Molecular basis of small-molecule binding to α-synuclein

Supporting Information

Paul Robustelli,1,2 Alain Ibanez-de-Opakua,3 Cecily Campbell-Bezat,1 Fabrizio Giordanetto,1 Stefan Becker,4 Markus Zweckstetter,3,4,5,* Albert C. Pan,1,* and David E. Shaw1,6,*

1 D. E. Shaw Research, New York, NY 10036, USA.

2 Department of Chemistry, Dartmouth College, Hanover, NH 03755, USA.

3 German Center for Neurodegenerative Diseases (DZNE), 37077 Göttingen, Germany.

4 Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany.

5 DFG Research Center for Nanoscale Microscopy and Molecular Physiology of the Brain (CNMPB), 37073 Göttingen, Germany.

6 Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA.

* To whom correspondence should be addressed.

Albert C. Pan
E-mail: Albert.Pan@DEShawResearch.com
Phone: (212) 403-8664
Fax: (646) 873-2664
Markus Zweckstetter
E-mail: Markus.Zweckstetter@dzne.de
Phone: +49 551 39 61181

David E. Shaw
E-mail: David.Shaw@DEShawResearch.com
Phone: (212) 478-0260
Fax: (212) 845-1286
Simulation Methods

Systems were equilibrated with GPU/Desmond. Production runs at 300 K were performed in the NPT ensemble with Anton, a special-purpose machine capable of running very long MD simulations. To further increase computational efficiency, we modified all the hydrogen (H) and water oxygen (O$_w$) masses (H: 4 Da; O$_w$: 10 Da); this allowed us to integrate the equations of motion using a RESPA scheme with an inner time step of 4.5 fs and an outer time step of 9 fs. Previous simulations of a well-characterized protein, villin, showed that using a larger time step and altering the masses of hydrogen and water oxygens did not have any substantial effect on the protein’s kinetics or the thermodynamics. Bonds involving hydrogen atoms were restrained to their equilibrium lengths using the M-SHAKE algorithm. Nonbonded interactions were truncated at 10 Å, and long-range electrostatic interactions were computed using the u-series approach. Proteins, water molecules, and ions were parameterized with the a99SB-*disp force field. All ligands were parameterized using the generalized Amber force field (GAFF) and the AM1-BCC charge model. Simulations of full-length α-syn were performed in a cubic water box with 108-Å sides. The initial protein structure was prepared as described in previous work. Na or Cl ions were added to a concentration of 25 mM for Ligands 23 and 47, and to a concentration of 50 mM for fasudil. Simulations of α-syn-C-term were performed in a cubic water box with ~42-Å sides (corresponding to protein and small molecule concentrations of ~0.020 M), and simulations of α-syn 29–49 were performed in a cubic water box with ~53-Å sides (corresponding to protein and small molecule concentrations of ~0.011 M). Na or Cl ions were added to each system to achieve charge neutrality. The α-syn 29–49 fragment was capped with ACE and NME groups at the N terminus and C terminus, respectively.

To calculate the dissociation constant ($K_D$) we used the expression $11$

$$K_D = \frac{P_u}{P_b} (\nu N_{Av})^{-1},$$
where $P_u$ is the fraction of simulation time in which the protein and ligand are unbound, $P_b$ is the fraction of simulation time in which the protein and ligand are bound, $v$ is the volume of the simulation box, and $N_A$ is Avogadro’s number. To partition trajectories into bound and unbound states, we calculated the closest distance between the protein atoms and the ligand heavy atoms, and simulation frames in which this distance was less than 6 Å were considered bound.

We define aromatic stacking by considering the mutual arrangement of two aromatic groups as defined by a vector $\mathbf{R}$ connecting the centroids of the aromatic groups, and the angles $\alpha$ and $\beta$, defined by the normal of each aromatic plane and the vector $\mathbf{R}$ (Figure 2C). A conformation was classified as aromatic stacking if the length of $\mathbf{R}$ was <5 Å, $\alpha$ was >135° or <45°, and $\beta$ was>135° or <45°. To ensure consistency of the direction of the normal vector assigned to each aromatic plane, we utilized three reference atoms in each aromatic ring, which we call A, B, and C. We defined two reference vectors using the coordinates of these reference atoms (reference_vector1 = $[x,y,z]_A - [x,y,z]_B$, reference_vector2 = $[x,y,z]_A - [x,y,z]_C$). We calculated a reference direction vector by taking the cross product of reference_vector1 and reference_vector2. The normal of each aromatic plane was assigned such that it points in the same direction as the reference vector direction for each ring, ensuring that the same direction conventions and corresponding stacking-quadrant definitions were used in all frames.

Potential hydrogen bond donors were defined as all nitrogen, oxygen, or sulfur atoms with an attached hydrogen. Hydrogen bonds were identified with a distance cutoff of 3.5 Å between the donor hydrogen and heavy-atom acceptor, and a donor-hydrogen-acceptor angle >150°. Charge contacts were defined as being when any two atoms that contain opposite formal charges were within 5 Å. Hydrophobic contacts were defined as being when carbon atoms were within 4 Å (excluding backbone Cα atoms).
Mutual information (MI) was calculated using binary contact probabilities between all intermolecular interactions according to Shannon and Weaver\textsuperscript{13}

\[
I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p_{(X,Y)}(x, y) \log\left( \frac{p_{(X,Y)}(x,y)}{p_X(x) p_Y(y)} \right),
\]

where \(p_{(X,Y)}\) is the joint probability mass function of \(X\) and \(Y\); where \(p_X\) and \(p_Y\) are the marginal probability mass functions of \(X\) and \(Y\) respectively; and where \(X\) and \(Y\) represent binary contact probabilities of intermolecular interactions. Because the MI is strictly positive and bounded from above by the entropy of the joint distribution, the MI between low-entropy contact pairs will necessarily appear to be low even if they are perfectly coupled. To better resolve these contact pairs, we normalized the MI to the joint entropy of the contact pair,\textsuperscript{14} which bounds the MI between 0 and 1.

### Error Estimation from Blocking Analyses

Error bars for simulated properties used a blocking analysis following Flyvbjerg and Petersen.\textsuperscript{15} In this procedure, the trajectory is divided into a given number of equally sized “blocks,” an average value is computed for a simulated property in each block, and the standard error is calculated of the average value across all blocks. Final error estimates for a given quantity can be obtained by calculating the standard errors of block averages with increasingly large block sizes until the standard error estimate converges within a noise threshold, or one can attempt to select an optimal block size in order to minimize the estimated error of the calculated variance.\textsuperscript{16} In this work, to reduce the chance of underestimating error bars and provide an estimate that is more likely to over- rather than underestimate the error, we report the largest standard error observed among all block sizes for each simulated quantity. That is, we have computed the error for equally sized blocks of all possible sizes, and, rather than selecting a block size that
minimizes the estimated error of the calculated variance\textsuperscript{16}, we report the largest blocking error obtained among all block sizes tested as our error estimate.

**Assessment of Simulation Convergence and Accuracy**

In order to assess the convergence of our MD simulations, we examined the autocorrelation functions of several simulated properties. As a general test of the sampling of the conformational space of simulations of full length α-synuclein and the α-syn-C-term fragment in the presence of fasudil, we examined the time-course and autocorrelation functions of the distance between the Cα atoms of the first and last residue of each construct (Cα end-to-end distance) and the RMSD of each simulation frame to the first frame of the simulation (Cα RMSD). As shown in Figure S13, for the 1500-µs simulation of full-length α-synuclein, we observe that the autocorrelation time of the Cα end-to-end distance is <1 µs, and the autocorrelation time of the Cα RMSD is <10 µs, indicating that these quantities were well sampled in simulations of this length (our simulations are >1500x and >150x the length of these autocorrelation times, respectively). In the case of the 200 µs simulation of the α-syn-C-term fragment in the presence of fasudil, we find that the autocorrelation time of the Cα end-to-end distance is <.01 µs and the autocorrelation time of the Cα RMSD (Figure S14) is <0.1 µs, and that this simulation is thus >20,000x and >2,000x the length of these autocorrelation times, respectively. These analyses suggest that the conformational space of α-synuclein and α-syn-C-term are well sampled in simulations of this length. We can also see this by calculating the secondary structure-propensities along with error bars (computed by blocking, as described above) for simulations of α-syn-C-term in the presence of fasudil and Ligand 47; see Figure S15. We find that for these 200-µs simulations, the error bars of secondary-structure propensities are relatively small, suggesting that our simulations are well converged.
We further assessed the convergence of the simulations of α-syn-C-term in its apo form and in the presence of fasudil and Ligand 47 by examining the autocorrelation of additional simulated properties. For each frame in the simulation, we calculated the cosine distance between the value of the dihedral angle of each residue and the dihedral basin centers of α-helical 

\[
((\phi_{\text{ref}},\psi_{\text{ref}})) = (-60.2, -45.3),
\]

β-sheet 

\[
((\phi_{\text{ref}},\psi_{\text{ref}})) = (135.3, -135.3),
\]

and polyproline II 

\[
((\phi_{\text{ref}},\psi_{\text{ref}})) = (-75.1, 155.4)
\]

secondary-structure elements according to:

\[
D_{\text{hedral distance}}(X) = \frac{1}{2} \sum_{i=1}^{N} \left( 1 + \cos(\phi_i - \phi_{\text{ref}}) + \cos(\psi_i - \psi_{\text{ref}}) \right)
\]

We compare the autocorrelation functions of the dihedral distance terms to the autocorrelation functions of the radius of gyration and Cα RMSD for α-syn-C-term in its apo form and in the presence of fasudil and Ligand 47 in Figure S14, and find that all of these simulated properties have autocorrelation times <~0.1 µs, indicating that these 200-µs simulations are >~2,000x the length of these autocorrelation times.

We also calculated the distribution of residence times for each binding event in the simulations of α-syn-C-term in the presence of fasudil and Ligand 47, where binding events are defined as a sequence of consecutive frames in which there is at least one contact (defined as a non-hydrogen atom-pair with a distance <6.0 Å) between the protein and the ligand (Figure S17). Based on this metric, there were 48904 binding events between fasudil and α-syn-C-term, with an average residence time of 3.02 ns, and 32673 binding events between Ligand 47 and α-syn-C-term, with an average residence time 5.09 ns. In Table S5, we also report the number of binding events with residence times above various thresholds. Finally, we examined the convergence of the $K_D$ values for simulations of α-syn-C-term in the presence of fasudil and Ligand 47 in Figure S16, showing that these quantities converge relatively quickly to their final equilibrium values.
Agreement of Simulations of α-Synuclein with Previously Reported NMR Data from Apo α-Synuclein

**Full-length α-synuclein**

A large amount of experimental data has been previously reported to characterize the conformational ensemble of apo α-synuclein. Chemical shift perturbations (CSPs) measured in the presence of fasudil and Ligand 47 were very small, with CSPs on the order ~0.01 ppm for backbone HN shifts and ~0.1 ppm for backbone N shifts (Figures 1 and 4). Based on the small magnitude of CSPs, it is unlikely that these compounds have a large effect on the backbone conformations of α-synuclein, and thus that they would have a significant effect on the values of RDCs, PREs, and scalar couplings. As an assessment of the accuracy of our full-length α-synuclein simulations in the presence of fasudil and Ligand 47, we thus compared the simulated values of NMR chemical shifts, RDCs, PREs and scalar couplings to previously reported NMR data\textsuperscript{20–27} for apo α-synuclein, and found excellent agreement with this data, on par with the best agreement observed among simulations performed with a large number of different force fields\textsuperscript{8,28} (Table S6; Figs. S18 and S19). Comparisons of the calculated and experimental chemical shifts for MD simulations of full-length α-synuclein in its apo form and in the presence of ligands are shown in Figure S18. Comparisons of the calculated and experimental scalar couplings and RDCs are shown in Figure S19.

**α-syn-C-term**

We compared the agreement of NMR chemical shifts calculated from simulations of α-syn-C-term in the presence and absence of small molecules to backbone chemical shifts previously measured in full-length apo α-synuclein (Table S7, Fig. S20). The agreement between calculated and experimental chemical shifts is excellent, suggesting that our simulations of α-syn-C-term
provide an accurate description of the conformational ensemble of residues 121–140 of α-synuclein in the context of the full protein. The similarity between the calculated chemical shifts also strongly suggests that the conformational ensemble of α-syn-C-term is not substantially perturbed by fasudil and Ligand 47.

**Experimental Methods**

The α-synuclein protein was expressed at 37 °C in *Escherichia coli* strain BL21(DE3) in M9 minimal medium supplemented with $^{15}$NH$_4$Cl (Cambridge Isotope Laboratories), by induction with 1 mM IPTG. The cell pellet from 1 L cell culture was resuspended in 40 mL lysis buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA, 0.5 M PMSF). Cell lysis was performed by three freeze/thaw cycles followed by eight ultrasonication cycles (20 sec at 90% power) on ice. The cell lysate was incubated at 96 °C in a water bath for 15 min. The supernatant was collected by centrifugation (Beckman Coulter, JA-25.5 rotor, 48,000 g at 4 °C for 45 min). Streptomycin sulfate was added to the supernatant at a final concentration of 10 mg ml$^{-1}$ and incubated at 4 °C for 30 min. The supernatant was collected by centrifugation (JA-25.5, 45 min, 48,000 g) and ammonium sulfate was added to a final concentration of 360 mg mL$^{-1}$ while stirring on ice for 30 min. After a final centrifugation step, the protein pellet was obtained and dialyzed against 25 mM Tris-HCl, pH 7.7 overnight. The dialysate was applied to a 30 mL POROS 20 HQ anion exchange column (Thermo Scientific) and eluted using a salt gradient from 0 to 1 M NaCl in the same buffer. In a final step the protein was further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 prep grade column (GE Healthcare) equilibrated with 50 mM HEPES, 100 mM NaCl, pH 7.4, 0.02% NaN$_3$ using an ÄKTA Purifier system (GE Healthcare). The final stock at 300 μM was filtered through sterile 0.22-μm filters and stored at −80 °C.

One-dimensional (1D) 1H NMR experiments and two-dimensional (2D) 1H-15N heteronuclear single quantum coherence (HSQC) experiments of α-synuclein in absence and presence of
different concentrations of ligands were acquired at 288 K on Bruker 700 MHz and 800 MHz spectrometers equipped with triple-resonance 5 mm cryogenic probes. The protein concentration was 30 μM and all samples were in HEPES 50 mM pH 7.4, NaCl 100 mM, D₂O 5%, NaN₃ 0.01% and DSS 50 μM. 1D experiments were used to calibrate different parameters. The signals of HEPES were always checked to correct pH changes caused by the addition of ligand powder through the titration. Signals from aromatic protons were used to measure the real concentration of each ligand. DSS methyl signal was used to calibrate the chemical shift. Spectra were processed with TopSpin 3.6 (Bruker) and analyzed using Sparky. The CSP error is based on the resolution of the spectra.

Ligands 2, 5, 23, 47, and fasudil were purchased from commercial sources and used without additional purification. The (1D) 1H NMR characterization of these ligands is shown in Figure S12.

We note that, whereas the NMR experiments were carried out at 288 K, the simulations were performed at 300 K, a temperature at which the greatest amount of reliable experimental data was available when optimizing the MD force field. We believe that possible discrepancies due to this small difference in temperature are minor: Points along protein melting curves calculated in simulation (and in experiment), for example, are often within statistical error of each other in the temperature range from 280 K to 300 K. Moreover, the predictions from the MD simulations for the NMR contact probabilities and small-molecule affinities agree well in this work despite the small difference in temperature between the simulations and experiment.
Supporting References

1. Bergdorf, M., Baxter, S., Rendelemen, C. A. & Shaw, D. E. Desmond/GPU performance as of November 2016. D. E. Shaw Research Technical Report DESRES/TR--2016-01 (2016)

2. David E. Shaw, J.P. Grossman, Joseph A. Bank, Brannon Batson, J. Adam Butts, Jack C. Chao, Martin M. Deneroff, Ron O. Dror, Amos Even, Christopher H. Fenton, Anthony Forte, Joseph Gagliardo, Genette Gill, Brian Greskamp, C. Richard Ho, Douglas J. Ierardi, Lev Iserovich, Jeffrey S. Kuskin, Richard H. Larson, Timothy Layman, Li-Siang Lee, Adam K. Lerer, Chester Li, Daniel Killebrew, Kenneth M. Mackenzie, Shark Yeuk-Hai Mok, Mark A. Moraes, Rolf Mueller, Lawrence J. Nociolo, Jon L. Peticolas, Terry Quan, Daniel Ramot, John K. Salmon, Daniele P. Scarpazza, U. Ben Schafer, Naseer Siddique, Christopher W. Snyder, Jochen Spengler, Ping Tak Peter Tang, Michael Theobald, Horia Toma, Brian Towles, Benjamin Vitale, Stanley C. Wang, and Cliff Young, "Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer," Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC14), Piscataway, NJ: IEEE, 2014, pp. 41–53.

3. Tuckerman, M., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. J. Chem. Phys. 97(3), 1990–2001 (1992)

4. Feenstra, K. A., Hess, B. & Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. J. Comput. Chem. 20(8), 786–798 (1999)

5. Piana, S., Lindorff-Larsen, K. & Shaw, D. E. Atomic-level description of ubiquitin folding. Proc. Natl. Acad. Sci. U. S. A. 110(15), 5915–5920 (2013).

6. Kräutler, V., van Gunsteren, W. F. & Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. J. Comput. Chem. 22, 501–508 (2001).
7. C. Predescu, A. K. Lerer, R. A. Lippert, B. Towles, J. P. Grossman, R. M. Dirks, D. E. Shaw, The $\mu$-series: A separable decomposition for electrostatics computation with improved accuracy. J. Chem. Phys. 152, 084113 (2020).

8. Robustelli, P., Piana, S., and Shaw, D.E. Developing a molecular dynamics force field for both folded and disordered protein states. Proc Natl Acad Sci U S A. 2018 May 22;115(21):E4758-E4766.

9. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. Journal of computational chemistry. 2004 Jul 15;25(9):1157-74

10. S. Piana, A. G. Donchev, P. Robustelli, D. E. Shaw, Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. J. Phys. Chem. B. 119, 5113–5123 (2015).

11. De Jong, D.H., Schäfer, L.V., De Vries, A.H., Marrink, S.J., Berendsen, H.J.C. and Grubmüller, H. (2011), Determining equilibrium constants for dimerization reactions from molecular dynamics simulations. J. Comput. Chem., 32: 1919-1928.

12. Chelli, R., Gervasio, F.L., Procacci, P. and Schettino, V., 2002. Stacking and T-shape competition in aromatic– aromatic amino acid interactions. Journal of the American Chemical Society, 124(21), pp.6133-6143.

13. Shannon, C.E. and Weaver, W. (1949). The mathematical theory of communication. University of Illinois Press, Urbana, Illinois.

14. Malvestuto, F.M., 1986. Statistical treatment of the information content of a database. Information Systems, 11(3), pp.211-223

15. Flyvbjerg, H., Petersen, H. G. Error estimates on averages of correlated data. The Journal of Chemical Physics. 91, 461–466 (1989). (1989)

16. Wolff, U. Monte Carlo errors with less errors. Computer Physics Communications. 156, 143–153 (2004).
17. Lindorff-Larsen K, Trbovic N, Maragakis P, Piana S, Shaw DE. Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. J Am Chem Soc. 2012 134(8):3787–3791.

18. Salvi N, Abyzov A, Blackledge M. Solvent-dependent segmental dynamics in intrinsically disordered proteins. Sci Adv. 2019 5(6):eaax2348.

19. Robustelli P, Trbovic N, Friesner RA, Palmer AG 3rd. Conformational dynamics of the partially disordered yeast transcription factor GCN4. J Chem Theory Comput. 2013 9(11):10.1021/ct400654r.

20. Sung YH, Eliezer D. Residual structure, backbone dynamics, and interactions within the synuclein family. J Mol Biol. 2007 372(3):689–707.

21. Rao JN, Kim YE, Park LS, Ulmer TS (2009) Effect of pseudorepeat rearrangement on α-synuclein misfolding, vesicle binding, and micelle binding. J Mol Biol 390:516–529.

22. Bertoncini CW, Jung YS, Fernandez CO, Hoyer W, Griesinger C, Jovin TM, Zweckstetter M (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. Proc Natl Acad Sci USA 102(5):1430–1435.

23. Lee JH, Li F, Grishaev A, Bax A (2015) Quantitative residue-specific protein backbone torsion angle dynamics from concerted measurement of 3 J couplings. J Am Chem Soc 137(4):1432–1435.

24. Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM (2005) Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations. J Am Chem Soc 127(2):476–477.

25. Salmon L, Nodet G, Ozenne V, Yin G, Jensen MR, Zweckstetter M, Blackledge M (2010) NMR characterization of long-range order in intrinsically disordered proteins. J Am Chem Soc 132(24):8407–8418.
26. Esteban-Martín S, Silvestre-Ryan J, Bertoncini CW, Salvatella X (2013) Identification of fibril-like tertiary contacts in soluble monomeric α-synuclein. Biophys J 105(5):1192–1198.

27. Schwalbe M, Ozenne V, Bibow S, Jaremko M, Jaremko L, Gajda M, Jensen MR, Biernat J, Becker S, Mandelkow E, Zweckstetter M, Blackledge M (2014) Predictive atomic resolution descriptions of intrinsically disordered hTau40 and α-synuclein in solution from NMR and small angle scattering. Structure 22(2):238–249.

28. Piana S, Robustelli P, Tan D, Chen S, Shaw DE. Development of a force field for the simulation of single-chain proteins and protein-protein complexes. J Chem Theory Comput. 2020 16(4):2494–2507.

29. Lee, W., Tonelli, M. & Markley, J.L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31, 1325-1327 (2015).

30. LeVine MV, Weinstein H (2014) NbIT - A New Information Theory-Based Analysis of Allosteric Mechanisms Reveals Residues that Underlie Function in the Leucine Transporter LeuT. PLOS Computational Biology 10(5): e1003603.

31. Y. Shen, A. Bax, SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. J Biomol NMR. 48, 13–22 (2010).
## Supporting Tables

### Fasudil

| Side chain          | Quadrant | Max/Min |
|---------------------|----------|---------|
|                     | 1        | 2       | 3        | 4 |
| Y125                | 0.226 ± 0.015 | 0.259 ± 0.020 | 0.272 ± 0.024 | 0.243 ± 0.016 | 1.20 |
| Y125 with D135 contact | 0.179 ± 0.060 | 0.427 ± 0.171 | 0.245 ± 0.097 | 0.148 ± 0.055 | 2.88 |
| Y133                | 0.251 ± 0.016 | 0.229 ± 0.014 | 0.254 ± 0.015 | 0.266 ± 0.021 | 1.16 |
| Y133 with D135 contact | 0.194 ± 0.037 | 0.176 ± 0.038 | 0.171 ± 0.031 | 0.459 ± 0.173 | 2.68 |
| Y136                | 0.227 ± 0.016 | 0.245 ± 0.015 | 0.274 ± 0.019 | 0.254 ± 0.018 | 1.21 |
| Y136 with D135 contact | 0.300 ± 0.062 | 0.257 ± 0.048 | 0.229 ± 0.063 | 0.213 ± 0.050 | 1.40 |

### Ligand 47

| Side chain          | Quadrant | Max/Min |
|---------------------|----------|---------|
|                     | 1        | 2       | 3        | 4 |
| Y125                | 0.227 ± 0.018 | 0.245 ± 0.024 | 0.264 ± 0.026 | 0.264 ± 0.026 | 1.16 |
| Y125 with D135 contact | 0.205 ± 0.046 | 0.453 ± 0.200 | 0.212 ± 0.064 | 0.130 ± 0.039 | 3.48 |
| Y133                | 0.288 ± 0.026 | 0.241 ± 0.018 | 0.253 ± 0.021 | 0.218 ± 0.016 | 1.32 |
| Y133 with D135 contact | 0.252 ± 0.046 | 0.200 ± 0.030 | 0.263 ± 0.054 | 0.285 ± 0.055 | 1.42 |
| Y136                | 0.219 ± 0.016 | 0.271 ± 0.022 | 0.251 ± 0.034 | 0.259 ± 0.045 | 1.24 |
| Y136 with D135 contact | 0.103 ± 0.021 | 0.193 ± 0.050 | 0.314 ± 0.133 | 0.391 ± 0.173 | 3.81 |
Table S1. Protein-ligand aromatic stacking orientations. For all conformations where the ring centers of a ligand and a tyrosine phenol group were within 5 Å, we calculated the angles (α, β) of the normal vectors of each ring plane relative to a vector connecting the two ring centers (illustrated in Figure 2C). To quantify asymmetries in the distributions of these angles, we report the ratio of the populations of the most populated quadrant to the least populated quadrant.
Table S2. Simulated $K_D$ values of fasudil, Ligand 47, and Ligand 23 to full-length α-synuclein, and two subregions containing tyrosine residues. We note that the rank order of small-molecule binding affinities to the N-terminal region of full-length α-syn differs from that of the isolated residue 29–49 fragment system. We believe this difference results from the fact that residues 29–49 are located in the relatively hydrophobic NAC region of α-synuclein: We observe that our simulations of full-length α-synuclein sample conformations in which this region interacts with other regions of the protein, changing the protein-ligand binding landscape in ways that differ from that of the isolated 29–49 fragment.

| Ligand | $K_D$ full length (mM) | $K_D$ 29–49 (mM) | $K_D$ 121–140 (mM) |
|--------|------------------------|------------------|-------------------|
| fasudil| 2.98 +/− 0.022         | 15.9 +/− 0.15    | 6.80 +/− 0.080    |
| 47     | 2.26 +/− 0.039         | 10.3 +/− 0.18    | 5.0 +/− 0.12      |
| 23     | 6.4 +/− 0.13           | 24 +/− 1.2       | 25.5 +/− 0.98     |
| Ligand | SMILES string of compound | $K_D$ α-syn 29–49 (mM) | Blocking Error | $K_D$ α-syn-C-term (mM) | Blocking Error |
|--------|---------------------------|------------------------|----------------|------------------------|----------------|
| 1      | S(=O)(=O)(N1CCNCCC1)c1c2c(ccc1)cccc2 | 35                     | 1.6            | 8.6                    | 0.33           |
| 2      | S(=O)(=O)(N1CCNCCC1)c1cc(C(=O)NC(C)c2cccccc2)ccc1 | 31                     | 1.2            | 7.2                    | 0.28           |
| 3      | S(=O)(=O)(N1CCNCCC1)c1cc2c(cc1)cccc2            | 28                     | 1.2            | 6.1                    | 0.37           |
| 4      | Clc1c(Cl)cccc(S(=O)(=O))N2CCNCCC2)c1            | 35                     | 1.1            | 8.0                    | 0.26           |
| 5      | Clc1c(Cl)cccc1S(=O)(=O)N1CCNCCC1               | 36                     | 1.3            | 7.8                    | 0.27           |
| 6      | S(=O)(=O)(N1CCNCCC1)c1ccc(C2CCCCC2)cc1         | 31                     | 1.3            | 7.3                    | 0.27           |
| 7      | S(=O)(=O)(N1CCNCCC1)c1c2nc(C)cc2ccc1           | 39                     | 1.5            | 8.3                    | 0.28           |
| 8      | S(=O)(=O)(N1CCNCCC1)c1c2c(c(OC)cc1)cccc2       | 32                     | 1.3            | 7.0                    | 0.45           |
| 9      | S(=O)(=O)(N1CCNCCC1)c1ccc(Oc2ccc(OC)cc2)cc1    | 24                     | 1.1            | 6.0                    | 0.31           |
| 10     | Clc1cc2c(S(=O)(=O)N3CCNCCC3)csc2cc1            | 29                     | 1.3            | 6.6                    | 0.38           |
| 11     | S(=O)(=O)(N1CCNCCC1)c1c2c(c(C)c1)nccc2         | 28                     | 1.2            | 6.3                    | 0.34           |
| 12     | S(=O)(=O)(N1CCNCCC1)c1cc2sc(C(C)(C)C)nc2cc1    | 28                     | 1.2            | 5.0                    | 0.25           |
| 13     | Clc1cc2c(Scc(S(=O)(=O)N3CCNCCC3)c2)ccc1        | 25                     | 1.6            | 5.5                    | 0.21           |
| 14     | S(=O)(=O)(N1CCNCCC1)c1cc2C(=O)NCCc2cc1         | 37                     | 1.9            | 7.2                    | 0.24           |
| 15     | Clc1cc2c(sc(S(=O)(=O)N3CCNCCC3)c2)cc1          | 28                     | 1.6            | 5.5                    | 0.32           |
| 16     | S(=O)(=O)(N1CCNCCC1)c1ccc(Oc2cc(C)cc2)cc1      | 28                     | 1.3            | 6.3                    | 0.35           |
| 17     | S(=O)(=O)(N1CCNCCC1)c1cn(-c2cccccc2)nc1        | 35                     | 1.6            | 7.1                    | 0.27           |
| 18     | S(=O)(=O)(N1CCNCCC1)c1c2c(c(C)(=O)OC)ncc2ccc1  | 27                     | 1.3            | 4.5                    | 0.29           |
| 19     | S(=O)(=O)(N1CCNCCC1)c1c(C)n(C)c21cccc2         | 36                     | 1.4            | 8.8                    | 0.33           |
| 20     | S(=O)(=O)(N1CCNCCC1)c1cc(-c2ccc(C)n2)ccc1      | 29                     | 1.2            | 5.5                    | 0.30           |
| 21     | S(=O)(=O)(N1CCN(C)CC1)c1c2c(cncc2)ccc1         | 38                     | 2.0            | 8.5                    | 0.37           |
| 22     | S(=O)(=O)(N1CC(N)CC1)c1c2c(cncc2)ccc1          | 36                     | 1.6            | 7.8                    | 0.29           |
| 23     | S(=O)(=O)(N1CCC(O)CC1)c1c2c(cncc2)ccc1         | 24                     | 1.0            | 18.4                   | 0.58           |
|   | Molecular Structure                                                                 |   |   |   |
|---|-----------------------------------------------------------------------------------|---|---|---|
| 24 | \(\text{S(O)(O)(N1CCC(CN)CC1)c1c2c(cncc2)ccc1} \)                             | 29 | 1.2 | 6.6 | 0.24 |
| 25 | \(\text{S(O)(O)(N1CCC(N)CC1)c1c2c(cncc2)ccc1} \)                             | 34 | 1.5 | 7.6 | 0.38 |
| 26 | \(\text{S(O)(O)(N1CC(CN)CC1)c1c2c(cncc2)ccc1} \)                             | 30 | 1.2 | 7.0 | 0.31 |
| 27 | \(\text{S(O)(O)(N1CC(CN)CC1)c1c2c(cncc2)ccc1} \)                             | 35 | 1.4 | 6.9 | 0.30 |
| 28 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 26 | 1.0 | 18.5 | 0.72 |
| 29 | \(\text{S(O)(O)(N1CC(N)CC1)c1c2c(cncc2)ccc1} \)                             | 27 | 1.0 | 6.2 | 0.28 |
| 30 | \(\text{S(O)(O)(N1CCOC1)c1c2c(cncc2)ccc1} \)                               | 36 | 1.9 | 8.5 | 0.29 |
| 31 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 41 | 1.9 | 9.6 | 0.37 |
| 32 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 27 | 1.4 | 6.1 | 0.46 |
| 33 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 42 | 2.2 | 10.4 | 0.32 |
| 34 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 49 | 1.8 | 11.7 | 0.33 |
| 35 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 37 | 1.7 | 7.6 | 0.41 |
| 36 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 25 | 1.0 | 6.0 | 0.33 |
| 37 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 37 | 1.4 | 7.9 | 0.16 |
| 38 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 40 | 1.6 | 8.9 | 0.26 |
| 39 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 29 | 1.7 | 4.5 | 0.17 |
| 40 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 30 | 1.4 | 5.7 | 0.28 |
| 41 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 27 | 1.3 | 5.3 | 0.29 |
| 42 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 37 | 2.1 | 8.9 | 0.48 |
| 43 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 36 | 2.3 | 8.0 | 0.34 |
| 44 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 32 | 1.3 | 5.3 | 0.29 |
| 45 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 30 | 1.4 | 7.1 | 0.30 |
| 46 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 39 | 1.5 | 8.1 | 0.30 |
| 47 | \(\text{S(O)(O)(N1CCNCC1)c1c2c(cncc2)ccc1} \)                               | 46 | 1.7 | 10.6 | 0.34 |
| 56 | S(=O)(=O)(N1CCNCCC1)c1cc2NC(=O)Oc2cc1 | 36 | 2.3 | 6.2 | 0.32 |

**Table S3.** $K_D$ values of all small-molecule compounds simulated in this work with fragments of α-syn containing either residues 29–49 or residues 121–140 (α-syn-C-term). All simulations were 60 µs in length, except for those of fasudil and Ligand 47 with α-syn-C-term which were 200 µs in length.
| Ligand      | Length (µs) | Notes                          |
|------------|------------|-------------------------------|
| 1 41 (Fasudil) | 1500       | Full-length α-syn             |
| 2 47        | 1500       | Full-length α-syn             |
| 3 23        | 1000       | Full-length α-syn             |
| 4 1         | 60         | α-syn 29–49 fragment         |
| 5 2         | 60         | α-syn 29–49 fragment         |
| 6 3         | 60         | α-syn 29–49 fragment         |
| 7 4         | 60         | α-syn 29–49 fragment         |
| 8 5         | 60         | α-syn 29–49 fragment         |
| 9 6         | 60         | α-syn 29–49 fragment         |
| 10 7        | 60         | α-syn 29–49 fragment         |
| 11 8        | 60         | α-syn 29–49 fragment         |
| 12 9        | 60         | α-syn 29–49 fragment         |
| 13 10       | 60         | α-syn 29–49 fragment         |
| 14 11       | 60         | α-syn 29–49 fragment         |
| 15 12       | 60         | α-syn 29–49 fragment         |
| 16 13       | 60         | α-syn 29–49 fragment         |
| 17 14       | 60         | α-syn 29–49 fragment         |
| 18 15       | 60         | α-syn 29–49 fragment         |
| 19 16       | 60         | α-syn 29–49 fragment         |
| 20 17       | 60         | α-syn 29–49 fragment         |
| 21 18       | 60         | α-syn 29–49 fragment         |
| 22 19       | 60         | α-syn 29–49 fragment         |
| 23 20       | 60         | α-syn 29–49 fragment         |
| 24 21       | 60         | α-syn 29–49 fragment         |
| 25 22       | 60         | α-syn 29–49 fragment         |
| 26 23       | 60         | α-syn 29–49 fragment         |
| 27 24       | 60         | α-syn 29–49 fragment         |
| 28 25       | 60         | α-syn 29–49 fragment         |
| 29 26       | 60         | α-syn 29–49 fragment         |
| 30 27       | 60         | α-syn 29–49 fragment         |
| 31 28       | 60         | α-syn 29–49 fragment         |
| 32 29       | 60         | α-syn 29–49 fragment         |
| 33 30       | 60         | α-syn 29–49 fragment         |
| 34 33       | 60         | α-syn 29–49 fragment         |
| 35 34       | 60         | α-syn 29–49 fragment         |
| 36 35       | 60         | α-syn 29–49 fragment         |
| 37 36       | 60         | α-syn 29–49 fragment         |
| 38 37       | 60         | α-syn 29–49 fragment         |
| 39 38       | 60         | α-syn 29–49 fragment         |
| 40 39       | 60         | α-syn 29–49 fragment         |
| 41 | 40 | 60 | α-syn 29–49 fragment |
|----|----|----|---------------------|
| 42 | 41 (Fasudil) | 60 | α-syn 29–49 fragment |
| 43 | 46 | 60 | α-syn 29–49 fragment |
| 44 | 47 | 60 | α-syn 29–49 fragment |
| 45 | 48 | 60 | α-syn 29–49 fragment |
| 46 | 49 | 60 | α-syn 29–49 fragment |
| 47 | 50 | 60 | α-syn 29–49 fragment |
| 48 | 51 | 60 | α-syn 29–49 fragment |
| 49 | 52 | 60 | α-syn 29–49 fragment |
| 50 | 53 | 60 | α-syn 29–49 fragment |
| 51 | 54 | 60 | α-syn 29–49 fragment |
| 52 | 55 | 60 | α-syn 29–49 fragment |
| 53 | 56 | 60 | α-syn 29–49 fragment |
| 54 | apo | 100 | α-syn-C-term fragment |
| 55 | 1 | 60 | α-syn-C-term fragment |
| 56 | 2 | 60 | α-syn-C-term fragment |
| 57 | 3 | 60 | α-syn-C-term fragment |
| 58 | 4 | 60 | α-syn-C-term fragment |
| 59 | 5 | 60 | α-syn-C-term fragment |
| 60 | 6 | 60 | α-syn-C-term fragment |
| 61 | 7 | 60 | α-syn-C-term fragment |
| 62 | 8 | 60 | α-syn-C-term fragment |
| 63 | 9 | 60 | α-syn-C-term fragment |
| 64 | 10 | 60 | α-syn-C-term fragment |
| 65 | 11 | 60 | α-syn-C-term fragment |
| 66 | 12 | 60 | α-syn-C-term fragment |
| 67 | 13 | 60 | α-syn-C-term fragment |
| 68 | 14 | 60 | α-syn-C-term fragment |
| 69 | 15 | 60 | α-syn-C-term fragment |
| 70 | 16 | 60 | α-syn-C-term fragment |
| 71 | 17 | 60 | α-syn-C-term fragment |
| 72 | 18 | 60 | α-syn-C-term fragment |
| 73 | 19 | 60 | α-syn-C-term fragment |
| 74 | 20 | 60 | α-syn-C-term fragment |
| 75 | 21 | 60 | α-syn-C-term fragment |
| 76 | 22 | 60 | α-syn-C-term fragment |
| 77 | 23 | 60 | α-syn-C-term fragment |
| 78 | 24 | 60 | α-syn-C-term fragment |
| 79 | 25 | 60 | α-syn-C-term fragment |
| 80 | 26 | 60 | α-syn-C-term fragment |
| 81 | 27 | 60 | α-syn-C-term fragment |
|    |     |   |                        |
|----|-----|---|-----------------------|
| 82 | 28  | 60| α-syn-C-term fragment |
| 83 | 29  | 60| α-syn-C-term fragment |
| 84 | 30  | 60| α-syn-C-term fragment |
| 85 | 33  | 60| α-syn-C-term fragment |
| 86 | 34  | 60| α-syn-C-term fragment |
| 87 | 35  | 60| α-syn-C-term fragment |
| 88 | 36  | 60| α-syn-C-term fragment |
| 89 | 37  | 60| α-syn-C-term fragment |
| 90 | 38  | 60| α-syn-C-term fragment |
| 91 | 39  | 60| α-syn-C-term fragment |
| 92 | 40  | 60| α-syn-C-term fragment |
| 93 | 41  | 200| α-syn-C-term fragment |
| 94 | 46  | 60| α-syn-C-term fragment |
| 95 | 47  | 200| α-syn-C-term fragment |
| 96 | 48  | 60| α-syn-C-term fragment |
| 97 | 49  | 60| α-syn-C-term fragment |
| 98 | 50  | 60| α-syn-C-term fragment |
| 99 | 51  | 60| α-syn-C-term fragment |
|100 | 52  | 60| α-syn-C-term fragment |
|101 | 53  | 60| α-syn-C-term fragment |
|102 | 54  | 60| α-syn-C-term fragment |
|103 | 55  | 60| α-syn-C-term fragment |
|104 | 56  | 60| α-syn-C-term fragment |

**Table S4.** List of simulations.
Table S5. Number and residence times of binding events observed in simulations of α-syn-C-term and Fasudil and α-syn-C-term and Ligand 47. For fasudil, if we filter out transient binding events and only consider events in which there are no protein-ligand contacts for at least 1 ns, followed by any persistent protein-ligand contact for at least 1 ns, we observed 6,444 binding events. Doing the same filtering for Ligand 47, we observed 3,966 binding events.

| Minimum Residence Time (ns) | Fasudil  | Ligand 47 |
|-----------------------------|----------|-----------|
| 0.18                        | 48904    | 32673     |
| 1                           | 16709    | 10714     |
| 2                           | 11647    | 7432      |
| 5                           | 6741     | 4654      |
| 10                          | 3679     | 3151      |
| 50                          | 319      | 734       |
| 100                         | 58       | 239       |
| 200                         | 3        | 55        |
| 500                         | 0        | 8         |
| 1000                        | 0        | 3         |
| 2000                        | 0        | 1         |
|                      | α-synuclein + Fasudil | α-synuclein + Ligand 47 | α-synuclein (apo)³⁸ |
|----------------------|----------------------|------------------------|---------------------|
| Cα CS                | 0.50                 | 0.56                   | 0.51                |
| HN CS                | 0.13                 | 0.14                   | 0.14                |
| N CS                 | 1.53                 | 1.65                   | 1.46                |
| C' CS                | 0.34                 | 0.37                   | 0.31                |
| Cβ CS                | 1.12                 | 1.14                   | 1.04                |
| RDC (Q)              | 0.37                 | 0.54                   | 0.41                |
| PRE                  | 0.19                 | 0.24                   | 0.17                |
| ³JHNHA               | 0.98                 | 1.08                   | 1.11                |
| ³JCC                 | 0.16                 | 0.17                   | 0.18                |

**Table S6.** Comparison of the RMSD between experimental NMR measurements previously performed on full length α-synuclein²⁰–²⁷ and simulations of α-synuclein in the presence and absence of small molecules. The apo α-synuclein simulation was previously reported in Robustelli et al.⁸ Chemical shift (CS) RMSDs are reported in ppm, residual dipolar coupling (RDC) and J-coupling RMSDs are reported in Hz, and paramagnetic relaxation enhancement (PRE) RMSDs are unitless.
Table S7. Comparison of the RMSD between experimental NMR chemical shift measurements previously performed on full-length α-synuclein and simulations of α-syn-C-term in the presence and absence of small molecules. Chemical shift (CS) RMSDs are reported in ppm.
Figure S1. Correlation of the contact probability between each residue of α-synuclein and fasudil observed in an unbiased 1.5-ms MD simulation (Fig. 1A) and NMR chemical shift perturbations observed in the presence of 2.7 mM fasudil (Fig. 1B).
**Figure S2.** No substantial difference exists between the conformational ensembles of bound and unbound states of fasudil and α-syn for either the ligand or the protein. Backbone conformation distributions, quantified by the magnitude of the cross product of the vectors between sequential backbone carbon alphas, $|\mathbf{C}_\alpha \text{ vector}| = |(\mathbf{C}_\alpha_i - \mathbf{C}_\alpha_{i-1}) \times (\mathbf{C}_\alpha_i - \mathbf{C}_\alpha_{i+1})|$, are unchanged between bound (A) and unbound ensembles (B). The color bars in (A) and (B) indicate the percentage probability that the $|\mathbf{C}_\alpha \text{ vector}|$ is the given magnitude ($y$-axis) for each residue ($x$-axis). The
bound and unbound distributions of Y133 Cα and Y136 Cα, are shown in (C), where solid lines represent the unbound distributions [U] and dotted lines represent the bound distributions [B]. The distributions of the unbound and bound distributions of |Cα vector| magnitudes at 14 Å$^2$ and 15 Å$^2$ are shown in (D). Finally, bound and unbound distributions of four central dihedrals of fasudil, highlighted in red in the ligand cartoons, are shown in (E) through (H).
Figure S3. A clustering algorithm was applied to all frames containing interactions between fasudil and residues 121–140 of α-synuclein to find regions of local backbone ordering. The four largest clusters with 1 or 2 ordered residues (A) or at least 4 ordered residues (B) are depicted here. Ordered residues are those with |Cα vector|-associated stability for 180 ns, calculated as described in the caption to Figure S2. Trajectory frames in which protein conformations had identical sets of stable residues were considered part of a cluster and overlaid in the figure. Backbone conformations with stable residues are shown as white tubes, and representative conformations of side chains with highest fraction of interactions with fasudil in each cluster are shown (tyrosine residues are colored blue, and D135 is colored green). Representative orientations of the isoquinoline ring of fasudil are shown in red, illustrating the highly heterogeneous ensemble of binding modes observed in each cluster. Binding modes are
very heterogeneous even in the most stable α-syn conformations with at least four stable residues.
Figure S4. Simulations with the α-syn-C-term construct produced a similar contact probability with fasudil when compared to full-length α-syn. Contact probability for fasudil and truncated α-syn-C-term (red) is overlaid onto the full-length contact probability from Fig. 1a. The inset is zoomed into residues 100–140.
Figure S5. Protein-protein contact maps with fasudil unbound and bound for full length α-syn
(A, B), α-syn-C-term (C, D), and apo α-syn-C-term (E). Two protein residues are in contact if their Cα-Cα distance is less than 6 Å. The color bar denotes the percentage of simulations frames that contain the contact.
Figure S6. No substantial difference exists between the conformational ensembles of bound and unbound states of fasudil and α-syn-C-term for either the ligand or the protein (same as Fig. S2, but for α-syn-C-term). Backbone-conformation distributions, quantified by the magnitude of the cross product of the vectors between sequential backbone carbon alphas (i.e., $|Cα_{vector}| = |(Cα_i - Cα_{i-1}) \times (Cα_i - Cα_{i+1})|$), are unchanged between the bound ensemble (A) and unbound ensembles (B) from a simulation of α-syn-C-term in the presence of fasudil, and are also unchanged compared to the ensemble from an apo simulation of α-syn-C-term (C). The
color bars in (A), (B), and (C) indicate the percentage probability that the $|\text{C}_\alpha \text{ vector}|$ is the given magnitude ($y$-axis) for each residue ($x$-axis). The bound and unbound distributions of Y133 C\(\alpha\) and Y136 C\(\alpha\) are shown in (D), where solid lines represent the unbound distributions [U] and dotted lines represent the bound distributions [B]. The distributions of the unbound and bound distributions of $|\text{C}_\alpha \text{ vector}|$ magnitudes at 14 Å\(^2\) and 15 Å\(^2\) are shown in (E). Finally, bound and unbound distributions of four central dihedrals of fasudil, highlighted in red in the ligand cartoons, are shown in (F) through (I).
**Figure S7.** Comparison of mutual information for all interactions observed in the bound ensembles for fasudil (left) and Ligand 47 (right). For each residue, the MI between each potential ligand interaction (hydrophobic contacts, aromatic stacking, charge-charge, and hydrogen bonding) and all other potential ligand interactions is shown. We calculated the MI using binary contact probabilities according to ref. 30. Because the MI is strictly positive and bounded from above by the entropy of the joint distribution, the MI between low-entropy contact pairs will necessarily appear to be low even if they are perfectly coupled. To better resolve these contact pairs, we normalized the MI to the joint entropy of the contact pair, which bounds the MI between 0 and 1.
Figure S8. Ligand 47 orients stacking orientations more strongly than fasudil when charge contacts are made. Stacking orientation between the aromatic rings of fasudil (top row) and Ligand 47 (bottom row) with the side chain of Y136. The plots correspond to orientations of the ligand rings with Y136 when aromatic stacking interactions are formed (first column, see Fig. 2C), when D135 charge interactions are formed (second column), and when both aromatic stacking and D135 charge interactions are formed simultaneously (third column).
Figure S8B. Temporal correlations for fasudil forming π-stacking interactions with Y133 and Y136; charge-charge interactions with D135 persist for as long as 100 ns. The time correlation function, \( C(\Delta t) = \frac{P(A; t_0 + \Delta t|B; t_0)}{P(A)} \), measures the probability of an interaction between fasudil and residue A at time \( t_0 + \Delta t \), given there being an interaction between fasudil and residue B at time \( t_0 \). The normalization is such that as \( \Delta t \rightarrow 0 \), \( C(\Delta t) \rightarrow \left( \frac{P(A|B)}{P(A)} \right) \), which is a measure of the correlation of the interactions with residues A and B (a value of 1 indicates that the probabilities of fasudil interacting with residue A and residue B are not correlated). As \( \Delta t \rightarrow \infty \), \( C(\Delta t) \rightarrow \frac{P(A)}{P(A)} = 1 \), because, at long times, there is no longer any correlation between interactions with A and with B. The plot above shows \( C(\Delta t) \) for the conditional probability of the interactions listed in the legend. Note that simultaneous interactions do not contribute significantly to these correlations, since the persistence of such interactions is only on the order of several hundred picoseconds: The average persistence times of simultaneous protein-ligand
interactions with Y133 and D135, Y136 and D135, and Y133 and Y136, are 570 ps, 350 ps, and 300 ps, respectively. These calculations were performed with fasudil interacting with α-syn-C-term.
Figure S9. Compounds selected to probe how chemical modifications affected the affinity of the interaction and the simulated binding mechanisms. Each compound was simulated with a fragment of α-synuclein containing residues 121 to 140. Commercially available compounds were chosen; Ligand 41 is fasudil.
Figure S10. Individual CSP titration curves for Y125, Y133, and Y136, as a function of compound concentration for Ligand 47, fasudil, Ligand 5, Ligand 2, and Ligand 23.
Figure S11. (A) Contact probabilities (red) computed from an unbiased MD simulation plotted against NMR chemical shift perturbations of Ligand 47 with α-synuclein. (B) Correlation between the contact probability of each residue of α-synuclein and Ligand 47 and NMR chemical shift perturbations observed in the presence of 2.4 mM Ligand 47.
Figure S12. 1D NMR spectra of measured compounds. Spectra were acquired during titration experiments in the presence of α-synuclein, which had a low signal relative to the compounds, and HEPES buffer, which had a high signal relative to the compounds, so only the region of aromatics and amide nitrogens between 10 and 7 ppm is plotted, as HEPES has no signal there. The concentration of the compounds are the maximum used in the titration experiments (same as in the HSQC figures in Fig. 4). Note that the spectra of fasudil and Ligand 23 are very similar because their aromatic moiety is the same.
Figure S13. Time-series of the Cα end-to-end distance (the distance between the Cα atom of the C-terminal and N-terminal residues) and Cα RMSD (the RMSD of each frame to the starting structure of the simulation) for full-length α-synuclein in the presence of fasudil (A) and their autocorrelation functions (B).
**Figure S14.** Autocorrelation functions of Cα RMSD (relative to the first frame), radius of gyration, and the dihedral distance from ideal α-helical, β-sheet, and polyproline II conformations for apo α-syn-C-term, α-syn-C-term with fasudil, and α-syn-C-term with Ligand 47.
Figure S15. Secondary structure propensities for simulations of apo α-syn-C-term, α-syn-C-term with fasudil, and α-syn-C-term with Ligand 47 computed with DSSP. Error bars are calculated by blocking.
Figure S16. Convergence of calculated $K_D$ values for simulations of α-syn-C-term in the presence of fasudil (Red) and Ligand 47 (Blue). Error estimates at each time point were calculated from a blocking analysis of the bound fraction of each compound.
Figure S17. Distribution of residence times for binding events between α-syn-C-term and fasudil (red), and α-syn-C-term and Ligand 47 (blue). Binding events are defined as a sequence of consecutive frames in which there was at least one contact (defined as a non-hydrogen atom-pair with a distance $<6.0 \text{ Å}$) between the protein and the ligand.
Figure S18. Comparison of per-residue backbone secondary chemical shifts from NMR
measurements performed on full-length apo α-synuclein with those calculated from MD simulations of full-length α-synuclein with and without ligands. Chemical shift predictions were made with the SPARTA+ chemical shift prediction algorithm. The left column plots the shifts themselves, and the right column plots the differences between the shifts predicted from simulation with the experimental shifts. The dashed black lines indicate the chemical shift prediction error of SPARTA+ on its training database of folded protein structures for each shift type.
**Figure S19.** Comparison of per-residue coupling constants from NMR on full length α-synuclein with those calculated from MD simulations of full length α-synuclein with and without ligands. The left column plots the coupling constants themselves, and the right column plots the differences between the coupling constants predicted from simulation with the experimental coupling constants.
Figure S20. Comparison of per-residue backbone secondary chemical shifts from NMR
measurements performed on full-length apo α-synuclein with those calculated from MD simulations of α-syn-C-term with and without ligands. Chemical shift predictions were made with the SPARTA+ chemical shift prediction algorithm. The left column plots the shifts themselves, and the right column plots the differences between the shifts predicted from simulation with the experimental shifts. The dashed black lines indicate the chemical shift prediction error of SPARTA+ on its training database of folded protein structures for each shift type.