1

Supplementary Information

Title: Stimulus-responsive Self-Assembly of Enzymatic Fractals by Computational Design

Authors:

Nancy E. Hernández[1,2*], William A. Hansen[2*], Denzel Zhu[3], Maria E. Shea[4], Marium

Khalid[5], Viacheslav Manichev[1,6], Matthew Putnins[2,5], Muyuan Chen[7], Anthony G. Dodge[8],

Lu Yang[1], Melissa Banal[9], Torgny Gustafsson[6,10], Leonard C. Feldman[6,10], Sang-Hyuk

Lee[10], Lawrence P. Wackett[8,11], Wei Dai[2,9] and Sagar D. Khare[1,2].

nancy.hernandez@rutgers.edu, wah49@scarletmail.rutgers.edu, denzel.zhu@rutgers.edu,
mes373@scarletmail.rutgers.edu, mariumkhalid105@gmail.com, v.manichev@rutgers.edu,
mputnins@gmail.com, muyuan.chen@bcm.edu, dodg0001@umn.edu,
ly118@scarletmail.rutgers.edu, mmb298@scarletmail.rutgers.edu, gustaf@physics.rutgers.edu,
l.c.feldman@rutgers.edu, shlee@physics.rutgers.edu, wacke003@umn.edu,
wdai@proteomics.rutgers.edu, sagar.khare@rutgers.edu

This PDF file includes:

Supplementary Methods

Supplementary Discussion

Supplementary References

Figs. S1 to S31

Tables S1

Movies S1-S3

Supplementary Methods

INDEX

SI 1.X Computational Design

SI 2.X Experimental Characterization

1 **SI 3.X Microscopy Experiments**

2   **SI 3.1** Transmission electron microscope (TEM)

3   **SI 3.2** Atomic force microscopy (AFM)

4   **SI 3.3** Helium ion microscopy (HIM)

5   **SI 3.4** High-resolution fluorescence microscopy

6   **SI 3.5** Cryo-EM Tomographic tilt series acquisition and reconstruction

7   **SI 3.6** Cryo-EM AtzAM1 and AtzCM1 model fitting and statistical analysis

8 **SI 4.X Enzymatic Assays**

9   **SI 4.1** Enzymatic activity was measured using the Berthelot assay

10   **SI 4.2** Temperature stress activity assays

11   **SI 4.3** Shaking stress activity assay

12   **SI 4.4** Construction and assay of Basotect® polymer foam with trapped assemblies and
13 free enzymes

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

(SI 1.1) Preparation of a two-component scaffold library – Crystal structure files for AtzA (PDB:4V1X) and AtzC (PDB:2QT3) were subject to several preparatory scripts to clean, symmetrize, and process the files for Rosetta Design[1-3]. The processed crystal structure files were then subject to a Rosetta Fast Relax[4] protocol to obtain starting structures of sufficiently low Rosetta Energy to serve as starting structures and ideal wild-type models. We created a two-component (AtzA:monomer and AtzC:monomer) scaffold library where the rigid-body position of AtzC:tetramer altered with respect to AtzA:hexamer along aligned C_2 symmetry axes via rotation and translation. To prepare the scaffold library we first aligned the proteins along paired C_2 symmetry axes (A+B chains for both AtzA and AtzC). We then translated AtzC along the aligned C_2 symmetry axis until the backbone atoms of each structure were at least 3Å apart to find the minimum starting distance (125Å). From the minimum starting distance we translated AtzC(monomer) in intervals of 1Å to a maximum distance of 145Å. To complete the two component scaffold library, for each translated AtzC(monomer) position we rotated the AtzC(monomer) about the C_2 symmetry axis by 360° in intervals of 5° for a total of 1440 library members.

(SI 1.2) RosettaMatch: simultaneous fusion domain and peptide pair stitching – After visual inspection of the two-component scaffold library, we noted the accessibility of the AtzA N-terminus and the AtzC C-terminus along the C_2 symmetry axis (chains A+B). Therefore, we decided to fuse the N-terminus of an fyn-SH2 super-binder (PDB:1A0T) to the C-terminus of AtzC and the C-terminus of the fyn-SH2 peptide binding partner to the N-terminus of AtzA. To achieve the simultaneous fusion, we converted the SH2-peptide crystal structure into an all-Cα 'ligand' file and used RosettaMatch[5] with geometric constraints to sample all sterically feasible rigid body placements of the SH2-peptide between each AtzA-AtzC pair in the two-component scaffold library. The geometric constraints used to coordinate the SH2 domain for simultaneous fusion were derived from a non-redundant protein library generated by the RCSB-PDB[6]. From N to C terminus, regardless of secondary structure we collected distances and angles between backbone atoms (Cα, nitrogen, and carboxyl carbon) up to and including 7 residues downstream (sequence-space) of each residue along the primary structure. The averages and standard deviations of these distributions were used to place matching constrains between residues of the AtzA-AtzC termini and the all-Cα SH2-peptide ligand. The full-atom SH2-peptide crystal structure was re-threaded back onto each of the matched SH2-peptide ligands creating 7,005 models with paired termini in proximally close and geometrically favorable positions. Rosetta GeneralizedKIC (kinematic loop closure)[7] was used to covalently link the paired termini and generate 3 potential linker-models for each matched SH2-peptide model, creating a library of 21,015 fused and bound AtzA-AtzC pairs.

(SI 1.3) Rosetta Design: novel interface design – A Rosetta FastRelax protocol was used to design novel interfaces generated in the previous steps. For each round of the FastRelax protocol we allowed all residues to sample every rotameric degree of freedom. In addition to rotameric

sampling, novel interface residues with a maximum Cα-Cα interface distance of 6Å and linker residues were allowed to change residue identity before energy minimization. All but linker backbone atoms were constrained with atom-coordinate constraints to favor the SH2-peptide placements determined in the RosettaMatch step. A final visual inspection was made to confirm the validity of each mutation made during this protocol. Mutations alleviating steric clashes were widely accepted; spurious mutations with little benefit were reverted to native residue identities before a subsequent round of repack and energy minimization[8].

(SI 1.4) Stochastic fractal assembly simulation summary – In order to better predict the supramolecular structure and topology we created a stochastic fractal assembly simulation that utilizes Boltzmann weighted probability distributions for an ensemble of predicted low-energy binding modes along the C_2-symmetry axes of the AtzA-AtzC pairs. The algorithm operates by starting with one oligomer (AtzA for this study) and attaches each complementary oligomer in layers. The Boltzmann probability distribution was used to decide how the oligomers in each layer were placed. A few key assumptions were made during the simulations that were based on chemical intuition. We assumed: 1) The symmetric divalent connection along a C_2-symmetry axis (two chains of pY-AtzA bound two chains of AtzC-SH2) would be energetically more likely than the monovalent connection formed between just one chain from each oligomer—reducing the probability of monovalent connection to an insignificant value. 2) Flexibility in the linker region would only lead to variations along the C_2-symmetry axis via the translation and rotation parameters used to create the two-component library—maintaining the inherent symmetry found in either oligomer. 3) Symmetry could but is not required to extend to 3 or 4 component substructures. Mixed vertex-centered and edge-centered species could occur around a single AtzA. This would lead to a substructure where two AtzC oligomers have a 180° bound-angle about AtzA, different from the more symmetric 120° bound-angles. 4) Changes in size and topology would arise from concentration changes of the enzyme and would need to be represented in the algorithm. 5) During fractal growth it is possible (and likely) that oligomers in one layer could come within 125Å (minimum connected distance) of other oligomers within another layer even if they are not connected. The details of this algorithm are described below.

(SI 1.5) Coarse-graining AtzA-C oligomers for stochastic fractal growth simulations – We predicted that fractal growth could continue indefinitely in all directions. To reduce the computational load and file size of particle models exceeding 100s or even 1000s of oligomers, we thought to coarse-grain our symmetric oligomers by reducing each chain to just 10 representative points in space (60 and 40 for whole hexamer and tetramer respectively). To coarse-grain we used a K-means-style clustering algorithm to place the 10 points at locations with the highest concentration of Cα atoms in each monomer (chain A). We then calculated and applied the symmetric transform to the 10 representative points to obtain a coarse-grained representation of each oligomer (hexamer and tetramer). When each point is converted into a sphere with a 12Å radius, the coarse-grained model shows agreement with the overall shape and size of the full-atom model.

(SI 1.6) Stochastic fractal assembly simulation – After experimental analysis revealed the best pY-AtzA and AtzC-SH2 variants we repeated the above Rosetta FastRelax protocol on all 21,015 fused AtzA-AtzC pairs while forcing the sequence identity of the best pY-AtzA and AtzC-SH2 pair. We generated an energy profile (Figure 1E-F) for conformations whose evaluated energy scored better than the wild-type components (504 models). Each conformation was represented by three parameters, translation (d), rotation (θ), and axis-binding preference (vertex or edge centered). The conformations were assigned Boltzmann weighted probabilities which were used to randomly propagate the coarse grained A-C components during simulation. We varied the kT term to obtain a total of 5 different Boltzmann weighted probability distributions (kT = 1, 3, 5, 7, and 9). Propagation was achieved by alternating layers of AtzA and AtzC components starting from an initial seed component (pY-AtzA in this study) which would continue until either placement of new components was determined either impossible or improbable or an external criterion was met (number of layers, size of particle, etc.). The propagation algorithm can be broken into 10 steps at any given layer:

1) Choose the number of components in the previous layer (or the seed component) based on a variable fraction with which new complementary oligomers would be placed.

2) Randomly select individuals from the chosen pool (1) to place new components.

3) Based on a random generated number from 0.0-1.0, select a matching d-θ-axis conformation via the probability of the conformation.

4) Randomly select available C_2-symmetry axes of the individual selected in (2) compatible with the conformation chosen in (3).

5) Choose whether or not to keep the selected C_2-symmetry axis (4) based on a variable null probability.

6a) If (5) passes the null, apply the rigid body transformation (d and θ) to the new member of the current layer.

6b) If (5) fails the null, mark the C_2-symmetry axis (4) of the individual selected in (2) as unviable and continue.

7) Repeat 3-6b until all C_2-symmetry axes of individual (2) are exhausted.

6

8) Perform a coarse grid-based clash check to ensure new layer members are sterically feasible.

9) Repeat 2-8 until all of the individuals chosen in (1) are exhausted.

10) Move to the next layer.

(SI 1.7) Temperature, fraction, and null parameter sweep – Varying the fraction (1) and null (5-6b) parameters gave rise to changes in topology and structure. We created 100 fractal models for each combination of fraction (range: 0.1-1.0, interval: 0.1) and null (range: 0.0-0.9, interval: 0.1) using the 5 different Boltzmann weighted probability distributions (with varying temperature)—creating 50,000 total fractal assemblies. An external criterion (15 layer limit) was set during the simulation stage to reduce the computational load of the simulation program as well as on the downstream data processing software. We analyzed each particle's individual size, number of layers, AtzA branch ratio (number of AtzC units bound to a unit of AtzA), lacunarity, and dimensionality (Df) from a 2D image. For every combination of temperature, fraction, and null we averaged the data across the 100 fractal assemblies. The results can be found in Figure S1 and S2.

(SI 1.8) Preparing fractal models for image analysis – Each fractal assembly was passed through a deterministic PyMOL script that would color the assembly black, convert the background white, show as spheres of scale 12Å, orient the image such that the longest diameters are in the X-Y plane, remove the glossy lighting and shine from the sphere models, and finally ray-trace render the image.

(SI 1.9) Determining fractal lacunarity and 2-D fractal dimension with ImageJ - The FracLac package[9] designed for ImageJ[10] was used to determine both the 2D lacunarity and fractal dimension (Df). With FracLac mode on, outside of the standard parameters, we checked the 'alternate random generator' box and allowed the minimum pixel size to be 1, and the color code was turned off. We then ran in batch-mode to process all of the fractal images. ImageJ outputs four files: summary, box count per grid, scan types, and batch data. Lacunarity and dimension were taken from the summary file for the parameter sweep while the 2D log vs log plot values were taken from the box counting grid file (ε and F).

(SI 1.10) Computational comparison of simulation and tomography fractals from Cryo-EM – Fitting of the experimentally computed protein density (Cryo-EM tomography) resulted in Cartesian coordinates representing the center of mass of the oligomeric components. To compare the experimental results to simulation we ran the simulation until at least a total of 5000 components were present in the model and calculated the geometric centers for all oligomeric components in the coarse-grained assembly to create new center-of-mass models. Using the

1 experimentally derived Cartesian coordinates and the center-of-mass models we performed a

2 computational analysis (see Cryo-EM fitting and statistical analysis below) to evaluate the fractal

3 size, nearest component neighbor distances, and relative AtzA-AtzC ratio (Fig. 3H,I). We

4 analyzed the 3D fractal dimension (Fig. 3J) with a 3D box counting program that counts the

5 number of geometric centers within a scaling (doubling) box size. The 2D fractal dimension (Fig.

6 3J) was calculated in the same way as previously mentioned. We found highest agreement of

7 simulations with $kT = 9$, $P_{null} = 0.1$, and $C_{frac} = 1.0$. An array of fractal images that represent the

8 average fractal for each value of P_{null} and C_{frac} at kT = 9 can be found in Figure S2.

9

10 **(SI 2.1) Creation of the designed AtzA, AtzB, and AtzC fusion constructs –** The DNA

11 sequence of the full-length *atzA* was amplified from the *pMD4::atzA*; *atzB* amplified from

12 *pAAJLS3::atzB;* and *atzC* was amplified from *pKK223-3::atzC*.[11-14] The Src kinase activator

13 phosphopeptide sequence, EPQYEEIPIYL, was created by ordering two complementary primers

14 that formed a linear fragment encoding the peptide sequence, used with the amplified *atzA* gene

15 and inserted into the linearized *pET15b+* vector through Gibson Assembly.[15] The Fyn SH2

16 superbinder gene was ordered as a gBlock fragment[15,16] and inserted into *pET29b+* (linearized with

17 *Nde*I and *Xho*I) using Gibson Assembly.The Fyn SH2 amplified gene was designed to be placed

18 on the C-terminal side of the *pET15b+::atzB* and *pET29b+::atzC* with a flexible GSS linker

19 between the proteins. The Fyn SH2 superbinder amplified gene *SH2* and the *atzC* amplified gene

20 were both inserted into the *pET29b+* linear vector using Gibson Assembly. The *atzBSH2* fusion

21 gene was ordered as a Gibson fragment[15] and inserted into the *pET15b+* linear vector using

22 Gibson Assembly. Point mutations were introduced using the QuickChange Site-Directed

23 Directed Mutagenesis Kit (Agilent Technologies) to create the final designs for AtzA and AtzC

24 models. DNA sequencing was used to confirm proper insertion and mutations (Genscript).

25

26 **(SI 2.2) AtzA and AtzC expression and purification –** The *pET15b+::atzApep* and

27 *pET29b+::atzCSH2* plasmids were co-transformed into *Escherichia coli* BL21 (DE3) with *pAG*

28 plasmid containing genes for the chaperone proteins, *groEL* and *groES* [17].For expression of the

29 AtzA models a 10 mL LB culture with 30 μg/mL of chloramphenicol and 100 μg/mL of

30 ampicillin was inoculated with a single colony and incubated overnight at 37°C and 250 rpm.

31 For the expression of the AtzC models a 10 mL LB culture with 30 μg/mL of chloramphenicol

32 and 50 μg/mL of kanamycin was inoculated. After growing overnight, the 10 mL cultures of the

33 AtzA and AtzC models were used to inoculate 500 mL of LB media, which was grown at 37°C

34 to an OD_{600} of 0.5-0.6, at which point the expression of chaperones was induced with the addition

35 of 1% (wt/vol) L-arabinose and grown for an additional 1-2 hours at 16°C . Expression of the

36 AtzA and AtzC models was then induced with 0.1mM IPTG (isopropyl-β-D-thiogalacto-

37 pyranoside) and grown overnight at 16°C. All subsequent steps were performed at 4°C. Cells

38 were centrifuged at 6,000 x g for 30 min. Cell pellets were re-suspended in 30 mL of 25 mM

39 HEPES, 200 mM NaCl, 5% glycerol, 40 mM imidazole, pH 7.5, and lysed by sonication. Cell

40 extracts were obtained by centrifugation at 50,000 x *g* for 30 min at 4°C. Protein purification was

41 performed using 5 mL Ni-NTA agarose resin (Qiagen) equilibrated with 10 mL of 25 mM

42 HEPES, 200 mM NaCl, 5% glycerol, 40 mM imidazole, pH 7.5. The lysate was applied to the

resin, the resin was washed with 45 mL of the same buffer, and the protein eluted with 20 mL of 25 mM HEPES, 200 mM NaCl, 5% glycerol, 400 mM imidazole, pH 7.5,. The purified protein was buffer exchanged (PD10-desalting column, GE Healthcare #17085101) into 50 mM HEPES, 100 mM NaCl, 5% glycerol, pH 7.4 (HNG). AtzA was expressed in high yields and precipitated if the buffer was not exchanged quickly. Proteins were frozen using liquid nitrogen and stored at -80°C. All proteins precipitated if dialyzed in HNG for 2 hours.

(SI 2.3) AtzB expression and purification – The *pET15b+::atzBSH2* plasmid was transformed into *E.coli* BL21 (DE3) cells. For expression of AtzB, a 10 mL LB culture with 100 μg/mL of ampicillin was inoculated overnight at 37°C and 250 rpm. The 10 mL overnight culture was used to inoculate 500 mL of LB media which was grown to an OD_{600} of 0.5-0.7 and induced with 1 mM IPTG and grown overnight at 16°C. The same purification protocol for the AtzA and AtzC models was used for AtzB. AtzBSH2 did not express if grown with zinc sulfate, as had been done customarily in previous literature.[13]

(SI 2.4) Src human kinase, super binder SH2 domain, SH2-DhaA expression and purification – The expression plasmid for Src human kinase[18] (gift from John Chodera, Nicholas Levinson, and Markus Seeliger. Addgene plasmid # 79700 was co-transformed with the expression plasmid for *Yersinia* YopH protein tyrosine phosphatase (PTPase)[18] (gift from John Chodera, Nicholas Levinson, and Markus Seeliger, Addgene plasmid # 79749) into *E. coli* Rosetta2 (DE3) (Novagen). For Src kinase expression a 10 mL LB culture with 50 μg/mL spectinomycin and 100 μg/mL of ampicillin was inoculated with a single colony and incubated overnight at 37°C, 250 rpm. The overnight culture was used to inoculate 500 mL of LB media which was grown to an OD_{600} of 0.5-0.7 and induced with 1mM IPTG and grown overnight at 18°C. The super binder SH2 domain and SH2-DhaA were transformed into E. coli BL21 (DE3) and expressed in the same way as the Src kinase above. Purification for the Src kinase was performed similarly and with the same buffers as AtzAM1, AtzBSH2, and AtzCM1. While, the super binder SH2 domain and SH2-DhaA were purified with the same purification protocol but with the following buffers: a wash buffer containing 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4, 20 mM imidazole and an elution buffer containing 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4, 200 mM imidazole. All proteins were buffer exchanged into HNG, frozen in liquid nitrogen and stored at -80°C.

(SI 2.5) YopH phosphatase construct, expression, and purification – The linear catalytic domain *YopH* gene (residues 164-468) was amplified from *pET13S-A::YopH*[18] and inserted with Gibson Assembly into a linearized pET15b+ vector. A 10 mL LB culture with 100 μg/mL of ampicillin was inoculated with a single colony and incubated overnight at 37°C. The expression and purification protocol is the same as the protocol used for the Src kinase.

(SI 2.6) Biuret hydrolase and cyanuric acid hydrolase expression and purification – Biuret hydrolase (BH)[19] expression strain (*E. coli* DH5α) and the *Moorella* Cyanuric acid hydrolase (CAH)[20] strain (*E.coli* BL21 (DE3)) were provided by Dr. Larry Wackett. A 10 mL culture with 50 μg/mL of kanamycin was inoculated for both BH and CAH and incubated at 37°C until OD$_{600}$ of 0.5-0.7 and induced with 1 mM IPTG for 4 hours at 37°C, 250 rpm. The expression and purification protocol is the same as the protocol used for the Src kinase.

(SI 2.7) Enzyme-linked immunosorbent assay (ELISA) – Phosphorylated AtzAM1 (pY-AtzAM1) was loaded onto clear flat-bottom immuno 96-well plates (Thermo Scientific item # 442404) at 20μg/mL and 1.25μg/mL in 50μL 1X PBS (Gibco pH 7.4, #10010023) overnight at 4∘C. Plates were rinsed twice in 200μL 1X TBS (Biorad #1706435). 1% BSA in TBS 0.05% Tween 20 was used to block wells at 200μL block solution for 1.5hr at 25°C under gentle agitation. Anti-phosphotyrosine 4G10 Platinum HRP conjugate (EMD #16-316) was diluted 1:5000 in 1% BSA TBS 0.05% Tween 20 and loaded onto the well at 25°C for 1.5hr under gentle agitation. Excess anti-phosphotyrosine was washed off with 200μL of TBS 0.05% Tween 20 in triplicate. To detect bound antibody, 100μL of TMB substrate reagent (Biolegend #421101) was added to each well and incubated for 5 minutes at 25°C. 100μL of TMB stop solution (Biolegend #423001) was added to the wells. Absorbance was read at 450nm using the Tecan Infinite M200 Pro plate reader.

(SI 2.8) Bio-layer interferometry (BLI) – AtzAM1 was phosphorylated using the conditions described below. pY-AtzAM1 was then biotinylated at 10mM Sulfo-NHS-Biotin (APExBIO) for 30min at 25°C. Excess biotin was buffer exchanged with a PD-10 desalting column (GE Healthcare) equilibrated with HNG. Biotinylated pY-AtzAM1 was loaded onto streptavidin (SA) coated biosensors (ForteBio) and used for BLI. AtzCM1 was flowed in from 4nM to 4μM. BLI experiments were performed using the BLItz System (ForteBio).

(SI 2.9) Phosphorylation, assembly formation, and disassembly – The phosphorylation protocol was based upon Src kinase activity assay by Sigma (Catalog # S1076). In a final reaction volume of 150μL, 3μM AtzAM1 was mixed into 1X Kinase Activity Buffer (4mM MgCl$_2$, 2.5mM MnCl$_2$, 0.25mM DTT, 5mM MOPS, 2.5mM glycerol-2-phosphate, 1mM EGTA, 400nM EDTA, pH 7.6), 2.5 mM MnCl$_2$,HNG, 2 mM ATP, 800ng Src kinase, and incubated for 7 – 16 hr at 25°C for phosphorylation to occur. After phosphorylating, AtzCM1 was added to a final 2μM concentration. Assembly was allowed to form at 2hr 25°C. Disassembly was performed by adding 4.8μg of YopH phosphatase into the 150μL reaction mixture after assembly formation occurred. Size measurements using DLS were performed to determine assembly formation/disassembly.

(SI 2.10) Dynamic light scattering (DLS) – 50 μL of an assembly sample was used for size determination using a Malvern Zetasizer and a quartz cuvette (ZEN2112, Malvern). Ten spectra measures were recorded for eleven replicates at 25 °C. The standard operating procedure accounted for 5% glycerol in solution.

(SI 2.11) DLS Inhibition Experiment - 6 μM pY-AtzAM1 was phosphorylated (1X KAB, 2 mM ATP, 1 mM DTT, HNG, 1 μg Src kinase) in a reaction volume of 75 μL. Incubation time was overnight at 25°C. SH2 or SH2-DhaA was added to each sample at 0 μM, 3 μM, 6 μM, 9 μM, 12 μM, 15 μM, 18 μM final concentration and allowed to "block" binding sites on the pY-AtzAM1 for 1 hr at 25°C. AtzCM1 was added to each sample at 2 μM final concentration. Therefore, the final concentrations of all components was 3 μM pyAtzA, 1 μM AtzCM1, 0 μM - 18 μM SH2 or SH2-DhaA. The sample was incubated for 2 hr at 25°C. DLS was performed to analyze assembly sizes. DLS was performed at 25°C, 50 μL/sample volume, in a low-volume quartz sizing cuvette (Malvern; ZEN2112) using a Zetasizer Nano ZS (Malvern). Measurements were performed in triplicates while each sample was read and averaged 15 times. This protocol was repeated at a final concentration of 1 μM pyAtzA, 0.66 μM AtzCM1, 0 μM -6 μM SH2-DhaA. Curve fitting was performed in MATLAB (R2016b; Mathworks) using the general model:

$$f(x) = \frac{A}{1 + e^{-k*(x-x_0)}} + B$$

where A, B, k, x_0 are constants. Adjusted R^2 was used to determine model validity. Inhibition concentration 50 (IC50) was determined based upon concentration of inhibitor that resulted in assembly size of 100nm measured.

(SI 2.12) DLS Titration Experiment – 6 μM, 3 μM, 1.5 μM, 0.5 μM, 0.1 μM pyAtzA was phosphorylated (as described previously) with an incubation time of overnight at 25°C. Either AtzCM1 wildtype (WT) or AtzCM1 superbinder (SB) was added to each sample at 2 μM, 1 μM, 0.5 μM, 0.25 μM, 0.50 μM final concentration. The sample was allowed to incubate for 2 hr at 25°C. Therefore, the final concentrations of all components was from 3 μM – 0.05 μM pyAtzA, 2 μM – 0.05 μM AtzCM1-WT or AtzCM1-SB. DLS was performed at 25°C, 50 μL/sample volume, in a low-volume quartz sizing cuvette (Malvern; ZEN2112) using a Zetasizer Nano ZS (Malvern). Measurements were performed in duplicate with each sample read and averaged 15 times.

(SI 2.13) DLS Kinetics (varying ATP) Experiment – An assembly mixture of 3 μM non-pyAtzA and 2 μM AtzCM1 was prepared (as described previously) and syringe-filtered at 0.22 μm. To each 50 μL reaction volume, 1.2 μg of src kinase was added. Size was monitored continuously for 30 min at 25°C in a low-volume quartz sizing cuvette (Malvern; ZEN2112) using a Zetasizer Nano ZS (Malvern) at 50 μL/sample. Measurements were performed in triplicates. Each sample was read and averaged five times over the course of 25 seconds for a

single time point. Curve fitting was performed in MATLAB (R2016b; Mathworks) using sloping spline function, with varying smoothing parameters. Adjusted R^2 was used to determine model validity.

(SI 3.1) Transmission electron microscope (TEM) – Assembly (3 μM pY-AtzAM1 and 2 μM AtzCM1) and non-assembly (3 μM non-pyAtzA and 2μM AtzCM1) samples were mixed, and diluted ten-fold in deionized water. The diluted samples were applied to the carbon-coated FCF400-Cu grids (Electron Microscopy Sciences, Hatfield, PA) which were glow-discharged for two hours under UV light to render the grids hydrophilic and adsorptive. A drop of sample (~5uL) was added on a piece of wax film and the grid was placed onto the sample droplet for absorption for two minutes. Excess sample solution was removed with a filter paper. A drop (~5uL) of 1% uranyl acetate was dropped on the wax paper and the grid was placed onto the staining solution droplet for two minutes to stain. Excess staining solution was removed by blotting with a filter paper, the grids were allowed to air dry for two minutes. Images were collected on JEOL 1200EX electron microscope with AMT-XR41 digital camera.

(SI 3.2) Atomic force microscopy (AFM) – The assemblies were directly visualized by non-contact mode atomic force microscopy (AFM) Parks Systems. Samples were prepared by depositing 20 μls of sample on silicon wafer and incubated for 5 minutes. After incubation, the silicon was washed with deionized water to remove salt and air dried overnight at 25°C. Assemblies were visualized by anAFM (Parks System). The AFM was used in non-contact mode (330 kHz resonant frequency and 42 N/m spring constant, PPP-NCHR Park Systems, #610-1051). Images were taken with 2048x2048 pixels with scan rates of 2 μm/s to 30 μm/s. The AFM images analysis was performed using Gwyddion software[21].

(SI 3.3) Helium ion microscopy (HIM) – The AFM sample preparation on a silicon wafer was used for HIM. Imaging was done on the Carl Zeiss Orion Plus Helium Ion Microscope (Carl Zeiss Microscopy, Peabody, MA) operating at 30 KeV acceleration voltage with a beam currents of about 1 pA. Most samples did not exhibit significant charging therefore electron flood gun was not used for charge neutralization. The vacuum reading in the analysis chamber during imaging was 2×10^7 torr.

(SI 3.4) High-resolution fluorescence microscopy – For the growth video, 20 μL of 3 μM AtzAM1 and 2 μM AtzCM1 sample (with all the required buffers as described previously) was deposited on a glass cover and 0.2 μm of Src kinase was added to the sample to allow for assembly formation to occur. The sample was monitored for an hour. For the 3-component assembly image (3 μM pY-AtzAM1, 1 μM AtzBSH2, 2 μM AtzCM1) the AtzBSH2 protein was dye labeled with the Alexa Fluor™ 647 NHS Ester (Succinimidyl Ester, ThermoFisher Scientific

#A2006) and buffer exchanged into HNG with a PD10-desalting column. Fluorescent images along with bright-field images were collected. Images were captured using a Nikon Ti-E inverted microscope. A Coherent Genesis laser at 567 and Coherent Obis Laser at 647 were used for fluorescent imaging, using 1mW power.

(SI 3.5) Cryo-EM Tomographic tilt series acquisition and reconstruction – For cryo-electron tomography, an AtzAM1 and AtzCM1 assembly sample was mixed with 10 nm gold fiducial markers to facilitate alignment in data processing. An aliquot of 3.5ml sample was applied to 2.0/1.0mm Quantifoil holey grids (Quantifoil, Germany) and plunge frozen using a Leica EM GP plunger (Leica). Tomographic tilt series acquisition was performed on a Talos Arctica microscope (Thermal Fisher) operated at an acceleration voltage of 200kV. This microscope was equipped with a field-emission gun, Volta phase plates, Gatan postcolumn energy filter and a K2 summit direct electron detector. Tilt series were collected at 39,000x microscope magnification with -0.5 μm defocus using FEI Tomography software. The sampling of the data was calibrated to be 3.49 Å/pixel. Typically, a tilt series ranged from -60° to 60° at 3° step increment. The accumulated dose for each tilt series was 60 electrons/Å2. Tilt series were aligned based on fiducial gold markers using the IMOD package[22]. 3D tomograms were obtained by weighted backprojection of aligned tilt series. Visualization and annotation of the 3D volumes were done in Chimera[23].

(SI 3.6) Cryo-EM AtzAM1 and AtzCM1 model fitting and statistical analysis – AtzAM1 and AtzCM1 complex subtomograms were extracted from 3D tomograms and bandpass filtered to reduce high frequency noises and low frequency gradient from ice thickness variation. Centers of AtzAM1 and AtzCM1 densities were identified as peaks within solid voxel clusters that were approximately sizes of an AtzAM1 hexamer, or an AtzCM1 tetramer. Potential free AtzAM1 or AtzCM1 complexes that were too close to a neighboring voxel peak (<120A) were removed. Assignment of AtzAM1 or AtzCM1 to an identified voxel cluster was done by applying the condition that AtzAM1 and AtzCM1 alternate in a chain. Densities that had three or more linkers to neighbors were assigned to be AtzAM1. Linear, unbranched assemblies were assigned by first determining identity of one end based on cross-correlation scores between the end peak densities and AtzAM1 or AtzCM1 models computed from their PDB structures. Assignment conflicts were resolved by pruning along the branches in the order of intensity values. The above protocol was first applied to a small assembly, and optimized and validated by human visual inspection before it was used on larger assemblies. Coordinates and connection information of each AtzAM1 or AtzCM1 complex in an assembly were extracted and used for statistical analysis and for comparison to simulation data. The volume of the assembly is defined by the volume of the convex hull that encloses all determined AtzAM1 or AtzCM1 molecule.

(SI 4.1) Enzymatic activity was measured using the Berthelot assay – Assembled enzyme samples (1.5 μM AtzAM1, 0.5 μM AtzBSH2, and 1 μM AtzCM1) were made by incubating the enzymes in 1X kinase activity buffer (with no DTT), 2.5 mM MnCl2, HNG, 0.2 μM Src kinase, and 2 mM ATP in a total volume of 500 μl at 25°C for four hours. The unassembled enzyme

samples were prepared using the same conditions, except no ATP was added to the sample. DLS was performed to verify assembly formation. 10 μL of 20 mM Atrazine dissolved in methanol was added to each 500 μL sample, for a final concentration of 400 μM atrazine, and another sample with the same conditions had no substrate added in order to establish a baseline measurement. Each condition was done in triplicate. After the addition of substrate, the samples are shaken at 100 RPM for 1.5 hr at 25°C. 140 μL of each sample is transferred to PCR tubes, then boiled at 99°C for 1.5 minutes, and then cooled at 4°C. The 140 μl were transferred to 1.5 mL microcentrifuge tubes and spun down at 20,000 rcf for 20 minutes to remove precipitated protein. 80 μl of the supernatant was used for the following steps. 1μg per 20 μL of sample of CAH and 1μg per 20 μL of sample of BH was added to each sample. The samples were incubated at 25°C for 2 hours to allow for the complete conversion of the cyanuric acid to ammonia by CAH and BH. The Berthelot assay was performed in triplicate on the resulting samples to determine the production of ammonia. For every mole of cyanuric acid produced, one mole of ammonia was assumed to have been produced. 20 μL of each sample was added to a 96-well plate (Greiner half area clear #675101). 60 μL of solution A (0.05 g/L sodium nitroprusside and 10g/L phenol) was added and mixed into every sample. Then 80 μL of solution B (5 g/L NaOH and 8.4 mL/L bleach) was added and mixed into every sample. The samples were incubated for 30 minutes at 25°C for a blue color to develop. The absorbance at 630 nm was read using Tecan Infinite M200 Pro plate reader. The extinction coefficient was determined using standards of cyanuric acid at known concentrations in the enzyme activity buffer that had been reacted with the BH and CAH for 2 hours.

(SI 4.2) Temperature stress activity assays – Assembled and unassembled enzyme samples were made as described above and incubated at 25°C for 4 hours to allow full assembly formation. The assemblies were then incubated at the following temperatures: 25°C, 40°C, 45°C, 50°C, 55°C, and 60°C for fifteen minutes, and cooled back to 25°C before the addition of 400 μM atrazine. After atrazine was added, the enzyme activity assay was performed as described above.

(SI 4.3) Shaking stress activity assay – Assembled and unassembled enzyme samples were made as described above and incubated at 25°C 4 hours. Both samples were shaken at 50, 100, 150, 200, 225, and 250 RPM 25°C for 1 hour before any addition of atrazine. 400 μM atrazine was added to the samples and shaking continued at their respective shaking speeds for 1.5 hour. The rest of the activity assay protocol was conducted the same as described above.

(SI 4.4) Construction and assay of Basotect® polymer foam with trapped assemblies and free enzymes – Hydrolyzed TEOS was prepared by combining 7 ml TEOS (Aldrich #131903), 3 ml water, and 0.04 ml 0.1N hydrochloric acid and stirring the solution for 2 h at room temperature[24]. Basotect® polymer foam (Procter and Gamble UPC# 0 37000 43515 0) was cut into 2.0 x 2.0 x 0.3 cm squares with a razor and 0.250 ml of assemblies or free enzyme solution was spotted onto each 2 x 2 cm face of the foam squares. Aliquots (1.0 or 0.5 ml) of hydrolyzed TEOS were diluted with HNG buffer to a final volume of 10 ml (10% or 5% TEOS). A single

1   application of 5% or 10% hydrolyzed TEOS solutions was done with a small paint brush

2   (Richeson 95822). The TEOS was allowed to set for 2 h, and then liquid was squeezed out of

3   each foam square and total protein concentration in the liquid was measured with the Bradford

4   assay (BioRad #500-0006). To assay activity in the embedded foam, 1 ml of 150 μM atrazine in

5   1X phosphate buffered saline (pH 7.4) was soaked into the foam squares and incubated for 1.5

6   hour at 25°C. Liquid was squeezed out after incubation and boiled as above to inactivate eluted

7   enzymes. Cyanuric acid produced during the incubation was assayed as described except that the

8   Berthelot reactions were conducted in 10 x 4 x 45 mm cuvettes (Sarstedt #67-742) and read

9   using a Beckman DU 640 spectrophotometer.

10

Supplementary Discussion

Fractal dimension from image analysis

13   The fractal (Hausdorff-Besicovitch) dimension[25], a concept introduced in 1918 to measure the

14   dimensions and local size of a shape, has been used to characterize simulated fractal patterns[26,27] as

15   well as peptide-based fractals obtained on a surface and imaged with AFM[28]. The Hausdorff-

16   Besicovitch dimension equation, defined by the divider formula

17

$$D_f = \lim_{r \to 0} \left(\frac{\log N(r)}{\log(r)} \right)$$

36

18   simply compares the length of a uniform line segment (r), used to outline an image, to the size

19   of the shape created by the line segments, N(r) across scaled values of r. With the development

20   of imaging technology, image analysis tools have been implemented to determine the fractal

21   dimension with greater accuracy[29-33] as well as measure the lacunarity[29] — a measure of the

22   'openness' a particular shape has. In place of line segments, image analysis tools (e.g., ImageJ)

23   place uniform boxes on the image and compare the number of boxes total, log(N), to boxes that

24   contain pixels, $\log(L/L_0)$, across scaled box size values, where L is the size of the box at each

25   iteration and L_0 is the size of the largest box size in the image. The slope of this relationship gives

26   an accurate measure of the dimensionality of the imaged object, D_f.

27

28   For comparing HIM (main text Figure 3) or TEM (Fig. S22) images to our simulated fractal

29   assemblies, we derived the fractal dimension (slope) and scalability (linear range) from 2D

30   image analysis using ImageJ. When comparing the Cryo-ET data to the computational

31   simulation results, projections were made to be analyzed with the same 2D image analysis.

32   Additionally, 3D-fractal dimension analysis was performed with an in-house 3D box-counting

33   algorithm that works in the same way that 2D image analysis does except the two-dimensional

34   boxes are replaced with three-dimensional cubes (voxels) during the scaling analysis.

Comparison of control (GS-rich-linker containing) and designed assembly topologies

38   Although we could differentiate the density of the fractal assemblies at all scales from μm to nm

scale (Fig. S22A and Fig. S22C, respectively), the globular (GS-rich-linker containing) assemblies varied too greatly in structure topology across samples for meaningful analysis—the majority of these images were dominated by dark shadowy particles too dense to obtain meaningful assignments of density to individual protein components (Fig. S22B). However, a few images (<10%) from the GS-linker rich set had small resolvable nm-scale regions where density could be interpreted and assigned to individual protein components (Fig. S22D). For these images. we compared the average monomer-monomer distance across 5 control (GS-rich) and 5 fractal-shaped assemblies (Fig. S31) on the nm-scale. In the fractal-shaped assemblies the inter-monomer distance is tightly clustered (134 ± 2 Å) among images of large (>25 nm size) assemblies (~40% of the set), suggesting uniformity of inter-component connections in agreement with the design conception. In contrast, in the resolvable parts of the control assembly tomograms (<10% of the entire imaged sample), we see three different types of structures: dispersed assembly (inter-monomer distance ~157Å), fractal-similar assemblies (~134Å), and densely packed globular ball-like structures (~125 Å). The robust catalytic activity of the control assembly (Fig. S30) demonstrates that the observed topologies in the control tomograms are not the result of protein unfolding but are in fact, mediated by the engineered SH2 domain-pY peptide interactions.

We note that in our analyses, our fractals formed by the same components can vary in shape and dimension from island to island on surfaces as well as in solution. However, despite inter-island variations, every island or fractal in solution is self-similar (with the same fractal dimension) from a few protein connections to micron-sized particle scales. Similar topological diversity was also found in studies of silk protein sericin[33], where variation in fractal dimension of observed protein islands was detected depending on the surface conditions but each island was self-similar.

Supplementary Information References

1. Dantas, G., Kuhlman, B., Callender, D., Wong, M. & Baker, D. A large scale test of computational protein design: Folding and stability of nine completely redesigned globular proteins. J. Mol. Biol. 332, 449–460 (2003).

2. Tinberg, C. E. et al. Computational design of ligand-binding proteins with high affinity and selectivity. Nature 501, 212–6 (2013).

3. Meiler, J. & Baker, D. ROSETTALIGAND: Protein-small molecule docking with full side-chain flexibility. Proteins Struct. Funct. Genet. 65, 538–548 (2006).

4. Tyka, M. D. et al. Alternate states of proteins revealed by detailed energy landscape mapping. J. Mol. Biol. 405, 607–618 (2011).

5. Zanghellini, A. et al. New algorithms and an in silico benchmark for computational enzyme design. Protein Sci. 15, 2785–2794 (2006).

6. Berman, H. M. et al. The protein data bank. Nucleic Acids Res. 28, 235–242 (2000).

7. Mandell, D. J. & Kortemme, T. Backbone flexibility in computational protein design. Current Opinion in Biotechnology 20, 420–428 (2009).

8. Fleishman, S. J. et al. Rosettascripts: A scripting language interface to the Rosetta Macromolecular modeling suite. PLoS One 6, (2011).

9. Karperien, A. FracLac for ImageJ. User's Guide for FracLac, V. 2.5 (2013). doi:10.13140/2.1.4775.8402

10. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with imageJ. Biophotonics International 11, 36–41 (2004).

11. Mandelbaum, R. T., Allan, D. l. & Wackett, L. P. Isolation and Characterization of a Pseudomonas sp . That Mineralizes the s -Triazine Herbicide Atrazine. Appl. Environ. Microbiol. 61, 1451–1457 (1995).

12. Seffernick, J. L., Johnson, G., Sadowsky, M. J. & Wackett, L. P. Substrate specificity of atrazine chlorohydrolase and atrazine-catabolizing bacteria. Appl. Environ. Microbiol. 66, 4247–4252 (2000).

13. Seffernick, J. L. et al. Hydroxyatrazine N-ethylaminohydrolase (AtzB): An amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. J. Bacteriol. 189, 6989–6997 (2007).

14. Shapir, N., Osborne, J. P., Johnson, G., Sadowsky, M. J. & Wackett, L. P. Purification, substrate range, and metal center of AtzC: The N-isopropylammelide aminohydrolase involved in bacterial atrazine metabolism. J. Bacteriol. 184, 5376–5384 (2002).

15. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–5 (2009).

16. Kaneko, T. et al. Superbinder SH2 Domains Act as Antagonists of Cell Signaling. Sci. Signal. 5, ra68-ra68 (2012).

17. Shapir, N. et al. TrzN from Arthrobacter aurescens TC1 is a zinc amidohydrolase. J. Bacteriol. 188, 5859–5864 (2006).

18. Parton, D. L. et al. An open library of human kinase domain constructs for automated bacterial expression. bioRxiv (2016).

19. Cameron, S. M., Durchschein, K., Richman, J. E., Sadowsky, M. J. & Wackett, L. P. New family of biuret hydrolases involved in s -triazine ring metabolism. ACS Catal. 1, 1075–1082 (2011).

20. Li, Q., Seffernick, J. L., Sadowsky, M. J. & Wackett, L. P. Thermostable cyanuric acid hydrolase from Moorella thermoacetica ATCC 39073. Appl. Environ. Microbiol. 75, 6986–6991 (2009).

21. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. Open Phys. 10, 181–188 (2012).

22. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. 116, 71–6 (1996).

23. Pettersen, E. F. et al. UCSF Chimera - A visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).

24. Mutlu, B. R., Yeom, S., Wackett, L. P. & Aksan, A. Modelling and optimization of a bioremediation system utilizing silica gel encapsulated whole-cell biocatalyst. Chem. Eng. J. 259, 574–580 (2015).

25. Hausdorff, F. Dimension und äußeres Maß. Math. Ann. 79, 157–179 (1919).

26. Meakin, P. Formation of fractal clusters and networks by irreversible diffusion-limited aggregation. Phys. Rev. Lett. 51, 1119–1122 (1983).

27. Meakin, P. Diffusion-controlled cluster formation in 2-6 dimensional space. Phys. Rev. A 27, 1495–1507 (1983).

28. Lomander, A., Hwang, W. & Zhang, S. Hierarchical self-assembly of a coiled-coil peptide into fractal structure. Nano Lett. 5, 1255–1260 (2005).

29. Kirkby, M. J. The fractal geometry of nature. Benoit B. Mandelbrot. W. H. Freeman and co., San Francisco, 1982. No. of pages: 460. Price: £22.75 (hardback). Earth Surf. Process. Landforms 8, 406–406 (1983).

30. Smith, T. G., Lange, G. D. & Marks, W. B. Fractal methods and results in cellular morphology - Dimensions, lacunarity and multifractals. Journal of Neuroscience Methods 69, 123–136 (1996).

31. Fairbanks, M. S., McCarthy, D. N., Scott, S. A., Brown, S. A. & Taylor, R. P. Fractal electronic devices: Simulation and implementation. Nanotechnology 22, (2011).

32. Murr, M. M. & Morse, D. E. Fractal intermediates in the self-assembly of silicatein filaments. Proc. Natl. Acad. Sci. 102, 11657–11662 (2005).

33. Khire, T. S., Kundu, J., Kundu, S. C. & Yadavalli, V. K. The fractal self-assembly of the silk protein sericin. Soft Matter 6, 2066 (2010).

1

2

3

4

5

6
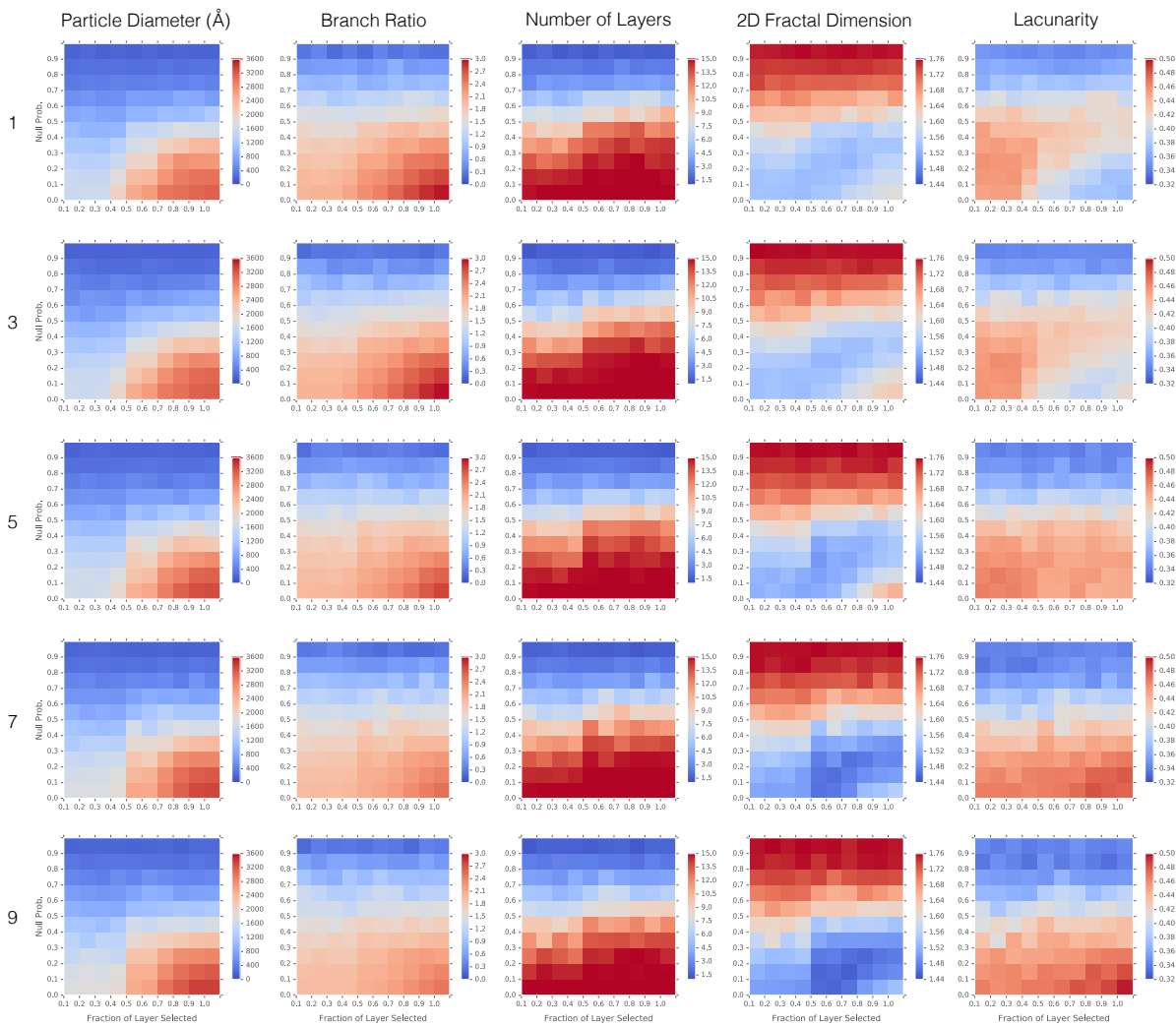
7

8

9

10

11

12

13

Fig. S1. Computational parameter sweep of kT (major y-axis), P_{null} (minor y-axis), and C_{frac} (minor x-axis). The various fractal topologies (limited to 15 layers) were evaluated by their particle diameter, branch ratio, layer count, 2D fractal dimension (D_f), and Lacunarity. We observe size, shape, and composition trends with varying P_{null} and C_{frac}. Less obvious trends in topology via lacunarity and D_f are also observed with changing kT. P_{null} values above 0.4 (0.5-0.9) and C_{frac} values below 0.5 (0.0-0.4) show a steep decline in particle size and number of total layers on average—terminating growth during simulation (unlike experimental data). For non-terminating values of P_{null} (0.0-0.4) and C_{frac} (0.5-1.0), D_f is high (~1.7) when the connection probability is high—more isotropic fractal—and low (~1.6) when the connection probability is low—more anisotropic fractal shapes. When the kT increases we notice that the relative difference between high and low connection probability is maintained, however, the overall D_f decreases (~1.6 and ~1.5) respectively. This can be attributed to the flatter probability landscape allowing for more 180° bound-angle (mixed vertex and edge centered connections around AtzA)—linearizing the branch connections on average and subsequently decreasing the fractal dimension.
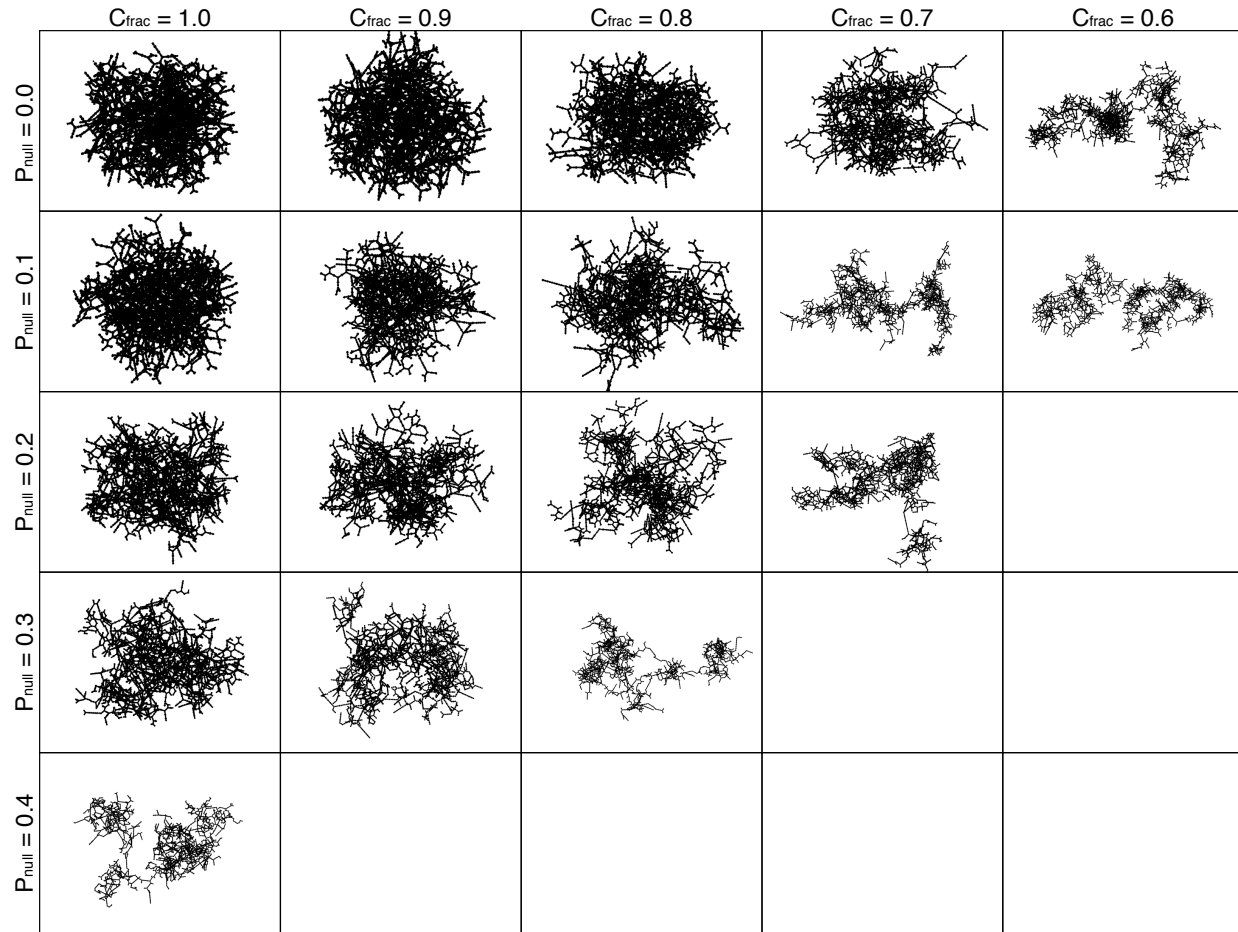
Fig. S2. Representative simulated fractal images (approx. 5000 components each and $kT = 9$) that possess the average layer count and branch ratio for varying values of P_{null} (y-axis) and C_{frac} (x-axis) of 100 models. We observe qualitatively, the number of layers and branch ratio decreases on average as the connection probability decreases. These results are qualitatively similar to varying concentrations of [pY-AtzA] in Figure 3E.
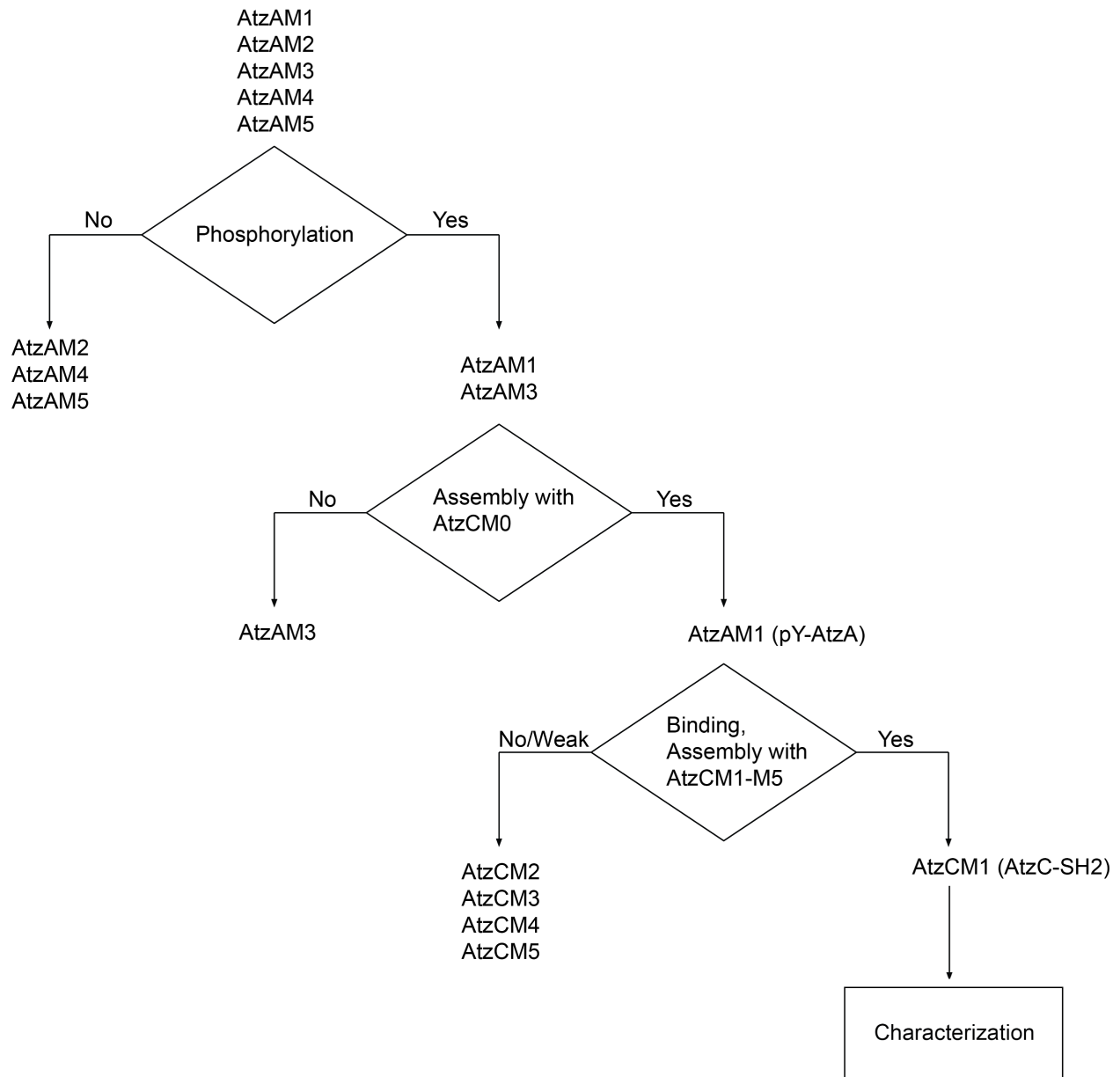
1
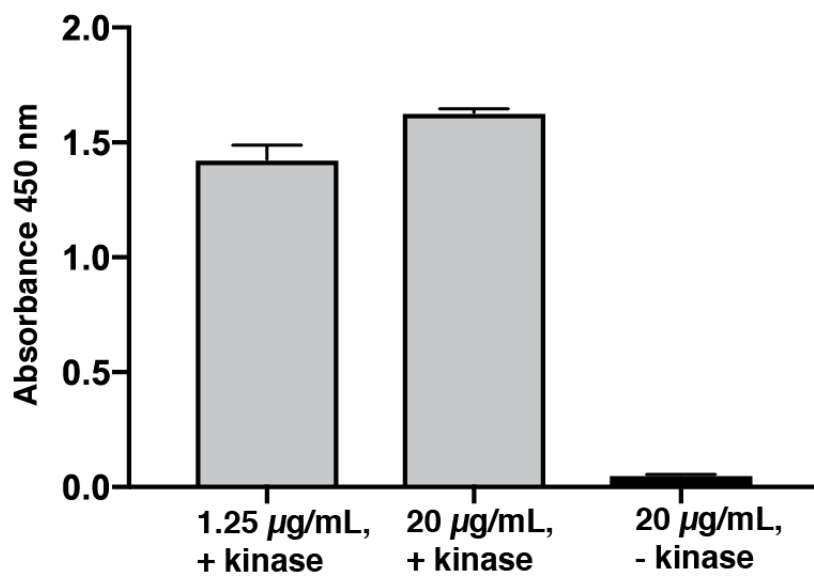
Fig. S3. Phosphorylation of SH2 peptide AtzA fusion (pY-AtzA) by Src kinase.
In order to verify phosphorylation of AtzA by Src kinase into phosphorylated SH2 peptide AtzA fusion (pY-AtzA), ELISA with
(1:4000 dilution) antiphosphotyrosine-horseradish peroxidase conjugate was performed on pY-AtzA samples either with Src
kinase (+) or without Src kinase (-), in phosphorylation reaction buffer at 1.25 μg/mL pY-AtzA or 20 μg/mL pY-AtzA. Data is
presented as mean ± 1 standard deviation.

AtzAM1
AtzAM2
AtzAM3
AtzAM4
AtzAM5

No ← **Phosphorylation** → Yes

AtzAM2
AtzAM4
AtzAM5

AtzAM1
AtzAM3

No ← **Assembly with AtzCM0** → Yes

AtzAM3

AtzAM1 (pY-AtzA)

No/Weak ← **Binding, Assembly with AtzCM1-M5** → Yes

AtzCM2
AtzCM3
AtzCM4
AtzCM5

AtzCM1 (AtzC-SH2)

Characterization

1

2 **Fig. S4. Experimental selection process for pY-AtzA and AtzC-SH2.** Five N-terminal SH2 binding peptide AtzA fusions
3 (AtzAM1-AtzAM5) and five C-terminal SH2 binding domain AtzC fusions (AtzCM1-AtzCM5) were selected, cloned,
4 expressed, and purified. AtzAM1-M5 were screened for having the ability to be phosphorylated via ELISA with anti-
5 phosphotyrosine. Only two AtzA designs, AtzAM1 and AtzAM3, showed strong phosphorylation. The ability for assembly
6 formation to occur with a direct C-terminal SH2 binding domain AtzC fusion (no mutations; AtzCM0) was used to select the best
7 AtzA design. AtzAM1 was chosen for superior assembly formation ability, becoming pY-AtzA. The five AtzC designs AtzCM1-
8 AtzCM5 were screened for the ability to effectively bind and assemble with pY-AtzA. The combination of pY-AtzA and
9 AtzCM1 (which we call AtzC-SH2) showed the strongest binding and the most robust assembly formation. This pair was then
10 chosen for further characterization.
11

1
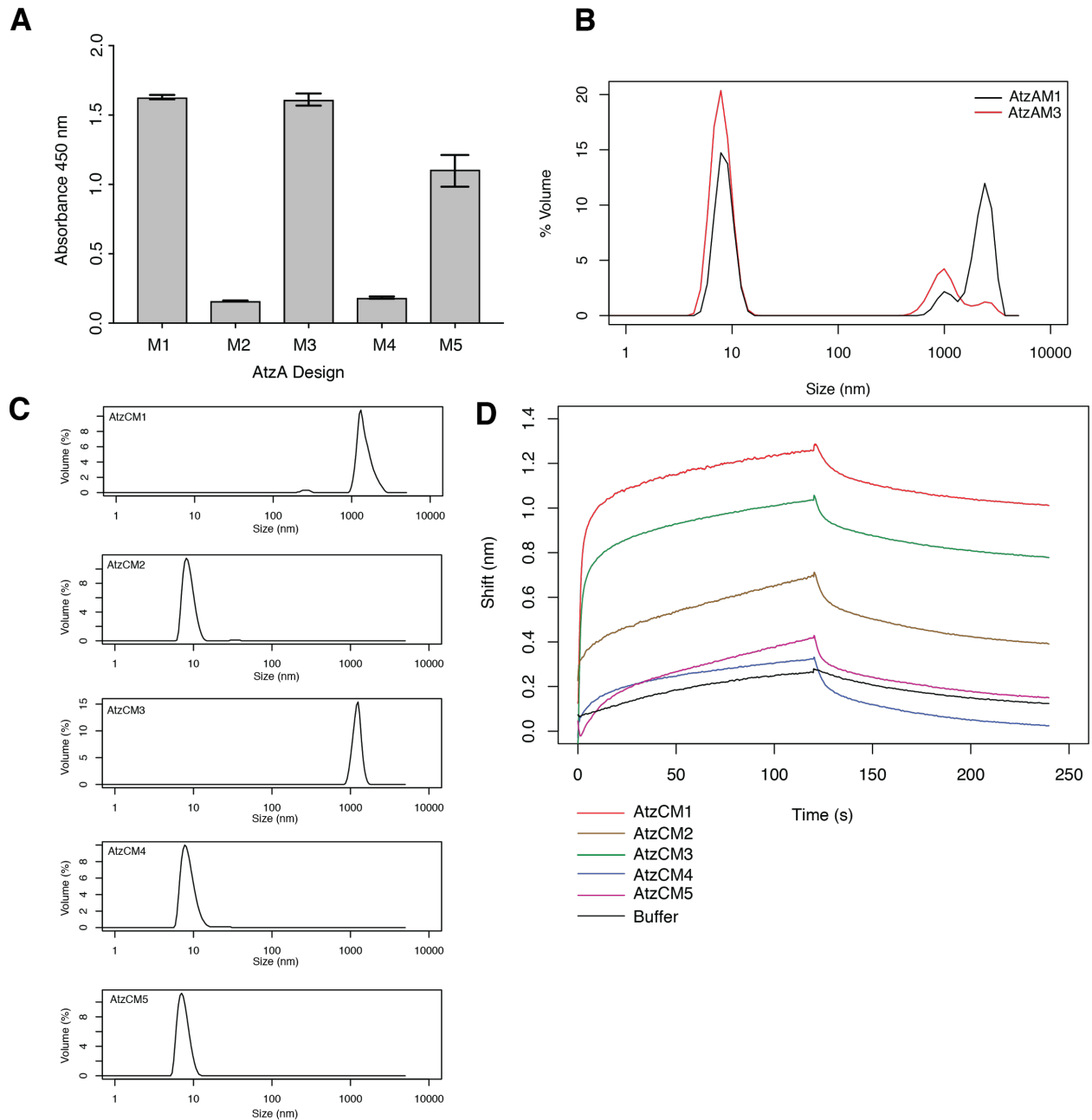
Fig. S5. Experimental selection of AtzA, AtzC subunits for characterization. (**A**) ELISA screening of AtzA designs. (**B**) DLS size distribution of AtzA designs with AtzCM0. (**C**) DLS size distribution of AtzC-SH2 designs with pY-AtzA. Samples prepared at 3 μM pY-AtzA, 2 μM AtzC-SH2 design. Only AtzCM1 and AtzCM3 showed assembly formation with pY-AtzA. Volume distribution reported. (**D**) BLI binding traces of AtzC-SH2 designs with pY-AtzA. AtzC-SH2 designs were screened for binding with BLI, using pY-AtzA as the load. Out of all AtzC-SH2 designs prepared, AtzCM1 had the highest binding affinity to pY-AtzA. Based on the assembly formation and binding data, AtzCM1 was chosen for further investigation.

8

24

A

Shift (nm)

— 4 μM AtzC-wtSH2
— 1 μM AtzC-wtSH2
— 0.25 μM AtzC-wtSH2
— 0.0625 μM AtzC-wtSH2
— 0.0040 μM AtzC-wtSH2

B

Shift (nm)

— 4 μM AtzCM1
— 1 μM AtzCM1
— 0.25 μM AtzCM1
— 0.0625 μM AtzCM1

1
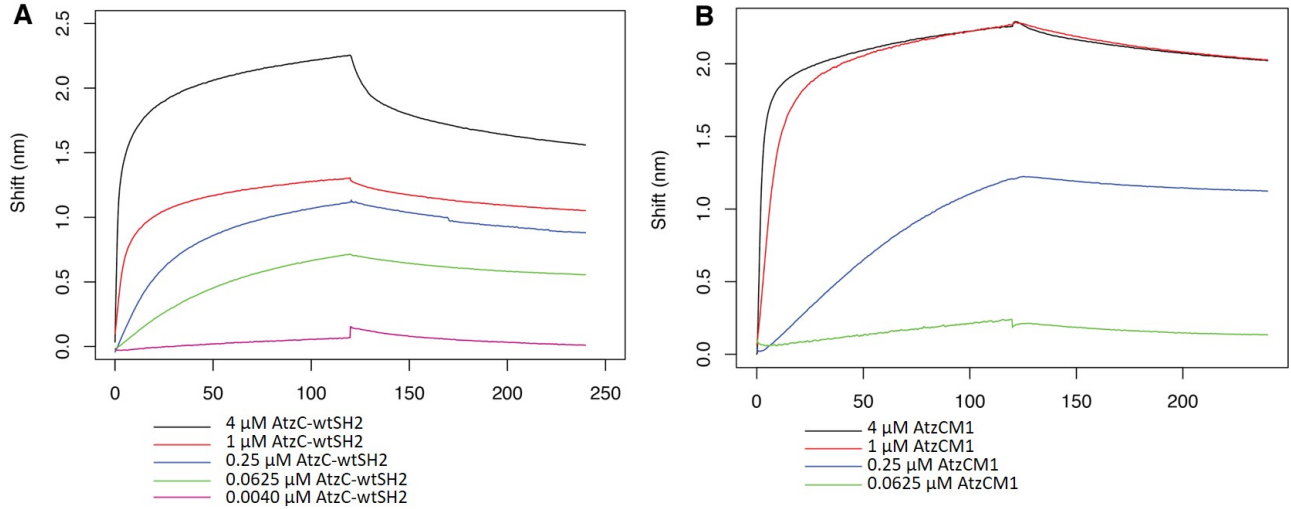
Fig. S6. Biolayer interferometry (BLI) binding profiles of AtzC wildtype SH2 fusion (AtzC-wtSH2) and AtzC superbinder SH2 fusion (AtzC-SH2) to phosphorylated SH2 binding peptide AtzA fusion (pY-AtzA). (**A**) Binding profile of AtzC-wtSH2 to pY-AtzA. PY-AtzA was loaded onto the biosensor via a streptavidin-biotin interaction. AtzC-wtSH2 was flowed into the sample. KD = 41.79 ± 0.32 nM. (**B**) Binding profile of AtzCM1 (superbinder) to pY-AtzA. PY-AtzA was loaded onto the biosensor via a streptavidin-biotin interaction. AtzC-SH2 was flowed into the sample. KD = 7.67 ± 0.52 nM.

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

```
AtzCM0  MSKDFDLIIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVKDEIDAKGNLVSPGFVDAH  60
AtzCM1  MSKDFDLIIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVKDEIDAKGNLVSPGFVDAH  60
AtzCM2  MSKDFDLIIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVKDEIDAKGNLVSPGFVDAH  60
AtzCM3  MSKDFDLIIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVGDEIDAKGNLVSPGFVDAH  60
AtzCM4  MSKDFDLIIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVKDEIDAKGNLVSPGFVDAH  60
AtzCM5  MSKDFDLIIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVKDEIDAKGNLVSPGFVDAH  60

AtzCM0  THMDKSFTSTGERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTL  120
AtzCM1  THMDKSFTSTGERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTL  120
AtzCM2  THMDKSFTSTGERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTL  120
AtzCM3  THMDKSFTSTGERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTL  120
AtzCM4  THMDKSFTSTGERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTL  120
AtzCM5  THMDKSFTSTGERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTL  120

AtzCM0  YTRTHVDVDSVAKTKAVEAVLEAKEELKDLIDIQVVAFAQSGFFVDLESESLIRKSLDMG  180
AtzCM1  YTRTHVDVDSVAKTKAVEAVLEAKEELKDLIDIQVVAFAQSGFFVDLESESLIRKSLDMG  180
AtzCM2  YTRTHVDVDSVAKTKAVEAVLEAKEELKDQIDIQVVAFAQSGFFVDLESESLIRKSLDMG  180
AtzCM3  YTRTHVDVDSVAKTKAVEAVLEAKEELKDSIDIQVVAFAQSGFFVDLESESLIRKSLDMG  180
AtzCM4  YTRTHVDVDSVAKTKAVEAVLEAKEELKDLIDIQVVAFAQSGFFVDLESESLIRKSLDMG  180
AtzCM5  YTRTHVDVDSVAKTKAVEAVLEAKEELKDLIDIQVVAFAQSGFFVDLESESLIRKSLDMG  180

AtzCM0  CDLVGGVDPATRENNVEGSLDLCFKLAKEYDVDIDYHIHDIGTVGVYSINRLAQKTIENG  240
AtzCM1  CDLVGGVDPATRENNVEGSLDLCFKLAKEYDVDIDYHIHDIGTVGVYSINRLAQKTIENG  240
AtzCM2  CDLVGGVDPATRENNVEGSLDLCFKLAKEYDVDIDYHIHDIGTVGVYSINRLAQKTIENG  240
AtzCM3  CDLVGGVDPATRENNVEGSLDLCFKLAKEYDVDIDYHIHDIGTVGVYSINRLAQKTIENG  240
AtzCM4  CDLVGGVDPATRENNVEGSLDLCFKLAKEYDVDIDYHIHDIGTVGVYSINRLAQKTIENG  240
AtzCM5  CDLVGGVDPATRENNVEGSLDLCFKLAKEYDVDIDYHIHDIGTVGVYSINRLAQKTIENG  240

AtzCM0  YKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFSSTPPTMPVIKLLEAGINLGC  300
AtzCM1  YKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFSSTPPTMPVIKLLEAGINLGC  300
AtzCM2  YKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFSSTPPTMPVIKLLEAGINLGC  300
AtzCM3  YKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFSSTPPTMPVIKLLEAGINLGC  300
AtzCM4  YKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFSSTPPTMPVIKLLEAGINLGC  300
AtzCM5  YKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFSSTPPTMPVIKLLEAGINLGC  300

AtzCM0  ASDNIRDFWVPFGNGDMVQGALIETQRLELKTNRDLGLIWKMITSEGARVLGIEKNYGIE  360
AtzCM1  ASDNIRDFWVPFGNGDMVQGALIETQRLELKTNRDLGLIWKMITSEGARVLGIEKNYGIE  360
AtzCM2  ASDNIRDFWVPFGNGDMVQGALIETQRLELKTNRDLGLIWKMITSEGARVLGIEKNYGIE  360
AtzCM3  ASDNIRDFWVPFGNGDMVQGALIETQRLELKTNRDLGLIWKMITSEGARVLGIEKNYGIE  360
AtzCM4  ASDNIRDFWVPFGNGDMVQGALIETQRLELKTNRDLGLIWKMITSEGARVLGIEKNYGIE  360
AtzCM5  ASDNIRDFWVPFGNGDMVQGALIETQRLELKTNRDLGLIWKMITSEGARVLGIEKNYGIE  360

AtzCM0  VGKKADLVVLNSLSPQWAIIDQAKRLCVIKNGRIIVKDEVIVASIQAEEWYFGKLGRKDA  420
AtzCM1  VGKKADLVVLNSLSPQWAIIDQAKRLCVIKNGRIIVKDEVIVAGGSAEEWYFGKLGRKDA  420
AtzCM2  VGKKADLVVLNSLSPQWAIIDQAKRLCVIKNGRIIVKDEVIGAGVAAEEWYFGKLGRKDA  420
AtzCM3  VGKKADLVVLNSLSPQWAIIDQAKRLCVIKNGRIIVKDEVIIASGAAEEWYFGKLGRKDA  420
AtzCM4  VGKKADLVVLNSLSPQWAIIDQAKRLCVIKNGAIIVKDEYILAGGSAEEWYFGKLGRKDA  420
AtzCM5  VGKKADLVVLNSLSPQWAIIDQAKRLCVIKNGSICVKDEAIMASGSAEEWYFGKLGRKDA  420
```

Fig. S7a. Sequence alignment of AtzC-SH2 designs AtzCM0-AtzCM1.

```
AtzCM0  ERQLLSFGNPRGTFLIRESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYITTRA  480
AtzCM1  ERQLLSFGNPRGTFLIRESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYITTRA  480
AtzCM2  ERQLLSFGNPRGTFLIRESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYITTRA  480
AtzCM3  ERQLLSFGNPRGTFLIRESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYITTRA  480
AtzCM4  ERQLLSFGNPRGTFLIRESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYITTRA  480
AtzCM5  ERQLLSFGNPRGTFLIRESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYITTRA  480


AtzCM0  QFETLQQLVQHYSERAAGLSSRLVVPSHKLEHHHHHH                          517
AtzCM1  QFETLQQLVQHYSERAAGLSSRLVVPSHKLEHHHHHH                          517
AtzCM2  QFETLQQLVQHYSERAAGLSSRLVVPSHKLEHHHHHH                          517
AtzCM3  QFETLQQLVQHYSERAAGLSSRLVVPSHKLEHHHHHH                          517
AtzCM4  QFETLQQLVQHYSERAAGLSSRLVVPSHKLEHHHHHH                          517
AtzCM5  QFETLQQLVQHYSERAAGLSSRLVVPSHKLEHHHHHH                          517
```

Fig. S7b. Sequence alignment of AtzC-SH2 designs AtzCM0-AtzCM1 (con't).
Sequence alignment of AtzC-SH2 designs prepared. AtzCM0 is a direct fusion of AtzC and superbinder SH2 domain without mutations. Mutations made are highlighted in black or grey (similar residues). The red box highlights the region where the superbinder SH2 domain is located.

```
AtzAM3  MGSSHHHHHHSSGLVPRGSHMEPQYEEIPNYGGLSIQHGTLVTMDQYRRVLGDSWVHVQD  60
AtzAM4  MGSSHHHHHHSSGLVPRGSHMEPQYEEIPDYGGLSIQHGTLVTMDQYRRVLGDSWVHVQD  60
AtzAM2  MGSSHHHHHHSSGLVPRGSHMEPQYEEIPDYGTLSIQHGTLVTMDQYRRVLGDSWVHVQD  60
AtzAM5  MGSSHHHHHHSSGLVPRGSHMEPQYEEIPDYGTLSIQHGTLVTMDQYRRVLGDSWVHVQD  60
AtzAM1  MGSSHHHHHHSSGLVPRGSHMEPQYEEIPYYQTLSIQHGTLVTMDQYRRVLGDSWVHVQD  60

AtzAM3  GRIVALGVHAESVPPPADRVIDARGKVVLPGFINAHTHVNQILLRGGPSHGRQFYDWLFN  120
AtzAM4  GRIVALGVHAESVPPPADRVIDARGKVVLPGFINAHTHVNQILLRGGPSHGRQFYDWLFN  120
AtzAM2  GRIVALGVHAESVPPPADRVIDARGKVVLPGFINAHTHVNQILLRGGPSHGRQFYDWLFN  120
AtzAM5  GRIVALGVHAESVPPPADRVIDARGKVVLPGFINAHTHVNQILLRGGPSHGRQFYDWLFN  120
AtzAM1  GRIVALGVHASSVPPPADRVIDARGKVVLPGFINAHTHVNQILLRGGPSHGRQFYDWLFN  120

AtzAM3  VVYPGQKAMRPEDVAVAVRLYCAEAVRSGITTINENADSAIYPGNIEAAMAVYGEVGVRV  180
AtzAM4  VVYPGQKAMRPEDVAVAVRLYCAEAVRSGITTINENADSAIYPGNIEAAMAVYGEVGVRV  180
AtzAM2  VVYPGQKAMRPEDVAVAVRLYCAEAVRSGITTINENADSAIYPGNIEAAMAVYGEVGVRV  180
AtzAM5  VVYPGQKAMRPEDVAVAVRLYCAEAVRSGITTINENADSAIYPGNIEAAMAVYGEVGVRV  180
AtzAM1  VVYPGQKAMRPEDVAVAVRLYCAEAVRSGITTINENADSAIYPGNIEAAMAVYGEVGVRV  180

AtzAM3  VYARMFFDRMDGRIQGYVDALKARSPQVELCSIMEETAVAKDRITALSDQYHGTAGGRIS  240
AtzAM4  VYARMFFDRMDGRIQGYVDALKARSPQVELCSIMEETAVAKDRITALSDQYHGTAGGRIS  240
AtzAM2  VYARMFFDRMDGRIQGYVDALKARSPQVELCSIMEETAVAKDRITALSDQYHGTAGGRIS  240
AtzAM5  VYARMFFDRMDGRIQGYVDALKARSPQVELCSIMEETAVAKDRITALSDQYHGTAGGRIS  240
AtzAM1  VYARMFFDRMDGRIQGYVDALKARSPQVELCSIMEETAVAKDRITALSDQYHGTAGGRIS  240

AtzAM3  VWPAPATTTAVTVEGMRWAQAFARDRAVMWTLHMAESDHDERIHGMSPAEYMECYGLLDE  300
AtzAM4  VWPAPATTTAVTVEGMRWAQAFARDRAVMWTLHMAESDHDERIHGMSPAEYMECYGLLDE  300
AtzAM2  VWPAPATTTAVTVEGMRWAQAFARDRAVMWTLHMAESDHDERIHGMSPAEYMECYGLLDE  300
AtzAM5  VWPAPATTTAVTVEGMRWAQAFARDRAVMWTLHMAESDHDERIHGMSPAEYMECYGLLDE  300
AtzAM1  VWPAPATTTAVTVEGMRWAQAFARDRAVMWTLHMAESDHDERIHGMSPAEYMECYGLLDE  300

AtzAM3  RLQVAHCVYFDRKDVRLLHRHNVKVASQVVSNAYLGSGVAPVPEMVERGMAVGIGTDNGN  360
AtzAM4  RLQVAHCVYFDRKDVRLLHRHNVKVASQVVSNAYLGSGVAPVPEMVERGMAVGIGTDNGN  360
AtzAM2  RLQVAHCVYFDRKDVRLLHRHNVKVASQVVSNAYLGSGVAPVPEMVERGMAVGIGTDNGN  360
AtzAM5  RLQVAHCVYFDRKDVRLLHRHNVKVASQVVSNAYLGSGVAPVPEMVERGMAVGIGTDNGN  360
AtzAM1  RLQVAHCVYFDRKDVRLLHRHNVKVASQVVSNAYLGSGVAPVPEMVERGMAVGIGTDNGN  360

AtzAM3  SNDSVNMIGDMKFMAHIHRAVHRDADVLTPEKILEMATIDGARSLGMDHEIGSIETGKRA  420
AtzAM4  SNDSVNMIGDMKFMAHIHRAVHRDADVLTPEKILEMATIDGARSLGMDHEIGSIETGKRA  420
AtzAM2  SNDSVNMIGDMKFMAHIHRAVHRDADVLTPEKILEMATIDGARSLGMDHEIGSIETGKRA  420
AtzAM5  SNDSVNMIGDMKFMAHIHRAVHRDADVLTPEKILEMATIDGARSLGMDHEIGSIETGKRA  420
AtzAM1  SNDSVNMIGDMKFMAHIHRAVHRDADVLTPEKILEMATIDGARSLGMDHEIGSIETGKRA  420

AtzAM3  DLILLDLRHPQTTPHHHLAATIVFQAYGNEVDTVLIDGNVVMENRRLSFLPPERELAFLE  480
AtzAM4  DLILLDLRHPQTTPHHHLAATIVFQAYGNEVDTVLIDGNVVMENRRLSFLPPERELAFLE  480
AtzAM2  DLILLDLRHPQTTPHHHLAATIVFQAYGNEVDTVLIDGNVVMENRRLSFLPPEHELAFLE  480
AtzAM5  DLILLDLRHPQTTPHHHLAATIVFQAYGNEVDTVLIDGNVVMENRRLSFLPPEHELAFLE  480
AtzAM1  DLILLDLRHPQTTPHHHLAATIVFQAYGNEVDTVLIDGNVVMENRRLSFLPPERELAFLE  480

AtzAM3  EAQSRATAILQRANMVANPAWRSL  540
AtzAM4  EAQSRATAILQRANMVANPAWRSL  540
AtzAM2  EAQSRATAILQRANMVANPAWRSL  540
AtzAM5  EAQSRATAILQRANMVANPAWRSL  540
AtzAM1  EAQSRATAILQRANMVANPAWRSL  540
```
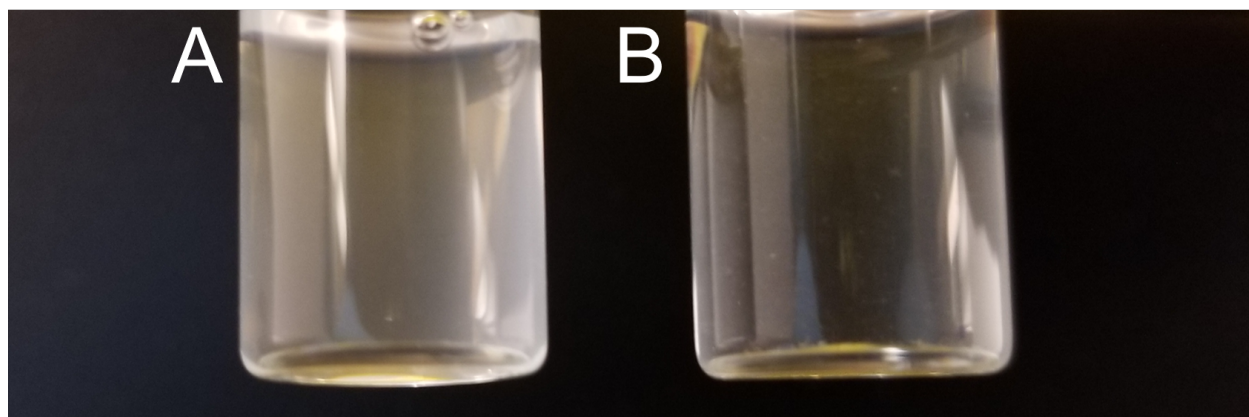
Fig. S8. Sequence alignment of pY-AtzA designs.

Sequence alignment of pY-AtzA designs prepared. AtzAM0 is a direct fusion of AtzA and SH2 binding peptide without mutations. Mutations made are shown in black.

Fig. S9. (**A**) 3 μM pY-AtzAM1 and 2 μM AtzCM1, shows a turbid solution that represents the assembly formed. (**B**) 3 μM non-pY-AtzAM1 and 2 μM AtzCM1, shows a clear solution with no assembly formation.
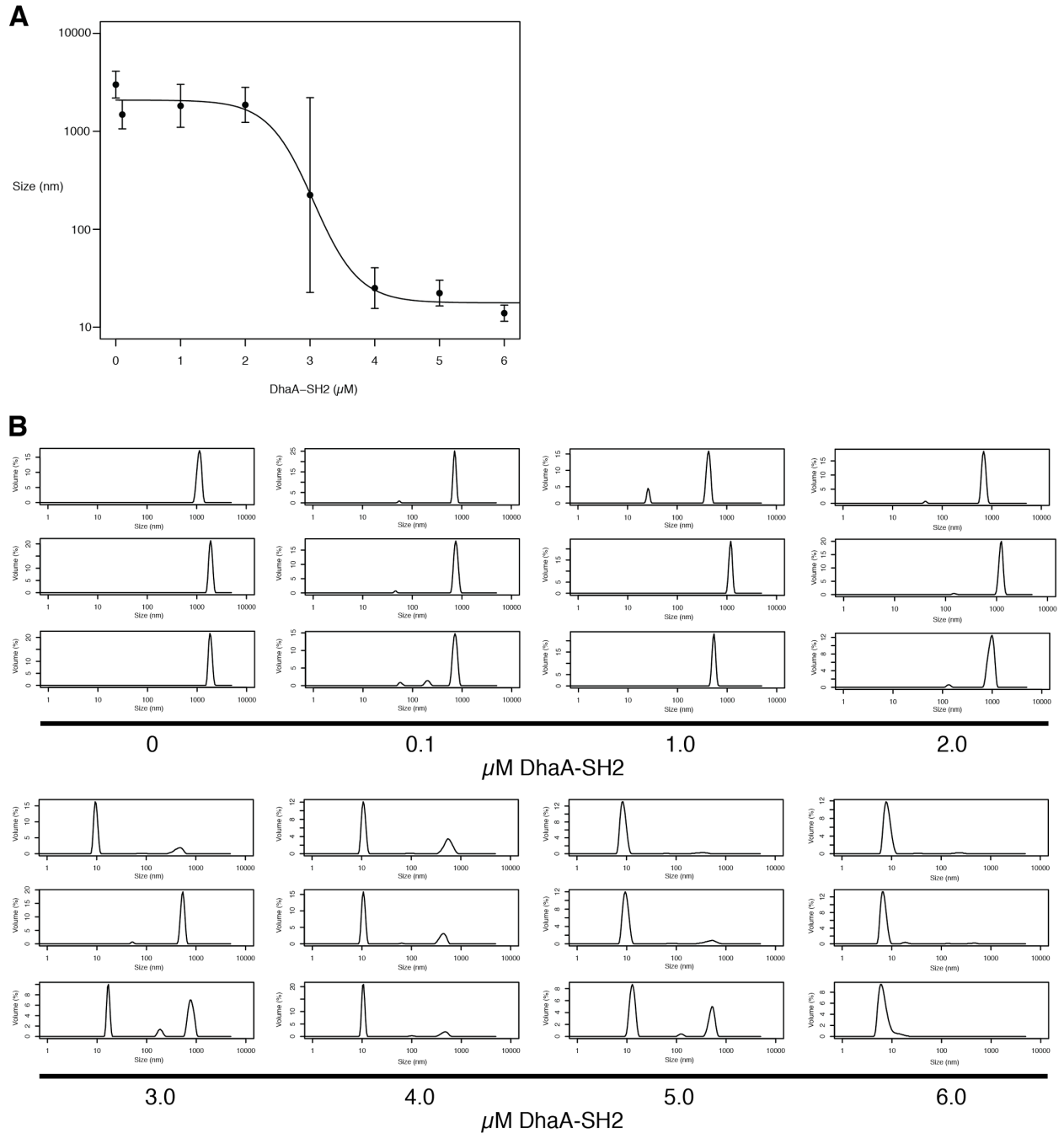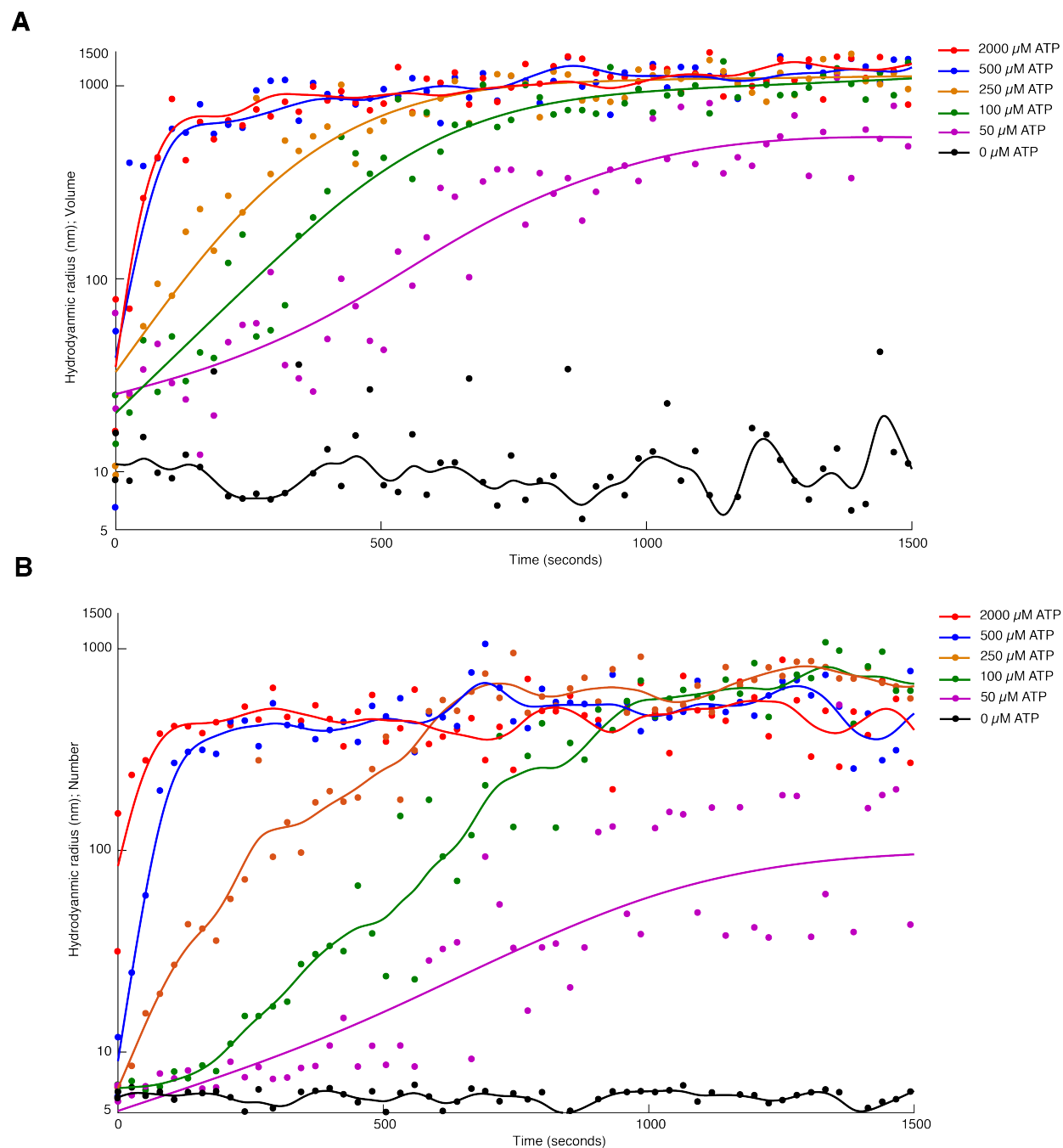
Fig. S10. Inhibition of assembly at 0.66 μM AtzC-SH2, 1 μM pY-AtzA, 0-6 μM SH2-DhaA. (A) Inhibition graph of SH2-DhaA on 0.66 μM AtzC-SH2, 1 μM pY-AtzA assembly. Size recorded represents most predominant DLS sizing peak. Data are presented as mean ± 1 standard deviation. IC50 = 3.05 μM. Adjusted R^2 = 0.98. (B) DLS traces of assembly from 0 - 6 μM SH2-DhaA. DLS traces are of triplicates.

1

2 **Fig. S11. Inhibition of assembly at 2 *µ*M AtzC-SH2, 3 *µ*M pY-AtzA with 0-15 *µ*M inhibitor.**

3

4

C

0 3.0 6.0

*µ*M DhaA-SH2

9.0 12.0 15.0

*µ*M DhaA-SH2

Fig. S12. Inhibition of assembly at 2 *µ*M AtzC-SH2, 3 *µ*M pY-AtzA with 0-15 *µ*M inhibitor (con't).
All DLS traces were performed in triplicate (**A**) Inhibition graph of SH2-DhaA of 2 *µ*M AtzC-SH2, 3 *µ*M pY-AtzA assembly.
Size recorded represents most predominant DLS sizing peak. Data are presented as mean ± 1 standard deviation. IC50 (SH2) =
6.18 *µ*M, IC50 (SH2-DhaA) = 6.13 *µ*M. Adjusted R^2 (SH2) = 0.97. Adjusted R^2 (SH2-DhaA) = 0.99. (**B**) DLS traces of assembly
from 0-15 *µ*M SH2. (**C**) DLS traces of assembly from 0-15 *µ*M SH2-DhaA.

Fig. S13. Rate of assembly formation is dependent on ATP concentration. (**A**) Volume mean of sample from 0 – 1500 sec. Each point represents average of triplicates. (**B**) Number mean of sample from 0 – 1500 sec. Each point represents average of triplicates. Curve fitting performed using sloping spline with smoothness parameter (p) and adjusted R^2 value given in table SX.

Distribution	ATP µM	Adjusted R-square	p
Vol	2000	0.7889	1.31E-05
Vol	500	0.888	1.31E-05
Vol	250	0.898	3.25E-08
Vol	100	0.9374	3.25E-08
Vol	50	0.867	3.25E-08
Vol	0	0.2638	0.000182922
Num	2000	0.5044	2.16E-05
Num	500	0.9303	2.16E-05
Num	250	0.9678	2.16E-05
Num	100	0.9543	2.16E-05
Num	50	0.7189	7.25E-09
Num	0	0.3338	0.000110956

1

2 **Table S1. Curve fitting data for Figure S13.** Adjusted R^2 and smoothing parameter (p) value given for curve fitting done on
3 assembly kinetics data.
4

5

6

7

8

9

10

11
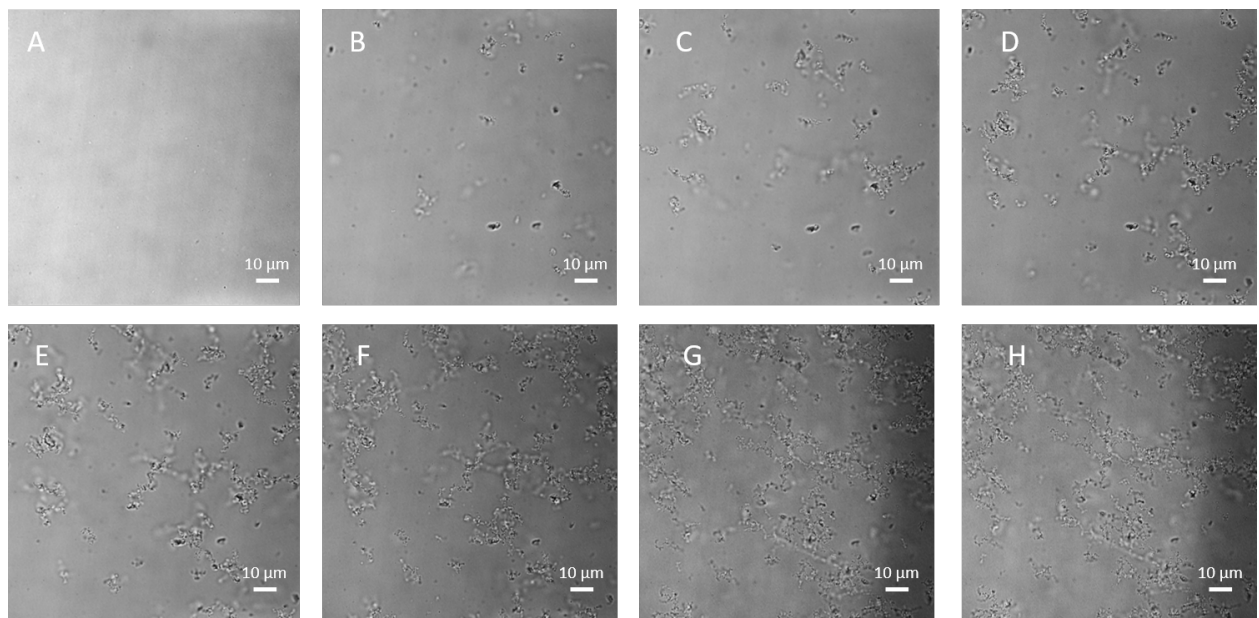
Fig. S14. Bright-field view of the assembly growing after the addition of Src kinase. (**A**) 3 minutes after addition of Src kinase, no assemblies shown. (**B**) 14 minutes after addition of Src kinase, small assemblies shown. (**C**) 18 minutes after addition of Src kinase, small 10 μm assemblies start to grow (**D**) 24 minutes after the addition of Src kinase, growth continues. (**E**) 30 minutes after addition of Src kinase, over 50 μm size assemblies form. (**F**) 35 minutes after addition of Src kinase, 100 μm size assemblies appear. (**G**) 40 minutes after addition of Src kinase, assemblies continue to grow. (**H**) 50 minutes after addition of Src kinase, assemblies have fully matured into fractal-like structures.

1   **Movie S1**

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17
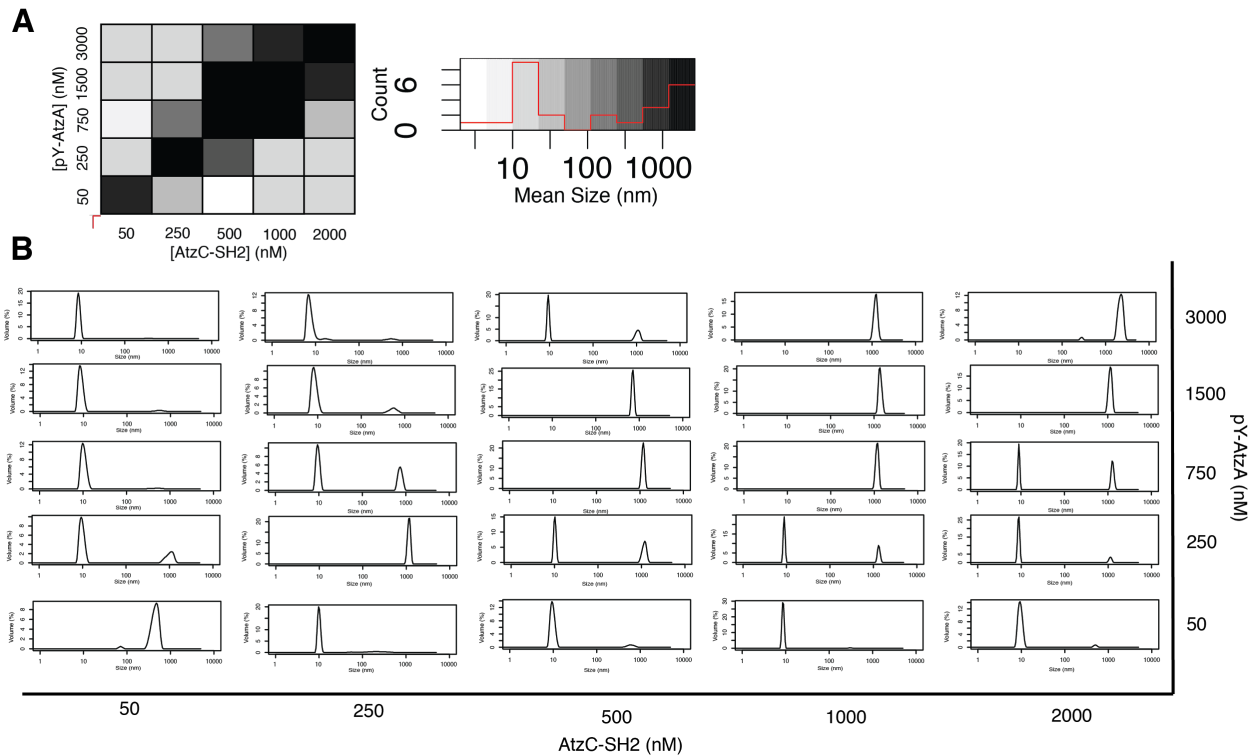
Fig. S15. Average size of particle formed by pY-AtzA and wild type AtzC-SH2. (**A**) Heat map showing volume-weighted mean size of particles found from 50-3000 nM pY-AtzA and 50-2000 nM AtzC-SH2. Value shown is average of two physical samples. Histogram illustrates distribution of sizes found on heatmap. (**B**) Volume distributions of heat map. Distributions shown are representative of other traces in the sample.
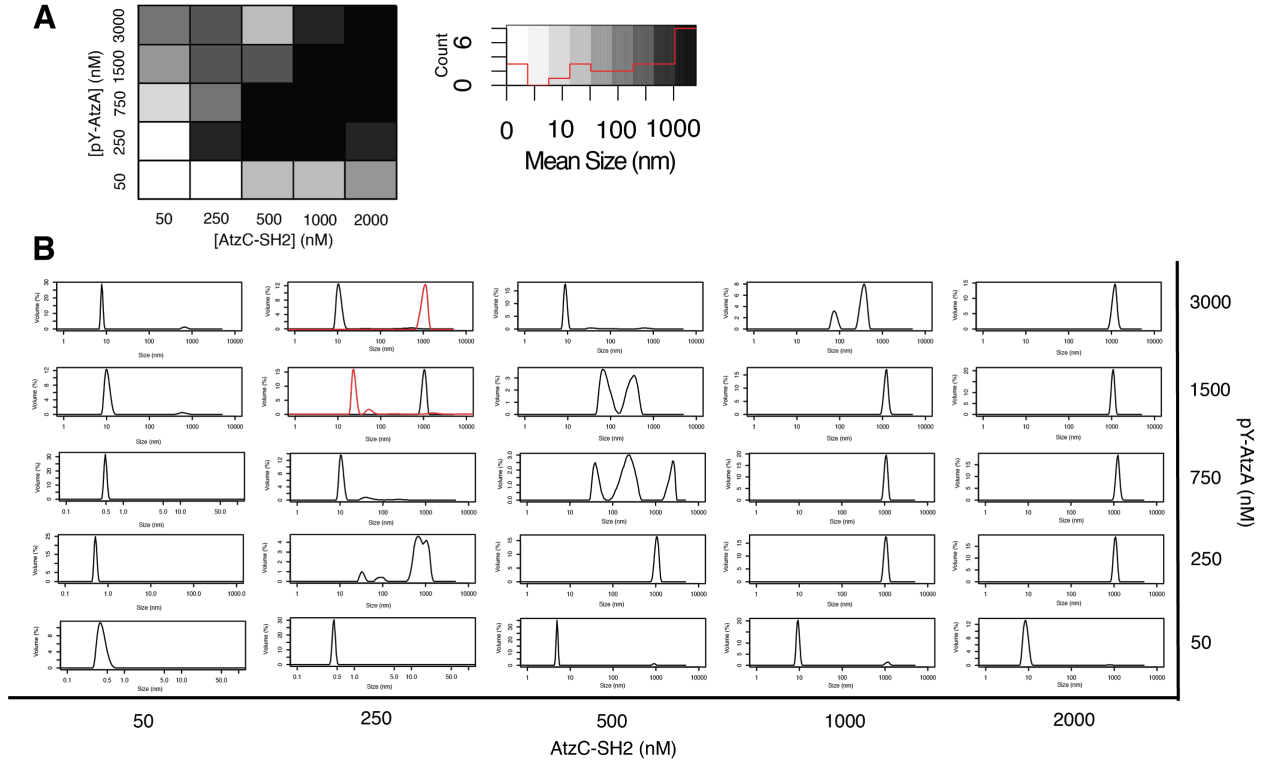
Fig. S16. Average size of particle formed by pY-AtzA and super-binder AtzC-SH2. (**A**) Heat map showing volume-weighted mean size of particles found from 50-3000 nM pY-AtzA and 50-2000 nM AtzC-SH2. Value shown is average of two physical samples. Histogram illustrates distribution of sizes found on heatmap. (**B**) Volume distributions of heat map. Distributions shown are representative of other traces in the sample.
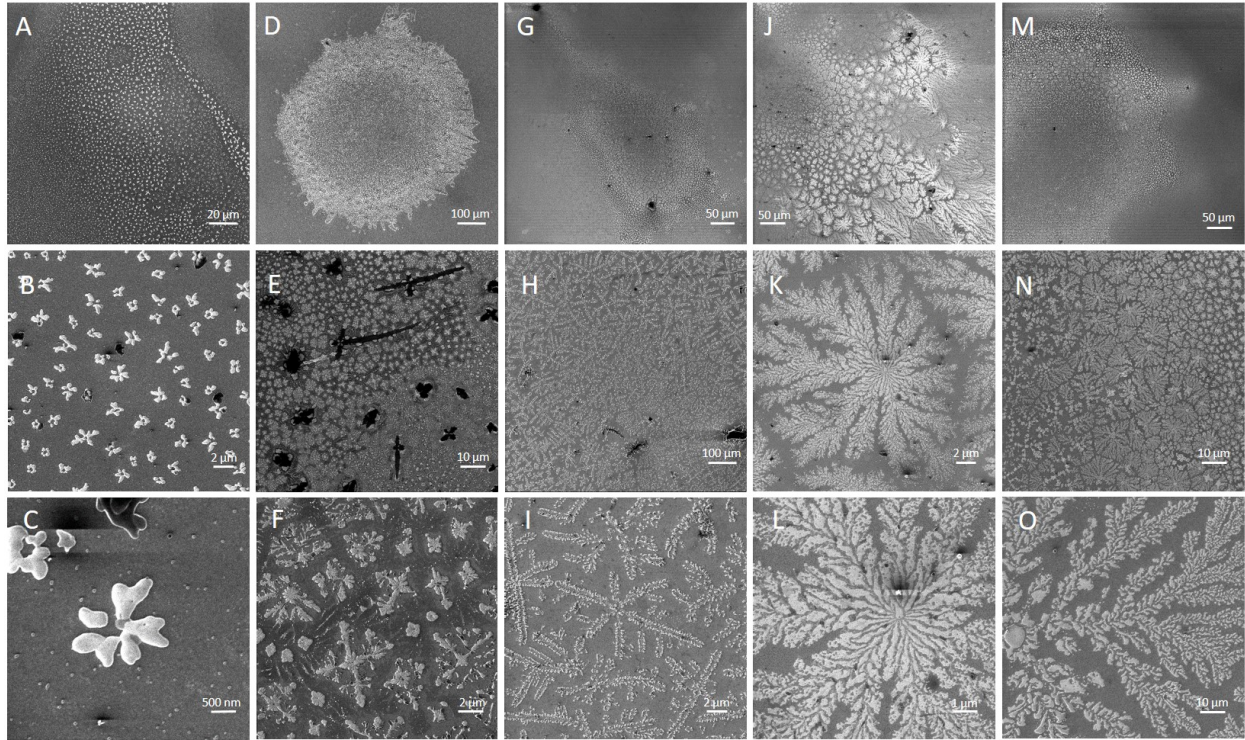
Fig. S17. Helium Ion Microscopy (HIM) depict fractal-like assembly with increasing AtzA concentrations. (**A** to **C**) 0.250 μM AtzAM1 and 2 μM AtzCM1. (**D** to **F**) 0.950 μM AtzAM1 and 2 μM AtzCM1 (**G**-**I**) 1.5 μM AtzAM1 and 2 μM AtzCM1. (**J** to **L**) 3 μM AtzAM1 and 2 μM AtzCM1. (**M** to **O**) 3 μM AtzAM1 and 1 μM AtzCM1.
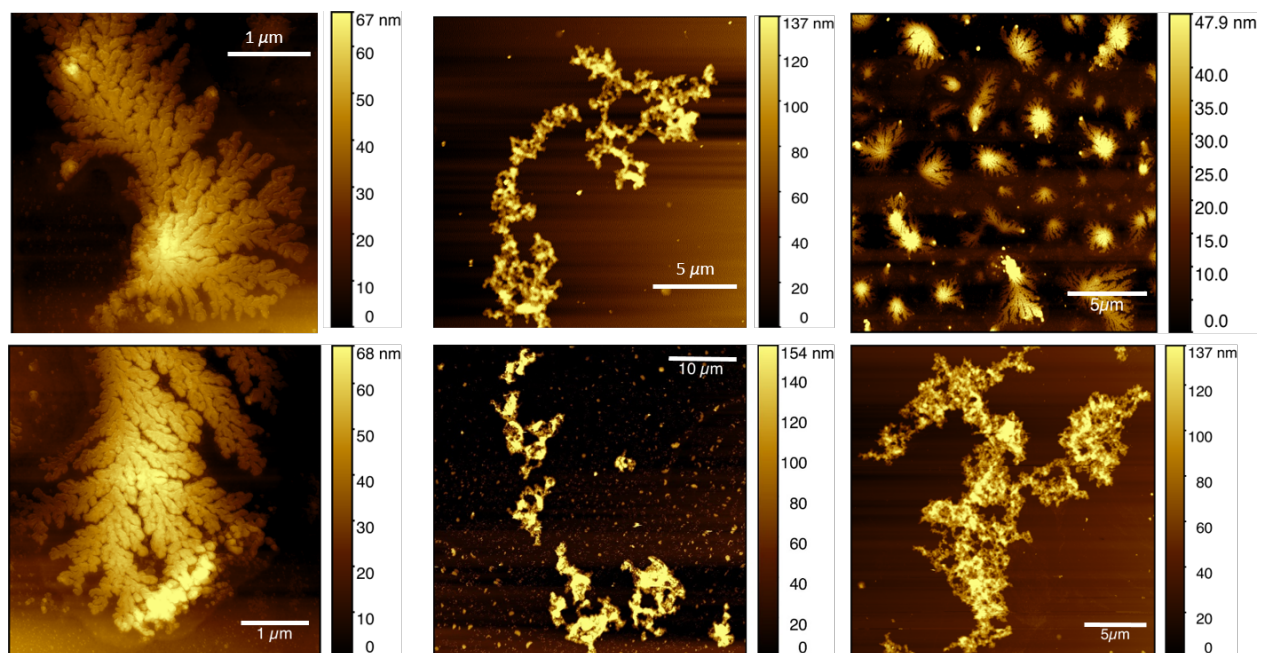
Fig. S18. Atomic Force Microscopy (AFM) images show fractal-like structures, fern-like, and petal-like structures, similar to HIM.
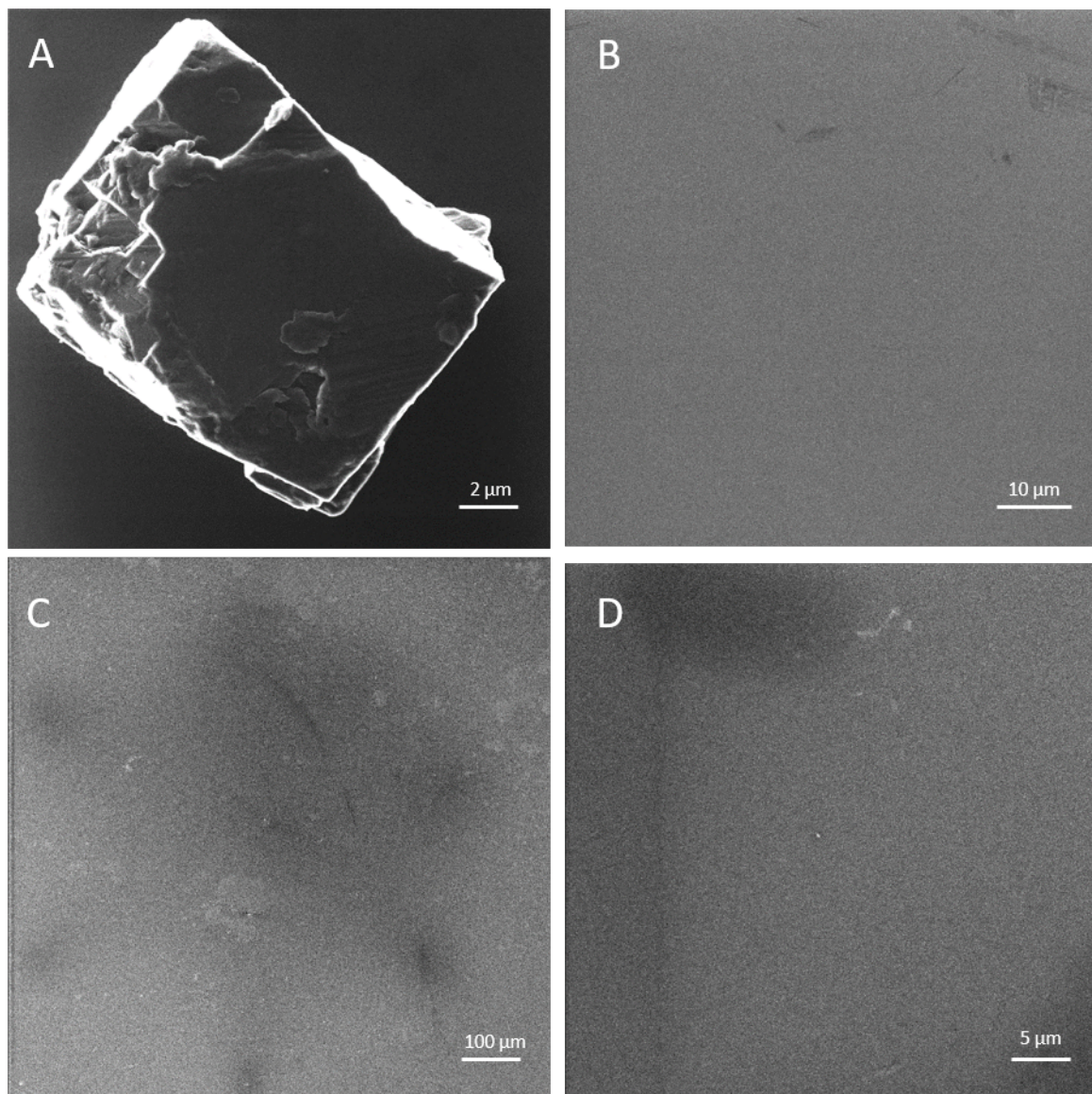
Fig. S19. Helium Ion Microscopy (HIM) buffer and non-phosphorylated controls preclude sale precipitation. In order to determine that our proteins were forming fractal-like patterns and it was not salt inducing the patterns, a buffer and non-phosphorylated proteins sample controls were used to preclude salt precipitation. (**A**) Usual HIM square salt crystals on a glass surface. (**B**) Deposited HNG buffer (50 mM Hepes, 100 mM NaCl, 5% glycerol, pH.7.4, buffer proteins are stored in) on silicon wafer shows no structures on the surface. (**C**) 3 μM non-pY-AtzAM1 and 2 μM AtzCM1 control shows no fractal-like structures. (**D**) 3 μM non-pY-AtzAM1 and 1 μM AtzCM1 show no fractal-like structures. All controls demonstrate that fractal structures are formed by phosphorylated protein components.

1
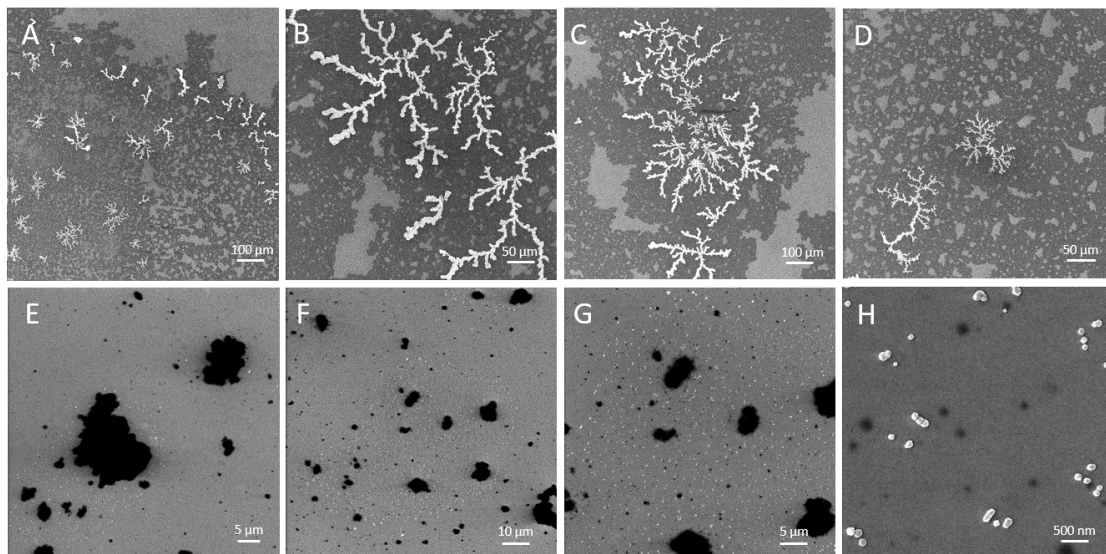
Fig S20. Helium Ion Microscopy comparison of fractal assembly and globular assembly. HIM Images depict fractal-like assembly with 3 uM AtzAM1 and 2 uM AtzCM1 final concentrations (**A** to **D**), while the 3 uM AtzAM1-ExtendedLinker and 2 uM AtzCM1-ExtendedLinker final concentrations show both large and small globular shape proteins on the silicon surface (**E** to **H**).

1

2
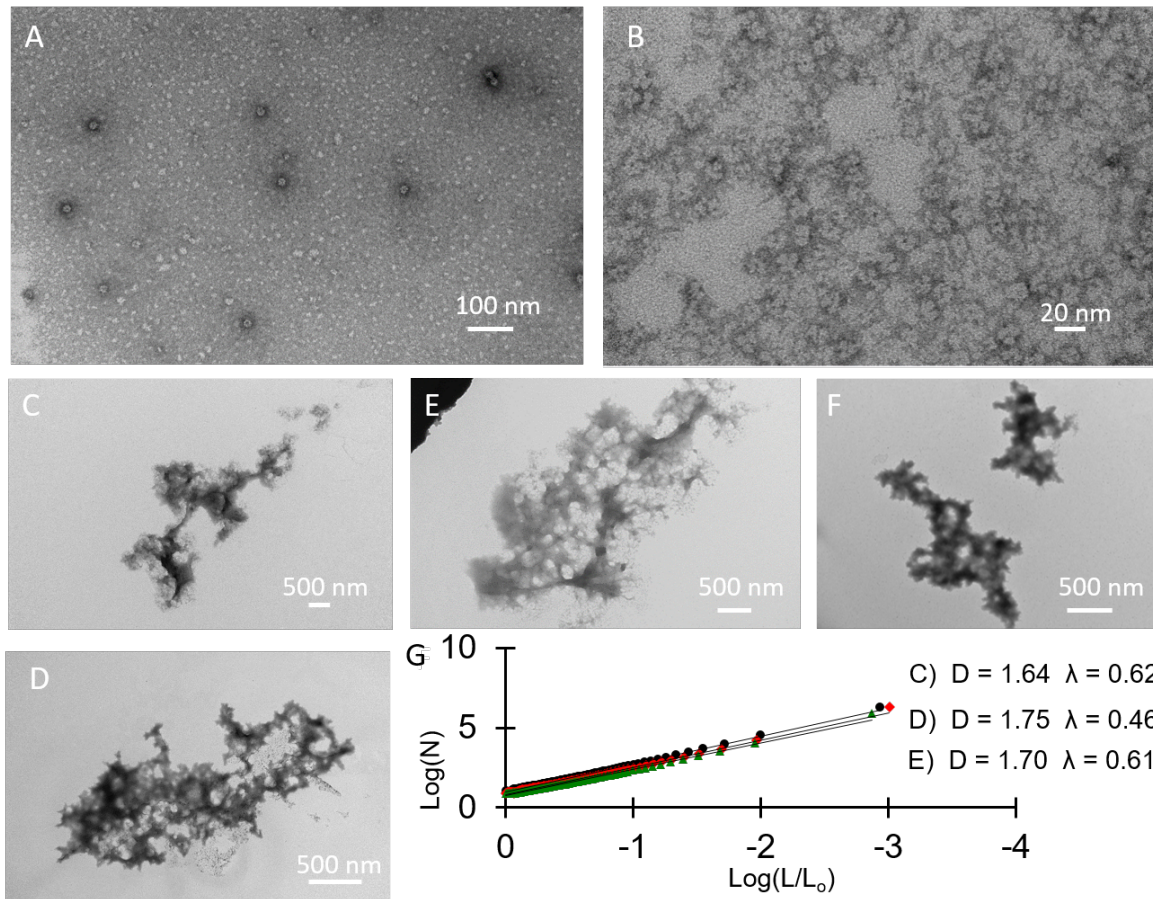
Fig. S21. Transmission Electron Microscopy (TEM) depicts fractal-like assemblies in the phosphorylated samples while the non-phosphorylated samples depict individual proteins. (**A** and **B**) ten-fold dilution of 3 μM non-pY-AtzAM1 and 2 μM AtzCM1, which shows the individual proteins. (**C** to **F**) Various assembly images of the ten-fold dilution of 3 μM pY-AtzAM1 and 2 μM AtzCM1 sample which form the fractal-like assembly consistently. (**G**) Image analysis (2D) using box counting yields the expected fractal dimension of ~1.7 for the C, D, and E, TEM images.
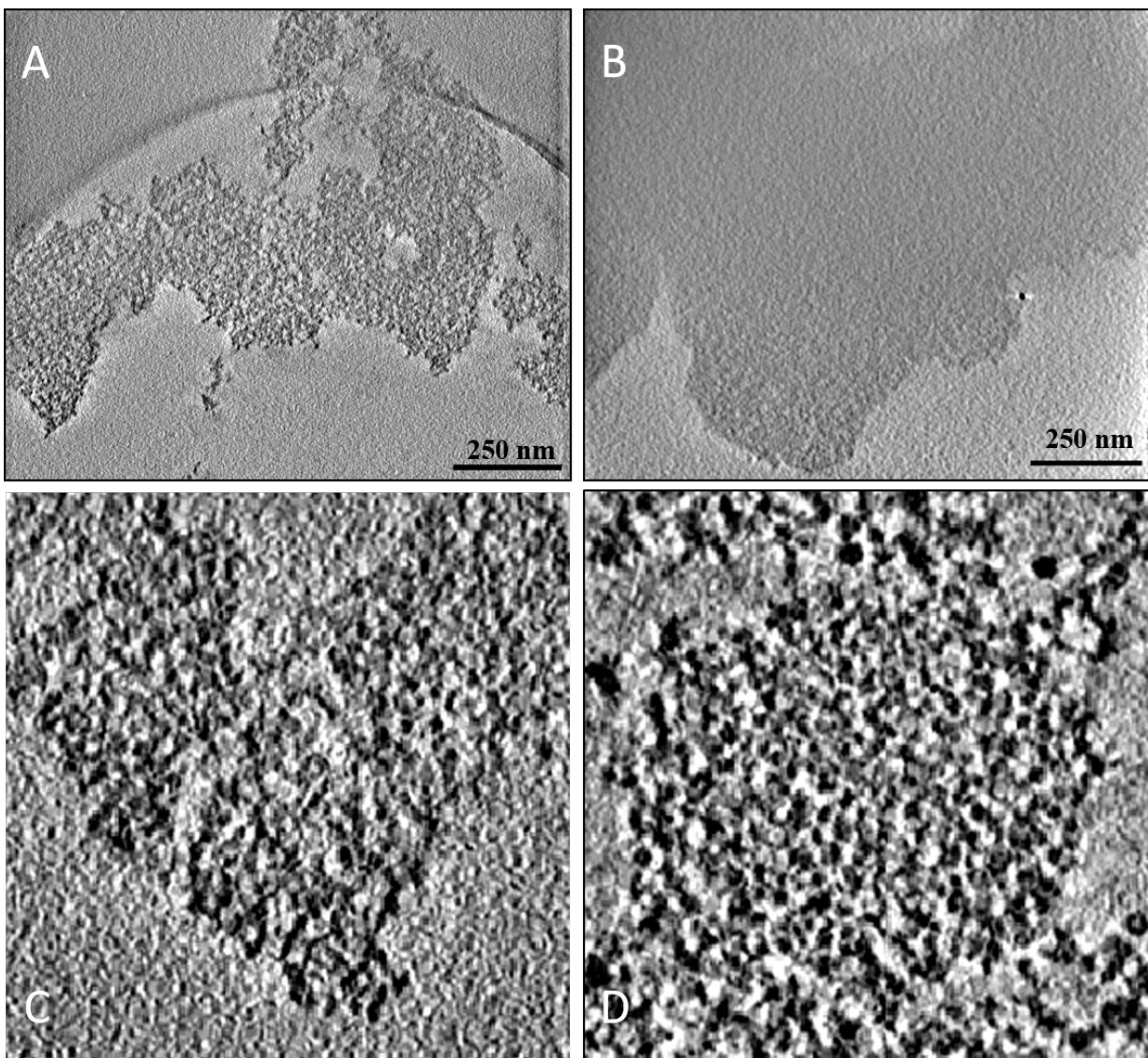
Fig S22. Comparison of the fractal assembly CryoEM tomograms and the extended linker globular assemblies. CryoEM tomograms of the fractal-like assemblies (**A**) and the extended linker assemblies (**B**) show a difference in the overall topology of the two different assemblies. Zoomed in versions of the images show representatives of a fractal assembly (**C**) and of a very dense and globular structure (**D**).

1

2

3   **Movies S2,S3.** CryoET-derived structure videos.

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

1

2

3

4
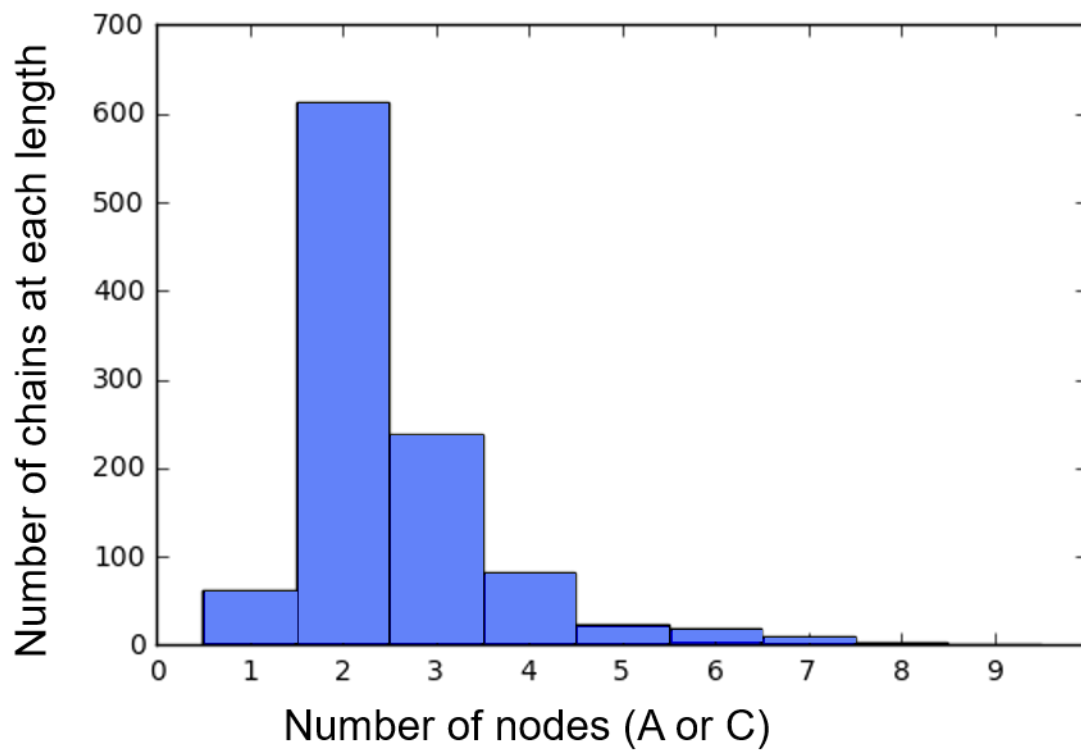
5 **Fig S23.** Length distribution of short chains that are not included in the large assembly.

6

7

8

9

100 nm
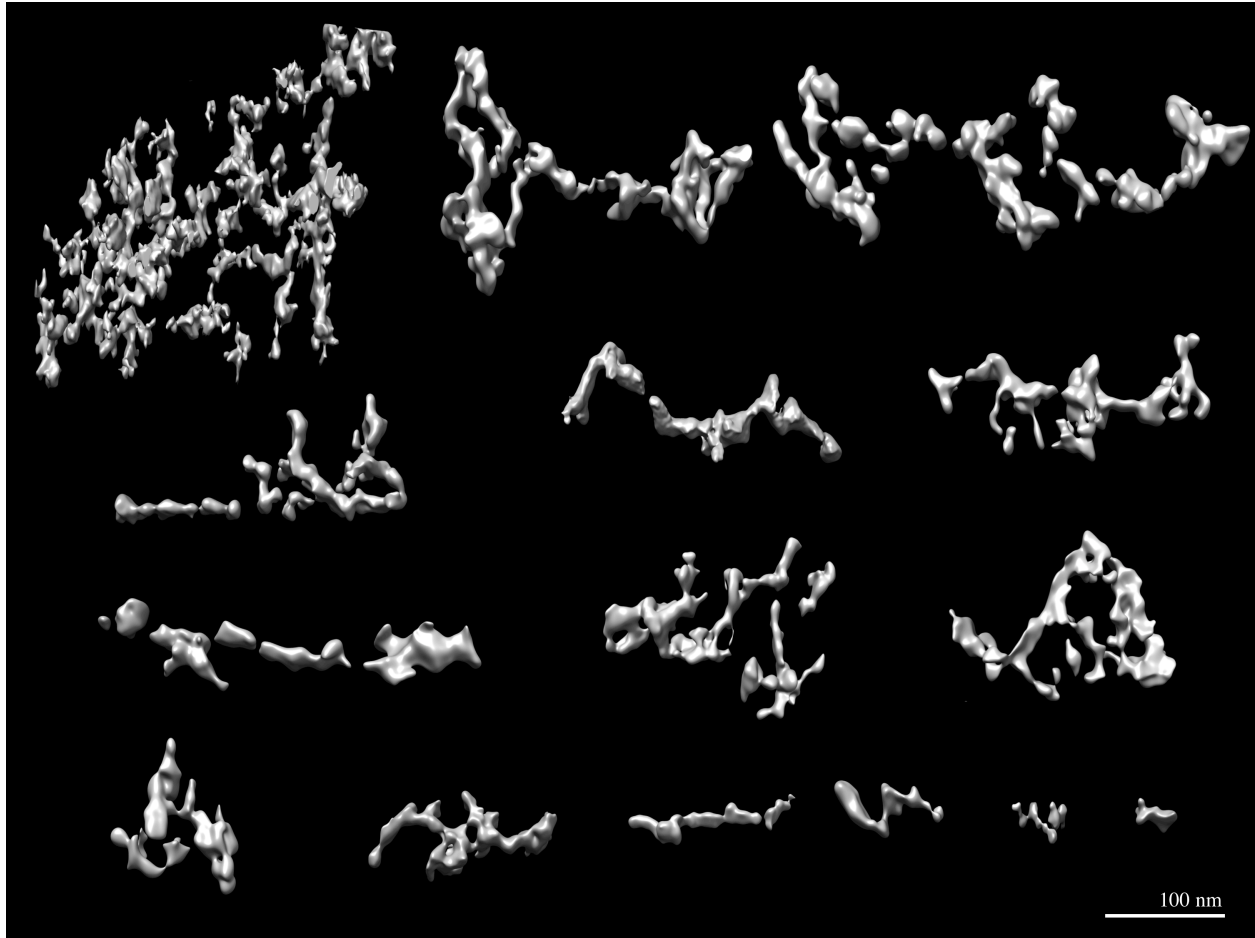
Fig S24. Isosurface views of the assembly tomograms, from large to small.

Sample	AtzAM1 µM	AtzBSH 2µM	AtzCM 1 µM
1	3.0	0.5	1.5
2	3.0	1.0	1.0
3	3.0	1.5	0.5
4	3.0	1.0	2.0

1
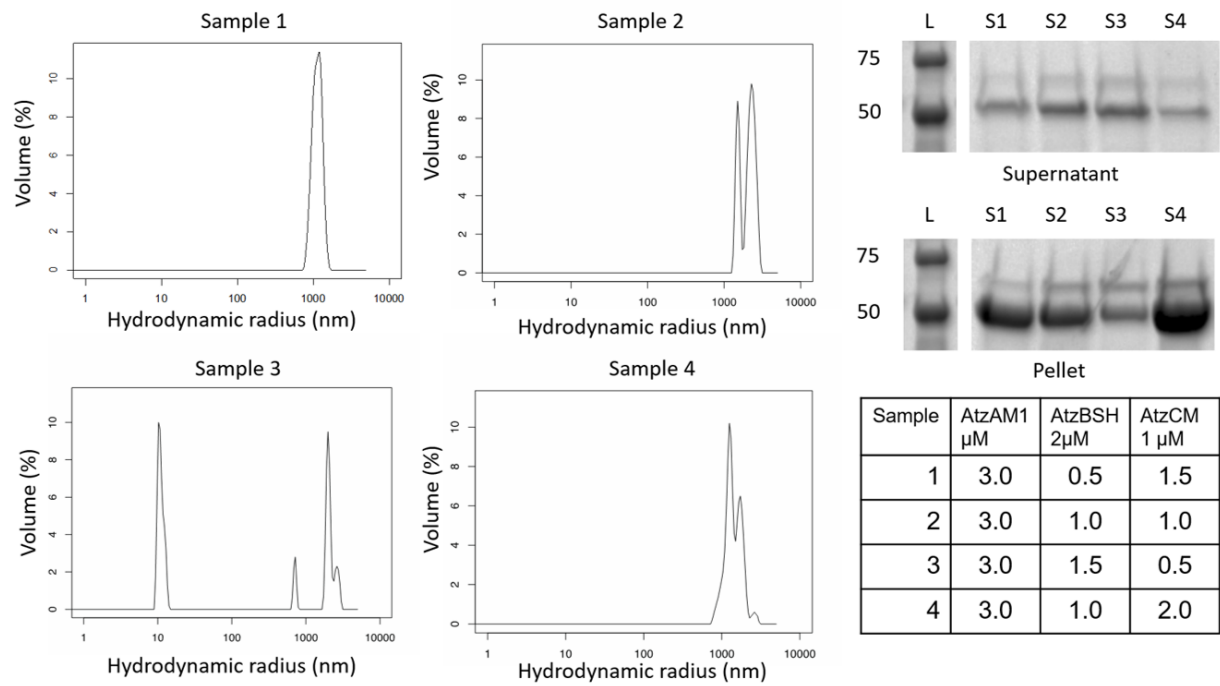
Fig. S25. DLS and SDS PAGE confirm AtzBSH2 incorporation into the 3-component assembly. AtzAM1, AtzBSH2, and AtzCM1 were added and allowed to incubate at various concentrations, then analyzed with DLS which showed that the addition of AtzBSH2 continues to have an assembly at ~1 μm. The SDS Page gel samples were a pelleted sample of the three components assembly and supernatant. If AtzBSH2 is incorporated into the assembly, there should not be any left in the supernatant. The pellet shows that the expected MW weight of AtzBSH2 ~69kda is seen in the pellet with increasing AtzBSH2 concentrations, this indicates that the AtzBSH2 was incorporated into the assembly since it became insoluble and does not appear in the supernatant.
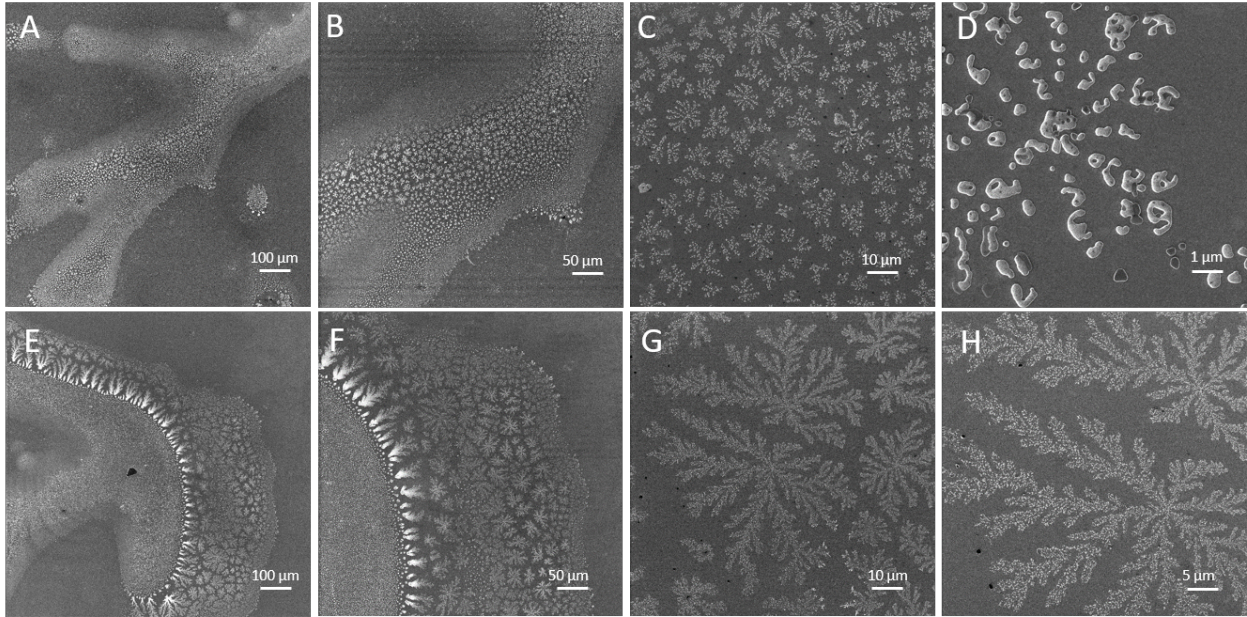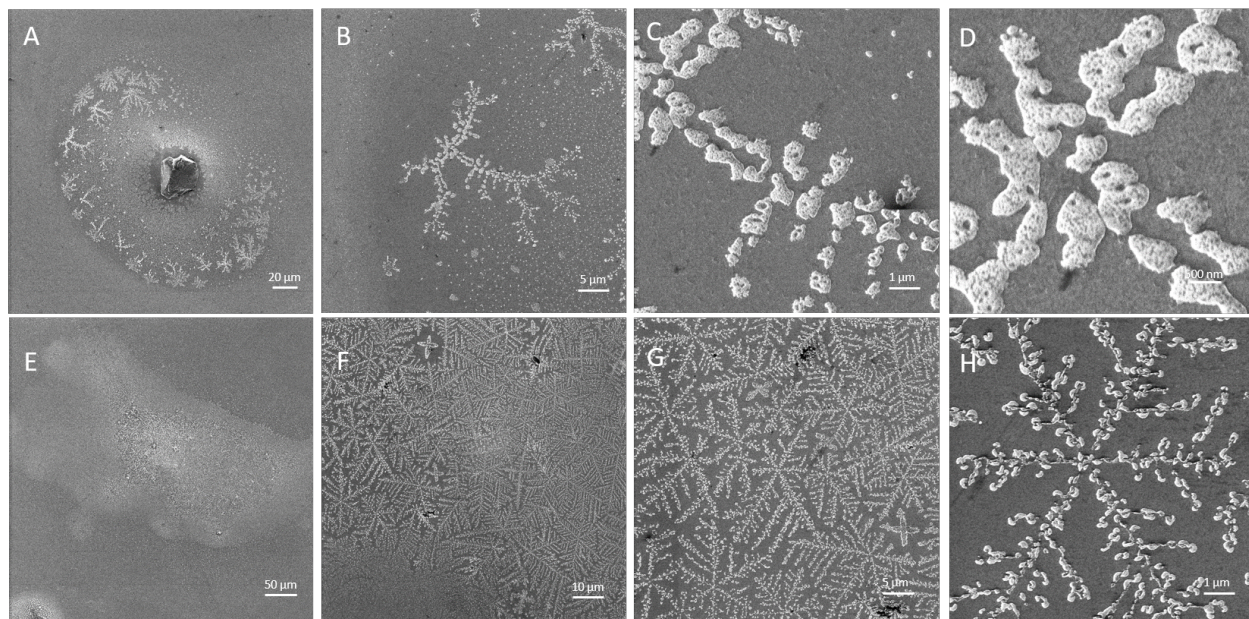
Fig. S26. Helium Ion Microscopy (HIM) images depict fractal-like assembly with 3 μM AtzAM1, 1 μMAtzBSH2, 1 μM AtzCM1 final protein concentrations. (**A** to **D**) Various views of the fractal-like 3-component assembly are shown.

Fig. S27. Helium Ion Microscopy (HIM) images depict fractal-like assembly with 3 μM AtzAM1, 1 μMAtzBSH2, 2 μM AtzCM1 final concentrations. (**A** to **H**) Various views of the 3-component assembly with fractal-like structures are shown.

1

2
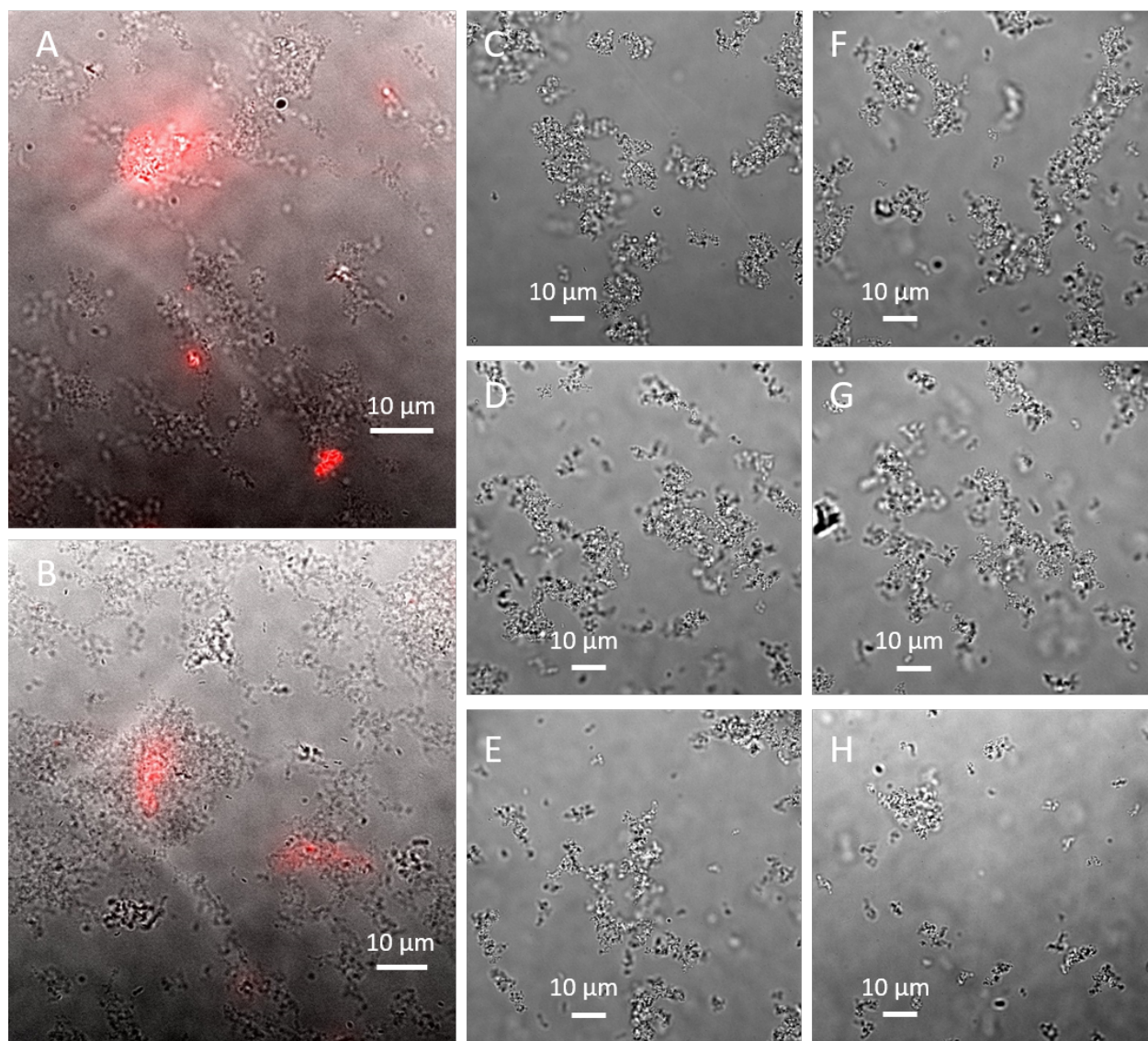
Fig. S28. Fluorescence microscopy and bright-field images of the 3-component assembly confirm incorporation of AtzBSH2 into assembly while bright-field images confirm the fractal-like nature of the 2-component assembly. (**A** and **B**) 3 μM AtzAM1, 1 μM AtzBSH2 dye labeled with Alexa Fluor™ 647, 2 μM AtzCM1 image shows AtzBSH2 incorporation into 3-component assembly at various locations (**C** to **H**) 3 μM AtzAM1 and 2 μM AtzCM1 assembly images depict fractal-like assembly structure.

8

9

10

11

12

13

1

2

3

4

5
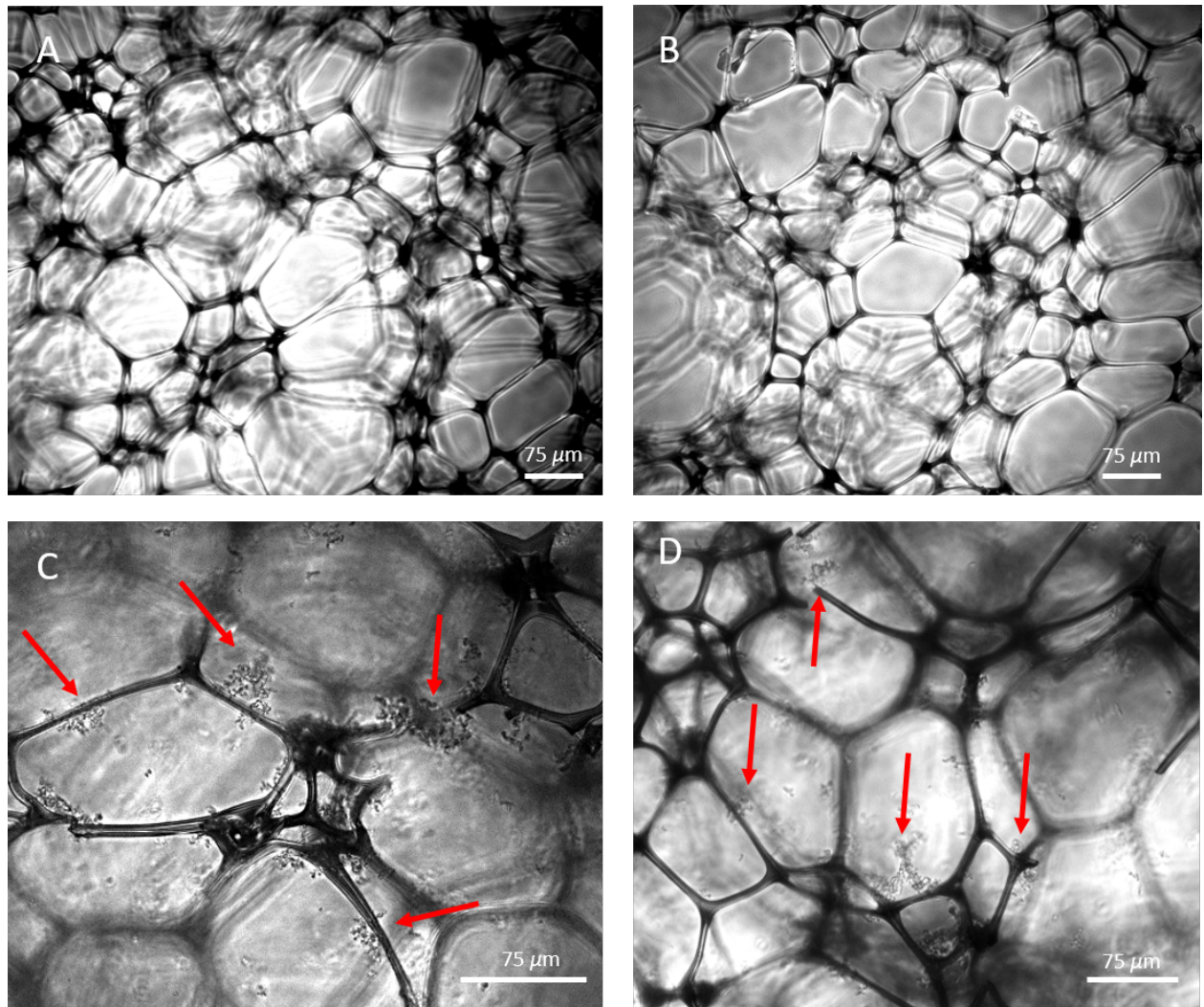
Fig. S29. Phase contrast micrographs of the Basotect® polymer foam with and without assemblies. (**A** and **B**) The microporous polymer foam with no assemblies. (**C** and **D**) The assemblies have been immobilized into the polymer foam, red arrows depict locations with assemblies. Images were taken with a Leica DM4000 B LED microscope, 10X objective (100X total magnification).

10

11

12

13

14

1

2

3
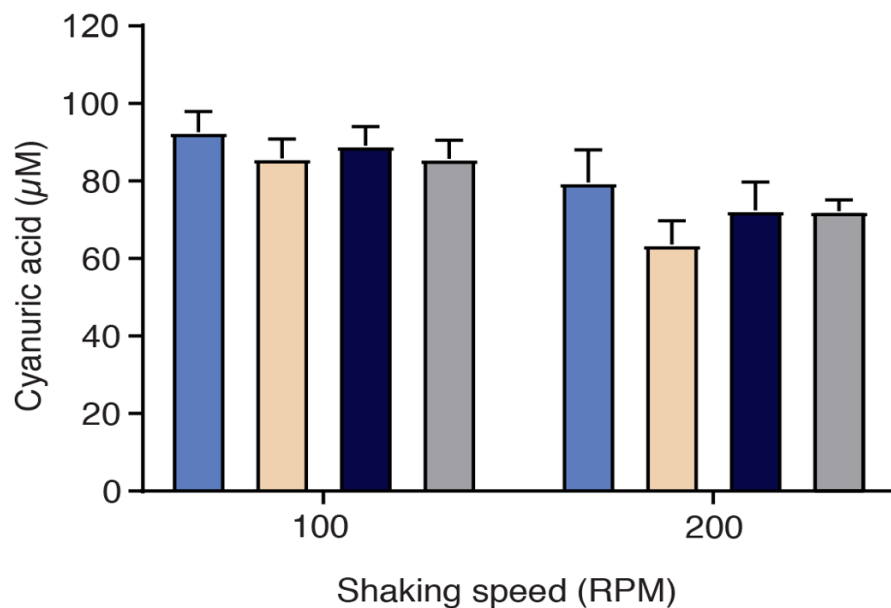
Fig. S30. The fractal-like assemblies (Reg-Assembly) and the extended linker globular assemblies (ExtLinker-Assembly) enzymatic conversion of atrazine to cyanuric acid demonstrates no enzymatic benefit of a globular assembly. AtzB was incorporated into the two-component assembly as an SH2-domain fusion as previously described to create the three-component assembly for both the fractal and globular assemblies. The activity of the fractal-like assembly was higher than the extended linker assemblies under high shaking speeds of 200 rpm.

9

10

11

12

13

14

15

16

17

18

19

20

1
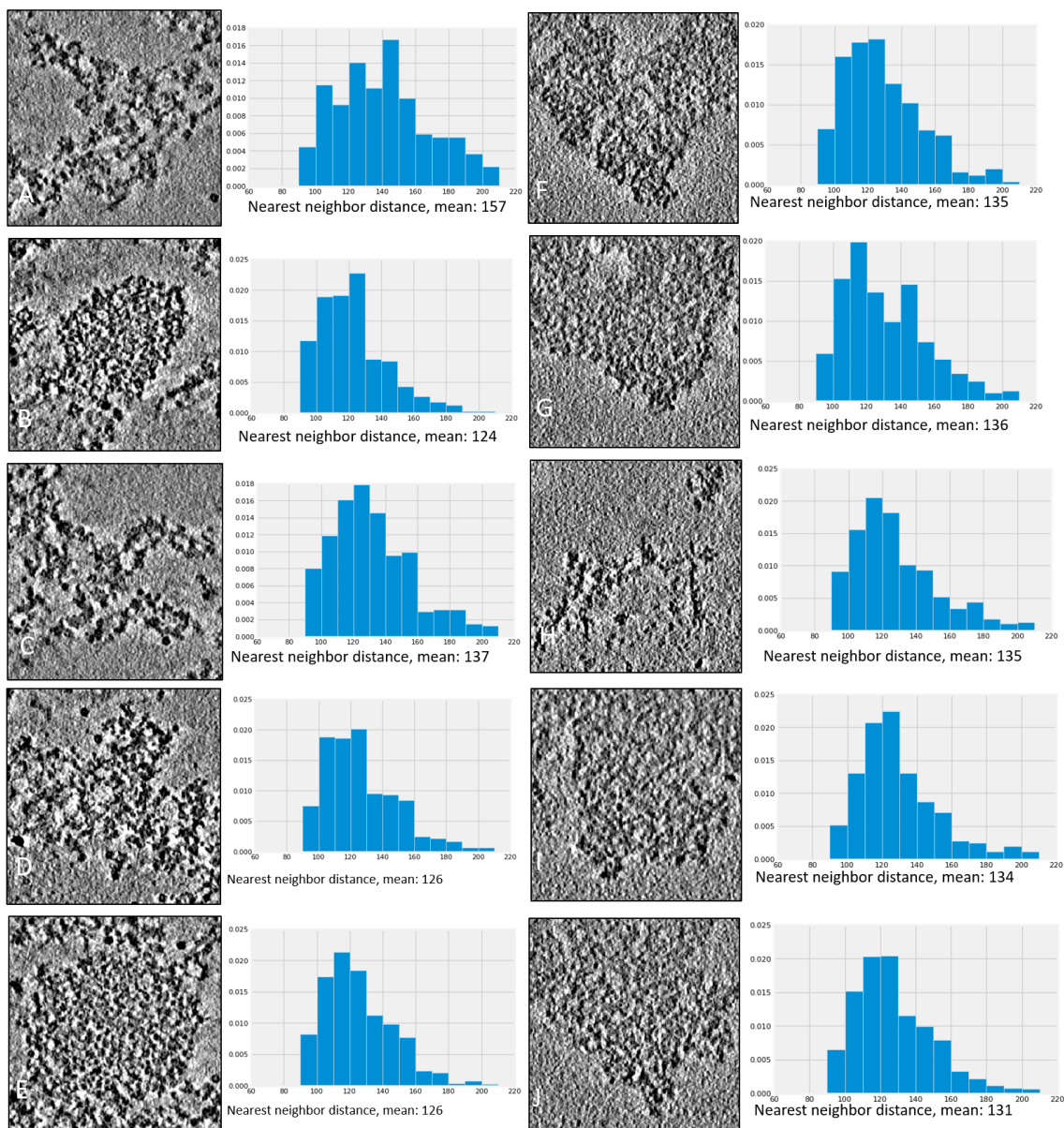
Fig S31. Analysis of the fractal assembly CryoEM tomograms and the extended linker globular assemblies. CryoEM
tomograms of the fractal-like assemblies (A-E) and the extended linker assemblies (F-J) next to the calculated nearest neighbor
distance and mean average distance are shown.

5

6