Structure of the Tuberous Sclerosis Complex 2 (TSC2) N Terminus Provides Insight into Complex Assembly and Tuberous Sclerosis Pathogenesis*

Reinhard Zech1, Stephan Kiontke1, Uwe Mueller2, Andrea Oeckinghaus3, and Daniel Kümmel4

From the 1Structural Biology Section, FBS Biology/Chemistry, University of Osnabrück, 49076 Osnabrück, Germany, 2Macromolecular Crystallography (BESSY-MX), Helmholtz-Zentrum Berlin für Materialien und Energie, 12489 Berlin, Germany, and 3Institute of Molecular Tumor Biology, Medical Faculty of the WWU Münster, 48149 Münster Germany

Tuberous sclerosis complex (TSC) is caused by mutations in the TSC1 and TSC2 tumor suppressor genes. The gene products hamartin and tuberin form the TSC complex that acts as GTPase-activating protein for Rheb and negatively regulates the mammalian target of rapamycin complex 1 (mTORC1). Tuberin contains a RapGAP homology domain responsible for inactivation of Rheb, but functions of other protein domains remain elusive. Here we show that the TSC2 N terminus interacts with the TSC1 C terminus to mediate complex formation. The structure of the TSC2 N-terminal domain from Chaetomium thermophilum and a homology model of the human tuberin N terminus are presented. We characterize the molecular requirements for TSC1–TSC2 interactions and analyze pathological point mutations in tuberin. Many mutations are structural and produce improperly folded protein, explaining their effect in pathology, but we identify one point mutant that abrogates complex formation without affecting protein structure. We provide the first structural information on TSC2/tuberin with novel insight into the molecular function.

In eukaryotic cells growth factor signaling is responsible to adjust proliferation and metabolism to environmental and developmental cues. Growth factor receptors regulate, among other targets, activation of the mammalian target of rapamycin complex 1 (mTORC1),2 considered the master regulator of cellular growth (1). This pathway involves the protein kinase B/AKT, the tuberous sclerosis complex (TSC), and the small GTPase Rheb (2). Under resting conditions, Rheb is kept inactive by TSC, which represents the cognate GTPase-activating protein (GAP) (3). Rheb undergoes the classical GTPase conversion and cycles between an active GTP-bound and an inactive GDP-bound state. Active Rheb is required for full activation of mTORC1 at the lysosomal membrane (4). Upon stimulation of growth factor signaling and the downstream kinase AKT, TSC gets inactivated by phosphorylation, thus leading to increasing levels of active Rheb and mTORC1.

The TSC complex consists of the 130-kDa subunit hamartin (encoded by TSC1) (5), the 200-kDa protein tuberin (TSC2 gene product) (6), and the recently identified TBC1D7 (7, 8), which form high oligomeric assemblies in cells (9) and have a similar architecture as the RalGAP complexes (10, 11). Tuberin contains a RapGAP-like domain that catalyzes the hydrolysis of Rheb-GTP to Rheb-GDP (12). In addition to the GAP domain, hamartin and tuberin show no sequence homology to other known proteins, but the N terminus of tuberin has been shown to mediate interaction with hamartin (13–16). TBC1D7 is a RabGAP domain-containing protein with specificity for Rab17 involved in trafficking from endosomes to cilia (17). In the context of the TSC complex, TBC1D7 stabilizes the hamartin dimerization and the interaction with tuberin and increases the GAP activity of the complex (7, 18, 19). Activity of TSC is in part regulated by its recruitment to Rheb on the lysosome, which has been reported to be promoted by Rag GTPases upon amino acid starvation (20). In contrast, insulin-dependent AKT phosphorylation of TSC2 at Ser939 and Ser981 is responsible for interactions between TSC2 and 14-3-3 scaffold proteins, leading to a translocation of the TSC complex from the lysosomal membrane to the cytosol and thus physically away from its target Rheb (21, 22).

Structural information on the TSC complex is limited, but recently, the crystal structure of the N-terminal HEAT repeat domain of TSC1 from Schizosaccharomyces pombe (TSC1-N) has been reported (23). TSC1-N forms a dimer in solution and a decameric assembly in the crystal, suggesting an involvement in oligomerization of TSC. Furthermore, TSC1-N has been proposed to interact with membranes. In addition, the complex structure of TBC1D7 with a short coiled-coil fragment of hamartin revealed how TBC1D7 stabilizes hamartin dimerization (18, 19).

Mutations in the TSC1 and TSC2 genes lead to tuberous sclerosis, a genetic disorder that affects 0.01–0.02% of all births and is characterized by the formation of benign tumors in skin, brain, heart, kidneys, and multiple other organs (24). Further clinical manifestations are epilepsy, learning difficulties, and behavioral problems (25). Mutations that cause TSC are spread over the entire sequences of both TSC1 and TSC2 (26–29), but...
mechanistic understanding of the functional consequences is precluded by the lack of biochemical and structural information. Also, the diverse spectrum of pathological phenotypes depending on different mutations remains elusive.

Orthologues of TSC1 and TSC2 exist in a range of eukaryotes (30), including the thermophilic fungus *Chaetomium thermophilum*, which we used for recombinant production of TSC proteins. These proteins show increased solubility and stability as compared with mammalian homologues and are thus better suited for structural studies (31). We find that the TSC2 N-terminus (TSC2-N) forms a HEAT repeat domain and interacts with the C-terminal half of TSC1 (TSC1-C). The structure helps to explain the effect of pathogenic missense mutations in tuberin (we will use TSC1/2 and hamartin/tuberin to distinguish between *C. thermophilum* and human proteins in the following), and we identify a point mutation that specifically blocks TSC complex formation without destroying protein folding. Our structural and biochemical characterization provides important molecular insight into TSC complex function and also represents a step forward in understanding tuberous sclerosis pathogenesis.

**Results**

**TSC2 N Terminus Interacts with TSC1 C Terminus**—Initial trials for recombinant expression of hamartin and tuberin in bacteria revealed problems with protein solubility and stability of the tested constructs. We then turned to TSC1 and TSC2 from *C. thermophilum*, which have an identical domain organization as their mammalian counterparts (Fig. 1a) and show sequence conservation over the entire length (supplemental Fig. S1). Studies on TSC1 and TSC2 can therefore serve as a model for hamartin and tuberin.

In our efforts to characterize the structure and assembly of the TSC complex, we focused on the N terminus of TSC2/
tuberin, which had been implicated in mediating the interaction with hamartin (13, 14, 16). We initially expressed a construct comprising residues 70–800 of TSC2 that was designed based on primary sequence analysis (32). In a limited proteolysis experiment with this protein (Fig. 2a), we identified a stable and soluble proteolytic fragment of TSC2 comprising residues 70–530 (referred to as TSC2-N; Fig. 1a). When co-expressed recombinantly in Escherichia coli (Fig. 1b), a stable complex of full-length TSC1 (TSC1-fl, residues 1–820) and TSC2-N could be co-purified with affinity chromatography, showing that TSC2-N confers interaction with TSC1. In contrast, no complex between TSC1-N (1–415) and TSC2-N was observed. Note that much more TSC2-N can be purified if no TSC1 is bound, likely because of the reduced solubility of the complex compared with the isolated protein. We also observe that TSC1-fl, which is sensitive to C-terminal degradation when expressed alone, is stabilized by co-expression with TSC2-N. This is indirect evidence for TSC2-N binding to the C terminus of TSC1.

Because with recombinant complex purification we were limited in the choice of fragments by the insolubility of some constructs, we further mapped the TSC1-TSC2 interaction in a cell culture system. HEK293 cells were transiently transfected with FLAG-TSC1 and HA-TSC2 constructs. Both full-length TSC2 (TSC2-fl) and TSC2N 70–530 bound to TSC1-fl with comparable efficiency (Fig. 1c). This confirms that the identified fragment represents the TSC1 interaction domain of TSC2. We next determined the interaction domain in TSC1 required for TSC2 association (Fig. 1d). Like hamartin, TSC1 from C. thermophilum contains a predicted N-terminal HEAT repeat domain and a C-terminal coiled-coil domain that are connected by a putative disordered linker. No interaction was detected between TSC1-N and full-length TSC2, but TSC1-C with and without the linker region (415–820 and 501–820, respectively) bound robustly. Additionally, we performed co-immunoprecipitations with human tuberin and hamartin and also observe that the N terminus of tuberin and the C terminus of hamartin are sufficient for complex formation (Fig. 1e). Thus, the interactions mapped are conserved for C. thermophilum proteins and the human homologues, suggesting identical complex architectures.

Crystal Structure of TSC2 N Terminus (TSC2-N)—The mapped TSC1-binding domain TSC2-N expressed well in E. coli and could be purified, yielding homogenous monomeric

Crystal Structure of TSC2 N Terminus

The mapped TSC1-binding domain TSC2-N expressed well in E. coli and could be purified, yielding homogenous monomeric

FIGURE 2. Structure of TSC2-N. a, limited proteolysis of TSC2 70–800 with trypsin led to the identification of a stable fragment, marked by an asterisk. b, TSC2-N (70–530) was loaded onto a Superdex 200 HiLoad 16/600 column equilibrated with 25 mM HEPES, 500 mM NaCl, 1 mM TCEP, pH 7.0, and 5% glycerol. The protein elutes as a monomer at 77 ml, corresponding to an apparent molecular mass of 67 kDa. c, top view (upper panel) and side view (lower panel) of the TSC2 N-terminal domain. Nine HRs and an additional C-terminal helix α19 are labeled. The structure can be divided into two subdomains, colored in light and dark cyan, respectively. d, representative 2Fo – Fc electron density map at the 2σ contour level of chain A HEAT repeat 2 (residues 124–159). e, structurally related proteins identified by a Dali search (33) in rainbow color code from N (blue) to C terminus (red): importin α3 (Protein Data Bank code 4UAE), TSC1-N (Protein Data Bank code 4KK1), and HspBP1 (Protein Data Bank code 1XQR).
TABLE 1
Data processing and refinement statistics for the structure of TSC2-N

|                  | TSC2-N (5HIU) | SeMet TSC2-N (peak) |
|------------------|---------------|---------------------|
| Space group      | P1            | P1                  |
| Wavelength (Å)   | 0.97626       | 0.97977             |
| Unit cell        |               |                     |
| a (Å)            | 42.7          | 43.7                |
| b (Å)            | 78.1          | 155.0               |
| c (Å)            | 153.8         | 78.8                |
| α (°)            | 89.990        | 90.0                |
| β (°)            | 89.9          | 101.4               |
| γ (°)            | 78.9          | 90.1                |
| Resolution (Å)   | 41.9–2.3 (2.42–2.3) | 43.0–2.5 (2.64–2.5) |
| Unique reflections | 83,768 (12,118) | 69,157 (10,004)     |
| Multiplicity     | 3.6 (3.5)     | 3.9 (3.8)           |
| I/σI             | 21.8 (2.7)    | 127.2 (2.9)         |
| Completeness (%) | 96.6 (95.8)   | 98.5 (97.2)         |
| Rmax (%)         | 3.6 (61.1)    | 5.6 (42.0)          |
| Wilson B factor (Å²) | 58.4        | 51.9                |
| Refinement (Å)   | 41.9–2.5      |                     |
| Reflections      | 62038         |                     |
| Rwork/Rfree (%)  | 23.3/26.2     |                     |
| rmsd bond distances (Å) | 0.008     |                     |
| rmsd bond angles (°) | 1.224     |                     |
| Mean B value (Å²) | 88.5         |                     |
| Ramachandran diagram (%) | 92.4     |                     |
| Most favored     | 92.4          |                     |
| Additionally allowed | 7.3         |                     |

* Values in parentheses refer to outer shell of reflections.

We obtained crystals that diffracted anisotropically to 2.3 Å, and selenomethionine-substituted crystals were grown for phase determination. The model was built and refined against native data to 2.5 Å (Table 1). The asymmetric unit contains four molecules with only small structural variations: molecules A and B show an rmsd of 0.016 Å, molecules C and D show an rmsd of 0.014 Å, and the rmsd between A or B versus C or D is 1.14 Å. The final models contain residues 88–527 with variable loop regions missing in the individual copies of TSC2-N because of the differential packing. Combining the information from the different copies of the asymmetric unit, we could generate a composite model lacking only the loop regions 232–242, 436–438, and 467–474 (Fig. 2, c, d, and e). TSC2-N forms an α-solenoid, consisting of 19 helices that are arranged in nine HEAT repeats (HR) and an additional C-terminal helix. Two subdomains can be defined that are separated by a short loop between HR 6 and 7, leading to a 75° tilt of the overall structure. A prediction for the N terminus of tuberin (32, 36) shows that, like the TSC2-N domain from C. thermophilum, it is probably composed of 19 α-helices. In a multiple sequence alignment, the predicted helices in tuberin N terminus (tuberin-N, residues 40–415) match very well with the secondary structure elements observed in TSC2-N (Fig. 3a). Based on this, we generated a model for tuberin-N with SWISS-MODEL/DeepView (37, 38) (Fig. 3, b and c). Tuberin-N was also modeled with nine HEAT repeats and an extra C-terminal helix but lacking the extensive loop regions in HR 4 and 9.

Membrane Interaction of TSC2-N—Active TSC is located on lysosomes, but its mechanism of membrane recruitment remains elusive. We generated an electrostatic potential map (39) of TSC2-N to identify potential membrane interaction interfaces, which often are represented by positively charged patches (Fig. 4a). Charges on TSC2-N are rather evenly distributed, with only one slightly basic patch at helix α14 that is also present at an equivalent position in tuberin-N (Fig. 4b). We did not observe a specific interaction of TSC2-N with a particular lipid in a protein lipid overlay assay (Fig. 4c). However, in a liposome sedimentation assay, TSC2-N bound specifically to a vacuolar lipid mix, but not to liposomes composed of neutral lipids (Fig. 4, d and e). The vacuolar SNARE Vam7, which binds to phosphoinositol 3-phosphate (PI(3)P) via a PX domain, served as a positive control, and TSC2-N binding to liposomes was inefficient in comparison. This indicates that TSC2-N has a weak affinity for negatively charged membranes.

Analysis of the TSC2-N Surface—To identify putative functionally important surface patches on TSC2-N, we mapped conserved residues on the structure (Fig. 5a). Most conserved residues are located in the interior of the protein and involved in stabilizing the domain fold. In addition, an accumulation of conserved residues can be found at the base of HEAT repeats 3–5, highlighting a potential functional interface (Fig. 5a). To test whether this interface might represent the binding site for TSC1 on TSC2-N, we designed three sets of mutations of conserved surface residues (set 1: K263E, F264A, and F307A; set 2: N194A and D260A; and set 3: P159T and R166A). There was no negative influence on the overall structure and stability of these mutants in comparison with wild-type TSC2-N detectable by CD spectroscopy (Fig. 5b) or differential scanning fluorimetry (Fig. 5c). The co-purification of TSC1 was impaired with the set 3 mutant (Fig. 5, d and e), suggesting that these residues play a role in mediating the contact between TSC1 and TSC2.

Characterization of Pathogenic Point Mutations in Tuberin—More than 2200 unique mutations in the TSC2 gene are reported in the tuberous sclerosis database of the Leiden Open Variation Database (29). We searched the database for missense point mutations reported or concluded as pathogenic or probably pathogenic and mapping to the region covered by our structure. Deletion, insertion, frame shifts, and stop mutations that likely lead to misfolded protein were excluded. The search yielded a list of 33 mutations (Table 2), which can be mapped on the model of tuberin-N and classified by their likely effect on protein structure (Fig. 6a). Six of the identified mutations (shown in green) do map to the surface of tuberin-N and should not affect folding. Such mutations can be suspected to interfere with molecular functions of tuberin, like its interaction with hamartin. In contrast, 11 changes (labeled in red) are of residues...
FIGURE 3. Modeling of the tuberin N terminus. a, a multiple sequence alignment of the N-terminal region of TSC2 from C. thermophilum, S. pombe, D. melanogaster, and humans was generated with ClustalW (57) and visualized with ESPript (60). Similar and conserved positions are marked by yellow and red boxes, respectively. Secondary structure elements of TSC2-N are labeled and shown above the alignment. Below the sequence of human tuberin-N, secondary structure is indicated as predicted by the PSIPRED server (32). Residues mutated in this study are marked by triangles (surface residues), asterisks (structural mutations), or numbers (for the corresponding set mutations). b, based on the x-ray structure of TSC2-N from C. thermophilum. c, a model for the N-terminal domain of human tuberin was created. The structures are rainbow color-coded from N terminus (blue) to C terminus (red).
located in $\alpha$-helices to proline. Such substitutions will be helix-breaking, might destroy the secondary structure, and might lead to destabilized or misfolded protein. Another class of structural mutants could affect intramolecular packing (labeled in yellow): exchange of residues that constitute the hydrophobic core of tuberin-N to amino acids with side chains of different shape or polarity also leads to destabilization and/or partial unfolding of the domain. Although amino acid changes in the protein core can be tolerated to some extent, we suggest that the disease-causing effect of these structural mutations could be explained by improper folding of tuberin-N and, consequently, dysfunctional protein.

To support our classification of pathogenic mutations, we generated substitutions in \textit{C. thermophilum} TSC2-N equivalent to the pathogenic surface mutations in tuberin-N (Table 2). These include the surface mutants L122G, S341W, R361A, R420A, and V464P/T465A/L466A. In addition to Val$^{464}$ $\rightarrow$ Pro, corresponding to Gln$^{373}$ in tuberin, we changed the subsequent highly conserved residues Thr$^{465}$ and Leu$^{466}$ to alanine because they might also be functionally important. We also tested a subset of seven mutations (L192R, N265I, S341P, G381E, A411P, L428R, and L497P) that were classified as structural and are conserved between \textit{C. thermophilum} and humans. Five of the tested structural mutants were insoluble, aggregated, or degraded during purification (Table 2). The variants L192R and N265I could be purified but showed reduced stability and a lower melting point in differential scanning fluorimetry (Fig. 6b). In contrast, the melting temperature of all surface mutants was comparable with wild-type TSC2-N (Fig. 6b) and CD spectra also did not indicate any severe structural defects (Fig. 6c), suggesting that their structures are unaffected by the mutations.
Because a structural defect of the surface mutants likely is not the cause of their pathogenicity, we asked whether binding to TSC1 might be impaired. Using our co-purification assay, we tested these mutants for their ability to interact with TSC1-fl (Fig. 6, d and e). Indeed, L122G showed a binding defect, but the remaining mutants bound TSC1 like wild type.

To confirm that the human pathogenic point mutation E75G, which is equivalent to L122G in C. thermophilum TSC2, indeed disrupts the complex, we tested binding of hamartin to tuberin in co-immunoprecipitation studies (Fig. 7, a and b). The interaction of tuberin E75G to hamartin was impaired when compared with the wild-type protein. Also, the mutation R261P, which we had classified as structural and where we showed that the equivalent substitution S341P destroys TSC2-N structure, had reduced interaction with hamartin. In contrast, the surface mutation tuberin R261W did not affect TSC1 association, consistent with the results for the corresponding mutant S341W in C. thermophilum TSC2.

Thus, despite their low sequence identity, the interaction site between TSC1 and TSC2 is conserved for tuberin and hamartin. When we map the location of mutations that affect TSC complex assembly in the structure of TSC2-N, strikingly, both the pathogenic mutation and set 3 residues are found in the same region of TSC2-N at the base of HEAT repeats 1–3 (Fig. 7c). Thus, we identified the binding site for TSC1 on TSC2 and pinpoint a pathogenic surface mutation to this region.

**Discussion**

The crystallographic analysis of the N-terminal domain of TSC2 from C. thermophilum provides the first structural infor-
mation of the TSC2 subunit of the TSC complex. Our work reveals that this domain adopts a HEAT repeat fold that also serves as a template to generate a model of the N terminus of human tuberin.

In the active state, TSC has to be recruited to its substrate Rheb onto the lysosomal membrane. A possible targeting mechanism is the interaction with specific membrane lipids. In the case of lysosomes, phosphoinositols 3,5-bisphosphate (PI(3,5)P_2) is considered a marker lipid, and PH (pleckstrin homology) domains have been reported as PI(3,5)P_2-binding modules (40, 41). TSC2-N does not show structural homology to PH domains but might represent a novel PI(3,5)P_2 interaction module (40, 41). TSC2-N does not show specificity for an organelle marker lipid, we expect that additional factors will play a role in membrane targeting could involve protein-protein interactions with a receptor on the lysosomal surface, which warrants further investigation. In this scenario, TSC2-N membrane binding could contribute to lysosomal attachment, which likely requires multiple TSC domains.

We further determined that TSC2-N represents an interaction domain for the assembly of the TSC complex. In both co-immunoprecipitation from cell culture and co-expression studies in E. coli, TSC2-N was sufficient to form a stable complex with TSC1. Vice versa, the C-terminal domain, but not the N terminus of TSC1, bound to TSC2, demonstrating that TSC1-C and TSC2-N represent the key domains responsible for TSC assembly. Importantly, we confirmed the same interaction domains in the human orthologues hamartin (C terminus) and tuberin (N terminus). Consistent with our findings, previous studies with tuberin and hamartin using yeast two-hybrid analysis (13, 14) also suggested that N-terminal regions of TSC2 are involved in binding TSC1. There was, however, a discrepancy over which elements of hamartin mediate complex assembly, suggesting either the C terminus (13), a fragment from the N-terminal half of hamartin (302–430) (14), or binding sites in both regions (16). Our study, probing for direct interactions and reconstituting a TSC subcomplex, shows that TSC1-C and hamartin-C are sufficient to form a complex with TSC2/tuberin (Fig. 1). We were unable to detect any stable interaction between TSC2-N and TSC1-N in co-immunoprecipitation, co-expression, or reconstitution experiments with individually purified proteins. However, because the TSC complex has been described as a higher oligomeric assembly, it may be stabilized by additional secondary interaction sites. Because our experiments were done at expression levels above the endogenous concentrations, these additional sites may be dispensable in vitro but could be of relevance in a physiological setting.

Based on our structure of TSC2-N and the model of tuberin-N, we ought to analyze pathogenic TSC mutations that map to this region. We classified mutations that map to the interior of the protein or proline mutations in secondary structure elements as structural and suggest that their pathogenicity might be explained by reduced protein stability or folding defects. We confirmed for seven positions that are conserved in TSC2 and tuberin that protein stability of these mutants is indeed impaired. Some conservative substitutions in the protein core might be tolerated, and we found for the core residue mutation L320F (L403F in tuberin) that the structure was indeed unaffected (Fig. 6c). However, our findings back the notion that for patient mutations in the protein core, the pathogenicity could be explained by problems with protein structure. Further support comes from the finding that the structural mutation R261P in tuberin also showed a functional defect and impaired binding to hamartin in co-immunoprecipitation assays.

In contrast, surface mutations on TSC2 did not affect protein structure. Their pathogenicity could therefore be caused by impaired molecular functions, like TSC1 binding. The pathogenic mutation L122G does not affect protein structure but specifically leads to loss of TSC1 binding, whereas other mutations (e.g. S341W) had no effect. The equivalent mutations in tuberin, E75G and R261W, respectively, had the same differential effect on complex assembly (Fig. 7a). Consistently, a functional assessment of TSC2 variants had shown that mTORC1 activity is elevated for the mutation E75G but comparable with wild type for R261W (27, 28).

Surprisingly, the residue of the pathogenic mutation E75G in the human protein is not conserved, and the chemical character of the equivalent position Leu^{122} is different in C. thermophi-
The fact that we nevertheless observe that the effect on complex assembly is preserved could be explained by the mutation to glycine, which does not affect the overall structure but could locally destabilize the loop region with detrimental consequences on TSC1 binding.

Previously, pathogenic and non-pathogenic mutations in tuberin had been tested for their influence on the interaction with hamartin (14). The pathogenic mutations G294E and the in-frame deletion I365del led to a loss of binding, and our tuberin model suggests that these changes likely produce unfolded protein and therefore have an indirect effect. Gly294 is located in helix α14 pointing toward the interior of the protein, and substitution with glutamate will disrupt the fold. Ile365 is in the middle of helix α17 and part of the hydrophobic core. Its deletion will result in a destabilizing shift in helix register. The non-pathogenic mutations R261W, M286V, and R367Q were reported as not effecting hamartin binding and, consistently, are surface-exposed residues in our model. R261W was described to not cause tuberous sclerosis (43) but was identified in patients with pulmonary lymphangioleiomyomatosis (44).

FIGURE 6. Role of pathogenic mutations on TSC2-N functionality. a, pathogenic missense mutations mapped onto the model of tuberin-N. Mutated residues are shown with spheres and are colored according to their structural classification: red mutations are helix breaking, yellow mutations affect intramolecular packing, and green mutations are surface-exposed residues. b, stability of selected pathogenic mutations in comparison with wild-type TSC2-N was tested by determining the melting temperature with differential scanning fluorimetry. Surface mutants are of solvent-exposed residues, whereas structural mutants are of residues involved in intramolecular interactions. c, protein folding of the pathogenic surface mutants compared with wild-type TSC2-N is not affected as monitored by circular dichroism spectroscopy. d, representative complex formation assay between TSC1-fl and TSC2-N by purification of the recombinant complex. TSC2-N wild type and indicated mutants were co-expressed with TSC1-fl and purified via GSH affinity pulldown. e, quantification of the complex purification assay from three biological repeats.
Because the mutation R261W is located on the surface of tuberin, does not affect recruitment of hamartin (Fig. 7a) (14), and does not alter mTORC1 activity (27), it could play a role in the interaction with an unknown ligand that is involved in the development of lymphangioleiomyomatosis.

We also identified a patch on the surface of TSC2-N that is required for interaction with TSC1 based on sequence conservation. This patch of set 3 mutations, like L122G, maps to the linker regions between HEAT repeats 1/2 and 2/3. The residues are located within 15 Å, identifying this area as one molecular surface that mediates the interaction with TSC1. However, because the interacting part of TSC1 is predicted to form a likely extended coiled-coil structure, it is possible that the actual interface is rather extensive.

Taken together, we conclude that our tuberin-N model is fundamentally correct, and we find that the insight obtained from the mutational studies of C. thermophilum proteins is highly relevant to understand pathological mutations in the human proteins. Structural studies on tuberin and hamartin have been problematic in the past. Working with proteins from C. thermophilum proves to be a useful approach to obtain structural information on these medically relevant proteins from a better suited organism. Although further investigations will be required, our results clearly demonstrate that structural studies on the TSC complex can guide characterizing the diverse and variable spectrum of symptoms observed in tuberous sclerosis.

**Experimental Procedures**

**Cloning and Mutagenesis**—Constructs of TSC1 and TSC2 were amplified from codon optimized synthetic genes (GeneScript) using Q5 Polymerase (New England Biolabs). DNA fragments were cloned into a modified pCDF (Novagen) vector for expression with an N-terminal GST tag and PreScission protease cleavage site or a modified pET28 (Novagen) with N-terminal HisC-SUMO (small ubiquitin-like modifier) tag. For cell culture expression, pCDNA3 vectors with 3× FLAG or HA tag were used. Mutations were generated according to the megaprime protocol (45).

**Interaction Studies**—For mammalian expression, constructs were cloned into pCDNA3 vectors with 3× FLAG or HA tag and introduced into HEK293 cells through calcium phosphate-mediated transfection. After 24 h, transfected HEK293 cells were lysed in co-immunoprecipitation buffer (25 mM HEPES, pH 7.5, 150 mM NaCl, 0.2% Nonidet P-40, 10% glycerol, 1 mM DTT plus protease inhibitors), and protein complexes were precipitated for 4 h with anti-FLAG antibody (M2; Sigma) and protein A/G-agarose. Immunoprecipitates and lysate samples were run on SDS-PAGE, and proteins were detected after immunoblotting with anti-FLAG (M2; Sigma) or anti-HA (Y-11; Santa Cruz) antibodies.

To probe for complex formation of recombinant proteins, HisC-SUMO-TSC1-fl (full-length) and GST-TSC2-N wild type or mutants were co-expressed in E. coli Rosetta (DE3) (Novagen) after cold shock overnight at 16 °C. Selenomethionine substituted TSC2-N was produced with the same conditions using feedback inhibition of methionine synthesis (46). Cells were resuspended in lysis buffer, and the soluble supernatant was incubated with GSH-agarose (Thermo Scientific). After washing, complexes were eluted off the beads by the addition of SUMO protease and PreScission. Equal amounts of eluted TSC2-N were loaded on SDS-PAGE gels, and proteins were visualized with Coomassie staining.

Western blots and Coomassie-stained gels were scanned with an Odyssey system (LI-COR) and quantified with Image Studio Light. Significance analysis was performed by t-test statistics (*, p < 0.05; **, p < 0.01; ***, p < 0.001).

**Protein Expression and Purification**—Proteins were expressed alone or in combination in E. coli Rosetta (DE3) (Novagen) after cold shock overnight at 16 °C. Selenomethionine substituted TSC2-N was produced with the same conditions using feedback inhibition of methionine synthesis (46). Cells were resuspended in lysis buffer (50 mM NaH2PO4, pH 7.5, 300–500 mM NaCl, 10 mM imidazole) supplemented with EDTA-free protease inhibitor, DNase, and lysozyme (0.3 mg ml−1), and lysis was performed with a microfluidizer (M110L; Microfluidics) or sonication (Branson Sonifier 250). His-tagged proteins were captured via gravity flow chromatography using nickel-nitrilotriacetic acid resin, GST-tagged proteins were purified by applying supernatant to glutathione superflow agarose (Thermo Scientific). The
matrix was washed with 100 ml of lysis buffer. Proteins bound to nickel–nitrioltriacetic acid-agarose were additionally washed with lysis buffer containing increasing imidazole concentration (20 and 50 mM). GST-tagged proteins were incubated overnight at 4 °C with PreScission- and His-tagged proteins with SUMO proteases to cleave off the tags. Target proteins were eluted and dialyzed against 10 ml of lysis buffer including 20 mM GSH. PIP MicroStrips were incubated with trypsin at different ratios (1:1600, 1:800, 1:400, 1:160, 1:80, and 1:40 (w/w)) in 25 mM Hepes, 300 mM NaCl, 1 mM DTT (pH 7.0), and peak fractions were pooled.

Limited Proteolysis—The construct TSC2 70–800 (0.8 mg ml⁻¹) was incubated with trypsin at different ratios (1:1600, 1:800, 1:400, 1:160, 1:80, and 1:40 (w/w)) in 25 mM Hepes, 300 mM NaCl, 1 mM DTT (pH 7.0), and 5% glycerol for 1 h at 37 °C. The reaction was stopped by adding SDS loading dye and boiling the samples for 3 min, and samples were analyzed via SDS-PAGE. A stable degradation fragment was cut out of the gel, destained, and digested with trypsin (sequencing grade; Promega) overnight at 37 °C. Peptides were identified with the Mascot search engine tool after applying HPLC and electrospray ionization-electron-transfer dissociation-ion trap analysis (Bruker).

Crystallization and Structure Determination—Initial crystallization conditions were identified with sitting drop vapor diffusion experiments using a semi-automated dispensing system (Gryphon; Art Robinson). The best crystals were obtained after several rounds of microseeding (47) with a protein concentration of 6 mg ml⁻¹ and 100 mM Tris (pH 7.0), 14% PEG 6000, 15% glycerol, 200 mM MgCl₂ as reservoir solution at 20 °C. Selonemethionine-derivative crystals were also obtained by microseeding. Prior to data collection, the crystals were briefly soaked in a cryo-protection condition containing 24% instead of 14% PEG 6000 and flash cooled in liquid nitrogen. A native data set of TSC2-N was measured at European Synchrotron Radiation Facility ID23-1 (Grenoble, France), and anomalous data were collected at BESSY II BL14.1 (Helmholtz-Zentrum Berlin, Germany) (48). The collected data were processed using XDSAPP (49, 50) and SCALA (51). Initial phases were determined with Phenix/AutoSol (52), followed by automated model building using ARP/wARP (53). The final model was obtained through iterative cycles of model building in COOT (54) and refinement against native data using REFMAC5 (55, 56). Structure factors and the final model have been deposited in the Protein Data Bank (code 5HIU). A homology model of the human tuberin N terminus was generated based on a ClustalW (57) sequence alignment with a DeepView project (38) in SWISS-MODEL (37). All images of molecular structures were created using PyMOL (58).

PIP Strip Assay—GST-tagged TSC2-N, as well as GST alone, were expressed and purified like described above but eluted with 10 ml of lysis buffer including 20 mM GSH. PIP MicroStrips were incubated with trypsin at different ratios (1:1600, 1:800, 1:400, 1:160, 1:80, and 1:40 (w/w)) in 25 mM Hepes, 300 mM NaCl, 1 mM DTT (pH 7.0), and peak fractions were pooled. The reaction was stopped by adding SDS loading dye and boiling the samples for 3 min, and samples were analyzed via SDS-PAGE. A stable degradation fragment was cut out of the gel, destained, and digested with trypsin (sequencing grade; Promega) overnight at 37 °C. Peptides were identified with the Mascot search engine tool after applying HPLC and electrospray ionization-electron-transfer dissociation-ion trap analysis (Bruker).

Crystal Structure of the TSC2 N Terminus

References
1. Shimobyashi, M., and Hall, M. N. (2014) Making new contacts: the mTOR network in metabolism and signalling crosstalk. Nat. Rev. Mol. Cell Biol. 15, 155–162
2. Dibble, C. C., and Cantley, L. C. (2015) Regulation of mTORC1 by PI3K signaling. Trends Cell Biol. 25, 545–555
3. Inoki, K., Li, Y., Xu, T., Guan, K.-L. (2003) Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. Genes Dev. 17, 1829–1834
tuberous sclerosis complex (TSC1)/TSC2 in C2C12 myotubes.

**Crystal Structure of the TSC2 N Terminus**

4. Zoneu, R., Efeyan, A., and Sabatini, D. M. (2011) mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat. Rev. Mol. Cell Biol.* **12**, 21–35
5. van Slegtenhorst, M., de Hoogt, R., Hermans, C., Nellist, M., Janssen, B., Verhoef, S., Lindhout, D., van den Ouweland, A., Halley, D., Young, J., Burley, M., Jeremiah, S., Woodward, K., Nahmias, J., Fox, M., et al. (1997) Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science* **277**, 805–808
6. European Chromosome 16 Tuberous Sclerosis Consortium (1993) Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* **75**, 1305–1315
7. Dibble, C. C., Els, W., Menon, S., Qin, W., Klekota, J., Asara, J. M., Finan, P. M., Kwiatkowski, D. J., Murphy, L. O., and Manning, B. D. (2012) TBC1D7 is a third subunit of the TSC1-TSC2 complex upstream of mTORC1. *Mol. Cell* **47**, 535–546
8. Nakashima, A., Yoshino, K., Miyamoto, T., Eguchi, S., Oshiro, N., Kikkawa, U., and Yonezawa, K. (2007) Identification of TBC7 having TBC domain as a novel binding protein to TSC1-TSC2 complex. *Biochem. Biophys. Res. Commun.* **361**, 218–223
9. Hoogeveen-Westerveld, M., van Unen, L., van den Ouweland, A., Halley, D., Hoogeveen, A., and Nellist, M. (2012) The TSC1-TSC2 complex consists of multiple TSC1 and TSC2 subunits. *BMC Biol.* **13**, 18
10. Shirakawa, R., Fukui, S., Kawato, M., Higashi, T., Kondo, H., Ikeda, T., Nakayama, E., Okawa, K., Nureki, O., Kimura, T., Kita, T., and Horiechi, H. (2009) Tuberous sclerosis tumor suppressor complex-like complexes act as GTPase-activating proteins for Ral GTPases. *J. Biol. Chem.* **284**, 21580–21588
11. Gridley, S., Chavez, J. A., Lane, W. S., and Lienhard, G. E. (2006) Adipocytes contain a novel complex similar to the tuberous sclerosis complex. *Cell Signal.* **18**, 1626–1632
12. Scrima, A., Thomas, C., Deaconescu, D., and Wittinghofer, A. (2008) The Rap-RapGAP complex: GTP hydrolysis without catalytic glutamate and arginine residues. *EMBO J.* **27**, 1145–1153
13. van Slegtenhorst, M., Nellist, M., Nagelkerken, B., Cheadle, J., Snell, R., van den Ouweland, A., Reuser, A., Sampson, J., Halley, D., and van der Sluijs, P. (1998) Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products. *Hum. Mol. Genet.* **7**, 1053–1057
14. Hodges, A. K., Li, S., Maynard, J., Parry, L., Braverman, R., Cheadle, J. P., DeClue, J. E., and Sampson, J. R. (2001) Pathological mutations in TSC1 and TSC2 disrupt the interaction between hamartin and tuberin. *Hum. Mol. Genet.* **10**, 2899–2905
15. Momose, S., Kobayashi, T., Tada, N., Itoyama, S., and Hino, O. (2007) N-terminal hamartin-binding and C-terminal GAP domain of tuberin can separate in vivo. *Biochem. Biophys. Res. Commun.* **356**, 693–698
16. Santiago Lima, A. J., Hoogeveen-Westerveld, M., Nakashima, A., Maat-Kievit, A., van den Ouweland, A., Halley, D., Kikkawa, U., and Nellist, M. (2014) Identification of regions critical for the integrity of the TSC1-TSC2-TBC1D7 complex. *PloS One* **9**, e93940
17. Yoshimura, S., Egerer, J., Fuchs, E., Haas, A. K., and Barr, F. A. (2007) Functional dissection of Rab GTPases involved in primary cilium formation. *J. Cell Biol.* **178**, 363–369
18. Gal, Z., Chu, W., Deng, W., Li, W., Li, H., He, A., Nellist, M., and Wu, G. (2016) Structure of the TBC1D7-TSC1 complex reveals that TBC1D7 stabilizes dimerization of the TSC1 C-terminal coiled coil region. *J. Mol. Biol. Cell Biol.* **10.1093/jmcb/mjw001**
19. Qin, J., Wang, Z., Hoogeveen-Westerveld, M., Shen, G., Gong, W., Nellist, M., and Xu, W. (2016) Structural basis of the interaction between tuberous sclerosis complex 1 (TSC1) and Tre2-Bub2-Cdc16 domain family member 7 (TBC1D7). *J. Biol. Chem.* **291**, 8591–8601
20. Demetriades, C., Dounpam, N., and Telemann, A. A. (2014) Regulation of TORC1 in response to amino acid starvation via lysosomal recruitment of TSC2. *Cell* **156**, 786–799
21. Menon, S., Dibble, C. C., Tallbot, G., Hoxjah, G., Valvezan, A. J., Takanashi, H., Cantly, L. C., and Manning, B. D. (2014) Spatial control of the TSC complex integrates insulin and nutrient regulation of mTORC1 at the lysosome. *Cell* **156**, 771–785
22. Miyazaki, M., McCarthy, J. J., and Esser, K. A. (2010) Insulin like growth factor-1-induced phosphorylation and altered distribution of
Analysis of all exons of TSC1 and TSC2 genes for germline mutations in Japanese patients with tuberous sclerosis: report of 10 mutations. *Am. J. Med. Genet.* **90**, 123–126

Jones, A. C., Shyamsundar, M. M., Thomas, M. W., Maynard, J., Idziaszczyk, S., Tomkins, S., Sampson, J. R., and Cheadle, J. P. (1999) Comprehensive mutation analysis of TSC1 and TSC2 and phenotypic correlations in 150 families with tuberous sclerosis. *Am. J. Hum. Genet.* **64**, 1305–1315

Badri, K. R., Gao, L., Hyjek, E., Schuger, N., Schuger, L., Qin, W., Chekaluk, Y., Kwiatkowski, D. J., and Zhe, X. (2013) Exonic mutations of TSC2/TSC1 are common but not seen in all sporadic pulmonary lymphangiectiomatosis. *Am. J. Respir. Crit. Care Med.* **187**, 663–665

Ke, S. H., and Madison, E. L. (1997) Rapid and efficient site-directed mutagenesis by single-tube “megaprimer” PCR method. *Nucleic Acids Res.* **25**, 3371–3372

Doublié, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530

D’Arcy, A., Villard, F., and Marsh, M. (2007) An automated microseed matrix-screening method for protein crystallization. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 550–554

Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G., and Weiss, M. S. (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J. Synchrotron Radiat.* **19**, 442–449

Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132

Krug, M., Weiss, M. S., Heinemann, U., and Mueller, U. (2012) XDSAPP: A graphical user interface for the convenient processing of diffraction data using XDS. *J. Appl. Crystallogr.* **45**, 568–572

Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol Crystallogr.* **50**, 760–763

Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L. W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 582–601

Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamm, V. S. (2001) ARP/wARP and molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1445–1450

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

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255

Joosten, R. P., Long, F., Murshedov, G. N., and Perrakis, A. (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680

Delano, W. L. (2003) The PyMOL Molecular Graphics System, Delano Scientific, San Mateo, CA

Cabrera, M., Nordmann, M., Perz, A., Schmidt, D., Gerondopulos, A., Barr, F., Piehler, J., Engelbrecht-Vandré, S., and Ungermann, C. (2014) The Mon1-Ccz1 GEF activates the Rab7 GTPase Ypt7 via a longin-fold-Rab interface and association with PI3P-positive membranes. *J. Cell Sci.* **127**, 1043–1051

Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDSrver. *Nucleic Acids Res.* **42**, W320–W324