OS R PR-00001 14.08.2018
Crystal structure determination of *Ct*UGGT-N *aka Ct*UGGT S180C:T742C

Summary. Two crystal structures of the double Cys mutant *Ct*UGGT-N *aka Ct*UGGT S180C:T742C have been determined by X-ray crystallography. The trigonal one, in space group $P3_212$ was determined at 4.7 Å resolution by molecular replacement using PDB ID 5NV4 as a search model, against the data collected by Mario Hensen and Roberta Ibba on I24@Diamond on 08.08.2018. The orthorhombic one, in space group $P2_12_12_1$, was determined to 4.6 Å resolution by molecular replacement using the $P3_212$ *Ct*UGGT-N model as a search model, against the data collected by Pietro Roversi and Roberta Ibba on I03@Diamond on 05.05.2018. Both structures show the disulphide bridge across the engineered double mutation S180C:T742C, and the TRXL1 and TRXL3 domains clamped shut. In the $P3_212$ form the orientation of TRXL2 with respect to the rest of the protein is the most opened ever observed so far (a rotation of 60 degrees away from the TRXL2 in PDB ID). The orthorhombic form proves that the clamped shut structure (*Ct*UGGT-N $P2_12_12_1$) is compatible with the bent shut structure (*Ct*UGGT-H, $P2_12_12_1$, PDB ID 5NV) *i.e.* both clamping and bending movements can close at the same time. A quadruple *Ct*UGGT mutant, containing both D611C:G1050C and S180C:T742C, should be feasible and its human UGGT1 equivalent could provide the best rigidification of the human structure conceivable so far.

Crystal growth and cryoprotection
The trigonal $P3_212$ *Ct*UGGT-N crystal comes from drop E9_1 of Tray 4, bar code label MC004659, set up by Johan C. Hill on 26.04.2018, using *Ct*UGGT-N purified by Pietro Roversi on 25.04.2018. Protein at OD_{280}=7.29 in HEPES 20 mM pH 6.5, 50 mM NaCl, 5 mM UDP-Glc, 1 mM $CaCl_2$. The condition is E9 of the MORPHEUS2 screen [2 mM Lanthanides, 0.1 M Buffer System 6 (1.0M, pH 8.5 at 20 °C, Gly-Gly, AMPD), pH 8.5, 50 % v/v Precipitant Mix 5 (30% w/v PEG 3000, 40% v/v 1, 2, 4- Butanetriol, 2% w/v NDSB 256)] mixed in protein:mother liquor ratio 100 nL:100 nL. The crystal (nicknamed "Mario") grew between day 57 and day 71, at 18 °C (Figure 1). The crystal was cooled down in liquid N_2 in August 2018. Self-cryoprotected as it comes from a MORPHEUS screen.

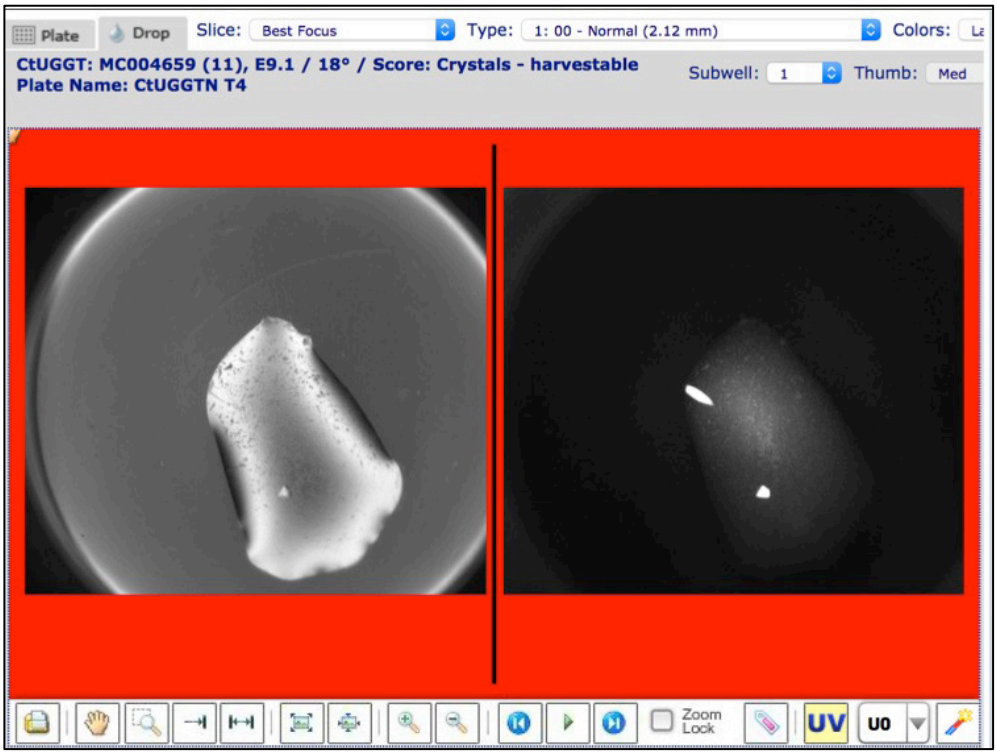

Figure 1. *Ct*UGGT-N, P3$_2$12, T4-E9_1_MH32, "Mario". from the Minstrel server at Day 71 (07.06.2018): left-hand side, visibile light; right-hand side, UV light.

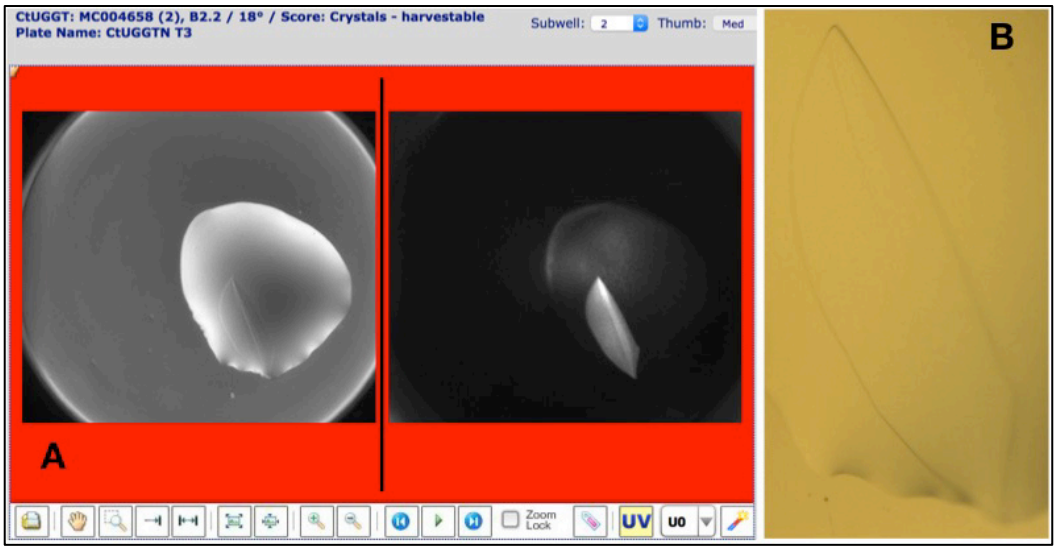

Figure 2. *Ct*UGGT-N, P212121, T3-B2_2, the "camera glare". A: from the Minstrel server at Day 2 (28.04.2018): left-hand side, visibile light; right-hand side, UV light. B: from the optical microscope, immediately before fishing (01.05.2018)

The orthorhombic $P2_12_12_1$ *Ct*UGGT-N crystal comes from drop B2_2 of Tray 3, bar code label MC004658, set up by Johan Hill on 26.04.2018, using *Ct*UGGT-N purified by Pietro Roversi on 25.04.2018. Protein at OD_{280}=7.29 in HEPES 20 mM pH 6.5, 50 mM NaCl, 5 mM UDP-Glc, 1 mM $CaCl_2$. The condition is B2 of the JCSG+ screen (0.2 M Sodium thiocyanate, 20% w/v PEG 3350) mixed in protein:mother liquor ratio 133 nL:66 nL. The crystal (nicknamed "the camera glare", Figure 2) grew in two days at 18 °C and broke into smaller pieces upon fishing. The crystal was cooled down in liquid N_2 by Roberta Ibba and Pietro Roversi on 01/05/2018. Cryoprotecting

solution made by mixing 2 µL of ethylene glycol in 8 µL of mother liquor (i.e. 20% EG).

X-ray diffraction.
X-ray diffraction from the crystal *Ct*UGGT-N T3-B2_2-xtal3 was collected by Pietro Roversi and Roberta Ibba on I03@Diamond on 01.05.2018, (file path: `/dls/i03/data/2018/mx18069-31/Zitzmann/CtUGGTN/T3-B2-2_3/T3-B2-2_3_1_####.cbf`);

X-ray diffraction from the crystal *Ct*UGGT-N T3-B2_2-xtal3 was collected by Mario Hensen and Roberta Ibba on I24@Diamond on 08.08.2018 (file path: `/dls/i24/data/2018/mx18069-40/CtUGGT/MH_32/MH_32_1_####.cbf`).

See Figure 3 for the data collection parameters.

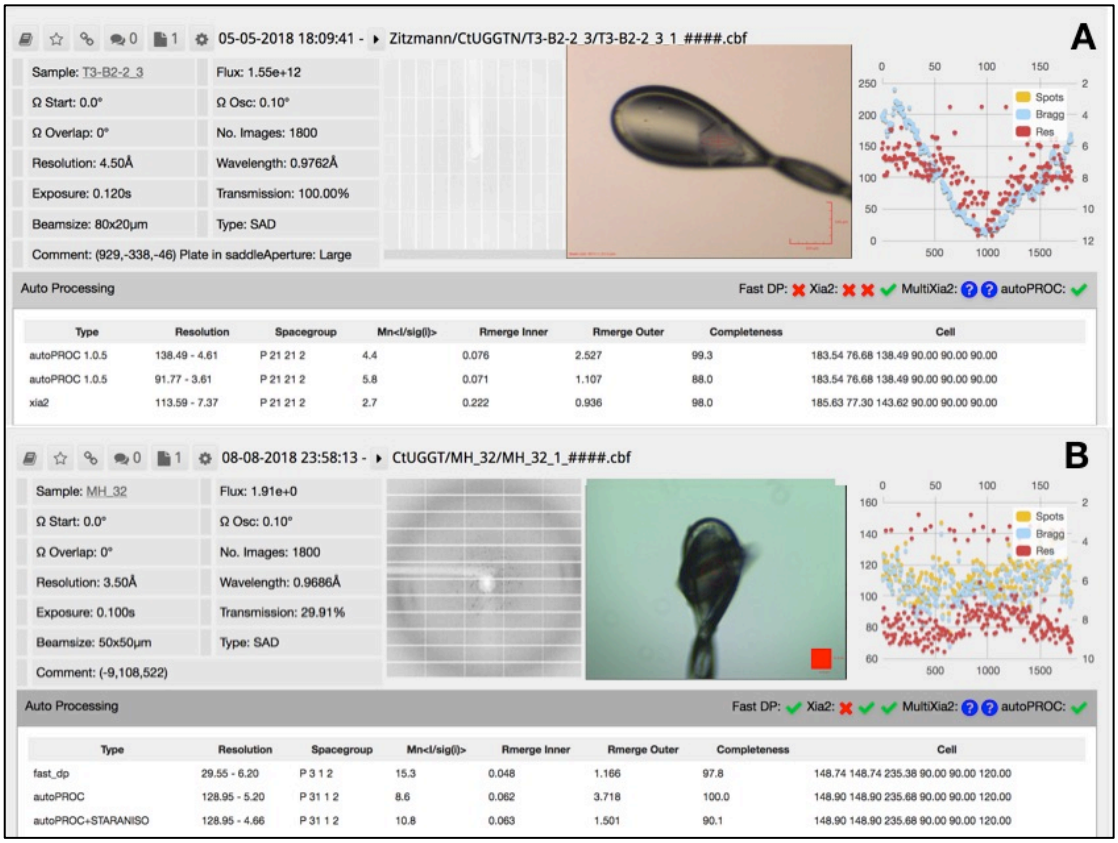

Figure 3. X-ray data collection of *Ct*UGGT-N. A: T3-B2_2-xtal3 on I03@Diamond 05.05.2018. **B.** T4-E9_1_MH32 on I24@Diamond 08.08.2018.

Data processing
Both datasets were processed with the `autoPROC` suite of programs (1). Table 1 contains the data processing statistics.

T4-E9_1_MH32: SRF and extinctions indicate either of $P3_212/P3_112$
The T4-E9_1_MH32 crystal indexes in a primitive hexagonal cell a=b= 148.858, c=235.610. The data scaled in $P3_1$ (so as not to bias the symmetry towards extra twofold axes) have a self rotation function compatible with

point group 312 (Figure 4): this restricts the possible space groups to trigonal primitive space groups in point group 312 *i.e.* either P312, $P3_112$ or $P3_212$.

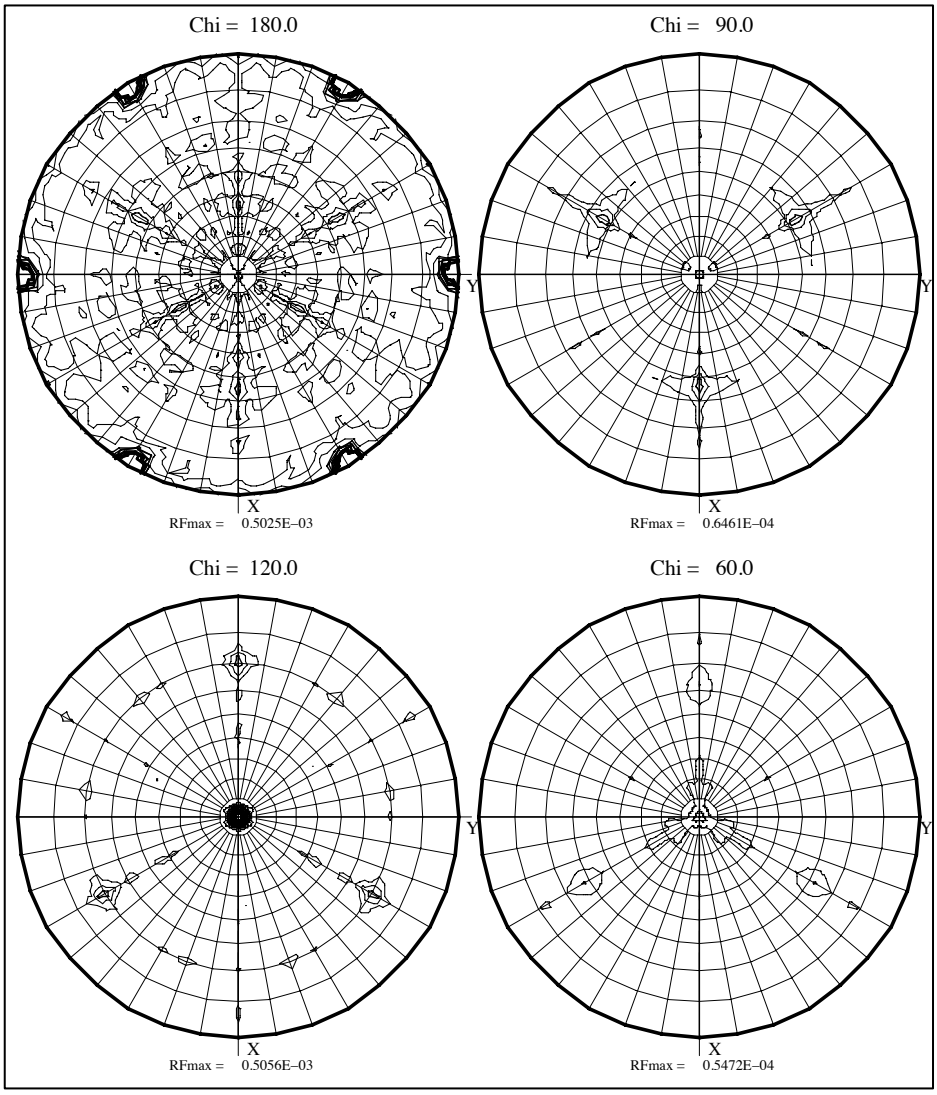

Figure 4. *Ct*UGGT-N T4-E9_1_MH32 displays 312 point symmetry. Sections at κ=180°, 120°, 90° and 60° of the self-rotation function computed from data scaled in $P3_1$. The peak at at θ=90°, ϕ=90° κ=180° (of the same intensity as the threefold at θ=0°, ϕ=0° in the κ=120° section) and the absence of a sixfold axis in the κ=60° section indicate point group 312.

More in detail, on the basis of the absence of a peak in the κ=60° self-rotation function section we can discard all hexagonal space groups (there is no sixfold axis in the κ=60° section along **c** so the crystal cannot be hexagonal); we can also discard primitive trigonal space groups of point group 321 or 3, because of the set of three twofold axes visible in the κ =180° section, which are directed as the long diagonal of the cell and the two threefold-related directions, indicating point group 312.

Then, the pattern of extinctions along the **c*** axis shows an alternation of strong 00L reflexions with L even, and weak or null intensity 00L reflexions with L odd: this is only compatible with either a 3_1 or a 3_2 axis along **c**, eliminating P312 and leaving only $P3_112$ or $P3_212$.

Table 1. *Ct*UGGT-N data processing statistics. In parentheses the values referring to the highest resolution shell.

PDB ID	**TBA**	
Space Group (Z)	$P2_12_12_1$ (4)	$P3_212$ (6)
X-ray source	I03@DLS	I24@DLS
Detector	Pilatus	
Wavelength (Å)	0.97622	0.96861
a (Å)	76.678	148.90
b (Å)	138.490	148.90
c (Å)	183.535	235.68
Resolution Limits (Å)	138.49-4.61	128.95-4.66
Completeness (%)	99.3 (99.1)	100.0 (100.0)
Measured Reflections	59,127 (2,665)	155,474 (45,237)
Unique Reflections	11,219 (541)	16,192 (4,574)
Multiplicity	5.3 (4.9)	9.6 (9.9)
$CC_{1/2}$	0.98 (0.43)	0.99 (0.22)
R_{meas}	0.46 (2.8)	0.22 (4.92)
< I/σ(I) >	4.4 (1.6)	6.4 (0.5)

T4-E9_1_MH32: solvent content analysis, likely 1 copy/asu and 72% solvent

`CCP4-Matthews` was run in $P3_112$ for the T4-E9_1_MH32 data**,** declaring a molecule of MW~174,000 Dalton. The analysis indicates either 1 copy/asymmetric unit with 72% solvent or two copies/asymmetric unit with 43% solvent. No extra NCS twofolds are visible in the SFR in the K=180° section (only crystallographic axes are visible), so it is highly likely that it is 1 copy/asu and 72% solvent, which also makes sense given the poor diffraction (this crystal may have high solvent content and therefore lousy packing).

T4-E9_1_MH32 initial phases: molecular replacement indicates $P3_212$, 1 copy/asu and 72% solvent

`CCP4-Molrep` was run in $P3_112$ and $P3_212$ searching with a copy of PDB ID 5NV4 from which all three TRXL1,2,3 were removed (leaving only TRXL4, BS1,BS2 and GT24 domains). The results were clearly better in $P3_212$ ($P3_212$ has wR=0.606, Score=0.435, TF/sigma=10.35, Contrast=9.21 *vs.* $P3_112$ wR=0.637, Score=0.372, TF/sigma=5.61, Contrast=4.76).

T4-E9_1_MH32 refinement and model building

The first `autoBUSTER` refinement was run with the following options:

```
refine -m CtUGGT-N-MH_32-P3212.mtz \
-p 5nv4.noTRXL1-TRXL2-TRXL3_molrep2.pdb \
-TLS January18.noTRXL123.AB.tls \
-RB January18.noTRXL123.AB.rb \
-target ../5nv4.pdb -d P3212-noTRXL1-TRXL2-TRXL3 \
StopOnGellySanityCheckError=no
```

The first map obtained in autoBUSTER from this MR model shows strong density for the TRXL3 domain and the Ca^{2+} site in the GT24 domain. The TRXL3 domain was added by superposing PDB ID 5NV4 onto the model, and real-space fitting the domain to the Fo-Fc map in `CCP4-coot`. After one more round of refinement, the TRXL1 domain was added in the same way. Finally, the TRXL2 domain was added by molecular replacement with `CCP4-Molrep` in presence of the rest of the structure. The structure was refined in `autoBUSTER (2)` with one TLS body per domain, one rigid body per domain, with external restraints (3) to PDB ID 5NV4. Table 2 reports the Rfactors and geometry statistics during the model building and refinement.

Model	R	Rfree	$rmsd_{bonds}$ (Å)	$rmsd_{angles}$ (°)
Delta TRXL1,2,3	0.397	0.467	0.007	1.0
Delta TRXL1,2	0.384	0.390	0.007	1.1
Delta TRXL2	0.311	0.320	0.007	1.1
Whole molecule	0.179	0.243	0.01	1.3

Table 2. Refinement and model building of T4-E9_1_MH32 in $P3_212$.

Fo-Fc residuals on two sites (the catalytic site and a crystal contact between TRXL2 and one of its symmetry mates) suggested a lanthanide ion from the crystallisation mix (which contains Y^{3+}, Tb^{3+}, Er^{3+}, Yb^{3+}. It is likely either Er^{3+} or Tb^{3+}, which are known to substitute for Ca^{2+} and Mn^{2+} in protein coordination sites (4-6). At the wavelenght of data collection, =0.96861, f'_{Er3+}=-1.7235 and f''_{Er3+}=8.2682, while f'_{Tb3+}=-1.046 and f''_{Tb3+}=6.9753. Peaks at +9.4 and +7.4 sigmas are indeed visible at these two sites in the anomalous Fourier difference map. The ions were modelled as Tb^{3+}, with a Tb^{3+}-O distance of 2.4 ± 0.3 Å (coordinating residues: site 1: D1302, D1304, D1435; site 2: E774 from a symmetry mate, E713, E716 and D818).

T3-B2_2 molecular replacement confirms $P2_12_12_1$, 1 copy/asu and 72% solvent

`Phaser` was used to determined the structure of the *Ct*UGGT-N T3-B2_2 orthorhombic xtal form by Molecular Replacement, searching in all primitive orthorhombic spacegroups, using the T4-E9_1_MH32 $P3_212$ structure as a

search model but having deleted the TRXL2 domain (declaring a RMSD of 2.0 Å - `Phaser` refined it to 0.9 Å) . Top scores: LLG=620, TFZ=28.2. The model was completed by addition of the TRXL2 domain, by molecular replacement using `CCP4-Molrep`. The lattice is the same as PDB ID 5NV4, i.e. the *Ct*UGGT-H, D611C:G1050C double mutant. The structure was refined in `autoBUSTER (2)` with one TLS body per domain, one rigid body per domain, with external restraints (3) to PDB ID 5NV4. Current R=0.209, Rfree=0.285; $rmsd_{bonds}$=0.01 Å; $rmsd_{angles}$=1.26°.

The orhtorhombic form proves that the clamped shut structure (*Ct*UGGT-N $P2_12_12_1$) is compatible with the bent shut structure (*Ct*UGGT-H, $P2_12_12_1$, PDB ID 5NV) i.e. both clamping and bending movements can close at the same time. A quadruple *Ct*UGGT mutant D611C:G1050C:S180C:T742C should therefore be feasible and its human UGGT1 equivalent could provide the best rigidification of the human structure conceivable so far

Figure 5 shows the *Ct*UGGT-N T4-E9_1_MH32 $P3_212$ structure (cyan) overlaid using `Theseus`(7) onto PDB ID 5NV4 (*aka* *Ct*UGGT-H, D611C:G1050C**,** green). The TRXL1 and TRXL3 domains move together because of the engineered disulphide S180C:T742C. The orientation of TRXL2 with respect to the rest of the protein is the most opened ever observed so far (a rotation of 60 degrees away from the TRXL2 in PDB ID 5NV4).

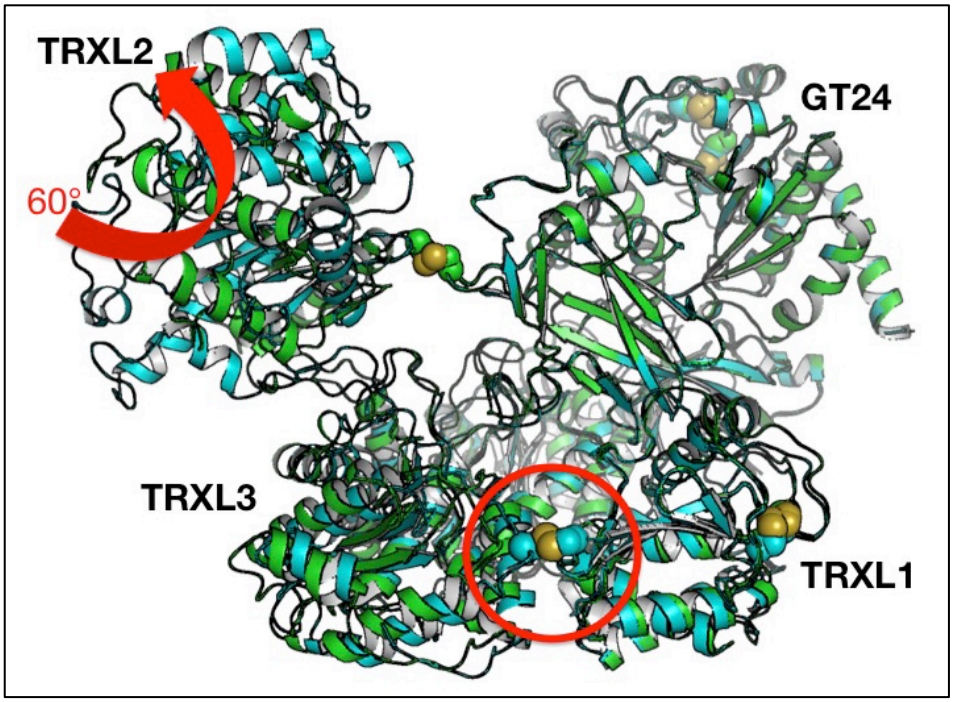

Figure 5. *Ct*UGGT-N and *Ct*UGGT-H structures. T4-E9_1_MH32 $P3_212$ structure (cyan) overlaid onto PDB ID 5NV4 (*aka* *Ct*UGGT-H, D611C:G1050C, green). Disulphide bridges in spheres representation. The TRXL1 and TRXL3 domains move together because of the engineered disulphide S180C:T742C (red circle). The orientation of TRXL2 with respect to the rest of the protein is the most opened ever observed so far (a rotation of 60 degrees away from the TRXL2 in PDB ID, red arrow).

References

1. Vonrhein C, et al. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):293–302.

2. Bricogne G, et al. (2017) BUSTER 2.10.3.

3. Smart OS, et al. (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):368–380.

4. Colman PM, Weaver LH, Matthews BW (1972) Rare earths as isomorphous calcium replacements for protein crystallography. *Biochemical and Biophysical Research Communications* 46(6):1999–2005.

5. Miake-Lye RC, Doniach S, Hodgson KO (1983) Anomalous x-ray scattering from terbium-labeled parvalbumin in solution. *Biophysical Journal* 41(3):287–292.

6. Hsiao Y-Y, et al. (2009) Crystal structure of CRN-4: implications for domain function in apoptotic DNA degradation. *Mol Cell Biol* 29(2):448–457.

7. Theobald DL, Steindel PA (2012) Optimal simultaneous superpositioning of multiple structures with missing data. *Bioinformatics* 28(15):1972–1979.