Journal of Peptide Science
J. Peptide Sci. **9**: 145–155 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.441

JOURNAL OF PEPTIDE SCIENCE

Computational Screening of Branched Cyclic Peptide Motifs as Potential Enzyme Mimetics

ATHANASSIOS STAVRAKOUDIS, SEVASTI MAKROPOULOU, VASSILIOS TSIKARIS,*
MARIA SAKARELLOS-DAITSIOTIS, CONSTANTINOS SAKARELLOS and IOANNIS N. DEMETROPOULOS*

Department of Chemistry, University of Ioannina, 451 10, Ioannina, Greece

Received 22 July 2002
Accepted 18 September 2002

Abstract: In a previous work we described the design, synthesis and catalytic activity of a branched cyclic peptide as a serine protease mimic. To maximize its catalytic activity we present now a systematic search of a large number of homologous peptides for potential enzyme activity as revealed by the topological arrangement of the catalytic triad residues. This process is accomplished by applying a combined molecular mechanics and molecular dynamics conformational search of about 200 molecules. Starting from a previously synthesized compound that showed some hydrolytic activity several analogues were modelled by amino acid substitutions in the main molecular framework using the Insight II molecular modelling environment with some script automation. Also presented is an algorithm that: (a) generates all possible combinations of residue substitutions, (b) scans the conformational space for each molecule via high temperature molecular dynamics, (c) picks the set of molecules the trajectories of which retained, to a considerable degree, the catalytic triad molecular arrangement, (d) subjects the selected molecules to layer solvation and energy minimization and chooses the molecules, the conformations of which could preserve the catalytic triad arrangement. Finally, a modelling with periodic boundary conditions, was performed to further support the reported algorithm. We found that at least one of the analogues could be a potential serine protease mimic, as revealed by the root-mean-square comparison between the catalytic triad in two molecular dynamics trajectories of the peptide and the corresponding residues in the crystal structure of trypsin. The most promising model candidate was synthesized and tested for its catalytic activity.

Keywords: branched cyclic peptide; computational screening; cyclization on solid support; molecular dynamics; serine protease models

INTRODUCTION

Serine proteases are members of a large class of proteolytic enzymes found in many organisms that have an important role in many biochemical processes, such as protein catabolism, digestion, blood pressure regulation, etc. All the necessary residues for catalysis form the active site of the enzyme, whilst the residues responsible for substrate binding and specificity constitute the binding site of the enzyme [1]. This is the so-called two site approximation (or split site model) for the enzyme function [2], which explains many features of enzyme reactions. In fact, almost all

Abbreviations: Ac, acetyl; Alloc, allyloxycarbonyl; BCL, Biosym Command Language; Boc, *tert*-butoxycarbonyl; DCM, dichloromethane; DIEA, diisopropylethylamine; DMF, *N*,*N*-dimethylformamide; ES-MS, electrospray ionization mass spectrometry; Fmoc, 9-fluorenylmethoxycarbonyl; HOBt, 1-hydroxy-1,2,3-benzotriazole; HPLC, high performance liquid chromatography; HTMD, high temperature molecular dynamics; MD, molecular dynamics; PBC, periodic boundary conditions; pNA, *para*-nitroaniline; pNP, *para*-nitrophenoxy; RMS, root mean square; SPSS, solid phase peptide synthesis; Suc, succinyl; TBTU, 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium; TFA, trifluoroacetic acid.

* Correspondence to: Dr Vassilios Tsikaris, Department of Chemistry, University of Ioannina, 45 110 Ioannina, Greece;
e-mail: btsikari@cc.uoi.gr
** Correspondence to: Ioannis N. Demetropoulos, Department of Chemistry, University of Ioannina, 45 110 Ioannina, Greece;
e-mail: idimitr@cc.uoi.gr

efforts to construct chemical models of enzymes are guided by the split site model approximation. Many compounds have been synthesized aimed at mimicking protease function [3–7]. Although the field of enzyme mimics has been criticized, it persists in a continuous flow of attempts to design potent catalysts based on enzyme structure by the correct arrangement of organic reactive groups in small constructs [8].

Some years ago, Corey and Corey [9] reviewed the field of *de novo* design of peptides as biocatalysts and listed some reasons for the failure to reproduce enzyme-like activities. They concluded that without a fixed peptide conformation there is no catalytic activity. Their conclusions agree with our point of view that any *de novo* design must be combined with a conformational search, especially for small peptides, which generally possess a very low level of ordered structure in solution.

The most obvious evidence supporting the above thoughts comes from the work of Marrone and McCammon [10] who performed MD simulation studies for a serine protease model. The simulation showed that the peptide has no conformational stability and that the reactive groups remain apart from each other. The experimental proof [11,12] for this situation, that came almost simultaneously, demonstrates the need for a conformational search after designing a potentially bioactive peptide.

In this study we try to expand our previous work [7] by applying a fast conformational search to 192 peptide candidates as serine protease mimicking compounds, as well as extended MD calculations with the PBC convention (NVT ensemble). The 'symmetry' of a series of branched cyclic peptides (Figure 1) is checked (via D/L replacements of its residues) in order to define optimal conditions for directing the three residues that constitute the active site of serine proteases in close contact, so that the peptide could potentially act as a catalyst. The best model candidate was synthesized and its catalytic activity was found to be substantially higher than that of our previous model [7].

MATERIALS AND METHODS

Computational Procedure

All computations were performed on a SGI computer running the Biosym Insight II/Discover software.

Leader compounds. The initially designed molecule is depicted in Figure 1. We made a series of amino acid replacements as follows. The chirality of the residues into cycle positions i and $i+3$ were substituted simultaneously each time. Therefore, Lys^{1}–Asp^{4}, Xaa^{2}–Lys^{5}, Lys^{3}–Xaa^{6} made three pairs. Xaa^{2} and Xaa^{6} were simultaneously substituted with the same residue. This operation is necessary in order to retain symmetry in the molecule. These changes produce $2^3 = 8$ combinations. The chiralities of Asp, His and Ser of the catalytic triad changed independently and produced $2^3 = 8$ new combinations, so that the total number of combinations was $8 \times 8 = 64$. Three amino acid substitutions were tested in place of Xaa: Pro for its tendency to form turns, Gly for its flexibility, and Ala for its medium-sized, neutral side chain. Finally, the Gly-Ala-Ser sequence was introduced to mimic the oxyanion hole found in serine proteases [1]. Since we have had three residues in place of Xaa, it turns out that the total number of molecules studied in this work is $3 \times 64 = 192$. A home-made code produced a file containing the D/L flags as input to a Biosym Command Language script in order to have all molecules built automatically with the right combination of chiralities. This file is summarized in Table 1 and is used throughout this study in order to assign the individual runs of the 192 molecules. For instance, the gly12 run is the simulation of the peptide in Figure 1 with Gly in place of Xaa and residue chiralities listed in line 12 of this Table. The Ala15 run is the simulation of the peptide in Figure 1 that has Ala in place of Xaa and residue chiralities denoted in line 15 of this Table, etc. The extensions vac, sol and pbc to these names refers to: (a) energy minimization *in vacuo* of the 100 frames derived from HTMD for vac extension; (b) energy minimization after addition of a layer consisting of 150 water molecules to each one of the 100 frames

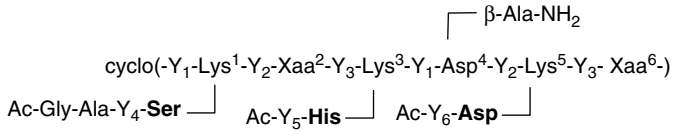

Figure 1 Peptides modelled as serine protease mimics. Active site residues Ser, His and Asp are in bold. Y_i denotes the D/L chirality (see Table 1).

Table 1 Chiralities of the Amino Acids Used to Construct the Set of 64 Peptides

Peptide number	Y_i						Peptide number	Y_i					
	1	2	3	4	5	6		1	2	3	4	5	6
1	L	D	L	L	D	L	*33*	D	D	L	L	D	L
2	L	D	L	L	L	L	*34*	D	D	L	L	L	L
3	L	D	L	L	D	D	*35*	D	D	L	L	D	D
4	L	D	L	L	L	D	*36*	D	D	L	L	L	D
5	L	D	L	D	D	L	*37*	D	D	L	D	D	L
6	L	D	L	D	L	L	*38*	D	D	L	D	L	L
7	L	D	L	D	D	D	*39*	D	D	L	D	D	D
8	L	D	L	D	L	D	*40*	D	D	L	D	L	D
9	L	L	L	L	D	L	*41*	D	L	L	L	D	L
10	L	L	L	L	L	L	*42*	D	L	L	L	L	L
11	L	L	L	L	D	D	*43*	D	L	L	L	D	D
12	L	L	L	L	L	D	*44*	D	L	L	L	L	D
13	L	L	L	D	D	L	*45*	D	L	L	D	D	L
14	L	L	L	D	L	L	*46*	D	L	L	D	L	L
15	L	L	L	D	D	D	*47*	D	L	L	D	D	D
16	L	L	L	D	L	D	*48*	D	L	L	D	L	D
17	L	D	D	L	D	L	*49*	D	D	D	L	D	L
18	L	D	D	L	L	L	*50*	D	D	D	L	L	L
19	L	D	D	L	D	D	*51*	D	D	D	L	D	D
20	L	D	D	L	L	D	*52*	D	D	D	L	L	D
21	L	D	D	D	D	L	*53*	D	D	D	D	D	L
22	L	D	D	D	L	L	*54*	D	D	D	D	L	L
23	L	D	D	D	D	D	*55*	D	D	D	D	D	D
24	L	D	D	D	L	D	*56*	D	D	D	D	L	D
25	L	L	D	L	D	L	*57*	D	L	D	L	D	L
26	L	L	D	L	L	L	*58*	D	L	D	L	L	L
27	L	L	D	L	D	D	*59*	D	L	D	L	D	D
28	L	L	D	L	L	D	*60*	D	L	D	L	L	D
29	L	L	D	D	D	L	*61*	D	L	D	D	D	L
30	L	L	D	D	L	L	*62*	D	L	D	D	L	L
31	L	L	D	D	D	D	*63*	D	L	D	D	D	D
32	L	L	D	D	L	D	*64*	D	L	D	D	L	D

derived from HTMD, properly cooled to 300 K for sol extension; and (c) simulation applying periodic boundary condition with minimum image for pbc extension. Details of each simulation are described in Table 2.

RMS comparison and C^{α} distances of catalytic residues. RMS distances between the heavy atoms of the side chains of the catalytic triad residues in the enzyme and the corresponding heavy atoms of the peptide model were measured in order to assign a similarity score. This comparison refers to 12 atoms (although others restrict this number to six [13]) belonging to the side chain heavy atoms of the three residues in the active site: Ser, His and Asp. We believe that it is necessary to include the C^{β} atoms in such comparisons in order to take into account the directionality of the chemical groups belonging to the side chains. All RMS measurements mentioned throughout this paper refer to the above definition. This means that the His side chain contributes up to half of the total value (6 atoms), while Asp (4 atoms) and Ser (2 atoms) side chains have minor contributions. Indeed, this is consistent with the significant role of imidazole in catalysis; its proper orientation towards the hydroxyl group of Ser and the leaving group of the cleaved bond is crucial. We also measured the distances between the C^{α} atoms of the active site residues throughout our simulation, as an additional criterion of similarity.

Table 2 Details of Computer Simulation Steps

	Model build	HTMD	min-vac	min-sol	pbc1	pbc2	pbc3
Steps of minimization	1500	—	1000	2500	5000	5000	5000
Max. der. (kcal/mol*Å)	0.01	—	0.1	0.01	0.01	0.01	0.01
Cutoff (Å)	—	—	12	10	10	10	10
Dielectric constant	1*r	1*r	1*r	1	1	1	1
Equilibration period in MD (ps)	—	20	—	—	10	10	10
MD duration (ps)	—	200	—	—	100	100	100
Frames collected	1	100	1	1	100	100	100
Box dimensions	—	—	—	—	25^3	25^3	28^3
Radius of layer solvation* (Å)	—	—	—	~5	—	—	—
Water molecules	—	—	—	150	433	440	652
Force field	cvff	cvff	cvff	cff91	cff91	cff91	cff91

Hereby, there is a summary of the geometrical parameters of the positional arrangements of the catalytic triad collected from the crystal structures of serine proteases. X-Ray diffraction structures were purchased from PDB. The template was extracted from the crystal structure of trypsin (entry code 2ptn) as a typical representative of the serine protease family. Table 3 summarizes crystallographic data of a group of serine proteases for the characteristic distances between the C^{α} atoms of the catalytic triad amino acids.

Template forcing. Template forcing was used for guiding each peptide (the moving molecule) to adopt a specific conformation, similar to the trypsin active site, referred to as the template. This is accomplished by adding a penalty function term to the potential energy of the system as defined by the force field. The extra term is described in the following equation:

$$E = V\left[\sum_{pairs}^{N} K_i \frac{\left(R_i^{moving} - R_i^{template}\right)^2}{N}\right]^{1/2}$$

where any atom from the moving molecule with coordinates R_i^{moving} is forced to adopt the coordinates of the atoms of the template molecule $R_i^{template}$, with force constant K_i equal to 100 kcal/Å^2 in our simulations. The template was the side-chain heavy atoms of the catalytic triad residues (Asp, His, Ser). Coordinates were taken from the Protein Data Bank, entry code 2ptn (trypsin).

The similarity was tested of peptide conformations derived from HTMD to the template structure, as well as the minimized conformations. The

Table 3 Distances in Å of C^{α} Atoms of Catalytic Triad Residues in Selected Serine Proteases (X-Ray Structures from PDB)

Enzyme	E.C. number	PDB entry	Ser:C^{α}-His:C^{α}	His:C^{α}-Asp:C^{α}
α-Chymotrypsin	3.4.21.1	5cha	8.48	6.4
Elastase	3.4.21.11	2est	8.1	6.56
α-Lytic protease	3.4.21.12	2lpr	8.36	6.16
Proteinase A	3.4.21.3	3sga	8.3	6.2
HLE	3.4.21.37	1ppg	8.3	6.27
Trypsin	3.4.21.4	2ptn	8.25	6.49
Proteinase B	3.4.21.48	3sgb	8.41	6.3
Thrombin	3.4.21.5	1ets	8.2	6.47
Kallikrein A	3.4.21.8	2pka	8.27	6.45

computed value:

$$Dif_i = |\mathrm{RMS}[\mathrm{frame}_i^{\mathrm{HTMD}}, \mathrm{TEMPLATE}] - \mathrm{RMS}[\mathrm{frame}_i^{\mathrm{vac}}, \mathrm{TEMPLATE}]| \forall i = 1, 1000$$

which is the absolute difference of the RMS values for all 1000 frames produced by the HTMD run of a peptide with Xaa = Pro (Figure 1) and all L- residues with respect to the template of trypsin, minus the RMS values of these frames after energy minimization. 145 frames record Dif values less than 0.1, while 815 frames have Dif_i values less than 1 Å. A comparison of these two trajectories (1 ns length) demonstrates that it is possible to extract the desired information without necessarily performing an energy minimization, which is computationally expensive. Therefore, at the first stage of simulations, we examined the conformational space of peptides only at 900 K.

Flowchart. The course of the peptide selection that attempts to retain conformations similar to the enzyme structure is described in Figure 2. This flowchart of the simulation runs three consecutive times, for each substitution of Xaa.

Initially a home coded program filled up a matrix with 384 (64×6) binary values as the input to a BCL script-driver of the biopolymer module of insight II appropriate to build the peptide, with the appropriate chirality's assignments. Each peptide required six values, three for side residues and three for the in circle residues. The file was updated at the end of each HTMD run, so the next simulation got a new serial input. The peptide was energy minimized and template forcing was applied to achieve the desired geometry consisting of the heavy atoms belonging to the side chains of Ser^{195}, His^{57}, Asp^{102} and the backbone atoms of the Gly^{193}-Ser^{195} sequence. At this stage all ω angles were constrained to 180°. The resulting structure was subjected to HTMD for the conformational search.

The set of 19 200 frames generated from the 192 molecular modifications of the original design has been classified with respect to the RMS of the heavy atoms of each side chain belonging to the catalytic triad *vis-a-vis* the corresponding atoms of the trypsin molecule.

Peptides that recorded low RMS values for at least one of the frames of the HTMD trajectories were treated with molecular mechanics energy minimization *in vacuo* and subsequently with layer solvation in order to obtain more reliable structure prediction than those obtained by HTMD. One of these peptides was systematically investigated by applying three MD runs with explicit full solvent treatment.

Molecular dynamics and mechanics description. HTMD were carried out for a 200 ps time interval (following a 20 ps equilibration period) using a time step-size of 1 fs with application of SHAKE. The temperature was maintained at 900 K by the weak coupling technique. Structures were saved every 2 ps, so that 111 conformations were stored and the last 100 were used for analysis.

Energy minimization *in vacuo* was carried out with 1500 steps (or until the maximum derivative value was less than 0.1 Kcal/mole*Å) of the conjugate gradients method. A linear (1*r) distance dependent dielectric constant was applied to mimic the bulk of the solvent for all *in vacuo* simulations.

For the layer solvation simulations the 100 structures derived from HTMD were cooled for a period of 1 ps by gradually reducing the temperature of the system from 900 K to 300 K and subsequently were subjected to minimization in the presence of explicit solvent molecules. Water solvation was performed with a radius adjusted to produce 150 water molecules around peptide centre of mass. The minimization was carried out in two steps. During the first step (1000 iterations of the conjugate gradient minimization algorithm) only the solvent atoms were allowed to move in order to eliminate any bad contacts and to allow them to maximize their hydrogen bonding to solute. In the second step (1500 conjugate gradient iterations) all atoms were free to move.

Periodic boundary condition simulations. For the gly12 molecule three simulations at 300 K were performed by applying periodic boundary conditions (minimum image) with water as solvent. The starting conformation in the first simulation (pbc1) was that with the lowest energy conformation from the 100 minimized structures in the presence of 150 water molecules. The second simulation (pbc2) started from the conformation with the lowest RMS, and the third one (pbc3) from the same conformation but an additional template force minimization was applied before the addition of solvent. All three simulations were carried out for 100 ps (following a 10 ps equilibration period) with 1 ps intervals. Structures were stored every 1 ps and the last 100 frames were used for analysis.

Peptide Synthesis

Amino acid derivatives were purchased from Bachem. Alloc-L-Lys(Fmoc)-OH was synthesized as

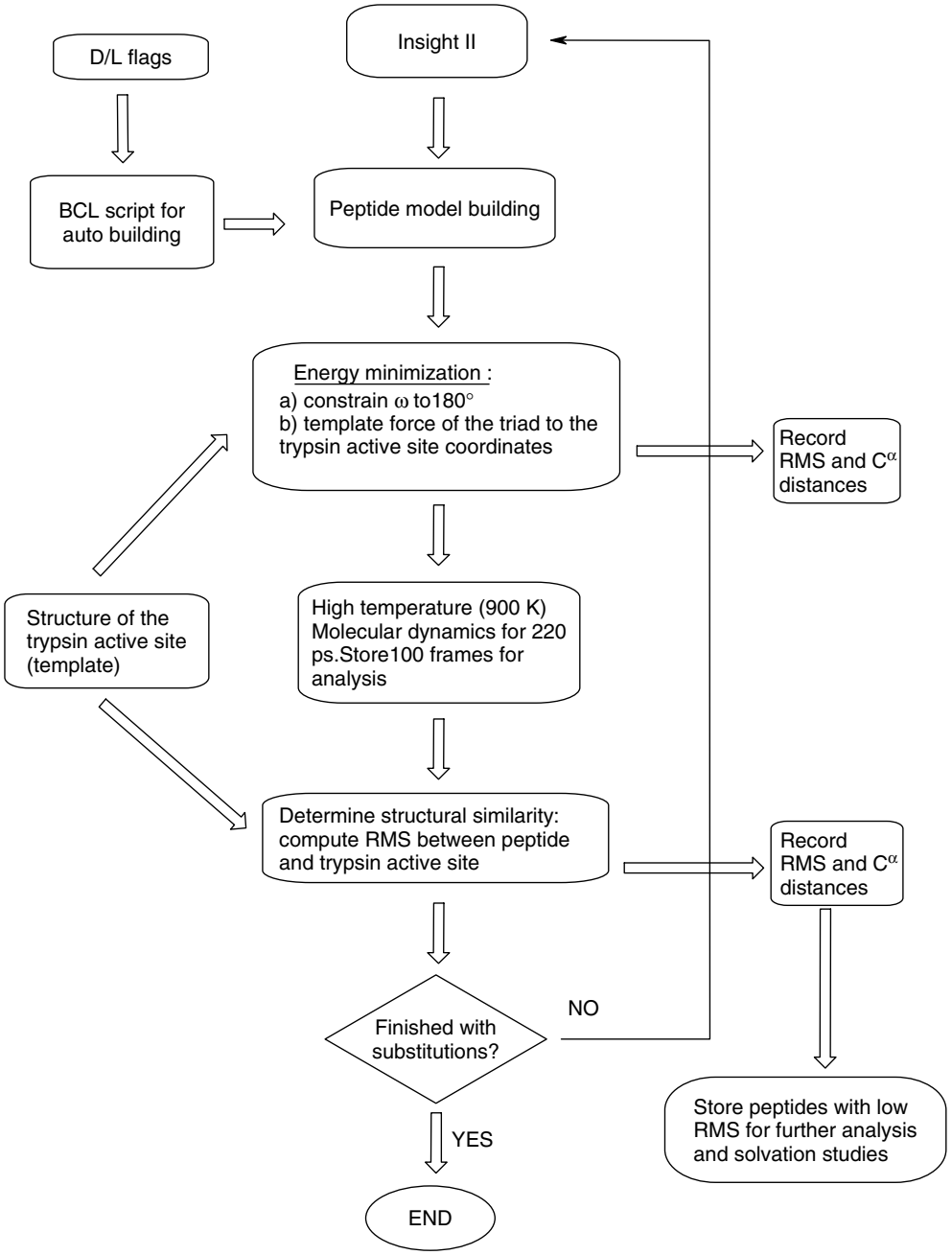

Figure 2 Flowchart of methodology of best peptide selection *via* HTMD simulations (see text for details).

described elsewhere [7]. Solvents were obtained from Merck and Fluka. DMF, pyridine and DIEA were purified by standard protocols. Suc-Ala-Ala-Ala-pNA and Boc-Ala-pNP (purchased from Sigma) were used as substrates.

The solid-phase synthesis of the peptide started by anchoring Fmoc-β-Ala-OH on a pMBHA resin (0.25 mmol/g). Substitution was estimated by measuring the Fmoc-piperidine adduct at 300.6 nm, after Fmoc removal with piperidine in DMF of small weighed aliquots of the peptide resin, followed by acetylation of the unreacted amino groups. A low substitution level of the resin was desired in order to facilitate the last cyclization step, while β-Ala was included to avoid steric hindrance of the β-carboxylate group of the Asp with the polymer benzene rings. Figure 3 describes the step-by-step synthetic procedure of the model peptide. To three

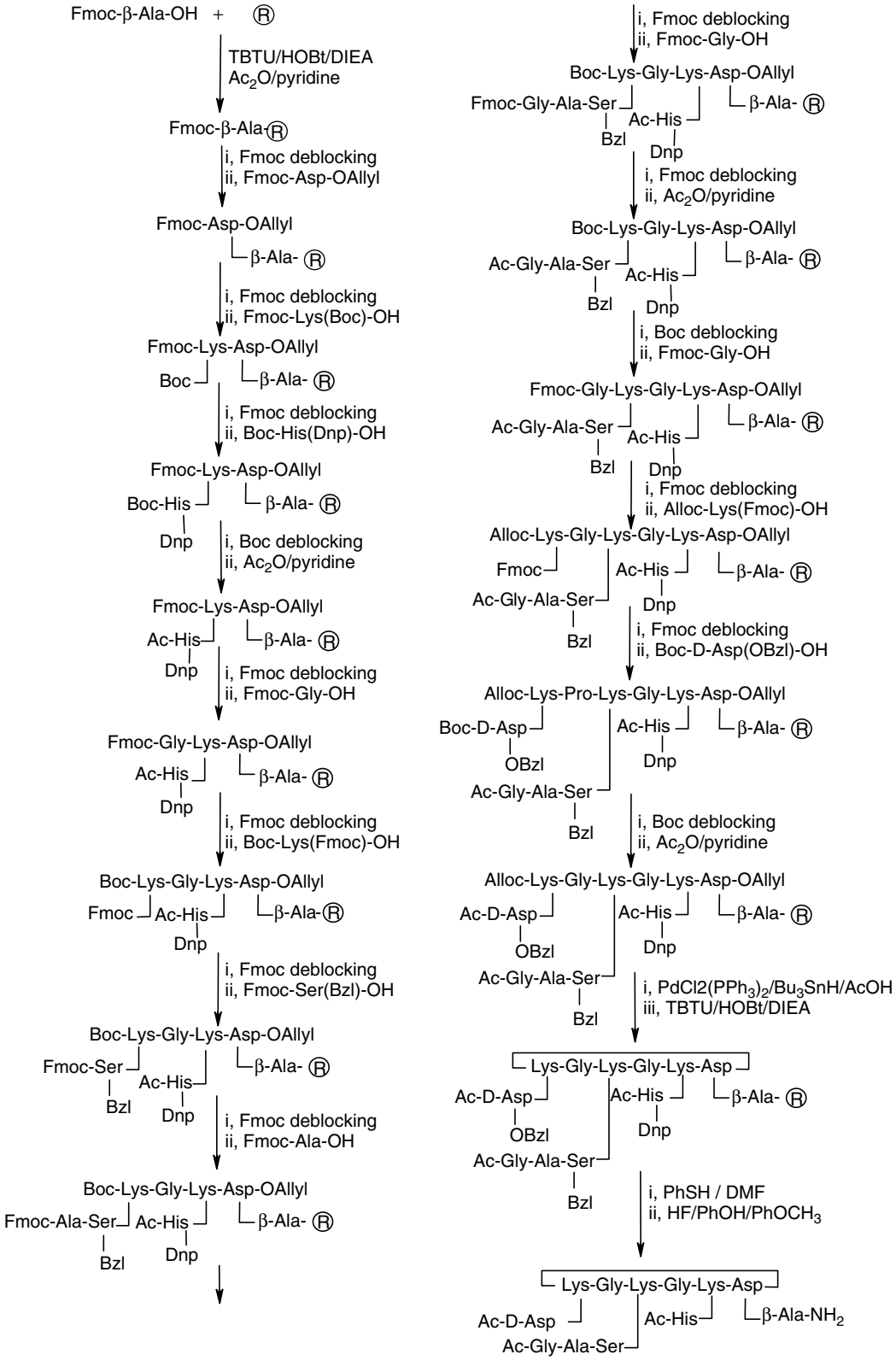

Figure 3 SPSS procedure followed for the synthesis of the selected peptide model. Fmoc, Boc and Alloc/Allyl deprotections were performed with (a) 40% piperidine in DMF, (b) 40% TFA in DCM, and (c) $PdCl_2(PPh_3)_2/Bu_3SnH/AcOH$ in DMF, respectively.

equivalents of the properly protected amino acid and three equivalents of HOBt and TBTU [14] dissolved in 15 ml of a DMF/DCM (1 : 1, v/v) solution six equivalents of DIEA was added. The mixture was stirred for 10 min and added to the resin. Couplings were performed for 30–120 min, and repeated if necessary. Fmoc deprotection was carried out with 40% piperidine in DMF, while Boc deprotection was performed with 40% TFA in DCM. Alloc and allyl groups were removed simultaneously with 0.04/3/3 $PdCl_2(PPh_3)_2/Bu_3SnH/AcOH$ [15] per allyl group in DMF (2x30 min). Cyclization [16] was performed for 3 d in the presence of TBTU/HOBt/DIEA (equimolar quantities to carboxylate). The yields of all steps were checked by the ninhydrin test. Deprotection and cleavage of the peptide from the resin was achieved by anhydrous HF treatment in the presence of phenol and anisole as scavengers. The crude material (yield 77.8%) was subjected to semi-preparative HPLC, using a C_{18} 25 cm × 1 cm, Hypercil reverse-phase column with solvent A, H_2O/0.1% TFA and B, CH_3CN/0.1% TFA. Programmed gradient elution (4.7 ml/min) was applied (95 : 5–70 : 30), with an elution time of 30 min. The yield of purified product was 14%. ES-MS of the purified product gave $[M + H]^+$ 1277.18; $[M + H]^+$ calcd 1277.36 (Figure 4).

Kinetic Studies

The hydrolytic activity of the model peptide was followed on a Shimadzu UV2100 spectrophotometer. A constant concentration of the model peptide of 200 μM was left to hydrolyse various concentrations of the Suc-Ala-Ala-Ala-pNA (release of 4-nitroaniline was followed at $\lambda = 410$ nm) or the Boc-Ala-pNP (release of 4-nitro-phenol was followed at $\lambda = 400$ nm) substrates in 0.1 M Tris and 0.5 M NaCl buffer adjusted to pH 7.5 at 25 °C. Buffer hydrolysis of the substrates was directly subtracted by placing the same substrate concentration in the reference cuvette. Data were analysed with the adapted Shimadzu software to calculate the initial slopes in the first 10–20 min. All experiments were performed in duplicate.

RESULTS AND DISCUSSION

Computer Simulations

A total of 192 peptides were tested for structural similarity to the active site of serine proteases (trypsin was used as a template for comparison) with HTMD. Four of them were selected as potential candidates and their sets of conformations derived from HTMD were subjected to energy minimization, *in vacuo* and in the presence of solvent (water layer

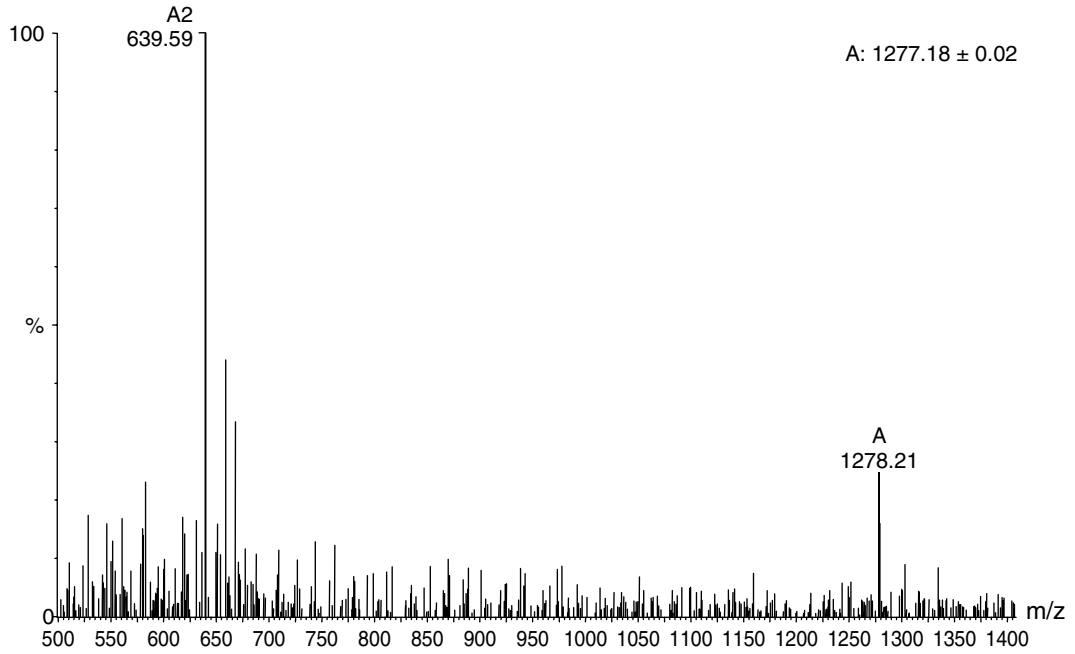

Figure 4 ES-MS of the gly12 synthesized peptide. $[M + H]^+$ 1277.18; $[M + H]^+$ calcd. 1277.36.

solvation). There was a preference for the Gly residue in place of Xaa that could be due to the Gly flexibility.

The simulations of the water-layered molecules are discussed hereby since they provide the main discriminatory operation of this study in favour of the active compounds.

Vacuum and solvent energy minimization. The distance Ser:C^{α}-His:C^{α} recorded for the gly12 molecule shows small fluctuations. This distance is around 8–9 Å, a value very close to those extracted from crystallographic data (Table 3). On the other hand, gly58-sol molecular ensemble frames, treated with energy minimization, produced an equally promising average value with the exception of large fluctuations of individual frames. Hence, it would not be expected to be a competitive good candidate for a serine protease mimetic. It should be stressed that a distance much shorter than that reported from the x-ray diffraction structures is not desirable since it is a sign of different side-chain directionality (reorganization that is affected by to the proximity of the two residues). Trajectories of gly20-sol and ala15-sol runs had a small percentage of frames with values for the distances Ser:C^{α}-His:C^{α} close to those of the x-ray structures (data not shown), while His:C^{α}-Asp:C^{α} distances are closer to those found in the crystal structures. Molecule gly12 shows a very moderate fluctuation of this distance, with values around 5–6.5 Å, very close to those inferred from the x-ray literature. In addition, this molecular design shows a considerable number of frames in close agreement (small RMS) to the trypsin template (see also Materials and Methods, Template Forcing). These findings prompted us to conclude that this molecule is the most appropriate to mimic the serine protease since the critical geometrical elements of the Ser-His-Asp triad are similar and remain dynamically similar to those of the enzymes. No other tested molecule exhibited such a propensity. The stabilization phenomenon of the C^{α} atom triad distances is a much stronger indicator for a promising enzyme mimetic compound than some isolated small RMS values. This statement has a justification that, despite the side-chain flexibility, the molecule is anchored to the favoured conformational region.

The credibility of these arguments has been tested by a further series of simulations with pbc conditions where three different initial conformations of the gly12 peptide yielded three trajectories. The main findings are discussed below.

Periodic boundary condition simulations. Figure 5 depicts the variation of the potential energy and RMS as a function of the consecutive frames of two pbc runs, gly12-pbc2 and gly12-pbc3. These two energy trajectories are similar, while the RMS *versus* frame curves are different. The gly12-pbc2 run records lower RMS values (each frame with respect to the template) (a smaller than 2 Å RMS value is generally regarded as acceptable [13]). The

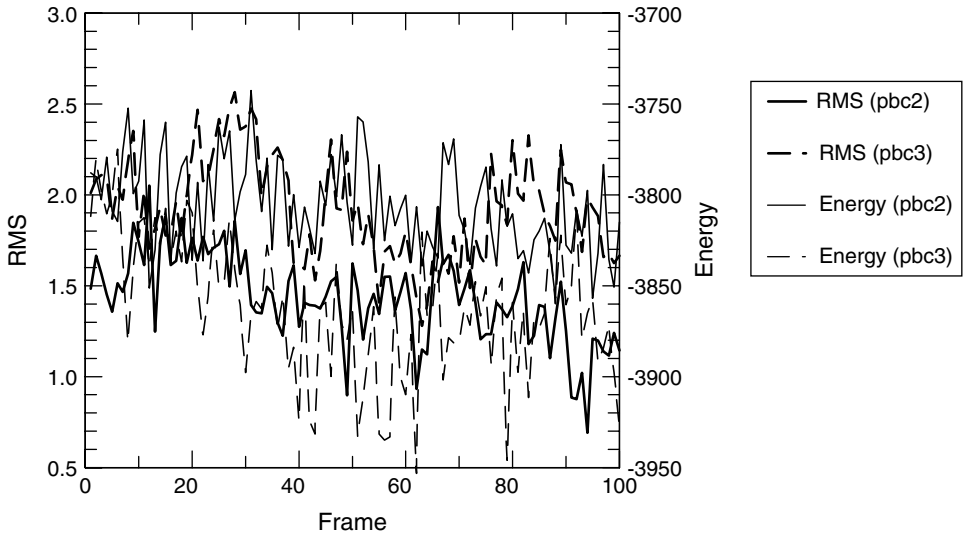

Figure 5 Variation of RMS (Å) and potential energy (Kcal/mol) in pbc2 and pbc3 molecular dynamics runs of the gly12 peptide. A constant value of 2200 Kcal/mol has been added to the energy term of gly12-pbc3 run in order to fit in graph settings.

gly12-pbc1 run has been omitted since the molecule did not converge to the targeted geometry. The MD simulation with periodic boundary conditions was applied to gly12 molecule for 100 ps and generated a trajectory with a considerable amount of frames possessing the right positioning of the three residues as to justify the formation of catalytic triad.

Figure 6 depicts the distances of Ser:C^{α}-His:C^{α} and His:C^{α}-Asp:C^{α} for the gly12-pbc2 and gly12-pbc3 runs. The gly12-pbc2 run yielded many frames with these distances close to those reported by the x-ray diffraction studies (8–9 Å for the former and 5–6.5 Å for the latter). The values recorded from the gly-pbc3 run diverge from the template value, in particular the His:C^{α}-Asp:C^{α} distance value.

We found that the distances between the C^{α} atoms of the catalytic triad residues are within the recorded range for the crystal structures of serine proteases (Table 3), for approximately 20%–30% of the frames. These findings provide additional support (other than the RMS) to the hypothesis that the cyclic part of the peptide has the tendency to bring the branches in proximity, at least for a considerable amount of time. It is clear that the His and Asp residues exhibited a tendency to interact. However, this tendency is not evident for the Ser and His dyad. It is also notable that Asp interacts also with Ser, which is expected from the symmetry of the molecule. In some frames Ser is hydrogen bonded to Asp. These structures cannot be attributed as 'good structures', since this hydrogen bond abolishes the ability of Ser to interact simultaneously with His, thus forming the catalytic triad. It will be of interest to make attempts to eliminate this interaction, although it appears to be less strong when the simulation is carried out in an explicit water environment.

Kinetic Studies

The hydrolytic activity of the enzyme model compound was followed using the Suc-Ala-Ala-Ala-pNA and Boc-Ala-pNP substrates. The model peptide showed substantial activity against Suc-Ala-Ala-Ala-pNA. Using the Boc-Ala-pNP substrate, the kinetic constants of the enzymatic model were estimated: $k_m = 0.492\ mM^{-1}$, $V_{max} = 264\ \mu M\ min^{-1}$ and $k_{cat} = 0.414\ min^{-1}$. This reaction is ~4.6 times faster than that with our previous serine protease model [7]. The achieved hydrolytic activity of the constructed compound indicates that the designed molecule sufficiently mimics the serine protease active site. Taking into account that the binding site of the enzyme was not included in our model compound, as well as the fact that the Ser side chain showed a limited preference for interacting with the His side-chain, we could explain the substantial hydrolysis of the Suc-Ala-Ala-Ala-pNA. This hydrolytic activity further supports the idea that computer simulations can lead to significant improvement in designing more active biomolecules. In this case a 4.6 times improvement in the catalytic rate, compared with

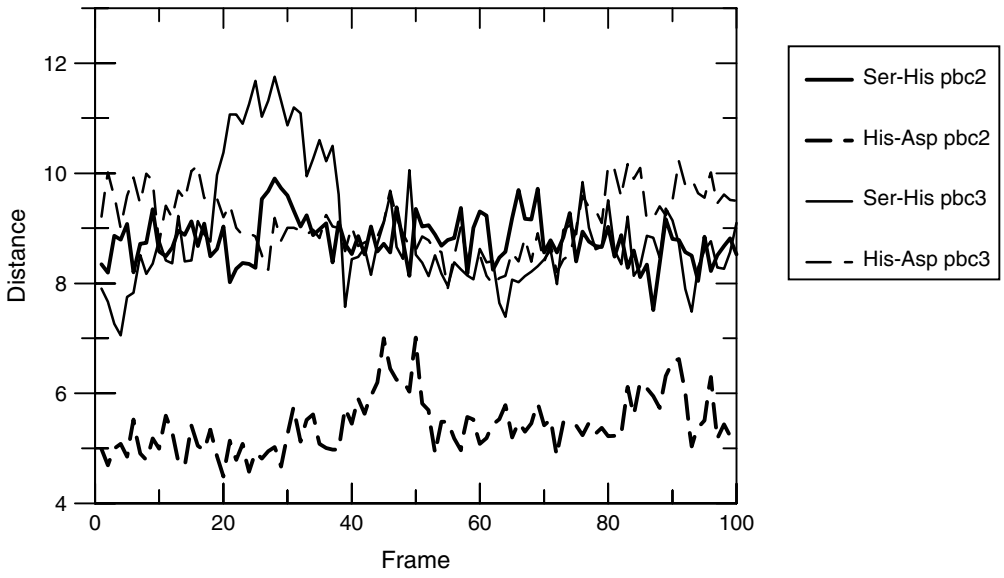

Figure 6 Distance (Å) evaluation between Ser:C^{α}-His:C^{α} and His:C^{α}-Asp:C^{α} in gly12-pbc2 and gly12-pbc3 molecular dynamics runs.

our previous model, was achieved at no laboratory cost.

CONCLUSIONS

A serine protease model peptide has been successfully derived from computer modelling and experimentally tested for enzymatic activity. Chemically or biochemically active compounds can be generated through structure-based design and modelling techniques. The approach that is proposed in this study can be summarized as follows: a molecular motif, a branched-cyclic peptide bearing three Lys residues, was scrutinized for effective enzyme mimetic action using computational methods. The above mentioned process generated approximately 200 molecules. A HTMD quick scan of the conformational space for each molecule coupled with a geometrical similarity measure (RMS) to an enzyme template were the operation tools of choice. The filtered through molecules exhibited a significant percentage of conformations with the property of geometrical similarity to the template catalytic triad. Since the enzyme activity was thought to be a function of distance and time (this hypothesis was termed spatiotemporal [2]), the model compound gly12 satisfies the correct placement of the catalytic triad for a large percentage of the simulation time. Furthermore, the computational operation is flexible and expandable. Hence, other molecular motifs can be investigated in a similar manner.

The successful synthesis of the target peptide, including cyclization, solely by the use of SPSS, demonstrates the feasibility of our proposed method for the preparation of branched-cyclic peptides on solid support. This methodology is expected to be useful for other applications of protein/enzyme mimicking approaches.

The peptide derived from computer modelling had better catalytic properties than our previous model [7]. To our knowledge, this is the first time that an enzyme model has been improved by an extensive conformational search.

REFERENCES

1. Kraut J. How do enzymes work? *Science* 1988; **242**: 533–540.
2. Menger FM. Enzyme reactivity from an organic perspective. *Acc. Chem. Res.* 1993; **26**: 206–212.
3. Atassi MZ, Mansouri T. Design of peptide enzymes (pepzymes): Surface-simulation synthetic peptides that mimic the chymotrypsin and trypsin active sites exhibit the activity and specificity of the respective enzyme. *Proc. Natl Acad. Sci. USA* 1993; **90**: 8282–8286.
4. Hahn KW, Klis WA, Stewart JM. Design and synthesis of a peptide having chymotrypsin-like activity. *Science* 1993; **248**: 1544–1547.
5. Muynck HD, Madder A, Farcy N, DeClercq PJ, Perez-Payan MN, Ohberg LM, Davis AP. Application of combinatorial procedures in the search for serine-protease-like activity with focus on the acyl transfer step. *Angew. Chem. Int. Ed. Engl.* 2000; **39**: 145–148.
6. Breslow R. Biomimetic chemistry and artificial enzymes: Catalysis by design. *Acc. Chem. Res.* 1995; **28**: 146–153.
7. Stavrakoudis A, Demetropoulos IN, Sakarellos C, Sakarellos-Daitsiotis M, Tsikaris V. Design, synthesis and catalytic activity of a serine protease synthetic model. *Lett. Pept. Sci.* 1997; **4**: 481–487.
8. Han MJ, Yoo KS, Chang JY, Ha K-T. 5-(β-Cyclodextrinylamino)-5-deoxy-α-D-riboses as models for nuclease, ligase, phosphatase, and phosphorylase. *Angew. Chem. Int. Ed. Engl.* 2000; **39**: 347–349.
9. Corey MJ, Corey E. On the failure of *de novo*-designed peptides as biocatalysts *Proc. Natl. Acad. Sci. USA* 1996; **93**: 11 428–11 434.
10. Marrone TJ, McCammon JA. Pepzyme dynamics and conformation. A molecular dynamics study in water. *J. Am. Chem. Soc.* 1994; **116**: 6987–6988.
11. Corey DR, Phillips MA. Cyclic peptides as proteases. A reevaluation. *Proc. Natl Acad. Sci. USA* 1994; **91**: 4106–4109.
12. Wells JA, Fairbrother WJ, Otlewski J, Laskowski M Jr, Burnier JA. A reinvestigation of a synthetic peptide (TrPepz) designed to mimic trypsin. *Proc. Natl Acad. Sci. USA* 1994; **91**: 4110–4114.
13. Wallace AC, Laskowski RA, Thornton JM. Derivation of 3D coordinate templates for searching structural databases. Application to Ser-His-Asp catalytic triads in the serine proteinases and lipases. *Protein Sci.* 1996; **5**: 1001–1013.
14. Schnoelzer M, Alewood P, Jones A, Alewood D, Kent SBH. *In situ* neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *Int. J. Pept. Protein Res.* 1992; **40**: 180–193.
15. Guibe F, Dangles O, Balavoine G. Palladium-catalyzed reaction of tributyltin hydride. Selective and very mild deprotection of allyl and allyloxycarbonyl derivatives of amino-acids. *Tetrahedron Lett.* 1986; **27**: 2365–2368.
16. Kates SA, Sole NA, Johnson CR, Hudson D, Barany G, Albericio F. A novel convenient, three-dimensional orthogonal strategy for solid-phase synthesis of cyclic peptides. *Tetrahedron Lett.* 1993; **34**: 1549–1552.