nature chemistry

ARTICLES

https://doi.org/10.1038/s41557-019-0277-y

Stimulus-responsive self-assembly of protein-based fractals by computational design

Nancy E. Hernández[1,2,12], William A. Hansen[2,12], 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], Ileana Marrero-Berríos[5], Melissa Banal[9], Phillip Rechani[10], 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]*

Fractal topologies, which are statistically self-similar over multiple length scales, are pervasive in nature. The recurrence of patterns in fractal-shaped branched objects, such as trees, lungs and sponges, results in a high surface area to volume ratio, which provides key functional advantages including molecular trapping and exchange. Mimicking these topologies in designed protein-based assemblies could provide access to functional biomaterials. Here we describe a computational design approach for the reversible self-assembly of proteins into tunable supramolecular fractal-like topologies in response to phosphorylation. Guided by atomic-resolution models, we develop fusions of Src homology 2 (SH2) domain or a phosphorylatable SH2-binding peptide, respectively, to two symmetric, homo-oligomeric proteins. Mixing the two designed components resulted in a variety of dendritic, hyperbranched and sponge-like topologies that are phosphorylation-dependent and self-similar over three decades (~10 nm–10 μm) of length scale, in agreement with models from multiscale computational simulations. Designed assemblies perform efficient phosphorylation-dependent capture and release of cargo proteins.

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[1,2]. Fractal forms are ubiquitous in geology (for example, in the architecture of mountain ranges), as well as in coastlines, snow formations and in physiology (for example, in neuronal and capillary networks and nasal membranes, where highly efficient molecular exchange occurs due to a fractal-induced high surface area to volume ratio[3]). Fabrication of fractal-like nanomaterials affords high physical connectivity within patterned objects[4], ultrasensitive detection of target binding moieties by patterned nanosensors[5], and rapid exchange and dispersal of energy and matter[6]. An intimate link between structural fractal properties of designed, nanotextured materials and functional advantages (for example, detection sensitivity) has been demonstrated[5], 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[7,8]. Many fractal fabrication efforts have relied on top-down patterning of surfaces[9]. The bottom-up design of supramolecular fractal topologies—both deterministic (for example, Sierpinski's triangles)[10,11] and stochastic (for example, arborols)[12,13] fractals—has been performed with small-molecule building blocks such as inorganic metal–ligand complexes or synthetic dendritic polymers utilizing coordinate or covalent bonds, respectively. Self-similar quasi-fractal shapes built with DNA origami have been reported[14–16]; however, fractal topologies have not been designed with proteins, which possess a wide range of functionalities and biocompatibility, and whose properties are dynamically controllable by reversible post-translational modifications[17]. Although fractal-like topologies have been detected as intermediates in the formation of natural protein-based biomaterials such as biosilica and silk[18,19], and have also been observed in peptide assemblies[20–22], 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[23] provides a general framework for the controllable and bottom-up fabrication of novel biomaterials with chosen supramolecular topologies, but these approaches have thus far only been applied to the design of integer (two or three)-dimensional ordered patterns such as layers, lattices and polyhedra[24–30]. Although external triggers such as metal ions and redox conditions have been used to trigger synthetic protein and peptide assemblies[20,21,31–34], 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, such as L-systems[35,36], in which the probability of branching, length and number of branches, and branching angle ranges at each iteration determine the emergent topology (Fig. 1a). Theoretical and simulation studies on the self-assembly of 'patchy' colloidal particles[37,38] have shown that a variety of topologies, including fractal-like topologies[39–41], can result from stochastic self-assembly processes involving strong, anisotropic short-range

[1]Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ, USA. [2]Institute for Quantitative Biomedicine, Rutgers University, Piscataway, NJ, USA. [3]Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, USA. [4]Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ, USA. [5]Department of Biomedical Engineering, Rutgers University, Piscataway, NJ, USA. [6]Institute of Advanced Materials, Devices and Nanotechnology, Rutgers University, Piscataway, NJ, USA. [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, USA. [8]BioTechnology Institute, University of Minnesota, St Paul, MN, USA. [9]Department of Cell Biology and Neuroscience, Rutgers, University, Piscataway, NJ, USA. [10]Department of Physics and Astronomy, Rutgers University, Piscataway, NJ, USA. [11]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, St Paul, MN, USA. [12]These authors contributed equally: Nancy E. Hernández, William A. Hansen. *e-mail: sagar.khare@rutgers.edu

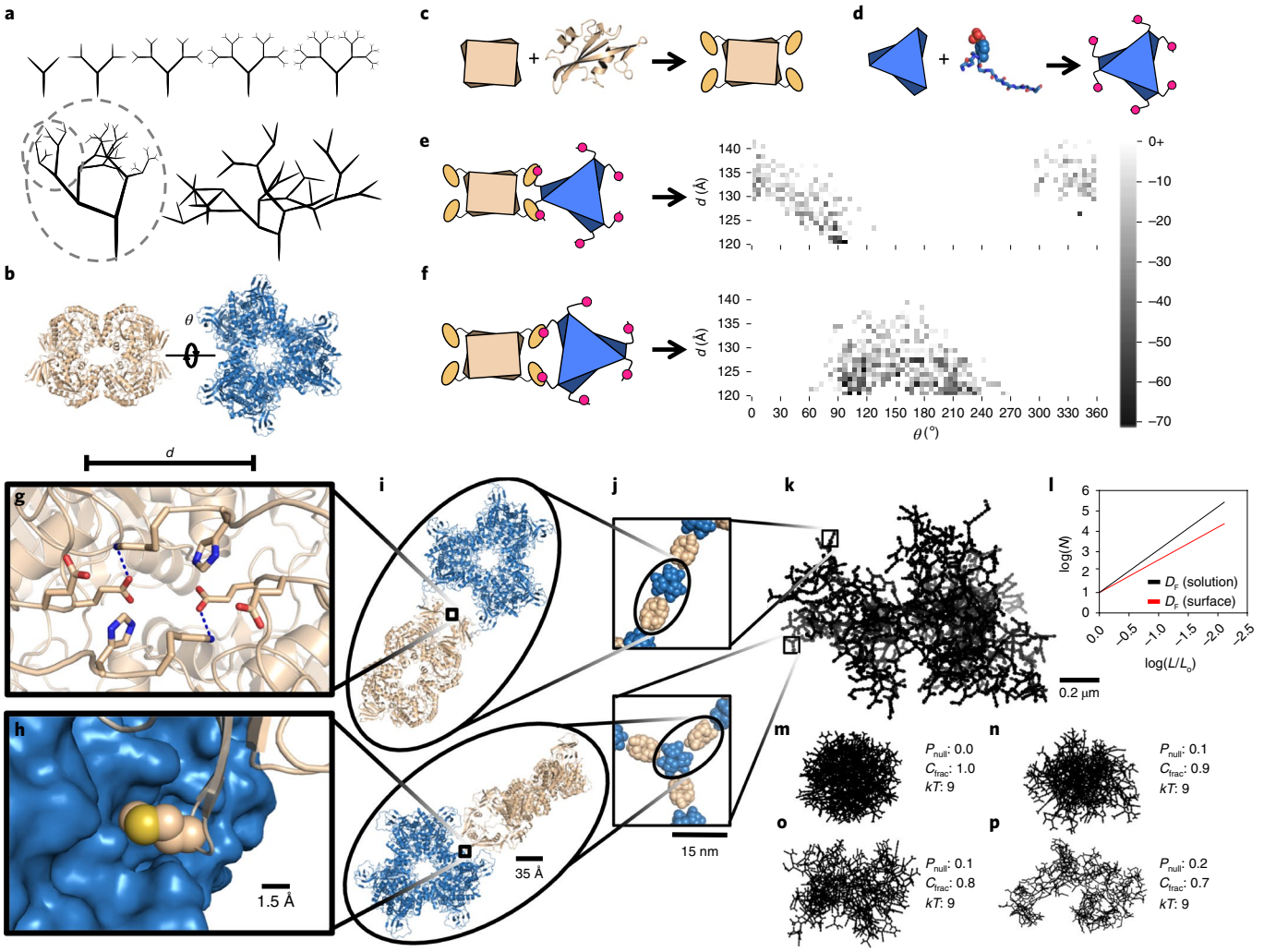


Fig. 1 | Multiscale computational design approach for fractal assembly design. a, Cartoon representations of an ordered self-similar scaling fractal, an unordered self-similar scaling fractal—note concentric circles are self-similar at different scales—and an unordered statistically self-similar fractal. **b**, A 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**,**d**, Design and modelling 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. The linker between the SH2 domain and AtzC was designed to ensure symmetric binding between the hexamer and tetramer leading to propagation. **e**,**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 was calculated using the Boltzmann distribution for two binding modes: vertex (**e**) and edge (**f**). Boltzmann-weighted connection probabilities were utilized in a stochastic chain-growth program with a coarse-grained protein model to generate emergent structures. **g**–**k**, Atomic interactions that stabilized novel interfaces formed from physically connected components (**g**,**h**) dictate the rotation along the *C*-symmetric axis between components (**i**,**j**), which ultimately produce combinations of orientations that lead to fractal-like topologies (**k**) on the micrometre scale. **l**, Representation of expected fractal dimension (slope) for fractals analysed in solution and on surfaces. The fractal dimension is calculated using a box-counting approach in which the image is divided into boxes of increasing length. *N* is the number of boxes that contain pixels at each box length, *L*, and L_o is the largest possible box (starting box) that fits the image. **m**–**p**, Examples of fractal simulation output across varying termination probabilities (P_{term}) and fraction of components (C_{frac}) at fixed *kT*.

forces[42–46]. Under conditions where the inter-molecular interaction energy is much larger (more negative) than the thermal energy, emergent large-scale aggregates are expected to be out-of-equilibrium kinetically trapped states rather than (usually crystalline) globally stable thermodynamic minima. As the reorganization of aggregate morphologies, once formed, is expected to be unfavourable, we reasoned that these kinetic traps can be utilized to produce a tunable and responsive structural (and thus functional) diversity of self-assembled protein-based systems[47].

Results

To implement a general approach for tunably designing arboreal fractal morphologies using triggerable self-assembly of protein building blocks, we envisioned the need for (1) a set of branching components whose binding to each other would lead to propagation of the assembly (Fig. 1a), (2) a modular system for connecting, with high affinity, these components reversibly in response to a chosen chemical trigger and (3) degeneracy of protein–protein binding modes (geometries), such that stochastic, but anisotropic, directional propagation of multiple branching geometries leads to emergent fractal-like supramolecular topologies (Supplementary Fig. 1). We chose (1) the oligomeric enzymes AtzA (hexameric) and AtzC (tetrameric) of the atrazine biodegradation pathway[48] featuring dihedral (D_3 and D_2, respectively) symmetry (Fig. 1b), (2) a phosphorylatable peptide (pY) tag with its corresponding engineered high-affinity 'superbinder' Src homology 2 (SH2) domain[49]

and (3) linker segments that can stabilize multiple binding orientations, respectively, as design elements encoding these properties (Fig. 1b–d). We have previously utilized a similar binding domain–peptide fusion strategy to design non-propagating multicomponent enzyme complexes[50].

Computational design and multiscale modelling of assembly formation. The sequences of the designed protein components were obtained using a procedure implemented in the Rosetta macromolecular modelling program[51], aimed at making a maximum of three divalent connections between each AtzA and AtzC mediated by SH2 domain-phosphopeptide binding (Fig. 1c,d). Divalent connections between components were sought to enable avidity, leading to strong, directional, short-range interactions ('aeolotropic interactions'[44]) that would promote fractal growth (Fig. 1e,f). We also reasoned that geometric degeneracy in the form of multiple propagatable (but still anisotropic) binding modes would favour fractal structures (Supplementary Fig. 1). In the first step of the design procedure, 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 centre-of-mass distance d and rotation angle about the aligned axis of symmetry—were sampled (Fig. 1b,e,f). For every value of d we sampled several discrete values of that, if uniformly adopted, were predicted to lead to an infinitely propagatable integer-dimensional lattice (Supplementary Fig. 1). The resulting propagatable 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 (Supplementary Fig. 2). Loop closure of successful SH2 domain and phosphopeptide placements was performed using Rosetta kinematic loop closure (Supplementary Fig. 2). Next, optimization of the new intra- and inter-component interfaces was performed using RosettaDesign (Supplementary Fig. 3). Five AtzA–AtzC fusion protein pairs were chosen for experimental characterization based on removal of steric clashes (as reflected by the calculated Rosetta energy, Supplementary Table 1), tight interface packing between the SH2 domain and AtzC, and visual examination of design models (Supplementary Fig. 3). We found that short, flexible linker sequences (for example, Gly-Gly-Ser) between the SH2 domain and AtzC led to the most efficient interface packing in designs, while still potentially allowing multiple binding modes: several mutations were common among design models obtained at different (single) values of (d,), suggesting geometric degeneracy in binding by each variant (Supplementary Table 1) would be feasible. Indeed, several other values from the propagatable angle set are energetically feasible for each designed AtzC-SH2 variant (Supplementary Table 1).

To fully evaluate the predicted geometric degeneracy and anisotropy of binding in designed inter-component interactions, the conformational landscape over all (d,) pairs (Fig. 1e,f) was constructed using Rosetta SymmetricFastRelax simulations for a designed hexamer–tetramer complex, and the calculated energies were Boltzmann-weighted (using a simulation temperature parameter, T) to obtain a probability distribution $P(d,)$ for the 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 micrometre length scale. Similar hierarchical approaches have previously been developed for modelling protein crystallization[52] and colloidal particle[43] and protein self-assembly[45]. In our approach, preferred inter-component interaction modes at the sub-nanometre scale (Fig. 1g,h) guide the emergence of higher-order structures on the nanometre (Fig. 1i,j) and micrometre length scales (Fig. 1k). For comparison with experiments, ~100s of emergent structures in the resulting ensemble were analysed to determine the fractal dimension (D_F) using the box counting image processing technique (Fig. 1l). The fractal dimension of an object is a measure of how its mass or shape scales as a function of length scale (Supplementary Fig. 4): an object is considered fractal if this scaling exponent is non-integer and typically less than the Euclidean dimension in which the object is placed. For example, the D_F of vasculature patterns on the two-dimensional (2D) surface of the human retina[53] is 1.7, and a diffusion-limited aggregation cluster in 3D space has a D_F of 2.3 (ref. [54]). In our simulations, a variety of assembly sizes and fractal dimensions, D_F, could be obtained by varying three parameters: (1) the fraction of growth sites selected at each growing layer allowed to continue propagation (c_{frac}), which reflects the stoichiometry of the two components, (2) the probability of termination at any chosen propagatable branching point (P_{term}), which reflects the affinity of interactions (the lower the affinity, the higher the P_{term}) and (3) the Boltzmann factor (k_BT), which determines the sampling of inter-component conformational diversity calculated from Rosetta simulations (Fig. 1m–p and Supplementary Figs. 5 and 6).

Experimental characterization of designed assemblies. Genes encoding the designed AtzA and AtzC variants and the corresponding fusions of wild-type domains were constructed and cloned into an *Escherichia coli* BL21(DE3) strain harbouring 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 by enzyme-linked immunosorbent assay (ELISA; Supplementary Fig. 7); binding and assembly formation with purified AtzC-SH2 domain fusions was assessed using biolayer interferometry and dynamic light scattering (DLS), respectively. Phosphorylation, binding and complete conversion of monomers into 1–10 m particles upon mixing was best detected with the proteins pY-AtzAM1 and AtzCM1 (Supplementary Figs. 8–10). Either phosphorylation levels were lower or inter-component binding was weaker (Supplementary Fig. 9) with other designs, so we chose the pY-AtzAM1:AtzCM1 design pair for further characterization of assembly–disassembly processes (Fig. 2a). Apart from fusion of pY-tag and SH2 domain, these proteins feature two and three substitutions compared to their wild-type parent, respectively (Supplementary Table 1 and Supplemetary Figs. 11 and 12).

Assembly formation by a mixture of the two components and Src kinase enzyme was adenosine triphosphate (ATP)-dependent (Fig. 2b) and was accompanied by the visible and spectrophotometrically measurable (Supplementary Fig. 13) appearance of turbidity, which could be reversed by adding a phosphatase (YopH) enzyme. The resulting distribution of particle sizes was detected by measuring the hydrodynamic radii using DLS (Fig. 2c). On completion of assembly formation, the apparent size of the particles as measured by DLS was between 1 and 10 m; however, this range represents the upper limit of measurement for the instrument, and actual particle sizes are expected to be larger. Addition of monovalent competitive inhibitors, that is, 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 half-maximum inhibitory concentration for the observed inhibition was ~2 [AtzA-pY] (measured as monomers) at two different concentrations of the components (Fig. 2d and Supplementary Figs. 14–16), 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.

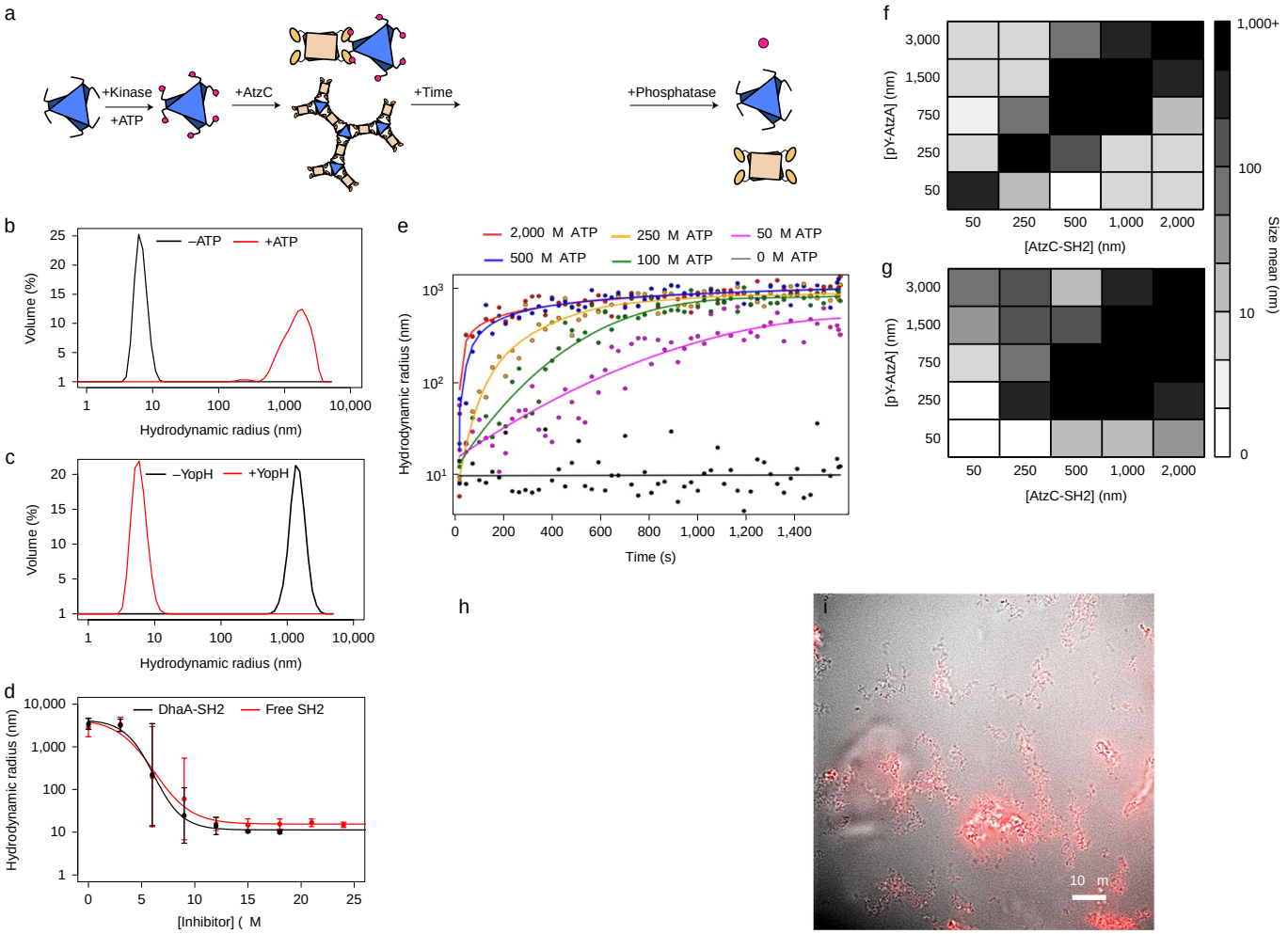
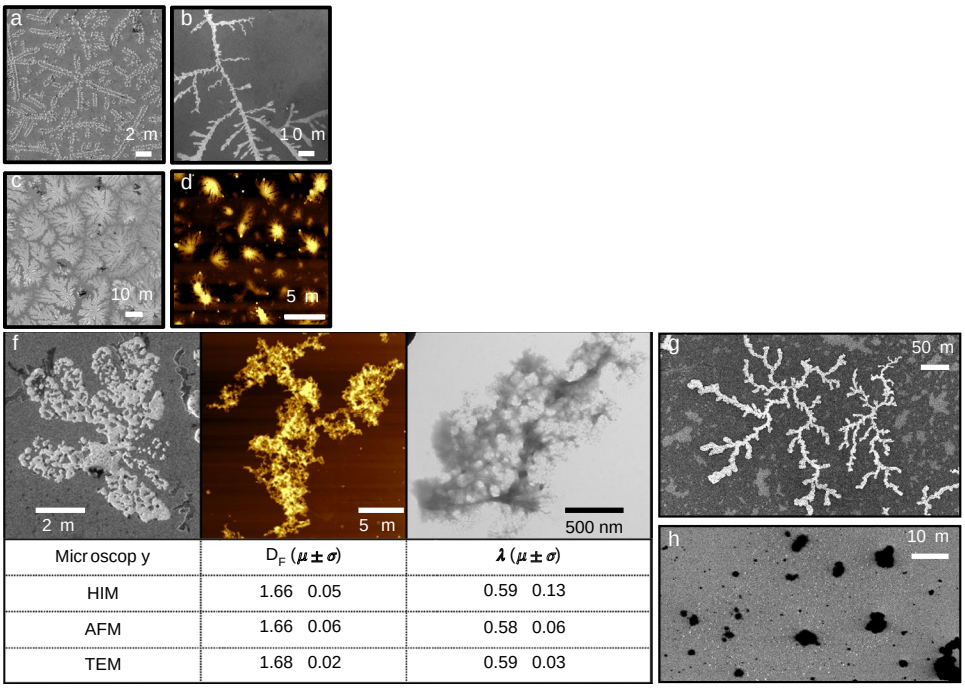


Fig. 2 | Assembly formation, dissolution and inhibition in vitro. a, Using Src kinase, the AtzAM1 can be phosphorylated (pY-AtzAM1) and incubated with AtzCM1-SH2 to form an assembly. The phosphatase (YOP) enzyme can be used for disassembly. **b**,**c**, Assemblies were expected to form (**b**) and dissolve (**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 (highest concentration of ATP to lowest, starting from top to bottom at time 0) **f**,**g**, Assembly formation is highly sensitive to the stoichiometry of the components. Varying the stoichiometry (**f**,**g**) 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**, Fractal-like structure observed by light microscopy. **i**, Fluorescence microscopy image of assembly formed by Alexa Fluor 647-labelled AtzCM1-SH2 and pY-AtzAM1.

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 M and 2 M, respectively, [ATP] 250 M led to complete conversion of monomers to assemblies within 5 min, whereas significantly slower rates of conversion were observed with lower [ATP] (Fig. 2e, Supplementary Fig. 17 and Supplementary Table 2). Visualization of assemblies using optical and fluorescence microscopy (with Alexa-647-labelled AtzC-SH2) revealed the existence of large (10 m) 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 (Supplementary Video 1 and Supplementary Fig. 18).

Apparent hydrodynamic radius (Fig. 2f,g) and polydispersity measured with DLS (Supplementary Figs. 19 and 20) 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 constants K_D for the weaker and tighter interactions were measured as ~40 and ~7 nM, respectively; Supplementary Fig. 10), whereas when one component is present in excess, assembly formation is inhibited, as expected from our branch propagation design model (Fig. 1). Assembly formation by non-stoichiometric concentration combinations with the higher-affinity SH2 domain variant (Fig. 2f,g) indicates that the inhibition caused by an excess of the binding partner is dynamic. Inhibition of assembly formation due to stoichiometric excess can also be overcome in an affinity-dependent manner: the zone of stoichiometries where assembly formation occurs is larger for the higher-affinity SH2 domain variant (Fig. 2g) compared to the lower-affinity variant (Fig. 2f). These results highlight the importance of high affinity in stabilizing the designed kinetically trapped aggregate state: under conditions of uneven stoichiometry (for example, 250 nM AtzA-pY; 1,000 nM AtzC-SH2) and in the absence of kinetic traps, all AtzA components should be bound by an excess of AtzC-SH2 domains, and no



Microscopy	D_F ($\mu \pm \sigma$)	λ ($\mu \pm \sigma$)
HIM	1.66 0.05	0.59 0.13
AFM	1.66 0.06	0.58 0.06
TEM	1.68 0.02	0.59 0.03

Fig. 3 | Assembly formation and characterization with helium ion microscopy, AFM and transmission electron microscopy. All methods reveal fractal-like topologies on a surface. **a–d**, Longer fractal-like structures and branch-like and flower-like structures are seen in helium ion microscopy (HIM) (**a**–**c**) and AFM (**d**). **e**, Representative HIM images for assemblies obtained at different concentrations of pY-AtzAM1 (250 nM–3 M) while maintaining a fixed concentration of AtzCM1-SH2 (2 M). Increasing concentrations of pY-AtzM1 result in larger assemblies with higher fractal dimensions. **f**, D_F and (the fractal dimension and lacunarity of the images) are similar for images obtained from different microscopy techniques. **g**,**h**, HIM images show fractal-like assembly formation with pY-AtzAM1 and AtzCM1-SH2 (**g**), while the Gly-Ser-rich linker-containing variants form globular assemblies under these conditions (**h**).

assemblies should result (expected particle size is 50 nm). This behaviour is observed for the weaker-affinity SH2 variant at this stoichiometry (Fig. 2f). In stark contrast, for the high-affinity SH2 variant (Fig. 2g), we observe micrometre-sized assemblies indicating the presence of aeolotropic kinetic trapping[44] and network formation by clusters of tightly bound AtzA-pY-AtzC-SH2 assemblies (Supplementary Video 1).

Structural characterization of surface-adsorbed assemblies. We next investigated if the dynamic and dendritic structures observed in solution by optical and fluorescence microscopy (Fig. 2h,i) could form fractals on solid surfaces, and if the topology of the surface-adsorbed assemblies could be controlled by varying the component stoichiometry. Due to the substantial increase of surface area derived from fractal patterns, surface-adsorbed fractals at the nanometre–micrometre scale are attractive design targets for applications in many fields, such as catalysis, fractal electronics and the creation of nanopatterned sensors[4,5]. 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 (AFM) reveals striking, intricately textured patterns that coat areas of up to 100 m^2. Various morphologies on the micrometre scale—including rod-like, tree-like, fern-like and petal-like—were observed (Fig. 3a–e); image analysis revealed fractal dimensions between 1.4 and 1.5 (Fig. 3a,b) and to the more diffusion limited aggregation (DLA)-like 1.78 (Fig. 3c,d and Supplementary Figs. 21 and 22). Assembly sizes and fractal dimensions could be tuned by varying the stoichiometry of the components (Fig. 3f), although some heterogeneity in morphologies was present in each sample. At 1:1 stoichiometry of the two components, DLA-like topologies with ~10 m size were observed, whereas more dendritic assemblies were observed when unequal stoichiometry samples were used (Fig. 3f). Similarly, smaller assembly sizes resulted when the concentration of one component became limiting.

Fractal patterns were not observed at any component stoichiometry without the addition of ATP and Src kinase, with unphosphorylated proteins or on drying the buffer (to preclude precipitation-induced assembly formation by the salt in the buffer), demonstrating that fractal structures are formed by designed components (Supplementary Fig. 23). 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. 3g,h and Supplementary Fig. 24). Thus, the surface-induced patterns observed with designed AtzC and AtzA are selectively formed following 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.

Structural characterization of assemblies using cryo-electron tomography. Transmission electron microscopy (TEM) of designed AtzA-AtzC proteins also revealed branching, dendritic networks reminiscent of fractal intermediates observed in biosilica formation[14] (Supplementary Fig. 25). To further 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. 4 and Supplementary Videos 2 and 3).

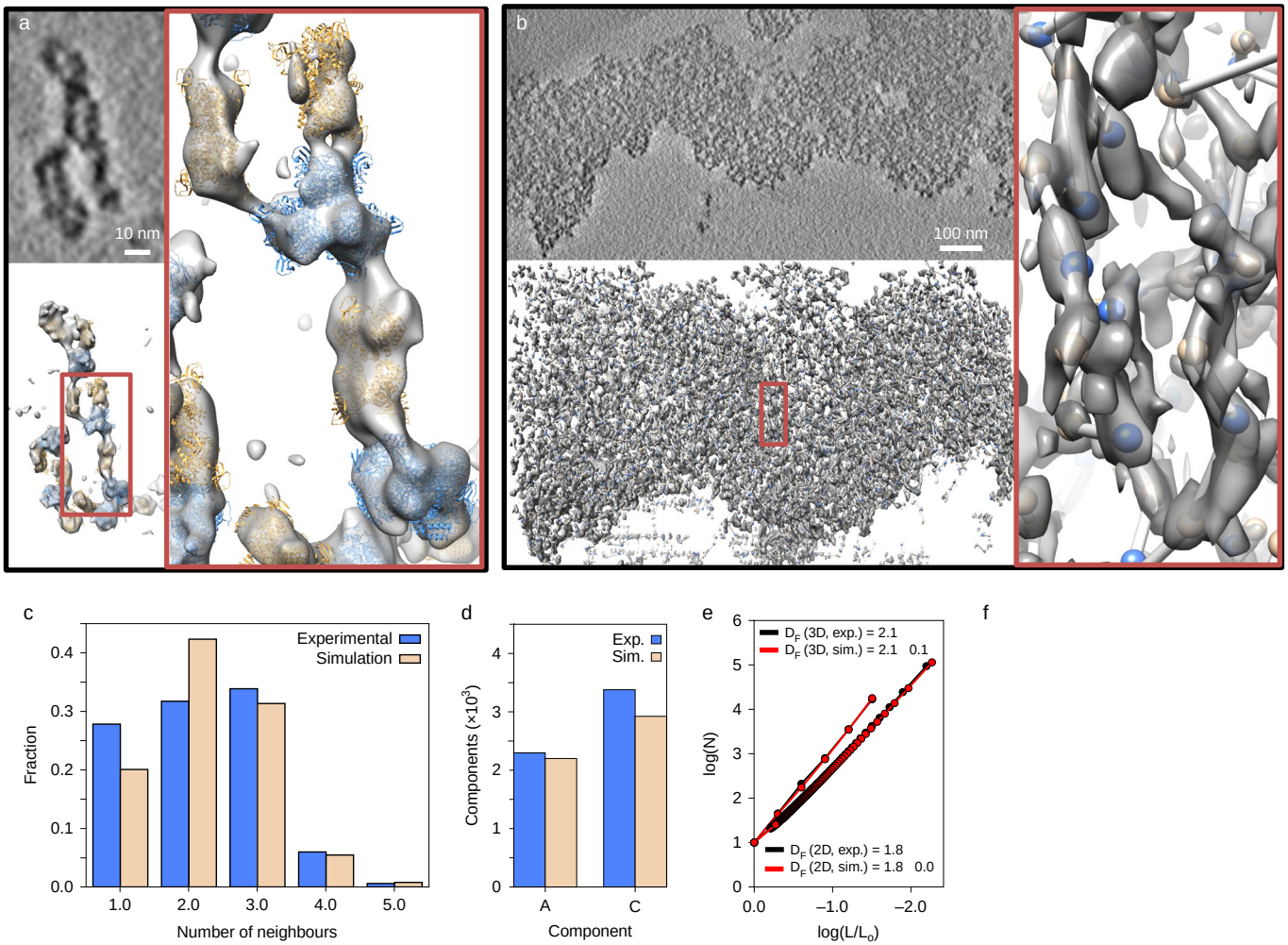


Fig. 4 | Assembly characterization with cryo-ET. a,**b**, Observed topologies in solution for a small (**a**) and large (**b**) assembly, in which the subtomograms were extracted and fit with Rosetta models. For the small assemblies, atomic-resolution models of pY-AtzAM1 (blue) and AtzCM1-SH2 (tan) were fitted to reveal the intercomponent connections along assembly branches. For the large assembly, due to the lower resolvability in this region of the sample, only geometric centres of density were used to assign to individual components (blue and tan spheres; Supplementary Section 3.6). **c**, Spatially proximal neighbour distribution from the cryo-ET-derived images compared to simulated assemblies. **d**, Relative component distribution in the cryo-ET image and from simulations. **e**,**f**, Image analysis (2D), using a box counting method, of the cryo-ET tomography subtomograms converted into 2D projections shows a similar fractal dimension. The standard deviation in the simulation is calculated from 100 simulations composed of ~5,000 components. 3D box counting revealed a similar fractal dimension (slope) between the experimentally observed and simulated assemblies. Parameters for the simulation that match the experimental data are P_{term} 0.1, C_{frac} 1.0 and kT 9.0. Two representative 2D projections with the matching parameters are shown (**f**).

Assemblies generated by mixing 3 M pY-AtzA and 2 M 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 set-up, pY-AtzA and AtzC-SH2 complexes in assembly tomograms were easily identified as density clusters. In contrast, constructs with Gly-Ser-rich linkers connecting the pY and SH2 domain with AtzA and AtzC did not form porous clusters but instead (~90% of the sample) formed large, dense globular clumps (Supplementary Fig. 26) where individual components were not resolvable. These large topology changes on the micrometre scale (as observed by both cryo-ET and HIM) upon conformational flexibility changes at the nanometre scale further reinforce 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. 4a). The topology of the largest, nearly fully interconnected assembly based on electron density (Fig. 4b), consisting of ~6,000 individual protein components, was further analysed and compared with an ensemble of simulated structures with approximately the same number of components. We compared the observed distributions of nearest-neighbour counts for AtzA-pY (Fig. 4c and Supplementary Fig. 27), relative numbers of component types incorporated (Fig. 4d), angular distribution (C-A-C connections, Supplementary Fig. 28) and the observed fractal dimension (Fig. 4e) of the assemblies with ensembles of structures generated using computational modelling (Fig. 4f) and found good agreement between the data and our simulations performed at specific parameter values (P_{term} 0.1, C_{frac} 1.0, kT 9.0). The observed nearest-neighbour distribution for the AtzA-pY component shows that a large majority of these proteins are connected to one, two or three neighbouring 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

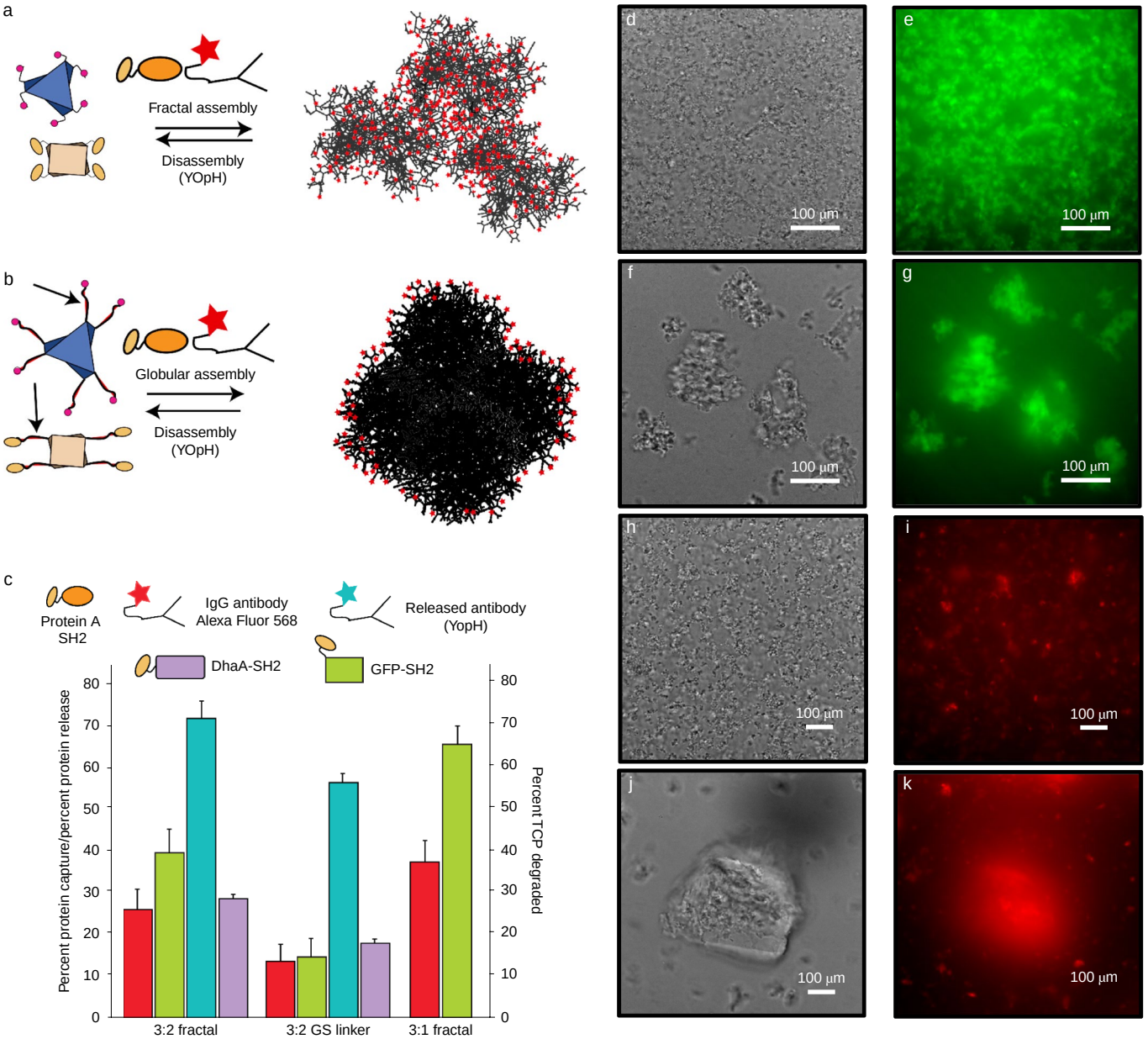


Fig. 5 | Fractal assemblies capture and release greater amounts of cargo compared to globular assemblies. a,**b**, Scheme for the envisioned reversible capture of cargo proteins for fractal (**a**) and globular (**b**) structures. Red stars denote an example cargo protein (antibody). **c**, Percent protein capture was measured for the 3:2 fractal (assemblies obtained with 3 AtzA-pY: 2 AtzC-SH2), 3:2 Gly-Ser (GS) linker (globular assemblies obtained at the same stoichiometry with fusion proteins containing long GS-rich linkers) and 3:1 fractal (assemblies obtained with 3 AtzA-pY: 1 AtzC-SH2). The 3:2 fractal captured more IgG antibody (red bars) and GFP-SH2 (green bars), and degraded more substrate TCP (purple bar; reflecting the higher capture efficiency of enzyme DhaA-Sh2) than the 3:2 GS linker assemblies. In addition, the 3:2 fractal released more captured antibody compared to the 3:2 GS linker when incubated with YopH phosphatase (blue bars). The plotted error is the standard deviation of the percent protein capture/release/TCP degradation obtained from three independent measurements. The error used was a standard deviation of the percent capture/release/TCP degradation over the triplicates. **d**–**g**, Confocal fluorescence microscopy images of the three-component assembly with GFP-SH2 showing the topology of incorporation of GFP-SH2 into the fractal assembly (**d**,**e**) and the incorporation of GFP-SH2 into the globular assembly (**f**,**g**). **h**–**k**, IgG antibody Alexa Fluor 568 incorporation into the fractal assembly (**h**,**i**) and incorporation into the globular assembly (**j**,**k**).

number of AtzA-pY proteins have four AtzC neighbours 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, although a small number of monovalent connections cannot be definitively ruled out in the cryo-ET images (Fig. 4c). We found that the fractal dimensions from the cryo-ET images and simulations (2.1) show good agreement (Fig. 4e). 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 anisotropic nature[32–34] of the underlying protein–protein interactions as encoded in the design approach (Fig. 1). 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 M, an ~125-fold increase compared to their bulk concentration (3 M AtzA-pY and 2 M AtzC-SH2). The particle

density in the fractal assembly is ~70,000 particles m ^{3} whereas the calculated densities of 2D and 3D crystalline lattices of similar volumes are estimated to be ~4,000 and ~40,000 particles m ^{3} (Supplementary Fig. 1). The high particle density in the fractal while maintaining porosity leads to a high effective surface area, a characteristic feature of macroscopic fractal objects such as trees and sponges. Although we could assign individual density clusters to individual components, the thickness of the ice in this region of the sample lowers resolvability and precludes direct measurement of orientation of the AtzA-pY and AtzC-SH2 components with respect to each other for comparison with Rosetta-calculated landscapes (Fig. 1). Although there is significant heterogeneity in assembly sizes (~60% of the proteins adsorbed on the cryo-ET grid are parts of smaller assemblies) and topologies (Supplementary Fig. 29), the observed increase in the effective concentrations concomitant with a high effective surface area with numerous solvent channels (Fig. 4a–f) indicates that induced fractal-like structure formation is a viable strategy to engineer protein assemblies with favourable sponge-like properties.

Functional characterization of designed assemblies. We next investigated if the observed textured, sponge-like topology, resulting in a high surface area:volume ratio in the fractal assembly, endows it with similar enhanced material capture ('soaking up') properties on the nanoscale as observed for macroscopic sponges. We reasoned that the lacunarity ('gappiness') of the fractal structure and use of an excess AtzA-pY component under fractal-forming stoichiometries would lead to several phosphopeptide sites on AtzA being open and accessible. The observed large pore sizes (Fig. 4b) would enable access to these sites for molecular capture of nanometre-sized, macromolecular moieties bearing SH2 domains. In contrast, due to their dense, globular structure, amorphous assemblies generated with Gly-Ser-rich linker-containing components would have less available binding sites, resulting in a lower loading capacity (Supplementary Figs. 24 and 26). To test the molecular capture properties of assemblies, we first used two fusion proteins in which macromolecular cargo proteins were fused to an SH2 domain: SH2-GFP, SH2-DhaA (an engineered DhaA enzyme for the degradation of the groundwater pollutant 1,2,3-trichloropropane, TCP), and measured the amount of cargo proteins captured by fractal and globular assemblies generated using identical amounts of component proteins (Fig. 5). Indeed, fractal assemblies captured greater amounts of cargo, as evidenced by fluorescence (green fluorescent protein, GFP) and enzymatic activity (DhaA) measurements, respectively (Fig. 5). Fluorescence microscopy of SH2-GFP containing assemblies revealed that, as anticipated from cryo-ET studies, the immobilized cargo protein was distributed throughout the assembly and localized to the surface for fractal and globular assemblies, respectively (Fig. 5d–g). To develop a more broadly applicable approach for exploiting the efficient molecular capture properties of fractal assemblies, we generated and utilized a SH2-protein A fusion protein to capture a fluorescent immunoglobulin-G (IgG) antibody. As observed for SH2-GFP and SH2-DhaA, fractal assemblies can efficiently capture this antibody (Fig. 5h–k and Supplementary Fig. 30). Furthermore, incubation of antibody-loaded assemblies with YopH phosphatase enzyme permits release of captured cargo antibodies (Fig. 5). As all full-length IgG antibodies universally have the binding sites for protein A (their Fc-domains), antibody-loaded fractal assemblies should enable (1) efficient molecular capture of a variety of macromolecular and small-molecule antigens and (2) phosphorylation-dependent antibody purification[55–57].

In our design framework, fractal loading capacity is determined by the number and accessibility of open phosphopeptide binding sites in the assembly. Thus, assemblies formed by 3 (AtzA-pY):1 (SH2-AtzC) are expected to have a greater loading capacity than those formed by 3 (AtzA-pY):2 (SH2-AtzC). Indeed, as anticipated, more antibody was captured and released by the former than the latter (Fig. 5c), demonstrating that customized optimization of molecular capture-and-release of specific nanoscale objects should be possible by varying the component stoichiometry to obtain the fractal properties on nanometre to micrometre scales. Finally, we asked if the observed functional advantages of a fractal topology over a globular one would extend to the capture and transport of small molecules within the assembly by measuring the efficacy of atrazine degradation. As cargo we incorporated AtzB—the third pathway enzyme (apart from AtzA and AtzC) required to convert atrazine to the relatively benign metabolite cyanuric acid (Supplementary Figs. 31 to 34). Although both the fractal and globular assemblies appear to be more robustly active under harsh reactions compared to unassembled enzymes (Supplementary Fig. 35), and when immobilized on a Basotect polymer foam (Supplementary Fig. 36), both globular and fractal assemblies are equally active (Supplementary Fig. 37). The significantly small size of atrazine (radius of gyration, R_g 1 nm) and other metabolic pathway intermediates probably allows them to diffuse equally efficiently in either assembly as the smaller solvent channel size in the globular assembly may not be an impediment for a small molecule guest as opposed to macromolecular guest molecules. Constructing fractal-like shapes with smaller sized proteins may allow access to smaller solvent channels.

Discussion

Our results demonstrate an approach by which fusion proteins may be designed to self-assemble into fractal-like morphologies on the 10 nm–10 m 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, anisotropy and geometric degeneracy of binding can be used to controllably generate a broad range of sizes and morphologies of fractal shapes with proteins. In contrast with computational design of integer-dimensional protein assemblies where considerable remodelling of protein surfaces is necessary to meet the exacting geometric requirements for inter-component binding[24,25], our design approach to obtain fractal-like morphologies involves few substitutions on protein surfaces. Instead, design goals are encoding high affinity via fusion of binding domains, and binding anisotropy and geometric degeneracy via short, flexible loops (see Supplementary Discussion). Although we used SH2 domain-pY peptide fusions as the high-affinity 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[58]. Our results show that this rich phase behaviour of proteins[44] 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 for molecular capture, further development in the design of protein-based fractals described here is expected to enable the production of novel classes of bionanomaterials and devices.

Methods

Descriptions of the methods for computational design, multiscale modelling, protein expression and characterization, and structural and functional characterization of assemblies are available in the Supplementary Information.

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary Information files. Raw data used to generate all figures are available via Figshare (https://figshare.com/projects/Stimulus-responsive_Self-Assembly_of_Protein-Based_Fractals_by_

Computational_Design/61976), and cryo-ET maps are available in the EMDB (accession codes EMD-20062 and EMD-20063).

Code availability

Scripts and input files used for Rosetta simulations and code used for coarse-grained simulations are available from a GitHub repository (https://github.com/sagark101/Nchem-Fractal-Assembly).

Received: 21 October 2018; Accepted: 29 April 2019;
Published online: 17 June 2019

References

1. Mandelbrot, B. B. *The Fractal Geometry of Nature* (W. H. Freeman & Company, 1982).
2. Stanley, H. E. & Meakin, P. Multifractal phenomena in physics and chemistry. *Nature* 335, 405–409 (1988).
3. Losa, G. A. *Fractals in Biology and Medicine* Vol. IV (Birkhäuser, 2005).
4. Fairbanks, M. S., McCarthy, D. N., Scott, S. A., Brown, S. A. & Taylor, R. P. Fractal electronic devices: simulation and implementation. *Nanotechnology* 22, 365304 (2011).
5. Soleymani, L., Fang, Z. C., Sargent, E. H. & Kelley, S. O. Programming the detection limits of biosensors through controlled nanostructuring. *Nat. Nanotechnol.* 4, 844–848 (2009).
6. Ge, J., Lei, J. D. & Zare, R. N. Protein–inorganic hybrid nanoflowers. *Nat. Nanotechnol.* 7, 428–432 (2012).
7. Zhang, P. C. & Wang, S. T. Designing fractal nanostructured biointerfaces for biomedical applications. *ChemPhysChem* 15, 1550–1561 (2014).
8. Lim, B. et al. Pd–Pt bimetallic nanodendrites with high activity for oxygen reduction. *Science* 324, 1302–1305 (2009).
9. Cerofolini, G. F., Narducci, D., Amato, P. & Romano, E. Fractal nanotechnology. *Nanoscale Res. Lett.* 3, 381–385 (2008).
10. Newkome, G. R. et al. Nanoassembly of a fractal polymer: a molecular 'Sierpinski hexagonal gasket'. *Science* 312, 1782–1785 (2006).
11. Shang, J. et al. Assembling molecular Sierpinski triangle fractals. *Nat. Chem.* 7, 389–393 (2015).
12. Newkome, G. R. & Moorefield, C. N. From 1 3 dendritic designs to fractal supramacromolecular constructs: understanding the pathway to the Sierpinski gasket. *Chem. Soc. Rev.* 44, 3954–3967 (2015).
13. 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).
14. Tikhomirov, G., Petersen, P. & Qian, L. Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* 552, 67–71 (2017).
15. Zhang, F., Nangreave, J., Liu, Y. & Yan, H. Reconfigurable DNA origami to generate quasifractal patterns. *Nano Lett.* 12, 3290–3295 (2012).
16. Rothemund, P. W., Papadakis, N. & Winfree, E. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol.* 2, e424 (2004).
17. Astier, Y., Bayley, H. & Howorka, S. Protein components for nanodevices. *Curr. Opin. Chem. Biol.* 9, 576–584 (2005).
18. Murr, M. M. & Morse, D. E. Fractal intermediates in the self-assembly of silicatein filaments. *Proc. Natl Acad. Sci. USA* 102, 11657–11662 (2005).
19. Khire, T. S., Kundu, J., Kundu, S. C. & Yadavalli, V. K. The fractal self-assembly of the silk protein sericin. *Soft Matter* 6, 2066–2071 (2010).
20. Lomander, A., Hwang, W. M. & Zhang, S. G. Hierarchical self-assembly of a coiled-coil peptide into fractal structure. *Nano Lett.* 5, 1255–1260 (2005).
21. 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 disulfide bond formation. *Macromolecules* 38, 3909–3916 (2005).
22. Li, B. et al. Nonequilibrium self-assembly of pi-conjugated oligopeptides in solution. *ACS Appl. Mater. Interfaces* 9, 3977–3984 (2017).
23. McManus, J. J., Charbonneau, P., Zaccarelli, E. & Asherie, N. The physics of protein self-assembly. *Curr. Opin. Colloid* 22, 73–79 (2016).
24. King, N. P. et al. Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science* 336, 1171–1174 (2012).
25. Hsia, Y. et al. Design of a hyperstable 60-subunit protein dodecahedron [corrected]. *Nature* 535, 136–139 (2016).
26. Suzuki, Y. et al. Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* 533, 369–373 (2016).
27. Sinclair, J. C., Davies, K. M., Venien-Bryan, C. & Noble, M. E. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nat. Nanotechnol.* 6, 558–562 (2011).
28. 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. USA* 98, 2217–2221 (2001).
29. Zhang, J., Zheng, F. & Grigoryan, G. Design and designability of protein-based assemblies. *Curr. Opin. Struct. Biol.* 27, 79–86 (2014).
30. Subramanian, R. H. et al. Self-assembly of a designed nucleoprotein architecture through multimodal interactions. *ACS Cent. Sci.* 4, 1578–1586 (2018).
31. Churchfield, L. A. & Tezcan, F. A. Design and construction of functional supramolecular metalloprotein assemblies. *Acc. Chem. Res.* 52, 345–355 (2019).
32. Sontz, P. A., Song, W. J. & Tezcan, F. A. Interfacial metal coordination in engineered protein and peptide assemblies. *Curr. Opin. Chem. Biol.* 19, 42–49 (2014).
33. 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).
34. Ringler, P. & Schulz, G. E. Self-assembly of proteins into designed networks. *Science* 302, 106–109 (2003).
35. Lindenmayer, A. Mathematical models for cellular interactions in development. II. Simple and branching filaments with 2-sided inputs. *J. Theor. Biol.* 18, 300–315 (1968).
36. Lindenmayer, A. Mathematical models for cellular interactions in development. I. Filaments with 1-sided inputs. *J. Theor. Biol.* 18, 280–299 (1968).
37. Glotzer, S. C. & Solomon, M. J. Anisotropy of building blocks and their assembly into complex structures. *Nat. Mater.* 6, 557–562 (2007).
38. Zhang, Z. & Glotzer, S. C. Self-assembly of patchy particles. *Nano Lett.* 4, 1407–1413 (2004).
39. Kartha, M. J. & Sayeed, A. Phase transition in diffusion limited aggregation with patchy particles in two dimensions. *Phys. Lett. A* 380, 2791–2795 (2016).
40. Nicolas-Carlock, J. R., Carrillo-Estrada, J. L. & Dossetti, V. Fractality á la carte: a general particle aggregation model. *Sci. Rep.* 6, 19505 (2016).
41. Guesnet, E., Dendievel, R., Jauffres, D., Martin, C. L. & Yrieix, B. A growth model for the generation of particle aggregates with tunable fractal dimension. *Physica A* 513, 63–73 (2019).
42. Mansbach, R. A. & Ferguson, A. L. Patchy particle model of the hierarchical self-assembly of pi-conjugated optoelectronic peptides. *J. Phys. Chem. B* 122, 10219–10236 (2018).
43. Bianchi, E., Tartaglia, P., Zaccarelli, E. & Sciortino, F. Theoretical and numerical study of the phase diagram of patchy colloids: ordered and disordered patch arrangements. *J. Chem. Phys.* 128, 144504 (2008).
44. Lomakin, A., Asherie, N. & Benedek, G. B. Aeolotopic interactions of globular proteins. *Proc. Natl Acad. Sci. USA* 96, 9465–9468 (1999).
45. Vacha, R. & Frenkel, D. Relation between molecular shape and the morphology of self-assembling aggregates: a simulation study. *Biophys. J.* 101, 1432–1439 (2011).
46. Bianchi, E., Tartaglia, P., La Nave, E. & Sciortino, F. Fully solvable equilibrium self-assembly process: fine-tuning the clusters size and the connectivity in patchy particle systems. *J. Phys. Chem. B* 111, 11765–11769 (2007).
47. Yan, Y., Huang, J. & Tang, B. Z. Kinetic trapping—a strategy for directing the self-assembly of unique functional nanostructures. *Chem. Commun.* 52, 11870–11884 (2016).
48. Wackett, L. P., Sadowsky, M. J., Martinez, B. & Shapir, N. Biodegradation of atrazine and related s-triazine compounds: from enzymes to field studies. *Appl. Microbiol. Biotechnol.* 58, 39–45 (2002).
49. Kaneko, T. et al. Superbinder SH2 domains act as antagonists of cell signaling. *Sci. Signal.* 5, ra68 (2012).
50. Yang, L. et al. Computation-guided design of a stimulus-responsive multienzyme supramolecular assembly. *Chembiochem* 18, 2000–2006 (2017).
51. Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* 77, 363–382 (2008).
52. Pellegrini, M., Wukovitz, S. W. & Yeates, T. O. Simulation of protein crystal nucleation. *Proteins* 28, 515–521 (1997).
53. Masters, B. R. Fractal analysis of the vascular tree in the human retina. *Annu. Rev. Biomed. Eng.* 6, 427–452 (2004).
54. Witten, T. A. & Sander, L. M. Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.* 47, 1400–1403 (1981).
55. Swartz, A. R. & Chen, W. SpyTag/SpyCatcher functionalization of E2 nanocages with stimuli-responsive Z-ELP affinity domains for tunable monoclonal antibody binding and precipitation properties. *Bioconjug. Chem.* 29, 3113–3120 (2018).
56. Bilgicer, B. et al. A non-chromatographic method for the purification of a bivalently active monoclonal IgG antibody from biological fluids. *J. Am. Chem. Soc.* 131, 9361–9367 (2009).
57. Handlogten, M. W., Stefanick, J. F., Deak, P. E. & Bilgicer, B. Affinity-based precipitation via a bivalent peptidic hapten for the purification of monoclonal antibodies. *Analyst* 139, 4247–4255 (2014).
58. Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* 11, 899–904 (2015).

Acknowledgements

The authors acknowledge support from the NSF (1330760 to S.D.K. and L.W.; DGE-1433187 to N.E.H.; 1429062 to S.D.K.) and the NIH (R01GM080139 to M.C.). Cryoelectron microscopy was supported by the Rutgers New Jersey CryoEM/ET Core Facility. The authors thank J. Chodera for providing Src kinase and YopH phosphatase plasmids, V. Nanda, K.-B. Lee, G. Montelione, H. Cho, M. Liu, A. Permaul, O. Dineen,

I. Patel and R. Patel for experimental assistance, and E. Tinberg, V. Nanda and D. Baker for helpful discussions.

Author contributions

N.E.H. and W.A.H. contributed equally to this work. N.E.H., W.A.H. and S.D.K. designed the research. W.A.H. developed the computational methodology for design and image analyses. N.E.H., D.Z., M.E.S. and M.K. expressed, purified and assayed proteins. N.E.H., V.M., T.G. and L.C.F. performed HIM. N.E.H., M.P., P.R. and S.-H.L. performed optical and fluorescence microscopy. D.Z. performed DLS, M.K. performed BLI and L.Y. performed TEM. I.M.-B. performed confocal microscopy. A.G.D. and L.P.W. performed polymer foam immobilization and activity assays. W.D., M.B., M.C. and W.A.H. performed cryo-ET and analyses. S.D.K., N.E.H. and W.A.H. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41557-019-0277-y.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to S.D.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.