Atomic structure of a toxic, oligomeric segment of SOD1 linked to amyotrophic lateral sclerosis (ALS)

Smriti Sangwan, Anni Zhao, Katrina L. Adams, Christina K. Jayson, Michael R. Sawaya, Elizabeth L. Guenther, Albert C. Pan, Jennifer Ngo, Destaye M. Moore, Angela B. Soria, Thanh D. Do, Lukasz Goldscheid, Rebecca Nelson, Michael T. Bowers, Carla M. Koehler, David E. Shaw, Bennett G. Novitch, and David S. Eisenberg

Fibrils and oligomers are the aggregated protein agents of neuronal dysfunction in ALS diseases. Whereas we now know much about fibril architecture, atomic structures of disease-related oligomers have eluded determination. Here, we determine the corkscrew-like structure of a cytotoxic segment of superoxide dismutase 1 (SOD1) in its oligomeric state. Mutations that prevent formation of this structure eliminate cytotoxicity of the segment in isolation as well as cytotoxicity of the ALS-linked mutants of SOD1 in primary motor neurons and in a Danio rerio (zebrafish) model of ALS. Cytotoxicity assays suggest that toxicity is a property of soluble oligomers, and not large insoluble aggregates. Our work adds to evidence that the toxic oligomeric entities in protein aggregation diseases contain antiparallel, out-of-register β-sheet structures and identifies a target for structure-based therapeutics in ALS.

Significance

More than 170 mutations in superoxide dismutase 1 (SOD1) are linked to inherited forms of ALS, and aggregates of this protein are a pathological feature associated with this disease. Although it is accepted that SOD1 gains a toxic function in the disease state, a molecular understanding of the toxic species is lacking. Here, we identify a short segment of SOD1 that is both necessary and sufficient for toxicity to motor neurons. The crystal structure of the segment reveals an out-of-register β-sheet oligomer, providing a structural rationale for the toxic effects of mutant SOD1 in ALS.


Reviewers: J.P.T., St. Jude Children’s Medical Hospital; and P.W., University of California, San Francisco.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes SOD1 and SIW).

1To whom correspondence should be addressed. Email: david@mbi.ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1705091114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1705091114
sheets do in amyloid fibrils, but instead contact weakly through polar and charged side chains scattered over the exterior of the core. Val29, Val31, Ser34, and water-mediated contacts. Hence, unlike amyloid fibrils, the corkscrew has no dry interface between sheets to stabilize its assembly. Instead, the corkscrew assembly is stabilized by weaker hydrophobic forces arising from the concave interior filled with aliphatic side chains of Val29, Val31, Ile35, and Leu38 (Fig. 1 and SI Appendix, Fig. S3).

The absence of a stable, amyloid-like, dry-sheet-sheet interface suggests that fragmentation of the corkscrew could be relatively facile and that its subsets of various sizes could fit the definition of soluble oligomers. Indeed, ion mobility MS experiments confirm that this SOD1 segment 28–38 forms low-molecular-weight oligomers in solution similar in cross-section to the crystal structure of the corkscrew (SI Appendix, Fig. S6), supporting our hypothesis that the corkscrew represents the structure of a soluble oligomer. Furthermore, the β-sheet–rich nature of the corkscrew is a property shared in common with other amyloid-related oligomers, such as α-synuclein, amyloid-β, and HET-s (8, 9, 24). In fact, several amyloid oligomers have been reported to share antiparallel, out-of-register β-strand architecture (24–26).

The corkscrew shares several structural features with another soluble oligomer, cylindrin, from the nonpathogenic amyloid-forming protein, αB-crystallin (27). Both oligomers are composed of antiparallel β-strands, shifted out-of-register by two residues. Both sheets are highly curved, and opposite faces of the sheet display different sets of side chains not related by symmetry (28). In both oligomers, individual strands hydrogen-bond to neighboring strands through alternating weak and strong interfaces (SI Appendix, Fig. S5A). The strong interface of the corkscrew is composed of nine interchain hydrogen bonds, whereas the corkscrew, although highly curved, is not entirely closed and is likely to exist in a range of oligomeric sizes.

The role of the corkscrew in ALS is supported by a model accommodating full-length SOD1. We modeled the remainder of SOD1 around the corkscrew scaffold, keeping the tertiary structure of SOD1 intact everywhere except near the corkscrew and avoiding steric conflict (Fig. 1B). In our model, strands 2 and 3 detach from the native fold, exposing the corkscrew-forming residues 28–38. This local unfolding may be triggered by dissociation of the SOD1 dimer and metal depletion. Biochemical studies have noted that metal-depleted monomer is prevalent in patient tissues (29, 30) as well as in mouse models, and is an intermediate in the conversion of native holo-SOD1 to pathological aggregates (31). This model agrees with H/D exchange and with MD and MS studies showing that most familial SOD1 mutants have minimal change in their secondary structure and contain a partially unfolded β-barrel at physiological temperature with local unfolding in β-strand 3 (20, 31).

We probed the role of the corkscrew in SOD1-associated cytotoxicity by introducing mutations to disrupt corkscrew architecture. The absence of a bulky side chain at position 33 appears essential to form the concave inner surface of the corkscrew (Fig. 1C). Mutation of Gly33 to a larger residue, such as tryptophan (Fig. 1D) or valine, would introduce severe steric clashes with Val29 and Val31, destabilizing the corkscrew structure. Consistent with this observation, our all-atom MD simulations revealed that the corkscrew was less stable for W33 than for G33, regardless of whether the N terminus was proline or lysine (Fig. 1E and SI Appendix, Fig. S4C). Thus, if the corkscrew were a cytotoxic motif, we would expect G33 mutants to alleviate toxicity of SOD1 familial mutants.

**Segment 28-38 Is Necessary and Sufficient for Toxicity.** Corkscrew-disruptive mutations alleviated toxicity of segment 28–38. We assayed cytotoxicity by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reduction in embryonic stem-cell-derived motor neurons expressing GFP to facilitate visualization of neuron morphology (32). The corkscrew-forming segment was aggregated, applied to motor neurons, and incubated overnight. We found that
viability was reduced by 60% compared with buffer-treated cells at physiological concentrations of 8–100 µM in a dose-dependent manner (Fig. 2A). In contrast, neither the segment harboring the corkscrew-disruptive G33W mutation (Fig. 2B) nor the less bulky G33V mutant (Fig. 2B) induced toxicity in cells at any concentration tested. The same trends were observed with the native 28–38 segment (P28) and the corresponding G33W and G33V mutations (SI Appendix, Fig. S7).

Corkscrew-disruptive mutations also alleviated toxicity of full-length SOD1 familial mutants A4V and G93A. These proteins were recombinantly expressed, purified, and aggregated by demetallation and agitation at 37 °C for 12 h, which produced a mixture of fibrils and insoluble fibrils. We tested the cytotoxic effects of the non–fibril-forming mutant (I104P) (33). Even though it did not form any mature fibrils (Fig. S4), it was toxic to motor neurons, and addition of the corkscrew-disrupting substitution, G33W, alleviated the cytotoxicity (Fig. 3B and C). These results suggest that fibril formation is not essential for cytotoxicity.

To identify which species of SOD1 aggregate is toxic, we monitored the toxicity of various SOD1 mutants as their aggregates

Toxicity of Full-Length SOD1 Derives from Soluble Oligomers. We asked if toxicity of full-length SOD1 derives from soluble oligomers or

Fig. 2. Corkscrew-forming segment 28–38 is necessary and sufficient for cytotoxicity. (A) Cell viability of motor neurons measured by an MTT reduction assay shows that the corkscrew segment (KVKVWGSIKGL) is toxic to primary motor neurons in a dose-dependent manner. Results are shown as mean ± SD (n = 3). Symbols represent individual values of triplicates, and bars represent average values. Statistical significance was analyzed using two-tailed t tests with Welch’s correction. (B) Corkscrew-forming segment (28–38) harboring single-point substitutions at Gly33 (G33V and G33W) is nontoxic to motor neurons. All peptide segments were prepared identically, and motor neurons were treated with different final concentrations. The statistical significance of G33V and G33W mutants was compared with segment 28–38 by two-way ANOVA. (C) Hb9-GFP-labeled motor neurons treated with 8 µM aggregated full-length familial mutants (A4V and G93A) lose neurites, but the corresponding corkscrew-disrupting mutants (G93A/G33V, G93A/G33W, A4V/G33V, and A4V/G33W) are nontoxic and neurites look healthy. (Scale bars, 20 µm.) (D) Cell viability measured by an MTT reduction assay confirming that the familial mutants A4V and G93A are toxic and that substitution of Gly33 with valine or tryptophan renders the protein nontoxic. Results are shown as mean ± SD (n = 3). Symbols represent individual values of triplicates, and bars represent average values. Statistical significance was analyzed by one-way ANOVA (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001).

Fig. 3. Toxicity of full-length SOD1 derives from soluble oligomers. (A) Electron micrographs of a non–fibril-forming SOD1 mutant (I104P) and the corresponding double mutant (I104P/G33W) show some aggregates but no large fibrils. (B) Motor neurons treated with I104P lose neurites and have shrunken cell bodies (Left), but I104P/G33W-treated cells look healthy (Right). (Scale bars, 20 µm.) (C) Cell viability measured by an MTT reduction assay confirmed that I104P is toxic and I104P/G33W is nontoxic. Statistical significance was analyzed using a two-tailed t test with Welch’s correction. (D and E) Toxic properties of SOD1 mutants depend on the duration of aggregation. The A4V and G93A mutants aggregated for 12–16 h are toxic to motor neurons, whereas extended agitation for 72 h renders the proteins nontoxic. The corkscrew-disrupting proteins (A4V/G33W and G93A/G33W) are nontoxic irrespective of the duration of aggregation. Results are shown as mean ± SD (n = 3). Symbols represent individual values of triplicates, and bars represent average values. Statistical significance was analyzed by two-tailed t tests with Welch’s correction (*P < 0.05, **P < 0.01). (F) Representative electron micrographs of various preparations of the familial mutants A4V and G93A and the double mutants A4V/G33W and G93A/G33W. Some large aggregates can be seen at 2- to 16-h time points, but no fibrils can be seen, whereas all constructs show large fibril loads at 72 h. (G) Immunoblots of the familial mutants aggregated for different time points. Samples aggregated for 12–16 h are A11-positive, and both proteins lose A11 reactivity when aggregated for 72 h. SOD100 was used as a loading control.
evolved over time. We found that both familial mutants (A4V and G93A) were toxic to cultured neurons when aggregated for 12–16 h but that extended aggregation for 72 h rendered the protein nontoxic irrespective of the duration of aggregation. Negative-stain EM showed abundant fibrils in the samples aggregated for 72 h (Fig. 3F), and immunoblotting with the conforma-tional oligomer-specific antibody (A11) suggested that samples aggregated for 72 h contained no oligomers (Fig. 3G). These results suggest that toxicity of aggregated SOD1 mutants derives from oligomers.

Whereas the relevance of our work to SOD1-related ALS seems clear, the relevance to non–SOD1-related ALS is less clear. A few reports suggest that wild-type SOD1 (WT) aggregates in sporadic ALS (34–36) analogous to mutant SOD1 aggregates in familial ALS; however, it is unknown if WT and mutant SOD1 aggregate by the same mechanism or have the same toxicity. We compared the aggregation and cytotoxic properties of WT and mutant SOD1. We expressed A4V SOD1 with and without the corkscrew-disrupting substitution at Gly33 in a zebrafish TDL6 line in which the primary motor neurons are labeled with GFP and the mitochondria are labeled with dsRed (41). We analyzed the axons at 2 d postfertilization. The A4V mutation caused an 8% reduction in axon length (Fig. 4A and B), as has been previously shown (39), but axon lengths of A4V/G33V-injected fish were significantly longer (Fig. 4A and B). Additionally, we observed that 30% of A4V-injected fish were severely deformed and could not be imaged, suggesting that an acute phenotype is lethal (SI Appendix, Fig. S9C). In contrast, the WT and double-mutant-expressing fish did not display large mortality. We observed a similar phenotype upon expression of the G93A familial mutant; G93A-expressing zebrafish have 5% reduction in axon lengths, whereas axon lengths of G93A/G33V-expressing zebrafish were significantly longer (Fig. 4C and D).

Defects in mitochondrial assembly and trafficking, along with vacuolation (42, 43) and abnormal clustering in neuronal processes (44–46), are established pathological phenotypes observed in the C. elegans model of familial SOD1-linked ALS (37–39). A4V/G33W and G93A/G33W were toxic to cultured neurons with or without corkscrew-disrupting mutations. Although we observed a similar phenotype upon expression of the G93A familial mutant; G93A-expressing zebrafish have 5% reduction in axon lengths, whereas axon lengths of G93A/G33V-expressing zebrafish were significantly longer (Fig. 4C and D).

Corkscrew Disruption Alleviates Defects in a Danio rerio (Zebrafish) ALS Model. To determine whether corkscrew-disrupting mutations alleviate axonopathies caused by familial SOD1 mutants in vivo, we conducted experiments using a zebrafish model of ALS (37–40). We expressed A4V SOD1 with and without the corkscrew-disrupting substitution at Gly33 in a zebrafish TDL6 line in which the primary motor neurons are labeled with GFP and the mitochondria are labeled with dsRed (41). We analyzed the axons at 2 d postfertilization. The A4V mutation caused an 8% reduction in axon length (Fig. 4A and B), as has been previously shown (39), but axon lengths of A4V/G33V-injected fish were significantly longer (Fig. 4A and B). Additionally, we observed that 30% of A4V-injected fish were severely deformed and could not be imaged, suggesting that an acute phenotype is lethal (SI Appendix, Fig. S9C). In contrast, the WT and double-mutant-expressing fish did not display large mortality. We observed a similar phenotype upon expression of the G93A familial mutant; G93A-expressing zebrafish have 5% reduction in axon lengths, whereas axon lengths of G93A/G33V-expressing zebrafish were significantly longer (Fig. 4C and D).

Defects in mitochondrial assembly and trafficking, along with vacuolation (42, 43) and abnormal clustering in neuronal processes (44–46), are established pathological phenotypes observed in the C. elegans model of familial SOD1-linked ALS (37–39). A4V/G33W and G93A/G33W were toxic to cultured neurons with or without corkscrew-disrupting mutations. Although we observed a similar phenotype upon expression of the G93A familial mutant; G93A-expressing zebrafish have 5% reduction in axon lengths, whereas axon lengths of G93A/G33V-expressing zebrafish were significantly longer (Fig. 4C and D).
in transgenic mice, patient-derived cells, and other models. However, they have not been reported in any zebrafish model of ALS thus far. Therefore, we analyzed the mitochondrial morphology upon expression of SOD1 familial mutants. Expression of A4V and G93A mutant protein caused remarkable mitochondrial pathology characterized by abnormal diffused clustering at the branch points indicative of defective mitochondria (Fig. 4E), whereas A4V/G33V-expressing fish had a mitochondrial network similar to WT fish. These defects were quantified by measuring the size and fluorescence intensity of the mitochondria, confirming that A4V-expressing fish displayed enlarged mitochondria (Fig. 4F and SI Appendix, Fig. S9), which were fewer in number (Fig. 4G and SI Appendix, Fig. S9). We also observed similar A4V-G33V zebrafish (Fig. 4H and SI Appendix, Fig. S9). Thus, disrupting the corkscrew segment alleviates ALS-linked axonopathies and mitochondrial defects in this in vivo model.

Discussion

Our experiments suggest that segment 28–38 of SOD1 is important for SOD1-mediated toxicity. The crystal structure of this segment revealed an oligomer composed of antiparallel, out-of-register β-strands, which assemble into a corkscrew-like structure. The G33V and G33W point mutants, which were designed to disrupt the observed oligomer, alleviated toxicity of both the isolated peptide and full-length SOD1. In a zebrafish model of ALS, G33V prevented axonopathies and mitochondrial defects, two characteristic features of ALS-linked pathology. Taken together, these results suggest that the corkscrew structure is critical for SOD1-mediated cytotoxicity.

The corkscrew structure explains its oligomeric state and suggests the identity of its potential interacting partners in the cell. The corkscrew is composed of a single twisted sheet rather than pairs of tightly mated sheets, as observed in αS stereic structure published thus far. A clue about the identity of the corkscrew’s interacting partners in the cell is offered by examining the functions of its structural homologs. A search for corkscrew homologs in the Protein Data Bank using the DALI server (47) yielded matches with other highly twisted β-sheets proteins, such as membrane receptor proteins, enzymes, and bactericidal-permeability increasing (BPI) protein (SI Appendix, Fig. S10). The twisted sheet seen in the crystal structure of BPI has been shown to bind lipids and destabilize membranes (48). It is conceivable that the cleft seen in the corkscrew structure is important for cytotoxicity, potentially as a binding site for lipids. The cleft of corkscrew is accessible to lipids and small molecules. In contrast, cylinder has no accessible cleft, which could explain its lower cytotoxicity relative to corkscrew.

Our results demonstrate that toxicity derives from the corkscrew oligomers rather than from fibrils (SI Appendix, Fig. S11). Previously, we have shown that out-of-register oligomers are likely off-pathway from in-register fibril formation due to the large energetic cost of rearrangement of out-of-register oligomers into in-register fibrils (49). Although we cannot ascertain whether the fibrils observed in our experiments are in-register, our results are consistent with this hypothesis. The corkscrew-disrupting mutations of G33V, G33W attenuate cytotoxicity but do not attenuate fibril formation.

Cytoxicity assays of the non–fibril-forming mutant (I104P) and the time-course assays with the familial mutants (A4V and G93A) suggest that toxicity is a property of soluble oligomers, and not of large insoluble fibrils. These findings for SOD1 align with the hypotheses proposed by others for amyloid-β and huntingtin that large misfolded aggregates are not directly toxic to neurons.

From a molecular perspective, it would be unlikely to find ALS-linked mutations in the 28–38 segment of SOD1, given its structural importance for mediating toxicity. Indeed, compared with other regions of SOD1, this segment contains few familial mutations, and no mutations are found in the core of this segment spanning residues 32–36 (19). Notably, familial mutants, including G37R and the rare mutants V29A and V31A, are found near the ends of this segment. From our crystal structure, we infer that all these mutations are compatible with the corkscrew structure, although it is unclear if they actively promote oligomer assembly.

In summary, we have identified an 11-residue segment in ALS-associated SOD1 that is necessary for its cytotoxicity. Our data support the hypothesis proposed by others for amyloidogenic diseases.

Materials and Methods

Crystals of SOD1(28–38) with P28K substitution were grown by hanging drop vapor diffusion using VDX plates (Hampton Research). Lyophilized peptide at 98% purity (GenScript, Inc.) was dissolved to 50 mg/ml in 50 mM Tris base buffer. The reservoir solution contained 0.2 M sodium citrate (pH 5) and 13% PEG 6000. All SOD1 constructs were expressed recombinantly in Escherichia coli. Hb9-eGFP mouse embryonic stem cells were maintained and differentiated into motor neurons as previously described (32). Aggregated protein preparations were added to cultured neurons at the given final concentration, and viability was measured by MTT reduction assay. Details are provided in SI Appendix, Materials and Methods.

All zebrafish (Danio rerio) were maintained in accordance with standard laboratory conditions (23). The University of California, Los Angeles Chancellor’s Animal Research Committee approved all experiments performed on zebrafish.

ACKNOWLEDGMENTS. We thank Lisa Johnson, David Borchelt, and Joan Valentine for discussions; Hamilton Trinh, Michael Collazo, Duliu Casioc, and staff at Argonne Photon Source, Northeastern Collaborative Access Team beamline 24-ID-E. We thank Hynek Wichterle (Columbia University) for the gift of HB-9-eGFP embryonic stem cells. We are grateful for the support to D.S.E. from the Howard Hughes Medical Institute, Department of Energy, and a grant from the National Institutes of Health (NIH) (AG029430). B.G.N. was supported by the University of California, Los Angeles (UCLA) Broad Center of Regenerative Medicine and Stem Cell Research, the Rose Hills Foundation, and a grant from the National Institute of Neurological Disorders and Stroke (NS072804). Muscular Dystrophy Association (92901), and the California Institute for Regenerative Medicine (CIRM) (RB1-01367 and RBS-07480). C.M.K. was supported by grants from CIRM (RT307678) and the National Institute of General Medical Sciences (GM61721). M.T.B. was supported by grants from the NIH (AG047116) and the National Science Foundation (CHE-130132 and CHE-1565941). S.S. was supported by a Whitcome Pre-Doctoral fellowship; K.L.A. was supported by the UCLA Cellular and Molecular Biology Training program (Ruth L. Kirschstein NIH Grant GM007185), a UCLA-California Institute for Regenerative Medicine Training Grant, and a UCLA Graduate Division Dissertation Year Fellowship; R.N. supported by a Larry L. Hillblom Foundation Fellowship; and C.K.J. was supported by a Beckman Research Scholarship.

7. Ben-Shalom T, et al. (2012) SOD1(28–38) with P28K substitution were grown by hanging drop vapor diffusion using VDX plates (Hampton Research). Lyophilized peptide at 98% purity (GenScript, Inc.) was dissolved to 50 mg/ml in 50 mM Tris base buffer. The reservoir solution contained 0.2 M sodium citrate (pH 5) and 13% PEG 6000. All SOD1 constructs were expressed recombinantly in Escherichia coli. Hb9-eGFP mouse embryonic stem cells were maintained and differentiated into motor neurons as previously described (32). Aggregated protein preparations were added to cultured neurons at the given final concentration, and viability was measured by MTT reduction assay. Details are provided in SI Appendix, Materials and Methods.


Supporting Information:

Materials and Methods

SOD1(28-38) structure determination
Crystals of SOD1(28-38) with P28K substitution were grown by hanging drop vapor diffusion using VDX plates (Hampton Research, Aliso, Viejo, CA). Lyophilized peptide at 98% purity (Genscript Inc.) was dissolved to 50 mg/ml in 50 mM Tris-base buffer. The reservoir solution contained 0.2 M sodium citrate pH 5 and 13% PEG 6000. Crystallization drops were prepared by mixing peptide with reservoir in a 2:1 ratio, in a total volume of 2 µL. Thick needle-like crystals appeared overnight and harvested after 2-3 days growth at ambient temperature. For isomorphous replacement, the crystals were soaked for 30 seconds in a solution composed of approximately 60% KI stock solution (prepared by dissolving solid KI to 0.5 M in reservoir), 30% reservoir solution, and 10% glycerol stock for cryoprotection. The crystals were subsequently flash frozen in liquid nitrogen. Single crystals were mounted with CrystalCap HT Cryoloops (Hampton Research, Aliso Viejo, CA). All data were collected at the Advanced Photon Source (Chicago, IL) on beamline 24-ID-E, which has a 5 µm microfocus beam suitable for the crystals. All data were processed using DENZO and SCALEPACK or XDS (1). SIRAS phasing produced an interpretable map and the structure was built using COOT (2). Model refinement was performed using REFMAC (3) and BUSTER (4). Figures were generated using PyMol (5). Unless stated otherwise, SOD1(28-38) refers to KVKVWGSIKGL.

Molecular Dynamics simulations
We performed all-atom MD simulations of SOD1 and αB crystallin on Anton (6, 7), a special-purpose computer designed to accelerate standard molecular dynamics simulations by orders of magnitude. Protein, water, and ions were represented explicitly, using the AMBER 99SB*-ILDN (8–10) force field and the TIP3P water model (11). Simulations of the SOD1 corkscrew were initiated from the crystal structure comprised of an eight-stranded corkscrew with either a wild-type (KVKVWGSIKGL and PVKVWGSIKGL) or G33W mutant (KVKVWWSIKGL and PVKVWWSIKGL) monomer sequence. Initial structures were placed in a cubic box 76 Å on each side and solvated with approximately 13,000 water molecules and 150 mM NaCl. The systems were minimized and then equilibrated in the NPT ensemble at 1 bar and 300 K for 50 ns with 5 kcal mol⁻¹ Å⁻² harmonic position restraints applied to all non-hydrogen atoms of the protein; these restraints were tapered off linearly over 50 ns. Production runs were initiated from the final snapshot of the equilibration run and performed in the NVT ensemble using a 2.5 fs time step. Non-bonded interactions were truncated at 12 Å. The wild-type (KVKVWGSIKGL and PVKVWGSIKGL) and G33W mutant (KVKVWWSIKGL and PVKVWWSIKGL) corkscrew simulations were run for 16.4 µs, 13.8 µs, 16.6 µs and 15.2 µs, respectively. RMSDs in Figs. 1E and S4C were calculated on the hydrophobic core of the corkscrew (Ca atoms of residues 3 to 9, excluding the chains at the end).

SOD1 oligomer assembly simulations were initiated from three monomers of a tandem repeat structure (KVKVWGSIKGL--GG--KVKVWGSIKGL) built in homology to a monomer of the tandem repeat structure of αB crystallin cylindrin (PDB ID 3SGR (12)). These monomers were separated in space by at least 30 Å,
placed in a box 62 Å on each side, and solvated with approximately 7,600 water molecules and 150 mM NaCl. System preparation and simulations then proceeded as described above for the SOD1 corkscrew simulations. Each monomer was weakly restrained such that its Cα root-mean-square deviation (RMSD) was within approximately 2 Å of the starting structure. In particular, a flat-bottom harmonic restraint, $U$, was applied to each monomer such that $U(R) = \frac{1}{2}k(R - R_0)^2$ for $R > R_0$, and $U(R) = 0$ otherwise. Here, $R$ is the RMSD from the initial structure, $R_0 = 2$ Å, and $k = 50$ kcal mol$^{-1}$ Å$^{-2}$. The flat-bottom RMSD restraints limit the monomers’ sampling of very collapsed structures and structures that have significantly reduced β-sheet content, and allow simulations to focus on discovering possible amyloid-like oligomers. We ran 20 assembly simulations beginning from the same starting conformation with different initial velocities drawn from a Maxwell-Boltzmann distribution. Of the twenty simulations, four were run for 4 µs and the rest were run for 1 µs.

The αB crystallin assembly simulations were initiated from three monomers of the tandem repeat structure of αB crystallin (GKLKVLGDVIEV--GG--KLKVLGDVIEV, PDB ID 3SGR). System preparation and simulation protocol were as above, except that we ran 10 assembly simulations in this case. Of these 10 simulations, two were run for 5 µs, two were run for 2 µs, and the rest were run for 1 µs.

**Full-length model of corkscrew oligomers**

We tested the hypothesis that full-length SOD1 is compatible with a corkscrew oligomer assembly by constructing a model. We aimed to achieve the following three goals: (i) maintain residues 28-38 in the exact corkscrew conformation as observed in the crystal structure of this segment; (ii) maintain the tertiary structure of native SOD1 intact for all remaining residues; (iii) maintain good stereochemistry and avoid steric overlap throughout. Coordinates of the native SOD1 fold were obtained from PDB ID code 2C9S, the 1.24 Å resolution structure of Zn-Zn human superoxide dismutase (13).

It was obvious by inspection that residues 28-38 must detach from the native SOD1 fold to expose its main chain for hydrogen bonding with the identical segment of other protomers and thereby form the twisted β-sheet that constitutes the corkscrew spine. Residues 28-38 encompass the third β-strand (β3) of the native SOD1 8-stranded β-barrel fold. This strand is hydrogen-bonded on both edges with its neighboring strands in the barrel, β2 and β6. The least disruptive way that we could find to expose β3 was to detach both β2 and β3 from the native barrel. This pair of strands is connected by only a short hairpin, so β3 would not be able to fully detach from the barrel without the co-detachment of β2. The alternative possibility, which is to detach strands 3 and 4 from the native fold, would disrupt more hydrogen bonds and van der Waals contacts.

We positioned and oriented the native SOD1 barrels in close proximity to the fixed coordinates of the corkscrew crystal structure (residues 28-38). The degrees of freedom for positioning and orienting the SOD1 barrels were limited by the constraints that barrel residues 27 and 39 be placed within a covalent bond length of residues 28 and 38 of the corkscrew, and that there be no steric clashes between
neighboring protomers. To aid in the positioning and orienting of the SOD1 barrels, we defined the helical symmetry operations of the corkscrew structure (pitch 71 Å, 16 strands per turn, two-fold symmetry axes perpendicular to helix axis), and applied these operations to a single SOD1 barrel. We interactively manipulated a single SOD1 barrel in the graphics program “O” (14), which simultaneously propagated each manipulation to the helical symmetry-related copies. Thus we were able to visually explore the orientations and positions of the SOD1 barrel that met the constraints of close proximity and steric clash avoidance while maintaining the helical symmetry of the corkscrew.

After suitably positioning and orienting the SOD1 barrel around the corkscrew spine, we covalently linked barrel residues 27 and 39 to corkscrew residues 28 and 38, respectively, using the graphics program Coot (15). At the N-terminal connection (residue 27 to 28), we note that we reoriented β2 (residues 15-22) to disrupt its main chain hydrogen bonds with β3, as was necessary to permit copies of β3 to self-associate as a corkscrew. To recover some of the main chain hydrogen bonds that were lost when β2 was separated from β3, we positioned pairs of β2 strands from neighboring protomers to form main chain hydrogen bonds. To make the connections stereochemically plausible, it was necessary to slightly perturb the conformation of the three hairpin forming residues at the C-terminus of the corkscrew β-strand (residues 36-38). One full turn, 16 protomers, were energy minimized with the program CNS (16), with hydrogen bond potentials included (17).

Expression and purification of SOD1 constructs

All SOD1 constructs were expressed recombinantly in E. coli. All recombinant proteins are based on a SOD1 background strain that carries a C6A, C111S double mutation to simplify the purification. The SOD1 gene was inserted into the pET22b vector from Novagen with NcoI and SalI restriction enzyme digestion sites (gift from Professor Joan Valentine’s lab at UCLA). Mutations in the SOD1 gene were made using the QuickChange Site-Directed mutagenesis Kit (Stratagene). The plasmid was transformed into the BL21(DE3)Gold (Agilent Technologies) expression strain. For expression, 10 ml LB + Amp (100 µg/mL) was inoculated from frozen stock and grown overnight. 10 ml of starting culture was added to a 2 L flask of 1 L LB + Amp (100 µg/mL) and grown for 3 hours at 37 °C to OD₆₀₀ = 0.6. IPTG was then added to 1 mM and Zn²⁺ was added to 0.05 mM to induce protein expression, which continued for an additional 3 hours. The bacterial pellet was collected by centrifugation at 4000 rpm for 10 mins.

To purify SOD1 from the bacteria, osmotic shock was used to release proteins including SOD1 from the periplasm. First, the cell pellet was resuspended in 30 mL of 30 mM Tris-HCl pH 8 with 20% sucrose and stirred slowly for 20 mins. The cells were collected by centrifugation at 10,000 × g at 4 °C for 10 mins. The pellet was then resuspended in 30 mL of ice-cold water and stirred slowly for 20 mins on ice to release periplasmic proteins. Next, cell debris was removed by centrifugation at 4 °C for 10 mins at 10,000 × g. The contaminant proteins were then removed by precipitation with ammonium sulfate. 0.326 g/ml (NH₄)₂SO₄ was added to the supernatant, and the solution was stirred at 4 °C for 45 mins. The protein was purified out of ammonium sulfate by separation over a phenyl sepharose column. The column was equilibrated with buffer A (2 M (NH₄)₂SO₄, 0.15 M NaCl, 0.05 M NaH₂PO₄, pH 7.0) and buffer B (0.15 M NaCl, 0.05 M NaH₂PO₄, pH 7.0). The protein eluted at 6-
30% buffer B. The protein was then concentrated and further purified by size exclusion chromatography on a silica G3000 column (Tosoh Bioscience). The column buffer comprised 0.1 M sodium sulfate, 25 mM sodium phosphate, and 1 mM sodium azide, pH 6.5.

**SOD1 demetallation (required for aggregation)**
The protein constructs purified by size exclusion chromatography were dialyzed in four buffers at 4°C to remove the bound metals. Protein preparations were put in a dialysis cassette (10 kDa molecular weight cutoff) and exchanged with buffer 1 twice (100 mM Na Acetate pH 3.8, 10 mM EDTA pH 8.0), followed by exchange with buffer 2 (100 mM Na Acetate pH 3.8, 100 mM NaCl, 1 mM EDTA pH 8.0), buffer 3 (100 mM Na Acetate pH 5.5, 100 mM NaCl, 1 mM EDTA pH 8.0), and finally buffer 4 (20 mM K Phosphate pH 7.0, 1 mM EDTA pH 8.0). Each exchange lasted 12 hours. To avoid metal contamination, all stock solutions (except for EDTA) were treated with Chelex 100 resin (Bio-Rad) to remove trace metals and all beakers were rinsed with metal-free water.

**Protein and peptide aggregation**
Full-length SOD1: Following metal removal, proteins were concentrated using a centrifugal concentrator (10 kDa molecular weight cutoff) to 800 µM (monomer) stock concentration and stored at -80 °C in 50 µL aliquots until further use. Aggregation assays were carried out in 1.5 mL microcentrifuge tubes (Fisher Scientific) under acidic conditions with 80 µM SOD1, 50 mM sodium acetate, 1 mM EDTA and 1 mM TCEP, pH 3.5. Preparations were agitated at 900 rpm (in Torrey Pine Shakers) at 37 °C for indicated times. These samples were then dialysed with 10 mM potassium phosphate and 1 mM EDTA, pH 7.0 for 4 hours at 4 °C.

Peptide: Peptide segments were aggregated by incubation at 37 °C for 12 hours with agitation in 50 mM tris-base buffer and added to cultured neurons.

**Differentiation of mESCs to motor neurons**
Hb9:eGFP mESCs were maintained and differentiated into motor neurons as previously described (18). Briefly, mESCs were first plated on gelatin to remove mouse embryonic fibroblasts (MEFs) prior to differentiation, and then plated in 60 mm bacterial petri dishes in core motor neuron medium (DMEM/F12, Neurobasal, 10% Knockout Serum Replacement, Pen-Strep, Glutamax) to induce embryoid body (EB) formation. Two days later, N2 supplement (1x), retinoic acid (1 µM; Sigma), and SAG (1 µM; Calbiochem) were added to the EBs. Media was changed every two days. After five days post-retinoic acid and SAG addition, Hb9:GFP EBs were dissociated using ice-cold 0.25% Trypsin-EDTA for 6 minutes at room temperature, followed by trituration in L-15 medium (Hyclone). Dissociated motor neurons were plated in core motor neuron medium with GDNF, BDNF, CNTF (all 10 ng/ml; Peprotech) on 96-well plates that were previously coated with poly-L-ornithine (0.01%; Sigma) and laminin (5 µg/ml; BD Biosciences) and incubated for 12-16 hrs before treatment with the SOD1 peptides or proteins. Approximately 80,000 cells were plated per well. All reagents are from Invitrogen, unless otherwise noted.

**Cytotoxicity Assays**
Aggregated protein preparations were added to cultured neurons at the given final concentration. Unless stated otherwise, all full-length proteins were aggregated for 12 hours before adding to motor neurons at 8 µM final concentration. Motor neurons were incubated for 12-16 hrs and imaged (Axio Observer.D1 microscope, Zeiss), followed by cell viability assay. Cell viability was measured using the commercially available MTT assay (Promega Cat#G4100) according to the manufacturer’s protocol with 4 hour incubation with the MTT reagent. All data were normalized using buffer as 100% viability and 0.2% SDS as 0% viability. Experiments were performed in triplicates and repeated a minimum of three times.

**Electron Microscopy**
Each 3-5 µL sample was spotted directly on freshly glow-discharged carbon-coated electron microscopy grids (Ted Pella, Redding, CA). After 4 min incubation, grids were rinsed twice with 5 µL distilled water and stained with 5 µl of 2% uranyl acetate for 1 min. Excess uranyl acetate was removed by blotting and specimens were examined on a T-12 electron microscope at an accelerating voltage of 80 kV. Images were recorded digitally by wide angle (top mount) BioScan 600W 1 × 1K digital camera (Gatan, Pleasanton, CA).

**Dot Blot Assay**
3-5 µL sample was spotted on a nitrocellulose membrane (Trans-Blot, Bio-Rad) followed by blocking with 10% fat-free milk in 1X TBST buffer (50 mM Tris, 150 mM NaCl, 0.05% Tween20) for 30 mins. Membranes were incubated with primary antibody – SOD100 (Enzo Life Sciences) at 1:1000 dilution, A11 (EMD Millipore) at 1:500 diluted in 5% fat free milk in TBST buffer at room temperature for 1 hour. The membranes were washed three times in TBST buffer before incubating with HRP-linked secondary antibody - anti-mouse HRP secondary antibody (Life Technologies) at 1:10000 dilution, or anti-rabbit HRP secondary antibody (Life Technologies) at 1:10000 diluted in 5% fat-free milk in TBST buffer at room temperature for 1 hour. Pierce ECL Western Blotting Substrate, (Thermo Scientific, #32209), a chemiluminescent substrate was used for the detection of horseradish peroxidase (HRP) on the immunoblots following the manufacturer’s instructions. A detailed protocol can be found at the following URL, https://tools.thermofisher.com/content/sfs/manuals/MAN0011536_Pierce_ECL_West__Blot_Subs_UG.pdf

**Thioflavin-T (ThT) Assay**
Fibril formation assays were performed with 80 µM protein concentration in 50 mM sodium acetate buffer, 1 mM EDTA, 1 mM TCEP, pH 3.5 and 10 µM ThT – conditions identical to those used for aggregating proteins for the cytotoxicity assays, but with the addition of ThT. All assays were carried out in black Nunc 96-well optical bottom plates (Thermo Scientific). Plates were agitated at 300 rpm in 3-mm rotation diameter in a Varioskan microplate reader (Thermo) at 37 °C. Fluorescence measurements were recorded every 30 mins using λ<sub>ex</sub> = 444 nm, λ<sub>em</sub> = 482 nm, with an integration time of 200 µs.

**Mass Spectrometry**
Peptide segments were dissolved in 20 mM ammonium acetate buffer (pH = 7) to a final concentration of 200 µM and incubated at room temperature without agitation.
Experiments were carried out at time $t_0$ and every day for a week and ion mobility was used to calculate the cross sections of oligomers. The peptide samples were nano-electrosprayed on a home-built, high-resolution ion mobility mass spectrometer (19) consisting of a source ion-funnel, a two-meter long drift cell, an exit funnel, a quadrupole mass analyzer and detectors. The ions were stored and focused in the source funnel and subsequently pulsed into the drift cell filled with helium gas at high pressure. A drift voltage was applied across the cell to create a weak electrical field. The ions moved forward under the influence of the electrical field, and at the same time experienced the drag force due to multiple collisions with helium buffer gas. An ion traveled with a constant velocity that reflected its size and shape. At the end of the drift cell, the ions were collected by an exit funnel where they were steered and focused into a quadrupole mass analyzer and detected as a function of arrival time in the arrival time distributions (ATDs). By measuring ATDs at different pressure to voltage (P/V) ratios, the absolute mobility of the ion of interest could be measured:

$$t_A = \frac{l^2}{K_0 \cdot 760 \text{ torr} \cdot T} \cdot \frac{P}{V} + t_0 \quad \text{(Eq. 1)}$$

In Eq. 1, $l$ is the length of the drift cell, $P$ is the gas pressure in torr, $V$ is the voltage across the cell, and $t_0$ is the time the ions spend outside the drift cell before reaching the detector. The collision cross section $\sigma$ can then be approximated from $K_0$ (20)

$$\sigma \approx \frac{3q}{16N} \left(\frac{2\pi k_B T}{\mu m}\right)^{\frac{1}{2}} \frac{1}{K_0} \quad \text{(Eq. 2)}$$

Here, $N$ is the buffer gas density, $\mu$ is the reduced mass of the collision system (He+ion) and $k_B$ is Boltzmann’s constant. The cross section contains information about the three dimensional structure of the ion. In the ATDs shown in Fig. S6, underneath each experimental peak is a dotted line representing the peak shape expected for a single conformer (21).

**Zebrafish**

All zebrafish (Danio rerio) were maintained in accordance with standard laboratory conditions (22). The University of California, Los Angeles (UCLA) Chancellor’s Animal Research Committee approved all experiments performed on zebrafish. The TDL6 (Tübingen driver line) zebrafish used in all experiments were identified in a screen for developmentally regulated enhancers that drive tissue-specific expression. Gal4-driven GFP expression marks primary motor neurons. Mitochondria were marked by a TOL2-mediated insertion of a Gal4-UAS-MLSDsRed construct (23).

**RNA production and injections**

SOD1 constructs used for zebrafish expressions were the full-length sequence with Cys6 and Cys111 intact. The constructs (WT, G93A, G93A/G33V, A4V, A4V/G33V were subcloned into pcGlobin2 vector and the constructs were linearized using Xbal to generate RNA. The Translation efficiency of the RNA of all constructs was tested by TnT® Quick Coupled Transcription/Translation System (Promega) using the manufacturer’s instructions. Construct G93A/G33V has an additional mutation of I18G designed to prevent steric clashes due to the presence of G33V in the native SOD1 barrel. As a control, we also tested the cytotoxicity of this construct in cell culture and found that it was non-toxic similar to G33V construct (Fig. S9B). RNA was transcribed using the mMessage Machine T7 kit (Ambion) and the mRNA was diluted to 500 ng/µL and 1000 ng/µL aliquots and stored at -80°C. 1 µL mRNA at 500 ng/µL (WT, G93A, G93A/G33V) and at 1000 ng/ µL (WT, A4V, A4V/G33V) was
injected into TDL6 embryos from timed mating at the 1-4 cell stage. Embryos were incubated in 1x E3 buffer at 28.5 °C.

**Analysis of axonopathy**

At 2 dpf (days post fertilization) zebrafish larvae microinjected with SOD1 constructs were anesthetized in 0.01% tricaine and imaged using a Leica MZ16F fluorescent stereoscope at 11.5X magnification. Observers were blinded and the length from the exit point of the spinal cord to the tip of the motor axon was measured for 5 axons per embryo corresponding to axons 12-16 - using Fiji Simple Neurite Tracer. Experiments were repeated multiple times and a minimum of 72 embryos were assessed for each construct.

**Analysis of Mitochondrial defects**

The mitochondrial network from the base of the spinal cord to the tip of the motor axon was imaged on a Leica TCS SP8 confocal microscope at 10X magnification with a 1.5X optical zoom. Images were processed with the Leica AF software and mitochondria size was measured for 5 axons per embryo using Fiji Image J Particle Analysis with the described algorithm (24). Before quantification, fluorescence from the spinal cord and regions outside of the larvae body were excluded. The program allowed automatic detection of isolated fluorescent spots by threshold and watersheding of binary formatted images without introducing user-bias. Spots with a circularity of 0.5-1 were selected to have their intensity and area quantified, where 1 is a complete circle. Intensity was quantified for each spot via overlapping gray levels.

**Protein expression in zebrafish**

Protein expression was confirmed by western blotting (Fig. S9). Zebrafish embryos were pooled from each group at 2 dpf and washed twice in Ringer’s buffer (NaCl 135 mM, KCl 5 mM, MgSO4 1mM, K2HPO4 0.4 mM, Glucose 5.5 mM, HEPES 20 mM pH 7.4). They were subsequently deyolked and dechorionated and lysed in RIPA buffer (Sigma) (volume equals twice the number of embryos in µL) with Halt protease inhibitor cocktail EDTA-free (Thermo Scientific) added. Lysates were homogenized by sonication and protein concentration was determined using the BCA protein assay kit (Thermo Fisher Cat #23252). Equal amounts of protein were loaded on 12% Tris-glycine PAGE gels and transferred onto a PVDF membrane using an iBlot dry blotting system (Thermo Fisher) at 20 V for 6 mins. Membranes were washed with methanol followed by water and then blocked with 10% fat free milk in TBST for 1 hr at room temperature and washed three times with TBST buffer. Membranes were incubated with primary antibodies – SOD100 (Enzo Life sciences) at 1:500 dilution in 5% milk in TBST and incubated for 1hr at room temperature or anti β-actin (Abcam) at 1:1000 diluted in 5% milk in TBST and incubated overnight at 4°C. Membranes were first probed with SOD100 and then stripped using Restore PLUS Western Blot stripping buffer (Thermo Scientific) and subsequently probed with anti β-actin. Blots were washed thrice with 1X TBST and subsequently incubated with the secondary anti-rabbit HRP secondary antibody (Sigma) at a 1:10000 dilution in 5% milk in TBST for 1hr at room temperature. Pierce ECL Western Blotting Substrate, (Thermo Scientific, #32209), a chemiluminescent substrate was used for the detection of horseradish peroxidase (HRP) on the immunoblots following the manufacturer’s instructions.
**Statistical Significance**
All statistical analyses were done using GraphPad Prism 7.0.

**Antibodies used**
The following antibodies were used in this study. SOD100 (EnzoLife sciences) at 1:1000 dilution, A11 (EMD Millipore) at 1:500 dilution, anti β-actin (Abcam) at 1:1000, anti-mouse HRP secondary antibody (Life technologies) 1:10000 dilution, anti-rabbit HRP secondary antibody (Life technologies) 1:10000 dilution.
Supporting Text

**Prediction of oligomer forming segments in SOD1**

Several reports have suggested that oligomers formed by amyloid proteins have a pore or barrel shape and are composed of antiparallel β-strands (25, 26). To investigate segments in SOD1 that may be critical for oligomer formation, we applied a combination of 3D-profiling and manual sequence fitting using a model of antiparallel β-sheet structure that we had previously discovered, termed cylindrin (12) (Fig. S1). In order to predict which 11-residue segments of SOD1 might form a cylindrin-like structure, we analyzed the cylindrin structure by visual inspection and computational modeling. We examined the structural features of the native cylindrin, and compared these with cylindrin models of segments from other proteins. Our analysis followed three stages:

1. **Computational prediction of cylindrin-compatible segments by 3D profiling**
   Using a process analogous to our 3D profile method for predicting steric-zipper-forming segments (27), we used the Rosetta modeling suite (28) for high-throughput computational modeling and prediction of cylindrin-forming segments (Fig. S1A). Starting from the 6-stranded cylindrin barrel structure (from PDB 3SGO), we stripped off all side-chain atoms, then threaded on each 11-residue sequence segment of several amyloid proteins (The same 11-residue sequence segment was threaded onto each of the six strands of the cylindrin.). Side chains matching the amino acid sequence of the segment being threaded were attached to the backbone and repacked; this selects an energetically favorable rotamer at each position. The backbone was then randomly perturbed by small sub-Angstrom moves, followed by repacking for up to 1000 iterations, or until energetic convergence. We then used the final, minimized Rosetta energy scores to rank the threaded protein segments—those with the lowest energy show the best compatibility with the cylindrin structure. The default Rosetta energy function (Talaris2013) was used.

2. **Analysis of computational predictions and visual comparison with the native cylindrin**
   Among the lowest-energy computational predictions, one sequence feature immediately stood out – glycine at position 6. A closer look at the native cylindrin structure revealed why: glycine’s lack of side chain provides room for the adjacent, inward-facing valine side chains (at position 4) to pack. As shown in Figure S1C, unrolling the cylindrin barrel to view the interior surface reveals a series of ‘knobs’ (valine 4 side chains, cyan) that can pack into ‘holes’ (green) provided by glycine. Without these glycine ‘holes’, the interior of the barrel would be too full, forcing the β-strands apart. A swap in the positions of knobs and holes (that is, moving glycine to position 4, and valine to position 6) might also be expected to allow the barrel to close. However, the computational modeling never gave low scores to segments with swapped knobs and holes. The reason we found is that the cylindrin backbone torsion angles at position 6 favor glycine.

Our analysis of the cylindrin’s interior packing also highlighted the importance of the other inward-facing side chains, which pack in symmetric layers (Fig. S1B). From
the center, and moving outward, these are valine 4, valine 8, and valine 2. The valines form a hydrophobic core, which together with hydrophilic side chains on the exterior likely promote assembly of the cylindrin. Based on these observations, our expectation was that good candidates for cylindrin formation would have three features: (1) a glycine at either position 4 or 6; (2) segregated hydrophobic and hydrophilic residues, in inward- and outward-facing positions, respectively; and (3) appropriately sized inward-facing residues, to optimize packing in the narrow interior of the barrel.

We therefore went on to analyze the predicted structures of the lowest-energy models, to see whether our expectations matched the computational predictions. As mentioned above, all of the lowest-energy segments had glycine at position 6. Regarding segregation of hydrophobic and hydrophilic residues, there was no clear pattern. Some low-energy segments were nearly completely hydrophobic. Others had polar, or even charged, residues positioned in the interior. Some of these polar residues, such as asparagine, were positioned to largely fulfill potential hydrogen bonds. Others, such as serine and histidine, were generally not able to hydrogen bond with adjacent side chains, suggesting their segments are unlikely to form a cylindrin. We concluded that while Rosetta seemed well-able to handle steric considerations when packing side chains, it had more difficulty handling charge and hydrogen-bonding considerations.

The sizes and shapes of interior side chains in low-energy models fit better with our expectations. Valine, and similarly sized side chains like threonine and isoleucine, packed quite well (Fig. S1D). Slightly larger side chains of methionine and histidine also fit (Fig. S1E, S1F). Even larger side chains, such as phenylalanine, tyrosine, and tryptophan, were absent in top-scoring segments. Side chains smaller than valine, such as alanine, serine, and even glycine, also came out in top-scoring segments (Fig. S1G, S1H). Although these did not give rise to steric clashes, they did leave visible gaps in the interior, which we hypothesized would be less favorable than if the holes were filled (similar to the original cylindrin).


Although the 3D profile method provided a number of reasonable predictions, there were enough deviations from our expectations that we also performed a manual search for cylindrin-compatible segments (Fig. S1I). We began by casting a wide net, using a sequence-based search to pull out any segments that matched either a basic pattern found in the cylindrin core sequence or in variations that we predicted were possible (for example, swapped ‘knob’ and ‘hole’ positions). Adjacent to the glycine ‘hole’, we restricted possible ‘knob’ residues to five: valine, isoleucine, leucine, methionine, and threonine. Sequences matching at least one of these basic patterns were then screened for clearly unfavorable features at other positions, and removed accordingly.

From an initial group of 143, 11-residue segments spanning the full sequence of SOD1, our screen yielded 4 overlapping sequences from the N-terminal region: segments 28-38, 30-40, 32-42, and 34-44. One of these, 32-42, was also predicted by our computational threading. A re-examination of the 28-38 segment showed that the small backbone movements allowed in our computational threading could not relieve
a clash between the proline at position 1 and the position 11 leucine of the neighboring strand, without breaking apart the structure.

As mentioned above inward facing glycines are critical for the hydrophobic packing. Among the 4 segments, 28-38 was the only segment with glycine at position 6 that is conserved in ALS patients and contains a central stretch of amino acids that are conserved in ALS patients (Fig. S2). Interestingly, the segment has also been found to be prone to local unfolding by H/D exchange and mass spectrometry studies and small molecule inhibitors of aggregation are found to bind this segment (29–31). Keeping these reports in mind, the segment 28-38 was chosen for further characterization.
Mass assignment by ion-mobility mass spectrometry (IM-MS)

A detailed experimental workflow for IM-MS experiments has been described elsewhere (32, 33). In this experiment, the ions were generated by n-ESI, stored in an ion-funnel, and subsequently pulsed into a drift cell filled with helium gas at $P = 13$ torr. The ions move through the drift cell toward the quadrupole mass analyzer under the influence of both a weak electrical field and a drag force caused by collisions with buffer gas. For the same oligomer with different conformations, those that are more extended will experience more collisions with the buffer gas, thus travel slower than those that are more compact. For oligomers with different sizes, species with higher charges are affected more by the electrical field (‘drift’ voltage) and travel faster than species with smaller charges. For example, a dimer with $z = +2$ charge will travel faster than a monomer with $z = +1$ charge, and the ATD (arrival time distribution) feature corresponding to the dimer will have a shorter arrival time than that of the monomer (34). By measuring the arrival times at different pressure to drift-voltage ratios, experimental collision cross sections can be measured for all features in the ATDs. The values can be compared to the cross sections obtained from theoretical modeling and x-ray crystallography.

The mass spectra of the cylindrin and corkscrew forming peptides are shown in Figure S6, panels A and B. Despite the difference in charge states in solution, major peaks with $n/z = +1/2, +2/3, +1/1$ and $+5/4$ were observed in the mass spectra for both peptides. SOD1(28-38) has additional peaks of $n/z = +1/3$ and $+3/4$. In addition, since SOD1(28-38) is highly charged, salt-adduct peaks are also observed, which are not observed in the mass spectra of the cylindrin segment.

We focused our analysis on the $n/z = 1/1$ mass spectral peaks of both segments since they contain multiple features that may correspond to large oligomers. First, we analyzed the ATD of the cylindrin segment, which has been shown to form hexamers by x-ray crystallography. The ATD shows multiple features (Fig. S6A), and we can unambiguously identify the monomer ($\sigma = 268$ Å$^2$) and dimer ($\sigma = 440$ Å$^2$) at 108.95 and 90.40 ms, respectively. The $z = +1$ monomer observed at this ATD has a similar cross section to $z = +2$ monomer at 599 m/z ($\sigma = 263 – 274$ Å$^2$, data not shown). Similarly, the cross section of $z = +2$ dimer is similar to $z = +3$ dimer at 799 m/z ($\sigma = 442 – 452$ Å$^2$, data not shown). The next peak at 85.5 ms is assigned a trimer with $\sigma = 625$ Å$^2$, as it does not fit a dimer. The spacing between the 75.75 ms peak and its adjacent peak at 85.5 ms (10.2 ms) is much larger than the spacing between dimer and trimer (4.85 ms). Thus, it cannot be classified as a tetramer and fits a hexamer with a cross section of 1122 Å$^2$. To resolve the size of the peak at 69.65 ms, we compared the mass spectra at different time points. The ATD of cylindrin after 24 hours shows a rearrangement in the peaks (Fig. S6D). The dimer and trimer peaks are essentially gone, while the hexamer peak at 75.75 ms is relatively unchanged. The peak at 69.65 ms increased in intensity, and no further peaks at shorter times have appeared. Hence the 69.65 ms peak is most likely the terminal stable soluble oligomer. The best assignment for the 69.95 ms peak is a compact hexamer with a cross section of 1020 – 1023 Å$^2$. If we use the X-ray coordinates of the cylindrin, we obtain a cross section of $\sigma_{av} = 1029$ Å$^2$ ($\sigma_{av}$ is the average cross section obtained from the trajectory and projected superposition approximation methods) (33, 35, 36), in excellent agreement with this assignment. Thus, it appears there are two stable hexamers of cylindrin, one open and one compact.
Next, we analyzed the mass spectra of SOD1(28-38). In this case, the n/z = 1/1 mass spectral peak initially appears to have a single, broad feature as its ATD (Fig S6E). After 24 hours, a more intense ATD is observed and now two features are resolved (Fig. S6F). Given our results for the cylindrin segment discussed above, the arrival times of these two features correspond to open (1189 Å$^2$) and compact (1113 Å$^2$) hexamers. The X-ray coordinates of SOD1(28-38) yield a cross section of $\sigma_{av} = 1182$ Å$^2$ that increases to 1226 Å$^2$ after 20-ns of MD relaxation, in good agreement with experiment. Similar to the cylindrin segment, the terminal species appears to be a hexamer for SOD1(28-38). In addition, the compact feature in the ATD is broad suggesting numerous “cylindrin-like” structures. If we thread the sequence of SOD1(28-38) on the backbone of the cylindrin structure, we obtain a cross section of 1107 Å$^2$, which is near the $\sigma_{av} = 1113$ Å$^2$ experimental cross section calculated above. This calculation supports our assignment of the broad peak centered near 74 ms as a compact, cylindrin-like hexamer. Thus, in solution, SOD1(28-38) segment forms oligomers similar in cross section to the observed crystal structure of the corkscrew and a more closed, cylindrin-like structure.
References:


Supplemental figures:

A. Computational prediction of cyldrin-compatible segments by 3D profiling

Visual analysis of the native cyldrin structure

Analysis of computational cyldrin predictions

I. Manual selection of cyldrin-compatible segments of SOD1

1. Define cyldrin-compatible sequence patterns

2. Search SOD1 for sequence patterns

3. Discard segments with problematic residues

4. Target remaining segments

(28-38): PVKWSERSKLGL
(30-40): KVWSERSKLGLTE
(32-42): WGSERSKLGLTE
(34-44): SIKGTEGLHTG

Pattern type:
- Standard cyldrin
- Shifted cyldrin
- Swapped knob & hole

Residue position:
- 1: D, E, K, R
- 2: D, E, K, R
- 3: D, E, K, R

Amino acid pattern:
- V, G
- I, L
- M, T

Discard segments with problematic residues:
- Proline at positions 2-10
- Charged, bulky, or small residues in core
- 5+ glycines
- Cysteine

Target remaining segments:
- (28-38): PVKWSERSKLGL
- (30-40): KVWSERSKLGLTE
- (32-42): WGSERSKLGLTE
- (34-44): SIKGTEGLHTG
Fig. S1. Prediction of cylindrin-compatible segments.

(A) 3D profile method for predicting cylindrin-compatible segments in amyloid proteins. An example of threading with the amyloid-β peptide is shown.

(B,C) Visual analysis of the cylindrin structure revealed a layering of side chains in the core (b; backbone trace with even-numbered, ‘inward-facing’ side chains shown as sticks, and glycine Cα shown as a small sphere), and a 'knob in hole' packing of valine and glycine residues (c; surface rendering) that allows for the small circumference of the barrel.

(D-H) Structures resulting from the 3D profile predictions were analyzed for packing in the barrel core. Representation is similar to panel B, but side-chain atoms are shown as spheres with van der Waals radii. The protein of origin and segment sequence are given, with even-numbered, ‘inward-facing’ residues in bold, and the critical position 4 residue underlined. (D) The original cylindrin structure, after Rosetta-based side-chain repacking and energy minimization. Notice the packing of valine 4, cyan, near the center. Although there is a gap between the two layers (arrowhead), it is small. (E) In this model of a segment from the human prion protein, methionine (cyan) occupies the central position 4, filling the space more completely than valine of the cylindrin sequence. (F) A histidine side chain at position 4 can also fill the central space, but it is not able to satisfy hydrogen bonding groups. (G) Alanine at position 4 does not fill the core as well, leaving gaps (arrowheads). (H) Adjacent alanine and glycine, at positions 2 and 8 respectively, leave large gaps in outer layers of the cylindrin model core (arrowheads), in spite of a native valine packing in the middle at position 4.

(I) Manual selection of cylindrin-compatible segments. First, three sequence-based criteria, which define the glycine position and the position of an adjacent, appropriate side chain (with cylindrin-compatible size and charge) in the core, were used in the initial selection of segments. The resulting sequences were further culled based on the presence of undesirable amino acids in specific positions. Prolines were avoided because of their tendency to disrupt β-structure. Charged and bulky inward-facing residues were avoided because of likely disruption to the packing. Small inward-facing residues were avoided because of gaps in packing, e.g. as in panel H. Large numbers of glycines were avoided because of the tendency for these sequences to be flexible rather than adopt a fixed conformation. Cysteines were avoided for experimental purposes. Four overlapping segments of SOD1 resulted from this manual selection.
Fig. S2. Frequency distribution of ALS-related mutations in SOD1. (A) Histogram representation of the number of familial-disease mutations found at each residue position. Gray connected line refers to the total number of patients known to carry mutations at that residue. The segment 28-38 is expanded. Val29 and Val31 each have one known familial mutation and Gly37 and Leu38 have two known familial mutations. Notice that the segment 28-38 has low frequency of individuals with ALS linked mutations with none in residues 32-36. A total of 4 individuals are known to carry mutations in this region. Data were obtained from http://alsod.iop.kcl.ac.uk (B) Ribbon diagram of dimeric SOD1 (PDB ID: 2C9S) with the segment 28-38 highlighted in blue. Zinc and copper atoms are shown as magenta spheres, and the familial mutation residues, A4V and G93A are colored red.
Fig. S3. Hydrophobicity mapped on the surface of the corkscrew crystal structure (left), and a lengthwise cross-section (right). The cleft interior is strongly hydrophobic (red) made up of valine and isoleucine side chains of Val29, Val31, and Ile35, whereas the exterior is strongly hydrophilic (blue) arising from Lys28, Lys30, Ser34, and Lys36.
Fig. S4. Lysine substitution at position 28 contributes weakly to the crystal packing and not to the stability of the corkscrew structure. (A) Crystal packing of the corkscrew structure. Two asymmetric units forming crystal contacts are shown in blue and yellow from neighboring corkscrews and N-terminal lysines are shown in sphere representation. Among the eight N-terminal lysines in each asymmetric unit, 5 form weak van der Waals interactions with a lysine residue and a tryptophan residue of the neighboring corkscrew. These interactions contribute weakly (~9%) of the total buried surface area in the crystal contacts. Buried surface area in the crystal contacts was calculated by AREAIMOL. (B) Lysine at position 28 is not essential for corkscrew
formation and replacing it with the native proline residue does not disrupt the potential for hydrogen bonding. The N-terminal lysine forms a hydrogen bond with the carbonyl oxygen of Ser34 in an adjacent β-strand (left). Substituting the N-terminal lysine with the native proline residue (gray) maintains the hydrogen bond (right). (C) The blue and red curves correspond to Cα root-mean-square deviations (RMSD) from the corkscrew crystal structure in MD simulations of the 8-chain wild-type (PVKVWGSIKGL) and G33W mutant (PVKVWWSIKGL) corkscrews, respectively. The G33W mutant deviates from the corkscrew structure, whereas the wild-type structure remains stable throughout the length of the simulation. The simulations suggest that the dynamics of the backbone at residue 33 may contribute to the instability of the G33W mutant. The G33 residue in the corkscrew structure occupies a region of Ramachandran space different from the allowable region for tryptophan and other non-glycine amino acids. In the wild-type simulations, the backbone dihedral angles of G33 tend to remain in the lower right quadrant of a Ramachandran plot, whereas those of W33 in the mutant quickly relax to the β-sheet/α-helix regions, leading to deviations from the corkscrew structure. The extra hydrophobicity of G33W may also have an effect, making the mutant corkscrew more likely to collapse.
Fig. S5. **Structural comparison of corkscrew and cylindrin.** (A) Hydrogen bonding network of the corkscrew structure (left). Shown here the orange β-strand has a strong interface with the β-strand below composed of 9 hydrogen bonds and a weak interface with the β-strand above composed of 7 hydrogen bonds. Hydrogen bonding network in the cylindrin structure (Right, PDB ID: 3SGO) shares a similar architecture where the orange strand has a stronger interface with the β-strand below composed of 12 hydrogen bonds.
bonds and a weak interface with the β-strand above composed of 8 hydrogen bonds. Br2 refers to the non-natural amino acid 2-bromoallyl-glycine that was used for phasing. (B) MD simulations of weakly restrained monomers of SOD1 spontaneously assembled into a corkscrew-like structure. A snapshot of an assembled corkscrew-like structure from the MD simulations (red) is overlaid onto the crystal structure (blue). Out of 20 simulations, three successfully assembled to a corkscrew-like structure. Interestingly, in trajectories where a corkscrew did not form, individual monomers often formed interfaces composed of out-of-register β-sheets. As a control, we found that monomers of αB crystallin spontaneously assembled into a cylindrin structure using the same simulation protocol. A snapshot of an assembled cylindrin from the MD simulation (red) is overlaid onto the crystal structure (blue, PDB ID 3SGR). Out of 10 simulations, one successful assembly to cylindrin was obtained. As in the SOD1 assembly runs, in αB crystallin assembly trajectories where a cylindrin did not form, we often observed dimers composed of out-of-register β-sheet interfaces.
Fig. S6. The corkscrew-forming segment, SOD1(28-38) forms oligomers in solution with cross section similar to the crystal structure. (A, B) n-ESI-q mass spectrum of 200 µM cylindrin and corkscrew peptides in 20 mM ammonium acetate buffer (pH = 7.0), respectively. Each mass spectral peak is annotated with n/z where n is the oligomer number and z is the charge. (C-F) Representative ATDs of n/z = 1/1 peaks of the cylindrin and corkscrew peptides incubated for t = 0 and 24 hours. Each ATD feature is annotated with n/z and an experimental collision cross section (M = monomer, D = dimer, Tr = trimer, Hex = hexamer). The narrow dashed lines are the peak shapes predicted for a single conformer of the cross sections given.
Fig. S7. Segment (28-38) with native proline is toxic and substitution at Gly33 renders it non-toxic. (A) Peptides were prepared identically by solubilizing in 50 mM tris-base buffer and overnight incubation at 37 °C with agitation and ES-derived motor neurons were treated with 100 µM final concentration. P28(28-38) is toxic while the peptides with substitutions G33V and G33W are non-toxic. Symbols represent individual values and bars represent average values. Statistical significance was measured by ANOVA. (**p < 0.01, ***p < 0.001). (B) Electron micrographs of all three peptides showed similar fibrils suggesting that substitution at Gly33 does not change the fibrillation propensity.
Fig. S8. WT and mutant SOD1 show similar kinetics of aggregation and cytotoxicity. (A) Thioflavin T assay of 80 µM WT protein shows a lag time of 2-4 hours.
before fibrils first appear (left). Cell viability (right) of ES-derived primary motor neurons incubated with protein aggregated for indicated times shows WT is only toxic when shaken up to 12 hours, when abundant fibrils first appear. (B) Similar to WT, the mutant G93A shows a lag time of 2-4 hours in a thioflavin T assay (left). Cell viability (right) of ES-derived primary motor neurons incubated with protein aggregated for indicated times shows G93A is also toxic when shaken up to 12 hours, when abundant fibrils first appear. (C) Thioflavin T assay (left) shows G33W, the corkscrew-disrupted mutant protein behaves similar to WT and G93A with a lag phase of 2-4 hours. However, it is non-toxic at all time points (right). Results shown as Mean±SD (n=3). Statistical significance was analyzed using a two-tailed T-test with Welch’s correction (*p < 0.05, ***p < 0.001, ns not significant). (D) Electron micrographs of SOD1 constructs aggregated for different lengths of time. Samples aggregated up to 12 hours tend to show some aggregates and sparse fibrils while samples aggregated for 16 hours or more show a large fibril load. Scale bar 200 nm.
Fig. S9. Expression of A4V and G93A mutant proteins causes axon shortening and abnormal mitochondrial clustering in zebrafish model. (A) Western blotting of zebrafish embryos to confirm expression of the mRNA injected. SOD100 was used to detect SOD1 and β-actin was used as loading control. UI refers to uninjected zebrafish. (B) Construct G93A/G33V/I18G is non-toxic in cell culture model similar to the G93A/G33V mutant. (C) Zebrafish at 2 dpf were imaged and higher percentage of A4V (30%) and G93A (10%) injected zebrafish displayed significant deformation compared to WT and the G33V expressing fish. (D) Average axon length (normalized to WT construct) and mitochondrial size of the different constructs. n refers to the total number of measurements made for each construct.
Fig. S10. Structural comparison of corkscrew with other β-sheet proteins. The DALI server was used to search the Protein Data Bank for proteins with a similar structural fold. Three main categories of proteins were found – big barrel proteins, small β-sheet proteins that bind ligands and open twisted β-sheet proteins that bind ligands. (A) Overlay of corkscrew (purple) with the structure of BamA (PDB: 4K3C), a β-barrel membrane protein (gray). (B) Overlay of corkscrew (purple) with the structure of AbyU (PDB: 5DYV), a diels alderase (gray) with HEPES (red) bound in the barrel. (C) Overlay of corkscrew (purple) with BPI (PDB: 1BP1), a bactericidal protein with phosphocholine bound. Notice that the twisted sheet creates a groove for binding phosphocholine (red). RMSD – root mean square deviation of Ca, LALI – length of aligned residues.
Fig. S11. Model of SOD1 aggregation pathway. SOD1 forms a stable dimer that is resistant to aggregation. Its destabilization into monomers renders it prone to aggregation into β-sheet rich fibrils. The corkscrew structure of segment 28-38 (purple) determined here, suggests that there is an off-pathway for oligomer formation independent of the in-register fibrils. The mutations, G33W/G33V specifically prevent corkscrew formation, an out-of-register oligomer. Fibril formation assays suggest that G33W/G33V do not affect the in-register fibril formation pathway and large fibrils are non-toxic.
<table>
<thead>
<tr>
<th></th>
<th>KVKVWGSIKGL</th>
<th>KVKVWGSIKGL (iodide)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beamline</strong></td>
<td>APS 24-ID-E</td>
<td>APS 24-ID-E</td>
</tr>
<tr>
<td><strong>Resolution Å</strong></td>
<td>1.9</td>
<td>2.1</td>
</tr>
<tr>
<td><strong>Total unique reflections</strong></td>
<td>8378</td>
<td>11489</td>
</tr>
<tr>
<td><strong>Total reflections observed</strong></td>
<td>47323</td>
<td>43500</td>
</tr>
<tr>
<td><strong>Unit cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>33.2, 44.4, 71.2</td>
<td>33.1, 44.4, 71.4</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.0, 90.0, 90.0</td>
<td>90.0, 90.0, 90.0</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P2;2;2;1</td>
<td>P2;2;2;1</td>
</tr>
<tr>
<td><strong>R_{merge}</strong></td>
<td>16.5% (48.4%)</td>
<td>9.5% (41.8%)</td>
</tr>
<tr>
<td><strong>I/σ</strong></td>
<td>8.4 (4.3)</td>
<td>11.1 (3.6)</td>
</tr>
<tr>
<td><strong>Completeness</strong></td>
<td>95.5 %</td>
<td>96.7 %</td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>0.9791</td>
<td>0.9791</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>35.5-2.0</td>
<td>19.3-2.0</td>
</tr>
<tr>
<td><strong>Reflections for refinement</strong></td>
<td>6690</td>
<td>6249</td>
</tr>
<tr>
<td>(after merging Friedel Pairs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R_{free}/R_{work} (%)</strong></td>
<td>25.08/21.09</td>
<td>26.2/21.2</td>
</tr>
<tr>
<td><strong>Molecules per asymmetric unit</strong></td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><strong>Solvent content (%)</strong></td>
<td>54.6</td>
<td>54.6</td>
</tr>
<tr>
<td><strong>Matthews coefficient (Å³/Da)</strong></td>
<td>2.71</td>
<td>2.71</td>
</tr>
<tr>
<td><strong>Total water molecules</strong></td>
<td>34</td>
<td>52</td>
</tr>
<tr>
<td><strong>Total iodide atoms</strong></td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Glycerol molecule</strong></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Rmsd bond length (Å)</strong></td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td><strong>Rmsd angles (°)</strong></td>
<td>1.022</td>
<td>0.88</td>
</tr>
<tr>
<td><strong>Ramachandran plot</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Allowed</strong></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Generous</strong></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Disallowed</strong></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table S2. Comparison of shape complementarity (Sc) and buried surface area (A_b) of corkscrew with cylindrin and steric zippers.

<table>
<thead>
<tr>
<th></th>
<th>Corkscrew</th>
<th>Cylindrin*</th>
<th>Steric Zipper**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sc</td>
<td>0.79</td>
<td>0.74</td>
<td>0.68</td>
</tr>
<tr>
<td>A_b (Å²)</td>
<td>984</td>
<td>943</td>
<td>1034</td>
</tr>
<tr>
<td>A_b/Residue</td>
<td>89</td>
<td>86</td>
<td>94</td>
</tr>
</tbody>
</table>

* PDB ID: 3SGO  
** PDB ID: 4RIL

For the Sc calculation, we examined the interface between one chain and the remaining chains of the assembly. A_b values were calculated using AREAIMOL. The area buried was calculated by subtracting the solvent accessible surface area of one chain of the assembly from the total solvent accessible surface area of an isolated chain.