Opening mechanism of a DNA modyfying enzyme investigated by metadynamics

Blasco Morozzo della Rocca[a], Andrea Coletta [b], Federico Iacovelli[a], Alessandro Desideri[a]

[a] Structural Biology Lab, Biology Dept, Università di Roma TorVergata, Roma, 00133, Italy
[b] Department of Chemistry, Aarhus University, Aarhus C, 8000, Denmark
e-mail: desideri@uniroma2.it

Keywords: DNA topoisomerase, MD simulation, metadynamics, protein dynamics,

DNA topoisomerases are important enzymes that catalyze the breakage and rejoining of the DNA during the progression of cellular processes such as transcription, replication and recombination. Topoisomerases cleavage either one or both DNA strands and all have a catalytic mechanism which involves a nucleophilic attack of a DNA phosphodiester bond by a tyrosine residue. Human topoisomerase IB (hTopIB) is a monomeric enzyme, constituted by 765 residues that are divided into four domains (see Fig 1) arranged in two lobes. The enzyme binds double-stranded DNA as a protein clamp and transiently cleaves a single strand forming a 3'-phosphotyrosyl linkage between the active site tyrosine, Tyr723 located in the C-terminal domain, and the nicked DNA. Torsional strain in the DNA then drives the unwinding of the supercoiled strands and finally the free 5'-OH religates the phosphodiester backbone of the DNA rebuilding an intact strand. The 3D structures show that the protein is bi-lobed and clamps around the DNA substrate with the "Cat" (for catalytic) lobe and the other named "Cap" and lip1 and lip2 interact through a non-covalent interaction between the carboxylic group of Glu497 (lip2) and the amino group of Lys369 (lip1) (Figure 1).
The opening of the protein has never been investigated in the presence of DNA and this can provide useful information since it represents the last step of the catalytic process when the relaxed DNA substrate is released.
In this work the protein opening process has been investigated for the first time in presence of the DNA substrate by means of metadynamics.
Metadynamics is an enhanced sampling technique that enables the exploration of high-energy regions of the Free Energy Surface (FES) by progressively discouraging already visited conformations. This is achieved by the addition, to the total energy of the system, of a biasing history-dependent term of the form of a summation of Gaussian hills, centered on the explored point of the subspace identified by a set of user-defined collective variables (Cvs), so the system is slowly pushed away from the low energy regions of the conformational landscape. At the same time a reconstruction of the FES is obtained as sum of the added Gaussian hills. The well-tempered variant of metadynamics is an adaptive-biased method that, through a progressive reduction of Gaussian hills height, achieves a correct convergence of FES calculation [1].
In the simulation, two CVs have been used to describe the clamp opening:
a) the distance between the center of mass of Cα atoms of lip1 (residues 363-371) and of lip2 (residues 494-502); and b) the number of hydrogen bonds between the protein and DNA counted, as implemented in PLUMED 1.3 [2], via the coordination number function.
The protein structure of wild type hTopIB has been obtained as previously reported [3]. The hydrated strutcture was subsequently subjected to a 4 ns long NPT molecular dynamics simulation. This equilibration phase has been used to create a list ($list_{hb}$) of acceptor and donor pairs (protein and DNA) that are found, in at least 1 frame, at a distance lower to 6.0 Å. The hTopIB clamp around the DNA molecule has been destabilized by means of metadynamics using Gromacs-4.5.5 with the PLUMEDv1.3 patch, using Gaussian hills of height of 1.0 kJ/mole, and a width of 0.5 Å and 1.0 in the first and second CV dimension respectively, deposed every 250 frames (0.5 ps). The HB analysis have been separately performed for the conformations close to each local minimum, defined as the locus of CVs for which a grid search minimization path ends in the local minimum itself. The total trajectories have been divided in subtrajectories depending on the region (in the CVs subspace) on which each frame was positioned: the conformations with an energy difference lower than 20 kJ/mole have been assigned to the minimum and used for H-bond analysis. The RMSD-based clusterization has been carried out with the gromos method using a cutoff of 3.0 Å.
In the simulation the evolution of the collective variables as a function of time is shown (Figure 2) i.e. an increase of the lips distance followed by a decrease and then a final increase, corresponding to the definitive opening of the clamp. The number of the protein-DNA H-Bonds has a trend almost fully complementary to the lips distance, decreasing as soon as the lips distance increases.

The Free Energy surface representing the opening transition is L-shaped and contains about fifteen local minima. These can be reduced by smoothing algorithms to identify a fewer number of relevant macrostates representative of the transition. Then the minimum energy path connecting these states can be computed, along with the energy barrier between them (Figure 3).
The opening process can be described through the presence of four different states Figure 6A. In detail (Figure 3, bottom) the enzyme starts from a stable closed state with an energy of about -250 KJoule/mole (labelled C1), reaches a minimum still indicative of a closed state with an energy of -235 KJoule/mole (labelled C2), passes through a high energy transition state at -175 KJoule/mole (labelled Ts) and, finally goes back to an open state with an energy of -210 KJoule/mole (labelled O1).

In the final O1 conformation the "Cat" domain, containing the catalytic pentad, is not in contact with the DNA double helix and the remaining H-B between the protein and DNA occur at the interface between the "Cap" domain and the DNA double helix. These data are depicted in Figure 4 where the H-B persisting for more than 50% of the total frames present in each state are reported by a full line and those between 25% and 50% by a dashed line.
The metadynamics simulation aimed at understanding the mechanism of opening of the human topoisomerase Ib proteins can provide some explanation for the functional behaviour of the enzyme. The contacts with the "Cap" region in the open state is compatible with a non- catalytic diffusion of the enzyme on the one dimensional DNA strand, thereby allowing quick search for supercoil accumulation that could then trigger clamp closing and catalytic removal of torsional stress.

[1] Laio A. and gervasio F.L., Rep Prog Phys 71 (2008) 126601
[2] Bonomi M. et al, Comp Phys Comm 180 (2009) 1961-72.
[3] Chillemi G. et al, Nucl Acid Res 35 (2007) 3032-8.

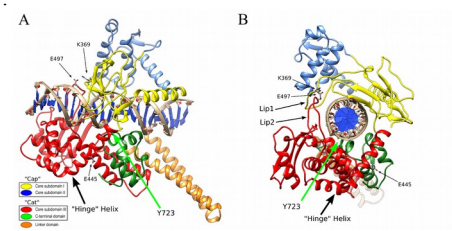

Figure 1.Structure of hTOP1 in complex with a 22-bp DNA. Orthogonal views are shown with subdomains in color and important residues highlighted.

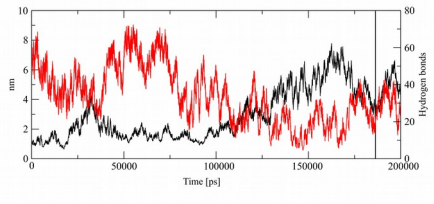

Figure 2. Time evolution of the collective variables used: distance betweeen the lips (in black, left Y axis) and coordination number, representing H bonds between protein and DNA (in red, right Y axis)

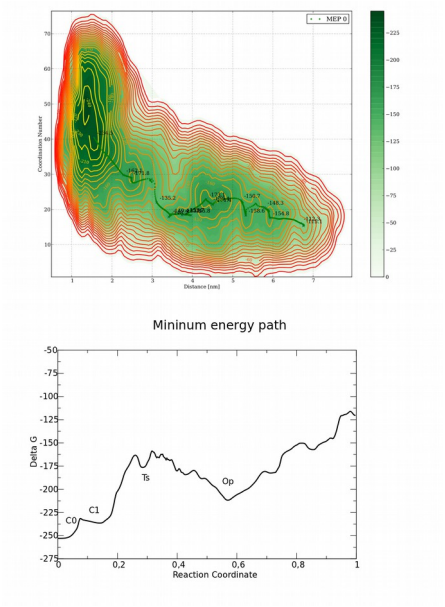

Figure 3. Top: Free energy surface reconstruction in coordination number and lips distance space. Energy values are color coded according to the scale on the right. The minimum energy path is also superimposed. Bottom: Representation of the minimum energy path as a function of reaction coorgdinate. Representative states, closed (C0 and C1), transition state (Ts) and the open state (Op) are marked.

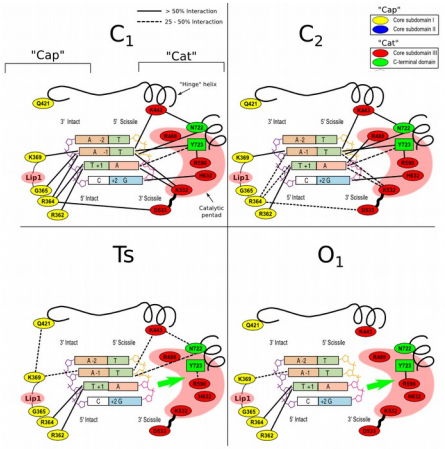

Figure 4. Schematic view of the interactions between protein residues and the DNA bases around the cleave site, in the four representative states. Residues are indicated by number and colored according to domain scheme of figure 1. Stable H-bonds are indicated with a solid line while more labile ones are dashed. The “hinge” helix is also shown.