A Meanfield Approach to the Thermodynamics of a Protein-Solvent System with Application to the Oligomerization of the Tumor Suppressor p53

Jaan Noolandi

Stanford University School of Medicine
300 Pasteur Drive, Boswell Bldg.
Stanford, CA 04305-5308 U.S.A.

Summary: The thermodynamic stability and oligomerization status of the tumor suppressor p53 tetramerization domain have been studied experimentally and theoretically. A series of hydrophilic mutations at Met-340 and Leu-344 of human p53 were designed to disrupt the hydrophobic dimer-dimer interface of the tetrameric oligomerization domain of p53. Meanfield calculations of the free energy of the solvated mutants as a function of interdimer distance were compared with experimental data on the thermal stability and oligomeric state [tetramer, dimer, or equilibrium mixture of both] of each mutant. The calculations predicted a decreasing stability and oligomeric state for the following amino acids at residue 340: Met [tetramer] > Ser Asp, His, Gln, > Glu, Lys [dimer], whereas the experimental results showed the following order: Met [tetramer] > Ser > Gln > His, Lys > Asp, Glu [dimers]. For residue 344, the calculated trend was Leu [tetramer] > Ala > Arg, Gln, Lys [dimer], and the experimental trend was Leu [tetramer] > Ala, Arg, Gln, Lys [dimer]. The discrepancy for the lysine side chain at residue 340 is attributed to the dual nature of lysine, both hydrophobic and charged. The incorrect prediction of stability of the mutant with Asp at residue 340 is attributed to the fact that within the meanfield approach, we use the wild-type backbone configuration for all mutants, but low melting temperatures suggest a softening of the α-helices at the dimer-dimer interface. This initial application of meanfield theory toward a protein-solvent system is encouraging for the application of the theoretical model to more complex systems.

Keywords: computer modeling; meanfield theory; p53; proteins; protein-solvent interactions

 CCC 1022-1360/00/$ 17,50+.50/0

Introduction

We present a meanfield theoretical approach for studying protein-solvent interactions[1,2]. Starting with the partition function of the system, we develop a field theory by introducing densities for the different components of the system. At this point, protein-solvent interactions are introduced following the inhomogeneous Flory-Huggins model for polymers. Finally, we calculate the free energy in a meanfield approximation. We apply this method to study the stability of the tetramerization domain of the tumor suppressor protein p53 when subjected to site-directed mutagenesis. The four chains of this protein are held together by hydrophobic interactions, and some mutations can weaken this bond while preserving the secondary structure of the single protein chains. p53 functions as a transcription factor that regulates genes involved in the control of the cell cycle in response to DNA damage[3,4]. In approximately half of all human cancers, it exists in a dysfunctional mutant form, but it can also be inactivated by the products of cellular or viral oncogenes. p53 normally functions as a tetramer held together by a short sequence of 30 amino acids near its C terminus. Oligomerization deficient mutants of p53 are also defective in many of p53's important biochemical activities[5,6], and the tetramerization domain of p53 can also help mutant inactive p53 to sequester active wild-type (WT) p53 into inactive heterotetramers[7,8]. Thus, knowledge of the structure and stability of the tetramerization domain is important for an understanding of normal and abnormal function of p53.

Theory

We coarse grain the protein by introducing two effective atoms for each amino acid: one at the C_α position representing the backbone, the other at the side chain center-of-mass position representing the side chain and its properties (e.g., hydrophobicity) [9,10]. Instead of confining the coarse-grained protein model to a discrete lattice, we use a continuum approach where the coordinates of each of the effective atoms as well as the solvent are represented by a density distribution. Applying a meanfield model to the system allows us to express the free energy of the system in a simple analytic form[11].

The free energy in the meanfield approximation is given by using the polymer-solvent system analogue[12],

$$\frac{F}{k_B t} = \sum_{n=1}^{N} \int dx \rho_n(x) \left[\ln \rho_n(x) + \frac{\omega_n(x) + W_n(x)}{2} \right] + C \qquad [1]$$

$$+ \sum_{n=1}^{N} \int dx \phi_n(x) \phi_S(x) + \rho_{0S} \int dx \phi_S(x) \ln \phi_S(x)$$

where the density distributions $\rho_n(\mathbf{x})$, the meanfields $\omega_n(\mathbf{x})$, and the volume fractions $\Phi_n(\mathbf{x})$ are defined

$$\rho_n(x) = e^{-\omega_n(x)} \Big/ \int dx e^{-\omega_n(x)},$$

$$\omega_n(x) = \sum_{n=1}^{N} \int dx' V_{nm}(x - x') \rho_n(x') + W_n(x),$$

$$\phi_n(x) = \int dy K_n(x - y) \rho_n(y), \qquad [2]$$

and C is a constant representing the kinetic energies of all of the particles in the system, and ρ_{0S} is the solvent bulk density. The kernel $K_n(\mathbf{x})$ is defined as

$$K_n(x) = \begin{cases} K_{norm}, & x \le R_n \\ 0, & else. \end{cases} \qquad [3]$$

The densities $\rho_n(\mathbf{x})$ describe the possible fluctuations of the respective effective atoms around their zero-temperature position. The mapping from these densities $\rho_n(\mathbf{x})$ to the volume fractions $\Phi_n(\mathbf{x})$ projects these fluctuations onto the outer shell of the effective atoms.

So far we have included no bond length and bond angle constraints. In a first approximation, we assume that the center-of-mass position of each amino acid side chain n remains at the same position R_{0n} (relative to its neighbors) as in the wild-type (WT) configuration (whereas its

distribution function may change with the separation of the two dimers from each other), i.e., we add to Eq. [2] the constraint

$$\int dx x \rho(x) = R_{0,n}. \qquad [4]$$

Because the side chains of the mutant amino acids differ in size from the WT amino acid, we must relax this constraint. However, we know that the side chains can exist only in a finite number of conformations (the so-called rotamers or rotameric states) because of geometric constraints[13,14]. Using tabulated torsion angles for the amino acid side chains [15,16] and well-known values for the bond lengths and angles for each of the 20 amino acids, we construct a table of center-of-mass positions for all the rotamers and include a proper bias function into the free energy functional Eq. [1] [12] for each of the mutant amino acids.

The side chain volume fraction $\Phi_n(\mathbf{x})$ is obtained by mapping the $\rho_n(\mathbf{x})$ onto a sphere just big enough to fit in side chain n. Because of the incompressibility of the system, i.e.,

$$\sum_{n=1}^{N} \phi_n(x) + \phi_S(x) = 1, \qquad [5]$$

The protein-solvent interactions in Eq. [2] yield nonzero contributions only where $\Phi_n(\mathbf{x})$ and $\Phi_S(\mathbf{x})$ overlap, i.e., at their mutual interface.

In our model, the interactions between the solvent and the amino acid side chains are characterized by parameters χ_n, Eq. [2] , which describe the hydrophobic character of amino acid n. Hydrophobic indices $\chi_F(n)$ have been estimated by Fauchère *et a.* [17] by comparing the distribution coefficients of amino acid amides in water and n-octanol and fixing the absolute scale such that χ_F [Gly] = 0. Eisenberg and McLachlan[18] used these values to estimate solvation free energies $\chi_E = 2.3\ RT\chi_F$ [kcal/mol]. For application of this method to p53tet, we are interested mainly in the change in the free energy during exposure of the hydrophobic dimer-dimer interface to the surrounding solvent. Therefore, we have included only the interface amino acid side chains in the meanfield calculations, while keeping all other amino acid side chains (as well as all

backbone C_αs) fixed at their experimentally observed positions. Free energies are then calculated for different distances between the two dimers when separated along the normal of their mutual mirror plane. Starting from an initial guess for the densities $\rho_n(\mathbf{x})$ and meanfields $\omega_n(\mathbf{x})$, we solve Eq.(2) self consistently up to a desired accuracy.

Results

Fig. 2 shows the calculated relative free energies of the p53tet mutants as function of the dimer-dimer distance obtained from the meanfield calculations. The relative inter-dimer distance measures the separation of the two dimers along the axis perpendicular to the dimer-dimer interface [see Fig.1]. A separation of zero (d = 0) corresponds to the experimentally observed dimer-dimer distance for WT p53.

Fig. 2 [A] shows results of free energy calculations for mutations at site 340. The WT structure shows a pronounced minimum in free energy in agreement with the observed stable tetramer. All of the mutations at site 340 also show a free energy minimum near the WT dimer-dimer distance, although their binding energy is consistently smaller (up to 50%) than that of the WT.

Fig. 2 [B] shows similar free energy results for mutations at site 344. The calculations suggest that all of the mutations considered are less stable than the WT protein. Only the mutant with Ala has a well-defined minimum in the free energy that is comparable to the weakest bound mutants at site 340. For the mutants with Arg, Gln, and Lys, the free energy decreases with increasing dimer-dimer distance. The calculations predict a decreasing stability of the tetrameric state for the following amino acids at residue 340: Met (tetramer) > Ser, Asp, His, Gln > Glu, Lys, whereas the experimental results gave the following order: Met(tetramer) > Ser > Gln > His, Lys > Asp, Glu (dimers). For residue 344, the calculated trend was Leu (tetramer) > Ala > Gln, Arg, Lys (dimer), and the experimental trend was Leu (tetramer) > Ala, Arg, Gln, Lys (dimer). It appears that the inclusion of protein-solvent interactions, which play an important role in the tetramer-dimer dissociation, compensate for the coarse-grained description of the protein structure used in the meanfield calculations.

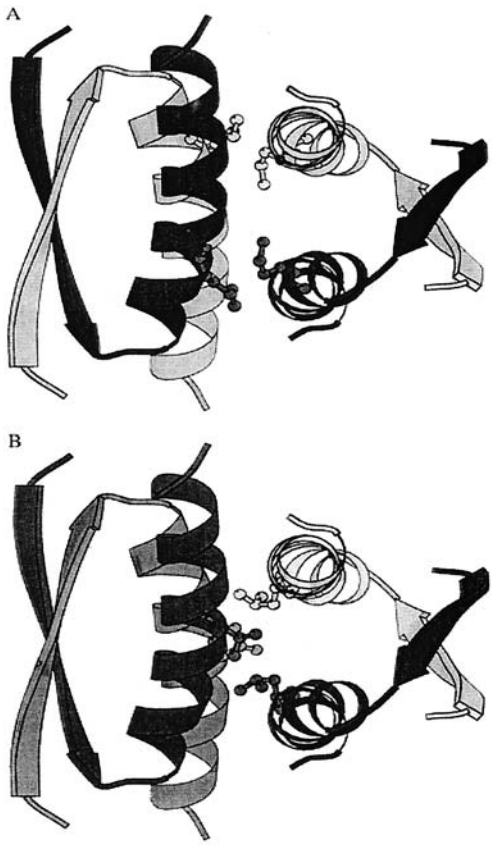

Figure 1. p53 tetramerization domain showing association of two primary dimers; black/gray. Amino acids targeted for mutagenesis at the dimer-dimer interface: [A] Met-340 [B] Leu-344.

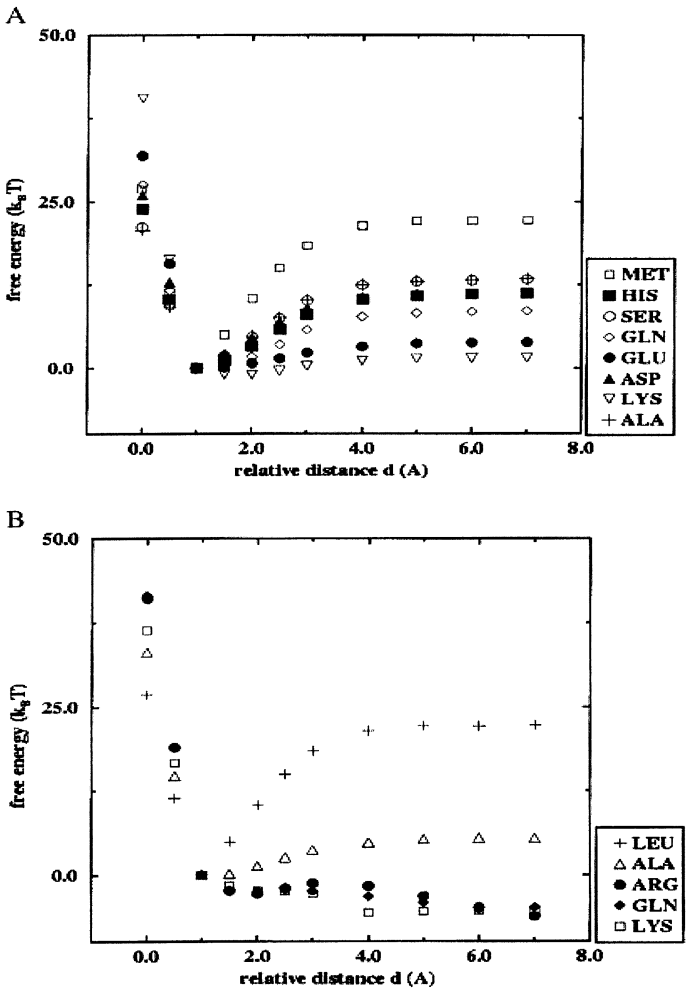

Figure 2. Calculated free energies as a function of the relative dimer-dimer distance for p53tet mutants: [A] mutants at site 340; [B] mutants at site 344.

The only major discrepancy between the simulation results and the experiments involves the Asp mutation at side 340: whereas the experiments show that this mutant dissociates into dimers, the meanfield calculation predicts a stable tetrameric configuration. However, for the numerical calculations, we assume that all mutants retain the WT backbone structure, which might no longer be true for this mutant as indicated by its low melting temperature and molecular dynamics calculations. We also find a discrepancy between the theoretical results and the experiments involving the Lys mutation at residue 340, which is the least stable mutant according to the calculations. This may be because lysine has a rather long and flexible side chain with both hydrophobic and hydrophilic character.

Conclusions

We have introduced a method to calculate protein-solvent interactions based on a meanfield approach to the partition function of the system. An advantage of this approach is the introduction of fields for the different residues of the system. This reduces the many solvent molecules needed for an atomistic calculation to a single field and allows the use of a Flory-Huggins type of contact interaction between the protein and the solvent. We have applied this method to study the stability of the tetramerization domain of the tumor suppressor p53 when subjected to single-point mutations at sites 340 or 344. Here, the secondary structure of the protein survives the mutation, but not necessarily the quaternary structure, leading to a dissociation of the protein into two dimers. This dissociation is mainly driven by the amino acid-solvent interactions, which prefer an exposed dimer-dimer interface for the more hydrophilic amino acids at sites 340 or 344.

Comparison with recent experiments on the stability of the p53 tetramerization domain shows that our meanfield representation of a solvent-protein system can predict the major features of a small self-associating protein system with reasonable accuracy.

[1] J. Noolandi, T.S.Davison, A.R.Volkel, X.-F. Nie, C. Kay, C.H. Arrowsmith *Proc.Natl.Acad.Sci. USA* **2000**, *97*, 9955.
[2] A. R. Volkel, J. Noolandi, *Biophys.J.* **2001**, *80*, 1524.
[3] L.J. Ko, C. Prives, *Genes Dev*. **1996,** *10* , 1054.
[4] D.P. Lane, *Nature* **1992**, *35 8*, 15.
[5] M. Tarurina, J.R. Jenkins, *Oncogene* **1993**, *8*, 3165.
[6] J.M. Singerland, J.R. Jenkins, S. Benchimol, *EMBO J* .**1993**, *1 2,* 1029.
[7] T.D. Halazonetis, A.N. Kandil, *EMBO J*. **1993**, *12* , 5057.
[8] J.Milner, E.A. Medcalf, *Cell* **1991**, *65* , 765.
[9] A. Kolinski, J. Skolnick, *J. Chem. Phys*. **1992**, *97*, 9412.
[10] A. Kolinski, A. Godzik, J. Skolnick, *J. Chem. Phys*. **1993**, *98* , 7420.
[11] A. R. Volkel, J. Noolandi, *J. Comput. Aided Mat. Des.* **1997**, *4* , 1.
[12] J. Noolandi, A.-C. Shi, P. Linse, *Macromolecules* **1996**, *29* , 5907.
[13] J. Ponder, F. M. Richards, *J. Mol. Biol.* **1987**, *193* , 775.
[14] R. L. Dunbrack, M. Karplus, *J. Mol. Biol.* **1993**, *230* , 543.
[15] P. Tuffrey, C. Etchbest, S. Hazout, R. Lavery, *J. Biomol. Struct. Dyn*. **1991**, *8* , 1267.
[16] P. Tuffrey, C. Etchbest, S. Hazout, R. Lavery, *J. Comp. Chem*. **1993**, *14* , 790.
[17] J.-L. Fauchere, V. Pliska, *Eur. J. Med. Chem. Chim. Ther.* **1983**, *18* , 369.
[18] D. Eisenberg, A. D. McLachlan, *Nature* **1986**, *319* , 199.