Engineering of brick and staple components for ordered assembly of synthetic repeat proteins

Jessalyn Miller, Agathe Urvoas, Benoît Gigant, Malika Ouldali, Ana Arteni, Agnès Mesneau, Marie Valerio-Lepiniec, Franck Artzner, Erik Dujardin, Philippe Minard

Engineering of brick and staple components for ordered assembly of synthetic repeat proteins.

Jessalyn Miller[1,2], Agathe Urvoas[2], Benoit Gigant[2], Malika Ouldali[2], Ana Arteni[2], Agnes Mesneau[2], Marie Valerio-Lepiniec[2], Franck Artzner[3], Erik Dujardin[4a], and Philippe Minard[2a]

1. Emory University Department of Chemistry, 1515 Dickey Drive, Atlanta, GA 30322 (present address)
2. Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, F-91198 Gif-sur-Yvette CEDEX, France
3. Institut de Physique de Rennes (IPR), CNRS, UMR 6251, Université de Rennes 1, F-35042 Rennes, France
4. Laboratoire Interdisciplinaire Carnot de Bourgogne, CNRS UMR 6303, Université de Bourgogne Franche-Comté, 21000 Dijon, France

a. To whom correspondence may be addressed. Email: philippe.minard@i2bc.paris-saclay.fr or erik.dujardin@cnrs.fr

Highlights

- "Brick" proteins assembled by "Staple" proteins form superhelices
- Brick and Staple proteins are synthetic repeat proteins
- The length of the assembly can be controlled by stop-Bricks
- The periodicity of staple protein along the super helix can be tuned
- Predefined assemblies can be formed either by inside or outside staple

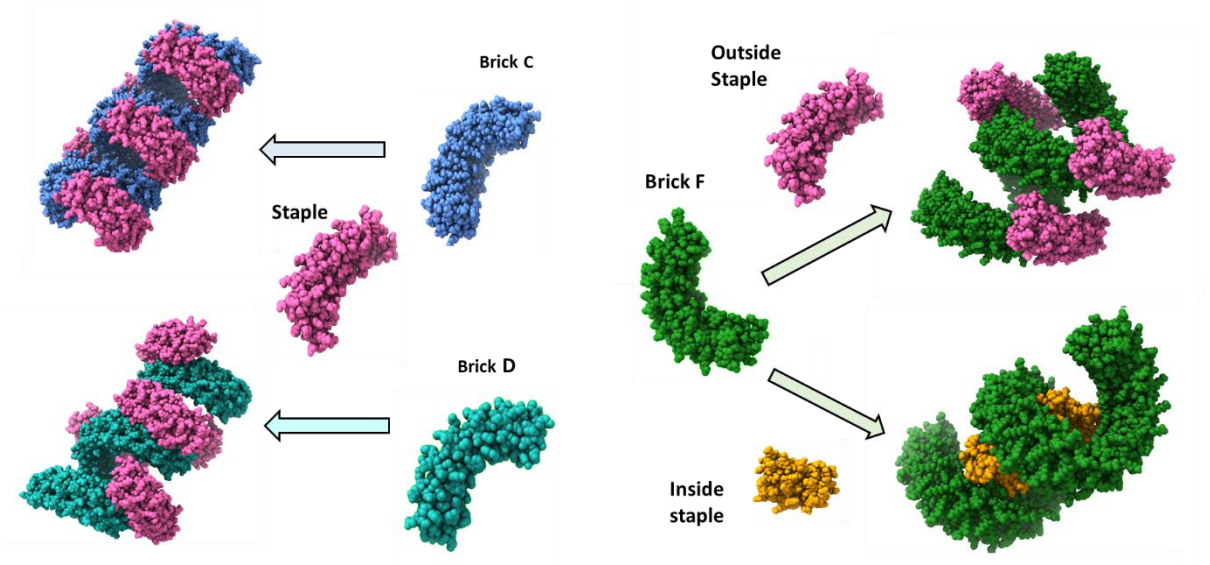
Abstract:

Synthetic αRep repeat proteins are engineered as Brick and Staple protein pairs that together self-assemble into helical filaments. In most cases, the filaments spontaneously form supercrystals. Here, we describe an expanded series of αRep Bricks designed to stabilize the interaction between consecutive Bricks, to control the length of the assembled multimers, or to alter the spatial distribution of the Staple on the filaments. The effects of these Brick modifications on the assembly, on the final filament structure and on the crystal symmetry are analyzed by biochemical methods, electron microscopy and small angle X-ray scattering. We further extend the concept of Brick/Staple protein origami by designing a new type of "Janus"-like Brick protein that is equally assembled by orthogonal staples binding its inner or outer

surfaces and thus ending inside or outside the filaments. The relative roles of longitudinal and lateral associations in the assembly process are discussed. This set of results demonstrates important proofs-of-principle for engineering these remarkably versatile proteins toward nanometer-to-micron scale constructions.

Keywords: *Tandem repeat protein, protein Origami, nanomaterial, self-assembly, protein design*

Graphical Abstract



Introduction

Repeat proteins classified as “solenoids” or class III proteins (Kajava, 2012) have naturally evolved from simple structural motifs such as α-α or α–β units (Alva and Lupas, 2018). Once concatenated in a single sequence, each motif interacts favorably with its neighboring motif giving rise to highly regular tertiary structure. Most side chains occupying solvent-accessible surface positions are not involved in folding or stability and tend to be highly variable. The juxtaposition of the variable side chains from

consecutive repeats in the folded protein generate a large variable and potential interaction surface. Natural repeat protein families such as LRR, ankyrin or HEAT (Andrade et al., 2001) repeats have naturally evolved to interact with many unrelated protein partners. This evolutionary process can be reconstituted by creating libraries of repeat proteins based on optimized repeats (Kajander et al., 2006). Highly diverse libraries such as Darpins (Binz et al., 2004), αReps (Guellouz et al., 2013; Urvoas et al., 2010), and repebodies (Lee et al., 2012) have been demonstrated to generate tight and specific protein binders (Boersma and Pluckthun, 2011). Due to their regularity and modularity, repeat proteins offer new possibilities for engineering protein assemblies and self-assembling nanostructures (Beloqui and Cortajarena, 2020; Bethel et al., 2022; Brunette et al., 2020; Brunette et al., 2015; Gidley and Parmeggiani, 2021; Parmeggiani et al., 2015).

We have recently described a new type of self-assembling protein structure based on αReps (Moreaud et al., 2023). αReps are a family of artificial proteins based on HEAT-like repeat that fold as curved solenoids (Urvoas et al., 2010). The concave surface is formed by hypervariable side chains and all structures of αRep binder complexes show that the target protein is bound by interaction with this surface. We have generated an αRep pair by selecting from the library a “Backbinder” protein, which specifically binds the convex, or “back”, surface of a designed αRep used as a “bait” protein and solved the crystal structure of the bait/Backbinder complex (PDB code 8AW4). The three repeats (I1, I2, I3) of the bait protein that contact the Backbinder are then split and appended at the termini of a chosen αRep protein hereafter named “Brick”. I1 and I2 are positioned at the C-terminus and I3 at the N-terminus of the Brick protein. Thus, when the Backbinder is mixed with the Brick protein, the Backbinder recruits two Brick proteins to reconstitute its cognate partner surface (I1-I2/I3), acting as staple that links two Brick proteins. Since each Brick can interact with another Brick at either extremity, the co-assembly can in principle reach infinite lengths. Experimental observations show that the isolated proteins are folded and soluble but rapidly form a macroscopically organized assembly upon mixing at room temperature.

Detailed characterization by SAXS, negative stain transmission TEM, cryoEM and tomography indicate helical assembly with pseudo C2 symmetry. Each Brick protein comprises 8 repeats, corresponding to one half-turn of the helix. There are therefore two Bricks and two Backbinders per turn. We also observed that these helical filaments spontaneously associate to form highly regular crystalline arrays. The Backbinders protruding from one side of a filament fit between two Backbinders on the neighboring filament. This interdigitation of Backbinders acts as driving force to form and extend the crystalline organization.

The principles used to design this new type of protein origami have proven to be effective. The experimental characterization of the resulting assemblies suggests a new set of questions:

1) The residues facing into the lumen of the assembly are not directly involved in the Brick/Backbinder interaction and have so far been considered neutral toward filament stability. Is it possible to stabilize the Brick/Brick interactions from the lumenal surface?

2) Is it possible to modulate the length of the assembly by using N- or C-terminal Stop proteins that can join the filament, but do not support further extension?

3) What is the effect of modifying the periodicity of the Backbinder on the resulting superhelix crystals?

4) Is it possible to extend this principle of assembly through the design of other orthogonal Brick/Staple pairs of proteins?

In this work, we find experimental answers to these questions and more generally, we explore the rules underlying these synthetic protein assemblies. We describe the rational design of new building elements such as new Brick and Stop proteins and report how these elements affect the co-assembly process as determined by SDS-PAGE, spectrophotometric techniques, X-ray scattering, and cryo EM.

Results

Stabilizing the assembly by optimizing Brick:Brick binding interface

The original Brick B design, described in (Moreaud et al., 2023), has been improved by introducing new H-Bonds and salt bridges at the Brick/Brick interface between the last repeat of one Brick protein and the first repeat of the next one in the assembly. This was done by substituting positions 22, 23, 26 and 30 of the first and last repeat of the cleaved brick protein. Side chains at these positions are oriented toward the concave face of the αRep, which forms the inside lumen of assembled superhelices. These positions are known to tolerate a diversity of side chains without compromising repeat stability (Urvoas et al., 2010). The relative positions of these residues from two neighboring repeats are predictable from the known αRep structures. We therefore explored side chains combinations in these specific positions that fulfill the following criteria: The side chains should be polar, with favorable helix propensity, and can establish H-bonds and/or salt bridges between two consecutive bricks. Side chains combinations maximizing interbricks interactions were manually explored using Fold-It Standalone as a structure manipulation tool based on Rosetta (Kleffner et al., 2017). A favorable predicted combination of side chains in these positions of first and last repeats was retained to design Brick C sequence and is shown in (Fig. 1). This new Brick design is hereafter named Brick C.

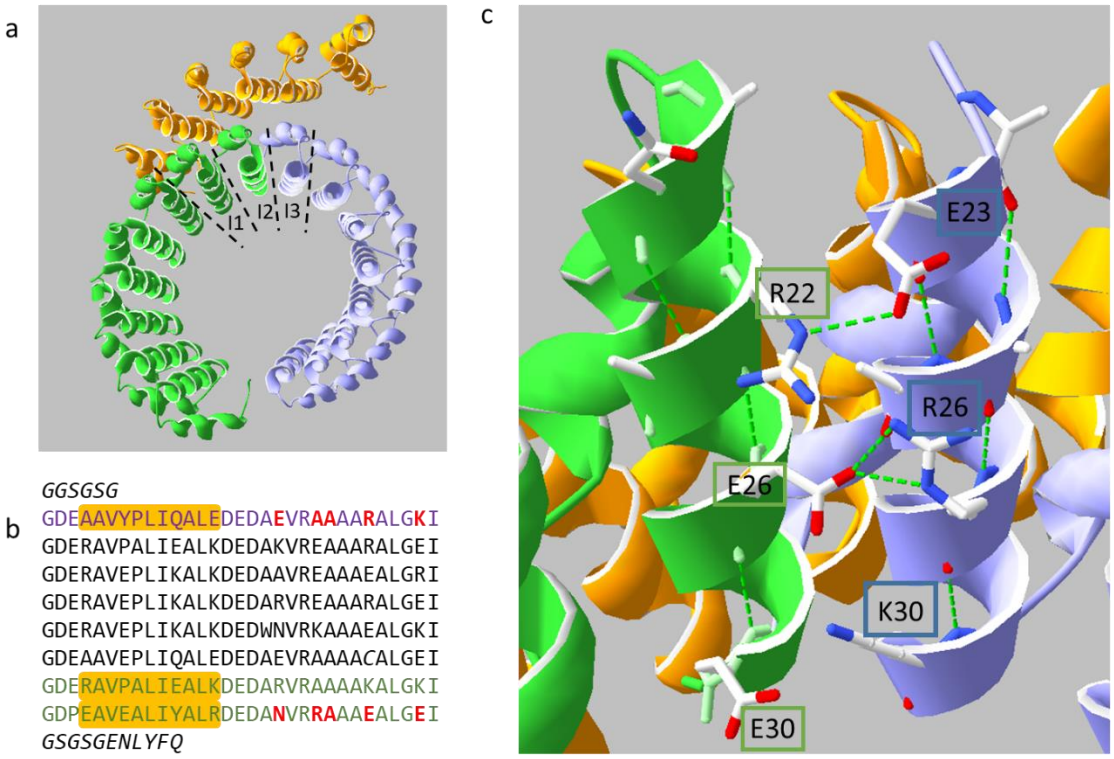


Fig. 1. Model of assembled Bricks C.

a) Model of a Backbinder protein (orange) bound on two Bricks C proteins (green and blue). The repeats of the Brick proteins I1 I2 and I3 interacting with the back binder proteins are indicated. b) Sequence of Brick C protein. The protein is initially expressed with two additional repeats named N and C caps, linked to 8 internal repeats through linkers and TEV cleavage sites. Cleavage by TEV protease leaves the sequence shown. The linker sequences on each extremity and TEV cleavage sites on the C term extremity are indicated in italics. Repeat I3 (blue), I1 and I2 (green) interact with the Backbinder protein. The Backbinder protein interacts with helix 1 of repeats I1, I2 and I3 in the zone boxed in orange. The positions whose side chains are oriented toward the lumen of the assembly correspond to positions 19, 22, 23 ,26 and 30 of each repeats. c) Closer view of the variable side chains on the last repeats of the first Brick C (green) and the second Brick C (blue). Relatively to the previously described Brick B the substitutions introduced on I2 (green) are K22R, A26E,Q30E on I3 (blue) D23E, E26R, R30K (numbers refer to sequence positions in each repeat). These amino-acids were substituted to introduce potential new salt bridges and H-bonds between consecutive Bricks.

To verify Brick C/Backbinder binding and determine a lower concentration limit for assembly, the two proteins were incubated together at 8, 4, 2, and 1 µM each for 16 hours at 37°C. After incubation, a cloudy white sediment was observed that pelleted easily upon centrifugation at 10000×*g* for 30 minutes. Analysis

of mix, supernatant, and pellet fractions by SDS-PAGE revealed that both proteins were present in the pellet, which indicates that pellets are composed of co-assembled protein (Fig. 2a). The co-assembly is detected for all examined concentrations, which suggests a critical concentration below 1 µM in contrast to a value greater than 4 µM for the Brick B/Backbinder system. The significant decrease of the critical assembly concentration strongly suggests the improvement Brick/Brick interaction intended by the sequence changes from Brick B to Brick C.

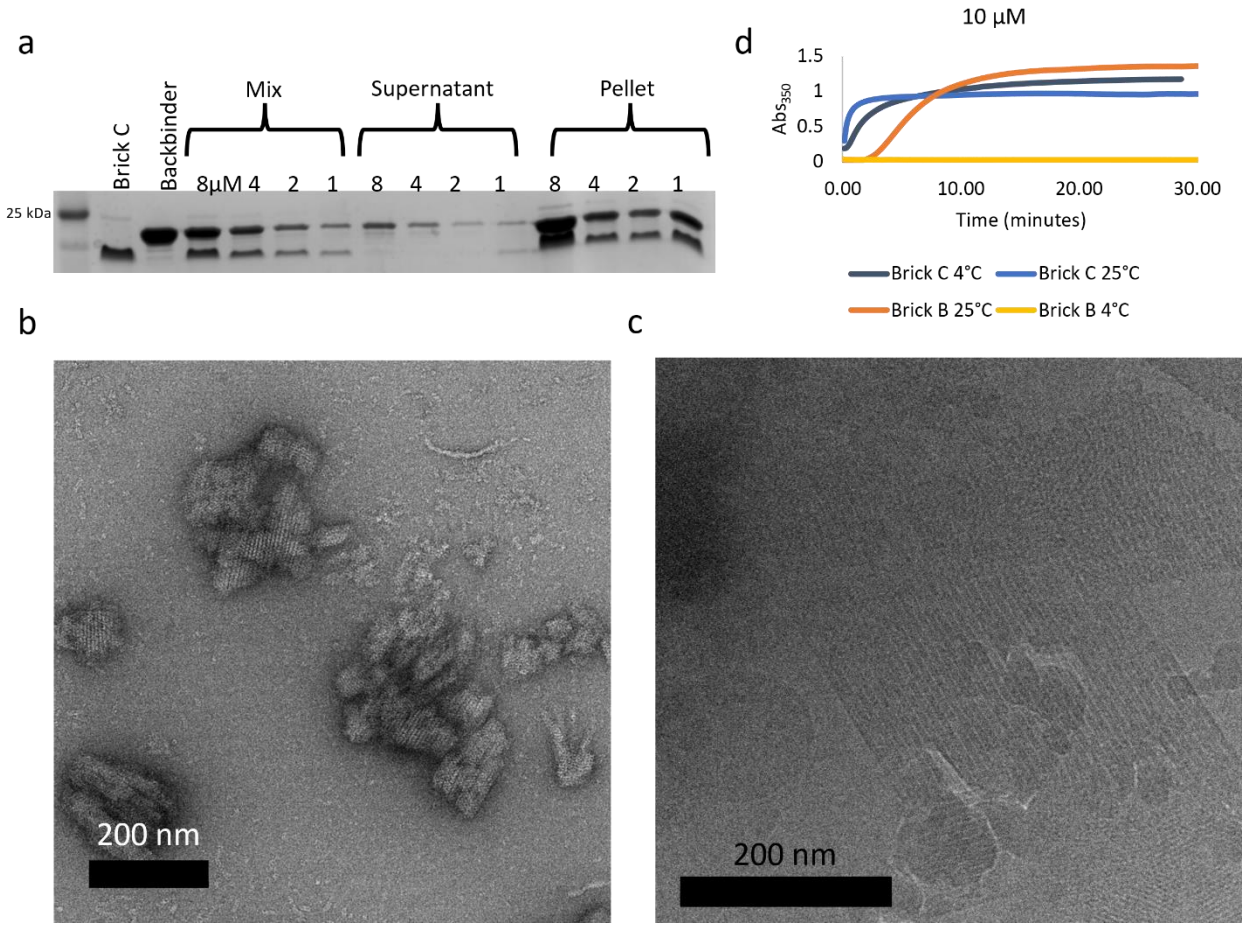


Fig. 2 Characterization of Brick C assembly

a) SDS PAGE analysis Brick C/Backbinder assembly of showing both Brick C and Backbinder in the pellet fraction after assembly at concentrations of 8, 4, 2, or 1 µM each protein. b) Negative-stain TEM micrograph showing crystalline super-assemblies obtained by assembly of Brick C and Backbinder (8 µM

each). c) Cryo EM image of crystalline super-assemblies of Brick C and Backbinder. d) Assembly curves in a temperature-controlled spectrophotometer monitored by absorbance at 350 nm as a function of time after mixing Brick and Backbinder proteins (10 µM). Kinetics with Brick B previously reported is shown for comparison. Brick C assembles faster than Brick B with less temperature dependance.

Transmission electron microscopy (TEM) observations of the 8 µM pellet fraction stained with uranyl acetate reveals that it is composed of nanotubular structures that are free or more generally closely packed into parallel crystalline bundles (Fig 2b). The structure of the pristine self-assembly is better observed without stain-induced electrostatic disruption of the bundles by cryo EM as shown in Figure 2c. The nanotubular structures appear aligned into highly ordered crystals similar to the fully indexed crystals obtained with Brick B (Moreaud et al., 2023).

The sequence change is also anticipated to have a marked impact on the assembly kinetics. Time-resolved light scattering monitored at 350 nm in temperature-controlled conditions is shown in Figure 2d. The evolution of samples containing Brick C contrasts with that of samples with Brick B in two ways: They reach the maximal absorbance in half the time (within 5 minutes rather than 10 minutes) and the rate of crystal formation almost unaffected by the temperature (Brick B does not assemble at all at 4°C while Brick C is only slightly slowed down). Note that no lag time due to manual mixing is detectable.

Taken together, these observations suggest that the interactions introduced on the terminal surface of the Brick proteins can be engineered to contribute favorably to the interactions between consecutive Bricks and hence to the Backbinder-driven assembly.

Changing the protein origami shape by tuning the Brick morphology

The crystalline organization of Bricks in the supramolecular assemblies is driven by the Brick/Backbinder interactions leading to superhelical filaments. In the case of Bricks B and C, we have demonstrated that the interdigitation of the Backbinder proteins from neighboring filaments led to higher order, crystalline organization. This interdigitation is favored by the positioning of one Backbinder every

half turn due to the 8-repeat length of Bricks B and C, which coincides with half of one superhelical pitch. Deletion of a complete repeat from a 8 repeat protein was expected to produce a shorter folded protein but not to alter the foldability of the resulting protein, since αReps as short as 3 repeats were shown to be fully folded proteins (Urvoas et al., 2010). Furthermore, if the deleted repeat is chosen from the central part of the protein that is not involved in the interface with the Backbinder, the assembly induced by the Backbinder will not be affected. However, as the relative positions of successive Backbinders along the superhelix are dictated by the number of repeats in the Brick protein, a shorter brick is expected to form filaments with a modified distribution of successive Backbinders along the superhelix. On this principle, we have designed a shorter, 7-repeat protein (Brick D) (Fig 3a), which bears the same terminal repeats as Brick C and so should bind the Backbinder and form filaments but staggers Backbinder around the filament 157.5˚ apart rather than 180° (Fig 3b). The resulting pronounced screw symmetry of the Backbinder locations is expected to impact the higher order crystallization of the superhelical filaments.

The Brick D protein was expressed, cleaved and purified. Identical volumes of 30 µM solutions of Brick D and Backbinder were mixed and incubated overnight at 37˚C. After centrifugation, pellet and supernatant fractions were diluted to 4 µM for imaging by TEM.

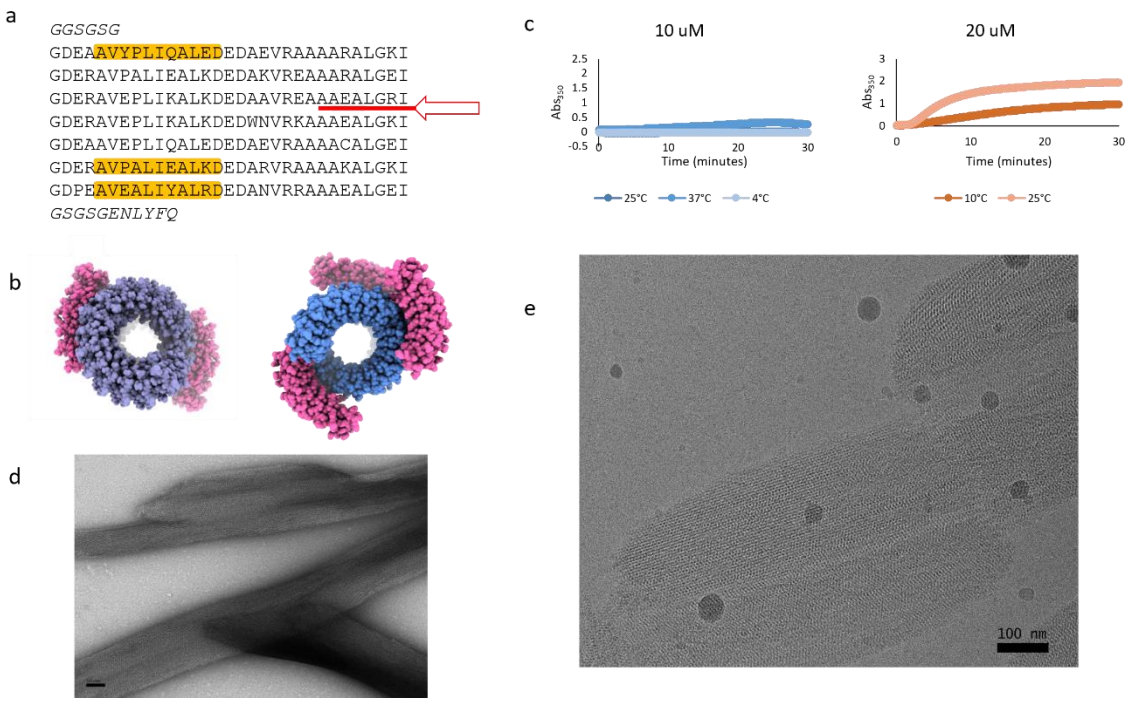


Fig. 3 Characterization of BrickD assembly

a) Sequence of Brick D. The fourth repeat of Brick C has been deleted (red arrow) leaving a Brick with 7 repeats. The interaction area with Backbinder (orange) is not modified. b) Comparative model of Brick C or D/Backbinder assembly. Axial view along the long axis of the filaments for Brick C (left) or Brick D (right) shows that with Brick D succesive Backbinders are not aligned on the axis of the filament, suggesting that with Brick D , the interdigitation of Backbinders is not possible without filament distortion. c) Assembly curves of Brick D and Backbinder at 10 (left) or 20 (right) µM concentration showing temperature and concentration dependence of assembly rate. d) Negative-stain TEM image of Brick D/Backbinder co-crystals. Crystals have a markedly different shape than Brick C/Backbinder crystals. e) CryoEM image of Brick D/Backbinder co-crystal with distinct moiré patterning.

Thermostability and Kinetics

Fig. 3c shows the time evolution of the light scattering of mixtures of Brick D and Backbinders. When the final concentration of both proteins is 10 µM, no scattering is observed at 4°C or 25°C and a very modest absorbance maximum is reached after 25 minutes at 37°C. However, when the final concentration was increased to 20 µM, the assembly proceeds slowly and monotonously at 10°C without reaching any plateau. At 25°C, we recover a sigmoidal behavior already observed for Brick B and a quasi-plateau is

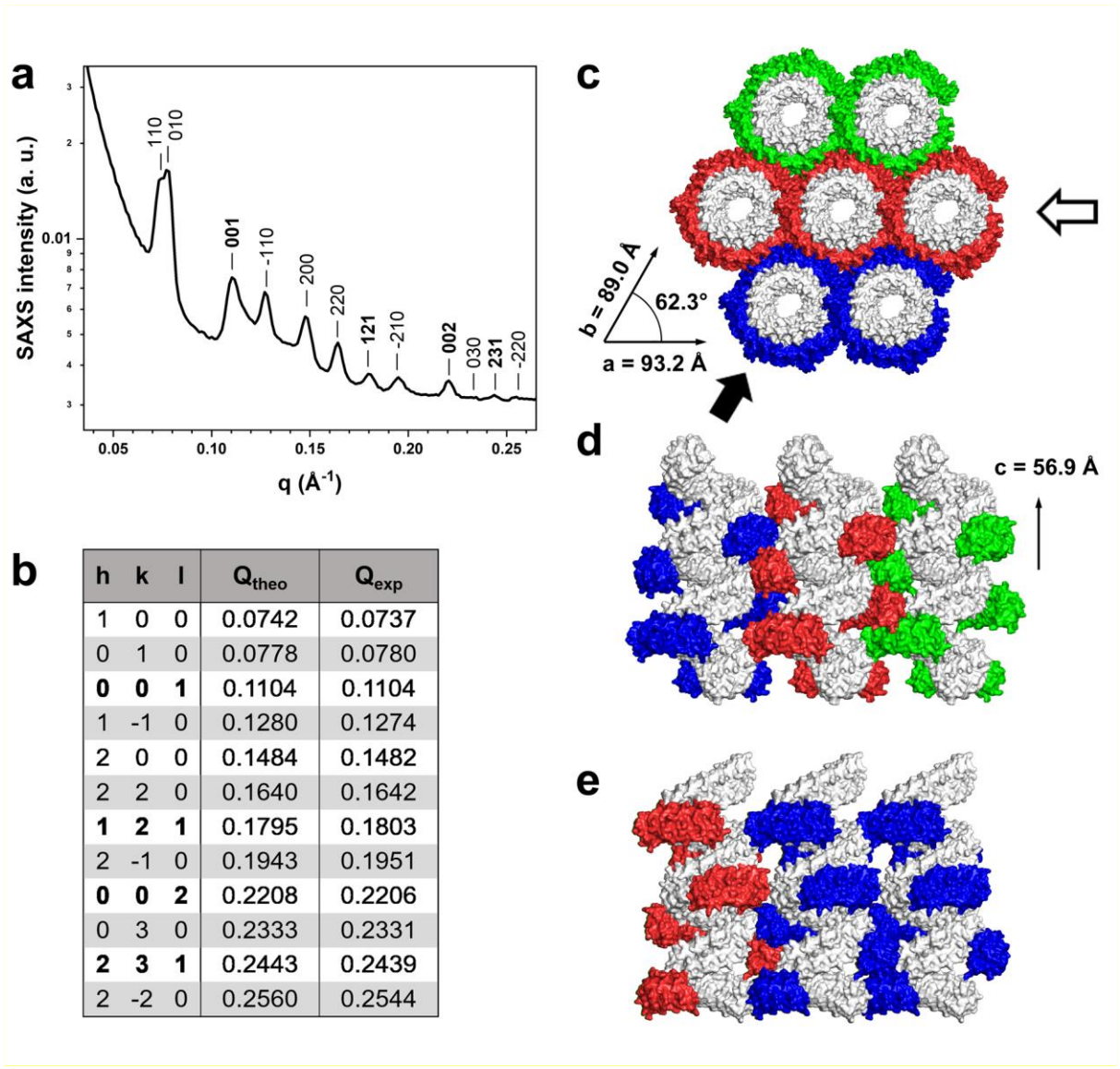
reached after 30 minutes. In spite of bearing the same terminal alterations that favor the Brick/Brick interactions, it therefore appears, Brick D does not form large objects able to scatter 350 nm-wavelength light as fast and at as low concentration as Brick C. At low temperature (below 10°C) and low concentrations (below 20 µM), nanometric assembly may form but higher order crystallization seems to be disfavored compared to the 8-repeat Bricks. The crystallization remains slow, which could indicate that the lateral interactions between neighboring filaments are less favorable with Brick D than with Brick C. This might be related to a frustration in the alignment of the staggered Backbinders.

TEM analysis of Brick D/Backbinder assemblies

Direct inspection by negative-stain and cryoEM revealed dispersed filaments (Fig S5) with length shorter than 75 nm in the supernatant and needle-like crystalline super-assemblies (Fig 3d,e) in the pellet fraction. No crystallites were found in the supernatant fraction nor long loose filaments in the pellet fraction.

Although Brick D was designed to frustrate the Backbinder interdigitation between filaments, it did not fully prevent the filaments from forming large crystalline organizations: Higher temperature and slower kinetics still lead to the pellet that shows higher order crystallization.

However, the structure and morphology of the crystals derived from Brick D are markedly different from the crystals obtained from Brick B and C. The crystals appear as flat rafts in cryoEM (Fig. 3e) that may curl up in the presence of the stain. This could be due to a partial charge neutralization by the ionic stain (uranyl acetate) letting the attractive polar and van der Waals inter-filament forces drive the compaction of the helical filaments.



b

h	k	l	Q_{theo}	Q_{exp}
1	0	0	0.0742	0.0737
0	1	0	0.0778	0.0780
0	**0**	**1**	0.1104	0.1104
1	-1	0	0.1280	0.1274
2	0	0	0.1484	0.1482
2	2	0	0.1640	0.1642
1	**2**	**1**	0.1795	0.1803
2	-1	0	0.1943	0.1951
0	**0**	**2**	0.2208	0.2206
0	3	0	0.2333	0.2331
2	**3**	**1**	0.2443	0.2439
2	-2	0	0.2560	0.2544

Fig4: *SAXS diffraction characterization of Brick D/Backbinder assembly*

(a) *SAXS diffraction pattern and (b) indexation table of a 10 µM supramolecular assembly of Brick D and Backbinder suspended in pure water. The peaks are indexed to the $P2_1$ monoclinic space group, with elementary cell parameters **a** = 93.2 Å, **b** = 89.0 Å and **c** = 56.9 Å. γ = 62.3°. In (b), Q_{theo} is the theoretical Bragg peak position and Q_{expe} is the experimental Bragg peak position determined by second derivative after parabolic smoothing. The theoretical peak positions are indicated in (a) by vertical lines. (c-e) Structural model of the superhelix 2D crystal derived from the SAXS data and viewed along the (c) {001}, (d) {100}, and (e) {010} zone axis. The rotational position of the superhelices with respect to their main axis along **c** is qualitatively estimated to minimize steric hindrance.*

SAXS study of Brick D/Staple co assembly

The self-assembly of Brick D and Backbinder was monitored by Small Angle X-ray Scattering (SAXS). The X-ray scattering is a well-defined set of diffraction peaks shown in Fig 4a that demonstrates a highly ordered crystalline structure rather than isolated monodisperse objects. The X-ray pattern remains unchanged at 70°C, indicating the high stability of the complex macroscopic assembly. Interestingly, the intense peak at $\boldsymbol{q}$ = 0.1104 $Å^{-1}$ corresponds to a distance of 56.9 Å that matches the helical pitch formed by the Bricks along the superhelical model as observed previously with Brick C and Backbinder, for which the pitch is 58.2 Å (Moreaud et al., 2023). The entire set of peaks is in perfect agreement with a monoclinic unit cell **a** = 93.2 Å, **b** = 89.0 Å and **c** = 56.9 Å. γ = 62.3°. The comparison with Brick C and Backbinder self-assembly confirms the nanotubes formation along the ***c*** axis as shown in the 3D model (Fig. 4c).

Interestingly, the packing of the superhelices is different by replacing Brick C with Brick D. First, the orthogonal rectangular packing observed with Brick C (Moreaud et al., 2023, Fig. 6A) turns into a close to hexagonal packing with Brick D. The nearest neighbor center-to-center distances are 89.0, 93.2 Å in the case of the superhelices with Brick D, which is significantly larger than with Brick C (64.9, 79.5 Å). Finally, superhelices made from Brick C exhibit a head-to-tail packing, which can be excluded with Brick D because there is only one nanotube per unit cell. All these results suggest that the role of the Backbinder interdigitation, which dictates the crystalline order of the superhelices made from Brick C is not observed with Brick D. In the present case, the packing should rather be considered as a compact 2D alignment of cylindrical objects. The side views of the 3D model along the ***a*** (Fig. 4d) and ***b*** (Fig. 4e) directions suggest that partial Backbinder interdigitation remains possible but never involves all Backbinders present between two Brick superhelices. The rotation of the superhelix around its main ***c*** axis is not locked by the interdigitation and thus probably reaches an optimum via multiple non-covalent and steric interactions.

This confirms that the reduction from 8 to 7 repeats in the Brick does not change the superhelical nanotube formation. Yet, in the absence of a strong interdigitation-driven inter-nanotube locking, the crystal force field imposes a quasi-hexagonal close packing and twofold symmetries that can distort the

nanotubular self-assembly. This fully accounts for the observation of that mixing Brick D / Backbinder produces tubular superhelices with less propensity to crystallize in the absence of Backbinder interdigitation. Yet, close-packed crystals of the highly anisotropic tubular superstructures are formed albeit with a slower kinetics and requiring a higher protein concentration than with Brick C.

Stochastic distribution of Bricks prevents lateral interactions and Backbinder interdigitation

Exploiting the rotational degree of freedom observed in pure Brick D/Backbinder assembly, we hypothesized that the crystallization of the superhelices could be prevented by mixing Bricks C and D prior to self-assembly upon exposure to the Backbinder. Indeed, the stochastic insertion of the two Bricks would results in a random distribution of the 157° and 180° relative positions of successive Backbinders, therefore reducing the opportunity for regular superhelix interdigitation. The superhelix would then remain an individualized soluble object. Assemblies obtained from mixtures of 8 µM total protein concentration with Brick C:Brick D molar ratios 50:50, 20:80 and 5:95 were examined. With a 50% Brick C content, crystals identical to those assembled from pure 4 µM Brick C formed. When the Brick C content is reduced to 20%, many short filaments and a few crystals were observed. Crystals completely disappeared upon further reduction of the Brick C content to 5%. Only numerous free filaments could then be observed (Fig.5). Diluting the total concentration of assembled proteins to 4 µM and 0.8 µM eventually yielded very small protein clusters rather than filaments (SI Figure S6). These observations suggest that irregularities in the geometry of distribution of the Backbinders due to a stochastic distribution of Brick C (8 repeats) and D (7 repeats) along the superhelical axis reduces the number and strength of inter-filament lateral interactions. Moreover, these observations suggest that if crystalline assemblies of filaments stabilized by lateral interactions are kinetically stable and do not dissociate upon dilution, the filaments are more dynamic and concentration-dependent.

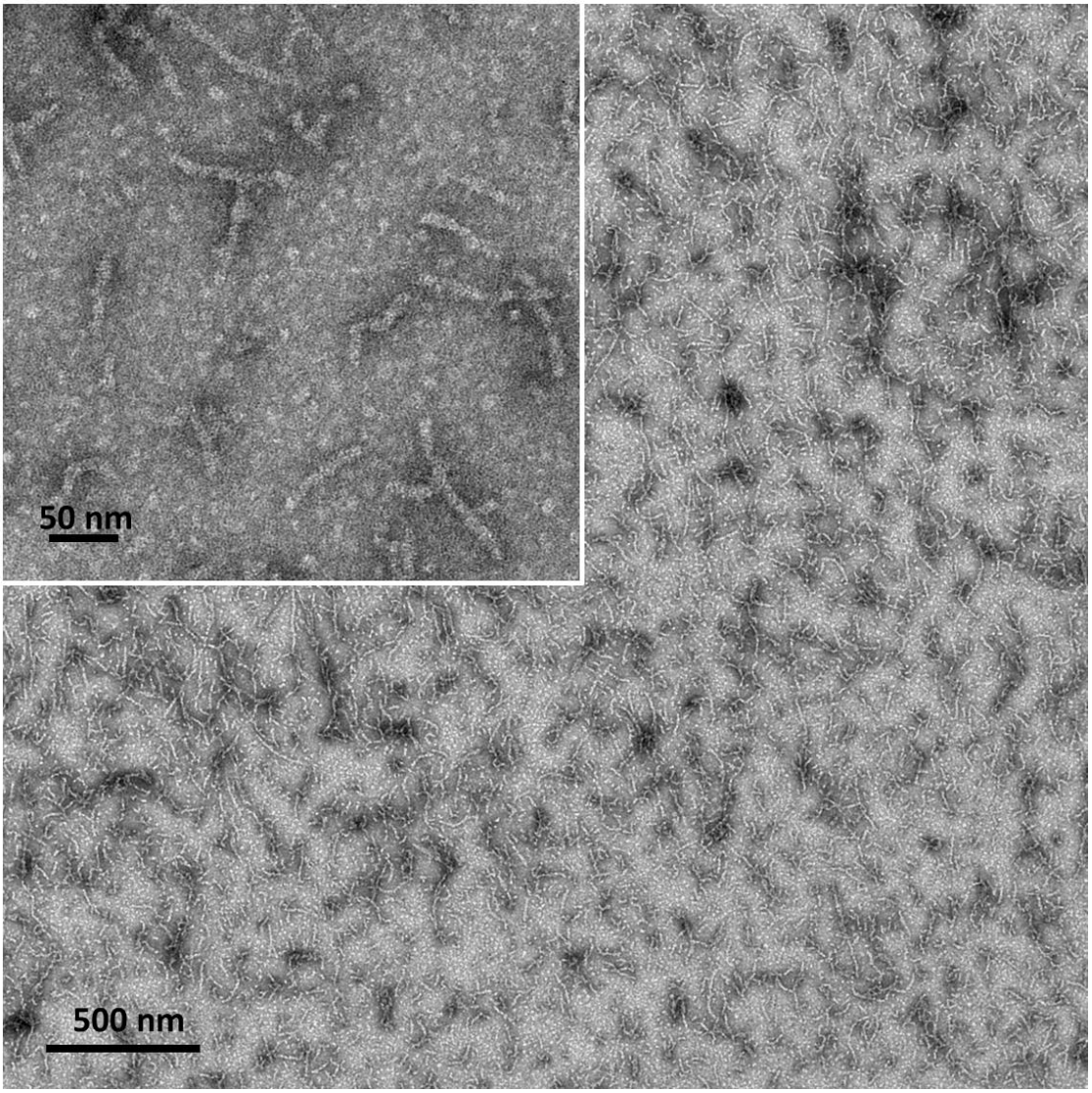


Fig. 5. Filaments produced by BrickC/BrickD mix assembled by Backbinder.

5% C, 95% D, Backbinder mix imaged at at 4 µM. No crystals were found in this sample, only soluble filaments.

Controlling the origami extension with N- and C-Stop Bricks

The Brick B, C and D proteins have the appropriate repeat sequences at their N- and C-termini to interact with the Backbinder. The average length of the super-helical assemblies could be programmed by designing Bricks with only one Backbinder binding extremity. Such N-Stop or C-Stop proteins were designed by replacing the corresponding terminal repeat with standard αRep terminal repeats, which are not recognized by the Backbinder (Fig S2). Once incorporated in a growing filament, a N- or C-Stop Brick cannot interact with a second Backbinder protein and the filament can no longer be extended. We expected that the length distribution of the filaments could be downshifted as we introduce a higher fraction of N- and/or C- stop Bricks alongside fully functional ditopic Bricks.

The effect of the Stop Bricks was examined by preparing equimolar mixtures of C-Stop and N-Stop and mixing them with either Brick C or Brick D at a chosen molar ratio before inducing the assembly by adding the Backbinder. The negative stained TEM images in Figures 6a and 6c demonstrate that the presence of N-Stop and C-Stop Bricks directly affects the extent of the superhelical assembly for both Brick C and D at Stop:Brick molar ratio of 1:4 and 1:16 respectively. However, when the content of Stop protein is reduced to 1:16, for Brick C, and 1:32, for Brick D, regular superhelix crystals do form similarly to mixture made from pure ditopic Bricks (Figs 6b and 6d).

With Brick D, a 1:16 molar ratio (Stop-Brick:BrickD) is sufficient to completely prevent the formation of insoluble materials and only produce isolated filaments. The only difference between Brick C and Brick D based superhelices is the periodicity of the Backbinder. These observations further confirm that the distribution of Backbinders along the superhelix is less favorable to lateral interaction with Brick D than with Brick C. With Brick D, the Backbinders are not aligned along the long axes of the filaments as they are with Brick C, but are staggered along the filaments. Consequently, interdigitation of Backbinders from Brick D filaments is only possible by a structural adaptation of the filaments, while the interdigitation of

aligned Backbinder from Brick C filaments does not generate frustration with lateral assembly. In other words, even very short filaments of Brick C/Backbinder tend to interact laterally, while short filaments of Brick D/Backbinder are able to form crystalline organization only above a critical length.

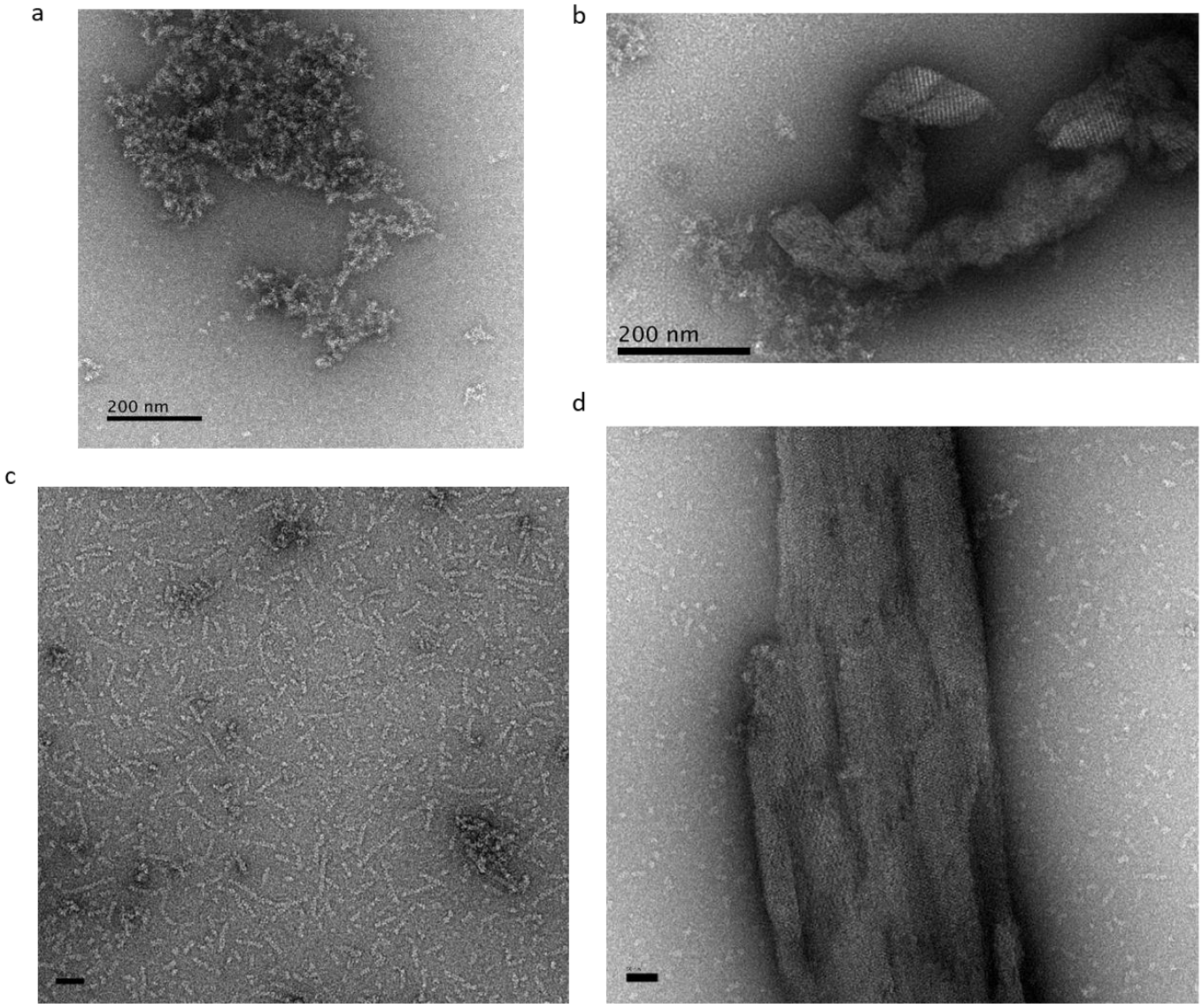


Fig. 6. Negative stained TEM images of protein assemblies resulting from the binding of the Backbinder with mixtures of C-Stop and N-Stop Bricks with regular Brick C or Brick D at different molar ratio.

(a) At a 1:4 molar ratio of Stop Bricks to Brick C, macroscopic crystal formation is abolished and only small aggregates proteins are observed. (b) At 1:16 or lower ratios of Stop Bricks to Brick C, large and ordered crystals are observed. (c) At 1:16 Stop-Brick to Brick D molar ratio, no crystal is observed but the sample is exclusively composed of linear filaments of uniform width. (d) At 1:32 molar ratio of Stop Bricks

to Brick D, short filaments coexist with macroscopic crystals with a different morphology than those obtained from Brick C.

Designing Backbinders for lumenal stapling of nanotubular assemblies

In the initial design, the Backbinder protein staples two consecutive Bricks by binding to their convex surfaces. In principle, a similar association effect could be obtained by a binding partner interacting with the concave surfaces of two consecutive Bricks, therefore driving the assembly, from inside the lumen of the filaments rather than from outside. To test this hypothesis, we have identified a tightly bound protein partner small enough to fit within the lumen of the superhelix. A small αRep comprising only three repeats, including the N- and C-Stop repeats, was previously identified to bind tightly (K_d in the low nM range) with another 6-repeat αRep named A3 (Guellouz et al., 2013). This small protein, originally named bA3-2, will hereafter be referred to as αRep2. The structure of the αRep2/A3 complex was solved (Guellouz et al., 2013) and was used to design a new Brick protein, named Brick F. The αRep2-binding surface of A3 was split in two parts and appended to the two extremities of the concave surface of Brick F (Fig. 7). Associating two Brick F proteins together reconstitutes the A3/αRep2 binding site, and it is therefore expected that αRep2 could act as an "inside Staple" protein for Brick F. The two convex and concave surfaces of αReps are independent so that it is possible to design a Brick F as a "Janus particle," able to interact with αRep2 on the inside and with the Backbinder on the outside. The convex face of Brick F was made identical to that of Brick C.

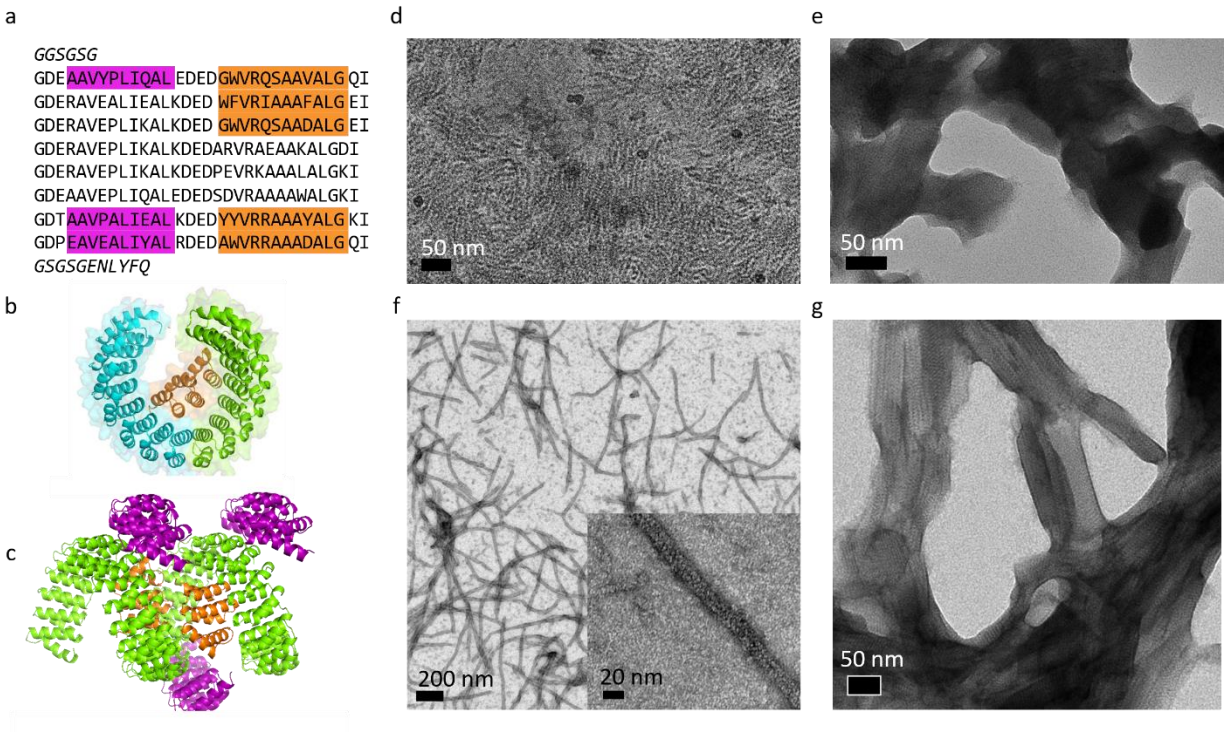


Fig. 7 Interactions of BrickF with αRep2 and Backbinder.

(a) Protein sequence of Brick F. In orange the domain of the helices that form the αRep2 binding surface. In purple, the domain of the helices responsible for Backbinder binding. (b, c) Axial view of the 3D model of two Brick F proteins (cyan, green) assembled by the interaction with an αRep2 protein (orange) bound on the concave surface. (c) Side view of a superhelix formed by Brick F (green) with Backbinder proteins (purple) interacting on the outside surface and αRep2 proteins (orange) acting as inside staple proteins. (d) *Cryo EM image of Brick F/Backbinder/αRep2 condensate. (e) Negative stain TEM image of BrickF/Backbinder crystals. (f) Long, twisted filaments of Brick F and α2 only. Inset: Close-up from same sample as A showing two or three filaments in each bundle. (g) Tape-like crystals of Brick F and αRep2 formed after several days.*

Brick F was exposed to either αRep2 or Backbinder alone, or to pre-mixed equimolar αRep2/Backbinder mixture. All proteins were mixed at a final concentration of 8 μM. All combinations involving Brick F became turbid as soon as added to binding partner (Backbinder and/or αRep2). As a control experiment, Brick C solution was mixed with αRep2 and remained as a clear solution. SDS-PAGE confirms that the pellets from turbid samples contain all constituent proteins (SI Figure S4). In negative-stain electron microscopy, the Brick F/αRep2/Backbinder sample appears to form 100-200 nm free filaments as well as poorly ordered bundles but no ordered crystals are observed. In cryoEM, no crystals

were found, but a condensate of flexible filaments with uniform morphology was observed (Fig. 7d). Brick F/Backbinder appears as crystals, identical to Brick C/Backbinder (Fig. 7e). Brick F/αRep2 produced flexible free and aggregated flexible filaments (Fig. 7f) over 1-µm long, with an average width of 16.6 nm. Each single Brick F/αRep2 filament is expected to be about 6.3 nm thick according to the model, so this corresponds to 2 or 3 intertwined filaments. After incubating at room temperature for several days, the F/αRep2 filaments form high-aspect-ratio tape-like assemblies that appear crystalline (Fig. 7g).

The assembly of Brick F with Backbinder is identical to Brick C with Backbinder, as expected. These results indicate that αRep2 is incorporated into the Brick F/Backbinder assembly, validating the concept that a single αRep can act as a Janus particle, binding one partner on each side. Importantly, our observation shows that αRep2 and Brick F are sufficient to form a stable superhelical assembly.

When both interior (αRep2) and exterior (Backbinder) staples are present in the Brick F filaments, crystals do not form, despite the presence of the Backbinders (Fig. 7g). This unexpected observation could possibly result from several causes. First, even though in all known αRep/target complexes the αRep fold adopts a curvature identical to that of an isolated αRep, it is possible that inter-filaments association may induce different structural distortions between bBE3- and αRep2-driven superhelices, thus preventing large scale formation of superhelix crystals. Alternatively, the presence of both inside and outside staples could result in a kinetic partitioning into two modes of assembly. Indeed, αRep2 may not fit within already formed Backbinder-driven filaments. Similarly, αRep2-stabilized filaments may result into the observed twisted superstructures, where bBE3 binding sites might not be accessible or geometrically altered, thus preventing bBE3 docking. As a result, Backbinder may not be evenly and systematically distributed along the filaments, therefore reducing the interdigitation-related interaction that drives the crystallization.

At present, our experimental data can not provide quantitative information with sufficient accuracy to detect alteration in the stoichiometry and to discriminate these hypotheses. Further experiments will be required to gain more precise insight in the interplay between inside and outside stapling.

Discussion

Two independent binding surfaces on the Brick proteins

We have recently described a strategy to create inducible protein assembly with pre-defined 3-dimensional positional and orientation geometry based on the interactions of two artificial repeat proteins: Brick proteins are organized in space under the control of a Staple protein. This process based on modular and structurally well-defined αRep is potentially versatile, as higher order geometries can be reached by the design of new Brick proteins based on different αRep size or specificity thus making this concept a generic approach to the assembly of synthetic protein origami. Our next goal is therefore to explore experimentally the behavior of elementary variations of the Brick protein in order to uncover the rules that govern these synthetic protein assemblies.

The assembly between Bricks is induced by the "Staple" effect of the Backbinder proteins. A fraction of the convex or "back" surface of the Brick is dedicated to the Brick/Backbinder interactions. However, the residues located in the concave, or "inner", surface can be freely modified so long as they do not disrupt the helical structures of the αRep. We have used this degree of freedom to design a new Brick protein (Brick C) that promotes interaction between two neighboring Bricks, once assembled by the Backbinder Staple. Indeed, the assemblies with Brick C form more rapidly, at lower protein concentration, and with less dependence on the incubation temperature than the first generation origami using Brick B proteins.

We also provide a clear example of the modularity and versatility of αRep design as one face of the protein can be altered independently of the opposite surface. The design of Brick F includes a recognition surface for a new protein partner (αRep2), unrelated to the Backbinder that is embedded in the inner surface of the superhelix. The binding surface for αRep2 is split over two consecutive Bricks. This protein

acts as an another, interior Staple and induces assembly via interactions in the lumen of the assembling superhelix. This directly suggests that Janus-type Bricks with alternative orthogonal “outside” and “inside” Staples could be potentially designed (or selected) and used to build more elaborate 3 dimensionally ordered target structures.

More generally, as repeat protein libraries are very efficient sources of tight and specific binding reagents, this principle of organized assembly could presumably be extended for other type of protein repeats based on solenoid architectures, although this should be confirmed by further experimental work.

Interdependence of longitudinal and lateral interactions

The assembled structure results directly from the readily programmable Brick/Staple interactions. A more subtle set of assembly rules emerge from the importance of inter-filament or “lateral” interactions that are instrumental to the final assembly but are much less predictable than the longitudinal Brick/Staple interactions. The highly regular and crystalline organization mediated by the Backbinder interdigitation previously observed with Brick B (Moreaud et al., 2023) is also observed, as expected, with Brick C, which leads to the same spatial distribution of the Backbinders. When the regularity of the Backbinder positioning within the assembly is disrupted by the slightly shorter Brick D, no inter-filament interdigitation is possible yet a different but highly regular organization appears leading to micron-scale supercrystals. Moreover; even when the protruding Backbinders are completely absent, such as in the Brick F/αRep2 assembly, no Backbinder interdigitation is possible, but higher order assembly nevertheless takes place whereby 2 or 3 intertwined filaments form a massive “rope”-like superstructure. Although it follows a lower symmetry, the superhelical filaments are highly regular objects that favor periodically distributed interactions that can be indefinitely repeated. Thus, even moderately attractive forces can accumulate in a regular way to create a significant inter-filament order. Further high resolution imaging

will provide better insight in these hard-to-predict interactions and should enable higher order construction models to emerge.

The difficulty of predicting the exact outcome of multimolecular repeated association of helical peptides has been recently discussed by Conticello and coworkers (Hughes et al., 2019). While difficult to predict, these interfilament interactions nevertheless play a decisive role in determining the macroscopic architecture that finally emerges from the associated proteins.

Regular multifilament structures could be desired for future applications such as structurally organized protein or nanomaterial self-assembly (Uribe et al., 2021). Alternatively, isolated filaments can be obtained by two efficient ways to inhibit interfilament assembly. The first one is to disrupt the regularity of the Backbinder decoration along filaments by using two different Bricks (such as C and D). The resulting stochastic Backbinder decoration on the periphery of the Brick superhelix makes it less likely to allow multiple weak but identical interactions regularly repeated along the filaments. The second one is to control the length of the superhelical as shown by using Stop Bricks with Brick C and Brick D. Shorter filaments are less prone to associate than longer ones since fewer elementary interactions are engaged. This is clearly visible by comparing the effects of the stop-Bricks on Brick C versus Brick D filaments. Brick D filaments have a lower tendency to interact laterally than do Brick C filaments, and thus need to be longer in order to form stable crystalline assemblies. Consequently, Brick D gives isolated filaments at lower concentrations of Stop Bricks than Brick C filaments, which can only be dissociated at a higher Stop/Brick ratio.

Note that in the absence of a Staple, the isolated Bricks are soluble proteins, that do not form filaments, indicating that the lateral interactions between Bricks is not sufficient to induce association and formation of superhelices. But if the Staple-induced longitudinal associations are critical, the lateral associations do contribute to the stability of the final assembled structures. Isolated filaments can

disassociate when diluted out of equilibrium, while fully formed crystals of Bricks B, C, or D with Backbinder may be rinsed without apparent dissolution.

Our present work also suggests possible concepts for future functionalization of these organized assemblies for metabolic engineering, structural biology or nanoconstruction. Functionalization of crystalline superstructures with a cargo (e.g., fusion protein, molecule, nanoparticle) could take profit of the improved assembly propensity based on Brick C, as long as the cargo does not introduce steric hindrance within and between filaments. Finally, the dynamic behavior of associated superstructures with a range of alternative partners also opens routes towards dynamically controlled superstructures, and is currently a recurring theme in the dynamics of natural protein superstructures such as microtubules (Knossow et al., 2020).

Materials and Methods

Expression plasmids

Proteins were expressed in *E. Coli* using synthetic genes cloned pUR22 and pUR23 expression plasmids (Rohweder et al., 2018). These expression plasmids adapted to Golden Gate cloning were designed and kindly provided by R. Sterner. These plasmids allow simple and efficient cloning using Golden Gate assembly and efficiently expression based on IPTG inducible T5 promoter.

Plasmid preparation. Plasmids were isolated with a NucleoSpin Plasmid mini plasmid prep kit (Macherey Nagel, Hoerdt, France) according to kit instructions. Insert sequences were verified by Eurofins Genomics (Germany). After sequence verification, plasmids were transformed into BL21 (Gold) *E. coli*.

Protein expression and preparation. One colony was selected for an overnight culture in 50 mL 2YT containing 100 $\mu g.mL^{-1}$ ampicillin at 37°C. A new 1L culture, was inoculated from the preculture to a final OD_{600} of 0.15. This culture was shaken at 37°C until OD_{600}=0,6 to 0.8. When this OD was reached, IPTG was

added to a final concentration of 1 mM and temperature was reduced to 30°C for protein production. After 4 hours of induction, the cells were pelleted in a centrifuge at 10000×*g* for 30 minutes at 4°C. Protein production was controlled by SDS-PAGE. Pellet was resuspended in 25 mL Tris Buffered Saline (Tris HCl 50 mM, NaCl 150 mM, pH 7.5) with one tablet complete™ protease inhibitor (Roche Boulogne-Billancourt, France) and stored at -20°C.

2 µL DNase (Thermo Scientific) was added to the pellet as it thawed in a tepid water bath. Once thawed, the pellet was sonicated on ice 30 seconds on/30 seconds off for 5 minutes. The lysed pellet was clarified by centrifuging at 10000×*g* 60minutes at 4°C. Supernatant was carefully removed and passed through a 0.2 µM syringe filter to remove cells.

Protein was purified via nickel-NTA immobilized metal affinity column chromatography (Protino® Ni-NTA Agarose, Machery-Nagel, Hoerdt, France) or by either by gravity (Protino® Ni-NTA Agarose, Machery-Nagel, Hoerdt, France) or by FPLC Ni-NTA IMAC (His-Trap FF Crude, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Ni-NTA purification was followed by size exclusion chromatography. Collected fractions were analyzed by SDS-PAGE. To cleave the 6x His tag and capping motifs, the protein was incubated with tobacco etch virus protein (TEV) at 1:50 TEV:protein, OD:OD in TBS with 1 mM dithiothreitol, 2 mM EDTA. This cleavage cocktail was incubated with shaking at 4°C overnight or 30°C for 2.5 hours. Cleaved protein was isolated by passing through Ni-NTA resin so that cleaved His tag and TEV protein are retained on the column. Cleavage was verified by SDS-PAGE. Control analysis of bricks C, D and F by size exclusion chromatography indicate that, in the absence of the Backbinder protein, the cleaved Brick proteins do not associate before or after TEV cleavage.

Filament assembly. Unless specified otherwise, proteins were incubated at a 1:1 molar ratio in TBS buffer at 37°C for assembly, and stored at -20°C.

Kinetics of assembly was monitored by light scattering. In the absence of molecular absorption for wavelength larger than 300 nm, the time evolution of the light scattering intensity was monitored by

recording the optical density at 350nm, OD_{350}. OD_{350} was monitored using a Varian Cary 50 Bio Spectrophotometer (Agilent, Santa Clara, CA, USA) equipped with temperature-controlled cell holder. For an experiment with 4 µM each protein, 50 µL of 8 µM Backbinder were placed in the cuvette and the absorbance reading at 350 nm was set to zero to compensate for background absorbance due to the cuvette. 50 µL of 8 µM Brick protein were then added to the cuvette, pipetted up and down to mix the data collection begun. Data was collected every 0.2 seconds for at least 15 minutes.

Transmission electron microscopy (TEM): Samples were analyzed by conventional electron microscopy using the negative staining method. 3 µL of sample suspension were deposited on an air glow-discharged 400 mesh copper carbon-coated grid for 1 minutes. The excess of liquid was blotted, and the grid rinsed with 2 % w/v aqueous uranyl acetate. The grids were visualized at 100 kV with a Tecnai 12 Spirit transmission electron microscope (Thermo Fisher, New York NY, USA) equipped with a K2 Base 4000 x 4000 camera (Gatan, Pleasanton CA, USA). Nominal magnification was at 15000x corresponding to a pixel size of 0.25 nm.

Cryo electron microscopy: The cryoEM grids were prepared using a Vitrobot Mark IV (ThermoFisher) at 20 °C and 100% humidity. 3 µL of sample were applied onto freshly glow-discharged Quantifoil grids (R2/2), 200 mesh grids. The grids were blotted for 10s s with blot force 2, then plunge-frozen in liquid-nitrogen-cooled ethane. The grids were mounted in a 626 Gatan holder using its cryo-transfer device.

CryoEM images were observed in a Tecnai G2 FEG electron microscope (ThermoFischer) operating at 200 kV and equipped with a DDC K2 Summit direct-detection camera (Gatan Inc.). Images were recorded at 15 000 × magnification, with a pixel size of 0.26nm at the specimen level and 20 e−/Å2.

SAXS measurements. SAXS measurements were carried out on three types of samples by varying the mixing process: homogenization or contact between the pure protein solutions. The final concentration of the self-assembled proteins mixture is 10 µM. All preparations were inserted into a capillary of 1.4 mm

in diameter. The capillaries are probed in a homemade SAXS-WAXS Guinier beam line with a bidimensional Pilatus detector placed at 273 mm of the sample (Bizien et al., 2015). The beam is generated by a 30 µm X-ray Copper source (Xenocs, Grenoble, France). The beam is focused on the detector (150 mm) and monochromatized (λ= 1.541 Å) by a toroidal multilayer mirror (Xenocs). 2D concentric scattering rings are radially integrated as a function of $\boldsymbol{q} = 4\pi.\sin \theta/\lambda$, where 2θ is the scattering angle. All samples exhibit a similar pattern representative of standard assembly.

References

Alva, V., Lupas, A.N., 2018. From ancestral peptides to designed proteins. Curr Opin Struct Biol 48, 103-109.

Andrade, M.A., Perez-Iratxeta, C., Ponting, C.P., 2001. Protein repeats: structures, functions, and evolution. J Struct Biol 134, 117-131.

Beloqui, A., Cortajarena, A.L., 2020. Protein-based functional hybrid bionanomaterials by bottom-up approaches. Curr Opin Struct Biol 63, 74-81.

Bethel, N.P., Borst, A.J., Parmeggiani, F., Bick, M.J., Brunette, T., Nguyen, H., Kang, A., Bera, A.K., Carter, L., Miranda, M.C., Kibler, R., Lamb, M., Li, X., Sankaran, B., Baker, D., 2022. Precisely patterned nanofibers made from extendable protein multiplexes. bioRxiv, 2022.2010.2014.511843.

Binz, H.K., Amstutz, P., Kohl, A., Stumpp, M.T., Briand, C., Forrer, P., Grutter, M.G., Pluckthun, A., 2004. High-affinity binders selected from designed ankyrin repeat protein libraries. Nat Biotechnol 22, 575-582.

Bizien, T., Ameline, J.C., Yager, K.G., Marchi, V., Artzner, F., 2015. Self-Organization of Quantum Rods Induced by Lipid Membrane Corrugations. Langmuir 31, 12148-12154.

Boersma, Y.L., Pluckthun, A., 2011. DARPins and other repeat protein scaffolds: advances in engineering and applications. Curr Opin Biotechnol 22, 849-857.

Brunette, T.J., Bick, M.J., Hansen, J.M., Chow, C.M., Kollman, J.M., Baker, D., 2020. Modular repeat protein sculpting using rigid helical junctions. Proc Natl Acad Sci U S A 117, 8870-8875.

Brunette, T.J., Parmeggiani, F., Huang, P.S., Bhabha, G., Ekiert, D.C., Tsutakawa, S.E., Hura, G.L., Tainer, J.A., Baker, D., 2015. Exploring the repeat protein universe through computational protein design. Nature 528, 580-584.

Gidley, F., Parmeggiani, F., 2021. Repeat proteins: designing new shapes and functions for solenoid folds. Curr Opin Struct Biol 68, 208-214.

Guellouz, A., Valerio-Lepiniec, M., Urvoas, A., Chevrel, A., Graille, M., Fourati-Kammoun, Z., Desmadril, M., van Tilbeurgh, H., Minard, P., 2013. Selection of specific protein binders for pre-defined targets from an optimized library of artificial helicoidal repeat proteins (alphaRep). PLoS One 8, e71512.

Hughes, S.A., Wang, F., Wang, S., Kreutzberger, M.A.B., Osinski, T., Orlova, A., Wall, J.S., Zuo, X., Egelman, E.H., Conticello, V.P., 2019. Ambidextrous helical nanotubes from self-assembly of designed helical hairpin motifs. Proc Natl Acad Sci U S A 116, 14456-14464.

Kajander, T., Cortajarena, A.L., Regan, L., 2006. Consensus design as a tool for engineering repeat proteins. Methods Mol Biol 340, 151-170.

Kajava, A.V., 2012. Tandem repeats in proteins: from sequence to structure. J Struct Biol 179, 279-288.

Kleffner, R., Flatten, J., Leaver-Fay, A., Baker, D., Siegel, J.B., Khatib, F., Cooper, S., 2017. Foldit Standalone: a video game-derived protein structure manipulation interface using Rosetta. Bioinformatics 33, 2765-2767.

Knossow, M., Campanacci, V., Khodja, L.A., Gigant, B., 2020. The Mechanism of Tubulin Assembly into Microtubules: Insights from Structural Studies. iScience 23, 101511.

Lee, S.C., Park, K., Han, J., Lee, J.J., Kim, H.J., Hong, S., Heu, W., Kim, Y.J., Ha, J.S., Lee, S.G., Cheong, H.K., Jeon, Y.H., Kim, D., Kim, H.S., 2012. Design of a binding scaffold

based on variable lymphocyte receptors of jawless vertebrates by module engineering. Proc Natl Acad Sci U S A 109, 3299-3304.
Moreaud, L., Viollet, S., Urvoas, A., Valerio-Lepiniec, M., Mesneau, A., Li de la Sierra-Gallay, I., Miller, J., Ouldali, M., Marcelot, C., Balor, S., Soldan, V., Meriadec, C., Artzner, F., Dujardin, E., Minard, P., 2023. Design, synthesis, and characterization of protein origami based on self-assembly of a brick and staple artificial protein pair. Proc Natl Acad Sci U S A 120, e2218428120.
Parmeggiani, F., Huang, P.S., Vorobiev, S., Xiao, R., Park, K., Caprari, S., Su, M., Seetharaman, J., Mao, L., Janjua, H., Montelione, G.T., Hunt, J., Baker, D., 2015. A general computational approach for repeat protein design. J Mol Biol 427, 563-575.
Rohweder, B., Semmelmann, F., Endres, C., Sterner, R., 2018. Standardized cloning vectors for protein production and generation of large gene libraries in Escherichia coli. Biotechniques 64, 24-26.
Uribe, K.B., Guisasola, E., Aires, A., Lopez-Martinez, E., Guedes, G., Sasselli, I.R., Cortajarena, A.L., 2021. Engineered Repeat Protein Hybrids: The New Horizon for Biologic Medicines and Diagnostic Tools. Acc Chem Res 54, 4166-4177.
Urvoas, A., Guellouz, A., Valerio-Lepiniec, M., Graille, M., Durand, D., Desravines, D.C., van Tilbeurgh, H., Desmadril, M., Minard, P., 2010. Design, production and molecular structure of a new family of artificial alpha-helicoidal repeat proteins (alphaRep) based on thermostable HEAT-like repeats. J Mol Biol 404, 307-327.

Acknowledgments

This work was funded by the Agence Nationale de la Recherche (ANR, Contracts No, ANR-18-CE44-0013-Scaffold-Art, ANR-21-CE09-0045-ProteOrigami). J. Miller acknowledges the support of the Office for Science & Technology of the French Embassy in the United States through a Chateaubriand Fellowship. This work has benefited from the Electron Microscopy, CryoEM, Photonic Microscopy and Protein Interactions Platforms of I2BC, supported by French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05-05. F. A. acknowledges the Région Bretagne and Rennes Métropole for the X-ray setup support. E. D acknowledges the Région Bourgogne-Franche-Comté for financial support (contract ANER-2022-Y13835 Piccolo).

Fig S1: Sequence and repeat organization of Brick proteins

Bricks proteins are produced as recombinant proteins with N and C caps modules that can be cleaved by TEV protease. Two His-tags located at each end of the protein allow to purified the fully cleaved protein from cap modules and his tag TEV protease by Ni NTA chromatography

Schematic design of the Brick C Protein sequence:

Bricks proteins are expressed as recombinant proteins with two external repeats named N and C cap (respectively in green and red). These repeats can be cleaved by TEV protease, as indicated by arrows.

The same color code is used in the sequence shown below

Sequence of Nascent Brick C: The different components (repeats, tags, cleavage site) of the Brick protein sequence are indicated in by a colored code and noted in separate lines, although all these sequence elements are located on a single polypeptide chain.

with a single Tryptophan residue (indicated in green) and a single cysteine residue (indicated in yellow) for future specific chemical labelling although this possibility was not used in the present work

N-Histag	MHHHHHHLDM
N-cap repeat	TDPEKVDMYIENLRDEDPEVRARAAEALGKI
Linker and N TEV clevage site	GSGSGENLYFQ/GGSGSG
Repeat I3	GDEAAVYPLIQALEDEDAEVRAAAARALGKI
Internal repeat	GDERAVPALIEALKDEDAKVREAAARALGEI
	GDERAVEPLIKALKDEDAAVREAAAEALGRI
	GDERAVEPLIKALKDEDARVREAAARALGEI
	GDERAVEPLIKALKDEDWNVRKAAAEALGKI
	GDEAAVEPLIQALEDEDAEVRAAAACALGEI
Repeat I1	GDERAVPALIEALKDEDARVRAAAAKALGKI
Repeat I2	GDPEAVEALIYALRDEDANVRRAAAEALGEI
Linker and TEV cleavage site	GSGSGENLYFQ/GGSGSG
C cap repeat	GDPRAEEALRRAREDEDPEVQKEAEKAEGEI
C-Histag	GSGSGHHHHHHKSLIS

Brick C, cleaved:

GGSGSG
GDEAAVYPLIQALEDEDAEVRAAAARALGKI
GDERAVPALIEALKDEDAKVREAAARALGEI
GDERAVEPLIKALKDEDAAVREAAAEALGRI
GDERAVEPLIKALKDEDARVREAAARALGEI
GDERAVEPLIKALKDEDWNVRKAAAEALGKI
GDEAAVEPLIQALEDEDAEVRAAAACALGEI
GDERAVPALIEALKDEDARVRAAAAKALGKI
GDPEAVEALIYALRDEDANVRRAAAEALGEI
GSGSGENLYFQ

Brick D, cleaved:

GGSGSG
GDEAAVYPLIQALEDEDAEVRAAAARALGKI
GDERAVPALIEALKDEDAKVREAAARALGEI
GDERAVEPLIKALKDEDAAVREAAAEALGRI
GDERAVEPLIKALKDEDWNVRKAAAEALGKI
GDEAAVEPLIQALEDEDAEVRAAAACALGEI
GDERAVPALIEALKDEDARVRAAAAKALGKI
GDPEAVEALIYALRDEDANVRRAAAEALGEI
GSGSGENLYFQ

Figure S2 : Schematic design of Stop-N and Stop-C Bricks.

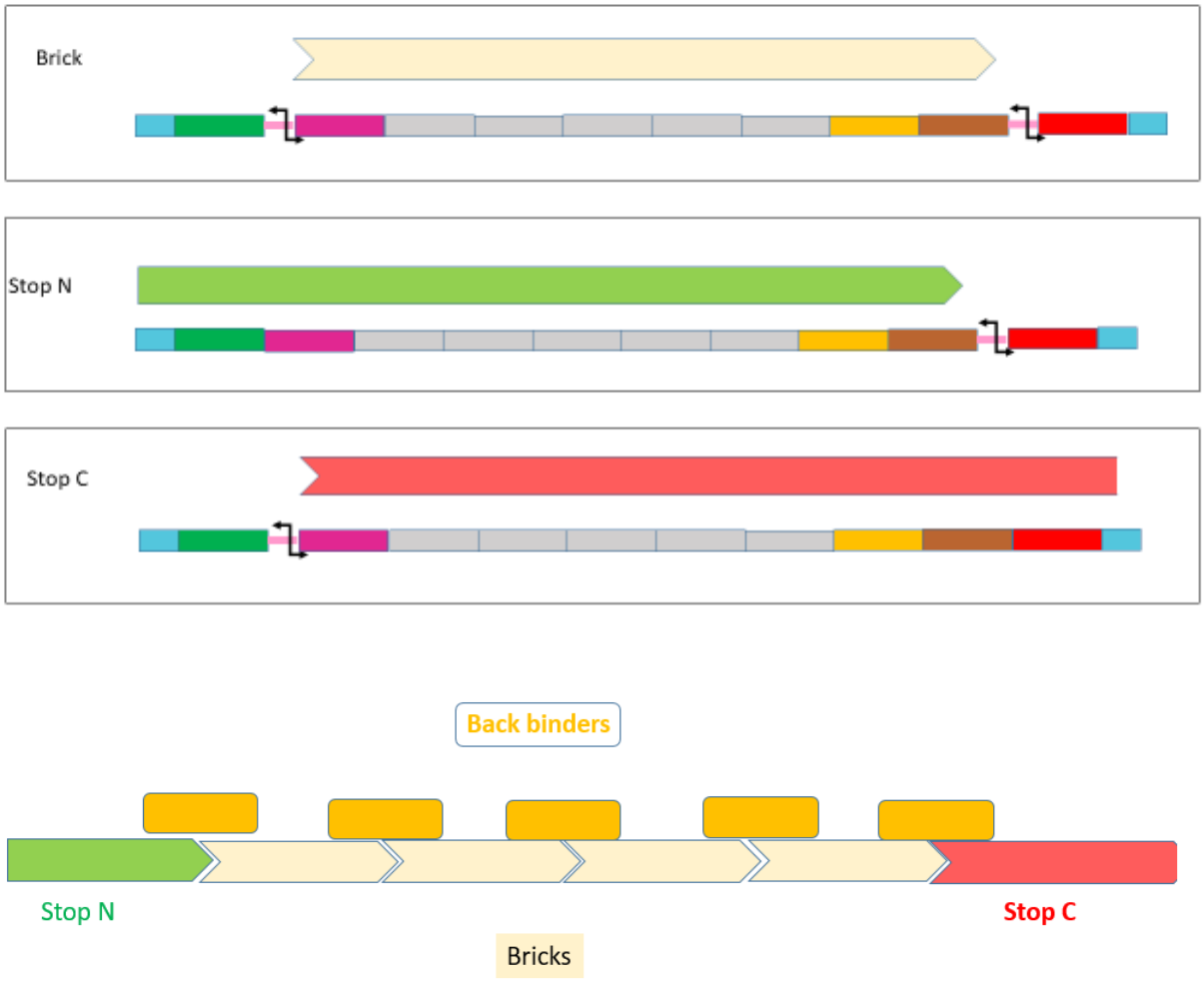


Fig S3 Sequences of proteins used

Brick N-Stop nascent:
MTDPEKVDMYIENLRDEDPEVRARAAEALGKI
GDERAVEPLIKALKDEDAKVREAAARALGEI
GDERAVEPLIKALKDEDAAVREAAAEALGRI
GDERAVEPLIKALKDEDARVREAAARALGEI
GDERAVEPLIKALKDEDWNVRKAAAEALGKI
GDEAAVEPLIQALEDEDAEVRAAAACALGEI
GDTAAVPALIEALKDEDARVRAAAAKALGKI
GDPEAVEALIYALRDEDANVRRAAAEALGEI
GSGSGENLYFQ/GGSGSG
GDPRAEEALRRAREDEDPEVQKEAEKAEGEI
LEHHHHHH

Brick N-Stop cleaved:
MTDPEKVDMYIENLRDEDPEVRARAAEALGKI
GDERAVEPLIKALKDEDAKVREAAARALGEI
GDERAVEPLIKALKDEDAAVREAAAEALGRI
GDERAVEPLIKALKDEDARVREAAARALGEI
GDERAVEPLIKALKDEDWNVRKAAAEALGKI
GDEAAVEPLIQALEDEDAEVRAAAACALGEI
GDTAAVPALIEALKDEDARVRAAAAKALGKI
GDPEAVEALIYALRDEDANVRRAAAEALGEI
GSGSGENLYFQ

Brick C Stop nascent:
MHHHHHHLDM
TDPEKVDMYIENLRDEDPEVRARAAEALGKIGSGS
GENLYFQ/GGSGSG
GDEAAVYPLIQALEDEDAEVRAAAARALGKI
GDERAVEALIEALKDEDAKVREAAARALGEI
GDERAVEPLIKALKDEDAAVREAAAEALGRI
GDERAVEPLIKALKDEDARVREAAARALGEI
GDERAVEPLIKALKDEDWNVRKAAAEALGKI
GDEAAVEPLIKALKDEDAEVRAAAACALGEI
GDERAVEPLIKALKDEDARVRAAAAKALGKI
GDPRAEEALRRAREDEDPEVQKEAEKAEGEI
Brick C Stop cleaved:
GGSGSG
GDEAAVYPLIQALEDEDAEVRAAAARALGKI
GDERAVEALIEALKDEDAKVREAAARALGEI
GDERAVEPLIKALKDEDAAVREAAAEALGRI
GDERAVEPLIKALKDEDARVREAAARALGEI
GDERAVEPLIKALKDEDWNVRKAAAEALGKI
GDEAAVEPLIKALKDEDAEVRAAAACALGEI
GDERAVEPLIKALKDEDARVRAAAAKALGKI
GDPRAEEALRRAREDEDPEVQKEAEKAEGEI

Brick F, nascent:
MRGSHHHHHH
TDPEKVDMYIENLRDEDPEVRARAAEALGKI
GSGSGENLYFQ/GGSGSG
GDEAAVYPLIQALEDEDGWVRQSAAVALGQI
GDERAVEALIEALKDEDWFVRIAAAFALGEI
GDERAVEPLIKALKDEDGWVRQSAADALGEI
GDERAVEPLIKALKDEDARVRAEAAKALGDI
GDERAVEPLIKALKDEDPEVRKAAALALGKI
GDEAAVEPLIQALEDEDSDVRAAAAWALGKI
GDTAAVPALIEALKDEDYYVRRAAAYALGKI
GDPEAVEALIYALRDEDAWVRRAAADALGQI
GSGSGENLYFQ/GGSGSG
GDPRAEEALRRAREDEDPEVQKEAEKAEGEI
LEHHHHHH

Brick F, cleaved:
GGSGSG
GDEAAVYPLIQALEDEDGWVRQSAAVALGQI
GDERAVEALIEALKDEDWFVRIAAAFALGEI
GDERAVEPLIKALKDEDGWVRQSAADALGEI
GDERAVEPLIKALKDEDARVRAEAAKALGDI
GDERAVEPLIKALKDEDPEVRKAAALALGKI
GDEAAVEPLIQALEDEDSDVRAAAAWALGKI
GDTAAVPALIEALKDEDYYVRRAAAYALGKI
GDPEAVEALIYALRDEDAWVRRAAADALGQI
GSGSGENLYFQ

α2:
MRGSHHHHHH
TDPEKVEMYIKNLQDDSSVVRKAAAVALGEI
GDERAVEPLIKALKDEDQFVRIAAAWALGKI
GGERVRAAMEKLAETGTGFARKVAVNYLETHKSLIS

Backbinder E3 nascent
MHHHHHHL
DENLYFQ/
GTDPEKVEMYIKNLQDDSIVVRYSAASALGKI
GDERAVEPLIKALKDEDGYVRQAAALALGQ
IGDERAVEPLIKALKDEDSTVRIRAARALGK
IGDERAVEPLIKALKDEDWQVRLSAASALGKI
GDERAVEPLIKALKDEDPSVRMAAANALGQI
GGERVRAAMEKLAETGTGFARKVAVNYLETH
PSETRGVPHIVMVDAYKRYK

Backbinder E3 included a C terminal Spy-tag3 sequence shown in dark green, although this tag was not used in the present work

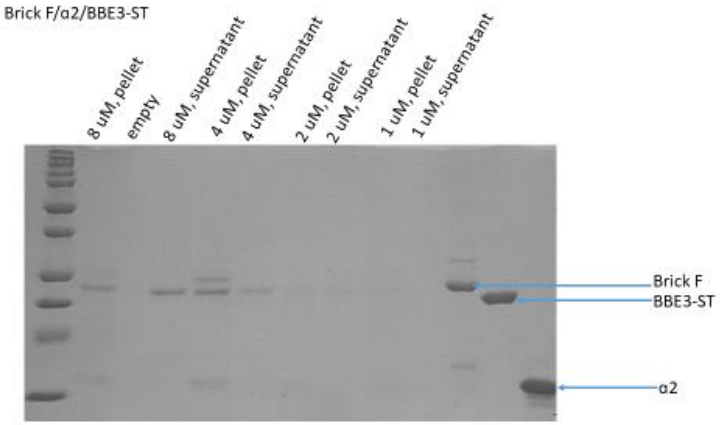


fig S4 SDS Page analysis of assembly resulting from BrickF in presence of both αRep2 and back binder proteins. Both inside and outside staple proteins proteins are present in the pellet fraction

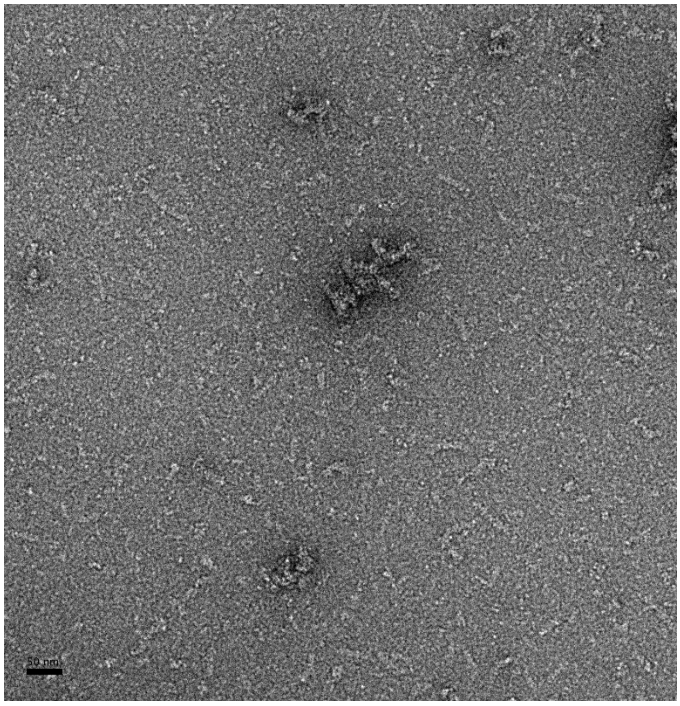


Figure S5. TEM image from supernatant fraction of Brick D with Backbinder showing short soluble filaments.

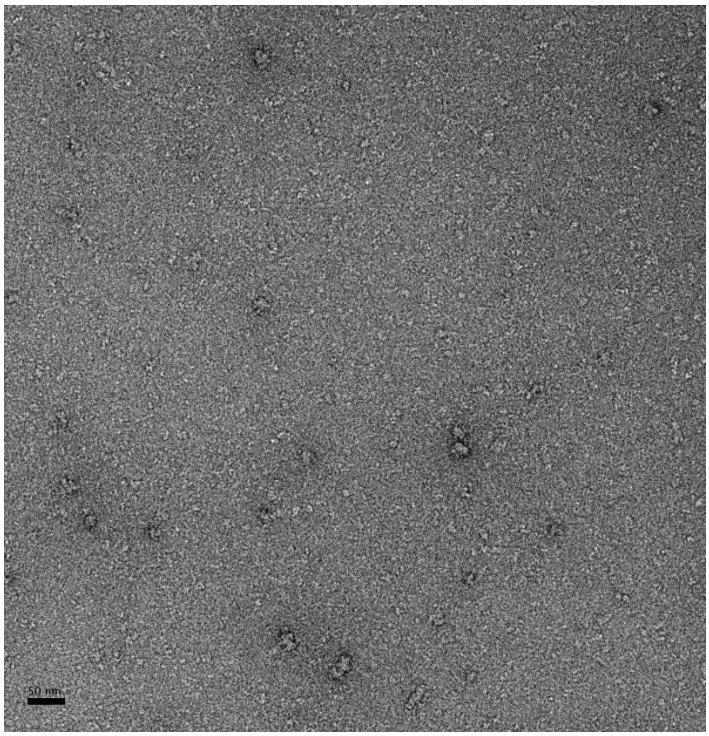


Fig S6. TEM micrograph showing 95% Brick D, 5% Brick C, and Backbinder, assembled at 8 μM and diluted to 0.8 μM.

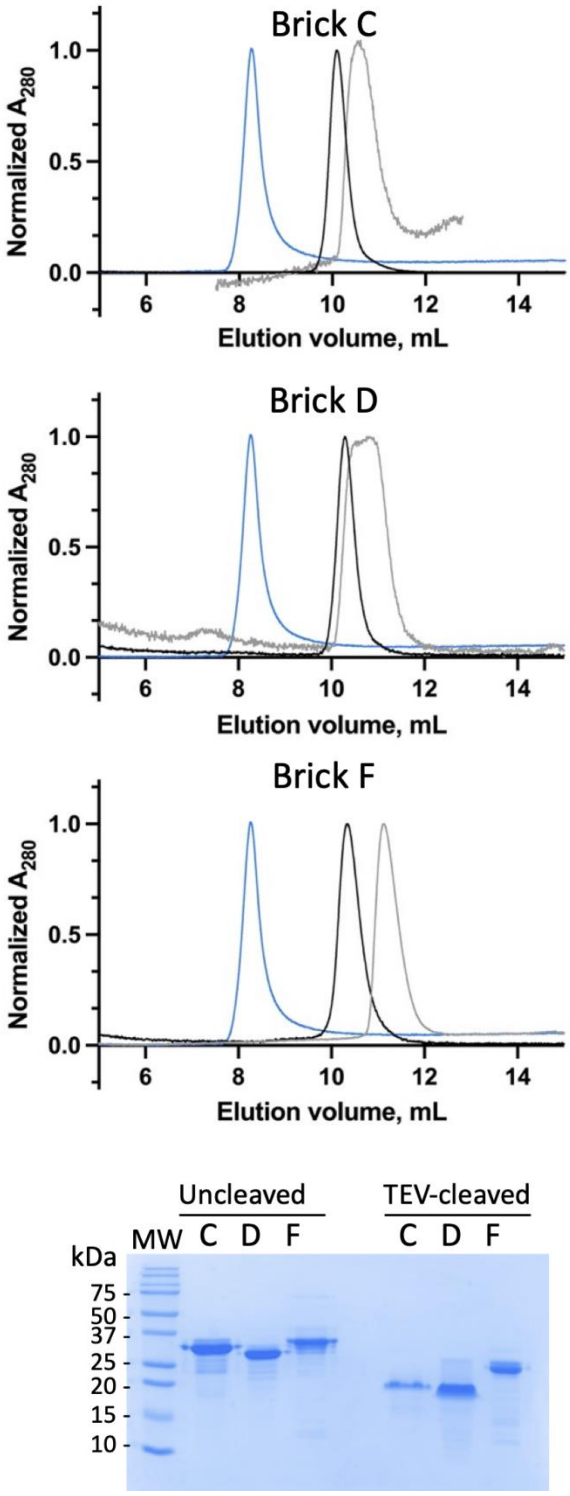


Fig S7

Analytical Size Exclusion Chromatography analysis of C, D and F bricks.

100 μL Solutions of uncleaved Brick C (110 μM), D (81 μM) or F (49 μM) were injected into an analytical 24 mL Superdex 75 (10/300) column equilibrated in TBS. For the TEV-cleaved bricks the fractions recovered from the Ni-NTA column flowthrough were injected on the Superdex 75 column. Blue Curves correspond to Dextran, a polymer eluted in the exclusion volume. For each panel, the Black curve corresponds to un-cleaved Bricks (C, D or F) and the grey curve corresponds to TEV-cleaved Brick (C, D or F). Bottom panel: SDS-PAGE analysis of un-cleaved and TEV-cleaved proteins injected on the size exclusion column. 5 ug of each protein was analyzed on a 15% polyacrylamide gel under denaturing conditions.

For each brick, the TEV-cleaved form is eluted in a larger volume than the un-cleaved form, indicating that the TEV digestion was efficient, as observed by SDS-PAGE. No peak is observed in the exclusion volume, showing that the bricks do not assemble or aggregate once cleaved with TEV protease. The assembly process requires the addition of protein staples.