Stereoelectronic and steric effects in side chains preorganize a protein main chain

Matthew D. Shoulders[a], Kenneth A. Satyshur[b], Katrina T. Forest[b,1], and Ronald T. Raines[a,c,1]

[a]Department of Chemistry, [b]Department of Bacteriology, and [c]Department of Biochemistry, University of Wisconsin, Madison, WI 53706

Edited by David A. Tirrell, California Institute of Technology, Pasadena, CA, and approved November 4, 2009 (received for review September 9, 2009)

Preorganization is shown to endow a protein with extraordinary conformational stability. This preorganization is achieved by installing side-chain substituents that impose stereoelectronic and steric effects that restrict main-chain torsion angles. Replacing proline residues in $(\mathbf{ProProGly})_7$ collagen strands with 4-fluoroproline and 4-methylproline leads to the most stable known triple helices, having T_m values that are increased by >50 °C. Differential scanning calorimetry data indicate an entropic basis to the hyperstability, as expected from an origin in preorganization. Structural data at a resolution of 1.21 Å reveal a prototypical triple helix with insignificant deviations to its main chain, even though 2/3 of the residues are nonnatural. Thus, preorganization of a main chain by subtle changes to side chains can confer extraordinary conformational stability upon a protein without perturbing its structure.

collagen triple helix | nonnatural amino acid | preorganization | protein stability | x-ray crystallography

The three-dimensional structures of proteins are stable because of a delicate balance of forces (1). Increases in the conformational stability of a protein—desirable in many contexts (2)—can be realized by enhancing the hydrophobic effect (3, 4), introducing or shielding a hydrogen bond (5) or electrostatic interaction (6–8), or introducing a disulfide crosslink (9–11) or metal ion-binding site (12). These strategies are often frustrated by enthalpy–entropy compensation, whereby enthalpic stabilization is offset entirely by entropic destabilization, or vice versa (13). Moreover, the strategies often lack subtlety and can lead to a misshapen and hence dysfunctional protein.

As elaborated by Cram (14), the principle of preorganization states that, "the more highly hosts and guests are organized for binding and low solvation prior to their complexation, the more stable will be their complexes." In a typical manifestation, a conformational constraint is imposed upon a receptor or ligand during its chemical synthesis. This principle can also yield biopolymers with increased conformational stability (Fig. 1). For example, the stability of a DNA duplex correlates with the helicity of its single strands (15, 16), and a bicyclic ("locked") derivative of ribose enhances that stability (17, 18). The stability of β-turns within protein structures can be increased by incorporating *D*-proline (19) and certain β-amino acids (20) that bias the conformations occupied by the unfolded polypeptide. Likewise, proteins containing a *cis*-peptide bond can be stabilized by constrained amides or isosteres (21–23). The utility of these substitutions is limited, however, by the difficulty of modifying the backbone of proteins as well as concomitant effects on structure and function (24–27).

We suspected that alterations to proline residues could provide a particularly incisive means to preorganize the conformation of a polypeptide chain. In classic work, Matthews and coworkers substituted proline, which is the least flexible residue, for an amino acid residue in bacteriophage T4 lysozyme with appropriate $\phi(C'_{i-1}\text{-}N_i\text{-}C^{\alpha}_i\text{-}C'_i)$ and $\psi(N_i\text{-}C^{\alpha}_i\text{-}C'_i\text{-}N_{i+1})$ torsion angles (28). The value of T_m increased by 2.1 °C, and the enhanced stability had an entropic origin. The conformation of proline is restricted because two of its main-chain atoms are confined within a pyrrolidine ring (29). We reasoned that constraints within its side chain would be transmitted with precision to its main chain.

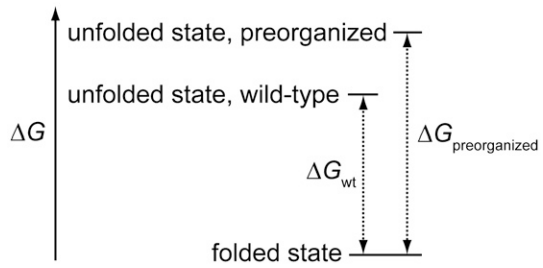

Fig. 1. Protein stabilization by preorganization. The unfolded state is destabilized because preorganization reduces its entropy, whereas the folded state is not perturbed. The net effect is a more favorable ΔG for protein folding.

Previously, we demonstrated that the pucker of a pyrrolidine ring can be biased by either stereoelectronic or steric effects (Fig. 2). (2*S*,4*S*)-4-fluoroproline (flp) and (2*S*,4*R*)-4-fluoroproline (Flp) prefer C^{γ}-endo and C^{γ}-exo ring puckers, respectively (Table S1) (30). This conformational bias is attributable to a stereoelectronic effect—the *gauche* effect. (2*S*,4*R*)-4-methylproline (mep) and (2*S*,4*S*)-4-methylproline (Mep) prefer C^{γ}-endo and C^{γ}-exo ring puckers, respectively (Table S1) (31). This bias is attributable to a steric effect, owing to the preference of the methyl groups to adopt pseudoequatorial positions. Importantly, main-chain torsion angles correlate with ring pucker. Prolines with C^{γ}-endo puckers have torsion angles of $\phi \approx -75°$ and $\psi \approx 164°$, whereas those with C^{γ}-exo puckers have $\phi \approx -60°$ and $\psi \approx 152°$ (30, 32).

As a model system for testing the feasibility of stabilizing a protein by preorganizing the backbone torsion angles of proline residues, we chose the collagen triple helix. This multimeric, tightly packed coiled coil contains many proline residues and requires precise backbone torsion angles for proper folding (33). These characteristics, along with the absence of surface loops and a hydrophobic core (34), lead to an inordinate entropic penalty for triple-helix formation. We suspected that we could obviate the high entropic cost for triple-helix folding by preorganizing individual collagen strands.

A crystal structure of a triple helix formed from the collagen-related polypeptide (CRP) $(\text{ProProGly})_{10}$ provides a roadmap for the rational use of preorganization. In this structure (35), the

Author contributions: M.D.S. and R.T.R. designed research; M.D.S., K.A.S., and K.T.F. performed research; M.D.S., K.A.S., K.T.F., and R.T.R. analyzed data; and M.D.S., K.A.S., K.T.F., and R.T.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3ipn).

[1]To whom correspondence may be addressed. E-mail: rtraines@wisc.edu or forest@bact.wisc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0909592107/DCSupplemental.

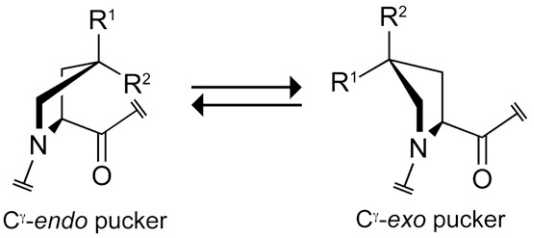

Fig. 2. Ring conformations of 4-substituted prolines. The C^{γ}-endo conformation is favored strongly by stereoelectronic effects if $R_1 = H$, $R_2 = F$ (as in flp) and by steric effects if $R_1 = Me$ (mep), $R_2 = H$. The C^{γ}-exo conformation is favored strongly by stereoelectronic effects if $R_1 = F$ (Flp), $R_2 = H$ and by steric effects if $R_1 = H$, $R_2 = Me$ (Mep).

conformation of the proline ring correlates with the first (Xaa) or second (Yaa) position in the XaaYaaGly triplet repeat. Residues in the Xaa position all have a C^{γ}-endo pucker, whereas those in the Yaa position are C^{γ}-exo (Fig. 2). This observation suggested to us that proline derivatives preferring the C^{γ}-endo ring pucker (and thus the proper backbone torsion angles) could preorganize triple-helix formation when in the Xaa position, whereas those that prefer the C^{γ}-exo ring pucker could preorganize triple-helix formation when in the Yaa position (31, 36–38).

Here, we integrate stereoelectronic and steric effects to preorganize a polypeptide chain. The result is the most stable known collagen triple helix, by far. We show by calorimetric analyses that the conferred hyperstability derives from entropy, as expected if the hyperstability is due to the proper preorganization of backbone torsion angles. Finally, we use high-resolution crystallographic analysis to demonstrate that the incorporation of nonnatural derivatives of proline does not perturb the structure of the triple helix. Based on our findings, we put forth the principle of preorganization and its manifestation via stereoelectronic and steric effects within amino acid side chains as a powerful means to enhance the conformational stability of proteins.

Results

Circular Dichroism Spectroscopy. The peptides $(ProProGly)_7$, $(ProHypGly)_7$, $(flpMepGly)_7$, and $(mepFlpGly)_7$ were prepared by segment condensation by using Fmoc-based solid-phase peptide synthesis (see *SI Text*). Hyp refers to (2*S*,4*R*)-4-hydroxyproline, which is prevalent in the Yaa position of natural collagen. CD spectroscopy indicated that all the CRPs except $(ProProGly)_7$ folded into a triple helix (Fig. 3*A*). Upon heating, the three folded CRPs underwent cooperative transitions (Fig. 3*B*). The triple helix formed from the "natural" collagen sequence, $(ProHypGly)_7$, unfolded with $T_m = 34\,°C$. In contrast, triple helices formed from $(flpMepGly)_7$ and $(mepFlpGly)_7$ were hyperstable with T_m values of 51 °C and 58 °C, respectively (Table 1). These $[(flpMepGly)_7]_3$ and $[(mepFlpGly)_7]_3$ triple helices are the most stable triple helices of their length prepared to date, by a wide margin.

Thermodynamic Analyses. To discern the origin for the hyperstability of $[(flpMepGly)_7]_3$ and $[(mepFlpGly)_7]_3$ triple helices, we undertook a complete thermodynamic analysis of their folding–unfolding equilibria, along with those for $(ProProGly)_7$ and $[(ProHypGly)_7]_3$. Triple helices were analyzed by differential scanning calorimetry (DSC) at a scan rate of 6 °C/h to ensure that the scan rate did not exceed the rate of triple-helix unfolding (Fig. 4) (39). Because $(ProProGly)_7$ does not form a stable triple helix, an endotherm was not observed (Fig. 4*A*). For peptides that displayed endotherms, the DSC data were fitted by using an association–dissociation equilibrium model with ΔC_P dependent on temperature (39, 40). (Equations are derived and the fitting methodology is described in *SI Text*.) This model permits calculation and comparison of ΔG, ΔH, ΔS, and ΔC_P values at any desired temperature. Because the calculated thermodynamic parameters are most accurate near the T_m value for each peptide, we analyzed each of the thermodynamic parameters at $T = 46\,°C$, which is halfway between the T_m for $[(ProHypGly)_7]_3$ and that for $[(mepFlpGly)_7]_3$. The results are summarized in Table 1.

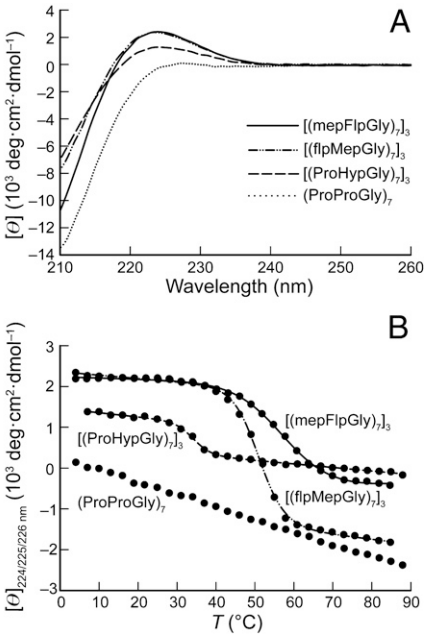

Fig. 3. Conformational analysis of triple-helical CRPs. (*A*) CD spectra at 4 °C of solutions of CRPs (200 μM in 50 mM HOAc) after their incubation at ≤4 °C for ≥24 h. (*B*) Effect of temperature on the molar ellipticity (224, 225, or 226 nm) of the solutions of *A*. Data were recorded at 3 °C intervals after a 5 min. equilibration. Two-state fits are displayed as lines.

X-Ray Crystal Structure of the $[(mepFlpGly)_7]_3$ Triple Helix. To ascertain whether prototypical triple helices are formed from these modified CRPs, we used single crystal x-ray diffraction to obtain a structure of the $[(mepFlpGly)_7]_3$ triple helix at a resolution of 1.21 Å (Table 2). High mobility of atoms near the ends of each peptide prevented accurate determination of the structure at the N- and C-termini, a complication in nearly all other triple-helix structures solved to date (41). Nonetheless, the central fifty-four amino acids of both $[(mepFlpGly)_7]_3$ triple helices in the

Table 1. Thermodynamic parameters for folding of triple-helical CRPs

peptide	T_m (°C)*	ΔG (kcal mol^{-1})†	ΔH (kcal mol^{-1})†	$-T\Delta S$ (kcal mol^{-1})†	ΔC_p (kcal mol^{-1} K^{-1})†
$(ProProGly)_7$	<5	—	—	—	—
$(ProHypGly)_7$	34	−2.1	−51.9	49.8	−0.3
$(flpMepGly)_7$	51	−4.6	−37.8	33.2	−0.3
$(mepFlpGly)_7$	58	−4.4	−25.3	20.9	−0.2

*Values of T_m (±1 °C) are from CD data (Fig. 3*B*).
†Values of ΔG, ΔH, $T\Delta S$, and ΔC_p are at 46 °C. The largest source of error (±5%) was in the determination of [peptide].

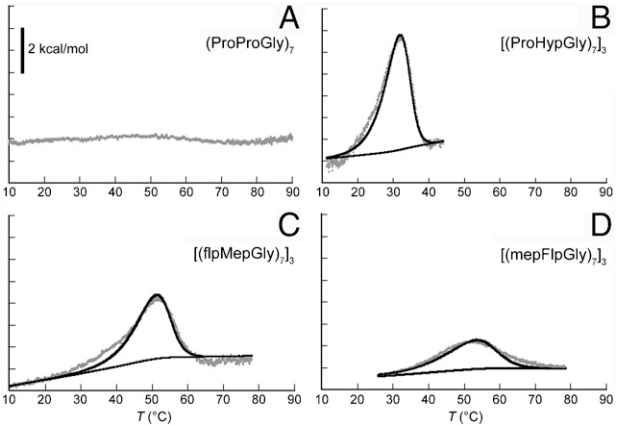

Fig. 4. Differential scanning calorimetry of triple-helical CRPs. Scans of (*A*) $(ProProGly)_7$, (*B*) $[(ProHypGly)_7]_3$, (*C*) $[(flpMepGly)_7]_3$, and (*D*) $[(mepFlpGly)_7]_3$ in 50 mM HOAc (aq). Peptides were heated from 5–98 °C at 6 °C/h. Data are in gray; fits are in black.

asymmetric unit were exceptionally well-defined and led to a high-quality structure, as seen in the fit to the electron density (Fig. S2).

Several aspects of the $[(mepFlpGly)_7]_3$ crystal structure are significant. First and foremost, this structure is of a protein in which fully 2/3 of the residues are nonnatural. Second, the structure is the first of a collagen triple helix containing a nonnatural derivative of proline. Third, $[(mepFlpGly)_7]_3$ is the shortest triple helix to succumb to analysis by x-ray diffraction. Finally, to the best of our knowledge, it has the highest resolution amongst all known nonaveraged collagen structures.

The overall crystal packing arrangement of $[(mepFlpGly)_7]_3$ is noteworthy. The two triple helices in each asymmetric unit pack in an antiparallel orientation, an unusual observation for collagen peptides. This packing creates a ladder of pockets of water alternating with close approaches of 3.3–3.6 Å of the adjacent triple helices (Fig. 5).

Remarkably, this largely nonnatural CRP adopts an overall fold that is indistinguishable from natural collagen sequences

Table 2. X-Ray crystallographic parameters for $[(mepFlpGly)_7]_3$

Data collection	
Wavelength (Å)	0.9785
Resolution (Å)	31.0 – 1.21 (1.5–1.21)
No. of unique reflections	21822 (4151)*
Completeness (%)	86.6 (60.4)
Redundancy	8.2 (4.5)
R_{symm} (%)	8.9 (14.6)
I/σ_I	9.9 (5.3)
Wilson *B* value (Å^2)	8.6
Space group	$P2_1$
a, *b*, *c* (Å)	26.38, 25.31, 61.93
α, β, γ (deg)	90.0, 90.05, 90.0
Refinement	
R-factor (%)	18.7
Free *R*-factor (%)	25.8
Protein atoms	800
Solvent atoms	194
Heteroatoms	16
B overall	19.3
B protein atoms	17.9
B solvent atoms	26.0
B heteroatoms	15.5
RMSD bonds (Å)/angles (°)	0.011/2.0

*The value for the highest resolution shell is shown in parentheses.

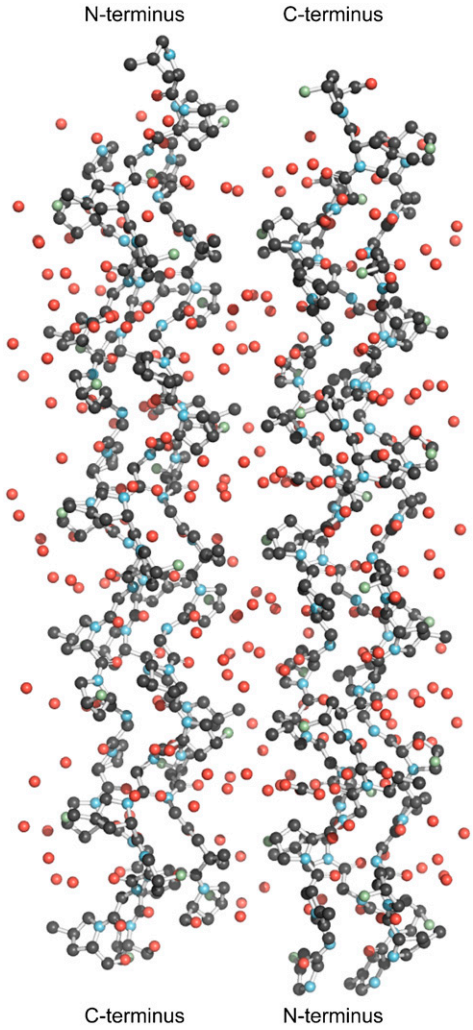

Fig. 5. Asymmetric unit of a crystalline $[(mepFlpGly)_7]_3$ triple helix. The two triple helices are antiparallel. Terminal glycine residues are included for illustration, but their coordinates are not deposited in the Protein Data Bank.

studied to date (Table 3). The rmsd of the main-chain atoms from their position in triple-helical $(ProHypGly)_{10}$ [1v7h (42)] or in $(ProProGly)_{10}$ (35) is 0.2 Å. This value holds for comparison with both 1k6f (35), our molecular replacement model, and 1a3i (43), an independent ProProGly structure, providing assurance against phasing model bias in our structure determination. Similar values have been reported for comparisons between various ProProGly and ProHypGly structures (43). This conservation of main-chain structure is evident from the congruence of ϕ, ψ, and ω with those in natural CRPs (Table 3).

Within the triple-helix crystal structure, the nonnatural amino acids mep and Flp adopt exclusively those side-chain or ring-pucker geometries that are favored by thermodynamics for the individual residues (Table S1 in the *SI Text*), namely mep as C^γ-endo (Fig. 6*A*) and Flp as C^γ-exo (Fig. 6*B*).

Discussion

The principle of preorganization has led to remarkable advances in host–guest chemistry, culminating in the work of Pedersen (44), Cram (14), Lehn (45), and others. This powerful principle has been used to stabilize the conformation of biopolymers, but only via the structurally disruptive engineering of their backbones. Here, we test the efficacy of an alternative strategy—modifying the side chain of a biopolymer to impose favorable conformational constraints upon its main chain.

Table 3. Structural parameters of $[(\text{mepFlpGly})_7]_3$ and other crystalline triple helices

Parameter	$[(\text{mepFlpGly})_7]_3$*	$[(\text{ProHypGly})_{10}]_3$†	$[(\text{ProProGly})_{10}]_3$‡
resolution (Å)	1.21	1.26	1.3
interstrand hydrogen bond (Å)	2.97 ± 0.06	2.88	3.04 ± 0.01
helical pitch	7/2	7/2	7/2
ϕ, Xaa position (deg)	-76.0 ± 4.1	-71.3 ± 1.4	-74.5 ± 2.9
ψ, Xaa position (deg)	164.7 ± 3.9	161.5 ± 1.1	164.3 ± 4.1
ω, Xaa position (deg)	180.12 ± 2.7	172.3 ± 1.0	176.0 ± 2.5
ϕ, Yaa position (deg)	$-60. \pm 1.5$	-56.9 ± 1.3	-60.1 ± 3.6
ψ, Yaa position (deg)	151.3 ± 3.6	150.0 ± 1.1	152.4 ± 2.6
ω, Yaa position (deg)	176.0 ± 1.3	174.7 ± 1.1	175.4 ± 3.4
ϕ, Gly (deg)	-71.5 ± 2.6	-71.3 ± 1.6	-71.7 ± 3.7
ψ, Gly (deg)	176.8 ± 4.3	174.2 ± 1.2	175.9 ± 3.1
ω, Gly (deg)	175.3 ± 2.3	178.8 ± 1.1	179.7 ± 2.0
rmsd (Å)§	—	0.2195	0.2320

*Values are averaged from the central thirty-six amino acids in the center of each triple helix in the asymmetric unit.

†From ref. 42.

‡From ref. 35.

§rmsd $= \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$, where n is the number of main-chain atoms (N, C^{α}, C', and O') in the central core 7 residues of each strand, and x is their absolute position.

Strands of collagen with the sequence $(\text{ProProGly})_7$ do not form a stable triple helix at 4 °C. We synthesized $(\text{flpMepGly})_7$ and $(\text{mepFlpGly})_7$ strands with subtle stereospecific changes to the prolyl side chains. Their triple helices have T_m values of 51 and 58 °C, respectively (Fig. 3 and Table 1). What is the origin of this extraordinary stability?

Previously, we used density functional theory calculations to demonstrate that Ac-mep-OMe and Ac-flp-OMe prefer the C^{γ}-endo ring pucker over the C^{γ}-exo ring pucker by 1.4 and 0.61 kcal/mol, respectively, whereas Ac-Mep-OMe and Ac-Flp-OMe prefer the C^{γ}-exo ring pucker by 1.7 and 0.85 kcal/mol, respectively (Table S1 in the *SI Text*) (30, 31). In contrast, Ac-Pro-OMe has only a slight preference for the C^{γ}-endo ring pucker (0.41 kcal/mol) as shown in Table S1 in the *SI Text* (30), whereas Hyp has a relatively weak preference for the C^{γ}-exo ring pucker (46). Further, crystalline Ac-Mep-OMe and Ac-Flp-OMe both display the C^{γ}-exo ring pucker, whereas crystalline Ac-Pro-OMe displays the C^{γ}-endo ring pucker (Table S1 in the *SI Text*) (47). Appropriate combinations of mep, Mep, flp, and Flp could, therefore, stabilize the triple helix by preorganization when in the Xaa and Yaa positions of triple helices, assuming that these positions require the ϕ and ψ angles provided by the C^{γ}-endo and C^{γ}-exo ring puckers, respectively.

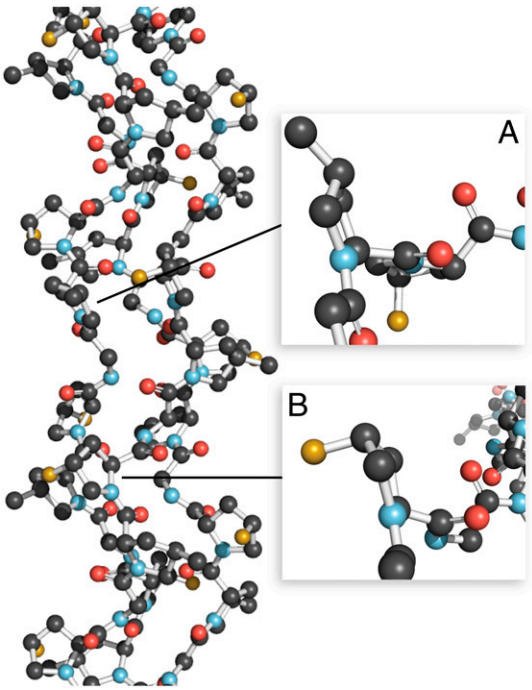

Fig. 6. Segment of a crystalline $[(\text{mepFlpGly})_7]_3$ triple helix. (*A*) C^{γ}-Endo ring pucker of a mep residue. (*B*) C^{γ}-Exo ring pucker of a Flp residue.

Preorganization affects the thermodynamics of protein folding by reducing the entropy of the unfolded state (Fig. 1). Preorganization can also have favorable and unfavorable enthalpic effects on protein folding. We turned to a calorimetric analysis of the folding–unfolding equilibria of the $[(\text{flpMepGly})_7]_3$ and $[(\text{mepFlpGly})_7]_3$ triple helices to provide insight on the underlying basis of their hyperstability, as a preorganizational origin would be manifested in a significant reduction in the entropic cost of triple-helix folding.

The results of our DSC analyses are summarized in Table 1. Our model for fitting the DSC data permits ΔC_p to be dependent on temperature, and assumes an association–dissociation equilibrium. We note that the same conclusions were obtained from baseline subtraction and integration of the excess C_p curve, though our model enhances accuracy and provides the ability to compare the results at the same temperature (39). For convenience, we chose to compare the thermodynamic parameters at $T = 46$ °C (which is midway between the highest and lowest T_m values), but our conclusions hold as well at other temperatures within the examined range. Values of ΔG for the unfolding of the triple helices increased in the order $[(\text{ProHypGly})_7]_3 < [(\text{flpMepGly})_7]_3 \approx [(\text{mepFlpGly})_7]_3$ (Table 1). The much more favorable $T\Delta S$ values for $[(\text{flpMepGly})_7]_3$ and $[(\text{mepFlpGly})_7]_3$ triple helices relative to the $[(\text{ProHypGly})_7]_3$ triple helix are consistent with the hyperstability being due to entropic stability provided by preorganization.

The less favorable ΔH values for $[(\text{flpMepGly})_7]_3$ and $[(\text{mepFlpGly})_7]_3$ triple-helix formation are likely due to the loss of polar hydration to side-chain hydroxyl groups. This loss of polar hydration is entropically favorable. Nonetheless, the hyperstability of $[(\text{flpMepGly})_7]_3$ and $[(\text{mepFlpGly})_7]_3$ triple helices cannot be attributed to this phenomenon, as the $(\text{ProProGly})_7$ CRP does not form a stable triple helix, despite an analogous loss of polar hydration. In addition, triple helices formed from the singly substituted CRPs $(\text{mepProGly})_7$, $(\text{ProMepGly})_7$, $(\text{flpProGly})_7$, and $(\text{ProFlpGly})_7$ (31, 38) (which all have a similar loss of polar hydration) are much less stable than those formed from $(\text{flpMepGly})_7$ and $(\text{mepFlpGly})_7$. Finally, the value of ΔC_p is similar in magnitude for all three triple helices, indicating that

the hyperstability is not due to an enhanced hydrophobic effect (34). Thus, we are compelled to attribute the hyperstability of $[(\text{flpMepGly})_7]_3$ and $[(\text{mepFlpGly})_7]_3$ triple helices to a high degree of structural preorganization.

Our conclusion that flp, mep, Flp, and Mep stabilize triple helices via preorganization relies on the assumption that these nonnatural proline derivatives do not alter the structure of the folded triple helix. Low-resolution structural data from CD spectroscopy are not sufficient to support such a conclusion. To address this critical issue, we obtained a high-resolution crystal structure of a $[(\text{mepFlpGly})_7]_3$ triple helix and compared that structure to those of two natural collagens (Table 3). Changes to the ϕ, ψ, and ω angles at each position were undetectable. We conclude that these side-chain-modified, nonnatural proline derivatives do not perturb triple-helix structure.

The crystal structure of a $[(\text{mepFlpGly})_7]_3$ triple helix demonstrates that the preferred ring puckers are indeed observed in the folded triple helix. Throughout the triple helix, mep residues display the expected C^γ-endo ring pucker (Fig. 6*A*); Flp residues display the C^γ-exo ring pucker (Fig. 6*B*). By biasing these preferred ring puckers for the Xaa and Yaa positions of the triple helix in the unfolded state by using steric and stereoelectronic effects, we were able to stabilize this multimeric protein structure by using side-chain modification to implement the principles of preorganization. Intriguingly, Pro, which slightly prefers the C^γ-endo pucker, is abundant in the Xaa position of natural collagen, whereas Hyp, which prefers the C^γ-exo pucker, is abundant in the Yaa position (33). Thus, Nature seems to have evolved a similar mechanism for stabilizing the collagen triple helix but does not produce the even more preorganized amino acids available through chemical synthesis.

Conclusion

We have employed the principle of preorganization to confer remarkable stability upon a multimeric protein. This approach yielded collagen triple helices with thermal stability greater than any other of similar length. As collagen is the most abundant protein in animals and constitutes 3/4 of human skin, such hyperstable collagens have potential as therapeutic agents.

Our findings are applicable to other proteins as well. The thermodynamic and crystallographic analyses reveal how proline derivatives can be employed favorably in other protein folds. Although we have used chemical synthesis, proline derivatives can be incorporated into proteins by heterologous expression by using proline auxotrophs of *Escherichia coli* or by semisynthetic methods (48–54). Derivatives other than 4-fluoroproline and 4-methylprolines can be used to bias ring pucker (55–57). The applicability is not limited to replacing proline residues—replacing other residues with appropriate ϕ and ψ torsion angles (Table 3 and Table S1 in the *SI Text*) can also be beneficial. Hence, the principle of preorganization of main chains by subtle changes to side chains provides a powerful and versatile means to enhance conformational stability.

Materials and Methods

Procedures for the synthesis of $(\text{ProProGly})_7$, $(\text{ProHypGly})_7$, $(\text{flpMepGly})_7$, and $(\text{mepFlpGly})_7$, and their analysis by CD spectroscopy and DSC, and procedures for the crystallization and structure determination of Ac-Mep-OMe are provided in *SI Text* on the PNAS web site.

Single crystals of $[(\text{mepFlpGly})_7]_3$ were grown at 4 °C by the hanging-drop vapor diffusion method from solutions containing HOAc and PEG-4000. The highest quality crystals were obtained from drops formed from 1 μL of ~10 mg/mL solution of $(\text{mepFlpGly})_7$ in 10% v/v HOAc(aq) and 1 μL of reservoir solution equilibrated against a reservoir containing 35% w/v PEG-4000 in water. Cubic crystals with average dimensions of $0.05 \times 0.05 \times 0.05$ mm grew at 4 °C in 10–14 d. These were vitrified without additional cryoprotectants.

X-ray diffraction data were collected in two 360° scans (0.5° images for 1 s and 1° images for 5 s) at 100 K on beamline ID-D at the Life Sciences Collaborative Access Team at the Advanced Photon Source. The data contained both very strong and very weak reflections, as had been observed previously for crystalline triple helices (58). Data were processed with HKL2000 (59), and initial phases were obtained by molecular replacement by using a truncated model of a crystalline $[(\text{ProProGly})_{10}]_3$ triple-helix structure determined by Berisio, Vitagliano, Mazzarella, and Zagari (Protein Data Bank entry 1k6f) (35). The $[(\text{mepFlpGly})_7]_3$ structure was refined initially with REFMAC5 (60, 61) and fitted manually to density maps with Coot (62). Refinement parameters for mep and Flp were prepared that did not constrain the proline ring pucker but allowed it to adopt the conformation that best fitted the electron density (Fig. S2). Residues in the central region of the triple helix were fitted readily to the electron density, although the ends of each helix were ordered poorly. Hence, we limited comparative calculations to the most well-ordered central four or five triplet repeats in each strand. Final rounds of refinement, including anisotropic *B* values, were carried out in Shelx97 (Table 2) (63).

ACKNOWLEDGMENTS. We thank Y.-S. Lin for both contributive discussions and practical aid regarding thermodynamic analyses. We also thank A. Choudhary, D.R. McCaslin, and F.W. Kotch for contributive discussions, and I.A. Guzei for solving the crystal structure of Ac-Mep-OMe. M.D.S. was supported by graduate fellowships from the US Department of Homeland Security and the Division of Medicinal Chemistry of the American Chemical Society. Experiments made use of the Biophysics Instrumentation Facility established by Grants BIR-0-9512577 (NSF) and S10 RR13790 (NIH), and the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 RR02301 (NIH). Use of the Advanced Photon Source was supported by the Office of Science, Office of Basic Energy Sciences of the US Department of Energy, under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology TriCorridor for the support of this research program (Grant 085P1000817). This work was supported by Grant AR044276 (NIH).

1. Srinivasan R, Rose GD (1999) A physical basis for protein secondary structure. *Proc Natl Acad Sci USA*, 96:14258–14263.
2. Eijsink VGH, et al. (2004) Rational engineering of enzyme stability. *J Biotechnol*, 113:105–120.
3. Tang Y, et al. (2001) Fluorinated coiled-coil proteins prepared in vivo display enhanced thermal and chemical stability. *Angew Chem Int Ed*, 40:1494–1496.
4. Horng J-C, Raleigh DP (2003) *F*-Values beyond the ribosomally encoded amino acids: Kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein. *J Am Chem Soc*, 125:9286–9287.
5. Garcia AD, Sanbonmatsu KY (2002) α-Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci USA*, 99:2782–2787.
6. Nicholson H, Becktel WJ, Matthews BW (1988) Enhanced protein thermostability from designed mutations that interact with α-helix dipoles. *Nature*, 336:651–656.
7. Yokum TS, Bursavich MG, Gauthier T, Hammer RP, McLaughlin ML (1998) 3_{10}-Helix stabilization via side-chain salt bridges. *Chem Commun* 1801–1802.
8. Olson CA, Spek EJ, Shi Z, Vologodskii A, Kallenbach NR (2001) Cooperative helix stabilization by complex Arg-Glu salt bridges. *Proteins: Struct Funct Genet*, 44:123–132.
9. Perry LJ, Wetzel R (1984) Disulfide bond engineered into T4 lysozyme: Stabilization of the protein toward thermal inactivation. *Science*, 226:555–557.
10. Matsumura M, Signor G, Matthews BW (1989) Substantial increase of protein stability by multiple disulfide bonds. *Nature*, 342:291–293.
11. Matsumura M, Becktel WJ, Levitt M, Matthews BW (1989) Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc Natl Acad Sci USA*, 86:6562–6566.
12. Kuroki R, et al. (1989) Design and creation of a calcium binding site in human lysozyme to enhance structural stability. *Proc Natl Acad Sci USA*, 86:6903–6907.
13. Liu L, Guo Q-X (2001) Isokinetic relationship, isoequilibrium relationship, and enthalpy-entropy compensation. *Chem Rev*, 101:673–696.
14. Cram DJ (1988) The design of molecular hosts, guests, and their complexes. *Science*, 240:760–767.
15. Chalikian TV, Völker J, Plum GE, Breslauer KJ (1999) A more unified picture for the thermodynamics of nucleic acid duplex melting: A characterization by calorimetric and volumetric techniques. *Proc Natl Acad Sci USA*, 96:7853–7858.
16. Holbrook JA, Capp MW, Saecker RM, Record MT Jr. (1999) Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: Interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry*, 38:8409–8422.
17. Koshkin AA, et al. (1998) LNA (locked nucleic acid): An RNA mimic forming exceedingly stable LNA:LNA duplexes. *J Am Chem Soc*, 120:13252–13253.
18. Kaur H, Arora A, Wengel J, Maiti S (2006) Thermodynamic, counterion, and hydration effects for the incorporation of locked nucleic acid nucleotides into DNA duplexes. *Biochemistry*, 45:7347–7355.
19. Haque TS, Gellman SH (1997) Insights on β-hairpin stability in aqueous solution from peptides with enforced type I′ and type II′ β-turns. *J Am Chem Soc*, 119:2303–2304.
20. Arnold U, et al. (2002) Protein prosthesis: A semisynthetic enzyme with a β-peptide reverse turn. *J Am Chem Soc*, 124:8522–8523.
21. Arnold U, et al. (2003) Protein prosthesis: A nonnatural residue accelerates folding and increases stability. *J Am Chem Soc*, 125:7500–7501.

22. Che Y, Marshall GR (2006) Impact of cis-proline analogs on peptide conformation. *Biopolymers*, 81:392–406.
23. Tam A, Arnold U, Soellner MB, Raines RT (2007) Protein prosthesis: 1,5-Disubstituted [1, 2, 3] triazoles as cis-peptide bond surrogates. *J Am Chem Soc*, 129:12670–12671.
24. Deechongkit S, et al. (2004) Context-dependent contributions of backbone hydrogen bonding to β-sheet folding energetics. *Nature*, 430:101–105.
25. Jenkins CL, Vasbinder MM, Miller SJ, Raines RT (2005) Peptide bond isosteres: Ester or (*E*)-alkene in the backbone of the collagen triple helix. *Org Lett*, 7:2619–2622.
26. Gao J, Kelly JW (2008) Toward quantification of protein backbone-backbone hydrogen bonding energies: An energetic analysis of an amide-to-ester mutation in an α-helix within a protein. *Protein Sci*, 17:1096–1101.
27. Dai N, Wang XJ, Etzkorn FA (2008) The effect of a trans-locked Gly-Pro alkene isostere on collagen triple helix stability. *J Am Chem Soc*, 130:5396–5397.
28. Matthews BW, Nicholson H, Becktel WJ (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci USA*, 84:6663–6667.
29. Ramachandran GN, Sasisekharan V, Ramakrishnan C (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol*, 7:95–99.
30. DeRider ML, et al. (2002) Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. *J Am Chem Soc*, 124:2497–2505.
31. Shoulders MD, Hodges JA, Raines RT (2006) Reciprocity of steric and stereoelectronic effects in the collagen triple helix. *J Am Chem Soc*, 128:8112–8113.
32. Vitagliano L, Berisio R, Mazzarella L, Zagari A (2001) Structural bases of collagen stabilization induced by proline hydroxylation. *Biopolymers*, 58:459–464.
33. Shoulders MD, Raines RT (2009) Collagen structure and stability. *Annu Rev Biochem*, 78:929–958.
34. Shoulders MD, Kamer KJ, Raines RT (2009) Origin of the stability conferred upon collagen by fluorination. *Bioorg Med Chem Lett*, 19:3859–3862.
35. Berisio R, Vitagliano L, Mazzarella L, Zagari A (2002) Crystal structure of the collagen triple helix model $[(\text{Pro-Pro-Gly})_{10}]_3$. *Protein Sci*, 11:262–270.
36. Holmgren SK, Taylor KM, Bretscher LE, Raines RT (1998) Code for collagen's stability deciphered. *Nature*, 392:666–667.
37. Holjmgren SK, Bretscher LE, Taylor KM, Raines RT (1999) A hyperstable collagen mimic. *Chem Biol*, 6:63–70.
38. Hodges JA, Raines RT (2003) Stereoelectronic effects on collagen stability: The dichotomy of 4-fluoroproline diastereomers. *J Am Chem Soc*, 125:9262–9263.
39. Nishi Y, et al. (2005) Different effects of 4-hydroxyproline and 4-fluoroproline on the stability of the collagen triple helix. *Biochemistry*, 44:6034–6042.
40. Johnson CR, Morin PE, Arrowsmith CH, Freire E (1995) Thermodynamic analysis of the structural stability of the tetrameric oligomerization domain of p53 tumor suppressor. *Biochemistry*, 34:5309–5316.
41. Bella J, Eaton M, Brodsky B, Berman HM (1994) Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science*, 266:75–81.
42. Okuyama K, et al. (2004) Crystal structures of collagen model peptides with Pro-Hyp-Gly repeating sequence at 1.26 Å resolution: Implications for proline ring puckering. *Biopolymers (Pept Sci)*, 76:367–377.
43. Kramer RZ, et al. (1998) X-ray crystallographic determination of a collagen-like peptide with the repeating sequence (Pro-Pro-Gly). *J Mol Biol*, 280:623–638.
44. Pedersen CJ (1988) The discovery of crown ethers (Noble Lecture). *Angew Chem Int Ed*, 27:1021–1027.
45. Lehn J-M (2002) Toward complex matter: Supramolecular chemistry and self-organization. *Proc Natl Acad Sci USA*, 99:4763–4768.
46. Improta R, Benzi C, Barone V (2001) Understanding the role of stereoelectronic effects in determining collagen stability. 1. A quantum mechanical study of proline, hydroxyproline, and fluoroproline dipeptide analogues in aqueous solution. *J Am Chem Soc*, 123:12568–12577.
47. Panasik N, Jr., Eberhardt ES, Edison AS, Powell DR, Raines RT (1994) Inductive effects on the structure of proline residues. *Int J Pept Protein Res*, 44:262–269.
48. Renner C, et al. (2001) Fluoroprolines as tools for protein design and engineering. *Angew Chem Int Ed*, 40:923–925.
49. Kim W, George A, Evans M, Conticello VP (2004) Cotranslational incorporation of a structurally diverse series of proline analogues in an *Escherichia coli* expression system. *ChemBioChem*, 5:928–936.
50. Nilsson BL, Soellner MB, Raines RT (2005) Chemical synthesis of proteins. *Annu Rev Biophys Biomol Struct*, 34:91–118.
51. Lummis SCR, et al. (2005) Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature*, 438:248–252.
52. Kim W, McMillan RA, Snyder JP, Conticello VP (2005) A stereoelectronic effect on turn formation due to proline substitution in elastin-mimetic polypeptides. *J Am Chem Soc*, 127:18121–18132.
53. Horng J-C, Raines RT (2006) Stereoelectronic effects on polyproline conformation. *Protein Sci*, 15:74–83.
54. Steiner T, et al. (2008) Synthetic biology of proteins: Tuning GFPs folding and stability with fluoroproline. *PLoS One*, 3:e1680.
55. Shoulders MD, Guzei IA, Raines RT (2008) 4-Chloroprolines: Synthesis, conformational analysis, and effect on the collagen triple helix. *Biopolymers*, 89:443–454.
56. Cadamuro SA, et al. (2008) Conformational properties of 4-mercaptoproline and related derivatives. *Angew Chem Int Ed*, 47:2143–2146.
57. Kotch FW, Guzei IA, Raines RT (2008) Stabilization of the collagen triple helix by *O*-methylation of hydroxyproline residues. *J Am Chem Soc*, 130:2952–2953.
58. Berisio R, Vitagliano L, Mazzarella L, Zagari A (2001) Crystal structure of a collagen-like polypeptide with repeating sequence Pro-Hyp-Gly at 1.4 Å resolution: Implications for collagen hydration. *Biopolymers*, 56:8–13.
59. Otwinowski Z, Minor W (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol*, 276:307–326.
60. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Cryst D*, 50:760–763.
61. Vagin AA, et al. (2004) REFMAC5 dictionary: Organisation of prior chemical knowledge and guidelines for its use. *Acta Cryst D*, 60:2284–2295.
62. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Cryst D*, 60:2126–2132.
63. Sheldrick GM (2008) A short history of SHELX. *Acta Cryst A*, 64:112–122.