A Quantum Mechanics / Molecular Mechanics Study of the Protein-Ligand Interaction of two potent inhibitors of human *O*-GlcNAcase: PUGNAc and NAG-thiazoline

Jeronimo Lameira,[1,2] Cláudio Nahum Alves,[1,*] Vicent Moliner,[2,*] Sergio Martí,[2] Natalia Kanaan,[2] Iñaki Tuñón[3]

[1] *Laboratório de Planejamento e Desenvolvimento de Fármacos, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, CP 11101, 66075-110, Belém, PA, Brazil.*

[2] *Departament de* Química Física i Analítica*, Universitat Jaume I, 12071 Castellón, Spain.*

[3] *Departament de Química Física, Universitat de Valencia, 46100 Burjasot, Valencia, Spain.*

Corresponding authors:

V. Moliner; moliner@uji.es Phone: (+34) 964-728084. Fax: (+34) 964-728066

C. N. Alves; nahum@ufpa.br Phone: (+55) 91-32017999. Fax: (+55) 91-32011635

O-glycoprotein 2-acetamino-2-deoxy-β-D-glucopyranosidase (*O*-GlcNAcase) hydrolyzes 2-acetamido-2-deoxy-β-D-glucopyranose (O-GlcNAc) residues of serine/threonine residues of modified proteins. O-GlcNAc is present in many intracellular proteins and appears to have a role in the etiology of several diseases including cancer, Alzheimer's disease, and type II diabetes. In this work, we have carried out molecular dynamics (MD) simulations using a hybrid Quantum Mechanics / Molecular Mechanics (QM/MM) approach, to determine the binding of two potent inhibitors, PUGNAc and NAG, with a bacterial *O*-GlcNAcase. The results of these simulations show that Asp-401, Asp-298 and Asp-297 residues play an important role in the protein-inhibitor interactions. These results might be useful to design compounds with more interesting inhibitory activity on the basis of its three-dimensional structure.

Keywords: PUGNAc, NAG-thiazoline, quantum mechanical/molecular mechanical (QM/MM), molecular dynamics (MD), DFT/B3LYP/MM

Introduction

O-glycoprotein 2-acetamino-2-deoxy-β-D-glucopyranosidase (*O*-GlcNAcase) is a member of the family of glycoside hydrolases-84 (GH84)[1] that hydrolyzes 2-acetamido-2-deoxy-β-D-glucopyranose (O-GlcNAc) residues from posttranslationally modified serine/threonine residues of nucleocytoplasmic protein.[1-3] O-GlcNAc is abundant in the brain, particularly on cytoskeletal proteins,[4-5] which has supported hypothesis that disruptions of O-GlcNAc proteins may contribute to certain neurodegenerative disorders such as Parkinson dystonia[6] and Alzheimer's disease.[5] O-GlcNAc is also present in many transcription regulatory proteins. Mutations of many of these proteins contribute to the oncogenic phenotype. Some of these mutations may exert their effects in part by interference of O-GlcNAc-mediated regulation of these proteins. In this regard, recent studies have clearly established that the conversion of glucose into glucosamine, which is catalyzed by glutamine, is essential for the development of insulin resistance, drawing a link between O-GlcNAc misregulation and type II diabetes.[7-8]

O-GlcNAcase catalyzed reaction takes place through a two-step mechanism, forming a transient oxazoline intermediate (see structure **3** in Figure 1) that it is subsequently broken. Oxazoline formation is a substrate-assisted process, where the 2-acetomido group is activated by an Asp residue, while departure of the aglycone leaving group is assisted by means of a proton transfer from a protonated Asp (see the hypothetical transition structure **2** in Figure 1). Inhibitors of *O*-GlcNAcase activity have been developed trying to mimic the geometrical and/or electronic properties of structures appearing during the reaction process, either transition structures or intermediates. *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc, **4**) and 1,2-dideoxy-2'-methyl-α-D-glucopyranoso-[2,1-*d*]-Δ2'-thiazoline (NAG-thiazoline, **5**) are potent and the best characterized inhibitors of human *O*-GlcNAcase.[9] Recently, the crystal structure of a bacterial *Clostridium perfringens* homologue (CpGH84H) in complex with PUGNAc was reported,

with significant sequence homology to the human *O*-GlcNAcase.[10] Moreover, similarities between the inhibition constants of human *O*-GlcNAcase and CpGH84H were observed, making this bacterial enzyme a good model of the human enzyme.[10] The analysis of enzyme-PUGNAc complex suggests that this molecule inhibits *O*-GlcNAcase by mimicking the transition state of the *O*-GlcNAcase-catalyzed hydrolysis of N-acetylglucosaminide by virtue of its sp^2 anomeric C1, similar to the oxocarbenium ion-like transition state.[10] On the other hand NAG-thiazoline is, in principle, geometrically similar to the oxazoline intermediate. More recently, Whitworth and co-workers have shown that NAG-thiazoline is a transition state analogue for *O*-GlcNAcase, while PUGNAc is a poor transition state analogue.[11]

The aim of this work is to understand the molecular mechanism by which such molecules inhibit the catalytic activity of its target enzyme, which will be useful for designing potent inhibitors of *O*-GlcNAcase. Recently, we have successfully used molecular dynamics (MD) simulations, using a combined quantum mechanical/molecular mechanical (QM/MM) approach,[13-15] to study the relationship between protein-ligand interaction potential energies and HIV-1 IN inhibitor activity.[12] In the present work, electrostatic binding free energy of CpGH84H complexed with both PUGNAc and NAG-thiazoline have been computed. Nevertheless, the problem of the protonation state of the inhibitors and the ionisable key aminoacids of the enzyme active site has to be sort out in advance. Thus, as shown in Figure 1, two proton transfer process occur between the substrate and the enzyme (Asp-297 and Asp-298 in CpGH84H) and these can be more or less advanced in the transition state of the reaction. Correct molecular modelling of the inhibitors should consider different positioning of these protons either on the inhibitor or the enzyme side. Finally, a detailed analysis of the interactions of the two inhibitors mentioned above with the key residues inside the binding pocket has been carried out in order to understand the differences between these two families of inhibitors.

Computational Methods

The initial coordinates for the QM/MM MD calculations were taken from the trimeric crystal structure of CpGH84H-PUGNAc complex, with PDB code 2CBJ.[10] This structure has been also used as a starting point for NAG-thiazoline inhibitor, by changing the PUGNAc substrate molecule by the NAG-thiazoline. Since standard pK_a values of ionizable groups can be shifted by local protein environments,[16] an accurate assignment of the protonation states of all these residues at pH=7 was carried out by recalculating the standard pK_a values of the titratable amino acids using the "cluster method",[17] as implemented by Field and co-workers.[16b] According to this method, each titratable residue in the protein is perturbed by the electrostatic effect of the protein environment. We have also performed calculation of pKa values of aminoacids within the empirical *propKa* program of Jensen et al.[18] Results obtained with both methods are qualitatively in agreement, predicting the same protonation state for all residues at the selected pH of the simulations (see Supporting Information). Thus, according to these results, most residues were found at their standard protonation state, except Asp-298 residue that seems to be protonated at this pH. Interestingly, Asp-298 has been proposed to be the catalytic acid, protonating the glycosidic bond in the first step of the *O*-GlcNAcase-catalyzed hydrolysis of N-acetylglucosaminides, while Asp-297 would be a general base in a typical acid/base catalysis.[19] Therefore, in this work we have modelled the PUGNAc-enzyme complex protonating the Asp-298 (A form in Figure 2) or protonating the oxime nitrogen of PUGNAc (B form in Figure 2). This latter complex, B, would resemble a more advanced stage of the reaction process, with the leaving group being protonated. On the other side, in order to model NAG-thiazoline-enzyme complex, we must keep in mind that the pK_a of the nitrogen atom at C2 position of the substrate is expected to change during the reaction, reaching in the intermediate state a value close to that of aspartate.[19] A strong hydrogen bond is then expected to be found between these two moieties. Thus, we modelled the enzyme-inhibitor complex protonating Asp-298 and Asp-297 (C form in Figure 2) or protonating Asp-

298 and the nitrogen of thiazoline ring of the NAG-thiazoline (D form in Figure 2). Finally, for comparative purposes, we also considered an enzyme-inhibitor complex protonating only Asp-298, while no proton was considered either in the C2 nitrogen atom or in Asp-297 (E form in Figure 2).

After adding the hydrogen atoms to the structure, series of optimization algorithms (steepest descent conjugated gradient and L-BFGS-B[20]) were applied. In order to avoid a denaturalization of the protein structure, all the heavy atoms of the protein and the inhibitor were restrained by means of a Cartesian harmonic umbrella with a force constant of 1000 $kJ \cdot mol^{-1} \cdot Å^{-2}$. Afterwards the system was fully relaxed, but restraining the peptidic backbone with a lower constant of 100 $kJ \cdot mol^{-1} \cdot Å^{-2}$.

Then the optimized protein was placed in a cubic box of pre-equilibrated water molecules (80 Å side), using the principal axis of the protein-inhibitor complex as the geometrical center. Any water with an oxygen atom lying in a radius of 2.8 Å from a heavy atom of the protein was deleted. The remaining water molecules were then relaxed using optimization algorithms. Afterwards, 100 ps of hybrid QM/MM Langevin-Verlet molecular dynamics (MD) at 300K and in a canonical thermodynamic ensemble (NVT) were used to equilibrate de model. For the hybrid QM/MM MD calculations, only the atoms of the inhibitor were selected to be treated by QM, using a semiempirial AM1 hamiltonian.[21] The rest of the system (protein plus water molecules) were described using the OPLS-AA[22] and TIP3P[23] force fields, respectively, as implemented in the DYNAMO library.[24]

Due to the amount of degrees of freedom, any residue 20 Å apart from any of the atoms of the initial inhibitor was selected to be kept frozen in the remaining calculations. Cutoffs for the non-bonding interactions were applied using a switching scheme, within a range radius from 14.5 to 16 Å.

Once the system was pre-equilibrated, 1500ps of QM/MM MD were run at temperature of 300 K for each system. The computed RMSD for the protein during the last

400 ps renders a value always below 0.9 Å, while the RMS of the temperature along the different equilibration steps was always lower than 2.5 K and the variation coefficient of the potential energy during the dynamics simulations was never higher than 0.3%.

Finally, QM/MM optimizations were carried out from structures derived from the AM1/MM MD simulations, using AM1 hamiltonian and B3LYP functional[25] together with the 6-31G* basis set, to describe the QM region of the system.

The potential energy of our scheme is derived from the standard QM/MM formulation:

$$E_{QM/MM} = \left\langle \Psi \left| \hat{H}_o \right| \Psi \right\rangle + \left(\sum \left\langle \Psi \left| \frac{q_{MM}}{r_{e,MM}} \right| \Psi \right\rangle + \sum\sum \frac{Z_{QM}\, q_{MM}}{r_{QM,MM}} \right) + E^{vdW}_{QM/MM} + E_{MM} \qquad (1)$$

$$\mathsf{E}_{\mathsf{QM/MM}} = \mathsf{E}_{\mathsf{vac}} + \mathsf{E}^{\mathsf{elect}}_{\mathsf{QM/MM}} + \mathsf{E}^{\mathsf{vdW}}_{\mathsf{QM/MM}} + \mathsf{E}_{\mathsf{MM}} \qquad (2)$$

where E_{MM} is the energy of the MM subsystem terms, $E^{vdW}_{QM/MM}$ the van der Waals interaction energy between the QM and MM subsystems and $E^{elect}_{QM/MM}$ includes both the coulombic interaction of the QM nuclei (Z_{QM}) and the electrostatic interaction of the polarized electronic wave-function (Ψ) with the charges of the protein (q_{MM}). In order to take advantage of powerful optimizations algorithms (see for example ref 26), we have substituted the electrostatic interaction term by a pure coulombic one, in which the charges of the QM atoms are obtained by means of electrostatic potential fit methods based on the electronic density.[27] Thus the final potential energy can be approximated by:

$$\mathsf{E}_{\mathsf{QM/MM}} = \left\langle \Psi \left| \hat{\mathsf{H}}_{\mathsf{o}} \right| \Psi \right\rangle + \sum\sum \frac{\mathsf{q}^{\mathsf{ChelpG-fit}}_{\mathsf{QM}}\, \mathsf{q}_{\mathsf{MM}}}{\mathsf{r}_{\mathsf{QM,MM}}} + \mathsf{E}^{\mathsf{vdW}}_{\mathsf{QM/MM}} + \mathsf{E}_{\mathsf{MM}} \qquad (3)$$

The interaction energy between the inhibitor and the environment, computed by residue, is evaluated as the difference between the QM/MM energy and the energies of the separated, non-interacting, QM and MM subsystems with the same geometry. Considering that the MM

part is described using a non-polarizable potential the interaction energy contribution of each residue of the protein is given by the following expression:

$$E_{QM/MM}^{Int} = \sum\sum \frac{q_{QM}^{ChelpG-fit} q_{MM}}{r_{QM,MM}} + E_{QM/MM}^{vdW} \quad (4)$$

Equation 4 is valid since the polarization effect cannot be decomposed in individual contributions. The global polarization effect can be obtained from the gas phase energy difference between the polarized, Ψ, and unpolarized, Ψ^o, wave functions.

Finally, in order to evaluate the electrostatic binding free energy, series of QM/MM MD simulations have been carried out, introducing a λ parameter in the electrostatic QM/MM interaction:

$$E_{QM/MM} = \left\langle \Psi \middle| \hat{H}_o \middle| \Psi \right\rangle + \lambda \left(\sum \left\langle \Psi \middle| \frac{q_{MM}}{r_{e,MM}} \middle| \Psi \right\rangle + \sum\sum \frac{Z_{QM}\, q_{MM}}{r_{QM,MM}} \right) + E_{QM/MM}^{vdW} + E_{MM} \quad (5)$$

Then λ smoothly changes between values of 1, corresponding to a full MM charge-wave function interaction, and 0 whereas no electrostatic interaction with the force field is introduced. Note that the van der Waals interaction is always calculated.

The calculation of the free energy difference of two consecutive windows is performed by means of free energy perturbation (FEP) methods, which for the case of non-polarizable classical force fields reduces to:

$$\Delta F_{Elect-QM/MM} \approx -\frac{1}{\beta} \sum_i \ln \left\langle e^{-\beta \left[\lambda_{i+1} \left(\Sigma \left\langle \Psi \middle| \frac{q_{MM}}{r_{e,MM}} \middle| \Psi \right\rangle + \Sigma\Sigma \frac{Z_{QM}\, q_{MM}}{r_{QM,MM}} \right) - \lambda_i \left(\Sigma \left\langle \Psi \middle| \frac{q_{MM}}{r_{e,MM}} \middle| \Psi \right\rangle + \Sigma\Sigma \frac{Z_{QM}\, q_{MM}}{r_{QM,MM}} \right) \right]} \right\rangle_{\lambda_i} \quad (6)$$

We have used a total of 100 windows, from λ=0 (no electrostatic interaction) to λ=1 (full interaction), which turns into a δλ of 0.01. The total free energy term is then evaluated as the sum of all the windows. The procedure can be applied both forwards (i→i+1) and backwards (i+1→i) as long as the comparison provides information about the convergence of the

process. In each window a total of 5 ps of relaxation followed by 10 ps of production QM/MM MD have been performed. Since the use of high level Hamiltonians for performing hybrid MD is actually prohibitive, the electrostatic free energy has been evaluated using the semiempirical AM1 hamiltoninan.

Results and Discussion

Figure 2 shows the B3LYP/MM optimized structures obtained from the 1.5 ns AM1/MM MD simulations, while the contribution of individual residues to the total protein-substrate interaction energy is displayed in Figure 3. In this later figure, negative values correspond to stabilizing effects. From Figure 2 it can be shown that the principal structural difference between forms **A** and **B** in PUGNAc-enzyme complex is found in the orientation of residue Asp-298. In **A** form the exocyclic oxime nitrogen of PUGNAc interacts through hydrogen bond with Tyr-335 residue, while in **B** this substrate atom is forming a hydrogen bond with Asp-298 residue. X-ray crystallographic data[10] is, in general, very close to structure **B**, although the exocyclic oxime nitrogen of PUGNAc forms two hydrogen bonds with both Asp-298 and Tyr-335 residues in the experimentally determined structure. The strong interaction with Asp-298 residue favor the stabilization of the protein-inhibitor complex as can be observed in Figure 3, in accordance with the observation that the protonated nitrogen engaged in an adventitious electrostatic interaction with the general acid/base catalytic residue.[28] Furthermore, it was found that mutation of Asp298 to Asn reduces k_{cat} of enzyme catalyzed reaction by 8100-fold.[10] Other important residues are Asp-297 and Asp-401. Asp-297 presents hydrogen bond interactions with nitrogen of amide nitrogen, being this interaction responsible of stabilizing the conformation of acetamido group and, as observed in Figure 3, presenting a favourable interaction with the ligand. Regarding Asp-401, it presents favourable interactions with O4 and O6 hydroxyl oxygen atoms of the ligand. Interestingly, it is observed that mutation of Asp-401 has significant effects on both, the ability of the enzyme

to bind the inhibitor and catalysis (10-fold increase in K_m and reduces k_{cat} 2400-fold).[10] Tyr335, Trp394 and Trp490 seem to be stabilizing the cavity created by the three aforementioned residues. Finally, due to the change of substrate charge between both complexes, Lys184' (a residue belonging to a different protein chain) presents a electrostatically favourable interaction with the substrate in **A** complex, while it is unfavourable in **B**.

Regarding the NAG-thiazoline inhibitor, **C**, **D** and **E** complexes can be compared with previously experimentally determined *Bacteroides thetaiotaomicron* GH84 (BtGH84)[29] structure bound to a derivative of this inhibitor. Thus, X-ray crystallographic data[10] show that Asp-297 (Asp-242 in BtGH84) is positioned to form a hydrogen bond with the nitrogen of the thiazoline ring of the inhibitor, in accordance with **D** structure. In general, this complex is the closest one to the crystallographic structure and, therefore, only this structure will be discussed in this paper. Notably, PUGNAc and NAG-thiazoline pattern of interactions within the CpGH84H active site are very similar (compare **B** and **D** in Figure 3). Differences are observed in their interactions with Asp-297 and Asp298. While PUGNAc, in its **B** form, interacts strongly through Asp-298, NAG-thiazoline presents a larger stabilizing interaction with Asp-297 in the **D** form. These differences are consequence of the different pK_a's of the inhibitors. While oxime nitrogen atom of PUGNAc is able to accept a proton from Asp-298 and thus to establish a salt bridge with the protein, the smaller value of the pK_a of the 2-nitrogen atom in NAG-thiazoline makes the hydrogen bond with unprotonated Asp-297 stronger in this last case while Asp298 remains protonated. This result is in accordance with studies of Whitworth and co-workers, who reported a comparative study of substrate-inhibitor hydrogen bond interactions concluding that interactions between residues with closely matched pKa values are stronger than those between residues whose pKa values differ greatly.[11] Finally, last residue establishing a strong interaction with both inhibitors is Asp-401, that presents a similar orientation to the one observed in the BtGH84 crystal structure.[29b]

The comparison of three-dimensional molecular electrostatic potential (MEP) surfaces for compounds depicted in Figure 2 is depicted in Figure 4. MEPs surfaces were generated, with Gauss View program,[30] from the B3LYP/MM optimized structures. These surfaces correspond to an isodensity value of 0.002 a.u. The most nucleophilic regions (negative electronic potential) are shown in red, while and the most electrophilic regions (positive electrostatic potential) are shown in blue. The similarity of the MEP calculated for both inhibitors in their correct protonation states (**B** and **D** forms) is evident. Both cases display a large positive (blue) regions matching with the positions of Asp-297, Asp-298 and Asp-401. It is interesting to note that these positive potentials are a consequence either of the inhibitor itself (the NAG-thiazoline case) or as a consequence of a proton transfer from the protein (PUGNAc). These MEPs can be then used as templates for the design of new potent inhibitors. In addition this analysis can be used to get some clues about the transition state of the catalyzed reaction. One could conclude that in this transition state the proton transfer from Asp-298 should be completed while the transfer to Asp-297 is most probably in a much earlier stage. The principal difference between the two most efficient inhibitors, **B** and **D,** is in phenyl carbamate moiety, which is absent in NAG-thiazoline.

Electrostatic binding free energies obtained from the FEP calculations at the AM1/MM level are shown in Table 1, Analysis of the results reveals that NAG-thiazoline-enzyme **B** complex has a binding free energy which is close to the value calculated for the PUGNAc-enzyme **D** complex. The small difference observed between these two inhibitors (19 kJ/mol) is in agreement with the experimental, i. e., PUGNAc binds only 0.66-fold more tightly than NAG-thiazoline in human *O*-GlcNAcase.[9] The loss of ligand-protein interaction, mainly trough Asp-298 residue in NAG-thiazoline complex, may be responsible for the differences in the activity between these inhibitors. By comparison of calculated binding free energies and the experimental measurements of K_I, it can be observed a direct relationship between both properties, which could be used as a guide in the design of new inhibitors.

In summary, our results suggest that design of an effective inhibitor should resemble the TS of the enzyme catalyzed reaction and theoretical calculations can be used to guide its synthesis.

Acknowledgements

We thank DGI for project BQU2003-04168-C03, Universitat Jaume I - BANCAIXA Foundation for projects P1·1B2005-13, P1·1B2005-15 and P1·1B2005-27 and SEUI for financial support of a Hispano-Brasileño collaboration project (PHB2005-0091-PC). We are also grateful to Dr M. Field and Dr P. Amara for supporting us in the pKa calculations and discussions. The authors also acknowledge the Servei d´Informatica, Universitat Jaume I for generous allotment of computer time. Lameira would like to thank CAPES for their financial support and the warm hospitality during the research stay at Departament de Química Física i Analítica, Universitat Jaume I.

Supporting Information Available

Comparative analysis of pKa values of aminoacids within the "cluster method" *versus* the empirical *propKa* program. This information is available free of charge via the Internet at http://pubs.acs.org.

References

1. Torres, C. R.; Hart, G. W. *J. Biol. Chem.* **1984**, *259*, 3308-3317.
2. Wells, L.; Vosseller, K.; Hart, G. W. *Science* **2001**, *291*, 2376-2378.
3. Hanover, J. A. *FASEB J*. **2001**, *15*, 1865-1876.
4. Dong, D. L.; Xu, Z. S.; Hart, G. W.; Cleveland, D. W. *J. Biol. Chem.* **1996**, *271***,** 20845-20852.
5. Arnold, C.S.; Johnson, G.V.; Cole, R.N.; Dong, D. Lee, M.; Hart, G.W. *J. Biol. Chem.* **1996**, *271*, 28741-28744.
6. Shafi, R.; Iyer, S. P. N. L.; Ellies, G.; O'Donnell, N.; Marek, K. W.; Chui, D.; Hart, G. W.; Marth, J. D. *Proc. Natl. Acad. Sci.* **2000**, *97*, 5735-5739.
7. Akimoto, Y.; Hart, G.W.; Hirano, H.; Kawakami, H. *Med. Mol. Morphol.* **2005**, *38*, 84-91.
8. Vosseller, K.; Wells, L.; Lane, M. D.; Hart, G. W. *Proc. Natl. Acad. Sci*. **2002**, *99*, 5313-5318.
9. Macauley, M. S.; Whitworth, G. E.; Debowski, A. W.; Chin, D.; Vocadlo, D. J. *J. Biol. Chem.* **2005**, *280*, 25313-25322.
10. Rao, F.V.; Dorfmueller, H. C.; Villa, F.; Allwood, M.; Eggleston, I.; van Aalten, D. M. *The EMBO Journal* **2006**, *25*, 1569-1578.
11. Whitworth, G. E.; Macauley, M. S.; Stubbs, K. A.; Dennis, R. J.; Taylor, E. J.; Davies, G. J.; Greig, I. R.; Vocadlo, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 635-644.
12. Alves, C. N.; Martí, S.; Castillo, R.; Andrés, J.; Moliner, V.; Tuñón, I.; Silla, E. *Chem. Eur. J.* **2007**, *13*, 7715-7724.
13. Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227-249.
14. Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700-733.
15. Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298-305.
16. Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415-436.
17. Field, M.; David, L.; Rinaldo, D. *Personal Communication*
18. Hui, L.; Robertson, A. D.; Jensen, J. H. Proteins, **2005**, 61, 704-721.
19. Çetinbas, N.; Macauley, M. S.; Stubbs, K. A.; Drapala, R.; Vocadlo, D. J. *Biochemistry* **2006**, *45*, 3835-3844.
20. Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. *J. Sci. Comp*. 1995, 16, 1190-1208.
21. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.
22. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.
23. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.

24. (a) Field, M. J. *A practical Introduction to the Simulation of Molecular Systems*; Cambridge University Press: Cambridge, *U.K.* **1999**. (b) Field, M. J.; Albe, M.; Bret, C.; Proust-de Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088-1100.

25. (a) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098-3100. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785-789.

26. (a) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483-3492. (b) Marti, S.; Moliner, V.; Tunon, I. *J. Chem. Theory and Comput.* **2005**, *1*, 1008-1016.

27. Breneman, C.M.; Wiberg, K.B. *J. Comput. Chem*. **1990**, *11*, 361-373

28. Mark, B. L.; Vocadlo, D. J.; Zhao, D. L.; Knapp, S.; Withers, S. G.; James, M. N. G. *J. Mol. Biol.* **2001**, *276*, 42131-42137.

29. (a) Dennis, R. J.; Taylor, E. J.; Macauley, M. S.; Stubbs, K. A.; Turkenburg, J. P.; Hart, S. J.; Black, G.; Vocadlo, D. J.; Davies, G. J.; *Nat. Struct. Mol. Biol.* **2006**, 13, 365-371. (b) Shanmugasundaram, B.; Debowski, A.W.; Dennis, R.J.; Davies, G. J.; Vocadlo, D.J. *Chem. Commun.* **2006**, *42*, 4372- 4374.

30. GaussView, Version 3.0, Dennington II, R.; Keith, T.; Millam, J.; Eppinnett, K.; Hovell, W. L.; Gilliland, R.; Semichem Inc., Shawnee Mission, KS, **2003**.

Table 1: Binding free energies ($\Delta F_{Elect-QM/MM}$) obtained from the electrostatic FEP calculations at the AM1/MM level (in kJ·mol^{-1}). Experimental ligand-protein dissociation constants (K_I in μM, from ref 9) are reported for comparisons purposes.

Ligand	$\Delta F_{Elect-QM/MM}$	K_I
B	-348.4	0.07
D	-367.4	0.046

Figure captions:

Fig. 1: Schematic representation of the first reaction step in *O*-GlcNAcase from reactants (**1**) to the oxazoline intermediate (**3**) through an hypothetical transition state (**2**). Structures of PUGNAc (**4**) and NAG-thiazoline (**5**) inhibitors are also shown. Numbering of residues corresponds to bacteria Clostridium perfringens.

Fig. 2: B3LYP/MM optimized structures obtained after 1.5 ns of AM1/MM MD simulations for the following *O*-GlcNAcase-inhibitor structures: (A) PUGNAc-enzyme, protonating Asp-298 (B) PUGNAc-enzyme, protonating the oxime nitrogen of PUGNAc; (C) NAG-thiazoline-enzyme, protonating Asp-297 and Asp-298; (D) NAG-thiazoline-enzyme, protonating Asp-298 and the nitrogen of thiazoline ring of the NAG-thiazoline; (E) NAG-thiazoline-enzyme, protonating the Asp-298. Selected hydrogen bond distances in Angstroms.

Fig. 3: Contributions of individual amino acid residues to inhibitor binding energy (in $kcal \cdot mol^{-1}$) computed for the B3LYP/MM optimized structures obtained after 1.5 ns of AM1/MM MD simulations for the different *O*-GlcNAcase-inhibitor structures depicted in Figure 1.

Fig. 4: Electrostatic potential (MEP) surfaces derived from BLYP/6-31G*/MM calculations for the compounds shown in Figure 1. The increase of negative charges goes from the blue (positive) to red (negative).

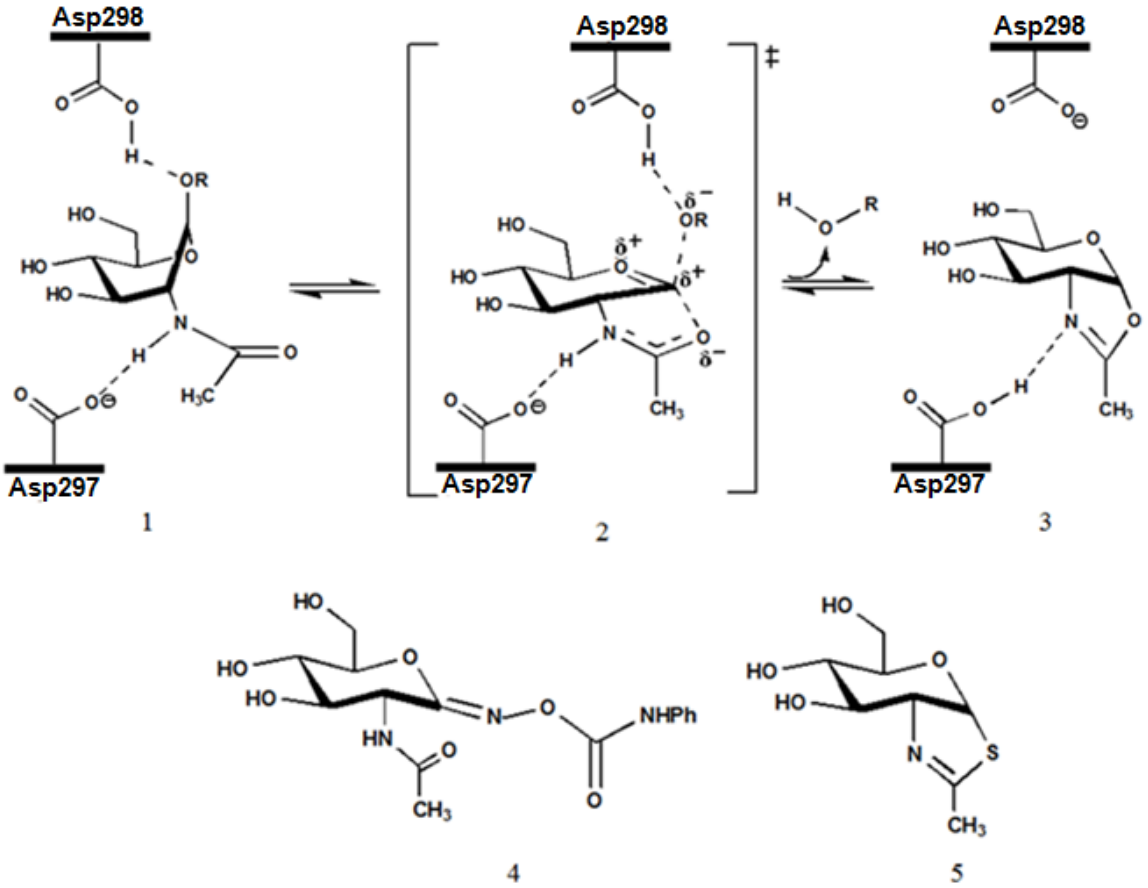
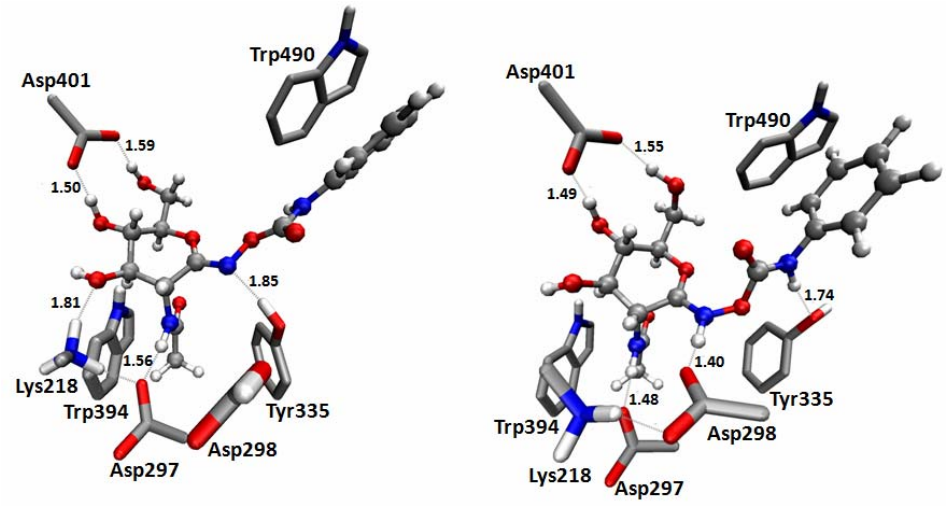
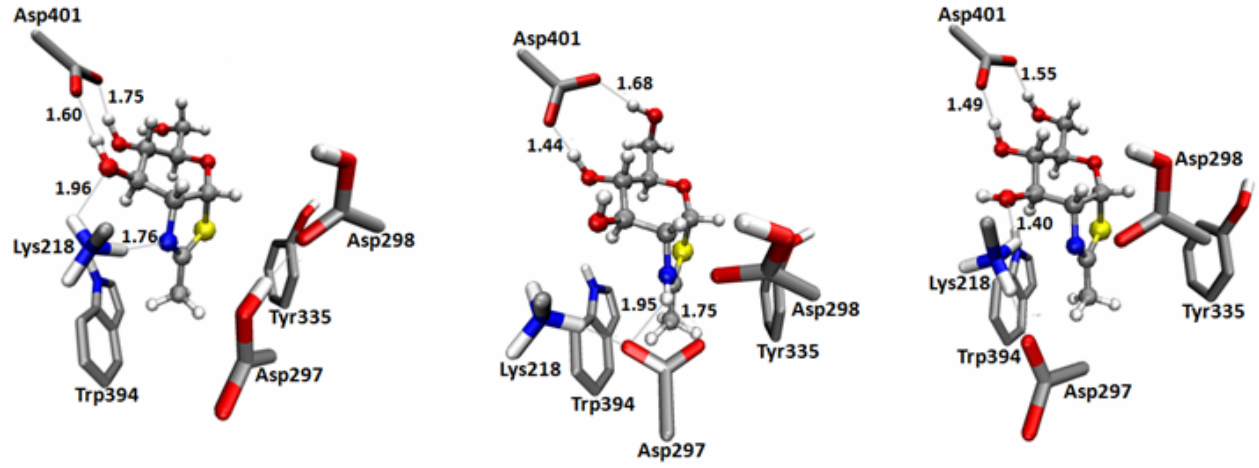


Figure 1



A **B**



C **D** **E**

Figure 2

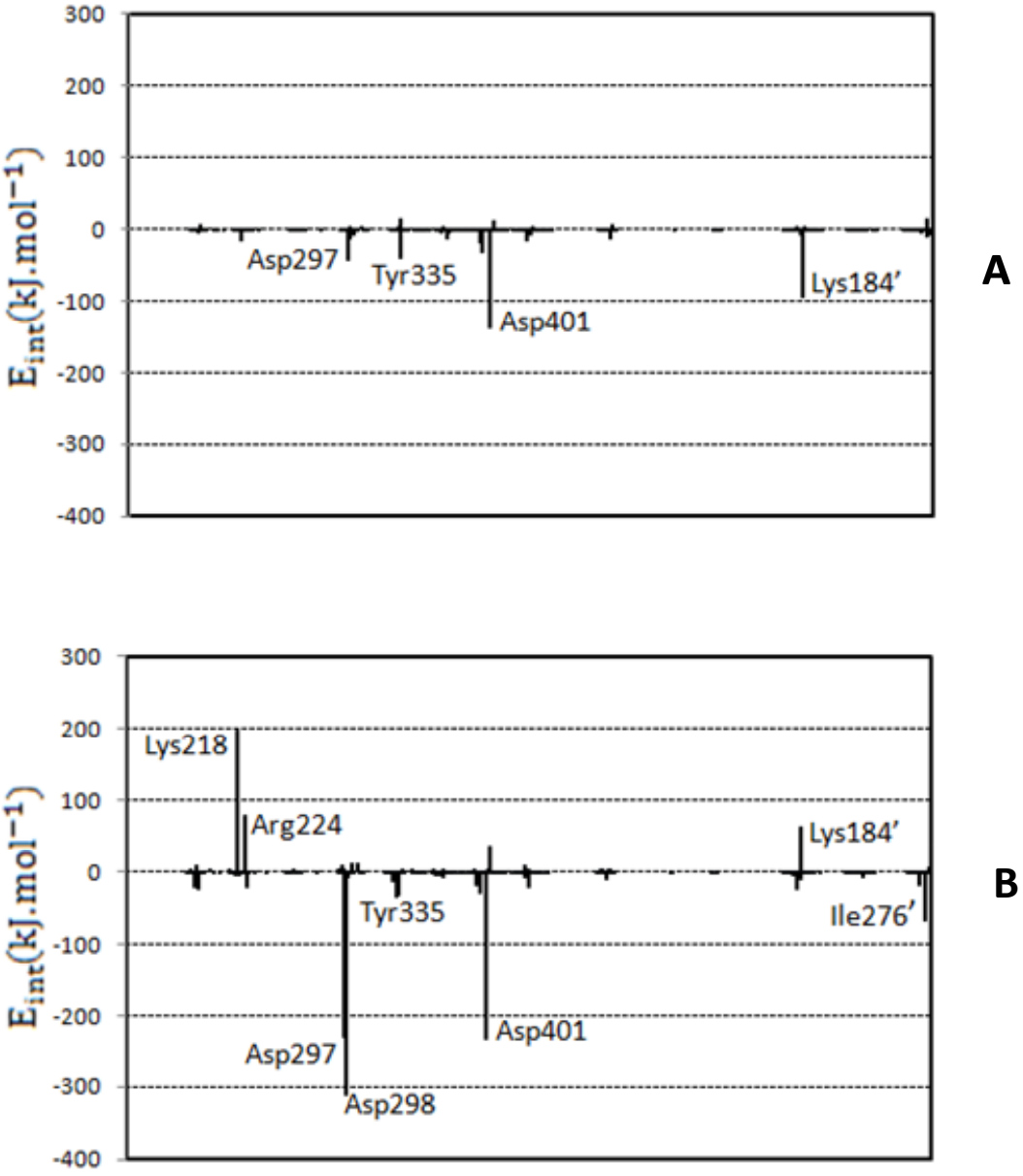

300
200
100
0
-100
-200
-300
-400
$E_{int}(kJ.mol^{-1})$
Asp297
Tyr335
Asp401
Lys184'
A
300
200
100
0
-100
-200
-300
-400
$E_{int}(kJ.mol^{-1})$
Lys218
Arg224
Tyr335
Asp297
Asp298
Asp401
Lys184'
Ile276'
B

Figure 3a

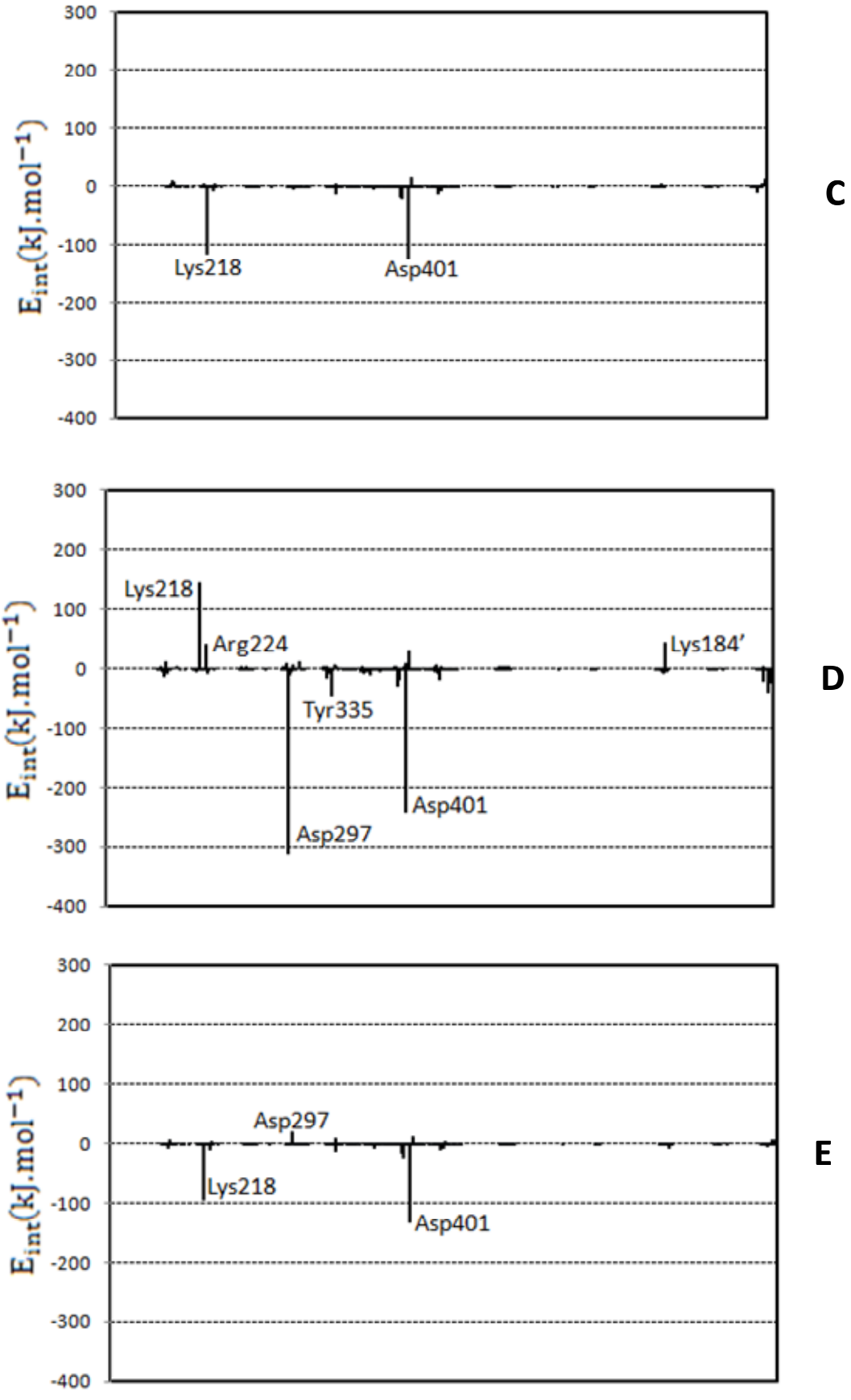

300
200
100
0
-100
-200
-300
-400
$E_{int}(kJ.mol^{-1})$
Lys218
Asp401
C
Lys218
Arg224
Tyr335
Asp297
Asp401
Lys184'
D
Asp297
Lys218
Asp401
E

Figure 3b

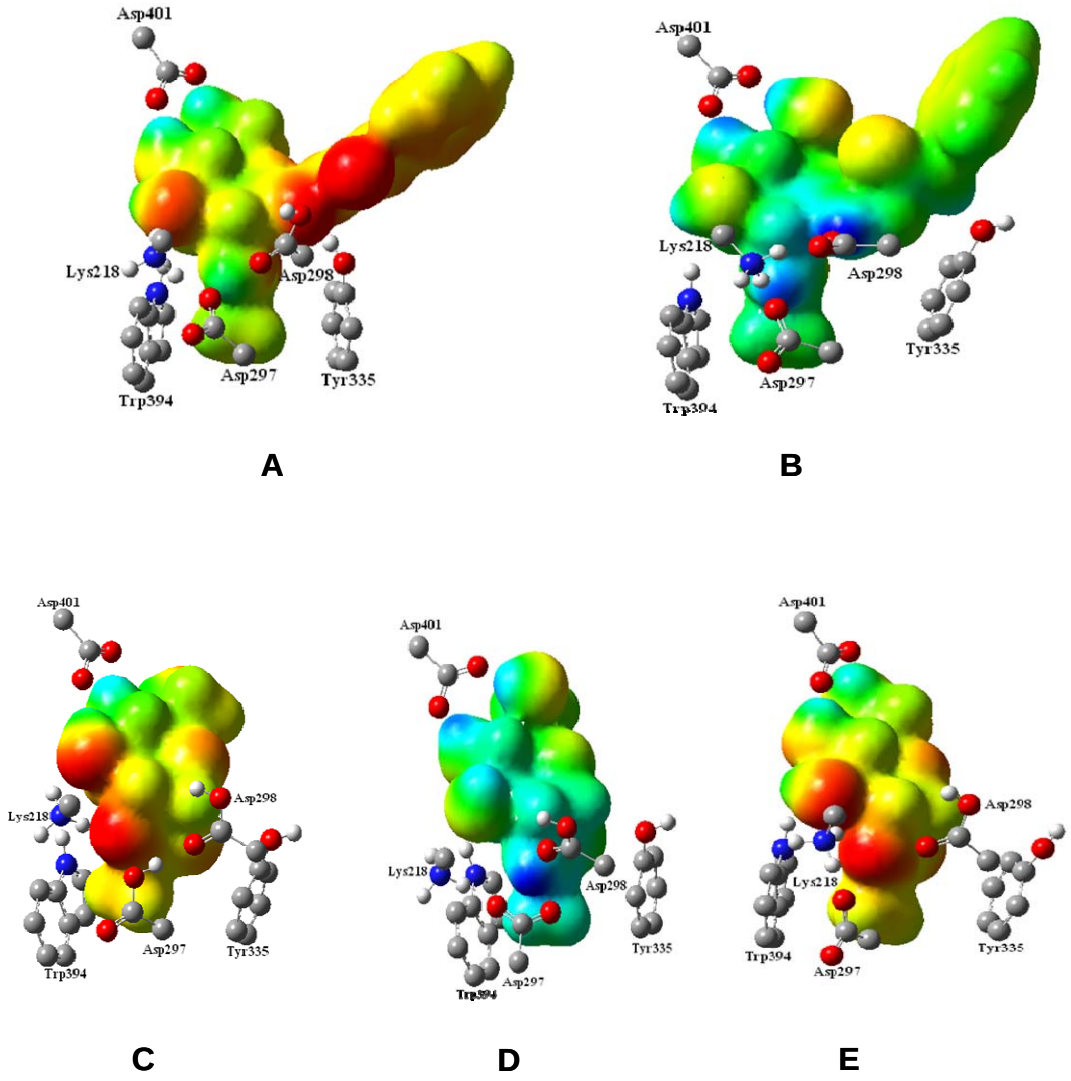

Asp401
Lys218
Asp298
Asp297
Tyr335
Trp394
A
Asp401
Lys218
Asp298
Tyr335
Asp297
Trp394
B
Asp401
Asp298
Lys218
Asp297
Tyr335
Trp394
C
Asp401
Lys218
Asp298
Tyr335
Asp297
Trp394
D
Asp401
Asp298
Lys218
Tyr335
Trp394
Asp297
E

Figure 4