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

Abstract: Fractal topologies, which are statistically self-similar over multiple length scales, are pervasive in nature. The recurrence of patterns at increasing length scales in fractal-shaped branched objects, *e.g.*, trees, lungs, and sponges, results in high effective surface areas, and provides key functional advantages, *e.g.*, for molecular trapping and exchange. Mimicking these topologies in designed protein-based assemblies will provide access to novel classes of functional biomaterials for wide ranging applications. Here, we describe a modular, multi-scale computational design method for the reversible self-assembly of proteins into tunable supramolecular fractal-like topologies in response to phosphorylation. Computationally-guided atomic-resolution modeling of fusions of symmetric, oligomeric proteins with Src homology 2 (SH2) binding domain and its phosphorylatable ligand peptide was used to design iterative branching leading to fractal-like assembly formation by enzymes of the atrazine degradation pathway. Structural characterization using various microscopy techniques and Cryo-electron tomography revealed a variety of dendritic, hyperbranched, and sponge-like topologies which are self-similar over three decades (~10nm-10μm) of length scale, in agreement with models from

[1]Department of Chemistry & Chemical Biology, Rutgers University, Piscataway, NJ.[2]Institute for Quantitative Biomedicine, Rutgers University, Piscataway, NJ.[3]Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ.[4]Department of Molecular Biology & Biochemistry, Rutgers University, Piscataway, NJ.[5]Department of Biomedical Engineering, Rutgers University, Piscataway, NJ. [6]Institute of Advanced Materials, Devices, and Nanotechnology, Rutgers University, Piscataway NJ. [7]Program in Structural and Computational Biology and Molecular Biophysics, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX. [8]BioTechnology Institute, University of Minnesota, St. Paul, MN[9]Department of Cell Biology and Neuroscience, Rutgers, University, Piscataway, NJ.[10]Department of Physics and Astronomy, Rutgers University, Piscataway, NJ.[11]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, St. Paul, MN
*Contributed equally to this work

multi-scale computational simulations. We demonstrate control over mesoscale topology (by linker design), formation dynamics, and functional enhancements due to dynamic multi-component assemblies constructed with three atrazine degradation pathway enzymes. The described design method should enable the construction of a variety of novel, spatiotemporally responsive catalytic biomaterials featuring fractal topologies.

Main

Fractional-dimensional (fractal) geometry – a property of shapes that are invariant or nearly invariant to scale magnification or contraction across many length scales – is a common feature of many natural objects[1]. Fractal forms are ubiquitous in geology, *e.g.*, in the architecture of mountain ranges, coastlines, snowflakes, and in physiology, *e.g.,* neuronal and capillary networks, and nasal membranes, where highly efficient molecular exchange occurs due to a fractal-induced high surface area:volume ratio[2]. Fabrication of fractal-like nanomaterials affords high physical connectivity within patterned objects[3], ultrasensitive detection of target binding moieties by patterned nanosensors[4], and rapid exchange and dispersal of energy and matter[5]. An intimate link between structural fractal properties of designed, nanotextured materials and functional advantages (e.g., detection sensitivity) has been demonstrated[4], and synthetic fractal materials are finding applications in sensing, molecular electronics, high-performance filtration, sunlight collection, surface charge storage, and catalysis, among myriad other uses[6,7]. Many fractal fabrication efforts have relied on top-down patterning of surfaces[8]. The bottom-up design of supramolecular fractal topologies – both deterministic (e.g., Sierpinski's triangles)[9,10] and stochastic fractals (e.g., arborols)[11,12]– has been performed with small molecule building blocks such as inorganic metal-ligand complexes or synthetic dendritic polymers utilizing co-ordinate or covalent bonds,

respectively. However, fractal topologies have not been designed with biomacromolecules, which possess a wide range of functionality, biocompatibility, and whose properties are dynamically controllable by reversible non-covalent forces[13]. While fractal-like topologies have been detected as intermediates in the formation of natural protein-based biomaterials such as biosilica and silk[14,15], and observed in peptide assemblies[16,17], their tunable construction by utilizing reversible non-covalent interactions between protein building blocks under mild conditions remains a fundamental design challenge.

Self-assembly of engineered proteins[18] provides a general framework for the controllable and bottom-up fabrication of novel biomaterials with chosen supramolecular topologies but these approaches have, thus far, been applied to the design of integer (two or three)-dimensional ordered patterns such as layers, lattices, and polyhedra[19–24]. While external triggers such as metal ions and redox conditions have been used to trigger synthetic protein and peptide assemblies[16,17,25,26], phosphorylation – a common biological stimulus used for dynamic control over protein function – has yet to be utilized for controlling protein assembly formation.

Among stochastic fractals, an arboreal (tree-like) shape is an elementary topology that can be generated using stochastic branching algorithms, *e.g.*, L-systems[27], in which the probability of branching, length and number of branches, and branching angle ranges at each iteration determine the emergent topology (Fig. 1A). To implement a general approach for tunably building arboreal fractal morphologies using triggerable self-assembly of protein building blocks, we envisioned the need for three design elements: (a) multiply branching components, (b) a modular system for connecting these components reversibly in response to a chosen chemical trigger, and (c) limited conformational flexibility at protein-protein connection points, such that stochastic but directional

propagation of multiple branching geometries leads to emergent fractal-like supramolecular topologies. We chose (a) the oligomeric enzymes AtzA (hexameric) and AtzC (tetrameric) of the atrazine biodegradation pathway[28] featuring dihedral (D_3 and D_2, respectively) symmetry (Fig. 1B), (b) a phosphopeptide (pY) tag with its corresponding engineered high-affinity "superbinder" Src homology 2 (SH2) domain[29,30], and (c) short designed linker segments as these design elements, respectively (Fig. 1B,C,D). The sequences and conformational landscapes of the designed protein components were obtained using a procedure implemented in the Rosetta macromolecular modeling program aimed at making a maximum of three divalent connections between AtzA and AtzC mediated by SH2 domain-phosphopeptide binding: first, one of the C_2 axes of the crystallographic structures of the two components were aligned. (Fig. 1B). Two alignments (Fig. 1E,F), obtained by rotating AtzA (hexamer) by 180° about its C_3 axis, were considered, and the remaining two symmetry-compatible degrees of freedom for placement – the inter-component center-of-mass distance d and rotation angle θ about the aligned axis of symmetry – were varied (Fig. 1B,E,F). The resulting placements were evaluated using RosettaMatch[31] for geometrically feasible fusion to the SH2 domain and phosphopeptide with the C-terminal AtzC and N-terminal of AtzA, respectively. Loop closure and optimization of the new intra- and inter-component interfaces generated by fusion and placement, respectively, were performed using Rosetta Kinematic Loop Closure and RosettaDesign. Five AtzA-AtzC design pairs were chosen for experimental characterization based on calculated interface energies in the designed conformation, number of residue insertions in connecting loops (zero), total number of substitutions (<5), and visual examination of design models.

To evaluate the energetically favorable emergent structures upon assembly formation dictated by designed inter-component interactions, the conformational landscape over all (d, θ) pairs (Fig. 1E,F) was constructed using Rosetta SymmetricFastRelax simulations for a designed hexamer-tetramer complex, and the calculated energies (Figs. 1E,F, S1) were Boltzmann-weighted (using a simulation temperature parameter, T) to obtain a probability distribution $\boldsymbol{P}(d, \theta)$ for branching geometry. This distribution, in turn, was used as input for a coarse-grained stochastic chain-growth tree generation algorithm for predicting ensembles of emergent topologies on the micrometer length scale (Fig. 1G-K, S2). For comparison with experiments, ~100s of emergent structures in the resulting ensemble were analyzed for fractal (Hausdorff) dimension (D_F) using the box counting image processing technique (Fig. 1L; see Methods). A variety of assembly sizes and fractal dimensions, D_F, could be obtained by varying three simulation parameters (also see Supplementary Discussion): the fraction of the two components at each growing layer (c_{frac}), the probability of termination at any propagatable connection point (p_{null}), and the Boltzmann factor (k_BT), which determines the sampling of inter-component conformational diversity calculated from Rosetta simulations (Fig. 1M-P).

Genes encoding the designed AtzA and AtzC variants and the corresponding fusions of wild type domains were constructed and cloned into an *E. coli* BL21(DE3) strain harboring a second plasmid for the inducible expression of GroEL/ES chaperones to aid protein yields. Purified AtzA designs were each phosphorylated using Src kinase and the presence of phosphotyrosine was confirmed using ELISA assays (Fig. S3); binding and assembly formation with purified AtzC-SH2 domain fusions was assessed using Biolayer Interferometry and Dynamic Light Scattering, respectively. Phosphorylation, binding and complete conversion of monomers

into 1-10 μm-sized particles upon mixing was best detected with the proteins pY-AtzAM1 and AtzCM1 (Fig. S4, S5, S6), and we chose this design pair for further characterization of assembly-disassembly processes (Fig. 2A). Apart from fusion of pY-tag and SH2 domain, these proteins feature 1 and 4 substitutions compared to their wild type parent, respectively (Fig. S7 and S8).

Assembly formation by a mixture of the two components and Src kinase enzyme was ATP dependent (Fig. 2B), was accompanied by the visible and spectrophotometrically measurable (Fig. S9) appearance of turbidity, which could be reversed by adding a phosphatase (YopH) enzyme. The resulting distribution of particle sizes was detected by measuring hydrodynamic radii using Dynamic Light Scattering (DLS) (Fig. 2C). Upon completion of assembly formation, the apparent size of the particles as measured by DLS was between 1-10 μm; however, this range represents the upper limit of measurement for the instrument; actual particle sizes were expected to be larger. Addition of monovalent competitive inhibitors, *i.e.* isolated SH2 domain or SH2 domain fused to an unrelated monovalent protein (SH2-DhaA) inhibited assembly formation in a concentration-dependent manner, demonstrating that the SH2-pYtag binding interaction underlies assembly formation. The apparent IC_{50} for the observed inhibition was ~2×[AtzA-pY] (measured as monomers) at two different concentrations of the components (Fig. 2D, S10 to S12), and in each case ~3×[AtzA-pY] was required for complete inhibition. According to our design model, each pY-AtzA (hexamer) makes at least two and at most three divalent connections for assembly propagation (Fig. 1E,F); thus, the observed inhibition stoichiometries are consistent with the existence of the designed divalent connections between AtzA-pY and AtzC-SH2 in the assemblies.

As the phosphorylation reaction requires ATP, assembly formation rates could be controlled by varying the concentration of added ATP. For [AtzA-pY] and [AtzC-SH2] of 3 μM

and 2 μM, respectively, [ATP] > 250 μM led to complete conversion of monomers to assemblies within 5 minutes, whereas significantly slower rates of conversion were observed with lower [ATP] (Fig. 2E, S13, Table S1). Visualization of assemblies using optical and fluorescence microscopy (with Alexa-647-labeled AtzC) revealed the existence of large (>10 μm) dendritic structures (Fig. 2F, G), whose formation could be observed in real time by adding kinase and ATP to a mixture of the two component proteins (Movie S1, Fig. S14).

Apparent hydrodynamic radius (Fig. 2F, G) and polydispersity measured with DLS (Fig. S15 and S16) could be controlled by varying the relative stoichiometry of the two components, and by using a weaker binding affinity variant of the SH2 domain fused to AtzC. A comparison of assembly formation trends for the lower (Fig. 2F) and higher affinity (Fig. 2G) SH2-domain-containing constructs shows that robust assembly formation is observed at nearly equal concentrations of the two components. Assemblies can be formed at concentrations as low as 50 nM (Dissociation constant, K_D, for the weaker and tighter interactions were measured as ~40 and ~7 nM, respectively; Fig. S6), whereas when one component is present in excess, assembly formation is inhibited, as expected from our branch propagation design model (Fig. 1). The existence of greater assembly formation by "off-diagonal" non-stoichiometric concentration combinations (particularly at low concentrations of AtzA-pY) for the tighter binding variant compared to the weaker-binding variant (Fig. 2F, G) indicates that the inhibition caused by an excess of the binding partner is dynamic and can be overcome using multivalency (especially for AtzA-pY which makes three connections according to the design model) in an affinity-dependent manner.

We next investigated if the dynamic and dendritic structures observed in solution by optical and fluorescence microscopy (Fig. 2H, I) could form surface-induced fractals, and if the topology of the surface-directed assemblies could be controlled by varying component stoichiometry. Due to the substantial increase of surface area derived from fractal patterns, surface-induced fractals at the nanometer-micrometer scale are attractive design targets for applications in many fields like catalysis, fractal electronics, and the creation of nanopatterned sensors[3,4]. Assemblies with a chosen stoichiometry of components were generated in buffer, dropped on the surface of a silicon (or mica) chip, and the solvent was evaporated at room temperature (298 K) under a dry air atmosphere. Visualization of these coated surfaces using Helium Ion and Atomic Force microscopy reveals striking, intricately textured patterns that coat up to 100 μm^2 areas (Fig. 3A-D). Various morphologies on the micron scale including rod-like, tree-like, fern-like, and petal-like were observed (Fig. 3A-D); image analysis revealed fractal dimensions between 1.4-1.5 (Fig. 3A,B) to the more Diffusion Limited Aggregation (DLA)-like 1.78 (Figs.3C,D, S17, and S18). Assembly sizes and fractal dimensions could be tuned by varying the stoichiometry of components (Fig. 3E), although some heterogeneity in morphologies was present in each sample. At 1:1 stoichiometry of the two components, DLA-like topologies with ~10 μm size were observed, whereas more dendritic assemblies were observed when unequal stoichiometry samples were used (Fig. 3E). Similarly, smaller assembly sizes resulted when the concentration of one component became limiting.

Fractal patterns were not observed at any component stoichiometry without addition of ATP and Src kinase, with unphosphorylated proteins, or upon drying the buffer (to preclude precipitation-induced assembly formation by the salt in the buffer) demonstrating that fractal

structures are formed by designed components (Fig. S19). Similarly, fractal topologies were not detected when long ($(GSS)_{10}$), conformationally flexible Gly-Ser-rich linkers were used to fuse the SH2 domain and pY tag to AtzC, and AtzA, respectively. In mixtures of these proteins, a densely packed globular topology was detected with HIM, typical of amorphous precipitates (Fig. S20). Thus, the surface-induced patterns observed with designed AtzC and AtzA are selectively formed upon inter-component association in the designed geometries but not upon isotropic, random association as expected for the highly flexible Gly-Ser-rich linker-containing variants.

Transmission electron microscopy of designed AtzA-AtzC proteins also revealed branching, dendritic networks reminiscent of fractal intermediates observed in biosilica formation[14] (Fig. S21). However, the low resolution of these images precludes identification and examination of individual protein components and their connectivity in the fractal structures. To investigate the conformations of designed assemblies in solution and to obtain sufficiently high-resolution structures to test the validity of our design approach, we characterized the assemblies using cryo-electron tomography (cryo-ET; Fig. 3F,G, S22, Movies S2, S3). Assemblies generated by mixing 3 μM pY-AtzA and 2 μM AtzC-SH2 (or corresponding AtzA and AtzC fusions with Gly-Ser-rich linkers as controls) were blotted on a grid, frozen, and visualized on a cryo-electron microscope. Due to the increased image contrast from Volt phase plates in our microscope setup, pY-AtzA and AtzC-SH2 complexes in assembly tomograms were easily identified as density clusters. In contrast, constructs with Gly-Ser-rich linkers connecting pY and SH2 domain with AtzA and AtzC did not form porous clusters but instead (~90% of the sample) formed large, dense globular clumps (Fig S22B) where individual components were not resolvable (also see Supplementary Discussion). These large topology changes on the micron scale (as observed by both cryo-ET and HIM) upon conformational flexibility changes at the nanometer scale, further

re-inforce the importance of directional association in our modular fractal assembly design framework.

Computational annotation of the density clusters formed by designed components in cryo-ET-derived images was performed based on individual molecular envelopes of components derived from Rosetta models of pY-AtzA and AtzC-SH2, respectively, to identify inter-component connections along assembly branches (Fig. 3F). The topology of the largest, nearly fully interconnected assembly based on electron density (Fig. 3G), consisting of approximately 6000 individual protein components, was further analyzed, and compared with an ensemble of simulated structures with approximately the same number of components. We compared the observed distributions of nearest-neighbor counts for AtzA-pY (Fig. 3H, Fig S23), relative numbers of component types incorporated (Fig. 3I) and the observed fractal dimension (Fig. 3J) of the assemblies with ensembles of structures generated using computational modeling (Fig. 3K) and found good agreement between the data and our simulations performed at specific parameter values (Fig. S2). The observed nearest neighbor distribution for the AtzA-pY component shows that a large majority of these proteins are connected to 1, 2, or 3 neighboring AtzC-SH2, in agreement with the divalent connections envisioned in the design model and implemented in the simulated assemblies (Fig. 1). Additionally, a small but significant number of AtzA-pY proteins have 4 AtzC neighbors in both the computational ensemble and the cryo-ET images, which indicates physically unconnected components being proximal to each other in space due to the packing in the assembly (Fig. 3H). We found that the fractal dimensions from the cryo-ET images and simulations (2.1) show good agreement (Fig. 3J,K). The expected fractal dimension for a DLA-like cluster, which results from isotropic interactions, is 2.3 and the observed decreased fractal dimension (2.1) indicates the non-isotropic nature[32–34] and/or lack of diffusion-limited

association of the underlying protein-protein interactions. Particle counting (and volume estimation) in a convex hull enclosing the largest assembly component yields an approximate local concentration of the proteins as ~600-700 μM, a ~125-fold increase compared to their bulk concentration (3 μM AtzA-pY and 2 μM AtzC-SH2). While there is significant heterogeneity in assembly sizes (~60% of the proteins adsorbed on the cryo-ET grid are parts of smaller assemblies) and topologies (Fig. S24), the observed increase in the effective concentrations concomitant with a large effective surface area with numerous solvent channels (Fig. 3F,G) indicates that induced fractal-like structure formation is a viable strategy to engineer protein assemblies with favorable sponge-like properties.

To investigate if sponge-like fractal assembly formation endows component proteins with a functional advantage over their unassembled counterparts, we measured the catalytic performance of component atrazine degrading enzymes (AtzA-pY and AtzC-SH2) subjected to various environmental stresses, and when immobilized on a macroscale three-dimensional melamine sponge scaffold. Three (AtzA, AtzB, and AtzC) of the six enzymes (AtzA-F) in the atrazine mineralization pathway sequentially catalyze the conversion of the groundwater contaminant and endocrine disruptor atrazine to the relatively benign cyanuric acid (Fig. 4A). While the crystal structure of AtzB is not available, it is expected to be a dimer[35]. We reasoned that by fusing AtzB to an SH2 domain, and by using sub-stoichiometric amounts of AtzA-pY and AtzC-SH2 that can still yield assemblies (Fig. 2H,I), AtzB can be effectively incorporated in the assembly in a host-guest strategy (Fig. 4B). Assays with DLS, SDS-PAGE (Fig. S25), HIM imaging (Fig. 4C, S26, and S27) and fluorescence microscopy (Fig. 4D, S28) confirmed that AtzB was incorporated in the assembly. Given the large size (several micrometers) and high

concentrations of proteins in the assembled state that were visualized in the cryo-ET data, we reasoned that assembled proteins would be protective towards environmental stresses such as increased temperature and mechanical shear forces that expose proteins to the air-water interface and cause denaturation. Indeed, assembled proteins are more thermotolerant (Fig. 4E) and withstand high shaking speeds (Fig. 4F) better than their unassembled counterparts, as evidenced by greater retention of activity under stress. Interestingly, the activity of assembled proteins is the highest at 200 rpm (Fig. 4F), indicating possibly enhanced mass transfer of substrates and intermediates from the bulk under these conditions.

To investigate the potential for manufacturing a flow-through atrazine water treatment material, we immobilized assemblies within a commercial open cell foam polymer matrix (Basotect®) that was protected with a silicon oxide polymer surface generated from tetraethoxysiloxane (TEOS), using previously described hydrolysis methods[36]. The enzyme assemblies were well-retained in the matrix (Fig. S29) and upon incubation with 150 μM atrazine, produced higher levels of cyanuric acid than parallel incubations with free enzymes that had been similarly immobilized (Fig. 4G). As the assemblies are also more robust to thermal and mechanical stress, these results indicate that it may be feasible to use the designed assemblies with the foam matrix for efficient atrazine removal from contaminated water. Finally, under shaking stress, the fractal-like designed assemblies produced more cyanuric acid than those formed with components with extended (Gly-Ser-rich) linkers (Fig. S30), indicating that a random, non-fractal-shaped assembly (Fig. S20, S22) is less functionally advantageous.

Our results establish a modular design framework by which fusion proteins may be designed to self-assemble into fractal-like morphologies on the 10 nm-10 μm length scale. The

design strategy is conceptually simple, modular, and should be applicable to any set of oligomeric proteins featuring cyclic, dihedral, and other symmetries, such that multivalent connections along with designed semi-flexible loops can be used to controllably generate a broad range of sizes and morphologies of fractal shapes with proteins. Although we used SH2 domain-pY peptide fusions as the modular connecting elements to endow phosphorylation responsiveness, the same design strategy should be applicable for the incorporation of other peptide recognition domains, responsive to other chemical or physical stimuli. The combination of multivalency and chain flexibility is a key determinant of other recently discovered phases formed by proteins, including droplets formed by liquid-liquid phase separation[37]. Our results show that this rich phase behavior of proteins also includes fractal-like morphologies that form colloidal particles with constituent microscopic molecular networks which may be visualized at high resolution using cryo-ET. Given the wide-ranging applications of fractal-like nanomaterials, further development in the design of protein-based fractals described here is expected to enable the production of novel classes of bionanomaterials and devices.

Figures

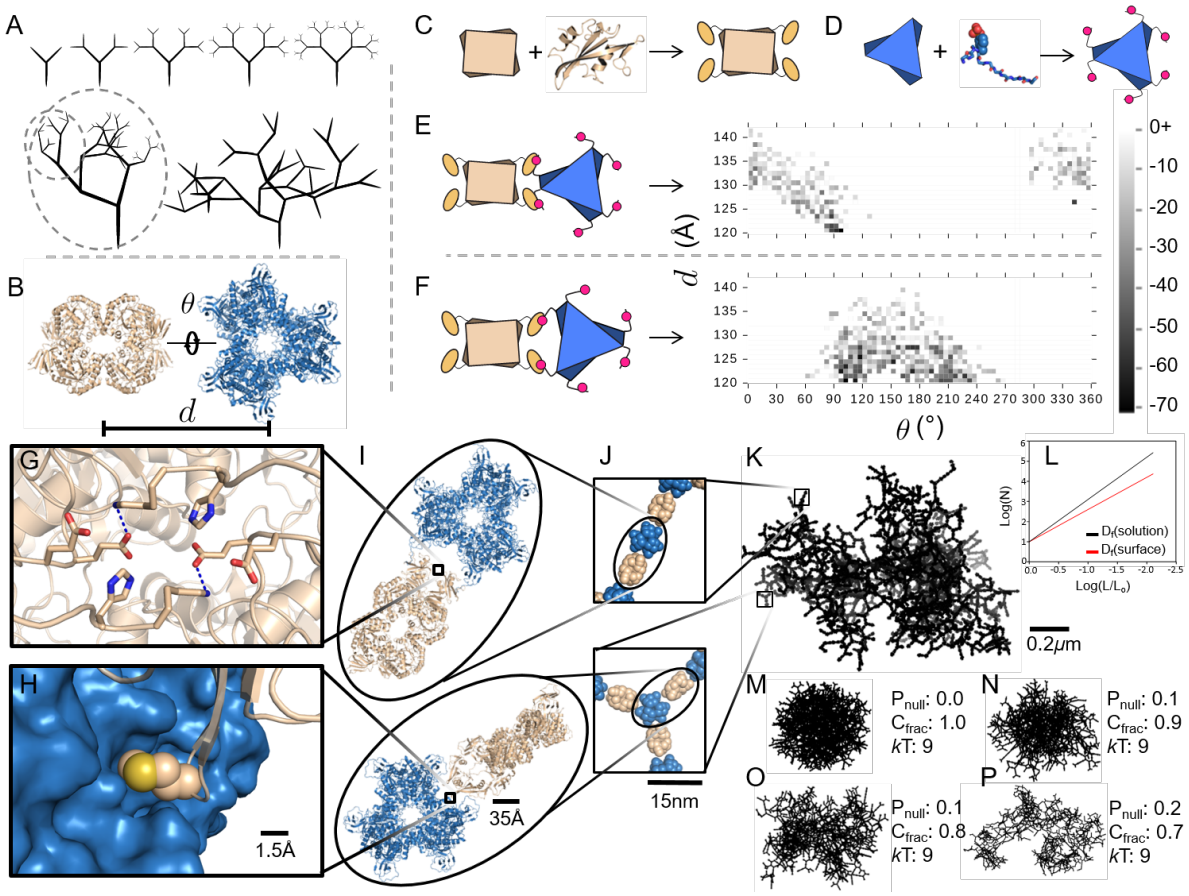


Figure 1: Multi-scale Computational Design Approach for fractal assembly design with pY-AtzA and AtzC-SH2. a, Cartoon representations of an ordered self-similar scaling fractal, an unordered self-similar scaling fractal—note concentric circles that are self-similar at different scales—and an unordered statistically self-similar fractal. **b,** Two-component library of AtzC (tan) and AtzA (blue) positions was generated by varying the rigid body degrees of freedom along paired C_2 symmetry axes. **c** and **d,** Design and modeling of assembly at the molecular scale was performed by fusing an SH2 binding domain and its corresponding phosphorylatable peptide to AtzC (C) and AtzA (D) respectively. Linker between the SH2 domain and AtzC was designed

to ensure near-symmetric binding between the hexamer and tetramer leading to propagation. **e** and **f,** Flexibility analysis was performed by evaluation of the Rosetta energy landscape of symmetrical connections and the probability of observing different connection distances and angles were calculated using the Boltzmann distribution for two binding modes: vertex (E) and Edge (F). **g** and **k,** Boltzmann weighted connection probabilities were utilized in a stochastic chain-growth program with a coarse-grained protein model to generate emergent structures. (**L**) Representation of expected fractal dimension (slope) for fractals analyzed in solution and on surfaces. **m** to **p,** Fractal simulation output across varying null probabilities (P_{null}) and fraction of components (C_{frac}) at fixed kT.

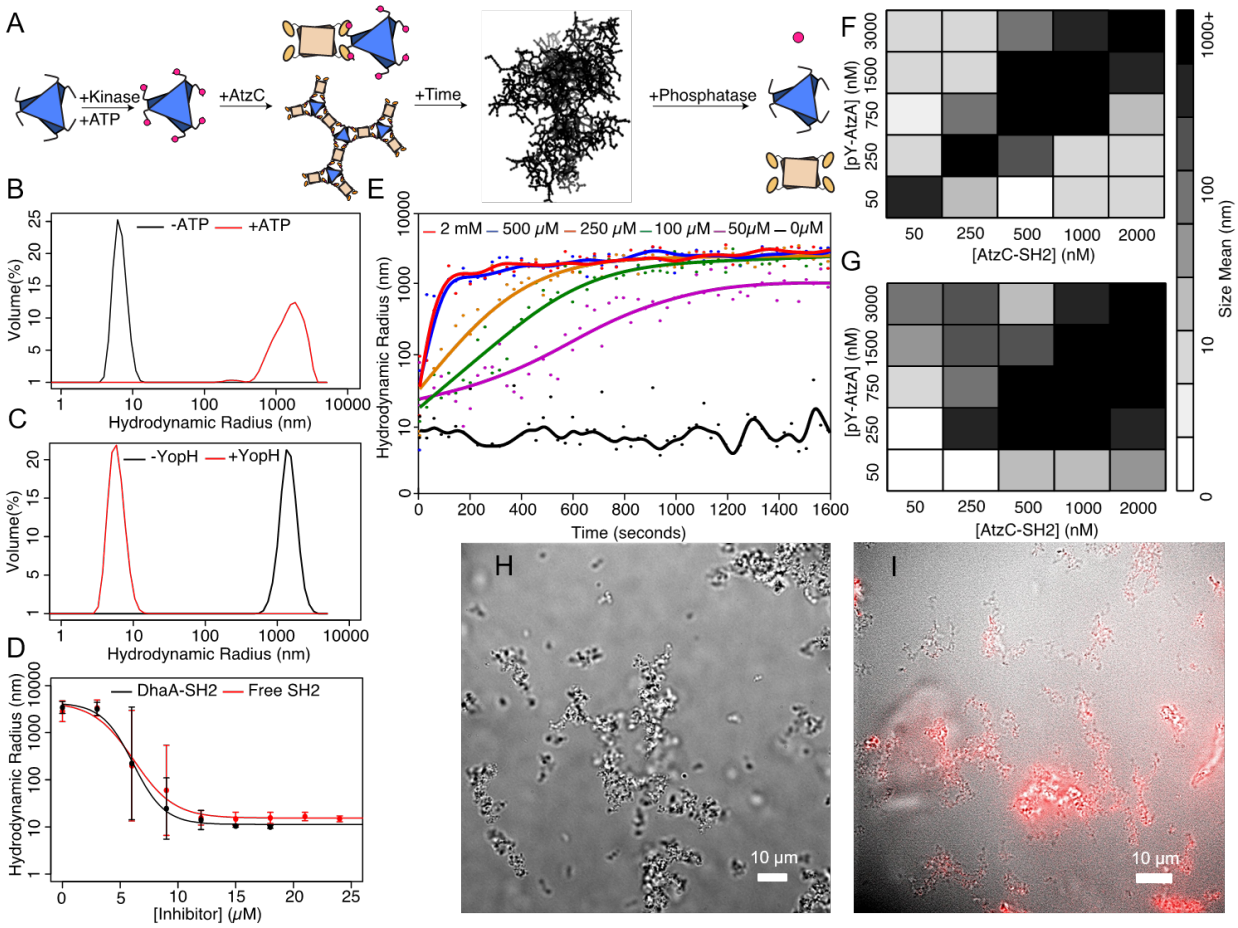


Figure 2: Assembly Formation, Dissolution and Inhibition *in vitro*. (**A**) AtzAM1 can be phosphorylated using purified Src kinase (pY-AtzAM1) and incubated with AtzCM1-SH2 to form an assembly. Likewise, the phosphatase (YOP) enzyme can be used to disassemble these structures. **b** and **c,** Assemblies were expected to form (B) and dissolve (C), respectively, as confirmed by DLS measurements. **d,** Incubation of assembling components with various concentrations of free SH2 domain and a different (monovalent) SH2 fusion protein led to robust inhibition. **e,** ATP concentration was shown to control the rate of assembly formation. **f** and **g,** Assembly formation is highly sensitive to stoichiometry of the components. Varying the stoichiometry and the use of a weaker-binding SH2-peptide interaction (F) leads to a perturbation of the assembly formation zone compared to the "superbinder" SH2 (G). **h,** The

dendritic structure in solution as observed by bright field microscopy. **i,** Fluorescence microscopy image of dye-labeled (Alexa Fluor 647^{TM}) AtzCM1-SH2 and a non-dye labeled pY-AtzAM1 shows the dendritic structures in solution.

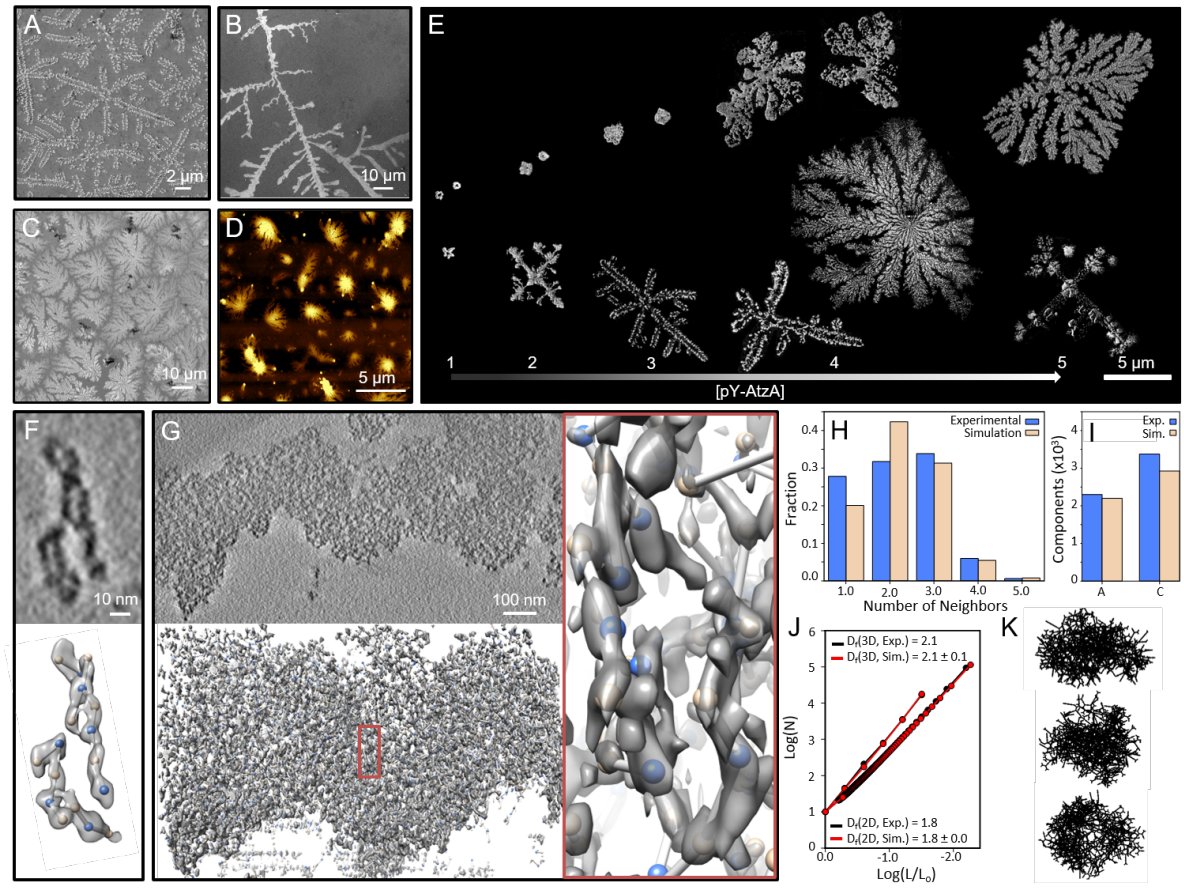


Figure 3: Assembly formation and characterization with Helium Ion Microscopy (HIM), Atomic Force Microscopy (AFM), and Cryo-electron Tomography, all reveal fractal-like topologies on a surface and in solution, **a** to **d**, Longer fractal-like structures, branch-like, and **c** to **d,** flower-like structures are seen in HIM and AFM. **e,** Representative HIM images for assemblies obtained at different concentrations of pY-AtzAM1 (250 nM- 3 µM) while maintaing a fixed concentration of AtzCM1-SH2 (2 µM). Increasing concentrations of pY-AtzM1 result in

larger assemblies that appear more lacunar, fractal-like, and demonstrate the impact of stoichiometry on assembly topology and size. **f** and **g,** A small and large tomogram respectively. Subtomograms were extracted and fitted with pY-AtzAM1 (blue spheres) and AtzCM1-SH2 (tan spheres) models. **h,** Connection information between AtzA and AtzC complexes was used for statistical analysis of the number of neighbors, which was compared to simulations. **i,** The nearest-neighbor component distribution obtained from the tomograms was found to closely match the component distributions from the simulation. (**J**) Image analysis (2D), using a box counting method, of the subtomograms converted into 2D projections show similar fractal dimension (slope) with 2D projections of structures from the simulations. 3D box counting revealed similar fractal dimension (slope) between the subtomograms and simulations. **k,** Simulation parameters found to match the experimental data include: P_{null}: 0.1, C_{frac}: 1.0, and kT: 9.0; three representative 2D projections with the matching parameters are shown.

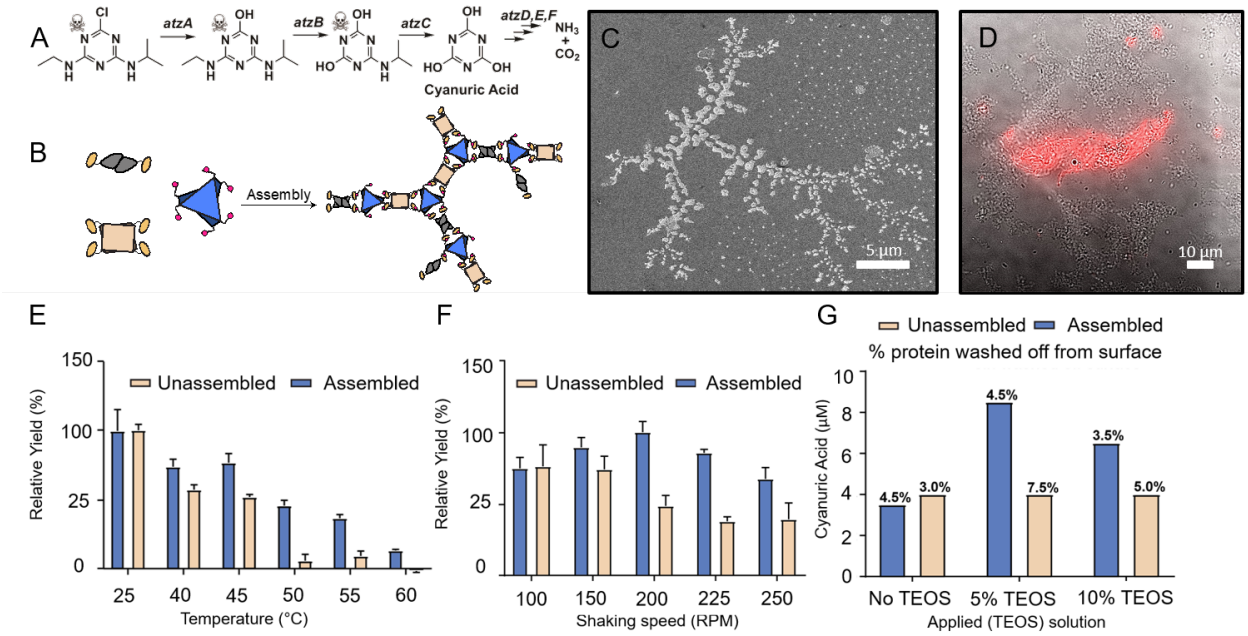
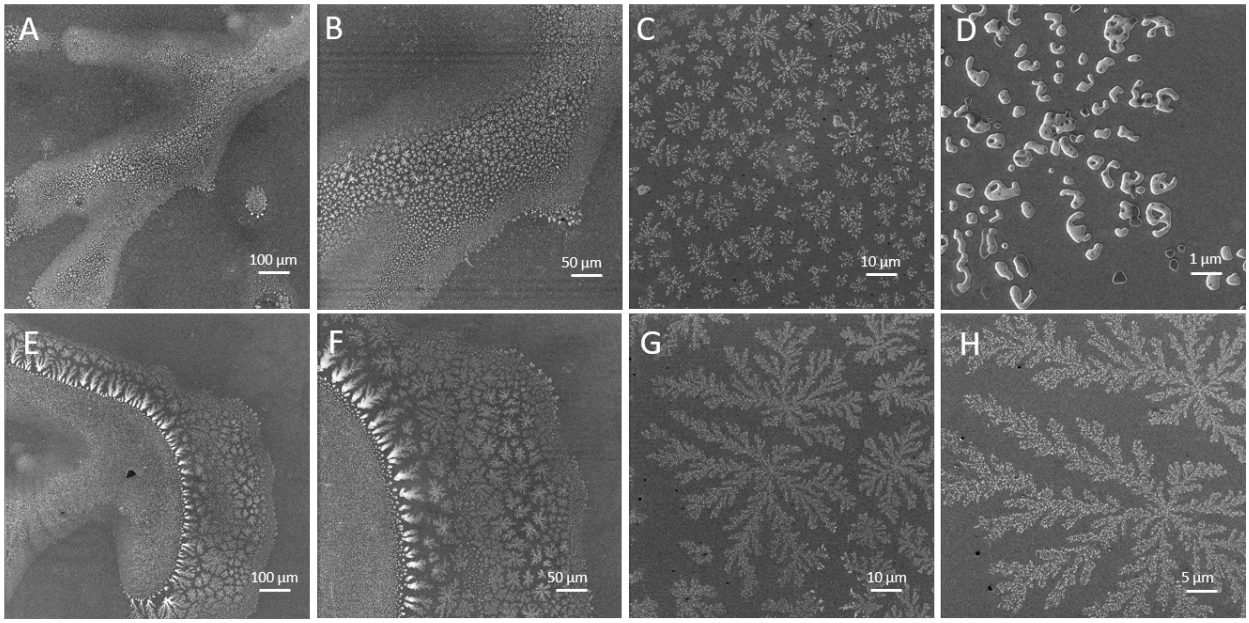
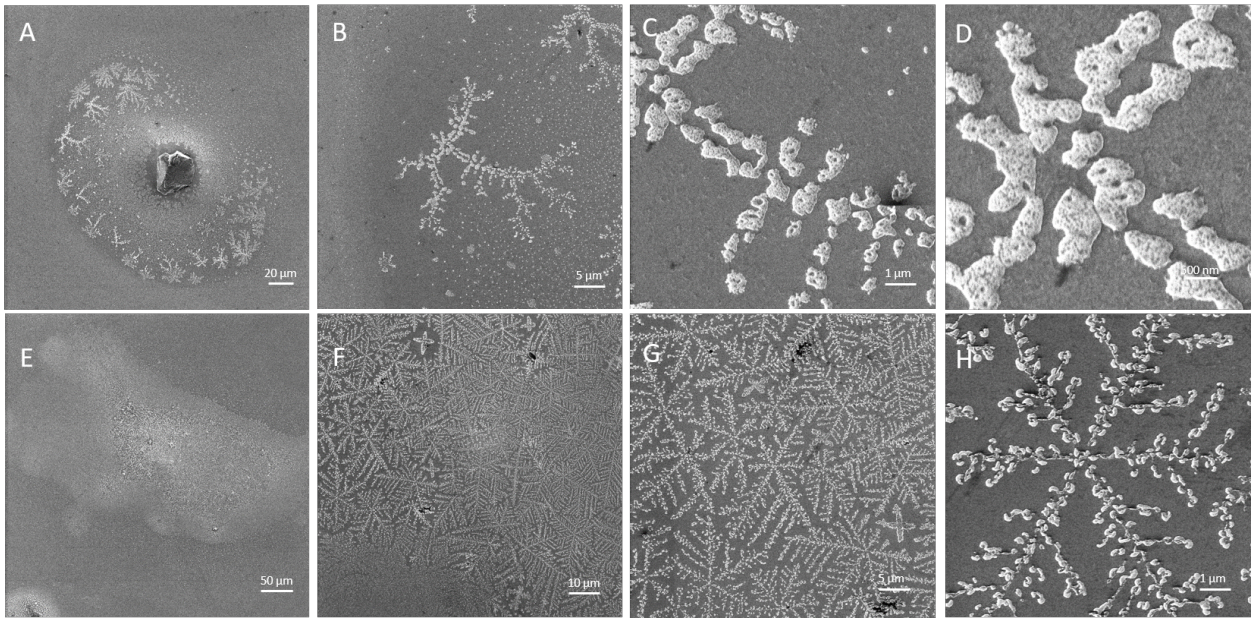
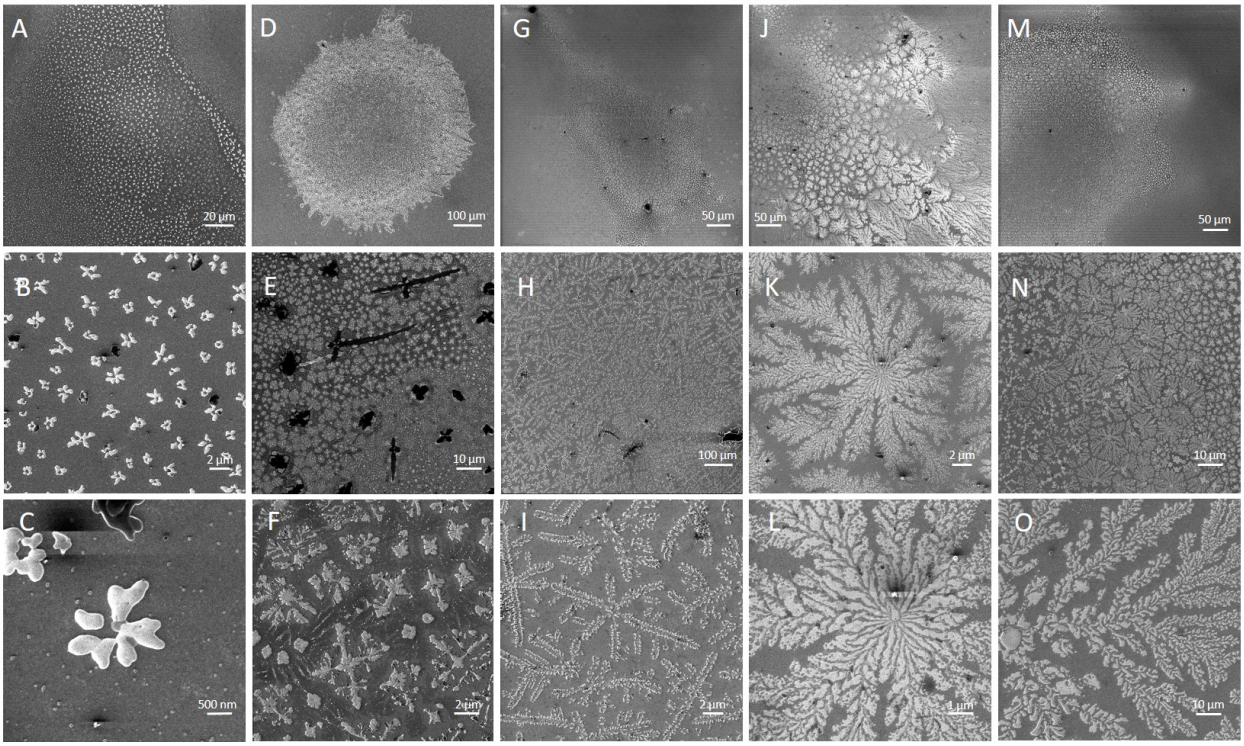


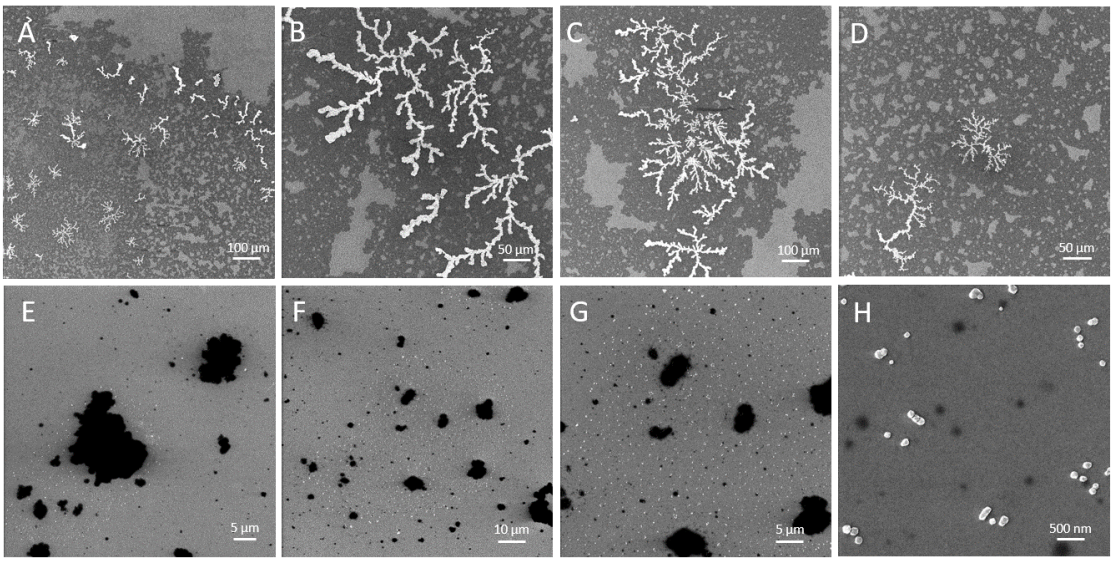
Figure 4: AtzBSH2 incorporation to construct a three-enzyme assembly. **a,** Atrazine degradation pathway, enzymatic conversion of atrazine to cyanuric acid, and further enzymatic conversion to NH_3 and CO_2. **b,** AtzB was added as an SH2-domain fusion to the two-component (AtzA-AtzC) assembly. **c** and **d,** Three-component assembly formation was validated using HIM (C), and the incorporation of AtzB was confirmed with fluorescence microscopy (using an Alexa-658-labeled AtzB) (D). **e** and **f,** Assemblies were found to be more thermotolerant, as detected by incubation at a given temperature for 30 min followed by activity assays, and more robust to mechanical shearing forces, as detected by ability to withstand shaking. **g,** Assemblies and free enzymes were incorporated into a Basotect® polymer foam with different TEOS % layers, to trap proteins, and assayed for cyanuric acid production. Proteins can be lost during the wash step after crosslinking and the % of protein lost under each condition is indicated on top of the bars.

Extended data (microscopy images; see SI for more information):









References:

1. Mandelbrot, B. The fractal geometry of nature. 406 (1983).
2. Losa, G.A., Merlini, D., Nonnenmacher, T.F., Weibel, E. . in *Fractals in biology and medicine. Volume IV*. vii, 314 (Birkhäuser, Basel, 2005).
3. Fairbanks, M. S., McCarthy, D. N., Scott, S. A., Brown, S. A. & Taylor, R. P. Fractal electronic devices: Simulation and implementation. *Nanotechnology* **22,** (2011).

4. Soleymani, L., Fang, Z., Sargent, E. H. & Kelley, S. O. Programming the detection limits of biosensors through controlled nanostructuring. *Nat. Nanotechnol.* **4,** 844–848 (2009).
5. Ge, J., Lei, J. & Zare, R. N. Protein-inorganic hybrid nanoflowers. *Nat. Nanotechnol.* **7,** 428–432 (2012).
6. Zhang, P. & Wang, S. Designing fractal nanostructured biointerfaces for biomedical applications. *ChemPhysChem* **15,** 1550–1561 (2014).
7. Lim, B. *et al.* Pd-Pt bimetallic nanodendrites with high activity for oxygen reduction. *Science (80-.).* **324,** 1302–1305 (2009).
8. Cerofolini, G. F., Narducci, D., Amato, P. & Romano, E. Fractal nanotechnology. *Nanoscale Res. Lett.* **3,** 381–385 (2008).
9. Newkome, G. R. *et al.* Nanoassembly of a fractal polymer: A molecular 'Sierpinski hexagonal gasket'. *Science (80-.).* **312,** 1782–1785 (2006).
10. Shang, J. *et al.* Assembling molecular Sierpiński triangle fractals. *Nat. Chem.* **7,** 389–393 (2015).
11. Newkome, G. R. & Moorefield, C. N. From 1 → 3 dendritic designs to fractal supramacromolecular constructs: understanding the pathway to the Sierpiński gasket. *Chem. Soc. Rev.* **44,** 3954–3967 (2015).
12. Shin, S. *et al.* Polymer Self-Assembly into Unique Fractal Nanostructures in Solution by a One-Shot Synthetic Procedure. *J. Am. Chem. Soc.* **140,** 475–482 (2018).
13. Astier, Y., Bayley, H. & Howorka, S. Protein components for nanodevices. *Current Opinion in Chemical Biology* **9,** 576–584 (2005).
14. Murr, M. M. & Morse, D. E. Fractal intermediates in the self-assembly of silicatein filaments. *Proc. Natl. Acad. Sci.* **102,** 11657–11662 (2005).
15. 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).
16. Lomander, A., Hwang, W. & Zhang, S. Hierarchical self-assembly of a coiled-coil peptide into fractal structure. *Nano Lett.* **5,** 1255–1260 (2005).
17. Shen, W., Lammertink, R. G. H., Sakata, J. K., Kornfield, J. A. & Tirrell, D. A. Assembly of an artificial protein hydrogel through leucine zipper aggregation and bisulfide bond formation. *Macromolecules* **38,** 3909–3916 (2005).
18. McManus, J. J., Charbonneau, P., Zaccarelli, E. & Asherie, N. The physics of protein self-assembly. *Curr. Opin. Colloid Interface Sci.* **22,** 73–79 (2016).
19. King, N. P. *et al.* Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science (80-.).* **336,** 1171–1174 (2012).
20. Hsia, Y. *et al.* Design of a hyperstable 60-subunit protein icosahedron. *Nature* **535,** 136–139 (2016).
21. Suzuki, Y. *et al.* Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **533,** 369–373 (2016).
22. Sinclair, J. C., Davies, K. M., Vénien-Bryan, C. & Noble, M. E. M. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nat. Nanotechnol.* **6,** 558–562 (2011).
23. Padilla, J. E., Colovos, C. & Yeates, T. O. Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proc. Natl. Acad. Sci.* **98,** 2217–2221 (2001).
24. Zhang, J., Zheng, F. & Grigoryan, G. Design and designability of protein-based

assemblies. *Current Opinion in Structural Biology* **27,** 79–86 (2014).
25. Brodin, J. D. *et al.* Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nat. Chem.* **4,** 375–382 (2012).
26. Ringler, P. & Schulz, G. E. Self-assembly of proteins into designed networks. *Science (80-.).* **302,** 106–109 (2003).
27. Lindenmayer, A. Mathematical models for cellular interactions in development II. Simple and branching filaments with two-sided inputs. *J. Theor. Biol.* **18,** 300–315 (1968).
28. Wackett, L., Sadowsky, M., Martinez, B. & Shapir, N. Biodegradation of atrazine and related s-triazine compounds: From enzymes to field studies. *Applied Microbiology and Biotechnology* **58,** 39–45 (2002).
29. Kaneko, T. *et al.* Superbinder SH2 Domains Act as Antagonists of Cell Signaling. *Sci. Signal.* **5,** ra68-ra68 (2012).
30. Yang, L. *et al.* Computation-Guided Design of a Stimulus-Responsive Multienzyme Supramolecular Assembly. *ChemBioChem* **18,** 2000–2006 (2017).
31. Richter, F., Leaver-Fay, A., Khare, S. D., Bjelic, S. & Baker, D. De novo enzyme design using Rosetta3. *PLoS One* **6,** (2011).
32. Van Anders, G., Ahmed, N. K., Smith, R., Engel, M. & Glotzer, S. C. Entropically patchy particles: Engineering valence through shape entropy. *ACS Nano* **8,** 931–940 (2014).
33. Zhang, Z. & Glotzer, S. C. Self-assembly of patchy particles. *Nano Lett.* **4,** 1407–1413 (2004).
34. Nicolas-Carlock, J. R., Carrillo-Estrada, J. L. & Dossetti, V. Fractality a la carte: A general particle aggregation model. *Sci. Rep.* **6,** (2016).
35. Seffernick, J. L. *et al.* Hydroxyatrazine N-ethylaminohydrolase (AtzB): An amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *J. Bacteriol.* **189,** 6989–6997 (2007).
36. 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).
37. Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **11,** 899–904 (2015).

Supplementary Information. Contains methods, supplementary discussion, supplementary information references, Figures S1-S31, Table S1, Movies S1-S3.

Acknowledgments. SDK and LWP acknowledge support from the NSF (grants MCB1330760, MRI142962). NEH acknowledges the NSF Graduate Research Fellowship (DGE-1433187). We thank J. Chodera for providing *E. coli* expression-optimized genes for Src kinase and Yop phosphatase; I. Marrero-Berríos, H. Cho, M. Liu, A. Permaul, O. Dineen, and R. Patel, for

experimental assistance; K-B Lee, G. Montelione, and V. Nanda for technical advice, and V. Nanda for helpful discussions and comments on the manuscript.

Author Contributions. NEH, WAH, and SDK, designed the research. WAH designed the proteins using Rosetta, constructed the fractal growth simulations and analyzed microscopy images. NEH, DZ, MES, and MK expressed and purified all the proteins used in the study. NEH, VM, TG, LPF performed Helium Ion Microscopy. NEH, MP and S.-H, Lee performed fluorescence microscopy and bright-field microscopy. DZ and WAH performed the DLS experiments. MK performed the BLI experiments. LY performed the TEM experiments. NEH and MES performed enzyme activity assays. AGD, LWP performed the polymer foam immobilization and activity assays. WD and MB performed the Cryo-electron tomography experiments and analyses. WAH and MC performed computational analyses of the Cryo-ET data and compared to simulations. SDK, NEH and WAH wrote the manuscript. All authors commented on the manuscript.