

Supplemental Methods

SOLUTION STRUCTURE OF THE GUANINE NUCLEOTIDE-BINDING STAS DOMAIN OF SLC26-RELATED SulP PROTEIN Rv1739c FROM *M. Tuberculosis*

Alok K. Sharma^{1,3,4}, Liwen Ye^{2,3,4}, Christina E. Baer⁵, Kumaran Shanmugasudaram^{1,3,4}, Tom Alber⁵, Seth L. Alper^{1,2,3,4*}, Alan C. Rigby^{1,3,4*}

¹Division of Molecular and Vascular Medicine, ²Renal Division, ³Center for Vascular Biology Research, Beth Israel Deaconess Medical Center, and ⁴Department of Medicine, Harvard Medical School, Boston, MA 02215; ⁵Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA

Running head: Structure of GTP-binding *M. tb* Rv1739c STAS domain

NMR Spectroscopy- Unlabeled and uniformly ¹³C/¹⁵N or ¹⁵N labeled STAS-6His *NMR* samples of 0.7-1.0 mM concentration were prepared in a 92% H₂O/8%D₂O solvent system comprised of buffer A (50 mM Na phosphate, 275 mM NaCl, pH 7.2) with 2-3 mM DTT-d₁₀, 0.05% (w/v) NaN₃, and 0.25 mM DSS as an internal standard as described previously (1).

All *NMR* data were acquired in gradient-selected, sensitivity-enhanced mode. Frequency discrimination in the indirect dimensions of all multidimensional spectra was achieved using the States-TPPI mode of quadrature detection (2). The directly and indirectly detected time domain of 2D and 3D data were processed using a 90° phase-shifted squared sinebell or a Gaussian filter weighting function. Data sets were zero-filled once in each dimension prior to Fourier transformation. *NMR* data were processed on an Intel PC workstation running RedHat Linux versions 7.1 and/or RedHat Enterprise 5. Temperature was kept constant throughout all *NMR* data acquisition. Chemical shifts were referenced to the DSS internal standard (3).

Protein backbone and side-chain resonances of ¹H, ¹³C, and ¹⁵N were assigned using 2D spectrum data including ¹H-¹⁵N HSQC, ¹H-¹³C CT-HSQC, [¹H-¹H] TOCSY ($\tau_m = 30$ and 70 ms), and [¹H-¹H] NOESY ($\tau_m = 120$ ms); triple resonance correlation spectrum data including HNCACB, HN(CO)CACB, HNCA, HN(CO)CA, HNCO, HN(CA)CO, HBHA(CO)NH, H(CCCO)NH-TOCSY, and C(CCO)NH-TOCSY; data from 3D spectra of ¹⁵N-edited NOESY-HSQC ($\tau_m = 120$ ms) and of ¹³C-edited NOESY-HSQC ($\tau_m = 120$ ms); and data from double-resonance spectra of HCCH-TOCSY and ¹⁵N-TOCSY-HSQC ($\tau_m = 50$ ms) (4-6). Scalar coupling constants were determined using 3D HNHA spectrum (4). Ring protons of the aromatic residues were assigned using a combination of ¹³C-edited NOESY-HSQC spectrum, and ¹H-¹H 2D TOCSY and NOESY spectra acquired on a sample of Rv1739c STAS in 99% D₂O. Steady state {¹H}-¹⁵N NOE spectra were recorded at 600 MHz on ¹⁵N-labeled STAS using Bruker standard pulse program *invinoef3gpsi*. Data were acquired in an interleaved manner using 5 s recycle delay, with and without 3s of ¹H saturation. Data were separated into NOE (with ¹H saturation) and NONOE spectra (without ¹H saturation) using the *edau* program in XWin-NMR. Separated NOE and NONOE experiments were processed using NMRPipe/NMRDraw, and heteronuclear NOEs were determined as the ratio of crosspeak heights in NOE and NONOE experiments. Duplicate spectra were acquired to estimate the uncertainty in NOE values.

Structure Determination of Rv1739c STAS domain- Integrated NOE crosspeak intensities in 2D and 3D NOESY spectra (as described in main text) were converted into distance restraints classified as strong (respective upper bound 3.5 Å), medium (4.0 Å), weak (5.0 Å), and very weak (6.0 Å). Backbone dihedral angles (ϕ , ψ) were determined with TALOS, the CaAnglerestraints routine in CYANA-2.1, and from scalar coupling constant as described earlier (7). Hydrogen bond restraints were generated for residues that were involved in slow amide-proton exchange (on the *NMR* timescale) in the regions of secondary structure, as determined from medium-range NOEs, in 2D ¹H-¹⁵N HSQC-based H-D exchange

experiments performed with ^{15}N labeled STAS lyophilized in D_2O . Upper bounds for hydrogen bond distances were set at 2.2 Å for HN-O and 3.2 Å for N-O, with a constant lower bound of 1.80 Å. Three-dimensional structures were calculated using program CYANA-2.1 based on torsion-angle dynamics approach. Four hundred random conformers were subjected to 10,000 annealing steps to obtain first round structures, and then hydrogen bond restraints were included in the restraints input data. After successive structure calculation cycles, final structure calculation used a distance-corrected restraint file with 40,000 annealing steps. Energy minimization and water refinement was performed with ensemble of 30 best structures. Molecular dynamic simulations in explicit water for water refinement were carried out with program CNS. Out of 30 refined structures 25 structures of highest refinement were selected to represent ensemble of solution structure of Rv1739c STAS.

Heteronuclear NMR investigation of Rv1739c STAS-nucleotide interaction- Chemical shift perturbation (CSP) studies were carried out by acquiring a series of 2D ^1H - ^{15}N HSQC spectra on ^{15}N labeled Rv1739c STAS (0.3-0.7 mM) following stepwise addition of increasing concentrations of GTP or GDP, from stocks in Buffer A in the presence of 2-3 mM DTT- d_{10} , to a saturating concentration of 20 mM nucleotide. CSPs were calculated as a weighted chemical shift ($\Delta\delta_{\text{weighted}}$) for each STAS residue. Perturbations with $\Delta\delta_{\text{weighted}} \geq 0.05$ were considered significant (8). The majority of HSQC crosspeaks exhibited increased line-broadening in the presence of nucleotide. Addition of GTP or GDP also elicited new resonances with nearly identical (^1H , ^{15}N) HSQC coordinates. These new resonances may be amide proton signals of residues undetected in unliganded STAS ^1H - ^{15}N HSQC due to fast exchange of amide proton with bulk solvent (1), now detectable due to nucleotide-induced motional rigidity. Alternately, these resonances may reflect slow conformational exchange of select residues, or to nucleotide-induced partial dimerization of STAS (9). GTP addition increased rigidity/order of some residues (including A10, D22, S57, and L114) but decreased rigidity/order of others (Y21 and L102). GDP addition increased structural rigidity/order of an expanded surface including residues D5, Y7, A10, R12, S57, and L114, but slightly decreased rigidity/order for residues V59, R79, and K89.

In silico docking of STAS-nucleotide complexes- *In silico* docking calculations were carried out using AutoDock Vina (10), which uses an Iterated Local Search Global Optimizer (10) for local conformation searching and optimization. The algorithm uses the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method (11,12) for local optimization. Complex structures were visualized by Autogrid4. Docking grids of 50 Å x 50 Å x 50 Å were used for nucleotide in complex with Rv1739c STAS, with 0.50 Å grid spacing of x, y, z coordinates (13.3663, 2.1575, -7.087).

Photoaffinity labeling of Rv1739c STAS domain- 5 μg Rv1739c STAS was incubated 15 min on ice in a 50 μL final volume containing 30 μM 8- N_3 -GTP(γ)-biotin-LC-PEO-Amine (Affinity Photoprobes, Lexington, KY) in 50mM Na phosphate, pH 7.2, in the presence or absence of 2 mM GTP. The mixture was irradiated 2 min at 254 nm by Mineralight UV lamp at a distance of 2 cm, followed by reaction termination with 10 μL 5x SDS load buffer containing 5% β -mercaptoethanol. 2/3 of the sample was subjected to SDS-PAGE and stained with Coomassie Blue. The remaining 1/3 of the electrophoresed sample was transferred to nitrocellulose (Optitran, Whatman/Sigma) and developed with streptavidin-coupled horseradish peroxidase (GE Healthcare/Amersham). Human RhoA-GST fusion protein (gift of Steen Hansen, Harvard Medical School) served as positive control for photoaffinity labeling.

Assay of nucleotidase activity- Rv1739c STAS (0, 2 or 11.5 μg) or human RhoA-GST (5 μg) was mixed with 33 or 64 nM [γ - ^{32}P]-GTP or 640 nM [γ - ^{32}P]-ATP (PerkinElmer, Waltham, MA) in a final volume of 20 μL containing (in mM) 100 NaCl, 2 MgCl_2 , and 50 HEPES pH 7.2. After 1-2 hrs incubation, 5 μL of the reaction mix was stopped with 5 μL 2.5% SDS, and 2 μL of the stopped reaction mix was spotted in triplicate on polyethyleneimine (PEI)-Cellulose F plates (Analtech, Newark, DE) for

thin layer chromatography (TLC). TLC plates were chromatographed in 1 M formic acid and 0.5 M LiCl for 1-2 hrs to separate hydrolyzed inorganic [³²P]-phosphate from unhydrolyzed [γ -³²P]-GTP, then dried and autoradiographed (HyBlot CL, Denville Scientific, Metuchen, NJ). The autoradiographs guided excision of spots containing inorganic [³²P]-phosphate. Excised spots on their plastic backing were immersed in scintillation fluid (Ecoscint A, National Diagnostics, Atlanta, GA) and counted (Packard Tri-Carb 2200CA) at 97% efficiency compared to isotope in solution.

Protein phosphorylation assays- Recombinant kinase domains of *M. tuberculosis* protein serine-threonine kinases PknB and PknL were PCR-amplified and cloned into expression vector pLIC-HMK. Kinase domains of PknA, PknD, PknE, PknF, PknH, PknJ, PknL, and full-length genes encoding PknI and the anti-sigma factors Rv0941, Rv1364, Rv3221, and Rv3287 were inserted into Gateway vector pHMGWA. Full-length PknG was cloned into pET-28b. Proteins were overexpressed in *E. coli* and purified as described (13-15).

Rv1739c STAS expression in M. smegmatis- N-terminal FLAG-Rv1739c STAS cDNA in pMS3 was expressed constitutively from a GroEL/HSP60 promoter in *M. smegmatis* MC2155. Cultures pelleted 10 min at 3,000 x g were resuspended in lysis buffer (1% Triton X-100 containing (in mM) 20 EDTA, 50 Na HEPES, pH 7.5, 50 NaF, 1 AEBSF) and lysed by bead beating. Centrifugation-clarified soluble fractions were analyzed by SDS-PAGE for Coomassie Blue staining and for transfer to PVDF membrane for dual immunoblot with polyclonal anti-pThr (Invitrogen) and monoclonal anti-FLAG M2 (Sigma).

References for Supplemental Methods

1. Sharma, A. K., Ye, L., Zolotarev, A. S., Alper, S. L., and Rigby, A. C. (2009) *Biomol NMR Assign* **3**(1), 99-102
2. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J Magn Reson* **85**(2), 393-399
3. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wuthrich, K. (1998) *J Mol Biol* **280**(5), 933-952
4. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press Inc., San Diego, CA.
5. Bax, A., and Grzesiek, S. (1993) *Acc. Chem. Res.* **26**, 131-138
6. Muhandiram, D. R., Farrow, N. A., Xu, G. Y., Smallcombe, S., and H., a. K., L. E. (1993) *J. Magn. Reson., Ser. B* **102**, 317-321
7. Sharma, A. K., Sharma, S. K., Surolia, A., Surolia, N., and Sarma, S. P. (2006) *Biochem* **45**(22), 6904-6916
8. Proteomics Sample Preparation, Wiley-VCH, 2008, ISBN: 978-3-527-31796-7; edited by Jorg von Hagen
9. Gasper, R., Meyer, S., Gotthardt, K., Sirajuddin, M., Wittinghofer, A. (2009) *Nat Rev Mol Cell Biol* **10**(6), 423-429
10. Trott, O., and Olson, A. J. (2010) *J Comput Chem* **31**(2), 455-461
11. Baxter, J. (1981) *J Oper Res Soc* **32**(9), 815
12. Blum, C., Maria, Roli, A., and Sampels, M. (2008) *Hybrid Metaheuristics: An Emerging Approach to Optimization*, Springer
13. Greenstein, A. E., Grundner, C., Echols, N., Gay, L. M., Lombana, T. N., Miecskowski, C. A., Pullen, K. E., Sung, P. Y., and Alber, T. (2005) *J Mol Microbio Biotechnol* **9**(3-4), 167-181
14. Greenstein, A. E., MacGurn, J. A., Baer, C. E., Falick, A. M., Cox, J. S., and Alber, T. (2007) *PLoS pathogens* **3**(4), e49
15. Alber, T. (2009) *Curr Opin Struct Biol* **19**(6), 650-657

Supplemental Figure Legends

Figure S1. Purification of Rv1739c STAS domain from *M. tuberculosis*. Fast protein liquid chromatography (FPLC) profile of purified STAS domain. The chromatogram shows a single elution peak of purified protein recorded at 214 nm (blue), 280 nm (green), and 260 nm (red). Inset: SDS-PAGE (16% gel) of purified STAS domain; lane 1: molecular weight marker, lanes 2 and 3: STAS domain in presence (+) or absence (-) of DTT. STAS domain migration was also unaffected by DTT when tested in separate gels (not shown).

Figure S2. Sequential Backbone assignment of Rv1739c STAS domain. *A & B.* Overlay of HNCACB (red and black contours for $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$, respectively) and HN(CO)CACB (green and blue contours for $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$, respectively) for amino acids A28 to D44 comprising first α -helix (in *A*), and amino acids R78 to V88 comprising third β -strand (in *B*). *C.* Assigned 2D ^1H - ^{15}N HSQC spectrum showing proton-nitrogen correlations from backbone and side chain pairs. The crosspeaks are labeled with their corresponding amino acid residue numbers. Proton-nitrogen pairs from side chains of Asn and Gln are connected by horizontal lines. Crosspeaks marked with * represent His residues from 6His-tag present at C-terminus of STAS amino acid sequence. Reproduced from Sharma *et al.* 2009.

Figure S3. Structural Properties of Rv1739c STAS domain. *A.* Electrostatic surface potential of Rv1739c STAS domain, showing acidic (red) and basic (blue) residues. *B.* Backbone $\{^1\text{H}\} - ^{15}\text{N}$ heteronuclear NOEs recorded at 298K, plotted as function of STAS domain amino acid sequence. Processed data were obtained for 108 resonances out of 119 non-proline resonances. Arrayed above is the secondary structure deduced from the calculated tertiary structure.

Figure S4. Intrinsic fluorescence quench of Rv1739c STAS by nucleotides. *A.* Three-dimensional representation of STAS domain hydrophobic residues shows Trp indole ring (red) oriented away from solvent into the non-polar environment of the neighbouring hydrophobic cluster, predicted to reduce the emission λ_{max} . *B, C, E, & G.* Representative experiments showing quench of intrinsic STAS domain Trp-Tyr fluorescence (12-20 μM) by increasing concentrations of GTP (*B*), GTP- γ -S (*C*), GDP- β -S (*E*), and ATP (*G*). λ_{ex} was 280 nm; λ_{em} was scanned between 300 and 400 nm. *D, F, & H.* Nucleotide concentration dependence for quench of peak Rv1739c STAS intrinsic fluorescence by GTP- γ -S (*D*; $K_{1/2} = 326 \pm 131 \mu\text{M}$; $R^2 = 0.98$), GDP- β -S (*F*; $K_{1/2} = 152 \pm 53 \mu\text{M}$; $R^2 = 0.98$), and ATP (*H*; $K_{1/2} = 756 \pm 232 \mu\text{M}$; $R^2 = 0.94$). Values in panels *D, F,* and *H* are means \pm s.e.m., $n=3$.

Figure S5. GTPase activity associated with Rv1739c STAS domain. *A.* PEI-cellulose thin layer electrophoretic separation of γ - ^{32}P -GTP substrate (lower spots at origin) from $^{32}\text{P}_i$ product (upper spots) in reactions (2 hrs at 20°C, 64 nM total GTP) containing human RhoA-GST (5 μg) or Rv1739c STAS (11.5 μg). Rv1739c STAS-associated GTPase activity was $353 \pm 39 \text{ pmol hr}^{-1} \mu\text{mol}^{-1}$ protein (mean \pm s.e.m., $n=17$ triplicate experiments with carrier-free (33 nM) γ - ^{32}P -GTP). *B.* Rv1739c STAS domain-mediated GTP hydrolysis is temperature-dependent, protease-sensitive, and inhibited by SDS. The triplicate 60 min reactions contained 6.6 μM STAS and 33 nM GTP. Values are means \pm s.d. of triplicate determinations.

Figure S6. Nucleotide-induced chemical shift perturbation plotted as weighted-average chemical shifts ($\Delta\delta_{\text{weighted}}$) for residues of Rv1739c STAS domain obtained from 2D HSQC overlay. *A & B.* Zoomed regions of residues significantly perturbed by GDP as in Figure 6B. *C.* GTP-induced changes measured in Figure 6A. *D.* GDP-induced changes measured in Figure 6B. Red bars in Figures C and D indicate residues perturbed ≥ 0.05 ppm (dashed horizontal line).

Figure S7. *In silico* docking of guanine nucleotides onto Rv1739c STAS domain structure. The best 9 lowest-energy docked poses for GTP (left) and GDP (right) bound to Rv1739c STAS, as determined using AutoDock Vina. GTP is predicted equally bound to the N-terminal and C-terminal regions of the STAS domain, whereas GDP is predicted preferentially bound to the C-terminal site. Figure was generated from PDB structure 2KLN using Benchware 3D (TRIPOS).

Figure S8. Rv1739c STAS domain is not a substrate for *in vitro* phosphorylation by recombinant *M. tuberculosis* protein kinases or anti- σ factors. All kinase reactions were conducted at room temperature for 30 min with 1 μ Ci γ - 32 P-ATP in the presence of 0.5 mM MnCl₂. *A.* 32 P-autoradiograph showing lack of phosphorylation of Rv1739c STAS domain by the indicated recombinant kinases in the presence of 50 μ M ATP. *B.* 32 P-autoradiograph showing phosphorylation of positive control kinase substrate GarA by the same recombinant kinases in the presence of 50 μ M ATP. *C.* 32 P-autoradiograph showing lack of phosphorylation of Rv1739c STAS domain by the same recombinant kinases in the combined presence of 50 μ M ATP and 1 mM GTP. *D.* Lack of phosphorylation of Rv1739c STAS domain by the indicated recombinant *M. tuberculosis* anti- σ factors in the presence of either 50 μ M ATP (left 4 lanes of each panel) or in the combined presence of 50 μ M ATP and 1 mM GTP (right 4 lanes of each panel). Left panel, 32 P-autoradiograph; right panel, same gel stained with Coomassie Blue. The M_r of the radioactive band in the Rv1364 kinase domain lanes is too high to be Rv1739c STAS.

Table S1. Selected reported $K_{1/2}$ values for low-affinity GTP binding proteins.

GTP-binding protein	Species	$K_{1/2}$ (GTP)	Function	Ref.
eRF3	human	200 μ M	ribosome-associated G protein	(a)
eIF5/IF2	<i>E. coli</i>	70 μ M	ribosome-associated G protein	(a)
EF-G	<i>E. coli</i>	40 μ M	ribosome-associated G protein	(a)
FeoB	<i>E. coli</i>	130 μ M	iron sensor	(b)
UsfX	<i>M. tb.</i>	200 μ M ^a	anti- σ factor antagonist	(c)
CodY	<i>B. subtilis</i>	1100 μ M	transcription factor	(d)

^a $K_{1/2}$ value for ATP = 1400 μ M (b).

- a. Haurlyiuk, V., Hansson, S., and Ehrenberg, M. (2008) *Biophysical journal* **95**(4), 1704-1715
- b. Eng, E. T., Jalilian, A. R., Spasov, K. A., and Unger, V. M. (2008) *J Mol Biol* **375**(4), 1086-1097
- c. Malik, S. S., Luthra, A., Srivastava, S. K., and Ramachandran, R. (2008) *Biochem Biophys Res Commun* **375**(3), 465-470
- d. Handke, L. D., Shivers, R. P., and Sonenshein, A. L. (2008) *J Bacteriol* **190**(3), 798-806

Figure S1.

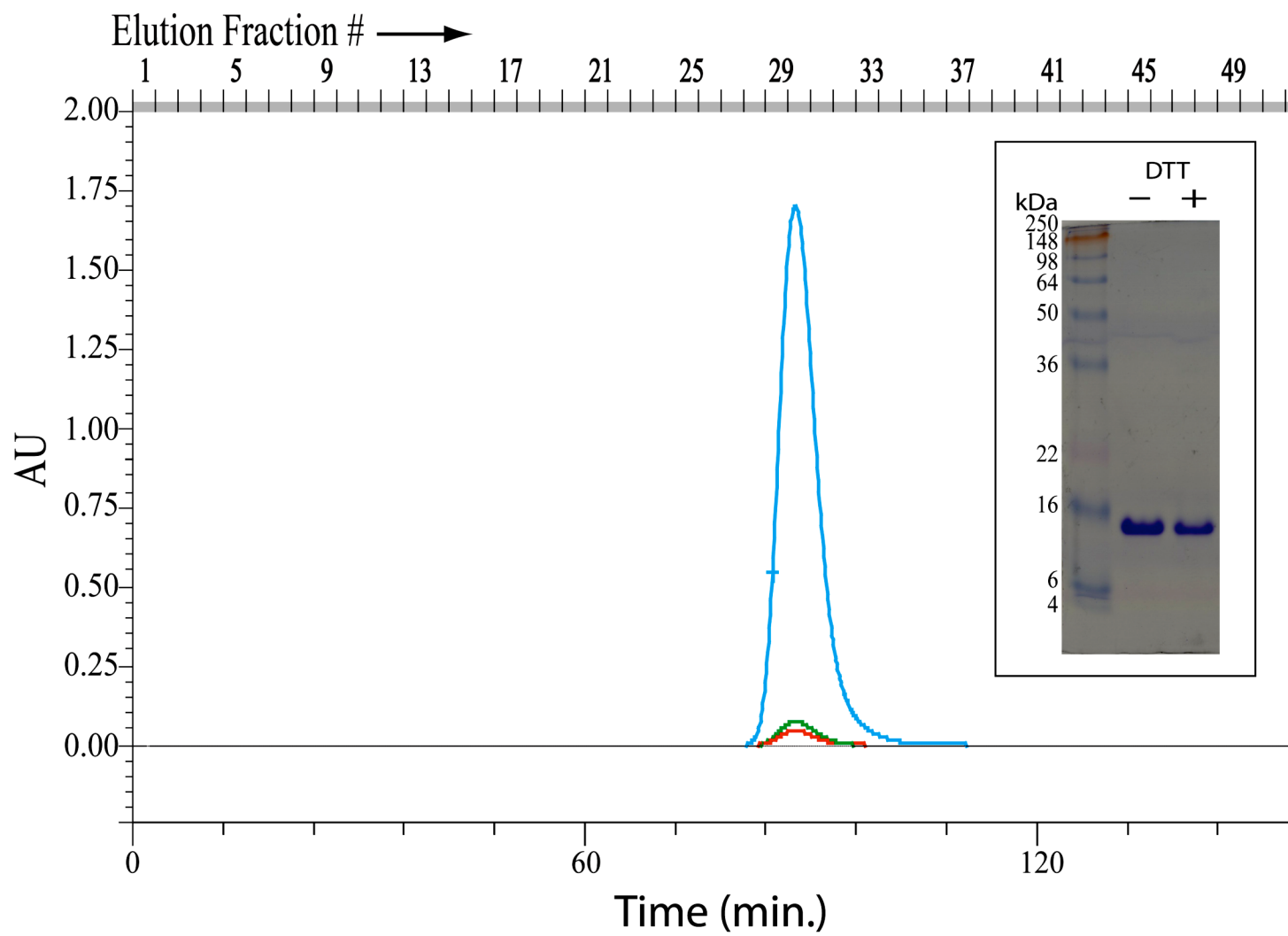
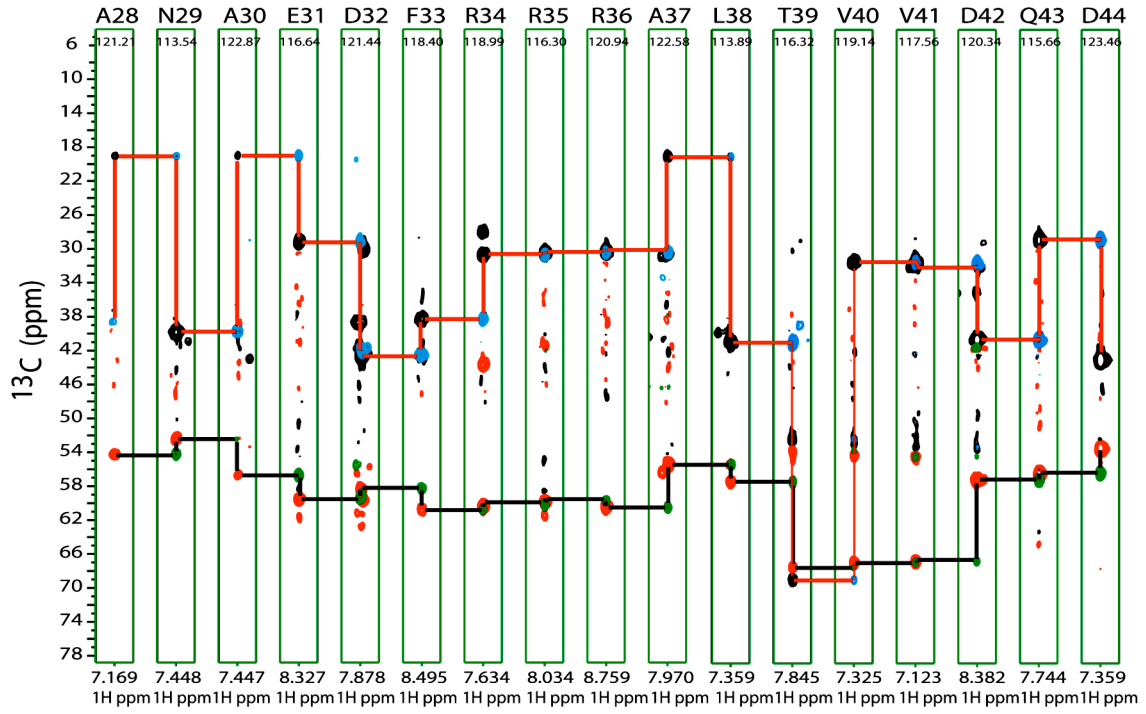


Figure S2.

A



B

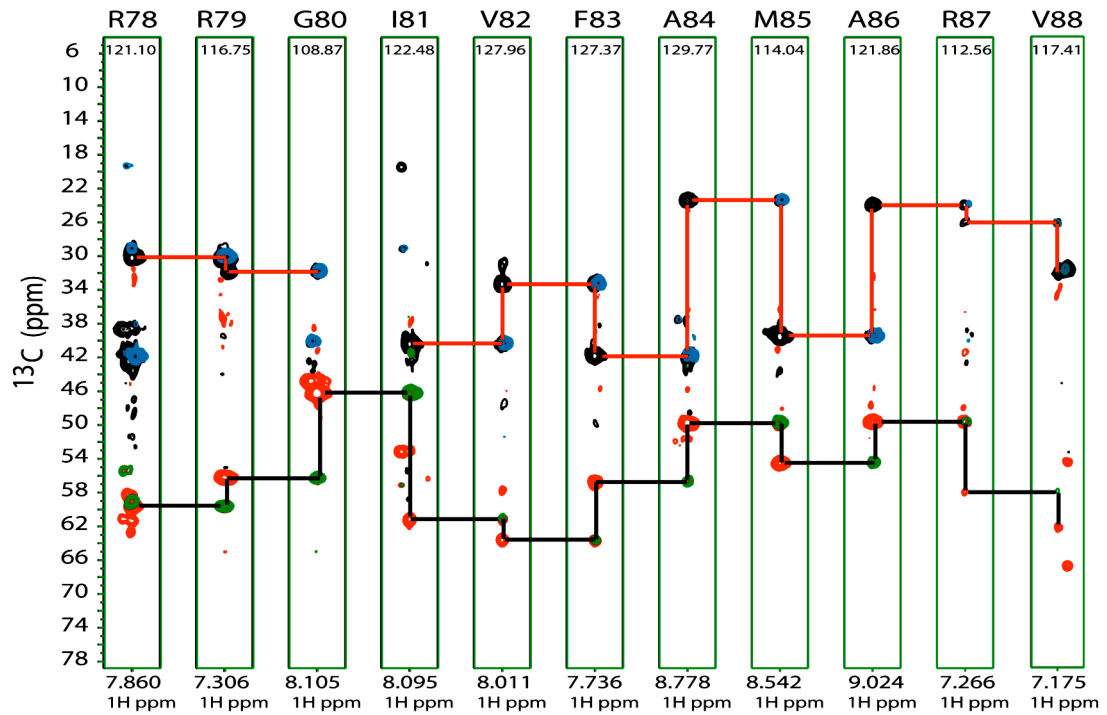
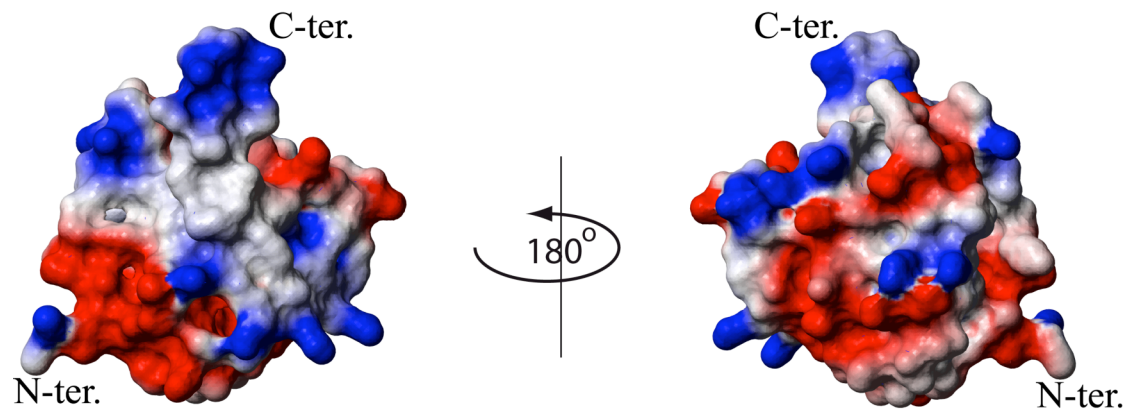


Figure S2 contd...

Figure S3.

A



B

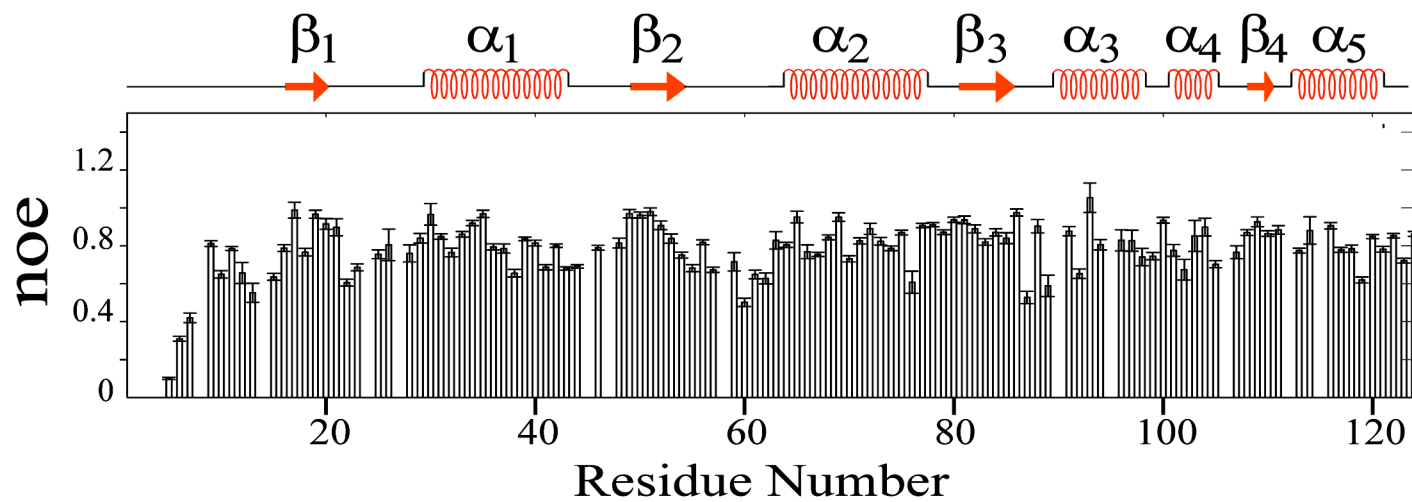
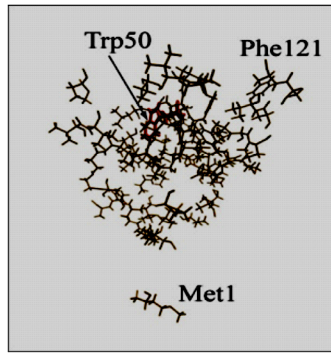
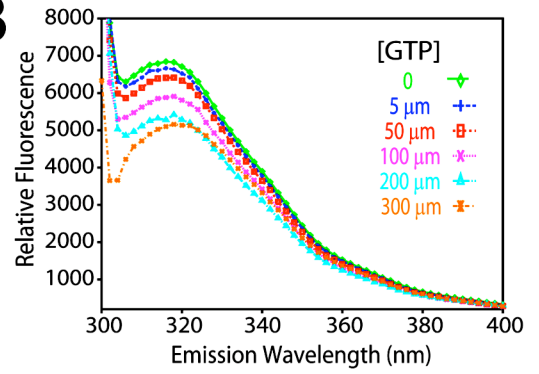


Figure S4.

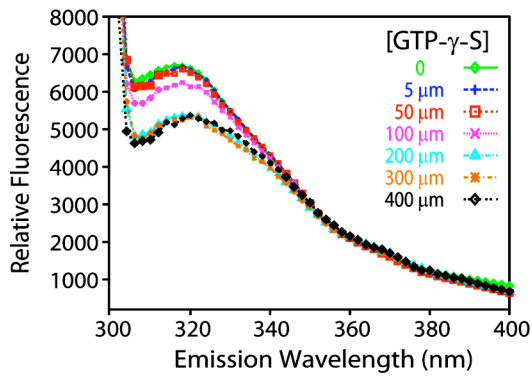
A



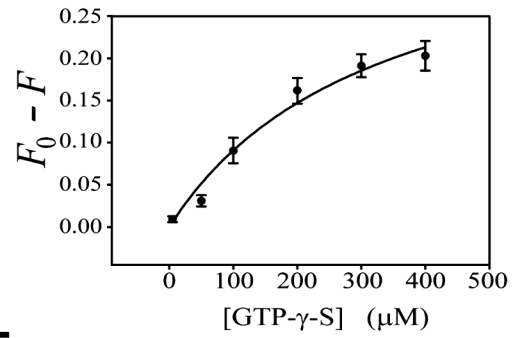
B



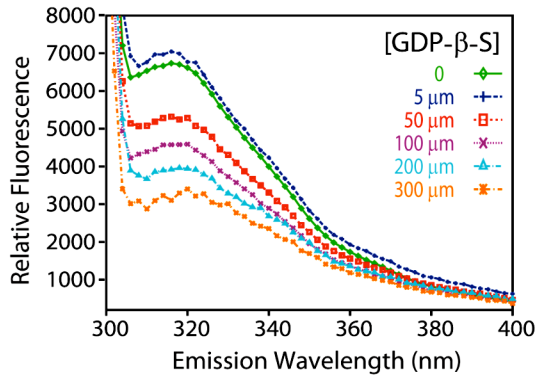
C



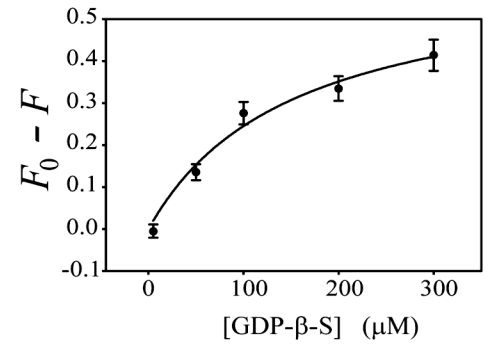
D



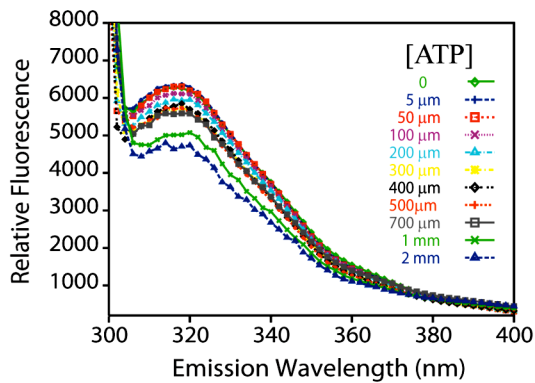
E



F



G



H

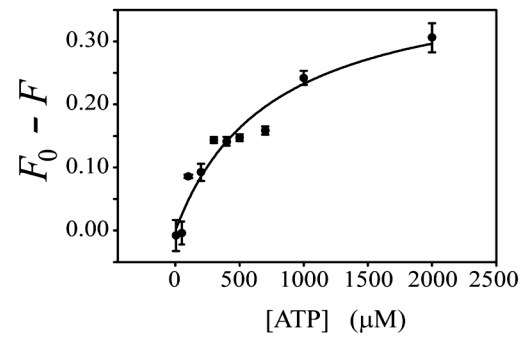


Figure S5.

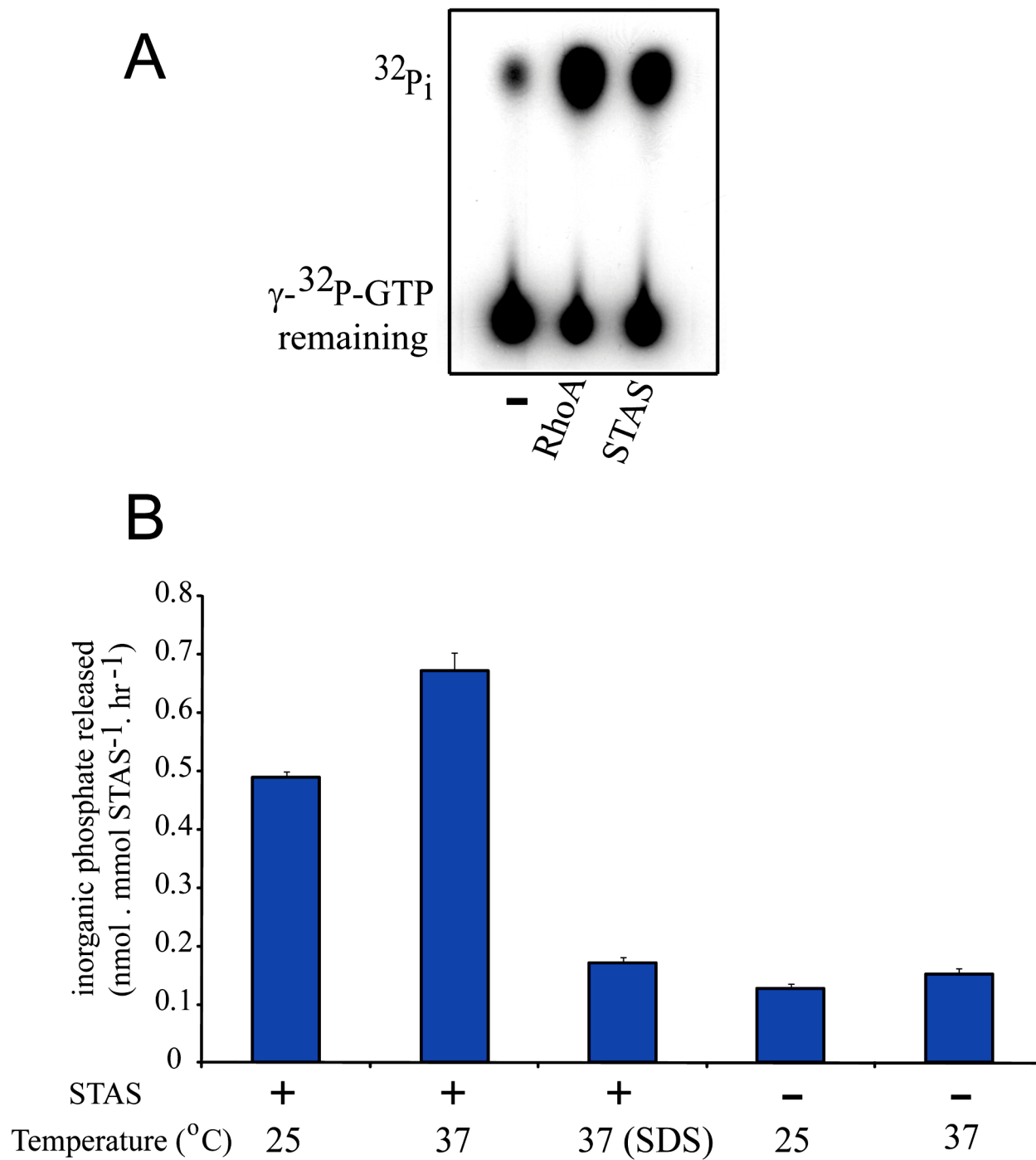


Figure S6.

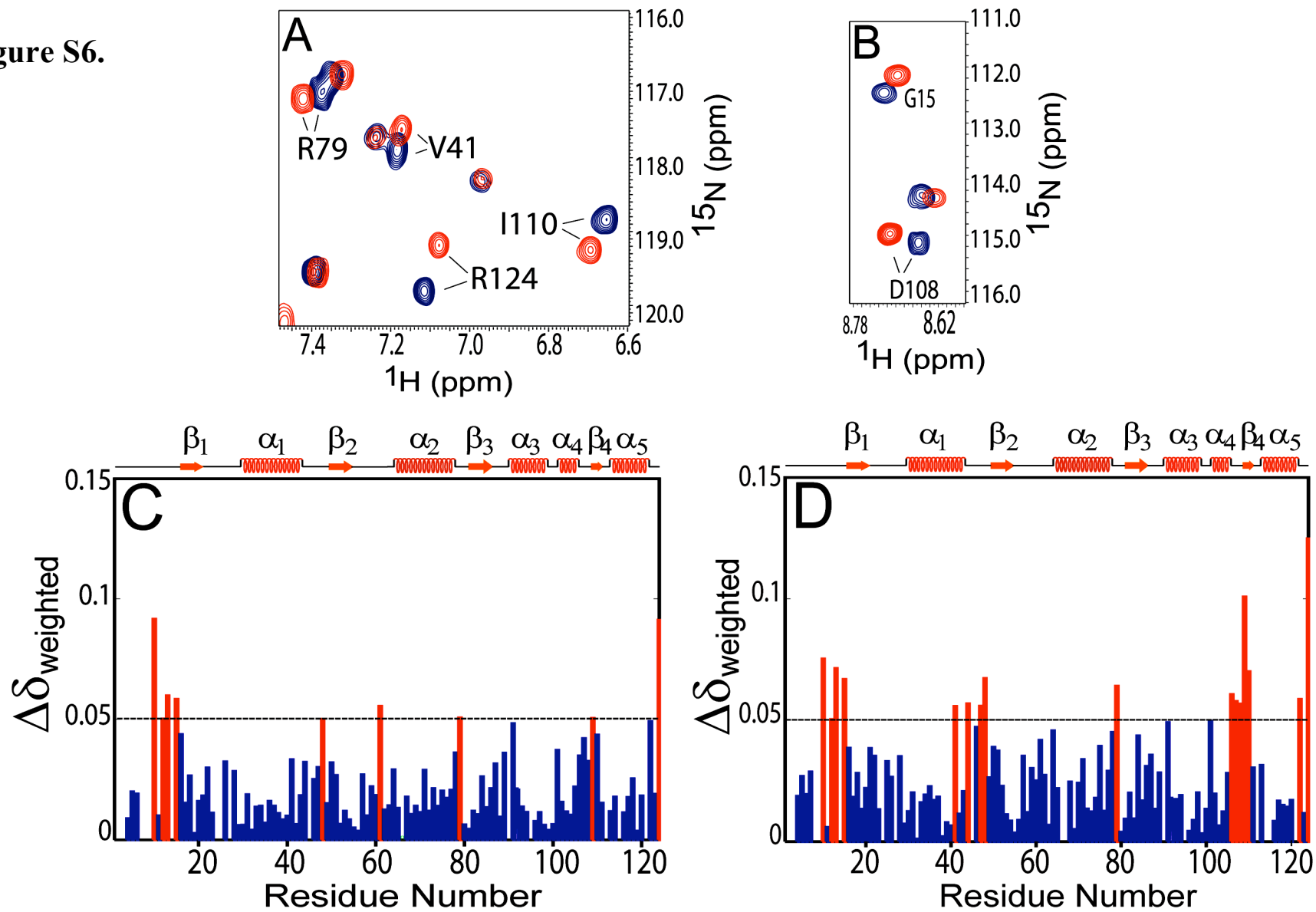
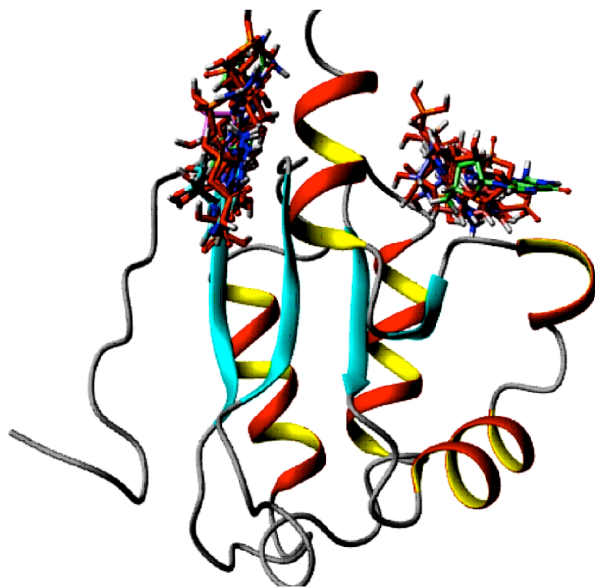


Figure S7.

GTP



GDP

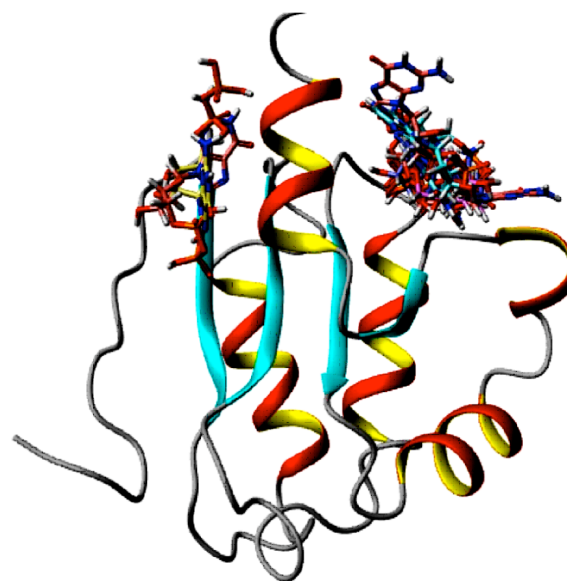


Figure S8.

