Supporting Information

**Bioinspired Design of Highly Wet-Adhesive and Elastic Photocrosslinkable Recombinant Human Protein Networks for Surgical Applications**

*Anna Resch1, Matthias C. Huber1,2,3,\* Alexander Resch4,\* Andreas Schreiber1,2, Thabo Lapp4, Uwe Jonas,1 Adrianna Kolberg5, Sandra Haas6, Mathaeus Tschaikowsky5,8, Jürgen Hubbuch6, Daniel Böhringer7, Thomas Reinhard7, Bernd Rolauffs8, Bizan N. Balzer3,5, Günther Schlunck7 & Stefan Schiller1,3,8,9,10 (\*equally contributing authors)*

1Center for Biological Systems Analysis, University of Freiburg, Habsburgerstrasse 49, 79104 Freiburg, Germany

2Faculty of Biology, University of Freiburg, Schänzlestrasse 1, 79104 Freiburg, Germany

3Cluster of Excellence livMatS @ FIT, Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, Georges-Köhler-Allee 105, 79110 Freiburg, Germany

4Chair for Information Systems Research, University of Freiburg, Platz der Alten Synagoge, 79098 Freiburg, Germany

5nstitute of Physical Chemistry, Albertstrasse 21, University of Freiburg, 79104 Freiburg

6Karlsruhe Institute of Technology, Institute of Functional Interfaces & Institute of Engineering in Life Sciences, Section IV: Molecular Separation Engineering, Karlsruhe Institute of Technology, Fritz-Haber-Weg 2, 76131 Karlsruhe, Germany

7Eye Center, Medical Center, Faculty of Medicine, University of Freiburg, 79110 Freiburg, Germany

8G.E.R.N. Research Center for Tissue Replacement, Regeneration & Neogenesis, Department of Orthopedics and Trauma Surgery, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

9BIOSS Centre for Biological Signalling Studies, University of Freiburg, Schänzlestrasse 18, 79104 Freiburg, Germany

10IMTEK Department of Microsystems Engineering, University of Freiburg, Georges-Köhler-Allee 103, 79110 Freiburg, Germany

Content

[Supplementary Information 3](#_Toc92462530)

[1. Genetic engineering of ULD-ELP-ULD building blocks 3](#_Toc92462531)

[*2.* Sequence homology of ULD-variants from different species compared with the human ULD-sequence *PDB 3TUO* 16](#_Toc92462532)

[3. Protein expression, purification and quantification 23](#_Toc92462533)

[3.1. Materials and Methods 23](#_Toc92462534)

[3.2. Protein purity and yield 24](#_Toc92462535)

[4. ULD tetramerization 24](#_Toc92462537)

[5. PyMol 25](#_Toc92462538)

[6. Dynamic Light Scattering (DLS) 25](#_Toc92462539)

[7. Mass spectrometry 26](#_Toc92462540)

[7.1. LC-MS-MS Analysis of tryptic peptides 26](#_Toc92462541)

[7.2. Tryptic peptides: Data Processing and Bioinformatic Analysis 27](#_Toc92462542)

[7.3. LC-MS-MS Analysis of intact proteins 27](#_Toc92462543)

[7.4. Intact proteins: Data Processing and Bioinformatic Analysis 27](#_Toc92462544)

[7.5. LC-MS²-characterization results of building blocks 27](#_Toc92462545)

[8. Hydrogel fabrication and crosslinking 29](#_Toc92462546)

[9. Dityrosine content 30](#_Toc92462547)

[10. Water content 32](#_Toc92462548)

[11. Nanoindentation 32](#_Toc92462549)

[12. AFM Analysis of Mechanical Properties - Young's Modulus and Resilience 33](#_Toc92462560)

[13. Dynamic Mechanical Analysis (DMA) 34](#_Toc92462561)

[13.1. Experimental settings 34](#_Toc92462562)

[13.2. DMA data processing and evaluation 34](#_Toc92462563)

[14. Sealing of corneal tissue ruptures and stromal defects 35](#_Toc92462564)

[14.1. Filling and sealing of cornea defects in porcine eyes 35](#_Toc92462565)

[14.2. Filling and sealing of human corneal explants 36](#_Toc92462566)

[References 36](#_Toc92462567)

# Supplementary Information

### Genetic engineering of ULD-ELP-ULD building blocks

The DNA sequence for the ubiquitin-like domain (ULD) of the human global gene organizer SATB1 (Special AT-rich sequence-binding protein 1; M97287)1 was codon optimized for *E. coli* K12 expression2 and for the implementation into the One-Vector-Toolbox-Platform (OVTP)3. ULD was further complemented with additional DNA sequences for the adaption to the NMBL-linker region of the OVTP system (see DNA sequence below). The synthetic gene was designed with Geneious 6.0.4 software and ordered as a synthetic gene from Invitrogen (Thermofisher Scientific Inc.). The synthesized ULD was cloned initially via BamHI and HindIII restriction digest into a pET28-NMBL-His vector to generate a pET28-NMBL-ULD-His plasmid. Afterwards the ULD sequence could be subcloned as an insertion module via SacI and EarI restriction digest or supplemented via BspQI and SacI restriction digest as described in detail in Huber et al. 20143. In this work the pET28-NMBL-ULD-His plasmids were completed with compatible elastin-like protein (ELP) spacer modules (ELP: V10, V20, V40, V80, V120, D20, H20, S20, R20, R40) and additional linker sequences shown below to pET28-NMBL-ELP-ULD-His (e.g., ELP spacer = (VPGVG)20 = V20 and further linker protein sequences shown below) followed by the addition of another compatible ULD module to complete the pET28-NMBL-ULD-Linker-ULD-His (for example: pET28-NMBL-ULD-V20-ULD-His) plasmid. The ELP linker could be replaced by any other (OVTP-compatible) functional or structural protein domain encoding DNA sequence (e.g., for HSA, ELPs, recombinant resilin, SpyCatcher, recombinant spider silk etc.) to introduce additional functionalities and to adjust the physicochemical properties of the resulting hydrogel for the aimed purposes. The resulting pET28-NMBL-ULD-Linker-ULD-His vectors were directly transformed into *E. coli* expression cells (here BL21(DE3) or BLR etc.) as described in Huber et al. 20143 and expressed and purified as described elsewhere4. Exemplary DNA sequences and resulting amino acid (aa) sequences of the applied ULD-ELP-ULD-His fusion proteins are listed below.

**DNA and aa sequences of ULD and plasmid linker region and examined ULD fusion proteins**

**Examples for ULD-ELP-ULD constructs**

**Color Codes**

**blue**: ULD aa sequence and encoding DNA sequences

**green**: Linker aa sequences and encoding DNA sequences

**black**: bridging aa (or aa sequences) and encoding DNA sequences originated from the

OVTP cloning platform used to create the respective constructs

**Human ubiquitin-like domain (ULD)**: Codon optimized DNA sequence of human ULD(**bold sequence**) with 5' BamHI and 3' HindIII restriction sites (underlined)

5'GGATCCTATGAGCTCTTCT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGGAAGAGAAGCTT3'

**Human ULD:** aa sequence of ULD (blue **bold sequence**) with flanking linker sequences

DPMSSS**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GGREA

**NMBL-linker** **region** from pET28-NMBL-His plasmid



**ULD-***V20***-ULD**-His from pET28-NMBL-ULD-V20-ULD-His plasmid

DNA sequence (975 bp)

5'ATGGATCCTATGAGCTCTTCT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA3'

aa sequence (324 aa / 33.3 kDa)

MDPMSSS**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGV**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GGREASSHHHHHH

**ULD-***V40***-ULD-His** from pET28-NMBL-ULD-V40-ULD-His plasmid

DNA sequence (1275 bp)

5'ATGGATCCTATGAGCTCTTCT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA3'

aa sequence (424 aa / 41.4 kDa)

MDPMSSS**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGV**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GGREASSHHHHHH

**ULD-***V80***-ULD-His** from pET28-NMBL-ULD-V80-ULD-His plasmid

DNA sequence (1875 bp)

5'ATGGATCCTATGAGCTCTTCT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTTGGTGTTCCGGGTGTT**GGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCT**GGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA3'

aa sequence (624 aa / 57.8 kDa)

MDPMSSS**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGV**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**GGREASSHHHHHH

Additional linker proteins of the ULD-Linker-ULD system expanding functional ELPs with soluble charged ELP-Tyr-variants ((DSY)8 and (VRY)8) and expression of ELP-linker substitutes: ULD-Linker-ULD constructs representing different properties by linker = Spider Silk, Resilin, SpyCatcher, mEGFP, HSA, EYFP-TEVRc-mEGFP and EYFP-TEVRc:

***Table S-1****:* ***ULD-Linker-ULD-His*** *constructs with different physicochemical, mechanical and functional properties encoded in the pET28-NMBL-ULD-Linker-ULD-His plasmid and expressed in E. coli BL21(DE3) cells*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Linker** | **abbreviation/specification** | **Mw**  kDa | **length**  aa | **designated property** |
| (DSY)8 | (VPGDG/VPGSG/VPGYG)8 | 35,4 | 344 | crosslinkable linker, - charge |
| (VRY)6 | (VPGVG/VPGRG/VPGYG)6 | 33,2 | 314 | crosslinkable linker, + charge |
| mEGFP | monomeric enhanced GFP | 52,1 | 464 | fluorescent protein (FP) linker |
| EYFP-TEVrec-mEGFP | FP-TEVprotease recogn.site-FP | 80,0 | 711 | cleavable double FP linker |
| EYFP-TEVrec | FP-TEVprotease recogn.site | 52,9 | 471 | cleavable double FP linker |
| mCHR-TEVrec-mEGFP | FP-TEVprotease recogn.site-FP | 79,7 | 708 | cleavable double FP linker |
| spisi10 | recombinant spider silk (10 x) | 53,8 | 574 | specific mechanical strength |
| spisi20 | recombinant spider silk (20 x) | 82,6 | 915 | specific mechanical strength |
| resi10 | recombinant resilin (10 x) | 38,4 | 374 | specific resilient properties |
| SpyCatcher | SpyCatcher (/SpyTag-System) | 34,3 | 310 | covalently modifiable linker |
| HSA | Human Serum Albumin | 91,6 | 810 | different functional properties |

All constructs were cloned in a high-expressing pET28-NMBL-vector as pET28-NMBL-ULD-Linker-ULD-His construct (see Huber et al., Biomaterials 2014)3. ULD-spisi10/20-ULD and ULD-resi10-ULD comprise potentially fibrous and/or elastic and, especially for the latter, mechanically resilient linker sequences. ULD-SpyCatcher-ULD contains an adapter linker sequence that can be conjugated with SpyTag (short protein) functionalized molecules allowing for site-selective bio-orthogonal protein functionalization. The ULD-mEGFP-ULD construct contains a globular green fluorescence protein as linker for visualization purposes. ULD-(DSY)8-ULD and ULD-(VRY)6-ULD include charged disordered and photo-crosslinkable linker sequences to adjust protein-tissue adhesion via glue charge and mechanical strength of the hydrogel via additional Tyr-crosslinking sites. ULD-EYFP-TEVrc-mEGFP-ULD as well as ULD-EYFP-TEVrc-ULD comprise linker sequences combining fluorescent properties with possible network degradability properties via the TEVprotease recognition sequence motif (TEVrec). The ULD-HSA-ULD construct combines protein glue properties with bio-resorbable linker properties of human derived serum albumin characteristics.

**Examples for ULD-Linker-ULD constructs with different properties**

**Color Codes**

**blue**: ULD aa sequence and encoding DNA sequences

**green**: Linker aa sequences and encoding DNA sequences

**black**: bridging aa (or aa sequences) and encoding DNA sequences originated from the

OVTP cloning platform used to create the respective constructs

I. Linker domain: resilin-like ELPs *(ELP = elastin-like proteins)*

ULD-(DSY)8-ULD / aa (35.4 kDa, 344 aa) *((DSY)1 = ((VPGDG)1(VPGSG)1(VPGYG)1 ELP)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGVPGDGVPGSGVPGYGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-(DSY)8-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTGTTCCGGGTGATGGTGTTCCGGGTAGCGGTGTTCCGGGTTACGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

ULD-(VRY)6-ULD / aa (33.2 kDa, 314 aa) *((VRY)1 = ((VPGVG)1(VPGRG)1(VPGYG)1 ELP)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGVPGVGVPGRGVPGYGVPGVGVPGRGVPGYGVPGVGVPGRGVPGYGVPGVGVPGRGVPGYGVPGVGVPGRGVPGYGVPGVGVPGRGVPGYGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-(VRY)6-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGTTCCGGGTGTTGGTGTTCCGGGTCGCGGTGTTCCGGGTTACGGTGTTCCGGGTGTTGGTGTTCCGGGTCGCGGTGTTCCGGGTTACGGTGTTCCGGGTGTTGGTGTTCCGGGTCGCGGTGTTCCGGGTTACGGTGTTCCGGGTGTTGGTGTTCCGGGTCGCGGTGTTCCGGGTTACGGTGTTCCGGGTGTTGGTGTTCCGGGTCGCGGTGTTCCGGGTTACGGTGTTCCGGGTGTTGGTGTTCCGGGTCGCGGTGTTCCGGGTTACGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

II. Linker domain: fluorescent proteins and proteolytic cleavage sites

ULD-mEGFP-ULD / aa (52.1 kDa, 464 aa) *(mEGFP = monomeric enhanced GFP)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-mEGFP-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

ULD-EYFP-*TEVrec*-mEGFP-ULD / aa (80.0 kDa, 711 aa) *(EYFP = enhanced yellow fluorescent protein; TEVrec = TEV protease recognition sequence)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG*ENLYFQ*GMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-EYFP-*TEVrec*-mEGFP-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGT*GAAAACCTGTACTTTCAG*GGTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

ULD-EYFP-*TEVrec*-ULD / aa (52.9 kDa, 471 aa)

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKG*ENLYFQ*GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-EYFP-*TEVrec*-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

ULD-mCHR-*TEVrec*-mEGFP-ULD / aa (79.7 kDa, 708 aa) *(mCHR = monomeric mCherry; TEVrec = TEV protease recognition sequence)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKG*ENLYFQ*GMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-mCHR-*TEVrec*-mEGFP-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGGGT*GAAAACCTGTACTTTCAG*GGTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

III. Linker domain: (resilient) structural recombinant protein domains

ULD-spisi10-ULD / aa (53.8 kDa, 574 aa) *(spisi10 = recombinant spider silk 10x)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGRGGLGGQGAGMAAAAAMGGAGQGGYGGLGSQGTSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-spisi10-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTCGTGGTGGTCTGGGTGGTCAGGGTGCTGGTATGGCTGCTGCTGCTGCTATGGGTGGTGCTGGTCAGGGTGGTTACGGTGGTCTGGGTTCTCAGGGTACCTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

ULD-resi10-ULD / aa (38.4 kDa, 374 aa) *(resi10 = recombinant resilin 10x)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGGRPSDSYGAPGGGNGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-resi10-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTGGTCGACCTTCTGATTCTTACGGTGCTCCTGGTGGTGGTAATGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

IV. Linker domain: human functional protein domain

ULD-HSA-ULD / aa (91.6 kDa, 810 aa) *(HSA = recombinant Human Serum Albumin)*

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGLGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-HSA-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGACGCTCACAAATCTGAAGTTGCTCACCGTTTCAAAGACCTGGGTGAAGAAAACTTCAAAGCTCTGGTTCTGATCGCTTTCGCTCAGTACCTGCAGCAGTGCCCGTTCGAAGACCACGTTAAACTGGTTAACGAAGTTACCGAATTTGCTAAAACCTGCGTTGCTGACGAATCTGCTGAAAACTGCGACAAATCTCTGCACACCCTGTTCGGTGACAAACTGTGTACCGTTGCTACCCTGCGTGAAACCTACGGTGAAATGGCTGACTGCTGCGCTAAACAGGAACCGGAACGTAACGAATGCTTCCTGCAGCACAAAGACGACAACCCGAACCTGCCGCGTCTGGTTCGTCCGGAAGTTGACGTTATGTGTACCGCTTTCCACGACAACGAAGAAACCTTCCTGAAAAAATACCTGTACGAAATCGCTCGTCGTCACCCGTACTTCTACGCTCCGGAACTGCTGTTCTTCGCTAAACGTTACAAAGCTGCTTTCACCGAATGCTGCCAGGCTGCTGACAAAGCTGCTTGCCTGCTGCCGAAACTGGACGAACTGCGTGACGAAGGTAAAGCGTCTTCTGCTAAACAGCGTCTGAAATGCGCTTCTCTGCAGAAATTCGGTGAACGTGCTTTCAAAGCGTGGGCTGTTGCTCGTCTGTCTCAGCGTTTCCCGAAAGCTGAATTTGCTGAAGTTTCTAAACTGGTTACCGACCTGACCAAAGTTCACACCGAATGCTGCCACGGTGACCTGCTGGAATGCGCTGACGACCGTGCTGACCTGGCTAAATACATCTGCGAAAACCAGGACTCTATCTCGTCTAAACTGAAAGAATGCTGCGAAAAACCGCTGCTGGAAAAATCTCACTGCATCGCTGAAGTTGAAAACGACGAAATGCCGGCTGACCTGCCGTCTCTGGCTGCTGACTTCGTTGAATCTAAAGACGTTTGCAAAAACTACGCTGAAGCTAAAGACGTTTTCCTGGGTATGTTCCTGTACGAATACGCTCGTCGTCACCCGGACTACTCTGTTGTTCTGCTGCTGCGTCTGGCTAAAACCTACGAAACCACCCTGGAAAAATGCTGCGCTGCTGCTGACCCGCACGAATGCTACGCTAAAGTTTTCGACGAGTTCAAACCGCTGGTTGAAGAACCGCAGAACCTGATCAAACAGAACTGCGAACTGTTCGAACAGCTGGGTGAATACAAATTCCAGAACGCTCTGCTGGTTCGTTACACCAAAAAAGTTCCGCAGGTTTCTACCCCGACCCTGGTTGAAGTTTCTCGTAACCTGGGTAAAGTTGGTTCTAAATGCTGCAAACACCCGGAAGCTAAACGTATGCCGTGCGCTGAAGACTACCTGTCTGTTGTTCTGAACCAGCTGTGCGTTCTGCACGAAAAAACCCCGGTTTCTGACCGTGTTACCAAATGCTGCACCGAATCTCTGGTTAACCGTCGTCCGTGCTTCTCTGCTCTGGAAGTTGACGAAACCTACGTTCCGAAAGAGTTCAACGCTGAAACCTTCACCTTCCACGCTGACATCTGCACCCTGTCTGAAAAAGAACGTCAGATCAAAAAACAGACCGCTCTGGTTGAACTGGTTAAACACAAACCGAAAGCTACCAAAGAACAGCTGAAAGCTGTTATGGACGACTTCGCTGCTTTCGTTGAAAAATGCTGCAAAGCTGACGACAAAGAAACCTGCTTCGCTGAAGAAGGTAAAAAACTGGTTGCTGCTTCTCAGGCTGCTCTGGGTCTGGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

V. Linker domain: modular modifiable domain (SpyTag/SpyCatcher conjugation system)

ULD-SpyCatcher-ULD / aa (34.3 kDa, 310 aa)

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGSGDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-SpyCatcher-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTTCAGGTGATAGTGCTACCCATATTAAATTCTCAAAACGTGATGAGGACGGCAAAGAGTTAGCTGGTGCAACTATGGAGTTGCGTGATTCATCTGGTAAAACTATTAGTACATGGATTTCAGATGGACAAGTGAAAGATTTCTACCTGTATCCAGGAAAATATACATTTGTCGAAACCGCAGCACCAGACGGTTATGAGGTAGCAACTGCTATTACCTTTACAGTTAATGAGCAAGGTCAGGTTACTGTAAATGGCGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

VI. Linker domain: Additional ELP domains (charged and water soluble)

ULD-(VPGKG)20-ULD / aa (33.8 kDa, 324 aa)

MDPMSSSGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGVPGKGTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHSGGREASSHHHHHH

ULD-(VPGKG)20-ULD / DNA

ATGGATCCTATGAGCTCTTCTGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTGTTCCGGGTAAAGGTACCATGCTGCCGGTTTTCTGCGTTGTTGAACACTACGAAAACGCTATCGAATACGACTGCAAAGAAGAACACGCTGAATTTGTTCTGGTTCGTAAAGACATGCTGTTCAACCAGCTGATCGAAATGGCTCTGCTGTCTCTGGGTTACTCTCACAGCTCTGCTGCTCAGGCTAAAGGTCTGATCCAGGTTGGTAAATGGAACCCGGTTCCGCTGTCTTACGTTACCGACGCTCCGGACGCTACCGTTGCTGACATGCTGCAGGACGTTTACCACGTTGTTACCCTGAAAATCCAGCTGCACTCTGGTGGAAGAGAAGCTTCATCACACCACCACCACCACCACTGA

****

***Figure S-1:*** *SDS-PAGE of expressed ULD-Linker-ULD constructs cloned in pET28-NMBL-His plasmid and expressed in E. coli BL21(DE3) cells. The headings on top of the protein gel lanes denominate for the linker sequence between two ULD protein domains. The (+) symbols on top of each lane mark IPTG-induced single clones of the respective construct. (–) symbols mark non-induced negative controls to induced clones in the neighboring lane on the right. The arrows on the left and right sides mark and denote the Mw of the different expressed fusion proteins. A. From left to the right expressed proteins for the ULD-Linker-ULD constructs are shown with the linkers: recombinant spider silk (spisi10), recombinant resilin (resi10), SpyCatcher protein, mEGFP and Human Serum Albumin (HSA) for different expressed single E. coli clones. B. From left to the right expressed proteins are shown for the ULD-Linker-ULD constructs with the linkers: EYFP-TEVrec-mEGFP, EYFP-TEVrec, (VPGDG/VPGSG/VPGYG)8, abbrev. for (DSY)8 and (VPGVG/VPGRG/VPGYG)6, abbrev. for (VRY)6 for different expressed single E. coli clones.*

### Sequence homology of ULD-variants from different species compared with the human ULD-sequence *PDB 3TUO*

Amino acids sequence: human ULD domain as shown for *PDB ENTITY SEQ 3TUO*:

**GTMLPVFCVVEHYENAIEYDCKEEHAEFVLVRKDMLFNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNPVPLSYVTDAPDATVADMLQDVYHVVTLKIQLHS**

yellow – **Cys** – potential disulfide formation

green – **Tyr** – potential C-C coupling via (Photo)redox reactionen

red – Lys-Gln – potential transglutaminase susceptible side chains

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Dog**, (*Canis lupus familiaris and Canis lupus dingo*, 100% sequence identity for DNA-binding protein **SATB1** (*Canis lupus familaris*: XP\_02264105.1)*)*:

DNA-binding protein **SATB1** *Canis lupus familaris*: XP\_022264105.1 , 100% sequence identity:

1 mdhlneapqg kehsdmsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmktnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvnfg qqpvpgntae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qiqspspttl 601 gtgdsrgvfl pglltpapwl gaapqvsghk lqvtytskrg ggiktcslgq nfqtckghti 661 tvvvardhca sgfslvlisw schgtclhss tegtplprrg glqtlsaqld lskdtiikff 721 qnqryylkhh gklkdnsgle vdvaeykeee llkdleesvq dknantlfsv kleeelsveg 781 ntdinadlkd

**human/dog 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Cat** (*Felis catus*, DNA-binding protein **SATB1** isoform X4, 100% sequence identity (XP\_011284516.1):

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmktnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvnfg qqpvpgntae qppspaqlsh gnqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qiqqpqaqaq 601 qqpappppqa qaqqpagprl pprqptvasp aeaedenrqk trprtkisve algilqsfiq 661 dvglypdeea iqtlsaqldl pkytiikffq nqryylkhhg klkdnsglev dvaeykeeel 721 lkdleesaqd knantlfsvk leeelsvegn tdintdlkd

**human/ cat 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Horse**, (*Equus caballus*, DNA-binding protein **SATB1** isoform X1, 100% sequence identity (XP\_014587330.2):

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmktnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvnfg qqpvpgntae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qiqqpppqpq 601 qppppppqpq aqqpagprlp prqptvaspa eaedenrqkt rprtkisvea lgilqsfiqd 661 vglypdeeai qtlsaqldlp kytiikffqn qryylkhhgk lkdnsglevd vaeykeeell 721 kdleesvqdk nantlfsvkl eeelsvegnt dinadlkd

**human/ horse 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Wild boar** (*Sus scrofa*, DNA-binding protein **SATB1**, 100% sequence identity (XP\_003358373.2)

DNA-binding protein **SATB1**, 100% sequence identity (XP\_003358373.2):

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmksnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledxp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvnfg qqpvpgntve qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qiqqppaqaq 601 pqpapppaqa qaqqpqagpr lpprqptvaa paeaeedsrq karprtkisv ealgilqsfi 661 qdvglypdee aiqtlsaqld lpkytiikff qnqryylkhh gklkdnsgle vdvaeykeee 721 llkdleesvq dknantlfsv kledelaveg ntdinadvkd

**human/ Wild boar 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Cow** (*Bos taurus*, DNA-binding protein **SATB1**, 100% sequence identity (NP\_001095510.1))

DNA-binding protein **SATB1**, 100% sequence identity (NP\_001095510.1):

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmksnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvsfg qqpapgstae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qiqppqappq 601 pqpapppapa qappppagpr lparqppaaa paeaedesrp karprtkisv ealgilqsfm 661 qdvglhpdee aiqtlsaqld lpqhtvikff qnqryhlkhh grlkdhagle vdvadykdde 721 llqdpdegap gagasplfav kledepaadg stdasadard

**human/ cow 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Sheep** (*Ovis aries*, DNA-binding protein **SATB1**, 100% sequence identity (XP\_027817855.1))

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmksnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvsfg qqpapgstae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snaghhhgdr pphvihvpae qmqspsptal 601 gqgesrgirp rgwycstgeq aqppqappqp qpapppapaq appppagprl parqppaaap 661 aeaedesrpk arprtkisve algilqsfmq dvglhpdeea iqtlsaqldl pqhtvikffq 721 nqryhlkhhg rlkdhaglev dvadykddel lqdpdegapg agasplfavk leeepaadgs 781 tdasadard

**human/ sheep 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Guinea pig** (*Cavia porcellus*, DNA-binding protein **SATB1** isoform X1, 100% sequence identity (XP\_005008358.1))

1 mdhlnettqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmktnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvnfg qqpvpgntae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 serngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qlqqqqqqqq 601 qqqqqppapp qpqpqpqpqa qagprlpprq ptvaspaesd eenrqksrpr tkisvealgi 661 lqsfiqdvgl ypdeeaiqtl saqldlpkyt iikffqnqry ylkhhgklkd nsglevdvae 721 ykeeellkdl eesvldknan tlfsvkleee lsvegntdid adlkd

**human/ Guinea pig 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Rabbit** (*Oryctolagus cuniculus,* DNA-binding protein **SATB1** 100% sequence identity (XP\_008264272.1))

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmktnlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqganhvnfg qqpvpgntae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pphiihvpae qiqspspttl 601 gkgesrgvfl pglptpapwl gaapqqpqpq qqqqqqpapp ppqpqpqpqa gprlpprqpt 661 vaspaesdee nrqktrprtk isvealgilq sfiqdvglyp deeaiqtlsa qldlpkytii 721 kffqnqryyl khhgklkdns glevdvaeyk eeellkdlee svqdknantl fsvkleeels 781 vegntdisad lkd

**human/ rabbit 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Rodents** (prairie vole, *Microtus ochrogaster,* DNA-binding protein **SATB1** isoform X1 sequence identity 100 %, (XP\_005368803.1))

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstga kmqgvplkhs 61 ghlmkanlrk gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkktkdmm vemdslsels qqgtnhvnfg qqpvpgntae qppspaqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp apllstppsr ppqvktatla 481 terngkpenn tmninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq serdaiyeqe snavhhhgdr pphiihvpae qiqspspstl 601 gkgesrcasl pglltpapwp saapqqqqqq qqqqqqpppp pqpqpqpqag prlpprqptv 661 assaesdeen rqktrprtki svealgilqs fiqdvglypd eeaiqtlsaq ldlpkytiik 721 ffqnqryylk hhgklkdnsg levdvaeyke eellkdlees vqdknantlf svklegelsv 781 egstdinadl kd

**human/ prairie vole 100% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Chicken**, *Gallus gallus*, DNA-binding protein **SATB1** (NP\_001186573.1) 98% sequence identity),

1 mdhlneatqg kehsemsnnv sdpkgppaki arleqngspl grgrlgstgt kmqgvplkhs 61 ghlmktnirk gsmlpvfcvv ehyenaieyd skeehaefvl vrkdmlfnql iemallslgy 121 shssaaqakg liqvgkwnpv plsyvtdapd atvadmlqdv yhvvtlkiql hscpkledlp 181 peqwshttvr nalkdllkdm nqsslakecp lsqsmissiv nstyyanvsa akcqefgrwy 241 khfkkakdmm vemdslsels qqganhvnfg qqpvpgntae qppspvqlsh gsqpsvrtpl 301 pnlhpglvst pispqlvnqq lvmaqllnqq yavnrllaqq slnqqylnhp ppvsrsmnkp 361 leqqvstnte vsseiyqwvr delkragisq avfarvafnr tqgllseilr keedpktasq 421 sllvnlramq nflqlpeaer driyqderer slnaasamgp aplistppsr ppqvktatia 481 terngktenn smninasiyd eiqqemkrak vsqalfakva atksqgwlce llrwkedpsp 541 enrtlwenls mirrflslpq perdaiyeqe snavhhhgdr pshiihvpae qiqqqqqqqq 601 qqqqqqqqqq pgprlpprqp tvaspaesed enrqkprprt kisvealgil qsfiqdvgly 661 pdeeaiqtls aqldlpkyti ikffqnqryy lkhhgklkdn sglevdvaey keeellkdle 721 dsiqdknant lfsvkleeel svegnteina elkd

**human/ chicken 98% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gsmlpvfcvv ehyenaieyd skeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Amphibian**

*Alligator mississippiensis*, DNA-binding protein **SATB2** 98% sequence identity, Seq. ID (KYO26889.1)

1 mrtnirkgsm lpvfcvvehy enaieydske ehaefvlvrk dmlfnqliem allslgyshs 61 saaqakgliq vgkwnpvpls yvtdapdatv admlqdvyhv vtlkiqlhsc pkledlppeq 121 wshttvrnal kdllkdmnqs slakecplsq smissivnst yyanvsaakc qefgrwykhf 181 kkakdmmvem dsiselpqqg anhvnfgqqp vpgntaeqpp spvqlshgsq psvrtplpnl 241 hpglvstpis pqlvnqqlvm aqllnqqyav nrllaqqsln qqylnhpppv srsmnkpleq 301 qvstntevss eiyqwvrdel kragisqavf arvafnrtqg llseilrkee dpktasqsll 361 vnlramqnfl qlpeaerdri yqderersln aasamgpapl istppnrppq vktatiater 421 nvktenntmn inasiydeiq qemkrakvsq alfakvaatk sqgwlcellr wkedpspenr 481 tlwenlsmir rflslpqper daiyeqesna vhhhgdrpsh iihvpaeqiq spsptthvke 541 ehrgaflpgl lttgpwlgat pqqqqqqqqq qqqqqqqpqq qqpgprlppr qptvaspaes 601 edenrqkprp rtkisvealg ilqsfiqdvg lypdeeaiqt lsaqldlpky tiikffqnqr 661 yylkhhgklk dnsglevdva eykeeellnd ledsiqdkna ntlfsvklee elsvegntel 721 ntelkd

**human/ alligator 98% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gsmlpvfcvv ehyenaieyd skeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

**Fish**

rainbow trout, *Oncorhynchus mykiss*, DNA-binding protein SATB1-like isoform X6, (XP\_021450053.1) with Seq. Identity 95%

DNA-binding protein SATB1-like isoform X6, XP\_021450053.1 , 95% Seq. Identity

1 mmdhlseacl gkencdlvns mgdrdskapp aklarleqng splgrarlgs tgaklagvpy 61 kpsghllknc hkrgnmlpvf cvvehyespi efdskeehae fvlvrkdmlf nqliemalls 121 lgyshssaaq akgliqvgkw npvplsyvtd apdatvadml qdvyhvvtlk iqlhscpkle 181 dlppeqwths tvrnalkell kdmnqsslak ecplsqsmis sivnstyyan vsaakcqefg 241 rwykhfkktk dmmgmrqpsq legylgtcvl hdsprantla dmdglsdhsp pgpnhnhnhl 301 tfsqqpipgn taeqppsspv pplpppshgg gqtpgrpaql ppglhhpglv stpispqlvn 361 qqlvmaqiln qqyavnrlla qqslsqqyln hppvnrnalt kplepqvasn tevsmeiyqw 421 vrdelkragi sqavfarvaf nrtqgllsei lrkeedpkta sqsllvnlra mqnflqlpea 481 erdriyqder ersltaasam gpaplistpp trpvqvwneq qprredcnnv rpedwnprip 541 vgispipesr kpgsitlfsn pfadtpypgq vkasplpsew ngktesciln inssiydeiq 601 qemkrakvsq alfakvaask sqgwlcellr wkedpspenr tlwenlsmir rflslaqger 661 daiyeqesta gqqqhhadrp phmmhmstdp mqsqthqppq qqqqhqppms lqhpsqppls 721 qqpmqqqqqq ppmgprlppr qpstaspaet edqarggggt karsgggkis qealgilqsf 781 iqdvgmypde eaihtlsaql dlpkltiikf fqnqrfyvnh pakppkepsn ssssspafee 841 elsdfragsa ellkeleest qtnstifsik ieehlarsea lsesdhetke

**human/rainbow trout 95% sequence homology:**

gtmlpvfcvv ehyenaieyd ckeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

gnmlpvfcvv ehyespiefd skeehaefvl vrkdmlfnql iemallslgy shssaaqakg liqvgkwnpv plsyvtdapd

atvadmlqdv yhvvtlkiql hs

atvadmlqdv yhvvtlkiql hs

\_\_\_ \_\_\_\_ \_\_\_\_ \_\_\_\_

### Protein expression, purification and quantification

#### Materials and Methods

We used *E. coli* strain BL21(DE3) for protein expression. Starter cultures of 2YT or Terrific Broth (TB) medium with 40 µg/mL kanamycin were inoculated with cells from glycerol stocks. We grew bacteria while shaking at 250 rpm and 37 °C overnight. The next morning, we centrifuged 2 mL starter culture for 4 min at 3,220 g and resuspended it in 1 mL fresh 2YT or TB medium as described by Quiroz and Chilkoti5. 400 mL fresh medium was inoculated with the resuspended starter culture and bacteria were grown for ca. 8 hours at 37 °C while shaking at 180 rpm. We induced protein expression by adding IPTG to a final concentration of 1 mM and temperature was reduced to 20 °C. Cells were then harvested after approx. 20 hours by centrifugation at 3,220 g and 4 °C for 30 min. For cell lysis, we used a lysis buffer containing 50 mM Tris-HCl, 500 mM NaCl, 4 M urea and 20 mM imidazole, pH 7.5, to resuspend the cell pellet. We then added lysozyme to a final concentration (f.c.) of 0.1 mg/mL, PMSF (f.c. 1 mM) and TCEP (f.c. 1 mM) and incubated the resuspended cells for 30 min on ice. After 2 times of freezing the suspension in liquid nitrogen and thawing and incubation with DNAse I in between the freezing cycles, we sonified the suspension and centrifuged at 10,000 g and 4 °C for 30 min. The supernatant containing the desired protein was then purified on a HisTrap FF crude column using Äkta FPLC purifier system connected to a fraction collector (F9-R, GE Healthcare). Absorbance was measured at 280 nm and analyzed using Unicorn 6.3 software. The column was equilibrated with Tris-buffer (50 mM Tris-HCl, 500 mM NaCl, 4 M urea, 20 mM imidazole). Following the loading of the lysate, washing was performed (50 mM Tris-HCl, 500 mM NaCl, 4 M urea, 20 mM imidazole). His-tagged proteins were eluted using an elution buffer containing 500 mM imidazole and elution fractions were pooled. After dialysis of the pooled elution fractions into 1 mM Tris-HCl buffer pH 8 to remove salts and remaining compounds from His-tag purification, protein aliquots were lyophilized. For purity analysis and identification, standard SDS-PAGE (Tris/glycine,10 %) was performed. Lyophilized proteins were stored at -20 °C. For the most frequently used proteins we determined the protein concentration prior to dialysis via absorbance at 280 nm using the spectrophotometer Nanodrop 1,000 (Peqlab Biotechnology GmbH) and applying Beer-Lambert law. For all other proteins, required protein amounts were weighed immediately before hydrogel fabrication.

#### Protein purity and yield

SDS-PAGE of His-tag based purification via FPLC (**Figure S-2**) shows resulting protein solutions of acceptable purity. For all elution fractions of ULD-V80-ULD, an additional band of approx. 30 kDa is visible. Mass spectrometry confirmed that this additional band contains ULD; ELP sequences cannot be distinguished by MS. Given the length of 80 ELP repeats and given the fact that no additional bands of this size are seen in other ULD-ELP-ULD products, we assume that ULD-V80-ULD is able to form a circular shape through intramolecular dimerization of both ULD tails. Such circular shapes are assumed to show divergent electrophoretic kinetics in SDS-PAGE. Chain lengths of 20 or 40 repeats in ULD-V20-ULD and ULD-V40-ULD might not provide enough flexibility to form such circular shapes.

ULD-V20-ULD and ULD-V40-ULD provided the highest protein yield in shake flask culture, approximately 200 mg/L and 100 mg/L, respectively.

***2***

### ULD tetramerization

Wang and colleagues demonstrated that ULD peptides form stable tetramers, even under harsh conditions (2M NaCl and 10 % glycerol)1. As we routinely used urea buffers for protein purification (and in some cases for hydrogel preparation), we performed SDS-PAGE to investigate the extent of tetramerization under various conditions, including increasing urea concentrations, addition of ethanol and heating. A concentrated solution of purified ULD was mixed with varying buffers containing different mixtures of water, urea and ethanol and incubated for 2 hours before 1-4 µL 5x SDS loading buffer was added (see Figure below: 4 µL in lanes 1-10; 1 µL in lanes 11-18). Loading buffer contained 0,02 % Bromophenol blue, 30 % glycerol, 250 mM Tris-HCl pH 6.8, 10 % SDS and 0.5M DTT. If 1µL 5x SDS LB was used, we added 0.5 µL 60 % Glycerol additionally. Selected samples were then heated and transferred to the SDS-PAGE immediately or after 2-24 hours of cooling. ULD in water without heating served as control (**Figure S-3**, lane 13).

Ein Bild, das Text, weiß, Bild enthält.

Automatisch generierte Beschreibung

***Figure S-3: ULD tetramers and tetramer disruption with heating or addition of ethanol and urea, visualized via SDS-PAGE.*** To disrupt ULD tetramers effectively - even in the presence of SDS - either heating to 98 °C in water, heating to 90 °C in 4 M urea or the addition of 30 % ethanol to 4 M urea buffer was necessary. Neither up to 60 % glycerol, nor acidic or basic conditions rendered ULD monomers in aqueous solutions.

### PyMol

The crystallographic ULD structure was accessed via the Protein Data Bank with accession code 3TUO2 and loaded into PyMOL v2.3.3, Schrodinger, LLC. Using the dot representation, we visualized the estimated solvent accessible protein surface by increasing the van der Waals radii by 1.4 [Å](https://de.wikipedia.org/wiki/%C3%85)ngström, the approximate radius of a water molecule rolling over the molecule surface. Tyrosine pairs, i.e., two tyrosine residues in close proximity, were colored to visualize solvent (and catalyst) accessible tyrosine pairs.

### 

***4***

### Mass spectrometry

#### LC-MS-MS Analysis of tryptic peptides

Samples for LC-MS/MS analysis were prepared by trypsin digestion using a standard protein digestion protocol6-8. Extracted tryptic peptides were separated by LC chromatography (Agilent 1200 SL G1312B system) on C18 column AdvanceBio Peptide Mapping, 2.1 x 150 mm, 2.7 µm, LC column, kept at 45 °C. Peptides were separated over a linear gradient from 10-45 % acetonitrile in 0.1 % formic acid at a flow rate of 200 µl/min. Including column loading and washing steps, the total time for an LC-MS/MS run was 50 min. Online MS analysis was performed on Impact HD UHR Time-of Flight Mass Spectrometer System ion trap (Bruker Daltonics) equipped with Apollo II ion funnel electrospray source. The voltage applied to the capillary was set to 4.5 kV, the end plate offset was 500 V, the nebulizer pressure 1.8 bar dry gas 8 l/min and a dry temperature of 200 °C. Peptides were measured in positive MS ion mode using an enhanced resolution scan in a mass range of 150 – 2800 m/z, MS-MS/MS cycle time was set to 3 sec with 0.5 sec for MS and 0.25 sec summation time for MS/MS spectra. Singly-charged ions were excluded from fragmentation. MS/MS spectra were obtained using CID fragmentation, recorded in mass range initiating at 150-2800 m/z. The collision energy was adjusted between 23–65 eV as a function of the m/z value. Absolute threshold for MS/MS fragmentation was set to 1755 counts and active exclusion was set to 2 spectra which were released again after 1 min or if peak intensity increased 3-fold. Internal calibration via infusion of ESI Low Concentration Tune Mix was set at the start of the run.

#### Tryptic peptides: Data Processing and Bioinformatic Analysis

Data were analysed after recalibration using Data Analysis 4.2 SR2, Biotools 3.2 Peptide Editor 3.2 (Bruker Daltonics, Bremen) and MASCOT 2.5 (Matrix Science, London, UK) search engine. MASCOT 2.5 scored peptides for identification based on a search with an initial allowed mass deviation of the precursor ion of up to 15 ppm and allowed fragment mass deviation of 50 mDa. The search engine was used for the MS/MS spectra search against the UniProtKB *E. coli* database (downloaded Dec 21, 2015, containing 4305 entries, 116 contaminants and the added target proteins). The contaminants list used was downloaded from ftp.the pgm.org as of Jan 21, 2016. To determine the number of false positive peptide hits, data were searched against the full-length decoy database using the MASCOT 2.5 percolator function. The false discovery rate was set below 1 % to exclude false positive hits. Enzyme specificity was set as C-terminal to Arg and Lys, also allowing cleavage at proline bonds and a maximum of two missed cleavages. Deamidation of asparagine and glutamine carbamylation of lysine residues and the N-terminus and methionine oxidation were set as variable modifications.

#### LC-MS-MS Analysis of intact proteins

Intact proteins were separated by LC chromatography (Agilent 1200 SL G1312B system) XDB-C8 Zorbax Eclipse 4.6 mm × 30 mm, 3.5 μm, LC column, kept at 60 °C. Proteins were separated over a linear acetonitrile gradient in 0.1 % formic acid from 5-24 % over 3 min and 24 – 40% over 30 min at a flow rate of 250 µl/min. Including column loading and washing steps, the total time for an LC-MS/MS run was 66 min. Online MS analysis was performed on Impact HD UHR Time-of Flight Mass Spectrometer System ion trap (Bruker Daltonics) equipped with Apollo II ion funnel electrospray source. The voltage applied to the capillary was set to 4.5 kV, the end plate offset was 500 V, the nebulizer pressure 1.8 bar dry gas 8 l/min and a dry temperature of 200 °C. Internal calibration via infusion of ESI Low Concentration Tune Mix was set at the start of each run.

#### Intact proteins: Data Processing and Bioinformatic Analysis

After automated internal calibration, data were processed, and protein peaks were searched based on the 214 nm UV chromatogram and as extracted ion chromatograms. Subsequently, identified protein peaks were deconvoluted in the range of 30,000 – 40,000 Da using Bruker software Maximum Entropy Data Analysis 4.2 SR 2. Based on the given amino acid sequence the average mass and the isotopic pattern of ULD-V20-ULD and ULD-V40-ULD proteins were generated using Bruker data analysis.

#### LC-MS²-characterization results of building blocks

aywaed**5**

***5****edswe*

***6***

### 

### Hydrogel fabrication and crosslinking

Lyophilized protein aliquots with known protein quantities were dissolved in ddH2O, 4 M urea or PBS, depending on the protein and the desired application. In case of hydrogel pads for nanoindentation, ddH2O(for ULD-V20-ULD and ULD-V40-ULD) or 4 M urea (for ULD-V80-ULD and ULD-V120-ULD) was used. For corneal sealing experiments, ddH2O or PBS was used. As soon as the protein was dissolved homogeneously, riboflavin-phosphate (stock solution 50 mM) or Tris(bipyridine)ruthenium(II) chloride (ru(II)bpy; stock solution 10 mM) was added to a final concentration of 2.5 mM riboflavin or 0.1 mM ru(II)bpy, followed by brief vortexing and centrifugation, then ammonium peroxodisulfate (APS) (stock solution 1 M) was added to a final concentration of 30 mM, again followed by vortexing and centrifugation. For exposure of the non-photocrosslinked protein-solution and subsequent crosslinking, we used the Prizmatix / Mountain Photonics UHP-T-DI LED, a high power LED light source with a power of 5.5 W and a collimated light beam of 460 nm wavelength that is connected to a power control element to adjust the desired intensity. Exposure time, distance to the light source and light intensity were chosen according to the non-photocrosslinked protein-solution’s requirements and desired application. The exposure energy density can be calculated using ,

where *E* is the exposure energy, *A* is the exposed area, *I* is the chosen intensity, and *t* is the exposure time. The exposed area was calculated based on the sample-lamp distance and the LED diameter. The set intensity ranged from 40 % to 80 %, depending on the gel composition.

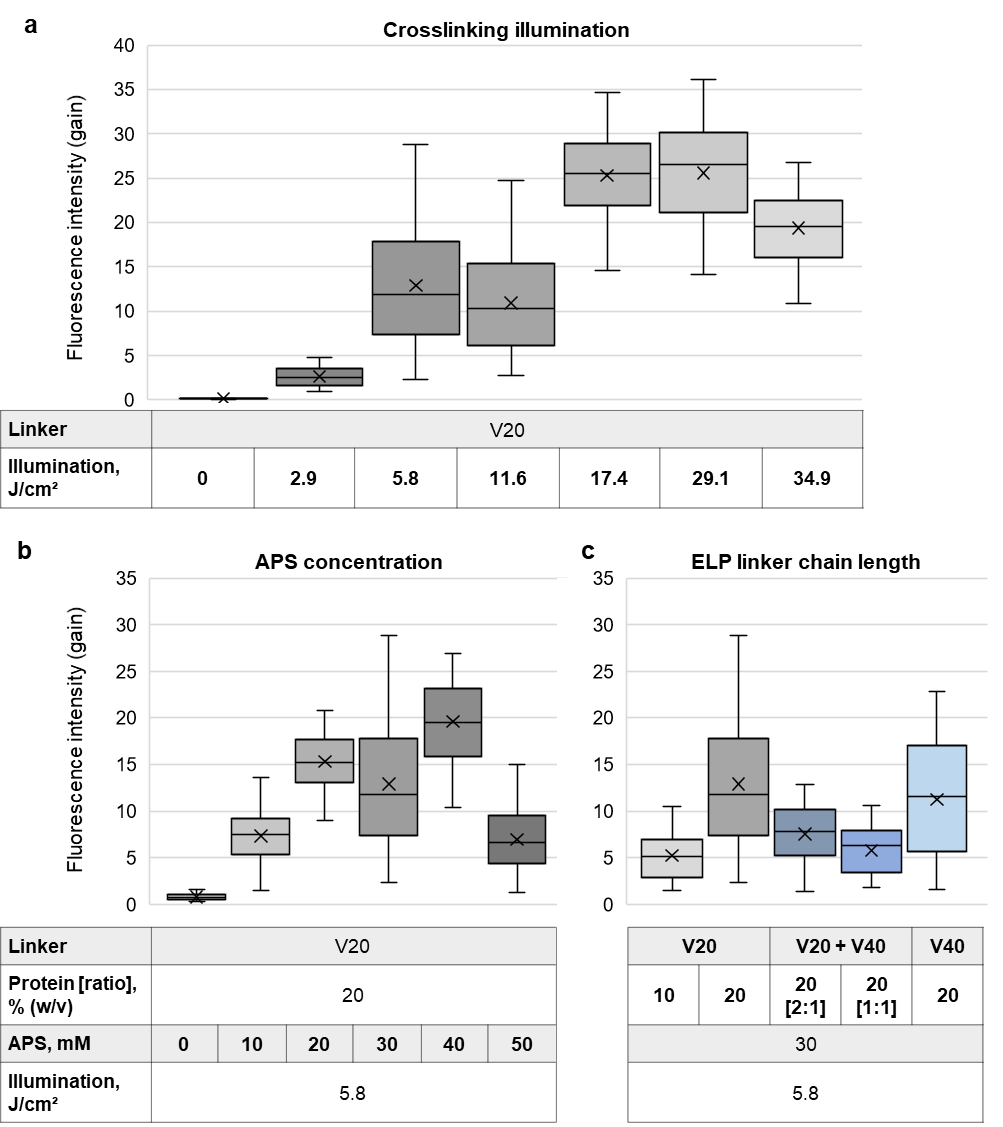
Ein Bild, das Text, Königin, Computer, Schreibtisch enthält.

Automatisch generierte Beschreibung

***Figure S-7: Preparation of hydrogel samples for characterization****.* For nanoindentation (A) and dynamic stretch experiments (B), we casted the non-photocrosslinked protein solution between two microscope slides and exposed the solution from both sides. To prevent sticking to the glass, the microscope slides were coated with a tetrafluoroethylene hexafluoropropylene copolymer foil. For fluorospectrometric measurements, we pipetted the non-photocrosslinked protein solution in a black 384-well plate.

### Dityrosine content

We determined dityrosine formation by utilizing its autofluorescence properties9-10. Fluorescence spectroscopy was performed using the SpectraMax iD5 Microplate Reader (Molecular Devices, LCC). 6 µL of non-photocrosslinked protein solution were pipetted in each well of a 384 well plate and subsequently exposed to light to induce crosslinking. Suitable excitation and emission wavelengths were identified based on a series of emission sweeps at excitation wavelengths of 260 nm to 340 nm, using a sample of ULD-V20-ULD, 20 % (w/v) protein in ddH2O, 2.5 mM riboflavin and 30 mM APS. As a result, we identified appropriate excitation and emission wavelengths at 320 nm and 455 nm, respectively. Applying these wavelengths, we determined relative amounts of dityrosine bonds within crosslinked hydrogels to identify targets for adjustment of the hydrogel’s properties11 (**Figure S-7**).

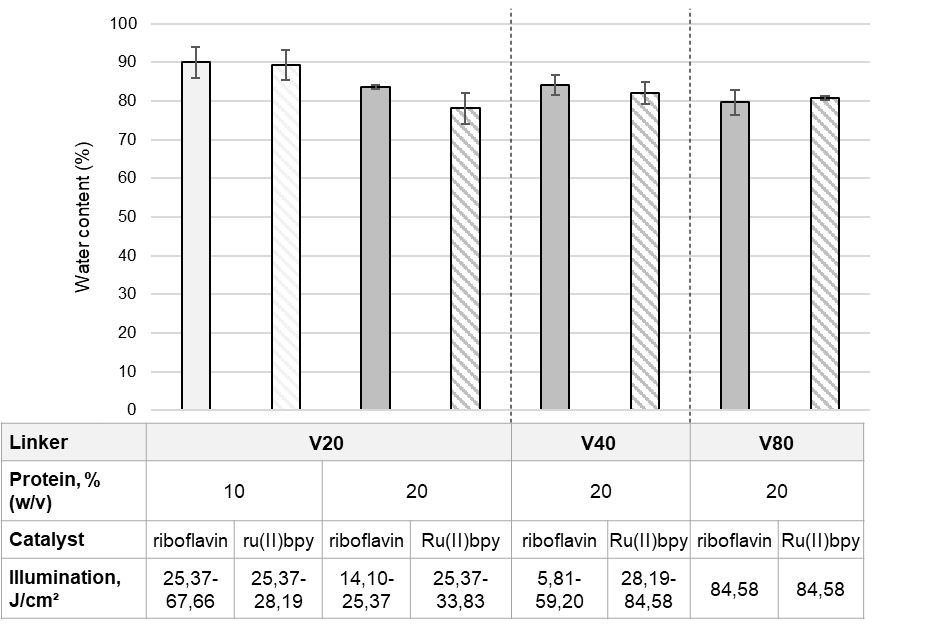
******

***Figure S-8: Dityrosine crosslink formation in ULD-Linker-ULD hydrogels depends on illumination energy, APS concentration and hydrogel composition****. For each measurement, 3-6 samples were prepared independently. The fluorescence intensity gain was measured at 12 central spots of each sample, as defined by a fixed matrix. a)**ULD-V20-ULD (20 % (w/v) protein prepared in ddH2O, 2.5 mM riboflavin, 30 mM APS) was exposed at 10 cm sample-lamp distance and exposure intensity 40 % for 15 to 180 seconds (equivalent to a total exposure energy density of 2.9 J/cm² to 34.9 J/cm²). b) ULD-V20-ULD (20 % protein prepared in ddH2O water, 2.5 mM riboflavin, exposed to 5.8 J/cm², 10 cm sample-lamp distance, exposure intensity 40 %, 30 sec) was crosslinked at different APS concentrations. Dityrosine formation benefits from rising APS concentrations up to 40 mM, higher APS concentrations seem to be detrimental. c) All hydrogel variants contained 2.5 mM riboflavin, 30 mM APS and were exposed to 5.8 J/cm². Mixed gels were composed of a 2:1 or 1:1 ratio of ULD-V20-ULD and ULD-V40-ULD, resulting in 20 % protein (w/v) in total.*

### Water content

19 µLprotein solution were pipetted on a coverslip, placed between two microscope slides with spacers and exposed to light from both sides. After letting the gels equilibrate in water for at least one day, we weighed the coverslip with the adherent gel after removing excess water using a lint-free cloth. The gel was then lyophilized, weighed again and removed from the coverslip to determine the coverslip’s weight. Water content was then calculated as

.

******

***Figure S-9: Water content of different ULD-ELP-ULD hydrogels, crosslinked with riboflavin or ru(II)bpy.*** For each linker type, 3-4 hydrogel samples were prepared and weighed in the hydrated and lyophilized state. There was no considerable difference in water content between gels crosslinked with riboflavin and with ru(II)bpy.

### Nanoindentation

For nanoindentation experiments, 19 µl protein solution was pipetted on a coverslip, placed between two microscope slides with spacers and exposed from both sides. The fabricated gel pads were 810 µm high and had a diameter of approx. 5.4 mm. Hydrogel pads on coverslips were measured while being submerged in water at ambient temperature using the Bioindenter (UNHT³ Bio, Anton Paar GmbH, Peseux, Switzerland) equipped with a ruby spherical tipped indenter (radius 500 µm). A load-controlled mode was chosen with nominal maximum load of 100 – 300 µN and constant loading and unloading rates of 300 µN/min, with or without a hold period of 60 s to observe creep behavior (time-dependent flow of incorporated fluid). The surface contact point was determined automatically, if a stiffness threshold of 2 µN/µm was exceeded and the measured normal force was at least 25 µN at the same time. For every measurement, correct determination of the contact point was confirmed and corrected manually, where necessary. For ULD-V20-ULD (10 %, 15 % and 20 % protein) as well as ULD-V40-ULD and ULD-V80-ULD (20 % each), measurements were carried out in triplicates. Every sample was measured at least 15 times in total. For each sample, 5 different surface contact points were determined based on a programmed matrix, and 3 indentation measurements were recorded in close proximity of each contact reference (matrix dimensions 1 mm x 1 mm, distance between measurements and respective contact reference 200 µm). Measurements were discarded, if no onset was evident (i.e., if measurement started after having already reached the sample surface), or if disturbances were noticed (e.g., if measurements were distorted by air bubbles or if the tip slipped off the edge of the gel pad). Young’s moduli were analyzed by the built-in Indentation software (v7.2.6) using the Hertz fit method12 for the loading portion of each load-displacement indentation curve and standard settings (fitting between 10 % and 98 % of maximum load (*F*max)). The Hertz formula yields the Young’s modulus according to the formula13:

Where *E* is the (reduced) Young’s modulus, *F* is the indentation load, *R* is the radius of the indenter tip, and *h* is the indentation depth. As the ruby tip is non-compressible, *E* equals the Young’s modulus.

***Table S-3: Fabrication parameters of ULD-ELP-ULD hydrogels measured via nanoindentation, number of measurements and resulting Young’s moduli from nanoindenter measurements.***

### AFM Analysis of Mechanical Properties - Young's Modulus and Resilience

Atomic force microscopy (AFM) measurements were performed with a MFP-3D-Bio (Asylum Research, an Oxford Instruments Company, Santa Barbara, USA). The samples were fixed in a fluid cell which is compatible for experiments in liquid with the solvent resistant and transparent UV-glue NOA63 (Edmund Optics GmbH, Mainz, Germany). The experiments were performed in ultrapure water (18.2 MΩ cm, Purelab Chorus 1, Elga LabWater, Celle, Germany). Between experiments the samples were dried or stored in 0.1% NaN3 (Sodium azide, 99%, Ferak laborant GmbH, Berlin, Germany) solution at 4 °C. Indentation experiments were performed using biosphere B1000-NCH (40.000 N/m) cantilevers (nanotools GmbH, München, Germany) comprising a spherical tip with a radius of 1 µm consisting of diamond like carbon (DLC). Before each experiment, the cantilever was calibrated by obtaining the inverse optical lever sensitivity (InvOLS) against a solid surface such as shown in Kolberg et al.15, while the nominal force constant of 40.000 N/m was taken. SiOx was taken from a wafer (Prime CZ-Si wafer, n-type (Phosphor) TTV<10 µm , MicroChem- icals GmbH, Ulm, Germany) and was cut into pieces using a diamond knife and immersed separately in microcentrifuge tubes filled with ethanol. The silicon oxide wafer pieces were sonicated (Elmasonic S15, Elma, Singen, Germany) for 10 min, rinsed twice in ethanol and dried under a nitrogen flow. The surfaces were cleaned just before the experiment started and used immediately.

For obtaining the Young’s modulus, up to two force maps (a grid of 10 x 10 points covering an area of 10 µm x 10 µm, velocities of 1 µm/s and 2 µm/s) as well as force-extension curves at the same sport (at least 100 curves covering a velocity range of 0.1 to 10 µm/s) were taken on different spots of the respective sample using a trigger force of 1 µN, a dwell time of 0, a sampling rate of 5 kHz and approx. 25°C. The Young’s modulus was determined by applying a Hertz fit for a spherical indenter12, 16 to the approach part of the force-extension curve representing an indentation of the cantilever tip into the underlying material. The Poisson ratio of the cantilever was set to 0.2 for the DLC tip17. For each sample altogether 300 indentation curves were taken for evaluation, fitting the first approx. 200 nm of the indentation curve (approach curve), which accounts to about 200-400 nm of indentation. This evaluation was performed by the custom-written procedures based in Igor Pro (Wave Metrics, USA) using a sequential search scheme via χ2-minimization for finding the contact point as given in Tschaikowsky et al.18 Proof-of-concept for the detection of early osteoarthritis pathology by clinically applicable endomicroscopy and quantitative AI-supported optical biopsy, Tschaikowsky, M. et al.18. The resilience was determined by calculating the ratio of the areas under the approach (indentation portion) and retraction curve as described in19. Differences in the (reduced) Young’s moduli reflected velocity dependencies and local variations (force maps). The resilience values were constant for the investigated velocity range and part of the sample.

***Table S-2:******AFM Data for the (reduced) Young’s modulus and resilience of ULD-V20-ULD (ru(II)bpy or riboflavin as catalyst)****. The error represents the sample standard deviation.*

|  |  |  |
| --- | --- | --- |
| **ULD-Linker-ULD** | **Young’s Modulus (kPa)** | **Resilience (%)** |
| ULD-V20-ULD Ru cat |  |  |
| 0.1 µm/s | 226 (± 33) | 79 (± 1) |
| 0.5 µm/s | 260 (± 37) | 80 (± 1) |
| 1.0 µm/s | 263 (± 27) | 78 (± 1) |
| 5.0 µm/s | 317 (± 40) | 80 (± 1) |
| 10.0 µm/s | 337 (± 62) | 79 (± 1) |
| Force Map (1.0 µm/s) | 302 (± 43) | 79 (± 1) |

|  |  |  |
| --- | --- | --- |
| ULD-V20-ULD Riboflavin cat |  |  |
| 0.1 µm/s | 148 (± 40) | 78 (± 1) |
| 0.5 µm/s | 205 (± 7) | 76 (± 1) |
| 1.0 µm/s | 199 (± 42) | 79 (± 1) |
| 5.0 µm/s | 301 (± 14) | 80 (±1) |
| 10.0 µm/s | 298 (± 13) | 80 (±1) |
| Force Map (2 µm/s) | 137 (± 53) | 79 (±1) |

### Dynamic Mechanical Analysis (DMA)

#### Experimental settings

Uniaxial tensile stretch experiments were carried out for ULD-V20-ULD, ULD-V20-RGD-ULD, ULD-V40-ULD and ULD-V40-RGD-ULD (20 % (w/v) protein each), crosslinked with 2.5 mM riboflavin and 30 mM APS and 67.7 J/cm². As high amounts of protein were required to cast these gels, proteins were dissolved in 4 M urea to ensure fast and homogeneous dissolving. For each of the four linker types, 5-6 samples were prepared independently. 580 µm thick gel films were cut to pieces of approx. 10 mm x 3-5 mm and measured at 37 °C in PBS buffer using the incubator-housed ElectroForce 5210 BioDynamic-Test-System (Texas Instruments, Dallas, Texas, USA; formerly Bose, Minnesota, USA). Measurements were started 5 minutes after fixing the sample chamber to the test system inside the incubator, to allow for temperature equilibration. Oscillating tensile stress was then applied at frequencies of 0.1 Hz (5 cycles) and 1 Hz (50 cycles), with the maximum displacement set to 20 %. To determine ultimate tensile stress and strain, gel pads were stretched at 50 µm/s until rupture.

#### DMA data processing and evaluation

Both the rupture experiment data and the dynamical stretch data were provided as raw data in csv-format, containing strain data as the independent variable and noisy stress data as the dependent variable. Measured force (stress) values were low, resulting in a low signal-to-noise ratio and requiring a denoising process to allow for data visualization and analysis. Data were denoised using a Savitzky-Golay Filter of window size and order . This specific filter type was chosen due to its extrema-preserving nature14. In the context of rupture experiments, the corresponding exact maximum strain value was then extracted by identifying the maximum index of the denoised stress data and allocating the corresponding strain value.

With regard to the sinusoidal dynamic stretch experiment data, the initial 4000 points were dropped for each data set to discard the mechanical setup phase (including calibration steps) and to allow gel samples to equilibrate in terms of protein chain re-arrangements and entanglements. This resulted in a total of 8000 data points per measurement, summing up to 200 stretching cycles.

The mean Young’s modulus was calculated for each subgroup based on the slope of the secant connecting the zero point and the maximum points in the stress-strain curve.

***Figure S-10:******Sequence of data processing****. Data sets for each measurement contained at least 12,000 data points, including the data points already recorded during the calibration and equilibration phase. The initial 4000 data points were hence discarded, leaving 4 subgroups of 2000 data points each, corresponding to 50 cycles per subgroup. Recorded force values were denoised via fast Fourier transform for each subgroup. Based on denoised force values, the corresponding stress was calculated and plotted as a function of the given displacement (= strain). Thus, for every recorded force-strain data point, the corresponding stress-strain value was plotted accordingly. Young’s modulus was calculated based on the slope of the secant of the resulting stress-strain curve for each subgroup of a given sample. Mean Young’s moduli (average of Young’s moduli of the four subgroups) are shown for each hydrogel sample in Table 1.*

After 200 stretch cycles, only slight deviations in the Young’s modulus were observed. Proceeding in time, both deviations to higher and to lower Young’s moduli have been observed, giving rise to the assumption that it’s not physical fatigue being shown, but rather external effects.

### Statistical analyses

Statistical analyses were performed using GraphPad Prism v9.2.0 or higher. Nanoindentation data were analyzed via nested one-way ANOVA (individual samples were nested per hydrogel type). Šídák’s post-hoc multiple comparison test was performed for prespecified group pairs to assess inter-group differences. DMA data were first analyzed via two-way ANOVA to assess the effect of different oscillatory frequencies (0.1 Hz and 1 Hz) and linker types (V20 vs. V40 vs. V20-RGD vs. V40-RGD). While the linker type yielded a significant difference (p=0.0078), the frequency did not (p=0.401), with pinteraction=0.370. Therefore, Young’s moduli obtained under different frequencies were pooled per linker type and data were analyzed via one-way ANOVA employing a Welch test and Dunnett’s T3 post-hoc multiple comparison test.

### Sealing of corneal tissue ruptures and stromal defects

#### Filling and sealing of cornea defects in porcine eyes

Extracted porcine eyes were obtained from a local butcher and were either frozen and thawed or were used on the day of slaughter. Fresh eyes were de-epithelialized before the procedure. The cornea was either incised with a scalpel or lamellar defects were introduced. To prevent liquid from leaking out through the applied incision, a 1 % solution of sodium hyaluronate (Z-HYALIN®) was injected directly underneath the cornea. Next, a paracentesis was set, and a small amount of air was injected into the anterior chamber, right behind the incision. Protein aliquots of ULD-V20-ULD and ULD-V40-ULD were dissolved in PBS buffer with 2.5 mM riboflavin and 30 mM APS, 2-4 µL of the freshly prepared non-photocrosslinked protein-solution were pipetted onto the corneal defect and illuminated with a total exposure energy density of 5.8 J/cm², comparable to values that are reported to be applied during keratoconus treatments20-21. An infusion bag was connected to an intraocular pressure sensor to measure and control intraocular pressures by varying the height of the infusion bag.

**1**

#### Filling and sealing of human corneal explants

Human corneae were received from the Lions Cornea Bank Baden-Württemberg (Freiburg, Germany) and were kindly provided by the Ophthalmology Department of the University Hospital Freiburg. Ethics approval for the use of corneal tissue was granted by the local Ethics Committee (Research Ethics Committee of the University of Freiburg). Explants were obtained from donors who had voluntarily agreed to donate organs, including eyes and parts thereof. Explants were only used for these experiments, if they were not fit for transplantation due to a low endothelial cell count, but otherwise intact.

Cornea explants were mounted on an artificial anterior eye chamber. Ulcer-like surface abrasions or penetrating vertical incisions were inflicted and subsequently sealed with ULD-V40-ULD with the catalyst riboflavin and crosslinked at a total exposure energy density of 5.8 J/cm². Sealed cornea explants were then kept in Modified Eagle’s Medium culture for 7 days (MEM, containing 10 % fetal bovine serum and penicillin/streptomycin). Cornea were embedded in paraffin and microtome cuts were prepared. The cuts underwent hematoxylin-eosin (HE) staining and were subsequently scanned to visualize cellular and extracellular matrix structures as well as the hydrogel filling.











### 

### 

### 

### 

### 

### 

# References

1. Wang, Z.; Yang, X.; Chu, X.; Zhang, J.; Zhou, H.; Shen, Y.; Long, J., The structural basis for the oligomerization of the N-terminal domain of SATB1. *Nucleic Acids Research* **2012,** *40* (9), 4193-4202.

2. Grote, A.; Hiller, K.; Scheer, M.; Münch, R.; Nörtemann, B.; Hempel, D. C.; Jahn, D., JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research* **2005,** *33* (suppl\_2), W526–W531.

3. Huber, M. C.; Schreiber, A.; Wild, W.; Benz, K.; Schiller, S. M., Introducing a combinatorial

DNA-toolbox platform constituting defined protein-based biohybrid-materials. *Biomaterials* **2014,** *35* (31), 8767-8779.

4. Sambrook, J. R., D. W. , *Molecular Cloning: A Laboratory Manual.* . 3rd ed.; Cold Spring Harbor Laboratory Press. : New York, 2001.

5. Quiroz, F. G.; Chilkoti, A., Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nature Materials* **2015,** *14* (11), 1164-1171.

6. Shevchenko, A. T., H.; Havli, J.; Olsen, J. V; Mann, M., In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes. *Nat. Protoc.* **2006,** *1* (6), 2856–2860.

7. Speicher, K. D. K., O.; Harper, S.; Speicher, D. W. . , Systematic Analysis of Peptide Recoveries from In-Gel Digestions for Protein Identifications in Proteome Studies. *J. Biomol. Tech.* **2000,** *11* (2), 74–86.

8. Geissinger, S. E.; Schreiber, A.; Huber, M. C.; Schiller, S. M., Adjustable Bioorthogonal Conjugation Platform for Protein Studies in Live Cells based on Artificial Compartments. *ACS Synthetic Biology* **2020,** *9* (4), 827-842.

9. Harms, G. S.; Pauls, S. W.; Hedstrom, J. F.; Johnson, C. K., Fluorescence and Rotational Dynamics of Dityrosine. *Journal of Fluorescence* **1997,** *7* (4), 283-292.

10. Partlow, B. P.; Applegate, M. B.; Omenetto, F. G.; Kaplan, D. L., Dityrosine Cross-Linking in Designing Biomaterials. *ACS Biomaterials Science and Engineering* **2016,** *2* (12), 2108-2121.

11. Fang, J.; Li, H., A facile way to tune mechanical properties of artificial elastomeric proteins-based hydrogels. *Langmuir* **2012,** *28* (21), 8260-8265.

12. Hertz, H., H. Hertz, Über die Berührung fester elastischer Körper, Journal für die reine und angewandte Mathematik 92, 156-171 (1881). *Journal für die reine und angewandte Mathematik* **1881,** *171*, 156-171.

13. Nohava, J.; Swain, M.; Lang, S. J.; Maier, P.; Heinzelmann, S.; Reinhard, T.; Eberwein, P., Instrumented indentation for determination of mechanical properties of human cornea after ultraviolet-A crosslinking. *Journal of Biomedical Materials Research - Part A* **2018,** *106* (5), 1413-1420.

14. Savitzky, A.; Golay, M. J. E., Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **1964,** *36* (8), 1627-1639.

15. Kolberg, A.; Wenzel, C.; Hugel, T.; Gallei, M.; Balzer, B. N., Covalent Attachment of Single Molecules for AFM-based Force Spectroscopy. *Journal of Visualized Experiments* **2020,** *157*, e60934.

16. Rico, F.; Roca-Cusachs, P.; Gavara, N.; Farré, R.; Rotger, M.; Navajas, D., Probing mechanical properties of living cells by atomic force microscopy with blunted pyramidal cantilever tips. *Physical Review E* **2005,** *72* (2), 021914.

17. Miyoshi, K., *Structures and mechanical properties of natural and synthetic diamonds.* NASA/TM--107249: Cleveland OH: Lewis Research Center, 1998.

18. Tschaikowsky, M.; Selig, M.; Brander, S.; Balzer, B. N.; Hugel, T.; Rolauffs, B., Proof-of-concept for the detection of early osteoarthritis pathology by clinically applicable endomicroscopy and quantitative AI-supported optical biopsy. *Osteoarthritis and Cartilage* **2021,** *29* (2), 269-279.

19. Huson, M. G.; Maxwell, J. M., The measurement of resilience with a scanning probe microscope. *Polymer Testing* **2006,** *25* (1), 2-11.

20. Wollensak, G., Crosslinking treatment of progressive keratoconus: New hope. *Current Opinion in Ophthalmology* **2006,** *17* (4), 356-360.

21. Randleman, J. B.; Khandelwal, S. S.; Hafezi, F., Corneal cross-linking. *Survey of Ophthalmology* **2015,** *60* (6), 509-523.

22. Sun, S., Inhibition of Protein Carbamylation in Urea Solution Using Ammonium Containing Buffers Shisheng. *Anal Biochem.* **2008,** *6* (9), 2166-2171.