DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES S
Volume **4**, Number **5**, October **2011**

doi:10.3934/dcdss.2011.4.1033

pp. **1033–1046**

THE VES HYPOTHESIS AND PROTEIN MISFOLDING

Leonor Cruzeiro

CCMAR and FCT, Universidade do Algarve
Campus de Gambelas
Faro 8005-139, Portugal

Abstract. Proteins function by changing conformation. These conformational changes, which involve the concerted motion of a large number of atoms are classical events but, in many cases, the triggers are quantum mechanical events such as chemical reactions. Here the initial quantum states after the chemical reaction are assumed to be vibrational excited states, something that has been designated as the VES hypothesis. While the dynamics under classical force fields fail to explain the relatively lower structural stability of the proteins associated with misfolding diseases, the application of the VES hypothesis to two cases can provide a new explanation for this phenomenon. This explanation relies on the transfer of vibrational energy from water molecules to proteins, a process whose viability is also examined.

1. **Introduction.** Proteins are the macromolecules of life since they mediate most of the processes that take place inside a living cell. In many cases their function is related to conformational changes which are concerted movements of large numbers of atoms, from a few hundreds to tens or even hundreds of thousands [17, 1, 3]. Although the initial and final structures of the proteins are sometimes known, the sequential chain of events that takes a protein from the initial conformation to the final conformation is largely unknown. Since these conformational changes involve the movement of protein domains with a large number of atoms they can be considered as classical events and have consequently been modelled with classical molecular dynamics (MD) potentials [27, 6]. Two approaches that have been applied are Normal Mode Analysis (NMA) [36] and Principal or Essential Dynamics (ES) [25], both of which are linearized schemes that neglect the nonlinear character of atom-atom interactions in proteins. Furthermore, although NMA and ES have identified modes that can be related to some observed conformational changes, it has not yet been possible to define initial conditions which, together with the atom-atom interactions, lead to those conformational changes and the only successful manner to go from an initial protein conformation to a final one, in a computer simulation, has so far been through Steered or Targeted Molecular Dynamics, in which artificial forces are added to drive the protein to a specific final conformation [33].

The classical MD simulations of protein conformational changes also tend to neglect the fact that the initial trigger is a delivery of energy to a small region of the protein - the active site. In many cases this trigger is a quantum mechanical event, namely, the chemical reactions of hydrolysis of Adenosinetriphosphate (ATP) or of

2000 *Mathematics Subject Classification.* Primary: 92C05, 82C10; Secondary: 92C37.
Key words and phrases. Davydov/Scott model, neurodegenerative diseases, ice worm.

Guanosidetriphosphate (GTP). Thus, to understand how a conformational change takes place we must know how the energy released by the chemical reaction eventually leads to the large scale domain motion that constitutes the conformational change, that is, we must know how the initial quantum state that is the immediate outcome of the chemical reaction eventually leads to the classical event constituted by the concerted motion of a large group of atoms. Here we concentrate on the initial quantum event and our main hypothesis is that the initial quantum state is a vibrational excited state of the peptide group, something that, in previous studies [7, 9], has been designated as *the VES hypothesis*. Although this designation was first used in 2005, the possibility that vibrational excited states have a role in protein function has a much longer history and was first proposed in 1973, by McClare, in the context of a "crisis in bioenergetics" [23]. McClare's idea was taken up by Davydov [12] who was interested in the conformational changes responsible for muscle contraction that are triggered by the hydrolysis of ATP. Davydov's assumption was that the energy released in the chemical reaction is stored in a well-known quantum mode of the peptide group, the amide I mode, which consists essentially of the stretching of the C=O bond [20]. In the Davydov/Scott model the interaction of the amide I mode with the vibrations of the associated hydrogen bonds leads to a self-trapped state known in the literature as the Davydov soliton [12, 34]. This is the state that arises at low temperatures. At high temperatures the Davydov soliton is not stable but computer simulations have shown [10, 11] that the amide I states are still localized (even more so than the Davydov soliton), not because of self-trapping, but because of static and dynamic irregularities (Anderson localization [2]). These numerical results have been confirmed by experimental measurements in the organic crystal of acetanilide (ACN) [14].

The classical descriptions of protein dynamics and of protein conformational changes are also unable to provide explanations for the cases in which the latter processes lead to non-native structures and protein aggregation, as happens in the so-called amyloid or misfolding diseases [13]. Two examples of misfolding diseases are prion diseases [29, 30, 32] and Huntington's [28, 21]. Prions are the proteins associated with scrapie in sheep, with bovine spongiform encephalopathy in cows and with variant-Creutzfeldt-Jakob disease in humans [29, 30, 32], all of which are neurologic disorders that lead to nervous system degeneracy. It is known that prions can fold into a native, fully functioning state, designated by $[\mathrm{PrP}^{C}]$, and without suffering any amino acid mutations and in the same thermodynamic conditions, they may also acquire another, pathogenic conformation, named $[\mathrm{PrP}^{Sc}]$, which has a greater percentage of β-sheet in its secondary structure [31]. One particularity of the misfolded form, $[\mathrm{PrP}^{Sc}]$, is that it seems to be able to induce misfolding in correctly folded proteins [29, 30, 32]. In spite of many studies, the causes of these harmful conformational changes remain unknown. Similarly, Huntington's disease, another neurodegenerative disease, continues to baffle researchers. Although it is now known that the primary cause is an expansion of stretches of the amino acid glutamine (GLN) in the huntingtin protein [28, 21], the structure of this protein is not yet fully resolved and the causes for protein aggregation remain obscure.

In section 3, it is demonstrated how the VES hypothesis can provide a general explanation for the structural instability of the proteins associated with misfolding diseases. This explanation makes use of the fact that the amide I energy of the peptide groups is approximately the same as the energy of the bending mode of water molecules where the latter consists of the angle bending, with respect to each

other, of the two O-H bonds in the same water molecule. Indeed, while the amide I energy is approximately 1660 cm^{-1} [20], the energy of the bending mode of water is approximately 1640 cm^{-1} [35]. The consequence of the resonance between these two vibrational modes is that a bending mode excitation in a water molecule that is close to a peptide group of a protein can jump to this peptide group, taking the form of an amide I excitation, and vice-versa, i.e., the amide I mode and the bending mode excitations can mix, something that has been observed experimentally [35]. The calculations presented in section 3 show that the mixing between the amide I and the bending mode can lead to an enhanced transfer of energy from the water to the proteins associated with misfolding diseases and thus explain their greater structural instability compared to other proteins.

The paper is organized as follows: in the next section the Davydov/Scott Hamiltonian, generalized in order to take into account the full atomic structure of a protein, is introduced and the equations of motion are derived; in section 3 two applications of the VES hypothesis to protein misfolding are presented and further arguments for the viability of this hypothesis are put forward; finally, the article ends with a discussion of the relation of the results presented to a possible cause of the harmful protein conformational changes in neurodegenerative disorders.

2. **Theory.** The phenomenon to be considered here is the propagation of a quantum of vibrational excitation in a protein-water system. The vibrational excited state may be an amide I vibration when the excitation is in a peptide group or a quantum of bending mode when the excitation is in a water molecule. To describe the propagation of these excitations, a generalized version of the Davydov/Scott Hamiltonian [12, 34] is used that has three terms, as in the original studies:

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

where $\hat{H}_{\mathrm{ex}}$, the quantum excitation Hamiltonian, describes the storage and transfer of vibrational excitations (either amide I or bending mode) among their respective sites (which for amide I modes are the C=O groups of amino acids and for the bending mode are the water molecules); H_{at}, the atomic Hamiltonian, describes the motions of all the atoms in the protein and $\hat{H}_{\mathrm{int}}$, the interaction Hamiltonian, describes the interaction of the quantum amide I excitation in a given C=O group with the deviation, from its equilibrium length, of the hydrogen bond connected to it, when such a bond exists (the definition of this hydrogen bond is specified in the two paragraphs after eqn(5)).

The Hamiltonian (1) constitutes a generalization of the original Davydov/Scott Hamiltonian for the reasons detailed below. The quantum excitation Hamiltonian, $\hat{H}_{\mathrm{ex}}$, is given by:

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

where $\hat{a}_n^{\dagger}(\hat{a}_n)$ are the creation(annihilation) operators for an amide I excitation or a bending mode of water at site n, N is the total number of amide and bending mode sites where a vibrational excitation can be found and $\epsilon_n = 1660\ \mathrm{cm}^{-1}$ when n stands for a C=O group and $\epsilon_n = 1640\ \mathrm{cm}^{-1}$ when n stands for a water molecule. A first generalization with respect to the original Davydov/Scott Hamiltonian is the term in ξ_n, where ξ_n is a random variable with a Gaussian distribution centred on zero and with a standard deviation of $10\ \mathrm{cm}^{-1}$. This term was added to take

into account the effect of the different local environments in which the amide I or the bending mode excitations can find themselves in.

The first term in (2) includes the operator for the number of excitations in each site, $\hat{a}_n^\dagger \hat{a}_n$, which, in this study, will be one. The second term in (2) describes the transfer of the excitations from amide site n to amide site m, which is the more probable the greater the magnitude of the dipole-dipole interaction between the sites, V_{nm}. A second generalization with respect to the original Davydov/Scott Hamiltonian is that the excitation Hamiltonian (2) goes beyond the nearest neighbor approximation used in most applications [12, 34, 11] and considers the dipole-dipole interactions between *all* N excitation sites, as has also been done with the discrete self-trapping equation for the crystal of ACN [15] and for the excitations in a globular protein [16]. The dipole-dipole interaction can be found in any textbook on electromagnetism, for instance [19], and has the form:

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

where $\epsilon_0 = 8.8542\times 10^{-12}$ F/m is the electric permittivity of the vacuum, k is the dielectric constant of the medium, $\vec{\mu}_n$ is the transition dipole moment of the excitation in site n, $\vec{u}$ is the unit vector directed from the center of the dipole in site n to center of the dipole in site m, $\vec{e}_n$ is the unit vector that defines the direction of the transition dipole moment in site n and R_{nm} is the distance between the centers of dipoles in sites n and m. In the following, 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 same peptide group [20, 26]. Nevskaya and Chirgadze [26] estimate that, in α-helices, the intensity of the transition dipole moment of the Amide I excitation is 0.3 D and that it makes an angle of 17^o away from the CO bond and in the direction of the CN bond, a value that is also within the range of 15^o to 25^o determined by Krimm and Bandekar [20]. On the other hand, in the absence of any values for the transition dipole moment of the bending mode of water, either theoretical or experimental, its orientation was tentatively assumed to be from the oxygen atom to the center of mass of the two hydrogen atoms in each water molecule and, to keep the number of parameters in (2) as small as possible, its strength was taken to be equal to that of the amide I vibration.

Thirdly, a most important generalization in Hamiltonian (1) with respect to the original Davydov/Scott Hamiltonian [12, 34] is the fact that full atomic structure of the protein-water systems is taken into account; in this spirit, the interactions between all the atoms is given by the classical potential AMBER [27, 6]:

$$\begin{aligned} H_{\rm at} = &\sum_{\rm bonds} K_d(d-d_{eq})^2 + \sum_{\rm angles} K_\theta(\theta-\theta_{eq})^2 + \\ &+ \sum_{\rm dihedrals} \frac{V_n}{2}[1+\cos(n\phi-\gamma)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}}\right], \end{aligned} \tag{4}$$

where $R_{ij} = |\vec{R}_i - \vec{R}_j|$, $\vec{R}_j$ being the three-dimensional position of atom j in the protein-water system. While in the original Davydov/Scott Hamiltonian [12, 34] the site motions are described by a harmonic potential, the atomic Hamiltonian above includes also nonlinear terms. Indeed, although covalent bonds between two atoms (first term) and angle bending between two consecutive covalent bonds (second

term) are represented by harmonic potentials, torsions (third term) are represented by a truncated Fourier series and hydrogen bonds and other nonbonded interactions (fourth term) are represented by a Lennard-Jones potential with the electrostatic interactions represented by a Coulomb potential. The nonlinear atomic Hamiltonian (4) depends on many empirical parameters (K_d, d_{eq}, K_θ, θ_{eq}, V_n, n, γ, A_{ij}, B_{ij}, q_j) which have been determined by fittings to experimental data and by comparisons with fully quantum calculations [27, 6]. In fact, the development of these atomic molecular dynamics potentials is still being pursued but they have already been successfully applied by the pharmaceutical and biotechnology industries to rational design of drugs. It should also be noted that, as in many previous studies [12, 34, 11], the motions of all the atoms in the protein-water system are treated classically.

Finally, the third term in (1), the interaction Hamiltonian, $\hat{H}_{\rm int}$, is as follows:

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

As in the original Davydov/Scott Hamiltonian χ is the nonlinearity parameter whose value is taken to be 62 pN , as in many other studies [12, 34, 11]. This Hamiltonian represents the effect that changes in the hydrogen bond length $\left(|\vec{R}_n^O - \vec{R}_m^N| \right)$ have on the amide I energy ϵ_n. How are the hydrogen bonds defined ? For a given oxygen atom in the C=O group of amino acid n, a search is made over the nitrogen atoms of the NH groups of all the other amino acids $m \neq n$ to find whether one of these nitrogens is at a distance 4.5 Å or less away from the oxygen, in which case the C=O group of amino acid n is considered to be hydrogen-bonded to the NH group of amino acid m; the equilibrium value, d_{eq}, of these hydrogen bonds was estimated from the equilibrium ensemble of conformations of the protein-water system to be 2.98 Å and the term $\left(|\vec{R}_n^O - \vec{R}_m^N| - d_{eq} \right)$ is thus the deviation from equilibrium values of the hydrogen bonds, as usual [12, 34, 11].

Another extension with respect to the original Davydov/Scott model is the factor $\cos(\theta_{nm})$ which takes into account the strong directionality that characterizes hydrogen bonds, θ_{nm} being the angle between the C=O bond of amino acid n makes with the H-N bond of amino acid m. When this angle is greater than 30^o the hydrogen bond between the C=O and the NH groups is considered to be broken and the corresponding term in (5) is set to zero. It is found that, together with the threshold value of 4.5 Å in d_{eq}, the angle dependence makes for a relatively smooth variation of the interaction Hamiltonian with the positions of the oxygen and nitrogen atoms of the amino acids.

Since the atom motions are classical, the components of the positions and velocities of all atoms are real; on the other hand, the motion of the vibrational excitation (either an amide I of a peptide group or a bending mode of a water molecule) is quantum mechanical, a difference that is marked by the hats above the excitation (2) and the interation (5) Hamiltonians. In this approach, the exact wavefunction for one quantum of excitation is:

$$\Psi \left(\{\vec{R}_n\}, t \right) = \sum_{r=1}^{N} \varphi_r \left(\{\vec{R}_n\}, t \right) \hat{a}_r^\dagger \, |0> \tag{6}$$

where φ_r is the probability amplitude for an amide I excitation in the peptide group r, or for a bending mode excitation in water molecule r, as the case may be, whose specific dependence on the conformation of the protein-water system and on time

is not specified *a priori* and will be determined by the equations of motion. The latter are derived by substituting the wavefunction (6) in the Schrödinger equation for the Hamiltonian (1-5) and using Hamilton's equations for the classical part, and are as follows:

$$\hat{H}\,\Psi = E\,\Psi \tag{7}$$

$$M_r\,\ddot{\vec{R}}_r = -\vec{\nabla}H_{\rm at}\left(\{\vec{R}_n\}\right) - \chi\,\cos\left(\theta_{rm}\right)\,\mid\varphi_r\mid^2\,\vec{u}_r + \vec{F}_r(t) - \Gamma\,\dot{\vec{R}}_r \tag{8}$$

where $\vec{u}_r$ is the unit vector from the oxygen atom of the C=O group of amino acid r to the hydrogen atom of the NH group of amino acid m to which amino acid r is hydrogen bonded (when this bond exists) and only affects directly the dynamics of those two atoms, i.e. this term is zero for all other atoms.

Eqn.(8) is a classical Langevin equation and describes the coupling of the protein-water system to a thermal bath at temperature T. $\vec{F}_r(t)$ is a stochastic force applied to atom r due to the thermal fluctuations and $-\Gamma\,\dot{\vec{R}}_r$ is a damping term, Γ being the strength of the damping. For eqn.(8) to lead to the correct thermal statistics of classical systems, the stochastic forces and the damping terms must obey the fluctuation-dissipation theorem $< F_r(t)\,F_m(t') >= 2M_r\Gamma k_B T\delta_{rm}\delta(t-t')$, k_B being the Boltzmann constant.

Without the second term, eqn(8) is that which is used for molecular dynamics of proteins in water with the AMBER force field [27, 6]. The second term, on the other hand, represents the influence that a quantum excitation in amino acid r has on the position of the oxygen and hydrogen atoms of the C=O and NH groups, respectively. Since it is found that, at $T = 300$ K, the other terms are, an average, ten to hundred times stronger than the second term, in the calculations presented in section 3 the second term was neglected.

It should be pointed out that the equations of motion (7) and (8) represent the coupling of a *stationary* Schrödinger equation for the quantum excitation to *classical* equations of motion for the atoms in the protein-water system. The reasons for this choice are explained in detail in references [10, 11]. A short version is as follows. It has been found that when the dynamical thermalization of a mixed quantum-classical system is accomplished by coupling the classical part to a classical bath, as is done in (8), the result is that the quantum part will obey classical statistics as well [10, 11, 22]. The reason is that when the time-dependent Schrödinger equation is coupled to the classical Langevin equation without any precautions, the classical bath makes the quantum states diffuse in the quantum phase space in a manner that violates the quantum statistics rule according to which only eigenstates matter for the quantum thermal average. One solution to this problem is that implemented in (7) and (8), in which the classical Langevin equation for the atoms in the protein-water system is coupled to the stationary Schrödinger equation for the quantum excitations (which ensures that only eigenstates of the latter are considered). This solution is only strictly valid when the quantum excitation responds very fast to any changes in the classical conformation, an approximation which is assumed valid here.

Finally, it is important to explain how the equations of motion (7) and (8) are integrated. The initial condition is a given conformation of the protein-water system (a set of positions for all the atoms) and a set of velocities e.g. obeying a Boltzmann distribution at a given temperature. From the positions of all oxygens in the C=O groups and of all hydrogen and all nitrogens in the NH groups, the dipole-dipole

interactions (3) are calculated and inserted in the stationary Schrödinger equation (7); then, the eigenvalue problem is solved numerically and a new quantum state state is selected in a Monte Carlo step, to make sure that in the end of many such selections the quantum states have been sampled with the proper Boltzmann weight; also, to satisfy Franck-Condon factors a new state is only accepted when its overlap with the old state is sufficiently large; once a new state is selected, it can be substituted in the Langevin equation (8) and the integration of this classical equation advances one step and leads to a new conformation, with which the same procedure is applied all over again. As explained above, the second term in (8) has been neglected and thus, while the conformation of the protein-water system influences strongly the states assumed by the quantum excitation, the latter does not have an influence in those conformations.

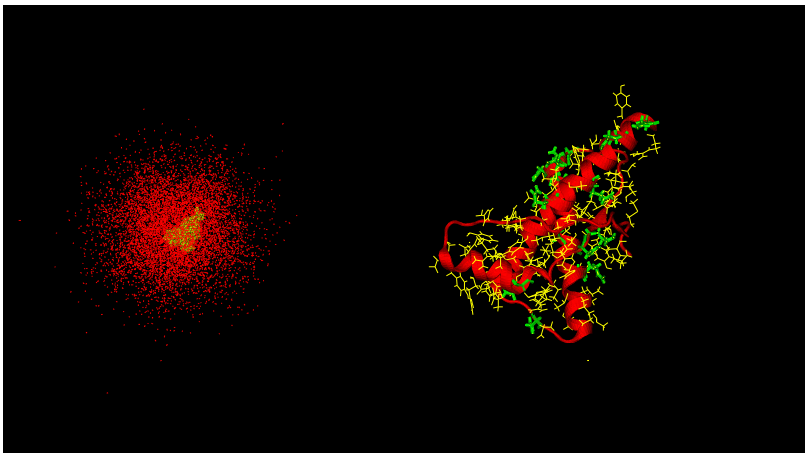

FIGURE 1. **Left figure**: Three dimensional display of the prion-water system; the water molecules are in red and the protein is in yellow. **Right figure**: Detail of the three-dimensional structure of the prion; the backbone is in red, the Glutamine and Asparagine side chains are in thick green and the remaining side chains are in thin yellow.

3. **Results.** The proteins considered here are associated with two misfolding diseases, namely, prion diseases and Huntington's diseases. The proteins associated with these diseases have one common feature in their amino acid sequences (which is in fact shared by all proteins associated with misfolding diseases): they both have regions rich in either the amino acids GLN and/or Asparagine (ASN). In section 3.2 the relevance of this fact to protein misfolding, in the light of the VES hypothesis, is emphasized.

The structure of the prion has been solved experimentally and the atom coordinates that can be found in file PDB1QLX [37], publicly available in the Protein Data Bank [4], have been used as an initial condition for the prion protein. On the other hand, the structure of the protein huntingtin is still unknown; what is known is that the disease is caused by mutations in huntingtin in which extra stretches of GLN amino acids are added to the normal primary sequence. For instance, huntingtin with glutamine repeats constituted by more than 37 GLNs lead to misfolding

of huntingtin and aggregation [18]. Thus, also studied here is a helix constituted by 46 GLN amino acids.

Both the prion and the polyglutamine helix were immersed in an explicit water bath, using the Leap module of AMBER [27, 6] and the resulting full protein-water systems are displayed in figures 1 and 2. The total number of atoms is 23521 for the prion-water system and 18398 for the polyglutamine-water system, the prion itself contributing 1696 atoms and the poly-glutamine helix alone possessing 785 atoms.

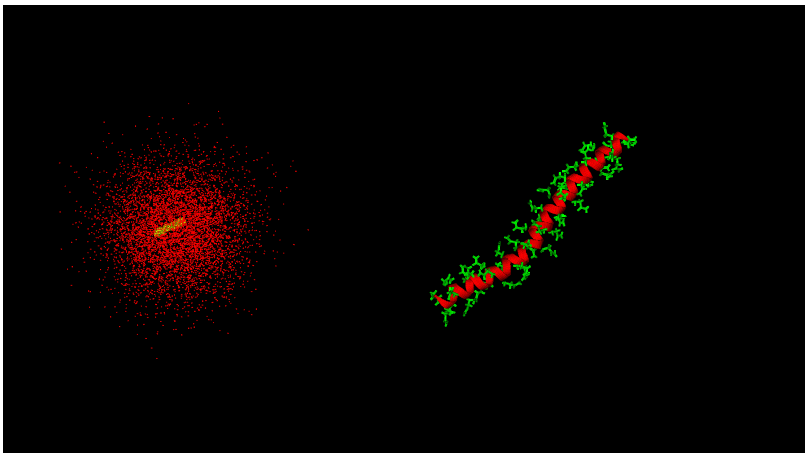

FIGURE 2. **Left figure**: Three dimensional display of the poly-glutamine-water system; the water molecules are in red and the protein is in yellow. **Right figure**: Detail of three dimensional structure of the poly-glutamine helix; the backbone is in red and the Glutamine side chains are in thick green.

3.1. **Classical molecular dynamics simulations.** The two protein-water systems were equilibrated at 300 K in a classical MD simulation, by integrating Eqn.(8) with the AMBER force field [27, 6]. The root mean square deviation per atom (RMSD) with respect to the initial structure is calculated to gauge the structural instability of the protein during the MD simulation and is displayed in figure 3. The larger the RMSD the greater the changes in protein conformation at a given time. Figure 3 shows that, even if the poly-glutamine helix has less than half of the atoms of the prion, its structure changes more with respect to the initial structure than the prion. This is because the initial structure for the poly-glutamine helix is a very open structure, only constrained by the internal hydrogen bonds of the helix, and by the interactions with the water molecules, while the prion structure is maintained by intra residue-residue interactions as well. The general conclusion is, however, that both proteins are structurally stable throughout the simulation. In fact, Figures 1 and 2 display the structures of the two proteins at the end of the MD simulations and confirm that both of them have retained the average structure they had at the beginning.

3.2. **Quantum energy transfer.** While the MD simulations in the previous section did not reveal any signs of the structural instability that prions and poly-glutamine stretches are known to have, here a new possible cause for this instability,

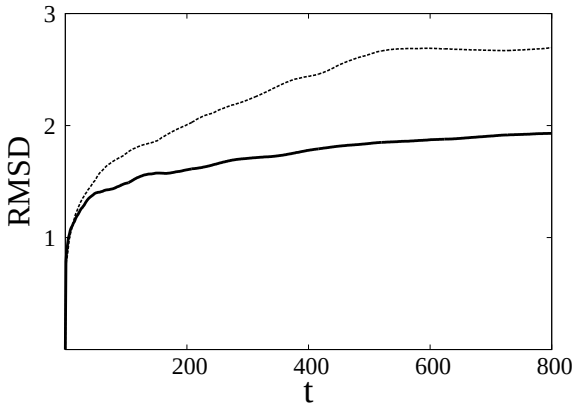

FIGURE 3. Variation of the cumulative RMSD with respect to the corresponding initial structure. The solid line is for the prion and the dashed line is for the poly-glutamine helix. The RMSD values are in Å and time is in picoseconds.

is explored. This new cause is based on the VES hypothesis. While from the point of view of the classical force fields the amino acids GLN and ASN, which occur more frequently in the proteins associated with misfolding diseases, are not very different from many other amino acids, from the point of view of the VES hypothesis they are truly unique: they are *the only two amino acids* that can have amide I excitations also in their side chains. I.e. while all amino acids can have amide I excitations in their main chain (that constitutes the protein backbone), GLN and ASN are the only two amino acids can also have amide I excitations in their side chains [20]. This means that they can divert to their side chains amide I excitations that propagate through the protein backbone and vice versa. Also, because they are polar amino acids, they tend to be on the surface of proteins, in close contact with water, and therefore very well positioned to extract bending mode excitations from the water molecules to their side chains.

In this section, the net increase in the energy transferred from water to the protein, due to the presence of GLN and ASN is calculated in the following manner. We start from an initial condition corresponding to a bending mode excitation in a water close to the protein and, using eqns.(7 and 8), calculate the probability, $P_P = \sum_{\text{amide I sites}} \mid \varphi_r \mid^2$, that, after a time t, this excitation has been transferred to the protein; and we calculate this probability P_P in the presence of GLN and ASN and in the absence of those two amino acids (what is meant by absence here is that the GLN and ASN side chain sites for amide I excitation are not taken into account in the calculation of the quantum states, leaving only the backbone sites); the ratio between these two probabilities is the net enhancement in the energy transferred from the water to the protein due to the presence of GLN and ASN. If this ratio is greater than 1 there is *more* energy transferred from the water to protein in the presence of GLN and ASN and if it is equal to 1 there is no enhancement.

To calculate the net increase of energy transferred to water (if it takes place), we can start instead with an initial quantum state corresponding to an amide I excitation in the protein and calculate the probability, $P_W = \sum_{\text{waters r}} \mid \varphi_r \mid^2$, that, after a time t, this excitation has become a bending mode excitation of a water

molecule; and we can calculate this probability P_W in the presence of GLN and ASN and in the absence of those two amino acids and make the ratio between these two probabilities, which is now the net enhancement in the energy transferred from the protein to the water.

The two probability ratios, calculated for the prion and for the poly-glutamine helix, are displayed in figure 4. The thin solid line corresponds to the probability ratio for energy transfer from the protein to the water and is 1 at all times, for both protein-water systems (there is only one thin line in the figure because they are superimposed), i.e., there is *no* enhancement in the transfer of energy from the protein to the water for both the prion and the polyglutamine helix. On the other hand, the presence of GLN amino acids leads to a strong enhancement in the transfer of energy from the water to the protein in the case of the polyglutamine helix (dashed line in figure 4). Indeed, the amount of energy transferred from water to the polyglutamine helix is 15 times greater in the first picosecond, when GLN amino acids are present. Because the poly-glutamine helix is a very open structure, in which the GLN side chains are in permanent contact with water, the enhancement factor later decreases, but it is still around 3 after 5 picoseconds. The presence of GLN and ASN amino acids also leads to an enhancement of the energy transferred from the water to the prion (solid line in figure 4), but not as markedly as for the poly-glutamine helix.

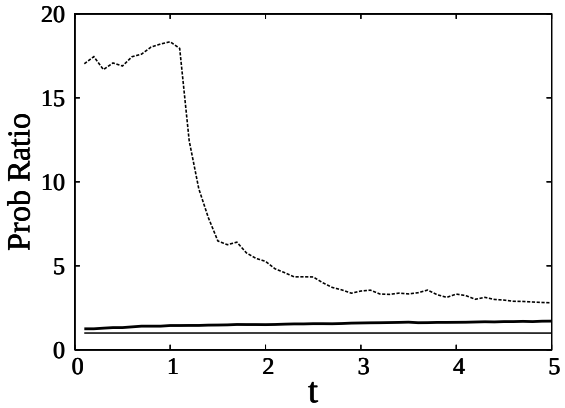

FIGURE 4. Ratio of the probability of a vibrational excitation to jump from the water to the protein (in the presence and in the absence of GLN and ASN), for the prion (solid line) and for the poly-glutamine helix (dashed line), and from the protein to the water in both proteins (thin line) (see text). Time is in ps.

3.3. **Viability of quantum energy transfer.** The new cause proposed here for the structural instability of the prion and of the poly-glutamine helix relies on the existence of a population of bending mode excitations in the waters that surround proteins. Since the bending mode (and the amide I mode) has an energy that is approximately 8 times larger than k_BT, for $T = 300$ K, it may be thought that a bending mode excitation is too rare an event to make this transfer of energy a viable cause for the structural instability of the prion and of mutated huntingtin. In general, proteins last around 48 hours in cells, after which period they are broken

back into amino acids and recycled. Let us then estimate the number of the bending mode excitations that arise, on average, around a protein, during 48 hours.

Let P be the probability of a bending mode excitation in 1 water molecule, when that water molecule is in equilibrium at a temperature T. If, on average, there are N_W hydration waters around the protein, then the average number of bending mode excitations around a protein, at any instant, when the system is in equilibrium, is $N_1 = P\,N_W$. Now, vibrational excitations have finite lifetimes, normally of the order of a few picoseconds. To maintain the average number of bending mode excitations that must be present at thermal equilibrium, new bending mode excitations must be constantly created to compensate for those which are annihilated. The number of bending mode excitations that are created per second can be estimated from the expression $B\,\rho_\nu N_0$, where B is Einstein's absorption coefficient, $\rho_\nu = \frac{8\pi h\nu^3}{c^3}\,\frac{1}{\exp{(h\nu/k_B T)}-1}$ is Planck's blackbody frequency distribution and N_0 is the number of water molecules in the ground state [5]. Since the rate of spontaneous decay, A, is related to B by $A = \frac{8\pi h\nu^3}{c^3}B$ and the population of water molecules in the ground state, N_0, and the population of water molecules with one bending mode excitation, N_1, is related by $N_1/N_0 = \exp\left(-\frac{\epsilon_w}{k_B\,T}\right)$, the rate of creation of bending mode excitations is given by $N_1\,A\,\frac{\exp\left(\frac{\epsilon_w}{k_B\,T}\right)}{\exp\left(\frac{\epsilon_w}{k_B\,T}\right)-1}$, where $A = 1/\tau$, τ being the lifetime of the bending mode vibration.

In order to determine the probability P for a bending mode excitation in the protein-water systems we considered here we need to know the partition function for the excitations described by the Hamiltonian (1-5), which is a very difficult task indeed. Instead, to get an estimate of the order of magnitude for this probability, let us consider the bending mode as an isolated harmonic oscillator. In this case, the partition function is $Z = \exp\left(-\frac{\epsilon_w}{2\,k_B\,T}\right) / \left(1-\exp\left(-\frac{\epsilon_w}{k_B\,T}\right)\right)$ and the probability for a water molecule to have Q quanta of excitation, $P(Q) = \frac{\exp\left(-\frac{(Q+1/2)\,\epsilon_w}{k_B\,T}\right)}{Z} = \exp\left(-\frac{Q\epsilon_w}{k_B\,T}\right)\left(1-\exp\left(-\frac{\epsilon_w}{k_B\,T}\right)\right)$. Since the energy of a bending mode excitation is approximately 1640 cm^{-1} [35], at $T = 300$ K, the probability that one water molecule has a bending mode excitation is $P \approx 0.00038$. The number of water molecules less than 4 Å away from a C=O group in the protein, calculated from the snapshots of the MD simulations, is 238 for the prion and 300 for the polyglutamine helix. Using an average lifetime of 5 ps for the bending mode of water we calculate that, on average, in 1 s, some 10^{+11} bending mode excitations arise close to a protein and in the 46 hour period that proteins generally last in cells, some 10^{16} bending mode excitations are created close to a protein! Even if this number was calculated after a number of approximations the conclusion is that the mechanism for the structural destabilization of proteins proposed here is certainly viable!

4. **Discussion.** The starting point of this study is the fact that the triggers of protein conformational changes are quantum events, and thus quantum states. This is irrefutable when the triggers are chemical reactions since chemical reactions are quantum processes and the immediate outcome of a quantum process is a quantum state. The crucial question then is what form these quantum states take. According to the VES hypothesis [7, 9] adopted in this study, the quantum states are vibrational excited states and according to the Davydov/Scott model [12, 34], which here

has been generalized to take into account the full atomic composition of a protein-water system, the quantum states are amide I excitations. The ultimate aim of the calculations presented in this paper is to show that the VES hypothesis can explain the greater structural instability of the proteins associated with misfolding diseases.

Two curious facts about the proteins associated with misfolding diseases are first, that they have stretches rich in two particular amino acids, namely, GLN and ASN [18] and secondly that such stretches enhance the propensity for protein aggregation [24]. The two proteins investigated here, namely the prion and the polyglutamine helix, are examples of proteins that misfold and aggregate [29, 30, 32, 28, 21]. Although the MD simulations presented in section 3 are too short to make definitive conclusions about the structural stability of the two proteins considered, the fact of the matter is that classical potentials, such as AMBER, have difficulty in explaining why GLN and ASN should make a protein structurally unstable since the force fields associated with these two amino acids are similar to those of many other amino acids [27, 6] that lead to structurally stable proteins. On the other hand, as is pointed out in section 3, from the point of view of the VES hypothesis, GLN and ASN are distinct from all other amino acids in that they are *the only two amino acids that can have amide I excitations in their side chains* and thus they are the only two amino acids that can interfere with the amide I propagation that can take place along the backbone of all proteins [7, 9, 8].

In section 3 the extra energy that can pass from the bending mode of water to the amide I vibration of the protein because of the presence of GLN and ASN was calculated for the prion and for the polyglutamine helix and, in both cases, an enhancement was found. According to the VES hypothesis vibrational excited states are the triggers of conformational changes and proteins which have greater amounts of energy in that form will have a stronger probability of undergoing a conformational change. The enhancement found can therefore explain why proteins with greater amounts of GLN and ASN are structurally less stable than normal proteins.

Figure 4 also shows that the enhancement in the energy transferred from the water to the protein is more pronounced for the polyglutamnine helix than for the prion. This difference may explain why Huntington's disease, which is associated with extra stretches of GLN amino acids added to the protein huntingtin, occurs in humans at a much younger age (around 20 years of age or younger) than prion diseases, which tend to arise in elderly people.

Finally, the calculations in section 3.3 indicate that the number of bending mode excitations created by thermal fluctuations alone make the mechanism proposed here for the structural de-stabilization of proteins certainly viable. Although the estimated number of bending mode excitations that can be created around one protein during its existence in cells was arrived at after a number of approximations, its magnitude (10^{16}!) suggests that, not only is the mechanism for protein destabilization proposed here certainly viable, but also that proteins and cells may well have to defend themselves against it. From this point of view, it is interesting that the proteins of thermophiles (microorganisms that live in thermal vents at the bottom of the ocean, where the temperature reaches the boiling point of water) have virtually no GLN- or ASN-rich regions. Although a quantitative investigation of the influence of temperature on this energy transfer mechanism has not yet been performed, it is possible to anticipate that its efficiency increases with temperature and that, at the boiling point of water, even normal proteins, with smaller amounts

of GLNs and ASNs, may become structurally unstable. And, to end with a prediction, according to the VES hypothesis, ice worms, that live in ice environments and melt at 4 C, must have proteins with greater amounts of GLN and ASN, in order to function properly. This can be proved or disproved by an analysis of the genome of the ice worm.

Acknowledgments. This work was funded in part by the Foundation for Science and Technology (FCT, Portugal) and by POCI 2010 and the European Community fund, FEDER. Part of the computer simulations were performed at the Laboratory for Advanced Computing, (LCA), University of Coimbra, Portugal.

REFERENCES

[1] J. Abrahams, A. Leslie, R. Lutter and J. Walker, *Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria*, Nature, **370** (1994), 621–628.
[2] P. W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev., **109** (1958), 1492–1505.
[3] H. C. Berg, *The rotary motor of bacterial flagella*, Annu. Rev. Biochem., **72** (2003), 19–54.
[4] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *The protein data bank*, Nuc. Acids Res., **28** (2000), 235–242.
[5] P. F. Bernath, "Spectra of Atoms and Molecules," 1^{st} edition, Oxford University Press, New York, Oxford, 1995.
[6] D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. III Cheatham, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner and P. A. Kollman, AMBER 6 (software), University of California, San Francisco, 1999.
[7] L. Cruzeiro, *Why are proteins with glutamine- and asparagine-rich regions associated with protein misfolding diseases?*, J. Phys.: Condens. Matter, **17** (2005), 7833–7844.
[8] L. Cruzeiro, *Influence of the nonlinearity and dipole strength on the amide I band of protein α-helices*, J. Chem. Phys., **123** (2005), 234909-1–7.
[9] L. Cruzeiro, *Protein's multi-funnel energy landscape and misfolding diseases*, J. Phys. Org. Chem., **21** (2008), 549–554.
[10] L. Cruzeiro-Hansson, *Dynamics of a mixed quantum-classical system at finite temperature*, Europhys. Lett., **33** (1996), 655–659.
[11] L. Cruzeiro-Hansson and S. Takeno, *Davydov model: The quantum, quantum-classical, and full classical model*, Phys. Rev. E, **56** (1997), 894–906.
[12] A. S. Davydov, "Solitons in Molecular Systems," 2^{nd} edition, Kluwer Academic Publ., Dordrecht, 1991.
[13] C. M. Dobson, *Protein folding and misfolding*, Nature, **426** (2003), 884–890.
[14] J. Edler and P. Hamm, *Self-trapping of the amide I band in a peptide model crystal*, J. Chem. Phys., **117** (2002), 2415–2424.
[15] J. C. Eilbeck, P. S. Lomdahl and A. C. Scott, *Soliton structure in crystalline acetanilide*, Phys. Rev. B, **30** (1984), 4703–4712.
[16] H. Feddersen, *Localization of vibrational-energy in globular protein*, Phys. Lett. A, **154** (1991), 391–395.
[17] M. Gerstein, A. M. Lesk and C. Chothia, *Structural mechanisms for domain movements in proteins*, Biochemistry, **33** (1994), 6739–6749.
[18] J. F. Gusella and M. E. Macdonald, *Molecular genetics: Unmasking polyglutamine triggers in neurodegenerative disease*, Nature Rev. Neurosci., **1** (2000), 109–115.
[19] J. D. Jackson, "Classical Electrodynamics," 2^{nd} edition, John Wiley & Sons, Inc., New York-Toronto, 1962.
[20] S. Krimm and J. Bandekar, *Vibrational Spectroscopy and conformation of peptides, polypeptides and proteins*, Adv. Prot. Chem., **38** (1986), 181–364.
[21] L. Masino and A. Pastore, *Glutamine repeats: Structural hypotheses and neurodegeneration*, Biochem. Soc. Trans., **30** (2002), 548–551.
[22] F. Mauri, R. Car and E. Tosatti, *Canonical statistical averages of coupled quantum-classical systems*, Europhys. Lett., **24** (1993), 431–436.
[23] C. W. F. McClare, *Resonance in bioenergetics*, Ann. N. Y. Acad. Sci., **227** (1974), 74–97.

[24] M. D. Michelitsch and J. S. Weissman, *A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions*, Proc. Natl. Acad. Sci. USA, **97** (2000), 11910–11915.
[25] D. Narzi, I. Daidone, A. Amadei and A. Di Nola, *Protein folding pathways revealed by essential dynamics sampling*, J. Chem. Theory Comput., **4** (2008), 1940–1948.
[26] N. A. Nevskaya and Yu. N. Chirgadze, *Infrared spectra and resonance interactions of amide-I and II vibrations of α-helix*, Biopolymers, **15** (1976), 637–648.
[27] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. III Cheatham, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman, *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*, Comp. Phys. Commun., **91** (1995), 1–41.
[28] M. F. Perutz and A. H. Windle, *Cause of neural death in neurodegenerative diseases attributable to expansion of glutamine repeats*, Nature, **143** (2001), 143–144.
[29] S. B. Prusiner, *Novel proteinaceous infectious particles cause scrapie*, Science, **216** (1982), 136–144.
[30] S. B. Prusiner, *Molecular biology of prion diseases*, Science, **252** (1991), 1515–1522.
[31] S. B. Prusiner, *Molecular biology and pathogenesis of prion diseases*, TIBS, **21** (1996), 482–487.
[32] S. B. Prusiner, *Prion diseases and the BSE crisis*, Science, **278** (1997), 245–251.
[33] J. Schlitter, M. Engels and P. Kruger, *Targeted molecular dynamics: a new approach for searching pathways of conformational transitions*, J. Mol. Graphics, **12** (1994), 84–89.
[34] A. Scott, *Davydov's soliton*, Phys. Rep., **217** (1992), 1–67.
[35] G. Sieler and R. Schweitzer-Stenner, *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** (1997), 1720–1726.
[36] M. Tirion, *Large amplitude elastic motions in proteins from a single parameter, atomic analysis*, Phys. Rev. Letters, **77** (1996), 1905–1908.
[37] R. Zahn, A. Liu, T. Luhrs, R. Riek, C. Von Schroetter, F. L. Garcia, M. Billeter, L. Calzolai, G. Wider and K. Wuthrich, *NMR solution structure of the human prion protein*, Proc. Nat. Acad. Sci. USA, **97** (2000), 145–150.

Received September 2009; revised October 2009.

E-mail address: lhansson@ualg.pt