University of Trento

Department of Physics

Thesis Submitted for the Degree of

Doctor of Philosophy

Generating and Validating Transition Path Ensembles of Protein Folding

---

Supervisor:
Prof. Pietro Faccioli

Author:
Simone Orioli

Academic Year 2017/2018

Contents

Introduction

Long chains of amino acid residues, called proteins, are ubiquitously present in a living organism. They play fundamental roles to sustain life, e.g. interactions with other macromolecules (DNA and RNA), catalysis of chemical reactions, molecule transport and cell signaling [1]. Proteins are formed in the cell by ribosomes, which expel them in a coil-like conformation, called the unfolded state. In order to fulfill their biological role, they need to rearrange in a unique three-dimensional conformation, called the native or folded state. Protein folding, the process by which a protein spontaneously goes from the denatured to the native state [2], is an extraordinarily complex phenomenon: it amounts to a stochastic search of a free energy minimum over a rugged and multidimensional free energy landscape. As a result of this complexity, proteins spend most of the folding time thermally oscillating in their initial reactant state, which is usually separated by high free energy barriers from the native one. It is not a surprise, then, that a general agreement on the molecular mechanism of protein folding has yet to be reached [3, 4, 5, 6].

In the last years, molecular dynamics (MD) [7] emerged as the tool of choice to tackle complex biomolecular problems from a computational perspective. MD consists of all the sets of algorithms and technical knowledge that allows one to integrate the Newton's equations of motion for a system of choice. The main reason behind the success of MD is that it employs empirically derived potentials (the so called force fields) and that it allows one to simultaneously track the motion of all atoms in a molecule with femtosecond time resolution. To date, MD provides the only possible strategy to characterize the dynamics of a macromolecule at the atomic level. However, the intrinsic complexity of a molecule's dynamics makes it challenging to exploit the full potential of MD. During its dynamics, a wide range of characteristic time scales is spanned by the molecule, requiring a very small integration step (usually 1 fs) to capture the fastest motions. Therefore, if one wanted to simulate the folding of a protein, one would need to integrate the equations of motion of hundreds of thousands of atoms for a time that is comparable with the typical folding time, which is usually much longer than milliseconds and might even reach minutes. This means that there are approximately 15 orders of magnitude of timescale sep-

aration between the shortest and the longest processes to simulate, not to mention that one would need to collect an ensemble of folding transitions to characterize the folding mechanism. The problem of timescale decoupling made the sampling of proteins configuration space virtually impossible, until extended cloud infrastructures [8] and special purpose supercomputers [9] were designed specifically to overcome this limitation, at least for a special class of mini-proteins (called *fast folders*) that fold with sub-millisecond timescale. These technological advancements were key in proving that MD simulations were actually able to produce multiple spontaneous folding and unfolding, and that the features extracted from folding simulations had a substantial overlap with experimental knowledge [10, 11, 12, 13, 14]. However, the second timescale still lied far and out of reach and, to the best of our knowledge, no successful attempt was made to simulate protein folding using MD alone ever since.

On a theoretical standpoint, an important effort was made to make the use of MD available to the largest possible class of biological systems. Enhanced sampling methods [15] were result of this endeavor. All of those methods exploit different levels theory and/or approximations to accelerate the dynamics of the system under study. To give just a few examples, replica exchange schemes stochastically increase or decrease the temperature of the system in order to help the system overcome free energy barriers [16]; path-sampling methods, instead, just focus on the reactive portion of the simulation, thus avoiding the need to simulate uninteresting and long thermal oscillations in the reactant basin [17, 18]. Enhanced sampling methods proved to be an extraordinarily powerful tool to study complex conformational transitions [19, 20]. However, their success was modest when dealing with protein folding, as their computational cost confined their application to fast folders. The configurational sampling of medium- and large-sized proteins, the characterization of their folding pathways and the relevant conformations populated along it remained unbearable task for computational bimolecular studies.

An important step forward in this direction was made by Ratchet-and-Pawl molecular dynamics [21, 22] and the related Bias Functional approach [23]. BF method is a path-sampling strategy that employs a biasing force to direct the system towards a given target state, and employs a variational strategy to assess how much a biased pathway is considered realistic, i.e. how probably the path is generated by unbiased MD. The bias force amounts for a harmonic and history-dependent contribution, defined as a function of a predetermined reaction coordinate. Every time the protein spontaneously progresses along the reaction coordinate, the force remains latent, while if the protein tries to backtrack towards more unfolded configurations the bias force adds up to the physical one to help the molecule advance in the conformational transition. This method was used to efficiently compute folding

pathways of proteins that fold in the second timescale or beyond [24, 25] and its results have been widely tested and validated in recent years, both against unbiased MD simulations [23] and experiments [24, 26, 27]. Unfortunately, this algorithm suffers of some notable disadvantages. The first is that it requires a choice of a reaction coordinate, which is made empirically: A priori, one has no simple way to understand if a collective variable is able to fully describe the folding process but still a sub-optimal choice of it might lead to artifacts that are hard to detect. The second one is hidden in the definition of the bias force: The history-dependence of this force breaks the microscopic reversibility of the dynamics. Thus, trajectories computed using the BF method cannot be easily employed to compute folding rates or to assess the Boltzmann weights of the sampled configurations.

This thesis proposes to provide a unified and systematic strategy to overcome the second timescale in protein folding, by exploiting qualities and drawbacks of the Bias Functional Method and proposing new theoretical approaches to overcome its limitations. The first half of the thesis is dedicated to the development of theoretical solutions to the problems discussed above: dependence on an a-priori defined collective coordinate and microscopic non-reversibility of the dynamics. In particular, the **first chapter** is devoted to the definition of a self-consistent scheme that iteratively optimizes the reaction coordinate employed in BF simulations. The iterative scheme naturally emerges as a mean-field approximation of Langevin dynamics and provides two new reaction coordinates that are not defined a-priori, but rather are self-consistently improved. This new strategy goes under the name of Self-Consistent Path Sampling (SCPS) [28] and displays several advantages with respect to the standard BF method, as shown through its application on two examples: the Langevin dynamics of a point in a 2-dimensional potential and the folding dynamics of the WW-domain of FIP35. In both cases, the results of SCPS are quantitatively consistent with the ones predicted by long MD simulations [10]. Even if SCPS enhances the reliability of BF simulations, downgraded by this approach to the initial guess of an optimization process, it only moderately increases the original computational complexity. Thus, self-consistent path sampling provides a sound alternative to the standard BF approach.

The **second chapter** concerns with the problem of recovering kinetic information from SCPS simulations [29]. This is done by means of Transition Path Theory (TPT) [30, 31]. In this theory, the reactive process through which a system performs a transitions between two (or more) states is completely characterized by three stochastic descriptors: the committor function, the reactive probability density and the reactive current. The knowledge of the committor function is key to the determination the other two functions and it is therefore the main ingredient of the whole theory. In

its original formulation, TPT relies on the ergodic assumption, i.e. on the possibility of generating very long trajectories which visit both the reactant and product states many times. On the other hand, rMD trajectories are intrinsically out of equilibrium events, and correspond to short irreversible transitions. Therefore, first of all, we will show how to generalize the results of TPT to the case where, instead of an infinitely long trajectory, one is provided with many short off-equilibrium simulations, e.g. the results of many runs of an enhanced sampling method. A further analysis of SCPS will reveal that its collective variables and the committor function are related: based on this observation, we will show how it is possible to use the results of the self-consistent procedure to approximate the committor function. Finally, we will present two possible algorithms, each one with its pros and cons, that exploit the knowledge of the committor function to exactly sample the transition region of the conformational reaction. We will test the performances of these combined approaches on a 2-dimensional system, showing that it provides results that are quantitatively comparable with TPT ones.

The second part of the manuscript is devoted to applications of the BF method on two different proteins: Canine milk lysozyme and alpha1-antitrypsin (A1AT). Canine milk lysozyme, discussed in the **third chapter**, is a 120 residues protein that folds within the second timescale. Its folding dynamics has been characterized [32] by means of circular dichroism (CD), and experimental technique that measures the differential emission of a protein solution upon circularly polarized light impinging [33, 34]. These experiments proved the existence of two metastable states of this protein, one of which differs from the native one just by the formation of some key tertiary contacts. Because of the chemical properties of the tryptophans present in the protein chain, the formation of these contacts can be monitored by looking at variations of the CD spectrum in the near-UV region (240-300 nm). Unfortunately, the experimental CD signal is an average over an ensemble of molecules and does not provide any atomic level information: the interpretation of this signal has always to rely on some a priori knowledge of the folding process. This third chapter will be devoted to show how it is possible to use the BF method, coupled to quantum chemical calculations [35, 36, 37], to enhance experimental sensitivity and provide atomic level description of the different metastable states of the protein We will show how to this just by assuming the knowledge of the protein's crystal structure and the CD spectrum of the metastable configurations along the folding pathway.

Finally, the **fourth chapter** will be concerned with the study of alpha1-antitrypsin, a 394 residues protein that belongs to the superfamily of serine protease inhibitors. The inhibition mechanism of this protein is unique and proceeds through a complex conformational modification, but it is easily hacked by pathological mutations

[38, 39]. In particular, the so called Z-mutation deeply destabilizes the native state of the protein, leading to an inevitable misfolding of the chain. In this chapter we will show that the BF method is sensible to point mutations in a protein chain, thus it can be used as a tool to study protein misfolding [25]: we will consider many different point mutations of A1AT, some of which are known to cause pathological consequences while some others are instead known to compensate for the destabilizing effects of other mutations. The reported results are consistent with the experimental knowledge and also provide new predictions that can be tested in a wet laboratory.

Remarks on Notation and Useful Observables

In this paragraph we want to anticipate some notation that will consistently be used throughout the whole manuscript and provide the definition of some observables that will be useful when analyzing protein folding results.

Heaviside Theta Function We will always assume the following convention for the Heaviside Theta function θ

$$\theta(x) = \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

For the alternative convention, assuming $\theta(0) = 1/2$, we will use the symbol $\tilde{\theta}$. Therefore:

$$\tilde{\theta}(x) = \begin{cases} 1 \text{ if } x > 0 \\ 1/2 \text{ if } x = 0 \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

Notation In the manuscript, we will employ the following notation:

1. x refers to a path. Depending on the situation the path might be in the full phase space or just in configuration space;

2. $\mathbf{x} = x(t)$ refers to a phase space point or a configuration space point, corresponding to the value assumed by the path at time t;

3. $\mathbf{x}_i$ refers to a 3-dimensional vector, specifying the position of a given atom i. Thus, for a system composed by N atoms, $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$.

Root Mean Square Deviation (RMSD) Given a configuration $\mathbf{x}$ and a reference one $\mathbf{y}$, both of them composed by N atoms with masses m_i, the RMSD of $\mathbf{x}$

with respect to $\mathbf{y}$ is calculated by first fitting $\mathbf{x}$ to $\mathbf{y}$ [40] and then by computing

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} m_i ||\mathbf{x}_i - \mathbf{y}_i||^2}{\sum_{i=1}^{N} m_i}} \tag{3}$$

RMSD is used as a measure of distance in configuration space. RMSD values have been computed either with GROMACS 4.6.5 [41] or with MDTraj [42].

Root Mean Square Fluctuation (RMSF) The RMSF of an atom i with position $\mathbf{x}_i$ with respect to its average position $\langle \mathbf{x}_i \rangle$ is given by

$$\mathrm{RMSF}_i = \sqrt{\langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle)^2 \rangle} \tag{4}$$

RMSF is usually used to determine the flexibility of a chain. RMSF values have been computed using GROMACS 4.6.5 [41].

Radius of Gyration (RG) Given a configuration $\mathbf{x}$ composed by N atoms, each one of mass m_i and located at distances $\mathbf{r}_i$ from the molecule's center of mass, the RG is defined as

$$\mathrm{RG} = \left(\frac{\sum_{i=1}^{N} m_i \mathbf{r}_i^2}{\sum_{i=1}^{N} m_i} \right) \tag{5}$$

RG values have been computed using GROMACS 4.6.5 [41].

Fraction of Native Contacts (Q) Given a configuration $\mathbf{x}$, the alpha-carbon contact map is the matrix defined as

$$C_{ij}^{\alpha}(\mathbf{x}) = \frac{1 - \left(\frac{||\mathbf{x}_i^{\alpha} - \mathbf{x}_j^{\alpha}||}{r_0} \right)^6}{1 - \left(\frac{||\mathbf{x}_i^{\alpha} - \mathbf{x}_j^{\alpha}||}{r_0} \right)^{10}} \tag{6}$$

where $\mathbf{x}_i^{\alpha}$ are the positions of the alpha-carbons in $\mathbf{x}$ and $r_0 = 0.75$ nm is a reference distance that defines a contact. The quantity $Q^0 = \sum_{i>j} C_{ij}^{\alpha}(\mathbf{x}_P)$, where $\mathbf{x}_P$ is the folded configuration, provides a continuous definition of the total number of contacts in the native protein. The fraction of native contacts is thus defined as

$$Q(\mathbf{x}) = \frac{1}{Q^0} \sum_{i>j} C_{ij}^{\alpha}(\mathbf{x}) \tag{7}$$

Q values have been computed adapting MDTraj [42] functions.

Plots and Structure Visualization Plots have been made with Matplotlib [43] or Gnuplot [44], while structures have been rendered using PyMOL [45] or VMD [46].

Chapter 1

Sampling Rare Transitions

The use of special purpose supercomputers [9], distributed cloud computing facilities [8] and the most widespread enhanced sampling techniques proved that folding simulations of biologically relevant proteins [1] are still unfeasible for modern computational resources. The next exascale generation of supercomputers is awaited to mitigate this problem, but millisecond timescale simulations are not expected to become routine for every scientist soon. The Bias Functional method [23] provides a valid instrument to approach this problem. This method is based on the so called *ratchet-and-pawl molecular dynamics* (rMD) and applies a variational principle to assess which trajectory, among a set of trial pathways, is the most realistic one. rMD simulations apply a history-dependent harmonic force to a system only if it tries to backtrack towards the reactant, while they apply no force at all if the system spontaneously progresses along a predefined reaction coordinate. It is clear that if the reaction coordinate employed in an rMD calculation was the exact one, a sampling scheme like the Bias Functional method would provide the correct folding mechanism. However, rMD simulations are usually performed along a predetermined reaction coordinate, related to the fraction of native contacts and not necessarily optimal. The variational principle is expected to soften the effects of possible errors introduced by the sub-optimal choice of the reaction coordinate. This is however not a definitive answer, as the presence of systematic errors is not easily detected.

In this chapter, we will introduce a novel enhanced sampling approach that is able to generate folding simulations without relying on a specific choice of reaction coordinate. Instead, the reaction coordinate will represent the output of a self-consistent optimization of an initial guess, thus providing further insight into the reaction of interest. This new scheme is the result of a mean-field approximation of two auxiliary variables introduced in the path integral representation of Langevin dynamics. Its computational cost remains modest and affordable to every laboratory

[1] the average dimension of a human protein is approximately 350 amino acids [47]

having access to a medium- or small-sized clusters.

The chapter is organized as follows. Section 1.1 is devoted to a brief review of the fast-paced quest to break the millisecond barrier in protein folding simulations, from year 2000 to our days. We will focus on some of the most common enhanced sampling algorithms and shortly describe their contributions and results, highlighting the necessity of new strategies to tackle the problem of folding. Section 1.2 will focus on the theoretical and technical foundations of the Bias Functional method. Section 1.3 introduces the new self-consistent strategy: we will prove that the new algorithm comes from a mean-field approximation of Langevin dynamics and discuss how to implement it in practice to perform protein folding simulations. Section 1.4 is instead devoted to a first application of this algorithm to a simple 2-dimensional system, where we will prove the ability of the algorithm to systematically improve the quality of a poor guess of the reactive pathway. In section 1.5 we will show the application of the new methodology to the folding of a fast-folding protein, the WW-domain of FIP35. Direct comparison with unbiased MD simulations will show that the folding mechanism provided by the self-consistent algorithm is indeed realistic and reliable. Finally, in section 1.6 we will draw some conclusions and discuss the work that still has to be done to validate the algorithm.

1.1 Breaking the Millisecond Barrier in Protein Folding

In 2010, the introduction of a dedicated supercomputer called *Anton* [9] allowed scientists at D. E. Shaw Research to run the first unbiased protein folding simulations [10]. The work focused on two proteins: the WW-domain of FIP35, a three-stranded β-sheet, and the villin headpiece C-terminal fragment, a three-helix boundle protein. Both of them folded in the μs regime, and many spontanous folding and unfolding events were detected. The following year, Lindorff-Larsen and his collaborators [11] relied again on the supercomputing capabilities of Anton to fold a representative set of 12 fast-folding proteins, ranging from 10 to 80 residues in length and from 104 to 2936 μs of total simulation time. Besides the exciting scientific goal that was reached, these results showed for the first time that hardware and software technologies were mature enough to run many ultra-long MD simulations and reach the millisecond time scale. Moreover, it showed that the existing forece fields were accurate enough to recognize protein native structures. However, it was just the tip of the iceberg. Many years before 2010, researchers had already performed pioneering protein folding simulations based on *enhanced sampling techniques*. We call enhanced sampling techniques all the theoretical methodologies and algorithms that have been designed to accelerate the performances of MD simulations. Several

methods, gounded on different levels of accuracy and efficiency, have been developed to address the long-standing problem of protein folding, and their widespread diffusion through open source MD simulation packages (GROMACS [48], NAMD [49], OpenMM [50] and Plumed [51] to cite just a few) made their use become routine. Unfortunately, millisecond time scale simulations are not yet routine, even employing advanced sampling strategies.

We want to dedicate this section to a brief historical review of the advances in enhanced sampling from year 2000 to our days. We will just focus on methodologies for which applications to protein folding have been reported in the literature. For this reason, many well-known methods (e.g. milestoning [52], finite temperature string method [18], wighted ensembles [53] and many others) will not be discussed here. Furthermore, we decided to restrict ourselves to methods that have been applied to all-atom systems, and we on purpose avoided any discussion about coarse-graining techniques, which define a wide and important field on their own [54, 55]. More details about the ideas behind the different methods and further in-depth analyses can be found in several brilliant reviews, not just on protein folding but on the more general problem of conformational transitions [56, 57, 58, 59, 60, 20, 61, 62, 15].

Replica Exchange Molecular Dynamics (REMD) In REMD [16], several replicas of the system are simulated at different temperatures. At regular time intervals, a swap between replicas close in temperature is attempted, according to a Metropolis criterion. The idea behind this method is to use the temperature as a parameter to help the system efficiently escape local minima. Given the non-Hamiltonian, intrinsically stochastic nature of temperature swaps, REMD cannot be employed to retrieve dynamical information, even though some attempts have been made to retrieve kinetic data from them [63].

TrpCage miniprotein (20 amino acids) was the first protein whose folding was tackled using REMD. In particular, Pitera and Swope [64] and Zhou [65, 66] reported successful folding events, respectively, in implicit and explicit solvent. In the same year, Garcia and Onuchic [67] folded the 46 amino acid fragment of the fragment B of staphylococcal protein A, while Rao and Caflish [68] showed that REMD allowed to reach a statistically relevant number of folding events of a three-stranded β-sheet within the microsecond aggregate simulation time scale. Andrec and collaborators [69] combined the use of REMD and random walk models to fold the C-terminal peptide of the B1 domain on protein G. At a later time, Liu and collaborators [70] proposed a reportedly more efficient REMD scheme (hydrophobic aided REMD), which was benchmarked on the folding of 3K(I) α-helix, while Ding and his colleagues [71] showed that coupling discrete molecular dynamics [72]

with REMD makes it even more efficient. Using the latter, it was possible to fold six fast-folding proteins composed by 20 to 60 residues. Later studies, conducted by Pashek and collaborators, focused on the stability diagram of TrpCage [73, 74]. Finally, recent applications of REMD in protein folding regard the benchmarks of custom force-fields for protein folding, where 17 fast-folding proteins were correctly folded in implicit [75] and explicit [76] solvent, or the study of the folding of TrpZip2 close to a graphene surface [77].

Metadynamics In metadynamics [78], the exploration of the conformational space occurs by evolving a set of collective coordinates and depositing, at a given rate, Gaussian potentials along the trajectory. If the system visits close-by points along the trajectories, Gaussians start to accumulate filling the potential wells, up to the point where free energy barriers are overcome and all the relevant configuration space is explored. The free energy along the given employed collective coordinates is exactly retrieved, up to fluctuations and an additice constant, from the sum of the deposited Gaussian potentials. In its *well-tempered* formulation [79], the parameters of the Gaussians can be tuned along the simulation to reach a better convergence, while in its *bias-exchange* formulation multiple replicas of the system are simulated along different reaction coordinates, allowing replica swapping at any given time acoording to a Metropolis criterion [80].

The study of protein folding free energy landscapes with metadynamics was first attempted by Bussi and collaborators [81], where metadynamics was coupled with REMD to fold the C-terminal β-hairpin of protein G. The development of the bias-exchange technique by Piana and Laio [80] allowed to study the folding of TrpCage in explicit solvent by means of very short (ns) simulations. A similar technique was also employed the following year by Piana and collaborators [82] to study the effect of point mutations in the folding pathway of villin headpiece. In later years, Kimanius and collaborators [83] showed that it was possible to employ SAXS data as collective coordinates in bias-exchange metadynamics to fold the TrpCage miniprotein. Particular focus on the characterization of metastable states in protein folding was given in recent works by Bonetti and collaborators [84] and Singh and colleagues [85], where they studied, respectively, the folding of 49 residues measles virus X-domain and the one of the insulin monomer. We also note that the works on prion misfolding [86] and on the computation of the free energy landscape of an intrinsically disordered protein [87], both employing metadynamics, provide relevant contributions to protein folding field.

Umbrella Sampling Similarly to metadynamics, umbrella sampling [88, 89, 90] requires as an input a set of collective variables that are supposedly able to describe important motions of the system. Restraint potentials, usually harmonic (*umbrellas*), are applied along the given collective variables in order to force the system to sample relevant regions in the configuration space. Particular care should be payed to provide a reasonable intersection between umbrellas, which facilitates the reconstruction of the transition between them, and the extraction of simulation data, that is usually performed through a weighted histogram analysis method (WHAM) [91].

Most of the applications of umbrella sampling to protein folding date to years before 2000, but it is still worth mentioning the most relevant ones. Young and Brooks [92] studied the folding free-energy of an alanine α-helix in explicit solvent at different chain lengths. Two years later, Sheinerman and Brooks [93] applied the same strategy to study the folding pathway of the B1 fragment of streptococcal protein G, using the total number of native contacts as a collective variable. The subsequent year, Bursulaya and Brooks [94] studied the folding free energy of a 20-residues three-stranded β-sheet in explicit water, by employing the total number of contacts and the radius of gyration as collective coordinates. Recently, Meshkin and Zhou [95], used the fraction of native contacts as a reaction coordinate to fold TrpCage and the zinc-finger motif (BBA).

Markov State Models-guided Sampling (MSM) Markov State Models (MSMs) [96, 97, 98, 99, 100] are a statistical framework within which the complex dynamics of a molecule is reduced to a memory-less (*Markovian*) transition network. The dynamics at MSM level is described using a master equation formalism, which defines the transition probabilities between *microstates*. Each microstate is learned from the input trajectories by applying optimal dimensional reduction techniques, such as time-lagged independent component analysis (tICA) [101, 102]. The major strength of MSMs is their ability to integrate several independent simulations: the higher the number of microstates in the model, the shorter are the trajectories that need to be simulated to cover the conformational space. Luty and McCammon [103] were the first ones to realize this: they created a small Markovian model by integrating many short simulations of a bimolecular system, showing for the first time that MSMs can be actually used as an enhanced sampling tool. In recent years, this idea has been widely exploited and converged into the development of the Folding@Home distributed computing platform [8].

Among the first MSM-driven folding simulations we include the work by Noé and collaborators [104], where the folding pathway of PinWW domain was studied by integrating many off-equilibrium simulations using MSMs and transition path

theory. In the same year, Ensign and Pande [105] simulated the folding of the WW-domain of FIP35 in implicit solvent, using the Folding@Home distributed computing facility and Bowman and collaborators [106] interpreted the obtained results using MSMs. The subsequent year, the Folding@Home facility was employed to simulate many folding transitions of the NTL9 protein in implicit solvent, a millisecond folder of 39 amino acids, as reported by Voelz and collaborators [12]. Bowman, Voeltz and Pande [13] also studied the folding of the 80 residue fragment of the λ-repressor, a simulation which captures dynamics beyond the millisecond time scale. Two years later, Voelz and his colleagues [14] published a study on the folding of the acyl-coenzyme A binding protein, which folds in 10 ms. To date, this is considered the longest time scale ever reached by unbiased simulations. Finally, in recent years Weber and collaborators [107] studied the misfolding properties of seven fast-folding proteins, and how beta-sheet rich configurations in the folding landscape might slow the folding dynamics of these systems.

Adaptive Sampling Methods Adaptive sampling methods are enhanced sampling techniques that guide the exploration of the conformational space via an on-the-fly analysis of the results of the simulation. These analyses help understanding which regions have been undersampled and direct the dynamics towards them. The main complexity behind these kinds of methods is the choice of an optimal way to reduce the data dimensionality and/or find lower dimensional embeddings where the relevant dynamics is happening.

Rao and Caflish [108] exploited this idea by mapping the folding landscape of a 20-residue antiparallel β-sheet on a complex network (a precursor of MSMs), while Preto and Clementi [109] used diffusion maps [110] (in their *locally scaled* formulation [111]) to find an optimal low-dimensional embedding of the dynamics and guide the sampling of the conformations of an alpha-helix. Also essential dynamics, e.g. dynamics driven by principal component analysis (PCA) [112], has been used to fold a protein: in particular, Daidone and collaborators [113] were able to use it to fold Horse Heart Cytochrome C protein. We note here that MSM-guided dynamics can be considered an adaptive sampling method.

Accelerated Molecular Dynamics (AMD) AMD [114] is an ehnaced sampling technique that helps a system escape deep energy minima by locally flattening them. When the system potential becomes smaller than a threshold, a bias is added to help the system leave the local well: the shape of the biasing energy and the threshold energy, thus the overall flattening of the energy landscape, are controlled by the user. Boltzmann statistics is recovered by means of a reweighting procedure [115, 116].

AMD was successfully used to study protein folding. In particular, we report here the two works by Yang and collaborators [117, 118], who were able to fold the TrpCage and TrpZip2 miniproteins, together with the results of Miao and colleagues [119], who showed that AMD can be employed to sensibly accelerate the dynamics and access the folding time scales of the villin headpiece and the WW-domain of FIP35.

Path-Sampling and Optimization Methods Path-sampling methods focus on the generation of the reactive portion of a trajectory, in order to avoid the problem of simulating long thermal oscillations in the reactant and/or in the product of the conformational reaction. The most famous method that embodies this idea is transition path sampling (TPS) [17], a Monte Carlo method that, given an initial guess for path joining the reactant to the product, is able to retrieve the transition path ensemble of the reaction. TPS revealed to be a useful tool to tackle protein folding. Bolhuis [120, 121] showed how it was possible to fold the C-terminal β hairpin of G-B1 protein using TPS, while Evans and Wales [122] reported a performance increase of the same calculation when performed using *discrete path sampling* [123]. Two years later, Juraszek and collaboratos [124] used TPS to fold the TrpCage miniprotein in explicit solvent, and also showed how it was possible to obtain more accurate results by means of improved versions of TPS [125, 126], namely *forward flux sampling* [127] and *transition interface sampling* [128]. The same year, Chopra and collaborators [129] studied the folding of a 54-residues polyglutamine chain.

Other path-optimization techniques have been successfull in simulating protein folding, and it is worth mentioning some more results here. Eastman and collaborators [130] studied the folding of the 36-residues avian pancreatic polypeptide using a reaction path-annealing algorithm, while recently, Lee and collaborators [131] designed a new algorithm to increase the computational efficiency of the Onsager-Machlup action optimization and computed the folding pathway of the 28-residues FSD-1 protein.

1.2 The Bias Functional Method

Consider a set of N solute particles (e.g. the atoms of a protein), whose positions are given by $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, immersed in a solvent (e.g. water with a physiological concentration of salt) whose M atoms are located at $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_M)$. The motion of this system can be modeled as a set of $3(N + M)$ coupled Langevin equations:

$$\begin{aligned} m_i \ddot{\mathbf{x}}_i &= -\gamma_i \dot{\mathbf{x}}_i - \boldsymbol{\nabla}_i U(\mathbf{x}, \mathbf{y}) + \boldsymbol{\eta}_i(t) \qquad i = 1, \ldots, N \\ \mu_j \ddot{\mathbf{y}}_j &= -\sigma_j \dot{\mathbf{y}}_j - \boldsymbol{\nabla}_j U(\mathbf{x}, \mathbf{y}) + \boldsymbol{\xi}_j(t) \qquad j = 1, \ldots, M \end{aligned} \tag{1.1}$$

where m_i and μ_j provide, respectively, the solute and solvent atom masses, γ_i and σ_j represent the corresponding friction coefficients. $\mathbf{F}_i(\mathbf{x},\mathbf{y}) = -\nabla_i U(\mathbf{x},\mathbf{y})$ represents the force-field and $\boldsymbol{\eta}_i(t)$ and $\boldsymbol{\xi}_j(t)$ are white-noise forces satisfying the fluctuation-dissipation theorem:

$$\begin{aligned} \langle \boldsymbol{\eta}_i(t) \cdot \boldsymbol{\eta}_j(\tau) \rangle &= 6 m_i \gamma_i k_{\mathrm{B}} T \delta_{ij} \delta(t-\tau) \qquad i,j = 1,\ldots,N \\ \langle \boldsymbol{\xi}_l(t) \cdot \boldsymbol{\xi}_h(\tau) \rangle &= 6 \mu_l \sigma_l k_{\mathrm{B}} T \delta_{lh} \delta(t-\tau) \qquad l,h = 1,\ldots,M \end{aligned} \tag{1.2}$$

The conditional probability density for the solute to perform a transition from a *reactant* configuration $\mathbf{x}_{\mathrm{R}} = x(0)$ to a *product* configuration $\mathbf{x}_{\mathrm{P}} = x(t)$ in a time t, where x is the path connecting the two points, can be written as the following path-integral:

$$p(\mathbf{x}_{\mathrm{P}}, t | \mathbf{x}_{\mathrm{R}}, 0) = \frac{1}{Z} \int \mathrm{d}\mathbf{y}_{\mathrm{F}} \int \mathrm{d}\mathbf{y}_{\mathrm{I}} \int_{\mathbf{y}_{\mathrm{I}}}^{\mathbf{y}_{\mathrm{F}}} \mathcal{D}y \int_{\mathbf{x}_{\mathrm{R}}}^{\mathbf{x}_{\mathrm{P}}} \mathcal{D}x \; e^{-S_{\mathrm{OM}}[x,y]} e^{-\beta U(\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{I}})} \tag{1.3}$$

In Eq. (1.3), Z represents the system's partition function, $\mathbf{y}_{\mathrm{I}}$ and $\mathbf{y}_{\mathrm{F}}$ are, respectively, the initial and final configurations of the solvent particles and S_{OM} is the Onsager-Machlup action (see appendix A for details):

$$\begin{aligned} S_{\mathrm{OM}}[x,y] \equiv \frac{\beta}{4} \int_0^t d\tau \Bigg[& \sum_{i=1}^{N} \gamma_i^{-1} \left(m_i \ddot{\mathbf{x}}_i + \gamma_i \dot{\mathbf{x}}_i + \boldsymbol{\nabla}_i U(\mathbf{x},\mathbf{y}) \right)^2 \\ & + \sum_{j=1}^{M} \sigma_j^{-1} \left(\mu_j \ddot{\mathbf{y}}_j + \sigma_j \dot{\mathbf{y}}_j + \boldsymbol{\nabla}_j U(\mathbf{x},\mathbf{y}) \right)^2 \Bigg] \end{aligned} \tag{1.4}$$

If one removes the integration over the solute paths in Eq. (1.3), one obtains the probability density associated to a path x:

$$\mathcal{P}[x] = \frac{1}{Z} \int \mathrm{d}\mathbf{y}_{\mathrm{F}} \int \mathrm{d}\mathbf{y}_{\mathrm{I}} \int_{\mathbf{y}_{\mathrm{I}}}^{\mathbf{y}_{\mathrm{F}}} \mathcal{D}y \; e^{-S_{\mathrm{OM}}[x,y]} e^{-\beta U(\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{I}})} \tag{1.5}$$

At this point it is useful to recall that the average value of an observable $O(x,y)$ along the solute path is defined as follows:

$$\langle O(x,y) \rangle = \frac{\int \mathrm{d}\mathbf{y}_{\mathrm{F}} \int \mathrm{d}\mathbf{y}_{\mathrm{I}} \int_{\mathbf{y}_{\mathrm{I}}}^{\mathbf{y}_{\mathrm{F}}} \mathcal{D}y \; O(x,y) e^{-S_{\mathrm{OM}}[x,y]} e^{-\beta U(\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{I}})}}{\int \mathrm{d}\mathbf{y}_{\mathrm{F}} \int \mathrm{d}\mathbf{y}_{\mathrm{I}} \int_{\mathbf{y}_{\mathrm{I}}}^{\mathbf{y}_{\mathrm{F}}} \mathcal{D}y \; e^{-S_{\mathrm{OM}}[x,y]} e^{-\beta U(\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{I}})}} \tag{1.6}$$

The statistically more relevant paths are the ones which maximize Eq. (1.5), thus obey:

$$\frac{\delta \mathcal{P}[x]}{\delta x} \sim 0 \tag{1.7}$$

By explicitly carrying out the functional derivative one obtains an equivalent condition:

$$\left\langle \frac{\delta S[x,y]}{\delta x} \right\rangle \sim 0 \tag{1.8}$$

In principle it is possible to obtain reactive pathways consistent with Eq. (1.8) by functionally optimizing the Onsager-Machlup action or related functionals [62, 131, 132]. However, as discussed in section 1.1, this strategy becomes unfeasible for systems with biologically relevant size, e.g. proteins. For this reason, in this section we shall look for a different solution to this problem. In particular, we are going to recover the so-called *Bias Functional Method* [23].

Consider a stochastic dynamics in which one introduces a time-dependent bias force $\mathbf{F}_i^{\text{bias}}(X,t)$ that acts exclusively on the solute atoms. In this new dynamics, the probability associated to the path x becomes

$$\mathcal{P}[x] = \frac{1}{Z}\int \mathrm{d}\mathbf{y}_{\mathrm{F}} \int \mathrm{d}\mathbf{y}_{\mathrm{I}} \int_{\mathbf{y}_{\mathrm{I}}}^{\mathbf{y}_{\mathrm{F}}} \mathcal{D}y \; e^{-S_{\text{bias}}[x,y]} e^{-\beta U(\mathbf{x}_{\mathrm{R}},\mathbf{y}_{\mathrm{I}})} \tag{1.9}$$

where

$$\begin{aligned} S_{\text{bias}}[x,y] \equiv \frac{\beta}{4}\int_0^t d\tau \Bigg[& \sum_{i=1}^{N} \gamma_i^{-1} \left(m_i \ddot{\mathbf{x}}_i + \gamma_i \dot{\mathbf{x}}_i + \boldsymbol{\nabla}_i U(\mathbf{x},\mathbf{y}) - \mathbf{F}_i^{\text{bias}}(\mathbf{x},t) \right)^2 \\ & + \sum_{j=1}^{M} \sigma_j^{-1} \left(\mu_j \ddot{\mathbf{y}}_j + \sigma_j \dot{\mathbf{y}}_j + \boldsymbol{\nabla}_j U(\mathbf{x},\mathbf{y}) \right)^2 \Bigg] \end{aligned} \tag{1.10}$$

Of course, the Onsager-Machlup action in Eq. (1.4) is recovered from Eq. (1.10) in case of vanishing bias force. Notice that the *unbiased* probability $\mathcal{P}[x]$ can be written as a function of the *biased* probability $\mathcal{P}_{\text{bias}}[x]$, just by multiplying the integrand in Eq. (1.5) by $e^{-S_{\text{bias}}[x,y]} e^{S_{\text{bias}}[x,y]}$:

$$\mathcal{P}[x] = \left\langle e^{-\Delta S[x,y]} \right\rangle_{\text{bias}} \mathcal{P}_{\text{bias}}[x] \tag{1.11}$$

where $\Delta S[x,y] = S_{\text{OM}}[x,y] - S_{\text{bias}}[x,y]$ and $\langle\cdot\rangle_{\text{bias}}$ is the average computed in the biased theory. Note that $\Delta S[x,y]$ does not depend extensively on $\mathbf{y}$ anymore, because the bias force acts only on the solute atoms and the solvent portion of the actions simplifies away. However, a residual dependence, mediated by the force $\boldsymbol{\nabla}_i U(\mathbf{x},\mathbf{y})$, remains and, as we will see in a moment, one can take care of it at the level of a saddle point approximation.

Using Eq. (1.11) one can recast Eq. (1.7) in the following exact form:

$$\frac{\delta \mathcal{P}_{\text{bias}}[x]}{\delta x} \left\langle e^{-\Delta S[x,y]} \right\rangle_{\text{bias}} + \left\langle \frac{\delta}{\delta x} e^{-\Delta S[x,y]} \right\rangle_{\text{bias}} \mathcal{P}_{\text{bias}}[x] \sim 0 \tag{1.12}$$

One can choose now to restrict the search of the paths satisfying Eq. (1.12) among the paths $\bar{x}$ generated within the biased dynamics, so that

$$\frac{\delta \mathcal{P}_{\text{bias}}[\bar{x}]}{\delta \bar{x}} \sim 0 \tag{1.13}$$

and the left term in Eq. (1.12) vanishes. Since $\mathcal{P}_{\text{bias}}[\bar{x}]$ is non-zero, one obtains

$$\left\langle \frac{\delta}{\delta \bar{x}} e^{-\Delta S[\bar{x},y]} \right\rangle_{\text{bias}} \sim 0 \tag{1.14}$$

The result in (1.14) is not exact anymore. Indeed, the search of the optimal path has been restricted only among the ones computed in the presence of the bias. Eq. (1.14) can be used to estimate which, among the paths computed in the biased theory, is the one with the highest probability to be obtained in the unbiased theory. To do so, it is easy to see that Eq. (1.14) is equivalent to

$$\frac{\delta \Delta S[\bar{x}, y]}{\delta \bar{x}} \sim 0 \tag{1.15}$$

Expanding Eq. (1.15) one finds

$$\frac{\delta}{\delta \bar{x}} \sum_{i=1}^{N} \gamma_i^{-1} \int_0^t d\tau \left[-2\mathbf{F}_i^{\text{bias}}(\bar{\mathbf{x}}, \tau) \cdot \left(m_i \ddot{\bar{\mathbf{x}}}_i + \gamma_i \dot{\bar{\mathbf{x}}}_i - \boldsymbol{\nabla}_i U(\bar{\mathbf{x}}, \mathbf{y}) \right) + \left| \mathbf{F}_i^{\text{bias}}(\bar{\mathbf{x}}, \tau) \right|^2 \right] \sim 0 \tag{1.16}$$

where $\bar{\mathbf{x}} = \bar{x}(\tau) = (\bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_N)$ and all the other terms vanish exactly. In order to simplify Eq. (1.16) even further, one can exploit the saddle point equations of motion of the solvent atoms in the biased theory, requiring $\delta S_{\text{bias}} \sim 0$ [23]:

$$m_i \ddot{\bar{\mathbf{x}}}_i + \gamma_i \dot{\bar{\mathbf{x}}}_i + \boldsymbol{\nabla}_i U(\bar{\mathbf{x}}, \mathbf{y}) \sim \mathbf{F}_i^{\text{bias}}(\bar{\mathbf{x}}, t) \tag{1.17}$$

Plugging Eq. (1.17) into Eq. (1.16) one finally obtains

$$\frac{\delta T[\bar{x}]}{\delta \bar{x}} \equiv \frac{\delta}{\delta \bar{x}} \sum_{i=1}^{N} \gamma_i^{-1} \int_0^t d\tau \, |\mathbf{F}_i^{\text{bias}}(\bar{\mathbf{x}}, \tau)|^2 \sim 0 \tag{1.18}$$

where T is called the *penalty functional*. The condition in Eq. (1.18) states that the biased path which has the highest probability to be realized in the unbiased dynamics is the one for which the time-averaged squared modulus of the bias force is smaller. As previously anticipated, the penalty functional is just a function of the path $\bar{x}$ and it is not affected by the solvent fluctuations.

The method discussed here proposes to simulate pathways linking a configuration $\mathbf{x}_{\text{R}}$ to a configuration $\mathbf{x}_{\text{P}}$ by employing a time-dependent bias force and variationally select a posteriori which, among the many simulations started from $\mathbf{x}_{\text{R}}$, is the most realistic one. The variational nature of this approach makes its accuracy critically dependent on the choice of the model subspace in which the optimal solution is looked for, i.e. the space of biased pathways. In the following, we shall discuss how to choose a bias force which is sufficiently good to justify the use of a variational approximation.

1.2.1 The Ratchet-and-Pawl Force

Suppose we are provided with a good reaction coordinate $z(\mathbf{x})$, which is minimal in the product state $\mathbf{x}_{\text{P}}$ and maximal in the reactant state $\mathbf{x}_{\text{R}}$. A *smart* bias force can be built from $z(\mathbf{x})$ by asking for the following properties. The bias force should [133]:

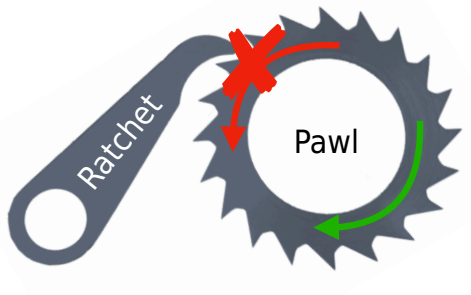

Figure 1.1: Representation of a ratchet-and-pawl system. The green arrow represents the direction of the allowed pawl rotation, while the red arrow represents the direction of the rotation forbidden by the action of the ratchet.

1. act in such a way to minimize the reaction coordinate;
2. vanish when the system spontaneously progresses towards the product along the reaction coordinate;
3. act only when the system tries to backtrack to higher values of the reaction coordinate.

A force that satisfies points 1. - 3. is given by:

$$\mathbf{F}_i^{\mathrm{rMD}}(\mathbf{x}, z_m) = -\frac{k_{\mathrm{rMD}}}{2}\boldsymbol{\nabla}_i z(\mathbf{x})(z(\mathbf{x}) - z_m(t))\theta(z(\mathbf{x}) - z_m(t)) \tag{1.19}$$

where k_{rMD} is an energy constant (units of kJ/mol), and the reaction coordinate is assumed to be of the kind

$$z(\mathbf{x}) = ||F(\mathbf{x}) - F^0||^2 \tag{1.20}$$

where $F(\mathbf{x})$ is a feature function extracted from the instantaneous value of the path $x(\tau) = \mathbf{x}$, $F^0 = F(\mathbf{x}_P)$ is the value of the feature in the product state, $||\cdot||$ is a norm in the corresponding feature space, and

$$z_m(t) = \min_{\tau \in [0,t]} z(x(\tau)) \tag{1.21}$$

is the smallest value reached by $z(x(\tau))$ up to time t. The force in Eq. (1.19) is called *ratchet-and-pawl force* [134, 21], and a dynamics implementing this additional force is called *ratchet-and-pawl molecular dynamics* (rMD):

$$\begin{aligned} m_i\ddot{\mathbf{x}}_i &= -\gamma_i\dot{\mathbf{x}}_i - \boldsymbol{\nabla}_i U(\mathbf{x},\mathbf{y}) + \mathbf{F}_i^{\mathrm{rMD}}(\mathbf{x}, z_m) + \boldsymbol{\eta}_i(t) \qquad & i = 1,\dots,N \\ \mu_j\ddot{\mathbf{y}}_j &= -\sigma_j\dot{\mathbf{y}}_j - \boldsymbol{\nabla}_j U(\mathbf{x},\mathbf{y}) + \boldsymbol{\xi}_j(t) & j = 1,\dots,M \end{aligned} \tag{1.22}$$

The biasing force in Eq. (1.19) is designed so that no work is done on the system in order to push it in a given direction, just like in the case of a ratchet-and-pawl mechanism (see Fig. 1.1), because every time the system spontaneously progresses

along $z(\mathbf{x})$ the force vanishes exactly. This is different from what happens, for example, in the case of steering and pulling forces, where the biasing force acts continuously along the whole trajectory. Just like in the case of steering algorithms, however, the starting and final configurations of the system should be known a-priori and cannot be computed from rMD.

rMD is completely determined by the choice of a ratchet constant k_{rMD} and the feature function $F(\mathbf{x})$. Thus, one should be careful in selecting the two, as a sub-optimal choice might generate highly unlikely pathways (see for example the supplementary material in [21]). The use of a soft biasing scheme, i.e. an harmonic one like in Eq. (1.19), should help reduce the artifacts introduced by a wrong choice of the feature (and thus of the reaction coordinate). The choice of the constant k_{rMD}, instead, should be guided by two important considerations. On the one hand, choosing a very small k_{rMD} evidently reduces rMD to conventional MD: In this case, the dynamics will be less affected by a sub-optimal choice of the reaction coordinate but will be also less accelerated by the biasing force. On the other hand, provided that z can represent only an approximation of the exact reaction coordinate, when high values of the ratchet constant are set, the dynamics will be greatly accelerated but will also lead to highly unlikely reaction pathways. For this reason, k_{rMD} should be chosen as a compromise between computational performance and physical meaning of the pathways.

The bias force in Eq. (1.19) has not a form $\mathbf{F}_i^{\mathrm{bias}}(\mathbf{x}, t)$, which is the one used in the proof of the variational principle in Eq. (1.18), but rather a form

$$\mathbf{F}_i^{\mathrm{bias}} = \mathbf{F}_i^{\mathrm{bias}}(\mathbf{x}, \xi(t)) \tag{1.23}$$

This difference introduces some modifications in the original path integral presented in Eq. (1.9). Indeed, if the dynamics of the auxiliary field $\xi(t)$ can be expressed as

$$\dot{\xi} = f(\xi) \tag{1.24}$$

then the path probability $P[x]$ in Eq. (1.9) becomes [23, 28]

$$\begin{aligned}\mathcal{P}[x] = \frac{1}{Z} \int \mathrm{d}\mathbf{y}_{\mathrm{F}} \int \mathrm{d}\mathbf{y}_{\mathrm{I}} \int_{\mathbf{y}_{\mathrm{I}}}^{\mathbf{y}_{\mathrm{F}}} \mathcal{D}y \int \mathcal{D}\xi \; e^{-\beta U(\mathbf{x}_{\mathrm{R}}, \mathbf{y}_{\mathrm{I}})} \cdot \\ \cdot \, e^{-S_{\mathrm{bias}}[x,y]} \delta \left(\xi(\tau) - \int_0^{\tau} d\tau' f(\xi(\tau')) \right) J[\bar{X}, \xi]\end{aligned} \tag{1.25}$$

where $J[\bar{x}, \xi]$ is a Jacobian which is introduced to make sure that the statistical weight of the paths is not affected by the measure $\mathcal{D}\xi$. It can be shown (see appendix B.1 and Ref. [28]) that $J[\bar{x}, \xi] = 1$ and that Eq. (1.18) holds unchanged also in the case of the path-integral in Eq. (1.25) [23]. This means that the rMD force can be used in conjunction with the variational condition in Eq. (1.18), and a possible algorithm combining the two approaches is the following:

1. prepare the system in an initial configuration $\mathbf{x}_{\mathrm{R}}$;
2. run N rMD simulations starting from $\mathbf{x}_{\mathrm{R}}$ towards a product configuration $\mathbf{x}_{\mathrm{P}}$;
3. apply the variational principle in Eq. (1.18) to the $M \leq N$ simulations that reached a configuration $\mathbf{x}_{\mathrm{F}}$ close to $\mathbf{x}_{\mathrm{P}}$, i.e. configurations for which $||\mathbf{x}_{\mathrm{P}} - \mathbf{x}_{\mathrm{F}}|| < \varepsilon$ with $||\cdot||$ a suitable metric and ε an arbitrarily small threshold;
4. among these M simulations, the one which minimizes the penalty function is called the *least biased* (LB) trajectory and provides the best variational guess of the reaction pathway between $\mathbf{x}_{\mathrm{R}}$ and $\mathbf{x}_{\mathrm{P}}$ generated via rMD.

Throughout the whole manuscript we will refer to the algorithm 1. - 4. as the *Bias Functional* (BF) method. It should be noted that the BF method is not exactly a variational algorithm, but rather it is inspired from a variational principle. Indeed, the minimum of the penalty function T is never found: It is approximated starting from a limited sample of bias trajectories, provided by the N trial rMD pathways.

1.2.2 Application to Protein Folding

Let us specialize now to the case of protein folding simulations. The reactant state of our system is provided by the unfolded state of the protein, while the product state is the native configuration of the molecule. As previously discussed, the choice of the feature function $F(\mathbf{x})$ employed to characterize the protein structure is crucial. A possible choice is the one proposed by the authors of Ref. [21]

$$F(\mathbf{x}) = C_{ij}(\mathbf{x}) = \begin{cases} \frac{1-\left(\frac{r_{ij}}{r_0}\right)^N}{1-\left(\frac{r_{ij}}{r_0}\right)^M} & \text{if } r_{ij} < r_c \\ \frac{N}{M} & \text{if } r_{ij} = r_c \\ 0 & \text{otherwise} \end{cases} \tag{1.26}$$

called the protein's *contact map*, where $r_{ij} = ||\mathbf{x}_i - \mathbf{x}_j||$ is the distance between the i-th and the j-th protein atom, $r_0 = 0.75$ nm is a reference distance that defines an atomic contact, $r_c = 1.2$ nm is a cutoff distance and $N < M$ are integers (usually set to $N = 6$ and $M = 10$). The natural norm in this feature space is provided by the *Frobenius norm*:

$$||F(x) - F^0||^2 = ||C(x) - C^0||^2_{\mathcal{N}} \equiv \frac{1}{\mathcal{N}} \sum_{i<j}^{N} \left(C_i j(x) - C^0_{ij}\right)^2 \tag{1.27}$$

where $\mathcal{N}$ is a normalization factor, here chosen to be $\mathcal{N} = 1$. This choice of feature was guided by the fact that a reaction coordinate based on a similar same smooth representation of the contact map was previously used in order to efficiently sample the free-energy landscape of the C-Terminal GB1 β-hairpin [135].

In the case of protein folding, the BF method can be better schematized in the following way (see also Fig. 1.2):

1. *Generation of the unfolded configuration*: the experimental structure of the native state (x-ray crystal or NMR structures) is used as a starting point of a short ($t \leq 3$ ns) high-temperature ($T > 700$ K) MD simulation. In this way, all the native contacts are destroyed ($Q < 0.1$) and the configuration doesn't preserve any knowledge of the original three-dimensional native structure. The final point of the high-temperature run is relaxed at the desired temperature (usually $T = 300$ K) and then used as a starting point for the folding simulation;

2. *Generation of the folding pathways*: given the initial condition, a fixed number of trial rMD trajectories are simulated at the temperature of interest and for a fixed amount of time. For each simulation, the rMD force is always defined employing protein contact maps as features. Some of the trajectories reach the native state of the protein (the so called *productive trajectories*), while other ones are not able to fold in the correct three-dimensional structure within the simulation time. These last trajectories should not be discarded: they provide important information on the presence of free energy barriers and, as we will discuss in great detail in Chapter 4, in particular cases they can be used to assess the misfolding propensity of a protein;

3. *Application of the variational principle*: the variational principle in Eq. (1.18) is finally applied to the ensemble of biased productive pathways in order to determine the LB trajectory, the one which has the highest probability to be realized in the absence of the bias.

Points 1. - 3. are then repeated in order to produce several LB trajectories. As previously discussed, when rMD is employed the initial and final points of the dynamics should be known a priori. This means that rMD is not able to predict the folded state of a protein, but rather its goal is to be able to predict the intermediate configurations connecting the unfolded to the native state.

The BF algorithm has been widely tested in recent years. In Ref. [23], the authors used it to simulate the folding of the WW-domain of FIP35 and the Villin Headpiece Domain, and systematically compared the results against unbiased MD folding simulations of the same proteins [10, 11]. The comparison highlighted that the algorithm is capable of reproducing the main features of the folding pathway of the two proteins and also to correctly predict the order in which the native contacts form during the folding process. The results of the BF method were also validated against experimental results: in a recent work [26], the algorithm was successfully

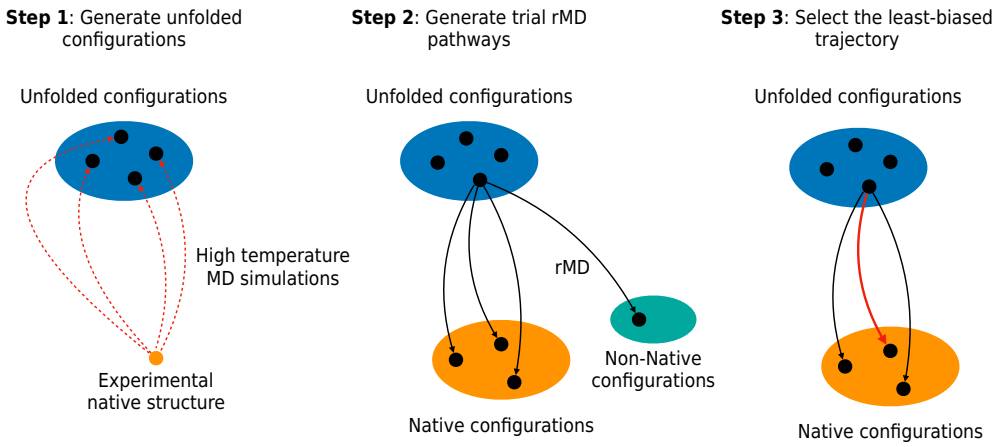

Figure 1.2: Sketch of the Biasf Functional algorithm applied to protein folding. In the first step, unfolded configurations are generated by means of high temperature MD simulations (dashed red lines); in the second step, many trial rMD simulations (solid lines) are started from the same initial configuration, some of which reach the native state and some of which do not; in the last step, among all the productive trajectories, the least-biased trajectory (solid red line) is chosen by means of the variational principle.

applied to characterize on the atomic-level the differences in the folding pathway between two similar ($\sim 60\%$ sequence similarity) proteins, IM7 and IM9. These two molecules approximately share the same native structure but show, respecitvely, a 3-state and a 2-state folding mechanism. Besides folding, the BF method was also used to study the latency transition of PAI-1 serpin [27].

The main advantage of the BF method is its extreme efficiency in generating folding pathways. The folding dynamics of a protein is usually accelerated by several order of magnitudes by the biasing force. To give just an explicit example, in chapter 4 we study a protein that needs approximately 10 minutes to fold: he BF method permits to simulate folding pathways of this protein in approximately $\sim$ 10 ns of simulation time, for an approximate acceleration of 11 orders of magnitude. This strong increase in simulation speed, however, comes at a price. The main problem resides in the fact that the rMD force is *history dependent*: indeed, the auxiliary field $z_m(t)$ depends on the previous values assumed by the reaction coordinate $z(\mathbf{x})$. The introduction of a history dependent bias force makes rMD dynamics not microscopically reversible, thus making it virtually impossible to retrieve kinetic observables, e.g. transition rates, directly from BF results.

A possible way to interpret the content of BF trajectories is to project them onto some qualitatively relevant collective variables $\{\theta_i\}$ and define the corresponding

kinetic free-energy landscape:

$$G_{\text{bias}}(\{\theta_i\}) = -k_{\text{B}}T \log P_{\text{bias}}(\{\theta_i\}) \tag{1.28}$$

where $P_{\text{bias}}(\{\theta_i\})$ is the frequency histogram of the configurations, in the collective variables space, visited by the simulations. G_{bias} should not be interpreted as an equilibrium free energy but can anyhow provide qualitative insight on the existence of metastable configurations along the folding pathway. Indeed, let us suppose we believe $z(\mathbf{x})$ to be a resonably good reaction coordinate: if so the protein would visit, in the correct order, all the metastable states along the folding pathway that would be visited by an unbiased simulation. This has been proved to hold for different protein systems [23, 21]. Moreover, since a bias force is acting on the protein, if kinetic free energy barriers are present even more so barriers would have been present in the unbiased free energy landscape. This qualitative argument tells us that G_{bias} provides a crude approximation of the true free energy profile, where barriers are systematically reduced and the metastable states are not correctly Boltzmann-weighted because of the presence of a history-dependent bias force. However, the reaction coordinate employed rMD calculations is basically arbitrary and even though the one proposed in Eq. (1.26) proved to perform well on a wide class of proteins (see e.g. Ref.s [26, 27, 24, 25], where BF results were successfully compared against experiments), it would be better to rely on some coordinate which is not chosen a priori and that assures an exhaustive sampling of the folding pathway.

The two main theoretical contributions of this manuscript go in the direction of overcoming the limitations we mentioned here. On the one hand, we will define a scheme that permits one to self-consistently optimize the rMD reaction coordinate; on the other hand, we will establish a strong connection between the latter self-consistent algorithm and *transition path theory* [30, 31, 136]. This will permit to manipulate simulation results in order to recover important kinetic quantities and to consistently describe the protein folding mechanism. These new theoretical tools make it clear that the BF algorithm should not be considered as an ultimate answer, but rather as a first approximation that is informative enough to admit a systematic optimization. In the next section we are going to start from the first aspect, which is the definition of a novel algorithm to iteratively improve the quality of the rMD reaction coordinate.

1.3 Deriving the Biased Dynamics from a Mean-Field Approximation

In this section we are going to present the basic theory behind the self-consistent optimization of the rMD reaction coordinate. Briefly, the idea behind it is to recover a particular form of rMD from a *mean-field* approximation of the exact Langevin dynamics. This new rMD dynamics will depend on two forces, biasing along the so-called *path variables* [137]. The peculiar feature of the path variables emerging from this calculation, however, is that they do not depend on a predefined path, but rather they depend on an *average path* which is obtained from a previous simulation. This naturally defines a self-consistent scheme, were the average path is recomputed at each iteration and it is passed to the next one, until convergence is attained. We will test this approach on a 2-dimensional toy model and we will finally apply it to the folding of the WW-domain of FIP35, a benchmark protein previously studied using unbiased MD simulations [10, 11, 105, 138].

1.3.1 Mean-Field Representation of Langevin Dynamics

Let us introduces two dummy variables $s_m(\tau)$ and $w_m(\tau)$ in the conditional probability in Eq. (1.3):

$$\begin{aligned} p(\mathbf{x}_\mathrm{P}, t|\mathbf{x}_\mathrm{R}, 0) = \frac{1}{Z}\int \mathrm{d}\mathbf{y}_\mathrm{F} \int \mathrm{d}\mathbf{y}_\mathrm{I} \int_{\mathbf{y}_\mathrm{I}}^{\mathbf{y}_\mathrm{F}} \mathcal{D}y \int_{\mathbf{x}_\mathrm{R}}^{\mathbf{x}_\mathrm{P}} \mathcal{D}x \; e^{-S_{\mathrm{OM}}[x,y]} e^{-\beta U(\mathbf{x}_\mathrm{R},\mathbf{y}_\mathrm{I})} \cdot \\ \cdot \int_{\bar{w}(0)} \mathcal{D}w_m \; \delta\left(w_m - \int_0^\tau d\tau' \dot{\bar{w}}(\tau')\theta(w_m(\tau') - \bar{w}(\tau'))\right) \cdot \\ \cdot \int_{\bar{s}(0)} \mathcal{D}s_m \; \delta\left(s_m - \int_0^\tau d\tau' \dot{\bar{s}}(\tau')\theta(s_m(\tau') - \bar{s}(\tau'))\right) \end{aligned} \tag{1.29}$$

The dynamics of the two dummy variables is controlled by the one of two additional time-dependent functions, $\bar{s}(\tau)$ and $\bar{w}(\tau)$:

$$\begin{aligned} \dot{s}_m(\tau) &= \dot{\bar{s}}(\tau)\theta(s_m(\tau) - \bar{s}(\tau)) \\ \dot{w}_m(\tau) &= \dot{\bar{w}}(\tau)\theta(w_m(\tau) - \bar{w}(\tau)) \end{aligned} \tag{1.30}$$

Eq. (1.30) shows that the dynamics of $s_m(\tau)$ and $w_m(\tau)$ remains frozen, respectively, every time $s_m(\tau) > \bar{s}(\tau)$ and $w_m(\tau) > \bar{w}(\tau)$. Thus, if we employ initial conditions

$$\begin{aligned} s_m(0) &= \bar{s}(0) \\ w_m(0) &= \bar{w}(0) \end{aligned} \tag{1.31}$$

$s_m(\tau)$ and $w_m(\tau)$ will represent the minimum values, respectively, of $\bar{s}(\tau)$ and $\bar{w}(\tau)$ up to time τ. As previously discussed in section 1.2 and proved in appendix B.1, the introduction of $s_m(\tau)$ and $w_m(\tau)$ in the path integral account for a trivial Jacobian

$$J\left[w_m, X\right] = J\left[s_m, X\right] = 1 \tag{1.32}$$

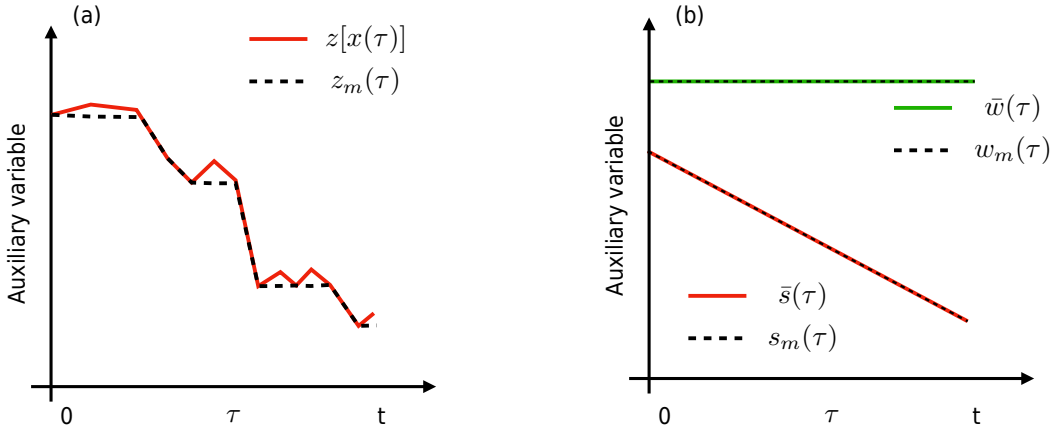

Figure 1.3: Adapted from Ref. [28], with the permission of AIP Publishing. Sketch of the different dynamics of the auxiliary variables introduced in (a) rMD and (b) in the derivation of the self consistent algorithm.

so that no modification with respect to the original functional measure in the path integral is performed.

It is worth specifying that the path integral in Eq. (1.29) remains exact independently on the specific functional form of the two functions $\bar{s}(\tau)$ and $\bar{w}(\tau)$. Let us then specialize to a specific choice for the additional functions:

$$\begin{aligned} \bar{s}(\tau) &= 1 - \frac{\tau}{t} \\ \bar{w}(\tau) &= w_0 \end{aligned} \tag{1.33}$$

Both of them are non-increasing, so for every value of $\tau \in [0, t]$ we have (see Fig. 1.3)

$$\begin{aligned} s_m(\tau) &= \bar{s}(\tau) \\ w_m(\tau) &= \bar{w}(\tau) \end{aligned} \tag{1.34}$$

Notably, the functions in Eq. (1.33) can be obtained via two useful limits:

$$\begin{aligned} \bar{s}(\tau) &= \lim_{\lambda\to\infty} s_\lambda[x,\tau] \equiv 1 - \frac{1}{t}\lim_{\lambda\to\infty} \frac{\int_0^t dt\, t' e^{-\lambda||F(\tau)-F(t')||^2}}{\int_0^t dt'\, e^{-\lambda||F(\tau)-F(t')||^2}} \\ \bar{w}(\tau) &= \lim_{\lambda\to\infty} w_\lambda[x,\tau] \equiv w_0 - \lim_{\lambda\to\infty} \frac{1}{\lambda}\log \int_0^t dt'\, e^{-\lambda||F(\tau)-F(t')||^2} \end{aligned} \tag{1.35}$$

where F is a generic feature describing the system at a given time, $F(\tau) = F(x(\tau))$. To prove Eq. (1.35), let us discretize the interval $[0, t]$ in N slices with step $\Delta t = t/N$. Calling $\tau = l\Delta t$ the intermediate steps, for s_λ we have

$$\begin{aligned} s_\lambda[x, l\Delta t] &= 1 - \frac{1}{N\Delta t}\frac{\sum_{k=1}^N \Delta t^2\, k e^{-\lambda||F(l\Delta t)-F(k\Delta t)||^2}}{\sum_{k=1}^N \Delta t\, e^{-\lambda||F(l\Delta t)-F(k\Delta t)||^2}} \\ &= 1 - \frac{1}{N}\frac{\sum_{k=1}^N k e^{-\lambda||F(l\Delta t)-F(k\Delta t)||^2}}{\sum_{k=1}^N e^{-\lambda||F(l\Delta t)-F(k\Delta t)||^2}} \end{aligned} \tag{1.36}$$

For $\lambda \to \infty$, only $l = k$ terms in the sum survive, because

$$\lim_{\lambda\to\infty} \lambda||F(l\Delta t) - F(k\Delta t)||^2 = \begin{cases} \infty \text{ if } k \neq l \\ 0 \text{ if } k = l \end{cases} \tag{1.37}$$

Thus we have

$$\lim_{\lambda\to\infty} s_\lambda[x, l\Delta t] = 1 - \frac{l}{N} = 1 - \frac{\tau}{\Delta t}\frac{\Delta t}{t} = 1 - \frac{\tau}{t} \tag{1.38}$$

which is exactly the upper equation in Eq. (1.33). For $w_\lambda[X, \tau]$ the procedure is almost identical. Let us discretize the integral inside the logarithm to obtain:

$$\begin{aligned} w_\lambda[x, \tau] &= w_0 - \frac{1}{\lambda}\log\sum_{k=1}^{N} \Delta t e^{-\lambda||F(k\Delta t)-F(l\Delta t)||^2} \\ &= \left(w_0 - \frac{1}{\lambda}\log\Delta t\right) - \frac{1}{\lambda}\log\sum_{k=1}^{N} e^{-\lambda||F(k\Delta t)-F(l\Delta t)||^2} \end{aligned} \tag{1.39}$$

If we use Eq. (1.37) in Eq. (1.39), we are left with

$$\lim_{\lambda\to\infty} w_\lambda[x, \tau] = w_0 - \lim_{\lambda\to\infty}\frac{1}{\lambda}\log\Delta t = w_0 \tag{1.40}$$

which is, again, the expected result reported in Eq. (1.33).

Let us employ Eq. (1.35) to rewrite the conditional probability in Eq. (1.29) as

$$\begin{aligned} p(X_\mathrm{P}, t|X_\mathrm{R}, 0) &= \lim_{\lambda\to\infty} p_\lambda(X_\mathrm{P}, t|X_\mathrm{R}, 0) \\ &= \lim_{\lambda\to\infty}\frac{1}{Z}\int \mathrm{d}\mathbf{y}_\mathrm{F}\int \mathrm{d}\mathbf{y}_\mathrm{I}\int_{\mathbf{y}_\mathrm{I}}^{\mathbf{y}_\mathrm{F}} \mathcal{D}y\int_{\mathbf{x}_\mathrm{R}}^{\mathbf{x}_\mathrm{P}} \mathcal{D}x e^{-S_\lambda[x,y,s_m,w_m]} e^{-\beta U(\mathbf{x}_\mathrm{R},\mathbf{y}_\mathrm{I})}\cdot \\ &\cdot \int_{\bar{w}(0)} \mathcal{D}w_m\,\delta\left(w_m - \int_0^\tau d\tau' \dot{w}_\lambda(\tau')\theta(w_m(\tau') - w_\lambda(\tau'))\right)\cdot \\ &\cdot \int_{\bar{s}(0)} \mathcal{D}s_m\,\delta\left(s_m - \int_0^\tau d\tau' \dot{s}_\lambda(\tau')\theta(s_m(\tau') - s_\lambda(\tau'))\right) \end{aligned} \tag{1.41}$$

One should note that the action in Eq. (1.41) is now explicitly dependent on the dummy variables through two new force terms:

$$\begin{aligned} S_\lambda[x, y, s_m, w_m] \equiv \frac{\beta}{4}\int_0^t d\tau \Bigg[&\sum_{i=1}^{N}\gamma_i^{-1}\left(m_i\ddot{\mathbf{x}}_i + \gamma_i\dot{\mathbf{x}}_i + \boldsymbol{\nabla}_i U(\mathbf{x}, \mathbf{y})\right. \\ &\left. - \mathbf{F}_i^s[x, s_m] - \mathbf{F}_i^w[x, w_m]\right)^2 \\ &+ \sum_{j=1}^{M}\sigma_j^{-1}\left(\mu_j\ddot{\mathbf{y}}_j + \sigma_j\dot{\mathbf{y}}_j + \boldsymbol{\nabla}_j U(\mathbf{x}, \mathbf{y})\right)^2\Bigg] \end{aligned} \tag{1.42}$$

The exact functional form of the forces is arbitrary $\mathbf{F}_i^s[x, s_m]$ and $\mathbf{F}_i^w[x, w_m]$, as long as they exactly vanish in the limit $\lambda \to \infty$. In particular, they can be expressed as

$$\mathbf{F}_i^\alpha[x, \alpha_m] = \zeta[x, \alpha_\lambda]\xi(\alpha_\lambda[x, t] - \alpha_m(t)) \qquad \alpha = s, w \tag{1.43}$$

where

$$\lim_{\lambda\to\infty} \xi(\alpha_\lambda[x,t] - \alpha_m(t)) = 0 \qquad \alpha = s, w \tag{1.44}$$

and $\zeta[x, \alpha_\lambda, \alpha_m]$ is a generic functional of the path and the dummy variables. Evidently, a possible choice for this forces is provided by a ratchet force:

$$\mathbf{F}_i^\alpha[x, \alpha_m] = -k_\alpha \boldsymbol{\nabla}\alpha_\lambda[x,t](\alpha_\lambda[x,t] - \alpha_m(t))\theta(\alpha_\lambda[x,t] - \alpha_m(t)) \qquad \alpha = s, w \tag{1.45}$$

where

$$\zeta[x, \alpha_\lambda, \alpha_m] = -k_\alpha \boldsymbol{\nabla}\alpha_\lambda[x,t] \qquad \alpha = s, w \tag{1.46}$$

and

$$\begin{aligned}\lim_{\lambda\to\infty} \xi(\alpha_\lambda[x,t] - \alpha_m(t)) &= \lim_{\lambda\to\infty} \theta(s_\lambda[x,t] - s_m(t))(s_\lambda[X,t] - s_m(t)) \\ &= \lim_{\lambda\to\infty} \theta(w_\lambda[x,t] - w_m(t))(w_\lambda[X,t] - w_m(t)) \\ &= 0 \qquad \alpha = s, w\end{aligned} \tag{1.47}$$

Given the way in which the forces in Eq. (1.43) are defined, the original path integral in Eq. (1.3) and the one in Eq. (1.41) are exactly equivalent. Let us now introduce a *mean-field approximation*: we substitute the values of $F(t')$ in Eq. (1.35) with their average values $\langle F(t')\rangle$:

$$\langle F(t')\rangle = \frac{\int \mathrm{d}\mathbf{y}_\mathrm{F} \int \mathrm{d}\mathbf{y}_\mathrm{I} \int_{\mathbf{y}_\mathrm{I}}^{\mathbf{y}_\mathrm{F}} \mathcal{D}y \int_{\mathbf{x}_\mathrm{R}}^{\mathbf{x}_\mathrm{P}} \mathcal{D}x\ F(x(t'))e^{-S_\lambda[x,y]}e^{-\beta U(\mathbf{x}_\mathrm{R},\mathbf{y}_\mathrm{I})}}{\int \mathrm{d}\mathbf{y}_\mathrm{F} \int \mathrm{d}\mathbf{y}_\mathrm{I} \int_{\mathbf{y}_\mathrm{I}}^{\mathbf{y}_\mathrm{F}} \mathcal{D}y \int_{\mathbf{x}_\mathrm{R}}^{\mathbf{x}_\mathrm{P}} \mathcal{D}x\ e^{-S_\lambda[x,y]}e^{-\beta U(\mathbf{x}_\mathrm{R},\mathbf{y}_\mathrm{I})}} \tag{1.48}$$

so that

$$\begin{aligned} s_\lambda[x,\tau] \to s_\lambda(x(\tau)) &= 1 - \frac{1}{t}\frac{\int_0^t dt\ t' e^{-\lambda||F(\tau)-\langle F(t')\rangle||^2}}{\int_0^t dt'\ e^{-\lambda||F(\tau)-\langle F(t')\rangle||^2}} \\ w_\lambda[x,\tau] \to w_\lambda(x(\tau)) &= w_0 - \frac{1}{\lambda}\log\int_0^t dt\ t' e^{-\lambda||F(\tau)-\langle F(t')\rangle||^2} \end{aligned} \tag{1.49}$$

With this replacement, the two variables stop depending functionally on the path x and become standard collective coordinates, which means they just depend on temporal realizations of the path $x(\tau)$. Owing to the mean-field approximation, even in the limit $\lambda \to \infty$ the dynamics is no longer exact, since the forces $\mathbf{F}_i^s$ and $\mathbf{F}_i^w$ no longer vanish. Instead, they define a new type of ratchet-and-pawl dynamics. Now the rMD force biases along two collective variables which depend on a pre-computed average path that can be systematically improved, within the mean-field approximation, through a self-consistent procedure. The two collective variables s_λ and w_λ are called *path* (or *tube*) *variables*, and were first introduced in Ref. [137]. A non trivial result of our derivation is that ratchet-like forces (forces that satisfy Eq.s (1.43) and (1.44)) along these two collective variables emerge naturally from a mean-field

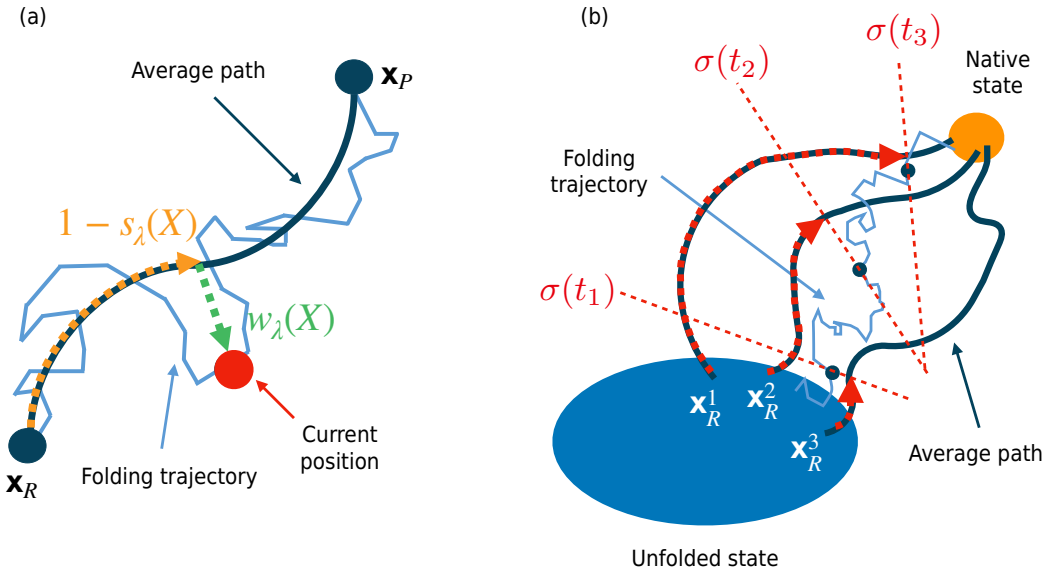

Figure 1.4: Adapted from Ref. [28], with the permission of AIP Publishing. Geometrical interpretation of (a) s_λ, w_λ and (b) σ variables.

approximation of the exact Langevin dynamics. Thus, any other predetermined collective variable employed in rMD calculations is necessarily sub-optimal with respect to the choice of the path variables. In chapter 2 we will further investigate this point and we will show that the reason why s_λ and w_λ have such interesting properties is hidden behind a deep connection with the committor function.

The systematic error coming from the mean-field approximation can be minimized by resorting to the variational principle of the BF approach, Eq. (1.18). Indeed, among all the productive pathways $\bar{x}$ computed in presence of the two biasing forces $\mathbf{F}_i^s$ and $\mathbf{F}_i^w$, the one which has the highest probability to be realized in the *absence* of the bias is the one which satisfies the following equality:

$$\frac{\delta}{\delta \bar{x}} \sum_{i=1}^{N} \gamma_i^{-1} \int_0^t d\tau \, |\mathbf{F}_i^s(\bar{x}(\tau)) + \mathbf{F}_i^w(\bar{x}(\tau))|^2 \sim 0 \tag{1.50}$$

1.3.2 Self-Consistent Path Sampling Algorithm

Given the theoretical considerations presented in section (1.3.1), we propose the following algorithm to sample protein folding pathways, which we shall call *Self-Consistent Path Sampling* (SCPS):

1. *Generation of the unfolded configuration*: the initial configuration $\mathbf{x}_R$ is generated in the same way presented in section (1.2.2), point 1.;

2. *Generation of the guess average path*: given a total simulation time t, run many independent rMD simulations starting from $\mathbf{x}_R$. Isolate the produc-

tive ones and use them to compute the guess average path in feature space $\{\langle F(\tau)\rangle\}^0_{\tau\in[0,t]}$;

3. *Iteration*: use the set $\{\langle F(\tau)\rangle\}^0_{\tau\in[0,t]}$ of average features to run a new ensemble of folding trajectories, now employing $\mathbf{F}^s_i$ and $\mathbf{F}^w_i$ as biasing forces and choosing a large value for the parameter λ. From this new ensemble, isolate the productive trajectories and use them to compute the new set $\{\langle F(\tau)\rangle\}^1_{\tau\in[0,t]}$.

4. *Convergence*: point 3. is repeated until, at the n-th iteration, a convergence criterion is met;

5. *Application of the variational principle*: the variational principle in Eq. (1.50) is finally applied to the converged ensemble of productive pathways in order to determine the least biased trajectory.

Steps 1.-5. are repeated to generate many independent LB trajectories starting from U different unfolded conformations $\mathbf{x}^1_R, \ldots, \mathbf{x}^U_R$. Given the results of all these independent folding calculations, it is possible to define a reaction coordinate $\sigma(\mathbf{x})$ that uses the information coming from all the simulations to measure the overall progress of folding. This can be done by combining the tube variables $s^1_\lambda, \ldots, s^U_\lambda$ obtained from each least biased trajectory and compute

$$\sigma(\mathbf{x}) = \frac{1}{Nt}\sum_{k=1}^{U} \frac{\int_0^t d\tau\ \tau e^{-\lambda||F(x)-\langle F(\tau')\rangle_k||^2}}{\int_0^t d\tau\ e^{-\lambda||F(x)-\langle F(\tau')\rangle^k||^2}} \tag{1.51}$$

where $\langle F(\tau')\rangle_k$ is the path obtained by averaging the folding trajectories started in $\mathbf{x}^k_R$. To better understand the physical meaning behind Eq. (1.51), it is worth discussing the geometrical interpretation of the $s_\lambda(\mathbf{x})$ and $w_\lambda(\mathbf{x})$ variables. $s_\lambda(\mathbf{x})$ measures the progress of the folding reaction along the average path in feature space, while instead $w_\lambda(\mathbf{x})$ measures the distance of $\mathbf{x}$ from the average path (see Fig. 1.4 (a)). Thus, $\sigma(\mathbf{x})$ assigns to each point on a folding path the closest point among the ones on the U average pathways, as shown in Fig. 1.4.

The algorithm leaves some freedom in the choice of a whole set of parameters and hyper-parameters: the specific choice of the features F and the norm $||\cdot||$ used to compute the distance in feature space, the way in which the average path is computed and the value of λ. The choice of the features used to represent a protein configuration is in principle arbitrary, but some of them can lead to better folding performances than others. To give an example, using the full configuration space as feature space and computing the distances between configurations by means of their RMSD proved to be an inefficient strategy that didn't lead to successful folding events[2]. A better choice (the same one that we adopted in the original work [28])

[2]We thank Carlo Camilloni for this observation.

is to consider contact maps as features and use the Frobenius norm as a distance between contact maps

$$||F(x) - \langle F(\tau')\rangle_k||^2 = ||C(x) - \langle C(\tau')\rangle^k||^2_{||C^0||^2} \tag{1.52}$$

where $!!C^0!!^2$, the squared norm of the native contact map, is used to normalize the distance. Once the features have been defined, one should compute the average path in that space. There is a relevant problem concerning this point: the algorithm doesn't scale well with the number of points in the average path[3]. For this reason, first the average path has to be computed as the time-average over all the folding pathways starting from the same initial condition $\mathbf{x}_R^k$. Then one should define a way to systematically downsample it in order to feed it to the next iteration. There are two possible ways to do so: the first one is to select only those points in the path that separated by a fixed distance d, while the second one (which is the strategy adopted in the original paper [28]) consists in selecting points in the path that are equispaced in time. The two approaches show comparable results on a simple 2d-model (compare the results reported in section 1.4 with the ones reported in Ref. [29]), but the first one has a nice advantage: in that case, the λ parameter can be simply defined as

$$\lambda = \frac{1}{d^2} \tag{1.53}$$

The choice in Eq. (1.53) approximately satisfies the limit $\lambda \to \infty$, required in (1.41), because the protein feels the dependence on a very small number of points in the average path at the time. In the second strategy, instead, there is no evident way in which λ should be defined and indeed in the original paper [28] the choice was made following an empirical argument. Because the algorithm requires λ to be set in such a way that $\lambda \gg 1$, one can look at behavior of the function

$$\xi(t, \tau) = \exp(-\lambda||C(x(t)) - \langle C(\tau)\rangle||^2) \tag{1.54}$$

as the time t evolves, see how many points in the average path contribute to the exponential and set λ in such a way to approximately minimize this number for every t.

In the illustrative application in section 1.4, the average path has been constructed by fixing a distance d and selecting only equidistant points in the average path. Differently, the results in section 1.5 have been obtained by considering points in the average path which are equidistant in time.

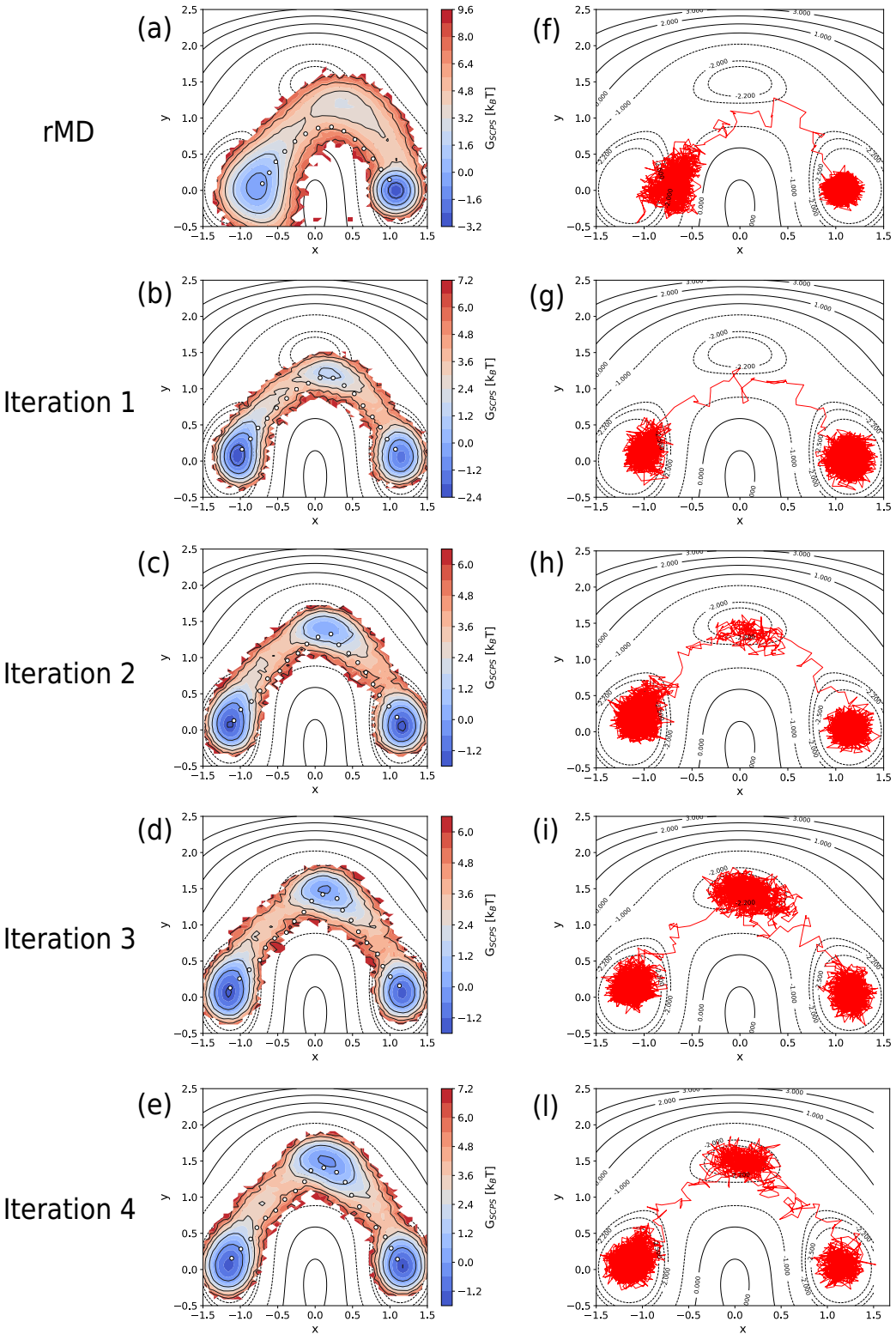

Figure 1.5: (a,b,c,d,e) Evolution of the G_{SCPS} biased free energy for increasing SCPS iterations. White dots correspond to the average path obtained from the corresponding iteration; (f,g,h,i,l) typical trajectories obtained from the corresponding iterations. In all the pictures, background contour lines represent the potential in Eq. (1.55).

1.4 Illustrative Application on a 2D Potential

To validate the SCPS alogorithm, it is useful to start from a simple example where the system's configuration space is under control and the features of problem can be thoroughly investigated. We considered as an illustrative example the diffusion of a particle in a 2D potential provided by

$$U(x,y) = u_0 \left[e^{-(x^2+y^2)} - a_0 e^{-(x^2+(y-y_0)^2)} - e^{-((x-x_0)^2+y^2)} - e^{-((x+x_0)^2+y^2)} \right] + u_1 \left[x^4 + (y-y_1)^4 \right] \tag{1.55}$$

where $x_0 = 1$, $a_0 = \frac{3}{5}$, $y_0 = \frac{5}{3}$, $u_1 = \frac{1}{5}$ and $y_1 = \frac{1}{3}$ in the appropriate units. This potential contains three minima, respectrively centered at $\mathbf{x}_R = (x_R, y_R) = (-1, 0)$, $\mathbf{x}_I = (x_I, y_I) = (0, 1.5)$ and $\mathbf{x}_P = (x_P, y_P) = (1, 0)$, and can be considered as a simplified version of the one studied in Ref. [139]. The reactant R and product P regions are defined, respectively, by the points around $\mathbf{x}_R$ and $\mathbf{x}_P$ for which $U(x,y) \leq -2.5 k_B T$. At low temperature, transitions from R to P and vice-versa occur through a single channel, which is the one passing from the intermediate state around $\mathbf{x}_I$. Thus we chose to study the system at $k_\mathrm{B}T = 0.15$.

The goal of this section is the following: we want to provide the SCPS algorithm a guess average path that contains systematic errors, and see if the iterative scheme is able to correct for them so to provide a faithful characterization of the $R \to P$ transition.

1.4.1 Generation of the Trial Pathway

As discussed above, we want to define an initial set of pathways that do not fully capture the nature of the reaction occurring between R and P. To do this, we can use rMD along an ill-defined reaction coordinate and a strong bias constant. As a reaction coordinate we used

$$d(\mathbf{x}) = ||\mathbf{x} - \mathbf{x}_P|| \tag{1.56}$$

which measures the instantaneous Euclidean distance of a point $\mathbf{x}$ to the product state, while the ratchet constant was set to $k_R = 50$. The choice of the ratchet constant ensures that the biasing force is approximately twice as big as the physical force at any time during the simulation. We integrated the rMD equations of motion with a time step of $dt = 0.02$ for a total time of $t = 4 \times 10^3 dt$ and we ran 1000 independent simulations starting from $\mathbf{x}_R$. All the simulations coverged to the product state, measured as $d(\mathbf{x}) < 0.01$, within time t and were thus used to compute

[3] This is because the heavy computation of the differences at the exponent of Eq. (1.49) has to be done for every contact map at each time step in the simulation.

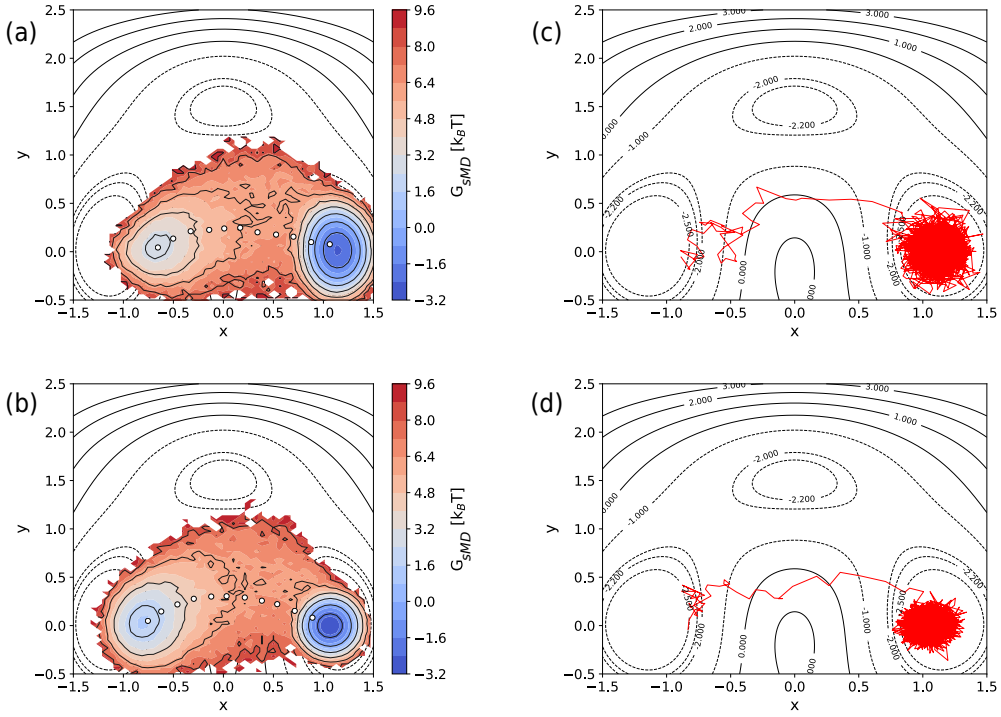

Figure 1.6: Biased free energies obtained using steered MD with bias force (a) $\mathbf{F}_1 = k_1 \boldsymbol{\nabla} d(\mathbf{x})$ and (b) $\mathbf{F}_2 = k_2 d(\mathbf{x}) \boldsymbol{\nabla} d(\mathbf{x})$. In both cases, white points represent the corresponding average path; (c,d) typical trajectories obtained, respectively, using the biasing forces in (a) and (b).

the guess average path $\langle x(\tau)\rangle^0_{\tau\in[0,t]}$. Some technicalities about the calculation of the average path are in order, and are discussed in detail in appendix B.2. The results of this calculation are reported in Fig. 1.5 (a) and (f). Fig. (1.5) (a) shows the distribution of the points sampled by rMD, while Fig. 1.5 (f) reports a typical rMD trajectory. In spite of the very high biasing force employed in the calculation, rMD data reproduce at a qualitative level the main features of the transition paths ensemble. Still, the results in Fig. 1.5 (a) displays evident systematic errors: the data distribution doesn't reflect the symmetry of the underlying potential, because of the strong push towards the product state provided by the biasing force; the existence of an intermediate state does not emerge because, as clearly highlighted by Fig. 1.5 (b), typical rMD trajectories neglect it; finally, the average path doesn't intersect the intermediate state around $\mathbf{x}_I$.

A few comments about the guess average pathway are in order here. Even though $\langle x(\tau)\rangle^0_{\tau\in[0,t]}$ fails to provide several important features of the reaction mechanism, the rMD average path is good enough to be successfully employed as an initial ingredient of the SCPS algorithm. This is not true, for example, for pathways obtained from steered MD (sMD), even if the forces employed in the simulations are comparable in intensity with the rMD one. This can be seen from Fig. 1.6 (a) and (b), where it is clear that the steered dynamics is ignoring the correct reaction pathway. Indeed, typical trajectories (see Fig. 1.6 (c) and (d)) overcome the barrier instead of passing around it. As a consequence, the first SCPS iteration performed with the steered average path doesn't result in any productive transition.

1.4.2 Self-Consistent Iterations

We used the average rMD path $\langle x(\tau)\rangle^0_{\tau\in[0,t]}$ to run the first self-consistent iteration. We ran 5000 simulations, biased along s_λ and w_λ, with a time step $dt = 0.02$ and for a total time of $t = 4 \times 10^4 dt$. We set $k_s = k_w = 50$ and $\lambda = d^{-2} = 30$. Fig. 1.5 (b) and (g) show the results of this first iteration. The reactive channel is better captured with respect to the rMD case and the trajectories start to accumulate in the vicinity of the I state. These features become more evident going on with the iterations. Eventually, after the third iteration (see Fig. 1.5 (d) and (i)), the reactive channel is correctly characterized and the average path shows all the desired properties: it is approximately symmetric and it connects the three basins in the potential. A further iteration, Fig. 1.5 (e) and (l), does not substantially change the picture.

Being SCPS a self-consistent procedure, one should worry about assessing its convergence. A possible way to check for convergence is to measure wether the L_2-norm between the n-th iteration's average path and the one at iteration $n+1$ is

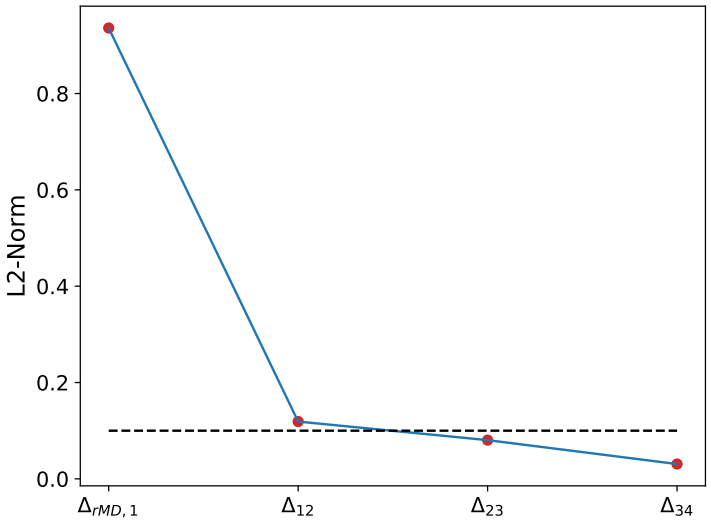

Figure 1.7: Convergence of the average pathway as a function of the number of SCPS iterations, estimated using Eq. (1.57) with $\varepsilon = 0.1$. $\Delta_{i,i+1}$ in the x-axis refer to the comparison between iterations i and $i+1$.

smaller than a given threshold ε:

$$||\langle \mathbf{x}(\tau) \rangle^n - \langle \mathbf{x}(\tau) \rangle^{n+1}||^2 < \varepsilon \tag{1.57}$$

In this simple example we can set $\varepsilon = 0.1$, and from Fig. (1.10) one can notice that after the third iteration convergence has been reached. In this method, the choice of the threshold is rather arbitrary and should be considered as a trade-off between the computational time that one wants to invest in the simulations and the quality of the resulting reaction pathway.

1.4.3 Reaction pathway in the s_λ - w_λ space

For high dimensional systems it is in general not trivial to find a relevant subspace on which to project the dynamics of the system in order to obtain information on the existence of metastable states. For this reason, it is interesting to investigate if the two variables s_λ and w_λ can provide a reasonable description of the reaction channel. Of course for such a simple system, projecting the configurations onto the s_λ - w_λ space doesn't amount for a dimensional reduction, but it is still a useful exercise which could be helpful in understanding more complex situations.

Fig. 1.8 shows the kinetic free energy G_{SCPS} in the s_λ - w_λ space of the system for the first and the last SCPS iterations. Three features of the plot immediately emerge: three basins are always present; the free energy lanscape of the last iteration is approximately symmetric around $s_\lambda = 0.5$; the fluctuations along the average path, governed by w_λ, decrease for increasing iterations. The first two features

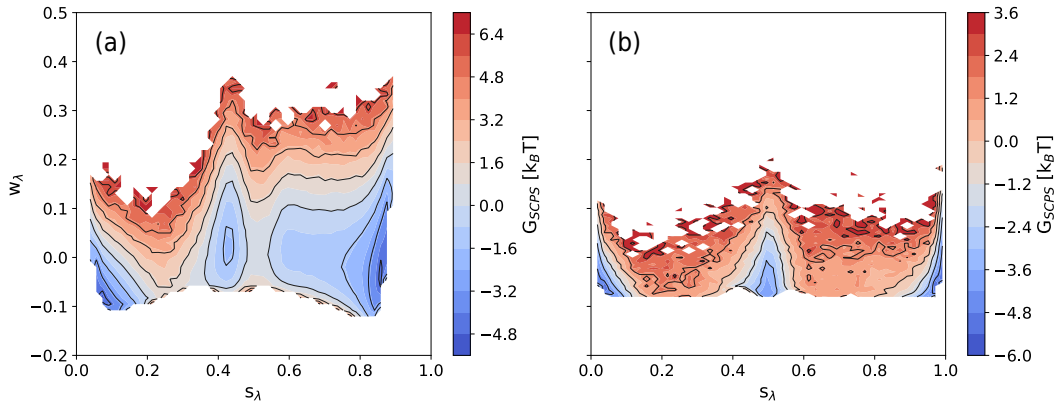

Figure 1.8: Kinetic free energy G_{SCPS} represented in the s_λ-w_λ space in the case of the (a) first and (b) fourth SCPS iterations.

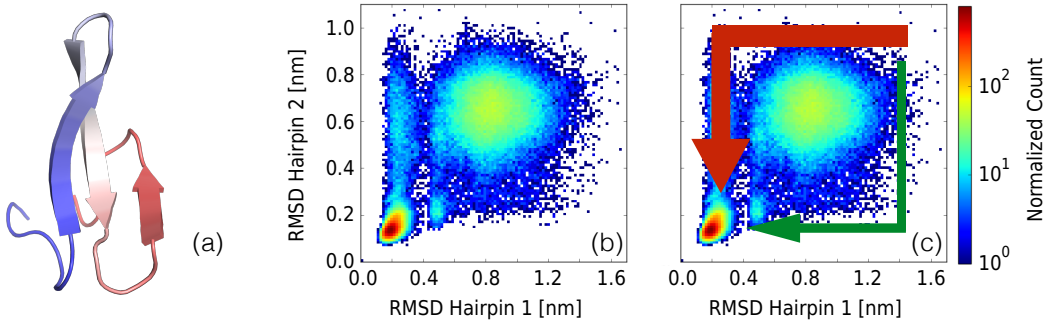

Figure 1.9: (a) Crystal structure of the WW-Domain of FIP35. Hairpin 1 is formed by the blue and white β-strands, while hairpin 2 is formed by the white and blue ones; (b) configurations of the WW-Domain of FIP35, obtained from 2 independent 100 μs long MD simulations [9], projected on the RMSD to native of the two hairpins; (c) schematic representation of the main folding pathway (red arrow), where hairpin 1 is formed first, and the sub-dominant pathway (green arrow) where the hairpin 2 is formed at the initial stages in the simulation.

highlight the fact that s_λ and w_λ are able to capture the essential properties of the underlying potential, together with an intriguing similarity with the committor function ($s_\lambda \sim 0.5$ in the correspondence of the intermediate state) which will be better exploited and deeply explored in chapter 2. The third feature is a robust check of the algorithm's stability: the fact that fluctuations around the average path are decreasing for increasing iterations means that the different self-consistent steps of the algorithm are able to define an increasingly better average path which, ultimately, reflects all the relevant information encoded in the potential.

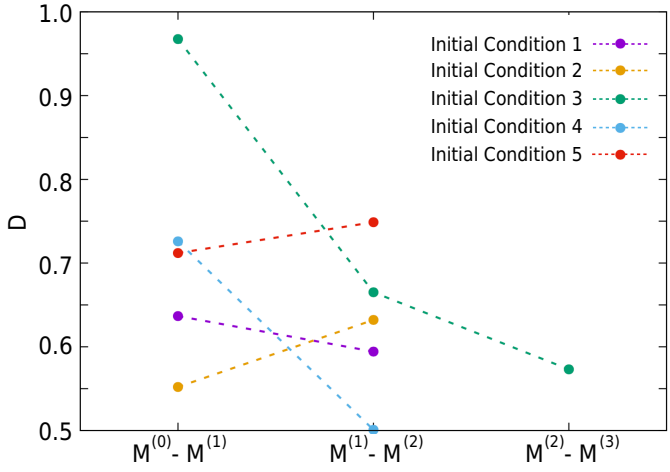

Figure 1.10: Adapted from Ref. [28], with the permission of AIP Publishing. Convergence of the SCPS simulations starting from 5 independent initial conditions

1.5 Application to the Folding of WW-Domain of FIP35

The WW-domain of FIP35 (shortly FIP35) is a 35 amino acid protein composed by two hairpins forming a 3-stranded β-sheet. Its crystal structure is reported in Fig. 1.9 (a). It is a fast-folding protein, which folds in the timescale of microseconds [140]. It has been widely studied both in implicit [105] and in explicit solvent [10, 11]. Simulations in both the solvent models agree on the fact that folding of FIP35 proceeds along two main pathways: a dominant one (*pathway 1*), where hairpin 1 is formed first and then the formation of hairpin 2 follows, and a sub-dominant one (*pathway 2*) where hairpins form in the reversed order (see Fig.s 1.9 (b) and (c)). Further analyses of the simulations [141] also revealed the existence of an intermediate state along pathway 1, where the first hairpin is completed and the second one has yet to be formed.

Because of the wide computational and experimental information about it, this protein constitutes a perfect benchmark for the SCPS algorithm. Thus, this section will be devoted to the analysis of the algorithm performance and the comparison of its results with the ones of standard rMD.

1.5.1 Simulations Setup and Convergence

SCPS simulations of FIP35 have been setup as follows:

1. starting from the energy minimized crystal structure, we generated 5 unfolded conditions by running 5 independent MD simulations at $T = 800$ K, each one lasting 100 ps;

2. for each initial condition, the guess folding pathways have been computed by

running 20 independent rMD trajectories, employing a ratchet constant $k_R = 10^{-4}$ kJ/mol. Each simulation lasted 500 ps and the temperature was set to $T = 350$ K;

3. the points along the average pathway in contact map space $\langle C(\tau) \rangle$ have been computed every 7 ps, using only productive trajectories. A trajectory reached the folded state and was considered productive if the RMSD to native of the last frame was smaller than 4 Å. The average path obtained in this way was used to define the path variables in Eq. (1.49). The value of λ was set empirically, as discussed in section 1.3.2, to $\lambda = 13.5$;

4. for each unfolded configuration, we ran 20 SCPS simulations, keeping the same temperature and simulation length of the guess rMD simulations. The SCPS force constants were set to $k_s = 2.5$ kJ/mol and $k_w = 10^{-4}$ kJ/mol;

5. we simulated two iterations, repeating steps 3. and 4., for four out of the five initial conditions, while we simulated three iterations the for the fifth.

All the simulations have been performed using the Amber99fs-ILDN force field [142] together with the implicit solvent model implemented in Gromacs 4.6.5 [41]. To thermostat the system, we employed the Bussi algorithm [143]. Finally, we used Plumed 2.0.2 [51] to implement the collective variables and the bias forces.

The convergence of the iterative algorithm has been assessed by means of the following heuristic and protein-specific strategy. Let us consider the feature space spanned by the RMSD to native of hairpin 1 and 2, which, based on the results of the MD simulations reported in Ref. [9] (see Fig. 1.9), extends from 0 to 1.6 nm for hairpin 1 and from 0 to 1 nm for hairpin 2. Let us call $p^{(I)}$ the two-dimensional histogram obtained by binning the folding trajectories, simulated during the I-th iteration, in the space of hairpin RMSDs: each $p_{ij}^{(I)}$ corresponds to the number of times the folding trajectories assumed the value $x = i\mathrm{d}x$ and $y = j\mathrm{d}y$, with $\mathrm{d}x = 0.8$Å , $d\mathrm{y} = 0.5$Å and $i, j = 1, \ldots, 20$. Conventionally, we refer to the rMD simulations as iteration $I = 0$. To identify the regions populated by the trajectories that fold in the I-th iteration, we consider a binary matrix defined as

$$M_{ij}^{(I)} = \begin{cases} 1 & \text{if } p_{ij}^{(I)} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.58}$$

which is then normalized to unit norm. To compare different iterations, we compute

$$D\left(M^{(I)}, M^{(J)}\right) = \left|\left|M^{(I)} - M^{(J)}\right|\right| \tag{1.59}$$

and we declare that convergence has been assessed at iteration I if

$$1 - \varepsilon \leq \frac{\left|\left|M^{(I)} - M^{(I-1)}\right|\right|}{\left|\left|M^{(I-1)} - M^{(I-2)}\right|\right|} \leq 1 \tag{1.60}$$

where $\varepsilon = 0.1$. Convergence of the results are reported in Fig. 1.10, where it can be seen that simulations starting from initial conditions $1, 2, 3$ and 5 do actually converge, while the one starting from initial condition 4 doesn't. This methodology has of course some limitations. The most evident one is that the binary matrix in Eq. (1.58) doesn't distinguish among highly and poorly populated regions, but rather it just globally focuses on the regions of feature space that are sampled by the folding simulations. Another possible way to assess convergence would be, for example, to project the average pathway on some relevant collective coordinates (in this case the RMSD of the two hairpins, for example) and measure the L2-norm of the distance between average pathways among the different iterations. Since SCPS simulations do not simply push the protein forward along the average path but also constraint the bundle of trajectories within a tube controlled by the variable w_λ, we expect from the convergence of the I-th average path to follow also the convergence of the corresponding $M^{(I)}$ matrix.

1.5.2 Comparison with MD results

In Fig. 1.11 we report the plots of the RMSD to native of the two hairpins of FIP35 obtained by means of MD simulations (Fig. 1.11 (a)), rMD productive trajectories (Fig. 1.11 (b)), the first SCPS iteration (Fig. 1.11 (c)) and converged SCPS productive trajectories (Fig. 1.11 (d)). At the beginning of this section, we discussed that the folding of FIP35 takes two possible routes. rMD simulations report however a different scenario: pathways 1 and 2 are both sub-dominant and an alternative dominant folding route emerges (*pathway 3*), where the two hairpins form cooperatively (see the dashed red box in Fig. 1.11 (a)). These discrepancies are artifacts introduced by the sub-optimal reaction coordinate used to guide rMD simulations: indeed, no sign of cooperativity is visible in unbiased MD simulations (see Fig.s 1.9 (b), (c) and 1.11 (a)). We expect SCPS to be able to start from this sub-optimal set of trajectories and optimize the reaction coordinate in such a way that only the correct folding pathways are selected.

Fig. 1.11 (c) shows the results of the first SCPS iteration. As it is clear from the position of the red dashed box in the picture, already after a single iteration pathway 3 has become sub-dominant. The converged SCPS simulations, reported in Fig. 1.11 (d), show that pathway 1 is actually captured as the dominant one (see the dashed purple box), and pathway 2 is still present but less traveled by the folding simulations. Pathway 3, instead, has completely disappeared at convergence, showing one more time the ability of SCPS to correct for systematic errors in the guess rMD simulations.

Fig. 1.12 reports the folding landscape of FIP35 projected on the plane spanned

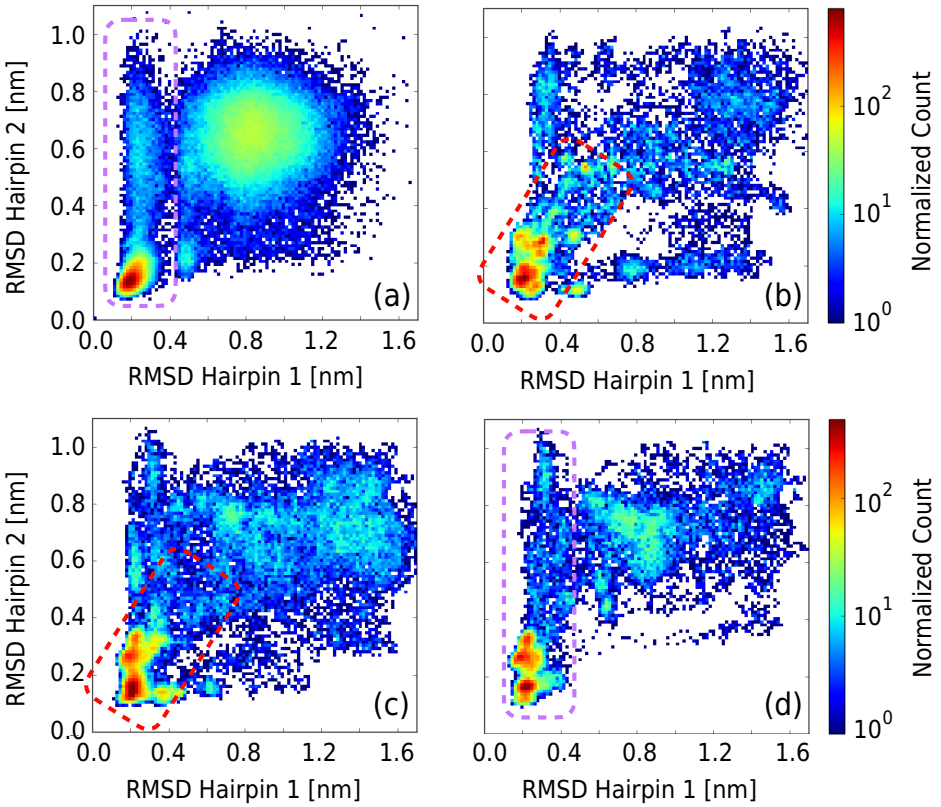

Figure 1.11: (a) RMSD to native of the two hairpins of FIP35, obtained from 2 independent 100 μs long MD simulations [10]. The dashed purple box highlights the region where hairpin 2 is formed after the completion of hairpin 1; (b) RMSD to native of the two hairpins of FIP35, obtained by running rMD simulations started from 5 independent initial conditions. The dashed red box shows the presence of a folding pathway where the two hairpins fold cooperatively; (c) RMSD to native of the two hairpins of FIP35, obtained from the first SCPS iteration. The dashed red box, in the same position as in (b), highlights that the cooperative folding pathway as become at least sub-dominant; (d) RMSD to native of the two hairpins of FIP35, obtained from the converged SCPS simulations. The dashed purple box, in the same position as in (a), displays that the folding pathway where hairpin 2 forms after hairpin 1 has become the dominant one.

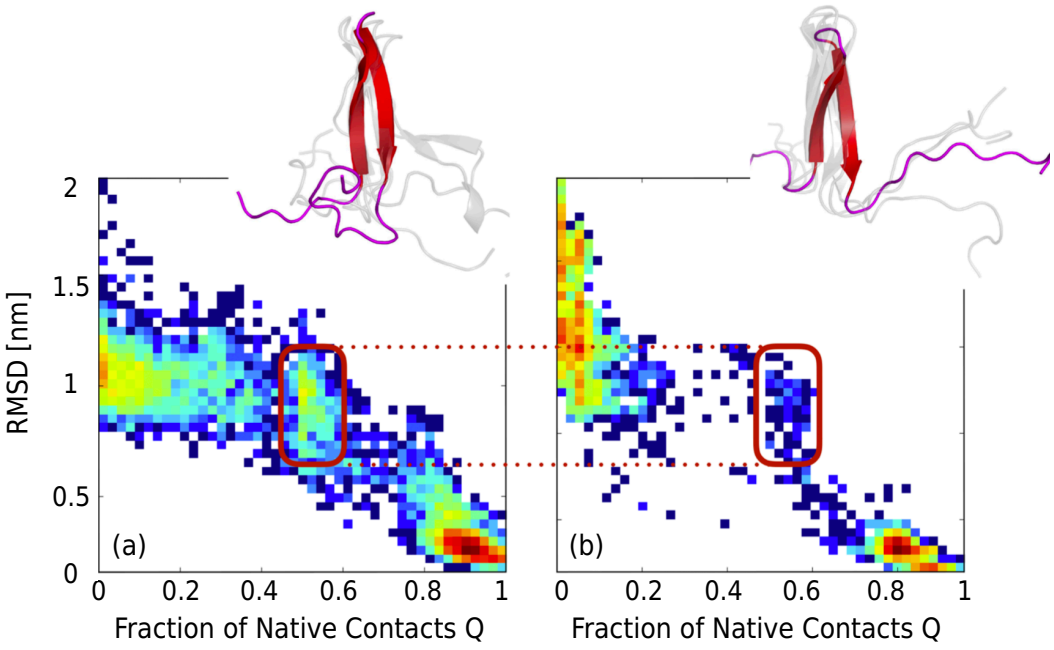

Figure 1.12: Adapted from Ref. [28], with the permission of AIP Publishing. Plot of the fraction of native contacts Q against the RMSD to native of FIP35 obtained from (a) MD simulations and (b) converged SCPS simulations. The red boxes highlight the position of an intermediate state along the folding pathway. On top of panels (a) and (b) we show some configurations extracted from the intermediate states.

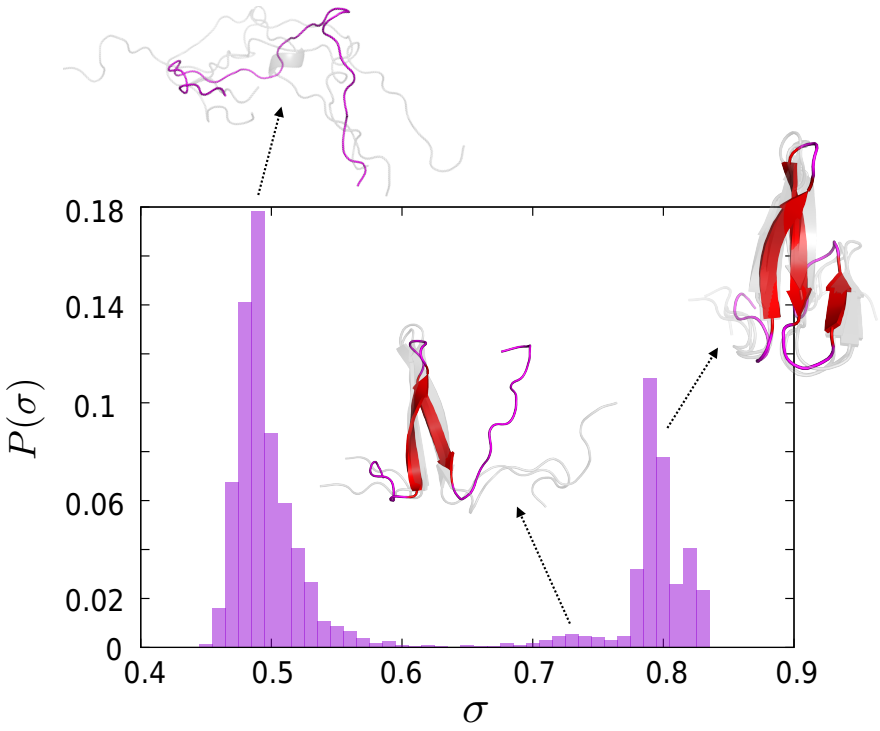

Figure 1.13: Adapted from Ref. [28], with the permission of AIP Publishing. Probability density of the σ variable. Reported protein configurations have been randomly extracted among the ones populating of the highest peaks in the graph.

by the RMSD to native of the whole protein and the fraction of native contacts Q. In panel (a) we report the result obtained from unbiased MD simulations, while in panel (b) we show the same graph obtained from converged SCPS simulations. In both cases, the folding processes progresses by populating an intermediate state, which corresponds to a configuration where only hairpin 1 is formed (see the red boxes in Fig. 1.12 and the protein configurations therein). The existence of an intermediate state was first noted in Ref. [141], where unbiased MD simulations of FIP35 [10] were analyzed using an optimized reaction coordinate. The same metastable state is visible when plotting the probability distribution $P(\sigma)$, in Fig. 1.13. This result has been obtained by histogramming the values of the collective coordinate σ computed from the converged SCPS trajectories. The plot in Fig. 1.13 should not be interpreted as an equilibrium probability density, but nonetheless the three main peaks of the distribution represent three relevant configurations of the FIP35 folding pathway: the unfolded state (around $\sigma \sim 0.5$), the intermediate state (the small kink around $\sigma \sim 0.72$) and the native one (a broad peak centered around $\sigma \sim 0.8$).

A further validation of the results of the SCPS algorithm is provided by the analysis of the order of native contact formation. A way to quantify this concept is given by *path similarity* [21, 144]. Abstractly, folding can be considered as a sequence of steps where native contacts are formed. Therefore, the similarity between two pathways can be deduced by looking at the temporal order in which all the native contacts between alpha-carbons are formed during the two folding events. Quantitatively, given the time of formation of the i-th contact in the k-th trajectory t_{ik}, we build the following tensor:

$$M_{ijk} = \tilde{\theta}(t_{jk} - t_{ik}) \tag{1.61}$$

where $\tilde{\theta}$ is the Heaviside theta-function in Eq. (2). For each pair of pathways k, k', the path similarity $s(k, k')$ is defined as

$$s(k, k') = \frac{1}{N_C(N_C - 1)} \sum_{i \neq j} \delta_{M_{ijk}, M_{ijk'}} \tag{1.62}$$

where N_C represents the total number of native contacts and δ_{hl} is the Kroneker delta-function. $s(k, k')$ is identically 1 if all the native contacts in trajectories k and k' are formed exactly at the same time, while it vanishes for pathways with completely different order of contact formation. In the following, it will be useful to distinguish between *self-* and *cross-similarity*. We will refer to self-similarity when k, k' are pathways taken from the same protein variants or same simulation datasets (e.g. comparing simulations coming from the same SCPS iteration). Differently, cross-similarity will refer to the case where k and k' are folding pathways representing different protein mutations or pertain to different simulation datasets (e.g. comparing

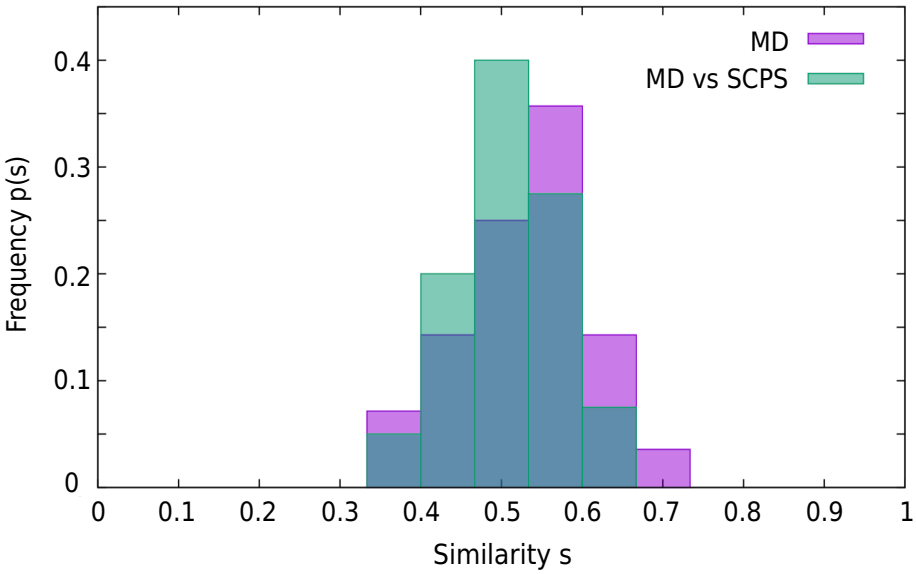

Figure 1.14: Path self-similarity distribution of MD folding trajectories (purple histogram) compared to the path cross-similarity between MD and SCPS simulations (green histogram).

rMD with SCPS simulations). In general, this distinction is possible as long as the number of alpha-carbons is conserved among the two variants, otherwise similarity in Eq. (1.62) is ill-defined. Supposing we are comparing a set of M simulations with another set of N simulations, the path (self- or cross-) similarity distribution can be obtained as

$$p(s) = \frac{1}{NM} \sum_{k<k'} \delta_{s,s(k,k')} \tag{1.63}$$

It can be proved numerically that the similarity distribution of pseudo-random series of contact formations is sharply peaked around $s \sim 0.3$.

We proceeded as follows: in order to quantify how much the folding pathway is intrinsically heterogeneous, we computed the self-similarity of the MD folding simulations of FIP35 (see the purple histogram in Fig. 1.14); secondary, we calculated the cross-similarity between MD and SCPS trajectories (see the green histogram in Fig. 1.14). The substantial overlap between the two histograms in Fig. 1.14 indicates that the differences in the order of contact formation between SCPS and MD lie within the statistical fluctuations intrinsic to the protein dynamics. The result of this last analysis, together with all the previous results, confirms that SCPS and MD simulations predict the same folding mechanism for FIP35.

1.6 Chapter Conclusions

This chapter was dedicated to the problem of sampling rare transitions, with a focus on protein folding. Among all the possible techniques that were used to tackle this

complex problem we concentrated on the Bias Functional method [23]. This strategy is, to our knowledge, the only one that has been employed to study the folding of proteins whose relaxation timescale exceeds hundreds of milliseconds (see chapters 3, 4 and Ref.s [24, 25]). However, the high efficiency of the method doesn't come without a cost. In particular, BF simulations break microscopic reversibility and thus cannot be employed to recover thermodynamic quantities. Another drawback of this method is that it employs an a-priori reaction coordinate that guides the protein towards its native state: it is intended that a sub-optimal choice of the reaction coordinate would inevitably lead to sub-optimal folding pathways. We showed that this last problem can be solved by employing a new type of rMD, which uses the results of the BF method as an initial guess and self-consistently optimizes the reaction coordinate along which the folding simulation is biased. We obtained this result using the path-integral description of Langevin dynamics and performing a *mean-field* approximation on two additional auxiliary fields. The theoretical formulation of the problem let us conclude that rMD, when performed along the so-called path variables, represents a mean-field approximation of the exact Langevin dynamics.

To validate the algorithm, first of all we studied a simple 2-dimensional system. Despite the simplicity of the example, the ability of SCPS to correct for the existence of systematic errors in the initial guess emerged clearly from it. The new method was also tested on the folding of a benchmark protein, the WW-domain of FIP35. Analyses of the simulations confirmed that the pictures described by SCPS and unbiased MD simulations are consistent and provide the same folding mechanism for this protein.

The advantages of using SCPS instead of the BF method come at a relatively moderate cost. Indeed, each SCPS trajectory is simulated using two biasing forces that still strongly accelerate the dynamics, so that every simulation is just 2 to 3 times slower with respect to its rMD counterpart. This performance degradation comes from the fact that many contact map distances have to be computed at each step, differently from what happened in rMD simulations, where just a single matrix norm was computed per step. Because the computation of the elements of a matrix is distributed among many cores, a possible way to increase the efficiency of the algorithm is by distributing the computation of matrix norms and relying on GPUs for the calculation of matrix elements. This modification to the code has yet to be implemented. Moreover, one should take into account that SCPS simulations have to be ran until convergence is reached. Convergence is usually attained (at least for small proteins such as FIP35) within 2-3 iterations. Altogether, SCPS is approximately one order of magnitude slower with respect to rMD. One order of magnitude increase in the computational time is a relatively small price to pay:

indeed, rMD simulations usually require 2 to 3 hours on 2 cores for very small proteins (e.g. < 30 amino acids) and not more than 24 hours on 16-32 cores for medium sized proteins (e.g. > 350 amino acids), in implicit solvent. Therefore, SCPS proposes as a good candidate to study large conformational transitions in an efficient way. There is no theoretical constraint that limits the applications of SCPS to protein folding, so every process of biological interest for which it is possible to compute reasonable guess reaction pathway, e.g. allosteric transitions or inward to outward conformational transitions of membrane proteins, can be systematically optimized using this self-consistent approach.

Finally, we note that extensive and careful testing is still required. In particular, it is imperative to put to proof SCPS in the presence of explicit solvent molecules. A possible way to adequately study the behaviour of this algorithm on a wide class of proteins would be to consider the dataset of folding simulations presented in Ref. [11] and systematically replicate the reported results.

Chapter 2

Transition Path Theory from Self-Consistent Path Sampling

Chapter 1 was fully devoted to the problem of sampling rare transitions using the BF method [23] or SCPS [28], its self-consistent counterpart. Computational simulations of biological processes, however, are a complex puzzle and sampling is just one of its pieces. In parallel with the refinement of numerical algorithms and computational shortcuts for sampling, one needs to develop theoretical methods to reduce the vast amount of data generated by simulations. Of particular interest are, e.g. a robust characterization of the reaction mechanism of the system of interest and its corresponding kinetic observables. *Transition Path Theory* (TPT) [30, 31, 139, 145] is a theoretical framework that permits to achieve these kind of results. On the one hand, it provides a rigorous identification of the reactive probability density $m_T(\mathbf{x})$ and the reactive current $\mathbf{J}_T(\mathbf{x})$ and it shows how it is possible to fully characterize a reactive process by means of these two distributions. On the other hand, TPT introduces the concept of (backward and forward) *committor function* $q(\mathbf{x})$ as optimal reaction coordinate and provides a simple representation of $m_T(\mathbf{x})$ and $\mathbf{J}_T(\mathbf{x})$ as a function of $q(\mathbf{x})$ and the Boltzmann distribution $e^{-\beta U(\mathbf{x})}$. At last, TPT is also considered the natural extension of Transition State Theory [146, 147, 148] to high-dimensional and rugged energy landscapes.

The main difficulty in applying TPT to every system of interest is the computation of the committor function. In principle, q is easily estimated from an infinitely long and ergodic trajectory connecting a reactant state R to a product state P, but in practice this is rarely, if never, feasible. Among all the numerical schemes that have been developed to estimate the committor (for some relevant results see Ref.s [110, 137, 149, 150, 151]), one of the most famous and successful ones is the *String Method* (SM). This method comes in two flavors: a *zero-temperature* one (ZTFM) [152, 153, 154], which is the most elementary among the two and provides reliable

results in the limit $\beta \to \infty$; a *finite-temperature* one (FTSM) [18, 145, 155, 156], theoretically more sophisticated but more effective in estimating q when $\beta \ll \infty$. Both these methods revolve around the introduction of a guess curve (or string) that connects R to P, and try and optimize this curve until it satisfies certain properties. In particular, in the finite temperature case, each point of the string (called *principal curve*) represents the average position of the system on iso-committor surfaces.

The string method and its derivates [157] are extremely valuable and allowed researchers to perform actual TPT calculations on several interesting systems, like for example the conformational activation of Src kinase [158, 159], a conformational transition of the nitrogen regulatory protein C receiver domain [18, 157, 160], the hydrophobic collapse of a coarse-grained chain [161], a conformational transition of myosin VI [162], the ion-dependent conformational dynamics of a transporter [163] and even to the study of critical nuclei in capillary condensation [164]. Like many other path-optimization methods [17, 23, 130, 165, 166, 167, 168, 169, 170, 171], however, SM might suffer from the difficulty to converge to a string which doesn't depend on the specific choice of the initial guess. This problem might become particularly severe for long and complex transitions such as protein folding.

On the other hand, in chapter 1 we argued that it is possible to use rMD (in both its *bare* [23] and self-consistent [28] formulations) to efficiently compute trajectories describing complex transitions. For these reasons, in this chapter we propose a novel algorithm to compute TPT ingredients, $q(\mathbf{x})$, $m_T(\mathbf{x})$ and $\mathbf{J}_T(\mathbf{x})$, by exploiting the the information embedded in SCPS results. First of all, we prove that SCPS results can be employed to approximate the forward committor function. Because the numerical performance of SCPS is comparable with the one of rMD, this method can provide valuable insight in the study of protein folding. Once the committor is known, we show how it is possible to capitalize on this knowledge to efficiently sample the transition path ensemble. We propose two different algorithms to do so: one bases on a *state-to-state* generalization of the *point-to-point* Langevin bridges [172, 173], and the other based on a novel type of rMD, where the committor is used as an optimal reaction coordinate. Both these algorithms show positive sides and shortcomings that need to be carefully taken into account. Finally, the knowledge of $q(\mathbf{x})$ and $m_T(\mathbf{x})$ will be used to derive the transition tubes through the computation of $\mathbf{J}_T(\mathbf{x})$ streamlines. This is done by using an algorithm which is well known in literature [139].

The chapter is structured as follows. In section 2.1 we review some of the basic concepts and derivations of TPT. Section 2.2 is instead devoted to summarize the main features behind the string method. In section 2.3 we introduce one of the main concepts of the chapter: we define a suitable range of timescales, which we refer

to as the *Steady Current Regime*, which permits us to approximate the results of TPT using, instead of an infinitely long and ergodic trajectory, many short and non-ergodic simulations that last in this time regime. This very technical step is needed to bridge the gap between the numerical results coming from enhanced sampling methods (short productive trajectories) and the theoretical requirements between TPT (infinitely long equilibrium trajectory). Section 2.4 is devoted to the algorithmic part of this chapter: we propose a method to approximate the committor function by means of SCPS and we present two algorithms that use this information to sample the transition path ensemble. In section 2.5 we show how it is possible to apply the machinery developed in section 2.4 to a simple 2D system, where we obtain promising results in a reasonably low computational time. Finally, in section 2.6 we will draw some final conclusions and discuss the advantages, and also the possible difficulties, to the application of this algorithm to biological systems.

2.1 Recap of Transition Path Theory

Let us suppose we have a generic state space Ω and that we observed an ergodic and infinitely long trajectory x:

$$x = \{x(t) \ : \ -\infty < t < +\infty\} \tag{2.1}$$

We also provide the definition of two regions in Ω, called *reactant state* R and *product state* P that satisfy:

$$P \subset \Omega \quad R \subset \Omega \quad P \cap R = \emptyset \tag{2.2}$$

The region $\Omega_T = \Omega/(R \cup P)$ represents the reactive part of the state space, also called *transition region*. Trajectory x, being infinitely long and ergodic, will join R and P an infinite amount of times. Given these ingredients, we want to characterize the reaction mechanism between R and P. TPT is the theoretical framework designed to do it, by addressing the following three questions:

1. what does it mean for a trajectory to be reactive?
2. what is the probability density associated to $x(t)$, conditioned to x to be reactive?
3. what is the probability current associated to reactive trajectories?

In this section we are going to see how TPT answers all these three questions, provided our first assumptions on x. To do this, we will introduce the fundamental concepts of *reactive trajectory*, *committor function*, *reactive distribution* and *reactive current*, and we will closely follow the dissertation presented in [31].

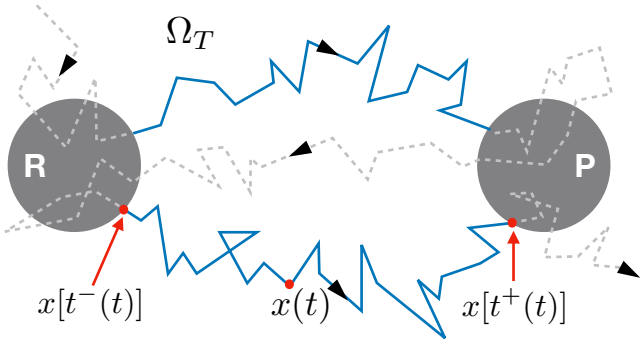

Figure 2.1: Sketch of an infinitely long trajectory x. The reactive portions of the trajectory are colored as solid blue lines, while non-reactive parts are represented as dashed grey lines. The figure also shows the positions of $x(t^+(t))$ and $x(t^-(t))$ associated to an arbitrary point $x(t)$.

2.1.1 Reactive trajectories

One starts by introducing two auxiliary quantities:

$$\begin{aligned} t^+(t) &= \min_{t' \geq t} \{t' \ : \ x(t') \in R \cup P\} \\ t^-(t) &= \max_{t' \leq t} \{t' \ : \ x(t') \in R \cup P\} \end{aligned} \tag{2.3}$$

$t^+(t)$ provides the smallest time instant after t for which the trajectory was either located in the reactant or in the product states. Similarly, $t^-(t)$ defines the biggest time interval before t for which the trajectory was, again, in the reactant or in the product states. The ensemble of *reactive trajectories* $\mathbb{E}_T$ or *transition path ensemble* is defined as the set of continuous segments of x for which the following properties hold:

$$\mathbb{E}_T \equiv \left\{ x(t) \ : \ x(t) \in \Omega_T, \ x(t^-(t)) \in R \text{ and } x(t^+(t)) \in P \right\} \tag{2.4}$$

This means that a portion of trajectory x is reactive if: (i) it is continuous; (ii) all its points live in the transition region, $x(t) \in \Omega_T$; (iii) the last time it accessed the transition region, it was by leaving the reactant state, $x(t^-(t)) \in R$; the next time it will leave the transition region, it will be by entering the product state, $x(t^+(t)) \in P$. A pictorial representation of the reactive portions of an infinitely long trajectory x is given in Fig. 2.1.

Reactive trajectories defined as in Eq. (2.4) possess a clear directionality: only continuous portions of x that go from R to P are reactive, while those *regressing* from P to R are excluded from $\mathbb{E}_T$. If the system's dynamics is microscopically

reversible[1], the ensemble

$$\mathbb{E}_T^* \equiv \left\{ x(t^*) \ : \ x(t^*) \in \Omega_T, \ x(t^-(t^*)) \in P \text{ and } x(t^+(t^*)) \in R \right\} \tag{2.5}$$

is also reactive if one considers $t^* = -t$. Microscopic reversibility is not a required assumption in TPT, which is general enough to study systems which evolve with a non-time reversible dynamics. Nonetheless, in what follows we shall assume that microscopic reversibility holds, as it sensibly simplifies the interpretation of some results.

2.1.2 The Committor Function

One of the pivotal concepts of TPT is the *committor function* q, which is usually considered to be the optimal reaction coordinate [151, 174]. One distinguishes between the *forward committor function*

$$q^+(\mathbf{x}) = \text{probability that } x(t) \text{ reaches } P \text{ before } R, \text{ given that } x(0) = \mathbf{x} \tag{2.6}$$

and the *backward committor function*

$$q^-(\mathbf{x}) = \text{probability that } x(t) \text{ reaches } R \text{ before } P, \text{ given that } x(0) = \mathbf{x} \tag{2.7}$$

Note that since we are assuming time-reversibility, one has

$$q^+(\mathbf{x}) = 1 - q^-(\mathbf{x}) \tag{2.8}$$

so we might as well drop the superscript $+$ and refer to the committor function as $q(\mathbf{x})$. This notation, unless differently specified, will hold for the rest of this manuscript.

A crucial property of the committor function is that it satisfies the backward-Kolmogorov equation (reported and discussed in appendix A):

$$D(\boldsymbol{\nabla}^2 - \beta \boldsymbol{\nabla} U(\mathbf{x}) \cdot \boldsymbol{\nabla}) q^{\pm}(\mathbf{x}) \equiv H_{\mathrm{FP}}^{\dagger} q^{\pm}(\mathbf{x}) = 0 \tag{2.9}$$

provided that suitable boundary conditions are satisfied. In particular, one equips $q^{\pm}(\mathbf{x})$ with absorbing boundary conditions at the boundaries of the reactant ∂R and the product ∂P. This will make sure that trajectories crossing the boundaries will be annihilated before entering the states. Conventionally, one provides the following values of the committor at the boundaries:

$$q(\mathbf{x})|_{\partial P} = 1 \quad q(\mathbf{x})|_{\partial R} = 0 \tag{2.10}$$

[1] the probability $P[x(t)]$ of observing a trajectory is the same of observing the time-reversed trajectory $P[x(-t)]$

The reason why Eq. (2.9) holds is briefly reviewed in appendix C.1. It should be noted that when the dynamics is deterministic, both $q^+(\mathbf{x})$ and $q^-(\mathbf{x})$ are rather uninformative functions, because, depending on the initial conditions, they are automatically assigned 0 or 1. For this reason, we shall assume one more time the dynamics to be stochastic.

In the following sections we will see how to employ the committor function to define the concepts of reactive distribution and reactive current.

2.1.3 Reactive Probability Density

The ergodicity assumption requires the trajectory x to be ergodic with respect to some equilibrium probability distribution $m(\mathbf{x})$. This means that the average value of an observable O along trajectory x can be evaluated either as a time average or as an ensemble average:

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt\; O(x(t)) = \int_{\Omega} d\mathbf{x}\; m(\mathbf{x}) O(\mathbf{x}) \tag{2.11}$$

At this point one can try and answer question 2.: what is the probability distribution $m_T(\mathbf{x})$ associated to a point $\mathbf{x} = x(t) \in \Omega_T$, conditioned to x being reactive? The trivial guess would be that $m_T(\mathbf{x})$ is just provided by the equilibrium distribution $m(\mathbf{x})$ restricted to the reactive portion, but this is incorrect:

$$m_T(\mathbf{x}) \neq \frac{m(\mathbf{x})|_{\Omega_T}}{\int_{\Omega_T} d\mathbf{x}\; m(\mathbf{x})} \tag{2.12}$$

Indeed, $m(\mathbf{x})|_{\Omega_T}$ takes into account non-reactive contributions coming from points sampled by trajectories that leave R and re-enter R before reaching P and trajectories that leave P and re-enter P before reaching R. To correctly identify it, one starts by defining

$$h_A[x(t)] = \begin{cases} 1 \text{ if } x(t) \in A \subset \Omega \\ 0 \text{ otherwise} \end{cases} \tag{2.13}$$

the characteristic function of set A. One can see that a point $x(t)$ belongs to a reactive portion of trajectory x if and only if

$$h_{\Omega_T}[x(t)] h_P[x(t^+(t))] h_R[x(t^-(t))] \equiv h_T[x(t)] = 1 \tag{2.14}$$

Eq. (2.14) is equivalent to the definition of $\mathbb{E}_T$ in Eq. (2.4). Given the ergodicity assumption, it is possible to write the reactive distribution as the probability density $m_T(\mathbf{x})$ for which

$$\lim_{T\to\infty} \frac{\int_{-T}^{T} dt\; O(x(t)) h_T(x(t))}{\int_{-T}^{T} dt\; h_T(x(t))} = \int_{\Omega_T} d\mathbf{x}\; m_T(\mathbf{x}) O(\mathbf{x}) \tag{2.15}$$

Similarly, probability density $m_R(\mathbf{x})$ for $\mathbf{x} = x(t)$ to be reactive at time t is simply given by

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt\; O(x(t))h_T(x(t)) = \int_{\Omega_T} d\mathbf{x}\; m_R(\mathbf{x})O(\mathbf{x}) \tag{2.16}$$

Definitions in Eq.s (2.15) and (2.16) are rather implicit: one would like to have a more simple and direct way to express those distributions. Let's see how one can do it.

$m_R(\mathbf{x})$ is, by definition, the product of the probability density for x to be in $\mathbf{x} = x(t)$ at time t, thus $m(\mathbf{x})$, and the probability $P_R(\mathbf{x})$ for that trajectory to be reactive. The most suitable way to express $P_R(\mathbf{x})$ is by resorting to the committor function. Evidently probability $P_R(\mathbf{x})$ can be written as

$$P_R(\mathbf{x})) = q(\mathbf{x})(1 - q(\mathbf{x})) \tag{2.17}$$

Thus

$$m_R(\mathbf{x}) = m(\mathbf{x})q(\mathbf{x})(1 - q(\mathbf{x})) \tag{2.18}$$

The reactive distribution is just a normalization factor away from Eq. (2.18). Indeed, one has that

$$\int_{\Omega_T} d\mathbf{x}\; q(\mathbf{x})(1 - q(\mathbf{x}))m(\mathbf{x}) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt\; h_T[x(t)] \equiv Z \tag{2.19}$$

thus comparing Eq.s (2.15), (2.16) and (2.19) one finally obtains

$$m_T(\mathbf{x}) = Z^{-1}m(\mathbf{x})q(\mathbf{x})(1 - q(\mathbf{x})) \tag{2.20}$$

which is the answer to question 2. and one of the main results of TPT [175]. In the NVT ensemble, Eq. (2.20) becomes

$$m_T(\mathbf{x}) = \frac{e^{-\beta U(x)}}{Z} q(\mathbf{x})(1 - q(\mathbf{x})) \qquad Z = \int_{\Omega_T} d\mathbf{x}\; e^{-\beta U(\mathbf{x})} q(\mathbf{x})(1 - q(\mathbf{x})) \tag{2.21}$$

$m_T(\mathbf{x})$ provides the first important information about the reaction $R \to P$, because it allows one to compute the fraction of time spent in a particular region of the transition space. In particular, calling this region C, one can compute this time by computing the average of $O(\mathbf{x}) = h_C[\mathbf{x}]$:

$$t_C = \int_{\Omega_T} d\mathbf{x}\; h_C[\mathbf{x}]m_T(\mathbf{x}) = \int_C d\mathbf{x}\; m_T(\mathbf{x}) \tag{2.22}$$

In a protein folding setting, this would be useful, i.e. to determine how much time is spent in a metastable state *during* the folding process, so while the conformational transition is still happening and not in equilibrium conditions. Moreover, *transition states* of the reaction are simply identified by peaks in $m_T(\mathbf{x})$: in this sense, TPT provides the natural extension of transition state theory [176, 177, 178].

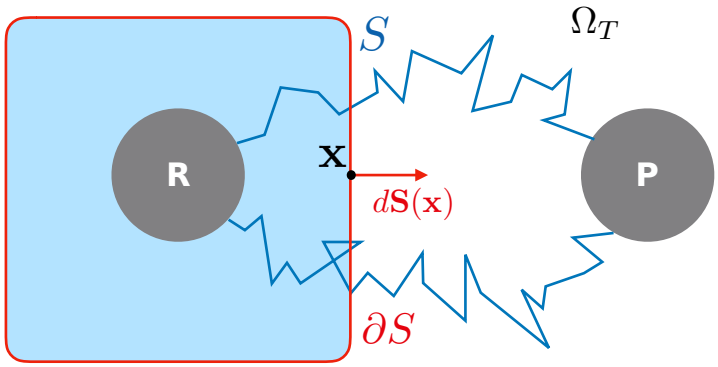

Figure 2.2: Schematic representation of the setting introduced in Eq. (2.23)

2.1.4 Reactive Current

In Ref. [139], it is reported a series of toy models on which the authors apply the concepts of transition path theory. Looking at those examples, one thing becomes immediately clear: the reactive distribution $m_T(\mathbf{x})$ doesn't encode all the information about the reaction mechanism. In particular, for multi-channel reactions $m_T(\mathbf{x})$ doesn't provide any information about the most probable pathway that connects the reactant to the product. This information is instead encoded in a $3N$-dimensional vector field called *reactive current*. Let us consider a surface S enclosing the reactant state, with border ∂S. $d\mathbf{S}(\mathbf{x})$ is the oriented surface element in $\mathbf{x}$, conventionally directed outwards the surface border (see Fig. 2.2). Exploiting ergodicity, the current can be implicitly defined as

$$\lim_{\tau\to 0^+}\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} dt\; h_R[x(t^-(t))]h_P[x(t^+(t))]\cdot \\ \cdot\left(h_S[x(t)]h_{\Omega/S}[x(t+\tau)] - h_{\Omega/S}[x(t)]h_S[x(t+\tau)]\right) = \int_{\partial S} d\mathbf{S}(\mathbf{x})\cdot\mathbf{J}_T(\mathbf{x}) \tag{2.23}$$

It can be proved [31, 136] that, if we assume time-reversibility the reactive current can be expressed as

$$\mathbf{J}_T(\mathbf{x}) = DZ^{-1}m(\mathbf{x})\boldsymbol{\nabla}q(\mathbf{x}) \tag{2.24}$$

Again, in the NVT ensemble $m(\mathbf{x})$ is replaced with the Boltzmann distribution and one finds

$$\mathbf{J}_T(\mathbf{x}) = D\frac{e^{-\beta U(\mathbf{x})}}{Z}\boldsymbol{\nabla}q(\mathbf{x}) \tag{2.25}$$

A first important property that one can notice is that this current is divergence-less:

$$\boldsymbol{\nabla}\cdot\mathbf{J}_T = D\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}q(\mathbf{x})\frac{e^{-\beta U(\mathbf{x})}}{Z}\right) = D\left(\boldsymbol{\nabla}^2 q(\mathbf{x}) - \boldsymbol{\nabla}U(\mathbf{x})q(\mathbf{x})\right)\frac{e^{-\beta U(\mathbf{x})}}{Z} = 0 \tag{2.26}$$

where in the last equality one uses the fact that the committor function satisfies the backward-Kolmogorov equation, Eq. (2.9). Using the divergence theorem, one also

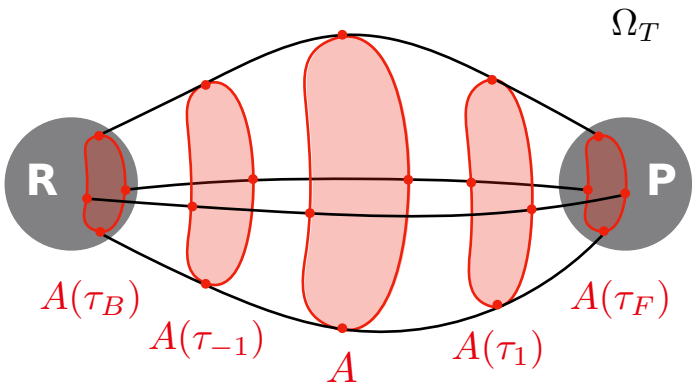

Figure 2.3: Schematic representation of the artificial surface dynamics in Eq. (2.30)

finds that

$$\int_S \boldsymbol{\nabla} \cdot \mathbf{J}_T(\mathbf{x}) dS = \int_{\partial S} d\mathbf{S}(\mathbf{x}) \cdot \mathbf{J}_T(\mathbf{x}) = 0 \tag{2.27}$$

which means that is $\partial S \subset \Omega_T$ is a closed surface, the reactive probability flux through S is zero. If ∂S_D, instead, is a dividing surface, i.e. a surface separating Ω in two disjoint regions, one containing R and one containing P, then the flux is constant and defines the *reaction rate*:

$$k = \int_{\partial S_D} d\mathbf{S}_D(\mathbf{x}) \cdot \mathbf{J}_T(\mathbf{x}) \tag{2.28}$$

This is necessary because, by construction, R is the only source and P is the only sink of reactive trajectories, so every reactive trajectory leaving R eventually reaches P. We stress the fact that the reactive current is non-vanishing at equilibrium differently from the Fokker-Planck current (see appendix A).

Of particular interest about $\mathbf{J}_T(\mathbf{x})$ are its streamlines, which can be used to determine the so called *transition tubes*. These tubes are defined as the regions in Ω_T where reactive events occur more likely. A possible algorithm based on $\mathbf{J}_T(\mathbf{x})$ to compute them is schematized in Fig. 2.3 and it works as follows. Given an arbitrary dividing surface ∂S, one can identify a portion $\partial A \subset \partial S$ which amounts for a given percentage p of the total flux:

$$p \int_{\partial S} d\mathbf{S}(\mathbf{x}) \cdot \mathbf{J}_T(\mathbf{x}) = \int_{\partial A} d\mathbf{A}(\mathbf{x}) \cdot \mathbf{J}_T(\mathbf{x}) \qquad p \leq 1 \tag{2.29}$$

Each point $\mathbf{z} \in \partial A$ is then evolved using the artificial dynamics

$$\frac{d\mathbf{z}(\tau)}{d\tau} = \mathbf{J}_T(\mathbf{z}(\tau)) \tag{2.30}$$

where τ is an artificial time not to be confused with the physical one. Each point is evolved forward in time, defining new surfaces $\partial A(\tilde{\tau})$ at each time $\tau = \tilde{\tau}$, until it reaches P, and subsequently the same points on ∂A are propagated backwards

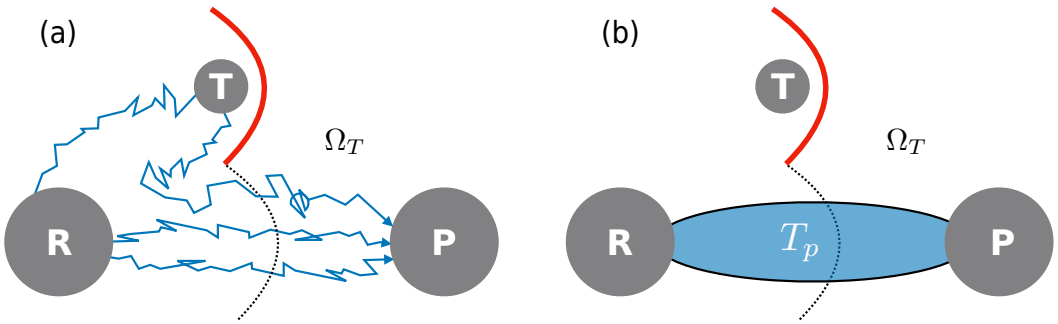

Figure 2.4: Schematic representation of a potential showing a kinetic trap T. (b) The transition tube T_p doesn't feel a net contribution from $R \to T$ transitions because (a) reactive trajectories that reach T have to backtrack to progress towards P.

in time until they reach R. Therefore, forward integration is repeated until time τ_F for which $\mathbf{z}(\tau_F) \in \partial P$, while backward integration is carried out for τ_B, where $\mathbf{z}(\tau_B) \in \partial R$. Note τ_F and τ_B are not necessarily similar and that their values may vary from point to point. Given Eq. (2.28), the flux carried through each surface $\partial A(\tilde{\tau})$ is constant. It means that the final tube

$$T_p = \bigcup_{\tau} \partial A(\tau) \tag{2.31}$$

connecting R to P will amount for the p percentage of the total probability flux.

Is t here a set of optimal surfaces which can be used to compute fluxes? Yes there is, and with no surprises these surfaces are the iso-committo ones. Let us call $\partial S_{\overline{q}}$ a surface for which $q(\mathbf{x}) = \overline{q} = \text{const.}$, i.e. an iso-committor surface. On this surface, $\boldsymbol{\nabla} q(\mathbf{x}) = \boldsymbol{\nabla} \overline{q} = 0$ in every direction except for the one normal to the surface. For this reason we have $\boldsymbol{\nabla} q(\mathbf{x}) \parallel d\mathbf{S}_{\overline{q}}(\mathbf{x})\ \forall \mathbf{x}$, which translates into the following property:

$$k = \int_{\partial S_{\overline{q}}} d\mathbf{S}_{\overline{q}}(\mathbf{x}) \cdot \mathbf{J}_T(\mathbf{x}) = D \int_{\partial S_{\overline{q}}} d\mathbf{x}\, |\boldsymbol{\nabla} q(\mathbf{x})| \frac{e^{-\beta U(\mathbf{x})}}{Z} \equiv D \int_{\partial S_{\overline{q}}} d\mathbf{x}\, j(\mathbf{x}) \tag{2.32}$$

The interesting thing about Eq. (2.32) is that

$$j(\mathbf{x}) = |\boldsymbol{\nabla} q(\mathbf{x})| \frac{e^{-\beta U(\mathbf{x})}}{Z} \tag{2.33}$$

is the probability density associated to the event of passing through $\partial S_{\overline{q}}$ for the first time by hitting $\mathbf{x}$ [179].

It is important to stress that reactive trajectories and streamlines of the reactive current should not be confused. Indeed, the first ones are stochastic while the second ones result from an artificial and deterministic dynamics that averages out local irrelevant features. To given an example, potentials describing high-dimensional systems can show kinetic traps, i.e. local metastable basins where the reactive trajectories spend a lot of time going back-and-forth, because the reaction cannot progress by

passing through them (see Fig. 2.4 (a) for a schematic representation). These processes however, do not contribute to the current and its flows (see Fig. 2.4 (b)), but rather they contribute to the definition of the reactive distribution.

2.2 The String Method

In section 2.1 we presented the committor fuction and discussed its importance in TPT, being the main ingredient needed to provide a robust statistical description of reactive pathways. In this section we will instead discuss how it is possible to compute it and the other related quantities (the reactive distribution and current) for a system of interest. TPT is based on the assumption that the transition path ensemble is obtained by suitably cutting segments of an infinitely long trajectory connecting a state R to a state P. This is impractical to say the least, and one should find smarter ways to compute TPT ingredients for a given system. An attractive possibility would be to directly solve Eq. (2.9) to obtain the committor function, but the high dimensionality of biological systems makes this equation intractable with e.g. finite element methods.

In this section we will discuss some approximations that can be made to compute the committor function, which altogether go under the name of *String Method* (SM) [18, 145, 152, 153, 154, 155, 156]. The main point is to assume that there exists a suitable curve that carries most of the reactive probability flux, and instead of solving Eq. (2.9) one can compute the committor on the curve and use this information to approximate it in the vicinity of this curve. As we will see, this is much easier than solving a high-dimensional non-linear differential equation or simulating an infinite stochastic trajectory.

2.2.1 Minimum Energy Paths and the Zero-Temperature String Method

Let us suppose we have a curve γ that is obtained by connecting the centers of the intersections between transition tubes and iso-committor surfaces. Let us also assume that this curve carries most of the reactive probability flux, i.e. there is a dominant reaction channel. We parametrize γ by means of a function $\boldsymbol{\phi} : [0,1] \to \Omega$ and a curvilinear coordinate $s \in [0,1]$:

$$\gamma = \{\boldsymbol{\phi}(s) \ : \ s \in [0,1]\} \tag{2.34}$$

We also define an auxiliary function, $s_\gamma : \Omega \to [0,1]$ such that $s_\gamma(\mathbf{x})$ identifies the point on γ which is closer to $\mathbf{x} \in \Omega_T$:

$$|\mathbf{x} - \boldsymbol{\phi}(s_\gamma(\mathbf{x}))| = \min_{s\in[0,1]} |\mathbf{x} - \boldsymbol{\phi}(s)| \tag{2.35}$$

We want to approximate the committor function under the assumption that $s_\gamma(\mathbf{x})$ is a parametrization of the committor $q(\mathbf{x})$ through an unspecified function f:

$$q(\mathbf{x}) = f(s_\gamma(\mathbf{x})) \qquad f : [0,1] \to [0,1] \tag{2.36}$$

Function f tells us the exact value of the committor on the curve and approximates $q(\mathbf{x})$ for $\mathbf{x} \notin \gamma$ by computing the value of the committor on the point $\boldsymbol{\phi}(s)$ closer to $\mathbf{x}$. This assumption is valid only in the vicinity of the curve γ, where it guarantees that iso-s_γ are, up to relabeling, iso-q surfaces.

The fact that γ carries most of the reactive probability flux is equivalent to ask for the points on the curve to maximize $j(\mathbf{x})$, defined in Eq. (2.33), on every iso-committor function $\partial S_{\overline{q}}$. By using Eq. (2.36) we can re-write Eq. (2.33) as

$$j(\mathbf{x}) = \frac{e^{-\beta U(\mathbf{x})}}{Z} f'(s_\gamma(\mathbf{x}))|\boldsymbol{\nabla} s_\gamma(\mathbf{x})| \tag{2.37}$$

and we should ask for $j(\mathbf{x})$ to be maximum for $\mathbf{x} = \boldsymbol{\phi}(s)$. It is also true that, along the curve, the gradients of $j(\mathbf{x})$ and $s_\gamma(\mathbf{x})$ are parallel [136]:

$$\boldsymbol{\nabla}\left(e^{-\beta U(\boldsymbol{\phi}(s))}|\boldsymbol{\nabla} s_\gamma(\boldsymbol{\phi}(s))|\right) \parallel \boldsymbol{\nabla} s_\gamma(\boldsymbol{\phi}(s)) \tag{2.38}$$

Manipulating these equations one finds well-known equation defining a *minimum energy path* (MEP)

$$(\boldsymbol{\nabla} U)^\perp [\boldsymbol{\phi}^*] = 0 \tag{2.39}$$

where one introduces the operator $\perp$ such that

$$\mathbf{f}^\perp[\boldsymbol{\phi}] = \mathbf{f} - (\mathbf{f} \cdot \hat{n}_\parallel^\gamma)\hat{n}_\parallel^\gamma \tag{2.40}$$

with $\hat{n}_\parallel^\gamma$ the unit vector tangent to γ. Eq. (2.39) tells us that the curve $\boldsymbol{\phi}^*(s)$ that carries the maximum flux is everywhere tangent to the gradient of the potential, i.e. it is a minimum energy path. One can also prove that function f is given by [136]

$$f(s) = \frac{\int_0^s ds'\, e^{\beta F(s')}}{\int_0^1 ds'\, e^{\beta F(s')}} \tag{2.41}$$

where $F(s)$ is the free energy along s, which is not known a priori and should be suitably sampled.

A possible strategy to compute the MEP connecting a given reactant state R to a product state P is given by the so-called *zero-temperature string method* [152, 153, 154]. The general idea of the string method is to evolve a guess *string* (the curve γ) in such a way that its parametrization is always satisfied and that it eventually converges to the MEP. The evolution of a string towards the MEP can be compactly described as

$$\mathbf{v}_n = -(\boldsymbol{\nabla} U)^\perp \tag{2.42}$$

where $\mathbf{v}_n$ is the normal velocity of the curve. Using a suitable parametrization of the curve, provided by

$$\gamma(t) = \{\boldsymbol{\phi}(t, \alpha) \ : \ \alpha \in [0, 1]\} \tag{2.43}$$

one can rewrite Eq. (2.42) as

$$\dot{\boldsymbol{\phi}} = -(\boldsymbol{\nabla} U)^{\perp}[\boldsymbol{\phi}] + \lambda \frac{\boldsymbol{\phi}'}{|\boldsymbol{\phi}'|} \tag{2.44}$$

where the dot denotes a time derivative, the air quote denotes the derivative with respect to the arc-length α and λ is a Lagrange multiplier that constraints the curve length to be constant. Eq. (2.44) can be solved through a two-step procedure: (i) the curve is evolved by following the potential, $\dot{\boldsymbol{\phi}} = -(\boldsymbol{\nabla} U)^{\perp}[\boldsymbol{\phi}]$; (ii) the parametrization is enforced on the evolved curve.

Once the MEP is given, one needs to sample the free-energy in its vicinity to get the function $f(s)$ in Eq. (2.41), using e.g. Blue Moon sampling [180, 181]. At that point, the job is done: iso-q surfaces are approximated as iso-s ones and transition tubes and the reactive distribution are recovered based on this assumption [136].

Besides it simplicity and simple interpretation, MEPs have many disadvantages [18, 136]. The main issue is that, by definition, the MEP depends on local features of the potential at a resolution higher than $k_{\mathrm{B}}T$, since the fine details are leveled out by the dynamics.

A possible way to go beyond the concept of MEP, without losing its conceptual simplicity, is by replacing the state space Ω with a space spanned by suitable collective coordinates $\{\theta\}$. This would average out unimportant features and also account for some non-local effects. In such a way, one would obtain the so called *minimum free energy path* (MFEP) [182] in a similar way to the MEP:

$$(\boldsymbol{\nabla} F(\{\theta\}))^{\perp}\,[\boldsymbol{\phi}^*] = 0 \tag{2.45}$$

The MFEP $\boldsymbol{\phi}^*$ is everywhere parallel to the gradient of the free-energy landscape. Again, this is in general not enough, because MFEP is still a local object and might ignore important features perpendicular to itself. Given these observations, it was proposed to abandon MEPs and MFEPs and resort to the concept of *principal curve* [145, 155, 156], to which we dedicate the next section.

2.2.2 Principal Curves and the Finite Temperature String Method

Let us consider a curve γ in Ω and the set of planes intersecting each point of the curve so that they are locally normal to the curve (see Fig. 2.5). If the position of the intersection coincides with the average position of the system along these planes, the curve is called a *principal curve* [18, 136, 183]. To put it in a more quantitative

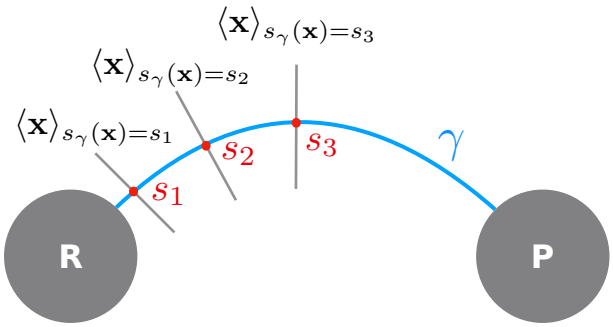

Figure 2.5: Schematic representation of a principal curve connecting R to P, as defined in Eq.s (2.46) and (2.47).

fashion, let us go back to Eq. (2.35) where we defined function $s_\gamma(\mathbf{x})$: γ is a principal curve with respect to the canonical density $Z^{-1}e^{-\beta U(\mathbf{x})}$ if

$$\boldsymbol{\phi}(s) = \langle \mathbf{x} \rangle_{s_\gamma(\mathbf{x})=s} \qquad \forall s \in [0,1] \tag{2.46}$$

where the average is a canonical average

$$\langle \mathbf{x} \rangle_{s_\gamma(\mathbf{x})=s} = \frac{\int_\Omega d\mathbf{x}\, \mathbf{x} e^{-\beta U(\mathbf{x})} \delta\left(s_\gamma(\mathbf{x}) - s\right)}{\int_\Omega d\mathbf{x}\, e^{-\beta U(\mathbf{x})} \delta\left(s_\gamma(\mathbf{x}) - s\right)} \tag{2.47}$$

A principal curve does not depend on some local features of the potential or the free energy along some suitable reaction coordinates, therefore it is a more global concept than MEP and MFEP. Indeed, the principal curve is defined by an average over the hypersurfaces defined by $s_\gamma(\mathbf{x}) = s$. Interestingly, the principal curve collapses to the MEP for $\beta \to \infty$ (that's why the method to compute MEP is called *zero-temperature* string method). Indeed, the expectation value in Eq. (2.47), for $\beta \to \infty$, is dominated by the point where $U(\mathbf{x})$ is minimum:

$$\lim_{\beta\to\infty} \langle \mathbf{x} \rangle_{s_\gamma(\mathbf{x})=s} = \mathbf{x}^* \tag{2.48}$$

where

$$U(\mathbf{x}^*) = \min_{\mathbf{x}, s_\gamma(\mathbf{x})=s} U(\mathbf{x}) \tag{2.49}$$

Evidently, Eq. (2.49) is equivalent to ask for

$$(\boldsymbol{\nabla} U(\mathbf{x}^*))^\perp = 0 \tag{2.50}$$

which means that for $\beta \to \infty$ the principal curve is equivalent to a MEP.

Conversely, for $\beta < \infty$ a principal curve can be arbitrarily different from a MEP and for rugged potentials a MEP might even lose any meaning (see the example on the rough Müller potential in Ref. [18]). Therefore one needs an evolution of the zero-temperature string method that allows one to compute the principal curve connecting

R to P. Given a guess string, this is discretized in $N+1$ *images*, $\{\boldsymbol{\phi}_\alpha\}_{\alpha=0,\dots,N}$. To each image one associates its corresponding Voronoi cell, defined as the sets $B_\alpha \subset \Omega$ containing all the points that are closer to a given $\boldsymbol{\phi}_\alpha$ with respect to any other $\boldsymbol{\phi}_\beta$:

$$B_\alpha = \{\mathbf{x} \,:\, |\mathbf{x}-\boldsymbol{\phi}_\alpha| < |\mathbf{x}-\boldsymbol{\phi}_\beta| \;\forall \alpha \neq \beta\} \tag{2.51}$$

Voronoi cells permit to easily approximate the expectation value in Eq. (2.47) for $s = \alpha/N$ as [18]:

$$\langle \mathbf{x} \rangle_{s_\gamma(\mathbf{x}) \sim \frac{\alpha}{N}} = \frac{\int_{B_\alpha} d\mathbf{x}\; \mathbf{x} e^{-\beta U(\mathbf{x})}}{\int_{B_\alpha} d\mathbf{x}\; e^{-\beta U(\mathbf{x})}} \tag{2.52}$$

In Eq. (2.52), Voroni cells play the role of thickened iso-s hyperplanes, and indeed the two concepts are related for $N \to \infty$ [18]. Eq. (2.52) suggests a simple algorithm to compute principal curves, which goes under the name of *finite-temperature string method*: (i) sample the Voronoi cells by running stochastic trajectories and imposing reflecting boundary conditions on B_α borders; (ii) update images $\{\boldsymbol{\phi}_\alpha\}$ positions and their corresponding cells B_α according to the estimate of $\langle \mathbf{x} \rangle_{s_\gamma(\mathbf{x}) \sim \frac{\alpha}{N}}$ obtained by sampling; (iii) enforce string parametrization. Since centroids of the Voronoi cells are optimized on the go, (i)-(iii) have to be iterated until some convergence criterion is met. At the last iteration, the algorithm returns a set of images $\{\boldsymbol{\phi}_\alpha^{\text{last}}\}$ that, once interpolated, define the principal curve.

One of the big advantages of this method is that the free energy needed to compute $f(s)$ in Eq. (2.41) is obtained as a byproduct of step (i). Indeed, one can prove [18] that $F(\alpha/N)$ is related to the probability of finding the system in cell B_α.

2.3 TPT from Non-Ergodic Trajectories

In section 2.1 we briefly reviewed the main concepts of TPT by following the constructive route marked by Ref. [136]. The statistical mechanics framework introduced by TPT strongly depends on the concept of committor function and thus all the sophisticated theoretical machinery becomes inapplicable if there is not practical way to estimate it. For this reason, in section 2.2 we summarized the main ideas behind the string method and discussed how it is possible to use it to practically apply TPT to a given system. We saw that SM avoids the problem of computing an infinitely long trajectory by converting it into a path-optimization task that doesn't require to actually sample the transition path ensemble. An important question still remains open: is TPT applicable to enhanced path-sampling techniques? In particular, the strength of such methods is to avoid simulating thermal oscillations in the reactant and product states by promoting the generation of short, non-ergodic productive trajectories. If ergodicity is not satisfied, the equivalences in Eq.s (2.11) and (2.23)

are not valid and thus TPT becomes inapplicable. This problem was faced in Ref. [104] and was solved by introducing a suitable partition of the configuration space and building a Markov state model connecting the different regions.

In this section we want to do introduce a different approach that can in principle be applied to every enhanced sampling technique, without the need to partition the configuration space. We are going to show that, under specific assumptions, short and non-ergodic trajectories can be used to recover TPT results, thus allowing one to use enhanced path-sampling results to approximate $q(\mathbf{x})$, $m_T(\mathbf{x})$ and $\mathbf{J}_T(\mathbf{x})$. This fact will be our stepping stone for the development of an efficient algorithm, based on SCPS, to compute the committor function, the reactive probability density and the reactive current for a given system.

2.3.1 Steady Current Regime

We start from the FP equation, presented in appendix A. In the case of activated transitions, the spectrum K of the FP operator shows a *gap*, in the sense that there exist two contiguous eigenvalues $k_F, k_S \in K$ such that

$$k_S \ll k_F \tag{2.53}$$

In the following, we will call *modes (or frequencies or eigenfrequencies) below the gap* all the eigenvalues k_i for which $k_i \leq k_S$. Conversely, *modes above the gap* will be those eigenvalues k_j for which $k_j \geq k_F$. $o(k_S)$ and $o(k_F)$ will provide, respectively, the typical scale for eigenfrequencies below and above the gap.

Let us prepare our system of interest in the reactant state R at time $t = 0$, and let us evolve it by integrating the FP equation. After a time $t \sim k_F^{-1}$ the system will relax into the local reactant well and it will distribute with the corresponding equilibrium distribution restricted to R:

$$P(\mathbf{x}, t \sim t_F^{-1}) \equiv \rho_0(\mathbf{x}) = \frac{e^{-\beta U(\mathbf{x})}}{Z_R} h_R[\mathbf{x}] \qquad Z_R = \int_\Omega d\mathbf{x}\, e^{-\beta U(\mathbf{x})} h_R[\mathbf{x}] \tag{2.54}$$

Letting the system evolve for $t \gtrsim k_F^{-1}$, the probability current

$$\mathbf{J}(\mathbf{x}) = -D(\boldsymbol{\nabla} + \beta \boldsymbol{\nabla} U(\mathbf{x}))P(\mathbf{x}, t) \tag{2.55}$$

will start to flow from R to P. This current will exist until equilibrium is reached, that is for $t \gtrsim k_S^{-1}$, where

$$P(\mathbf{x}, t) = P(\mathbf{x}) = \frac{e^{-\beta U(\mathbf{x})}}{Z} \tag{2.56}$$

and

$$\mathbf{J}(\mathbf{x}) = -D(\boldsymbol{\nabla} + \beta \boldsymbol{\nabla} U(\mathbf{x})) \frac{e^{-\beta U(\mathbf{x})}}{Z} = 0 \tag{2.57}$$

Timescale $t \gtrsim k_S^{-1}$ and the limit $T \to \infty$ used in TPT represent the same kinetic regime, which we previously referred to as ergodic. Now, we want to avoid the necessity to reach the ergodic limit, and focus instead on the single-barrier crossing timescales typical of enhanced sampling methods. This is made possible by the fact that we assumed to have timescale decoupling, which allows us to access a timescale τ which is simultaneously long enough to ensure local relaxation and short enough to avoid that equilibration has been attained:

$$k_F^{-1} \ll \tau \ll k_S^{-1} \tag{2.58}$$

For $t \sim \tau$, the system had time to cross the barrier at most a single time and the dynamics of the current is frozen in time. To see this, let us expand the current on the basis of the right eigenstates of the FP equation (see appendix A, Eq. (**??**))

$$\begin{aligned} \mathbf{J}(\mathbf{x}, t) &= \sum_{n=0}^{\infty} c_n e^{-k_n t} \left[-D(\boldsymbol{\nabla} + \beta \boldsymbol{\nabla} U(\mathbf{x})) R_n(\mathbf{x}) \right] \\ &\equiv \sum_{n=1}^{\infty} c_n e^{-k_n t} \mathbf{J}_n(\mathbf{x}) \end{aligned} \tag{2.59}$$

where the term $n = 0$ is evidently equal to zero because of Eq. (2.57). Let us consider the first dominant terms of the expansion:

$$\mathbf{J}(\mathbf{x}, t) = c_1 e^{-k_1 t} \mathbf{J}_1(\mathbf{x}) + c_2 e^{-k_2 t} \mathbf{J}_2(\mathbf{x}) + o\left(e^{-k_3 t}\right) \tag{2.60}$$

If $t \sim \tau$ is in the SCR, we have

$$\begin{aligned} &\tau \gg k_2^{-1} \sim o(k_F^{-1}) \Rightarrow e^{-k_i \tau} \sim 0 \quad \forall i \geq 2 \\ &\tau \ll k_1^{-1} \sim o(k_S^{-1}) \Rightarrow e^{-k_1 \tau} \sim 1 \end{aligned} \tag{2.61}$$

Putting together Eq.s (2.61) and (2.60) we find

$$\mathbf{J}(\mathbf{x}, \tau) = \mathbf{J}(\mathbf{x}) \sim c_1 \mathbf{J}_1(\mathbf{x}) \tag{2.62}$$

Eq. (2.62) means that the probability current is steady in time if τ is chosen like in Eq. (2.58). For this reason, we will refer to the time interval in Eq. (2.58) as the *Steady Current Regime* (SCR). In the following sections we are going to use the SCR to prove that it is possible to approximate the TPT results by employing short, non-ergodic trajectories, e.g. productive trajectories generated from enhanced sampling techniques.

2.3.2 Committor Function in the Steady Current Regime

Throughout this chapter we are going to implement the following notation. Let $A \subset \Omega$ represent a given state and let ∂A be its surface. We will use the superscript

(A) to label a generic function $f(\mathbf{x})$ on which we are imposing absorbing boundary conditions in ∂A:

$$f^{(A)}(\mathbf{x})\Big|_{\mathbf{x}\in\partial A} = 0 \tag{2.63}$$

In this subsection we want to show how it is possible to approximate the committor function on the SCR. Let us start by introducing two Green functions of the FP equation:

$$\left(\frac{\partial}{\partial t} - H_{\mathrm{FP}}\right) P^{(A)}(\mathbf{x}, t|\mathbf{x}_i, t_i) = \delta(\mathbf{x}-\mathbf{x}_i)\delta(t-t_i) \qquad A = R, P \tag{2.64}$$

The two functions obey the same equation but differ in the choice of the boundary conditions. We can use the two Green functions in Eq. (2.64) to build two distributions

$$\begin{aligned} Q^{(R)}(\mathbf{x}, t) &= \int_\Omega d\mathbf{x}_f \; h_P[\mathbf{x}_f] P^{(R)}(\mathbf{x}_f, t_f|\mathbf{x}, t_f - t) \\ Q^{(P)}(\mathbf{x}, t) &= \int_\Omega d\mathbf{x}_f \; h_R[\mathbf{x}_f] P^{(P)}(\mathbf{x}_f, t_f|\mathbf{x}, t_f - t) \end{aligned} \tag{2.65}$$

where we introduced an intermediate time $t \in [0, t_f]$ and both t and t_f are chosen in the SCR. It should be noted that the characteristic functions in each integral are chosen in such a way that $Q^{(A)}$ functions satisfy the imposed boundary conditions for each $\mathbf{x}_f$. The two distributions in Eq. (2.65) are time-independent in the SCR and they are strictly related to the forward and backward committor functions. To see why, let us focus on $Q^{(R)}(\mathbf{x}, t)$ (the same arguments hold for $Q^{(P)}(\mathbf{x}, t)$ as well) and expand it as an infinite series. One should note that $P^{(R)}(\mathbf{x}_f, t_f|\mathbf{x}, t_f - t)$ is solution of the backward-Kolmogorov equation (see appendix A), thus it should be expanded on the basis of the left-eigenstates of the FP operator [184]:

$$\begin{aligned} P^{(R)}(\mathbf{x}_f, t_f|\mathbf{x}, t_f - t) &= e^{\beta U(\mathbf{x}_f)} \sum_{n=0}^{\infty} L_n^{(R)}(\mathbf{x}) L_n^{(R)}(\mathbf{x}_f) e^{-k_n(t_f - t_f + t)} \\ &= \sum_{n=0}^{\infty} L_n^{(R)}(\mathbf{x}) R_n^{(R)}(\mathbf{x}_f) e^{-k_n t} \end{aligned} \tag{2.66}$$

Plugging Eq. (2.66) into Eq. (2.65) we have

$$\begin{aligned} Q^{(R)}(\mathbf{x}, t) &= \int_\Omega d\mathbf{x}_f \; h_P[\mathbf{x}_f] \sum_{n=0}^{\infty} L_n^{(R)}(\mathbf{x}) R_n^{(R)}(\mathbf{x}_f) e^{-k_n t} \\ &= \sum_{n=0}^{\infty} r_n^{(R)} e^{-k_n t} L_n^{(R)}(\mathbf{x}) \end{aligned} \tag{2.67}$$

where $r_n^{(R)} = \int \mathbf{x}_f \; h_P[\mathbf{x}_f] R_n^{(R)}(\mathbf{x}_f)$. Let us focus on the zeroth component of Eq. (2.67). Absorbing boundary conditions at ∂R make the first eigenfrequency $k_0 \sim k_S$, because every time a trajectory hits the reactant, it gets annihilated by its boundary.

The corresponding right eigenstate $R_0^{(R)}(\mathbf{x})$ corresponds to the equilibrium distribution, but restricted to the product basin because of the boundary conditions:

$$R_0^{(R)}(\mathbf{x}) \sim \frac{e^{-\beta U(\mathbf{x})}}{Z_P} h_P[\mathbf{x}] \qquad Z_P = \int_\Omega d\mathbf{x}\, e^{-\beta U(\mathbf{x})} h_P[\mathbf{x}] \tag{2.68}$$

The expansion coefficient is thus provided by

$$r_0^{(R)} = Z_P^{-1} \int \mathbf{x}_f\, h_P^2[\mathbf{x}_f] e^{-\beta U(\mathbf{x})} = Z_P^{-1} Z_P = 1 \tag{2.69}$$

where we used the fact that $h_A^n[\mathbf{x}] = h_A[\mathbf{x}]\ \forall n > 0$. For what concerns instead the lowest left-eigenstate, we notice that the the orthonormality condition

$$\int d\mathbf{x}\, R_0^{(R)}(\mathbf{x}) L_0^{(R)}(\mathbf{x}) = 1 \tag{2.70}$$

must hold, so $L_0^{(R)}(\mathbf{x})$ is uniformly equal to one within P and it decreases in the transition region to eventually vanish at the boundary of R. Finally, if we choose t in the SCR, we plug Eq. (2.69) into Eq. (2.67) and we recall the limits in Eq. (2.61) we find that

$$Q^{(R)}(\mathbf{x}, t) \sim L_0^{(R)}(\mathbf{x}) \tag{2.71}$$

Similar arguments lead us also to

$$Q^{(P)}(\mathbf{x}, t) \sim L_0^{(P)}(\mathbf{x}) \tag{2.72}$$

As a last step of our derivation, we introduce the time-integrals of $Q^{(R)}$ and $Q^{(P)}$:

$$\begin{aligned} q_{\mathrm{SCR}}^{+}(\mathbf{x}) &= \frac{1}{t_f - t_0} \int_{t_0}^{t_f} dt\, Q^{(R)}(\mathbf{x}, t) \\ q_{\mathrm{SCR}}^{-}(\mathbf{x}) &= \frac{1}{t_f - t_0} \int_{t_0}^{t_f} dt\, Q^{(P)}(\mathbf{x}, t) \end{aligned} \tag{2.73}$$

where t_0 and t_f are introduced in such a way that the integration range is limited to timescales in the SCR, which means that t_0 and t_f are any times that satisfy

$$(t_0 k_F)^{-1} \ll 1 \qquad t_f k_S \ll 1 \tag{2.74}$$

Eq.s (2.71), (2.72), (2.73) and (2.74) let us finally conclude that

$$q_{\mathrm{SCR}}^{+/-}(\mathbf{x}) \sim L_0^{(R)/(P)}(\mathbf{x}) \tag{2.75}$$

which also means that

$$H_{\mathrm{FP}}^{\dagger} q_{\mathrm{SCR}}^{+/-}(\mathbf{x}) \sim 0 \tag{2.76}$$

Eq. (2.75) provides the approximation of the committor function in the SCR, thus it will be our basic ingredient to build the reactive distribution and current in this same time regime. Indeed, $q_{\mathrm{SCR}}^{+/-}(\mathbf{x})$ is solution of the backward-Kolmogorov equation

and it also satisfies the same boundary conditions of the usual committor function. Our approximation is *rigorous* in the sense that it is controlled by two well-defined expansion parameters k_F and k_S. Because of the properties of $L_0^{(R)/(P)}$ we also find that

$$q^-_{\mathrm{SCR}}(\mathbf{x}) = 1 - q^+_{\mathrm{SCR}}(\mathbf{x}) \tag{2.77}$$

Therefore, we will drop the + superscript and always use the more concise notation $q_{\mathrm{SCR}}(\mathbf{x}) \equiv q^+_{\mathrm{SCR}}(\mathbf{x})$, unless differently specified.

The statement about the boundary conditions could use some further comments. There are some evident differences in the way in which boundary conditions are enforced here with respect to how the committor function is usually defined. When we were discussing the general definition of the committor function in section 2.1.2, we imposed absorbing boundary conditions on *both the product and the reactant state*, while here q^+_{SCR} is defined by just requiring absorbing boundary conditions on ∂R (the opposite holds for q^-_{SCR}.) This choice is potentially dangerous, because in $Q^{(P)}(\mathbf{x})$, for example, trajectories that enter ∂R and leave it again might provide a sensible contribution. Fortunately this is not the case, and the proof of this technical point is left to appendix C.3. We can finally conclude that, up to contributions in the SCR expansion parameters k_F and k_S,

$$q_{\mathrm{SCR}}(\mathbf{x})|_{\mathbf{x}\in\partial R} \sim 0 \qquad q_{\mathrm{SCR}}(\mathbf{x})|_{\mathbf{x}\in\partial P} \sim 1 \tag{2.78}$$

2.3.3 Reactive Probability Density and Current

Let us now discuss how it is possible to retrieve the reactive probability density defined in Eq. (2.21) in the SCR. Reactive pathways in the SCR perform single barrier-crossings and sample a density distribution given by

$$\begin{aligned} m_{\mathrm{SCR}}(\mathbf{x}) \equiv \frac{1}{t_f - t_0} \int_{t_0}^{t_f} dt \int_\Omega d\mathbf{x}_i \int_\Omega d\mathbf{x}_f \; h_P[\mathbf{x}_f] P^{(R)}(\mathbf{x}_f, t_f | \mathbf{x}, t) \cdot \\ \cdot P^{(P)}(\mathbf{x}, t | \mathbf{x}_i, 0) \rho_0(\mathbf{x}_i) h_R[\mathbf{x}_i] \end{aligned} \tag{2.79}$$

where $\rho_0(\mathbf{x}_i)$ was introduced in Eq. (2.54). Because both $P^{(R)}$ and $P^{(P)}$ are included in the integral in Eq. (2.79), we are sure that the only trajectories that contribute in $m_{\mathrm{SCR}}(\mathbf{x})$ are the reactive ones, i.e. trajectories that do not backtrack to the product state up to corrections of the order of the SCR expansion parameters. We want to show that Eq. (2.79) can be expressed as

$$m_{\mathrm{SCR}}(\mathbf{x}) \propto e^{-\beta U(\mathbf{x})} q_{\mathrm{SCR}}(\mathbf{x})(1 - q_{\mathrm{SCR}}(\mathbf{x})) \tag{2.80}$$

which is the SCR equivalent of Eq. (2.21). To see this, first of all we introduce the distribution

$$P^{(P)}(\mathbf{x}, t) = \int_\Omega d\mathbf{x}_i \; P^{(P)}(\mathbf{x}, t | \mathbf{x}_i, 0) \rho_0(\mathbf{x}_i) h_R[\mathbf{x}_i] \tag{2.81}$$

The detailed balance condition ensures that

$$P^{(P)}(\mathbf{x},t|\mathbf{x}_i,0) = e^{\beta U(\mathbf{x}_i)} P^{(P)}(\mathbf{x}_i,t|\mathbf{x},0) e^{-\beta U(\mathbf{x})} \tag{2.82}$$

Plugging the right-hand side of Eq. (2.82) into Eq. (2.81) we find

$$\begin{aligned} P^{(P)}(\mathbf{x},t) &= e^{-\beta U(\mathbf{x})} \int_\Omega d\mathbf{x}_i \; P^{(P)}(\mathbf{x},t|\mathbf{x}_i,0) e^{\beta U(\mathbf{x}_i)} \rho_0(\mathbf{x}_i) h_R[\mathbf{x}_i] \\ &= e^{-\beta U(\mathbf{x})} \int_\Omega d\mathbf{x}_i \; P^{(P)}(\mathbf{x},t|\mathbf{x}_i,0) e^{\beta U(\mathbf{x}_i)} \frac{e^{-\beta U(\mathbf{x}_i)}}{Z_R} h_R^2[\mathbf{x}_i] \\ &= \frac{e^{-\beta U(\mathbf{x})}}{Z_R} Q^{(P)}(\mathbf{x},t) \end{aligned} \tag{2.83}$$

where in the last step we used the definition in Eq. (2.65). Using the result of Eq. (2.83) in Eq. (2.79) we obtain

$$\begin{aligned} m_{\mathrm{SCR}}(\mathbf{x}) &= \frac{1}{t_f - t_0} \int_{t_0}^{t_f} dt \int_\Omega d\mathbf{x}_f \; h_P[\mathbf{x}_f] P^{(R)}(\mathbf{x}_f,t_f|\mathbf{x},t) \frac{e^{-\beta U(\mathbf{x})}}{Z_R} Q^{(P)}(\mathbf{x},t) \\ &= \frac{e^{-\beta U(\mathbf{x})}}{Z_R} \left[\frac{1}{t_f - t_0} \int_{t_0}^{t_f} dt \; Q^{(R)}(\mathbf{x},t_f - t) Q^{(P)}(\mathbf{x},t) \right] \end{aligned} \tag{2.84}$$

Recalling that $Q^{(R)/(P)}(\mathbf{x},\tau) \sim Q^{(R)/(P)}(\mathbf{x})$ if τ is chosen in the SCR and using the definition of $q_{\mathrm{SCR}}(\mathbf{x})$ in Eq. (2.73), we finally obtain the result in Eq. (2.80), up to an irrelevant normalization factor.

Similarly to what we did in Eq. (2.79), the reactive current in the SCR can be readily defined as

$$\mathbf{J}_{\mathrm{SCR}}(\mathbf{x}) = -\frac{D}{t_f - t_0} \int_{t_0}^{t_f} dt \; Q^{(R)}(\mathbf{x},t_f - t) \left(\overrightarrow{\boldsymbol{\nabla}} - \overleftarrow{\boldsymbol{\nabla}} + \beta \boldsymbol{\nabla} U(\mathbf{x}) \right) P^{(P)}(\mathbf{x},t) \tag{2.85}$$

where the arrows on $\overrightarrow{\boldsymbol{\nabla}}$ and $\overleftarrow{\boldsymbol{\nabla}}$ indicate, respectively, that the gradient acts on right and on the left. $-\overleftarrow{\boldsymbol{\nabla}}$ term is included to account for re-crossing contributions. Resorting to the same arguments that led us from Eq. (2.79) to Eq. (2.80), we obtain

$$\mathbf{J}_{\mathrm{SCR}}(\mathbf{x}) \propto D e^{-\beta U(\mathbf{x})} \boldsymbol{\nabla} q_{\mathrm{SCR}}(\mathbf{x}) \tag{2.86}$$

which is the realization of the reactive current in the SCR and it is the equivalent of the TPT result in Eq. (2.25). Again, the normalization factor is practically irrelevant. The detailed calculation to obtain Eq. (2.86) can be found in appendix C.2.

2.4 Exploiting q_{SCR} to Sample the Transition Region

Since the results obtained in the SCR correctly approximate the ones of TPT, the committor function continues to play a pivotal role in our theory. It is fundamental

at this point to find a computationally efficient way to estimate it, even for complex and ultra-long conformational transitions like protein folding. In this section, we are going to show that SCPS is the key to do it. This ie because two main reasons: The SCR analog of the committor function can be recovered from the knowledge of $s_\lambda(\mathbf{x})$ and the average path provides a mean-field approximation of a principal curve. In this sense, SCPS might be considered as an algorithm to approximate principal curves.

The residual problem of SCPS is that the pathways generated from it do sample exactly the transition region. Indeed, the best we can get from SCPS is a mean-field approximation of the transition path ensemble, where the single trajectories are not endowed with any kinetic meaning. In order to correct for this inaccuracy, in this section we will present two possible algorithms to sample the transition region that exploit the knowledge of q_{SCR} coming from SCPS calculations.

The first algorithm takes advantage of the fact that it is possible to exactly introduce a position-dependent biasing force in the Langevin equation by accounting for absorbing boundary conditions. This effective Langevin equation defines a stochastic process whose stationary distribution is the reactive probability density $m_T(\mathbf{x})$. The formalism is based on a generalization of the concept of *Langevin bridges* [172, 173], in a way that will become clearer in the following. The generation of reactive pathways in high-dimensional and rough energy landscapes might be however complicated, even with the help of the position dependent bias force. This is because reactive pathways can sensibly *detour* from the dominant reaction channels by, e.g. populating kinetic traps [19, 98].

To solve this problem, we propose a new algorithm that samples instead the Boltzmann distribution restricted to Ω_T by generating trajectories that travel only forward in q_{SCR}, thus avoiding the problem of detours. The main drawback of this approach is that such a dynamics is obtained by resorting to a history-dependent bias force, which breaks microscopic reversibility. In this sense, these new trajectories cannot be directly interpreted as physical reactive events.

The schemes we are introducing in this section provide a continuum generalization of the algorithms introduces in Ref.s [19, 98] in the framework of Markov jump processes.

2.4.1 Relationship Between the Average Path and Principal Curves

The average path defined by SCPS can be visualized as a mean-field approximation of a principal curve. To see why this is true, let us consider a generic parametrization

of the average path as a function of an arc-length α:

$$\gamma = \{\boldsymbol{\phi}(\alpha) \ : \ \alpha \in [0,1]\} \qquad \boldsymbol{\phi} : [0,1] \to \Omega \tag{2.87}$$

Let us also define a function $\tilde{s}_\gamma$ as

$$\tilde{s}_\gamma(\mathbf{x}) = 1 - s_\lambda(\mathbf{x}) \tag{2.88}$$

The natural way to parametrize the average path is with respect of time, because all the trajectories contributing to γ have the same length. We can define $\alpha = \tau/t$ where t is the total length of SCPS simulations. In this way, Eq. (2.88) can be rewritten as

$$\tilde{s}_\gamma(\mathbf{x}) = \frac{\int_0^1 d\alpha \ \alpha e^{-\lambda||C(\mathbf{x})-\boldsymbol{\phi}(\alpha)||^2}}{\int_0^1 d\alpha \ e^{-\lambda||C(\mathbf{x})-\boldsymbol{\phi}(\alpha)||^2}} \tag{2.89}$$

We recall that Eq. (2.89) is the result of a mean-field approximation (see section 1.3). Before the approximation we had

$$\tilde{s}_\gamma = 1 - s_\lambda = \frac{\tau}{t} = \alpha \tag{2.90}$$

This means that $\tilde{s}_\gamma(\mathbf{x})$ is the mean-field approximation of a function that associates to each configuration $\mathbf{x}$ the value of α corresponding to the point on γ closest to $\mathbf{x}$:

$$\tilde{s}_\gamma : \Omega \to [0,1] \tag{2.91}$$

This is exactly the definition of function $s_\gamma(\mathbf{x})$, reported in Eq. (1.33). Now, the SCPS average path is defined as

$$\boldsymbol{\phi}(\alpha) = \langle C(\mathbf{x}) \rangle_\alpha \sim \langle C(\mathbf{x}) \rangle_{\tilde{s}_\gamma(\mathbf{x})=\alpha} \tag{2.92}$$

Eq. (2.92) is the result we were looking for: apart from a mean-field approximation and provided that $\alpha = \tau/t$ parametrization is chosen to represent the curve, the SCPS average-path is an approximation of a principal curve.

2.4.2 Computing $q_{\mathbf{SCR}}$ using SCPS

We discussed in section 2.2.1 that the string method approximates the committor function under the assumption that $s_\gamma(\mathbf{x})$, in Eq. (2.35), is a re-parametrization of $q(\mathbf{x})$ through an unknown function $f(s)$. In this section we are going to do the same, but substituting $s_\gamma(\mathbf{x})$ with $\tilde{s}_\gamma(\mathbf{x})$ computed from SCPS. This choice is particularly convenient because it can be proved that iso-q_{SCR} and iso-$\tilde{s}$ hyperplanes coincide, up to relabeling, in the vicinity of the average path. This proof is technical and it is left to appendix C.4. In the following, we will assume that the interpretation of iso-s hyperplanes as iso-q_{SCR} ones holds throughout the whole transition region:

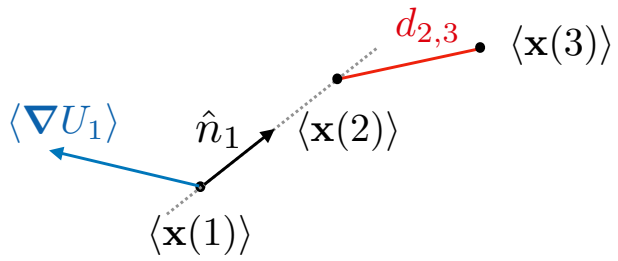

Figure 2.6: Schematic representation of the ingredients in Eq. (2.93).

Thus the problem becomes how to recover the values of $q(\mathbf{x})$ given $\tilde{s}_\gamma(\mathbf{x})$, i.e. *relabel* iso-$s$ hyperplanes. To tackle this problem, we recall that $q_{\mathrm{SCR}}(\mathbf{x})$ is a solution of the backward-Kolmogorov equation, provided that suitable absorbing boundary conditions are imposed. Since $\boldsymbol{\nabla}\tilde{s}_\gamma(\mathbf{x}) \parallel \boldsymbol{\nabla} q_{\mathrm{SCR}}(\mathbf{x})\ \forall \mathbf{x} \in \Omega_T$, if $\mathbf{x}$ is a point on the average path the only direction in which the two gradients are non-zero is the one along the average path itself. For this reason, if we compute the committor on the N points defining the mean path, $\{\langle \mathbf{x}(i) \rangle\}$, its discrete values $\{q_i = q(\langle \mathbf{x}(i) \rangle)\}$ will obey a discretized version of the backward-Kolmogorov equation:

$$\frac{1}{d_{i+1,i}} \left(\frac{q_{i+1} - q_i}{d_{i+1,i}} + \frac{q_{i-1} - q_i}{d_{i,i-1}} \right) + \frac{\beta}{d_{i+1,i}} (q_{i+1} - q_i) \hat{\mathbf{n}}_i \cdot \langle -\boldsymbol{\nabla} U_i \rangle \sim 0 \tag{2.93}$$

where $\hat{\mathbf{n}}_i$ is the vector tangent to the average path in q_i, $\langle -\boldsymbol{\nabla} U_i \rangle$ is the average force acting on the i-th frame of the path and $d_{i+1,i}$ is the Euclidean distance between $\langle \mathbf{x}(i+1) \rangle$ and $\langle \mathbf{x}(i) \rangle$ (see Fig. 2.6). The discrete set of points $\{\tilde{s} = \tilde{s}_\gamma(\langle \mathbf{x}(i) \rangle)\}$ obtained by computing $\tilde{s}_\gamma(\mathbf{x})$ on the points of the average path does not satisfy Eq. (2.93), because $\tilde{s}_\gamma(\mathbf{x}) \neq q_{\mathrm{SCR}}(\mathbf{x})$. However, $\{\tilde{s}_i\}$ can be optimized in such a way to satisfy the discrete backward-Kolmogorov equation just by looking for the minimum of the functional

$$I\left[\{a_i\}\right] = \sum_{i=0}^{N} \left[\frac{1}{d_{i+1,i}} \left(\frac{a_{i+1} - a_i}{a_{i+1,i}} + \frac{a_{i-1} - a_i}{d_{i,i-1}} \right) + \frac{\beta}{d_{i+1,i}} (a_{i+1} - a_i) \hat{\mathbf{n}}_i \cdot \langle -\boldsymbol{\nabla} U_i \rangle \right]^2 \tag{2.94}$$

The functional in Eq. (2.94) can be initialized with $\{\tilde{s}_i\}$ and it attains its minimum for $\{a_i\} = \{q_i\}$. The optimization $\delta I = 0$ returns a value of q for each iso-$\tilde{s}$ surface, thus defining a map $q(\tilde{s}_\gamma)$ which relabels iso-$\tilde{s}$ hyperplanes.

To summarize, the strategy we propose to compute the committor function is the following:

1. simulate the reactive process of interest using SCPS algorithm (see section 1.3.2), starting from many different configurations;

2. compute $\tilde{s}_\gamma(\mathbf{x})$ function for the points in the average path;

3. optimize the values of $\tilde{s}$ on the average path by minimizing the functional in Eq. (2.94);

4. assign to every iso-$\tilde{s}$ hyperplane the corresponding value of $q(\tilde{s})$ obtained from the minimization;

5. since the committor function is continuous, values of q for points not included in the average path can be obtained by interpolation.

We stress that this procedure is more general than the computation of the committor function just in the vicinity of the principal curve. Indeed, $\tilde{s}_\gamma(\mathbf{x})$ is computed over all the region of configuration space that is sampled by SCPS. The fact that, for a point on the average path, the gradient of the committor function vanishes in every direction except the one of the average path itself is used just to relabel iso-$\tilde{s}$ hypersurfaces. Of course we expect our estimation of the committor function to be less precise in the regions that are rarely sampled by SCPS.

As a final remark, it should be noted that $\tilde{s}_\gamma(\mathbf{x})$ is defined starting from a single initial condition. For high-dimensional systems, when entropy plays an important role, the description of the reaction mechanism coming from many independent simulations started from the same initial condition might not be sufficient. In this case, the optimization procedure should be repeated for every set of $\{\tilde{s}_i\}$ computed over all the average pathways resulting from the simulations. In general, average pathways will sensibly differ from one simulation to another. Hypersurfaces corresponding to the same value of the committor function can be finally interpolated to better define iso-q surfaces spanning over the whole accessible configuration space. The required non-linear interpolation is provided, for example, by the $\sigma(\mathbf{x})$ function defined in section 1.3.2.

2.4.3 Conditional Langevin Dynamics

We know that, for timescales in the steady current regime, the reactive probability distribution $m_T(\mathbf{x})$ can be rigorously approximated with its SCR analog $m_{\mathrm{SCR}}(\mathbf{x})$. Here we want to show that it is possible to sample $m_{\mathrm{SCR}}(\mathbf{x})$ by means of the following stochastic equation of motion:

$$\dot{\mathbf{x}} = D\left(-\beta\boldsymbol{\nabla}U(\mathbf{x}) + 2\frac{\boldsymbol{\nabla}q_{\mathrm{SCR}}(\mathbf{x})}{q_{\mathrm{SCR}}(\mathbf{x})}\right) + \boldsymbol{\eta}(t) \tag{2.95}$$

Eq. (2.95) is already known in the literature, and it was first derived in Ref.s [185], while in Ref.s [19, 98] its discrete analog was obtained in the context of Markov jump processes. In this section we will derive Eq. (2.95) again, in a way that clearly shows that it can be used to sample $m_{\mathrm{SCR}}(\mathbf{x})$. We start by re-writing the reactive

distribution in the SCR in a convenient way

$$\begin{aligned} m_{\mathrm{SCR}}(\mathbf{x}) &= \frac{1}{t_f - \tau_0} \int_{\tau_0}^{t_f} dt\; Q^{(R)}(\mathbf{x}, t_f - t) \frac{e^{-\beta U(\mathbf{x})}}{Z_R} Q^{(P)}(\mathbf{x}, t) \\ &= \frac{1}{t_f - \tau_0} \int_{\tau_0}^{t_f} dt\; Q^{(R)}(\mathbf{x}, t_f - t) P^{(P)}(\mathbf{x}, t) \\ &\sim \frac{1}{t_f - \tau_0} \int_{\tau_0}^{t_f} dt\; Q^{(R)}(\mathbf{x}, t) P^{(P)}(\mathbf{x}, t) \\ &\equiv \frac{1}{t_f - \tau_0} \int_{\tau_0}^{t_f} dt\; \mathbb{P}(\mathbf{x}, t) \end{aligned} \tag{2.96}$$

where we used the result in Eq. (2.83) and the fact that, in the SCR, Q and P functions are almost time-independent. Sampling the distribution $\mathbb{P}(\mathbf{x}, t)$ is identical to sample $m_{\mathrm{SCR}}(\mathbf{x})$ as long as t is chosen in the SCR. One can show (see appendix C.5) that $\mathbb{P}(\mathbf{x}, t)$ satisfies the following effective Fokker-Planck equation:

$$\frac{\partial \mathbb{P}(\mathbf{x}, t)}{\partial t} = D\boldsymbol{\nabla} \cdot \Big[\boldsymbol{\nabla} + \beta \boldsymbol{\nabla} U(\mathbf{x}) - 2\boldsymbol{\nabla} \log Q^{(R)}(\mathbf{x}, t)\Big] \mathbb{P}(\mathbf{x}, t) \tag{2.97}$$

The underlying microscopic dynamics satisfies detailed balance, because, as it is clear from the derivation in appendix C.5, the non-linear term $-2\boldsymbol{\nabla} \log Q^{(R)}(\mathbf{x}, t)$ comes just from imposing reflecting boundary conditions at the borders of the product and the reactant states. Distribution $\mathbb{P}(\mathbf{x}, t)$ can thus be sampled by an effective Langevin dynamics of the form

$$\dot{\mathbf{x}} = D\left(-\beta \boldsymbol{\nabla} U(\mathbf{x}) + 2\boldsymbol{\nabla} \log Q^{(R)}(\mathbf{x}, t)\right) + \boldsymbol{\eta}(t) \tag{2.98}$$

The dynamics in Eq. (2.98) is known in literature as *conditional Langevin dynamics* (CLD). In general, Eq. (2.98) is not straightforward to use because of the implicit dependence on $Q^{(R)}(\mathbf{x}, t)$, and the latter term is usually treated by resorting to some approximations [172, 173]. However, if the timescale is chosen in the SCR, as prescribed by the integration limits in Eq. (2.96), we know that $Q^{(R)}(\mathbf{x}, t)$ becomes time-independent and that it provides the SCR representation of the committor function. Thus, the conditional Langevin dynamics that can be employed to sample $m_{\mathrm{SCR}}(\mathbf{x})$ is precisely the one reported in Eq. (2.95). The position-dependent $2\boldsymbol{\nabla} \log q_{\mathrm{SCR}}(\mathbf{x})$ bias entering Eq. (2.95) is *optimal*, in the sense that conditional Langevin dynamics provides the same transition path ensemble of the original Langevin equation in the limit in which the committor $q_{\mathrm{SCR}}(\mathbf{x})$ is exact. Since the result of the SCPS calculation is a map $q(\sigma)$, we can use the chain-rule to rewrite the bias force as

$$\mathbf{F}_{\mathrm{CLD}}(\mathbf{x}) = 2\frac{\boldsymbol{\nabla} q_{\mathrm{SCR}}(\mathbf{x})}{q_{\mathrm{SCR}}(\mathbf{x})} = 2q'_{\mathrm{SCR}}(\sigma)\frac{\boldsymbol{\nabla}\sigma(\mathbf{x})}{q_{\mathrm{SCR}}(\sigma(\mathbf{x}))} \tag{2.99}$$

where $\boldsymbol{\nabla}\sigma(\mathbf{x})$ is a by-product of the SCPS calculation and both $q'_{\mathrm{SCR}}(\sigma) = dq_{\mathrm{SCR}}(\sigma)/d\sigma$ and $q_{\mathrm{SCR}}(\sigma(\mathbf{x}))$ result from the functional optimization in Eq. (2.94).

It is clear that, without an a-priori estimation of the committor, Eq. (2.95) is useless. Although we propose the use of the SCPS estimate of the committor function, every method that approximates $q(\mathbf{x})$ can be coupled with the conditional Langevin dynamics in Eq. (2.95) [110, 137, 149, 150, 151]. Just to give an example, diffusion maps-directed sampling [186] can be easily coupled with the scheme we are proposing.

It is important to stress at this point that Eq. (2.95) should not be used to generate infinitely long trajectories. Indeed, in that case one would converge to the distribution

$$\tilde{\mathbb{P}}(\mathbf{x}) = \frac{1}{N} q_{\mathrm{SCR}}^2(\mathbf{x}) e^{-\beta U(\mathbf{x})} \tag{2.100}$$

with N a suitable normalization factor. $\tilde{\mathbb{P}}(\mathbf{x}) \neq m_{\mathrm{SCR}}(\mathbf{x})$ and it has no direct physical interpretation. The correct way to sample $m_{\mathrm{SCR}}(\mathbf{x})$ from Eq. (2.95) is by implementing the following non-equilibrium process. We start by sampling initial conditions on the iso-committor hyperplane $q_{\mathrm{SCR}}(\mathbf{x}) = \varepsilon$, where ε is arbitrarily small. These points are provided, for example, by the configurations visited by SCPS on the $q = \varepsilon$ iso-committor surface. Starting from those points, we run several trajectories integrating Eq. (2.95). If, at time $t_{\bar{q}}$, a trajectory x hits the hyperplane $q_{\mathrm{SCR}}(x(t_{\bar{q}})) = \bar{q}$, where $\bar{q} \sim 1-\varepsilon$, the trajectory is absorbed by the boundary conditions on ∂P and thus the calculation is stopped. This non-equilibrium dynamics embeds the correct absorbing boundary conditions and samples $m_{\mathrm{SCR}}(\mathbf{x})$.

2.4.4 Ideal rMD

Let us define a new type of rMD, where instead of $z(\mathbf{x})$ or $s_\lambda(\mathbf{x})$ and $w_\lambda(\mathbf{x})$ we use the committor function as a collective coordinate. The ratchet force is given by

$$\mathbf{F}_{\mathrm{irMD}}(\mathbf{x},t) = k_R \; q'_{\mathrm{SCR}}(\sigma) \; \boldsymbol{\nabla}\sigma(\mathbf{x}) \; \theta\left(q^M_{\mathrm{SCR}}(t) - q_{\mathrm{SCR}}(\mathbf{x})\right) \; \xi\left(q^M_{\mathrm{SCR}}(t) - q_{\mathrm{SCR}}(\mathbf{x})\right) \tag{2.101}$$

where

$$q^M_{\mathrm{SCR}}(t) = \max_{\tau\in[0,t]} q_{\mathrm{SCR}}(x(\tau)) \tag{2.102}$$

and $\xi(y)$ is a non-negative function for every $y > 0$. The original Langevin dynamics can be modified to accommodate the bias force in Eq. (2.101):

$$\dot{\mathbf{x}} = D(-\beta\boldsymbol{\nabla}U(\mathbf{x}) + \mathbf{F}_{\mathrm{irMD}}(\mathbf{x},t)) + \boldsymbol{\eta}(t) \tag{2.103}$$

We refer to the dynamics in Eq. (2.103) as *ideal ratchet-and-pawl molecular dynamics* (irMD). Because of the Heaviside-theta function in Eq. (2.101), the functional form of $\xi(y)$ becomes relevant only when the system backtracks along the committor function. In the $k_R \to \infty$ limit, where backtracking is exactly forbidden, the functional shape of $\xi(y)$ becomes irrelevant.

The strength of this method resides in the fact that the mapping between σ and q is generally non-linear, especially when metastable states are present. Indeed, the committor function usually increases in a strict monotonic way until it reaches a metastable state, inside which the committor remains approximately constant. This behavior results from the fact that points in a basin are kinetically close, thus the probability to reach the product state does not sensibly vary within a well. The result of this non-linearity is that the bias force vanishes once an irMD trajectory reaches a metastable state, allowing for an exhaustive sampling of the state. Intuitively, this tells us that irMD is suited for sampling the Boltzmann distribution restricted to the transition region: This is proved in appendix C.6 in a more mathemcatically sound fashion. irMD defines a non-equilibrium process where short trajectories progress only forward along the committor function, avoiding the problem of detours present in the conditional Langevin dynamics (Eq. (2.95)). The price to pay to avoid detours is, not surprisingly, the loss of microscopic reversibility as a consequence of the history-dependent bias force in Eq. (2.101). This means that each single trajectory computed by irMD cannot be separately interpreted as a member of the transition path ensemble. We note that irMD with $\xi(y) = y$ is a continuum generalization of the *no-detour dynamics* introduced in Ref. [19, 98] in the context of discrete Markov jump processes.

As a final remark, we stress that the reactive probability density can be recovered from irMD simulations simply by multiplying the resulting distribution with the reactive factor $q(1-q)$:

$$e^{-\beta U(\mathbf{x})}\Big|_{\Omega_T} \to e^{-\beta U(\mathbf{x})}\Big|_{\Omega_T} q(\mathbf{x})(1-q(\mathbf{x})) = e^{-\beta U(\mathbf{x})} q(\mathbf{x})(1-q(\mathbf{x})) = m_T(\mathbf{x}) \quad (2.104)$$

2.4.5 A Summary Pipeline

Fig. 2.7 summarizes the steps that have to be performed to apply the pipeline presented in this chapter. First of all, a preliminary simulation is done using bare rMD, which efficiently generates productive events from a given reactant R to a given product P. The result of rMD is a set of average configurations $\{\langle \mathbf{x}(i)\rangle_{\mathrm{Guess}}\}$ that provide a first, sub-optimal approximation of the principal curve connecting R to P. In principle the mean path can be specified by a list of specific features $\{\langle F(i)\rangle_{\mathrm{Guess}}\}$: in the original formulation of SCPS in Ref. [28] the features were provided by contact maps, $F = C_{jk}$, while in Ref. [137] the features were RMSDs computed along a pre-determined path, $\langle F(i)\rangle = F(i) = \mathrm{RMSD}(i)$. As we discussed in section 1.3.2 and in the original work [28], contact maps represent a convenient choice for systems as big as fast-folding proteins, but in principle one is free to use whatever feature better describes the conformational transition.

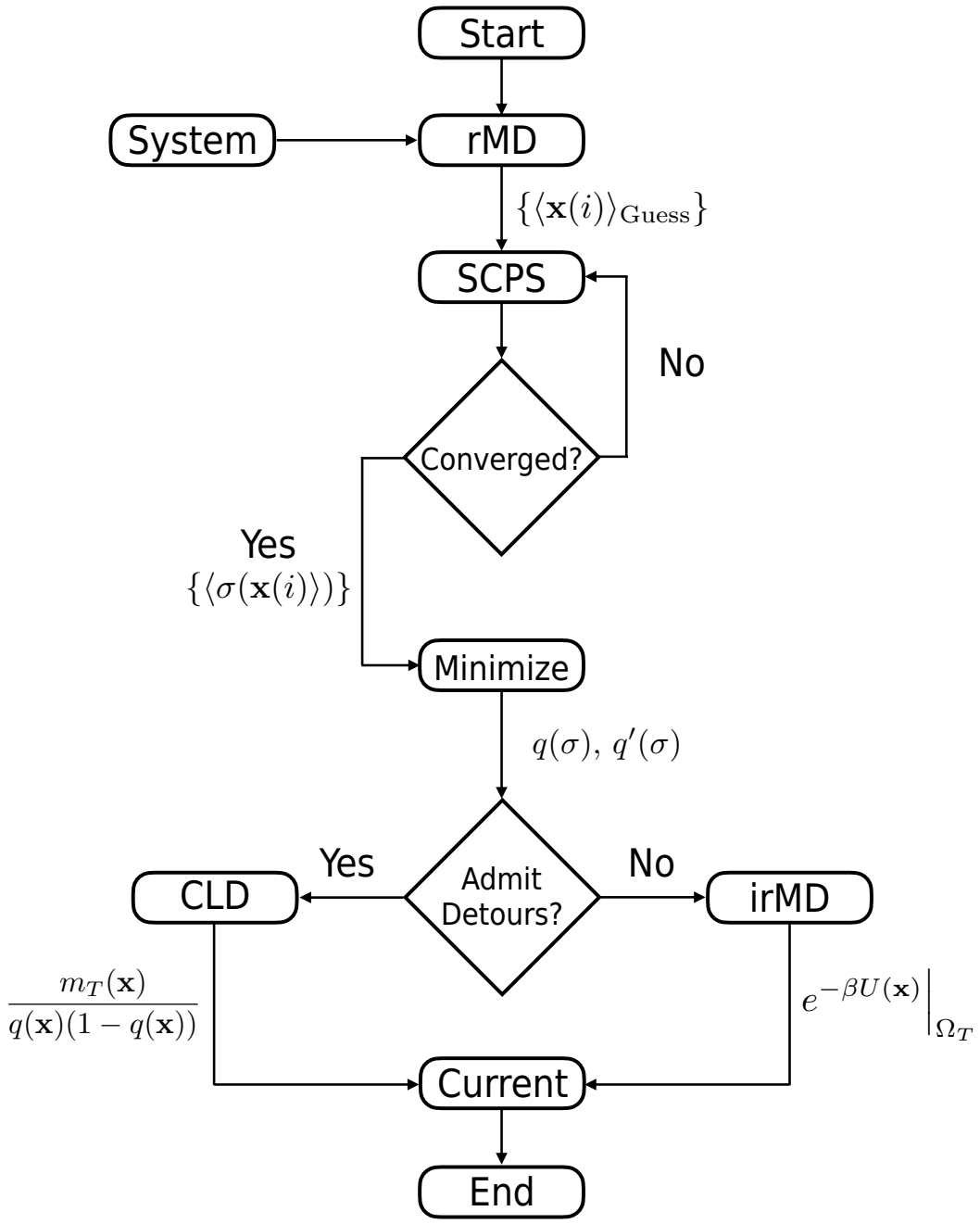

Figure 2.7: Schematic representation of the algorithm that allows to recover TPT results from SCPS simulations.

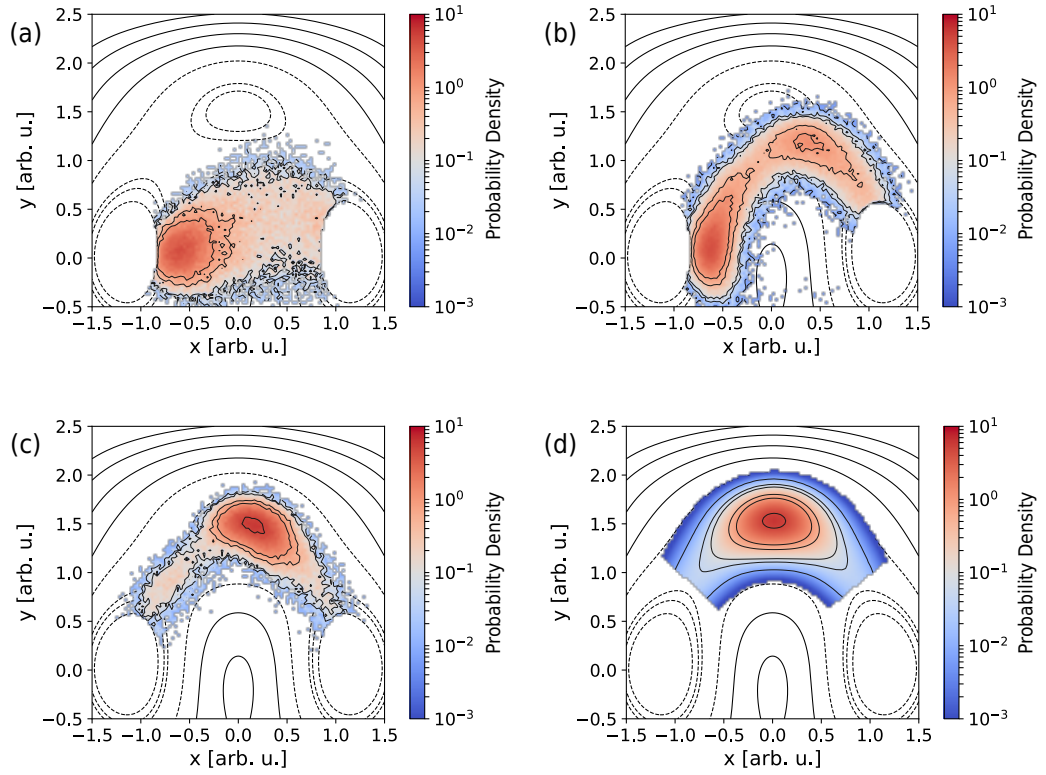

Figure 2.8: Reactive probability densities obtained by means of (a) steered MD, (b) rMD, (c) SCPS compared to (d) the exact one.

The guess mean-path is used as the input of a SCPS calculation, which, at convergence, provides two important results: a mean-field estimation $\{\langle \mathbf{x}(i) \rangle\}$ of the principal curve connecting R to P and the values of $\sigma(\mathbf{x})$ computed on this same curve, $\{\sigma(\langle \mathbf{x}(i) \rangle)\}$. The latter are given as an input to the functional in Eq. (2.94).

Minimization of Eq. (2.94) is a trivial computational task (functional is one-dimensional) and it returns the relabeling map $q(\sigma)$. The derivative $q'(\sigma)$ is readily obtained from $q(\sigma)$. At this point, one is finally ready to sample the reactive probability distribution $m_T(\mathbf{x})$. Depending on the complexity of the system and on the typical timescales of detours, one should decide whether CLD or irMD is better suited for its problem. CLD is more computationally demanding because of detours, but trajectories computed with Eq. (2.95) are true reactive trajectories. irMD, instead, capitalizes on the computational efficiency of rMD to sample the Boltzmann distribution in the transition region. However, each single trajectory does not carry kinetic meaning because of the presence of a history-dependent biasing force in the equations of motion.

Once $m_T(\mathbf{x})$ or $e^{-\beta U(\mathbf{x}}\big|_{\Omega_T}$ are known, they can be fed to the algorithm presented in section 2.1.4 to finally compute the reactive current. At that point, the whole kinetic information about the system is reconstructed and the conformational

transition is fully characterized.

2.5 Illustrative Application on a 2D Potential

To test our algorithm for the calculation of the committor function and sampling of the transition region, we employ the results on the 2-dimensional potential presented in section 1.4. In this simple setting, there are two fundamental simplifications: (i) it is possible to run CLD trajectories without caring about the problem of detours. In general this is not true a priori and one should carefully consider which method, among irMD and CLD, is best suited for its application; (ii) the small entropy within the reactant state avoids the necessity to run trajectories starting from many different initial configurations. Thus, the reaction mechanism can be fully characterized by means of $\tilde{s}_\gamma$.

In the following section we will use the mean path $\{\langle \mathbf{x}(i) \rangle\}$ obtained from the last iteration of the SCPS calculation to compute $\tilde{s}_\gamma(\mathbf{x})$, and optimize its values along the average path to obtain the corresponding committor function $q(\tilde{s}_\gamma)$. Then we will apply the knowledge of the committor function to integrate the CLD, Eq. (2.95), and irMD, Eq. (2.103), equations of motion. Finally, we will employ the algorithm discussed in section 2.1.4 to compute the transition tubes and identify the reaction channel.

In Fig. 2.8 we report the transition regions sampled by (a) sMD, (b) rMD and (c) SCPS compared with (d) the exact one. The exact result has been computed by means of the TPT definition in Eq. (2.20), where the committor function has been estimated using the procedure described in appendix C.7. The other results (sMD, rMD and SCPS) have been obtained by considering the whole ensemble of trajectories and selecting only the reactive portions of the simulations. Evidently, sMD fails in determining the location of the transition region, and populates only the portion of Ω surrounding a transition state that is never visited by reactive trajectories. rMD simulations, instead, do a much better job in determining the shape of the transition region, but present however many defects, as already discussed in section 1.4.1. On the contrary, SCPS provides a reasonable approximation of the exact transition region. Two main inaccuracies emerge: on the one hand, the peak of the distribution is slightly shifted to the right with respect to the exact one, possibly as a residual result of the strong biasing force employed in the rMD guess simulations; on the other hand, the tails of the distribution are not exhaustively sampled, probably as a consequence of the bias force acting along the w_λ variable, which controls the width of the tube around the mean path.

Before going on, it is worth discussing the properties of the average pathways

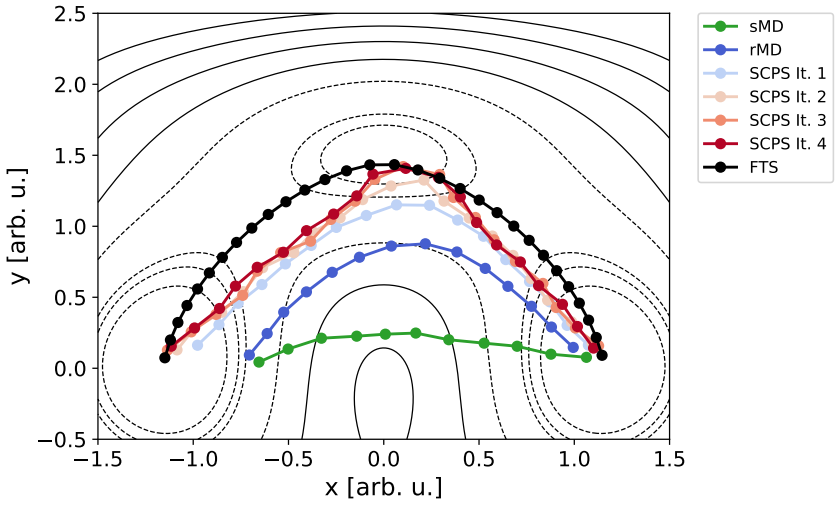

Figure 2.9: Average pathways obtained from sMD, rMD and the four different iterations of SCPS, together with the principal curve computed using the FTS method. All the pathways are superimposed to the level curves of the potential in Eq. (1.55).

obtained from the different simulations strategies we considered in section 1.4. The average paths computed from sMD, rMD and all the SCPS iterations are reported in Fig. 2.9, superimposed to the potential level curves. Our results are compared, on the same picture, with the principal curve obtained from FTS method[2]. It is clear that the sMD and the rMD average pathways fail in approximating the principal curve: this is expected, as none of the two possesses the expected theoretical properties to do so. SCPS iterations, instead, approximate the finite temperature string with increasing precision. A well-defined feature of the SCPS average paths becomes evident from Fig. 2.9: they do not manifest the *semi-circular* shape of the exact principal curve, but rather they seem to connect the different states in the potential with *straight lines*. This feature comes from the fact that each point is defined as a time-average. Qualitatively speaking, if a trajectory x is found in R at time t and another trajectory y is found in I at the same time t, the average $1/2(x(t) + y(t))$ will be found on the straight line connecting R to I. For this reason, we expect the average path to approximate the principal curve by connecting the relevant states in the potential with straight lines.

2.5.1 Computing the Committor Function

In Fig. 2.10 (a,b,c) we report the $\tilde{s}_\gamma$ functions in the region $\Omega = [-1.5, 1.5] \times [-0.5, 2.5]$, computed from, respectively, the sMD, rMD and SCPS average pathways.

[2]We thank E. Vanden-Eijnden for making available a MATLAB version of the 2-dimensional FTS method on his website: `https://cims.nyu.edu/~eve2/string.htm`

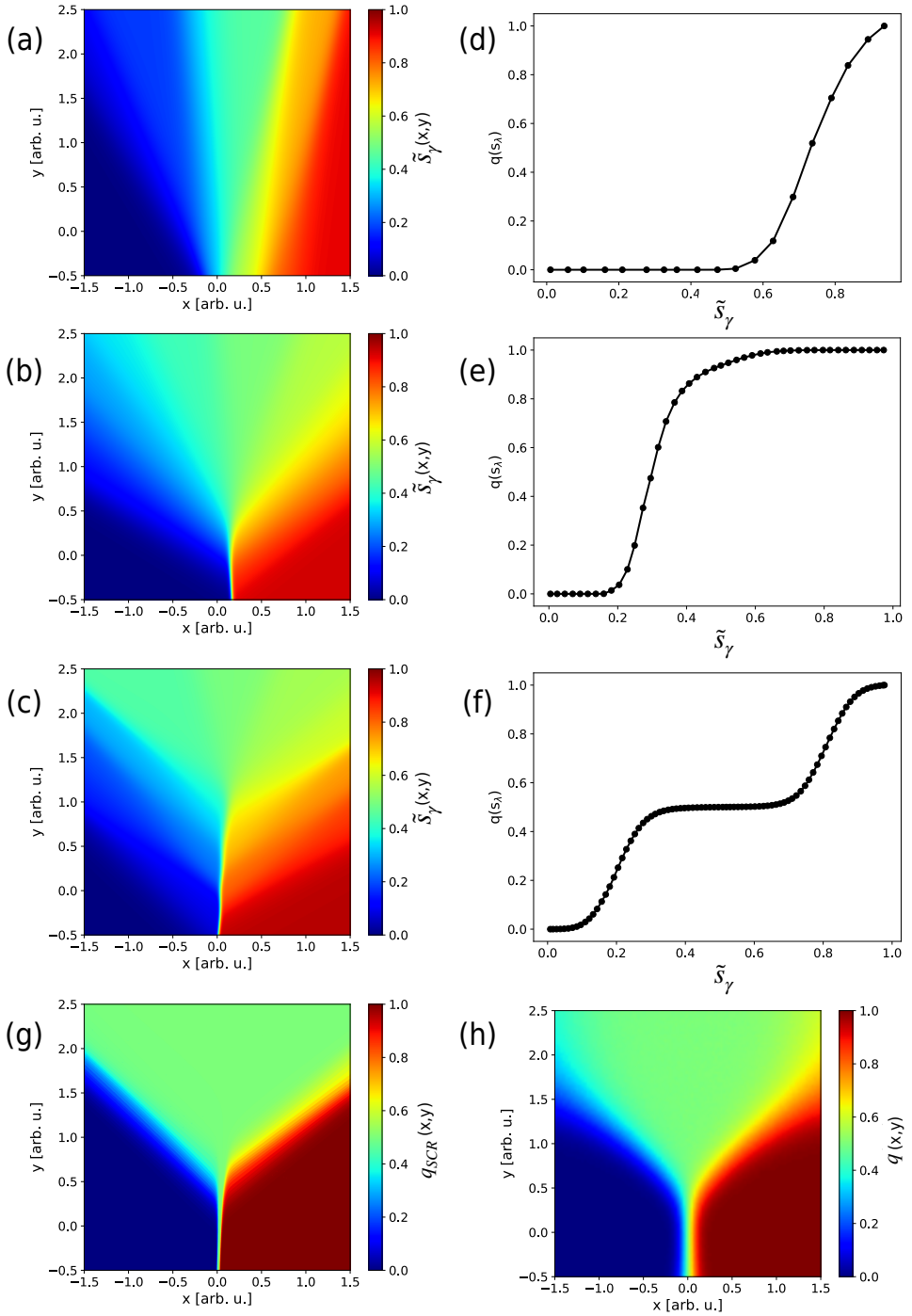

Figure 2.10: (a,b,b) $\tilde{s}_\gamma$ computed from the average paths obtained, respectively, by steered MD, rMD and SCPS; (d,e,f) $q(\tilde{s}_\gamma)$ functions resulting from the minimization of I functional in Eq. (2.94) in the case of, respectively, steered MD, rMD and SCPS; (g) committor function obtained by applying the map $q(\tilde{s}_\gamma)$ in (f) on the $\tilde{s}_\gamma$ in (e); (h) exact committor function.

These results should be compared with the exact committor, reported in Fig. 2.10 (h) and computed using the prescription in appendix C.7. As expected, sMD results represent a misleading representation of the reaction coordinate. In particular, it doesn't indicate the presence of a potential barrier separating R from P, which, instead, is present in the exact committor in the form of a sharp bend in the vicinity of $\mathbf{x} = (0, 0)$. rMD, instead, feels the presence of a barrier, but is strongly reduces its extension. This means that rMD simulations provide a small, but finite, probability to reach P by passing through the barrier. For what concerns SCPS, it provides results similar to exact ones: as expected, iso-q and iso-$\tilde{s}_\gamma$ curves in the transition region are approximately parallel. Moreover, both the two functions show the sharp bend indicating the existence of a potential barrier, with comparable extension along the y direction. Overall, results reported in Fig. (2.10) (c) and (h) imply that $\tilde{s}_\gamma(\mathbf{x})$ represents a reasonable approximation of the committor function. Nevertheless, some differences can still be noticed: the most evident one is that iso-$\tilde{s}_\gamma$ lines are much more spaced with respect to iso-q ones, indicating the existence of a non-trivial mapping connecting the two functions.

The estimation of $q(\mathbf{x})$ from $\tilde{s}_\gamma(\mathbf{x})$ is done by minimizing the functional in Eq. (2.94). The functional was initialized with $\{\tilde{s}^i_\gamma = \tilde{s}_\gamma(\langle \mathbf{x}(i) \rangle)\}$ and was equal to $I[\{\sigma_i\}] \sim 800$. After the minimization, performed using the Sequential Least SQuares Programming (SLSQP) algorithm [187] implemented in SciPy [188], the functional dropped to $I[\{q_i\}] \sim 2 \times 10^{-4}$. This indicates that our estimation of the committor function satisfies the BK equation up to corrections of order $o(10^{-4})$. The resulting map $q(\tilde{s}_\gamma)$ obtained from the minimization procedure is reported in Fig. 2.10 (f). Notably, this function shows a plateau corresponding to the location of the intermediate state, associated to $q = 0.5$. This was expected both as a result of the symmetry of the problem and also because the committor function should be constant within metastable regions. The SCPS estimate of q, obtained by re-labeling the iso-$\tilde{s}_\gamma$ curves in Fig. (2.10) (c) through the function $q(\tilde{s}_\gamma)$ represented Fig. 2.10 (f), is reported in Fig. 2.10 (g). This result is compatible with the exact one in Fig. (2.10) (h), with a reduced spacing between iso-surfaces with respect to Fig. (2.10) (f). Sensible differences between the estimated and the exact committor functions can be seen only in the regions of Ω which are rarely or never sampled by SCPS trajectories.

It is of particular interest to observe the $q(\tilde{s}_\gamma)$ maps obtained by optimizing the $\tilde{s}_\gamma$ values on the sMD and rMD average pathways, reported in Fig. 2.6 (d) and (e). None of the two maps show the presence of a plateau region in correspondence of $q \sim 0.5$. This is expected, because both sMD and rMD simulations are incapable to extensively populate the intermediate state, because of the high biasing force exerted

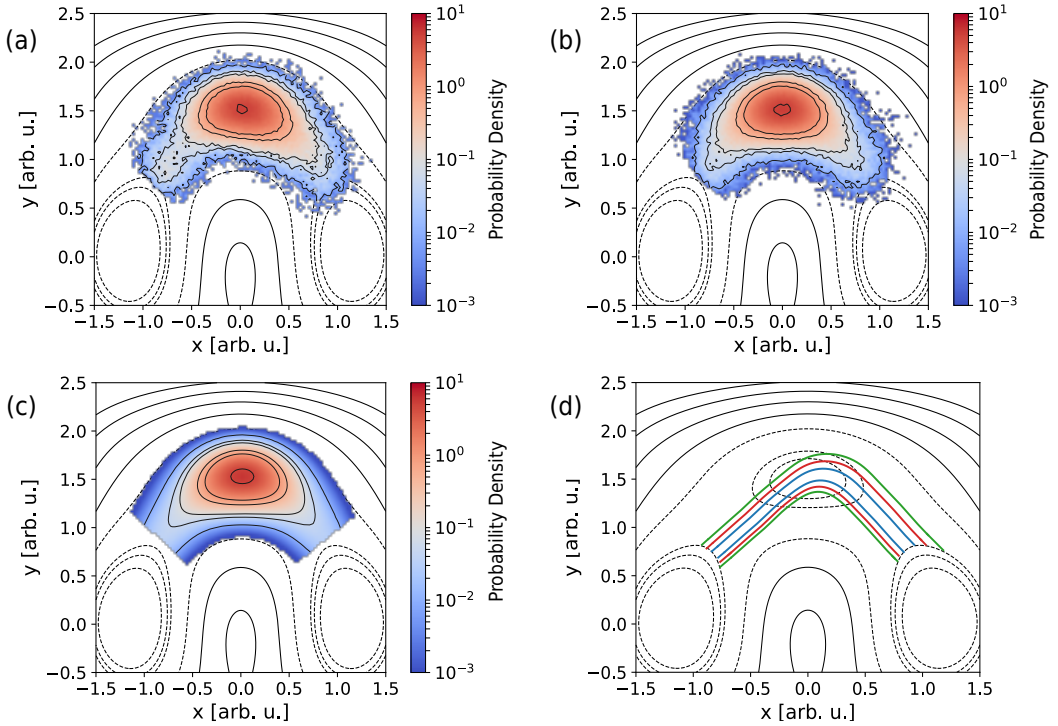

Figure 2.11: Reactive probability densities obtained by means of (a) CLD and (b) irMD compared to (c) the exact result. (d) Transition tubes corresponding to the 30 (blue lines), 60 (red lines) and 90% (green lines) of the current flow through iso-committor surfaces.

along an unreliable reaction coordinate. We can thus conclude the following. Let us suppose that, in a realistic high-dimensional situation, the only way to generate pathways connecting R to P is to resort to a very high biasing force along a coordinate whose quality cannot be clearly estimated. In this case, SCPS is able to sensibly correct for the errors in the original biased dynamics and return an approximation of a principal curve which can be further optimized to provide a reasonable estimation of the committor function.

2.5.2 Reactive Probability Density and Transition Tubes

We previously discussed the fact that SCPS provides a mean-field estimate of the transition probability density $m_T(\mathbf{x})$, which is in reasonable agreement with the exact result. However, we can exploit the knowledge of $q_{\mathrm{SCR}}(\mathbf{x})$ coming from the SCPS simulations to systematically improve the quality of the results. The way to do this is to use the algorithms proposed in sections 2.4.3 and 2.4.4, respectively CLD and irMD. The quality of the estimation of $m_T(\mathbf{x})$ by these algorithms is affected only by possible errors in the knowledge of the committor function.

We ran 1000 independent CLD and irMD simulations, each one $t = 4 \times 10^3 dt/\gamma$ long. In the irMD simulations we employed an elastic constant $k_R = 50$, exactly like

in the rMD and SCPS simulations, and we set $\xi(y) = y$ in Eq. (2.101). In both CLD and irMD trajectories were initialized on the iso-q curve $q(\mathbf{x}) = \varepsilon = 0.01$ and then stopped on the surface $q(\mathbf{x}) = 1 - \varepsilon = 0.99$. Results in Fig. 2.11 (a) and (b) show that the $m_T(\mathbf{x})$ distributions sampled by CLD and irMD are quantitatively comparable with the exact one, despite the presence of small errors in the estimation of the committor function. The latter point can be justified with the fact that the committor is incorrectly estimated only far from the main reaction channel, where SCPS simulations do not spend much time.

Finally, to conclude the description of the reactive process, we computed the transition tubes using the algorithm presented in section 2.1.4. In particular, the integration of Eq. (2.30) was initialized on the $q(\mathbf{x}) = 0.5$ surface and then propagated forward and backward in the artificial time parameter τ. Our result in Fig. 2.11 (d) shows the transition tubes corresponding to the 30, 60 and 90% of the current flow through iso-committor surfaces. The reaction channel is correctly located, but the bending point of the streamlines is slightly shifted to the right: this might be induced by small deviations in the committor that propagate in the integration of the differential equation in Eq. (2.30).

2.6 Chapter Conclusions

The main focus of this chapter was on how to link transition path theory, which is rigorously defined only in the case of an infinite and ergodic trajectory, with the practical need to simulate many short out of equilibrium trajectories using enhanced sampling methods. This was done by exploiting the properties of kinetics in a suitable time range, which we called steady current regime, in which TPT results can be approximated with averages performed over short, non-ergodic trajectories. A second point of great focus of this chapter was the fact that the self-consistent path sampling method, introduced in chapter 1, can be actually thought as an algorithm that approximates the principal curve connecting the reactant to the product states. This fact comes together with the realization of the strong connection between the variable s_λ, defined in section 1.3.1, and the committor function. In particular, the committor function can be obtained from s_λ by minimizing the functional defined in Eq. (2.94) along the approximate principal curve defined by SCPS.

Next, we discussed how it is possible to capitalize on the knowledge of the committor function by defining two schemes to sample the reactive probability density. The first one is a conditional Langevin dynamics that exactly samples the transition path ensemble but suffers for the existence of possible detours from the main reactive pathway. The second one is a variation of the original ratchet-and-pawl MD which

employs the committor function as a reaction coordinate (thus named *ideal* rMD), which uses a strong biasing force along q to hinder possible detours but does not satisfy microscopic reversibility.

The last part of this chapter was dedicated to the illustration and validation of our algorithm on a simple 2d model. Future applications of this work should then necessarily focus on more high-dimensional systems, as the usefulness and power of this method on complex structural transitions still needs to be thoroughly assessed. A possiblilty would be to validate the results of the algorithm on proteins on which TPT has already been applied, e.g. PinWW in Ref. [104]. Indeed, while the efficiency of SCPS on the folding of proteins has already been studied [28], irMD and CLD still needs to be extensively tested. Finally, the theory presented here can be extended in its validity by repeating the calculations without assuming the Langevin dynamics to be overdamped.

Chapter 3

Enhancing Experimental Sensitivity with the Bias Functional Method

Benchamarks of the BF method against both MD and experiments show that, despite its intrinsic limitations, the algorithm provides sensible results and may be used to study a wide class of proteins [26, 27, 144]. The same approach might be also useful in order to support the interpretation of experimental data. Up to now, the simultaneous use of experimental techniques and computer simulation to achieve a deep understanding of the folding problems has proved to be quite challenging [189, 190]. The problems come from both the computational and the experimental sides.

From a computational perspective, protein simulations can achieve extremely high spatiotemporal resolutions. However, as we wildely discussed in previous chapters, the computational workload needed to simulate the folding of a protein is extremely high. It was estimated that molecular dynamics simulations alone, performed on special purpose supercomputers, are effective in modeling the folding on just the 10% of the Protein Data Bank [191, 192]. These results call for further advancements: better force fields, which are constantly being improved [191, 193, 194, 195], better described water-protein interactions [196, 197, 195] and faster ways to sample proteins conformations.

From the experimental side, protein studies have reached a sensible level of reliability, but most of the experimental techniques either lack spatial resolution or can probe with high accuracy only distances between specific residues. Let us discuss some examples. Hydrogen/deuterium (H/D) exchange detected by mass spectroscopy can be used to detect solvent exposed regions in the protein chain [198]

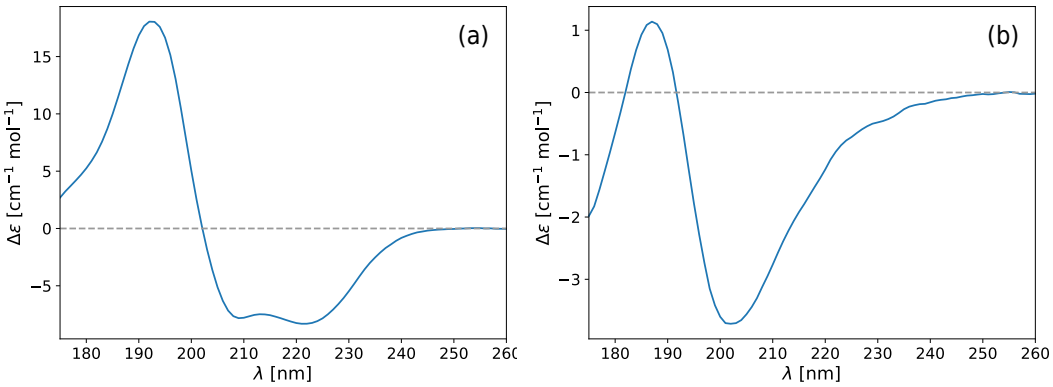

Figure 3.1: Representative CD spectra for (a) a α-helix and (b) a three-stranded β-sheet, computed using Pdb2CD [213].

and small-angle x-ray scattering (SAXS) measures the overall degree of protein compactness [199]. Even though these methods can be applied to kinetically follow the folding reaction, e.g. SAXS can be combined with the stopped-flow technique [200], they cannot be used to assess the quality of atomistic simulations as they lack spatial resolution (surface accessibility and degree of compactness are rather coarse-grained information) and they cannot be applied to study sub-millisecond transitions [201]. Alternative techniques, that focus on single molecules, have been developed that can reach much higher temporal and spatial resolutions. Among the wide variety of possible methodologies [202], a few notable examples are provided by single molecule Förster Resonance Energy Transfer (FRET) experiments [203] and atomic force microscopy (AFM) measurements [204]. The former can measure variations in the distances between two suitably located cromophores with subnanometric precision and within millisecond timescale [205, 206], while the latter measures the distance between two residues on which an external stress is applied, again with subnanometric resolution [207, 208]. Single molecule techniques represented a fundamental step forward in experimental biophysics [209, 210], but unfortunately they require to actively alter the chemical composition of the protein chain, by mutating specific residues [211] or attaching fluorescent probes [212]. The natural question arises whether manipulations of the chemical nature of the protein sensibly alter their dynamics and biological function.

Differently from the experimental strategies we discussed up to now, spectroscopic methods do not require any chemical modifications of the protein, and circular dichroism (CD) measurements can be easily performed to monitor the type and amount of secondary structures of an equilibrium ensemble of proteins in solution [33, 34]. CD spectra come from the differential absorbance of left- and right-circularly polarized light. In the far-UV region (190-230 nm), CD measurements are dominated by the

signal coming from secondary structures, and empirical algorithms can be used to deconvolve the different α and β contributions and estimate their average content [214, 215, 216]. Time-resolved CD experiments are also possible, and are used together with the stopped-flow technique to monitor the kinetics of secondary structure formation [214, 217, 218, 219, 220]. A typical CD spectrum of an α-helix shows a strong positive peak around $\lambda = 190$ nm and a negative doublet in $\lambda = 208$ and 220 nm (see Fig. 3.1 (a)), while β-sheets signals are less well defined with usually a negative peak around $\lambda \sim 180$ nm, a positive band in $\lambda \sim 195$ nm and a negative peak around $\lambda \sim 210 - 220$ nm [221] (see also Fig. 3.1 (b)). However, the insight on the folding process coming from this analysis is limited, in the sense that it provides structural information only at the level of secondary structures that, in most of the cases, are already formed at the initial stages of the conformational reaction.

There is a further spectral region which can be investigated using the CD technique, which is the so-called near UV (240-300 nm) region. As a matter of fact, CD in this region is sensitive to the $\pi - \pi^*$ excitations occurring in the side-chains of aromatic residues (Tyr, Phe, Trp; see appendix D for further details) [222, 223]. It should be also noted that fingerprints of aromatic couplings start to emerge already for wavelengths $\lambda < 230$ nm, but in those regions the signal is too dominated by amide electronic transitions in the protein backbone [224]. Near-UV CD signal represents a highly spatially resolved probe to detect local rearrangements in folding, which however cannot unequivocally predict three-dimensional rearragements of the protein without relying on some prior models [220].

In this chapter we are going to to shine new light on experimental CD results through a novel hybrid and *ab-initio* algorithm based on the ability of the BF method to robustly predict folding pathways and on the possibility to compute CD spectra by means of computational chemistry calculations. This algorithm is *ab-initio* in the sense that CD spectra are computed just be employing all-atom calculations. Briefly, our proposed pipeline is the following: (i) use the BF method to compute the folding pathway of a protein of interest; (ii) identify, using suitable reaction coordinates, the metastable states along the folding pathway; (iii) harvest representative conformations from the metastable basins; (iv) perform quantum chemical calculations based on an exciton model [35, 36, 37] to compute the near-UV CD spectrum from the harvested conformations; (v) compare the simulated CD spectra with experimental results. When near-UV CD spectra are available for a protein of interest, this novel algorithm will be able to bridge the resolution gap between simulations and experiments, providing atomic-level structures that can help the interpretation of experimental data.

We will test our approach on three systems: the canine milk lysozyme, a globular

protein composed by 129 residues that folds in the time scale of seconds [32]; a synthetic version of the same protein, where the disulphide bonds, present in the crystal structure, have been removed; IM7 protein, an α protein composed by 86 residues that folds in the time scale of milliseconds [225]. In particular, the first system represents a *best-case scenario*, where we can access CD data of a protein with many tryptophans, no mutation or chemically altered environment is needed in order to stabilize its intermediate conformations and it is particularly easy to find reaction coordinates to determine the presence of metastable basins. The second system, namely the modified version of the lyszoyme where all the disulphide bonds have been reduced, is used to assess the sensitivity of our algorithm to chemical modifications in the protein chain. Finally, IM7 protein represents a *worst-case scenario* system, where the CD signal comes from the interaction of a single Trp residue with all the other aromatic ones, mutations are needed in order to stabilize its intermediate state and folding progresses along non-trivial reaction coordinates. Remarkably, our algorithm provides quantitative results for all the above mentioned cases, showing that near-UV CD signals coupled with enhanced sampling techniques and quantum chemical calculations are indeed a powerful tool to detect subtle changes in the tertiary structure of a protein and provide a full, atomic-level picture of its folding pathway.

The chapter is structured as follows. In section 3.1 we will briefly introduce the concept of CD spectrum in a theoretical fashion and describe how it is possible to compute it using the so-called matrix method. In section 3.2 we will first discuss the experimental knowledge about the canine milk lysozyme, then apply our pipeline and discuss the corresponding results. Section 3.3 is devoted to the sensitivity assessment of our pipeline, which is applied to the canine milk lysozyme without disulphide bonds. Section 3.4 is instead dedicated to the application of our algorithm to IM7 protein and the corresponding results. Finally, in section 3.5 we will draw our conclusions about the methodology and discuss some possible further improvements.

3.1 Calculation of CD Spectra

When we let circularly polarized light impinge on a solution of chiral objects, like proteins, it will show different refractive indices for left- and right-circularly polarized light [224]. As a consequence, the extinction coefficient ε[1] will be different for left- and right-circularly polarized light, $\varepsilon_L \neq \varepsilon_R$. This effect is called *circular dichroism*,

[1] a coefficient which measures how strongly a chemical species attenuates light at a given wavelength

and the CD spectrum is provided by the differential absorbance

$$\Delta\varepsilon(\lambda) = \varepsilon_L(\lambda) - \varepsilon_R(\lambda) \tag{3.1}$$

as a function of the probe wavelength λ [226, 227, 228]. The integral

$$R^{ij} = \frac{hc}{32\pi^2 N_A} \int_{\Delta\lambda_{ij}} d\lambda \, \frac{\Delta\varepsilon(\lambda)}{\lambda} \tag{3.2}$$

where h is the Planck constant, c the speed of light and N_A the Avogadro number, computed over a range of wavelengths $\Delta\lambda_{ij}$ associated with the electronic transition $i \to j$ between states i and j, is known as the *rotational strenght* of that transition [224, 228].

From a theoretical perspective, the rotational strength R^{0k} of the transition $0 \to k$ from the ground-state to a generic excited state k is provided by the Rosenfeld equation [229]:

$$R^{0k} = \mathrm{Im}\left[\langle\psi^0|\boldsymbol{\mu}|\psi^k\rangle\langle\psi^k|\mathbf{m}|\psi^0\rangle\right] \tag{3.3}$$

where $|\psi^0\rangle$ and $|\psi^k\rangle$ represent, respecitvely, the ground- and the k-th excited states, $\boldsymbol{\mu}$ is the electric transition dipole moment

$$\boldsymbol{\mu} = q\mathbf{r} \tag{3.4}$$

where q is the particle charge, $\mathbf{r}$ is its position and $\mathbf{m}$ is the magnetic transition dipole moment

$$\mathbf{m} = -i\mu_B \mathbf{r} \times \boldsymbol{\nabla} \tag{3.5}$$

with μ_B is the Bohr magneton and i is the imaginary unit. A qualitative interpretation of the CD spectrum comes from the following considerations. While the $\boldsymbol{\mu}$ operator describes a linear charge displacement, the $\mathbf{m}$ operator is an angular momentum and thus characterizes a circulation of charge. The combination of the two provides a helical displacement of charge, which leads to different interactions with left- and right-circularly polarized light [224].

From Eqs. (3.2) and (3.3) it is clear that it is possible to use the rotational strength to compute the CD signal of a protein. However, the main problem comes from the fact that $|\psi^0\rangle$ and $|\psi^k\rangle$ cannot be computed ab-initio for molecules as big as proteins, and one needs to resort to some approximation. Several methods can be employed to approximate the calculation of these quantities [230], but one of the the most successful ones is provided by the *matrix method* [231, 232, 233].

3.1.1 The Matrix Method

Let us assume we can split the protein into M independent regions sensible to the light probe, which we will call *chromophores*. The i-th cromophore, in an electronic

state s, will be described by a state $|\phi_{is}\rangle$. The k-th excited state of the protein can be thus approximated as a linear combination of multi-cromophore states $|ia\rangle$ describing a situation in which cromophore i is in an excited state $s = a$ and all the other ones are in the ground state $s = 0$:

$$|ia\rangle = |\phi_{10}\rangle|\phi_{20}\rangle \dots |\phi_{ia}\rangle \dots |\phi_{j0}\rangle \dots |\phi_{M0}\rangle \tag{3.6}$$

Supposing each cromophore can access n excited states, the protein k-th excited state can be written as

$$|\psi^k\rangle = \sum_{i=0}^{M} \sum_{a=0}^{n} c_{ia}^k |ia\rangle \tag{3.7}$$

The first sum in Eq. (3.7) is over the M cromophores, while the second one is over the N excited states accessible to each cromophore. The c_{ia}^k coefficients are unknown and need to be determined. The protein excited states and their corresponding energies are solution of the Schrödinger equation:

$$\hat{\mathcal{H}}|\psi^k\rangle = E^k|\psi^k\rangle \tag{3.8}$$

where $\hat{\mathcal{H}}$ is the Hamiltonian governing the system's dynamics. This Hamiltonian can be chosen to be the Frenkel exciton Hamiltonian [234]:

$$\hat{\mathcal{H}}_{\text{ex}} = \sum_{i=0}^{M} \sum_{a=0}^{n} \mathcal{E}_i^a |ia\rangle\langle ia| + \sum_{i,j=0}^{M} \sum_{a,b=0}^{n} V_{ij}^{ab} |ia\rangle\langle jb| \tag{3.9}$$

where $\mathcal{E}_i^a$ are the excitation energies of the i-th non-interacting chromophore in the a-th excited state, called *site energies*, and V_{ij}^{ab} are the electronic couplings between two transitions of different chromophores.

Assuming the Hamiltonian in Eq. (3.9), Eq. (3.8) can be solved by diagonalization in order to obtain both the energy levels of exciton states E^k and the expansion coefficients c_{ia}^k. The rotatory strength of the $0 \to k$ transition can be rewritten as a function of these ingredients [33, 235]:

$$R^{0k} = -\frac{\pi E^k}{2hc^2} \sum_{i,j,a,b} c_{ia}^k c_{jb}^k \mathbf{R}_{ij} \cdot (\boldsymbol{\mu}_{i0a} \times \boldsymbol{\mu}_{j0b}) \tag{3.10}$$

where $\mathbf{R}_{ij}$ measures the distance between the centers of the i-th and the j-th cromophores. For all the technical details about, for example, the spectral broadening of the transition energies or the estimation of the parameters in the Hamiltonian, we refer the reader to the specialized literature [222, 228, 236, 237].

3.2 Application to the Canine Milk Lysozyme

The first system of interest is the canine milk lysozyme (see Fig. 3.2), a 129 residue globular protein which has a folding time in the range of seconds [32]. This protein

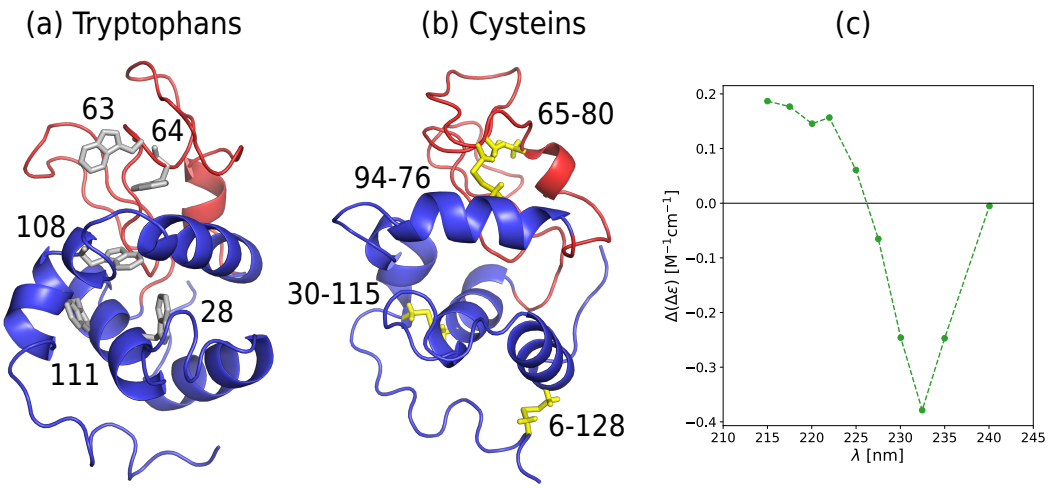

Figure 3.2: Crystal structure (PDB entry: 1EL1) of the canine milk lysozyme. Water molecules and ions have been removed from the original structure. The red-colored region represents the β-domain, while the blue-colored one represents the α-domain of the protein. (a) The five protein tryptophans are represented as grey sticks; (b) the four disulphide bonds in the protein crystal are represented as yellow sticks; (c) the experimental kinetic CD difference spectrum between **I-Second** and **Native** states, adapted from [32].

possesses several attractive features that make it a perfect study case: (i) its folding timescale is inaccessible to standard MD techniques; (ii) its topology is relatively simple; (iii) different experiments [238, 32] confirm the existence of a folding intermediate; (iv) its intermediate states have been characterized using the stopped-flow CD technique [32].

The authors of Ref. [32] performed a kinetic CD measurement of the Escherichia coli recombinant[2] canine milk lysozyme, inducing a refolding reaction by a denaturant concentration jump. The resultant kinetics was then followed by rapid CD measurements, which show the presence of a *burst* (**I-Burst**) intermediate, populated within the dead-time of the stopped flow apparatus (25 ms) and that can be interpreted as the unfolded state of the protein in absence of denaturant [239], and a second (**I-Second**) intermediate state, populated at $\sim$ 22 s^{-1} rate. The system finally reaches the **Native** state at a rate of $\sim$ 0.5 s^{-1}. Further interpretation of the CD data suggests that the **I-Burst** state has already a significant fraction of secondary structures formed and that in the transition between the **I-Burst** and **I-Second** the content of secondary structures is further increased. Differently, the main differences between **I-Second** and the **Native** state are due to changes in the tertiary structures. In particular, the authors suggest that the change in CD spec-

[2]A host cell in which a gene is copied that encodes a specific protein expresses a so-called *recombinant protein*

trum occurring during the latter transition is caused by the exciton coupling of the tryptophan residues in the protein (see Fig. 3.2 (a)). This was determined by looking at the so-called *kinetic difference CD spectrum*, which is computed as the difference between the **Native** and the **I-Second** CD signals. As shown by the green dotted curve in Fig. (3.2) (c), the experimental kinetic CD spectrum is a bisignate signal with a positive peak around ~ 210 nm and a negative one in ~ 230 nm. The zero crossing occurs at 226 nm. It is experimentally well known that both α-helices and β-sheets provide a negative contribution to the CD spectrum in the region $210-240$ nm [221]. For this reason, secondary structures cannot be responsible for a bisignate couplet in the kinetic CD spectrum with a zero-crossing at 226 nm. This spectral feature of the kinetic difference CD signal is necessarily due to exciton coupling of tryptophan residues [32]. In the following section we are going to thoroughly assess these experimental results by adding atomic-level information on the lysozyme folding pathway.

As a final remark, it has to be noted that the canine milk lysozyme has 4 disulphide bonds[3] (see Fig. 3.2 (b)) and that refolding experiments has been conducted in an oxidative environment [32], so that convalent bonds between cysteins are formed correctly.

3.2.1 Simulation Setup

rMD simulations of the folding of the canine milk lysozyme were set up as follows:

1. a set of 10 independent unfolded configurations was generated by thermally unfolding the energy-minimized crystal structure of the protein. In particular, we performed 5×10^5 MD time steps with a time step of $dt = 1$ fs at a temperature of 800 K;

2. starting from each unfolded configuration, we ran 20 rMD simulations, employing a timescale of $dt = 1$ fs and performing 5×10^5 rMD steps at $T = 350$ K. The harmonic constant of the bias force was set to $k_{\mathrm{rMD}} = 2.5 \times 10^{-4}$ kJ/mol, making sure that the bias force was at least two orders of magnitude smaller than the total physical force at any time.

3. out of each of the 20 rMD simulations, the LB trajectory is selecting by implementing the variational principle in Eq. (1.18).

All simulations (MD and rMD) were ran in GROMACS 4.6.5 [41] using the Amber ff99SB-ILDN force field [142], the Bussi thermostat [143] and the Generalized Born

[3] covalent bonds formed between cysteins. In an MD simulations, they play the role of further topological constraint, just as polypeptide bonds.

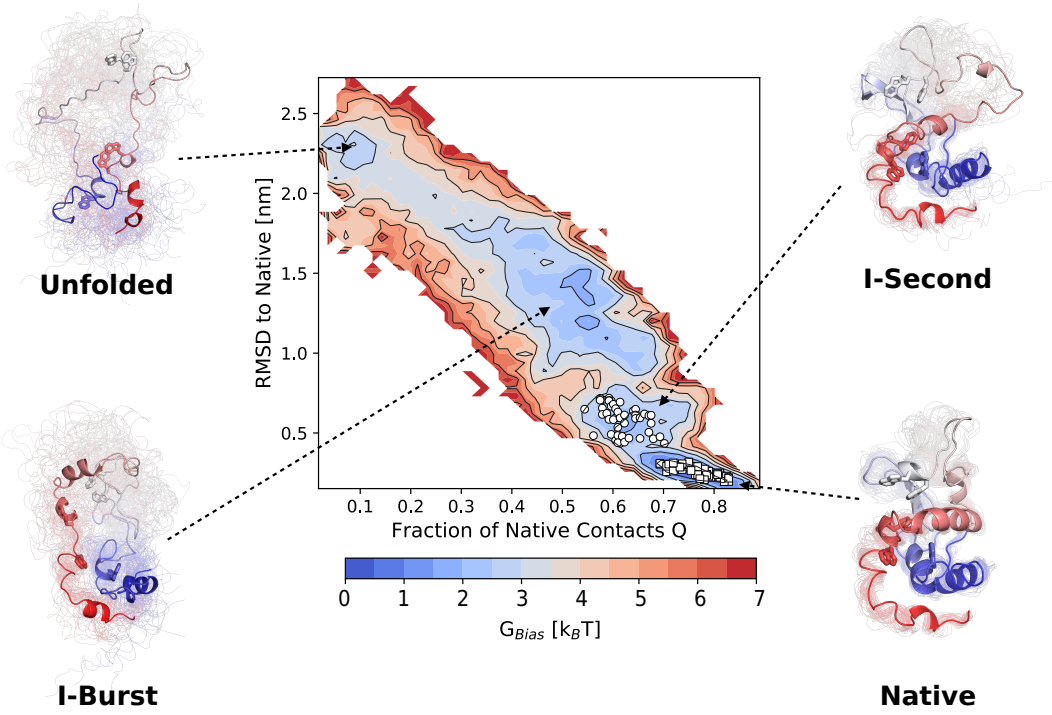

Figure 3.3: Kinetic free energy profile obtained by projecting the whole ensemble of canine milk lysozyme rMD trajectories onto the RMSD to native and the fraction of native contacts Q. Low free energy regions reveal the existence of metastable states, respectively identified as **Unfolded**, **I-Burst**, **I-Second** and **Native**. Protein structures represent conformations that have been harvested from the minima, while white markers, when present, represent their corresponding location on the plot.

implicit description of the solvent model [240, 241]. The collective variable and the corresponding bias force needed to perform BF calculations were implemented in Plumed 2.0.2. [51].

For what concerns the setup of the quantum chemical calculations, we advise the reader to refer to the specialized literature [228, 236, 242, 243, 244, 245, 246, 247, 248] and to the original paper [24].

3.2.2 The folding pathway

In Fig. 3.3 we show the kinetic free energy landscape $G_{\text{Bias}}(Q, R)$ obtained by projecting all the 200 rMD simulations onto the plane spanned by the fraction of native contacts Q and the RMSD to the native crystal structure R.

Inspection of $G_{\text{Bias}}(Q, R)$ reveals the existence of four long-lived states along the folding pathway. To obtain an atomic-level description of these states we should extract some configurations from the basins. In order to harvest a statistically significant number of configurations and do it in a least biased way, we proceeded as

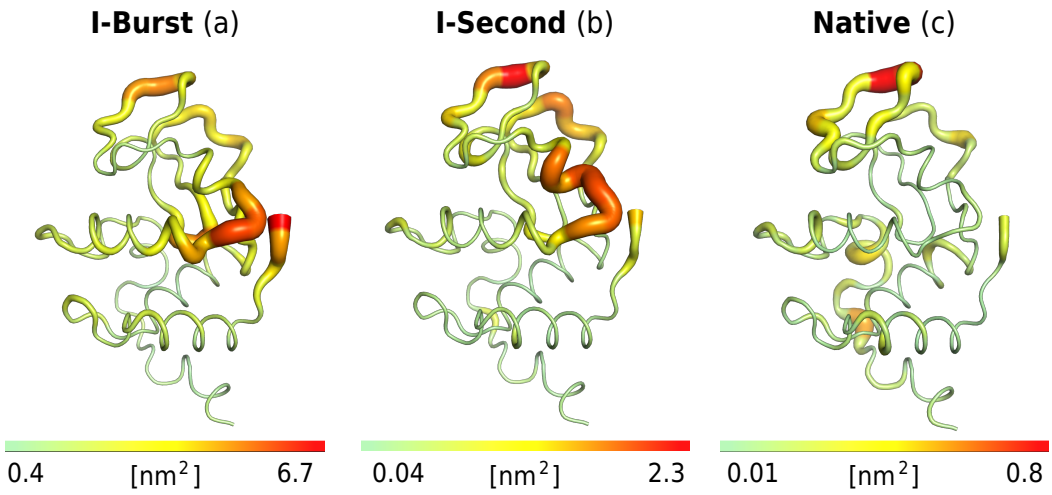

Figure 3.4: Putty representation of the protein crystal structure, colored using the B-factors. B-factors have been computed as $b(a) = 8\pi^2 \mathrm{MSF}(a)$ [249], where $\mathrm{MSF}(a)$ is the mean square fluctuation of residue a in the (a) **I-Burst**, (b) **I-Second** and (c) **Native** basin.

follows. In the following, we will refer to this as the *harvesting prescription*. A structure $\overline{\mathbf{x}}$ is harvested from a basin if it satisfies the following two criteria:

1. $G_{\mathrm{Bias}}(Q(\overline{\mathbf{x}}), R(\overline{\mathbf{x}})) < g$, where g is a given threshold;

2. $\overline{\mathbf{x}}$ belongs to a LBT.

The first criterion is used to select conformations which are as close as possible to the local free energy minimum, while the second criterion helps reducing the systematic error introduced by the bias force.

Harvested configurations of the lysozyme, obtained by using $g = 2k_{\mathrm{B}}T$, are shown in Fig. 3.3. We refer to the top-left well of the picture as the **Unfolded** basin, which is populated by highly denatured states. Configurations in this region are never sampled by spontaneous folding-unfolding events, but only by high-temperature MD simulations. In this sense, they provide an approximation of the chemically denatured states reached during the experiments.

The first metastable state, found approximately in the middle of the $P(Q, R)$ projection, is highly heterogeneous and occupies a wide region both in RMSD and in fraction of native contacts. We identify it with the experimental **I-Burst** state. Looking at the root mean square fluctuations (RMSF) of the protein residues in this basin, Fig. 3.4 (a), it is evident that the chain is particularly flexible, especially around the beta region. This is can be interpreted by the fact that alpha-helices are already partially formed in this stage, and the residual mobility comes from

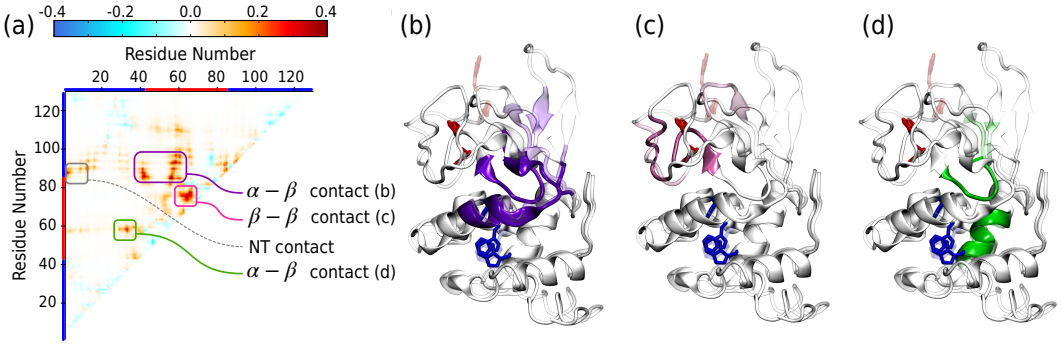

Figure 3.5: Adapted with permission from [24]. Copyright 2018 American Chemical Society. (a) Plot of Eq. (3.11), where the key tertiary contacts formed during the **I-Second** $\rightarrow$ **Native** have been highlighted; (b,c,d) superimposed **I-Second** (transparent) and crystal (white) structures, showing as colored regions the locations involved in the formation of the key tertiary contacts. Colors from (a) have been preserved in (b,c,d) for sake of clarity.

the absence of tertiary packing, which is also confirmed by visual inspection of the harvested conformations in Fig. 3.3.

The formation of the secondary and tertiary contacts in the alpha domain is completed in a further metastable state, which can be thought to approximate the experimental **I-Second** configuration. Differently, the beta domain preserves a high mobility, as can be seen from the selected configurations in Fig. 3.3 and the B-factors represented in Fig. 3.4 (c).

The last basin corresponds, with no surprise, to the protein's **Native** state. To understand what key contacts are formed in the **I-Second** $\rightarrow$ **Native** transition we can make use of the harvested configurations. In particular, we computed the difference

$$C_{ij}^{\mathbf{I\text{-}S}\rightarrow\mathbf{N}} = \langle C^{\mathbf{I\text{-}S}} \rangle_{ij} - \langle C^{\mathbf{N}} \rangle_{ij} \tag{3.11}$$

between the average C_α contact map $\langle C^{\mathbf{I\text{-}S}} \rangle_{ij}$ of the structures harvested in the **I-Second** basin and the one obtained from selected native configurations, $\langle C^{\mathbf{N}} \rangle_{ij}$. Matrix in Eq. (3.11) is reported in Fig. 3.5 and shows that the main structural difference between **I-Second** and **Native** states is the tertiary packing between α and β domains. This conclusion is consistent with the time-resolved CD spectrum analysis reported in Ref. [32].

3.2.3 The CD signal

Now that the existence of the metastable states is assessed and the main structural differences have been characterized, we can move to the calculation of the near-UV CD spectra. We are mainly interested in the CD spectra of the **Native** and the **I-**

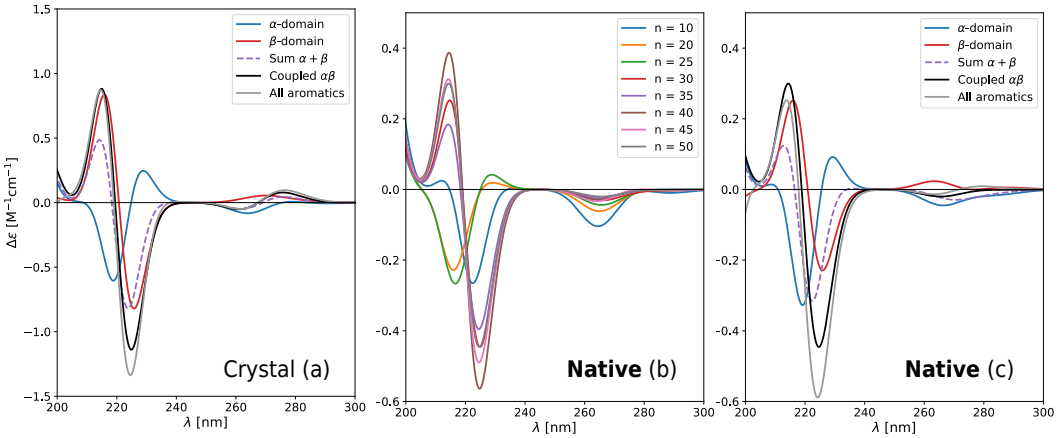

Figure 3.6: Simulated CD spectra: (a) crystal structure; (b) convergence of the **Native** signal upon inclusion of an increasing number of configurations in the calculation; (c) average signal computed from all the harvested configurations in the **Native** basin. In (a) and (c), black solid lines correspond to the CD spectrum computed considering all and only the protein tryptophans as cromophoric units; blue solid lines show the spectrum coming from tryptophans in the the α-domain alone (Trp 28-108-111) and the red solid lines instead refer to the signal coming from tryptophans the β-domain (Trp 63-64); purple dotted lines represent the sum of the α and β uncoupled contributions; grey solid lines show the signal obtained considering all the aromatic contributions.

Second states in order to use them to compute the kinetic CD difference spectrum, which can be directly compared with the experimental results in Ref. [32]. In order to do that, we can apply the matrix method, described in section 3.1.1, to the harvested configurations representing the **I-Second** and **Native** states. The cromophoric units are provided here by the aromatic residues, phenalanine (Phe), tyrosine (Tyr) and tryptophans (Trp), but, as we will see and coherently with experimental observations [32], tryptophans alone provide the dominant contribution to the signal.

In Fig. 3.6 (a) we show the CD signal computed from the protein crystal structure by using only Trps as cromophoric units. The spectrum shows two main features: a weak positive band around $\sim$ 290 nm, which is typical of lysozymes [32, 250, 251, 252], and a strong bisignate couplet at shorter wavelengths. This last couplet in experiments is hidden by an even stronger signal coming from secondary structures.

To understand the difference sources contributing to the shape of the CD signal, we computed the decoupled spectra coming from the α and β regions of the protein separately, shown respecitevly as red and blue curves in Fig. 3.6 (a). The difference between the two is sensible, as they show opposite signs and are shifted relatively to each other. As it turns out, the sum of the two uncoupled spectra significantly differs from the fully coupled one (respectively shown as black and dotted lines in

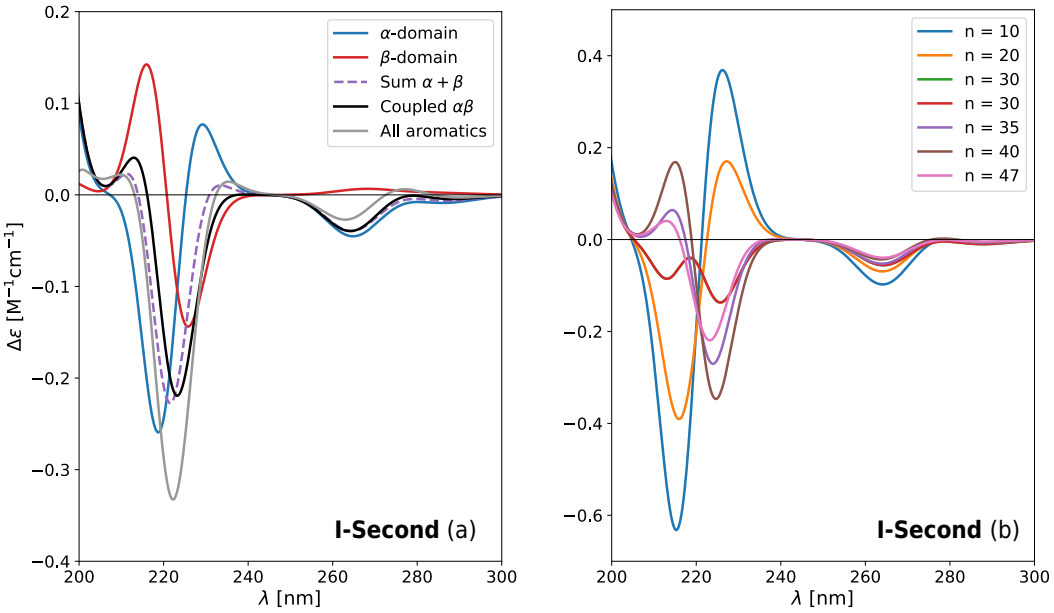

Figure 3.7: Simulated CD spectra: (a) average signal computed from all the harvested configurations in the **I-Second** basin. Black solid lines correspond to the CD spectrum computed considering all and only the protein tryptophans as cromophoric units; blue solid lines show the spectrum coming from tryptophans in the the α-domain alone (Trp 28-108-111) and the red solid lines instead refer to the signal coming from tryptophans the β-domain (Trp 63-64); purple dotted lines represent the sum of the α and β uncoupled contributions; grey solid lines show the signal obtained considering all the aromatic contributions; (b) convergence of the **I-Second** signal upon inclusion of an increasing number of configurations in the calculation.

Fig. 3.6 (a)), underlining the importance of the α-β domain coupling in order to determine the CD spectrum and the sensitivity of the CD signal itself to the specific arrangement of aromatic residues. Indeed, the introduction of the coupling accounts for half of the signal intensity in the lysozyme band around ~ 290 nm.

The CD spectrum computed from the crystal structure alone might not be statistically significant. Differently from single molecule experiments such a FRET, CD signals represent ensemble averages over an extremely high number of molecules. Thus, explicitly introducing structural fluctuations in our calculation is important to better simulate experimental conditions. In order to do so, we can average the signal computed from the harvested **Native** configurations. The number of protein molecules in a solution used for CD experiments is order of magnitudes bigger than the number of configurations we harvested from the native basin. Therefore, it is important to see if our simulated CD spectrum converges by systematically incrementing the number of structures used to compute the average. This indeed happens, as reported in Fig. 3.6 (e).

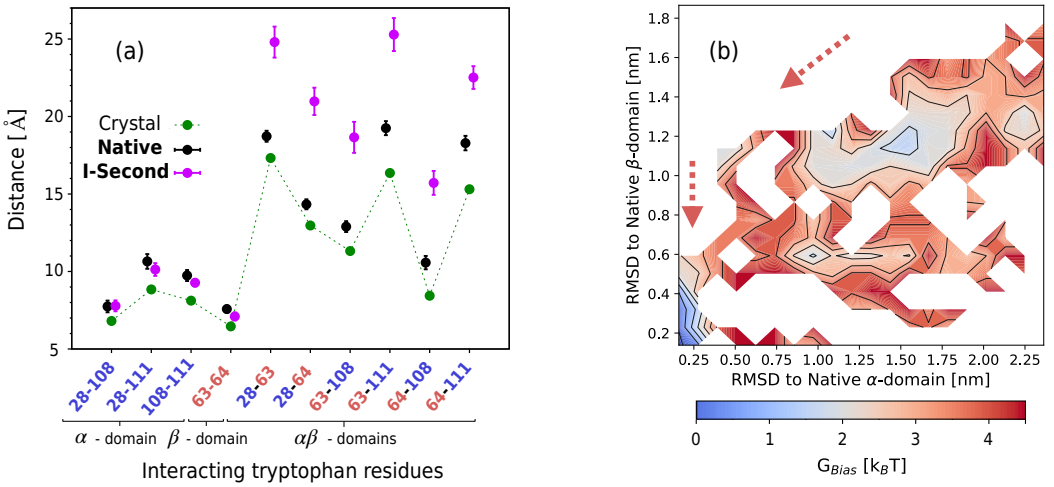

Figure 3.8: (a) Adapted with permission from [24]. Copyright 2018 American Chemical Society. Average pairwise distances between Trp couples in the **I-Second** state (purple dots), in the **Native** state (black dots) and in the crystal (green dots). Colors in the x-label represent the different regions in the protein where Trps are found: blue labels refer to residues in the α-domain, while red labels refer instead to amino acids in the β-domain. Error bars indicate the standard error on the average value; (b) Kinetic free energy lanscape of the canine milk lysozyme projected on the RMSD to native of the α and β domains. The red arrow schematically shows the main folding pathway in this plane.

Given the B-factors reported in Fig. 3.4 (c), one would expect to notice important differences between the Crystal and the **Native** signals mainly in the β domain, which is more subject to structural fluctuations. The results of the quantum chemical calculations reported in Fig. 3.6 (b) confirm our anticipations. Indeed, the α signal remains almost unaltered with respect to the Crystal one, while the β one is suppressed by more than a factor 2 as a result of structural disorder affecting coupligns and orientation Anyway, all the **Native** signals preserve the same spectral feature observed in the Crystal spectrum, besides of the lysozyme band around ~ 290 nm, which becomes extremely weak, as it is suppressed by the negative contributions around ~ 260 nm.

If we explicitly take into account all the aromatic residues in the Crystal and in the **Native** CD spectra, we obtain the results reported as a grey line in Fig. 3.6 (c). It is evident that tryptophans provide the main contribution to the signal, and introducing further couplings with other aromatic residues just slightly increases the overall signal intensity.

In Fig. 3.7 (a) we report the CD spectrum obtained by averaging the signals coming from the harvested **I-Second** structures. Also in this case we took care of assessing the convergence of the spectrum for an increasing number of configurations,

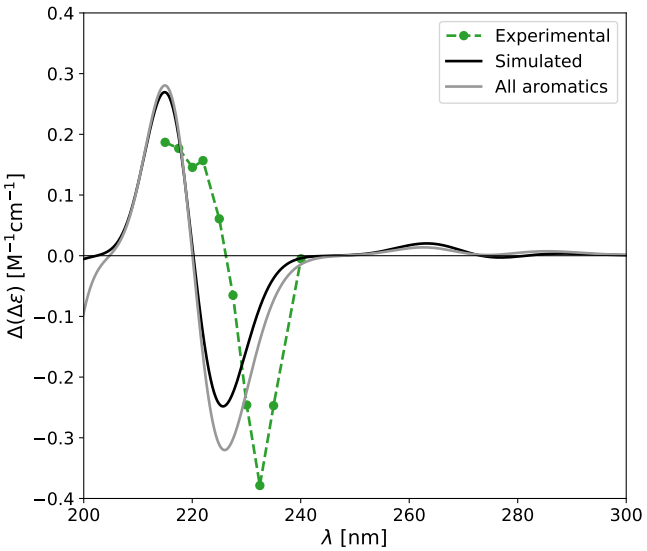

Figure 3.9: Kinetic difference CD spectrum, computed as the difference between the **Native** and **I-Second** CD signals compared against the experimental result (green dashed line) from Ref. [32]. The black solid line represents the signal computed using only tryptophans as cromophoric units, while the grey solid curve is obtained by the contribution of all the aromatic residues.

as reported in Fig. 3.7 (b). The part of the spectrum related to the α domain is approximately unaltered with respect to the **Native** one, while the signal coming from the β domain is broadened and reduced in intensity. As a consequence, the overall signal is sensibly different from the **Native** one throughout the whole range of considered wavelengths. Fig. 3.7 (a) also provides an interesting structural information about the **I-Second** state: the sum of the spectra of the two separate domains is basically indistinguishable from the fully coupled one, meaning that in the **I-Second** state the domains are decoupled. This fact is in contrast with what happens in the **Native** state, but it is coherent with our previous observations about the conformation of the metastable state. A further interpretation of this signal comes from the distribution of the Trp residues pairwise distances. As reported in Fig. (3.8), interdomain distances are approximately identical for both **Native** and **I-Second** state, while intradomain distances are much higher in the **I-Second** case.

The overall picture emerging from our analyses of the BF folding trajectories is the following: the transition **I-Second** $\rightarrow$ **Native** starts once the α and β domains are separately formed (secondary structures and tertiary intradomain contacts are almost fully completed) and occurrs by packing the two foldons through the formation of specific tertiary contatcs. This is also confirmed by looking at the results in Fig. 3.8 (b), where we show the kinetic free energy profile obtained by projecting

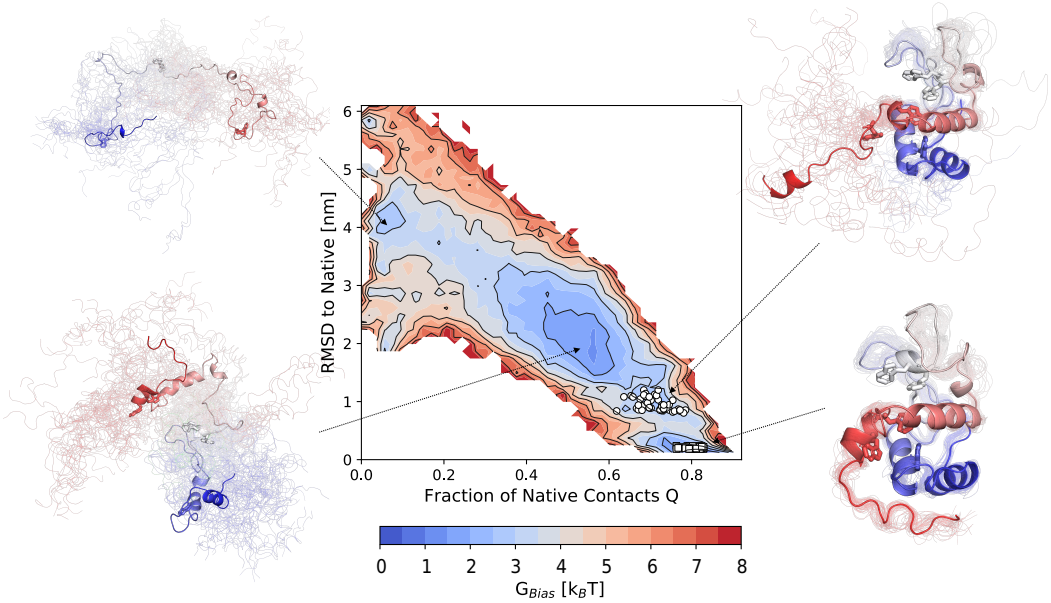

Figure 3.10: Kinetic free energy profile $G_{\mathrm{Bias}}(\mathrm{RMSD}_\alpha, \mathrm{RMSD}_\beta)$, obtained by projecting the whole ensemble of rMD trajectories of canine milk lysozyme without disulphide bonds onto the RMSD to native and the fraction of native contacts Q. Low free energy regions reveal the existence of metastable states. Protein structures represent conformations that have been harvested from the minima, while white dots, when present, represent their corresponding location on the plot.

all the frames of the rMD simulations onto the plane spanned by the RMSD to native of the α and β domains, respectively. $G_{\mathrm{Bias}}(\mathrm{RMSD}_\alpha, \mathrm{RMSD}_\beta)$ shows that the formation of the two foldons proceeds in a rather cooperative way, and the **Native** state is formed once the β domain packs into the fully native α one.

Finally, to provide a direct comparison with experiments, we computed the kinetic CD difference spectrum of the protein. At the beginning of this chapter we argued that secondary structures don't play a role in the definition of the signal, because both the β and the α spectra are negative in the region of interest.

For this reason, the CD difference spectrum can be computed by simply taking the difference between the **Native** and the **I-Second** signals, respectively signal in Fig. (3.6) (c) and Fig. (3.7) (a). The results are reported in Fig. (3.9) (a). In spite of a small 5 nm blue-shift, the computed signal quantitatively captures the shape of the experimental one. The remarkable similarity between the experimental and simulated results shows that the structures visited by the BF method represent a good approximation of the experimental **I-Second** state. As reported in Fig. 3.9, solid grey line, considering all the aromatic residues introduces a small intensity correction in the region around ~ 230 nm.

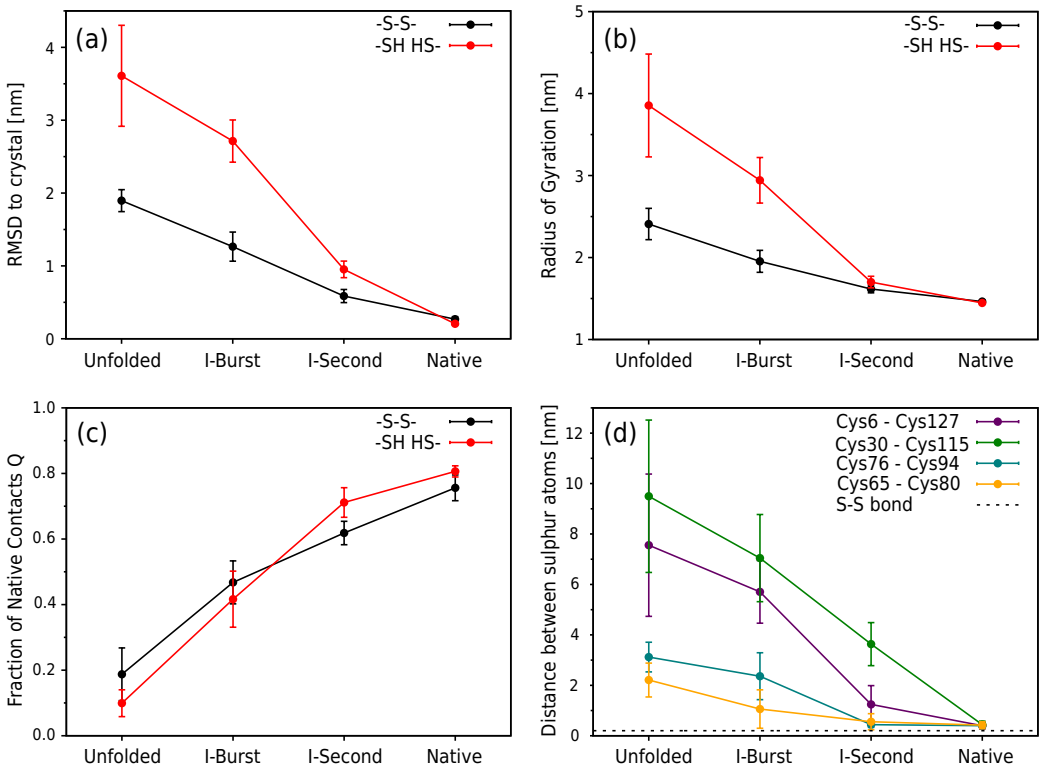

Figure 3.11: Adapted with permission from [24]. Copyright 2018 American Chemical Society. For all the harvested configurations, in presence and absence of disulphide bonds: (a) Average RMSD to native; (b) average fraction of native contacts; (c) average radius of gyration. (d) Average Cys-Cys distance for reduced disulphide bonds compared to the average distance between bonded cysteins (~ 2 Å).

3.3 Sensitivity Assessment

Even though the results in Fig. 3.9 are quite promising, the study of a single protein doesn't provide much information about the sensitivity of the method. It is of fundamental importance to understand whether our ab-initio method is sensible enough to detect subtle modifications in the chemical composition of the protein. In order to understand this, we applied our pipeline to a synthetic system: we reduced the disulphide bonds (Cys30 - Cys115, Cys127 - Cys6, Cys94 - Cys76, Cys65-Cys80) present in the crystallized protein (see Fig. 3.2 (b)), we generated new initial conditions via high-temperature MD and used them as a starting point for our folding calculations. Identically to the previous case, we generated an ensemble of 200 rMD trajectories, which we projected on the plane spanned by the fraction of native contacts and the RMSD to native. The plot of the new kinetic free energy landscape $G_{\text{bias}}^{\text{no DB}}(\text{RMSD}_\alpha, \text{RMSD}_\beta)$ is shown in Fig. 3.10. Fig. (3.10) and Fig. (3.3) are

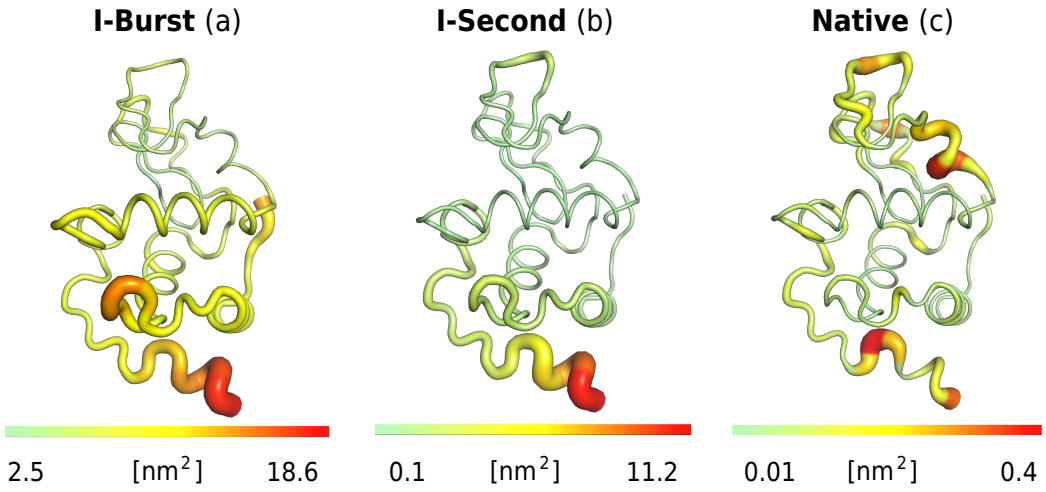

Figure 3.12: Putty representation of the protein crystal structure, colored using the B-factors. B-factors have been computed as $b(a) = 8\pi^2 \mathrm{MSF}(a)$ [249], where $\mathrm{MSF}(a)$ is the mean square fluctuation of residue a in the (a) **I-Burst**, (b) **I-Second** and (c) **Native** basin in the absence of disulphide bonds.

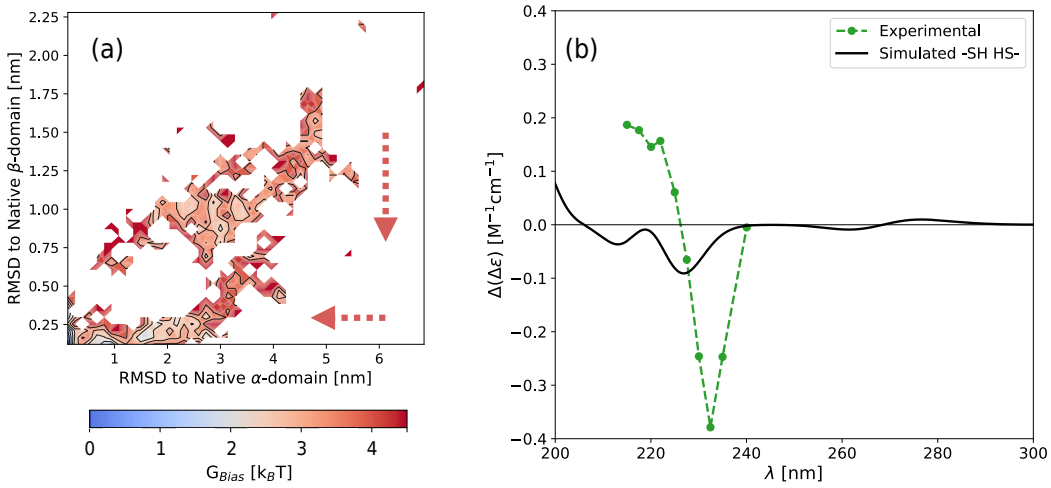

Figure 3.13: (a) Kinetic free energy landscape of the synthetic version of the canine milk lysozyme, where the disulphide bonds have been reduced, projected on the RMSD to native of the α and β domains. The red arrows schematically indicate the main folding pathway in this plane; (b) Kinetic difference CD spectrum, computed as the difference between the **Native** and **I-Second** CD signals in the absence of disulphide bonds.

very similar on a qualitative level. Indeed, both of them reveal the existence of four long-lived states along the folding pathway: an **Unfolded** one, around $Q \sim 0.1$ and RMSD $\sim$ 4 nm, a broad region around $Q \sim 0.5$ and RMSD $\sim$ 2 nm (**I-Burst**), another intermediate state centered in $Q \sim 0.7$ and RMSD $\sim$ 1 nm (**I-Second**) and finally a **Native** state.

To understand the main difference between the states along the folding pathway in presence and absence of disulphide bonds we computed the average RMSD to native, the average radius of gyration, the average fraction of native contacts, the average Cys-Cys distance and the RMSF and B-factors for all configurations harvested from the two histograms. Results are summarized in Fig.s 3.11 and 3.12.

Unfolded and **I-Burst** states obtained in the simulations where disulphide bonds have been reduced are very different from the ones computed in the presence of bonded cysteins. Indeed, the lack of five topological constraints allows the protein to explore much more extended conformations during high temperature MD simulations. This is also confirmed by the high B-factors reported in Fig. 3.12. For what concerns the **I-Second** states, no clear differences emerge at the level of structural compactness of the two proteins (see Fig. 3.11 (a) and (b)). However, three main differences are evident by looking at Fig.s 3.11 (c) and (d), Fig. (3.4) and Fig. 3.12: the couples Cys30-Cys115 and Cys6-Cys127 are always found at a distance much bigger than 2Å, which is the reference distance for a disulphide bond, and the α-domain is much more flexible than the β one in the absence of disulphide bonds. Notably, these two cystein pairs are responsible, in the crystal structure, for the packing of the C-terminal loop to the main protein body (see Fig. 3.2 (b)). Furthemore, the fraction of native contacts in the reduced **I-Second** configuration is higher with respect to the bonded **I-Second**, suggesting that the configurations lacking disulphide bonds are more **Native**-like.

The absence of the five topological constraints introduces a key difference in the protein's folding pathway. While in the presence of disulphde bonds, the protein formed the α and β foldons almost cooperatively, in the absence of those restraints the β domain systematically forms before the α one, as it is clear from Fig. 3.13 (a), where we report the kinetic free energy landscape $G_{\text{bias}}^{\text{no DB}}(\text{RMSD}_\alpha, \text{RMSD}_\beta)$. This folding pathway is made possible by the absence of the Cys30-Cys115 and Cys6-Cys127 bonds, which allow for a higher protein flexibility in the C-terminal region. The three-dimensional structure of the **I-Second** state is thus dramatically different with respect to the one in presence of disulphide bonds (compare Fig. 3.3 and 3.10), and this difference is reflected in kinetic CD difference spectrum computed from the new **I-Second** and **Native** states, shown in Fig. (3.13) (b). Coherently with the analyses we performed, the very low intensity and spectral shape of the

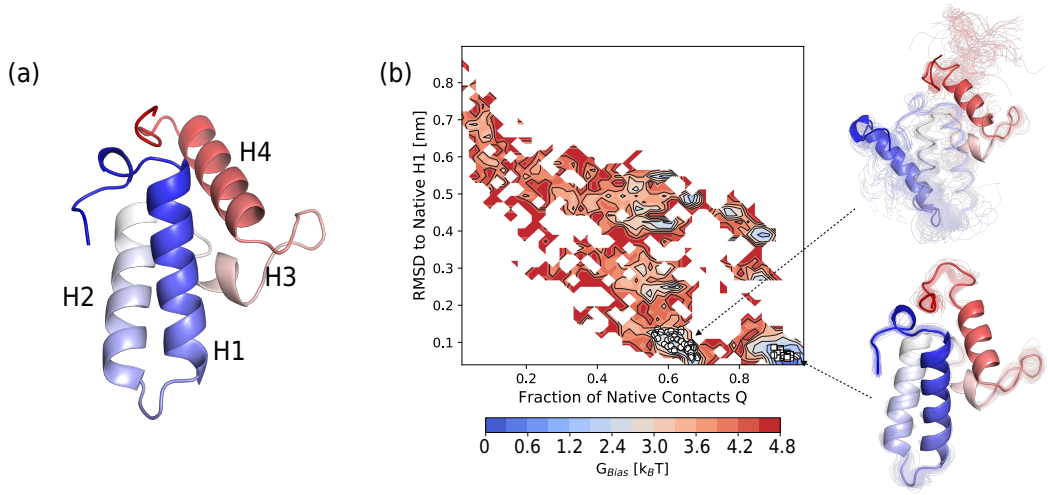

Figure 3.14: (a) Crystal structure of IM7 protein (PDB code: 1CEI), together with the standard nomenclature for its secondary structures; (b) Kinetic free energy profile obtained by projecting the whole ensemble of rMD trajectories of IM7 protein onto the RMSD to native of helix 1 and the fraction of native contacts Q. Low free energy regions reveal the existence of metastable states. Protein structures represent conformations that have been harvested from the minima, while markers dots, when present, represent their corresponding location on the plot.

simulated kinetic CD difference signal suggest that the **I-Second** state in absence of the disulphide-bonds is much more **Native**-like with respect to the one in presence of such topological constraints.

To summarize, we can state that our pipeline is sensible to subtle changes in the chemical modifications of the protein chain that alter its folding pathway. Indeed, experiments have been conducted in the presence of disulphide bonds [32] and our ab-initio algorithm confirms that the kinetic CD difference spectrum of a simulation where the disulphide bonds have been reduced is not coherent with the experimental result.

3.4 Application to the IM7 Protein

As a last application, we tested our ab-initio strategy on the IM7 protein, a small four helix bundle belonging to the family of colicin immunity binding proteins of *E. Coli* (see Fig. 3.14 (a)). The folding of this protein was studied in Ref. [26] using the BF method, therefore in this section we will employ the data generated in that work as an input for our quantum chemical part of the pipeline (for details about the simulation setup, please refer to the original article [26]). This protein provides a perfect additional test, as the CD spectrum in the near-UV region does not come

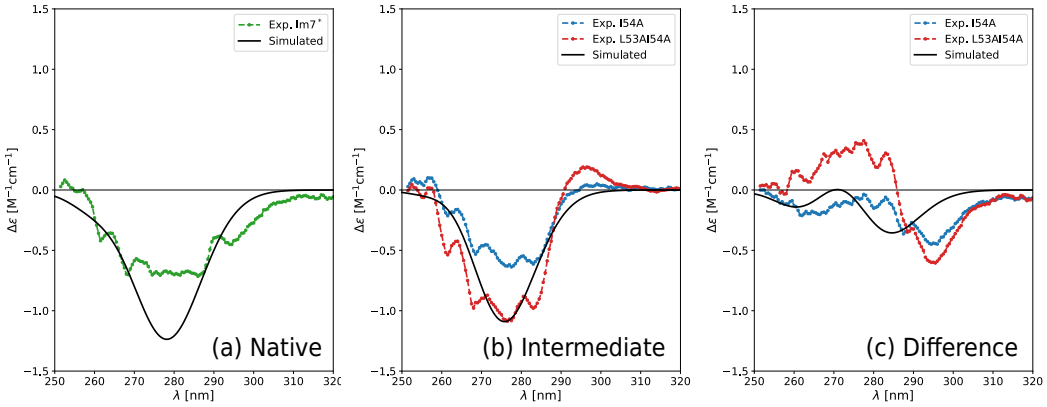

Figure 3.15: CD spectra of IM7 computed as the average over all the harvested (a) Native and (b) Intermediate configurations. Experimental curves are shown as green (native), red and blue (intermediates) dotted lines. (c) Kinetic CD difference spectrum computed as the difference between the Native and the Intermediate signals. In both (b) and (c), two experimental curves are shown, corresponding to I54A and L53A/I54A.

from the coupling of tryprophans alone, but rather it comes from the interaction of the only tryptophan in the protein (Trp75) with all the other aromatic residues [253].

The theoretical analysis in Ref. [26] resolved a metastable state in the folding pathway of IM7, coherently with a wide experimental evidence [253, 254, 255, 256, 257, 258]. From the study it also emerged that the folding of the protein chain can be followed by observing the degree of nativity of Helix 1 (see Fig. 3.14 (a)). Indeed, coherently with previous studies [259, 260], Helix 1 should be already completed in the intermediate state while the rest of the protein is not yet fully native. For this reason, the intermediate should be visible by looking at the kinetic free energy landscape along the fraction of native contacts Q and the RMSD to native of Helix 1, reported in Fig. 3.14 (b). It can be clearly seen that, as expected, a metastable state emerges at $\mathrm{RMSD}_{\mathrm{H1}} \sim 0.1$ nm and $Q \sim 0.6$. Using the harvesting prescription described in section 3.2.2, we selected from the Intermediate and the Native basins a number of configurations sufficient to attain convergence of the CD signal.

As reference experimental data, we refer to the work in Ref. [253], where the folding of IM7 was thoroughly studied by employing CD in both far- and near-UV regions. Differently from the lysozyme case, however, comparison with experimental data is not immediate. First of all, the native state of the protein, called IM7*, is tagged with a hexa-histidine tag. Secondarily, the intermediate state of IM7 doesn't live long enough to be detected by CD experiments. Thus, authors of Ref. [253] proposed a set of mutations that selectively destabilize the native state in favor of the intermediate one. In the following, we will postulate that the intermediate

state obtained from BF simulations describes the experimental intermediate states attained by Ile54Ala and Leu53Ala/Ile54Ala mutations[4].

The experimental near-UV CD spectra are reported in Fig.s 3.15 (a) and (b) using dotted green and grey curves. Native IM7* shows two negative peaks around $\lambda \sim$ 270 nm and $\lambda \sim 295$ nm, while both the selected intermediates show a substantial negative absorbance in the $270 - 280$ nm wavelength window. Our simulated CD spectra are reported in Fig.s 3.15 (a) and (b) using solid black lines. CD curves have been computed by taking into account the contributions from all the aromatic residues in the protein. Our results for the native state average out the two negative peaks and produce a single, large negative peak around 280 nm (see Fig. 3.15 (a)). The negative band around $\lambda \sim 280$ nm of the intermediate states is also predicted by our simulations (see Fig. 3.15 (b)). We stress that, despite the fact that the two spectra (Native and Intermediate) look very similar, they are actually sensibly different. This can be clearly seen by looking at the kinetic CD difference spectrum in Fig. 3.15 (c). Our simulated spectra correctly identify the differences between the Native and the Intermediate signals, predicting a signal which is almost quantitatively compatible with the Ile54Ala experimental one. The main difference between the measured and the computed CD signals resides in a ~ 10 nm blue-shift, which could be imputed to the lack of vibronic effects [261] but was nonetheless similarly present in the case of lysozyme (see Fig. 3.9).

3.5 Chapter Conlusions

Circular dichorism is routinely being used to monitor the level of formation of secondary structures in a protein. The synergy between CD and stopped-flow apparatuses makes it possible to obtain information on the content of α-helices and β-sheets while the folding process is happening. Unfortunately, CD data alone do not provide information at atomic-level resolution, but rather a coarse-grained information about the chain packing. In this chapter we discussed the fact that the CD signal coming from aromatic residues provides a powerful probe of changes in the protein's tertiary structure and we showed that the combination of the BF method and quantum chemical calculations yields an atomic-level interpretation of these modifications. Our independent analyses of the canine milk lysozyme and IM7 protein confirm the robustness and the accuracy of the proposed approach. Furthermore, the study of a synthetic variant of canine milk lysozyme, where all the disulphide bonds have been

[4]This is a standard notation to describe protein mutations. For example, Ile54Ala (or I54A) tells that isoleucine 54 has been mutated with an alanine. When two mutations are separated by a "/" it means that they are both implemented in the same protein.

reduced, confirms the high sensitivity of the method: even though the same number of metastable states are detected, the corresponding CD signals do not match the experimental observations.

Even if we report a quantitative accuracy against experimental results, it is still worth noticing that the method shows some limitations. The first one comes of course from the fact that the BF biasing scheme is applied to make the folding calculations feasible. In particular, the bias force, presented in section 1.2, is determined a-priori and might lead to systematic errors in the calculation of folding paths. A possibility would be then to use the self-consistent path sampling algorithm (see section 1.3.2), which iteratively optimizes the reaction coordinate. In the original paper [28], a first attempt to fold the canine milk lysozyme with SCPS was proposed, showing that the computational cost of simulating proteins that fold in the time scale of seconds is still sustainable with small clusters. A second limitation comes from the fact that our determination of intermediate states is reaction coordinate dependent. This fact is clearly highlighted by the differences in the analyses of the canine milk lysozyme and IM7 protein. Unfortunately, powerful dimensional reduction methods such as MSMs or tICA cannot be trivially employed on BF data, because one of the main assumptions of these methods, i.e. microscopic reversibility, is not satisfied. Approaching to the problem using SCPS would be, once again, a possible solution to the problem thanks to the relationship between the σ variable and the committor function established in section 2.4.2. Indeed, the committor function would provide an optimal reaction coordinate on which to project the data and select configurations in an unbiased fashion.

A possible development of this methodology would be to couple it with ensemble refinement methods [190, 262], which would help selecting the best configurations compatible with experimental data. Therefore, a possible way to modify our approach would be the following: (i) simulate folding trajectories by means of SCPS; (ii) use the strategy discussed in section 2.4.5 to relate σ to the committor function; (iii) detect the existence of metastable states as plateau regions in the $q(\sigma)$ function; (iv) employ techniques borrowed from integrative structural biology (e.g. Maxim Entropy methods [190, 263, 264]) to select a-posteriori the best protein configurations compatible with experimental data.

Still, without any further modifications, the method proposed here looks promising and its application should not be limited to proteins whose folding pathways are already known. In the future, this kind of analysis could be employed to deal with protein mutations and design, having the goal of determining the reaction mechanism and identifying of metastable states along the folding pathway.

Chapter 4

Case Study: Folding and Misfolding of α1-Antitrypsin

Serpin (**ser**ine **p**rotease **in**hibitor) proteins constitute a superfamily of similar proteins whose biological role is protease[1] inhibition [265]. Serpins are widely known because of their peculiar mechanism of action, in which their reactive center loop is cleaved by the interaction with the target enzyme and the protein performs a large conformational transition in order to kinetically trap the target [266]. These proteins are known to fold to a metastable state, and finally reach their fully native state only after making a transition to a biologically inhert state (latency transition) [267]. Despite the advantages of the conformational transition mechanism with respect to the usual competitive mechanism of proteases, serpins suffer from a wide variety of pathological mutations [38, 39]. These mutations induce protein misfolding and alter the secreted levels of secreted inhibitors, leading to severe diseases called serpinopathies [268]. Being able to characterize with atomic detail how mutations alter folding propensities of protein is a fundamental step in the quest for finding therapeutical remedies against pathological conditions induced by misfolding.

In this chapter we will focus on α1-antitrypsin (A1AT), a serin protease inhibitor whose mutations result in the most common serpinopathies: *A1AT deficiencies* (A1ATD, [269]). A1AT is a 394 residue, single-domain protein composed by 3 beta sheets (from A to C in Fig. 4.1) and 9 helices (from A to I in Fig. 4.1), which surround a β-sheet scaffold. The reactive center loop (RCL in Fig. 4.1) protrudes from the main protein body and contains a breakable bond mediating A1AT inhibitory specificity. A1AT is encoded by SERPINA1 gene [269] and it is one of the major anti-protease molecule in humans. It is secreted in hepatocytes[2] and transported through the circulatory system in lungs, where it regulates the levels of neutrophil

[1] Enzymes that break peptide bonds in proteins.

[2] Liver cells

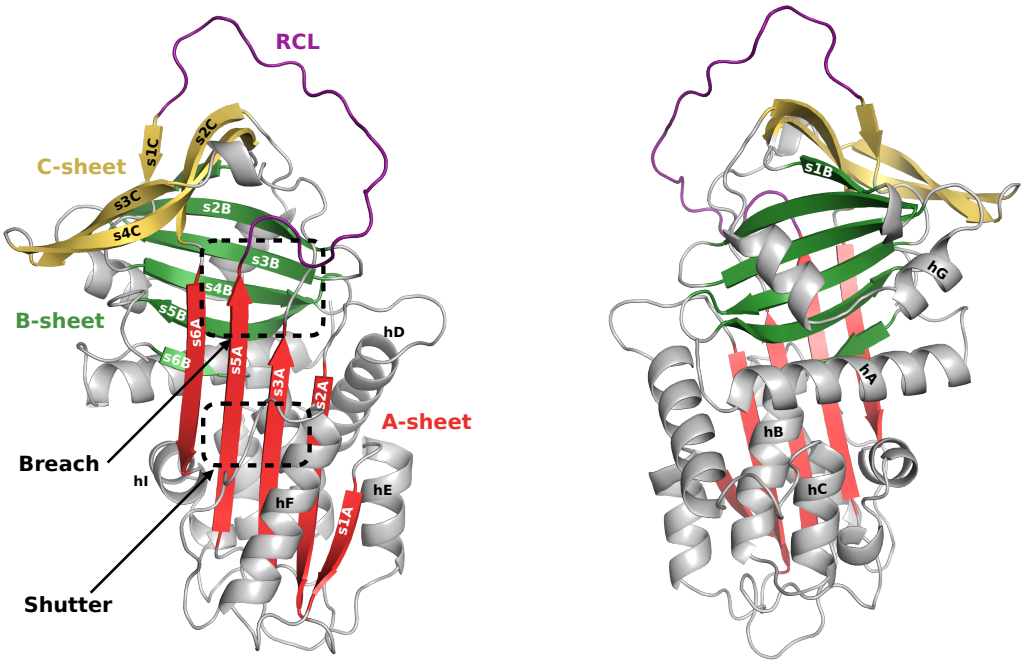

Figure 4.1: Structure of A1AT (PDB code: 1QLP). Relevant structural elements have been colored for clarity. Nomenclature of the secondary structure elements has been taken from [273].

elastase [267]. SERPINA1 is a highly polymorphic gene: more than 150 mutations have already been described [269]. The most common one is a single nucleotide mutation that leads to a severely pathological A1AT variant, in which a glutamic acid is substituted with a lysine (Glu342Lys, Fig. 4.2 (a) and (b)). The resulting mutated protein is usually called *Z-mutant*. This mutant is known to fold into non-native aberrant configurations, and results into the accumulation of A1AT in the endoplasmatic reticulum of hepatocytes [270], causing cellular damage, like hepatocellular carcinoma, and liver diseases, such as cirrhosis and hapatitis [271, 272]. A further effect of Z-mutant retention in hepatocytes is the lack of A1AT in the circulating system, which leads to lung structure breakdown (emphysema) due to aberrant concentration of neutrophil elastase [268].

In spite of the many experimental studies that have been conducted since the '90s [270, 274, 275, 276, 277, 278], detailed structural information about misfolded A1AT is still lacking, slowing down the research on possible ways to rescue the folding process or to increase the propensity of misfolded states to be targeted for degradation [279, 280, 281]. The reason why it is so difficult to determine the three-dimensional structure of Z-mutant is its polymerization propensity [270, 282]. To date, no wide consensus has been reached on the structure of the Z-mutant polymers. The first polymerization model, model A, suggested that polymers can be created

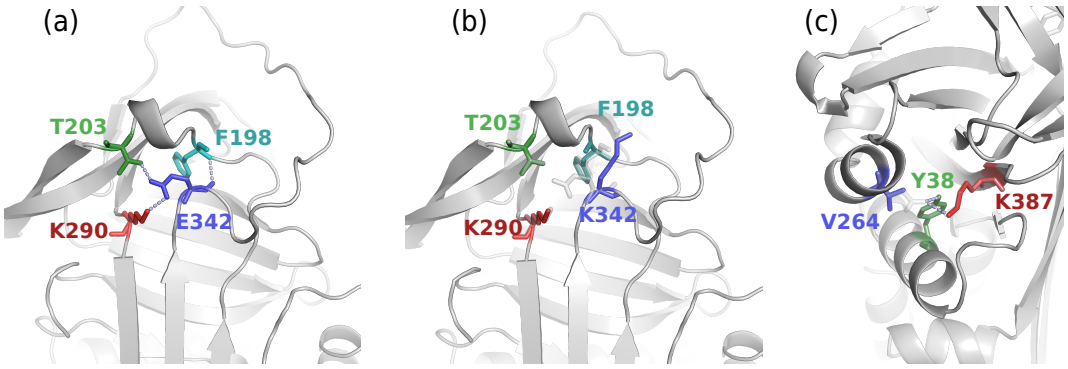

Figure 4.2: (a) Network of interactions formed around Glu342 in the WT protein; (b) disruption of the interaction network around residue 342 after the Glu342Lys mutation; (c) network of interactions around residue 264 after Glu264Val mutation. In both (b) and (c), transparent sticks show the positions of the residues before the mutations.

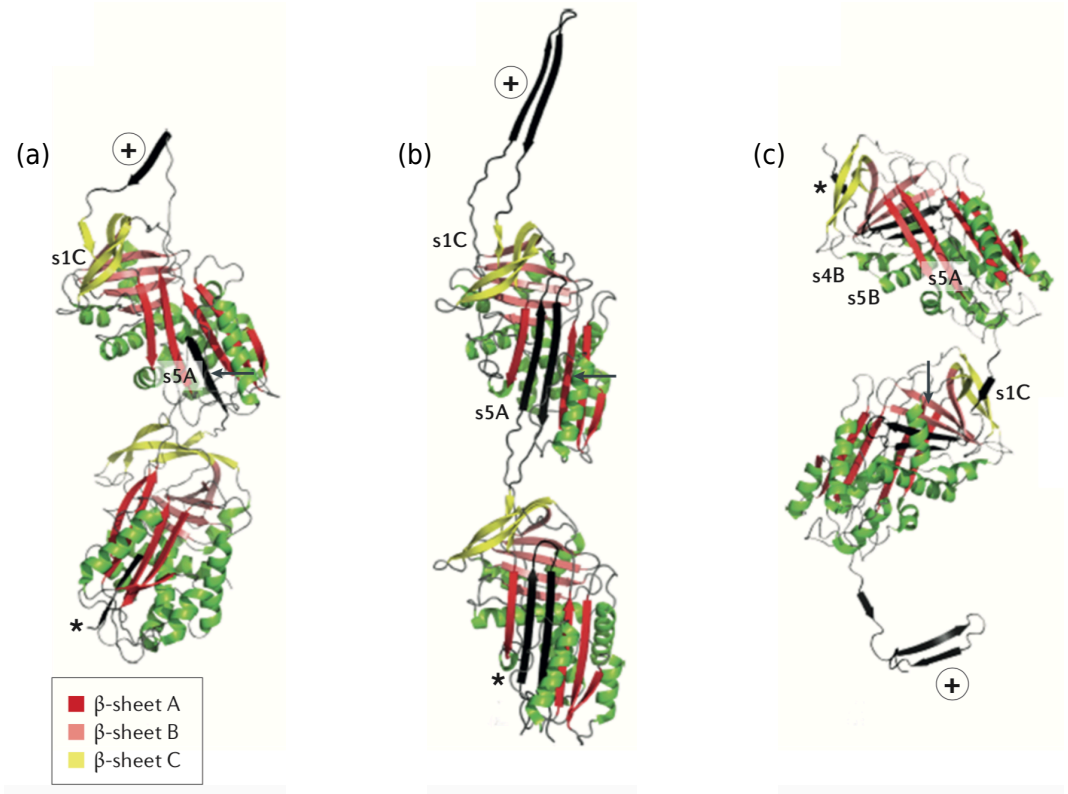

Figure 4.3: Adapted by permission from Springer Nature: Macmillan Publisher Limited, part of Springer Nature, Nature Review Disease Primers [269], Copyright 2016. The three main polymerization models for Z-mutant: (a) reactive center loop of a protein inserts into sheet A of another one, mimicking the completion of sheet A after the protein's inhibitory transition; (b) the β-hairpin formed by the RCL of a protein links to the sheet s5A of another one; (c) a flexible linker formed by strands s1C, s4B and s5B inserts in the C-terminal region of another protein.

by the linkage between the RCL and β-sheet A [270] (see Fig. 4.3 (a)). Alternative conformations where also proposed, suggesting either a model, model B, where the linker is a β-hairpin formed by the RCL and strand 5A (see Fig. 4.3 (b)) or where, model C, the linker is instead formed by the C-terminal β-hairpin (strands 4B and 5B) and s1C (see Fig. 4.3 (c)). Model A is supported by A1AT chemical denaturation and refolding experiments [273] and by ion-mobility mass spectroscopy [283], while model C is supported by SAXS data [284]. Model B is instead more typical of other serpins, e.g. antithrombin [285] and indeed it was never observed in A1AT experiments.

Z-mutant has a very high polymer formation rate, which is proportional the destabilization effect induced on the protein by the mutation [286]. Other species, with lower rate of polymer formation, thus less toxic, have also been studied. Among those, the *S-mutant* (Glu264Val, Fig. 4.2 (a) and (c)) [287] results in a decrease of circulating A1AT in blood and it is generally recognized as a mildly misfolding variant [288].

To date, no definitive therapeutical strategy exists for A1ATD, and the only clinical possibilities to tackle this condition are liver transplant [289] and A1AT augmentation therapy [290]. While the former solves the problem at its core by replacing all the liver cells in the patient, the latter proved to be effective in slowing down emphysema progression [290, 291, 292]. The biggest issues with A1AT augmentation therapy concern the dosage, which to date is still arbitrarily decided [269], and the fact that it is a lifelong therapy that requires repeated infusions. New approaches are currently under examination, although none of them has been approved for human use so far [293, 294, 295]. We briefly summarize here some of the most promising ones (not in numerical order):

1. **SERPINA1 silencing** proposes to prevent mutants secretion as a whole by targeting the corresponding gene, with encouraging pre-clinical results [296]. Preventing all A1AT production, however, would increase lung damage, so this therapy would need to be integrated with standard augmentation therapy;

2. **Mutant degradation in cell** can be carried out by enhancing cell autophagy[3], thus alleviating the presence of polymers and reducing liver damage [293, 294, 295]. Tests are still carried out to assess the optimal dose of autophagy-inducing drugs, with however inconclusive results [269];

3. **Inhibition of A1AT polymerization** aims at stabilizing functional conformers and reducing the polymerization propensity of A1AT mutants using small

[3]The process by which the cell can recycle organels. It is usually activated to remove damaged components in the cytoplasm.

molecules [297, 298, 299, 300, 301, 302, 303]. However, most the compounds showed undesired behavior in further tests and it proved very difficult to devise drugs for animal models [269];

4. **Improvement of A1AT folding** should be in principle possible by means of chemical chaperons, and this approach would simultaneously benefit liver and lungs. However, promising results on animal models [304] came to nothing on a pilot clinical trial [305].

It is evident at this point that the lack of an atomic-level misfolded structure of S- and Z-mutants is reducing the probability to find a possible therapeutic strategy to cure A1ATD. In particular, approaches 3. and 4. would extremely benefit on the knowledge of these structures. In previous chapters, we widely discussed the attractive possibility to approach problems like this with molecular dynamics: in principle, one would be interested in simulating the folding of A1AT and its variants to assess the differences between the mutated species. However, we also conveyed that for biological interesting proteins (except a few but relevant exceptions, e.g. Ref. [306]) this is a hopeless effort, because of the problem of timescale decoupling. Serpins makes no exception in this case: they are composed by ~ 400 residues and need tens of minutes to fold [274, 277, 278]. For this reason, in this last chapter we will rely on the BF method to study the folding of A1AT and its mutated variants, in the quest to provide useful atomic-level structures that might be useful for medical research. The chapter is structured as follows: in section 4.1 we will discuss the details of the simulation setup; section 4.2 is devoted to the discussion of the results of wild type (WT, not mutated) A1AT folding; in section 4.3 we will discuss the folding of the two pathogenic mutations Z and S and we will try and elucidate the reasons behind some experimental evidences on these two protein variants by simulating the folding of three suitable *rescue mutations*, namely Glu/Lys (Lys290Glu/Glu342Lys), Glu/Glu (Lys290Glu/Glu342Lys) and Ser/Lys (Lys290Ser/Glu342Lys). Particular focus will be given to the comparison with WT and experimental results; finally in section 4.4 we will draw some general conclusions on A1AT folding.

4.1 Simulation Setup

The reference structure of our BF calculations is the x-ray crystal structure of WT A1AT [307] (PDB code: 1QLP). To our knowledge, the first 22 residues of all the published A1AT crystal structures are disordered: for this reason, the sequence we employed in our simulations starts from Phe23. From now on, the 1QLP structure with the first 22 residues removed will be referred to as the WT native structure. As

we anticipated at the beginning of this chapter, this A1AT state is only metastable. However, the purpose of this chapter is such that no ambiguity between the biological native state and the state used as a target for the BF simulations should emerge.

All the A1AT mutated variants were generated by means of the mutator function in VMD [46]. After all the corresponding simulations were already completed, a crystal structure of Z mutation was published (PDB code: 5IO1, [308]), indistinguishable from the one we obtained by in silico mutation. Indeed, the RMSD between the two structures is approximately ~ 0.6Å, while no sensible difference emerges at the level of contact maps.

For every A1AT species, the starting configurations of each BF simulation were generated by running 1 ns MD at $T = 1600$ K with an integration timestep of $dt = 1$ fs and subsequently relaxing the conformation corresponding to the last frame for 10 ns at $T = 300$ K. This procedure permits us to sample the region of conformation space corresponding to $Q < 0.1$, which we identify with the unfolded state of the protein. From each initial condition, we ran 12 rMD simulations at $T = 300$ K, each of them lasting 30 ns with and integration timestep of $dt = 1$ fs and a ratchet constant of $k_R = 0.02$ kJ/mol. All the rMD simulations were performed using the Amber ff99SB-ILDN force field [142], thermostatted using the Bussi algorithm [143] while solvation effects were taken into account by means of the Generalized Born implicit solvation model implemented in Gromacs 4.5.2 [41].

Before we proceed and discuss the results of our simulations, we want to add an important remark. In some BF trajectories, the final state of the simulation displays spurious strand-crossings that are not present in the crystal structures. This should not be imputed to protein mutations but rather to the intrinsic ambiguity of the contact map that doesn't sensibly change in presence of the cross-over. An example is reported in appendix E.2.

Finally, we anticipate that, in the incoming sections, protein structures are colored by foldon, where each foldon has been defined a posteriori after the simulations.

4.2 Folding of A1AT WT

We begin our investigation of A1AT by studying the folding pathway of WT. The complexity of this protein makes it easier to understand the folding process by first analyzing a typical folding trajectory, Fig. 4.4 (a), and then by looking at the global folding landscape emerging from all the simulations, Fig. 4.4 (b). Folding starts by forming local foldons (Fig. 4.4 (a), stage 1) and then progresses through the formation of interactions between amino acids at the top of strand 5A and 6A and the ones in the B-C barrell (Fig.4.4 (a), stage 2). In particular, in this stage

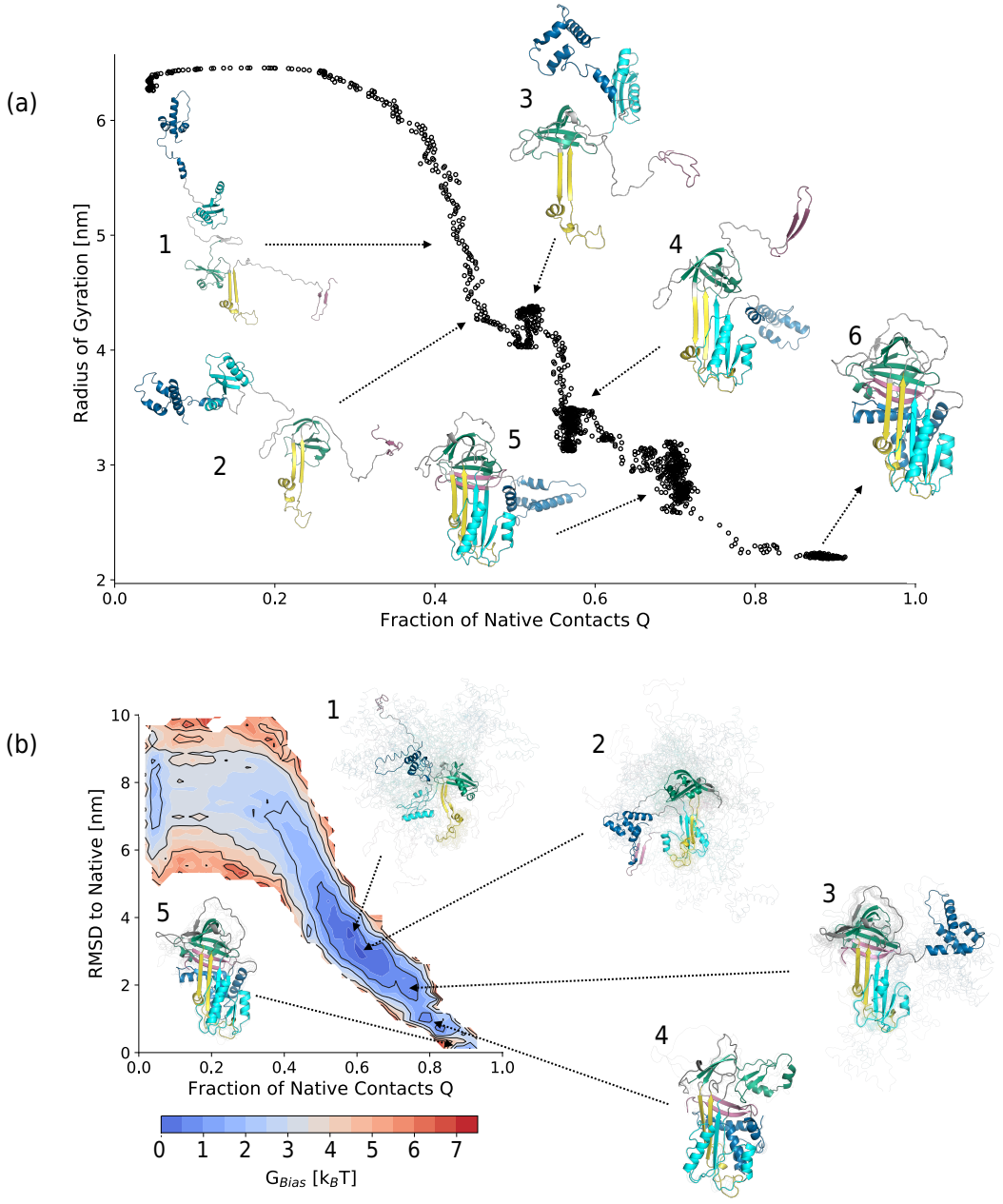

Figure 4.4: (a) Typical successfull WT folding pathway, described by the evolution of its radius of gyration, R_g, as a function of the fraction of native contacts, Q. The structures represent snapshots of the folding trajectory; (b) kinetic free energy landscape obtained by projecting all the WT folding simulations on the plane defined by the fraction of native contacts Q and the RMSD to native. For each kinetic free energy minimum, an ensemble of representative structures is shown, which was obtained by following the harvesting prescription described in section 3.2.2.

Glu342 and Thr203 form an hydrogen bond, while Pro289, Met220 and Met221 form a network of VdW interactions (see Fig. 4.5). At this point, the loop at the top of

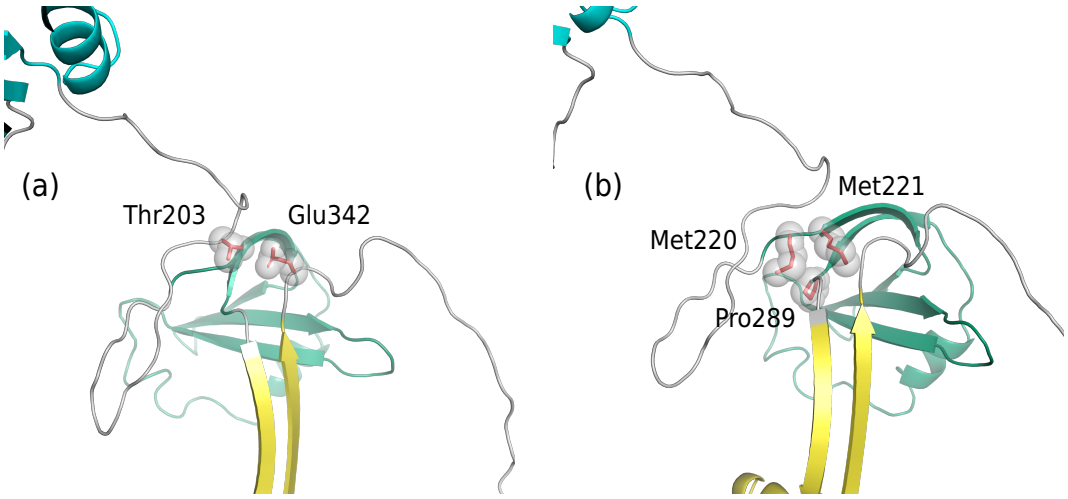

Figure 4.5: Close up of the interactions formed in stage 2 of Fig. 4.4 between (a) Thr203 and Glu342 and (b) Met220, Met221 and Pro289.

strand 3A interacts in a non-native fashion with β-strands 2B and 3B, preventing in this way the docking of strands that form sheet A and also preventing the correct positioning of the B-C barrell (Fig. 4.4 (a), stage 3). The rupture of the non-native contacts finally permits the positioning of B-C barrell and the formation of sheet A (Fig. 4.4 (a), stage 4). In the last stages of folding, the C-terminal hairpin and the N-terminal helices are still solvent exposed: folding completes by docking first the C-terminal hairpin to strands 1B-3B (Fig. 4.4 (a), stage 5) and finally packing the N-terminal helices to the back of the A strands (Fig. 4.4 (a), stage 6).

Let us focus now on the patterns emerging from the projection of all the folding trajectories on the fraction of native contacts Q and the RMSD to native (see Fig. 4.4 (b)). The predicted metastable states in Fig. 4.4 (b), identified as low kinetic free energy regions, are coherent with the inspection of the R_g graph in Fig. 4.4 (a), and provide a visual representation of the structural heterogeneity inside each state. The main difference between the representations in Fig. 4.4 (a) and (b) is provided by state **4** in Fig. 4.4 (b). The transition **4**→**5** represents a sub-dominant pathway where the B-C barrel consolidates as a last part of folding. A main prediction of the BF method is thus that the folding pathway of WT A1AT systematically progresses by docking foldons, which are formed in the early stages of the simulation, in a predetermined order. Furthermore, folding progresses along a single dominant pathway, summarized as follows: (i) formation of local foldons; (ii) development of non-native interactions between foldons; (iii) completion of sheet A while N- and C-terminal regions of the protein remain solvent-exposed; (iv) docking of C-terminal β-hairpin; (v) packing of N-terminal helices.

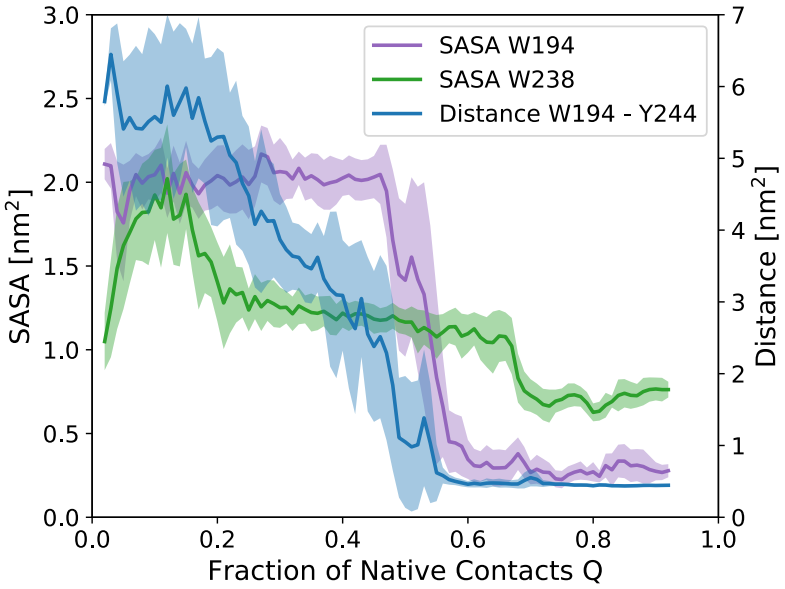

Figure 4.6: SASA of Trp194, Trp238 and the Trp194-Tyr244 distance computed along the WT folding simulations as a function of the fraction of native contacts Q.

4.2.1 Comparison with Experiments

BF results on the folding of WT A1AT qualitatively agree with several experimental findings. First of all, the docking of C- and N-terminal regions has been studied through fragment complementation [276], showing that a fragment containing the C-terminal β-hairpin docks the N-terminal helices fragment only in presence of sheet s5A. This is coherent with the results of the BF calculation, predicting that the C-terminus is incorporated into the protein only after the formation of sheet A has been completed. This is also coherent with kinetic refolding experiments monitored by hydrogen/deuterium exchange, which concluded that s5A and s4B are among the last regions to lose solvent exposure [277, 278].

Extensive experimental research has also been carried out to understand why the folding of A1AT needs to pass through its metastable (biologically active) state before reaching the native (latent) state [277, 278]. These studies revealed that insertion of the RCL into sheet A in early stages of folding is prevented. This is the result of two complementary effects: on the one hand, the B-C barrell forms early during folding and anchors the RCL into a solvent-exposed configuration; on the other hand the fast hydrophobic collapse of sheet A hinders the insertion of the RCL between s3A and s5A. These results agree with our BF simulations, where the B-C barrell and sheet A form cooperatively, so that the RCL remains solvent exposed throughout the whole simulation. One has to notice, however, that the insertion of the RCL between sheets s3A and s5A is by no means expected to happen in our BF simulations, as this

information is not encoded in the reference contact map C_0. Nevertheless, the way in which BF folding progresses is coherent with experimental observations reported in Ref.s [277, 278].

Another observable which is suitable to be experimentally monitored is the distance between cysteins that form disulphide bonds. Unfortunately, A1AT contains only a single cysteine, but other serpins have been studied (ovalbumin and antithrombin III) that form a disulphide bond once the N-terminal packs to the protein's main body. In these experiments [309, 310] it was shown that the N-terminal disulphide bond forms only after the docking of the C-terminal hairpin. BF simulations thus provide a testable explanation for these experimental results, predicting that ovalbumin, antithrombin III and A1AT serpins share the subsequent docking of C- and N-terminal regions as the last step in their folding pathway.

A1AT contains two tryptophans: Trp238, located in strand B and Trp194 found in the breach region. Equilibrium unfolding studies of WT show that Trp194 performs two transitions, a native-to-intermediate and an intermediate-to-unfolded one, while Trp238 displays a single, broad transition [311]. Trp194 fluorescence studies also reported that WT refolding progresses at least through three phases: a fast collapse ($\sim$ 50 ms), a slower transition ($\sim$ 500 ms) and an ultra-long phase (hundreds of seconds) [274]. It is known that the two main sources that can alter Trp194 fluorescence are solvent accessibility and quenching by Tyr244 [312]. These effects can be monitored in our BF simulations too, by computing Trp194 and Trp238 solvent accessible surface areas (SASA) and Tyr244-Trp194 distance. Results are reported in Fig. (4.6) (a): SASA values have been computed using the Shrake-Rupley algorithm implemented in MDTraj 1.9.0 [313, 42], setting a radius of 0.144 nm and 100 sphere points. Coherently with experimental results, Tyr244-Trp194 distance shows two transitions: the first one corresponds to the formation of the local foldons (Fig.s 4.4 (a) stage 1, and (b) basin **1**) while the second one occurs during the docking of sheet A to the B-C barrel (Fig.s 4.4 (a) stage 3, and (b) basin **2**). Trp238 SASA shows a globally decreasing behavior, qualitatively consistent with the experimental evidence of a single broad transition. Instead, Trp194 SASA does not show the three distinct phases observed in equilibrium unfolding experiments, but rather it performs a single and rapid drop around $Q \sim 0.55$. Looking at Fig. (4.6) one can notice that Trp194 reaches its native SASA during the sheet A and B-C barrel docking: this means that our simulations are probably missing some subtle effect around Trp194 in the early stages of folding. It is rather difficult to understand what kind of effect are we neglecting, but the fact that Tyr244-Trp194 distance is correctly captured might suggest that the process we are ignoring is indeed subtle and has no global repercussions on folding.

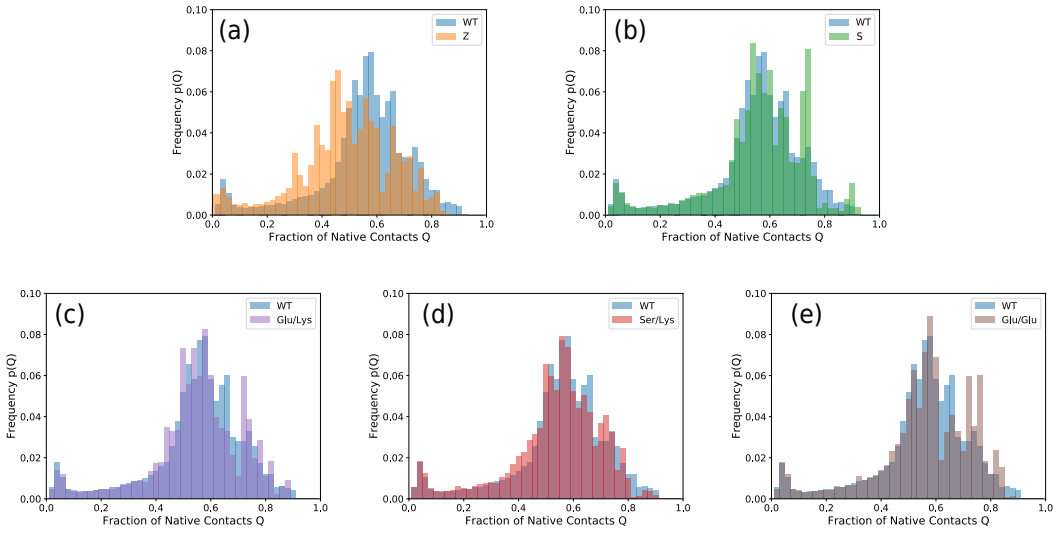

Figure 4.7: Biased probability densities $P(Q)$ of the fraction of native contacts sampled by the folding and misfolding BF simulations of the five different A1AT variants. For each variant, N simulations were available and the histograms were computed using all the frames of each trajectory. The green distribution, representing WT, is shown as a background in each graph and dark shaded regions represent the overlap between the histograms.

During the writing of the reference article, Ref. [25], a paper was published on the folding of A1AT using a Structure Based (Gō) Model [314, 315, 316] (for applications of Gō models on protein folding see Ref.s [317, 318]). Because of the fundamental differences between the two implemented models (i.e. on the one hand a room-temperature all-atom biased simulation and on the other hand a coarse-grained high-temperature one) it is very difficult to compare the results of the two studies, which however broadly agree on several conclusions.

4.3 Folding of A1AT Mutants

After discussing the folding of WT, let us focus on the differences emerging between the folding pathways of WT and its mutations.

4.3.1 Native Contact Probability

As stated at the beginning of this chapter, Z and S mutants of A1AT result in pathogenic conformations because of their misfolding propensity [270, 288]. To understand if BF simulations allow us to see any evidence of an increased misfolding propensity induced by mutations, we considered this simple argument: to a higher misfolding propensity it must correspond a smaller probability to populate high values of the fraction of native contacts. With this in mind, we simulated an ensemble

of trajectories for S and Z and used them to compute the histograms $P_{\text{variant}}(Q)$, corresponding to the probability, in the presence of the biasing force, to populate a given value of the fraction of native contacts Q. Results for WT, Z and S and are reported in Fig.s 4.7 (a) and (b). The comparison between $P_{\text{WT}}(Q)$ and $P_{\text{Z}}(Q)$ immediately reveals that Z mutant is sistematically less able to populate high values of Q, which is a strong measure of its misfolding propensity and thus its pathological activity. The statistical significance of this statement can be assessed by performing a two-sample Kolmogorov-Smirnov test [319], a non-parametric statistical test that assumes as its null hypothesis that the values of Q obtained from the simulations of WT and Z are sampled from the same distribution (see details in appendix E.1). In all our tests, we rejected the null hypothesis with a p-value $p < 2 \times 10^{-16}$, thus confirming our assumption that $P_{\text{WT}}(Q)$ and $P_{\text{Z}}(Q)$ are statistically independent. The reason why Z-mutant shows such inability to populate high values of Q should be looked for in the nature of the mutation. As reported in Fig. 4.2 (a) and (b), the substitution of Glu342 with a lysine destroys a robust network of hydrogen-bonds and electrostatic interactions at the top of sheet A. In section 4.2 we noticed that, in WT, the formation of sheet A and the B-C barrel occur cooperatively and at the early stages of the folding pathway. If the breach region cannot be stabilized because of the lack of the salt-bridge Glu342-Lys290, which is replaced by a Lys342-Lys290 charge repulsion and steric hindrance, then the formation of the B-C barrell incurs in problems as well (because the non-native interaction between strand 5/6A and B-C barrell cannot be formed) leading to the final state where β-sheets are poorly formed and the protein is misfolded. We will go back to this point in section 4.3.2, when we will discuss the kinetic free energy landscape of Z and the other mutations.

For what concerns the more benign S-mutant instead, we see thath $P_{\text{S}}(Q)$ shows a much more WT-like distribution, and populates high values of Q. This again correlates very well with experimental results and with the mildly toxic behavior of S mutation.

It is experimentally well known [320] that misfolding induced by Z mutation can be rescued by applying another mutation (*rescue mutation*), Lys290Glu/Glu342Lys (Glu/Lys). This mutation substitues Lys290 with another Glu thus restoring the original salt-bridge but in a reversed order. We simulated an esemble of folding pathways of this mutated variant of A1AT and, consistently with experimental evidence, we obtained a $P_{\text{Glu/Lys}}(Q)$ distribution which is WT-like (Fig. 4.7 (c)).

In principle, Glu342Lys mutation produces two separate contributions that hinder the correct folding of the chain: an electrostatic one, coming from the breakage of the Glu342-Lys290 salt-bridge, and a steric one due to the higher side-chain length of lysine with respect to the one of glutamic acid. A possible way to test the relative

importance of these two contributions is to leave Glu342 in place and mutate Lys290 with another glutamic acid: Lys290Glu (Glu/Glu). This mutation preserves the electrostatic repulsion already present in Z-mutant but reduces the steric hindrance coming from the length of lysine side-chain. Notably, our BF folding simulations of Glu/Glu predict a $P_{\mathrm{Glu/Glu}}(Q)$ that correlates very well with $P_{\mathrm{WT}}(Q)$ (Fig. 4.7 (d)). This is coherent with cellular studies reporting a $\sim$ 20% secretion efficiency for Z and a $\sim$ 75% secretion efficiency for Glu/Glu compared with WT one [321]. Glu/Glu can thus be considered as another rescue mutation for Z.

Mutation Glu342Lys/Lys290Ser (Lys/Ser) provides an independent test of this hypothesis, as it further alleviates steric clashes. Again, BF simulations predicts Lys/Ser to be again a rescue mutation, because of its distribution $P_{\mathrm{Lys/Ser}}(Q)$ that is more native like than Z-like (Fig. 4.7 (e)).

The substitution of a salt-bridge with an electrostatic repulsion has always been advocated as the reason behind the misfolding propensity of Z-mutant [276]. To our knowledge, this fact was not thoroughly tested, neither experimentally nor theoretically. Indeed, our results on rescue mutations suggest that electrostatic repulsion coming from the Lys342-Lys290 interaction does not play a main role in the misfolding of Z-mutant. Rather the key effect comes from the steric hindrance introduced by the different length between Lys and Glu sidechains. This is a strong but testable statement that can be experimentally assessed through protein mutagenesis and *in-vitro* refolding experiments.

4.3.2 Insight on the Folding Pathways

Further insight on the differences in the reactive mechanism between WT and its pathological mutations can be obtained by comparing their kinetic free energy landscapes. Results are reported in Fig. 4.8 - 4.12.

Z-mutant

As already evident from $P_{\mathrm{Z}}(Q)$ shown in Fig. 4.7 (a), Z mutant is systematically less able than WT to populate highly native conformations. In particular, none of the Z trajectories reached the fully folded state, but rather they got stuck in states with $Q < 0.8$ and RMSD > 1 nm. The most native-like structures reached by Z trajectories do not correctly form the B sheet, leaving part of it solvent exposed (Fig. 4.8, basin **5**), but the more populated basin **4** suggests that most of the configurations remain stuck in much more unfolded states (RMSD $\sim$ 2 nm). Indeed, structures in basin **4** in Fig. 4.8 show the inability of Z to correcly form sheets A and B. This suggests that Z trajectories early diverge from WT ones during folding by hindering the stabilization of β-sheet A.

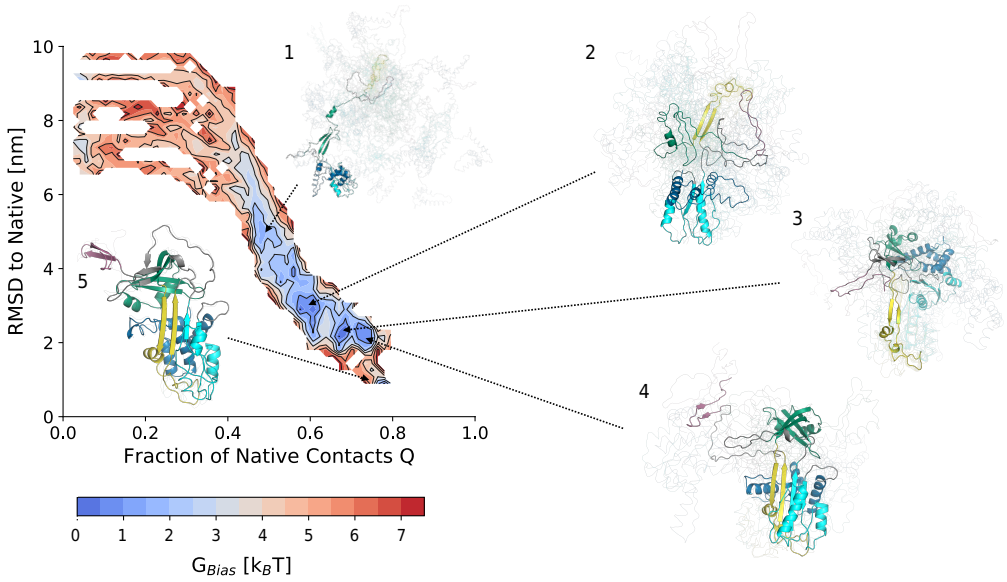

Figure 4.8: Kinetic free energy landscape of Z-mutant, together with ensembles of representative configurations extracted from the predicted metastable regions.

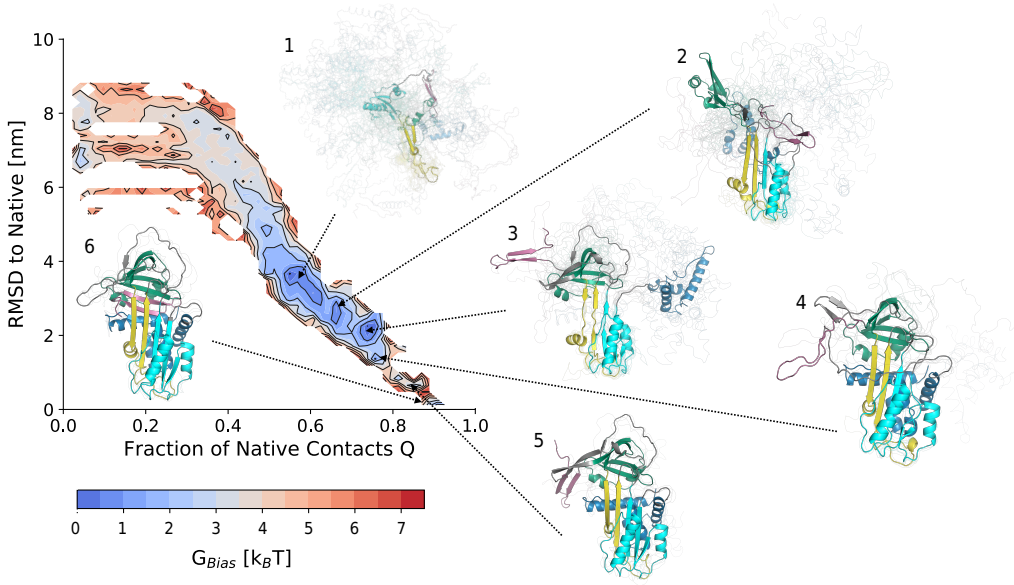

Figure 4.9: Kinetic free energy landscape of S-mutant, together with ensembles of representative configurations extracted from the predicted metastable regions.

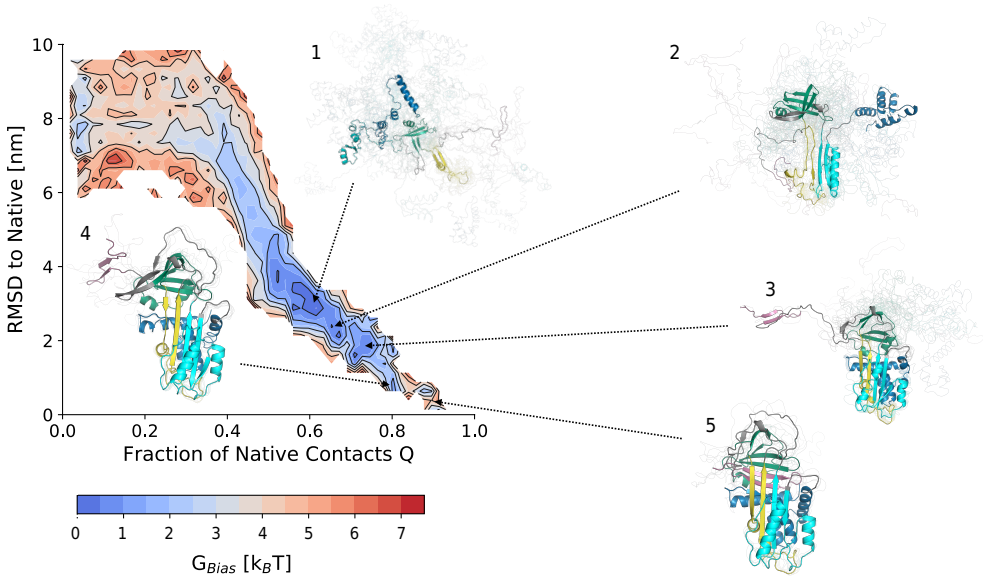

Figure 4.10: Kinetic free energy landscape of Lys290Ser/Glu342Lys mutant, together with ensembles of representative configurations extracted from the predicted metastable regions.

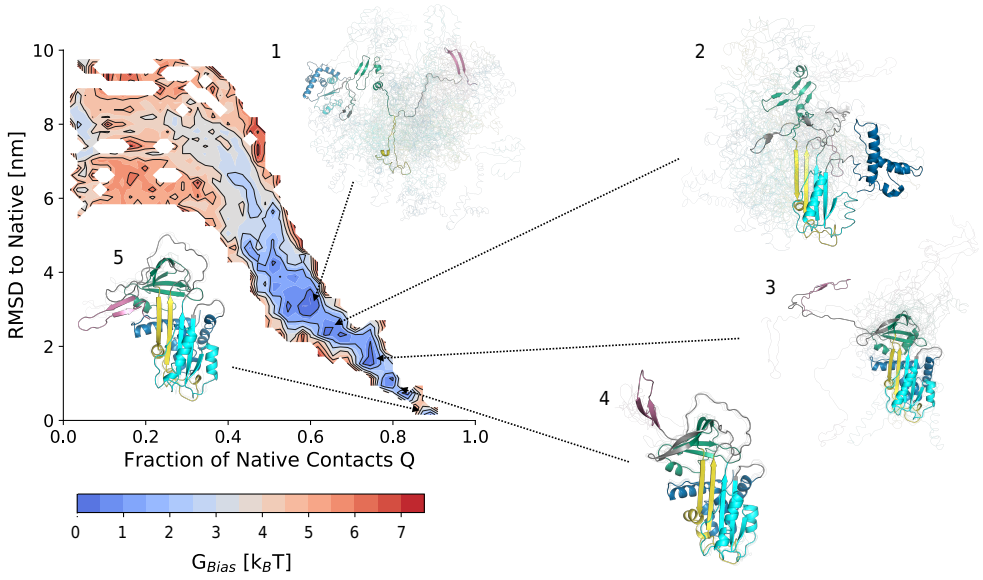

Figure 4.11: Kinetic free energy landscape of Lys290Glu/Glu342Lys mutant, together with ensembles of representative configurations extracted from the predicted metastable regions.

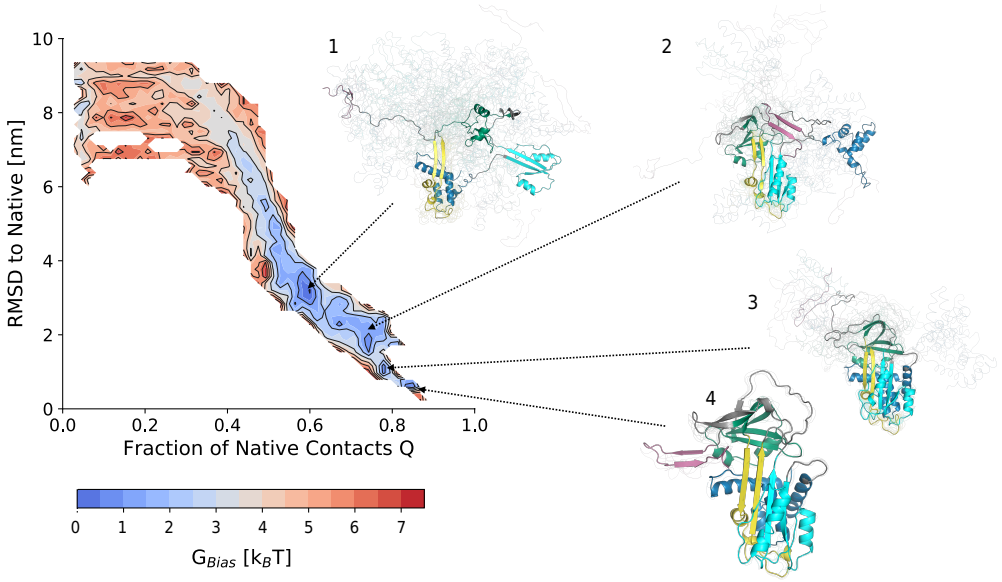

Figure 4.12: Kinetic free energy landscape of Lys290Glu mutant, together with ensembles of representative configurations extracted from the predicted metastable regions.

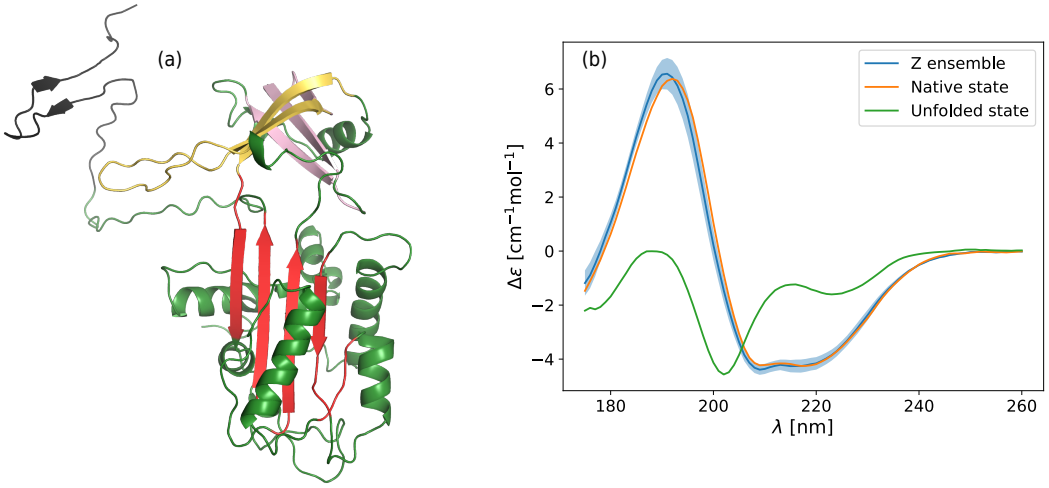

Figure 4.13: (a) Lowest energy conformer of Z-mutant, harvested from basin **4** in Fig. 4.8 (b). Color scheme of the protein is chosen to make it easier to compare it with Fig. 4.3; (b) far-UV circular dichroism spectra of the native state of Z (orange curve), the unfolded states (green curve) and the average spectrum of the states in basin **4**. Shaded regions represent the standard deviation. All the signals have been obtained using the PDB2CD webserver [213].

It is experimentally well known that the polymerization prone state of Z mutant is partially folded [322, 323]. As we widely discussed at the beginning of this chapter, the detailed conformational properties of the chain are still reason of debate (see Fig. 4.3 and Ref.s [273, 283, 284, 285]). BF simulations should in principle be able to provide a potential candidate for the polymerization prone state of Z-mutant. Here we propose as a model the ensemble of conformations harvested from basin **4** in Fig. 4.8. In Fig. 4.13 (a) we show the lowest energy conformer extracted from that basin, which shows a partially malformed β-sheet A, an incomplete and un-docked B-C barrell and solvent exposed C-terminal β hairpin and s1C strand. This conformation is coherent with the polymerization model C in Fig. 4.3 (c), where the hairpin formed by s4B and s5B, connected to the protein main body through s1C, is used as an anchor to dock the adjacent misfolded protein. Because of its polymerization prone nature, structural information about Z-mutant are very poor, so it difficult to validate our model against robust data. Nonetheless we can try to understand if our ensemble of Z misfolded structures is coherent with two experimentally determined properties: (i) the polimerization-prone structure of Z is not globally unfolded, but rather partially folded [270]; (ii) its collisional cross-section (CCS), measured by ion mobility mass spectroscopy [324, 325], is $118 \pm 17\%$ higher then the native CCS. To address the first property, we used an empirical algorithm to predict the far-UV circular dichorism spectrum [213] of the misfolded Z ensemble: Our results, reported in Fig. 4.13 (b) show that the CD signal is more typical of folded proteins than a random coil. We also computed an estimation of the CCS of the misfolded Z ensemble using IMPACT [325], obtaining a CCS $137 \pm 4\%$ higher than the native one. Both our results are consistent with experimental expectations and we can thus conclude that BF simulations predicted Z-mutant configurations compatible with general experimental properties of the misfolded chain.

S-mutant

In evident contrast with the results on Z-mutant, the kinetic free energy landscape of the mildly pathogenic mutation S is much more WT like (see Fig. 4.9). Notably, many trajectories are able to reach the fully native structures, as it can be seen from Fig. 4.9, basin **6**. The main difference between S and WT folding pathways can be found in the docking order of N- and C-termini: in S, N-terminal helices pack before the docking of the C-terminal β-hairpin. A possible explanation of this fact resides in the lack of Glu264-Lys387 salt bridge, which is present in WT and it is destroyed by the Glu256Val mutation (see Fig. 4.2 (c)). In WT, this salt-bridge links helix G with the C-terminal hairpin, helping to anticipate the docking of the latter before the packing of N-terminal helices. The absence of Glu264 also breaks the WT

hydrogen bond with Tyr38 (see Fig. 4.2 (c) and Ref. [288]). The breakage of both the salt-bridge and the important Glu264-Tyr38 hydrogen bond makes it easier for the C-terminus to fit the protein also after the N-terminus has already packed.

For what concerns S-mutant misfolding, in our simulations it seems due to a high barrier separating basin **4** from basin **6**. In **4**, a failure in correctly positioning the B-C barrell (see Fig. 4.9) hinders the docking of the C-terminus and partly sequesters the RCL, reducing its solvent exposure. The reduced SASA of the RCL would explain the reduced biological activity of S-mutant. Moreover, the interesting difference with Z-mutant misfolded conformation is that, in basin **4** of S, s1C is correctly formed and it is not free to move like in the Z conformer in Fig. (4.13) (a). A possible effect of this is the reduced mobility of the anchor formed by C-terminal hairpin and thus a reduced polymerization efficiency. Moreover, the barrier limiting **4**→**6** transition should be high enough to compete with the polymerization process. All these facts together would provide an atomic level explanation of the observations that S-mutant: (i) polymerizes (C-terminus is solvent exposed and can be used to dock an adjacent misfolded protein); (ii) is less efficient in polymerizing than Z (the anchor is much less mobile); (iii) is less toxic than Z (because part of the secreted proteins are actually able to reach the folded state) [268, 323].

Unfortunately, experimental information about the misfolding of S-mutant is very poor, as most of the research is carried out on its more pathogenic relative, Z-mutant. Therefore, our statement is nothing but an interpretation of our simulation data. To make our conclusions testable, one could in principle estimate the transition rate $k_{\mathbf{4}\to\mathbf{6}}$ by means of the TPT-based algorithm presented in chapter 2, measure experimentally the rate of polymer formation [269] and compare them.

Rescue mutants

Besides slight differences in the kinetic free energy landscapes, the folding pathways of Ser/Lys and Glu/Lys and Glu/Glu look very similar to the S one (see, respectively, Fig.s 4.10, 4.11 and 4.12). In particular, the last steps in folding always involve the packing of N-terminal helices followed by docking of the C-terminal β-hairpin. Conversely to S, the reason behind the difference in the order of foldons packing between rescue mutants and WT is not completely clear. Notable differences with respect to S folding pathway appear at early stages in the Ser/Lys mutant folding, where the formation of sheet A and the B-C barrel looks not cooperative, or at very late stages in folding of Glu/Glu, which never reaches a fully folded state leaving the C-terminal hairpin solvent exposed, and of Glu/Lys, which either fails in docking the C-terminus or shows defects in the formation of sheet C. All these results show that, while rescue mutations are able to dramatically reduce the misfolding propensity

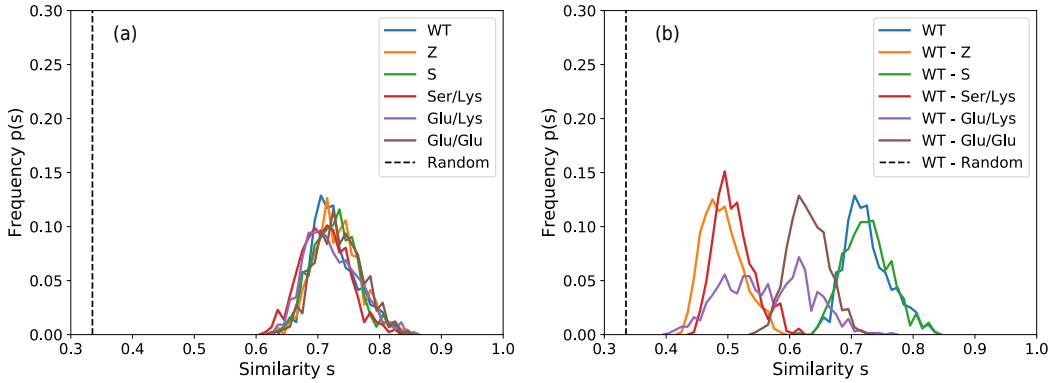

Figure 4.14: (A) Distribution of the self-similarity coefficient for all the A1AT variants. (B) Distribution of the cross-similarity coefficient between WT and the other A1AT mutants, compared against WT self-similarity. In both figures, the dashed vertical curve represents the value of self-similarity obtained by a pseudo-random sequence of contact formations.

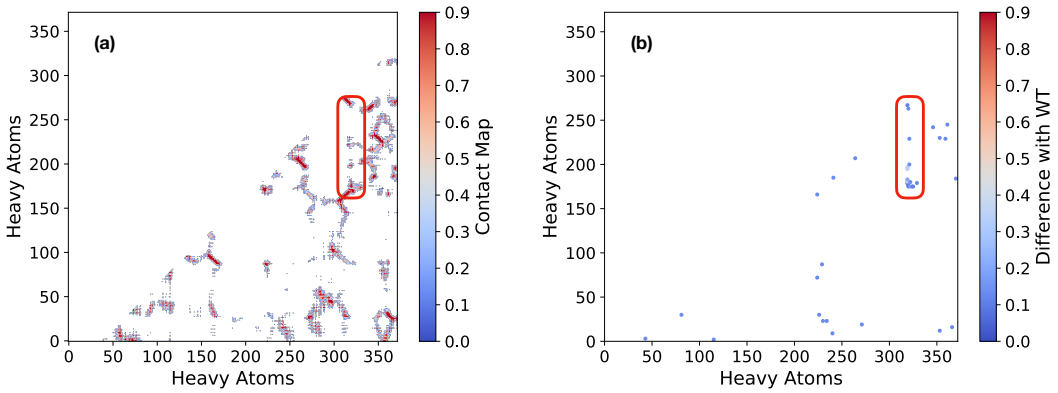

Figure 4.15: (a) Heavy atoms contact map of native WT, computed using Eq. (1.26); (b) difference between the heavy atoms contact maps of WT and Z. In both pictures, the red box shows the location around Glu342.

of Z, the network of hydrogen bonding and electrostatic interactions mediated by Glu342 formed at the top of sheet A is key to provide a fully formed native state.

4.3.3 Order of Contact Formation

From the analysis based on the kinetic free energy profiles, it emerges that rescue mutations are able to lead to highly native configurations. However, Fig.s 4.4, 4.10, 4.11 and 4.12 also show that the original WT folding pathway is not restored after rescue mutations. Path similarity, introduced in section 1.5.2, can help quantify how much the folding pathways of the different A1AT variants deviate from the WT one.

The self-similarities of all the A1AT simulated variants are shown in Fig. 4.14 (a), while cross-similarities are reported in Fig. 4.14 (b). Notably, all the self-

similarities are peaked around $s \sim 0.75$: this is a strong indicator that the folding of all the protein variants progresses along a single, preferential route [144, 23], which might however be different for the different species. This result is consistent with the previous kinetic free energy analysis. Cross-similarities, instead, differ significantly. WT and S pathways are similar $p_{\text{WT}}(s) \sim p_{\text{WT-S}}(s)$, while strong differences emerge between WT and Z and WT and Ser/Lys. A strong similarity between S and WT folding pathway might result as unexpected because of the different N- and C-termini docking order, but it is due to the fact that the early stages of the folding pathway occur in a consistent way, with the cooperative formation of sheet A and the B-C barrel. Strong differences between WT and Z folding pathways were instead revealed in advance by the analysis of the kinetic free energy profiles, as well as the differences between WT and Lys/Ser. Indeed, the former are due to the high misfolding propensity of Z which disrupts an important contact network at the top of β-sheet A (see Fig. 4.2 (b)) and prevents the protein chain to reach the native conformation. The latter, instead, come from the non-cooperative formation of sheet A and the B-C barrel. The remaining mutations show intermediate cross-similarities, with $p_{\text{WT-Glu/Glu}}(s)$ peaked around $s \sim 0.6$ and $p_{\text{WT-Glu/Lys}}(s)$ loosely broadened around the same value.

Some final important comments are in order here. Point mutations do not significantly alter the contact map of the WT native state. This is exemplified by Fig. (4.15), where we show in (a) the WT native contact map and in (b) the difference between the native WT contact map and the Z native contact map. The red boxes in both Fig.s 4.15 (a) and (b) enclose the region around Glu342, which is the residue mutated in Z. As expected, the mutation locally changes the network of interactions around that glutamic acid, but apart from these modifications the two contact maps are indistinguishable[4]. This means that the collective variable used to bias the simulations, given by

$$z(X) = ||C(X) - C^0||^2 \tag{4.1}$$

is substantially identical for every protein variants because $C^0_{\text{WT}} \simeq C^0_{\text{mutants}}$. Thus, all the differences in the folding pathways emerge as a result of the physical force field and and seem to be unrelated to the introduction of a non-physical biasing force.

[4]Some collateral and less intense differences are anyway present, and are mostly due to small changes induced by the further energy minimization step required after applying the in silico mutation.

4.4 Chapter Conclusions

In this chapter we focused on the study of folding and misfolding of A1AT, a protein belonging to the serpin family whose mutations are associated to sever diseases. Our simulations were based on the BF method described in section 1.2 and provided a coherent and experimentally reasonable picture of the folding and misfolding of this protein. In particular, we realized that folding proceeds by the formation of stable foldons that dock in a predefined order along a main folding pathway. The conformational change that concludes the folding process is the ordered docking, respectively, of the C- and N-terminal regions to the main protein body. This is supported by numerous experimental observations [274, 275, 276, 278, 309, 311, 320, 321], and the atomically detailed results of our simulations provide further predictions that can be tested in a wet laboratory. We put particular emphasis in the discussion of the misfolding propensity of mutation Glu342Lys, called Z-mutant. Coherently with experimental evidence, our simulations show that Z-mutant folding pathway deviates very soon from the WT one, leading the protein to a polymerization prone state. In this configuration, a flexible linker formed by the C-terminal β-hairpin and strand s1C remains solvent exposed, coherently with one of the models presented in literature [284]. The all-atom structure extracted from our simulations can be used in princple to aid the development of therapeutical strategies that focus on the degradation of mutants in cell [293, 294, 295] or the inhibition of A1AT polymerization [297, 298, 299, 300, 301, 302, 303]. Misfolding of S-mutant, instead, occurs late along the folding pathway and allows to reach more compact conformations, consistently with a reduced toxicity compared to Z-mutant [288]. Finally, the study of rescue mutations allowed us to question the common paradigm under which electrostatic repulsion between Lys290 and Lys342 plays a key role in Z-misfolding [276]. Simulation data suggest instead that the main actor in Z-mutant misfolding is the steric hindrance introduced by the increased side-chain length of residue 342 in the Glu342Lys mutation. This is a new prediction of our model and can be validated by *in-vitro* studies of the proposed rescue mutations.

Overall, aggregated data show a consistent folding and misfolding picture that can be systematically assessed through experiments. However, it is worth discussing some weak points emerging from our analysis. First of all, our folding simulations are conducted biasing along a pre-determined reaction coordinate, which might be sub-optimal and might lead to systematic errors in our interpretation of the process. Because of the dimension and the typical folding time scale of the protein ($\sim$ 400 amino acids and $\sim$ 10 minutes folding time) it is impossible to validate our results against bare MD trajectories or replicate them using different enhanced sampling

techniques (see the discussion in section 1.1). A possibility would be then to use self-consistent path sampling (see chapter 1 and Ref. [28]) to optimize the reaction coordinate and obtain even more reliable results. In practice this would be possible because, in terms of computational workload, each SCPS iteration performs as good as an rMD simulation. Furthermore, studies performed to date suggest that typically 3-4 iterations may be sufficient to reach convergence. The use of SCPS would also help in defining an optimal reaction coordinate on which to project the configurations. Indeed, the relationship discussed in section 2.4.2 between σ collective variable and the committor function would be extremely helpful in elucidating the reaction mechanism in a quantitative fashion.

A final concern regards the use of an implicit solvation method in our simulations. Even though increasing evidence supports the idea that BF method performs reasonably good in implicit solvent (see for example chapter 3 and Ref.s [24] and [26]), it would be important to test the validity of our conclusions by repeating folding and misfolding simulations in the presence of explicit solvent molecules. In particular, long range interactions are more carefully taken into account in explicit solvent models, providing an ideal and conclusive testing ground to assess the role of electrostatic repulsion in the misfolding of Z-mutant.

Conclusions

Molecular dynamics has gone a long way since its first application by McCammon, Gelin and Karplus in 1977 [326]. From there on, an ever increasing amount of biological systems has been chracterized using molecular simulations, ranging from proteins [327] to RNA [328]. This resounding growth has been possible thanks to widespread free and scalable softwares [48, 49, 51, 50] and to the development of increasingly accurate force fields [191, 193, 194, 195]. Still, the problem of time scale inaccessibility survives, especially for complex phenomena like protein folding [58]. Indeed, the folding of most of the biologically relevant proteins is still far out of reach for modern high performance computing facilities. Neither hardware advancements, such as special purpose supercomputers [9] and de-localized cloud infrastructures [8], nor intense theoretical studies, culminated in the development of enhanced sampling techniques [15, 20, 56, 57, 58, 59, 60, 61, 62], were able to provide an ultimate answer to this problem. Millisecond folding simulations are still not routine for the majority of the laboratories.

The main goal of this manuscript was to build a framework within which folding simulations of medium sized proteins become feasible without requiring extended computational resources. The main ingredient behind this scheme is the Bias Functional (BF) method [23], an enhanced sampling method that proved to be effective in studying proteins with folding timescales above hundreds of milliseconds. The manuscript was structured in two main parts: in the first two chapters we proposed two main theoretical advancements of the BF method, aiming at solving its main drawbacks [28, 29]; the second two chapters were instead dedicated to the application of the BF method to two average-sized proteins [24, 25].

In the **first chapter** we presented the theory behind the BF method, with great focus on its qualities and drawbacks. In particular, we realized a fundamental limitation of the approach: this algorithm employs a predefined reaction coordinate to bias the folding dynamics towards the protein's native state. Given the nature of the method, a sub-optimal choice of such reaction coordinate might lead to the computation of highly unlikely reaction pathways. To overcome this limitation, in this chapter we developed and presented the self-consistent path sampling (SCPS)

algorithm. The latter does not rely on any specific choice of reaction coordinate, but rather it iteratively optimizes a guess collective variable within a mean-field approximation. We proved the performance of this algorithm on a 2-dimensional toy model and on a standard benchmark protein, the WW-domain of FIP35. In both the cases, we concluded that the performances of SCPS were superior to the ones of rMD and, most importantly, in accordance with the results of unbiased MD simulations.

The **second chapter** was devoted instead to the solution of a main problem shared by the BF method and SCPS. The introduction in the dynamics of a history-dependent bias force breaks microscopic reversibility, so that nor BF neither SCPS simulations can be employed to recover kinetic information on the reactive process. To solve this problem, we proved the existence of a connection between SCPS and Transition Path Theory (TPT) [30, 31]. This connection was established through two main steps: the first step required to define a time range (the *steady current regime*) within which short and non-ergodic simulations can be employed to approximate the results of TPT; in the second step we proved that the collective variables of SCPS are strongly related to the committor function and that the average path resulting from the converged self-consistent calculation is a mean-field approximation of a principal curve. Finally, we developed two possible strategies to employ the committor to sample the transition path ensemble. The whole machinery was tested on a 2-dimensional toy model, with promising results.

The first two chapters provide the main theoretical contributions of this manuscript. In particular, the second chapter proposes a general pipeline to start from a crystal structure of a protein and obtain all the TPT observables corresponding to the folding process. In principle, this pipeline should be applicable to medium-sized proteins even with limited computational resources, because all the required simulations are performed using rMD. Extensive testing is however still needed: in particular, SCPS has yet to be investigated in the presence of explicit solvent molecules, while the full pipeline still needs to be applied to the folding of a protein and accurately analyzed. Future works on this subject will go in the direction of trying and closing this gap between theory, hopefully providing robust proof of principles.

The second part of the manuscript was dedicated to two applications of the BF method to medium-sized proteins. The **third chapter** focused on the possibility to use the BF method to enhance experimental sensitivity. In particular, we proved that the synergy of BF folding simulations, quantum chemical calculations and experimental data obtained from near-UV circular dichroism (CD) can help elucidate the atomic level conformations of the folding intermediates of the canine milk lysozyme and IM7 protein. In this chapter we developed a scheme to extract relevant conformations from the BF simulations and compute the near-UV signal coming from

these structures using ab-initio quantum chemical methods. Finally, the simulated signals were compared against the experimental ones to retroactively assess the quality of the atomic-level configurations. In both the cases, we noticed a quantitative agreement of the simulated spectra with the experimental ones.

The **fourth** and final **chapter** focused on a case study, alpha-1 antitrypsin. This protein belongs to the family of serin protease inhibitors (serpins), and its inhibitory mechanism amounts for a complex conformational transition. Its biological function can be altered by several mutations, some of which lead to protein misfolding and consequent severe pathologies, e.g. lung emphysema and hepatic cirrhosis. In this final part of the manuscript we used the BF method to study the folding of the wild type (WT) protein and two of its most common pathological mutations: Z and S. We realized that the folding of WT occurs along a main single pathway, where foldons are formed early in the conformational reaction and then dock in a predefined order. In particular, one of our main predictions is that the N-terminal group of helices docks to the main protein body only after the C-terminal β-hairpin has reached its native conformation. The study of the two mutations revealed to be coherent with experimental expectations: misfolding of Z-mutant occurs early in the simulations and all the trajectories get stuck in a highly aberrant conformation; S-mutant, instead, misfolds late and is able to reach more native conformations and misfolds, coherently with its reduced toxicity with respect to Z. Notably, BF simulations predict an all-atom conformation of misfolded Z-mutant which is coherent with SAXS data. Finally, we also studied the so-called rescue mutations of Z, i.e. mutations expected to restore the folding capability of Z-mutant. The main conclusion of these studies was that, contrary to previous assumptions, the role of electrostatic repulsion in Z-mutation is sub-dominant with respect to the steric one.

The last two chapters show two practical examples of the application to realistic cases of the BF method. In both the studies, comparison with experimental data has always been the main focus. Our analyses always provided at least qualitative, sometimes also quantitative, agreement with pre-existent literature on the subjects. Still, the methodologies presented here need to be further refined: in particular, the choice of the collective variables employed in the study of the processes of interests needs to be taken in an unsupervised fashion. In chapters **one** and **two** we proposed a strategy to overcome this limitation: if BF simulations are further optimized using SCPS, we can use the relationship linking variable s (or σ, in the multidimensional case) to the committor to define an optimal reaction coordinate on which to project our data. To this end, preliminary results on FIP35 and on a 2-dimensional model provided promising results. Extensive testing of these ideas on increasingly complex systems will be the focus of future works.

Appendices

Appendix A

The Fokker-Planck Equation, in a Nutshell

This appendix is devoted to a brief introduction to the tools on which all the theoretical arguments of this manuscript are based.

Let us consider a set of N particles, whose positions are defined by $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ and corresponding masses given by $m_1, \ldots, m_N$. The equations of motion describing the dynamics of the system in the NVE ensemble are simply provided by the Newton's equations:

$$m_i \ddot{\mathbf{r}}_i + \boldsymbol{\nabla} U(\mathbf{x}) = 0 \tag{A.1}$$

where $F(\mathbf{x}) = -\boldsymbol{\nabla} U(\mathbf{x})$ is the force-field. If one wants to thermostat the dynamics to a given temperature T, a possible way is to introduce a damping effect proportional to a friction coefficient γ (for sake of simplicity we will assume that every particle in the system shares the same friction coefficient: $\gamma_i = \gamma \ \forall i = 1, \ldots, N$) and a set of random forces $\boldsymbol{\xi}_i$ that simulate the effect of the hidden degrees of freedom of an infinite thermal reservoir. In this way one obtains the so-called *Langevin equation*:

$$m_i \ddot{\mathbf{r}}_i + \gamma \dot{\mathbf{r}}_i + \boldsymbol{\nabla} U(\mathbf{x}) = \boldsymbol{\xi}_i(t) \tag{A.2}$$

The random force $\boldsymbol{\xi}$ should satisfy the fluctuation-dissipation theorem:

$$\langle \boldsymbol{\xi}_i(t) \cdot \boldsymbol{\xi}_j(\tau) \rangle = 6\gamma_i k_{\mathrm{B}} T \delta_{ij} \delta(t - \tau) \qquad i, j = 1, \ldots, N \qquad \forall t, \tau \tag{A.3}$$

The parameters in Eq. (A.2) naturally define an intrinsic timescale $\tau = \gamma^{-1}$. If one is interested in studying systems whose typical dynamics occurs at timescales $t \gg \tau$, then the Langevin equation can be approximated with its *overdamped* form:

$$\dot{\mathbf{r}}_i = -D\beta \boldsymbol{\nabla} U(\mathbf{x}) + \boldsymbol{\eta}_i(t) \tag{A.4}$$

where $D = \gamma^{-1} k_{\mathrm{B}} T$ is the *diffusion coefficient*, $\beta = (k_{\mathrm{B}} T)^{-1}$ and $\boldsymbol{\eta}_i = \gamma^{-1} \boldsymbol{\xi}_i$ is a random velocity. The autocorrelation of the random velocity is obtained as:

$$\langle \boldsymbol{\eta}_i(t) \cdot \boldsymbol{\eta}_j(\tau) \rangle = 6D\delta_{ij}\delta(t-\tau) \qquad i,j = 1,\ldots,N \qquad \forall t,\tau \tag{A.5}$$

The points computed by integrating Eq. (A.4) sample a time-dependent distribution which is solution of the so-called *Fokker-Planck equation* [184]:

$$\frac{\partial}{\partial t} P(\mathbf{x},t) = -H_{\mathrm{FP}} P(\mathbf{x},t) \equiv D \sum_{i=1}^{N} \boldsymbol{\nabla}_i \cdot \left[\boldsymbol{\nabla}_i + \beta \boldsymbol{\nabla}_i U(\mathbf{x}) \right] P(\mathbf{x},t) \tag{A.6}$$

where H_{FP} is the *Fokker-Planck* (FP) *operator*. One can immediately realize that the stationary distribution of the Fokker-Planck equation is the Boltzmann distribution:

$$P(\mathbf{x},t) \xrightarrow[t\to\infty]{} P(\mathbf{x}) = \frac{1}{Z} e^{-\beta U(\mathbf{x})} \qquad Z = \int d\mathbf{x}\, e^{-\beta U(\mathbf{x})} \tag{A.7}$$

It should be noted that the FP equation can be written in the form of a continuity equation:

$$\frac{\partial}{\partial t} P(\mathbf{x},t) + \boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{x},t) = 0 \tag{A.8}$$

where the *FP current* is given by

$$\mathbf{J}(\mathbf{x},t) = -D \left[\boldsymbol{\nabla} + \beta \boldsymbol{\nabla} U(\mathbf{x}) \right] P(\mathbf{x},t) \tag{A.9}$$

Again, the stationary solution is the Boltzmann distribution. Using this fact, one can easily see that the FP current vanishes at equilibrium:

$$\mathbf{J}(\mathbf{x},t) \xrightarrow[t\to\infty]{} \mathbf{J}(\mathbf{x}) = 0 \tag{A.10}$$

The Fokker-Planck operator is not Hermitian, so it admits non-coincident left- and right-eigenstates:

$$H_{\mathrm{FP}} R_n(\mathbf{x}) = \lambda_n R_n(\mathbf{x}) \qquad L_n(\mathbf{x}) H_{\mathrm{FP}} = H_{FP}^{\dagger} L_n(\mathbf{x}) = \lambda_n L_n(\mathbf{x}) \tag{A.11}$$

where $H_{FP}^{\dagger}$ is the adjoint of the FP operator and right- and left-eigenstates define the same spectrum. Before going on with the spectral properties of H_{FP}, given its importance in the framework of stochastic dynamics it is worth mentioning that the adjoint of the FP operator is also known as the *backward Kolmogorov operator*:

$$H_{FP}^{\dagger} = -D(\boldsymbol{\nabla}^2 - \beta \boldsymbol{\nabla} U(\mathbf{x}) \boldsymbol{\nabla}) \tag{A.12}$$

and defines a corresponding *backward Kolmogorov equation*:

$$\frac{\partial}{\partial t} Q(\mathbf{x},t) = H_{FP}^{\dagger} Q(\mathbf{x},t) = -D(\boldsymbol{\nabla}^2 - \beta \boldsymbol{\nabla} U(\mathbf{x}) \boldsymbol{\nabla}) Q(\mathbf{x},t) \tag{A.13}$$

The spectrum of the FP operator is negative and ordered, in the sense that:

$$\lambda_0 = 0 > -|\lambda_1| > -|\lambda_2| > -|\lambda_3| > \ldots \tag{A.14}$$

Right- and left-eigenstates satisfy the following orthonormality condition:

$$\int d\mathbf{x}\; R_n(\mathbf{x}) L_m(\mathbf{x}) = \delta_{nm} \tag{A.15}$$

The zero-th eigenstates, corresponding to $\lambda_0 = 0$, show a peculiar property. In particular, it is evident that

$$R_0(\mathbf{x}) = \frac{1}{Z} e^{-\beta U(\mathbf{x})} \tag{A.16}$$

thus, for the orthonormality condition in Eq. (A.15) and from the properties of the stationary distribution described in Eq. (A.7), one has

$$L_0(\mathbf{x}) = 1 \tag{A.17}$$

The spectral properties of the Fokker-Planck operator allow one to express the formal solution of the FP equation by means of an expansion on the basis of the right-eigenstates:

$$P(\mathbf{x}, t) = \sum_{n=0}^{\infty} c_n R_n(\mathbf{x}) e^{-\lambda_n t} \qquad c_n = \int d\mathbf{x}\; L_n(\mathbf{x}) P(\mathbf{x}, 0) \tag{A.18}$$

Because of the normalization condition, the first coefficient c_0 is evidently equal to 1, so the series can be also expressed as

$$P(\mathbf{x}, t) = \frac{e^{-\beta U(\mathbf{x})}}{Z} + \sum_{n=1}^{\infty} c_n R_n(\mathbf{x}) e^{-\lambda_n t} \tag{A.19}$$

Eq. (A.19) embeds an extremely important interpretation: when starting from a distribution $P(\mathbf{x}, 0)$ different from the equilibrium one, $R_n(\mathbf{x})$ with $n > 0$ represent modes that need to be exponentially suppressed in order to actually reach equilibrium.

The FP operator can be made Hermitian by means of a non-unitary transformation [184]:

$$\mathcal{H} \equiv e^{\frac{\beta}{2} U(\mathbf{x})} H_{\mathrm{FP}} e^{-\frac{\beta}{2} U(\mathbf{x})} = -D\boldsymbol{\nabla}^2 + \frac{D\beta^2}{4} \left[(\boldsymbol{\nabla} U(\mathbf{x}))^2 - \frac{2}{\beta} \boldsymbol{\nabla}^2 U(\mathbf{x}) \right] \tag{A.20}$$

where the probability distribution becomes

$$\psi(\mathbf{x}, t) \equiv e^{-\frac{\beta}{2} U(\mathbf{x})} P(\mathbf{x}, t) \tag{A.21}$$

In this way, the Fokker-Planck equation is transformed into a Wick-rotated Schrödinger equation:

$$-\frac{\partial}{\partial t} \psi(\mathbf{x}, t) = \mathcal{H} \psi(\mathbf{x}, t) \tag{A.22}$$

This analogy is particularly useful, because it permits to easily prove that the conditional probability of starting in $\mathbf{x}_0$ at time 0 and reaching $\mathbf{x}$ at time t in the

overdamped Langevin dynamics can be expressed as a path-integral. In particular, we recall that the path-integral representing the Wick-rotated quantum propagator (or conditional probability) is provided by [329]:

$$P(\mathbf{x}, t|\mathbf{x}_0, 0) = \langle \mathbf{x}|e^{-\mathcal{H}\tau/\hbar}|\mathbf{x}_0\rangle = \mathcal{N}\int \mathcal{D}x(\tau) \exp\left[-\frac{1}{\hbar}S_E[x(\tau)]\right] \tag{A.23}$$

where $\mathcal{N}$ is a normalization factor, $\mathcal{D}x(\tau)$ is the path measure, $\hbar$ is the reduced Planck constant and $S_E[x(\tau)]$ is the Euclidean action

$$S_E[x(\tau)] = \int_0^\tau dt\ \left(\frac{m}{2}\dot{\mathbf{x}}^2 + U(\mathbf{x})\right) \equiv \int_0^\tau dt\ \mathcal{L}_E(\mathbf{x}, \dot{\mathbf{x}}) \tag{A.24}$$

defined as the time-integral of the Euclidean Lagrangian $\mathcal{L}_E$. The introduction of the effective Hamiltonian in Eq. (A.20) allows one to rewrite the conditional probability Eq. (A.23) as

$$P(\mathbf{x}, t|\mathbf{x}_0, 0) = e^{-\frac{\beta}{2}(U(\mathbf{x})-U(\mathbf{x}_0))}\mathcal{N}\int \mathcal{D}x(\tau)\ \exp\left[-\int_0^\tau dt\ \left(\frac{\dot{\mathbf{x}}^2}{4D} + V_{\text{eff}}[x(\tau)]\right)\right] \tag{A.25}$$

with

$$V_{\text{eff}} = \frac{D\beta^2}{4}\left[(\boldsymbol{\nabla}U(\mathbf{x}))^2 - \frac{2}{\beta}\boldsymbol{\nabla}^2 U(\mathbf{x})\right] \tag{A.26}$$

A property similar to Eq. (A.25) can be proved also for the underdamped Langevin equation in Eq. (A.2). It can be shown [330, 331, 332] that in this case one has

$$P(\mathbf{x}, t|\mathbf{x}_0, 0) = e^{-\beta U(\mathbf{x}_0)\int \mathcal{D}x(\tau) - S_{\text{OM}}[\mathbf{x}]} \tag{A.27}$$

where

$$S_{\text{OM}}[\mathbf{x}] = \frac{\beta}{4}\sum_{i=1}^{N}\frac{1}{\gamma_i m_i}\int_0^t d\tau\ (m_i\ddot{\mathbf{r}}_i + \gamma_i\dot{\mathbf{r}}_i + \boldsymbol{\nabla}_i U(\mathbf{x}))^2 \tag{A.28}$$

is the so-called *Onsager-Machlup action*.

As a last comment on conditional probabilities, we stress that also $P(\mathbf{x}, t|\mathbf{x}_0, 0)$ satisfies the FP equation, which in this case is known as the *forward Kolmogorov equation* [184]:

$$\frac{\partial}{\partial t}P(\mathbf{x}, t|\mathbf{x}_0, 0) = -H_{\text{FP}}P(\mathbf{x}, t|\mathbf{x}_0, 0) \tag{A.29}$$

Instead, if we consider $\mathbf{x}_0$ as the variable of $P(\mathbf{x}, t|\mathbf{x}_0, 0)$, the conditional probability satisfies the backward Kolmogorov equation [184]:

$$\frac{\partial}{\partial t}P(\mathbf{x}, t|\mathbf{x}_0, 0) = -H^\dagger_{\text{FP}}(\mathbf{x}_0)P(\mathbf{x}, t|\mathbf{x}_0, 0) \tag{A.30}$$

where

$$H^\dagger_{\text{FP}}(\mathbf{x}_0) \equiv -D(\boldsymbol{\nabla}^2_{\mathbf{x}_0} - \beta\boldsymbol{\nabla}_{\mathbf{x}_0}U(\mathbf{x}_0)\boldsymbol{\nabla}_{\mathbf{x}_0}) \tag{A.31}$$

Appendix B

Appendix to Chapter 1

B.1 The Jacobian $J[\bar{\mathbf{x}}, \xi]$ is identical to one

In this section we want to show that the integral

$$I = \int \mathcal{D}\xi(t)\, \delta\left(\xi(\tau) - \int_0^\tau d\tau' f(\xi(\tau'))\right) \tag{B.1}$$

is identical to 1, which is equivalent to say that the Jacobian $J[\bar{\mathbf{x}}, \xi]$ introduced in Eq. (1.25) is also equal to 1. A possible way to prove this statement is by discretizing the path-integral in Eq. (B.1) in N slices with $N \to \infty$:

$$\begin{aligned} I &= \lim_{N\to\infty} \left[\prod_{i=0}^{N} d\xi(t_i)\right] \prod_{k=0}^{N} \delta\left[\xi(t_k) - \Delta t \sum_{l=0}^{k-1} f(\xi(t_l))\right] \\ &= \lim_{N\to\infty} \left[\prod_{i=0}^{N} d\xi(t_i)\right] \delta\left[\xi(t_0) - \Delta t f(\xi(t_0))\right] \cdot \\ &\qquad \delta\left[\xi(t_1) - \Delta t\{f(\xi(t_0)) + f(\xi(t_1))\}\right] \cdot \ldots \\ &= \lim_{N\to\infty} \left[\prod_{i=0}^{N} d\xi(t_i)\right] \delta\left[\xi(t_0) - \Delta t f(\xi(t_0))\right] R_{N-1}(t_0, \ldots, t_N) \end{aligned} \tag{B.2}$$

where $R_{N-1}(t_0, \ldots, t_N)$ is the product of all the other $N-1$ delta terms. From the last step in Eq. (B.2) we see that we can easily integrate out the first variable $\xi(t_0)$ and substitute it with $\Delta t f(\xi(t_0))$ as required by the Dirac delta. Thus we have

$$I = \lim_{N\to\infty} \left[\prod_{i=1}^{N} d\xi(t_i)\right] \delta\left[\xi(t_1) - 2\Delta t f(\xi(t_0))\right] R_{N-2}(t_1, \ldots, t_N) \tag{B.3}$$

Again, Eq. (B.3) can be used to integrate away $\xi(t_1)$. In general, the integral can be expressed as

$$I = \lim_{N\to\infty} \left[\prod_{i=\alpha}^{N} d\xi(t_i)\right] \delta\left[\xi(t_\alpha) - \alpha\Delta t f(\xi(t_0))\right] R_{N-\alpha}(t_\alpha, \ldots, t_N) \tag{B.4}$$

for every $\alpha = 1, \ldots, N$. The integral expression in Eq. (B.4) can be used to iteratively integrate out all the $\xi(t_\alpha)$ terms up to $\xi(t_N)$, thus closing the proof.

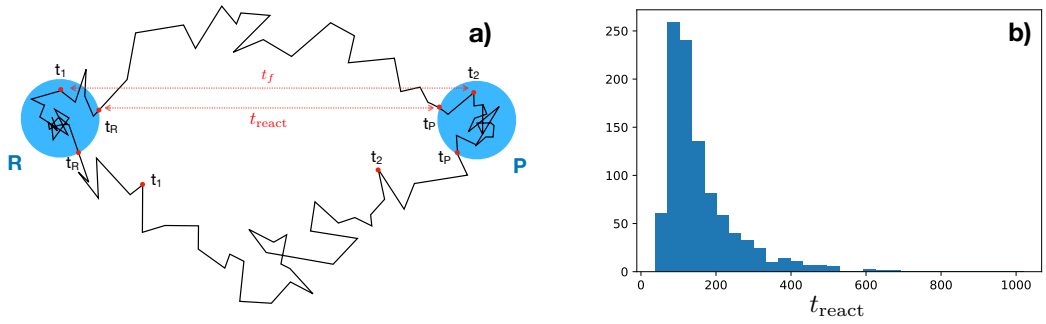

Figure B.1: Reproduced from Ref. [29], with the permission of AIP Publishing. (a) Sketch of two trajectories and their reactive component. The one on the top is retained, because $t_f > t_{\text{react}}$, while the bottom one is not; (b) histogram of the reactive times obtained from the rMD simulations.

B.2 The Average Path

In this appendix we will discuss the strategy adopted to compute the average path needed for 2-dimensional SCPS simulations. Reactant and the product states are identified as the ones around $x_R = (-1, 0)$ and $x_P = (1, 0)$ respectively, for which the potential is smaller than the threshold $U(\mathbf{x}, \mathbf{y}) = -2.5 k_B T$. The average path is defined by considering only the reactive part of each simulation: for this reason, we only retain the portion of the trajectory which lies outside the boundaries of R and P. On the other hand, the average path requires trajectories that have the same time-length to be defined. Therefore, we adopted a strategy to define an optimal duration of the pathways to be averaged over. For each trajectory, we stored $t_{\text{react}} = t_P - t_R$, the difference between times at which they enter the product and exit the reactant, respectively. This quantity represents the time that each trajectory spent in the reactive region. We computedτ and $\Delta\tau$, respectively the average and the standard deviation of the distribution of reactive times, and fixed the time window equal to $t_f = \tau + \Delta\tau$. Every trajectory that within a time t_f is not able to connect the reactant to product state is automatically discarded from the calculation of the average. Finally, from each trajectory we select only a portion that lasts from $t_1 = t_P - \frac{t_f - t_{\text{react}}}{2}$ to $t_2 = t_R + \frac{t_f - t_{\text{react}}}{2}$ and average these portions to obtain the final average path.

Appendix C

Appendix to Chapter 2

C.1 The committor function satisfies the backward Kolmogorov equation

To see why Eq. (2.9) holds (and, wlog we will do this just for the forward case), in this appendix we introduce an alternative definition of the committor function [333]:

$$q(\mathbf{x}) = \int_0^\infty dt \int_{\partial P} d\mathbf{S}(\mathbf{x}') \cdot \mathbf{J}(\mathbf{x}', t|\mathbf{x}, 0) \tag{C.1}$$

where $\mathbf{x} \in \Omega_T$, $\mathbf{x}' \in \partial P$, ∂P is the border of the product state, $\mathbf{J}(\mathbf{x}', t|\mathbf{x}, 0)$ is the Fokker-Planck current associated to the conditional probability $P(\mathbf{x}', t|\mathbf{x}, 0)$ and finally $d\mathbf{S}(\mathbf{x}')$ is the surface element in $\mathbf{x}'$, conventionally pointing inward the product state. Even though it is not immediately evident, Eq.s (2.6) and (C.1) define the same function. Indeed, we see that the integral of the Fokker-Planck current in Eq. (C.1) provides the crossing rate through the product surface, $k_{\partial P}(\mathbf{x}, t)$. The integral $\int_0^\infty dt\, k_{\partial P}(\mathbf{x}, t)$ provides the fraction of trajectories, starting at $\mathbf{x}$, that leave the transition state for the first time by passing through ∂P. The fact that they are leaving Ω_T for the first time comes from the boundary conditions on the states borders. This is indeed identical to the definition of the forward committor in Eq. (2.6).

Let us now apply the backward-Kolmogorov operator $H^\dagger_{\mathrm{FP}}(\mathbf{x})$ to Eq. (C.1):

$$\begin{aligned} H^\dagger_{\mathrm{FP}}(\mathbf{x}) q(\mathbf{x}) &= \int_0^\infty \int_{\partial P} d\mathbf{S}(\mathbf{x}') \cdot H^\dagger_{\mathrm{FP}}(\mathbf{x}) \mathbf{J}(\mathbf{x}', t|\mathbf{x}, 0) \\ &= \int_0^\infty dt\, \frac{\partial}{\partial t} \int_{\partial P} d\mathbf{S}(\mathbf{x}') \cdot \mathbf{J}(\mathbf{x}', t|\mathbf{x}, 0) \\ &= \int_{\partial P} d\mathbf{S}(\mathbf{x}') \cdot \mathbf{J}(\mathbf{x}', t|\mathbf{x}, 0)\big|_{t=0}^{t=\infty} = 0 \end{aligned} \tag{C.2}$$

where in the first equality we used the definition of the conditional current, in the second equality the fundamental theorem of calculus and in the last equality we used

the fact that at $t = \infty$ all the particles have been removed from the transition region while at time $t = 0$ the probability density is exactly localized in $\mathbf{x}$.

C.2 Reactive Current in the SCR

In this appendix we will prove explicitly the equality presented in Eq. (2.86):

$$
\begin{aligned}
\mathbf{J}_{\mathrm{SCR}}(\mathbf{x}) &= -\frac{D}{t_f-\tau_0}\int_{\tau_0}^{t_f} dt\; Q^{(R)}(\mathbf{x},t_f-t)\left(\overrightarrow{\boldsymbol{\nabla}}-\overleftarrow{\boldsymbol{\nabla}}+\beta\boldsymbol{\nabla}U(\mathbf{x})\right)P^{(P)}(\mathbf{x},t)\\
&= -\frac{D}{t_f-\tau_0}\int_{\tau_0}^{t_f} dt\;\Big[Q^{(R)}(\mathbf{x},t_f-t)\boldsymbol{\nabla}P^{(P)}(\mathbf{x},t)\\
&\qquad\qquad -P^{(P)}(\mathbf{x},t)\boldsymbol{\nabla}Q^{(R)}(\mathbf{x},t_f-t)\\
&\qquad\qquad +\beta Q^{(R)}(\mathbf{x},t_f-t)\boldsymbol{\nabla}U(\mathbf{x})P^{(P)}(\mathbf{x},t)\Big]\\
&= -\frac{D}{t_f-\tau_0}\int_{\tau_0}^{t_f} dt\;\left[Q^{(R)}(\mathbf{x},t_f-t)\boldsymbol{\nabla}\left(\frac{e^{-\beta U(\mathbf{x})}}{Z_R}Q^{(P)}(\mathbf{x},t)\right)\right.\\
&\qquad\qquad -\frac{e^{-\beta U(\mathbf{x})}}{Z_R}Q^{(P)}(\mathbf{x},t)\boldsymbol{\nabla}Q^{(R)}(\mathbf{x},t_f-t)\\
&\qquad\qquad \left.+\beta Q^{(R)}(\mathbf{x},t_f-t)\boldsymbol{\nabla}U(\mathbf{x})\frac{e^{-\beta U(\mathbf{x})}}{Z_R}Q^{(P)}(\mathbf{x},t)\right]\\
&= -\frac{D}{t_f-\tau_0}\int_{\tau_0}^{t_f} dt\;\left[-\beta\boldsymbol{\nabla}U(\mathbf{x})Q^{(R)}(\mathbf{x},t_f-t)\frac{e^{-\beta U(\mathbf{x})}}{Z_R}Q^{(P)}(\mathbf{x},t)\right.\\
&\qquad\qquad +Q^{(R)}(\mathbf{x},t_f-t)\frac{e^{-\beta U(\mathbf{x})}}{Z_R}\boldsymbol{\nabla}Q^{(P)}(\mathbf{x},t)\\
&\qquad\qquad -\frac{e^{-\beta U(\mathbf{x})}}{Z_R}Q^{(P)}(\mathbf{x},t)\boldsymbol{\nabla}Q^{(R)}(\mathbf{x},t_f-t)\\
&\qquad\qquad \left.+\beta Q^{(R)}(\mathbf{x},t_f-t)\boldsymbol{\nabla}U(\mathbf{x})\frac{e^{-\beta U(\mathbf{x})}}{Z_R}Q^{(P)}(\mathbf{x},t)\right]\\
&= -\frac{D}{t_f-\tau_0}\frac{e^{-\beta U(\mathbf{x})}}{Z_R}\int_{\tau_0}^{t_f} dt\;\Big[Q^{(R)}(\mathbf{x},t_f-t)\boldsymbol{\nabla}Q^{(P)}(\mathbf{x},t)\\
&\qquad\qquad -Q^{(P)}(\mathbf{x},t)\boldsymbol{\nabla}Q^{(R)}(\mathbf{x},t_f-t)\Big]\\
&= -D\frac{e^{-\beta U(\mathbf{x})}}{Z_R}\left(\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})\boldsymbol{\nabla}(1-\mathbf{q}_{\mathrm{SCR}}(\mathbf{x}))-(1-\mathbf{q}_{\mathrm{SCR}}(\mathbf{x}))\boldsymbol{\nabla}\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})\right)\\
&= -D\frac{e^{-\beta U(\mathbf{x})}}{Z_R}\left(-\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})\boldsymbol{\nabla}\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})-\boldsymbol{\nabla}\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})+\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})\boldsymbol{\nabla}\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})\right)\\
&= D\frac{e^{-\beta U(\mathbf{x})}}{Z_R}\boldsymbol{\nabla}\mathbf{q}_{\mathrm{SCR}}(\mathbf{x})
\end{aligned}
$$

which provides the expected result.

C.3 ∂R crossing contributions in $q^+_{\mathbf{SCR}}$ are negligible

In this appendix we want to prove that, even if q^+_{SCR} is defined by imposing absorbing boundary conditions *only* on ∂P, the contributions to the committor function coming from trajectories that enter R and escape it are negligible. To do so we consider two points, separated by a long diffusion distance, $\mathbf{x}_f$ and $\mathbf{x}_0$ and we call $P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ the conditional probability to start in $\mathbf{x}_0$ at time $t = 0$ and reach $\mathbf{x}_f$ at time $t = t_f$. Let us also consider a generic volume S enclosed by a surface ∂S: $P^{(S)}(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ represents the same conditional probability, but we explicitly impose absorbing boundary conditions over ∂S on it. Finally, we call $G(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ the probability density associated to a transition from $\mathbf{x}_0$ to $\mathbf{x}_f$ in time $t = t_f$, conditioned to touch ∂S at least once. The three objects are trivially linked by

$$P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) = P^{(S)}(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) + G(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) \tag{C.3}$$

It is also true that $P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ is equal to the sum between the conditional probability to perform the transition $\mathbf{x}_0 \to \mathbf{x}_f$ without passing through ∂S and the conditional probability to reach a point $\mathbf{x}' \in \partial S$ from $\mathbf{x}_0$ at time $t = t'$ and then reach $\mathbf{x}_f$ from it in a time $t = t_f - t'$. In general, the number of re-crossing events of ∂S might be arbitrarily large. Using this argument, we can rewrite $P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ as

$$\begin{aligned} P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) &= P^{(S)}(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) + \int_0^t dt' \int_{\partial S} d\mathbf{x}'\, G(\mathbf{x}_f, t_f|\mathbf{x}', t') P^{(S)}(\mathbf{x}', t'|\mathbf{x}_0, 0) \\ &= P^{(S)}(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) + \\ &+ \int_0^t dt' \int_{\partial S} d\mathbf{x}' \left[P(\mathbf{x}_f, t_f|\mathbf{x}', t') - P^{(S)}(\mathbf{x}_f, t_f|\mathbf{x}', t')\right] P^{(S)}(\mathbf{x}', t'|\mathbf{x}_0, 0) \end{aligned} \tag{C.4}$$

where in the second equality we used Eq. (C.3). Eq. (C.4) is a Dyson series that defines $P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ as an expansion over terms proportional to $P^{(S)}(\mathbf{x}', t'|\mathbf{x}_0, 0)$. Now we can use the assumption that $\mathbf{x}_0$ and $\mathbf{x}_f$ are separated by a high diffusion distance: this means that $P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) \ll 1$ for every choice of t_f in the SCR. For this reason, $P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0)$ can be approximated with the lowest order term of the Dyson series, thus providing

$$P(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) \sim P^{(S)}(\mathbf{x}_f, t_f|\mathbf{x}_0, 0) \tag{C.5}$$

If we choose $\partial S \equiv \partial R$ we can use Eq. (C.5) to state that the contributions coming from trajectories that fluctuate within R and then leave it are suppressed at first order in the SCR expansion parameters.

C.4 Iso-σ and iso-q_{SCR} hyperplanes coincide

In this appendix we are going to show that iso-σ and iso-q hypersurfaces are parallel. To do so, we exploit a path-integral representation of the distributions $Q^{(R)}$ and $P^{(P)}$ to show that the reactive current in $\mathbf{x}$ in the SCR can be expressed as a function of the average velocity in of the transition paths in $\mathbf{x}$. Let us see how to obtain it. Following the discussion in Ref. [334], we can re-write the reactive current as

$$\mathbf{J}_{\mathrm{SCR}}(\mathbf{x}) = \frac{1}{t_f - t_0}\int_{t_0}^{t_f} dt\, \frac{1}{2}\left(\langle \mathbf{v}(\mathbf{x},t)\rangle^{(P)} + \langle \mathbf{v}(\mathbf{x},t)\rangle^{(R)}\right) Q^{(R)}(\mathbf{x}, t_f - t) P^{(P)}(\mathbf{x},t) \tag{C.6}$$

In Eq. (C.6), $\langle \mathbf{v}(\mathbf{x},t)\rangle^{(P)}$ and $\langle \mathbf{v}(\mathbf{x},t)\rangle^{(R)}$ denote, respectively, the average velocity of the pathways at point $\mathbf{x}$ at time t, and are obtained by resorting to the path-integral expression of $\boldsymbol{\nabla} Q^{(R)}(\mathbf{x}, t_f - t)$ and $\boldsymbol{\nabla} P^{(P)}(\mathbf{x}, t_f)$:

$$\langle \mathbf{v}(\mathbf{x},t)\rangle^{(P)} \equiv \frac{\int_\Omega d\mathbf{x}_i\, \rho(\mathbf{x}_i) \int_{\mathbf{x}_i}^{\mathbf{x}} \mathcal{D}\mathbf{x}\, \dot{\mathbf{x}}(t) \exp\left[-\int_0^t d\tau\, \left(\mathcal{L}_{\mathrm{OM}}[\mathbf{x}] + \Omega_R[\mathbf{x}]\right)\right]}{\int_\Omega d\mathbf{x}_i\, \rho(\mathbf{x}_i) \int_{\mathbf{x}_i}^{\mathbf{x}} \mathcal{D}\mathbf{x}\, \exp\left[-\int_0^t d\tau\, \left(\mathcal{L}_{\mathrm{OM}}[\mathbf{x}] + \Omega_P[\mathbf{x}]\right)\right]} \tag{C.7}$$

$$\langle \mathbf{v}(\mathbf{x},t)\rangle^{(R)} \equiv \frac{\int_\Omega d\mathbf{x}_f\, h_P[\mathbf{x}_f] \int_{\mathbf{x}}^{\mathbf{x}_f} \mathcal{D}\mathbf{x}\, \dot{\mathbf{x}}(t) \exp\left[-\int_0^t d\tau\, \left(\mathcal{L}_{\mathrm{OM}}[\mathbf{x}] + \Omega_R[\mathbf{x}]\right)\right]}{\int_\Omega d\mathbf{x}_f\, h_P[\mathbf{x}_f] \int_{\mathbf{x}}^{\mathbf{x}_f} \mathcal{D}\mathbf{x}\, \exp\left[-\int_0^t d\tau\, \left(\mathcal{L}_{\mathrm{OM}}[\mathbf{x}] + \Omega_R[\mathbf{x}]\right)\right]} \tag{C.8}$$

where

$$\mathcal{L}_{\mathrm{OM}}[\mathbf{x}] = \frac{1}{4D}\left(\dot{\mathbf{x}} + \beta D \boldsymbol{\nabla} U(\mathbf{x})\right)^2 \tag{C.9}$$

and $\Omega_A[\mathbf{x}]$ functions are defined as

$$\Omega_A[\mathbf{x}] = \begin{cases} \infty & \text{if } \mathbf{x} \in A \subset \Omega \\ 0 & \text{otherwise} \end{cases} \tag{C.10}$$

Functions in Eq. (C.10), when A is substituted with R and P, effectively define the absorbing boundary conditions in the corresponding path-integrals. If the times t_0 and t_f are chosen within the SCR, the current becomes virtually time-independent and Eq. (C.6) can be rewritten as

$$\mathbf{J}_{\mathrm{SCR}}(\mathbf{x}) \sim \langle \mathbf{v}(\mathbf{x})\rangle m_{\mathrm{SCR}}(\mathbf{x}) \tag{C.11}$$

where

$$\langle \mathbf{v}(\mathbf{x})\rangle = \frac{1}{2}\left(\langle \mathbf{v}(\mathbf{x})\rangle^{(P)} + \langle \mathbf{v}(\mathbf{x})\rangle^{(R)}\right) \tag{C.12}$$

If we choose $\mathbf{x}$ on the SCPS average path, the reactive current computed in $\mathbf{x}$ is parallel to the time derivative of $\langle \mathbf{x}\rangle$, i.e. the average position which enters the definition of $s_\lambda(\mathbf{x})$ and $\sigma(\mathbf{x})$ (see Eq.s (1.49) and (1.51)). Since Eq. (2.23) prescribes that the reactive current is parallel to $\boldsymbol{\nabla} q(\mathbf{x})$, we can finally conclude that iso-σ and iso-q hypersurfaces coincide up to re-labeling.

C.5 Conditional Langevin Dynamics

In this appendix we will prove the equality in Eq. (2.97). In the following, for sake of notational simplicity we will call $Q = Q^{(R)}(\mathbf{x}, t)$, $P = P^{(P)}(\mathbf{x}, t)$, $U = U(\mathbf{x})$ and $\mathbb{P} = \mathbb{P}(\mathbf{x}, t)$:

$$\begin{aligned}
\frac{\partial \mathbb{P}}{\partial t} &= \frac{\partial}{\partial t}(QP) \\
&= P\frac{\partial Q}{\partial t} + Q\frac{\partial P}{\partial t} \\
&= PH_{\mathrm{FP}}^{\dagger}Q + QH_{\mathrm{FP}}P \\
&= P(\boldsymbol{\nabla}^2 - \beta\boldsymbol{\nabla}U\boldsymbol{\nabla})Q + Q\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U)P \\
&= P\boldsymbol{\nabla}^2 Q - P\beta\boldsymbol{\nabla}U\boldsymbol{\nabla}Q + Q\boldsymbol{\nabla}^2 P + Q\beta\boldsymbol{\nabla}U\boldsymbol{\nabla}P + Q\beta\boldsymbol{\nabla}^2 UP \\
&= \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}(QP) - 2\boldsymbol{\nabla}Q\boldsymbol{\nabla}P - 2P\beta\boldsymbol{\nabla}U\boldsymbol{\nabla}Q + P\beta\boldsymbol{\nabla}U\boldsymbol{\nabla}Q \\
&\quad + Q\beta\boldsymbol{\nabla}U\boldsymbol{\nabla}P + Q\beta\boldsymbol{\nabla}^2 UP \\
&= \boldsymbol{\nabla}\cdot[\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U](QP) - 2\boldsymbol{\nabla}Q\boldsymbol{\nabla}P - 2P\beta\boldsymbol{\nabla}U\boldsymbol{\nabla}Q \\
&= \boldsymbol{\nabla}\cdot[\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U](QP) - 2\boldsymbol{\nabla}Q\boldsymbol{\nabla}P - 2P\beta\boldsymbol{\nabla}^2 Q \\
&= \boldsymbol{\nabla}\cdot[\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U](QP) - 2\boldsymbol{\nabla}\cdot(P\boldsymbol{\nabla}Q) \\
&= \boldsymbol{\nabla}\cdot[\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U](QP) - 2\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\nabla}Q}{Q}(PQ)\right) \\
&= \boldsymbol{\nabla}\cdot[\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U - 2\boldsymbol{\nabla}Q](QP) \\
&= \boldsymbol{\nabla}\cdot[\boldsymbol{\nabla} + \beta\boldsymbol{\nabla}U - 2\boldsymbol{\nabla}\log Q]\,\mathbb{P}
\end{aligned}$$

that proves our statement.

C.6 irMD Samples the Boltzmann Distribution Restricted to the Transition Region

In this appendix we will show that the non-equilibrium process proposed in Eq. (2.103) can be used to sample the Boltzmann distribution in the transition region. To do set, let us set $\xi(x) = 1$ in Eq. (2.101) for the moment. Let us use Eq. (2.103) to generate many reactive trajectories that are stopped once they overcome the boundaries of the product state. Let us consider only points $\mathbf{x}$ in the transition region that are far enough from the reactant and the product, so that transitions from the product to $\mathbf{x}$ and transition from $\mathbf{x}$ to the product can be neglected. In this case, the probability density obtained using irMD satisfies a master equation of the form

$$\frac{\partial}{\partial t}P(\mathbf{x}, t) = \int d\mathbf{y}\,[T(\mathbf{x}|\mathbf{y})P(\mathbf{y}, t) - T(\mathbf{y}|\mathbf{x})P(\mathbf{x}, t)] \tag{C.13}$$

where $T(\mathbf{y}|\mathbf{y})$ is the transition rate, in the irMD dynamics, between $\mathbf{x}$ and $\mathbf{y}$. It can be expressed as follows:

$$T(\mathbf{y}|\mathbf{x}) = T_0(\mathbf{y}|\mathbf{x})\theta(q(\mathbf{y}) - q(\mathbf{x})) + T_B(\mathbf{y}|\mathbf{x})\theta(q(\mathbf{x}) - q(\mathbf{y})) \tag{C.14}$$

where $T_0(\mathbf{y}|\mathbf{x})$ is the transition rate of the unbiased dynamics and $T_B(\mathbf{y}|\mathbf{x})$ is instead the one of the dynamics in the presence of the bias. Eq. (C.14) means that every time the dynamics backtracks along the committor, the rate is provided by the original Langevin dynamics, while if the dynamics proceeds by increasing the committor the rate is provided by the biased dynamics. We can associate a Fokker-Planck equation to the master equation in Eq. (A.6) by computing the corresponding Kramers-Moyal coefficients [184] (here discussed only in one dimension for sake of simplicity):

$$\begin{aligned} a_1(x) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int dy\ (y-x)T(y|x) \\ &= \lim_{\Delta t\to 0,\ x(t)\to x}\frac{1}{\Delta t}\left[\langle (x(t+\Delta t)-x(t))\,\theta(q(x(t+\Delta t))-q(x(t))\rangle_0\right. \\ &\left.+\ \langle (x(t+\Delta t)-x(t))\,\theta(q(x(t))-q(x(t+\Delta t))\rangle_B\right] \end{aligned} \tag{C.15}$$

$$\begin{aligned} a_2(x) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int dy\ (y-x)^2T(y|x) \\ &= \lim_{\Delta t\to 0,\ x(t)\to x}\frac{1}{\Delta t}\left[\left\langle (x(t+\Delta t)-x(t))^2\,\theta(q(x(t+\Delta t))-q(x(t))\right\rangle_0\right. \\ &\left.+\ \left\langle (x(t+\Delta t)-x(t))^2\,\theta(q(x(t))-q(x(t+\Delta t))\right\rangle_B\right] \end{aligned} \tag{C.16}$$

$\langle\cdot\rangle_0$ and $\langle\cdot\rangle_B$ denote, respectively, the averages over the occurrences of the random noise in the unbiased and the biased dynamics. The limits in Eq.s (C.15) and (C.16) can be computed explicitly:

$$\begin{aligned} a_1(x) &= -\frac{D}{k_{\mathrm{B}}T}\left(\nabla U(x) - \frac{k_R}{2}\nabla q(x)\right) \\ a_2(x) &= 2D \end{aligned} \tag{C.17}$$

so that the corresponding Fokker-Planck equation is given by

$$\frac{\partial}{\partial t}P(x,t) = D\nabla\left(\nabla + \beta U(x) - \frac{k_R}{2}\nabla q(x)\right)P(x,t) \tag{C.18}$$

Eq. (C.18) does not depend on any specific choice of the constant k_R. It is easy to see that the Boltzmann distribution is the stationary distribution of Eq. (C.18): we stress however that the stationary distribution should not be looked for as the stationary distribution of an equilibrium process, but rather as the one of a non-equilibrium process where trajectories are shoot from the border of the reactive basin and are stopped once they overcome the border of the product one. The corresponding current is given by

$$J_{\mathrm{irMD}}(x) = -D\nabla\left(\nabla + \beta U(x) - \frac{k_R}{2}\nabla q(x)\right)\frac{e^{-\beta U(x)}}{Z} = \frac{k_R}{2}\nabla q(x)\frac{e^{-\beta U(x)}}{Z} \tag{C.19}$$

which is proportional to the one of TPT. Because Eq. (C.18) does not depend on the choice of k_R, we can consider the limit $k_R \to \infty$. In this case, the specific choice of the function $\xi(x)$ becomes irrelevant because backtracks along the committor occur with negligible rates. For this reason, provided that k_R is chosen big enough, irMD exactly samples the Boltzmann distribution restricted to the transition state, irregardless of the choice of $\xi(x)$.

C.7 Calculation of the committor from unbiased Langevin dynamics

In this appendix we will show how we computed the reference committor function appearing in Fig. 2.10. Region $\Omega = [-1.5, 1.5] \times [-0.5, 2.5]$ was split into 100×100 bins. From each bin, we sampled 2×10^5 points , each of which were used as starting points for Langevin dynamics simulations. Langevin dynamics simulations consisted of $4 \times 10^3 dt/\gamma$ time steps and were performed using $dt = 0.02$, $\gamma = 1$ and $k_\mathrm{B}T = 0.15$. A trajectory was considered productive if, at some point in the simulation, it reached $U(x, y) < -2.5$ $\mathrm{k}_B\mathrm{T}$ and $\sqrt{(x - x_P)^2 + (y - y_P)^2} < 0.02$. Every trajectory meeting this condition before the end of the simulation was stopped. The value of the committor function in bin (i, j) was computed from these trajectories as

$$q(i, j) = \frac{\text{Number of productive trajectories started from } (i, j)}{\text{Total number of simulations started from } (i, j)} \tag{C.20}$$

Appendix D

Brief Chemical Vocabulary

π-MO (or simply π-orbital) molecular orbital formed by the overlap between two adjacent p_x or p_y atomic p-orbitals. It shares the same nodal structure of the original p-orbitals, with vanishing electron density in the plane connecting the bonded nuclei. π-orbitals can be *bonding* or *antibonding* (called in this case π^*-orbitals). Antibonding molecular orbitals weaken the bond between two atoms and help increasing the molecule's energy. They show an additional nodal plane with respect to the corresponding bonding oribital, located in the region between the bonded nuclei.

Aromaticity describes a cyclic, flat molecule with a ring of resonance bonds. The peculiar disposition of p-bonds makes the molecule *conjugated*, meaning that electrons are delocalized among all the bonds, which form two π-orbitals surrounding the whole molecule (see Fig. D.1 (a)). This makes aromatic molecules particularly stable and little reactive.

Heterocyclic compound cyclic molecule with at least two different atomic species forming its ring.

Indole aromatic heterocyclic compound with a bicyclic structure: the first ring is a benzene and the second one is a pyrrole (see Fig. D.1 (b)).

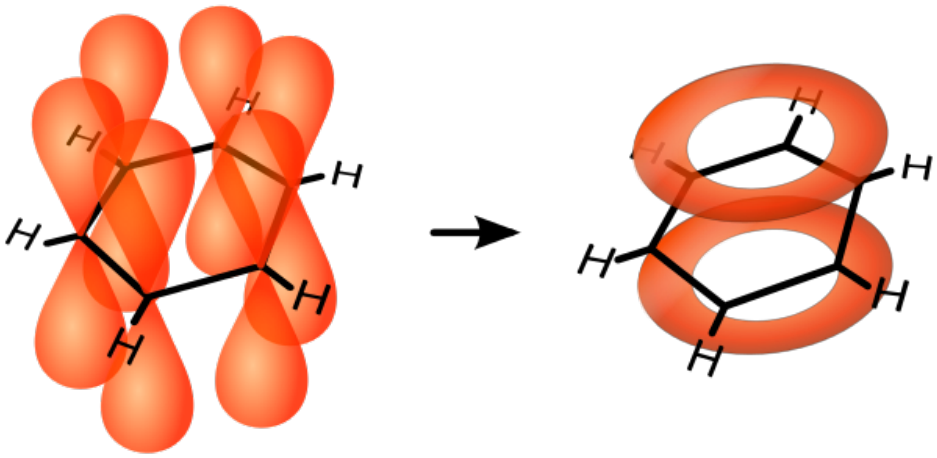

Figure D.1: (a) Molecular Orbital structure of benzene; (b) structure of an indole.

Appendix E

Appendix to Chapter 4

E.1 Kolmogorov-Smirnov test

The statistical significance of the probability distributions $P_{\text{mutant}}(Q)$ discussed in section 4.3.1 sampled by BF simulations can be assessed by a two-sample Kolmogorov-Smirnov (KS) test [319]. In this test we compared the values of Q obtained by the simulations of WT protein and all its corresponding mutations. The null hypothesis of this test is that the values of Q obtained running BF simulations with different protein variants are sampled from the same distribution. Because the number of simulations for each variant are different, before applying the KS test we boostrapped from each set of Qs a number of points which is equal to half the dimension of the smallest dataset, i.e. the one containing only 29 trajectories. Moreover, for each dataset we repeated the bootstrapping procedure ten times and performed the KS test again. In all the KS tests we performed, the null hypothesis was rejected with a p-value $p < 2 \times 10^{-16}$. This result allows us to conclude that each $P_{\text{mutant}}(Q)$ is statistically independent from all the other ones.

E.2 Topological Inconsistencies in Folded Conformers

In this appendix we show an examples of the strand cross-overs discussed in section 4.1. In Fig. E.1, s4C and s3C strands cross in an opposite way with respect to the crystal structure.

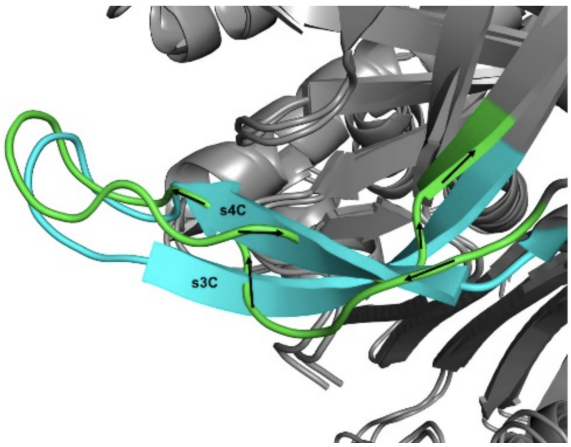

Figure E.1: From [25]. Non-native topology in the gate region of A1AT. The crystal structure of A1AT is shown in cyan, while the final result of a BF simulations is shown in grey.

Bibliography

[1] J. M. Berg, J. L. Tymoczko, and L. Stryer, "Biochemistry, ; w. h," 2002.

[2] C. B. ANFINSEN, E. HABER, M. SELA, and F. H. WHITE, "The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain.," *Proceedings of the National Academy of Sciences*, vol. 47, no. 9, pp. 1309–1314, 1961.

[3] K. A. Dill and J. L. MacCallum, "The protein-folding problem, 50 years on.," *Science*, vol. 338, no. 6110, pp. 1042–1046, 2012.

[4] S. W. Englander and L. Mayne, "The case for defined protein folding pathways," *Proceedings of the National Academy of Sciences*, vol. 114, no. 31, pp. 8253–8258, 2017.

[5] S. W. Englander and L. Mayne, "Reply to eaton and wolynes: How do proteins fold?," *Proceedings of the National Academy of Sciences*, vol. 114, no. 46, pp. E9761–E9762, 2017.

[6] R. M. Abaskharon and F. Gai, "Meandering Down the Energy Landscape of Protein Folding: Are We There Yet?," *Biophys. J.*, vol. 110, no. 9, pp. 1924–1932, 2016.

[7] M. Karplus and J. A. McCammon, "Molecular dynamics simulations of biomolecules," *Nature Structural and Molecular Biology*, vol. 9, no. 9, p. 646, 2002.

[8] M. Shirts and V. S. Pande, "Screen savers of the world unite!," *Science*, vol. 290, no. 5498, pp. 1903–1904, 2000.

[9] D. E. Shaw, R. O. Dror, J. K. Salmon, J. Grossman, K. M. Mackenzie, J. A. Bank, C. Young, M. M. Deneroff, B. Batson, K. J. Bowers, *et al.*, "Millisecond-scale molecular dynamics simulations on anton," in *Proceedings of the conference on high performance computing networking, storage and analysis*, p. 39, ACM, 2009.

[10] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers, "Atomic-level characterization of the structural dynamics of proteins," *Science*, vol. 330, no. 6002, pp. 341–346, 2010.

[11] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, "How fast-folding proteins fold," *Science*, vol. 334, no. 6055, pp. 517–520, 2011.

[12] V. A. Voelz, G. R. Bowman, K. Beauchamp, and V. S. Pande, "Molecular simulation of ab initio protein folding for a millisecond folder ntl9 (1- 39)," *Journal of the American Chemical Society*, vol. 132, no. 5, pp. 1526–1528, 2010.

[13] G. R. Bowman, V. A. Voelz, and V. S. Pande, "Atomistic folding simulations of the five-helix bundle protein λ6- 85," *Journal of the American Chemical Society*, vol. 133, no. 4, pp. 664–667, 2010.

[14] V. A. Voelz, M. Jäger, S. Yao, Y. Chen, L. Zhu, S. A. Waldauer, G. R. Bowman, M. Friedrichs, O. Bakajin, L. J. Lapidus, *et al.*, "Slow unfolded-state structuring in acyl-coa binding protein folding revealed by simulation and experiment," *Journal of the American Chemical Society*, vol. 134, no. 30, pp. 12565–12577, 2012.

[15] T. Maximova, R. Moffatt, B. Ma, R. Nussinov, and A. Shehu, "Principles and overview of sampling methods for modeling macromolecular structure and dynamics," *PLoS computational biology*, vol. 12, no. 4, p. e1004619, 2016.

[16] Y. Sugita and Y. Okamoto, "Replica-exchange molecular dynamics method for protein folding," *Chemical physics letters*, vol. 314, no. 1-2, pp. 141–151, 1999.

[17] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, "Transition path sampling: Throwing ropes over rough mountain passes, in the dark," *Annual review of physical chemistry*, vol. 53, no. 1, pp. 291–318, 2002.

[18] E. Vanden-Eijnden and M. Venturoli, "Revisiting the finite temperature string method for the calculation of reaction tubes and free energies," *The Journal of chemical physics*, vol. 130, no. 19, p. 05B605, 2009.

[19] M. Cameron and E. Vanden-Eijnden, "Flows in complex networks: theory, algorithms, and application to lennard–jones cluster rearrangement," *Journal of Statistical Physics*, vol. 156, no. 3, pp. 427–454, 2014.

[20] R. C. Bernardi, M. C. Melo, and K. Schulten, "Enhanced sampling techniques in molecular dynamics simulations of biological systems," *Biochimica et Biophysica Acta (BBA)-General Subjects*, vol. 1850, no. 5, pp. 872–877, 2015.

[21] C. Camilloni, R. A. Broglia, and G. Tiana, "Hierarchy of folding and unfolding events of protein g, ci2, and acbp from explicit-solvent simulations," *The Journal of Chemical Physics*, vol. 134, no. 4, p. 045105, 2011.

[22] G. Tiana and C. Camilloni, "Ratcheted molecular-dynamics simulations identify efficiently the transition state of protein folding," *The Journal of Chemical Physics*, vol. 137, no. 23, p. 235101, 2012.

[23] S. a Beccara, L. Fant, and P. Faccioli, "Variational scheme to compute protein reaction pathways using atomistic force fields with explicit solvent," *Phys. Rev. Lett.*, vol. 114, p. 098103, Mar 2015.

[24] A. Ianeselli, S. Orioli, G. Spagnolli, P. Faccioli, L. Cupellini, S. Jurinovich, and B. Mennucci, "Atomic detail of protein folding revealed by an ab initio reappraisal of circular dichroism," *Journal of the American Chemical Society*, vol. 140, no. 10, pp. 3674–3682, 2018. PMID: 29473417.

[25] F. Wang, S. Orioli, A. Ianeselli, G. Spagnolli, S. a Beccara, A. Gershenson, P. Faccioli, and P. L. Wintrode, "All-atom simulations reveal how single-point mutations promote serpin misfolding," *Biophysical Journal*, vol. 114, no. 9, pp. 2083 – 2094, 2018.

[26] F. Wang, G. Cazzolli, P. Wintrode, and P. Faccioli, "Folding mechanism of proteins im7 and im9: Insight from all-atom simulations in implicit and explicit solvent," *The Journal of Physical Chemistry B*, vol. 120, no. 35, pp. 9297–9307, 2016. PMID: 27532482.

[27] G. Cazzolli, F. Wang, S. a Beccara, A. Gershenson, P. Faccioli, and P. L. Wintrode, "Serpin latency transition at atomic resolution," *Proceedings of the National Academy of Sciences*, vol. 111, no. 43, pp. 15414–15419, 2014.

[28] S. Orioli, S. a Beccara, and P. Faccioli, "Self consistent calculation of protein folding pathways," *Preprinted at arXiv:*, 2017.

[29] G. Bartolucci, S. Orioli, and P. Faccioli, "Transition path theory from biased simulations," *The Journal of chemical physics*, vol. 149, no. 7, p. 072336, 2018.

[30] E. Weinan and E. Vanden-Eijnden, "Towards a theory of transition paths," *Journal of statistical physics*, vol. 123, no. 3, p. 503, 2006.

[31] E. Vanden-Eijnden, "Transition path theory," in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, pp. 453–493, Springer, 2006.

[32] M. Nakao, K. Maki, M. Arai, T. Koshiba, K. Nitta, and K. Kuwajima, "Characterization of kinetic folding intermediates of recombinant canine milk lysozyme by stopped-flow circular dichroism," *Biochemistry*, vol. 44, no. 17, pp. 6685–6692, 2005.

[33] N. Berova, P. L. Polavarapu, K. Nakanishi, and R. W. Woody, *Comprehensive Chiroptical Spectroscopy*. Hoboken, NJ: Wiley, 2012.

[34] R. W. Woody and K. A. Dunker, "Aromatic and cystine side-chain circular dichroism in proteins," in *Circular Dichroism and the Conformational Analysis of Biomolecules* (J. T. Yang and G. D. Fasman, eds.), ch. 4, pp. 109–157, New York, NY: Plenum Press, 1996.

[35] S. Jurinovich, G. Pescitelli, L. Di Bari, and B. Mennucci, "A TDDFT/MMPol/PCM model for the simulation of exciton-coupled circular dichroism spectra.," *Phys. Chem. Chem. Phys.*, vol. 16, pp. 16407–18, 2014.

[36] D. Padula, S. Jurinovich, L. Di Bari, and B. Mennucci, "Simulation of Electronic Circular Dichroism of Nucleic Acids: From the Structure to the Spectrum," *Chem. - A Eur. J.*, vol. 22, no. 47, pp. 17011–17019, 2016.

[37] F. Segatta, L. Cupellini, S. Jurinovich, S. Mukamel, M. Dapor, S. Taioli, M. Garavelli, and B. Mennucci, "A Quantum Chemical Interpretation of Two-Dimensional Electronic Spectroscopy of Light-Harvesting Complexes," *Journal of the American Chemical Society*, vol. 139, no. 22, pp. 7558–7567, 2017.

[38] R. W. Carrell and D. A. Lomas, "Conformational disease," *The Lancet*, vol. 350, no. 9071, pp. 134 – 138, 1997.

[39] S. M. Janciauskiene, R. Bals, R. Koczulla, C. Vogelmeier, T. Köhnlein, and T. Welte, "The discovery ofα1-antitrypsin and its role in health and disease," *Respiratory Medicine*, vol. 105, no. 8, pp. 1129 – 1139, 2011.

[40] B. H. D. van der Spoel, E. Lindahl and the GROMACS development team, "Gromacs user manual version 4.6.5." www.gromacs.org, April 2014.

[41] B. Hess and C. Kutzner and D. van der Spoel and E. Lindahl, "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," *Journal of Chemical Theory and Computation*, vol. 4(3), 2008.

[42] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, and V. S. Pande, "Mdtraj: a modern open library for the analysis of molecular dynamics trajectories," *Biophysical journal*, vol. 109, no. 8, pp. 1528–1532, 2015.

[43] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing In Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[44] T. Williams, C. Kelley, and many others, "Gnuplot 4.6: an interactive plotting program." `http://gnuplot.sourceforge.net/`, April 2013.

[45] Schrödinger, LLC, "The PyMOL molecular graphics system, version 1.8." November 2015.

[46] W. Humphrey, A. Dalke, and K. Schulten, "Vmd: Visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33 – 38, 1996.

[47] L. Brocchieri and S. Karlin, "Protein length in eukaryotic and prokaryotic proteomes," *Nucleic acids research*, vol. 33, no. 10, pp. 3390–3400, 2005.

[48] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1-2, pp. 19 – 25, 2015.

[49] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with namd," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802.

[50] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande, "Openmm 7: Rapid development of high performance algorithms for molecular dynamics," *PLOS Computational Biology*, vol. 13, pp. 1–17, 07 2017.

[51] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "Plumed 2: New Feathers for an Old Bird," *Computer Physics Communications*, vol. 185, 2014.

[52] J. M. Bello-Rivas and R. Elber, "Exact milestoning," *The Journal of Chemical Physics*, vol. 142, no. 9, p. 03B602_1, 2015.

[53] E. Suarez, S. Lettieri, M. C. Zwier, C. A. Stringer, S. R. Subramanian, L. T. Chong, and D. M. Zuckerman, "Simultaneous computation of dynamical and equilibrium information using a weighted ensemble of trajectories," *Journal of chemical theory and computation*, vol. 10, no. 7, pp. 2658–2667, 2014.

[54] W. Noid, "Perspective: Coarse-grained models for biomolecular systems," *The Journal of chemical physics*, vol. 139, no. 9, p. 09B201_1, 2013.

[55] R. Potestio, C. Peter, and K. Kremer, "Computer simulations of soft matter: Linking the scales," *Entropy*, vol. 16, no. 8, pp. 4199–4245, 2014.

[56] S. A. Adcock and J. A. McCammon, "Molecular dynamics: survey of methods for simulating the activity of proteins," *Chemical reviews*, vol. 106, no. 5, pp. 1589–1615, 2006.

[57] J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror, and D. E. Shaw, "Long-timescale molecular dynamics simulations of protein structure and function," *Current opinion in structural biology*, vol. 19, no. 2, pp. 120–127, 2009.

[58] T. J. Lane, D. Shukla, K. A. Beauchamp, and V. S. Pande, "To milliseconds and beyond: challenges in the simulation of protein folding," *Current opinion in structural biology*, vol. 23, no. 1, pp. 58–65, 2013.

[59] C. Abrams and G. Bussi, "Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration," *Entropy*, vol. 16, no. 1, pp. 163–199, 2013.

[60] H. Fujisaki, K. Moritsugu, Y. Matsunaga, T. Morishita, and L. Maragliano, "Extended phase-space methods for enhanced sampling in molecular simulations: a review," *Frontiers in bioengineering and biotechnology*, vol. 3, p. 125, 2015.

[61] Y. Miao and J. A. McCammon, "Unconstrained enhanced sampling for free energy calculations of biomolecules: A review," *Molecular simulation*, vol. 42, no. 13, pp. 1046–1055, 2016.

[62] R. Elber, "Perspective: Computer simulations of long time dynamics," *The Journal of chemical physics*, vol. 144, no. 6, p. 060901, 2016.

[63] D. van der Spoel and M. M. Seibert, "Protein folding kinetics and thermodynamics from atomistic simulations," *Physical review letters*, vol. 96, no. 23, p. 238102, 2006.

[64] J. W. Pitera and W. Swope, "Understanding folding and design: Replica-exchange simulations of"trp-cage"miniproteins," *Proceedings of the National Academy of Sciences*, vol. 100, no. 13, pp. 7587–7592, 2003.

[65] R. Zhou, "Trp-cage: folding free energy landscape in explicit water," *Proceedings of the National Academy of Sciences*, vol. 100, no. 23, pp. 13280–13285, 2003.

[66] R. Zhou, "Exploring the protein folding free energy landscape: coupling replica exchange method with p3me/respa algorithm," *Journal of Molecular Graphics and Modelling*, vol. 22, no. 5, pp. 451–463, 2004.

[67] A. E. García and J. N. Onuchic, "Folding a protein in a computer: an atomic description of the folding/unfolding of protein a," *Proceedings of the National Academy of Sciences*, vol. 100, no. 24, pp. 13898–13903, 2003.

[68] F. Rao and A. Caflisch, "Replica exchange molecular dynamics simulations of reversible folding," *The Journal of chemical physics*, vol. 119, no. 7, pp. 4035–4042, 2003.

[69] M. Andrec, A. K. Felts, E. Gallicchio, and R. M. Levy, "Protein folding pathways from replica exchange simulations and a kinetic network model," *Proceedings of the National Academy of Sciences*, vol. 102, no. 19, pp. 6801–6806, 2005.

[70] P. Liu, X. Huang, R. Zhou, and B. J. Berne, "Hydrophobic aided replica exchange: an efficient algorithm for protein folding in explicit solvent," *The Journal of Physical Chemistry B*, vol. 110, no. 38, pp. 19018–19022, 2006.

[71] F. Ding, D. Tsao, H. Nie, and N. V. Dokholyan, "Ab initio folding of proteins with all-atom discrete molecular dynamics," *Structure*, vol. 16, no. 7, pp. 1010–1018, 2008.

[72] E. A. Proctor, F. Ding, and N. V. Dokholyan, "Discrete molecular dynamics," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 1, no. 1, pp. 80–92, 2011.

[73] D. Paschek, S. Hempel, and A. E. García, "Computing the stability diagram of the trp-cage miniprotein," *Proceedings of the National Academy of Sciences*, vol. 105, no. 46, pp. 17754–17759, 2008.

[74] D. Paschek, R. Day, and A. E. García, "Influence of water–protein hydrogen bonding on the stability of trp-cage miniprotein. a comparison between

the tip3p and tip4p-ew water models," *Physical Chemistry Chemical Physics*, vol. 13, no. 44, pp. 19840–19847, 2011.

[75] H. Nguyen, J. Maier, H. Huang, V. Perrone, and C. Simmerling, "Folding simulations for proteins with diverse topologies are accessible in days with a physics-based force field and implicit solvent," *Journal of the American Chemical Society*, vol. 136, no. 40, pp. 13959–13962, 2014.

[76] F. Jiang and Y.-D. Wu, "Folding of fourteen small proteins with a residue-specific force field and replica-exchange molecular dynamics," *Journal of the American Chemical Society*, vol. 136, no. 27, pp. 9536–9539, 2014.

[77] E. K. Peter, M. Agarwal, B. Kim, I. V. Pivkin, and J.-E. Shea, "How water layers on graphene affect folding and adsorption of trpzip2," *The Journal of chemical physics*, vol. 141, no. 22, p. 12B611_1, 2014.

[78] A. Laio and M. Parrinello, "Escaping free-energy minima," *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12562–12566, 2002.

[79] A. Barducci, G. Bussi, and M. Parrinello, "Well-tempered metadynamics: A smoothly converging and tunable free-energy method," *Phys. Rev. Lett.*, vol. 100, p. 020603, Jan 2008.

[80] S. Piana and A. Laio, "A bias-exchange approach to protein folding," *The journal of physical chemistry B*, vol. 111, no. 17, pp. 4553–4559, 2007.

[81] G. Bussi, F. L. Gervasio, A. Laio, and M. Parrinello, "Free-energy landscape for β hairpin folding from combined parallel tempering and metadynamics," *Journal of the American Chemical Society*, vol. 128, no. 41, pp. 13435–13441, 2006.

[82] S. Piana, A. Laio, F. Marinelli, M. Van Troys, D. Bourry, C. Ampe, and J. C. Martins, "Predicting the effect of a point mutation on a protein fold: the villin and advillin headpieces and their pro62ala mutants," *Journal of molecular biology*, vol. 375, no. 2, pp. 460–470, 2008.

[83] D. Kimanius, I. Pettersson, G. Schluckebier, E. Lindahl, and M. Andersson, "Saxs-guided metadynamics," *Journal of chemical theory and computation*, vol. 11, no. 7, pp. 3491–3498, 2015.

[84] D. Bonetti, C. Camilloni, L. Visconti, S. Longhi, M. Brunori, M. Vendruscolo, and S. Gianni, "Identification and structural characterization of an intermediate in the folding of the measles virus x domain," *Journal of Biological Chemistry*, pp. jbc–M116, 2016.

[85] R. Singh, R. Bansal, A. S. Rathore, and G. Goel, "Equilibrium ensembles for insulin folding from bias-exchange metadynamics," *Biophysical journal*, vol. 112, no. 8, pp. 1571–1585, 2017.

[86] A. Barducci, R. Chelli, P. Procacci, V. Schettino, F. L. Gervasio, and M. Parrinello, "Metadynamics simulation of prion protein: β-structure stability and the early stages of misfolding," *Journal of the American Chemical Society*, vol. 128, no. 8, pp. 2705–2710, 2006.

[87] G. H. Zerze, C. M. Miller, D. Granata, and J. Mittal, "Free energy surface of an intrinsically disordered protein: comparison between temperature replica exchange molecular dynamics and bias-exchange metadynamics," *Journal of chemical theory and computation*, vol. 11, no. 6, pp. 2776–2782, 2015.

[88] G. M. Torrie and J. P. Valleau, "Monte carlo free energy estimates using non-boltzmann sampling: application to the sub-critical lennard-jones fluid," *Chemical Physics Letters*, vol. 28, no. 4, pp. 578–581, 1974.

[89] B. Roux, "The calculation of the potential of mean force using computer simulations," *Computer physics communications*, vol. 91, no. 1-3, pp. 275–282, 1995.

[90] C. Bartels and M. Karplus, "Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1450–1462, 1997.

[91] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method," *Journal of computational chemistry*, vol. 13, no. 8, pp. 1011–1021, 1992.

[92] W. S. Young and C. L. Brooks III, "A microscopic view of helix propagation: N and c-terminal helix growth in alanine helices," *Journal of molecular biology*, vol. 259, no. 3, pp. 560–572, 1996.

[93] F. B. Sheinerman and C. L. Brooks III, "Calculations on folding of segment b1 of streptococcal protein g1," *Journal of molecular biology*, vol. 278, no. 2, pp. 439–456, 1998.

[94] B. D. Bursulaya and C. L. Brooks, "Folding free energy surface of a three-stranded β-sheet protein," *Journal of the American Chemical Society*, vol. 121, no. 43, pp. 9947–9951, 1999.

[95] H. Meshkin and F. Zhu, "Thermodynamics of protein folding studied by umbrella sampling along a reaction coordinate of native contacts," *Journal of chemical theory and computation*, vol. 13, no. 5, pp. 2086–2097, 2017.

[96] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. Elsevier, 1992.

[97] C. Schütte and M. Sarich, "Metastability and markov state models in molecular dynamics, volume 24 of courant lecture notes," *American Mathematical Society*, 2013.

[98] F. Noé, J. Chodera, G. Bowman, V. Pande, and F. Noé, "An introduction to markov state models and their application to long timescale molecular simulation, vol. 797 of advances in experimental medicine and biology," 2014.

[99] W. Wang, S. Cao, L. Zhu, and X. Huang, "Constructing markov state models to elucidate the functional conformational changes of complex biomolecules," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 8, no. 1, p. e1343, 2018.

[100] B. E. Husic and V. S. Pande, "Markov state models: From an art to a science," *Journal of the American Chemical Society*, vol. 140, no. 7, pp. 2386–2396, 2018.

[101] C. R. Schwantes and V. S. Pande, "Improvements in markov state model construction reveal many non-native interactions in the folding of ntl9," *Journal of chemical theory and computation*, vol. 9, no. 4, pp. 2000–2009, 2013.

[102] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, "Identification of slow molecular order parameters for markov model construction," *The Journal of chemical physics*, vol. 139, no. 1, p. 07B604_1, 2013.

[103] B. A. Luty and J. A. McCammon, "Simulation of bimolecular reactions: synthesis of the encounter and reaction steps," *Molecular simulation*, vol. 10, no. 1, pp. 61–65, 1993.

[104] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, "Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 19011–19016, 2009.

[105] D. L. Ensign and V. S. Pande, "The fip35 ww domain folds with structural and mechanistic heterogeneity in molecular dynamics simulations," *Biophysical journal*, vol. 96, no. 8, pp. L53–L55, 2009.

[106] G. R. Bowman, K. A. Beauchamp, G. Boxer, and V. S. Pande, "Progress and challenges in the automated construction of markov state models for full protein systems," *The Journal of chemical physics*, vol. 131, no. 12, p. 124101, 2009.

[107] J. K. Weber, R. L. Jack, C. R. Schwantes, and V. S. Pande, "Dynamical phase transitions reveal amyloid-like states on protein folding landscapes," *Biophysical journal*, vol. 107, no. 4, pp. 974–982, 2014.

[108] F. Rao and A. Caflisch, "The protein folding network," *Journal of molecular biology*, vol. 342, no. 1, pp. 299–306, 2004.

[109] J. Preto and C. Clementi, "Fast recovery of free energy landscapes via diffusion-map-directed molecular dynamics," *Physical Chemistry Chemical Physics*, vol. 16, no. 36, pp. 19181–19191, 2014.

[110] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and computational harmonic analysis*, vol. 21, no. 1, pp. 5–30, 2006.

[111] W. Zheng, M. A. Rohrdanz, M. Maggioni, and C. Clementi, "Polymer reversal rate calculated via locally scaled diffusion map," *The Journal of chemical physics*, vol. 134, no. 14, p. 144109, 2011.

[112] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[113] I. Daidone, A. Amadei, D. Roccatano, and A. Di Nola, "Molecular dynamics simulation of protein folding by essential dynamics sampling: folding landscape of horse heart cytochrome c," *Biophysical journal*, vol. 85, no. 5, pp. 2865–2871, 2003.

[114] D. Hamelberg, J. Mongan, and J. A. McCammon, "Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules," *The Journal of chemical physics*, vol. 120, no. 24, pp. 11919–11929, 2004.

[115] W. Sinko, Y. Miao, C. A. F. de Oliveira, and J. A. McCammon, "Population based reweighting of scaled molecular dynamics," *The Journal of Physical Chemistry B*, vol. 117, no. 42, pp. 12759–12768, 2013.

[116] Y. Miao, S. E. Nichols, and J. A. McCammon, "Free energy landscape of g-protein coupled receptors, explored by accelerated molecular dynamics," *Physical Chemistry Chemical Physics*, vol. 16, no. 14, pp. 6398–6406, 2014.

[117] L. Yang, M. P. Grubb, and Y. Q. Gao, "Application of the accelerated molecular dynamics simulations to the folding of a small protein," *The Journal of chemical physics*, vol. 126, no. 12, p. 03B616, 2007.

[118] L. Yang, Q. Shao, and Y. Q. Gao, "Thermodynamics and folding pathways of trpzip2: an accelerated molecular dynamics simulation study," *The Journal of Physical Chemistry B*, vol. 113, no. 3, pp. 803–808, 2008.

[119] Y. Miao, F. Feixas, C. Eun, and J. A. McCammon, "Accelerated molecular dynamics simulations of protein folding," *Journal of computational chemistry*, vol. 36, no. 20, pp. 1536–1549, 2015.

[120] P. G. Bolhuis, "Transition-path sampling of β-hairpin folding," *Proceedings of the National Academy of Sciences*, vol. 100, no. 21, pp. 12129–12134, 2003.

[121] P. G. Bolhuis, "Kinetic pathways of β-hairpin (un) folding in explicit solvent," *Biophysical journal*, vol. 88, no. 1, pp. 50–61, 2005.

[122] D. A. Evans and D. J. Wales, "Folding of the gb1 hairpin peptide from discrete path sampling," *The Journal of chemical physics*, vol. 121, no. 2, pp. 1080–1090, 2004.

[123] D. J. Wales, "Discrete path sampling," *Molecular physics*, vol. 100, no. 20, pp. 3285–3305, 2002.

[124] J. Juraszek and P. Bolhuis, "Sampling the multiple folding mechanisms of trp-cage in explicit solvent," *Proceedings of the National Academy of Sciences*, vol. 103, no. 43, pp. 15859–15864, 2006.

[125] J. Juraszek and P. G. Bolhuis, "Rate constant and reaction coordinate of trp-cage folding in explicit water," *Biophysical Journal*, vol. 95, no. 9, pp. 4246 – 4257, 2008.

[126] J. Juraszek, G. Saladino, T. Van Erp, and F. Gervasio, "Efficient numerical reconstruction of protein folding kinetics with partial path sampling and pathlike variables," *Physical review letters*, vol. 110, no. 10, p. 108106, 2013.

[127] R. J. Allen, D. Frenkel, and P. R. ten Wolde, "Forward flux sampling-type schemes for simulating rare events: Efficiency analysis," *The Journal of chemical physics*, vol. 124, no. 19, p. 194111, 2006.

[128] T. S. Van Erp and P. G. Bolhuis, "Elaborating transition interface sampling methods," *Journal of computational Physics*, vol. 205, no. 1, pp. 157–181, 2005.

[129] M. Chopra, A. S. Reddy, N. Abbott, and J. De Pablo, "Folding of polyglutamine chains," *The Journal of chemical physics*, vol. 129, no. 13, p. 10B603, 2008.

[130] P. Eastman, N. Grønbech-Jensen, and S. Doniach, "Simulation of protein folding by reaction path annealing," *The Journal of Chemical Physics*, vol. 114, no. 8, pp. 3823–3841, 2001.

[131] J. Lee, I.-H. Lee, I. Joung, J. Lee, and B. R. Brooks, "Finding multiple reaction pathways via global optimization of action," *Nature Communications*, vol. 8, p. 15443, 2017.

[132] P. Faccioli, M. Sega, F. Pederiva, and H. Orland, "Dominant pathways in protein folding," *Physical review letters*, vol. 97, no. 10, p. 108101, 2006.

[133] M. Marchi and P. Ballone, "Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems," *The Journal of Chemical Physics*, vol. 110, no. 8, pp. 3697–3702, 1999.

[134] E. Paci and M. Karplus, "Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations11edited by a. r. fersht," *Journal of Molecular Biology*, vol. 288, no. 3, pp. 441 – 459, 1999.

[135] M. Bonomi, D. Branduardi, F. L. Gervasio, and M. Parrinello, "The unfolded ensemble and folding mechanism of the c-terminal gb1 beta-hairpin," *Journal of the American Chemical Society*, vol. 130, no. 42, pp. 13938–13944, 2008. PMID: 18811160.

[136] E. Vanden-Eijnden *et al.*, "Transition-path theory and path-finding algorithms for the study of rare events.," *Annual review of physical chemistry*, vol. 61, pp. 391–420, 2010.

[137] D. Branduardi, F. L. Gervasio, and M. Parrinello, "From a to b in free energy space," *The Journal of chemical physics*, vol. 126, no. 5, p. 054103, 2007.

[138] P. L. Freddolino, F. Liu, M. Gruebele, and K. Schulten, "Ten-microsecond molecular dynamics simulation of a fast-folding ww domain," *Biophysical journal*, vol. 94, no. 10, pp. L75–L77, 2008.

[139] P. Metzner, C. Schütte, and E. Vanden-Eijnden, "Illustration of transition path theory on a collection of simple examples," *The Journal of chemical physics*, vol. 125, no. 8, p. 084110, 2006.

[140] F. Liu, D. Du, A. A. Fuller, J. E. Davoren, P. Wipf, J. W. Kelly, and M. Gruebele, "An experimental survey of the transition between two-state and downhill

protein folding scenarios," *Proceedings of the National Academy of Sciences*, vol. 105, no. 7, pp. 2369–2374, 2008.

[141] S. V. Krivov, "The free energy landscape analysis of protein (fip35) folding dynamics," *The journal of physical chemistry B*, vol. 115, no. 42, pp. 12315–12324, 2011.

[142] K. Lindorff-Larsen and S. Piana and K. Palmo and P. Maragakis and J. L. Klepeis et al., "Improved side-chain torsion potentials for the Amber ff99SB protein force field," *Proteins: Structure, Function, Bioinformatics*, vol. 78, 2010.

[143] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of chemical physics*, vol. 126, no. 1, p. 014101, 2007.

[144] S. a Beccara, T. Škrbić, R. Covino, C. Micheletti, and P. Faccioli, "Folding pathways of a knotted protein with a realistic atomistic force field," *PLOS Computational Biology*, vol. 9, pp. 1–9, 03 2013.

[145] E. Weinan, W. Ren, and E. Vanden-Eijnden, "Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes," *Chemical Physics Letters*, vol. 413, no. 1-3, pp. 242–247, 2005.

[146] H. Eyring, "The activated complex and the absolute rate of chemical reactions.," *Chemical Reviews*, vol. 17, no. 1, pp. 65–77, 1935.

[147] J. Horiuti, "On the statistical mechanical treatment of the absolute rate of chemical reaction," *Bulletin of the Chemical Society of Japan*, vol. 13, no. 1, pp. 210–216, 1938.

[148] E. Wigner, "The transition state method," *Transactions of the Faraday Society*, vol. 34, pp. 29–41, 1938.

[149] J.-H. Prinz, M. Held, J. C. Smith, and F. Noé, "Efficient computation, sensitivity, and error analysis of committor probabilities for complex dynamical processes," *Multiscale Modeling & Simulation*, vol. 9, no. 2, pp. 545–567, 2011.

[150] Y. Khoo, J. Lu, and L. Ying, "Solving for high dimensional committor functions using artificial neural networks," *arXiv preprint arXiv:1802.10275*, 2018.

[151] S. V. Krivov, "Protein folding free energy landscape along the committor-the optimal folding coordinate.," *Journal of chemical theory and computation*, 2018.

[152] E. Weinan, W. Ren, and E. Vanden-Eijnden, "String method for the study of rare events," *Physical Review B*, vol. 66, no. 5, p. 052301, 2002.

[153] W. Ren *et al.*, "Higher order string method for finding minimum energy paths," *Communications in Mathematical Sciences*, vol. 1, no. 2, pp. 377–384, 2003.

[154] E. Weinan, W. Ren, and E. Vanden-Eijnden, "Simplified and improved string method for computing the minimum energy paths in barrier-crossing events," *Journal of Chemical Physics*, vol. 126, no. 16, p. 164103, 2007.

[155] E. Weinan, W. Ren, and E. Vanden-Eijnden, "Finite temperature string method for the study of rare events," *J. Phys. Chem. B*, vol. 109, no. 14, pp. 6688–6693, 2005.

[156] W. Ren, E. Vanden-Eijnden, P. Maragakis, and W. E, "Transition pathways in complex systems: Application of the finite-temperature string method to the alanine dipeptide," *The Journal of chemical physics*, vol. 123, no. 13, p. 134109, 2005.

[157] A. C. Pan, D. Sezer, and B. Roux, "Finding transition pathways using the string method with swarms of trajectories," *The journal of physical chemistry B*, vol. 112, no. 11, pp. 3432–3440, 2008.

[158] W. Gan, S. Yang, and B. Roux, "Atomistic view of the conformational activation of src kinase using the string method with swarms-of-trajectories," *Biophysical journal*, vol. 97, no. 4, pp. L8–L10, 2009.

[159] D. Shukla, Y. Meng, B. Roux, and V. S. Pande, "Activation pathway of src kinase reveals intermediate states as targets for drug design," *Nature communications*, vol. 5, p. 3397, 2014.

[160] J. L. Adelman and M. Grabe, "Simulating rare events using a weighted ensemble-based string method," *The Journal of chemical physics*, vol. 138, no. 4, p. 01B616, 2013.

[161] T. F. Miller, E. Vanden-Eijnden, and D. Chandler, "Solvent coarse-graining and the string method applied to the hydrophobic collapse of a hydrated chain," *Proceedings of the National Academy of Sciences*, vol. 104, no. 37, pp. 14559–14564, 2007.

[162] V. Ovchinnikov, M. Karplus, and E. Vanden-Eijnden, "Free energy of conformational transition paths in biomolecules: The string method and its application to myosin vi," *The Journal of chemical physics*, vol. 134, no. 8, p. 02B631, 2011.

[163] C. Zhao and S. Y. Noskov, "The molecular mechanism of ion-dependent gating in secondary transporters," *PLoS computational biology*, vol. 9, no. 10, p. e1003296, 2013.

[164] C. Qiu, T. Qian, and W. Ren, "Application of the string method to the study of critical nuclei in capillary condensation," *The Journal of chemical physics*, vol. 129, no. 15, p. 154711, 2008.

[165] H. Jónsson, G. Mills, and K. W. Jacobsen, "Nudged elastic band method for finding minimum energy paths of transitions," in *Classical and quantum dynamics in condensed phase simulations*, pp. 385–404, World Scientific, 1998.

[166] R. Elber and D. Shalloway, "Temperature dependent reaction coordinates," *The Journal of Chemical Physics*, vol. 112, no. 13, pp. 5539–5545, 2000.

[167] M. K. Kim, R. L. Jernigan, and G. S. Chirikjian, "Efficient generation of feasible pathways for protein conformational transitions," *Biophysical journal*, vol. 83, no. 3, pp. 1620–1630, 2002.

[168] A. Ghosh, R. Elber, and H. A. Scheraga, "An atomically detailed study of the folding pathways of protein a with the stochastic difference equation," *Proceedings of the National Academy of Sciences*, vol. 99, no. 16, pp. 10394–10398, 2002.

[169] D. R. Weiss and M. Levitt, "Can morphing methods predict intermediate structures?," *Journal of molecular biology*, vol. 385, no. 2, pp. 665–674, 2009.

[170] P. Faccioli, M. Sega, F. Pederiva, and H. Orland, "Dominant pathways in protein folding," *Physical review letters*, vol. 97, no. 10, p. 108101, 2006.

[171] M. Sega, P. Faccioli, F. Pederiva, G. Garberoglio, and H. Orland, "Quantitative protein dynamics from dominant folding pathways," *Physical review letters*, vol. 99, no. 11, p. 118102, 2007.

[172] H. Orland, "Generating transition paths by langevin bridges," *The Journal of chemical physics*, vol. 134, no. 17, p. 174114, 2011.

[173] M. Delarue, P. Koehl, and H. Orland, "Ab initio sampling of transition paths by conditioned langevin dynamics," *The Journal of Chemical Physics*, vol. 147, no. 15, p. 152703, 2017.

[174] R. Elber, J. M. Bello-Rivas, P. Ma, A. E. Cardenas, and A. Fathizadeh, "Calculating iso-committor surfaces as optimal reaction coordinates with milestoning," *Entropy*, vol. 19, no. 5, p. 219, 2017.

[175] G. Hummer, "From transition paths to transition states and rate coefficients," *The Journal of chemical physics*, vol. 120, no. 2, pp. 516–523, 2004.

[176] J. B. Anderson, "Predicting rare events in molecular dynamics," *Advances in Chemical Physics*, vol. 91, pp. 381–432, 1995.

[177] C. Bennett, "Algorithms for chemical computations," in *ACS symposium Series*, vol. 46, pp. 63–97, Am. Chem. Soc Washington, 1977.

[178] D. Chandler, "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation," *The Journal of Chemical Physics*, vol. 68, no. 6, pp. 2959–2970, 1978.

[179] E. Vanden-Eijnden, M. Venturoli, G. Ciccotti, and R. Elber, "On the assumptions underlying milestoning," *The Journal of chemical physics*, vol. 129, no. 17, p. 174102, 2008.

[180] G. Ciccotti and M. Ferrario, "Blue moon approach to rare events," *Molecular simulation*, vol. 30, no. 11-12, pp. 787–793, 2004.

[181] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden, "Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics," *ChemPhysChem*, vol. 6, no. 9, pp. 1809–1814, 2005.

[182] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, "String method in collective variables: Minimum free energy paths and isocommittor surfaces," *The Journal of chemical physics*, vol. 125, no. 2, p. 024106, 2006.

[183] T. Hastie and W. Stuetzle, "Principal curves," *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 502–516, 1989.

[184] H. Risken, "Fokker-planck equation," in *The Fokker-Planck Equation*, pp. 63–95, Springer, 1996.

[185] J. Lu and J. Nolen, "Reactive trajectories and the transition path process," *Probability Theory and Related Fields*, vol. 161, no. 1-2, pp. 195–244, 2015.

[186] W. Zheng, M. A. Rohrdanz, and C. Clementi, "Rapid exploration of configuration space with diffusion-map-directed molecular dynamics," *The journal of physical chemistry B*, vol. 117, no. 42, pp. 12769–12776, 2013.

[187] D. Kraft, "A software package for sequential quadratic programming," *Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt*, 1988.

[188] E. Jones, T. Oliphant, and P. Peterson, "{SciPy}: open source scientific tools for {Python}," 2014.

[189] K. A. Dill and J. L. MacCallum, "The protein-folding problem, 50 years on," *Science*, vol. 338, no. 6110, pp. 1042–1046, 2012.

[190] S. Bottaro and K. Lindorff-Larsen, "Biophysical experiments and biomolecular simulations: A perfect match?," *Science*, vol. 361, no. 6400, pp. 355–360, 2018.

[191] S. Piana, J. L. Klepeis, and D. E. Shaw, "Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations," *Current opinion in structural biology*, vol. 24, pp. 98–105, 2014.

[192] D. N. Ivankov, S. O. Garbuzynskiy, E. Alm, K. W. Plaxco, D. Baker, and A. V. Finkelstein, "Contact order revisited: influence of protein size on the folding rate," *Protein science*, vol. 12, no. 9, pp. 2057–2062, 2003.

[193] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, and D. E. Shaw, "Systematic validation of protein force fields against experimental data," *PloS one*, vol. 7, no. 2, p. e32131, 2012.

[194] K. A. Beauchamp, Y.-S. Lin, R. Das, and V. S. Pande, "Are protein force fields getting better? a systematic benchmark on 524 diverse nmr measurements," *Journal of chemical theory and computation*, vol. 8, no. 4, pp. 1409–1414, 2012.

[195] P. Robustelli, S. Piana, and D. E. Shaw, "Developing a molecular dynamics force field for both folded and disordered protein states," *Proceedings of the National Academy of Sciences*, p. 201800690, 2018.

[196] P. S. Nerenberg and T. Head-Gordon, "New developments in force fields for biomolecular simulations," *Current opinion in structural biology*, vol. 49, pp. 129–138, 2018.

[197] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell Jr, "Charmm36m: an improved force field for folded and intrinsically disordered proteins," *nature methods*, vol. 14, no. 1, p. 71, 2016.

[198] Y. Bai, "Protein folding pathways studied by pulsed- and native-state hydrogen exchange," *Chemical Reviews*, vol. 106, no. 5, pp. 1757–1768, 2006. PMID: 16683753.

[199] D. Nettels, S. Müller-Späth, F. Küster, H. Hofmann, D. Haenni, S. Rüegger, L. Reymond, A. Hoffmann, J. Kubelka, B. Heinz, K. Gast, R. B. Best, and B. Schuler, "Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins," *Proceedings of the National Academy of Sciences*, vol. 106, no. 49, pp. 20740–20745, 2009.

[200] J. F. Eccleston, J. P. Hutchinson, and H. D. White, "Stopped-flow techniques," *Protein Ligand Interactions: Structure and Spectroscopy. A Practical Approach Series(Harding, SE, and Chowdry, BZ, Eds.) pp*, pp. 201–237, 2001.

[201] B. Peterman, "Measurement of the dead time of a fluorescence stopped-flow instrument," *Analytical Biochemistry*, vol. 93, pp. 442 – 444, 1979.

[202] H. Miller, Z. Zhou, J. Shepherd, A. J. M. Wollman, and M. C. Leake, "Single-molecule techniques in biophysics: a review of the progress in methods and applications," *Reports on Progress in Physics*, vol. 81, no. 2, p. 024601, 2018.

[203] D. W. Piston and G.-J. Kremers, "Fluorescent protein fret: the good, the bad and the ugly," *Trends in biochemical sciences*, vol. 32, no. 9, pp. 407–414, 2007.

[204] R. B. Best, D. J. Brockwell, J. L. Toca-Herrera, A. W. Blake, D. A. Smith, S. E. Radford, and J. Clarke, "Force mode atomic force microscopy as a tool for protein folding studies," *Analytica Chimica Acta*, vol. 479, no. 1, pp. 87–105, 2003.

[205] B. Schuler and W. A. Eaton, "Protein folding studied by single-molecule fret," *Current Opinion in Structural Biology*, vol. 18, no. 1, pp. 16 – 26, 2008. Folding and Binding / Protein-nucleic acid interactions.

[206] D. K. Sasmal, L. E. Pulido, S. Kasal, and J. Huang, "Single-molecule fluorescence resonance energy transfer in molecular biology," *Nanoscale*, vol. 8, pp. 19928–19944, 2016.

[207] K. C. Neuman and A. Nagy, "Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy," *Nature Methods*, vol. 5, pp. 491 EP –, 05 2008.

[208] A. B. Churnside and T. T. Perkins, "Ultrastable atomic force microscopy: Improved force and positional stability," *FEBS Letters*, vol. 588, no. 19, pp. 3621 – 3630, 2014. SI: Single molecule techniques - Applications in biology.

[209] X. Michalet, S. Weiss, and M. Jäger, "Single-molecule fluorescence studies of protein folding and conformational dynamics," *Chemical Reviews*, vol. 106, no. 5, pp. 1785–1813, 2006. PMID: 16683755.

[210] P. R. Banerjee and A. A. Deniz, "Shedding light on protein folding landscapes by single-molecule fluorescence," *Chem. Soc. Rev.*, vol. 43, pp. 1172–1188, 2014.

[211] C. P. Toseland, "Fluorescent labeling and modification of proteins," *Journal of Chemical Biology*, vol. 6, pp. 85–95, Jul 2013.

[212] K. L. Holmes and L. M. Lantz, "Chapter 9 protein labeling with fluorescent probes," in *Cytometry*, vol. 63 of *Methods in Cell Biology*, pp. 185 – 204, Academic Press, 2001.

[213] L. Mavridis and R. W. Janes, "Pdb2cd: a web-based application for the generation of circular dichroism spectra from protein atomic coordinates," *Bioinformatics*, vol. 33, no. 1, pp. 56–63, 2017.

[214] N. J. Greenfield, "Using circular dichroism spectra to estimate protein secondary structure," *Nat. Protoc.*, vol. 1, no. 6, pp. 2876–2890, 2007.

[215] A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos, "Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 24, pp. E3095–103, 2015.

[216] A. Lobley, L. Whitmore, and B. Wallace, "Dichroweb: An interactive website for the analysis of protein secondary structure from circular dichroism spectra," *Bioinformatics*, vol. 18, pp. 211–212, 2002.

[217] H. Roder, K. Maki, and H. Cheng, "Early Events in Protein Folding Explored by Rapid Mixing Methods," *Chemical Reviews*, vol. 106, no. 5, pp. 1836–1861, 2006.

[218] B. Ranjbar and P. Gill, "Circular Dichroism Techniques: Biomolecular and Nanostructural Analyses- A Review," *Chemical Biology & Drug Design*, vol. 74, no. 2, pp. 101–120, 2009.

[219] D. S. Kliger, E. Chen, and R. A. Goldbeck, "Nanosecond time-resolved natural and magnetic chiroptical spectroscopies," in *Comprehensive Chiroptical Spectroscopy* (N. Berova, P. L. Polavarapu, K. Nakanishi, and R. W. Woody, eds.), vol. 1, ch. 7, Hoboken, NJ: Wiley, 2012.

[220] F. Hache, "Time-resolved circular dichroism: What can we learn on conformational changes?," in *SPIE OPTO* (C. E. Tabor, F. Kajzar, T. Kaino, and Y. Koike, eds.), pp. 93600K–6, SPIE, 2015.

[221] J. Jiang, D. Abramavicius, B. M. Bulheller, J. D. Hirst, and S. Mukamel, "Ultraviolet spectroscopy of protein backbone transitions in aqueous solution: Combined qm and mm simulations," *The Journal of Physical Chemistry B*, vol. 114, no. 24, pp. 8270–8277, 2010.

[222] H. E. Auer, "Far-ultraviolet absorption and circular dichroism spectra of L-tryptophan and some derivatives," *J. Am. Chem. Soc.*, vol. 95, no. 9, pp. 3003–3011, 1973.

[223] R. W. Woody, "Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins," *Eur. Biophys. J.*, vol. 23, no. 4, pp. 253–262, 1994.

[224] B. M. Bulheller, A. Rodger, and J. D. Hirst, "Circular and linear dichroism of proteins," *Phys. Chem. Chem. Phys.*, vol. 9, pp. 2020–2035, 2007.

[225] A. L. Hansen and L. E. Kay, "Quantifying millisecond time-scale exchange in proteins by cpmg relaxation dispersion nmr spectroscopy of side-chain carbonyl groups," *Journal of biomolecular NMR*, vol. 50, no. 4, pp. 347–355, 2011.

[226] "K. nakanishi, n. berova, r.w. woody (eds.): Circular dichroism, principles and applications, vch, weinheim, new york, cambridge, isbn 1-56081-618-x, 1994.," *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 99, no. 2, pp. 230–230.

[227] C. Cantor and P. Schimmel, *Biophysical Chemistry: Part I: The Conformation of Biological Macromolecules*. W. H. Freeman, 1980.

[228] N. Sreerama and R. W. Woody, "On the analysis of membrane protein circular dichroism spectra.," *Protein Sci.*, vol. 13, no. 1, pp. 100–12, 2004.

[229] L. Rosenfeld, "Quantenmechanische theorie der natürlichen optischen aktivität von flüssigkeiten und gasen," *Zeitschrift für Physik*, vol. 52, pp. 161–174, Mar 1929.

[230] I. Tinoco, "Theoretical Aspects of Optical Activity. Part two: Polymers," *Adv. Chem. Phys.*, vol. 4, p. 113, 1962.

[231] P. M. Bayley, E. B. Nielsen, and J. A. Schellman, "Rotatory properties of molecules containing two peptide groups: theory," *The Journal of Physical Chemistry*, vol. 73, no. 1, pp. 228–243, 1969.

[232] R. W. Woody, "Improved calculation of the $n\pi^*$ rotational strength in polypeptides," *J. Chem. Phys.*, vol. 49, no. 11, pp. 4797–4806, 1968.

[233] R. W. Woody and I. Tinoco, "Optical rotation of oriented helices. iii. calculation of the rotatory dispersion and circular dichroism of the alpha- and 310-helix," *The Journal of Chemical Physics*, vol. 46, no. 12, pp. 4927–4945, 1967.

[234] A. Davydov, *Theory of Molecular Excitons*. Springer US, 2013.

[235] N. Harada and K. Nakanishi, *Circular dichroic spectroscopy: exciton coupling in organic stereochemistry*. Mill Valley, CA: Univ Science Books, 1983.

[236] S. Arulmozhiraja and M. L. Coote, "1 L a and 1 L b States of Indole and Azaindole: Is Density Functional Theory Inadequate?," *J. Chem. Theory Comput.*, vol. 8, pp. 575–584, feb 2012.

[237] Y. Yamamoto and J. Tanaka, "Polarized Absorption Spectra of Crystals of Indole and Its Related Compounds," *Bull. Chem. Soc. Jpn.*, vol. 45, no. 5, pp. 1362–1366, 1972.

[238] L. A. Morozova-Roche, J. A. Jones, W. Noppe, and C. M. Dobson, "Independent nucleation and heterogeneous assembly of structure during folding of equine lysozyme," *Journal of Molecular Biology*, vol. 289, no. 4, pp. 1055 – 1073, 1999.

[239] G. Wildegger and T. Kiefhaber, "Three-state model for lysozyme folding: Triangular folding mechanism with an energetically trapped in- termediate," *J. Mol. Bio.*, vol. 270, pp. 294–304, 1997.

[240] A. Onufriev, D. Bashford, and D. A. Case, "Exploring protein native states and large-scale conformational changes with a modified generalized Born model," *Proteins: Structure, Function, Bioinformatics*, vol. 55, 2004.

[241] M. Schaefer, C. Bartels, and M. Karplus, "Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solvation model," *Journal of Molecular Biology*, vol. 284, 1998.

[242] C.-P. Hsu, G. R. Fleming, M. Head-Gordon, and T. Head-Gordon, "Excitation energy transfer in condensed media," *The Journal of Chemical Physics*, vol. 114, no. 7, pp. 3065–3072, 2001.

[243] M. F. Iozzi, B. Mennucci, J. Tomasi, and R. Cammi, "Excitation energy transfer (EET) between molecules in condensed matter: A novel application of the polarizable continuum model (PCM)," *J. Chem. Phys.*, vol. 120, no. 15, pp. 7029–7040, 2004.

[244] T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp)," *Chem. Phys. Lett.*, vol. 393, no. 1, pp. 51–57, 2004.

[245] E. Cancès, B. Mennucci, and J. Tomasi, "A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics," *The Journal of Chemical Physics*, vol. 107, no. 8, pp. 3032–3041, 1997.

[246] A. D. Becke, "Density-functional thermochemistry. iii. the role of exact exchange," *The Journal of Chemical Physics*, vol. 98, no. 7, pp. 5648–5652, 1993.

[247] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, "Gaussian 09 Revision D.01." Gaussian Inc. Wallingford CT 2009.

[248] S. Jurinovich, L. Cupellini, C. A. Guido, and B. Mennucci, "Exat: Excitonic analysis tool," *Journal of Computational Chemistry*, vol. 39, no. 5, pp. 279–286.

[249] K. Trueblood, H.-B. Bürgi, H. Burzlaff, J. Dunitz, C. Gramaccioli, H. Schulz, U. Shmueli, and S. Abrahams, "Atomic dispacement parameter nomenclature. report of a subcommittee on atomic displacement parameter nomenclature," *Acta Crystallographica Section A*, vol. 52, no. 5, pp. 770–781, 1996.

[250] H. Nakatani, K. Maki, K. Saeki, T. Aizawa, M. Demura, K. Kawano, S. Tomoda, and K. Kuwajima, "Equilibrium and Kinetics of the Folding and Unfolding of Canine Milk Lysozyme †," *Biochemistry*, vol. 46, no. 17, pp. 5238–5251, 2007.

[251] M. Mizuguchi, M. Arai, Y. Ke, K. Nitta, and K. Kuwajima, "Equilibrium and kinetics of the folding of equine lysozyme studied by circular dichroism spectroscopy," *J. Mol. Biol.*, vol. 283, no. 1, pp. 265–277, 1998.

[252] K. Sasahara, M. Demura, and K. Nitta, "Partially Unfolded Equilibrium State of Hen Lysozyme Studied by Circular Dichroism Spectroscopy," *Biochemistry*, vol. 39, no. 21, pp. 6475–6482, 2000.

[253] G. R. Spence, A. P. Capaldi, and S. E. Radford, "Trapping the on-pathway folding intermediate of Im7 at equilibrium," *J. Mol. Biol.*, vol. 341, no. 1, pp. 215–226, 2004.

[254] N. Ferguson, A. P. Capaldi, R. James, C. Kleanthous, and S. E. Radford, "Rapid folding with and without populated intermediates in the homologous four-helix proteins im7 and im911edited by a. r. fersht," *Journal of Molecular Biology*, vol. 286, no. 5, pp. 1597 – 1608, 1999.

[255] A. P. Capaldi, M. C. R. Shastry, C. Kleanthous, H. Roder, and S. E. Radford, "Ultrarapid mixing experiments reveal that im7 folds via an on-pathway intermediate," *Nature Structural Biology*, vol. 8, pp. 68 EP –, 01 2001.

[256] S. A. Gorski, A. P. Capaldi, C. Kleanthous, and S. E. Radford, "Acidic conditions stabilise intermediates populated during the folding of im7 and im911edited by c. r. matthews," *Journal of Molecular Biology*, vol. 312, no. 4, pp. 849 – 863, 2001.

[257] C. T. Friel, D. Alastair Smith, M. Vendruscolo, J. Gsponer, and S. E. Radford, "The mechanism of folding of im7 reveals competition between functional and kinetic evolutionary constraints," *Nature Structural &Amp; Molecular Biology*, vol. 16, pp. 318 EP –, 03 2009.

[258] A. M. Figueiredo, G. R. Moore, and S. B.-M. Whittaker, "Understanding how small helical proteins fold: conformational dynamics of im proteins relevant to their folding landscapes," *Biochemical Society Transactions*, vol. 40, no. 2, pp. 424–428, 2012.

[259] J. Gsponer, H. Hopearuoho, S. B.-M. Whittaker, G. R. Spence, G. R. Moore, E. Paci, S. E. Radford, and M. Vendruscolo, "Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein im7," *Proceedings of the National Academy of Sciences*, vol. 103, no. 1, pp. 99–104, 2006.

[260] E. Paci, C. T. Friel, K. Lindorff-Larsen, S. E. Radford, M. Karplus, and M. Vendruscolo, "Comparison of the transition state ensembles for folding of im7 and im9 determined using all-atom molecular dynamics simulations with φ value restraints," *Proteins: Structure, Function, and Bioinformatics*, vol. 54, no. 3, pp. 513–525.

[261] Z. Li and J. D. Hirst, "Quantitative first principles calculations of protein circular dichroism in the near-ultraviolet," *Chem. Sci.*, vol. 8, no. 6, pp. 4318–4333, 2017.

[262] M. Bonomi, G. T. Heller, C. Camilloni, and M. Vendruscolo, "Principles of protein structural ensemble determination," *Current opinion in structural biology*, vol. 42, pp. 106–116, 2017.

[263] E. T. Jaynes, "Information theory and statistical mechanics," *Physical review*, vol. 106, no. 4, p. 620, 1957.

[264] W. Boomsma, J. Ferkinghoff-Borg, and K. Lindorff-Larsen, "Combining experiments and simulations using the maximum entropy principle," *PLoS computational biology*, vol. 10, no. 2, p. e1003406, 2014.

[265] G. A. Silverman, P. I. Bird, R. W. Carrell, F. C. Church, P. B. Coughlin, P. G. W. Gettins, J. A. Irving, D. A. Lomas, C. J. Luke, R. W. Moyer, P. A. Pemberton, E. Remold-O'Donnell, G. S. Salvesen, J. Travis, and J. C. Whisstock, "The serpins are an expanding superfamily of structurally similar but functionally diverse proteins: Evolution, mechanism of inhibition, novel functions, and a revised nomenclature," *Journal of Biological Chemistry*, vol. 276, no. 36, pp. 33293–33296, 2001.

[266] J. C. Whisstock and S. P. Bottomley, "Molecular gymnastics: serpin structure, folding and misfolding," *Current Opinion in Structural Biology*, vol. 16, no. 6, pp. 761 – 768, 2006. Catalysis and regulation / Proteins.

[267] J. A. Huntington, R. J. Read, and R. W. Carrell, "Structure of a serpin–protease complex shows inhibition by deformation," *Nature*, vol. 407, pp. 923 EP –, 10 2000.

[268] B. Gooptu and D. A. Lomas, "Conformational pathology of the serpins: Themes, variations, and therapeutic strategies," *Annual Review of Biochemistry*, vol. 78, no. 1, pp. 147–176, 2009. PMID: 19245336.

[269] C. M. Greene, S. J. Marciniak, J. Teckman, I. Ferrarotti, M. L. Brantly, D. A. Lomas, J. K. Stoller, and N. G. McElvaney, "α1-antitrypsin deficiency," *Nature Reviews Disease Primers*, vol. 2, pp. 16051 EP –, 07 2016.

[270] D. A. Lomas, D. LI-Evans, J. T. Finch, and R. W. Carrell, "The mechanism of z α1-antitrypsin accumulation in the liver," *Nature*, vol. 357, pp. 605 EP –, 06 1992.

[271] T. Sveger, "Liver disease in alpha1-antitrypsin deficiency detected by screening of 200,000 infants," *New England Journal of Medicine*, vol. 294, no. 24, pp. 1316–1321, 1976. PMID: 1083485.

[272] S. Eriksson, J. Carlson, and R. Velez, "Risk of cirrhosis and primary liver cancer in alpha1-antitrypsin deficiency," *New England Journal of Medicine*, vol. 314, no. 12, pp. 736–739, 1986. PMID: 3485248.

[273] B. Krishnan and L. M. Gierasch, "Dynamic local unfolding in the serpin α-1 antitrypsin provides a mechanism for loop insertion and polymerization," *Nature structural & molecular biology*, vol. 18, no. 2, p. 222, 2011.

[274] D. Kim and M.-H. Yu, "Folding pathway of human α1-antitrypsin: Characterization of an intermediate that is active but prone to aggregation," *Biochemical and Biophysical Research Communications*, vol. 226, no. 2, pp. 378 – 384, 1996.

[275] E. L. James, J. C. Whisstock, M. G. Gore, and S. P. Bottomley, "Probing the unfolding pathway of α1-antitrypsin," *Journal of Biological Chemistry*, vol. 274, no. 14, pp. 9482–9488, 1999.

[276] K. Dolmer and P. G. W. Gettins, "How the serpin α1-proteinase inhibitor folds," *Journal of Biological Chemistry*, vol. 287, no. 15, pp. 12425–12432, 2012.

[277] Y. Tsutsui, R. Dela Cruz, and P. L. Wintrode, "Folding mechanism of the metastable serpin α1-antitrypsin," *Proceedings of the National Academy of Sciences*, vol. 109, no. 12, pp. 4467–4472, 2012.

[278] B. B. Stocks, A. Sarkar, P. L. Wintrode, and L. Konermann, "Early hydrophobic collapse of α1-antitrypsin facilitates formation of a metastable state: Insights from oxidative labeling and mass spectrometry," *Journal of Molecular Biology*, vol. 423, no. 5, pp. 789 – 799, 2012.

[279] N. Novoradovskaya, J. Lee, Z. X. Yu, V. J. Ferrans, and M. Brantly, "Inhibition of intracellular degradation increases secretion of a mutant form of

alpha1-antitrypsin associated with profound deficiency.," *The Journal of Clinical Investigation*, vol. 101, pp. 2693–2701, 6 1998.

[280] J. H. Teckman and D. H. Perlmutter, "Retention of mutant α1-antitrypsin z in endoplasmic reticulum is associated with an autophagic response," *American Journal of Physiology-Gastrointestinal and Liver Physiology*, vol. 279, no. 5, pp. G961–G974, 2000. PMID: 11052993.

[281] T. Kamimoto, S. Shoji, T. Hidvegi, N. Mizushima, K. Umebayashi, D. H. Perlmutter, and T. Yoshimori, "Intracellular inclusions containing mutant α1-antitrypsin z are propagated in the absence of autophagic activity," *Journal of Biological Chemistry*, vol. 281, no. 7, pp. 4467–4476, 2006.

[282] E. Miranda, J. Pérez, U. I. Ekeowa, N. Hadzic, N. Kalsheker, B. Gooptu, B. Portmann, D. Belorgey, M. Hill, S. Chambers, J. Teckman, G. J. Alexander, S. J. Marciniak, and D. A. Lomas, "A novel monoclonal antibody to characterize pathogenic polymers in liver disease associated with α1-antitrypsin deficiency," *Hepatology*, vol. 52, no. 3, pp. 1078–1088.

[283] U. I. Ekeowa, J. Freeke, E. Miranda, B. Gooptu, M. F. Bush, J. Pérez, J. Teckman, C. V. Robinson, and D. A. Lomas, "Defining the mechanism of polymerization in the serpinopathies," *Proceedings of the National Academy of Sciences*, vol. 107, no. 40, pp. 17146–17151, 2010.

[284] M. A. Behrens, T. J. Sendall, J. S. Pedersen, M. Kjeldgaard, J. A. Huntington, and J. K. Jensen, "The shapes of z-α1-antitrypsin polymers in solution support the c-terminal domain-swap mechanism of polymerization," *Biophysical journal*, vol. 107, no. 8, pp. 1905–1912, 2014.

[285] M. Yamasaki, W. Li, D. J. Johnson, and J. A. Huntington, "Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization," *Nature*, vol. 455, no. 7217, p. 1255, 2008.

[286] J. A. Irving, I. Haq, J. A. Dickens, S. V. Faull, and D. A. Lomas, "Altered native stability is the dominant basis for susceptibility of α1-antitrypsin mutants to polymerization," *Biochemical Journal*, vol. 460, no. 1, pp. 103–119, 2014.

[287] M. Owen, R. Carrell, and S. Brennan, "The abnormality of the s variant of human α-1-antitrypsin," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 453, no. 1, pp. 257–261, 1976.

[288] P. R. Elliott, P. E. Stein, D. Bilton, R. W. Carrell, and D. A. Lomas, "Structural explanation for the deficiency of s α1-antitrypsin," *Nature Structural and Molecular Biology*, vol. 3, no. 11, p. 910, 1996.

[289] J. H. Teckman and A. Jain, "Advances in alpha-1-antitrypsin deficiency liver disease," *Current Gastroenterology Reports*, vol. 16, p. 367, Dec 2013.

[290] K. R. Chapman, J. G. W. Burdon, E. Piitulainen, R. A. Sandhaus, N. Seersholm, J. M. Stocks, B. C. Stoel, L. Huang, Z. Yao, J. M. Edelman, and N. G. McElvaney, "Intravenous augmentation treatment and lung density in severe α1 antitrypsin deficiency (rapid): a randomised, double-blind, placebo-controlled trial," *The Lancet*, vol. 386, no. 9991, pp. 360 – 368, 2015.

[291] A. Dirksen, J. H. Dijkman, F. Madsen, B. Stoel, D. C. Hutchison, C. S. Ulrik, L. T. Skovgaard, A. Kok-Jensen, A. Rudolphus, N. Seersholm, *et al.*, "A randomized clinical trial of α1-antitrypsin augmentation therapy," *American journal of respiratory and critical care medicine*, vol. 160, no. 5, pp. 1468–1472, 1999.

[292] A. Dirksen, E. Piitulainen, D. Parr, C. Deng, M. Wencker, S. Shaker, and R. Stockley, "Exploring the role of ct densitometry: a randomised study of augmentation therapy in alpha-1 antitrypsin deficiency," *European Respiratory Journal*, 2009.

[293] T. Hidvegi, M. Ewing, P. Hale, C. Dippold, C. Beckett, C. Kemp, N. Maurice, A. Mukherjee, C. Goldbach, S. Watkins, *et al.*, "An autophagy-enhancing drug promotes degradation of mutant α1-antitrypsin z and reduces hepatic fibrosis," *Science*, vol. 329, no. 5988, pp. 229–232, 2010.

[294] S. Kaushal, M. Annamali, K. Blomenkamp, D. Rudnick, D. Halloran, E. M. Brunt, and J. H. Teckman, "Rapamycin reduces intrahepatic alpha-1-antitrypsin mutant z protein polymers and liver injury in a mouse model," *Experimental biology and medicine*, vol. 235, no. 6, pp. 700–709, 2010.

[295] R. Mahadeva, T. R. Dafforn, R. W. Carrell, and D. A. Lomas, "6-mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerization implications for the prevention of z α1-antitrypsin-related cirrhosis," *Journal of Biological Chemistry*, vol. 277, no. 9, pp. 6771–6774, 2002.

[296] S. Guo, S. L. Booten, M. Aghajan, G. Hung, C. Zhao, K. Blomenkamp, D. Gattis, A. Watt, S. M. Freier, J. H. Teckman, *et al.*, "Antisense oligonucleotide treatment ameliorates alpha-1 antitrypsin–related liver disease in mice," *The Journal of clinical investigation*, vol. 124, no. 1, pp. 251–261, 2014.

[297] M. Mallya, R. L. Phillips, S. A. Saldanha, B. Gooptu, S. C. Leigh Brown, D. J. Termine, A. M. Shirvani, Y. Wu, R. N. Sifers, R. Abagyan, *et al.*, "Small molecules block the polymerization of z α1-antitrypsin and increase the clearance of intracellular aggregates," *Journal of medicinal chemistry*, vol. 50, no. 22, pp. 5357–5363, 2007.

[298] B. Gooptu, E. Miranda, I. Nobeli, M. Mallya, A. Purkiss, S. C. L. Brown, C. Summers, R. L. Phillips, D. A. Lomas, and T. E. Barrett, "Crystallographic and cellular characterisation of two mechanisms stabilising the native fold of α1-antitrypsin: implications for disease and drug design," *Journal of molecular biology*, vol. 387, no. 4, pp. 857–868, 2009.

[299] Y.-P. Chang, R. Mahadeva, W.-S. W. Chang, S.-C. Lin, and Y.-H. Chu, "Small-molecule peptides inhibit z α1-antitrypsin polymerization," *Journal of cellular and molecular medicine*, vol. 13, no. 8b, pp. 2304–2316, 2009.

[300] Y.-P. Chang, R. Mahadeva, A. O. Patschull, I. Nobeli, U. I. Ekeowa, A. R. McKay, K. Thalassinos, J. A. Irving, I. Haq, M. P. Nyon, *et al.*, "Targeting serpins in high-throughput and structure-based drug design," in *Methods in enzymology*, vol. 501, pp. 139–175, Elsevier, 2011.

[301] A. O. Patschull, B. Gooptu, P. Ashford, T. Daviter, and I. Nobeli, "In silico assessment of potential druggable pockets on the surface of α1-antitrypsin conformers," *PLoS One*, vol. 7, no. 5, p. e36612, 2012.

[302] M. P. Nyon and B. Gooptu, "Therapeutic targeting of misfolding and conformational change in α1-antitrypsin deficiency," *Future medicinal chemistry*, vol. 6, no. 9, pp. 1047–1065, 2014.

[303] M. P. Nyon, T. Prentice, J. Day, J. Kirkpatrick, G. N. Sivalingam, G. Levy, I. Haq, J. A. Irving, D. A. Lomas, J. Christodoulou, *et al.*, "An integrative approach combining ion mobility mass spectrometry, x-ray crystallography, and nuclear magnetic resonance spectroscopy to study the conformational dynamics of α1-antitrypsin upon ligand binding," *Protein Science*, vol. 24, no. 8, pp. 1301–1312, 2015.

[304] J. A. Burrows, L. K. Willis, and D. H. Perlmutter, "Chemical chaperones mediate increased secretion of mutant α1-antitrypsin (α1-at) z: a potential pharmacological strategy for prevention of liver injury and emphysema in α1-at deficiency," *Proceedings of the National Academy of Sciences*, vol. 97, no. 4, pp. 1796–1801, 2000.

[305] J. H. Teckman, "Lack of effect of oral 4-phenylbutyrate on serum alpha-1-antitrypsin in patients with α-1-antitrypsin deficiency: A preliminary study," *Journal of pediatric gastroenterology and nutrition*, vol. 39, no. 1, pp. 34–37, 2004.

[306] S. Piana, K. Lindorff-Larsen, and D. E. Shaw, "Atomistic description of the folding of a dimeric protein," *The Journal of Physical Chemistry B*, vol. 117, no. 42, pp. 12935–12942, 2013.

[307] E. P. R., P. X. Y., D. T. R., and L. D. A., "Topography of a 2.0åstructure of α1-antitrypsin reveals targets for rational drug design to prevent conformational disease," *Protein Science*, vol. 9, no. 7, pp. 1274–1281.

[308] X. Huang, Y. Zheng, F. Zhang, Z. Wei, Y. Wang, R. W. Carrell, R. J. Read, G.-Q. Chen, and A. Zhou, "Molecular mechanism of z α1-antitrypsin deficiency," *Journal of Biological Chemistry*, vol. 291, no. 30, pp. 15674–15686, 2016.

[309] M. Onda, K. Nakatani, S. Takehara, M. Nishiyama, N. Takahashi, and M. Hirose, "Cleaved serpin refolds into the relaxed state via a stressed conformer," *Journal of Biological Chemistry*, vol. 283, no. 25, pp. 17568–17578, 2008.

[310] K. Chandrasekhar, H. Ke, N. Wang, T. Goodwin, L. M. Gierasch, A. Gershenson, and D. N. Hebert, "Cellular folding pathway of a metastable serpin," *Proceedings of the National Academy of Sciences*, vol. 113, no. 23, pp. 6484–6489, 2016.

[311] D. J. Tew and S. P. Bottomley, "Probing the equilibrium denaturation of the serpin α1-antitrypsin with single tryptophan mutants; evidence for structure in the urea unfolded state11edited by c. r. matthews," *Journal of Molecular Biology*, vol. 313, no. 5, pp. 1161 – 1169, 2001.

[312] A. W. McMillan, B. L. Kier, I. Shu, A. Byrne, N. H. Andersen, and W. W. Parson, "Fluorescence of tryptophan in designed hairpin and trp-cage miniproteins: measurements of fluorescence yields and calculations by quantum mechanical molecular dynamics simulations," *The Journal of Physical Chemistry B*, vol. 117, no. 6, pp. 1790–1809, 2013.

[313] A. Shrake and J. Rupley, "Environment and exposure to solvent of protein atoms. lysozyme and insulin," *Journal of molecular biology*, vol. 79, no. 2, pp. 351–371, 1973.

[314] V. V. H. Giri Rao and S. Gosavi, "On the folding of a structurally complex protein to its metastable active state," *Proceedings of the National Academy of Sciences*, 2018.

[315] B. J. D., O. J. Nelson, S. N. D., and W. P. G., "Funnels, pathways, and the energy landscape of protein folding: A synthesis," *Proteins: Structure, Function, and Bioinformatics*, vol. 21, no. 3, pp. 167–195.

[316] J. N. Onuchic and P. G. Wolynes, "Theory of protein folding," *Current Opinion in Structural Biology*, vol. 14, no. 1, pp. 70 – 75, 2004.

[317] C. Clementi, H. Nymeyer, and J. N. Onuchic, "Topological and energetic factors: what determines the structural details of the transition state ensemble and en-route intermediates for protein folding? an investigation for small globular proteins," *Journal of Molecular Biology*, vol. 298, no. 5, pp. 937 – 953, 2000.

[318] C. Hyeon and D. Thirumalai, "Capturing the essence of folding and functions of biomolecules using coarse-grained models," *Nature Communications*, vol. 2, pp. 487 EP –, 09 2011.

[319] M. A. Stephens, "Edf statistics for goodness of fit and some comparisons," *Journal of the American Statistical Association*, vol. 69, no. 347, pp. 730–737, 1974.

[320] M. Brantly, M. Courtney, and R. Crystal, "Repair of the secretion defect in the z form of α1-antitrypsin by addition of a second mutation," *Science*, vol. 242, no. 4886, pp. 1700–1702, 1988.

[321] R. Sifers, C. Hardick, and S. Woo, "Disruption of the 290-342 salt bridge is not responsible for the secretory defect of the piz α1-antitrypsin variant," *Journal of Biological Chemistry*, vol. 264, no. 5, pp. 2997–3001, 1989. cited By 30.

[322] M.-H. Yu, K. N. Lee, and J. Kim, "The z type variation of human α1-antitrypsin causes a protein folding defect," *Nature Structural Biology*, vol. 2, pp. 363 EP –, 05 1995.

[323] D. A. Lomas, J. R. Hurst, and B. Gooptu, "Update on alpha-1 antitrypsin deficiency: New therapies," *Journal of Hepatology*, vol. 65, no. 2, pp. 413 – 424, 2016.

[324] M. F. Bush, Z. Hall, K. Giles, J. Hoyes, C. V. Robinson, and B. T. Ruotolo, "Collision cross sections of proteins and their complexes: A calibration frame-

work and database for gas-phase structural biology," *Analytical Chemistry*, vol. 82, no. 22, pp. 9557–9565, 2010. PMID: 20979392.

[325] E. G. Marklund, M. T. Degiacomi, C. V. Robinson, A. J. Baldwin, and J. L. Benesch, "Collision cross sections for structural proteomics," *Structure*, vol. 23, no. 4, pp. 791 – 799, 2015.

[326] J. A. McCammon, B. R. Gelin, and M. Karplus, "Dynamics of folded proteins," *Nature*, vol. 267, no. 5612, p. 585, 1977.

[327] R. O. Dror, R. M. Dirks, J. Grossman, H. Xu, and D. E. Shaw, "Biomolecular simulation: a computational microscope for molecular biology," *Annual review of biophysics*, vol. 41, pp. 429–452, 2012.

[328] Y. Hashem and P. Auffinger, "A short guide for molecular dynamics simulations of rna systems," *Methods*, vol. 47, no. 3, pp. 187–197, 2009.

[329] A. Zee, *Quantum field theory in a nutshell*. Princeton university press, 2010.

[330] L. Onsager and S. Machlup, "Fluctuations and irreversible processes," *Physical Review*, vol. 91, no. 6, p. 1505, 1953.

[331] S. Machlup and L. Onsager, "Fluctuations and irreversible process. ii. systems with kinetic energy," *Physical Review*, vol. 91, no. 6, p. 1512, 1953.

[332] J.-P. Blaizot, D. D. Boni, P. Faccioli, and G. Garberoglio, "Heavy quark bound states in a quark–gluon plasma: Dissociation and recombination," *Nuclear Physics A*, vol. 946, pp. 49 – 88, 2016.

[333] C. W. Gardiner, "Handbook of stochastic methods for physics, chemistry and the natural sciences, + berlin," *Springer-Verlag Gibson, JJ (1986). The ecological approach to visual perception. Hillsdale, Nj: Lawrence Erlbaum Associates, Inc.(Original work published 1979) Hosking, JRM (1981), Fractional differencing. Biometrika*, vol. 68, pp. 165–176, 1985.

[334] G. Mazzola, S. a. Beccara, P. Faccioli, and H. Orland, "Fluctuations in the ensemble of reaction pathways," *The Journal of chemical physics*, vol. 134, no. 16, p. 164109, 2011.