

# SCIENTIFIC REPORTS

OPEN

## Changes in the folding landscape of the WW domain provide a molecular mechanism for an inherited genetic syndrome

Received: 11 April 2016

Accepted: 01 July 2016

Published: 26 July 2016

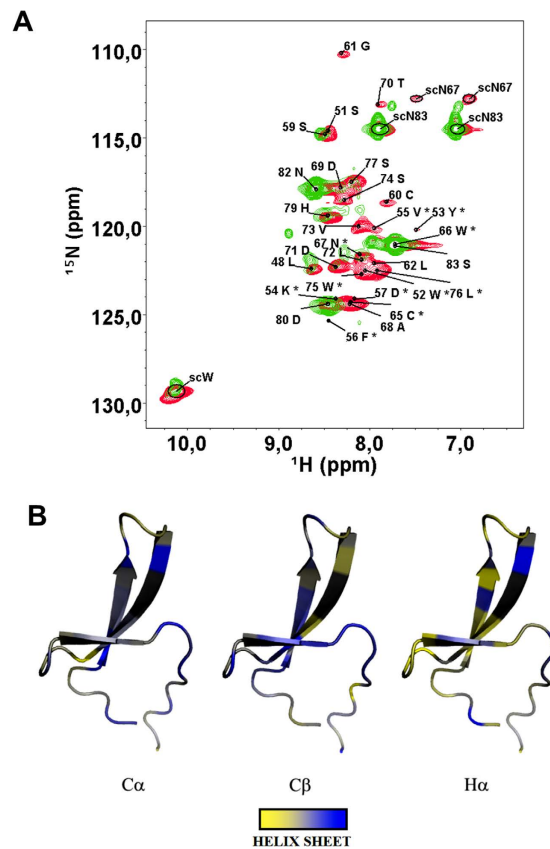
Encarna Pucheta-Martinez<sup>1,\*</sup>, Nicola D'Amelio<sup>2,\*</sup>, Moreno Lelli<sup>3</sup>, Jorge L. Martinez-Torrecuadrada<sup>4</sup>, Marius Sudol<sup>5,6,7</sup>, Giorgio Saladino<sup>1,2</sup> & Francesco Luigi Gervasio<sup>1,2</sup>

WW domains are small domains present in many human proteins with a wide array of functions and acting through the recognition of proline-rich sequences. The WW domain belonging to polyglutamine tract-binding protein 1 (PQBP1) is of particular interest due to its direct involvement in several X chromosome-linked intellectual disabilities, including Golabi-Ito-Hall (GIH) syndrome, where a single point mutation (Y65C) correlates with the development of the disease. The mutant cannot bind to its natural ligand WBP11, which regulates mRNA processing. In this work we use high-field high-resolution NMR and enhanced sampling molecular dynamics simulations to gain insight into the molecular causes the disease. We find that the wild type protein is partially unfolded exchanging among multiple beta-strand-like conformations in solution. The Y65C mutation further destabilizes the residual fold and primes the protein for the formation of a disulphide bridge, which could be at the origin of the loss of function.

The WW is the smallest domain found in proteins (about 40 amino-acids) with a variety of functions within the cell<sup>1–3</sup>. Just as the SH3 domain, it recognizes proline-rich regions in cognate proteins<sup>4</sup>. Its diverse and regulated localization within the cell (both in the nucleus and the cytoplasm) stresses the biological importance of WW domain-containing proteins and explains why signalling via WW domain complexes is implicated in several human diseases including muscular dystrophy, Alzheimer and Huntington diseases, Liddle's syndrome of hypertension, cancer and X chromosome linked intellectual disabilities<sup>2,5–10</sup>. The Golabi-Ito-Hall (GIH) syndrome, in particular, is an X-chromosome linked disease caused by a missense mutation in the WW domain of the Polyglutamine Binding Protein 1 (PQBP1), which is widely expressed in various organs but enriched in the brain. The WW domain of PQBP1 mediates the interaction with the nucleocytoplasmic shuttling splicing factor SIPP1 (previously known as NpWBP and WBP11), which regulates mRNA processing and transcription<sup>11</sup>, by recognizing the proline-rich sequence of SIPP1<sup>12,13</sup>. Mutations of PQBP-1 have also been reported in several other X-chromosome-linked intellectual disability disorders (XLID) and progressive neuro-degenerative diseases<sup>6,8,14,15</sup>. Possible molecular causes linking WW mutations to the GIH syndrome have been investigated by Sudol and coworkers<sup>11</sup>. In their study, the authors observed a moderate loss of signaling in the GIH-causing Y65C mutant and suggested that the fold of the WW domain might be compromised by the mutation, with consequent loss of interaction with its partners in the splicing complex.

The fold of WW domains is in general well known, consisting of a stable, triple stranded beta sheet<sup>16–34</sup>. The solution NMR structures of several WW domains have been determined revealing a common fold but also

<sup>1</sup>Department of Chemistry, University College London, London WC1E 6BT, United Kingdom. <sup>2</sup>Research Institute of Structural and Molecular Biology, University College London, London WC1E 6BT, United Kingdom. <sup>3</sup>Institute of Florence, Department of Chemistry, Magnetic Resonance Center (CERM), 50019 Sesto Fiorentino (FI), Italy. <sup>4</sup>Crystallography and Protein Engineering Unit, Spanish National Cancer Research Centre (CNIO), C/Melchor Fernandez Almagro 3, 28029, Madrid, Spain. <sup>5</sup>Institute of Molecular and Cell Biology A\*STAR, 61 Biopolis, Singapore 138673, Republic of Singapore. <sup>6</sup>Mechanobiology Institute, 5A Engineering Drive 1, Singapore 117411, Republic of Singapore. <sup>7</sup>National University of Singapore, Department of Physiology, The Yong Loo Li School of Medicine, 2 Medical Drive, Singapore 117597, Republic of Singapore. \*These authors contributed equally to the work. Correspondence and requests for materials should be addressed to G.S. (email: [g.saladino@ucl.ac.uk](mailto:g.saladino@ucl.ac.uk)) or F.L.G. (email: [f.l.gervasio@ucl.ac.uk](mailto:f.l.gervasio@ucl.ac.uk))



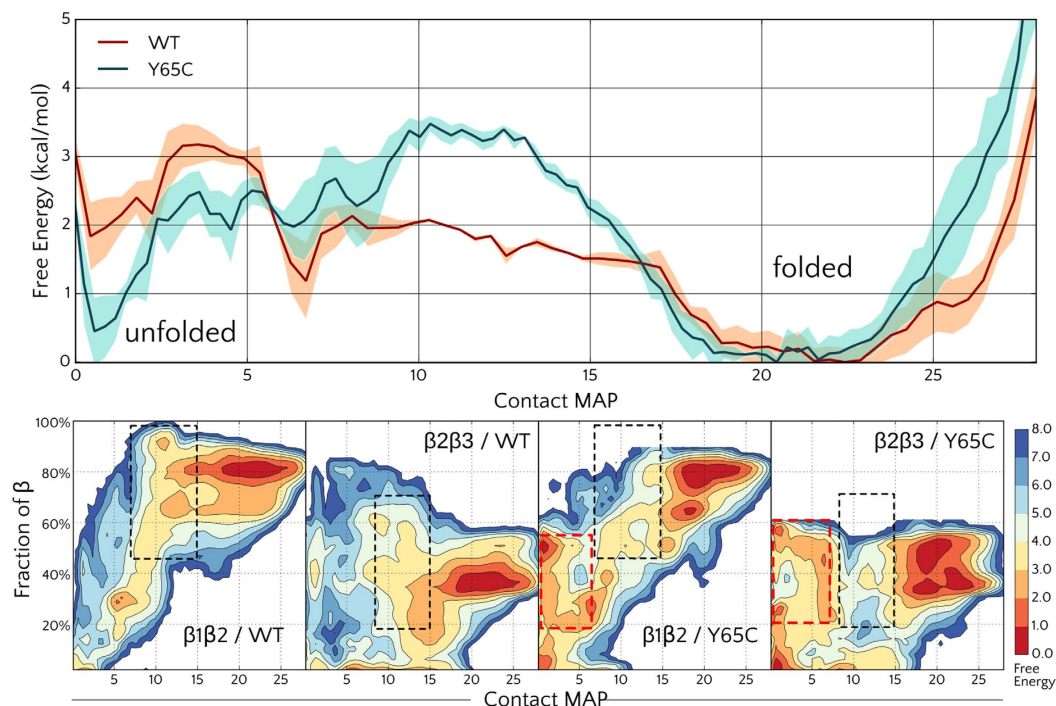
**Figure 1.** (A)  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectrum of WT WW domain (green) superimposed to the one of its Y65C mutant (red), for which also the assignment is shown (500 MHz of  $^1\text{H}$  Larmor frequency,  $T = 278$  K). Residues labeled with a star (\*) are tentatively assigned. (B) Pictorial representation of raw values of Chemical Shift Index onto the structural model of Y65C WW domain from PQBP. Beta strand-like structures (blue) are predominant. Deviations of opposite sign are displayed in yellow. Residues for which data are not available (unassigned) are colored in gray.

different degrees of conformational stability. While in general the domain is remarkably well ordered<sup>16–34</sup>, in some cases it presents conformational exchange<sup>19,21,32</sup>. The structure has been also studied in the presence of a binding peptide which might stabilize the fold<sup>17,19,28,30–32</sup>. Recently, the X-ray structure of the C-terminus of PQBP1 has been determined in complex with spliceosomal protein U5-15kD<sup>35</sup>, showing how a YxxPxxVL motif in PQBP1 is recognized. The WW domain, however, was not included in the protein sequence. Here we investigate the underlying causes of the GIH disease by using a combination of high-field solution NMR and state-of-the-art enhanced sampling simulations to determine the effect of the Y65C mutation on the structure and dynamics of the WW domain of PQBP1.

## Results and Discussion

### The WW domain, from Poly-glutamine binding protein (PQBP), exchanges among different conformations in solution.

The  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum of the PQBP1 WW domain (Fig. 1A), reveals that the protein exchanges among multiple conformations in solution, a behavior which has been reported for a few other WW domains<sup>19,21,32</sup>. The dispersion of signals in the proton dimension approaches the one expected for intrinsically disordered proteins. However, the large line-width is not expected for a random coil behavior. Despite the small size of the protein and the use of a doubly-labeled sample, the assignment of signals was particularly demanding (see Table S1 in Supporting Information). Excluding formation of large aggregates as a possible cause (the protein is purified by size exclusion chromatography, yielding one single peak consistent with a monomer) such extreme broadening could be explained by the presence of conformational exchange in the micro to milliseconds time scales. Severe line broadening is indicating the presence of dynamic exchange approaching the intermediate regime in the NMR time scale. This is observed when the difference in the frequency shifts of the exchanging resonances is comparable with the exchange rate. Thus changing pH, temperature (that affect the exchange rates) or changing the magnetic field (that scales the difference in the frequency shifts) could help in displacing the intermediate exchange regime towards a slow or a fast exchange regime that has narrower NMR line-width. We tried to modify the experimental conditions in order to get sharper lines for a detailed structural determination based on NOE analysis. Unfortunately the quality of the spectrum does not improve when the pH is lowered from 7.4 to 6.5 or by changing NMR field (500, 600, 700 and even 1000 MHz). Also the temperature has very little effect. The spectrum of the WT slightly improves at 313 K but the line-width remains so broad that



**Figure 2.** (Top) Free Energy profile of the WW domain WT and of the Y65C mutant as a function of the native contact map (CMAP); the shaded region represents the error. (Bottom) Projection of the WT and Y65C folding free energy profile along the contact map and the fraction of  $\beta$  structure of the first ( $\beta1\beta2$ ) and second ( $\beta2\beta3$ ) hairpins. The  $\beta1\beta2$  element appears to be the most stable and is well structured also in partially folded conformations (black dotted boxes, CMAP between 10 and 15). The mutant partially folded conformations however, appear less structured, and misfolded conformations (with low CMAP and some residual  $\beta$  structure) appear (red boxes).

a complete assignment is challenging. This is probably due to the low stability of the native fold combined with a broad energy basin, which might lead to a conformational exchange between various structures.

NMR allows the detection of secondary structure elements residue by residue even in the presence of conformational exchange. This is achieved by measuring the chemical shift of backbone carbon atoms (carbonyl, carbons in positions  $\alpha$  and  $\beta$  but also  $H\alpha$ ), whose deviations from the tabulated random coil values, is predictive of alpha-helices or beta-strands character<sup>36</sup>. The accuracy of the Chemical Shift Index (CSI)<sup>36</sup> has been recently improved to 90%<sup>37</sup>. In cases of fast exchange, the value of the deviation is expected to be a population weighted average of the deviations expected for each structure. This decreases the absolute value of the CSI but it can still give an indication of the most populated secondary structure conformations in solution. In the present case, the calculated deviations are compatible with a beta fold for both the wild type (WT) and the mutant (Fig. 1B). Among the three predictors ( $C\alpha$ ,  $C\beta$  and  $H\alpha$ ),  $C\beta$  is the most reliable in detecting beta-strands. According to its values, beta structure is present where expected from the canonical WW fold, however it is not very stable and undergoes a significant conformational exchange dynamics. Similar conclusions have been reached by circular dichroism<sup>38</sup>. The presence of beta structures in the folded ensemble of the WW is expected, as the typical fold for this domain consists in a triple stranded beta sheet both in crystal structures and in solution<sup>16–34</sup>. A NMR-based model of our specific amino-acid sequence also yields a similar fold (Fig. 1B). The hydrophobic core consisting of two tryptophan residues (giving the name to the domain) is well packed with other conserved hydrophobic residues (Y65, N67, P78)<sup>11</sup>.

In order to explore the folding mechanism and better understand the reasons for the intrinsic instability of the fold, we performed enhanced sampling simulations with Parallel Tempering Metadynamics (PTMetaD)<sup>39</sup> of both the wild-type and the Y65C mutant. Since the timescales needed to observe folding phenomena are frequently beyond the current capabilities of standard molecular dynamics (MD) simulations, enhanced-sampling algorithms are often used to study protein folding with atomistic models. Metadynamics-based approaches are among the most successful enhanced sampling methods and have been extensively used to study the folding free energy of small proteins<sup>39–43</sup>. Taking advantage of the recently developed well-tempered ensemble (WTE) approach<sup>44</sup>, we were able to use a more limited amount of replicas, namely 4, in the 300–400 K temperature range. The folding free energy landscape for the WT protein as a function of the selected collective variables (CV) is reported in Fig. 2A. The WT protein reported a folding free energy  $\Delta G^{U \rightarrow F}$  around  $-2.0$  kcal/mol, compatible with the reported melting temperature<sup>38</sup> of 45 °C. A broad main minimum is observed for the folded structure in the free energy (FE) profile, suggesting that some degree of heterogeneity is present within the folded state itself. Several intermediate states where either the first ( $\beta1\beta2$ ) or the second ( $\beta2\beta3$ ) hairpins are formed were also observed with energies 1–2 kcal/mol higher than the folded state. Molecular simulations thus give an explanation of the intrinsic

instability (values of  $\Delta G^{U \rightarrow F}$  in small folded proteins, are commonly between  $-5$  and  $-15$  kcal/mol) and provide a structural view of the interchanging conformers.

**The WW domain Y65C mutant appears to be more disordered than the WT protein.** The folding landscape of the WW Y65C mutant shares many similarities with the WT one, but the stability of the folded state is even lower, with a  $\Delta G^{U \rightarrow F}$  around  $-0.5$  kcal/mol (Fig. 2A). The folding barrier is instead predicted to be higher (3.5 kcal/mol). The free energy profiles projected along the  $\beta 1\beta 2$  contacts (Fig. 2B) show that, in addition to a stabilization of the unfolded structures, the Y65C considerably destabilized partially folded intermediates, where the  $\beta 1\beta 2$  or  $\beta 1\beta 2$  hairpins are formed. Indeed, a lower stability of the mutant is expected as the mutation destabilizes the well conserved hydrophobic core involving residues W52, Y65, N67 and P78<sup>11</sup>.

This is in agreement with the  $^1\text{H}, ^{15}\text{N}$  HSQC spectrum, whose signals are sharper, indicating higher flexibility, and with previous data<sup>38</sup> based on far and near UV CD spectra, reporting a destabilizing effect of the mutation. The authors of Ref. 38 also found a more cooperative folding for the WT, regardless the similar melting temperature (45 °C).

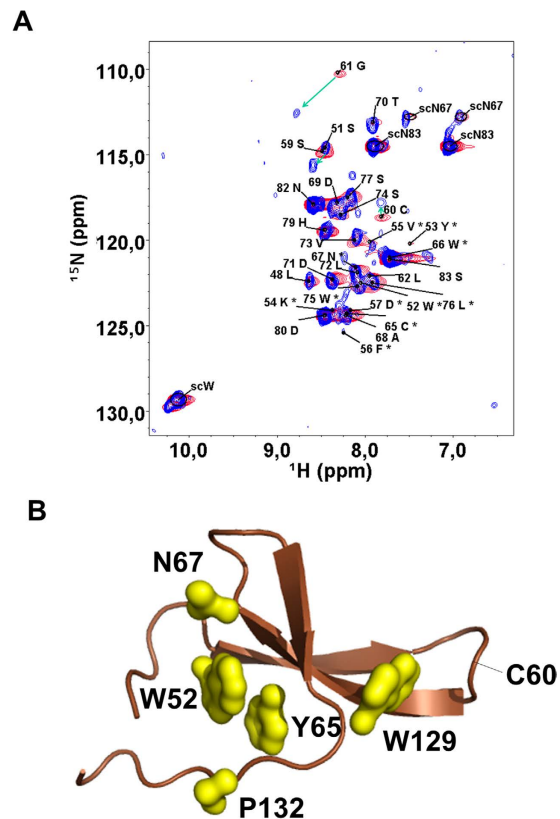
The sharper peaks and better quality of the spectrum allowed us to record a 3D  $^1\text{H}, ^{15}\text{N}$ -NOESY-HSQC spectrum with good quality cross peaks. In well-folded proteins, the inter-proton Nuclear Overhauser Effect (NOE) is a major source of structural restraints to obtain molecular structures in solution by NMR. However, the accuracy of the distance estimated by NOE is strongly reduced in presence of conformational exchange dynamics. The changes in the inter-proton distances averages the NOE and thus the intensity of the NOESY cross-peak is not anymore directly related to the a distance value. In many cases, the amplitude of the conformational dynamics brings the inter-proton distances to values where the NOE is negligible, with a consequent weakening of NOESY cross-peaks, especially for long-range inter-proton contacts. For this reason, the absence of long-range NOESY cross-peaks observed in the 3D  $^1\text{H}, ^{15}\text{N}$  NOESY-HSQC spectrum of WW Y65C provides further evidence of an extended conformational dynamics. The sharper signals in the Y65C mutant HSQC are suggestive of a different dynamical regime with respect to the wild-type; it is possible that a higher flexibility in the Y65C mutant draws the dynamics toward a fast exchange in the NMR timescale with narrower lines. However, it should be noted that our data are not consistent with an intrinsically disordered protein. Extensive broadening is still present in the spectra of the mutant; and both the Chemical Shift Index (Fig. 1B) and simulations (Fig. 2), indicate the presence of beta structures in the ensemble. Furthermore, a dramatic difference of signal intensity is observed in both proteins along the sequence (see Figure S1 in SI). This is more consistent with a model where some residual structured elements are in conformational exchange among different forms, rather than the hypothesis of a random coil peptide, that would display narrow and intense lines for almost all its amino-acids.

**The Y65C mutant of WW domain easily oxidizes *in vitro*.** The destabilization of the WW fold could, in principle, impair the binding of the WW molecular partners and be, on its own, a direct cause of the GIH syndrome. However, the oxidation of Cys<sub>65</sub> in the Y65C mutant has also been proposed as the molecular cause of a loss of function in PQBP1<sup>11</sup>. *In vitro*, the protein dimerizes<sup>38</sup>, this might be due to the formation of an intermolecular disulphide bridge.

Indeed, our data show that the purified Y65C mutant of the WW domain is a soluble protein where Cys<sub>60</sub> is clearly oxidized. This is unambiguously demonstrated by the chemical shift of Cys<sub>60</sub>  $\beta$ -carbon, which is resonating at 41.6 ppm (the expected shift in reduced cysteines is around 30 ppm)<sup>45</sup> The oxidation state of Cys<sub>65</sub> is less evident because of its very weak amide signal, which makes possible only a tentative assignment. As Cys<sub>60</sub> is in a loop (see the structure in Fig. 3B) and is directly connected to a flexible glycine residue, its involvement in inter-molecular disulphide bonds cannot be excluded, but this bond should also be formed in the WT, and we do not find any evidence for that. Addition of DTT promptly reduces Cys<sub>60</sub> (Fig. 3A) and changes in the spectrum are limited to nearby residues (Gly<sub>59</sub> and Ser<sub>61</sub>). A disulphide bond connecting two cysteines inter-molecularly is consistent with the limited changes observed in the  $^1\text{H}, ^{15}\text{N}$ -HSQC spectrum upon reduction. However, albeit small variations generally imply little change in structure, a partial loss of folding could be masked in the ensemble of conformational dynamics. Moreover, at variance with the WT, the mutant elutes as two peaks compatible with a monomer and a dimer, difficult to separate. In the NMR spectrum of the purified protein only the oxidized form is visible.

We thus cannot exclude the formation of an intramolecular disulphide bond in competition with the intermolecular one *in vitro*. In cells, however, the latter would have a stronger biological meaning, since the domain is part of a much larger protein in the cell. A way to further address this question is to ascertain if the formation of an intramolecular disulphide bridge is compatible with the proposed folding ensemble. To assess the WW propensity to form structures compatible with the formation of the disulphide bridge, we re-weighted<sup>46</sup> the FE profiles along the distance between the C $\beta$  of residues Cys<sub>60</sub> and Cys<sub>65</sub> (Tyr<sub>65</sub> in the WT). A more heterogeneous ensemble of distances was observed for the mutant, revealing an increased flexibility of the WW domain (Fig. 4).

Interestingly, structures compatible with the formation of a disulphide bond (around 4 Å) were clearly observed in the simulations. The structures with a short Cys60-X65 distance are considerably more populated in the mutant. However, short distances were only observed in the unfolded ensemble, suggesting that residues 60 and 65 are not able to approach one another when the protein is correctly folded. Indeed, the CD spectra of the mutant indicate a lower content in beta sheet with respect to the WT protein<sup>38</sup>, corroborating the hypothesis that the intramolecular disulphide bridge stabilizes a partially different fold. For further proof, we extracted a structure with low Cys<sub>60</sub>-Cys<sub>65</sub> distance from the free energy landscape of the Y65C mutant and tested the effect of the formation of a disulphide bridge *in-silico*. We performed a new PTMetaD with the same setup of the previous simulations and obtained a new folding profile (Fig. 4B). The protein explored several unfolded conformations and formed, transiently, the  $\beta 1\beta 2$  sheet, but was never able to adopt the native fold.



**Figure 3.** (A)  $^1\text{H}$ ,  $^{15}\text{N}$ HSQC spectrum of oxidized Y65C WW domain (blue, 500 MHz, 278 K) superimposed to the one of its reduced form (red). The assignment refers to the oxidized form. Green arrows indicate large chemical shift variations. (B) The hydrophobic core is disrupted by Y65C.

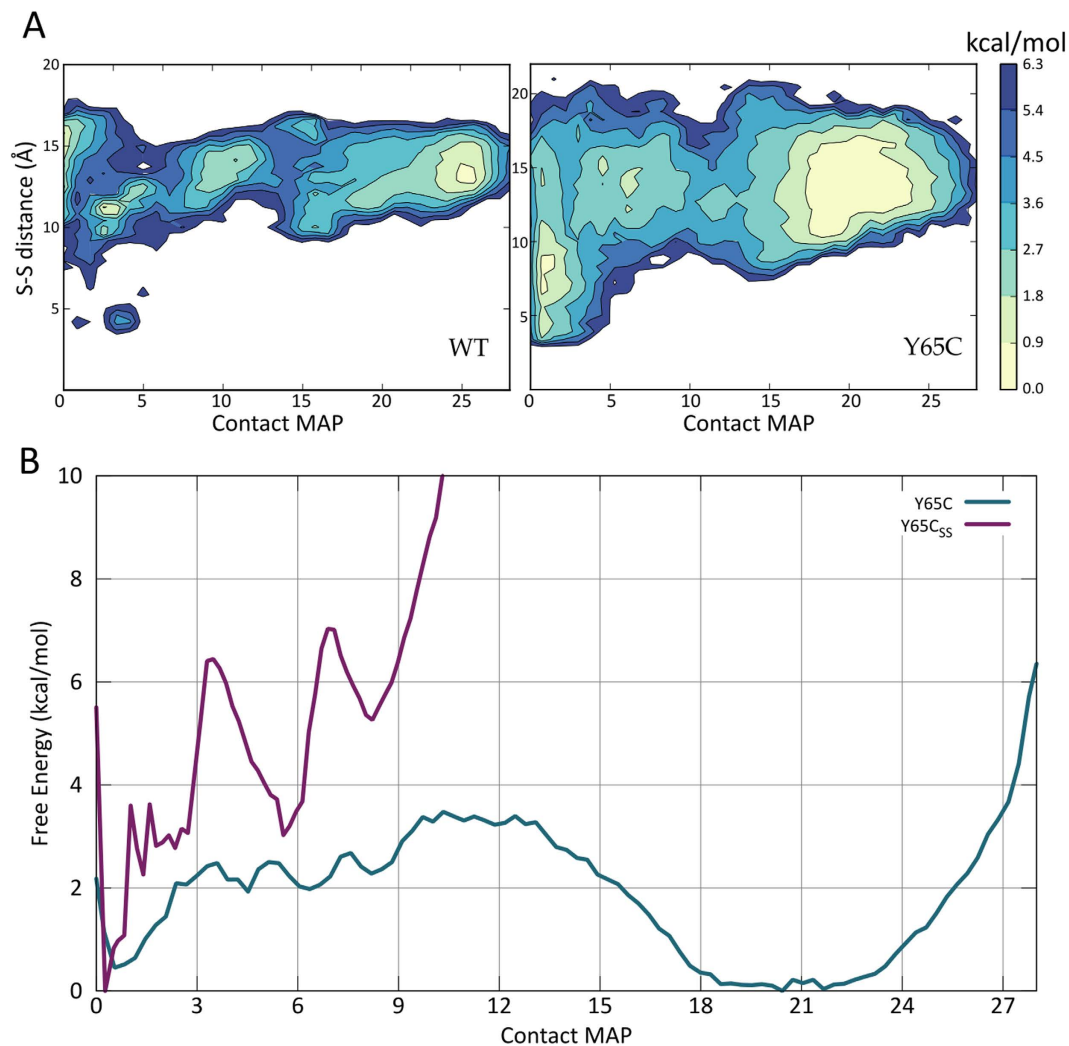
## Conclusions

Understanding the folding behavior of the WW domain in solution may be essential to unravel the molecular causes of the GIH syndrome<sup>11,38</sup>. The WW domain from PQBP1 appears to have an unstable and disordered fold, at variance with most other WW domain structures determined so far in solution<sup>16–34</sup>. Although uncommon, conformational exchange between open and closed forms have been reported for other WW domains<sup>19,21,32</sup>.

The Y65C mutation of PQBP1 affects the hydrophobic core of the protein and may make it more prone to oxidation quenching the interaction with its biological partner. The domain, however, does not have the characteristic of an intrinsically disordered protein. NMR data are consistent with extensive exchange in the micro to milliseconds time scale between random coil and beta structures, and the chemical shift index reveals a significant amount of beta-sheet character. The hydrophobic residues W52, N67, Y65, W75 and P78 form a stable core, allowing the formation of three antiparallel beta-strands. Parallel Tempering Metadynamics explains the low stability of such fold, estimating a folding free energy around  $-2.0$  kcal/mol. These values are in good agreement with thermal unfolding data, which report a melting temperature of  $45^\circ\text{C}$ <sup>38</sup>. Mutation of Y65C is expected to destabilize significantly the residual structure of the protein, as the mutated tyrosine is right in the center of the hydrophobic core. Indeed, parallel Tempering Metadynamics predicts an even lower folding free energy (around  $-0.5$  kcal/mol), while thermal unfolding suggests a more cooperative unfolding process with similar melting temperature<sup>38</sup>. Consistently, both CD<sup>38</sup> and NMR suggest increased disorder and a larger contribution from random coil conformations reflected in a narrower line-width.

Conformational variability is essential for the accessibility of cysteine residues and therefore for their oxidation. The mutant tends to form dimers *in vitro*<sup>38</sup>. Our experimental data are consistent with the presence of dimers but also monomers. In our NMR spectra, line broadening prevents the observations of the more structured residues of the protein, including Cys<sub>65</sub>. However, the signal of Cys<sub>60</sub> (located in a loop of the structural model) is well visible and clearly oxidized. This allowed us to show that, unlike the WT protein, the mutant is able to form disulphide bridges through Cys<sub>60</sub>. We cannot exclude the formation of both an intramolecular and intermolecular disulphide bridge. But since Cys<sub>60</sub> is exposed in both the WT and the mutant, in the WT it does not tend to oxidize, and in cells the WW domain is part of a large multi-domain protein, we propose that the biologically relevant disulphide bridge is formed intra-molecularly, as shown in the simulations. Indeed, Parallel Tempering Metadynamics provides a clear picture of the folding and misfolding process, showing that not only an intramolecular bridge can be formed but it would also affect the folding landscape. The model accounts for both the increased disorder caused by the mutant and the large linewidth observed after oxidation.

Reversible oxidation of cysteine has been reported in physiological<sup>47</sup> and pathological conditions<sup>11,47</sup>. As the formation of an intramolecular disulphide bridge would prevent the binding of the regulatory WW



**Figure 4.** (A) Reweight of the Free Energy along the distance between the sulfur atoms of residues Cys60 and Cys65 (Tyr65 in the WT). The population of structures with a low distance between the two residues is increased in the Y65C mutant, with values as low as 4 Å in the unfolded state (CMAP < 5) of both proteins. (B) Free Energy of the Y65C mutant with and without the disulphide bridge between Cys60 and Cys65 as a function of the contact map.

domain to its molecular partner, our model provides a new example of the role of cysteine oxidation and a plausible molecular mechanism at the basis of the GIH disease.

### Methods Summary

NMR experiments were performed on two different constructs of the WW domain of the PQBP1: a 37 amino acid long construct with sequence GLPPSWYKVFDPSCGLPY Y WNADTDLWSLSPHPNS and the corresponding Y65C (Y19C) mutant. NMR spectra were recorded on 0.5 mM samples at 278 K in 20 mM phosphate buffer at pH 6.5, 0.17 mM NaCl on Bruker Avance 500 MHz, 600 MHz, 700 MHz and 1000 MHz spectrometers. The simulations were performed with the CHARMM22\* force field<sup>48</sup> and explicit TIP3P water molecules<sup>49</sup> and the PT-MetaD<sup>39,50,51</sup> enhanced-sampling algorithm, using the software Gromacs-4.5.5<sup>52</sup> combined with the PLUMED<sup>53</sup> plugin. The simulations were run until a satisfactory convergence of the free energy reconstruction was reached (650 ns per replica)<sup>54</sup>. Full materials and methods are reported in the SI.

### References

1. Kay, B. K., Williamson, M. P. & Sudol, M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **14**, 231–241 (2000).
2. Sudol, M. & Hunter, T. NeW Wrinkles for an old domain. *Cell* **103**, 1001–1004 (2000).
3. Sudol, M., Sliwa, K. & Russo, T. Functions of WW domains in the nucleus. *FEBS Lett.* **490**, 190–195 (2001).
4. Macias, M. J., Wiesner, S. & Sudol, M. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett.* **513**, 30–37 (2002).
5. Cossée, M. *et al.* Exonic microdeletions in the X-linked PQBP1 gene in mentally retarded patients: a pathogenic mutation and in-frame deletions of uncertain effect. *Eur. J. Hum. Genet. EJHG* **14**, 418–425 (2006).

6. Kalscheuer, V. M. *et al.* Mutations in the polyglutamine binding protein 1 gene cause X-linked mental retardation. *Nat. Genet.* **35**, 313–315 (2003).
7. Kleefstra, T. *et al.* Genotype-phenotype studies in three families with mutations in the polyglutamine-binding protein 1 gene (PQBP1). *Clin. Genet.* **66**, 318–326 (2004).
8. Lubs, H. *et al.* Golabi-Ito-Hall syndrome results from a missense mutation in the WW domain of the PQBP1 gene. *J. Med. Genet.* **43**, e30 (2006).
9. Martinez-Garay, I. *et al.* A two base pair deletion in the PQBP1 gene is associated with microphthalmia, microcephaly, and mental retardation. *Eur. J. Hum. Genet. EJHG* **15**, 29–34 (2007).
10. Stevenson, R. E. *et al.* Renpenning syndrome comes into focus. *Am. J. Med. Genet. A.* **134**, 415–421 (2005).
11. Sudol, M., McDonald, C. B. & Farooq, A. Molecular insights into the WW domain of the Golabi-Ito-Hall syndrome protein PQBP1. *FEBS Lett.* 1–5, doi: 10.1016/j.febslet.2012.03.041 (2012).
12. Komuro, A., Saeki, M. & Kato, S. Association of Two Nuclear Proteins, Npw38 and NpwBP, via the Interaction between the WW Domain and a Novel Proline-rich Motif Containing Glycine and Arginine. *J. Biol. Chem.* **274**, 36513–36519 (1999).
13. Kunde, S. a. *et al.* The X-chromosome-linked intellectual disability protein PQBP1 is a component of neuronal RNA granules and regulates the appearance of stress granules. *Hum. Mol. Genet.* **20**, 4916–4931 (2011).
14. Germanaud, D. *et al.* The Renpenning syndrome spectrum: new clinical insights supported by 13 new PQBP1-mutated males. *Clin. Genet.* **79**, 225–235 (2011).
15. Jensen, L. R. *et al.* Hybridisation-based resequencing of 17 X-linked intellectual disability genes in 135 patients reveals novel mutations in ATRX, SLC6A8 and PQBP1. *Eur. J. Hum. Genet.* **19**, 717–720 (2011).
16. Bayer, E. *et al.* Structural analysis of the mitotic regulator hPin1 in solution: insights into domain architecture and substrate binding. *J. Biol. Chem.* **278**, 26183–26193 (2003).
17. Bobby, R. *et al.* Structure and dynamics of human Nedd4-1 WW3 in complex with the  $\alpha$ ENaC PY motif. *Biochim. Biophys. Acta.* **1834**, 1632–1641 (2013).
18. Fuller, A. A. *et al.* Evaluating beta-turn mimics as beta-sheet folding nucleators. *Proc. Natl. Acad. Sci. USA* **106**, 11067–11072 (2009).
19. Gao, Y.-G., Yang, H., Zhao, J., Jiang, Y.-J. & Hu, H.-Y. Autoinhibitory structure of the WW domain of HYPB/SETD2 regulates its interaction with the proline-rich region of huntingtin. *Struct. (London, Engl. 1993)* **22**, 378–386 (2014).
20. Kato, Y., Hino, Y., Nagata, K. & Tanokura, M. Solution structure and binding specificity of FBP11/HYPA WW domain as Group-II/III. *Proteins* **63**, 227–234 (2006).
21. Kowalski, J. A., Liu, K. & Kelly, J. W. NMR solution structure of the isolated Apo Pin1 WW domain: comparison to the x-ray crystal structures of Pin1. *Biopolymers* **63**, 111–121 (2002).
22. Luh, L. M. *et al.* Molecular crowding drives active Pin1 into nonspecific complexes with endogenous proteins prior to substrate recognition. *J. Am. Chem. Soc.* **135**, 13796–13803 (2013).
23. Macias, M. J. *et al.* Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* **382**, 646–649 (1996).
24. Macias, M. J., Gervais, V., Civera, C. & Oschkinat, H. Structural analysis of WW domains and design of a WW prototype. *Nat. Struct. Biol.* **7**, 375–379 (2000).
25. Ng, C. A., Kato, Y., Tanokura, M. & Brownlee, R. T. C. Structural characterisation of PinA WW domain and a comparison with other group IV WW domains, Pin1 and Ess1. *Biochim. Biophys. Acta.* **1784**, 1208–1214 (2008).
26. Ohnishi, S. *et al.* Solution structure of an atypical WW domain in a novel beta-clam-like dimeric form. *FEBS Lett.* **581**, 462–468 (2007).
27. Ohnishi, S. *et al.* Structural basis for controlling the dimerization and stability of the WW domains of an atypical subfamily. *Protein Sci. A Publ. Protein Soc.* **17**, 1531–1541 (2008).
28. Pires, J. R. *et al.* Solution structures of the YAP65 WW domain and the variant L30K in complex with the peptides GTPPPPYTVG, N-(n-octyl)-GPPPY and PLPPY and the application of peptide libraries reveal a minimal binding epitope. *J. Mol. Biol.* **314**, 1147–1156 (2001).
29. Qin, H., Pu, H. X., Li, M., Ahmed, S. & Song, J. Identification and structural mechanism for a novel interaction between a ubiquitin ligase WWP1 and Nogo-A, a key inhibitor for central nervous system regeneration. *Biochemistry* **47**, 13647–13658 (2008).
30. Ramirez-Espain, X., Ruiz, L., Martin-Malpartida, P., Oschkinat, H. & Macias, M. J. Structural characterization of a new binding motif and a novel binding mode in group 2 WW domains. *J. Mol. Biol.* **373**, 1255–1268 (2007).
31. Schelhorn, C., Martín-Malpartida, P., Suñol, D. & Macias, M. J. Structural Analysis of the Pin1-CPEB1 interaction and its potential role in CPEB1 degradation. *Sci. Rep.* **5**, 14990 (2015).
32. Smet-Nocca, C., Launay, H., Wieruszkeski, J.-M., Lippens, G. & Landrieu, I. Unraveling a phosphorylation event in a folded protein by NMR spectroscopy: phosphorylation of the Pin1 WW domain by PKA. *J. Biomol. NMR* **55**, 323–337 (2013).
33. Webb, C. *et al.* Structural features and ligand binding properties of tandem WW domains from YAP and TAZ, nuclear effectors of the Hippo pathway. *Biochemistry* **50**, 3300–3309 (2011).
34. Wiesner, S., Stier, G., Sattler, M. & Macias, M. J. Solution structure and ligand recognition of the WW domain pair of the yeast splicing factor Prp40. *J. Mol. Biol.* **324**, 807–822 (2002).
35. Mizuguchi, M. *et al.* Mutations in the PQBP1 gene prevent its interaction with the spliceosomal protein U5-15 kD. *Nat. Commun.* **5**, 3822 (2014).
36. Wishart, D. S., Sykes, B. D. & Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647–1651 (1992).
37. Hafsa, N. E. & Wishart, D. S. CSI 2.0: a significantly improved version of the Chemical Shift Index. *J. Biomol. NMR* **60**, 131–146 (2014).
38. Tapia, V. E. *et al.* Y65C missense mutation in the WW domain of the Golabi-Ito-Hall syndrome protein PQBP1 affects its binding activity and deregulates pre-mRNA splicing. *J. Biol. Chem.* **285**, 19391–19401 (2010).
39. Bussi, G., Gervasio, F. L., Laio, A. & Parrinello, M. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J. Am. Chem. Soc.* **128**, 13435–13441 (2006).
40. Marinelli, F., Pietrucci, F., Laio, A. & Piana, S. A kinetic model of trp-cage folding from multiple biased molecular dynamics simulations. *PLoS Comput. Biol.* **5**, e1000452 (2009).
41. Bonomi, M., Branduardi, D., Gervasio, F. L. & Parrinello, M. The unfolded ensemble and folding mechanism of the C-terminal GB1 beta-hairpin. *J. Am. Chem. Soc.* **130**, 13938–13944 (2008).
42. Saladino, G., Marenchino, M., Pieraccini, S., Sironi, M. & Gervasio, F. L. A Simple Mechanism Underlying the Effect of Protecting Osmolytes on Protein Folding. *J. Chem. Theory Comput.* **7**, 3846–3852 (2011).
43. Saladino, G., Marenchino, M. & Gervasio, F. L. Bridging the Gap between Folding Simulations and Experiments: The Case of the Villin Headpiece. *J. Chem. Theory Comput.* **7**, 2675–2680 (2011).
44. Bonomi, M. & Parrinello, M. Enhanced Sampling in the Well-Tempered Ensemble. *Phys. Rev. Lett.* **104**, 1–4 (2010).
45. Kraulis, P. J. Protein three-dimensional structure determination and sequence-specific assignment of <sup>13</sup>C and <sup>15</sup>N-separated NOE data. A novel real-space ab initio approach. *J. Mol. Biol.* **243**, 696–718 (1994).
46. Bonomi, M., Barducci, A. & Parrinello, M. Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J. Comput. Chem.* **30**, 1615–1621 (2009).

47. Ostman, A., Frijhoff, J., Sandin, A. & Böhmer, F.-D. Regulation of protein tyrosine phosphatases by reversible oxidation. *J. Biochem.* **150**, 345–356 (2011).
48. Piana, S., Lindorff-Larsen, K. & Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **100**, L47–L49 (2011).
49. Jorgensen, W. L. Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J. Am. Chem. Soc.* **103**, 335–340 (1981).
50. Sutto, L., Marsili, S. & Gervasio F. L. New advances in metadynamics *Wiley Interdiscip Rev Comput Mol Sci* **2**, 771–779 (2012).
51. Masetti, M., Cavalli, A., Recanatini, M. & Gervasio, F. L. Exploring Complex Protein–Ligand Recognition Mechanisms with Coarse Metadynamics. *J Phys Chem B.* **113**(14), 4807–4816 (2009).
52. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
53. Bonomi, M. *et al.* PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **180**, 1961–1972 (2009).
54. Lovera, S. *et al.* Towards a Molecular Understanding of the Link between Imatinib Resistance and Kinase Conformational Dynamics *PLoS. Comput. Biol.* **11**, e1004578 (2015).

## Acknowledgements

We acknowledge the PRACE Research Infrastructure (Hornet, Curie & MareNostrum), the e-Infrastructure South (Emerald), HecBioSim (EPSRC grant EP/L000253/1) for computational resources, and the access to large-scale NMR facilities in Lyon through the European Project Bio-NMR (Project nr. 261863). Support for this work was provided by Engineering and Physical Sciences Research Council Grant EP/M013898/1 (to FLG and GS). M.S. has been supported by Seed Grants from IMCB and NUS.

## Author Contributions

F.L.G., G.S. and M.S. conceived and designed the research. E.P.-M, N.D'A., M.L. and J.L.M.-T. performed the experiments. G.S. performed the simulations. F.L.G., G.S., N.D'A. and M.S. wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Pucheta-Martinez, E. *et al.* Changes in the folding landscape of the WW domain provide a molecular mechanism for an inherited genetic syndrome. *Sci. Rep.* **6**, 30293; doi: 10.1038/srep30293 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016