Structure of the SUN domain defines features of a molecular bridge in the nuclear envelope

Zhaocai Zhou[1,2,6], Xiulian Du[1,6], Zheng Cai[1,6], Xiaomin Song[1], Hongtao Zhang[1], Takako Mizuno[4], Emi Suzuki[4], Marla Rosanne Yee[1], Alan Berezov[1,4], Ramachandran Murali[1,4], Shiaw-Lin Wu[5], Barry L. Karger[5], Mark I. Greene[1,4*], and Qiang Wang[1,3,4*]

[1]Department of Pathology and Laboratory Medicine, University of Pennsylvania. Perelman School of Medicine, Philadelphia, PA, USA 19104-6082

[2current address-]State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China 200031

[3 current address-]Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, [4]Department of Biomedical Sciences, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048

[5]Barnett Institute, Northeastern University, Boston, MA, USA 02115

[6]These authors contributed equally to this work.

* Corresponding authors:

Qiang Wang and Mark I. Greene, 252 John Morgan Building, 36th Street and Hamilton Walk, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104-6082, Phone: 215-898-2868, Fax: 215-898-2401, Email: qiang.wang@cshs.org; greene@reo.med.upenn.edu

MATERIALS AND METHODS

Constructs Preparation

The SUN domain (V519-H716) was amplified with standard PCR protocol using human SUN2 cDNA as a template. The forward primer is AAA AAA CAT ATG GTG ATT GGA GTG ACA GAG G; the reverse primer is CCG CTC GAG CTA GTG ATG GTG ATG GTG GTG GGC GGG CTC CCC ATG C. The fragment was cloned into pET21a(+) vector with Nde1 and Xho1 sites, which introduced a C-terminal 6xHis tag.

The QuikChange site-directed mutagenesis kit (Stratagene) was used to make the SUN domain mutants. PCR was performed using the following primers:

ZZSun2F91Af   ggggccacaaggcgccgccgtggtccgc

ZZsun2F91Ar   gcggaccacggcggcgccttgtggcccc

ZZSun2A128Ef gcccccaaggacttcgagatctttgggtttgacg

ZZSun2A128Er cgtcaaacccaaagatctcgaagtccttgggggc

ZZSun2L142Af ctgcagcaggaggggacagcccttggcaagttc

ZZSun2L142Ar gaacttgccaagggctgtcccctcctgctgcag

ZZSun2R174Af caggtggtggagctggcgatcctgactaactg

ZZsun2R174Ar cagttagtcaggatcgccagctccaccacctg

ZZSun2L176Af tggagctgcggatcgcgactaactggggcc

ZZSun2L176Ar ggccccagttagtcgcgatccgcagctcca

D541Nfor cgctacagtgagaaccgcatcgggctggca

D541Nrev TGCCAGCCCGATGCGGTTCTCACTGTAGCG

G608Dfor TTCCAGGGGCCAGACTTCGCCGTGGTCCGC

G608Drev GCGGACCACGGCGAAGTCTTGTGGCCCCTGGAA

Protein Expression and Purification

The *E.coli* BL21(DE3) cells containing the SUN domain gene were grown in LB medium supplemented with Ampicillin (50μg/mL). when OD_{600} reached 0.6, 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) was added to the culture and the induced expression was maintained for 16 h at 15°C.

Cells were sonicated in a buffer containing 20mM Hepes, 300 mM KCl, pH7.0, 10 mM imidazole, 1% NP-40, and then centrifuged at 12,000 g for 30mins. The supernatant was mixed with Talon metal affinity Resin (Clontech Laboratories Inc), incubating for 2 hours at 4°C. The beads were washed with 20mM Hepes, 300 mM KCl, pH7.0, 20mM imidazole for three times. The protein was eluted with 20 mM Hepes, 300 mM KCl, pH7.0, 100 mM imidazole. The fractions containing SUN domain were pooled and concentrated and further purified by Superdex HR 200 gel filtration column using 20mM Hepes, 300 mM KCl as running buffer. The purified protein was concentrated to approximately 5 mg/ml for crystallization.

The following standard proteins were used to calibrate the column, with the elution volume of the proteins shown in ml.

Protein Standards	Chymotrypsin	Ovalbumin	Albumin
MW (kD)	25	43	67
Volume (ml)	12.7	10.8	10

Crystallization, Data Collection and Structure Determination

Crystallization screening was set up using freshly prepared SUN domain protein. Hanging-drop vapor-diffusion method was used at 4°C by mixing equal amounts (typically 1 μl) of protein sample and well solution. Small crystals appeared in a week from the drop containing 0.1M imidazole, 1.0M sodium acetate, pH6.5. Further optimization by adding 10 mM YCl_3 as additive yielded bigger crystals suitable for diffraction data collection. Heavy atom derivatives were

prepared by soaking native crystals with 30mM K_2PtCl_4 for 30mins or 1mM $LuCl_3$ overnight. Crystals were flash-frozen in liquid nitrogen by using the well buffer containing an additional 30% glycerol as cryoprotectant. Data were collected on a Rigaku Raxis IV++ diffraction system, and processed in the *HKL2000* package (Otwinowski & Minor, 1997). Statistics are shown in Table S1.

Multiple isomorphous replacement and anomalous scattering (MIRAS) were used to obtain phases in the *Solve* and *Resolve* packages (Terwilliger, 2003; Terwilliger & Berendzen, 1999). Initial map allowed identification of partial secondary structural elements including one α-helix and several β-strands. Multiple rounds of manual tracing and model building in *Coot* (Emsley & Cowtan, 2004) and *O* (Jones et al, 1991) were performed combining with refinement in *CNS* (Brunger et al, 1998) and *Refmac* in *CCP4* package (1994). The final model includes residues 520-716 of human Sun2 protein. The asymmetric unit contains only one molecule and it forms a trimer with two adjacent symmetry-related molecules. Model quality was evaluated with *Procheck* (Laskowski et al, 1993). Refinement statistics are shown in SI Table 1. Buried surface area was calculated using the programs *Areaimol* in the *CCP4* package (1994).

Structural Modeling

The peptide "SSEEDYSCTQANNFARSFYPMLRYTNGPPPT" was synthesized by a commercial company. The protein sequence was submitted to NCBI Blast for a homology search against proteins from the PDB database. Homologous fragments with 40% or higher identity were taken from structures corresponding to PDB codes 2OVX, 1FVP and 2RC7 for structural modeling of KASH domain luminal part. Modeling was performed in program *modeller* (Sali & Blundell, 1993) followed manual building. This model was then applied for the building of SUN-KASH complex. The structures of SUN domain and KASH domain were treated as rigid bodies, and the relative orientation between them were manually adjusted based on the hydrophobic

packing and shape complementarity between the putative active site of SUN domain and the C-shaped terminal loop of KASH domain.

Pull Down Assay

The lysate of 293T cells transfected with GFP-SUN2 or FLAG-SUN2 expressing plasmids was incubated with glutathione beads coupled with either GST (G) or GST-NSP2-KASH-LD (N) fusion protein. The proteins bound to the beads were separated by SDS-PAGE, followed by Western Blot using anti-GFP or anti-FLAG antibody.

Isothermal Titration Calorimetry (ITC)

Binding of NSP2 peptide to SUN2 SUN domain protein was investigated by Isothermal titration calorimetry (ITC) using a Microcal VP-ITC (MicroCal Inc., USA) instrument.

0.02 mM recombinant SUN2 SUN domain and 0.4 mM synthesized NSP2 peptide were dialysed against 20 mM HEPES pH 7.3 and 100 mM KCl for 3 times. Then the experiments were initiated by injecting 23x8ul injections of NSP2 peptide from the syringe into the calorimetric cell containing 1.4 ml of SUN2 SUN domain at a fixed temperature of 25°C. Titration curves were fitted to the data using the *Microcal Origin 7.0* (OriginLab Corporation) yielding the stoichiometry N, the binary equilibrium constant K_a (= K_d^{-1}) and the enthalpy of binding. The entropy of binding ΔS was calculated from the relationship $\Delta G = -RT \ln K_a$ and the Gibbs-Helmholtz equation. To calculate the the stoichiometry of SUN-NSP2, the 'One set of Sites' model was used do fitting of the ITC data.

Circular dichroism (CD)

Circular dichroism measurements were carried out on an AVIV Circular Dichroism Spectrometer Model 202. Far-UV CD spectra were measured in the range of 200–250 nm at a protein

concentration of 15 μM with a 0.1 cm path-length cuvette. All the CD measurements were made at 20 °C with 10 seconds averaging time and 1.0 nm wavelength step. Mean residue ellipicity were calculated.

REFERENCE:

(1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50:** 760-763

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54:** 905-921

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-2132

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2):** 110-119

Laskowski R, MacArthur M, Moss D, Thornton J (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26:** 283-291

Otwinowski Z, Minor W (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276:** 307-326

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234:** 779-815

Terwilliger TC (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr D Biol Crystallogr* **59:** 38-44

Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55:** 849-861

Wolf E, Kim PS, Berger B (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci* **6:** 1179-1189

SUPPLEMENTAL FIGURE LEGENDS:

Figure S1. Sequence alignment of the SUN domains.
The C-terminal SUN domain regions of human Sun1-3 and yeast UNC84 were aligned with the secondary structural elements corresponding to human Sun2 labeled above the sequences. Conserved amino acids were shaded. Black squares represent core residues important for the three-helix bundle formation. Black stars represent residues on the CHEF β-sheet surface forming part of the putative KASH binding site. Black dots represent residues involved in the juncture that connects the cloverleaf-like β-sandwich to the stalk-like α-helix.

Figure S2. Trimerization of SUN domain. Panel A shows the shape complementarity between the protomers of SUN domain trimer. Panel B is a space-filling model of the stem-like three-helix bundle. Panel C is a stereo view of the stem with critical residues labeled.

Figure S3. Comparison between SUN domain trimer and TRAF2 trimer. Surface models of SUN domain trimer and TRAF2-CD40 complex were pictured in panel A and B respectively. Two view directions were presented with the active sites highlighted for comparison.

Figure S4. AA'-loop is involved in crystal packing. A stereo view of the interactions between SUN domain trimers in the crystal lattice shows that the extended AA'-loop (in red) mediate high density of crystal contacts.

Figure S5. Pull down assay. Wild-type SUN2 protein binds with nesprin 2G luminal domain (NSP2-KASH-LD), while deletion of the first coiled coil motif (T398-S332) or the C-terminus (R707-H716) of the SUN domain rendered SUN2 incapable of binding to NSP2-KASH-LD. 293T cells were transfected with GFP-SUN2 (A) or the FLAG-SUN2 constructs (B). The cell lysate was incubated with glutathione beads coupled with either GST (G) or GST-NSP2 -KASH-LD fusion protein (N). The proteins bound to the beads were separated by SDS-PAGE, followed by Western Blot using anti-GFP or anti-FLAG antibody. (C) Schematic representations of the SUN2 constructs used in the pull down assay. I: Input (12.5% of total lysate); G: GST; N: GST-NSP2-KASH-LD

Figure S6. ITC analysis of SUN-KASH binding. Panel A shows the binding between the wild type SUN domain protein and the longer version of KASH peptide corresponding to the complete luminal domain of nesprin 2G. Panel B shows the binding between the SUN domain mutant R692A.

Figure S7. D541N and G608D mutations do not affect trimerization. (A) Size exclusion chromatography (SEC) profile of SUN2 SUN domain with D541N mutation. (B) SDS-PAGE analysis of fractions of SEC. Lanes 1-9 are sequential elution fractions from the beads. (C) SEC profile of SUN2 SUN domain with G608D mutation. (D) SDS-PAGE analyses of G608D mutant. Lanes 3-10are sequential elution fractions from the beads. (E) Circular dichroism profiles of wild-type and mutant SUN2 SUN domain proteins.

Figure S8. Oligomeric states of coiled coil motifs. The mutlicoil (Wolf et al, 1997) program was used to predict the oligomeric sates of the potential coiled coil motifs located within the luminal domain of human SUN2 protein. The dimeric and trimeric probabilities of the two potential coiled coil motifs (labeled as CC1 and CC2 in the primary sequence cartoon) were rendered as blue dash line and red solid line respectively.

<u>Figure S9. A model of the SUN2 protein network.</u> Based on the current knowledge of the SUN domain structure and the two coiled coil coils, the trimeric complex is brought together by the SUN domain and the adjacent coiled coil motif. The inter-trimer interaction is mediated by the distal coiled coil, which is predicted to be dimeric. The network can extend all around the nuclear envelope.
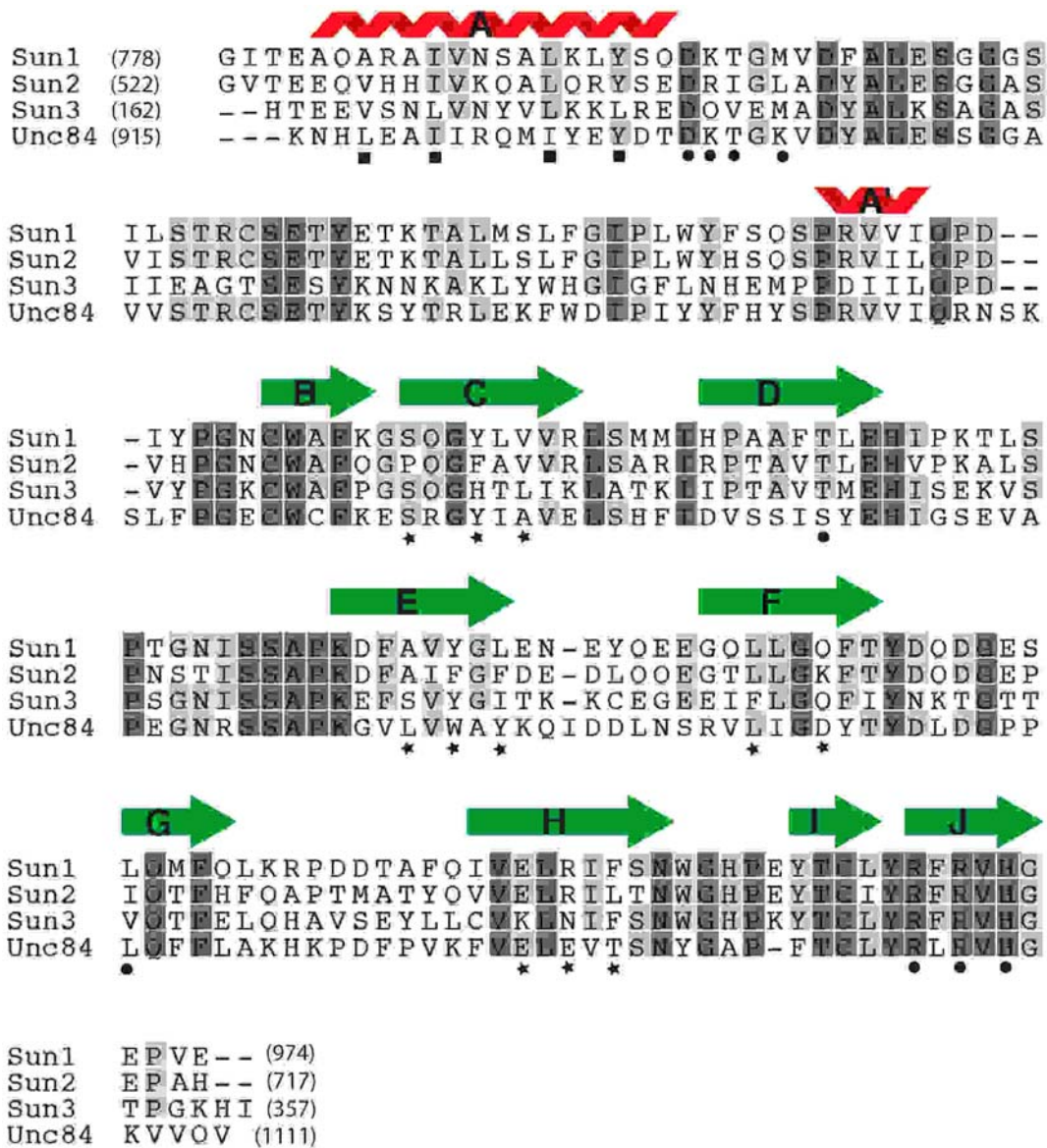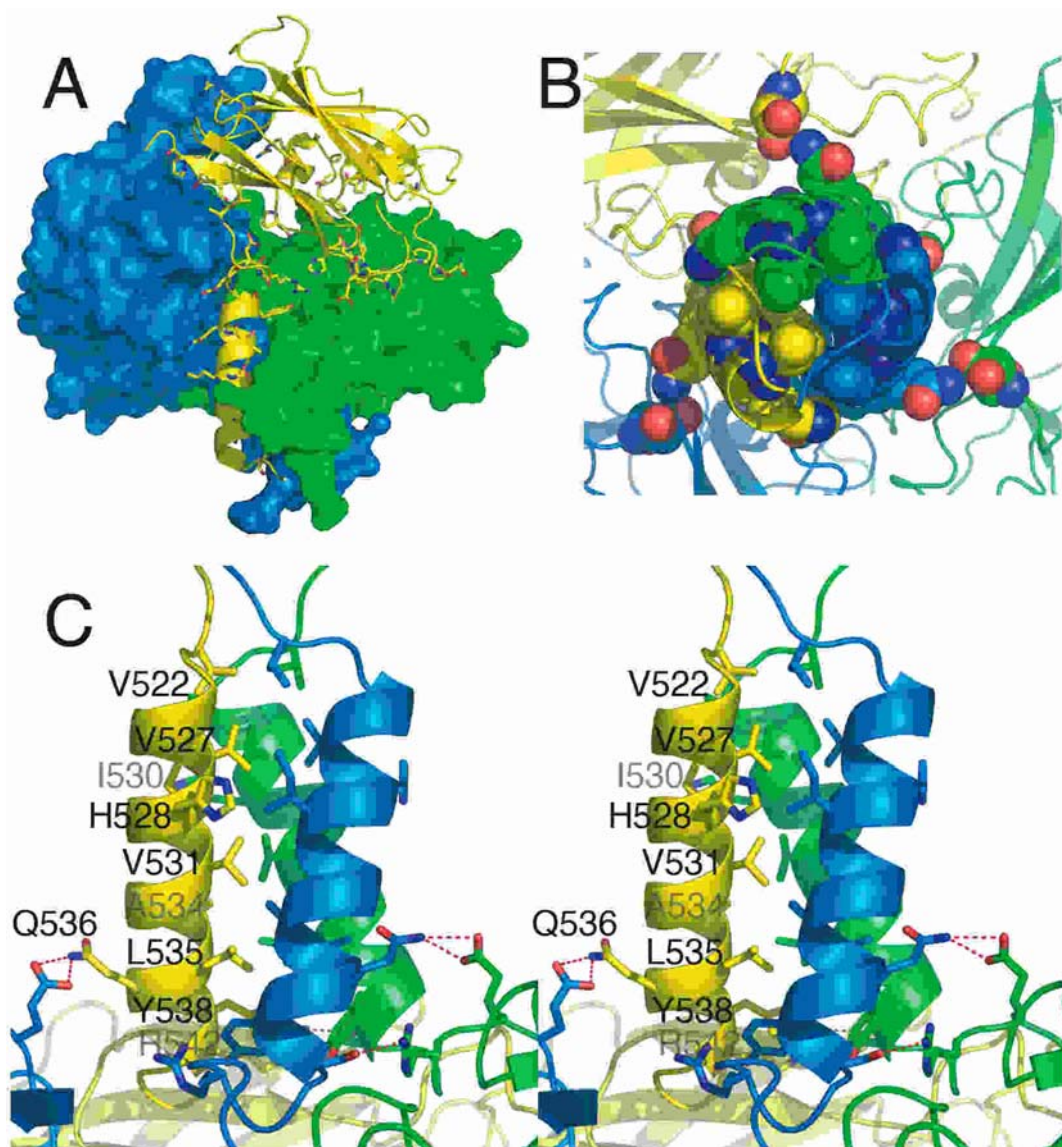
```
                            AAAAAAAAAAAAAAAA  A
Sun1  (778)  GITEAQARAIVNSALKLYSQDKTGMVDFALESGGGS
Sun2  (522)  GVTEEQVHHIVKQALQRYSEDRIGLADYALESGLAS
Sun3  (162)  --HTEEVSNLVNYVLKKLREDOVEMADYALKSALAS
Unc84 (915)  ---KNHLEAIIRQMIYEYDTDKTGKVDYATESSLGA
                   ▪     ▪      ▪       ▪    ● ● ● ●

                                         A'
Sun1   ILSTRCSETVETKTALMSLFGIPLWYFSQSPRVVIQPD--
Sun2   VISTRCSETVETKTALLSLFGIPLWYHSQSPRVILQPD--
Sun3   IIEAGTSESYKNNKAKLYWHGIGFLNHEMPFDIILQPD--
Unc84  VVSTRCSETYKSYTRLEKFWDIPIYYFHYSPRVVIQRNSK

                  ───B──▶      ───C──▶        ───D──▶
Sun1   -IYPGNCWAFKGSQGYLVVRLSMMTHPAAFTLFHIPKTLS
Sun2   -VHPGNCWAFQGPQGFAVVRLSARIRPTAVTLFHVPKALS
Sun3   -VYPGKCWAFPGSOGHTLIKLATKLIPTAVTMEHISEKVS
Unc84  SLFPGECWCFKESRGYIAVELSHFTDVSSISYEHIGSEVA
                ★     ★  ★             ●

                 ───E──▶              ───F──▶
Sun1   PTGNISSAPKDFAVYGLEN-EYQEEGQLLGQFTYDQDGES
Sun2   PNSTISSAPKDFAIFGFDE-DLQQEGTLLGKFTYDQDGEP
Sun3   PSGNISSAPKEFSVYGITK-KCEGEEIFLGQFIYNKTGTT
Unc84  PEGNRSSAPKGVLVWAYKQIDDLNSRVLIGDYTYDLDGPP
                ★  ★ ★              ★   ★

            ──G──▶            ───H──▶      ─I─▶ ──J──▶
Sun1   LGMFQLKRPDDTAFQIVELRIFSNWGHPEYTCLYRFRVHG
Sun2   IGTFHFQAPTMATYQVVELRILTNWGHPEYTCIYRFRVHG
Sun3   VQTFELQHAVSEYLLCVKLNIFSMWGHPKYTCLYRFHVHG
Unc84  LQFFLAKHKPDFPVKFVELEVTSNYGAP-FTCLYRLHVHG
          ●             ★  ★ ★            ● ● ●


Sun1   EPVE-- (974)
Sun2   EPAH-- (717)
Sun3   TPGKHI (357)
Unc84  KVVQV  (1111)
```

Figure S1

Figure S2

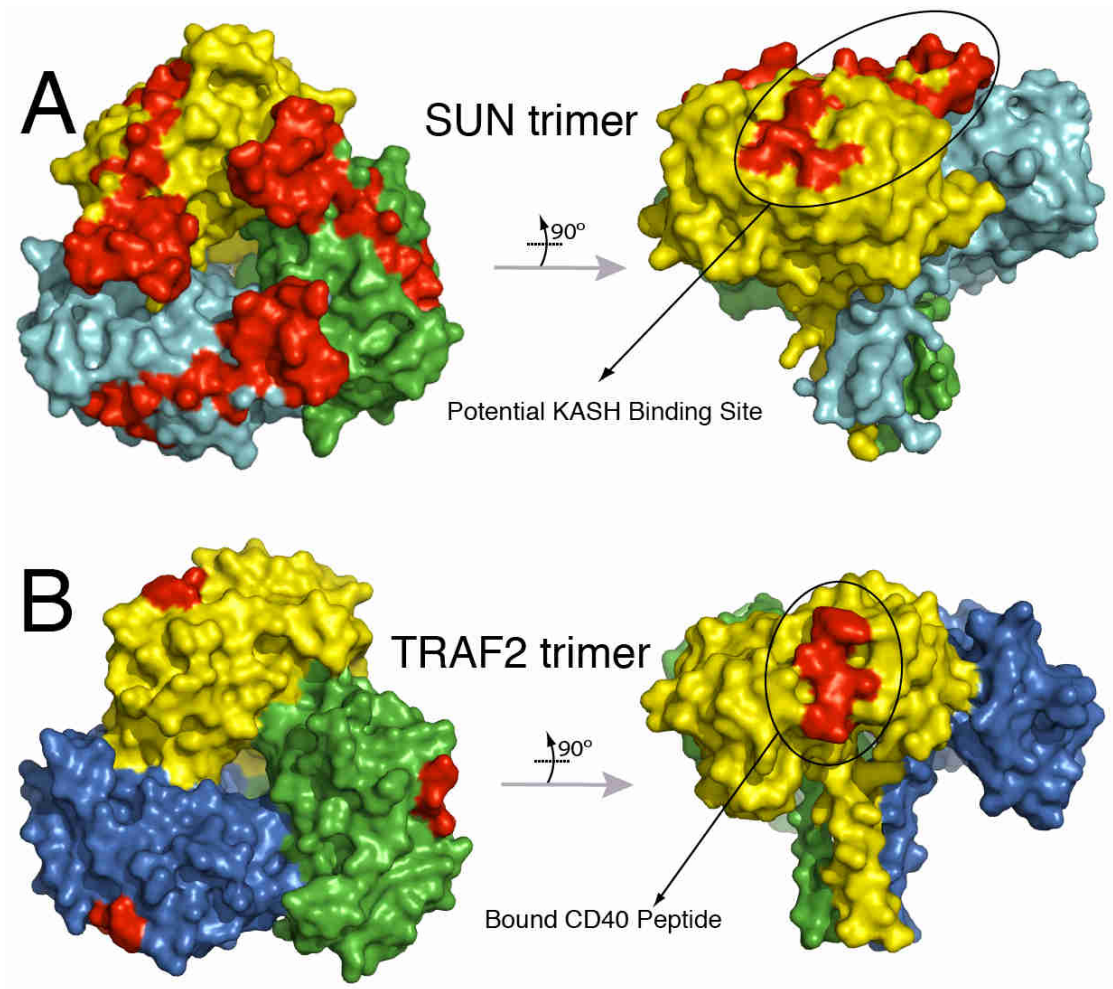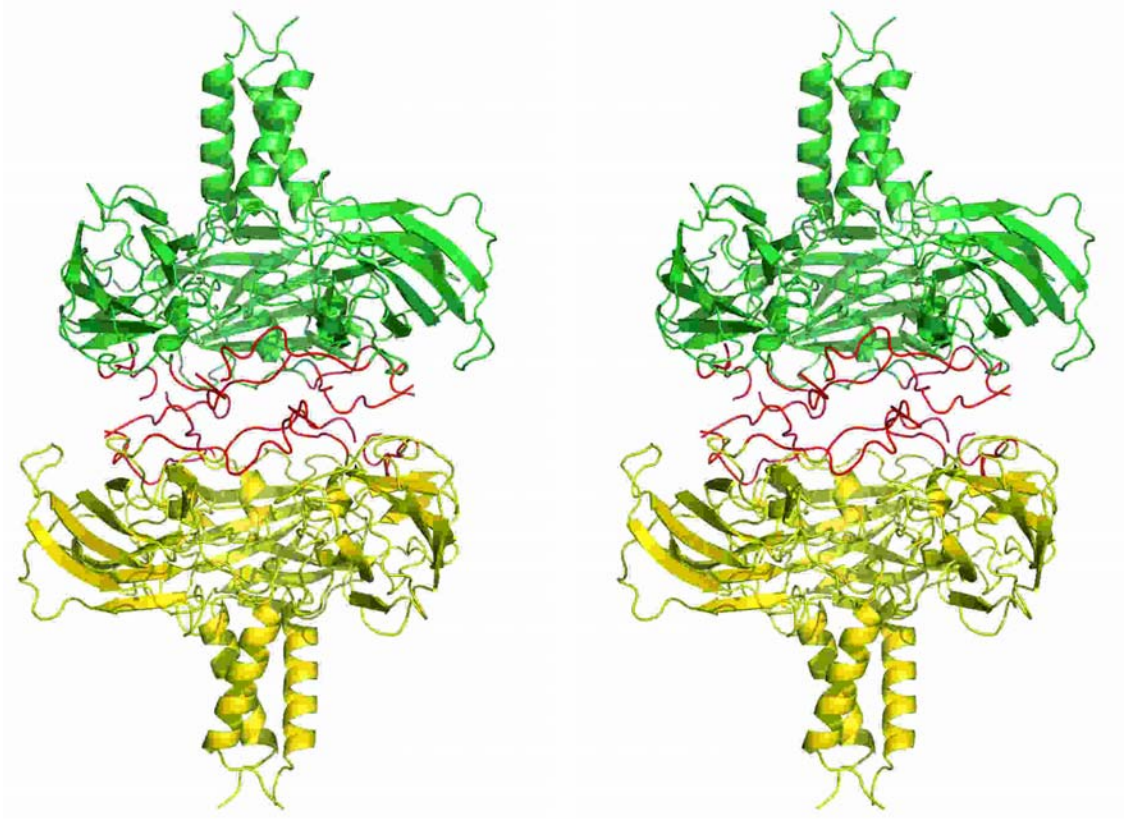A   SUN trimer

90°

Potential KASH Binding Site

B   TRAF2 trimer

90°

Bound CD40 Peptide

Figure S3

Figure S4

Figure S5

Figure S6
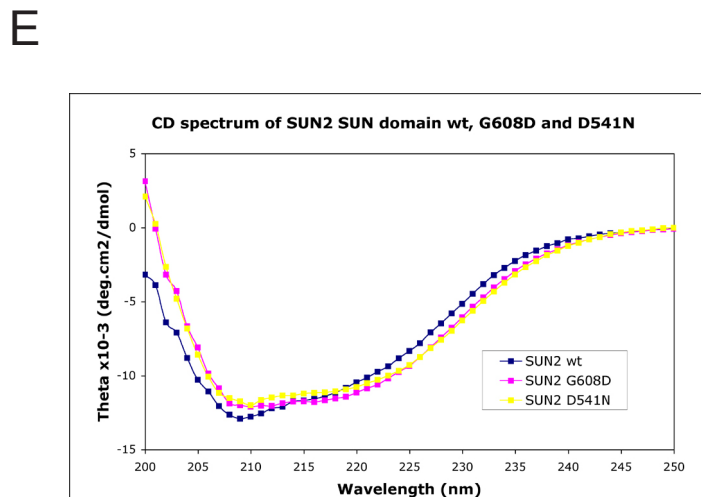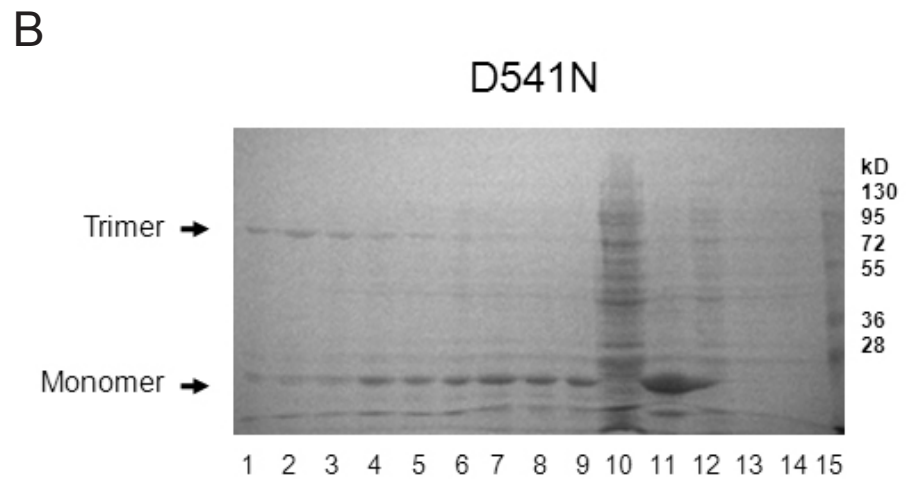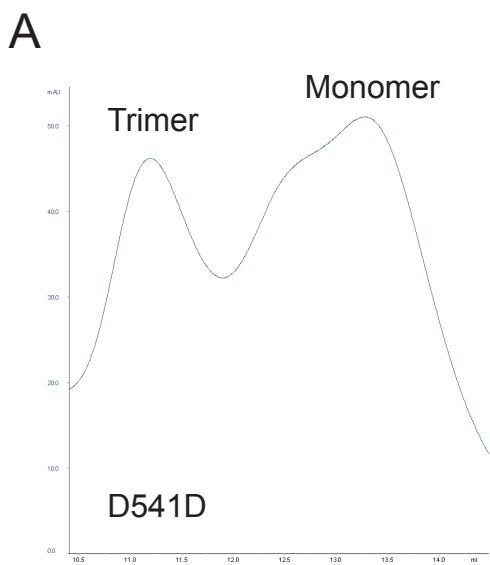
A

Trimer · Monomer

D541D

B

D541N

Trimer →
Monomer →

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

kD
130
95
72
55
36
28

C

Trimer · Monomer

G608D

D

G608D

kD
250
130
95
72
55
36
28

← Trimer
← Monomer

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

E

CD spectrum of SUN2 SUN domain wt, G608D and D541N

SUN2 wt
SUN2 G608D
SUN2 D541N
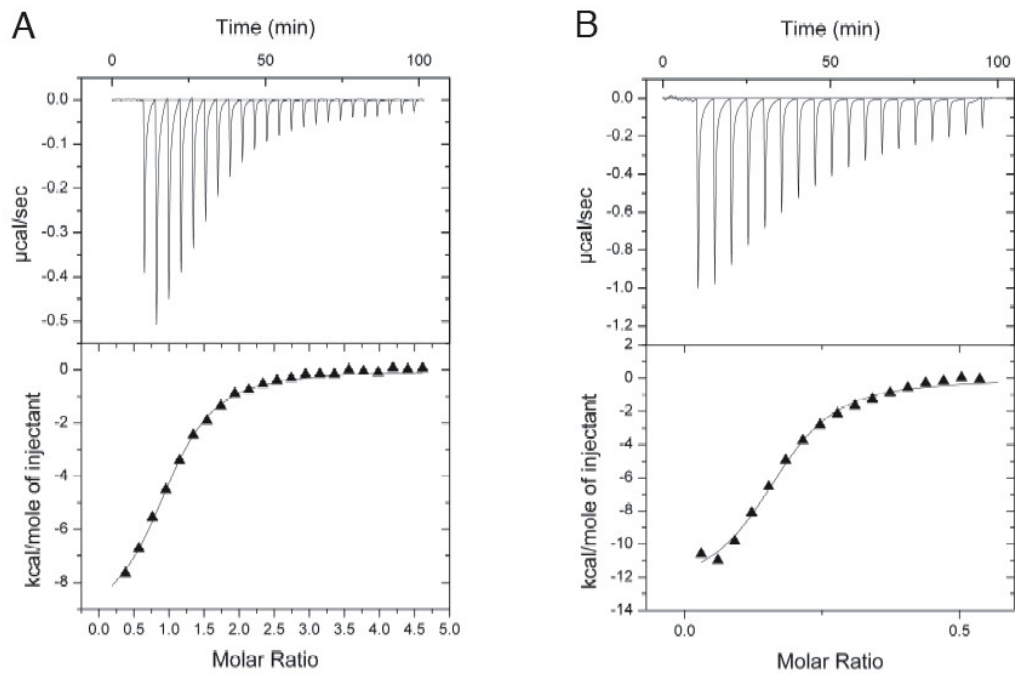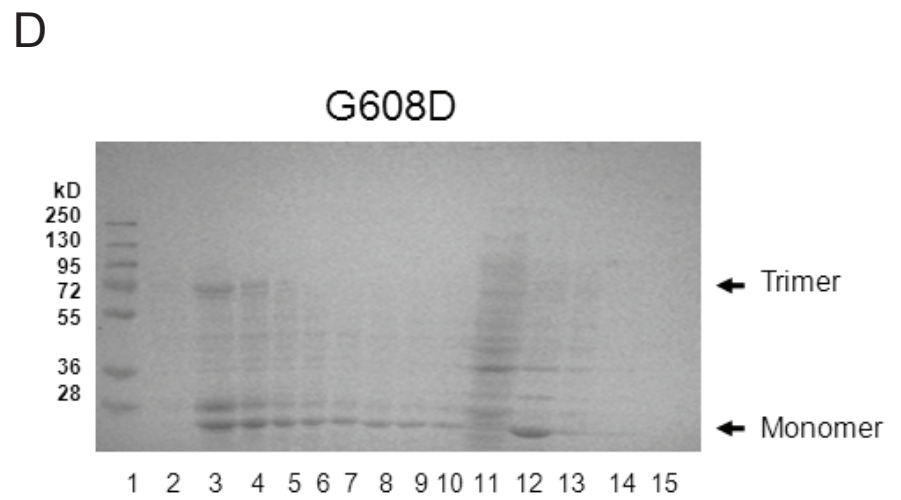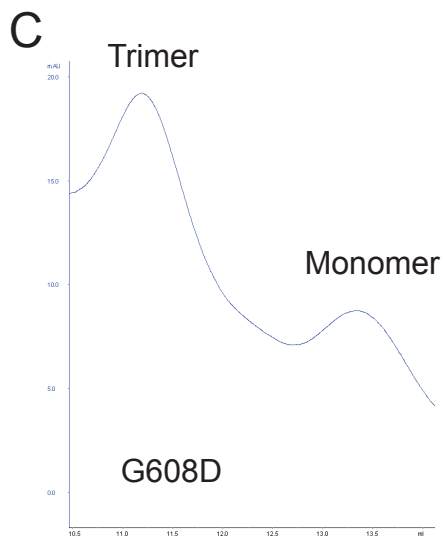
Figure 7S

Figure S8

Figure S9