Available online at www.sciencedirect.com
SCIENCE DIRECT
Carbohydrate Research 341 (2006) 2233–2245
Carbohydrate RESEARCH

Docking studies on glycoside hydrolase Family 47 endoplasmic reticulum α-(1→2)-mannosidase I to elucidate the pathway to the substrate transition state

Chandrika Mulakala,[a] Wim Nerinckx[b] and Peter J. Reilly[a,*]

[a]*Department of Chemical and Biological Engineering, 2114 Sweeney Hall, Iowa State University, Ames, IA 50011, USA*
[b]*Department of Biochemistry, Physiology and Microbiology, Ghent University, 9000 Ghent, Belgium*

Received 30 January 2006; received in revised form 3 May 2006; accepted 11 May 2006
Available online 27 June 2006

Abstract—α-(1→2)-Mannosidase I from the endoplasmic reticulum (ERManI), a Family 47 glycoside hydrolase, is a key enzyme in the *N*-glycan synthesis pathway. Catalytic-domain crystal structures of yeast and human ERMan1s have been determined, the former with a hydrolytic product and the latter without ligands, with the inhibitors 1-deoxymannojirimycin and kifunensine, and with a thiodisaccharide substrate analog. Both inhibitors were bound at the base of the funnel-shaped active site as the unusual $^{1}C_{4}$ conformer, while the substrate analog glycon is a $^{3}S_{1}$ conformer. In the current study, AutoDock was used to dock α-D-mannopyranosyl-(1→2)-α-D-mannopyranose with its glycon in chair ($^{1}C_{4}$, $^{4}C_{1}$), half-chair ($^{3}H_{2}$, $^{3}H_{4}$, $^{4}H_{3}$), skew-boat ($^{O}S_{2}$, $^{3}S_{1}$, $^{5}S_{1}$), boat ($^{2,5}B$, $^{3,O}B$, $B_{1,4}$, $B_{2,5}$), and envelope (^{3}E, ^{4}E, E_{3}, E_{4}) conformations into the yeast ERManI active site. Both docked energies and forces on docked ligand atoms were calculated to determine how the ligand distorts to the transition state. From these, we can conclude that (1) both $^{1}C_{4}$ and $^{O}S_{2}$ can be the starting conformers; (2) the most likely binding pathway is $^{1}C_{4} \rightarrow {}^{3}H_{2} \rightarrow {}^{O}S_{2} \rightarrow {}^{3,O}B \rightarrow {}^{3}S_{1} \rightarrow {}^{3}E$; (3) the transition state is likely to be close to a ^{3}E conformation.

Keywords: AutoDock; Carbohydrate conformation; Docking; Enzyme mechanism; GH47; Mannosidase; Structure–function relationship; Transition state

1. Introduction

N-Glycan synthesis in eukaryotes begins in the endoplasmic reticulum (ER) with the transfer of a pre-formed oligosaccharide precursor, usually $Glc_3Man_9GlcNAc_2$, from dolichyl phosphate to an Asn/X/Ser(Thr) sequence on newly synthesized polypeptides. Various ER glycosidases trim $Glc_3Man_9GlcNAc_2$ to $Man_8GlcNAc_2$. Specifically, ER α-(1→2)-mannosidase I (ERManI) produces Isomer B of $Man_8GlcNAc_2$ from $Man_9GlcNAc_2$ by removing a mannosyl residue from the middle branch of the oligosaccharide structure.[1–3]

ERManIs belong to glycoside hydrolase Family 47[4] (GH47) by sequence homology. Catalytic-domain crystal structures of human[5,6] and *Saccharomyces cerevisiae*[7] ERManIs have been determined; they have an unusual $(\alpha,\alpha)_7$-barrel structure, with a C-terminal β-hairpin protruding into the barrel from one side to plug it. The other end of the barrel forms a ~25 Å-wide funnel-shaped cavity that narrows to ~10 Å at the funnel-tube neck. A Ca^{2+} ion sits at its end.

In the *S. cerevisiae* ERManI crystal structure, the hydrolytic product $Man_8GlcNAc_2$ isomer B, N-linked to one protein molecule, extends into the barrel of the adjacent symmetry-related molecule, interacting with its active site.[7] Human ERManI structures have been determined with no ligands,[5] with the inhibitors 1-deoxymannojirimycin (DMJ) and kifunensine (KIF),[5] and with the thiosaccharide substrate analog methyl

* Corresponding author. Tel.: +1 515 294 5968; fax: +1 515 294 2689; e-mail: reilly@iastate.edu

0008-6215/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2006.05.011

2-*S*-(α-D-mannopyranosyl)-2-thio-α-D-mannopyranoside (S-Man_2).[6] Both inhibitors bind at the base of the funnel-shaped active site in the unusual $^1C_4$ conformation, with their positions corresponding to the expected position of the middle-arm mannosyl residue of $Man_9GlcNAc_2$ had it not been hydrolyzed in the *S. cerevisiae* ERManI crystal structure. The S-Man_2 glycon is found in the novel $^3S_1$ conformation.

The combined information of the human and yeast structures can be extrapolated through computational docking to predict the bound conformation of the substrate $Man_9GlcNAc_2$ in the yeast ERManI active site. We used AutoDock,[8] a small-molecule docking program, to do this.[9] Docking α-D-mannopyranosyl-(1→2)-α-D-mannopyranose (Man_2) in the active site helped to establish the identity of the catalytic base (Glu435) of this enzyme, whose assignment was ambiguous due to the absence of the middle-arm mannosyl residue in the yeast ERManI crystal structure. The E_4 conformation of the glycon mannosyl residue was predicted as the transition state (TS) based on the available structural information of the ligands in the human and yeast crystal structures. However, this conformation has been suggested as an unlikely TS candidate due to a syn–axial interaction between the bulky C-5 hydroxymethyl group and the C-3 hydroxyl group,[10] the $^3H_4$ conformation instead being a better candidate.[6,10] Also, stereoelectronic theory requires that the scissile glycosidic C-1–O-G bond be antiperiplanar to a lone pair of electrons on the ring oxygen atom for subsequent ring distortion to the TS.[11] Since the glycosidic bond is equatorial in the $^1C_4$ conformer and is not antiperiplanar to the ring oxygen lone pair, the glycon would have to adopt a skew-boat conformation that agrees with the antiperiplanar requirement en route to the TS. The aim of the present work, therefore, is twofold—to use computational docking by AutoDock to establish the TS conformation, and to also determine the conformational itinerary of the glycon as it progresses to the TS.

AutoDock predicts where a ligand binds on the surface of a macromolecule, such as a protein or DNA, whose tertiary structure is known.[8] AutoDock treats the macromolecule as rigid, while the ligand is allowed torsional flexibility. Although conformational changes are often observed upon ligand binding to enzymes, the treatment in this case is a reasonable one, since human and yeast ERManI active sites are practically identical, and as binding of DMJ and KIF in the human ERManI active site causes insignificant side-chain rearrangements in their respective crystal structures.[5]

AutoDock computes the nonbonded interaction energy between ligand and macromolecule, the problem therefore being searching the ligand conformational space in the vicinity of the macromolecule to find the conformation with the lowest interaction energy. The AutoDock suite provides four different ways to search this conformational space: simulated annealing algorithm (SAA), genetic algorithm (GA), Lamarckian genetic algorithm (LGA), and local search (LS), the last based on the Solis and Wets (SW) method.[12]

The SAA is slow[8] and therefore it was not employed in this work. The GA is based on the Darwinian principles of selection, random mutation, and crossover. In the AutoDock implementation of the GA, the genes are a string of real values representing the translation, orientation, and torsional angles for the various ligand torsions. An initial random population with a user-defined number of ligand conformations is generated, and this is then subjected to selection, mutation, and crossover, resulting in a new population constituting the next generation of individuals. The process is repeated over a user-defined number of generations, and the individual with the lowest binding energy is finally reported by AutoDock. The LGA is an extension of the GA, and is so called because the processes of selection, mutation, and crossover on every generation of population are followed by a LS, and the changes due to the LS are inherited by the next generation. Of the SAA, GA, and LGA, the last is the most efficient in searching the ligand conformational space for the best docking energy.[8]

Ab initio studies on model compounds[13,14] and on sugar and sugar analogs in solution,[15,16] as well as under specific enzymatic configurations,[17,18] have helped to establish many important aspects of the glycolysis reaction, such as the nature of the oxocarbenium cation that forms the TS as well as the ring distortions that accompany this transition. Since AutoDock primarily docks by packing van der Waals spheres between a conformationally fixed sugar ligand (torsional changes are allowed in the hydroxyl groups and the C-5 carboxymethyl group, but the ring pucker remains unchanged) and a rigid enzyme, it cannot model the continuous transition of the substrate to the TS. However, the reduced computational expense compared to ab initio studies allows us to easily dock many putative intermediates in the conformational itinerary to the TS, to aid in understanding the glycolysis mechanism.

In addition to the ERManI study, we have used AutoDock to understand enzyme structure–function relationships in glucoamylase,[19] β-amylase,[20] surfactant protein D,[21] and phospholipase D.[22] Recently, substrate binding energies on docked substrates were complemented with computed forces on substrate atoms in crystal structures of cellulases Cel7A and Cel7B.[23,24] The forces give insights on substrate dynamics in the active site, which cannot be inferred from the binding energies that AutoDock generally outputs.

In the present work, Man_2 with 16 different glycon mannosyl conformers was docked in the yeast ERManI (PDB 1DL2) active site, and docked energies (E_{Total}) as well as forces were determined to gain further insights

into its function. The 16 conformers were chosen as follows: Eight ($^{2,5}B, B_{2,5}, {}^{3}E, E_3, {}^{4}E, E_4, {}^{3}H_4$, and $^{4}H_3$) satisfy the requirement that the C-2, C-1, O-5, and C-5 atoms be planar for formation of the oxocarbenium ion. Even though ERManI was expected to act on a TS with either an E_4 or a $^{3}H_4$ glycon conformation, all eight were docked to test the reliability of the computational method. Another conformer, $^{O}S_2$, of comparable steric energy in solution to the $^{1}C_4$ conformer,[25] which also should clear the narrow opening of the ERManI active site,[9] was also docked as a possible starting conformation for the hydrolysis reaction. Three conformers, $^{O}S_2$, $^{3}S_1$, and $^{5}S_1$, satisfy the antiperiplanar requirement[11] for formation of the oxocarbenium ion TS. Therefore, the pseudorotational series $^{O}S_2$, $^{3,O}B$, $^{3}S_1$, $B_{1,4}$, and $^{5}S_1$ was also docked. In addition, the $^{3}H_2$ conformer was docked, as it is the most likely intermediate for the transformation of $^{1}C_4$ to a skew-boat conformation. The low-energy $^{4}C_1$ and $^{1}C_4$ conformers were also docked. Since the inhibitors DMJ and KIF bind at the base of the funnel-shaped active site of human ERManI in the unusual $^{1}C_4$ conformation,[5] and the yeast and human forms of this enzyme have practically identical active sites,[5] these two inhibitors were docked into the yeast ERManI active site in this form.

2. Theory and procedures

2.1. Iterative minimization

As the different docked ligand conformers have identical chemical compositions, and as some of these conformers are quite close to each other in the spatial arrangement of their constituent atoms, the search for the optimal docked conformer had to be very rigorous to ensure that it was as close as possible to the global minimum and was not in a local minimum. This was accomplished through the iterative minimization protocol outlined in Figure 1. The starting positions of the docked Man_2 ligands were obtained by superimposing their aglycons on the middle-arm mannosyl residue, Man605, of the enzyme product, $Man_8GlcNAc_2$, in the crystal structure of yeast ERManI. This starting conformer was subjected to three rounds of docking using the LGA and another three rounds using the pseudo-Solis and Wets (pSW) LS algorithm, which is a modified implementation of the SWLS method. The lowest energy conformers of these six docking runs were then subjected to further minimization using an iterative SWLS procedure, in which the best docked conformation of an LS run is used as the starting conformation for the next LS iteration. Thirty iterations were conducted in an iterative minimization run, which was performed five times on the best docked conformations of each of the six LGA and pSW runs (Fig. 1).

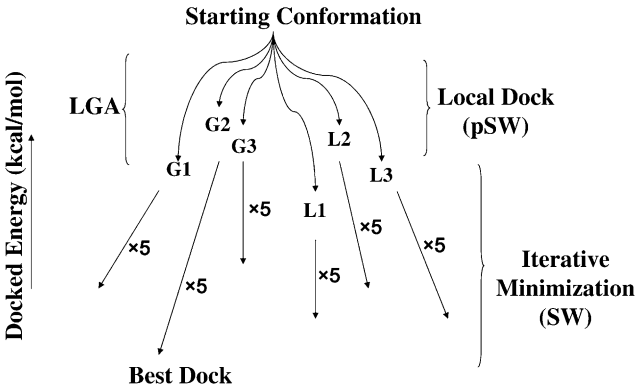

Figure 1. Iterative minimization protocol. In this example, three different starting conformations are subjected to the LGA, and another three starting conformations are subjected to local minimization by a pSW technique. Of the many resulting conformations, the ones from each starting conformation with the most negative docked energies are then each subjected five times to 30 rounds of iterative minimization by the SW technique.

The pyranosyl ring structure results in high densities of hydroxyl groups capable of forming hydrogen bonds in protein–carbohydrate interactions. Also important are hydrophobic stacking interactions, often with aromatic residues.[26] Since hydrogen-bonding interactions are an order of magnitude stronger than van der Waals interactions and have directional dependence, thorough local optimization of the ligand is essential to accurately predict docked energies. This is evident from the results of our iterative minimization protocol. For example, the decrease of docked energy with progressive iteration for the best of the five iterative minimizations for each of the six starting conformations of docked Man_2-$^{1}C_4$ is plotted in Figure 2. Docked energies for all six starting conformations (Fig. 1) decrease significantly upon minimization, with energies falling most steeply in the first five iterations. Also, not all ligands reach the same minimum, necessitating several starting points for the docked ligands to increase the probability of finding the global minimum. For carbohydrate ligand docking with AutoDock, therefore, a thorough iterative local minimization is recommended for promising candidates obtained from a global docking procedure for an accurate prediction of the docked energy. In the present case, since the docked ligands are all the same molecules with different glycon conformations, the difference in ligand docked energy is expected to be small; approaching as close as possible to the global minimum is therefore very important for comparing the different docking results.

As a control, we docked the $^{1}C_4$ forms of DMJ and KIF[5] into the yeast ERManI active site[7] using the iterative minimization protocol, giving docked energies of −97.81 and −116.50 kcal/mol and RMSDs of 0.26 and 0.40 Å from their crystal-structure positions, respectively. The iterative minimization protocol yielded improved docking positions over our earlier results for

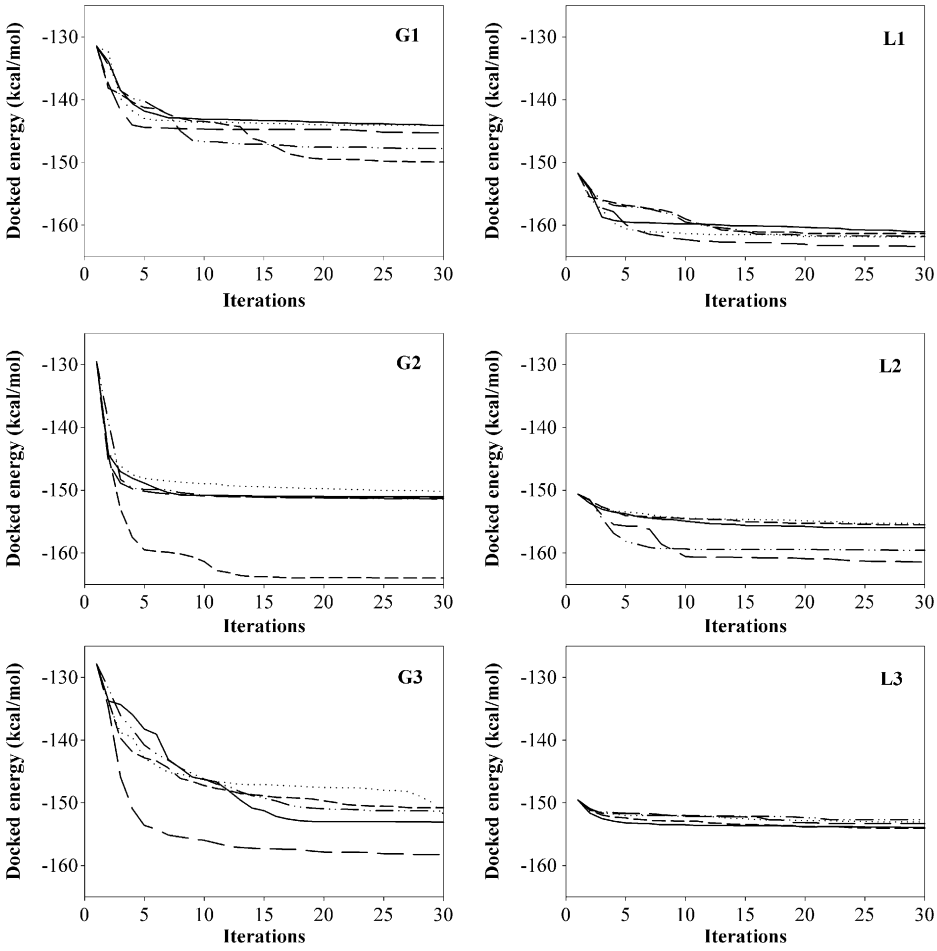

Figure 2. Changes of E_{Total} with successive iterative minimizations of α-Man$_2$-$^1C_4$.

yeast ERManI,[9] as RMSDs then were 0.62 Å for DMJ and 0.72 Å for KIF. Also, Man$_2$-$^3S_1$ docked into the yeast ERManI active site compares very well with the crystal-structure S-Man$_2$ in the human ERManI active site, whose glycon is also in the $^3S_1$ conformation,[6] with an RMSD between the two of 0.72 Å. This is an especially good agreement, since their glycosidic bond angles and lengths are slightly different (115° and 1.4 Å for Man$_2$-$^3S_1$ vs 106° and 1.8 Å for S-Man$_2$). No water molecules were included in docking Man$_2$, since all the crystal-structure ligands were accurately reproduced without inclusion of any crystal-structure water molecules. The docked energies for the 16 docked Man$_2$ conformers are shown in Table 1. The best docked conformation can result from any one of the six starting conformations, implying that most of the ligands were trapped in local minima, further emphasizing the value of different starting points.

2.2. Calculation of docked energies and RMSDs

The docked ligand energy (E_{Total}) reported by AutoDock is a sum of the enzyme–ligand interaction energy (E_{Inter}) and the ligand internal energy (E_{Intra}). Since torsional flexibility is modeled in the ligand by assigning random changes to torsions, evaluation of E_{Intra} is necessary so that AutoDock can energetically penalize ligands generated with unreasonable geometries. AutoDock evaluates both E_{Inter} and E_{Intra} as a sum of van der Waals and electrostatic nonbonded interaction energies. Although this simple energy function formulation is essential to keep energy computation inexpensive, it is insufficient to accurately estimate carbohydrate ligand internal energies, which require a more elaborate energy function to capture hydrogen-bond geometries and exo-anomeric affects.[27] Also, there is no way to accurately determine E_{Intra} consistent with AutoDock's method for determining E_{Inter}. This hampers a direct comparison of the different docked ligand conformations based on E_{Total}. However, values of E_{Inter}, a measure of enzyme–ligand complementarity for docked ligands with different ring puckering, can be directly compared, and therefore we have done that here (Table 1). This is reasonable, since all the docked conformers have the same chemical composition, and E_{Inter} is a sum of the interaction energies of each ligand atom with the enzyme.

Table 1. E_{Total} and E_{Inter} values (kcal/mol) and RMSDs (Å) for optimal conformations

Conformer	E_{Total} from global (Lamarckian) search			E_{Total} from local search			E_{Total} from local iteration on global search			E_{Total} from local iteration on local search			E_{Inter}	RMSD
	G1	G2	G3	L1	L2	L3	1	2	3	1	2	3		
$B_{1,4}$	−144.12	−134.00	−133.83	−148.81	−148.15	−146.23	**−170.43**[a]	−157.60	−167.99	−158.69	−157.09	−161.86	−126.26	1.95
$^{2,5}B$	−153.20	−147.84	−145.13	−163.89	−163.42	−162.83	−171.61	**−177.18**	−167.54	−175.92	−174.04	−171.50	−140.37	1.87
$B_{2,5}$	−140.14	−124.75	−124.15	−144.92	−138.92	−136.20	**−160.61**	−148.53	−144.55	−150.16	−151.70	−146.62	−128.60	1.57
$^{3,O}B$	−155.40	−148.66	−129.17	−163.86	−158.78	−155.83	−165.10	**−170.77**	−165.21	−167.33	−164.48	−161.54	−138.85	1.63
$^{1}C_4$	−131.45	−129.52	−127.85	−151.66	−150.62	−149.52	−149.95	**−164.01**	−158.41	−163.46	−161.45	−154.17	−123.78	1.61
$^{4}C_1$	−131.78	−126.09	−118.52	−146.69	−144.93	−144.66	−147.20	−160.16	−136.53	**−161.78**	−153.04	−155.55	−130.51	2.37
^{3}E	−151.89	−148.02	−145.80	−153.32	−152.47	−151.86	−172.54	−159.71	**−176.57**	−156.82	−157.28	−156.78	**−147.16**	0.79
E_3	−134.77	−119.63	−115.80	−151.70	−149.23	−148.97	−149.28	−149.34	−137.03	−155.62	−159.98	**−161.93**	−134.65	1.98
^{4}E	−141.74	−136.84	−134.92	−145.77	−141.60	−138.87	−150.17	−158.00	−144.74	−154.14	**−160.18**	−152.74	−119.40	2.82
E_4	−143.41	−142.71	−142.43	−170.67	−169.99	−169.73	−173.04	−168.22	−174.57	−174.50	−174.68	**−175.32**	−146.60	2.43
$^{3}H_2$	−137.39	−136.59	−135.69	−151.32	−150.90	−150.76	−152.56	−148.90	−151.07	−153.29	−155.64	**−156.60**	−137.30	1.64
$^{3}H_4$	−135.28	−132.14	−131.46	−150.18	−149.84	−148.68	−151.96	−146.65	−148.74	−151.98	**−154.32**	−152.17	−136.37	1.28
$^{4}H_3$	−121.92	−121.70	−116.53	−148.37	−144.50	−144.44	−132.12	−148.01	−129.18	**−161.50**	−151.15	−150.36	−134.09	2.09
$^{3}S_1$	−129.64	−125.26	−116.15	−138.79	−136.56	−134.91	−145.24	−143.60	−144.55	−141.82	−139.35	**−151.63**	−140.20	1.07
$^{5}S_1$	−153.61	−152.07	−150.71	−155.86	−153.30	−151.41	−171.07	**−172.27**	−167.22	−167.07	−161.92	−167.82	−128.97	1.86
$^{O}S_2$	−162.33	−147.54	−133.77	−153.50	−152.55	−152.16	**−168.62**	−161.51	−154.88	−155.84	−156.20	−155.85	−138.44	1.63

[a] Bolded numerals signify the most negative values of E_{Total} and E_{Inter}.

Any true comparison of docked conformers has to be based on E_{Total}, since rankings of conformers could be affected by decreases in E_{Inter} compensated for by increases in E_{Intra}. We considered obtaining E_{Intra} values for the final docked conformers more accurately using the MM3 molecular mechanics force field or through quantum mechanical calculations and then adding those values, appropriately scaled, to AutoDock's E_{Inter} values. However, different scaling schemes alter the relative energies of the various docked conformers and therefore their relative rankings, and there seems to be no way of validating any chosen scaling scheme. A detailed computation of E_{Intra} values for α-Man is available as MM3 isoenergy contour maps on a two-dimensional Cremer–Pople space[25] (Fig. 3). Although these values cannot substitute for AutoDock's E_{Intra} values, the isoenergy map has important information about α-Man conformational energies, such as the $^{4}C_1$ conformation as the global minimum and the existence of two comparable local energy minima for the $^{1}C_4$ and $^{O}S_2$ conformations. It also displays the geographical features of the isoenergy surface; the hills and valleys help to identify probable low-energy conformational pathways of the ligand en route to the TS.

To circumvent the absence of accurate E_{Intra} values with which to compute E_{Total} values, we have compared the docked conformers on the basis of their E_{Inter} values along with their RMSDs from crystal-structure ligands. Since DMJ, KIF, and S-Man$_2$ dock very closely to each other (within 0.59 Å RMSD of each other for superimposable atoms), we believe that our predicted TS should also dock very closely to them, and therefore we consider RMSD values to be very important for drawing conclusions about the TS conformation.

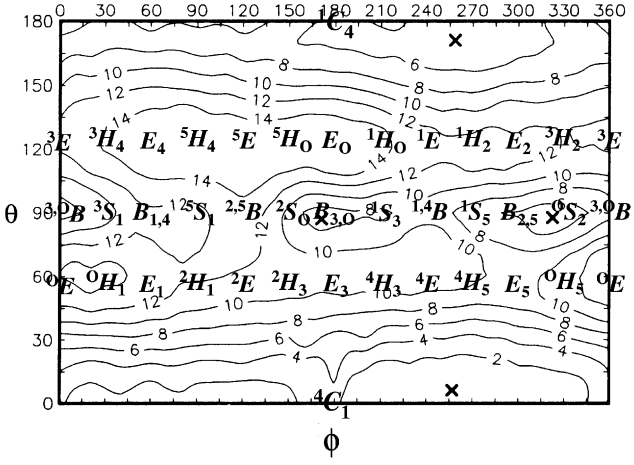

Figure 3. Cremer–Pople isoenergy contour map for α-Man, showing E_{Intra} values in kcal/mol and positions of different conformers.[25] ×'s show local minima, with the global minimum being located near $\theta = 5°$, $\phi = 255°$.

The RMSDs were computed from crystal-structure DMJ for their glycon residues and from the middle-arm mannosyl residue, Man605, of $Man_8GlcNAc_2$ product found in the yeast ERMan1 crystal structure for their aglycon residues. We consider Man605 to be a better basis for aglycon RMSD calculations than crystal-structure S-Man_2 because docked Man_2 represents the terminal middle-arm residues of $Man_9GlcNAc_2$, ERManI's natural substrate. Therefore, we expect that the −1 and +1 subsites of the ERManI active site would accommodate the Man_2 TS conformer without significant deviation of its aglycon residue from crystal-structure Man605, as a significant deviation would further strain the substrate and would lead to further energy expenditure, which cannot be accounted for by docked Man_2. A further consideration favoring the Man605/DMJ combination over S-Man_2 as a basis for RMSD computation is that the glycosidic angle and bond lengths of S-Man_2 differ from those of Man_2. Conformers along the predicted pathway to the TS should therefore exhibit decreasing E_{Inter} and RMSD values, should comply with the antiperiplanar lone-pair hypothesis of Deslongchamps,[11] and should be on a low-energy route on the MM3 isosurface, a condition necessary for efficient enzyme operation.

It should be noted here that self-consistent Lennard–Jones potentials of AutoDock 1.0 before multiplication by free-energy model coefficients were used to evaluate nonbonded interaction energies, since they best reproduce the crystal-structure ligands and also because unweighted parameters are necessary to calculate forces on docked ligands. The E_{Inter} and E_{Intra} values reported here therefore represent binding enthalpies and not binding free energies.

2.3. Transition-state pathway

Dowd et al.[25] reported MM3 isoenergy contour maps for ring conformations of various aldopyranoses on a two-dimensional Cremer–Pople θ–ϕ space, where θ and ϕ represent the relative orientation of puckering about the ring. Ideal chair forms are at $\theta = 0°$ and 180°, with boat and skew-boat forms at $\theta = 90°$ and envelope and half-chair forms at $\theta = 60°$ and 120°. The isoenergy contour lines based on the low-energy $^4C_1$ conformation on the Cremer–Pople space for α-D-mannopyranose are reproduced here for convenience (Fig. 3).

In human ERManI crystal structures, the mannosyl analog DMJ is bound in the active site in the unusual $^1C_4$ conformation.[5] For α-D-mannopyranose, the $^1C_4$ and $^OS_2$ conformations have MM3 steric energies of 4.38 and 4.24 kcal/mol, respectively, higher than that of the $^4C_1$ conformation.[25] This implies that their equilibrium concentrations in solution at 298 K would be ~0.06% of the total α-D-mannopyranose concentration in solution. In a previous docking study,[9] we suggested that due to the narrowness of the neck of the funnel-like active-site opening, the $^4C_1$ conformation with a predominantly equatorial orientation of its ring substituents would be too wide to enter the narrow active site, and therefore the enzyme selectively binds the higher-energy $^1C_4$ conformation, which can enter it, as may the $^OS_2$ conformer. Also, since the terminal mannosyl residue of the $Man_9GlcNAc_2$ substrate cleaved by ERManI is attached to a group of greater bulk, its angle of entry is determined by its C-1 substituent orientation. If axial (as in $^4C_1$), the ring has to enter the neck breadth-wise, while if equatorial ($^1C_4$) or isoclinal ($^OS_2$), the ring can ease sideways into the active-site opening. Both $^1C_4$ and $^OS_2$ were therefore examined as the possible starting conformers bound selectively by this enzyme.

Stereoelectronic theory suggests that a β-chair ($^1C_4$) would have to undergo a conformational change to a β-skew-boat, where the ring oxygen has an electron lone pair antiperiplanar to the scissile bond, to reach the TS.[11] For α-D-mannopyranose, three skew-boat conformations, $^OS_2$, $^3S_1$ and $^5S_1$, satisfy this requirement. Since S-Man_2 crystallized with its glycon in the $^3S_1$ conformation in human ERManI,[6] this conformer is the more likely skew-boat conformation adopted by the mannosyl substrate before twisting to the TS.

As both $^OS_2$ and $^1C_4$ are likely starting conformations bound by ERManI, it is necessary to determine the low-energy pseudorotational pathway from these to the $^3S_1$ or $^5S_1$ conformers, which can be determined from the MM3 isocontour map (Fig. 3). The transition from $^OS_2$ to $^3S_1$ can occur via the $^{3,O}B$ conformer. There are two possible routes for the transition of $^1C_4$ to $^3S_1$—via $^3H_4$ or via $^3H_2$ to $^OS_2$, which then flips to $^3S_1$ via $^{3,O}B$. The indirect pathway via $^3H_2$ has a ~2 kcal/mol lower energy difference compared to the $^3H_4$ pathway (Fig. 3) and may therefore be preferred. The transition from $^3S_1$ to $^5S_1$ occurs via $B_{1,4}$. Also possible is the conversion of $^OS_2$ directly to the TS, since it meets the Deslongchamps antiperiplanarity requirement.[11] This, however, is less likely, since there is no experimental evidence suggesting it, such as the crystal-structure $^3S_1$ conformation found by Karaveg et al.[6]

Substrate hydrolysis in ERManI occurs with configurational inversion at the anomeric center[1] via a borderline S_N1–S_N2 process with a TS geometry whose C-1 and O-5 atoms must be sp^2 hybridized,[11] causing the sugar-ring C-5, O-5, C-1, and C-2 atoms to be in the same plane. Eight conformers satisfy this requirement—$^{2,5}B$, $B_{2,5}$, $^3H_4$, $^4H_3$, ^3E, E_3, ^4E, and E_4. Two of these, E_4 and $^3H_4$,[6,26] have been previously predicted as TSs based on the principle of least motion. All these conformers were therefore docked to determine the conformational itinerary of the substrate to the TS.

3. Results and discussion

3.1. Energy and RMSD values

E_{Inter} and RMSD values, the latter from crystal-structure DMJ and Man605[5,7] for the glycon and aglycon, respectively, were compared for the 16 conformations optimally docked based on their E_{Total} values. The docked conformers can broadly be classified into two groups: those that are likely to be involved in the α-glycoside ($^4C_1$) TS pathway ($^4C_1, {}^4H_3, {}^4E, E_3$, and $B_{2,5}$) (Set A), and those in the β-glycoside ($^1C_4$) pathway ($^1C_4, {}^{2,5}B, {}^3E, {}^3H_2, {}^3H_4, {}^{3,O}B, {}^3S_1, {}^5S_1, B_{1,4}, E_4$, and $^OS_2$) (Set B). Set B conformers nearly always dock with lower energies and RMSDs than the Set A conformers (Fig. 4a), ruling out the $^4C_1$ conformer as a possible starting conformation. The TS appears to be a ^3E conformer, based on its low combination of E_{Inter} and RMSD values. Also, based on decreases in both E_{Inter} and RMSD values combined with the MM3 isoenergy maps of Figure 3, three different pathways are possible, $^1C_4 \rightarrow {}^3H_2 \rightarrow {}^OS_2 \rightarrow {}^{3,O}B \rightarrow {}^3S_1 \rightarrow {}^3E$, $^1C_4 \rightarrow {}^3H_2 \rightarrow {}^OS_2 \rightarrow {}^3E$, and $^1C_4 \rightarrow {}^3H_4 \rightarrow {}^3S_1 \rightarrow {}^3E$ (Fig. 4). Since only E_{Inter} values were used to compare docked conformers, and since AutoDock's evaluation method for E_{Intra} is not accurate, we also retrieved docked conformers with the lowest E_{Inter} values, which in most cases were not the conformers with the lowest E_{Total} values. These E_{Inter} values and their RMSDs from the crystal structures were also plotted (Fig. 4b). General trends in energies and RMSDs for Set B conformers are the same in both cases, and again Set B conformers dock better than Set A conformers. For conformers predicted to be in the three possible TS pathways, E_{Inter} differences between conformers with the lowest E_{Total} and E_{Inter} values are <1 kcal/mol for $^3H_2$, $^OS_2$, and E_4, 2.37 kcal/mol for $^{3,O}B$, and 3.74 kcal/mol for $^3S_1$. Differences of 7.10 and 10.29 kcal/mol occur for $^3H_4$ and $^1C_4$, respectively, implying that they differ vastly in their internal energies. The ^3E conformer has the lowest values of both E_{Total} and E_{Inter} (Table 2). Even though the E_4 conformer has E_{Total} and E_{Inter} values almost as low as those of ^3E, it is not likely to be the TS, since its RMSD from the crystal-structure ligands is high.

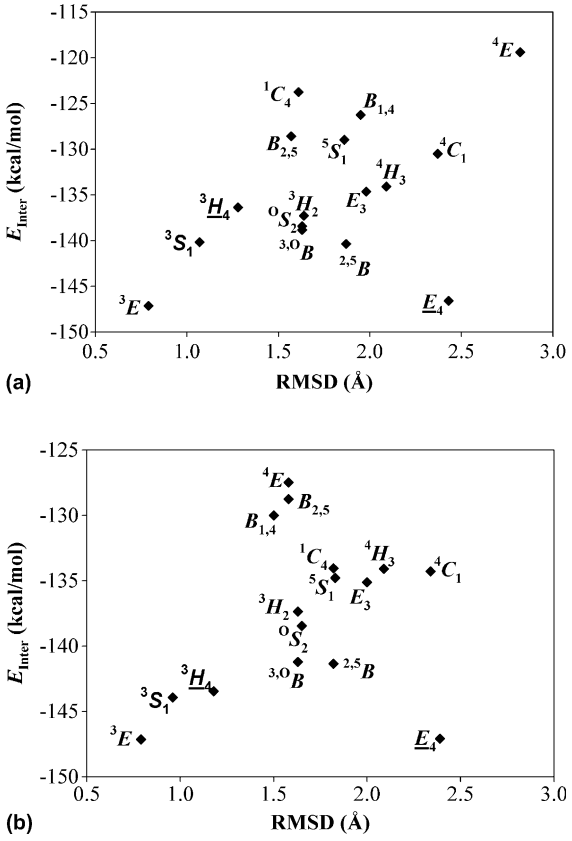

Figure 4. E_{Inter} versus RMSD for conformers with (a) lowest overall E_{Total} and (b) lowest E_{Inter}. Conformers on the suggested pathway to the TS (in turn): $^1C_4$, $^3H_2$, $^OS_2$, $^{3,O}B$, $^3S_1$, ^3E. Underlined: the two previously predicted TS conformers.

3.2. Force computations

Computed forces on ligand atoms can be useful in gaining information on the ligand dynamics in the active site,[23] since forces are vectors and capture information that scalar energies do not. The effect of the forces on docked ligands was studied in two ways—scalar sums of all the forces on ligand atoms were computed to obtain the total distorting force of the enzyme on the docked ligands (Table 2), and individual forces on ligand hydroxyl groups were qualitatively analyzed to find the direction of the conformational twist caused by these forces.

The total distortion forces on the conformers thought to be part of the TS pathway are listed in Table 2 for those with the lowest docked E_{Inter} and E_{Total} values. For all but the $^3H_4$ and E_4 conformers, the force is the same or significantly higher for the conformer with the lowest docked E_{Inter} value, which implies that greater interaction with the enzyme (lower E_{Inter} values) leads to greater stress (higher E_{Intra} values) on those conformers. This is similar to the process by which the high-energy substrate TS is stabilized in the enzyme active site through increased enzyme–substrate interactions, evidently achieved by greater steric complementarity with the TS compared to the relaxed conformation of the substrate. For conformers with the lowest docked E_{Total} values, the total distortion force increases with decreasing E_{Inter} values, a trend consistent with expecting increased stress on the ligand as it approaches the TS. Since no such trend is observed in the total distortion force for the ligands with lowest docked E_{Inter} values, these probably do not represent optimal docked conformations. The energy function of AutoDock, therefore, seems to be capable of capturing changes to the E_{Intra}

Table 2. Forces (pN), E_{Inter} values (kcal/mol), and RSMDs (Å) of conformations in the predicted transition-state pathway

Conformer	Derived from conformers with lowest E_{Inter} values			Derived from conformers with lowest E_{Total} values		
	Force	E_{Inter}	RMSD	Force	E_{Inter}[a]	RMSD[a]
$^{1}C_4$	1176	−134.07	1.82	486	−123.78	1.61
$^{3}H_2$	709	−137.55	1.63	708	−137.30	1.64
$^{O}S_2$	770	−138.46	1.65	754	−138.44	1.63
$^{3,O}B$	1101	−141.22	1.63	831	−138.85	1.63
$^{3}S_1$	1183	−143.94	0.96	825	−140.20	1.07
^{3}E	1043	−147.16	0.79	1043	−147.16	0.79
$^{3}H_4$[b]	637	−141.22	1.18	749	−136.37	1.28
E_4[b]	1123	−147.09	2.39	1291	−146.60	2.43

[a] Obtained from Table 1.
[b] Not part of predicted pathway.

value of the ligand when it is allowed torsional flexibility. In the following analysis of the conformational twist induced by forces on the ligand atoms, we therefore have chosen ligands with the lowest docked E_{Total} values.

Hydrogen-bonding forces are an order of magnitude stronger than van der Waals forces[23] and are therefore key to generating the conformational changes that lead to the TS. We have demonstrated how computation of forces along with energies can give useful insights into enzyme function.[23,24] We showed, for example, that GH7 cellulases owe their processivity to a force on the substrate by the enzyme that pulls it along the active site. We also showed how these forces are involved in stabilizing the TS, or in keeping the highly specific enzyme–substrate interaction from clogging the active site.

In this work, since our goal was to predict the conformational itinerary to the TS, we were interested to see if the forces on individual atoms of docked Man_2 conformers would carry information about how they were twisted. The problem here, therefore, is to gauge the trajectory of a conformer that is acted upon by several forces in different directions, each of these forces acting upon different torsional axes of rotation. To simplify this problem, we have used a simple and intuitive scheme to help us obtain a qualitative sense for the direction of conformational twist. We have listed the orientations of all the ring substituents (defined in Fig. 5) of all 38 Man_2 conformers in Figure 6. Only sufficiently large forces (>10 pN) were considered, which automatically ruled out van der Waals forces. Components of the hydrogen-bonding forces orthogonal to the ring torsional axis were visually observed to predict the effect of that force on the substituent orientation, which is shown by arrows next to individual substituents of docked ligands (Fig. 6). As an example, forces on optimally docked $^{3}H_2$ ligand are shown in Figure 7. The information in Figure 6 was then used to assign weights to any given conformational transformation. For example, the $^{3}H_2$ to $^{O}S_2$ shift is valued at four points because forces on O-1, O-2, O-3, and O-6 atoms twist the $^{3}H_2$ conformer in the direction of the $^{O}S_2$ conformer. It should be borne in mind that the effect of these external forces would cause internal rearrangements within the mannosyl ring, most likely causing movement toward the nearest local minimum on the conformational

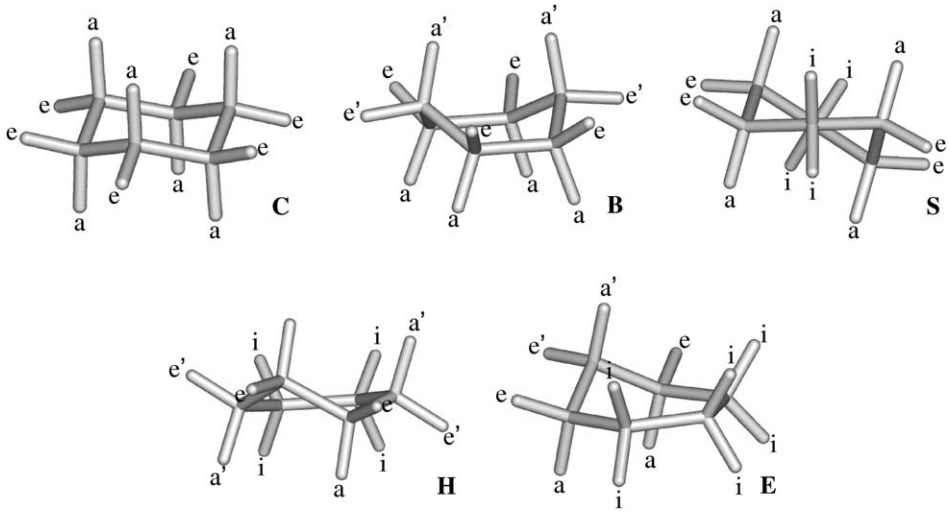

Figure 5. Substituent positions on boat (B), chair (C), envelope (E), half-chair (H), and skew-boat (S) conformations: a = axial, a′ = quasi-axial, e = equatorial, e′ = quasi-equatorial, and i = isoclinal.

	$^{1}C_{4}$
O-1	e→e
O-2	e→n
O-3	a→a
O-4	a→a'
O-6	a→n

	^{3}E	$^{3}H_{4}$	E_{4}	$^{5}H_{4}$	^{5}E	$^{5}H_{O}$	E_{O}	$^{1}H_{O}$	^{1}E	$^{1}H_{2}$	E_{2}	$^{3}H_{2}$	^{3}E
O-1	i→a'	i→a'	i→a'	i	i	e'	e	e	e'	e	e	e'→i	i→a'
O-2	e→n	e'→i	i→n	i	i	i	i	e'	e	e	e'	e→n	e→n
O-3	a'→n	a→n	a→i	a'	i	i	i	i	i	a'	a	a→a	a'→n
O-4	a→a	a→a'	a'→a'	a	a	a'	i	i	i	i	i	a'→a	a→a
O-6	i→i	a'→a	a→a	a	a'	a	a	a'	i	i	i	i→e'	i→i

	$^{3,O}B$	$^{3}S_{1}$	$B_{1,4}$	$^{5}S_{1}$	$^{2,5}B$	$^{2}S_{O}$	$B_{3,O}$	$^{1}S_{3}$	$^{1,4}B$	$^{1}S_{5}$	$B_{2,5}$	$^{O}S_{2}$	$^{3,O}B$
O-1	a→a	a→a	a'→a	a→n	a→a	i	e	e	e'	e	e→e'	i→a'	a→a
O-2	e→e'	i→a'	a→n	a→n	a'→a'	a	a	i	e	e	e'→n	e→n	e→e'
O-3	a'→a'	a→n	a→n	i→n	e→i	e	e'	e	e	i	a→a	a→a	a'→a'
O-4	a→a	a→n	a'→n	a→a'	a→a	i	e	e	e'	e	e→e'	i→a'	a→a
O-6	e→e	i→i	a→n	a→a	a'→n	a	a	i	e	e	e→n	e→e	e→e

	^{O}E	$^{O}H_{1}$	E_{1}	$^{2}H_{1}$	^{2}E	$^{2}H_{3}$	E_{3}	$^{4}H_{3}$	^{4}E	$^{4}H_{5}$	E_{5}	$^{O}H_{5}$	^{O}E
O-1	a	a	a'	a	a	a'	i→a'	i→e'	i→e'	i	i	a'	a
O-2	i	a'	a	a	a'	a	a→a	a'→a	i→n	i	i	i	i
O-3	i	i	i	e'	e	e	e'→e	e→e'	e→e'	e'	i	i	i
O-4	i	i	i	i	i	e'	e→e'	e→e	e'→e'	e	e	i	i
O-6	e	a'	i	i	i	i	i→a'	e'→n	e→i	e	e'	e'	e

	$^{4}C_{1}$
O-1	a→n
O-2	a→a
O-3	e→e'
O-4	e→n
O-6	e→e

Figure 6. Hydroxyl group orientations for conformations in Cremer–Pople space. Arrows indicate the direction of twist of the hydrogen-bonding force on the hydroxyl group. The overall force on the aglycon was used to determine the direction of the force on O-1. n = no appreciable force.

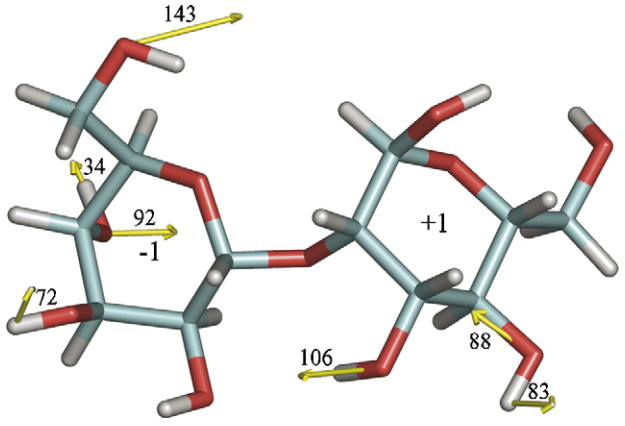

Figure 7. Hydrogen-bonding forces in pN on docked α-Man$_2$-$^{3}H_{2}$. Light blue: carbon, red: oxygen, silver: hydrogen. Mannosyl residues bound in subsites −1 and +1 are labeled. Figure created with PyMOL (DeLano Scientific, South San Francisco, CA).

energy surface. However, it is expected that the additive effect of several external forces on the hydroxyl groups is necessary to force the glycon conformation to move toward its high-energy TS. Therefore, a higher assigned weight for a given transformation points to a higher probability of its occurrence. Points can similarly be assigned for conformer stabilization, an important consideration for the putative TS conformations, since enzymes are believed to operate through TS stabilization. The assigned points for individual transformations derived from Figure 6 are indicated in a pseudorotational map of pyranose ring conformations, along with conformer stabilization points for ^{3}E, $^{3}H_{4}$, and E_{4}, the three putative TSs (Fig. 8).

Several patterns emerge from this information. Clearly, the $^{O}S_{2}$ conformer is pushed toward the $^{3}S_{1}$ structure via the $^{3,O}B$ conformer. The $^{1}C_{4}$ conformer seems to be pushed toward the $^{3}H_{2}$ structure rather than the $^{3}H_{4}$ conformer. This is reasonable since the steric energy of the $^{3}H_{2}$ structure is ~2 kcal/mol lower than that of the $^{3}H_{4}$ conformer,[25] and the former transition therefore offers a lower energy barrier for its eventual conformational twist to the $^{3}S_{1}$ structure. Transformation from the $^{3}S_{1}$ structure to the TS, however, is not as clear. It seems to be pushed more toward the $B_{1,4}$ conformer than to the $^{3}H_{4}$, E_{4}, or ^{3}E putative TSs. This

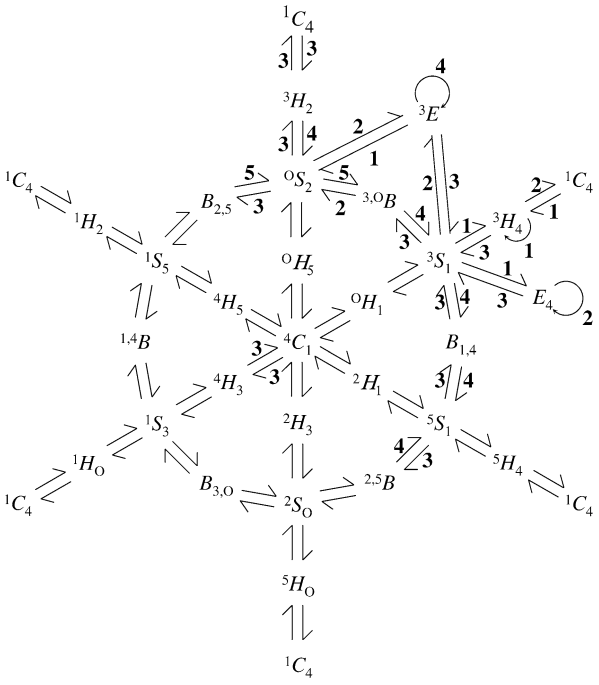

Figure 8. Pseudorotational map of pyranose ring conformations. The numbers adjacent to the arrows are the assigned weights from calculated forces for conformational transformations. Numbers adjacent to semicircles denote weights of stabilizing forces.

may be because conformational sampling is discontinuous. It appears that the skew-boat conformation satisfying the stereoelectronic antiperiplanarity requirement[11] for eventual transformation to the TS lies between the $^{3}S_1$ and $B_{1,4}$ conformers. However, the $^{3}S_1 \rightarrow {}^{3}E$ transition seems to be slightly preferred over $^{3}S_1 \rightarrow {}^{3}H_4$ or $^{3}S_1 \rightarrow E_4$ transitions. Also favoring this transition is the fact that the ^{3}E conformation is more stabilized by the active site, since its stabilization has four points compared to two for E_4 and one for $^{3}H_4$. The enzyme, therefore, seems to operate through the $^{1}C_4 \rightarrow {}^{3}H_2 \rightarrow {}^{O}S_2 \rightarrow {}^{3,O}B \rightarrow {}^{3}S_1 \rightarrow {}^{3}E$ pathway, while the $^{1}C_4 \rightarrow {}^{3}H_4 \rightarrow {}^{3}S_1 \rightarrow {}^{3}E$ pathway seems less likely. The direct transformation of $^{O}S_2$ to a ^{3}E-like TS is also possible and is very attractive due to its agreement with the principle of least motion. However, both the presence of the crystal-structure $^{3}S_1$ conformer of S-Man$_2$[6] and the forces on $^{O}S_2$ (Fig. 8) appear to favor a $^{3}S_1$-based TS pathway. A consideration of the interaction of the three skew-boat conformers, $^{O}S_2$, $^{3}S_1$, and $^{5}S_1$, with the incoming nucleophile helps to explain why $^{3}S_1$ is preferred over $^{O}S_2$ or $^{5}S_1$. The $^{5}S_1$ conformation results in a 1,2 syn–diaxial interaction between the mannosyl 2-OH and the incoming nucleophile. This interaction is partly relieved in $^{3}S_1$ and disappears in $^{O}S_2$, since the 2-OH of the latter is equatorial. However, $^{O}S_2$ has a 1,3 syn–diaxial interaction between the 3-OH and the incoming nucleophile that is partly relieved in $^{3}S_1$. Therefore, of the three skew-boat conformations, $^{3}S_1$ seems to have the least steric hindrance with the incoming nucleophile and may therefore be the preferred conformation en route to the TS.

These results are also consistent with our earlier hypothesis that the enzyme selectively binds Man$_2$-$^{1}C_4$ over Man$_2$-$^{4}C_1$, even though the latter docks with a lower E_{Inter} than the former. Also, the forces on the latter (Fig. 6) indicate that it cannot distort to the TS. The Ca^{+2} at the base of the active site is coordinated to the O-2 and O-3 hydroxyl groups of all the conformers predicted to be in the TS pathway and is therefore prominently involved in the distortion of the substrate to its TS.

It should be noted that the pathway from the TS back to the ground-state conformation of product β-mannopyranose-^{3}E complexed to ERManI is not necessarily the same as the pathway of complexed Man$_2$-$^{1}C_4$ to the TS. Although it is probable that β-mannopyranose-$^{1}C_4$ will be attained, the presence of a β-monosaccharide rather than an α-linked disaccharide leads to significantly different steric energy contours[25] and molecular dimensions.

The force analysis also raises several pertinent questions about the enzyme mechanism. The force calculations suggest that the enzyme active site may help to distort the substrate to its TS. Most enzymes have concave active sites with their amino acid residue side chains closely packed together, leading to fairly rigid conformations. Do enzymes, therefore, achieve their remarkable rate enhancements through carefully positioned hydrogen-bonding residues that direct conformational change to the TS? Is the 1043 pN (Table 2) strain generated on the TS sufficient to counter its internal strain? Vector sums of forces on all the atoms of the docked inhibitors DMJ and KIF are 286 and 298 pN, respectively (Fig. 9), the former with a significant outward component and the latter with an inward component. This agrees with KIF's being a much stronger inhibitor of ERManI, with an IC_{50} of 0.2 μM compared to 20 μM for DMJ.[5] It is likely that inhibitor strength is inversely dependent on its turnover rate, which in turn may depend upon how strongly the enzyme pulls the molecule into, or pushes it away from, its active site. Thus, visualizing the forces on docked ligands may yield information about the binding kinetics of a ligand that cannot be inferred from binding enthalpies alone.

3.3. Catalytic water molecules

Incorporating structural water molecules complicates docking in several ways, since it is not possible to dock two or more ligands in the active site simultaneously with AutoDock. Minimizing a water molecule before introducing an organic ligand can result in wrong binding energies, since the hydrogen-bonding pattern of the included water molecule can change upon introducing the ligand into the active site. Therefore, it is a standard

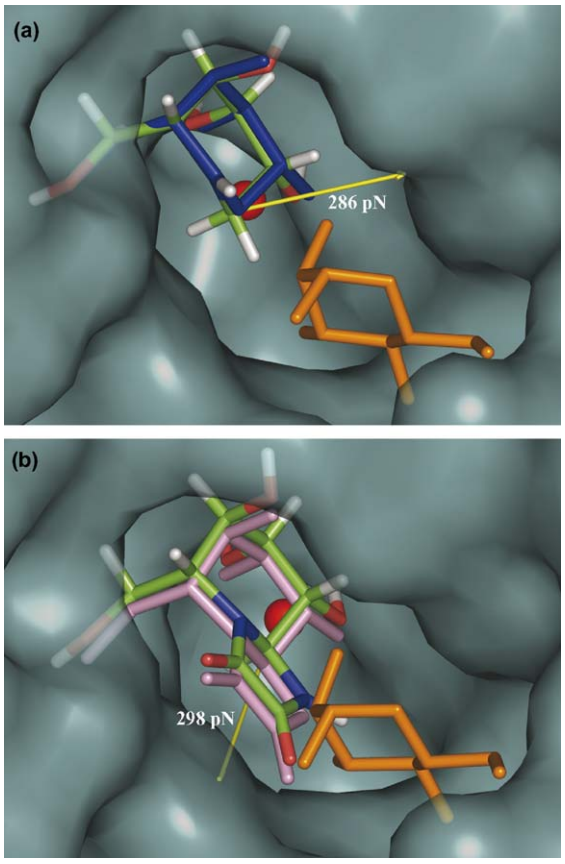

Figure 9. Overall forces on docked DMJ (a) and KIF (b) in the active site, with green: carbon, red: oxygen, silver: hydrogen, dark blue: nitrogen. Crystal structures: dark blue: DMJ, pink: KIF, orange: Man605, red ball: Ca^{2+}. Figure created with PyMOL (DeLano Scientific, South San Francisco, CA).

practice to not include any water molecules during docking, unless their presence is necessary to reproduce a crystal ligand by docking as a control, or unless they are essential for understanding the catalytic mechanism. In the latter case, it would be better to dock the more mobile water molecules after ligand docking to correctly predict hydrogen bonding in the ligand's presence.

Two water molecules have been implicated in the yeast ERManI catalytic mechanism. W195 appears to mediate proton donation by the catalytic acid, while W54 is activated by the catalytic base for nucleophilic attack on the C-1 atom of the mannosyl residue in subsite −1 of Man_2.[7] To gain insights into their functions, we earlier docked these catalytic water molecules in the active site after docking Man_2-E_4 there.[9] Here, since Man_2-^3E is the most likely TS conformer, we docked it and then W54 and W195. The two water molecules docked with energies of −29.70 and − 29.44 kcal/mol, respectively, and with RMSDs of 0.77 and 0.36 Å to crystal structures W54 and W195, respectively. As we suggested earlier,[9] Glu435 seems to be the catalytic base and W195 seems to mediate proton donation.

4. Conclusions

The growing importance of carbohydrate-binding proteins as potential drug targets has led to increased interest in modeling carbohydrate–protein interactions.[27] Knowledge of TS geometry is not only essential for understanding enzyme mechanism but is also useful in inhibitor design. In this study, we report the use of the docking software AutoDock to identify the conformational itinerary of a substrate in an enzyme active site. Although studies such as this one can only capture the shape and electrostatic complementarity of ligand–enzyme geometries, they are nevertheless attractive because of their reduced computational expense compared to more detailed energy calculational methods like molecular dynamics or ab initio techniques. It is important to note here that the local charge distribution of the TS, especially at the ring oxygen atom, differs substantially from that of the ground state,[28,29] and this difference is not taken into account when modeling with AutoDock. However, since the portion of E_{Inter} due to electrostatic interactions is typically <5%, it is unlikely that this charge redistribution will affect the docking results significantly. In other words, in AutoDock shape complementarity involving van der Waals and hydrogen-bonding interactions contributes more significantly to E_{Inter} than does electrostatic complementarity. This method is also limited since ring transformations cannot be continuously modeled; it therefore depends upon careful experimental design. However, this study proves that important clues about enzyme activity lie in the active-site shape alone.

A thorough local energy minimization of the ligand, achieved with the SW algorithm implemented in AutoDock, seems necessary to resolve distinctions between structurally closed docked conformations, although an improved evaluation of E_{Intra} would be necessary to unambiguously compare docked conformations based on E_{total} values. The TS pathway predicted with AutoDock clearly agrees with experimental results by being $^1C_4$-based rather than $^4C_1$-based, and also with stereochemical theory by having $^3S_1$ for a skew-boat intermediate. The final TS predicted, ^3E, is a close conformational neighbor of the $^3H_4$-like TS predicted by Nerinckx et al.[10] and Karaveg et al.,[6] and therefore lends validity to their predictions.

5. Computational methods

5.1. Conformational modeling

Different glycon conformations of Man_2 were generated using PCModel (Serena Software, Bloomington, IN) by constraining the planar ring atoms for a given conformation and then minimizing with its MM3 force field.[30]

For example, $^{3}S_{1}$ was generated by restricting C-2, C-4, C-5, and O-5 to a plane through torsional constraints and then minimizing the conformation with these constraints.

5.2. Energy calculations

Both polar and nonpolar hydrogens were explicitly modeled in both protein and ligands, the former with Lennard–Jones 12–10 hydrogen-bonding parameters and the latter with 12–6 parameters. Hydrogen atoms were added to the yeast ERManI (PDB ID 1DL2) using the What If molecular modeling program.[31] All water molecules were removed while docking. All-atom charges of the Amber force field[32] were used to assign partial charges to the protein atoms. Hydrogen atoms were added to the ligand using Babel,[33] and partial charges were generated with MOPAC[34] using the PM3 Hamiltonian. The grid maps for the van der Waals and electrostatic energies were generated with $101 \times 101 \times 101$ grid-points spaced at 0.375 Å and were centered on the Ca^{2+} ion at the base of the active site. The effect on substrate binding of the Ca^{2+} ion was recognized by using the same parameters for it as those used by Allen et al.[35] Self-consistent Lennard–Jones potentials of AutoDock 1.0 before multiplication by the free-energy model coefficients were used to evaluate nonbonded interaction energies, since they best reproduce the crystal ligands and also because unweighted parameters are necessary to calculate forces on docked ligands. Electrostatic interactions were evaluated using a distance-dependent dielectric constant to model solvation effects. Only the two water molecules mentioned in Section 3.3 were explicitly modeled, and only to confirm their function after Man_2-^{3}E was docked.

For the LGA, the initial population had 50 individuals, the maximal number of energy evaluations was 1.5×10^{5}, the maximal number of generations was 6, the number of top individuals that survived into the next generation was 1, the probability that a gene would undergo random mutation was 0.02, the crossover probability was 0.80, and the average worst energy was calculated over a 10-generation window. The local search component of the LGA was done using the pSW algorithm. A maximum of 300 iterations were allowed per local search, the probability of performing a local search on an individual was 1.0, the maximal number of consecutive successes or failures before doubling or halving the step size of the local search was 4, and the local search was terminated when the step size was <0.01. A total of 100 LGA dockings were performed in a docking run. For the three pSWLS runs on the starting conformation, the local search parameters were the same as those used in the LS component of the LGA. A total of 200 pSWLS dockings were performed in a docking run. The SWLS parameters in the iterative minimization were the same as those used for the pSWLS runs.

5.3. Force calculations

Since AutoDock outputs an energy grid, forces can be computed from them by $-\nabla E = F$, where ∇ is the gradient operator

$$\nabla = \frac{\partial}{\partial x} i + \frac{\partial}{\partial y} j + \frac{\partial}{\partial z} k,$$

and where i, j, and k are unit vectors in the x, y, and z directions, E is the potential energy, and F is the force experienced by the ligand atom. For a sufficiently fine grid, the difference between the potentials of adjacent grid points divided by the distance between them can be used to compute the force on the atom in the x, y, and z directions to obtain a force vector on the atom. A grid distance of 10^{-8} Å led to convergence of force values. Also, the precision of the output energies in the grid files had to be changed from float to double to obtain sufficient resolution for force-value convergence. Forces due to van der Waals as well as electrostatic interactions were evaluated. To minimize this computation, computed subgrids were separately centered around every individual ligand atom.

Acknowledgments

We thank Professor Arthur Olson (Scripps Research Institute) for donating AutoDock and gratefully acknowledge support from the US Department of Agriculture through the Biotechnology Byproducts Consortium and from the US National Science Foundation.

References

1. Herscovics, A. *Biochim. Biophys. Acta* **1999**, *1426*, 275–285.
2. Herscovics, A. *Biochim. Biophys. Acta* **1999**, *1473*, 96–107.
3. Herscovics, A. *Biochimie* **2001**, *83*, 757–762.
4. Coutinho, P. M.; Henrissat, B. 1999, http://afmb.cnrs-mrs.fr/CAZY/.
5. Vallée, F.; Karaveg, K.; Herscovics, A.; Moremen, K. W.; Howell, P. L. *J. Biol. Chem.* **2000**, *275*, 16197–16207.
6. Karaveg, K.; Siriwardena, A.; Tempel, W.; Liu, Z.-J.; Glushka, J.; Wang, B.-C.; Moremen, K. W. *J. Biol. Chem.* **2005**, *280*, 29837–29848.
7. Vallée, F.; Lipari, F.; Yip, P.; Sleno, B.; Herscovics, A.; Howell, P. L. *EMBO J.* **2000**, *19*, 581–588.
8. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
9. Mulakala, C.; Reilly, P. J. *Proteins* **2002**, *49*, 125–134.
10. Nerinckx, W.; Desmet, T.; Claeyssens, M. *FEBS Lett.* **2003**, *538*, 1–7.
11. Deslongchamps, P. *Pure Appl. Chem.* **1993**, *65*, 1161–1178.

12. Solis, F. J.; Wets, R. J.-B. *Math. Oper. Res.* **1981**, *6*, 19–30.
13. Andrews, C. W.; Fraser-Reid, B.; Bowen, J. P. *J. Am. Chem. Soc.* **1991**, *113*, 8293–8298.
14. Smith, B. J. *J. Am. Chem. Soc.* **1997**, *119*, 2699–2706.
15. Winkler, D. A.; Holan, G. *J. Med. Chem.* **1989**, *32*, 2084–2089.
16. Stubbs, J. M.; Marx, D. *J. Am. Chem. Soc.* **2003**, *125*, 10960–10962.
17. Tvaroska, I.; Andre, I.; Carver, J. P. *J. Am. Chem. Soc.* **2000**, *122*, 8762–8776.
18. Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752–755.
19. Coutinho, P. M.; Dowd, M. K.; Reilly, P. J. *Ind. Eng. Chem. Res.* **1998**, *37*, 2148–2157.
20. Laederach, A.; Dowd, M. K.; Coutinho, P. M.; Reilly, P. J. *Proteins* **1999**, *37*, 166–175.
21. Allen, M. J.; Laederach, A.; Reilly, P. J.; Mason, R. J.; Voelker, D. R. *Glycobiology* **2004**, *14*, 693–700.
22. Aikens, C. L.; Laederach, A.; Reilly, P. J. *Proteins* **2004**, *57*, 27–35.
23. Mulakala, C.; Reilly, P. J. *Proteins* **2005**, *60*, 598–605.
24. Mulakala, C.; Reilly, P. J. *Proteins* **2005**, *61*, 590–596.
25. Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1994**, *264*, 1–19.
26. Vyas, N. *Curr. Opin. Struct. Biol.* **1991**, *1*, 732–740.
27. Laederach, A.; Reilly, P. J. *Proteins* **2005**, *60*, 591–597.
28. Kajimoto, T.; Liu, K. K.-C.; Pederson, R. L.; Zhong, Z.; Ichikawa, Y.; Porco, J. A.; Wong, C.-H. *J. Am. Chem. Soc.* **1991**, *113*, 6187–6196.
29. Nerinckx, W.; Desmet, T.; Piens, K.; Claeyssens, M. *FEBS Lett.* **2005**, *579*, 302–312.
30. Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
31. Rodriguez, R.; Chinea, G.; Lopez, N.; Pons, T.; Vriend, G. *CABIOS* **1998**, *14*, 523–528.
32. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
33. Walters, P.; Stahl, M., 1992. http://smog.com/chem/babel/.
34. Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–105.
35. Allen, M. J.; Laederach, A.; Reilly, P. J.; Mason, R. J. *Biochemistry* **2001**, *40*, 7789–7798.