



Pergamon

SCIENCE @ DIRECT®

Tetrahedron: *Asymmetry* 14 (2003) 3385–3399

TETRAHEDRON:  
ASYMMETRY

# Studying enzyme enantioselectivity using combined *ab initio* and free energy calculations: $\alpha$ -chymotrypsin and methyl *cis*- and *trans*-5-oxo-2-pentylpyrrolidine-3-carboxylates

F. Felluga,<sup>a</sup> G. Pitacco,<sup>a</sup> E. Valentin,<sup>a</sup> A. Coslanich,<sup>b</sup> M. Fermeglia,<sup>b</sup> M. Ferrone<sup>b</sup> and S. Prichl<sup>b,\*</sup>

<sup>a</sup>Department of Chemical Sciences, University of Trieste, Via Licio Giorgieri 1, 34127 Trieste, Italy

<sup>b</sup>Computer-aided Systems Laboratory, Department of Chemical, Environmental and Raw Materials Engineering, DICAMP, University of Trieste, Piazzale Europa 1, 34127 Trieste, Italy

Received 18 July 2003; accepted 1 September 2003

**Abstract**—The application of a computational approach, based on molecular dynamics (MD) simulations and quantum mechanical-free energy (QM-FE) calculations, to explain the different substrate specificity and enantioselectivity of  $\alpha$ -chymotrypsin ( $\alpha$ -CT) in the hydrolysis of methyl *cis*- and *trans*-5-oxo-2-pentylpyrrolidine-3-carboxylates is described. By applying a combination of molecular mechanics energy derived from MD simulations in explicit solvent, and solvation free energy derived from a continuum solvation model, we have calculated reasonable absolute free energies of binding ( $\Delta G_{\text{bind}}$ ) for each  $\alpha$ -CT/enantiomer complex formation, and elucidated the balanced nature of the factors contributing to  $\Delta G_{\text{bind}}$ . Furthermore, our calculations based on QM-FE techniques have yielded an insight into the major issues affecting the observed enantioselectivity in the hydrolysis of substrate ester bonds by  $\alpha$ -chymotrypsin.

© 2003 Elsevier Ltd. All rights reserved.

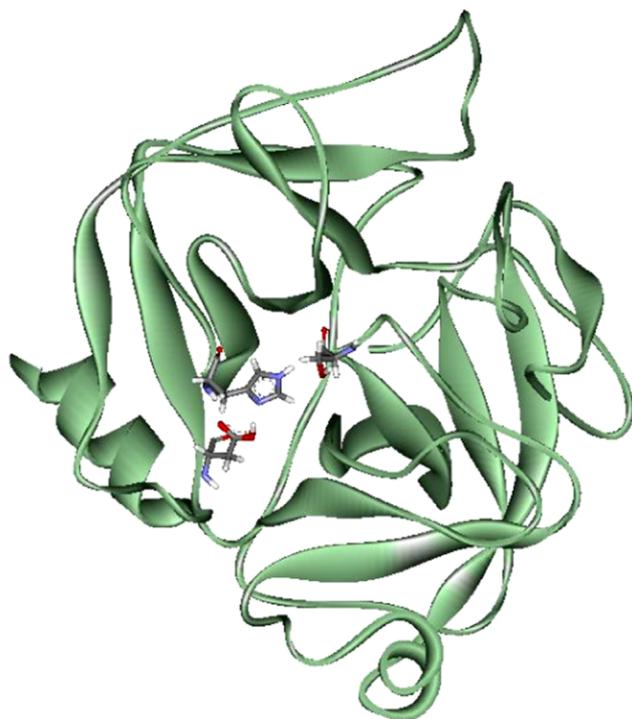
## 1. Introduction

Enzymes have the ability to recognize and selectively act on their substrates in a stereospecific manner.<sup>1,2</sup> Exquisite enantioselectivity is among the most important properties of enzymes to synthetic chemists. Under this perspective, the ultimate goal would be the ability to tailor or design enzyme–solvent systems to catalyze any specific chemical reaction at will.

Notwithstanding the large amount of knowledge at our disposal, usually no definitive evidence can be derived for enzyme reaction mechanisms from the applications of experimental techniques alone.<sup>1</sup> At this point, the role of physical/theoretical chemists comes into play, since they can provide a link between a given enzyme structure, as contributed by the above mentioned experimental techniques, and the relevant function, as determined by experiments on enzyme-driven reactions.<sup>3,4</sup>

In the field of polymers, and specifically of biopolymers, theoretical structure–function investigations have proven very challenging for a plethora of reasons. The first when considering protein molecules, is their size; indeed, the number of atoms is usually relatively large and computer experiments based on molecular dynamics on such systems are rather time consuming and hampered by several complications, even when the intermolecular potential functions employed are rather analytically simple. Even with an ‘ideally perfect’ classical force field, able to describe very accurately the energy of system, the adequate sampling of the conformational space available to an enzyme (or to a physically bonded enzyme–substrate complex) would remain a sort of chimera.<sup>5</sup> A further complexity is intrinsic in the nature of classical force fields: their inability to simulate the processes of bonds breaking and/or forming. Chemical reactions are among the most important phenomena to which quantum chemical approaches can be applied; accordingly, quantum mechanical (QM) calculations have played an essential role in understanding gas-phase reaction energetics and dynamics. In combination with either explicit molecular mechanical (MM) models or continuum models, QM approaches have also been applied to simulate chemical reactions in the condensed phase.<sup>6,7</sup> Even more challenging has been

\* Corresponding author. Tel.: +39-040-5583750; fax: +39-040-569823; e-mail: [sabrinap@dicamp.units.it](mailto:sabrinap@dicamp.units.it)



**Figure 1.** Energy-relaxed model of three-dimensional structure of  $\alpha$ -CT (ribbon model). The catalytic amino acid triad is highlighted in color using a stick style.

the effort to simulate chemical reactions in complex heterogeneous environments such as in enzymes.<sup>8</sup> Unfortunately, the computational time required for QM calculations rises steeply with the number of atoms and electrons in the system, and therefore it is very difficult to apply QM calculations to large biological molecular assemblies.

In the late 70's, Warshel and Levitt introduced the concept of combining molecular quantum mechanical (QM) and molecular mechanical (MM) methods.<sup>9</sup> This approach limits the quantum chemical description to the reaction center and uses a computationally efficient classical treatment for the remainder of the molecule. Several other QM/MM approaches have been reported since that pioneering work,<sup>4,10–23</sup> which differ substantially in the type of QM and MM method used, and in the treatment of the QM/MM interactions. Accordingly, the level of quantum mechanical theory involved varied from computer-demanding but fairly accurate ab initio Hartree–Fock and density functional techniques to semiempirical approaches and empirical valence bond (EVB) descriptions.<sup>24</sup> For the time being, the most practical QM/MM approach to study enzyme catalysis is the quantum mechanical/free energy perturbation method (QM/FE) proposed by Kollman et al.,<sup>25</sup> and based on the original work of Jorgensen et al.<sup>26,27</sup> The main feature of Kollman's method is the static (gas phase) description of the reacting solute. The static geometry allows for a completely classical simulation of a chemical reaction in solution to calculate the relative free energies. This approach assumes, however, that

solvation has a little if no influence on the course of the reaction, and that it only affects the energy profile.

Although several successful applications give support to this assumption for reactions of ordinary organic solutes,<sup>28–33</sup> three major challenges must be overcome before this approach can be used for complex enzymatic reactions:

- Model system fragments can not usually simply be optimized in the gas phase to obtain a relevant reaction pathway, due to the preorganization of the enzyme, which fixes the relative geometry of the fragments and keeps them from moving freely relative to one another as would happen on the gas phase reaction pathway;
- The so-called 'link atom problem', the alias given in the absence of an obvious way to describe, in a correct fashion, the energies at the junction between covalently bonded molecular mechanical and quantum mechanical atoms, which almost always occur in enzymatic reactions. The simple organic reactions are free of this concern because there are no covalent bonds between the reacting solutes (whose energies are evaluated quantum mechanically) and the solvent molecules (whose noncovalent interactions with each other and the solute are described by molecular mechanics);
- The generation of charges for the quantum mechanical atoms, in order to calculate their interaction with the molecular mechanical atoms.

Concerning point a), the general QM/FE approach proposed by Kollman et al.<sup>25</sup> meets the challenge since it relies on the fact that non-covalent interactions involve a much less stringent directionality than covalent interactions. For point b), the problem is mitigated in the method through a suitable treatment of the charges used on the QM atoms and a judicious choice of the restraints. Finally, for point c) the technique of choice is the restrained electrostatic potential (RESP) approach,<sup>34–37</sup> due to its notable features: (i) That it is identical to the approach used to derive molecular mechanical electrostatic charges for the protein, thus automatically leading to balanced protein–protein and protein–substrate interactions and (ii) because the Lagrangian constraints in the RESP method can be employed in a general way, which is an essential component in the strategy employed to circumvent the link atom problem outlined above.

$\alpha$ -Chymotrypsin is a prototypical serine protease enzyme that has been studied most extensively in this respect.<sup>38,39</sup> In nature, this enzyme cleaves the peptidic bond of aminoacids characterized by hydrophobic side chains.<sup>1</sup> It also hydrolyzes ester bonds.<sup>40,41</sup> The hydrolytic reaction of  $\alpha$ -CT is invariably initiated by a nucleophilic attack on the scissile amide or ester carbonyl bond by the deprotonated hydroxyl group of Ser 195, resulting in an acylated enzyme intermediate.<sup>42,43</sup> Besides Ser 195, the catalytic activity of  $\alpha$ -CT depends on two other amino acid residues: Histidine 57 and Aspartate 102. These three amino acids are distant from one another in the primary structure of the protease, but close together in the folded, native con-

formation (see Fig. 1). In this context, the reaction mechanism of  $\alpha$ -chymotrypsin was deduced, in part, from its three-dimensional structure determined by X-ray crystallography.<sup>44</sup>

At the active site of  $\alpha$ -chymotrypsin there is an extended binding region composed of several subsites, of which one in particular, denoted as S1, plays a dominant role in substrate recognition. At S1 we find three distinct loci that interact with a given substrate in the formation of the enzyme–substrate complex, in addition to the nucleophilic OH group of Ser 195 mentioned above. First, a large, hydrophobic cleft (*hc*), a crevice in the surface of the protein that is bordered by the side chains of several hydrophobic amino acid residues. This pocket comprises residues 189–194 on one side, and 214–220 on the other. S1 also contains a hydrogen-bonding site (*hb*) where, in the case of peptide bond hydrolysis, the amide hydrogen of the substrate is directed towards the backbone peptide C=O of Ser 214, and a cavity of restricted volume (*rv*), in which the  $\alpha$ -H of the substrate is directed towards the  $\alpha$ -H of Met 192.<sup>45,46</sup>

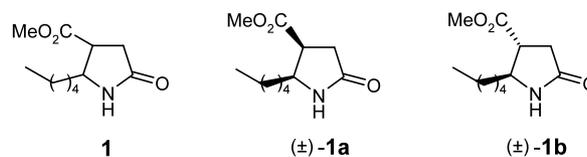
The transition states for the acylation of  $\alpha$ -CT by its substrates are thought to be similar to the tetrahedral intermediates for the corresponding reactions.<sup>47</sup> Whereas the tight binding of substrates to proteases is nowadays fully understood,<sup>48,49</sup> it is not clear why, in some cases, the amide or the ester bonds of the substrates remain intact in the complexes and are hydrolyzed so slowly if at all, especially in the case of enantiomeric substrates. Two main reasons can be envisaged to explain the slow to no hydrolysis of serine proteases substrates:

1. The enzyme may bind the substrate in such a way that favorable noncovalent bonds in the enzyme–substrate complex have to be disturbed upon formation of an enzyme–transition-state complex. In this way, the large binding energy of the Michaelis–Menten-like complex (MMC) is used to increase the energy of the transition state for bond cleavage.<sup>50</sup>
2. Following the cleavage of the amide/ester bond, the newly formed termini are held in close proximity in the modified substrate–protease acyl–enzyme complex (AEC), greatly favoring the reformation of the cleaved bond. The tight binding inhibits the hydrolysis of the covalent acyl–enzyme complex by preventing a water molecule from entering the active site.<sup>51–53</sup>

Herein, we have applied an ab initio quantum mechanics (QM) and molecular dynamics/free energy (MD-FE) calculations based approach (QM/MD-FE)<sup>25</sup> to investigate the enantioselectivity of  $\alpha$ -chymotrypsin towards ester bond cleavage in  $\beta$ -methoxycarbonyl- $\gamma$ -*n*-pentyl- $\gamma$ -lactams.<sup>54</sup> The  $\gamma$ -lactam nucleus (pyrrolidin-2-one) characterizes many compounds possessing biological and pharmaceutical activities.<sup>55</sup> Among the compounds containing the  $\gamma$ -lactam moiety, lactacystin<sup>56,57</sup> plays a major role as a potent 20S proteasome peptidase inhibitor, and constitutes a synthetic challenge for researchers owing to the presence of four contiguous

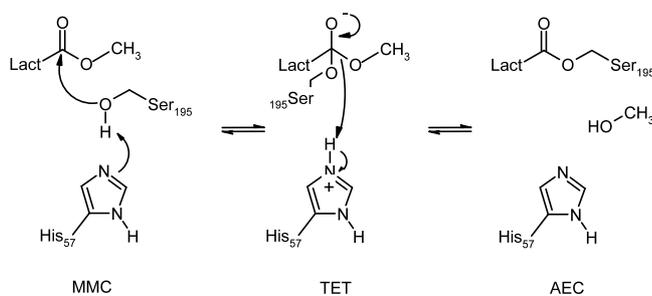
stereocenters.<sup>58,59</sup> Further examples are represented by pilolactam,<sup>60</sup> recently patented by Garst et al.<sup>61</sup> as a muscarinic active principle, and Rolipram, an antidepressant and phosphodiesterase inhibitor synthesized by two different research groups,<sup>62,63</sup> and currently manufactured by Schering Plough.

Among the plethora of substances featuring the lactam ring moiety as a structural component,  $\beta$ -carboxy- $\gamma$ -lactams are interesting compounds since they can be considered as aza analogues of paraconic acids. The latter form an interesting, small class of biologically active trisubstituted  $\gamma$ -butyrolactones,<sup>64–66</sup> carrying a carboxylic group at the  $\beta$  position. Accordingly, two years ago some of us reported the resolution of a series of methyl esters of 1-alkyl-5-oxo-pyrrolidine-3-carboxylic acids by chemo-enzymatic hydrolysis of the ester function.<sup>67</sup> Among the suitable enzymes available on the market,  $\alpha$ -CT turned out to be the choice for the enantiomeric resolution of these lactams with high enantiomeric excess (ee). The specificity of the enzyme and the high enantiopreference observed experimentally were successively fully rationalized by means of molecular mechanics/dynamics simulations on the corresponding enzyme–substrate complexes.<sup>50</sup> Following this work, we studied<sup>68</sup> the  $\alpha$ -CT mediated kinetic resolution of diastereomeric *cis*- and *trans*-methyl 5-oxo-2-pentylpyrrolidine-3-carboxylates of type **1** (see Fig. 2). The aim of this work was to obtain compounds **1a** and **1b** in enantiomerically pure form, as promising candidates for biological and toxicological activity.



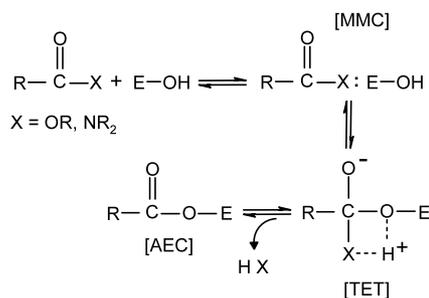
**Figure 2.** Methyl 5-oxo-2-pentylpyrrolidine-3-carboxylates considered for enantioselective hydrolysis by  $\alpha$ -CT.

Both experiment and theory<sup>69</sup> suggest that, for hydrolysis of an amide or ester bond by  $\alpha$ -CT, the rate-determining step is the formation of the tetrahedral intermediate during the first step (TET), which is acylation by the enzyme (see Fig. 3a). Furthermore, there is much evidence that both the structure and free energy of the tetrahedral intermediate during acylation are very similar to those of the transition state for its formation from the initial noncovalent (Michaelis–Menten) complex (MMC).<sup>69,70</sup> Given this, we have calculated the activation free energies for the acylation step of ester bond cleavage (i.e. from MMC to TET) in  $\alpha$ -CT/(+)- and (–)-**1a**, and  $\alpha$ -CT/(+)- and (–)-**1b** complexes. Comparison of the ester bond hydrolysis of similar reactive-site sequences in  $\alpha$ -CT–substrate complexes had allowed us to investigate factors responsible for the observed enantiopreference of the enzyme. Molecular dynamics simulations of 500 ps were also carried out for the  $\alpha$ -CT/(+)- and (–)-**1a**, and  $\alpha$ -CT/(+)- and (–)-**1b** tetrahedral intermediates [TET-(+)- and (–)-**1a**, and TET-(+)- and (–)-**1b**], and for the  $\alpha$ -CT/(+)-



**Figure 3.** (a, top) Schematic representation of the reaction pathway of serine protease catalysis of amide and ester bonds. E-OH represents the  $\alpha$ -CT hydroxyl group on Ser 195. According to text, the following abbreviations are used: MMC, Michaelis–Menten complex; TET, tetrahedral intermediate; AEC, acyl–enzyme intermediate. (b, bottom) Reaction mechanism of the acylation step of  $\alpha$ -CT-catalyzed hydrolysis and the structures of MMC, TET and AEC studied in this work. LACT stands for the residual lactam ring.

and (–)-**1a**, and  $\alpha$ -CT/(+)- and (–)-**1b** acyl–enzyme intermediates [AEC-(+)- and (–)-**1a** and AEC-(+)- and (–)-**1b**] (see Fig. 3b).



Calculations for the TET-1 intermediates were performed to verify how the catalytic groups are located into the active site for the hydrolysis of the ester bond. Acyl–enzyme calculations were carried out with the purpose of seeing how the newly formed termini in the substrate and the acyl group are located within the enzyme active site for the possible reformation of the cleaved ester bond.

## 2. Methods

### 2.1. The QM/MD-FE approach

The theoretical background of the adopted quantum mechanical/molecular dynamics-free energy approach is described in detail in the original paper by Kollman et al.<sup>25</sup> Therefore, in this section we will give only a brief outline of the method. According to the basic principle of the method, *ab initio* calculations for the reactive part of the enzyme–substrate complex are combined with the classical treatment of the interaction free energy between the quantum mechanical system and the classical environment, as well as within the classical environment itself. In the first step, the residues of the complex that are crucial in the enzymatic reaction are

cut out, and the relevant dangling bonds are saturated with hydrogens (referred to as ‘link atoms’). Geometry optimizations of this QM model system is then performed, with constraints on a few selected internal coordinates, which ensures that the relative orientation of the model fragments stays close to the preorganized geometry in the enzyme. It is important that these constraints are included in the QM treatments of enzymatic reaction centers because pure gas phase optimizations, in which the fragments are allowed to move freely, often yield an energetically very different reaction pathway.<sup>25</sup> On the basis of the QM-optimized structures, RESP fitting methodology<sup>34–37</sup> was applied to derive atomic partial charges, and the model fragments were reinserted into the protein coordinate frame. Resorting to the use of the standard force field representation for the QM atoms, extended by a few additional potential terms to restrain the reinserted model system to its QM-optimized structure, allowed the free energy of interaction to be calculated. The global free energy change,  $\Delta G_{\text{tot}}$ , between two points along the reaction pathway can be approximated by the following relationship:

$$\Delta G_{\text{tot}} = \Delta E_{\text{QM}} + \Delta G_{\text{int}} \quad (1)$$

in which  $\Delta E_{\text{QM}}$  represents the difference in *ab initio* energy, and  $\Delta G_{\text{int}}$  is the difference of the free energy of interaction.

### 2.2. Starting structure definition of the noncovalent (Michaelis–Menten) complexes

The starting 3-D model of  $\alpha$ -chymotrypsin ( $\alpha$ -CT) was based on the X-ray crystallographic structure of the  $\alpha$ -CT–eglin C complex (PDB code: 1ACB<sup>71</sup>). Water molecules in the coordinate file were included, and hydrogens were added to the protein backbone and side chains with the PARSE module of the AMBER 6.0 package.<sup>72,73</sup> All ionizable residues were considered in the standard ionization state at neutral pH. The all-atom force field (FF) parameters by Cornell et al.<sup>36</sup> (in *parm94.dat* file of the AMBER 6.0 code) was applied for protein relaxation. The primary cut-off distance for non-bonded interaction was set to 12 Å, the cut-off taper for the Coulomb and van der Waals interactions were 1.2 and 2, respectively. The GB/SA continuum solvation model<sup>74,75</sup> was used to mimic a water environment. Geometry refinement was carried out using the SANDER module via a combined steepest descent–conjugate gradient algorithm, using as a convergence criterion for the energy gradient of the root-mean-square of the Cartesian elements of the gradient equal to 0.01 kcal/(mol Å). As expected, no relevant structural changes were observed between the active site of the  $\alpha$ -CT relaxed structure and the original 3-D structure.

The model structures of the four possible enantiomers of the  $\beta,\gamma$ -disubstituted- $\gamma$ -lactams **1** were generated using the 3-D sketcher tool of Materials Studio.<sup>76</sup> All the molecules were subjected to an initial energy minimization using Discover.<sup>77</sup> In this case, the convergence

criterion was set to  $10^{-4}$  kcal/(mol Å). The conformational search was carried out using a combined molecular mechanics/molecular dynamics simulated annealing (MDSA) protocol.<sup>78</sup> Accordingly, the relaxed structures were subjected to 5 repeated temperature cycles (from 298 K to 1000 K and back) using constant volume/constant temperature (NVT) MD conditions. At the end of each annealing cycle, the structures were again energy minimized to converge below  $10^{-4}$  kcal/(mol Å), and only the structures corresponding to the minimum energy were used for further modelling. The electrostatic charges for the geometrically optimized lactam molecules were obtained by RESP,<sup>34–37</sup> and the electrostatic potentials were produced by single-point quantum mechanical calculations at the Hartree–Fock level with a 6-31+G\* basis set.<sup>79</sup>

The optimized structures of the enantiomeric substrates were docked into the  $\alpha$ -CT active site according to a procedure validated for *N*-substituted  $\beta$ -methoxycarbonyl- $\gamma$ -lactams.<sup>50</sup> To proceed with the docking simulation, all non-polar hydrogen atoms of the small organic molecules were deleted and their charges automatically added to those of the corresponding carbon atom by the program AutoTors included in the suite AutoDock 3.0.<sup>80</sup> The relevant grids of affinity potentials used by AutoDock were calculated by running the program AutoGrid. In order to encase a reasonable region of the protein surface and interior volume, centered on the crystallographic identified binding site, the grids were 60 Å on each side. Grid spacing (0.375 Å), and 120 grid points were applied in each Cartesian direction so as to calculate mass-centered grid maps. Amber 12-6 and 12-10 Lennard-Jones parameters were used in modeling van der Waals interactions and hydrogen bonding (N–H, O–H and S–H), respectively. In the generation of the electrostatic grid maps, the distance dependent relative permittivity of Mehler and Solmajer<sup>81</sup> was applied.

For the docking of each lactamic enantiomer to the protein, three hundred Monte Carlo/Simulated Annealing (MC/SA) runs were performed, with 100 constant temperature cycles for simulated annealing. Translation, quaternion parameters and torsions were set at random before SA runs. Each cycle had a maximum of 20,000 accepted or rejected moves, the minimal energy structure being passed to the next cycle. The temperature was reduced by a 0.95 factor per cycle from an initial value of  $RT=100$  cal/mol. For these calculations, the GB/SA implicit water model<sup>74,75</sup> was again used to mimic the solvated environment. The rotation of the angles  $\phi$  and  $\varphi$ , and the angles of side chains were set free during the calculations. All other parameters of the MC/SA algorithm were kept as default. Following the docking procedure, all structures of each lactam enantiomer were subjected to cluster analysis with a tolerance of 1 Å for an all-atom root-mean-square (RMS) deviation from a lower-energy structure representing each cluster family. The structure with the lowest interaction energy was selected for further evaluation.

Each best enantiomeric substrate/ $\alpha$ -chymotrypsin complex resulting from the automated docking procedure was further refined in the AMBER suite using the quenched molecular dynamics method (QMD). In this case, 100 ps MD simulation at 298 K were employed to sample the conformational space of the substrate–enzyme complex in the GB/SA continuum solvation environment.<sup>74,75</sup> The integration step was equal to 1 fs. After each ps, the system was cooled to 0 K, the structure extensively minimized, and stored. To prevent global conformational changes of the enzyme, the backbone of the protein binding site was constrained by a harmonic force constant of 100 kcal/Å, whereas the amino acid side chains and the ligands were allowed movement without any constraint.

The best energy configuration of each complex resulting from the previous step was solvated by adding a sphere of TIP3P water molecules<sup>82</sup> with a 20 Å radius from the  $O_\gamma$  of the catalytic Ser 195 with the use of the *cap* option of the LEAP module of AMBER 6.0. The protein complex was neutralized adding a suitable number of counter ions ( $\text{Na}^+$  and  $\text{Cl}^-$ ) in the positions of largest electrostatic potential, as determined by the module CION of the AMBER 6.0 platform. The counter ions, which had distances larger than 16 Å from the active site, were fixed in space during all simulations to avoid artifactual long range electrostatic effects on the calculated free energies. After energy minimization of the water molecules for 1500 steps, and MD equilibration of the water sphere with fixed solute for 20 ps, further unfavorable interactions within the structures were relieved by progressively smaller positional restraints on the solute (from 25 to 0 kcal/(mol Å<sup>2</sup>) for a total of 4000 steps. Each system was gradually heated to 298 K in three intervals, allowing a 5 ps interval per each 100 K, and then equilibrated for 50 ps at 298 K, followed by 400 ps of data collection runs, necessary for the estimation of the free energy of binding (vide infra). After the first 20 ps of MD equilibration, additional TIP3P water molecules were added to the 20 Å water cap to compensate for those that were able to diffuse into gaps of the enzyme. The MD simulations were performed at constant  $T=298$  K using the Berendsen coupling algorithm,<sup>83</sup> an integration time step of 1 fs, and the applications of the SHAKE algorithm<sup>84</sup> to constrain all bonds to their equilibrium values, thus removing high frequency vibrations. Long-range non-bonded interactions were truncated by using a 20 Å residue-based cut-off.

### 2.3. Free energy of binding of the non-covalent complexes in water

For the calculation of the binding free energy between  $\alpha$ -CT and the four chiral lactam substrates in water, a total of 400 snapshots were saved during the MD data collection period as described above, one snapshot per each 1 ps of MD simulation. The binding free energy,  $\Delta G_{\text{bind}}$ , of each complex in water was calculated according to the procedure proposed by Srinivasan et al.<sup>85</sup> and, since its application to  $\alpha$ -CT/lactam complexes has already been reported in details in our previous paper,<sup>50</sup>

it will be only briefly described. According to this method,  $\Delta G_{\text{bind}}$  is calculated as:

$$\Delta G_{\text{bind}} = \Delta G_{\text{MM}} + \Delta G_{\text{sol}}^{\text{C}} - \Delta G_{\text{sol}}^{\text{L}} - \Delta G_{\text{sol}}^{\text{P}} - T\Delta S \quad (2)$$

where  $\Delta G_{\text{MM}}$  is the interaction energy between the ligand and the protein,  $\Delta G_{\text{sol}}^{\text{C}}$ ,  $\Delta G_{\text{sol}}^{\text{L}}$  and  $\Delta G_{\text{sol}}^{\text{P}}$  are the solvation free energy for the complex, the ligand and the protein, respectively, and  $-T\Delta S$  is the conformational entropy contribution to the binding. All energetic analyses were done for only a single MD trajectory of the lactam/protein complex considered, with unbound protease and lactam snapshots taken from the snapshots of that trajectory.

$\Delta G_{\text{MM}}$  can be obtained from the molecular mechanics (MM) interaction energies as:

$$\Delta G_{\text{MM}} = \Delta G_{\text{MM}}^{\text{ele}} + \Delta G_{\text{MM}}^{\text{vdW}} \quad (3)$$

where  $\Delta G_{\text{MM}}^{\text{ele}}$  and  $\Delta G_{\text{MM}}^{\text{vdW}}$  are the electrostatic and van der Waals contributions to the interaction energy between the ligand and the receptor. It is worthy of note that the molecular mechanics energy  $\Delta G_{\text{MM}}$  in Eq. (3) effectively consists of a valence part,  $\Delta G_{\text{MM}}^{\text{val}}$ , but, as the structure of the protein in its bound and unbound state is the same, the contribution of this term to the binding free energy is zero. Accordingly, this term has been omitted in Eq. (2). In our case, all other quantities were calculated with the *anal* and *carol* modules from the AMBER 6.0 suite. The infinite cut-offs for all interactions and the *parm94* force field parameters<sup>36</sup> were applied. The total solvation energy,  $\Delta G_{\text{sol}}$ , is divided in two parts: the electrostatic contribution,  $\Delta G_{\text{sol}}^{\text{ele}}$ , and the non-polar term,  $\Delta G_{\text{sol}}^{\text{np}}$ .

$$\Delta G_{\text{sol}} = \Delta G_{\text{sol}}^{\text{ele}} + \Delta G_{\text{sol}}^{\text{np}} \quad (4)$$

The polar component of  $\Delta G_{\text{sol}}$  was evaluated with the PB approach.<sup>86</sup> This procedure involves using a continuum solvent model, which represents the solute as a low dielectric medium (i.e. of dielectric constant  $\epsilon=1$ ) with embedded charges and the solvent as a high dielectric medium ( $\epsilon=80$ ) with no salt. All atomic charges were taken from the Cornell et al. force field,<sup>36</sup> since these are consistent with the MM energy calculations. However, as suggested by Chong et al.,<sup>87</sup> the atomic radii were taken from the PARSE parameter set<sup>88</sup> instead of the *parm94* FF set because of the small size of the hydrogens in the latter. The dielectric boundary is the contact surface between the radii of the solute and the radius (1.4 Å) of a water molecule. The numerical solution of the linearized Poisson–Boltzmann equations were solved on a cubic lattice by using the iterative finite-difference method implemented in the *DelPhi* software package.<sup>89</sup> The grid size used was 0.5 Å. Potentials at the boundaries of the finite-difference lattice were set to the sum of the Debye–Hückel potentials.

The non-polar contribution to the solvation energy,  $\Delta G_{\text{sol}}^{\text{np}}$ , was calculated from the following equation:<sup>88</sup>

$$\Delta G_{\text{sol}}^{\text{np}} = \gamma SA + b \quad (5)$$

in which  $\gamma=0.00542 \text{ kcal}/\text{\AA}^2$ ,  $b=0.92 \text{ kcal/mol}$ , and SA is the solvent-accessible surface estimated with the MSMS<sup>90</sup> program.

To complete the estimate of the free energy of binding, we also calculated the entropy components arising from the solute degrees of freedom. Given that our goal was a qualitative comparison of  $\Delta G_{\text{bind}}$  for the different enantiomeric molecules, we assumed that the entropies were similar in magnitude for the close-structured lactams. This assumption seemed reasonable, given that our previous work on similar compounds<sup>50</sup> and Kuhn and Kollman's calculated values of  $T\Delta S$  for various ligands binding to avidin.<sup>91</sup> Accordingly, the heavily computational-demanding entropy variation calculations were not carried out.

#### 2.4. Definition of the QM model

For this purpose, a model system must be carefully chosen in order to balance the need for including all the atoms of residues that are affecting the electronic structure of the reaction center at an accurate theoretical level and the computational time required for ab initio calculations. The partitioning of the system into a reaction zone (the QM region) and environment (the MM region) is entirely artificial. Indeed, the a priori definition of an 'enzyme environment' has rather an obscure meaning. In the case of  $\alpha$ -CT, the catalytic triad of His 57, Asp 102 and Ser 195, the oxyanion hole residues and the substrate need to be considered for inclusion in the QM model. The decision of which residues or atoms to include inevitably determines the level of ab initio theory that can be applied. Of the residues mentioned above, Ser 195, His 57, and the substrate are certainly the most important, as they are actively involved in the bond breaking and forming processes (see Fig. 3). To allow for the highest possible level of QM calculations, we decided to reduce further our model system by excluding the aliphatic chain substituent of the lactam ring from computations.

To obtain reasonable orientations of the active site residues from QM optimizations,<sup>25</sup> the geometry of MMC-1 and TET-1 were partly optimized keeping the  $C_{\beta}$  and  $C_{\gamma}$  atoms of His 57, the  $C_{\beta}$  and  $O_{\gamma}$  atoms of Ser 195, and the O=C–O atoms of the substrate fixed. The calculations were conducted at the HF/6-31+G\* level. The gas-phase reaction energies ( $\Delta E_{\text{QM}}$ ) were calculated, starting from the HF geometries, at the MP2/AUG-cc-pVDZ level.<sup>92</sup> The atomic point charges were obtained with RESP fitting,<sup>74,75</sup> using Lagrangian restraints to fix the net charge of the quantum mechanical atoms to  $-0.07 \text{ e}$ . This ensures that the sum of the charges of the quantum mechanical atoms plus the charges of the molecular mechanical atoms of His 57, Ser 195, and the substrate atoms is exactly zero.<sup>93</sup>

Concerning the tetrahedral intermediates TET-1, the initial models were generated starting from the refined structures of the corresponding MMCs by forming an appropriate C–O bond between Ser 195 and the substrate, and by transferring the Ser 195 H<sub>γ</sub> proton to the

NE2 of His 57. The geometry of the resulting structures was then relaxed from internal strains by 1000 steps of conjugate gradient algorithm. Finally, ab initio geometry optimizations were conducted by analogy to the procedure described above for MMC-1 although, in this case, only two calculations were needed to be carried out: One for the protonated His 57 and one for the covalent complex between the substrate and Ser 195. In the latter case, only the carbonyl oxygen and the neighboring methyl carbon of the adduct complex were kept fixed, by analogy with the procedure adopted for the MMC-1.

Figure 4 shows the quantum mechanically optimized structures of MMC-1 and TET-1 for the four enantiomers of **1**.

### 2.5. Free energy of interaction by thermodynamic integration

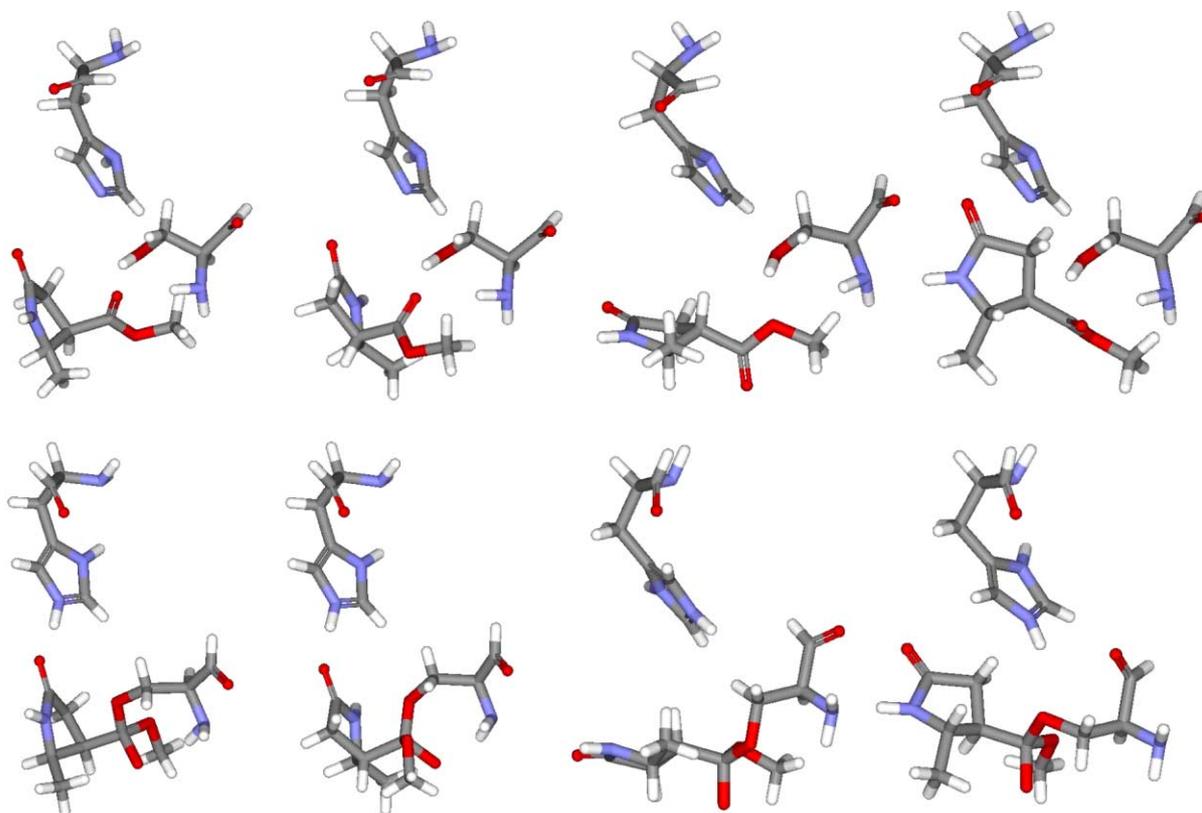
Free energy (FE) simulations<sup>94,95</sup> were employed to calculate, using the thermodynamic integration (TI) method, the relative free energy  $\Delta G_{\text{int}}$  between MMC and TET for each enantiomer of **1** in the enzyme environment. In FE calculations, the QM model system is a function of a perturbation parameter  $\lambda$ . When  $\lambda=1$ , the charge distribution and the bond topology of the model corresponds to MMC. On the contrary, when  $\lambda=0$  the charges and the topology are those of TET. In FE simulations, any free energy variation

within the perturbed group is neglected, as they are included in the gas-phase energy term ( $\Delta E_{\text{QM}}$ ). Accordingly, the second term in Eq. (1),  $\Delta G_{\text{int}}$ , accounts for the environmental effects on the ab initio QM fragments and, when combined with the difference in energies between the QM fragments, the overall free energy  $\Delta G_{\text{tot}}$  can be obtained by means of Eq. (1).

In the TI method, the total free energy change between the two systems A and B can be represented by a summation of the force field terms  $V$ , such as bond lengths and angles, torsions and non-bonded (i.e. van der Waals and electrostatic) interactions:

$$\Delta G = \int_0^1 \langle \partial V / \partial \lambda \rangle_{\lambda} d\lambda \approx \sum_i \langle \Delta V \rangle_{\lambda_i} \Delta \lambda \quad (6)$$

By virtue of its nature, free energy can be also represented as a sum of different parts of the system. Accordingly herein, we decomposed  $\Delta G_{\text{int}}$  into components for each amino acid, solvent molecules, and substrate atoms. The free energy components represent the interactions between the structural components of the environment and the QM active site. At this point, it has to be emphasized that, although  $\Delta G$  is a function of state, its components are not; thus, they depend on the pathway chosen in going from A to B. Therefore, the components cannot be taken as quantitative predictions, but rather as representative of the magnitude of the interactions and, as such, useful tools for getting physical insight into the system considered.



**Figure 4.** Quantum mechanically optimized structures of MMC-1 and TET-1 for the four enantiomers of **1**: (a) MMC-(2*R*,3*R*)-**1a**; (b) MMC-(2*S*,3*S*)-**1a**; (c) MMC-(2*R*,3*S*)-**1b**; (d) MMC-(2*S*,3*R*)-**1b**; (e) TET-(2*R*,3*R*)-**1a**; (f) TET-(2*S*,3*S*)-**1a**; (g) TET-(2*R*,3*S*)-**1b**; (h) TET-(2*S*,3*R*)-**1b**.

In performing FE calculations, we adopted the following conditions: After having reinserted the model fragments into the enzyme, and equilibrated the system for 50 ps, the perturbation simulations were run for 75 ps using a dual cut-off method,<sup>96</sup> with 12 Å primary and 22 Å secondary cut-off. No interactions between QM atoms were included in the  $\Delta G_{\text{int}}$ , as they were already contained in the corresponding  $\Delta E_{\text{QM}}$  energy term [see Eq. (1)] with each run used 101 windows with each window comprised of 200 steps of equilibration and 300 steps of data collection. In order to gather information about errors involved in the calculations from hysteresis, both forward and backward runs were performed.

### 2.6. MD simulations of TET and AEC intermediates of $\alpha$ -CT and **1** enantiomers

The solvated, neutralized and energy-minimized structures of the MMCs of  $\alpha$ -CT and the four possible enantiomers of **1** were selected as good starting points for tetrahedral (TET) and acyl-enzyme intermediates (AEC) MD simulations of  $\alpha$ -CT/**1** enantiomers. The program InsightII<sup>97</sup> was used to modify the active site structure to that of the tetrahedral and acyl-enzyme intermediates, respectively.

After a cycle of energy minimization (1000 steps), a preliminary, MD equilibration phase of 50 ps was carried out, followed by a data acquisition run of 500 ps at 298 K, using a 16 Å cut-off and an integration time step of 1.5 fs. MD frames were collected every 100 steps for further analysis. While in the TET MD simulations, the parameters were the same as in FE calculations, for the ester group of the AEC the torsion parameters and the atomic charges were taken from the work of Peräkylä and Kollman,<sup>93</sup> whereas the bond, angle, improper torsions and non-bonded parameters were taken for the work of Fox et al.<sup>98</sup>

## 3. Results and discussion

### 3.1. Free energy of binding of the noncovalent complexes in water

The experimental results<sup>68</sup> obtained from the enzymatic hydrolysis of the diastereomeric lactams ( $\pm$ )-**1a** and ( $\pm$ )-**1b** indicate that the hydrolysis of the *cis*-diastereomers **1a** by  $\alpha$ -CT proceeded with a complete lack of stereoselectivity, leading to the corresponding racemic lactamic acids. On the other hand, the *trans*-diastereomers **1b** were successfully resolved, allowing the isolation, at 27% conversion, of the lactamic acid of (2*S*,3*R*)-**1b** with 45% ee (23% yield), and the recovery of (2*R*,3*S*)-**1b**, at 80% conversion, with 99% ee (18% yield).<sup>68</sup>

Not surprisingly, the application of molecular modelling techniques to all the stereoisomers of **1** provides a challenge in terms of the specificity of the protocol used and the ability of the energy evaluation to distinguish between correct and incorrect orientations of the bound

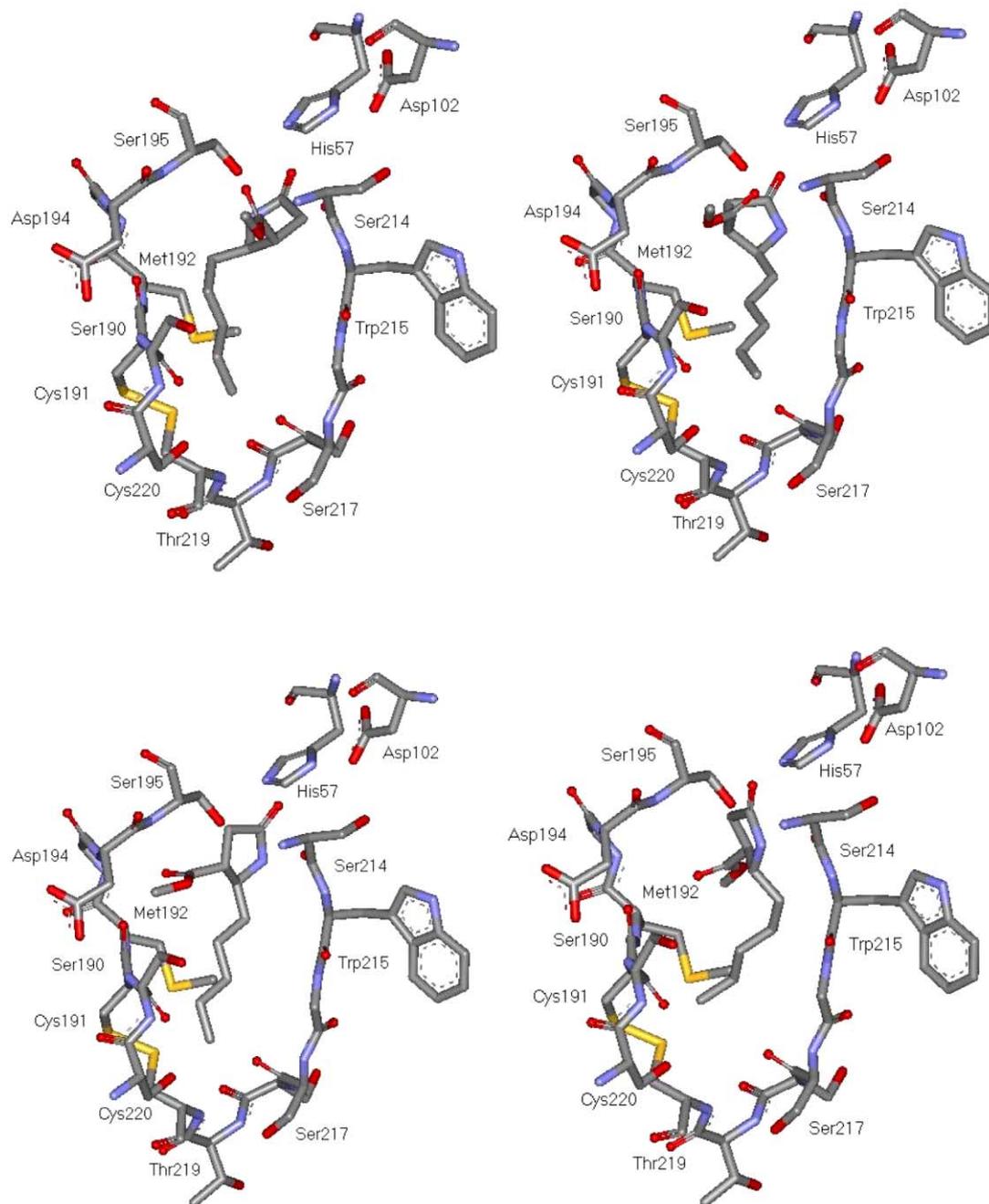
substrate. In the generation of the initial enzyme-substrate model, our previous results in fine-tuning AutoDock procedures for the binding of ligands to proteic receptors<sup>50,99,100</sup> indicate that the cluster of similar conformations with the lowest energy docked structure always reproduced very closely the crystallographic binding mode.

While interaction with a receptor will certainly perturb the conformational energy of a flexible ligand, high affinity suggests that the ligand would not be highly distorted upon binding. In this particular case, the docking study shows that the four stereoisomers (+)- and (–)-**1a**, and (+)- and (–)-**1b** bind effectively in the selected protein region, with no noticeable energy penalty. In fact, the differences in conformational energy between the bound and unbound state of lactams (+)- and (–)-**1a**, and (+)- and (–)-**1b** were confined within 2.5 kcal/mol (1.80 kcal/mol for (2*R*,3*R*)-**1a**, 1.88 kcal/mol for (2*S*,3*S*)-**1a**, 2.20 kcal/mol for (2*R*,3*S*)-**1b** and 2.00 for (2*S*,3*R*)-**1b**, respectively). Following the analysis of the representative clusters (data not shown), the structure believed to represent a significant binding mode in the  $\alpha$ -CT binding site was selected for further optimization, according to the QM-MD procedure as described in the previous section.

Figures 5(a–d) show the resulting molecular models of the docked  $\alpha$ -CT/(+)- and (–)-**1a** and  $\alpha$ -CT/(+)- and (–)-**1b** complexes. In both optimized models for (+)- and (–)-**1a**, the aliphatic chain substituent at C2 of the lactam ring fills the hydrophobic binding site *hc* from Ser 195 downwards a distance of roughly 8.6 Å in the case of the (2*R*,3*R*)-**1a**, and 8.3 Å in the case of (2*S*,3*S*)-**1a**. Furthermore, the hydrolyzing COOCH<sub>3</sub> group assumes, again for both enantiomers (+)- and (–)-**1a**, an orientation in space which is favorable for the interaction with the catalytic triad His 57, Asp 102 and Ser 195. On the contrary, while the (2*R*,3*S*)-**1b** enantiomer extends in the *hc* pocket for approximately 9.2 Å downwards from Ser 195, and orientates the methoxycarbonyl group in a favorable position for hydrolysis by the catalytic triad, the alternative enantiomer (2*S*,3*R*)-**1b** cannot achieve a similar spatial arrangement into the enzyme active site, because of the different orientation of the *n*-alkyl substitute chain. Any attempt to eliminate this unfavorable situation in (2*S*,3*R*)-**1b** by rotation of the aliphatic chain caused it to penetrate the enzyme backbone, and was thus discarded.

The analysis of the trajectories of the MD simulations for both  $\alpha$ -CT/**1a** complexes indicates that for (2*R*,3*R*)-**1a** there is a constant presence of a H-bond which involves the carbonyl oxygen atom of the ester group of the lactam and the OH group of Ser 195. The average dynamic length (ADL) of this H-bond is 2.20 Å. Furthermore, in this situation the catalytic couple His 57–Asp 102 presents a stable H-bond, of ADL equal to 2.57 Å.

For the alternative complex  $\alpha$ -CT/(2*S*,3*S*)-**1a**, whilst the same H-bonds between the C=O group of the lactam



**Figure 5.** Optimized molecular model of the docked  $\alpha$ -CT/**1a** and  $\alpha$ -CT/**1b** complexes: (a, top left)  $\alpha$ -CT/(2*R*,3*R*)-**1a**; (b, top right)  $\alpha$ -CT/(2*S*,3*S*)-**1a**; (c, bottom left)  $\alpha$ -CT/(2*R*,3*S*)-**1b**; (d, bottom right)  $\alpha$ -CT/(2*S*,3*R*)-**1b**. For the sake of simplicity, only the amino acids forming the catalytic triad (His 57, Asp 102, and Ser 195), as well as those pertaining to the binding site are shown. Hydrogen atoms are also omitted for clarity.

ester and the hydroxyl group of Ser 195, and between His 57 and Asp 102 are still present (ADL = 2.27 Å and 2.59 Å, respectively) another, alternative intermolecular interaction between the oxygen atom of the OCH<sub>3</sub> group of the lactam and the hydroxy group of Ser 195 may form intermittently throughout the MD simulation, but it does not persist during the entire trajectory.

The aliphatic, linear chain substituent at C2 of the lactam ring is positioned within the aryl binding site *hc* where, in both cases, it favorably interacts with the

hydrophobic side chain of Met 192. For (2*R*,3*R*)-**1a**, during the simulation, the two chains contact each other at a C–C average distance of about 4.1 Å, thereby increasing mutually their hydrophobic contact area and hence stabilizing the complex. A similar situation is encountered for the other enantiomer, where the distance between the two chains is absolutely comparable (4.0 Å).

Considering the trajectories of both  $\alpha$ -CT/**1b** complexes unveils a different situation for the two *trans*-stereois-

mers. In the case of (2*S*,3*R*)-**1b**, two stable H-bonds are detected: One that links the hydroxyl group of Ser 195 and the oxygen atom of the OCH<sub>3</sub> moiety of the substrate (ADL=2.38), and the other that connects the NH group of His 57 to the carbonyl oxygen of the lactam ring (ADL=2.83 Å). The C=O group of the ester also involved in other intermolecular interaction of this type with the NH of the peptide bond between Gly 193 and Met 192 (ADL=2.41 Å). For the alternative enantiomer (2*R*,3*S*)-**1b** however, the simulation shows only the presence of a H···O=C interaction between Ser 195 and the hydrolyzing ester group of the lactam, and another H-bond between the NH group of the heterocyclic substrate and the NH group of the peptidic bond between Val 213 and Ser 214 (ADL=2.28 Å). Finally, in both α-CT/**1b** complexes, the catalytic triad exhibits an N–H···O interaction between His 57 and Asp 102, of ADL equal to 2.60 Å and 2.62 Å, respectively.

Although the automated docking procedure was able to place both stereoisomers (+)- and (–)-**1b** in the protease active site without unreasonable conformational strain, the deeper penetration and the more favorable orientation of the *n*-alkyl chain of (2*S*,3*R*)-**1b** into the hydrophobic pocket *hc* of α-CT helps this substrate in maximizing interactions with the side chain of Met 192 with respect to its alternative enantiomer (2*R*,3*S*)-**1b**. Indeed, while in the former case the two chains contact each other at a C–C average distance of about 3.8 Å during the MD simulation, in the latter case the reasonably larger distance (i.e. 6.5 Å) makes these interactions weaker.

All the energy terms and estimated free energy of binding for the two set of complexes of lactams (+)- and (–)-**1a** and (+)- and (–)-**1b** and α-CT, obtained from the molecular dynamics simulations performed on the Michaelis–Menten optimized complexes according to the procedure described in details in the previous sections, are reported in Table 1. It is worth reminding at this point,  $T\Delta S$  that the term for substrate binding is not included in Table 1. Given that our goal is a qualitative comparison of  $\Delta G_{\text{bind}}$  for different stereoisomers, as in our previous work<sup>50</sup> we assume that the entropies are similar in magnitude for the close-structured lactams.

**Table 1.** Energy terms and binding free energies<sup>a</sup> (kcal/mol) of α-CT/lactam enantiomers (+)- and (–)-**1a** and (+)- and (–)-**1b**

	(2 <i>R</i> ,3 <i>R</i> )- <b>1a</b>	(2 <i>S</i> ,3 <i>S</i> )- <b>1a</b>	(2 <i>R</i> ,3 <i>S</i> )- <b>1b</b>	(2 <i>S</i> ,3 <i>R</i> )- <b>1b</b>
$\Delta G_{\text{int}}^{\text{vdW}}$	–10.1±0.1	–9.5±0.1	–10.0±0.1	–11.2±0.1
$\Delta G_{\text{int}}^{\text{ele}}$	–71.5±0.2	–70.7±0.2	–71.5±0.1	–72.2±0.1
$\Delta G_{\text{MM}}$	–81.6	–80.2	–81.5	–83.4
$\Delta G_{\text{sol}}^{\text{np}}$	–2.6±0.0	–2.4±0.0	–2.4±0.0	–2.5±0.0
$\Delta G_{\text{sol}}^{\text{ele}}$	79.5±0.2	78.2±0.2	81.1±0.1	80.5±0.1
$\Delta G_{\text{sol}}$	76.9±0.2	75.8±0.2	78.7±0.1	78.0±0.1
$\Delta G_{\text{tot}}^{\text{ele}}$	5.4	5.1	7.2	5.8
$\Delta G_{\text{bind}}$	–4.7	–4.4	–2.8	–5.4

<sup>a</sup>  $T\Delta S$  not included (see text).

As we can see from Table 1, the calculated free energies of the α-CT/(2*R*,3*R*)-**1a** ( $\Delta G_{\text{bind}} = -4.7$  kcal/mol) and of the α-CT/(2*S*,3*S*)-**1a** complex formation ( $\Delta G_{\text{bind}} = -4.4$  kcal/mol) are almost equal, in harmony with the experimentally verified lack of enantioselectivity of this protease with respect to the *cis*-configured substrates.<sup>67</sup> Accordingly, the components of the free energy of binding for one complex are essentially unchanged from those of the corresponding alternative complex. This result is again not surprising, since the relative position and the number and type of intermolecular interactions developed by each lactam enantiomer within the protein active site are equivalent. On the other hand, (2*S*,3*R*)-**1b** binds more tightly to α-CT than its enantiomer (2*R*,3*S*)-**1b**. This molecule has not only the most favorable van der Waals interaction with the protein ( $\Delta\Delta G^{\text{vdW}}[(2*S*,3*R*)-(2*R*,3*S*)] = -1.2$  kcal/mol), but also the most favorable total electrostatic contribution ( $\Delta\Delta G_{\text{tot}}^{\text{ele}}[(2*S*,3*R*)-(2*R*,3*S*)] = -1.4$  kcal/mol).

It is very important to consider the electrostatic component of the molecular mechanics energy  $\Delta G_{\text{int}}^{\text{ele}}$  together with the electrostatic contribution to solvation  $\Delta G_{\text{sol}}^{\text{ele}}$  when examining the role of electrostatics in any protein/ligand complex formation. In fact, as proven by several studies,<sup>101–107</sup> electrostatics generally disfavor the docking of a ligand/receptor couple because the unfavorable change in the electrostatic of solvation is mostly, but not fully, compensated by the favorable electrostatics within the resulting host–guest complex. Indeed, the total electrostatic energy contributions,  $\Delta G_{\text{tot}}^{\text{ele}}$ , to the binding free energy for both α-CT/(+)- and (–)-**1a** and α-CT/(+)- and (–)-**1b** complexes are not favorable, with values ranging from 5.1 to 7.2 kcal/mol (see Table 1).

In particular, the formation of the complex between α-CT and the (2*S*,3*R*)-**1b** lactam is less unfavorable than the corresponding, opposite enantiomer complex because of a less positive, total electrostatic term, in which the penalty paid by the electrostatic of solvation is better compensated by favorable electrostatic interaction within the complex. This calculation suggests that a crucial factor for the enantioselective hydrolysis of these γ-lactams by α-CT is to achieve an optimal electrostatic interaction between the ligand and the protein active site but also to suffer less desolvation penalty. Thus, even though electrostatics tend to overall destabilize all complex formation, it is the optimized balance of opposing electrostatic contributions and a more favorable dispersion term that leads to a tighter binding of the *trans*-(2*S*,3*R*)-**1b** enantiomer to α-CT.

### 3.2. Gas-phase energies, reaction activation and interaction free energies

As explained above, the reaction activation energies  $\Delta G_{\text{tot}}$  for hydrolysis of the ester bond (MMC→TET) were estimated from the gas-phase energies of the corresponding QM models ( $\Delta E_{\text{QM}}$ ) and the free energies of interaction ( $\Delta G_{\text{int}}$ ) between the quantum-mechanical models and the environment obtained from free energy perturbation calculations (see Eq. (1)). The calculated

gas-phase energies ( $\Delta E_{\text{QM}}$ ) of the QM models involving the four different enantiomers of the lactams **1a** and **1b** are reported in Table 2.

**Table 2.** Calculated gas-phase energies  $\Delta E_{\text{QM}}$  (kcal/mol) for MMC-1→TET-1 model systems<sup>a</sup> obtained from ab initio calculations

MMC→TET	$\Delta E_{\text{QM}}$
(2 <i>R</i> ,3 <i>R</i> )- <b>1a</b>	55.3
(2 <i>S</i> ,3 <i>S</i> )- <b>1a</b>	55.0
(2 <i>R</i> ,3 <i>S</i> )- <b>1b</b>	55.2
(2 <i>S</i> ,3 <i>R</i> )- <b>1b</b>	54.8

<sup>a</sup> MMC=Michaelis–Menten complex; TET=tetrahedral intermediate.

As we can see from this Table, the values for  $\Delta E_{\text{QM}}$  are close to each other for all four QM models. Accordingly, the values of the interaction energies  $\Delta G_{\text{int}}$  are to be responsible for the different  $\Delta G_{\text{tot}}$  values and, hence, for the different enantioselectivity of  $\alpha$ -CT towards these compounds. The corresponding, calculated values of  $\Delta G_{\text{int}}$  and  $\Delta G_{\text{tot}}$  [ $=\Delta E_{\text{QM}}+\Delta G_{\text{int}}$ , Eq. (1)] are reported in Table 3, from which we readily see that the formation of the acyl–enzyme intermediate has approximately the same free activation energy in the case of the two *cis*-enantiomers **1a** ( $\Delta\Delta G_{\text{tot}}=\Delta G_{\text{tot}}(\alpha\text{-CT}/(2*R*,3*R*)-\mathbf{1a})-\Delta G_{\text{tot}}(\alpha\text{-CT}/(2*S*,3*S*)-\mathbf{1a})=-0.3$  kcal/mol) whereas, for the alternative couple of lactams **1b**, the global free energy change is significantly lower in the  $\alpha\text{-CT}/(2*S*,3*R*)-\mathbf{1b}$  complex than in the  $\alpha\text{-CT}/(2*R*,3*S*)-\mathbf{1b}$  assembly ( $\Delta\Delta G_{\text{tot}}[\alpha\text{-CT}/(2*R*,3*S*)-\mathbf{1b}-\alpha\text{-CT}/(2*S*,3*R*)-\mathbf{1b}]=+3.9$  kcal/mol).

**Table 3.** Interaction free energies  $\Delta G_{\text{int}}$  (kcal/mol) and total free energy difference  $\Delta G_{\text{tot}}$  (kcal/mol) for MMC-1→TET-1 model systems<sup>a</sup>

	$\Delta G_{\text{int}}^{\text{b}}$			$\Delta G_{\text{tot}}^{\text{c}}$
	Forward	Reverse	Average	
(2 <i>R</i> ,3 <i>R</i> )- <b>1a</b>	-33.9	-32.8	-33.4	21.9
(2 <i>S</i> ,3 <i>S</i> )- <b>1a</b>	-33.3	-32.2	-32.8	22.2
(2 <i>R</i> ,3 <i>S</i> )- <b>1b</b>	-34.2	-33.4	-33.8	21.4
(2 <i>S</i> ,3 <i>R</i> )- <b>1b</b>	-37.0	-37.6	-37.3	17.5

<sup>a</sup> MMC=Michaelis–Menten complex; TET=tetrahedral intermediate.

<sup>b</sup> The values from forward and reverse runs, and the corresponding averages are reported.

<sup>c</sup> Calculated from Eq. (1), using average values of  $\Delta G_{\text{int}}$ , and the  $\Delta E_{\text{QM}}$  values reported in Table 2.

Resorting to the thermodynamic integration (TI) method allowed us to decompose the total values of the interaction energies  $\Delta G_{\text{int}}$  into components involving different parts of the system. As anticipated in the Methods section, in this work we considered the contribution to  $\Delta G_{\text{int}}$  from the enzyme, substrate (i.e. all atoms of the lactam enantiomers not included in the QM model), lactam, link atoms and all water molecules. All these values are reported in Table 4.

**Table 4.** Components of the interaction free energy  $\Delta G_{\text{int}}$  (kcal/mol) for MMC-1→TET-1 model systems<sup>a</sup>

	(2 <i>R</i> ,3 <i>R</i> )- <b>1a</b>	(2 <i>S</i> ,3 <i>S</i> )- <b>1a</b>	(2 <i>R</i> ,3 <i>S</i> )- <b>1b</b>	(2 <i>S</i> ,3 <i>R</i> )- <b>1b</b>
Enzyme	-52.3	-52.8	-53.0	-55.1
Substrate <sup>b</sup>	8.3	8.4	8.5	6.6
Link atoms <sup>c</sup>	5.1	5.3	5.3	5.2
Water	-3.9	-3.3	-4.0	-3.7
Lactam	9.4	9.6	9.4	9.7

<sup>a</sup> MMC=Michaelis–Menten complex; TET=tetrahedral intermediate.

<sup>b</sup> Atoms of the lactam substrate not included in the QM model, excluding link atoms.

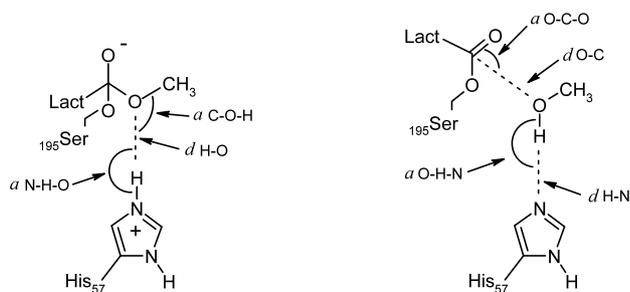
<sup>c</sup> Atoms connecting the reacting center atoms of the lactam and the protein to the remaining atoms, not included in the QM model.

The peptide residues of the enzyme exert a stabilizing contribution to  $\Delta G_{\text{int}}$  (MMC→TET) in excess of 2.1 kcal/mol in the case of the  $\alpha\text{-CT}/(2*S*,3*R*)-\mathbf{1b}$  system with respect to its alternative enantiomer, whereas, for the *cis*-couple, the difference in contribution due to the protein aminoacids is confined to 0.5 kcal/mol. The contributions brought about by the substrate and link atoms are all unfavorable (see Table 4). In particular, the substrate component is 1.9 kcal/mol less favorable for  $\alpha\text{-CT}/(2*R*,3*S*)-\mathbf{1b}$  than for the corresponding, alternative  $\alpha\text{-CT}/(2*S*,3*R*)-\mathbf{1b}$ .

Interestingly, water lowers the activation free energy of ester bond cleavage in all cases. This can be ascribed to the fact that water molecules have almost the same access to the active site, independent of the relative position of the aliphatic chain substituent of the lactam ring within the  $\alpha$ -chymotrypsin hydrophobic binding site *hc*. Further to this, the atoms of the lactam belonging to the QM model have an unfavorable, average contribution of 9.5 kcal/mol to the MMC→TET reaction for all four diastereomers. Summarizing the information reported in Table 4, whilst the  $\Delta G_{\text{int}}$  components of  $\alpha\text{-CT}$ , substrate, link atoms, water and lactam are equivalent in the case of the *cis*-stereoisomers **1a**, for the alternative pair of compounds **1b** they all contribute to increase the total free energy change  $\Delta G_{\text{tot}}$  for the MMC→TET reaction in the case of the  $\alpha\text{-CT}/(2*R*,3*S*)-\mathbf{1b}$  complex by 3.9 kcal/mol with respect to the corresponding, alternative  $\alpha\text{-CT}/(2*S*,3*R*)-\mathbf{1b}$  assembly (see Table 3).

### 3.3. MD average structures of TET and AEC intermediates

As mentioned above, molecular dynamics simulations of the tetrahedral intermediates of  $\alpha$ -chymotrypsin with both couples of enantiomeric lactams (TET-**1a** and TET-**1b**) were performed to determine how the reacting groups involved (i.e. the N–H bond of the protonated His 57 and hydrolyzing O–CH<sub>3</sub> bond) are oriented (see Fig. 6a). As illustrated in Figure 3, in ester hydrolysis reactions catalyzed by  $\alpha$ -chymotrypsin, the tetrahedral intermediate forms an acyl–enzyme intermediate by



**Figure 6.** Definition of the geometrical parameters for the (a, left)  $\alpha$ -CT/**1** tetrahedral intermediates and (b, right)  $\alpha$ -CT/**1** acyl-enzyme intermediates.

proton transfer from the protonated His 57 to the oxygen of the scissile ester bond. We can sensibly assume that such H-transfer is rather fast if the involved H atom is close to the oxygen of the O-CH<sub>3</sub> hydrolyzing group ( $d$  H-O distance, Fig. 6a) and also if the N-H bond of protonated His 57 is directed towards the same O atom; that is the (His57)N-H-O(CH<sub>3</sub>) angle is proximate to the optimum value of 180°. Moreover, the direction along which the transferring proton approaches the hydrolyzing OCH<sub>3</sub> group should be suitable for reaction; in other words, the angle C-O-H (see Fig. 6a) should be close to 120°.

Table 5 reports the results obtained from the geometrical analysis conducted over the relevant MD trajectories. An inspection of this Table reveals that, for both *cis*-lactams (+)- and (-)-**1a**, all parameters assume similar values compatible with ester bond cleavage. In the case of the *trans*-lactams (+)- and (-)-**1b**, the average

**Table 5.** Values<sup>a</sup> of the geometrical parameters<sup>b</sup> for active-site groups obtained from MD simulations of the four tetrahedral intermediates of TET( $\alpha$ -CT/**1a**) and TET( $\alpha$ -CT/**1b**) complexes

	Maximum	Minimum	Average	ASD <sup>c</sup>
<b><math>\alpha</math>-CT/(2<i>R</i>,3<i>R</i>)-<b>1a</b></b>				
$d$ H-O	4.20	2.00	2.83	0.35
$a$ N-H-O	179.3	107.2	148.3	9.4
$a$ C-O-H	132.2	85.8	105.9	7.0
<b><math>\alpha</math>-CT/(2<i>S</i>,3<i>S</i>)-<b>1a</b></b>				
$d$ H-O	4.15	2.01	2.80	0.33
$a$ N-H-O	177.9	105.6	147.9	10.3
$a$ C-O-H	135.1	88.2	106.4	6.8
<b><math>\alpha</math>-CT/(2<i>R</i>,3<i>S</i>)-<b>1b</b></b>				
$d$ H-O	4.03	1.95	2.77	0.31
$a$ N-H-O	180.5	104.9	154.6	9.90
$a$ C-O-H	130.3	89.1	106.3	7.2
<b><math>\alpha</math>-CT/(2<i>S</i>,3<i>R</i>)-<b>1b</b></b>				
$d$ H-O	4.00	1.96	2.69	0.30
$a$ N-H-O	180.8	109.3	156.9	9.3
$a$ C-O-H	133.1	83.2	107.9	8.6

<sup>a</sup> Distances ( $d$ ) are expressed in Å and angles ( $a$ ) in degrees.

<sup>b</sup> Refer to Figure 6a and text for the definitions.

<sup>c</sup> ASD=average standard deviation.

distance H-O (2.69 Å), and average N-H-O (156.9°) and C-O-H (107.9°) angles relative to the TET( $\alpha$ -CT/(2*S*,3*R*)-**1b**) complex are globally more favorable for the reaction than for the TET( $\alpha$ -CT/(2*R*,3*S*)-**1b**) assembly.

Finally, the dynamics simulations on the AEC( $\alpha$ -CT/(+)- and (-)-**1a**) and AEC( $\alpha$ -CT/(+)- and (-)-**1b**) acyl-enzyme intermediates were conducted to explore the relative orientation of the newly formed methanol molecule and the carbonyl moiety of the AEC ester group (see Fig. 6b). In this respect, we can once again safely assume that the eventual reformation of the ester bond and the corresponding Michaelis-Menten complex can easily occur if one of the oxygen lone pairs of CH<sub>3</sub>OH is in a suitable distance ( $d$  O-C distance, see Fig. 6b) with the C=O group of the AEC, and if the direction by which it approaches the carbon atom is along perpendicular to the plane generated by the C=O group itself (i.e.  $a$  O-C-O, see Fig. 6b). As further requirements, (i) the leaving hydrogen of the OH group in methanol should be in close proximity to the NE2 of His 57, which will be acting as a base during the reformation reaction (quantified by the  $d$  H-N distance, see Fig. 6b), and (ii) the same O-H bond should be directed towards the NE2 of the catalytic His 57, i.e. the angle O-H-N should approach the ideal value of 180°.

Table 6 lists the values of the geometrical parameters described above obtained from the MD trajectories of

**Table 6.** Values<sup>a</sup> of the geometrical parameters<sup>b</sup> for active-site groups obtained from MD simulations of the four acyl-enzyme intermediates of AEC( $\alpha$ -CT/**1a**) and AEC( $\alpha$ -CT/**1b**) complexes

	Maximum	Minimum	Average	ASD <sup>c</sup>
<b><math>\alpha</math>-CT/(2<i>R</i>,3<i>R</i>)-<b>1a</b></b>				
$d$ O-C	6.81	3.69	4.87	0.43
$d$ H-N	7.64	3.44	5.21	0.75
$a$ O-C-O	110.8	60.8	86.3	7.4
$a$ O-H-N	144.2	37.5	76.1	19.5
<b><math>\alpha</math>-CT/(2<i>S</i>,3<i>S</i>)-<b>1a</b></b>				
$d$ O-C	6.72	3.71	4.81	0.41
$d$ H-N	7.88	3.37	5.30	0.81
$a$ O-C-O	112.0	58.9	88.2	7.1
$a$ O-H-N	145.2	30.6	77.7	21.0
<b><math>\alpha</math>-CT/(2<i>R</i>,3<i>S</i>)-<b>1b</b></b>				
$d$ O-C	3.86	2.69	3.08	0.32
$d$ H-N	4.36	1.99	2.87	0.62
$a$ O-C-O	111.8	55.3	86.8	7.3
$a$ O-H-N	177.3	96.3	159.7	18.2
<b><math>\alpha</math>-CT/(2<i>S</i>,3<i>R</i>)-<b>1b</b></b>				
$d$ O-C	6.44	3.81	4.96	0.40
$d$ H-N	7.23	3.41	5.33	0.58
$a$ O-C-O	111.9	62.3	84.1	7.1
$a$ O-H-N	144.8	36.2	76.9	17.9

<sup>a</sup> Distances ( $d$ ) are expressed in Å and angles ( $a$ ) in degrees.

<sup>b</sup> Refer to Figure 6b and text for the definitions.

<sup>c</sup> ASD=average standard deviation.

the corresponding four AEC( $\alpha$ -CT/**1a**) and AEC( $\alpha$ -CT/**1b**) acyl–enzyme intermediates. In accordance with the results obtained for the TET intermediates, the values of the geometrical parameters for the AEC complexes formed by the *cis*-enantiomers (+)- and (–)-**1a** are similar and, in both cases, the general balance is unfavorable to a reverse reaction (see Table 6). On the contrary, in the case of the *trans*-**1b** couple of lactams, two opposite situations are revealed by an inspection of the geometrical parameters reported in Table 6. Indeed, these values show that, in the case of the AEC( $\alpha$ -CT/(2*S*,3*R*)-**1b**), the carbonyl moiety of the acyl group and the nitrogen NE2 of His 57 are both quite far apart from the newly formed methanol molecule (4.96 Å and 5.33 Å, respectively), and hence in an unsuitable position to favor backward reaction. In the alternative AEC( $\alpha$ -CT/(2*R*,3*S*)-**1b**) assembly, the oxygen atom of CH<sub>3</sub>OH remains, for the vast majority of time, at approximately 3 Å from the carbon atom of the C=O group and within 30° from the ideal value of 90° of the attack angle (see Table 6).

Moreover, both the distance involving His 57 NE2 and the hydrogen atom of the alcoholic moiety (*d* H–N, see Fig. 6b) and the relevant angle formed by the same two atoms and the methanolic oxygen (*a* O–H–N, see Fig. 6b) are suitable to backward reaction, being equal, on average, to 2.87 Å and 160°, respectively (see Table 6). Accordingly, the relative orientation of the acyl intermediate and the leaving methanol molecule are appropriate for reformation of the broken bond.

#### 4. Conclusions

The interaction of two enantiomeric couples of methyl 5-oxo-2-pentylpyrrolidine-3-carboxylates, namely the *cis*-(+)- and (–)-**1a** and *trans*-(+)- and (–)-**1b**, with the  $\alpha$ -chymotrypsin active site has been simulated using a fine-tuned automated docking procedure, subsequently refined by quenched molecular dynamics. By applying a combination of molecular mechanics energy derived from MD simulations in explicit solvent, and solvation free energy derived from a continuum solvation model, we have calculated reasonable absolute free energies of binding for all  $\alpha$ -CT/enantiomer complex formations.

In general, the energetic analyses reveal that the van der Waals interactions and the nonpolar contributions to solvation always provide the basis for the favorable absolute free energy of binding. On the other hand, a delicate balance also exists between the always favorable gas-phase electrostatics term and the unfavorable change in electrostatic contribution to the solvation. Indeed, by counteracting the favorable electrostatic interactions that form between the lactam and the protein binding site, the desolvation of the protein residues plays an important role in determining the effect of the electrostatics, as a whole, on the formation of any  $\alpha$ -CT/lactam enantiomer complexes.

Furthermore, we have presented the results obtained by the application of the quantum mechanical/free energy

perturbation method (QM/FE), coupled with free energy component analysis, aimed at investigating the roles played by different parts of the catalytic site in the experimentally observed enantioselective power of  $\alpha$ -chymotrypsin towards the  $\beta,\gamma$ -disubstituted- $\gamma$ -lactamic esters ( $\pm$ )-**1a** and ( $\pm$ )-**1b**.<sup>68</sup> These computational procedures yield insight into the major factors affecting the enantioselectivity of ester bond hydrolysis by  $\alpha$ -chymotrypsin.

According to the resulting evidence, we speculated that the enantioselectivity in the hydrolysis of  $\alpha$ -CT towards the (2*R*,3*S*)-**1b** and (2*S*,3*R*)-**1b** compounds may arise also from the fact that, in the case of the  $\alpha$ -CT/(2*R*,3*S*)-**1b** assembly, the ester bond, once cleaved, is much more likely to re-ligate to form the intact lactam ester rather than form the acyl–enzyme and have the newly formed methyl alcohol molecule diffuse away. We supported this speculation by means of molecular dynamic simulations of the tetrahedral intermediate of both couples of enantiomeric lactams **1a** and **1b**, and simulations of the cleaved lactam ester attached to  $\alpha$ -chymotrypsin. These simulations showed that, through much of the 500 ps trajectory, the –OH end of methanol is in an excellent position to attack the acyl C=O and regenerate the ester bond in the case of the (2*R*,2*S*)-**1b** derivative. In contrast to this, for the alternative enantiomer (2*S*,3*R*)-**1b**, and for both the enantiomers of *cis*-**1a**, the alcoholic moiety of the corresponding cleaved lactamic ester substrate moved away from the vicinity of the acyl C=O group, favoring hydrolysis of the acyl–enzyme intermediate.

The application of free energy component analysis to the enantioselective hydrolysis of the  $\beta,\gamma$ -disubstituted- $\gamma$ -lactamic esters ( $\pm$ )-**1a** and ( $\pm$ )-**1b** further suggested that, whilst the amino acid residues of the enzyme, substrate and link atoms, which are structural groups common to all  $\alpha$ -CT/lactam complexes, afforded a similar contribution in the case of the *cis*-stereoisomers **1a**, they all increase the free energy of forming the transition state in the  $\alpha$ -CT/(2*R*,3*S*)-**1b** reaction by 3.9 kcal/mol compared to this process in  $\alpha$ -CT/(2*S*,3*R*)-**1b** alternative assembly. Although we must note that free energy components are sensitive to the pathway chosen in the corresponding free energy calculations and, therefore, care should be taken to interpret their values, we consider that, if properly used, free energy analysis is a valuable computational tool, which gives useful physical information on the system studied.

#### Acknowledgements

The authors thank M. M. Garrett Morris and Peter J. Goodford for provision of the Software AutoDock v. 3.0. We are also much indebted to the group at UCSF (especially to Jim Caldwell) for providing us with Amber 6.0. This work was supported financially by the Italian Ministry for University and Research (MIUR, Rome, Italy, PRIN 2001 and FIRB 2003 to S.P.) and by the University of Trieste (Trieste, Italy, Special

Grant for Research to S.P. and M.F.): Both are gratefully acknowledged.

### References

- Fersht, A. R. *Enzyme Structure and Mechanism*; W. H. Freeman: New York, 1985.
- Roberts, S. M.; Turner, N. J.; Willets, A. J.; Turner, M. K. *Introduction to Biocatalysis Using Enzymes and Microorganisms*; Cambridge University Press: Cambridge, 1995.
- Knowles, J. R. *Science* **1987**, *236*, 1252–1258.
- Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.
- Kollman, P. A.; Kuhn, B.; Peräkylä, M. *J. Phys. Chem. B* **2002**, *106*, 1537–1542.
- Hinchliffe, A. *Chemical Modeling: from Atoms to Liquids*; John Wiley & Sons: New York, 1999.
- Young, D. C. *Computational Chemistry*; John Wiley & Sons: New York, 2001.
- Theoretical Biochemistry—Processes and Properties of Biological Systems*; Eriksson, L. A., Ed.; Elsevier Science: Amsterdam, The Netherlands, 2001.
- Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *193*, 227–249.
- Singh, U.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.
- Bash, P. A.; Field, M.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8092–8094.
- Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184–189.
- Field, M.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
- Stanton, R. V.; Hartsough, D. S.; Merz, K. M., Jr. *J. Comput. Chem.* **1995**, *16*, 113–128.
- Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1995**, *99*, 11266–11275.
- Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298–305.
- Monard, G.; Loos, M.; Thery, V.; Baka, K.; Rivail, J. *Int J. Quantum Chem.* **1996**, *58*, 153–159.
- Matsubara, T.; Sieber, S.; Morukuma, K. *Int J. Quantum Chem.* **1996**, *60*, 1101–1109.
- Thompson, M. A. *J. Phys. Chem.* **1996**, *100*, 14492–14507.
- Ke, T.; Tidor, B.; Klibanov, A. M. *Biotechnol. Bioeng.* **1998**, *57*, 741–746.
- Kazlauskas, R. *J. Curr. Opin. Chem. Biol.* **2000**, *4*, 81–88.
- Kollman, P. A.; Kuhn, B.; Peräkylä, M. *J. Phys. Chem. B* **2002**, *106*, 1537–1542 and references cited therein.
- Kollman, P. A.; Kuhn, B.; Donini, O.; Peräkylä, M.; Stanton, R. V.; Bakowies, D. *Acc. Chem. Res.* **2001**, *34*, 72–79.
- Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 3448–3457.
- Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154–162.
- Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 2974–2975.
- Madura, J. D.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 2517–2527.
- Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 7430–7432.
- Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 7535–7542.
- Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 10966–10968.
- Jorgensen, W. L.; Blake, J. F.; Lim, D. C.; Severance, D. L. *J. Chem. Soc., Farad. Trans.* **1994**, *90*, 1727–1732.
- Schreiner, P. R.; Severance, D. L.; Jorgensen, W. L.; van Schleyer, P. R.; Schaefer, H. F. *J. Am. Chem. Soc.* **1995**, *117*, 2663–2664.
- Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.
- 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.
- Cieplak, P.; Cornell, W. D.; Bayly, C. I.; Kollman, P. A. *J. Comp. Chem.* **1995**, *16*, 1357–1377.
- Poppe, L.; Novak, L. *Selective Biocatalysis*; VCH Publishers: New York, 1992.
- Sheldon, R. A. *Chirotechnology: Industrial Synthesis of Optically Active Compounds*; Marcel Dekker: New York, 1993.
- Serocke, J. E.; Neuroth, H. *J. Biol. Chem.* **1950**, *182*, 577–584.
- Faber, K. *Biotransformations in Organic Chemistry*, 4th ed.; Springer Verlag: Berlin, 2000.
- Hein, G. E.; Niemann, C. *J. Am. Chem. Soc.* **1962**, *84*, 4487–4494.
- Blow, D. M. *Acc. Chem. Res.* **1976**, *9*, 145–152.
- Tsukada, H.; Blow, D. M. *J. Mol. Biol.* **1985**, *184*, 703–711.
- Birtktoft, J. J.; Blow, D. M. *J. Mol. Biol.* **1972**, *68*, 187–240.
- Blow, D. M. *Israel J. Chem.* **1974**, *12*, 483–493.
- Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629–3637.
- Bode, W.; Huber, R. *Eur. J. Biochem.* **1992**, *204*, 433–451.
- Conte, L. L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177–2198.
- Felluga, F.; Fermeglia, M.; Ferrone, M.; Pitacco, G.; Pricl, S.; Valentin, E. *Tetrahedron: Asymmetry* **2002**, *13*, 475–489.
- Longstaff, C.; Campbell, A. F.; Fersht, A. R. *Biochemistry* **1990**, *29*, 7339–7347.
- Perona, J. J.; Tsu, C. A.; Craick, C. S.; Fletterick, R. J. *J. Mol. Biol.* **1993**, *230*, 919–933.
- Castro, M. J. M.; Anderson, S. *Biochemistry* **1996**, *35*, 11435–11446.
- Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331–358.
- Katzung, B. G. *Basic and Clinical Pharmacology*, 7th ed.; Appleton & Lange: Stamford, CT, USA, 1998.
- Corey, E. J.; Li, W.-D. *Z. Chem. Pharm. Bull.* **1999**, *47*, 1–10.
- Corey, E. J.; Choi, S. *Tetrahedron Lett.* **1993**, *34*, 6969–6972.
- Uno, H.; Baldwin, J. E.; Russell, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2139–2140.

59. Nagamitsu, T.; Sunazuka, T.; Tanaka, H.; Omura, S.; Sprengeler, P. A.; Smith, A. B., III *J. Am. Chem. Soc.* **1996**, *118*, 3584–3590.
60. Ghelfi, F.; Bellesia, F.; Forti, L.; Ghirardini, G.; Grandi, R.; Libertini, E.; Montemaggi, M. C.; Pagnoni, U. M.; Pinetti, A.; De Buyck, L.; Parsone, A. F. *Tetrahedron* **1999**, *55*, 5839–5852.
61. Dolky, L. J.; Fedorank, N. A.; Esfandiani, S.; Garst, M. E. US Pat. 5,675,019, 1997.
62. Mulzer, J.; Zuhse, R.; Schmiechen, R. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 870–871.
63. Meyers, A. I.; Snyder, L. *J. Org. Chem.* **1993**, *58*, 36–42.
64. Forzato, C.; Nitti, P.; Pitacco, G.; Valentin, E. In *Targets in Heterocyclic Systems, Chemistry and Properties X*; Attanasi, O.; Spinelli, D., Eds.; Società Chimica Italiana: Rome, 1999.
65. Maier, M. S.; Gonzalez Marimon, D. I.; Stortz, C. A.; Adler, M. T. *J. Nat. Prod.* **1999**, *62*, 1565–1567.
66. Forster, A.; Fitremam, J.; Renaud, P. *Tetrahedron Lett.* **1998**, *39*, 7097–7100.
67. Felluga, F.; Pitacco, G.; Prodan, M.; Pricl, S.; Visintin, M.; Valentin, E. *Tetrahedron: Asymmetry* **2001**, *12*, 3241–3249.
68. Felluga, F.; Fermeglia, M.; Ferrone, M.; Pitacco, G.; Pricl, S.; Valentin, E. *Helv. Chim. Acta* **2002**, *85*, 4046–4054.
69. Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 5712–5726.
70. Weiner, S. J.; Singh, U. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1985**, *107*, 2219–2229.
71. Frigerio, F.; Coda, A.; Pugliese, L.; Lionetti, C.; Menegatti, E.; Amiconi, G.; Schnebli, H. P.; Ascenzi, P.; Bolognesi, M. *J. Mol. Biol.* **1992**, *225*, 107–123.
72. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER* 6, 1999, University of California, San Francisco, USA.
73. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G. L.; Kollman, P. A. *Comp. Phys. Commun.* **1995**, *91*, 1–41.
74. Jayaram, B.; Sprous, D.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102*, 9571–9576.
75. Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comp. Chem.* **1999**, *20*, 217–230.
76. Materials Studio Program Package (v. 2.2), Accelrys Inc., San Diego, CA, USA.
77. Discover Program Package, as implemented in Ref. 76.
78. Fermeglia, M.; Pricl, S. *AIChE J.* **1999**, *45*, 2619–2627.
79. Hehre, W. J.; Radom, L.; van Schleyer, P. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.
80. 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.
81. Mehler, E. L.; Solmajer, T. *Protein Eng.* **1991**, *4*, 903–910.
82. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Comput. Phys.* **1983**, *79*, 926–935.
83. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Comput. Phys.* **1984**, *81*, 3684–3690.
84. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
85. Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401–9409.
86. Sharp, K. A.; Honig, B. H. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301–312.
87. Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. *Proc. Natl. Acad. Sci.* **1999**, *96*, 14330–14335.
88. Sitkoff, D.; Sharp, K. A.; Honig, B. H. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
89. Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1987**, *9*, 327–335.
90. Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*, 305–320.
91. Kuhn, B.; Kollman, P. A. *J. Med. Chem.* **2000**, *43*, 3786–3791.
92. Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
93. Peräkylä, M.; Kollman, P. A. *J. Am. Chem. Soc.* **2000**, *122*, 3436–3444.
94. Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
95. Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395–2417.
96. van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992–1023.
97. InsightII Program Package (v. I2000.1), Accelrys Inc., San Diego, CA, USA.
98. Fox, T.; Scanlan, T. S.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 11571–11577.
99. Fermeglia, M.; Ferrone, M.; Pricl, S.; Spada, P. *Food Hydrocoll.* **2002**, *16*, 557–564.
100. Manfredini, S.; Solaroli, N.; Angusti, A.; Nalin, F.; Durini, E.; Vertuani, S.; Pricl, S.; Ferrone, M.; Spadari, S.; Foche, F.; Verri, A.; DeClercq, E.; Balzarini, J. *Antivir. Chem. Chemother.* **2003**, in press.
101. Novotny, J.; Sharp, K. A. *Progr. Biophys. Mol. Biol.* **1992**, *58*, 203–224.
102. Misra, V. K.; Sharp, K. A.; Friedman, R. A.; Honing, B. H. *J. Mol. Biol.* **1994**, *238*, 245–263.
103. Misra, V. K.; Hecth, J. L.; Sharp, K. A.; Friedman, R. A.; Honing, B. H. *J. Mol. Biol.* **1994**, *238*, 264–280.
104. Sharp, K. A. *Biophys. Chem.* **1996**, *61*, 37–49.
105. Shen, J.; Wendoloski, J. *J. Comput. Chem.* **1996**, *17*, 350–357.
106. Novotny, J.; Brucoleri, R. E.; Davis, M.; Sharp, K. A. *J. Mol. Biol.* **1997**, *268*, 401–411.
107. Brucoleri, R. E.; Novotny, J.; Davis, M. E. *J. Comput. Chem.* **1997**, *18*, 268–276.