JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. 2018, 140, 6391−6399
pubs.acs.org/JACS

Synthesis of Cyclic Megamolecules

Justin A. Modica, Yao Lin,[†] and Milan Mrksich*

Departments of Chemistry and Biomedical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60201, United States

Supporting Information

ABSTRACT: This paper describes the synthesis of giant cyclic molecules having diameters of 10−20 nm. The molecules are prepared through the reactions of a fusion protein building block with small molecule linkers that are terminated in irreversible inhibitors of enzyme domains present in the fusion. This building block has N-terminal cutinase and C-terminal SnapTag domains that react irreversibly with *p*-nitrophenyl phosphonate (pNPP) and benzylguanine (BG) groups, respectively. We use a bis-BG and a BG-pNPP linker to join these fusion proteins into linear structures that can then react with a bis-pNPP linker that joins the ends into a cyclic product. The last step can occur intramolecularly, to give the macrocycle, or intermolecularly with another equivalent of linker, to give a linear product. Because these are coupled first- and second-order processes, an analysis of product yields from reactions performed at a range of linker concentrations gives rate constants for cyclization. We determined these to be 9.7×10^{-3} s^{-1}, 2.3×10^{-3} s^{-1}, and 8.1×10^{-4} s^{-1} for the dimer, tetramer, and hexamer, respectively. This work demonstrates an efficient route to cyclic macromolecules having nanoscale dimensions and provides new scaffolds that can be generated using the megamolecule approach.

INTRODUCTION

The synthesis of megamolecules, which we define as discrete (monodisperse) molecular structures having masses exceeding 100 000 Da, and with dimensions greater than 10 nm, is challenging. Strategies based on oligonucleotide assembly—as illustrated by the range of structures prepared as DNA origami—represent an exciting approach to generating well-defined megamolecules.[1] The use of proteins as building blocks, by extension, is significant because of their wide structural diversity and because proteins display enormously more functional properties than oligonucleotides. The assembly of proteins, however, is constrained by their chemical complexity and the lack of general reactions (unlike the hybridization and ligation reactions used for DNA) that can efficiently join them into larger structures. In the present paper, we build on earlier work that demonstrated the use of an enzyme-mediated protein assembly method and report the design and construction of cyclic macromolecules having molecular weights between 100−300 kDa and diameters of 10−20 nm. We also characterize the rates for cyclization in these large structures and introduce the SnapTag enzyme as a new orthogonal building block that expands the range of megamolecule structures that can be prepared.

Pioneering work by Schulz and Wagner demonstrated supramolecular protein assemblies based on protein−ligand interactions to give extended lattices[2−4] and nanorings.[5,6] These structures are typically polydisperse and, because they are based on noncovalent connections, are not stable to manipulation and dilution.[7] Yeates, Baker, and others have used protein−protein interactions to construct cages,[8−10] polyhedra,[11,12] modified viral particles,[13,14] and artificial enzymes.[15] These approaches harness multivalent contacts to drive the formation of discrete structural assemblies, yet they often require the evaluation of a large number of designs to realize the desired products.

Methods based on the covalent assembly of protein building blocks, in contrast, could be used to prepare structurally well-defined and stable structures. To be useful, these approaches require chemistries that are efficient, regioselective, and that use functional groups that are present (or can easily be incorporated in) proteins. Recent work using sortase enzymes,[16] "dock-and-lock" methods,[17] engineered bacterial adhesins,[18] and split inteins[19] are important advances toward this goal, while nonsense codon engineering to site-specifically introduce non-natural amino acids has enabled the incorporation of functional groups for chemoselective ligations including azides, alkynes, aldehydes and others.[20−22] In the latter case, the preparation of modified proteins can be challenging and the subsequent bimolecular conjugation reactions can require many hours or days to prepare products in acceptable yield.[23] Often, rate constants for these ligation reactions are modest and the high molecular weight and limited solubility of proteins limits reaction concentrations to the micromolar regime.[24] This, in turn, translates to observed reaction rates that are 3−4 orders of magnitude slower than those of the corresponding small molecule reactions.

To overcome the poor rates for covalent joining of proteins, we reported a method based on the reaction of enzymes with mechanism-based irreversible inhibitors. With this strategy, the inhibitors first bind to the active site of the protein, overcoming the slow rates that would characterize the corresponding bimolecular reaction. Further, because the reactions are promoted by a specific active site geometry and occur at a

Received: March 12, 2018
Published: May 3, 2018

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

single active site residue, they undergo little reaction at other sites, giving defined products. We reported the stepwise synthesis of a 240 kDa molecule prepared from the assembly of five HaloTag-cutinase fusion proteins with four bifunctional linkers.[25] HaloTag is an engineered haloalkane dehalogenase that is irreversibly inhibited by primary alkyl chlorides through a nucleophilic displacement reaction with an aspartic acid residue in its active site.[26] Cutinase is a serine esterase that is irreversibly inhibited by phosphonate ligands via phosphonylation of an active site serine residue.[27]

In the present paper we demonstrate the use of this megamolecule approach to prepare large cyclic molecules. The route begins with the preparation of a linear oligomer of fusion proteins and then treats this intermediate with a bifunctional linker designed to react with each of the terminal domains to give a cyclized product. We demonstrate the preparation of cyclic molecules having molecular weights between 90 and 280 kDa. Further, we characterize the first order rate constants for the cyclization reactions. We find that they are in the range of those observed for cyclization of small molecules and that there is a modest decrease in rate as the cyclic molecules increase in size. This work establishes a route toward cyclic molecules that are difficult to prepare through other means and expands the diversity of structural scaffolds that can be prepared with the megamolecule approach.

RESULTS AND DISCUSSION

Approach. We used two enzyme-inactivator pairs to assemble the cyclic molecules (Figure 1). The first is cutinase,

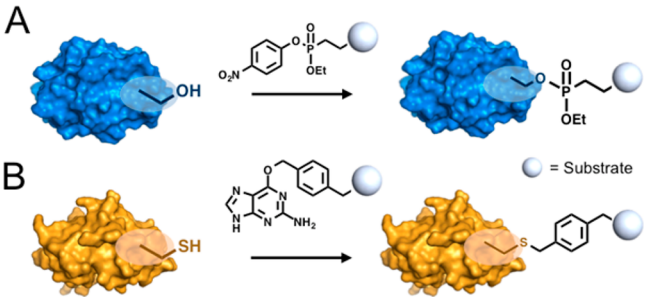

Figure 1. We used two enzyme−inhibitor reactions to assemble fusion proteins into cyclic megamolecules. (A) Cutinase, a serine esterase, reacts selectively and irreversibly with electrophilic phosphonate derivatives at its nucleophilic Ser120 residue to yield serine phosphonate esters. This results in a stable protein−substrate complex. (B) SnapTag, an O^6-alkylguanine alkyltransferase reacts irreversibly with benzylguanine (BG) derivatives at its catalytic Cys145.

which we described earlier, and is covalently inhibited by a *p*-nitrophenyl phosphonate (pNPP) group.[28] The second is SnapTag, an engineered 19 kDa O^6-alkylguanine alkyltransferase developed by Johnsson and co-workers, that reacts with benzylguanine (BG) derivatives at its catalytic Cys145 residue to yield a stable thiobenzyl adduct.[29] These substrate-enzyme reactions do not require cofactors and proceed rapidly with observed rate constants up to 10^4 M^{-1} s^{-1}, and therefore can be rapidly carried out at low concentration of protein. Furthermore, both enzymes are relatively small (<30 kDa), are efficiently expressed in *E. coli* hosts, and both are quite soluble and stable for reasonable periods of storage. We constructed an expression plasmid based on pET21d that encodes an N-terminal cutinase fused to a C-terminal SnapTag with an $(EAAAK)_4$ linker and a C-terminal 6× His-tag, to yield a construct we refer to as **CS** (for cutinase-SnapTag wherein the bolded "**C**" denotes the N-terminal fusion partner). We selected this helical linker because it has been shown to more effectively separate fusion domains than do flexible polypeptide linkers.[30] We expressed this construct in Origami B(DE3) *E. coli* with induction by IPTG. Yields were typically 4−6 mg/L of culture after isolation by cobalt immobilized metal chelate affinity chromatography (Co-IMAC) and size-exclusion chromatography (SEC). We synthesized two homobifunctional linkers that could either symmetrically join two SnapTag domains or two cutinase domains, and we also synthesized a heterobifunctional linker that can join one of each. The syntheses of these linkers were based on straightforward manipulation and are shown in Figure 2. We used water-soluble polyethylene glycol (PEG) precursors to construct a bis-BG linker **1** through standard amide coupling of 4-aminomethyl-BG[31] to *O*,*O*′-bis(carboxyethyl)-dodecaethylene glycol and **4**, a bis(ethyl *p*-nitrophenyl phosphonate) linker in 3 steps from dodecaethylene glycol. We prepared the heterobifunctional linker **8** from two amino-terminal fragments bearing each functional group. Each linker contains ≥12-ethylene glycol repeats as we have shown previously that this length is sufficient to span the active sites of the tethered proteins.

Our synthetic route began by linking the SnapTag domains of the CS monomer with **1** to yield a symmetric dimer with terminal cutinase domains. This step was followed by treatment of the dimer with **4** to afford the cyclic product. We chose this approach for two reasons. First, had we chosen to synthesize a dimer of the form **C**S−**C**S using heterobifunctional linker **8**, it would have been necessary to "protect" the first cutinase domain with a blocking ligand to prevent early cyclization of an intermediate or undesired oligomerization. Second, for the structures described later in the manuscript, we can purify intermediates using SEC, which is straightforward, but requires a significant mass difference between the product and reactants. By adding two fusion proteins to both ends of the symmetrical intermediates, our synthesis aids in their purification. Both of these principles are important to the retrosynthetic analysis of related megamolecule structures. We note that in early approaches, we employed a synthesis that relied on coupling of two HaloTag domains and found those reactions to proceed with lower yields, for reasons that we still do not understand.

Cyclic Dimer. We first demonstrate the synthesis of a cyclic dimer molecule prepared from two **C**S fusion proteins and two bifunctional linkers (Figure 3). We started by treating the recombinant **C**S monomer (50 μM) in Tris-buffered saline (TBS; 25 mM Tris-HCl, 150 mM NaCl, pH 7.4) with 0.5 equiv, a stoichiometric amount, of bis-BG linker **1** to give the symmetrical dimer **C**S−S**C**. The reaction was allowed to proceed for 15 min and then purified by SEC to give the desired product in 40% yield. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of **C**S−S**C** showed a band near 90 kDa, consistent with the calculated mass of 92 kDa. We then treated this intermediate at 500 nM in TBS with 1.5 equiv of bisphosphonate linker **4**. We isolated the major reaction peak using SEC and compared this product to **C**S and **C**S−S**C** by SDS-PAGE (Figure 3B). Both **C**S and **C**S−S**C** appear near the expected molecular weights of 45 and 92 kDa respectively while the apparent cyclic species, $(\mathbf{CS}-\mathbf{SC})_{cyc}$, migrates with lower mobility than the linear precursor. Comparison of the size exclusion chromatographs of **C**S−S**C** and the product of the cyclization reaction showed, for the

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

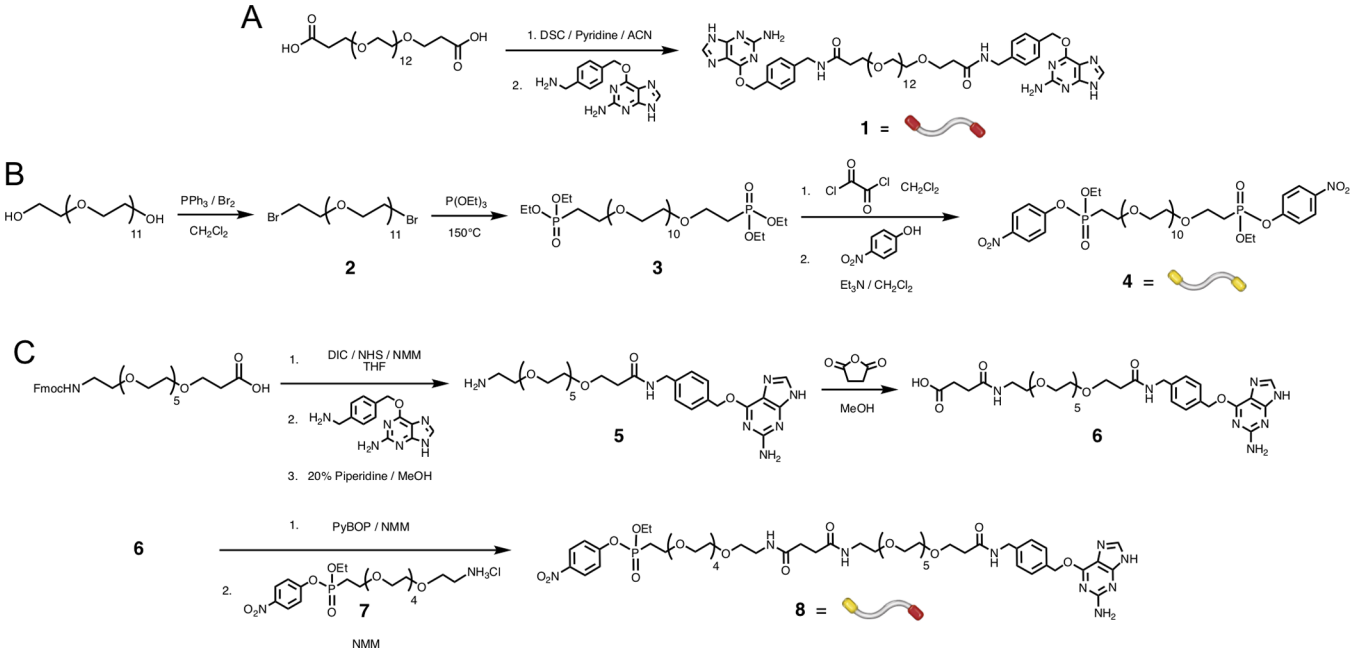

Figure 2. Synthetic scheme for the preparation of linkers used in the assembly of cyclic megamolecules. (A) The synthesis of a homobifunctional linker **1** with terminal BG groups for joining SnapTag domains. Here the red ends of the stylized linker represent BG groups. (B) The synthesis of symmetric linker **4** with terminal ethyl *p*-nitrophenyl phopshonate groups for joining cutinase domains. The yellow ends of the stylized linker represent pNPP groups. (C) Synthetic scheme for heterobifunctional linker **8** used later in the work to prepare larger oligomers of **CS**.

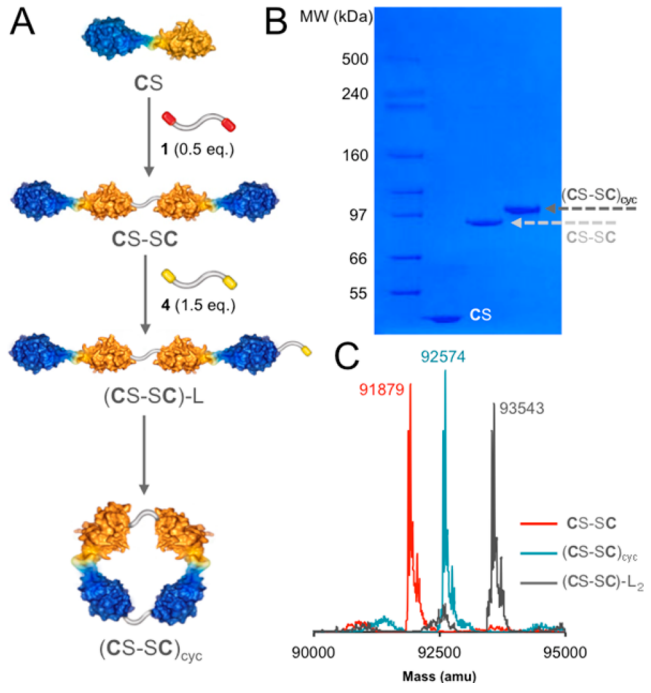

Figure 3. Synthetic scheme and characterization of linear and cyclic forms of the **CS** dimer, **CS**−**SC**. (A) **CS** is first reacted with 0.5 equiv of linker **1** to dimerize the SnapTag domains of two **CS** proteins. This dimer is then treated with linker **4** to afford first a monofunctionalized intermediate, (**CS**−**SC**)−L, which then undergoes intramolecular reaction to afford the cyclic dimer $(\mathbf{CS}-\mathbf{SC})_{cyc}$. (B) SDS-PAGE characterization of the **CS** monomer; and linear and cyclic dimer. The monomer and linear dimer show bands near the calculated molecular weights of 45 and 92 kDa, whereas the cyclic species shows a band at higher apparent molecular weight than the calculated value of 93 kDa. (C) Deconvoluted ESI-MS spectra of linear, cyclic and linear double functionalized forms of the protein.

latter, a peak having a longer retention time than that for the linear precursor, indicating a species with a smaller hydrodynamic radius than the starting material (Figure S1).

Another possible product would derive from reaction of 2 equiv of **4** with **CS**−**SC** to give a symmetric molecule that is then unable to cyclize. To verify that the lower mobility band was indeed the cyclic species and not the bis-reacted protein, we intentionally prepared the latter species, (**CS**−**SC**)−L_2, by treating **CS**−**SC** with a 50-fold excess of **4** and then purified this species using a micro gel filtration column. We then determined the masses of **CS**−**SC**, the apparent cyclic product, and (**CS**−**SC**)−L_2 via electrospray ionization mass spectrometry (ESI-MS). For **CS**−**SC**, we observed a deconvoluted mass of 91 879 Da, in excellent agreement with the calculated mass of 91 878 Da. The spectra of the putative cyclic product had a deconvoluted mass of 92 574 Da, which corresponds exactly to the expected mass shift of ΔMW = 696 Da from the linear precursor resulting from the addition of 1 equiv of **4** less the two *p*-nitrophenylates that are displaced in the cyclization reaction. For the dimer that was reacted with 50 equiv of **4**, we obtained a deconvoluted mass of 93 543 Da, corresponding to the addition of 2 equiv of **4** less the two liberated pNPP groups at ΔMW = 1664 Da from the linear starting material. These data are consistent with the proposed structures.

Partitioning between Cyclization and Oligomerization. The cyclization reaction described above can also give products that derive from intermolecular reaction of two **CS**−**SC** proteins to give the longer tetrameric molecule, **CS**−**SC**−**CS**−**SC**. This partitioning between intramolecular and intermolecular reactivity is a common feature of cyclization reactions and the dependence of the relative yield for the two products on the concentration of the linker allows determination of the first order rate constant for the cyclization reaction. For example, in our first attempts at preparing the cyclic dimer, we performed reactions with **CS**−**SC** at a

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

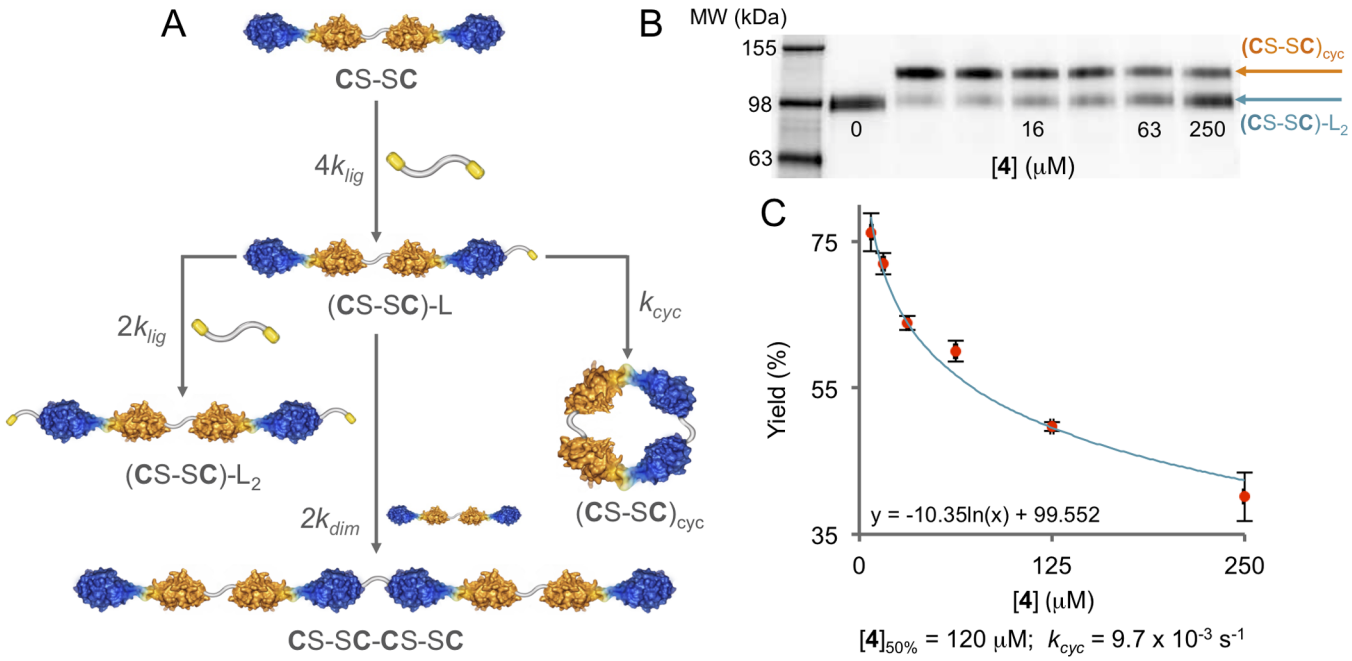

Figure 4. Kinetic partitioning in the reaction of CS−SC with **4** and the determination of the intramolecular rate constant for cyclization. (A) Upon reaction with an equivalent of **4**, CS−SC is converted to a monofunctionalized intermediate (CS−SC)−L that can partition into three reaction channels: one with another equivalent of **4** to yield the double functionalized product (CS−SC)−L_2, one with another equivalent of CS−SC to yield a CS−SC dimer, and an intramolecular channel wherein the free cutinase domain reacts with the pendant phosphonate to generate the cyclic species with first-order rate constant k_{cyc}. (B) The dependence of yield of cyclic dimer on the concentration of **4** can be determined by treating 500 nM aliquots of a dye-labeled dimer with varying linker concentrations, resolving the products of reaction by SDS-PAGE and quantifying the bands using gel fluorimetry. (C) These yields are then plotted vs [**4**] and fit to a logarithmic function. This function is then used to determine the concentration of **4** at which the linear and cyclic species are produced in equal amounts. This concentration is then used to determine k_{cyc} from the relationship in eq 3.

concentration of 5 μM. At this higher concentration, we found that the initially formed (CS−SC)−L led to an appreciable amount of the CS−SC dimer due to the intermolecular addition of an equivalent of CS−SC to (CS−SC)−L.

In Figure 4A, we outline the kinetic scheme of the cyclization reaction. The linker **4** first reacts with CS−SC to give an intermediate (CS−SC)−L that can then react by three distinct pathways. First, a second equivalent of **4** can react with the remaining cutinase domain to yield (CS−SC)−L_2; this reaction is described by a second-order rate constant $2k_{lig}$. This rate constant contains the statistical factor 2 due to the number of reactive phosphonates per linker molecule. Second, reaction of the monofunctionalized intermediate with another equivalent of CS−SC will form CS−SC−CS−SC in an intermolecular reaction with rate constant $2k_{dim}$. Finally, an intramolecular reaction of the pendant phosphonate ligand with the internal cutinase domain gives the desired cyclic product. This reaction is described by a first-order rate constant k_{cyc}. As described earlier, we found that reaction concentrations of CS−SC at or below 500 nM were sufficient to effectively eliminate dimer formation. Under these conditions, (CS−SC)−L largely partitions into only two of the three reaction channels. Accordingly, the following equations approximately describe rate of formation of both (CS−SC)−L_2 and $(CS-SC)_{cyc}$:

$$\text{Rate of formation of } (CS-SC)-L_2 = 2k_{lig}[(CS-SC)-L][L] \quad (1)$$

$$\text{Rate of formation of } (CS-SC)_{cyc} = k_{cyc}[(CS-SC)-L] \quad (2)$$

It follows that when the two products are formed in equal amounts, the rates of the two reactions are equal and therefore

$$k_{cyc} = 2k_{lig}[L] \quad (3)$$

Using this relationship, we can then determine the rate constant for cyclization k_{cyc} by determining the second-order rate constant $2k_{lig}$, and the concentration of **4** at which the two products are formed in equal amounts.

We first determined $2k_{lig}$, the rate constant for the reaction between (CS−SC) and **4**. To eliminate the possibility of the competing cyclization reaction interfering in the determination of this rate constant, we used site directed mutagenesis to introduce a mutation (S120A) into the cutinase-SnapTag construct. This mutation renders the cutinase domain catalytically inactive and only allows reaction to proceed at the SnapTag domain. We expressed and purified this mutant (C^XS), treated it with 50 equiv of linker **1** to form a BG-functionalized monomer, C^XS−L, purified this intermediate via SEC and then reacted the complex with an equivalent of CS to form the monofunctional dimer (C^XS−SC). We then treated purified C^XS−SC (500 nM), with **4** (50 μM) and followed the release of *p*-nitrophenol at 401 nm using UV−vis absorption spectroscopy. We used standard analyses to determine an effective second order rate constant of $2k_{lig}$ = 81 ± 6 M^{-1} s^{-1} (Figure S2). We then determined the concentration at which the cyclic and linear species are formed in equal amounts by treating a dye-labeled CS−SC at a concentration of 500 nM with **4** in concentrations ranging from 8 to 250 μM for 24 h to ensure complete reaction and then resolved the products with SDS-PAGE. The CS−SC used in these reactions was labeled

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

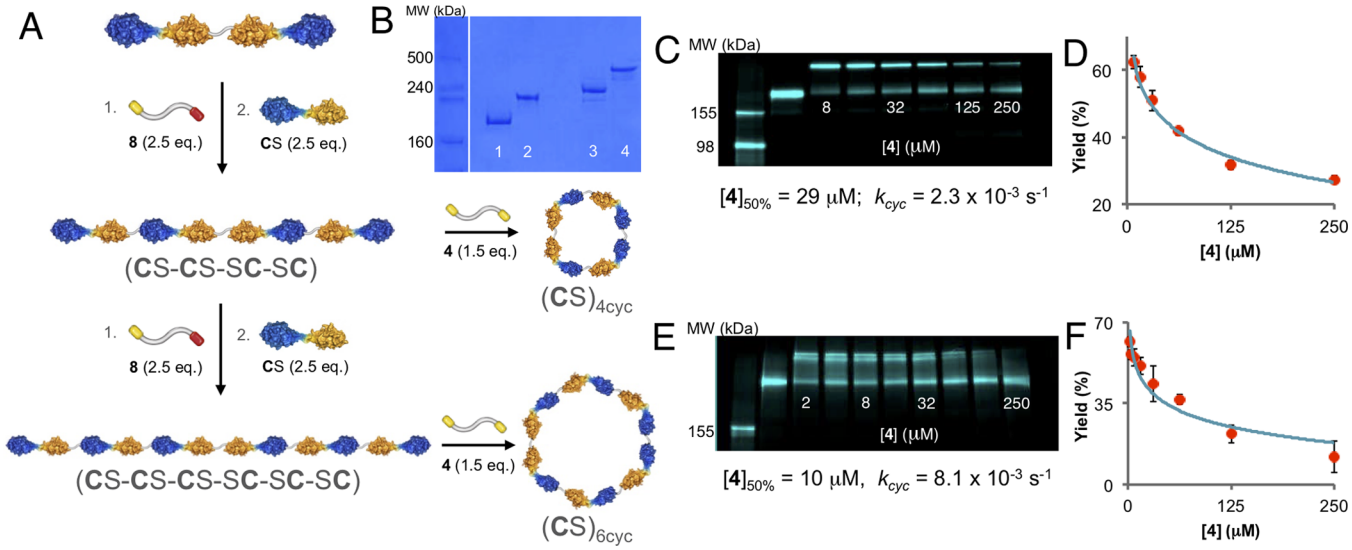

Figure 5. Preparation of longer oligomers of CS and determination of their rate constants for cyclization. (A) A tetramer of CS, CS−CS−SC−SC, is prepared by treating the dimer with 2.5 equiv of **8** followed by 2.5 equiv of CS. A hexamer, CS−CS−CS−SC−SC−SC, is prepared in the same fashion from the tetramer. These linear oligomers are then cyclized using **4**. (B) SDS-PAGE analysis of the purified products of these reactions. Lanes 1 and 2 show the linear and cyclic tetramer, and lanes 3 and 4, the linear and cyclic hexamer, respectively. The linear products show bands near the calculated molecular weights of 185 and 278 kDa respectively, whereas the cyclic species migrate with lower mobility more pronounced than in the dimer case. (C) Determination of dependence of tetramer cycle yield on [**4**] by fluorescent SDS-PAGE. (D) Yields of cyclic tetramer plotted vs [**4**] and fit to a logarithmic function to determine the concentration of linker where products partition equally. (E,F) The same analysis as in D and E but for the hexamer.

with an AlexaFluor 488 dye and purified by SEC. The fluorescent label allowed us to quantitate the products of reaction with gel electrophoresis followed by fluorimetry of the bands using a gel scanner (Figure 4B). Yields for the cyclic products were determined using the following equation:

$$\%\text{Yield } (\mathbf{CS}{-}\mathbf{SC})_{cyc} = (I_{cyc}/(I_{df} + I_{cyc})) \times 100 \quad (4)$$

where I_{cyc} and I_{df} are the total fluorescence intensities for the $(\mathbf{CS}{-}\mathbf{SC})_{cyc}$ and $(\mathbf{CS}{-}\mathbf{SC}){-}L_2$ bands, respectively. We then plotted the yields for each product, $(\mathbf{CS}{-}\mathbf{SC})_{cyc}$ and $(\mathbf{CS}{-}\mathbf{SC}){-}L_2$ against the concentration of **4**. The relationship between observed yields and the linker concentration is complex as the rate laws governing product distribution involve consecutive second-order reactions,[32] therefore, we treated the data empirically and found they were reasonably well fit by a logarithmic function. Solving this function for y = 50% yields a concentration of **4** of 120 μM that gives equal amounts of the two products (Figure 4C). Together with the value we obtained for $2k_{lig}$, and eq 3, we determined the first order rate constant for cyclization, k_{cyc} to be 9.7×10^{-3} s^{-1} ($t_{1/2} \sim 1$ min).

Cyclization of Larger Rings. The approach we described above for preparing the cyclic dimer is well-suited for preparing larger cyclic molecules simply by synthesizing longer linear oligomers that can then be cyclized with an appropriate bifunctional linker. We next describe the preparation of two larger cyclic molecules and we again measure the rate constants for cyclization. Assembly of the larger oligomers of CS was accomplished using a hetero-bifunctional linker **8** (Figure 2C) that allows extension of the core CS−SC dimer via the process shown in Figure 5A. We prepared this linker by condensation of an amino-hexa(ethylene glycol)-(ethyl-*p*-nitrophenyl)-phosphonate) (**7**) with a succinate-terminated hexa(ethylene glycol)-benzylguanine amide (**6**). To assemble the tetramer, CS−CS−SC−SC, we first treated CS−SC (40 nmol, 50 μM) with 2.5 equiv of asymmetric linker **8** for 5 h to functionalize each cutinase domain with a pendant BG functional group. Excess linker was then removed via SEC and the product peak fractions pooled and concentrated. We removed excess **8** from this intermediate by performing SEC. In general, we favor purifying the intermediates to prevent reaction of CS in the subsequent step with free linker as well as other side reactions that can lower yields. The double functionalized intermediate was then treated with 2.5 equiv of the CS protein (100 nmol, 100 μM) for 1.5 h. The resulting tetramer, CS−CS−SC−SC, was purified by SEC and isolated in 60% yield. We used the same process starting from the tetramer to prepare a hexameric product, CS−CS−CS−SC−SC−SC, in 53% isolated yield. To cyclize these expanded linear oligomers, we separately treated CS−CS−SC−SC and CS−CS−CS−SC−SC−SC at 500 nM with 1.5 equiv of **4**, and purified the products by SEC to isolate the cyclic tetramer, $(\mathbf{CS}{-}\mathbf{CS}{-}\mathbf{SC}{-}\mathbf{SC})_{cyc}$, and the cyclic hexamer, $(\mathbf{CS}{-}\mathbf{CS}{-}\mathbf{CS}{-}\mathbf{SC}{-}\mathbf{SC}{-}\mathbf{SC})_{cyc}$. Analysis of the purified products by SDS-PAGE (Figure 5B) showed bands near the calculated molecular weights of 185, and 277 kDa for the linear tetramer and hexamer, respectively, while the corresponding cyclic products again showed lower mobility in the gel with bands near 240 for $(\mathbf{CS}{-}\mathbf{CS}{-}\mathbf{SC}{-}\mathbf{SC})_{cyc}$, and ~400 kDa for $(\mathbf{CS}{-}\mathbf{CS}{-}\mathbf{CS}{-}\mathbf{SC}{-}\mathbf{SC}{-}\mathbf{SC})_{cyc}$. The hexamer cycle has 6638 atoms in its cyclic chain. For comparison, the cyclic polyethylene molecules prepared by Grubbs and co-workers using ring-opening metathesis polymerization (ROMP) had an average molecular weight of 200 kDa, corresponding to approximately 16 500 carbon atoms in the backbone.[33]

We determined the rate constant k_{cyc} for CS−CS−SC−SC as described earlier, by first labeling this tetramer with AlexaFluor 488, removing the excess dye with micro gel-filtration column, and then treating the protein at 500 nM with concentrations of bifunctional linker **4** ranging from 8−250 μM for 24 h. We then separated the products of each reaction via SDS-PAGE (Figure

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

5C) and quantified their intensities using a gel scanner to obtain reaction yields using eq 3. The yields were plotted against the concentration of **4**, and fit to a logarithmic function to identify the concentration at which equal amounts of products deriving from intra- and intermolecular reaction are observed (Figure 5D). Using our previous measure of $2k_{lig}$, we determined k_{cyc} (tetramer) = 2.3×10^{-3} s^{-1} ($t_{1/2} \sim 6$ min). This procedure was repeated with dye labeled **CS**−**CS**−**CS**−**SC**−**SC**−**SC** in reactions of the protein with **4** ranging from 1 to 250 μM (Figure 5E). Analysis of these reaction data afforded k_{cyc} (hexamer) = 8.1×10^{-4} s^{-1} ($t_{1/2} \sim 14$ min). We assume, in our use of the same $2k_{lig}$ value for each analog that this rate constant does not depend on the molecular weight or some other property of the larger complexes. We have examined this effect in previous work and found that the value of this constant varies less than 10% even when the molecular weight of the complex increases by 2 fold.

Our experimental cyclization rate constants depend on the size of the linear precursor and generally compare favorably to those found for the cyclization of small peptides,[34,35] and the intramolecular reaction of α,ω-functionalized hydrocarbons,[36,37] even for the 278 kDa hexamer. A log−log plot of our cyclization rate constants vs monomer number (Figure 6A)

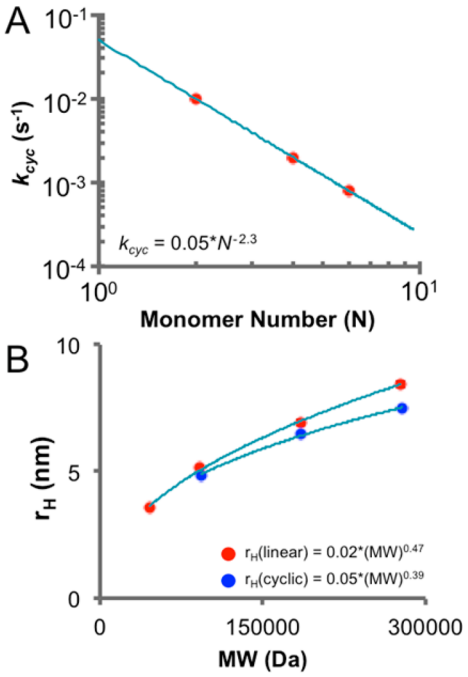

Figure 6. (A) A log−log plot of the dependence of cyclization rate constant k_{cyc} on monomer number, N. A linear fit to this data yields a scaling relationship of $N^{-2.3}$. (B) A plot of the dependence of hydrodynamic radius on molecular weight for the linear and cyclic species. These data are fit to a power law and yield scaling coefficients of $\nu = 0.47$ and 0.39 for the linear and cyclic species, respectively.

yields a straight line with $k_{cyc} \propto N^{-2.3}$. This scaling relationship between cyclization rate and monomer number is in agreement with values obtained from theory ($N^{-2.2}$) and simulation ($N^{-2.2\pm0.1}$) for freely jointed polymers with excluded volume effects.[38,39]

Our ability to separate cyclic products from linear precursors by SEC reflects a smaller hydrodynamic volume of the cyclic species. Hence, we sought to quantitatively compare the hydrodynamic radii of the products using dynamic light scattering (DLS) and to investigate how these radii depend on molecular weight. We obtained DLS spectra for the megamolecules at ~0.5 mg/mL at 25 °C in TBS pH 7.4 on a DynaPro NanoStar instrument. These data are listed in Table S1 along with the calculated and experimental masses of each species. We plotted the hydrodynamic radii of the cyclic and linear products vs molecular weight and fit each data set to a power law to obtain scaling coefficients of $\nu = 0.47$ and 0.39 for the linear and cyclic products, respectively (Figure 5B). Power law relationships between hydrodynamic radius and molecular weight are characteristic of polymeric macromolecules and the magnitude of the scaling coefficient gives information about the shape of the polymer in solution.[40] Theory predicts limiting values of this coefficient to be 0.33 for spherical molecules and 1.0 for rigid rods, with intermediate values representing other geometries. For example, coefficients of $\nu = 0.29$ and 0.30 have been determined for globular proteins and dendrimers, while segments of double stranded DNA have yielded values of $\nu = 0.60$. Therefore, our scaling coefficients illustrate that the products of the cyclization reactions have a significantly more globular architecture in solution than their linear precursors, in accordance with the proposed cyclic geometry.

CONCLUSION

This work introduces a modular strategy for preparing structurally well-defined cyclic molecules having sizes greater than 10 nm. The preparation of nanomaterials (structures having dimensions of 10−100 nm) whose covalent structures are perfectly defined is challenging, particularly for those that require the assembly of multiple proteins. Our use of enzyme-inactivator pairs overcomes many of these challenges by providing rapid and regiospecific reactions that can be sequenced to modularly prepare a range of large structures. Furthermore, the use of orthogonally reactive building blocks and stepwise synthesis permits the isolation of discrete complexes in contrast with those afforded by polymerization reactions. Here, we demonstrate the preparation of linear and cyclic protein megamolecules of 90−280 kDa in high yield in short times under dilute physiological conditions.

This method builds on our previous work and provides a new route to the efficient preparation of cyclic molecular scaffolds with systematic control over protein organization and composition. The cyclic megamolecules—and analogous linear and branched structures—offer new opportunities to address several problems. For example, we believe these methods will allow the preparation of antibody mimics for therapeutic applications. Unlike the bivalent IgG molecules that are now used, such mimics can be designed to present multiple Fab domains, including the use of different Fabs, that can be and mix-and-matched to alter mode of action or tune therapeutic efficacy. We also expect that the spatial arrangement of affinity domains can have an important influence on their biological activity—for example, how conformational dynamics may allow for "walking" behavior on a membrane.[41] We also believe the megamolecule scaffolds may be useful in arranging enzymes that constitute a biosynthetic pathway; recent work has shown that such systems can "channel" intermediates for more efficient multistep syntheses in vitro and in vivo.[42−44] Finally, the megamolecules may represent a route toward synthetic extracellular matrix components, because of their modularity and large sizes.[45]

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

EXPERIMENTAL METHODS

Materials. Hi-mark prestained high molecular weight and Benchtop Fluorescent protein markers, NuPage antioxidant buffer additive, NuPage Novex 3−8% Tris-Acetate precast polyacrylamide gels, NuPage Novex SDS Tris-Acetate running buffer, AlexaFlour488-TFP ester labeling kit, Microbiospin 6 and 30 columns (Pierce) and 1 M Tris buffer, pH 8.0 were purchased from Thermo Fisher Scientific.

Megamolecule Syntheses. All reactions were carried out in the following buffer: 25 mM Tris-HCl, 150 mM NaCl, 0.05% NaN_3, pH 7.8 @ 25 °C (TBS-N_3). Stock solutions of linker **1** were prepared in 1:1 DMSO:50 mM NaOAc buffer pH 5, and of linkers **4** and **8**, in DMSO. Size-exclusion purification steps of products were carried out using TBS-N_3 as the mobile phase at 4 °C on an Äkta FPLC (GE Healthcare). All abosorbance measurements were made using a NanoDrop UV−vis spectrophotometer (Thermo Fisher). The concentration of all protein samples was determined by dividing the measured absorbance at 280 nm by the calculated extinction coefficient based on the peptide sequence of **CS** using Protparam (www.expasy.org). Yields listed below correspond to the amount of oligomeric protein products isolated as >90% pure and not crude yields for the reaction.

Cutinase-SnapTag Dimer (CS−SC). 120 nmol CS (1.2 mL, 100 μM) was treated with compound **1** (60 nmol, 60 μL, 10 mM) for 10 min at r.t. After this period, the reaction was diluted to 2 mL and applied to a Hi-Load Superdex 200 16/60 gel filtration column using a flow rate of 1 mL/min. Fractions containing pure **CS−SC**, as determined by SDS-PAGE were pooled and concentrated to ∼1 mL using a 5 mL Amicon ultrafiltration unit. ε = 71 196 M^{-1} cm^{-1}. (24 nmol, 40%) ESI-MS: 91 879.27, calcd for $C_{4074}H_{6452}N_{1150}O_{1231}S_{20}$ 91 878.83 g/mol.

Cutinase-SnapTag Dimer Cycle ((CS−SC)$_{cyc}$). CS−SC (7.5 nmol, 15 mL, 500 nM) was treated with compound **4** (11.3 nmol, 11.3 μL, 1 mM) for 24 h at r.t.. After this period, the reaction was concentrated to 2 mL in a 5 mL Amicon ultrafiltration unit and then applied to a Hi-Load Superdex 200 16/60 gel filtration column using a flow rate of 1 mL/min. Fractions containing pure (**CS−SC**)$_{cyc}$, as determined by SDS-PAGE were pooled and concentrated to ∼0.5 mL using a 5 mL Amicon ultrafiltration unit. ε = 71 196 M^{-1} cm^{-1}. (4 nmol, 53%) ESI-MS: 92 573.50, calcd for $C_{4102}H_{6508}N_{1150}O_{1246}P_2S_{20}$ 92 573.51 g/mol.

Double Functionalized Cutinase-SnapTag dimer ((CS−SC)−L$_2$). CS−SC (1 nmol, 100 μL, 10 μM) was treated with 50 equiv of **4** (50 nmol, 5 μL, 100 mM) for 30 min at r.t.. Excess linker was removed using a MicroBioSpin 6 column. ESI-MS: 93 542.97 Da calcd for $C_{4142}H_{6570}N_{1152}O_{1267}P_4S_{20}$ 93 542.38 g/mol.

Monofunctional Cutinase-SnapTag Dimer (C^{X}S−SC). C^{X}S (20 nmol, 2 mL, 10 μM) was treated with compound **1** (1 μmol, 6 μL, 10 mM) for 5 min at r.t. This reaction resulted in a BG functionalized cutinase-deactivated monomer (C^{X}S−BG). This was then purified on a Hi-Load Superdex 200 16/60 gel filtration column to remove excess linker and any homodimer formed in the reaction. Fractions corresponding to the major peak were then concentrated to 1 mL in a 5 mL 10 kDa Amicon ultrafiltration unit into which free CS (30 nmol, 300 μL, 100 μM) was added and allowed to react for 30 min. After this period, the reaction was diluted to 2 mL then reapplied to a Hi-Load Superdex 200 16/60 gel filtration column using a flow rate of 1 mL/min. Fractions containing pure (C^{X}S−SC), as determined by SDS-PAGE were pooled and concentrated to ∼1 mL using a 5 mL Amicon ultrafiltration unit. ε = 71 196 M^{-1} cm^{-1} (8.4 nmol, 42%).

Cutinase-SnapTag Tetramer (CS−CS−SC−SC). CS−SC (40 nmol, 800 μL, 50 μM) was treated with compound **8** (100 nmol, 10 μL, 10 mM) 5 h at r.t.. After this period, the reaction was diluted to 2 mL and applied to a Hi-Load Superdex 200 16/60 gel filtration column using a flow rate of 1 mL/min to remove excess ligand. Fractions corresponding to the major peak were collected, concentrated to 1 mL, and treated with **CS** (100 nmol, 1.0 mL, 100 μM) for 30 min. After this period, the reaction was applied to a Hi-Load Superdex 200 16/60 column in series with a Hi-Prep Sephacryl S300 26/60 gel filtration column using a flow rate of 0.3 mL/min. Fractions containing pure **CS−CS−SC−SC**, as determined by SDS-PAGE were pooled and concentrated to 1 mL using a 5 mL Amicon ultrafiltration unit. ε = 142 392 M^{-1} cm^{-1}. (24 nmol, 60%) MALDI-MS: m/z 181 366, calcd for $C_{8184}H_{12964}N_{2304}O_{2479}P_2S_{40}$ 184 640.48 g/mol.

Cutinase-SnapTag Tetramer Cycle ((CS−CS−SC−SC)$_{cyc}$). CS−CS−SC−SC (6.8 nmol, 13.6 mL, 500 nM) was treated with compound **4** (10.2 nmol, 1 μL, 10 mM) 24 h at 25 °C. After this period, the reaction was concentrated to 2 mL in a 5 mL Amicon ultrafiltration unit and then applied to a Hi-Load Superdex 200 16/60 column in series with a Hi-Prep Sephacryl S300 26/60 gel filtration column using a flow rate of 0.3 mL/min. Fractions containing pure (**CS−CS−SC−SC**)$_{cyc}$, as determined by SDS-PAGE were pooled and concentrated to ∼0.5 mL using a 5 mL Amicon ultrafiltration unit. ε = 142 392 M^{-1} cm^{-1}. (3 nmol, 44%) MALDI-MS: m/z 182 121, calcd for $C_{8212}H_{13020}N_{2304}O_{2494}P_4S_{40}$ 185 335.16 g/mol.

Cutinase-SnapTag Hexamer (CS−CS−CS−SC−SC−SC). CS−CS−SC−SC (15 nmol, 1.5 mL, 10 μM) was treated with compound **8** (37.5 nmol, 3.8 μL, 10 mM) for 5 h 25 °C. After this period, the reaction was diluted to 2 mL with running buffer and applied to a Hi-Load Superdex 200 16/60 gel filtration column using a flow rate of 1 mL/min to remove excess ligand. Fractions corresponding to the major peak were collected, concentrated to 1 mL and treated with **CS** (37.5 nmol, 375 μL, 100 μM) for 30 min. After this period, the reaction was applied to a Hi-Load Superdex 200 16/60 column in series with a Hi-Prep Sephacryl S300 26/60 gel filtration column using a flow rate of 0.3 mL/min. Fractions containing pure **CS−CS−CS−SC−SC−SC**, as determined by SDS-PAGE were pooled and concentrated to 1.0 mL using a 5 mL Amicon ultrafiltration unit. ε = 213 588 M^{-1} cm^{-1}. (8 nmol, 53%) MALDI-MS: m/z 272 903, calcd for $C_{12294}H_{19480}N_{3458}O_{3727}P_4S_{60}$ 277 406.17 g/mol.

Cutinase-SnapTag Hexamer Cycle ((CS−CS−CS−SC−SC−SC)$_{cyc}$). 2.2 nmol CS−CS−CS−SC−SC−SC (2.2 nmol, 4.4 mL, 500 nM) was treated with compound **4** (3.3 nmol, 33 μL, 10 mM) 1.5 d at 25 °C. After this period, the reaction was concentrated to 2 mL and then applied to a Hi-Load Superdex 200 16/60 column in series with a Hi-Prep Sephacryl S300 26/60 gel filtration column using a flow rate of 0.3 mL/min. Fractions containing pure (**CS−CS−CS−SC−SC−SC**)$_{cyc}$ as determined by SDS-PAGE were pooled and concentrated to 0.5 mL using a 5 mL Amicon ultrafiltration unit. ε = 213 588 M^{-1} cm^{-1}. (1 nmol, 45%) MALDI-MS: m/z 272 928, calcd for $C_{12322}H_{19536}N_{3458}O_{3742}P_6S_{60}$ 278 100.85 g/mol.

AlexaFluor488 CS Oligomer Conjugates. CS−SC, CS−CS−SC−SC, and CS−CS−CS−SC−SC−SC (2 nmol, 1 mg/mL) were labeled with AlexaFluor 488 using an AlexaFluor488-TFP ester protein labeling kit in 1× PBS buffer according to the manufacturers instructions. Labeled conjugates were subsequently purified by SEC at 4 °C via Äkta FPLC on a Superdex 200 10/300 GL column using TBS-Azide at a flow rate of 0.5 mL/min as the mobile phase. Fractions were analyzed by scanning of SDS-PAGE gels using a Typhoon 4800 gel imager (GE Healthcare). Pure fractions were pooled and concentrated to 5 μM using a Microcon 500 μL concentrator. The following extents of labeling were determined according to the manufacturer's protocol. The extent of labeling for each species: **CS−SC**: ∼5 dyes/molecule; **CS−CS−SC−SC**: ∼7 dyes/molecule; **CS−CS−CS−SC−SC−SC**: ∼8 dyes/molecule.

Gel Electrophoresis. All protein samples were first denatured and reduced by treating each with Laemmli sample buffer (50 mM Tris-HCl, 2% (w/v) sodium dodecyl sulfate, 100 mM dithiothreitol (DTT), 50% (v/v) glycerol, 0.2 mg/mL bromophenol blue, pH 6.8). All samples were run on a 3−8% precast Novex NuPAGE polyacrylamide gels using NuPAGE Tris-Acetate SDS running buffer containing NuPAGE antioxidant buffer additive in the upper buffer chamber with Hi-Mark prestained high molecular weight standards as molecular weight markers. Electrophoresis was carried out at 125 V for 70 min at r.t. Staining was done using Coomassie blue and visualized using a MultiDoc-It Digital Imaging System (UVP). For the determination of cyclic and linear product yields using AlexaFluor488-labeled proteins, gels were run using the same protocol with Benchtop fluorescent protein standards as molecular weight

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

markers. These gels were scanned using a Typhoon 4800 gel imager and bands quantified using the ImageQuant TL (GE Healthcare) software package.

Determination of Cyclization Rate Constants of CS-Megamolecules. To determine the concentration at which equal partitioning between linear and cyclic products occurred, labeled linear **CS**−**SC** and **CS**−**CS**−**SC**−**SC** were treated in six separate reactions (12.5 pmol/rxn) with 6.2, 3.1, 1.6, 0.8, 0.4, and 0.2 nmol compound **4** in TBS-N_3 (25 μL final volume; 500 nM final protein concentration; 250, 125, 63, 31, 16, 8 μM final ligand concentration). These reactions were allowed to proceed for 24 h. For **CS**−**CS**−**CS**−**SC**−**SC**−**SC**, two additional reactions with **4** (0.1 nmol and 0.5 nmol corresponding to 4 and 2 μM final linker concentration) were performed. Each set of reactions was repeated in triplicate. The products of these reactions were then resolved by SDS-PAGE, the double functionalized and cyclic product bands visualized using a Pharos FX gel scanner, and then quantified using the Quantity One software package. Yields for cyclic products were determined using eq 4 in the text. These yields were plotted vs [**4**] and logarithmic fits to the data performed using Origin software (Origin). The resulting logarithmic equation was solved for 50% cycle yield to determine the concentration of **4** at which equipartitioning occurred. Using this value, the rate constant determined for the reaction of $\mathbf{C}^{\mathbf{X}}$**S**−**SC** + compound **4** (81 M^{-1} s^{-1}), and eq 4, rate constants were determined for each of the cyclization reactions.

Mass Spectrometry. ESI mass spectra were obtained on an Agilent 1200 series HPLC connected to an Agilent 6210A time-of-flight (TOF) mass spectrometer. Samples were prepared by dilution of stock protein solution to 1 μM in water. A 10 μL injection of each sample was captured on a C18 trap column (Waters) and eluted using a gradient from 5% to 95% acetonitrile and 0.1% formic acid in water with a flow rate of 0.25 mL/min. Data was analyzed with Agilent MassHunter Qualitative Analysis B.04.00 and spectra were deconvoluted using a maximum entropy deconvolution calculation.

Dynamic Light Scattering. All dynamic light scattering (DLS) measurements were made on a DynaPro Nanostar instrument (Wyatt Technology) at the University of Chicago Biophysics Core Facility. Time-dependent intensity fluctuations of the scattered light by the sample are used to generate a second-order autocorrelation function. A fit of the autocorrelation function is done in real time by the AstraV software package to generate a diffusion coefficient (D) for the interrogated species. The hydrodynamic radius (r_H) of the species is then determined from the diffusion coefficient using the Stokes−Einstein equation. Buffer offsets were made for each sample condition, and assigned to each set of measurements using those conditions. Offset thresholds were typically <0.02% of the mean baseline intensity. All measurements were made on samples at 0.5 mg/mL protein concentration in TBS-N_3. Each measurement was comprised of 20 × 5 s acquisitions to yield a total data collection time of 100 s. Cutoff times of 1.00 and 100 μs were used for fitting each autocorrelation function. Each set of 100 s acquisitions was repeated three times and these values averaged over the three trials to yield the reported r_H for each species.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b02665.

Detailed information regarding the cloning, sequence, expression and purification of the cutinase-SnapTag monomers and syntheses of the linkers (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*milan.mrksich@northwestern.edu

ORCID

Yao Lin: 0000-0001-5227-2663

Milan Mrksich: 0000-0002-4964-796X

Present Address

[†]Department of Chemistry, University of Connecticut, 55 North Eagleville Road, Storrs, Connecticut 06269, United States.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Kevin Metcalf for obtaining the protein ESI-MS spectra and Meghan Carrell for her invaluable assistance in creating protein graphics. We thank Dr. Elena Solomaha at the University of Chicago Biophysics Core Facility for her help with DLS experiments. This work made use of the IMSERC at Northwestern University, which has received support from the State of Illinois. This work was also supported by the Northwestern University Keck Biophysics Facility and a Cancer Center Support Grant (NCI CA060553). This material is based on research sponsored by the Air Force Research laboratory under agreement number FA8650−15−2−5518.

■ REFERENCES

(1) Castro, C. E.; Kilchherr, F.; Kim, D. N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. *Nat. Methods* **2011**, *8* (3), 221−229.

(2) Ringler, P.; Schulz, G. E. *Science* **2003**, *302* (5642), 106−9.

(3) Oohora, K.; Onoda, A.; Hayashi, T. *Chem. Commun. (Cambridge, U. K.)* **2012**, *48* (96), 11714−26.

(4) Sinclair, J. C.; Davies, K. M.; Venien-Bryan, C.; Noble, M. E. *Nat. Nanotechnol.* **2011**, *6* (9), 558−62.

(5) Carlson, J. C.; Jena, S. S.; Flenniken, M.; Chou, T. F.; Siegel, R. A.; Wagner, C. R. *J. Am. Chem. Soc.* **2006**, *128* (23), 7630−8.

(6) Bai, Y.; Luo, Q.; Zhang, W.; Miao, L.; Xu, J.; Li, H.; Liu, J. *J. Am. Chem. Soc.* **2013**, *135* (30), 10966−9.

(7) Bastings, M. M. C.; de Greef, T. F. A.; van Dongen, J. L. J.; Merkx, M.; Meijer, E. W. *Chem. Sci.* **2010**, *1* (1), 79−88.

(8) Lai, Y. T.; Cascio, D.; Yeates, T. O. *Science* **2012**, *336* (6085), 1129−1129.

(9) Lai, Y. T.; Reading, E.; Hura, G. L.; Tsai, K. L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O. *Nat. Chem.* **2014**, *6* (12), 1065−1071.

(10) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D. *Nature* **2014**, *510* (7503), 103.

(11) Gradisar, H.; Bozic, S.; Doles, T.; Vengust, D.; Hafner-Bratkovic, I.; Mertelj, A.; Webb, B.; Sali, A.; Klavzar, S.; Jerala, R. *Nat. Chem. Biol.* **2013**, *9* (6), 362.

(12) Kim, Y. E.; Kim, Y. N.; Kim, J. A.; Kim, H. M.; Jung, Y. *Nat. Commun.* **2015**, DOI: 10.1038/ncomms8134.

(13) Wen, A. M.; Steinmetz, N. F. *Adv. Healthcare Mater.* **2014**, *3* (11), 1739−44.

(14) Su, Z.; Wang, Q. *Angew. Chem., Int. Ed.* **2010**, *49* (52), 10048−50.

(15) Song, W. J.; Tezcan, F. A. *Science* **2014**, *346* (6216), 1525−1528.

(16) Proft, T. *Biotechnol. Lett.* **2010**, *32* (1), 1−10.

(17) Rossi, E. A.; Goldenberg, D. M.; Chang, C. H. *Bioconjugate Chem.* **2012**, *23* (3), 309−323.

(18) Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (12), E690−E697.

(19) Harvey, J. A.; Itzhaki, L. S.; Main, E. R. G. *ACS Synth. Biol.* **2018**, *7* (4), 1067−1074.

(20) Neumann, H.; Wang, K. H.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464* (7287), 441−444.

(21) ElSohly, A. M.; Francis, M. B. *Acc. Chem. Res.* **2015**, *48* (7), 1971−1978.

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399

(22) Andersen, K. A.; Raines, R. T. *Methods Mol. Biol.* **2015**, *1248*, 55−65.
(23) Kim, C. H.; Axup, J. Y.; Dubrovska, A.; Kazane, S. A.; Hutchins, B. A.; Wold, E. D.; Smider, V. V.; Schultz, P. G. *J. Am. Chem. Soc.* **2012**, *134* (24), 9918−9921.
(24) Dommerholt, J.; van Rooijen, O.; Borrmann, A.; Guerra, C. F.; Bickelhaupt, F. M.; van Delft, F. L. *Nat. Commun.* **2014**, *5*, 5378.
(25) Modica, J. A.; Skarpathiotis, S.; Mrksich, M. *ChemBioChem* **2012**, *13* (16), 2331−2334.
(26) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chem. Biol.* **2008**, *3* (6), 373−82.
(27) Kwon, Y.; Han, Z.; Karatan, E.; Mrksich, M.; Kay, B. K. *Anal. Chem.* **2004**, *76* (19), 5713−20.
(28) Hodneland, C. D.; Lee, Y. S.; Min, D. H.; Mrksich, M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (8), 5048−52.
(29) Johnsson, K.; Chidley, C.; Mosiewicz, K. *Bioconjugate Chem.* **2008**, *19* (9), 1753−1756.
(30) Arai, R.; Ueda, H.; Kitayama, A.; Kamiya, N.; Nagamune, T. *Protein Eng., Des. Sel.* **2001**, *14* (8), 529−532.
(31) Pauly, G. T.; Loktionova, N. A.; Fang, Q. M.; Vankayala, S. L.; Guida, W. C.; Pegg, A. E. *J. Med. Chem.* **2008**, *51* (22), 7144−7153.
(32) Lente, G. *J. Math. Chem.* **2015**, *53* (4), 1172−1183.
(33) Bielawski, C. W.; Benitez, D.; Grubbs, R. H. *Science* **2002**, *297* (5589), 2041−2044.
(34) Zhang, C.; Dai, P.; Spokoyny, A. M.; Pentelute, B. L. *Org. Lett.* **2014**, *16* (14), 3652−5.
(35) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. *Biochemistry* **2001**, *40* (24), 7099−108.
(36) Galli, C.; Illuminati, G.; Mandolini, L.; Tamborra, P. *J. Am. Chem. Soc.* **1977**, *99* (8), 2591−2597.
(37) Casadei, M. A.; Galli, C.; Mandolini, L. *J. Am. Chem. Soc.* **1984**, *106* (4), 1051−1056.
(38) Oono, Y.; Freed, K. F. *J. Chem. Phys.* **1981**, *75* (2), 1009−1015.
(39) Podtelezhnikov, A.; Vologodskii, A. *Macromolecules* **1997**, *30* (21), 6668−6673.
(40) Waigh, T. A. Biomacromolecules. In *Applied Biophysics*; John Wiley & Sons, Ltd., 2007; pp 171−203.
(41) Preiner, J.; Kodera, N.; Tang, J. L.; Ebner, A.; Brameshuber, M.; Blaas, D.; Gelbmann, N.; Gruber, H. J.; Ando, T.; Hinterdorfer, P. *Nat. Commun.* **2014**, DOI: 10.1038/ncomms5394.
(42) Lim, S. I.; Yang, B.; Jung, Y.; Cha, J.; Cho, J.; Choi, E. S.; Kim, Y. H.; Kwon, I. *Sci. Rep.* **2016**, *6*, 6.
(43) Dueber, J. E.; Wu, G. C.; Malmirchegini, G. R.; Moon, T. S.; Petzold, C. J.; Ullal, A. V.; Prather, K. L. J.; Keasling, J. D. *Nat. Biotechnol.* **2009**, *27* (8), 753−U107.
(44) Conrado, R. J.; Wu, G. C.; Boock, J. T.; Xu, H. S.; Chen, S. Y.; Lebar, T.; Turnsek, J.; Tomsic, N.; Avbelj, M.; Gaber, R.; Koprivnjak, T.; Mori, J.; Glavnik, V.; Vovk, I.; Bencina, M.; Hodnik, V.; Anderluh, G.; Dueber, J. E.; Jerala, R.; DeLisa, M. P. *Nucleic Acids Res.* **2012**, *40* (4), 1879−1889.
(45) Wu, J. H.; Li, P. F.; Dong, C. L.; Jiang, H. T.; Xue, B.; Gao, X.; Qin, M.; Wang, W.; Chen, B.; Cao, Y. *Nat. Commun.* **2018**, DOI: 10.1038/s41467-018-02917-6.

DOI: 10.1021/jacs.8b02665
J. Am. Chem. Soc. 2018, 140, 6391−6399