

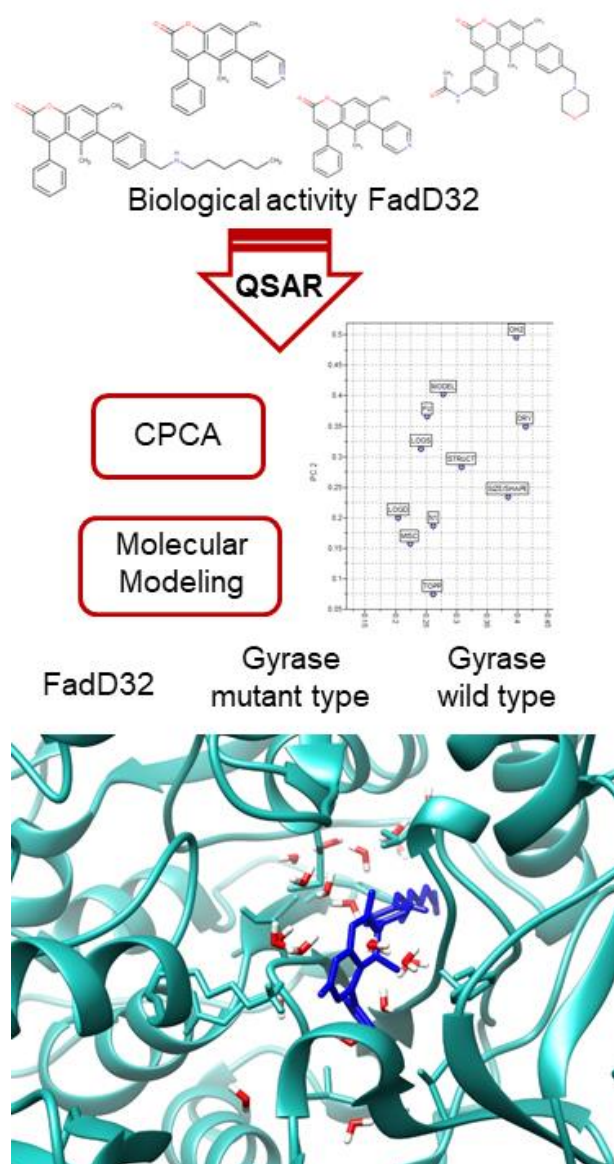
Chemometrics and Molecular Modeling applied to coumarin derivatives as potential multitarget inhibitors in *Mycobacterium tuberculosis*

Jéssika de Oliveira Viana ^a, Mayara Santos Maia ^a, Gabriela Cristina Soares Rodrigues ^a, Marcus Tullius Scotti ^a, Luciana Scotti ^{a,b}

^a Program of Natural and Synthetic Bioactive Products (PgPNSB), Health Sciences Center, Federal University of Paraíba, João Pessoa-PB, Brazil; E-mail: viana_jess@hotmail.com

^b Teaching and Research Management - University Hospital, Federal University of Paraíba, Campus I, 58051-900, João Pessoa-PB, Brazil

Graphical Abstract.



Abstract.

Tuberculosis (TB) is still a worldwide health problem caused, in large part, by *Mycobacterium tuberculosis* (MTB). The urgent need for the discovery of new antitubercular (anti-TB) agents has revealed the activity of coumarin derivatives against MTB. Therefore, a series of 36 coumarin derivatives was used to evaluate their antitubercular activity *in silico* against three proteins of *Mycobacterium tuberculosis*, known as FadD32, wild type and mutant DNA gyrase. The compounds had their energies minimized. Calculations of molecular descriptors were performed to carry out chemometric study of CPCA, docking and molecular dynamics. Our results showed that the major influence is the amphiphilic character, with predominance of the hydrophilic character, which corroborated with the docking results in the cavity of the active site for proteins. In homology studies it has been observed that mutation changes contribute to protein destabilization and may interfere with compound activity. The Score values and the interactions for the three proteins of the study were also favorable, evidencing the promising activity of coumarin derivatives anti-TB. We observed that the greatest divergence of interactions occurred for GyrA, whereas for GyrB the mutations did not influence the interaction of coumarins. In Molecular dynamics studies showed similar degree of instability the ligand PDB and coumarin, showing that certain residues are involved in the conformational alteration of the protein. Therefore, we conclude that the lead compound 2c showed potentiality of the coumarin derivative class as new candidate antituberculosis.

Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, is an infectious disease with high levels of mortality worldwide, currently with approximately 6.3 million new cases per year that often present resistance to both first- and second-line drugs [1]. These high rates of incidence are due to several factors including bacterial resistance, AIDS cases and latent tuberculosis that can reoccur in the patient.

All methods used in this search for new tuberculosis drugs were *in silico*. The CADD (computer-aided drug design) studies are increasingly being employed in industry and universities. Among the tools included in CADD is Molecular Docking, which is based on a computational technique that seeks to determine the interaction affinity of a ligand (a molecule) with the active site of its receptor (a macromolecular target). From this complex it is possible to infer characteristics related to the binding of the ligand within the active site of the receptor, in addition to inferring the best pose, that is, the best conformation and orientation of the ligand according to its energetic value [5].

In the studies conducted by Ananthan et al. (2009) [2], the authors identified through chemical high-throughput screen that coumarin derivatives showed anti-tubercular activity. Further studies have also demonstrated the use of these compounds as inhibitors of the fatty acyl ACP synthetase activity (FadD32) enzyme *in vitro*. In addition, in the molecular docking studies the compounds were potential inhibitors of the enzymes MtbFtsZ [3] and MtbDprE1 [4]. In addition to docking, studies involving molecular dynamics have become an attractive method widely employed in medicinal chemistry to recognize molecular information at the atomic level of a protein with its specific target on a computationally simulated scale of time and pressure [6].

Therefore, the aim of our study was to evaluate the potential multi-target character of the coumarins derivatives in two enzymes validated in the literature for *Mycobacterium tuberculosis*: FadD32 and DNA Gyrase wild and mutant using CPCA analysis, Docking and Molecular Dynamics.

Materials and Methods

Dataset and Optimization Geometry

From Kawate (2013) [7] we selected 36 coumarins derivatives, which had been described to inhibit the protein FadD32 (Fatty acid degradation protein D32) *in vitro* models of TB. The authors calculated the Inhibitory Concentration (IC₉₀) values, which inhibit specific biological population by 90%, on a micromolar scale (μM). All structures were drawn in HyperChem for Windows v. 8.0.5 [8] and the compounds had their molecular geometries minimized using the molecular mechanics MM+ force field and semi-empirical method AM1 (Austin Model 1) [9].

Molecular Descriptors

Three-dimensional (3-D) structures in .SDF format were used as input data in the Volsurf+ program v. 1.0.7 [10] and were subjected to molecular interaction fields (MIFs) to generate descriptors using the following probes: N1 (amide nitrogen–hydrogen-bond donor probe), O (carbonyl oxygen–hydrogen-bond acceptor probe), OH2 (water probe) and DRY (hydrophobic probe). Additional non-MIF-derived descriptors were generated to create a total of 128 descriptors. After calculation of the descriptors, the chemometric analysis of CPCA (Concensus Principal Component Analysis) the values of variance of the data were performed and observed.

Design of mutations in proteins

Three mutations in DNA GyrA (A90V, D94G and D89N) and three in DNA GyrB (D500N, D533A and E540D) have been described in the literature to make amino acid modifications of the two crystallized DNA Gyrase enzymes available in the RCSB Protein Data Bank database (PDB ID: 5BS8 and PDB ID: 2ZJT). The mutations were performed using the UCSF Chimera 1.12 software [11] with the Structure Editing option.

Docking

The structures of tree *M. tuberculosis* proteins: DNA Gyrase (PDB ID: 5BTC), Long-chain-fatty-acid-AMP Ligase FadD32 (PDB ID: 5HM3) and DNA Gyrase B subunit (PDB ID: 2ZJT) were downloaded from the Protein Data Bank (PDB). The docking was also performed for the mutant DNA gyrase, obtained by means of the homology model. All 36 structures of coumarins were submitted to molecular docking using the Molegro Virtual Docker v. 6.0.1 (MVD) [12]. All the water molecules were deleted from the enzyme structure. The enzymes and compounds structures were prepared using the same default parameter setting in the same software package. For all proteins docking was performed using the active site of the enzyme with the standard drug. In addition to the 36 coumarins, 3 compounds were used as controls, for comparative data, as Levofloxacin, Ciprofloxacin and, to DNA Gyrase B subunit, the compound Novobiocin. The MolDock score (GRID) algorithm was used as the score function [13]. The methodology was validated performing redocking of the ligand reported in PDB crystal structure for all *M. tuberculosis* proteins used in this study.

Molecular Dynamics Simulations

Molecular dynamics simulations were performed to estimate the stability of interactions between proteins and ligands using the Gromacs 5.0 software [14-15]. The topology of the ligands was prepared using the PRODRG topology generator (<http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrgrg/submit.html>) [16] applying the

GROMOS43a1 force field. The DM simulations were performed in 50.000 cycles at 1 ns. To determine the stability of the structure and if the complex is stable near the experimental framework, the mean square root displacement (RMSD) of all heavy atoms was calculated in relation to the starting structures. Residual Fluctuations (RMSF) were also analyzed to understand the role of residues near the receptor binding site.

Results and Discussion

CPCA analysis

The exploratory analysis, CPCA, was performed without any selection as to the test atoms or probes (probes) and descriptors available in the VolSurf + program. Thus, 13 blocks of descriptors were used, grouping 128 variables. We could observe the values obtained for the data variance and the first two main components explained more than 71% of the total variance, whereas only the OH2 block obtained 81% of data variance.

The blocks of the variables OH2 (hydrophilic regions), DRY (hydrophobic regions), MODEL (descriptors related to the molecular form) presented greater weight. The predominance of the hydrophilic characteristics can be observed by the contribution of the OH2 descriptors, which are calculated ratios between the hydrophilic surfaces and the total surfaces of the molecule. Therefore, we can conclude that the amphiphilic influence is positive for the antituberculosis activity of coumarin derivatives.

Design of mutations in proteins

M. tuberculosis DNA gyrase is resistant to quinolones, which are important bactericides used in the treatment of tuberculosis. In the search of new drugs for the treatment of tuberculosis, it is important to know the biology of the organism and its targets in order to obtain good responses to treatment without resistance. The exchange of amino acids with different chemical and structural properties may increase or decrease the affinity with the linkers. Substitutions of Alanine by Valine at position 90 of GyrA and glutamic acid by Aspartate at position 40 of GyrB may not undergo so many variations, since both have equal chemical properties with only subtle differences in their chain. Already the D94G, D89N, D500N and D533A substitutions may contribute to the stability or instability of receptor-ligand interaction, since hydrophobic and hydrophilic amino acid exchanges. Thus, stability or interaction instability will depend on the compatibility of the chemical properties of the linker in the mutation regions.

Docking

The docking analysis was performed to elucidated key interactions between ligand and receptor and to study the differences in the binding mode of FadD32 and DNA gyrase wild and mutant type. Docking of FadD32 showed interaction values very close to those of the PDB ligand, with -175kcal/mol for coumarin 3f and -194kcal/mol for the ligand PDB. The residue interaction analysis between the ligand and amino acids of FadD32 forming the active site could provide the quantitative explanation for the observed difference in binding affinity. The residue interaction analysis shows that the steric contacts were more prevalent over the hydrogen contribution in 2c, 3d and 3b, with essential interactions for activity observed with residues Ser321, Gly351 and Val397 in the region 5,7-dimethyl-2H-chromen-2-one, being the potential pharmacophore proposed for the coumarins antitb. The analysis of the hydrophobic region in the active site demonstrated the presence of the hydrophilic characteristic on its surface, corroborating with the CPCA studies. The molecular docking was validated by redocking of the native ligand back into the active site of FadD32 and DNA Gyrase through root-mean-square deviation (RMSD) between the pose of native ligand obtained by docking and the observed X-ray crystallographic conformation.

To DNA Gyrase A subunit, docking with the wild type revealed highest binding efficacy to coumarin derivatives, with values similar to the control drugs, such as Ciprofloxacin and Levofloxacin. However, with respect to DNA gyrase mutant subunit A, control drugs decreased their interactions affinities, while some coumarins demonstrated an increase in their interaction with the enzyme, presenting better scores (-45 kcal/mol to 3f) when compared to the control drugs (-7 kcal/mol to Levofloxacin). Derivatives 2c, 3f, 3d and 4c showed the best interactions among the coumarins in the study. Therefore, decrease in the docking score for the double mutants suggests poor affinity of LFX and CFX toward double mutants as compared to the wild type, thus hindering the process of blocking DNA replication. For GyrA, the presence of the Asn499 and Ser90 interactions were observed as key residues for the activity of these compounds, while the mutant GyrA showed the same interaction with Asn499, but a new interaction with Gly501.

In DNA Gyrase B subunit, the interactions were favorable for wild-type GyrB, with energy values close to Novobiocin, the reference drug. However, when submitted to mutant GyrB, coumarin derivatives presented better Scores, especially derivative 2c (-127 kcal/mol) when compared to the reference drug (-90 kcal/mol). To DNA Gyrase B subunit were found the essential interactions with the residues: Asp500, Phe486, Asp575 and Asp571, while in the mutant B subunit presented the same interactions presented for the wild type, demonstrating little influence in the cases of mutation reported for the subunit.

The residue interaction analysis as well as the score and the energy suggest that 2c interacts relatively strongly with enzyme FadD32 and wild and mutant DNA gyrase than other derivatives coumarins, with interactions very close to literature drugs, which agrees with the observed antitubercular activity.

Molecular Dynamics Simulations

The enzyme FadD32 increased its instability according to the increase of time, requiring greater mobility of the ligands. RMSD analysis also showed that the ligand pdb and coumarin have a similar degree of instability, showing higher interaction with the protein up to 0.3ns. In addition, RMSF calculations revealed that residues 20, 294 and 529 to 536 contribute more to the conformational change of the protein.

Conclusions

In this study were analyzed the Molecular docking of 36 coumarins derivatives in three enzymes, the FadD32 and DNA gyrase wild and mutant type, besides chemometric studies in the identification of the most predominant physicochemical characteristics of coumarin derivatives. In the homology studies it was observed that the mutation changes contribute to protein destabilization due to hydrophobic and hydrophilic amino acid exchanges, which may interfere in the interaction with the bioactive compounds. Among the coumarin derivatives, the compounds 2c and 3d were shown to have the best profiles against the reported characteristics, with good energy values and stability of binding with amino acid residues. In addition to its amphiphilic characteristic, to promote better connection with the active site of FadD32, corroborating with the activities of IC₉₀ reported in the literature. Therefore, the lead compound shows the potentiality of the class of coumarin derivatives as promising new candidates for the development of anti-tuberculosis drugs.

References

1. WHO, World Health Organization. Global Tuberculosis Report 2016. http://www.who.int/tb/publications/global_report/en/ (accessed July 15, 2017).
2. Ananthan, S., Faaleolea, E.R., Goldman, R.C., Hobrath, J.V., Kwong, C.D., Laughon, B.E, et al. High-throughput screening for inhibitors of *Mycobacterium tuberculosis* H37Rv. *Tuberculosis*, 89(5), 334-353, 2009.
3. Sridevi, D.; Sudhakar, K.U.; Ananthathatmula, R.; Nankar, R.P.; Doble, M. Mutation at G103 of MtbFtsZ Altered their Sensitivity to Coumarins. *Front. Microb.*, 8(578), e.12, 2017.
4. Shaikh, M.H.; Subhedar, D.D.; Shingate, B.B.; Khan, F.A.K.; Sangshetti, J.N.; Khedkar, V.M. et al. Synthesis, biological evaluation and molecular docking of novel coumarin incorporated triazoles as antitubercular, antioxidant and antimicrobial agents. *Med. Chem. Research*, 25(4), 790-804, 2016.
5. Yuriev, E.; Holien, J.; Ramsland, P.A. Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. *J. Mol Recognition*, 28(10), 581-604, 2015.
6. Perilla, J.R.; Goh, B.C.; Cassidy, C.K.; Liu, B.; Bernardi, R.C.; Rudack, T.; Yu, H.; Wu, Z.; Schulten, K. Molecular dynamics simulations of large macromolecular complexes. *Current Opinion In Structural Biology*, 31, 64-74, 2015.
7. Kawate, T.; Iwase, N.; Shimizu, M.; Stanley, S.A.; Wellington, S.; Kazyanskaya, E. et al. Synthesis and structure–activity relationships of phenyl-substituted coumarins with anti-tubercular activity that target FadD32. *Bioorg Med Chem Letters*, 23(22), 6052-6059, 2013.
8. Hyperchem Program Release 8.0.5 for Windows. Hypercube, Inc.: Gainesville, FL, USA, 2008.
9. Dewar, M.J.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J. Development and use of quantum mechanical molecular models. 76. AM1: a new general-purpose quantum mechanical molecular model. *J Amer Chem Soc*, 107(13), 3902-3909, 1985.
10. Cruciani, G.; Crivori, P.; Carrupt, P.A.; Testa, B. Molecular fields in quantitative structure–permeation relationships: the VolSurf approach. *J Mol Struct: Theo Chem*, 503(1), 17-30, 2000.
11. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C. et al. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J Comput Chem*, 25, 1605-1612, 2004.
12. Molegro Virtual Docker (MVD) for Windows. CLC Bio Company: Aarhus, Denmark, 2013.
13. Thomsen, R.; Christensen, M.H. MolDock: a new technique for high-accuracy molecular docking. *J. Med Chem*, 49(11), 3315-3321, 2006.
14. MU Abraham, T Murtola, R Schulz, S Páll, JC Smith, B Hess, et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software X*. 2015; 1-2: 19-25.
15. HJC Berendsen, D Van der Spoel, R Van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*. Vol 91, Issues 1-3, 1995.
16. A. W. Schüttelkopf and D. M. F. van Aalten. "PRODRG: a tool for high-throughput crystallography of protein-ligand complexes". *Acta Crystallogr. D60*, 1355–1363, 2004.