The 8th Alpbach Workshop on: COILED-COIL, FIBROUS & REPEAT PROTEINS

Der Böglerhof – pure nature resort, Alpbach, Austria

Sunday, 4th September – Friday, 9th September 2022

8th Alpbach Workshop on: Coiled-coil, fibrous and repeat proteins

We gratefully acknowledge the generous sponsorship of

Generate : Biomedicines

A Flagship Pioneering Company

for this meeting. In particular we want to highlight the Travel Awards they made possible, which go to:

Fabio Pirro
Jakub Ptacek
Jessalyn Miller
Jessica Cross
Kateryna Maksymenko
Katherine Albanesse
Prasun Kumar
Sophia Tan
Adrian Dobbelstein

We thank Meytal Landau and Gustav Oberdorfer for selecting the award winners.

SUNDAY SEPTEMBER 4

Arrivals, Reception and Dinner

MONDAY SEPTEMBER 5

09:00 – 10:00 **New developments in coiled coils** (Andrei Lupas)

Andrei Lupas (Tübingen) – The structural diversity of natural coiled coils
Birte Hernandez Alvarez (Tübingen) – Towards a coiled coil with a polar core
Italo Cavini (São Carlos/Bristol) – X-ray structure of the metastable septin heterodimeric coiled coil reveals hendecad repeats, aromatic residues in the core and unusual *d-ag* layers

10:00 – 10:25 **Software resources for modelling and analysis** (Stanisław Dunin-Horkawicz)

Vikram Alva (Tübingen) – Protein Sequence Analysis Using the MPI Bioinformatics Toolkit

10:25 – 10:50 **Tea break**

10:50 – 12:30
Gevorg Grigoryan (Generate Biomedicines/Dartmouth) – Computational Protein Design: then and now
Stanisław Dunin-Horkawicz (Tübingen/Warsaw) – Coiled-coil domains and AlphaFold2 - friends or foes?
Prasun Kumar (Bristol) – Socket2 and CC+ 2022: Bioinformatics resources for assigning and analyzing coiled-coil structures and models
Mohammad ElGamacy (Tübingen) – Damietta: A Tensorised Protein Design Engine, and Its Application in Therapeutic Protein Design

12:30 - **Lunch**
Afternoon free

17:00 – 17:40 **Flash Presentations for Posters**

Adrian Dobbelstein (Tübingen) – Variety of domains in trimeric autotransporter adhesins
Bastien Canonica (Zürich) – Selection of a repeat protein recognizing mNeonGreen with picomolar affinity
Evangelia Notari (Edinburgh) – Computational design of multi-state coiled coils using Molecular Dynamics simulations
Frances Gidley (Bristol) – Specific Protein-Protein Interfaces: Towards multi-component modular protein systems
Horst Lechner (Graz) – Designing Proteins to Protect Fluorescent Dyes in Light-Emitting Diodes
Jessalyn Miller (Emory) – Diverse assembly strategies for nanotubes from αRep-derived proteins
Paulina Dubiel (Bristol) – Computational design of photoactive de novo flavoproteins

Prasun Kumar (Bristol) – Constructing synthetic-peptide assemblies from de novo designed 3_{10} helices
Sophia Tan (San Francisco) – Computational de novo design of phospho-responsive molecular switches
Tadej Satler (Ljubljana) – Protein binders for functionalization of proteins with coiled-coils
Wenzhao Dai (Zürich) – Protein Origami with Armadillo Repeat Fragments
Žiga Strmšek (Ljubljana) – Incorporation of folded protein domains into coiled-coil protein origami

17:40 – 18:30 **Posters**

18:30 – 20:00 **Dinner**

20:00 – 21:30 **Repeat-protein Design** (Fabio Parmeggiani)

Lynne Regan (Edinburgh) - TBA
Fabio Parmeggiani (Bristol) – Designed repeat proteins: longer, faster and more creative
Oliver Zerbe (Zürich) – Fast dynamics in complexes of picomolar binders
Agathe Urvoas (Paris) – AlphaReps artificial repeat proteins as tools for modular protein assemblies

TUESDAY SEPTEMBER 6

09:00 – 10:30 **Coiled-coil Design I: structure and dynamics** (Dek Woolfson)

Dek Woolfson (Bristol) – From rational to computational de novo design of coiled-coil assemblies and back again
Roman Jerala (Ljubljana) – Coiled-coil modules for designed protein folds and cellular logic circuits
Franziska Thomas (Heidelberg) – The WW domain as a model system for the design of miniaturized proteins
Elise Naudin and **Katherine Albanese** (Bristol) – From peptides to proteins: coiled-coil tetramers to single-chain 4-helix bundles

10:30 – 11:00 **Tea break**

11:00 – 12:30 **Coiled-coil Design II: function** (Anna Peacock)

Bill DeGrado (UCSF) – De novo design of water-soluble and membrane-soluble proteins inspired by coiled coils
Anna Peacock (Birmingham) – Coiled coils as ligands for inclusion in the inorganic chemist's toolbox
Fabio Pirro (Bristol) – Computational design of single-chain alpha-helical barrel proteins
Erich Michel (Zürich) – Improved repeat protein stability by computational protein design

12:30 - **Lunch**
Afternoon free

18:30 – 20:00 **Dinner**

20:00 – 21:20 **Thematically open session** (Andrei Lupas)

Russell Wilson (Linz) – Coiled Coil Molecular Force Sensors for Measuring Cellular Forces & Attachment
Timo Ullrich (Tübingen) – Computational design of granulopoietic proteins
Katya Maksymenko (Tübingen) – Protein design of growth factor inhibitors
Javier Montenegro (Santiago de Compostela) – The exo-helical symmetry of the α-Helix

WEDNESDAY SEPTEMBER 7

09:00 – 10:25 **Larger assemblies and materials** (Vince Conticello)

Vince Conticello (Emory) – Self-assembly of Peptide-based Nanomaterials: Structural Insights into the Self-Assembly of Helical Peptide Filaments and Tubes from CryoEM Analysis
Abhishek Jalan (Bayreuth) – Molecular Clamps Chaperone Collagen Folding
Giel Stalmans (Leuven) – The integrative structural study of nuclear lamin assembly
Katherine Albanese (Bristol) – *De novo* designed self-assembling modules that function at cell membranes

10:25 – 10:55 **Tea break**

10:55 – 12:30 **Bundles and barrels** (Gustav Oberdorfer)

Allon Hochbaum (Irvine) – TBA
Gustav Oberdorfer (Graz) – TBA
Birte Höcker (Bayreuth) – Navigating the structure and stability landscape of de novo $(\beta\alpha)_8$- barrels by protein design and engineering
Will Dawson (Bristol) – Mimicking mammalian olfaction with arrays of de novo designed peptide assemblies

12:30 - **Lunch**
Afternoon free

17:00 - 18:30 **Of α-fibers and β-fibers** (Meytal Landau)

Meytal Landau (Technion/Hamburg) – Virulent and Antimicrobial Amyloids in Infections and Neurodegeneration
Andrei Lupas (Tübingen) – Highly repetitive genomic ORFs: a source of new fibrous proteins?
Antoine Schramm (Tübingen) – Cryptic coiled coil-forming potential in ribosomal proteins
Mikel Martinez-Goikoetxea (Tübingen) – Here Be Coiled-Coil Dragons: Computational Search for New Hendecad Coiled Coils

18:30 – 20:00 **Dinner**

20:00 – 21:30 **Coiled coils at the membrane** (Alexander Kros and Ai Niitsu)

Ai Niitsu (RIKEN) – *De novo* design of membrane coiled-coil barrels
Alexander Kros (Leiden) – Coiled-coil peptide-mediated RNA delivery
Ben Hardy (Bristol) – Computational Design of a de novo Transmembrane Cytochrome
Wenyue Dai (Leeds) – Bacterial toxins assembled using coiled-coil peptides mediate membrane fusion

THURSDAY SEPTEMBER 8

09:00 – 10:40 **Natural coiled coils in signal transduction** (Andreas Möglich)

Heikki Takala (Jyväskylä) – Structure, function, and application of a model bacterial phytochrome
Andreas Möglich (Bayreuth) – Signal Transduction in Photoreceptor Histidine Kinases
Murray Coles (Tübingen) – Revisiting HAMP in the AlphaFold age
Jakub Ptáček (Tübingen) – Archaeal receptor Af1503 – an ideal model for structural and functional studies of transmembrane signaling

10:40 – 11:10 **Tea break and Group Photograph**

11:10 – 12:50 **Natural coiled coils in complex biological systems** (Mark Dodding)

Mark Dodding (Bristol) – Molecular architecture of the autoinhibited kinesin-1 lambda particle
Marcus Jahnel (Dresden) - EEA1 and Rab5 form a two-component molecular motor at the early endosome
Guto Rhys (Bayreuth/Cardiff) and **Jess Cross** (Bristol) – *De novo* designed peptides for cellular delivery and subcellular localisation
Owen Davies (Edinburgh) – Structural basis of meiotic chromosome synapsis by the synaptonemal complex

13:00 - **Lunch**
Afternoon Free

17:00 – 19:30 **The quest for specific binding proteins** (Dek Woolfson and Andrei Lupas; Andreas Plückthun)

Andrei Lupas (Tübingen) - Introduction
Andreas Plückthun (Zürich) - Binding proteins: Evolution of design and design of evolution
Sarel Fleishman (Weizmann) – An atomistic and machine-learning paradigm for designing large repertoires of functional variants
Brian Kuhlman (Chapel Hill) – Improving enzyme efficiency via designer incorporation of a substrate recruitment domain
Tanja Kortemme (UCSF) – Computational design of new protein shapes and functions

20:00 – **Gourmet Dinner**

FRIDAY SEPTEMBER 9

Departures

De novo designed self-assembling modules that function at cell membranes

Katherine I. Albanese[1,2*], Elise A. Naudin[1], Bram Mylemans[1,2], Kate W. Kurgan[1], Muziyue Wu[3], Orion D. Weiner[3], Derek N. Woolfson[1,2,4,5*]

[1]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[2]Max Planck-Bristol Centre for Minimal Biology, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[3]Cardiovascular Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, 555 Mission Bay Blvd. South, San Francisco, CA 94158, USA

[4]School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK

[5]BrisEngBio, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

*To whom correspondence should be addressed: katherine.albanese@bristol.ac.uk, D.N.Woolfson@bristol.ac.uk

Abstract: Protein modules that are mutable, orthogonal, portable, and dynamic are essential tools for synthetic biology. For the work described herein, we aimed to mimic nature's means of sensing and accessing membrane curvatures by constructing fully synthetic systems that self-assemble into specific geometries at cell membranes. In doing so, we aim to help understand mechanisms of cell morphogenesis that are otherwise difficult to characterize *in vivo*.

While advances in computational protein-structure prediction and design are generating many functional *de novo* peptides and proteins,[1-3] the resulting polypeptides are often hyper-stable and not well suited as mimics of natural dynamic assemblies. Therefore, we have taken a rational *peptide-to-protein* design approach starting from α-helical coiled coils to generate single-chain protein modules with desired properties. Coiled coils mediate protein interactions in many biological processes, have relatively straightforward sequences, tuneable interaction strengths, directed specificity, and can be designed *de novo* to generate components that are orthogonal to those in nature. Here, we define clear sequence-to-structure relationships to deliver specific, robust, coiled-coil tetramers from which a single-chain protein can be derived and used in cells.[4]

Our single-chain protein module is soluble, monomeric, and stable in solution. We have determined the X-ray crystal structure for the module, which overlays with the coiled-coil tetramer parent structure and the AlphaFold2 prediction. The successful implementation of this *peptide-to-protein* design approach demonstrates the utility of well-understood design rules in underpinning the robust design of full-length *de novo* proteins. Moreover, in cells, when appended to an *E. coli* membrane-targeting sequence, the full protein construct is well distributed across the membrane, and it can alter the shape and size of the cells. Next, this pipeline will be applied to design higher-order coiled-coil and protein oligomers that can then be incorporated in heteromeric 2D assemblies to remodel membranes *in vitro* and in living cells.

1. Korendovych, I. V.; DeGrado, W. F., De novo protein design, a retrospective. *Quarterly Reviews of Biophysics* **2020,** *53*, e3.
2. Pan, X.; Kortemme, T., Recent advances in de novo protein design: Principles, methods, and applications. *Journal of Biological Chemistry* **2021,** *296*.
3. Woolfson, D. N., A Brief History of De Novo Protein Design: Minimal, Rational, and Computational. *Journal of Molecular Biology* **2021,** *433* (20), 167160.
4. Naudin, E. A., Albanese, K. I., Smith, A. J., Mylemans, B., Baker, E. G., Weiner, O. D., Andrews, D. M., Tigue, N., Savery, N. J., Woolfson, D. N., From peptides to proteins: 1. coiled-coil tetramers to single-chain 4-helix bundles. *Manuscript submitted.*

Protein Sequence Analysis Using the MPI Bioinformatics Toolkit

Vikram Alva
Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Germany

Abstract
The MPI Bioinformatics Toolkit (https://toolkit.tuebingen.mpg.de [1,2]) provides interactive access to a wide array of the best-performing bioinformatics tools and databases, including the state-of-the-art protein sequence comparison methods HHblits and HHpred. In this talk, I will discuss the application of the Toolkit for detecting remote homologs, inferring protein function, and visualizing relationships in large sequence datasets. I will demonstrate these using, as an example, the coiled-coil segment-containing OmpM protein superfamily, whose members tether the outer membrane to the peptidoglycan layer in many Gram-negative bacteria [3].

References

[1] Zimmermann L, Stephens A, Nam SZ, et al. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol. 2018.*

[2] Gabler F, Nam SZ, Till S, et al. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr Protoc Bioinformatics. 2020.*

[3] von Kügelgen A, van Dorst S, Alva V, & Bharat TAM (2022) A multi-domain connector links the outer membrane and cell wall in deep-branching bacteria. *Proc Natl Acad Sci U S A. 2022.*

Selection of a repeat protein recognizing mNeonGreen with picomolar affinity

Bastien B. Canonica, Erich Michel, Andreas Plückthun

Department of Biochemistry, University of Zurich, Zurich, Switzerland

b.canonica@bioc.uzh.ch

The well-known limitations of antibodies have led researchers to design existing proteins to bind specific targets and epitopes in a controlled manner. From this quest, multiple protein scaffolds have emerged, such as the designed Armadillo repeat protein (dArmRP). In our lab, we work with the dArmRP scaffold to acquire specific and high affinity binders to any desired target by exploiting its modular property to bind unstructured peptides. In selections for linear epitopes we normally prevent, by proper counterselections, the binding to folded proteins, and found that they do not cross-react.

Here, we did allow selection for a folded protein and thus show that the scaffold can also bind a fluorescent protein, exploiting exposed Trp residues on the randomized surface of a particular dArmRP member.

Fluorescent fusion proteins are widely used for cell localization, protein immobilization, folding assessment, binding energy determination and guiding directed evolution. Consequently, we have observed over the past years a tremendous increase in the development and engineering of new fluorescent proteins in an attempt to cover the entire visible light spectrum. Specific binders able to recognize these extensively used proteins will reveal themselves essential to the scientific community.

mNeonGreen is one of these emerging fluorophores displaying great biophysical properties [1]. In our group, we have selected an dArmRP targeting mNeonGreen with high affinity. We have observed a dependence of the binding affinity on the number of modules present in the dArmRP and we could obtain a well-defined crystal structure helping us characterizing key features necessary to the binding interactions. The development of this binder opens a wide range of possibilities for further applications involving mNeonGreen.

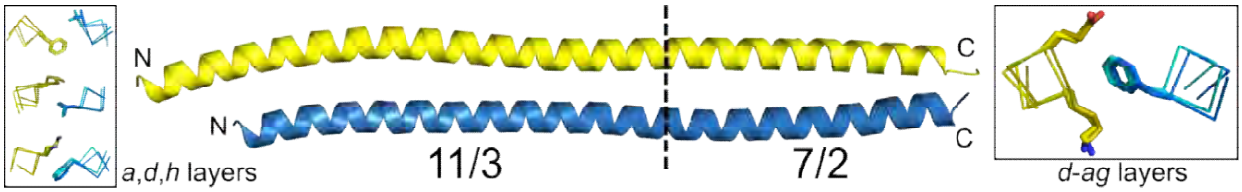
1. Shaner, N., Lambert, G., Chammas, A. *et al.* A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat Methods* **10,** 407–409 (2013). https://doi.org/10.1038/nmeth.2413

X-ray structure of the metastable septin heterodimeric coiled coil reveals hendecad repeats, aromatic residues in the core and unusual *d-ag* layers

Italo A. Cavini,[a,b] Humberto Pereira,[a] Richard C. Garratt[a]

[a]*São Carlos Institute of Physics, University of São Paulo, São Carlos, Brazil*
[b]*School of Chemistry, University of Bristol, Bristol, United Kingdom*

Septins are membrane-associated, GTP-binding proteins present in most eukaryotes that, as part of the cytoskeleton, play important roles as scaffolds and/or diffusion barriers.[1] They polymerize into long filaments composed of up to four different septin subunits, which are held together by two alternated interfaces: the G-interface, formed by the GTP-binding domain, and the NC-interface, composed of the two N- and C-terminal domains, the latter usually having coiled coil (CC) prediction. In heterodimeric NC-interfaces, the formation of a putative parallel heterodimeric CC, although not essential, seems to help the interaction and/or discriminate the correct septin partners.[2] However, none of the septin structures released so far depicts the C-terminal domains due to their intrinsic flexibility. In a recent work, we solved the homodimeric CC structures of the human paralogs SEPT6 and SEPT8.[3] Both are antiparallel CCs and could explain the interfilament cross-bridges seen in different septin microscopy studies. Here, we present the first structure of a heterodimeric septin coiled coil, the SEPT14-SEPT7 human heterodimer solved by X-ray diffraction at 1.8 Å resolution. SEPT14 is a paralog of SEPT6 and SEPT8, all believed to occupy the same position within filaments. SEPT14-SEPT7 CC has two long parallel helices as anticipated. Several aromatic residues (four from each chain, mainly phenylalanines) populate the core, but they do not pair against each other; instead, they pair with other hydrophobic residues and interact with them through weak CH-π contacts. Surprisingly, the N-terminal half of the CC is unusually right-handed due to 11/3 repeats, with three types of core positions (*a*,*d*,*h*). This right-handed twist is annulled by the C-terminal half which is made of 7/2 repeats. Interestingly, the C-terminal part of the structure possesses uncommon *d-ag* layers in which the hydrophobic at *d* is almost entirely buried between the hydrophilics *a* and *g* of the other helix. In the entire structure, *a* positions are occupied mostly by hydrophilic residues, leaving *d* and *h* as the genuine hydrophobic core positions. All this taken together emphasizes the metastability built in both sequences of this coiled coil, allowing them to transition between the parallel heterodimer and the antiparallel homodimer states.



References

[1] Mostowy S & Cossart P (2012) Septins: the fourth component of the cytoskeleton. *Nature Reviews Molecular Cell Biology*, 13(3), 183-194.
[2] Cavini IA et al. (2021) The structural biology of septins and their filaments: An update. *Frontiers in Cell and Developmental Biology*, 3246.
[3] Leonardo DA†, Cavini IA†, et al. (2021) Orientational ambiguity in septin coiled coils and its structural basis. *Journal of Molecular Biology* 433(9), 166889.

Revisiting HAMP in the AlphaFold age

Murray Coles, Stanisław Dunin-Horkawicz, Marcus D. Hartmann, Andrei N. Lupas
Department of Protein Evolution, MPI for Biology, 72076 Tübingen, Germany;
andrei.lupas@tuebingen.mpg.de

HAMP domains are small, dimeric, four-helical bundles found exclusively in a class of signalling proteins in prokaryotes, unicellular eukaryotes, fungi, and plants. For over a decade - and several Alpbach meetings - we have been pursuing a model of HAMP-mediated signal transduction via transitions in coiled-coil packing modes. But how would this project have played out in the alphaFold era? Do alphaFold models have the level of detail needed to resolve such mechanistic issues? Here I review the HAMP project under the alphaFold paradigm and in particular the case of polyHAMP arrays, tracts of up to 57 contiguous HAMP domains that occur in chemotaxis and histidine kinase signalling proteins. Experimental structures have characterised polyHAMP arrays as tightly organised, rod-shaped proteins with an alternating pattern of two distinct HAMP forms along the array. These forms are distinguished by coiled-coil packing modes, including variants of complementary x-da packing with opposite rotation states to those previously observed in canonical HAMPs. Modelling with AlphaFold2 suggests that this alternating pattern is the hallmark of poly-HAMP arrays, supporting our model of signalling via axial helix rotation.

Title: Self-assembly of Peptide-based Nanomaterials: Structural Insights into the Self-Assembly of Helical Peptide Filaments and Tubes from CryoEM Analysis.

Ordy Gnewou,[1] Fengbin Wang,[2] Chunfu Xu,[1] Charles Modlin,[1] Gevorg Grigoryan,[3] Edward Egelman,[2] and Vincent Conticello[1]

[1]Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, GA, USA 30322
[2]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22908
[3]Department of Computer Science, Dartmouth College, Hanover, NH 03755

Abstract: Structurally defined materials on the nanometer length-scale have been historically the most challenging to rationally construct and the most difficult to structurally analyze. Sequence-specific biomolecules, i.e., peptides and nucleic acids, have advantages as design elements for construction of these types of nano-scale materials in that correlations can be drawn between sequence and higher order structure, potentially affording ordered assemblies in which functional properties can be controlled through the progression of structural hierarchy encoded at the molecular level. However, the predictable design of self-assembled structures requires precise structural control of the interfaces between peptide subunits (protomers). In contrast to the robustness of protein tertiary structure, quaternary structure has been postulated to be labile with respect to mutagenesis of residues located at the protein-protein interface. In addition, self-assembling peptide systems have been shown to be exquisitely sensitive to preparative conditions. We have employed self-assembling peptide systems based on α-helical coil-coil and β-strand structural motifs to interrogate the concept of designability of interfaces within the structural context of filaments and nanotubes. These peptide systems provide a framework for understanding how minor sequence changes in evolution or preparative conditions can translate into very large changes in supramolecular structure, which provides significant evidence that the designability of protein interfaces is a critical consideration for control of supramolecular structure in self-assembling systems.

1. Egelman, E.; Xu, C.; DiMaio, F.; Magnotti, E.; Modlin, C.; Yu, X.; Wright, E.R.; Baker, D.; Conticello, V.P. Structural plasticity of helical nanotubes based on coiled-coil assemblies, *Structure* **2015**, *23*, 280-9.
2. Wang, F.; Gnewou, O.; Modlin, C.; Xu, C.; Juneja, P.; Grigoryan, G.; Egelman, E.H.; Conticello, V.P. Structural analysis of cross α-helical nanotubes provides insight into the designability of filamentous peptide nanomaterials. *Nat. Commun.* **2021**, *12*, 407.
3. Wang, F.; Gnewou, O.; Wang, S.; Osinski, T.; Zuo, X.; Egelman, E.H.; Conticello, V.P. Deterministic chaos in the self-assembly of β-sheet nanotubes from an amphipathic oligopeptide. *Matter (Cell Press)* **2021**, *4*, 3217-31.
4. Wang, F.; Gnewou, O.; Solemanifar, A.; Conticello, V.P.; Egelman, E.H. Cryo-EM of helical polymers. *Chem. Rev.* **2022**, https://doi.org/10.1021/acs.chemrev.1c00753.
5. Rogers, J.; Hughes, S.A.; Modlin, C.; Conticello, V.P. Synthetic helical Filaments and Tubes based on peptide and peptido-mimetic polymers. *Quarterly Reviews of Biophysics* (invited review), **2022**, 1-103, doi: 10.1017/S0033583522000014.

Bacterial toxins assembled using coiled-coil peptides mediate membrane fusion

Wenyue Dai [a] Erik Kempmann,[b] Francesca Rosato,[b] George Heath,[a] Maria Nikolova,[a] Paul Beales, [a] Winfried Römer,[b] Ralf Richter, [a] Michael Webb,[a] Bruce Turnbull,[a]*

[a] School of Chemistry and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK. Email: w.b.turnbull@leeds.ac.uk
[b] Faculty of Biology, Albert Ludwigs University of Freiburg, Schänzlestraße 1, 79104 Freiburg, Germany. Email: winfried.roemer@bioss.uni-freiburg.de

Cholera toxin B-subunit (CTB_5) binds to GM1 glycosphingolipids at the surface of cells or artificial synthetic membrane such as giant unilamellar vesicles (GUV), causing the glycolipids to cluster and induce membrane curvature as a first step to achieve endocytic internalization[1]. However, as a single CTB_5 has multiple carbohydrate-binding sites facing towards the same side, it cannot lead to membrane crosslinking unlike other lectins, e.g., LecA which has carbohydrate binding sites that face in opposing directions[2]. Recent research has shown that multimeric complexes of biotinylated CTB_5 and streptavidin can mediate GUV crosslinking and membrane fusion processes [3].

To further understand how the orientation and number of carbohydrate-binding sites in CTB_5 assemblies achieve membrane fusion, we developed a new way of assembling CTB_5 toxins through either antiparallel or parallel coiled-coils as rigid linkers. Genes for a series of parallel and anti-parallel coiled-coil constructs were designed and expressed, and the purified protein products were characterised by SDS-PAGE and mass spectrometry for purity and molecular weight. Their oligomerization state was verified through size exclusion chromatography and mass photometry, and the topology of the protein assembly in solution and membrane is being studied by ISAMBARD, molecular simulation, atom force microscopy and QCMD seperately.

GUV confocal microscopy experiments have been used to study membrane-perturbing behaviours of the protein constructs. Heavy crosslinking was immediately observed when at least one of the protein complex was added, followed by an elongated interface between two GUVs and fusion events (Figure 1).

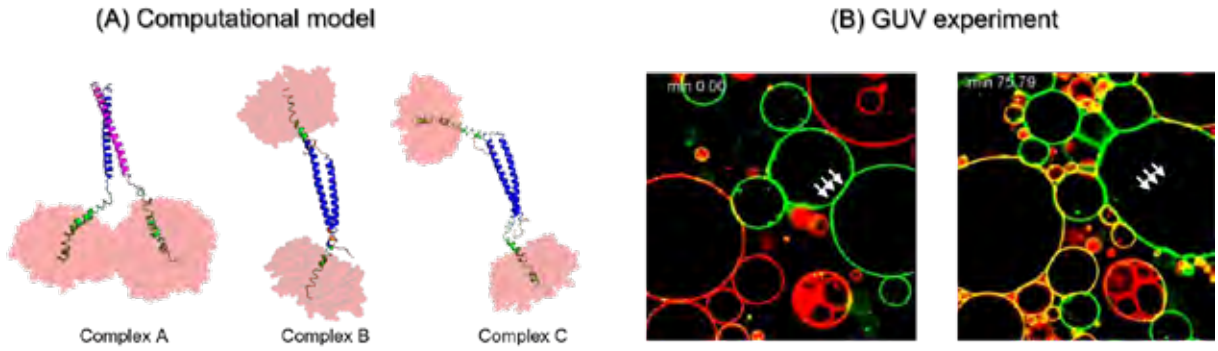


Figure 1. (a) Computational model of different coiled coil linked-CTB_5 assembly used in membrane fusion experiment. (b) Crosslinking of multiple GUVs (5 mol% GM1 with either red Atto 647 or no membrane dye) form elongated interfaces, after the addition of Alexa fluor 488 labelled protein complexes (green). Fusion events were captured by confocal microscopy.

Acknowledgement:

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement numbers 814029 (synBIOcarb ITN). COST Action CA18132 GlycoNanoProbes is acknowledged.

References:

1. Romer, W., et al., *Actin dynamics drive membrane reorganization and scission in clathrin-independent endocytosis.* Cell, 2010. **140**(4): p. 540-53.
2. Villringer, S., et al., *Lectin-mediated protocell crosslinking to mimic cell-cell junctions and adhesion.* Sci Rep, 2018. **8**(1): p. 1932.
3. Wehrum, S. et al., Membrane fusion mediated by non-covalent binding of re-engineered cholera toxin assemblies to glycolipids. ChemRXiv, 2021. https://doi.org/10.26434/chemrxiv-2021-j9stg

Protein Origami with Armadillo Repeat Fragments

Wenzhao Dai
Department of Chemistry
Universität Zürich, Switzerland
wenzhao.dai@chem.uzh.ch

Building macromolecules of complex shapes has become an important branch of synthetic biology, and an intellectual challenge for the involved scientists. The purpose of this study is to create a protein assembly that allows placing molecules in defined geometric relations relative to each other, to investigate effects that depend on proximity.

Combining biochemical methods and computational tools like pyRosetta, this study aims to design a protein complex built from rather extended armadillo repeat proteins (dArmRP) and ankyrin repeat proteins (DARPin) that can rigidly assemble into a one-dimensional array. The DARPins themselves bind to GFPs thereby forming a protein complex assembled from three different proteins, in which GFP is equally spaced to itself. So far, dArmRPs of more than 100kDa are built, which are long enough to bind two or more DARPin-ligand fusions with high affinity. Rigidly connected DARPins fused to the dArmRP-binding ligands are used to load the cargo with adjustable geometry. The availability of DARPins against a multitude of cargos ensures that many different sytems can be studied with this system.

During this study I designed a cloning method that can create large repeat proteins by stepwise cloning. Optimization is underway presently to redesign the protein assembly so that every available DARPin does bind a cargo without decreased affinity.

Mimicking mammalian olfaction with arrays of *de novo* designed peptide assemblies

William M. Dawson[1,†], Kathryn L. Shelley[1,†], Jordan M. Fletcher[1,2], D. Arne Scott[1,2], Lucia Lombardi[1], Guto G. Rhys[1,3], Tania J. LaGambina[2], Ulrike Obst[2], Antony J. Burton[1,4], Jessica A. Cross[1], George Davies[1], Freddie J.O. Martin[1], Francis J. Wiseman[1], R. Leo Brady[1], David Tew[5], Christopher W. Wood[6] and Derek N. Woolfson[1]

[1]University of Bristol, [2]Rosa Biotech, [3]University of Bayreuth, [4]AstraZeneca, [5]GlaxoSmithKline, [6]University of Edinburgh

Mammalian olfaction—the sense of smell—discriminates between many odorant molecules. It achieves this using 300 – 2000 G-protein coupled receptors (GPCRs). Rather than making specific receptor-odorant interactions, each receptor responds to a variety of molecules. The composite response is interpreted by the brain as a smell. Differential sensing attempts to mimic this. However, GPCRs are membrane-spanning proteins, making them difficult to manipulate. Indeed, attempts to use them in sensing have met with limited success.

We have developed an array of computationally designed *de novo* coiled coils as synthetic receptors for sensing (Fig. 1). Specifically, we use self-assembling α-helical barrels (αHBs) with central channels that can be altered predictably to vary their sizes, shapes and chemistries.[1] The channels accommodate environment-sensitive dyes that fluoresce upon binding.[2] Challenging arrays of dye-loaded barrels with analytes causes differential dye displacement. The resulting colorimetric fingerprints are used to train machine-learning models that relate the patterns to the analytes. We show that this system discriminates between a range of biomolecules, biomarkers, and diagnostically relevant biological samples. As αHBs are robust and chemically diverse, the system has potential to sense many analytes in various settings.

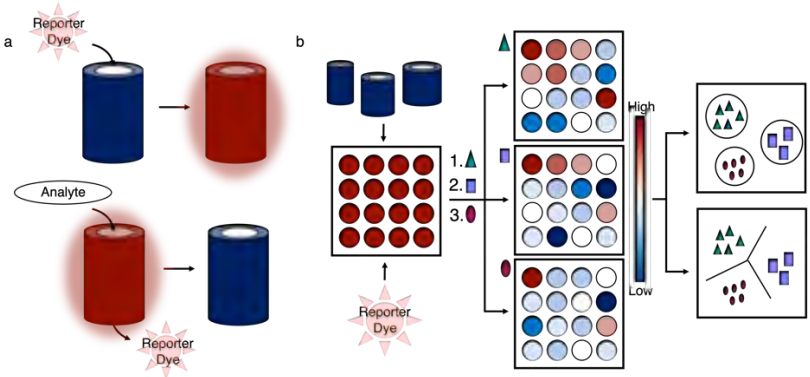


Figure 1. Concept for the *de novo* designed α-helical-barrel differential sensor. a, Top: α-Helical barrels (αHB) are loaded with an environment-sensitive dye giving a fluorescent signal. **Bottom:** The dye is displaced by an analyte causing a loss of fluorescence that can be measured. **b, Left**: Different αHBs are combined with the environment sensitive dye in multi-well plates. **Middle**: The resulting array is challenged with different analytes, which can be pure compounds or complex mixtures. Depending on the relative binding strengths of the dye and the analytes for each αHB, dye is displaced differentially across the array to give a 'fingerprint' for each analyte. **Right**: Statistical and machine-learning methods are used to classify the different fingerprints and relate them to the analytes. The resulting models can be used as predictive classifiers for naïve samples.

1. Thomson *et al.*, 2014, Science, 346, 485; Rhys *et al.*, 2018, Nat Commun, 9, 4132; Dawson *et al.*, 2021, Chem Sci, 12, 6923. 2. Thomas *et al.*, 2018, ACS Synth Biol, 7, 1808

Variety of domains in trimeric autotransporter adhesins

Adrian Dobbelstein, Andrei N. Lupas, Vikram Alva
Department of Protein Evolution, MPI for Biology, 72076 Tübingen, Germany;
adrian.dobbelstein@tuebingen.mpg.de

Trimeric autotransporter adhesins (TAAs) are a family of highly repetitive, homotrimeric surface proteins found in many Gram-negative bacteria. In pathogenic bacteria, TAAs play multiple roles during host infection, like adhesion, biofilm formation, and immune evasion. Structurally, TAAs exhibit a modular architecture, sharing a common anchor domain and possessing a variable number of stalk and head domains. The modularity of TAAs has been used for the bioinformatic and experimental characterization of their constituent domains and full-length homotrimeric complexes. However, many TAAs remain poorly annotated because of their extreme sequence diversity and the challenges underlying their experimental characterization.

To produce a comprehensive annotation of TAAs, we identified all TAAs within the RefSeq protein sequence database and predicted the homotrimeric structures of ~9000 representative TAAs using AlphaFold-multimer. As TAAs vary from under 200 amino acids to over 3000, we used their modular architecture for refined predictions based on sequence and structure analysis. We identified many uncharacterized stalk motifs and domains, variations of previously known stalk domains, and many structurally diverse head domains.

Molecular architecture of the autoinhibited kinesin-1 lambda particle

Johannes F. Weijman[1], Sathish K.N. Yadav[1], Katherine J. Surridge[1], Jessica A. Cross[1,2], Ufuk Borucu[3], Judith Mantell[1], Derek N. Woolfson[1,2,4], Christiane Schaffitzel[1,4], and **Mark P. Dodding[1,4]**

[1]School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, BS8 1TD, UK
[2]School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK
[3]GW4 Facility for High-Resolution Electron Cryo-Microscopy, University of Bristol
[4]Bristol BioDesign Institute, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol, BS8 1TQ, UK

Abstract: Despite continuing progress in kinesin enzyme mechanochemistry and emerging understanding of the cargo recognition machinery, it is not known how these functions are coupled and controlled by the alpha-helical coiled coils encoded by a large component of kinesin protein sequences. Here, we combine computational structure prediction with single-particle negative stain electron microscopy to reveal the coiled-coil architecture of heterotetrameric kinesin-1, in its compact state. An unusual flexion in the scaffold enables folding of the complex, bringing the kinesin heavy chain-light chain interface into close apposition with a tetrameric assembly formed from the region of the molecule previously assumed to be the folding hinge. This framework for autoinhibition is required to uncover how engagement of cargo and other regulatory factors drive kinesin-1 activation.

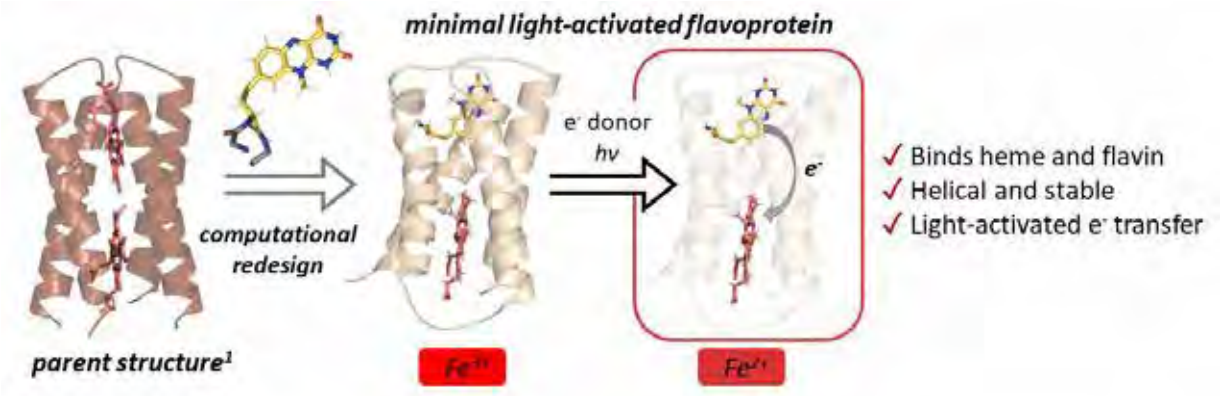
Computational design of photoactive *de novo* flavoproteins

Paulina Dubiel[1]; Marc van der Kamp[1], J. L. Ross Anderson[1]
[1]*Department of Biochemistry, University of Bristol, BS8 1TD*

Corresponding author: paulina.dubiel@bristol.ac.uk

Abstract

Electron transfer between redox centres is at the core of biochemical processes essential for life, such as respiration and photosynthesis. However, studying these processes can be challenging due to complex nature of bioenergetic proteins, in which functionality is often obscured by the protein's evolutionary legacy. This can be addressed by designing proteins *de novo*, with redox active cofactors, such as hemes or flavins, placed within a simple protein scaffold. Flavins are of particular interest as they are capable of one and two electron transfer reactions, and can be photoactivated, enabling light-induced electron transfer and catalysis. Here, we engineer a flavin-binding site into a *de novo* four helix bundle, using computational methods to design a binding site with covalently anchored flavin. We demonstrate that this approach represents a successful strategy for incorporating flavins into several *de novo* heme-containing proteins. Using this strategy, we obtain two proteins which both successfully bind heme and covalently anchored riboflavin in the desired stoichiometry, whilst maintaining highly helical, coiled-coil fold and high thermal stability. We demonstrate these proteins' ability to perform light-activated heme reduction in the order of minutes, as well as light-induced electron transfer to a *de novo* designed heme-containing membrane protein. Furthermore, we investigate how natural flavin transferases can be exploited to covalently attach natural flavins to *de novo* protein scaffolds, with our preliminary results indicating that this is a viable strategy. The work described here creates a foundation for incorporating covalently bound flavins into designed coiled-coil scaffolds, enabling creation of minimal light-activated redox proteins. Future efforts will involve utilising flavin transferases to attach flavins *in vivo*, as well as using the described strategies to create a photoactivated *de novo* catalyst.



References

[1] Hutchins, G. H. et al. Precision design of single and multi-heme de novo proteins. *bioRxiv* 2020.09.24.311514 (2020).

Structural basis of meiotic chromosome synapsis by the synaptonemal complex

James M. Dunce, Orla M. Dunne, Lucy J. Salmon and Owen R. Davies

Wellcome Centre for Cell Biology, Institute of Cell Biology, University of Edinburgh, UK.

The synaptonemal complex (SC) is a supramolecular protein assembly that mediates synapsis between homologous chromosomes during meiosis. The SC is formed of three self-assembling coiled-coil systems. Firstly, an SYCP3 coiled-coil tethers together chromatin loops and forms a paracrystalline lattice that compacts the chromosome axis (Syrjanen, Pellegrini, & Davies, 2014). Secondly, an SYCP1-SYCE3 coiled-coil complex assembles into a lattice-like array that acts as the molecular zipper that binds together homologous chromosomes (Crichton et al., 2022; Dunce et al., 2018). Finally, SYCE2-TEX12 is a coiled-coil structure that undergoes hierarchical assembly from a 2:2 complex, to 4:4 complex and then into 10-nm fibres that intertwine into 40-nm bundles, providing a fibrous backbone that enables SC elongation to lengths of up to 24-um (Dunce, Salmon, & Davies, 2021). This assembly mechanism bears striking resemblance with intermediate filament proteins vimentin, lamin and keratin. The discovery of these mechanisms though integrated structural biology approaches, and their implications for the wider coiled-coil field will be discussed.

Crichton, J. H., Dunce, J. M., Dunne, O. M., Salmon, L. J., Devenney, P. S., Lawson, J., . . . Davies, O. R. (2022). Structural maturation of SYCP1-mediated meiotic chromosome synapsis through conformational remodelling by molecular adapter SYCE3. *bioRxiv*, 2022.2003.2006.483192. doi:10.1101/2022.03.06.483192

Dunce, J. M., Dunne, O. M., Ratcliff, M., Millan, C., Madgwick, S., Uson, I., & Davies, O. R. (2018). Structural basis of meiotic chromosome synapsis through SYCP1 self-assembly. *Nat Struct Mol Biol, 25*(7), 557-569. doi:10.1038/s41594-018-0078-9

Dunce, J. M., Salmon, L. J., & Davies, O. R. (2021). Structural basis of meiotic chromosome synaptic elongation through hierarchical fibrous assembly of SYCE2-TEX12. *Nat Struct Mol Biol, 28*(8), 681-693. doi:10.1038/s41594-021-00636-z

Syrjanen, J. L., Pellegrini, L., & Davies, O. R. (2014). A molecular model for the role of SYCP3 in meiotic chromosome organisation. *Elife, 3*. doi:10.7554/eLife.02963

Coiled-coil domains and AlphaFold2 – friends or foes?

Rafał Madaj[1], Kamil Kamiński[1], Aleksander Wiński[1], Jan Ludwiczak[1], and Stanisław Dunin-Horkawicz[1,2]

[1] Institute of Evolutionary Biology, University of Warsaw, Warsaw, Poland
[2] Department of Protein Evolution, Max Planck Institute for Biology, Tübingen, Germany

Deep learning-based tools revolutionized the field of protein structure prediction. We present our recent attempts to use one of such methods, AlphaFold2 (AF2), for the modeling of coiled-coil domains.

We benchmarked AF2 using a set of experimental coiled-coil structures. For each benchmark case, we applied the following procedure: first, we determined local structural parameters **[1]** for the experimental structure and the corresponding AF2 model. Then, by comparing these structural parameters, we defined a quality index that reflects the discrepancy between the experimental structure and the model.

Knowing the advantages and limitations of AF2, we performed more focused analyses of **(i)** HAMP domains **[2]**, a family of coiled coils whose conformational changes are key for prokaryotic signal-transducing proteins, and **(ii)** the applicability of AF2 models for the augmentation of data sets used for training of machine learning methods.

[1] Szczepaniak K, Bukala A, da Silva Neto AM, Ludwiczak J, Dunin-Horkawicz S. A library of coiled-coil domains: from regular bundles to peculiar twists. *Bioinformatics*. 2020;36(22-23):5368-5376.

[2] Dunin-Horkawicz S, Lupas AN. Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction. *J Mol Biol*. 2010;397(5):1156-1174.

Damietta: A Tensorised Protein Design Engine, and Its Application in Therapeutic Protein Design

Kateryna Maksymenko[1], Birte Hernandez Alvarez[1], Timo Ullrich[1], Julia Skokowa[2], Andrei Lupas[1], Mohammad ElGamacy[1,2,†]

[1]Max Planck Institute for Developmental Biology, 72076 Tübingen (Germany)
[2]University Hospital Tübingen, Division of Translational Oncology, 72076 Tübingen (Germany)

Despite the groundbreaking successes of computational design in recent years, simultaneous improvements of design throughput and accuracy are continually needed to achieve better experimental success rates, and tackle more difficult design problems. We describe Damietta[1]; a novel protein design framework that maximises computational efficiency by tensorising energy calculations, and improves accuracy by relying on a self-consistent scoring function. This scoring function is not trained or contaminated by any learnt parameters, but relies purely on physics-based force field. We deploy these design concepts to tackle three design problems with different levels of difficulty, yielding agonists and antagonists of growth factor signaling pathways with therapeutic potential. First, we use Damietta to design epidermal growth factor (EGF) inhibitors based on an EGF receptor template structure. Testing only two designs, they were capable of binding EGF and inhibiting its signaling in cells. Second, we also use Damietta to create stabilised variants of metal binding proteins, leading to greatly improved high-density metal binders that we are currently further developing for radiotracing and immunoradiotherapy applications. Third, we design a bispecific, single-domain cytokine, capable of engaging two different cytokine receptors (here, we start by a IL3-Rα/G-CSFR combination). Such a "*novokine*" possesses a novel fold, and can serve as a non-natural cytokine with novel function. These applications exemplify the design of proteins with therapeutic potential, and demonstrate Damietta to be applicable for a range of protein design and engineering problems.

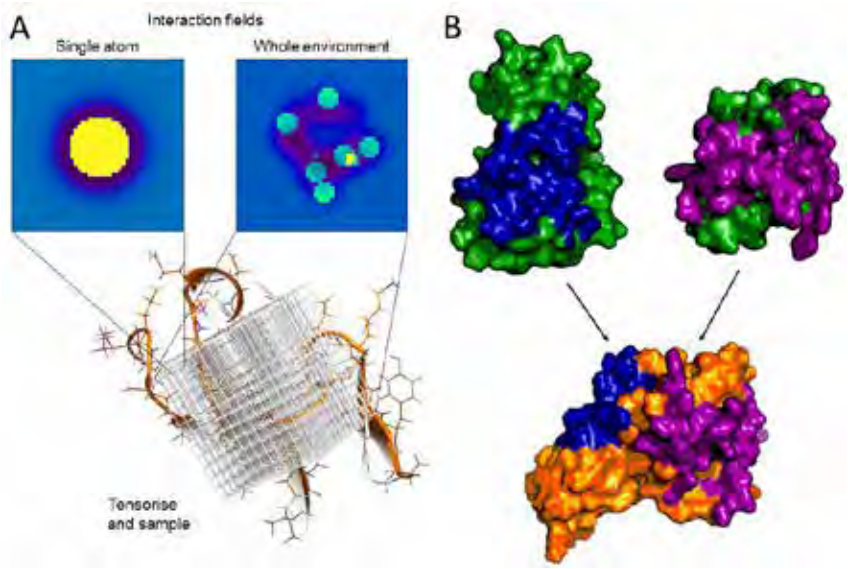


Figure 1. The concept and applications of the Damietta design framework. (A) A tensorised representation of the non-bonded interaction fields can accelerate protein design calculations through i) *single instruction, multiple data* processing paradigm, ii) precomputing of rotamer fields, and iii) enabling parallel implementation. (B) An example application presented here is the design of a *novokine*; a novel cytokine that bind and co-localise a non-native combination of cytokine receptor subunits.

[1] A pre-release of the Damietta software is available at: https://bio.mpg.de/damietta/

An atomistic and machine-learning paradigm for designing large repertoires of functional variants

Sarel J. Fleishman
Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel; sarel@weizmann.ac.il

We recently developed methods that combine phylogenetic analysis and Rosetta atomistic design calculations to design highly optimized variants of natural proteins[1,2]. Our methods have been used by thousands of users worldwide to generate stable therapeutic enzymes[3,4], vaccine immunogens[5,6], and highly active enzymes for a range of needs in basic and applied research[7–11]. We now present a machine-learning strategy to design and economically synthesize millions of active-site variant that are likely to be stable, foldable and active. We applied this approach to the chromophore-binding pocket of GFP to generate more than 16,000 active designs that comprise as many as eight mutations in the active site. The designs exhibit extensive and potentially useful changes in every experimentally measured parameter, including brightness, stability and pH sensitivity. We also applied this strategy to design millions of glycoside hydrolases that exhibit significant backbone changes in the active site. Here too, we isolated more than 10,000 catalytically active and very diverse designs. Contrasting active and inactive designs illuminates areas for improving enzyme design methodology. This new approach to high-throughput design allows the systematic exploration of sequence and structure spaces of enzymes, binders and other functional proteins.

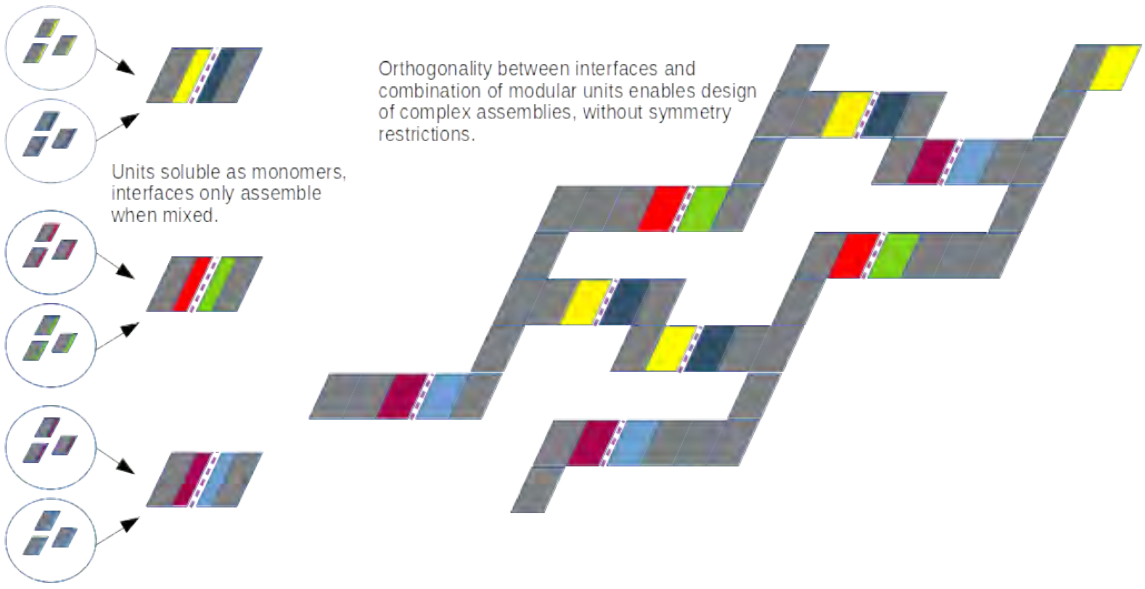
References

1. Khersonsky, O. & Fleishman, S. J. What Have We Learned from Design of Function in Large Proteins? *BioDesign Research* **2022**, 1–11 (2022).
2. Weinstein, J., Khersonsky, O. & Fleishman, S. J. Practically useful protein-design methods combining phylogenetic and atomistic calculations. *Curr. Opin. Struct. Biol.* **63**, 58–64 (2020).
3. Goldenzweig, A. *et al.* Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Mol. Cell* **63**, 337–346 (2016).
4. Hettiaratchi, M. H. *et al.* Reengineering biocatalysts: Computational redesign of chondroitinase ABC improves efficacy and stability. *Sci Adv* **6**, eabc6378 (2020).
5. Campeotto, I. *et al.* One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 998–1002 (2017).
6. Malladi, S. K. *et al.* One-step sequence and structure-guided optimization of HIV-1 envelope gp140. *Curr Res Struct Biol* **2**, 45–55 (2020).
7. Allouche-Arnon, H. *et al.* Computationally designed dual-color MRI reporters for noninvasive imaging of transgene expression. *Nat. Biotechnol.* (2022) doi:10.1038/s41587-021-01162-5.
8. Vos, P. D. *et al.* Computationally designed hyperactive Cas9 enzymes. *Nat. Commun.* **13**, 3023 (2022).
9. Barber-Zucker, S. *et al.* Stable and Functionally Diverse Versatile Peroxidases Designed Directly from Sequences. *J. Am. Chem. Soc.* **144**, 3564–3571 (2022).
10. Khersonsky, O. *et al.* Automated Design of Efficient and Functionally Diverse Enzyme Repertoires. *Mol. Cell* **72**, 178–186.e5 (2018).
11. Bengel, L. L. *et al.* Engineered enzymes enable selective N-alkylation of pyrazoles with simple haloalkanes. *Angew. Chem. Int. Ed Engl.* **60**, 5554–5560 (2021).

Specific Protein-Protein Interfaces: Towards multi-component modular protein systems

Frances Gidley, Fabio Parmeggiani
University of Bristol, School of Chemistry

Protein assemblies facilitate many essential roles in biological systems, from structural support to communication and catalysis. Designed self-assembling modular protein structures with specific geometric and spatial properties can be applied to novel biomaterial needs, including in the influencing of cell behaviour through the organised display of extracellular signal molecules. Use of computational design and modeling methods to design novel protein-protein interfaces and large protein scaffold assemblies will enable production and design of specific, functionalisable, multi-component assemblies and production of tuneable nanomaterials for bio-applications and beyond.
A set of novel and orthogonal designed helical repeat protein dimers are described here. These dimeric interfaces are modular building blocks, and larger designs capable of self-assembly into larger protein assemblies are currently undergoing experimental characterisation.



Computational Protein Design: then and now

Gevorg Grigoryan[1,2]

[1] *Generate Biomedicines, Inc., Somerville, MA, USA*
[2] *Dartmouth College, Hanover, NH, USA*

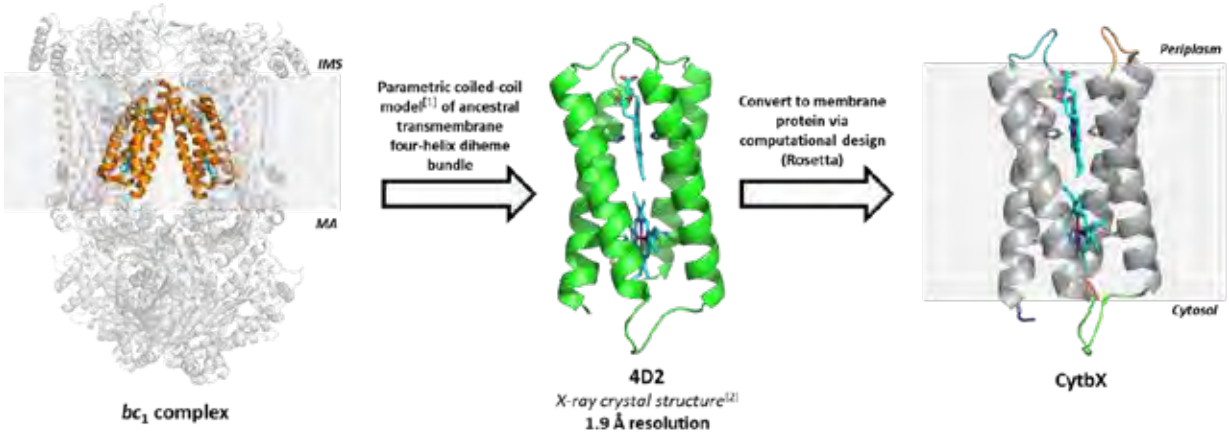
Successful attempts to design novel protein sequences to carry out pre-defined structural tasks date back to at least the 1980's. As a field, we have come a long way since then, with any number of immensely impressive milestones achieved. At the same time, certain key elements have been lacking in our computational approaches—namely robustness, reproducibility, and predictivity. This has meant that many biological problems remained out of reach of routine protein design. In the meantime, we are seeing a burgeoning of machine-learning (ML) techniques applied to problems of protein science generally and protein design specifically. It is exciting to see the beginning of this increased participation in the field. At the same time, it is interesting to ask what is fundamentally going on. In this talk, I will attempt to build intuition as to some of the reasons to expect that ML-based methods may do better than traditional techniques, and in what way. I will refer to some of the work ongoing at Generate, earlier work at Dartmouth, as well as broader results in the field. I will argue that 1) the universe of stable protein conformations is a highly patterned space, 2) ML methods are allowing us to depart, in a systematic manner, from the confines of low-order approximate potentials we have traditionally used to describe this space, and 3) new data-driven approaches to understanding proteins are pointing to a new paradigm for conducting molecular science. I will conclude with a perspective on what we may expect going forward and what the new challenges in our field may be.

Computational Design of a *de novo* Transmembrane Cytochrome

Benjamin J. Hardy[1,2*], Alvaro Martin Mermosilla[1], Dinesh K. Chinthapalli[3], Carol V. Robinson[3], J. L. Ross Anderson[1,2], Paul Curnow[1,2]

[1]School of Biochemistry, University of Bristol, UK; [2]BrisSynBio, Life Sciences Building, Tyndall Avenue, Bristol, UK; [3]Department of Chemistry, University of Oxford, UK

**Corresponding email: ben.hardy@bristol.ac.uk*



The *de novo* design of bioenergetic proteins provides insight into the underlying principles of protein folding and assembly, permits the study of electron transport in minimal biomimetic systems, and produces new components for synthetic biology. Many respiratory and photosynthetic complexes share an ancestral four-helix bundle architecture at their core that positions two b-type hemes for transmembrane electron transport. This fundamental scaffold is an ideal basis for creating man-made bioenergetic membrane proteins. We report here the design, recombinant production and characterisation of such an artificial protein through the computational transformation of an existing water-soluble coiled-coil design (4D2)[1,2] into a sequence capable of membrane integration. This design process retains the key interactions that govern bundle assembly and heme binding but otherwise the resulting construct, which we call cytochrome b xenoprotein (CytbX), shares no sequence similarity with any natural metalloprotein. CytbX can be recombinantly expressed in *E. coli and S. cerevisiae*, is efficiently inserted into cellular membranes, and recruits heme *in vivo*. This artificial protein can be successfully purified from membranes using mild non-ionic detergents and exhibits exceptional thermal stability (T_M > 95 °C). Absorbance spectroscopy and native mass spectrometry confirm that CytbX binds to two hemes via bis-histidine coordination, and redox potentiometry reveals splitting of heme reduction potentials reminiscent of natural transmembrane cytochromes. The hemes of CytbX are accessible to small diffusive redox partners enabling this protein to engage in a range of biomimetic electron transport reactions. These findings suggest that coiled-coil architectures provide a route to novel transmembrane bioenergetic assemblies that can capture the core attributes of natural proteins and engage with cellular biology.

References:

(1) Ghirlanda, G.; Osyczka, A.; Liu, W.; Antolovich, M.; Smith, K. M.; Dutton, P. L.; Wand, A. J.; DeGrado, W. F. De Novo Design of a D2-Symmetrical Protein That Reproduces the Diheme Four-Helix Bundle in Cytochrome Bc1. *J. Am. Chem. Soc.* **2004**, *126* (26), 8141–8147. https://doi.org/10.1021/ja039935g.

(2) Hutchins, G. H.; Noble, C. E. M.; Blackburn, H.; Hardy, B.; Landau, C.; Parnell, A. E.; Yadav, S.; Williams, C.; Race, P. R.; Sofia, A.; et al. Precision Design of Single and Multi-Heme de Novo Proteins. *bioRxiv* **2020**, 2020.09.24.311514. https://doi.org/10.1101/2020.09.24.311514.

Towards a coiled coil with a polar core

Birte Hernandez Alvarez
Max Planck Institute for Biology, Department of Protein Evolution
Max-Planck-Ring 5, 72076 Tübingen, Germany

Naturally occurring coiled coils often deviate from the canonical model, which is characterized by a regular sequence pattern built from heptad repeats (*abcdefgh*) with hydrophobic residues in the core positions *a* and *d*. The occupation of single core positions of one or more consecutive heptads by polar residues is a rather common deviation from the standard. Such polar core residues locally destabilize the coiled coil structure, which in turn is often crucial for functionality, as has been shown for many examples.

We have looked at the coiled-coil domains from various phage proteins and from trimeric autotransporter adhesins of Gram-negative bacteria. In these coiled coils, polar residues are rather frequent and found in up to 50 % of the core positions, in rare cases they even occupy both core positions of a single repeat. Using bioinformatic analyses, we identified polar core motifs, such as the anion-binding motif N@*d* (hxxNxx) and a Ca^{2+}-binding motif with N@*a*, Q@*d* and D@*e (*NxxQDx*)*, and characterized them biophysically and structurally. Analysis of both motifs in the GCN4 background showed that the coordination of ions in the core of trimeric coiled coils provides structural specificity. Furthermore, we designed a trimeric coiled coil comprising a central segment of identical, consecutive heptads with exclusively polar core residues flanked by GCN4-pII adaptors. Starting from this idealized construct of high sequence symmetry, we created a second design more similar to naturally occurring sequences by increasing the sequence length, replacing individual residues, inserting specific salt bridges, and introducing a trigger motif. Finally, we shortened the flanking GCN4 adaptors to assess the impact of flanking canonical repeats on the folding propensity and stability of coiled-coil segments with polar cores.

The structural and biophysical data obtained in this work provide a clear picture of the strategies used by trimeric coiled coils to accommodate polar side chains in their cores.

Navigating the structure and stability landscape of *de novo* $(\beta\alpha)_8$-barrels by protein design and engineering

S. Romero-Romero[1], S. Kordes[1], D.A. Fernandez-Velasco[2], B. Höcker[1]

[1]Biochemistry III, University of Bayreuth, 95440 Bayreuth, Germany

[2]Faculty of Medicine, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico

Birte.Hoecker@uni-bayreuth.de

The $(\beta\alpha)_8$-barrel or TIM-barrel fold is one of the most abundant topologies in nature and a common functional scaffold [1]. It is thought to have evolved from smaller βαβ-units through duplication and fusion and it has been possible to design an idealized four-fold symmetric version *de novo* [2]. Starting from this first version, we build a collection of stable *de novo* TIM barrels (DeNovoTIMs) using a computational fixed-backbone and a modular approach based on improved hydrophobic packing [3] and the introduction of salt-bridges [4]. DeNovoTIMs were subjected to thorough biochemical and folding analyses to explore their structure and stability. We find that DeNovoTIMs navigate a region of the stability landscape different to natural proteins, with variations spanning 60 degrees in melting temperature and 22 kcal per mol in conformational stability throughout the designs. Significant non-additive or epistatic effects were observed in their stability and structural features when stabilizing mutations from different barrel regions were combined. Salt-bridge variants exhibit important differences in comparison with the parental proteins, both in conformational stability and structural properties (Figure). The design of stable proteins increases the applicability of *de novo* proteins and provides important information on the molecular determinants of sequence-structure-stability relationships - an essential step towards fine-tuned modulation of protein stability by protein design and engineering.

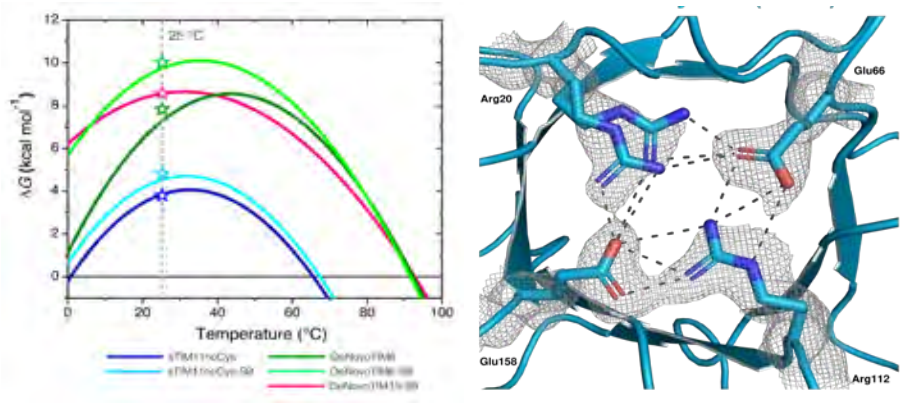


Figure:
Left: Stability curves of the salt bridge variants. Constructed from DSC experiments and Gibbs-Helmholtz eq. Symbols indicate Δ*G* values at 25 °C determined by chemical unfolding.
Right: Structural conformations of salt bridge cluster in sTIM11noCys-SB (pdb id: 7OSU). Salt bridge residues highlighted in sticks. 2Fo–Fc electron density map contoured at 1σ shown as gray mesh.

1. Romero-Romero, Kordes, Michel, Höcker (2021) Curr Opin Struct Biol ***68****, 94-104.*

2. Huang, Feldmeier, Parmeggiani, Fernández-Velasco, Höcker, Baker (2016) Nature Chem Biol ***12****, 29-34.*

3. Romero-Romero, Costas, Silva Manzano, Kordes, Rojas-Ortega, Tapia, Shanmugaratnam, Rodríguez-Romero, Baker, Höcker, Fernández-Velasco (2021) J Mol Biol ***433****, 167153.*

4. Kordes, Romero-Romero, Lutz, Höcker (2022) Protein Sci ***31****, 513-527.*

EEA1 and Rab5 form a two-component molecular motor at the early endosome

Marcus Jahnel

Physics of Life Cluster of Excellence and BIOTEC / TU Dresden, Dresden, Germany

A recurring theme exists in vesicle trafficking: unusually long membrane-bound coiled-coil tethering proteins are often paired with specific small globular GTPases on incoming vesicles. Examples include Golgin-245 and Arl1, GMAP-210 and Arf1, or EEA1 and Rab5, with size differences of more than 50-fold between partners. What is the purpose of this uneven molecular pairing?

On early endosomes, binding of the small GTPase Rab5 induces large conformational changes within the coiled-coil tethering protein EEA1, which switches from an extended to a more flexible collapsed state. Earlier, we showed that this entropic collapse of EEA1 gives rise to an effective entropic force that can pull tethered membranes closer together [I]. However, it remained unclear whether EEA1 could spontaneously recover from the collapsed state after interacting with the GTPase. Here, using fluorescence correlation spectroscopy, we show that EEA1 can undergo multiple cycles of this flexibility transition in a GTPase-dependent manner without needing additional factors [II]. Based on these observations, we developed a semi-flexible polymer model (Fig. 1) to describe the mechano-chemical cycle that drives this two-component molecular motor and could analyse its efficiency. We hope our work sheds new light on active GTP-driven molecular systems that take advantage of bistable coiled-coil domains in organising eukaryotic vesicle transport.

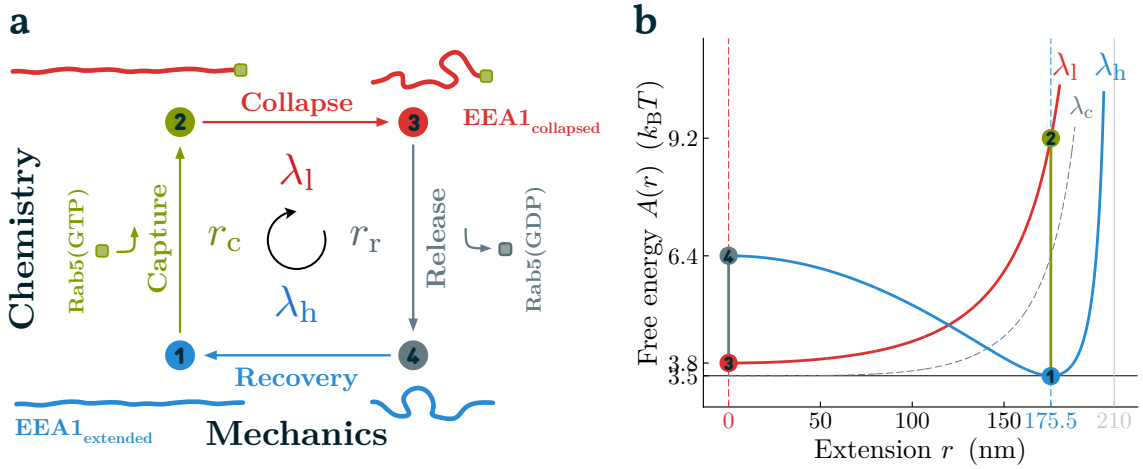


***Figure**: a) In an idealised two-state model for tethering at the early endosome, the chemical states of the signalling molecule (active/inactive) Rab5 are coupled to mechanical states (rigid/flexible) of the long coiled-coil tether EEA1. b) Our coarse-grained semi-flexible polymer model estimates the free energy differences between the individual states and allows for a calculation of the motor efficiency.*

References:

1. D. H. Murray *et al.*, An Endosomal Tether Undergoes an Entropic Collapse to Bring Vesicles Together, Nature 537, 107 (2016).

2. A. Singh *et al.*, GTPase dependent cyclic flexibility transitions drive the two-component EEA1-Rab5 molecular motor. *bioRxiv* (2022).

Molecular Clamps Chaperone Collagen Folding – Abhishek Jalan

Collagens are highly abundant human proteins. They are composed of three polypeptides containing Gly-Xaa-Yaa triplet repeats that supercoil to form a triple-helical structure. It has been proposed that human collagens fold via a nucleation-zipper mechanism, in which three peptides nucleate at the C-termini and then propagate towards the N-terminal like a zip-chain.[1] A collagen-specific chaperone heat shock protein HSP47 is proposed to bind unique sites during the propagation phase providing local stability and offsetting the entropic cost of folding like a zip-chain.[2] However, the HSP47-chaperoned paradigm does not explain several features of collagen folding. First, high affinity HSP47 binding sites are overwhelmingly located towards the N-terminal half of some of the most abundant human collagens.[3–5] The implication is that HSP47 binds high affinity sites after half the triple-helix has already folded. Second, HSP47 does not recognize collagen type XII.[3] Third, human collagens come in two flavours, those that contain uninterrupted Gly-Xaa-Yaa repeats and those in which this perfect repeat pattern is interrupted by non-collagenous sequences.[6] Interruptions dramatically perturb the triple-helical structure resulting in decreased stability and slow folding rate.[7] HSP47 binding sites are not located close to interruption sites. Thus, how collagens re-nucleate after an interruption and compensate for the loss in stability and folding rate is not understood. And finally, polypeptides in triple-helices adopt a staggered alignment with respect to each other. This is critical for correct supramolecular assembly and protein-protein recognition. However, incorrect slippage of polypeptides within the triple helix can be readily introduced.[8] The HSP47-chaperoned folding does not account for how native collagens avoid similar slippages during folding.

The information for correct folding of most proteins is encoded in their amino acid sequence. We demonstrate this to be true for collagens as well. We show that all 28 known human collagens contain specialized molecular motifs capable of geometrically specific interchain interactions. These motifs are anomalously abundant in all collagens at a level not explained by the principle of random distribution. Incorporating these motifs into collagen triple-helical peptides increases folding rate but decrease the unfolding rate compared to control sequences. Importantly, these motifs are selectively enriched on either side of the interruption sites in all collagens. The fast folding and slow unfolding attributes of the motifs could provide a mechanism for how triple-helices re-nucleate and compensate for the loss in stability near the interruption sites. Furthermore, incorrect alignment of polypeptides within the triple-helices abrogates the molecular interaction within the motifs. Most revealingly, residues present in these motifs colocalize with several known disease-causing mutation across diverse collagen types, testifying to their biological relevance. To summarize, the high general abundance, ability to form geometrically specific interactions, increased folding rate, high kinetic stability, association with disease-causing mutations and selective enrichment close to the interruption sites suggests that the motifs act as molecular clamps that chaperone collagen folding. These results open up new locations for therapeutic intervention during mutation-induced collagen pathologies.

[1] J. Engel, D. J. Prockop, *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 137.
[2] T. Koide, Y. Nishikawa, S. Asada, C. M. Yamazaki, Y. Takahara, D. L. Homma, A. Otaka, K. Ohtani, N. Wakamiya, K. Nagata, K. Kitagawa, *J. Biol. Chem.* **2006**, *281*, 11177.
[3] A. Köhler, M. Mörgelin, J. M. Gebauer, S. Öcal, T. Imhof, M. Koch, K. Nagata, M. Paulsson, M. Aumailley, U. Baumann, F. Zaucke, G. Sengle, *FASEB J.* **2020**, *34*, 12040.
[4] H. Cai, P. Sasikumar, G. Little, D. Bihan, S. W. Hamaia, A. Zhou, J. M. Gibbins, R. W. Farndale, *Biomolecules* **2021**, *11*, 983.
[5] E. T. Abraham, S. Oecal, M. Mörgelin, P. W. N. Schmid, J. Buchner, U. Baumann, J. M. Gebauer, *J. Biol. Chem.* **2021**, 101169.
[6] S. Ricard-Blum, *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, 1.
[7] X. Sun, Y. Chai, Q. Wang, H. Liu, S. Wang, J. Xiao, *Biochemistry* **2015**, *54*, 6106.
[8] A. A. Jalan, K. A. Jochim, J. D. Hartgerink, *J. Am. Chem. Soc.* **2014**, *136*, 7535.

Coiled-coil modules for designed protein folds and cellular logic circuits

Roman Jerala

Department of synthetic biology and immunology, National institute of chemistry, Ljubljana, Slovenia.

Coiled-coil (CC) dimers are ubiquitous building modules in natural proteins. CC dimers can be concatenated into longer chains for the construction of new modular protein folds based on topological principles distinct from natural proteins. Coiled-coil protein origami (CCPO) structures have interesting properties and are highly designable, enabling multiple use of the same building modules, design of the folding pathways and other functional properties. Although CC dimers occur frequently in natural proteins, designed CC dimers can also be used for cell regulation. We have designed new CC pairs that can fused to other proteins, enabling new type of regulation of biological processes, including localization multiplexing within cells, augmented transcriptional response based on chemical regulators (CCCtag) and faster kinetics based on combination of split proteases with coiled-coil modules (SPOC logic). Those building blocks have also been used to construct two genetically encoded orthogonal secretion systems for fast secretion of proteins based on the pool of ER-retained proteins. A generally applicable regulation platform INSRTR (inserted peptide structure regulator), is based on regulation of selected protein activity by the addition of a peptide that forms a rigid coiled-coil dimer with an inserted peptide allosterically disrupts the protein function. This platform was developed to enable the construction of ON/OFF protein switches, their regulation by small molecules, and logic functions with a rapid response in mammalian cells. INSRTR was demonstrated on several diverse proteins with versatile activities including enzymes, signaling mediators, DNA binders/transcriptional regulators, fluorescent proteins, and antibodies implemented as a sensing domain of anticancer chimeric antigen receptors on T cells. INSRTR platform presents extraordinary potentials for regulating biological systems and therapeutic applications.

Computational design of new protein shapes and functions

Tanja Kortemme

University of California San Francisco (UCSF), San Francisco, CA, USA,

http://kortemmelab.ucsf.edu

tanjakortemme@gmail.com

There has been exciting progress in the computational design of proteins with new structures [1], highlighting the potential to advance many applications in biological engineering, as well as to provide insights into the design principles of natural protein functions. Many significant challenges remain, both in the accuracy of current computational approaches, and in the complexity of protein geometries and functions that can be designed at present. I will discuss our recent progress with new computational methods and their applications. Our work includes (i) reshaping of protein conformations for reprogrammed functions [2], (ii) engineering proteins to detect and respond to new small molecule signals in living cells [3], and (iii) controlling protein shapes to create fold families for new functions [4]. Emerging machine learning methods provide opportunities and challenges.

[1] Pan X, Kortemme T. Recent advances in de novo protein design: principles, methods, and applications. *Journal of Biological Chemistry*. 2021 Mar 18:100558.

[2] Krivacic C*, Kundert K*, Pan X*, Pache RA*, Liu L, Conchúir SO, Jeliazkov JR, Gray JJ, Thompson MC, Fraser JS, & Kortemme T. Accurate positioning of functional residues with robotics-inspired computational protein design. *Proc Natl Acad Sci USA* 2022; 119(11), e2115480119.

[3] Glasgow AA*, Huang YM*, Mandell DJ*, Thompson M, Ritterson R, Loshbaugh AL, Pellegrino J, Krivacic C, Pache RA, Barlow KA, Ollikainen N, Jeon D, Kelly MJS, Fraser JS, Kortemme T. Computational design of a modular protein sense-response system. *Science* 2019, 366(6468):1024-1028

[4] Pan X, Thompson M, Zhang Y, Lin L, Fraser JS, Kelly MJ, Kortemme T. Expanding the space of protein geometries by computational design of de novo fold families. *Science* 2020, 369(6507):1132-1136.

Coiled-coil peptide-mediated RNA delivery

Alexander Kros

RNA therapy relies on delivery of exogenous (therapeutic) RNA molecules, such as messenger (mRNA) or small interfering RNA (siRNA), to control disease-relevant gene expression. For efficient functional cytosolic delivery to, and release within, target cells, these highly charged, immunogenic and membrane-impermeable RNA molecules require the use of delivery systems. To this end, ionizable lipid nanoparticles (LNPs) serve as state-of-the-art vehicles that can package, protect and release RNA molecules inside cells. LNPs have realized the translation of RNA therapeutics to the clinic, highlighted by the approval of Onpattro®, enabling RNA interference (RNAi) therapy for the treatment of polyneuropathies resulting from transthyretin-mediated amyloidosis. In addition, this platform has been successfully expanded for the delivery of other RNA molecules, yielding safe and effective mRNA-based vaccines for SARS-CoV-2.

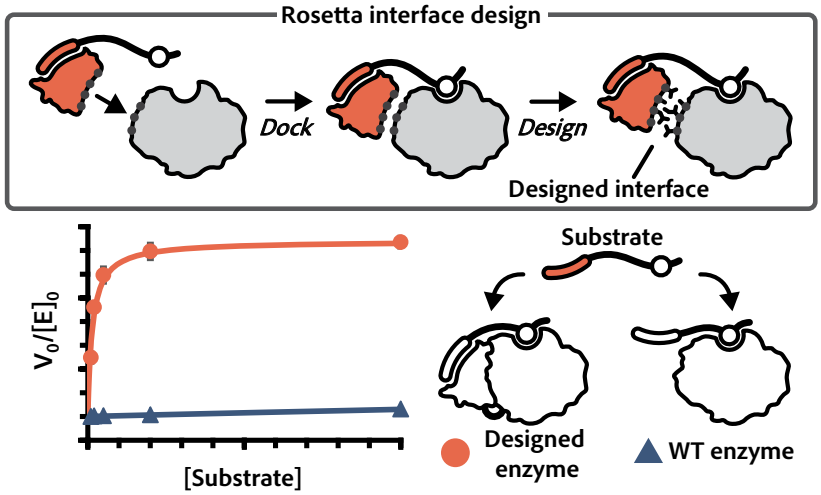
However, like most nonviral delivery systems, LNPs are very inefficient in escaping the endosome resulting in very low amounts of RNA reaching the cytosol (<5 %). Therefore a highly efficient gene delivery system using fusogenic coiled-coil peptides was developed. We modified LNPs encapsulating GFP-encoding mRNA, and cells with complementary coiled-coil lipopeptides. Coiled-coil formation between these lipopeptides induced fast nucleic acid uptake and enhanced GFP expression. The cellular uptake of coiled-coil modified LNPs is likely driven by membrane fusion thereby omitting typical endocytosis pathways. This direct cytosolic delivery circumvents the problems commonly observed with the limited endosomal escape of mRNA. Therefore fusogenic coiled-coil peptide modification of existing LNP formulations to enhance nucleic acid delivery efficiency could be beneficial for several gene therapy applications.

Title: **Improving enzyme efficiency via designer incorporation of a substrate recruitment domain**

Rodney Park[1], Chayanid Ongpipattanakul[2], Satish K. Nair[2,3,4], Albert A. Bowers[5*], Brian Kuhlman[1,6*]

[1]Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599
[2]Department of Biochemistry, University of Illinois at Urbana–Champaign, Urbana, IL 61801
[3]Center for Biophysics and Computational Biology, University of Illinois at Urbana–Champaign, Urbana, IL 61801
[4]Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL 61801
[5]Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27599
[6]Lineburger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Abstract: Promiscuous enzymes that modify peptides and proteins are powerful tools for labeling biomolecules; however, directing these modifications to desired substrates can be challenging. Here, we use computational interface design to install a substrate recognition domain adjacent to the active site of a promiscuous enzyme, catechol O-methyltransferase (COMT). This effectively decouples substrate recognition from the site of catalysis and promotes modification of peptides recognized by the recruitment domain. We determined the crystal structure of this novel multi-domain enzyme, SH3-588, which shows that it closely matches our design. SH3-588 methylates directed peptides with catalytic efficiencies exceeding the WT enzyme by over 1000-fold, whereas peptides lacking the directing recognition sequence do not display enhanced efficiencies. The designer enzyme also prefers to modify directed substrates over undirected in competition experiments, suggesting that we can use protein design to direct posttranslational modifications to specific sequence motifs on target proteins in complex multi-substrate environments.



Socket2 and CC$^+$ 2022:
Bioinformatics resources for assigning and analyzing coiled-coil structures and models

Prasun Kumar and Derek N. Woolfson
University of Bristol, UK

Nearly all biological activities depend on protein-protein interactions. Many such interactions are facilitated by α-helical coiled coils (CC), which account for up to 5% of all proteomes. These CCs range in oligomeric states from dimers to 15-mers, and come in different lengths and various topologies. Thus, the ability to extract, visualize and analyze CC structures aids biological research in general, and protein design and engineering specifically. Socket is one such programme that finds CC motifs in protein structures by identifying their signature knobs-into-holes (KIH) packing between the component helices.[1] This was presented by Walshaw and Woolfson lab in 2001. Since then, the number and variety of structurally defined CCs deposited in the RCSB PDB has increased considerably. Therefore, to address some gaps in Socket that might fail to capture all of these, we have made changes to Socket to give Socket2.[2] Socket2 can recognize all CC types based on the 20 proteinogenic residues or including other amino acids. The program also has a new and easy-to-use web server, and a GUI with additional features accessed freely at http://coiledcoils.chm.bris.ac.uk/socket2/home.html.

Using Socket2, we have generated a beta-version update to the CC database, CC$^+$,[3] from all of the KIH-positive structures from the PDB. CC$^+$ is a searchable repository of CC structures. For the new 2022 version, we have included CCs identified in tertiary models predicted by AlphaFold2 and held at EBI (https://alphafold.ebi.ac.uk/). This includes CCs predicted from 48 genomes. We believe that the insights from this expanded database will, for example, enable hypothesis-driven cell-biology research, and aid the determination and understanding of further sequence-and-structure relationships for CCs to guide protein engineering and design.

1. Walshaw, J. and Woolfson, D.N., 2001. Socket: a program for identifying and analysing coiled-coil motifs within protein structures. ***J Mol Biol 307***(5), pp.1427-1450.
2. Kumar, P. and Woolfson, D.N., 2021. Socket2: a program for locating, visualizing and analyzing coiled-coil interfaces in protein structures. ***Bioinformatics 37***(23), pp.4575-4577.
3. Testa, O.D., Moutevelis, E., and Woolfson, D.N., 2009. CC plus : a relational database of coiled-coil structures. ***Nucleic Acids Res* 37**, D315-D322.

Constructing synthetic-peptide assemblies from de novo designed 3_{10} helices

P Kumar[1], NG Paterson[2], J Clayden[1], DN Woolfson[1,3,4]

[1]School of Chemistry, University of Bristol, Bristol, UK
[2]Diamond Light Source, Didcot, UK
[3]School of Biochemistry, University of Bristol, Bristol, UK
[4]BrisSynBio, University of Bristol, Bristol, UK

Compared with the iconic α helix, 3_{10} helices occur less frequently in protein structures. The different 3_{10}-helical parameters lead to energetically less favourable internal energies, and a reduced tendency to pack into defined higher-order structures. Consequently, in natural proteins, 3_{10} helices rarely extend past 6 residues, and do not form regular supersecondary, tertiary, or quaternary interactions. Here, we show that despite their absence in nature, synthetic protein-like assemblies can be built from 3_{10} helices (Fig. 1). We report the rational design, solution-phase characterisation, and X-ray crystal structures of water-soluble bundles of 3_{10} helices with consolidated hydrophobic cores[1]. The design uses 6-residue repeats informed by analysis of short 3_{10} helices in natural proteins and incorporates α-aminoisobutyric acid residues. Design iterations reveal a tipping point between α-helical and 3_{10}-helical folding, and identify features required for stabilising assemblies in this unexplored region of protein-structure space.

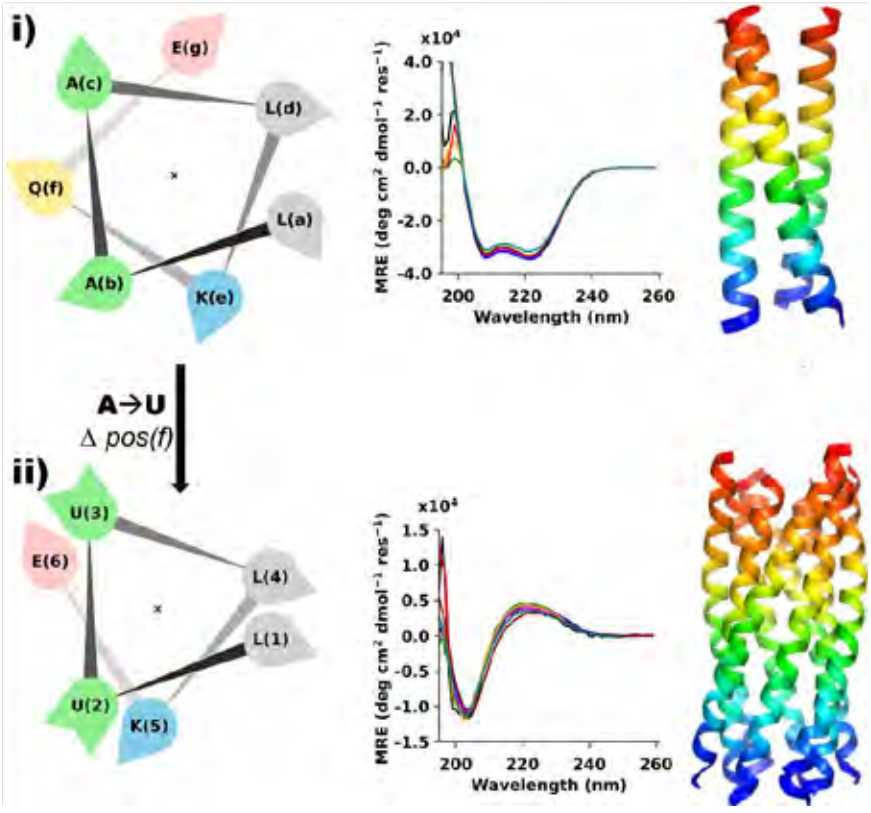


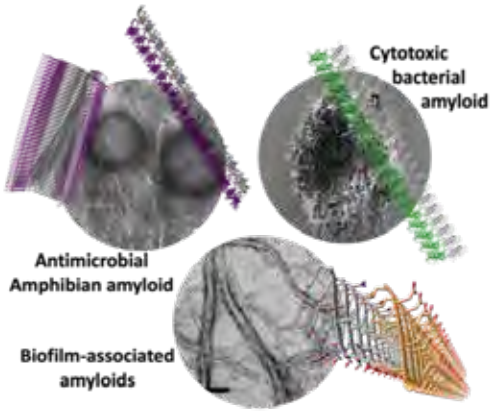
Figure 1: Design principles, biophysical characterization, and X-ray crystal structures for *de novo* helical bundles. From left to right: helical wheels, circular dichroism spectra, and cartoons of X-ray structures for (i) a *de novo* trimeric coiled coil (PDB ID: 7qdk), and (ii) a completely new 3_{10}-helical assembly (PDB ID: 7qdi).

1 Kumar, P., Paterson, N. G., Clayden, J. & Woolfson, D. N. De novo design of discrete, stable 310-helix peptide assemblies. *Nature*, doi:10.1038/s41586-022-04868-x (2022).

Virulent and Antimicrobial Amyloids in Infections and Neurodegeneration

Meytal Landau
Associate Professor, Department of Biology, Technion-Israel Institute of Technology, Haifa, Israel.
Visiting Group Leader, The European Molecular Biology Laboratory (EMBL), Hamburg, Germany
Associate member, Center for Structural Systems Biology (CSSB), Hamburg, Germany

Amyloids are protein fibers with robust structures, which are known mainly in the context of neurodegenerative diseases yet are secreted by species across kingdoms of life to carry out physiological function and help survival and activity. For example, several microbial amyloids serve as key "weapons" making infections more aggressive. Thereby, they exposed new routes for the development of novel antivirulence drugs, which may elicit less resistance as the evolutionary pressure on the microbe is less profound compared to bactericidal drugs. Our laboratory published the first structures of bacterial amyloid fibrils involved in virulent activities. Our findings thus far exposed an extreme structural diversity, extending beyond canonical amyloid cross-β structures, and encoding different activities. In particular, the discovery of a novel class of cross-α amyloid fibrils of toxic peptides presented a unique protein architecture, offered drug targets and leads, and opened a fresh perspective to study amyloid-related toxicity. Moreover, we revealed that amyloids secreted by bacteria highly abundant in the microbiome and food sources show similarities in molecular structures to human amyloids involved in neurodegenerative diseases such as Alzheimer's and Parkinson's. This might raise concerns about the involvement of microbes in facilitating these diseases, similar to prion proteins transmitted by contaminated meat that elicit the Creutzfeldt-Jakob disease. In addition, we identified peptides produced across species that provide antimicrobial protection that form amyloid fibrils and determined their first high resolution structures. This amyloid-antimicrobial link signifies a physiological role in neuroimmunity for human amyloids. Such antimicrobial fibrils can facilitate the design of functional and stable nanostructures to serve as a stable coating for medical devices or implants, industrial equipment, food packing and more.



Atomic structures of amyloid fibrils determined by X-ray crystallography and cryogenic electron microscopy (cryo-EM) of microbial and antimicrobial amyloids. A scanning electron micrograph shows cells damaged by a cytotoxic bacterial amyloid peptide, and transmission electron micrographs display fibrils of antimicrobial peptides covering bacterial cells, and of massive fibrils formed by biofilm-associated amyloids.

References

1. E. Tayeb-Fligelman, O. Tabachnikov, A. Moshe, O. Goldshmidt-Tran, M.R. Sawaya, N. Coquelle, J-P. Colletier, and M. Landau. *Science* 355(6327): 831-833; 2017
2. N. Salinas, A. Moshe, J-P. Colletier, and M. Landau. *Nat. Commun.* 9(1):3512; 2018.
3. S. Perov, O. Lidor, N. Salinas, N. Golan, E. Tayeb-Fligelman, M. Deshmukh, D. Willbold, and M. Landau. *PLoS Pathog* 15(8): e1007978; 2019
4. E. Tayeb-Fligelman, N. Salinas, O. Tabachnikov, and M. Landau. *Structure* S0969-2126(19)30445-9; 2020
5. Y. Engelberg and M. Landau. *Nat. Commun.* 11, 3894; 2020
6. N. Salinas, E. Tayeb-Fligelman, M. Sammito, D. Bloch, R. Jelinek, D. Noy, I. Uson, and M. Landau. *PNAS* 118 (3) e2014442118; 2021
7. Y. Engelberg, P. Ragonis-Bachar and M. Landau. *Biomacromolecules* 2022
8. R. Bücker, C. Seuring, C. Cazey, K. Veith, M. García-Alai, K. Grünewald, and M. Landau. The Cryo-EM Structures of two Amphibian Antimicrobial Cross-β Amyloid Fibrils. *Nat. Commun 13:* 4356; 2022
9. P. Ragonis-Bachar*, B. Rayan*, E. Barnea, Y.Engelberg, A. Upcher, M. Landau. Natural Antimicrobial Peptides Self-assemble as alpha/beta Chameleon Amyloids. *Biomacromolecules, in press*

Designing Proteins to Protect Fluorescent Dyes in Light-Emitting Diodes

Horst Lechner[a], Liher García[b], Aitziber L. Cortajarena[b], Sara Ferrara[c], Rubén D. Costa[c], Claudia Barolo[d], Pedro B. Coto[e], Gustav Oberdorfer[a]

[a] *Graz University of Technology, Institute of Biochemistry, Petersgasse 12, 8010 Graz, Austria*
[b] *Center for Cooperative Research in Biomaterials (CIC biomaGUNE), Parque Científico y Tecnológico de Gipuzkoa, Paseo Miramón 194 edificio C, 20014 San Sebastián, Spain*
[c] *Technical University of Munich, Chair of Biogenic Functional Materials, Schulgasse 22, 94315 Straubing, Germany*
[d] *Department of Chemistry and NIS Interdepartmental Centre, University of Turin, Via P. Giuria 7, 10125 Turino, Italy*
[e] *Spanish National Research Council (CSIC) and Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal 5, 20018 San Sebastián, Spain*

We redesign proteins to be employed as artificial fluorescent proteins. These fluorescent proteins are used as down-converting filters for the emitting source of (blue) light-emitting-diodes (LEDs) to create white Bio-LEDs. They are used as hosts for small organic compounds (dyes) with fluorescent properties, similar to the ones used in organic LEDs (OLEDs). The protein should protect these compounds from non-radiative deactivation and aggregation - central problems with these kinds of fluorescent molecules rendering them less attractive for the use in white LEDs. Additionally, the derived proteins will be embedded in a polymer matrix to increase their stability against pH changes, temperature and the high photon-flux occurring when the blue LED-emitter is covered with this biological filter.

Our major aim in this multi-disciplinary project is to computationally identify and redesign proteins that can accommodate and specifically bind the dyes, while retaining high (thermal) stability.

To achieve this goal, we curated a library of several hundred protein scaffolds, and selected the best hits after docking or placing the dyes and redesigned them using different design strategies as implemented in the Rosetta suite. Additionally, some fluorescent compounds might be anchored into the binding pocket of the protein via a covalent attachment using a non-canonical amino acid or a cysteine residue. To accomplish this, a suitable anchoring point in the pocket was identified during design.

We currently have a few dozen designs for several of our envisaged dyes expressed and selected ones were evaluated concerning their binding, spectroscopic and biophysical properties. One design was characterized in greater detail, crystal structures were obtained and its stability on an LED was determined to be more than 70 days.

The structural diversity of natural coiled coils

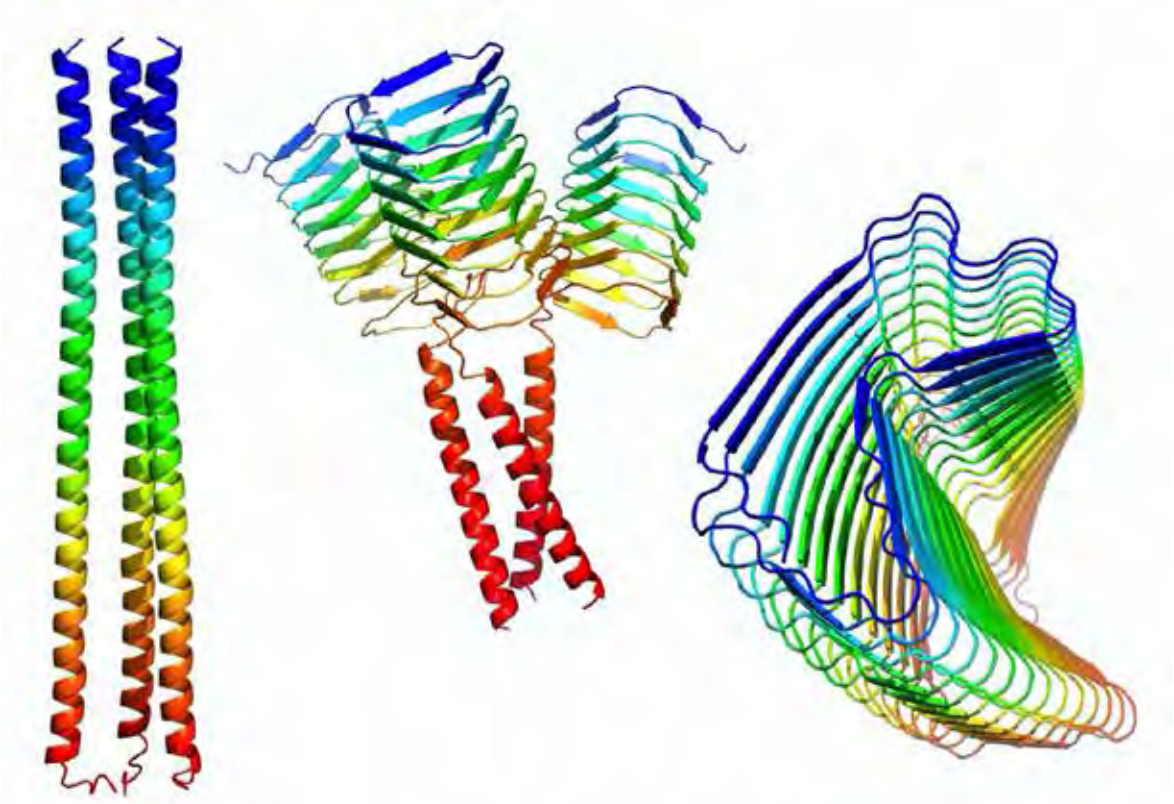
Andrei N. Lupas, Stanisław Dunin-Horkawicz, Mikel Martinez Goikoetxea, Marcus Hartmann, Birte Hernandez Alvarez
Department of Protein Evolution, MPI for Biology, 72076 Tübingen, Germany; andrei.lupas@tuebingen.mpg.de

Seventy years ago, Francis Crick introduced the coiled coil to account for the structural properties of proteins referred to as 'k-m-e-f', for keratin, myosin, epidermin, and fibrinogen. His model envisaged 2 or 3 α-helices wound around each other in parallel orientation, systematically interlocking their side chains along the fiber in a pattern that would repeat every seven residues (the heptad repeat), specifying a supercoil of opposite handedness to that of the α-helix. Since then, the coiled-coil fold has been found to be vastly more diverse, encompassing structures of between two and more than 20 helices in parallel or antiparallel orientation, which may form fibers, levers, tubes, funnels, sheets, spirals, and rings. It covers periodicities leading to both left-handed and right-handed supercoils (and to straight helical bundles in between); and includes local departures from α-helical structure such as 3_{10}-helices, π-turns, and even short β-strands, as in the α/β-coiled coils. Here we will describe how a few simple biophysical rules can produce such a seemingly endless diversity.

Highly repetitive genomic ORFs: a source of new fibrous proteins?

Andrei N. Lupas
Department of Protein Evolution, MPI for Biology, 72076 Tübingen, Germany;
andrei.lupas@tuebingen.mpg.de

The proteome of life is mainly formed by proteins belonging to a few thousand domain families. However, besides these, each genome appears to encode proteins that have little or no sequence similarity to proteins in other genomes, often not even in different strains of the same species. These proteins, variously referred to as singletons or ORFans, may in some cases just correspond to untranscribed open reading frames, but proteomic studies show that many are in fact real proteins. Indeed, there is rapidly growing evidence that new proteins continuously arise in previously non-coding DNA sequence, particularly in multicellular eukaryotes. So far, none of the ones that have been studied show a tendency to fold into a defined tertiary structure, but at least in prokaryotes, which are not particularly tolerant of unstructured polypeptides in their cytosol and typically have mutation rates that allow them to clear junk efficiently from their genomes, it is not inconceivable that some are partly or entirely structured. Proceeding from the observation that the simplest mechanism to obtain a longer open reading frame is the amplification of a shorter sequence lacking stop codons and that the dominant mechanism to obtain a folded protein is the repetition of a shorter unstructured peptide, I have been collecting instances of such new repetitive proteins that are likely to fold. Analysis with AlphaFold unsurprisingly now indicates that most have the propensity to form fibrous folds.



Examples of potentially structured repetitive ORFs. *Left*: an ORF from Ehrlichia ruminantium with a 7-residue periodicity, *middle*: an ORF from Escherichia coli E110019 with two different 7-residue periodic segments, and *right*: an ORF from Nematostella vectensis with a 43-residue periodicity. These are AlphaFold models, colored blue to red from N- to C-terminus.

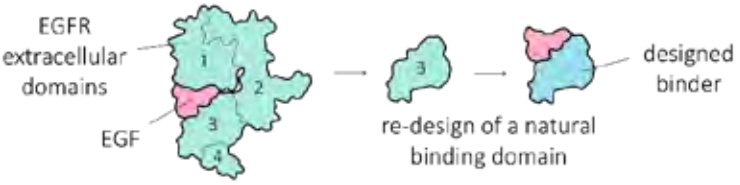
Protein design of growth factor inhibitors

Kateryna Maksymenko[1], Julia Skokowa[2], Patrick Müller[3], Andrei Lupas[1], Mohammad ElGamacy[1,2]

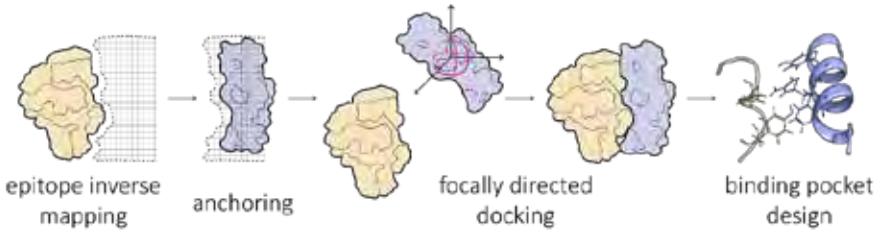
[1]Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Germany
[2]Division of Translational Oncology, University Hospital Tübingen, Germany
[3]University of Konstanz, Germany

Growth factors are signaling molecules coordinating the complex functionality of multicellular organisms during development and homeostasis. Since aberrant expression of growth factors can cause diverse disorders such as cancer, autoimmune and cardiovascular diseases, growth factors and their receptors are central targets for therapeutic modulation. One of the options to manipulate signaling interactions is to use protein-based binders that are highly specific and able to target various molecular surfaces. Here, we present two different strategies of computational protein design to obtain inhibitors against growth factors which are key modulators of tumor progression. The first approach requires the structure of a native growth factor : growth factor receptor complex and aims to re-engineer a natural binding domain to make it more soluble, more stable, or more affine. In contrast, the second approach relies only on the structure of a target epitope and takes advantage of a new software for massive-scale docking of a target site against a protein structure database to select the high shape complementary scaffolds. Adopting the first approach, we designed inhibitors of epidermal growth factor (EGF) using a single domain of EGF receptor as a template. Experimental evaluation of only two designed candidates revealed that both of them are solubly expressed and bind EGF with nanomolar affinities (i.e. 5-fold stronger than a native domain). Furthemore, one design supresses EGF-induced proliferation of epidermoid carcinoma cells, and shows specific inhibition of EGFR signaling in zebrafish embryos. Using the second strategy, we designed inhibitors of vascular endothelial growth factor (VEGF) based on two different scaffolds. The binding affinities of the designs (16 candidates) to VEGF range from nano- to micromolar levels. X-ray structure determination of one of the candidates showed atomic-level agreement with the design model. Moreover, the best designs showed the ability to inhibit proliferation of VEGF-dependent cells *in vitro* and *in vivo*. Thus, our results demonstrate the feasibility of the rational and generalizable approaches to design high-affinity protein binders against predefined conformational motifs.

1. Binder re-design (EGF inhibitors)



2. *De novo* binder design (VEGF inhibitors)



Here Be Coiled-Coil Dragons: Computational Search for New Hendecad Coiled Coils

Mikel Martinez-Goikoetxea, Andrei Lupas
Max Planck Institute for Biology, Tübingen, Germany
mikel.martinez@tuebingen.mpg.de

Repetition is a powerful mechanism to generate innovation in proteomes, which ultimately enables the adaptation of biological organisms to their environment. No other protein motif exemplifies this better than the coiled coil motif, which owing to its relatively short repeat size, high level of degeneracy and periodically reocurring interactions that could potentially extend indefinitely, can be found in virtually every proteome. This wealth of natural examples as well as their uniquely parameterizable structure have allowed for the development of robust prediction and design protocols for coiled coils, granting them the consideration of the best understood protein fold.

The vast majority of coiled coils in nature consist of heptad (7) repeats, but other repeat types such as hendecads (11) or pentadecads (15) are also compatible with coiled-coil structure. Typically, these non-canonical coiled-coil repeats, of which the most common is the hendecad, are found in low-copy number interspersed between heptads, where they introduce local changes in the packing of the side-chains as well as distortions in the protein backbone. In contrast, there are very few examples of coiled-coil sequences composed mainly or entirely of non-heptad repeats. This motivated us to search for new hendecads, in order to expand our coiled-coil prediction and design protocols even further.

Here, we present a search strategy based on tandem repeat detection which we have employed to obtain a reliable set of hendecad coiled coils. These new hendecad families illustrate the immense versatility of coiled coils, in the form of phage Tail Measure Proteins, membrane-bound bacterial proteins, fungal kinesins, and the largest known hendecad coiled coil which evolved from a cell-division protein broadly conserved in Gram-positive bacteria.

Improved repeat protein stability by computational protein design

Erich Michel[1]*, Stefano Cucuzza[2†], Peer R.E. Mittl[1], Oliver Zerbe[2], Andreas Plückthun[1]*

[1]Department of Biochemistry, University of Zurich, Zurich, Switzerland
[2]Department of Chemistry, University of Zurich, Switzerland
[†]Current address: Novartis Pharma AG, Basel, Switzerland
*Contact: e.michel@bioc.uzh.ch, plueckthun@bioc.uzh.ch

High protein stability is an important feature for proteins used as therapeutics, diagnostics and in basic research. We have previously employed consensus design to engineer optimized Armadillo repeat proteins (ArmRPs) for sequence-specific recognition of linear epitopes with a modular binding mode. These designed ArmRPs (dArmRPs) feature high stability and are composed of M-type internal repeats that are flanked by N- and C-terminal capping repeats which protect the hydrophobic core from solvent exposure.
While the overall stability of the designed ArmRPs is remarkably high, subsequent biochemical and biophysical experiments revealed that the N-capping repeat assumes a partially unfolded, solvent-accessible conformation for a small fraction of time that renders it vulnerable to proteolysis and aggregation. To overcome this problem, we have designed new N-caps starting from an M-type internal repeat using the Rosetta software. The superior stability of the computationally refined models was experimentally verified by CD and NMR spectroscopy. A crystal structure of a dArmRP containing the novel N-cap revealed that the enhanced stability correlates with an improved packing of this N-cap to the hydrophobic core of the dArmRP. Hydrogen exchange experiments further show that local unfolding of the N-cap is reduced by several orders of magnitude, resulting in increased resistance to proteolysis and reduced aggregation.

Abstract for Alpbach Coiled-coil 2022

Jessalyn Miller, Emory University (presenting work from a Chateaubriand Fellowship at I2BC)

Diverse assembly strategies for nanotubes from αRep-derived proteins

α-helical tandem repeat proteins derived from the Minard lab's αRep library have exhibited a wide variety of binding and assembling patterns. Here we present divergent strategies for assembling members of the αRep family into soluble filaments, as well as stability and mechanical insights. A method of co-assembling αRep nanotubes by "stapling" core units ("Bricks") together with a specially designed "staple" αRep has recently been demonstrated. Bricks that differ by a single repeat are strikingly different in stability and solubility of the assembled nanotubes. Assemblies with a high propensity for crystallization can be solubilized by randomizing the location of the protruding "staple" unit or by shortening the filament. Our observations indicate that αRep self-assembling nanotubes offer a rich potential diversity of form and function.

Signal Transduction in Photoreceptor Histidine Kinases

Andreas Möglich
Department of Biochemistry, University of Bayreuth, 95447, Bayreuth, Germany.
andreas.moeglich@uni-bayreuth.de

Sensory photoreceptors mediate sensation to light across all kingdoms of life. Light absorption by a chromophore, bound within the photosensor module of the receptor, elicits photochemical reactions that channel into changes in structure, dynamics, and function of a linked effector module. Prokaryotic photoreceptors frequently exert histidine-kinase activity and therefore belong to the two-component signaling systems (TCS). These photoreceptor histidine kinases are usually parallel oriented homodimers and feature alpha-helical coiled coils as linkers and conduits between photosensor and effector. To substantial extent, the photosensors and effectors can be exchanged for other input and output modules, respectively, which hints at general signal-transduction mechanism. As a case in point, a given histidine-kinase effector could be regulated by either a light-oxygen-voltage or a bacteriophytochrome sensor, thus establishing control by blue and red light, respectively, over TCS activity and bacterial gene expression. Deliberate modification of linker sequence and length provides insight into signal transduction and a means of constructing novel receptors with hitherto unavailable, yet highly desirable, traits.

The exo-helical symmetry of the α-Helix

J. Montenegro, J. Bergueiro, J. M. Martinez-Parra, R. Gomez-Ojea, G. A. Daudey, M. Calvelo, H. Fernández-Caro

[1]Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela; 15705, Spain. Javier.montenegro@usc.es

Abstract: The α-helix (3.6_{13}) is a peptide secondary structure used by nature in the construction of proteins.[1] The understanding of this architecture constituted a breakthrough that allowed us to understand the disposition of the chemical information of the proteins on a tridimensional space.[2] The lateral side chains of the α-helix define an unlimited number of topologies around the peptide backbone. These topologies are directly related to the biological interactions and the functional properties of numerous peptides and proteins. We have studied the potential exo-helical patterns defined by sequential repetition of amino acids residues and the resulting symmetries symmetries with for the peptide topology. We focus our efforts on the spectroscopic elucidation of the possible topologies generated by the amino acid repetition patterns in the peptide sequence. To achieve this goal, we used as model a fluorogenic non canonical amino acid incorporated into an α-helix biased skeleton.[3] Theoretical, computational, and spectroscopical elucidation of the principal exo-helical topologies of the α-helix is reported for model peptides modified with a non-canonical chromophore residue. The direct observation of these exo-helical topologies provides experimental evidence of their structural properties and construction variability. The topological information is essential for the understanding of peptide interactions and *de novo* design of synthetic surrogates.

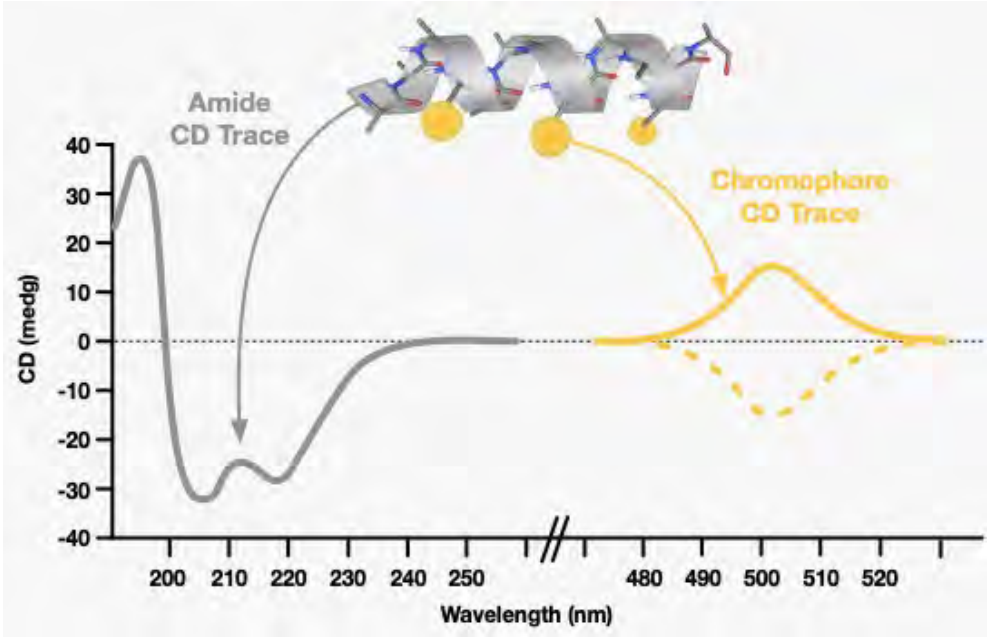


Figure. Fluorogenic non canonical amino acids allowed the experimental observation of the different exo-helical topologies defined by the α-helix.

References

[1] L. Pauling, R. B. Corey, H. R. Branson, *Proc. Nat. Acad. Sci*. **1951**, *37*, 205-211.
[2] A. L. Boyle, D. N. Woolfson, *Chem. Soc. Rev*. **2011**, *8*, 4295-4306.
[3] J. M. Finke, P. A. Jennings, et al. *Biopolymers*. **2007**, *86*, 193-211.

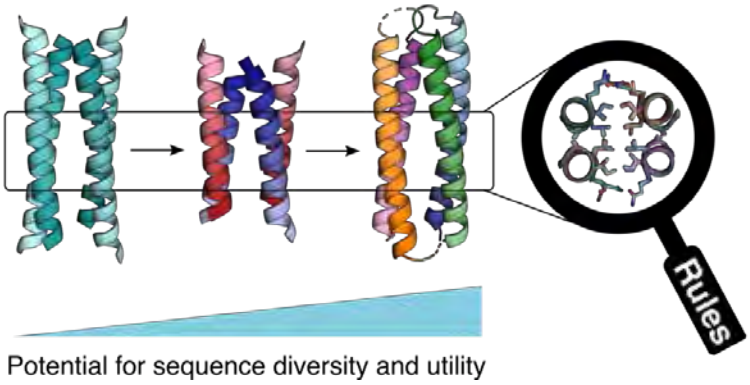
From peptides to proteins: coiled-coil tetramers to single-chain 4-helix bundles

Elise A. Naudin,[1] Katherine I. Albanese,[1] Abigail J. Smith,[2] Bram Mylemans,[1] Emily G. Baker,[1, 2] Nigel J. Savery,[2, 3] Derek N. Woolfson[1, 2, 3]

[1]School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK;
[2]School of Biochemistry, University of Bristol, University Walk, Bristol, BS8 1TD, UK;
[3]BrisSynBio, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK;

The design of completely synthetic proteins from first principles—*de novo* protein design—is challenging. This is because, despite recent advances in computational protein-structure prediction and design, we do not understand fully the sequence-to-structure relationships for protein folding, assembly, and stabilization.[1, 2] Antiparallel 4-helix bundles are amongst the most studied scaffolds for *de novo* protein design.[3, 4] We set out to re-examine this target, and to determine clear sequence-to-structure relationships, or design rules, for the structure. Our aim was to determine a common and robust sequence background for designing multiple *de novo* 4-helix bundles, which, in turn, could be used in chemical and synthetic biology to direct protein-protein interactions and as scaffolds for functional protein design.

Our approach starts by analyzing known antiparallel 4-helix coiled-coil structures to deduce design rules. These are implemented to deliver three new *de novo* structures: antiparallel homotetramer and heterotetramer, and a single-chain 4-helix bundle. All of the assembled designs are confirmed in aqueous solution using biophysical methods, and ultimately by determining high-resolution X-ray crystal structures. Our route from peptides to proteins provides an understanding of the role of each residue in each design.[5]



References:

1. Korendovych, I. V.; DeGrado, W. F., *Q. Rev. Biophys.* **2020,** *53*, e3.
2. Woolfson, D. N., *J. Mol. Biol.* **2021,** *433* (20), 167160.
3. Hill, R. B.; Raleigh, D. P.; Lombardi, A.; DeGrado, W. F., *Acc. Chem. Res.* **2000,** *33* (11), 745-754.
4. Lombardi, A.; Pirro, F.; Maglio, O.; Chino, M.; DeGrado, W. F., *Acc. Chem. Res.* **2019,** *52* (5), 1148-1159.
5. Naudin, E. A.; Albanese, K. I.; Smith, A. J.; Mylemans, B.; Baker, E. G.; Weiner, O. D.; Andrews, D. M.; Tigue, N.; Savery, N. J.; Woolfson, D. N., *bioRxiv* **2022**, 2022.08.04.502660.

De novo design of membrane coiled-coil barrels

Ai Niitsu[1], Andrew R Thomson[2], Alistair Scott[3], Jason T Sengel[4], Kozhinjampara R Mahendran[5], Yuji Sugita[1], R Leo Brady,[3] Mark I Wallace[4], Hagan Bayley[5], and Derek N Woolfson[3,6,7]

[1]Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Japan
[2]School of Chemistry, University of Glasgow, UK
[3]School of Chemistry, University of Bristol, UK
[4]Department of Chemistry, King's College London, UK.
[5]Department of Chemistry, University of Oxford, UK.
[6]School of Biochemistry, University of Bristol, UK
[7]Bristol BioDesign Institute, UK

Transmembrane α-helical coiled-coil barrels are fundamental and yet rare structures in known membrane proteins.[1,2] One way to access these structures is *de novo* protein design. However, the *de novo* design of membrane-spanning peptides and proteins that form autonomous barrels remains challenging. To address this challenge, building on our understanding of water-soluble and membrane coiled-coil barrels,[3,4] we design a series of transmembrane coiled-coil peptide barrels *de novo*. Through computational peptide design, structure modelling, and molecular dynamics simulations along with experimental structural analysis and electrical recordings, we have discovered and characterised coiled-coil peptide channels that conduct ions. They shed light on the sequence-to-structure/stoichiometry relationships of membrane coiled coils and their potential applications in engineering artificial ion channels.

1. A. Niitsu, J. W. Heal, K. Fauland, A. R. Thomson and D. N. Woolfson, *Phil. Trans. R. Soc. B*, 2017, **372**, 8-13.
2. K. R. Mahendran, A. Niitsu, L. Kong, A. R. Thomson, R. B. Sessions, D. N. Woolfson and H. Bayley, *Nat. Chem.*, 2017, **9**, 411-419.
3. A. R. Thomson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady and D. N. Woolfson, *Science*, 2014, **346**, 485-488.
4. A. J. Scott, A. Niitsu, H. T. Kratochvil, E. J. M. Lang, J. T. Sengel, W. M. Dawson, K. R. Mahendran, M. Mravic, A. R. Thomson, R. L. Brady, L. Liu, A. J. Mulholland, H. Bayley, W. F. DeGrado, M. I. Wallace and D. N. Woolfson, *Nat. Chem.*, 2021, **13**, 643-650.

Computational design of multi-state coiled coils using Molecular Dynamics simulations

Evangelia Notari[1*], Christopher W. Wood[2], Julien Michel[1]

[1]*EaStCHEM School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh EH9 3FJ, United Kingdom*

[2]*School of Biological Sciences, University of Edinburgh, C.H. Waddington Building, Max Born Crescent, Edinburgh EH9 3BF, United Kingdom*

*e.notari@sms.ed.ac.uk

Abstract

Proteins found in Nature have only sampled a small fraction of the available protein sequence space to solve problems faced during evolution. *De novo* protein design has so far mostly focused on design algorithms that maximise the free energy difference between a single conformational state and other plausible macrostates, resulting in rigid and hyper-stable folds that lack functionality. However, proteins often interconvert between different conformational states in order to execute their functions. Future progress in the field of *de novo* protein design depends on advances in designing multi-state proteins, i.e. proteins that can switch between different conformations upon an external perturbation.

Coiled coils are an attractive scaffold for *de novo* protein design, given that they are one of the few folds in Nature whose structure can be accurately described with parametric, simple mathematical equations. These clear sequence-to-structure rules can assist in the selection of sequences that can assemble into a desired oligomerisation state, e.g. dimer, trimer, tetramer etc. However, unexpected oligomerisation states can often be adopted upon external stimuli, e.g. a change in the pH [1].

In our work, we aim to develop a robust and reproducible computational pipeline for designing multi-state coiled coils. To this end, we are employing atomistic Molecular Dynamics simulations to capture the fine balance of intermolecular interactions that dictate the oligomerisation state of coiled coils. The use of enhanced sampling methods such as Metadynamics further enables the estimation of the free energy of aggregation of the designs, which can indicate whether specific oligomerisation states are more thermodynamically favourable over others.

References

[1] Rhys, Guto G., Wood, Christopher W., Beesley, Joseph L., Zaccai, Nathan R., Burton, Antony J., Brady, R. Leo, Thomson, Andrew R. and Woolfson, Derek N. (2019). *Journal of the American Chemical Society*, **141**(22), 8787-8797.

Designed repeat proteins: longer, faster and more creative

Fabio Parmeggiani

Schools of Chemistry and Biochemistry, Bristol BioDesign Institute
University of Bristol, UK

Repeat proteins, and in particular designed repeat proteins, have seen a progressive rise in interest and use in the last 20 years[1,2], as scaffolds for binders[3] or as parts of large protein architectures[4,5]. Despite advances in design, there are still limitations on the experimental level, especially for repetitive genes synthesis.
My group exploits the intrinsic modularity of repeat proteins[6–8] to develop methods to design and produce novel proteins in a faster and more reliable manner. In my talk I will describe our software tools for designing proteins from modular units (Elfin/ElfinUI[9,10]), our work on DNA assemblies of repetitive units and the design of novel interfaces for formation of repeat protein complexes.
The goal is to simplify and speed up design of complex protein architectures to be used as scaffolds for novel functional nanomaterials and nanoparticles.

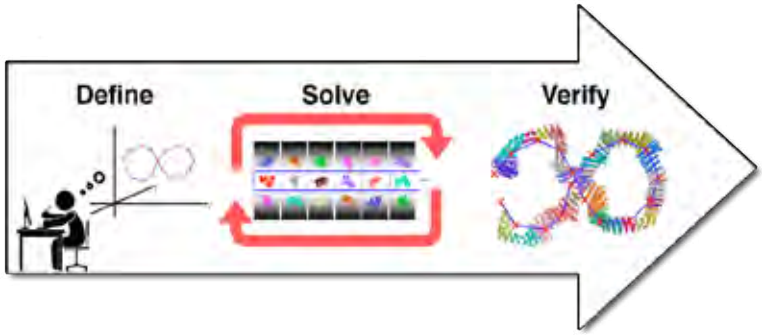


Figure: ELFIN, our modular approach to design of custom protein architectures

References

1. Parmeggiani, F. & Huang, P.-S. Designing repeat proteins: a modular approach to protein design. *Curr. Opin. Struct. Biol.* 45, 116–123 (2017). **2.** Gidley, F. & Parmeggiani, F. Repeat proteins: designing new shapes and functions for solenoid folds. *Curr. Opin. Struct. Biol.* 68, 208–214 (2021). **3.** Plückthun, A. Designed Ankyrin Repeat Proteins (DARPins): Binding Proteins for Research, Diagnostics, and Therapy. *Annu. Rev. Pharmacol. Toxicol.* 55, 489–511 (2015). **4.** Hsia, Y. *et al.* Design of multi-scale protein complexes by hierarchical building block fusion. *Nat. Commun.* 12, 2294 (2021). **5.** Divine, R. *et al.* Designed proteins assemble antibodies into modular nanocages. *Science* 372, eabd9994 (2021). **6.** Brunette, T. J. *et al.* Exploring the repeat protein universe through computational protein design. *Nature* 528, 580–584 (2015). **7.** Brunette, T. J. *et al.* Modular repeat protein sculpting using rigid helical junctions. *Proc. Natl. Acad. Sci.* 117, 8870–8875 (2020). **8.** Fallas, J. A. *et al.* Computational design of self-assembling cyclic protein homo-oligomers. *Nat. Chem.* 9, 353–360 (2017). **9.** Yeh, C.-T., Brunette, T., Baker, D., McIntosh-Smith, S. & Parmeggiani, F. Elfin: An algorithm for the computational design of custom three-dimensional structures from modular repeat protein building blocks. *J. Struct. Biol.* 201, 100–107 (2018). **10.** Yeh, C.-T., Obendorf, L. & Parmeggiani, F. Elfin UI: A Graphical Interface for Protein Design With Modular Building Blocks. *Front. Bioeng. Biotechnol.* 8, (2020).

Coiled coils as ligands for inclusion in the inorganic chemist's toolbox

Anna F. A. Peacock

[1] *School of Chemistry, University of Birmingham, Edgbaston, UK,* a.f.a.peacock@bham.ac.uk

Proteins are versatile and powerful ligands for metal ions, capable of achieving unusual coordination chemistries and therefore chemical properties. Often these features can be challenging to reproduce with small molecules ligands, and as such there is much interest in the development of robust miniature protein scaffolds as novel ligands for use within inorganic chemistry. That begs the question, might the metal ions not utilised by Nature, not also benefit from proteins as ligands?

With this in mind, work within our group has focused on the coordination of lanthanide ions to such a new class of ligands. Our ligands consist of a supercoil of helices, a coiled coil, into the hydrophobic core of which is engineered a binding site suitable for lanthanide ion coordination.[1] We have embarked on an investigation into how, with protein design, one can design lanthanide sites with a high degree of precision and control.[2,3,4] The design strategies employed will be presented and how these can be used to tune the desired coordination chemistry, and in turn the chemical properties (e.g. the use of gadolinium coiled coils as MRI contrast agents), will be presented. As will our work to develop multimetallo designs (see Figure),[5] featuring either the same, or distinct metal binding sites. Thereby presenting an opportunity to design multifunctional complex systems.

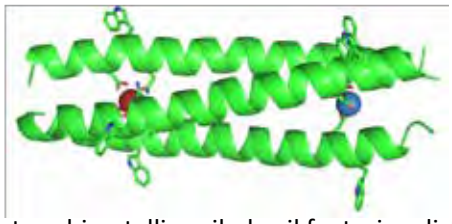


Fig. 1. *De novo* designed hetero bimetallic coiled coil featuring distinct lanthanide sites.

ACKNOWLEDGEMENTS

The authors thank the University of Birmingham, the School of Chemistry, the EPSRC and Horizon2020 for generous financial support.

REFERENCES

[1] Berwick, M. R., Lewis, D. J., Pikramenou, Z., Jones, A. W., Cooper, H. J., Wilkie, J., Britton, M. M., Peacock, A. F. A. (2014) De Novo Design of Ln(III) Coiled Coils for Imaging Applications, *J. Am. Chem. Soc.*, 136: 1166-9.

[2] Berwick, M. R., Slope, L. N., Smith, C., King, S. M., Newton, S. L., Gillis, R., Adams, G., Rowe, A., Harding, S., Britton, M. M., Peacock, A. F. A. (2016) Location dependent coordination chemistry and MRI relaxivity, in de novo designed lanthanide coiled coils, *Chem. Sci.*, 7: 2207-16.

[3] Slope, L. N.; Hill, M. G.; Smith, C. F.; Teare, P.; de Cogan, F. J.; Britton, M. M.; Peacock, A. F. A. (2020) Tuning Coordination Chemistry Through the Second Sphere in Designed Metallocoiled Coils, *Chem. Comm.*, 56:3729-32.

[4] Slope, L. N.; Daubney, O. J.; Campbell, H.; White, S. A.; Peacock, A. F. A., (2022) Location Dependent Lanthanide Selectivity Engineered into Structurally Characterized Designed Coiled Coils, *Angew. Chem. Int. Ed.*, 60: 24473-7.

[5] Teare, P., Smith, C. F., Adams, S. J., Sellamuthu, A., Ciani, B., Jeuken, L. J. C., Peacock, A. F. A. (2018) pH dependent binding in *de novo* heterobimetallic coiled coils, *Dalton Trans.,* 47: 10784-90.

Computational design of single-chain alpha-helical barrels proteins

Fabio Pirro[1*], Graham Leggett[2], Derek N. Woolfson[1, 3,4*]

[1]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK
[2]Department of Chemistry, University of Sheffield, Brook Hill, Sheffield S3 7HF, UK
[3]School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK
[4]BrisEngBio, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

*To whom correspondence should be addressed: fabio.pirro@bristol.ac.uk, D.N.Woolfson@bristol.ac.uk

Abstract: All the possible combination of protein sequences and structures is often referred as the *Protein Universe*.[1] In the last 10 years, *de novo* protein design has proved to be a powerful tool to investigate regions of this universe apparently unexplored by nature, the so-called *Dark Matter of Protein Space*.[2,3]

An example of such dark-matter proteins—or at least, of structures rarely used by nature—are the α-helical barrels (αHBs). These are usually cyclically symmetric coiled coils of five or more helices with central accessible channels.[4-7]Through empirical and computational studies, sequence-to-structure relationships for αHBs have been developed.[4,8,9] In turn, these have allowed access to oligomerization non found yet in nature,[6] and to the functionalization of αHBs for small-molecule binding,[10] catalysis,[11] and as membrane-spanning ion channels.[12-13] However, most of these examples of αHBs are oligomers of short peptides. To date, no αHBs, natural or designed from scratch, have been reported in which all the helices are part of the same chain.

Here, we report the development of a computational pipeline for the design of single-chain αHB proteins. The computed sequences and structures were screened *in silico* to maximize protein expression, minimizing either aggregation propensity or sequence repetitivity. The first generation of these αHB proteins have been produced with high levels of expression, they are monomeric in solution, α helical, hyperstable, and they function as small-molecule binders as effectively as their oligomeric peptide counterparts.

References:

1. Kolodny, Rachel, et al. *Annu Rev Biophys* 42.1 (2013): 559-82.
2. Woolfson, Derek N., et al. *Current opinion in structural biology* 33 (2015): 16-26.
3. Taylor, William R., et al. *Structure* 17.9 (2009): 1244-1252.
4. Zaccai, Nathan R., et al. *Nature chemical biology* 7.12 (2011): 935-941
5. Thomson, Andrew R., et al. *Science* 346.6208 (2014): 485-488
6. Liu, Jie, et al. *Proceedings of the National Academy of Sciences* 103.42 (2006): 15457-15462.
7. Malashkevich, Vladimir N., et al. *Science* 274.5288 (1996): 761-765.
8. Dawson, William M., et al. *Chemical science* 12.20 (2021): 6923-6928.
9. Rhys, Guto G., et al. *Journal of the American Chemical Society* 141.22 (2019): 8787-8797.
10. Thomas, Franziska, et al. " *ACS Synthetic Biology* 7.7 (2018): 1808-1816.
11. Burton, Antony J., et al. " *Nature chemistry* 8.9 (2016): 837-844.
12. Xu, Chunfu, et al. *Nature* 585.7823 (2020): 129-134
13. Scott, Alistair J., et al. *Nature chemistry* 13.7 (2021): 643-650

The quest for specific binding proteins

Andreas Plückthun, Dept of Biochemistry, University of Zurich, Switzerland
plueckthun.bioc.uzh.ch plueckthun@bioc.uzh.ch

Specific binding molecules play an ever-increasing role in biomedical research, clinical diagnostics and therapy. At the same time, their creation has been a key motivation to advance the development of technologies both in directed evolution and in structure-based protein design.

In the first example, the learnings from the immune system will be discussed. To recreate a synthetic system, a library of stably engineered scaffolds, containing constant and randomized residues where relevant, and a selection system of true evolution, i.e., capable of introducing "somatic" mutations, are all needed. After recreating antibody selection, this concept was transferred to Designed Ankyrin Repeat Proteins (DARPins), using the in vitro evolution system of Ribosome Display. DARPins are very stable, can be expressed in any cell compartment, and have been used in research, diagnostics and late-stage clinical trials.

In the second example, binding proteins are used to advance structural biology and drug design of recalcitrant drug targets. G protein-coupled receptors are an extremely important class of drug targets, but structure-based drug design has remained challenging because of the difficulty of obtaining high-resolution structures. Using special DARPins that favor crystal packing, in conjunction with directed evolution of the GPCRs themselves to stabilize them, several pairs of agonist/antagonist complexes could be structurally determined, clarifying the critical design features of agonists and antagonists.

In the third example, the paradigm of selection from large universal libraries to obtain binding proteins will be challenged. When it comes to linear epitopes, it should be possible to exploit the periodicity of peptide bonds and create a completely modular system, based on a binding protein design that shares the same periodicity. Using Armadillo Repeat Proteins, specific binding pockets for a range of different amino acid side chains have been created, and their correct assembly has been demonstrated. Using orthogonal approaches of design, selection, evolution, biophysical testing, and structure determination, significant progress has been reached toward the creation of a system of binding proteins that are modular and complementary to a given peptide sequence.

References are all available as PDFs, sorted according to topic or year:
https://plueckthun.bioc.uzh.ch/publications/

Archaeal receptor Af1503 – an ideal model for structural and functional studies of transmembrane signaling

Jakub Ptacek, Reinhard Albrecht, Stanisław Dunin-Horkawicz, Joachim E. Schultz, Murray Coles, Marcus D. Hartmann, Andrei N. Lupas

Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Max-Planck-Ring 5, D-72076 Tübingen, Germany

Transmembrane receptors allow cells to monitor their environment and initiate appropriate responses. The molecular mechanism of signal transduction is not yet fully understood, but the successful construction of functional chimeras shows that an entire class of receptors, comprising histidine kinases, chemoreceptors, adenylate cyclases and phosphatases, shares the same mechanism. One of the smallest known receptors to date, Af1503 from *Archaeoglobus fulgidus,* has proven particularly suitable for modular design and structural studies. Its HAMP domain was the first to be solved at high resolution and led to a model for signal transduction based on axial helix rotation. However, further progress was stalled by our inability to solve the full-length structure of the protein, until AlphaFold2 provided us with a sufficiently good model for molecular replacement. In contrast to other receptors from the same class, Af1503 lacks a C-terminal effector domain, making it difficult to identify its natural ligand and establish its functional role. To address this problem, we engineered a functional chimera, consisting of Af1503 fused to the EnvZ histidine kinase domain. This chimera is well folded and fully functional *in vivo*, enabling us to determine its substrate specificity. In conjunction with crystallography studies, we could establish that the native ligands of Af1503 are fatty acids of 12 to 18-carbon atom chains, opening the way to study the mechanism of signal transduction at atomic resolution.

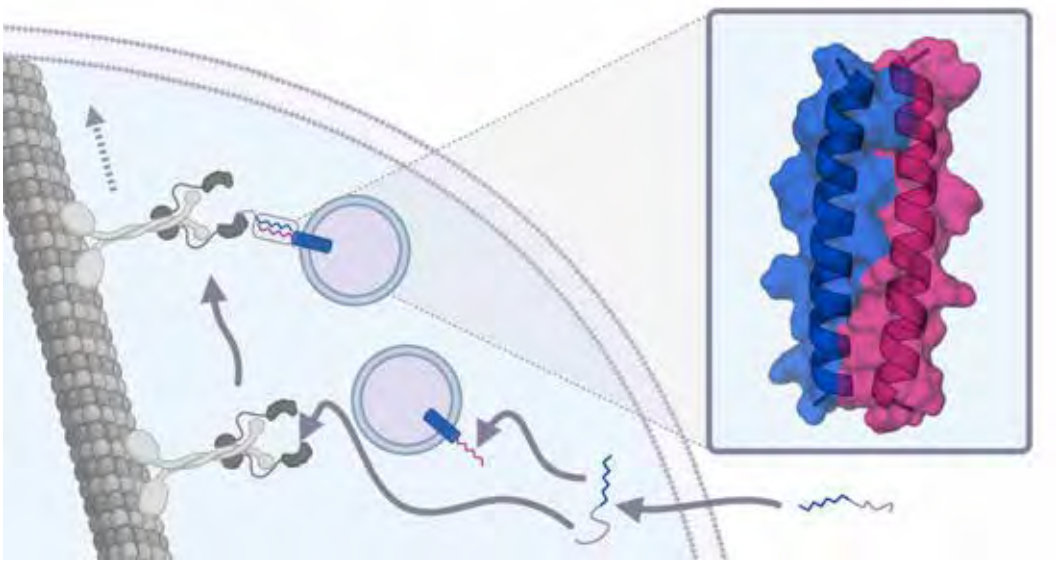
De novo designed peptides for cellular delivery and subcellular localisation

Guto G. Rhys[1,2,†], Jessica A. Cross[3,4,†], William M. Dawson[3,†], Harry F. Thompson[3,4], Sooruban Shanmugaratnam[1], Nigel J. Savery[4,5], Mark P. Dodding[4], Birte Höcker[1,*], and Derek N. Woolfson[3,4,5,*]

[1]Department of Biochemistry, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany
[2]School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff, CF10 3AT, United Kingdom
[3]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom
[4]School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, United Kingdom
[5]BrisSynBio, Life Sciences Building, University of Bristol, Bristol, BS8 1TQ, United Kingdom

[†]These authors contributed equally.
[*]Corresponding authors.

Abstract
Increasingly, it is possible to design peptide and protein assemblies *de novo* from first principles or computationally. This provides new routes to functional synthetic polypeptides, including designs to target and bind proteins of interest. Much of this work has been developed *in vitro*. Therefore, a challenge is to deliver *de novo* polypeptides efficiently to sites of action within cells. Here, we describe the design, characterization, intracellular delivery, and subcellular localisation of a *de novo* synthetic peptide system. This comprises a dual-function basic peptide, programmed both for cell penetration and target binding, and a complementary acidic peptide that can be fused to proteins of interest and introduced into cells using synthetic DNA. The designs are characterized *in vitro* using biophysical methods and X-ray crystallography. The utility of the system for delivery into mammalian cells and subcellular targeting is demonstrated by marking organelles and actively engaging functional protein complexes.



Created with BioRender.com

Protein binders for functionalization of proteins with coiled-coils

Satler Tadej[1], Jerala Roman[1,2]

[1] Department of Synthetic Biology and Immunology, National Institute of Chemistry, Ljubljana, Slovenia. [2] EN-FIST Centre of Excellence, Ljubljana, Slovenia.

Coiled-coil is a structural motif comprised of two or more intertwined α-helices that associate through hydrophobic and electrostatic interactions. They are abundant among natural and *de novo* proteins, making them an attractive target for protein functionalization.

Here we present a system for grafting binding sites of small protein binders to the coiled-coil structural motif. Modified residues were mostly surface residues (positions b, c and f) of coiled-coils that do not interfere with the coiled-coil formation. The presented protein binders can be easily produced in bacteria and used for many different applications, including facilitating structural characterization of small proteins with coiled-coils.

We scanned the PDB database for protein complexes where small protein binders preferably interacted with a single α-helix of their protein target. Binding epitopes of several selected binders were then manually or computationally grafted to the coiled-coil structural motif and experimentally tested with NativePAGE and ITC. Successful binding to the coiled-coil dimer was achieved with one nanobody Nb49 (source PDB ID: 7A48) and one Intrabody Ib3 (source PDB ID: 3LRH), with the latter having higher affinity, effortless grafting and being less disruptive. Binding epitopes were also successfully transferred to the proteins, using trigonal protein from coiled-coils as a model. With a slight loss of affinity, Ib3 epitope grafting allowed us to modify only the surface residues of the coiled-coils, which are generally not involved in coiled-coil interaction and therefore do not disturb the specificity of coiled-coil pairing. Additionally, to show the modularity of Ib3 system, we grafted two binding epitopes to the same coiled-coil motif. Experiments confirmed the binding of Ib3 to both sites, with one being preferred over the other.

The presented system offers robust, modular, and effortless epitope grafting and coiled-coil functionalization, making it attractive tool for medical, biotechnological and research applications.

Cryptic coiled-coil forming potential of ribosomal protein fragments

A. Schramm, F. Springer, C. Oberist, J. Wilhelm, J. Martin, M.D. Hartmann, A.N. Lupas
Department of Protein Evolution, Max Planck Institute for Biology, 72076 Tübingen, Germany

Repetition of building blocks, from single amino acids to entire domains, is a central mechanism in the emergence of new proteins and is likely to have played a critical role in the emergence of the first folded proteins during the transition from the RNA-peptide to the DNA-protein world. In order to explore the ability of polypeptides that are structured on an RNA scaffold to achieve autonomous folding, independently of the scaffold, we studied the potential of ribosomal protein fragments to form coiled-coils. For this, we used the ribosomal proteins of *Thermus thermophilus* as a source of such fragments and chose six that showed an elevated coiled-coil propensity in computational predictions. We took two approaches in parallel: we inserted the fragments between GCN4 N16V adaptors in order to establish their basic compatibility with coiled-coil structure, and we amplified them by repetition in order to explore whether they could yield autonomously folded forms. GCN4 N16V is a leucine zipper that can equally form dimers and trimers, thus not obscuring the oligomeric preference of the insert. Fragments inserted between such adaptors yielded highly stable oligomeric species with a high content of α-helical structure, mostly trimeric, but, in the case of a fragment from the L29 protein, dimeric. These results showed that the fragments we chose were fundamentally compatible with the coiled-coil fold. Concurrently, we investigated their ability to fold autonomously upon repetition. Of the six fragments, two (from L10 and L29) were able to fold autonomously upon duplication. In agreement with our previous results, the duplicated L10 fragment formed trimers in solution and the duplicated L29 fragment dimers. Since the native L29 protein forms a monomeric, aniparallel helical hairpin with local knobs-into-holes packing in the context of the ribosome, the ability of an L29 fragment to form parallel dimers is maybe not as surprising as the ability of the L10 fragment to form a parallel trimer, substantially different from its structure in the ribosome. These results show that there is indeed cryptic potential to form new structures in ribosomal proteins.

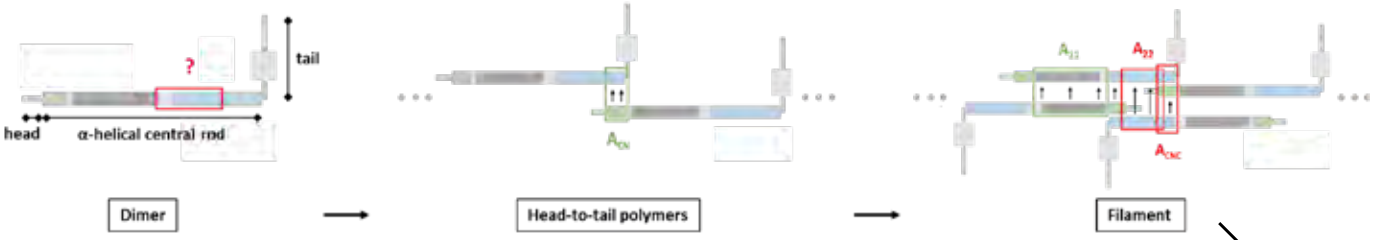
Giel Stalmans

Abstract
Alpbach Workshop 'Coiled-coil, Fibrous & Repeat proteins'
September 4-9, Alpbach, Austria, 2022

The integrative structural study of nuclear lamin assembly.

Lamin proteins (A, B1, B2 and C) are part of the intermediate filament (IF) family just like the well-known keratins. They are , however, the only IF proteins found within the nucleus of cells. Here, they play a vital role in normal cell functioning by not only providing mechanical stability but also being involved normal mitosis, chromatin organization and transcription, DNA replication and repair, etc.

At molecular level, lamins form 3.5 nm thick filaments which arise in a defined, hierarchal self-assembly manner (see scheme 1). It starts with an elementary **coiled-coil building block**, a dimer of two parallel monomers (each 646 AA). In more detail, the latter consist of an unstructured head domain, an α-helical rod domain and a (partially) unstructured tail domain. Next, there is a longitudinal interaction, called A_{CN}, that leads to head-to-tail polymers. In turn, those polymers assemble in a half-staggered, antiparallel fashion, called A_{11} and A_{22}, to form mature filaments. Eventually, in specific conditions, these filaments can somehow assemble laterally in paracrystals. Unfortunately, it remains to a great extent unknown how these filaments and paracrystals are actually formed at molecular level.



Scheme 1: lamin filament assembly pathway.

Hence, our lab started an *in vitro* integrative structural approach in which X-ray crystallography, chemical cross-linking mass spectrometry and electron microscopy are complementary used to solve the lamin assembly. Together those two former techniques have already allowed us to build in the **first molecular A_{CN} interaction model** (see figure 1 below). This model has been presented in a publication '*Stalmans G, et al. Addressing the Molecular Mechanism of Longitudinal Lamin Assembly Using Chimeric Fusions. Cells. 2020; 9(7)*)'. More recently, I solved **two more novel X-ray crystallographic structures** (unpublished). In general, these novel structures reveal the last missing part of the dimeric building block at atomic level. Furthermore, they provide additional information about filament assembly as well.

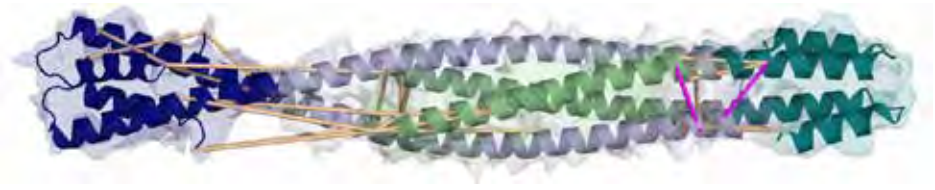
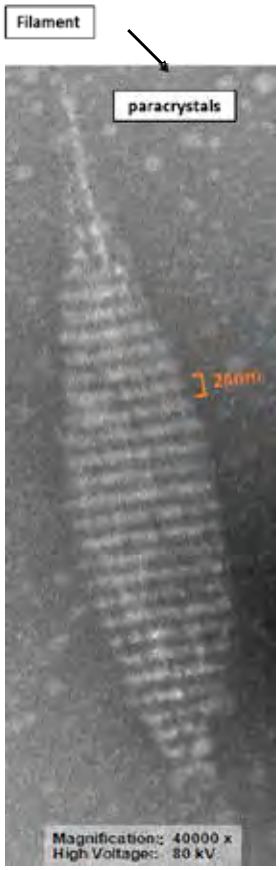


Figure 1: lamin A_{CN} interaction model based on two individual crystal structures and 31 cross-links between both capped fragments.

In parallel, investigation of the central α-helical rod domain of vimentin, a cytoplasmic intermediate filament, using circular dichroism combined with SEC-MALS provided a detailed stability study of the different central α-helical rod domain regions with their specific coiled-coil characteristics (published soon).



Structure, function, and application of a model bacterial phytochrome

Heikki Takala

Department of Biological and Environmental Science, Nanoscience Center,
University of Jyväskylä, Finland
Email: heikki.p.takala@jyu.fi

Phytochrome photoreceptors sense red and far-red light and cycle between the spectroscopically, structurally, and functionally distinct Pr and Pfr states. Upon light activation, the structural changes of the photosensory module of phytochrome are relayed to the enzymatic output module, which occurs through coiled coil-forming signaling helices. Bacterial phytochromes commonly belong to two-component systems that transmit environmental stimuli to a response regulator (RR) protein through histidine kinase (HK) activity.

We have recently revealed with cryo-EM that minor structural changes in the photosensory module of the paradigm phytochrome from *Deinococcus radiodurans* (*Dr*BphP) cause a large change in its output HK module [1]. We have also demonstrated biochemically that *Dr*BphP exclusively acts as a phosphatase, but that its photosensory module can control HK activity of homologous receptors [2]. This knowledge has been applied to generate optogenetic tools, called pREDusk and pREDawn, that allow over 100-fold regulation of gene expression under red light [3]. The triggering by red light enables high spatial resolution, safe light intensities, and superior tissue penetration compared to blue light-responsive systems.

[1] Wahlgren W, et al. (2022), *under review*
[2] Multamäki, et al. (2021) *Nat Commun* 12, 4394. https://doi.org/10.1038/s41467-021-24676-7
[3] Multamäki E, et al. (2022) JMB-D-22-00464, Available at *SSRN*: https://ssrn.com/abstract=4108992.

Computational *de novo* design of phospho-responsive molecular switches

Sophia Tan, William DeGrado
University of California, San Francisco

Molecular switches enable detection and control over a wide range of cellular behavior. Though there have been numerous successes in designing switches that respond to ligand binding or photoactivation, there are fewer examples of *de novo* switches that respond to the addition or removal of post-translational modifications (PTMs). In this project, we attempt computational design of molecular switches that are activated by kinase phosphorylation. Our designed switches comprise peptides that self-assemble into coiled coils when phosphorylated, but remain monomeric in an unstructured state when unphosphorylated.

We are using these molecular switches to design kinase reporters in collaboration with the Shu Lab at UCSF; by grafting a kinase substrate sequence onto our molecular switch peptides, we can induce phospho-dependent high-order oligomerization that will condense into a phase-separated readout (e.g. GFP-dense liquid droplet formation). This technique will facilitate imaging studies that track spatiotemporal *in vivo* kinase activity under various conditions.

Our designed molecular switches are inspired by seminal studies in the elucidation of stability determinants of coiled coils, and these switches also build on previous work from our lab whereby a designed peptide switches from an unstructured monomer to a structured 4-helix bundle upon Protein Kinase A (PKA) phosphorylation. This current study relies on design principles set forth in the previous work, but we now have the advantage of using structural bioinformatics to improve interhelical interactions and make the designs more sensitive. We're excited to ultimately apply this computational methodology to design highly specific orthogonal switches for other kinases, and to even expand to other PTMs.

The WW domain as a model system for the design of miniaturized proteins

Franziska Thomas

Institute of Organic Chemistry, Heidelberg University
Im Neuenheimer Feld 270, 69120 Heidelberg, Germany, franziska.thomas@oci.uni-heidelberg.de

De novo designed biomolecules are traditionally used as models to mimic natural systems or to create new functional assemblies that operate under physiological conditions. In this context, small protein folding motifs are highly interesting because, ideally, the function of each amino acid residue is understood and the effects of changes on structure and/or function can be readily determined. Because of the well-understood sequence-structure relationships, the coiled-coil motif is probably most commonly used in synthetic biology efforts;[1] however, other protein folding motifs are also becoming more widely used.

Our current research focus is the WW domain, a small β-sheet motif that recognizes proline-rich peptide sequences and is known for its high sequence capacity. Our goal is to install new functions by rational design or combinatorial approaches. For example, we use split-WW domains bound to an antiparallel heterodimeric coiled-coil for reconstitution of folding (Figure 1). Based on this approach, split-WW domain libraries are readily accessible in which the fragments of the different library members can be combined following a split-and-mixing procedure . In this way, we identified a WW domain that binds ATP in the lower µM regime.[2]

In addition to this combinatorial approach, we use peptide engineering to modulate the function of WW domains. For example, using human carbonic anhydrase II as a natural model, we installed a Zn(II) coordination site on the surface of a WW domain (WW-CA). WW-CA is specific for Zn(II), exhibits Zn(II)- and pH-dependent switchability, and is, to our knowledge, the first WW domain scaffold with this type of binding activity.[3] Such systems can be advanced to develop peptide sensors.[4] In long-term studies, we aim to design a basic scaffold on which multiple functional sites could be imprinted, based on the analysis of sequence-structure relationships.

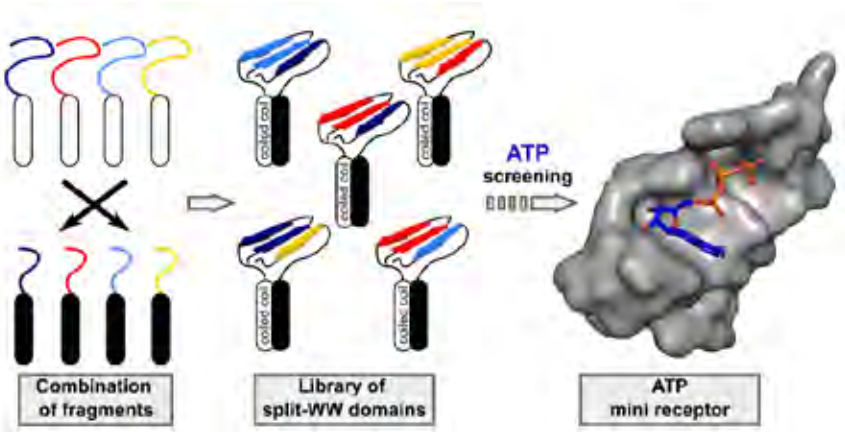


Figure 1. Concept of a combinatorial approach towards WW domains with new function.

References:

[1] W. M. Rink, F. Thomas, *Chem. Eur. J.* **2019**, *25*, 1665–1677. [2] H. Neitz, N. B. Paul, F. R. Häge, C. Lindner, R. Graebner, M. Kovermann, F. Thomas, *Chem. Sci.* **2022**, DOI: 10.1039/D2SC01078J. [3] T. L. Pham, M. Kovermann, F. Thomas, *ACS Synth. Biol.* **2022**, 254-264. [4] M. Werner, J. Pampel, T. L. Pham*, F. Thomas, *Chem. Eur. J.* **2022**, DOI: 10.1002/chem.202201339.

Computational design of granulopoietic proteins

Timo Ullrich[1], Julia Skokowa[2], Andrei Lupas[1] and Mohammad ElGamacy[1,2],
[1] Max-Planck-Institute of Biology, Department Protein Evolution, Tübingen
[2] Translational oncology, Internal Medicine II, University Hospital Tübingen

De novo protein design has evolved into a powerful tool providing a new generation of potential therapeutics. Recently, we have designed novel proteins which act as granulocyte-colony stimulating factor receptor (G-CSFR) agonists utilizing a topological rescaffolding (1) and refactoring strategy (2). Unlike the native receptor ligand (G-CSF), which represents an important immunotherapeutic, these proteins are small, highly stable, and can be produced with high yields in *Escherichia coli*. In order to optimize such promising molecules, we developed a pipeline that is able to quickly and efficiently affinity-maturate *de novo* designed proteins. In essence, a bacterial display is used in combination with computationally-designed libraries, utilizing the new design software Damietta. Affinity enhanced variants were investigated as tandem fusions in a cell-based assay, from which the most active variant nearly reached the activity of recombinant human G-CSF itself. Additionally, high affinity variants were utilized to design competitive G-CSFR antagonists which can serve as a protein-based drug to treat G-CSF dependent types of cancer.

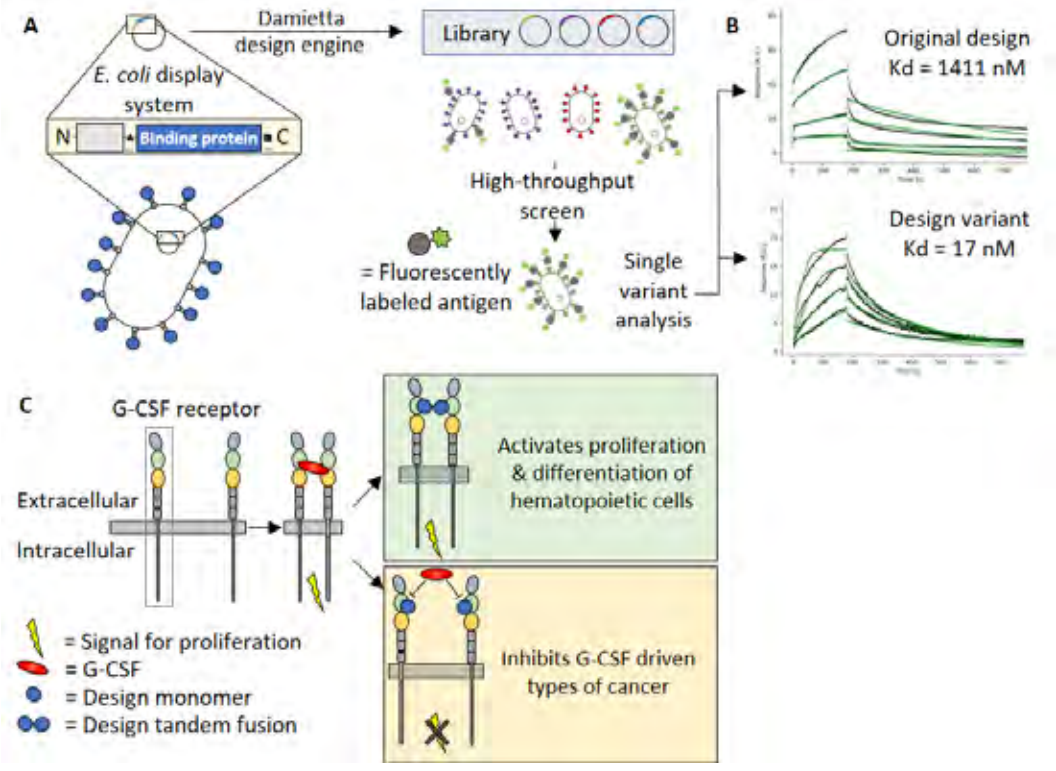


Figure 1: A) Affinity maturation pipeline, B) surface plasmon resonance of an original granulopoietic protein design and an affinity enhanced variant, C) granulocyte-colony stimulating factor (G-CSF) receptor activation and respectively mode of action of granulopoietic protein designs as monomers or tandem fusions

(1) Hernandez Alvarez, Birte, et al. "Design of novel granulopoietic proteins by topological rescaffolding." *PLoS biology* 18.12 (2020): e3000919.
(2) Skokowa, Julia, et al. "A topological refactoring design strategy yields highly stable granulopoietic proteins." *Nature communications* 13.1 (2022): 1-17.

AlphaReps artificial repeat proteins as tools for modular protein assemblies

A. Urvoas, V. Campanacci, B. Gigant, S. Thébault, B. Delmas, T. Di Meo, K. Kariyawasam, R. Ricoux, J.P. Mahy, M. Valerio-Lepiniec, P. Minard

Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.

A family of artificial repeat proteins, named αRep, was designed from the concatenation of a 31-residue consensus helical motif with 5 hypervariable positions. αReps from the highly diverse library (10^9 variants) present a concave diversified binding surface and a variable number of repeats. αRep proteins are easily produced and highly stable[1]. Using phage display methods, it is possible to select protein binders with high selectivity and affinity for a variety of protein targets (K_D from nM to µM). We have explored αRep applications in the fields of structural and cellular biology[2,3], as crystallization helpers and intracellular as tracers in living cells but also, in the fields of chemistry and physics with the development of biohybrid artificial metalloenzymes[4,5], generic biosensors[6] and bionanomaterials[7]. Recently, αReps used as crystallization helpers for tubulin[2,3] provided structural insights for microtubule nucleation process; αRep nanoligands targeting the Spike protein neutralize SARS-CoV2 variants[8] and reduce infection severity in hamster; coupling a Manganese porphyrin complex into a bidomain αRep scaffold resulted into an artificial metalloenzyme with inducible oxidase activity[5]. As compared with molecular Lego® bricks, αReps can further be exploited to engineer modular protein assemblies.

1. Urvoas A, Guellouz A, Valerio-Lepiniec M, Graille M, Durand D, Desravines DC, van Tilbeurgh H, Desmadril M, Minard P. (2010) Design, production and molecular structure of a new family of artificial alpha-helicoidal repeat proteins (αRep) based on thermostable HEAT-like repeats. *J Mol Biol*. 404, 307-27.
2. Campanacci V, Urvoas A, Cantos-Fernandes S, Aumont-Nicaise M, Arteni AA, Velours C, Valerio-Lepiniec M, Dreier B, Plückthun A, Pilon A, Poüs C, Minard P, Gigant B. (2019) Insight into microtubule nucleation from tubulin-capping proteins. *Proc Natl Acad Sci U S A* 116(20):9859-9864.
3. Campanacci V, Urvoas A, Ammar-Khodja L, Aumont-Nicaise M, Noiray M, Lachkar S, Curmi PA, Minard P Gigant B (2022) Structural convergence for tubulin binding of CPAP and vinca domain microtubule inhibitors, *Proc Natl Acad Sci U S A* 119 (19) e2120098119
4. Di Meo T, Kariyawasam K, Ghattas W, Valerio-Lepiniec M, Sciortino G, Maréchal JD, Minard P, Mahy JP, Urvoas A, Ricoux R. (2019) Functionalized artificial bi-domain proteins based on an α-solenoid protein repeat scaffold: a new class of artificial Diels-Alderases. *ACS omega* 4, 4437-4447.
5. Kariyawasam K, Di Meo T, Hammerer F, Valerio-Lepiniec M, Sciortino G, Maréchal JD, Minard P, Mahy JP, Urvoas A and Ricoux R. (2020) An artificial hemoprotein with inducible peroxidase- and monooxygenase-like activities. *Chem. Eur. J.* 26, 1-10.
6. Léger C, Yahia-Ammar A, Susumu K, Medintz I, Urvoas A, Valerio-Lepiniec M, Minard P, Hildebrandt N. (2020) Picomolar Biosensing and Conformational Analysis Using Artificial Bidomain Proteins and Terbium-to-Quantum Dot FRET. *ACS Nano* 14(5):5956-5967.
7. Prasad J, Viollet S, Gurunatha KL, Urvoas A, Fournier AC, Valerio-Lepiniec M, Marcelot C, Baris B, Minard P, Dujardin E. (2019) Directed evolution of artificial repeat proteins as habit modifiers for the morphosynthesis of (111)-terminated gold nanocrystals. *Nanoscale*. 11(37):17485-17497
8. Thébault S, Lejal N, Dogliani A, Urvoas A, Valerio-Lepiniec M, Lavie M, da Costa B, Bourgon C, Donchet A, Fraysse A, Saint-Albin-Deliot A, Klonjkowski B, Dubuisson J, Roussel A, Minard P, Le Poder S, Meunier N, Delmas B. (2022) Alpha-helicoidal HEAT-like Repeat Proteins (αRep) neutralize SARS-CoV-2 variants and reduce infection severity in a hamster model. Accepted in *PLOS Pathogens*

Coiled Coil Molecular Force Sensors for Measuring Cellular Forces & Attachment

Russell J. Wilson[a,b], Melis Goktas[b], Zeynep Atris[b], Kerstin Blank[a,b]

a Institute of Experimental Physics, Johannes Kepler University Linz, Linz, Austria
b Max Planck Institute of Colloids and Interfaces, Potsdam, Germany

Molecular force sensors (MFSs) are small molecular probes used to observe piconewton forces and have proven an effective tool to investigate biomechanical phenomena. Unfortunately, the current MFS technology (dsDNA) has limited and expensive functionalization strategies and a maximal rupture force plateau at 65 pN. Protein structural motifs (such as the coiled coil) provide a plethora of options for functionalization and have unique mechanical behaviors that can be exploited to make MFSs. We have designed a pair of heterodimeric coiled coil-based molecular forces sensors with different force loading geometries (terminal and central). The coiled coil MFSs were characterized with AFM-based single molecule force spectroscopy (SMFS) and showed different rupture force/loading rate profiles depending on loading geometry. The mechanically calibrated coiled coils (rupture forces= 20-50 pN) were utilized to fabricate mechanoresponsive surfaces that discriminate forces transmitted across cell integrin-ligand attachments. As a proof of concept, fibroblast cells were adhered to surfaces containing RGDS-functionalized coiled coil sensors or covalently coupled RGDS (control). We show that all tested force sensors have initial attachment within 30 min after cell seeding. At time points between 60-120 min, distinct cell spreading behavior is observed for coiled coil MFSs with different thermodynamic and mechanical stabilities. These results aid the future design of novel coiled coil-based MFSs with greater functionality such as multiplexing ligands, higher order assembly (3-5) and/or fusion proteins.

From rational to computational *de novo* design of coiled-coil assemblies and back again

Dek Woolfson
Schools of Chemistry and Biochemistry, & Bristol BioDesign Institute
University of Bristol, UK

Protein design—*i.e.*, the construction of entirely new protein sequences that fold into prescribed structures—has come of age: it is now possible to generate many stable protein folds from scratch using rational and/or computational approaches.[1] One challenge for the field is to move past natural protein structures and target the so-called 'dark matter of protein space'; that is, protein structures that should be possible in terms of chemistry and physics, but which biology seems to have overlooked or not used prolifically. This talk will illustrate what is currently possible in this nascent field using *de novo* helical peptides as building blocks.[2]

First, my talk will survey our understanding of,[2] design methods for,[3,4] our current "toolkit" of *de novo* α-helical coiled coils, Figure.[5,6] This includes dark-matter protein structures such as α-helical barrels, Figure (middle).[7] Briefly, I will mention how these barrels can be adapted to make new nanotube materials,[8] rudimentary catalysts,[9] membrane-spanning pores,[10] components of a new types of sensing devices,[11] and proteins that switch conformational state.[12] Finally, I will describe the rational design, construction, and characterisation of a completely new coiled-coil-like assembly made from unprecedented, long, 3_{10}-helical peptides.[13]

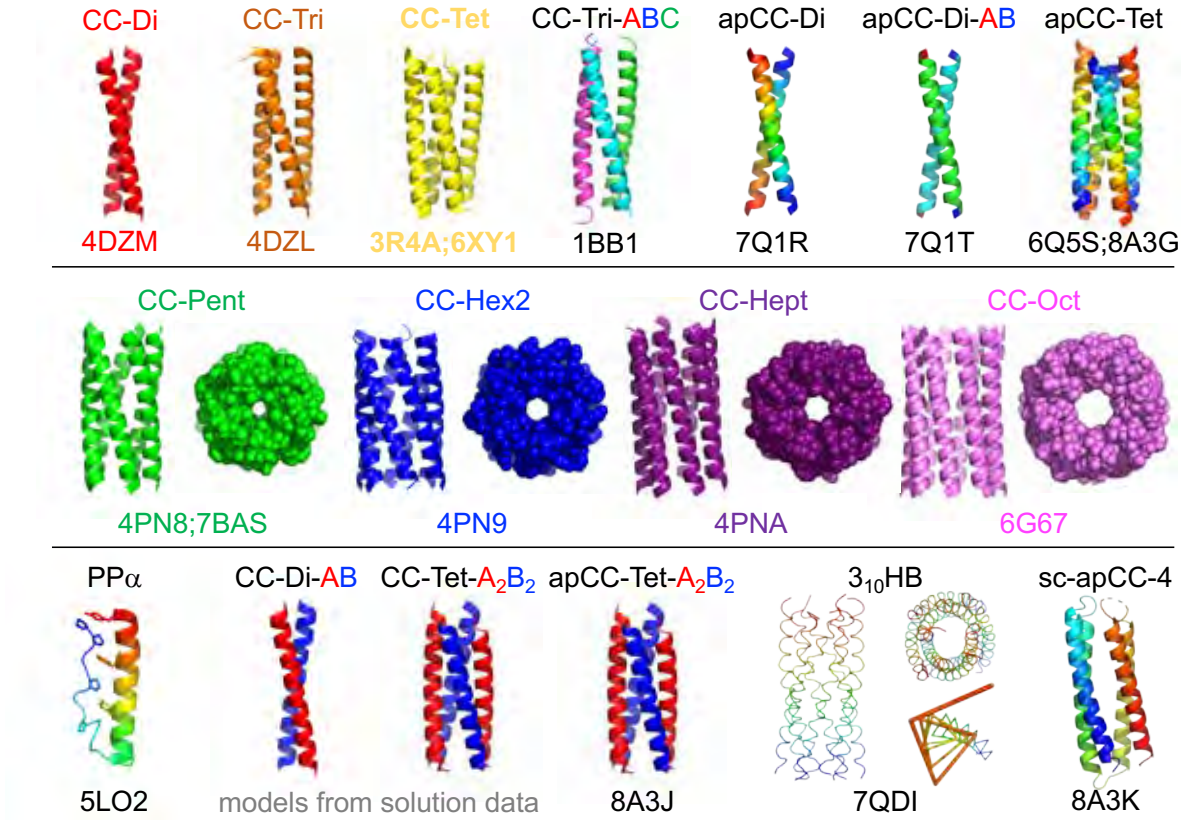


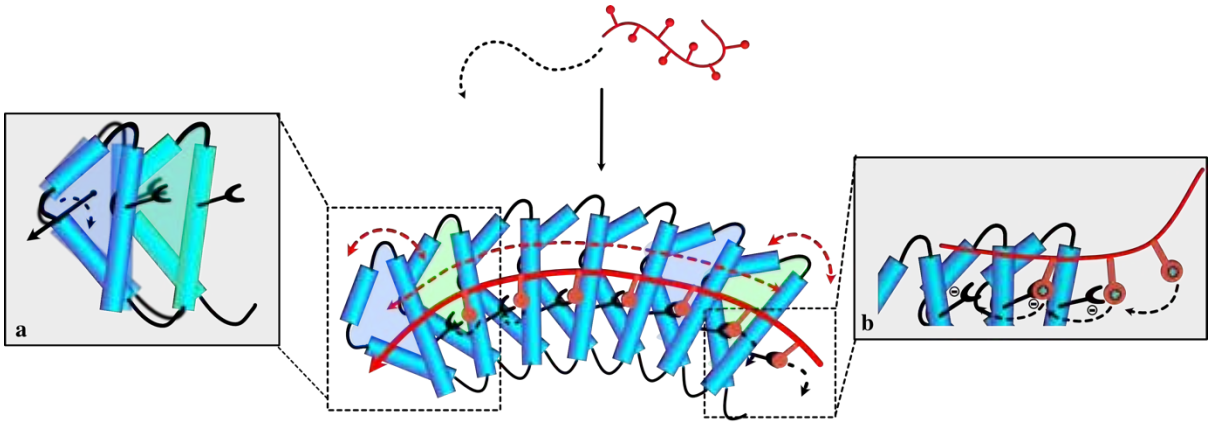
Figure: Cartoons of the structures for the de novo coiled-coil toolkit with systematic names and PDB codes.

References: 1. A brief history of *de novo* protein design: minimal, rational, and computational. DN Woolfson. ***J Mol Biol* 433**, ARTN: 167160 (2021). **2.** Coiled-coil design: updated and upgraded. DN Woolfson. ***Subcellular Biochemistry 82***, 35-61 (2017). **3.** CCBuilder: an interactive web-based tool for building, designing and assessing coiled-coil-protein assemblies. CW Wood *et al.* ***Bioinformatics 30***, 3029-3035 (2014). **4.** ISAMBARD: an open-source computational environment for biomolecular analysis, modelling and design. CW Wood *et al.* ***Bioinformatics 33***, 3043–3050 (2017). **5.** A basis set of *de novo* coiled-coil peptide oligomers for rational protein design and synthetic biology. JM Fletcher *et al.* ***ACS Synth Biol 1***, 240-250 (2012). **6.** A set of *de novo* designed parallel heterodimeric coiled coils with quantified dissociation constants in the micromolar to sub-nanomolar regime. F Thomas *et al.* ***J Am Chem Soc 135***, 5161-5166 (2013). **7.** Computational design of water-soluble α–helical barrels. AR Thomson *et al.*, ***Science 346***, 485-488 (2014). **8.** Modular design of self-assembling peptide-based nanotubes. NC Burgess *et al.*, ***J Am Chem Soc 137,*** 10554-10562 (2015). **9.** Installing hydrolytic activity into a completely *de novo* protein framework. AJ Burton *et al.*, ***Nature Chemistry 8***, 837-844 (2016). **10.** A monodisperse transmembrane α-helical peptide barrel. AJ Scott *et al.*, ***Nature Chemistry 13***, 643-50 (2021). **11.** *De novo*-designed α-helical barrels as receptors for small molecules. F Thomas *et al.*, ***ACS Synth Biol 7***, 1808-1816 (2018). **12.** Structural resolution of switchable states of a *de novo* peptide assembly. WM Dawson WM *et al.* ***Nature Communications 12***, ARTN:1530 (2021). **13.** *De novo* design of discrete, stable 3_{10}-helix peptide assemblies. P Kumar *et al.* ***Nature* 607**, 387-392 (2022).

Fast dynamics in complexes of picomolar binders

Stefano Cucuzza, Malgorzata Sitnik, Simon Jurt, Daniel Häussinger, Andreas Plückthun, Oliver Zerbe*

Peptides solely built from the (KR) dipeptide sequence bind with very high affinity to designed Armadillo repeat proteins of the general format NM_nC. Therein, N and C are the terminal caps, and n sequence-identical internal repeats M. Each internal module can bind one KR unit. Here, we investigated $(KR)_n$ peptides with n=4-7 binding to NM_4C and NM_7C. In particular I will focus on two features of the binding reaction: i) the structural response of the protein to peptide binding, and ii) the dynamics of the peptide in the ArmRP-bound state. To describe the protein structure in its free and complexed form we have used pseudocontact shifts (PCS) to generate a structure of high precision based on backbone data only. Given the sequence identity of internal repeats we have resolved the enormous challenge of obtaining sequence-specific assignments by segmental isotope labeling. The structures, that in the unliganded state display significant variations in supercoil between the sequential repeats, are generally regularized so that the binding interface now is very similar for every dipeptide unit:



Interestingly, we observed broad lines of protein resonances at the peptide-binding interface and of loop residues that connect individual modules in the peptide complex even for binders with sub-picomolar affinities. We have extracted rate constants for overall binding and unbinding reactions as well as for internal dynamics of the bound state from titrations followed by heteronuclear NMR. We believe that such dynamics will escape in an analysis using other biophysical methods commonly applied to derive dissociation constants. I will discuss implications of our results and also differences in conformational state(s) of these proteins when derived from solution and solid state data.

Incorporation of folded protein domains into coiled-coil protein origami

Strmšek Žiga[1,2], Tušar Anja[1], Vidmar Sara[1], Pavko Neža[1], Aupič Jana[1], Lapenta Fabio[1], Lainšček Duško[1], Trefalt Gregor[3], Jerala Roman[1,2]

[1] Department of Synthetic Biology and Immunology, National Institute of Chemistry, Ljubljana, Slovenia. [2] EN-FIST Centre of Excellence, Ljubljana, Slovenia. [3] Department of Inorganic and Analytical Chemistry, University of Geneva, Geneva, Switzerland.

Natural protein folds have large diversity with a broad range of 3D structures and functions, however, proteins could adopt almost endless diversity that could not have been explored by nature. New types of protein assemblies, therefore, represent an opportunity for biochemists to rationally design protein assemblies based on principles not explored by nature. Coiled-coil protein origami (CCPO) is a *de novo*, rationally designed type of protein folds, composed by concatenating coiled-coils (CC) building modules into a single polypeptide chain, with the ability to self-assemble into precisely defined polyhedral nanocages. Many different protein polyhedra have been designed and characterized, such as a trigon, tetrahedron and rectangular pyramid, up to bipyramid and trigonal prism that are among the largest *de novo* designed polypeptide chains[1–5].

Here we present a system for introduction of protein domains at the vertices of tetrahedral CCPOs by direct genetic fusion, as a way of functionalization of protein origami. As a proof-of-principle studies, tetrahedral CCPO TET12SN, decorated with up to four GFP or RFP protein domains have been designed, isolated and characterized. Analysis with CD spectroscopy confirmed the expected secondary structure content, analytical size exclusion indicated correct size and SAXS revealed the expected shape. Based on those results, immunization trials were conducted, with RFP chosen as a target antigen. Mice were immunized with equal amount of (i) monomeric antigen (RFP), (ii) CCPO scaffold decorated with the antigen (TET12SN-RRRR), (iii) CCPO scaffold (TET12SN) or (iv) buffer. Mice immunized with an antigen, presented on a scaffold (TET12SN-RRRR), developed IgG antibodies against RFP faster and in higher quantities compared to mice immunized with a monomeric antigen.

To expand the capabilities of the developed system of decorating CCPO, rLuc protein domain was presented on TET12SN on one, two or three vertices, with the remaining vertices decorated with GFP and RFP protein domains so all four vertices were occupied in all cases. Again, CD spectroscopy, analytical size exclusion chromatography and SAXS measurements confirmed the success of the developed strategy. Moreover, the catalytic activity of rLuc was not effected upon integration into TET12SN.

During the analysis of *in silico* generated models, the enhanced flexibility of the protein domain, presented at the vertex IV was observed – to remedy that, the insertion site for the protein domain, presented at the vertex IV into the TET12SN was modified. Seven additional protein constructs with varying ratios of GFP and RFP protein domain and with different positions of insertion were designed, isolated and characterized. Following the same principle, seven additional structures similar to the previously described ones with varying ratios of GFP to RFP were designed, but this time TET12SN(22CC), a tetrahedral CCPO constructed from only four unique CC pairs, two of them repeated two times[4], was used as a scaffold. A protein design, built on TET12SN(22CC) and decorated with four RFP protein domains was also investigated with AFM and cryoEM, providing further evidence of the successfully decorating vertices of CCPO.

The developed system of decorating vertices of tetrahedral CCPOs with protein domains offers a robust, versatile tool for future applications, where precise distance and stoichiometry of presented protein domains are required, which could be used in medical or biotechnological applications.

1. Gradišar, H. *et al.* Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat. Chem. Biol.* **9**, 362–366 (2013).
2. Ljubetič, A. *et al.* Design of coiled-coil protein-origami cages that self-assemble in vitro and in vivo. *Nat. Biotechnol.* **35**, 1094–1101 (2017).
3. Lapenta, F. *et al.* Self-assembly and regulation of protein cages from pre-organised coiled-coil modules. *Nat. Commun.* (2021) doi:10.1038/s41467-021-21184-6.
4. Aupič, J. *et al.* Designed folding pathway of modular coiled-coil- based proteins. *Nat. Commun.* (2021) doi:10.1038/s41467-021-21185-5.
5. Božič Sabina, Gradišar, H., Aupič, J., Round, A. R. & Jerala, R. Triangular in Vivo Self-Assembling Coiled-Coil Protein Origami. *Cite This ACS Chem. Biol* **16**, 310–315 (2021).