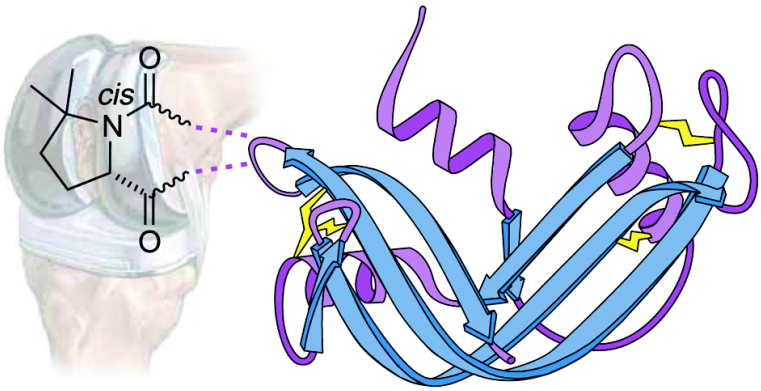
Communication

Protein Prosthesis: A Nonnatural Residue Accelerates Folding and Increases Stability

Ulrich Arnold, Matthew P. Hinderaker, Jens Kditz, Ralph Golbik, Renate Ulbrich-Hofmann, and Ronald T. Raines

J. Am. Chem. Soc., **2003**, 125 (25), 7500-7501• DOI: 10.1021/ja0351239 • Publication Date (Web): 30 May 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 05/30/2003

Protein Prosthesis: A Nonnatural Residue Accelerates Folding and Increases Stability

Ulrich Arnold,[†] Matthew P. Hinderaker,[‡] Jens Köditz,[†] Ralph Golbik,[†] Renate Ulbrich-Hofmann,[†] and Ronald T. Raines*[,‡,§]

Department of Biochemistry/Biotechnology, Martin-Luther University, 06099 Halle, Germany, and Departments of Chemistry and Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

Received March 12, 2003; E-mail: raines@biochem.wisc.edu

The chemical basis for the folding of amino acid chains into particular conformations and the stability of those conformations has been studied for decades.[1] A major goal of this effort is to increase the rate at which proteins attain their native structures and enhance the stability of those structures. Here, we demonstrate that a nonnatural amino acid can be installed into a protein to both accelerate its folding and enhance its conformational stability. Moreover, installation of this residue provides new insight into the pathway by which the protein folds into its native structure.

A polypeptide chain must reverse its direction to fold into a compact structure. An especially simple means to that end is the formation of a cis (that is, *E*) peptide bond.[2] Only 0.03% of $Xaa_{i-1}-nonPro_i$ peptide bonds are in the cis conformation in folded proteins.[3] This prevalence increases to 5.2% for $Xaa_{i-1}-Pro_i$ peptide bonds. Accordingly, turns often contain a proline residue preceded by a cis peptide bond.[3b,4]

The nonnatural amino acid 5,5-dimethyl-L-proline (dmP) is unusual in that an $Xaa_{i-1}-dmP_i$ peptide bond is almost exclusively in the cis conformation.[5] We reasoned that the rational use of this constraint could accelerate the folding of a protein as well as enhance its conformational stability. The incorporation of nonnatural residues such as dmP via recombinant DNA techniques is made problematic by intrinsic limitations in both the genetic code and the substrate specificity of the ribosomal peptidyl transferase. Expressed protein ligation (EPL) has emerged as a tool to overcome these limitations.[6] EPL combines the prodigious biosynthetic capability of the ribosome with the entire palette of synthetic chemistry.

As a model system, we chose ribonuclease (RNase) A, which has been the object of much seminal work in protein chemistry.[7] Residues Gly112−Asn113−Pro114−Tyr115 of RNase A form a β-turn in which the peptide bond preceding Pro114 is in the cis conformation. We used EPL to replace Pro114, which is one of four proline residues in RNase A, with dmP. Specifically, we used biosynthesis to produce RNase A fragment 1−94 with a C-terminal thioester.[8] We then synthesized dmPOH and used it to incorporate a dmP residue at position 114 of a peptide corresponding to residues 95−124. The two fragments were ligated, and the product was folded and purified to yield intact dmP114 RNase A (Figure 1; *m*/*z* 13 845; expected: 13 841).

Enzymatic catalysis provides an extremely stringent measure of native protein structure.[9] This measure is especially useful herein, as the synthetic portion of dmP114 RNase A contains His119, which contributes 10^5-fold to the value of k_{cat}/K_M.[10] The catalytic activity of dmP114 RNase A ($k_{cat}/K_M = 1.6 \pm 0.2 \times 10^7$ M^{-1} s^{-1}) is indistinguishable from that of the wild-type enzyme ($k_{cat}/K_M = 1.5 \pm 0.2 \times 10^7$ M^{-1} s^{-1}).[11] We conclude that the addition of two methyl groups to residue 114 has no significant effect on the structure of RNase A.

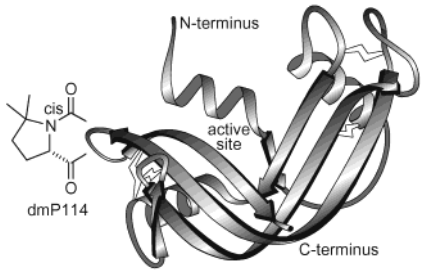
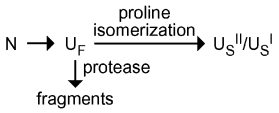


Figure 1. Ribbon diagram of ribonuclease A showing the site of installation of a 5,5-dimethyl-L-proline residue.

Scheme 1



The dmP residue enhances the conformational stability of RNase A. The value of T_m, which is the temperature at the midpoint of the thermal transition, increases from 63.6 ± 0.2 °C in wild-type RNase A[8] to 66.4 ± 0.2 °C in the dmP114 variant. This increase corresponds to $\Delta\Delta G_m = 4.6 \pm 0.4$ kJ mol^{-1}.[12] In marked contrast, all attempts to replace Pro114 of RNase A with natural residues result in a substantial decrease in stability.[13] The installation of two β-amino acid residues, (*R*)-nipecotic acid−(*S*)-nipecotic acid, at positions 113−114 does enhance the stability of RNase A ($\Delta T_m = 1.2 \pm 0.3$ °C),[8] but less than that of a single dmP residue at position 114 ($\Delta T_m = 2.8 \pm 0.3$ °C).

The enhanced stability of dmP114 RNase A could arise from its slower unfolding or faster folding (or both).[14] To dissect the effect of the dmP114 residue, we determined the rate constants for the unfolding and folding of both dmP114 RNase A and the wild-type protein. Unfolding rate constants (k_U) were determined by limited proteolysis with thermolysin.[15] The two proteins unfold with the same rate constant at 47.5−57.5 °C. This result is in accordance with the generally accepted model for the unfolding pathway of RNase A.[16] As a result of prolyl peptide bond isomerizations,[17] RNase A exists in its unfolded state as a mixture of three major unfolded species, which fold at decreasing rates: $U_F > U_S^{II} > U_S^{I}$ (Scheme 1).[18] Prolyl peptide bond isomerization occurs subsequently to the fast unfolding of the native enzyme N to U_F, which is a substrate for proteolysis.[15] Thus, as expected, dmP114 has no effect on the value of k_U.

Folding rate constants (k_F) were determined by using stopped-flow fluorescence spectroscopy to monitor the recovery of catalytic activity upon folding.[11] In 4.5 M guanidine−HCl (Gdn−HCl), neither wild-type RNase A nor its dmP114 variant has detectable activity, and both are unfolded completely according to their intrinsic fluorescence. Upon dilution, dmP114 RNase A regains its activity 6-fold faster than does wild-type RNase A at all

[†] Martin-Luther University.
[‡] Department of Chemistry, University of Wisconsin, Madison.
[§] Department of Biochemistry, University of Wisconsin, Madison.

10.1021/ja0351239 CCC: $25.00

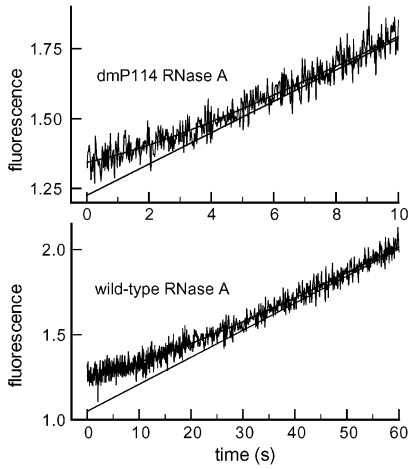


Figure 2. Reactivation of unfolded dmP114 RNase A and wild-type RNase A at 5 °C.[11] Raw data were fitted to the single-exponential plus linear (i.e., steady-state) equation: fluorescence $(t) = Ae^{-k_F t} + mt + n$.

Table 1. Values of k_F (±25%) and $\Delta\Delta G^{\ddagger}$ for Reactivation of Wild-type RNase A and Its dmP114 Variant at 5, 15, and 25 °C

	k_F (s^{-1})		
T (°C)	wild-type RNase A	dmP114 RNase A	$\Delta\Delta G^{\ddagger}$ (kJ mol^{-1})
5	0.0472	0.275	4.1
15	0.0873	0.617	4.7
25	0.117	0.665	4.3

temperatures investigated (Figure 2; Table 1). Accordingly, $\Delta G^{\ddagger}$ for the reactivation of dmP114 RNase A is lower by $\Delta\Delta G^{\ddagger} = 4.4 \pm 0.3$ kJ mol^{-1}. This value of $\Delta\Delta G^{\ddagger}$ is indistinguishable from that of $\Delta\Delta G_m$, which indicates that the enhanced stability of dmP114 RNase A arises solely from faster folding.[19] Within the dead-time of mixing,[20] 30−35% of the final activity was recovered, in accordance with a previous study of wild-type RNase A using CD spectroscopy and dynamic light scattering.[21] Because U_F folds with a rate constant of ~50 s^{-1} under these conditions (0.17 M Gdn−HCl)[13b,22] and accounts for ~20% of all unfolded molecules,[23] the gain of activity within the dead-time of mixing can be attributed to the formation of native enzyme from U_F. The values of k_F (Table 1) likely correspond, then, to the formation of active enzyme from U_S^{II}.[24,25] The accelerated folding of dmP114 RNase A contradicts the prior designation of Pro114 as a residue that does not affect the overall folding rate.[26] Rather, Pro114 is a "type II" proline residue[27]—one that can affect but not block folding.

We propose two explanations for the faster folding of dmP114 RNase A. First, dmP114 could give U_S^{II} a structure that is more similar to that of the native protein. An early intermediate with largely intact secondary structure has been reported in the folding of U_S^{II}.[16,22b] The formation of this intermediate could be facilitated by the preorganization of C-terminal residues proximal to dmP114. Second, dmP114 could eliminate slower folding subspecies within U_S^{II}, thereby increasing the fraction of the faster folding ones. The overall rate of U_S^{II} folding is an ensemble average from all of its subspecies, which have heterogeneous structures and varying prolyl peptide bond isomers.[28] Evidence for various folding pathways within U_S^{II} was provided by the detection of an early folding intermediate, which occurs for only ~40% of all U_S^{II}.[16,22b] Apparently, a transient barrier exists to the folding of these U_S^{II} subspecies, which could have a trans peptide bond preceding Pro114. Fixing this peptide bond in the cis conformation would eliminate this barrier.

We conclude that dmP is a superb tool for protein chemists. The replacement of a cis proline residue with dmP endows RNase A with faster folding and enhanced conformational stability without compromising native structure. We envision that the rational use of dmP and other prosthetic segments could enable chemotherapeutic proteins (such as variants and homologues of RNase A[29]) to survive longer in vivo or retain activity after oral administration.

Acknowledgment. U.A. was supported by the Arthur B. Michael postdoctoral fellowship. M.P.H. was supported by NIH Chemistry−Biology Interface Training Grant GM08506. J.K. was supported by the Max-Buchner-Forschungsstiftung. This work was supported by Grants GM44783 (NIH), INT-0129163 (NSF), and D/0104622 (DAAD).

Supporting Information Available: Procedures and additional data for syntheses and analysis reported herein (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Richards, F. M. *Cell. Mol. Life Sci.* **1997**, *53*, 790−802.
(2) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425−1436.
(3) (a) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol.* **1990**, *214*, 253−260. (b) Weiss, M. S.; Jabs, A.; Hilgenfeld, R. *Nat. Struct. Biol.* **1998**, *5*, 676.
(4) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397−412.
(5) (a) An, S. S. A.; Lester, C. C.; Peng, J.-L.; Li, Y.-J.; Rothwarf, D. M.; Welker, E.; Thannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J. Am. Chem. Soc.* **1999**, *121*, 11558−11566. (b) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188−1194. For other means to stabilize the cis isomer, see: (c) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777−778. (d) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem., Int. Ed.* **2001**, *40*, 923−925.
(6) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249−289.
(7) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045−1065.
(8) Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. *J. Am. Chem. Soc.* **2002**, *124*, 8522−8523.
(9) Knowles, J. R. *Science* **1987**, *236*, 1252−1258.
(10) Thompson, J. E.; Raines, R. T. *J. Am. Chem. Soc.* **1994**, *116*, 5467−5468.
(11) Assays of catalytic activity used the fluorogenic substrate: 6-carboxyfluorescein-dArU(dA)$_2$-6-carboxytetramethylrhodamine (Kelemen, B. R.; Klink, T. A.; Behlke, M. A.; Eubanks, S. R.; Leland, P. A.; Raines, R. T. *Nucleic Acids Res.* **1999**, *27*, 3696−3701).
(12) The value of $\Delta\Delta G_m$ was calculated by using $\Delta H_m = 574$ kJ mol^{-1} and $\Delta C_p = 9.4$ kJ mol^{-1} K^{-1} for wild-type RNase A (ref 15a).
(13) (a) Schultz, D. A.; Baldwin, R. L. *Protein Sci.* **1992**, *1*, 910−916. (b) Dodge, R. W.; Scheraga, H. A. *Biochemistry* **1996**, *35*, 1548−1559. (c) Schultz, L. W.; Hargraves, S. R.; Klink, T. A.; Raines, R. T. *Protein Sci.* **1998**, *7*, 1620−1625.
(14) Fersht, A. R.; Matouschek, A.; Serrano, L. *J. Mol. Biol.* **1992**, *224*, 771−782.
(15) (a) Arnold, U.; Ulbrich-Hofmann, R. *Biochemistry* **1997**, *36*, 2166−2172. (b) Arnold, U.; Ulbrich-Hofmann, R. *Eur. J. Biochem.* **2001**, *268*, 93−97.
(16) Udgaonkar, J. B.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8197−8201.
(17) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953−4963.
(18) (a) Schmid, F. X.; Blaschek, H. *Eur. J. Biochem.* **1981**, *114*, 111−117. (b) Garel, J. R.; Nall, B. T.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1853−1857.
(19) Because the U_F:U_S^{II}:U_S^{I} ratio did not change significantly, changes in this ratio cannot be responsible for the increase observed in $\Delta G^{\ddagger}$.
(20) Dead-times: dmP114 RNase A, 20 ms; wild-type RNase A, 40 ms.
(21) Noeppert, A.; Gast, K.; Zirwer, D.; Damaschun, G. *Fold. Des.* **1998**, *3*, 213−221.
(22) (a) Garel, J. R.; Nall, B. T.; Baldwin, R. L. *Biochemistry* **1995**, *34*, 4088−4096. (b) Udgaonkar, J. B.; Baldwin, R. L. *Biochemistry* **1995**, *34*, 4088−4096.
(23) Garel, J. R.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 3347−3351.
(24) The active form is actually a native-like intermediate, I_N (ref 18a). The very slow folding of U_S^{I}, which folds with a rate constant of 7×10^{-3} s^{-1} and accounts for ~15% of all unfolded molecules (ref 23), is not detected (or not resolved) herein.
(25) Liu, W.; Tsou, C. L. *Biochim. Biophys. Acta* **1987**, *916*, 465−473.
(26) Adler, M.; Scheraga, H. A. *Biochemistry* **1990**, *29*, 8211−8216.
(27) Levitt, M. *J. Mol. Biol.* **1981**, *145*, 251−263.
(28) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. *Biochemistry* **2002**, *41*, 14637−14644
(29) Leland, P. A.; Raines, R. T. *Chem. Biol.* **2001**, *8*, 405−413.

JA0351239