JBC Papers in Press. Published on June 8, 2008 as Manuscript M710576200
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M710576200

ON THE SIGNIFICANCE OF TOC-GTPASE HOMODIMERS*

Patrick Koenig[1], Mislav Oreb[2], Karsten Rippe[3], Claudia Muhle-Goll[4], Irmgard Sinning[1], Enrico Schleiff[2] and Ivo Tews[1,5]

From the [1]Heidelberg University Biochemistry Center (BZH), Im Neuenheimer Feld 328, 69120 Heidelberg, Germany; [2]JWGU Frankfurt am Main, Cluster of Excellence Macromolecular Complexes, Department of Biosciences, Max-von-Laue Str. 9, 60439 Frankfurt, Germany; [3]Deutsches Krebsforschungszentrum and BIOQUANT, Research Group Genome Organization & Function, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany; [4]Institut für biologische Grenzflächen 2 (IBG 2), Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

Running title: Dimerization behavior of Toc G-Domains

[5]Address correspondence to: Ivo Tews, +49-6221-5447-85 (phone), -90 (fax), ivo.tews@bzh.uni-heidelberg.de

Precursor protein translocation across the outer chloroplast membrane depends on the action of the Toc complex, containing GTPases as recognizing receptor components. The G-domains of the GTPases are known to dimerize. In the dimeric conformation an arginine contacts the phosphate moieties of bound nucleotide *in trans*. Kinetic studies suggested that the arginine in itself does not act as an arginine finger of a reciprocal GTPase activating protein (GAP). Here we investigate the specific function of the residue in two GTPase homologues. Arginine to alanine replacement variants have significantly reduced affinities for dimerization, compared with wild-type GTPases. The amino acid exchange does not impact on the overall fold and nucleotide binding, as seen in the monomeric X-ray crystallographic structure of the *Arabidopsis* Toc33 arginine-alanine replacement variant at 2.0 Å. We probed the catalytic centre with the transition state analogue GDP/AlF_x using NMR and analytical ultracentrifugation. AlF_x binding depends on the arginine, suggesting the residue can play a role in catalysis, despite the non-GAP nature of the homodimer. Two –non exclusive– functional models are discussed: (1) the coGAP hypothesis, in which an additional factor activates the GTPase in homodimeric form; (2) the switch hypothesis, in which a protein, presumably the large Toc159 GTPase, exchanges with one of the homodimeric subunits, leading to activation.

The vast majority of GTPases serve as molecular switches that regulate various signaling and transport processes within the cell. GTPases bind and hydrolyze GTP, and the nucleotide is recognized by five loops of specific function, called G1- to G5-loop (1). Typically, GTPases have only low intrinsic GTPase rates and rely on auxiliary proteins, such as GTPase activating proteins (GAPs) and guanosine nucleotide exchange factors (GEFs) (2). Regulation of hydrolytic activity can in various ways also be achieved through dimerization of the GTPase. For the different studied cases of dimeric GTPases (3-8), differences exist with respect to interaction mode or the function of dimerization.

The small GTPases of the Toc34 type (9) and the multi-domain GTPases of the Toc159 type (9,10) are subunits of the membrane inserted Toc-complex which transports precursor proteins from the cytoplasm across the outer chloroplast envelope membrane (11-13). While Toc GTPases can homo- and hetero-dimerize *in vitro* (14-23), mechanistic models of protein import consider a Toc34/Toc159 interaction (24). Previous 3D structures show the *ps*Toc34 GTPase from *Pisum sativum* in the GDP (17) and in the GMPPNP (25) bound states as dimers. The functional homologue *at*Toc33 from *Arabidopsis thaliana* (14,26) is a monomer in both nucleotide loading states (16,25). Both GTPases can homodimerize in solution (15,16,25), but *at*Toc33 has a lower association constant compared with *ps*Toc34 (16,25,27).

It is not entirely clear how dimerization and hydrolytic activity are linked (16,18,23,25). This is surprising as the dimerization interface not only involves a number of Toc-specific insertions

(17,25) but also several G-loops that bind the nucleotide. An arginine contacting nucleotide *in trans* in dimeric GTPase complexes has been assigned a function in dimer formation (15,23), in nucleotide recognition (16) and in catalysis (16,17,23). To decipher the specific role of this residue, we studied the GTPases *ps*Toc34 and *at*Toc33 as well as arginine to alanine replacement variants *ps*Toc34^{R133A} and *at*Toc33^{R130A}. We conclude on the physiological role of Toc GTPases Toc33/34 homodimers, which are in abundance in the Toc complex (28-30).

Experimental Procedures

Cloning and protein purification - Mutants of *at*Toc33 and *ps*Toc34 were generated by PCR using *at*Toc33 (aa 1-251) (14) and *ps*Toc34 (aa 1-267) (25) as template. Constructs were cloned into pET21d (Novagen, Madison, WI, USA) to generate *at*Toc33^{R130A} and *ps*Toc34^{R133A} with C-terminal hexa-histidine tag.

Recombinant proteins were purified using nickel affinity chromatography (GE-Healthcare, Freiburg, Germany) in 50 mM Tris buffered at pH 7.4, containing 500 mM NaCl, 10 mM imidazole, 5 mM $MgCl_2$, 10 % glycerol and 0.7 mM β-mercaptoethanol as running buffer; elution buffer additionally contained 500 mM imidazole.

For crystallization, the protein was further purified by gel filtration, using a Superdex 75 prep grade 26/60 column (GE-Healthcare) with 20 mM HEPES buffered at pH 7.4, containing 150 mM KCl, 3 mM $MgCl_2$ and 0.7 mM β-mercaptoethanol as running buffer.

For analytical ultracentrifugation, *ps*Toc34 and *ps*Toc34^{R133A} were further purified after nickel purification by size exclusion chromatography using a Superdex 75 26/60 column and 20 mM Tris buffer at pH 8.5, containing 100 mM NaCl and 3 mM $MgCl_2$.

For nucleotide exchange of *ps*Toc34, buffer exchange after nickel affinity purification with 20 mM Tris at pH 8.5, containing 100 mM NaCl and 3 mM $MgCl_2$, was performed on a PD10 column (GE-Healthcare). The protein was incubated with 2 mM GMPPNP and 100 U alkaline phosphatase (New England Biolabs, Frankfurt, Germany) for 10 hours at 15 °C. A subsequent purification step by nickel affinity chromatography and size exclusion chromatography was performed to remove alkaline phosphatise.

Crystallography - *at*Toc33^{R133A} was crystallized at a concentration of 0.8 mM using sitting drop vapor diffusion and 2 µl drop size at 19 °C. Crystals were obtained within 3 days in 20% PEG 3350 with 0.2 M NH_4Cl, subsequently frozen in liquid nitrogen and stored using mother liquor containing an additional 20 % glycerol as cryo-protectant. Data were collected at ESRF, Grenoble, France, on beamline ID14-4 at a wavelength of 0.933 Å on an ADSC Quantum-q4 CCD imaging device.

Data were integrated and scaled with the HKL-software (31). Data-reduction, Free-R assignment and all further data manipulation were carried out with the CCP4 suite of programs (32). The structure was determined by molecular replacement using the program MOLREP (33) with *at*Toc33 as a search model (PDB entry 3BB3, (25)). Iterative model building and refinement were carried out with the programs "coot" (34) and REFMAC5 (35) cycled with ARP (36).

NMR spectroscopy - ^{19}F NMR spectra were measured on a DRX300 spectrometer (Bruker, Rheinstetten, Germany) operating at 270 MHz. The spectra were acquired at 25° C, using protein at a concentration of 0.5 mM in 20 mM Tris buffered at pH 7.0, containing 75 mM NaCl, 3 mM $MgCl_2$, 10 mM NaF, 1 mM $AlCl_3$ and 10 % D_2O added prior to acquisition. The spectra were referenced to external trifluoracetate. A 90 degree pulse was used with a repetition rate of 2s. 4096 free induction decays were summed up. The spectra were processed with TOPSPIN (Bruker, Germany).

Biochemical and biophysical assays - For analytical ultracentrifugation, nucleotide load of the protein sample was controlled by RP-HPLC, as described (25). A preparation of GDP loaded GTPase was split, and to one half of the preparation 10 mM NaF and 1 mM $AlCl_3$ were added. Both samples were incubated overnight at 4°C. The final protein concentration for analytical ultracentrifugation on a Beckman Optima XL-A ultracentrifuge equipped with absorbance optics and an An60 Ti rotor (Beckman Coulter, Fullerton, CA) was adjusted to 45 µM. Sedimentation velocity runs were carried out at 20 °C at 40,000

rpm, using the size exclusion buffer as described above as reference. Buffer density (1.00314 ml g^{-1}), buffer viscosity (1.002 mPa sec) as well as the partial specific volume of *ps*Toc34 based on amino acid sequence ($\bar{v}$ = 0.7410 ml g^{-1}) were calculated using the program SEDNTERP, version 1.05 (J. Philo, D. Hayes, and T. Laue, www.jphilo.mailway.com/download.htm). Sedimentation coefficients were determined from the c(s) distribution using the program SEDFIT (37,38), normalized for water and 20 °C.

Dimerization behaviour of nickel affinity purified *at*Toc33 and *at*Toc33^{R130A} was analyzed by size exclusion chromatography using a Superdex75 16/100 (GE-Healthcare) equilibrated with 20 mM Tris buffered at pH 8.5, containing 75 mM NaCl, and 3 mM $MgCl_2$. For molecular weight determination, 100 µl of nickel affinity purified protein at a concentration of approximately 0.6 mM was loaded onto a Superdex75 HR 10/300 gel filtration column (GE-Healthcare), equilibrated with 20 mM Tris buffered at pH 7, containing 75 mM NaCl, and 3 mM $MgCl_2$. Alternatively, the buffer contained 10 mM NaF and 1 mM $AlCl_3$. For in-line detection, a Mini Dawn light scattering instrument (Wyatt Technology, Santa Barbara, CA, USA) and a refractory index detector (WGE Dr. Bures, Dallgow, Germany) were used. Data were evaluated using the AstraV software (Wyatt Technology).

GTPase activities of *at*Toc33, *at*Toc33^{R130A}, *ps*Toc34 and *ps*Toc34^{R133A} were determined by a HPLC based hydrolysis assay as described previously (25). Protein was used at a concentration of ~0.8 mM in 20 mM Tris-HCl buffered at pH 8.0, containing 75 mM KCl and 5 mM $MgCl_2$.

RESULTS

Homodimerization and nucleotide recognition of small Toc-GTPases – To understand the function of dimerization of the two homologous GTPases *at*Toc33 and *ps*Toc34, we first analyzed the previously published 3D structures of *ps*Toc34 (17,25). Five G-loops bind the nucleotide, as in other GTPases (1), but in addition the G-loops G2, G3 and G4 are involved in dimerization (grey, blue and green in Figure 1A). Furthermore, the conserved box loop (CB-loop) contributes to dimerization ((25), red in Figure 1A). This sequence motif is unique to Toc and the so called Aig-like GTPases (39). Also at the dimer interface is helix α5 (green in Figure 1A).

We investigated how nucleotide recognition and dimerization are linked, since the G-loops are located at the interface. We note a change in function of the G4 loop which has lost properties of nucleotide recognition to gain properties in dimerization. Small Toc GTPases have a conserved histidine in the G4 loop (sequence motif THAQ), not present in the G4 loop of canonical small GTPases like the GTPase Ras p21 (sequence motif NKxD (17,40,41), Figure 1B). In *ps*Toc34, *ps*His163 of the G4-loop is in hydrogen bonding distance to *ps*Tyr132 in the CB-loop of the homodimerization partner *in trans* (Figure 1C). *ps*His163 further makes a π-stacking interaction with the guanine ring of the base. The Toc34 G4 loop is deprived of a central aspartate residue that in small GTPases normally would specify the nucleotide through direct interaction with N1 and the 2-amino group of the guanine base (Figure 1B-D). Consequently, in *ps*Toc34 other interactions are responsible for nucleotide recognition: the G5 glutamate contacts guanosine-N1 directly and the guanosine 2-amino group *via* one bridging water (WAT, Figure 1C). The lack of specificity explains the reported hydrolysis of XTP (42,43).

The structural impact of R130A exchange on nucleotide recognition – The CB loop not only provides *ps*Tyr132 but more importantly the two adjacent arginines, *ps*Arg128 and *ps*Arg133, for dimerization. Replacement of any of these arginines with alanine leads to abrogation of dimerization (15-17,23). A structural study carried out at 3.2 Å resolution on the *at*Toc33^{R130A} reported a monomeric structure of the GTPase (PDB entry 2J3E, (16)). Interestingly, the guanosine moiety was fitted in an unusual conformation in which the guanidine group was turned by 150° (16). Given the analysis presented above on incomplete nucleotide recognition by the G4/G5 loops, one might suppose that this unusual nucleotide conformation could occur, even though it would be in disagreement with other GTPase structures.

The earlier data prompted us to initiate a crystallographic study with the aim to collect atomic resolution data on *at*Toc33^{R130A}. The expression

construct encoded amino acids 1-251 rather than amino acids 1-256 as in the previous study (16). Further, crystallisation conditions were slightly different: we used 20% PEG 3350, 0.2 M NH_4Cl, pH 7.4 instead of 30% PEG 4000, 0.2 M ammonium acetate, 0.05 M sodium acetate, pH 4.8 with added NADH (16). The crystals diffracted to a Bragg spacing below 2 Å, using synchrotron light. We determined the structure by molecular replacement using the model of native *at*Toc33 (PDB entry 3BB3,(25)). Space group symmetry and molecular packing are identical to the earlier study on *at*Toc33^{R130A} (16), but different from wild-type *at*Toc33. At position 130, no side chain electron density is seen, consistent with the arginine-alanine exchange. All three structures are observed in the GDP bound form and are highly similar (with respective rmsd values of 0.58 Å over 231 C_α-positions and of 1.1 Å over 233 C_α-positions for comparison of *at*Toc33^{R130A} with native *at*Toc33, PDB entry 3BB3, (25), and with the earlier structure of *at*Toc33^{R130A}, PDB entry 2J3E, (16)). Importantly, structural changes reported to occur in *at*Toc33^{R130A} (16) are not confirmed.

Initial refinement without nucleotide resulted in clear negative F_{obs}-F_{calc} density for the nucleotide, as shown in Figure 2A. Compared with the earlier report (16), the syn–conformation of the glycosidic bond in the GDP molecule is not confirmed (Figure 2a, white nucleotide). Instead, the common anti-conformation is observed (Figure 2a, black nucleotide). Thus, while it was previously suggested that a change in nucleotide conformation might have occurred either by lack of dimerization or as a result of the *at*R130A exchange (16), comparison with the wild-type structure and the high resolution structure of *at*Toc33^{R130A} presented here rule out this possibility, demonstrating that nucleotide binding is unaffected.

Effect of arginine to alanine exchange on the homodimerization of Toc34 - In the structure of dimeric *ps*Toc34 an arginine (Arg133, the equivalent of *at*Arg130) contacts the β- and γ-phosphate groups *in trans*. This interaction is suggestive of a function as arginine finger, often found in GAP-GTPase interactions (Figure 1C) (44). This has led to the proposal that the *ps*Toc34 homodimer could be a self activating GAP-complex (16,17). Furthermore, this interaction has been described to be pivotal for dimerization (15,16,23).

We established the dimerization properties of *at*Toc33 and of *at*Toc33^{R130A} using size exclusion chromatography (Figure 2B). As the chromatogram shows, both proteins dimerize. However, the dimerization behavior of *at*Toc33^{R130A} is impaired. The difference in migration behaviour of both, monomeric and dimeric species, can be explained by a presumed difference in dimerization behaviour: a fast dimerization equilibrium in *at*Toc33 would lead to a decreased apparent size of the dimer; in turn, it would increase the apparent size of the monomer. This is supported by static light scattering data, given below. While dimeric and monomeric species don't baseline separate for *at*Toc33, *at*Toc33^{R130A} is different and shows baseline separation.

Effect of arginine to alanine exchange on the hydrolysis rate of Toc34 – Previously, it was shown in multiple turnover assays that arginine-alanine replacement impacts on GTP hydrolysis. *at*Toc33^{R130A} shows a minor reduction in hydrolysis rate compared with *at*Toc33 (15,16), and *ps*Toc34^{R133A} shows a loss of hydrolytic activity (23).

We investigated the effect of an arginine to alanine replacement on GTP hydrolysis using an HPLC based single turnover assay for determination of enzymatic rates (45). The assay was carried out at higher concentrations than the multiple turnover experiments, allowing for dimer formation of wild type proteins (see Experimental Procedures for details). *at*Toc33^{R130A} showed 0.6-fold hydrolytic activity compared with wild-type protein (k_{cat} = 2.9 x 10^{-5} s^{-1} for *at*Toc33^{R130A} and k_{cat} = 4.4 x 10^{-5} s^{-1} for wild-type *at*Toc33). Similarly, *ps*Toc34^{R133A} showed 0.3-fold hydrolytic activity of the wild-type GTPase (k_{cat} = 2.4 x 10^{-5} s^{-1} for *ps*Toc34^{R133A} and k_{cat} = 8.4 x 10^{-5} s^{-1} for wild-type protein). This establishes for one that *ps*Toc34^{R133A} possesses hydrolytic activity and on the other hand demonstrates that the arginine-alanine exchange is only of minor influence on GTP hydrolysis in either GTPase. Apparent differences in determined hydrolysis rates with earlier reports (23) are likely explained by the different experimental setup of single and multiple turnover measurements.

Binding of AlF_x to psToc34 – Since exchange of *at*Arg130/*ps*Arg133 with alanine has only limited influence on the GTP hydrolysis rate, we tested whether the respective arginines can act at all as arginine fingers, employing aluminum fluoride as a probe. Aluminum fluoride exists as an equilibrium of different species in solution, and is thus abbreviated here as AlF_x. AlF_x can act as a transition state mimicry of phoshoryl-transfer reactions (46) and has been shown to directly bind to the Gα proteins (47) that contain an intrinsic domain for stimulation of catalysis. Intrinsic stimulatory domains are absent in small GTPases such as Ras or Toc34, and they instead require a GAP for activation. For Ras_{GDP}, AlF_x binding depends on the presence of the RasGAP proteins (48).

GTPase/GAP/AlF_x complexes show AlF_x binding in the active site in place of the γ-phosphate, and they thus require GDP loaded GTPase subunits. The GAP arginine finger is often present as binding partner. If *ps*Arg133 would act as an arginine finger in the *ps*Toc34 dimer, AlF_x binding to *ps*Toc34 can be expected.

AlF_x binding to *ps*Toc34 was tested using ^{19}F-NMR. A buffer solution containing $AlCl_3$ and NaF shows peaks at -77.0 ppm and at -41.7 ppm, corresponding to AlF_x and free F^- (Figure 3A). After addition of *ps*$Toc34_{GDP}$, a peak shifted by -24.5 ppm from the resonance signal of free F^-, is observed at -66.2 ppm (Figure 3D). The chemical shift, varying between -20 ppm and -22.4 ppm in previous studies (47,49,50), is indicative of AlF_x binding to nucleotide binding proteins, and has been described before for *ps*Toc34 (17). To verify the specificity of the interaction in the catalytic centre, *ps*Toc34 loaded with non-hydrolysable GTP analogue GMPPNP was used. Since the binding site is occupied by the γ-phosphate of GMPPNP, specific binding of AlF_x to the γ-phosphate site can be excluded. Indeed, no binding of AlF_x is detected when the GMPPNP loaded GTPase is investigated (Figure 3E).

The effect of AlF_x on psToc34 dimerization – We next tested the stability of the *ps*Toc34/GDP/AlF_x complex. Size exclusion chromatography of *ps*$Toc34_{GDP}$ in the presence of AlF_x in the buffer indicated stabilization of the dimer, as evidenced by a shift to a higher molecular weight species (data not shown).

To quantify AlF_x induced oligomerization, analytical ultracentrifugation was employed. Two samples were compared, distinguished by presence of AlF_x. Without AlF_x treatment, *ps*Toc34 is present in monomer-dimer equilibrium (Figure 4A, (23,25)). With AlF_x treatment, *ps*Toc34 was exclusively dimeric (Figure 4A). Thus, addition of AlF_x leads to stabilization of the *ps*Toc34 homodimer as reported for classical GTPase-GAP interactions like Ras-RasGAP (48).

When analytical ultracentrifugation was repeated with *ps*$Toc34^{R133A}$, no dimeric species was observed, regardless of AlF_x treatment (Figure 4B). This demonstrates that the effects seen before with wild-type *ps*Toc34 are specific and require the presence of *ps*Arg133, in line with the NMR data (Figure 3).

Influence of AlF_x on the dimerization of atToc33 – We then assayed the effect of AlF_x binding on the dimerization behavior of *at*Toc33. Since the protein exhibits a lower K_a for dimerization, analytical ultracentrifugation is impractical due to the high protein concentrations that would be required. Instead, we employed a setup where size exclusion chromatography was coupled with static light scattering and a refractive index detector to determine absolute molecular weights. This method does not require use of internal standards (51).

Similar to previous reports (15-17,23), two molecular species were observed for wild-type *at*Toc33 in the absence of AlF_x (Figure 5A). The analysis of static light scattering gave a signal yielding a molecular weight of 63 kDa for protein fractions in the first peak, which fits well with the value of 60 kDa for an *at*Toc33 dimer. However, protein fractions of the second peak displayed a molecular weight of 45 kDa; he tail of this second peak was fitted with a molecular weight of 36 kDa, likely to represent the monomer. The 45 kDa species likely results from a dynamic equilibrium between dimeric and monomeric species. Thus, *at*Toc33 exists as a fast equilibrium between the two states (compare Figure 2B). This is consistent with data on *ps*Toc34 (23). When *at*$Toc33^{R130A}$ was investigated, a single peak fitted to a molecular weight of 30-31 kDa is obtained, corresponding to monomeric protein (Figure 5B).

When AlF_x was present in the buffer (Figure 5C), *at*Toc33 shifted to a higher molecular weight

species. Light scattering data were fitted to a molecular weight of 66 kDa, corresponding to the molecular weight of dimeric *at*Toc33 (60 kDa). Thus, stabilization of the dimeric species occurs with *at*Toc33 upon addition of AlF_x. However addition of AlF_x does not affect *at*Toc33^{R130A} which remains also in the presents of AlF_x monomeric (Figure 5D). Light scattering data were fitted to a molecular weight of 25 kDa.

The dimerization behaviour of *at*Toc33 (Figure 5) and of *ps*Toc34 (Figure 5) are similar. The dimeric state of both proteins is stabilized by AlF_x. With the arginine-alanine exchange mutants, it can be shown that binding is specific, since stabilization of the dimer requires the presence of *at*Arg130/*ps*Arg133.

DISCUSSION

Dimerization of Toc GTPases is generally assumed to be a feature of the assembly of the Toc apparatus (24). GTPase dimerization is recurrent, and the Toc GTPases are thus not exceptional in this respect. Documented examples of dimeric GTPases are for instance the SRP GTPases FtsY, Ffh (4,5) and FlhF (3), the GTPases of the dynamin type, e.g. hGBP (7), the GTPase MnmE involved in tRNA modification (8) and the metal-binding GTPase HyB (6). However, the dimerization interface is different between these GTPases, and so is the functional significance of dimerization of these various GTPases.

The isolated G-domains of *ps*Toc34 and *at*Toc33 both dimerize, but they differ with respect to their dimerization properties (16,25). The K_d of the *at*Toc33 dimer is about one order of magnitude higher than that for *ps*Toc34. Both K_ds are in the sub-millimolar range (25,27). While these figures, determined for the isolated GTPase, seem high, dimerization may still occur in the physiological context on the membrane or within the Toc complex through elevated local concentrations. Interaction may be helped by the C-terminal membrane anchor, not present in the protein analyzed here.

The dimerization interface itself is preserved between different Toc34 GTPases (16) and involves the CB-motif (25) as well as the G4 / G5 loops, with G4 performing a dual role in nucleotide recognition and dimerization (Figure 1C). The CB motif carries the arginine that contacts the nucleotide *in trans*, the function of which is controversially discussed with respect to dimerization, nucleotide binding and catalysis (15-17,23,25). The function of this arginine requires thus further clarification, since it is the key to elucidate the task of the Toc34 homodimer.

The role of *ps*Arg133/*at*Arg130 in dimerization was previously investigated using a variety of techniques, including native PAGE, analytical ultracentrifugation and size exclusion (15,16,23). It was shown that *at*Toc33^{R130A} and *ps*Toc34^{R133A} are unable to dimerize. In contrast, we show that *at*Toc33^{R130A} forms dimers at high protein concentrations using size exclusion chromatography (Figure 2B, Figure 5A,C). This suggests that Arg130 in *at*Toc33 is a key –but not the sole– player in homodimerization.

A function of *ps*Arg133/*at*Arg130 in nucleotide recognition was suggested on the basis of the previous 3.2 Å structure of *at*Toc33^{R130A} (16) in a monomeric state with an unusual nucleotide conformation. However, the limited resolution and the lack of a wild-type reference GTPase structure did not allow concluding whether the amino acid exchange directly affected the structure, or whether the effect was indirect and caused by the lack of dimerization. Based on the 2 Å resolution structure presented here, an altered nucleotide conformation can be excluded. This is also evident from superposition with the now available monomeric wild-type *at*Toc33 structure (25).

Finally, participation of *ps*Arg133 in catalysis was previously proposed on the basis of the crystal structure of *ps*Toc34 that showed this residue in a conformation similar to the classic GAP arginine finger (17). The transition state mimic GDP/AlF_x binds to the *ps*Toc34 homodimer but not to the the *ps*Arg133 mutant or of the GMPPNP loaded GTPase (Figure 3). This implies that AlF_x indeed acts as a transition state mimicry, demonstrating the arginine is in an appropriate conformation to act during catalysis. AlF_x binding also stabilizes the homodimers in *ps*Toc34 (Figure 4) and in *at*Toc33 (Figure 5). AlF_x induced dimerization suggests a composite binding site formed by both dimerization partners, involving *ps*Arg133 / *at*Arg130.

Summing up from our studies on *ps*Arg133/*at*Arg130 and the literature, the following evidences doubt or directly contradict that small Toc GTPases form self associating GAP-complexes (17): (i) While it would be expected that a GAP complex favours the GTP state, it is observed that GMPPNP and GDP loaded states of the GTPase both dimerize with similar efficiency (17,25). (ii) Despite the stabilization of G-loops in the dimerization interface, switch I retains some flexibility which would be unexpected for a GAP complex (25). (iii) Despite the presence of the stabilizing arginine, reminiscent of the classic GAP arginine finger, the catalytic center is incomplete as no residue for the positioning of the catalytic water is present (25); instead the catalytic site is accessible for solvent by a short tunnel. (iv) Kinetic data argue against the formation of a GAP-like complex: the acceleration of hydrolysis in GTPase-GAP complexes is typically in the order of 2-5 magnitudes (44). However, no significant catalytic activation is observed after dimer formation (23,25). This is supported by data from arginine-alanine exchange proteins that show that mutation of *ps*Arg133/*at*Arg130 have only a minor effect on the hydrolysis rate of Toc33/34.

Thus, small Toc GTPases represent a paradox since *ps*Arg133/*at*Arg130 seems poised properly to act as an arginine finger, similar to that of a GAP, but hydrolysis rates are not accelerated. Acceleration of hydrolysis in Ras-like GTPases and their respective GAPs is mainly due to the positioning of a catalytic residue to polarize a water for hydrolytic attack (52). For instance, replacement of this crucial residue in a Ras:RasGAP system leads to abrogation of hydrolysis, even when an arginine finger interaction is present (53). Our structural analysis of the Toc34 homodimer shows that despite presence of *ps*Arg133 or *at*Arg130 the catalytic machinery remains incomplete, and an essential catalytic residue for the positioning of the catalytic water is required, explaining the minor effect of dimerization on hydrolysis rates (25).

We have previously shown that the nucleotide binding pocket is accessible in the *ps*Toc34$_{GMPPNP}$ homodimer, and suggested a catalytic residue could be inserted into the catalytic centre. This would functionally define the homodimer as a coGAP complex that requires both, the homodimeric interaction and a third protein (coGAP hypothesis) (25). In addition, the third protein may also be required to stabilize and organize the catalytic centre, then giving it its true GTPase/GAP type character. *ps*Arg133/*at*Arg130 thus would fulfill the role of an arginine finger in catalysis only if the coGAP is present.

In a second proposal, the homodimer has to dissociate to become functional in the physiological context (switch hypothesis). Catalytic data suggest that the monomeric species requires interaction with another protein for activation. The large GTPase Toc159 is an obvious candidate for this interaction. Toc159 can supply an arginine, similar to what is seen in the homodimer but probably with slightly different geometry (25). The heterodimer is thus not only asymmetric but also different from the homodimer.

We conclude there may well be a physiological role for the small Toc33/34 GTPase homodimer, complementing the postulated heterodimer of small and large Toc GTPases. This is further supported by the stoichiometry of the Toc complex, where small GTPase subunits are in molar excess over large GTPase subunits. Hence, two differentially regulated events in the Toc mediated chloroplast protein import cycle would exist.

REFERENCES

1. Sprang, S. R. (1997) *Annu. Rev. Biochem.* **66**, 639-678
2. Vetter, I. R., and Wittinghofer, A. (2001) *Science* **294**(5545), 1299-1304
3. Bange, G., Petzold, G., Wild, K., Parlitz, R. O., and Sinning, I. (2007) *Proc. Natl. Acad. Sci. U S A* **104**(34), 13621-13625
4. Egea, P. F., Shan, S. O., Napetschnig, J., Savage, D. F., Walter, P., and Stroud, R. M. (2004) *Nature* **427**(6971), 215-221

5. Focia, P. J., Shepotinovskaya, I. V., Seidler, J. A., and Freymann, D. M. (2004) *Science* **303**(5656), 373-377

6. Gasper, R., Scrima, A., and Wittinghofer, A. (2006) *J. Biol. Chem.* **281**(37), 27492-27502

7. Prakash, B., Praefcke, G. J., Renault, L., Wittinghofer, A., and Herrmann, C. (2000) *Nature* **403**(6769), 567-571

8. Scrima, A., and Wittinghofer, A. (2006) *Embo J.* **25**(12), 2940-2951

9. Kessler, F., Blobel, G., Patel, H. A., and Schnell, D. (1994) *Science* **266**(5187), 1035-1039

10. Chen, K., Chen, X., and Schnell, D. J. (2000) *Plant Physiol.* **122**(3), 813-822

11. Jarvis, P., and Robinson, C. (2004) *Curr. Biol.* **14**(24), R1064-1077

12. Kessler, F., and Schnell, D. J. (2004) *Trends Cell. Biol.* **14**(7), 334-338

13. Soll, J., and Schleiff, E. (2004) *Nat. Rev. Mol. Cell Biol.* **5**(3), 198-208

14. Jelic, M., Soll, J., and Schleiff, E. (2003) *Biochemistry* **42**(19), 5906-5916

15. Weibel, P., Hiltbrunner, A., Brand, L., and Kessler, F. (2003) *J. Biol. Chem.* **278**(39), 37321-37329

16. Yeh, Y. H., Kesavulu, M. M., Li, H. M., Wu, S. Z., Sun, Y. J., Konozy, E. H., and Hsiao, C. D. (2007) *J. Biol. Chem.* **282**(18), 13845-13853

17. Sun, Y. J., Forouhar, F., Li Hm, H. M., Tu, S. L., Yeh, Y. H., Kao, S., Shr, H. L., Chou, C. C., Chen, C., and Hsiao, C. D. (2002) *Nat. Struct. Biol.* **9**(2), 95-100

18. Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., and Schleiff, E. (2004) *Embo J.* **23**(3), 520-530

19. Smith, M. D., Hiltbrunner, A., Kessler, F., and Schnell, D. J. (2002) *J. Cell Biol.* **159**(5), 833-843

20. Wallas, T. R., Smith, M. D., Sanchez-Nieto, S., and Schnell, D. J. (2003) *J. Biol. Chem.* **278**(45), 44289-44297

21. Hiltbrunner, A., Bauer, J., Vidi, P. A., Infanger, S., Weibel, P., Hohwy, M., and Kessler, F. (2001) *J. Cell Biol.* **154**(2), 309-316

22. Ivanova, Y., Smith, M. D., Chen, K., and Schnell, D. J. (2004) *Mol. Biol. Cell* **15**(7), 3379-3392

23. Reddick, L. E., Vaughn, M. D., Wright, S. J., Campbell, I. M., and Bruce, B. D. (2007) *J. Biol. Chem.* **282**(15), 11410-11426

24. Oreb, M., Tews, I., and Schleiff, E. (2008) *Trends Cell. Biol.* **18**(1), 19-27

25. Koenig, P., Oreb, M., Höfle, A., Kaltofen, S., Rippe, K., Sinning, I., Schleiff, E., and Tews, I. (2008) *Structure* **16**(4), 585-596

26. Jarvis, P., Chen, L. J., Li, H., Peto, C. A., Fankhauser, C., and Chory, J. (1998) *Science* **282**(5386), 100-103

27. Oreb, M., Höfle, A., Mirus, O., and Schleiff, E. (2008) *J. Exp. Bot. in press*

28. Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R. (2003) *J. Cell Biol.* **160**(4), 541-551

29. Kikuchi, S., Hirohashi, T., and Nakai, M. (2006) *Plant Cell Physiol.* **47**(3), 363-371

30. Vojta, A., Alavi, M., Becker, T., Hormann, F., Kuchler, M., Soll, J., Thomson, R., and Schleiff, E. (2004) *J. Biol. Chem.* **279**(20), 21401-21405

31. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In. *Macromolecular Crystallography, Pt A*

32. Collaborative Computational Project, N. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**(Pt 5), 760-763

33. Vagin, A. A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022-1025

34. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**(Pt 12 Pt 1), 2126-2132

35. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53**(Pt 3), 240-255

36. Lamzin, V. S., and Wilson, K. S. (1997) Automated refinement for protein crystallography. In. *Macromolecular Crystallography, Pt B*

37. Dam, J., and Schuck, P. (2004) *Methods Enzymol.* **384**, 185-212

38. Schuck, P. (2000) *Biophys. J.* **78**(3), 1606-1619

39. Nitta, T., and Takahama, Y. (2007) *Trends Immunol.* **28**(2), 58-65

40. Bourne, H. R., Sanders, D. A., and McCormick, F. (1991) *Nature* **349**(6305), 117-127

41. Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) *J. Mol. Biol.* **317**(1), 41-72

42. Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., and Schleiff, E. (2002) *Biol. Chem.* **383**(12), 1875-1883

43. Aronsson, H., Combe, J., and Jarvis, P. (2003) *FEBS Lett.* **544**(1-3), 79-85

44. Scheffzek, K., and Ahmadian, M. R. (2005) *Cell. Mol. Life Sci.* **62**(24), 3014-3038

45. Hemsath, L., and Ahmadian, M. R. (2005) *Methods* **37**(2), 173-182

46. Wittinghofer, A. (1997) *Curr. Biol.* **7**(11), R682-685

47. Higashijima, T., Graziano, M. P., Suga, H., Kainosho, M., and Gilman, A. G. (1991) *J. Biol. Chem.* **266**(6), 3396-3401

48. Ahmadian, M. R., Mittal, R., Hall, A., and Wittinghofer, A. (1997) *FEBS Lett.* **408**(3), 315-318

49. Maruta, S., Henry, G. D., Sykes, B. D., and Ikebe, M. (1993) *J. Biol. Chem.* **268**(10), 7093-7100

50. Praefcke, G. J., Geyer, M., Schwemmle, M., Robert Kalbitzer, H., and Herrmann, C. (1999) *J. Mol. Biol.* **292**(2), 321-332

51. Margarit, S. M., Sondermann, H., Hall, B. E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., and Kuriyan, J. (2003) *Cell* **112**(5), 685-695

52. Scrima, A., Thomas, C., Deaconescu, D., and Wittinghofer, A. (2008) *Embo J.*

53. Bollag, G., and McCormick, F. (1991) *Nature* **351**(6327), 576-579

54. Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W., and Wittinghofer, A. (1990) *Embo J.* **9**(8), 2351-2359

FOOTNOTES

* We thank Anja Höfle and Broder Schmidt for experimental support and Jens Radzimanowski, Klemens Wild and the staff at the ESRF synchrotron for help with data collection. P.K. is supported by a fellowship of the interdisciplinary PhD program "Molecular machines: mechanisms and functional inter-

connections" of the Land Baden-Württemberg to I.S. The work was supported by grants from the Deutsche Forschungsgemeinschaft (SFB594-B11) to E.S. and by the Volkswagenstiftung to E.S. and K.R.

Abbreviations used: *at*, *Arabidopsis thaliana*; GAP, GTPase activating protein; GEF, guanosine nucleotide exchange factor; *ps*, *Pisum sativum*; Toc/Tic, translocon at the outer/inner chloroplastic envelope membrane.

FIGURE LEGENDS

Figure 1: Analysis of dimerization specific features of the GTPase *ps*Toc34 from *Pisum sativum* (GMMPNP bound state, PDB entry 3BB4 (25)). *A.* Left hand side: view onto the dimerization face of a single monomer. Right hand side: in the crystal structure, one monomer, in grey, interacts with a second monomer, in white. Between the two views, the grey monomer is turned by 90 ° around a vertical axis. The molecules are shown as surface representations; critical elements involved in dimerization are visualized in color and shown as cartoon (G2: grey, G3: blue, G4: green and CB-loop: red). Residues discussed in the text are numbered and shown in stick representation. *B.* Alignment of the G4 loop region of Toc34 GTPases with Ras p21, a representative of canonical small GTPases. Sequences used are: atToc33 *Arabidopsis thaliana* Toc33 NP_171730; atToc34 *Arabidopsis thaliana* Toc34 NP_196119; Bnap1 *Brassica Nappus* Toc33 AAQ17548; Mtru *Medicago truncatula* Toc34 gb ABD28666.1; Oluc *Ostreococcus lucimarinus* predicted small Toc GTPase CCE9901 XP_001417009.1; Otau *Ostreococcus tauri* Toc34 emb CAL53037.1; Ovio *Orychophragmus violaceus* Toc33-like protein gb AAM77647.1; Ppat1 *Physcomitrella patens* Toc34-1 gb AAS47580.1; Ppat2 *Physcomitrella patens* Toc34-2 gb AAS47581.1; Ppat3 *Physcomitrella patens* Toc34-3 gb AAS47582.1; psToc34 *Pisum sativum* Toc34 Q41009; Ptri1 *Populus trichocarpa* small Toc GTPase LG_XIV0229; Ptri2 *Populus trichocarpa* small Toc GTPase LG_II1667; Stub *Solanum tuberosum* GTPbinding-like protein gb ABB16976.1; Vvin *Vitis vinifera* hypothetical protein emb CAN63847.1; Zmay1 *Zea mays* Toc34-1 emb CAB65537.1; Zmay2 *Zea mays* Toc34-2 emb CAB77551.1. hsRas/p21 *Homo sapiens* H-Ras p21 P01112. *C.* The G4 and G5 loops of *ps*Toc34 interact with the nucleotide. The 2-amido group of the nucleotide is only in indirect contact with the protein via a water molecule (WAT). Thus guanosine and xanthosine nucleotides cannot be distinguished. *ps*His163 makes a hydrogen bonding contact with *ps*Tyr132' of the CB-loop of the interacting homodimerization partner. Also shown is *ps*Arg133', interacting with phosphate moieties of the dimerization partner. *D.* A similar representation as in (C) for the Ras p21 protein (GMMPNP bound state, PDB entry 5p21, (54)). The conserved Asp119 in the G4 loop recognizes GTP specifically by interacting with N1 and the 2-amide.

Figure 2: *A* Conformation of GDP in the *at*Toc33^{R130A} structure, with the GDP molecule shown in black. Difference density (F_{obs}-F_{calc}) obtained after structure refinement without nucleotide is shown as red mesh. An altered GDP conformation was described previously (white GDP molecule, PDB entry 2J3E, (16)). *B* Size exclusion chromatography of *at*Toc33 (solid line) and *at*Toc33^{R130A} (dashed line), using a Superdex75 16/100 column. Peak fractions were analyzed by SDS-PAGE (inset; molecular weight marker – lane (M) from top to button: 200, 150, 120, 100, 85, 70, 60, 50, 40, 30, 25, 20, 15 and 10 kDa).

Figure 3: Binding of AlF_x to *ps*Toc34 by ^{19}F-NMR. *A.* ^{19}F-NMR spectrum of buffer containing $AlCl_3$ and NaF shows two peaks (-77.0 ppm and -41.7 ppm) which has been assigned to free F^- and to AlF_x. *B, C.* The addition of nucleotides (GDP, GMPPNP) shows no change compared with sample containing only buffer. *D.* Addition of *ps*Toc34$_{GDP}$ to buffer containing $AlCl_3$ and NaF shows appearance of an additional peak at -66.2 ppm *E.* Addition of *ps*Toc34$_{GMPPNP}$ to buffer containing $AlCl_3$ and NaF shows the two peaks representative of free F^- and AlF_x (-77.0 ppm and -41.7 ppm).

Figure 4: Analysis of dimerization properties of *ps*Toc34 by analytical in the presence (dashed line) and absence of AlF_x (solid line). The c(s) sedimentation coefficient distribution is shown. *A.* Wild-type

*ps*Toc34 protein with peaks at 2.7 S, corresponding to the monomeric protein species, and at 3.6 S, corresponding to the dimeric protein species. *B. ps*Toc34^{R133A} with a single peak at 2.4 S, corresponding to the monomeric protein species.

Figure 5: Effect of AlF_x on the homodimerization of *at*Toc33, using size exclusion chromatography on a Superdex75 HR 10/300 column with UV-detection (dashed line, left hand scale) and in-line static light scattering (solid line, right hand scale). Areas averaged for size determination are indicated by vertical lines and annotated with the fitted molecular weights, as indicated by the gray triangle. *A. at*Toc33. *B. at*Toc33^{R130A}. *C. at*Toc33 in the presence of AlF_x. *D. at*Toc33^{R130A} in the presence of AlF_x.

TABLES

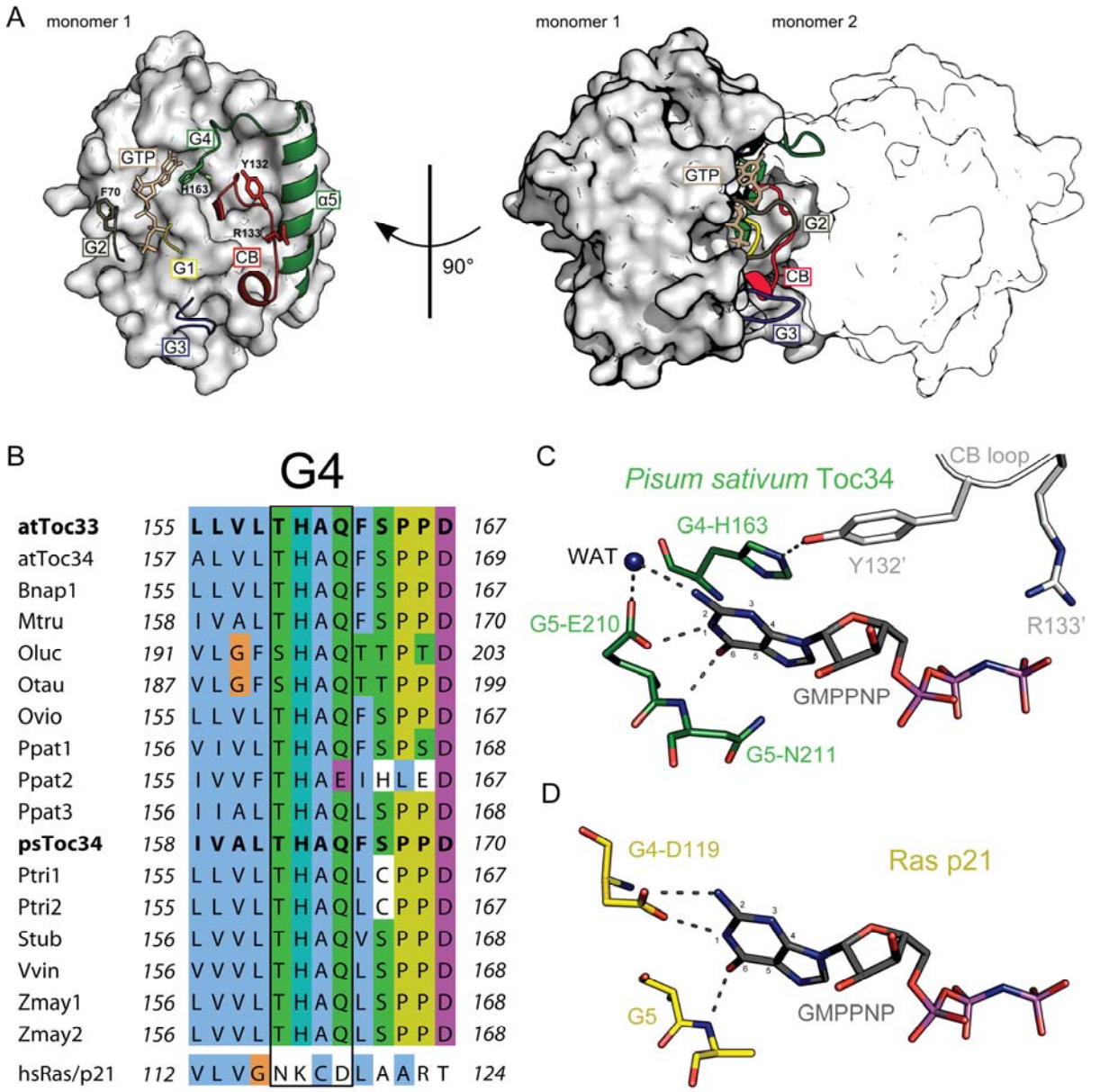
Table 1: Crystallographic analysis

	*at*Toc33^{R130A}:GDP
Space group	$P4_32_12$
Unit cell a=b, c (Å)	71.44, 112.46
Solvent content (%)	43
# mol in AU	1
Resolution (Å)	30.00 – 1.96
Average B (Å^2)	25
Unique reflections	21984
Mosaicity (°)	0.92
R_{sym} (%)*	3.7
Completeness (%)	97.5
<I> / <sigI>	37.2
Redundancy	6.9
HR shell (Å)	1.99 - 1.96
HR R_{sym} (%)*	28.4
HR Completeness (%)	96.5
HR <I> / <sigI>	4.4
Redundancy	6.2
Amino acids	2-68, 72-250
Total protein atoms (including double conformations)	2175
Water	212
ligand atoms	GDP, Mg^{2+}
RMSD bonds (Å)	0.017
RMSD angles (°)	1.598
R_{free} (%)‡	24.07
R_{work} (%)†	19.60

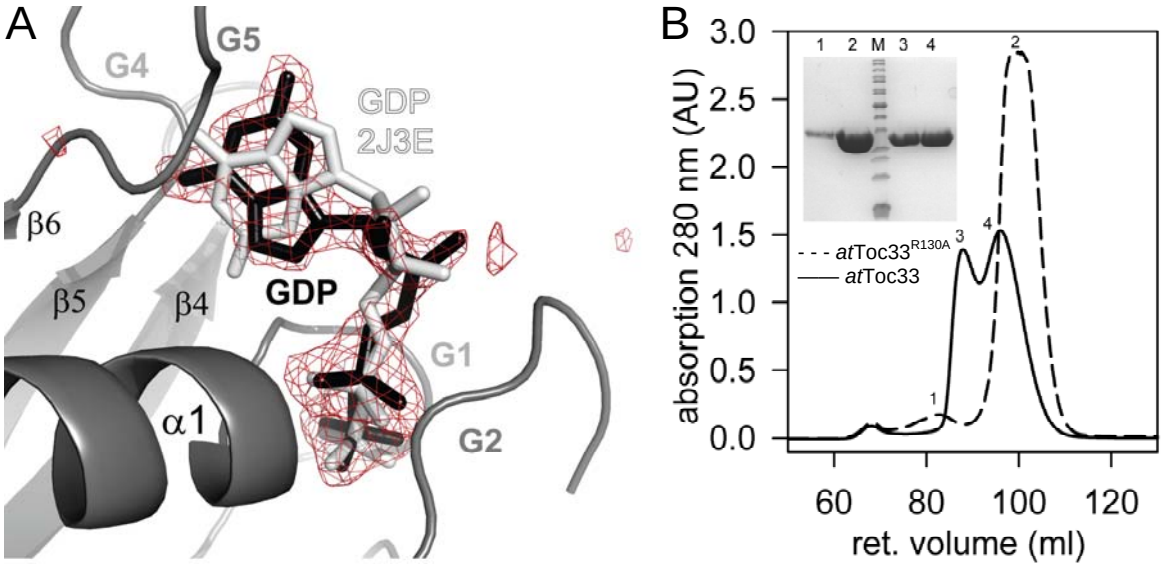
*$R_{sym} = \Sigma_h\Sigma_i |I(h) - I(h)_i|/\Sigma_h\Sigma_i I(h)_i$, where I(h) is the mean intensity

†$R_{work} = \Sigma_h||F_{obs}(h)| - | F_{calc}(h)||/ \Sigma h|F_{obs}(h)|$, where $F_{obs}(h)$ and $F_{calc}(h)$ are observed and calculated structure factors, respectively

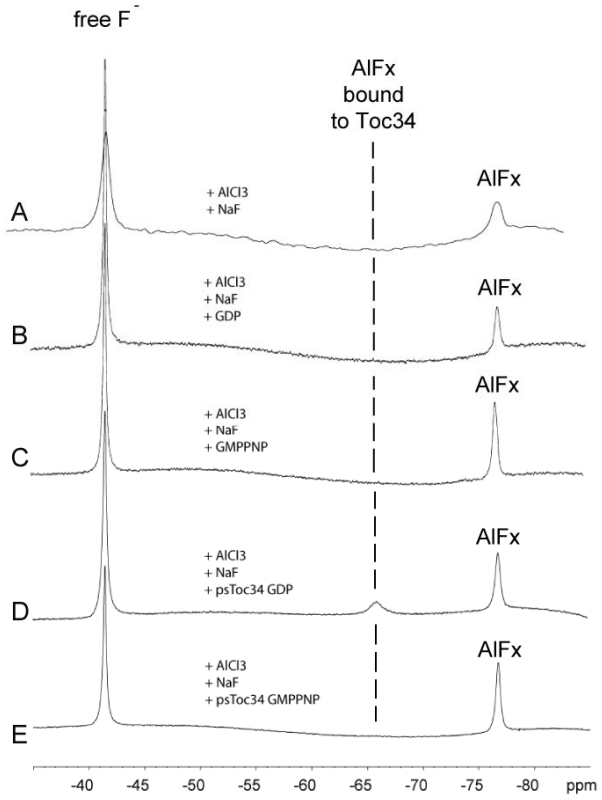
‡5 % of the data were excluded to calculate R_{free}



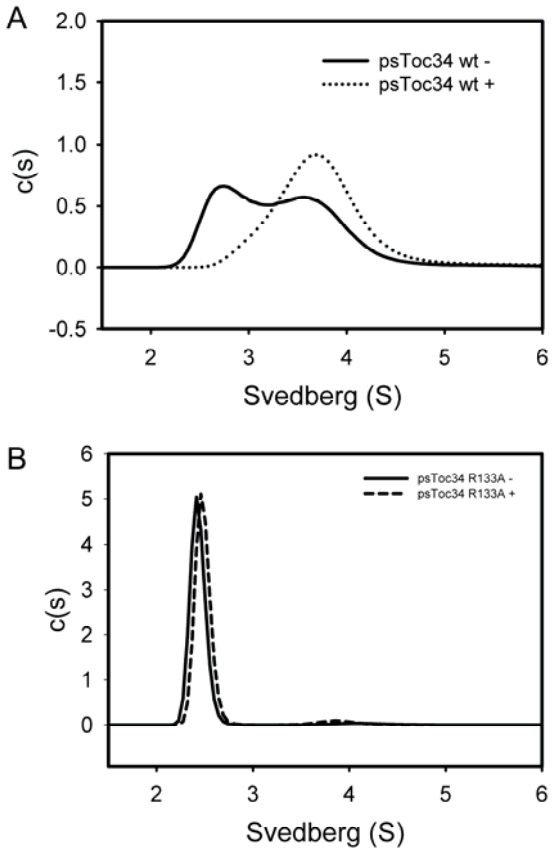
Koenig *et al.*, Figure 1



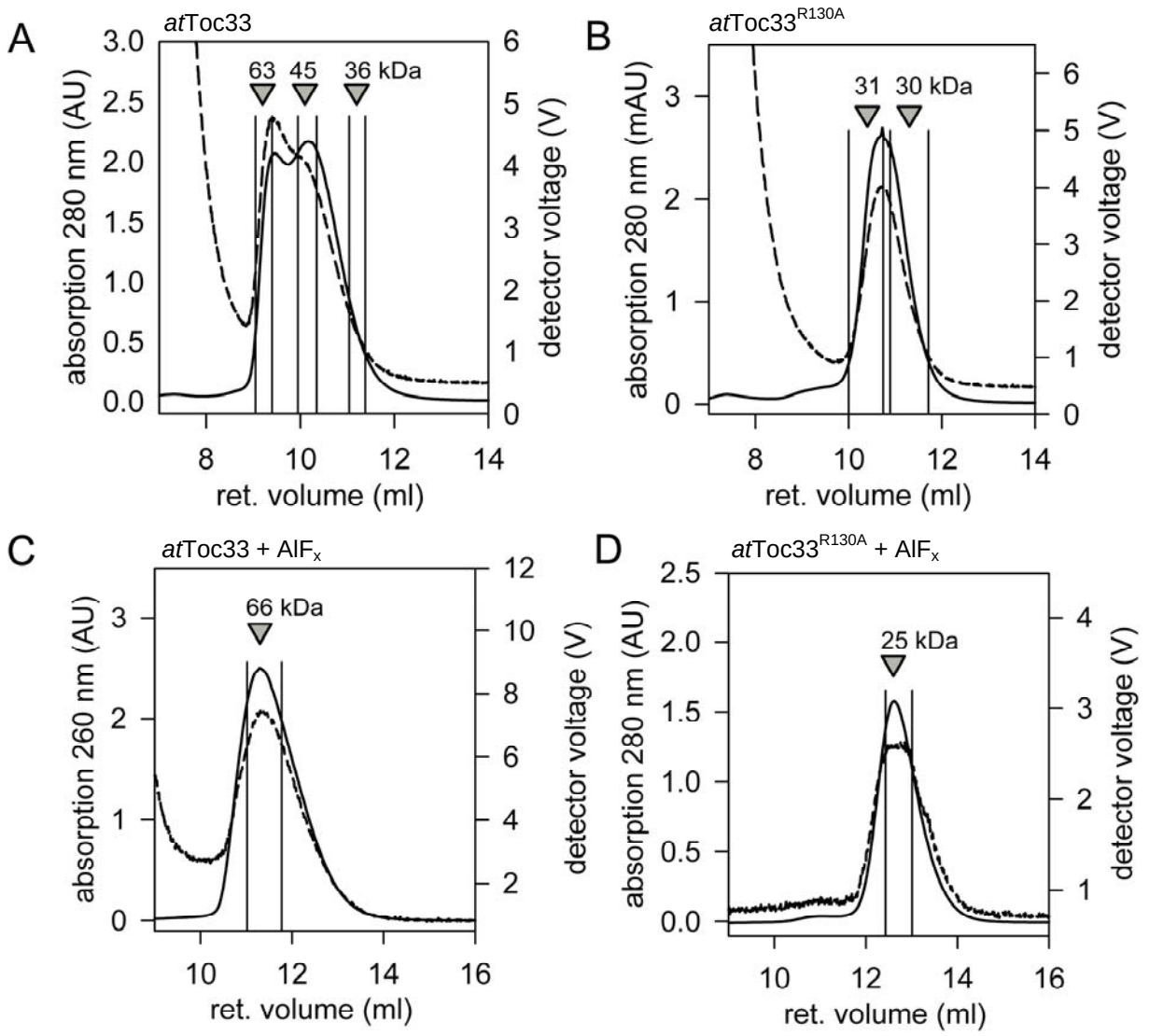
Koenig et al. Figure 2



Koenig *et al.*, Figure 3



Koenig *et al.*, Figure 4



Koenig *et al.*, Figure 5