



# Identification of compounds that bind the centriolar protein SAS-6 and inhibit its oligomerization

Received for publication, June 12, 2020, and in revised form, August 26, 2020. Published, Papers in Press, September 1, 2020. DOI 10.1074/jbc.RA120.014780

Julia M. C. Busch<sup>1</sup>, Minos-Timotheos Matsoukas<sup>2</sup>, Maria Musgaard<sup>1</sup>, Georgios A. Spyroulias<sup>2</sup>, Philip C. Biggin<sup>1</sup>, and Ioannis Vakonakis<sup>1,\*</sup>

From the <sup>1</sup>Department of Biochemistry, University of Oxford, Oxford, United Kingdom and the <sup>2</sup>Department of Pharmacy, University of Patras, Patras, Greece

Edited by Enrique M. De La Cruz

Centrioles are key eukaryotic organelles that are responsible for the formation of cilia and flagella, and for organizing the microtubule network and the mitotic spindle in animals. Centriole assembly requires oligomerization of the essential protein spindle assembly abnormal 6 (SAS-6), which forms a structural scaffold templating the organization of further organelle components. A dimerization interaction between SAS-6 N-terminal “head” domains was previously shown to be essential for protein oligomerization *in vitro* and for function in centriole assembly. Here, we developed a pharmacophore model allowing us to assemble a library of low-molecular-weight ligands predicted to bind the SAS-6 head domain and inhibit protein oligomerization. We demonstrate using NMR spectroscopy that a ligand from this family binds at the head domain dimerization site of algae, nematode, and human SAS-6 variants, but also that another ligand specifically recognizes human SAS-6. Atomistic molecular dynamics simulations starting from SAS-6 head domain crystallographic structures, including that of the human head domain which we now resolve, suggest that ligand specificity derives from favorable Van der Waals interactions with a hydrophobic cavity at the dimerization site.

Centrioles are organelles that are widespread in eukaryotes (1–3) and coordinate a broad spectrum of cellular activities. In animal cells, a pair of centrioles close to the nucleus compose the structured core of the centrosome, which acts as the major microtubule-nucleating center of the cell (4, 5). Centrosomes nucleate microtubules and direct formation of their network, which is essential for intracellular transport processes and organelle positioning relative to the nucleus, and organize the mitotic spindle during cell division, thereby assisting the correct segregation of genetic material. Moreover, centrioles in nonproliferating animal cells and in unicellular eukaryotes dock to the cell membrane, where they act as templating platforms for cilia and flagella involved in motility and sensing (6). The number of centrioles in cells is maintained constant through new organelle assembly at each cell cycle (7, 8), which starts with recruit-

ment of the spindle assembly abnormal protein 6 (SAS-6) to a site adjacent to each preexisting centriole.

SAS-6 is an essential protein for centriole formation (9–13) and a key component of a scaffold-like structure known as the “cartwheel” (14) or “central tube” (15). The cartwheel forms early in the centriole assembly process and subsequently organizes downstream organelle components (7, 8). Cartwheel formation by SAS-6 depends on large-scale oligomerization of this protein. Crystallographic, EM, atomic force microscopy, and biophysical studies of SAS-6 variants from multiple species have shown that SAS-6 oligomerization is driven by two dimerization interfaces mediated by the protein “head” domain and a homodimeric coiled-coil domain (16–23). These dimerization interfaces cooperate to form closed rings of SAS-6 proteins in most species or open-ended spiral arrangements in nematodes, both of which have been reconstituted *in vitro* (16–19, 21–23). In either case, WT SAS-6 rings and spirals feature 9-fold radial symmetry on average, which matches the radial symmetry of cartwheels and centrioles, whereas engineered SAS-6 variants that form oligomers with altered symmetry lead to the assembly of non9-fold-symmetric organelles (23). Thus, SAS-6 oligomers are thought to impart correct radial symmetry to centrioles, thereby ensuring the formation of functional organelles.

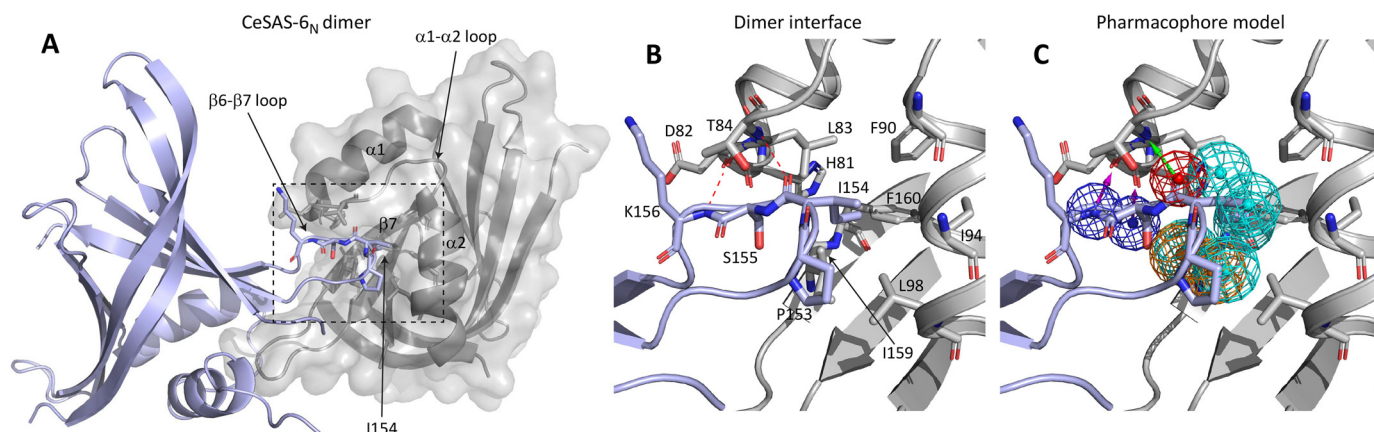
Centriole assembly is a tightly regulated process because the creation of too many or too few organelles can lead to ciliopathies, developmental abnormalities, genomic instability because of unipolar or multipolar mitotic spindles, and potentially cancer (24–28). The polo-like kinase 4, which initiates centriole duplication by controlling SAS-6 localization to the site of organelle assembly, is considered a promising target for cancer therapy where the underlying causes of the disease involve centrosome overamplification (29). To date, three different small molecule inhibitors selectively target polo-like kinase 4 with nanomolar affinity (30–32) and block cell proliferation in a number of human cancer lines, including osteosarcoma, cervical carcinoma, and breast, lung, and colon cancer cells (32–34), likely via centrosome depletion. However, expanding the repertoire of molecular targets for anti-cancer therapeutics is clearly desirable to combat cell desensitization to specific drugs. SAS-6 oligomerization may be such a molecular target, because SAS-6 variants disrupting protein oligomerization prevented centriole assembly in nematodes, algae, insects, and human cells (17–20).

Here, we evaluate the SAS-6 head domain as a target of small molecules that block domain dimerization and, thus, protein

This article contains [supporting information](#).

✂ Author's Choice—Final version open access under the terms of the Creative Commons [CC-BY](#) license.

\* For correspondence: Ioannis Vakonakis, [ioannis.vakonakis@bioch.ox.ac.uk](mailto:ioannis.vakonakis@bioch.ox.ac.uk). Present address for Maria Musgaard: Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Ontario, Canada



**Figure 1. A pharmacophore model of CeSAS-6<sub>N</sub> dimerization.** *A*, the dimer of CeSAS-6<sub>N</sub>, where the  $\beta 6$ - $\beta 7$  loop of one domain (light blue) protrudes into a hydrophobic pocket of the second domain (gray) formed by residues on  $\alpha 1$ ,  $\alpha 2$ , and  $\beta 7$ . Derived from PDB ID 3PYI. *B*, magnified view of the dimerization interface, corresponding to the boxed area in (*A*), with key residues annotated and hydrogen bonding interactions show as dashed red lines. *C*, a structure-based pharmacophore model depicting the interactions mapped. Hydrophobic features are depicted in cyan, aromatic in orange, hydrogen bond donors in blue, and a hydrogen bond acceptor in red. Arrows depict the direction of the hydrogen bond vectors (purple from donors and green from acceptor).

oligomerization. Guided by the SAS-6 head domain structure, we derived a library of chemical fragments that may bind to this domain and tested it against the human, nematode, and algae SAS-6 variants. We discovered one small molecule binding to all three SAS-6 variants, but also another that interacts only with human SAS-6. Combined with crystallographic, NMR, and atomistic molecular dynamics simulations, our work provides a starting point toward developing a specific inhibitor of SAS-6 oligomerization.

## Results

### A virtual screening approach to identify CeSAS-6 head domain interactors

The molecular architecture of SAS-6 proteins is remarkably conserved across species despite large sequence divergence (1), comprising an N-terminal globular head, a 30–50 nm-long coiled coil, and an unstructured C-terminal extension (17, 19). The SAS-6 coiled coil forms a homodimer with the sub- $\mu\text{M}$  dissociation constant ( $K_d$ ) via an extensive interaction interface and is sufficiently stable to mediate protein dimerization in cells (35). The SAS-6 head domain also self-associates but with lower affinity ( $K_d$  of 50–100  $\mu\text{M}$  in vertebrate, nematode, and algae SAS-6) and a smaller dimerization interface compared with the coiled-coil domain (17, 19). Thus, the optimal target for small molecules aimed at disrupting SAS-6 oligomerization may be the weaker and smaller dimerization interface of the SAS-6 head domain.

To identify compounds that would disrupt the SAS-6 head domain dimerization, we established a virtual screening process starting from the head domain dimer described for *Caenorhabditis elegans* SAS-6 (CeSAS-6<sub>N</sub>, residues 1–168) (17). Comparison of SAS-6 crystallographic structures suggests that the head domain mode of dimerization is highly conserved (Fig. S1). In the case of CeSAS-6<sub>N</sub>, dimerization involves the domain  $\beta 6$ - $\beta 7$  loop, which is inserted deeply into a hydrophobic cavity of the second monomer (Fig. 1, *A* and *B*). At the tip of  $\beta 6$ - $\beta 7$ , a single hydrophobic amino acid (Ile-154) is key for protein association. Ile-154

substitution by charged residues or glycine disrupts CeSAS-6<sub>N</sub> dimerization and abrogates protein function (17); conversely, Ile-154 substitution by a larger hydrophobic amino acid (I154W) enhances protein association by  $\sim 20$ -fold (16). Further elements of the dimerization interface are backbone interactions between loop residues Ser-155 and Lys-156 and between residues Asp-82 and Thr-84 in the  $\alpha 1$ - $\alpha 2$  linker that forms the upper rim of the dimerization hydrophobic pocket. Additionally, Pro-153 likely stabilizes Ile-154 positioning (Fig. 1*B*).

These elements of the dimerization interaction were used to inform a structure-based pharmacophore model in which the CeSAS-6 tetrapeptide Pro-Ile-Ser-Lys (residues 153–156) of the  $\beta 6$ - $\beta 7$  loop was considered as “ligand” and the second protein monomer as “receptor”. The pharmacophore model comprises a set of features, mapped in space, that if satisfied by a small molecule would be suggestive of CeSAS-6<sub>N</sub> binding. Features mapped included the two hydrophobic contacts of the Ile-154 aliphatic group with the side chains of Phe-90, Ile-94, Leu-98, Ile-159, and Phe-160, which comprise the hydrophobic cavity (Fig. 1*C*). Two hydrogen bond donors were mapped to interact with the carbonyl atoms of His-81 and Asp-82, as well as a hydrogen bond acceptor directed toward the Thr-84 NH atom. The pharmacophore was complemented with one more hydrophobic feature in the cavity, close to Leu-83, and two mixed aromatic ring/hydrophobic features interacting with Leu-98 and Ile-159. We prepared a ligand data set comprising  $\sim 6$  million commercially available compounds from the ZINC database, with 250 alternative conformations generated for each compound. All conformers were rigid body-fitted to subgroups of at minimum five features of the pharmacophore model (see “Experimental procedures”) and ranked by goodness of fit. The top-scored  $\sim 2,000$  compounds were further filtered according to chemical diversity using the Tanimoto coefficient (36) and by visual inspection to derive a final library of 37 compounds (A1–D1, Table S1) that may associate with the hydrophobic cavity at the CeSAS-6<sub>N</sub> dimerization interface.

## Inhibitors of SAS-6 oligomerization

### NMR-based screen confirms small molecule binding to CeSAS-6<sub>N</sub>

We initially assessed the solubility of compounds A1–D1 in aqueous buffer by visual inspection of precipitation following compound transfer from 100% v/v DMSO to 5% v/v DMSO and 2 mM final concentration. Further, we recorded <sup>1</sup>H NMR spectra of each compound in 5% v/v DMSO. Of the 37 compounds tested, eight were insoluble in aqueous buffer as judged by the absence of resonances in <sup>1</sup>H NMR spectra, whereas five additional compounds showed decreased solubility as evidenced by precipitation but were still detectable in <sup>1</sup>H NMR spectra (Table S1).

We then recorded <sup>1</sup>H-<sup>15</sup>N heteronuclear single quantum coherence (<sup>1</sup>H-<sup>15</sup>N HSQC) NMR spectra of <sup>15</sup>N-labeled CeSAS-6<sub>N</sub> in isolation and in the presence of 2 mM nominal concentration of compounds. In such spectra, the position of peaks, corresponding to individual protein amino acids, is sensitive to the amino acid chemical environment and, hence, to binding events. These binding events can be mapped back to the protein structure if the identity of NMR peaks has been assigned, as is the case for CeSAS-6<sub>N</sub> (37). To lessen the likelihood that compounds targeting the hydrophobic cavity of CeSAS-6<sub>N</sub> are out-competed by the β6-β7 loop as a result of protein dimerization, we employed a monomeric variant of this protein, CeSAS-6<sub>N</sub> Δ103–130 I154E. Previous work has shown that the I154E mutation abolishes dimerization (17), whereas deletion of residues 103–130, which form a long unstructured loop, does not affect the domain structure (16) and improves the resolution of NMR spectra (37). Thus, these protein changes leave the hydrophobic pocket of the dimerization interface unaltered and solvent-exposed.

Comparison of <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectra from CeSAS-6<sub>N</sub> Δ103–130 I154E alone and with compounds allowed an initial qualitative differentiation of compound affinity. Fig. 2, A–F show exemplar NMR spectra obtained upon addition of compounds A11, A12, and B1 to CeSAS-6<sub>N</sub> Δ103–130 I154E and the per-residue quantification of spectral changes. As seen there, addition of compounds A11 and B1 to CeSAS-6<sub>N</sub> Δ103–130 I154E caused notable changes in NMR peak positions, indicative of potential binding, whereas addition of A12 did not have a visible effect on NMR spectra. Mapping the most strongly perturbed amino acids to the CeSAS-6<sub>N</sub> structure reveals that A11 binding affects residues at the vicinity of the hydrophobic cavity at the dimerization interface (Fig. 2, G and H), that A12 causes few observable changes to the protein (Fig. 2I), and that residues affected by B1 binding are not in the target pocket (Fig. 2J). Overall, among the 29 soluble or partly soluble compounds tested, A11 was the only compound that perturbed NMR spectra in a pattern indicating binding to the CeSAS-6<sub>N</sub> target pocket (Fig. S2).

We attempted to estimate the binding affinity of A11 to CeSAS-6<sub>N</sub> by recording a series of <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectra and assessing the extent of peak changes as a function of A11 concentration. As shown in Fig. S3, NMR spectral perturbations continued to increase in A11 titrations and did not reach saturation even at 2 mM compound concentration, suggesting a binding affinity in the millimolar range. Furthermore, we exam-

ined whether A11 binding altered the dynamic behavior of CeSAS-6<sub>N</sub>. We recorded heteronuclear {<sup>1</sup>H}-<sup>15</sup>N NOE ratios, which are sensitive to picosecond-nanosecond timescale motions (38), in the presence of A11. As shown in Fig. S4, addition of A11 did not have a significant impact on overall protein flexibility, which may be because the interaction of A11 with CeSAS-6<sub>N</sub> is very weak (Fig. S3) and, thus, rather transient.

### Compound A11 disrupts CeSAS-6 oligomerization

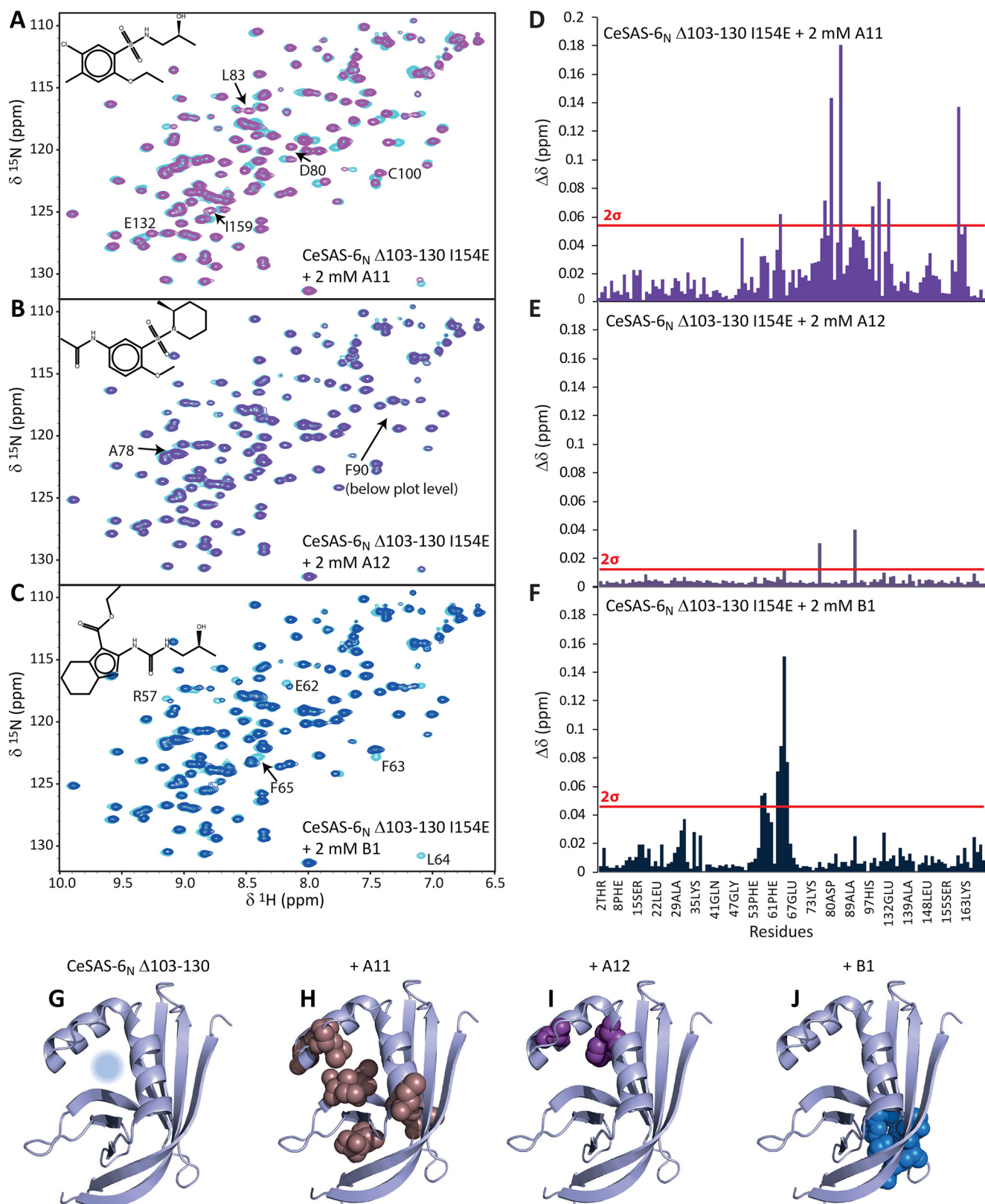
A11 is the first small molecule shown to bind the CeSAS-6<sub>N</sub> hydrophobic cavity at the dimerization interface. To assess whether such binding disrupts protein oligomerization, we performed sedimentation velocity analytical ultracentrifugation (AUC) experiments monitored by optical interference. For these experiments, a CeSAS-6 construct that includes both the N-terminal domain and a section of the coiled-coil domain (CeSAS-6<sub>N-CC</sub>; S123E/I154W, residues 1–215) was used, which can form oligomers if both dimerization interfaces engage in self-association interactions (Fig. 3A). Furthermore, the construct bore two mutations that stabilized the protein in solution (S123E) (17) and enhanced the head domain dimerization affinity (I154W) (16), thereby leading to stronger oligomers. The latter was deemed essential because addition of even small amounts of DMSO, in which compounds were dissolved, had a negative effect on higher-order protein oligomerization (Fig. S5A). We observed that addition of 2 mM A11 in AUC assays reduced the prevalence of large protein oligomers such as tetramers derived from two CeSAS-6<sub>N-CC</sub> coiled-coil dimers interacting via their head domains (Fig. 3A). In contrast, addition of 2 mM A12, which did not show CeSAS-6<sub>N</sub> binding in NMR assays, had no significant effect on CeSAS-6<sub>N-CC</sub> oligomerization (Fig. S5B).

To further assess the effect of A11 on CeSAS-6 oligomerization, we imaged the formation of open-ended spiral assemblies by this protein using EM. We previously showed that such assemblies form readily on carbon-coated grids, likely assisted by protein adsorption on the grids (16). As shown in Fig. 3, B and E, samples of CeSAS-6<sub>N-CC</sub> S123E/I154W alone produced multiple long-spiral assemblies, whereas addition of 5% v/v DMSO slightly reduced both the number and length of these assemblies (Fig. 3, C and E). However, addition of 2 mM A11 in these assays strongly suppressed formation of spirals, resulting in far fewer and shorter assemblies (Fig. 3, D and E). We conclude that A11 prevents CeSAS-6 oligomerization likely by competing with the CeSAS-6<sub>N</sub> β6-β7 loop for binding to the hydrophobic pocket of the head domain dimerization interface.

### Binding of library compounds to SAS-6 orthologues

The pharmacophore features used to select compounds for binding to the CeSAS-6<sub>N</sub> dimerization interface, which included hydrogen bonds directed against protein backbone atoms and the hydrophobic nature of the interface, are common across SAS-6 orthologues. We thus posited that compounds selected using the pharmacophore model of CeSAS-6<sub>N</sub> may also bind head domains from other variants of this protein. To test this hypothesis, we recorded <sup>1</sup>H-<sup>15</sup>N HSQC NMR





spectra of *Chlamydomonas reinhardtii* and human SAS-6 head domains (CrSAS-6<sub>N</sub>, residues 1–159 and HsSAS-6<sub>N</sub>, residues 1–152, respectively) in the presence of 2 mM of library compounds. Similar to earlier assays of CeSAS-6<sub>N</sub>, we employed

dimerization-impaired variants of these domains (*C. reinhardtii* SAS-6<sub>N</sub> (CrSAS-6<sub>N</sub>) F145E and HsSAS-6<sub>N</sub> F131E) (17) to ensure that the hydrophobic cavity of the dimerization interface is solvent-exposed.

## Inhibitors of SAS-6 oligomerization

Analysis of CrSAS-6<sub>N</sub> F145E NMR spectra showed that, similar to CeSAS-6<sub>N</sub>, A11 yielded the strongest perturbations of amino acid resonances, suggesting binding of this compound to the head domain (Fig. 4, A–C). We assigned the NMR spectra of CrSAS-6<sub>N</sub>, allowing us to quantitate these changes on a per-residue basis (Fig. S6, A–C and Fig. S7A) and locate the amino acids affected by A11 binding on the domain structure. As shown in Fig. 4, D and E, addition of A11 perturbed amino acids surrounding the hydrophobic cavity of the CrSAS-6<sub>N</sub> dimerization site, whereas addition of other compounds, such as A12 or B1, produced only small changes in this domain (Fig. 4, F and G). Similar analysis of HsSAS-6<sub>N</sub> F131E NMR spectra revealed that A11 also bound the human orthologue of the SAS-6 head domain (Fig. 4H). However, in contrast to previous titrations in CeSAS-6<sub>N</sub> and CrSAS-6<sub>N</sub>, we observed that A12 bound to HsSAS-6<sub>N</sub> and yielded perturbations in the NMR spectra comparable with those of A11 (Fig. 4I). Other compounds in the library, such as B1 (Fig. 4J), did not produce analogous changes on HsSAS-6<sub>N</sub> F131E spectra.

To elucidate the location of A11 and A12 binding to HsSAS-6<sub>N</sub>, we assigned the NMR spectra of this domain and determined the crystallographic structure of HsSAS-6<sub>N</sub> F131E at 1.46 Å resolution. Comparison of the HsSAS-6<sub>N</sub> head domain structure to that of other resolved SAS-6 variants revealed that the human protein is most similar to CrSAS-6<sub>N</sub> and the *Danio rerio* SAS-6 head domain, with sub-0.7 Å root mean square deviation (RMSD) of C<sub>α</sub> atom positions (Fig. S8). In contrast, the HsSAS-6<sub>N</sub> structure diverged more strongly from that of CeSAS-6<sub>N</sub> (~2.8 Å C<sub>α</sub> RMSD), largely driven by repositioning of the α1 and α2 helices. We assigned the HsSAS-6<sub>N</sub> F131E NMR spectra and quantified the spectral changes upon addition of compounds on a per-residue basis (Fig. S6, D–F and Fig. S7B), which allowed us to locate the most affected amino acids on the domain structure. As shown in Fig. 4, K–M, both A11 and A12 perturb amino acids surrounding the dimerization interface, whereas addition of other compounds did not produce a similar pattern (Fig. 4N). We conclude that A11 binds at the hydrophobic cavity of the SAS-6 head domain dimerization interface in a promiscuous manner but that, interestingly, A12 displays species selectivity binding only at the relevant cavity of the human orthologue.

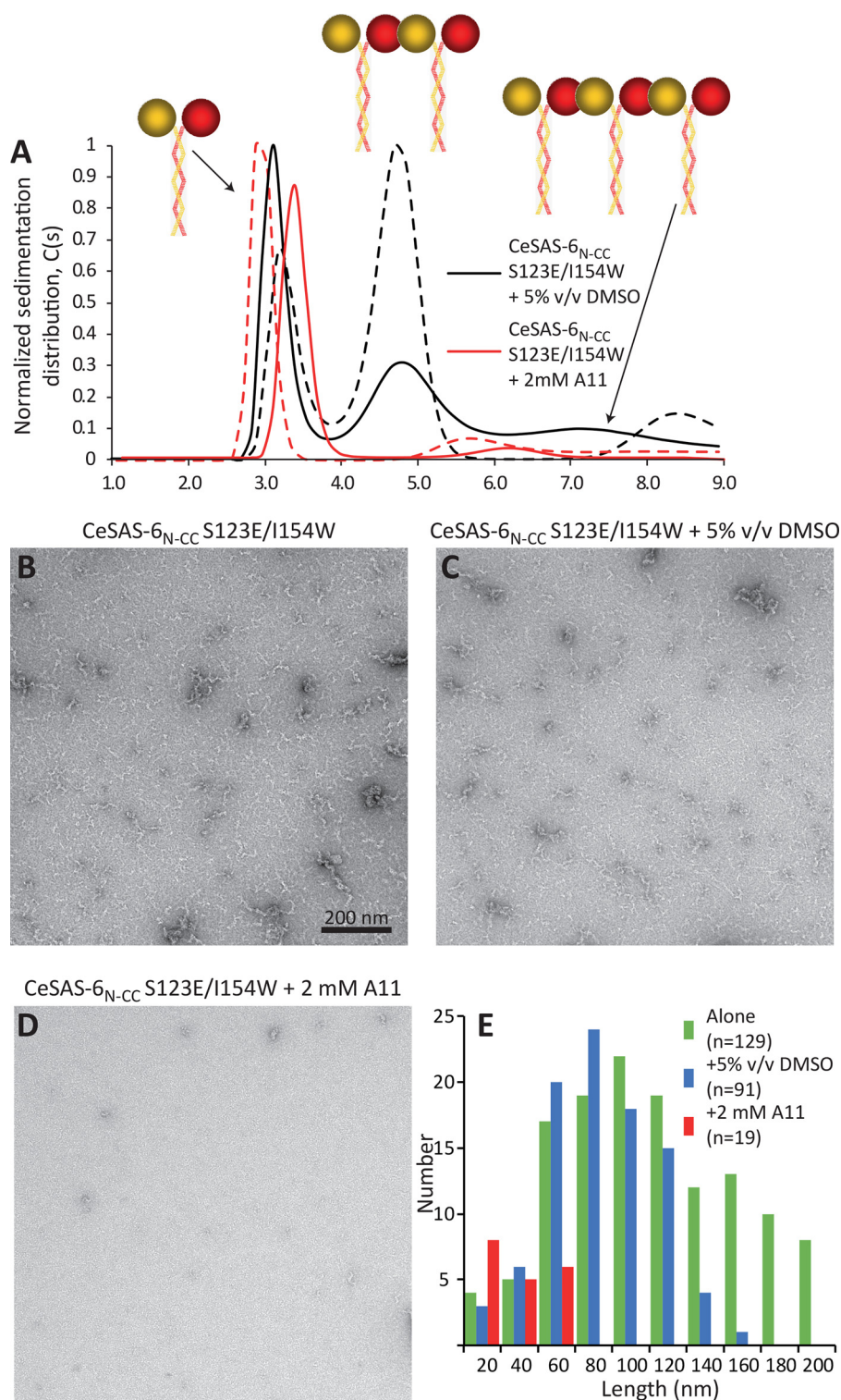
### Molecular dynamics simulations suggest binding determinants of A11 and A12 to SAS-6 head domains

To understand the binding of compounds A11 and A12 to SAS-6 head domains, and the origins of A12 species selectivity, we performed atomistic molecular dynamics (MD) simulations. A11 and A12 were docked in the hydrophobic cavity of the CeSAS-6<sub>N</sub> dimerization site in an initial binding pose consistent with the pharmacophore model, and to those of CrSAS-6<sub>N</sub>

and HsSAS-6<sub>N</sub> using docking approaches. We performed four replicate MD simulations, each 0.5 μs long, in an aqueous environment. A11 simulations revealed a broadly stable and consistent binding pose of the compound in the dimerization cavity (Fig. 5A, Fig. S9, A–C, and Movies S1–S3), although excursions from this position were observed in HsSAS-6<sub>N</sub> (Fig. S9B and Movie S2). The most common A11 binding pose involved the compound aromatic group inserted deep in the hydrophobic cavity led by the chloride atom and forming close interactions with the sidechains of Val-85, Val-93, and Ile-159 in CeSAS-6<sub>N</sub> (Fig. 5B). Similar hydrophobic interactions were observed in MD simulations of A11 with CrSAS-6<sub>N</sub> and HsSAS-6<sub>N</sub> (Fig. 5, C and D). Further, the simulations suggested the presence of stable hydrogen bonds forming between A11 and the α1–α2 loop of SAS-6 head domains. Specifically, the A11 sulfonyl oxygens participated in two hydrogen bonding interactions with the backbone amide and hydroxyl hydrogens of CeSAS-6<sub>N</sub> Thr-84, and the A11 hydroxyl group formed a third hydrogen bond with the backbone carbonyl of Asp-82 (Fig. 5B). Although the Thr-84 equivalent residue in CrSAS-6<sub>N</sub> and HsSAS-6<sub>N</sub> is a leucine (Leu-96 and Leu-77, respectively), two hydrogen bonding interactions between an A11 sulfonyl oxygen and the backbone amide of this leucine, and between the A11 hydroxyl group and the backbone carbonyl of Gly-75 (HsSAS-6<sub>N</sub>) or Gly-94 (CrSAS-6<sub>N</sub>), were sufficient to maintain a stable binding pose (Fig. 5, C and D). This suggests that the CeSAS-6<sub>N</sub> Thr-84 side-chain-mediated hydrogen bond is not critical for A11 binding. Excursions from this binding pose in the MD simulations of A11 with HsSAS-6<sub>N</sub> maintained the interactions of the compound aromatic group with the hydrophobic cavity at the dimerization site but involved the formation of new hydrogen bonding interactions with residues of helix α2 (Movie S2).

Compared with A11, no stable binding pose was observed in MD simulations of A12 with CeSAS-6<sub>N</sub>, as evidenced by high and fluctuating ligand RMSD values (Fig. S9D and Movie S4), which is consistent with the lack of A12 binding to this SAS-6 head domain in NMR assays (Fig. 2, B and D). In contrast, A12 adopted a semi-stable binding pose to HsSAS-6<sub>N</sub> (Fig. S9E and Movie S5), in agreement with binding observed by NMR (Fig. 4I and Fig. S6E). A12 binding involved fluctuating interactions of the methylpiperidine moiety with multiple residues of the HsSAS-6<sub>N</sub> dimerization cavity (Leu-76, Val-78, Phe-83, Lys-86, Phe-87, Leu-90, His-136, and Leu-137; Fig. 5E), compared with A11 where similar but tighter interactions were formed by the compound aromatic group (Fig. 5D). We postulate that the methoxy and acetamide decorations of the A12 aromatic group prevented its stable insertion in the HsSAS-6<sub>N</sub> hydrophobic cavity. The A12 methylpiperidine moiety was somewhat mobile in the hydrophobic cavity as a result of rotations around the

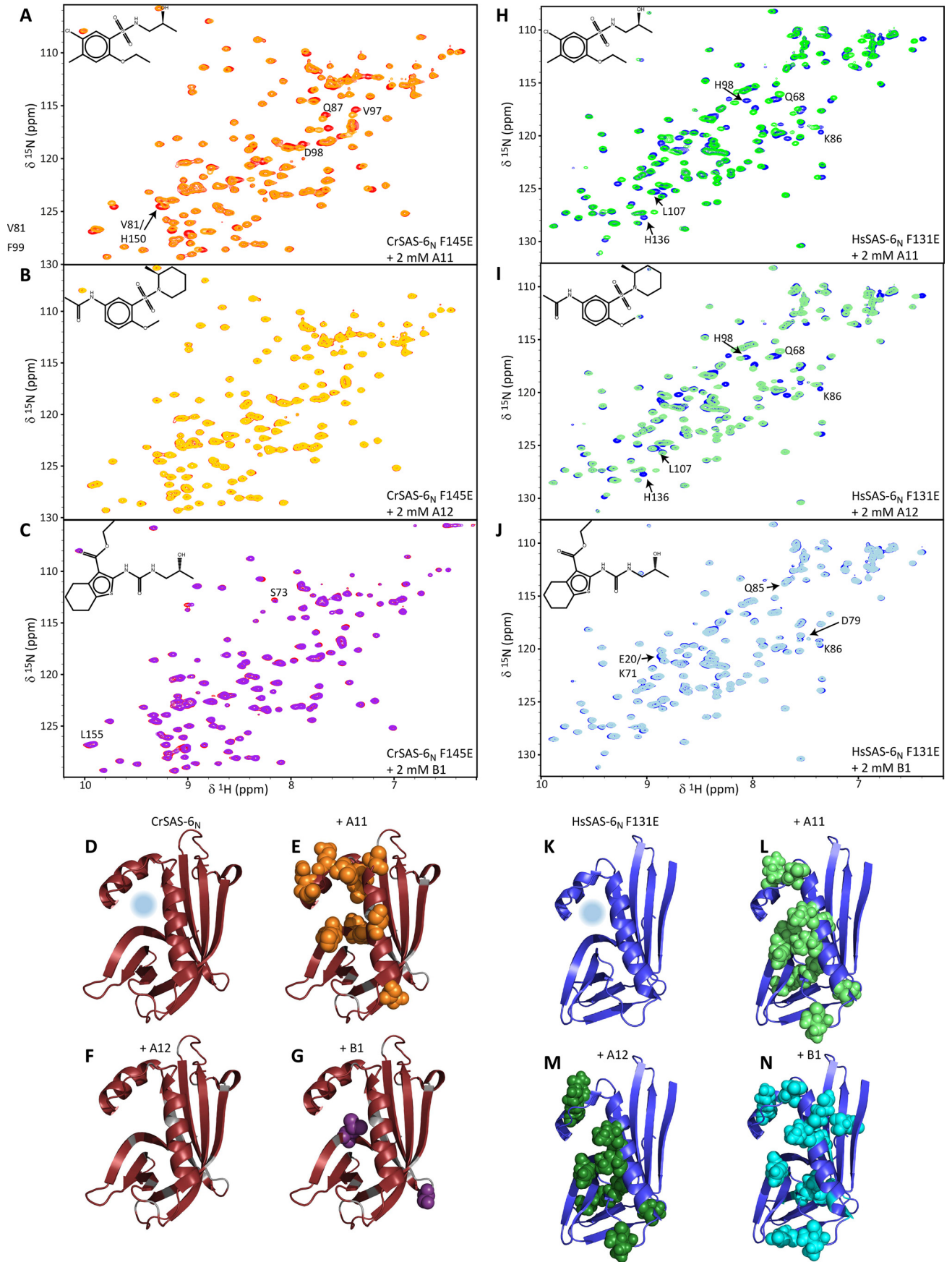
**Figure 2. NMR assay of CeSAS-6<sub>N</sub> small molecule interactors.** A–C, overlay of NMR <sup>1</sup>H–<sup>15</sup>N HSQC spectra produced by CeSAS-6<sub>N</sub> Δ103–130 I154E alone (cyan) or in the presence of 2 mM A11 (A), A12 (B), or B1 (C) compounds. The chemical structures of the compounds are inlaid, and CeSAS-6<sub>N</sub> resonances exhibiting the most significant changes upon addition of compounds are labeled. D–F, per-residue quantification of combined changes in <sup>1</sup>H and <sup>15</sup>N chemical shifts of CeSAS-6<sub>N</sub> Δ103–130 I154E resonances upon addition of 2 mM A11 (D), A12 (E), or B1 (F) compounds. A measure of two standard deviations of all changes observed is shown as a red line, indicating amino acids that experienced the strongest perturbations upon compound addition. G, CeSAS-6<sub>N</sub> Δ103–130 monomer structure with the dimerization site targeted by compounds indicated by a light blue circle, derived from PDB ID 4G79. H–J, amino acid residues strongly perturbed by addition of compounds A11 (H), A12 (I), or B1 (J), shown in sphere representation. Only A11 produces changes in amino acids surrounding the targeted site, suggesting compound binding at the hydrophobic cavity of the dimerization interface.

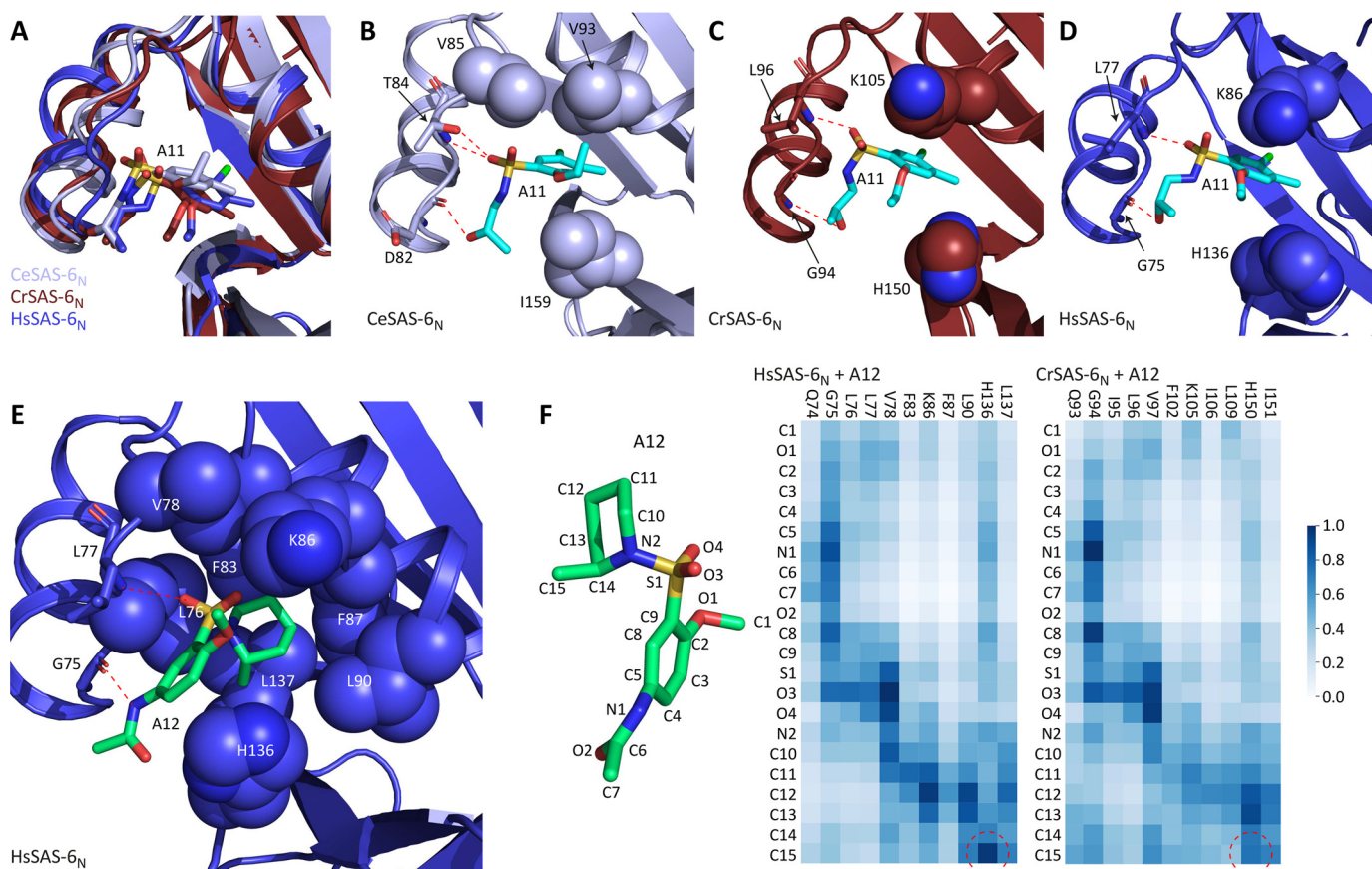


**Figure 3. Assays of CeSAS-6<sub>N-CC</sub> oligomerization.** A, sedimentation velocity analysis of CeSAS-6<sub>N-CC</sub> S123E/I154W in the presence of 5% v/v DMSO (black solid and dashed lines) or upon addition of 2 mM A11 (red solid and dashed lines). Solid versus dashed lines correspond to two independent repeats of the assay. The molecular architecture of CeSAS-6<sub>N-CC</sub> and the oligomeric state of different sedimentation species are shown schematically on top (spheres: head domains; zigzag lines: coiled-coil segments). Stable CeSAS-6 dimers are mediated by the coiled-coil segment. Addition of A11 reduces the prevalence of higher oligomeric species, which necessitate interactions between head domains. B–D, representative electron micrographs of CeSAS-6<sub>N-CC</sub> S123E/I154W spiral assemblies formed at 1 mg/ml protein concentration and standard buffer (B) or in the presence of 5% v/v DMSO (C) or 2 mM A11 (D). Common scale bar in panel B 200 nm. E, plot of assembly length versus number of occurrences from the CeSAS-6<sub>N-CC</sub> S123E/I154W samples shown in (B–D). Addition of A11 strongly reduces the number and length of visible assemblies.



**Inhibitors of SAS-6 oligomerization**





**Figure 5. MD simulations of A11 and A12 binding to SAS-6 head domains.** A, superposition of representative states from MD simulations of CeSAS-6<sub>N</sub> (light blue), CrSAS-6<sub>N</sub> (red), and HsSAS-6<sub>N</sub> (blue) with A11. The ligand, shown as sticks and with carbon atoms colored similar to the respective protein, adopts a consistent binding pose relative to head domains (schematic representation), with the compound aromatic group buried in the hydrophobic cavity of the head domain dimerization site. B, key interactions forming between A11 and CeSAS-6<sub>N</sub> seen in MD simulations. Hydrogen bonds are shown as red dashed lines. A11 and protein amino acids participating in hydrogen bonds are shown in stick representation. The sidechains of amino acids involved in hydrophobic interactions are shown as spheres. The ligand is shown as sticks with carbon atoms in cyan. C–E, similar views of interactions between A11 and CrSAS-6<sub>N</sub> (C) or HsSAS-6<sub>N</sub> (D), and A12 with HsSAS-6<sub>N</sub> (E). The A12 ligand carbon atoms are colored light green. F, ligand contact maps from the HsSAS-6<sub>N</sub> and CrSAS-6<sub>N</sub> MD simulations with A12. A12 atom nomenclature is shown alongside a stick representation of the ligand (right panel). The HsSAS-6<sub>N</sub> (middle panel) and CrSAS-6<sub>N</sub> (left panel) maps show close contacts between protein residues and ligand atoms as annotated. Different densities (blue hues) denote the occurrence of close contacts during the simulation, normalized to the most frequently observed contact (assigned value of 1, dark blue). Dashed circles highlight close contacts of the A12 methyl group in the HsSAS-6<sub>N</sub> simulation that are absent in CrSAS-6<sub>N</sub>.

sulfonyl-methylpiperidine bond (Fig. S9D and Movie S5), albeit with a preference on methyl group orientation (see below). We observed the formation of two stable hydrogen bonds between the A12 sulfonyl group oxygen and the HsSAS-6<sub>N</sub> Leu-77 amide and between the A12 acetamide and the Gly-75 carbonyl atoms (Fig. 5E).

Although dimerization cavities of SAS-6 head domains are consistently hydrophobic, we noted that only three of eight HsSAS-6<sub>N</sub> amino acids involved in interactions with the A12 methylpiperidine moiety are strictly conserved in CeSAS-6<sub>N</sub> (Fig. S8G). These changes in amino acid sidechains may account for the lack of favorable methylpiperidine packing in the dimerization cavity of CeSAS-6<sub>N</sub>, leading to high com-

ound mobility in simulations. Consistent with this analysis, MD simulations of A12 in complex to CrSAS-6<sub>N</sub>, where five of eight methylpiperidine-binding amino acids are conserved, showed relatively stable compound binding (Fig. S9F and Movie S6). However, in these CrSAS-6<sub>N</sub> simulations the A12 methylpiperidine moiety showed no clear positioning preference, whereas in HsSAS-6<sub>N</sub> simulations the moiety methyl group was primarily directed toward Leu-76, His-136, and Leu-137 (Fig. 5F). Two of these three HsSAS-6<sub>N</sub> amino acids are substituted in CrSAS-6<sub>N</sub> (Leu-76 versus Ile-95 and Leu-137 versus Ile-151; Fig. S8G), which may be responsible for the increased methylpiperidine mobility in CrSAS-6<sub>N</sub> MD simulations compared with those of HsSAS-6<sub>N</sub>. At the same time, we

**Figure 4. NMR assays of CrSAS-6<sub>N</sub> and HsSAS-6<sub>N</sub> compound binding.** A–C, overlay of NMR <sup>1</sup>H-<sup>15</sup>N HSQC spectra produced by CrSAS-6<sub>N</sub> F145E alone (red) or in the presence of 2 mM A11 (A), A12 (B), or B1 (C) compounds. The chemical structure of compounds is inlaid, and CrSAS-6<sub>N</sub> resonances exhibiting the most significant changes upon addition of compounds are labeled. D, CrSAS-6<sub>N</sub> monomer structure with the dimerization site indicated by a light blue circle, derived from PDB ID 3Q0Y. E–G, amino acid residues strongly perturbed by addition of compounds A11 (E), A12 (F), or B1 (G), shown in sphere representation. Only A11 produces changes in amino acids surrounding the targeted site, suggesting compound binding at the hydrophobic cavity of the dimerization interface. H–N, similar NMR spectra of HsSAS-6<sub>N</sub> F131E (H–J, blue) with compounds A11, A12, and B1, and structure representations of the HsSAS-6<sub>N</sub> dimerization site (K) and residues affected by compound binding (L–N). The HsSAS-6<sub>N</sub> structure was resolved as part of this study (Fig. S8 and Table S2). In contrast to CeSAS-6<sub>N</sub> and CrSAS-6<sub>N</sub>, both A11 and A12 bind the dimerization site of HsSAS-6<sub>N</sub>.



## Inhibitors of SAS-6 oligomerization

note that CrSAS-6<sub>N</sub>-A12 simulations clearly underestimate the fluctuations in the binding pose of this compound, because no relevant complex formation was observed *in vitro*. (Fig. 4B and Fig. S6B). Similarly, A11 binding in MD simulations appears less dynamic than experimental evidence would suggest. Although A11 titrations show that ligand binding and dissociation occur in the fast (sub- $\mu$ s) NMR timescale (e.g. Fig. S3A), we observed no instance of A11 clearly leaving the binding pocket in any of the 0.5  $\mu$ s-long simulations (Fig. S9, A–C and Movies S1–S3).

### Discussion

We presented here work toward the identification of small molecules targeting the SAS-6 head domain and preventing its dimerization. SAS-6 is the main component of the cartwheel scaffold in centrioles and in this role serves to ensure correct assembly of this organelle. Because SAS-6 oligomerization, including via self-association interactions of the protein head domain, is essential for centriole formation, a high-affinity and high-specificity compound blocking head domain dimerization could have a wide range of possible uses, from cell biological tool to putative anti-cancer therapeutic. However, further steps are necessary to develop such a potent and specific compound starting from this work.

Previously, a bromoindole derivative termed PK9119 was shown to bind the dimerization site of *Leishmania major* SAS-6 head domains with millimolar-level affinity and potentially also that of *D. rerio* SAS-6, although with significantly smaller potency (18). PK9119 did not bind CeSAS-6<sub>N</sub> in our NMR assays (Fig. S10). Rather, in combined *in silico* and *in vitro* assays we demonstrated that compound A11 associates equally well with the head domains of nematode, algae, and human SAS-6, as judged by perturbations of NMR spectra (Fig. 2D and Fig. S5, A and D) and persistence in MD simulations (Fig. S9, A–C), whereas compound A12 shows binding specificity for the human orthologue. Thus, we consider that A11 may provide a starting point for developing a compound that would promiscuously target SAS-6 head domains. Furthermore, the presence of two SAS-6-binding molecules in our library of just 24 highly soluble compounds suggests that the pharmacophore model used captures preferences common across SAS-6 head domains. Thus, our relatively small compound library may comprise a good initial set of molecules for identifying binders of SAS-6 orthologues in general.

Binding of A11 to CeSAS-6<sub>N</sub> was notably weak, with millimolar-level affinity (Fig. S3), which is likely responsible for the lack of large-scale changes in the CeSAS-6<sub>N</sub> structure and dynamics inferred from MD simulations and  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE NMR experiments (Fig. S4 and Movie S1). However, we were unable to sufficiently resolve peaks in the NMR spectra for protein dynamics from many residues in the CeSAS-6<sub>N</sub>  $\alpha$ 1- $\alpha$ 2 loop, including residues potentially forming hydrogen bonds to A11; thus, we cannot exclude the possibility that localized changes in  $\alpha$ 1- $\alpha$ 2 loop dynamics may occur upon A1 binding. Nevertheless, because A11 was sufficiently potent to prevent protein oligomerization in AUC and EM assays (Fig. 3), we considered how compound affinity may be improved. We noted

that, first, both A11 and A12 include a sulfonyl group that participates in hydrogen bonding interactions with residues of the protein  $\alpha$ 1- $\alpha$ 2 loop (Fig. 5). In A11, an aminopropanol moiety extends from the sulfonyl group and provides an additional hydrogen bond via interactions with the  $\alpha$ 1- $\alpha$ 2 loop, whereas in A12 similar functionality is offered by a methoxy-benzylacetamide moiety. Sulfonyl groups were present in several ligands identified as putative CeSAS-6<sub>N</sub> binders (Table S1), stemming from the selection for hydrogen bond acceptors in the pharmacophore mode. However, some of these ligands were poorly soluble (A1, C8, and C9), which limited screening of sulfonyl derivatives. Nevertheless, we believe that neither the sulfonyl group nor its extensions present in A11 and A12 are optimal for the protein-ligand interactions observed, because of high levels of ligand mobility permitted by bond rotations around these groups. Rather, we propose that replacing the sulfonyl with a semi-rigid ester group, extended with a short alkylamine or alcohol moiety, may prove to be more favorable by reducing the number of possible conformational states for the ligand.

Both A11 and A12 feature aromatic groups, but only that of A11 interacted with the hydrophobic cavity of SAS-6 head domains (Fig. 5, A–D), likely because of hydrophilic decorations of the A12 aromatic moiety. The relevance of aromatic groups in SAS-6 head domain dimerization is well understood, not least because most SAS-6 proteins use a phenylalanine amino acid at the tip of the  $\beta$ 6- $\beta$ 7 loop to mediate the key hydrophobic interaction (17, 19, 20), whereas a tryptophan substitution at the same site enhances the interaction strength (16, 23). Thus, an aromatic group with minimal and hydrophobic decorations, similar to that of A11, is likely essential for deriving a promiscuous binder of SAS-6 head domains. However, potentially more exciting is the discovery that cycloalkane groups, such as the methylpiperidine of A12, may also be suitable for interacting with the SAS-6 head domain dimerization cavity and confer specificity in the interaction. As seen in MD simulations of HsSAS-6<sub>N</sub> with A12 (Fig. 5E), the available hydrophobic interface in this orthologue is larger than originally thought on the basis of other SAS-6 structures, and this interface could be exploited by “growing” the methylpiperidine moiety, thereby gaining human SAS-6-specific enthalpic contributions to the interaction. Further, at the periphery of the dimerization interface, two residues, Lys-86 and His-136 in HsSAS-6<sub>N</sub>, are only partly exploited by our currently ligand designs and offer opportunities for hydrogen bonding. Lys-86, in particular, is little-conserved among SAS-6 orthologues (Fig. S8G), suggesting that it could be targeted for interactions that would increase ligand specificity to human SAS-6.

In conclusion, we show here that a pharmacophore model derived from the dimerization interface of SAS-6 head domains can successfully identify compounds binding to this interface with the potential to block SAS-6 oligomerization. Intriguingly, although the pharmacophore was designed on the basis of general SAS-6 preferences, we observed the emergence of interaction specificity toward human SAS-6 already in the context of a relatively small library of compounds, likely as a result of packing contributions at the dimerization cavity. We propose that these contributions can be further optimized toward a strong and specific binder of human SAS-6.

## Experimental procedures

### Protein production and purification

Constructs encoding fragments of *C. elegans* SAS-6 (UniProt ID: O62479) residues 1–168 (CeSAS-6<sub>N</sub>) or 1–215 (CeSAS-6<sub>N-CC</sub>) were cloned in pHis vector (17), which provides an N-terminal His<sub>6</sub>-tag and thrombin cleavage site for removal of the cloning tag. A construct encoding *C. reinhardtii* SAS-6 (UniProt ID: A9CQL4) residues 1–159 (CrSAS-6<sub>N</sub>) was cloned in pSTCm1 vector (17), which also provides N-terminal His<sub>6</sub>-tag and thrombin cleavage elements. A construct encoding human SAS-6 (UniProt ID: Q6UVJ0) residues 1–152 (HsSAS-6<sub>N</sub>) was cloned in pFLOAT2 vector (39), which provides an N-terminal His<sub>6</sub>-tag and a human rhinovirus 3C protease cleavage site. Site-specific mutations were introduced in these constructs by the QuikChange method (Agilent Technologies). The identity of constructs was verified by sequencing (Eurofins Scientific).

*Escherichia coli* strain BL21(DE3) cells bearing these genetic constructs were grown at 37 °C in lysogeny broth, or M9 minimal medium supplemented with <sup>15</sup>N-enriched NH<sub>4</sub>Cl or <sup>15</sup>N-enriched NH<sub>4</sub>Cl and <sup>13</sup>C-enriched glucose, and appropriate antibiotics. Protein expression was induced at optical density at 600 nm of ≥0.6 by addition of 0.25 mM isopropyl β-D-1-thiogalactopyranoside and let to proceed overnight at 18 °C. The cells were harvested by centrifugation and resuspended in 150 mM NaCl, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.5, buffer supplemented with 1% v/v Triton X-100 and cOmplete mini EDTA-free protease inhibitor mixture (Roche). The cell suspension was lysed by sonication and clarified by centrifugation at 30,000 × *g* at 4 °C for 30 min.

Clarified supernatants were loaded in pre-packed HiTrap Talon columns (GE Lifesciences) equilibrated in lysis buffer. The columns were washed in lysis buffer supplemented with 5 mM imidazole, and bound proteins were eluted from columns by a gradient of lysis buffer supplanted with 250 mM imidazole. His<sub>6</sub>-tags were removed by addition of thrombin (Sigma-Aldrich) or homemade human rhinovirus 3C protease, and protein samples were further purified by size exclusion chromatography using pre-packed Sephadex 75 or Sephadex 200 HiLoad 16/600 columns (GE Healthcare Life Sciences) equilibrated in NMR or crystallization buffer (see below). Protein-containing fractions were concentrated using spin ultrafiltration and Amicon Ultra (Millipore) devices at 4 °C. Protein quality was assessed by SDS-PAGE chromatography, protein identity was confirmed by electrospray MS (Department of Biochemistry, University of Oxford, UK), and protein concentration was established by UV spectroscopy using a calculated extinction coefficient (40) and a NanoDrop spectrophotometer (Thermo Fisher Scientific).

### NMR

Sequence-specific NMR resonance assignments were performed as described previously (41). Briefly, NMR experiments were performed at 25 °C using Bruker Avance III spectrometers with cryogenic TCI probeheads and 11.7–17.6-T magnetic field strengths. We used samples of <sup>13</sup>C/<sup>15</sup>N-enriched proteins at 0.5 mM concentration for resonance assignments and <sup>15</sup>N-enriched proteins at 0.1 mM concentration for binding assays.

Samples were in 150 mM NaCl, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.5 buffer supplemented with 5% v/v D<sub>2</sub>O, 0.02% w/v NaN<sub>3</sub>, and 50 μM 4,4-dimethyl-4-silapentane-1-sulfonic acid for reference purposes. Samples for binding assays contained either 5% v/v DMSO or the indicated concentration of compound supplemented with DMSO to reach a 5% v/v final concentration. Assignment experiments were performed using 3D CBCA(CO)NH, CBCANH, and HNCA pulse sequences. NMR data were processed using NMRpipe (42) and analyzed using PIPP (43) and Sparky (44). Where necessary, chemical shift assignments were transferred between WT and mutant constructs by overlaying spectra. For the calculation of combined chemical shift perturbations upon addition of compounds we summed up the <sup>1</sup>H and <sup>15</sup>N contributions using the following empirically derived equation:

$$\Delta\delta = \sqrt{\frac{1}{2}(\delta_H^2 + 0.14\delta_N^2)}$$

Chemical shift changes were considered notable if they exceeded at least twice the intrinsic experimental error ( $2\sigma_{\text{exp}}$ ) or twice the standard deviation ( $2\sigma$ ) of chemical shift perturbation values of all residues in the experiment. The intrinsic experimental error was calculated from the digital resolution of spectra to be ~0.01 ppm. <sup>1</sup>H-<sup>15</sup>N NOE ratios were recorded using 0.5 mM protein concentration samples and analyzed as described previously (41).

### Analytical ultracentrifugation

AUC sedimentation velocity experiments were performed at the Research Complex at Harwell (Harwell, UK) using a Beckman Coulter ProteomeLab XL-I ultracentrifuge and An-50Ti or An-60Ti rotors (Beckman Coulter) with sapphire window inserts. Experiments were conducted at 15 °C and 35,000 rpm with data recorded simultaneously at a 280 nm wavelength and via interference optics over 250 complete scans. Protein samples were at 50 mM concentration in 150 mM NaCl, 20 mM HEPES buffer, pH 7.5, with 2 mM Tris(2-carboxyethyl)phosphine and supplemented with 5% v/v DMSO or small molecule compounds at 2 mM final concentration as indicated. Data were analyzed using Sedfit (45).

### EM

Electron micrographs were recorded as described previously (16). Briefly, freshly purified protein samples at 1 mg/ml concentration were extensively dialyzed against a 150 mM NaCl, 20 mM HEPES, pH 7.5 buffer, supplemented with 5% v/v DMSO or 2 mM A11 as necessary, and applied to carbon-coated grids, which were glow discharged using a Leica EM ACE200 vacuum coater. The samples were incubated on grids for 2 min prior to blotting and staining with 2% w/v uranyl acetate for 30 s before being left to dry in the dark for 10 min. Electron micrographs were obtained using a Tecnai12 transmission electron microscope and a Gatan OneView CMOS camera.

## Inhibitors of SAS-6 oligomerization

### X-ray crystallography

Crystals of HsSAS-6<sub>N</sub> F131E were produced by the hanging drop vapor diffusion method. A Mosquito robot (TTP Labtech) was used to set up 100-nl drops with a 1:1 volume ratio of protein in 50 mM NaCl, 20 mM Tris(hydroxymethyl)aminomethane (Tris)-Cl, pH 7.5 buffer at 98 mg/ml concentration and diverse mother liquors. Crystals with plate-like morphology developed in ~3 days using 30% w/v PEG 6000 and 0.1 M HEPES, pH 7 as mother liquor and were cryoprotected by brief immersion in mother liquor supplemented with 20% v/v glycerol prior to cooling in liquid nitrogen. Diffraction data were recorded at the Diamond Light Source (Harwell, UK) beamline I03 to 1.46 Å maximum resolution. Data were processed with Xia2 (46), and the structure was solved by molecular replacement using Phaser (47) and a homology model of HsSAS-6<sub>N</sub> based on the previous *D. rerio* SAS-6 head domain structure (PDB ID: 2Y3V; 19). Iterative structure refinement was performed using Phenix (48) and model building in Coot (49). Crystallographic data quality and refinement statistics are shown in Table S2.

### Pharmacophore preparation and virtual screening

Discovery Studio 3.5 software (50) was used for the generation of the pharmacophore model using the CeSAS-6<sub>N</sub> dimer crystallographic structure (PDB ID: 3PYI; 17) as the starting point. Water molecules in the structure were removed prior to analysis. Interaction patterns between the two domains were determined with a minimum feature distance of 1.5 Å. The pharmacophore model obtained was manually edited to contain a total of 10 features. We screened 100 different pharmacophore feature combinations, including at least five of these features in all combinations. At minimum, pharmacophore features included two hydrophobic or aromatic features close to Leu-98 and Ile-159 and one hydrogen bond donor. Further, we added exclusion features representing the van der Waals surface of CeSAS-6<sub>N</sub>.

The ~13 million commercially available compounds of the Clean Drug-like subset of the ZINC database (51) were filtered according to molecular mass (300 to 500 Da), hydrogen bond donors ( $\geq 1$ ), and aromatic groups ( $\geq 1$ ) using OpenBabel (52), yielding a screening subset of ~6 million molecules. 250 conformations of each compound were generated, fitted to the 100 different pharmacophore models with Discovery Studio 3.5, and ranked according to their fitting score as described before (53).

### Molecular dynamics simulations

A monomer from the crystal structure of CeSAS-6<sub>N</sub> Δ103–130 (PDB ID: 4G79; 16) was used for preparation of the CrSAS-6<sub>N</sub> starting model, with A11 and A12 inserted in the cavity of the dimerization site according to the pharmacophore model. For the CrSAS-6<sub>N</sub> (PDB ID: 3Q0Y; 17) and HsSAS-6<sub>N</sub> starting models with A11 and A12, we used monomer copies of the proteins and the best docking poses obtained by AutoDock Vina (54), using a 25 × 25 × 25-Å grid box, a search space of 10 binding modes, and a search parameter of 5. MD simulations were performed using GROMACS version 2018.3 (55) and the

AMBER99SB-ILDN force field (56). The protein-ligand complexes were inserted in a pre-equilibrated box containing water implemented using the TIP3P water model (56). Force field parameters for A11 and A12 were generated using the general AMBER force field and HF/6–31G\*–derived Restrained Electrostatic Potential atomic charges (57). The reference system consisted of the protein, the ligand, ~16,900 water molecules, and 47 Na and 47 Cl ions in an 80 × 80 × 80-Å simulation box, resulting in a total number of ~53,200 atoms. Each system was energy-minimized and subsequently subjected to a 10-ns MD equilibration, with an isothermal-isobaric ensemble using isotropic pressure control (58) and positional restraints on protein and ligand coordinates. The resulting equilibrated systems were replicated four times and independent 500-ns MD trajectories were produced at a constant temperature of 300 K, using separate v-rescale thermostats (58) for the protein, ligand, and solvent molecules. A time-step of 2 fs was used, and all bonds were constrained using the LINCS algorithm (59). Lennard-Jones interactions were computed using a cutoff of 10 Å and electrostatic interactions were treated using particle mesh Ewald (60) with the same real-space cutoff.

### Notes

Small molecule compounds were purchased from commercial vendors as indicated in Table S1 and dissolved in 100% v/v DMSO to a 40 mM final concentration. Structure figures were prepared using PyMOL (61). Sequence alignments were performed using Clustal Omega (62).

### Data availability

The HsSAS-6<sub>N</sub> F131E structure model and associated data have been deposited in the RCSB Protein Data Bank under accession number 6Z4A. NMR assignments were deposited in BioMagResBank under accession numbers 50300 (CrSAS-6<sub>N</sub> F145E) and 50308 (HsSAS-6<sub>N</sub> F131D).

**Acknowledgments**—We are grateful to Nick Soffe and Christina Redfield for maintenance of the solution NMR facility, to Leanne Slatter for assistance with protein production, to Anna Füzéry for recording NMR data, and to Jemma Harris and David Scott for supervision of the AUC assays. We thank Georgios Hatzopoulos and Pierre Gönczy for assistance with SAS-6 inhibition assays, the Diamond Light Source and SOLEIL synchrotrons for provision of macromolecular crystallography beamtime, and the Advanced Research Computing (Oxford) for additional computer time.

**Author contributions**—J. M. C. B., M.-T. M., and I. V. conceptualization; J. M. C. B., M.-T. M., and I. V. formal analysis; J. M. C. B. and I. V. validation; J. M. C. B., M.-T. M., and I. V. investigation; J. M. C. B., M.-T. M., and I. V. visualization; J. M. C. B., M.-T. M., and I. V. methodology; M.-T. M. and I. V. data curation; M.-T. M., M. M., G. A. S., P. C. B., and I. V. writing-review and editing; M. M., G. A. S., P. C. B., and I. V. supervision; G. A. S., P. C. B., and I. V. funding acquisition; I. V. writing-original draft; I. V. project administration.

**Funding and additional information**—This work was supported by Medical Research Council UK Grant MR/N009274/1 and the



Cephalosporin Fund Grant CF 329 (to I. V.). J. M. C. B. was funded by a PhD studentship from the Biotechnology and Biological Sciences Research Council UK to the Oxford Interdisciplinary Bioscience Doctoral Training Partnership. This work was also supported by the Greek State Scholarships Foundation Grant MIS-5033021 (to M. T. M.), the Wellcome Trust Grant 094872/Z/10/Z (to I. V.), the Engineering and Physical Sciences Research Council Grant EPSRC, EP/R029849/1 (to I. V.), the Wellcome Institutional Strategic Support Fund, the EPA Cephalosporin Fund, and the John Fell OUP Research Fund. Partial support of this work was provided by the project “INSPIRED - The National Research Infrastructures on Integrated Structural Biology, Drug Screening Efforts and Drug target functional characterization” Grant MIS 5002550, implemented under the Action “Reinforcement of the Research and Innovation Infrastructure” funded by the Operational Programme “Competitiveness, Entrepreneurship and Innovation” Grant NSRF 2014-2020 and co-financed by Greece and the European Union (European Regional Development Fund). This project made use of ARCHER via HECBioSim (RRID:SCR\_015854), supported by EPSRC Grant EP/R029407/1 (to P. C. B.). We thank Leonardo Pardo for access to high-performance computing facilities and use of software.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: SAS, spindle assembly abnormal protein; AUC, analytical ultracentrifugation;  $^1\text{H}$ - $^{15}\text{N}$  HSQC,  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single quantum coherence; MD, molecular dynamics; Plk, polo-like kinase; CeSAS, *Caenorhabditis elegans* SAS; CeSAS-6<sub>N</sub>, CeSAS-6 N-terminal domain; CeSAS-6<sub>N-CC</sub>, CeSAS-6 construct including N-terminal and coiled-coil domains; RMSD, root mean square deviation; MD, molecular dynamics; HsSAS-6<sub>N</sub>, human SAS-6N; CrSAS-6N, *Chlamydomonas reinhardtii* SAS-6N.

## References

- Hodges, M. E., Scheumann, N., Wickstead, B., Langdale, J. A., and Gull, K. (2010) Reconstructing the evolutionary history of the centriole from protein components. *J. Cell Sci.* **123**, 1407–1413 [CrossRef Medline](#)
- Marshall, W. F. (2009) Centriole evolution. *Curr. Opin. Cell Biol.* **21**, 14–19 [CrossRef Medline](#)
- Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J. B., and Bettencourt-Dias, M. (2011) Evolution: tracing the origins of centrioles, cilia, and flagella. *J. Cell Biol.* **194**, 165–175 [CrossRef Medline](#)
- Conduit, P. T., Wainman, A., and Raff, J. W. (2015) Centrosome function and assembly in animal cells. *Nat. Rev. Mol. Cell Biol.* **16**, 611–624 [CrossRef Medline](#)
- Bornens, M. (2012) The centrosome in cells and organisms. *Science* **335**, 422–426 [CrossRef Medline](#)
- Marshall, W. F. (2008) Basal bodies platforms for building cilia. *Curr. Top. Dev. Biol.* **85**, 1–22 [CrossRef Medline](#)
- Firat-Karalar, E. N., and Stearns, T. (2014) The centriole duplication cycle. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **369**, 20130460 [CrossRef Medline](#)
- Gönczy, P., and Hatzopoulos, G. N. (2019) Centriole assembly at a glance. *J. Cell Sci.* **132**, jcs228833 [CrossRef Medline](#)
- Nakazawa, Y., Hiraki, M., Kamiya, R., and Hirono, M. (2007) SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole. *Curr. Biol.* **17**, 2169–2174 [CrossRef Medline](#)
- Leidel, S., Delattre, M., Cerutti, L., Baumer, K., and Gönczy, P. (2005) SAS-6 defines a protein family required for centrosome duplication in *C. elegans* and in human cells. *Nat. Cell Biol.* **7**, 115–125 [CrossRef Medline](#)
- Dammermann, A., Müller-Reichert, T., Pelletier, L., Habermann, B., Desai, A., and Oegema, K. (2004) Centriole assembly requires both centriolar and pericentriolar material proteins. *Dev. Cell* **7**, 815–829 [CrossRef Medline](#)
- Dobbelaere, J., Josué, F., Suijkerbuijk, S., Baum, B., Tapon, N., and Raff, J. (2008) A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in *Drosophila*. *PLoS Biol.* **6**, e224 [CrossRef Medline](#)
- Strnad, P., Leidel, S., Vinogradova, T., Euteneuer, U., Khodjakov, A., and Gönczy, P. (2007) Regulated HsSAS-6 levels ensure formation of a single procentriole per centriole during the centrosome duplication cycle. *Dev. Cell* **13**, 203–213 [CrossRef Medline](#)
- Hirono, M. (2014) Cartwheel assembly. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **369**, 20130458 [CrossRef Medline](#)
- Pelletier, L., O’Toole, E., Schwager, A., Hyman, A. A., and Müller-Reichert, T. (2006) Centriole assembly in *Caenorhabditis elegans*. *Nature* **444**, 619–623 [CrossRef Medline](#)
- Hilbert, M., Erat, M. C., Hachet, V., Guichard, P., Blank, I. D., Flückiger, I., Slater, L., Lowe, E. D., Hatzopoulos, G. N., Steinmetz, M. O., Gönczy, P., and Vakonakis, I. (2013) *Caenorhabditis elegans* centriolar protein SAS-6 forms a spiral that is consistent with imparting a ninefold symmetry. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11373–11378 [CrossRef Medline](#)
- Kitagawa, D., Vakonakis, I., Olieric, N., Hilbert, M., Keller, D., Olieric, V., Bortfeld, M., Erat, M. C., Flückiger, I., Gönczy, P., and Steinmetz, M. O. (2011) Structural basis of the 9-fold symmetry of centrioles. *Cell* **144**, 364–375 [CrossRef Medline](#)
- van Breugel, M., Wilcken, R., McLaughlin, S. H., Rutherford, T. J., and Johnson, C. M. (2014) Structure of the SAS-6 cartwheel hub from *Leishmania major*. *eLife* **3**, e01812 [CrossRef Medline](#)
- van Breugel, M., Hirono, M., Andreeva, A., Yanagisawa, H. A., Yamaguchi, S., Nakazawa, Y., Morgner, N., Petrovich, M., Ebong, I. O., Robinson, C. V., Johnson, C. M., Vepintsev, D., and Zuber, B. (2011) Structures of SAS-6 suggest its organization in centrioles. *Science* **331**, 1196–1199 [CrossRef Medline](#)
- Cottee, M. A., Muschalik, N., Johnson, S., Leveson, J., Raff, J. W., and Lea, S. M. (2015) The homo-oligomerization of both Sas-6 and Ana2 is required for efficient centriole assembly in flies. *eLife* **4**, e07236 [CrossRef Medline](#)
- Nievergelt, A. P., Banterle, N., Andany, S. H., Gönczy, P., and Fantner, G. E. (2018) High-speed photothermal off-resonance atomic force microscopy reveals assembly routes of centriolar scaffold protein SAS-6. *Nat. Nanotechnol.* **13**, 696–701 [CrossRef Medline](#)
- Pfreundschuh, M., Alsteens, D., Hilbert, M., Steinmetz, M. O., and Müller, D. J. (2014) Localizing chemical groups while imaging single native proteins by high-resolution atomic force microscopy. *Nano Lett.* **14**, 2957–2964 [CrossRef Medline](#)
- Hilbert, M., Noga, A., Frey, D., Hamel, V., Guichard, P., Kraatz, S. H., Pfreundschuh, M., Hosner, S., Flückiger, I., Jaussi, R., Wieser, M. M., Thielges, K. M., Deupi, X., Müller, D. J., Kammerer, R. A., et al. (2016) SAS-6 engineering reveals interdependence between cartwheel and microtubules in determining centriole architecture. *Nat. Cell Biol.* **18**, 393–403 [CrossRef Medline](#)
- Bettencourt-Dias, M., Hildebrandt, F., Pellman, D., Woods, G., and Godinho, S. A. (2011) Centrosomes and cilia in human disease. *Trends Genet.* **27**, 307–315 [CrossRef Medline](#)
- Venghateri, J. B., Jindal, B., and Panda, D. (2015) The centrosome: a prospective entrant in cancer therapy. *Expert Opin. Ther. Targets* **19**, 957–972 [CrossRef Medline](#)
- Gönczy, P. (2015) Centrosomes and cancer: revisiting a long-standing relationship. *Nat. Rev. Cancer* **15**, 639–652 [CrossRef Medline](#)
- Thornton, G. K., and Woods, C. G. (2009) Primary microcephaly: do all roads lead to Rome? *Trends Genet.* **25**, 501–510 [CrossRef Medline](#)
- Nigg, E. A., and Holland, A. J. (2018) Once and only once: mechanisms of centriole duplication and their deregulation in disease. *Nat. Rev. Mol. Cell Biol.* **19**, 297–312 [CrossRef Medline](#)
- Zhao, Y., and Wang, X. (2019) PLK4: a promising target for cancer therapy. *J. Cancer Res. Clin. Oncol.* **145**, 2413–2422 [CrossRef Medline](#)

## Inhibitors of SAS-6 oligomerization

30. Mason, J. M., Lin, D. C., Wei, X., Che, Y., Yao, Y., Kiarash, R., Cescon, D. W., Fletcher, G. C., Awrey, D. E., Bray, M. R., Pan, G., and Mak, T. W. (2014) Functional characterization of CFI-400945, a polo-like kinase 4 inhibitor, as a potential anticancer agent. *Cancer Cell* **26**, 163–176 [CrossRef Medline](#)
31. Wong, Y. L., Anzola, J. V., Davis, R. L., Yoon, M., Motamedi, A., Kroll, A., Seo, C. P., Hsia, J. E., Kim, S. K., Mitchell, J. W., Mitchell, B. J., Desai, A., Gahman, T. C., Shiau, A. K., and Oegema, K. (2015) Cell biology. Reversible centriole depletion with an inhibitor of polo-like kinase 4. *Science* **348**, 1155–1160 [CrossRef Medline](#)
32. Lei, Q., Xiong, L., Xia, Y., Feng, Z., Gao, T., Wei, W., Song, X., Ye, T., Wang, N., Peng, C., Li, Z., Liu, Z., and Yu, L. (2018) YLT-11, a novel PLK4 inhibitor, inhibits human breast cancer growth via inducing maladjusted centriole duplication and mitotic defect. *Cell Death Dis.* **9**, 1066 [CrossRef Medline](#)
33. Kawakami, M., Mustachio, L. M., Zheng, L., Chen, Y., Rodriguez-Canales, J., Mino, B., Kurie, J. M., Roszik, J., Villalobos, P. A., Thu, K. L., Silvester, J., Cescon, D. W., Wistuba, I. I., Mak, T. W., Liu, X., et al. (2018) Polo-like kinase 4 inhibition produces polyploidy and apoptotic death of lung cancers. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 1913–1918 [CrossRef Medline](#)
34. Suri, A., Bailey, A. W., Tavares, M. T., Gunosewoyo, H., Dyer, C. P., Gruppenmacher, A. T., Piper, D. R., Horton, R. A., Tomita, T., Kozikowski, A. P., Roy, S. M., and Sredni, S. T. (2019) Evaluation of protein kinase inhibitors with PLK4 cross-over potential in a pre-clinical model of cancer. *Int. J. Mol. Sci.* **20**, 2112 [CrossRef Medline](#)
35. Keller, D., Orpinell, M., Olivier, N., Wachsmuth, M., Mahen, R., Wyss, R., Hachet, V., Ellenberg, J., Manley, S., and Gönczy, P. (2014) Mechanisms of HsSAS-6 assembly promoting centriole formation in human cells. *J. Cell Biol.* **204**, 697–712 [CrossRef Medline](#)
36. Chen, X., and Reynolds, C. H. (2002) Performance of similarity measures in 2D fragment-based similarity searching: comparison of structural descriptors and similarity coefficients. *J. Chem. Inf. Comput. Sci.* **42**, 1407–1414 [CrossRef Medline](#)
37. Busch, J. M. C., Erat, M. C., Blank, I. D., Musgaard, M., Biggin, P. C., and Vakonakis, I. (2019) A dynamically interacting flexible loop assists oligomerization of the *Caenorhabditis elegans* centriolar protein SAS-6. *Sci. Rep.* **9**, 3526 [CrossRef Medline](#)
38. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 [CrossRef Medline](#)
39. Rogala, K. B., Dynes, N. J., Hatzopoulos, G. N., Yan, J., Pong, S. K., Robinson, C. V., Deane, C. M., Gönczy, P., and Vakonakis, I. (2015) The *Caenorhabditis elegans* protein SAS-5 forms large oligomeric assemblies critical for centriole formation. *eLife* **4**, e07410 [CrossRef Medline](#)
40. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423 [CrossRef Medline](#)
41. Mayer, C., Slater, L., Erat, M. C., Konrat, R., and Vakonakis, I. (2012) Structural analysis of the *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *J. Biol. Chem.* **287**, 7182–7189 [CrossRef Medline](#)
42. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 [CrossRef Medline](#)
43. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Res.* **95**, 214–220 [CrossRef Medline](#)
44. Goddard, T. D., and Kneller, D. G. SPARKY 3, University of California, San Francisco
45. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 [CrossRef Medline](#)
46. Winter, G. (2010) Xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 [CrossRef](#)
47. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 [CrossRef Medline](#)
48. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C., and Adams, P. D. (2008) Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426**, 419–435 [CrossRef Medline](#)
49. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 [CrossRef Medline](#)
50. Accelrys (2012) Discovery Studio Environment. *Release 3.5 Ed.*, Accelrys Software Inc. San Diego
51. Sterling, T., and Irwin, J. (2015) ZINC 15—ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 [CrossRef Medline](#)
52. O’Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011) Open Babel: An open chemical toolbox. *J. Cheminform.* **3**, 33 [CrossRef Medline](#)
53. Matsoukas, M.-T., Aranguren-Ibáñez, Á., Lozano, T., Nunes, V., Lasarte, J. J., Pardo, L., and Pérez-Riba, M. (2015) Identification of small-molecule inhibitors of calcineurin-NFATc signaling that mimic the PxIxIT motif of calcineurin binding partners. *Sci. Signal.* **8**, ra63 [CrossRef Medline](#)
54. Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 [CrossRef Medline](#)
55. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 [CrossRef](#)
56. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 [CrossRef Medline](#)
57. Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 [CrossRef](#)
58. Bussi, G., Zykova-Timan, T., and Parrinello, M. (2009) Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *J. Chem. Phys.* **130**, 074101 [CrossRef Medline](#)
59. Hess, B., Bekker, H., Berendsen, H. J., and Fraaije, J. G. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 [CrossRef](#)
60. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An  $N_s^2$  log( $N$ ) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 [CrossRef](#)
61. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, U.S.A.
62. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 [CrossRef Medline](#)