SUPPLEMENTARY INFORMATION:

Proofreading exonuclease on a tether: the complex between the *E. coli* DNA polymerase III subunits α, ε, θ and β reveals a highly flexible arrangement of the proofreading domain

Kiyoshi Ozawa[1,2], Nicholas P. Horan[1], Andrew Robinson[1], Hiromasa Yagi[2], Flynn R. Hill[1], Slobodan Jergic[1], Zhi-Qiang Xu[1], Karin V. Loscha[2], Nan Li[1], Moeava Tehei[1], Aaron J. Oakley[1], Gottfried Otting[2], Thomas Huber[2] and Nicholas E. Dixon[1,*]

[1] School of Chemistry, University of Wollongong, NSW 2522, Australia

[2] Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia

SUPPLEMENTARY METHODS

Plasmids for protein expression

α270-His$_6$ and α270: The *dnaE$^+$* plasmid pND517 (60) was used as PCR template (with primers 891 and 1563, below) to produce a fragment encoding α270 with a C-terminal His$_6$-tag between NdeI and EcoRI sites. After restriction digestion, the fragment was inserted between the same sites in vector pETMCSI (33) to generate the T7-promoter vector pKO1367. Plasmid pKO1439, containing the gene encoding untagged α270 was made similarly (with primers 891 and 1962).

891	5'-TTTTTTTTCATATGTCTGAACCACGTTTCGTACACC-3' (36-mer)
1563	5'- TTTGAATTCTTAGTGATGGTGATGGTGATGGGTTACGTTACAGCGTTTGGCGATCTC-3' (57-mer)
1962	5'-TTTGAATTCTTAGGTTACGTTACAGCGTTTGGC-3' (33-mer)

εCTS$_{59}$ and the εCTS$_{59}$-His$_6$ amber mutant: The T7 expression construct for εCTS$_{59}$ (containing Met185–Ala243 of ε, preceded by a biotinylation tag; Figure 1B) was made in two steps. Vector pKO1316 was made by annealing complementary oligonucleotides 907 and 908 and inserting them between the NdeI and MluI sites of pETMCSI, eliminating the original NdeI site and creating a new one after codons for a biotinylation sequence (61). Next the *dnaQ$^+$* plasmid pSH1017 (36) was used as template with primers 975 and 988; insertion of the resulting fragment between the NdeI and EcoRI sites of pKO1316 gave pKO1422.

A T7 promoter vector was also made for expression of a version of εCTS$_{59}$ with a C-terminal unnatural amino acid (UAA) followed by a His$_6$-tag. The constructs encoding εCTS$_{59}$-UAA-His$_6$ was created by PCR with pKO1422 as template, using primers 514 and 1996. The amplified fragments were digested with NdeI and EcoRI for subsequent insertion between the corresponding sites of pETMCSI to give pεCTS-1996.

514	5'-CGACTCACTATAGGGAGACCACAAC-3' (25-mer)

907	5'-TAATGGCTGGTCTGAACGACATCTTCGAAGCTCAGAAAATCGAATGGCACGAACATATGA-3' (60-mer)
908	5'-CGCGTCATATGTTCGTGCCATTCGATTTTCTGAGCTTCGAAGATGTCGTTCAGACCAGCCAT-3' (62-mer)
975	5'-TTTGAATTCTTATGCTCGCCAGAGGCAAC-3' (29-mer)
988	5'-TTTTTTTTTCATATGGCTTTTGCGATGGAAGGAGAGACACAACAG-3' (44-mer)
1996	5'-TTTGAATTCTTAATGGTGATGGTGATGGTGCTATGCTCGCCAGAGGCAAC-3' (50-mer)

εCTS–α270 fusion proteins (constructs A): The gene encoding the $εCTS_{35}$–α270 fusion protein (construct A, comprising residues 209–243 of the εCTS fused via a 9 amino acid linker [sequence: TRESGSIGS] to the N-terminus of α270; Figure 1C) was made by strand-overlap extension PCR following amplification in the first step of a *dnaQ* fragment from pSH1017 (primers 2224 and 2225) and a *dnaE* fragment from pKO1367 (primers 1962 and 2223). The two fragments were combined and used as template in the second step (primers 1962 and 2225). The fragment was digested with NdeI and EcoRI and inserted at the same sites in pETMCSI to give pKO1479.

A version of this construct with a C-terminal His_6-tag (plasmid pKO1497) was made similarly (template pKO1479 and primers 1563 and 2225). Subsequent DNA sequencing revealed that pKO1479 and pKO1497 contained a PCR-generated mutation in *dnaE(270)*, encoding α270(Leu21Pro). Consequently, the wild-type versions of both plasmids, named pKO1479wt and pKO1497wt, were made similarly. A longer version of this construct, $εCTS_{44}$–α270 (comprising residues 200–243 of the εCTS fused to the N-terminus of the Leu21Pro mutant of α270) was made using primers 564 and 566 to amplify the α270 fragment from pKO1479, and primers 565 and 568 to amplify the *dnaQ(200–243)* fragment from pSH1017. The second PCR step used primers 564A and 568. After cleavage with NdeI and EcoRI, the fragment was inserted between these sites in pETMCSI to give plasmid pNH1543.

564	5'-TTAGGTTACGTTACAGCG-3' (24-mer)
564A	5'-AAGAATTCTTAGGTTACGTTACAGCG-3' (32-mer)
565	5'-TGCTCGCCAGAGGCAACT-3' (24-mer)
566	5'-AGTTGCCTCTGGCGAGCA-3' (24-mer)
568	5'-AAAAAAAACATATGGCAACAATTCAGCGCATTG (33-mer)
2223	5'-ACGCGTGAGTCTGGCTCTATCGGATCCATGTCTGAACCACGTTTCGTACACCTGCGG-3' (57-mer)
2224	5'-GGATCCGATAGAGCCAGACTCACGCGTTGCTCGCCAGAGGCAACTTCCGCCTTTCTTCTG-3' (60-mer)
2225	5'-TTTTTTTTTCATATGGCAAGTAAGTTACGCGTTGTTTTTGCGACAGATG-3' (48-mer)

α270–εCTS fusion protein (constructs B) The fusion construct B (α270–$εCTS_{35}$) comprises residues 209–243 of εCTS fused to the C-terminus of α270 via the same 9-residue linker as used for constructs A (Figure 1C). The α270-encoding portion was PCR-amplified from pND517 using primers 891 and 2308

and the segment encoding residues of the εCTS was amplified from pKO1422 using primers 975 and 2307. The products were gel purified, combined in equimolar ratio and used as template for a second PCR with primers 891 and 975. Digestion with NdeI and EcoRI, gel-purification and insertion into pETMCSI created the expression vector pKO1559. A version of α270–εCTS with an N-terminal His$_6$-tag was created by insertion of the ~1 kb NdeI–EcoRI fragment from pKO1599 between the same sites of pETMCSIII (33) to give pKO1560.

975	5'-TTTGAATTCTTATGCTCGCCAGAGGCAAC-3' (29-mer)
2307	5'- ACGCGTGAGTCTGGCTCTATCGGATCCGCAAGTAAGCTGCGCGTTGTTTTTGCG-3' (54-mer)
2308	5'-GGATCCGATAGAGCCAGACTCACGCGTGGTTACGTTACAGCGTTTGGCGATCTC-3' (54-mer)

ε193, ε$_Q$193 and ε$_L$193: Plasmid pSH1017 (36) was used as template for PCR amplification of a DNA fragment containing the 5'-terminal 193 codons of the *dnaQ* gene using the vector primer pET3 designed to preserve the BamHI site that precedes the ribosome-binding site of *dnaQ* and primer 405 designed to incorporate an EcoRI site just following a TAA stop codon. The PCR product was isolated following digestion with BamHI and EcoRI and inserted between the same restriction sites in pETMCSII (33) to place the gene under transcriptional control of the phage T7 φ10 promoter in plasmid pSJ1447. This vector directs overproduction of the ε193 protein that contains an intact clamp-binding motif (CBM; Figure 1A). The similar plasmids pSJ1448 and pSJ1449 that direct expression of ε$_Q$193 (Gln182Ala mutant; weakened CBM) and ε$_L$193 (Thr183Ala/Met185Leu/Ala186Pro/Phe186Leu mutant; strengthened CBM), respectively, were constructed by an identical procedure, but using plasmids pSJ1445 and pSJ1446 (25) as templates for PCR.

pET3	5'-CGACTCACTATAGGGAGACCACAAC-3' (25-mer)
405	5'-TTTGAATTCTTATGTCTCTCCTTCCATCGC-3' (30-mer)

Full-length α, His$_6$-α and His$_6$-α$_{GL}$: The *dnaE* gene was amplified from pND517 using PCR primers 891 and 897. The resulting PCR fragment was digested with NdeI and EcoRI and the ~3570 bp product isolated from an agarose gel. Insertion into the corresponding sites of pETMCSI yielded the plasmid pKO1341, which directs the expression of *dnaE* under transcriptional control of the T7 promoter. The *dnaE*⁺ NdeI–EcoRI fragment from pKO1341 was isolated and then inserted between the same sites in pETMCSIII to give plasmid pKO1533, which directs expression of the α subunit with an N-terminal His$_6$-tag.

For construction of a similar vector expressing the His$_6$-α$_{GL}$ (V832G/A921L/M923L) triple mutant (Figure 1D), *dnaE*⁺ gene fragments of 665, 2005 and 366 bp were amplified using pSJ1392 (25) as template and the primer pairs 2476 and 2483, 2481 and 2482, and 2479 and 2480, respectively. The mutated gene was reassembled from the three gel-isolated fragments by strand overlap extension PCR with primers 2476 and 2479. The resulting 2.4 kb fragment from the second PCR coded for the triple mutant V832G/A921L/M923L. After digestion with XhoI and StuI, the product (2.4 kb) was isolated from a gel and inserted between the same two sites in pKO1533 to give pKO1536.

897	5'-TTTGAATTCTTAGTCAAACTCCAGTTC -3' (27-mer)
2476	5'-TATTTTCTCGAGCTGATCCGCACCGGCAGGCCG-3' (33-mer)

2479	5'-TCAGGTACAGGCCTAACGTTTCACGTTCCCCATC-3' (34-mer)
2480	5'-AACGACGACGGCGAAATCGGTTATGGTATTGGCGCGATC-3' (39-mer)
2481	5'-CGCGCCAATACCATAACCGATTTCGCCGTCGTCG-3' (34-mer)
2482	5'-GACGTTGACTTCTGTATGGAGAAACGCGATCAG-3' (33-mer)
2483	5'-CTGATCGCGTTTCTCCATACAGAAGTCAACGTC-3' (33-mer)

Primers used to construct Bpa mutants of α270

In the table below, codons of mutated amino acids are underlined and mutated bases are in bold.

P1[a]	5'-PO$_4$-TTAGCTGGTCGATCCCGCGAAATTAATACG-3' (30-mer)
P2[a]	5'-PO$_4$-CCAGCTAACAAAAAACCCCTCAAGACCCG-3' (29-mer)
P3[a]	5'-PO$_4$-TCGATCCCGCGAAATTAATACG-3' (22-mer)
P4[a]	5'-PO$_4$-CAAAAAACCCCTCAAGACCCG-3' (21-mer)
1375[b]	5'-TCGATCCCGCGAAATTAATACGACTCAC-3' (28-mer)
1376[b]	5'-CAAAAAACCCCTCAAGACCCGTTTAGAG-3' (28-mer)
1577[c]	5'-PO$_4$-CCGCATATGTCTGAA**TAG**CGTTTCGTACACCTGCGG-3' (36-mer)
1578[c]	5'-PO$_4$-AGGGATGTGAAAGTTAAACAAAATTATTTCTAG-3' (33-mer)
1579[d]	5'-PO$_4$-GGCCTGGCCAAAACC**TAG**CCGTTGGTAAAAAAGGCG-3' (36-mer)
1580[d]	5'-PO$_4$-ATCGATCATCGAGTAGTCGCTGTGCACCCG-3' (30-mer)
1581[e]	5'-PO$_4$-GATTTTAACGTCCAGTGC**TAG**CTGCTGGGTGATGAG-3' (36-mer)
1582[e]	5'-PO$_4$-TGCCCCGACGATAGGCTTAATCCCTGCGCCATG-3' (33-mer)
1583[f]	5'-PO$_4$-CTGATCTCAAAAGCGTAT**TAG**CGCGGGTACGGTG-3' (34-mer)
1584[f]	5'-PO$_4$-CAACGTCAGATTCTGATAGCCGGTATTGTTCGC-3' (33-mer)
1585[g]	5'-PO$_4$-ACCCTCGACGATCCT**TAG**CGCCCGCGTAACTATTCG-3' (36-mer)
1586[g]	5'-PO$_4$-AAAGCCGTCGTGGATCGCGACGCGGATTTC-3' (30-mer)
2360[h]	5'-CGCCCGCGTAAC**TAG**TCGCCGCAGCAATATATG-3' (33-mer)
2361[h]	5'-CTGCGGCGA**C**TAGTTACGCGGGCGTTTAGG-3' (30-mer)
2412[i]	5'-ATCCGCACCGGC**TA**GCCGGATGAAGAAAGC-3' (30-mer)
2413[i]	5'-TTCTTCATCCGGC**TA**GCCGGTGCGGATCAG-3' (30-mer)
2414[j]	5'-AACTATTCGCCG**TAG**CAATATATGCGTAGC-3' (30-mer)

2415[j]	5'-ACGCATATATTGCT**A**CGGCGAATAGTTACG-3' (30-mer)

[a] Primers used for circularization of PCR-generated templates for cell free synthesis of Bpa mutants of α270.

[b] Outside primers used for PCR-generated templates for cell free synthesis of Bpa mutants of α270.

[c] Forward and reverse primers to generate the Pro4Bpa mutant of α270.

[d] Forward and reverse primers to generate the Ala25Bpa mutant of α270.

[e] Forward and reverse primers to generate the Asp75Bpa mutant of α270.

[f] Forward and reverse primers to generate the Gln106Bpa mutant of α270.

[g] Forward and reverse primers to generate the Lys229Bpa mutant of α270.

[h] Forward and reverse primers to generate the Tyr234Bpa mutant of α270.

[i] Forward and reverse primers to generate the Arg175Bpa mutant of α270.

[j] Forward and reverse primers to generate the Gln237Bpa mutant of α270.

Purification of εCTS–α270 fusion proteins

The εCTS$_{35}$–α270(Leu21Pro) fusion protein (construct A; Figure 1C) used for crystallography was purified as follows. *E. coli* BL21(λDE3)*recA* (27) containing plasmids pLysS (62) and pKO1479 was grown at 37°C in LB broth containing 50 mg l^{-1} thymine, 200 mg l^{-1} ampicillin and 35 mg l^{-1} chloramphenicol to A_{600} = 0.8–1.0, whereupon IPTG was added to 0.5 mM and growth continued for a further 3 h. A suspension of cells (19.9 g from 6 liters of culture) in 300 ml of lysis buffer [50 mM Tris.HCl (pH 7.6), 2 mM EDTA, 2 mM dithiothreitol, 20 mM spermidine, 0.5 mM phenylmethyl sulfonyl fluoride, 10% (w/v) sucrose] was passed twice through a French Press at 10,000 psi, before being clarified by centrifugation. Proteins in the supernatant were precipitated by addition of ammonium sulfate (0.2 g per ml of lysate) followed by gentle agitation over 45 min at 4°C, and harvested by centrifugation. The pellet was resuspended in 80 ml of Buffer E [50 mM Tris.HCl (pH 7.6), 1 mM dithiothreitol, 1 mM EDTA, 20% (v/v) glycerol] and dialyzed against Buffer E+160 mM NaCl before being passed through a column (2.5 x 12 cm) of DEAE-Toyopearl 650M that had been equilibrated with the same buffer. The εCTS–α270 protein did not bind to the column and was collected in the flow through fractions; this step removes nucleic acids that otherwise interfere with later steps. Fractions were pooled (120 ml), dialysed against 2 changes of 1 l of Buffer E (without salt), and reloaded onto a similar column equilibrated with Buffer E. Proteins were eluted in a gradient (500 ml) of 0–0.25 M NaCl in Buffer E. The εCTS–α270 protein eluted at ~60 mM NaCl.

Ten 10 ml fractions retained from this step were frozen separately at –80°C for storage. Two of the peak fractions (20 ml) were later thawed, diluted with 30 ml of Buffer E and split into 5 aliquots that were separately applied to a MonoQ 10/100 GL column (GE Healthcare) and eluted in gradients (64 ml) of 30–400 mM NaCl in Buffer E. On each occasion, the highly purified εCTS–α270 eluted at ~130 mM NaCl. The combined yield was ~43 mg. Its molecular mass was confirmed by ESI-MS after dialysis into 0.1% formic acid (observed: 34,454.8 ± 0.2; calculated: 34,438.1, assuming loss of the N-terminal f-Met). The α270–εCTS (construct B) and other εCTS–α270 (construct A) fusion proteins were purified similarly.

Purification of His$_6$-α$_{GL}$

His$_6$-α$_{GL}$ was expressed in vivo and purified as follows. BL21(λDE3)*recA*/pLysS/pKO1536 was grown at 30°C in 5 liters of Z-medium supplemented with 100 mg l^{-1} ampicillin and 33 mg l^{-1} chloramphenicol in a 20-liter fermenter (9). Overexpression of α$_{GL}$ was induced by addition of 1 mM IPTG at A_{595} = 1.4 and the culture was incubated for another 11 h at 24°C, yielding 54.4 g of cells. The cells were resuspended in 300 ml of lysis buffer (9) and disrupted using a French press at 12,000 psi. The lysate was clarified and

proteins in the supernatant (325 mL) were precipitated by addition of ammonium sulfate (97.5 g) at 4°C. The precipitate was collected by centrifugation and dissolved with 75 mL of buffer A [50 mM HEPES-KOH (pH 7.5), 300 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol]. The solution was divided into two batches for purification on a 5 mL column of Ni-NTA in buffer A with a linear imidazole gradient (20 to 500 mM). The total eluate (100 ml) was dialyzed against buffer B [20 mM Tris-HCl (pH 7.55), 1 mM EDTA, 2 mM dithiothreitol, 5% (v/v) glycerol] and loaded onto a DEAE-Toyopearl 650M column (2.5 x 13 cm); His$_6$-α_L was eluted in a NaCl gradient (0–1 M in buffer B). Pooled fractions containing His$_6$-α_L (115 ml) were dialyzed against buffer P [20 mM sodium phosphate (pH 6.6), 4 mM EDTA, 1 mM dithiothreitol, 10% (v/v) glycerol] for fractionation on a freshly activated and packed phosphocellulose column (2.5 x 9 cm) with a linear gradient of 0 to 1 M NaCl in buffer P. The purity of α_{GL} in collected fractions was carefully assessed by SDS-PAGE and only samples with <5% impurities were retained. Less pure fractions were successfully turned into high quality samples by re-chromatography on the phosphocellulose column. More than 60 mg of the mutant α_{GL} was obtained in high purity.

Molecular modeling

Modeling of the structure of the structure of the *E. coli* (*Eco*) $\alpha\epsilon\theta$:β_2:DNA complex in the polymerization mode commenced with (i) docking of the X-ray structure of the *Eco* β_2:DNA complex (PDB: 3BEP) onto that of the full-length *Taq* α complex with primer-template DNA (PDB: 3E0D), which brings the internal CBM of α into proximity with one of the protein-binding sites of β_2, essentially as described by Wing et al. (55); (ii) structural alignment and replacement of corresponding domains in the higher-resolution structure of full-length *Taq* α (PDB: 2HPI) with those of *Eco* α917 (PDB: 2HNH) followed by homology modeling of the remaining domains (residues 918–1160) of *Eco* α (using PDB: 3E0D); and (iii) minor alignment of the NMR structure of θ in the ϵ186:θ complex (16) based on the position of HOT in the X-ray structure of ϵ186:HOT (PDB: 2IDO). The domain orientations in model (ii) were then adjusted to overlay optimally with those of model (i) to give a model for the *Eco* α:β_2:DNA complex (iv). Next, our 1.7 Å structure of ϵCTS$_{35}$–α270(L21P) (Figure 2) was overlaid on the α PHP domain in model (iv) and used to replace it in all regions of close identity, i.e., ignoring changes around the L21P mutation in α270, to give model (v). The additional region of α–ϵ contact in ϵCTS$_{44}$–α270(L21P) was then built into this model (vi). Next, the CBMs in ϵ (residues 182–187) and α (residues 920–924) were modeled into the two protein-binding sites of β_2, based on the most closely related structures of β_2-peptide complexes; PDB: 1JQL (63) and 3Q4L (64), respectively, and the side chain conformations of the β_2 pocket and the CBM in ϵ were refined using Rosetta FlexPepDock (65). At this point, it was possible to demonstrate (vii) that a random coil peptide corresponding to the flexible Q-linker of ϵ (residues 188–201) is able to flexibly bridge the gap between the CBM bound to β_2 and the site of (weak) interaction of residues 202–205 of ϵ with the region around His183 and Asp252 in the PHP domain of α, and (viii) that the $\epsilon\theta$ sub-complex is still able, without steric clashes, to fit between its binding sites on β_2 and the PHP domain of α to form a compact structure that orients the active site of ϵ inwards to be placed near the double-stranded DNA that runs between the central channel of β_2 and the active site of α (e.g., in Figure 6A).

Files to enable detailed visualization of models are available at the website: http://comp-bio.anu.edu.au/huber/Pol3.

SUPPLEMENTARY TABLES

Table S1. X-ray data collection and refinement statistics (molecular replacement).

	$\varepsilon CTS_{35}-\alpha270(L21P)$	$\varepsilon CTS_{44}-\alpha270(L21P)$
PDB ID:	4GX8	4GX9
Data collection		
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	84.47, 56.63,138.01	83.02, 56.98,135.07
α, β, γ (˚)	90.0, 93.52, 90.0	90.0, 93.78, 90.0
Resolution (Å)	24–1.7 (1.8–1.7)[1]	35–2.15 (2.19–2.15)[1]
R_{sym} or R_{merge}	10.7 (39.6)	9.5 (49.4)
$I/\sigma I$	4.4 (1.8)	10.4 (1.9)
Completeness (%)	98.1 (97.2)	97.2 (96.5)
Redundancy	7.0 (6.8)	3.1 (2.9)
Refinement		
Resolution (Å)	24–1.7	35–2.15
No. reflections	140,410	67,126
R_{work} / R_{free}	21.3 / 24.6	22.6 / 29.1
No. atoms		
Protein	9883	9801
Ligand/ion	4	0
Water	1129	277
B-factors		
Protein	18.1	21.9
Ligand/ion	31.2	-
Water	26.1	21.8
R.M.S. deviations		
Bond lengths (Å)	0.009	0.019
Bond angles (°)	1.12	1.75
Ramachandran plot		
residues in: core region	92.5%	89.2%
additional allowed region	7.2%	10 / 4%
generously allowed region	0.4%	0 / 3%
disallowed region	0%	0%

[1] Values in parentheses are for highest-resolution shell.

Table S2. Backbone resonance assignments of α270 in the absence and presence of εCTS_{59} determined by combinatorial isotope labeling*

Residues	Free $\delta(^{1}H)$/ppm	Free $\delta(^{15}N)$/ppm	Bound $\delta(^{1}H)$/ppm	Bound $\delta(^{15}N)$/ppm
V7	6.64	124.1	6.51	124.3
T24	8.14	115.3	8.15	115.1
A25	8.91	120.2	8.90	120.2
K30	8.08	121.5	8.04	121.2
M36	7.95	119.8	7.79	119.2
C48	8.99	122.3	8.94	121.8
H58	8.30	121.2	8.26	121.4
I62	7.94	116.4	7.96	116.4
K63	8.19	127.6		
V66	9.61	128.3	9.57	128.7
C74	9.57	122.1	9.56	122.0
D75	8.84	124.8	8.83	124.9
H83	7.50	119.6	7.50	119.6
N96	8.51	120.5	8.63	120.8
S102	7.98	113.1	7.97	113.2
K103	8.40	122.1	8.46	122.1
D119	8.42	113.3	8.44	113.3
W120	8.22	123.1	8.21	123.1
N125	6.63	113.3	6.63	113.2
G127	8.73	109.4	8.79	109.0
L130	9.04	126.5	9.15	126.7
G133	8.49	116.1	8.47	115.5
G134	7.46	109.6	7.50	109.7
M136	6.58	112.7	6.60	112.8
N147	7.42	117.1	7.45	117.1
A149	8.25	124.7	8.27	124.6
D152	7.89	117.1	7.88	117.1
V155	8.80	118.2	8.78	117.9
Y166	7.76	120.7	7.66	120.9
F167	8.07	124.4	7.92	124.3
L168	8.48	122.1	8.47	122.0
E169	7.85	117.0		
S180	8.36	116.9	8.43	117.4
Y181	8.36	122.3	8.39	122.4
L182	7.97	123.7	8.03	123.9
H183	8.19	116.8	8.30	117.3
R192	8.32	112.4	8.34	112.5
V197	8.99	118.7	8.99	118.6
T199	8.40	109.8	8.39	110.1
I205	9.28	118.4	9.17	118.3
S207	8.49	120.9		
A212	8.01	122.9	8.01	122.9

Residue			
D226	8.15	120.3	8.15	120.3
D227	7.25	123.4	7.23	123.4
K229	8.92	117.4	8.91	117.4
Y234	8.45	118.0	8.45	117.9
Q237	7.88	113.9	7.88	113.8
A251	6.94	122.5	6.95	122.7
A264	7.72	120.1	7.72	120.1
C267	7.08	117.1	6.87	115.6
N268	8.73	126.2	8.80	127.0

[1] At 25°C in 20 mM Tris.HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (NMR buffer). The assignments rely on the uniqueness of amino acid pairs in the amino acid sequence of α270.

Table S3. Backbone resonance assignments of εCTS_{59} in the complex with α270 determined by combinatorial isotope labeling[1]

Residue[2]	$\delta(^1H)$/ppm	$\delta(^{15}N)$/ppm	$\delta(^{13}C\alpha)$/ppm
M			
A			
S			
M	8.00	121.0	53.0
T	8.16	114.4	59.3
G	8.41	111.5	42.8
G	8.28	108.9	42.6
L	8.15	121.6	52.7
N	8.39	118.4	50.7
D	8.09	120.5	51.9
I	7.83	120.1	59.1
F			
E	8.13	121.1	54.6
A			
Q	8.16	118.6	53.4
K	8.08	122.5	53.9
I	8.03	122.3	58.6
E			
W	8.17	122.8	55.2
H	7.75	120.7	53.4
E	8.13	122.2	54.1
H			
M185			
A186			
F187	7.93	118.7	54.9
A188	8.04	125.0	49.8
M189			

E190	8.36	122.0	54.3
G191	8.36	110.1	42.6
E192	8.23	120.7	54.0
T193	8.20	115.3	59.6
Q194	8.39	122.7	53.4
Q195/196/197[3]	8.36	121.6	53.4
Q195/196/197[3]	8.41	121.7	53.4
Q195/196/197[3]	8.43	121.7	53.3
G198	8.37	110.1	42.6
E199	8.23	120.7	53.9
A200	8.31	124.8	49.9
T201	8.03	113.6	59.4
I202	7.98	122.7	58.7
Q203	8.19	119.5	52.8
R204	8.28	122.6	53.0
I205	8.02	122.3	58.7
V206	8.34	126.4	59.7
R207	8.38	125.9	52.9
Q208	8.59	122.6	53.1
A209	8.05	124.4	50.1
S210			
K211			
L212			
R213			
V214			
V215			
F216			
A217			
T218			
D219			
E220			
E221	8.36	122.1	54.3
I222			
A223			
A224			
H225			
E226			
A227			
R228			
L229			
D230			
L231			
V232			
Q233			
K234			
K235			
G236			
G237	8.10	110.0	42.9

S238	
C239	
L240	
W241	
R242	
A243	

[1] At 25˚C in 20 mM Tris.HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol. The assignments rely on the uniqueness of amino acid pairs in the amino acid sequence of εCTS$_{59}$ and on 3D HNCA and HN(CO)CA experiments with combinatorially ^{13}C- and uniformly ^{15}N-labeled samples.

[2] The table lists the full amino-acid sequence of the εCTS$_{59}$ construct. Residue numbers pertain to ε and are omitted for residues preceding the ε segment. Blank lines identify residues for which no peaks could be identified in the ^{15}N-HSQC spectrum.

[3] Three peaks were resolved in the ^{15}N-HSQC spectra for residues Q195, Q196 and Q197 but their precise sequence-specific assignment could not be established due to insufficient spectral resolution in the 3D NMR experiments.

Table S4. H-Bonded and electrostatic contacts between εCTS and α270 in the crystal structures of constructs A [εCTS$_{35}$–α270(L21P) and εCTS$_{44}$–α270(L21P)].

ε residue	donor/ acceptor	α residue	donor/ acceptor	donor–acceptor distance (Å)[1]
εCTS$_{35}$–α270(L21P): PDB 4GX8 (1.7 Å)				
ε Arg213	O	α Asn259	ND2	3.14, 3.10, 3.08, 3.15
	NH1	α Asp164	OD1	3.38, 3.33, 3.31, 3.38
	NH2	α Asp164	OD2	3.40, 3.46, 3.60, 3.55
ε Val215	N	α Asn259	OD1	2.99, 2.90, 2.94, 2.95
ε Ala217	N	α Glu262	OE1	2.96, 2.93, 2.96, 2.94
ε Glu221	OE2	α Arg266	NE	2.96, 2.91, 2.83, 2.91
ε His225	ND1	α Lys63	NZ	2.92, 2.90, 2.87, 2.88
ε Arg228	NE	α Val269	O	2.76, 2.82, 2.81, 2.77
	NH1	α His58	NE2	3.02, 3.08, 3.06, 3.04
	NH2	α His58	NE2	3.02, 3.02, 3.06, 2.96
ε Leu240	N	α Gly61	O	2.84, 2.79, 2.84, 2.81
ε Trp241	NE1	α Glu3	OE1	2.93, 2.97, 2.97, 2.93
εCTS$_{44}$–α270(L21P): PDB 4GX9 (2.15) Å[2]				
ε Gln203	O	α Glu179	OE2	----, 2.78, 2.95, ----
	O	α His183	NE2	----, 2.82, 2.65, ----
	O	α Arg172	NH2	----, 3.27, ----, ----
	NE2	α His183	ND1	----, 3.47, 3.41, ----
ε Arg204	O	α Asp252	OD2	----, 3.47, ----, ----
	NH1	α Leu249	O	----, 3.24, 3.01, ----
	NH1	α Asp252	OD1	----, 2.78, 2.75, ----
	NH2	α Glu248	O	----, 3.39, 3.39, ----
ε Ile205	N	α Asp252	OD2	----, ----, 2.66, ----
ε Ser210	O	α Glu255	OE2	----, 3.04, 2.76, ----

[1] Distances between donor and acceptor atoms in molecules A, B, C, D in the unit cell (as defined in the PDB files), respectively.
[2] Contacts additional to those seen in εCTS$_{35}$–α270(L21P).

SUPPLEMENTARY FIGURES
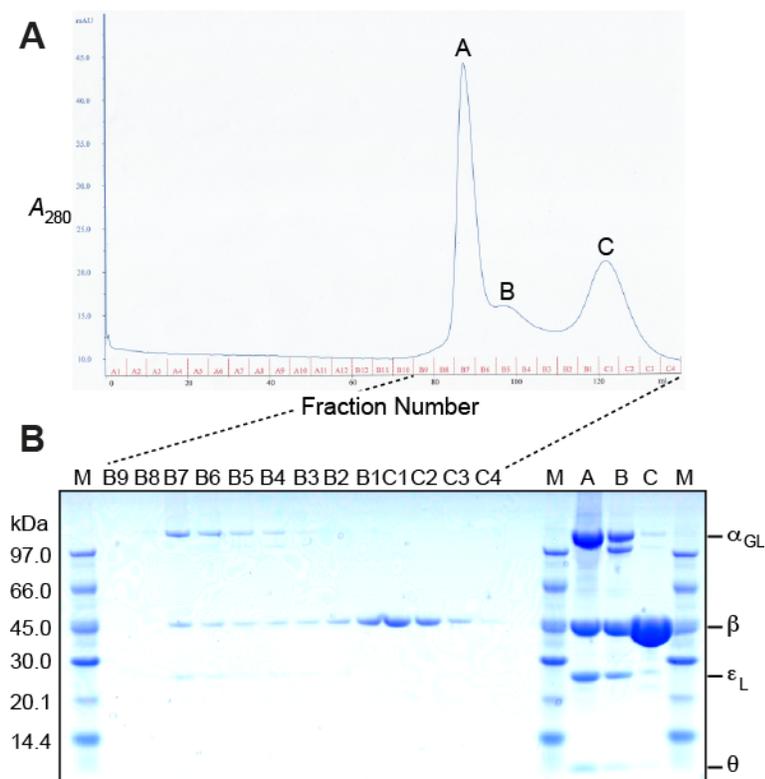
Figure S1. Gel filtration of His$_6$-α_{GL}: ^{15}N-ε_L:θ:β_2 shows that the complex is sufficiently stable for purification. **(A)** A mixture of 100 μM His$_6$-α_{GL}: ^{15}N- ε_L:θ and 400 μM β_2 was gel filtered through a column (2.5 x 47 cm) of Sephacryl S200 in NMR buffer. **(B)** Analysis of the three peaks of the chromatogram in (A) using a 4–12% NuPAGE gradient gel with MES-SDS running buffer (Coomassie blue staining). The gel shows that peak A (pooled fractions B6–B7) contains the α_{GL}:ε_L:θ:β_2 complex, whereas peak B (fractions B2–B6) also shows a truncated form of α_{GL}. Peak C (fractions B1–C4) is primarily β_2.
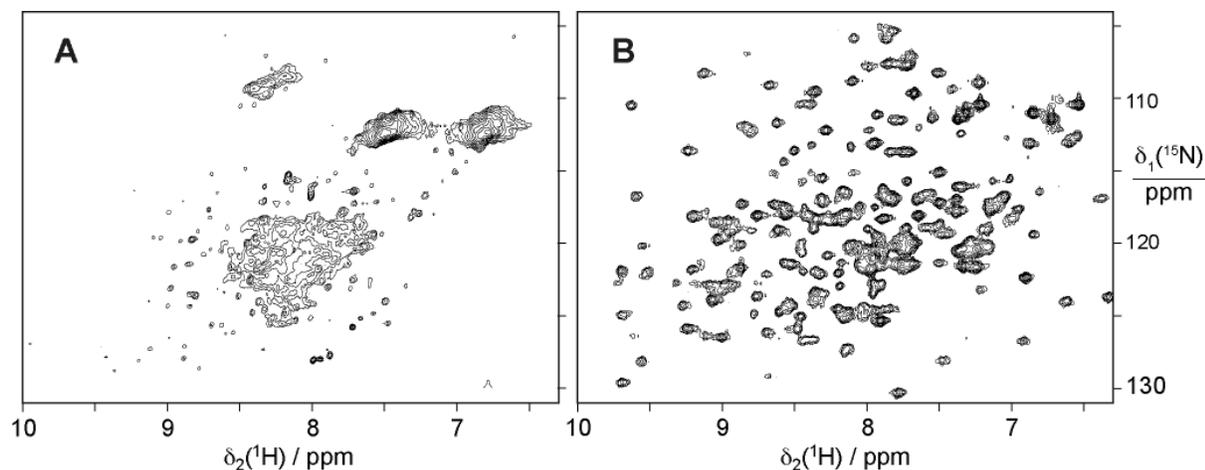
Figure S2. Properly folded α270 is readily obtained by cell-free protein synthesis but not by in vivo expression; 15N-HSQC spectra of **(A)** ^{15}N-α270 made in vivo. **(B)** ^{15}N-α270 made by cell-free expression.

B
```
MSEPRFVHLR VHSDYSMIDG LAKTAPLVKK AAALGMPALA ITDFTNLCGL VKFYGAGHGA
GIKPIVGADF NVQCDLLGDE LTHLTVLAAN NTGYQNLTLL ISKAYQRGYG AAGPIIDRDW
LIELNEGLIL LSGGRMGDVG RSLLRGNSAL VDECVAFYEE HFPDRYFLEL IRTGRPDEES
YLHAAVELAE ARGLPVVATN DVRFIDSSDF DAHEIRVAIH DGFTLDDPKR PRNYSPQQYM
RSEEEMCELF ADIPEALANT VEIAKRCNVT HHHHHH
```
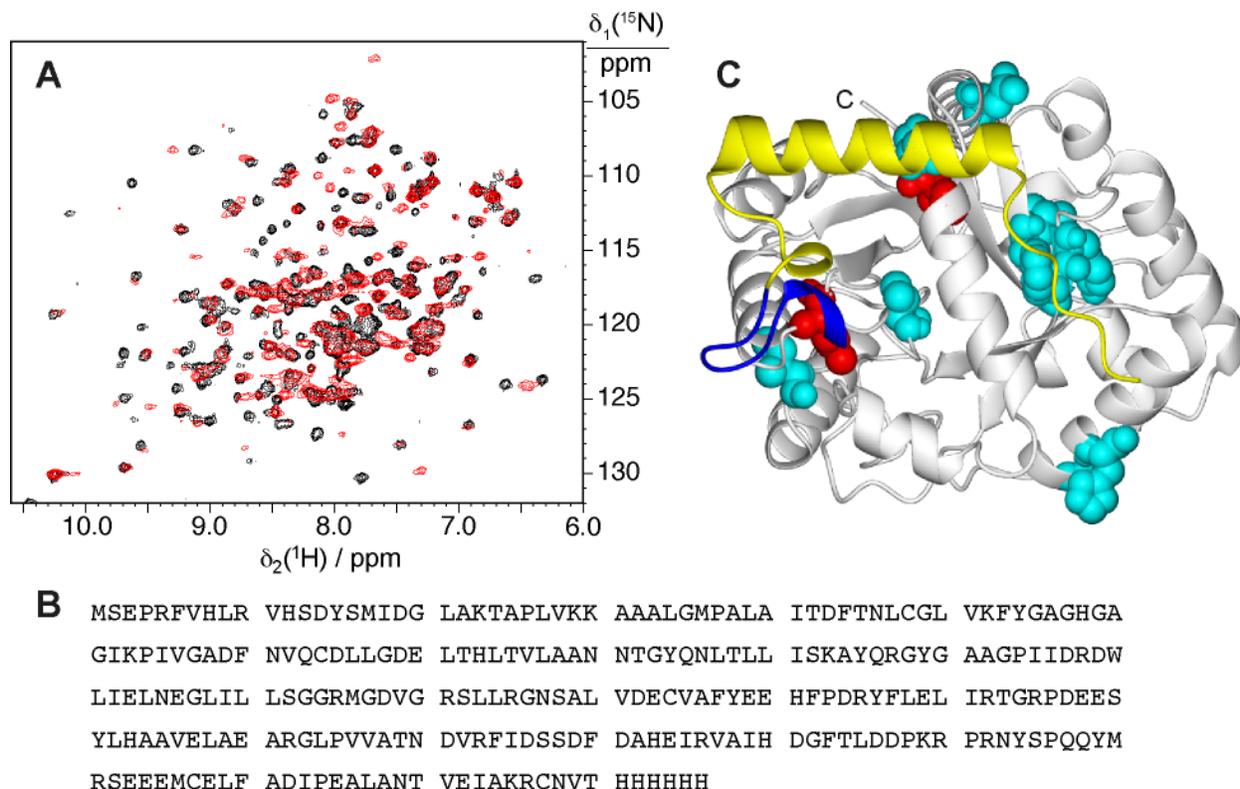
Figure S3. The εCTS binds to the PHP domain of α, causing local changes in the chemical environment of α. **(A)** Superimposition of ^{15}N-HSQC spectra of uniformly ^{15}N/^{13}C-labeled samples of α270 with (red) and without (black) εCTS$_{59}$ shows that the chemical shifts of some but not all cross-peaks change due to the presence of the εCTS, indicating that the fold of the PHP domain remains intact upon binding of εCTS. **(B)** Amino acid sequence of the α270 construct. **(C)** Structure of construct A (Figure 2) highlighting the residues for which chemical shift changes were observed for α270 in the complex with εCTS$_{59}$ compared to α270 alone. Residues for which amides displayed chemical shift changes >0.08 ppm in the presence versus the absence of εCTS$_{59}$ are shown in space filling representation. The two residues with the largest chemical shift change (>0.15 ppm, highlighted in red) are in proximity to the εCTS peptide (shown in yellow). The linker peptide connecting the εCTS with the N-terminus of α270 in construct A is shown in blue. The C-terminus of α270 is indicated.
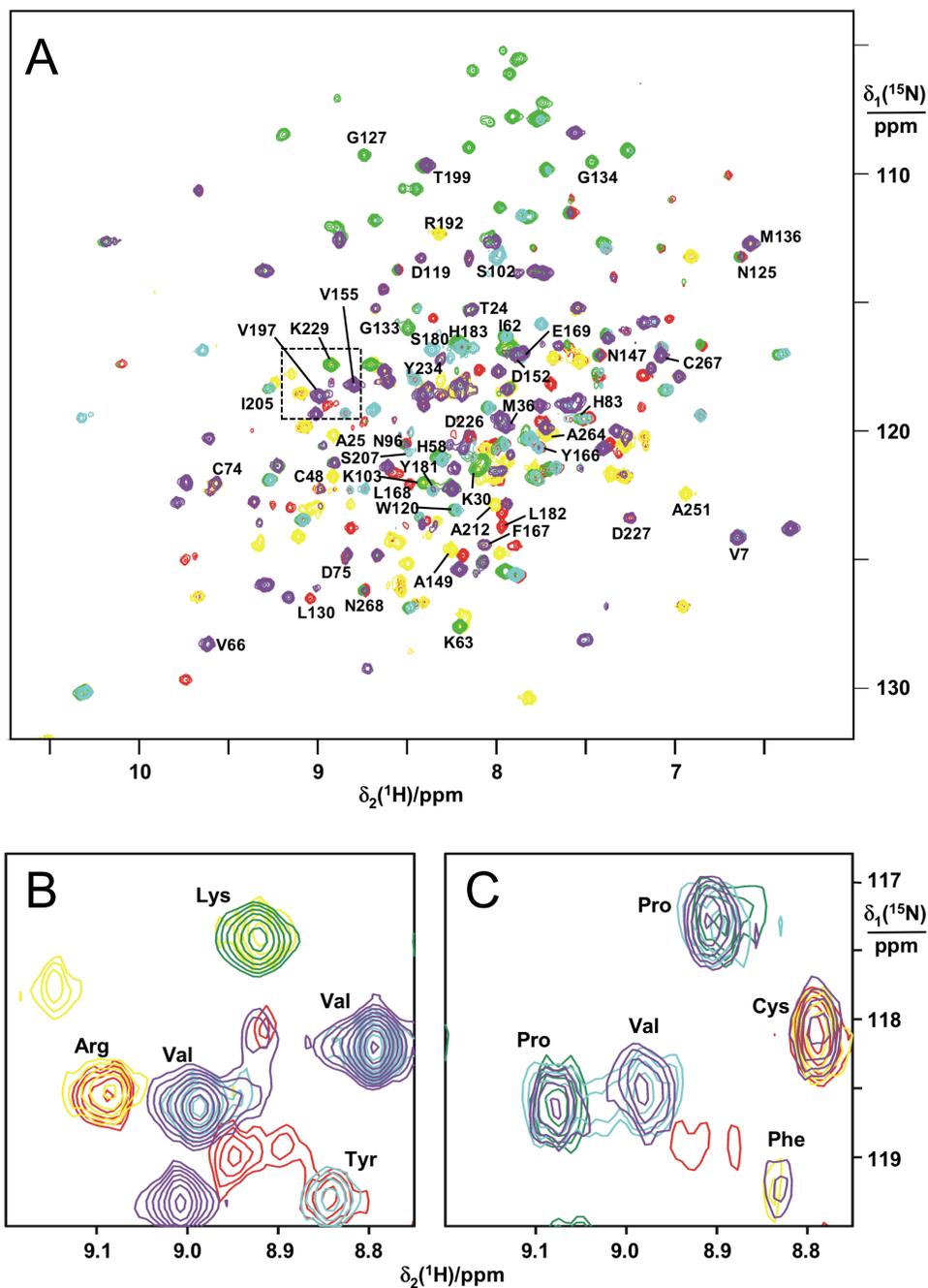
Figure S4. NMR resonance assignments of backbone amides of α270 by combinatorial isotope labeling. **(A)** Superimposition of ^{15}N-HSQC spectra of five combinatorially ^{15}N-labeled samples of α270. The five spectra are distinguished by color (red, yellow, green, cyan, purple). Cross-peaks that could be assigned to backbone amides of specific residues, based on their own amino acid type and on the amino acid type of their preceding residue, are labeled with their assignment (see Supplementary Table S2). **(B)** Spectral region marked by a box in (A). Residue types identified by the combinatorial labeling are indicated. **(C)** Superimposition of the same region of the 2D HN(CO) spectra of five combinatorially ^{13}C-labeled samples produced with uniform ^{15}N-labeling to identify the preceding residue types (41). Residue types of preceding residues are indicated. For three of the peaks, the amino acid pairs identified in this way are unique in the amino acid sequence, leading to the sequence-specific resonance assignment shown in (A).
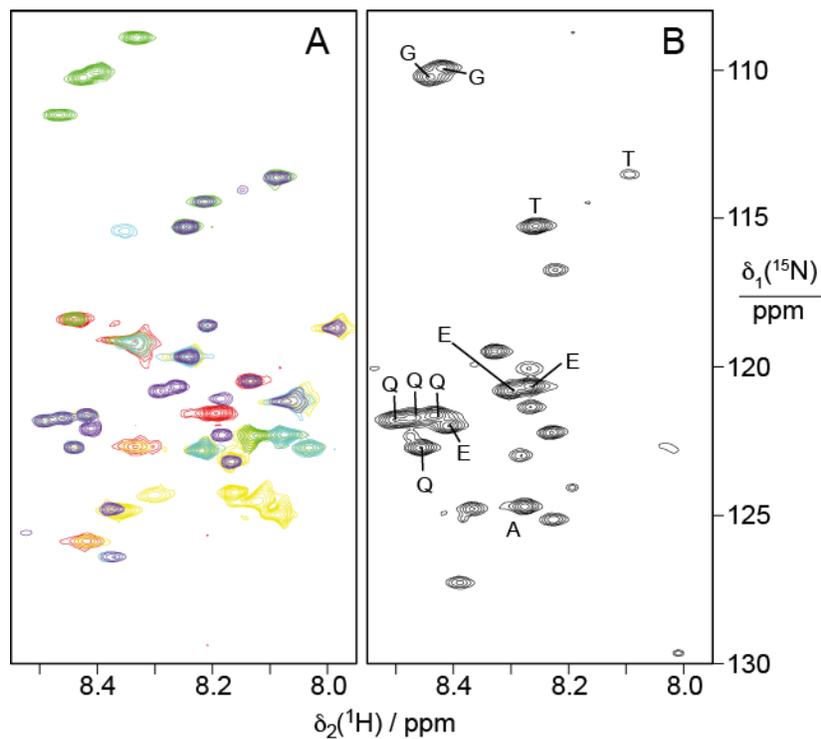
Figure S5. Resonance assignments of residues of εCTS_{59} in the complex with $\alpha270$ and evidence for enhanced mobility of residues comprising the Q-linker in the εCTS in full-length ε_L in complex with θ and α_{GL}. **(A)** Superimposition of the ^{15}N-HSQC spectra obtained by combinatorial ^{15}N-labeling of εCTS_{59}. Sequence specific resonance assignments were obtained by 3D HNCA and HN(CO)CA spectra. **(B)** ^{15}N-HSQC spectrum of ^{15}N-labeled ε_L in complex with unlabeled θ and α_{GL}. Peaks are labeled with the residue types that could be identified by comparison with the spectra in (A). The close similarity of the chemical shifts with those in (A) demonstrates that many residues of the εCTS are mobile also in the 66.5 kDa complex of $\alpha270$ with full-length ε and θ.
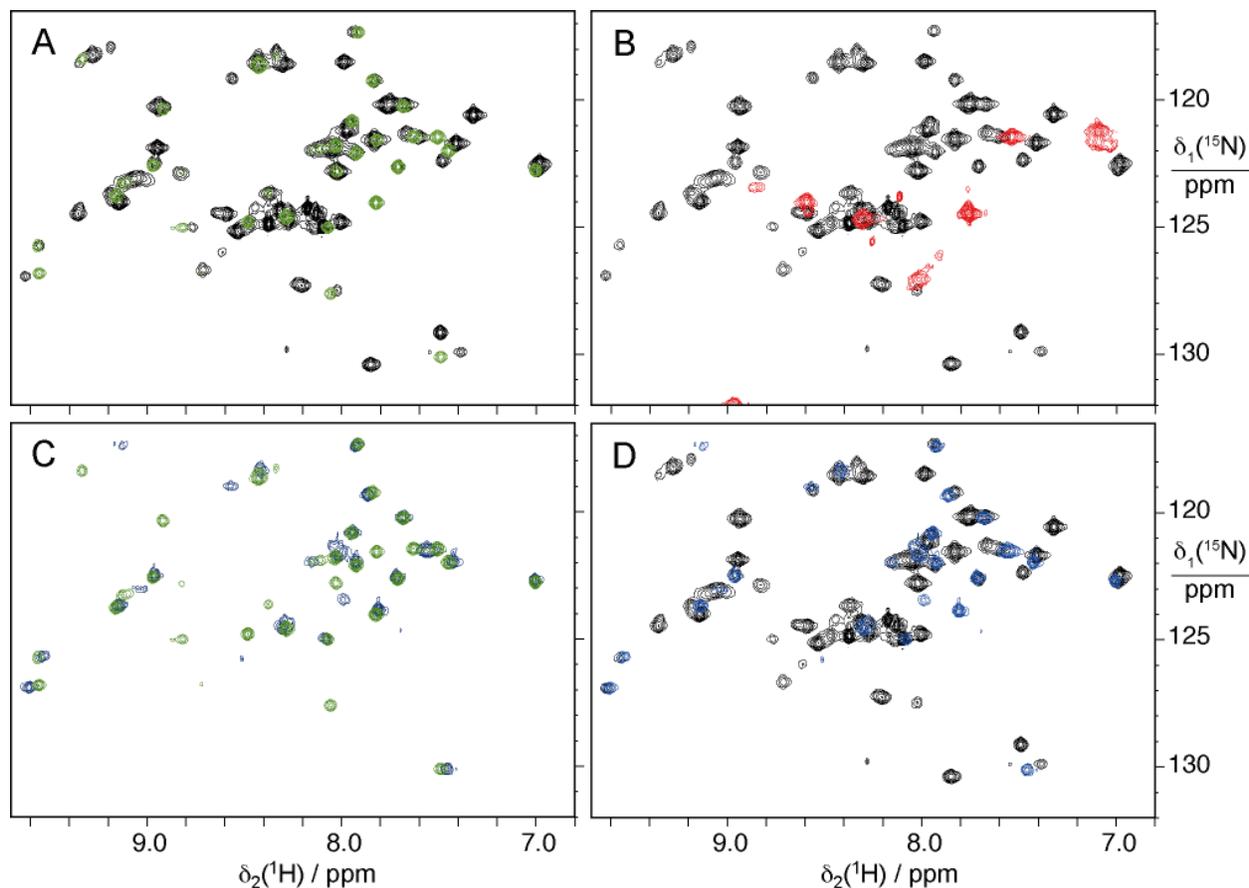
Figure S6. Comparison of the structural integrity of α270 following fusion with the εCTS N-terminal (construct A) or C-terminal (construct B) of the α270 domain (Figure 1C). For improved spectral resolution, all samples were selectively labeled with ^{15}N-alanine by cell-free expression. **(A)** Superimposition of the ^{15}N-HSQC spectrum of construct A (green) onto the ^{15}N-HSQC spectrum of the α270:εCTS$_{59}$ complex (black) produced by co-expression of the two proteins. **(B)** Same as (A), except that the ^{15}N-HSQC spectrum of construct B is superimposed (red). The NMR spectrum shows much less similarity to the α270:εCTS$_{59}$ spectrum than that of construct A, showing that construct A preserves the structure of α270 better. **(C)** Impact of the Leu21Pro mutation on the structural integrity of α270. Comparison between the spectrum of construct A with (blue) and without (green) the Leu21Pro mutation shows that the mutation perturbs a significant number of cross-peaks but leaves most peaks intact. (D) Superimposition of the spectrum of the Leu21Pro construct A (blue) on that of the α270:εCTS$_{59}$ complex (black).

17
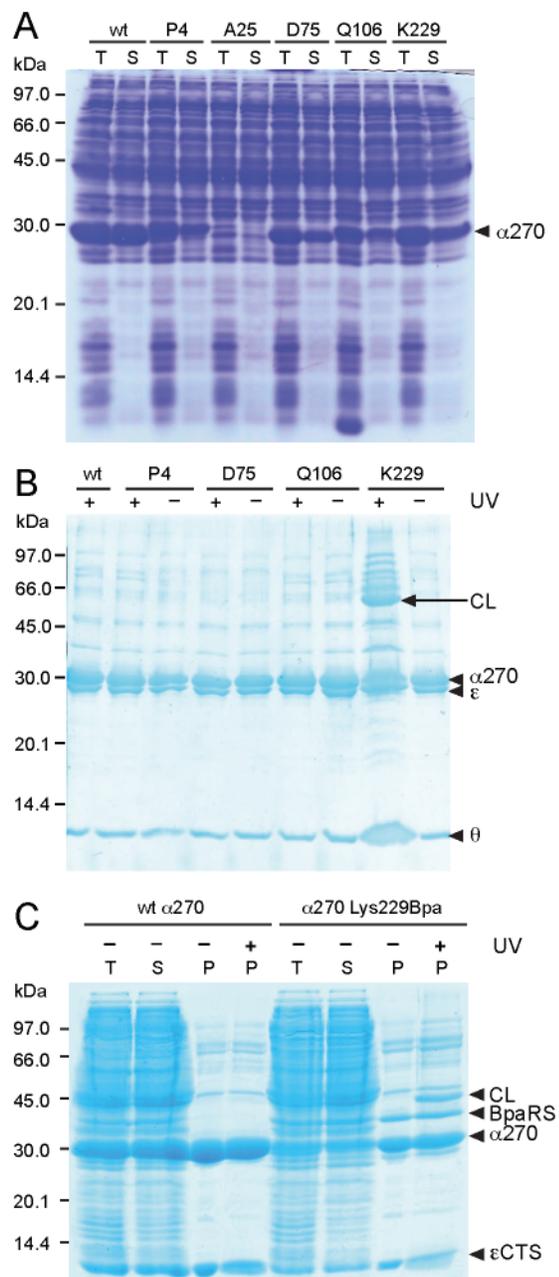
Figure S7. Photo-crosslinking of benzoylphenylalanine (Bpa), that had been introduced into α270 at different sites, with ε. **(A)** Cell-free production of the α270 mutants Pro4Bpa, Ala25Bpa, Asp75Bpa, Gln106Bpa, and Lys229 Bpa; 15% SDS-PAGE with Coomassie blue staining. The band of α270 is indicated; T, total reaction mixture; S, soluble fraction. **(B)** The complex of α270 with ε and θ can be photo-crosslinked for the mutant Lys229Bpa but not for the mutants Pro4Bpa, Asp75Bpa, Gln106Bpa, or the wild-type; 15% SDS-PAGE with Coomassie blue staining. The bands of α270, ε and θ and the crosslinked product of the Lys229Bpa mutant (CL) are indicated on the right. The + and – symbols identify the samples prepared with and without UV irradiation at 312 nm. **(C)** The Lys229Bpa mutant of α270 can similarly be photo-crosslinked with εCTS$_{59}$ instead of the ε:θ complex, indicating that the crosslink is with the εCTS also in full-length ε. The control with wild-type α270 shows that crosslinking depends on incorporation of Bpa; 15% SDS-PAGE with Coomassie blue staining. The bands of α270 and εCTS$_{59}$ and the crosslinked product of the Lys229Bpa mutant (CL) are indicated. Purification using the His$_6$ tag of α270 also co-purified the His$_6$ tagged aminoacyl-tRNA synthetase (BpaRS) used for cell-free production of the Lys229Bpa mutant of α270. The + and – symbols identify the samples prepared with and without UV irradiation at 312 nm. T, total cell free reaction product; S, soluble fraction; P, purified complex.
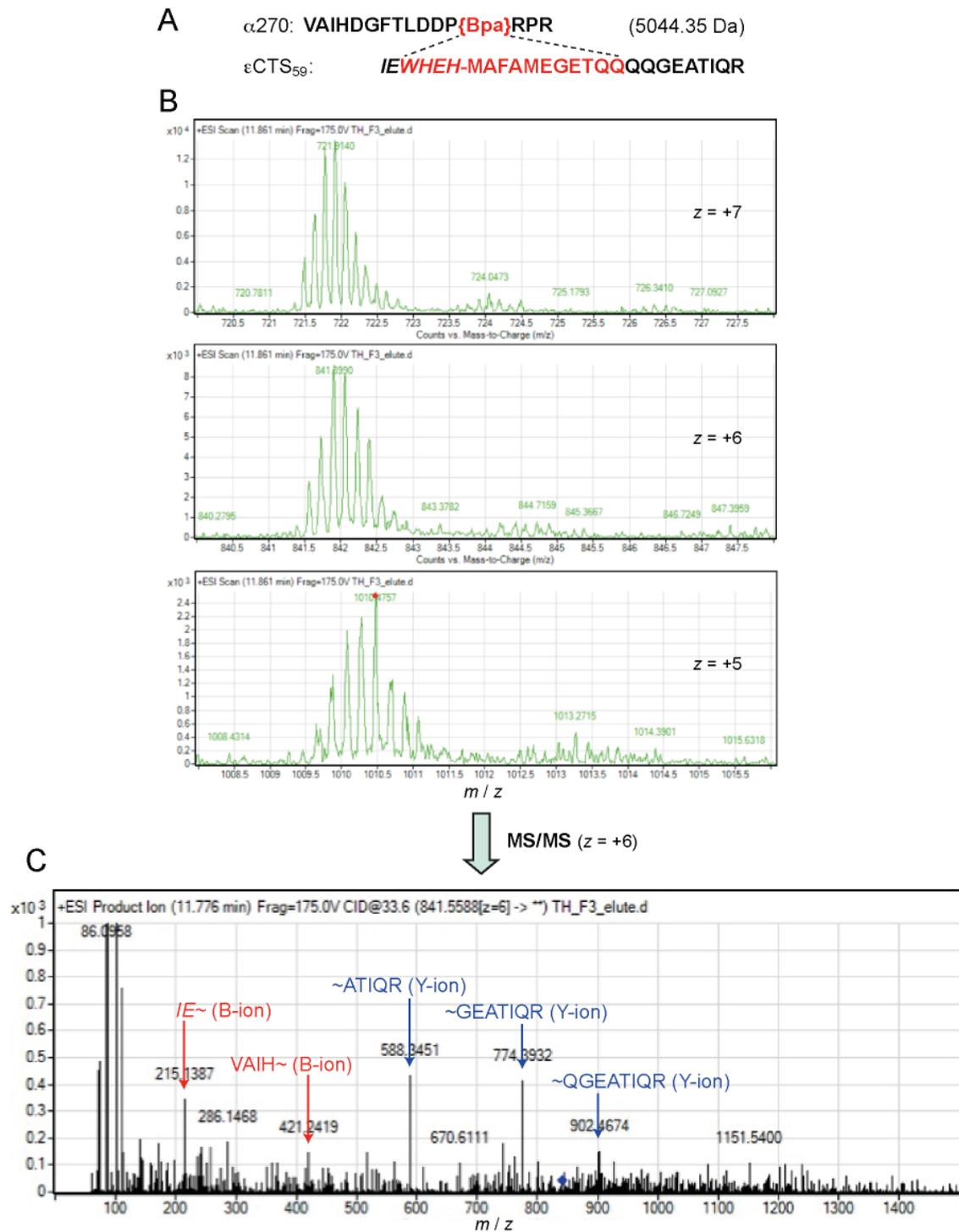
Figure S8. MS analysis of an in-gel tryptic digest of the crosslinked α270(Lys229Bpa)–εCTS$_{59}$ product narrows the cross-linking site to a peptide N-terminal of the glutamine-rich segment in εCTS. **(A)** Sequence of the crosslinked peptide (5044.35 Da), identified by **(B)** high-resolution LS/ESI-MS; 5+, 6+ and 7+ ions are shown in different regions of the spectrum. **(C)** MS/MS analysis of the 6+ ion from (B) limits the site of crosslinking to the region preceding Gln196 of εCTS$_{59}$, highlighted in red in (A). cf. Figure 1B.
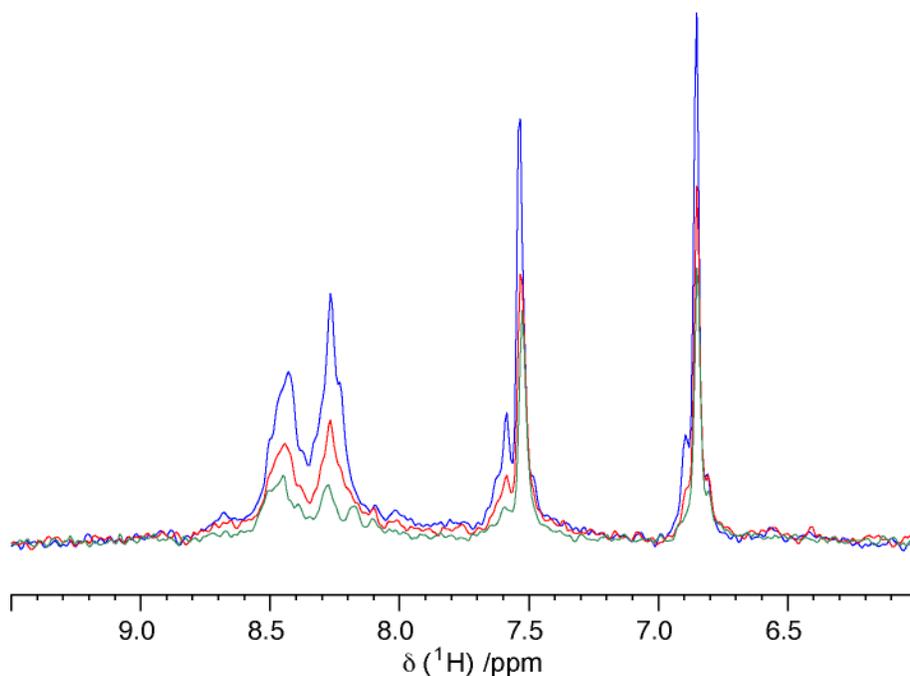
Figure S9. The β_2 clamp binds tightly to the α_{GL}:ϵ_L:θ complex, yet the high mobility of some of the residues of ϵ_L is only moderately attenuated. The first FIDs of ^{15}N-HSQC spectra were recorded of ^{15}N-labeled ϵ_L in complex with unlabeled His$_6$-α_{GL} and θ, followed by the addition of increasing amounts of β_2. The concentration of the α_{GL}:ϵ_L:θ complex was 0.1 mM. Blue: His$_6$-α_{GL}:[15]N-ϵ_L:θ complex; red: His$_6$-α_{GL}:[15]N-ϵ_L:θ with one equivalent of β_2 (as dimer); green: His$_6$-α_{GL}:[15]N-ϵ_L:θ with two equivalents of β_2. A spectrum recorded with four equivalents of β_2 was identical to that with two equivalents (not shown). The failure to attenuate all signals from the mobile segment of [15]N-ϵ_L indicates that the high mobility of some residues in the ϵCTS is attenuated but not abolished in the α_{GL}:[15]N-ϵ_L:θ:β_2 complex.

Supplementary References:

9.   Ozawa,K., Jergic,S., Park,A.Y., Dixon,N.E. and Otting,G. (2008) The proofreading exonuclease subunit ε of *Escherichia coli* DNA polymerase III is tethered to the polymerase subunit α via a flexible linker. *Nucleic Acids Res*., **36**, 5074–5082.

16.   Pintacuda,G., Park,A.Y., Keniry,M.A., Dixon,N.E. and Otting,G. (2006) Lanthanide labeling offers fast NMR approach to 3D structure determinations of protein-protein complexes. *J. Am. Chem. Soc*., **128**, 3696–3702.

25.   Jergic,S., Horan,N.P., Elshenawy,M.M., Mason,C.E., Urathamakul,T., Ozawa,K., Robinson,A., Goudsmits,J.M.H., Wang,Y., Pan,X., Beck,J.L., van Oijen,A.M., Huber,T., Hamdan,S.M. and Dixon,N.E. (2013) A direct proofreader–clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J*., in press. doi: 10.1038/emboj.2012.347.

27.   Williams,N.K., Prosselkov,P., Liepinsh,E., Line,I., Sharipo,A., Littler,D.R., Curmi,P.M.G., Otting,G. and Dixon,N.E. (2002) *In vivo* protein cyclization promoted by a circularly-permuted *Synechocystis* sp. PCC6803 DnaB mini-intein. *J. Biol. Chem*., **277**, 7790–7798.

33.   Neylon,C., Brown,S.E., Kralicek,A.V., Miles,C.S., Love,C.A. and Dixon,N.E. (2000) Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: A model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry*, **39**, 11989–11999.

36.   Hamdan,S., Bulloch,E.M., Thompson,P.R., Beck,J.L., Yang,J.Y., Crowther,J.A., Lilley,P.E., Carr,P.D., Ollis,D.L., Brown,S.E and Dixon,N.E. (2002) Hydrolysis of the 5'-*p*-nitrophenyl ester of TMP by the proofreading exonuclease (ε) subunit of *Escherichia coli* DNA polymerase III. *Biochemistry*, **41**, 5266–5275.

41.   de la Cruz,L., Nguyen,T.H.D., Ozawa,K., Shin,J., Graham,B., Huber,T. and Otting,G. (2011) Binding of low-molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease – fold analysis by pseudocontact shifts. *J. Am. Chem. Soc*., **133**, 19205–19215.

55.   Wing,R.A., Bailey,S. and Steitz,T.A. (2008) Insights into the replisome from the structure of a ternary complex of the DNA polymerase III α-subunit. *J. Mol. Biol*., **382,** 859–869.

60.   Wijffels,G., Dalrymple,B.P., Prosselkov,P., Kongsuwan,K., Epa,V.C., Lilley,P.E., Jergic,S., Buchardt,J., Brown,S.E., Alewood,P.F., Jennings,P.A. and Dixon, N.E. (2004) Inhibition of protein interactions with the β_2 sliding clamp of *Escherichia coli* DNA polymerase III by peptides from β_2-binding proteins. *Biochemistry*, **43**, 5661–5671.

61.   Beckett,D., Kovaleva,E. and Schatz,P.J. (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci*., **8**, 921–929.

62.   Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*., **185**, 60–89.

63.   Jeruzalmi,D., Yurieva,O., Zhao,Y., Young,M., Stewart,J., Hingorani,M., O'Donnell,M. and Kuriyan,J. (2001) Mechanism of processivity clamp opening by the δ subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell*, **106**, 417–428.

64.   Wolff,P., Olieric,V., Briand,J.P., Chaloin,O., Dejaegere,A., Dumas,P., Ennifar,E., Guichard,G., Wagner,J. and Burnouf,D. (2010) A three step model for peptide ligand binding onto the *E. coli* processivity ring. DOI: 10.2210/pdb3q4l/pdb.

65.   London,N., Raveh,B., Cohen,E., Fathi,G. and Schueler-Furman,O. (2011) Rosetta FlexPepDock web server – high resolution modeling of peptide-protein interactions. *Nucleic Acids Res*., **39**, W249–253.