Effect of Catenation on Protein Folding Stability

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · SEPTEMBER 2003

Impact Factor: 12.11 · DOI: 10.1021/ja0355978 · Source: PubMed

1 AUTHOR:

Huan-Xiang Zhou
Florida State University
259 PUBLICATIONS **10,085** CITATIONS

SEE PROFILE

J|A|C|S
COMMUNICATIONS

Published on Web 07/11/2003

Effect of Catenation on Protein Folding Stability

Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306

Received April 11, 2003; E-mail: hxzhou@csit.fsu.edu

The catenane topology, consisting of two interlocked rings, is common in DNAs.[1] In proteins, the first examples of catenanes have just been reported.[2] Like covalent linking between two subunits of a dimeric protein, catenation changes the folding process of the protein from bimolecular to unimolecular. Here I present a theoretical model for the effect of catenation on folding stability. Catenation has two significant consequences (see Figure 1). First, both subunits are circularized. Second, the relative motion of the subunits in the unfolded state is constrained because the two circularized subunits are interlocked. Accounting for the dual effects leads to the conclusion that catenation is 10- to 1000-fold more powerful in stabilizing the folded structure than covalent linking, thus rationalizing the dramatic stabilization of a protein catenane observed by Blankenship and Dawson.[2c]

The theory builds on previous studies of the effects of covalent linking and backbone cyclization on protein stability.[3,4] The effect of catenation can be measured by the ratio of the folding equilibrium constants, $K^{\rm lin}$ and $K^{\rm cat}$, for the linear dimeric and the catenated proteins. This ratio defines an effective concentration:

$$K^{\rm cat}/K^{\rm lin} = C_{\rm cat} \tag{1}$$

When the protein concentration is less than the effective concentration, the folded fraction of the dimeric variant is less than that of the catenated variant. For covalent linking of a dimeric protein, the effective concentration $C_{\rm cl}$ was found to be reasonably predicted by the probability density $p(\mathbf{r})$ for the end-to-end vector of the peptide linker connecting the C terminal of one subunit to the N terminal of the other, with the linker modeled as a wormlike chain.[3] Values of $C_{\rm cl}$ ranged from 0.001 to 0.1 M.

To calculate $C_{\rm cat}$, it is useful to introduce a fictitious intermediate state, consisting of two uninterlocked, unfolded circular chains, which are free to move relative to each other (see Figure 1B, lower branch).[5] The equilibrium constant from the interlocked unfolded chains to this intermediate is an effective concentration, $C_{\rm il}$, arising from the constraint of relative motion between the subunits by the interlocking (see below). The step from the intermediate state to the folded state is equivalent to the folding of the linear dimeric protein (shown in Figure 1A with equilibrium constant $K^{\rm lin}$), except for the effects of the backbone cyclization of the two subunits. The equilibrium constant for this step is thus $K^{\rm lin}q_{\rm cyc,A}q_{\rm cycl,B}$, where $q_{\rm cycl}$ is the enhancement in folding stability by backbone cyclization. This enhancement was found to be given by[4]

$$q_{\rm cycl} = p(d)/\int p(\mathbf{r})P_{\rm u}(\mathbf{r})d^3\mathbf{r} \tag{2}$$

where $P_{\rm u}(\mathbf{r})$ is the probability density for the vector between the N and C termini of a linear subunit and d is the distance between the termini in the folded structure (see Figure 1A).[6]

The effective concentration $C_{\rm il}$ due to the interlocking of the unfolded chains was calculated by treating each chain as a rigid circular ring. One ring, with radius a, was centered at the origin and oriented within the xy plane. The other, with radius b, was randomly centered and oriented. Specifically, its center was randomly distributed within a rectangular box with sides $2(a+b)$, $2(a+b)$, and $2b$ along the x, y, and z directions, respectively. This box was centered at the origin. For each configuration of the second ring, the two points of intersection with the xy plane were found. If one intersection point was inside the first ring while the other intersection point was outside, then the two rings were interlocked. The fraction, $f_{\rm il}$, of interlocked configurations was computed from the random sampling. The effective concentration due to the interlocking of the two rings was then obtained from

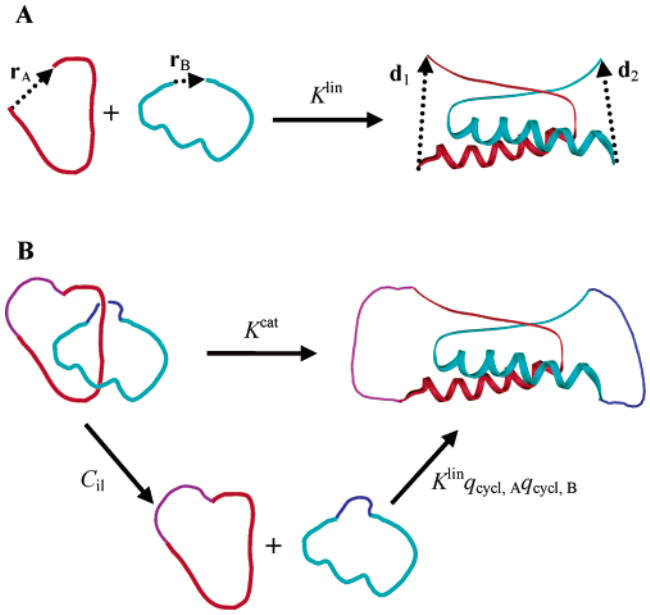

Figure 1. Folding equilibria of a linear dimeric protein (A) and a catenated protein (B). In the lower branch of panel B, a fictitious intermediate state is introduced. This state consists of two uninterlocked unfolded chains, which are free to move relative to each other.

$$C_{\rm il}^{-1} = 6.022 \times 10^{-4} \times 8(a+b)^2 b f_{\rm il}\ {\rm M}^{-1} \tag{3}$$

where a and b are in units of Å. Taken the two steps together, the overall effective concentration is

$$C_{\rm cat} = C_{\rm il}q_{\rm cycl,A}q_{\rm cycl,B} \tag{4}$$

A fuller account of the theory for $C_{\rm cat}$ can be found in Supporting Information.

The theory was applied to the protein catenane designed by Blankenship and Dawson.[2c] This design, called p53cat$^{\rm dim}$, was based on the M340E/L344K dimer mutant of the p53 tetramerization domain.[7] The dimer mutant is shown in Figure 1A as the folded state. In p53cat$^{\rm dim}$, four residues each were added to the N and C termini (at E326 and K357, respectively) of each subunit. The new

10.1021/ja0355978 CCC: $25.00

termini were then ligated, leading to an interlocked topology between the two identical subunits. The linear subunit and the peptide linker thus have $N = 31$ and $l = 9$ peptide bonds, respectively.

The unfolded circular chain of each subunit has a circumference of $(N + l) \times 3.8$ Å $= 152$ Å. If this chain forms a circular ring, the radius would be 24 Å. However, most configurations of the unfolded chain will be much more compact than the fully extended ring. In addition, the two interlocked chains will not be able to come into rim-to-rim contact because of excluded volume. Therefore, a more reasonable estimate for the "effective" radius of the unfolded chain is perhaps half of the maximum radius, i.e., 12 Å. With $a = b = 12$ Å, the fraction of configurations with interlocking configurations was found to be 0.25. The effective concentration due to the interlocking of the two unfolded chains was then $C_{il} = 0.12$ M. Increasing the rigid ring radii to 19 Å would reduce C_{il} by a factor of 4.

In the ensemble of 20 NMR-derived structures of the linear dimer (Protein Data Bank entry 1hs5),[7] the distances between the C_α atoms of E326 and K357 ranged from 11.1 to 24.3 Å, with an average of 17.2 Å. With $N = 31$, $l = 9$, and $d_1 = d_2 = 17.2$ Å, the enhancement in folding stability by the backbone cyclization of both subunits was found to be $q_{cyl,A}q_{cyl,B} = 2.13^2 = 4.5$. The overall effective concentration by catenation was then 0.12×4.5 M $= 0.54$ M. Blankenship and Dawson[2c] measured $k_BT \ln K^{lin}$ and $k_BT \ln K^{cat}$ to be 8.7 and 9.0 kcal/mol, respectively, at $T = 277$ K. The experimental result for the effective concentration was thus $C_{cat} = K^{cat}/K^{lin} = \exp[(9.0-8.7)/k_BT]$ M $= 1.7$ M. The calculated result is in reasonable agreement with experiment.[8]

In the previous study on the effect of backbone cyclization,[4] it was found that the stabilization can be maximized by varying the linker length. For p53cat^{dim}, with $d_1 = d_2 = 17.2$ Å and $N = 31$, $q_{cycl,A}$ $(= q_{cycl,B})$ was maximized to 2.51 when the linker length was increased from 9 to 15.[9] However, increasing the linker length also leads to an increase in the volume in which the two unfolded chains can have relative motion, and hence a decrease in C_{il}.[9] The overall effective concentration was predicted to change from 0.54 M to 0.33, 0.57, 0.55, 0.50, and 0.43 M when the linker length was changed from 9 to 7, 11, 13, 15, and 17, respectively. The predicted dependence of the effective concentration on the linker length can be tested experimentally.

In this study the effect of interlocking the two unfolded chains (first step of the lower branch in Figure 1B) was obtained by treating each chain as a rigid ring. A much more realistic treatment of interlocking has already been developed for studying catenated DNAs[1b] and can be adapted to study catenated proteins (the adaptation requires elaborate fine-tuning of several parameters). One important aspect neglected in the rigid-ring treatment is that, in an interlocked configuration, many conformations otherwise accessible to each unfolded chain will be eliminated because they lead to topologies other than the catenane (e.g., uninterlocked or multiple chain-crossing). This aspect is akin to what happens when an unfolded chain is near the boundary of a confining space.[10] It serves to increase the effective concentration C_{il}. Its neglect in the rigid-ring treatment perhaps explains the underestimate of K^{cat}/K^{lin} presented above.

Only one example of naturally catenated proteins has been reported thus far.[1a] However, there is no reason to believe that this report is just a rare exception. The first backbone-cyclized protein was obtained by chemical ligation.[11] After nearly 20 years, a number of naturally circular proteins were identified.[12] If circular proteins offer a clue, we should expect to see more naturally catenated proteins. Moreover, a large number of dimeric proteins have been found to have intertwined interfaces such as that in the p53 dimer mutant (see Supporting Information). These dimeric proteins can be readily catenated by chemical ligation.

In summary, a theory for the stabilizing effect of the catenane topology has been developed. There are two distinct contributions: backbone cyclization of the two subunits and constrained relative motion of the subunits in the unfolded state by the interlocking. The predicted effect for catenating a p53 dimer mutant was in reasonable agreement with the experimental result. The effective concentration due to covalent linking is in the range of 0.001−0.1 M, whereas the effective concentration due to catenation has been found to be ~1 M. Hence, catenation is 10−1000-fold more powerful in stabilizing the folded structure. This dramatic stabilization offers a strong incentive for further exploiting catenation as a protein design strategy.

Acknowledgment. This work was supported in part by NIH Grant GM58187.

Supporting Information Available: Two examples of dimeric proteins ready for catenation and full derivations of the theory (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Wang, J. C. *Nat. Rev. Mol. Cell. Biol.* **2002**, *3*, 430−440. (b) Vologodskii, A. V.; Zhang, W.; Rybenkov, V. V.; Podtelezhnikov, A. A.; Subramanian, D.; Griffith, J. D.; Cozzarelli, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3045−3049.
(2) (a) Wikoff, W. R.; Liljas, L.; Duda, R. L.; Tsuruta, H.; Hendrix, R. W.; Johnson, J. E. *Science* **2000**, *289*, 2129−2133. (b) Yan, L. Z.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2001**, *40*, 3625−3627. (c) Blankenship, J. W.; Dawson, P. E. *J. Mol. Biol.* **2003**, *327*, 537−548.
(3) Zhou, H.-X. *J. Am. Chem. Soc.* **2001**, *123*, 6730−6731.
(4) Zhou, H.-X. *J. Mol. Biol.*, **2003**. In press.
(5) Topologically it is not possible to pass from an interlocked to an uninterlocked configuration. However, physically the fictitious intermediate state is a special case of the unfolded state, where all interactions between the two circular chains are turned off.
(6) In calculating q_{cycl}, $p(\mathbf{r})$ was generated by modeling the peptide linker as a wormlike chain while $P_u(\mathbf{r})$ was approximated by a Gaussian distribution.[4]
(7) Davison, T. S.; Nie, X.; Ma, W.; Lin, Y.; Kay, C.; Benchimol, S.; Arrowsmith, C. H. *J. Mol. Biol.* **2001**, *307*, 605−617.
(8) The equilibrium of unimolecular (catenated) folding is concentration independent, but the equilibrium of bimolecular (dimeric) folding is concentration dependent. Hence, care should be exercised in comparing the energetics of the two equilibria. Here this issue was circumvented by working with equilibrium constants, K^{cat} and K^{lin}. By definition K^{lin} does not depends on concentration; it actually determines the concentration dependence of the folded fraction. The experimental results considered here are those extrapolated to zero denaturant. Blankenship and Dawson[2c] found a significant difference in the dependences of K^{cat} and K^{lin} on denaturant concentration. However, this difference may arise primarily from the different interactions with the denaturant by the catenated and dimeric variants in the unfolded state.
(9) For linker lengths of 7, 9, 11, 13, 15, and 17, $q_{cycl,A}$ and $q_{cycl,B}$ were found to be 1.55, 2.13, 2.35, 2.47, 2.51, and 2.49, respectively. On the other hand, assuming that a and b scale with the lengths of the circular chains, C_{il} were predicted to be 0.14, 0.12, 0.10, 0.09, 0.08, and 0.07 M, respectively.
(10) Zhou, H.-X.; Dill, K. A. *Biochemistry* **2001**, *40*, 11289−11293.
(11) Goldenberg, D. P.; Creighton, T. E. *J. Mol. Biol.* **1984**, *179*, 527−545.
(12) Trabi, M.; Craik, D. J. *Trends Biochem. Sci.* **2002**, *27*, 132−138.

JA0355978