Structural bioinformatics

DDT - Drug Discovery Tool: a fast and intuitive graphics user interface for Docking and Molecular Dynamics analysis

Simone Aureli[1], Daniele Di Marino[1,2], Stefano Raniolo[1], Vittorio Limongelli[1,3*]

[1]Università della Svizzera italiana (USI), Faculty of Biomedical Sciences, Institute of Computational Science - Center for Computational Medicine in Cardiology, via G. Buffi 13, CH-6900 Lugano, Switzerland
[2]Polytechnic University of Marche, Department of Life and Environmental Sciences, Via Brecce Bianche, 60131 Ancona, Italy.
[3]Department of Pharmacy, University of Naples "Federico II", via D. Montesano 49, I-80131 Naples, Italy

*To whom correspondence should be addressed.

Abstract

Motivation: The ligand/protein binding interaction is typically investigated by docking and molecular dynamics (MD) simulations. In particular, docking-based virtual screening (VS) is used to select the best ligands from database of thousands of compounds, while MD calculations assess the energy stability of the ligand/protein binding complexes. Considering the broad use of these techniques, it is of great demand to have one single software that allows a combined and fast analysis of VS and MD results. With this in mind, we have developed the Drug Discovery Tool (DDT) that is an intuitive graphics user interface able to provide structural data and physico-chemical information on the ligand/protein interaction.
Results: DDT is designed as a plugin for the Visual Molecular Dynamics (VMD) software and is able to manage a large number of ligand/protein complexes obtained from AutoDock4 (AD4) docking calculations and MD simulations. DDT delivers 4 main outcomes: i) ligands ranking based on an energy score; ii) ligand ranking based on a ligands' conformation cluster analysis; iii) identification of the aminoacids forming the most occurrent interactions with the ligands; iv) plot of the ligands' center-of-mass coordinates in the cartesian space. The flexibility of the software allows saving the best ligand/protein complexes using a number of user-defined options.
Availability: *DDT_site_1* (alternative *DDT_site_2*); the DDT tutorial movie is available clicking here
Contact: vittoriolimongelli@gmail.com
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1. Introduction

Computer-aided drug design plays today a pivotal role in the identification of new drug candidates. Among the most successful approaches is the docking-based virtual screening (VS), which is used to predict the interaction between a ligand, typically small molecules, and its protein or nucleic acid target (hereafter protein for the sake of simplicity) (Morris, G.M. et al, 2009). The resulting ligand/protein complexes are evaluated and ranked according to a scoring function, which estimates the strength of the binding interaction. However, the approximation of the docking sampling produces many false positives, thus molecular dynamics calculations (MD) are usually performed to validate the docking results (D'Amore C et al, 2014; Di Leva FS et al, 2019). Despite the several software available for the analysis of the VS and MD results, none of them allows handling both the docking complexes and the corresponding MD trajectories. This operation requires indeed the use of different programs, resulting not user-friendly

and causing delay in the drug design pipeline. The Drug Discovery Tool (DDT) addresses this need allowing the even inexpert user to retrieve structural and energetics data on the ligand/protein complexes obtained from VS and MD calculations. DDT has a user-friendly graphics user interface (GUI) that is implemented in the worldwide used VMD software (Humphrey, W. et al, 1996).

2. Methods and Features

The DDT pipeline is shown in Figure 1 and is composed of two modules: the *VS package* (red) and the *MD package* (green).
VS package - DDT processes the AutoDock4 (AD4) docking results retrieving from the AD4 .dlg file the most occurred ligand binding mode, that is the largest cluster family of the ligand bound conformations (LC), and its binding energy score (ES). In addition, DDT easily allows the identification of the residues closer to the docked ligands (NA), showing as histograms their occurrence frequency in the docked poses. DDT also plots in the cartesian space the coordinates of the ligands' centre-of-mass
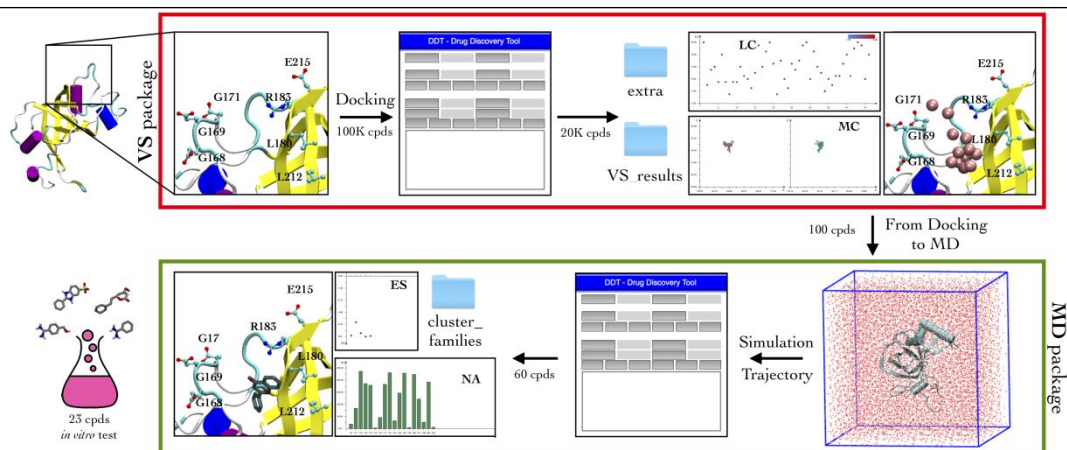
Fig 1. Schematic representation of the drug design campaign carried out targeting the TEL-patch of TPP1 and managed through DDT. *VS package (Upper)* - The analysis of ~100K docking calculations using DDT leads to the selection of the best 100 hits based on ligand conformation cluster analysis (LC), binding energy score (ES), location of the ligand in the TEL-patch (MC) and interactions engaged with the protein residues (NA). *MD package (Lower)* - The analysis of the atomistic MD simulations on the 100 best complexes with DDT leads to select the 60 ligands engaging long-lasting interactions with the TEL-patch residues (NA). From these, the 23 compounds with the most stable ligand binding modes (CLUSTER) and the best binding energy scores (ES) were selected for the *in vitro* tests.

for each docked ligand (MC), which can be displayed as spheres in the protein's binding site (SHOW MC). All these features make DDT a most useful tool in drug design, especially as explorative instrument in case in which the precise location of the binding site is unknown.

MD package - This module of DDT is designed for the post-processing analysis of trajectories of ligand/protein complexes obtained from MD simulations or related techniques (e.g., Monte Carlo, Funnel-Metadynamics). Similarly to the VS package, the residues closer to the ligand are identified from histograms accounting for their frequency of occurrence in the trajectory (NA); while the ligand's centre-of-mass coordinates along the trajectory are plotted in the cartesian space (MC) and displayed as spheres in the protein binding site (SHOW MC). DDT can also perform the clustering in families of the conformations assumed by the ligand along the trajectory (LC) and evaluate each cluster using an AD4-like ligand/protein binding energy score (ES). The latter is computed using the atoms' parameters reported in the GROMACS topology file. Therefore, topology files from different software (e.g., NAMD, AMBER) should be first converted in the GROMACS format before using ES. All the analysis can be performed on the whole trajectory or on a user-defined selection of the trajectory.

More details and a list of all the features of DDT can be found in the Supplementary Information as tutorial and manual document.

3. Use-case

We present in this section a compelling use-case of DDT in the identification of the first ligand-candidates of the shelterin protein TPP1. In this protein, 7 amino acids (E168, E169, E171, R180, L183, L212, E215), named TEL patch (Fig. 1), have been identified as responsible for the binding of TPP1 with telomerase and thus for the activation of the telomere maintenance mechanism (Nandakumar J. et al, 2012). For this reason, TPP1 represents a novel and promising anticancer molecular target. On this target we carried out a VS campaign followed by MD calculations to screen over 7 millions of compounds. We used DDT to analyse the results and select the best ligands for the pharmacological evaluation. The fact of using one single software (i.e., VMD) and the versatility of DDT with its several features, allow performing the analysis of the VS and MD results in a straightforward way, especially considering the large amount of structures and data analysed. Specifically, we used the VS package to analyse ~100k potential TPP1 ligands obtained from VS calculations, select those binding to the TEL-patch and those interacting with its 7 residues using the MC and NA

functions, respectively. The ~20k selected compounds were further filtered based on the frequency of occurrence of their binding conformation and their binding energy score, using the LC and ES values, respectively. Each of the resulting 100 complexes underwent 0.5µs MD simulations in order to assess their energy stability. The obtained trajectories were analysed using the DDT MD package. In detail, the NA function allowed identifying the ligands interacting for longer time with the TEL-patch residues. Then, the LC command was applied to the 60 selected trajectories to perform a ligand conformation cluster analysis, and finally each cluster family was scored based on the DDT ligand/protein binding energy score (ES). The 23 compounds with the lowest ES scores were finally proposed for the currently ongoing pharmacological tests.

4. Conclusions

DDT is an intuitive and user-friendly GUI to analyse the structural and energetics properties of ligand/protein complexes. The several user-defined options make DDT a versatile tool for expert and novel investigators in computer-aided drug design studies.

References

Abraham, M.J. *et al.* (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software-X*, **1**: 19-5.

D'Amore C. *et al.* (2014) Design, Synthesis, and Biological Evaluation of Potent Dual Agonists of Nuclear and Membrane Bile Acid Receptors. *J Med Chem*,**57**:937-954.

Di Leva F.S. *et al.* (2019) Discovery of ((1, 2, 4-oxadiazol-5-yl) pyrrolidin-3-yl) ureidyl derivatives as selective non-steroidal agonists of the G-protein coupled bile acid receptor-1, *Sci Rep*, **9**: 2504

Morris, G.M. *et al.* (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, **30**: 2785-2791.

Heyer L.J. K *et al.* (1999) Exploring Expression Data: Identification and Analysis of Coexpressed Genes. *Genome Res.*, **9**: 1106-1115.

Hill, D.A. and Reily, P.J. (2015) Scoring Functions for Autodock. *Methods in Molecular Biology*, **1273**: 467-474.

Humphrey, W. *et al.* (1996) VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, **14**: 33-38.

DDT - Drug Discovery Tool: a graphical user interface for Docking and MD fast analysis.

Nandakumar, J. et al. (2012) The TEL patch of telomere protein TPP1 mediates
    telomerase recruitment and processivity, *Nature*, **492**: 285-289