Università degli Studi di Milano-Bicocca

Dipartimento di Biotecnologie e Bioscienze

Dottorato di Ricerca in Tecnologie Convergenti per i Sistemi Biomolecolari

Ciclo XXXIII

Structural modelling of biological macromolecules: the cases of neurofibromin, bifurcating Electron Transferring Flavoprotein and Amyloid-β (1-16) peptide

Candidato: Fabio Rizza

Matricola: 823672

Supervisor: Dr. Renata Tisi

Tutor: Prof Luca De Gioia

Coordinatore: Prof. Paola Branduardi

ANNO ACCADEMICO 2019/2020

Table of content

Summary

In this thesis, three independent projects were addressed, sharing the computational approach based on molecular modelling and in particular molecular dynamics.
In the first project, the Sec14-PH domain of neurofibromin (NF1) was investigated. The Sec14 domains have been identified in many different proteins, from prokaryotes to humans, serving as exchangers of lipid molecules between membranes, by means of a pocket whose opening is allowed by the motion of a specific alpha-helix (called lid helix). The crystal structure of the NF1-Sec14 domain (of both the wild type and some mutants associated with the onset of neurofibromatosis pathology) has revealed its peculiarity of being structurally coupled to a PH domain that strongly interacts with the lid helix through a long loop (called lid-lock loop). On this basis, a mechanism for the opening of the Sec14 lipid pocket was formulated which would involve a concerted movement of the lid-lock loop, but this movement has actually never been shown. Guided by available experimental data on the thermal denaturation of Sec14-PH domain of NF1, both on the wild type and some neurofibromin-related mutants, several simulations at high temperature were carried out to compare the dynamics of the wild type domain with a pathological mutant associated with the onset of neurofibromatosis. Our simulations lead us to suggest an opening mechanism for the lid helix and provide a hypothesis for the structural and dynamic basis of the onset of the disease in the case of the specific mutant.
The second project addressed the study of a protein called EtfAB which catalyzes a recently discovered process known as Flavin-Based Electron Bifurcation (FBEB). This mechanism is only exploited by some anaerobic microorganisms as a third way of energy coupling. So far, four unrelated protein families are known that are able to catalyze FBEB. Among these, EtfAB, catalyzes the electron transfer between the two FAD molecules bound to it. Surprisingly, the distance between these two FADs, as observed in the crystal structure of EtfAB, is 18 Å, whereas biological electron transfer is considered more likely to occur at a maximal distance of 14 Å. To explain this, a possible mechanism has been suggested that could bring the two FAD molecules closer together. Using molecular dynamics, it was possible to test, and discard, the proposed mechanism. Furthermore, with the Density Functional Theory (DFT), it was possible to provide an interpretation to some spectroscopic data regarding the possible electron transfer between the two FAD molecules.
In the third project, I collaborated with Prof. Luca Bertini on a project on the production and propagation of some reactive oxygen species (ROS) in the context of

the amyloid-beta peptide involved in the pathogenesis of Alzheimer's. In the amyloid hypothesis on the onset of Alzheimer's disease, an important role has been attributed to the damage caused by ROS, produced by a metal ion coordinated to the amyloid peptide itself, in particular by the hydroxyl radical (OH.-). However, the details of how these radicals propagate and react have not yet been clarified. While Prof. Bertini's DFT calculations addressed the oxidative capacities of the hydroxyl radical and the possible reaction products in the context of the amyloid-beta peptide, my molecular dynamics simulations provided an overview on which possible targets of the hydroxyl radical, coordinated to the ion Cu of the complex, could actually react with the hydroxyl radical due to the dynamic motions of the peptide.

Introduction

Molecular modelling

Computational modelling is the creation of a computable description of a system. This model must comprise a way to express the state of the system and a set of rules for transforming that state under the effect of specific perturbations. The main goal is to predict the dynamical behavior of the system and eventually simulate the results of an experiment without physically performing it in the lab, with the advantage of bypassing technical obstacles and the practical effect of saving time and, maybe, money.
A model may be, for example, a cell. Its state is the list of all (or a subset) of its constituents, with their own concentration, and the rules are a set of kinetic equations that describe the temporal evolution of the concentration of its constituent. This chemical approach is used for example in systems biology to study metabolic pathways or cellular processes.
Molecular modelling is the application of modelling principles to molecules. Molecules are actually physical systems whose description depends on the level of theory employed. The most fundamental level is the quantum level, in which the system is described in terms of quantum concepts, like wavefunctions, and obeying the rules of quantum mechanics (QM), like the Schrödinger equation and its derivative forms. QM is the most accurate level of theory we have to describe molecular systems, but its complexity makes it computationally expensive and feasible only for systems composed of up to a few tens or hundreds of atoms. It is usually employed to predict the shape of small molecules and study chemical reactions, being the only level of theory able to adequately treat electrons.

Basics of molecular mechanics

Another way to describe a molecular system is the molecular mechanics formalism. At this level of theory, the system is described with a classical mechanics language: the state is the complete set of positions and velocities of its constituents (atoms or groups of atoms). Molecular mechanics provides a way of expressing the potential energy (V) of a system as a function of the positions of its atoms (r). This potential energy V(r) is necessary to derive the forces acting on the system and eventually to predict its evolution according to some dynamical law, being it deterministic or probabilistic. This expression of the potential energy is called a force field.

Force fields describe the energy of a system as a combination of different kinds of interaction that its atoms establish with each other. These are usually divided into bonding and non-bonding terms.

$$V(\boldsymbol{r}) = V_{bonding} + V_{non-bonding} \tag{1}$$

$V_{bonding}$ is the sum of different energies contributions deriving from *bonds stretching*, *angles bending*, *improper* (or out-of-plane) *dihedrals stiffness*, and *torsions (dihedrals)*. The first three terms are expressed by a quadratic function of the respective coordinate while the torsion potential is a sinusoidal function to account for the multiplicity of rotamers.

$$\begin{aligned} V_{bonding} = & \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 \\ & + \sum_{\substack{improtper \\ dihedrals}} k_\varphi(\varphi - \varphi_0)^2 \\ & + \sum_{dihedrals} \sum_{1}^{6} k_{\Phi,n}(1 + \cos(n\Phi - \delta_n)) \end{aligned} \tag{2}$$

Non-bonding energy depends on the electrostatic interactions between charges and from the Van der Waals forces. The former is expressed by the classical coulombic potential while the latter are usually condensed in the Lennard-Jones potential.

$$\begin{aligned} V_{non-bonding} = & \sum_{\substack{coulomb\ pairs \\ i,j}} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \\ & + \sum_{\substack{vdw\ pairs \\ i,j}} \epsilon_{i,j} \left[\left(\frac{R_{\min i,j}}{r_{i,j}} \right)^{12} - 2 \left(\frac{R_{\min i,j}}{r_{i,j}} \right)^{6} \right] \end{aligned} \tag{3}$$

This is a "minimal force field", the essential elements that a molecular mechanical force field must have but many different force fields have been developed, each with its own specific form and set of parameters. For example, a force field may add an explicit term for hydrogen bonds or a correction term for angles bending (for example the Urey-Bradley potential in the CHARMM force fields). Each force field was developed to reproduce specific data, being them quantum mechanical calculations or experimental results, on specific targets: small organic molecules, biological macromolecules, or even

inorganic compounds and non-biological polymers. This means that each force field has its own area of application and for biological macromolecules the most popular are the AMBER and CHARMM families of force fields.

Molecular dynamics: concepts and principles

In classical physics the deterministic evolution of the state of a system is made possible by solving Newton's equations of motion:

$$m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2} = -\frac{\partial V}{\partial \boldsymbol{r}_i} \tag{4}$$

The simulation of a molecular system by numerical integration of the (4), when the V is represented by a force field, is called Molecular Dynamics (MD).
MD, by tracking the movement of every atom in the system, is one of the most powerful approaches we have to get deeper insights into the biochemical functions of macromolecules.
The output of a MD simulation is a trajectory which serves as a statistical ensemble of replicas of the system over the phase space (the set of all the possible positions and momenta of the system). From the ensemble average $\langle A_i \rangle$ of a given property *i* we can calculate the macroscopic thermodynamics observables that can be experimentally measured.

$$\langle A_i \rangle = \frac{\iint A e^{\beta H(\boldsymbol{q},\boldsymbol{p})}\, d\boldsymbol{q} d\boldsymbol{p}}{\iint e^{-\beta H(\boldsymbol{q},\boldsymbol{p})} d\boldsymbol{q} d\boldsymbol{p}} \tag{5}$$

This quantity can be tricky, if not impossible, to calculate but thanks to the ergodic theorem we can use molecular dynamics to compute the time average $\overline{A_i}$ of the same property and use it in place of the ensemble average, in the limit of a reasonable amount of time.

$$\overline{A_i} = \frac{1}{T} \int_0^T A[\boldsymbol{q}(t), \boldsymbol{p}(t)]\, dt \tag{6}$$

$$\lim_{t \to \infty} \overline{A_i} = \langle A_i \rangle \tag{7}$$

Depending on which thermodynamic variables are kept constant a statistical ensemble can be defined. To match the experimental conditions the most widely used in molecular dynamics are the canonical ensemble, NVT (constant number of particles, volume and temperature), and isothermal-isobaric ensemble, NPT (constant number of particles, pressure and temperature).
Molecular dynamics was first applied to an inorganic system in the '50 and to a biological system (the bovine pancreatic trypsin inhibitor) in the '70 by J.A. McCammon, B.R. Gelin and M. Karplus. It was simulated in the void for only 9 ps, but it was enough to show the fluid-like character of the interior of proteins and the diffusional mode of local atom motions.

Today, thanks to the constant increase in computational power that allows ever longer simulation times, up to the millisecond scale, MD can be applied to a variety of situations. For example, it can be used to test the accuracy, and refine, experimental structural models, ab initio prediction of small proteins fold, or a system response to the binding of a ligand or to a post-translational modification, or even the current flowing through a channel. Further, if the energy of the system is adequately tricked, it is also possible to overcome the problem of the "reasonable amount of time" and get free energy profiles of the so-called rare events or protein-protein binding.

Molecular dynamics: a typical setup

A MD simulation always starts with a structure of the molecule to study, possibly an experimentally resolved one, that provides the starting positions, and a force field chosen to describe the molecule. The molecule is then put into a simulation box that will be shaped so to be replicated infinite times in all directions by applying periodic boundary conditions. This is necessary to avoid the system to interact with void. Then water is put into the system and ions are added to neutralize net charges. Before the real simulation (or "productive run") starts, it is necessary to prepare the system in order to avoid subsequent instabilities. For this, an initial minimization step is usually required, mainly to remove possible defects in the starting structure that might result in unacceptable large force values that eventually cause the blowing up (or failure) of the simulation. After the minimization one or more equilibration steps are usually required to relax the system to the target values of temperature, pressure, or volume, depending on the chosen ensemble. Firstly, temperature is provided by assigning velocities to the atom according to the Maxwell distribution at the target temperature. Temperature control is achieved by an algorithm that mimics a thermostat, for example by rescaling the velocities at each integration step, even though other approaches are possible. Pressure control is usually by modulating the box sizes, and then the volume.

MD simulations are speeded up by reducing the number of interactions to calculate. Specific algorithms do that with the aim of not losing the accuracy of the simulation. For that, VdW interactions are usually cutoff beyond a given distance threshold and long-range electrostatics (beyond another distance cutoff) electrostatic are not treated explicitly, usually making use of the Particle Mesh Ewald (PME) approach.

First project: Molecular dynamics simulations reveal structural interconnection defects in mutant Sec14-PH bipartite domain from human neurofibromin

Authors: Fabio Rizza, Jacopo Vertemara, Renata Tisi.
This chapter is based on the manuscript we submitted to FEBS letters.

Abstract

Neurofibromin, the main RasGAP in the nervous system, is a 2818 aa protein with several poorly characterized functional domains. The only well-studied domains are the GTPase Regulating Domain (GRD) and the Sec14-PH domain, which was found to mediate protein-protein interactions. Mutations in NF1 encoding gene lead to an autosomal dominant syndrome, neurofibromatosis, with an incidence of 1 out of 3000 newborns. Missense mutations are not always localized in the GAP domain but spread all over the exons, in the Sec14-PH encoding sequences as well. The molecular issues caused by these pathogenic mutations have yet to be characterized and structural data could not highlight the defect in mutant K1750Δ Sec14-PH functionality. In order to investigate the dynamic properties of the wild type and a mutant version of NF1 Sec14-PH domain, we performed molecular dynamics simulations either of the PH module alone or of the entire domain at different temperatures. The wild-type Sec14-PH domain showed global overall stability throughout all the replicas from 300K to 400K, with only minor changes due to local secondary structure loss at the higher temperatures. On the other hand, an increase in the flexibility in the lid-lock loop was observed for the K1750Δ mutant already at 380K. This led to the loss of the tight binding between the two subdomains with consequent complete impairment of the module fold. Our data demonstrate that the lid-lock is fundamental for the structural interdependence of NF1 bipartite Sec14-PH domain.

Introduction

Neurofibromatosis type 1 (NF-1) is an autosomal dominant disease with an incidence of 1 over 3000 births arising from mutations in the Neurofibromin encoding gene, NF1. Affected patients manifest typical symptoms like cafe-au-lait spots on the skin and are likely to develop tumors in the peripheral nervous system and learning disabilities [1].
Neurofibromin (NF1) is a 2818 residue protein, mainly expressed in the brain, whose pivotal function is to act as a GAP (GTPase-activating Protein) in order to quench growth signals [2]. This function is carried out by a 300-residue domain located in a central region of the sequence and called GRD (GAP-Related Domain) that provides a binding groove for Ras and a typical arginine finger that assists GTP hydrolysis in the active site of Ras: detailed insights into the mechanism of the process came from both structural and biochemical studies on the isolated GRD [3]. Apart from the catalytic GRD, Neurofibromin is a poorly understood protein. There are several known cellular partners, but the biological meaning and molecular features of most interactions have not been elucidated yet [4,5]. Besides, NF1-associated mutations are widespread in the gene rather than focused on the GAP-domain encoding region [6].
Very little information is available about neurofibromin structural features. The only structurally investigated domains up to date are the GRD and the Sec14-PH domains, located just downstream of the GRD, due to the difficulty in purifying other portions of the protein [7].
The X-ray structure determination of the bipartite Sec14-PH, which hosts many of the NF1 associated mutations, was unfortunately not resolutive to understand the mechanism at the base of the pathogenicity of the investigated mutations [8]. As the name suggests, this region consists of two subdomains, structurally similar to two well-known folds, a Sec14 and a PH module. The first Sec14-containing protein was characterized in yeast as a lipid exchanger between membranes. Many proteins have hitherto been characterized in eukaryotes that possess a Sec14 module [9]. It consists of a globular fold with a lipid cage which is gated by an α-helix (also referred to as the lid helix) which can undergo an extended conformational change, as it is suggested by both crystal structures and simulations [10,11]. The PH domain is a typical protein module involved in signal transduction. PH domains from different proteins were reported to either bind lipids or mediate protein-protein interactions [12].
In the context of neurofibromin, the Sec14 and the PH domain are fused in a unique 250 residue structure. Its main feature is the structural interaction between the two parts. A long-loop protrusion (also called lid-lock) shaped as a beta-hairpin, in the PH module bends upon the lid helix in the Sec14 forcing it to adopt a closed conformation with a lipid molecule stuck inside [8]. Since the NF1 Sec14-PH domain can exchange lipids in

vitro, its unique structure raises the question of the mechanism of the process [13]. Indeed, for it to achieve this task an extended movement of the beta protrusion has been proposed for a concerted opening of the gating helix in the Sec14 domain. It has also been suggested that the water-exposed surface of the beta protrusion might serve as a docking site for a binding partner which could in turn regulate the process [8], [14]. Hitherto, the only characterized functions for the Sec14PH have been in protein-protein interaction. In fact, Sec14-PH has been characterized in vitro and in vivo as the interacting partner of the endogenous opioid μ- (MORs) receptors activated Gβγ subunits, which are able to inhibit NF1 RasGAP function [15]. The PH module was reported to drive neurofibromin interaction with the serotonin 5 hydroxytryptamine 6 (5-HT6) receptor, in order to regulate the basal activity of the receptor itself [16].

Based on the difficulty in expressing the protein, the structures of only three mutant Sec14-PH domains have been solved: a missense mutant (I1584V), a deletion mutant (K1750 deleted), and a tandem repeat mutant (TD). None of these mutants show any dramatic effect on the global fold of the module, but only minor local changes [8]. The I1584V mutant is structurally identical to the wild type and exhibits the same thermal denaturation behavior. The K1750Δ carries a deletion in the lid-lock protrusion resulting in a local backbone deviation and a flip in the orientation of the side chains of a few subsequent residues. In contrast with the WT and the I1584V, the K1750Δ mutant showed a lower denaturation temperature. The TD mutant carries a duplication in the linker between the Sec14 and the PH subdomains resulting in a longer loop that could not be solved by the X-ray data. Its thermal behavior resembles that of the K1750Δ.

To date, the concerted motion of the lid-lock and the lid helix is the proposed model for the lipid exchange function of this module. Starting from crystal structures and guided by thermal denaturation data we performed molecular dynamics simulations to compare the dynamics of the wild type with the K1750Δ mutant, which revealed that the lid-lock is fundamental for structural integrality of the bipartite Sec14PH module.

Materials and methods

Systems setup

The coordinates of the neurofibromin Sec14-PH domain were obtained from the Protein Data Bank [17] (PDB ID: 2E2X for the wild type [13] and PDB ID: 3PG7 for the K1750Δ mutant [8]). Since the two crystal structures started at different positions, the initial residues were removed so that both WT and mutant started at L1569. Starting point coordinates for the PH domain alone were obtained from the correspondent Sec14-PH structure by removing the Sec14 moiety (from 1569 to 1714). For all systems, N-terminus was then capped with an acetyl (ACE) group and the C-terminus was capped with N-methyl amide (NME) group. All systems setup and subsequent simulations were performed with Gromacs 2018.2 [18]. Hydrogen atoms to both amino acid residues and water molecules were added using the pdb2gmx module of Gromacs. PH domains were solvated in a cubic box with dimensions of 73.66 x 73.66 x 73.66 Å and filled with CHARMM-TIP3P water molecules [19]. The genion module was also used to add sodium and chloride ions at 150 mM [NaCl] to neutralize the systems. Sec14-PH proteins were solvated in a cubic box with dimensions of 92.06, 92.06, 92.06 Å, filled with CHARMM-TIP3P water molecules, and neutralized with 150 mM NaCl.

Molecular dynamics simulations and analysis

All MD simulations were performed with the mdrun module of GROMACS 2018.2 with the CHARMM36m-nov2018 forcefield [20] for protein and the Charmm General Force Field (CGenFF) [21] for the lipid molecule. For PH domains, a single simulation at 300K was performed for both the WT and the mutant. For the Sec14-PH domains six independent simulations, each with a different temperature, were performed for both the WT and the mutant (300K, 330K, 340K, 360K, 380K, 400K). Before all the production simulations, the systems were minimized with the steepest descent method until maximum force went below 1 kJ/mol/nm. The minimization was followed by a 1 ns NVT equilibration phase at the target temperature with the V-rescale algorithm [22] to stabilize the temperature and a 1 ns NPT equilibration phase with the Berendsen [23] algorithm to stabilize the pressure (1 atm). Both equilibration phases were performed with a 1 fs time-step and with position restraints on heavy atoms of protein and lipid. The production runs continued in the NPT ensemble without position restraints for 1 μs, with the V-rescale method for stabilizing the temperature, and the Parrinello-Rahman barostat [24] for the pressure (1 atm). Snapshots of the coordinates were written out every 100 ps. A 2-fs time step was used throughout the production simulations. For both short-range electrostatics and Lennard–Jones a 12 Å cutoff was

applied, in the latter case with a force-switch at 10 Å. Long-range electrostatic interactions were calculated with the Particle Mesh Ewald method [25,26] with a grid space of 12 Å and bonds involving hydrogen atoms were constrained using the LINCS algorithm [27]. Periodic boundary conditions were applied. All computation was performed on the CINECA Marconi supercomputer.

All subsequent analysis (RMSD, RMSF, Cluster analysis, Principal Component Analysis, distances, and secondary structures) was performed with Gromacs 2018.2 on the 10000 frames trajectories coming out of the simulations. Graphs were plotted with Xmgrace. All protein images and videos were rendered with Ucsf Chimera 1.14 [28].

Results

PH portion of Sec14PH domain is independently stable and shows scarce flexibility

Preliminary simulations at 300 K on the entire Sec14-PH module, for both WT and K1750Δ mutant (data not shown) did not reveal any hypothesized functional motion in the PH domain.
Expression of an isolated WT neurofibromin PH domain has been shown to rescue 5-HT6 serotonin receptor-operated cAMP signaling in mutant cells carrying specific NF1-associated mutations in the PH sequence [16]. We then postulated that it could retain structural integrity albeit in the absence of the Sec14 moiety.
Molecular dynamics simulations were performed for a total of 1 μs each on the WT and the K1750Δ isolated PH domain, in order to evaluate the effect of K1750 deletion in the lid-lock protrusion (alias lid-lock loop) on the intrinsic flexibility of the isolated PH domain at 300 K.

Wild Type

Except for the C-terminal tail, the lid-lock beta protrusion is the most flexible region (Fig. 1A, movie S1) in the PH portion of the Sec14-PH domain. To track the position of the lid-lock loop relative to the core, we calculated the Root Mean Square Deviation (RMSD) of the backbone atoms of the loop throughout the simulation, after eliminating the roto-translational motions by superposition of all backbone atoms on the first frame of the trajectory. The protein behaves as a two-state system: a dominant conformation at RMSD of about 4 Å from the starting structure alternates with another one with a RMSD of 6/8 Å (Fig. 1B). A cluster analysis based on the loop RMSD relative to the whole protein backbone (with a cut-off of 2.5 Å) elicited a series of representative snapshots (cluster centroids) of the simulation. Fig. S1 and Table S1 show the distribution and numerosity of the 25 resulting clusters. The first five clusters cover 92% of the structures sampled by the simulation and therefore their centroids capture the main conformations of the protein. Cluster 1 (70% of total frames) dominates the trajectory and alternates with cluster 2 (9% representativeness), which is enriched when the RMSD is higher. Centroids for these two clusters can be considered the models for state 1 and state 2, respectively (Fig. 1C). These two states differ in the conformation of the N-terminal region of the lid-lock loop. In cluster 1, similarly to the crystal structure, the hydrogen bond between backbone NH of A1746 with backbone carbonyl O of K1731 and the hydrogen bond between backbone NH of S1733 with backbone carbonyl O of Thr1744 stabilize the lid-lock conformation. In cluster 2, the former bond

is broken and the second is substituted by the H-bond between NH of S1733 with the sidechain hydroxyl OH of T1744. This conformational change is due to a rotation in the φ angle of T1744 and results in a slight rotation of the loop towards the outside (Fig. 1C, see detail).

Hydrophobic contacts and two specific hydrogen bonds between the lid-lock loop and the lid helix in the Sec-14 domain are lost in the isolated PH domain, resulting in a loop slightly rotated in cluster 1, suggesting that the Sec14 portion is involved in stabilizing the conformation of the loop observed in the Sec14-PH crystal. This little "left-right" swing of the lid-lock loop, seen from the perspective of the (missing) Sec14 domain, represents the whole intrinsic dynamics of the PH domain in the absence of the Sec14 domain, indicating a generally scarce flexibility in this portion of NF1 protein.

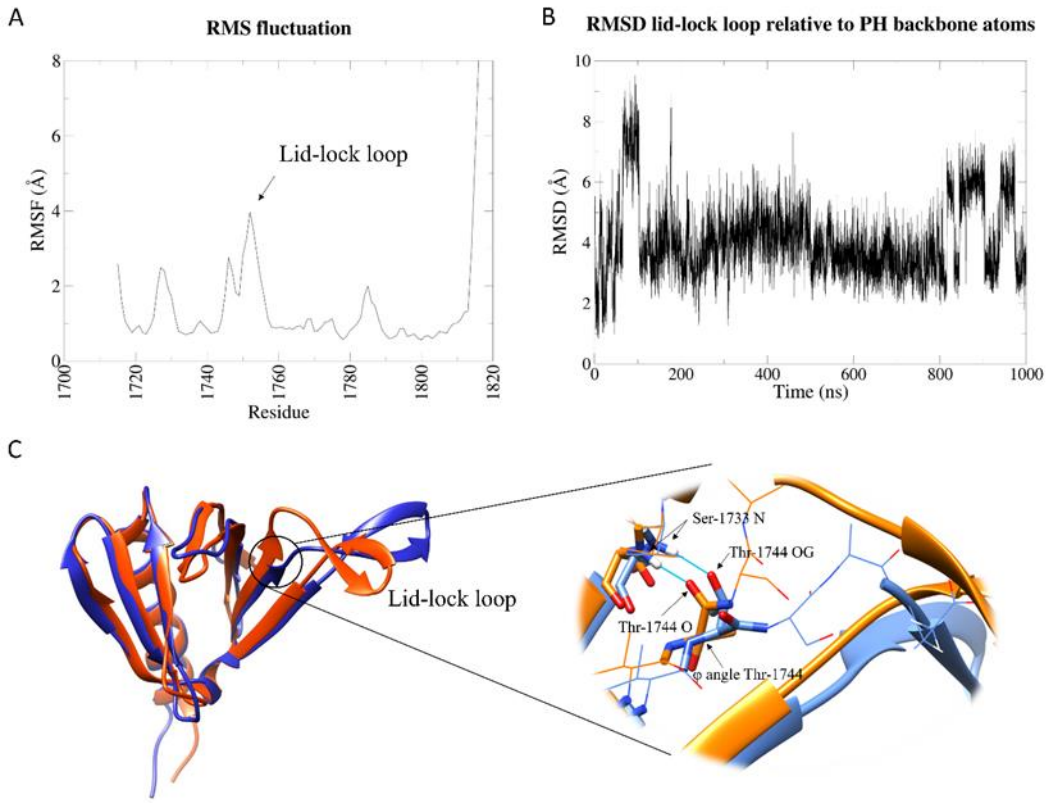


Fig. 1 | Analysis of the WT PH domain simulation. A: Root Mean Square Fluctuation of backbone atoms of the PH domain. B: Root Mean Squared Deviation of the backbone atoms of the lid-lock loop after superposition of the all the backbone atoms of the domain. C (left): cluster 1, orange, and cluster 2, blue, superimposed. The two structures differ by a conformational change at the N-terminal part of the lid-lock loop, resulting in a left-right swing of the loop. C (right): zoom on the structural differences between cluster 1 and cluster 2. The φ angle of Thr-144 is rotated in cluster 2 resulting in a changing in the direction of the loop. The carbonyl oxygen of Thr-1744 moves away and is substituted by the hydroxyl oxygen (OG) of Thr-1744 sidechain in forming h-bond with Ser-1733.

K1750Δ

Differently from the wild type protein, the K1750Δ mutant shows a higher intrinsic flexibility in the lid-lock loop in the isolated PH domain (Fig 2A, movie S2). According to the crystal structure of the mutant [8], the K1750 deletion only causes local structural changes. Due to the shortening of the loop, A1746 in the mutant is not structurally aligned with A1746 in the WT, causing the weakening of the H-bond between backbone NH of A1746 and backbone carbonyl O of K1731. As mentioned, this interaction was found to stabilize state 1 in the WT simulation. Furthermore, starting from E1747 the

structural alignment is recovered but the sequences are shifted by one position up to WT K1750 and mutant T1749. Taken together these two structural differences observed in the crystal structures of the WT and mutant PH domain, a higher flexibility of the mutant is expected, especially in the absence of the Sec14 domain.

While retaining the general compactness of the domain core, in the mutant the beta-protrusion explores a higher number of different conformations compared to the wild type, some of them compatible with the formerly invoked open configuration [14], as suggested by the lid-lock loop RMSD relative to the whole PH domain (Fig. 2B). Cluster analysis with a 4 Å cut-off reveals many different conformations are explored during the simulation (Fig. 2C). Like for the WT, a dominating cluster 1 (64% of the total) resembles the crystal structure. Cluster 4 represents a structure analogous to WT state 2, with the same conformational change in the N-terminal region of the lid-lock loop, but the dynamics of the mutant protein cannot be described by a two-state model. Cluster 2 and 3 are the largest among the clusters that explore conformations that could not be observed in the WT protein dynamics. In particular, the most dramatic conformational change is shown in cluster 3, which is mainly explored between 600 ns and 700 ns. In the centroid structure, the lid-lock loop moves upwards and flips its bending orientation.

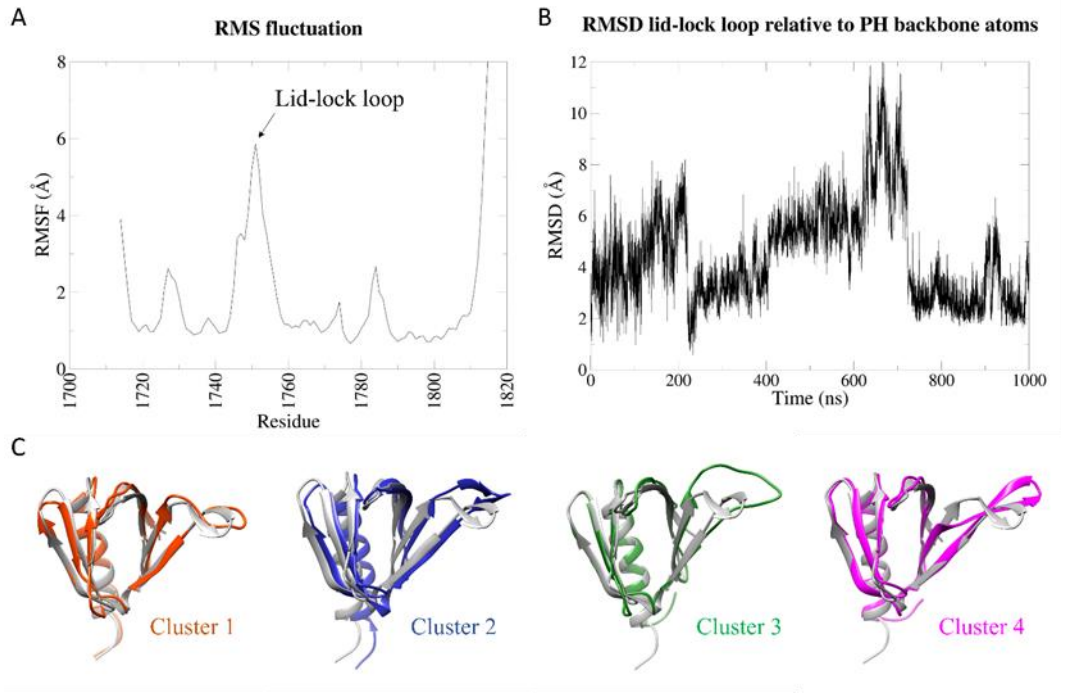


Fig. 2 | Analysis of the K1750Δ PH domain simulation. A: Root Mean Square Fluctuation of backbone atoms of the PH domain. B: Root Mean Squared Deviation of the backbone atoms of the lid-lock loop after superposition of the all the backbone atoms of the domain. C: centroids of the first four clusters aligned against the crystal structure. Cluster 1 and cluster 4 resemble state 1 and state 2 of the WT and the left-right swing due to conformational change in the beta strand before the loop. Cluster 2 and cluster 3 show the "up motion" of the loop with the flip in the bending orientation (cluster 3) that could not be observed in the WT.

Dynamic behavior of the whole wild-type Sec14-PH domain

Due to the interactions between the lid-lock loop with the Sec14 upper lid helix, any of the intrinsic dynamics similar to the ones observed for the isolated PH domain could be observed in a preliminary 1 μs-MD simulation at 300K for the whole Sec14-PH domain (movie S3). In order to explore more possible states for the Sec14-PH domain, we performed a series of simulations at progressively higher temperatures (300K, 330K, 340K, 360K, 380K, 400K). The molecule of phosphatidylethanolamine buried in the Sec14 lipid pocket in the crystal was maintained.

The Root Mean Square Fluctuation (RMSF) of the wt Sec14-PH domain shows an increased flexibility within specific regions when temperature raises to 400K (Fig. 3A), in line with a loss in secondary structure content in those regions corresponding to peaks in RMSF. Accordingly, the RMSD shows a similar pattern throughout the

simulation from 300K to 380K (Fig. 3B) (compare with movies S3 and S4), but at 400K it seems to deviate more from the starting conformation (movie S5). The most affected part was the linker peptide between Sec14 and PH subdomains (aa 1700-1714). Overall, the PH module (aa 1715-1816) shows to retain its structure, while Sec14 appears more susceptible to temperature: in fact, increased flexibility can be observed in the helices that constitute the lipid cage. Lid-lock loop did not show the same flexibility observed in the simulations of isolated PH domain. In the context of the whole module, the lid-lock loop appears to be one of the more stable regions, due to the stabilizing interaction with the Sec14 upper lid helix.

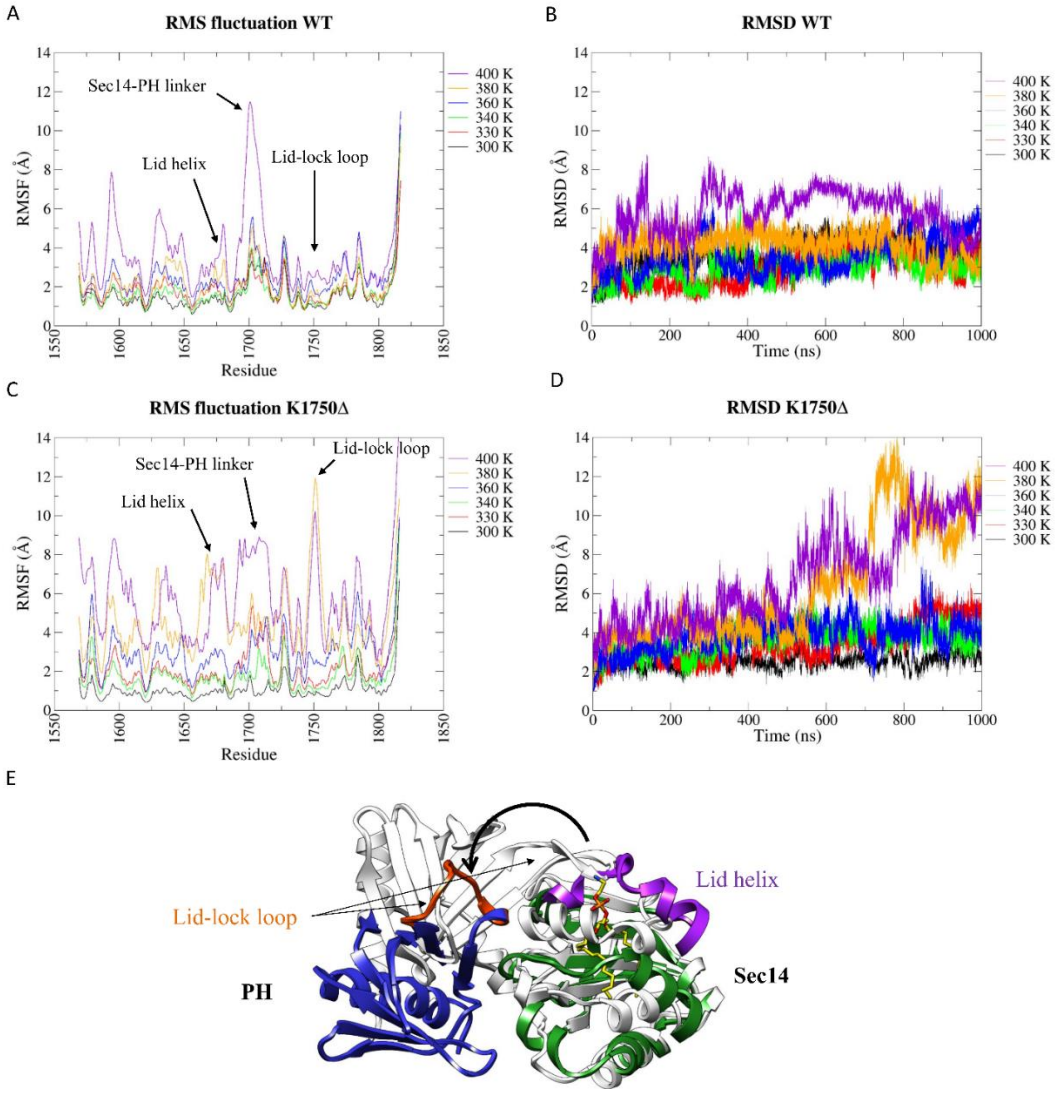


Fig. 3 | Analysis of molecular simulations at increasing temperatures. A: Root Mean Square Fluctuation of WT Sec14-PH. B: Root Mean Square Fluctuation of mutant Sec14-PH. C: Root Mean Square Deviation of WT Sec14-PH. D: Root Mean Square Deviation of WT Sec14-PH. E: a frame from the 380K simulation of the mutant Sec14-PH. The lid-lock loop lost its secondary structure and consequently the PH domain detached from the Sec14 moiety.

The K1750Δ mutation weakens the interconnection between the two portions of the bipartite Sec14-PH domain

A completely different behavior was observed during the MD simulation for the K1750Δ mutant Sec14-PH domain. A dramatic increase in flexibility was observed already at 380K and furthermore at 400K, as can be observed in the RMSF graphs (Fig. 3C) (compare movies S7 and S8 to S6). Differently from the wild type, the RMSD plots at both 380K and 400K show that the tertiary structure dramatically deviated from the starting conformation (crystal structure) at around 500 ns (Fig. 3D). Interestingly, in both the simulations at these temperatures, the lid lock loop has a pick in flexibility (Fig. 3C). These results agree with experimental thermal denaturation data reported in literature [8], describing a CD signal, suggesting a lower melting temperature for the mutant relative to the wt. The analysis of the secondary structure percentage in the structures explored during the 380K and 400K simulations confirms that the K1750Δ mutant Sec14-PH loses more secondary structures than the wild-type (Fig. S3).
The analysis of the structures sampled in the second half of the simulations performed at 380K and 400K revealed that the lid-lock loop exhibits the same dynamic behavior as in the isolated PH domain at room temperature, with a flip in the bending direction (Fig. 3E). This resulted in the loss of contacts between Sec14 and PH portions and, eventually, in the physical separation between the two subdomains (see movies S7 and S8).

Opening of the lipid pocket

Since Sec14 domains are associated with lipid exchange both in vitro and in vivo, we took advantage of our simulations to check for a structure compatible with an open conformation, suited for lipid exchange. Ryan and collaborators [11] used molecular dynamics to study the closing of the lid helix in the yeast Sec14p which was crystallized in the open conformation (PDB 1AUA [10]). They monitored the distance between the alpha-carbon of Arg-195 and Phe-231 (on the lid helix). In the crystal structure that distance was 23 Å and in their simulation it dropped to around 10 Å in agreement with a closed conformation. A closed conformation was crystallized later for the yest Sec14 homologue protein (Sfh1, PDB 3B74) revealing a distance between Lys-197 and Phe-233 (corresponding to Arg-195 and Phe-231 on Sec14p, respectively) of 6.9 Å. Structural alignment of Nf1-Sec on Sfh1 shows that the corresponding residues to R195/K197 and F231/F233 are Ser-1639 (on α3 helix) and Leu-1675 (on the lid helix), respectively, whose alpha-carbon distance is 6.6 Å. Inspired by the work of Ryan and collaborators, we applied the same approach on neurofibromin.

Wild type

Fig. S4 shows the distributions of the distance between S1639Cα and L1675Cα for the WT. From 300K to 360K we can see a sharp distribution with a peak at 7.5 Å, close to the 6.6 Å value of the crystal structure, in agreement with the overall stability suggested by RMSD and RMSF graphs. Rising the temperature to 380K and 400K the pick shifts towards around 12 Å with more spread distributions.
At 400K the S1639Cα--L1675Cα distance reaches values even over 23 Å, but they are mainly associated with loss in secondary structures that occur at 400K, causing large distortions in the internal coordinate of the protein. In the WT those frames in which secondary structures are preserved never show a S1639Cα--L1675Cα distance over 20 Å. Anyway, we could select a representative structure for a putative open conformation with a 19 Å opening (fig. 4). The lid-lock loop is slightly shifted relative to its position in the crystal structure and moves in a fashion like that observed in the PH domain alone.

K1750⊿

The mutant protein follows the same trend as the WT regarding the opening of the lipid cage. Up to 360K there is no significant deviation from the starting value (fig 4). At 380K and 400K the pattern is like the WT, with the distributions getting more spread. As shown earlier, the mutant domain suffers from high instability at 380K and 400K with the lid-lock loop opening and the PH domain losing contact with Sec14. This in turn causes a drop in the stability of many helical segments of the Sec14, which unfold to coiled coil. Anyway, before the conformational transition occurred it was possible to pick some frames that satisfied the helical component of the lipid cage and a S1639Cα--L1675Cα distance compatible with a hypothesized open conformation. A representative frame in which such distance is 25 Å is shown in fig. 4. We can see that also for the mutant protein the lid-lock loop does not need to alter its internal conformation but is moved away from its original position by a rigid body rotation of the PH domain.

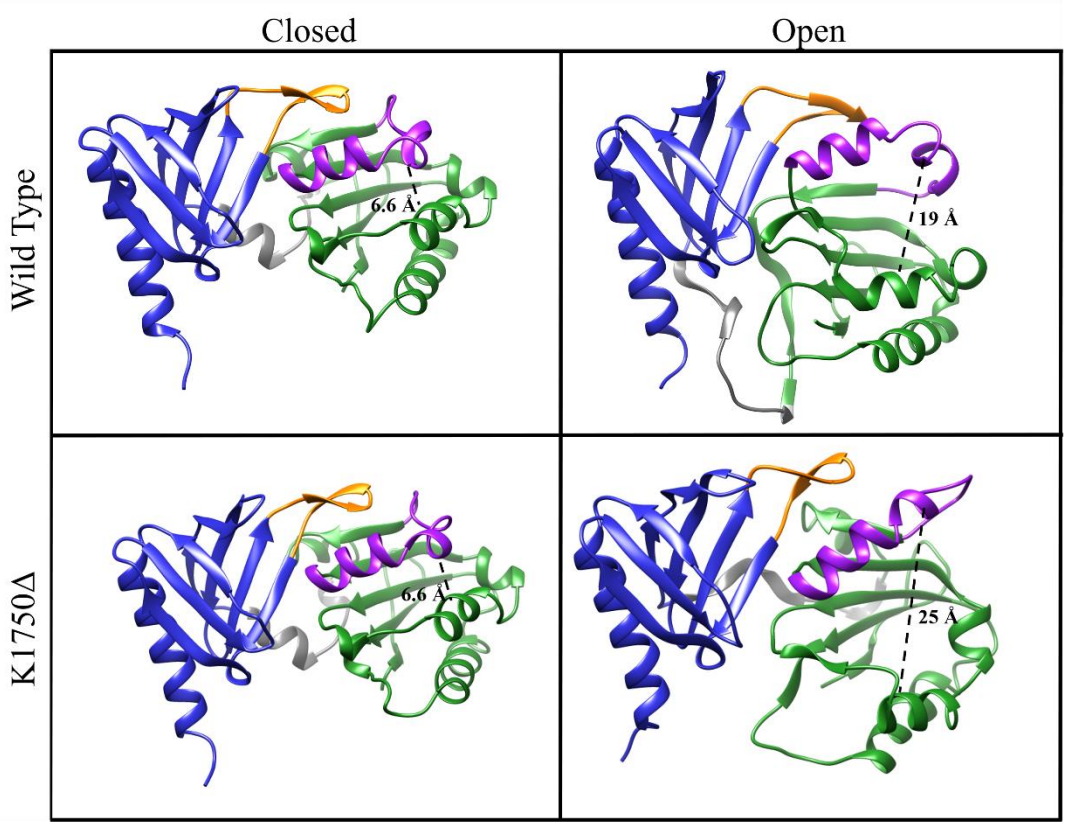


Fig. 4 | Comparison of the closed and open conformation for the Sec14 domain of neurofibromin. The distance between S1639Cα and L1675Cα is shown with a dashed line. Those frames were chosen from trajectories at 400K in which they were satisfied both the secondary structure content, especially in the helices of the lipid pocket, and a S1639Cα--L1675Cα distance compatible with an open conformation, using as a reference the 23 Å in the yeast Sec14p.

Discussion

The molecular bases of the pathogenesis of neurofibromatosis have been thoroughly investigated in the last years, thanks to the availability of several animal models [29]. Unfortunately, no progress was made in understanding the molecular mechanisms regulating NF1 functions, including RasGAP activity, with the outcome that effective therapies are still unavailable. This is mostly related to the intrinsic difficulty in NF1 manipulation, due to the huge size, the unknown structure of most of the protein, the impossibility to express and purify full-length or large fragments in E. coli, due to the incompatibility of the full cDNA with this host [7]. Structural data on NF1 protein portions are consequently scarce and precious.

Scheffzeck and co-workers [8] applied circular dichroism to investigate thermal stability of both WT and a series of mutants in Sec14-PH domain and concluded that the mutant K1750Δ loses its structure at a lower temperature relative to the wt. Our results confirm that, although the mutant K1750Δ Sec14-PH domain can adopt a conformation similar to the wild-type domain, which is proved by structural data and confirmed by the analyses of the conformations explored during the simulation, the mutation still provokes increased flexibility in the lid-lock loop, both in the isolated PH domain (Fig. 2A and B) and in the whole Sec14-PH domain (Fig. 3C and D). This flexibility could impinge on the overall stability of the Sec14-PH domain, both in terms of secondary structure loss and most importantly in terms of structural dependence of the two subdomains. In fact, while WT PH domain is characterized by a structurally robustness both alone and within the bipartite domain (Fig. 1A and Fig. 3A), the Sec14 moiety is characterized by a higher flexibility which matches the flexibility of the lid-lock loop in the mutant structure (Fig. 3A and 3B). The loss of contacts between the PH lid-lock loop and the Sec14 lid helix has consequences on the stability of the lid helix and of the overall stability of the Sec14 lipid cage. Taken together, our results suggest that the structural integrity of the Sec14-PH domain requires a tight interconnection within the two modules, and that its loss severely damages structural integrity of Sec14 moiety as well. It is worth noting that a pathogenic mutant with a tandem repeat insertion in the linker between Sec14 and PH is known. Given our results, we hypothesize that altering the flexibility of this linker may affect the tight structural interconnection between Sec14 and PH domain and so the overall stability of the Sec14-PH module.

Finally, our simulations suggest that the main role of the lid-lock loop would be to retain the global fold of the double module rather than to modulate the putative lipid exchange function, as previously proposed [14]. As a matter of fact, neurofibromin Sec14 subdomain was shown to exert lipid exchange activity in vitro [13], but no biological function was hitherto attributed to this activity. In contrast, physiological protein-

protein functions were reported for either the Sec14-PH domain or the PH subdomain itself [15,16] that would be dramatically affected by a loss in the overall Sec14-PH domain conformation sturdiness. The observed lipid exchange activity would require the opening of the Sec14 module, but this would rather be achieved without disturbing the connection of the lid-lock with the lid helix, with a joint motion of the whole PH domain together with the lid helix with respect to the rest of Sec14 lipid cage.

References

1. Brosseau J-P, Liao C-P & Le LQ (2020) Translating current basic research into future therapies for neurofibromatosis type 1. *Br J Cancer* **123**, 178–186.

2. Martin GA, Viskochil D, Bollag G, McCabe PC, Crosier WJ, Haubruck H, Conroy L, Clark R, O'Connell P & Cawthon RM (1990) The GAP-related domain of the neurofibromatosis type 1 gene product interacts with ras p21. *Cell* **63**, 843–849.

3. Rabara D, Tran TH, Dharmaiah S, Stephens RM, McCormick F, Simanshu DK & Holderfield M (2019) KRAS G13D sensitivity to neurofibromin-mediated GTP hydrolysis. *Proc Natl Acad Sci U S A* **116**, 22122–22131.

4. Kobayashi D, Tokuda T, Sato K, Okanishi H, Nagayama M, Hirayama-Kurogi M, Ohtsuki S & Araki N (2019) Identification of a Specific Translational Machinery via TCTP-EF1A2 Interaction Regulating NF1-associated Tumor Growth by Affinity Purification and Data-independent Mass Spectrometry Acquisition (AP-DIA). *Mol Cell Proteomics* **18**, 245–262.

5. Li X, Gao M, Choi JM, Kim B-J, Zhou M-T, Chen Z, Jain AN, Jung SY, Yuan J, Wang W, Wang Y & Chen J (2017) Clustered, Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9-coupled Affinity Purification/Mass Spectrometry Analysis Revealed a Novel Role of Neurofibromin in mTOR Signaling. *Mol Cell Proteomics* **16**, 594–607.

6. D'Angelo F, Ceccarelli M, Tala, Garofano L, Zhang J, Frattini V, Caruso FP, Lewis G, Alfaro KD, Bauchet L, Berzero G, Cachia D, Cangiano M, Capelle L, de Groot J, DiMeco F, Ducray F, Farah W, Finocchiaro G, Goutagny S, Kamiya-Matsuoka C, Lavarino C, Loiseau H, Lorgis V, Marras CE, McCutcheon I, Nam D-H, Ronchi S, Saletti V, Seizeur R, Slopis J, Suñol M, Vandenbos F, Varlet P, Vidaud D, Watts C, Tabar V, Reuss DE, Kim S-K, Meyronet D, Mokhtari K, Salvador H, Bhat KP, Eoli M, Sanson M, Lasorella A & Iavarone A (2019) The molecular landscape of glioma in patients with Neurofibromatosis 1. *Nat Med* **25**, 176–187.

7. Bonneau F, Lenherr ED, Pena V, Hart DJ & Scheffzek K (2009) Solubility survey of fragments of the neurofibromatosis type 1 protein neurofibromin. *Protein Expr Purif* **65**, 30–37.

8. Welti S, Kühn S, D'Angelo I, Brügger B, Kaufmann D & Scheffzek K (2011) Structural and biochemical consequences of NF1 associated nontruncating mutations in the Sec14-PH module of neurofibromin. *Hum Mutat* **32**, 191–197.

9. Saito K, Tautz L & Mustelin T (2007) The lipid-binding SEC14 domain. *Biochim Biophys Acta* **1771**, 719–726.

10. Sha B, Phillips SE, Bankaitis VA & Luo M (1998) Crystal structure of the Saccharomyces cerevisiae phosphatidylinositol- transfer protein. *Nature* **391**, 506–510.

11. Ryan MM, Temple BRS, Phillips SE & Bankaitis VA (2007) Conformational dynamics of the major yeast phosphatidylinositol transfer protein sec14p: insight into the mechanisms of phospholipid exchange and diseases of sec14p-like protein deficiencies. *Mol Biol Cell* **18**, 1928–1942.

12. Balla T (2005) Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci* **118**, 2093–2104.

13. Welti S, Fraterman S, D'Angelo I, Wilm M & Scheffzek K (2007) The sec14 homology module of neurofibromin binds cellular glycerophospholipids: mass spectrometry and structure of a lipid complex. *J Mol Biol* **366**, 551–562.

14. D'Angelo I, Welti S, Bonneau F & Scheffzek K (2006) A novel bipartite phospholipid-binding module in the neurofibromatosis type 1 protein. *EMBO Rep* **7**, 174–179.

15. Xie K, Colgan LA, Dao MT, Muntean BS, Sutton LP, Orlandi C, Boye SL, Boye SE, Shih C-C, Li Y, Xu B, Smith RG, Yasuda R & Martemyanov KA (2016) NF1 Is a Direct G Protein Effector Essential for Opioid Signaling to Ras in the Striatum. *Curr Biol* **26**, 2992–3003.

16. Deraredj Nadim W, Chaumont-Dubel S, Madouri F, Cobret L, De Tauzia M-L, Zajdel P, Bénédetti H, Marin P & Morisset-Lopez S (2016) Physical interaction between neurofibromin and serotonin 5-HT6 receptor promotes receptor constitutive activity. *Proc Natl Acad Sci U S A* **113**, 12310–12315.

17. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242.

18. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25.

19. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FT, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D & Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**, 3586–3616.

20. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmüller H & MacKerell AD Jr (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **14**, 71–73.

21. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I & Mackerell AD (2009) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, NA–NA.

22. Bussi G, Donadio D & Parrinello M (2007) Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126**, 014101.

23. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81**, 3684–3690.

24. Parrinello M & Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182–7190.

25. Darden T, York D & Pedersen L (1993) Particle mesh Ewald: AnN·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089–10092.

26. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H & Pedersen LG (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577–8593.

27. Hess B, Bekker H, Berendsen HJC & Johannes G E (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463–1472.

28. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.

29. Nix JS, Blakeley J & Rodriguez FJ (2020) An update on the central nervous system manifestations of neurofibromatosis type 1. *Acta Neuropathol* **139**, 625–641.

Supplementary materials

Table S1. Number of structures in each cluster for the wild type isolated PH domain. Cluster based on backbone atoms RMSD with a cutoff of 2.5 Å.

Cluster	# of structures
1	6998
2	936
3	667
4	466
5	177
6	149
7	138
8	126
9	125
10	52
11	29
12	28
13	22
14	19
15	18
16	17
17	11
18	6
19	5
20	3
21	3
22	2
23	2
24	1
25	1

Table S2. Number of structures in each cluster for the K1750Δ isolated PH domain. Cluster based on backbone atoms RMSD with a cutoff of 4 Å.

Cluster	# of structures
1	6372
2	842
3	707
4	586
5	347
6	245
7	207
8	203
9	116
10	99
11	53
12	43
13	32
14	26
15	21
16	17
17	15
18	14
19	12
20	10
21	8
22	5
23	5
24	3
25	3
26	3
27	3
28	2
29	1
30	1

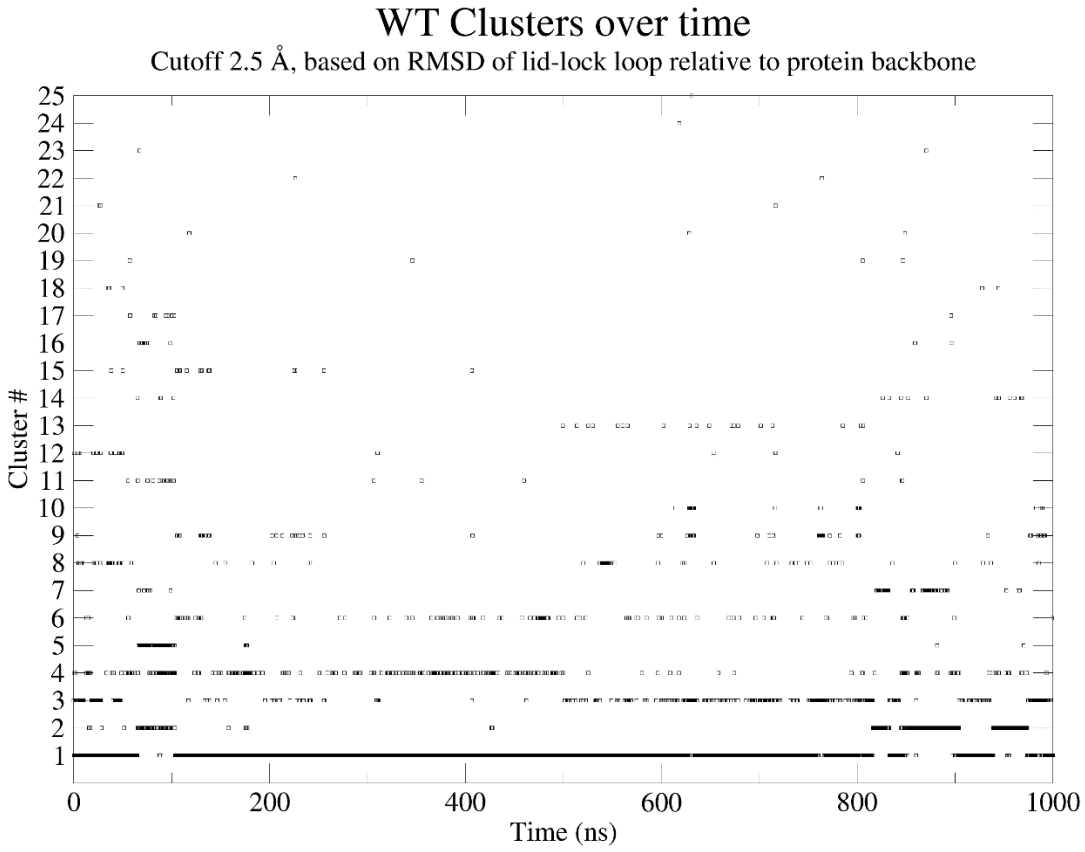


Fig S1 | Wild Type clusters distribution over the course of the simulation.

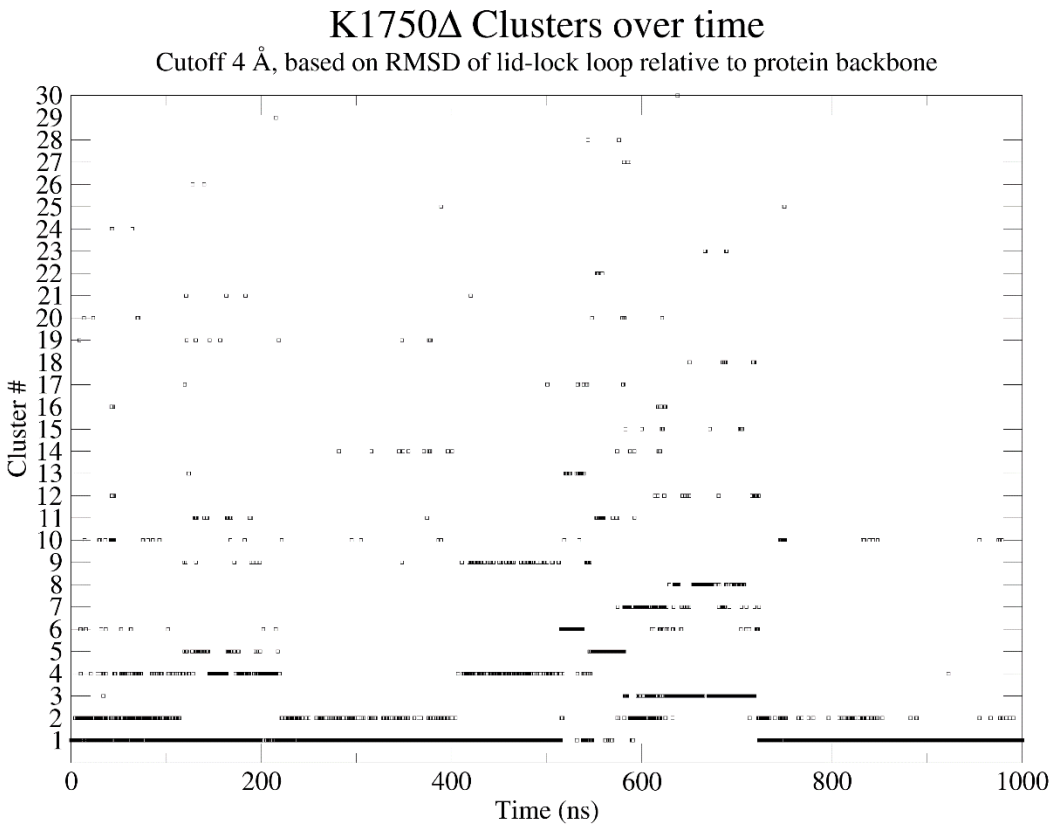


Fig S2 | Mutant K1750Δ clusters distribution over the course of the simulation.

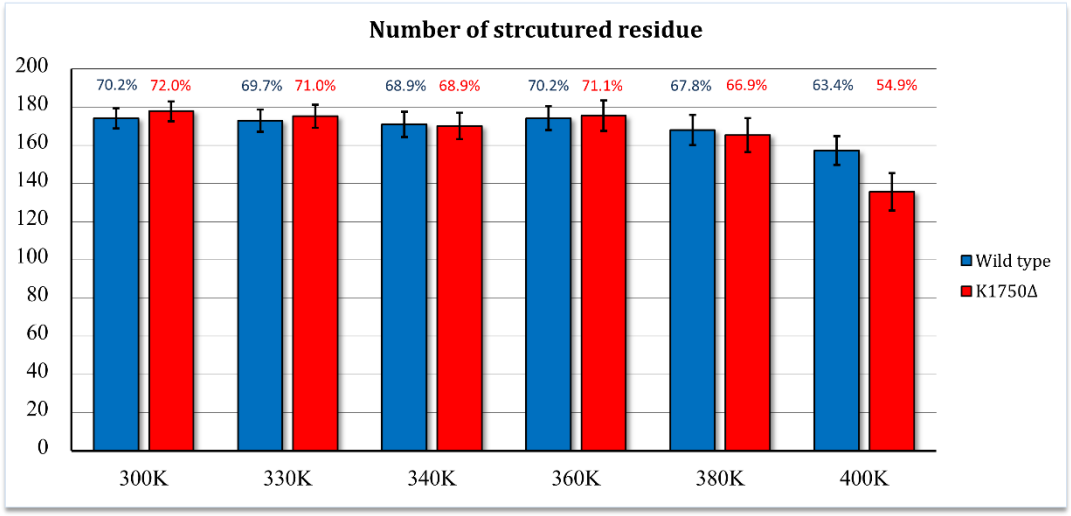


Fig S3 | Percentage of secondary structure at different temperatures, for WT and mutant Sec14-PH.

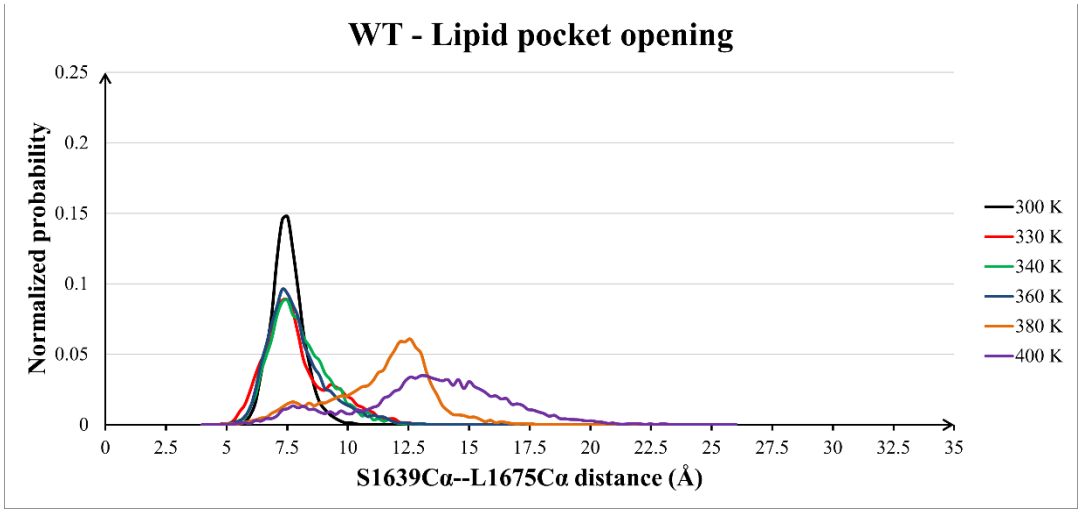


Fig S4 | Distributions of the distance between S1639Calpha and L1675Calpha as a probe for the opening of the lipid pocket in the wild type Sec1-PH.

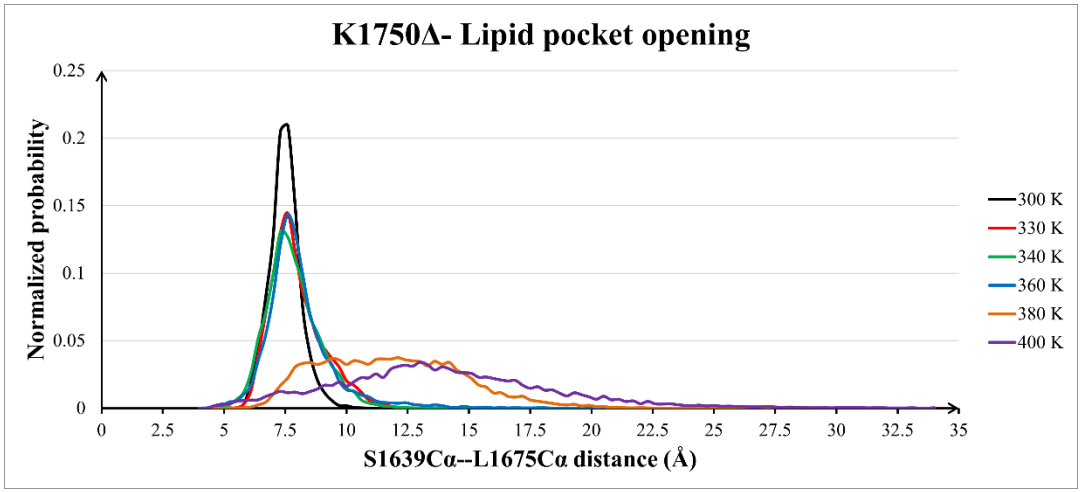


Fig S5 | Distributions of the distance between S1639Calpha and L1675Calpha as a probe for the opening of the lipid pocket in the K1750Δ mutant Sec1-PH.

Second project: Molecular modelling of a bifurcating Electron Transferring Flavoprotein

Authors: Fabio Rizza, Federica Arrigoni, Renata Tisi, Luca De Gioia

Introduction

Flavin-based electron bifurcation

Flavin-Based Electron Bifurcation (FBEB) is a recently [1,2] discovered energy coupling mechanism exploited by many prokaryotes to couple the exergonic reduction of a molecule (high reduction potential) with the endergonic reduction of another (low potential). This mechanism works by splitting the route of two electrons outgoing from the same two-electron reductant donor: one will follow the exergonic branch of the reaction and the other the endergonic branch (fig. 1). The two-electron donor is a flavin-containing molecule (usually FAD). The reaction usually starts with NAD(P)H which donates two electrons, in the form of hydride (H-), to the flavin. This flavin acts as a two- to one-electron converter. The electron following the high-potential route will reach a final flavin molecule that acts (after two cycles) as a one- to two-electron converter that will reduce the final acceptor (e.g., NAD(P)+, Crotonyl-CoA). The other electron, in the endergonic route, will eventually reduce the low-potential acceptor ferredoxin. At the end of the two cycles there will be one reduced high-potential acceptor (e.g., Crotonyl) and two low-potential reduced ferredoxins. Reduced ferredoxin then provides the organism with free energy needed to do useful work, for example to generate ion gradients [3,4].

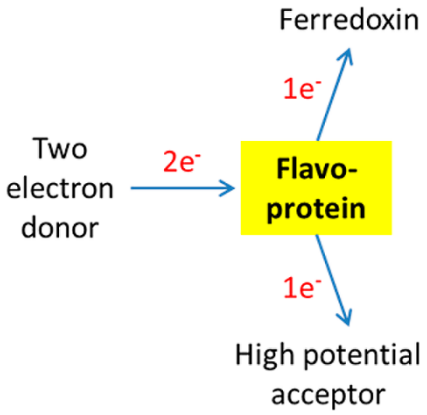


Fig. 1 | General scheme of Flavin-Based Electron Bifurcation. Adapted from [3].

To be suitable for electron bifurcation, a flavin must have a highly unstable semiquinone (radical) state. In particular, it must be characterized by a so-called crossed redox potential: the second reduction, from the radical semiquinone to the fully reduced hydroquinone, is "easier" (higher potential) relative to the first, from the oxidized to the semiquinone form. This is also a feature of free flavins but in the bifurcating flavins the potential gap can reach up to 1.2 V. This is clearly due to the protein environment which destabilizes the semiquinone form. However, how this is achieved is still a mystery. In the case of Etf and Nfn (another family of bifurcating enzymes) an important role in this function has been attributed to a conserved Arginine (also present in non-bifurcating Etf) donating a H-bond to β-FAD N5.
Apart from its biochemical and evolutionary point of view, FBEB can inspire the scientific community to develop new technological devices to reduce low potential acceptors and get useful electro-chemical free energy [5].

Electron transferring flavoproteins

FBEB relies on sophisticated enzymatic systems, grouped into four unrelated families [3]. These are Electron transferring flavoproteins (EtfAB), NADH-dependent ferredoxin:$NADP^+$ reductase (NfnAB), NAD(P)H dehydrogenase (NuoF homologues) and heterodisulfide reductase (HdrABC).
The Electron-transferring flavoprotein Butyryl-CoA Dehydrogenase complex (EtfAB-Bcd) was the first bifurcating enzyme to be discovered [2]. EtfAB is a heterodimeric flavoprotein conserved across all domains of life until mammals, but only some anaerobic organisms use it as an electron bifurcation device. In mammals [6–8], where they do not catalyze electron bifurcating events, they are still involved in transferring electrons coming from the oxidation of fatty acid to the quinone pool in the

mitochondrial electron transport chain [9]. Mutations in the Etf genes result in fatal metabolic diseases [10].

The structure of many Etf proteins, both bifurcating and non-bifurcating, has been solved [11–13]. They all share the same architecture, made up of two subunits, namely A and B. Subunit A consists of two globular domains, called domain I and domain II. The latter is homologous to flavodoxin and binds a FAD molecule called α-FAD. Subunit B is composed of a single globular domain (called domain III) and a long tail (EtfB arm) with a C-terminal helix strongly interacting with domain II. At the interface between domain III and domain I there is the binding site of the bifurcating FAD, called β-FAD (fig. 2). This site is occupied by an AMP molecule in non-bifurcating Etfs.

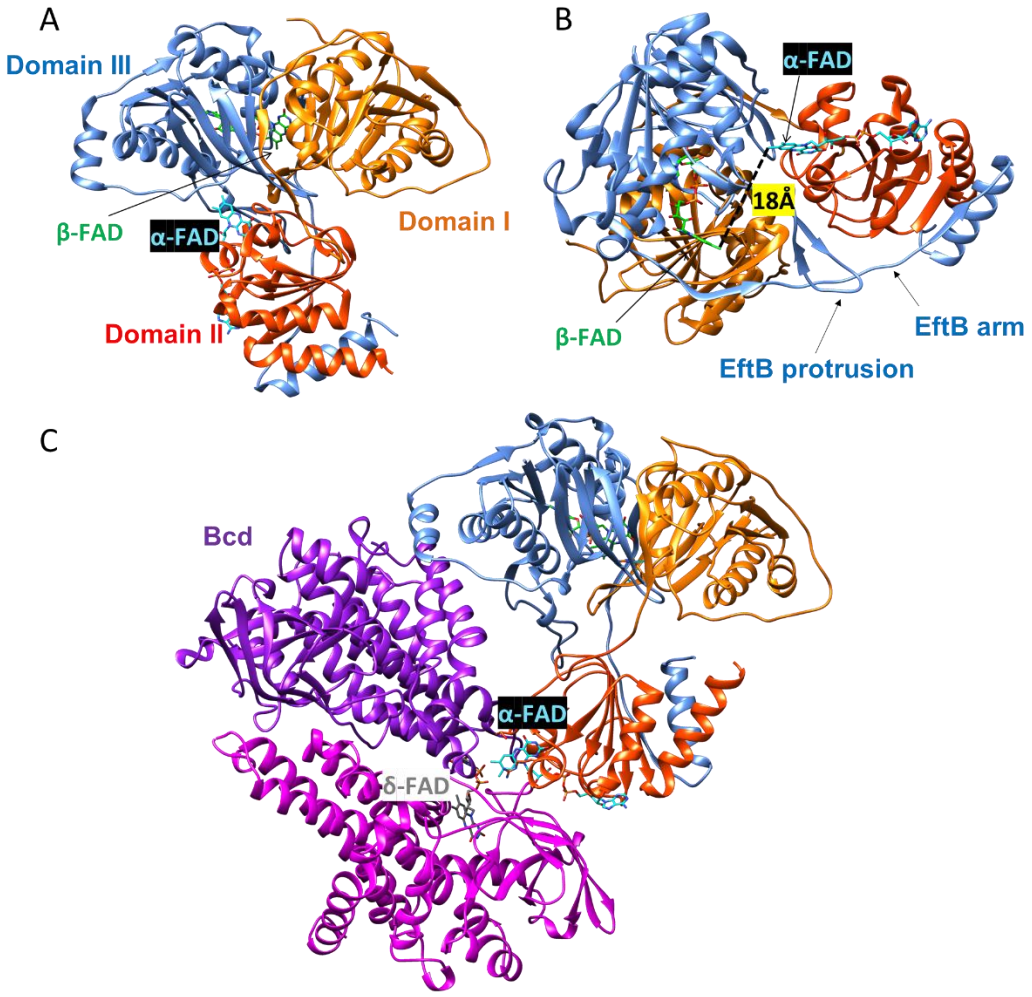


Fig. 2 | Structure of EtfAB-Bcd. A and B: EtfAB is composed of two subunits, namely A and B: Subunit A consists of domain I and II while subunit B consists of domain III. Domain II binds α-FAD while β-FAD is at the interface between domain I and III. The distance between α-FAD and β-FAD is 18 Å and this conformation is called bifurcating-like state (B-like state, see text). C: EtfAB forms a complex with Butyryl-CoA Dehydrogenase (Bcd) whose δ-FAD will eventually reduce the final acceptor crotonyl-CoA to butyryl-CoA. To reduce δ-FAD, domain II must rotate by 80° into a conformation called dehydrogenase state (D-state).

β-FAD receives the electrons from NADH, binding on subunit B close to β-FAD, and sends one towards the anionic semiquinone α-FAD in the high-potential route, which becomes a fully reduced hydroquinone, and the other towards ferredoxin (Fd) in the low-potential route. After α-FAD has received the electron, domain II rotates by 80 degrees and swings towards Butyryl-CoA Dehydrogenase (Bcd) where α-FAD can reduce δ-FAD bound to Bcd. After this first cycle is completed, domain II moves back

so that α-FAD can receive another electron from β-FAD and give it to δ-FAD which can now reduce the final acceptor Crotonyl-CoA to Butyryl-CoA (fig. 3).

$$\text{Crotonyl-CoA} + 2\text{NADH} + \text{Fd}_{ox} \rightarrow \text{Butyryl-CoA} + 2\text{NAD}^+ + \text{Fd}_{red}$$

This functional model is based on the crystal structure of EtfAB in absence of Bcd from A. fermentans, in which α-FAD and β-FAD are close enough to allow the bifurcation step, and on the crystal structure of EtfAB in complex with Bcd from C. difficile in which α-FAD is close to δ-FAD bound to Bcd.

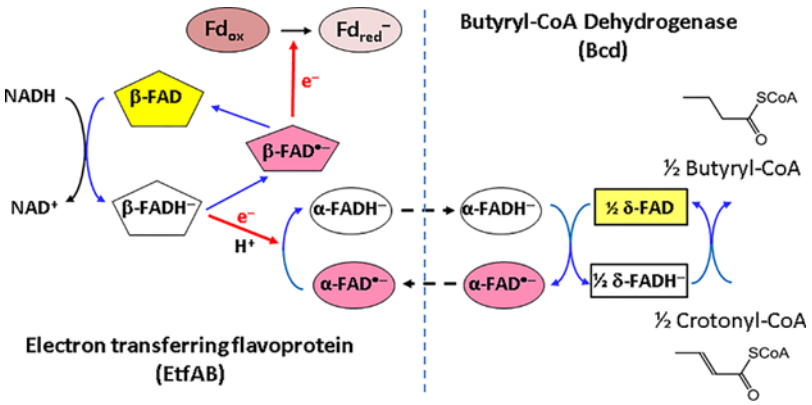


Fig. 3 | The bifurcating process catalyzed by EtfAB-Bcd. Adapted from [3].

This model can explain much of the experimental evidence but still leaves an important unsolved question about the electron transfer for β-FAD. For an efficient electron transfer to occur, the donor-acceptor distance should be below 14Å [14]. The crystal structure of EtfAB without Bcd in the bifurcating state shows a distance between β-FAD and α-FAD of 18Å. This conformation has then been renamed "bifurcating-like" state (B-like state). The authors postulated that a small (10°) rotation of domain II (different from the one, previously described, allowing the swing of domain II toward Bcd) would allow α-FAD to get closer to β-FAD to receive the electrons. The conformation in which the distance between β-FAD and α-FAD goes below 14Å would be the real Bifurcating state (or B-state). In this model, the conformation in which α-FAD is closer to δ-FAD and on Bcd is called dehydrogenase state, or D-state. However, this motion that would transform the B-like into the B-state has never been proved so far. Further, is this movement triggered by the binding of NADH? To answer these questions, we performed a series of molecular dynamics simulations with three different systems: EtfAB, EtfAB-NADH (with β-FAD in the oxidized neutral state) and EtfAB-NAD^+ (with β-FAD in the fully reduced anionic state).

Materials and methods

The coordinates of the bifurcating EtfAB from A. *fermentans* in the bifurcating-like state were obtained from the Protein Data Bank [15] (PDB ID: 4KPU [12]). The first two and last two residues, missing in the PDB file were manually added and no terminal cap was added. All systems setup and subsequent simulations were performed with Gromacs 2018.2 [16]. Hydrogen atoms to both amino acid residues and water molecules were added using the pdb2gmx module of Gromacs. The NADH molecule was docked in its binding site with the Ucsf Chimera interface for Vina, since the nicotine moiety of NAD was not visible in the PDB file.

Three different systems were created: EtfAB, EtfAB-NADH and EtfAB-NAD+. All of them were solvated in a dodecahedral box with dimensions of 11.14 x 11.14 x 11.14 Å and filled with CHARMM-TIP3P water molecules [17]. The genion module was also used to add sodium and chloride ions at 150 mM [NaCl] to neutralize the systems.

All MD simulations were performed with the mdrun module of GROMACS 2018.2 with the CHARMM36m-nov2018 force field [18] for protein, the Charmm General Force Field for NAD/NADH and a novel set of Charmm-compatible parameters for flavins published in [19]. Three independent, 1 μs long, simulations for each system were performed at 300K, for a total of 9 μs. Before the production simulations, all the systems were minimized with the steepest descent method until maximum force went below 1 kJ/mol/nm. The minimization was followed by a 2 ns NVT equilibration phase at the target temperature with the V-rescale algorithm [20] to stabilize the temperature and a 2 ns NPT equilibration phase with the Berendsen [21] algorithm to stabilize the pressure (1 atm). Both equilibration phases were performed with a 1 fs time-step and with position restraints on heavy atoms of the system. The production simulations continued in the NPT ensemble without position restraints for 1 μs, with the V-rescale method for stabilizing the temperature, and the Parrinello-Rahman barostat [22] for the pressure (1 atm). Snapshots of the coordinates were written out every 100 ps. A 2 fs time step was used throughout the production simulations. For both short-range electrostatics and Lennard–Jones a 12 Å cutoff was applied, in the latter case with a force-switch at 10 Å. Long-range electrostatic interactions were calculated with the Particle Mesh Ewald method [23,24] with a grid space of 12 Å and bonds involving hydrogen atoms were constrained using the LINCS algorithm [25]. Periodic boundary conditions were applied. All calculations were performed on the CINECA Marconi supercomputer.

All subsequent analysis (RMSD, RMSF, distances, Cluster analysis and Principal Component Analysis) was performed with Gromacs 2018.2 on the 10000 frames

trajectories coming out of the simulations. Graphs were plotted with Xmgrace. All protein images and videos were rendered with Ucsf Chimera 1.14 [26].

Results

NADH docking

A docking calculation was performed to place the nicotine moiety, invisible in the PDB file, in the correct productive position, stacked with the flavin rings of β-FAD. As already observed [12,13], the sidechain of V223 and S226 block the stacking position of the nicotine moiety (fig. 4). A simple sidechain rotation would not be enough to allow a productive stacking between nicotine and flavin and a local backbone rearrangement is required. Indeed, our docking could not find a productive conformation of NADH. Among the calculated poses, we chose the one with the adenine moiety placed in its correct docking site, in agreement with the crystal structure. This is the starting point for the subsequent molecular simulations.

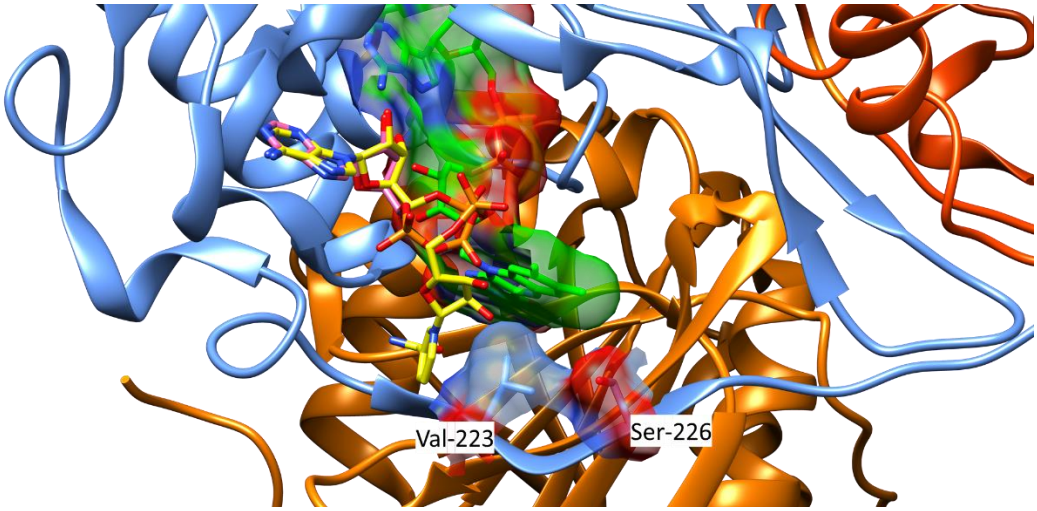


Fig. 4 | Docking of NADH. In green β-FAD, in pink it is the adenosine moiety visible in the crystal structure, in yellow the best pose of the whole NADH molecule. The sidechain of Val-223 and Ser-227 block the formation of a productive stacking interaction between the nicotine moiety of NADH and the flavin rings of β-FAD.

Etf dynamics

Three different systems were simulated: EtfAB-FADs, EtfAB-FADs-NADH and EtfAB-FADs-NAD^+. To obtain the last system, we started from EtfAB-FADs-NADH and switched the parameters of NADH to NAD^+ and of β-$FADH_2$ to β-FAD. α-FAD was always simulated in the deprotonated (anionic) semiquinone state, as experimentally observed. Three independent simulations for each system were performed.

Persistence of NAD^+ in the docking site

Molecular dynamics simulations should help to dock the NAD^+ molecule in the docking site, by allowing that backbone fluctuation needed to place NAD in stacking conformation. This has been achieved only during the second replica of the simulation with NADH and the second replica of the simulation with NAD^+ (fig. 5). In all other cases the thermal fluctuations break the interactions between subunit B of Etf and the adenine part of NAD^+.

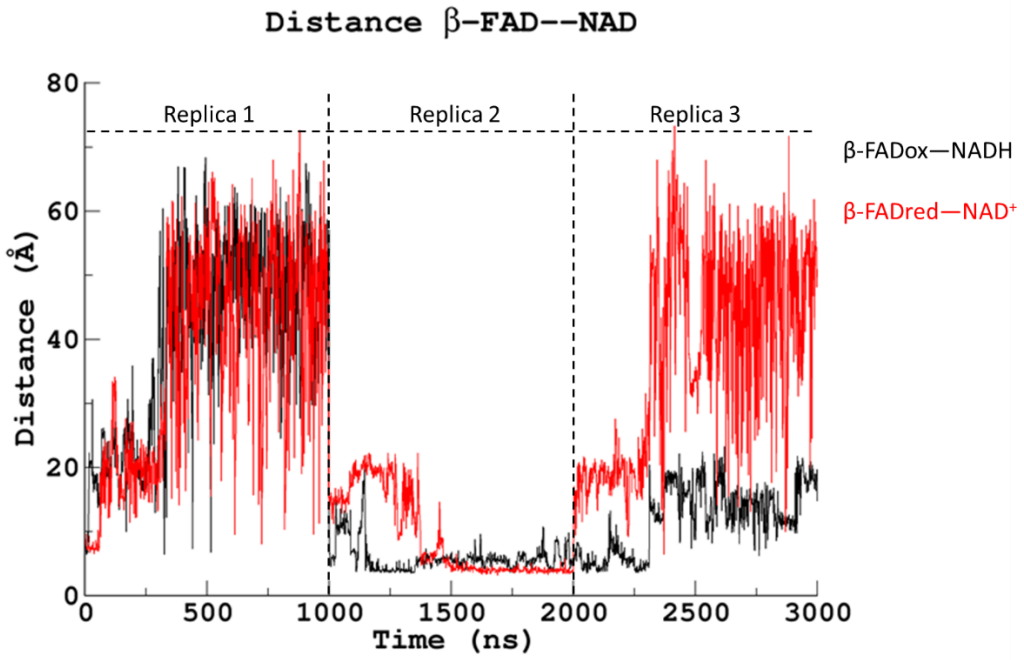


Fig. 5 | Persistence of NAD in the binding site. Only during the second replica for the two NAD-containing systems, the NAD molecule remained within its docking site.

β-FAD -- a-FAD distance

The main goal of these simulations was to assess the hypothesized rotation of domain II that would permit the two FAD molecules to approach below the 14Å threshold, to allow an efficient electron transfer. As we can see in figure 6, such a minimum distance between β-FAD and α-FAD was never achieved during the simulation in any of the system states. They fluctuate around the value observed in the crystal structure, independently of the presence of NAD(H).

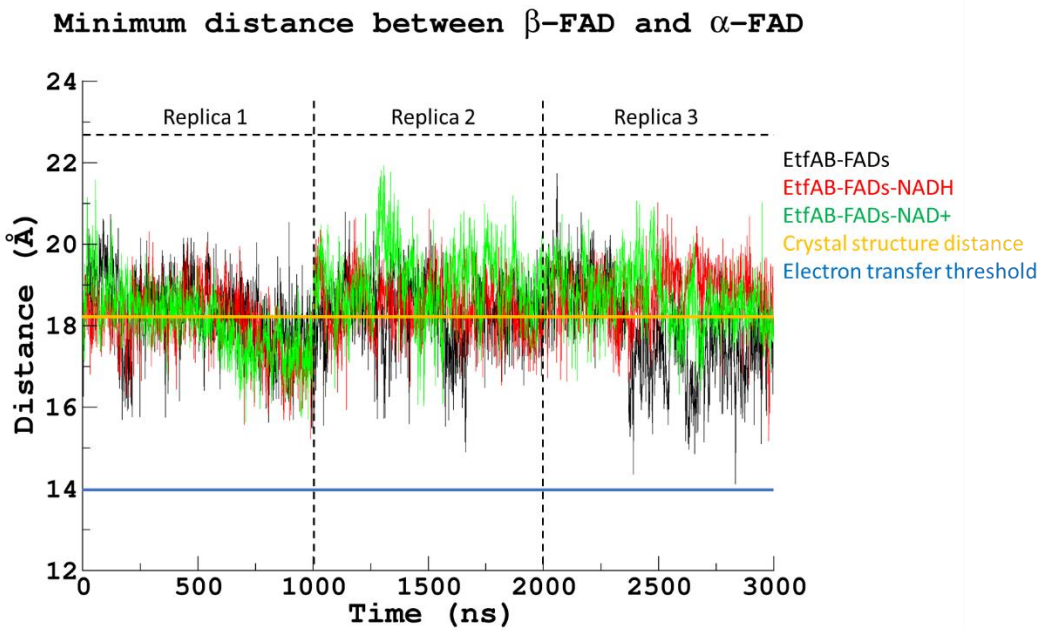


Fig. 6 | Distance between β-FAD and α-FAD. In our trajectories the distance between the two flavins fluctuate around the value observed in the crystal structure and never get below 14 Å.

The Root Mean Square Fluctuations (RMSF) analysis on the alpha-carbons suggests a limited overall internal flexibility, except for the EtfB protrusion, the EtfB arm and the C-terminal of domain II which is tightly coupled to the EtfB arm (fig. 7). For the two systems in complex with NAD(H) the analysis was performed on the portion of the trajectory in which the NAD(H) molecule is in a stacking conformation, to detect any NAD(H)-dependent increase in structural flexibility.

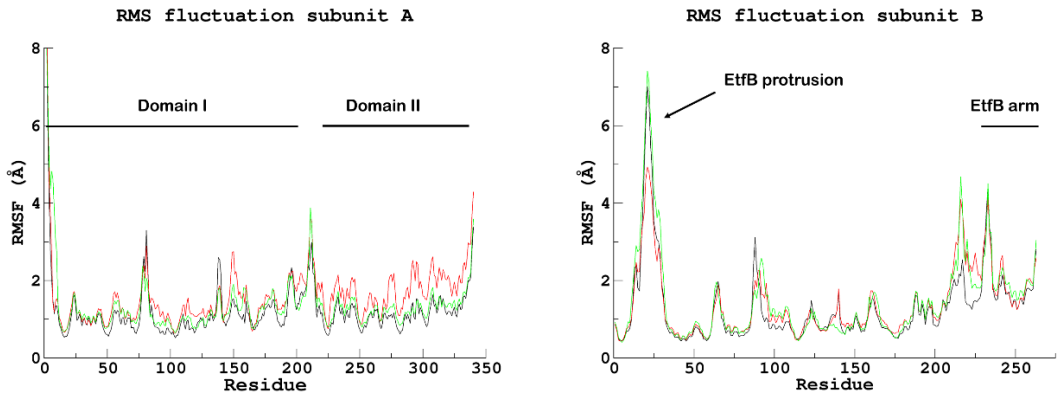


Fig. 7 | RMSF of EtfAB. Overall, the atomic fluctuations are small, with the exception of the EtfB protrusion, the EtfB arm and the C-termianl domain II.

The conformation of the EtfB protrusion represents the main structural difference between known bifurcating and non-bifurcating EtfAB. In the former, it interacts with the EtfB arm (open conformation), while in the latter it interacts with the globular portion of subunit B (closed conformation). If it were in the closed conformation also in the bifurcating EtfAB it would probably interfere with the docking of ferredoxin on subunit B to receive the very low potential electron from β-FAD. The role of this loop is not clear and has been proposed to guide the dynamics of domain II which tightly interacts with the EtfB arm. In our simulations the EtfB arm can indeed close, like in the non-bifurcating EtfAB but this motion is not coupled with a specific rotation of domain II.

To clean the trajectory from the noise of random fluctuations, and only leave the essential modes of motion with potential biological functions, we performed Principal Component Analysis (PCA) based on the covariance matrix of the positions of the alpha carbons, after superposition of the coordinates of subunit B, to capture, if present, the relative motions of domain II. To capture the effect of the NAD(H) molecule, the PCA was also performed on the EtfAB-FADs-NADH and EtfAB-FADs-NAD^+ simulations. In the system without NAD, for which we have a 3-μs trajectory, the fluctuations on the first three eigenvectors, accounting for the 60% of the total variance, still confirm that the EtfB protrusion is the most mobile region. At the same time, domain II is free to rotate around the axe connecting its center of mass with that of domain III. However, the direction of this rotation is not dependent on the direction of the swing of the EtfB protrusion. For example, in the first PC the open-to-close swing of the loop is coupled to an anti-clockwise rotation of domain II, while in the second PC the same swing of the loop is coupled to a clockwise rotation of domain II.

This would suggest that these two structural elements (EtfB protrusion and domain II) may not be dynamically coupled.
The more distant an element is from the interface between domain II and domain III the higher the fluctuation associated with that element: corresponding atoms on the C-terminal helix, tightly bound to the C-terminal helix of EtfB arm were translated by 7 Å moving between the two extremes of PC1.
Domain II has some mobility with respect to domain III allowed by the flexibility of the linker between domain I and II. However, this is not enough to guarantee the α-FAD to get closer to β-FAD by the 4 Å necessary to shorten the distance between the two flavins to the electron transfer threshold, as hypothesized. This is because of the strong electrostatic and H-bond interactions forming at the interface between domain II and domain III. In particular, the H-bonds between Arg253 on domain II with Tyr-40 on domain III and Asp-257 on domain II with Arg179 on domain III. Those interactions are localized just around the flavin rings of the α-FAD and represent the hinge of the rotations of domain II, thus preventing α-FAD to move from its position but allowing the distant parts of domain II to swing.

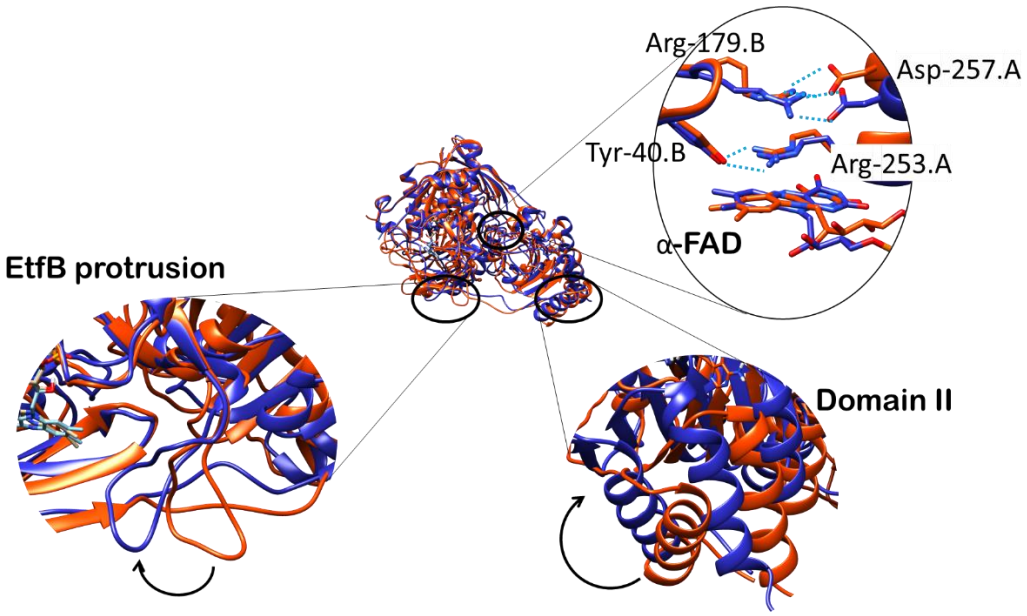


Fig. 8 | Effect of the first principal component. EtfB protrusion and the distal region of domain II are the most mobile elements of EtfAB both in the row trajectory and in the essential modes. Strong interactions at the interface with domain I serve as a hinge preventing α-FAD to move towards β-FAD.

The same essential modes could be observed also for the systems with NAD(H). However, the extent of the motions was smaller compared to the system without NAD because of the lower amount of data collected, since the NAD(H) persisted within its docking site only for a small fraction of the entire simulations (1 μs for EtfAB-FADs-NADH and 0,5 μs for EtfAB-FADs-NADH).
The only way the two flavins could get closer, even though never below 14 Å is by the fluctuations of β-FAD within its pocket, which shows high mobility. In particular, the H-bond between β-FAD-N5 and Arg-146 on domain I, which is considered important for destabilizing the semiquinone state, is not stable over the course of our simulations. We do not know if the low persistence of this interaction arises from a non-proper parameterization of the flavins or if it reproduces the correct behavior, but it would have interesting consequences on the understanding of the role of the chemical environment on the stability of the semiquinone state and will deserve further investigation.

Spectra calculations with DFT

Our molecular dynamics simulations suggest that domain II could not rotate to bring α-FAD close enough to β-FAD to permit the electron transfer. So, we investigated the possibility of an electron transfer event at distances over 14 Å. The next calculations were performed by my colleague Dr. Federica Arrigoni by means of Density Functional Theory (DFT), and I will briefly report them for the sake of completeness of the project.
The problem of the electron transfer in EtfAB has been recently studied experimentally in [27]. What the authors show is a charge transfer band at 726 nm during reductive titration of EtfAB, with a pick when α-FAD is fully reduced, and β-FAD is oxidized. They hypothesized that there must be a huge conformational change such to establish a stacking interaction between the two flavins. Since our simulations do not suggest any possible approach of the two FAD molecules, we tried to calculate the spectra in conditions compatible with the results of our simulations.

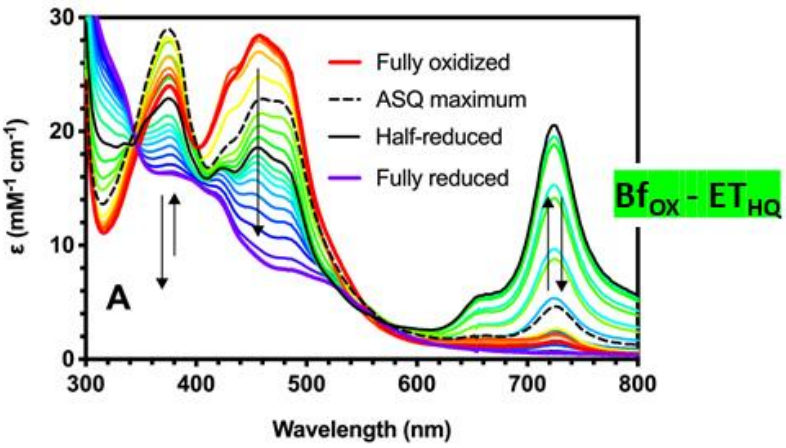


Fig. 9 | The charge transfer band observed in [27] during a reductive titration. The highest intensity is observed with the bifurcating FAD (β-FAD) in the oxidized state and the electron-transfer FAD (α-FAD) in the fully reduced state.

Initially, calculations were performed with the two flavins alone at 18 Å and then we gradually increased the complexity of the environment. We first added the adenosine moiety of the FAD molecules and then we added residues that have been suggested to play an important role in modulating the electrochemical potential of the flavins (like Arg-146 domain I for β-FAD and Ser-270 on domain II for α-FAD) or the electron transfer itself, like Tyr-37 on subunit B as suggested in [27]. Even in the more complex case, only a very weak band arose at around 700 nm.

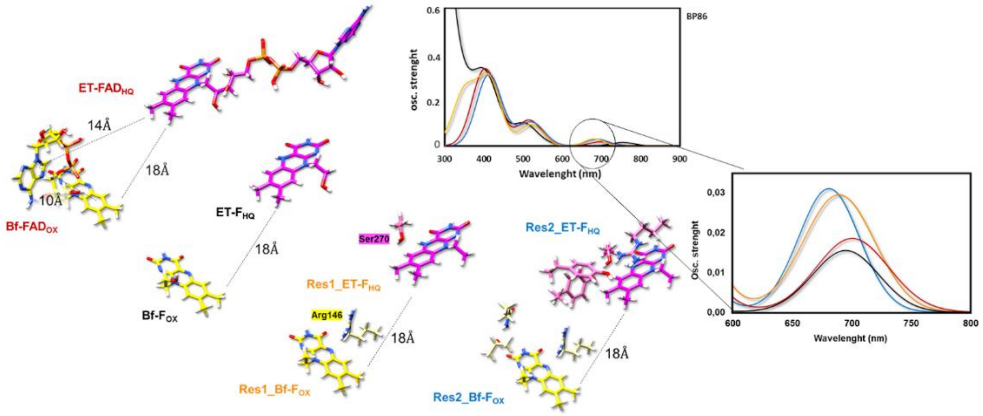


Fig. 10 | Starting with only the two flavins (black curve), more elements were gradually added. Adenosine moieties (red curve), FAD-N5 interacting residues (orange curve) and more residues that might have a role in electron transfer (blue curve). Even increasing the complexity of the systems, the intensity of the calculated band is far from the experimental intensity.

After that we tested the effect of the distance on the spectra due to local fluctuations in the position of β-FAD within its pocket. A frame from the simulations was chosen where the distance α-FAD--β-FAD was 15.5 Å but also in this case the band was very weak compared to the experimental one.
Another hypothesis we tested was that a second flavin could enter in the pocket of β-FAD and establish a stacking interaction. So, we performed a docking calculation of a FAD molecule in the β-FAD pocket by removing the NADH from a frame of our simulations. Different poses were found with parallel stacking originating an intense band.
Finally, we tested the hypothesis that a productive interaction within 14 Å could arise from a flavin on a different EtfAB. So we performed a protein-protein docking between subunitB-β-FAD and domainII-α-FAD but neither of the poses allowed the two flavins to approach below 18 Å.

Discussion

With our calculations we tested a model for the bifurcating EtfAB to allow an intramolecular electron transfer between two FAD molecules that are apparently too far away from each other. The previously proposed model is that a conformational change would occur at the interface between domain II and domain III such that domain II could bring its α-FAD 4 Å closer to β-FAD [28]. This would be allowed by a rotation of domain II around the linker with domain I and taking advantage of the soft solvent mediated interactions with domain III. When this model was proposed there was no direct evidence for it nor any calculations. More recently, a spectroscopy experiment has observed a charge transfer band, that is an electronic communication occurring between the two flavins. The authors invoked a drastic conformational change that would bring the two flavins in a stacking configuration [27,28].

However, our simulation could not capture any frame in which the distance between the flavins dropped below 14 Å. This is because the complete postulated motion for domain II is blocked by strong interactions at the interface with domain III which are localized very close to the binding pocket of α-FAD serving as hinges for a swing of the most distal regions of the domain, with the result that α-FAD is stuck in its position.

Also, our simulations showed the flexibility of the EtfB protrusion, a long loop that in the crystal structure interacts with the EtfB arm which in turn is tightly bound to the domain II. No hypothesis has been formulated so far regarding its possible function. In our simulations its back-and-forth swing is not coupled with a specific direction of the observed (non-functional) rotation of domain II, so its role is hard to interpret.

Since our simulations, within the limit of a few microseconds, suggest that the hypothesized functional motion may not occur, we switched to a DFT approach to try to get insights into the spectroscopy data. When the two flavins are more than 15 Å distant, a small band could be predicted, even though it is much less intense than the experimentally reported one. The only way to get a band comparable with the experimental one is to place a second FAD in the β-FAD pocket in a parallel stacking configuration, where usually a NAD(H) molecule would be. A scenario that was not considered by the authors of the experiments, who invoked a dramatic structural change, is that a free FAD in solution could enter the β-FAD pocket and establish a stacking interaction.

To conclude, the way electrons flow within bifurcating EtfAB is not solved yet. The original conformational change that was proposed on the base of the crystal structure is unlikely according to our simulations, and at least another interpretation of the spectroscopic results is possible. In future calculations, it will be important to extend both the molecular dynamics simulations, maybe with a free energy approach to better

explore hidden regions in the phase space, and to add elements to the DFT models, also including other residues that have been proposed to bridge the electron transfer, so far without any experimental evidence.

References

1. Li F, Hinderberger J, Seedorf H, Zhang J, Buckel W, Thauer RK. Coupled Ferredoxin and Crotonyl Coenzyme A (CoA) Reduction with NADH Catalyzed by the Butyryl-CoA Dehydrogenase/Etf Complex from Clostridium kluyveri. Journal of Bacteriology. 2008. pp. 843–850. doi:10.1128/jb.01417-07

2. Herrmann G, Jayamani E, Mai G, Buckel W. Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. J Bacteriol. 2008;190: 784–791

3. Buckel W, Thauer RK. Flavin-Based Electron Bifurcation, A New Mechanism of Biological Energy Coupling. Chem Rev. 2018;118: 3862–3886.

4. Thauer RK. Flavin-based electron bifurcation: A novel mechanism of energy coupling in anaerobic microorganisms. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2014. p. e7. doi:10.1016/j.bbabio.2014.05.130

5. Das A, Hessin C, Ren Y, Desage-El Murr M. Biological concepts for catalysis and reactivity: empowering bioinspiration. Chem Soc Rev. 2020;49: 8840–8867

6. Hauge JG, Crane FL, Beinert H. On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. III. Palmityl coA dehydrogenase. J Biol Chem. 1956;219: 727–733.

7. Crane FL, Beinert H. On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. II. The electron-transferring flavoprotein. J Biol Chem. 1956;218: 717–731

8. Crane FL, Mii S, Hauge JG, Green DE, Beinert H. On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. I. The general fatty acyl coenzyme A dehydrogenase. J Biol Chem. 1956;218: 701–706

9. Ruzicka FJ, Beinert H. A new iron-sulfur flavoprotein of the respiratory chain. A component of the fatty acid beta oxidation pathway. J Biol Chem. 1977;252: 8440–8445

10. Loehr JP, Goodman SI, Frerman FE. Glutaric acidemia type II: heterogeneity of clinical and biochemical phenotypes. Pediatr Res. 1990;27: 311–315.

11. Roberts DL, Frerman FE, Kim JJ. Three-dimensional structure of human electron transfer flavoprotein to 2.1-A resolution. Proc Natl Acad Sci U S A. 1996;93: 14355–14360

12. Chowdhury NP, Mowafy AM, Demmer JK, Upadhyay V, Koelzer S, Jayamani E, et al. Studies on the Mechanism of Electron Bifurcation Catalyzed by Electron Transferring Flavoprotein (Etf) and Butyryl-CoA Dehydrogenase (Bcd) ofAcidaminococcus fermentans. Journal of Biological Chemistry. 2014. pp. 5145–5157. doi:10.1074/jbc.m113.521013

13. Demmer JK, Pal Chowdhury N, Selmer T, Ermler U, Buckel W. The semiquinone swing in the bifurcating electron transferring flavoprotein/butyryl-CoA dehydrogenase complex from Clostridium difficile. Nat Commun. 2017;8: 1577

14. Page CC, Moser CC, Chen X, Leslie Dutton P. Natural engineering principles of electron tunnelling in biological oxidation–reduction. Nature. 1999. pp. 47–52. doi:10.1038/46972

15. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28: 235–242

16. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015. pp. 19–25. doi:10.1016/j.softx.2015.06.001

17. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B. 1998;102: 3586–3616

18. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods. 2017;14: 71–73

19. Aleksandrov A. A Molecular Mechanics Model for Flavins. J Comput Chem. 2019;40: 2834–2842

20. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. The Journal of Chemical Physics. 2007. p. 014101. doi:10.1063/1.2408420

21. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics. 1984. pp. 3684–3690. doi:10.1063/1.448118

22. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied Physics. 1981. pp. 7182–7190. doi:10.1063/1.328693

23. Darden T, York D, Pedersen L. Particle mesh Ewald: AnN·log(N) method for Ewald sums in large systems. The Journal of Chemical Physics. 1993. pp. 10089–10092. doi:10.1063/1.464397

24. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of Chemical Physics. 1995. pp. 8577–8593. doi:10.1063/1.470117

25. Hess B, Bekker H, Berendsen HJC, Johannes G E. LINCS: A linear constraint solver for molecular simulations. Journal of Computational Chemistry. 1997. pp. 1463–1472. doi:3.0.co;2-h">10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.0.co;2-h

26. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25: 1605–1612.

27. Duan HD, Lubner CE, Tokmina-Lukaszewska M, Gauss GH, Bothner B, King PW, et al. Distinct properties underlie flavin-based electron bifurcation in a novel electron transfer flavoprotein FixAB from. J Biol Chem. 2018;293: 4688–4701

28. Chowdhury NP, Mowafy AM, Demmer JK, Upadhyay V, Koelzer S, Jayamani E, et al. Studies on the mechanism of electron bifurcation catalyzed by electron transferring flavoprotein (Etf) and butyryl-CoA dehydrogenase (Bcd) of Acidaminococcus fermentans. J Biol Chem. 2014;289: 5145–5157.

Third project: On the propagation of the OH radical produced by CU-Amyloid Beta peptide model complexes. Insights from molecular modelling (published paper)

Federica Arrigoni, **Fabio Rizza**, Renata Tisi, Luca De Gioia, Giuseppe Zampella, Luca Bertini*

Department of Biotechnologies and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy. e-mail luca.bertini@unimib.it

Abstract

Oxidative stress and metal dyshomeostasis are considered crucial factors in the pathogenesis of the Alzheimer's disease (AD). Indeed, transition metal ions such as Cu(II) can generate Reactive Oxygen Species (ROS) via O_2 Fenton-like reduction, catalysed by Cu(II) coordinated to the Amyloid beta (Aβ) peptide. Despite intensive effort, the mechanisms of the ROS-induced molecular damage remain poorly understood. In the present paper, we investigate on the basis of molecular modelling computations the mechanism of the OH radical propagation toward the Aβ peptide, starting from the end-product of the OH radical generation by Cu(II)·Aβ. We evaluate i) the OH oxidative capacity, as well as the energetic of the possible Aβ oxidation residue targets, by quantum chemistry Density Functional Theory (DFT) on coordination models of Cu(II)/OH/Aβ and ii) the motion of the $OH^{\cdot}$ approaching to the Aβ target residues by classical Molecular Dynamics (MD) on the full peptide Cu(II)/OH/Aβ(1-16). The results show that the oxidative capacity of OH coordinated Cu(II)Aβ is significantly lower than that of free OH radical and that propagation toward Aβ Asp and His residues are favored over Tyr residues. These results are discussed based on the recent literature on *in vitro* Aβ metal-catalyzed oxidation and on the possible implications in the AD oxidative stress mechanism.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by extracellular accumulation of Amyloid beta peptide (Aβ) in senile plaques and the intracellular hyperphosphorylation of the tau protein. (1) The majority of AD cases occur in age groups over 65 years and is the most frequent cause of dementia, being globally the fifth leading cause of death. (2) Despite the enormous efforts of the scientific community the pathogenesis of AD has yet to be cleared, no effective treatment has yet been found and those currently available can only delay the disease.

At first glance on the scientific literature, the conspicuous number of hypotheses and mechanism proposed for the etiology and pathogenesis of AD are striking, a sign of its multifunctional nature. Since the first seminal paper by Hardy and Higgins (3) the so-called Amyloid Cascade Hypothesis has been considered the most popular hypothesis in explaining the aetiology and pathogenesis of AD. Aβ is a 39-42 AA peptide that is the results of a pathological proteolytic processing of the Amyloid precursor protein (APP), a transmembrane protein expressed in particular at the synapse of neurons. The observation that the main constituent of the senile plaques is the Aβ peptide led to the conclusion that Aβ deposits are "*the causative agent*" and an upstream event in the evolution of AD.

New scientific evidence and the almost complete failure clinical trial of the anti-amyloid treatments (4) led to a wider view of AD which is currently considered a multifactorial disorder and a number of key events and risk factors have been associated to AD. Among these, oxidative stress (5) and metal dyshomeostasis have been intensively investigated and now it is increasingly evident that these two events are closely related. (6) (7)

Oxidative stress, defined as an imbalance between free radicals production and antioxidant defences has been characterized in AD blood (8) and post-mortem brain samples by measuring the level of biomarkers which are intermediates or end-product for lipid peroxidation, protein and nucleic acids (nuclear and mitochondrial DNA and RNA) oxidation induced by free radicals.

Lipid peroxidation is a non-enzymatic process under which free radicals attack lipids containing carbon-carbon double bond(s) of the polyunsaturated fatty acids, resulting in cell damage. (9) (10) The intermediates of this process are highly toxic aldehydes such as 4-hydroxynonenal (HNE) or malondialdehyde (11) whose level have been found higher in AD brain samples compared to control samples. (12) (13) Similarly, the level of 8-hydroxyguanine, one of the products of nucleic acid oxidation (14), have been found significantly elevated in AD samples. (15) Finally, redox proteomics techniques

(16) has been used to identify proteins oxidatively modified by the Aβ peptide (in particular protein carbonylation) *in vitro* and *in vivo*.

Transition metal dyshomeostasis and accumulation are common features in AD as other neurodegenerative disorders (17) and the *metal hypothesis* was proposed besides the amyloid cascade one. (18) Focusing on copper, this hypothesis starts from the clinical evidence of the Cu dyshomeostasis (19) in AD and from the experimental evidence that senile plaques are metal sinks for Cu accumulation. A Significant increases of Cu(II) ions was detected in AD samples in comparison with healthy age-matched controls (390 μM vs 70 μM) (20) (21) and for this the role of copper dyshomeostasis in AD that has been subject of an intense research. (22) (23)

Copper in AD can be considered as a pivotal element that is able of accounting for both senile plaque formation and oxidative stress.

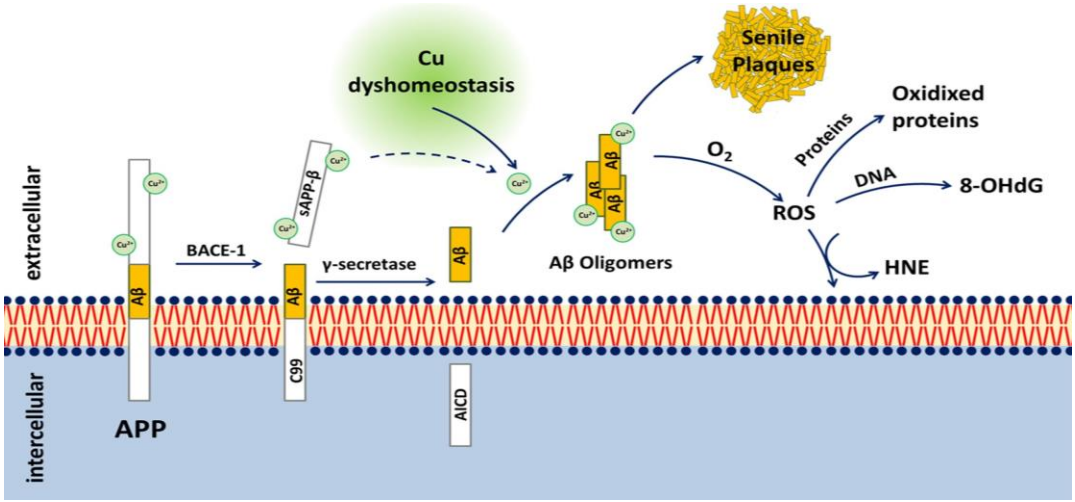


Figure 1. From left to right, the connection among amyloid cascade, metal dyshomeostasis and oxidative stress hypothesis. Upon amyloidogenic proteolytic processing of APP by β- (BACE-1) and γ-secretase, three sub-systems are formed, namely the soluble APP-β, the amyloid Intracellular domain (AICD) and the Aβ peptide. This latter starts to form oligomers, possibly favored or assisted by Cu(II) (which is also initially bound to APP), until the formation of the senile plaques. Small Cu(II) Aβ oligomers catalyze ROS formation in presence of molecular oxygen and assisted by the extracellular reducing environment. The oxidation of proteins, lipids and nucleic acids by highly oxidant species such as the OH radical is implicated in the brain oxidative stress and can be detected by formation of radical propagation end-products such as 4-hydroxynonenal (HNE), (11) 8-hydroxyguanine (14) and oxidized protein products.

Following from left to right Figure 1, we resume the link among amyloid cascade, metal dyshomeostasis and oxidative stress hypothesis. It is well established that APP features three high affinity Cu binding sites (24) (25) (26) (27) (28) with a K_D in the range 0.001-10 nM and a putative role in Cu homeostasis. Following the amyloidogenic proteolytic cleavage by β- and γ-secretase (29), the Aβ peptide is released and forms at first small soluble oligomeric intermediates that are successively converted to fibrils. (30,31) The lack of selectivity of γ-secretase leads to the formation of Aβ peptides of different length in which 1-42, 4-42 and 1-40 are prevalent (32). Cumulative evidence document that the formation of the former is an early toxic event in the pathogenesis of AD. (33) (30)

The role of transition metals in the Aβ reactivity and aggregation is continuously under investigation although a definitive word on this issue has not been given yet. (34) Aβ peptide binds Cu(I) and Cu(II) with a 1:1 ratio. The Aβ(1-42) peptide is characterized by a 1-16 hydrophilic portion that hold 6 charged side chains (Asp1, Glu3, Arg5, Asp7, Glu11 and Lys16) and three His residues (His 6,13 and 14). 1-16 fragment can bind Cu(II) with high affinity (K_d = 0.1 nM (35)) with a pH dependent coordination of 3N2O type (36) that involved the N-terminal and carboxylate of Asp1, the three His residues (37) and likely the Asp1-Ala2 peptide bond carbonyl group. (38) The affinity of Cu(I) in its 2 His coordination is more debated (39) but recent measurement found it similar to that of Cu(II). (35)

In the so-called "Aβ oligomers hypothesis" (30) (40) copper plays a dual role in stabilizing the small Aβ aggregate by:

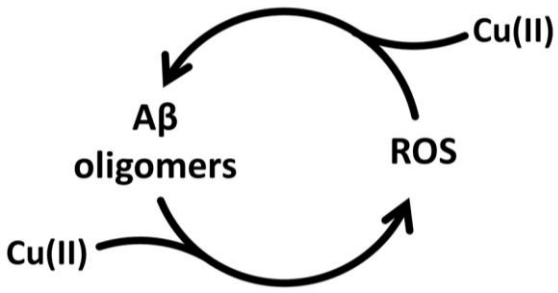
1) the formation of Aβ-Cu-Aβ bridges.
2) mediating the dityrosine cross-linking via ROS formation.

Single molecule atomic force spectroscopy experiments (41) evidence the increase in binding forces between two Aβ(1-42) in presence of Cu(II) peptide likely due to the formation of Cu bridges, stabilizing small Aβ oligomers thus increasing their lifetime. (42) Similarly Sitkiewicz et al. (43) observed that Cu(II) stabilizes a more compact form of small Aβ oligomers. In this field also *in silico* molecular modelling approach has given a important contribution (44) in particular regarding the dynamic of the Cu/Aβ dimers and the mechanism of copper influence on Aβ oligomers compactness. (45) (46)

Cu(II), besides assisting the formation of the toxic small Aβ oligomers, certainly promotes ROS formation, as clearly evidenced in the recent literature. *In vitro* and *in silico* experiments showed that Cu(II) coordinated to Aβ (i) can be reduced to Cu(I) using ascorbate as reductant with an inner sphere mechanism (47); (ii) in presence of oxygen and ascorbate can generate OH radicals (47,48) (49) (50) (51) through a Fenton-like mechanism. Moreover, emerging evidence clearly show that the radical production in AD can be related to copper dyshomeostasis mediated by Aβ oligomers, thus yielding

oxidative stress. This is justified by the brain reducing environment and the high concentration of molecular oxygen.
Finally, OH radical can propagate toward the membrane, other proteins and nucleic acids inducing oxidative stress, as confirmed by post-mortem AD patients' investigations.

The line that connects amyloid (oligomers), metal dyshomeostasis and oxidative stress hypotheses is far from clarified in all its aspects, but it can be resumed in the sketch below.



Cu(II) promotes the Aβ oligomerization through Cu bridges and the ROS formation that in turn promotes dityrosine cross-link stabilizing Aβ oligomers. Moreover, closing this circle, there are evidences that small Cu(II)·Aβ oligomers are the most redox active and therefore the *optimal ROS generators*. (52)
At this point, the investigation on Cu(II) Aβ-induced Aβ oxidation becomes crucial. In previous *in vitro* investigations, experiments were performed on Aβ peptide of different length (from 1-6 to 1-42) incubated with Cu(II) in presence of a reductant. In this condition, the reaction produces OH radicals that propagates causing Aβ oxidation, whose products have been investigated by various mass spectrometry techniques. In the first investigation on Aβ(1-28) and Aβ(1-40) (53) Schoneich and Williams reported the oxidation of His13 and His14 to 2-oxo-His, followed by His6 oxidation after prolonged oxidation, while Tyr10 modification was excluded. Inoue et al. (54) confirmed these results on Aβ(1-16) and Aβ(1-40) while they detected only Asp1 and His6 oxidation on Aβ(1-6). (55) Kowalik-Jankowska et al. (56) investigated Aβ(1-10) and Aβ(1-16) peptides showing that His13, His14 and Asp1 were targeted by oxidation, while again Tyr10 is not involved. Nadal et al. (57) monitoring with ^{1}H NMR identified the oxidation of His and Met35 residues for Aβ(1-42). Cassagnes et al. (58) focused on Aβ(1-28) with the aim of elucidating the Cu(II) coordination responsible for the ROS production. In their investigation, these authors detected Asp1

decarboxylation/deamination and His13/His14 oxidation, while neither His6 nor Tyr10 modification were observed. Cheignon et al. (59) investigated the Aβ(1-40), finding that the main target of copper-catalyzed ROS production is Asp1. Finally there are a number of investigations only focused on dityrosine cross-link detection on Aβ(1-16), (60) Aβ(1-40) (61) and Aβ(1-42) (62). In a very recent investigation by Vázquez et al (63) the formation of Tyr-Tyr cross-links has been quantified as 3% for Aβ(1-40) in presence of copper and H_2O_2. All these apparently contradictory results find a possible rationale in the computational investigation of La Penna and Li (64) in which the authors carried out molecular dynamics (MD) on $[Cu(II)\cdot A\beta(1\text{-}42)]_2$ dimers showing that the formation of Tyr-Tyr crosslink is favored when the metal ion makes link between different peptides.

In this paper, we try to elucidate by means of Density Functional theory (DFT) and Molecular Dynamics (MD) techniques some points regarding $OH^{\cdot}$ propagation, whose picture remains elusive in literature:

1) the molecular structure of the Cu(II)·OH·Aβ product that comes from the three step Fenton-like O_2 reduction cycle;
2) the redox capacity of the OH radical when coordinated to Cu(II);
3) the thermodynamic and kinetic aspects of the OH propagation toward Aβ;
4) the dynamic of the OH approach to the Aβ oxidation target residues.

Points 1-3 have been investigated at all-electron DFT level on Cu(II)·Aβ model and Cu(II)·Aβ(1-6) coordination's while point 4 has been elucidated by MD simulation on the full Cu(II)·OH·Aβ(1-16) coordination.

Computational details

DFT. The DFT computations were performed using the pure GGA BP86 ((65),(66)) DFT functional and RI technique (67) as implemented in TURBOMOLE 7.2 (68). Basis sets of triple-ζ plus polarization split valence quality (69) were adopted for all atoms. Solvent effect is accounted for by using the COSMO (Conductor-like Screening Model) approach.(70) Water solvation has been considered by setting the dielectric constant equal to 80. This computational setting provides ground state geome-try parameters in good agreement with experimental X-ray values. (71) Charge distributions were evaluated using Natural bond orbital analysis (NBO). Ground state geome-try optimizations were carried out with convergence criteria fixed to 10^{-6} hartree for the energy and 0.001 hartree·bohr^{-1} for the gradient norm vector. The effects of ZPE, thermal and entropic contributions on the purely electronic total energy values to compute free energies have been investigated by means of evaluation of the approximated roto-translational partition function of each molecular species, at T=298 K and P=1 bar. This level of theory has been already tested and adopted in the previous DFT investigation carried out in our laboratory. (72) (51)

Molecular dynamics. N-terminal Cu(II) coordinated DAEFRHDSGYEVHHQK peptide model was built from the corresponding Aβ(1-16) Zn(II) coordinated structure obtained from NMR pdb entry 1ZE9 (73). After substituting Zn(II) with Cu(II), 2 OH group were added with UCSF Chimera software version 1.13(74) to get the starting structures for molecular dynamics (MD) simulations with Gromacs 2018.2 74). We performed MD simulations with the CHARMM36m(75) forcefield for the protein and the CHARMM TIP3P forcefield (76) for water. The peptide·Cu(II)·$(OH)_2^-$ complex was centred in a dodecahedral box at a minimum distance of 1.6 nm from the edge of the box, applying periodic boundary conditions, resulting in a total 6000 water molecules. Two Na^+ ions were added to neutralize the system's charge. Histidines were simulated in the Nε2-H tautomer. Restraints were added to the topology before the minimization step to constraint the Cu(II) coordinations according to 4-coordinated most stable one obtained from DFT geometry optimization (see SI Scheme S4 and Table S7). Each system was energy minimized by steep descent method until maximum force went below 500 J/(mol·nm). We then equilibrated the systems with a first 500 ps NVT simulation at 300 K with the V-rescale thermostat followed by a second 500 ps NPT step to equilibrate the pressure with the Berendsen barostat. In each equilibration simulation the position of heavy atoms in the protein was restrained and the integration time step was 1 fs. Five independent 1 μs production simulations without restraining protein atoms were subsequently performed in NPT ensemble with the Parrinello-Rahman barostat and an integration time-step of 2 ps. Analysis was performed on the

merged trajectory of the five replicas with a 100 ps spacing between frames, for a total of 50000 frames.

Results

Stability of the Cu(II)(OH)$_2^-$·Aβ model coordinations. Previous investigations (48) (51) (72) proposed three Cu(II)/Cu(I) redox competent coordinations in the three electron Fenton-like O_2 reduction cycle as sketched in Figure 2. In each step, ascorbate reduces Cu(II) to Cu(I) that reduces the substrate ($^3O_2$, superoxide O_2^-, and hydroperoxide O_2H^-) until the formation of one OH^- and OH radical. The O-O bond dissociation of the peroxide yields three possible Cu(II)$(OH)_2^-$Aβ coordinations in which the OH radical and the OH^- ligand are bound to the metal ion.

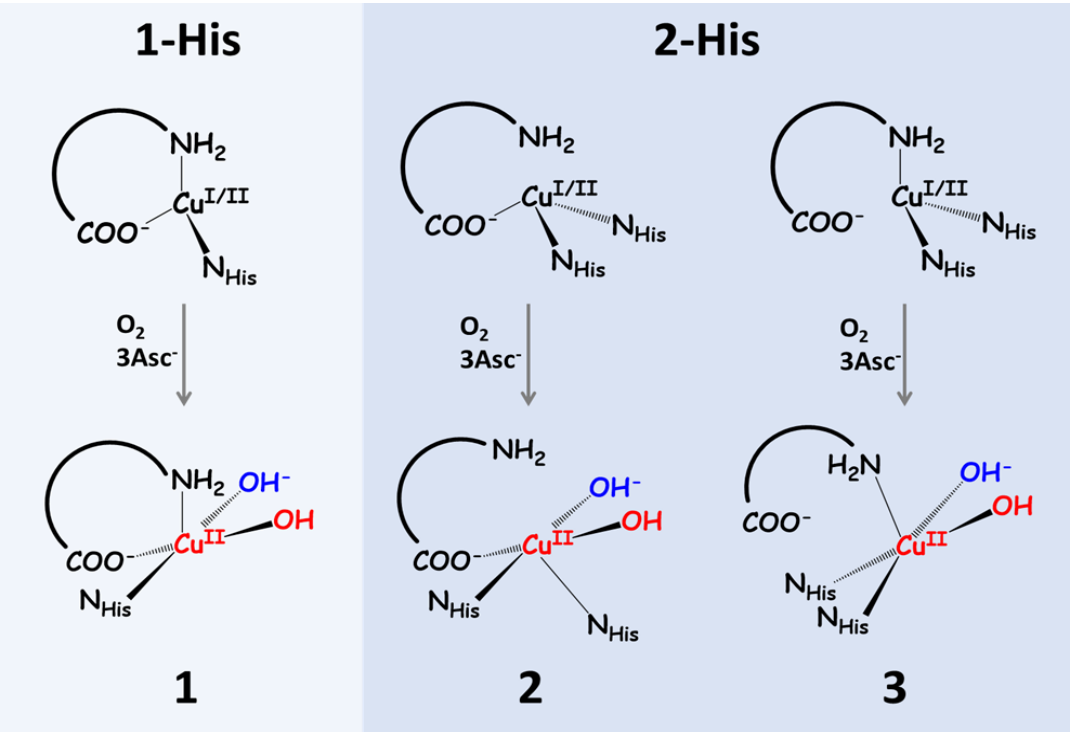


Figure 2. Cu(II)/Cu(I)·Aβ active coordinations in the $^3O_2$ three-step reduction cycle by ascorbate (Asc^-) along with the corresponding Cu(II)$(OH)_2^-$·Aβ final products. 1-His coordination (**1**) was proposed from experiments in ref. (48) while 2-His coordinations (**2** and **3**) were proposed according to DFT computations in ref. (51)

Starting from these results we investigated in detail the structure of final Cu(II)$(OH)_2^-$·Aβ coordinations, here simplified considering only the Asp1–Ala2 dipeptide and the

histidine side chains modelled as 4-methyl imidazole rings. We considered both mono (1-His, **1**) and bis histidine (2-His, **2** and **3**) models (Scheme S1 in SI) coordinated either at Nδ or Nε. The 1-His and 2-His most representative forms of the possible Cu(II) coordinations, as identified by DFT minimum geometry search, are sketched in Figure 3. Starting form model **1**, we identified 17 isomers in an energy range of 10.4 kcal/mol, while from **2** and **3** the found 9 structures in an energy range of 16 kcal/mol (see Table S1 and S2 in SI).

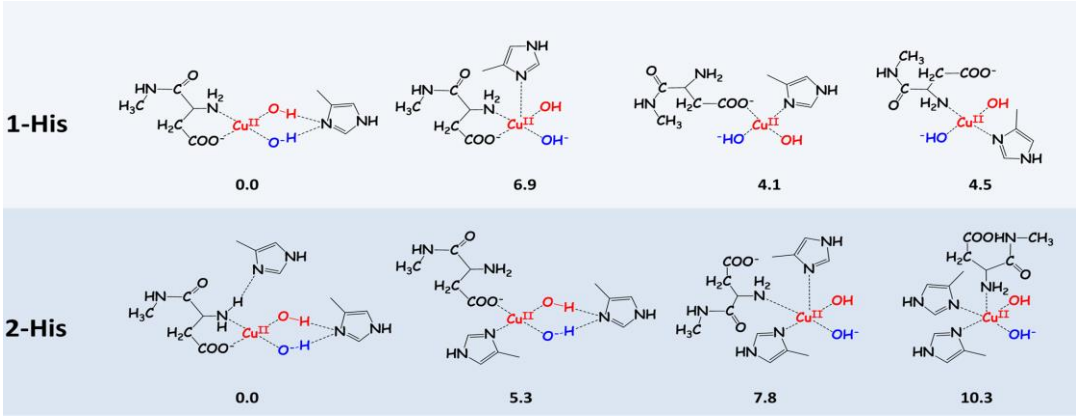


Figure 3. Most relevant 1-His and 2-His Cu(II)$(OH)_2^-$-Aβ models proposed according to DFT geometry optimizations. Energy differences against the most stable 1-His and 2-His are reported in kcal/mol.

The most stable form is always four-coordinated with no His side chains bound to the metal centre. The first genuine **1**-type penta-coordinated model is 6.9 kcal/mol higher in energy with His in apical coordination, while is 7.8 kcal/mol higher in energy for **2**- and **3**-type structures.

The ground state structure is a singlet state while the corresponding triplet state is always higher in energy and the singlet broken-symmetry solution collapses to the restricted singlet state.

The electronic structure of the ground state can be described as Cu(II)$(OH)_2^-$ in which one unpaired electron is delocalized over the two OH ligands since:

1. the optimized Cu-O bond distances are similar;
2. the negative charge is almost equally distributed on the two OH ligands;
3. in the corresponding triplet state one unpaired electron is localized on Cu and one on the two oxygen atoms of the OH ligands;

4. one or more occupied frontier MO result bonding with respect to both Cu-O interactions (see Scheme S2 in SI for details).

For the sake of completeness we further explore one **1**-type PES structure to evaluate (i) the relative stability of the Cu(II)(O^-)(H_2O)·Aβ oxyl form; (ii) propensity of the OH group to be protonated and detach from coordination with Cu. We observed that upon geometry optimization the oxyl form spontaneously evolves toward Cu(II)$(OH)_2^-$·Aβ form, while OH protonation is unfavorable by 19.8 kcal/mol. The details are reported in SI (Scheme S1a and S1b).

Stability of the Cu(II)$(OH)_2^-$·Aβ(1-6) coordinations.

In this section we explore at DFT level the conformational space of the full Cu(II)$(OH)_2^-$·Aβ(1-6) (the peptide is the C- terminus capped DAEFRH). Here we consider only a small number of structures and therefore these simulations are not intended to be exhaustive of the dynamical behavior of the system, but serve us to assess in a better detail:

(1) the effect of the Cu(II) second coordination sphere on the OH radical propagation toward tyrosine side chain (vide infra).
(2) the stability of a His6/Phe4 side chain pi-stacking interaction observed in the MD simulation.

On the basis of the mono histidine coordinations identified above, we carried out several restrained classical MD, imposing constraints on the Cu(II) coordination. From these MD runs we singled out 34 representative structures i) of the possible Cu(II) coordination; ii) of the possible interactions among Aβ(1-6) side chains. Each structure has been optimized first at classical level and then at DFT level (see Table S3a in SI for details) using the same level of theory adopted for small coordination model (but with def-TZVP basis set instead of def2-TZVP). The DFT optimized structures range in an energy interval of 20 kcal/mol and the first 10 are in an energy range of 5 kcal/mol. The relevant results are:

(1) lower energy structures are characterized by the formation salt bridge interaction that involve Arg5 with Asp1 or Glu3 side chains;
(2) among the first 10 structure we found four- and five-coordinated copper forms are almost isoenergetic;
(3) at least one OH ligand is exposed to solvent;
(4) in five-coordinated forms, in most of the case His side chain prefers to occupy apical position.

Cu(II)-Aβ coordinated OH radical is less oxidizing than solvated OH.

The OH radical in $Cu(II)(OH)_2^-$·Aβ is a weaker oxidant compared to free OH radical in water solution, due to the delocalization of the unpaired electron on both OH ligand through Cu(II), that decrease their radical character. To give a quantitative estimate of this, we evaluate the decreasing of the OH radical standard reduction potential E° in $Cu(II)(OH)_2^-$·Aβ model coordinations compared to solvated OH. Exploiting the results of a recent DFT study conducted in our laboratory (77) on water clusters at the same level of theory adopted in this paper, solvated OH latter has been modelled as a $OH(H_2O)_{20}$ cluster (see Figure S4 and Table S6 in SI for details) which is then reduced to $(H_2O)_{21}$ cluster. We compute the relative reduction potential ΔE°=E°(2)-E°(1) for the following half-reactions:

$$Cu(II)(OH)_2^-\cdot A\beta + e^- + H^+ \rightarrow Cu(II)(H_2O)(OH^-)\cdot A\beta \qquad E°(1)$$

$$OH(H_2O)_{20} + e^- + H^+ \rightarrow (H_2O)_{21} \qquad E°(2)$$

ΔΔE° does not depend on the potential value of the reference electrode. For the copper-coordinated OH radical computations we considered the four most stable 1-His four-coordinated forms and two five-coordinated forms, while for $OH(H_2O)_{20}$ and $(H_2O)_{21}$ clusters the lowest energy minimum structure found.

Depending on the $Cu(II)(OH)_2^-$·Aβ considered, we estimate a ΔE° value between -0.87 V to of -1.61 V, being for the most stable coordination -1.31 V. Interestingly, for the five-coordinated forms the ΔE° value is lower (-0.87 V) than four-coordinated because of the stabilization of the $Cu(II)(H_2O)(OH^-)$·Aβ reduced form, in which the water molecule leaves the Cu(II) coordination sphere.

Recalling that E° of the $OH^- + e^- + H^+ \rightarrow H_2O$ half-reaction is 2.730 V vs SHE (78), the E° value for $Cu(II)(OH)_2^-$·Aβ/$Cu(II)(H_2O)(OH^-)$·Aβ obtained as

$$E° = 2.73 - \Delta E°$$

between 1.12 and 1.88 V (1.31 V for the most stable form, see Figure 4 and Table S4 in SI), suggesting that OH^- would still be an oxidant species but with a much lesser oxidative capacity to solvated OH radical.

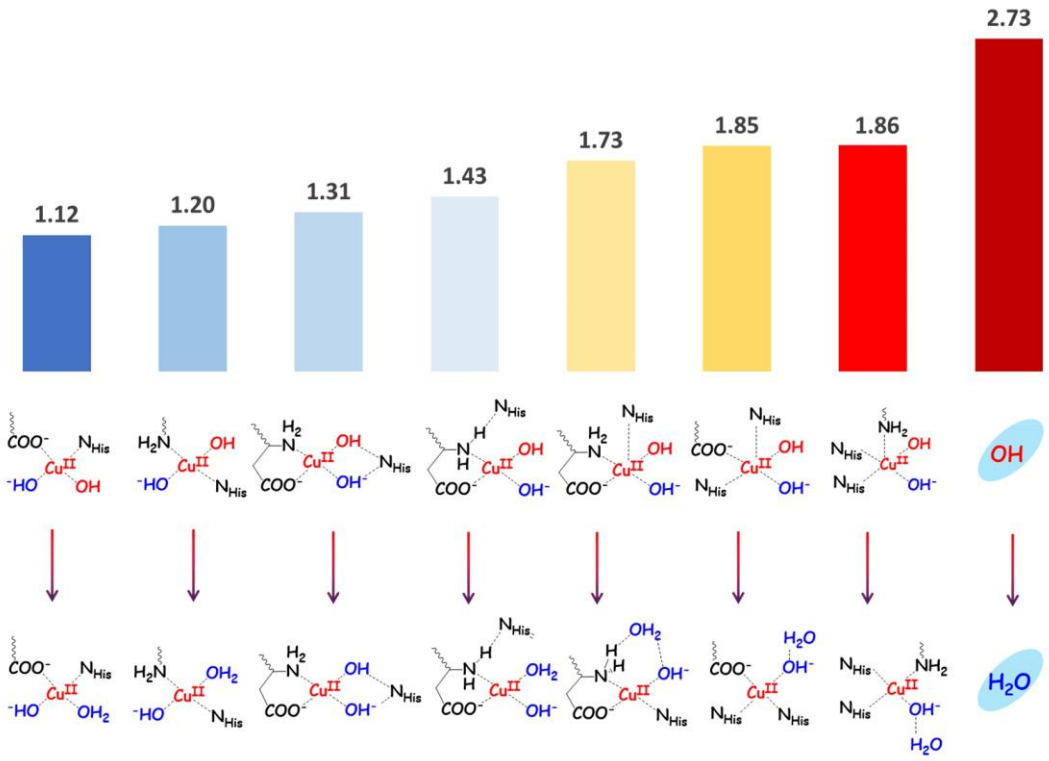


Figure 4. Estimated DFT values of the standard reduction potential E° (in V) for Cu(II)$(OH)_2^-$·Aβ/Cu(II)$(OH^-)(H_2O)$·Aβ half-reaction computed from relative ΔE° calculated with respect to the OH/H_2O half-reaction in which $OH(H_2O)_{20}$ and $(H_2O)_{21}$ clusters are adopted as model. The value of E° is then obtained by subtracting ΔE° to the tabulated OH/H_2O E° value of 2.73 V.

Upon Cu(II)$(OH)_2^-$·Aβ reduction and protonation 1) the Cu atomic charge results less positive by 0.12; 2) for five-coordinated forms the detach of the H_2O ligand from the Cu(II) coordination is observed, stabilizing the Cu(II)$(OH^-)(H_2O)$·Aβ structure, lowering the ΔE° values.

This results further corroborates the choice of Cu(II)$(OH)_2^-$·Aβ coordinations as a radical precursor. Indeed the estimated reduction potential difference between solvated and Cu(II)-coordinated OH excludes the possibility of the latter to leave the Cu(II) coordination. To better quantify this aspect we compute the energy difference for the following process

$$Cu(II)(OH)_2^-\cdot A\beta + H_2O \rightleftarrows Cu(II)(OH^-)\cdot A\beta + (OH)\cdot H_2O$$

modelling water as a $(H_2O)_{21}$ cluster while for solvated OH is used the $OH(H_2O)_{20}$ cluster considered above. We found this process thermodynamically unfavorable by 25.5 kcal/mol (See Scheme S1c in SI for details).

General OH radical propagation modelling.

OH is a highly electrophilic radical that propagates oxidizing electron-rich protein sites to pair its open shell electronic structure. (79)
In the present case, depending on the amino acid considered, we will observe the H abstraction (i.e the propagation of an electron/proton couple) from the residue side chain toward the OH radical with the formation of a H_2O molecule (as in the case of Asp and Tyr) or the hydroxylation of the side chain (as in the case of His). In general, we describe the process as:

$$Cu(II)(OH)_2^{-}\cdot A\beta + Res \rightleftarrows Cu(II)(OH)_2^{-}\cdot A\beta\cdot Res \rightleftarrows Cu(II)(OH)_2^{-}\cdot A\beta\cdot Res^{+\bullet}$$

Here we want to underline the formation of an inner sphere intermediate between $Cu(II)(OH)_2^{-}\cdot A\beta$ and the residue (Res) side chain which is characterized by the approach of the OH group to the atom undergoing the attack of the radical.

OH radical propagation and Tyrosyl radical formation.

The OH radical propagates toward tyrosine side chains with the formation of the Tyrosyl radicals. These latter are still reactive species that can react with each other with the formation of the dityrosyl bridges, as observed in the case of Aβ(1-40) experiments described above.
Recalling i) the evaluation of 1.31 V of the reduction potential for the $Cu(II)(OH)_2^{-}\cdot A\beta$ proposed above and ii) the standard reduction potential for the Tyr-O$^{\bullet}$/Tyr-OH half-reaction equal to 0.732 V vs SHE,(78) we can conclude, on the basis of the general relation $\Delta G=-nF\Delta E°$ ($F=23.06$ kcal$\cdot V^{-1}mol^{-1}$, n is the number of electrons involved in the redox process in this case equal to 1, $\Delta E°=0.58$ V), that this propagation process is in principle thermodynamically favoured by roughly 13.3 kcal/mol.
We investigate 10 different $Cu(II)(OH)_2^{-}\cdot A\beta\cdot Tyr$ model adducts, with a Tyr residue (here modelled as a 4-methylphenol molecule, see Table S5 in SI for details).
When triplet state MO occupancy is imposed to the S=0 ground state, the H abstraction is spontaneously observed during the geometry optimization, with a localization of one unpaired electron on the phenol group. When the S=0 BS MO occupancy is imposed, the geometry optimization converges to a more stable form compared to the S=1 one.

In Figure 5 is reported the scan of the Potential Energy Surface (PES) Tyr adduct along the H abstraction coordinate for the most stable 1-His model coordination. The process is slightly endothermic by 2.4 kcal/mol with a free energy barrier of 3.3 kcal/mol.

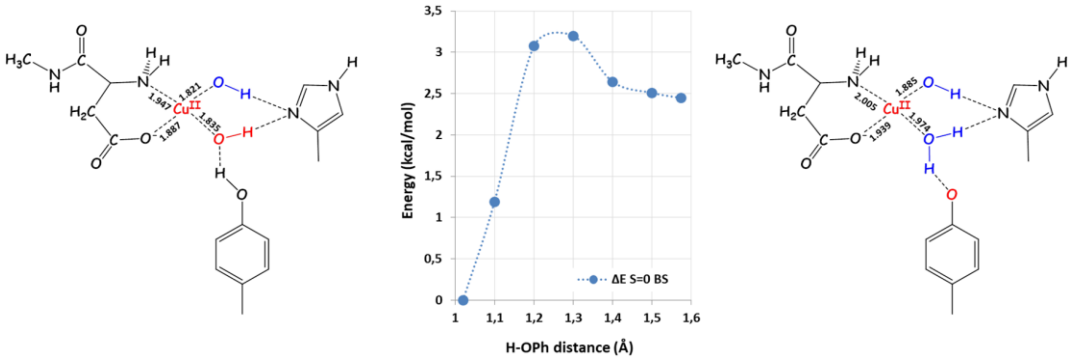


Figure 5. Scan of the most stable 4-coordinated Cu(II)$(OH)_2$·Aβ model in the Tyrosine (Tyr10) side chain adduct along the H abstraction coordinate. On the left S=0 Tyrosine adduct starting point; on the right the S=0 broken symmetry final product for the $OH^{\cdot}$ propagation toward Tyrosine side chain for the most stable Cu(II)$(OH)_2^{-}$·Aβ coordination. Bond distances in Å.

Upon H abstraction, all the bond distances of the Cu(II) coordination bonding network increase due i) to the loss of a charge-charge electrostatic interaction in the Cu/H_2O bonding and ii) to a charge-transfer from OH^- to Cu(II) of 0.1 e. For four-coordinated models, these destabilization effects do not seem sufficiently counterbalanced by the stabilization deriving from delocalization of the unpaired electron on the phenoxyl ring, and thus the overall process results slightly endothermic. For five-coordinated models, instead, the Cu/H_2O interaction destabilizes less the coordination and the process results slightly exothermic.

However, the electron transfer mediated by Tyr residues in copper bioinorganic system also depends on the Cu second sphere coordination environment, as recently shown in the copper monooxygenase active site. (80) To investigate this aspect, we repeated the above described investigation using the full Cu(II)$(OH)_2^{-}$·Aβ(1-6) DFT optimized models (see Table S3b in SI for details). Among the 34 computed structures, we considered the 10 most representative of the various possible Cu(II) coordinations, computing the Tyr adduct in the S=0, S=1 and S=0 BS state. On average, the propagation process results slightly exothermic by -3.9 kcal/mol, that corresponds to a

ΔE° of 0.17 V. This latter value, which is decreased compared to 0.578 V, could be the results of a further stabilization of the tyrosine side chain when involved in the adduct with $Cu(II)(OH)_2^-$·Aβ.

Histidine oxidation.

The reaction of OH radical with His side chain starts with the OH radical adduction at C2 of the imidazole ring, followed first by a reduction and then by a protonation with the formation of a 2-oxo imidazole ring (see Figure 6) (81) (82)

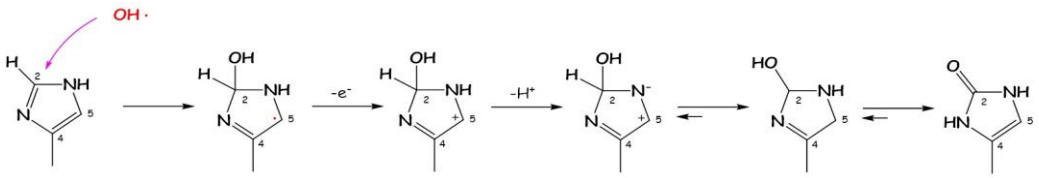


Figure 6. The mechanism of the 2-oxo histidine formation according to ref (81) and (82).

We explore the PES of three $Cu(II)(OH)_2^-$·Aβ·His adduct models sketched in Figure 7 along the his oxidation pathway by shortening of the C-O distance between C2 of the imidazole ring and the closest oxygen atom that belongs to OH group coordinated to copper.

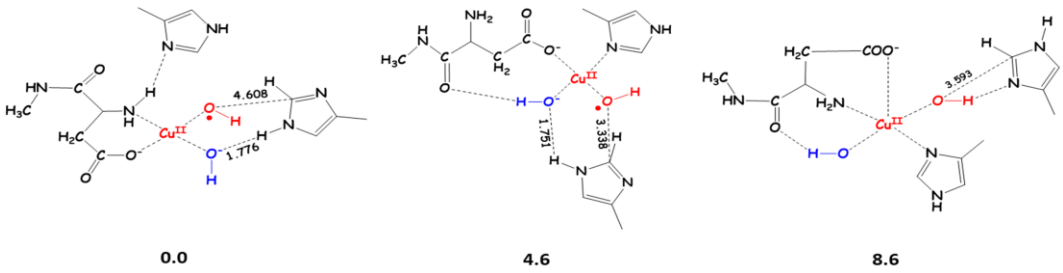


Figure 7. His adducts considered for the PES exploration along the His oxidation reaction coordinate. Energy differences in kcal/mol, distances in

The results of these PES scans are reported in Figures S1 in SI. In Figure 8 is reported a scheme summarizing PES scan results for the most stable His adduct.

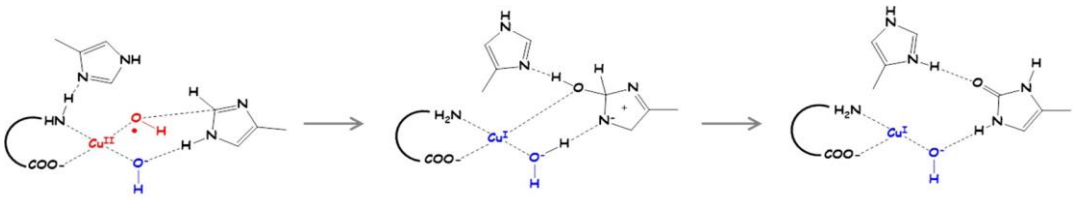


Figure 8. Most stable His adduct reactivity along the His oxidation reaction coordinate. The first step is the binding of OH to the C2 of the imidazole ring, that also causes the concomitant reduction of tetra-coordinated Cu(II) to three-coordinated Cu(I). The deprotonation step is assisted by a non-coordinating His residue. The final rearrangement at the C2 with two proton shifts yields the formation of the 2-oxo imidazole.

Starting from the genuine Cu(II) minimum forms, the scans show a maximum of energy at C-O distance around 1.75/2.00Å with a gain of density by copper, suggesting that its reduction has started. The computed free energy barrier is on average 15.9 ± 1.7 kcal/mol. Overcoming the energy barrier, the mechanism observed is the same for the three structures considered:

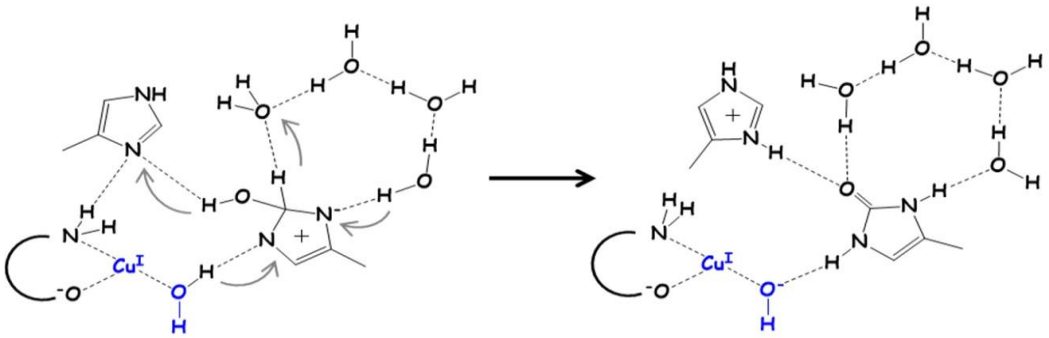
(1) oxidation of the imidazole ring by four-coordinated Cu(II) which is reduced to three- or bis-coordinated Cu(I);
(2) rearrangement at the imidazole ring C2 with deprotonation assisted by a non-coordinating His side chain.

In detail, we observed that the deprotonation of the imidazole NH function is performed by the OH^- coordinated to copper that is not involved in the C2 attack, forming a water molecule that can leave the copper coordination sphere. The final intermediate is characterized by a two-coordinated Cu(I) ion and a zwitterionic form of the 2-oxo imidazole. The next steps are:

(1) the rearrangement at the C2 with the shift of the proton to Nε;
(2) the shift of the proton on the OH on C2 to Nδ with the final formation of the 2-oxo imidazole ring.

In the two passages we observed a further gain of energy compared to the reactant and the overall process is highly exergonic (on average 49.3 ± 10.5 kcal/mol).

In order to shed light on the thermodynamic of the C2 rearrangement, we explore one of the PES of the zwitterionic form in which the deprotonated Nδ and the OH on C2 are explicitly solvated with 4 water molecules:
The idea is to scan the PES of the system along the coordinate of the concerted transfer of multiple protons from i) the C2 to the H_2O molecules network until saturation of the negative charge on Nδ ii) the oxygen atom of the OH group to the non-coordinating imidazole ring iii) the H_2O molecule coordinated to copper to the Nε of the imidazole (See Figure S2 in Si for details).



Using this model, we compute a free energy barrier for the C2 rearrangement of 13.1 kcal/mol.

In the mechanism proposed in Figure 7, we observed the active role of a non-coordinating histidine as protonable group in the C2 rearrangement. This role could also be played by a carboxylate group close to the Cu(II). To verify this possibility we investigate the His oxidation pathway for $Cu(II)(OH)_2^-$·Aβ model coordination in which a Glu/Asp side chain (modelled as a propionate anion) is interacting with the OH group or the His residues under attack. Computations (see SI) confirm that the carboxylate group is able to assist the proton-transfer from the imidazole ring under OH attack with free energy barriers of 15.2 kcal/mol.

Aspartate oxidation.

The Asp1 oxidative mechanism consists of two steps, which are sketched for the most stable $Cu(II)(OH)_2^-$·Aβ form in Figure 9.

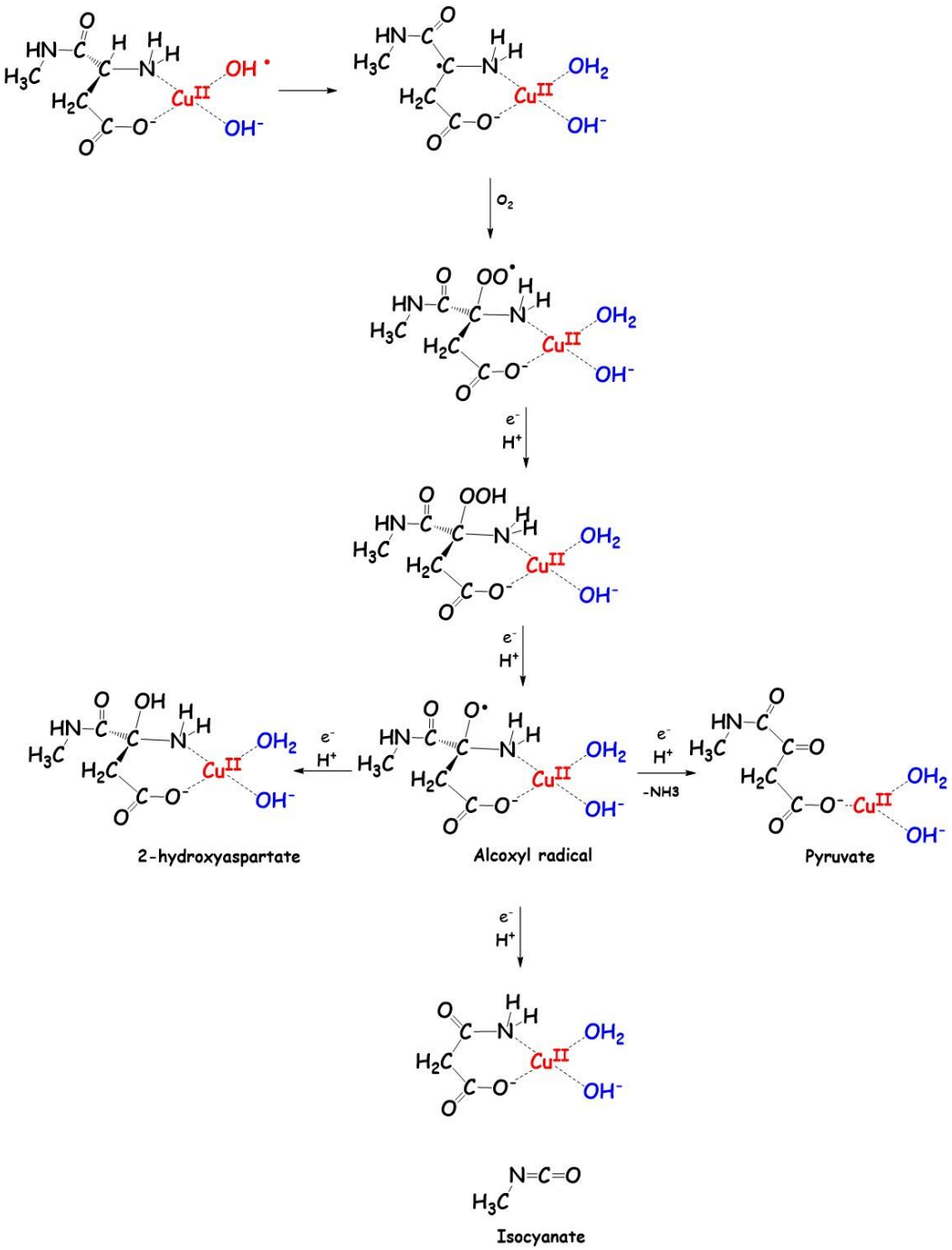


Figure 9. First step of the Asp1 oxidative mechanism with the formation of the alkoxyl radical upon O_2 addiction at the Asp1 Cα. Second step of the Asp1 oxidative mechanism. Starting from the alkoxyl radical, the three possible products are shown.

In the first step, an alkoxyl radical is formed through the initial hydrogen abstraction from Cα of Asp1 followed by 1) the O_2 addiction at the Cα, 2) a first reduction/protonation with the formation of a peroxyl radicals and finally 3) a second reduction/protonation with the formation of an hydroperoxide radical which is converted to alkoxyl radical by water elimination. In the second step. The alkoxyl radical can evolve to three different products namely i) an isocyanate through the Asp1-Ala2 peptide bond cleavage ii) a pyruvate through an oxidative deamination with the formation of NH_3. iii) 2-hydroxyaspartate.

We investigate the Cα H abstraction coordinate which is the slow step in the Asp1 oxidation. Starting from the S=1 O_2 Van der Waals adducts of a selection of Cu(II)$(OH)_2^-$·Aβ coordinations, we explore the corresponding PES along the Cα H abstraction pathway. Since for this H abstraction it is necessary that the OH group approaches the Cα H, this implies that not all the coordinations will be equally reactive. In Figure 10 the four considered coordinations are sketched (structures **1-4**).

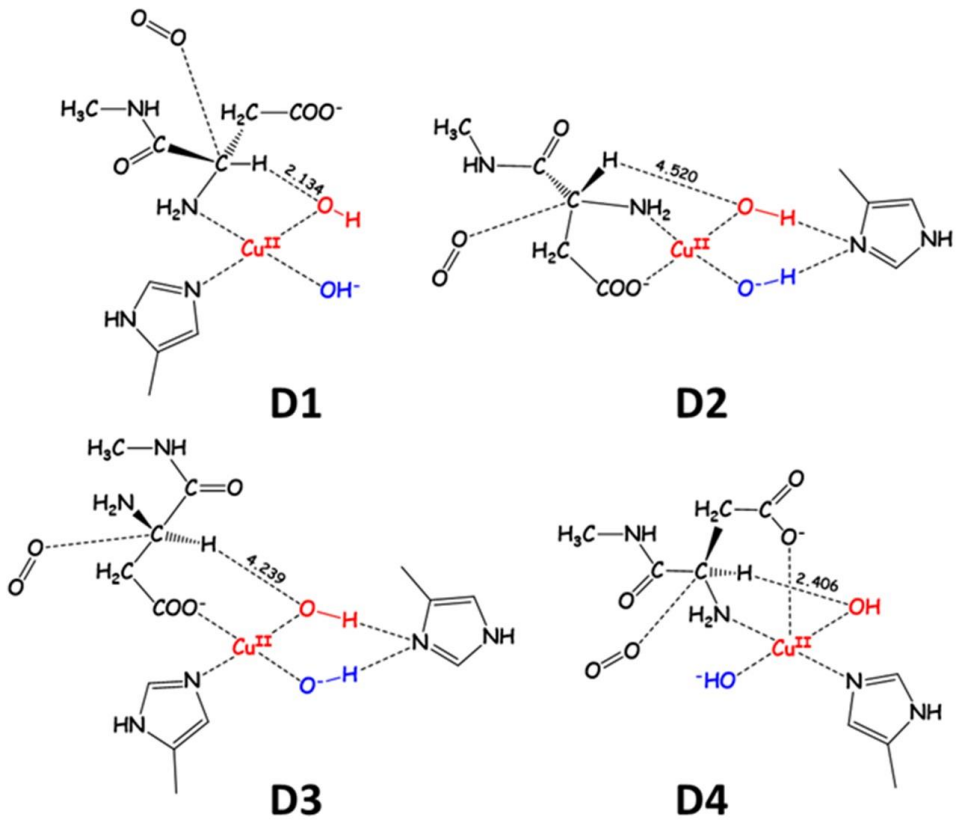


Figure 10. S=1 Cu(II)$(OH)_2$··Aβ·$^3O_2$ geometry optimized models considered for the PES exploration along the Asp1 oxidation reaction coordinate. H···O distance between H of Cα and O of OH (in Å).

The $^3O_2$ binding free energy is endergonic on average by 10.9 kcal/mol due to very weak charge–dipole interaction and entropic effects. The scans of the PES along the H abstraction coordinate are reported in the SI (Figure S3). The free energy barriers are between 6.6 to 10.0 kcal/mol (on average 7.9 ± 1.5 kcal/mol), except for structure **D2** (24.9 kcal/mol). This latter higher value is due to the significant molecular distortion induced by the H-O distance shortening. Finally, the overall stabilization due to the formation of the peroxyl radicals is 20.0 kcal/mol on average.

Cu(II)-$(OH)_2^-$ Aβ(1-16) full model investigated by MD.

The entire OH propagation process can be ideally decomposed in two steps: first the approaching of the copper bound OH radical to the side chain of the residue involved in the propagation, and second the actual propagation of the radical. The second step has been elucidated by DFT modelling. Indeed, in the previous sections we presented the mechanism of OH propagation toward histidine, aspartate and tyrosine side chains, on a Cu(II)$(OH)_2^-$·Aβ minimal model (the metal coordination plus the residues side chains involved in the propagation processes). However, this thermodynamic and kinetic characterization of the propagation processes by only focusing on the metal centre gives partial information regarding which is the most favoured propagation process, since it clearly leaves out the dynamic of the reactive centres approaching each other. DFT is not able, in fact, to give us information about the first step of the process. The approach adopted to shed light to the first point is the molecular dynamic of the Cu(II)$(OH)_2^-$·Aβ(1-16) full residue models with the aim of evaluating the type and the frequency of the intrapeptide OH radical/side chain approaches:

$$Cu(II)(OH)_2^-\cdot A\beta(1\text{-}16) \rightleftarrows Cu(II)(OH)_2^{2-}\cdot A\beta_{ox}(1\text{-}16)$$

Here we considered the Cu(II) N1O3 planar coordination (Asp1 N terminal amino group and carboxylate + 2OH group with restrains) while His residues are unconstrained. Since the low Aβ(1-16) concentration in the experiments (0.5/1 mM), we exclude at this level the interpeptide OH approaches.

We carried out five 1-μs long independent MD simulations. According to the analysis on the 5 μs merged trajectory, the Aβ(1-16) is very flexible and no secondary structures have been observed persisting for more than 5-10 ns, confirming the analysis on previous Cu(II)·Aβ MD simulations. (83) (84) (85) Among the possible opposite charge interactions, the most frequent one is Arg5-Asp1 and Arg5-Glu3 (18.9% and 13.0% of the whole simulation time, respectively). The average atomic distances between the hydrogens of the Cα that belong to the Asp1 side chain and the Oxygen atoms of the OH groups are 3.5 ± 0.1 Å and 4.3 ± 0.1 Å. The low standard deviation of these two values testifies the rigidity of this coordination and *therefore* we conclude that the Asp1 propagation process is always the favored form both from the thermodynamic and the kinetic point of view, in agreement with all the experimental investigations where the Asp1 oxidized products have been observed.

In the analysis of the MD trajectory, we classify a given frame as "reactive" if the distance between the two atoms involved in the propagation step is equal to the Lennard-Jones potential r_{min} parameter as defined in the force field (see SI for details). In detail the atoms considered are the oxygen atoms that belong to one of the OH groups bound to Cu(II) and i) the C2 atom of the His6, His13 and His14 side chain and ii) the hydrogen atom of the OH that belong to the Tyr10. The percentage of "reactive" frames as plotted in Figure 11 are low and decrease in the series His6 (0.7%) >His14 (0.4%) >His13 (0.2%) >Tyr10 (0.1%).

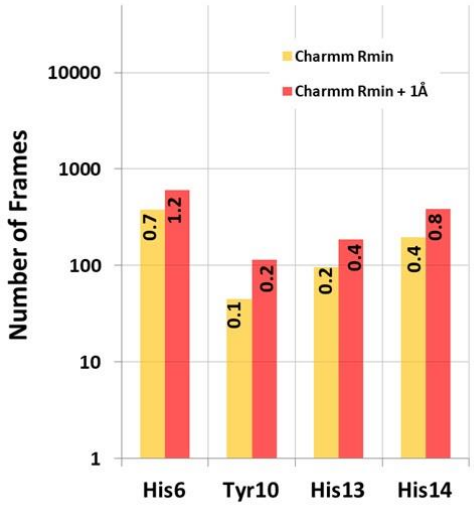


Figure 11. Number of "reactive" frames in $Cu(II)(OH)_2$-Aβ(1-16) 5 μs MD simulation in which His6/13/14 or Tyr10 the atomic distances between the atoms involved in the OH propagation process are less or equal to the force field Lennard-Jones r_{min} parameter (yellow) or r_{min}+1Å (orange). The trajectory has been sampled every 100 ps collecting 5000 frames. On each bar the percentage of the "reactive" frames is reported with respect to the number of analysed frames.

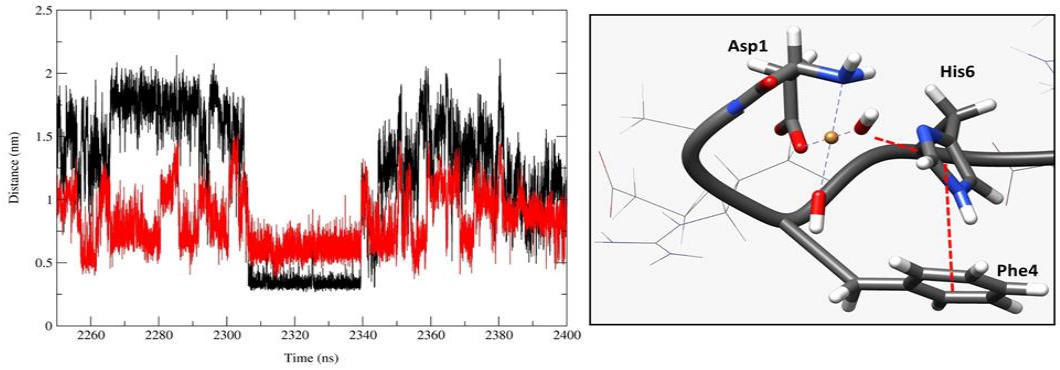


Figure 12. On the left: observed trend during the Cu(II)$(OH)_2^-$·Aβ(1-16) MD trajectory 35 ns portion related to the approaching of the centre-of-mass distance between Phe4 and His6 side chains (in black), which correlates with the approaching of the His6 C2 atom of the imidazole ring and the oxygen of one OH ligand on Cu(II) (in red). All distances are in nm. On the right: sketch of the structure of Cu(II) coordination environment in one of the MD frame along the 35 ns portion. Red dotted lines underlined the HO-C2 distance and the His6/Phe4 side chain distances.

The analysis of the trajectory reveals that the Lys16-Asp1/Glu3 interactions are not persistent during the simulation (5.4 and 4.9 %, respectively), but fairly correlated (0.78-0.80) with His13/His14 approaching to OH in terms of the radius of gyration (Rg). As one could expect, the formation of the coulombic interactions help the head and tail of Aβ(1-16) to get closer, thus making possible for the His13/His14 side chain to approach the OH radical at Asp1.

The case of His6 approaching seems not correlated to a salt-bridge formation but instead to a persistent π-stacking interaction (20% with Rg correlation equal to 0.35) between His6 and Phe4 aromatic rings, which has been also observed continuous for a time interval of 35 ns as shown in Figure 12. The T-shaped configuration observed (Figure 12 on the right) is typical of this weak interaction (86) which is able to assist the His6/OH approaching. To assess this, we carried out a DFT geometry optimization on the 1-6 portion (See SI for details) of a representative frame taken from the 35-ns MD interval, finding that Phe4 ring assists the rotation of His6 imidazole ring with the formation of the same type of NH/OH hydrogen bond reported in Figure 6 models.

Discussion and conclusion

The DFT and classical modelling led to a picture of the OH propagation toward protein oxidation in Cu(II)$(OH)_2^-$·Aβ, revealing new and interesting aspects.
Cu(II)$(OH)_2^-$·Aβ is the final product of the three-steps Fenton-like cycle of O_2 reduction with ascorbate as reducing agent, and the initial reactant of the OH propagation process. DFT minimum search for Cu(II)$(OH)_2^-$·Aβ minimal coordination model shows that four-coordinated Cu(II) forms are favoured at least by 6.6 kcal/mol compared with the most stable five-coordinated form. Cu(II) coordination is of N1O3 type, with both amino terminal and carboxylate groups of Asp1 and the two OH ligands bound to Cu(II). When a larger part of the secondary coordination sphere of Cu(II) is considered in the full Cu(II)$(OH)_2^-$·Aβ(1-6), four- and five-coordinated forms are very similar in energy.
The electronic structure of these coordinations can be described as a Cu(II) bound to two OH ligands that share one unpaired electron, thus suggesting that their oxidative power is almost halved compared with free solvated OH. Indeed, the most relevant feature revealed by these computations regards their oxidative power. We estimate that standard reduction potential of OH in Cu(II)$(OH)_2^-$·Aβ could be around 1.31 V compared to the 2.73 V value for solvated OH.
This fact potentially might have some important consequences in the chemistry of AD. Indeed *Aβ might possess a sort of double redox neuroprotective role* since 1) Aβ coordinated to Cu(II) is less efficient in OH production than free Cu(II) in solution; 2) OH coordinated to Cu(II)Aβ has a lower oxidative capacity compared to free OH in solution. However, at the same time, we recall that due to its high reactivity, OH radical half-life in biological system is extremely short (around 1 ns (9)) and for this reason is not able to reach high concentration in aqueous solution(87). *OH radical in Cu(II)$(OH)_2^-$·Aβ will be less oxidant but, at the same time, will have a much longer lifetime that allows it to survive long enough to diffuse into the cell membrane lipid bilayer, starting lipid peroxidation, or to reach and oxidize other proteins or nucleic acids*. Indeed, we estimate that OH standard reduction potential in Cu(II)$(OH)_2^-$·Aβ is still sufficiently high to induce radical propagation and protein oxidation (probably with a longer time scale). *Paradoxically, the initial redox protective role of Cu(II) coordinated to Aβ could be the causative agent of its neurotoxicity*.

We investigate the OH propagation toward Asp, Tyr and His side chains in two steps. In the first, focusing on the Cu(II) coordination, we computed at DFT level the energetic and the free energy barriers for the H/OH abstraction slow step. In the second we investigate using MD simulations the kinetic of the intrapeptide OH/side chain approach on the full Cu(II)$(OH)_2^-$·Aβ(1-16).

Since Asp1 is coordinated to Cu(II), Asp1 Cα is constantly close to OH and therefore Asp1 oxidation is not affected by the dynamic of the OH approaching to the reactive centres. Asp1 also has a low free energy barrier for H abstraction (7.9 kcal/mol) but to complete its oxidation O_2 addiction is required, which is thermodynamically unfavoured due to the low $Cu(II)(OH)_2^-{\cdot}A\beta/O_2$ affinity. We conclude that Asp1 oxidation is globally favoured in agreement with the literature evidence, since in the *in vitro* experiments the O_2 concentration can be considered high enough to allow Asp1 oxidation completion.

In Tyr and His oxidation, we observe two opposite trends. We evaluate from MD simulation that Tyr approaching to OH is more than 10 time less frequent than for one of the three His residues, but while free energy barrier for H abstraction from Tyr is only few kcal/mol, the one associated with OH attack to His side chain is 15.9 kcal/mol. MD simulations help us to discern between these two scenarios.

Indeed, we observe that His6 and and His13/His14 approaching are assisted is by Phe4/His6 π-stacking and Lys16/Asp1, Glu3 salt-bridge respectively, while Tyr10 approach seems much less sensitive to a particular non-bonding interaction and rather determined by the causality of the peptide dynamic. Therefore our results seems to confirm the finding from Cu(II)·Aβ(1-16) experiments where Tyr-Tyr cross-link through Tyr10 oxidation is not observed.

Finally, it is not straightforward to assess which is the most reactive His residue toward OH propagation, since MD results do not provide a clear-cut answer to this issue, but according to Figure 10 we tentatively conclude that His6 is preferred over His14 and His13.

We conclude with a final comment on the dityrosine cross-links. Although the results presented above essentially deal with the OH propagation for Aβ(1-16), some considerations can be made also in the case of longer Aβ peptide. The fact that the formation of the cross-links is observed in the experiments with Aβ(1-40) (according to ref(63) only 3% of the initial peptide concentration) can be explained as following. Cu(II) initially would promote the formation of Aβ dimers in which the metal coordination is still exposed to the solvent, allowing the interaction with O_2 and the reducing agent to form OH radicals. Among these dimers, the 1:1 $[Cu(II){\cdot}A\beta]_2$ are those likely prone to Tyr-Tyr cross-link formation, since

1. Cu(II) ions and the hydrophobic interactions at the Aβ C-terminus could constrain the Aβ flexibility enough to keep Aβ oxidation target residues close to the metal center, and therefore to the OH radical;
2. two tyrosinate radical are needed to form the Tyr-Tyr cross-link.

Once OH radicals are formed at the dimers, these have an attenuated oxidative capacity and could survive long enough to propagate. When OH approaches Tyr side chain, the

energy barrier for the propagation is low and the tyrosinate radical are formed, inducing their oxidative dimerization. The rate of the overall process depends on many factors that could slow it down. In order to clarify this suggested picture, the reactivity and dynamics of these dimers toward OH propagation are currently under investigation in our laboratory.

This paper is dedicated to the memory of our dear friend and co-worker Prof. Maurizio Bruschi who passed away on February 17, 2020.

References

1. Querfurth HW, LaFerla FM. Alzheimer's Disease [Internet]. Vol. 362, New England Journal of Medicine. 2010. p. 329–44.

2. Michael B, Ellul M. F1000Prime recommendation of Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016 [Internet]. F1000 - Post-publication peer review of the biomedical literature. 2019.

3. Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. Science. 1992 Apr 10;256(5054):184–5.

4. Ricciarelli R, Fedele E. The Amyloid Cascade Hypothesis in Alzheimer's Disease: It's Time to Change Our Mind [Internet]. Vol. 15, Current Neuropharmacology. 2017.

5. Butterfield DA, Halliwell B. Oxidative stress, dysfunctional glucose metabolism and Alzheimer disease. Nat Rev Neurosci. 2019 Mar;20(3):148–60.

6. Greenough MA, Camakaris J, Bush AI. Metal dyshomeostasis and oxidative stress in Alzheimer's disease [Internet]. Vol. 62, Neurochemistry International. 2013. p. 540–55. Available from: http://dx.doi.org/10.1016/j.neuint.2012.08.014

7. Cheignon C, Tomas M, Bonnefont-Rousselot D, Faller P, Hureau C, Collin F. Oxidative stress and the amyloid beta peptide in Alzheimer's disease. Redox Biol. 2018 Apr; 14:450–64.

8. Skoumalová A, Hort J. Blood markers of oxidative stress in Alzheimer's disease. J Cell Mol Med. 2012 Oct;16(10):2291–300.

9. Halliwell B. Free Radicals in Biochemistry and Medicine [Internet]. Encyclopedia of Molecular Cell Biology and Molecular Medicine. 2006. Available from: http://dx.doi.org/10.1002/3527600906.mcb.200300036

10. Butterfield DA. BRAIN LIPID PEROXIDATION AND ALZHEIMER DISEASE: SYNERGY BETWEEN THE BUTTERFIELD AND MATTSON LABORATORIES. Ageing Res Rev. 2020 Mar 20;101049.

11. Bradley-Whitman MA, Lovell MA. Biomarkers of lipid peroxidation in

Alzheimer disease (AD): an update. Arch Toxicol. 2015 Jul;89(7):1035–44.

12. Lovell MA, Ehmann WD, Mattson MP, Markesbery WR. Elevated 4-Hydroxynonenal in Ventricular Fluid in Alzheimer's Disease [Internet]. Vol. 18, Neurobiology of Aging. 1997. p. 457–61. Available from: http://dx.doi.org/10.1016/s0197-4580(97)00108-5

13. Praticò D, MY Lee V, Trojanowski JQ, Rokach J, Fitzgerald GA. Increased F2-isoprostanes in Alzheimer's disease: evidence for enhanced lipid peroxidation in vivo. FASEB J. 1998 Dec;12(15):1777–83.

14. Moreira PI, Nunomura A, Nakamura M, Takeda A, Shenk JC, Aliev G, et al. Nucleic acid oxidation in Alzheimer disease [Internet]. Vol. 44, Free Radical Biology and Medicine. 2008. p. 1493–505. Available from: http://dx.doi.org/10.1016/j.freeradbiomed.2008.01.002

15. Lovell MA, Soman S, Bradley MA. Oxidatively modified nucleic acids in preclinical Alzheimer's disease (PCAD) brain [Internet]. Vol. 132, Mechanisms of Ageing and Development. 2011. p. 443–8. Available from: http://dx.doi.org/10.1016/j.mad.2011.08.003

16. Butterfield DA, Boyd-Kimball D. Redox proteomics and amyloid β-peptide: insights into Alzheimer disease. J Neurochem. 2019 Nov;151(4):459–87.

17. Huang X, Moir RD, Tanzi RE, Bush AI, Rogers JT. Redox-active metals, oxidative stress, and Alzheimer's disease pathology. Ann N Y Acad Sci. 2004 Mar; 1012:153–63.

18. Bush AI, Tanzi RE. Therapeutics for Alzheimer's disease based on the metal hypothesis [Internet]. Vol. 5, Neurotherapeutics. 2008. p. 421–32. Available from: http://dx.doi.org/10.1016/j.nurt.2008.05.001

19. Kepp KP, Squitti R. Copper imbalance in Alzheimer's disease: Convergence of the chemistry and the clinic [Internet]. Vol. 397, Coordination Chemistry Reviews. 2019. p. 168–87. Available from: http://dx.doi.org/10.1016/j.ccr.2019.06.018

20. Lovell MA, Robertson JD, Teesdale WJ, Campbell JL, Markesbery WR. Copper, iron and zinc in Alzheimer's disease senile plaques [Internet]. Vol. 158, Journal of the Neurological Sciences. 1998. p. 47–52. Available from:

http://dx.doi.org/10.1016/s0022-510x(98)00092-6

21. James SA, Churches QI, de Jonge MD, Birchall IE, Streltsov V, McColl G, et al. Iron, Copper, and Zinc Concentration in Aβ Plaques in the APP/PS1 Mouse Model of Alzheimer's Disease Correlates with Metal Levels in the Surrounding Neuropil. ACS Chem Neurosci. 2017 Mar 15;8(3):629–37.

22. Hung YH, Bush AI, Cherny RA. Copper in the brain and Alzheimer's disease. J Biol Inorg Chem. 2010 Jan;15(1):61–76.

23. Sensi SL, Granzotto A, Siotto M, Squitti R. Copper and Zinc Dysregulation in Alzheimer's Disease. Trends Pharmacol Sci. 2018 Dec;39(12):1049–63.

24. Barnham KJ, McKinstry WJ, Multhaup G, Galatis D, Morton CJ, Curtain CC, et al. Structure of the Alzheimer's disease amyloid precursor protein copper binding domain. A regulator of neuronal copper homeostasis. J Biol Chem. 2003 May 9;278(19):17401–7.

25. Kong GK-W, Adams JJ, Cappai R, Parker MW. Structure of Alzheimer's disease amyloid precursor protein copper-binding domain at atomic resolution [Internet]. Vol. 63, Acta Crystallographica Section F Structural Biology and Crystallization Communications. 2007. p. 819–24. Available from: http://dx.doi.org/10.1107/s1744309107041139

26. Spoerri L, Vella LJ, Pham CLL, Barnham KJ, Cappai R. The Amyloid Precursor Protein Copper Binding Domain Histidine Residues 149 and 151 Mediate APP Stability and Metabolism [Internet]. Vol. 287, Journal of Biological Chemistry. 2012. p. 26840–53. Available from: http://dx.doi.org/10.1074/jbc.m112.355743

27. Wild K, August A, Pietrzik CU, Kins S. Structure and Synaptic Function of Metal Binding to the Amyloid Precursor Protein and its Proteolytic Fragments. Front Mol Neurosci. 2017 Jan 31; 10:21.

28. Young TR, Pukala TL, Cappai R, Wedd AG, Xiao Z. The Human Amyloid Precursor Protein Binds Copper Ions Dominated by a Picomolar-Affinity Site in the Helix-Rich E2 Domain [Internet]. Vol. 57, Biochemistry. 2018. p. 4165–76. Available from: http://dx.doi.org/10.1021/acs.biochem.8b00572

29. Cai H, Wang Y, McCarthy D, Wen H, Borchelt DR, Price DL, et al. BACE1 is

the major β-secretase for generation of Aβ peptides by neurons [Internet]. Vol. 4, Nature Neuroscience. 2001. p. 233–4. Available from: http://dx.doi.org/10.1038/85064

30. Cline EN, Bicca MA, Viola KL, Klein WL. The Amyloid-β Oligomer Hypothesis: Beginning of the Third Decade. J Alzheimers Dis. 2018;64(s1): S567–610.

31. Ahmed M, Davis J, Aucoin D, Sato T, Ahuja S, Aimoto S, et al. Structural conversion of neurotoxic amyloid-β1–42 oligomers to fibrils [Internet]. Vol. 17, Nature Structural & Molecular Biology. 2010. p. 561–7. Available from: http://dx.doi.org/10.1038/nsmb.1799

32. Stefaniak E, Bal W. CuII Binding Properties of N-Truncated Aβ Peptides: In Search of Biological Function [Internet]. Vol. 58, Inorganic Chemistry. 2019. p. 13561–77. Available from: http://dx.doi.org/10.1021/acs.inorgchem.9b01399

33. Fukumoto H, Tokuda T, Kasai T, Ishigami N, Hidaka H, Kondo M, et al. High-molecular-weight β-amyloid oligomers are elevated in cerebrospinal fluid of Alzheimer patients [Internet]. Vol. 24, The FASEB Journal. 2010. p. 2716–26. Available from: http://dx.doi.org/10.1096/fj.09-150359

34. Rana M, Sharma AK. Cu and Zn interactions with Aβ peptides: consequence of coordination on aggregation and formation of neurotoxic soluble Aβ oligomers [Internet]. Vol. 11, Metallomics. 2019. p. 64–84. Available from: http://dx.doi.org/10.1039/c8mt00203g

35. Young TR, Kirchner A, Wedd AG, Xiao Z. An integrated study of the affinities of the Aβ16 peptide for Cu(i) and Cu(ii): implications for the catalytic production of reactive oxygen species [Internet]. Vol. 6, Metallomics. 2014. p. 505–17. Available from: http://dx.doi.org/10.1039/c4mt00001c

36. Atrián-Blasco E, Gonzalez P, Santoro A, Alies B, Faller P, Hureau C. Cu and Zn coordination to amyloid peptides: From fascinating chemistry to debated pathological relevance [Internet]. Vol. 371, Coordination Chemistry Reviews. 2018. p. 38–55. Available from: http://dx.doi.org/10.1016/j.ccr.2018.04.007

37. Alí-Torres J, Maréchal J-D, Rodríguez-Santiago L, Sodupe M. Three Dimensional Models of Cu2 -Aβ(1–16) Complexes from Computational Approaches [Internet]. Vol. 133, Journal of the American Chemical Society.

2011. p. 15008–14. Available from: http://dx.doi.org/10.1021/ja203407v

38. Drew SC, Masters CL, Barnham KJ. Alanine-2 Carbonyl is an Oxygen Ligand in Cu2 Coordination of Alzheimer's Disease Amyloid-β Peptide − Relevance to N-Terminally Truncated Forms [Internet]. Vol. 131, Journal of the American Chemical Society. 2009. p. 8760–1. Available from: http://dx.doi.org/10.1021/ja903669a

39. Alies B, Badei B, Faller P, Hureau C. Reevaluation of Copper(I) Affinity for Amyloid-β Peptides by Competition with Ferrozine-An Unusual Copper(I) Indicator [Internet]. Vol. 18, Chemistry - A European Journal. 2012. p. 1161–7. Available from: http://dx.doi.org/10.1002/chem.201102746

40. Bagheri S, Squitti R, Haertlé T, Siotto M, Saboury AA. Role of Copper in the Onset of Alzheimer's Disease Compared to Other Metals. Front Aging Neurosci. 2017; 9:446.

41. Hane F, Tran G, Attwood SJ, Leonenko Z. Cu2 Affects Amyloid-β (1–42) Aggregation by Increasing Peptide-Peptide Binding Forces [Internet]. Vol. 8, PLoS ONE. 2013. p. e59005. Available from: http://dx.doi.org/10.1371/journal.pone.0059005

42. Pedersen JT, Østergaard J, Rozlosnik N, Gammelgaard B, Heegaard NHH. Cu(II) Mediates Kinetically Distinct, Non-amyloidogenic Aggregation of Amyloid-β Peptides [Internet]. Vol. 286, Journal of Biological Chemistry. 2011. p. 26952–63. Available from: http://dx.doi.org/10.1074/jbc.m111.220863

43. Sitkiewicz E, Kłoniecki M, Poznański J, Bal W, Dadlez M. Factors Influencing Compact–Extended Structure Equilibrium in Oligomers of Aβ1–40 Peptide—An Ion Mobility Mass Spectrometry Study [Internet]. Vol. 426, Journal of Molecular Biology. 2014. p. 2871–85. Available from: http://dx.doi.org/10.1016/j.jmb.2014.05.015

44. Strodel B, Coskuner-Weber O. Transition Metal Ion Interactions with Disordered Amyloid-β Peptides in the Pathogenesis of Alzheimer's Disease: Insights from Computational Chemistry Studies [Internet]. Vol. 59, Journal of Chemical Information and Modeling. 2019. p. 1782–805. Available from: http://dx.doi.org/10.1021/acs.jcim.8b00983

45. Pham DQH, Li MS, La Penna G. Copper Binding Induces Polymorphism in

Amyloid-β Peptide: Results of Computational Models [Internet]. Vol. 122, The Journal of Physical Chemistry B. 2018. p. 7243–52. Available from: http://dx.doi.org/10.1021/acs.jpcb.8b03983

46. Huy PDQ, Van Vuong Q, La Penna G, Faller P, Li MS. Impact of Cu(II) Binding on Structures and Dynamics of Aβ42 Monomer and Dimer: Molecular Dynamics Study [Internet]. Vol. 7, ACS Chemical Neuroscience. 2016. p. 1348–63. Available from: http://dx.doi.org/10.1021/acschemneuro.6b00109

47. Balland V, Hureau C, -M. Saveant J. Electrochemical and homogeneous electron transfers to the Alzheimer amyloid- copper complex follow a preorganization mechanism [Internet]. Vol. 107, Proceedings of the National Academy of Sciences. 2010. p. 17113–8. Available from: http://dx.doi.org/10.1073/pnas.1011315107

48. Cheignon C, Jones M, Atrián-Blasco E, Kieffer I, Faller P, Collin F, et al. Identification of key structural features of the elusive Cu-Aβ complex that generates ROS in Alzheimer's disease. Chem Sci. 2017 Jul 1;8(7):5107–18.

49. Mirats A, Alí-Torres J, Rodríguez-Santiago L, Sodupe M, La Penna G. Dioxygen activation in the Cu-amyloid β complex. Phys Chem Chem Phys. 2015 Nov 7;17(41):27270–4.

50. Atrián-Blasco E, Del Barrio M, Faller P, Hureau C. Ascorbate Oxidation by Cu(Amyloid-β) Complexes: Determination of the Intrinsic Rate as a Function of Alterations in the Peptide Sequence Revealing Key Residues for Reactive Oxygen Species Production. Anal Chem. 2018 May 1;90(9):5909–15.

51. Arrigoni F, Prosdocimi T, Mollica L, De Gioia L, Zampella G, Bertini L. Copper reduction and dioxygen activation in Cu-amyloid beta peptide complexes: insight from molecular modelling. Metallomics. 2018 Nov 14;10(11):1618–30.

52. Tabner BJ, Mayes J, Allsop D. Hypothesis: Soluble AβOligomers in Association with Redox-Active Metal Ions Are the Optimal Generators of Reactive Oxygen Species in Alzheimer's Disease [Internet]. Vol. 2011, International Journal of Alzheimer's Disease. 2011. p. 1–6. Available from: http://dx.doi.org/10.4061/2011/546380

53. Schöneich C, Williams TD. Cu(II)-Catalyzed Oxidation of β-Amyloid Peptide Targets His13and His14over His6: Detection of 2-Oxo-histidine by HPLC-

MS/MS [Internet]. Vol. 15, Chemical Research in Toxicology. 2002. p. 717–22. Available from: http://dx.doi.org/10.1021/tx025504k

54. Inoue K, Garner C, Ackermann BL, Oe T, Blair IA. Liquid chromatography/tandem mass spectrometry characterization of oxidized amyloid beta peptides as potential biomarkers of Alzheimer's disease [Internet]. Vol. 20, Rapid Communications in Mass Spectrometry. 2006. p. 911–8. Available from: http://dx.doi.org/10.1002/rcm.2395

55. Inoue K, Nakagawa A, Hino T, Oka H. Screening Assay for Metal-Catalyzed Oxidation Inhibitors Using Liquid Chromatography−Mass Spectrometry with an N-Terminal β-Amyloid Peptide [Internet]. Vol. 81, Analytical Chemistry. 2009. p. 1819–25. Available from: http://dx.doi.org/10.1021/ac802162n

56. Kowalik-Jankowska T, Ruta M, Wiśniewska K, Łankiewicz L, Dyba M. Products of Cu(II)-catalyzed oxidation in the presence of hydrogen peroxide of the 1–10, 1–16 fragments of human and mouse β-amyloid peptide [Internet]. Vol. 98, Journal of Inorganic Biochemistry. 2004. p. 940–50. Available from: http://dx.doi.org/10.1016/j.jinorgbio.2004.03.001

57. Nadal RC, Rigby SEJ, Viles JH. Amyloid β−Cu2 Complexes in both Monomeric and Fibrillar Forms Do Not Generate H2O2Catalytically but Quench Hydroxyl Radicals† [Internet]. Vol. 47, Biochemistry. 2008. p. 11653–64. Available from: http://dx.doi.org/10.1021/bi8011093

58. Cassagnes L-E, Hervé V, Nepveu F, Hureau C, Faller P, Collin F. The Catalytically Active Copper-Amyloid-Beta State: Coordination Site Responsible for Reactive Oxygen Species Production [Internet]. Vol. 125, Angewandte Chemie. 2013. p. 11316–9. Available from: http://dx.doi.org/10.1002/ange.201305372

59. Cheignon C, Hureau C, Collin F. Real-time evolution of Aβ 40 metal-catalyzed oxidation reveals Asp1 as the main target and a dependence on metal binding site [Internet]. Vol. 472, Inorganica Chimica Acta. 2018. p. 111–8. Available from: http://dx.doi.org/10.1016/j.ica.2017.07.031

60. Mukherjee S, Kapp EA, Lothian A, Roberts AM, Vasil'ev YV, Boughton BA, et al. Characterization and Identification of Dityrosine Cross-Linked Peptides Using Tandem Mass Spectrometry. Anal Chem. 2017 Jun 6;89(11):6136–45.

61. Gu M, Bode DC, Viles JH. Copper Redox Cycling Inhibits Aβ Fibre Formation and Promotes Fibre Fragmentation, while Generating a Dityrosine Aβ Dimer. Sci Rep. 2018 Nov 1;8(1):16190.

62. Al-Hilaly YK, Williams TL, Stewart-Parker M, Ford L, Skaria E, Cole M, et al. A central role for dityrosine crosslinking of Amyloid-β in Alzheimer's disease [Internet]. Vol. 1, Acta Neuropathologica Communications. 2013. p. 83. Available from: http://dx.doi.org/10.1186/2051-5960-1-83

63. Vázquez G, Caballero AB, Kokinda J, Hijano A, Sabaté R, Gamez P. Copper, dityrosine cross-links and amyloid-β aggregation. J Biol Inorg Chem. 2019 Dec;24(8):1217–29.

64. Penna GL, La Penna G, Li MS. Computational models explain how copper binding to amyloid-β peptide oligomers enhances oxidative pathways [Internet]. Vol. 21, Physical Chemistry Chemical Physics. 2019. p. 8774–84. Available from: http://dx.doi.org/10.1039/c9cp00293f

65. Becke AD. Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A Gen Phys. 1988 Sep 15;38(6):3098–100.

66. Perdew JP. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. Phys Rev B Condens Matter. 1986 Jun 15;33(12):8822–4.

67. Eichkorn K, Weigend F, Treutler O, Ahlrichs R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. Theoretical Chemistry Accounts. 1997;97(1-4):119–24.

68. Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C. Electronic structure calculations on workstation computers: The program system turbomole. Chem Phys Lett. 1989;162(3):165–9.

69. Schäfer A, Huber C, Ahlrichs R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. J Chem Phys. 1994;100(8):5829–35.

70. Klamt A, Schüürmann G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc Perkin Trans 2. 1993;(5):799–805.

71. Bertini L, Bruschi M, Romaniello M, Zampella G, Tiberti M, Barbieri V, et al. Copper coordination to the putative cell binding site of angiogenin: a DFT investigation. In: Highlights in Theoretical Chemistry. 2012. p. 255–69.

72. Prosdocimi T, De Gioia L, Zampella G, Bertini L. On the generation of OH(·) radical species from H2O2 by Cu(I) amyloid beta peptide model complexes: a DFT investigation. J Biol Inorg Chem. 2016 Apr;21(2):197–212.

73. Zirah S, Kozin SA, Mazur AK, Blond A, Cheminant M, Ségalas-Milazzo I, et al. Structural changes of region 1-16 of the Alzheimer disease amyloid beta-peptide upon zinc binding and in vitro aging. J Biol Chem. 2006 Jan 27;281(4):2151–61.

74. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004 Oct;25(13):1605–12.

75. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods. 2017 Jan;14(1):71–3.

76. Price DJ, Brooks CL 3rd. A modified TIP3P water potential for simulation with Ewald summation. J Chem Phys. 2004 Nov 22;121(20):10096–103.

77. Villa AM, Doglia SM, De Gioia L, Bertini L, Natalello A. Anomalous Intrinsic Fluorescence of HCl and NaOH Aqueous Solutions. J Phys Chem Lett. 2019 Nov 21;10(22):7230–6.

78. Armstrong DA, Huie RE, Koppenol WH, Lymar SV, Merényi G, Neta P, et al. Standard Electrode Potentials Involving Radicals in Aqueous Solution: Inorganic Radicals (IUPAC Technical Report) [Internet]. IUPAC Standards Online. Available from: http://dx.doi.org/10.1515/iupac.87.0732

79. Hawkins CL, Davies MJ. Generation and propagation of radical reactions on proteins. Biochim Biophys Acta. 2001 Apr 2;1504(2-3):196–219.

80. Bertini L, Breglia R, Lambrughi M, Fantucci P, De Gioia L, Borsari M, et al. Catalytic Mechanism of Fungal Lytic Polysaccharide Monooxygenases Investigated by First-Principles Calculations. Inorg Chem. 2018 Jan 2;57(1):86–97.

81. Schöneich C. Mechanisms of metal-catalyzed oxidation of histidine to 2-oxo-

histidine in peptides and proteins. J Pharm Biomed Anal. 2000 Jan;21(6):1093–7.

82. Safaei Z, Shiroudi A, Zahedi E, Sillanpää M. Atmospheric oxidation reactions of imidazole initiated by hydroxyl radicals: kinetics and mechanism of reactions and atmospheric implications. Phys Chem Chem Phys. 2019 Apr 17;21(16):8445–56.

83. Raffa DF, Rauk A. Molecular dynamics study of the beta amyloid peptide of Alzheimer's disease and its divalent copper complexes. J Phys Chem B. 2007 Apr 12;111(14):3789–99.

84. Turner M, Mutter ST, Platts JA. Molecular dynamics simulation on the effect of transition metal binding to the N-terminal fragment of amyloid-β [Internet]. Vol. 37, Journal of Biomolecular Structure and Dynamics. 2019. p. 4590–600. Available from: http://dx.doi.org/10.1080/07391102.2018.1555490

85. La Penna G, La Penna G, Li MS. Towards High-Throughput Modelling of Copper Reactivity Induced by Structural Disorder in Amyloid Peptides [Internet]. Vol. 24, Chemistry - A European Journal. 2018. p. 5259–70. Available from: http://dx.doi.org/10.1002/chem.201704654

86. Karthikeyan S, Nagase S. Origins of the Stability of Imidazole–Imidazole, Benzene–Imidazole, and Benzene–Indole Dimers: CCSD(T)/CBS and SAPT Calculations [Internet]. Vol. 116, The Journal of Physical Chemistry A. 2012. p. 1694–700. Available from: http://dx.doi.org/10.1021/jp2109779

87. Gligorovski S, Strekowski R, Barbati S, Vione D. Addition and Correction to Environmental Implications of Hydroxyl Radicals (OH). Chem Rev. 2018 Feb 28;118(4):2296.

Conclusions and perspectives

In this thesis, I tackled three different projects, in which I investigated the structural dynamics of a protein, both physiology and pathology related, by mean of molecular dynamics simulations.
Molecular dynamics has proven a useful tool to investigate conformational changes that would be otherwise impossible to observe in classical *in vitro* experiments and to provide insights and explanations to the experimental data.
Nonetheless, like every computational approach, molecular dynamics suffers from drawbacks that must be considered when interpreting the outcomes of the simulations. Three major classes of issues can be identified:

i. inaccuracies in the calculations due to the empirical force fields used to describe the physical-chemical properties of molecules;
ii. limited computational time to explore the phase space of the system and then the inability to observe the phenomenon of interest;
iii. missing pieces in the model, due to either the constraint of insufficient computational power to study an arbitrary large system or to the lack of experimental data.

The development of ever more accurate and general force fields is a fervent area of research involving several computational physicists and chemists. As far as we know, the force field used in this thesis is at the state of the art in the simulations of macromolecules in solutions. Improving the accuracies of forcefields is not a short-term goal.

In the first project, we simulated the dynamics of the Sec14-PH domain of neurofibromin (NF1), a protein involved in the onset of neurofibromatosis type 1. The Sec14-PH is a lipid-binding module whose exact function is still mysterious. According to the proposed functional model for this protein, a loop in the PH subdomain would rearrange to guarantee the opening of the lipid task in the Sec14 moiety. This event may be regulated but there is no evidence of such regulation so far. The idea was to compare the dynamics of the wild-type protein with that of a pathological mutant (with the mutation localized in the afore mentioned loop), with known crystal structure, and that was shown to have a lower melting temperature in circular dichroism (CD) experiments. By simulating both wild type and the mutant NF1 Sec14-PH at ever increasing temperature we see a different behavior for the mutant form at the higher temperature,

in qualitative agreement with the interpretation of the CD data. Moreover, with our simulations we could get insights into the opening of the lipid pocket showing that it is not coupled with the opening of the loop, who seems rather involved in guaranteeing the structural integrity of the domain. Our calculations could address the intrinsic propensity of the two proteins to unfold under a stress condition (high temperature) but currently they can hardly quantify the relative stability *in vivo*. Further, we obviously cannot directly prove the role of hypothetical interactor in regulating this event.

In the second project, we studied an Electron Transferring Flavoprotein (EtfAB), which can catalyze an electron bifurcation reaction. Based on the crystal structures of this protein from different organisms, two conformations the protein adopts during the process are known, namely B (bifurcating) and D (dehydrogenase). A third conformation has been suggested as a requirement for the bifurcation step of the process since the 18 Å distance between the donor bifurcating flavin and the acceptor flavin as observed in the B-state is considered too long. The B-state has then been renamed B-like and the hypothesized active conformation would be the real B-state. To reach the B-state, a domain of EtfAB, called domain II, should rotate relatively to domain III and we used molecular dynamics to test this hypothesis. Our simulations did not prove the existence of a motion able to shorten the distance between the two flavins from 18 Å to 14 Å within the 3 μs of simulated time because the hinge of the rotation, composed of hydrogen bonds and salts bridges between domain II and domain III, is localized too close to the acceptor flavin to allow it to move toward the bifurcating flavin.
In our simulations, however, we could not take into account the possible effect of binding partners, because of both the absence of structural information about the interaction (ferredoxin-EtfAB) and the size of the system including Bcd that would make the system computationally unaffordable. Anyway, we do not exclude that the motion we were looking for might happen in the absence of binding partners or on a longer time scale, which cannot be treated with classical all-atom molecular dynamics. In the future we will try to overcome this obstacle with the use of a biased sampling approach like, for example, metadynamics.

In the third project in collaboration with prof. Luca Bertini, molecular dynamics simulations were used to investigate the conformational dynamics of the Amyloid-β (1-16) peptide in solution to help elucidate the mechanism of reactive oxygen species (ROS) propagation, which is considered as one of the concurrent causes of tissue damages in Alzheimer's disease. In particular, calculations provided the relative probabilities of collision between the hydroxyl radical ($OH^{\cdot}$) coordinated to the Cu(II) center and a series of sidechains which are considered to initiate the radical propagation.

Among them, His6 has the higher probability to collide with the $OH^{\cdot}$ in line with experimental observations. Furthermore, our simulation provided a mechanism for the reaction between His6 to start.

Our simulations, however, did not provide a full comprehension of what happens *in vitro* because only a single peptide was considered. This rules out any effect due to a possible cross-reaction between Aβ molecules but offer a model for dilute solution of the peptide. Including more copies of Aβ molecules will be a future task, as well as, simulating the Aβ peptide in a more physiological environment like a membrane.

Appendix: Timescales of electron transfer in bifurcating systems and some personal considerations on EtfAB

Basic concepts on Marcus theory

The rate constant for an Electron transfer (ET) reaction

$$D + A \rightarrow D^+ + A^-$$

between a donor D and an acceptor A, is best explained by the Marcus theory of electron transfer, built on the transition state theory [1,2]. The classical Marcus equation is:

$$k_{ET} = \frac{4\pi^2}{h} \frac{1}{\sqrt{4\pi\lambda k_B T}} |H_{DA}|^2 e^{-\frac{(\Delta G^0 + \lambda)^2}{4\lambda k_B T}} \quad (1)$$

where $|H_{DA}|^2$ is the tunneling coupling, a measure of the rate at which the reactants decade into products in a thermally activated DA complex and $(\Delta G^0 + \lambda)^2$ is the activation energy, which depends on both the reaction free energy G^0 and on a specific parameter of the theory, the reorganization energy λ, the energy change associated with the reorganization of the nuclei of the solvent and the complex DA when the reaction takes place.

In the weak-coupling limit, the tunneling coupling is an exponential function of the distance:

$$|H_{DA}|^2 = |H_{DA}^{max}|^2 e^{-\beta r} \quad (2)$$

where β is the tunneling dumping factor and is a property of the medium through which the electrons travel.

The covalent bonds are the faster medium with $\beta = 0.9$ Å^{-1}, whereas the slower propagation is through space with 2.8 Å^{-1} < β < 3.5 Å^{-1}. Within proteins, a value of β = 1.4 Å^{-1} has been measured.

Combining (1) and (2) and switching to logarithm we get:

$$\ln k_{ET} = \ln k_{ET}^{max} - \beta r + constant \tag{3}$$

A first empirical expression for the (3) (in base 10) was given by [3]

$$\log_{10} k_{ET} = 15 - 0.6R - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda} \tag{4}$$

where 0.6 is the transformation into base 10 of β = 1.4 Å^{-1}.
Since the maximum rate, when the reorganization energy λ is equal to $-\Delta G^0$, is at a distance of 3.6 Å (Van der Waals contact) and is 10^{13} s^{-1} [4], which is the vibrational frequency of the nuclei, for exergonic reaction the (4) can be rewritten as [5]

$$\log_{10} k_{ET}^{ex} = 13 - 0.6(R - 3.6) - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda} \tag{5}$$

To account for the atomic density ($0 < \varrho < 1$) within proteins, and so indirectly considering their structure, equation (5) can be modified in [5]:

$$\log_{10} k_{ET}^{ex} = 13 - (1.2 - 0.8\rho)(R - 3.6) - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda} \tag{6}$$

Where with a maximum density $\varrho = 1$, $\beta = 0.9$ Å^{-1} (1.2 – 0.8 = 0.4 Å^{-1} in base 10), and at the minimum density $\varrho = 0$, $\beta = 2.8$ Å^{-1} (1.2 – 0 = 1.2 Å^{-1} in base 10). The average density within protein has been estimated as equal to 0.76 which gives 1.2 – 0.8*0.76 = 0.6 Å^{-1} in base 10. For endergonic reactions, the authors suggest the next correction [5]:

$$\log_{10} k_{ET}^{en} = 13 - (1.2 - 0.8\rho)(R - 3.6) - 3.1\frac{(-\Delta G^0 + \lambda)^2}{\lambda} - 16.7\Delta G^0 \tag{7}$$

Electron transfer in NfnI

In a paper on NfnI, a member of the bifurcating NADH-dependent ferredoxin:NADP reductase (NfnAB) family for which a structure is available, [5,6] a half-life was

measured for the highly unstable semiquinone form of the bifurcating FAD (L-FAD) equal to 10 ps and then estimated a k_{ET} for the electron transfer between L-FAD and its low potential acceptor 4Fe-4S equal to 10^{11} s^{-1}. Considering a conventional value for $\lambda = 0.7$ eV [4], an L-FAD--4Fe4S distance of 5.4 Å (as seen in the crystal structure) and $\beta = 1.1$ $Å^{-1}$ (instead of $\beta = 1.4$ $Å^{-1}$) they solved the equation (5) (with the 0.6 factor transformed in 0.48 for the usage of $\beta = 1.1$ $Å^{-1}$) to get ΔG^0:

$$\log_{10} k_{ET}^{ex} = 13 - 0.6(R - 3.6) - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda}$$

$$\xrightarrow{yields} 11 = 13 - 0.48(5.4 - 3.6) - 3.1\frac{(\Delta G^0 + 0.7)^2}{0.7}$$

From which: $\Delta G^0 = -0.194$ eV

Being E^0 for 4Fe-4S = -718 mV, they got an E^0 for L-FAD ox/sq couple equal to (-718 – 194) = **-912 mV**.

The authors then used the same equation to get k_{ET} for the electron transfer from L-FAD hq to the high potential acceptor 2Fe-2S (E^0 di 80 mV). Being the E^0 of L-FAD hq/sq equal to 359 mV we have ΔG = - (80 – 359) = +280 meV = **0.28 eV.** The distance, as observed the crystal structure, is **13.1 Å.**

$$\log_{10} k_{ET}^{ex} = 13 - 0.48(R - 3.6) - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda} \xrightarrow{yields}$$

$$\log_{10} k_{ET}^{ex} = 13 - 0.48(13.1 - 3.6) - 3.1\frac{(0.28 + 0.7)^2}{0.7} = 4.2$$

$$10^{4.2}\, s^{-1} = 1.6x10^4 s^{-1} \rightarrow \tau_{ET} = \mathbf{62\ \mu s}$$

Note: since the reaction is actually endergonic ($\Delta G^0 = +0.28$ eV) authors should have used the

Electron transfer in EtfAB

I will now apply the same reasoning to EtfAB (equazione 6 and not 7).
ΔG^0 between β-FAD hq (E^0 = +171 mV) and α-FAD sq (E^0 = -136 mV) = - (-136 – 171) = **+0.307 eV.** Let's keep conventional **λ = 0.7 eV** and apply the **18 Å** distance. We get:

$$\log_{10} k_{ET}^{ex} = 13 - 0.48(R - 3.6) - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda}$$

$$\log_{10} k_{ET}^{ex} = 13 - 0.48(18 - 3.6) - 3.1\frac{(0.307 + 0.7)^2}{0.7} = 1.6$$

$k_{ET} = 10^{1.6} s^{-1} = 4x10^{1} s^{-1} \rightarrow \tau_{ET} = \mathbf{25\ ms}$

Is it slow or fast? We don't know but for sure it is a very long time for an electron transfer, three orders of magnitude longer than NfnI. If the distance between β-FAD and α-FAD dropped down to the 14 Å threshold [5], then also kET would drop to $3*10^3$ s^{-1}, maybe more realistic. In our simulations, on the microseconds scale, that motion has never been observed.

I would like to remember that the whole process in EtfAB requires a huge conformational change for passing from the B to the D. We don't know the actual time scale of this movement, but it is expected to be very slow (even slower than one millisecond). In this context, it might not be necessary for the electron transfer between β-FAD and α-FAD to occur faster than a millisecond, without being the kinetic limiting step.

References

1. Marcus RA. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. The Journal of Chemical Physics. 1956. pp. 966–978. doi:10.1063/1.1742723

1. Marcus RA, Sutin N. Electron transfers in chemistry and biology. Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics. 1985. pp. 265–322. doi:10.1016/0304-4173(85)90014-x

2. Moser C, Dutton P. Engineering protein structure for electron transfer function in photosynthetic reaction centers. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 1992. pp. 171–176. doi:10.1016/s0005-2728(05)80012-x

3. Moser CC, Keske JM, Warncke K, Farid RS, Leslie Dutton P. Nature of biological electron transfer. Nature. 1992. pp. 796–802. doi:10.1038/355796a0

4. Page CC, Moser CC, Chen X, Dutton PL. Natural engineering principles of electron tunnelling in biological oxidation-reduction. Nature. 1999;402: 47–52.

5. Duan HD, Diessel Duan H, Lubner CE, Tokmina-Lukaszewska M, Gauss GH, Bothner B, et al. Distinct properties underlie flavin-based electron bifurcation in a novel electron transfer flavoprotein FixAB fromRhodopseudomonas palustris. Journal of Biological Chemistry. 2018. pp. 4688–4701. doi:10.1074/jbc.ra117.000707

Acknowledgments

This thesis would have never been written without the precious help of many people that supported me during these years, both scientifically and personally. In particular, I owe a very special thanks to my tutor and main supervisor Dr. Renata Tisi for assisting (and standing) me in all the projects in this thesis, accepting me as her PhD student and teaching me to do research; Prof. Luca De Gioia for letting me work with his laboratory's machines and be part of his group, even giving me the opportunity to lead a work in the fascinating topic of electron bifurcation; Prof Luca Bertini for choosing me as a collaborator for his project on the Amyloid-beta peptide and for showing me so much mind-blowing quantum chemistry! Last, but of course not least, Prof. Giuseppe Zampella, Dr. Jacopo Vertemara, Dr. Federica Arrigoni and Dr. Alessia D'Aloia, for giving me much help in the lab and being so close to me when I was feeling down. All these people are friends, as well as colleagues. Finally, I cannot forget to mention Prof. Paola Branduardi and Dr. Isabella Mauri for their work in organizing and managing the TeCSBi PhD course.