Topology Engineering of Proteins in Vivo Using Genetically Encoded, Mechanically Interlocking SpyX Modules for Enhanced Stability

Dong Liu,†‡ Wen-Hao Wu,†‡ Ya-Jie Liu,† Xia-Ling Wu,‡ Yang Cao,† Bo Song,§ Xiaopeng Li,§ and Wen-Bin Zhang*†‡

†Key Laboratory of Polymer Chemistry & Physics of Ministry of Education, Center for Soft Matter Science and Engineering, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, P. R. China
‡Department of Chemistry, University of South Florida, Tampa, Florida 33620, United States

ABSTRACT: Recombinant proteins are traditionally limited to linear configuration. Herein, we report in vivo protein topology engineering using highly efficient, mechanically interlocking SpyX modules named AXB and BXA. SpyX modules are protein domains composed of p53dim (X), SpyTag (A), and SpyCatcher (B). The p53dim guides the intertwining of the two nascent protein chains followed by autocatalytic isopeptide bond formation between SpyTag and SpyCatcher to fulfill the interlocking, leading to a variety of backbone topologies. Direct expression of AXB or BXA produces protein catenanes with distinct ring sizes. Recombinant proteins containing SpyX modules are obtained either as mechanically interlocked obligate dimers if the protein of interest is fused to the N- or C-terminus of SpyX modules, or as star proteins if the protein is fused to both N- and C-termini. As examples, cellular syntheses of dimers of (GB1)2 (where GB1 stands for immunoglobulin-binding domain B1 of streptococcal protein G) and of four-arm elastin-like star proteins were demonstrated. Comparison of the catenation efficiencies in different constructs reveals that BXA is generally much more effective than AXB, which is rationalized by the arrangement of three domains in space. Mechanical interlocking induces considerable stability enhancement. Both AXB and BXA have a melting point ∼20 °C higher than the linear controls and the BXA catenane has a melting point ~2 °C higher than the cyclic control BX. Notably, four-arm elastin-like star proteins demonstrate remarkable tolerance against trypsin digestion. The SpyX modules provide a convenient and versatile approach to construct unconventional protein topologies via the “assembly-reaction” synergy, which opens a new horizon in protein science for stability enhancement and function reinforcement via topology engineering.

INTRODUCTION

In synthetic polymers, chain topology and molecular architecture have been important molecular parameters to tune their physical properties and functions. Controlling macromolecular topology is thus one of the most commonly practiced molecular engineering approaches. To a certain extent, it compensates for the lack of sequence control in traditional polymers. An extraordinary example is the star polymer, which emerged back in the early days of polymer science. Multiple polymer chains emanating from a central focal point constitute the basic structure of star polymers. This unique topology imparts intriguing properties to star polymers, including low viscosity, multiple surface functional groups, low crystallinity, etc. Nowadays, synthetic star polymers can be prepared with delicate control over the arm number, arm composition, molecular weight, polydispersity, etc. Sophisticated architectures with heterogeneous compositions are not uncommon. By contrast, recombinant proteins are usually synthesized by the cellular machinery in an extremely precise fashion. The stereochemistry, sequence, chain length, and even 3D folded structure are all determined on the genetic level and faithfully reproduced in the biological system. The topology is, however, usually restrained to linear configuration. To date, there are only a few examples of nonlinear proteins, including cyclic proteins, tadpoles, stars, H-shape proteins, and catenanes, yet its full potential remains to be explored. It is thus highly desired to further expand the topological dimension for protein engineering.

The significance of controlling the composition and topology of the protein backbone has been increasingly recognized. A purely synthetic approach to “backbone engineering” has been developed via introducing compositions other than amide bonds (e.g., esters and triazole linkages) into the polypeptide backbone to tune their assembly and properties. Knotted topology has been found to bring considerable stabilizing effects with completely native polypeptide backbones. Topological variations beyond linear configuration have included cyclics, tadpoles, stars, and other branched proteins. Among them, cyclic topology is the most commonly encountered nonlinear protein topology in nature and has also been extensively studied to date. Various technologies...
have been developed to prepare cyclic proteins, both *in vivo* and *in vitro*, including intein technology, SpyTag-SpyCatcher chemistry, sortase/butlase-mediated ligation, etc. Typical benefits from backbone cyclization include the dramatically improved stability against both thermal denaturation and proteolytic digestion. By contrast, there is only scarce reporting on star proteins and other branched proteins in the literature. A related example is the synthetic star polypeptides that exhibit robust and potent antimicrobial activities. This is in line with their highly exposed multivalent surface functionalities. Taking advantage of the highly reactive SpyTag-SpyCatcher chemistry, isomeric three-arm and four-arm star proteins have been prepared with high efficiency *in vitro*. However, direct expression of star proteins *in vivo* remains a challenge.

Supramolecular architecture is another important, yet less explored topic in protein science. Nature has elegantly demonstrated the power of supramolecular structure engineering in examples such as the viral capsids where the stability is greatly enhanced by mutually interlocked protein catenanes. Dawson et al. reported the solid state synthesis of a peptide catenane by native chemical ligation and carefully studied the thermodynamics of protein catenation. Recently, protein catenation has also been achieved *in vivo* based on an elastin-like model protein, facilitated by p53dim-guided chain intertwining and SpyTag-SpyCatcher ligation. Being genetically encoded, this method can be adapted to other proteins of interest as well. Protein catenation brings in significant improvement in thermal stability and proteolytic resistance.

Nevertheless, the catenation efficiency was only moderate (≈60%) in previous reports. We thus seek ways to further improve the catenation efficiency and create even more complex protein topologies.

In this contribution, we report the design and application of genetically encoded SpyX modules for the efficient cellular synthesis of diverse mechanically interlocked protein architectures including obligate dimers and star proteins (Figure 1). The topology was demonstrated by partial digestion experiments, and the catenation efficiency was evaluated for different constructs. In general, the SpyX modules have high catenation efficiencies even in the presence of fusion proteins at either end of the N- and C-termini or both. The consequences of topology engineering have also been examined. It was found that proteins with complex topologies exhibit significantly improved stability against thermal denaturation and trypsin digestion.

### RESULTS AND DISCUSSION

**Molecular Design of SpyX Modules.** The design of SpyX modules is based on p53dim (X, which forms an entwined dimer upon folding) and a genetically encoded SpyTag (A) and SpyCatcher (B) reactive pair (which spontaneously form an isopeptide bond upon reconstitution). Depending on their relative locations, there are two configurations, namely, AXB and BXA (Figure 1). To facilitate catenation, we consistently added short linkers with sequence of GSGS before X and of SGGSG after X in both constructs (Figures S1 and S2). Upon expression *in vivo*, the X domain shall form entwined dimers, bringing A and B in vicinity to react and form the covalent catenanes. Hence, AXB is expected to have concatenated rings with relatively smaller sizes, and BXA would have almost the entire SpyCatcher domain concatenated. Since the A-B complex is tightly folded, the overall sizes of AXB and BXA should be similar. They all possess two N-termini and two C-
termini, allowing the preparation of diverse mechanically interlocked protein topologies. When fusion proteins are tethered at one single terminus, obligate dimers will be obtained. When fusion proteins are tethered on both termini, four-arm star proteins will be produced (Figure 1). To probe the topology, a proteolytic site recognized and cleaved by tobacco etch virus (TEV) protease was included in both designs near the X domain. Two mutants were also designed as controls. The K344P mutant of the p53dim domain (X, where the lysine at position 344 is mutated to proline) could not form the entwined dimer and exists as a monomer. The D/A mutant of SpyTag (A) abolishes the chemical reactivity by changing the aspartic acid to nonreactive alanine (Figures S1 and S2). Therefore, AX′B′ and BX′A′ are expected to give cyclic monomer controls whereas AX′B and BX′A serve as linear monomer controls. We anticipated that by placing SpyTag, SpyCatcher, and p53dim in close neighborhood, the catenation efficiency can be dramatically improved. The resulting SpyX domains have relatively small overall sizes (∼24 kDa) and thus may serve as mechanically interlocking protein modules.

**Protein Catenanes from SpyX Modules.** The genes were designed, optimized, synthesized, and cloned to pQE80L vector. The plasmids were used to transform *Escherichia coli* BL21 strain and the expressions were all performed in 2XYT medium at 30 °C with IPTG induction. Purifications were performed using affinity chromatography following the standard protocol. The products were analyzed by size-exclusion chromatography (SEC) and SDS–PAGE. The characterization results of AXB and BXA as well as their mutants are summarized in Figure 2. It is clear that both AXB and BXA were mainly obtained as dimers in the crude products (Figure 2A). For AXB, there are some chain-extended oligomers, whereas, for BXA, it is almost exclusively dimer. After purification with SEC, the dimers of AXB and BXA show up as monomodal peaks at retention volumes much lower than the cyclic monomer mutants (AX′B and BX′A) (Figures 2B and 2C). Interestingly, the linear monomer mutants (AX′B and BX′A) elute at almost the same time as the catenanes. It indicates that, in solution, the linear mutants exist as dimers due to the entwining of X and the complexation of A′ and B′ regardless of covalent bond formation. It is also the preorganization of A and B by X that promotes the highly efficient catenation. The MALDI-TOF mass spectra also corroborated the dimer formation (Figures 2D and S3). For AXB, it gives mainly a molecular peak of [M + H]^+ at m/z of 47163 that is consistent with the expected value of 47097 for AXB dimer within the range of error. The A′XB and A′XB′ both show mainly one molecular peak assigned as [M + H]^+ at m/z of 23523 and 23524, respectively. For BXA, the molecular peak of [M + H]^+ at m/z of 49265 was only observed at relatively low abundance as compared to the major peak assigned as [M + 2H]^+ at m/z of 24646. We speculated that BXA may have strong binding affinity to proton and, thus, mainly forms doubly charged ions upon MALDI ionization.

The dimer formation is also evident in SDS–PAGE for both AXB and BXA (Figure 2E). Theoretically, assuming that the linkers between A and X and between X and B are sufficiently long and flexible, there will be four possible products, namely, [2]catenane, cyclic dimer, knotted cyclic dimer, and non-covalently associated rings (Figure S4). Among them, only the last one breaks into cyclic monomers in SDS–PAGE. We used gel densitometry to quantify the relative abundance of possible products. For AXB, the dimer is a clean band, accounting for ~92% of the mixture. There is much more cyclic monomer (~6%) than linear monomer (~1%). For BXA, the dimer band is smeared and there seem to be two distinct bands which represent ~61% and ~31% of the total mixture, respectively. The smearing could be due to a highly stable folded structure that resists denaturation even in the presence of SDS. While products of different topology may also contribute to multiple smearing bands, we will show in the following section by limited TEV digestion experiments that it is likely. Compared to AXB, the residual BXA monomer exists almost...
exclusively as cyclic monomers (∼8%). It indicates that the monomeric rings are the major side product in both cases, which could result either from premature ring closure or from the formation of noncovalently associated rings after chain entwining. The catenation efficiency is thus evaluated as the product of dimer contents determined individually from SEC profile and SDS–PAGE. It is (64% × 92%) for AXB and (96% × 92%) for BXA, respectively.

We further performed electrospray ionization mass spectrometry (ESI-MS) characterization combined with traveling wave ion mobility (TWIM) separation, which is useful to differentiate species with identical m/z but different charges and sizes. This powerful technique has been used to study supramolecular aggregates, peptide self-assembly, protein conformation, etc. In this work, the proteins were first ionized and the ionic species were subjected to ion mobility separation by their mass-to-charge (m/z) ratios and sizes/charges. The full ESI mass spectra of AXB and BXA with charge states denoted are shown in Figure 3. It should be noted that the signals of trace monomers may superimpose within that of dimers. After ion mobility separation, the dimers were indeed found to be the dominant species for both AXB and BXA (Figures 3C and 3D). The monomers were only detected in trace abundance. The signals of monomers and dimers extracted from TWIM-MS spectra are shown in Figure S5. The results are consistent with the findings in SDS–PAGE, again confirming high catenation efficiencies.

To prove that the dimers are catenanes rather than cyclic dimers, we performed partial digestion experiments. Since TEV cleavage site is placed inside the concatenated ring, treating proteins with TEV protease would first lead to both cyclic monomer and linear monomer upon limited proteolysis and further to linear monomer upon complete digestion. Meanwhile, digestion of cyclic monomer controls and linear monomer controls were also performed for comparison. The appearance of cyclic monomer upon partial digestion is a clear evidence of the interlocked catenane topology in both cases (Figures 2F and S6). By contrast, the cyclic and linear controls were converted to relinearized monomers and fragments, respectively, under identical conditions. The same process was also followed by UPLC–ESI-MS (Figures S7 and S8). Taking AXB for example, the purified sample shows up as a main peak at 47089, which matches the expected value of 47097 for catenane. Consistent with SDS–PAGE analysis, there are two very minor peaks with molecular weights of 23544 and 23562, which can be assigned as the cyclic and linear monomers, respectively. Upon partial digestion, the amounts of cyclic and linear monomers increased rapidly, especially the linear one. Eventually, only the peak assigned to the linear monomer can be observed (Figure S7). Throughout the process, we could not observe any peaks corresponding to the linear dimer with molecular weight of 47114, which is the partially digested product of both simple and knotted cyclic dimers. The UPLC–ESI-MS analysis of the partially digested products of BXA also gives similar results (Figure S8). It suggests that the cyclic dimers exist, if any, in extremely low abundance. Other indirect evidence comes from the comparison between the AXB dimer and AX'B dimer in SDS–PAGE (Figure 2E, lanes 2 and 4). The dimer of AX'B is most likely a cyclic dimer, but it appears in a distinct position as compared to the dimer of AXB. So the existence of cyclic dimers in AXB is less likely. Therefore, both AXB and BXA are indeed mechanically interlocked protein catenanes.

From Obligate Dimers to Star Proteins. Having established that AXB and BXA are both very efficient in forming protein catenanes, we further investigated how the introduction of various proteins would affect the catenation efficiency. The protein can be added either to the N-terminus

Figure 3. ESI-MS characterizations of (A) AXB and (B) BXA and two-dimensional TWIM-MS plot for (C) AXB and (D) BXA. For both samples, two species, namely, monomer and dimer, were observed with dimer being the overwhelming majority. The brightness of the color represents the abundance of signals at different charge states.
or to the C-terminus of SpyX modules, or both. As an example, immunoglobulin-binding domain B1 of streptococcal protein G, a small folded protein known as GB1, has been used as the model protein.\textsuperscript{50} Two consecutive domains have been added, leading to four constructs: (GB1)\textsubscript{2}-AXB, AXB-(GB1)\textsubscript{2}, (GB1)\textsubscript{2}-BXA, BXA-(GB1)\textsubscript{2} (Figure S9). The proteins were expressed under the same conditions as AXB and BXA. The SEC overlay of the crude expression products is shown in Figure 4A. It can be seen that both AXB and BXA effectively lead to the formation of dimers and the dimer content in the crude product is already very high (up to 97\% in BXA-(GB1)\textsubscript{2}). Since GB1 protein is known to exist as monomer, the products obtained here can be regarded as obligate dimers of (GB1)\textsubscript{2}. After SEC purification, the products show up as symmetric monomodal peaks, suggesting uniform composition. The SDS−PAGE analysis reveals that there are also trace amounts of monomers in the products, a scenario similar to that in AXB or BXA. The catenation efficiencies are summarized in Table 1, which are generally good. The MALDI-TOF mass spectra further prove the dimer formation (Figure S10). For example, AXB-(GB1)\textsubscript{2} shows a major molecular peak [M]\textsuperscript{+} of 74356, which matches well with the calculated value of 74341 within the range of error. The results suggest that the incorporation of folded protein domains, either to the N- or to the C-terminus, does not affect the mechanical interlocking of SpyX modules, leading to the formation of obligate dimers. The results have more implications for engineering protein activities that depend critically on the dimerization states, such as 3CL protease,\textsuperscript{51} or for creating multifunctional proteins that can be used to make protein-based hydrogels.\textsuperscript{52}

We further challenge the capability of SpyX modules to mechanically interlock fusion proteins by installing elastin-like protein domains onto both termini of the SpyX domain. Elastin-like proteins are intrinsically disordered proteins (IDPs) with random-coil-like conformations in solution.\textsuperscript{53−55} Unlike folded proteins such as GB1, IDPs usually have no stable ordered secondary structure in solution.\textsuperscript{56,57} They generally have much larger hydrodynamic volume than folded protein counterparts. Hence, they are expected to exert considerable steric hindrance on catenation owing to the excluded volume effects. They are ideal models for evaluating the influence of fusion proteins on catenation efficiency. An ELP with molecular weight of ~8 kDa was attached to each of the terminus. Since the center core is tightly folded, it is reasonable to consider the products as 4-arm star polymers. Two star protein constructs were designed and denoted as EAXBE and EBXAE, respectively (Figures 1, S11, and S12). Their cyclic and linear monomer controls were also designed using the KP mutant of p53dim (X\textsuperscript{′}) as well as the DA mutant of SpyTag (A\textsuperscript{′}). Unlike A\textsuperscript{′}XB and BXA\textsuperscript{′}, the linear controls in this case (EA\textsuperscript{′}XBE and EBX\textsuperscript{′}AE) are incapable of forming intertwined dimers in solution.

![Figure 4. Molecular characterizations of mechanically interlocked obligate dimers and star polymers: (A) SEC overlay of the crude expression products and (B) purified products of (GB1)\textsubscript{2}-AXB, AXB-(GB1)\textsubscript{2}, (GB1)\textsubscript{2}-BXA, BXA-(GB1)\textsubscript{2}. (C) SDS−PAGE analysis of purified products of (GB1)\textsubscript{2}-AXB, AXB-(GB1)\textsubscript{2}, (GB1)\textsubscript{2}-BXA, BXA-(GB1)\textsubscript{2}. (D) SEC overlay of the crude expression products of EAXBE and EBXAE. (E) SEC overlay and (F) SDS−PAGE analysis of EAXBE, EA\textsuperscript{′}XBE, EA\textsuperscript{′}XBE, EBXAE, EBX\textsuperscript{′}AE, EBX\textsuperscript{′}AE.]

### Table 1. Summary of Protein Catenation Efficiencies

| Samples       | SEC\textsuperscript{a} (%) | SDS\textsuperscript{b} (%) | Total\textsuperscript{c} (%) |
|---------------|----------------------------|----------------------------|-------------------------------|
| AXB           | 64                         | 92                         | 59                            |
| BXA           | 96                         | 92                         | 88                            |
| AXB-(GB1)\textsubscript{2} | 82                      | 80                         | 66                            |
| BXA-(GB1)\textsubscript{2} | 97                      | 86                         | 83                            |
| (GB1)\textsubscript{2}-AXB | 78                      | 93                         | 83                            |
| (GB1)\textsubscript{2}-BXA | 92                      | 83                         | 76                            |
| EAXBE         | 84                         | 85                         | 71                            |
| EBXAE         | 86                         | 89                         | 77                            |

\textsuperscript{a}Calculated from SEC profile based on the integration area of the dimer peak. \textsuperscript{b}Calculated from the SDS−PAGE analysis based on gel densitometry analysis. \textsuperscript{c}The overall catenation efficiency.
The crude products of EAXBE and EBXAE were first analyzed by SEC. Although their molecular weights are identical, the samples elute at distinct retention volumes, suggesting distinct hydrodynamic volumes in solution. After SEC purification, the monomodal peaks of these star proteins were compared to the corresponding cyclic and linear controls (Figure 4E). The results clearly demonstrate that both EAXBE and EBXAE are dimers appearing at much lower retention volume relative to the cyclic and linear mutants. Interestingly, the elution profiles of EAXBE and EA’X’BE overlap with each other, whereas the retention volume of EBX’AE is obviously larger than that of EBX’A’E. The SDS–PAGE analysis of the purified products reveals that both EAXBE and EBXAE are dimers with apparent molecular weights at ~100 and ~120 kDa, respectively, whereas the cyclic controls appear at ~40 kDa and the linear control appears at ~50 kDa. Although both proteins are 4-arm stars with identical molecular weights, there are still minor differences between their topologies. The difference is consistent with the SEC profiles, suggesting that BXA has a more expanded, much larger structure than AXB. This could be rationalized based on the arm distribution on the SpyX core, which will be further elaborated later. In MALDI-TOF mass spectra, only one peak could be observed for both EAXBE and EBXAE, which are assigned as the molecular peak with two charges [M + 2H]2+ (Figure S13). This is probably because the products may have strong binding affinity to proton and, thus, mainly form doubly charged ions upon MALDI ionization. It may also have something to do with the mechanically interlocked structure where each of the two noncovalently interlocked segments bears one charge. A similar phenomenon has been observed previously. This evidence demonstrates that star proteins can be effectively expressed in vivo using SpyX as an interlocking module to create the core.

Discussions on Catenation Efficiency and the Effects of Topology Engineering. We have shown that SpyX modules (AXB and BXA) are highly efficient in forming mechanically interlocked architectures regardless of the tethering position of the fused proteins. But the catenation efficiency varies as a direct consequence of the 3D arrangement of A, X, and B. In general, BXA is much more efficient than AXB in guiding the mechanical interlocking of protein chains. The dimer content in the crude products can be almost quantitative (up to 97% for BXA-(GB1)2). We analyzed the crystal structures of X (PDB: 1SAK) and A–B complex (PDB: 4MLS) and proposed possible models for AXB and BXA as shown in Figure 5. The chain topology of p53dim and SpyTag/SpyCatcher complex mandates their configuration in three-dimensional (3D) space. For p53dim, the intertwined structure makes the N-terminus and C-terminus be at same side of the folded structure. For A–B complex, since SpyTag and the first strand of SpyCatcher are arranged in parallel β-sheet configuration, the N-terminus of SpyTag and the N- and C- termini of SpyCatcher domain are crowded on the one end of the β-barrel. This is in perfect alignment with the configuration of BXA (Figure 5), leading to a more compact structure and a better positioning for efficient reaction. The catenation efficiency is thus much higher. However, the configuration of AXB puts A and B in antiparallel arrangement. It would require either one to fold back in the opposite direction for reconstitution and reaction (Figure 5). This slows down the reaction and decreases the catenation efficiency. The trend is consistent for all fusion proteins. For AXB alone, the slower reaction between A and B leads to the accumulation of telochemicalics, and thus, the intermolecular reactions are promoted to give more chain-extended products (Figure 2A). The overall catenation efficiency was merely ~59%. For fusion proteins containing AXB, the tethered proteins exert significant steric hindrance due to excluded volume effect to discourage the intermolecular over intramolecular reactions, leading to higher catenation efficiency. The effect is most evident if SpyTag is located between the protein of interest and p53dim domain. Therefore, (GB1)2-AXB has a higher catenation efficiency (~73%) than AXB-(GB1)2 (~66%). A similar trend has also been observed in the BXA module, as shown by the higher catenation efficiency in BXA-(GB1)2 (~83%) than (GB1)2-BXA (76%). This is somewhat counterintuitive, yet not surprising since changing the location of SpyTag to the middle of the chain has been previously shown to increase the yield of tadpole proteins. The 3D configuration of AXB and BXA is the origin of the difference in retention volumes of various fusion proteins. For AXB, both N- and C-termini point to the same side of the β-barrel and the tethered proteins are thus quite crowded, leading to a compact structure and a smaller hydrodynamic volume. For BXA, the N-terminus and C-terminus are pointing to the opposite ends of the β-barrel of the A–B complex, and the tethered proteins are thus more expanded in 3D space, leading to apparently higher hydrodynamic volume in SEC (Figure 2A). This difference is also consistently observed in each pair of SpyX-containing fusion proteins.

While the creation of different protein topologies itself presents an enormous challenge, a more fundamental question is “how does the topology influence the properties and functions of proteins?” To address this, we first examined the protein catenanes of AXB and BXA in comparison to the linear and cyclic controls. As revealed by DSC thermogram (Figure S14), both catenanes and cyclic monomers exhibit a dramatic increase in melting point (Tm) (~86 °C) as compared to the linear monomers (~62 °C). Thermal stability improvement from cyclic protein to catenane is relatively minor, probably because the protein is already very stable. The AXB and AX’B have almost the same Tm which is reasonable since the cyclized portion is relatively small and the influence is minor. By
In summary, we have developed highly efficient, genetically encoded, mechanically interlocking SpyX modules (AXB and BXA) for engineering protein topology in vivo. The expression of fusion proteins containing SpyX modules leads to the formation of a variety of mechanically interlocked protein topologies, including protein catenanes, obligate dimers, and star proteins. The clustering of p53dim with SpyTag/SpyCatcher ensures efficient protein chain intertwining and covalent fixation. The catenation efficiency was found to depend on the configuration of three components as well as the fused protein of interest. The configuration also influences the stability of the final product. Generally, BXA has better catenation efficiency than AXB and the products containing BXA are also more stable. There are many advantages in using SpyX modules for protein engineering. First of all, it leads to the formation of obligate dimers which would never disassociate even upon high dilution and does not require the tethered protein domain to possess specific protein–protein interactions. Second, it allows the formation of star proteins simply by direct expression, which is useful as multifunctional proteins. Finally, topology may enhance protein stability. For example, star proteins are shown to possess significantly enhanced stability against trypsin digestion. The capability of SpyX modules for in vivo protein topology engineering is an extraordinary example demonstrating the power of the “assembly-reaction” synergy. It expands the scope of protein topology beyond linear configuration and serves as a versatile platform to further engineer the functional properties of proteins.

**CONCLUSIONS**

In summary, we have developed highly efficient, genetically encoded, mechanically interlocking SpyX modules (AXB and BXA) for engineering protein topology in vivo. The expression of fusion proteins containing SpyX modules leads to the formation of a variety of mechanically interlocked protein topologies, including protein catenanes, obligate dimers, and star proteins. The clustering of p53dim with SpyTag/SpyCatcher ensures efficient protein chain intertwining and covalent fixation. The catenation efficiency was found to depend on the configuration of three components as well as the fused protein of interest. The configuration also influences the stability of the final product. Generally, BXA has better catenation efficiency than AXB and the products containing BXA are also more stable. There are many advantages in using SpyX modules for protein engineering. First of all, it leads to the formation of obligate dimers which would never disassociate even upon high dilution and does not require the tethered protein domain to possess specific protein–protein interactions. Second, it allows the formation of star proteins simply by direct expression, which is useful as multifunctional proteins. Finally, topology may enhance protein stability. For example, star proteins are shown to possess significantly enhanced stability against trypsin digestion. The capability of SpyX modules for in vivo protein topology engineering is an extraordinary example demonstrating the power of the “assembly-reaction” synergy. It expands the scope of protein topology beyond linear configuration and serves as a versatile platform to further engineer the functional properties of proteins.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscentsci.7b00104.

- Molecular cloning, protein expression and purification protocols, protein sequences, and other molecular characterizations (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

Tel: + 86 10 6276 6876. Fax: + 86 10 6275 1710. E-mail: wenbin@pku.edu.cn.

**ORCID**

- Bo Song: 0000-0002-4337-848X
- Xiaopeng Li: 0000-0001-9655-9551
- Wen-Bin Zhang: 0000-0002-8746-0792

**Author Contributions**

‡D.L. and W.-H.W. contributed equally to the work.

**Notes**

The authors declare no competing financial interest.
ACKNOWLEDGMENTS

We are grateful for the financial support from the 863 Program (2015AA020941), the National Natural Science Foundation of China (Grants 21474003, 91427304), and “1000 Plan (Youth)”.

REFERENCES

(1) Hadjichristidis, N. Complex Macromolecular Architectures: Synthesis, Characterization, and Self-assembly, John Wiley: Singapore, 2011.
(2) Matyjaszewski, K.; Grubbs, R.; Leibler, L. Macromolecular Engineering: Precise Synthesis, Materials Properties, Applications; Wiley-VCH: Weinheim, 2007.
(3) Hsieh, H. L.; Quirk, R. P. Anionic Polymerization: Principles and Practical Applications; Marcel Dekker: New York, 1996.
(4) Whittington, S. G.; Summers, D. W.; Lodge, T. Topology and Geometry in Polymer Science; Springer: New York, 1998.
(5) Hadjichristidis, N.; Pitsikalis, M.; Pispas, S.; Iatrou, H. Polymers with Complex Architecture by Living Anionic Polymerization. Chem. Rev. 2001, 101, 3747–3792.
(6) Zhang, W.-B.; He, J.; Yue, K.; Liu, C.; Ni, P.; Quirk, R. P.; Cheng, S. Z. D. Rapid and Efficient Anionic Synthesis of Well-Defined Eight-Arm Star Polymers Using OctavinyliPOSS and Poly(styryl)lithium. Macromolecules 2012, 45, 8571–8579.
(7) Ren, J. M.; McKenzie, T. G.; Fu, Q.; Wong, E. H.; Xu, J.; An, Z.; Shanmugam, S.; Davis, T. P.; Boyer, C.; Qiao, G. G. Star Polymers. Chem. Rev. 2016, 116, 6743–836.
(8) Wu, W.; Wang, W.; Li. Star polymers: Advances in biomedical applications. Prog. Polym. Sci. 2015, 45, 55–85.
(9) Matyjaszewski, K. Architecturally complex polymers with controlled heterogeneity. Science 2011, 333, 1045–1049.
(10) Hawker, C. J.; Wooley, K. L. The Convergence of Synthetic Organic and Polymer Chemistries. Science 2005, 309, 1200–1205.
(11) Whitford, D. Proteins: Structure and Function; J. Wiley & Sons: Hoboken, NJ, 2005.
(12) Zhang, W.-B.; Sun, F.; Tirrell, D. A.; Arnold, F. H. Controlling macromolecular topology with genetically encoded SpyTag-SpyCatcher chemistry. J. Am. Chem. Soc. 2013, 135, 13988–97.
(13) Wang, X.-W.; Zhang, W.-B. Cellular Synergism of Protein Catenates. Angew. Chem. Int. Ed. 2016, 55, 3442–3446.
(14) Dzheechocklit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W. Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. Nature 2004, 430, 101–105.
(15) Chen, Y.; Guan, Z. Bioinspired Modular Synthesis of Elastin-Mimic Polymers To Probe the Mechanism of Elastin Elasticity. J. Am. Chem. Soc. 2010, 132, 4577–4579.
(16) Sayre, T. C.; Lee, T. M.; King, N. P.; Yeates, T. O. Protein stabilization in a highly knotted protein polymer. Protein Eng, Des. Sel. 2011, 24, 627–630.
(17) King, N. P.; Jacobitz, A. W.; Sawaya, M. R.; Goldschmidt, L.; Yeates, T. O. Structure and folding of a designed knotted protein. Proc. Natl. Acad. Sci. U. S. A. 2010, 107, 20732–20737.
(18) Trabi, M.; Craik, D. J. Circular proteins - no end in sight. Trends Biochem. Sci. 2002, 27, 132–138.
(19) Montalban-Lopez, M.; Sanchez-Hidalgo, M.; Cebrian, R.; Maqueda, M. Discovering the bacterial circular proteins: bacteriocins, cyanobactins, and pilins. J. Biol. Chem. 2012, 287, 27007–27013.
(20) Conlan, B. F.; Gillon, A. D.; Craik, D. J.; Anderson, M. A. Circular proteins and mechanisms of cyclization. Biopolymers 2010, 94, 573–83.
(21) Above, T. L.; Camarero, J. A. Biological synthesis of circular peptides. J. Biol. Chem. 2012, 287, 27026–27032.
(22) Tam, J. P.; Wong, C. T. Chemical synthesis of circular proteins. J. Biol. Chem. 2012, 287, 27020–27025.
(23) Iwai, H.; Lingel, A.; Pluckthun, A. Cyclic green fluorescent protein produced in vivo using an artificially split PI-Pfalin intein from Pyrococcus furiosus. J. Biol. Chem. 2001, 276, 16548–54.
ruthenium(II)-terpyridine metallomacrocycles. J. Am. Chem. Soc. 2011, 133, 11967–76.

(46) Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplov, D. B.; Shea, J. E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer’s disease. Nat. Chem. 2009, 1, 326–31.

(47) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to beta-sheet in amyloid fibril formation. Nat. Chem. 2011, 3, 172–7.

(48) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. Science 2005, 310, 1658–61.

(49) Shi, L.; Holliday, A. E.; Shi, H.; Zhu, F.; Ewing, M. A.; Russell, D. H.; Clemmer, D. E. Characterizing intermediates along the transition from polyproline I to polyproline II using ion mobility spectrometry-mass spectrometry. J. Am. Chem. Soc. 2014, 136, 12702–11.

(50) Cao, Y.; Li, H. Polyprotein of GB1 is an ideal artificial elastomeric protein. Nat. Mater. 2007, 6, 109–14.

(51) Li, C.; Qi, Y.; Teng, X.; Yang, Z.; Wei, P.; Zhang, C.; Tan, L.; Zhou, L.; Liu, Y.; Lai, L. Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like protease. J. Biol. Chem. 2010, 285, 28134–40.

(52) Sun, F.; Zhang, W.-B.; Mahdavi, A.; Arnold, F. H.; Tirrell, D. A. Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. Proc. Natl. Acad. Sci. U. S. A. 2014, 111, 11269–74.

(53) Urry, D. W. Entropic Elastic Processes in Protein Mechanisms. I. Elastic Structure Due to an Inverse Temperature Transition and Elasticity Due to Internal Chain Dynamics. J. Protein Chem. 1988, 7, 1–34.

(54) Urry, D. W. Entropic Elastic Processes in Protein Mechanisms. II. Simple (Passive) and Coupled (Active) Development of Elastic Forces. J. Protein Chem. 1988, 7, 81–114.

(55) Urry, D. W.; Hugel, T.; Seitz, M.; Gaub, H. E.; Sheiba, L.; Dea, J.; Xu, J.; Parker, T. Elastin: a representative ideal protein elastomer. Philos. Trans. R. Soc. B 2002, 357, 169–84.

(56) Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N. Introducing protein intrinsic disorder. Chem. Rev. 2014, 114, 6561–88.

(57) Uversky, V. N. Intrinsically disordered proteins: Springer: New York, 2014.

(58) Hagan, R. M.; Bjornsson, R.; McMahon, S. A.; Schomburg, B.; Braithwaite, V.; Buhl, M.; Naismith, J. H.; Schwarz-Linek, U. NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. Angew. Chem., Int. Ed. 2010, 49, 8421–8425.