Journal of Molecular Graphics and Modelling 36 (2012) 30–35

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Engineered enzyme interactions with polycyclic aromatic hydrocarbons: A theoretical approach

Vito Librando[a,b,*], Matteo Pappalardo[a,b]

[a] Dipartimento di Scienze Chimiche, Università di Catania, Viale A. Doria 6, 95125 Catania, Italy
[b] Research Center for Analysis, Monitoring and Minimization Methods of Environmental Risk, Chemical Science Building, Viale A. Doria 6, 95125 Catania, Italy

ARTICLE INFO

Article history:
Received 14 November 2011
Received in revised form 20 February 2012
Accepted 28 February 2012
Available online 23 March 2012

Keywords:
PAH
Enzyme
Molecular dynamics
Docking
Bioremediation

ABSTRACT

In this paper, the techniques of modelling, docking and molecular dynamics were used to study eight single amino acid mutations of the enzyme PhnI to optimise its enzymatic degradation capability. The eight mutants were first equilibrated to avoid deformations of the secondary and tertiary structure and to minimise alterations in the functionality of the chimera enzymes that were obtained. For this purpose, we monitored the potential energy of the systems and the fluctuations of the backbone of the enzymes. The structures of mutant enzymes, at equilibrium, were subjected to docking calculations with selected PAHs. The results indicated a significant increase in the PAH-enzyme interaction with respect to the wild-type protein. The considerable computing resources offered by the GRID computing system made it possible to perform calculations on the entire enzyme system, consisting of six protein subunits, as highlighted in the recent literature.

1. Introduction

Environmental pollution in the form of polycyclic aromatic hydrocarbons (PAHs) is primarily due to human activities, such as the incomplete combustion of coal, oil, gas or timber. The accumulation of these products in the environment is major problem for human health [1,2]. It is well known that carcinogenesis by PAHs involves the production of PAH-hydro-diols through metabolic activation of the enzyme P450 mono-oxygenase, which is followed by the production of epoxide ring-diols that are able to fit easily in the grooves of the DNA and by a reaction with the DNA that results in carcinogenesis and genetic alterations.

Some mechanisms for the degradation of these molecules are available in nature, but their efficiency decreases with increasing PAH molecular weight. As a consequence, PAHs with a high molecular weight have the tendency to accumulate in the environment, which is the major problem [3,4]. To improve natural PAH degradation mechanisms and to increase our knowledge of the processes and reactions of molecular bio-degradation, many studies are currently in progress [5–11].

Recent advances in the in silico mutation techniques have led to a better understanding of enzymatic degradation by selective mutation of specific enzymes [12–18] and a better knowledge of the bio-degradation reactions. In addition, recent computer advances have led to an important approach that can be used in enzyme degradation studies: docking calculations applied to naphthalene dioxygenase (NDO) demonstrate that all PAHs can interact with the active site of NDO in a thermodynamically favourable manner. However, only the steric factor for both the enzyme and the PAH determines those molecules that can efficiently react with the enzyme [19]. Despite the high number of known PAHs, only a small number of PAHs are present in the environment at high concentrations, which allows us to focus on a small number of PAHs. As reported in the recent literature, the PAHs benz(a)anthracene, benzo(g,h,i)perylene, chrysene, fluoranthene, naphthacene and indeno(1,2,3-cd)pyrene are frequently found in contaminated environments [20].

Recently, Jakonic et al. have solved the three-dimensional structure [21] of a dioxygenase from Sphingomonas CHY-1 [22], known as PhnI and showed a good degradative capacity with respect to PAHs.

Recently, Librando and Pappalardo [23] used a theoretical approach to show that the enzyme PhnI can effectively remove the aforementioned PAHs but is ineffective for high molecular weight PAHs. This result highlights how the shape of a molecule represents an important parameter that determines the subsequent interaction between the pollutant and the active site. Using the mixed approach of molecular modelling, molecular dynamics and docking calculations, we generated new chimera enzymes with enhanced degradation activity with respect to PAHs.

* Corresponding author at: Dipartimento di Scienze Chimiche, Università di Catania, Viale A. Doria 6, 95125 Catania, Italy. Tel.: +39 095 7385201; fax: +39 095 330424.
E-mail address: vlibrando@unict.it (V. Librando).

1093-3263/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2012.02.005

2. Methods

The native enzyme is composed of three subunits and each subunit is composed of a Rieske-cluster domain and a catalytic domain hosting one mononuclear iron. Fig. 1 presents the amino acid composition of each active site. The structure of the native enzyme, obtained from the protein databank (PDB code: 2CKF), has been modified by "in silico" mutagenesis using the Accelrys software suite. Eight mutants were generated from the structure of the native enzyme: D204G, F201A, F350A, H207A, I260A, L356A, N200G and N295G. The choice of amino acid change was made based on the results of docking calculations of PAHs with the native protein, which were performed in a previous paper [23]. In particular, the amino acid selected was always located in the active site near the iron atom and chosen based on specific hydrophobic/hydrophilic ratios and steric factors. The mutations were chosen to cause as little change as possible in the secondary and tertiary structure of the protein and the active site. For these reasons, we decided to replace the amino acids of the native protein only with glycine or alanine. The molecular dynamics calculations were performed using the software NAMD7 [24] and CHARMM force field [25]. The final box size was $100 \times 110 \times 105$ Å and contained approximately 115,000 atoms. Solvated molecules were first minimised (conjugate gradient algorithm) and then gradually heated to 300 K and the molecules were equilibrated with a thermal bath at 300 K for 1 ns, adopting an NPT ensemble. The time step was set to 2 ps and we adopted the 'rigid bond' algorithm of NAMD. The interactions between atoms that were not covalently bonded have been treated by adopting a cut-off of 12 Å. The structures of the selected PAHs were constructed using the editor of the Accelrys software package and minimised using the CHARMM force field and adopting Gasteiger charges [26]. For the docking calculations, we adopted a grid large enough to completely cover the active site and to cover part of the corresponding protein subunits [27]. The grid, with an edge length of 60 Å, was centred on the mononuclear iron atom of the active site. The docking calculations were performed using the Autodock Vina software [28] and the structure was obtained using the average structure from the final 10 ps of the simulation.

The calculation of the electrostatic potential was performed using the APBS software [29]. In this work, we adopted the full Coulombic approximation. In particular, the solution of Poisson–Boltzmann equation for a single point with charge q_i placed at the origin of the axes has the Debye–Huckel form:

$$\phi_i = q_i \varepsilon_r^{-1} e^{-kr}$$

Here ϕ is the electrostatic potential in units of kT/e (e is the electron charge), ε is the dielectric constant of the water solvent, k is the inverse of the Debye–Huckel length and r represents the spatial coordinates of the i[th] point. The equation was solved using the method of finite approximations [30–32]. The grid used to determine the electrostatic potential has a density of 138 points.

All calculations were performed using the computing resources made available and maintained by the GRID COMETA consortium, which has 2000 CPU and 8 TB RAM.

3. Results

The known structures of the eight mutants were obtained by energy minimisation and the final structure was required to have a stable conformation and to be at equilibrium. To be confident that these conditions were achieved, we monitored the potential energy and RMSD. The potential energies (Fig. 2) for the mutants F350A, F201A, L356A, D204G, I260A and N200G gradually decrease until a constant average value is reached in the final part of the equilibrium phase of the calculation. Alternatively, the mutants H207A and N295G showed a drop in the potential energy at approximately 550 and 300 ps of the simulation time, respectively, after which the energy oscillated around the average value, as expected from a conformationally stable structure. To test the reliability of the punctiform mutation strategy, a control mutant was generated in which all eight amino acids mentioned were mutated simultaneously in the same model enzyme. Even in this case, the energy decreases slowly and reaches a constant value at approximately 800 ps into the simulation. The values of the fluctuations of the backbone (RMSD) determined in this work were calculated using the following equation:

$$RMSD_\alpha(t_i) = \sqrt{\frac{\sum_{\alpha=1}^{N_\alpha} (\vec{r}_\alpha(t_j) - (\vec{r}_\alpha))^2}{N_\alpha}}$$

where N_α is the number of atoms whose position must be compared, N_t is the time interval over which to compare, $\vec{r}_\alpha(t_j)$ is the position of the atom at time t_j and $\vec{r}_\alpha$ is the average position of the atom to which α is compared, which is defined as follows:

$$\left\langle \vec{r}_\alpha \right\rangle = \frac{q}{N_t} \sum_{j=1}^{N_c} \vec{r}_\alpha(t_j)$$

In this work, we defined a value corresponding to the average structure calculated for each mutant. The data reported in Fig. 3 include those values for all mutants and a decrease in the RMSD values of approximately three to one was observed, with fluctuations on the order of 0.5.

The results presented in Fig. 3 show that after the initial decrease of the RMSD, its value remains stable in all cases, indicating that the enzymes are conformationally stable. Even the mutant N295G, after an increase in the RMSD between 600 and 800 ps, is characterised by constant fluctuation of the structure. After showing that the overall three-dimensional structures of enzymes are stable and that they are at equilibrium, we focused our attention on the active site. We calculated the centre of mass (MC) of the active site for each mutant and the distance between the MC and the Cα amino acid ligands to obtained a set of distances that characterise the active site geometry. Because the absolute size of the active site is not useful in this case, we calculated the difference between the size of the catalytic site for each mutant and the size of the native enzyme. Fig. 4 shows the percentage difference of the distances as a function of the primary sequence.

The dotted band in the central part of the figure corresponds to a deviation of less than 5% and values that fall within this range are considered minor changes in the distance. In all mutants, we observed that approximately 50% of the distances are within the 5% region and only a few are beyond the 10% limit. The figure shows the mutated amino acids (black arrows) and the mutations distance values are located farther from the MC than the native enzyme, with the exception of the mutants F350A and F201A. The figure shows that, with the exception of the D204G mutant, the early amino acids (200–230) show the tendency to move away from the reference positions, indicating a possible opening of the active site. In the mutants F350A, H207A, L356A and N295G, a decrease in the distance of amino acid 356 from the MC was observed, with values lower than 15%. The control mutant, indicated in the figure by the number 8, shows that all mutations do not significantly perturb the active site, which is maintained at the average values obtained by analysing the individual mutants.

After determining that the mutants show no appreciable structural distortions, we performed docking calculations between the enzymes and the PAHs selected. Furthermore, to obtain comparable data, we monitored the percentage insertion of the PAH into the enzyme pocket, as identified in the literature [21,22]. In Table 1, we

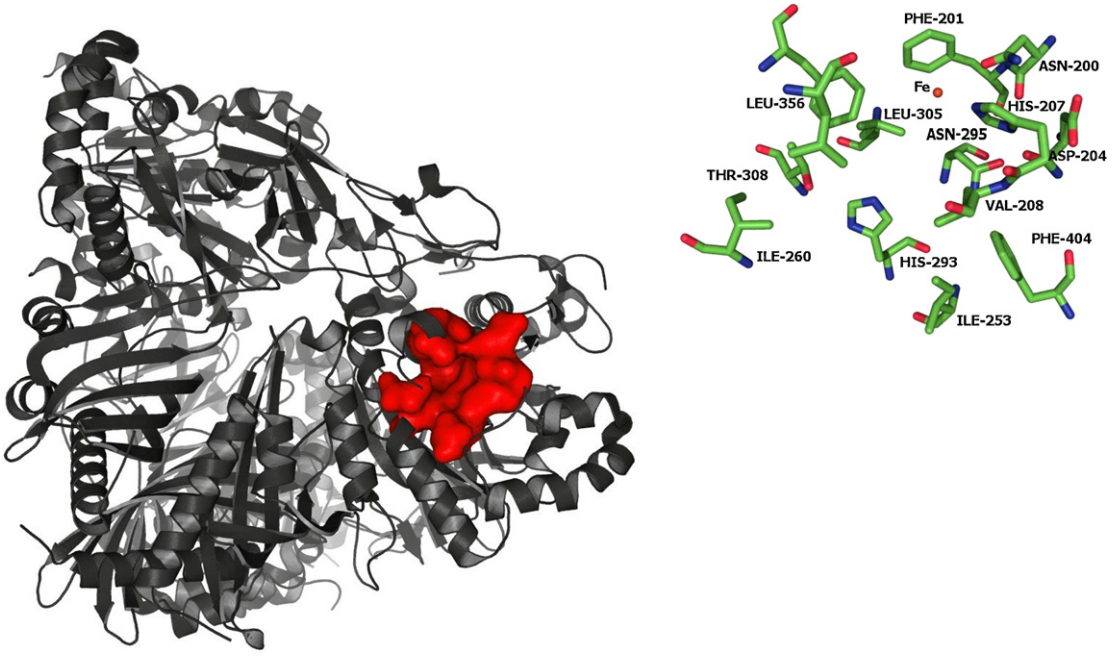

Fig. 1. 3D representation of enzyme studied here; the red surface indicates the active site location. The inset contains the amino acids of the active site, with the iron atom highlighted in orange. (For interpretation of references to colour in this figure legend, the reader is referred to the web version of this article.)

report the data described below and the reference values calculated for the native enzyme.

The results show the existence of a mutation that is active in all PAHs (F350A), except for indeno(1,2,3-cd)pyrene, which for only 50% of cases correctly recognised the catalytic site. In all other cases, there was a significant increase in molecular recognition relative to the native protein. These activity results are important because benzo(ghi)perylene, fluoranthene and naphtacene

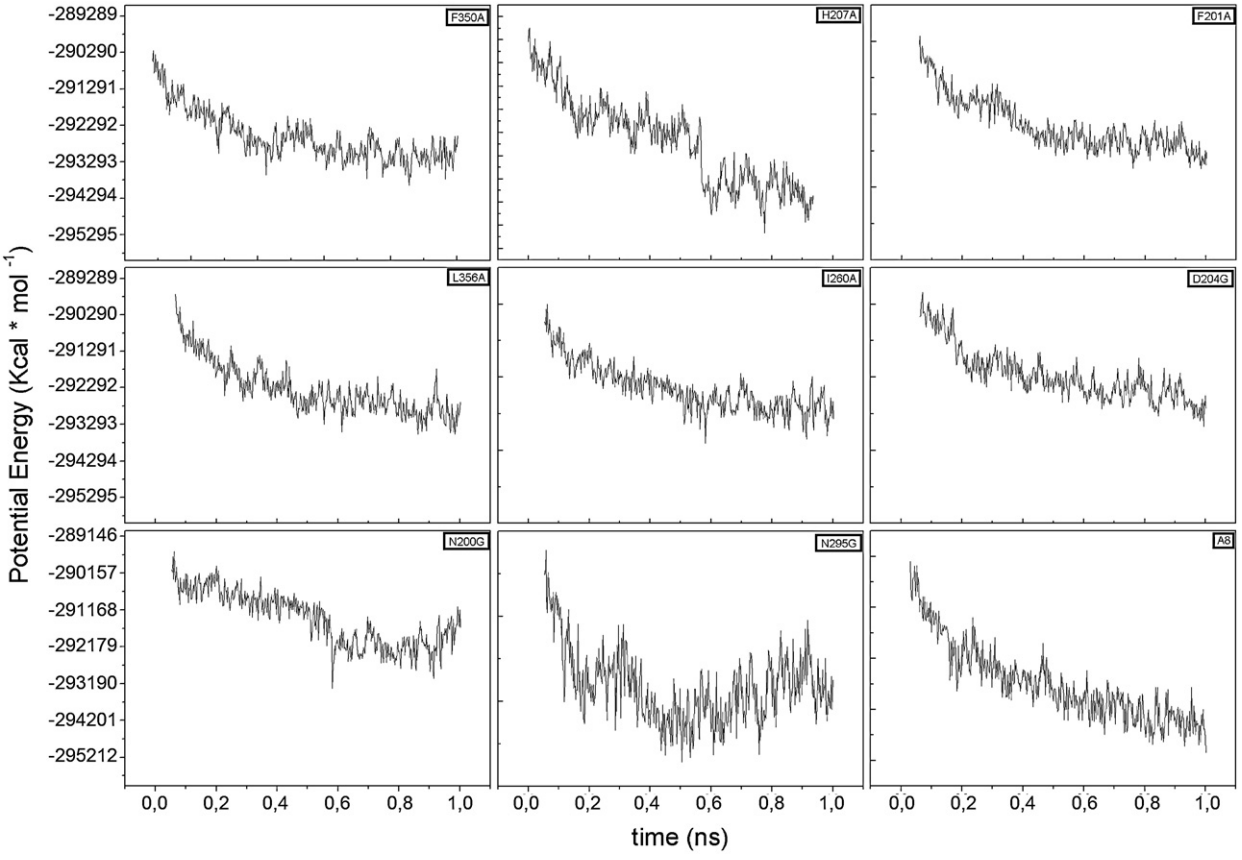

Fig. 2. Potential energy of the mutants investigated as a function of simulation time.

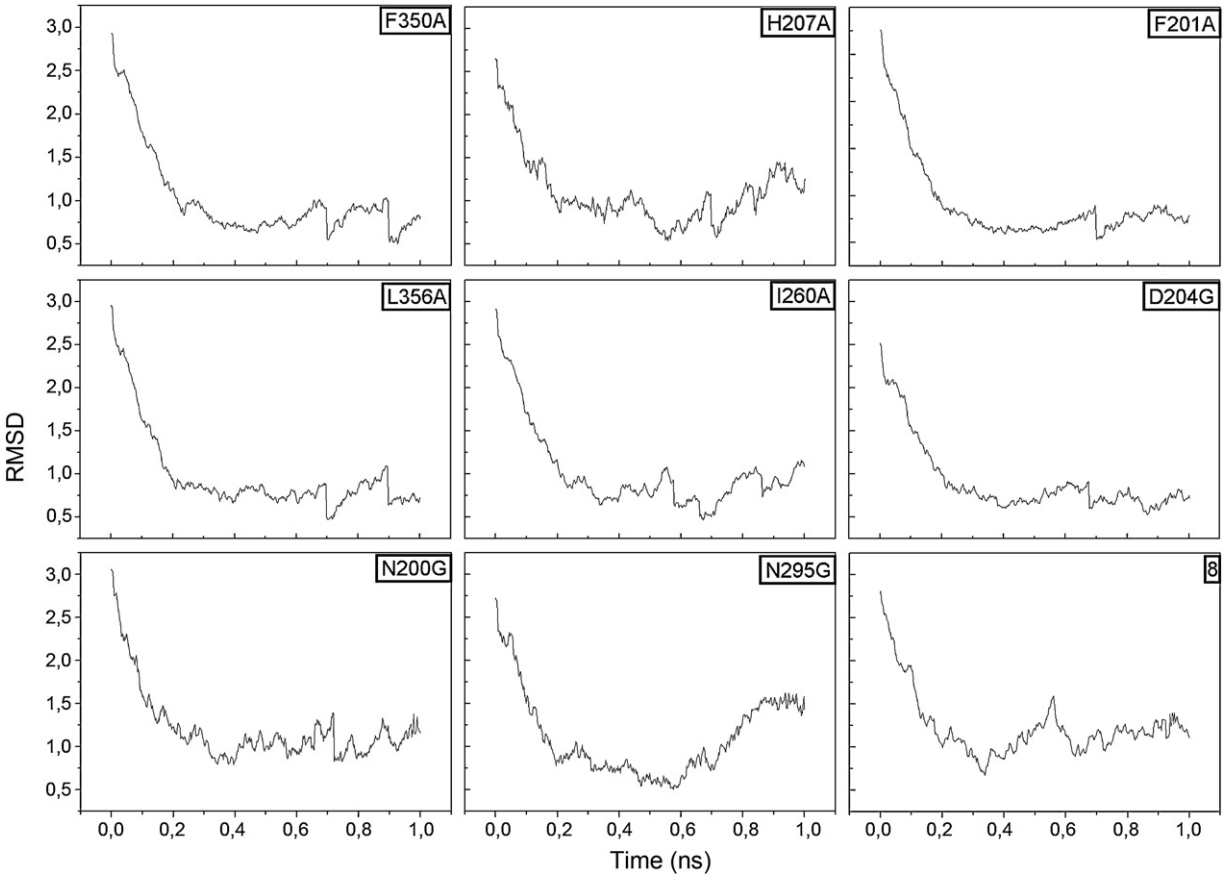

Fig. 3. RMSD of the mutants calculated as a function of simulation time.

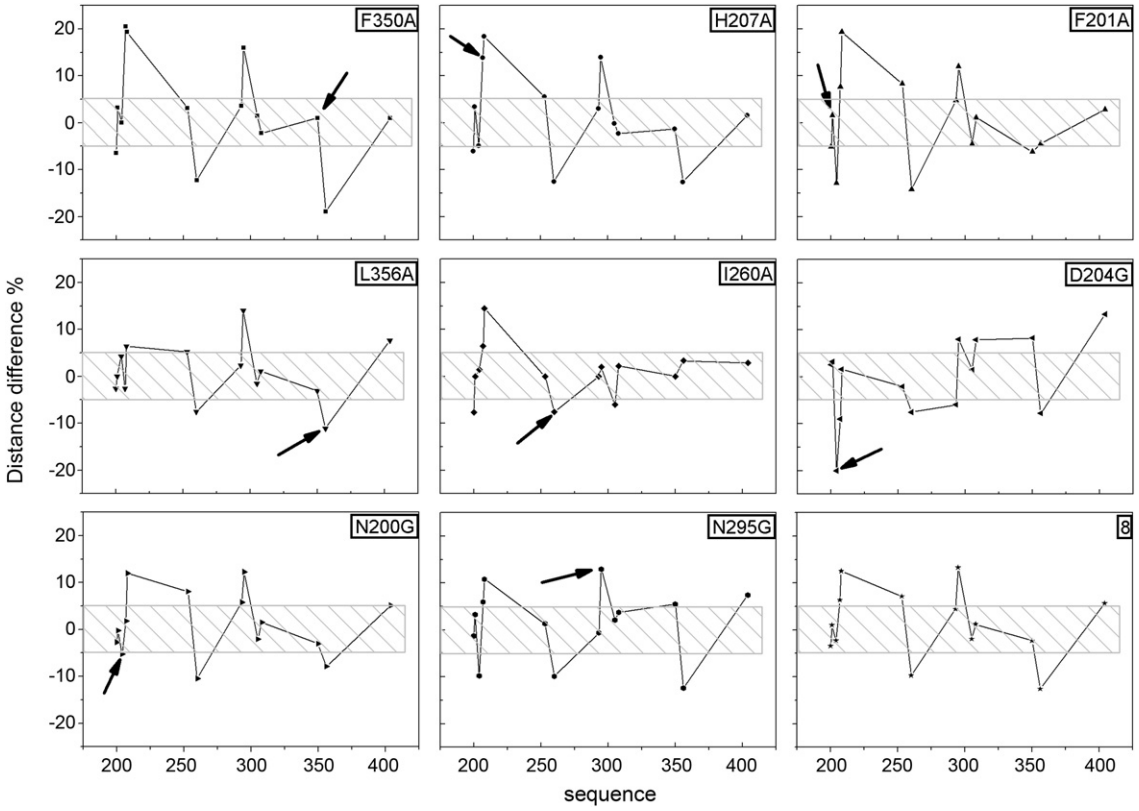

Fig. 4. Changes in the distance of the amino acids of the active site of the mutants compared with the native structure, depending on the primary sequence.

Table 1
Shows the percentages of successful docking of reference PAH with mutant. PAHs considered are: (1) benz(a)anthracene, (2) benzo(ghi)perylene, (3) crysene, (4) fluoranthene, (5) naphtacene, (6) indeno(1,2,3-cd)pyrene.

	D204G	F201A	F350A	H207A	I260A	L356A	N200G	N295G	WT
1	50.0	44.4	100.0	33.3	50.0	28.6	100.0	85.7	32.0
2	0.0	0.0	100.0	0.0	0.0	11.1	25.0	88.9	0
3	44.4	66.7	100.0	44.4	44.4	66.7	66.7	77.8	43.0
4	66.7	44.4	100.0	22.2	44.4	100.0	44.4	100.0	0
5	0.0	0.0	100.0	0.0	11.1	100.0	0.0	92.9	0.0
6	0.0	66.7	50.0	0.0	0.0	44.4	77.8	61.1	53.0

cannot fit within the active site of the wild-type enzyme. Another interesting mutant is N295G, which exhibited a higher number of correct insertions than the reference but was still inferior with respect to the aforementioned mutant. In all other cases, the docking percentages are similar to those of the reference. The I260A mutant has slightly higher values for inclusion for the reference molecules benz(a)anthracene and crysene naphtacene, whereas inclusion was identical for benzo(ghi)perylene and less for indeno(1,2,3-cd)pyrene.

4. Discussion

The docking results indicate that the F350A mutant is the most active molecule towards the selected PAHs because the three-dimensional structure is not distorted. Therefore, the basis for the increased interaction must lie in the active site. The RMSD of the backbone in the last 500 ps of the simulation is between 0.5 and 1 (Fig. 3), which implies that the structure of the mutant is rigid and is not perturbed by the mutation. On the whole, the geometry of the site remains unchanged and alanine 350 (mutated residue) remains at a similar position with respect to that of the native protein. However, residues 207, 208 and 295 move away from the MC of the active site by over 15%; in contrast, residues 260 and 356 move closer to the MC. If these small changes in the mutant significantly increase the free volume within the active site, high molecular weight PAHs may have enough space to fit within the catalytic site. This is detectable by comparing the free volume of the active site of mutant F350 (Fig. 5, red) with that of the native protein (Fig. 5, blue). The presented surfaces, reported as grids, correspond to areas of space occupied by the atoms that constitute the proteins. In Fig. 5, we report the amino acids in the active site and superimposed them on the surface.

The figure shows that the native protein has a small cavity in the centre of the active site (Fig. 5, blue structure) that contains the metal ion. Conversely, the central area of the F350A mutant is empty (structures in red, Fig. 5) and only the peripheral part of the active site was occupied. The data obtained for the mutant, when compared with the native protein, offer the first explanation for the increased efficiency of the enzyme F350A towards high molecular weight PAHs.

The above discussion is confirmed by comparing the volume of the active site of the native enzyme and mutant I260A, which represents the enzyme with the lowest ability to interact with the PAHs. A comparison with the native enzyme (Fig. 6, in yellow) shows that the mutant in question has a cavity volume that is similar to or lower than the reference. These data are also supported by the modest change in the distances between amino acids and the MC (Fig. 4). The distance data show that the amino acids of the active site move away from the MC, while the mutated amino acid (alanine 260) moves closer to the centre of the active site and causes a decrease in the free volume. In addition to the aforementioned steric effect, it is logical to think that energetic factors also affect the results. Because the guest molecules (PAHs) are purely hydrophobic, the active site should have adequate hydrophobicity. We calculated the hydrophobicity value using the method of Kyte-Doolittle (calculated as hydrophobic surfaces of catalytic cavity) and we compared

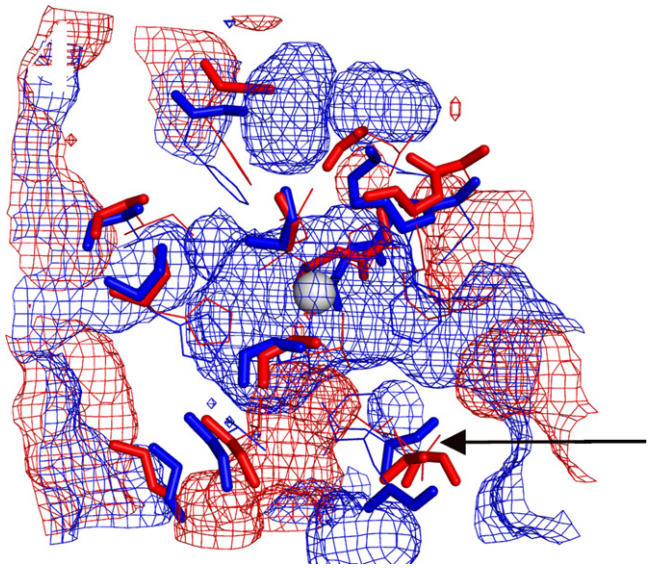

Fig. 5. Representation of the amino acids of the active site (F350A mutant red, blue native protein). At the centre in grey, we present the trivalent Fe atom. Empty areas are highlighted within the enzyme cavity (red mutant F350A, blue native protein). (For interpretation of references to colour in this figure legend, the reader is referred to the web version of this article.)

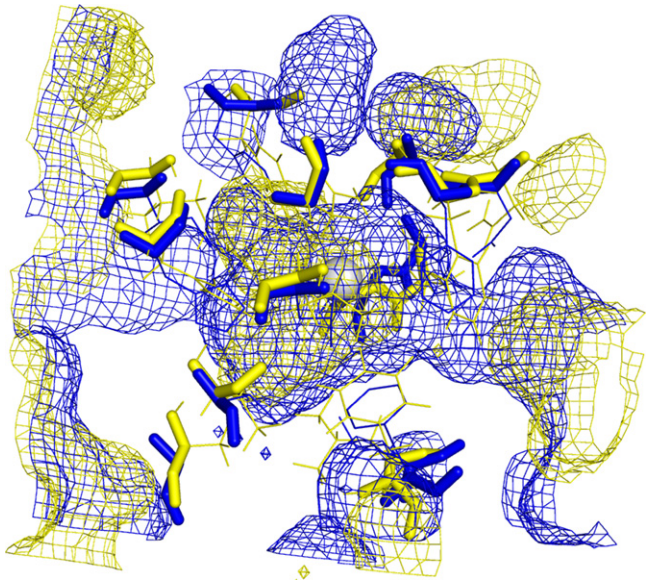

Fig. 6. Representations of the amino acids of the active site (I260A mutant yellow, blue native protein). At the centre in grey, we present the trivalent Fe atom. Empty areas are highlighted within the enzyme cavity (yellow mutant I260A, blue native protein). (For interpretation of references to colour in this figure legend, the reader is referred to the web version of this article.)

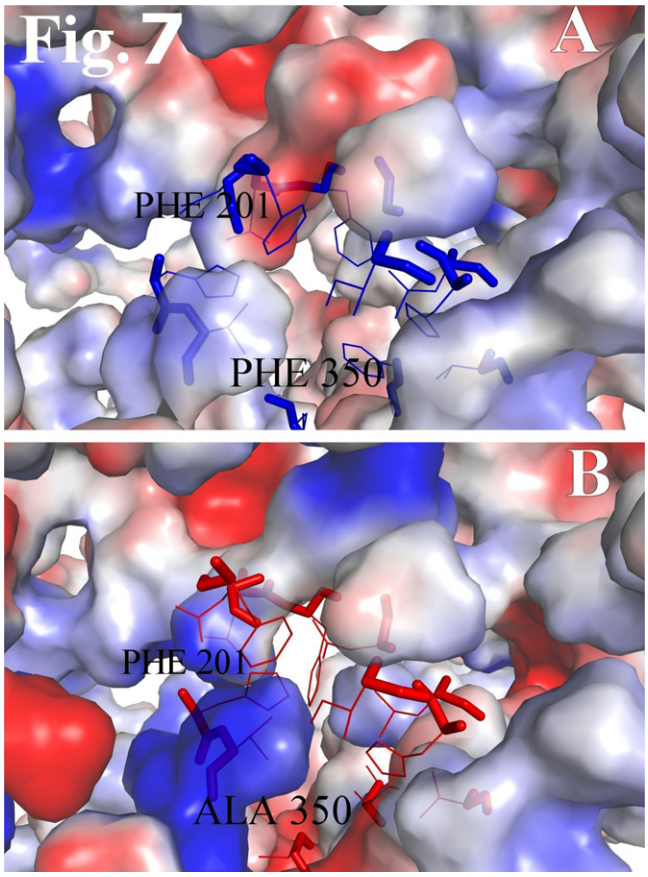

Fig. 7. The top panel shows the electrostatic potential for the native protein. The lower panel shows the electrostatic potential for the mutant F350A.

the obtained data to the native protein. The results did not show significant differences between the mutants and the native enzyme (data not shown). Another parameter that we measured to explain the different behaviour of the investigated mutants was the electrostatic potential. In Fig. 7, we present the calculated electrostatic potential within the active site for the native protein (panel A) and for the F350A mutant (panel B). The native protein has a catalytic site with an overall neutral electrostatic potential, with the exception of the slightly negative value around the phenyl-alanine 201. Conversely, the mutation of residue 350 causes a substantial change in the electrostatic potential of the protein (Fig. 7, panel b). Because the site of the F350A mutant is characterised by a strong positive electrostatic potential and PAHs [33] are characterised by a negative potential, it appears reasonable to assume that an enzyme, such as F350A, could serve as an attractor for a PAH molecule.

5. Conclusion

The results of this work confirm data previously reported in the literature and show that some mutants have the ability to effectively interact with higher molecular weight PAHs. The F350A mutation is the most effective among those selected and is characterised by a slight variation in the geometry of the active site, exhibiting interesting results with respect to PAH interactions. The increase in the free volume of the active site allows for higher molecular weight PAHs to enter the site. Moreover, the mutation causes a change in the electrostatic potential of the active site such that the mutant enzyme serves as an electrostatic attractor for PAHs. This paper highlights how an “in silico” approach can be a useful tool when planning protein engineering experiments.

Acknowledgments

This work has been supported by the Ministero dell’Istruzione, dell’Università e della Ricerca MIUR, Roma, Prin 2009 Program. Computational resources were provided by “Consorzio Cometa”, Catania.

References

[1] F.J. Zhang, C. Cortez, R.G. Harvey, J. Org. Chem. 65 (2000) 3952–3960.
[2] W. Guo, Y. Pei, Z. Yang, H. Chen, J. Hazard. Mater. 187 (2011) 441–449.
[3] S. Krivobok, S. Kuony, C. Meyer, M. Louwagie, J.C. Willison, Y. Jouanneau, J. Bacteriol. 185 (2003) 3828–3841.
[4] F.Y. Wu, X.Z. Yu, S.C. Wu, X.G. Lin, M.H. Wong, J. Hazard. Mater. 187 (2011) 341–347.
[5] A.K. Haritash, C.P. Kaushik, J. Hazard. Mater. 169 (2009) 1–15.
[6] B. Chen, M. Yuan, H. Liu, J. Hazard. Mater. 188 (2011) 436–442.
[7] A. Janbandhu, M.H. Fulekar, J. Hazard. Mater. 187 (2011) 333–340.
[8] A.M. Farnet, S. Criquet, S. Tagger, G. Gil, J. Le Petit, Can. J. Microbiol. 46 (2000) 189–194.
[9] E.C. Santos, R.J.S. Jacques, F.M. Bento, M.D.R. Peralba, P.A. Selbach, E.L.S. Sa, F.A.O. Camargo, Bioresour. Technol. 99 (2008) 2644–2649.
[10] G.J. Verrhiest, B. Clement, B. Volat, B. Montuelle, Y. Perrodin, Chemosphere 46 (2002) 187–196.
[11] C. Biache, T. Ghislain, P. Faure, L. Mansuy-Huault, J. Hazard. Mater. 188 (2011) 221–230.
[12] V. Librando, S. Forte, Biochem. Eng. J. 27 (2005) 161–166.
[13] J.C. Colombo, A. Barreda, C. Bilos, N. Cappelletti, M.C. Migoya, C. Skorupka, Environ. Pollut. 134 (2005) 267–276.
[14] M.D. Fang, C.L. Lee, C.S. Yu, Mar. Pollut. Bull. 46 (2003) 941–953.
[15] A.R. Mostafa, T.L. Wade, S.T. Sweet, A.K.A. Al-Alimi, A.O. Barakat, J. Marine Syst. 78 (2009) 1–8.
[16] I. Zrafi-Nouira, Z. Khedir-Ghenim, F. Zrafi, R. Bahri, I. Cheraeif, M. Rouabhia, D. Saidane-Mosbahi, Bull. Environ. Contam. Toxicol. 80 (2008) 566–572.
[17] Y. Wan, X.H. Jin, J.Y. Hu, F. Jin, Environ. Sci. Technol. 41 (2007) 3109–3114.
[18] V. Fernandez-Gonzalez, S. Muniategui-Lorenzo, P. Lopez-Mahia, D. Prada-Rodriguez, J. Chromatogr. A 1207 (2008) 136–145.
[19] K.H. Wammer, C.A. Peters, Environ. Sci. Technol. 39 (2005) 2571–2578.
[20] E.A.A. Gawad, M. Al Azab, M.M. Lotfy, Environ. Geol. 54 (2008) 1091–1102.
[21] J. Jakoncic, Y. Jouanneau, C. Meyer, V. Stojanoff, Biochem. Biophys. Res. Commun. 352 (2007) 861–866.
[22] J. Jakoncic, Y. Jouanneau, C. Meyer, V. Stojanoff, FEBS J. 274 (2007) 2470–2481.
[23] V. Librando, M. Pappalardo, J. Mol. Graph. Model. 29 (2011) 915–919.
[24] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, J. Comput. Chem. 26 (2005) 1781–1802.
[25] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, J. Chem. Phys. 79 (1983) 926–935.
[26] J. Gasteiger, M. Marsili, M.G. Hutchings, H. Saller, P. Low, P. Rose, K. Rafeiner, J. Chem. Inf. Comp. Sci. 30 (1990) 467–476.
[27] J. Wei, H. Li, W.L. Qu, Q.Z. Gao, Neurochem. Int. 55 (2009) 637–642.
[28] O. Trott, A.J. Olson, J. Comput. Chem. 31 (2010) 455–461.
[29] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 10037–10041.
[30] I. Klapper, R. Hagstrom, R. Fine, K. Sharp, B. Honig, Proteins 1 (1986) 47–59.
[31] M.K. Gilson, B.H. Honig, Nature 330 (1987) 84–86.
[32] A. Nicholls, K.A. Sharp, B. Honig, Proteins 11 (1991) 281–296.
[33] V. Librando, A. Alparone, G. Tomaselli, J. Mol. Model. 14 (2008) 489–497.