ADVANCED MATERIALS

Supporting Information

for *Adv. Mater.*, DOI: 10.1002/adma.201103316

Nanomechanics of Streptavidin Hubs for Molecular Materials

*Minkyu Kim , Chien-Chung Wang , Fabrizio Benedetti , Mahir Rabbi , Vann Bennett , and Piotr E. Marszalek **

Submitted to ADVANCED MATERIALS

DOI: 10.1002/adma.201103316

Supporting Information for:

Nanomechanics of Streptavidin Hubs for Molecular Materials

*By Minkyu Kim, Chien-Chung Wang, Fabrizio Benedetti, Mahir Rabbi, Vann Bennett, and Piotr E. Marszalek**

[*] Dr. M. Kim, Dr. M. Rabbi, and Prof. P. E. Marszalek
Department of Mechanical Engineering and Materials Science
Center for Biologically Inspired Materials and Material Systems
Center for Biomolecular and Tissue Engineering
Duke University
144 Hudson Hall, Durham, NC, 27708 (USA)
Fax: (+1) 919-660-5409
E-mail: pemar@duke.edu
Dr. C. Wang
Graduate Insititute of Biotechnology
National Chung-Hsing University
250 Kuo-Kaung Rd. Taichung 402, Taiwan (R.O.C)
F. Benedetti
Institute of Physics of Biological Systems
EPFL
CH-1015 Lausanne (Switzerland)
Prof. V. Bennett
Howard Hughes Medical Institute and Department of Cell Biology
Duke University Medical Center Durham, NC, 27710 (USA)

Table of Contents

Enhancing the solubility of streptavidin-based protein building blocks

Previously, when partner proteins were fused onto the N-terminus of streptavidin, the folding ability of the chimeric proteins was limited, and the chimeras were produced from inclusion bodies.[1] To prevent our target proteins from precipitating, we planned to follow one of the strategies to increase protein solubility by using soluble long polypeptide tags, which act to aid in solubilizing recombinant proteins[2, 3], for instance, maltose binding protein (MBP).[4]

Soluble long $I27_6$ [5, 6] and $(SNase\text{-}I27)_3$[7] polypeptides were fused onto the N-terminus of streptavidin to increase the solubility of recombinant streptavidin complexes during protein expression and purification. Based upon protein gel analysis (Figure S1-S3), we confirmed that designed chimeras are soluble and were produced without requiring a protein denaturing-refolding procedure. AFM imaging and AFM stretching measurements also support that our chimeras are folded correctly, providing results of proper functional intact biotin binding pockets of streptavidin (Figure 1d and Figure S4) and mechanical fingerprints of I27 modules and SNase domains (Figure 2b and 3b). Therefore, we conclude that our soluble tandemly repeated proteins domains enhanced the solubility of streptavidin-based complexes, and this fact allowed us to simply follow conventional protein expression and purification methods.

DNA engineering

To prepare $I27_6$-SM (streptavidin monomer) and $(SNase\text{-}I27)_3$-SM plasmids, the core streptavidin gene (residues 13-139) in the pET21a plasmid was amplified by polymerase chain reaction (PCR) with the following forward and reverse primers, which include Spe I and EcoR I restriction enzyme sites (underlined), respectively; 5'-GAAGAA<u>ACTAGT</u>GGCGCCGGCGCCGCTGAAGCTGGTATCACC-3' and 5'-GAAGAA<u>GAATTC</u>CTACTAATGATGATGATGATGGGCGCCGGAAGCAGCGGA

CGGTTT-3' (Integrated DNA Technologies, Coralville, IA, USA). Digested by Spe I and EcoR I restriction enzymes, and gel purified (QIAquick Gel Extraction kit, QIAGEN, Valencia, CA, USA, catalog # 28706) PCR products were cloned into the 7th and 8th positions of pAFM1-8 and pAFM1-8(SNase 2, 4, 6) plasmids[6, 7].

To construct the $I27_7$ plasmid, the following oligonucleotides which contain Mlu I and EcoR I restriction enzyme sites (underlined) were annealed, and then cloned into position 8 of the pAFM 1-8 plasmid; 5'- CGCGTGGCGCCGGCGCC CATCATCATCATCATCATTAGTAGG-3' and 5'- AATTCCTACTAATGATGATGATGATGATG GGCGCCGGCGCCA-3' (Integrated DNA Technologies).

Protein expression and purification

All plasmids were transformed into *E.coli* OverExpress C41(DE3) pLysS cells (Lucigen, Middleton, WI, catalog # 60444). A freshly-grown bacterial colony was inoculated in 3ml LB Broth medium with 1mM ampicillin at 37˚C overnight. For pre-culture, 50 µl of overnight culture of all samples were inoculated into 5ml EB™ (Zymoresearch, Orange, CA, USA, catalog # M3012) until $OD_{600} > 1$ at 37˚C. Proteins were expressed overnight at room temperature by adding 15ml of OB™ (Zymoresearch, catalog # M3013) in the presence of 0.5 mM IPTG. The cells were harvested by spinning down at 4k x g for 20 min, and then frozen at -80˚C. Thawed cells were resuspended with Lysis buffer including Lysozyme and Benzonase Nuclease (QIAGEN, Valencia, CA, catalog # 37900). The lysates were spun down at 14k x g for 30 min at 4˚C. All polyhistidine tagged proteins, $I27_8$, I27-$(SNase\text{-}I27)_3$, $I27_7$, $(I27_6\text{-}SM)_4$ and $[(I27\text{-}SNase)_3\text{-}SM]_4$ were purified by Ni-NTA Superflow Columns (QIAGEN, catalog # 30622). These polyproteins were eluted by the wash buffer in the presence of 250 mM imidazole. Eluted proteins were loaded into the Slide-A-Lyzer G2 Dialysis Cassette (Thermo Scientific, Rockford, IL, USA, product # 87735) and dialyzed against 2L of 100 mM

Tris-HCl, 150 mM NaCl buffer, and 0.5 mM Tcep (Thermo Scientific, Rockford, IL, product # 77720) (pH ~7.6) for overnight. Dialyzed samples were concentrated using Amicon Ultra-0.5 and 15 centrifugal filter devices (50 kDa for $I27_8$, I27-$(SNase\text{-}I27)_3$ and $I27_7$, and 100 kDa for $(I27_6\text{-}SM)_4$ and $[(I27\text{-}SNase)_3\text{-}SM]_4$) (Millipore, Billerica, MA, USA, catalog # UFC905024 and UFC510024). The size of all proteins were estimated by SDS-PAGE (Invitrogen, Camarillo, CA, catalog # EA0378) and $(I27_6\text{-}SM)_4$ and $[(I27\text{-}SNase)_3\text{-}SM)]_4$ were characterized by NativePAGE (Invitrogen, catalog # BN1003).

SDS-PAGE and Native PAGE analysis

To determine that $(I27_6\text{-}SM)_4$ are constructed as intended, we measured the molecular mass of denatured and intact proteins by performing SDS-PAGE and Native PAGE (Figure S2) on the samples. As expected, denatured proteins ran similar to the estimated molecular weight based on protein residues, such as ~78 kDa for $I27_6$-SM (Figure S2a). To examine whether $(I27_6\text{-}SM)_4$ formed or not, we ran the intact tetramers on a native gel. Since the mobility rate of the intact $(I27_6\text{-}SM)_4$ structures on a native gel, however, is not known, and the mobility rate can be affected by the shape of tandem repeats of an I27 domain, we ran $(I27_6\text{-}SM)_4$ together with $I27_8$ constructs as a control experiment. One clear single band was observed on the native gel for both $(I27_6\text{-}SM)_4$ and $I27_8$ (Figure S2b and S2d). However, when considering the expected molecular mass of both proteins based on their residues, $(I27_6\text{-}SM)_4$ = ~78 kDa x 4 and $I27_8$ = ~84 kDa, the molecular mass on the native gel for both constructs are overestimated. While carefully analyzing the result of $I27_8$ on the native gel (Figure S2d), we realized that tandem repeats of the I27 domain run slowly on the native gel. This effect also influenced the mobility of the tetrameric complexes in Figure S2b. Therefore, we confirmed that $I27_6$-SM (Figure 1a and Figure S2a) were clearly self-assembled into tetrameric structures, $(I27_6\text{-}SM)_4$ (Figure 1b and Figure S2b).

On the other hand, $[(SNase\text{-}I27)_3\text{-}SM]_4$ complexes showed consistent results on SDS-PAGE and a native gel with single bands (Figure S3). Thus, we also confirmed that designed streptavidin-based fusion proteins were clearly self-assembled into tetramers.

Producing size-controlled biotinylated DNA

To produce single-end biotin-labeled size-controlled double-stranded DNA (1.5kb), the β-catenin plasmid[8] was amplified by polymerase chain reaction (PCR), with *Pfu Turbo* Hotstart PCR master mix (Agilent Technologies, Santa Clara, CA, USA, catalog # 600600) with the following forward and reverse PCR primers where only forward primer contains 5' biotin (Integrated DNA Technologies, Coralville, IA, USA); 5'-biotin-TCAGGCTCAGGCATGCTGAAACATGCAGTTGTAAACTTGATT-3' and 5'-GCCTGAGCCTGAAGCAAGTTCACAGAGGACCCCTGCAGCTAC-3'. After performing PCR and DNA gel extraction and purification, 1.5 kb 5'-end biotinylated DNA were obtained.

AFM imaging on biotinylated DNA and chimeric streptavidin complexes

In order to verify the function of biotin binding pockets in streptavidin-based tetrameric structures, we performed AFM imaging on self-assembled biotinylated DNA to $(I27_6\text{-}SM)_4$ and pure streptavidin (SA) (product # 21122, Thermo Scientific, Rockford, IL) (Figure S4). The biotinylated DNA and $(I27_6\text{-}SM)_4$ or the DNA and SA were mixed with the molar ratio of 2:1 in the buffer containing 20mM Tris-HCl, 100mM potassium glutamate, 0.4 mM DTT and 10 mM $MgCl_2$ for more than 30 min at room temperature prior to incubating on mica. After 1 min incubation on a freshly cleaved mica, samples were thoroughly washed by deionized H_2O and air-dried before AFM imaging. Images were collected with the Nanoscope V MultiMode™ Scanning Probe Microscope (Veeco Instruments Inc., Santa Barbara, CA) using tapping mode with an E scanner. RTESP probes (Veeco, Santa Barbara, CA, USA) with resonance frequencies of ~300 kHz were used for air imaging. All images for self-assembled

biotinylated DNA/$(I27_6\text{-}SM)_4$ and the DNA/SA complexes were captured at a scan rate of ~3.0 Hz, with 1024 x 1024 pixel resolution, and with a scan size of 2 x 2 µm. Images were analyzed to determine the function of biotin binding pockets between $(I27_6\text{-}SM)_4$ and the pure streptavidin tetramer, by comparing the number of DNA molecules on each protein.

AFM stretching measurements

All force/extension measurements were carried out on custom-built AFM instruments as described previously[8-10]. To mimic similar working conditions when streptavidin-based supramolecular structures are fully connected with other biotinylated materials to possibly create networks, we maintain the presence of soluble 10 µM d-biotins (Avidity, Aurora, CO, product # BIO200) at all times for our experiments. All samples were picked up for stretching measurements with an untreated sharpened AFM tip with pulling speeds of around 500 nm/s (MSNL, Veeco, Santa Barbara, CA, USA). The spring constant of each cantilever was calibrated on a fresh mica substrate in the same buffer as the stock protein solution with additional biotins. By using the energy equipartition theorem[11], we obtained 12~18 $pN \cdot nm^{-1}$ cantilever spring constants in solution at room temperature. RMS force noise for the cantilever was ~12 pN in the 1-500 Hz bandwidth. 1–4 mg/ml of all stock protein samples were diluted to 4–40 µg/ml for $(I27_6\text{-}SM)_4$ and $I27_7$ constructs, and 50-100 µg/ml for $[(SNase\text{-}I27)_3\text{-}SM]_4$ and $I27\text{-}(SNase\text{-}I27)_3$ constructs with the same buffer for cantilever calibration. Diluted $I27_7$ and $(I27_6\text{-}SM)_4$ were incubated on a freshly cleaned glass substrate, while $I27\text{-}(SNase\text{-}I27)_3$ and $[(SNase\text{-}I27)_3\text{-}SM]_4$ were incubated on a gold substrate. After 40 min incubation, all samples were gently washed (1-2 times) with the buffer to remove weakly bound samples. During AFM experiments, the sample on the piezo was moved toward the cantilever and returned to its original position, and then shifted to the next location by the 3-axis piezoelectric stage. By this approach, single pulling experiments with a sharpened tip on all different pulling locations, we increase the possibility to test intact molecules.

Data selection criteria to determine rupture force events of the streptavidin

When analyzing force-extension curves, we selected criteria to clearly determine rupture force events of streptavidin tetramer. With the assumption that intermolecular interactions of streptavidin tetramer are stronger than mechanical unfolding of either I27 or SNase domains, we only choose force curves which contain 7 or more fingerprints of I27 (Figure 2b) or 4 or more fingerprints of SNase modules (Figure 3b). Although less than 7 I27 (or 4 SNase) force peaks are observed (Figure S5), it is still possible that rupture force events of streptavidin tetramer can be captured in force-extension curves (red circle in Figure S5a). However, most measured forces would be detachment forces (gray circles in Figure S5). Therefore, to reduce numerous possibilities of measuring detachment forces, we only consider data which match our data selection criteria, i.e. mechanical unfolding of 7 or more I27 domains or 4 or more SNase modules (Figure 2b and 3b).

Differentiating rupture events of streptavidin tetramers from detachment events

Even though numerous detachment forces during AFM force spectroscopy measurements can be excluded by the data selection criteria, two possible detachment forces still exist in force-extension curves between streptavidin-based supramolecular structures and the substrate or the AFM cantilever. Thus, to distinguish rupture events of the streptavidin tetramer from these two possible detachment events, we performed control experiments.

For control experiments, 7 tandem repeats of I27 domains, $I27_7$, and four I27 interspersed with three SNase, I27-$(SNase\text{-}I27)_3$, constructs are used as counterparts of $(I27_6\text{-}SM)_4$ and $[(I27\text{-}SNase)_3\text{-}SM]_4$, respectively. Final forces from control experiments (gray circle in Figure S6a and S7) only display detachment forces between cantilever and molecule or substrate and molecule (i → ii of Figure S6b). By comparing obtained all final force values

from streptavidin-based protein complexes and from control experiments, we could clearly identify the rupture force events of the streptavidin tetramer (Figure 2c and 3c and Figure S8).

Statistical analysis

To describe the probability density in a simple way, the histogram method is often utilized. One weak point of the method is that total distribution can be affected by the starting bin position. Contrarily, when optimal bandwidth is obtained in kernel density estimation[12-15] (KDE), since each kernel is located on the center of each data value, the probability density function (pdf) generated by the KDE method can well describe the total tendency of data. However, the KDE method also has a disadvantage that the method cannot capture the sharp features of data, such as a high kurtosis, because of the use of a universal bandwidth[16]. Since observed data of rupturing streptavidin interfaces showed high kurtosis (Figure S8), using only the KDE method might be misleading in differentiating rupture events from detachment events. Therefore, to complement each method, total tendency of density distributions and a significant kurtosis are rigorously analyzed by utilizing both histogram and KDE methods (Figure 2c and 3c and Figure S8).

References

[1] T. Sano, C. R. Cantor, *Method Enzymol* **2000**, *326*, 305-311.

[2] D. S. Waugh, *Trends Biotechnol* **2005**, *23*, 316-320.

[3] K. Terpe, *Appl Microbiol Biot* **2003**, *60*, 523-533.

[4] H. P. Sorensen, H. U. Sperling-Petersen, K. K. Mortensen, *Protein Expres Purif* **2003**, *32*, 252-259.

[5] M. Carrion-Vazquez, A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, J. M. Fernandez, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 3694-3699.

[6] A. Steward, J. L. Toca-Herrera, J. Clarke, *Protein Science* **2002**, *11*, 2179-2183.

[7] C.-C. Wang, T.-Y. Tsong, Y.-H. Hsu, P. E. Marszalek, *Biophysical Journal* **2011**, *100*, 1094-1099.

[8] M. Kim, K. Abdi, G. Lee, M. Rabbi, W. Lee, M. Yang, C. J. Schofield, V. Bennett, P. E. Marszalek, *Biophysical Journal* **2010**, *98*, 3086-3092.

[9] P. E. Marszalek, A. F. Oberhauser, Y. P. Pang, J. M. Fernandez, *Nature* **1998**, *396*, 661-664.

[10] M. Rabbi, and Marszalek, P.E., in *Single Molecule Techniques: A laboratory manual* (Ed: P. R. Selvin, and Ha, T.), Cold Spring Harbor Laboratory Press, **2008**, 371-394.

[11] E. L. Florin, M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V. T. Moy, H. E. Gaub, *Biosens Bioelectron* **1995**, *10*, 895-901.

[12] M. Rosenblatt, *Ann Math Stat* **1956**, *27*(3), 832-837.

[13] E. Parzen, *Ann Math Stat* **1962**, *33*(3), 1065-1076.

[14] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley & Sons, Inc., New York **1992**.

[15] D. W. Scott, *Biometrika* **1979**, *66*, 605-610.

[16] L. J. Yang, J. S. Marron, *J Am Stat Assoc* **1999**, *94*, 580-589.

[17] C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, *Science* **1994**, *265*, 1599-1600.

[18] T. Keenan, D. R. Yaeger, N. L. Courage, C. T. Rollins, M. E. Pavone, V. M. Rivera, W. Yang, T. Guo, J. F. Amara, T. Clackson, M. Gilman, D. A. Holt, *Bioorgan Med Chem* **1998**, *6*, 1309-1335.

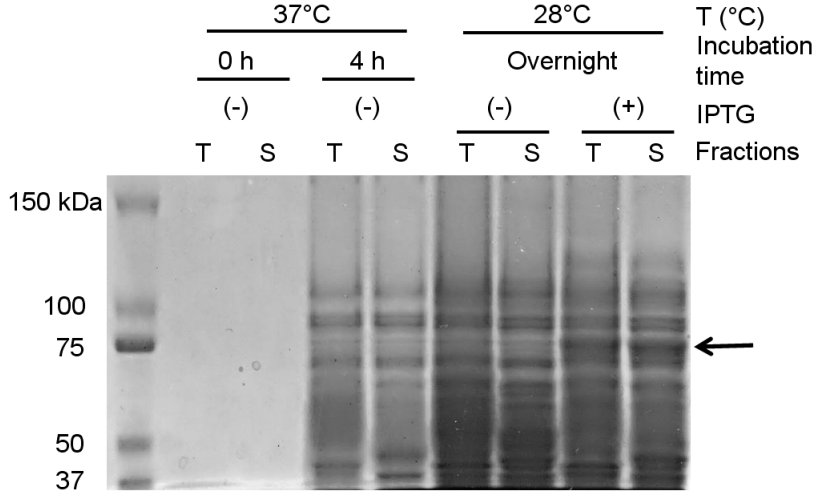

Figure S1. Protein expression of genetically fused six tandem repeats of the I27 domain and the streptavidin monomer, $I27_6$-SM. After cell lysis, all proteins from the cell crudes (T) and soluble fraction (S) were analyzed by SDS-PAGE. In the presence of 0.5 mM IPTG (+), $I27_6$-SMs (~78 kDa) were only induced on both T and S fractions (black arrow). Based on SDS-PAGE, most of our streptavidin-based protein complexes were expressed in soluble fractions.

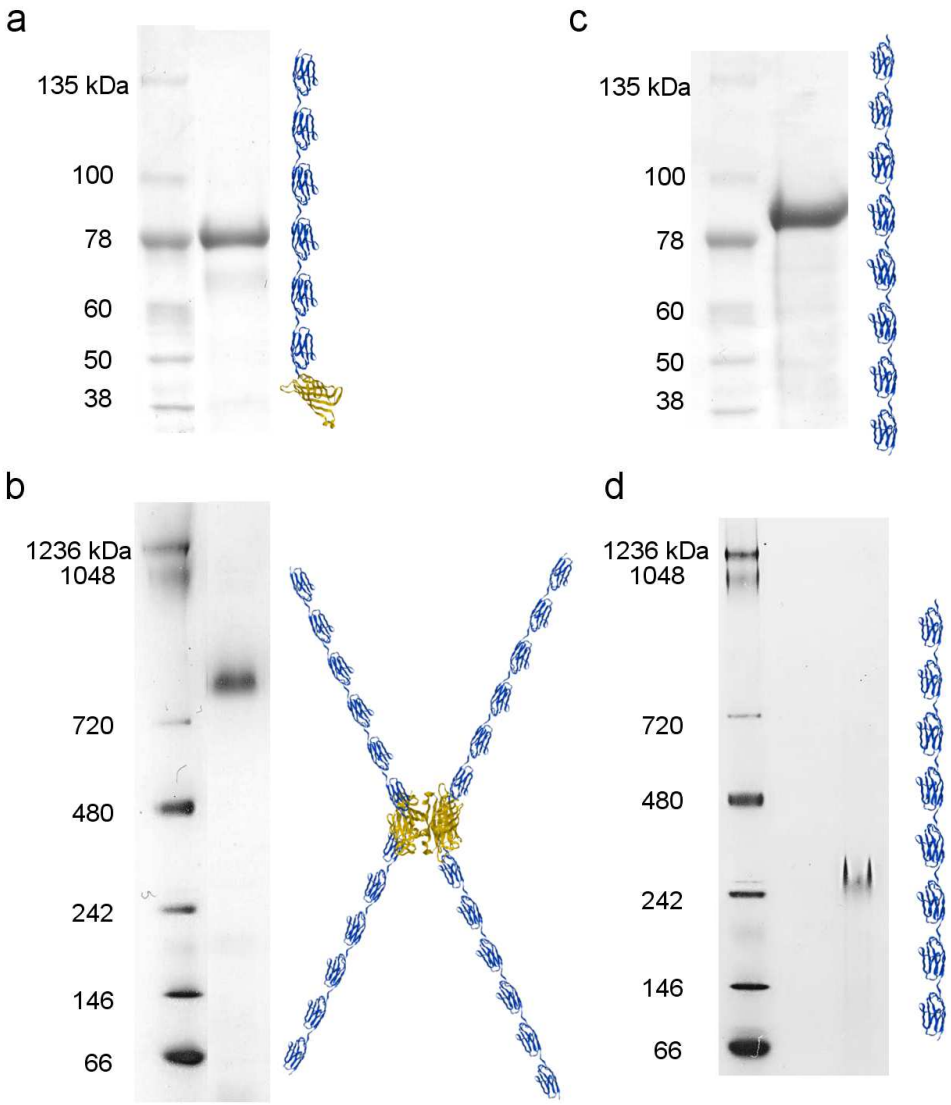

Figure S2. Protein gel analysis on denatured and intact structures of $(I27_6\text{-SM})_4$ and $I27_8$. a) $I27_6$-SM constructs on SDS-PAGE. b) $(I27_6\text{-SM})_4$ on a native gel. c) $I27_8$ constructs on SDS-PAGE. d) $I27_8$ constructs on a native gel. By comparing all results captured on SDS-PAGE and native gels (a-d), we found that tandem repeats of I27 domains decrease the mobility rate of $(I27_6\text{-SM})_4$ and $I27_8$ constructs on a native gel.

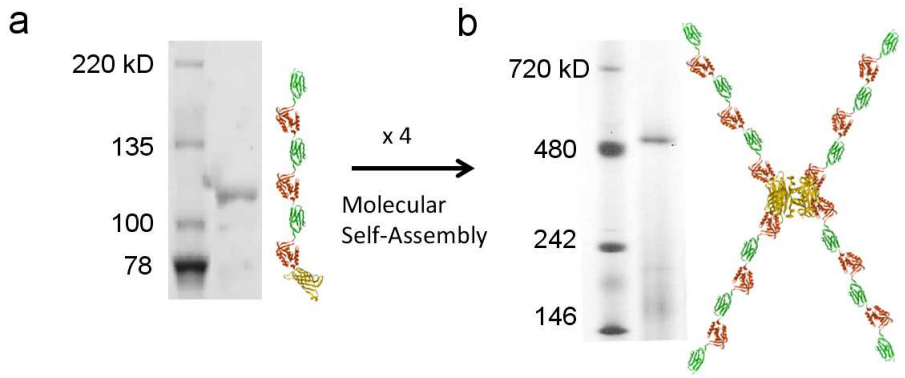

Figure S3. Protein gel analysis on denatured and intact structures of $[(I27\text{-}SNase)_3\text{-}SM]_4$. a) $(SNase\text{-}I27)_3\text{-}SM$ on SDS-PAGE. b) $[(SNase\text{-}I27)_3\text{-}SM]_4$ on a native gel.

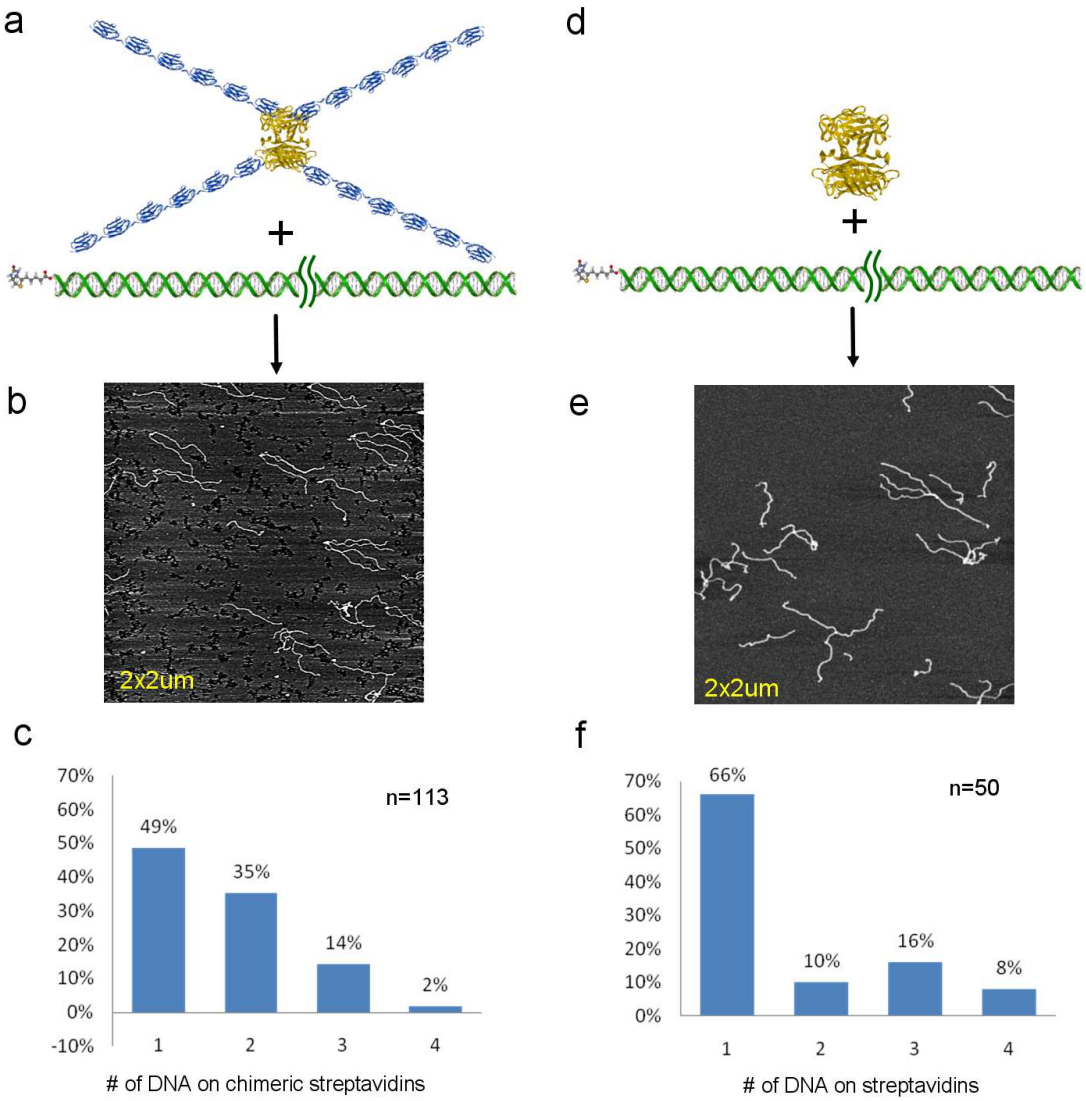

Figure S4. Comparing connecting abilities between the streptavidin with/without the presence of protein arms. a) $(I27_6\text{-}SM)_4$ and biotinylated DNA are mixed with 1: 2 molar ratio and deposited on the mica substrate for AFM imaging. b) Up to 4 DNA could bind to streptavidin-based supramolecular structures. c) DNA/$(I27_6\text{-}SM)_4$ complexes were classified by the number of DNA attached on the complex. d)-f) Pure streptavidins were also mixed with biotinylated DNA for AFM imaging. By comparing histograms in (c) and (f), we conclude that the function of biotin binding pockets in streptavidin tetramers with the presence of four protein arms is not reduced.

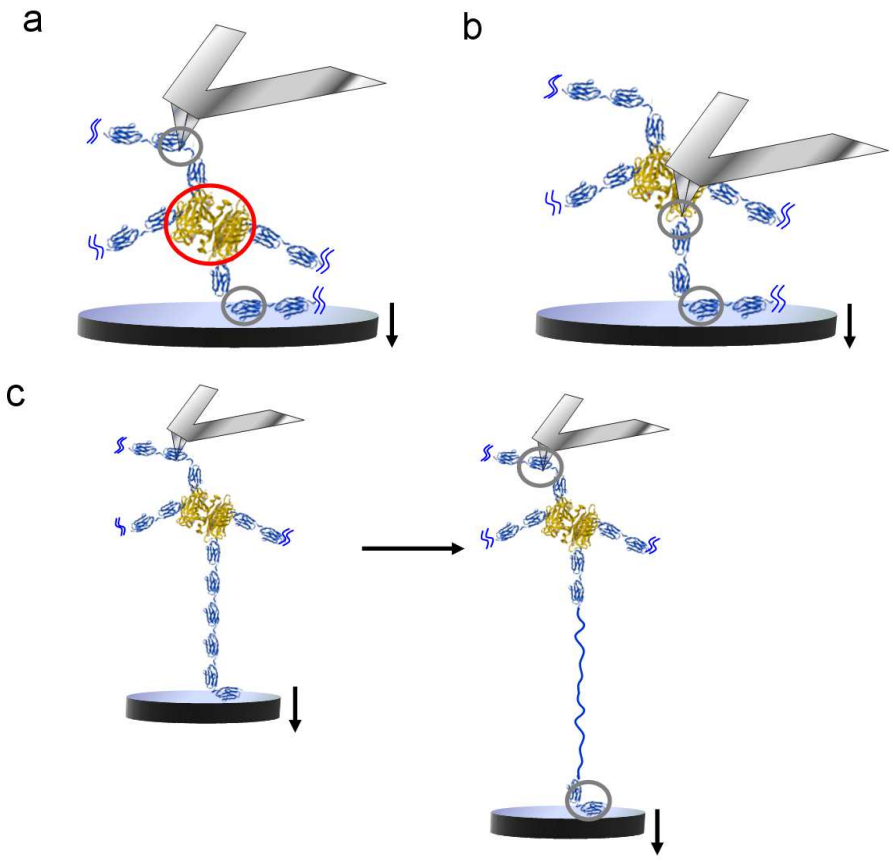

Figure S5. Possible experimental schematics of final forces in force-extension curves during AFM force spectroscopy measurements when less than 7 mechanical fingerprints of I27 domains are considered. a)-c) Three experimental geometries are possible when force-extension curves contain 6 or less I27 force peaks. With the assumption that rupture forces of streptavidin is stronger than unfolding forces of I27 domains, only one case in (a) can really measure the rupture force of the streptavidin (red circle). For easy identification of the true tetramer rupture forces, it is necessary to exclude various detachment forces (gray circles). For this reason, we only considered data which include 7 (or up to 12) I27 force peaks on force-extension curves (Figure 2b). Also, using the same analogy, when force-extension curves clearly show 4 (or up to 6) SNase force peaks, we analyzed final forces on $[(\text{SNase-I27})_3\text{-SM}]_4$ (Figure 3b).

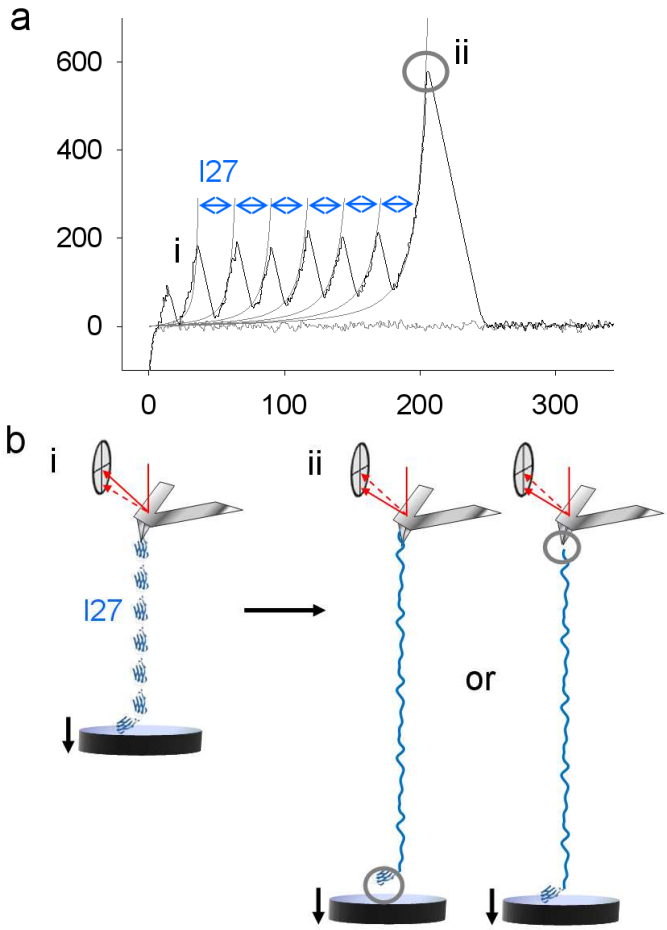

Figure S6. Detachment force measurements on $I27_7$ constructs. a) Force-extension curve shows mechanical unfolding of 6 I27 domains when the $I27_7$ construct was stretched by AFM. Multiple gray lines are worm-like chain (WLC) models[17] fits to the curve with contour length increment, ΔLc = 29.8 nm, and persistence length, p = 0.4 nm. b) Schematics illustrate experimental geometry while AFM measures detachment forces either the substrate to the molecule or the AFM cantilever to the molecule (gray circles).

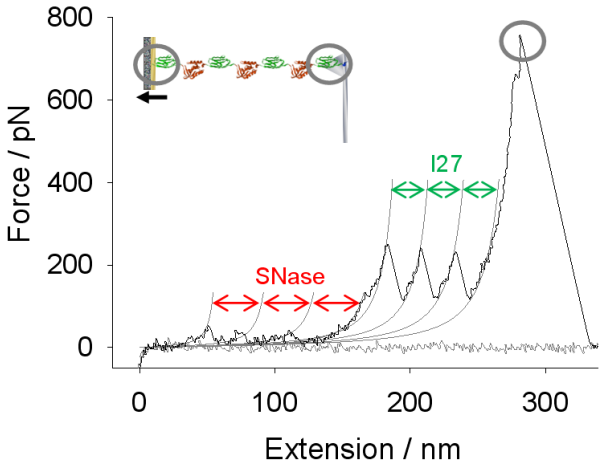

Figure S7. Detachment force measurements on I27-(SNase-I27)$_3$ constructs. The I27-(SNase-I27)$_3$ construct was stretched by using AFM. A detachment force between the molecule and the AFM tip or the substrate (gray circles in the inset) is measured on the force-extension curve before the force drops to zero (gray circle). WLC fits to the curve (gray lines) measure ΔLc = 43 nm with p = 0.6 nm for SNase modules and ΔLc = 28.5 nm with p = 0.4 nm for I27 domains.

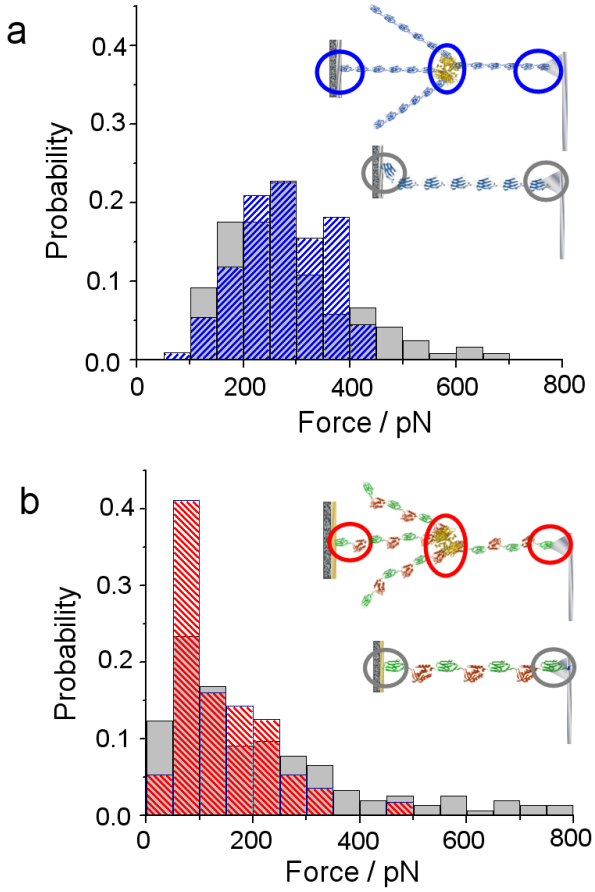

Figure S8. Differentiating rupture forces of the streptavidin tetramers from detachment forces by using the histogram method. a) Comparing histograms of final forces between $(I27_6\text{-SM})_4$ (rupture or detachment forces, blue circles, blue hashed bars) and $I27_7$ constructs (detachment forces, gray circles, gray bars). b) Comparing histograms of final forces of $[(I27\text{-SNase})_3\text{-SM}]_4$ (rupture or detachment forces, red circles, red hashed bars) and $I27\text{-}(SNase\text{-}I27)_3$ constructs (detachment forces, gray circles, gray bars). Rupture events of the streptavidin tetramer are observed at approximately 400 pN (a) and 100 pN (b), and results of pdf in Figure 2c and 3c also match with these observations.

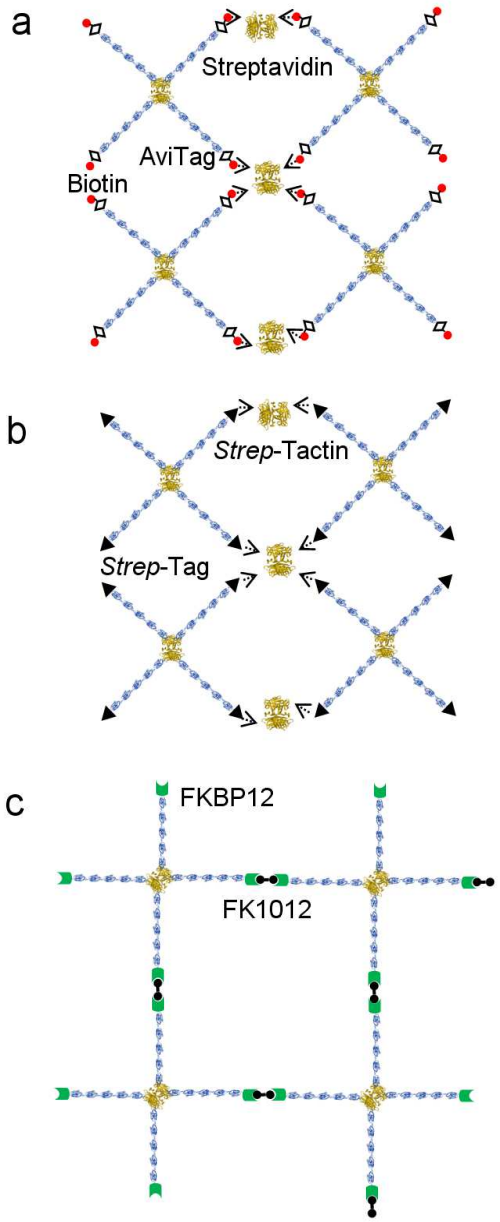

Figure S9. Possible self-assembled protein-based networks and materials. a) By fusing AviTag in the N-terminus of streptavidin-based complexes, biotins can enzymatically bind to the tag. Then, mixing soluble streptavidin can trigger the construction of either 2D networks or 3D materials. b) *Strep*-tagged protein complexes mixed with *Strep*-Tactin would build self-assembled materials. c) FKBP12 domain[18] can also be fused in the N-terminus of streptavidin-based complexes. By adding FK1012, a chemical inducer of dimerization, 2D networks will be created (adapted from ref. 18).