



**HAL**  
open science

## NMR secondary structure and interactions of recombinant human MOZART1 protein, a component of the gamma-tubulin complex

Cyprian D Cukier, Audrey Tourdes, Dounia El-Mazouni, Valérie Guillet, Julian Nomme, Lionel Mourey, Alain Milon, Andreas Merdes, Virginie Gervais

► **To cite this version:**

Cyprian D Cukier, Audrey Tourdes, Dounia El-Mazouni, Valérie Guillet, Julian Nomme, et al.. NMR secondary structure and interactions of recombinant human MOZART1 protein, a component of the gamma-tubulin complex. *Protein Science*, 2017, 26 (11), pp.2240 - 2248. 10.1002/pro.3282 . hal-03002130

**HAL Id: hal-03002130**

**<https://hal.science/hal-03002130>**

Submitted on 20 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# NMR secondary structure and interactions of recombinant human MOZART1 protein, a component of the gamma-tubulin complex

Cyprian D. Cukier,<sup>1</sup> Audrey Tourdes,<sup>1</sup> Dounia El-Mazouni,<sup>1</sup> Valérie Guillet,<sup>1</sup> Julian Nomme,<sup>1</sup> Lionel Mourey,<sup>1</sup> Alain Milon,<sup>1</sup> Andreas Merdes,<sup>2\*</sup> and Virginie Gervais<sup>1\*</sup>

<sup>1</sup>Institut de Pharmacologie et Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, Toulouse, France

<sup>2</sup>Centre de Biologie du Développement, Université de Toulouse, CNRS, UPS, Toulouse, France

Received 19 June 2017; Accepted 17 August 2017

DOI: 10.1002/pro.3282

Published online 29 August 2017 proteinscience.org

**Abstract:** Mitotic-spindle organizing protein associated with a ring of  $\gamma$ -tubulin 1 (MOZART1) is an 8.5 kDa protein linked to regulation of  $\gamma$ -tubulin ring complexes ( $\gamma$ TuRCs), which are involved in nucleation of microtubules. Despite its small size, MOZART1 represents a challenging target for detailed characterization *in vitro*. We described herein a protocol for efficient production of recombinant human MOZART1 in *Escherichia coli* and assessed the properties of the purified protein using a combination of size exclusion chromatography coupled with multiangle light scattering (SEC-MALS), dynamic light scattering (DLS), and nuclear magnetic resonance (NMR) experiments. MOZART1 forms heterogeneous oligomers in solution. We identified optimal detergent and buffer conditions for recording well resolved NMR experiments allowing nearly full protein assignment and identification of three distinct alpha-helical structured regions. Finally, using NMR, we showed that MOZART1 interacts with the N-terminus (residues 1–250) of GCP3 ( $\gamma$ -tubulin complex protein 3). Our data illustrate the capacity of MOZART1 to form oligomers, promoting multiple contacts with a subset of protein partners in the context of microtubule nucleation.

**Keywords:** MOZART1; GCP3; NMR; oligomeric state;  $\gamma$ -tubulin

**Abbreviations:** MOZART1, mitotic-spindle organizing protein associated with a ring of  $\gamma$ -tubulin 1;  $\gamma$ TuRC,  $\gamma$ -tubulin ring complex;  $\gamma$ TuSC,  $\gamma$ -tubulin small complex; GCP,  $\gamma$ -tubulin complex protein; SEC-MALS, size exclusion chromatography coupled with multi-angle light scattering; DLS, dynamic light scattering; NMR, nuclear magnetic resonance; LB, Luria-Bertani; IPTG, isopropyl- $\beta$ -D-1-thiogalactopyranoside; DTT, dithiothreitol; IMAC, immobilized metal affinity chromatography; PMSF, phenylmethylsulfonyl fluoride; DDM, N-dodecyl beta maltoside; NDSB-195, non-detergent sulfobetaine-195; LDAO, N,N-dimethyldodecan-1-amine oxide; SDS, sodium dodecyl sulfate; CSI, chemical shift index

Additional Supporting Information may be found in the online version of this article.

**Broader statement:** MOZART1 is a small protein linked to regulation of  $\gamma$ -tubulin ring complexes. Despite its small size, it represents a challenging target for detailed characterization *in vitro* due to oligomerization properties. Here, we show that the use of detergents can significantly improve MOZART1 behavior in solution allowing NMR resonance assignment. This study provides experimental conditions to investigate interactions between MOZART1 and its molecular partners.

Grant sponsor: Agence Nationale de la Recherche; Grant number: ANR-13-BSV8-0007-01; Grant sponsor: Fondation ARC pour la Recherche sur le Cancer; post-doctoral grant to JN.

\*Correspondence to: Virginie Gervais, Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, Toulouse, France. E-mail: virginie.gervais@ipbs.fr and Andreas Merdes, Centre de Biologie du Développement, Université de Toulouse, CNRS, UPS, France. E-mail: andreas.merdes@univ-tlse3.fr

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Introduction

Microtubules are a major component of the cytoskeleton in eukaryotic cells, involved in processes such as transport, cell division, and signaling.<sup>1</sup> They are highly dynamic polar structures capable of rapid rearrangements in response to stimuli.<sup>2–4</sup> Microtubules polymerize from heterodimers of  $\alpha$ - and  $\beta$ -tubulin. The linear arrangement of tubulin dimers into protofilaments creates a polarity, with  $\beta$ -tubulin oriented to the so-called plus-end of the microtubule, and  $\alpha$ -tubulin oriented toward the minus-end. Most microtubules in the cell contain 13 protofilaments, aligned laterally in a three-start helix with a hollow core. Assembly of microtubules involves hydrolysis of GTP to GDP. *In vitro*, polymerization of microtubules from purified  $\alpha/\beta$ -tubulin is kinetically unfavorable and produces a large percentage of polymers with a variable number of protofilaments. To ensure the correct diameter and geometry, assembly of microtubules in cells is nucleated from specific sites, involving multiprotein complexes called  $\gamma$ TuRCs ( $\gamma$ -tubulin ring complexes).<sup>1</sup> Depletion of  $\gamma$ TuRC components in cells leads to strong mitotic defects, such as aberrant mitotic spindles and error-prone segregation of chromosomes.<sup>5</sup>

$\gamma$ TuRCs are  $\sim 2.2$  MDa multiprotein complexes composed of at least six different proteins, with molecular masses ranging from  $\sim 70$  to 210 kDa and present in single or multiple copies in the complex.<sup>1</sup> The overall architecture of the  $\gamma$ TuRCs is conserved between distant organisms such as plants, yeast, and human, but in some lower organisms (e.g., *Saccharomyces cerevisiae*),  $\gamma$ TuRCs can have a reduced number of subunits.<sup>1</sup> The structural insight into  $\gamma$ TuRC organization comes from several cryo-electron microscopy studies<sup>6–12</sup> and from the X-ray structures of  $\gamma$ -tubulin<sup>13,14</sup> and  $\gamma$ -tubulin complex protein 4 (GCP4), a prototype of all GCP proteins.<sup>15</sup> Those studies provided important information on how  $\gamma$ TuRCs may perform their function, leading to models of  $\gamma$ TuRCs as structural templates for microtubule nucleation, by longitudinal binding of  $\alpha/\beta$ -tubulin dimers to the surface of  $\gamma$ -tubulin. However, it remains enigmatic how cells regulate microtubule nucleation spatially and temporally, as the majority of  $\gamma$ TuRCs are found in an inactive form, soluble in the cytoplasm.<sup>16</sup> Various proteins such as pericentrin, NEDD1/GCP-WD, and CDK5RAP2 have been implicated in the regulation of  $\gamma$ TuRCs.<sup>17–19</sup> In addition, recent studies revealed that MOZART1 (human mitotic-spindle organizing protein associated with a ring of  $\gamma$ -tubulin 1) protein may equally be involved in this regulation.<sup>20–27</sup>

MOZART1 is a relatively small protein consisting of 82 residues.<sup>25</sup> Similar to other components of  $\gamma$ TuRCs, it is conserved between different species from animals to plants and most fungi with the exception of

*S. cerevisiae*, and its fission yeast and plant homologs are called Mzt1 and GIP1, respectively (Fig. 1).<sup>20,25</sup> These proteins are predicted to have an  $\alpha$ -helical fold and would interact with a 10-amino-acid hydrophobic sequence motif present in the N-terminal extension of GCP3, GCP5, GCP6, and potentially of GCP2 proteins.<sup>26,27</sup> The binding of MOZART1 to a subset of GCPs may support the assembly of multiple  $\gamma$ TuSCs into  $\gamma$ TuRCs, and a cooperative binding of other proteins to the GCPs, such as proteins carrying CM1-motifs, and NEDD1.<sup>26,27</sup> Furthermore, recent studies have shown that MOZART1 together with Spc110 would act as oligomerization chaperones to cooperatively promote oligomerization of  $\gamma$ TuSCs into well-organized and active microtubule nucleation template.<sup>27</sup> Altogether, these data indicate that MOZART1 may contribute to the recruitment of  $\gamma$ TuRCs to microtubule-organizing centers (MTOCs), and to the activation of microtubule nucleation.<sup>20–27</sup> Consistently, defects in MOZART1 lead to abnormalities in mitotic spindle assembly and in cell division and development.<sup>20–22,24,25,27</sup>

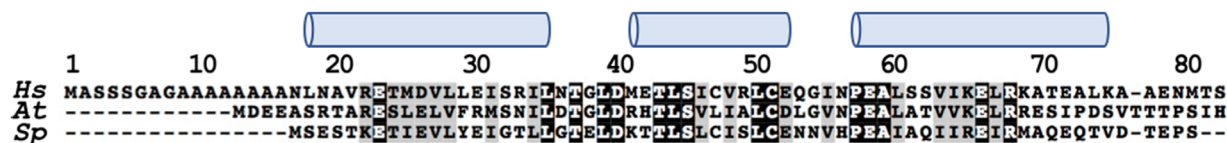
The biochemical and structural description of MOZART1 and its interaction with GCP3 and other GCPs may significantly advance our understanding of  $\gamma$ TuRC functionality. Unfortunately, recent studies on yeast and plant homologs of MOZART1 have shown that despite its small size, Mzt1/GIP1 protein represents a challenging target for detailed characterization *in vitro*.<sup>20,25</sup>

In this article, we report the expression and purification of recombinant human MOZART1 and the structural characterization of the purified protein using a combination of DLS, SEC-MALS, and NMR experiments. We performed <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N resonance assignments, which indicate the presence of three distinct helices. Furthermore, NMR chemical shift perturbation (CSP) analysis revealed that the three helices are involved in the interaction with GCP3. Together these data provide the first complete biochemical, biophysical, and structural characterization of human MOZART1 reported so far, paving the way for further binding studies with  $\gamma$ TuRCs components.

## Results

### Overexpression and purification of MOZART1 protein from prokaryotic cells

The production of MOZART1 protein in *Escherichia coli* cells proved to be challenging, failing to provide any recombinant protein when the wild-type version of the gene was used. To obtain quantities of protein required for *in vitro* characterization, we expressed the protein from a plasmid-containing codon-optimized version of MOZART1 gene (Supporting Information, Fig. S1). The latter gene allowed for production of soluble MOZART1 protein in a range of temperatures



**Figure 1.** Sequence alignment of human MOZART1 protein and its plant and yeast homologs (Hs: *Homo sapiens* MOZART1-NP001065243, At: *Arabidopsis thaliana* GIP1-AEE82764, Sp: *Schizosaccharomyces pombe* Mzt1-POCF96). The strictly conserved residues are in white on a black background and the similar residues are on a grey background. The alignment was done using Clustal Omega<sup>44</sup> and rendered with Boxshade. Location of helices as determined by NMR (see below) is indicated on the top of the alignment.

(20–37°C), with the highest yield for expression at 20°C for 24 h (up to 8 mg of protein per liter of LB or minimal medium culture).

### Biophysical characterization and preparation of samples for NMR studies

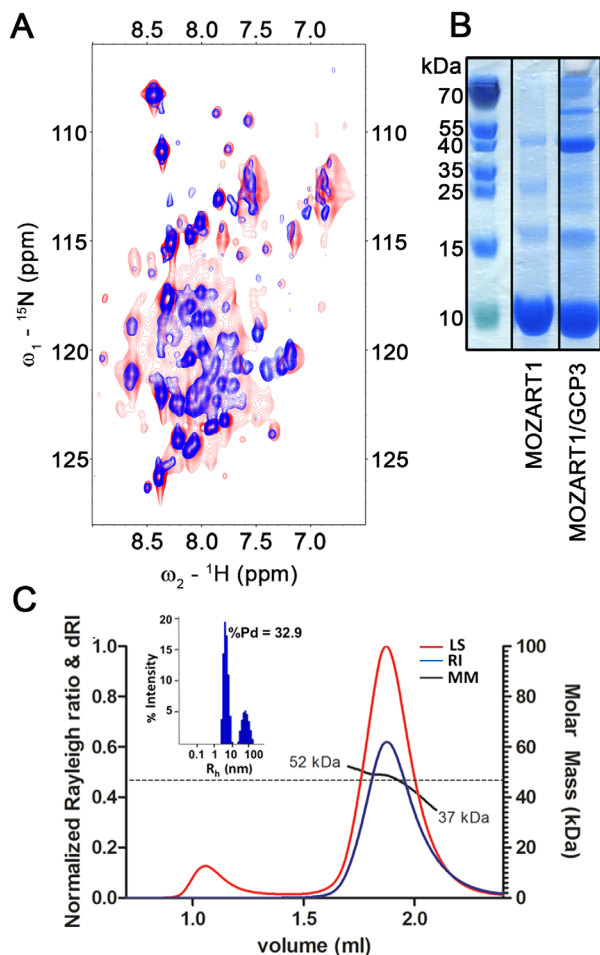
The initial NMR samples gave spectra of poor quality with many broad lines, making resonances unobservable, and with only a few (~20) sharp lines [Fig. 2(A)], a disappointing result considering the low molecular mass of recombinant MOZART1 protein (8.7 kDa). Sharper lines could be obtained by employing TROSY experiments, which are optimized for measuring macromolecules of large molecular weight. Nevertheless, the number of observed signals and the quality of spectra were still not satisfactory [Fig. 2(A)]. The extensive screening of experimental conditions including pH, ionic strength, temperature, and reducing agents could only partially overcome these limitations, but not sufficiently to make resonance assignment of the protein feasible.

To better understand the difficulties encountered, we studied the oligomeric state and homogeneity of the MOZART1 protein sample. We noticed that the protein eluted at ~60 mL on the Superdex S75 16/600 column used for purification, which corresponded to a smaller elution volume and larger size than expected for an 8.7 kDa monomer. To provide a reliable molar mass, we performed SEC-MALS analysis,<sup>28</sup> where light scattering and refractometry curves showed one elution peak giving an average molar mass of 46.6 kDa [Fig. 2(C)]. Such a value is consistent with the presence of pentameric forms. Furthermore, the molar mass range from 37 to 52 kDa raises the possibility that MOZART1 is in equilibrium between different oligomeric states (tetramer to hexamer). Notably, the presence of a smaller peak representing earlier eluted protein (0.1% of the total mass contributing to <10% of the light scattering signal) indicates a minor amount of higher molecular weight species, suggesting that MOZART1 is slightly prone to aggregation. In line with this, DLS analysis demonstrated that the MOZART1 sample was not perfectly monodispersed, revealing a polydispersity of 32.9% [Fig. 2(C)]. Altogether, these data indicate that MOZART1 does not exist in a single oligomeric state, but is rather polydisperse.

As already mentioned, many attempts to improve the NMR spectral quality were only partially successful and NMR spectra of MOZART1, despite employing TROSY-based experiments, were of poor quality. We then included a panel of detergent- and non-detergent-solubilizing agents into our optimization process (Table S1). The tested compounds affected MOZART1 NMR samples to several extents: from no effect (DDM, NDSB-195) to satisfactory enhancement of NMR signals (sulfobetaine-12, LDAO) [Fig. 3(A)]. We chose sulfobetaine-12 to be included in NMR sample of MOZART1 protein, because it gave slightly improved NMR spectra compared with LDAO. Subsequent experiments demonstrated that sulfobetaine-12 effect on MOZART1 spectra is concentration dependent and significant improvement is achieved at ~60:1 molar ratio of sulfobetaine-12:MOZART1 protein [Fig. 3(B)]. The presence of sulfobetaine-12 indeed decreased the heterogeneity of the MOZART1 sample as revealed by the polydispersity value reduced from 32.9% to 22.4% in DLS experiments, nevertheless, still higher when compared to sulfobetaine-12 micelles alone (10.9%) [Figs. 2(C) and 3(C)]. Hydrodynamic radii ( $R_h$ ) were calculated in the different conditions showing a reduction of  $R_h$  value from 5.05 nm for MOZART1 alone to 2.77 nm in sulfobetaine-12 [Figs. 2(C) and 3(C)].

### Resonance assignment and initial structure characterization

The optimized NMR conditions where the MOZART1 protein was solubilized in sulfobetaine-12 micelles allowed us to record the 3D NMR experiments necessary for protein resonance assignment. A completeness of 86% was achieved for backbone resonances and 82% for side-chain resonances with primarily missing assignments corresponding to residues located in the N-terminal alanine-rich region of the protein [Fig. 4(A,B)]. This region is not present in the homologs of MOZART1 (Fig. 1) and is most likely flexible as several sharp cross-peaks of alanines in <sup>1</sup>H-<sup>15</sup>N spectra could not be assigned [Fig. 4(A)]. The analysis of chemical shift values of C<sup>α</sup> and C<sup>β</sup> resonances demonstrated the presence of three α-helices within MOZART1 protein (residues L18-L35, M41-E52, P57-K75) [Fig. 4(C) and Supporting Information, Fig. S2].



**Figure 2.** Characterization of MOZART1 protein sample. (A)  $^{15}\text{N}$ -SOFAST-HMQC (red) and BEST-TROSY (blue) spectra of MOZART1 in 50 mM sodium phosphate pH 6.0, 150 mM NaCl recorded at 25°C and 600 MHz Larmor frequency. (B) Coomassie-stained SDS-PAGE gel of samples of the MOZART1 protein alone after IMAC and SEC chromatography (middle) and after co-purification with the SUMO-GCP3<sup>(1–250)</sup>-His<sub>6</sub> fusion protein on a HiTrap TALON<sup>®</sup> affinity column (right). (C) Elution profile examined by SEC-MALS. The traces of light scattering (red, LS), differential refractive index (blue, RI), and MALS calculated molar mass (black, MM) are shown. The black trace indicates the MALS calculated molar mass of the eluted protein that ranges from 52 to 37 kDa. The grey dashed line indicates the average molar mass value of 46.6 kDa (the theoretical molecular weight of recombinant MOZART1 is 8.7 kDa). The histogram of size distribution obtained from DLS analysis of purified protein is shown in the window where hydrodynamic radius ( $R_h$ ) versus intensity percentage is represented. The major peak is centered at an  $R_h$  value of 5.05 nm. The percentage of polydispersity (%Pd) is indicated.

Similar boundaries were obtained from the analysis of chemical shift values of  $\text{C}^\alpha$ ,  $\text{C}^\beta$ ,  $\text{C}'$ ,  $\text{H}^\text{N}$ ,  $\text{H}^\alpha$ , and N resonances with the TALOS+ software,<sup>29</sup> in which torsion angle values characteristic of  $\alpha$ -helix were observed for residues A16-L35, M41-E52, and P57-A76. The three helices were systematically predicted by structure prediction programs such as Lomets,<sup>30</sup> Phyre2,<sup>31</sup>

or CS-Rosetta<sup>32</sup> with comparable boundaries (Supporting Information, Fig. S2).

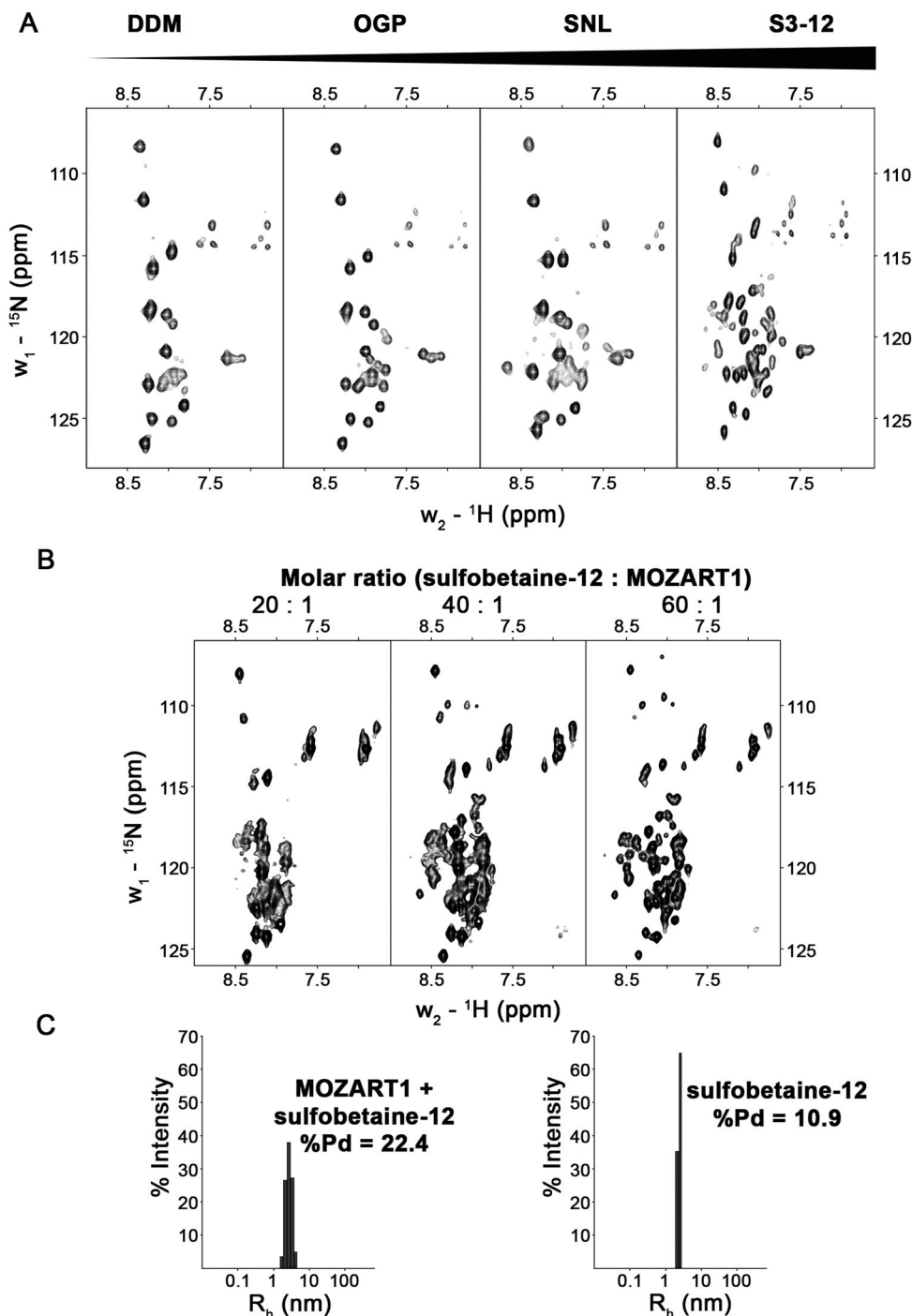
### NMR titration

Because MOZART1 was initially described as a GCP3-interacting protein,<sup>22,24</sup> we studied the binding between these two partners in more detail. Using MOZART1 and the N-terminal domain of GCP3 (residues 1–250), we recorded a series of SOFAST HMQC-based chemical shift perturbation (CSP) experiments upon addition of increasing concentrations of SUMO-GCP3<sup>(1–250)</sup>-His<sub>6</sub> fusion protein and we followed changes in the MOZART1 chemical shifts. The chemical shift variations observed in NMR titration experiments are due to fast exchange between bound and free forms, and are best suited for weak interactions and thus low affinity complexes (typically  $K_d > \mu\text{M}$ ). The addition of unlabeled GCP3 resulted in rather weak chemical shift changes in the HMQC spectra. Nevertheless, several residues of MOZART1 appeared to be more affected upon addition of GCP3, in terms of chemical shift changes (residues S32, K69, T71, and T81) and/or signal broadening (residues G6, L29, and M41). Most of these residues map onto the helices of MOZART1 [Fig. 4(D)].

Altogether, our NMR experiments indicate an interaction between MOZART1 and the N-terminal domain of GCP3 on the millisecond or faster NMR exchange regime.

### Discussion

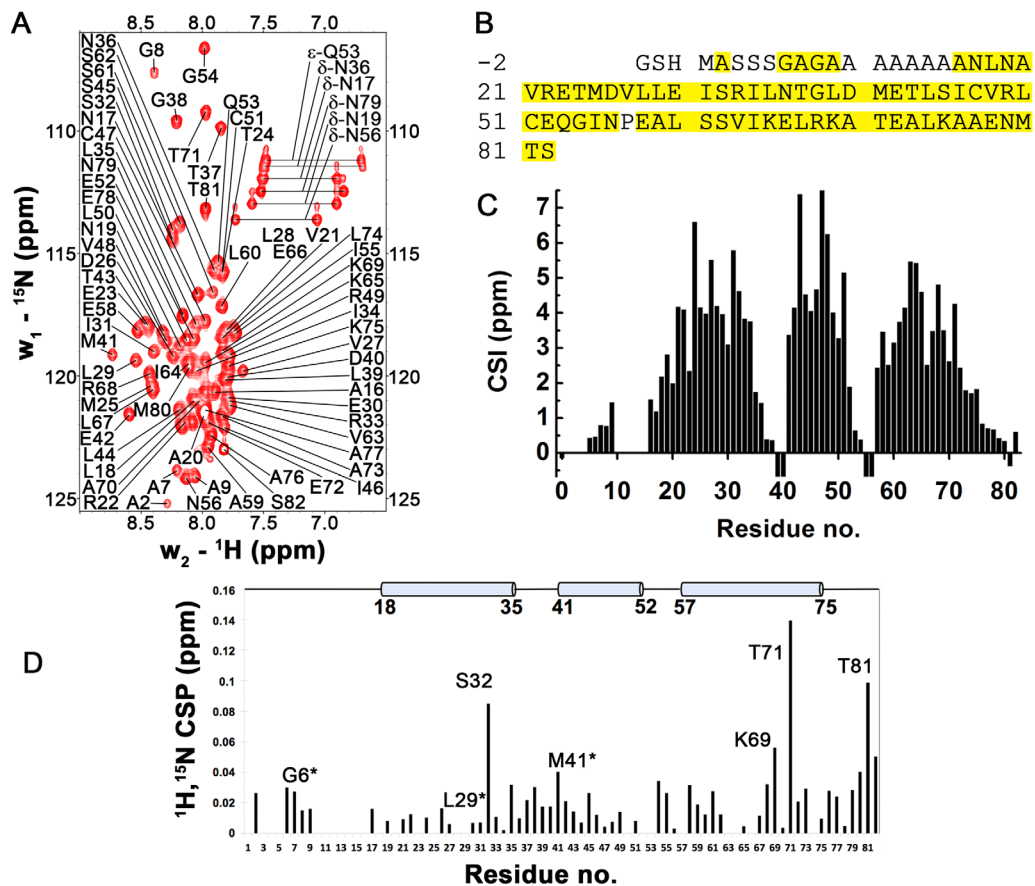
The small MOZART1 protein has recently gained more attention as an important player in the regulation of  $\gamma\text{TuRC}$  targeting and activity.<sup>21,26,27,33</sup> In this article, we describe the protocol for expression and purification of human recombinant MOZART1 protein from *E. coli*, which allows for obtaining sufficient amount of protein for biophysical characterization *in vitro* and for structural studies. Using a combination of DLS, SEC-MALS, and NMR, we demonstrate that MOZART1 displays self-association properties in solution, forming several different oligomers, which are in equilibrium. The single peak observed in the SEC-MALS experiments corresponds to a molar mass of  $46.6 \pm 0.1$  kDa, suggesting a pentamer on average. On the other hand, DLS experiments showed that the sample is heterogeneous. Thus, heterogeneity could explain the poor quality of the NMR spectra, which is likely due to chemical exchange between the different states of MOZART1 in the sample leading to NMR line broadening. The presence of MOZART1 oligomers in solution may reflect a role of this protein in the lateral assembly of multiple  $\gamma\text{TuSCs}$  into  $\gamma\text{TuRCs}$ , similar to the previously documented role of Spc110 in  $\gamma\text{TuRC}$  assembly.<sup>26,27,34–36</sup>



**Figure 3.** Screening of solubilizing agents on MOZART1 protein sample. (A) BEST-TROSY spectra of MOZART1 protein in the presence of different detergents recorded at 25°C and 600 MHz Larmor frequency. The example spectra with increased improvement from left to right are presented. DDM, N-dodecyl beta maltoside; OGP, octyl  $\beta$ -D-glucopyranoside; SNL, N-lauroyl-sarcosine; S3-12, sulfobetaine-12. (B) SOFAST-HMQC spectra of 150  $\mu\text{M}$  MOZART1 in the presence of different concentrations of sulfobetaine-12 detergent. The spectra were recorded at 37°C and 600 MHz Larmor frequency. (C) Histograms of size distribution obtained from DLS analysis for MOZART1 in the presence of sulfobetaine-12 detergent (left) and for sulfobetaine-12 detergent alone (right). %Pd, percent polydispersity.

However, the oligomeric state renders this protein family highly challenging for *in vitro* studies. Dhani et al. showed that yeast Mzt1 protein gives three distinct peaks on a gel filtration column, possibly corresponding to a dodecamer ( $\sim 130$  kDa),

a heptamer/hexamer ( $\sim 50$  kDa), and a tetramer/trimer ( $\sim 35$  kDa).<sup>25</sup> The authors also experienced poor behavior of Mzt1 protein in NMR experiments giving only about 25% of expected correlations in  ${}^1\text{H}$ - ${}^{15}\text{N}$  HSQC spectra.<sup>25</sup> Similarly, oligomeric forms



**Figure 4.** Resonance assignment and NMR secondary structure of MOZART1 protein. (A) Assigned  $^{15}\text{N}$ -BEST-HSQC spectrum of MOZART1 protein recorded at  $37^\circ\text{C}$  and 600 MHz Larmor frequency. (B) Primary sequence of MOZART1 protein with residues for which backbone amide resonances were assigned highlighted in yellow. (C) Chemical shift index (CSI =  $(\delta_{\text{C}}^{\alpha}, \text{MOZART1} - \delta_{\text{C}}^{\alpha}, \text{random coil}) - (\delta_{\text{C}}^{\beta}, \text{MOZART1} - \delta_{\text{C}}^{\beta}, \text{random coil})$ ). Random coil values were taken from Wishart et al.<sup>45</sup> The positive (CSI > 1.5) values indicate clearly three well-defined  $\alpha$ -helical regions in solution. (D) Chemical shift perturbation (CSP) of MOZART1 upon addition of the GCP3<sup>(1–250)</sup> fusion protein. The CSPs of MOZART1 are plotted versus residue numbers with positions of helices as determined by NMR indicated on top. The residues that display chemical shift changes or whose NMR signals are significantly broadened upon titration with GCP3 are labeled. Stars indicate residues displaying signal broadening (until the limit of detection in the case of L29).

were observed in *C. albicans* (CaMzt1) and plant (GIP1) MOZART1 homologs.<sup>20,27</sup>

We demonstrate here that the quality of MOZART1 protein sample can be significantly improved by including detergents in the sample. The best results were obtained for sulfobetaine-12, which in molar ratio of  $\sim 60:1$  (detergent:protein) provided significant improvement of NMR signals. Interestingly, the aggregation number (number of molecules forming individual micelles) of sulfobetaine-12 is 55–87,<sup>37</sup> suggesting that there is one MOZART1 molecule per detergent micelle. Our experimental conditions provided the best quality of NMR spectra obtained so far for a MOZART1/Mzt1/GIP1 protein, allowing us to report the first nearly complete chemical shift resonance assignment for MOZART1. The protein displays three distinct  $\alpha$ -helical regions, adding unambiguous and accurate structural data to the predicted 3D model of AtGIP1/MOZART1.<sup>20</sup> These regions largely correspond to conserved parts in all organisms.

Despite our attempts to measure at 950 MHz on a concentrated sample, we were unable to identify sufficient amount of long-range unambiguous NOE restraints for structure calculations. Therefore, the atomic-resolution structure of MOZART1 remains elusive.

Finally, we showed that MOZART1 is able to bind to the N-terminal extension of GCP3 in detergent solution, and identified residues perturbed upon interaction that are located in the helices of MOZART1, thereby participating to the binding interface. In our NMR experiments, the weak binding observed between these two proteins supports the fact that MOZART1 may interact with fully assembled  $\gamma$ TuRC rather than with individual GCP components,<sup>26</sup> likely resulting in a high-avidity interaction. Whether MOZART1 uses a unique interface to bind GCP3 and other components such as Spc110 remains to be addressed.

In summary, this study clearly establishes the tendency of MOZART1 to form oligomers in nondetergent

solutions, a fact that may be linked with its capacity to promote multiple contacts with GCPs and its key role as an oligomerization chaperone for correct microtubule nucleation template assembly. In this context, our work provides experimental conditions for the observation of good-quality NMR spectra, and will therefore contribute to future structural analysis of complexes between MOZART1 and its molecular partners, leading to a better understanding of the supramolecular organization of  $\gamma$ TuRC complexes.

## Materials and Methods

### **Cloning, protein expression, and purification**

The wild-type *MOZART1* gene was PCR amplified from a pGEX-MOZART1 plasmid.<sup>26</sup> Next, wild-type and codon-optimized (GenScript) versions of full-length human *MOZART1* gene were cloned into pET28a expression vector (Novagen) introducing N-terminal thrombin-cleavable HisTag. Recombinant unlabeled MOZART1 protein was produced in a BL21 (DE3) strain of *E. coli* (Novagen) cultivated at 30°C in LB medium supplemented with 30  $\mu$ g/mL kanamycin. Expression was induced with 0.5 mM isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) (Euro-medex) when the optical density at 600 nm reached 0.35–0.5. Cells were cultured for another 4 h at 30°C, harvested by centrifugation and stored at –20°C. For <sup>15</sup>N- and <sup>15</sup>N,<sup>13</sup>C-labeled samples of MOZART1, LB medium was substituted with M9 minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl (Sigma) and <sup>13</sup>C-D-glucose (Cambridge Isotope Laboratories) as the only source of nitrogen and carbon, respectively, and the temperature and time for expression were modified to 20°C and 24 h, respectively.

Frozen cell pellets were resuspended in equilibration buffer (50 mM sodium phosphate pH 8.0, 10 mM imidazole, 150 mM NaCl, 10% (v/v) glycerol, 1 mM dithiothreitol (DTT)) containing 0.1 mg/mL lysozyme, 0.03 mg/mL DNase I, 0.1% (v/v) Triton X-100 and a mini-Complete EDTA-free protease inhibitor cocktail tablet (Roche), sonicated on ice, and centrifuged (40,000g, 40 min, 4°C). Recombinant MOZART1 protein was purified from the soluble fraction using immobilized metal ion affinity chromatography column (GE Healthcare Hi-Trap™ IMAC). The bound protein was eluted using a gradient of 10–500 mM imidazole in 50 mM sodium phosphate pH 8.0, 150 mM NaCl, 10% (v/v) glycerol, and 1 mM DTT buffer. Subsequently, the fractions containing recombinant protein were pooled, concentrated to 2.5–3 mg/mL and treated with thrombin for 3 h at 20°C (Novagen, 0.9 U per 1 mg of recombinant protein) to remove the histidine tag. The thrombin was inactivated with 1 mM PMSF and the recombinant MOZART1 protein was further purified by size-exclusion chromatography (HiLoad Superdex S75 16/600 column, GE Healthcare) using 50 mM

sodium phosphate pH 8.0, 150 mM NaCl, and 2 mM DTT as a running buffer. The fractions with the recombinant MOZART1 protein were combined and the buffer was exchanged to 10 mM sodium phosphate pH 6.0, 50 mM NaCl, and 2 mM DTT by dialysis (NMR buffer). The final protein samples were concentrated to 300–400  $\mu$ M, flash-frozen in liquid nitrogen, and stored at –80°C. Protein concentration was determined using Bradford assay (Bio-Rad). To perform NMR titration with MOZART1, the N-terminus domain of human GCP3 (residues 1–250) was cloned into a pET14b plasmid and expressed as a SUMO-GCP3<sup>(1–250)</sup>-His<sub>6</sub> cleavable fusion protein. Owing to poor yield of GCP3 protein during purification preventing any NMR titration study, different strategies have been considered of which co-purification, which resulted in improved yields. Briefly, the recombinant GCP3<sup>(1–250)</sup> and MOZART1 proteins were first expressed separately, then co-lysed and co-purified using a HiTrap TALON® affinity column (GE healthcare). As the two proteins contain a histidine tag, they were first retained on the IMAC column and then separated using size-exclusion chromatography (HiLoad Superdex S200 16/600 column, GE healthcare) [Fig. 2(B) and Supporting Information, Fig. S3]. After purification, the recombinant SUMO-GCP3<sup>(1–250)</sup>-His<sub>6</sub> fusion protein was concentrated to a final concentration of 5 mg/mL (190  $\mu$ M) in the NMR buffer (with sulfobetaine-12, see below) to perform NMR titration of <sup>15</sup>N-labeled MOZART1 (purified separately).

### **Dynamic light scattering (DLS)**

Samples of 150  $\mu$ M <sup>15</sup>N-labeled MOZART1 protein in 50 mM sodium phosphate pH 6.0, 150 mM NaCl with or without 25 mM sulfobetaine-12, and 10 mM DTT were analyzed at 20.5°C using a DynaPro NanoStar instrument (Wyatt Technology) and 4  $\mu$ L disposable cuvettes (Wyatt Technology). The data were analyzed with Dynamics software (Wyatt Technology).

### **Size exclusion chromatography coupled with multiangle light scattering (SEC-MALS)**

<sup>15</sup>N-labeled MOZART1 protein buffered in 50 mM sodium phosphate pH 6.0, 150 mM NaCl was analyzed on a Superdex 200 Increase 5/150 GL (GE Healthcare) column with multiangle light scattering (MALS). The column was equilibrated in a 50 mM sodium phosphate 0.1  $\mu$ m filtered buffer (pH 6.0 and 150 mM NaCl) on an Agilent 1260 Infinity LC chromatographic system (Agilent Technology). Data were collected using a DAWN 8+ and Optilab T-rEX refractive index detector (Wyatt Technology). Of 150  $\mu$ M protein sample, 70  $\mu$ L was loaded on the column and the separation was performed at a flow rate of 0.4 mL/min at 15°C. Results were analyzed using the ASTRA 6.1 software (Wyatt Technology).



## Nuclear magnetic resonance (NMR) spectroscopy

For optimization of experimental conditions, samples of 50–350  $\mu\text{M}$   $^{15}\text{N}$ -labeled MOZART1 protein were prepared in 90%  $\text{H}_2\text{O}/10\%$   $^2\text{H}_2\text{O}$  solutions of buffers with different pH, ionic strengths, and additives. The quality of samples was evaluated by recording various  $^1\text{H}$ - $^{15}\text{N}$  correlation experiments (HSQC, TROSY, BEST-TROSY, BEST-HSQC, and SOFAST-HMQC) on a Bruker Avance III spectrometer equipped with TCI cryoprobe and operating at 600 MHz  $^1\text{H}$  frequency or on a Bruker Avance III spectrometer equipped with QCI probe and operating at 700 MHz  $^1\text{H}$  frequency. The range of temperature tested was 25–45°C.

For resonance assignment,  $^{15}\text{N}$ - or  $^{15}\text{N},^{13}\text{C}$ -labeled MOZART1 protein samples at concentrations in the range of 0.45–0.7 mM were prepared in 95%  $\text{H}_2\text{O}/5\%$   $^2\text{H}_2\text{O}$  solutions of 10 mM sodium phosphate pH 6.0, 50 mM NaCl, 2 mM DTT, and 30–50 mM sulfobetaine-12 (sulfobetaine-12:MOZART1 molar ratio was always >60:1). NMR spectra were recorded at 37°C on Bruker Avance III or Bruker Avance II spectrometers equipped with TCI cryoprobes and operating at 600 or 950 MHz  $^1\text{H}$  frequencies, respectively. All NMR spectra were processed with Topspin 3.2 (Bruker Biospin) and analyzed using NMRFAM-SPARKY.<sup>38</sup>  $^1\text{H}^{\text{N}}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}^{\alpha}$ ,  $^{13}\text{C}^{\beta}$ , and  $\text{C}'$  resonance assignments were obtained from HNCA, BEST-HN(CO)CACB, BEST-HNCACB, HNCO, and HN(CA)CO experiments.<sup>39,40</sup> The side-chain resonance assignments were obtained from  $^{15}\text{N}$ -TOCSY-HSQC (mixing time = 50 ms), (H)CCH-TOCSY (mixing time = 11 ms), H(C)CH-TOCSY (mixing time = 11 ms),  $^{15}\text{N}$ -NOESY-HSQC (mixing time = 80 or 120 ms), and  $^{13}\text{C}$ -NOESY-HSQC (mixing time = 80 or 120 ms).<sup>40–42</sup> The assignment was deposited in BMRB under accession number 27093.

For NMR titration and interaction with the N-terminus domain of GCP3, 2D SOFAST-HMQC experiments<sup>43</sup> of MOZART1 at the concentration of 12–20  $\mu\text{M}$  were recorded at 298 K after each incremental addition of unlabeled SUMO-GCP3<sup>(1–250)</sup>-His<sub>6</sub>. We could not use higher excess of GCP3<sup>(1–250)</sup> due to degradation and/or precipitation problems. Normalized chemical shift changes of MOZART1 (CSP) were calculated as  $\Delta\delta_{\text{weighted}} = [(\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{N}} \times 0.154)^2]^{1/2}$ .

## Acknowledgments

SEC-MALS, DLS equipments, and Bruker 500 MHz, 600 MHz, and 700 MHz NMR spectrometers belong to the IBSA Integrated Screening Platform of Toulouse (PICT, IPBS, CNRS - Université de Toulouse). The Région Midi-Pyrénées and European structural funds are acknowledged for funding the NMR equipment. The authors acknowledge TGIR-RMN for access to a Bruker 950 MHz NMR spectrometer in Grenoble and

Drs Bernhard Brutscher and Adrien Favier for their support and discussions. They thank Dr Jens Lüders (IRB Barcelona) for providing pGEX-MOZART1 plasmid.

## Conflict of Interest

The authors declare no conflict of interest in publishing the results of this study.

## References

1. Kollman JM, Merdes A, Mourey L, Agard DA (2011) Microtubule nucleation by gamma-tubulin complexes. *Nat Rev Mol Cell Biol* 12:709–721.
2. Nogales E, Whittaker M, Milligan RA, Downing KH (1999) High-resolution model of the microtubule. *Cell* 96:79–88.
3. Nogales E, Wolf SG, Downing KH (1998) Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* 391:199–203.
4. Akhmanova A, Steinmetz MO (2015) Control of microtubule organization and dynamics: two ends in the limelight. *Nat Rev Mol Cell Biol* 16:711–726.
5. Remy MH, Merdes A, Gregory-Pauron L (2013) Assembly of gamma-tubulin ring complexes: implications for cell biology and disease. *Prog Mol Biol Transl Sci* 117: 511–530.
6. Oegema K, Wiese C, Martin OC, Milligan RA, Iwamatsu A, Mitchison TJ, Zheng Y (1999) Characterization of two related *Drosophila* gamma-tubulin complexes that differ in their ability to nucleate microtubules. *J Cell Biol* 144:721–733.
7. Kollman JM, Polka JK, Zelter A, Davis TN, Agard DA (2010) Microtubule nucleating gamma-TuSC assembles structures with 13-fold microtubule-like symmetry. *Nature* 466:879–882.
8. Zheng Y, Wong ML, Alberts B, Mitchison T (1995) Nucleation of microtubule assembly by a gamma-tubulin-containing ring complex. *Nature* 378:578–583.
9. Moritz M, Braunfeld MB, Guenebaut V, Heuser J, Agard DA (2000) Structure of the gamma-tubulin ring complex: a template for microtubule nucleation. *Nat Cell Biol* 2:365–370.
10. Kollman JM, Zelter A, Muller EG, Fox B, Rice LM, Davis TN, Agard DA (2008) The structure of the gamma-tubulin small complex: implications of its architecture and flexibility for microtubule nucleation. *Mol Biol Cell* 19:207–215.
11. Choy RM, Kollman JM, Zelter A, Davis TN, Agard DA (2009) Localization and orientation of the gamma-tubulin small complex components using protein tags as labels for single particle EM. *J Struct Biol* 168:571–574.
12. Greenberg CH, Kollman J, Zelter A, Johnson R, MacCoss MJ, Davis TN, Agard DA, Sali A (2016) Structure of gamma-tubulin small complex based on a cryo-EM map, chemical cross-links, and a remotely related structure. *J Struct Biol* 194:303–310.
13. Aldaz H, Rice LM, Stearns T, Agard DA (2005) Insights into microtubule nucleation from the crystal structure of human gamma-tubulin. *Nature* 435:523–527.
14. Rice LM, Montabana EA, Agard DA (2008) The lattice as allosteric effector: structural studies of alphabeta- and gamma-tubulin clarify the role of GTP in microtubule assembly. *Proc Natl Acad Sci USA* 105:5378–5383.
15. Guillet V, Knibiehler M, Gregory-Pauron L, Remy MH, Chemin C, Raynaud-Messina B, Bon C, Kollman JM, Agard DA, Merdes A, Mourey L (2011) Crystal structure

- of gamma-tubulin complex protein GCP4 provides insight into microtubule nucleation. *Nat Struct Mol Biol* 18:915–919.
16. Moudjou M, Bordes N, Paintrand M, Bornens M (1996) gamma-Tubulin in mammalian cells: the centrosomal and the cytosolic forms. *J Cell Sci* 109:875–887.
  17. Luders J, Patel UK, Stearns T (2006) GCP-WD is a gamma-tubulin targeting factor required for centrosomal and chromatin-mediated microtubule nucleation. *Nat Cell Biol* 8:137–147.
  18. Haren L, Remy MH, Bazin I, Callebaut I, Wright M, Merdes A (2006) NEDD1-dependent recruitment of the gamma-tubulin ring complex to the centrosome is necessary for centriole duplication and spindle assembly. *J Cell Biol* 172:505–515.
  19. Choi YK, Liu P, Sze SK, Dai C, Qi RZ (2010) CDK5RAP2 stimulates microtubule nucleation by the gamma-tubulin ring complex. *J Cell Biol* 191:1089–1095.
  20. Batzenschlager M, Masoud K, Janski N, Houlne G, Herzog E, Evrard JL, Baumberger N, Erhardt M, Nomine Y, Kieffer B, Schmit AC, Chaboute ME (2013) The GIP gamma-tubulin complex-associated proteins are involved in nuclear architecture in *Arabidopsis thaliana*. *Front Plant Sci* 4:480.
  21. Masuda H, Mori R, Yukawa M, Toda T (2013) Fission yeast MOZART1/Mzt1 is an essential gamma-tubulin complex component required for complex recruitment to the microtubule organizing center, but not its assembly. *Mol Biol Cell* 24:2894–2906.
  22. Nakamura M, Yagi N, Kato T, Fujita S, Kawashima N, Ehrhardt DW, Hashimoto T (2012) *Arabidopsis* GCP3-interacting protein 1/MOZART 1 is an integral component of the gamma-tubulin-containing microtubule nucleating complex. *Plant J* 71:216–225.
  23. Janski N, Herzog E, Schmit AC (2008) Identification of a novel small *Arabidopsis* protein interacting with gamma-tubulin complex protein 3. *Cell Biol Int* 32:546–548.
  24. Janski N, Masoud K, Batzenschlager M, Herzog E, Evrard JL, Houlne G, Bourge M, Chaboute ME, Schmit AC (2012) The GCP3-interacting proteins GIP1 and GIP2 are required for gamma-tubulin complex protein localization, spindle integrity, and chromosomal stability. *Plant Cell* 24:1171–1187.
  25. Dhani DK, Goult BT, George GM, Rogerson DT, Bitton DA, Miller CJ, Schwabe JW, Tanaka K (2013) Mzt1/Tam4, a fission yeast MOZART1 homologue, is an essential component of the gamma-tubulin complex and directly interacts with GCP3(Alp6). *Mol Biol Cell* 24:3337–3349.
  26. Cota RR, Teixido-Travesa N, Ezquerro A, Eibes S, Lacasa C, Roig J, Luders J (2017) MZT1 regulates microtubule nucleation by linking gamma TuRC assembly to adapter-mediated targeting and activation. *J Cell Sci* 130:406–419.
  27. Lin TC, Neuner A, Flemming D, Liu P, Chinen T, Jakle U, Arkowitz R, Schiebel E (2016) MOZART1 and gamma-tubulin complex receptors are both required to turn gamma-TuSC into an active microtubule nucleation template. *J Cell Biol* 215:823–840.
  28. Wyatt PJ (1993) Light-scattering and the absolute characterization of macromolecules. *Anal Chim Acta* 272:1–40.
  29. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223.
  30. Wu ST, Zhang Y (2007) LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res* 35:3375–3382.
  31. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858.
  32. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci USA* 105:4685–4690.
  33. Lin TC, Neuner A, Schiebel E (2015) Targeting of gamma-tubulin complexes to microtubule organizing centers: conservation and divergence. *Trends Cell Biol* 25:296–307.
  34. Lin TC, Neuner A, Schlosser YT, Scharf AND, Weber L, Schiebel E (2014) Cell-cycle dependent phosphorylation of yeast pericentrin regulates gamma-TuSC-mediated microtubule nucleation. *Elife* 3:[ARTICLE #].
  35. Lynch EM, Grocock LM, Borek WE, Sawin KE (2014) Activation of the gamma-tubulin complex by the Mto1/2 complex. *Curr Biol* 24:896–903.
  36. Lyon AS, Morin G, Moritz M, Yabut KCB, Vojnar T, Zelter A, Muller E, Davis TN, Agard DA (2016) Higher-order oligomerization of Spc110p drives gamma-tubulin ring complex assembly. *Mol Biol Cell* 27:2245–2258.
  37. le Maire M, Champeil P, Moller JV (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta* 1508:86–111.
  38. Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* 31:1325–1327.
  39. Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J Magn Reson* 187:163–169.
  40. Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spect* 34:93–158.
  41. Schleucher J, Schwendinger M, Sattler M, Schmidt P, Schedletzky O, Glaser SJ, Sorensen OW, Griesinger C (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J Biomol NMR* 4:301–306.
  42. Kay LE, Xu GY, Singer AU, Muhandiram DR, Forman-Kay JD (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain <sup>1</sup>H and <sup>13</sup>C correlations in H<sub>2</sub>O samples of proteins. *J Magn Reson B* 101:333–337.
  43. Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33:199–211.
  44. Sievers F, Higgins DG (2014) Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 1079:105–116.
  45. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD (1995) H-1, C-13 and N-15 random coil NMR chemical-shifts of the common amino-acids. 1. Investigations of nearest-neighbor effects. *J Biomol NMR* 5:332–332.

## **Supplementary Material**

**Figure S1.** Sequence alignment of wild-type and codon-optimized genes encoding human MOZART1 protein. The identical residues are highlighted in black. The alignment was done using Clustal Omega [1] and rendered with ESPrnt 3.0 [2].

**Figure S2.** Localization of the three helices as determined by NMR (analysis of C<sup>α</sup> and C<sup>β</sup> chemical shift values), and comparison with various prediction programs (Jpred4 [3], Lomets [4], Phyre2 [5] and CS-Rosetta [6]). The percentage of predicted α-helix is given.

**Figure S3.** Co-purification of recombinant MOZART1 and SUMO-GCP3<sup>(1-250)</sup>-His<sub>6</sub> fusion proteins using size-exclusion chromatography (HiLoad Superdex S200 16/600 column, GE healthcare). Closed view on the elution profile obtained for both proteins and corresponding SDS-PAGE characterization. Elution fractions are marked in red.

**Table S1.** Overview of solubilizing agents tested.

- 
1. Sievers, F. & Higgins, D. G. (2014) Clustal Omega, accurate alignment of very large numbers of sequences, *Methods Mol Biol.* **1079**, 105-16.
  2. Robert, X. & Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.* **42**, W320-W324.
  3. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. (2015) JPred4: a protein secondary structure prediction server, *Nucleic Acids Res.* **43**, W389-94.
  4. Wu, S. T. & Zhang, Y. (2007) LOMETS: A local meta-threading-server for protein structure prediction, *Nucleic Acids Res.* **35**, 3375-3382.
  5. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc.* **10**, 845-858.
  6. Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsy, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D. & Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data, *Proc Natl Acad Sci U S A.* **105**, 4685-90.
-

| Reagent                          | IUPAC name  | Charge       | MW<br>g.mol <sup>-1</sup> | Critical<br>micelle<br>concentration<br>(mM) <sup>a</sup> | Aggregation<br>number <sup>a</sup> | Structure |
|----------------------------------|---|--------------|---------------------------|---|------------------------------------|-----------|
| <b>CHAPS</b>                     | 3-[dimethyl-[3-[[[(4R)-4-[(3R,5S,7R,8R,9S,10S,12S,13R,14S,17R)-3,7,12-trihydroxy-10,13-dimethyl-2,3,4,5,6,7,8,9,11,12,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl]pentanoyl]amino]propyl]azaniumyl]propane-1-sulfonate | Zwitterionic | 615                       | 3-10  | 4-14                               |           |
| <b>NDSB-195</b>                  | 3-[ethyl(dimethyl)azaniumyl]propane-1-sulfonate   | Zwitterionic | 195                       | NA  | NA                                 |           |
| <b>Sulfobetaine-12</b>           | 3-[dodecyl(dimethyl)azaniumyl]propane-1-sulfonate   | Zwitterionic | 336                       | 1.4-4   | 55-87                              |           |
| <b>LDAO</b>                      | N,N-dimethyldodecan-1-amine oxide   | Zwitterionic | 229                       | 2.2   | 69-73                              |           |
| <b>Octyl β-D-glucopyranoside</b> | (2R,3S,4S,5R,6R)-2-(hydroxymethyl)-6-octoxyoxane-3,4,5-triol  | Non-ionic    | 292                       | 19-25   | 90                                 |           |
| <b>N-dodecyl beta maltoside</b>  | (2R,3R,4S,5S,6R)-2-[(2R,3S,4R,5R,6R)-6-dodecoxy-4,5-dihydroxy-2-(hydroxymethyl)oxan-3-yl]oxy-6-(hydroxymethyl)oxane-3,4,5-triol   | Non-ionic    | 511                       | 0.18  | 110-140                            |           |
| <b>N-lauroyl-sarcosine</b>       | sodium 2-[dodecanoyl(methyl)amino]acetate   | Anionic      | 293                       | ~15   | 2                                  |           |
| <b>SDS</b>                       | sodium dodecyl sulfate  | Anionic      | 288                       | 1.2-7.1   | 62-101                             |           |

<sup>a</sup> Critical micelle concentration and aggregation numbers are taken from le Maire *et al.* [36] and Sanders *et al.* [46]

|                 |                 |                 |               |              |              |              |
|-----------------|-----------------|-----------------|---------------|--------------|--------------|--------------|
|                 | 1               | 10              | 20            | 30           | 40           | 50           |
| WT              | <b>ATGGCGAG</b> | TAGC <b>AGC</b> | <b>GGTGCT</b> | <b>GGGGC</b> | <b>GCGGC</b> | <b>GCCGC</b> |
| codon_optimized | <b>ATGGCGAG</b> | CTCT <b>AGT</b> | <b>GGTGC</b>  | <b>GGTGC</b> | <b>GCGGC</b> | <b>GCCGC</b> |

|                 |                  |               |            |                 |          |
|-----------------|------------------|---------------|------------|-----------------|----------|
|                 | 60               | 70            | 80         | 90              | 100      |
| WT              | <b>TCTGAATGC</b> | <b>GGTGCG</b> | <b>GGA</b> | <b>ACCATGGA</b> | <b>C</b> |
| codon_optimized | <b>TCTGAATGC</b> | <b>AGTGCG</b> | <b>GAA</b> | <b>ACCATGGA</b> | <b>T</b> |

|                 |                   |              |              |                |             |
|-----------------|-------------------|--------------|--------------|----------------|-------------|
|                 | 110               | 120          | 130          | 140            | 150         |
| WT              | <b>TTTTGAATAC</b> | <b>TGGCT</b> | <b>TAGAT</b> | <b>ATGGAAC</b> | <b>CTG</b>  |
| codon_optimized | <b>TCTTGAACAC</b> | <b>GGCT</b>  | <b>TGAC</b>  | <b>ATGGAAC</b> | <b>GCTG</b> |

|                 |                 |               |             |            |              |
|-----------------|-----------------|---------------|-------------|------------|--------------|
|                 | 160             | 170           | 180         | 190        | 200          |
| WT              | <b>TGTGAACA</b> | <b>AGGAAT</b> | <b>TAA</b>  | <b>CCA</b> | <b>GAAGC</b> |
| codon_optimized | <b>TGTGAACA</b> | <b>GGTAT</b>  | <b>CAAT</b> | <b>CCG</b> | <b>GAAGC</b> |

|                 |          |              |              |               |
|-----------------|----------|--------------|--------------|---------------|
|                 | 210      | 220          | 230          | 240           |
| WT              | <b>T</b> | <b>CGCAA</b> | <b>GCTAC</b> | <b>TGAAGC</b> |
| codon_optimized | <b>G</b> | <b>CGCAA</b> | <b>GCAC</b>  | <b>GAAGC</b>  |



