PROTEINS: Structure, Function, and Genetics 46:295–307 (2002)
DOI 10.1002/prot.10017

Docking Multiple Conformations of a Flexible Ligand Into a Protein Binding Site Using NMR Restraints

Adam P.R. Zabell[1] and Carol Beth Post[1,2]
[1]*Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, Indiana*
[2]*Department of Biological Sciences, Purdue University, West Lafayette, Indiana*

ABSTRACT **A method is described for docking a large, flexible ligand using intra-ligand conformational restraints from exchange-transferred NOE (etNOE) data. Numerous conformations of the ligand are generated in isolation, and a subset of representative conformations is selected. A crude model of the protein–ligand complex is used as a template for overlaying the selected ligand structures, and each complex is conformationally relaxed by molecular mechanics to optimize the interaction. Finally, the complexes were assessed for structural quality. Alternative approaches are described for the three steps of the method: generation of the initial docking template; selection of a subset of ligand conformations; and conformational sampling of the complex. The template is generated either by manual docking using interactive graphics or by a computational grid-based search of the binding site. A subset of conformations from the total number of peptides calculated in isolation is selected based on either low energy and satisfaction of the etNOE restraints, or a cluster analysis of the full set. To optimize the interactions in the complex, either a restrained Monte Carlo-energy minimization (MCM) protocol or a restrained simulated annealing (SA) protocol were used. This work produced 53 initial complexes of which 8 were assessed in detail. With the etNOE conformational restraints, all of the approaches provide reasonable models. The grid-based approach to generate an initial docking template allows a large volume to be sampled, and as a result, two distinct binding modes were identified for a fifteen-residue peptide binding to an enzyme active site. Proteins 2002;46:295–307.**

Key words: exchange-transferred NOE; peptide–protein modeling; simulated annealing; ligand docking; Monte Carlo; molecular dynamics; band 3; scoring functions

INTRODUCTION

Exchange-transferred nuclear Overhauser effect spectroscopy (etNOESY) is an accurate NMR method for determining the bound peptide structure in peptide–protein complexes.[1] Distance restraints estimated from etNOESY intensities are used to determine multiple structures of the peptide in isolation, that is, the set of in vacuo peptide structures. Such structural analysis of a peptide–protein complex is particularly important in the characterization of protein–protein associations when the intact protein complex is not amenable to direct structural analysis. The etNOESY experiment provides information about the interproton distances within the peptide, but little or no information on intermolecular contacts. As such, orientation of the peptide in the protein-binding site requires modeling the peptide–protein complex through computational approaches.

Considerable progress has been made in predicting the binding of a small molecule ligand to a protein receptor.[2–4] The development of efficient sampling algorithms and reliable scoring functions has lead to successful docking of small rigid ligands in a number of cases. One approach to molecular docking is through descriptors of the surface shape[5,6] with provision for a limited degree of flexibility.[4,7,8] A number of groups[2,3,9–18] have utilized all-atom models based on a molecular mechanics force field, with sampling of conformations by molecular mechanics, Monte Carlo, genetic algorithms, or distance embedding. More recent docking studies strive to include flexibility of the ligand and/or receptor. The majority of docking algorithms that allow flexibility have been tested on ligands with fewer than twenty rotatable bonds. The results produce a reasonably accurate ligand complex that is often within 1 to 2 Å rms deviation of the crystallographic structure.

Nonetheless, molecular docking in a receptor-binding site remains problematic for flexible peptides with more than a few residues.[19,20] The large number of internal coordinates is challenging for most sampling algorithms and an energy function based on physical terms alone is not a sufficiently discriminating scoring function. Addition of NMR distance restraints to a molecular mechanics force field makes it feasible to dock a flexible peptide with 15 to 20 residues, or approximately 100 dihedral angles. We report the use of distance restraints from etNOESY data combined with the CHARMM force field with different computational protocols to dock a fully flexible fifteen-

Grant sponsor: NIH; Grant numbers: R01-GM39478, K04-GM00661, 5T32-GM08296; Grant sponsor: Lucille P. Markey Foundation; Grant sponsor: Purdue University Academic Reinvestment Program.

*Correspondence to: Carol Beth Post, Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 47907-1333. E-mail: cbp@purdue.edu.

Received 6 March 2001; Accepted 10 August 2001

residue peptide derived from band 3 in the active site of aldolase.

Initial modeling of the band 3 peptide (B3P)-aldolase complex has been reported.[21] We extend the earlier results by exploring a number of docking protocols. The goal was to achieve a reliable model of the peptide–protein complex rather than to develop an efficient procedure that could be used for screening a large number of ligands. The protocols differ in how initial complexes are generated for multiple etNOESY peptide structures, how a subset of peptides is selected for docking from the large set of etNOESY structures, and the algorithm used for conformational sampling of the complex. Initial complexes were generated from either a manually built template model or from a grid-based search of the binding site. Cluster analysis for selecting representative structures from the large set of etNOESY peptide structures is compared to the usual procedure for selection based on low energy and agreement with the NMR restraints. The common sampling algorithm for NMR structure determination is simulated annealing molecular dynamics, and this technique is compared to Monte Carlo minimization (MCM).[22,23] Tests with three-residue peptides and other ligands of equivalent molecular weight demonstrate that MCM is effective for conformational sampling in docking flexible ligands[2,12,24,25] and is a reasonable procedure to test with larger systems. The combination of MCM and etNOESY distance restraints has been used to model a flexible 14-residue peptide in the binding site of thrombin.[17] Initial coordinates for the N-terminal ten peptide residues were taken from the crystallographic structure of thrombin complexed to a shorter version of the peptide, and these investigators developed a procedure for growing the rest of the peptide. The approach reported here is of more general application than that described by Maurer and co-workers: here, the full-length peptide structure is docked from the start rather than growing the peptide starting from a crystallographic structure of the complex, and therefore our approach does not depend on prior knowledge of a peptide complex structure. In addition, the quality of the peptide–protein complex is assessed from a number of structural features following the conformational search; this post-assessment is also described herein.

Band 3 has been shown to bind and inhibit three enzymes involved in glycolysis[26,27]: aldolase,[28,29] glyceraldehyde-3-phosphate dehydrogenase (G3PDH), and phosphofructokinase (PFK).[30,31] Tyrosine phosphorylation of band 3 at Tyr 8 prevents binding and thus mediates against inhibition of glycolysis. A synthetic peptide derived from the N-terminal fifteen-residues of band 3—MEELQDDYEDMMEEN—has been used to demonstrate inhibition of the three enzymes in vitro[28,31–33] and in vivo.[34,35] The structure of this band 3 peptide (B3P) when bound to aldolase[21] and G3PDH[36] determined by the etNOE method[1] forms a loop centered on Tyr 8 [see Fig. 1(A)]. An initial modeling study[21] suggests the loop binds deeply into the active site of aldolase, with the aromatic ring of tyrosine 8 pointing away from the active site and surrounded by the side chains of Leu 4 and Met 12 to form a Phosphorylation Sensitive Interaction (PSI) loop. A dearth of long-range etNOE interactions at both the N- and C-terminus of B3P [Fig. 1(B)] resulted in significant divergence in the peptide ends among the ligand solutions.

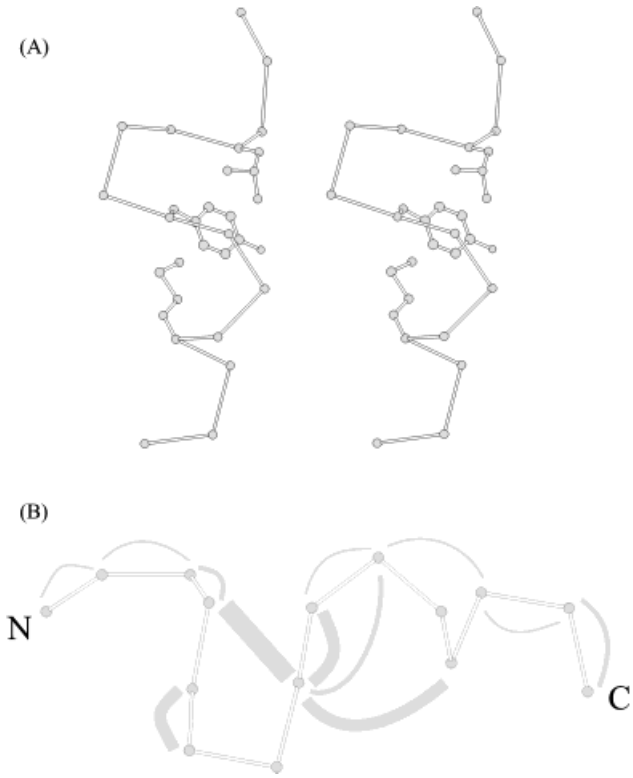


Fig. 1. **A:** Cα trace in stereo of a single B3P structure with the side chains of Leu 4, Tyr 8, and Met 12 drawn to highlight the PSI loop interactions. **B:** Same structure as in A with gray curves connecting those residues with interresidue etNOEs. Curve width is proportional to the number of etNOEs between the residue pair.

Docking of B3P to aldolase by the protocols described in this paper predicts two modes of B3P binding related by roughly a 180° rotation. Thus, the more exhaustive search procedure provides better information than the initial modeling.[21] While Monte Carlo minimization searched a larger region of space, simulated annealing finds an overall lower energy complex structure.

METHODS

The procedure for docking multiple conformations involved three steps: formation of an initial peptide:protein complex to overlay multiple sets of in vacuo-generated B3P coordinates, selection of these coordinates from the full set of 200 in vacuo structures resulting from the etNOESY data, and a restrained conformational search.

In Vacuo Peptide Structures

The structure of the 15-residue B3P (MEELQDDYEDMMEEN-NH_2) bound to aldolase was determined using etNOE data and restrained molecular dynamics, as described in detail elsewhere.[21] A total of 67 distance restraints classified as strong (1.8–2.7 Å), medium (1.8–3.3 Å), or weak (1.8–5.0 Å) were used to generate a set of 200

B3P structures in vacuo with a standard simulated annealing protocol. Multiple conformations were then selected from this set of in vacuo NMR structures to initiate docking.

Template Complex

An initial orientation of B3P bound to aldolase focused only on the interactions of residues 4–9, the PSI loop, since this loop is best defined by the etNOE restraints. Stoichiometry measurements, inhibition assays, and displacement assays indicate that B3P binds the active site of aldolase,[21] therefore docking targeted the active site vicinity. A model of aldolase and the PSI loop was used as a template for generating initial coordinates for multiple B3P:aldolase structures. Each peptide selected from the set of 200 in vacuo NMR structures was positioned into the active site of aldolase using a least-squares superposition of the PSI loop mainchain atoms. The template for residues 4–9 in the active site of aldolase was generated either manually using interactive graphics (IG) or computationally using a grid search (GS).

Interactive graphics (IG)

The PSI loop from the lowest energy structure of the 200 in vacuo peptides was manually docked into the energy minimized crystal structure of human muscle aldolase[37] (PDB access code: 1ALD) using the graphics software program QUANTA (Molecular Simulations Inc.). The PSI loop was translated and rotated as a rigid body using the QUANTA tools to optimize by eye the electrostatic and steric interactions with aldolase. The resulting placement of B3P is named the IG0 template.

Grid search (GS)

An exhaustive search of the space surrounding the protein active site was achieved using a discrete set of orientations of the PSI loop relative to aldolase. This grid-based search was initiated with the PSI loop further from the surface of aldolase than the IG0 template and in two orientations within the active site. One of the two starting orientations was aligned roughly the same as that in the IG0 template, while the other was rotated by approximately 180° about a primary axis defined by Asp 7 Cα of B3P and Arg 148 Cζ of aldolase. Arg 148, along with Lys 41, Arg 42, Lys 107, and Lys 146, bind the substrate of aldolase.[38] The active site region was systematically sampled through a series of rigid-body translations and rotations of the peptide relative to the enzyme with a calculation of the potential energy using CHARMM and version 23 of the all-hydrogen force field.[39,40] Two axes orthogonal to the primary axis were roughly tangential to the enzyme surface and defined a plane for translations. The peptide was translated in 1 Å increments to 18 positions: nine positions in the tangential plane for each of two translations along the primary axis. The first position in the primary axis was defined by the location of the PSI loop at the start of the search, the second position was 1 Å closer to aldolase Arg 148. At each of these locations, the peptide fragment was rotated ±30° in 10° increments about the tangential axes and 360° in 30° increments through the primary axis, for a total of 588 rotations. The interaction energy between the peptide fragment and aldolase was calculated for each position and the structure with the lowest interaction energy was selected as the new initial configuration for a subsequent round through the grid search. A new primary axis was defined based on this new configuration and the process repeated until the initial configuration returned the lowest interaction energy of the round. Each iteration took 11 hours of CPU time on an IBM RISC6000/370 workstation. This grid search resulted in the template structures GS0 and GS180 with roughly opposite orientations in the aldolase binding site (see Fig. 3).

Selection and Reduction of In Vacuo etNOE Structures

Two selection procedures were employed to reduce the number of free peptide coordinates from the 200 in vacuo NMR structures. The first procedure selected a subset of 20 free peptide structures based on their agreement with the NOE data (no NOE violations greater than 0.2 Å) and their low total energy (an average total energy equal to −152 kcal/mol).[21] These 20 peptide structures are the E_{lo} subset. The second procedure clustered the set of 200 in vacuo peptides by the Cα-Cα distance for residues Glu 2–Gln 5, Glu 2–Tyr 8, and Glu 2–Glu 13 and provided six major clusters.[41] Reversing the order in which the individual structures were input to the algorithm did not significantly change the size or membership of the six clusters. The members within each cluster were averaged, and the average structure was subjected to restrained energy minimization using conjugant gradient minimization and the CHARMM23 force field with distance restraints implemented as a soft-square-well potential with a force constant of 100 kcal mol^{-1} $Å.^{-2}$ The minimization was performed until the potential energy gradient was less than 0.5. These six structures are the *clu* subset.

Restrained Search Methods

Aldolase:B3P complexes were generated from an overlay of the 20 E_{lo} or the 6 *clu* B3P structures onto the template structures based on a least-squares superposition of the main chain atoms for residues 4 to 9. These starting complexes were conformationally relaxed by either a simulated annealing molecular dynamics or a Monte Carlo energy minimization procedure. NMR restraints were maintained in each search method.

Simulated annealing

Simulated annealing (SA) of each complex was done using CHARMM version 23 and the all hydrogen topology and parameter files modified[36] to ensure planarity in the tyrosine ring and the peptide bond at high temperature. The energy function included coulombic electrostatic terms, Lennard-Jones terms, the usual geometric terms, and NMR distance restraints. The distance restraints were implemented as a soft-square-well potential with a force constant of 100 kcal mol^{-1} $Å^{-2}$ for the 20 E_{lo} peptide

structures, and 150 kcal mol^{-1} $Å^{-2}$ for the six *clu* structures. The side chain coordinates for protein residues within 10 Å of the peptide were restrained with a harmonic force constant of 100 kcal mol^{-1} $Å^{-2}$. All mainchain atoms on the protein and all protein atoms on residues greater than 10 Å from the peptide were fixed. The initial structures for the complex were subjected to 450 steps of Powell energy minimization before SA. SA was executed for a decrease in temperature from 500 to 300 K over 5 psec with a 1-fsec time step. The final step of the SA protocol was 150 steps of energy minimization. Each SA run took 50 minutes of CPU time on an IBM RISC6000/370 workstation.

Monte Carlo

The basic Monte Carlo cycle[42] was used to dock a flexible peptide by sampling a specific number of dihedral angles as well as the overall orientation of peptide relative to enzyme. The structure was minimized following each Monte Carlo move and the Metropolis criterion applied to determine whether the structural changes were to be accepted or rejected. Because the PSI loop, residues 4 to 9, was well defined by the etNOE data, only the dihedral angles of the N- and C-terminal ends of the peptide were sampled by Monte Carlo. The number of dihedral angles to change per Monte Carlo cycle was chosen as a random number 2^{-n} where n is the number of angles to be perturbed. Mainchain and side chain dihedrals of the peptide termini were randomly modified from their value at the start of a cycle based on a Gaussian distribution with a standard deviation of 90° for ψ and 60° for ϕ and χ. Thus, three standard deviations covered the width of typically allowed regions for each backbone dihedral angle based on the Ramachandran plot (i.e., $\psi_{3\sigma} = 270°$ and $\phi_{3\sigma} = 180°$) while the side chain dihedral was permitted to move between low-energy rotameric states (i.e., $\chi_{3\sigma} = 180°$). For the rigid body translation and rotation of the peptide, a random number over a uniform distribution was applied, with upper bounds of 1.5 Å and 30°, respectively, each along a randomly chosen vector. Two hundred steps of restrained energy minimization were executed following each Monte Carlo step. The Metropolis criterion was applied to the minimized energy of the peptide to judge the acceptability of the step. The protein structure was fixed and the peptide was allowed full mobility during the energy minimization. The experimental etNOE distances were restrained with a force of 150 kcal mol^{-1} $Å^{-2}$.

The Monte Carlo minimization (MCM) with thermalization[12,23] was applied invoking a history independent change in the cutoff temperature of the Metropolis criterion to overcome local minima. The MCM procedure was implemented as a set of four units repeated three times for a total of 1,800 cycles: (1) 50 cycles with T = 10,000 K allowing Monte Carlo movements in the ϕ, ψ, and χ dihedral angles of B3P residues 1–3, (2) 250 cycles with T = 310 K allowing Monte Carlo movements in the ϕ, ψ, and χ dihedral angles of B3P residues 1–3, (3) 50 cycles with T = 10,000 K allowing Monte Carlo movements in the ϕ, ψ, and χ dihedral angles of B3P residues 10–15, (4) 250 cycles with T = 310 K allowing Monte Carlo movements in the ϕ, ψ, and χ dihedral angles of B3P residues 10–15. The entire procedure of 1,800 cycles took 17 hours on an IBM RISC6000/370 workstation.

Structure Evaluation

To ensure a meaningful comparison of the SA and the MCM search methods, the final structure from each procedure was subjected to restrained energy minimization under the same potential function until the potential energy gradient was less than 0.5. This usually involved 100 steps of Powell energy minimization. This minimization used an NOE restraint force constant of 150 kcal mol^{-1} $Å^{-2}$ and fixed all aldolase atoms.

Evaluation tools

Seven criteria were used to evaluate the structure quality of the modeled complex at the end of the search procedure. The internal energy of the peptide (E_{self}), the interaction energy between the peptide and the enzyme (E_{inter}), and the NOE violation energy (E_{NOE}) were calculated from the CHARMM22 force field. Internal bond energy terms as well as electrostatic and van der Waals interactions within the peptide were incorporated in E_{self}, while E_{inter} only considered the electrostatic and van der Waals interactions between the two molecules. The energy of the peptide (E_{B3P}) is the sum of E_{self} and E_{inter}. Acceptable ϕ/ψ values as defined by Procheck-NMR[43] and the tabulation of unsatisfied buried hydrogen bond donors and acceptors as defined by distance and angle defaults within QUANTA were used to measure the reliability of the peptide structure. Efficient packing of the two molecules was monitored with a measure of the buried surface area of B3P and the cavity volume between peptide and protein using GRASP.[44] The buried surface area was defined as the difference between the total surface area of the isolated peptide and the largest contiguous surface area of the peptide at least 1.4 Å away from the isolated aldolase surface. The intermolecular cavity volumes were defined as those spaces inaccessible to bulk solvent but beyond the contour of the molecular surfaces in the complex. Finally, an R factor calculation was used to measure the degree to which each structure satisfied the experimental restraints.

Several forms of an R factor calculation for structures solved by NMR have been reported.[45–47] In this study, we use a value analogous to the form used in X-ray crystallographic analysis:

$$R = \sum|(V_{obs}) - (V_{calc})| \Big/ \sum(V_{obs}) \qquad (1)$$

where V_{obs} and V_{calc} are the observed and calculated peak volumes, respectively. Crosspeak intensities were calculated for the fast-exchange limit using the program CORONA (Calculated OR Observed NOESY Analysis)[48] for a spectrometer frequency of 500 MHz, a mixing time of 150 msec, a correlation time for free peptide of 1.0 nsec and for the B3P:aldolase complex of 60.0 nsec, total molecular concentrations of 2.8 mM for the peptide and 0.28 mM for

TABLE I. Pairwise rmsd on Selected Sets of Atoms of B3P

	Mainchain[a]		Mainchain 4–9[b]		
Subset	1–15	1–3	4–9	10–15	1–15
200 in vitro	4.0	7.4	1.2	8.6	6.6
53 after docking	4.0	6.7	1.1	8.7	6.5
Final 8 complexes	3.8	6.4	1.0	6.7	5.3
Best 0° and 180°	2.2	2.1	0.7	3.6	2.5

[a]Least-squares superposition over all peptide mainchain (N, Cα, C) atoms.
[b]Least-squares superposition over mainchain atoms of residues 4–9 of the peptide.

aldolase, and a dissociation constant of 0.052 mM. These parameters correspond to the experimental conditions. For each of the B3P:aldolase complexes, the rate-matrix solution of the Bloch equations for magnetic relaxation and chemical exchange was obtained for B3P plus aldolase protons within 8.0 Å of B3P to give 107 B3P protons and between 240 and 416 aldolase protons in the matrix. The free peptide signal in the experimental data was found to contain only one crosspeak, $Y8_{\delta 1,2\text{-}\epsilon 1,2}$ (and its mirror across the diagonal). The volume of this peak in the free peptide spectrum was less than 10% of that in either bound peptide spectrum. Since this factor is within the noise level of the data, the etNOESY intensities were not corrected for the free peptide intensity. In order to calculate an R factor, it is necessary to scale together experimental intensities measured in H_2O or D_2O solvent, as well as experimental with calculated etNOESY peak intensities. The H_2O and D_2O etNOESY spectra were scaled by a least-squares fit of 18 crosspeaks chosen for their resolution and their spread throughout the spectrum. Peaks calculated by CORONA were scaled with a least-squares fit to the observed data using the 67 etNOEs that were applied as experimental restraints.

RESULTS AND DISCUSSION

Initial Peptide Dispersion

The distribution of interresidue etNOE interactions of B3P bound to aldolase is shown schematically in Figure 1(B); a curve drawn between Cα atoms illustrates an etNOE interaction between any proton in each of the two residues and the width of the curve represents the number of interactions. Relative to the number of interactions commonly observed in a protein, there are fewer interresidue NOE interactions per residue for a peptide bound at the protein surface. In the case of the B3P:aldolase complex, there is an absence of restraints between either of the two termini of B3P and the central PSI loop, which leads to many orientations of the termini for the 200 in vacuo NMR structures. The dispersion in the structural solutions for B3P (see Table I) is highest at the termini since all of the interresidue restraints are either sequential or involve Tyr 8. The average pairwise mainchain rmsd for all 200 in vacuo peptide structures for the entire length of the peptide is 4.0 Å, and is 1.2 Å for the PSI loop, residues 4 to 9. A mainchain overlay of the PSI loop gives rmsd values over the N- and C-terminal portions of the peptide of 7.4 Å and 8.6 Å, respectively. Nearly half of the 200 B3P structures satisfy the NMR restraints with less than one etNOE violation for a 0.1 Å threshold.

To model a complex starting with multiple in vacuo peptide structures, a reduction in the number of peptide conformations is desirable for practical reasons. Commonly used criteria to select structures from the large set of in vacuo NMR solutions are low energy, the absence of violated NOE restraints, and a favorable Ramachandran distribution. These criteria applied to the 200 in vacuo B3P structures yielded 20 E_{lo} conformations. Paring down the full set of peptide structures in this manner does not remove the dispersion in the peptide termini; the pairwise mainchain rmsd is 1.2 Å for the PSI loop and 4.0 Å for all residues, nearly the same as rmsd values for the initial set of 200 structures.

An alternative approach for selecting NMR structures is cluster analysis.[41] Clustering is on the full set of 200 NMR in vacuo peptide structures and considers a larger number of the etNOE in vacuo solutions than the 20 E_{lo} conformations. Clustering was based on three Cα-Cα distances: Glu 2–Gln 5, Glu 2–Tyr 8, and Glu 2–Glu 13 [see Fig. 2(A)]. These distances provide a direct measure of extension of the N-terminus (Glu 2–Gln 5), and the full peptide (Glu 2–Glu 13), and infer a measure of twist (Glu 2–Gln 5 plus Glu 2–Tyr 8) between the PSI loop and the N-terminus. Six major clusters, shown in Figure 2(B), were found that had at least ten structures with an average NOE energy less than 1.6 kcal mol^{-1} and fewer than 1.5 restraint violations per structure (see Table II). The structures within each cluster differed by an average pairwise mainchain rmsd between 2.1 and 3.1 Å.

The average structures of the major clusters are the six *clu* conformations used for docking to B3P. This method for reducing the set of etNOE in vacuo structures retains the conformational variation that the experimental data cannot distinguish, but it is condensed within six structures instead of the 20 E_{lo} structures or the total set of 200 structures. The six cluster average structures represent 90 out of the 200 etNOE structures and have reasonable agreement with the NMR data. The mainchain of any cluster-average structure was on average 4.2 Å rmsd from the other cluster-average structures. The comparison of cluster 2 (blue) to the other five clusters finds rmsd values all larger than 4.8 Å. Some individual cluster-averages are similar to each other. The smallest mainchain rms difference equals 1.7 Å for the average structure of cluster 1 [shown in green in Fig. 2(B)] compared to that of cluster 3 (red). Cluster 6 (orange) is also similar to both cluster 1 and cluster 3; the mainchain rmsd equals 2.9 and 2.6 Å, respectively. Clusters 4 (yellow) and 5 (purple) are similar, with an rmsd of 2.6 Å. The three Cα-Cα distance pairs shown in Table II reflect the trends in the rmsd values.

Docking B3P

The docking protocols outlined in Scheme I differ according to three steps involved in producing a complex: how to generate an initial template complex to overlay multiple B3P structures, how to select the subset of structures to be

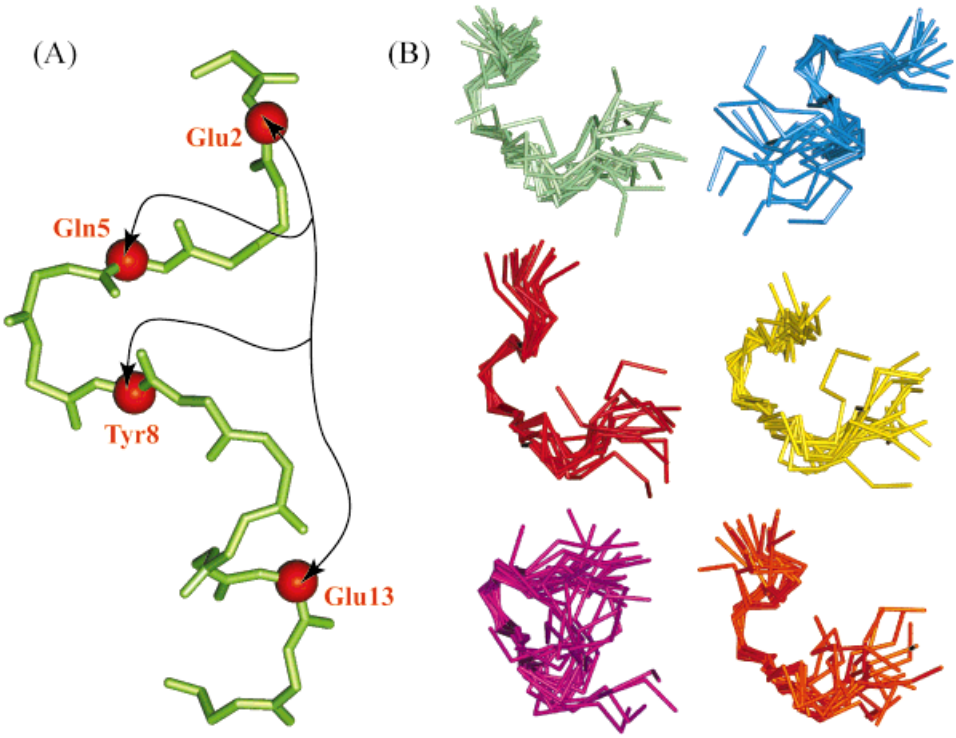


Fig. 2. **A:** Mainchain trace of a representative B3P structure with the four Cα atoms used for clustering shown by red spheres. Arrows indicate distances used for clustering. **B:** Cα trace of the structures from the in vacuo protocol grouped according to membership in each of the six clusters. Each cluster is shown with an rmsd overlay of the PSI loop and oriented so the PSI loop is in the same relative position. Strands are colored by cluster.

TABLE II. Clustering Results After the First SA Protocol

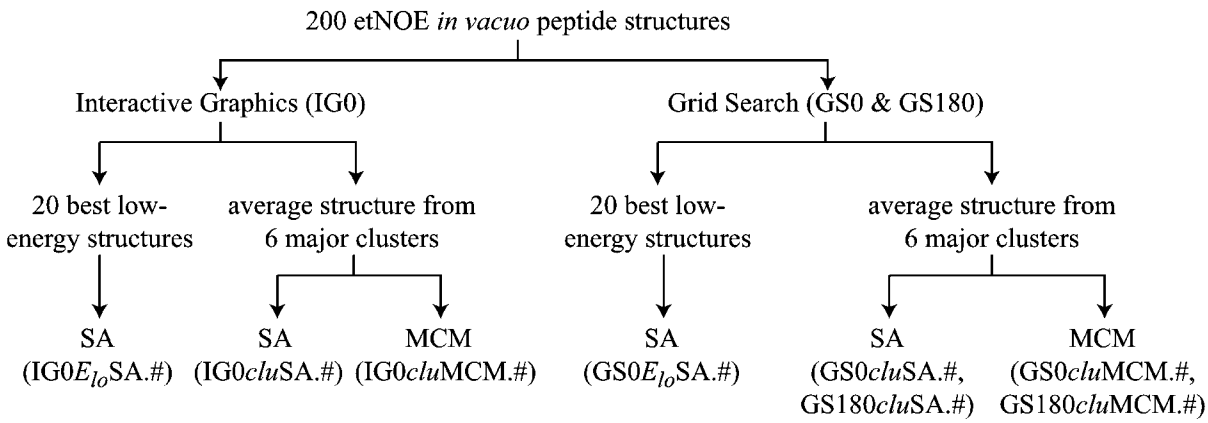
				RMS difference		Cα–Cα distance, Å		
Cluster	Structures per cluster	No. (NOE viol)[a]	$\langle E_{NOE} \rangle$[b]	MC[c]	AH[d]	2–5	2–8	2–13
1	20	0.85	0.83	3.0	4.2	6.21	12.18	18.20
2	16	1.06	1.14	3.1	4.4	7.91	8.59	18.13
3	12	1.08	1.25	2.4	3.7	6.25	11.82	16.41
4	15	1.40	1.41	2.1	3.5	8.84	9.95	7.97
5	12	1.33	1.56	2.6	3.9	8.88	9.12	8.22
6	15	1.33	1.53	3.0	4.4	7.45	11.40	16.38

[a]Average number of violations greater than 0.1 Å.
[b]NOE energy with a force constant of 150 kcal mol^{-1}.
[c]Pairwise rmsd using all mainchain atoms.
[d]Pairwise rmsd using all heavy atoms.

docked from the 200 in vacuo B3P structures, and how to perform the restrained conformational sampling. B3P: aldolase complexes are referred to by initial complex (IG0, GS0, or GS180), peptide reduced set (E_{lo} or *clu*), search method (MCM or SA), and run number (#).

Template complex

A single template complex for overlaying multiple in vacuo peptide structures was generated either by interactive graphics (IG) or a grid search (GS). Docking a rigid PSI loop fragment of B3P into the binding site by IG depends on manual skill and may suffer from human bias. The GS procedure utilizes a crude initial positioning of the PSI loop followed by a computer-based search of the volume that surrounds the binding site and is therefore less dependent on manual skill. The explicit series of translations and rotations that were performed by the GS sampled ~4,800 $Å^3$ around the binding site with each iteration. The step sizes only roughly mapped the contact potential surface, but a higher resolution search would be considerably more costly in computation time and was not considered necessary to provide an adequate initial position for the peptide. The grid search identified an N to C orientation of the PSI loop (GS0) similar to that found by IG docking (IG0). A second orientation found by the grid search (GS180) was rotated by approximately 180° around an axis normal to the enzyme surface (Fig. 3). The GS0 and GS180 template structures of the PSI loop had similar interaction



Scheme 1. Docking protocol.

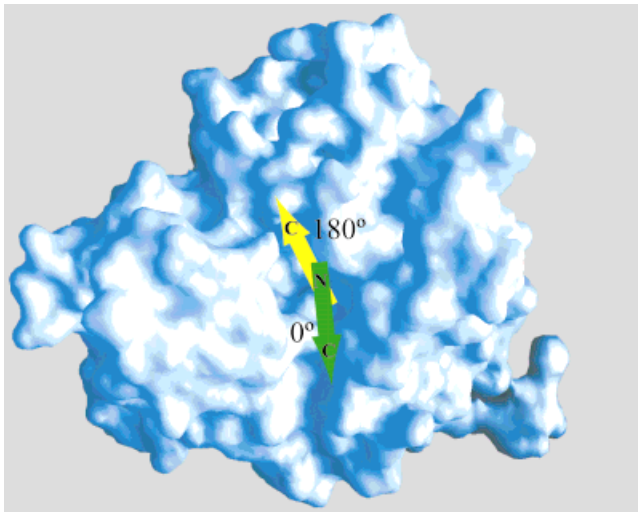


Fig. 3. The 0° (green) and 180° (yellow) orientations of the PSI loop in the binding site of aldolase. Each arrow points in an N- to C-terminus direction for a docked PSI loop. Aldolase is shown as a blue molecular surface drawn with GRASP.[44]

energies with aldolase, and both orientations of the peptide were used for the restrained conformational search.

In vacuo etNOE structure selection

Two subsets of structures, the E_{lo} and the *clu* reduced sets, were chosen as described in Methods. These reduced sets had significant overlap in regard to selected peptide structures. Each of the clusters contained at least one peptide from the E_{lo} set. A Cα rmsd comparison between the twenty E_{lo} structures and the six *clu* average structures shows that 12 of the 20 E_{lo} structures are within 3 Å of at least one *clu* average structure.

Selected in vacuo peptide structures were overlaid with respect to the PSI loop (residues 4 to 9) on the IG0, GS0, and GS180 templates (Scheme I). Twenty-six initial complexes were generated from the IG0 template, 26 complexes from the GS0 template, and six from the GS180 template using the *clu* average structures to make 58 initial complexes. Overlaying the PSI loop of the *clu* structures with the IG0 and GS0 templates led to steric clash of B3P termini with aldolase so that five of these complexes were rejected from further analysis. By comparison, the divergent peptide termini were more readily accommodated when B3P was docked in the 180° orientation and none of the GS180 complexes were rejected. Thus a total of 53 sets of initial coordinates were obtained for restrained conformational searches.

Restrained conformational search

Either restrained Monte Carlo minimization (MCM) or restrained simulated annealing (SA) were used to refine the 53 starting complexes. The evolution of E_{inter} is shown in Figure 4 for a typical SA and MCM run starting from the same initial coordinates. The graphs show relative differences in interaction energy from the lowest energy structure. It should be noted that the absolute energy values are not comparable since different potential functions were used for SA and MCM. The energy shows a rapid decrease to a plateau region for SA [Fig. 4(A)], while the Monte Carlo simulation is more stochastic by nature [Fig. 4(B)]. To characterize the degeneracies in energy sampled with MCM or SA, coordinates from snapshots in the simulation were compared with the minimum-energy structure. In the case of SA, a structure near the end of the simulation was the minimum-energy structure while for MCM, the minimum-energy structure occurred near step 300. Coordinate deviations were calculated after a superposition of either aldolase or B3P in order to monitor the changes of peptide orientation in the binding site as well as the internal changes in peptide configuration. Values for rms differences averaged over mainchain or all heavy atoms of the peptide are presented in Figure 4(C and D) for the simulations described in Figure 4(A and B). The contrast between the gradual, smooth sampling of SA and the stochastic sampling of MCM is apparent in the coordinate rmsd values, as it was in the energy values. Interestingly, the energy time evolution in SA decreases in the first 2 psec corresponding to a displacement of approximately 1.5 Å. The complex fluctuates in a fairly stable energy state from 2 psec to the end of the SA simulation. Comparison of any one structure (in this case, the last structure) in that

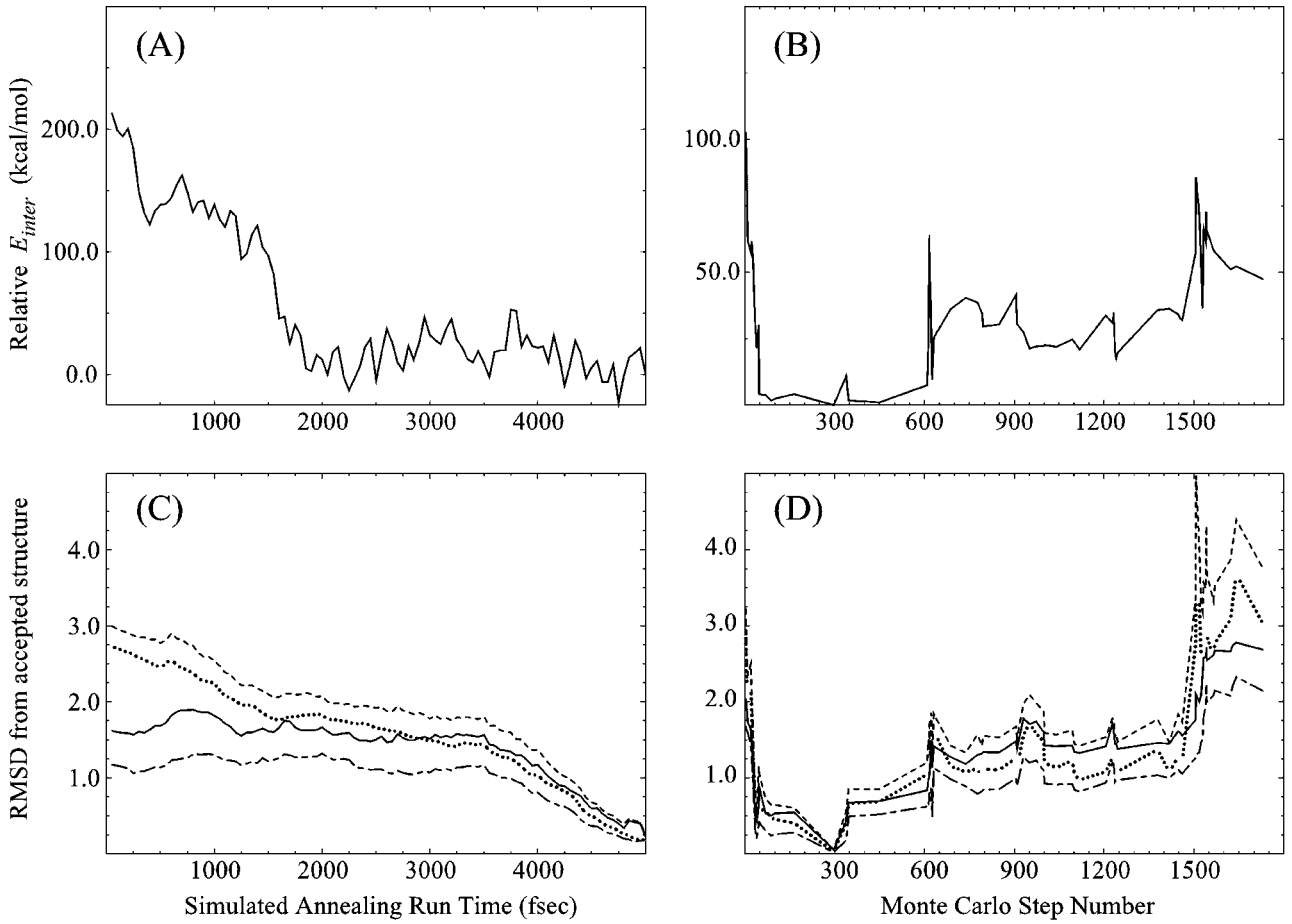


Fig. 4. **A:** Plot of the interaction energy of the peptide, E_{inter}, as a function of the step number from a typical simulated annealing run, GS0*clu*SA.2. **B:** Same information as in A but for the MCM run from the same starting structure, GS0*clu*MC.2. **C:** Plot of the rms difference for B3P using least square superposition of main chain atoms from either aldolase or B3P. The trajectory structures were compared to the minimum-energy structure near the end of the simulation. For superposition with regard to aldolase, the rms value is averaged over the mainchain atoms (dotted line) or all heavy atoms (dashed line) of B3P. For superposition with regard to B3P, the rms is averaged over mainchain atoms (dot-dashed line) or all heavy atoms (solid line) of B3P. **D:** Same information as in C but for the MCM run described in B. The MCM structures were compared to the minimum-energy structure near step 300.

well with the others suggests a sampling time for the well of about 1.4 psec. Rms comparisons referenced to other structures within the local energy well produce similar results (not shown) in terms of magnitude and time constant.

A measurement of the conformational space sampled by each technique was estimated using GRASP[44] to calculate the total volume of a peptide ensemble comprising B3P structures from SA and MCM runs that started from the same IG0, GS0, or GS180 template and the same B3P conformation. An example volume from one SA run is shown in Figure 5(A), and is approximately 0.7 times the size of the analogous MCM volume [Fig. 5(B)]: 3,650 Å^3 and 5,300 Å^3, respectively. The integrated volumes shown in Figure 5 reflect fluctuations of one orientation; there are no transitions during a single run between the two orientations of B3P shown in Figure 3 because of the large conformational change associated with such a transition. In all cases starting from the same initial coordinates, MCM explores more volume than SA.

The changes in B3P conformation and position within the binding site were assessed between an initial and final complex of the conformational search. Rmsd values averaged over the structures from different methods shown in Table III were calculated by either a superposition of peptide residues (mcB3P) to measure the change in peptide internal structure, or by a superposition of aldolase residues (mcALD) to measure the change in peptide orientation. A 1.4 to 2.5 Å change in the internal structure of B3P was observed, while including the position of B3P within the binding site results in larger values of 3.1 to 6.6 Å. The substantially smaller rmsd in the internal structure of B3P is due in part to the etNOE distance restraints.

Analysis of Docked Complexes

The analysis of the B3P:aldolase complexes is based on energy, structural quality, and agreement with experiment (see Table IV). Comparisons are made for groups according to initial complex (GS vs. IG), structure selection (*clu* vs. E_{lo}), and conformational search method (SA vs. MCM). The average value and standard deviation for six independent criteria are listed for each group. Structures with disallowed physical arrangements following the restrained conformational search (e.g., overlapped peptide:

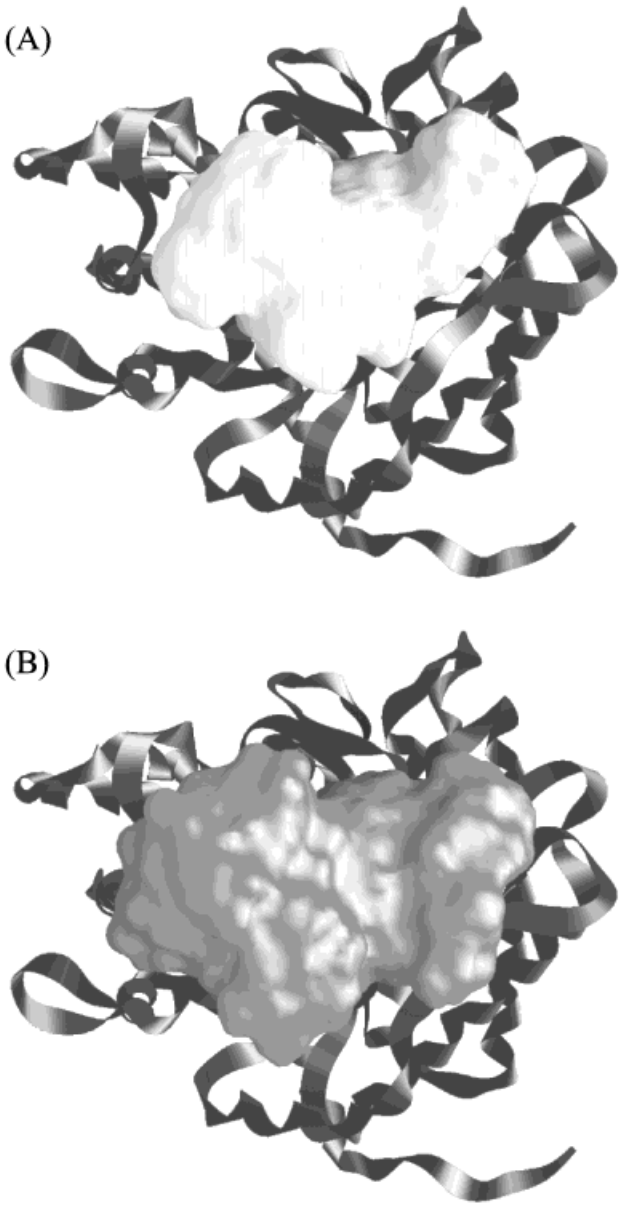


Fig. 5. Volume searched at the aldolase active site by the two runs described for Figure 4. The solid surface is the integrated volume searched by B3P during (**A**) simulated annealing rMD, GS0*clu*SA.2 and (**B**) Monte Carlo with minimization, GS0*clu*MC.2. The aldolase monomer is shown as a ribbon. The figure was made with GRASP.[44]

protein structures, inverted atom chirality) were excluded to give a total of 57 complexes obtained from all protocols.

Protocol comparisons

The energy and structural quality of B3P were evaluated for the complexes modeled by each protocol. Based on the averaged results shown in Table IV, all the protocols perform well with no one protocol being more effective overall. The two methods (IG or GS) to generate an initial template complex for overlaying the PSI loop of multiple in vacuo B3P structures each produced well docked structures. GS0 and GS180 templates, with reversed chain

TABLE III. Average rmsd Between Initial and Final Structures of B3P

Method	rmsd (mcALD)[a]	rmsd (mcB3P)[b]
IG0E_{lo}SA	4.8 ± 1.0	2.0 ± 0.7
IG0*clu*SA	3.1 ± 0.8	1.7 ± 0.8
IG0*clu*MCM	3.2 ± 1.7	2.0 ± 1.3
GS0E_{lo}SA	3.9 ± 1.1	2.1 ± 0.9
GS0*clu*SA	3.6 ± 0.8	1.4 ± 0.2
GS0*clu*MCM	4.6 ± 1.6	2.0 ± 0.6
GS180*clu*SA	5.1 ± 1.2	1.8 ± 0.3
GS180*clu*MCM	6.6 ± 2.1	2.5 ± 0.7

[a]Mainchain root mean square difference, aligning aldolase residues.
[b]Mainchain root mean square difference, aligning B3P residues.

directionality in the aldolase binding site, also lead to similar results on average as found by comparison of averaged features for GS0*clu*SA and GS0*clu*MCM with GS180*clu*SA and GS180*clu*MCM in Table IV. The E_{lo} selection might be expected to have lower E_{self} compared with the *clu* selection, but no such distinction was found; comparison of E_{self} in Table IV between E_{lo} or *clu* protocols that are otherwise equal (i.e., IG0*clu*SA vs. IG0E_{lo}SA and GS0*clu*SA vs. GS0E_{lo}SA) show that the average energy values are within the standard deviation. Comparison of SA with MCM finds similar E_{NOE} and structural quality terms. One distinction was found when examining these pairs of structures: the physical energy terms E_{B3P}, E_{inter}, and E_{self} are consistently lower for SA than MCM protocols, even after energy minimization against the same potential function. For example, values of E_{B3P} SA protocols range from approximately −650 to −790 kcal mol^{-1} while those from MCM are approximately −550 kcal mol^{-1}.

Individual structure comparison

On average, no one protocol showed an advantage for modeling the B3P:aldolase complex, so the complexes were assessed individually to identify the best models. Structures were filtered by acceptable angles in the Ramachandran plot, correct chirality, and a relatively low R factor (Equation 1). Those structures with a phi/psi angle pair in the disallowed region of the Ramachandran plot, wrong chirality at the end of the docking protocol, or with an R factor greater than 1.0 were excluded. The R factor was used rather than the number of restraint violations because the number of violations was not sufficiently discriminating.[1] Filtering based on these criteria reduced the number of structures to eight. The structures are listed in Table V in order of E_{inter} values, along with the other structural analysis terms. Each protocol variation for modeling the B3P:aldolase complex—IG vs. GS, E_{lo} vs. *clu*, and SA vs. MCM—is represented in this set of eight structures. Relatively few of the structures resulted from MCM or the 180° orientation.

The R factor averaged over the best eight structures is 0.92, and the minimum value is 0.83. These values are larger than those typically reported for protein structures.[49–51] Nonetheless, the R factor is found here to be a

TABLE IV. Average Structure Analysis Terms for Each of Eight Methods for Generating the B3P:Aldolase Complex

Method	E_{self}[a]	E_{inter}[b]	E_{B3P}[c]	E_{NOE}[d]	Unfavorable ϕ/ψ[e]	Non-Hbond[f]	Buried area, Å^2	Cavity volume, Å^3
IG0E_{lo}SA								
18 structures	−93.0 ± 45.8	−632.0 ± 110.3	−725.0 ± 84.6	12.7 ± 3.4	0.28 ± 0.6	4.28 ± 1.8	362.0 ± 82.1	170.4 ± 119.8
IG0*clu*SA								
5 structures	−102.8 ± 60.5	−687.2 ± 81.0	−790.0 ± 40.6	10.8 ± 2.3	0.60 ± 0.9	4.00 ± 1.6	353.4 ± 71.1	160.8 ± 17.1
IG0*clu*MCM								
4 structures	−77.4 ± 40.7	−483.4 ± 41.4	−560.7 ± 46.0	11.2 ± 2.6	0.50 ± 0.6	4.50 ± 3.9	306.5 ± 23.9	5.3 ± 10.6
GS0E_{lo}SA								
13 structures	−119.5 ± 38.7	−612.8 ± 70.8	−732.2 ± 58.7	13.4 ± 2.4	0.46 ± 0.5	2.77 ± 1.9	342.9 ± 49.4	52.0 ± 84.3
GS0*clu*SA								
3 structures	−69.9 ± 60.3	−675.5 ± 30.8	−745.4 ± 90.1	11.6 ± 2.5	0.33 ± 0.6	3.67 ± 3.1	351.5 ± 60.6	114.3 ± 112.1
GS0*clu*MCM								
3 structures	−24.1 ± 56.6	−517.8 ± 107.2	−541.9 ± 59.3	9.7 ± 2.9	0.33 ± 0.6	3.00 ± 1.7	359.5 ± 36.6	118.4 ± 177.9
GS180*clu*SA								
6 structures	−83.7 ± 23.0	−566.0 ± 135.1	−649.8 ± 112.9	10.6 ± 1.4	0.83 ± 0.8	3.83 ± 1.2	350.0 ± 94.7	92.6 ± 123.5
GS180*clu*MCM								
5 structures	−71.3 ± 19.1	−497.2 ± 35.1	−568.5 ± 30.3	12.1 ± 2.6	0.60 ± 0.6	3.00 ± 1.9	283.3 ± 55.3	17.6 ± 39.3

[a]Intrapeptide energy, kcal mol^{-1}.
[b]Intermolecular interaction energy, kcal mol^{-1}.
[c]Summation of intrapeptide and intermolecular energies, kcal mol^{-1}.
[d]Energy from NOE constraint violations, kcal mol^{-1}.
[e]Number of residues outside of generously allowed regions of the Ramachandran plot.
[f]Number of hydrogen bond groups not interacting in a hydrogen bond.

TABLE V. Structure Analysis Terms for the Eight Best B3P:Aldoalse Complexes[†]

Structure[a]	E_{self}	E_{inter}	E_{B3P}	E_{NOE}	R factor	Unfav. ϕ/ψ	Non-Hbond	Buried area, Å^2	Cavity volume, Å^3	rmsd (mcALD)	rmsd (mcB3P)
IG0E_{lo}SA.6	−53.1	−742.3	−795.5	15.4	0.94	0	4	425.4	205.7	5.9	2.7
IG0*clu*SA.3	−86.7	−731.1	−817.8	13.6	0.89	0	4	397.1	150.9	2.2	1.6
GS180*clu*SA.1	−73.0	−650.9	−724.0	11.4	0.93	0	3	390.8	312.0	5.7	1.7
GS0E_{lo}SA.6	−123.7	−640.6	−764.3	10.1	0.92	0	2	319.3	0.0	6.0	4.2
IG0E_{lo}SA.5	−69.1	−593.3	−662.4	14.5	0.97	0	3	321.5	217.6	5.0	2.5
GS0E_{lo}SA.9	−106.4	−570.6	−677.0	13.9	0.99	0	4	281.2	0.0	3.8	1.9
GS0E_{lo}SA.3	−175.4	−519.5	−695.0	16.5	0.88	0	4	327.8	0.0	2.5	1.2
IG0*clu*MCM.1	−68.5	−468.8	−537.2	11.9	0.83	0	4	304.4	0.0	4.4	4.0

[a]Terms defined in Tables III and IV.
[b]See Scheme 1 for the protocol used to calculate the individual complex listed.

useful indicator. Filtering based on other factors gave the same general results in terms of which protocols lead to the best docked complexes. The large R values for the B3P structures may arise because the relatively sparse number of restraints cannot overcome the deficiency inherent in applying loose strong/medium/weak restraints. The restraints have a common lower bound of 1.8 Å, allowing considerably closer distances than might be justified from the crosspeak volume. In a direct protein structure determination, the relatively larger number of restraints per residue can find the correct solution and thus lower R values in spite of loose restraints. In our experience with other exchange systems that have more etNOE interactions per residue, a lower value for the R factor is achieved. Additional work on this topic is currently under investigation. The cavity volume was calculated for each complex in order to measure the degree to which the molecules fit together in a lock-and-key arrangement. It was expected that those complexes with optimal intermolecular interactions would have negligible cavity volumes. However, four structures including the three with the lowest E_{inter} contained a single cavity between 150 and 310 Å^3. Visual examination of this cavity demonstrated that a pocket was created by the formation of hydrogen bonds from B3P residues Asp 6 and Asp 7 to several aldolase residues, primarily Arg 42 and Arg 148. This pocket has been shown to be the binding site for fructose 1,6-bisphosphate, the enzyme substrate.[52] If binding of B3P to aldolase does cover the heart of the binding site as predicted, it is likely that this cavity produced in computational docking is actually filled with water molecules.

All eight final structures had extended N- and C-terminal ends, allowing extensive interactions along the protein surface, while retaining the PSI loop seen in the in vacuo peptides. The final eight structures are shown in Figure 6 in their bound orientation after superposition of aldolase. Seven of the eight final structures are similar to each other and have a Cα rmsd less than 3 Å after a least-squares superposition of all mainchain atoms. These seven structures are most like the average structure from

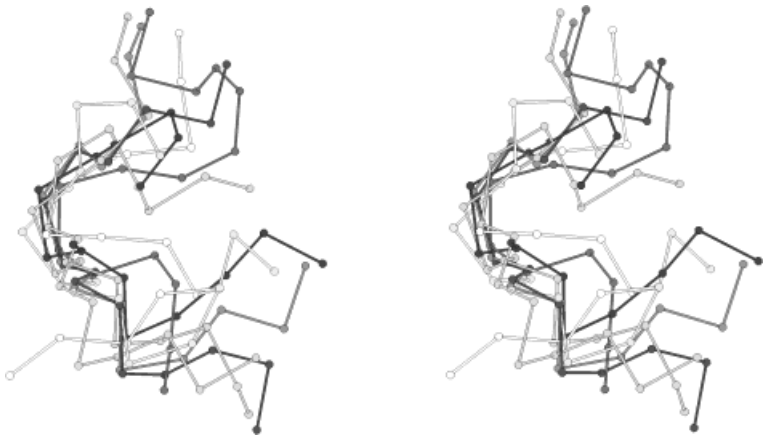


Fig. 6. Cα trace in stereo of the final eight B3P structures displayed in their orientation within the binding site of aldolase. Superposition is done with respect to aldolase, which is not drawn for clarity.

cluster 3 with an average Cα rmsd of 2.8 Å between this average cluster structure and the seven final structures. The structure that is not as similar to cluster 3, GS0E_{lo}SA.3, is closest to the average structure of cluster 2 (Cα rmsd equals 2.9 Å). The conformational heterogeneity found in the in vacuo 200 peptides is decreased for the eight structures (Table I), particularly in C-terminal residues. After superposition of B3P mainchain atoms in residues 4–9, the rmsd for the N-terminal residues decreased from 7.4 Å in the 200 in vitro structures to 6.4 Å in the final eight structures, and the C-terminal residues decreased from 8.6 to 6.7 Å.

Two binding orientations of B3P were predicted by the more extensive grid-based protocol described here for docking, in contrast to initial modeling work on this system, which resulted in only one orientation.[21] The B3P structure in either orientation (Fig. 7) agrees well with the NOE distance restraints (Table V) and is involved in interactions with many of the same aldolase residues. An rmsd value comparing these two structures (Table I) is significantly smaller than that for the final eight complexes, particularly with respect to the N- and C-terminal residues. In both orientations, the PSI loop has electrostatic contacts between Asp 6 and Asp 7 of B3P with aldolase residues Arg 42, Lys 107, and Arg 148. These residues are critical for binding the natural substrate of aldolase.[38,52] Arg 42 and Lys 107 bind the 6-phosphate of fructose 1,6-bisphosphate, and Arg 148 interacts with atoms at the substrate cleavage point. In the 0° orientation, Asp 6 forms hydrogen bonds with Lys 41 and Arg 42 of aldolase, while Asp 7 forms hydrogen bonds to Lys 107, Lys 146, and Arg 148. In the 180° orientation, the interactions are permuted but most of the residues on either molecule remain the same; Asp 6 binds to Lys 107 and Ser 35 while Asp 7 binds to Arg 148 and Arg 42. In both orientations, the side chain of Tyr 8 in B3P is directed away from aldolase and interacts with acidic side chains on the peptide, a feature considered important for phosphorylation control of the PSI loop.[21,36]

The model in Table V with the lowest E_{inter}, IG0E_{lo}SA.6, is nearly identical to the previous model for the B3P:

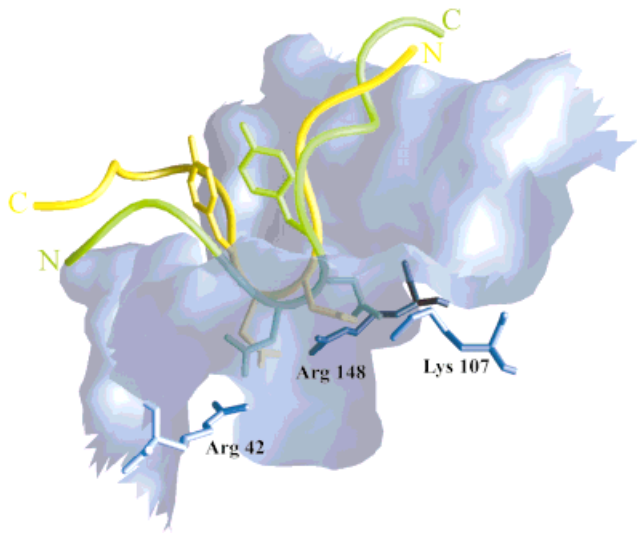


Fig. 7. The top structure in each orientation, highlighting the side chain interactions between peptide and protein. The B3P structures are a ribbon trace in green for IG0E_{lo}SA.6, or yellow for GS180*clu*SA.1. The binding site of aldolase is shown as a blue surface and selected aldolase residues that form hydrogen bonds to the peptide and that define the cavity volume described in the text. B3P side chain residues Asp 6 and Asp 7 are shown to bind in the same location, forming hydrogen bonds with aldolase residues Arg 42, Lys 107, and Arg 148.

aldolase complex.[21] Although it had the poorest measured E_{self}, the overall low energy derives from more intermolecular contact and a large buried surface area (425 Å^2).

CONCLUSIONS

The 15-residue peptide B3P was docked into the binding site of aldolase by using protocols that differed in the selection of a workable number of in vacuo etNOE structures from a large set, generation of an initial template for overlaying multiple peptide structures, and conformational sampling to a relaxed complex. A key result was the identification of two binding orientations of B3P (Fig. 7) from the initial grid search, whereas the second orientation was overlooked when the initial template was built

using interactive graphics. The orientations are related by ~180° rotation of B3P and reflect the near twofold symmetry of both the peptide sequence and the natural enzyme substrate. Both orientations satisfy the NOE distance restraints and form multiple hydrogen bonding interactions in the aldolase active site. Such a "reversible ligand" has been shown to exist in other systems as well, most notably peptidomimetic inhibitors of HIV protease[53] and proline-rich peptides that bind SH3 domains.[54] Given the conformational heterogeneity in B3P binding indicated by experimental NMR results, it is reasonable that both peptide orientations occur in solution.

All protocols generate reasonable models with the et-NOE restraints as determined from analysis of various structural features of the final docked B3P:aldolase complexes. A cluster analysis of the etNOE set of structures is a straightforward approach for selection of representative peptides from the large set of in vacuo etNOE structures, and may be useful when presented with significant conformational disorder. We find no complication starting from a set of pre-determined etNOE structures to dock the protein. Others suggested a procedure[17] whereby the etNOE restraints are applied de novo in the presence of the protein by building the peptide by fragments in the binding site. This fragment buildup is not necessary when following the protocols described here. We also find that an explicit grid-based search to find an initial B3P:aldolase model is a viable alternative to initial docking by interactive graphics and mostly avoids human intervention. Although all protocols generated reasonable complexes, more of the structures with the best score (listed in Table V) resulted from simulated annealing rather than Monte Carlo minimization. Similarly, another study that compared SA and MCM approaches[2] finds that SA provides structures with lower energies as well as structures closest to those in the crystallographic complexes.

The grid-based approach is more computationally costly than manual positioning by IG. If a thorough search of the docking site with a minimum of operator bias is desired, or if the protein has a poorly defined binding site, then the extra CPU time necessary for a grid-based search for an initial complex may be preferred. MCM searched a larger volume surrounding the initial position while SA provided a more conformationally relaxed structure with lower energy. Use of MCM followed by SA may be an improvement over the protocols explored by this study.

ACKNOWLEDGMENTS

This work was supported by a grant to C.B.P. from the NIH (R01-GM39478). C.B.P. was supported by a Research Career Development Award from the NIH (K04-GM00661) and A.P.R.Z. received a training fellowship from the NIH (5T32-GM08296). The computing facilities shared by the Structural Biology group are supported by grants from the Lucille P. Markey Foundation and the Purdue University Academic Reinvestment Program.

REFERENCES

1. Eisenmesser EZ, Zabell APR, Post CB. Accuracy of bound peptide structures determined by exchange transferred nuclear overhauser data: a simulation study. J Biomol NMR 2000;17:17–32.
2. Vieth M, Hirst JD, Dominy BN, Daigler H, Brooks CL, III. Assessing search strategies for flexible docking. J Comput Chem 1998;19:1623–1631.
3. Diller DJ, Verlinde CLMJ. A Critical evaluation of several global optimization algorithms for the purpose of molecular docking. J Comput Chem 1999;20:1740–1751.
4. Goldman BB, Wipke WT. QSD quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock). Proteins 2000;38:79–94.
5. Norel R, Lin SL, Wolfson HJ, Nussinov R. Molecular surface complementarity at protein-protein interfaces: The critical role played by surface normals at well-placed, sparce points in docking. J Mol Biol 1995;252:263–273.
6. Lin SL, Nussinov R. Molecular recognition via face center representation of a molecular surface. J Mol Graphics 1996;14:78–90.
7. Thorner DA, Wild DJ, Willet P, Wright PM. Similarity searching in files of three-dimensional chemical structures: flexible field-based searching of molecular electrostatic potentials. J Chem Inf Comput Sci 1996;36:900–908.
8. Sandak B, Nussinov R, Wolfson HJ. A method for biomolecular structural recognition and docking allowing conformational flexibility. J Comp Biol 1998;5:631–654.
9. Abagyan R, Totrov M, Kuznetsov D. ICM: a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem 1994;15:488–506.
10. Luty BA, Wasserman ZR, Stouten PFW, Hodge CN, Zacharias M, McCammon JA. A molecular mechanics/grid method for evalutaion of ligand-receptor interactions. J Comput Chem 1995;16:454–464.
11. Welch W, Ruppert J, Jain AN. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. Chem Biol 1996;3:449–462.
12. Caflisch A, Fischer S, Karplus M. Docking by Monte Carlo minimization with a solvation correction: application to an FKBP-substrate complex. J Comput Chem 1997;1997:6.
13. Aszodi A, Taylor WR. Hierarchic inertial projection: a fast distance matrix embedding algorithm. Comput Chem 1997;21:13–23.
14. Apostolakis J, Plückthun A, Caflisch A. Docking small ligands in flexible binding sites. J Comput Chem 1998;19:21–37.
15. Shoichet BK, Leach AR, Kuntz ID. Ligand solvation in molecular docking. Proteins 1999;34:4–16.
16. Makino S, Ewing TJA, Kuntz ID. DREAM++: Flexible docking program for virtual combinatorial libraries. J Comput Aid Mol Des 1999;13:513–532.
17. Maurer MC, Trosset J-Y, Lester CC, DiBella EE, Scheraga HA. New General Approach for determining the solution structure of a ligand bound weakly to a receptor: structure of a fibrinogen Aα-like peptide bound to thrombin(S195A) obtained using NOE distance constraints and an ECEPP/3 flexible docking program. Proteins 1999;34:29–48.
18. Pak Y, Wang S. Application of a molecular dynamics simulation method with a generalized effective potential to the flexible molecular docking problems. J Phys Chem B 2000;104:354–359.
19. Stigler R-D, Hoffmann B, Abagyan R, Schneider-Mergener J. Soft docking an L and a D peptide to an anticholera toxin antibody using internal coordinate mechanics. Structure 1999;7:663–670.
20. Moult J, Hubbard T, Fidelis K, Pedersen JT. Critical assessment of methods of protein structure prediction (CASP): round III. Proteins 1999;(Suppl. 3):2–6.
21. Schneider ML, Post CB. Solution structure of a band 3 peptide inhibitor bound to aldolase: a proposed mechanism for regulating binding by tyrosine phosphorylation. Biochemistry 1995;34:16574–16584.
22. Li Z, Scheraga HA. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. Proc Natl Acad Sci USA 1987;84:6611–6615.
23. Caflisch A, Niederer P, Anliker M. Monte Carlo docking of oligopeptides to proteins. Proteins 1992;13:223–230.
24. Caflisch A, Niederer P, Anliker M. Monte Carlo minimization with thermalization for global optimization of polypeptide conformations in Cartesian coordinate space. Proteins 1992;14:102–109.
25. Li Z, Scheraga HA. Structure and free energy of complex thermodynamic systems. J Mol Struct (Theochem) 1988;179:333–352.
26. Salhany JM. Erythrocyte Band 3 Protein. Boca Raton, FL: CRC Press, Inc. 1990; p 153–177.

27. Low PS. Structure and function of the cytoplasmic domain of band 3: center of erythrocyte membrane-peripheral protein interactions. Biochim Biophys Acta 1986;864:145–167.
28. Murthy SNP, Liu T, Kaul RK, Köhler H, Steck TL. the aldolase-binding site of the human erythrocyte membrane is at the NH_2 terminus of band 3. J Biol Chem 1981;256:11203–11208.
29. Strapazon E, Steck TL. Binding of rabbit muscle aldolase to band 3, the predominant polypeptide of the human erythrocyte membrane. Biochemistry 1976;15:1421–1424.
30. Jenkins JD, Badden DP, Steck TL. Association of phosphofructokinase and aldolase with the membrane of the intact erythrocyte. J Biol Chem 1984;259:9374–9378.
31. Tsai I-H, Murthy SNP, Steck TL. Effect of red cell membrane binding on the catalytic activity of glyceraldehyde-3-phosphate dehydrogenase. J Biol Chem 1982;257:1438–1442.
32. Harris SJ, Winzor DJ. Interactions of glycolytic enzymes with erythrocyte membranes. Biochim Biophys Acta 1990;1038:306–314.
33. Higashi T, Richards CS, Uyeda K. The interaction of phosphofructokinase with erythrocyte membranes. J Biol Chem 1979;254:9542–9550.
34. Rogalski AA, Steck TL, Waseem A. Association of glyceraldehyde-3-phosphate dehydrogenase with the plasma membrane of the intact human red blood cell. J Biol Chem 1989;264:6438–6446.
35. Low PS, Rathinavelu P, Harrison ML. Regulation of glycolysis via reversible enzyme binding to the membrane protein, Band 3. J Biol Chem 1993;268:14627–14631.
36. Eisenmesser EZ, Post CB. Insights into tyrosine phosphorylation control of protein–protein association from the NMR structure of a band 3 peptide inhibitor bound to glyceraldehyde-3-phosphate dehodrogenase. Biochemistry 1998;37:867–877.
37. Gamblin SJ, Cooper B, Millar JR, Davies GJ, Littlechild JA, Watson HC. The crystal structure of human muscle aldolase at 3.0 Å resolution. FEBS Lett 1990;262:282–286.
38. Gamblin SJ, Davies GJ, Grimes JM, Jackson RM, Littlechild JA, Watson HC. Activity and specificity of human aldolases. J Mol Biol 1991;219:573–576.
39. MacKerell ADJ, Bashford D, Bellott M, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
40. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4:187–217.
41. Karpen ME, Tobias DJ, Brooks CL. Statistical Clustering: techniques for the analysis of long molecular-dynamics trajectories. Analysis of 2.2-ns trajectories of YPGDV. Biochemistry 1993;32:412–420.
42. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH. Equation of state calculations by fast computing machines. J Chem Phys 1953;21:1087–1092.
43. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996;8:477–486.
44. Nicholls A, Sharp K, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
45. Gonzalez C, Rullmann JAC, Bonvin AMJJ, Boelens R, Kaptein R. Toward and NMR *R* Factor. J Magn Reson 1991;91:659–664.
46. Withka JM, Srinivasan J, Bolton PH. Problems with, and alternatives to, the NMR *R* Factor. J Magn Reson 1992;98:611–617.
47. Gronwald W, Kirchhofer R, Gorler A, et al. RFAC, a program for automated NMR R-factor estimation. J Biomol NMR 2000;17:137–151.
48. Zheng J, Post CB. Protein indirect relaxation effects in exchange-transferred NOESY by a rate-matrix analysis. J Magn Reson B 1993;101:262–270.
49. Desai P, Coutinho E, Srivastava S, Saran A. Structure of the epitope-fragment 579 to 601 in the GP41 transmembrane domain of HIV-1. Indian J Chem Sect A 2000;39:274–280.
50. Edmondson SP, Qiu LS, Shriver JW. Solution structure of the DNA-binding protein SAC7D from the hyperthermophile sulfolobus acidocaldarius. Biochemistry 1995;34:13289–13304.
51. Ohman A, Lycksell PO, Graslund A. A refined 3-dimensional solution structure of a carboxy-terminal fragment of apolipoprotein-CII. Eur Biophys J Biophys Lett 1993;22:351–357.
52. Dalby A, Dauter Z, Littlechilde JA. Crystal structure of human muscle aldolase complexed with fructose 1,6-bisphosphate: mechanistic implications. Protein Sci 1999;8:291–297.
53. Keinan S, Avnir D. Quantitative symmetry in structure–activity correlations: the near C-2 symmetry of inhibitor/HIV protease complexes. J Am Chem Soc 2000;122:4378–4384.
54. Saraste M, Musacchio A. Backwards and forwards binding. Nature Struct Biol 1994;1:835–837.