

Computational analysis of the aminic subsite of PGA explains the influence of amine structure on enantioselectivity

Alessandra Basso^a, Paolo Braiuca^a, Sergio Clementi^b, Cynthia Ebert^{a,1},
Lucia Gardossi^{a,*}, Paolo Linda^a

^a Dipartimento di Scienze Farmaceutiche, Università degli Studi, piazzale Europa 1, 34127 Trieste, Italy

^b Laboratorio di Chemiometria, Dipartimento di Chimica, Università degli Studi, via Elce di Sotto 10, 06123 Perugia, Italy

Received 7 September 2001; received in revised form 30 April 2002; accepted 14 May 2002

Abstract

PGA in toluene catalyses the resolution of aromatic amino acids with high enantiomeric excess (>90%) whereas for aliphatic amino acids the enantiomeric excess is far lower (<20%). Molecular modelling explains such a behaviour in terms of interactions in the active site and energies of the intermediate tetrahedral complexes. In this paper, the GRID program has found an innovative application in biocatalysis. The GRID analysis indicates that two different regions can be distinguished within the aminic subsite of PGA active site, the first being mainly hydrophobic and the second one hydrophilic. As a consequence, aromatic L-enantiomers are highly energetically favoured with respect to the D-enantiomers. The low difference in energy between the two enantiomers of aliphatic amino acids explains the poor enantioselectivity of PGA for this kind of compounds. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Biocatalysis in organic solvent; GRID; Penicillin G acylase; Molecular modelling; Enantioselectivity

1. Introduction

Penicillin G acylase (PGA) is an important enzyme used in the industrial production of 6-APA by hydrolysis of penicillin G. Its structure is well known and many crystallographic data are accessible [1,2].

Duggleby in 1995 determined the crystal structure of PGA from *E.coli* characterised by the best resolution (1.90 Å, PDB i.d: 1pnk) [2] and an extensive study of the binding of a series of phenylacetic acid derivatives was carried out by Done et al. [1].

Other contributions to the understanding of the structure of this enzyme come from the determination of the crystal structure of a slowly processing precursor mutant of PGA [3] or the crystal structure of a PGA from a mutant of *P. Rettgeri* [4]. Few information about the aminic site of the PGA are till now available and only one complex of PGA–penicillin G has been reported [5].

The particular structure of PGA, presenting a high selective pocket for the acyclic moiety and a wide tolerance for substrates in the aminic subsite, offers a valuable tool for synthetic applications. In the recent years, our group has demonstrated that PGA is highly active in organic solvent, and many aspects of its synthetic activity have been investigated [6–13]. In particular, PGA catalyses the enantioselective resolution of D,L-PhGlyOMe and D,L-(4OH)PhGlyOMe

* Corresponding author. Tel.: +39-40-5583110; fax: +39-40-52572.

E-mail addresses: ebert@univ.trieste.it (C. Ebert), gardossi@univ.trieste.it (L. Gardossi).

¹ Co-corresponding author.

[11], which are molecules of important pharmaceutical interest. In the present paper, the influence of the structure of the amino compound on enantioselectivity of PGA in organic solvent is presented.

Molecular modelling can be a useful technique when studying the interactions between small organic molecules and large proteic targets in order to predict selectivity or to explain experimental results. In this paper, the analysis with GRID program, which is widely employed for the design of biomolecules of pharmaceutical interest [14,15], has found a novel application in biocatalysis.

To gain more insights into PGA–substrate interactions, experimental results obtained in toluene are analysed on the basis of a computational analysis of the aminic subsite.

2. Experimental

2.1. Molecular modelling

2.1.1. Molecular modelling equipment

All calculations were performed on a Silicon Graphics O₂ workstation. The manipulation of molecules, the graphic visualisation, the energy minimisation and the dynamic simulation were performed using the molecular modelling program-package SYBYL version 6.6 [16]. GRID computations were performed with GRID program version 19 [17]. Quantum mechanical computations were performed by Spartan 4.1.2 [18].

2.1.2. Structures of penicillin G acylase

For all calculations, the X-ray crystal structures of the penicillin G amidase from *E.coli* (Brookhaven Protein Data Bank [19], entry number: 1pnk and 1ai6) were used for all the calculations. All water molecules, which are ‘stable’ in the crystal structures, on the basis of their β -factor ($<50 \text{ \AA}^2$), were included in the calculations and the hydrogen positions were calculated using the Biopolymer module of SYBYL.

2.1.3. Force field parameters

The AMBER ‘all-atom’ force field [20] in its SYBYL implementation has been used for all energy calculations, minimisations and dynamics simulations. All the calculations were carried out by

considering only amino acids residues and all the stable structures located—at least partially—within 10 Å radius sphere from O γ of Ser B:1. The minimisations were performed by means of the Powell method, using a distance dependent dielectric constant ($\epsilon_r = 4.5$). This value of dielectric constant mimics a non-polar medium like toluene and allows to obtain a structure similar to the crystallographic one without modelling solvent molecules [21].

2.1.4. Tetrahedral intermediates structures

All tetrahedral intermediates were sketched bonding the O γ of the serine residue (Ser B:1) and the amino group of the nucleophile to the reactive carbonyl carbon of the phenyl acetic residue. This carbon atom changes to a tetrahedral sp^3 -hybridized configuration. The partial charges of the tetrahedral intermediates were calculated by an *ab initio* algorithm with the 6-31G** basis set (Spartan). In the molecular mechanics calculations, the standard AMBER atom types were used for the atoms of the tetrahedral intermediates, while bond lengths, angles and torsions on tetrahedral carbon were constrained to the values obtained by the *ab initio* optimisation.

2.1.5. Molecular dynamics

A molecular dynamics calculation was carried out on each of the tetrahedral intermediate. The simulation proceeded by a brief *relaxation* for a period of 3 ps. Then, 10 cycles of a *simulated annealing* were performed on the complex by heating and cooling it from 700 to 0 K for a time period of 10 ps.

2.1.6. GRID

The grid dimension was set to include completely the aminic subsite of the PGA and the grid spacing was set to 1 Å. Calculations were performed with the different single-atom probes, representing the chemical groups of the nucleophiles: C1=, C3, DRY, O, OES, N2:, H, H₂O. The docking was performed by means of the GROUP program on a database of conformers generated by a systematic search of the nucleophiles using the DRY probe for carbon atoms.

2.2. Experimental

2.2.1. Materials

PGA-450 (Roche) has a water content of 62.3% and an activity of 122 U g⁻¹ and it was partially

dehydrated to 30% of water before use with the aid of Celite® R-640 (Fluka) according to a procedure already reported [8].

Methyl 4-hydroxyphenylacetate and all amines were obtained from Aldrich. AlaOMe, LeuOMe, PheOMe free amines were obtained from their corresponding hydrochlorides (see Section 2.2.4).

2.2.2. Determination of water activity

Water activity was measured using a hygrometer DARAI (Trieste, Italy). Measurements were carried out by sealing the sensor into the open end of the glass vials, thermostatted, until constant reading. All samples were previously equilibrated at 30 °C for at least 24 h in an air-bath type thermostatted orbital shaker (Thalassia-Trieste, Italy) [8].

2.2.3. Assay of PGA activity in water

Enzymatic activity of PGA-450 was assayed in sodium phosphate buffer 0.05 M by automated titration (TTT80 Radiometer, Denmark) of the phenylacetic acid formed during the hydrolysis of benzylpenicillin potassium salt (Fluka). One enzymatic unit corresponds to the amount of enzyme that hydrolyses 1 µmol of benzylpenicillin in 1 min at pH 8.0 and 37 °C [8].

2.2.4. Enzymatic acylation of D,L-amino acids

A total of 240 µmol of D,L-amino acid and 120 µmol of methyl 4-hydroxyphenylacetate were added to 6 ml of dry toluene with 50 mg of PGA-450 and incubated in a thermostatted orbital shaker at 30 °C.

For AlaOMe, LeuOMe and PheOMe, the corresponding hydrochloride and an equimolar amount of NaHCO₃·10H₂O were suspended in toluene and the mixture was stirred for 30 min at 30 °C. The free amine dissolved in the toluene was then transferred to the vial containing the enzyme and methyl 4-hydroxyphenylacetate.

After 24 h, the reactions were analysed simply by removing the organic phase. Amides were characterised by ¹H and ¹³C NMR (Varian: Gemini 200 Spectrometer at 200 MHz).

2.2.5. Determination of enantiomeric excess

Enantiomeric excess values (see Table 1) were calculated by analysing samples from the enzymatic reactions by LEC-HPLC or by ¹H NMR.

Table 1

Comparison of enantiomeric excess (experimental data) and difference of complex energy ($\Delta\Delta E$) (molecular modelling simulations)

Nucleophile	ee% ^a	$\Delta\Delta E$ (kcal mol ⁻¹)
D,L-AlaOMe	24	0.3
D,L-LeuOMe	20	1.2
D,L-PheOMe	>98	9.7
D,L-TyrOMe	94	10.1
D,L-TrpOMe	92	7.6
D,L-PhGlyOMe	>98	8.3
D,L-4-HO-PhGlyOMe	>98	7.4
R,S- α -Methylbenzylamine	0	0.9

^a The enantiomeric excess (ee%) was calculated at 50% conversion of the amine compound and the value was referred to the non-reacted amine. The ee% values for aromatic amines (PheOMe, TyrOMe and TrpOMe) were calculated by analysing samples withdrawals from the enzymatic reaction by LEC-HPLC. The enantiomeric excess for aliphatic amino acids (AlaOMe and LeuOMe) and α -methylbenzylamine was calculated by ¹H NMR analysis of the diastereomeric salts prepared with R-mandelic acid in 1 ml of CDCl₃. Enantiomeric excess for the reaction with D,L-phenylglycine methyl ester (D,L-PhGlyOMe) and D,L-4-hydroxyphenylglycine methyl ester (D,L-4-HO-PhGlyOMe) was calculated as previously reported [11].

LEC-HPLC is a chromatographic technique based on the separation of the two enantiomers of chiral amino acids by forming a diastereomeric complex between a copper ion-amino acid (present in the mobile phase) and the chiral stationary phase (*N*-alkyl L-hydroxyproline). This technique is particularly useful for the separation of non-derivatised D,L-aromatic amino acids. LEC-HPLC analyses (for PheOMe, TyrOMe and TrpOMe) were performed using a Regis® Davankov LEC Column (mobile phase methanol:water 15:85 containing 5×10^{-4} M of copper acetate monohydrate, isocratic conditions, flow 2 ml min⁻¹, $\lambda_{\max} = 254$ nm). Before the analysis, all samples underwent basic hydrolysis. Samples from the enzymatic reaction mixtures were mixed to an equal volume of NaOH (pH 12) and warmed for 10 min. The samples were subsequently acidified with HCl 1N to Congo red, diluted to 500 µl with ultrapure water and analysed by LEC-HPLC. The hydrolysis procedure is selective for ester group, leaving the amide bond intact, and does not cause racemisation.

The enantiomeric excess for aliphatic amino acids (AlaOMe and LeuOMe) and α -methylbenzylamine was calculated by forming diastereomeric salts with

R-mandelic acid in 1 ml of CDCl_3 and analysing them with ^1H NMR.

The values of enantiomeric excess were calculated as follows:

$$\text{ee\%} = \frac{c^{\text{D}} - c^{\text{L}}}{c^{\text{D}} + c^{\text{L}}} \times 100$$

where c^{D} and c^{L} correspond to the concentration of the enantiomer D and L, respectively (or *R* and *S*, respectively).

Enantiomeric excess for the reaction with D,L-phenylglycine methyl ester (D,L-PhGlyOMe) and D,L-4-hydroxyphenylglycine methyl ester (D,L-4-HO-PhGlyOMe) was calculated as previously reported [11].

3. Results and discussion

3.1. Enantioselectivity of penicillin G acylase: enzymatic reactions

In a recent study, we have reported that PGA in organic solvent accepts different L-amino acids. The present paper aims to get more insights into the steric and structural features of the nucleophile (namely the amino compound) which affect the enantioselectivity of penicillin G acylase.

Enzymatic reactions were performed in toluene using an equimolar mixture of the acyl donor and both the enantiomers of the nucleophiles (Scheme 1).

The enantiomeric excess was calculated after 24 h of reaction, at 50% of the amine conversion into the corresponding amide so that all the acyl donor was consumed. The obtained enantiomeric excess values indicate that, at equal degrees of conversion, PGA is highly enantiospecific (enantiomeric excess >90%)

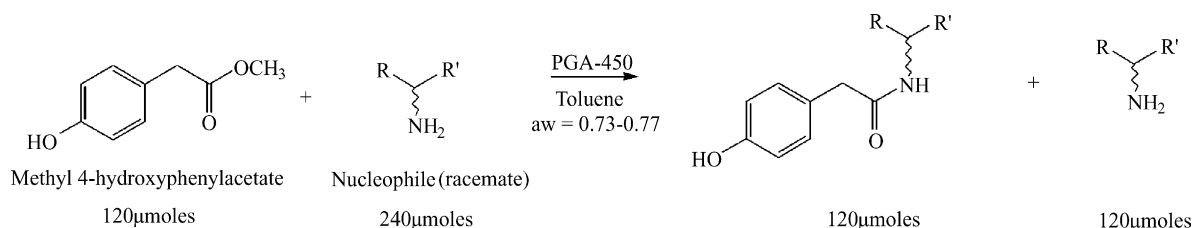
towards aromatic amino acids, whereas it is poorly enantioselective towards enantiomers of aliphatic amino acids. Analyses performed during the reaction time course confirm such low enantioselectivity.

3.2. Enantioselectivity of penicillin G acylase: computational

The experimental data were compared to a computational study of enzyme–substrate interactions. Tetrahedral intermediates formed during the acylation reaction between the D- (or L-) amino acid, the acyl donor (methyl 4-hydroxyphenylacetate) and the catalytic Ser B:1, were designed (crystallographic structure of PGA: PDB id: 1pnk [2]). A tetrahedral intermediate can be considered as the chemical species that best simulates, under both a geometric and an energetic point of view, the transition state of the reaction.

The 4-hydroxyphenylacetic derivative structure was sketched and its position taken directly from the crystallographic data (PDB id: 1ai6) [1]. A molecular dynamics simulation was performed to explore the conformational space starting from this crystallographic position.

Concerning the aminic subsite, just one crystallographic structure containing the aminic portion (complex of PGA with penicillin G) has been reported [5]. Hence, nucleophile conformations were calculated by molecular dynamics simulations on a database of conformations generated by a systematic conformational search. A systematic search on torsional bonds of the nucleophiles (either L- and D-enantiomers) with 30° steps was performed. The database was submitted to a bump check to discard unacceptable structures overlapping the protein or giving origin to atoms



Scheme 1. Scheme of the enzymatic acylation of some nucleophiles catalysed by PGA-450 in toluene. For relative structures of nucleophiles, see Table 1.

bump. The dynamic simulations were performed for the remaining structures and the final lowest energy conformation was chosen for the minimisation.

A further confirm came from the docked complexes calculated by GRID program by using a docking box designed to include the aminic subsite of the 1pnk PGA structure. Despite the slight differences in the atom positions (mainly due to the difference between the covalent tetrahedral species and the non-covalent docked complexes) the docked conformations are wholly consistent with the low energy conformations calculated by the dynamics simulation. Finally, the complete tetrahedral intermediates were minimised and their energy evaluated.

The enantioselectivity of an enzymatic reaction can be related to the difference of complex energies ($\Delta\Delta E$) of the two enantiomers [22]. The difference of complex energy ($\Delta\Delta E$) is given by

$$\Delta\Delta E = \Delta E_L - \Delta E_D$$

where ΔE_L and ΔE_D correspond to the complex energies of tetrahedral intermediate obtained either with

the L- or the D-enantiomer.

$$\Delta E_L = E_{L-tet} - (E_{enz} + E_{sub})$$

$$\Delta E_D = E_{D-tet} - (E_{enz} + E_{sub})$$

where E_{L-tet} is the energy of the tetrahedral intermediate with the L-enantiomer, E_{D-tet} the energy of the tetrahedral intermediate with the D-enantiomer, E_{enz} the energy of the free enzyme and E_{sub} is the energy of the free substrates. Since E_{sub} for the two enantiomers can be considered the same in a non-chiral environment and E_{enz} is the same, it follows that:

$$\Delta\Delta E = E_{L-tet} - E_{D-tet}$$

A value of $\Delta\Delta E$ around 5 kcal mol^{-1} is usually related to enzymatic enantiospecificity (Table 1) [23,24]. Molecular modelling results are in agreement with experimental data and clearly indicate the enantiospecificity of the enzyme towards aromatic L-amino acids.

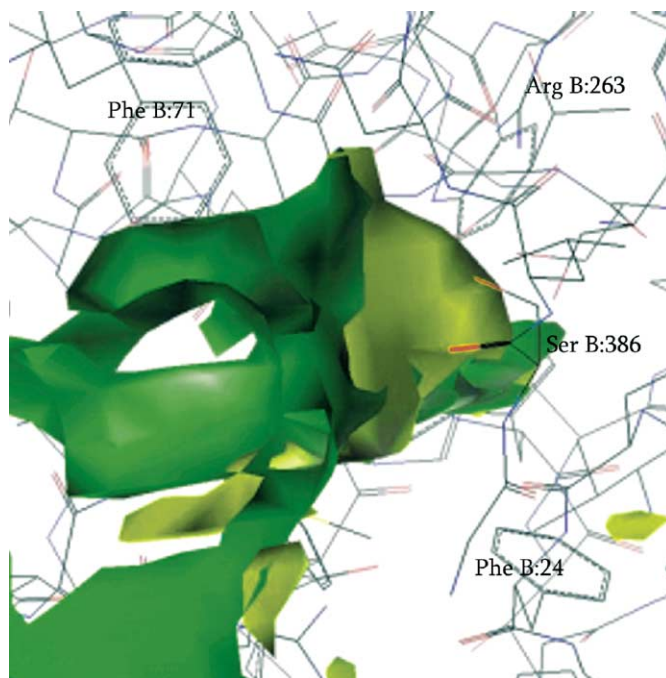


Fig. 1. Superimposition of the GRID surfaces generated by the hydrophobic probe C1= (green area) and the carbonyl oxygen probe O (yellow area).

3.3. Analysis of catalytic subsites of PGA with GRID

In order to understand the reasons of PGA variation in selectivity, the aminic subsite containing the amine portion of tetrahedral intermediates was

structurally analysed using the GRID program. GRID is a computational procedure, developed by Goodford in 1985, which detects energetically favourable binding sites on molecules of known three-dimensional structure [25–27]. This tool allows to estimate the

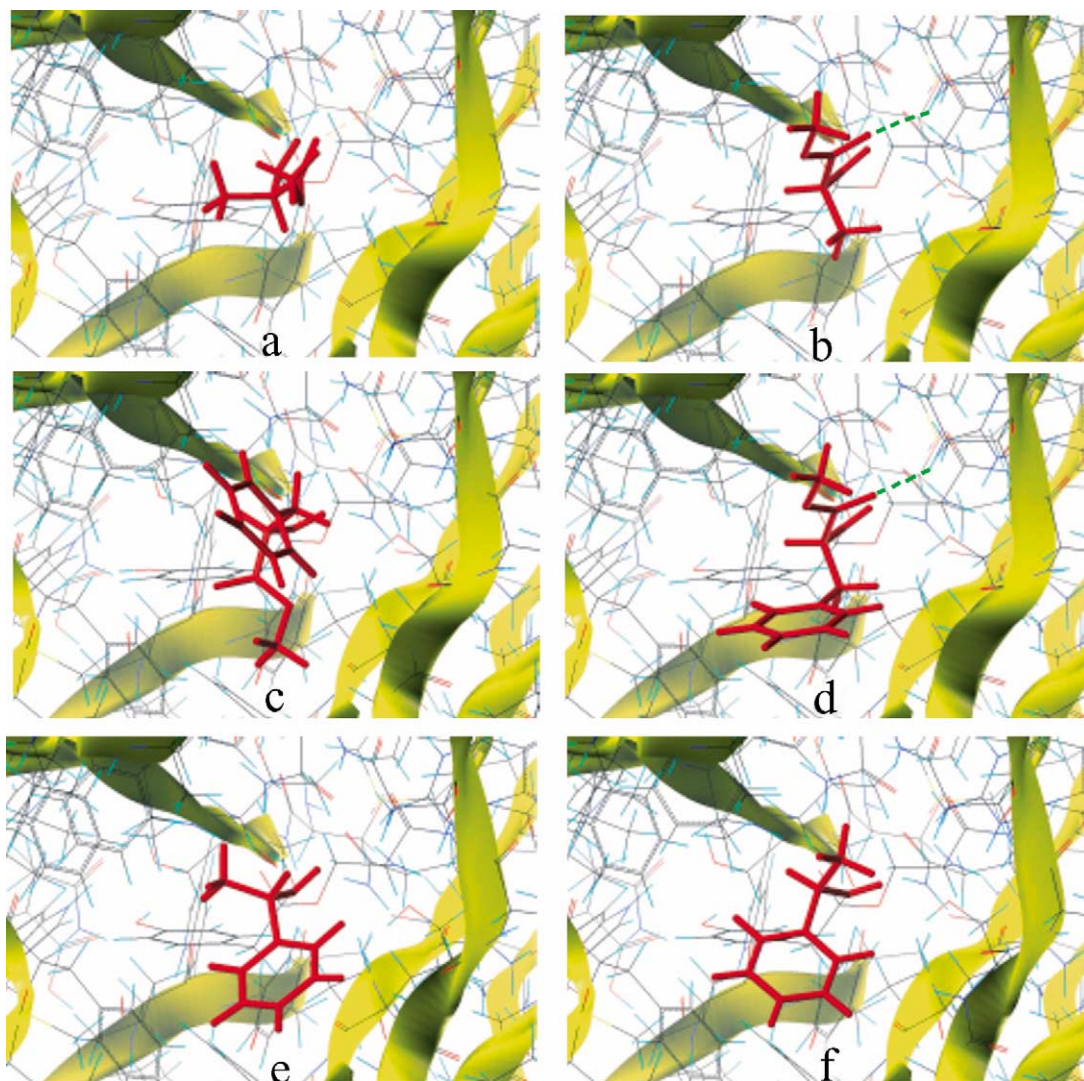


Fig. 2. Tetrahedral intermediates of some representative nucleophiles. Red bold structures represent the aminic portion of tetrahedral intermediates and green dotted lines the H-bonds between the carbonyl group and Arg B:263. D-AlaOMe (a) and L-AlaOMe (b) both form H-bond with Arg B:263 and the shift of the methyl group does not influence significantly the energetic situation. For LeuOMe, the situation is similar. In the case of aromatic amino acids (D-PheOMe (c) and L-PheOMe (d)) only L-enantiomers form H-bond with Arg B:263 with a consequent energetic stabilisation respect the D-enantiomers. R- and S- α -methylbenzylamine ((e) and (f), respectively), lacking in the ester group, can not take a good stabilisation in the aminic subsite. The consequence is that none of the two enantiomers is appreciably stabilised with respect to the other.

energies of interaction between small chemical groups, the *probes*, and a *target*, which is represented, in the present case, by the catalytic site of PGA.

The nature of the aminic subsite was analysed using two different probes, the aromatic carbon probe and the carbonyl oxygen probe (C1= and O) [17]. Fig. 1 shows the catalytic site that locates the nucleophile and the isoenergy surfaces generated by each of the two probes. GRID calculations indicate that the aminic subsite can be divided in two portions. The presence of many aromatic amino acids (Phe A:146, Phe B:71, Leu B:253, Phe B:256) generates an apolar zone represented by a green surface, whereas polar amino acids (Arg B:263, Ser B:386, Asn B:241, Asn B:388) lead to a second zone represented as a yellow surface.

3.4. Structure analysis of tetrahedral intermediates generated by D- and L-amino acids

Tetrahedral intermediates of D,L-enantiomers of both aromatic and aliphatic compounds were compared. All substrates locate the carbonyl group and hydrophobic portion as indicated by GRID analysis.

The interaction of the carbonyl of the ester group in the hydrophilic portion is characterised by a great energetic stabilisation (5 kcal mol^{-1}), whereas the interaction between the hydrophobic portion of the substrate with the enzyme is lower ($-1.5 \text{ kcal mol}^{-1}$). Aromatic L-amino acids (such as PheOMe or TrpOMe) set their aromatic ring in the hydrophobic portion of the aminic subsite (green area, Fig. 1), while the ester group places in the hydrophilic one (yellow area, Fig. 1). Furthermore, the structure is stabilised by a H-bond between the carbonyl group of the ester and Arg B:263. Corresponding D-enantiomers, due to the high steric hindrance of their side chain, are constrained to allocate their ester group outside the energetic favoured position (yellow area, Fig. 1). This generates a destabilisation of at least 5 kcal mol^{-1} .

Aliphatic substrates (AlaOMe and LeuOMe) have a higher conformational freedom, within the aminic subsite, compared to aromatic amino acids. Both enantiomers locate the ester group in the hydrophilic portion, and the small energetic differences observed are due to the shift of the side chain. Hydrogen bonding between the ester group and Arg B:263 is observed for both enantiomers, so that destabilisation

of the D-enantiomer is negligible with respect to the L-enantiomer (about 1 kcal mol^{-1}) (Fig. 2).

The behaviour of α -methylbenzylamine confirms such a finding. This molecule, that resembles the PhGlyOMe but has a methyl instead of an ester group, cannot form the H-bond with the Arg B:263. The negligible destabilisation of the S-enantiomer is experimentally confirmed by the measured enantiomeric excess.

4. Conclusions

Results obtained show that PGA, in toluene, catalyses the resolution of aromatic amino acids with high enantiomeric excess ($>90\%$) whereas for aliphatic amino acids the enantiomeric excess is far lower ($<20\%$).

Molecular modelling explains such a behaviour in terms of interactions and complex energies. The novel application of GRID analysis indicates that two different regions can be distinguished within the aminic subsite of PGA active site. The structural analysis of tetrahedral intermediates gives information about the location and conformation of aminic substrates in the subsite. Aromatic amino acids locate their aromatic rings and ester groups in such a position that L-enantiomers complexes are highly energetically favoured with respect to the D-enantiomers. The low difference in energy between the two enantiomers of aliphatic amino acids explains the poor enantioselectivity of PGA for this kind of compounds.

The authors would like to thank Gabriele Cruciani for helpful discussion. Thanks are due to E.U. contract number QLK3-CT-2000-00160 (COMBIOCAT project) for financial support to L.G. and to MURST (Rome) for financial support to P.L.

References

- [1] S.H. Done, J.A. Brannigan, P.C.E. Moody, R.E. Hubbard, J. Mol. Biol. 284 (1998) 463.
- [2] H.J. Duggleby, S.P. Tolley, C.P. Hill, E.J. Dodson, G. Dodson, P.C.E. Moody, Nature 373 (1995) 264.
- [3] L. Hewitt, V. Kasche, K. Lummer, R.J. Lewis, G.N. Murshudov, C.S. Verma, G.G. Dodson, K.S. Wilson, J. Mol. Biol. 302 (2000) 887.
- [4] M.A. McDonough, H.E. Kley, J.A. Kelly, Prot. Sci. 8 (1999) 1971.

- [5] W.B. Alkema, C.M. Hensgens, E.H. Kroezinga, E. De Vries, R. Floris, J.M. Van Der Laan, B.W. Dijkstra, D.B. Janssen, *Prot. Eng.* 13 (2000) 857.
- [6] C. Ebert, L. Gardossi, P. Linda, *Tetrahedron Lett.* 37 (1996) 9377.
- [7] C. Ebert, L. Gardossi, P. Linda, *J. Mol. Catal. B: Enzym.* 5 (1998) 241.
- [8] L. De Martin, C. Ebert, G. Garau, L. Gardossi, P. Linda, *J. Mol. Catal. B: Enzym.* 6 (1999) 437.
- [9] A. Basso, L. De Martin, C. Ebert, L. Gardossi, P. Linda, V. Zlatev, *J. Mol. Catal. B: Enzym.* 11 (2001) 851.
- [10] A. Basso, L. De Martin, C. Ebert, L. Gardossi, A. Tomat, M. Casarci, O. Li Rosi, *Tetrahedron Lett.* 41 (2000) 8627.
- [11] A. Basso, P. Braiuca, L. De Martin, C. Ebert, L. Gardossi, P. Linda, *Tetrahedron: Asymm.* 11 (2000) 1789.
- [12] A. Basso, S. Biffi, L. De Martin, C. Ebert, L. Gardossi, P. Linda, *Croatia Chem. Acta* 74 (2001) 757.
- [13] A. Basso, L. De Martin, C. Ebert, L. Gardossi, P. Linda, *J. Mol. Catal. B: Enzym.* 16 (2001) 73.
- [14] G. Cruciani, K. Watson, *J. Med. Chem.* 37 (1994) 2589.
- [15] M.A. Kastenholz, M. Pastor, G. Cruciani, E.E.J. Haaksma, T. Fox, *J. Med. Chem.* 43 (2000) 3033.
- [16] SYBYL (version 6.6) distributed by Tripos Ass., St. Louis, Missouri.
- [17] GRID (version 19) distributed by Molecular Discovery Ltd., London, UK.
- [18] Spartan (version 4.1.2) distributed by Wavefunction Inc., Irvine, California.
- [19] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, *Nucleic Acid Res.* 28 (2000) 235.
- [20] S.J. Weiner, P.A. Kollman, D.T. Nguyen, D.A. Case, *J. Comput. Chem.* 7 (1986) 230.
- [21] I.T. Christiansen, F.S. Jørgensen, *J. Biomol. Struct. Dyn.* 3 (1997) 473.
- [22] M. Norin, K. Hult, A. Mattson, T. Norin, *Biocatalysis* 7 (1993) 131.
- [23] A. Fersht, *Enzyme Structure Mechanism*, Freeman, New York, 1985, p. 316.
- [24] T. Ke, B. Tidor, A.M. Klibanov, *Biotechnol. Bioeng.* 57 (1998) 741.
- [25] P.J. Goodford, *J. Med. Chem.* 28 (1985) 849.
- [26] D.N.A. Boobbyer, P.J. Goodford, P.M. McWhinnie, R.C. Wade, *J. Med. Chem.* 32 (1989) 1083.
- [27] R. Wade, K.J. Clerk, P.J. Goodford, *J. Med. Chem.* 36 (1993) 140.