Biochimie 113 (2015) 54–58

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

# Taming molecular flexibility to tackle rare diseases

Maria Vittoria Cubellis [a, *], Marc Baaden [b, *], Giuseppina Andreotti [c]

[a] Dipartimento di Biologia, Università Federico II, 80126 Napoli, Italy
[b] Laboratoire de Biochimie Théorique, CNRS, UPR9080, Univ Paris Diderot, Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005 Paris, France
[c] Istituto di Chimica Biomolecolare-CNR, 80078 Pozzuoli, Italy

ARTICLE INFO

*Article history:*
Received 14 January 2015
Accepted 20 March 2015
Available online 2 April 2015

*Keywords:*
Molecular dynamics
Disease
Diagnostic use
Protein stability

ABSTRACT

Many mutations responsible of Fabry disease destabilize lysosomal alpha-galactosidase, but retain the enzymatic activity. These mutations are associated to a milder phenotype and are potentially curable with a pharmacological therapy either with chaperones or with drugs that modulate proteostasis. We demonstrate the effectiveness of molecular dynamics simulations to correlate the genotype to the severity of the disease. We studied the relation between protein flexibility and residual enzymatic activity of pathological missense mutants in the cell. We found that mutations occurring at flexible sites are likely to retain activity *in vivo*. The usefulness of molecular dynamics for diagnostic purposes is not limited to lysosomal galactosidase because destabilizing mutations are widely encountered in other proteins, too, and represent a large share of all the ones associated to human diseases.



## 1. Introduction

For most diseases different types of mutations exist and each type requires a specific therapeutic approach. Some mutations lower the free energy difference between the folded and the unfolded protein, shifting the equilibrium towards the latter one. Unstable proteins, although retaining the functional chemical groups needed for the biological activity, are sensitive to proteolysis and are cleared by the protein quality control systems in the cell. Hence, for these mutations, which represent a good share of all the ones associated with human diseases [1], the reduction of the protein concentration in the cell is the primary effect and the reduction of total activity is only a secondary effect. Small chemicals, which are known as pharmacological chaperones, bind preferentially to the folded state, thereby at least partially restoring the equilibrium between folded and unfolded states, and rescue these mutants [2]. They cannot be used for all the genotypes of a given disease, but in general are limited to those which retain residual activity. Nonetheless pharmacological chaperones offer advantages, low cost, oral administration and increased bio-availability. Beside pharmacological chaperones, other small molecules are being evaluated for therapy. They are not specific for a given mutated protein, but alter protein homeostasis [3].

Computational modeling, for instance molecular dynamics simulations, can be used to predict residual activity in the cell. This knowledge is important both for diagnosis and for therapy, because residual activity, the severity of the disease, and responsiveness to small molecule drugs are correlated.

Fabry disease represents a good example to show how conformational flexibility predictions can be used for designing original treatments for rare diseases. Among the many computational techniques that exist to predict protein flexibility such as normal mode analysis and distance geometry approaches, we will focus on molecular dynamics (MD). Fabry disease is X-linked and relatively frequent, 1–9 in 100000 (OMIM: 30150). Different mutations of the gene encoding lysosomal alpha-galactosidase A (AGAL) result in a wide phenotypic spectrum, with respect to age at onset, rate of disease progression, severity of clinical manifestations. Patients with the late onset or atypical form of Fabry disease retain some AGAL activity and are asymptomatic until adult age when they develop cardiac and kidney problems [4].

The treatment of Fabry disease with a pharmacological chaperone 1-deoxy-galactonojirimycin (DGJ) was first proposed by Fan et al., in 1999 [5]. The introduction in clinical practice of galactose to enhance AGAL activity in patients was reported by Frustaci et al., in 2001 [6]. Since then, responsiveness to pharmacological chaperones has been assessed for a huge number of AGAL mutations,

*Abbreviations:* AGAL, lysosomal alpha-galactosidase; DGJ, 1-deoxy-galactonojirimycin; RMSD, Root mean square deviation; RMSF, root mean square fluctuation.

* Corresponding authors.
*E-mail addresses:* cubellis@unina.it (M.V. Cubellis), baaden@smplinux.de (M. Baaden).

http://dx.doi.org/10.1016/j.biochi.2015.03.018

covering both early and late onset forms of Fabry disease (for a review please consult Fabry_CEP [7] and references therein). A relatively large proportion of mutants, in particular among mutations associated with the late onset form of Fabry disease, recover activity when treated with DGJ. In a few cases it was possible to prove that DGJ acts by enhancing thermodynamic stability of the mutants [8,9].

In this paper we correlate the flexibility of the sites where AGAL mutations occur with the residual activity in the cells. This result is useful for the evaluation of severity and the choice of a personalized therapy. The direct measure of residual activity in the cells for each case would be impractical because more than 520 missense/nonsense mutations have been described in the databank HGMD® professional [10] for Fabry disease and most of them are private, that means that they are seen in a single family.

## 2. Materials and methods

### 2.1. Molecular dynamics simulations

We used the structure of AGAL solved in the presence (3GXT) or in the absence (3GXN) of DGJ at pH 4.5 as input. We run a 50 ns MD simulation with the amber03 force field (a variant of the AMBER-99 one [11]) at the same pH at which crystals were grown. We used the Yasara program under default conditions combined with fully automatic optimized assignment of topology and parameters for the ligand using the AutoSMILES procedure [12,13]. All systems were solvated with explicit TIP3P water molecules and Na+ and Cl- counterions were added as background salt and to preserve overall electrical neutrality. Each system was energy minimized by using the steepest descent method to relax any steric conflicts before beginning the simulations. Simulations were carried out with periodic boundary conditions. Long-range electrostatic interactions were calculated by using PME with a direct-space cut-off of 7.86 Å. All simulations were performed by using an NVT ensemble at 298 K. A 2 fs/1 fs double-integration time step was used. Root mean square deviation (RMSD) between structures following least-squares fitting to the reference energy-minimized input structure, was calculated with Gromacs [14]. The RMSD value increases steadily in the first 25 ns. After the stabilization of the system, Root mean square fluctuations (RMSF) of alpha carbons of each residue were calculated with Gromacs [14].

### 2.2. Miscellaneous

Graph plotting was carried out with Kaleidagraph (Synergy Software, PA).

We assigned secondary structure with SEGNO [15]. Active site residues were identified with DrosteP [16]. The figure showing the AGAL structure colored by RMSF was produced with CHIMERA [17].

## 3. Results and discussion

### 3.1. Correlation between alpha-galactosidase flexibility and residual activity of missense mutants in the cell

An attempt to systematically classify the phenotype of Fabry disease based on structural features of AGAL was made by Saito [18]. In this previous work, mutations were divided into two categories defined as classic or atypical based on clinical signs. We believe that this classification is too simplistic because for most mutations few affected people are known and symptoms are variable even among patients with the same mutation [19]. In general, residual enzyme activity is associated with a less severe phenotype [20,21]. We decided to use a different approach and move from binary categories to real values, taking advantage of the fact that enzymatic activity of mutants has been measured in cells derived from patients or in cells transiently transfected with plasmid carrying a specific AGAL mutation.

We collected data from several laboratories [22–28] and examined 244 mutations altogether. The effect of mutations depends critically on the positions of the affected residue in the protein structure and secondarily on the type of substitution, conservative ones being more tolerated. In a few cases, lack of activity can be explained straightforwardly because the mutation affects the active site or disulphide bridges. We define active site residues as those that line the most conserved pocket in the protein structure: by this definition the active site residues in AGAL are W47, D92, D93, Y134, L168, D170, Y207, R227, E203, L206, D231, S297. When these amino acids are mutated, activity is null or proximal to zero. Disulphide bridges are essential to maintain the 3D-structure of the protein and it is not surprising that when a cysteine involved in a bridge (C52, C94; C56, C63; C142, C172; C202, C223; C378, C383) is mutated, fold, and consequently activity, is lost. It is worth observing that not all the cysteine residues in AGAL form disulphide bridges, the exception being C90 and C174.

When the active site is not severely compromised, the activity in the cells depends on two factors, the specific activity (U/mg) and the amount of the mature protein, which in turn depends on stability. Only in a very few cases it was possible to purify mutants and measure specific activity or stability independently, but in these cases it was shown that most missense mutations observed in Fabry patients do not affect the maximal velocity ($k_{cat}$) or the affinity for substrate ($K_M$) [29]. In order to analyze all mutations and try to correlate their effect in the cell with the position in the protein structure, we calculated an average residual activity per site. For example, we assign a value of 10.75% to Q327 because Q327K and Q327E have 0 or 21.5% residual activity, respectively. We divided the data in two sets. The first one includes 66 sites with zero average residual activity, the second one 92 sites with non null average residual activity. This distinction is motivated by the observation that the residual activity and stability in the cell correlates only for mutants that are not completely inactive.

We carried out MD simulations using the structure of AGAL solved in the absence (3GXN) or in the presence (3GXT) of DGJ at pH 4.5 [30] as input. We run a 50 ns simulation at the same pH at which crystals were grown and calculated root mean square fluctuation (RMSF) values as a measure of flexibility per residue. In Fig. 1, the structure of AGAL solved in the presence of DGJ (3GXT) was colored by RMSF with colors ranging from blue (low flexibility) to white (medium flexibility) to red (high flexibility). AGAL is a homo-dimer and each subunit is made up by two domains, a TIM barrel where the active site is located and an antiparallel beta domain. Inspection of Fig. 1 suggests that flexibility is minimal in the regions not exposed to solvent buried between subunits, between domains or inside the TIM barrel.

Indeed accessibility (percent of residue surface exposed to solvent) and flexibility follow the same trend but direct correlation between the two properties is relatively low, approximately 70%.

In Fig. 2 we show RMSF for the molecular dynamics run reported on the structure without the drug (blue line) or with the drug (black line).

There is little difference between the two. Notably the region where the highest difference is observed, spans aa 173–177. This observation is interesting because Asp 170 makes a salt bridge with the heterocyclic nitrogen in the drug molecule [30].

The beta strand of the TIM barrel at the tip of which are located the active site residues, forms the rigid core of the protein. The active site residues are rigid both in 3GXN and 3GXT. Circles represent sites where mutations are associated with null residual activity (first set).