FEBS Letters 580 (2006) 2488–2494

Water molecules as structural determinants among prions of low sequence identity

Alfonso De Simone[a,b], Guy G. Dodson[a,c], Franca Fraternali[a,d,*], Adriana Zagari[b,e,*]

[a] National Institute for Medical Research, The Ridgeway, Mill Hill, NW7 1AA, London, UK
[b] Dipartimento delle Scienze Biologiche, Sezione Biostrutture and CNISM, Università di Napoli Federico II, Via Mezzocannone 16, I-80134 Napoli, Italy
[c] York Structural Biology Laboratory, University of York, York YO10 5YW, UK
[d] The Randall Centre for Molecular Mechanisms of Cell Function, King's College London, New Hunt's House, Guy's Campus, London SE1 1UL, UK
[e] Istituto di Biostrutture e Bioimmagini, CNR, via Mezzocannone 16, I-80134 Napoli, Italy

Received 27 December 2005; revised 9 February 2006; accepted 9 February 2006

Available online 17 April 2006

Edited by Hans Eklund

Abstract **The nature of the factors leading to the conversion of the cellular prion protein (PrP^C) into its amyloidogenic isoform (PrP^{Sc}) is still matter of debate in the field of structural biology. The NMR structures of non-mammalian PrP^C (non-mPrP) from frog, chicken and turtle [Calzolai, L., Lysek, D.A., Perez, D.R., Guntert, P. and Wuthrich, K. (2005) Prion protein NMR structures of chickens, turtles, and frogs. Proc. Natl. Acad. Sci. USA 102, 651–655] have provided some new and valuable information on the scaffolding elements that preserve the PrP^C folding, despite their low sequence identity with the mammalian prions (mPrP). The present molecular dynamics study of non-mPrP^C focuses on the hydration properties of these proteins in comparison with the mammalian ones. The data reveal new insights in the PrP hydration and focus on the implications for PrP^C folding stability and its propensity for interactions. In addition, for the first time, a role in disfavoring the PrP^C aggregation is suggested for a conserved β-bulge which is stabilized by the local hydration.**

Keywords: Prion protein; Protein hydration; Molecular dynamics; Edge-to-edge aggregation; Structural comparison

1. Introduction

The cellular prion protein (PrP^C) is a monomeric glycoprotein (~210aa in the mature form) composed of a flexible N-terminal region and a globular C-terminal domain of three α-helices (H1, H2 and H3) and a short double stranded antiparallel β-sheet (β1 and β2). As yet the PrP^C function(s) is(are) not known [1]. Extracellular deposition of insoluble PrP amyloid fibrils apparently occurs at the onset of transmissible spongiform encephalopathy (TSE), a group of fatal neurological disorders also known as ''prion diseases'' [2,3].

It is now established that the normal (cellular) form of prion (PrP^C) converts into an amyloidogenic isoform (PrP^{Sc}) with structural differences [4] that favor its ready aggregation to amyloid fibrils [5–7]. However the structural pathway from PrP^C to PrP^{Sc}, and the molecular basis of the subsequent fibrils formation, are poorly understood [8–10]. One factor that is important in the process is the nature of the prion's hydration and its role in the stability of PrP. As elegantly showed by Fernandez et al., amyloidogenic proteins and especially PrP^C have a large number of defectively wrapped hydrogen bonds [11–14]. These backbone hydrogen bonds are poorly protected against water interaction by flanking hydrophobic residues.

High-pressure calorimetry studies have revealed that alterations in the PrP hydration occur by passing from PrP^C to PrP^{Sc} [15]. Interestingly, solvent environment has been showed to tune the amyloidogenesis of insulin [16].

In a previous study, based on molecular dynamics (MD) simulations, we pointed out the special hydration properties at the surface of the human (huPrP) and sheep (shPrP) prions [17]. The calculations characterized protein surfaces where tightly bound waters (referred as to ''sharp spots'') evidently add to the local structural stability. Equally the hydration maps identified specific surfaces where main chain H-bonds are surrounded by very mobile bulk-like water (referred as to ''smooth spots''). These regions might exhibit energetically close alternative patterns of H-bonding effectively modulating the local structural stability and thereby favoring unfolding and aggregation events [17].

Recently, Calzolai et al. [18] have resolved the NMR structures of PrPs from chicken (chPrP), turtle (tPrP), and frog (xlPrP). These three proteins share about 30% of sequence identity with the better known mammalian PrPs (mPrPs), which themselves form a conserved group of PrPs with about 90% of sequence identity. The newly resolved non-mammalian prion structures show the general features of the PrP^C-fold; that includes a mobile disordered N-term tail and a globular C-term domain. In particular the structural comparison shows that in non-mPrPs the secondary structure elements are moderately conserved, while the major structural variability is found in the H1 packing and in some connecting loops (see Supplementary Fig. 1). As result of a detailed analysis of the sequences and the structural comparisons, scaffolding residues, important for preserving the PrP^C-fold, were identified.

*Corresponding authors.
E-mail addresses: ffranca@nimr.mrc.ac.uk (F. Fraternali), zagari@chemistry.unina.it (A. Zagari).

Abbreviations: MD, molecular dynamics; MDHS, molecular dynamics hydration sites; PrP, prion protein; PrP^C, cellular PrP; PrP^{Sc}, scrapie PrP; huPrP, human PrP; shPrP, sheep PrP; chPrP, chicken PrP; tPrP, turtle PrP; xlPrP, frog PrP; mPrP, mammalian PrP; RMSD, root mean square deviation

0014-5793/$32.00
doi:10.1016/j.febslet.2006.02.083

In the present study we have extended our analysis of mPrP hydration to the three non-mammalian prions: chPrP, xlPrP and tPrP. Taking advantage from data calculated for various low sequence identity PrPs, our analysis points out significantly conserved hydration patterns. The results are discussed in terms of folding stability and propensity for protein–protein interaction.

2. Methods

2.1. Molecular dynamics set-up

Several MD simulations in explicit solvent have been performed on the C-term globular domains of different non-PrPs NMR structures: chPrP(126–242), xlPrP(125–226) and tPrP(119–225) (pdbcode: 1U3M, 1XU0 and 1U5L, respectively) [18]. As a reference for mPrPs, the shPrP(125–230) X-ray crystal structure (pdbcode: 1UW3) [19] was selected. Throughout this paper the huPrP residue numbering is adopted for mPrP (consistently with the 1UW3 crystal structure). The MD trajectories were used for calculating the water distribution around the proteins. All the simulations were performed with the GROMACS [20] package by using GROMOS96 [21] force field. A time step of 2 fs was used. The trajectories have been saved every 250 steps (0.5 ps). The systems were simulated in an NPT ensemble by keeping constant the temperature (300 K) and pressure (1 atm); a weak coupling [22] to external heat and pressure baths were applied with relaxation times of 0.1 ps and 1 ps, respectively. The initial shortest distance between the protein and the box boundaries was 1.5 nm. The remaining box volume was filled with extended single point charge (SPCE) water model [23]. Bonds were constrained by LINCS [24] algorithm. Non-bonded interactions were accounted by using the particle mesh Ewald method (PME, grid spacing 0.12 nm) [25] for the electrostatic contribution and cut-off (14 Å) for Van der Waals contribution. The protonation states of pH sensitive sidechains were as follows: Arg and Lys were positively charged, Asp and Glu were negatively charged and His was neutral. A system of neutral charge was achieved by adding Na^+ ions. The simulations have been carried out for a time of 10 ns. Root mean square deviations (RMSDs) vs. time have been calculated on C^{α}-atoms for all PrP structures with and without loops (see Supplementary Fig. 2). As can be seen from the plots, all the systems converge to a plateau after about 2 ns. The subsequent stationary sampling (~8 ns) has been considered sufficiently long for the purpose of a hydration analysis. Further simulation details are summarized in Table 1.

2.2. Water density function

Our hydration analysis is largely based on the solvent density map [17,26,27] whose maxima are assumed to be the molecular dynamics hydration sites (MDHS). The space surrounding the protein is divided in two shells: the first describes the water around the protein and comprises the region within a distance of 0.6 nm from the protein surface. The second shell extends from 0.6 nm to 0.8 nm from the protein surface and represents the bulk solvent shell. The solvent density calculation is grid based (step-size 0.05 nm). To avoid the noisiness produced by protein translation and rotation, for each frame the atom coordinates are transformed by superimposing the current model onto a reference one. The local maxima of the density function (MDHS) are searched following the restrictions to be the highest value in a radius of 0.14 nm with a minimum density of 1.7 times the value of bulk water. Subsequently, MDHS are classified on the basis of the water residence time calculated through the time autocorrelation function.

2.3. Time autocorrelation function and residence time

The time autocorrelation function $P(\tau)$ [26] provides the probability of finding the same water in the hydration site at two different times t, $t + \tau$. The adopted formula is:

$$P(\tau) = \sum_t \delta(W(t), W(t+\tau))$$

where the delta function $\delta(W(t), W(t+\tau))$ assigns 1 or 0 whether the same water is (or is not) found in the hydration site at times t and $t + \tau$. The $P(\tau)$ curve is then fitted by a single exponential decay providing the residence time.

2.4. Calculation of the free energy of water binding

The free energy difference between two states a and b of the system is calculated with the double-decoupling method [28]. This method divides the binding process into two steps; as first, a water molecule is transferred from the bulk solvent to the gas phase (no intermolecular interactions); subsequently, it is relocated into the binding site where it is allowed to optimize the native interactions. Both the contributions are calculated with the slow growth method.

The derivate of the free energy with respect to the reaction coordinate "λ" (which is 0 in the state a and 1 in the state b) is computed according to the formula:

$$\frac{dG}{d\lambda} = \left\langle \frac{\partial H(\boldsymbol{p}, \boldsymbol{q}, \lambda)}{\partial \lambda} \right\rangle_{\lambda}$$

where $H(\boldsymbol{p},\boldsymbol{q},\lambda)$ is the classical Hamiltonian and $\boldsymbol{p}$ and $\boldsymbol{q}$ are the Cartesian coordinates and the conjugate momenta, respectively.

2.5. Linear hydration function

The linear hydration function (LHF) assigns a certain hydration value to each residue of the protein. As first each MDHS is assigned to the nearest residue. Then, the LHF is calculated according to the formula:

$$\mathrm{LHF}_{aa} = N \sum_{j=aa-2}^{j\leqslant aa+2} \left(\sum_i \frac{\rho_i}{d_i} \Big/ \mathrm{SASA}_{aa} \right)$$

where N is a normalization coefficient introduced to ensure an adimensionality and a curve area of 1. The "j" sum indicates that plot is drawn by means of a sliding window of five residues. The "i" sum is computed over the hydration sites corresponding to the residue; "d" is the distance from the nearest protein atom and "ρ" is the relative density of the hydration site. SASA is the surface accessibility of the residue (POPS server [29]).

Table 1
Parameters of the simulations

	shPrP	chPrP	tPrP	xlPrP
Structure pdbcode	1UW3	1U3M	1U5L	1XU0
Charge neutralizing ions	2 Na^+	1 Na^+	2 Na^+	–
Starting box size (Å)	$64.3 \times 63.9 \times 81.3$	$74.5 \times 74.9 \times 65.3$	$77.9 \times 64.1 \times 58.9$	$78.3 \times 67.8 \times 59.0$
Water molecules	10489	11598	9286	9895
Energies (kJ/mol)				
Protein–protein (total)	−10292.79	−10561.47	−9690.13	−9446.58
Protein–protein (LJ)	−3901.27	−4047.42	−3713.12	−3653.62
Protein–protein (El)	−6391.52	−6514.05	−5977.01	−5792.96
Protein–solvent (total)	−10372.21	−11603.65	−10151.91	−10643.94