



**HAL**  
open science

## **A mixed molecular modeling-robotics approach to investigate lipase large molecular motions**

Sophie Barbe, Juan Cortés, Thierry Simeon, Pierre Monsan, Magali Siméon,  
Isabelle André

### **► To cite this version:**

Sophie Barbe, Juan Cortés, Thierry Simeon, Pierre Monsan, Magali Siméon, et al.. A mixed molecular modeling-robotics approach to investigate lipase large molecular motions. *Proteins - Structure, Function and Bioinformatics*, 2011, 79 (8), pp.2517-2529. <10.1002/prot.23075>. <hal-02651861v2>

**HAL Id: hal-02651861**

**<https://laas.hal.science/hal-02651861v2>**

Submitted on 15 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# **A Mixed Molecular Modelling - Robotics Approach to Investigate Lipase Large Molecular Motions**

Sophie Barbe<sup>1-3</sup>, Juan Cortés<sup>1,4</sup>, Thierry Siméon<sup>1,4</sup>,  
Pierre Monsan<sup>1-3,5</sup>, Magali Remaud-Siméon<sup>1-3</sup> and Isabelle André<sup>1-3\*</sup>

<sup>1</sup> Université de Toulouse; INSA, UPS, INP, ISAE ,UT1, UTM; LAAS, LISBP; F-31077 Toulouse, France

<sup>2</sup> CNRS, UMR5504, F-31400 Toulouse, France

<sup>3</sup> INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

<sup>4</sup> CNRS ; LAAS ; 7 avenue du colonel Roche, F-31077 Toulouse Cedex 4, France

<sup>5</sup> Institut Universitaire de France, 103 boulevard Saint-Michel, F-75005 Paris

\*Correspondence to Isabelle André

Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés

INSA

135, avenue de Ranguueil

31077 Toulouse cedex 04

France

Tel: +33 (0) 561 559 963

Fax: +33 (0) 561 559 400

E-mail: [isabelle.andre@insa-toulouse.fr](mailto:isabelle.andre@insa-toulouse.fr)

**Short running title:** Mixed molecular modelling-robotics investigation of lipase molecular motions

**Keywords:** *Burkholderia cepacia* lipase; molecular dynamics; solvent environment; lid movement; protein flexibility; conformational transition; path planning; robotics

## ABSTRACT

Large-scale conformational rearrangement of a *lid* sub-domain is a key event in the interfacial activation of many lipases. We present herein a study in which the large-scale “open-to-closed” movement of *B. cepacia* lipase *lid* has been simulated at the atomic level using a hybrid computational method. The two-stage approach combines path-planning algorithms originating from robotics and molecular mechanics methods. In the first stage, a path-planning approach is used to compute continuous and geometrically feasible pathways between two protein conformational states. Then, an energy minimization procedure using classical molecular mechanics is applied to intermediate conformations in the path. The main advantage of such a combination of methods is that only geometrically feasible solutions are prompted for energy calculation in explicit solvent, which allows the atomic-scale description of the transition pathway between two extreme conformations of BCL (open and closed states) within very short computing times (a few hours on a desktop computer). Of interest, computed pathways enable the description of intermediate conformations along the “open-to-closed” conformational transition of BCL *lid* and the identification of bottlenecks during the *lid* closing. Furthermore, consideration of the solvent effect when computing the transition energy profiles provides valuable information regarding the feasibility and the spontaneity of the movement under the influence of the solvent environment. This new hybrid computational method turned out to be well-suited for investigating at an atomistic level large-scale conformational motion and at a qualitative level, the solvent effect on the energy profiles associated with the global motion.

## INTRODUCTION

The catalytic activity of many lipases (triacylglycerol acylhydrolases, EC 3.1.1.3) is known to be highly enhanced at an oil-water interface or in non-polar solvents, a phenomenon termed “interfacial activation”<sup>1</sup>. Although the molecular determinants involved in this phenomenon are not yet well understood, it has been reported to be closely related to large conformational changes of a mobile sub-domain of lipases<sup>2,3</sup>, called *lid*, which is located at the entrance of the active site. Rearrangements of the *lid* are generally assumed to be induced by the adsorption of the enzyme to a hydrophobic interface<sup>4</sup> or by the binding of a substrate<sup>5,6</sup>, or in some cases involving both events. The induced conformational changes lead to a “switch” between two extreme conformations: a closed state in which the *lid* shields the active site rendering the enzyme inactive and an open state with an exposed active site corresponding to a fully active enzyme. Insights on these conformational changes have been gained from the many X-ray structures of lipases reported in the last few years<sup>7-30</sup>. They have clearly shown that lipases may be found either in closed and/or in open conformation (stabilized or not by a covalently bound inhibitor or a detergent molecule). Although these crystallographic studies have enhanced our knowledge on the mobile sub-domains of lipases, the detailed description of the conformational transitions remains unclear. Other biophysical techniques are thus needed to complete X-ray data and capture the conformational dynamics of these enzymes. NMR relaxation experiments<sup>31-35</sup>, EPR spectroscopy<sup>36</sup> as well as single-molecule Fluorescence Resonance Energy Transfer (FRET) measurements<sup>37-40</sup> have allowed to capture the dynamics of large-scale conformational movements. However, to date, no experimental studies have been able to measure the *lid* conformational transition of lipases.

To elude the lack of experimental information, computational approaches have increasingly become a useful means to predict and understand conformational changes. A number of molecular dynamics (MD) simulations have been carried out to get more insight into the *lid* rearrangements of different lipases<sup>30,41-58</sup>. However, most of these simulations have been performed on short time scales and thus have only allowed modelling the dynamic properties of equilibrium states (closed or open). Indeed, simulation times needed to capture the entire conformational event, particularly with explicit simulation of solvents, are usually shorter, relative to the characteristic time of the conformational change. Therefore, alternative methods such as essential dynamics<sup>43,44,46,47,51,53</sup>, or normal mode analysis<sup>49</sup> have been successfully used to selectively enhance conformational sampling along specific directions of motions and identify large collective motions for a number of lipases. Although requiring intensive computing resources, MD simulations have also been performed on large time scales to attempt unravelling at the atomic level the entire conformational transition between the closed and open conformation of lipases<sup>30,42,54-58</sup>. In few of these studies, lipase conformational transitions were fast enough to be captured within 20ns MD simulations<sup>54-56,58</sup>. In particular, we have recently investigated using extensive MD simulations the influence of the environment on the conformational rearrangements of *Burkholderia cepacia* lipase (BCL) and its role on enzyme activation<sup>56</sup>. However, the expensive CPU cost of these classical molecular modelling simulations drastically impairs their use for routine exploration of large protein motions in solution.

To circumvent these limitations and increase our ability to reliably simulate molecular motions that can occur on large spatial and temporal scales, novel computational methodologies are needed. Herein, we used a methodology which separates the search of *lid* conformational transition pathways in two stages aiming to

speed up the computation. The first stage consists in an exploration of geometrically feasible motions of BCL *lid* using the robotics-based approach while the second stage uses molecular mechanics for an energy evaluation of solutions found at the previous stage, while taking into account explicit simulation of solvents. Compared to previous work {Cortes, 2005 #812}, the novelty of our study resides in the more extensive exploration of conformational domain transition and the consideration of solvent effect on the dynamics of lipases. The key advantage of the robotics-based approach is that it combines the efficacy of a geometrical treatment of the main molecular constraints (for example, steric clash avoidance, and structural constraints acting on the molecular chain model such as kinematic loop closure constraints) with the performance of path-planning algorithms<sup>59,60</sup> which enable fast exploration of high-dimensional conformational spaces. Such conformational search method is then well-adapted for handling large molecular motions in a continuous way and within very short computing times. In the recent years, robotics-based algorithms have been successfully applied for studying various molecular motions problems such as ligand docking<sup>61-65</sup> and accessibility pathways in flexible receptors<sup>66-70</sup>, protein and RNA folding pathways<sup>61,65,71-76</sup> or conformational changes of proteins<sup>63,68-70,77-81</sup> due to loop motions, domain motions and transmembrane  $\alpha$ -helices motions. In our study, the robotic-based approach is used to search routes devoid of geometric obstructions that connect the open and closed states of BCL in order to gain molecular knowledge about the structures of intermediate conformations during the “open-to-closed” transition and identify bottlenecks in the *lid* closing. Additionally, consideration of the solvent in the computation of the transition energetic profiles allows investigating the influence of the solvent on the protein conformational reorganization. Compared to prior MD simulation results<sup>56</sup>, our mixed robotics-molecular mechanics approach showed a gain in performance of several orders of magnitude to compute the large

amplitude motions of BCL and evaluate the feasibility and the spontaneity of the movement under the influence of the solvent environment.

## **METHODS**

### **Homology modelling**

The crystal structure of BCL refined at 2.0 Å resolution was taken as initial conformation for the simulations (PDB: 3LIP)<sup>23</sup>. As no crystal structure of BCL in closed conformation was available, a homology model was built using the protein modelling server SWISS-MODEL. The crystallographic structure of the homologous *Chromobacterium viscosum* lipase in closed form (PDB: 1CVL)<sup>8</sup> was used as a template. The quality of the homology model was analyzed using the ProSA-web interface<sup>82</sup> which determined a Z-score of -6.1 and a local model quality similar to that of the CVL template which showed a Z-core of -6.7. Noteworthy, the Z-scores of both conformations, the closed homology model of BCL and the closed X-ray structure of CVL, were slightly better than that of the open BCL crystal structure (PDB: 3LIP) which was of -4.9.

### **Geometric path-planning-based conformational exploration**

The method relies on a mechanistic all-atom model in which proteins are represented as articulated mechanisms, by analogy to a robot description, and allowing thus the use of path-planning algorithms to explore the high-dimensional conformational space<sup>68</sup>.

#### *Mechanistic molecular model*

Proteins are described using all-atom, hard sphere models forming a complex poly-articulated chain. Groups of rigidly bonded atoms form the bodies of the mechanism and the articulations between bodies correspond to the bond torsions

(bond length and angles are considered fixed). The atoms of BCL are represented by rigid spheres at 85 % of their van der Waals radii. Figure 1 shows the open X-ray structure of BCL and represents the articulated mechanical model on a detailed view of a residue. Using a geometric interpretation of the van der Waals repulsive force, the spheres associated with non-bonded atom pairs cannot overlap. Steric clash avoidance is handled by the BioCD collision detection algorithm<sup>83</sup>, which allows efficient detection of self-collision and distance computations in highly articulated molecular models. Kinematic loop closure<sup>69,77</sup> constraints are introduced to simulate interactions such as hydrogen bonds as well as to maintain the chain integrity between the different segments of the protein whose backbone is either mobile or rigid.

#### *Conformational space exploration algorithm*

The conformational search is derived from the Rapidly-exploring Random Trees (RRT) algorithm<sup>84</sup>. The basic principle of RRT is to incrementally grow a random tree rooted at the initial conformation  $q_{init}$  (the open Xray structure of BCL) to explore the reachable conformational space and find a feasible path connecting  $q_{init}$  to the  $q_{goal}$  (the closed homology model of BCL) (Fig. 2). At each iteration, the tree is expanded toward a randomly sampled conformation  $q_{rand}$ . The nearest node  $q_{near}$  in the tree to the sample  $q_{rand}$  is selected, and an attempt is made to expand  $q_{near}$  in the direction of the straight path to  $q_{rand}$ . If the expansion is feasible (i.e the motion satisfies all geometric constraints), it leads to the generation of a new node  $q_{new}$  and a feasible local path  $p_{new}$ . The expansion process is iterated until the current expanded node  $q_{new}$  can be connected to  $q_{goal}$ . The key feature of the RRT algorithm is that, thanks to a simple node sampling and selection strategy, the tree expansion is implicitly biased toward unexplored regions of the search space. In this work, we have applied a recent variant of the RRT algorithm, called ML-RRT (Manhattan-Like Rapidly-exploring Random Tree)<sup>85</sup>, which shows a higher efficiency for dealing with

high-dimensional problem by partitioning the conformational parameters in either *active* or *passive* sets of parameters and computes the motion of the elements associated with both parameter types in a decoupled manner. Active parameters are essential for the system motion, and they are directly treated at each iteration of the algorithm. Passive parameters are of secondary importance. Indeed, they are only handled when they hinder the motion of active parts or other passive parts identified as blocking parts during the tree expansion. In the present application, the active parameters correspond to the torsion angles of flexible BCL backbone segments (residues 125-169), while the passive parameters correspond to the torsion angles of BCL side-chains.

### **Pathway energy refinement in solvent**

Each of the ten geometrical pathways generated using the path planning algorithm was subsequently subjected to energy minimizations. Snapshots were taken every time the RMSD of the *lid* increased by 0.1 Å along the different motion pathways. The saved conformations were minimized using the Sander program with the all-atom ff03 force field of the AMBER9 software package<sup>86</sup> while taking into account either explicit water or explicit octane environment. For each conformation, the calcium ion, which plays a structural role in BCL, was conserved and appropriately parameterized according to prior work<sup>56</sup>. The system was embedded in a rectangular parallelepiped solvent box that left a space of 1 nm around the solute. For explicit water minimizations, TIP3P water molecules (approximately 8,000) were added using the LEaP module integrated in the AMBER9 package<sup>86</sup>. An octane box was created to carry out minimizations in explicit octane. To obtain appropriate atomic charges for use in the simulation, an *ab initio* calculation was carried out on octane molecule using the Jaguar software<sup>87</sup>. The Hartree-Fock calculation was run at

6-31G\* level and Mulliken atomic charges were computed. The *gaff* force field<sup>88</sup> was then used to parameterize octane molecules. The octane box was then subjected to 2000 energy minimization steps to remove any unfavorable contacts between octane molecules. The system was then equilibrated from 100 to 310 K under constant volume condition over 100 ps and then it was turned on constant pressure over 100 ps to adjust the system density. The weak-coupling method<sup>89</sup> was used to couple the system to a thermal bath of 310 K and a barostat of 1 bar with coupling constants of 0.2 ps. Finally, the protein was embedded in a box filled with approximately 800 octane molecules. The procedure for energy minimization in both environments consisted of : (i) one cycle (100 steps of steepest-descent algorithm and 400 steps of conjugate gradient algorithm) where atomic positions of solute were restrained using a harmonic potential; (ii) four cycles (50 steps of steepest-descent algorithm and 100 steps of conjugate gradient algorithm) where a force constant applied on heavy atoms of the protein was progressively diminished along procedure from 20 to 1 kcal mol<sup>-1</sup> Å<sup>-2</sup>; one cycle (50 steps of steepest-descent energy minimizations and then conjugate gradient energy minimizations until rms gradient was less than 0.1 kcal mol<sup>-1</sup>) of unrestrained minimization.

## RESULTS AND DISCUSSION

### Extreme states of the BCL conformational transition

The search for a transition path between two conformational states using path-planning algorithms requires the definition of an *initial* and a *goal* state, respectively the open and the closed conformation of BCL. In all BCL crystal structures determined to date<sup>15-17,23,24</sup>, both in the presence and in the absence of a bound

substrate-like inhibitor, the enzyme was found in an open conformation. These structures are highly similar as they show a low root mean square deviation of their C $\alpha$  atoms (RMSD  $\approx$  0.3-0.5 Å). Therefore, the BCL structure labelled 3LIP in the Protein DataBank<sup>23</sup> was chosen as the *initial* conformation to carry out the simulations. As no experimental information was available about the closed conformation of BCL, homology modelling was used to build a closed model of BCL using as template the *Chromobacterium viscosum* lipase (CVL), which shares 84% sequence identity with BCL. This lipase is found in a closed conformation in its crystallographic structure<sup>8</sup> (PDB:1CVL), with an  $\alpha$ -helical domain covering the active site.

Comparison of the closed homology model and the open X-ray structure of BCL revealed that the largest difference resides in residues 129-166, which encompass the two helices  $\alpha$ 5 and  $\alpha$ 6 of the U1 domain (residues 118-166) (Fig. 3A). Indeed, the largest positional shift is observed for the  $\alpha$ 5 helix, which gets displaced by as much as 20.5 Å with an average atomic displacement of about 13.5 Å between the two conformations. Additionally, the  $\alpha$ 5 helix appears four residues shorter in the closed conformation (residues 137-149) than in the open conformation (residues 134-150). Reversely, the  $\alpha$ 6 helix slightly changes its orientation and appears longer in the closed conformation (residues 156-160) compared to the open form (residues 160-166). The interconnecting loop between the  $\alpha$ 4 and  $\alpha$ 5 helices, formed by residues 128-133 in the open state and 128-136 in the closed state, is displaced by about 8 Å in the closed conformation. Some residues of the U2 domain (214-261) facing the *lid*, also undergo a positional shift in the closed homology model compared to the open crystal structure (Fig. 3B). However, these residues correspond to unresolved regions in the X-ray structure of CVL as well as to a BCL region of high crystallographic B-

factors. Some minor changes can also be seen in regions 17-26, 51-54 and 233-236 when comparing the closed and the open conformations of BCL (Fig. 3B).

Overall, the conformational changes between the closed homology model and the open X-ray conformation mostly involve the motion of the  $\alpha 5$  helix (Fig. 3A). By rolling back and forth this element, the active site is either shielded in the closed form (closed *lid*) or exposed in the open conformation (open *lid*), controlling thus the accessibility of the active site to substrate molecules and the activation of BCL. The mobile elements identified to be involved in the BCL conformational change are in agreement with results from MD simulations<sup>53,55,56</sup>. While short MD simulations carried out on BCL in explicit water had already shown high fluctuations for the region comprising the  $\alpha 5$  helix<sup>53</sup>, our more recent long-scale MD simulations<sup>56</sup> clearly demonstrated that the BCL conformational transition involves a large displacement of about 13 Å of the  $\alpha 5$  helix toward the  $\alpha 9$  helix of the so-called U2 domain which faces the *lid*.

In the current work, the conformational search for feasible transition pathways between the open and the closed conformations was carried out by considering the flexibility of all protein side-chains as well as the backbone of amino acid residues from 125 to 169 which overlap with the region identified to be mobile by comparison of the closed homology model and the open X-ray structure of BCL.

### **Description of *lid* transition pathways**

The hybrid methodology combining the path-planning algorithm ML-RRT followed by energy minimization was used to compute ten conformational transition paths between the open and closed states of BCL. Note that due of the randomness of the conformational exploration algorithm, different transition pathways (if they exist) can be identified from different runs. Each path was generated within a few CPU

hours on a single Intel Pentium4 processor (3.2GHz). The analysis of the ten paths generated by the robotics approach showed that all computed transition pathways are highly similar. For illustration, Figure 4A displays the RMSD variation for BCL *lid* region (residues 125-169) along each computed transition pathway. During the conformational transition, the RMSD of the 125-169 region gradually increased by about 4.5 Å in a similar way for all trajectories. The final BCL conformation obtained upon the ten conformational explorations were found highly similar to the closed homology model (RMSD < 1 Å) and the closed BCL conformation obtained after 20 ns of MD simulations in explicit water<sup>56</sup> (RMSD  $\approx$  1.7 Å). To compare in more details the ten conformational pathways, we have plotted in Figure 4B the displacement of the C $\alpha$  atoms for the *lid* region of each pair of consecutive conformations along each of the ten trajectories. The displacements undergone by each residue of the *lid* are comparable along the ten trajectories. Notably, some amino acid residues appeared to be shifted by over 10 Å during the conformational transition. Interestingly, these residues belong to two well-defined regions of the *lid*: the first region located between residues 134 and 146 corresponds to the  $\alpha$ 5 helix while the second region, situated between residues 154 and 158, corresponds to the interconnecting loop between the  $\alpha$ 5 and  $\alpha$ 6 helices. Residues S136, T137, V138, A141 and F142 located at the N-end of the  $\alpha$ 5 helix experienced the largest displacements during the conformational change (displacements > 14 Å) what could be associated to the partial unfolding at the N-cap of the  $\alpha$ 5 helix (Fig. 6). Indeed, the  $\alpha$ 5 helix is predicted to shorten during the *lid* closing on the basis of the predicted homology model for the closed BCL form and of our previous results issued from molecular dynamics simulations carried out on BCL in explicit water<sup>56</sup>. Within the *lid* region, the smallest variations of the motion amplitude were found for the  $\alpha$ 6 helix

and the interconnecting loop between the  $\alpha 4$  and  $\alpha 5$  helices. For loop residues 128 to 133, the maximal displacements (peaks in Fig. 4B) were found within the [1 – 10 Å] range. Their displacement amplitudes varied depending on the distance of the residue to the  $\alpha 5$  helix. Indeed, the largest variation was observed for G133, the closest residue to the  $\alpha 5$  helix, whereas the smallest displacement was seen for A128, the most distant residue. This indicates that residues of the  $\alpha 4$ - $\alpha 5$  interconnecting loop are dragged by the  $\alpha 5$  helix motion. Similar trend was observed for the maximal displacements of the  $\alpha 6$  residues (160-166) which were found to vary within the [1 – 10 Å] range. These differences among the residues forming the  $\alpha 6$  helix reflect the tilt of the  $\alpha 6$  helix toward the  $\alpha 4$  helix observed during the *lid* closure. Whereas the  $\alpha 5$  helix structure was found to get shorter during the open-to-close transition, the length of the  $\alpha 6$  helix was not altered along the trajectories computed by robotics, in consistency with the observation made from molecular dynamics simulations<sup>56</sup>. These results indicate that the folding of the  $\alpha 6$  helix N-cap observed in the homology model of the closed BCL is thus not crucial for a complete shielding of the active site by the *lid*.

Overall, residue displacements occurred in a continuous and simultaneous way for all ten *lid* closure trajectories (Fig. 4B). For illustration, Figure 5 shows the displacement of the C $\alpha$  atoms of *lid* residues for each iteration step along one transition pathway. In this plot, two main colored regions can be observed which correspond to the most mobile regions of the BCL *lid*. These regions correspond respectively to the  $\alpha 5$  helix (residues 134-146) and the interconnecting loop between the  $\alpha 5$  and  $\alpha 6$  helices (residues 154-158). Noteworthy, these movements appear to occur simultaneously at the same iteration numbers and they gradually increase from 1 Å at the beginning of the movement to about 18 Å at the end of the closure

transition. The closure movement of BCL *lid* appears thus to be concerted within one single step. The *lid* conformational reorganization appears thus driven by a continuous rolling movement of the  $\alpha 5$  helix, giving rise to smaller displacements of the regions surrounding the helical structure. During this motion, the N-ter of the  $\alpha 5$  helix comes closer to the  $\alpha 4$  helix while the rest of the  $\alpha 5$  helix moves toward the  $\alpha 9$  helix (Fig. 6). This rearrangement is accompanied by a  $\sim 30^\circ$  tilt of the  $\alpha 6$  helix toward the  $\alpha 4$  helix, consequently burying deeper the  $\alpha 6$  helix into the protein core. To illustrate the conformational changes occurring during the *lid* closure, several snapshots taken along the transition pathway are shown in Figures 6A-6E. The active site is seen to become progressively covered by the  $\alpha 5$  helix rendering it inaccessible to ligand binding. During this *lid* closing, hydrophobic residues from the  $\alpha 5$  helix, which were initially exposed (Fig. 7A), became more buried inside the protein to form a hydrophobic core with other hydrophobic residues from  $\alpha 4$  and  $\alpha 9$  helices (Fig. 7B). The formation of such a hydrophobic region at the interface of the three helices prevents the residues from interacting with the external environment. The closing mechanism that we have observed in the robotics-computed paths and the sequence of structural rearrangements of the secondary elements are closely related to the putative mechanism described in our earlier molecular dynamics study of BCL<sup>56</sup>.

The BioCD algorithm<sup>83</sup> for collision detection integrated in the software prototype BioMove3D<sup>68</sup> was used to identify collisions between atoms of distinct amino acid residues of the protein along each solution pathway and the corresponding inter-atomic distances found below 85% of the van der Waals equilibrium distance. The collision plot shown in Figure 8 reports the relative frequency of contacts for residues whose atoms were detected to be in contact with atoms from other residues of the protein during the transition. For instance, an inter-residue contact with a frequency of

100 % means that atoms of one given residue were at one moment in time very close to atoms of another residue of the protein in all ten computed pathways. These contacts can involve either atoms of the backbone and/or side-chain of the amino acid residues. The collision plot allows outlining key regions in the structure by identifying the residues with the highest number of contacts and/or strongest interactions during the BCL *lid* transition pathway. Therefore, such information holds great potential to underline residues that could play a key role in BCL *lid* closing/opening mechanism by gearing the movement or establishing interactions. On the diagonal of the plot are shown the contacts occurring between amino acid sequence neighbors, *i.e.* within the same structural element such as an  $\alpha$ -helix. Numerous high contact frequencies are found on the diagonal. In particular, six amino acids which belong to the  $\alpha$ 5 helix of the *lid* display a remarkably high frequency of contacts: S135, S136, V138, N144, V145 and F146. Notably, these residues, involved in inter-residue contacts, were also found to undergo large displacements during the *lid* closure (Fig. 4B). Interestingly, residue V138 was identified in prior molecular dynamics studies<sup>56</sup> to play a key role in the BCL conformational shift and *in silico* mutations of V138 were shown to completely impede the enzyme closure movement. Both Serines, 135 and 136, belong to the N-ter of the  $\alpha$ 5 helix which gets unfolded during the *lid* closing. The S135 backbone is involved in a hydrogen bonding interaction with Asp130 which was previously described as stabilizing the conformation of the interconnecting loop between  $\alpha$ 4 and  $\alpha$ 5 helices<sup>55</sup>. The side chains of V145 and F146 hydrophobic residues are also considered to be involved in the formation of the hydrophobic interface between the  $\alpha$ 5 helix of the *lid* and the  $\alpha$ 9 helix of the U2 sub-domain facing the *lid* (Fig. 7A-7B) and are thus likely to play a role in gearing the *lid* movement under the influence of the environment. The other residues observed on the diagonal which

display high frequency of contacts belong to BCL structure elements like the loop between  $\alpha 5$  and  $\alpha 6$  helices of the *lid* (S153/N154), the  $\alpha 9$  helix of the U2 domain facing the *lid* (T251/G265) or the loop between  $\beta 1$  and b1 (L17/T18). The side-chains of almost all these residues undergo conformational rearrangement during the course of *lid* transition pathway. Note that L17, one of both residues of BCL oxyanion hole, was previously shown as undergoing conformational change during *lid* transition pathway<sup>56</sup>. Near the diagonal are found contacts between residues of three helices of the *lid* :  $\alpha 4$  and  $\alpha 5$  helices or  $\alpha 4$  and  $\alpha 6$  helices. Most of these residues are involved in the formation of the hydrophobic patch between these helices. Of utmost interest are the inter-residue contacts appearing the farthest from the diagonal as these residues are remote in the amino acid sequence. The pairs of colliding residues identified on the graph are thus seen to be brought in close vicinity during BCL conformational transition. Analysis of the figure reveals that ten inter-residue contacts display an exceptionally high frequency of contact (over 60%, colored in dark green in Fig. 8). Most inter-residue contacts result from the protein inner rearrangements occurring during the *lid* closing and the packing of the U1 and U2 domains. For illustration, the amino acid residues detected in the collisions are represented in the BCL structure on Figure 8.

In order to investigate the influence of solvent on the energetic of the closing pathways, each of the ten robotics-computed trajectories was further refined using a classical molecular mechanics force field and considering distinct environments, either water or octane, in which BCL activity is respectively known to vary drastically. The minimizations performed in taking account solvent conditions did not lead to significantly different *lid* conformational rearrangements during the enzyme closing. The RMSD variation of the *lid* along the transition pathways after

minimization in explicit water and octane environment are similar to the RMSD variation determined before minimization. The largest measured differences between the RMSD before and after minimization were around 1.1 Å. However, the variation of the total potential energy followed opposite trends depending on the solvent (Fig. 9). In explicit water simulations, the potential energy gradually decreased during the closing of BCL while the reverse tendency was observed in explicit octane simulation as the energy was gradually increased during the closing of the protein. This behaviour is consistent with our previous MD simulations<sup>56</sup> which showed that the opening of BCL occurs in octane while the closing happens in aqueous media. Moreover, as in MD simulations, no large energy barrier was observed during transition pathways. The transition pathway between the two BCL forms may thus occur through a spontaneous motion of the *lid* under the sole effect of the solvent. These results are in agreement with the “enzyme model” which assumes that *lid* conformational rearrangements are induced by adsorption of the enzyme to an organic phase.

The energetic refinement of the pathways in explicit solvent led to qualitative data allowing a better understanding of the influence of solvent on large scale molecular motions. This approach provided similar results to our previous MD-based work but within a significantly reduced computing time. Simulations have only required few hours of computing time on a single processor what is remarkably short compared to MD CPU times.

## CONCLUSIONS

Conformational rearrangements between open and closed states of *B. cepacia* lipase have been investigated in this manuscript by means of a novel hybrid computational method with a fully atomistic description. This approach combines the use of both path planning techniques, highly efficient for the exploration of high dimensional conformational spaces and able to investigate geometrically feasible transition pathways between protein conformations, and classical molecular mechanics to evaluate pathway energetics under the influence of solvent. The advantage of these path-planning algorithms is that they allow a non-local exploration and the identification of continuous, geometrically feasible, transition paths. This geometry filtering enables discarding unauthorized geometries before energy refinement to accelerate the computation of energy favourable pathways. Although, the method outperforms the computing time performances of MD methods by several orders of magnitude, the trade-off is the rougher treatment of the molecular system during the geometrical exploration. Nonetheless, the determination of the open-to-closed atomistic pathways and the associated energy profiles in different environments led to similar conclusions of prior MD studies<sup>56</sup> regarding the favourable spontaneous closing movement of BCL in water which is shown to be disfavoured in octane in favour of a spontaneous opening mechanism. Analysis of pathways allowed us to pinpoint key structural regions involved in large conformational changes of the mobile sub-domain. This information led to a comprehensive understanding of the molecular determinants triggering the conformational transition undergone by the enzyme during its activation. Therefore, a mechanistic approach to molecular simulations combined to post-refinement using classical energy minimizations in explicit solvent allows access to useful data on

relevant large molecular motions for enzyme activity. Overall, this novel approach offers new ways to investigate, at atomistic level, relevant large scale enzyme molecular motions which can play a key role on biological processes.

## **ACKNOWLEDGMENTS**

Authors are grateful Igor Tvaroska (Institute of Chemistry, Slovak Academy of Sciences, Bratislava, Slovakia) for providing atomic charges for octane molecules derived from *ab initio* calculations. Authors wish also to thank the Computing center of Region Midi-Pyrénées (CALMIP, Toulouse, France) and the Center for Computing Resources (CRI) of INSA-Toulouse for providing calculation resources and support. This work was supported by the Institut des Technologies Avancées du Vivant (ITAV), Toulouse Canceropole campus.

## REFERENCES

1. Verger R. Interfacial activation of lipases: facts and artifacts. *Trends Biotechnol* 1997;15:32-38.
2. Louwrier A, Drtina GJ, Klibanov AM. On the issue of interfacial activation of lipase in non aqueous media. *Biotechnol Bioeng* 1996;50:1-5.
3. Ferrato F, Carriere F, Sarda L, Verger R. A critical reevaluation of the phenomenon of interfacial activation. *Methods Enzymol* 1997;286:327-347.
4. Derewenda ZS. Structure and function of lipases. *Adv Protein Chem* 1994;45:1-52.
5. Thuren T. A model for the molecular mechanism of interfacial activation of phospholipase A2 supporting the substrate theory. *FEBS Lett* 1988;229:95-99.
6. Brockman HL, Law JH, Kezdy FJ. Catalysis by adsorbed enzymes. The hydrolysis of tripropionin by pancreatic lipase adsorbed to siliconized glass beads. *J Biol Chem* 1973;248:4965-4970.
7. Noble ME, Cleasby A, Johnson LN, Egmond MR, Frenken LG. The crystal structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate. *FEBS Lett* 1993;331:123-128.
8. Lang D, Hofmann B, Haalck L, Hecht HJ, Spener F, Schmid RD, Schomburg D. Crystal structure of a bacterial lipase from *Chromobacterium viscosum* ATCC 6918 refined at 1.6 angstroms resolution. *J Mol Biol* 1996;259:704-717.
9. Mancheno JM, Pernas MA, Martinez MJ, Ochoa B, Rua ML, Hermoso JA. Structural insights into the lipase/esterase behavior in the *Candida rugosa* lipases family: crystal structure of the lipase 2 isoenzyme at 1.97Å resolution. *J Mol Biol* 2003;332:1059-1069.
10. Brady L, Brzozowski AM, Derewenda ZS, Dodson E, Dodson G, Tolley S, Turkenburg JP, Christiansen L, Høge-Jensen B, Nørskov L. A serine protease triad forms the catalytic centre of a triacylglycerol lipase. *Nature* 1990;343(6260):767-770.
11. Derewenda U, Swenson L, Wei Y, Green R, Kobos PM, Joerger R, Haas MJ, Derewenda ZS. Conformational lability of lipases observed in the absence of an oil-water interface: crystallographic studies of enzymes from the fungi *Humicola lanuginosa* and *Rhizopus delemar*. *J Lipid Res* 1994;35:524-534.
12. Schrag JD, Cygler M. 1.8 Å refined structure of the lipase from *Geotrichum candidum*. *J Mol Biol* 1993;230:575-591.
13. Derewenda U, Swenson L, Green R, Wei Y, Dodson GG, Yamaguchi S, Haas MJ, Derewenda ZS. An unusual buried polar cluster in a family of fungal lipases. *Nat Struct Biol* 1994;1:36-47.
14. Jung SK, Jeong DG, Lee MS, Lee JK, Kim HK, Ryu SE, Park BC, Kim JH, Kim SJ. Structural basis for the cold adaptation of psychrophilic M37 lipase from *Photobacterium lipolyticum*. *Proteins* 2008;71:476-484.
15. Luic M, Tomic S, Lescic I, Ljubovic E, Sepac D, Sunjic V, Vitale L, Saenger W, Kojic-Prodic B. Complex of *Burkholderia cepacia* lipase with transition state analogue of 1-phenoxy-2-acetoxybutane: biocatalytic, structural and modelling study. *Eur J Biochem* 2001;268:3964-3973.
16. Mezzetti A, Schrag JD, Cheong CS, Kazlauskas RJ. Mirror-image packing in enantiomer discrimination molecular basis for the enantioselectivity of

- B.cepacia* lipase toward 2-methyl-3-phenyl-1-propanol. Chem Biol 2005;12:427-437.
17. Lang DA, Mannesse ML, de Haas GH, Verheij HM, Dijkstra BW. Structural basis of the chiral selectivity of *Pseudomonas cepacia* lipase. Eur J Biochem 1998;254:333-340.
  18. Grochulski P, Bouthillier F, Kazlauskas RJ, Serreqi AN, Schrag JD, Ziomek E, Cygler M. Analogs of reaction intermediates identify a unique substrate binding site in *Candida rugosa* lipase. Biochemistry 1994;33:3494-3500.
  19. Bobrowicz P, Davidson RC, Li H, Potgieter TI, Nett JH, Hamilton SR, Stadheim TA, Miele RG, Bobrowicz B, Mitchell T, Rausch S, Renfer E, Wildt S. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. Glycobiology 2004;14:757-766.
  20. Nardini M, Lang DA, Liebeton K, Jaeger KE, Dijkstra BW. Crystal structure of *pseudomonas aeruginosa* lipase in the open conformation. The prototype for family I.1 of bacterial lipases. J Biol Chem 2000;275:31219-31225.
  21. Uppenberg J, Ohrner N, Norin M, Hult K, Kleywegt GJ, Patkar S, Waagen V, Anthonsen T, Jones TA. Crystallographic and molecular-modeling studies of lipase B from *Candida antarctica* reveal a stereospecificity pocket for secondary alcohols. Biochemistry 1995;34:16838-16851.
  22. Egloff MP, Marguet F, Buono G, Verger R, Cambillau C, van Tilbeurgh H. The 2.46 Å resolution structure of the pancreatic lipase-colipase complex inhibited by a C11 alkyl phosphonate. Biochemistry 1995;34:2751-2762.
  23. Schrag JD, Li Y, Cygler M, Lang D, Burgdorf T, Hecht HJ, Schmid R, Schomburg D, Rydel TJ, Oliver JD, Strickland LC, Dunaway CM, Larson SB, Day J, McPherson A. The open conformation of a *Pseudomonas* lipase. Structure 1997;5:187-202.
  24. Kim KK, Song HK, Shin DH, Hwang KY, Suh SW. The crystal structure of a triacylglycerol lipase from *Pseudomonas cepacia* reveals a highly open conformation in the absence of a bound inhibitor. Structure 1997;5:173-185.
  25. Grochulski P, Li Y, Schrag JD, Bouthillier F, Smith P, Harrison D, Rubin B, Cygler M. Insights into interfacial activation from an open structure of *Candida rugosa* lipase. J Biol Chem 1993;268:12843-12847.
  26. Uppenberg J, Hansen MT, Patkar S, Jones TA. The sequence, crystal structure determination and refinement of two crystal forms of lipase B from *Candida antarctica*. Structure 1994;2:293-308.
  27. Eydoux C, Spinelli S, Davis TL, Walker JR, Seitova A, Dhe-Paganon S, De Caro A, Cambillau C, Carriere F. Structure of Human Pancreatic Lipase-Related Protein 2 with the Lid in an Open Conformation(.). Biochemistry 2008;47:9553-9564.
  28. van Pouderoyen G, Eggert T, Jaeger KE, Dijkstra BW. The crystal structure of *Bacillus subtilis* lipase: a minimal alpha/beta hydrolase fold enzyme. J Mol Biol 2001;309:215-226.
  29. Carrasco-Lopez C, Godoy C, de Las Rivas B, Fernandez-Lorente G, Palomo JM, Guisan JM, Fernandez-Lafuente R, Martinez-Ripoll M, Hermoso JA. Activation of bacterial thermoalkalophilic lipases is spurred by dramatic structural rearrangements. J Biol Chem 2009;284:4365-4372.
  30. Bordes F, Barbe S, Escalier P, Mourey L, Andre I, Marty A, Tranier S. Exploring the conformational states and rearrangements of *Yarrowia lipolytica* Lipase. Biophys J;99:2225-2234.

31. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol* 2004;11:945-949.
32. Boehr DD, McElheny D, Dyson HJ, Wright PE. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 2006;313:1638-1642.
33. Palmer AG, 3rd. NMR characterization of the dynamics of biomacromolecules. *Chem Rev* 2004;104:3623-3640.
34. Cui Q, Karplus M. Catalysis and specificity in enzymes: a study of triosephosphate isomerase and comparison with methyl glyoxal synthase. *Adv Protein Chem* 2003;66:315-372.
35. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D. Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 2005;438:117-121.
36. Belle V, Fournel A, Woudstra M, Ranaldi S, Prieri F, Thome V, Currault J, Verger R, Guigliarelli B, Carriere F. Probing the opening of the pancreatic lipase lid using site-directed spin labeling and EPR spectroscopy. *Biochemistry* 2007;46:2205-2214.
37. Myong S, Stevens BC, Ha T. Bridging conformational dynamics and function using single-molecule spectroscopy. *Structure* 2006;14:633-643.
38. Rothwell PJ, Berger S, Kensch O, Felekyan S, Antonik M, Wohrl BM, Restle T, Goody RS, Seidel CA. Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. *Proc Natl Acad Sci U S A* 2003;100:1655-1660.
39. Schuler B, Lipman EA, Eaton WA. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* 2002;419:743-747.
40. Zhang Z, Rajagopalan PT, Selzer T, Benkovic SJ, Hammes GG. Single-molecule and transient kinetics investigation of the interaction of dihydrofolate reductase with NADPH and dihydrofolate. *Proc Natl Acad Sci U S A* 2004;101:2764-2769.
41. Peters GH, Svendsen A, Langberg H, Vind J, Patkar SA, Toxvaerd S, Kinnunen PK. Active serine involved in the stabilization of the active site loop in the *Humicola lanuginosa* lipase. *Biochemistry* 1998;37:12375-12383.
42. Jensen M, Jensen T, Kjaer K, Bjornholm T, Mouritsen O, Peters G. Orientation and conformation of a Lipase at an Interface Studied by molecular Dynamics Simulations. *Biophysical Journal* 2002;83:98-111.
43. Peters G, Bywater R. Computational analysis of chain flexibility and fluctuations in *Rhizomucor miehei* lipase. *Protein Eng* 1999;12:747-754.
44. Peters G, Bywater R. Influence of a lipid interface on protein dynamics in a fungal lipase. *Biophys J* 2001;81:3052-3065.
45. Peters GH, Olsen OH, Svendsen A, Wade RC. Theoretical investigation of the dynamics of the active site lid in *Rhizomucor miehei* lipase. *Biophys J* 1996;71:119-129.
46. Peters GH, van Aalten DM, Edholm O, Toxvaerd S, Bywater R. Dynamics of proteins in different solvent systems: analysis of essential motion in lipases. *Biophys J* 1996;71(5):2245-2255.
47. Peters GH, van Aalten DM, Svendsen A, Bywater R. Essential dynamics of lipase binding sites: the effect of inhibitors of different chain length. *Protein Eng* 1997;10:149-158.

48. Norin M, Haeffner F, Hult K, Edholm O. Molecular dynamics simulations of an enzyme surrounded by vacuum, water, or a hydrophobic solvent. *Biophys J* 1994;67:548-559.
49. Jaaskelainen S, Verma CS, Hubbard RE, Linko P, Caves LS. Conformational change in the activation of lipase: an analysis in terms of low-frequency normal modes. *Protein Sci* 1998;7:1359-1367.
50. James JJ, Lakshmi BS, Raviprasad V, Ananth MJ, Kanguane P, Gautam P. Insights from molecular dynamics simulations into pH-dependent enantioselective hydrolysis of ibuprofen esters by *Candida rugosa* lipase. *Protein Eng* 2003;16:1017-1024.
51. James JJ, Lakshmi BS, Seshasayee AS, Gautam P. Activation of *Candida rugosa* lipase at alkane-aqueous interfaces: a molecular dynamics study. *FEBS Lett* 2007;581:4377-4383.
52. Tejo BA, Salleh AB, Pleiss J. Structure and dynamics of *Candida rugosa* lipase: the role of organic solvent. *J Mol Model* 2004;10:358-366.
53. Lee J, Suh S, Shin S. Computational studies of Essential Dynamics of *Pseudomonas cepacia* Lipase. *Journal of Biomol Struct & Dyn* 2000;18:297-309.
54. Cherukuvada S, Seshasayee A, Raghunathan K, Anishetty S, Pennathur G. Evidence of a double-lid movement in *Pseudomonas aeruginosa* lipase: Insights from molecular dynamics simulations. *PLoS Comp Biol* 2005;1:182-189.
55. Trodler P, Schmid RD, Pleiss J. Modeling of solvent-dependent conformational transitions in *Burkholderia cepacia* lipase. *BMC Struct Biol* 2009;9:38.
56. Barbe S, Lafaquiere V, Guieysse D, Monsan P, Remaud-Simeon M, Andre I. Insights into lid movements of *Burkholderia cepacia* lipase inferred from molecular dynamics simulations. *Proteins* 2009;77:509-523.
57. Rehm S, Trodler P, Pleiss J. Solvent-induced lid opening in lipases: a molecular dynamics study. *Protein Sci* 2010;19:2122-2130.
58. Wang Y, Wei DQ, Wang JF. Molecular dynamics studies on T1 lipase: insight into a double-flap mechanism. *J Chem Inf Model* 2010;50:875-878.
59. Choset H, Lynch S, Hutchinson S, Kantor G, Burgard W, Kavraki L, Thrun S. Principles of Robot Motion: Theory, Algorithms, and Implementations. In: Press CM, editor; 2005.
60. LaValle S-M. Planning Algorithms. In: Press NYCU, editor; 2006.
61. Apaydin MS, Brutlag DL, Guestrin C, Hsu D, Latombe JC, Varma C. Stochastic roadmap simulation: an efficient representation and algorithm for analyzing molecular motion. *J Comput Biol* 2003;10:257-281.
62. Singh AP, Latombe JC, Brutlag DL. A motion planning approach to flexible ligand binding. *Proc Int Conf Intell Syst Mol Biol* 1999:252-261.
63. Bayazit OB, Song G, Amato NM. Ligand Binding with OBPRM and Haptic User Input. Proceedings of the 2001 IEEE International Conference on Robotics and Automation (ICRA) 2001:954-959.
64. Apaydin MS, Guestrin CE, Varma C, Brutlag DL, Latombe JC. Stochastic roadmap simulation for the study of ligand-protein interactions. *Bioinformatics* 2002;18 Suppl 2:S18-26.
65. Apaydin MS, Singh AP, Brutlag DL, Latombe JC. Capturing Molecular Energy Landscapes with Probabilistic Conformational Roadmaps. . Proceedings of the 2001 IEEE International Conference on Robotics and Automation (ICRA) 2001:932-939.

66. Guieysse D, Cortes J, Puech-Guenot S, Barbe S, Lafaquiere V, Monsan P, Simeon T, Andre I, Remaud-Simeon M. A structure-controlled investigation of lipase enantioselectivity by a path-planning approach. *Chembiochem* 2008;9:1308-1317.
67. Lafaquiere V, Barbe S, Puech-Guenot S, Guieysse D, Cortes J, Monsan P, Simeon T, Andre I, Remaud-Simeon M. Control of lipase enantioselectivity by engineering the substrate binding site and access channel. *Chembiochem* 2009;10:2760-2771.
68. Cortes J, Simeon T, Ruiz de Angulo V, Guieysse D, Remaud-Simeon M, Tran V. A path planning approach for computing large-amplitude motions of flexible molecules. *Bioinformatics* 2005;21:i116-125.
69. Cortes J, Le DT, Iehl R, Simeon T. Simulating ligand-induced conformational changes in proteins using a mechanical disassembly method. *Phys Chem Chem Phys* 2010;12:8268-8276.
70. Cortes J, Barbe S, Erard M, Simeon T. Encoding Molecular Motions in Voxel Maps. *IEEE/ACM Trans Comput Biol Bioinform* 2011;8(2):557-563.
71. Amato NM, Dill KA, Song G. Using motion planning to map protein folding landscapes and analyze folding kinetics of known native structures. *J Comput Biol* 2003;10:239-255.
72. Tang X, Kirkpatrick B, Thomas S, Song G, Amato NM. Using motion planning to study RNA folding kinetics. *J Comput Biol* 2005;12:862-881.
73. Amato NM, Song G. Using motion planning to study protein folding pathways. *J Comput Biol* 2002;9(2):149-168.
74. Chiang TH, Apaydin MS, Brutlag DL, Hsu D, Latombe JC. Using stochastic roadmap simulation to predict experimental quantities in protein folding kinetics: folding rates and phi-values. *J Comput Biol* 2007;14(5):578-593.
75. Thomas S, Song G, Amato NM. Protein folding by motion planning. *Phys Biol* 2005;2(4):S148-155.
76. Tapia L, Tang X, Thomas S, Amato NM. Kinetics analysis methods for approximate folding landscapes. *Bioinformatics* 2007;23(13):i539-548.
77. Cortes J, Simeon T, Remaud-Simeon M, Tran V. Geometric algorithms for the conformational analysis of long protein loops. *J Comput Chem* 2004;25:956-967.
78. Kirillova S, Cortes J, Stefaniu A, Simeon T. An NMA-guided path planning approach for computing large-amplitude conformational changes in proteins. *Proteins* 2008;70:131-143.
79. Enosh A, Fleishman SJ, Ben-Tal N, Halperin D. Prediction and simulation of motion in pairs of transmembrane alpha-helices. *Bioinformatics* 2007;23:e212-218.
80. Thomas S, Tang X, Tapia L, Amato NM. Simulating protein motions with rigidity analysis. *J Comput Biol* 2007;14(6):839-855.
81. Tapia L, Thomas S, Amato NM. A motion Planning Approach to Studying Molecular Motions. *Communications in information and Systems* 2010;10(1):53-68.
82. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007;35(Web Server issue):W407-410.
83. Ruiz de Angulo V, Cortés J, Siméon T. BioCD : An efficient algorithm for self-collision and distance computation between highly articulated molecular models. *Robotics: Science and Systems* 2005.

84. LaValle SM, Kuffner JJ. Rapidly-exploring random trees: progress and prospects. Eds. Boston: A.K. Peters: Donald, B. Lynch, K. Rus, D.; 2001.
85. Cortes J, Jaillet L, Simeon T. Disassembly path planning for complex articulated objects. *IEEE Transactions on Robotics* 2008;24:475-481.
86. Case DA, Darden TE, Cheatham ITE, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA. AMBER 9, University of California, San Francisco. 2006.
87. Jaguar software, Schrödinger, Portland, US.
88. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general Amber force field. *J Comput Chem* 2004;25:1157-1174.
89. Berendsen HJ, Postma JP, Van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamic with coupling to an external bath. *J Chem Phys* 1984:3684-3690.

## FIGURE CAPTIONS

**Figure 1.** Representation of the mechanistic molecular model. Proteins are modelled as articulated mechanisms. Bonded atom groups form the bodies and the articulations correspond to bond torsions. Rotations between backbone atom groups are  $\Phi$ ,  $\Psi$  and  $\omega$ , and  $\gamma_1$  and  $\gamma_2$  for the side chains. Hydrogen atoms have been omitted for clarity purpose.

**Figure 2.** Illustration of one expansion step of a search tree using a RRT-based algorithm. The tree tends to cover  $CS_{\text{feasible}}$ : the feasible subset of the explored space.

**Figure 3.** Extreme conformations of the BCL conformational transition.

(A) Backbone superposition of the open X-ray structure and the closed homology model of BCL. The two conformations are shown in a cartoon representation. The U1 domain (residues 118-166) of the open crystal structure of BCL (PDB: 3LIP) is coloured in red and that of the closed homology model in orange. The catalytic triad (S87, D264 and H286) and amino acid residues involved in oxyanion hole stabilization (L17 and Q88) are shown in yellow stick. To illustrate the large displacement of the  $\alpha_5$  helix between both conformations, the distance between  $C\alpha$  atom of the residue T137 in open and closed structures is shown. The displacement of the interconnecting loop between  $\alpha_4$  and  $\alpha_5$  helices is displayed by the distance between  $C\alpha$  atom of the residue T132 in open and closed structures.

(B) Plot of the positional difference of  $C\alpha$  atoms between the open X-ray structure and the closed homology model versus BCL residue number. The secondary structure of open BCL is shown on the graph for reference.

**Figure 4.** (A) RMSD variation of BCL *lid* region (residues 125-169) along the ten transition pathways computed using the path-planning approach. (B) Displacement of the  $C\alpha$  atoms for the *lid* region of each pair of consecutive conformations along each of the ten trajectories. The secondary structure of open BCL is shown on the graph for reference.

**Figure 5.** Displacement of BCL *lid* residues along one conformational transition pathway. The plot shows the displacement of the residue  $C\alpha$  atoms for each iteration

step along the pathway. The displacement is colour-coded from 0 to 18 Å. Coloured regions correspond to the most mobile segments.

**Figure 6.** Different conformations adopted by BCL in going from the open to the closed state. (A) Superposition of BCL at different iterations. The *lid* region is colored in red. The two extreme conformations (closed and open states) of the *lid* are colored in green. From B to E, the figures show the progressive conformational changes occurring along the transition pathway.

**Figure 7.** Exposition to the solvent of the hydrophobic residues (coloured in orange) composing the *lid* region in BCL in the open crystal structure (A) and in the (B) closed homology model.

**Figure 8.** Plot of the relative frequency of contacts for residues whose atoms were detected to be in contact with atoms from other residues of the protein during the transition. Contacts over 60% are colored in dark green, between 40-60% in medium green, between 20-40% in light green. The diameter of the spheres reflects the % of contacts. Colliding residues are shown on BCL structure. High contacts (>60%) are mainly found between residues of the  $\alpha 4$ ,  $\alpha 5$  and  $\alpha 6$  helices.

**Figure 9.** Average potential energy profiles along BCL *lid* transition obtained from (A) the ten computed robotics pathways minimized in explicit water, (B) the ten computed robotics pathways minimized in explicit octane.

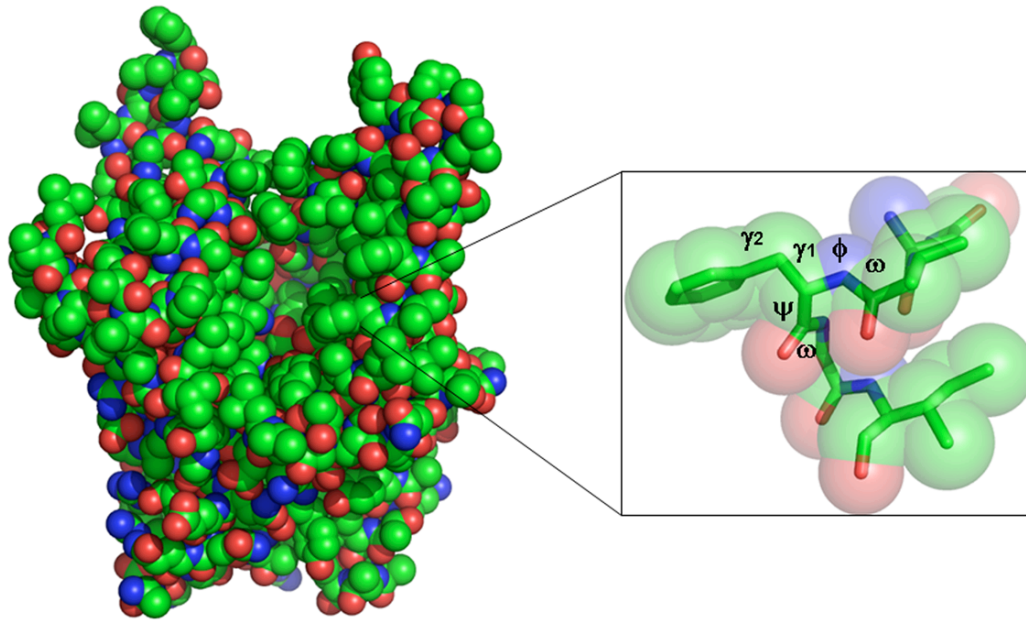


Fig. 1

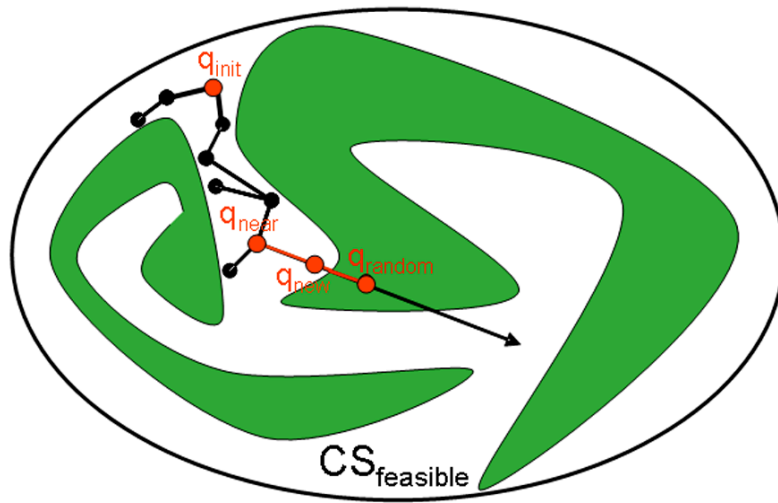
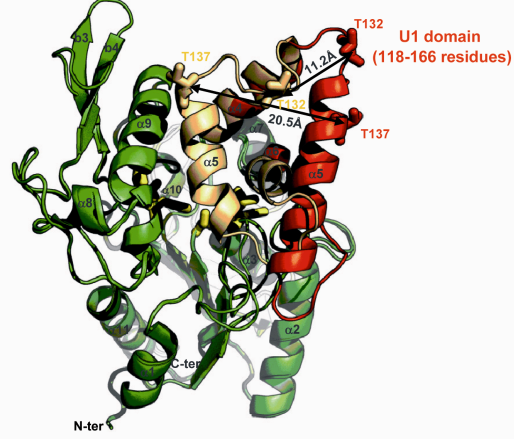
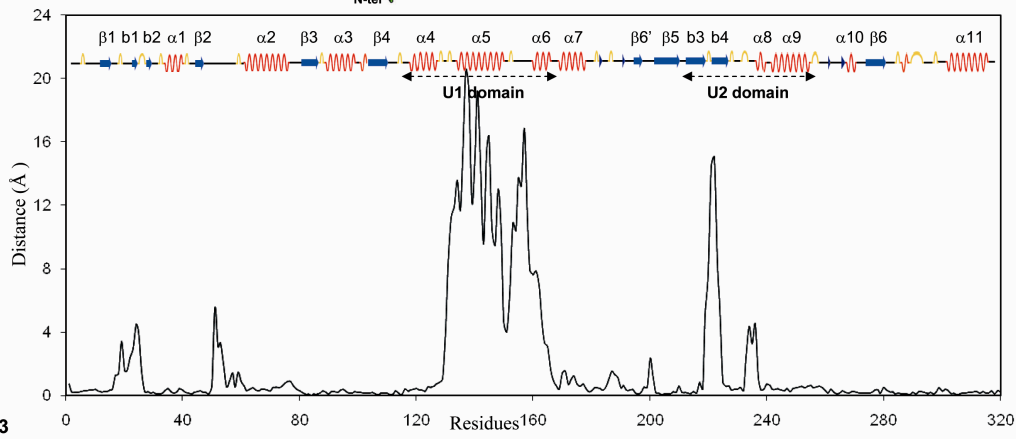


Fig. 2

**A**



**B**



**Fig. 3**

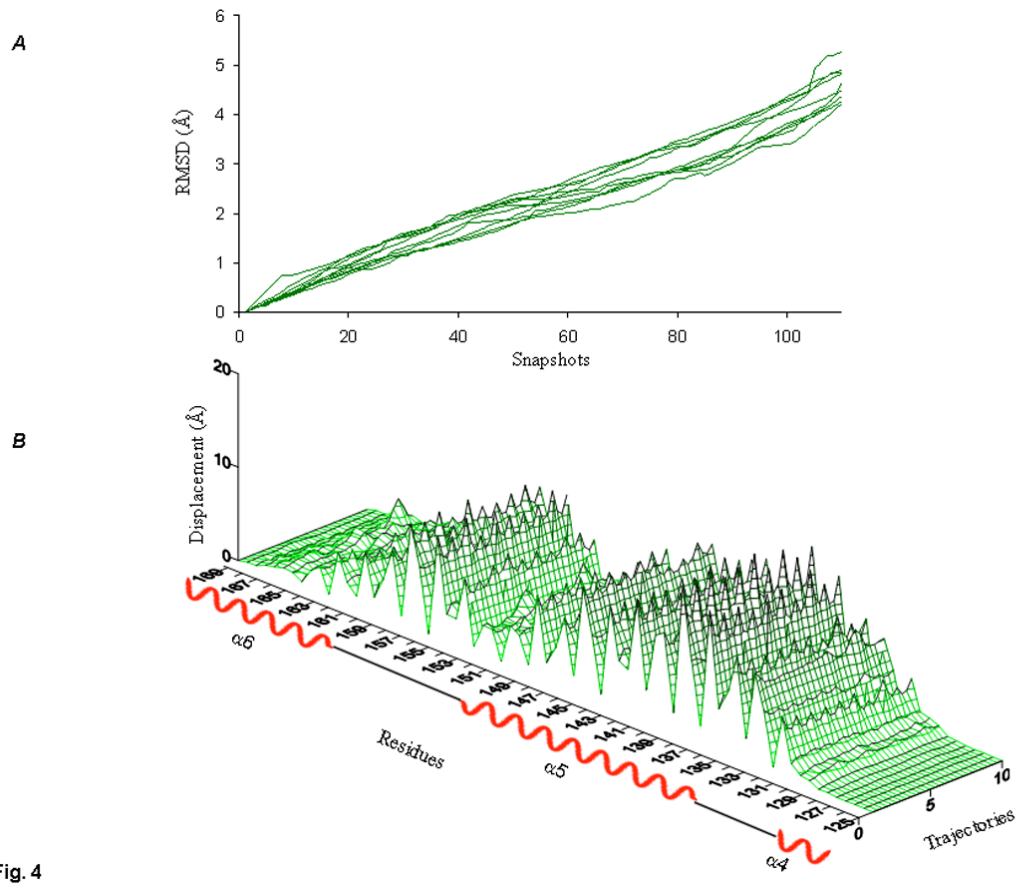


Fig. 4

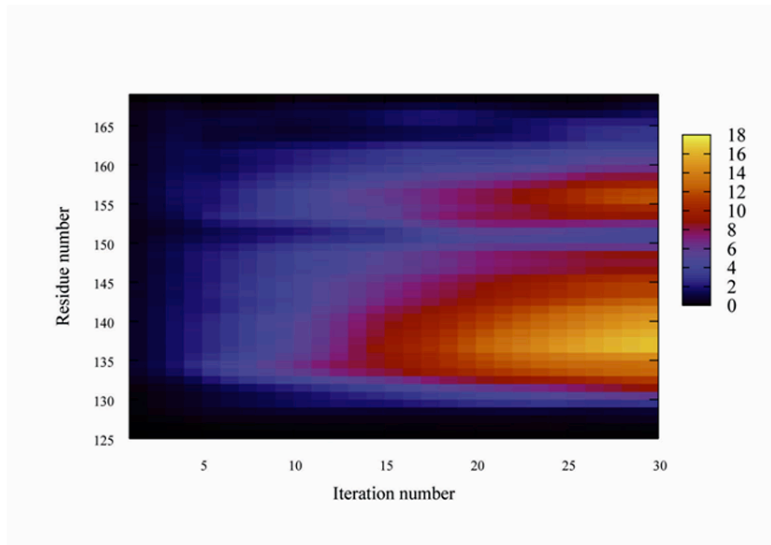


Fig. 5

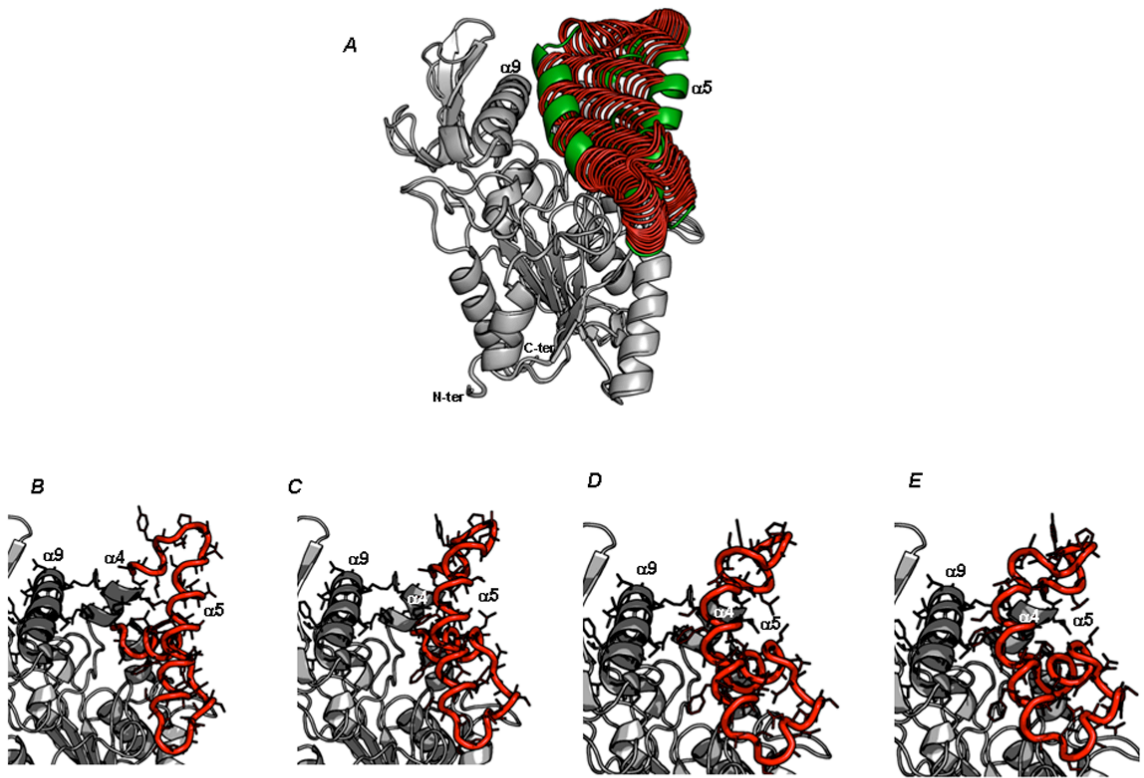
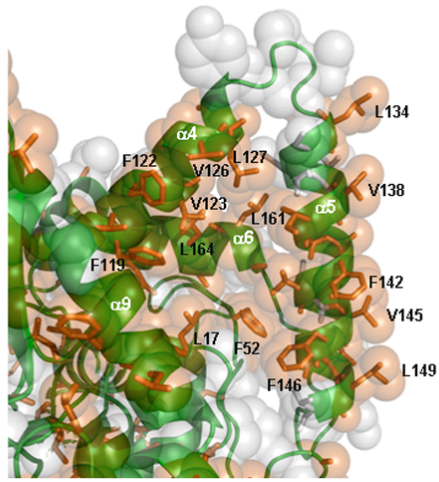
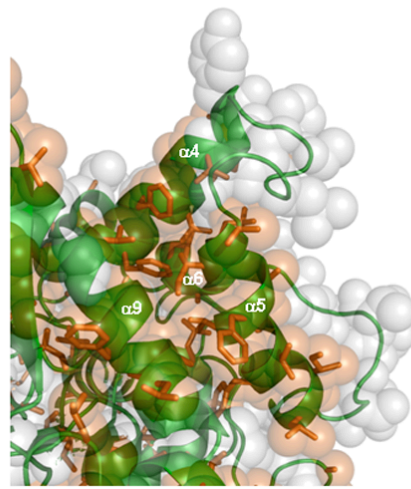


Fig. 6

**A**



**B**



**Fig. 7**



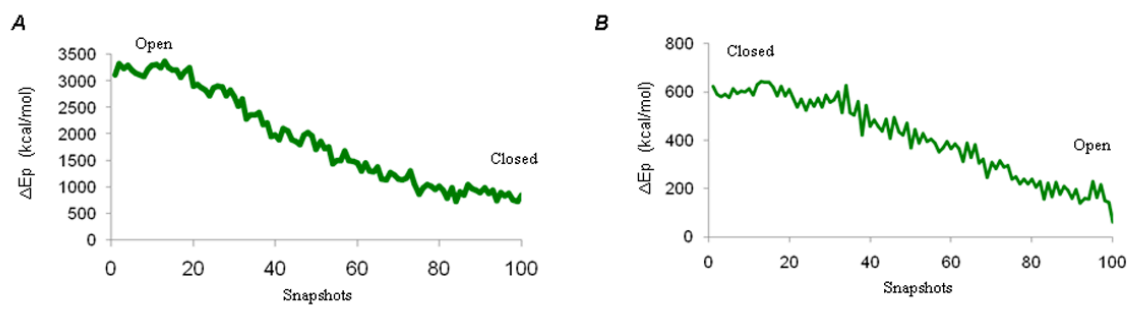


Fig. 9