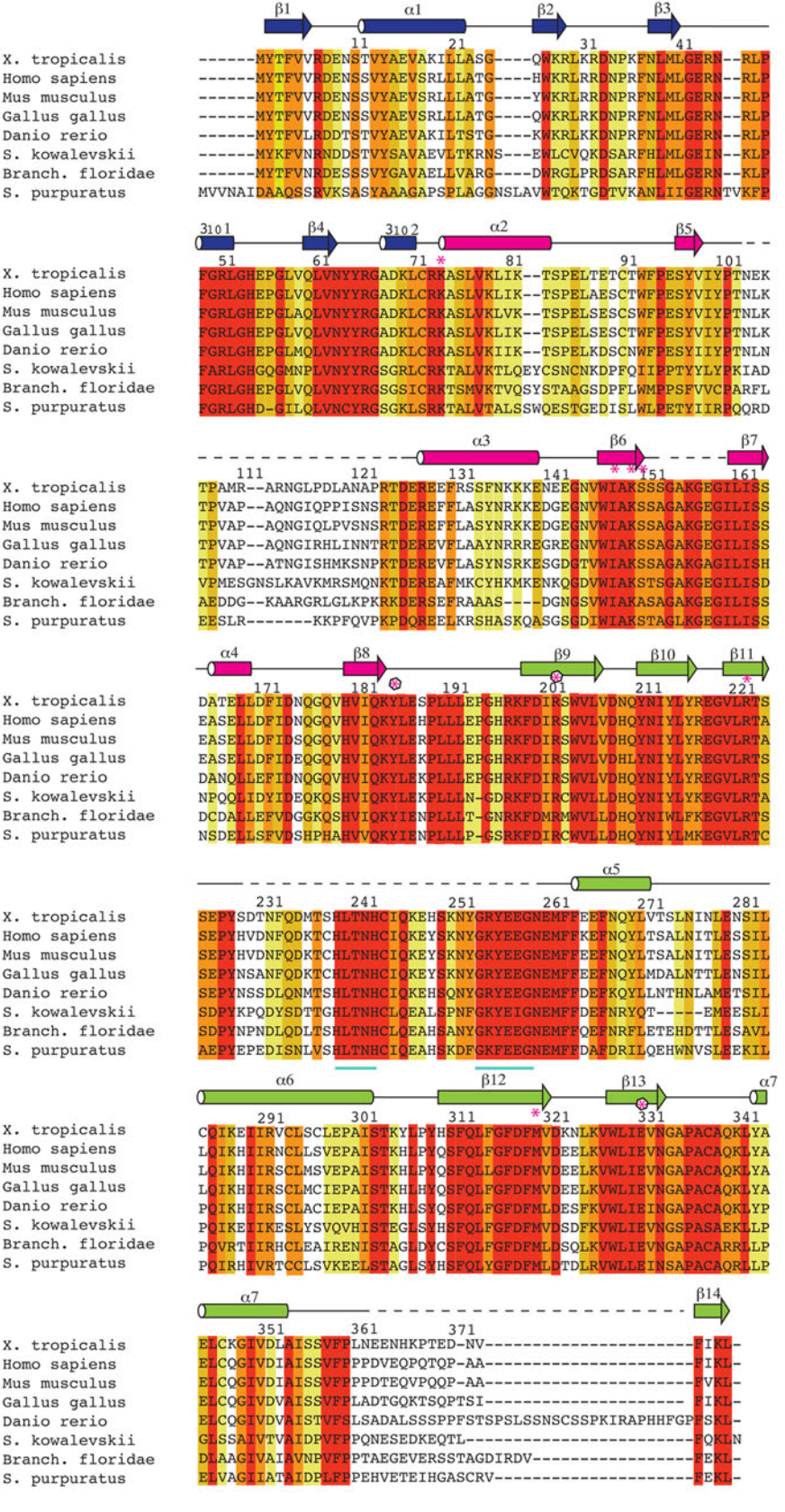
Supplementary Information

Tubulin tyrosine ligase structure reveals adaptation of an ancient fold to bind and modify tubulin

Agnieszka Szyk[1], Alexandra M. Deaconescu[2], Grzegorz Piszczek[3] and Antonina Roll-Mecak[1,3]

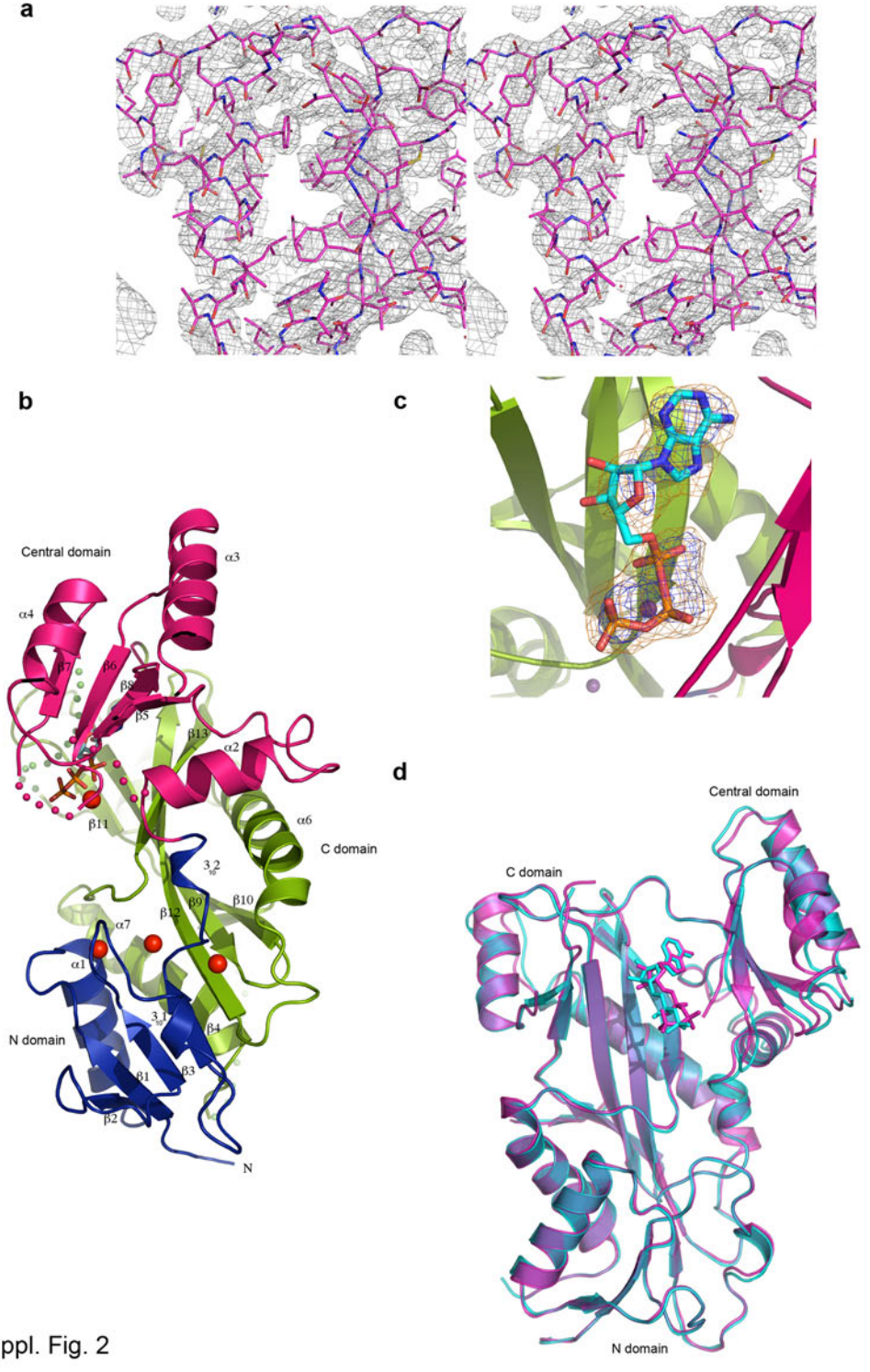

β1 α1 β2 β3
11 21 31 41
X. tropicalis ------MYTFVVRDENSTVYAEVAKILLASG----QWKRLKRDNPKFNLMLGERN--RLP 48
Homo sapiens ------MYTFVVRDENSSVYAEVSRLLLATG----HWKRLRRDNPRFNLMLGERN--RLP 48
Mus musculus ------MYTFVVRDENSSVYAEVSRLLLATG----YWKRLRRDNPRFNLMLGERN--RLP 48
Gallus gallus ------MYTFVVRDENSSVYAEVSRLLLATG----QWKRLRKDNPRFNLMLGERN--RLP 48
Danio rerio ------MYTFVLRDDTSTVYAEVAKILTSTG----KWKRLKKDNPRFNLMLGERN--RLP 48
S. kowalevskii ------MYKFVNRNDDSTVYSAVAEVLTKRNS---EWLCVQKDSARFHLMLGEIN--KLP 49
Branch. floridae ------MYTFVNRDESSSVYGAVAELLVARG----DWRGLPRDSARFHLMLGERN--KLP 48
S. purpuratus MVVNAIDAAQSSRVKSASYAAAGAPSPLAGGNSLAVWTQKTGDTVKANLIIGERNTVKFP 60
3101 β4 3102 α2 β5
51 61 71 81 91 101
X. tropicalis FGRLGHEPGLVQLVNYYRGADKLCRKASLVKLIK--TSPELTETCTWFPESYVIYPTNEK 106
Homo sapiens FGRLGHEPGLVQLVNYYRGADKLCRKASLVKLIK--TSPELAESCTWFPESYVIYPTNLK 106
Mus musculus FGRLGHEPGLAQLVNYYRGADKLCRKASLVKLVK--TSPELSESCSWFPESYVIYPTNLK 106
Gallus gallus FGRLGHEPGLVQLVNYYRGADKLCRKASLVKLIK--TSPELSESCTWFPESYVIYPTNLK 106
Danio rerio FGRLGHEPGLMQLVNYYRGADKLCRKASLVKIIK--TSPELKDSCNWFPESYIIYPTNLN 106
S. kowalevskii FARLGHGQGMNPLVNYYRGSGRLCRKTALVKTLQEYCSNCNKDPFQIIPPTYYLYPKIAD 109
Branch. floridae FGRLGHEPGLVQLVNYYRGSGSICRKTSMVKTVQSYSTAAGSDPFLWMPPSFVVCPARFL 108
S. purpuratus FGRLGHD-GILQLVNCYRGSGKLSRKTALVTALSSWQESTGEDISLWLPETYIIRPQQRD 119
α3 β6 β7
111 121 131 141 151 161
X. tropicalis TPAMR--ARNGLPDLANAPRTDEREEFRSSFNKKKENEEGNVWIAKSSSGAKGEGILISS 164
Homo sapiens TPVAP--AQNGIQPPISNSRTDEREFFLASYNRKKEDGEGNVWIAKSSAGAKGEGILISS 164
Mus musculus TPVAP--AQNGIQLPVSNSRTDEREFFLASYNRKKEDGEGNVWIAKSSAGAKGEGILISS 164
Gallus gallus TPVAP--AQNGIRHLINNTRTDEREVFLAAYNRRREGREGNVWIAKSSAGAKGEGILISS 164
Danio rerio TPVAP--ATNGISHMKSNPKTDEREVFLASYNSRKESGDGTVWIAKSSAGAKGAGILISH 164
S. kowalevskii VPMESGNSLKAVKMRSMQNKTDEREAFMKCYHKMKENKQGDVWIAKSTSGAKGEGILISD 169
Branch. floridae AEDDG--KAARGRLGLKPKRKDERSEFRAAAS----DGNGSVWIAKASAGAKGEGILISS 162
S. purpuratus EESLR-------KKPFQVPKPDQREELKRSHASKQASGSGDIWIAKSTAGLKGEGILISS 172
α4 β8 β9 β10 β11
171 181 191 201 211 221
X. tropicalis DATELLDFIDNQGQVHVIQKYLESPLLLEPGHRKFDIRSWVLVDNQYNIYLYREGVLRTS 224
Homo sapiens EASELLDFIDNQGQVHVIQKYLEHPLLLEPGHRKFDIRSWVLVDHQYNIYLYREGVLRTA 224
Mus musculus EASELLDFIDSQGQVHVIQKYLERPLLLEPGHRKFDIRSWVLVDHQYNIYLYREGVLRTA 224
Gallus gallus EASELLDFIDEQGQVHVIQKYLEKPLLLEPGHRKFDIRSWVLVDHLYNIYLYREGVLRTS 224
Danio rerio DANQLLEFIDNQGQVHVIQKYLEKPLLLEPGHRKFDIRSWVLVDHQYNIYLYREGVLRTS 224
S. kowalevskii NPQQLIDYIDEQKQSHVIQKYLEKPLLLN-GDRKFDIRCWVLLDHQYNIYLYKEGVLRTA 228
Branch. floridae DCDALLEFVDGGKQSHVIQKYIENPLLLT-GNRKFDMRMWVLLDHQYNIWLFKEGVLRTS 221
S. purpuratus NSDELLSFVDSHPHAHVVQKYIENPLLLP-GSRKFDIRCWVLLDHQYNIYLMKEGVLRTC 231
α5
231 241 251 261 271 281
X. tropicalis SEPYSDTNFQDMTSHLTNHCIQKEHSKNYGRYEEGNEMFFEEFNQYLVTSLNINLENSIL 284
Homo sapiens SEPYHVDNFQDKTCHLTNHCIQKEYSKNYGKYEEGNEMFFKEFNQYLTSALNITLESSIL 284
Mus musculus SEPYHVDNFQDKTCHLTNHCIQKEYSKNYGKYEEGNEMFFEEFNQYLTSALNITLESSIL 284
Gallus gallus SEPYNSANFQDKTCHLTNHCIQKEYSKNYGRYEEGNEMFFEEFNQYLMDALNTTLENSIL 284
Danio rerio SEPYNSSDLQNMTSHLTNHCIQKEHSQNYGRYEEGNEMFFDEFKQYLLNTHNLAMETSIL 284
S. kowalevskii SDPYKPQDYSDTTGHLTNHCLQEALSPNFGKYEIGNEMFFDEFNRYQT-----EMEESLI 283
Branch. floridae SDPYNPNDLQDLTSHLTNHCLQEAHSANYGKYEEGNEMFFQEFNRFLETEHDTTLESAVL 281
S. purpuratus AEPYEPEDISNLVSHLTNHCIQEAHSKDFGKFEEGNEMFFDAFDRILQEHWNVSLEEKIL 291
α6 β12 β13 α7
291 301 311 321 331 341
X. tropicalis CQIKEIIRVCLSCLEPAISTKYLPYHSFQLFGFDFMVDKNLKVWLIEVNGAPACAQKLYA 344
Homo sapiens LQIKHIIRNCLLSVEPAISTKHLPYQSFQLFGFDFMVDEELKVWLIEVNGAPACAQKLYA 344
Mus musculus LQIKHIIRSCLMSVEPAISTKHLPYQSFQLLGFDFMVDEELKVWLIEVNGAPACAQKLYA 344
Gallus gallus LQIKHIIRSCLMCIEPAISTKHLHYQSFQLFGFDFMVDEELKVWLIEVNGAPACAQKLYA 344
Danio rerio PQIKHIIRSCLACIEPAISTKHLSYQSFQLFGFDFMLDESFKVWLIEINGAPACAQKLYP 344
S. kowalevskii PQIKEIIKESLYSVQVHISTEGLSYHSFQLFGFDFMVDSDFKVWLIEVNGSPASAEKLLP 343
Branch. floridae PQVRTIIRHCLEAIRENISTAGLDYCSFQLFGFDFMLDSQLKVWLIEVNGAPACARRLLP 341
S. purpuratus PQIRHIVRTCCLSVKEELSTAGLSYHSFQLYGFDFMLDTDLRVWLLEINSAPACAQRLLP 351
α7 β14
351 361 371
X. tropicalis ELCKGIVDLAISSVFPLNEENHKPTED-NV---------------------FIKL- 377
Homo sapiens ELCQGIVDIAISSVFPPDVEQPQTQP-AA---------------------FIKL- 377
Mus musculus ELCQGIVDIAISSVFPPDTEQVPQQP-AA---------------------FVKL- 377
Gallus gallus ELCQGIVDVAISSVFPLADTGQKTSQPTSI---------------------FIKL- 378
Danio rerio ELCQGIVDVAISTVFSLSADALSSSPPFSTSPSLSSNSCSSPKIRAPHHFGPFSKL- 400
S. kowalevskii GLSSAIVTVAIDPVFPPQNESEDKEQTL---------------------FQKLN 376
Branch. floridae DLAAGIVAIAVNPVFPPTAEGEVERSSTAGDIRDV---------------------FEKL- 380
S. purpuratus ELVAGIIATAIDPLFPPEHVETEIHGASCRV---------------------FEKL- 386

Suppl. Fig. 1

Supplementary Figure 1

Alignment of selected tubulin tyrosine ligase sequences

Sequence numbering scheme corresponds to *X. tropicalis* TTL. Secondary structure elements are denoted as follows: α or 3_{10} helices, cylinders; β strands, arrows; random coil, lines. Dotted lines represent portions of the polypeptide chain that were not well resolved in the electron density maps. BLOSUM62[1] sequence similarity is color coded using a gradient from white (40% identity) to red (100% identity). Residues in the active site are marked by * and residues in the active site that cause a drastic loss of activity when mutated are marked by a circle. Cyan lines delineate the two conserved motifs in the substrate loop.


a
b
Central domain
α3
α4
β7
β6
β8
β5
β13
α2
β11
α6
C domain
3 10 2
β12
β9
β10
α7
α1
3 10 1
β4
N domain
β1
β3
β2
N
c
d
Central domain
C domain
N domain

Suppl. Fig. 2

Supplementary Figure 2

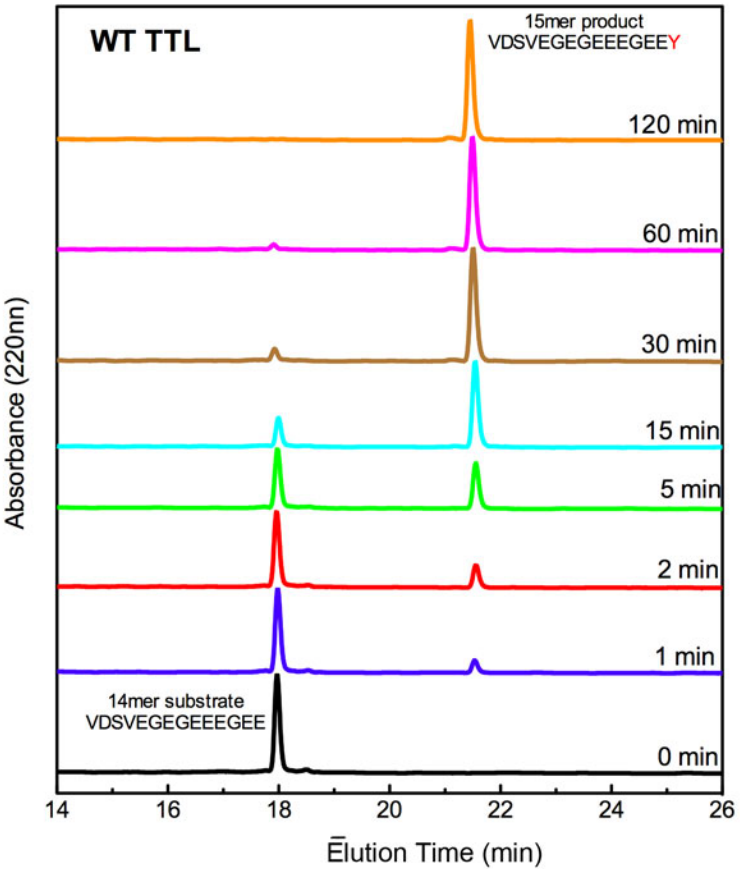
Nucleotide binding to TTL

a. Stereo view of a section of the electron density at 1σ after multi-wavelength anomalous dispersion (MAD) phasing and density modification. Shown is a region of α1.

b. Ribbon diagram of TTL color-coded as in Figure 1a and 90° rotated to show the continuous β-sheet between the N- and C-terminal domains

c. Close-up of the active site in the AMPPNP bound TTL structure showing the bound AMPPNP as a ball and stick representation and the Fo-Fc density contoured at 3σ (dark blue) before modeling the nucleotide and the 2Fo-Fc omit map density contoured at 1.2σ (gold). Secondary structure elements colored as in Figure 1a.

d. Ribbons representation of superimposed apo (grey), ADP bound TTL (cyan) and AMPPNP bound TTL (magenta). Bound nucleotide is shown as a ball and stick model. Residues in the β-sheet supporting the active site were used for the superposition. The superposition reveals there are no large conformational changes between the apo/ADP and AMPPNP crystal forms (r.m.s.d between TTL:AMPPNP and apo TTL or TTL:ADP are 1.04Å and 1.2Å, respectively for main chain atoms). The central domain shows the most plasticity possibly in response to the identity of the bound nucleotide, with residues in the loop connecting strands β6 and β7 moving up to 9Å to interact more intimately with the γ-phosphate. The central domain has the highest B-factors in the TTL structure, consistent with its conformational flexibility.

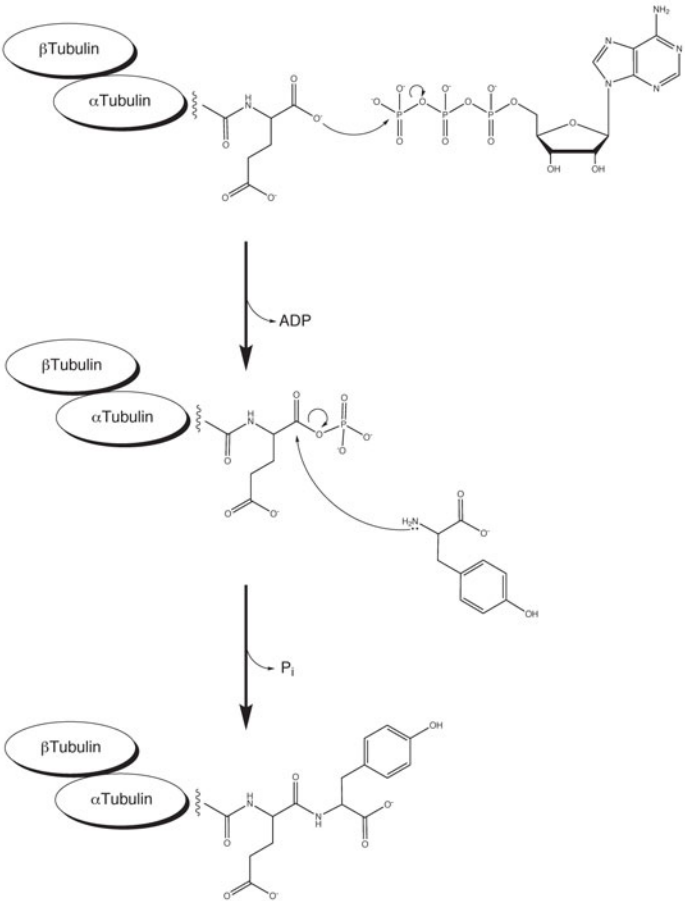

WT TTL
15mer product
VDSVEGEGEEEGEEY
120 min
60 min
30 min
15 min
5 min
2 min
1 min
14mer substrate
VDSVEGEGEEEGEE
0 min
Absorbance (220nn)
14
16
18
20
22
24
26
Elution Time (min)

Suppl. Fig. 3

Supplementary Figure 3

TTL Tyrosination activity assay using a peptide substrate

HPLC assay for α–tubulin peptide tyrosination. HPLC traces showing the progression of the tyrosination reaction monitored by the appearance and gradual increase in the absorbance of the elution peak for the tyrosinated peptide and the simultaneous decrease in the total absorbance of the elution peak of the substrate peptide from a C18 reversed phase high-performance liquid chromatography column.


βTubulin
αTubulin
ADP
βTubulin
αTubulin
Pi
βTubulin
αTubulin

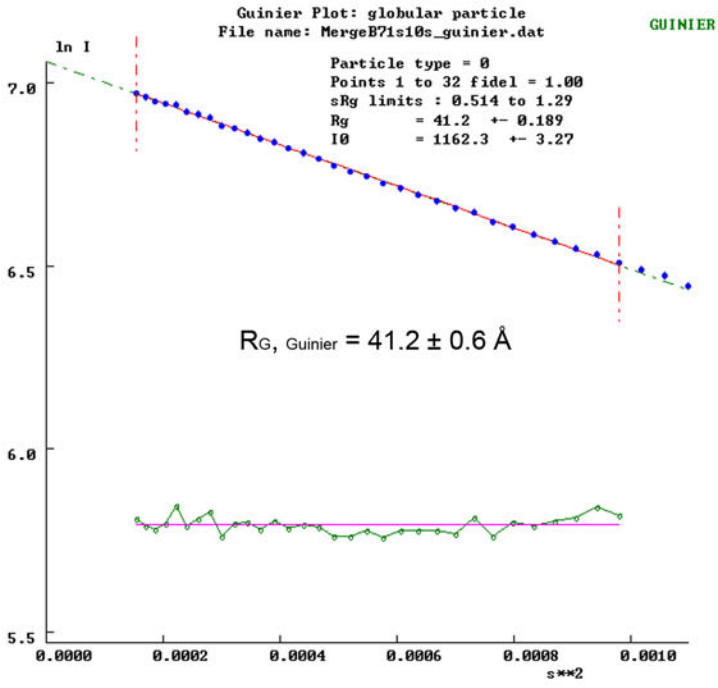
Suppl. Fig. 4

Supplementary Figure 4

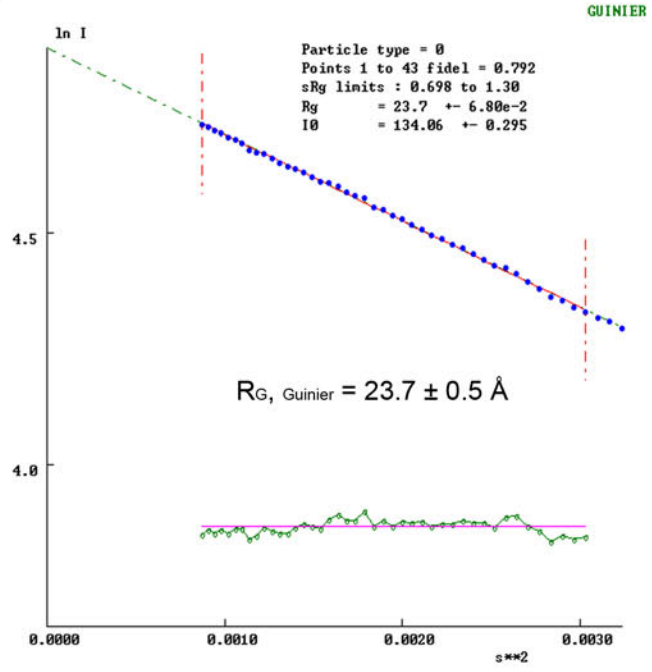
Proposed reaction mechanism for TTL

The reaction proceeds through formation of an acyl-phosphate intermediate by phosphorylation of the α-carboxylate of the terminal Glu of the α-tail with the γ-phosphate of ATP, followed by nucleophilic attack on the phosphorylated carboxylate by the amine of the incoming Tyr to form Tyr-tubulin.

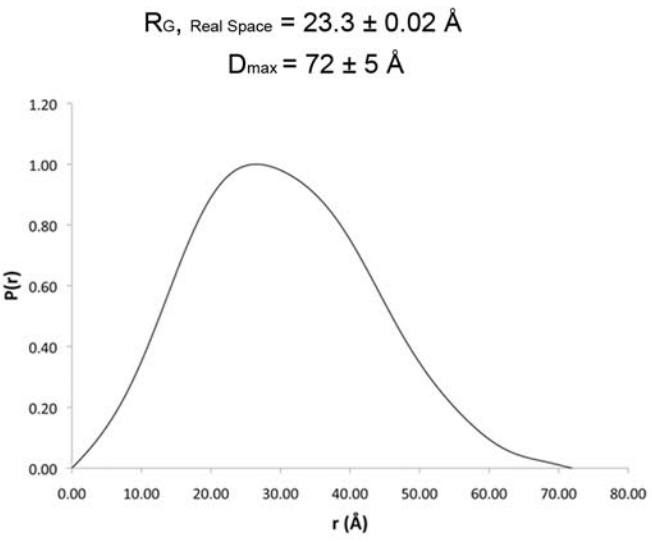
a



b



c



Suppl. Fig. 5

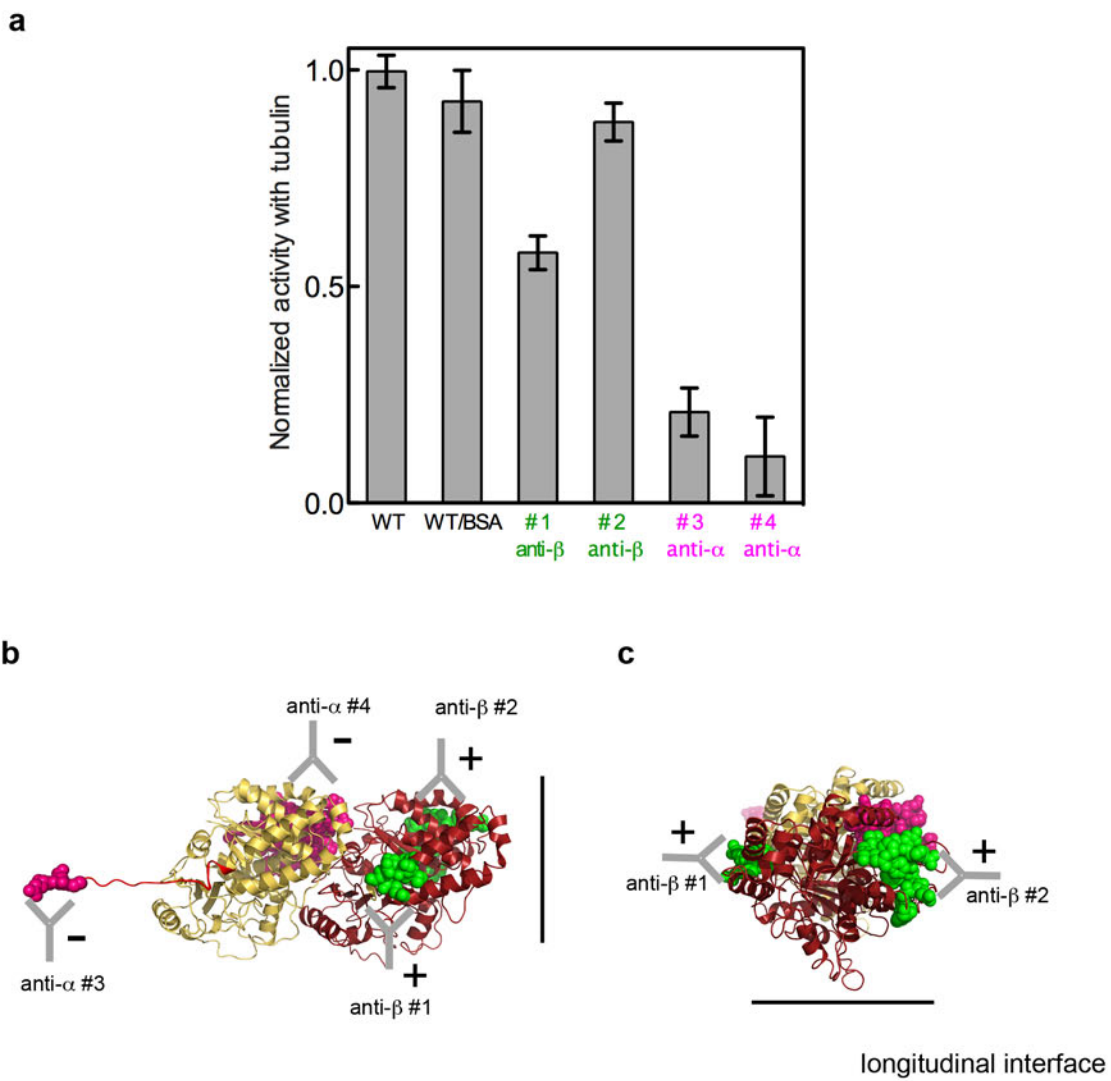
Supplementary Figure 5

Small-angle X-ray scattering Analysis of the TTL–tubulin complex and TTL

a. Guinier plot for the experimental data collected for the TTL–tubulin complex with linear fit (pink line). Graphical representation of the residuals is shown in green. The radius of gyration (R_G) was derived by the Guinier approximation $I(q) = I(0)\ \exp(-q^2R_{G2}/3)$ with the limits $qR_G < 1.3$

b. Guinier plot for the experimental data collected for TTL with linear fit (pink line). Graphical representation of the residuals is shown in green. The radius of gyration (R_G) was derived by the Guinier approximation $I(q) = I(0)\ \exp(-q^2R_{G2}/3)$ with the limits $qR_G < 1.3$.

c. Pair-distance distribution function [P(r)] computed from the experimental SAXS data for TTL.


a
Nomalized activity with tubulin
1.0
0.5
0.0
WT
WT/BSA
#1
anti-β
#2
anti-β
#3
anti-α
#4
anti-α
b
anti-α #4
anti-β #2
anti-α #3
anti-β #1
c
anti-β #1
anti-β #2
longitudinal interface

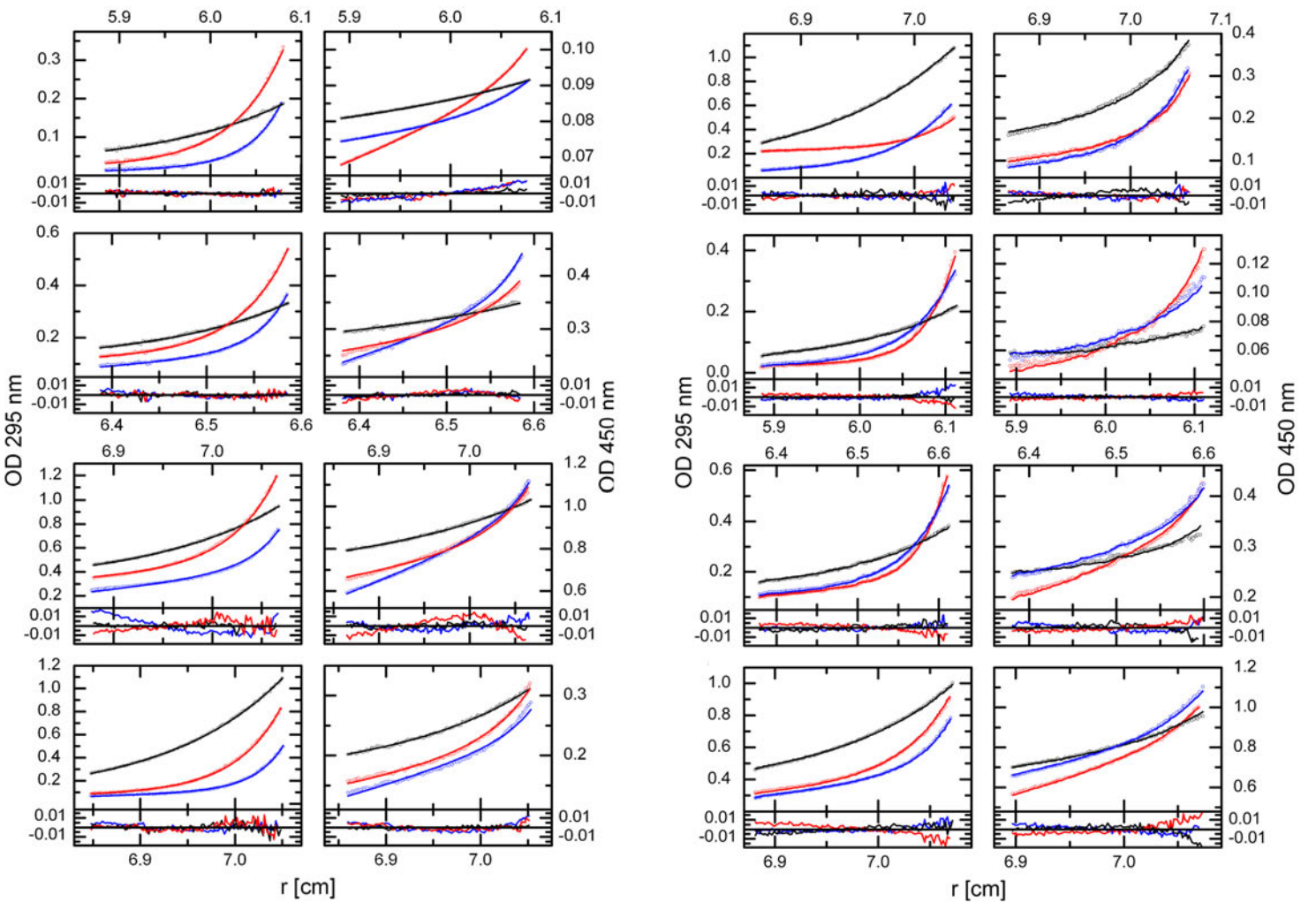
Suppl. Fig. 6

Supplementary Figure 6

Antibodies recognizing β-tubulin do not inhibit TTL tyrosination activity.

a. Tubulin tyrosination activity carried out in the presence of antibodies recognizing epitopes on α- and β-tubulin (#1 anti-β = monoclonal [D66] anti-β-tubulin (Abcam, Catalog # ab11307), epitope amino acids 427-432; #2 anti-β = monoclonal [TU-6] anti-β-tubulin (Abcam, Catalog # ab7792), epitope amino acids 81-95; #3 anti-α = monoclonal anti-Glu-α-tubulin (Synaptic Systems Catalog # 30020011), epitope XEE including the terminal α-carboxylate group; #4 anti-α = monoclonal [TU-1] anti-α-tubulin (Abcam, Catalog # ab7750), epitope amino acids 65-97).

b, c. Location of antibody epitopes on the tubulin heterodimer structure (**b**, orientation as in Figure 5e; **c**, 90° rotation of **b**). Epitopes recognized by the antibodies are shown as balls (α-tubulin, magenta, β-tubulin, green). (+) indicates no effect of antibody on tubulin tyrosination activity, (-) denotes inhibition.



Suppl. Fig. 7

Supplementary Figure 7

Analytical Ultracentrifugation Analysis of TTL binding to α-tubulin tail substrate and product peptides

a. The AUC equilibrium sedimentation profiles of four TTL-substrate peptide mixtures at three different speeds (15,000 rpm, black; 24,000 rpm, red; 30,000 rpm, blue) measured at two wavelengths, 295 nm scans on the left and 450 nm on the right column, respectively. Solid lines represent fits to the data. The peptide was labeled with fluorescein on its N-terminal amine to aid detection. Each row represents a different sample (panel 1: 5 μM TTL + 3.75 μM peptide; panel 2: 5 μM TTL +18.8 μM peptide; panel 3: 15 μM TTL + 56 μM peptide; panel 4: 40 μM TTL + 20 μM peptide). The lower narrow panel shows residual deviation of the fit.

b. The AUC equilibrium sedimentation profiles of four TTL-product peptide mixtures at three different speeds (15,000 rpm, black; 24,000 rpm, red; 30,000 rpm, blue) measured at two wavelengths, 295 nm scans on the left and 450 nm on the right column, respectively. Solid lines represent fits to the data. The peptide was labeled with fluorescein on its N-terminal amine to aid detection. Each row represents a different sample (panel 1: 40 μM TTL + 20 μM peptide; panel 2: 5 μM TTL +4 μM peptide; (3) panel 3: 5 μM TTL + 20 μM peptide; (4) panel 4: 15 μM TTL + 60 μM peptide). The lower narrow panel shows residual deviation of the fit.

Supplementary Methods

Protein Expression and Purification

Xenopus tropicalis tubulin tyrosine ligase (residues 2-377) was expressed in *E. coli* as a GST-fusion protein. Protein expression was induced in BL21 Rosetta (DE3) pLysS *E. coli* strain with 0.5 mM IPTG for 16 hours at 16C. The cells were lysed using a microfluidizer, and the cellular debris was removed by centrifugation at 100,000xg for 1 hour. The protein was purified with a 20 ml GST resin column (GE Healthcare) and then the N-terminal GST tag was cleaved with Tev protease. Following protease cleavage, TTL was further purified by ion exchange, hydrophobic interaction and size exclusion chromatography. The SeMet protein was expressed in a Met auxotroph strain and purified as the wild-type protein. Mutagenesis was performed using Quickchange (Stratagene). The structure guided TTL mutants used in this study were purified following the same protocol as for the wild-type protein. Purity and integrity of wild-type TTL as well as all engineered mutants were verified by mass spectrometry and measured masses agreed with the predicted masses within 2Da.

Crystallization and Structure Determination

Crystals of apo TTL (~ 8mg/ml) grew at room temperature by hanging drop vapor diffusion in 0.1 M Hepes pH 7.5, 8% (v/v) PEG8000, 8% ethylene glycol, 0.05 M $MgCl_2$ with the symmetry of space group C2 with one TTL copy per asymmetric unit (Unit Cell: $\underline{a}$=116.6 Å, $\underline{b}$=76.2 Å, $\underline{c}$=44.2 Å, β=90.6°; diffraction limit=2.8Å). More than 90% of the crystals were split and/or were weakly diffracting and were not suitable for structure determination. Se-Met TTL crystals

grew with symmetry C2 in 0.1 M Hepes pH 7.5, 8% (v/v) PEG8000, 12% ethylene glycol, 0.05 M $MgCl_2$ (Unit Cell: *a* =116.7 Å, *b* =75.6 Å, *c*= 44.1 Å, β=90.8°; diffraction limit 2.4Å). The crystals were flash frozen in liquid nitrogen and used for X-ray data collection [Advanced Light Source (ALS), Beamlines 8.2.1 and 8.2.2]. Se-Met MAD data were collected at two X-ray wavelengths, corresponding to the high energy remote (λ1) and inflection point (λ2) of the Se K-absorption edge. Diffraction data were reduced using HKL2000[3]. Three of the possible five Se sites were found using SOLVE[4]. The experimental electron density maps at this stage were not interpretable [mean overall figure of merit (FOM) 0.14], but density modification (solvent flipping and histogram matching) yielded maps in which ~35% of the protein could be built. Better quality maps were obtained by combining the MAD Se-Met phases with phases to 6 Å from Ta_6Br_{12} soaked crystals. The location of the Ta_6Br_{12} cluster was determined by anomalous difference Fourier synthesis using model phases. Density modification with SOLOMON[5] gave interpretable maps in which we were able to build another ~10% of the polypeptide chain; however, the rest of the electron density map was too noisy to allow further building.

Addition of $BaCl_2$ to crystallization conditions yielded larger, stronger diffracting crystals (Unit Cell: a=116.4, b=75.8, c=43.9,β=90.6°, diffraction limit 2.24Å) for which MAD data were collected at two X-ray wavelengths, corresponding to the high energy remote (λ1) and inflection point (λ2) of the Se K-absorption edge (Table 1). Phases were calculated in SHARP[6] using the three Se sites previously determined with SOLVE. The anomalous phasing power for the peak (between 30Å and 2.4Å) was 0.628 and the dispersive phasing power was 0.529. The SHARP experimental map (mean overall FOM=0.31) was interpretable only after density modification (mean overall FOM of 0.80; Supplementary Fig. 2a), but was of sufficient quality to build 60%

of the protein. Refinement of the partial model using CNS[7] and calculation of electron density maps with combined phases and later $(2|F_{obs}|-|F_{calc}|)\alpha_{model}$ difference Fourier syntheses allowed completion of protein structure building using COOT[8]. The current refinement model of TTL consists of 308 residues plus two magnesium ions and 36 water molecules. Four regions of the polypeptide chain (residues 103-126; residues 155-158; residues 226-259; residues 364-373) are not well resolved in the electron density map, and are presumed disordered (the integrity of the protein was checked by SDS gel electrophoresis and showed no proteolytic cleavage in the crystal). The working and free R factors for the TTL structure at 2.5Å resolution are 24.9% and 29.8%, respectively (Table 1). PROCHECK[9] revealed no unfavorable (ϕ,ψ) combinations, and main chain and side-chain structural parameters consistently better than average (overall G-factor=0.2).

Crystals of TTL:AMPPNP grew with space group symmetry $P2_1$ in 0.1M MES 6.0, 11% Peg3350, 5mM $MgCl_2$, 1mM AMPPNP, with two TTL copies per asymmetric unit (Unit Cell: a=44.4, b= 74.6, c=117.3, β=90.1°; diffraction limit=2.5Å). The structure was solved by molecular replacement using PHASER[10] by using the apo TTL structure as the search model (residues around the active site were excluded from the model). Difference electron density maps revealed unambiguous density for AMPPNP (Supplementary Fig. 2c). Several rounds of iterative model building and refinement were performed using COOT[8] and CNS[7]. The current refinement model of TTL:AMPPNP consists of 624 residues plus seven magnesium ions and 30 water molecules. Four regions of the polypeptide chain (residues 103-125; residues 154-157; residues 229-259; residues 363-371) are not well resolved in the electron density map in either of the molecules in the AU, and are presumed disordered. The working and free R factors for the TTL

structure at 2.5Å resolution are 25.4% and 28.5%, respectively (Table 1). PROCHECK[9] revealed no unfavorable (ϕ,ψ) combinations, and main chain and side-chain structural parameters consistently better than average (overall G-factor=0.2). Even though the crystals were grown in the presence of 0.5 mM α-tail peptide (EEEGEE-iodoY), no electron density features for the peptide or the iodo-Y could be identified in our maps (more below).

Crystals of TTL:ADP grew with space group symmetry C2 in 0.1M HEPES 7.0, 14% PEG8000, 5mM $MgCl_2$, 1mM ADP, with one TTL copy per asymmetric unit (Unit Cell: a=117.0, b=75.7, c=44.2, β=90.8°; diffraction limit=2.9Å). The structure was solved by molecular replacement using PHASER[10] using the apo TTL structure as the search model. Several rounds of iterative model building and refinement were performed using COOT[8] and CNS[7]. The current refinement model of TTL:ADP consists of 312 residues plus three magnesium ions and two water molecules. Four regions of the polypeptide chain (residues 103-126; residues 155-158; residues 225-260; residues 364-373) are not well resolved in the electron density map, and are presumed disordered. The working and free R factors for the current TTL:ADP structure at 2.9Å resolution are 23.1% and 31.4%, respectively (Table 1). PROCHECK[9] revealed no unfavorable (ϕ,ψ) combinations, and main chain and side-chain structural parameters consistently better than average (overall G-factor=0.1).

Several dozen data sets were collected and analyzed for TTL crystals grown in the presence of ADP, ATPγS or AMPPNP and α-tubulin tail peptides at 0.5 to 1mM concentrations (substrate: EEEGEE and VDSVEGEGEEEGEE as well as product: EEEGEEY and

VDSVEGEGEEEGEEY), but no clear electron density corresponding to peptide could be seen. Data sets were collected and analyzed for crystals grown in the presence of product peptides with iodo-Y, but no electron density features for the iodo-Y could be identified in our maps. This is not surprising given the low affinity that TTL has for the α-tail peptide (K_D=144 μM and 96μM for the substrate and product peptides, respectively). Structural figures were prepared with the programs PyMOL[11] and CHIMERA[12]. Maps of electrostatic surfaces were calculated using DELPHI[13].

Gel filtration Assays

Complex formation between TTL or TTL mutants and tubulin (Cytoskeleton) was monitored using gel filtration chromatography. Tubulin and TTL or TTL mutant were mixed at a molar 2:1 ratio (60 μM tubulin, 30 μM TTL) in the absence or presence of 1mM ADP, 1mM ATP or 1 mM ATPγS. After 20 min incubation at room temperature, the mixture was analyzed by gel filtration chromatography on a Superdex 75 10/300 GL column (GE Healthcare). Components of the mixture were eluted with a buffer consisting of 50 mM MES/K pH 6.8, 100 mM KCl, 10 mM $MgCl_2$, 5 mM DTT (0.2 mM ATP, 0.2 mM ATPγS or 0.2mM ADP were present in the gel filtration buffer where nucleotide dependence for tubulin binding was investigated) at a flow rate of 0.5 ml/min. The absorbance of the eluting solution was monitored at 280 nm and fractions were collected for SDS-PAGE analysis and stained with Sypro (Invitrogen) for visualization. The same experiment was repeated for TTL or TTL mutants alone. Complex formation between tubulin and TTL or TTL mutants was monitored both as disappearance of the TTL signal and co-elution of both components (Fig. 4).

X-ray Scattering Data Acquisition, Analysis and Modeling

SAXS data for the reconstructions of the *Xenopus tropicalis* TTL:tubulin complex were collected at the SIBYLS beamline at ALS using a Mar 165 CCD area detector (165 mm diameter). 15μl of sample was placed in a quartz cell maintained at 4 C°. To avoid condensation on the walls of the cuvette, the chamber was flooded with helium. The complex and matching buffer solutions (20 mM Hepes pH 7.0, 100 mM KCl, 1 mM EGTA, 1 mM $MgCl_2$, 100 μM GTP, 1 mM ATPγS, 5mM DTT) were exposed for 12s, subdivided into two 1s exposures, one 10s exposure followed by another 1s exposure. Scattering was measured at concentrations of 5.2 mg/ml, 2.6 mg/ml and there was no evidence for aggregation at the higher concentration. The raw scattering data were scaled and buffers were subtracted by using software written by Greg Hura (ALS SYBILS). Individual scattering curves were visually inspected for radiation damage prior to averaging. The scattering curves that were collected at different concentrations were normalized and merged together in PRIMUS[14] to yield a low-noise composite curve. The radii of gyration (R_g) were initially computed from the Guinier plot[15] (Supplementary Fig. 5). The pair distance distribution function, P(r), was calculated by using the indirect Fourier transform method as implemented in GNOM[14], and values of the R_g were computed from the second moment of the P(r) function. These values for R_g compared favorably to the initial values derived from Guinier plots, 43.0 ± 0.2 Å and 41.2± 0.6 respectively (Fig. 5a and Supplementary Fig. 5). Porod volume analysis, as implemented in PRIMUS, was used to estimate the molecular weight of the scattering species in solution. For the scattering curve obtained with the sample at 5.2 mg/ml, the calculated Porod volume was 250nm^3, corresponding to a molecular weight of 150kDa. For the scattering curve collected at a concentration of 2.6mg/ml, the Porod Volume was 243.9 nm^3 (corresponding to a

molecular weight of 146kDa). In order to determine the maximum dimension of the particle (D_{max}), the P(r) function was first computed with GNOM with no constraint on the function to become zero at D_{max}. This gave the initial estimate for D_{max} that was used in the subsequent analysis. For the final determination of P(r), the function was constrained to go to zero at D_{max}. Because the inverse scattering problem has no unique solution, fifteen *ab initio* simulations were performed using GASBOR[16] and minimization against the experimental scattering data rather than the pair-distribution function. GASBOR assigns a pseudo residue to each residue in the protein. The only information that goes into the shape reconstruction other than the X-ray curve is the number of residues in the complex as well as the D_{max} obtained from the X-ray data. A value of 165Å was used in the *ab initio* modeling. The dummy atom models resulting from fifteen individual GASBOR runs were aligned, averaged, and scored with a normalized spatial discrepancy (NSD) using the program package DAMAVER[17]. The criterion for inclusion in averaging procedures was NSD < mean NSD + 2 * variation. According to this criterion, one of the fifteen simulations was discarded from analysis. The resulting probability map was filtered with DAMFILT[17]. The fourteen selected *ab initio* models agree well, yielding 1.31 ±0.05 (NSD ± SD). The crystallographic α/β tubulin model (PDB ID 1JFF) was fit as a rigid-body into the SAXS envelope manually, and then locally adjusted using the fit-in-map algorithms implemented in CHIMERA[12].

SAXS data for *Xenopus tropicalis* TTL bound to ATPγS were collected at the SIBYLS beamline at ALS using a Mar 165 CCD area detector (165 mm diameter). 15µl of sample was placed in a Quartz cell maintained at 4C. To avoid condensation on the walls of the cuvette, the chamber was flooded with Helium. The protein and matching buffer solutions (20 mM Hepes pH 7.0, 100

mM KCl, 10 mM $MgCl_2$, 1 mM ATPγS, 2mM TCEP) were exposed for 8s, subdivided into two 1s exposures, one 6s exposure followed by another 1s exposure. The curves were measured at concentrations of 8, 4, and 2mg/ml, and there was no evidence for aggregation at higher concentrations. The raw scattering data were scaled, and buffers were subtracted by using software written by Greg Hura (ALS SYBILS). Individual scattering curves were visually inspected prior to averaging to insure that radiation damage was minimal. Individual scattering curves that were collected at different concentrations or exposure duration were scaled and merged together using PRIMUS[18] to yield a low-noise interference-free composite curve. Guinier plot analysis was employed to assess any potential aggregation for data points satisfying the requirement qR_g <1.3. Radii of gyration were initially computed based on the Guinier plot analysis in PRIMUS[18], and pair distribution functions P(r) and D_{max} values were calculated using the indirect Fourier transform as implemented in GNOM. To properly assess D_{max}, P(r) was initially not constrained to be zero. This gave us an initial estimate of D_{max} that was used in subsequent analysis. For the final processing of the data, P(r) was constrained to become zero at D_{max}. The radii of gyration were initially computed from the Guinier plot[15]. Real-space values of the R_g were computed from the second moment of the P(r) function. These R_g values were in excellent agreement with the initial values derived from Guinier plots, 23.3 ±0.02Å and 23.7 ±0.5Å respectively (Supplementary Fig. 5b, c).

Analytical Ultracentrifugation Experiments

Sedimentation Velocity (SV) AUC experiments were performed at 20 °C with the ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA) using absorption detection optics.

We used sedimentation velocity and not sedimentation equilibrium AUC because of the labile nature of tubulin, which tends to slowly denature and aggregate when kept for the extended time periods required for the latter experiment. Peak fractions of TTL, tubulin or TTL:tubulin from the size exclusion chromatography column were used to prepare samples at loading concentrations ranging from 0.3 μM to 5.25 μM of the tubulin dimer and from 0.3 μM to 22 μM of TTL. All samples and the running buffer (50 mM MES-K pH 6.8, 100 mM KCl, 10 mM $MgCl_2$, 1 mM TCEP; 40 μM GTP for the runs with tubulin) were loaded into the 12 mm path length, double sector cells (0.4 mL). After thermal equilibration at rest, the rotor was accelerated to 50,000 rpm and radial scans at 280 nm or 230 nm were taken continuously until no further boundary movement was observed. Data were analyzed in terms of continuous c(*s*) distributions using the SEDFIT program and with the discrete species model in the SEDPHAT program[19]. Buffer density and viscosity was measured using the DMA-58 densitometer and AMVn viscometer (Anton-PAAR KG, Graz, Austria), respectively. The partial specific volume of TTL was calculated using the SEDNTERP program[20] and for tubulin the partial specific volume of 0.736 ml/g was used[21]. In all types of analyses menisci and bottom positions as well as f/f_0 ratios were optimized during the fitting procedure. For all cases good fits were obtained with Root Mean Square Deviations (RMSD) in the 0.003 to 0.006 OD range for scans at 280 nm, and 0.006 to 0.009 OD RMSD range for the lowest protein concentrations scanned at 230 nm. Sedimentation coefficients were corrected to standard conditions at 20 °C in water, $s_{20,w}$.

The c(*s*) analysis of SV data obtained with TTL and tubulin alone shows in both cases single peak distributions representing the TTL monomer with $s_{20,w}$ of 3.36 S and the tubulin dimer with $s_{20,w}$ of 6.15 S, respectively (Fig. 4). The c(*s*) distributions obtained with protein mixtures

containing 1:1 tubulin:TTL ratios at increasing concentrations show two peaks, a small *s* value peak representing free TTL, with the area of the peak decreasing with increasing protein concentration, and the high *s* value peak representing the position of the tubulin-TTL complex reaction boundary. The position of the complex peak shifts to higher *s* values with increasing protein concentration, according to the predictions from Gilbert-Jenkins theory of bimodal reaction boundaries of rapidly interacting systems[22,23]. The width of the protein complex peak is broader than peaks obtained for either protein sedimenting alone, further indicating the presence of a dynamic equilibrium between free proteins and the TTL-tubulin complex. In the 5 to 6 μM protein concentration range the c(*s*) distributions exhibit particularly broad peaks of the protein complex, with the weight average $s_{20,w}$ values in the 6.7 S to 7.3 S range and peak areas representing over 90% of the protein mixture loading concentrations, indicating that the association reaction is fast relative to the time scale of the sedimentation experiment ($k_{off} > 10^{-2}$ s^{-1}; see reference 23). With increasing concentrations of the rapidly interacting components the position of the reaction boundary will asymptotically approach the *s* value of the complex. To obtain SV data at higher protein concentration and to avoid non-ideal sedimentation effects that can occur in protein samples above 1 mg/ml, we increased the concentration of TTL, the smaller protein, to 22 μM, but kept the tubulin concentration in the mixture at 5.25 μM. The c(*s*) distribution obtained with this sample shows two dominant narrow peaks (Fig. 4). The peak with weight average $s_{20,w}$ value of 3.37 S represents an excess of free TTL protein and the second peak, with weight average $s_{20,w}$ value of 7.29 S, represents the TTL:tubulin complex. The molecular mass of the complex species calculated from the c(*s*) distribution is 141,974 Da and is in very good agreement with the theoretical mass of the TTL:tubulin complex (143,584 Da). Since the c(*s*) distribution of this high concentration sample shows two peaks of approximately

the same area representing two solutes with possibly different shapes, but the c(*s*) method uses one average f/f_0 frictional ratio for the whole distribution, we performed additional data analysis using the discrete species model to make sure that this assumption does not affect the molecular mass determination. In the species analysis of the SV data the discrete solutes components have their own *s* and *M* parameters, and are not connected to the determination of a common f/f_0 ratio. To further constrain the species analysis fit we fixed known molecular masses of free proteins and their putative 1:1 complex in the fitting procedure. This analysis yielded $s_{20,w}$ values of 3.36 S and 7.27 S for the free TTL and for the TTL:tubulin complex, respectively. Importantly, this constrained fit has the same final RMSD value of 0.005 OD as the general c(s) distribution fit for that sample, indicating that within the experimental error the sample can be represented as a mixture of free TTL and the 1:1 TTL:tubulin complex.

Sedimentation equilibrium AUC experiments were performed at 10°C using 6 channel 1.2 cm path length cells. Absorption scans at 295 nm and 450 nm were taken after equilibrium was reached at 15,000; 24,000 and 30,000 rpm. The substrate peptide (VDSVEGEGEEEGEE) was labeled with fluorescein on its N-terminal α-amine to aid detection. Experiments were performed for each component separately (30 μM peptide and 20 μM TTL, respectively) and the following mixtures: (1) 5 μM TTL + 3.75 μM peptide; (2) 5 μM TTL +18.8 μM peptide; (3) 15 μM TTL + 56 μM peptide; (4) 40 μM TTL + 20 μM peptide. The product peptide (VDSVEGEGEEEGEEY) was also labeled with fluorescein on its N-terminal α-amine to aid detection. Experiments were performed for each component separately (30 μM peptide and 20 μM TTL, respectively) and the following mixtures: (1) 40 μM TTL + 20 μM peptide; (2) 5 μM

TTL +4 μM peptide; (3) 5 μM TTL + 20 μM peptide; (4) 15 μM TTL + 60 μM peptide. The SEDPHAT software[19] was used for all data analysis. The apparent buoyant mass and extinction values for TTL and fluorescein tubulin tail peptides were optimized by separate global analysis of data collected on samples containing TTL and labeled peptides alone. Those values were fixed in a final analysis used to obtain a best fit association constant by global analysis of data from four different reaction mixtures collected at two wavelengths and three rotor speeds. The fit residuals for the global analysis of 24 data sets for each of the experiments with either the substrate or product peptide did not exceed +/- 0.01 absorbance units.

Assay for the tyrosination of α−tubulin

Tyrosination of bovine brain tubulin (Cytoskeleton) was assayed in a 10μl reaction mixture composed of 50mM MES/K, 100 mM KCl, 10 mM $MgCl_2$, 5 mM DTT, 2.5 mM ATP, 0.3 mM L-tyrosine. Tubulin at the concentration of 5μM was incubated for 60 min at 37^{o}C with TTL at molar ratios 1:1. 1:0.02 and 1:0.002. Control reactions were performed in the same conditions, but lacking TTL, ATP or Tyr, respectively. The enzymatic reaction was terminated with 2x SDS-PAGE Sample Buffer (10 μl). Samples containing 0.05 μg of tubulin were subjected to Western blotting, using anti-Tyr-α-tubulin (1:500, Synaptic Systems) or anti-Glu-α-tubulin (1:5000, Synaptic Systems) and anti-mouse peroxidase-conjugated antibody (1:2000, Cell Signaling) as a secondary antibody. The relative intensities of Western blot bands were quantified using the program Image J (NIH).

Assay for tyrosination activity using a C-terminal α−tubulin peptide

The synthetic peptide, VDSVEGEGEEEGEE (Biosynthesis), was dissolved in 20 mM MEK/K pH 6.8, 100 mM KCl, 10 mM $MgCl_2$ to a concentration of 10 mM, and the pH of the solution was adjusted to neutral with 5M KOH. Tyrosination of the peptide by TTL was performed by mixing 250µl of solution consisting of 20 mM MES/K pH 6.8, 100 mM KCl, 10 mM $MgCl_2$, 5 mM DTT, 2.5 mM ATP, 0.3 mM L-tyrosine, 0.2 mM peptide and 1µM TTL or mutated enzyme[24]. After incubation for 0-120 min at room temperature, aliquots of the reaction mixture (25 µl) were mixed with equal volumes of 0.1% trifluoroacetic acid, and subjected to high-resolution liquid chromatography (HPLC) using a reversed phase (RP) analytical column from Vydac (218MS C18 5 mm, 25 x 0.46 cm). The substrate tertadecapeptide, corresponding to the C-terminal sequence of α−tubulin, is tyrosinated by TTL when incubated in the presence of L-tyrosine and ATP. The slowly hydrolyzable analogs ATPγS and AMPPNP do not support tyrosination. The product, VDSVEGEGEEEGEEY, contains a single C-terminal tyrosine and is more hydrophobic then the substrate. This difference aided the separation of both peptides using RP HPLC with an acetonitrile gradient (5-35% in 30min) in 0.05% trifluoroacetic acid as an eluent. The gradual conversion during the enzymatic reaction of a substrate peptide to the product is illustrated in Supplementary Fig. 3. Quantities of both substrate and product peptides were estimated from the areas of corresponding peaks on the UV detection absorbance spectrum (220 nm), and the percentages of product formation *versus* time were plotted for TTL and various mutants to obtain initial rates.

Tyrosination of α-tubulin in the presence of tubulin antibodies

Tyrosination of bovine brain tubulin (Cytoskeleton) in the presence of tubulin antibodies was assayed in a 10μl reaction mixture in 20mM MES/K, 100 mM KCl, 10 mM $MgCl_2$. Prior to the enzymatic reaction 0.2 μM tubulin was incubated for 10 min at 4°C with the following antibodies: 3.6 μM mouse monoclonal [D66] anti-β-tubulin (Abcam), 4.2 μM mouse monoclonal [TU-1] anti-α-tubulin (Abcam), 3.4 μM mouse monoclonal [TU-6] anti-β-tubulin (Abcam), 4.0 μM anti-Glu-α-tubulin (Synaptic Systems), and as a control 4.0 μM bovine albumin. After the incubation each sample was supplemented with 0.3 mM L-tyrosine and 0.02 μM TTL. The enzymatic reaction was started with 2.5mM ATP and incubated for 1 hour at room temperature. Control reaction for each condition lacked ATP. The reaction was terminated with 2x SDS-PAGE Sample Buffer (90 μl). Samples containing 0.01 μg of tubulin were subjected to Western blotting, using mouse monoclonal anti-Tyr-α-tubulin Clone TUB-1A2 (1:5000, Sigma-Aldrich) and anti-mouse peroxidase-conjugated antibody (1:2000, Cell Signaling) as a secondary antibody. Western blots were quantified using the program ImageJ (NIH) and normalized tyrosination activity is shown in Supplementary Fig. 6.

Microtubule Polymerization Assays

Turbidity experiments were performed with 20 μM tubulin and 3.5, 10 and 20 μM TTL (Fig. 6a) as well as 20 μM TTL mutants (Fig. 6b). Purified proteins were mixed with tubulin in the presence of 1mM ATP and incubated in assembly buffer on ice for 5 minutes (as described in[25] and then injected into 37C° heated quartz cuvettes. Absorbance at 350 nm was monitored at 0.5 sec intervals.

In vivo Microtubule Dynamics

Control USOS cells (mKusabira Orange-EB3 stable cell line) or cells expressing wild-type TTL N-terminally tagged with GFP or TTL mutant were imaged every 3 s on a spinning disk confocal microscope using a 100x 1.49NA objective lens on a inverted microscope (TE2000, Nikon) equipped with a Perfect Focus System (Nikon) and a spinning disc confocal scan head (CSU-X1; Yokogawa). Movies were recorded using a cooled charge-coupled device (CoolSnap HQ2; Photometrics) using 500 ms exposure times. Microscope system automation was controlled with MetaMorph software (MDS Analytical Technologies). MT dynamics were analyzed from EB3 movies using plusTipTracker[2], a Matlab-based, open-source software package that combines automated detection, tracking, analysis, and visualization tools for movies of fluorescently labeled MT plus end binding proteins.

References

1. Henikoff, S. & Henikoff, J.G. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89, 10915-9 (1992).

2. Matov, A. et al. Analysis of microtubule dynamic instability using a plus-end growth marker. *Nat Methods* 7, 761-8 (2010).

3. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276, 307-326 (1997).

4. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* 55, 849-61 (1999).

5. Abrahams, J.P. & Leslie, A.G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr* 52, 30-42 (1996).

6. Fortelle, E.d.l.a.B., G. Maximum-Likelihood Heavy-Atom Parameter Refinement for Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction Methods. in *Methods in Enzymology Macromolecular Crystallography Part A*, Vol. Vol. 276. (ed. Sweet, C.W.C.a.R.M.) 472-494 (Academic Press, San Diego, 1997).

7. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nat Protoc* 2, 2728-33 (2007).

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

9. Laskowski, R.A., Moss, D.S. & Thornton, J.M. Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231, 1049-67 (1993).

10. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674 (2007).

11. DeLano, W.L. The PyMOL Molecular Graphics System. (2002).

12. Pettersen, E. et al. UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612 (2004).

13. Rocchia, W. et al. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J Comput Chem* 23, 128-37 (2002).

14. Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25, 495-503 (1992).

15. Guinier, A.a.F., G. *Small Angle Scattering of X-rays*, (John Wiley, New York, 1955).

16. Svergun, D.I., Petoukhov, M.V. & Koch, M.H. Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80, 2946-53 (2001).

17. Volkov, V. & Svergun, D. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* 36, 860-864 (2003).

18. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. & Svergun, D.I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* 36, 1277-1282 (2003).

19. Lebowitz, J., Lewis, M.S. & Schuck, P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* 11, 2067-79 (2002).

20. Laue, T.M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. . Analytical Ultracentrifugation in Biochemistry and Polymer Sciences. (ed. Harding, S.E., Rowe, A. J., Horton, J. C.) 90-125 (Royal Society of Chemistry, UK, 1992).

21. Lee, J.C. & Timasheff, S.N. Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride. *Biochemistry* 13, 257-65 (1974).

22. Schuck, P. *Diffusion-deconvoluted sedimentation coefficient distributions for the analysis of interacting and non-interacting protein mixtures*., (The Royal Society of Chemistry, Cambridge, 2005).

23. Schuck, P. Diffusion of the reaction boundary of rapidly interacting macromolecules in sedimentation velocity. *Biophys J* 98, 2741-51 (2010).

24. Rüdiger, M., Wehland, J. & Weber, K. The carboxy-terminal peptide of detyrosinated alpha tubulin provides a minimal system to study the substrate specificity of tubulin-tyrosine ligase. *Eur J Biochem* **220**, 309-20 (1994).

25. Slep, K.C. & Vale, R.D. Structural basis of microtubule plus end tracking by XMAP215, CLIP-170, and EB1. *Mol Cell* 27, 976-91 (2007).