Probing Protein Stability with Non-natural Amino Acids

Nediljko Budisa[a]* and Greta Pifat[b]

*[a]Max-Planck Institut für Biochemie, Abteilung Strukturforschung,
Am Klopferspitz 18a, D-82152 Martinsried, Germany
(e-mail: budisa@biochem.mpg.de);*

*[b]Ruđer Bošković Institute, Bijenička 54, 10001 Zagreb, Croatia
(e-mail: pifat@rudjer.irb.hr)*

Quantitative replacement of methionine with its non-natural amino acid analogues, norleucine, selenomethionine and telluromethionine in human recombinant annexin V, is applied to study conformational and folding properties in solution. This procedure replaces each methionine sulphur atom with Se, Te or -CH_2-, providing single-atom exchanges, or »atomic mutations«. Using guanidine chloride as denaturant, the estimated stabilities of protein variants are not significantly changed. The denaturation midpoints are shifted towards lower values owing to the increase in the hydrophobicity of exchanged residues. Co-operativity expressed in terms of *m*-values is also affected by such exchanges and is highly correlated with the physical properties of methionine and its analogues in solution. This approach can contribute to a detailed understanding of interactions of particular amino acids and their implications on protein folding and protein-ligand interactions.

INTRODUCTION

Protein structures are stabilized by several factors, such as van der Waals interactions, hydrogen bonds, hydrophobic interactions, effect of hydration on non-polar groups and salt bridges.[1] A major class of proteins, globular

---

* Author to whom correspondence should be addressed.

proteins, fold into compact and unique configurations of the native (folded) state. The problem in protein folding is to understand the small net effects of large and diverse driving forces.

Over the past thirty years, intensive research was carried out on studies of detailed microscopic descriptions of the energetics and mechanisms of the folding transition process using site-directed mutagenesis. However, this approach usually produces large and undefined (complicated) local changes with unpredictable consequences on the protein stability and folding, thus making the quantification of detected changes difficult.

The importance of systematic replacement of methionine (Met) in proteins by its isosteric analogues was recently recognised by X-ray crystallographers.[2] In addition, similar strategy can be used in the folding studies of proteins. For example, Met isosteric analogues, such as selenomethionine (SeMet), telluromethionine (TeMet) or norleucine (Nle), have the same chain length and the same number of single rotatable bonds (Figure 1). Proteins substituted in this way have all Met sulphur atoms replaced by Se, Te or -CH_2- providing single-atom exchanges (»atomic mutations«). The advantage of an »atomic mutation« approach is that the deviations from the canonical values for bond lengths and angles are minimal. Such fine alterations can be a suitable model for precise studies of various interactions in proteins.
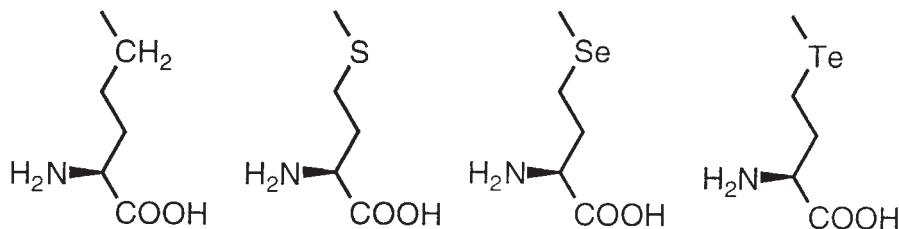

Figure 1. Natural amino acid methionine and its non-natural analogues which can be incorporated into proteins. The analogues in the series Nle → Met → SeMet → TeMet exhibit an increase in X-C (X = CH_2, S, Se, Te) bond length (1.5 Å → 1.8 Å → 2.1 Å → 2.4 Å). Therefore, the steric effects at the sites of replacement in proteins are plausible. However, they are regarded as isosteric because the replacement is completely isomorphous, as confirmed by X-ray analysis of annexin V and other substituted proteins.[2]

Human recombinant annexin V provides an ideal experimental model to test this approach to protein folding since it was the first protein labelled almost quantitatively with Nle and TeMet.[3] Such substitutions do not affect the biological activity of annexin V, while wt-protein and its variants exhibit

isomorphism in crystalline states.[2] To assess whether their behaviour is different in solution, we made conformational and folding studies of wt (wild-type) and substituted forms of annexin V. It is a typical α-helical protein with four domains consisting of α-helix bundles.[4] In recombinant form, it contains seven Met residues located in different parts of the molecule: surface exposed Met214, partially solvent accessible Met259 and Met299, and in hydrophobic environments Met28, Met80, Met152 and Met273 (Figure 2).
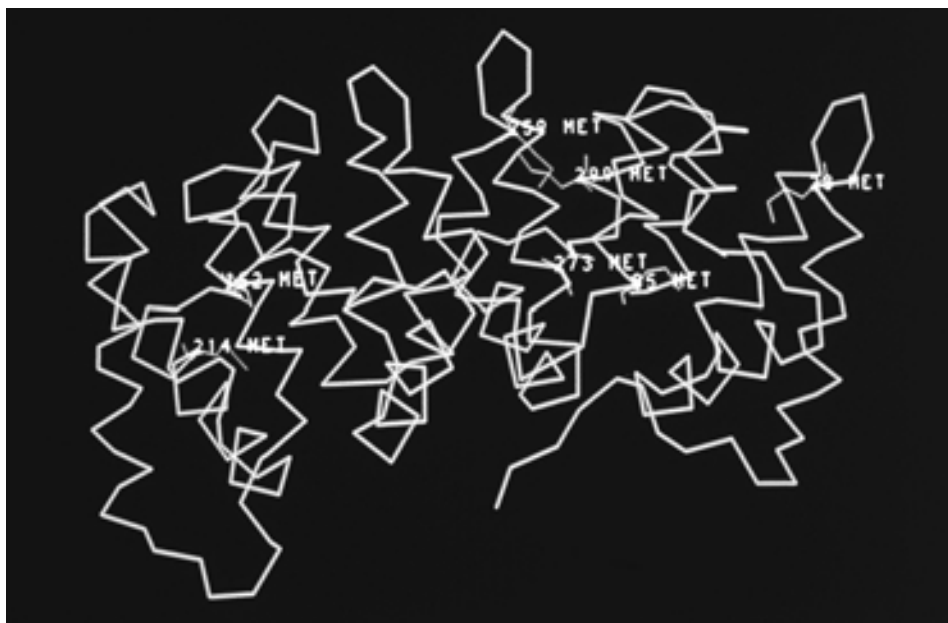

Figure 2. C^{α}-plot of human recombinant annexin V (side view) with distribution of Met-side chains over the structure. The figure was generated by program FRODO[19] using co-ordinates for annexin V deposited in the Brookheaven Protein Data Bank.

EXPERIMENTAL

Recombinant protein expression in *E. coli,* purification as well as the residue--specific replacements of Met Sδ atoms with Se, Te and -CH$_2$- was achieved as described elsewhere.[2,3] Purified protein samples were kept in phosphate buffered saline (PBS) solution[3] (pH 7.4).

Our equilibrium folding studies were performed by using guanidine chloride (GdmCl) as denaturant and assuming a two-state equilibrium model.[5] The data were calculated by the widely used linear free energy model of unfolding, which assumes linear dependence of the free energy in the function of denaturant concen-

trations.[5] The validity of the two-state model for annexin V was demonstrated elsewhere.[6–8] Possible complications of Met and its analogues oxidation are eliminated by the fact that refolded proteins retain their molecular mass, as assessed by mass spectrometry. However, in all our refolding experiments, TeMet-annexin V can be only partially refolded, possibly due to the mentioned high reactivity of the exposed TeMet side chains.[9]

Unfolding is monitored through measurements of a large fluorescence shift upon denaturation, caused by dequenching of the buried single tryptophan residue (W187) in the folded state. An amount (5 µL) of protein of a concentration of 0.2 mM in PBS was incubated overnight at 20 °C with 500 µL guanidinium hydrochloride solutions in PBS. Fluorescence (λ_{ex} = 278 nm, λ_{em}=360 nm) was integrated over a 3 s interval. The obtained values were corrected for buffer background fluorescence. Free energy of unfolding, m-values, pre- and post-unfolding baselines were analyzed using a non-linear least-squares fit procedure. A two-state model was assumed where correlation coefficients typically exceeded 0.999, while errors were consistently below 10%.

TABLE I

Changes in free energy of unfolding caused by substitution of Met with its analogues in annexin V, as determined by reversible GdmCl-induced denaturation followed by fluorescence change measurements. The data were calculated using a linear extrapolation method described in Ref. 5. Each value is an average of five measurements.

Protein	ΔG^0_{N-U}(kcal mol^{-1})[*]	m(kcal mol^{-1} M^{-1})[**]	$[D]_{1/2}$(M)[***]	ΔG^0_{N-U}(kcal mol^{-1})[****]
Wild-type	9.09 ± 2.08	5.41 ± 0.31	1.60 ± 0.03	——
SeMet	8.93 ± 1.39	6.56 ± 0.89	1.35 ± 0.12	– 0.16
Nle	8.96 ± 0.94	6.65 ± 0.67	1.29 ± 0.13	– 0.13
TeMet	7.27 ± 2.35	7.73 ± 0.94	0.90 ± 0.05	– 1.82

* free energy of folding determined by denaturation of proteins in PBS solution (pH 7.4) and linear extrapolation of the data to zero denaturant concentration.

** m-value is the slope of the linear denaturation plot.

*** $[D]_{1/2}$ is the concentration of denaturant at which 50% of the protein sample is unfolded.

**** the difference between of the native and the substituted proteins.

Here we demonstrate that the native annexin V is a highly stable protein with an estimated free energy of unfolding (ΔG^0_{N-U}) 9.09 ± 2.08 kcal mol^{-1}, which is in the range expected for single monomeric proteins.[1] The determined data are summarized in Table I and the curves are presented in Figure 3. The constant m is 5.41 ± 0.31 kcal mol^{-1} M^{-1} with a transition denaturation midpoint ($[D]_{1/2}$) of 1.60 ± 0.03 M. Substitution with SeMet decreases the denaturation midpoint by about 0.25 M, while the stability is not changed with a higher m-value (6.56 ± 0.89 kcal mol^{-1} M^{-1}). The Nle-annexin V exhibits similar behaviour. In fact, these two protein analogues have always very similar denaturation profiles (Figure 3B). The effect of TeMet
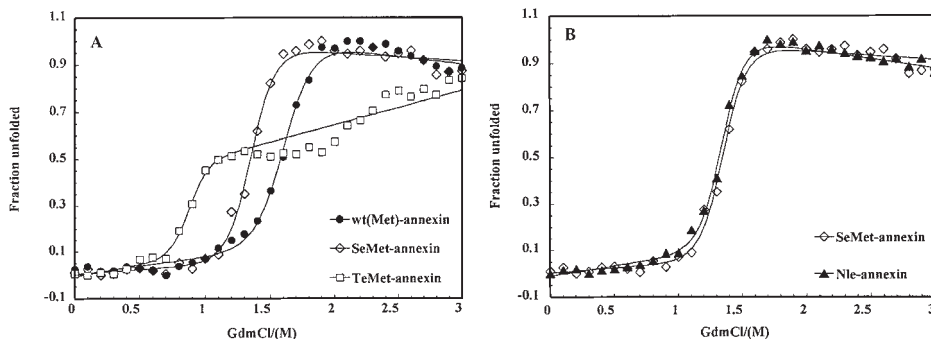
Figure 3. (A) Equilibrium unfolding curves for wt and substituted forms of human recombinant annexin V, plotted as fractions of unfolded protein *versus* concentrations of denaturant. (B) Unfolding profiles for Nle- and SeMet-annexin V forms.

incorporation is more dramatic, resulting in a further decrease in $[D]_{1/2}$ by about 0.7 M ($[D]_{1/2}$ = 0.90 ± 0.05 M). However, the m-value (7.73 ± 0.94 kcal mol^{-1} M^{-1}) is also dramatically increased by about 30% in comparison with that of the wt-protein. Moreover, the TeMet-protein form shows a different denaturation profile. After 2 M GdmCl (Figure 3A), the denatured state is reached with increasing unfolding on further denaturing. This might indicate the presence of certain amounts of residual secondary structure, which has been reported for other proteins as well.[10,11] This might also be the main reason for the estimated lower free energy of unfolding. The shape of this curve could indicate the presence of stable intermediates between the native and denatured states. However, other investigations (thermal and urea denaturation studies) indicate their absence.[6] Although TeMet-annexin V exhibits the estimated lower ΔG^0_{N-U} value, this value is close to the ΔG^0_{N-U} values for wt-, Nle-, and SeMet-annexin V forms (an average 8.56 kcal mol^{-1}) despite the dramatic changes in the denaturation midpoint values.
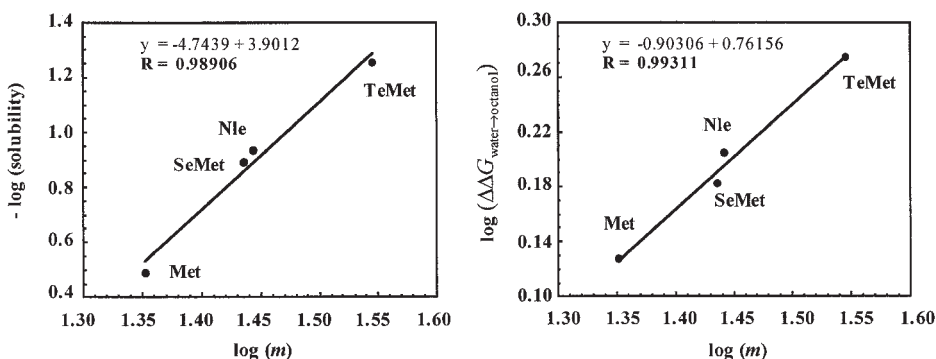

Figure 4. Correlated logarithmic values for solubility in water and free energy of transfer from water to n-octanol ($\Delta\Delta G_{\text{water}\rightarrow\text{octanol}}$) of Met and its analogues with the m-values found in the transition region during denaturation (Table I).

The *m*-values from unfolding curves are highly correlated with solubility and side-chain free energy of transfer from water to *n*-octanol ($\Delta\Delta G_{\text{water}\rightarrow\text{octanol}}$), as shown in Figure 4. Solubility data are (in M): 0.33 ± 0.02 (Met), 0.13 ± 0.01 (SeMet), 0.12 ± 0.01 (Nle), 0.056 ± 0.02 (TeMet). Values of the $\Delta\Delta G_{\text{water}\rightarrow\text{octanol}}$ are (in kcal mol^{-1}) 1.34 ± 0.04 (Met), 1.52 ± 0.14 (SeMet), 1.60 ± 0.01 (Nle), 1.88 ± 0.80 (TeMet). Both, $\Delta\Delta G_{\text{water}\rightarrow\text{octanol}}$ and solubility data were taken from Ref. 6.

DISCUSSION

Met analogues, although isosteric, exhibit some distinct features. The electronegativity of sulphur (2.55) is more similar to that of carbon (2.5) and selenium (2.58) than that of tellurium (2.0). This decrease in electronegativity from carbon to tellurium also includes a longer (and therefore weaker) covalent bond and larger electrovalent and co-ordination spheres. Preference of thiobutane for the *gauche* over *anti* conformation (observed in *n*-butane) in the gas phase was found, and an explanation that the longer S-C *vs.* C-C bond caused less induction of steric repulsions was offered.[12] On the contrary, it was shown that the Cγ-Cδ bond in Nle13 of ribonuclease S′ is also in the *gauche* conformation, leading to the conclusion that other factors than local residue geometry had to be equally or even more important.[13]

Conformational studies by fluorescence spectroscopy revealed small changes in the fluorescence emission of TeMet-annexin V (red shift by 6 nm) while Nle- and SeMet-substituted proteins exhibit no shift in fluorescence emission spectra in comparison with the wt-form (our unpublished material). On the other hand, the analysis by circular dichroism spectroscopy of wt-annexin V and its variants in phosphate buffered solution at room temperature reveals the shape and dichroic intensities typical of α-helical protein.[2]

The *m*-value is supposed to describe the steepness or co-operativity of the transition and was interpreted as a measure of the hydrophobic surface exposed upon unfolding.[5] It is now well established that *m*-values derived from the analysis of GdmCl denaturation are strongly correlated with the change in the accessible surface area upon unfolding.[14] In these terms, it is clear that the slopes of our unfolding curves representing *m*-values increase with increasing hydrophobicity of the exchange in the order: Met → SeMet ≈ Nle → TeMet. Therefore, the increase of *m*-values from our experiments can be also explained in terms of an increase of the specific volumes of atomic mutants (S → Se → Te). Furthermore, in our experiments, *m*-values clearly represent contributions of hydrophobic interactions to the protein overall stability. This also explains the relatively constant value of the estimated $\Delta G_{\text{N–U}}^{0}$ values: the increase in hydrophobicity compensates for the decrease in the denaturation transition midpoint.

However, it was argued that the hydrophobic effect alone seemed to be insufficient to account for the existence of a unique equilibrium structure in proteins since aggregates of oil, which form spontaneously in water, lack a specific internal architecture.[15] Moreover, the high ionic strength of GdmCl at high concentrations was proposed to mask the electrostatic interactions of proteins, measuring the relative contribution of hydrophobic interactions.[16] Our results seem to be in line with such reasoning since the decrease in polarity (S $\rightarrow$ Se $\rightarrow$ Te) does not significantly affect the determined ΔG_{N-U}^0 value.

Analyses of physical parameters, such as solubility and side-chain free energy of transfer from water to n-octanol ($\Delta\Delta G_{water\rightarrow octanol}$) of Met and its analogues, indicate a systematic decrease in polarity and increase in hydrophobicity from Met to TeMet.[6] This dual nature of »atomic mutants« confirms the importance of hydrophobic interactions in our folding experiments. Namely, the folding parameter m-value, which represents the co-operativity of unfolding and changes with systematic variations in atomic number, correlates well (higher than 98%) with the determined physical parameters of Met and its isosteric analogues in solution, as shown in Figure 4. This is an indication that the properties of amino acids determined in solutions should correlate with their properties when integrated and modulated in a highly structured protein molecule.

It has been reported for several other proteins that single mutations can produce large differences in the enthalpy, heat capacity or free energies of denaturation.[13,17,18] This is surprising, since it has been assumed that these mutations do not produce large structural perturbations of the native state.[11] For example, for removal of methylene equivalent, the average loss in protein stability is typically a few times larger than the free energy measured by partitioning between water and n-octanol.[18] Crucial differences between transferring a side chain from water to the interior of the protein were attributed to these observations.[17] It is difficult to make similar comparisons for annexin because (i) it is not clear whether the observed differences are the result of a synergistic effect or not, (ii) the validity of such comparisons is in question because the folding co-operativity is significantly changed in terms of m-values, and (iii) the estimated stability is constant among mutants and therefore not correlated to the changes in solubility and side-chain free energy of transfer from water to n-octanol.

Since the replaced Met residues in annexin V are not uniformly distributed, it is difficult to recognize the particular contribution of each determined parameter without further studies. It is also possible that the observed net changes result from a synergistic effect (*i.e.* that they are not a simple sum of all changes). To address this additivity problem, further studies should combine the »atomic mutation« approach with site-directed mutagenesis techniques.

In conclusion, we have demonstrated the validity of the non-natural amino acid bioincorporation technology for an »atomic mutation« approach to the problem of protein folding. Indeed, the main message is that the subtle and systematic exchanges in amino acid side chains of proteins and their high correlation with the physical properties of the same amino acids in solution might lead to predictions of the folding parameters of highly specific protein structures.

REFERENCES

1. K. A. Dill, *Biochemistry* **29** (1990) 7133–7155.
2. N. Budisa, W. Karnbrock, S. Steinbacher, A. Humm, L Prade, T. Neuefeind, L. Moroder, and R. Huber, *J. Mol. Biol*. **270** (1997) 616–623.
3. N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann, and R. Huber, *Eur. J. Biochem*. **230** (1995) 788–796.
4. S. E. Moss, H. C. Edwards, and M. J. Crumpton, *Novel Calcium-binding Proteins* C. W. Heizmann (Ed.), Springer-Verlag, Berlin, 1992, pp. 536–566.
5. C. N. Pace, *Method. Enzymol*. **131** (1986) 266–280.
6. N. Budisa, Ph. D. Thesis, Technische Universität, München, 1997.
7. S. Liemann, J. Benz, A. Burger, D. Voges, A. Hoffmann, R. Huber, and P. Göttig, *J. Mol. Biol.* **258** (1996) 551–561.
8. T. Vogl, C. Jatzke, H. J. Hinz, J. Benz, and R. Huber, *Biochemistry* **36** (1997) 1657–1668.
9. W. Karnbrock, E. Weyher, N. Budisa, R. Huber, and L. Moroder, *J. Am. Chem. Soc*. **118** (1996) 913–914.
10. D. Shortle, *FASEB J*. **10** (1996) 17–34.
11. J. H. Carra and P. L. Privalov, *FASEB J*. **10** (1996) 67–74.
12. S. Gellman, *Biochemistry* **30** (1991) 6633–6636.
13. J. Thompson, G. S. Ratnaparkhi, R. Varadarajan, J. M. Sturtevant, and F. M. Richards, *Biochemistry* **33** (1994) 8587–8593.
14. J. K. Myers, C. N. Pace, and J. M. Scholtz, *Protein Sci*. **4** (1995) 2138–2148.
15. G. D. Rose and R. Wolfenden, *Annu. Rev. Biophys. Biomol. Struct*. **22** (1993) 381–415.
16. O. D. Monera, C. M. Kay, and R. S. Hodges, *Protein Sci*. **3** (1994) 1984–1991.
17. J. T. Kellis, K. Nyberg, and A. R. Fersht, *Biochemistry* **28** (1989) 4914–4922.
18. D. Shortle, *Quat. Rev. Biophys*. **25** (1992) 205–250.
19. T. A. Jones, *J. Appl. Crystallog.* **15** (1978) 24–31.

S A Ž E T A K

Ispitivanje stabilnosti proteina sa neprirodnim aminokiselinama

Nediljko Budisa i Greta Pifat

Metioninski krajnji ostatci u humanom rekombinantnom aneksinu V zamijenjeni su s neprirodnim aminokiselinskim analozima norleucinom, selenometioninom i telurometioninom u svrhu studija konformacijskih promjena i smatanja proteina u otopini. Pri tom postupku, svaki metioninski sumpor zamijenjen je sa Se, Te ili -CH$_2$-, dajući zamjene do razine pojedinih atoma, tj. »atomske mutacije«. Stabilnost proteinskih varijanti određivana s pomoću denaturanta gvanidinijeva klorida nije se bitno promijenila. Središnje točke prijelaznog područja pomiču se pri denaturaciji prema nižim vrijednostima zbog porasta hidrofobnosti promijenjenih ostataka. Kooperativnost izražena u obliku m-vrijednosti također se mijenja s navedenim promjenama te je visoko korelirana s fizikalnim svojstvima metionina i njegovih analoga u otopini. Taj pristup može pridonijeti potanjem razumijevanju međudjelovanja pojedinih aminokiselina u proteinskoj molekuli kao i njihova utjecaja na smatanje proteina te na međudjelovanja protein-ligand.