Why are proteins with glutamine- and asparagine-rich regions associated with protein misfolding diseases?

2005 J. Phys.: Condens. Matter 17 7833

(http://iopscience.iop.org/0953-8984/17/50/005)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **17** (2005) 7833–7844 doi:10.1088/0953-8984/17/50/005

Why are proteins with glutamine- and asparagine-rich regions associated with protein misfolding diseases?

Leonor Cruzeiro

CCMAR and FCT, University of Algarve, Campus de Gambelas, 8000 Faro, Portugal

E-mail: lhansson@ualg.pt

Received 28 April 2005, in final form 27 September 2005
Published 2 December 2005
Online at stacks.iop.org/JPhysCM/17/7833

Abstract

The possibility that vibrational excited states (VESs) are the drivers of protein folding and function (the VES hypothesis) is explored to explain the reason why Gln- and Asn-rich proteins are associated with degenerative diseases. The Davydov/Scott model is extended to describe energy transfer from the water solution to the protein and vice versa. Computer simulations show that, on average, Gln and Asn residues lead to an initial larger absorption of energy from the environment to the protein, something that can explain the greater structural instability of prions. The sporadic, inherited and infectious character of prion diseases is discussed in the light of the VES hypothesis. An alternative treatment for prion diseases is suggested.

1. Introduction

The function of most proteins depends on a well defined average three-dimensional structure, and protein misfolding is associated with many different neurodegenerative diseases, such as Huntington's, Creutzfeldt-Jakob, Alzheimer's and Parkinson's in humans, as well as with scrapie in sheep and bovine spongiform encephalopathy in cows. It has been established that proteins rich in the amino-acids glutamine (Gln) and asparagine (Asn) have a greater propensity to *spontaneously* form the amyloid aggregates characteristic of misfolding diseases (Mitchelitsch and Weissman 2000). For instance, Gln repeats of lengths greater than 37 are associated with Huntington's disease (Gusella and Macdonald 2000, Perutz and Windle 2001). It has also been found that proteins that are not implicated in these diseases, such as myoglobin, can be put in conditions in which they, too, form aggregates (Dobson 2000). Aggregate formation is thus possibly a general property of proteins. However, this general property should be distinguished from the tendency prions have to fold to different average structures, even in normal cell conditions. Indeed, it is now known that prions can fold into a native, fully functioning state, [PrP^{C}], and without suffering any mutations and in the same thermodynamic conditions may also acquire another conformation, [PrP^{Sc}], which has a greater percentage of β-sheet in its secondary structure (Prusiner 1996). In spite of many studies, the

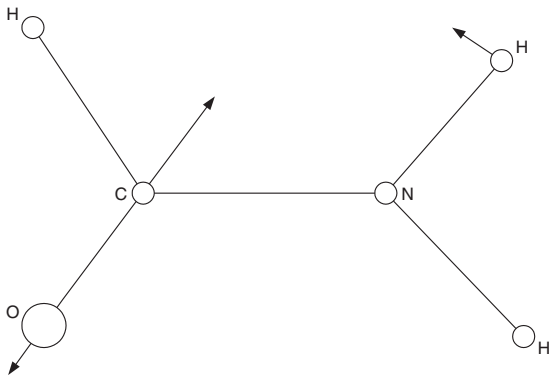

Figure 1. Illustration of the amide I of vibration of the peptide group. The arrows indicate the motions of the atoms in this normal mode. Adapted from figure 2(b) in Krimm and Bandekar (1986).

causes of these conformational changes, and also of the inherited, infectious and sporadic nature of prion diseases, remain unknown.

The main aim of this work is to put forward a possible cause for the conformational change suffered by prions and at the same time suggest a reason why greater amounts of Gln and Asn can induce such changes in proteins. This suggestion is closely connected with a model according to which protein function involves a step whereby energy is stored in the form of vibrational excited states (Scott 1992), something designated here as *the VES hypothesis*. Within the VES hypothesis it is readily understood why Gln and Asn can be so disruptive to protein folding and protein function as they are the *only two amino-acids* that can interfere directly with energy transfer in proteins. In section 5 this effect will be demonstrated in a quantitative manner.

2. The VES hypothesis

The possibility that vibrational excited states have a role in protein function was first proposed in 1973, by McClare, in the context of a 'crisis in bioenergetics' (McClare 1974). This idea was taken up by Davydov (1991) who was interested in the conformational changes responsible for muscle contraction, where the trigger and the energy donating reaction is the hydrolysis of adenosinetriphosphate (ATP). Davydov's assumption is that the first event after the hydrolysis of ATP is the storing of the energy released in the chemical reaction in the form of a vibrational mode of the peptide group, known as amide I, which consists essentially of the stretching of the C=O bond (see figure 1) and whose energy varies with secondary structure of the protein (Krimm and Bandekar 1986). In the Davydov/Scott model the interaction of the amide I mode with the vibrations of the associated hydrogen bonds leads to a localization of the amide I excitation in a few peptide groups, a mechanism known as self-trapping, that leads to a state designated in the literature as the Davydov soliton (Scott 1992). This is the state that arises at low temperatures. On the other hand, computer simulations showed that, at biological temperatures, the Davydov soliton is not stable and that amide I excitations are localized, not because of self-trapping, but because of static and dynamic irregularities (Cruzeiro-Hansson and Takeno 1997). The overall conclusion was nevertheless that the Davydov/Scott model can explain how the energy that is released in a chemical reaction at the active site can propagate, without dissipation, to other regions of the protein, where it is used for work, for instance, to produce a conformational change.

The first experimental evidence for a Davydov-like state was obtained in the crystal of acetanilide (ACN) (Careri *et al* 1984) which includes hydrogen bonded chains identical to those that stabilize the secondary structure of proteins and thus constitutes a possible model system in which to test the Davydov/Scott model. Careri and co-workers (Scott 1992) found an anomalous line, 15 cm^{-1} lower than the amide I excitation in that crystal, which was interpreted as a self-trapped, soliton-like state. This interpretation has been confirmed by Hamm and co-workers, who have applied nonlinear spectroscopic methods to the study of vibrational excited states in the organic crystals of ACN (Edler *et al* 2002, Edler and Hamm 2002, 2003, 2004) and N-methylacetamide (NMA) (Edler and Hamm 2003, 2004). Their measurements show that both the amide I and the NH stretch excitations in these crystals possess self-trapped states, with the NH stretch, whose energy is approximately twice that of amide I, being 20 times more stable. Their experimental data (Edler and Hamm 2002) also confirm the results obtained in computer simulations (Cruzeiro-Hansson and Takeno 1997), according to which at low temperature vibrational excitations are self-trapped, while at biological temperatures they are localized because of static and dynamical disorder.

Experimental support for the validity, in proteins, of the Davydov/Scott model has also been accumulating. Indeed, while an early objection to a biological role of vibrational excited states in proteins was their short lifetimes, thought to be in the subpicosecond range (McClare 1974), it is now known that, even in isolated amino-acids, the lifetime of the amide I mode can last approximately 1 ps and the first measurements of the lifetime of this mode in real proteins showed that, in myoglobin, the higher-energy side of the amide I band has a lifetime of 15 ps (Xie *et al* 2000). More recently, a low-lying band of amide I has also been found in myoglobin (Austin *et al* 2003). This state, which is thought to arise from self-trapping, has a longer lifetime, of 30 ps at 50 K, than the normal amide I band, which is 5 ps (Austin *et al* 2003). Finally, in a recent publication Hamm and co-workers study NH excitations in a model α-helix and not only find a self-trapped state for the NH stretch vibration in this system but also that it can only arise when the helical structure is intact (Edler *et al* 2004), i.e. when the C=O groups are hydrogen bonded to each other, as is assumed in the Davydov/Scott model.

In short, the experimental evidence for the kind of interactions first postulated by Davydov for amide I excitations in proteins is now widespread (Lindgard and Stoneham 2003). The lifetimes of these excitations vary from system to system and have been estimated to be 15 ps in myoglobin (Xie *et al* 2000) and 12 ps in ACN (Edler and Hamm 2002). Computer simulations show that in a few picoseconds the energy generated by the hydrolysis of ATP at the active site can travel anywhere in the protein (Cruzeiro-Hansson and Takeno 1997). Here, 5 ps simulations of energy transfer from hydration waters to the protein and vice versa will be shown. The question we are concerned with is *why proteins with greater amounts of Gln and Asn are more likely to form amyloid aggregates.* One possible answer, as detailed in the next section, is directly related to the VES hypothesis and Davydov/Scott's model.

3. What is so special about glutamine and asparagine?

Gln and Asn are polar amino-acids, a physical characteristic they share with aspartate, glutamate, serine, threonine, lysine, arginine and histidine. From the point of view of their vibrational spectra, however, Gln and Asn have a property that none of the other amino-acids, either polar or nonpolar, possess, namely, their residues can have amide I excitations (Krimm and Bandekar 1986). Thus, Gln and Asn are the only two amino-acids that can extract, to their residues, energy that is carried in the form of amide I excitations, along the protein backbone. *The main hypothesis in this article is that this fact underlies the potentially disruptive effects, to protein folding and to protein function, of large amounts of Gln and Asn in a protein sequence.*

While until now all applications of the Davydov/Scott model have included only amide I energy transfer along the protein backbone, here the possibility of amide I excitations in the residues of Gln and Asn will also be considered.

A second aspect of the simulations presented here that is different from all the simulations of the Davydov/Scott model performed before is the inclusion of the water molecules closest to the protein. Initially, Davydov/Scott's model was developed to describe the dynamics of amide I excitations generated by the hydrolysis of ATP. However, a protein in a cell, or a protein in a solution, responds to many other triggers, such as water molecules, ions, small molecules and even other proteins. All these actions may lead to local inputs of energy to the protein, that can also be in the form of amide I excitations. Indeed, in a first extension of the Davydov/Scott idea, Careri and Wyman calculated that the binding of a ligand to a protein can generate enough energy to produce two amide I vibrations in the protein and proposed that the so-called 'turning wheel' model of enzyme action (Wyman 1975) is kept running by amide I excitations (Careri and Wyman 1984). In the present work, the effect of neighbouring water molecules on the protein is considered. It is established experimentally that the amide I mode of peptides mixes with the bending mode of water (Sieler and Schweitzer-Stenner 1997). Energy can thus flow from vibrationally excited water molecules to the protein amide I vibrations, and thereby induce conformational changes in the protein, or, vice versa, energy that should be used for protein work can be diverted to water. *The main objective of the simulations presented in this article is to study the effect of the presence of Gln and Asn residues on this energy exchange between the prion and the hydration waters.*

4. Modelling vibrational energy transfer in a protein–water system

The general philosophy in this work is to obtain a representative set of equilibrium conformations of the protein–water system and assume that, at biological temperatures, the presence of one amide I excitation in the protein or of one bending mode excitation in the hydration waters does not significantly alter these conformations. This is certainly true at biological temperatures because the changes in the atomic positions induced by these excitations is at least ten times smaller than the changes induced by thermal agitation (Cruzeiro-Hansson and Takeno 1997). This being the case, vibrational energy transfer within the protein, or between the protein and water, or within the water molecules, is modelled with a Hamiltonian adapted from the Davydov/Scott model (Davydov 1991, Scott 1992), as detailed below (equations (1)–(5)).

The protein selected was a prion fragment Leu 125–Arg 228 of the human prion protein (see figure 2), whose atomic coordinates are given in file 1QLX (Zahn *et al* 2000) of the protein data bank (Berman *et al* 2000). The procedure to obtain the equilibrium conformations of the prion is as follows. The AMBER 6.0 package (Pearlman *et al* 1995, Case *et al* 1999) was used to protonate the prion structure obtained from file 1QLX and 7274 water molecules were added, to simulate a water solution. Using the AMBER 6.0 force field (Pearlman *et al* 1995, Case *et al* 1999), the system constituted by the protein and the 7274 water molecules was first energy minimized and, after a period of slow heating to 27 °C, it was equilibrated, at that temperature and at constant pressure, in a nanosecond molecular dynamics simulation. After equilibration, 1000 snapshots of the protein–water system were stored from a further 100 ps simulation, sampled at 0.1 ps intervals. The 1000 snapshots thus obtained constitute our ensemble of structures representative of the prion conformations at 27 °C.

Figure 3 shows the relative water accessibility of the C=O groups in the protein. This was estimated by finding the average number of water molecules that are within 4 Å of each carbonyl group of the protein, using the 1000 snapshots mentioned above. For example,

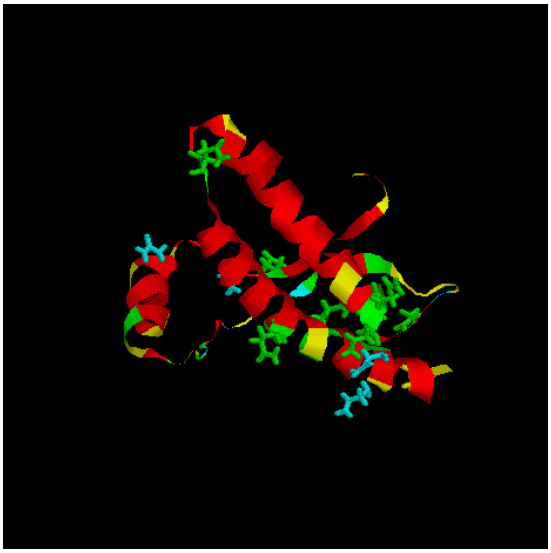

Figure 2. Three dimensional structure of the prion fragment 1QLX. Atoms in Gln and Asn residues are shown as sticks. In red are shown atoms with less than 0.3 water molecules within a 4 Å radius, in yellow those which had 0.3–0.6 water molecules within that distance, in green those which had one water molecule and in light blue those that had more than one molecule (see the text and figure 2). This figure was done with RasMol.

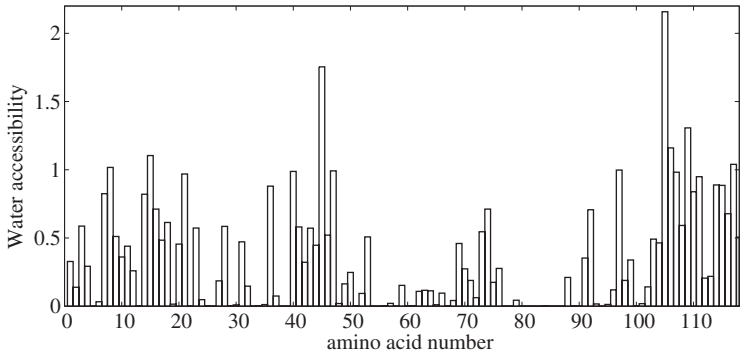

Figure 3. Histogram of the average number of water molecules within 4 Å of each carbonyl group of the prion fragment. Numbers between 105 and 114 correspond to carbonyl groups in the residues of Gln and Asn.

carbonyl groups for which there was one water molecule within a 4 Å radius in 300 of the 1000 snapshots have an average of 0.3, those for which there was one water molecule within that sphere in all of the snapshots have an average of 1, and so on. The data in figure 3 were used in the colour code of figure 2, and in red are the groups with an average number of water molecules less than 0.3. In this category are most of the backbone carbonyls in the helical parts of the prion fragment and also the carbonyls in Asn 57 and Gln 62. In yellow are the groups for which there are between 0.3 and 0.6 water molecules within a 4 Å radius, which are mainly the backbone carbonyls in the loops. In green are groups for which the average number of water molecules is between 0.6 and 1. Approximately half of the carbonyl groups in the Gln and Asn are in this category. Finally, in light blue are the groups with more than one

water molecule in a 4 Å radius. Only the carbonyl groups of the Gln and Asn residues satisfy this condition, which demonstrates the much greater contact with water that these carbonyls can have.

Vibrational energy transfer was modelled for the system constituted of all the protein C=O groups, together with the subset of all water molecules that were less than 4 Å away from a C=O group of the proteins (either in the backbone or in the Gln and Asn residues), in at least one of the 1000 snapshots sampled. This was found to amount to 238 water molecules. As the prion fragment includes 104 amino-acids and thus 103 backbone carbonyls, plus 14 extra C=O groups in the residues of its 14 Gln and Asn amino-acids, the number of sites for vibrational excitations of the whole system was a total of 355.

Vibrational energy transfer (of one amide I vibration or of one bending excitation of water) was modelled by the following Hamiltonian, $\hat{H}$:

$$\hat{H} = \hat{H}_{\rm ex} + H_{\rm at} + \hat{H}_{\rm int} \tag{1}$$

where $\hat{H}_{\rm ex}$, the exciton Hamiltonian, describes the transfer of the amide I excitation between protein sites and of the bending mode of water between water molecules, as well as the exchange of an amide I excitation in the protein with a bending excitation in a water molecule; $H_{\rm at}$, the atomic Hamiltonian, describes the fluctuations of all atoms in the system, as predicted by the AMBER force field (Pearlman *et al* 1995, Case *et al* 1999); and $\hat{H}_{\rm int}$, the interaction Hamiltonian as proposed by Davydov (1991), describes the interaction of the amide I excitation in a given C=O group with the deviations from equilibrium lengths of the hydrogen bond connected to it (if such a hydrogen bond exists, as explained below).

The exciton Hamiltonian differs from that used previously (Cruzeiro-Hansson and Takeno 1997) and is as follows:

$$\hat{H}_{\rm ex} = \sum_{n=1}^{\rm N} \epsilon_n \hat{a}_n^\dagger \hat{a}_n + \sum_{n<m=1}^{N} \left[V_{n,m} \left(\hat{a}_n^\dagger \hat{a}_m + \hat{a}_m^\dagger \hat{a}_n \right) \right] \tag{2}$$

where $N = 355$ is the total number of sites where a vibrational excitation can be found and in the first term, ϵ_n, for $n = 1, \ldots, 103$, is the energy of an *isolated* amide I vibration in the backbone of the prion. This energy depends on the conformation and local environment of the peptide groups and can be used to assess the secondary structure of proteins (Krimm and Bandekar 1986). Here, the amide I energy of an *isolated* excitation is taken to be 1660 cm $^{-1}$, as in previous simulations (Scott 1992, Cruzeiro-Hansson and Takeno 1997). For $n = 104, \ldots, 117$, ϵ_n is the energy of an isolated amide I vibration in the residues of Gln and Asn, which is taken to be 1650 cm^{-1} (Krimm and Bandekar 1986) and, for $n = 118, \ldots, 355$, it is the energy of the bending mode of water, of 1640 cm^{-1} (Sieler and Schweitzer-Stenner 1997). It should also be noted that the results are not dependent on the absolute values of these quantities but on their relative differences, i.e., the important factors are the energy difference, of 20 cm^{-1}, between the backbone amide I vibration and the bending mode of water and of 10 cm^{-1} between the amide I of Gln and Asn residues and the previous two.

The first term in (2) counts the total energy and number of these vibrational excitations, which in the simulations presented here is just one. The second term in (2) describes the transfer of the excitations from a C=O group n to a C=O group m, which is more probable the greater the magnitude of the dipole–dipole interaction, V_{nm}, between those sites, which has already been used in calculations for ACN (Eilbeck *et al* 1984) and for a globular protein (Feddersen 1991) and has the form

$$V_{nm} = \frac{1}{4\pi\epsilon_0 k} \frac{|\vec{\mu}_n||\vec{\mu}_n|}{R_{nm}^3} \left[\vec{e}_n \cdot \vec{e}_m - 3\,(\vec{u} \cdot \vec{e}_n)\,(\vec{u} \cdot \vec{e}_m) \right] \tag{3}$$

where $\epsilon_0 = 8.8542 \times 10^{-12}$ F m^{-1} is the permittivity of vacuum, k is the dielectric constant of the medium, $\vec{\mu}_n$ is the transition dipole moment in site n, $\vec{u}$ is the unit vector directed from the centre of one dipole to centre of the other, $\vec{e}_n$ is the unit vector which defines the direction of the transition dipole moment in site n and R_{nm} is the distance between the centres of dipoles in sites n and m. The positions and orientations of the transition dipole moments for the amide I excitations are calculated from the positions of the carbon and oxygen in the carbonyl groups and the position of the nitrogen in the corresponding amide groups. Nevskaya and Chirgadze (1976) estimate that, in α-helices, the transition dipole moment makes an angle of 17° away from the CO bond and in the direction of the CN bond, a value that is also within the range of 15° to 25° determined by Krimm and Bandekar (1986). The intensity of the transition dipole moment for amide I excitation in the backbone carbonyls is taken to be 0.30 D, as determined by Nevskaya and Chirgadze (1976). For the sake of simplicity, the values of the intensities of the transition dipole moments of the amide I excitations in the residues of Gln and Asn are taken to be the same values as for the backbone. With regard to the transition dipole moment of the bending mode of water, this author is not aware of any values, determined experimentally or estimated theoretically. Thus, its orientation was tentatively assumed to be from the oxygen atom to the centre of mass of the two hydrogen atoms in each water molecule, and its strength was taken to be equal to that of the amide I vibration.

The dipole–dipole interaction given by equation (3) also depends on the dielectric constant, k, of the medium. The dielectric constant of bulk water is 80 and recent calculations show that for the protein interior it is between 10 and 12 (Nguyen *et al* 2004). For the amide I excitation this dielectric constant was taken to be 10 and for the hydration waters it was taken to be 40. For the dipole–dipole couplings that correspond to an energy transfer between the protein and water this dielectric constant was taken to be the geometric average of these two values, namely, $\sqrt{10 \times 40} = 20$.

H_{at}, the atomic Hamiltonian, is a classical energy functional, which is used to study protein structure and dynamics, as well as its interactions with ligands:

$$H_{\mathrm{at}} = E\left(\{\vec{R}_n\}\right) \tag{4}$$

where $\{R_n\}$, represents the set of positions of all atoms in the protein–water system. As described by Pearlman *et al* (1995), Case *et al* (1999), this energy functional includes harmonic potentials for bond stretching and angle bending, a truncated Fourier series to represent torsions, a Lennard-Jones potential to represent nonbonded interactions and, most importantly, electrostatic interactions between atoms that are more than two covalent bonds away from each other. It is this potential that is used to select the 1000 snapshots of the prion–water system mentioned above.

Finally, the interaction Hamiltonian, $\hat{H}_{\mathrm{int}}$, is as first proposed by Davydov (1991), Scott (1992) and describes the change of the amide I energy with the length of the hydrogen bond connected with it, when such a bond exists (see below):

$$\hat{H}_{\mathrm{int}} = \chi \sum_{n=1}^{117}\left[\left(|\vec{R}_n^{\mathrm{O}} - \vec{R}_m^{\mathrm{N}}| - d_{\mathrm{eq}}\right)\cos(\theta_{nm})\,\hat{a}_n^{\dagger}\hat{a}_n\right] + \sum_{n=118}^{N}\xi_n\,\hat{a}_n^{\dagger}\hat{a}_n \tag{5}$$

$\chi = 62$ pN is the value determined experimentally (Scott 1992) and used in previous simulations (Scott 1992, Cruzeiro-Hansson and Takeno 1997), $(|\vec{R}_n^{\mathrm{O}} - \vec{R}_m^{\mathrm{N}}| - d_{\mathrm{eq}})$ is the deviation of the hydrogen bond length between the oxygen of the carbonyl group of amino-acid n and the nitrogen of the amine of amino-acid m from its equilibrium value, d_{eq}, which was determined from the set of the 1000 protein conformations mentioned above to be 2.86 Å. For each carbonyl oxygen, n, there was at most one such bond with another amino-acid, m, and when the distance $|\vec{R}_n^{\mathrm{O}} - \vec{R}_m^{\mathrm{N}}|$ between the oxygen and the nitrogen atoms was greater than 3.3 Å,

or the angle between C=O and H–N groups deviated by more than 30° from a straight line, the corresponding hydrogen bond was assumed to be broken and the corresponding term in (5) was zero. The factor $\cos(\theta_{nm})$, where θ_{nm} is the angle between the C=O bond of the carbonyl and the H–N bond of the amine, was included to take into account the strong directionality of the hydrogen bond.

The summation in the first term of (5) covers only the carbonyl sites. The water sites are taken care of by the second term where ξ_n is a Gaussian random variable with a standard deviation of 50 cm^{-1} that represents the influence of the local hydrogen connections on the energy of the bending mode of the water molecules.

In the work reported here the Hamiltonian defined by equations (1)–(5) is used to calculate the states of amide I excitations in the carbonyl groups or of the bending excitation in water molecules. Only the case of one quantum, which can either be an amide I or a bending mode, or indeed, be partially one and the other, is considered. The idea that Gln and Asn amino-acids enhance transient energy absorption from the environment is tested by comparing the absorption of energy by the protein from the water molecules, in the presence and in the absence of Gln and Asn amino-acids. Thus, energy transfer is simulated in two different proteins: one of them is the prion fragment in file 1QLX (Zahn *et al* 2000), which includes 104 amino-acids, 14 of which are Gln and Asn residues, and which consequently possesses 117 sites for amide I excitations (103 in the backbone and 14 in the side chains), as explained above, and the second protein is a 'mutation', i.e., it is the protein that would arise if all the Gln and Asn amino-acids were changed into other amino-acids, without a change in the backbone structure. This second protein only has 103 sites for amide I excitations. Together with the 238 water molecules closest to the protein this leads to one system (the prion fragment), which is labelled as QN_+, with 355 sites, and to a second system, the 'mutated' prion, with 341 sites, which is labelled QN_-. In the next section, the probability of absorption of energy from the water molecules, and of its retention in the proteins, is calculated for the $QN_\pm$ proteins.

5. Results

In order to calculate the probability of energy transfer from the water molecules to the protein we follow the path from a specific initial state in which the excitation is essentially in the form of a bending mode of water, located at a given water molecule in the hydration shell and calculate the average probability, $P_{\mathrm{wp}}^{QN_\pm}$, for such an excitation to jump to the protein, where QN_+ is for the prion fragment 1QLX and QN_- is for the mutation that does not include the 14 Gln and Asn amino-acids. The reverse situation is also considered, in which the excitation is in the form of an amide I mode, initially located at a protein site, and the average probability of it remaining there, $P_{\mathrm{pw}}^{QN_\pm}$, is estimated.

How are these average probabilities obtained? A particular site for the excitation is selected *a priori*, for instance the water that is closest to the protein. For each conformation of the QN_+ prion–water system we have 355 (or 341, for the QN_- system) possible states for one vibrational excitation, according to the Hamiltonian (1)–(5). From these 355 states that which has the highest probability for the excitation to be in the water selected is chosen. The time evolution from this initial state is as was done previously (Cruzeiro-Hansson and Takeno 1997), but using the new Hamiltonian given by equations (1)–(5). Averages are made over the 104 water molecules closest to the protein, taken as initial sites, and over 500 different starting conformations. In this way the average probability, $P_{\mathrm{wp}}^{QN_\pm}$, for transfer of energy from the water to the protein is calculated. On the other hand, the average probability, $P_{\mathrm{pw}}^{QN_\pm}$, for an excitation that arises in one of the 103 backbone carbonyls to remain there is calculated by a similar procedure, with the averages being made over the 103 backbone carbonyls, taken

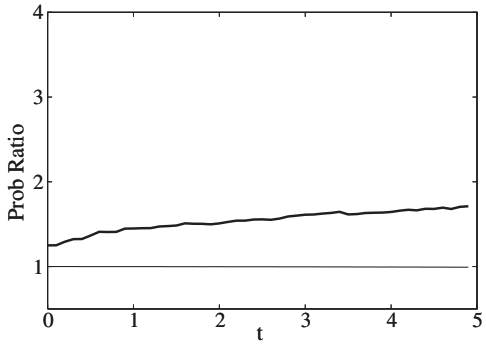

Figure 4. Ratio, $\frac{P_{\rm wp}^{QN+}}{P_{\rm wp}^{QN-}}$, of the probability of a vibrational excitation to jump from the water to the protein (thick line) and from the protein to the water, $\frac{P_{\rm pw}^{QN+}}{P_{\rm pw}^{QN-}}$, (thin line). Time is in ps.

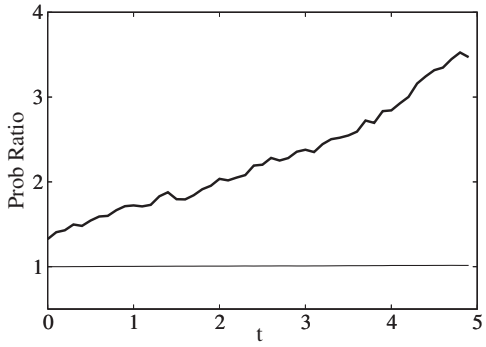

Figure 5. The same as figure 4 but for a subset of the conformations considered in figure 4.

as initial sites and also over the same 500 different starting conformations. In this way, all averages involve the same number of conformations of the protein–water system, i.e., the same number of points: 51 500.

Figure 4 shows the ratio, $\frac{P_{\rm wp}^{QN+}}{P_{\rm wp}^{QN-}}$ (thick line), of the average probability of absorption of energy from water by the prion $P_{\rm wp}^{QN+}$ and by the mutated protein $P_{\rm wp}^{QN-}$, as well as the ratio, $\frac{P_{\rm pw}^{QN+}}{P_{\rm pw}^{QN-}}$ (thin line), of the average probabilities for the energy that is generated in the proteins to remain there. Figure 4 demonstrates that, for this prion fragment, the initial absorption of energy from water to the protein is enhanced by a factor of 1.25 initially, and is approximately 1.7 after 5 ps. On the other hand, the amount of energy transferred from the protein to the water molecules is approximately the same in the presence and in the absence of Gln and Asn residues, at least for the first 5 ps. Taken together, these two results mean that the presence of Gln and Asn residues does lead to a transient increase in the vibrational energy absorbed by the prion from the hydration waters.

Figure 5 shows the same variables as in figure 4, but for a subset of prion conformations corresponding to only 50 initial conditions of the total 500. Figure 5 shows the importance of the protein conformation in this transient energy absorption. Indeed, for this particular set of protein conformations, the enhancement factor brought about by the presence of Gln and Asn amino-acids is approximately 1.3 initially but increases rapidly with time and is approximately 3.5 after 5 ps.

6. Discussion

Gln and Asn can lead to structural instability because of unfavourable interactions with neighbouring regions of the protein in which Gln and Asn find themselves. These are causes of structural instability that they share with the other amino acids. In this work, the focus is, however, on a feature that is unique to Gln and Asn, namely, that they alone can have amide I excitations in their residues.

The picture implicit in the calculations presented here is that proteins are continuously absorbing energy from the environment in the form of vibrational excited states, particularly amide I vibrations. In the previous section, results were presented for the *relative* probability of the prion fragment used in the simulations to extract energy from the bending mode of water to the amide I mode of the protein. The *absolute* probability for this extraction is likely to be very low and will depend on other factors not included in the calculation, such as the probability for the appearance of a bending vibration in a water molecule, the probability that this molecule is close enough to the protein for this excitation to have a chance of being transferred to the amide I mode of the protein, the average number of water molecules that constitute the relevant hydration shell of each protein, etc. Some of these parameters depend, in turn, on the sequence and conformation of the protein considered. What the simulations presented here show is that, when all these factors are equal, proteins with a greater number of Gln and Asn residues will have an increased capacity to extract energy from the environment in the form of amide I vibrations.

The prion fragment used here is a small protein, constituted by 104 amino-acids and a structure that is relatively open, with only 20 amino-acids for which no waters are found within a 4 Å radius, in the short time interval of 100 ps (see figure 3). But already for this small and open structure, in particular conformations, the average relative probability of transferring vibrational energy from the water molecules to the prion can be enhanced by a factor of up to 3.5 or more, when the amide I modes of the residues of Gln and Asn are taken into account (see figures 4 and 5). Similar calculations need to be done for the more real proteins found in cells, but the results here already indicate that this enhancement will depend on the conformation and that, generally, the greater the number of Gln and Asn residues, particularly at the protein surface, the greater the enhancement will be. I.e., Gln- and Asn-rich proteins are more efficient at extracting energy in this form from the environment, and the greater the percentage of these amino-acids in a protein the greater the amount of energy they will extract. It should be pointed out that this extraction is transient, that it takes place in the picosecond timescale and that the energy extracted in this way most often will just be returned to the environment, in the form of heat, and without producing any change in the average protein conformation. In the case of enzymes, it may even be beneficial and help speed up the reactions that they catalyse. But, occasionally, and particularly so for the proteins with excess amounts of Gln and Asn amino-acids, this energy can be used to promote harmful conformational changes, namely, turning α-helices into β-sheets, as happens in prions (Prusiner 1996). Cells have many defences against protein malfunction and diseases are normally the result of a series of failures of these defences. For instance, a misfolded protein will normally be degraded by the cell and even correctly folded proteins are digested within 48 h of their synthesis. But when, for some reason, a misfolding transformation is not counteracted, it may lead to the formation of protein aggregates and to disease. In this case, Gln/Asn-rich proteins have a higher probability of undergoing such processes.

This proposed interference of Gln and Asn residues in energy transfer in proteins can also explain the inherited, infectious and sporadic nature of prion diseases. The calculations here show that, in principle, *all* proteins have a finite probability of absorbing energy from

the environment and of undergoing harmful conformational changes. Let us suppose that the absolute probability of onset of disease because of a series of harmful conformational changes in a specific protein with an average number of Gln and Asn amino-acids is 60 years. Then, according to the results of the previous section, for proteins with greater amounts of Gln and Asn residues, this onset can be reduced to 20–30 years, or less, something that explains the sporadic cases of prion diseases and also the reason why they arise more frequently in older people. On the other hand, some people may have inherited prion genes with mutations that lead to a prion sequence or a prion structure for which the probability of extraction of energy is greater, so that the harmful conformational changes occur more often. For these individuals, the onset of the disease is more probable and occurs at younger ages. Finally, the infectious form can be understood if the non-native forms of the prions are even more efficient at extracting energy and thus are able to induce harmful conformational changes to the still native prions that come in contact with them. From this point of view, the species barrier is due to the fact that this energy transfer is optimal for identical molecules. On the other hand, there is also the odd possibility that a prion with a different amino-acid sequence may transfer energy reasonably efficiently to the host prion, something that will lead to the crossing of the species barrier.

The results in the previous section can also explain the findings of a search made on the genomes of 31 organisms that showed that Gln-/Asn-rich regions are essentially absent in the proteins of thermophilic bacteria (Mitchelitsch and Weissman 2000). Indeed, the efficiency of the energy transfer simulated in this work increases as the temperature increases and the results presented here suggest that Gln-/Asn-rich regions could make an average protein of a thermophile too prone to undergo harmful conformational changes. Therefore, organisms that must survive at higher temperatures cannot afford to possess proteins with such regions.

It should also be emphasized that the results of the previous section include several approximations whose validity must be tested. First, there are the inherent approximations of the AMBER potential (Pearlman *et al* 1995, Case *et al* 1999). Secondly, the parameters used for the isolated amide I energies, for the transition dipole moments of the amide I vibration and for the bending mode of water, the dielectric constants for the protein and for the hydration waters and the nonlinearity parameter χ must be estimated more accurately. Also, the influence of local environment on the energy of the bending mode of water should be calculated as for the amide I excitation. For these reasons, the relative probabilities presented above have a qualitative value and cannot be considered exact quantitative estimates for energy transfer from the water to the protein and vice versa. However, the uncertainties above do not affect the general conclusion that Gln and Asn residues enhance energy transfer from the water molecules to the protein and can thus affect its folding and function.

Theories should be judged by what they can explain and be tested by the predictions they lead to. A first prediction of the VES hypothesis and of the work reported here is that one of the distinctions of the different prion strains is in the number of Gln and Asn in positions accessible to the solution. A second prediction is that reducing the number of Gln and Asn at the protein surface, or chemically modifying them so that the amide I mode does not have an energy similar to that of the bending mode of water, should render prions less infectious. Such chemical modifications thus constitute an alternative mechanism for the treatment of prion diseases. A third prediction is that repeats of amino-acids other than Gln and Asn do not have the same pathological consequences.

Acknowledgment

This work was funded in part by the Portuguese Foundation for Science and Technology.

References

Austin R H, Xie A, van der Meer L, Shinn M and Neil G 2003 Self-trapped states in proteins? *J. Phys.: Condens. Matter* **15** 1–6

Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N and Bourne P E 2000 The protein data bank *Nucleic Acids Res.* **28** 235–42

Careri G, Buontempo U, Galluzzi F, Scott A C, Gratton E and Shyamsunder E 1984 Spectroscopic evidence for Davydov-like solitons in acetanilide *Phys. Rev.* B **30** 4689–702

Careri G and Wyman J 1984 Soliton-assisted unidirectional circulation in a biochemical cycle *Proc. Natl Acad. Sci. USA* **81** 4386–8

Case D A, Pearlman D A, Caldwell J W, Cheatham T E III, Ross W S, Simmerling C L, Darden T A, Merz K M, Stanton R V, Cheng A L, Vincent J J, Crowley M, Tsui V, Radmer R J, Duan Y, Pitera J, Massova I, Seibel G L, Singh U C, Weiner P K and Kollman P A 1999 *AMBER 6* University of California, San Francisco

Cruzeiro-Hansson L and Takeno S 1997 Davydov model: the quantum, quantum-classical, and full classical model *Phys. Rev.* E **56** 894–906

Davydov A S 1991 *Solitons in Molecular Systems* 2nd edn (Dordrecht: Kluwer–Academic)

Edler J and Hamm P 2002 Self-trapping of the amide I band in a peptide model crystal *J. Chem. Phys.* **117** 2415–24

Edler J and Hamm P 2003 Two-dimensional vibrational spectroscopy of the amide I band of crystalline acetanilide: Fermi resonance, conformational substates, or vibrational self-trapping? *J. Chem. Phys.* **119** 2709–15

Edler J and Hamm P 2004 Spectral response of crystalline acetanilide and N-methylacetamide: vibrational self-trapping in hydrogen-bonded crystals *Phys. Rev.* B **69** 214301

Edler J, Hamm P and Scott A C 2002 Femtosecond study of self-trapped vibrational excitons in crystalline acetanilide *Phys. Rev. Lett.* **88** 067403

Edler J, Pfister R, Pouthier V, Falvo C and Hamm P 2004 Direct observation of self-trapped vibrational states in α-helices *Phys. Rev. Lett.* **93** 106405

Eilbeck J C, Lomdahl P S and Scott A C 1984 Soliton structure in crystalline acetanilide *Phys. Rev.* B **30** 4703–12

Feddersen H 1991 Localization of vibrational-energy in globular protein *Phys. Lett.* A **154** 391–5

Fersht A 1999 *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding* (New York: Freeman) p 325

Gusella J F and Macdonald M E 2000 Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease *Nat. Rev. Neurosci.* **1** 109–15

Krimm S and Bandekar J 1986 Vibrational Spectroscopy and conformation of peptides, polypeptides and proteins *Adv. Prot. Chem.* **38** 181–364

Lindgard P A and Stoneham A M 2003 Selftrapping, biomolecules and free electron lasers *J. Phys.: Condens. Matter* **15** V5–9

McClare C W F 1974 Resonance in bioenergetics *Ann. New York Acad. Sci.* **227** 74–97

Mitchelitsch M D and Weissman J S 2000 A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions *Proc. Natl Acad. Sci. USA* **97** 11910–5

Nevskaya N A and Chirgadze Yu N 1976 Infrared spectra and resonance interactions of Amide-I and II vibrations of α-Helix *Biopolymers* **15** 637–48

Nguyen D M, Reynald R L, Gittis A G and Lattman E E 2004 X-ray and thermodynamic studies of staphylococcal nuclease variants I92E and I92K: insights into polarity of the protein *Interior J. Mol. Biol.* **341** 565–74

Pearlman D A, Case D A, Caldwell J W, Ross W S, Cheatham T E III, DeBolt S, Ferguson D, Seibel G and Kollman P 1995 AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules *Comput. Phys. Commun.* **91** 1–41

Perutz M F and Windle A H 2001 Cause of neural death in neurodegenerative diseases attributable to expansion of glutamine repeats *Nature* **412** 143–4

Prusiner S B 1996 Molecular biology and pathogenesis of prion diseases *Trends Biochem. Sci.* **21** 482–7

Scott A 1992 Davydov's soliton *Phys. Rep.* **217** 1–67

Sieler G and Schweitzer-Stenner R 1997 The amide I mode peptides in aqueous solution involves vibrational coupling between the peptide group and water molecules of the hydration shell *J. Am. Chem. Soc.* **119** 1720–6

Wyman J 1975 The turning wheel: a study in steady states *Proc. Natl Acad. Sci. USA* **72** 3983–7

Xie A, Meer L V D, Hoff W and Austin R H 2000 Long-lived amide I vibrational modes in myoglobin: breathers in biology *Phys. Rev. Lett.* **84** 5435–8

Zahn R, Liu A, Luhrs T, Riek R, Von Schroetter C, Garcia F L, Billeter M, Calzolai L, Wider G and Wuthrich K 2000 NMR solution structure of the human prion protein *Proc. Natl Acad. Sci. USA* **97** 145–50