Protein Engineering, Design & Selection vol. 24 no. 8 pp. 627–630, 2011
Published online June 13, 2011 doi:10.1093/protein/gzr024

SHORT COMMUNICATION

Protein stabilization in a highly knotted protein polymer

Tobias C.Sayre[1], Toni M.Lee[1], Neil P.King[1] and Todd O.Yeates[1,2,3]

[1]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA and [2]California Nanosystems Institute, University of California, Los Angeles, CA 90095, USA
[3]To whom correspondence should be addressed.
E-mail: yeates@mbi.ucla.edu

Received May 10, 2011; revised May 10, 2011; accepted May 11, 2011

Edited by David Eisenberg

The polypeptide backbones of a few proteins are tied in a knot. The biophysical effects and potential biological roles of knots are not well understood. Here, we test the consequences of protein knotting by taking a monomeric protein, carbonic anhydrase II, whose native structure contains a shallow knot, and polymerizing it end-to-end to form a deeply and multiply knotted polymeric filament. Thermal stability experiments show that the polymer is stabilized against loss of structure and aggregation by the presence of deep knots.
Keywords: biomaterials/disulfide/protein design/protein knot/topology

A few unusual proteins adopt folded configurations in which the polypeptide backbone is knotted—i.e. pulling the two termini would result in an internal knot (reviewed in Lua and Grosberg, 2006; Virnau *et al.*, 2006; King *et al.*, 2007; Taylor, 2007; Virnau *et al.*, 2011). The existence of such proteins in nature raises intriguing questions (Yeates *et al.*, 2007; Mallam, 2009; Shakhnovich, 2011). How do topologically complex proteins manage to reach their natively folded configurations? Experimental studies on a few different knotted proteins indicate that their folding pathways are complex, yet their folding rates are not prohibitively slow (Mallam and Jackson, 2005; Mallam and Jackson, 2006; Mallam and Jackson, 2007; Mallam *et al.*, 2008a; Andersson *et al.*, 2009; Mallam, 2009; King *et al.*, 2010). The physical mechanisms for knot formation remain largely unsettled. Computer simulations suggest that the knotted state tends to occur late in the folding process, typically when a terminal tail or looped segment passes through an opening created by other elements of the structure (Sulkowska *et al.*, 2009; Faisca *et al.*, 2010; Noel *et al.*, 2010); these events appear to be rate-limiting, though the introduction of favorable non-native interactions can dramatically speed folding (Wallin *et al.*, 2007). On the other hand, experiments suggest that loosely knotted topologies could form early during folding (Mallam *et al.*, 2008a, b), though it has been difficult in *in vitro* experiments to know whether protein molecules are being refolded from completely unfolded (i.e. fully extended and therefore unknotted) configurations (Mallam *et al.*, 2010).

Whether protein knots play important functional or stabilizing roles, or whether they simply represent accidental curiosities, is unclear. A few cases are suggestive. In the well-studied case of the knotted RNA methyltransferase family, the core of the knot helps form the binding site for the *S*-adenosyl methionine cofactor (Nureki *et al.*, 2002; Lim *et al.*, 2003). In the phytochrome protein, the knot sits where increased rigidity could help transmit light-driven conformational changes (Wagner *et al.*, 2005; Bornschlogl *et al.*, 2009). In alkaline phosphatase, a slip knot in the structure has been argued to contribute to that enzyme's unusual thermostability (King *et al.*, 2007). The smallest known protein that contains a slipknot is from a virus that infects thermophilic *Sulfolobus* archaebacteria (King *et al.*, 2007). The resistance of knotted proteins to mechanical unfolding has been examined by atomic force microscopy. For the knotted carbonic anhydrase enzyme, a much greater resistance to unfolding was observed when the pulling force would have led to knot tightening (Wang and Ikai, 1999), in contrast to a similar experiment in which pulling at different sites was able to extend the polypeptide chain into an unknotted configuration (Wang *et al.*, 2002). The mechanical stability of the phytochrome protein was also tested by single-molecule atomic force microscopy (Bornschlogl *et al.*, 2009). In that case, the extension force for unfolding (73 pN) was within the wide bounds set by normal (i.e. unknotted) proteins.

One challenge in investigating the role of knotting is constructing controls that make it possible to probe the effects of topology specifically (Wang *et al.*, 2002; King *et al.*, 2007, 2010; Yeates *et al.*, 2007; Mallam *et al.*, 2008b). In the present study, we investigated the effects of knotting in a protein after polymerizing it to form highly knotted protein filaments. The naturally knotted protein, human carbonic anhydrase II (HCAII), was chosen as the building block for polymerization. The native structure of HCAII is only shallowly knotted; just the very C-terminal end of the polypeptide backbone protrudes through a loop created by the rest of the protein (Fig. 1a). Because the knot is shallow, its effect on the stability of an individual protein molecule might be expected to be modest; the protein can unfold once the terminal few amino acids are unthreaded from the knot. However, if many protein molecules are connected together, the knots can be made much deeper, allowing the effects of knot depth to be analyzed.

Two topologically distinct polymers of HCAII were designed to specifically test the effects of topology on polymer stability: a 'serially knotted' polymer and a

© The Author 2011. Published by Oxford University Press. All rights reserved.
For Permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/peds/article/24/8/627/1507520 by University of Washington user on 23 October 2025

'superficially knotted' polymer (Fig. 1b). Polymerization was achieved by cysteine-based intermolecular disulfide bond formation under oxidizing conditions (Fig. 1), a strategy that has been used to study the unfolding of synthetic protein polymers in single-molecule force experiments (Dietz and Rief, 2006; Bornschlogl *et al.*, 2009). To create the building blocks for the two types of polymers studied here, two variants of the HCAII protein were constructed, each bearing two cysteine residues at strategic positions. To avoid competing reactions, these cysteines were placed in a form of HCAII from which the single native cysteine at residue 205 had been changed to serine. The first protein construct, used to build the serially knotted polymer, has two cysteine residues in short extensions at either end of the polypeptide chain. With cysteines at these positions, polymerization leads to a chain of protein subunits linked so that a continuous path through the polymer backbone contains multiple knots in series (Fig. 1b). In contrast to the shallowness of the knot in an individual monomeric subunit, the knots in this polymer can only be fully untied by unthreading great lengths of the polymer chain. The second protein construct, used to build the superficially knotted polymer, has one cysteine at the N-terminus of the protein and one at position 151, before the knot. Polymerization of this protein produces a chain of protein subunits, each of which bears a shallow knot, yet the continuous path through the polymer backbone is unknotted (Fig. 1b). This polymer served as a control, and allowed a direct assessment of the effects knot depth and complex topology have on the properties of the protein. It should be noted that the polymerization strategy offers no control over the polarity (i.e. head-to-tail vs. head-to-head or tail-to-tail connection) of the monomers. The polymers are therefore heterogeneous in that regard, but this does not affect considerations of topology or knotting.

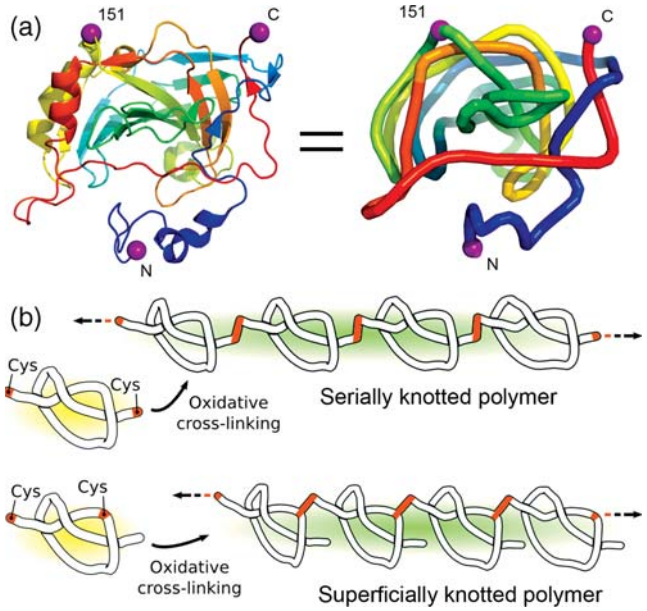

Fig. 1. Structure of HCAII and the polymerization of designed mutants into knotted polymer filaments. (**a**) A ribbon diagram of wild-type HCAII (PDB code 1lug). Highlighted are residue 151 and the terminal extensions, which were sites of cysteine insertion in different-designed constructs. On the right, a smoothed representation of the backbone clarifies the knotted topology (Norcross and Yeates, 2006). (**b**) Polymerization of two different protein constructs into different types of filaments. In the construct shown on the top, cysteines are present at both termini, and polymerization leads to a series of very deep knots. In the bottom construct, cysteines are present at the N-terminus and at residue 151, leading to a set of shallowly knotted proteins connected by a polymer backbone that is not knotted.

Each of the two engineered HCAII proteins was expressed recombinantly in *E. coli* and purified in the presence of dithiothreitol (DTT) to keep the protein in its reduced, monomeric form. Following oxidation, protein polymers of varying length were observed via sodium dodecyl sulfate polyacrylamide gel electrophoresis and size exclusion chromatography (Fig. 2). The degree of polymerization was similar for the two versions of the protein. The polymer chains ranged from dimers up to sizes too large to migrate through the gel, with the majority of material being longer than three subunits. The complexity of the gel migration patterns suggests the likely presence of both linear and cyclized forms of the polymers. Both proteins remained primarily monomeric under reducing conditions as expected.

Circular dichroism (CD) and aggregation were used to evaluate the thermal stability of the two polymerized samples. Attempts were not made to isolate species of specific length; the experiments therefore report on the behavior of polymers of mixed sizes. The room temperature CD spectra of the polymers were similar to the HCAII

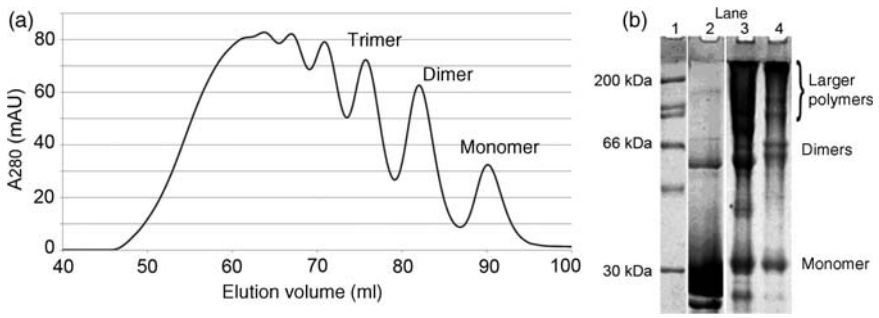

Fig. 2. Formation of HCAII polymers by oxidative cross-linking. (**a**) Size exclusion chromatography of the superficially knotted polymer. (**b**) SDS polyacrylamide gel electrophoresis of reduced (monomeric) and oxidized (polymerized) proteins. Lane 1—MW markers; lane 2—monomeric form (reduced); lane 3—superficially knotted polymer; lane 4—serially knotted polymer. In the polymerized forms, higher weight molecular species are evident at intervals above the monomeric form. Multiple conformations appear to be present for some oligomeric states, likely reflecting conformational heterogeneity, from partial knot retention in the presence of SDS for example. A minor fraction of the reduced sample appears as a dimeric form under the electrophoretic conditions. The molecular mass of the monomer, including the engineered tails, is 30.2 kDa. The reduced (monomeric) protein shown in panels a and b is the form used to construct the deeply knotted polymer (i.e. the form containing cysteines at both termini).

Downloaded from https://academic.oup.com/peds/article/24/8/627/1507520 by University of Washington user on 23 October 2025

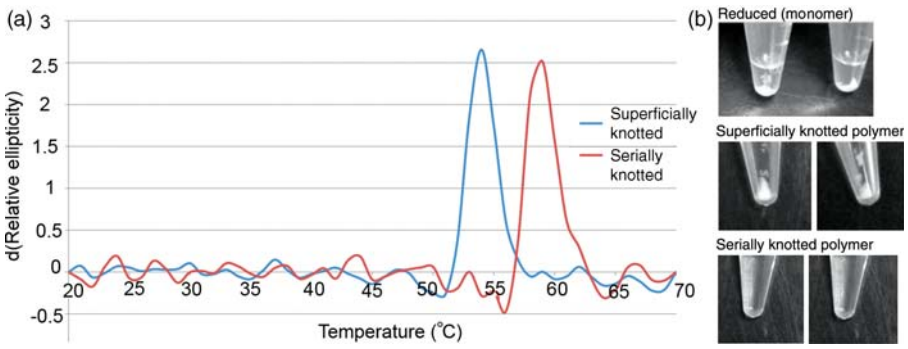

Fig. 3. Enhanced resistance of the serially knotted polymer to thermal denaturation. (**a**) First derivative melting curves are shown based on the CD signal at 210 nm. The melting transition peaks occur at 54°C for the superficially knotted polymer and at 59°C for the serially knotted polymer. (**b**) The serially knotted polymer shows enhanced resistance to thermal aggregation at 54°C. Shown in duplicate are samples of the monomer, the superficially knotted polymer, and the serially knotted polymer, after heating and centrifugation.

monomer, indicating that the individual subunits in the polymers were properly folded (not shown). The CD signal at 210 nm was used to monitor thermal melting at higher temperatures. The melting transition temperatures were evaluated from first-derivative curves of the spectra (Fig. 3). The melting temperature of the serially knotted polymer (T_m= 59°C) was clearly elevated compared with the superficially knotted polymer (T_m= 54°C). The difference in T_m was visualized by testing for aggregation. Samples were heated to 54°C for 5 min and then centrifuged. Only 5% of the superficially knotted polymer remained in solution. The result was similar for the monomeric HCAII. In contrast, 35% of the serially diluted protein remained in solution. The distinct melting and aggregation behaviors of the different constructs were reproducible over independent trials.

The findings highlight two key points. First, knotting in proteins can be a stabilizing feature, and the stabilizing effects of the knot can be enhanced by end-to-end polymerization of protein subunits, as this prevents the knots from becoming easily untied under denaturing conditions. The effects of knotting were evaluated separately from the effects of polymerization by designing a control experiment in which the same protein was polymerized in a way that did not produce multiple knots in series. Evidently, confining the overall mathematical topology of a protein chain stabilizes the native state against irreversible thermal denaturation. Studying the folding behavior of these polymers under reversible conditions might make it possible to extract thermodynamic parameters related to knotting. Second, the design experiments demonstrate a synthetic route to novel biomaterials with complex topologies and potentially useful features. The preliminary findings reported here serve to motivate further studies on topologically complex protein-based materials.

Materials and methods

Construction of wild-type and mutant HCAII genes

The wild-type HCAII gene, which codes for one cysteine at residue 205, was obtained in the pCMV-SPORT6 vector (Open Biosystems, accession number BC011949). The gene was cloned into the pET-M11 expression vector by amplifying the gene through polymerase chain reaction with Platinum Taq DNA polymerase (Invitrogen), performing a 2-h vector and insert double digestion, and ligating the insert between EcoRI and NcoI restriction sites at a 3:1 ratio of insert to vector. To make a construct containing cysteines at the two termini, cloning primers were designed to introduce a cysteine in a short extension having the sequence GGCAGGS at the N-terminus, following the $(His)_6$ tag and tobacco Etch virus (TEV) protease cleavage site encoded by the vector. A second cysteine was introduced in a short extension having the sequence SGGACGG at the C-terminus. The native cysteine at position 205 was removed using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene). Site-directed mutagenesis was also used to generate the second construct, which had the C-terminal cysteine removed, and a new cysteine placed at residue 151.

Protein expression, purification and polymerization

Plasmids containing wild-type or mutant HCAII genes were transformed into Rosetta 2(DE3) competent cells (Novagen). Cells were grown at 37°C to OD600 = 0.7 in Luria–Bertani media supplemented with 100 mg/l kanamycin before 0.5 mM isopropyl β-D-1-thiogalactopyranoside was added to induce protein expression. After five or more hours of expression at 20°C, cells were harvested by centrifugation at 7700 × *g* for 10 min, resuspended in lysis buffer (25 mM TRIS, 25 mM MOPS pH 7.2, 0.1% Tween 20) supplemented with protease inhibitor cocktail (Sigma), and lysed by sonication. After clearing lysates by centrifugation at 30 000 × *g* for 30 min, HCAII proteins were purified by immobilized metal affinity chromatography (IMAC) using a HisTrap column (GE Healthcare) and eluted with a linear gradient of imidazole up to 500 mM in a buffer containing 50 mM TRIS pH 7.2, 250 mM NaCl and 1 mM DTT. Fractions containing pure protein were pooled, diluted to 1 mg/ml, and dialyzed into 50 mM TRIS pH 8.0, 75 mM NaCl, 0.5 mM EDTA, 1 mM DTT. The N-terminal $(His)_6$ tag was cleaved off by addition of TEV protease (Kapust *et al.*, 2001) prepared in-house, to a ratio of 1 mg TEV:15 mg protein and incubation at room temperature overnight. Cleaved HCAII proteins were purified away from TEV protease and $(His)_6$ tags by a second round of IMAC, in which the cleaved proteins either flowed through the column, or eluted at a low imidazole concentration (probably due to the Zn-binding site in HCAII binding with low affinity to Ni^{2+} on the column). Fractions containing pure, cleaved HCAII proteins were pooled and dialyzed into 30 mM TRIS pH 7.2, 30 mM NaCl, 10 μM ZnCl2 before concentration to

Downloaded from https://academic.oup.com/peds/article/24/8/627/1507520 by University of Washington user on 23 October 2025

~15 mg/ml using Amicon centrifugal filter devices (Millipore). Protein concentrations were determined by measuring the absorbance of the solution at 280 nm, and using extinction coefficients calculated by ExPASy's protparam tool. Polymerization was achieved by the addition of 5 mM $CuSO_4$ to concentrated proteins, followed by extensive dialysis against 30 mM TRIS pH 7.2, 30 mM NaCl, 10 μM $ZnCl_2$ to remove the copper.

Thermal stability/aggregation assay

To test for differences in aggregation, both polymerized and reduced proteins were diluted to 0.5 mg/ml in 30 mM TRIS pH 7.2, 30 mM NaCl, 10 μM $ZnCl_2$, and heated to 54°C for 5 min, followed by centrifugation at 18 000 × *g* for 10 min to pellet aggregated material. Concentrations of soluble protein before and after heating/pelleting were determined by measuring the absorbance of the solution (or supernatant) at 280 nm. Samples were inspected visually to verify the presence or absence of pellets.

Circular dichroism

Far-UV CD spectra were recorded using a Jasco J-715 spectrometer with a Peltier temperature-controlled cell holder (Jasco). Samples were prepared by dilution of concentrated protein to a final concentration of 0.25 mg/ml in 30 mM TRIS pH 7.2, 30 mM NaCl, 1 mM DTT. Scans were taken between 200 and 260 nm at a temperature of 20°C and a scan rate of 20 nm/min using a bandwidth of 1 nm.

Acknowledgments

The authors would like to thank Dr Martin Phillips and Inna Pashkov for technical assistance, and Jesse Dill and Prof. Susan Marqusee for useful discussions. This work was supported by award R01GM081652 from the National Institutes of Health.

Conflict of interest

None declared.

References

Andersson,F.I., Pina,D.G., Mallam,A.L., Blaser,G. and Jackson,S.E. (2009) *FEBS J.*, **276**, 2625–2635.
Bornschlogl,T., Anstrom,D.M., Mey,E., Dzubiella,J., Rief,M. and Forest,K.T. (2009) *Biophys. J.*, **96**, 1508–1514.
Dietz,H. and Rief,M. (2006) *Proc. Natl Acad. Sci. USA*, **103**, 1244–1247.
Faisca,P.F., Travasso,R.D., Charters,T., Nunes,A. and Cieplak,M. (2010) *Phys. Biol.*, **7**, 16009.
Kapust,R.B., Tözsér,J., Fox,J.D., Anderson,D.E., Cherry,S., Copeland,T.D. and Waugh,D.S. (2001) *Protein Eng.*, **14**, 993–1000.
King,N.P., Jacobitz,A.W., Sawaya,M.R., Goldschmidt,L. and Yeates,T.O. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 20732–20737.
King,N.P., Yeates,E.O. and Yeates,T.O. (2007) *J. Mol. Biol.*, **373**, 153–166.
Lim,K., Zhang,H., Tempczyk,A., Krajewski,W., Bonander,N., Toedt,J., Howard,A., Eisenstein,E. and Herzberg,O. (2003) *Proteins*, **51**, 56–67.
Lua,R.C. and Grosberg,A.Y. (2006) *PLoS Comput. Biol.*, **2**, 350–357.
Mallam,A.L. (2009) *FEBS J.*, **276**, 365–375.
Mallam,A.L. and Jackson,S.E. (2005) *J. Mol. Bio.l*, **346**, 1409–1421.
Mallam,A.L. and Jackson,S.E. (2006) *J. Mol. Biol.*, **359**, 1420–1436.
Mallam,A.L. and Jackson,S.E. (2007) *J. Mol. Biol.*, **366**, 650–665.
Mallam,A.L., Morris,E.R. and Jackson,S.E. (2008a) *Proc. Natl Acad. Sci. USA*, **105**, 18740–18745.
Mallam,A.L., Onuoha,S.C., Grossmann,J.G. and Jackson,S.E. (2008b) *Mol. Cell*, **30**, 642–648.
Mallam,A.L., Rogers,J.M. and Jackson,S.E. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 8189–8194.
Noel,J.K., Sulkowska,J.I. and Onuchic,J.N. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 15403–15408.
Norcross,T.S. and Yeates,T.O. (2006) *J. Mol. Biol.*, **362**, 605–621.
Nureki,O., Shirouzu,M., Hashimoto,K., *et al*. (2002) *Acta. Crystallogr. D. Biol. Crystallogr.*, **58**, 1129–1137.
Shakhnovich,E. (2011) *Nat. Mater.*, **10**, 84–86.
Sulkowska,J.I., Sulkowski,P. and Onuchic,J. (2009) *Proc. Natl Acad. Sci. USA*, **106**, 3119–3124.
Taylor,W.R. (2007) *Comp. Chem. Bioinformatics.*, **31**, 151–162.
Virnau,P., Mirny,L.A. and Kardar,M. (2006) *PLoS Comput Biol*, **2**, 1074–1079.
Virnau,P., Mallam,A. and Jackson,S. (2011) *J. Phys. Condens. Matter*, **23**, 033101.
Wagner,J.R., Brunzelle,J.S., Forest,K.T. and Vierstra,R.D. (2005) *Nature*, **438**, 325–331.
Wallin,S., Zeldovich,K.B. and Shakhnovich,E.I. (2007) *J. Mol. Biol.*, **368**, 884–893.
Wang,T. and Ikai,A. (1999) *Jpn. J. Appl. Phys.*, **38**, 3912–3917.
Wang,T., Arakawa,H. and Ikai,A. (2002) *Ultramicroscopy*, **91**, 253–259.
Yeates,T.O., Norcross,T.S. and King,N.P. (2007) *Curr. Opin. Chem. Biol.*, **11**, 595–603.

Downloaded from https://academic.oup.com/peds/article/24/8/627/1507520 by University of Washington user on 23 October 2025