BIOORGANIC CHEMISTRY **26,** 295–308 (1998)
ARTICLE NO. BH981118

Searching for the Lowest Energy Conformation of Substrates in the Carboxypeptidase A Active Site Using Monte Carlo/Energy Minimization Techniques

Guiqing Liang and John F. Sebastian[1]

Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056

Received December 19, 1997

More than fifteen substrates of carboxypeptidase A (CPA) have been "docked" to the active site of the enzyme and searched for the lowest energy conformation of the substrates bound to the active site. The method employed combines Monte Carlo procedures with energy minimization (MC/EM procedures). The distances of $P_{1'}$–P_3 to $S_{1'}$–S_3 were measured. The computational results are consistent with the proposed binding interactions of CPA and its substrates and also with the promoted-water pathway of CPA catalysis. This study has demonstrated that MC/EM procedures are very useful for searching the conformations of substrates available in the active site of the enzyme.

INTRODUCTION

Bovine carboxypeptidase A (CPA) is a metallopeptidase containing a zinc ion bound to a single polypeptide chain of 307 amino acids. It catalyzes the hydrolysis of peptides/proteins with aromatic or large aliphatic side chain residues at the C-terminus (*1–3*). The important residues involved in the CPA catalysis are Arg-71, Arg-127, Asn-144, Arg-145, Tyr-248, Glu-270, Zn^{2+}, and the metal-bound water molecule (*3*). X-ray studies of bovine CPA–inhibitor complexes suggest that Arg-71 is hydrogen-bonded to the amide carbonyl oxygen in the P_2 site (*4*). The NH of the amide in the P_1 site is hydrogen-bonded to the hydroxyl oxygen of Tyr-248. The hydroxyl oxygen of Tyr-248 and the amide nitrogen of Asn-144 are both believed to be hydrogen bonded to the carboxylate oxygen in the $P_{1'}$ site. Arg-145 interacts with the carboxylate oxygens in the $P_{1'}$ site. It is also proposed that Arg-127 is hydrogen bonded to the amide oxygen in the P_1 site as a catalytic residue (*5*). In the promoted-water pathway, Glu-270 acts as a general base, abstracting a proton from zinc-bound water molecule (*3*).

We applied computational methodology to examine the binding interactions of CPA and its substrates which were proposed by Christianson and Lipscomb (*3*): Monte Carlo techniques along with energy minimization (MC/EM) using the Amber force field to search for the lowest energy conformations of the substrates available in the active site of CPA (*6*). The MC/EM techniques have been used to study the

[1] To whom correspondence and reprint requests should be addressed. Fax: (513) 529-5715. E-mail: SEBASTJF@miavx1.acs.muohio.edu.

0045-2068/98 $25.00

enzyme-inhibitor/substrate systems and the approach has turned out to be a useful tool for studies involving computer-aided design and the evaluation of enzyme inhibitors (*7*). In this paper we have made use of this methodology to study the binding interactions between searching for the lowest energy conformations of the substrates bound to the active site of the enzyme. The computational procedures are described in detail under Computational Methods.

COMPUTATIONAL METHODS

Protein Model

We used the X-ray structure of bovine carboxypeptidase A complexed with the phosphonate inhibitor *O*-[[(1*R*)-[[*N*-(phenylmethoxycarbonyl)-L-phenylalanyl] amino]isobutyl]hydroxyphosphinyl]-L-3-phenyllactate (FVF) (2.0 Å resolution) as the computational protein model (*8, 9*). The Protein Data Bank file, 7CPA, contains 307 amino acid residues, a Zn^{2+} ion, and the inhibitor FVF. When searching for the lowest energy conformation of substrates bound to the active site of the enzyme, the inhibitor was deleted from the complex and the substrates were manually docked to the active site. All of the computations were started with the enzyme coordinates given in 7CPA. The starting structures of the substrates were chosen at random. Two different manual dockings of the substrate to the active site of the enzyme have been applied to the conformational searching for each substrate. The final data shown in this paper are the average of the two trials. In most of the cases studied, the searched conformations of the substrate with the active site of the enzyme by two different manual dockings gave the same or very similar results.

Monte Carlo/Energy Minimization Procedures

We adopted the method described by Guida (*7*) to perform the conformational searches. The selected substructure, Amber force field, and the conformational search conditions were modified to fit to the carboxypeptidase A enzyme system and are described in detail below.

1. Substructure selection. The substructure included the residues in the enzyme's active site and the inhibitor or docked substrate and was subjected to computations. The selected residues in the active site are His-69, Arg-71, Glu-72, Arg-124, Arg-127, Asn-144, Arg-145, His-196, Ser-197, Tyr-198, Ser-199, Ile-201, Tyr-248, Ala-250, Gly-253, Ser-254, Ile-255, Asp-256, Thr-268, Glu-270, Phe-279, and zinc ion. There were more than 250 heavy atoms selected. During each Monte Carlo (MC) cycle, the inhibitor/substrates were unconstrained. The torsion angles of the inhibitor/substrates which were allowed to rotate during each MC step are shown in Fig. 1. In addition, the hydrogens of hydroxyl groups from tyrosine, serine, and threonine and the amino hydrogens were also unconstrained to allow for reorientation during energy minimization to obtain proper hydrogen-bonding geometries. All other residues were constrained by a force constant 100 kJ/Å^2, except Glu-270, which was constrained by a smaller force constant 10 kJ/Å^2 because of its putative movement during hydrolysis (*8, 10*).

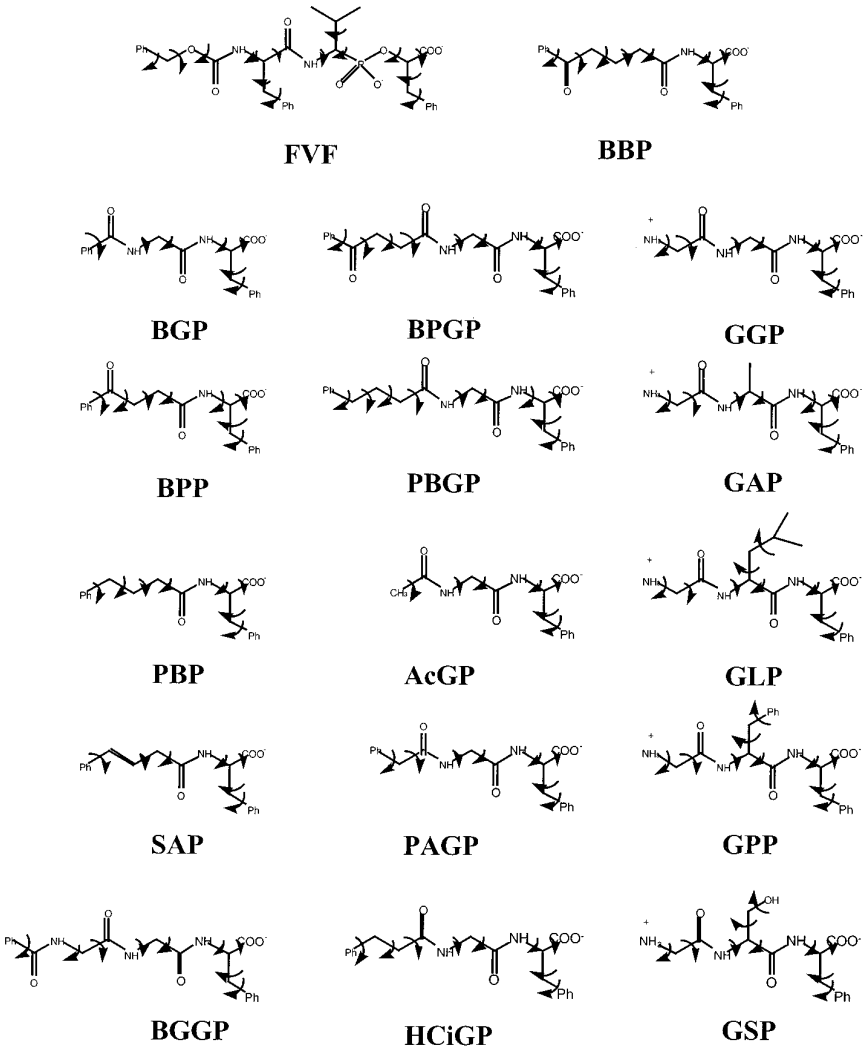

FIG. 1. The structures of the inhibitor FVF and all of the substrates subjected to the MC/EM conformational searches in the 7CPA active site model. Torsion angles rotated during each MC step are shown with the arrows: FVF; BBP, *N*-benzoylbutanoyl-L-Phe; BGP, *N*-benzoylglycyl-L-Phe; BPP, *N*-benzoylpropanoyl-L-Phe; PBP, 4-phenylbutanoyl-L-Phe; SAP, *N*-(*trans*-styrylacetyl)-L-Phe; BGGP, *N*-benzoylglycylglycyl-L-Phe; BPGP, *N*-(3-benzoylpropanoyl)glycyl-L-Phe; PBGP, *N*-(4-phenylbutanoyl) glycyl-L-Phe; AcGP, *N*-acetylglycyl-L-Phe; PAGP, *N*-(2-phenylacetyl)glycyl-L-Phe; HCiGP, *N*-hydrocinnamoylglycyl-L-Phe; GGP, Gly-Gly-Phe; GAP, Gly-Ala-Phe; GLP, Gly-Leu-Phe; GPP, Gly-Phe-Phe; GSP, Gly-Ser-Phe.

2. AMBER force field. Macromodel Version 4.5 was used for this computational study and the AMBER force field was employed for computations of the molecular energetics (*11–13*). There were some modifications made to the parameter set assigned to FVF, Zn, and its ligands (His-69, His-196, and Glu-72) by the BATCHMIN program (*7, 14–15*). In addition, explicit hydrogens were added to aryl moieties (phenyl rings from Phe-279, Tyr-198, Tyr-248, and inhibitor/substrates) in order to correctly model electrostatic interactions for aromatic residues (*16, 17*). The edge-to-face arrangement of Tyr-198 and Phe-279 is commonly observed in the crystal structures of CPA and its complexes (*18*). The detailed modified parameters are shown in the Appendix. Thus the aromatic hydrogens were assigned a charge of +0.15e with the corresponding aromatic carbons being assigned a charge of −0.15e. Due to a large amount of atoms (more than 250 atoms) subjected to conformational searching of CPA along with its substrates/inhibitors, the charge of Zn ion applied in the Amber force in MacroModel program is not very critical. The Zn atom was assigned a charge of +2.0e. Another charge on Zn (+1.7e) was also tried with the FVF–CPA complex; the result showed no significant difference from the +2.0e case.

3. Definition of conformational search conditions. In order to test the MC/EM procedures in searching for the conformations of substrates in the active site of the enzyme, a set of highly defined coordinates from the X-ray structure of 7CPA along with the inhibitor was applied for defining the best conformational search conditions as described by Guida (*7*). Two computational calculations were carried out. First, the selected substructure of 7CPA with the inhibitor was subjected to local energy minimization, which was terminated when the energy gradient RMS fell below 0.01 kJ/Å. Second, the inhibitor was removed and ''redocked'' to the enzyme's active site and the inhibitor along with the residues involved in the active site were selected as the substructure and subjected to the MC/EM procedures. The default nonbonded cutoff protocol employed by the BATCHMIN program was used except a van der Waals cutoff was changed to 8.0 Å. In general, all conformers that were different from the global minimum energy conformation by no more than 20 kJ/mol were saved. The energy minimization was terminated either when energy gradient RMS fell below 0.05 kJ/Å or after 500 iterations. After completion of the MC/EM conformational search, the lower energy conformers were subjected to further energy minimization to reduce the energy gradient RMS to below 0.01 kJ/Å. Water solvent was applied for both computations and the PRCG (Polak-Ribiere Conjugate Gradient) was used for the energy minimization. The computational structure(s) of the inhibitor FVF were compared to the X-ray structure. Only when computational structure(s) were visually the same as or close to the X-ray structure would the conformational search conditions used then be applied for probing the lowest energy conformation of substrates docked to the active site of the enzyme. Otherwise the conditions would be refined until the requirements were met. The approach described above has proved to be reliable for searching the conformations of substrates in the enzyme's active site.

4. Conformational search of substrates in the enzyme's active site. The inhibitor was deleted from the complex and the studied substrate was manually ''docked'' to the active site in a position similar to that of the inhibitor. Then the substructure

including enzyme's active site and the docked substrate was defined and subjected to the MC/EM procedures to perform the conformational searches.

Platform and Software

The computations were carried out on a Silicon Graphics-Indy workstation. Macromodel V4.5 was used. About 2000–5000 MC/EM steps were required for searching the "lowest" energy conformation of substrate in the active site of the enzyme which took 10–12 CPU days.

RESULTS

For the purpose of defining computational parameters, the crystal structure of the CPA–FVF complex (*8, 9*) was used as a starting point for testing the MC/EM procedures. The computational result of the substructure local energy minimization (Fig. 2) showed that the energy minimization resulted in a little change of the inhibitor FVF relative to the crystal structure at the $P_{1'}$ site and the peptide backbone at the P_1–P_3 sites. After the conformational search of the inhibitor bound to the active site of CPA, two conformers with the lowest energy were selected to compare with the crystal structure. The comparison shows that the two searched lowest energy conformers of the inhibitor FVF also have changed little relative to the crystal structure at the C-terminal residue and the peptide backbone in the P_1–P_3 sites (Fig. 3). A comparison of the crystal and computed structures of FVF based on the torsion angles is shown in Table 1. The results of the two computations showed that the side chains of FVF at the P_2–P_3 site have significant movement compared to the crystal structure of the inhibitor. This might be expected as the side-chain conformations have only slightly different energies. For this reason, the side-chain conformation that we got computationally could very well be quite different from the crystal structure. In our studies, we only focused on the substrate's C-terminal residue and the peptide backbone at the P_1–P_3 sites. Therefore we concluded that the computational parameters that we had defined above were suitable and could be applied to the lowest energy conformational searches of the substrates "docked" to the active site of CPA without further modification.

A typical result of the lowest energy conformational searching of the substrates bound to the enzyme's active site is shown in Fig. 4, using *N*-benzoylglycyl-L-phenylalanine (BGP) as an example. The values of the distances between substrate and enzyme's active site after the lowest energy conformational search are summarized in Table 2. For all the substrates studied, the distances between the terminal residue's carboxylate oxygens and Arg-145 nitrogens were less than 3 Å (**B**). The Tyr-248 hydroxyl oxygen and this carboxylate oxygen is about 2.6 Å away from each other (**C**). The Asn-144 α-amide nitrogen is approximately 3.7 Å away from the carboxylate oxygen, which indicates an unlikely hydrogen binding interaction (**A**). The Arg-127 nitrogen is about 2.7 Å away from the carbonyl oxygen at P_1 site (**E**). In general for dipeptide analogs, the Arg 71 nitrogen is less than 3 Å away from carbonyl oxygen at P_2 site (**F**) and the Tyr-248 hydroxyl oxygen is also less

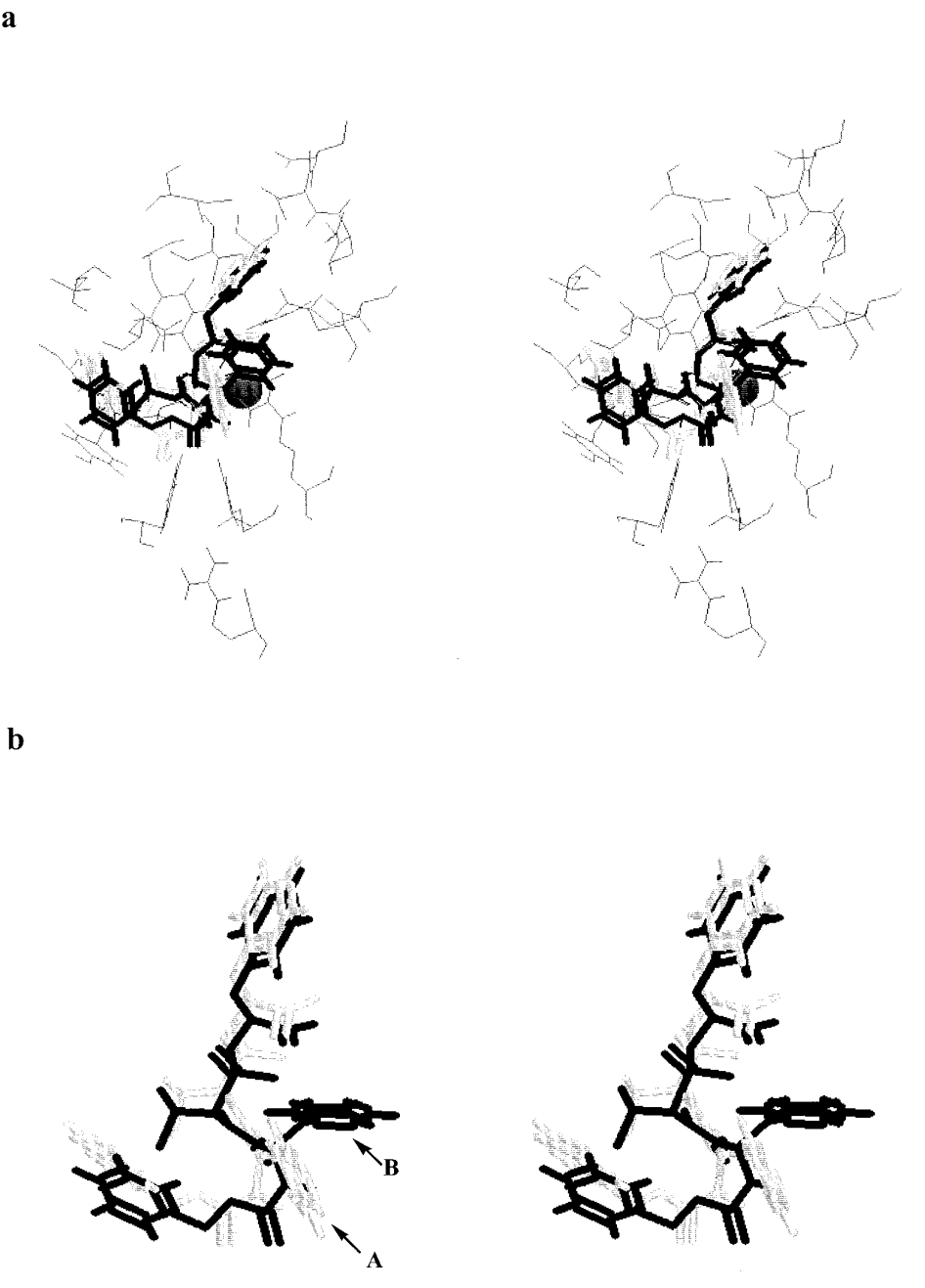

FIG. 2. The comparison of the X-ray crystal and computed structures of the inhibitor FVF after the substructure local energy minimization. (a) The stereoview comparison of FVF in the active site. Wireframe, the residues in the active site; Spacefill, Zn ion; Sticks, X-ray structure of FVF (light) and computed structure of FVF (dark). (b) The stereoview comparison of FVF with the deletion of the active site. A, X-ray crystal structure of FVF; B, the computed structure of FVF.

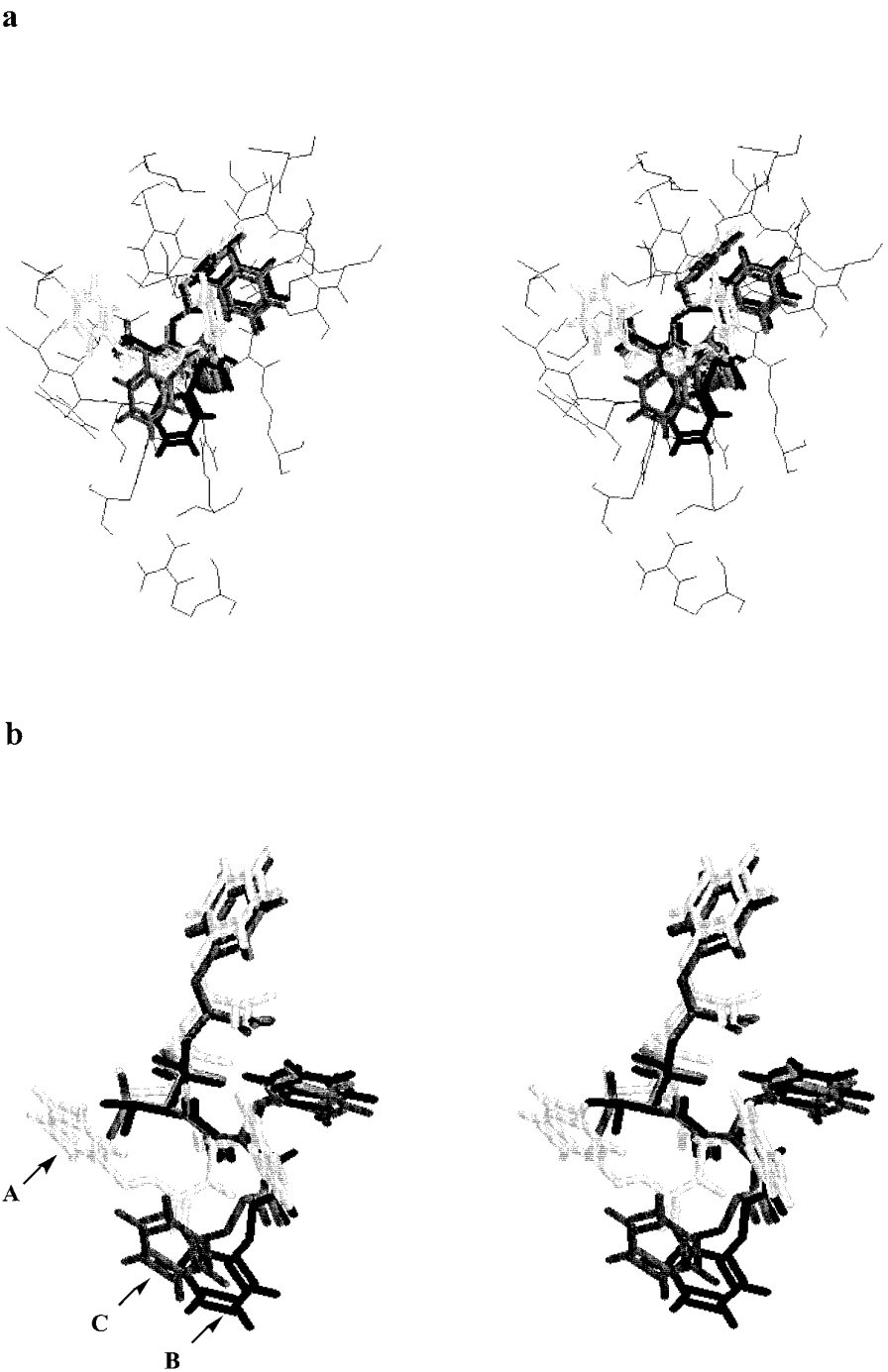

FIG. 3. The comparison of the X-ray crystal and computational structures of the inhibitor FVF after conformational search of the inhibitor in the active site of the enzyme. (a) The stereoview comparison of the computed structures of FVF and X-ray structure in the active site. Wireframe, the residues in the active site; Spacefill, Zn ion; Sticks, X-ray structure of FVF (light) and the two computed structures of FVF with the lowest energy (dark); (b) The stereoview comparison of FVF with the deletion of the active site. A, X-ray crystal structure of FVF; B and C, the two conformers of FVF searched with the lowest energy. B is the one with the global energy minimum.