Protein Engineering, Design & Selection vol. 24 no. 10 pp. 773–775, 2011
Published online July 1, 2011 doi:10.1093/protein/gzr032

Comparison of a molecular dynamics model with the X-ray structure of the N370S acid-β-glucosidase mutant that causes Gaucher disease

Marc N.Offman[1,6], Marcin Krol[2], Burkhard Rost[1], Israel Silman[3], Joel L.Sussman,[4] and Anthony H.Futerman[5]

[1]Institut für Informatik/I12, Technische Universität München, Boltzmannstr 3, 85748 Garching b. München, Germany, [2]Department of Bioinformatics and Telemedicine, Jagiellonian University, Collegium Medicum, Lazarza 16, 31-531 Krakow, Poland, [3]Departments of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel, [4]Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel and [5]Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel

[6]To whom correspondence should be addressed.
E-mail: offman@rostlab.org

Received March 12, 2011; revised May 25, 2011; accepted May 31, 2011

Edited by Rebecca Wade

Recently, two studies were published that examined the structure of the acid-β-glucosidase N370S mutant, the most common mutant that causes Gaucher disease. One study used the experimental tool of X-ray crystallography, and the other utilized molecular dynamics (MD). The two studies reinforced each other through the similarities in their findings, but each approach also added some unique information. Both studies report that the conformation of active site loop 3 changes, due to an altered hydrogen bonding network; however, the MD study produced additional data concerning the flexibility of loop 1 and the catalytic residues that are not observed in the other study.

Keywords: GCase/Gaucher disease/molecular dynamics/N370S

Proteins are intrinsically flexible (Stryer, 1981; Fenimore *et al.*, 2002), existing in groups of similar but distinct conformations (energy basins) (Henzler-Wildman and Kern, 2007). Protein function often results from flexibility (Fenimore *et al.*, 2002). Flexibility can depend on pH, on ligand/substrate binding, and be affected by interaction with other environmental factors such as membranes. Protein flexibility can be determined experimentally (Englander *et al.*, 2003), predicted from sequence (Schlessinger and Rost, 2005; Gu *et al.*, 2006), and simulated using molecular dynamics (MD), which can be applied both to experimental structures and to models derived from such structures (Karplus, 2002). Although MD was introduced in the 1980s (Brooks *et al.*, 1983), its success remained relatively limited partly due to inadequate computational resources. The computational power available in recent years has resulted in rapid development of the technique and in the generation of data of increasing functional relevance (Puklin-Faucher *et al.*, 2006; Nury *et al.*, 2010; Shaw *et al.*, 2010; Zhu and Hummer, 2010; Offman *et al.*, 2011). This new potential has made it even more important to understand the pitfalls of the tool that is at its best when applied to predicting flexibility from experimental high-resolution structures. Used in such a way, MD simulations provide an effective, reliable, affordable and fast approach to investigate protein flexibility.

Mutations in the enzyme acid-β-glucosidase (GCase; glucocerebrosidase) cause Gaucher disease (GD), the most common lysosomal storage disorder. GCase seems highly susceptible to sequence changes: over 200 mutants have been described (Hruska *et al.*, 2008). Eight years ago, researchers solved the first X-ray structure of GCase (Dvir *et al.*, 2003). Since then, several crystal structures have been obtained under different experimental conditions, of both the apo-enzyme and of complexes with various ligands (Premkumar *et al.*, 2005; Brumshtein *et al.*, 2006; Brumshtein *et al.*, 2007; Lieberman *et al.*, 2007; Kornhaber *et al.*, 2008; Brumshtein *et al.*, 2009; Brumshtein *et al.*, 2010). Briefly, the structure consists of a characteristic $(\alpha/\beta)_8$ triosephosphate isomerase barrel containing the catalytic residues, designated as domain III, and two smaller domains, I and II, which are composed mainly of β-sheets. Although their biological role is unknown, also mutations in both domains I and II can cause GD. Three loops, which control access to the active site, have been shown to adopt alternative conformations (Premkumar *et al.*, 2005).

No structural information concerning any of the mutations causing GD, including the most common mutant, N370S, was available until recently, when two independent studies appeared simultaneously. One used X-ray crystallography to determine the 3D crystal structure experimentally (Wei *et al.*, 2011), whereas the other used MD to predict what structural changes might occur in GCase as a consequence of the N370S mutation (Offman *et al.*, 2010), based on the original wild type (WT) crystal structure (Dvir *et al.*, 2003). Both studies analyzed the structure at physiological and lysosomal pH, but the MD study also examined both the structure of the apo-enzyme and that of a complex with *N*-butyl-deoxynojirimycin (NB-DNJ), a putative chemical chaperone (Sawkar *et al.*, 2006). The GCase crystal structure of the mutant N370S was resolved at a resolution of 2.7 Å and with a R-value of 0.175 (protein data bank – entry 3KE0); the GCase starting structure for the MD simulation of the mutant N370S was generated from the WT crystal structure (2V3D: 1.96 Å res, R 0.157) (Brumshtein *et al.*, 2007), changing the mutated amino acid side chains with the software SCWRL3 (Canutescu *et al.*, 2003); missing residues were modeled

Table I. Comparison of the crystal structure and of structures extracted from several MD trajectories of the N370S GCase mutant

	X-ray study[a]	MD simulations[b]
Agreement	In WT, two conformations of loop 3[c]; extended or α-turn. In N370S, only extended conformation H-bond W312-N370	In WT, two conformations of loop 3, both extended and α-turn. In N370S, loop 3 changes to extended conformation during simulations H-bond or π/H bond observed for W312-N370; Once NB-DNJ bound, H-bond is observed less often
	N370S more rigid Reduced, suboptimal distance between catalytic residues E235 and E340 can be measured but was not noted by the authors	N370S more rigid, especially at pH 7.4 Reduced, suboptimal distance between catalytic residues E235 and E340
	Overall, N370S at pH 7.4 is similar to WT at pH 4.5	Loop behavior of N370S at pH 7.4 is similar to that of WT at pH 4.5
Partial agreement	For N370S at acidic pH, H-bond W312-S366; at neutral pH, H-bond between W312-S370 WT is sensitive to pH change N370S not sensitive to pH change	For N370S at acidic and neutral pH, H-bond seen for W312-S366, but not for W312-N370 Both WT and N370S are pH-sensitive N370S is less affected, with small differences in H-bond occupancies of the active site and in flexible behavior of loops 1/3 between pH 7.4 and pH 4.5
	Loop 3 less flexible at neutral and acidic pH for N370S Increased stability of N370S at pH 4.5	Loop 3 less flexible at pH 7.4 for N370S, but not at acidic pH N370S is more stable at both pHs
No agreement	Loop 1 identical between WT and N370S In WT/NB-DNJ complex, H-bond between Y313-E235 For N370S H-bond between Y313-E235	Some destabilization of loop 1 observed for WT at pH 7.4, not for N370S In WT/NB-DNJ complex, H-bond for Y313-E340 formed more often In N370S, Y313-E235 H-bond hardly observed

[a]Wei *et al.* (2011)
[b] In the study of Offman *et al.* (2010), the findings reported were observed in a statistically significant number of structures in the MD trajectories.
[c]In the study of Wei *et al.* loop 1 is referred to as loop 3, and loop 3 is referred to as loop 1. In the table, as throughout the text, the original designation of the loops is used (Brumshtein *et al.*, 2007).

from another GCase WT crystal structure (1OGS: 2.0 Å res, R 0.195) (Dvir *et al.*, 2003). Both WT and mutant structures were simulated for 10 ns at 350 K, at pH 7.4 and 4.5, in five parallel and independent runs.

Table I shows that overall the two studies yielded rather similar structural information. The compared information was extracted from each study, where B-factors were made use of to study protein flexibility in the crystal structure, and root mean square fluctuations (RMSF) in the MD analysis. Furthermore, in both studies the mutant structures were compared with other known WT structures and visually inspected to gain further understanding of the impact of the mutations on protein structure/flexibility. For the MD analysis these changes were subsequently analyzed computationally for the WT and N370S mutant at both pH 7.4 and pH 4.5; using the trajectories of each of the five MD simulation runs, a total of 20 000 structures were analyzed (a further 30 000 structures were obtained for complexes of both WT GCase and the N370S mutant with NB-DNJ).

The most important observation agreed upon by both methods is that one of the loops adjacent to the active site, loop 3 (residues 312–319), exists only in an extended conformation in N370S, whereas in WT GCase both an extended and a helical conformation are found. This is a consequence of a change in H-bond network in and around the active site, since an H-bond between N370 and W312 present in the WT structure is absent in the N370S mutant. Both the crystal structure and the MD simulations of N370S suggest that both W312, which can interact with either S366 or N370 in the WT structure, and Y313, which can interact with E235, E340 and C342, contribute crucially to the stability and integrity of the active site. Both these residues display altered conformations and different side-chain orientations in N370S, thus causing a shift in loop 3. Both studies report an overall increased stability of N370S, and some structural similarities between the WT structure at pH 4.5 and the N370S structure at pH 7.4. In hindsight, it seems remarkable that it was possible to extract such a subtle feature from the MD simulations. With regard to the degrees of stabilization and flexibility of the three active site loops, the experimental and theoretical methods are in disagreement. Molecular dynamics predicts some differences between loop 1 (residues 341–350) in the WT structure and in N370S that are not seen in their crystal structures. The MD-derived prediction that loop 1 presents with an above-average flexibility (RMSF: ∼1.5 Å for loop 1; ∼1.05 Å for loop 3; ∼ 0.74 Å for the whole protein) is supported by a recent study on the dynamic behavior of GCase using hydrogen/deuterium exchange mass spectrometry (Bobst *et al.*, 2010).

For mutations that affect enzyme activity, the key issue is to understand how changes in flexibility and H-bond networks affect the active site. The MD study reported a decreased optimal distance between the two catalytic residues, E235 and E340, from ∼5.5–5.2Å to 4.5Å for N370S compared with WT GCase. A reduced distance of 4.8Å was observed in the N370S crystal structure at acidic pH, though not as pronounced as seen in the MD simulations, which are represented by a randomly chosen snapshot (e.g., in this case ∼5 ns) from one of the five MD trajectories (Fig. 1). Since this change is fairly subtle, it was not referred to in the experimental study (Wei *et al.*, 2011). Thus, MD can highlight crystallographic data that might otherwise be overlooked. Once this random snapshot is overlaid on the native and N370S crystal structure, the shifted and altered conformation of loop 3 can be seen in both the MD and experimental N370S structures (Fig. 1)—a feature that was shown to be statistically significant in the MD study. The similarity of loop 1 in the N370S MD and crystal structures is larger than for the corresponding WT structures, which can be explained by the significantly increased flexibility of loop 1. Furthermore, when comparing structural similarities between the snapshots for the WT and N370S loop 1 and the

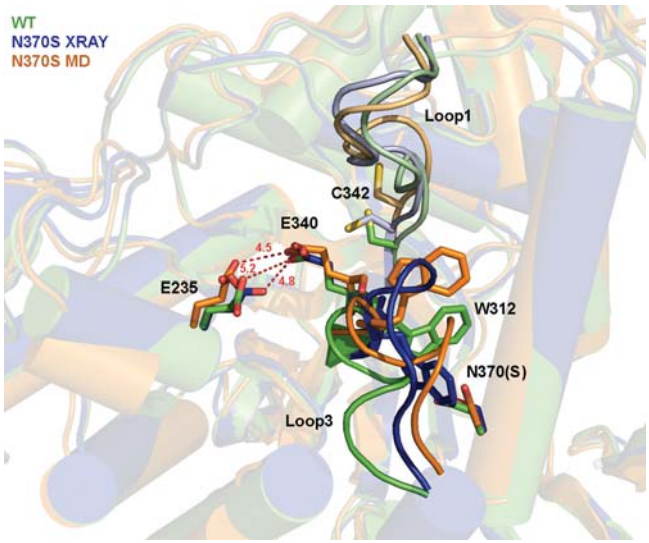

Fig. 1. Overlay of the WT GCase crystal structure, the N370S crystal structure and the N370S MD structure at pH 4.5. Displayed are close-up views of the active-site region for the WT crystal structure (green), the N370S crystal structure (blue) and the N370S MD structure (orange). Upper loop 1 is shown in light colors, and lower loop 3 in darker colors. The side chains of active-site residues (E235, E340 and C342), the important residue W312, and N370(S) are represented as stick models. The distances between E235 and E340 are shown as red dashed lines, with the distances indicated (angstroms).

respective crystal structures, using the TM-score (Zhang and Skolnick, 2004), the following observations can be made: both WT and N370S MD simulations start with the WT conformation for loop 1; of the five independent simulations for the WT, 45% of the snapshots are more similar to the WT configuration found in the crystal structure, and 25% more similar to the N370S mutant crystal structure. For the MD simulations of N370S, in 40% of the snapshots the conformation of loop 1 remains similar to that found in the WT crystal structure (starting structure); 45%, however, are more similar to the loop 1 conformation found in the N370S mutant crystal structure.

In summary, the comparison of the crystallographic and MD studies on GCase has shown that MD has the capacity to predict and enhance experimental results, and to extract features and characteristics that were not obvious in the crystal structures. Furthermore, MD allows rapid screening of the structural characteristics of many mutants, and could, in principle, be applied to all the known GD mutants, whereas their analysis by X-ray crystallography requires their expression, purification and crystallization. Similarly, MD allows investigation of the binding of both known and predicted pharmacological chaperones to both WT GCase and to its mutants.

Funding

This work was supported by the Alexander von Humboldt-Foundation. J.L.S. is the Morton and Gladys Pickman Professor of Structural Biology. A.H.F. is the Joseph Meyerhoff Professor of Biochemistry at the Weizmann Institute of Science.

References

Bobst,C.E., Thomas,J.J., Salinas,P.A., Savickas,P. and Kaltashov,I.A. (2010) *Protein Sci.*, **19**, 2366–2378.
Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) *J. Comput. Chem.*, **4**, 187–217.
Brumshtein,B., Aguilar-Moncayo,M., Garcia-Moreno,M.I. *et al.* (2009) *Chembiochem*, **10**, 1480–1485.
Brumshtein,B., Greenblatt,H.M., Butters,T.D., Shaaltiel,Y., Aviezer,D., Silman,I., Futerman,A.H. and Sussman,J.L. (2007) *J. Biol. Chem.*, **282**, 29052–29058.
Brumshtein,B., Salinas,P., Peterson,B., Chan,V., Silman,I., Sussman,J.L., Savickas,P.J., Robinson,G.S. and Futerman,A.H. (2010) *Glycobiology*, **20**, 24–32.
Brumshtein,B., Wormald,M.R., Silman,I., Futerman,A.H. and Sussman,J.L. (2006) *Acta. Crystallogr. D Biol. Crystallogr.*, **62**, 1458–1465.
Canutescu,A.A., Shelenkov,A.A. and Dunbrack,R.L. Jr. (2003) *Protein Sci.*, **12**, 2001–2014.
Dvir,H., Harel,M., McCarthy,A.A., Toker,L., Silman,I., Futerman,A.H. and Sussman,J.L. (2003) *EMBO Rep.*, **4**, 704–709.
Englander,J.J., Del Mar,C., Li,W., Englander,S.W., Kim,J.S., Stranz,D.D., Hamuro,Y. and Woods,V.L. Jr. (2003) *Proc. Natl Acad. Sci. USA*, **100**, 7057–7062.
Fenimore,P.W., Frauenfelder,H., McMahon,B.H. and Parak,F.G. (2002) *Proc. Natl Acad. Sci. USA*, **99**, 16047–16051.
Gu,J., Gribskov,M. and Bourne,P.E. (2006) *PLoS Comput. Biol.*, **2**, e90.
Henzler-Wildman,K. and Kern,D. (2007) *Nature*, **450**, 964–972.
Hruska,K.S., LaMarca,M.E., Scott,C.R. and Sidransky,E. (2008) *Hum. Mutat.*, **29**, 567–583.
Karplus,M. (2002) *Acc. Chem. Res.*, **35**, 321–323.
Kornhaber,G.J., Tropak,M.B., Maegawa,G.H., Tuske,S.J., Coales,S.J., Mahuran,D.J. and Hamuro,Y. (2008) *Chembiochem*, **9**, 2643–2649.
Lieberman,R.L., Wustman,B.A., Huertas,P., Powe,A.C. Jr, Pine,C.W., Khanna,R., Schlossmacher,M.G., Ringe,D. and Petsko,G.A. (2007) *Nat. Chem. Biol.*, **3**, 101–107.
Nury,H., Poitevin,F., Van Renterghem,C., Changeux,J.P., Corringer,P.J., Delarue,M. and Baaden,M. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 6275–6280.
Offman,M.N., Krol,M., Patel,N., Krishnan,S., Liu,J., Saha,V. and Bates,P.A. (2011) *Blood*, **117**, 1614–1621.
Offman,M.N., Krol,M., Silman,I., Sussman,J.L. and Futerman,A.H. (2010) *J. Biol. Chem.*, **285**, 42105–42114.
Premkumar,L., Sawkar,A.R., Boldin-Adamsky,S., Toker,L., Silman,I., Kelly,J.W., Futerman,A.H. and Sussman,J.L. (2005) *J. Biol. Chem.*, **280**, 23815–23819.
Puklin-Faucher,E., Gao,M., Schulten,K. and Vogel,V. (2006) *J. Cell. Biol.*, **175**, 349–360.
Sawkar,A.R., D'Haeze,W. and Kelly,J.W. (2006) *Cell Mol. Life Sci.*, **63**, 1179–1192.
Schlessinger,A. and Rost,B. (2005) *Proteins*, **61**, 115–126.
Shaw,D.E., Maragakis,P., Lindorff-Larsen,K. *et al.* (2010) *Science*, **330**, 341–346.
Stryer,L. (1981) *Biochem. Soc. Symp.*, **46**, 39–55.
Wei,R.R., Hughes,H., Boucher,S., Bird,J.J., Guziewicz,N., Van Patten,S.M., Qiu,H., Pan,C.Q. and Edmunds,T. (2011) *J. Biol. Chem.*, **286**, 299–308.
Zhang,Y. and Skolnick,J. (2004) *Proteins*, **57**, 702–710.
Zhu,F. and Hummer,G. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 19814–19819.