Inversion of Receptor Binding Preferences by Mutagenesis: Free Energy Thermodynamic Integration Studies on Sugar Binding to L-Arabinose Binding Proteins†

M. Zacharias,* T. P. Straatsma, and J. A. McCammon

Department of Chemistry, University of Houston, Houston, Texas 77204-5641

F. A. Quiocho

Howard Hughes Medical Institute and Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

Received January 21, 1993; Revised Manuscript Received April 29, 1993

ABSTRACT: The *Escherichia coli* L-arabinose-binding protein (ABP) participates as a specific receptor in the transport of L-arabinose, D-fucose, and D-galactose through the periplasmic space. The wild-type protein binds L-arabinose about 40 times more strongly than D-fucose. A mutation of the protein at position 108 (Met → Leu) causes a specificity change. The Met^{108}Leu ABP slightly prefers binding of D-fucose over L-arabinose. Molecular dynamics (MD) and thermodynamic integration (TI) computer simulations were performed to study the mechanism of sugar discrimination and specificity change based on the known high-resolution X-ray structures. The specificity change was evaluated by calculating the difference in free energy of L-arabinose versus D-fucose bound to wild-type and Met^{108}Leu ABP. The calculated free energy differences are consistent with the experimentally observed specificity of wild-type and Met^{108}Leu ABP. The simulations indicate that the specificity change of Met^{108}Leu ABP is accomplished mainly by reduced Lennard–Jones interactions of residue 108 with L-arabinose and improved interactions with D-fucose. In addition to MD/TI calculations on sugar binding, finite difference Poisson–Boltzmann calculations were performed to identify the most stable ionization state of buried ionizable residues in ABP.

The L-arabinose-binding protein (ABP) serves as a receptor for the transport of sugar molecules through the periplasmic space in Gram-negative bacteria (Furlong, 1987). It belongs to a class of transport proteins with very similar three-dimensional structures and similar functions as primary receptors in chemotaxis. Some other members of this class are the D-glucose/D-galactose binding protein (GGBP), the maltose binding protein (MBP), and the phosphate and sulfate binding proteins (PBP, SBP) (Quiocho, 1986; 1990). ABP contains 306 amino acids, folded into two similar domains each having a core of six-strand β-sheets and two flanking α-helices. The two domains are connected by three peptide segments (Quiocho & Vyas, 1984). The sugar binding site is located in the cleft between the two domains. ABP can bind L-arabinose, D-galactose (with 2-fold reduced affinity), and D-fucose (with 40-fold reduced affinity) (Vermersch et al., 1991). The three sugars differ at the sugar ring atom C5. The methyl group bound to C5 in D-fucose is replaced by a methylene hydroxyl group in case of D-galactose and by a hydrogen in L-arabinose. The high-resolution X-ray structures of the wild-type protein and two mutants in complexes with the three sugars have been determined recently (Quiocho et al., 1989; Vermersch et al., 1990, 1991). These show that sugar binding to the protein is mediated by several hydrogen bonds between sugar hydroxyl groups and polar residues of the protein as well as apolar contacts. At least three water molecules participate in hydrogen bonding of the ligand. Apolar contacts include the interaction of a tryptophan, a phenylalanine, and a methionine residue (Trp^{16}, Phe^{17}, and Met^{108}) with the sugar ring (Quiocho & Vyas, 1984; Quiocho et al., 1989). In one of the mutants the Met residue at position 108, located relatively close to the sugar C5 substituents, has been replaced by Leu (Vermersch et al., 1991). This mutation (Met^{108}Leu) alters the binding affinity and specificity from Ara > Gal ≫ Fuc (wild type) to Gal ≫ Fuc > Ara.

Inspection of the X-ray structures indicates that the Met^{108} Cϵ atom contacts the C5 atom of L-arabinose and moves slightly upon D-fucose or D-galactose binding to optimize the contact to the methyl or methylene hydroxyl group of these ligands (Figure 1). This conformational change is the only significant structural change in the binding pocket of the protein upon binding L-arabinose versus D-fucose or D-galactose. In case of D-galactose binding to ABP, the water structure in the binding pocket changes from that in the L-arabinose complex; one water molecule is replaced by the methylene hydroxyl group of the sugar. The water structure is similar for the other two sugars binding to ABP. Since the Leu^{108} residue in Met^{108}Leu ABP is slightly smaller than a Met residue, it was proposed that an optimal van der Waals contact with the C5 atom of L-arabinose is not possible, which in turn leads to a decrease in affinity (Vermersch et al., 1991). The increase in affinity of the mutant for D-fucose and D-galactose, however, was attributed to the difference in electronegativity of the sulfur atom in methionine compared to carbon in leucine (Vermersch et al., 1991).

Previous computational studies on ABP have concentrated on the structural changes that the protein undergoes upon ligand binding (Mao et al., 1982; Mao & McCammon, 1984). ABP is presumed to adopt an "open" conformation in solution with a wide cleft between the two domains (Mao et al., 1982).

† This work was in part supported by NSF and NCSA grants to J.A.M., R. A. Welch Foundation grants to J.A.M. and F.A.Q., an NIH grant to F.A.Q., and a grant from the W. M. Keck Foundation. M.Z. is a postdoctoral fellow supported by the Deutsche Forschungsgemeinschaft (DFG).

* To whom correspondence should be addressed.

0006-2960/93/0432-7428$04.00/0

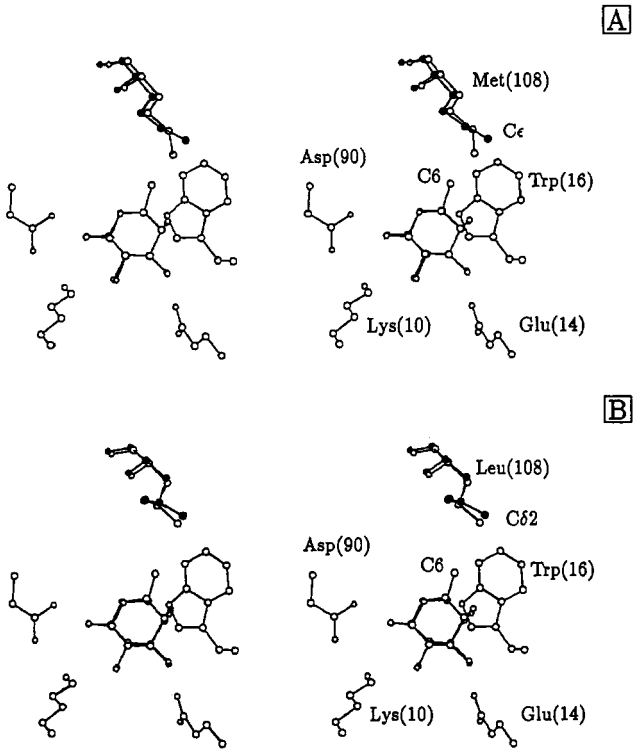

FIGURE 1: Conformation of Met[108] and Leu[108] in the X-ray structures of wild-type (A) and Met[108]Leu ABP (B) complexed with L-arabinose and D-fucose. Atoms of residue 108 in the complexes with D-fucose are represented by filled circles. The C6 atom of D-fucose as well as Met[108] Cε and Leu[108] Cδ2 are labeled. In addition, the figure shows the location of four amino acid side chains (Lys[10], Glu[14], Trp[16], and Asp[90]) which contact the ligand.

In this conformational state the ligand can easily approach the binding site; the protein subsequently undergoes a hinge-bending motion to adopt the close form with a narrow cleft between the domains (Mao & McCammon, 1984).

Here a theoretical study is described on the change in specificity for binding of D-fucose and L-arabinose to Met[108]-Leu ABP compared to wild-type ABP using molecular simulation techniques. The study requires calculations based only on the closed ligand-bound forms of wild-type and Met[108]-Leu ABP, for which high-resolution X-ray structures are available (Quiocho et al., 1989; Vermersch et al., 1991).

It is important to note that the change of receptor specificity by the replacement of a single amino acid is a phenomenon observed not only for ABP but also for other ligand–receptor complexes. One recent example is the discovery that a single amino acid substitution is responsible for the pharmacological distinction of rodent and human 5-hydroxytryptamine (5-HT) receptor (Oksenberg et al., 1992). Molecular dynamics/free energy difference calculations offer the possibility to study in atomic detail the mechanism for sugar discrimination and change in specificity of ABP due to a single amino acid replacement (McCammon & Harvey, 1987; Straatsma & McCammon, 1992).

In the course of a molecular dynamics simulation, a given ligand A bound to a protein is slowly converted into ligand B and the change in free energy of the system is recorded. D-Fucose was transformed into L-arabinose inside wild-type and Met[108]Leu ABP using thermodynamic integration, an established technique for free energy calculations (Beveridge & DiCapua, 1989; Straatsma & McCammon, 1991a). The calculated change in free energy of binding is consistent with experiment, and the calculation allows us to better understand the specificity change of the mutant versus wild-type protein for the two sugar molecules.

COMPUTATIONAL METHODS

Thermodynamic Integration. Thermodynamic integration is a computational technique allowing the calculation of the free energy difference between two states represented by two Hamiltonians, e.g., H_A and H_B. The two states are coupled by a parameter λ in such a way that for $\lambda = 0$ the system is represented by H_A and for $\lambda = 1$ by H_B.

In the course of a molecular dynamics computer simulation, the Hamiltonian of a system initially in state A is slowly converted into a Hamiltonian describing state B by varying λ between 0 and 1. In thermodynamic integration an analytic expression for the derivative of the Hamiltonian with respect to λ is evaluated at each time step of the simulation and integrated over the change in λ. The free energy difference is then given by

$$\Delta G = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda \qquad (1)$$

In a computer simulation, expression 2 for the free energy difference between A and B is approximated by a sum over ensemble averages for the derivative of the Hamiltonian versus λ at discrete steps in λ (Beveridge & DiCapua, 1989):

$$\Delta G = \sum_i \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda_i} \Delta\lambda \qquad (2)$$

In the present paper, thermodynamic integration is performed using the multiconfiguration thermodynamic integration technique (MCTI) (Straatsma & McCammon, 1991b), which offers the possibility of systematically improving the calculations for each step in the control parameter λ. Ensemble averages for the derivative of the Hamiltonian with respect to λ as well as other properties can be evaluated independently for each λ. In addition, statistical errors for the properties evaluated at each step in λ can be calculated (Straatsma & McCammon, 1991b). The accuracy of the free energy difference can be improved for each value of λ by continuing the simulation for the corresponding λ without repeating the whole simulation.

In the free energy simulation of binding of D-fucose and L-arabinose to wild-type and mutant ABP, the thermodynamic cycles shown in Figure 2 were used. The free energy of binding of a ligand to a protein could in principle be determined by directly simulating the binding process. However, the process of ligand binding involves large conformational changes, i.e., diffusion of the ligand to the binding region and closing the ABP structure around the ligand. It is not currently feasible to simulate processes on these time scales by molecular dynamics.

However, it is possible to determine the free energy difference of binding of the two ligands to the two proteins from processes 4 and 8 shown in Figure 2. These artificial transitions, only possible using computer simulations, may involve only small conformational changes in the case of the D-fucose to L-arabinose transition, in which one methyl group is replaced by a hydrogen atom. Processes 4 and 8 may then be efficiently sampled through a molecular simulation.

This approach was used here to calculate the change in specificity of ABP for L-arabinose and D-fucose due to the Met[108]Leu mutation. The change in specificity is defined as the difference in free energy of D-fucose versus L-arabinose comparing the Met[108]Leu ABP complex with wild-type ABP. This quantity is the difference in free energy of processes 8

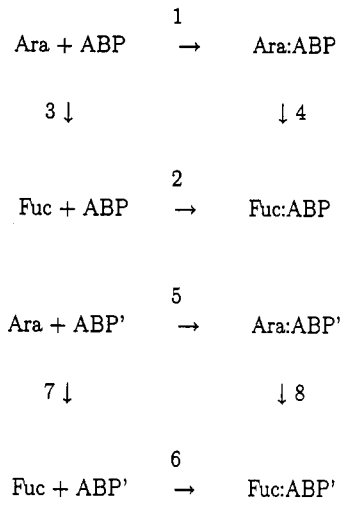

FIGURE 2: Thermodynamic cycles used in the calculation of sugar-binding specificity due to mutagenesis of ABP. Here, Ara represents L-arabinose, Fuc represents D-fucose, ABP represents the L-arabinose-binding protein, and ABP′ represents Met[108]Leu ABP. The quantity of interest is $(\Delta G_6 - \Delta G_5) - (\Delta G_2 - \Delta G_1)$.

and 4 shown in Figure 2. A negative specificity change indicates that the mutation improves binding of D-fucose compared to that of L-arabinose.

Computational Details. The ARGOS program package (Straatsma & McCammon, 1990) was used for all computer simulations, including structural preparation and modification steps. The GROMOS parameter set (van Gunsteren & Berendsen, 1987) was used. The force field consisted of bond, angle, proper dihedral torsion, improper dihedral torsion, Lennard–Jones, and electrostatic terms and had the following functional form:

$$V = \sum_{\text{bonds}} {}^1/_2 K_b (b - b_0)^2 + \sum_{\text{angles}} {}^1/_2 K_\theta (\theta - \theta_0)^2 + \sum_{\text{proper-torsion}} K_\phi [1 + \cos(n\phi - \delta)] + \sum_{\text{improper-torsion}} {}^1/_2 K_\xi (\xi - \xi_0)^2 + \sum_{i<j} \left(\frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^{6}} + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad (3)$$

Here, V is the potential energy of the system. The K_x (with $x = b$, θ, ϕ, and ξ) are appropriate force constants for the different bonded terms. C_{12} and C_6 are repulsive and attractive Lennard–Jones parameters, respectively, r_{ij} is the interatomic distance, ϵ_0 is the dielectric constant in vacuum, and q_i and q_j are the charges on atoms i and j. Water molecules were represented by the SPC/E model (Berendsen et al., 1987). All molecular dynamics/thermodynamic integration calculations were run under constant temperature (293 K) by coupling to an external heat bath (Berendsen et al., 1984). A temperature of 293 K for all simulations was chosen because it corresponds to the temperature used in the experimental studies of ABP. All bond lengths, including the perturbed bond, were constrained to their standard length using SHAKE (Ryckaert et al. 1984), which allowed a time step of 2.0 fs in the MD simulations. The cutoff radius for all nonbonded interactions between solute atoms and solvent atoms was 0.95 nm. Published high-resolution X-ray structures of ABP and Met[108]Leu ABP complexed with L-arabinose were used as starting structures for the molecular dynamics simulations of the sugar mutation inside the protein. Polar hydrogen coordinates were added to the X-ray structures and optimized by energy minimization (100 steps steepest descent minimization) while all other atoms of the structure were kept fixed.

The structure was solvated in a 2.4-nm radius sphere centered around the L-arabinose/D-fucose C5 atom by overlaying the X-ray structure with coordinates of equilibrated SPC/E water molecules and deleting all water molecules closer than 0.25 nm to any protein atom. After elimination of water molecules located in cavities of the protein where hydrogen bond formation is not possible, the water was relaxed (100 steps steepest descent minimization) while the entire protein was kept fixed. Following minimization, the water was equilibrated for 20 ps within an MD simulation.

During this MD run only, waters within 1.8 nm from the sugar C5 atoms were treated dynamically. The rest of the structure was kept fixed. The system was further equilibrated by first relaxing the protein and ligand atoms within 1.8 nm from sugar C5 and then starting the MD of the protein by heating it from 100 to 293 K (in 100-K steps) in 15 ps with velocity reassignment every 0.4 ps.

The full structure, treating all atoms within 1.8 nm from the sugar C5 atom dynamically, was subsequently equilibrated for 30 ps at 293 K. It should be mentioned that a reduction of the dynamically treated sphere to atoms within 1.4 nm from sugar C5 did not affect the TI results significantly.

Thermodynamic integration was performed in 21 × (2000 + 4000) × 0.002 ps = 252 ps. The notation used here is $N_\lambda \times (N_e + N_d) \times \Delta t = T$, where N_λ is the number of steps in λ, N_e and N_d are the number of equilibration and data-gathering time steps, respectively, per λ. T is the simulation time and Δt is the time per MD step. Forward (L-arabinose to D-fucose) and reverse (D-fucose to L-arabinose) thermodynamic integration in both proteins was performed in order to evaluate systematic errors and the hysteresis of all calculations.

In the course of the transformation of the methyl group in D-fucose to a dummy group in L-arabinose, the bond length between the group and the sugar C5 atom was reduced from 0.153 nm initially to 0.07 nm in the final stage of the simulation. Vice versa, the bond length was increased in the reverse simulation. This simulation protocol circumvents the problem of a discontinuity in the perturbation at the initial (or final stage) of the simulation due to the sudden creation of a Lennard–Jones singularity when creating or deleting an atom. The initial onset of the Lennard–Jones interaction (which is the cause of the discontinuity) takes place within the van der Waals sphere of the solute atom connected to the dummy group by an, initially, very short bond. TI without bond sprouting/desprouting can strongly disturb the molecular structure and causes large statistical (systematic) errors at the very beginning or end of the calculation (Straatsma et al., 1992). The same sprouting protocol was used for all thermodynamic integration simulations.

Poisson–Boltzmann Finite Difference Calculations. In order to elucidate the most probable ionization state of some buried and partially buried ionizable groups in ABP and Met[108]Leu ABP, electrostatic energy calculations were performed using the finite difference Poisson–Boltzmann approach as implemented in the UHBD program package (Davis et al., 1991). The method allows the calculation of the electrostatic energy of molecules of arbitrary shape. Water is not treated explicitly but as a high dielectric continuum surrounding the protein molecule, which is in turn treated as a region of low dielectric constant. The molecule is initially mapped onto the center of a grid. The charge on each atom is distributed over the eight neighboring grid points according to a trilinear weighting function (Davis & McCammon, 1991). All grid points interior and exterior to the molecule were

Table I: Electrostatic Stability of Ionizable Side Chains near the Ligand Binding Site

	ionization state[a]							
	1	2	3	4	5	6	7	8
Asp^{89}	–	n	–	n	–	n	–	n
Glu^{20}	–	–	n	n	–	–	n	n
His^{259}	n	n	n	n	+	+	+	+
desolvation	138.5	87.4	84.5	31.4	182.4	128.9	126.4	72.8
Coulomb energy	–20.1	–74.1	–80.3	–47.3	–102.9	–93.3	–128.0	–47.7
pK shift[b]	0.0	20.1	17.1	37.2	0.0	20.1	17.1	37.2
sum of contributions	118.4	33.4	21.3	21.3	79.5	55.7	15.5	62.3
relative to state 4[c]	97.1	12.1	0.0	0.0	58.2	34.4	–5.8	41.0

[a] The electrostatic stability of different ionization states of three amino acid residues (Glu^{20}, Asp^{89}, and His^{259}) located near the sugar binding site was calculated using the finite difference Poisson–Boltzmann approach (see text for details). The charge on each amino acid in each of the eight possible ionization states is indicated by (n) neutral, (+) positive, or (–) negative. Charges were distributed according to the GROMOS parameter set (van Gunsteren & Berendsen, 1987). Electrostatic energies are given in kilojoules per mole. [b] The term pK shift corresponds to the free energy of a neutral state of the ionizable group in aqueous solution for a given pH, $2.3RT(\mathrm{pH} - \mathrm{p}K_{\mathrm{intr}})$, where the $\mathrm{p}K_{\mathrm{intr}}$ corresponds to the pK of the ionizable group in aqueous solution. The stability of each ionization state was calculated for pH = 7.0. The $\mathrm{p}K_{\mathrm{intr}}$ values of Glu and Asp used in this study were 4.0 and 3.5, respectively. Since the $\mathrm{p}K_{\mathrm{intr}}$ of His is almost neutral, we did not include a penalty of neutralization at pH 7.0. [c] For the last row of the table the neutral state of all three ionizable groups was taken as the reference state.

assigned relative dielectric constants of $\epsilon = 4.0$ and $\epsilon = 78.0$, respectively. Since all calculations were performed at zero ion concentration, the PB equation reduces to the Poisson equation. The focusing technique was used to calculate boundary potentials for smaller grids from large grid calculations. By using the focusing technique (Gilson et al., 1987), the boundary potential of a fine grid (100^3, 0.3-Å grid spacing) was calculated from a large grid (100^3, 0.9-Å grid spacing) including the complete protein structure. The potential at the boundary of the large grid was calculated analytically by treating each charged atom as a Debye–Hückel sphere. The electrostatic energy of the system is given by the sum over the charges times the electrostatic potential. Further details of the method are given in Davis et al. (1991) and Gilson et al. (1987).

The stability of each ionization state of an ionizable group can change upon burial of the group inside a protein (Yang & Honig, 1992; Gilson, 1993). The change in stability is caused by electrostatic interactions with charges of the protein or due to desolvation (lower dielectric constant inside a protein) (Basford & Gewert, 1992). The process of burying a group can be separated into three steps illustrated in the following for an acidic residue in its neutral (protonated) form. The first contribution is the free energy of protonating the residue at a given pH in aqueous solution. The free energy of a protonated acidic group is $2.3RT(\mathrm{pH} - \mathrm{p}K_{\mathrm{intr}})$, where T is the temperature, R is the gas constant, and $\mathrm{p}K_{\mathrm{intr}}$ is the pK of the group in aqueous solution. In a second step, the group is buried inside the protein. The free energy change of this desolvation process can be obtained from finite difference PB calculations by taking the difference in electrostatic energy between the amino acid located inside the protein and in a high dielectric environment ($\epsilon = 78.0$). Third, the electrostatic (coulombic) interaction of the buried group can be calculated from the electrostatic energy of the protein charges in the field of the buried group. If one is interested in the influence of different ionization states of other groups on the stability of the buried group, step three needs to be repeated for each set of charges on the protein.

RESULTS AND DISCUSSION

Ionization State of Buried Ionizable Groups. The structures of ABP and $Met^{108}Leu$ ABP contain three ionizable amino acid residues located approximately 0.8 nm away from the ligand binding site. These residues are Glu^{20}, Asp^{89}, and His^{259}. All three side chains are partially (His) or fully (Glu and Asp) excluded from solvent and so could form salt bridges. Initial molecular dynamics calculations showed that the deviation from the X-ray structure depends on the ionization state of these three groups. As described under Computational Methods, the relative stability of each ionization state can be determined by electrostatic energy calculations using the finite difference Poisson–Boltzmann approach. The electrostatic energy of each ionization state of the system of three amino acids was calculated, thereby accounting for contributions due to desolvation and to electrostatic interactions among the three amino acids and with the rest of the protein. Calculations were performed using the wild-type relaxed ABP structure. An overall charge of –1 was assigned to all other Glu and Asp residues and +1 to all Lys and Arg residues. The two other His residues in ABP, which are outside the dynamics region in the MD/TI, were kept neutral.

The approximation is made that the amino acid adopts the same conformation in solution as in the protein. It is possible that the average structure of the amino acid side chain in solution differs from the structure in the protein, which in turn can influence its electrostatic energy. This approximation, however, does not affect the estimation of the relative stabilities of each ionization state.

Taking into account the intrinsic pK's of all three groups, the calculated electrostatic energies allow one to calculate the probability of each ionization state (or relative contribution) at any given pH; in other words, one can calculate the titration curve of each group (Gilson, 1993). However, in the context of our MD/TI calculations only the dominant ionization state for this system of ionizable groups is of interest. Table I shows the result of the finite difference calculations. It turns out that the state with both the His^{259} and Asp^{89} charged and a neutral Glu^{20} is the most stable and therefore predominant ionization state. Note, however, that the states with all three side chains neutral or only Asp^{89} ionized are only about 6 kJ mol^{-1} less stable than the most stable ionization state. Interestingly, in most cases where ionized groups were buried in the protein, the coulombic interaction among the three residues and with the protein does not compensate for the unfavorable desolvation of charged side chains. The electrostatic (coulombic) interaction energy compensates for the loss of solvation only in the case of the completely neutral state or the most stable state, with a salt bridge between a positive His^{259} and a negative Asp^{89}.

Table II: Multiconfiguration Thermodynamic Free Energy Difference Results

	ΔG (kJ mol^{-1})
Wild-Type ABP[a]	
MCTI (Ara/Fuc)	−6.9 (±0.6)
MCTI (Fuc/Ara)	6.4 (±0.5)
Met^{108}Leu ABP	
MCTI (Ara/Fuc)	−15.1 (±0.9)
MCTI (Fuc/Ara)	14.0 (±0.8)
	$\Delta\Delta G$ (kJ mol^{-1})
Specificity Change (Fuc/Ara) Wild Type/Mutant[b]	
experiment[c]	−10.0
MCTI	−7.9 (±1.3)
	ΔU (kJ mol^{-1})
Average Lennard–Jones Interaction of the Sugar with Residue 108[d]	
wild type (Ara)	−6.4 (±1.0)
wild type (Fuc)	−3.8 (±0.7)
Met^{108}Leu mutant (Ara)	−3.1 (±0.6)
Met^{108}Leu mutant (Fuc)	−5.3 (±0.6)

[a] The table shows the calculated free energy differences of D-fucose (Fuc) and L-arabinose (Ara) bound to wild-type and Met^{108}Leu ABP. [b] The change in specificity is defined as the difference in free energy of D-fucose versus L-arabinose comparing the Met^{108}Leu ABP complex with wild-type ABP. [c] The experimentally determined change in specificity was taken from Vermersch et al. (1991). [d] Average Lennard–Jones interactions were calculated from 15-ps MD simulations of wild-type and Met^{108}Leu ABP complexes.

Calculations with a smaller interior relative dielectric constant (ϵ = 2.0) predicted the same ionization state to be the most stable but changed the relative stabilities of other ionization states (data not shown). For all subsequently described molecular dynamics and thermodynamic integration calculations on wild-type and Met^{108}Leu ABP, the most stable ionization state was assumed. It is important to note that this choice resulted in an MD trajectory in close agreement with the X-ray structure. The sugar ligand moved from the position in the X-ray structure by less than 0.05 nm throughout the whole simulation.

Calculated Free Energy Differences. Free energy differences for the L-arabinose to D-fucose transition were calculated in wild-type and Met^{108}Leu ABP. The results for all calculations are given in Table II. The calculated specificity change is here defined as the difference in binding free energy of D-fucose versus L-arabinose in which Met^{108}Leu ABP is compared to wild-type ABP. It amounts to −7.9 (±1.3) kJ mol^{-1}. This result is in reasonable agreement with the corresponding experimentally determined value of −10.0 kJ mol^{-1} (see Table II). The forward and reverse protein simulations showed only a small drift of the calculated free energies after 1500–2500 data-gathering steps, implying that 2000–4000 data-gathering steps per λ are sufficient for the full convergence of the MCTI. The final drift in the calculated free energy was less than 0.5 kJ mol^{-1} ps^{-1} for all MCTIs. Figure 3 shows the change in free energy as a function of the perturbation parameter λ for the forward and reverse thermodynamic integrations. The hysteresis of the calculated free energy changes (the difference between forward and reverse MCTI) was largest for the MCTI on the mutant protein (1.1 kJ mol^{-1}) and was less than 0.5 kJ mol^{-1} for the calculations on wild-type ABP. This result further indicates good convergence of the calculated free energy changes. The derivative of the Hamiltonian describing the sugar–water interaction versus λ (λ derivative) had the same sign in both protein simulations (at each step in λ) and was by a factor of 10 smaller than the corresponding λ derivative for the sugar–protein Hamiltonian (data not shown). This result indicates that water molecules in the binding pocket do not contribute to the change in specificity of Met^{108}Leu ABP compared to wild type.

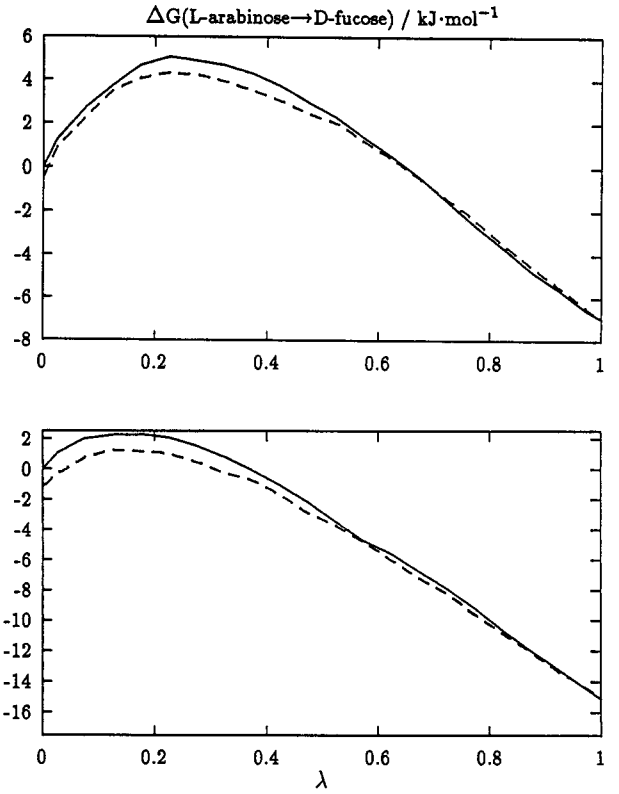

FIGURE 3: Calculated free energy change as a function of the perturbation parameter λ. The upper diagram corresponds to the simulation of wild-type ABP; the lower diagram to the simulation of Met^{108}Leu ABP. The state with λ = 0 corresponds to L-arabinose in complex with ABP/Met^{108}Leu ABP, and λ = 1.0 represents the D-fucose–protein complex. Both forward (solid line) and reverse MCTI results are drawn.

The Lennard–Jones (LJ) interaction of the sugar molecule with Met^{108} and Leu^{108} (in the case of the mutant) was calculated and averaged over a period of 15 ps from separate MD runs of the L-arabinose and D-fucose complexes, respectively. Table II shows that, for wild-type ABP, the average LJ interaction of L-arabinose with Met^{108} is 2.6 kJ mol^{-1} more negative than the interaction with D-fucose. In contrast, in the case of Met^{108}Leu ABP, the ligand interaction with residue 108 favors D-fucose over L-arabinose by about 2.2 kJ mol^{-1}. These results indicate that LJ interactions of sugar ligands with residue 108 contribute significantly to distinguishing between L-arabinose and D-fucose. Note that the methyl group of D-fucose, which is transformed to a dummy during the TI, has no partial charge so that it does not interact electrostatically with the protein.

In addition to the Met/Leu^{108}–sugar interaction, the distance between the sugar C5 atom and Met^{108} Cϵ and Leu^{108} Cδ2, respectively, was recorded. The resulting distance histograms are shown in Figure 4. For the Met^{108}Leu mutant the distance distributions look similar, independent of whether D-fucose or L-arabinose binds to the protein. The average distance between sugar C5 and Leu^{108}Cδ2 over the recorded time period was similar, 0.48 nm (L-arabinose) and 0.49 nm (D-fucose), respectively. An inspection of the X-ray structures of Met^{108}Leu ABP (Vermersch et al., 1991) revealed corresponding distances of 0.48 and 0.52 nm for L-arabinose and D-fucose complexes, respectively. The calculated distance

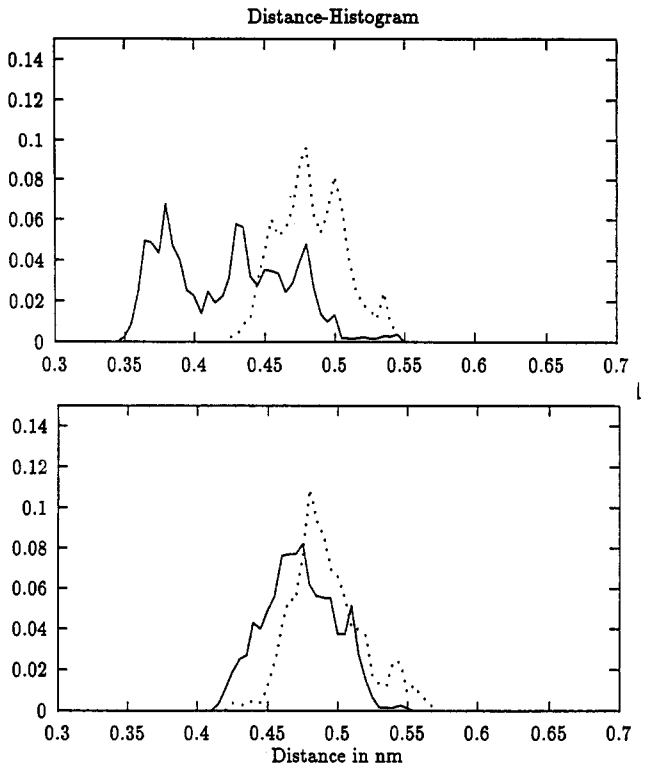

FIGURE 4: Histograms for the distance between the sugar C5 atom and the closest atom of residue 108 for the D-fucose and L-arabinose complexes. The upper histogram corresponds to MD simulations on wild-type ABP in complex with L-arabinose (solid line) and D-fucose (dotted line). The distance between sugar C5 and Met^{108} Cε was recorded at every step of a 15-ps MD simulation. In the case of Met^{108}Leu ABP (lower histogram), the distance between sugar C5 and Leu^{108} Cδ2 was recorded. Data were normalized to the number of records.

histograms for wild-type ABP, however, show significant differences depending on which sugar is bound. The average distance of Met^{108}Cε to the sugar C5 atom in the L-arabinose/ABP complex during the MD simulation was 0.42 nm, smaller by about 0.07 nm than in the D-fucose complex (0.49 nm). The smaller average distance between L-arabinose C5 atom and Met^{108}Cε is consistent with the X-ray study that also showed a smaller distance between the two atoms in the L-arabinose complex. The corresponding distances in the X-ray structures are 0.4 nm (L-arabinose) and 0.5 nm (D-fucose). The result further illustrates our finding that the average LJ interaction favors L-arabinose in wild-type ABP (the Met^{108} side chain conformation adjusts to improve LJ interactions with L-arabinose). It indicates that the Leu^{108} in Met^{108}Leu ABP disfavors L-arabinose compared to D-fucose due to an insufficient adjustment of the shorter, less flexible Leu side chain to improve LJ interactions with L-arabinose.

A second important result, apparent from Figure 4, is that in addition to the shift in average distance, the distance distribution is much broader in the case of L-arabinose bound to wild-type ABP than in the D-fucose complex. A narrow distance histogram indicates less conformational freedom of the Met^{108}Cε atom. A broader histogram, on the other hand, indicates larger conformational flexibility which is entropically favorable. Apparently, the larger conformational freedom only allowed for L-arabinose binding to wild-type ABP might add another (entropic) contribution to its preference as a ligand. The crystallographic *B*-factor, determined in the X-ray structural analysis of proteins, is a measure of the global and local mobility of atoms in the crystal. It is important to note that the crystallographic *B*-factor of the Met^{108}Cε is indeed larger in the L-arabinose complex (0.145 nm^2) than in the D-fucose complex (0.117 nm^2), whereas in the Met^{108}Leu mutant the *B*-factor of Leu^{108}Cδ2 in the L-arabinose complex (0.115 nm^2) is even slightly smaller than in the D-fucose complex (0.131 nm^2). This observation further supports the validity of the MD simulation.

CONCLUSIONS

The calculations described in the present study on the binding free energy difference of D-fucose and L-arabinose to wild-type and Met^{108}Leu ABP are consistent with experimental studies on the system. The Met^{108}Leu mutation in ABP leads to a calculated change in the relative free energy of binding the two sugars that is in almost quantitative agreement with experiment.

The study indicates that the change in specificity of the mutant ABP is largely due to improved LJ contacts between the methyl group of D-fucose and the Leu^{108} side chain compared to the Met^{108} side chain in wild-type ABP. As also pointed out in the X-ray study of the mutant (Vermersch et al., 1991), equally important for the change in specificity seems to be the loss of an apolar contact in case of L-arabinose binding to Met^{108}Leu ABP. The Leu^{108} side chain does not undergo sufficient conformational adjustments to improve LJ interactions to the smaller L-arabinose ligand. In wild-type ABP the average distance between Met^{108}Cε and sugar C5 decreases upon binding of L-arabinose, allowing improved LJ contacts.

The molecular dynamics simulations demonstrate a larger flexibility of the Met^{108} Cε atom in the wild-type L-arabinose complex compared to the D-fucose complex, not observed in the simulation of the mutant. This might add an entropic contribution in favor of the wild-type L-arabinose complex.

It was pointed out in an analysis of the X-ray structure of Met^{108}Leu ABP that the difference in electronic character of the sulfur in the wild-type Met^{108} might be responsible for the change in specificity (Vermersch et al., 1991). Although the calculations do not exclude this possibility, the specificity change can be largely accounted for by a change in LJ interactions.

A second conclusion of the MD/TI study is the finding that water molecules located in the binding cleft do not play a significant role in discriminating between the two ligands and are therefore not responsible for the specificity change. This result is consistent with the X-ray analysis that does not show a significant change in the water structures of the D-fucose or L-arabinose complexes due to the Met^{108}Leu mutation (Vermersch et al., 1991).

In addition to the thermodynamic integration results on L-arabinose versus D-fucose binding to ABP, the present study indicates the importance of controlling and choosing a reasonable charge state of ionizable side chains, in particular of residues which are partially or fully buried in the interior of the protein.

This is usually done in an ad hoc manner which can be very misleading. The attractive electrostatic contribution, for example, upon formation of a salt bridge inside a protein does not necessarily compensate for the loss of solvation. At least an approximate ordering of the energetics of different ionization states can be obtained by finite difference Poisson–Boltzmann calculations which in turn is helpful in identifying the most probable ionization state.

ACKNOWLEDGMENT

We thank Jian Shen, Brock A. Luty, Mike K. Gilson, and William R. Cannon for helpful discussions.

REFERENCES

Bashford, D., & Gewert, K. (1992) *J. Mol. Biol. 224*, 473.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984) *J. Chem. Phys.* *81, 3684.

Berendsen, H. J. C., Postma, J. P. M., & Straatsma, T. P. (1987) *J. Chem. Phys. 91*, 6269.

Beveridge, D. L., & DiCapua, F. M. (1989) *Annu. Rev. Biophys. Biophys. Chem. 18*, 431.

Davis, M. E., & McCammon, J. A. (1991) *J. Comput. Chem. 12*, 909.

Davis, M. E., Madura, J. D., Luty, B. A., & McCammon, J. A. (1991) *Comput. Phys. Commun. 62*, 187.

Furlong, C. E. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., et al., Eds.) pp 768–796, American Society of Microbiology, Washington, DC.

Gilson, M. K. (1993) *Proteins: Struct., Funct., Genet. 15*, 266.

Gilson, M. K., Sharp, K. A., & Honig, B. (1987) *J. Comput. Chem. 9*, 327.

Mao, B., & McCammon, J. A. (1984) *J. Biol. Chem. 259*, 4964.

Mao, B., Pear, M. R., McCammon, J. A., & Quiocho, F. A. (1982) *J. Biol. Chem. 257*, 1131.

McCammon, J. A., & Harvey, S. C. (1987) in *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, England.

Oksenberg, D., Marsters, S. A., O'Dowd, B. F., Jin, H., Havlik, S., Perovtka, J. S., & Ashkenazi, A. (1992) *Nature 360*, 161.

Quiocho, F. A. (1986) *Annu. Rev. Biochem. 55*, 287.

Quiocho, F. A. (1990) *Philos. Trans. R. Soc. London 326*, 341.

Quiocho, F. A., & Vyas, N. K. (1984) *Nature 310*, 381.

Quiocho, F. A., Wilson, D. K., & Vyas, N. K. (1989) *Nature 340*, 404.

Ryckaert, J., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327.

Straatsma, T. P., & McCammon, J. A. (1990) *J. Comput. Chem. 11*, 943.

Straatsma, T. P., & McCammon, J. A. (1991a) *Methods Enzymol. 202*, 497.

Straatsma, T. P., & McCammon, J. A. (1991b) *J. Chem. Phys. 95*, 1175.

Straatsma, T. P., & McCammon, J. A. (1992) *Annu. Rev. Phys. Chem. 43*, 407.

Straatsma, T. P., Zacharias, M., & McCammon, J. A. (1992) *Chem. Phys. Lett. 196*, 297.

van Gunsteren, W. F., & Berendsen, H. J. C. (1987) Groningen Molecular Simulation (GROMOS) Library Manual, BIOMOS B. V., Nijenborgh 16, 9747 AG Groningen, The Netherlands.

Vermersch, P. S., Tesmer, J. J. G., Lemon, D. D., & Quiocho, F. A. (1990) *J. Biol. Chem. 265*, 16592.

Vermersch, P. S., Lemon, D. D., Tesmer, J. J. G., & Quiocho, F. A. (1991) *Biochemistry 30*, 6861.

Yang, A. S., & Honig, B. (1992) *Curr. Opin. Struct. Biol. 2*, 40.