Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Modeling the effect of pathogenic mutations on the conformational landscape of protein kinases

Giorgio Saladino[1] and Francesco Luigi Gervasio[1,2]

Most proteins assume different conformations to perform their cellular functions. This conformational dynamics is physiologically regulated by binding events and post-translational modifications, but can also be affected by pathogenic mutations. Atomistic molecular dynamics simulations complemented by enhanced sampling approaches are increasingly used to probe the effect of mutations on the conformational dynamics and on the underlying conformational free energy landscape of proteins. In this short review we discuss recent successful examples of simulations used to understand the molecular mechanism underlying the deregulation of physiological conformational dynamics due to non-synonymous single point mutations. Our examples are mostly drawn from the protein kinase family.

Addresses
[1] Department of Chemistry, University College London, London WC1E 6BT, United Kingdom
[2] Research Institute of Structural and Molecular Biology, University College London, London WC1E 6BT, United Kingdom

Corresponding author: Gervasio, Francesco Luigi (f.l.gervasio@ucl.ac.uk)

Current Opinion in Structural Biology 2016, **37**:108–114

This review comes from a themed issue on **Theory and simulation**

Edited by **Modesto Orozco** and **Narayanaswamy Srinivasan**

http://dx.doi.org/10.1016/j.sbi.2016.01.005

Introduction

Proteins are dynamic macromolecules and their function is often dependent on their motions and conformational flexibility [1]. Recent high-resolution structural studies using X-ray crystallography, Cryo-EM and NMR have given direct evidences of how protein plasticity and dynamic behavior is crucial for their function [2–4]. This is particularly true in signaling proteins such as protein kinases that react to allosteric stimuli such as ligand or protein binding and post-translational modifications by switching to an active conformation and starting a signaling cascade that eventually controls the activity and fate of cells [5,6]. The current consensus is that the conformational switch is made possible due to an ensemble of different conformations being accessible to the protein, which moves on a complex *conformational landscape* [7••,8].

The substitution of one or more amino acids of a protein due to pathogenic non-synonymous mutations, may affect the conformational free energy (FE) landscape and alter the equilibrium between different conformations in an analogous way to physiological allosteric signals. Thus, mutations can also have an impact on the conformational transitions necessary for protein–protein interactions and small molecule binding [7••,8,9,10••]. Due to these dramatic effects, even single-point non-synonymous mutations of signaling proteins can lead to the development of several diseases and are often associated with cancer [11,12]. In this respect, understanding how mutations affect the conformational landscape of proteins (and their function) is an important step in predicting the effect of genetic mutations arising from genome-wide screening and in designing effective personalized therapies of complex multi-factorial diseases [13•].

Oncogenic and drug-resistant mutations may affect the equilibria of the various conformations assumed by these proteins, leading to a global change in the conformational dynamics that is not easily captured by static crystal structures. Atomistic molecular dynamics (MD) simulations, which are able to fully capture the dynamical nature of the proteins, are thus an increasingly useful tool in complementing structural data and understand the effect of the mutations on the conformational free energy landscape. As they provide a full atomistic description of the system and its dynamics, MD have been frequently described as a *computational microscope* [14–17]. However, MD have been limited by the accessible timescales. Typically an MD simulation lasts up to a few microseconds. This issue, which is commonly known as the 'timescale problem', prevents a statistically meaningful observation of conformational changes in conventional MD. Recently, the development of special-purpose hardware, such as Anton [18,19], and of new algorithms specifically aimed at enhancing the sampling of MD and solving the time-scale problem, made it possible to analyze in great detail the effects of mutations on the conformational landscape of biomolecules. An in-depth review of enhanced-sampling MD techniques is beyond the scope of the current review. In the reconstruction of conformational free energy landscapes of biomolecules, four types of algorithms (alone or in combination) are often used: (1) methods based on multiple replicas of the system such as parallel tempering or Hamiltonian replica

exchange [20,21]; (2) methods based on the reconstruction of the free energy profile along an optimal path connecting different states, such as milestoning [22], transition path sampling [23] or the path collective variables [24]; (3) methods that enhance the sampling and reconstruct the free energy profile along a set of relevant coordinates (collective variable or CV) such as umbrella sampling [25], Targeted Molecular Dynamics (TMD) [26], Metadynamics and its many derivatives [24,25]; (4) swarms of MD trajectories and Markov State Models [27]. Here we report on recent success in the molecular understanding of the effect of pathogenic mutations on the conformational landscape of protein kinases (one of the most important and studied drug target families) through long or enhanced-sampling atomistic MD simulations.

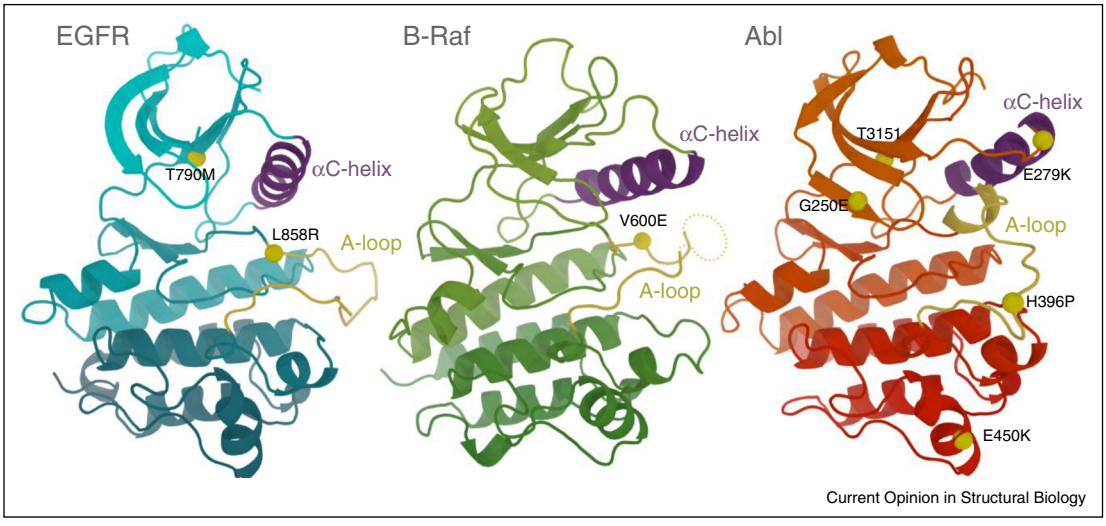
Oncogenic mutations

The regulatory proteins encoded by *oncogenes* and *tumor suppressor genes* play a fundamental role in the onset and progression of cancer [28]. Simulations helped the understanding of the molecular mechanism of oncogenic mutations in these crucial proteins, clarifying their mode of action in oncoproteins, such as Ras [29] and various protein kinases and in tumor suppressor genes such as p53 [28].

A prominent example is the epidermal growth factor receptor (EGFR), a tyrosine kinase that, due to its role in cancer, is one of the most studied signaling proteins. EGFR is a cell-surface receptor involved in the regulation of key cellular processes [30]. Single-point mutations in EGFR, such as L858R (see Figure 1) are among the most frequently observed in lung carcinoma [31]. *In vitro* experiments appear to suggest that EGFR mutants are more active than the wild type [32,33]. Several computational studies have addressed the molecular mechanism by which these mutations affect the complex EGFR regulation, complementing the abundant experimental data. Dixit and Verkhivker [26] simulated the activation of EGFR WT and of the L858R mutant. During activation, in EGFR, as in most protein kinases, the long *activation loop* (A-loop) assumes an extended configuration (see Figure 1). The authors used TMD to shift the loop from the inactive to the active conformations. In the simulations, they observed a two-step mechanism. First the functionally important αC-helix is repositioned to assume an active-like conformation.

The authors concluded that the mutation appear to stabilize the active state, by aiding the assembling of the so-called hydrophobic spine, and the transition of the αC-helix. Wan and Coveney [34] performed long multiple-replica MD simulations of wild-type and L858R EGFR starting from active and inactive crystal structures. They observed that, while both the A-loop and the αC-helix have fluctuations similar to the rest of the protein in the active state, their mobility changes significantly in the inactive one: the A-loop appears more flexible, while the αC is considerably more rigid. Analyzing the distribution of orientations of the αC-helix, the authors also discovered large orientational rearrangements in the active state. L858R changes this distribution and generates alternative orientations of the helix that increase the cavity between the kinase two lobes, thus easing the A-loop transition by removing sterical constraints. In agreement with previous reports, the authors suggest that while the inactive state of EGFR is energetically favored,

Figure 1



The structure of EGFR (cyan, active state), B-Raf (green, semi-active) and Abl (red, inactive) are shown with oncogenic and drug resistant mutations indicated by yellow spheres. The A-loop is open and fully extended in the active structure and closed in the inactive one. The important αC-helix is highlighted in purple.