ORIGINAL RESEARCH

Mutational Effect of Structural Parameters on Coiled-Coil Stability of Proteins

Amutha Selvaraj Maheshwari[1,2] and Govindaraju Archunan[1]

[1]Department of Animal Science, School of Life Sciences, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.
[2]Department of Biotechnology, Anna University—BIT Campus, Tiruchirappalli, Tamil Nadu, India.
Corresponding author email: archunan@bdu.ac.in, asmaheshwari@yahoo.com

Abstract: Understanding the parameters that influence the melting temperature of coiled-coils (CC) and their stability is very important. We have analyzed 45 CC mutants of DNA binding protein, electron transport protein, hydrolase, oxidoreductase, and transcription factors. Many mutants have been observed at Tm = 40 °C–60 °C with ΔS = 9–11 kcal/°C mol, ΔG = −400 to −450 kcal/mol, and Keq = 0.98–1.03. The multiple regression analysis of Tm reveals that influences of thermodynamic parameters are strong (R = 0.97); chemical parameters are moderate (R = 0.63); and the geometrical parameters are negligible (R = 0.19). The combination of all these three parameters exhibits a little higher influence on Tm (R = 0.98). From the analysis, it has been concluded that the thermodynamic parameters alone are very important in stability studies on protein coil mutants. Besides, the derived regression model would have been useful for the reliable prediction of the melting temperature of coil mutants.

Keywords: coil propensity, compactness, equilibrium constant, melting temperature, solvent accessibility, structural parameters, free energy change, van't Hoff enthalpy

Introduction

Among all the found coiled-coils (CC), the dimeric CC only populates 73.8%, whereas the other 26% are more complex.[1] The architecture of a particular CC domain determines its oligomerization state, rigidity, and ability to function as a molecular recognition system. Much progress has been made towards understanding the parameters that determine CC formation and stability.[2] As per the previous reports, by combining the most favorable inter- and intra-helical salt-bridge arrangements, it is possible to design CC oligomerization domains with improved stability properties.[3] It has been concluded that dynamic changes in the helical registry may be a general property of CC.[4]

The hydrophobic interactions that occur along the CC interface stabilize the α-helical CC, and the hydrophobic clustering stabilizes the CC structure.[5] There is preference for hydrophobic (Ala, Leu), positively- (Lys, Arg), and negatively- (Glu) charged residues in CC domains. The surrounding hydrophobicity of residues in CC domains is significantly less than that of residues in other regions of CC proteins. The residues in CC domains are more stable than those in other regions and are largely influenced by medium-range contacts. The long-range interactions play a dominant role in other regions of protein, as well as in non-CC helices.[6]

Stability measurements indicate that maintenance of all favorable electrostatic interactions, or the avoidance of two potentially repulsive interactions contributes approximately 2.1 kcal/mol to helix orientation preference.[7] The guanidine hydrochloride (GdnHCl) masks electrostatic repulsions due to its ionic nature. In addition, Glu and Gln in the e and g positions of the heptads repeat and have very similar effects on CC stability in the presence of GdnHCl.[8] Buried urea/urea contacts lead to extremely stable dimeric CC at melting temperatures between 63 °C and 79 °C. Core ureas can also form stable complexes with a variety of other polar groups, including guanidine, acids, and amides.[9]

Earlier studies suggest that the classified structural information is more dominant than the experimental conditions for understanding the stability of protein mutants. The comparison of amino acid properties with free-energy terms indicated that the energetic contribution explained the mutant stability better in the coil region.[6] The present work discusses how different parameters influence the melting temperature of protein coil mutants. This study also reveals that whether the degree of influence be large or small, it would have some effect on coil mutant stability at different solvent accessibility.

Materials and Methods
Datasets

The dataset contains different coil mutants (289 data) of 1RN1, 1A23, 1STN, 1RGG, 1CSP, 1HUE, 1FTG, 1RX4, 1RGG, and 1C9O at exposed, partially buried, and buried regions. These proteins belong to the classification of DNA binding protein, electron transport protein, hydrolase, oxidoreductase, and transcription factors. This dataset has been generated from the ProTherm database,[10] available at http://gibk26.bio.kyutech.ac.jp/jouhou/Protherm/protherm.html. The data redundancy check was carried out for reliable result prediction. The experimental data such as Tm, ΔTm, ASA (%), and pH has been obtained with conditions such as: two state proteins, single mutation, coil secondary structures, and thermal denaturation. The availability of parameters like ΔCp, activity, and so on, is insufficient to use in the present data analysis. In addition, the more significant parameters have not been ignored. Hence, the omissions of the abovementioned less significant parameters do not cause much effect on result prediction.

Amount of hydrogen atoms (pH)

The hydrogen atoms are positively charged and they attract the negative side of the polar amino acids. Some amino acids are polar and have positively- or negatively-charged side chains. Thus, the change in pH alters the stability of a protein structure.

Accessible surface area (ASA)

The solvent-accessible surface area, ASA or SASA, is often used for calculating the transfer-free energy, ΔGtr. This transfer-free energy is required to move a biomolecule from an aqueous solvent to a non-polar solvent, such as in a lipid environment. The ASA can be used to improve protein secondary structure prediction.[11,12]

Melting temperature (Tm)

The protein melting point is an important characteristic feature of a protein and is used for various purposes

such as drug development. In general, the protein will be more stable if the Tm is high. As per the previous report, the melting temperature of a protein can be predicted from its amino acid sequence.[13]

Van't Hoff analysis
The van't Hoff equation relates the equilibrium constant with temperature, melting temperature, and standard enthalpy change of the process.

$$Keq = \exp\left[\Delta HvH/R\left((1/Tm) - (1/T)\right)\right] \quad (1)$$

where, Keq—equilibrium constant; ΔHvH—van't Hoff enthalpy; R—gas constant; Tm—melting temperature; and T—temperature.

Chemical thermodynamic relationships
The Tm and ΔTm values could be determined experimentally; T and Keq values can be calculated using equation 2. From equation 3, the change in free energy (ΔG) is calculated. The change in entropy (ΔS) is calculated by equation 4.

$$\Delta Tm = Tm - T \quad (2)$$

$$\Delta G = -RT\ln Keq \quad (3)$$

$$\Delta G = \Delta H - T\Delta S \quad (4)$$

Free energy change (ΔG) and transfer-free energy change (ΔGtr)
The positive ΔG indicates that the native state is more stable than the denatured state. The hydrophobicity scale based on transfer Gibbs energy of residues from the non-aqueous phase to water has been widely used to estimate the conformational stability of proteins. As per the previous report, the molar volume-corrected octanol solubility scale provides the best agreement with changes in protein stability upon mutation.[14] The solubility scale based on octanol to water transfer energies for the amino acids has been used for our analysis.

Compactness (Z)
Compactness is calculated through a solvent-accessible surface area of a fragment (ASAsurf), divided by its minimum possible area. The minimum possible surface area is the surface area of a sphere with an equal volume as that of the fragment,[15] which

is calculated by an integration of all individually exposed solvent-accessible surface areas.[16]

$$Z = ASAsurf/(36 \times 3.14 \times vol2)1/3 \quad (5)$$

The ASA (Å2) values taken from ProTherm database include: Ala-110.2; Asp-144.1; Cys-140.4; Glu-174.7; Phe-200.7; Gly-78.7; His-181.9; Ile-185.0; Lys-205.7; Leu-183.1; Met-200.1; Asn-146.4; Pro-141.9; Gln-178.6; Arg-229.0; Ser-117.2; Thr-138.7; Val-153.7; Trp-240.5; and Tyr-213.7. The classification of amino acids used in this study (as given in the database) is as follows: buried if presenting an ASA with less than 20% accessibility; partially buried if the ASA is between 20% and 50%; and exposed if the ASA is more than 50%.

Coil propensity (Pc)
The propensity of amino acids for coil conformation is also an intrinsic property of amino acids. The amino acid propensities for secondary structures are more reliable depending on the degree of homogeneity of the protein dataset used to evaluate them.[17]

Single correlation (r) and multiple regression (R) analysis
The relationship between melting temperature and mutation-induced property changes, such as ASA, transfer free energy, van't Hoff enthalpy, equilibrium constant, and so on, have been calculated using Pearson's correlation. The multiple regression coefficients for all mutants are computed from all possible combinations of geometrical, chemical, and thermodynamic parameters. In addition, the regression models are derived for each category.

Results and Discussion
Coil mutations
When considering the free energy change due to mutation, many data lie in the range of ASA of 0%–80%, which is shown in Figure 1A. In Figure 1B, many mutants are found in the range of 30 °C–80 °C of melting temperature. The plot between pH and $\Delta\Delta G$ shows that many data are available at pH 2–3, pH 4–8, and near pH 7, and comparatively few data are also available at pH 8–10 (Fig. 1C). The correlations between ASA and $\Delta\Delta G$ seem to be high at
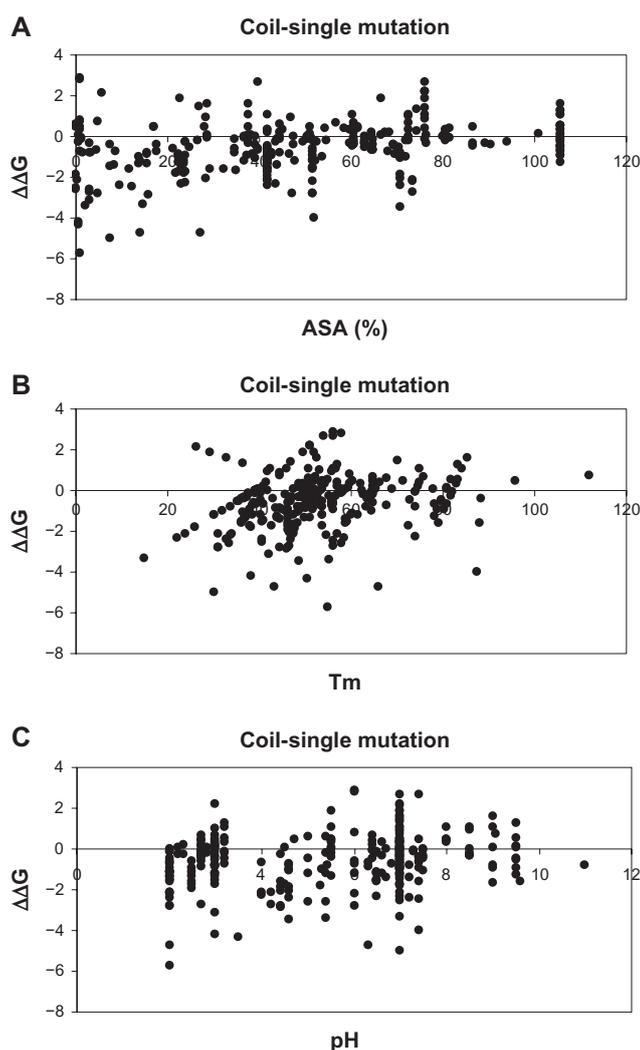
Figure 1 (**A**) Plot depicts the relationship between ASA (%) and free energy change due to various single mutations in the coil region. (**B**) Plot depicts the relationship between Tm and free energy change due to various single mutations in the coil region. (**C**) Plot depicts the relationship between pH and free energy change due to various single mutations in the coil region.

Table 1. Relationship between accessible surface area and free energy change due to mutation at various pH ranges.

r	ASA vs. ΔΔG	Number of data
All pH	0.22	222
pH 1.1–2.0	**0.49**	36
pH 2.1–3.0	**0.53**	69
pH 3.1–4.0	**−0.96**	12
pH 4.1–5.0	−0.26	27
pH 5.1–6.0	0.06	34
pH 6.1–7.0	0.50	116
pH 7.1–8.0	0.01	45
pH 8.1–9.0	0.30	13
pH 9.1–10.0	**0.06**	**13**

Table 2. Illustration of various parametrical relationships with melting temperature.

Parameters	r
Tm vs. ASA	0.34
Tm vs. ΔHvH	−0.17
Tm vs. ΔS	**−0.58**
Tm vs. ΔG	**−0.73**
Tm vs. Keq	**−0.59**
Tm vs. ΔGtr	−0.22
Tm vs. Z	−0.19
Tm vs. Pc	0.15

pH 6–7 (Table 1). In addition, considerable amounts of van't Hoff enthalpy values are also available for pH 7, so the relationship between Tm and ΔHvH at pH 7 at different ranges of solvent accessibility is analyzed.
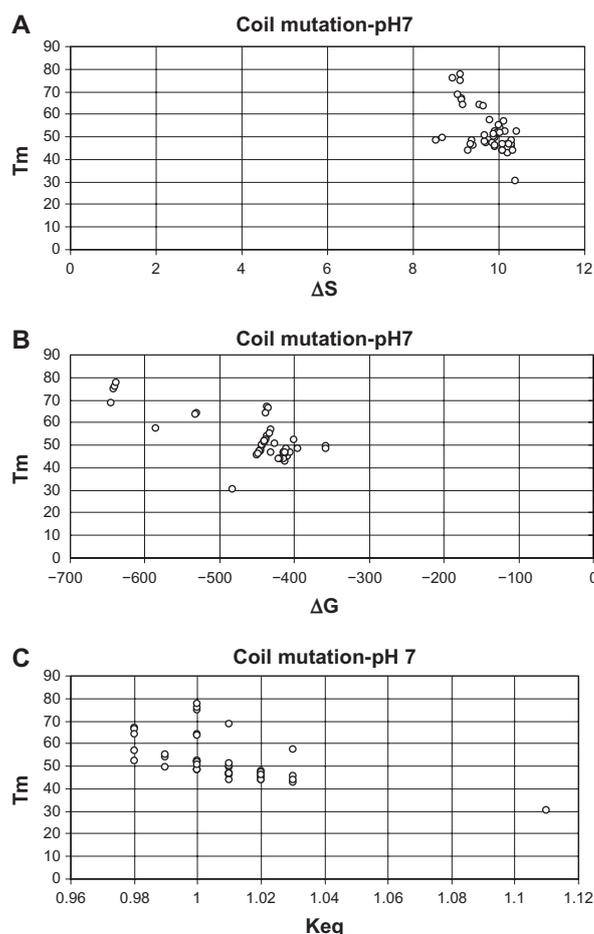
Figure 2 (**A**) Illustration of entropy change versus melting temperature with no data redundancy. (**B**) Illustration of free energy change versus melting temperature with no data redundancy. (**C**) Illustration of equilibrium constant versus melting temperature with no data redundancy.

Table 3. Illustration of derived regression models and regression coefficients.

Sl. no.	Parameters	Regression model	Regression coefficient
1	Thermodynamic	$Tm = 298.09 + 0.58\Delta HvH - 31.63\Delta S - 0.04\Delta G$	**0.97**
2	Chemical	$Tm = 331.63 - 276.06\ Keq - 0.98\Delta Gtr$	**0.63**
3	Geometrical	$Tm = 53.31 - 0.01Z + 0.94Pc$	0.19
4	Combined	$Tm = 308.17 + 0.17\Delta HvH - 10.78\Delta S - 0.09\Delta G - 199.67\ Keq + 0.69\Delta Gtr - 0.01Z + 2.40Pc$	**0.98**

Note: Bold data represents high correlation (>0.5).

However, multiple Tm values are found for the same entry in the ProTherm database under the same set of conditions. The collected data has been categorized into the dataset with low Tm values and the dataset with high Tm values, and appreciable correlations are not found for entries with high Tm values. So, the dataset with low Tm values is focused for further analysis, and the melting temperature shows a strong negative correlation with free energy changes. The melting temperature also gives an average negative correlation with entropy change and the equilibrium constant (Table 2).

The distribution of entropy change, free energy change, and the equilibrium constant-related data with melting temperature is illustrated in Figure 2A–C, in which $Tm = 40\,°C-60\,°C$ consists of many data with $\Delta S = 9-11$ kcal/$°C$ mol, $\Delta G = -400$ to -450 kcal/mol and $Keq = 0.98-1.03$. The multiple regression of Tm is high with thermodynamic parameters, average with chemical parameters, and negligible with geometrical parameters (Table 3).

Conclusion

Many mutants are observed at $Tm = 40\,°C-60\,°C$ with $\Delta S = 9-11$ kcal/$°C$ mol, $\Delta G = -400$ to -450 kcal/mol, and $Keq = 0.98-1.03$. The multiple regression analysis of coil mutants Tm revealed that the influences of thermodynamic parameters are strong ($R = 0.97$), chemical parameters are moderate ($R = 0.63$), and geometrical parameters are negligible ($R = 0.19$); however, the combination of all these three parameters ($R = 0.98$) exhibits slightly higher influences on Tm than the thermodynamic parameters. It has been concluded that the thermodynamic parameters alone are very important in stability studies on protein coil mutants. In addition, the derived regression model would have been useful for the reliable prediction of the melting temperature of coil mutants.

Author Contributions
Conceived of and designed the experiments: ASM. Analyzed the data: ASM. Wrote the first draft of the manuscript: ASM. Contributed to the writing of the manuscript: ASM, GA. Agree with manuscript results and conclusions: ASM. Jointly developed the structure and arguments for the paper: ASM, GA. Made critical revisions and approved final version: GA. All authors reviewed and approved of the final manuscript.

Competing Interests
Author(s) disclose no potential conflicts of interest.

Disclosures and Ethics
As a requirement of the publication, the author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship

and contributorship, conflicts of interest, privacy, and confidentiality, and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

References

1. Moutevelis E, Woolfson DN. A periodic table of coiled-coil protein structures. *J Mol Biol*. 2009;385(3):726–32.
2. Burkhard P, Stetefeld J, Strelkov SV. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol*. 2001;11(2):82–8.
3. Burkhard P, Ivaninskii S, Lustig A. Improving coiled-coil stability by optimizing ionic interactions. *J Mol Biol*. 2002;318(3):901–10.
4. Xi Z, Gao Y, Sirinakis G, Guo H, Zhang Y. Single-molecule observation of helix staggering, sliding, and coiled coil misfolding. *Proc Natl Acad Sci U S A*. 2012;109(15):5711–6.
5. Kwok SC, Hodges RS. Clustering of large hydrophobes in the hydrophobic core of two-stranded alpha-helical coiled-coils controls protein folding and stability. *J Biol Chem*. 2003;278(37):35248–54.
6. Gromiha MM, Parry DA. Characteristic features of amino acid residues in coiled-coil protein structures. *Biophys Chem*. 2004;111(2):95–103.
7. McClain DL, Binfet JP, Oakley MG. Evaluation of the energetic contribution of interhelical Coulombic interactions for coiled coil helix orientation specificity. *J Mol Biol*. 2001;313(2):371–83.
8. Kohn WD, Kay CM, Hodges RS. Protein destabilization by electrostatic repulsions in the two-stranded alpha-helical coiled-coil/leucine zipper. *Protein Sci*. 1995;4(2):237–50.
9. Diss ML, Kennan AJ. Simultaneous directed assembly of three distinct heterodimeric coiled coils. *Org Lett*. 2008;10(17):3797–800.
10. Kumar MD, Bava KA, Gromiha MM, et al. ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Res*. 2006;34(Database issue):D204–6.
11. Momen-Roknabadi A, Sadeghi M, Pezeshk H, Marashi SA. Impact of residue accessible surface area on the prediction of protein secondary structures. *BMC Bioinformatics*. 2008;9:357.
12. Adamczak R, Porollo A, Meller J. Combining prediction of secondary structure and solvent accessibility in proteins. *Proteins*. 2005;59(3):467–75.
13. Gorania M, Seker H, Haris PI. Predicting a protein's melting temperature from its amino acid sequence. *Conf Proc IEEE Eng Med Biol Soc*. 2010;2010:1820–3.
14. Sharp KA, Nicholls A, Friedman R, Honig B. Extracting hydrophobic free energies from experimental data: relationship to protein folding and theoretical models. *Biochemistry*. 1991;30(40):9686–97.
15. Bordo D, Argos P. Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J Mol Biol*. 1991;217(4):721–9.
16. Tsigelny IF. Protein structure prediction: bioinformatics approach. *International University Line*. 2002:493.
17. Costantini S, Colonna G, Facchiano AM. Amino acid propensies for secondary structures are influenced by the protein structural class. *Biochem Biophys Res Commun*. 2006;342(2):441–51.
18. Boice JA, Dieckmann GR, DeGrado WF, Fairman R. Thermodynamic analysis of a designed three-stranded coiled coil. *Biochemistry*. 1996;35(46):14480–5.
19. Kammerer RA. Alpha-helical coiled-coil oligomerization domains in extracellular proteins. *Matrix Biol*. 1997;15(8–9):555–65; discussion 567–8.
20. Kohn WD, Kay CM, Hodges RS. Salt effects on protein stability: two-stranded alpha-helical coiled-coils containing inter- or intrahelical ion pairs. *J Mol Biol*. 1997;267(4):1039–52.
21. McClain DL, Gurnon DG, Oakley MG. Importance of potential interhelical salt-bridges involving interior residues for coiled-coil stability and quaternary structure. *J Mol Biol*. 2002;324(2):257–70.
22. Yu YB. Coiled-coils: stability, specificity, and drug delivery potential. *Adv Drug Deliv Rev*. 2002;54(8):1113–29.
23. Lu SM, Hodges RS. Defining the minimum size of a hydrophobic cluster in two-stranded helical coiled-coils: effects on protein stability. *Protein Sci*. 2004;13(3):714–26.
24. Xu C, Joss L, Wang L, Pechar M, Kopecek J. The influence of fusion sequences on the thermal stabilities of coiled-coil proteins. *Macromol Biosci*. 2002;2:395–401.
25. Apgar JR, Gutwin KN, Keating AE. Predicting helix orientation for coiled-coil dimers. *Proteins*. 2008;72(3):1048–65.
26. Lee DL, Lavigne P, Hodges RS. Are trigger sequences essential in the folding of two-stranded alpha-helical coiled-coils? *J Mol Biol*. 2001;306(3):539–53.
27. Meisner KW, Sosnick TR. Fast folding of a helical protein initiated by the collision of unstructured chains. *Proc Natl Acad Sci U S A*. 2004;101(37):13478–82.
28. Steinkruger JD, Woolfson DN, Gellman SH. Side-chain pairing preferences in the parallel coiled-coil dimer motif: insight on ion pairing between cores and flanking sites. *J Am Chem Soc*. 2010;132(22):7586–8.
29. Swindells MB, MacArthur MW, Thornton JM. Intrinsic phi, psi propensies of amino acids, derived from the coil regions of known structures. *Nat Struct Biol*. 1995;2(7):596–603.
30. Grewal PS. *Numerical Methods of Statistical Analysis*. New Delhi: Sterling Publishers; 1997.
31. Sanders DH, Eng RJ, Murph F. *Statistics: A Fresh Approach,* 3rd ed. New York: McGraw Hill; 1985.