wealth of data for this system, there remain open problems regarding the structure and energetics of specific sugar recognition. The crystal structures of CV-N bound to each of two high-mannose oligosaccharides, Man_6 and Man_9, have been solved (Botos et al. 2002), but the carbohydrates are not fully resolved. While they are of reasonable resolution (2.4 and 2.5 Å), only one of the two binding sites is occupied, and the oligosaccharide structures deviate strongly from that expected; several mannoses are in the β configuration (where α anomers are expected), and numerous rings are in disfavored ring conformations. Overall, only the small portion of the oligosaccharide making the most intimate contact with the protein is particularly well structured. A solution structure with both sites bound to the disaccharide Manα(1–2)Man is also available (Bewley 2001). In this structure, however, few NOE constraints were available to accurately define the lower affinity site. Here, we present computationally refined models of α(1–2)-linked di- (Man_2) and trimannose (Man_3), representative of the D1 arm of Man_9, bound in both sites. These models, combined with molecular dynamics and continuum electrostatic analysis, capture the observed specificity of binding with semiquantitative accuracy.

Results and Discussion

Initially, explicit-solvent molecular dynamics simulations were carried out, beginning with the solution structure in complex with two mannose disaccharides (PDB 1iiy) (Bewley 2001). These simulations quickly revealed an instability in the lower affinity binding site; the disaccharide began to dissociate from cyanovirin-N within the first 500 ps of simulation (see Fig. 1, top panels). In the higher affinity site, however, the sugar remains stably bound for as long as we have simulated (currently upwards of 20 ns). Similar behavior has been noticed by Margulis, using

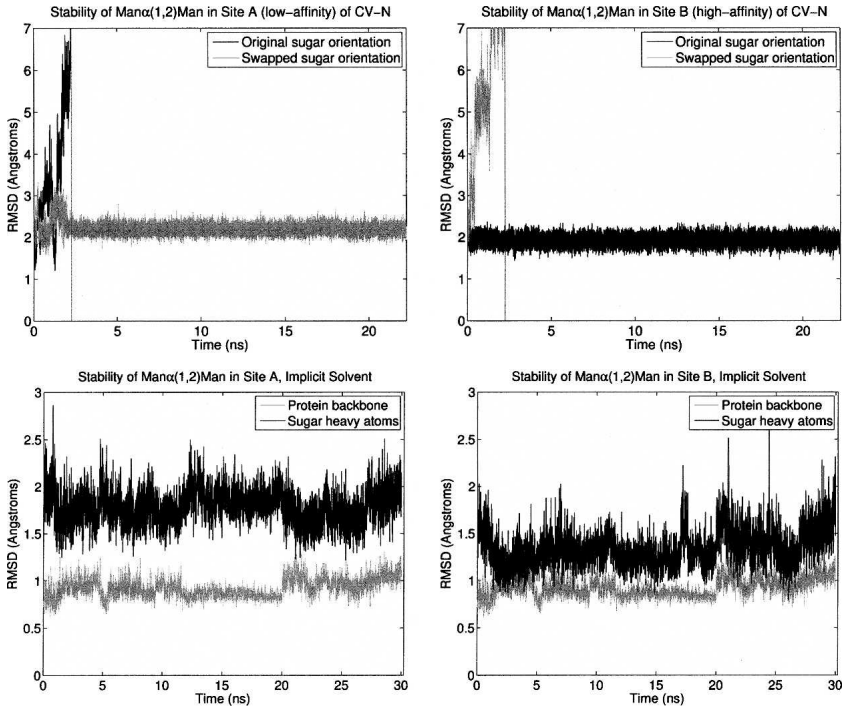

Figure 1. Stability of Man_2 in each CV-N binding site. (*Top*) The root-mean-square deviation (RMSD) of sugar-heavy atoms, relative to the initial coordinates, are shown over 22 ns of explicit-solvent molecular dynamics. Trajectories beginning from the published structure are in black; those beginning with the sugar placed in the orientation found in the other site are shown in gray. A vertical bar indicates the end of the equilibration phase; all analysis was done on frames following this point. (*Bottom*) The RMSD of both protein backbone atoms (in black) and sugar-heavy atoms (in gray), relative to the initial coordinates, are shown over 30 ns of implicit solvent (Generalized Born) molecular dynamics.

different parameters and simulation conditions (Margulis 2005), suggesting that the instability is not simply an artifact of the method.

The difference between the two sites in affinity for Man_2 is small (10-fold difference in K_a, as measured by a two-site fit to calorimetric data) (Bewley and Otero-Quintero 2001); thus, this behavior is not expected. The two binding sites of CV-N are pseudo-symmetric, but the orientation of the sugar in each site of the structure differs noticeably; the sugar in the low-affinity site makes less intimate contact with some binding-site residues (see Fig. 2, top panels). A higher symmetry model of dimannose binding in the low-affinity site was thus considered. The protein backbone of the high affinity binding site was superimposed on that of the lower affinity site, and the sugar from the high affinity site placed in the lower affinity pocket. The structure was then briefly minimized while all protein residues further than 4.0 Å (minimum distance) from the sugar were held fixed. Molecular dynamics simulation from this starting structure shows similar stability to that of the high-affinity site. As an additional test, this procedure was repeated in reverse—the low-affinity site was superimposed on the high-affinity site, and the low-affinity sugar orientation placed in the high-affinity pocket. The binding site was briefly relaxed, and then subjected to molecular dynamics. This model shows the same instability as was first observed in the lower affinity site. These data are displayed in Figure 1 (top panels). However, it remains possible that the observed stability of the refined structure is due to the constraints used in the simulation, which involved a sphere of solvent placed around each binding site, with the protein atoms outside this sphere held fixed. To evaluate this possibility, an unconstrained simulation using the Generalized Born implicit solvent model was run, beginning with the preferred sugar orientation in each site. The sugars remain stably bound throughout this simulation (Fig. 1, bottom panels), while similar simulations with the nonpreferred starting orientations show rapid dissociation (data not shown). Taken in combination, these results strongly suggest that the source of the instability is the initial orientation of the sugar.

Fewer NOEs were observed experimentally for the lower affinity site than for the higher, and thus the published model was not uniquely determined by experimental constraints (Bewley 2001). To further test the validity of the new model, we tracked the distances of all atoms involved in observed intermolecular NOEs throughout the first 10 ns of molecular-dynamics simulation (details in the Supplemental material); all experimentally observed contacts remain within 6.0 Å for the majority of the simulation, and most remain within 5.0 Å or less. In comparison, several contacts in the published model are beyond the largest constraint distance (6.0 Å). While the timescale of the simulation is much lower than that observed in the NOESY experiment, these data demonstrate the consistency of the refined model with available data.

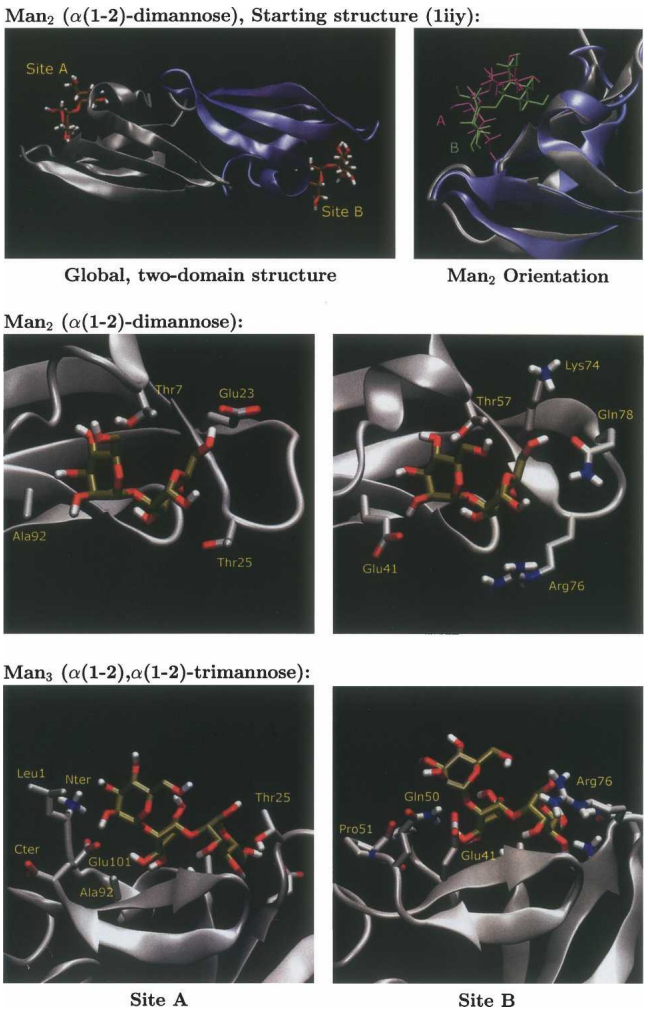

Figure 2. Structures of di- and trimannose bound to CV-N. (*Top left*) The overall structure of CV-N, showing two sugar binding sites in pseudo-symmetric domains (domain A in gray, domain B in blue). (*Top right*) A superposition of the two binding sites of CV-N, showing the difference in orientation of the original Man_2 model in the low-affinity site (pink/gray) from the structure of the high-affinity site (green/blue). (*Middle* and *bottom*) A representative frame from each dynamic simulation is displayed, with protein in gray and sugar in bronze. All amino acid side chains involved in significant electrostatic interactions are shown. The viewpoint perspective of Man_3 is rotated roughly 90° from that of Man_2. Figures generated with VMD (Humphrey et al. 1996).

These structures begin to explain the key determinants of oligosaccharide specificity in the two sites of CV-N, as shown in Figure 2 (middle panels). In particular, a strong electrostatic interaction in the high-affinity site (between Glu41 and Man2 OH2) is absent in the low-affinity site, where the corresponding residue is an alanine (Ala92).

There are other ways in which the two binding sites vary, but the overall differences in interactions are less pronounced. A hydrogen bond made between Glu23 and Man1 OH6 in Site A is replaced by a similar interaction with Gln78 in Site B; these residues are not in equivalent positions, but make equivalent interactions. Thr25 in Site A is replaced by Arg76 in Site B, but neither make close, specific interactions with the sugar. Interestingly, the Glu/Ala variation at sites 41/92 also may explain an intriguing feature of CV-N—α(1-2)-linked trimannose binds to these sites with a specificity reversed from that of the dimannose. That is, Site A (low affinity for dimannose) binds the trimannose with higher affinity than does Site B (high-affinity for dimannose). With an additional α(1–2)-linkage extending from Man2, the favorable hydrogen bond made by Glu41 in Site B would be lost, as the donating hydroxyl would become part of the glycosidic bond; in Site A the lack of an interaction with this hydroxyl would more easily accommodate this change.

To investigate this further, models of Man_3 in each site were generated by extending the dimannose by one unit. A new glycosidic bond formed either with OH2 of Man2 or OH1 of Man1 would result in the same α(1–2),α(1–2)-linked trimannose; the structure of Man_9—accurate in this region—indicates that the first of these configurations is preferred (Botos et al. 2002), and thus this was used in model building. Again, molecular dynamics simulations were performed to assess stability, and the sugar in both binding sites remained stably bound throughout. Representative structures are shown in Figure 2 (bottom). These structures support the analysis above: The incorporation of Man2 OH2 into the glycosidic bond makes Glu41 unable to interact favorably with the sugar in Site B. As well, two new interactions are made in Site A, explaining the increased affinity of this site for the trisaccharide: Glu101 (the C-terminal residue) receives a hydrogen bond from OH2 of the third sugar; and the charged N terminus donates one to OH3. In Site B, neither of these interactions are made: Gln50 (analogous to Glu101) is involved in satisfying Glu41 with an intramolecular interaction, and thus makes no interactions with the sugar; and a proline-containing loop replaces the chain termini, eliminating any hydrogen-bonding potential.

While the explanations based on the structural models seem reasonable, consideration of a bound state alone can be misleading, as the energetics of binding involves differences between bound and unbound states. Desolvation effects, in particular, have been clearly established as essential contributors to affinity and specificity (Sheinerman et al. 2000; Lazaridis 2002). To assess this, the binding free energies of each complex were computed, accounting for solvent with a Poisson–Boltzmann/Surface Area (PBSA) model (Gilson and Honig 1988; Sitkoff et al. 1994) recently optimized for use with carbohydrates (Green 2008); these energies were combined with van der Waals interactions and averaged over one thousand frames extracted from the MD simulations (the MD/PBSA approach) (Srinivasan et al. 1998). To assess convergence of these simulations, averages were also computed over the first and last halves of the trajectory independently, and with 5- and 10-fold coarser sampling; all these ensembles gave statistically equivalent results (details in the Supplemental material), suggesting strong convergence of the results, both in terms of length of simulation and sampling frequency. These calculations give three energetic terms: the intermolecular van der Waals interactions made in the bound state (ΔG^{vdw}), the hydrophobic solvation energy ($\Delta G^{h\phi}$, computed as proportional to the solvent-accessible surface area buried upon binding), and the electrostatic contribution (ΔG^{elec}, including both the solvent-screened intermolecular interactions and the cost of desolvating each molecule upon binding). The sum of these terms gives what may be termed a ''semirigid'' binding free energy; energies are averaged over multiple conformations of both sugar and protein, but the structural ensembles of the unbound state are identical to those of the bound state.

The unbound ensembles will not, in actuality, be identical to those of the bound complex, and thus several additional terms contribute to the true binding free energy. Two of these are strain energies—the energetic cost of perturbing the ensemble of structures found in the unbound state into that found in the bound state—with a contribution both from the sugar and from the protein. Strain energies are easily computed by running additional simulations of the unbound state; the difference in total energy (bond, angle, and dihedral strain, intramolecular van der Waals and Coulombic interactions, and solute–solvent interactions) between the unbound ensemble and the ensemble of structures taken from the bound state (with the binding partner removed) is the strain. This contribution was computed for each oligosaccharide (ΔG^{str}); the constraints used in the simulations led to artifacts in the protein strain term, which was thus neglected. Conceptually, this approach considers binding as a two-step process, in which the ligand ensemble first is perturbed into an ensemble that is ''pre-formed'' for binding (the carbohydrate strain energy), followed by rigid-body binding of each member of this ensemble to the corresponding member of the protein ensemble (semirigid binding energy).

Additionally, there are terms related to solute entropy (solvent entropy is included implicitly in the PBSA model), both due to configuration/vibrational flexibility of each solute as well as the translational and rotational degrees of freedom of each component. Configurational entropy is most often estimated using normal modes (as in the first applications of the MM/PBSA approach) or quasi-harmonic analysis, with the latter having an advantage of being

derived from the same simulation as other energetic values (Karplus and Kushick 1981). There are significant assumptions made in both these approaches, however, and thus the values are best considered estimates. Entropic contributions to binding from the carbohydrate estimated with a quasi-harmonic analysis are roughly 0.5 kcal/mol, and contributions to relative binding free energies are even smaller (<0.2 kcal/mol, see Supplemental material). These values have been neglected in subsequent analysis, but their inclusion would not change the results in any significant way. The small contributions may initially be nonintuitive, if one expects significant flexibility in the free sugars. However, analysis of glycosidic bond dihedrals throughout the unbound simulation suggests a highly restricted conformational space, even in the unbound state (Supplemental material); the small additional constriction upon binding incurs a relatively small entropic penalty. Of course, it is entirely possible that more complete conformational sampling of the unbound state requires much longer timescales than considered here, a caveat in all simulations. In keeping with the semirigid approximation for the protein (to avoid artifacts from constraints), protein entropy was neglected.

The terms neglected with a semirigid (protein) receptor model certainly contribute to the absolute binding free energies, and may contribute to relative energies as well. However, the qualitative agreement of the results with experiment suggests that these approximations are reasonable. In Site B, Man_2 and Man_3 make largely the same contacts with the protein, with the exception of a single hydrogen bond to E41; thus, the protein strain and entropy of these two states may be expected to be quite similar; in Site A, Man_3 makes additional contacts with the free N and C termini, and thus a larger unfavorable contribution might be expected than for Man_2 binding in this site, or for Man_3 binding in Site B. Site B additionally contains a large, flexible arginine at position 76, replaced with Thr 25 in Site A; this could lead to an increased unfavorable contribution in Site B for both sugars. Quantifying these effects is beyond the scope of this work, but the general trends should be considered in considering the following results.

The relative affinities of each state (presented in Table 1) are all computed with reasonable accuracy, further supporting the validity of the structural models; no attempt has been made here to modify our energetic model to give improved agreement with experiment. In all cases, the trends are computed correctly, both trimer/dimer specificity in each site and the relative affinities of either ligand in the two sites. However, the magnitude of the differences are somewhat overpredicted, a matter worthy of further consideration; for the two favorable values, the strain and entropic penalties paid by the protein on binding might be expected to modulate these toward lower magnitudes. Considering each energetic term suggests a dominant role for electrostatic interactions in defining the specificity at each site, as only this term shows the same trend as the overall free-energy differences. This is consistent with the structural analysis, in which all major differences involved electrostatic interactions.

Structure-guided mutagenesis and design are important tools, both for exploring function and for the development of complexes with enhanced affinity. For these to be successful, however, accurate models are essential. The models presented here thus provide an important reference for future work on cyanovirin-N. Ensembles of snapshots from the MD trajectories, as well as a minimized representative structure of each refined complex, are available from the authors.

Additionally, the results validate the use of the MD/PBSA approach in the study of carbohydrate-binding proteins. While this method has been used extensively in the study of protein–protein and protein–small molecule interactions, to date there have been few applications to carbohydrates. This work strongly motivates the pursuit of future studies on protein–carbohydrate recognition by these approaches.

Materials and Methods

Construction of increased symmetry binding models

The initial structure for all calculations was the solution structure of cyanovirin-N (CV-N) bound to Manα(1–2)Man (PDB 1iiy) (Bewley 2001). CV-N has two pseudo-symmetric binding domains; in order to construct binding models based on

Table 1. *Relative free energies of binding*[a]

	$\Delta\Delta G^{vdw}$	$\Delta\Delta G^{h\phi}$	$\Delta\Delta G^{elec}$	$\Delta\Delta G^{str}$	$\Delta\Delta G^{comp}$	$\Delta\Delta G^{expt}$
$\Delta\Delta G$ Site A−Site B						
Man_2	−0.9 (0.1)	−0.1 (0.0)	−2.3 (0.1)	+0.4 (0.2)	−2.8 (0.2)	−1.5
Man_3	−3.8 (0.1)	+0.3 (0.0)	+8.2 (0.1)	+2.1 (0.2)	+6.9 (0.3)	+1.7
$\Delta\Delta G$ Man_3–Man_2						
Site A	−1.3 (0.1)	−0.7 (0.0)	−3.7 (0.1)	−0.7 (0.3)	−6.4 (0.3)	−1.4
Site B	−4.2 (0.1)	−0.2 (0.0)	+6.5 (0.1)	+1.0 (0.3)	+3.3 (0.4)	+1.8

[a] Van der Waals, hydrophobic surface burial, electrostatics, carbohydrate strain, total computed, and experimental (Bewley and Otero-Quintero 2001; Bewley et al. 2002), all in kcal/mol. Errors are the standard error of the mean for the ensemble averages.

symmetry, the backbone atoms of equivalent residues in each site were superimposed by an RMSD fit. The coordinates of the sugar in the site were then replaced with those from the superimposed structure; visual analysis indicated that no major steric clashes were introduced by this procedure. This initial placement was then subjected to a short minimization, with all protein residues >4.0 Å from the sugar held fixed, to alleviate small clashes. These manipulations were done using the CHARMM software package (version 32B1) (Brooks et al. 1983).

Explicit-solvent molecular dynamics

Molecular dynamics calculations were performed with CHARMM, using the Param22 protein force field (MacKerell et al. 1998), the Carbohydrate Solution Force Field (csff) (Kuttel et al. 2002), and a TIP3P water model (Jorgensen et al. 1983). A 15.0 Å radius sphere of water molecules was centered on each ligand; waters with oxygens within 2.8 Å of any solute heavy atom were then removed. After 10 ps of simulation with fixed sugar and protein atoms, these steps were repeated to fill voids in the solvent. For further dynamics, all atoms outside the sphere were fixed, and a spherical boundary potential maintained waters in the droplet. Protein atoms in the outer 2 Å of the sphere were harmonically restrained (10 kcal/mol/Å^2) and subjected to Langevin dynamics with a moderate friction coefficient (10 ps^{-1}); waters in the outer 2 Å were unrestrained, but subjected to Langevin dynamics with a high friction coefficient (62 ps^{-1}). Atoms in the inner 13 Å radius sphere were propagated with Newtonian dynamics. The SHAKE algorithm was used to fix all bonds involving hydrogens, and the time step of integration was 2 fs. The system was equilibrated for 2 ns, following which 20 ns of dynamics were collected for further study.

Implicit-solvent molecular dynamics

Implicit-solvent molecular dynamics simulations were carried out with the CHARMM software package and force field parameters as detailed above. The GBSW module was used to provide Generalized Born-based solvation terms (Im et al. 2003), using atomic radii optimized for this approach (Nina et al. 1997; Green 2008). No constraints were applied, other than the use of the SHAKE algorithm to fix all bonds involving hydrogens. Simple Newtonian dynamic was used, with a time step of 2 fs.

Computation of binding free energies

Binding free energies were computed with an MD/PBSA model (Srinivasan et al. 1998). The explicit-solvent MD trajectories were sampled every 20 ps, for a total of one thousand frames per system. Rigid-body binding free energies were computed for each snapshot, and the results averaged over all frames. Electrostatic contributions were computed for each snapshot as the difference between the total electrostatic free energies of the bound complex and the two unbound components. These were obtained by solution of the linearized Poisson–Boltzmann (PB) equation (Gilson and Honig 1988), using a multigrid finite-difference solver (M.D. Altman and B. Tidor, unpubl.) distributed with the Integrated Continuum Electrostatics (ICE) software package (Green and Tidor 2003; D.F. Green, E. Kangas, Z.S. Hendsch, and B. Tidor, unpubl.). A solute dielectric constant of 2.0 and a solvent dielectric constant of 80.0 were used; the dielectric boundary was defined by the molecular surface using a 1.4 Å radius probe, with radii optimized for this purpose (Nina et al. 1997; Green 2008). The ionic strength was set to 0.145 M, with a 2.0 Å ion-exclusion layer. A 65^3-unit grid was used with overfocusing boundary conditions (the longest dimension of the molecule occupying first 23%, then 92%, and then 184% of one edge of the grid). Boundary potentials extracted from the previous calculation were used in each case, with Debye–Hückel boundary potentials used for the initial calculation. Energetic contributions from atoms falling outside the finest grid were taken from the middle-resolution grid. The total binding energy for each snapshot was the sum of the electrostatic contribution, the intermolecular van der Waals energy and a term proportional to the solvent-accessible surface area buried on binding. The area was computed with CHARMM, using a 1.4 Å probe radius, and the energetic contribution was given by $\Delta G^{h\Phi} = 0.005\Delta A + 0.86$ kcal/mol (Sitkoff et al. 1994).

Computation of sugar strain energies

Sugar strain energies were computed by comparing the total ensemble-averaged energies of sugar conformations extracted from bound-state molecular dynamics and those for conformations from unbound simulations. Electrostatic solvation free energies were computed with a Poisson–Boltzmann model, as the difference between a system with solvent dielectric constant of 80 (ionic strength of 0.145 M) and solute dielectric constant of 2, and a system of uniform dielectric constant of 2 (zero ionic strength). For these calculations, the largest grid contained the whole sugar (92% of the longest dimension). Hydrophobic solvation free energies were estimated with a term proportional to the total solute surface area, as described above. These energies were added to the molecular-mechanics energy of the solute, including all bonded terms (bonds, angles, dihedrals), intramolecular van der Waals, and intramolecular Coulombic (in uniform dielectric of 2) interactions. The sugar strain energy is given by the difference in the ensemble average of the total energy of the sugar in conformations extracted from the complex simulation and the similar average for conformations from a simulation of the free sugar. Thus, these values correspond to the energetic cost of perturbing the unbound conformational ensemble into the ensemble that is capable of binding, in a fully solvated context. Sugar entropies were computed from the same trajectories using the QUASI module of the CHARMM software package (Karplus and Kushick 1981), using the average structure as a reference.

Electronic supplemental material

Tables of all computed energies, of the persistence of observed NOE contacts, details on convergence assessment, and analysis of carbohydrate entropies are available as Supplemental material.

Acknowledgments

We thank Bruce Tidor for making available the MultigridPBE software. This work was supported by startup funding awarded to D.F.G. by the State University of New York, the College of Engineering and Applied Sciences, and the Department of Applied Mathematics and Statistics, and by a Simons Scholar summer research fellowship to R.N.T.

References

Bewley, C.A. 2001. Solution structure of a cyanovirin-N:Manα1–2Manα complex: Structural basis for high-affiinity carbohydrate-mediated binding. *Structure* **9:** 931–940.

Bewley, C.A. and Otero-Quintero, S. 2001. The potent anti-HIV protein cyanovirin-N contains two novel carbohydrate binding sites that selectively bind to Man_8D1D3 and Man_9 with nanomolar affinity: Implications for binding to the HIV envelope protein gp120. *J. Am. Chem. Soc.* **123:** 3892–3902.

Bewley, C.A., Gustafson, K.R., Boyd, M.R., Covell, D.G., Bax, A., Clore, G.M., and Gronenborn, A.M. 1998. Solution structure of cyanovirin-N, a potent HIV-inactivating protein. *Nat. Struct. Biol.* **5:** 571–578.

Bewley, C.A., Kiyonaka, S., and Hamachi, I. 2002. Site-specific discrimination by cyanovirin-N for α-linked trisaccharides comprising the three arms of Man_8 and Man_9. *J. Mol. Biol.* **322:** 881–889.

Botos, I., O'Keefe, B.R., Shenoy, S.R., Cartner, L.K., Ratner, D.M., Seeberger, P.H., Boyd, M.R., and Wlodawer, A. 2002. Structures of the complexes of a potent anti-HIV protein cyanovirin-*N* and high mannose oligosaccharides. *J. Biol. Chem.* **277:** 34336–34342.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., Currens, M.J., et al. 1997. Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: Potential applications to microbicide development. *Antimicrob. Agents Chemother.* **41:** 1521–1530.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4:** 187–217.

Gilson, M.K. and Honig, B. 1988. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins Struct. Funct. Genet.* **4:** 7–18.

Green, D.F. 2008. Optimized parameters for continuum solvation calculations with carbohydrates. *J. Phys. Chem. B* **112:** 5238–5249.

Green, D.F. and Tidor, B. 2003. Evaluation of electrostatic interactions. In *Current protocols in bioinformatics* (ed. G.E. Petsko), chapter 8.3. John Wiley & Sons, Inc., New York.

Humphrey, W., Dalke, A., and Schulten, K. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* **14:** 33–38.

Im, W., Lee, M.S., and Brooks III, C.L. 2003. Generalized born model with a simple smoothing function. *J. Comput. Chem.* **24:** 1691–1702.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79:** 926–935.

Karplus, M. and Kushick, J.N. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules* **14:** 325–332.

Kuttel, M., Brady, J.W., and Naidoo, K.J. 2002. Carbohydrate solution simulations: Producing a force field with experimentally consistent primary alcohol rotational frequencies and populations. *J. Comput. Chem.* **23:** 1236–1243.

Lazaridis, T. 2002. Binding affinity and specificity from computational studies. *Curr. Org. Chem.* **6:** 1319–1332.

Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N., and Gregory, T.J. 1990. Assignment of interchain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J. Biol. Chem.* **265:** 10373–10382.

Lifson, J., Coutré, S., Huang, E., and Engleman, E. 1986. Role of envelope glycoprotein carbohydrate in human immunodeficiency virus (HIV) infectivity and virus-induced cell fusion. *J. Exp. Med.* **164:** 2101–2106.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102:** 3586–3616.

Margulis, C.J. 2005. Computational study of the dynamics of mannose disaccharides free in solution and bound to the potent anti-HIV virucidal protein cyanovirin. *J. Phys. Chem. B* **109:** 3639–3647.

Nina, M., Beglov, D., and Roux, B. 1997. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. B* **101:** 5239–5248.

Scanlan, C.N., Offer, J., Zitzmann, N., and Dwek, R.A. 2007. Exploiting the defensive sugars of HIV-1 for drug and vaccine design. *Nature* **446:** 1038–1045.

Sheinerman, F.B., Norel, R., and Honig, B. 2000. Electrostatic aspects of protein–protein interactions. *Curr. Opin. Struct. Biol.* **10:** 153–159.

Sitkoff, D., Sharp, K.A., and Honig, B. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **98:** 1978–1988.

Srinivasan, J., Cheatham, T.E., Cieplak, P., Kollman, P.A., and Case, D.A. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **120:** 9401–9409.

Tsai, C.C., Emau, P., Jiang, Y., Tian, B.P., Morton, W.R., Gustafson, K.R., and Boyd, M.R. 2003. Cyanovirin-N as a topical microbicide prevents rectal transmission of SHIV89.6P in macaques. *AIDS Res. Hum. Retroviruses* **19:** 535–541.

Tsai, C.C., Emau, P., Jiang, Y., Agy, M.B., Shattock, R.J., Schmidt, A., Morton, W.R., Gustafson, K.R., and Boyd, M.R. 2004. Cyanovirin-N inhibits AIDS virus infections in vaginal transmission models. *AIDS Res. Hum. Retroviruses* **20:** 11–18.

Yang, F., Bewley, C.A., Louis, J.M., Gustafson, K.R., Boyd, M.R., Gronenborn, A.M., Clore, G.M., and Wlodawer, A. 1999. Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. *J. Mol. Biol.* **288:** 403–412.