Chimeric single α-helical domains as rigid fusion protein connections for protein nanotechnology and structural biology

Anna-Sophia Krebs[1], Gabriella Collu[1,2], Tobias Bierig[1,2], Niveditha Varma[1] and Roger M. Benoit[1*]

[1]Laboratory of Nanoscale Biology, Division of Biology and Chemistry, Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland

[2]Department of Biology, ETH Zürich, 8093 Zürich, Switzerland

[*]Correspondence to:
E-mail: roger.benoit@psi.ch

Abstract

Fusion proteins are essential tools for heterologous protein expression, purification and crystallization. Non-optimized connections between a fusion protein and the protein of interest are usually flexible. Here we show that the ER/K α-helical motif (1) can be seamlessly fused to terminal helices of proteins, forming a chimeric, partially free-standing rigid helix. This modular building block allows the rigid connection of two proteins. Through the intrinsic stability of the ER/K helix, the fusion protein and the protein of interest can be separated by a desired distance, assuring natural freedom of conformation in the latter. We solved the crystal structure of an engineered fusion protein at 1.9 Å resolution. The construct comprises an N-terminal calmodulin-binding peptide (2,3), an ER/K α-helical motif (1), yellow fluorescent protein (YFP)(4), a flexible linker and a C-terminal calmodulin (CaM)(2). The ER/K α-helical motif forms a free-standing rigid helix that is continuous with the N-terminal α-helix of YFP. Calmodulin is bound to the calmodulin-binding peptide, increasing the mass of the rigid fusion. This molecular biomimetic, protein nanotechnology approach enables new possibilites for the engineering of crystal contacts and for the design of rigid protein particles with increased mass, for example to enable structure elucidation of small proteins by cryo-electron microscopy.

Results and Discussion

Nature provides a myriad of proteins or protein assemblies that can be modified and exploited for nanotechnology (5). For example, virus capsid proteins can self-assemble into virus-like particles, which are highly organized spheres (6). In combination with recombinant DNA technology, protein self-assembly can be exploited to design proteinaceous nanoparticles that display antigens at a high local concentration, for example to increase the immunogenicity of vaccines, or alternatively as building blocks for nanomaterials. The orientation of the displayed proteins or peptides can be critical. Many applications even require rigidity throughout the complete fusion protein. Rigid connections are difficult to design without extensive experimental optimization, therefore, the engineering of rigid fusion proteins constitutes an important bottleneck in protein nanotechnology.

The engineering of large particles for the structure elucidation of small proteins by cryo-electron microscopy (cryo-EM) is a contemporary example of an application that is highly important for structure-based drug design and is just now coming of age. Recent developments, for which the Nobel Prize in Chemistry 2017 was awarded, enabled the structure elucidation of biological molecules at near-atomic resolution by cryo-electron microscopy (reviewed in Ref.(7)). A molecular mass of typically several hundred kDa is required to obtain a good signal-to-noise ratio, and clearly identifiable structural features are important to allow the correct angular orientation of the particles (8). The mass of many biologically important proteins is below 100 kDa. Simple fusion proteins or self-assembling fusion proteins (9) can be used to increase the size of the particles and to enable structure elucidation by cryo-EM. To allow straightforward interpretation of the EM data, the particles have to be rigid.

The aim of our investigation was to test whether it is possible to rationally design rigid chimeric fusion proteins that are connected by a segment of a naturally occurring, stable single α-helical domain (SAH). Such an SAH connection, when introduced as a seamless extension of an N- or C-terminal helix or as an extension of a helix at an intracellular loop, only contacts the protein of interest at a single site. Therefore, the conformational freedom of motion within the protein should not be affected. This represents an important advantage over the use of binders such as nanobodies (10) or designed ankyrin repeat proteins (11), which require three-dimensional epitopes for rigid binding and hence typically influence the conformation of the protein of interest.

Unlike typical α-helices, whose stability depends on contacts with other regions in the tertiary structure of proteins, SAHs are stabilized by dynamic charge interactions between the side-chains, resulting in intrinsic stability and rigidity (1,12). SAHs can for example be found in smooth muscle caldesmon (13), in myosin-X (14) and in the ribosomal protein L9 from *Bacillus stearothermophilus* (15). SAHs have been extensively characterized in solution using a wide variety of methods, including NMR (16), FRET studies (17) and small angle X-ray scattering (18). The stability of SAHs has furthermore been confirmed by crystallographic studies of such helices in their natural context (14,15).

We tested whether it is possible to take a segment of an SAH out of its natural context and to use it as a modular architectural element to attach rigid extensions to proteins. Unlike for the application as a simple spacer, to be useful as a tool for structural studies, the transitions between the rigid free-standing helix and the fusion proteins to which it is attached have to form a seamless, continuous chimeric helix, to assure rigidity throughout the fusion construct. We used a segment of the ER/K motif from myosin-X (14) to engineer a fusion protein containing a rigid linker. Upstream of the rigid ER/K motif, we fused a calmodulin-

binding peptide (CBP) from skeletal muscle myosin light-chain kinase (MLCK), which binds the protein calmodulin (CaM) with very high affinity (3). The unbound CBP is a random coil (19) that adopts an α-helical structure upon binding CaM, which wraps almost completely around the peptide (2). Because the CBP from sMLCK, unlike the ER/K motif, is not intrinsically helical and rigid (19, 20), it can be used as a flexible tag or linker that can later be rigidified and increased in mass by the addition of CaM. In our construct, the CBP-ER/K module was fused to the N-terminus of enhanced yellow fluorescent protein (EYFP, GFP-10C)(4), which was used as a test protein that can be replaced by any protein of interest. The color and fluorescence of YFP simplified protein purification and the identification of protein crystals. CaM was fused to the C-terminus of YFP via a flexible linker, to allow intramolecular binding to the N-terminal CBP. The overall structure of the fusion protein, henceforth referred to as YFPnano, is shown in **Fig. 1**, and the data collection and refinement statistics are listed in Table 1.

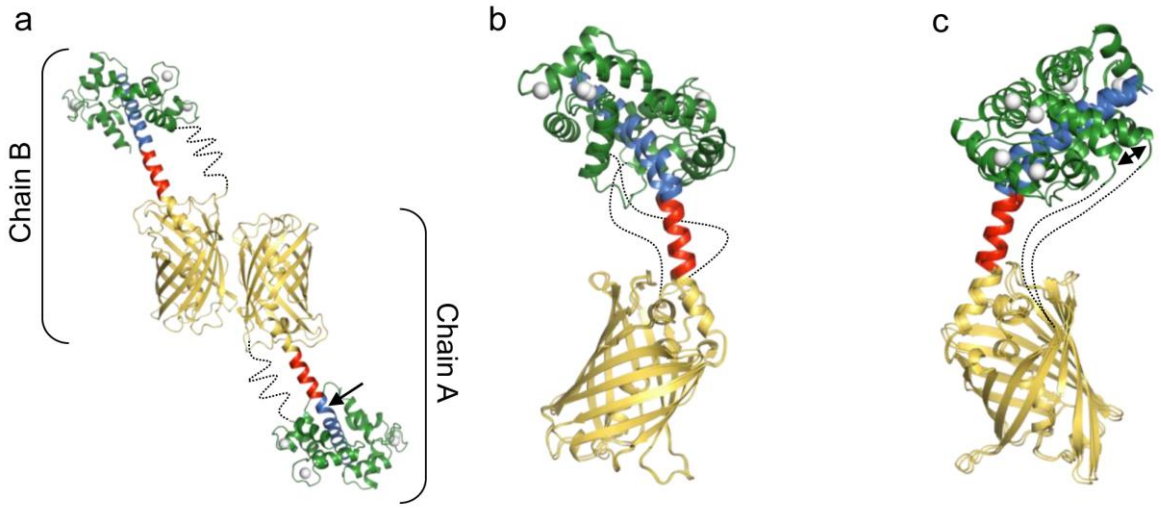

Fig. 1) Structure of the YFPnano fusion protein. Cartoon representation of the crystal structure of the CBP (blue) - ER/K motif (red) - YFP (yellow) - linker (dotted lines) - CaM fusion protein. Dotted lines indicate the flexible linker between YFP and CaM that is not visible in the structure. Calcium ions are shown as white spheres. **a,** The two molecules in the asymmetric unit are shown. The ER/K motif forms a stable, seamless extension of the short N-terminal helix of YFP. CaM is intramolecularly bound to the CBP in both molecules in the asymmetric unit. The CBP helix is continuous with the ER/K helix. In chain A, there is a kink within the CBP helix, indicating that this is the weakest region of the extended helix (arrow). This kink is probably induced by forces originating from the crystal packing. **b and c,** Superimposition of the two YFPnano chains in the structure from different views. The arrow in c indicates the difference in the positioning of the most N-terminal CaM helix just downstream of the flexible linker, which may be induced by pulling forces from the linker.

Table 1) Data collection and refinement statistics

Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	53.7, 117.8, 84.5
α, β, γ (°)	90, 99.9, 90
Resolution range (Å)	48.59-1.901 (1.97-1.90)
R_{meas}	0.10 (1.02)
$CC_{1/2}$	0.998 (0.696)
I/σI	14.12 (1.62)
Completeness (%)	99.88 (99.57)
Multiplicity	6.5 (5.1)
Total reflections	526,754 (41,413)
Unique reflections	81,149 (8,052)
R_{work} / R_{free}	0.1808 / 0.2138
No. atoms	
Macromolecules	6,496
Ligands	124
Solvent	366
B-factors	
Macromolecules	37.27
Ligands	39.98
Solvent	38.64
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	0.94

Highest-resolution shell in parentheses.

We designed a single YFPnano construct, without optimization of any linkers. Indeed, the architecture of the final structure was largely as expected, indicating that the use of rigid helix segments as linkers minimizes the need for construct optimization. Overall, the CaM-CBP binding mode in our structure (Fig. 2) is in good agreement with the solution NMR structure described by Ikura et al. (2), based on which our construct was designed. The calcium-sensing protein calmodulin consists of two domains. The N-terminal domain comprises amino acids 4 -74 and the C-terminal domain comprises residues 82-147. The two domains are connected by a loop in the center (amino acids 75 - 81).

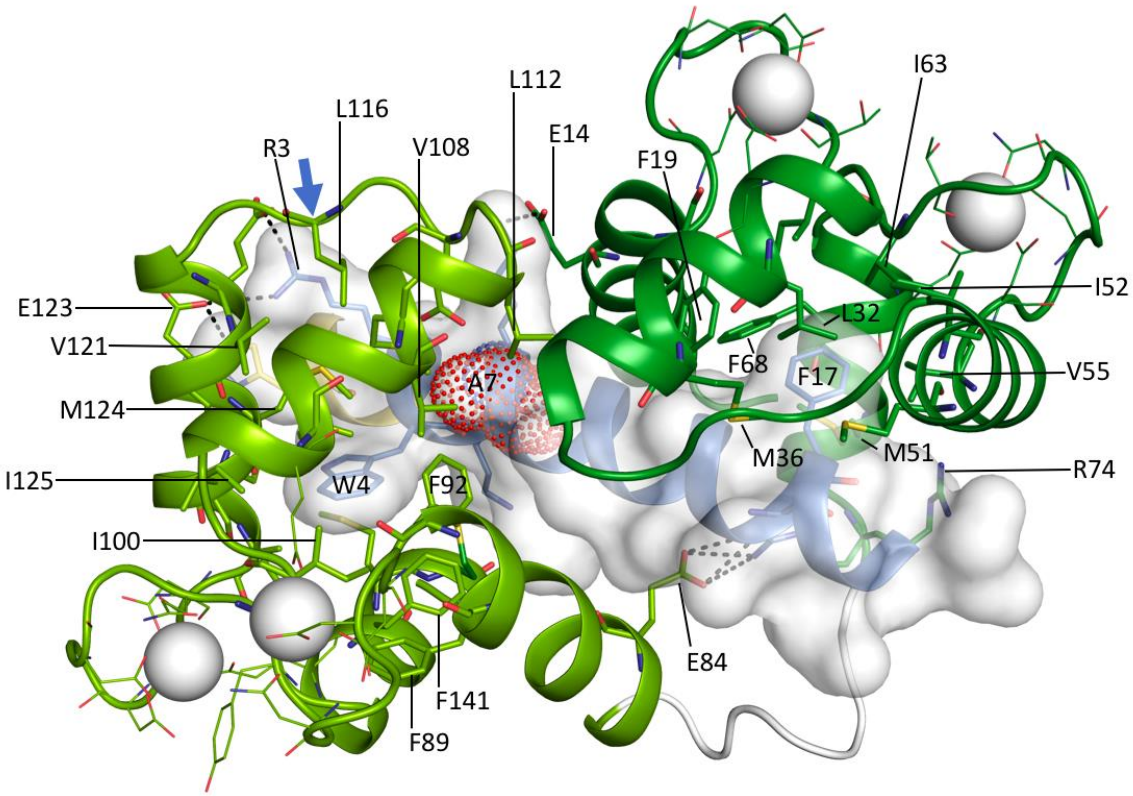

Fig. 2) CaM - CBP interactions in YFPnano. The calmodulin-binding peptide is shown as a cartoon (blue) and as a transparent surface (white). Alanine 7, the affinity-increasing mutation, is shown as red dots. Calmodulin is shown as a cartoon in green (light green = C-terminal domain, dark green = N-terminal domain). Residues forming the hydrophobic pockets of CaM in which the peptide residues W4 and F17 are anchored are shown as sticks (CaM C = green, CBP C = blue, N = dark blue, O = red, S = yellow). Calcium ions are depicted as white spheres, and the residues that coordinate the ions are shown as lines (C = green, N = dark blue, O = red). Polar interactions are indicated by black dotted lines. The loop connecting the two CaM domains is shown in white. The yellow part of the CBP helix indicates the two residues from the rhinovirus 3C protease cleavage site that remained after cleaving off the N-terminal 6xHistidine purification tag from the YFPnano fusion protein.

Each domain encompasses two EF hand helix-loop-helix motifs, each coordinating one calcium ion. One CaM molecule can hence bind up to four calcium ions. Peptide binding by CaM is calcium-dependent (3, 20). The two loops, each coordinating one calcium ion in one of the two EF hand motifs of a CaM domain, lie in an antiparallel direction to each other and engage in loop-to-loop, backbone-to-backbone hydrogen bonds between Ile 27 and Ile 63.

The sMLCK peptide is buried in a large hydrophobic channel of CaM, the CBP-CaM interaction is dominated by hydrophobic interactions.

Tryptophan 4 and phenylalanine 17 of the CBP form hydrophobic anchors that insert into hydrophobic pockets that are located between the four helices of the two EF-hand motifs of the C-terminal, respectively of the N-terminal domain of CaM (Fig. 2 and Ref.(2)). CaM residues F92, I100, L105, M124, A12, V136, M144, M145, and more distantly F89, M109, V121, I125 and F141 provide a hydrophobic environment around CBP W4, while CBP F17 is embedded in a pocked that is formed by the CaM residues F19, I27, L32, M36, M51, I52, V55, I63, F68 and M71.

Inspection of the YFPnano structure and electron density reveals that only few hydrogen bonds and salt bridges are involved in the CaM-CBP interaction, notably H-bonds between CaM M71 O and CBP R16 NH1, R74 O and R16 NH1, M72 O and R16 NH1, E54 OE2 and S21

OG, D50 O and S21 OG, and salt bridges between E123 OE1 and R3 NH1, E120 OE2 and R3 NH2, E14 OE1 and K6 NZ, E84 OE1 and R16 NH2, E84 OE1 and R16 NH2.

To assure formation of a very stable complex, we used a CBP sequence with a single-point amino acid mutation (Asn-to-Ala) in the MLCK M13 peptide that results in super-high affinity to CaM (K_d = 2.2(±1.4) × 10^{-12} M, corresponding to a ~1000-fold affinity improvement compared to the wt peptide)(3). To the best of our knowledge, our YFPnano structure is the first structure of CaM in complex with the M13 MLCK peptide containing this Asn-to-Ala single-point amino acid mutation. In YFPnano, this residue is in proximity of the hydrophobic CaM residues F92, V108, M109, L112 and L116 (Fig. 2) and of the negatively charged amino acid E114. Alanine fits better into this mostly hydrophobic environment than the corresponding polar asparagine from the wild type peptide.

In YFPnano, we included the CBP and the CaM within the same construct, at opposite termini of YFP, to promote an intramolecular interaction. When applying this strategy for purposes such as increasing the mass of a fusion protein, CaM can instead be produced separately and then be added to the CBP-tagged protein of interest after expression. In this way, the target protein only has to be altered minimally, by the addition of an N- or C-terminal CBP/(ER/K) tag.

When designing fusion proteins that contain both, CBP and CaM, it is of key importance to take into consideration the geometry of exposure of the CBP with respect to CaM. In YFPnano, the CBP and CaM were connected to opposite termini of YFP. YFP adopts a β-barrel structure with both termini exposed towards the same side of the protein. This geometry resulted in intramolecular CBP-to-CaM binding. If intramolecular binding is geometrically hindered or unfavorable, constructs would be expected to either form dimers or long oligomers from intermolecular CBP-CaM interactions, which could circularize at a certain length, depending on the flexibility and relative orientation of the linkers. While such self-assembling "protein strings" may be interesting for other applications in the field of protein nanotechology, for structural biology, this type of oligomerization would only be desirable if the final assembly is rigid.

The crystal structure of YFPnano unveiled another important property of the rigid helix extension: inspection of sections through the crystal packing (Fig. 3) reveals the central role of the rigid extension for crystal contact formation. The CaM-bound extended CBP-ER/K helix forms a pillar that enables the formation of crystal contacts at a distance from the protein of interest. This could open new possibilities for proteins that are otherwise difficult to crystallize, such as membrane proteins that are embedded in detergent micelles. Rigid extended helices as designed structural elements may hence also be important for the newest generation of crystallography techniques, including 3D electron crystallography and time-resolved crystallography using X-ray free-electron lasers.

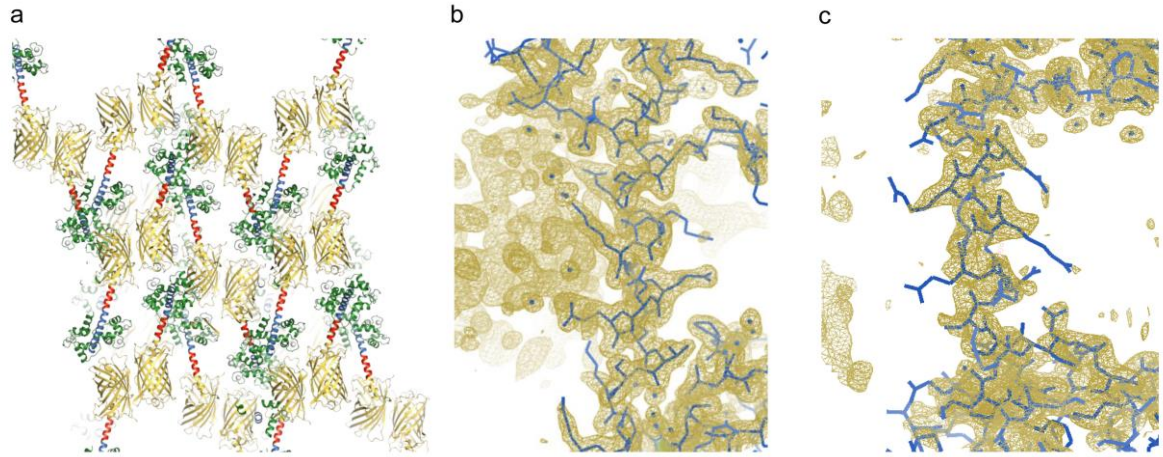

Fig. 3) The crystal lattice is built on rigid helix pillars. a, Section through the crystal packing of the YFPnano crystal, which belongs to space group $P2_1$. The view emphasizes the key role of the rigid helix pillars in the crystal packing. The use of such rigid extensions, here decorated with a helix-binding protein (CaM), enables the formation of crystal contacts at a significant distance from the protein of interest. The CaM as well as the exposed helix provide polar surface area for crystal contacts. This strategy should allow new possibilities for the crystallization of membrane proteins, which are typically surrounded by detergent micelles and often cannot easily make ordered crystal contacts. **b and c**, Electron density in the region of the ER/K helix, showing that the core of the helix is well defined in the structure (yellow = 2Fo-Fc map, blue = model). **b,** The ER/K helix in chain A is stabilized by a crystal contact. The exposed side chains are mostly well defined. **c,** The ER/K helix in chain B does not engage in direct crystal contacts and is fully exposed. The side chains are mostly undefined and B-factors are elevated in this region, however, there is clear electron density around the helix core, confirming that the ER/K motif indeed forms a stable, exposed helix.

In conclusion, our YFPnano structure demonstrates that it is possible to rationally design seamless transitions between an N-terminal α-helix of a protein of interest and a stable, free-standing ER/K single α-helical domain, forming an extended, chimeric helix. We furthermore extended the ER/K helix with a calmodulin-binding peptide and decorated the peptide with calmodulin. Our results show how stable single α-helical domains (SAHs) can be used as modular structural elements for the rigid connection of two protein domains. In solution or in single-molecule studies, the single helix also functions as a spacer that prevents undesired interactions between the protein of interest and the connected fusion protein. For crystallization, all exposed parts of the helix and CaM provide soluble surface area that can engage in crystal contact formation. This molecular biomimetic, protein nanotechnology approach of fusion protein design should be useful whenever rigid connections between proteins are required, not only for structural biology, but also as building blocks for a wide range of protein nanotechnology applications, such as the design of protein or peptide nanoparticles, protein cages or nanofibers.

Materials and Methods

Cloning and protein production

The cDNA encoding Calmodulin (UniProt P0DP23) was amplified by PCR using the primers 5'-AACTTCAAGGATCCGCTGCTGCTGCTGACCAGCTGACCGAAGAAC-3' and 5'-TGCTCGAGTGCGGCCGCTCATTATTTAGCGGTCATCATCTGAACGAACTC-3'. The PCR product was digested with BamHI and NotI and ligated into the BamHI / NotI sites of a modified pET vector containing eYFP. The eYFP-TEVsite-linker-CaM region of the resulting 6xHistag-PreScissionSite-eYFP-TEVsite-linker-calmodulin construct was amplified with the primers 5'-CGCGAAGAGGAAGAAGTGAGCAAGGGCGAGGAGCTGTTC-3' and 5'-CTCGAATTCGGATCCTCATTATTTAGCGGTCATCATCTGAACGAACTCTTCGTAG-3' and subcloned into a modified, PCR-linearized (21) pET vector, containing segments coding for an N-terminal calmodulin-binding site (3) and a rigid ER/K helix (14), by co-transformation cloning (22). Plasmid linearization was performed using primers 5'-TTCTTCCTCTTCGCGTTTGCGCTTTTCCTC-3' and 5'-GGATCCGAATTCGAGCTCCGTCGACAA-3'.

Protein expression and purification

The fusion protein was expressed in Rosetta2(DE3) *E. coli* cells (EMD Millipore Novagen) in 2x YT medium supplemented with 50 µg/ml kanamycin. The cells were grown at 37 °C, 130 rpm in a shaker-incubator until an OD_{600} of 0.75 was reached. Then, the temperature was reduced to 30 °C, $CaCl_2$ was added to a concentration of 0.2 mM and overexpression of the fusion protein was induced by the addition of IPTG to 1 mM. After 4 hours, the temperature was further reduced to 20 °C. The cells were harvested by centrifugation after overnight expression. The cell pellet was resuspended in 50 mM Tris, pH 7.8, 150 mM NaCl, supplemented with a tablet of complete EDTA-free protease inhibitor cocktail (Roche Diagnostics). The cells were lysed by ultrasonication. After incubation with lysozyme and DNase I at 4 °C for 30 minutes, the lysate was cleared by centrifugation at 16500 x g for 15 minutes at 4 °C. ß-mercaptoethanol was added to a final concentration of 5 mM to the cleared lysate. The protein was purified by Ni-NTA IMAC using Ni-Sepharose 6 Fast Flow resin (GE Healthcare). The wash buffer consisted of 50 mM Tris, pH 7.8, 150 mM NaCl and the elution buffer of 50 mM Tris, pH 7.8, 150 mM NaCl, 350 mM imidazole. To cleave off the 6xHis-tag, recombinant human rhinovirus 3C protease was added to the eluted protein, and the protein was dialyzed overnight against wash buffer. The cleaved fusion protein was further purified by gel filtration on a Superdex 75 column in 20 mM Tris, pH 7.8, 150 mM NaCl. $CaCl_2$ was added to the eluted protein to a concentration of 5 mM.

Crystallization and structure elucidation

For crystallization, the YFPnano fusion protein was concentrated to 20 mg/ml in a 10 kDa cutoff centrifugal filter device (Amicon). Initial crystals were obtained in the Qiagen PEGs Suite crystallization screen, well H2 (0.2 M di-Ammonium tartrate, 20% (w/v) PEG 3350). Optimized crystals were grown in 24% (w/v) PEG 3350, 0.2 M di-Ammonium tartrate, 10% glycerol in sitting drops at 20 °C.

A complete dataset to 1.9 Å resolution was collected from a single crystal at 100 K at the X06DA beamline of the Swiss Light Source synchrotron in Villigen, Switzerland, using a wavelength of 1.000 Å. XDS (23) was used for data processing. The high-resolution limit was chosen according to the cc1/2 (24), making sure that the I/Sigma in the highest resolution shell was not below 1. The space group of the crystal was P1 21 1. The structure of the fusion protein was solved by molecular replacement using a structure of YFP (pdb entry 3V3D (25)) as the search model. Two YFP molecules could be positioned in the asymmetric unit using Phaser (26). After the first refinement cycles, additional density for the extended helix and calmodulin became apparent.

A helix was placed into the density near the N-terminus of YFP chain A in Coot (27). In iterative rounds of refinement and model building, the helix was built, starting at the very characteristic electron density of tryptophan 4 and phenylalanine 17. Calmodulin was then placed by first building the most characteristic helix in Coot (27), followed by superposing the calmodulin structure from pdb

entry 2BBM (2) onto the CBP helix and the CaM helix in PyMOL (28). After three cycles of refinement, CaM regions that did not fit the electron density were deleted and then manually built in Coot (27) with iterative rounds of refinement and model building. The same strategy was then applied to build the extended helix and CaM in chain B. Refinement was performed using phenix.refine (29). The geometry and stereochemistry was validated using MolProbity (30). The final refined structure exhibits good geometry and stereochemistry. In chain B, the side chains of CaM amino acids 41-72 are poorly defined and amino acids 73 - 79 are undefined. The flexible linker connecting YFP and CaM is undefined in both chains. PyMOL (28) and Coot (27) were used for the preparation of structure figures. The structure was deposited in the PDB under the accession code: 6HR1.

Acknowledgments

We thank May Sharpe, Laura Vera, Meitian Wang and Akira Shinoda for support at the SLS beamline and at the crystallization facility, Martin Schärer, Tobias Weinert, Tim Grüne and Andrea Prota for important discussions about structural biology related aspects of the project, Takashi Ishikawa, Gebhard F.X. Schertler, Jan Pieter Abrahams and Dmitry Veprintsev for supporting the project and Hans Widmer and Sandra Jacob for interesting discussions about the project.
This work was supported by grants from Novartis FreeNovation and from UBS Promedica (1401/M) to R.M.B.

Author contributions

R.M.B. designed the research. A.S.K., R.M.B., G.C., T.B. and N.V. carried out the research. R.M.B., A.S.K., G.C. and T.B. analyzed the data. R.M.B. wrote the paper with important input from the other authors.

References

1 Sivaramakrishnan S, Spink BJ, Sim AY, Doniach S, & Spudich JA (2008) Dynamic charge interactions create surprising rigidity in the ER/K alpha-helical protein motif. *Proc Natl Acad Sci U S A* 105:13356-13361.

2 Ikura M*, et al.* (1992) Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* 256:632-638.

3 Montigiani S, Neri G, Neri P, & Neri D (1996) Alanine substitutions in calmodulin-binding peptides result in unexpected affinity enhancement. *J Mol Biol* 258:6-13.

4 Ormo M*, et al.* (1996) Crystal structure of the Aequorea victoria green fluorescent protein. *Science* 273:1392-1395.

5 Howorka S (2011) Rationally engineering natural protein assemblies in nanobiotechnology. *Curr Opin Biotechnol* 22:485-491.

6 Ludwig C & Wagner R (2007) Virus-like particles-universal molecular toolboxes. *Curr Opin Biotechnol* 18:537-545.

7 Shen PS (2018) The 2017 Nobel Prize in Chemistry: cryo-EM comes of age. *Anal Bioanal Chem* 410:2053-2057.

8 Nogales E & Scheres SH (2015) Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell* 58:677-689.

9 Coscia F*, et al.* (2016) Fusion to a homo-oligomeric scaffold allows cryo-EM analysis of a small protein. *Sci Rep* 6:30909.

10 Desmyter A, Spinelli S, Roussel A, & Cambillau C (2015) Camelid nanobodies: killing two birds with one stone. *Curr Opin Struct Biol* 32:1-8.

11 Sennhauser G & Grutter MG (2008) Chaperone-assisted crystallography with DARPins. *Structure* 16:1443-1453.

12 Swanson CJ & Sivaramakrishnan S (2014) Harnessing the unique structural properties of isolated alpha-helices. *J Biol Chem* 289:25460-25467.

13 Bryan J*, et al.* (1989) Cloning and expression of a smooth muscle caldesmon. *J Biol Chem* 264:13873-13879.

14 Ropars V*, et al.* (2016) The myosin X motor is optimized for movement on actin bundles. *Nat Commun* 7:12456.

15 Hoffman DW*, et al.* (1994) Crystal structure of prokaryotic ribosomal protein L9: a bi-lobed RNA-binding protein. *EMBO J* 13:205-212.

16 Liu D*, et al.* (2005) The N-terminal 26-residue fragment of human programmed cell death 5 protein can form a stable alpha-helix having unique electrostatic potential character. *Biochem J* 392:47-54.

17 Arai R, Ueda H, Kitayama A, Kamiya N, & Nagamune T (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng* 14:529-532.

18 Arai R, Wriggers W, Nishikawa Y, Nagamune T, & Fujisawa T (2004) Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. *Proteins* 57:829-838.

19 Ikura M & Bax A (1992) Isotope-filtered 2D NMR of a protein-peptide complex: Study of a skeletal muscle myosin light chain kinase fragment bound to calmodulin. *J Am Chem Soc* 114:2433-2440.

20 Miyawaki A*, et al.* (1997) Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature* 388:882-887.

21 Benoit RM, Wilhelm RN, Scherer-Becker D, & Ostermeier C (2006) An improved method for fast, robust, and seamless integration of DNA fragments into multiple plasmids. *Protein Expr Purif* 45:66-71.

22 Benoit RM*, et al.* (2016) Seamless Insert-Plasmid Assembly at High Efficiency and Low Cost. *PLoS One* 11:e0153158.

23 Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66:125-132.
24 Karplus PA & Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336:1030-1033.
25 De Meulenaere E*, et al.* (2013) Improving the second-order nonlinear optical response of fluorescent proteins: the symmetry argument. *J Am Chem Soc* 135:4061-4069.
26 McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.
27 Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
28 The PyMOL Molecular Graphics System, Schrödinger, LLC
29 Afonine PV*, et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68:352-367.
30 Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12-21.