Supplementary Information

Preorganized cyclic modules facilitate the self-assembly of protein nanostructures

*Jaka Snoj, Fabio Lapenta, Roman Jerala**
**Corresponding author: Roman Jerala, roman.jerala@ki.si*

This supplementary information includes:

- Supplementary Text
- Figs. S1 to S10
- Tables S1 to S7

Supplementary Text

Cloning

For cloning cySB6, cyRH1, cyRH2 a chimeric gene with a singular reading frame was constructed by positioning the gene for each subunit between the C-terminal and N-terminal parts of the split intein Gp41. Plasmid pCIRCgp41-1 a gift from Barbara Di Ventura & Roland Eils (Addgene plasmid # 74227; http://n2t.net/addgene:74227; RRID: Addgene_74227) containing C-terminal and N-terminal parts of the split intein Gp41 was amplified with primers to linearize it and introduce flanking regions CWE and RGK after C-terminal and before N-terminal parts of the split intein Gp41, respectively. For cloning cySB6 the backbone had complementary terminal sequences introduced using special primers creating pCIRCgp41-1a. Then the G block cySB6-TEV was cloned into the backbone by performing a Gibson reaction. Afterward, two more primers were used to remove the TEV sequence from one of the linkers and replace it with an "SGPG" sequence. G-blocks coding for either cyRH1 or cyRH2, each containing complementary terminal regions, were cloned into a linearized backbone pCIRCgp41-1b by performing a Gibson reaction.

To clone RH1 and RH2 constructs, plasmid pCIRCgp41-1 was amplified with primers in order to linearize it and remove the gp41 intein parts creating a backbone pCIRCgp41-1c. The flanking regions CWE and RGK were included at the termini of the inserts to ensure the aa sequence between linear and cyclic proteins was identical. G-block coding for either RH1 or RH2, each containing complementary terminal regions, were inserted into the backbone by performing a Gibson reaction. Additionally, g-block cyRH2-SS-Gp was inserted into the backbone pCIRCgp41-1c. Construct cyRH2-SS-Gp in comparison to cyRH2 contained a strep tag on either of the transcript's termini to enhance bacterial production. Constructs SB6, SB9b and SB9c were ordered as G-blocks containing appropriate complementary terminal sequences to be inserted in a linearized pET41a vector that was amplified by primers.

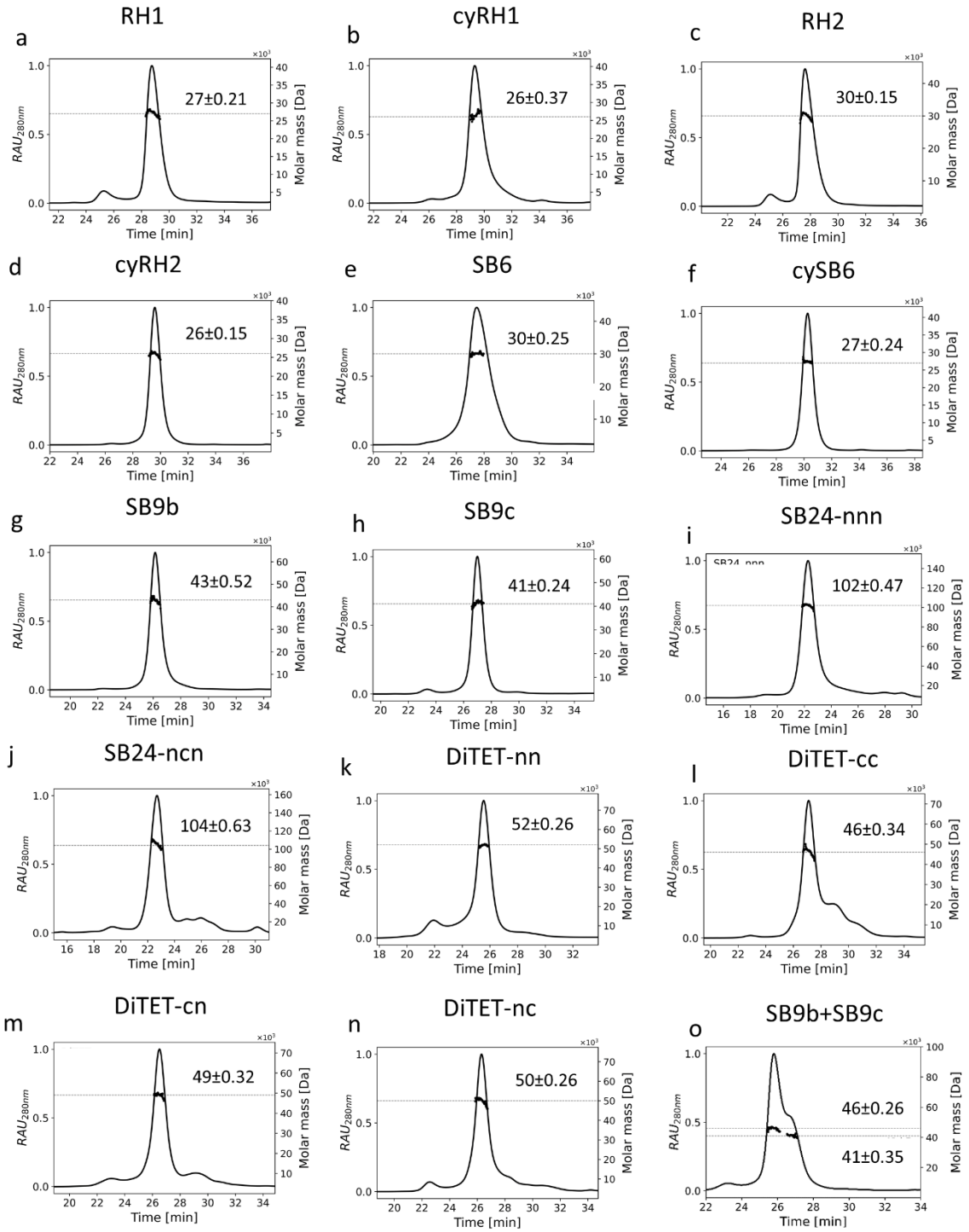

Figure S1. SEC-MALS chromatograms of individual proteins (a-h), protein complexes (i-n) and a mixture of two proteins (o). UV signal is reported in relative absorbance units (RAU). The molecular weight of the main peak calculated from light scattering is indicated on the panels (in kDa) and corresponds to the theoretical masses calculated from the amino acid sequence. Theoretical masses are listed in Table S1.

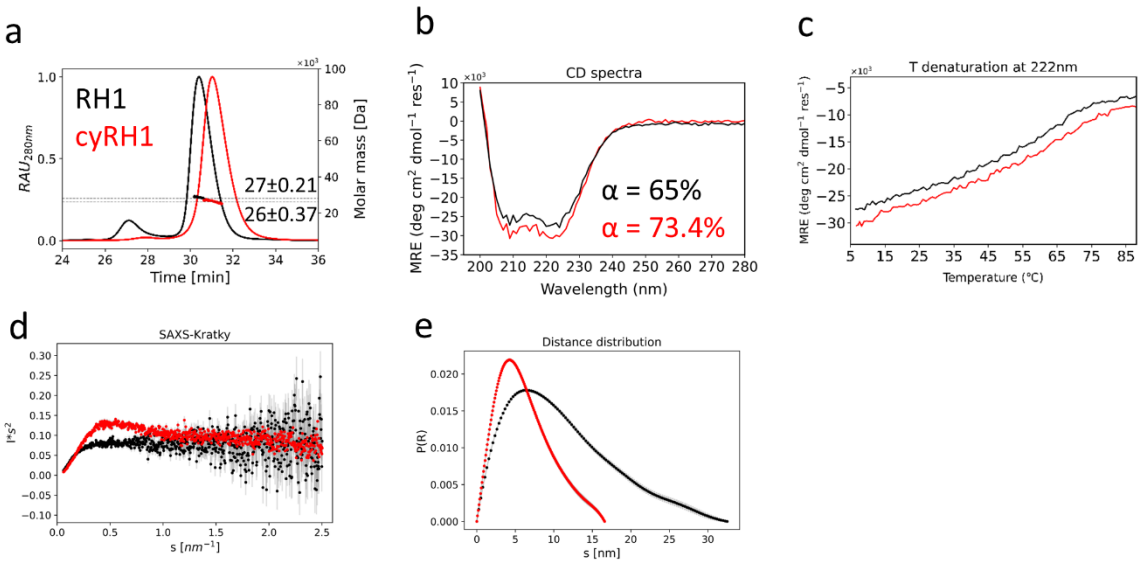

Figure S2. Biochemical characterization and comparison of the linear and cyclic subunit of a two-chain tetrahedron. On all the panels of the figure a noncyclic subunit RH1 is in black and a cyclic subunit cyRH1 is in red. (a) SEC-MALS chromatograms of RH1 and cyRH1 where the molecular weight of the peaks was calculated from light scattering and corresponds to the theoretical mass calculated from the amino acid sequence (theoretical Mw of RH1 =25.7 kDa and cyRH1 = 25.7 kDa). (b) Circular dichroism spectra of RH1 and cyRH1 at 5 °C. The helicity percentage is indicated on the panels. (c) CD signal at 222 nm of the proteins RH1 and cyRH1 during thermal denaturation. (d) Kratky plots as obtained by the SAXS experiments. Error bars in grey represent the standard deviation for each data point. (e) The pair-distance distribution function, P(r), as obtained by the SAXS experiments. The Dmax values of the proteins are RH1 = 32.5±0.1 nm and cyRH1 = 16.6±0.3 nm.

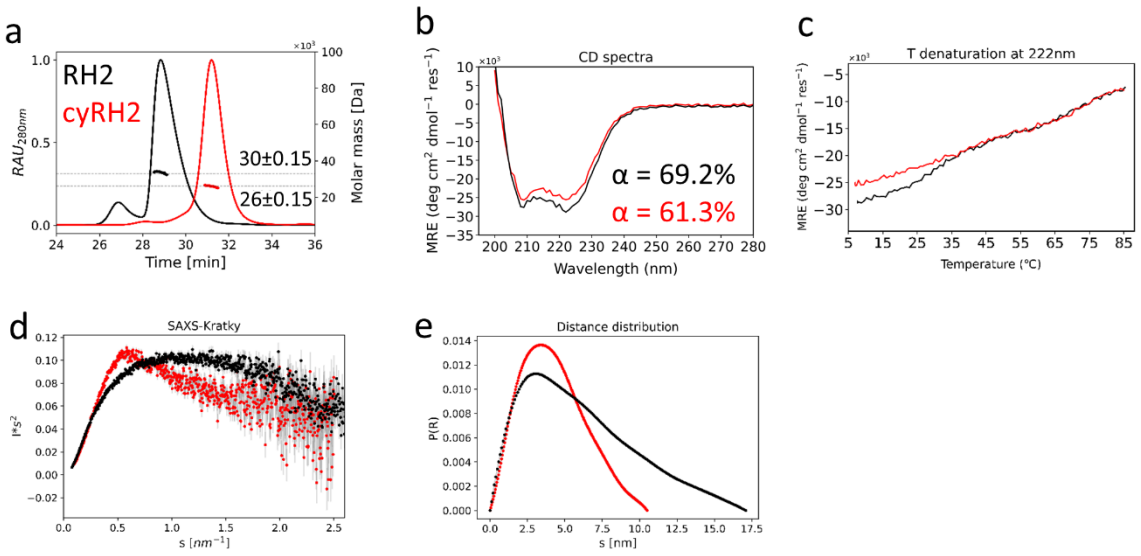

Figure S3. Biochemical characterization and comparison of the linear and cyclic subunit of a two-chain tetrahedron. On all the panels of the figure a noncyclic subunit RH2 is in black and a cyclic subunit cyRH2 is in red. (a) SEC-MALS chromatograms of RH2 and cyRH2 where the molecular weight of the peaks was calculated from light scattering and corresponds to the theoretical mass calculated from the amino acid sequence (theoretical Mw of RH2 =25.7 kDa and cyRH2 =25.7 kDa). (b) Circular dichroism spectra of RH2 and cyRH2 at 5 °C. The helicity percentage is indicated on the panels. (c) CD signal at 222 nm of the proteins RH2 and cyRH2 during thermal denaturation. (d) Kratky plots as obtained by the SAXS experiments. Error bars in grey represent the standard deviation for each data point. (e) The pair-distance distribution function, P(r), as obtained by the SAXS experiments. The Dmax values of the proteins are RH2 = 17.1±0.3 nm and cyRH2 = 10.5±0.1 nm.

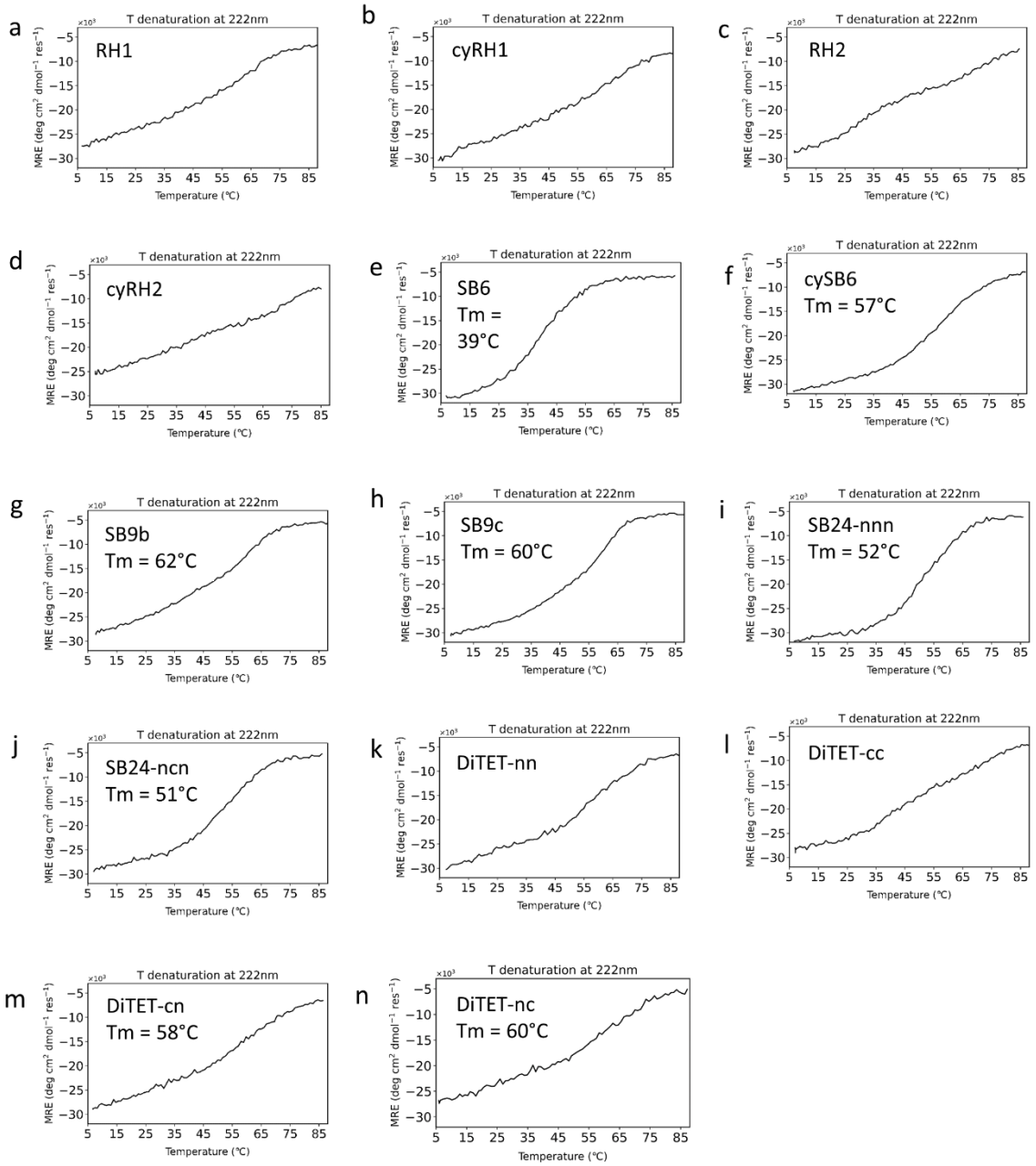

Figure S4. CD signal at 222 nm expressed in mean residue ellipticity (MRE) of individual proteins (a-h) and protein complexes (i-n) during thermal denaturation. If the melting temperature (Tm) could be determined, it is indicated in the panel.

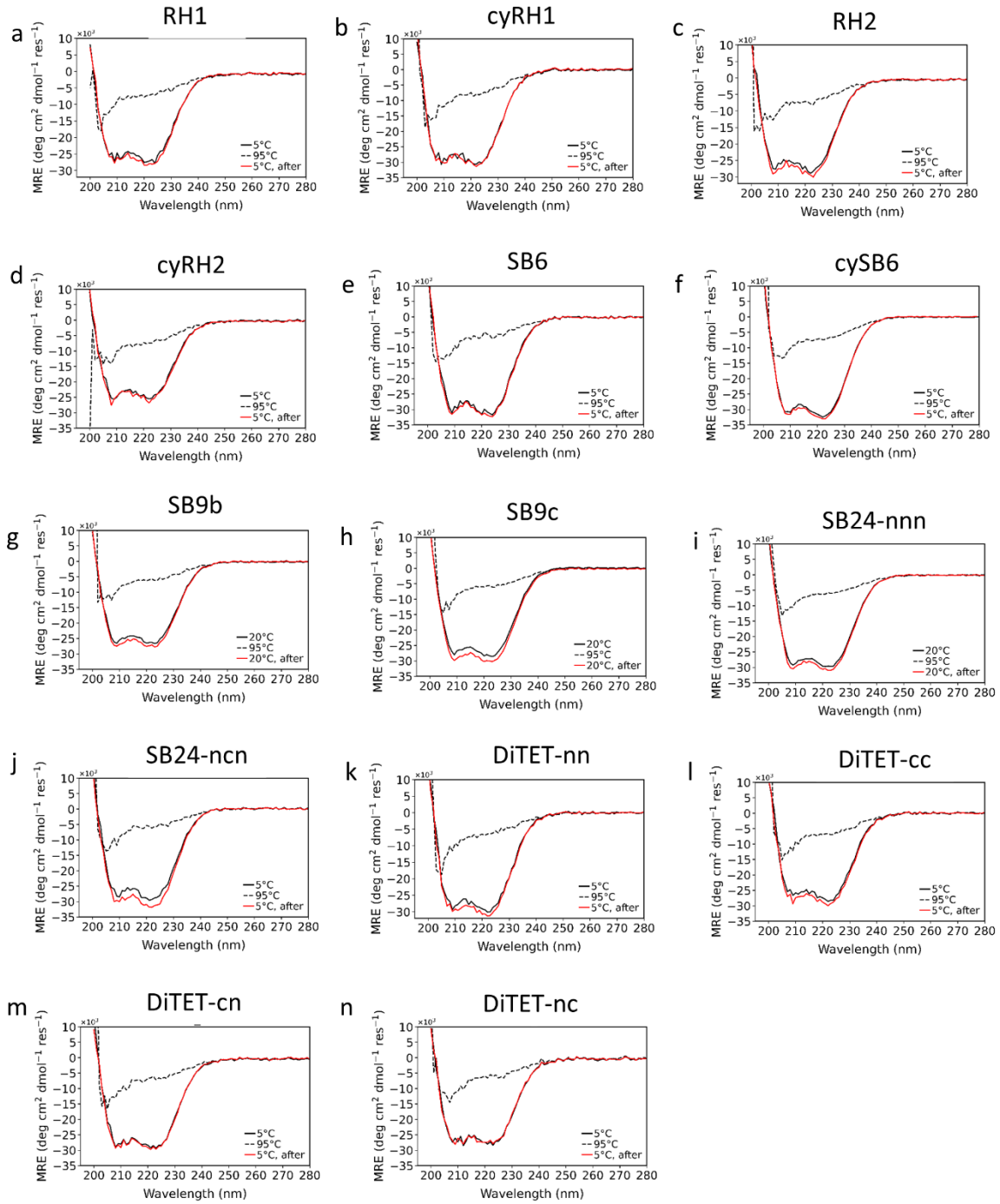

Figure S5. Circular dichroism spectra of individual proteins (a-h) and protein complexes (i-n). Spectra were measured at 20 °C (black), 95 °C (dotted black) and 20 °C (red) after refolding in panels g–i. All the rest spectra of the panels were measured at 5 °C (black), 95 °C (dotted black) and 5 °C (red) after refolding. The percentage of α helicity measured before thermal denaturation is listed in Table S1.

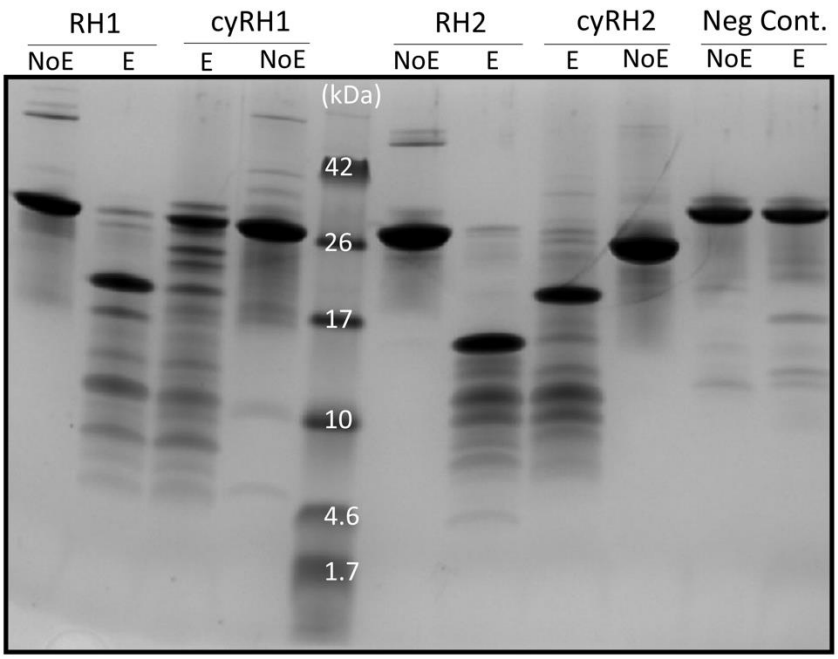

Figure S6. Protein cyclization with inteins confirmed by elastase digestion and electrophoretic analysis. Either cyclic (cyRH1, cyRH2) or linear protein (RH1, RH2) variants were cut with elastase peptidase and then ran on tricine SDS-PAGE. Elastase cut proteins at positions V or A and we observed the presence of a larger fragment in the case of cut cyclic variants and a smaller fragment in the case of a linear variant indicating successful cyclization. For the negative control (on the right of the gel) a CC-based 6-segment protein with no V or A was used.

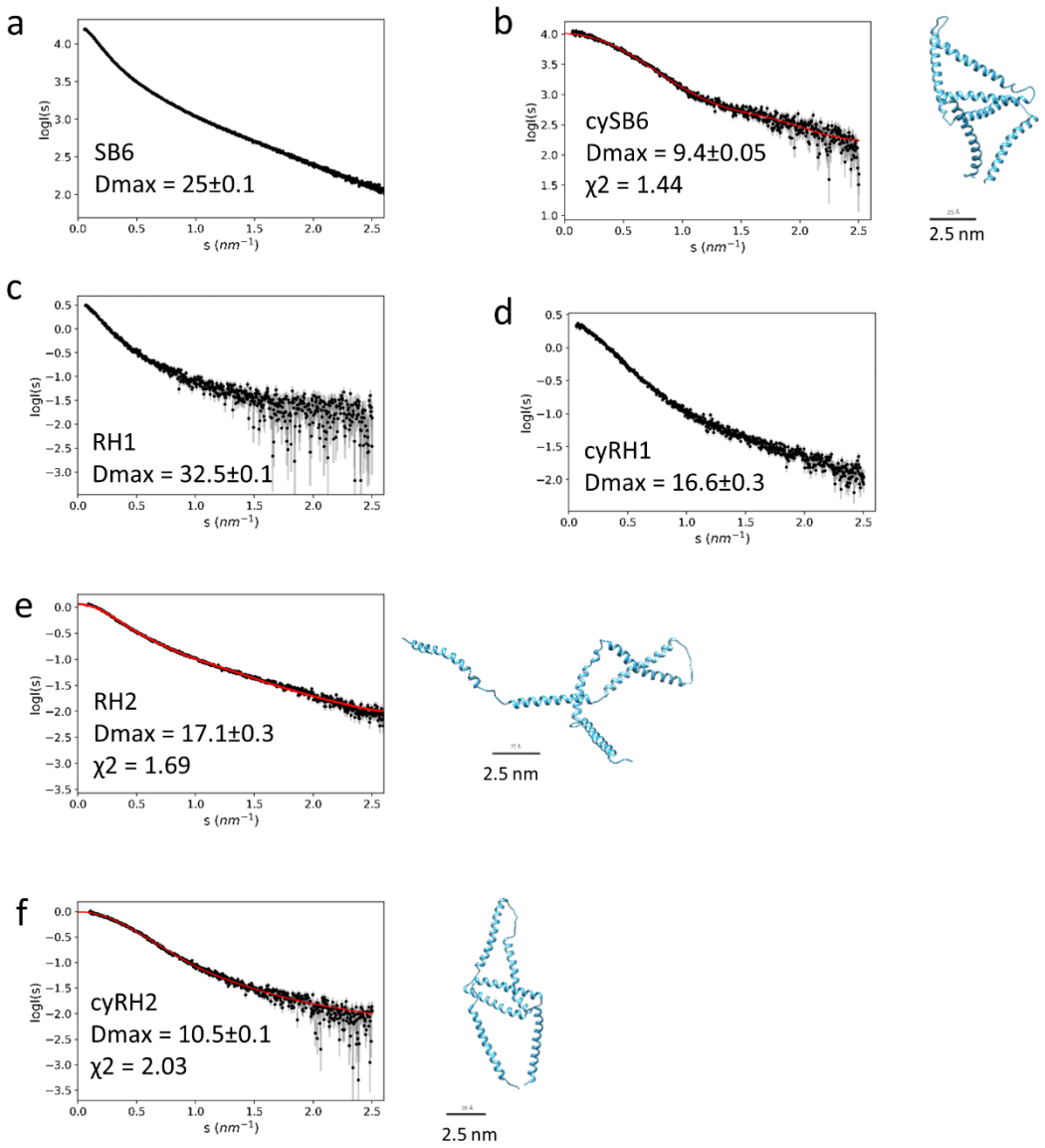

Figure S7. SAXS analysis of 6-segment long subunits. Panels show experimental SAXS profiles (black trace) of the proteins and Dmax calculated from a pair-distance distribution function. A good fit of the theoretical model scatter (red trace) to the experimental profile was determined in the case of cySB6 (b), RH2 (e), cyRH2 (f) with χ^2 = 1,44, χ^2 = 1.69 and χ^2 = 2.03, respectively. The theoretical model structures are shown next to the panels with fits.

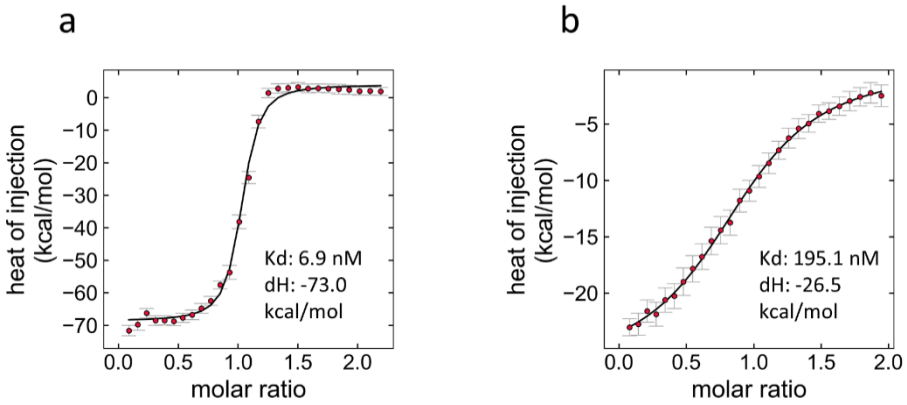

Figure S8. ITC measurements of SB24 complex formation. (a) Titrant (SB6) with a concentration of 12.36 µM was titrated into an equimolar mixture of SB9b and SB9c (each 1.33 µM). (b) Titrant (cySB6) with a concentration of 13.96 µM was titrated into an equimolar mixture of SB9b and SB9c (each 1.5 µM). Kd and enthalpy of the reaction are shown in the panel.

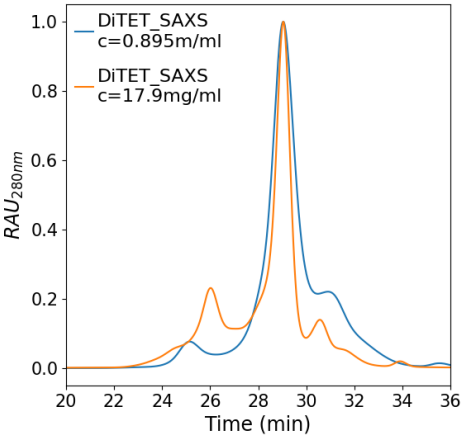

Figure S9. UV signal of SEC separation (increase 200) of complex DiTET-cc measured at 0.895 mg/ml (blue trace) and 17.9 mg/ml (orange trace). The comparison shows there is less shoulder to the right of the peak when protein complex is at high concentration.

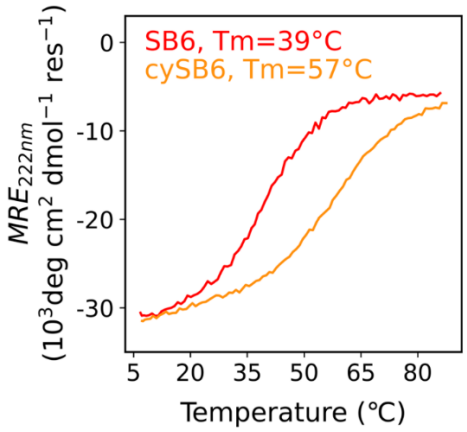

Figure S10. CD signal at 222 nm of the proteins SB6 and cySB6 during thermal denaturation. The melting temperatures (Tm) are indicated in the panel.

Table S1. List of proteins and complexes analyzed in this study. Theoretical mass was calculated from the amino acid sequence and then confirmed using SEC-MALS. α helicity was calculated from CD spectra of individual proteins measured at 20 °C (SB6, SB9b, SB9c) or 5 °C (all the rest). Dmax was determined from a pair-distance distribution function of SAXS results. N.a. indicates "not acquired".

Protein/complex	Theoretical molecular mass (kDa)	Mass determined by SEC-MALS (kDa)	α helicity determined by CD (%)	Dmax determined by SAXS (nm)
RH1	25.7	27±0.2	65.0	32.5±0.1
cyRH1	25.7	26±0.4	73.4	16.6±0.3
RH2	25.7	30±0.2	69.2	17.1±0.3
cyRH2	25.7	26±0.2	61.3	10.5±0.1
DiTET-cc	51.4	46±0.4	68.1	8.8±0.2
DiTET-nc	51.4	50±0.3	65.0	12.5±0.2
DiTET-cn	51.4	49±0.3	67.9	N.a.
DiTET-nn	51.4	52±0.3	71.6	14.7±0.1
SB6	26.1	30±0.3	73.7	25.1±0.1
cySB6	25.8	27±0.2	77.8	9.4±0.05
SB9b	42.2	43±0.5	62.2	N.a.
SB9c	40.6	41±0.2	67.1	N.a.
SB24-nnn	108.9	102±0.5	69.7	16.1±0.3
SB24-ncn	108.6	104±0.6	69.5	14.7±0.05

Table S2. Amino acid sequence of designed protein constructs.

Protein	Annotation and amino acid sequence
RH1 and cyRH1	Segments in order: AP10-GCN-P3mSN-P5SH-GCN-P7SH CWEMSPEDKLAQIKEKLQQIKEELAANEEKLQANKYGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGP GSPEDEIQQLEEEISQLEQKNSELKEKNQELKYGSHHHHHHSGSPEDENEKLEEKIWELKRKNEELKREIKEL EEGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGPGSPEDEIKELEWKNEELKREIKELEEKNEELKRR GK
RH2, cyRH2 and cyRH2-SS-Gp	Segments in order: P9mSN-GCN-AP4-P8SH-GCN-P6SH CWEMSPEDENQSLEQKNSQLKQEISQLEQEIQQLEYGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGP GSPEDELAANEEELQQNEQKLAQIKQKLQAIKYGSHHHHHHSGSPEDKIEELKRENEELEWKIEELKRENEEL EKGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGPGSPEDKNEELKREIKELEWENEELERKIEELKRR GK
cySB6-TEV	Segments in order: P6SHb-GCN-P6SHb-P8SHb-GCN-P8SHb CWEMSPEDKNEELKREIKELEWENEALERKIAELKRGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSE NLYFQGGSSPEDKNEELKREIKELEWENAELERKIEELKRGSHHHHHHGSSPEDKIAELKRENEELEYKIEEL KRENEALEKGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGPGSPEDKIEELKRENAELEYKIEELKRE NEALEKRGK
cySB6	Segments in order: P6SHb-GCN-P6SHb-P8SHb-GCN-P8SHb CWEMSPEDKNEELKREIKELEWENEALERKIAELKRGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGP GSPEDKNEELKREIKELEWENAELERKIEELKRGSHHHHHHGSSPEDKIAELKRENEELEYKIEELKRENEAL EKGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGPGSPEDKIEELKRENAELEYKIEELKRENEALEKR GK
SB6	Segments in order: P6SH-GCN-P6SH-P8SH-GCN-P8SH MSPEDKNEELKREIKELEWENEELERKIEELKRGSGPGQLEDKVEELLSKNYHLENEVERLKKLVGGSGPGSP EDKNEELKREIKELEWENEELERKIEELKRGSGPGSPEDKIEELKRENEELEWKIEELKRENEELEKGSGPGQ LEDKVEELLSKNYHLENEVERLKKLVGGSGPGSPEDKIEELKRENEELEWKIEELKRENEELEKLEHHHHHHH H
SB9b	Segments in order: P1mSN-P3mSN-P9SH-BCRmSH-P4mSN-P5SH-P2SmN-BCRmSH-P7SH MGHHHHHHHHHMENLYFQSGSGSPEDEIRQLEQENSQLERENQRLEQEIYQLERGSGPGSPEDEIQQLEEEISQ LEQKNSELKEKNQELKYGSGPGSPEDENEKLERKNEELKWEIKKLEREIKELERGSGPGDIEQELERAKQSIE ELEREVNQERSRMQYLQTRLSGSGPGSPEDKISQLKEKIQQLKQENQQLEEENSQLEYGSGPGSPEDENEKLE EKIWELKRKNEELKREIKELEEGSGPGSPEDKIEELKEKNSQLKEKNEELKQKIYELKEGSGPGDIEQELERA KQSIEELEREVNQERSRMQYLQTRLSGSGPGSPEDEIKELEWKNEELKREIKELEEKNEELKRLELE
SB9c	Segments in order: P1mSN-P3mSN-P7SH-BCRmSH-P4mSN-P5SH-P2mSN-BCRmSH-P10SH MSPEDEIRQLEQENSQLERENQRLEQEIYQLERGSGPGSPEDEIQQLEEEISQLEQKNSELKEKNQELKYGSG PGSPEDEIKELEWKNEELKREIKELEEKNEELKRGSGPGDIEQELERAKQSIEELEREVNQERSRMQYLQTRL SGSGPGSPEDKISQLKEKIQQLKQENQQLEEENSQLEYGSGPGSPEDENEKLEEKIWELKRKNEELKREIKEL EEGSGPGSPEDKIEELKEKNSQLKEKNEELKQKIYELKEGSGPGDIEQELERAKQSIEELEREVNQERSRMQY LQTRLSGSGPGSPEDKNKELKEENKELEWKIEELKEKIKELKELEHHHHHHHH
TRI6SN*	GSPEDEIRQLEQENSQLERENQRLEQEIYQLERGSGPGSPEDENSQLEEKISQLKQKNSELKEEIQQLEYGSG PGSPEDKISQLKEKIQQLKQENQQLEEENSQLEYGSGPGSPEDKIEELKEKNSQLKEKNEELKQKIYELKEGS GPGSPEDKNSELKEEIQQLEEENQQLEEKISELKYGSGPGSPEDEIQQLEEEISQLEQKNSELKEKNQELKY

*Negative control for elastase digestion (sequence does not contain amino acids V and A)

Table S3. Coiled-coil building-blocks. Orthogonal dimer-forming CC units used in constructs. The column on the right indicates in which construct was the particular building block used.

CC	Sequence	Constructs
P1mSN	SPED EIRQLEQ ENSQLER ENQRLEQ EIYQLER G	SB9b, SB9c
P2mSN	SPED KIEELKE KNSQLKE KNEELKQ KIYELKE G	SB9b, SB9c
P3mSN	SPED EIQQLEE EISQLEQ KNSELKE KNQELKY G	SB9b, SB9c, RH1, cyRH1
P4mSN	SPED KISQLKE KIQQLKQ ENQQLEE ENSQLEY G	SB9b, SB9c
AP4	SPED ELAANEE ELQQNEQ KLAQIKQ KLQAIKY G	RH2, cyRH2
P5SH	SPED ENEKLEE KIWELKR KNEELKR EIKELEE	SB9b, SB9c, RH1, cyRH1
P6SH	SPED KNEELKR EIKELEW ENEELER KIEELKR	SB6, RH2, cyRH2
P6SHb	SPED KNEELKR EIKELEW ENEALER KIAELKR	cyBS6
P7SH	SPED EIKELEW KNEELKR EIKELEE KNEELKR	SB9b, SB9c, RH1, cyRH1
P8SH	SPED KIEELKR ENEELEW KIEELKR ENEELEK	SB6, RH2, cyRH2
P8SHb	SPED KIAELKR ENEELEY KIEELKR ENEALEK	cySB6
P9mSN	SPED ENQSLEQ KNSQLKQ EISQLEQ EIQQLEY G	RH2, cyRH2
P9SH	SPED ENEKLER KNEELKW EIKKLER EIKELER	SB9b
P10SH	SPED KNKELKE ENKELEW KIEELKE KIKELKE	SB9c
AP10	SPED KLAQIKE KLQQIKE ELAANEE KLQANKY G	RH1, cyRH1
GCN	QLED KVEELLS KNYHLEN EVERLKK LVG	SB6, cySB6 RH1, cyRH1, RH2, cyRH2
BCRmSH	DIEQ ELERAKQ SIEELER EVNQERS RMQYLQT RLS	SB9b, SB9c

Table S4. List of DNA sequences of the backbones used for cloning constructs for split intein cyclization. "/ /" indicates the position where the g-block was cloned into the backbone, underlined nucleotides indicate the location of the introduced CWE and RGK flanking regions, the blue color indicates C-terminal part of the Gp41 intein, the orange color indicates N-terminal part of the Gp41 intein.

Backbone name	DNA sequence
pCIRCgp41-1a	TCCTTAGCTTTCGCTAAGGATGATTTCTGGAATTTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCGAATTCGCGGCCGCTTCTAGAGAATA ATTTTGTTTAACTTTAAGAAGGAGATACTAGATGATGCTGAAAAAAATCCTGAAAATCGAAGAGCTGGATGAACGTGAACTGATCGATATTGAGGTGTCCGGTA ACCACCTGTTTTACGCTAACGATATTCTGACCCACAACTGTTGGGAAATGTCTCCAGAAGACAAAAAC / / AAACGAGGCACTTGAAAAACGCGGTAAGTGCTTGGATCTGAAAACCCAGGTTCAGACCCCGCAGGGTATGAAGGAAATTTCCAACATCCAGGTCGGTGATCTG GTACTGAGCAACACGGGTTACAACGAAGTTCTGAACGTCTTCCCGAAATCTAAAAAAAAGTCTTACAAAATCACCCTGGAAGATGGCAAGGAAATCATCTGTTCC GAAGAACACCTGTTTCCGACGCAGACTGGTGAAATGAACATCTCCGGTGGTCTGAAAGAAGGTATGTGTCTGTATGTTAAAGAATAATAAATCGGTGAAATGC ACGACTGATAGTACTAGTAGCGGCCGCTGCAGTCCGGCAAAAAAGGGCAAGGTGTCACCACCCTGCCCTTTTTCTTTAAAACCGAAAAGATTACTTCG CGTTATGCAGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCC TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG AACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTT GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGG GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA GCTCGAGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTC ATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCT GCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGA ATGGCAAAAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGC GCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTC ACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGT CGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCG GGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCT GGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAAT GTAACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGCTCGAGTGCCACCTGACGTCTAAGAAACCATT ATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCAGAATTTCAGATAAAAAAAA
pCIRCgp41-1b	TCCTTAGCTTTCGCTAAGGATGATTTCTGGAATTTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCGAATTCGCGGCCGCTTCTAGAGAATA ATTTTGTTTAACTTTAAGAAGGAGATACTAGATGATGCTGAAAAAAATCCTGAAAATCGAAGAGCTGGATGAACGTGAACTGATCGATATTGAGGTGTCCGGTA ACCACCTGTTTTACGCTAACGATATTCTGACCCACAACTGTTGGGAAATG / / CGCGGTAAGTGCTTGGATCTGAAAACCCAGGTTCAGACCCCGCAGGGTATGAAGGAAATTTCCAACATCCAGGTCGGTGATCTGGTACTGAGCAACACGGGTTA CAACGAAGTTCTGAACGTCTTCCCGAAATCTAAAAAAAAGTCTTACAAAATCACCCTGGAAGATGGCAAGGAAATCATCTGTTCCGAAGAACACCTGTTTCCGAC GCAGACTGGTGAAATGAACATCTCCGGTGGTCTGAAAGAAGGTATGTGTCTGTATGTTAAAGAATAATAAATCGGTGAAATGCACGACTGATAGTACTAGTA GCGGCCGCTGCAGTCCGGCAAAAAAGGGCAAGGTGTCACCACCCTGCCCTTTTTCTTTAAAACCGAAAAGATTACTTCGCGTTATGCAGGCTTCCTCGC TCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCG ACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGC TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATC AAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGCTCGAGTCCCGTCAAGTCA GCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCA TATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACA TCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATGCATTT CTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATAC GCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTC TAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCG TCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAG ATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTT GAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAG ACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGCTCGAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTAT AAAAATAGGCGTATCACGAGGCAGAATTTCAGATAAAAAAAA

Table S5. List of DNA sequences of the backbones used for cloning noncyclic constructs. "/ /" indicates the position where the g-block was cloned into the backbone.

Name	DNA sequence
pCIRCgp41-1c	TCCTTAGCTTTCGCTAAGGATGATTTCTGGAATTTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCGAATTCGCGGCCGCTTCTAGAGAATAATTTTGTTTAACTTTAAGAAGGAGATACTAG / / ATCGGTGAAATGCACGACTGATAGTACTAGTAGCGGCCGCTGCAGTCCGGCAAAAAAGGGCAAGGTGTCACCACCCTGCCCTTTTTCTTTAAAACCGAAAAGATTACTTCGCGTTATGCAGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCACAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGCTCGAGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGGTTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCTAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCGATCTCGATCCCGCGAAATTAA	
pET41a	TACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACACCATGGG / / AGTTTGAGAAATAATTGATTAATACCTAGGCTGCTAAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGGTTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCTAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCGATCTCGATCCCGCGAAATTAA	

Table S6. List of DNA sequences of the inserted fragments.

Name	DNA sequence
cySB6-TEV	ATGTCTCCAGAAGACAAAAACGAGGAACTGAAACGTGAAATTAAGGAGCTTGAATGGGAAAACGAAGCCTTAGAACGTAAGATCGCAGAGTTGAAGCGCGGCTCTGGGCCA GGCCAATTAGAAGATAAGGTTGAGGAGCTGTTAAGCAAAAATTACCACTTGGAAAATGAAGTGGAACGTTTAAAAAAATTAGTTGGTGGTGGGTCGGAGAACTTGTATTTC CAGGGGGGCAGCAGTCCTGAGGATAAAAATGAAGAGCTGAAGCGTGAAATCAAGGAACTTGAGTGGGAGAATGCTGAGTTGGAACGCAAAATCGAGGAACTGAAGCGCGGT TCGCACCACCATCATCATCACGGCTCTTCGCCGGAAGACAAGATTGCAGAGCTTAAGCGCGAGAATGAAGAACTGGAGTATAAGATTGAAGAACTTAAACGCGAAAATGAG GCACTGGAAAAAGGTTCCGGCCCCGGGCAGCTGGAGGACAAAGTTGAGGAATTGCTGTCTAAGAATTATCACCTTGAAAATGAGGTGGAACGCTTAAAGAAGCTTGTCGGC GGCAGTGGGCCTGGTTCACCGGAAGACAAAATTGAGGAACTGAAACGCGAAAACGCAGAATTGGAATATAAAATTGAGGAGCTTAAGCGCGAAAACGAGGCACTTGAAAAA
cySB6	TGTTGGGAAATGTCTCCAGAAGACAAAAACGAGGAACTGAAACGTGAAATTAAGGAGCTTGAATGGGAAAACGAAGCCTTAGAACGTAAGATCGCAGAGTTGAAGCGCGGC TCTGGGCCAGGCCAATTAGAAGATAAGGTTGAGGAGCTGTTAAGCAAAAATTACCACTTGGAAAATGAAGTGGAACGTTTAAAAAAATTAGTTGGTGGTTCAGGACCTGGC AGTCCTGAGGATAAAAATGAAGAGCTGAAGCGTGAAATCAAGGAACTTGAGTGGGAGAATGCTGAGTTGGAACGCAAAATCGAGGAACTGAAGCGCGGTTCGCACCACCAT CATCATCACGGCTCTTCGCCGGAAGACAAGATTGCAGAGCTTAAGCGCGAGAATGAAGAACTGGAGTATAAGATTGAAGAACTTAAACGCGAAAATGAGGCACTGGAAAAA GGTTCCGGCCCCGGGCAGCTGGAGGACAAAGTTGAGGAATTGCTGTCTAAGAATTATCACCTTGAAAATGAGGTGGAACGCTTAAAGAAGCTTGTCGGCGGCAGTGGGCCT GGTTCACCGGAAGACAAAATTGAGGAACTGAAACGCGAAAACGCAGAATTGGAATATAAAATTGAGGAGCTTAAGCGCGAAAACGAGGCACTTGAAAAACGCGGTAAG
cyRH1	ACCCACAACTGTTGGGAAATGTCACCAGAAGATAAATTGGCACAAATCAAGGAAAAGTTACAGCAAATCAAAGAGGAGCTGGCTGCCAACGAGGAAAAGCTGCAGGCTAAC AAATATGGGAGTGGACCTGGGCAGTTGGAAGATAAGGTCGAGGAATTACTGAGTAAAAACTATCACTTGGAGAACGAGGTTGAACGCCTGAAAAAGTTGGTCGGCGGCAGT GGACCCGGCAGCCCGGAGGACGAGATCCAACAATTAGAAGAAGAAATTTCCCAACTGGAACAAAAGAACTCTGAATTGAAAGAGAAGAACCAGGAGTTGAAATACGGATCT CATCATCATCACCATCACGGGTCCTCACCAGAAGACGAAAACGAGAAGCTTGAAGAGAAGATTTGGGAGCTTAAGCGCAAGAACGAAGAGCTGAAGCGTGAGATTAAGGAG CTGGAGGAGGGATCTGGGCCTGGACAGTTAGAAGACAAAGTTGAAGAGTTGCTGTCCAAGAACTACCATTTGGAAAATGAGGTGGAGCGTCTGAAGAAACTGGTGGGTGGG AGCGGCCCGGGTAGTCCAGAGGACGAGATCAAAGAGTTGGAATGGAAAAATGAGGAATTGAAACGTGAAATCAAAGAGTTAGAGGAAAAAAACGAGGAGCTGAAACGCCGC GGTAAGTGCTTGGATCT
cyRH2	ACCCACAACTGTTGGGAAATGTCACCAGAAGACGAAAATCAGTCTTTAGAGCAAAAAAACAGCCAGCTTAAGCAAGAGATTAGTCAACTGGAGCAAGAAATTCAGCAGCTT GAATATGGGAGTGGACCTGGGCAGTTGGAAGATAAGGTCGAGGAATTACTGAGTAAAAACTATCACTTGGAGAACGAGGTTGAACGCCTGAAAAAGTTGGTCGGCGGCAGT GGACCCGGCTCACCTGAAGACGAATTAGCTGCTAACGAGGAAGAGTTACAGCAAAATGAACAAAAGTTGGCTCAAATCAAACAAAAGCTGCAGGCTATTAAATACGGATCT CATCATCATCACCATCACGGGTCCAGTCCCGAGGATAAGATTGAAGAACTGAAACGCGAGAATGAAGAGTTGGAGTGGAAAATCGAAGAATTAAAGCGTGAGAATGAGGAA TTGGAGAAGGGATCTGGGCCTGGACAGTTAGAAGACAAAGTTGAAGAGTTGCTGTCCAAGAACTACCATTTGGAAAATGAGGTGGAGCGTCTGAAGAAACTGGTGGGTGGG AGCGGCCCGGGTTCACCCGAAGATAAAAACGAGGAATTAAAGCGCGAAATCAAAGAATTAGAGTGGGAAAATGAAGAGCTTGAACGCAAGATTGAGGAGCTTAAGCGTCGC GGTAAGTGCTTGGATCT
cyRH2-SS-Gp	AATTTTGTTTAACTTTAAGAAGGAGATACTAGATGagcgcgtggagccatccgcagtttgaaaaaGGCTCAGGATCGGGAATGCTGAAAAAAATCCTGAAAATCGAAGAGC TGGATGAACGTGAACTGATCGATATTGAGGTGTCCGGTAACCACCTGTTTTACGCTAACGATATTCTGACCCACAACTGTTGGGAAATGTCACCAGAAGACGAAAATCAGT CTTTAGAGCAAAAAAACAGCCAGCTTAAGCAAGAGATTAGTCAACTGGAGCAAGAAATTCAGCAGCTTGAATATGGGAGTGGACCTGGGCAGTTGGAAGATAAGGTCGAGG AATTACTGAGTAAAAACTATCACTTGGAGAACGAGGTTGAACGCCTGAAAAAGTTGGTCGGCGGCAGTGGACCCGGCTCACCTGAAGACGAATTAGCTGCTAACGAGGAAG AGTTACAGCAAAATGAACAAAAGTTGGCTCAAATCAAACAAAAGCTGCAGGCTATTAAATACGGATCTCATCATCATCACCATCACGGGTCCAGTCCCGAGGATAAGATTG AAGAACTGAAACGCGAGAATGAAGAGTTGGAGTGGAAAATCGAAGAATTAAAGCGTGAGAATGAGGAATTGGAGAAGGGATCTGGGCCTGGACAGTTAGAAGACAAAGTTG AAGAGTTGCTGTCCAAGAACTACCATTTGGAAAATGAGGTGGAGCGTCTGAAGAAACTGGTGGGTGGGAGCGGCCCGGGTTCACCCGAAGATAAAAACGAGGAATTAAAGC GCGAAATCAAAGAATTAGAGTGGGAAAATGAAGAGCTTGAACGCAAGATTGAGGAGCTTAAGCGTCGCGGTAAGTGCTTGGATCTGAAAACCCAGGTTCAGACCCCGCAGG GTATGAAGGAAATTTCCAACATCCAGGTCGGTGATCTGGTACTGAGCAACACGGGTTACAACGAAGTTCTGAACGTCTTCCCGAAATCTAAAAAAAAGTCTTACAAAATCA CCCTGGAAGATGGCAAGGAAATCATCTGTTCCGAAGAACACCTGTTTCCGACGCAGACTGGTGAAATGAACATCTCCGGTGGTCTGAAAGAAGGTATGTGTCTGTATGTTA AAGAAGGTTCTGGTAGCGGAAGCGCTTGGAGCCATCCGCAATTCGAAAAATAATAAATCGGTGAAATGCACGACTG
RH1	AATTTTGTTTAACTTTAAGAAGGAGATACTAGATGTGTTGGGAAATGTCACCAGAAGATAAATTGGCACAAATCAAGGAAAAGTTACAGCAAATCAAAGAGGAGCTGGCTG CCAACGAGGAAAAGCTGCAGGCTAACAAATATGGGAGTGGACCTGGGCAGTTGGAAGATAAGGTCGAGGAATTACTGAGTAAAAACTATCACTTGGAGAACGAGGTTGAAC GCCTGAAAAAGTTGGTCGGCGGCAGTGGACCCGGCAGCCCGGAGGACGAGATCCAACAATTAGAAGAAGAAATTTCCCAACTGGAACAAAAGAACTCTGAATTGAAAGAGA AGAACCAGGAGTTGAAATACGGATCTCATCATCATCACCATCACGGGTCCTCACCAGAAGACGAAAACGAGAAGCTTGAAGAGAAGATTTGGGAGCTTAAGCGCAAGAACG AAGAGCTGAAGCGTGAGATTAAGGAGCTGGAGGAGGGATCTGGGCCTGGACAGTTAGAAGACAAAGTTGAAGAGTTGCTGTCCAAGAACTACCATTTGGAAAATGAGGTGG AGCGTCTGAAGAAACTGGTGGGTGGGAGCGGCCCGGGTAGTCCAGAGGACGAGATCAAAGAGTTGGAATGGAAAAATGAGGAATTGAAACGTGAAATCAAAGAGTTAGAGG AAAAAAACGAGGAGCTGAAACGCCGCGGTAAGTAATAAATCGGTGAAATGCACGACTG
RH2	AATTTTGTTTAACTTTAAGAAGGAGATACTAGATGTGTTGGGAAATGTCACCAGAAGACGAAAATCAGTCTTTAGAGCAAAAAAACAGCCAGCTTAAGCAAGAGATTAGTC AACTGGAGCAAGAAATTCAGCAGCTTGAATATGGGAGTGGACCTGGGCAGTTGGAAGATAAGGTCGAGGAATTACTGAGTAAAAACTATCACTTGGAGAACGAGGTTGAAC GCCTGAAAAAGTTGGTCGGCGGCAGTGGACCCGGCTCACCTGAAGACGAATTAGCTGCTAACGAGGAAGAGTTACAGCAAAATGAACAAAAGTTGGCTCAAATCAAACAAA AGCTGCAGGCTATTAAATACGGATCTCATCATCATCACCATCACGGGTCCAGTCCCGAGGATAAGATTGAAGAACTGAAACGCGAGAATGAAGAGTTGGAGTGGAAAATCG AAGAATTAAAGCGTGAGAATGAGGAATTGGAGAAGGGATCTGGGCCTGGACAGTTAGAAGACAAAGTTGAAGAGTTGCTGTCCAAGAACTACCATTTGGAAAATGAGGTGG AGCGTCTGAAGAAACTGGTGGGTGGGAGCGGCCCGGGTTCACCCGAAGATAAAAACGAGGAATTAAAGCGCGAAATCAAAGAATTAGAGTGGGAAAATGAAGAGCTTGAAC GCAAGATTGAGGAGCTTAAGCGTCGCGGTAAGTAATAAATCGGTGAAATGCACGACTG
SB9b	TCTAGATAATTTTGTTTAACTTTAAGAAGGAGATATACACCATGGGGCATCACCACCACCACCATCACCATATGGAAAACTTATATTTCCAAAGTGGTTCCGGGTCTCCTG AGGATGAGATCCGCCAGCTGGAGCAGGAAAACAGTCAGCTTGAACGCGAAAATCAGCGCCTTGAACAAGAGATCTACCAATTAGAGCGCGGTTCCGGGCCCGGGAGCCCGG AAGATGAAATTCAACAGCTTGAGGAGGAAATCTCTCAGCTGGAACAAAAAAATTCCGAACTGAAAGAAAAAAATCAGGAGCTTAAATACGGTTCCGGCCCGGGATCCCCAG AGGACGAAAATGAGAAGTTAGAACGCAAGAATGAGGAGTTGAAGTGGGAGATTAAGAAGCTTGAACGCGAAATCAAAGAATTAGAACGCGGCAGCGGACCAGGGGACATTG AGCAAGAGCTTGAGCGCGCAAAGCAGTCCATTGAAGAGTTGGAGCGCGAGGTAAACCAAGAACGTTCTCGCATGCAATATCTTCAAACCCGTTTGTCTGGGAGTGGACCGG GAAGCCCTGAAGATAAGATTTCGCAATTGAAGGAAAAGATCCAACAACTGAAACAAGAGAATCAGCAACTTGAGGAGGAAAACTCACAACTTGAGTATGGTAGCGGGCCCG GCTCTCCTGAGGACGAGAACGAGAAATTAGAAGAGAAAATCTGGGAGCTTAAGCGTAAAAATGAGGAGTTAAAGCGTGAAATCAAGGAATTAGAGGAAGGGTCCGGTCCAG GCAGTCCGGAGGATAAAATTGAAGAATTAAAGGAGAAAAACAGTCAGTTGAAAGAAAAAAACGAGGAATTGAAACAAAAAATCTATGAGTTAAAGGAGGGCTCTGGCCCAG GTGATATTGAACAGGAGTTGGAACGTGCCAAACAATCAATTGAAGAGCTGGAACGTGAGGTCAACCAGGAACGCAGCCGCATGCAGTATTTACAGACTCGTTTAAGCGGCT CCGGCCCCGGATCGCCTGAGGACGAAATCAAAGAGCTGGAATGGAAAAACGAGGAGCTGAAACGTGAAATTAAAGAATTGGAGGAGAAAAATGAAGAATTAAAACGCCTTG AACTCGAGTAATGACCTAGG
SB9c	TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCACCCGAAGACGAAATCCGCCAATTGGAGCAGGAGAACAGCCAATTGGAGCGTGAGAACCAGCGC TTAGAACAGGAGATCTATCAGCTGGAGCGCGGCTCTGGGCCAGGATCACCGGAAGATGAGATCCAACAACTGGAAGAGGAAATCTCGCAGTTAGAACAGAAAAATTCGGAA TTAAAGGAGAAGAACCAAGAGCTTAAATATGGTTCGGGTCCTGGATCCCCAGAAGACGAAATTAAGGAACTGGAGTGGAAGAACGAAGAGCTTAAACGCGAGATCAAAGAA CTGGAAGAGAAAAACGAGGAATTGAAACGTGGGTCTGGCCCAGGAGACATCGAACAGGAGCTTGAACGCGCTAAGCAATCCATTGAAGAGTTAGAGCGTGAGGTTAATCAA GAGCGCTCGCGTATGCAGTATCTTCAGACACGTCTGTCGGGTTCCGGTCCAGGTAGCCCCGAAGACAAAATCTCCCAGTTGAAAGAAAAAATTCAGCAACTGAAACAAGAA AATCAACAATTAGAAGAAGAGAACTCGCAGTTGGAGTATGGGAGTGGTCCCGGAAGCCCGGAGGACGAGAATGAGAAACTGGAGGAGAAGATTTGGGAACTGAAGCGTAAA AATGAGGAGTTAAAACGTGAAATCAAGGAACTGGAGGAGGGATCGGGGCCCGGTTCCCCAGAGGACAAGATTGAAGAATTGAAGGAAAAGAACTCGCAACTGAAAGAGAAG AATGAAGAATTAAAGCAAAAAATCTACGAGCTTAAGGAAGGTAGTGGGCCCGGTGACATTGAGCAAGAACTGGAGCGTGCCAAACAGAGCATCGAAGAATTAGAACGCGAG GTGAACCAAGAACGCTCTCGTATGCAATACCTGCAGACGCGTTTGTCAGGATCCGGGCCGGGCTCTCCTGAAGATAAGAATAAAGAGTTGAAAGAAGAAAACAAGGAACTG GAGTGGAAAATTGAAGAGTTGAAAGAGAAGATTAAAGAGCTTAAAGAACTCGAGCATCACCACCACCACCATCACCATTAATGACCTAGG
SB6	TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGCCCAGAAGATAAGAATGAGGAGCTTAAGCGCGAAATCAAGGAACTGGAGTGGGAAAATGAAGAG CTGGAGCGCAAAATTGAGGAGTTGAAACGTGAAGTGGCCCGGGCCAGCTCGAAGACAAAGTTGAGGAGCTTCTCAGCAAGAATTATCATTTGGAGAATGAAGTAGAACGC CTGAAGAAACTTGTGGGCGGCTCCGGCCCCGGCAGCCCCGAGGATAAGAACGAAGAGTTAAAACGTGAGATTAAGGAATTGGAGTGGGAGAATGAGGAACTGGAGCGCAAG ATCGAAGAGTTGAAACGCGGCAGTGGACCGGGTTCGCCAGAGGATAAAATTGAAGAGCTTAAACGCGAGAACGAGGAGCTTGAATGGAAGATTGAGGAACTGAAACGCGAG AACGAAGAATTAGAAAAGGGTAGTGGGCCTGGGCAACTGGAGGACAAGGTTGAGGAATTATTATCTAAGAACTATCATCTTGAGAACGAGGTTGAGCGTTTGAAGAAGTTA GTCGGGGGGTCAGGGCCCGGCTCTCCGGAAGATAAAATTGAAGAGTTGAAACGCGAAAACGAAGAACTTGAGTGGAAAATTGAAGAATTAAAGCGTGAGAATGAGGAGTTA GAGAAACTCGAGCATCACCACCACCACCATCACCATTAATGACCTAGG

Table S7. Primers used for cloning.

Description	Name	DNA sequence
Used for backbone pCIRCgp41-1a amplification	pCIRCgp41-1a_F	AAACGAGGCACTTGAAAAACGCGGTAAGTGCTTGGATCTGAAAACCCAGGTTCAG
	pCIRCgp41-1a_R	GTTTTTGTCTTCTGGAGACATTTCCCAACAGTTGTGGGTCAGAATATCGTTAGCGTAAA
Used for sequencing of constructs in backbones pCIRCgp41-1a, b, c	pCIRCgp41-1-seq_F	TCCTTAGCTTTCGCTAAGGATGATTTCTGGAATT
	pCIRCgp41-1-seq_R	CCTGCATAACGCGAAGTAATCTTTTCGGTTTTAA
Removal of TEV from cySB6-TEV	SB6-TEV-remov_F	TGGTGGTTCAGGACCTGGCAGTCCTGAGGATAAAAATGAAGAGCTGAAGCG
	SB6-TEV-remov_R	GCCAGGTCCTGAACCACCAACTAATTTTTTTAAACGTTCCACTTCATTTTCC
Used for backbone pCIRCgp41-1b amplification	pCIRCgp41-1b_F	CGCGGTAAGTGCTTGGATCTGAAAACC
	pCIRCgp41-1b_R	CATTTCCCAACAGTTGTGGGTCAGAATATCGTTA
Used for backbone pCIRCgp41-1c amplification	pCIRCgp41-1c_F	ATCGGTGAAATGCACGACTGATAGTAC
	pCIRCgp41-1c_R	CTAGTATCTCCTTCTTAAAGTTAAACAAAATTATTCTCTAGAAGC
Used for backbone pET41a amplification	pET41a_F	AGTTTGAGAAATAATTGATTAATACCTAGGCTGCTAAACAAAGCCCGAAAGGAAGCTGAG
	pET41a_R	CCCATGGTGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGG
Used for sequencing of constructs in backbone pET41a	pET41a_seq_F	TAATACGACTCACTATAGGG
	pET41a_seq_R	CTAGTTATTGCTCAGCGGT