Bipartite design of a self-fibrillating protein copolymer
with nanopatterned peptide display capabilities

Marc Bruning¹,⁴,⁵, Laurent Kreplak²,³, Sonja Leopoldseder¹, Shirley A. Müller², Philippe Ringler², Laurence Duchesne⁴,⁵, David G. Fernig⁴,⁵, Andreas Engel², Zöhre Ucurum-Fotiadis¹, Olga Mayans¹,⁴,⁵*

¹Division of Structural Biology & ²Maurice E. Müller Institute, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland; ³Department of Physics and Atmospheric science, Sir James Dunn Building, Dalhousie University, Halifax, NS, B3H 3J5, Canada; ⁴School of Biological Sciences & ⁵Liverpool Institute for Nanoscale Science Engineering and Technology, Biosciences Building, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK.

*Corresponding author: Olga Mayans
Tel: +44 (0)151 7954472 / Fax: +44 (0)151 7954406 / Olga.Mayans@liv.ac.uk

Supp. Material: Sections S1-S4
SECTION S1: Production of building-blocks and fiber samples

Methods

Cloning

The Z_{1212} tandem was constructed by adjoining two Z1Z2 fragments from human titin (residues 1–196; EMBL X90568) using overlap extension PCR. The C- and N-termini of Z2 and Z1, respectively, were adjoined so that this resulted in a GETTQ linker sequence. Cloning was in the pETM-11 vector (EMBL plasmid collection) using NcoI and SacI restriction sites, which incorporated a His_{6}-tag and a TEV protease cleavage site N-terminal to the target construct.

The Tel truncated variant (residues 1-87; Swiss-Prot O15273) was cloned into the vector pETM-11 using the NcoI and KpnI restriction sites. Residues C8, C15, C38 and C57 were mutated into serines to prevent unspecific aggregation due to oxidation.

Protein production

Overexpression of Z_{1212} and Tel was in *E. coli* Rosetta (Novagen). Cells were grown in Luria Bertani medium supplemented with 25 µg/mL kanamycin and 34 µg/mL chloramphenicol up to an OD_{600} of 0.6. Expression was induced with 1 mM IPTG and cultures grown further at 20°C overnight. Cells were harvested by centrifugation. Cell pellets corresponding to Z_{1212} expressions were resuspended in a lysis buffer containing 50 mM Tris-HCl pH 7.4, 100 mM NaCl, supplemented with 1 mg/mL Lysozyme and 0.01 mg/mL Dnase I. Cell pellets of Tel expressions were resuspended in a lysis buffer containing 25 mM Tris-HCl pH 7.4 and 8M Urea. In both cases, resuspended cells were lysed by sonication and lysates clarified by centrifugation. Supernatants containing either Z_{1212} or Tel were applied to a Ni^{2+}-NTA Hi-Trap column (GE Healthcare) equilibrated in the corresponding lysis buffer. Elution used 200 mM imidazole. For Z_{1212}, eluted fractions were dialyzed against lysis buffer in the presence of TEV protease and 1
mM DTT for tag removal. Further purification was by reverse metal affinity (Ni$^{2+}$-NTA Hi-Trap column, GE Healthcare) and size exclusion chromatography (Superdex 200 HR16/60 column; GE Healthcare) equilibrated in lysis buffer. Tel was purified using a Ni$^{2+}$-NTA Hi-Trap column (GE Healthcare) in the presence of urea, but neither tag removal nor gel filtration was performed. Finally, $Z_{1212}$ and Tel samples were concentrated to approx. 32 mg/mL and 11 mg/mL, respectively. $Z_{1212}$ was stored at 4 °C and Tel at RT until further use.

**Assembly**

To promote assembly, Tel and $Z_{1212}$ were mixed at a molar ratio of 3:1 until a final total protein concentration of 10 mg/mL had been attained in a total volume of 200 µL in 50 mM Tris-HCl pH 7.4, 100 mM NaCl (assembly buffer). Mixtures were then dialyzed against assembly buffer at 37°C for approx. 18 hours and later stored at RT for periods of up to 2 months. Fibers could be observed after 2 hrs and they were well developed by the final dialysis time of 18 hrs.

Regarding sample stability, assembly was also assayed at 18 °C and obtained comparable results. Sample storage at 4°C did not cause alterations to the assemblies. As standard, we used pH 7.4 in our protocols, but additional bibliographic data exist on the interaction of the biological complex Z1Z2/Tel that show that this is stable and structurally integral in the pH range 4.45 to 8.0. Since the polymerization of our fibers is strictly based on the cross-linking effect of this natural protein complex, it can be concluded that the stability of the fibers cannot be less than the stability of its building blocks and the interactions that they form.

**Results**

The $Z_{1212}$ tandem was obtained as a soluble product in high yields (in excess of 65 mg pure protein per gram of *E. coli* wet cell mass), could be concentrated to elevated values (>50 mg/mL), showed exceptional stability upon storage at 4°C and did not undergo degradation.
during the maximal period assayed, which was 6 months. In brief, Z_{1212} showed excellent properties for production and storage.

The cross-linking protein Tel was engineered and produced in yields of ~10 mg of pure protein per gram of *E. coli* wet cell mass and showed similar high stability during storage. Given that Tel is intrinsically unstructured and only folds by binding to Z1Z2, it was produced and stored under denaturing conditions (typically 6M urea) to retain solubility.
Fig S1: Production of building blocks and polymeric assemblies

a. SDS-PAGE of the purified Z\textsubscript{1212} tandem sample at several concentrations. The molecular mass (Mr) in kDa of protein markers is given; b. SDS-PAGE of soluble samples recovered from mixing experiments (various concentrations shown); c. Native-PAGE of assembled samples revealing multiple species of high Mr. Lanes 1-4 correspond to assembly experiments at increasing sample concentrations: 2, 5, 10, and 15 mg/mL, respectively.
SECTION S2: Estimation of the oligomeric state of engineered Z₁₂₁₂ tandems by size exclusion chromatography combined with multi-angle light scattering (SEC-MALS)

Methods

SEC-MALS measurements were performed on an ÄKTA basic system (GE Healthcare) connected to a tri-angle static light-scattering detector and a differential refractometer (miniDAWN Tristar and Optilab, respectively). A Superdex 200 HR 10/300 GL column (GE Healthcare) was used, equilibrated in assembly buffer and at a flow rate of 0.7 ml/min. Samples were injected at volumes of 100 µL and at a concentration of 2 mg/mL. A value of 0.185 mL/g was used as specific refractive index increment (dn/dc). Recording and processing of data used the ASTRA software (Wyatt Technology, Santa Barbara, CA, USA).

Results

The oligomeric state of the Z₁₂₁₂ tandem could influence the assembly of the projected fibers. States other than monomers could either hinder the productive binding of telethonin or favor the formation of fibril bundles. Thus, we investigated the association state of this tandem using SEC-MALS. This technique allows an accurate determination of the Mr of a protein sample without being influenced by molecular shape or hydrodynamic parameters. For Z₁₂₁₂, the measured Mr was 39.0 kDa (protein concentration estimated using refractive index), which is in excellent agreement with the theoretical value of 41.6 kDa calculated from sequence data for a monomer of this sample. Thus, it can be expected that the engineered tandems are suitable building blocks for the formation of mono-fibers with reduced or no tendency to bundling.
Fig S2: SEC-MALS measurement of $Z_{1212}$

The measured Mr (red) and normalized refractive index (continuous black line) are shown. The theoretical Mr of the monomeric state is given as a dashed horizontal line.
SECTION S3: Electron Microscopy studies

Methods

Transmission Electron Microscopy (TEM)

Samples were adsorbed onto copper grids directly as recovered from the assembly mixture or after diluting 10 and 100 times with assembly buffer. Glow-discharged 200 mesh carbon-coated copper grids were incubated with 5 µl sample volume for 60 sec and washed three times with distilled water. Following, grids were stained twice with 5 µl of 2% uranyl-acetate for 5 sec the first time and for 45 sec the second time. Excess fluid and stain was removed with filter paper. Imaging was carried out on a Philips CM-100 transmission electron microscope operating at 80 kV.

Gold nanoparticles (AuNPs) functionalized with TrisNiNTA groups were used to label the fibers. AuNPs (1 µL at 500 nM concentration) having an average of 1.5 TrisNiNTA tags per nanoparticle were added to 9 µL of the assembled mixture (undiluted or at dilutions 1:10, 1:100, 1:1000) and incubated overnight at room temperature in assembly buffer. 5 µL of the sample was then deposited on the grids as described above.

Scanning Transmission Electron Microscopy (STEM)

STEM was carried out on assembly products in 50 mM Tris-HCl pH 7.4, 100 mM NaCl. Samples were diluted 1000x in buffer and aliquots of 5 µl adsorbed onto glow discharged thin carbon film of STEM microscopy grids (thin carbon supported by a perforated thick carbon layer covering 200-mesh, gold-coated copper grids). The grids were blotted, washed on 8 droplets of quartz double-distilled water (blotting between each step), plunged into liquid
nitrogen, and freeze-dried at -80 °C and 5 × 10⁻⁸ torr overnight in the microscope. Tobacco mosaic virus (TMV) adsorbed onto a separate grid and air-dried served as mass standard. A Vacuum Generators STEM HB-5 interfaced to a modular computer system (Tietz Video and Image Processing Systems GmbH) was employed. Digital dark-field images were recorded from the unstained grids at an accelerating voltage of 80 kV, a nominal magnification of 200,000x and doses of 449±14 electrons/nm². Images were evaluated using the MASDET program package as detailed in Krzyzánka et al.⁴ The MPL values were corrected for the experimentally determined, beam-induced mass-loss, scaled as indicated by the mass standard, binned, displayed in a histogram, and described by a single Gaussian curve.

Results

Mass-per-length Analysis of Tape-like Fibers by Scanning Transmission Electron Microscopy (STEM)

To investigate the packing of building blocks within the fibers, unstained samples were imaged by STEM and their mass-per-unit length (MPL) calculated. The mass of the thinner fibers (ii) could not be estimated using this approach, because their curled up arrangements resulted in excessive overlap that prevented accurate mass calculations. STEM measurements of the thicker fibers (i) yielded a mass histogram with a broad peak at a MPL of 53±11 kDa/nm (Fig 3). Although the accuracy of this measurement was reduced by a high background caused by the presence of unassembled species on the STEM grids, the high mass of the Z₁₂₁₂/Tel building blocks (110 kDa) make the data suitable to define the assembly to a first approximation. The data indicated the presence of a full Z₁₂₁₂/Tel unit per 2 nm of fiber length. Based on molecular dimensions calculated from crystal structures (Fig 1, main text), the MPL data support a transverse organization where building blocks stack perpendicular to the fibril axis
with one single $Z_{1212}$/Tel unit acting as the translational element along the fibril length. A longitudinal arrangement, where $Z_{1212}$ tandems lie parallel to the fibrillar axis, would only satisfy the measured MPLs if ~9 $Z1Z2$/Tel units were laterally associated (longitudinal MPL of a $Z1Z2$/Tel unit $\approx 6$ kDa/nm), which would result in larger fiber diameters than those observed. Thus, STEM data support the computational models for this fiber morphology (i).

Figure S3: STEM measurements of thick fibers

Mass histogram (n=108 fiber segments) with an average mass-per-length (MPLAV) of 53±11 kDa/nm, where the main peak (n=99; delimited by dotted lined) corresponds to a MPLAV of 54±10 kDa/nm.
SECTION S4: Computer simulations of propagative fiber assembly modes

Methods

$Z_{1212}$ tandems were modeled based on the crystal structure of the Z1Z2/Tel complex (PDB code 1YA5)\(^1\) as follows: a Z1Z2 pair was extracted from the said complex and juxtaposed onto another identical pair by least-squares fitting of the N, C, O and C\(\alpha\) atoms of threonine residues common to the beginning and end of the Z1Z2 structures to be fused (T197 C-terminal to Z2 and T2 N-terminal to Z1, respectively). This led to a VQGETTQA linker sequence (residues in bold are integral part of the Ig folds of Z1 and Z2), where residues TT are predicted to be the main components of a mechanical hinge providing conformational freedom to the tandem. The resulting model (Fig S4a) was used as an initial template to generate a family of conformers with different linker arrangements using the program tCONCOORD\(^5\), which explored the degrees of freedom of the TT motif in the linker sequence. To only allow variations in the linker, the Z1Z2 doublets were maintained as rigid bodies by constraining the interatomic distances of their component C\(\alpha\) atoms. The 100 resulting models (Fig S4b) were energy minimized and subsequently used in the fiber assembly simulations.

Computational assembly was carried out using the coordinates of $Z_{1212}$ conformers so generated and the crystal structure of the Z1Z2/Tel complex. The latter was used as template to position Tel at the point where the tandems interconnect. An algorithm was developed in Perl scripting using routines from LSQKAB\(^6\) and PyMOL\(^7\) that performed model superposition and monitored intermolecular distances, respectively. In this algorithm, each cycle of assembly consisted of the following steps: \(i\) a given $Z_{1212}$ model was selected randomly; \(ii\) the Z1Z2/Tel complex was least-square superimposed onto one of the Z1Z2 pairs; \(iii\) a second $Z_{1212}$ tandem was superimposed onto the remaining Z1Z2 pair of Z1Z2/Tel; \(iv\) the Z1Z2 domains of the Z1Z2/Tel complex were removed leaving an assembly unit consisting of two $Z_{1212}$ tandems.
“glued” through a telethonin molecule. Repetitive cycles of this process allowed the building of a fiber model composed of any desired number of Z$\textsubscript{1212}$ units. Longitudinal and transversal assembly modes were analyzed.

**Fig S4: Models**

- **a.** Initial Z$\textsubscript{1212}$ model; **b.** Family of 100 unique conformers calculated using tCONCOORD; **c.** Tilt and twist angles between Z1Z2 components defining the overall arrangement of the 100 conformers (represented as open circles). Here, tilt angles were defined as the angle between the C$\alpha$ atoms of the N-terminal, mid-hinge (residue 196) and C-terminal residues. The twist angle was defined as the dihedral angle between the longitudinal planes containing the Z1Z2 pairs of a given tandem. Such planes were defined by the C$\alpha$ atoms of the first and last residues of the inter Z1Z2 linkers (residues 193 and 199) and the apices of the C’D loop of the adjoining Z1/Z2 domains (residues 148 and 244) of the Z1Z2 pair considered. Models for which fibril formation through stacking interactions is shown in **e.** are indicated as filled squares and their overall conformation is given in **d;** **e.** Models of transversal fiber assembly resulting from several tandem conformations (twist and tilt angles are given). It should be noted that the increase in fibril helicity correlates with the increase in twist angle between Z1Z2 pairs. In solution, a fiber so assembled could be expected to undergo dynamic transitions across the several permitted states.
References


