Non-canonical Lipoproteins with Programmable Assembly and Architecture

Md Shahadat Hossain,[†a] Corina Maller,[†a] Yinghui Dai,[b] Shikha Nangia,[b,c] and Davoud Mozhdehi*[a,c]
[†]These authors contributed equally to this work.

[a] Department of Chemistry, 1-019 Center for Science and Technology, Syracuse University, Syracuse, NY 13244 (USA), E-mail: dmozhdeh@syr.edu.
[b] Department of Biomedical and Chemical Engineering, 343 Link Hall, Syracuse University, Syracuse, NY 13244 (USA).
[c] BioInspired Syracuse, Syracuse University, Syracuse, NY 13244, United States

Electronic Supplementary Information

Table of Contents

1. Materials

The pETDuet-1 vector was purchased from EMD Millipore (Billerica, MA). The chemically competent Eb5alpha and BL21(DE3) cells, restriction enzymes, ligase, and corresponding buffers, as well as DNA extraction and purification kits, were purchased from New England Biolabs (Ipswich, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from A. G. Scientific (San Diego, CA). 12-bromododecanoic acid, apomyoglobin, adrenocorticotropic hormone (ACTH), sinapinic acid, alpha-cyano-4-hydroxycinnamic acid, copper (II) sulfate pentahydrate, (+)-sodium L-ascorbate, dipropargyl ether, propargyl alcohol, and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich (St. Louis, MO). High-performance liquid chromatography-(HPLC) grade acetonitrile, SnakeSkin™ dialysis tubing with 3.5K nominal molecular weight cut off (MWCO), mass spectroscopy grade Pierce™ trypsin protease, Tryptone, Yeast Extract, Sodium chloride, ampicillin, kanamycin, phosphate buffer saline (PBS), myristic acid, DMSO, Polyethylenimine (PEI), and ethanol were purchased from Thermo Fisher Scientific (Rockford, IL). Mini-PROTEAN® TGX Stain-Free™ Precast Gels, Precision Plus Protein™ All Blue Pre-stained Protein Standard, and Precision Plus Protein™ Unstained Protein Standards were purchased from Bio-Rad Laboratories, Inc. (Hercules, CA). AF488-DBCO was purchased from Fluroprobes (Scottsdale, AZ). The carbon-coated grid (CF300-Cu) was purchased from Electron Microscopy Sciences. Deionized water was obtained from a Milli-Q® system (Millipore SAS, France). Simply BlueTM SafeStain was purchased from Novex (Van Allen Way Carlsbad, CA). All chemicals were used as received without further purification.

2. Sequence of IDPP

We used N-myristoyl transferase from Saccharomyces cerevisiae (UniProtKB - P14743) in this study. The N-terminal peptide fragment of S. cerevisiae Arf2 (UniProtKB - P19146, highlighted in red below) was selected as the recognition sequence because of its high affinity for yeast NMT (K_m ~ 0.08 μM).[1] The high affinity of this peptide toward yeast NMT is advantageous because

Devadas et al. reported that binding of ADA-CoA to NMT reduces the affinity of a model octapeptide toward NMT:ADA-CoA complex by 7-fold, compared to the natural substrate (M-CoA).[2]

GLYASKLFSNLGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGV
PGVGVPGGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVP
GVGVPGGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVPG
VGVPGGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVPGV
GVPGGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVPGVG
VPGGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVPGVGV
PGGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVPGVGVP
GGKGVGVPGVGVPGAGVPGVGVPGVGVPGVGVPGVGVPGAGVPGVGVPGVGVPG
GKGY

3. Synthesis of 12-azidododecanoic acid (ADA)

Synthesis of 12-azidododecanoic acid (ADA) was conducted by modifying existing literature procedures.[3,4] Briefly, 12-bromododecanoic acid (400 mg, 1.43 mmol, 1 equiv.) was mixed with sodium azide (465 mg, 7.15 mmol, 5 equiv.) and DMF (10 mL). The heterogeneous reaction mixture was stirred for 12 h at 60 °C. After cooling the reaction mixture, the solvent was evaporated in vacuo at 30°C. The remaining solid was dissolved in 20 mL deionized water and extracted with ethyl acetate (20 ml x 3). The organic layer was washed with 0.1 M HCl (20 mL x 3). The solvent was removed in vacuo to obtain the product as a white waxy solid. Yield: 87%. ^{1}H-NMR (400 MHz, $CDCl_3$): δ [ppm] = 3.29 (t, 2H, N_3-CH_2), 2.38 (t, 2H, CH_2-COOH), 1.65 (m, 4H), and 1.31 (m, 14 H). ^{13}C-NMR (100 MHz, $CDCl_3$): δ [ppm] = 180.30, 51.58, 34.23, 29.44, 29.42, 29.35, 29.19, 29.12, 29.02, 28.84, 26.70, 24.66.

4. Protein Expression, post-translational modification, and purification

4.1 Expression

Protein expression was conducted in BL21(DE3) strains. A 4 mL of sterile 2x YT medium with a corresponding antibiotic (ampicillin or kanamycin, see Table S1) was inoculated with a single bacterial colony. The culture was then shaken at 37 °C on an orbital shaker

at 200 rpm. After overnight growth, this suspension was used to inoculate each 1 L of sterile 2x YT medium. The bacteria were grown in an orbital shaker incubator at 37 °C at 180 rpm. After reaching OD_{600} ~ 0.8, the culture media was supplemented with the fatty acid, myristic acid or ADA, at the final concentration of 500 μM. After 15 min, expression was induced by the addition of IPTG to a final concentration of 1 mM. For non-myristoylated constructs, no myristic acid was added, and the protein expression was induced by adding IPTG to the culture at the OD of 0.8. The expression of proteins was then continued for 4 hours.

Note on optimization of expression condition: Devadas et al. quantified the substrate preference of NMT to accept ADA or M by quantifying the amount of modified peptide substrate (GARASVLS) using an in vitro end-point assay.[2] In this assay, NMT exhibited a comparable preference for transferring ADA or M to this octapeptide. The observed comparable substrate preference is due to the contribution of several opposing factors:

1) This assay requires activation of M or ADA to their corresponding acyl-CoA using Pseudomonas acyl-CoA synthetase. This enzyme exhibited a strong preference toward ADA over M and produced 1.5-2 times more ADA-CoA compared to M-CoA in vitro.
2) NMT binds the peptide substrate only after binding acyl-CoA. In follow-up kinetic studies, NMT is shown to have a comparable affinity for M-CoA or ADA-CoA (K_m = ~ 4-6 μM). However, the binding of ADA-CoA reduced the NMT affinity for its peptide substrate. K_m for the peptide was increased to 66 μM, compared to 9 μM for binding of the peptide to NMT:Myr-CoA complex.
3) Once the ternary complex (NMT:acyl-CoA:Peptide) was formed, NMT exhibited higher efficiency in transferring the ADA-CoA to N-terminal glycine residue, V_m (ADA) = 1.5 × V_m (M).

Based on the kinetic data, NMT has a higher preference for M vs. ADA in vitro, based on V_m/K_m as a measure of catalytic efficiency. We note that the apparent substrate preference of NMT inside E.coli is influenced by the intracellular pool of ADA-CoA and M-CoA. The concentration of these activated fatty acids is influenced by several (time-dependent) factors[5]:

1) Transport of ADA across the cell envelope (diffusion or by fatty acid transporter, FadL)
2) The preference of E. coli acyl-CoA synthetase (FadD) to accept ADA as a substrate

3) The intracellular pool of Myr-CoA from the endogenous fatty acid biosynthesis pathway
4) Possible degradation (β-oxidation) of ADA-CoA in the fatty acid degradation pathway
5) Basal expression of NMT and IDPP in cells before induction and addition of ADA

Additionally, we cannot rule out that differences in the polypeptide substrate further changes the kinetics of acyl transfer in vivo. We used the N-terminal sequence of Arf2 fused to a large IDPP in our study, but Devadas et al. used a short octapeptide derived from $Pr55^{gag}$ polypeptide precursor of human immunodeficiency virus I.

In our hand, we have noted that the relative concentration of ADA-IDPP vs. M-IDPP decreases with increasing the induction time (Fig. S4). Therefore, we opted for using a higher concentration of ADA in the expression culture and reducing the expression time.

4.2 Purification

The cells were harvested by centrifugation at 3745 x g at 4 °C for 30 min. The bacterial pellet was resuspended in phosphate buffer saline (PBS, 15 mL per 1 L of expression culture). The cells were then lysed by two rounds of sonication (10s on, 60s off, total sonication time of 90s. Power: 60-80 w). To ensure complete dissipation of the heat generated during sonication cycles, the sample was kept on ice for 10 min after each cycle. Polyethylenimine was then added to the lysed culture (2 mL of 50% w/v solution per 1 L of expression culture) to remove the DNA fragments. The lysed bacterial solution was centrifuged (22,830 x g, 4 °C, 15 min) to separate the protein from insoluble cell debris and PEI-complexed nucleic acids. IDPP, M-IDPP, and ADA-IDPP were then purified from the endogenous *E.coli* proteins by leveraging their temperature-triggered liquid-liquid phase separation using inverse transition cycling.[6] Each protein was further purified by preparative HPLC to ensure purity (>95%) for the self-assembly studies. Reverse-phase HPLC (RP-HPLC) was performed with a Shimadzu HPLC system (Phenomenex Jupiter® 5 µm C4 300 Å, LC Column 250 × 10 mm, solvent A: H_2O + 0.1% TFA, solvent B: acetonitrile + 0.1% TFA). The percentage of the organic solvent in the mobile phase was increased from 0 to 90% over the course of 35 minutes. After HPLC purification, the organic solvent was removed by dialysis against water using SnakeSkin™ Dialysis Tubing (7000 MWCO, Thermo Scientific) overnight, followed by

lyophilization. Lyophilized proteins were kept at -20 °C for long-time storage. Purified protein yield per L of culture: IDPP (~10 mg); M-IDPP (~8 mg); ADA-IDPP (~3-5 mg).

4.3. Synthetic modification of ADA-IDPP

4.3.1 Synthesis of bolaamphiphile (BMT-IDPP$_2$)

Bolaamphiphile was synthesized by joining two ADA-IDPP chains via bis(methoxy-1,2,3-triazole), BMT, using a telechelic alkyne (dipropargyl ether), Fig. S12. ADA-IDPP (100 μM in water, 1.1 equiv) was mixed with dipropargyl ether (45 μM, 1.0 equiv) at room temperature. Aqueous solutions of $CuSO_4$ and the accelerating ligand, tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), were initially mixed in a separate tube for 5 minutes to form the THPTA–Cu^{2+} complex before addition to the reaction mixture (working concentrations: [$CuSO_4$] = 250 μM and [THPTA] = 1.25 mM). The reaction was then started by adding sodium ascorbate to reduce the copper complex in situ to start the copper catalyzed alkyne azide cycloaddition (CuAAC). The progress of the reaction was monitored by using SDS-PAGE as well as analytical HPLC. After 3-6 hours, excess EDTA was added to the reaction mixture to stop the reaction. Bolaamphiphile construct was separated from excess ADA-IDPP using preparative HPLC. SDS-PAGE (Fig. S12b) and Analytical HPLC (Fig. S12c) were used to confirm the successful purification of BMT-IDPP$_2$ (green trace, t_R = 29.9 min) from excess ADA-IDPP (red trace, t_R = 31.1 min).

Given that the LCST of bolaamphiphile (and its concentration dependence) is identical to ADA-IDPP (Table S3), we did not attempt to purify these two proteins from the reaction mixture using traditional inverse transitional cycling, i.e., by adding kosmotropic salts to the mixture to trigger the LCST phase transition isothermally. However, we note that unlike ADA-IDPP, bolaamphiphile undergoes a reversible phase transition at elevated temperatures (Fig. 2b and Fig. S14c). Consequently, we suggest that it may be possible to separate IBMT-IDPP$_2$ from the ADA-IDPP by thermal annealing at elevated temperatures. This condition results in the irreversible transition of ADA-IDPP into fibers with hysteric aggregation, which can then be separated from soluble BMT-IDPP$_2$ constructs at temperatures below LCST.

4.3.2. Synthesis of the control hydroxylmethyl-1,2,3-triazole-IDPP (HMT-IDPP)

A control construct (HMT-IDPP, single-tail amphiphile in which the terminal azide is converted to a triazole moiety) was synthesized using CuAAC by reacting ADA-IDPP (100 μM, 1 equiv) with excess propargyl alcohol (1.0 mM, 10 equiv) using a reaction protocol described for BMT-IDPP$_2$. HMT-IDPP was purified using HPLC, and its purity was confirmed using SDS-PAGE (Fig. S12b, lane 2) and analytical HPLC (Fig. S12c, purple trace, t_R= 29.0).

5. Computational method

The molecular structure of M-peptide and ADA-peptide molecules were built in two-steps. First, the three-dimensional structure of the recognition peptide (GLYASKLFSNL) was determined using the I-TASSER webserver.[7] In the second step, the M and ADA lipid tails were added to the glycine residue of the peptide using the CHARMM-GUI webserver.[8] The individual M-peptide and ADA-peptide structures were equilibrated at 300 K using the CHARMM36 all-atom force field available within the CHARMM-GUI workspace.

The equilibrated structures were used to construct two explicitly solvated systems containing: (i) 15 molecules of M-peptide and (ii) 15 molecules of ADA-peptide. Each system contained 150 mM NaCl solution and TIP3P water molecules in a cubic box with a box length of 10 nm. The systems were energy minimized and equilibrated at T = 300 K in the isothermal-isochoric (NVT) and at P = 1 bar in isothermal-isobaric (NPT) ensemble constraints in the CHARMM-GUI workflow. The equilibrated output from these runs was used to study self-assembly behavior.

The self-assembly molecular dynamics simulations were performed using the GROMACS 2019.4 simulation package.[9] The M-peptide and ADA-peptide molecules and ions were modeled using the CHARMM36 all-atom force field.[10] Water was modeled using TIP3P.[11] The NPT production runs were run for 400 ns using 2 fs timestep. The temperature was maintained at 303.15 K using the Nose-Hoover thermostat with τt = 1.0 ps.[12,13] The pressure was maintained at 1 bar using isotropic coupling using Parrinello-Rahman barostat with τP = 5 ps and compressibility of 4.5 x 10−5 bar^{-1}.[14,15] Periodic boundary conditions were applied in all three dimensions. The nonbonded van der Waals interactions were calculated using a cutoff of 1.2 nm. The long-range electrostatic interactions were studied with particle mesh Ewald (PME) algorithm at a 1.2 nm cutoff.[16]

Further, the behavior of the system during the heating cycle was simulated for both systems over 280−360 K temperature range (280, 290, 300, 310, 320, 330, 340, 350, and 360 K) over a total simulation time of 400 ns. Analyses of the results from thermal annealing simulations, including radius of gyration (Rg) and solvent accessible surface area (SASA), were performed using GROMACS built-in utilities. Molecular visualization was performed using VMD software.[17]

6. Characterization

6.1 Fourier Transform Infrared Spectroscopy

Fourier transform infrared spectroscopy (FT-IR) spectra were obtained on a Bruker Tensor 27 spectrophotometer with a MIR source and a DLaTGS detector. Spectra were recorded under ambient conditions at a resolution of 4 cm^{-1}. A total of 64 scans were recorded for each spectrum in addition to the background.

6.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

The purity, molecular weight, and fluorescent labelling of the purified proteins were first assessed using SDS-PAGE (10% cross-linked. TGX Stain-Free™ gels). The gels with fluorescently labelled proteins (see below) were initially visualized using a blue LED transilluminator (λ_{ex} = 465 nm) followed by visualization under UV-light (λ_{ex} = 350 nm) using a BioRad Gel Doc EZ imager. The gels were then stained with SimplyBlue™ SafeStain by following the manufacturer protocol and imaged using BioRad Gel Doc EZ imager.

6.3 Fluorescent Labelling

Strain-promoted alkyne-azide cycloaddition reaction was used to selectively label the proteins bearing bioorthogonal azide group (ADA-IDPP). Alexa Fluro 488 fluorophore conjugated to Dibenzocyclooctyne (AF488-DBCO) was dissolved in DMSO at the final concentration of 1 mg/mL. Protein solution (~ 25 μM, 10 μL) of ADA-IDPP or M-IDPP were mixed AF488-DBCO (~ 1.05 mM, 4 μL) in an Eppendorf tube. The reaction mixture was kept in the dark at room temperature on a tube rotator overnight. The proteins were then separated from unreacted AF_{488}-DBCO using SDS-PAGE. The labelled protein

bands were visualized as described above. Total protein content was assayed by staining the gel with SimplyBlue (Fig. S5a).

6.4 Analytical HPLC

Analytical RP-HPLC was performed on a Shimadzu instrument using a Phenomenex Jupiter® 5 µm C4 300 Å, 250 × 4.6 mm LC Column with a mobile phase consisting of a gradient of acetonitrile in water (Table S2) to analyze IDPP, M-IDPP, and ADA-IDPP, Fig. S5b. HMT-IDPP and BMT-$IDPP_2$ were analyzed by a Phenomenex Jupiter® 5 µm C18 300 Å, 250 × 4.6 mm LC Column using the solvent gradient in Table S2 (Fig. S12c). The proteins were analyzed using a photodiode array detector at wavelengths between 190 and 800 nm .

Table S2 The gradient mobile phase composition of analytical HPLC.

Time (min)	%B (CH_3CN + 0.1% TFA)
0	0
5	0
45	90

6.5 Matrix Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS)

MALDI-TOF-MS was conducted on Bruker Autoflex III with smart ion source. A saturated solution of sinapinic acid in 30% acetonitrile was used as the matrix. Samples for MALDI-TOF-MS analysis were prepared by mixing 2 µL of the protein solutions (50 µM) with the matrix (8 µL), followed by serial dilution. These solutions were plated onto a sample plate and dried at room temperature. Apomyoglobin (M_W = 16,952.27 Da) was used as standard.

6.6 Trypsin Digestion of Proteins

To confirm the regioselectivity of lipid modification, proteins were digested with trypsin, and the peptide fragments were analyzed using MALDI-TOF-MS. Trypsin digestion was conducted according to the manufacturer's protocol. Briefly, 9 µL of protein (50 µM in PBS) was added to 10 µL of 50 mM ammonium bicarbonate buffer (pH = 7.8) in an Eppendorf tube. To this mixture, 1 µL trypsin (reconstituted as 1 µg/µL in 50 mM acetic

acid) was added, and the reaction mixture was incubated at 37 °C. After 3 h, the peptide fragments were analyzed by MALDI-TOF-MS. α-cyano-4-hydroxycinnamic acid was used as a matrix and adrenocorticotropic hormone (M_W = 2,464.1989 Da) was used as calibrant for the analysis of the trypsin-digested peptide fragments.

6.7 Turbidimetry assay

Temperature-triggered phase separation of IDPP, M-IDPP, and ADA-IDPP were monitored using an Agilent UV-Vis Spectrophotometer (Cary100) equipped with a Peltier temperature controller by measuring the optical density of the solution at 350 nm (where all proteins had negligible extinction coefficients). Proteins were dissolved in PBS in three concentrations (50, 25, and 12.5 μM) and were equilibrated at 20 °C (below the transition temperature of all constructs) for 5 minutes before heating the solution at the rate of 1 °C/min. M-IDPP, ADA-IDPP, HMT-IDPP, and BMT-$IDPP_2$ were heated to 45 °C, while unmodified IDPP was heated to 60 °C as it had a higher transition temperature. At the end of the heating cycle, the protein solutions were cooled at the rate of 1 °C/min to 20 °C while monitoring the optical density at 350 nm. The transition temperatures were calculated by the method of first derivative using the optical density plotted against the temperature. Transition temperatures are defined as the inflection point (the maximum of the first derivative) in the optical density during the heating cycle. These data were fitted to the following model $T_t = -m \times \ln [IDPP] + T_c$ to derive critical transition temperature (T_c) and the concentration dependence of T_t (m), summarized in Table S3.

Table S3 Critical transition temperature (T_c) and the concentration dependence of T_t derived from turbidity plots.

Construct	m (90% CI)[a]	T_c (90% CI)[b]
IDPP	5.3 (2.8–7.9)	65.9 (57.6–74.2)
M-IDPP	1.1 (0.4–1.8)	37.1 (34.7–39.6)
ADA-IDPP	1.3 (0.4–2.2)	36.4 (33.4–39.4)
HMT-IDPP	2.9 (1.5–4.4)	45.6 (41.9–49.4)
BMT-$IDPP_2$	1.4 (0–2.8)	35.1 (31.5–38.7)

Derived values from fitting the T_t data to the following model: $T_t = -m \times \ln ([IDPP]) + T_c$. [a] °C/ln (μM/μM). [b] °C. 90% confidence intervals are calculated from the linear regression analysis using Graphpad prism 8.4.

6.8 Dynamic Light Scattering (DLS)

DLS experiments were conducted on Zetasizer Nano (Malvern Instruments) with 173° backscatter detector. Protein samples in PBS (50 μM) were filtered through a Millex®-LH Low protein Binding Hydrophilic LCR Membranes (0.45 μm) into the DLS cuvette. The measurements were conducted in triplicates at each temperature. The temperature was increased in 5 °C increments, and the samples were equilibrated at each temperature for 1 min. IDPP was analyzed in the temperature range of 15-60°C (IDPP), and lipidated IDPP (M-IDPP and ADA-IDPP) were analyzed in the temperature range of 15-45 °C. At the end of the heating cycle, the temperature was reduced to 15 °C. Scattering correlation function was analyzed using the Zetasizer software using the method of cumulants and CONTIN to calculate the hydrodynamic radii of each construct at each temperature. The size of IDPP unimer was estimated using the following equation[18], which is a scaling law based on Flory's mean field approach[19]:

$$R_g = \sqrt{\frac{2l_p b}{(2v+1)(2v+2)}}\, N^v = 8.8 \text{ nm}$$

Table S4. Parameters used for estimation of the size of IDPP unimer.

	Parameter	Value (unit)	Reference
l_p	Persistence length= ½ Kuhn length (l_k)*	1.05 nm	[20]
b	Monomer size*	0.365 nm	[20]
N	3 × (number of pentapeptide repeat)	240	[21]
v	Flory exponent for the expanded coil state	3/5	[19]

6.9 Transmission Electron Microscopy (TEM)

The TEM imaging was performed on JEOL-2100F operated at 200 kv and images were recorded by Gatan CCD camera. ADA-IDPP solution (10 μL) used in DLS experiment, i.e., after thermal annealing, were deposited onto carbon-coated grid (CF300-Cu, Electron Microscopy Sciences) or glow discharged carbon-coated copper grids for 5 minutes before blotting the excess solution. The grid was then stained with 10 μL of 1% uranyl acetate and the excess stain was blotted after 1 min. The grid was dried at room temperature for 2 hours followed by vacuum drying for 12 hours before imaging. BMT-$IDPP_2$ solution was heated to 37 °C before depositing a 10 μL of the sample onto the TEM grid. Images of M-IDPP and ADA-IDPP (below LCST, Fig. S8) were obtained in a

similar manner by dissolving the lyophilized protein in cold PBS just before deposition onto the grid.

6.10 Pyrene Assay

We used pyrene assay to determine the critical micellization concentration (CMC) of M-IDPP and ADA-IDPP at 20°C (below transition temperature).[22] The fluorescence of pyrene was measured using a Cary Eclipse Fluorescence Spectrophotometer at the excitation wavelength = 334 nm. The emission signal was recorded between 350-450 nm, at 1 nm intervals with the scan rate of 600 nm/min.

7. Statistical analysis

Statistical analysis was conducted using GraphPad Prism 8.4. The line representing the phase boundary between single-phase and two-phase regimes for each construct (Fig. 2d and Fig. S14b,d) was obtained by linear regression analysis of transition temperatures vs. natural log of concentration. The dashed band represents 90% confidence interval.

The error bars for all DLS measurements (Fig. 3b, Fig. S9, and Fig. S17c) represent standard deviation of three measurements. Two-way Analysis of Variance (ANOVA) in Fig. 3b was conducted using GraphPad prism software.

TEM images were analyzed using Fiji to determine the width of fibers. The result is reported as average ± standard deviation as well as the frequency distributions. The sample size is reported in parenthesis.

8. Supplementary Figures

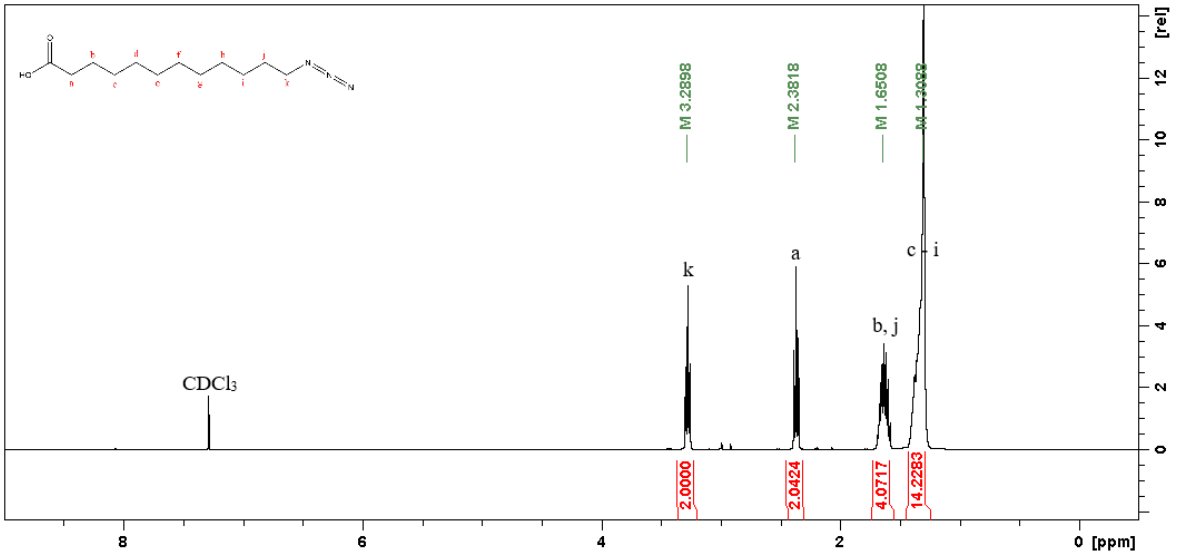

Fig. S1 ^{1}H NMR spectra of ADA.

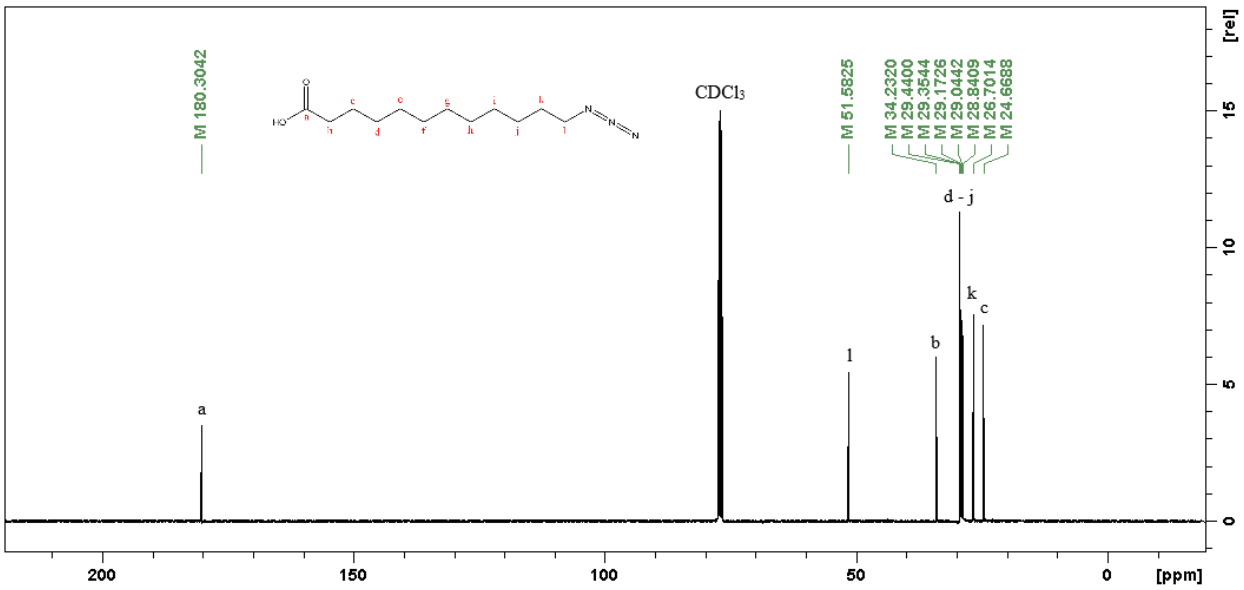

Fig. S2 ^{13}C NMR spectra of ADA.

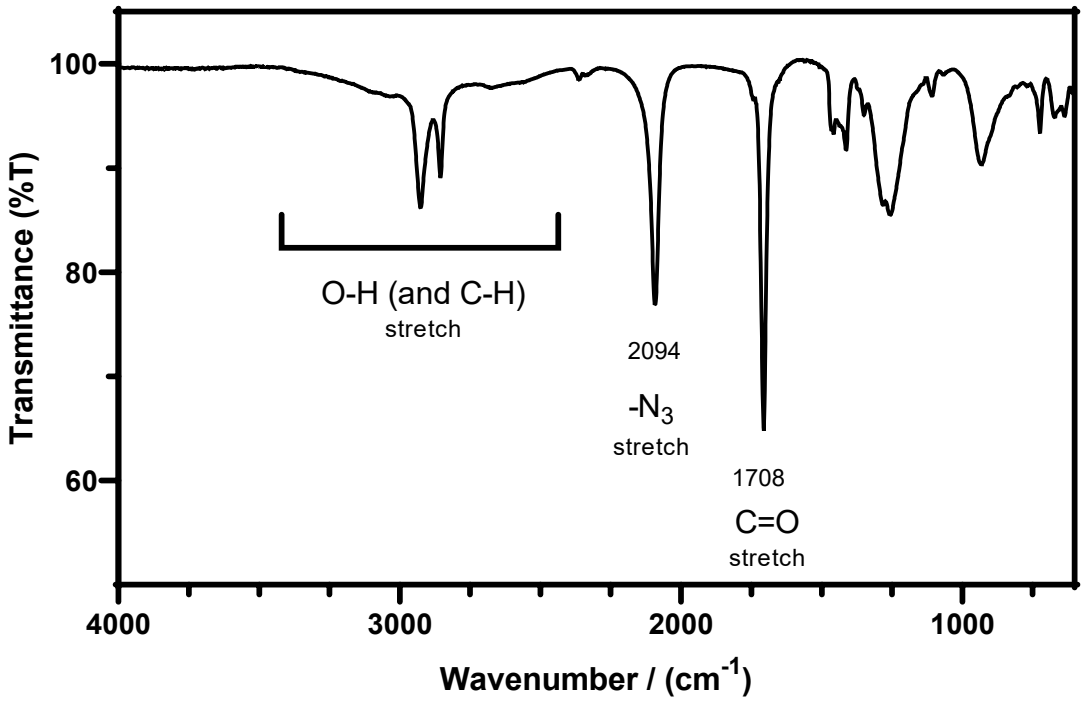

Fig. S3 Attenuated Total Reflection Fourier Transform Infrared Spectroscopy (ATR-FT-IR) spectra of ADA.

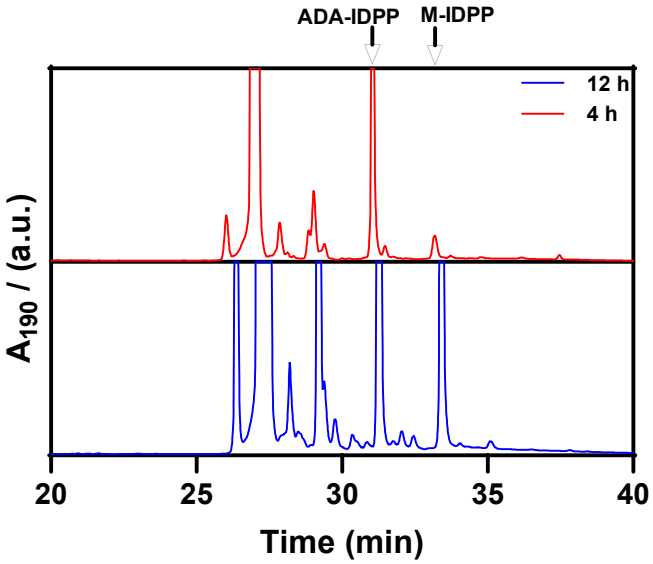

Fig. S4 HPLC chromatogram of "semi-purified" cellular lysate mixtures grown in the presence of ADA. Cells were collected 4 h (red) or 12 h (blue) post-induction. The ratio of ADA-IDPP to M-IDPP decreases with time. Due to the presence of endogenous proteins and metabolites, it is difficult to quantify the ratio of M-IDPP and ADA-IDPP in lysate without isotope labelling. Instead, we opted to first enrich the target lipoproteins by leveraging their phase transition using inverse transition cycling (ITC). This method uses the "reversible" phase behaviour of target proteins to separate them from the endogenous impurities. Therefore, the recovery of proteins with irreversible phase behaviour (ADA-IDPP) can be lower than their analogues with fully reversible phase behaviour (M-IDPP). To avoid this inherent bias, we used a high concentration of sodium chloride (3 M NaCl) to salt-out IDPP and then extracted the pellet with ethanol and water mixture (to break up the aggregates of ADA-IDPP). This ethanolic extract was analysed with analytical HPLC without further purification to compare the relative amount of ADA- and M-IDPP (marked with arrows). Other peaks in the chromatogram are due to the impurities extracted during this procedure.

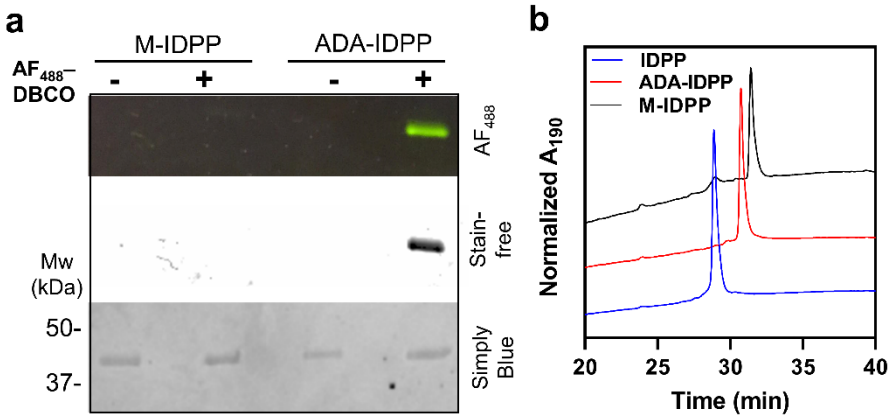

Fig. S5 Molecular characterization confirms the incorporation of ADA. a) Only ADA-IDPP reacts with a fluorophore bearing a dibenzocyclooctyne (DBCO), b) Reverse-phase HPLC was used to quantify the hydrophobicity of each construct by comparing their retention time (t_R). The observed trend—t_R (min) = IDPP (28.8) < ADA-IDPP (30.7) < M-IDPP (31.5)—is consistent with the increased hydrophobicity of M-IDPP compared to ADA-IDPP.

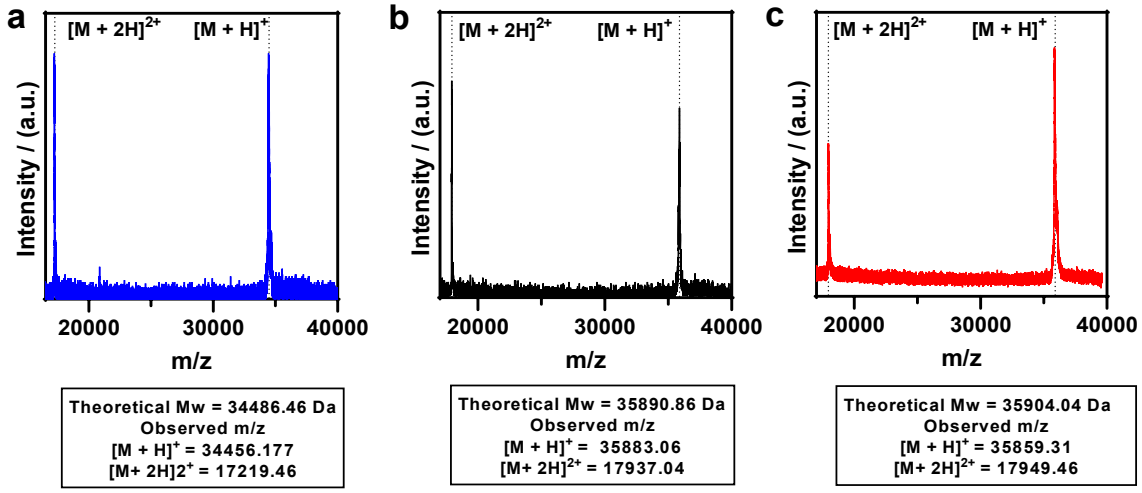

Fig. S6 MALDI-TOF-MS of IDPP (a), M-IDPP (b), and ADA-IDPP (c). Vertical dashed lines denote theoretical molecular weight.

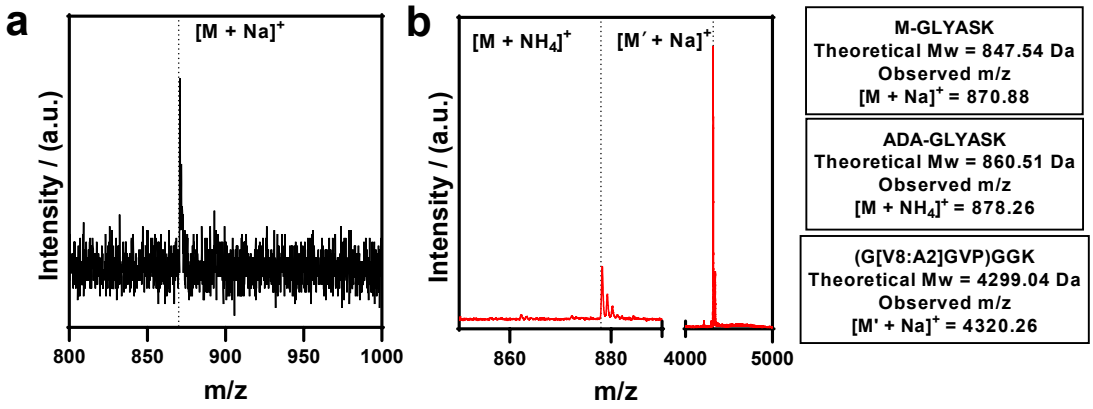

Fig. S7 MALDI-TOF-MS of N-terminal peptide fragment of M-IDPP (a) and ADA-IDPP (b) after digestion with trypsin to confirm the regioselective modification. There are 9 lysines in M-IDPP (and ADA-IDPP), 8 of which are distributed throughout the sequence of IDPP and one is located within the ARF2 recognition sequence. As shown in Fig. S7b, the ADA modification was only observed at the N-terminal, supported by the peak at m/z = 878.26, which was assigned to the ammonium adduct (present in the trypsin digest buffer). We do note that azides may decompose during the MALDI ionization process and have been found to form metastable ions with the observed m/z differing from the theoretical molecular weight.[23] For ADA-IDPP, none of the internal lysine residues were modified and only unmodified IDPP fragments were observed (m/z = 4320.26). Though MALDI-TOF-MS is not quantitative, this result supports the notion that ADA is not added to the ε-amine of the lysine residues and is instead added to the N-terminal of the ARF2 sequence. Vertical dashed lines denote theoretical molecular weight.

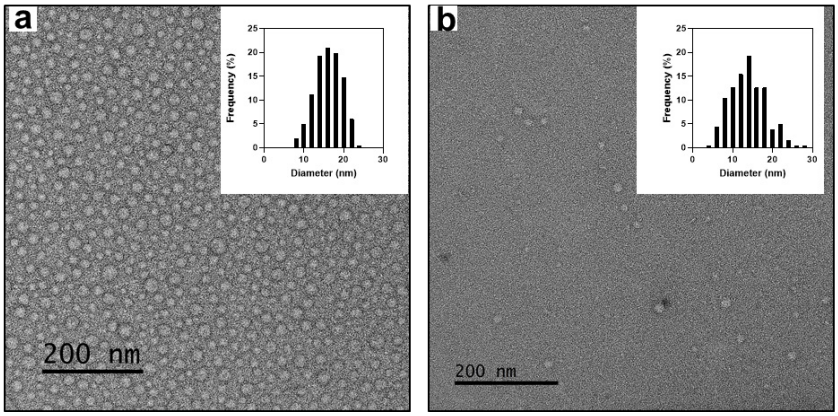

Fig. S8 TEM of negatively stained M-IDPP (a) and ADA-IDPP (b) before thermal annealing. M-IDPP forms spherical micelles with an average diameter of 15.3 ± 3.6 nm (N = 200); ADA-IDPP forms spherical micelles with an average diameter of 12.8 ± 4.5 nm (N = 181). Inset in each panel summarizes the distribution of measured diameters. The observed size of each construct in dried state (TEM) is smaller than the hydrodynamic radius in solution determined by DLS (cumulant method).

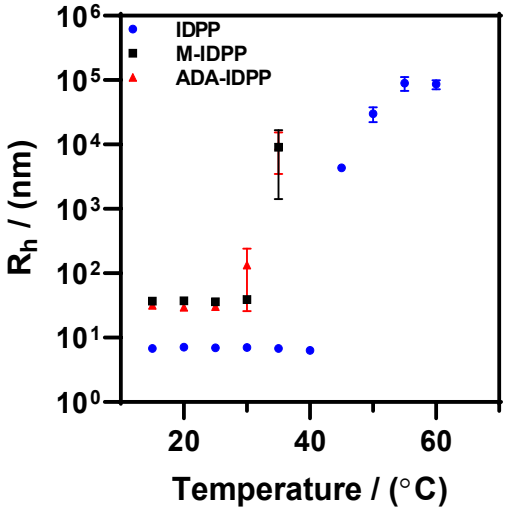

Fig. S9 DLS characterization of the self-assembly of IDPP, M-IDPP, and ADA-IDPP in solution at different temperatures. The hydrodynamic radius of the constructs remained unchanged below T_t, and all three constructs formed large macroscopic (> μm size) polymer-rich coacervates above T_t, consistent with the size of polymer-rich coacervates formed by IDPPs.[24] The temperature of the protein solution was increased in 5 °C steps, and the protein solution was incubated at each temperature for 1 min before the start of the DLS measurement. The observed T_t from DLS experiment is in excellent agreement with the T_t observed from the turbidimetry for IDPP, ADA-IDPP, and M-IDPP. The error bars represent standard deviation of three measurements.

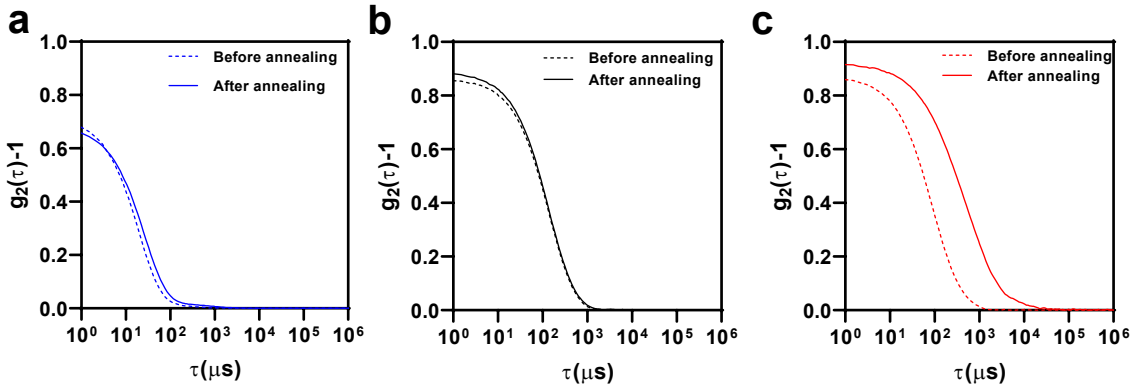

Fig. S10 The DLS auto-correlation functions of IDPP (a) and M-IDPP (b) and ADA-IDPP (c) before and after thermal annealing (dashed and solid lines).

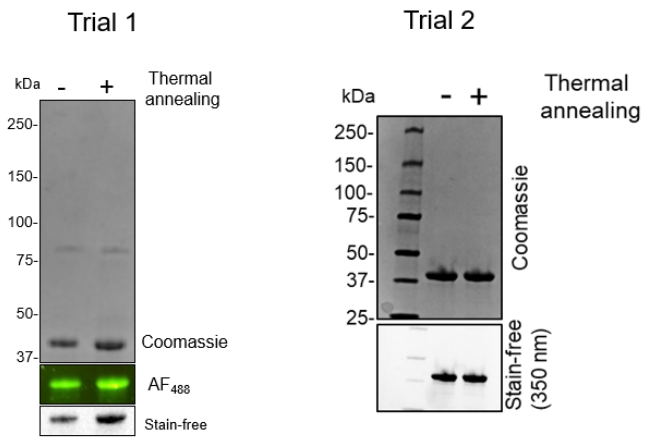

Fig. S11 Azide moiety in ADA-IDPP does not decompose during thermal annealing. To ensure that the azide group is still present in the self-assembled fibers after thermal annealing, AF_{488}-DBCO was added to ADA-IDPP after heating and cooling the protein solution above and below LCST (thermal cycle: 20°C → 40°C → 20°C). Dynamic light scattering was used to confirm the formation of larger hysteric aggregates after the completion of the thermal cycle. As a control, a solution of ADA-IDPP before thermal annealing, i.e., stored in PBS at T = 20 °C < LCST), was also mixed with the fluorophore. Both control and thermally annealed samples (trial 1, lane 1 and 2) exhibited similar labeling with AF_{488}-DBCO dye, consistent with the stability of the terminal azide under the thermal annealing conditions. Additionally, the thermally annealed sample does not contain any higher molecular weight oligomers that are expected form the cross-linking reaction mediated by nitrenes (reactive intermediates formed during azide decomposition).[25] Similar results were observed when the protein solution was heated to 60 °C, trial 2, and cooled to 20°C. These two observations support our hypothesis that the aliphatic azide of ADA is thermally stable in the experimental range and does not decompose during the assembly.

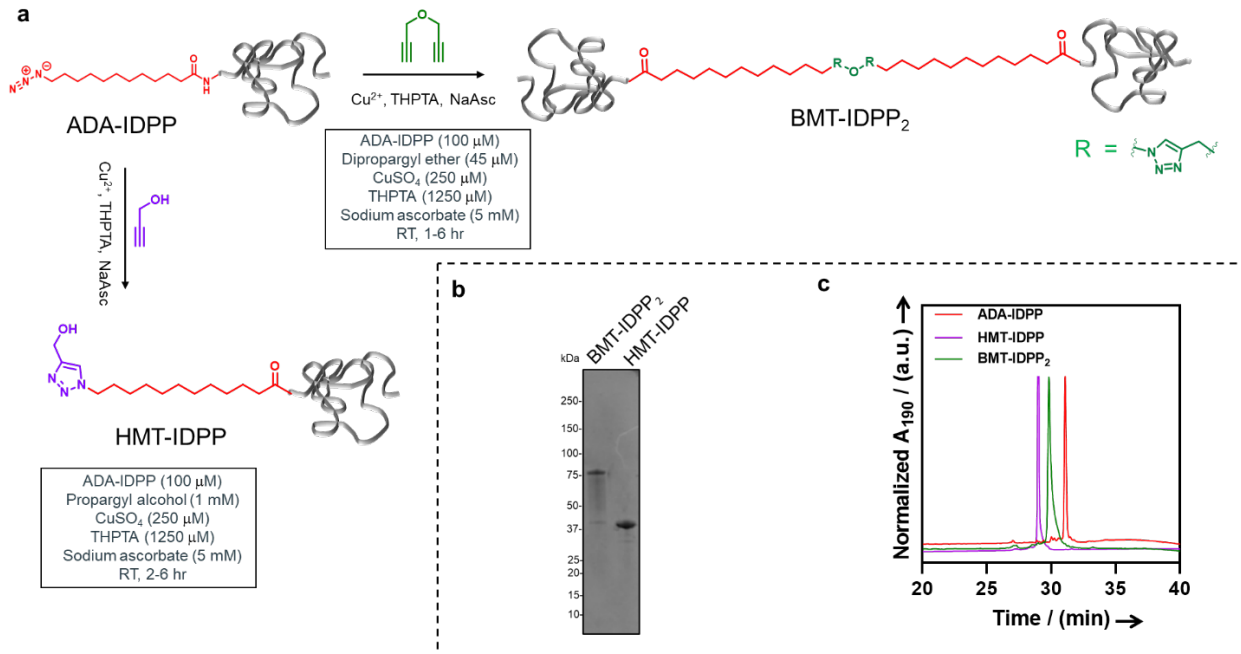

Fig. S12 Copper assisted alkyne azide cycloaddition (CuAAC) can be used to functionalize the ω-azide moiety of the non-canonical lipoprotein. a. Schematic of reactions of ADA-IDPP with a telechelic alkyne (dipropargyl ether, green) to convert the single tail amphiphile into a bolaamphiphile (BMT-IDPP$_2$, BMT: bis(methoxy-1,2,3-triazole); and the reaction of ADA-IDPP with propargyl alcohol (purple) to convert azide into functionalized triazole (HMT-IDPP, HMT: hydroxylmethyl-1,2,3-triazole). b. SDS-PAGE analysis of purified BMT-IDPP$_2$ and HMT-IDPP. c. RP-HPLC chromatogram of ADA-IDPP, HMT-IDPP, BMT-IDPP$_2$ using a gradient of water and acetonitrile on a C18 column. The retention time of each construct is inversely correlated to the polarity of the lipid chain end.

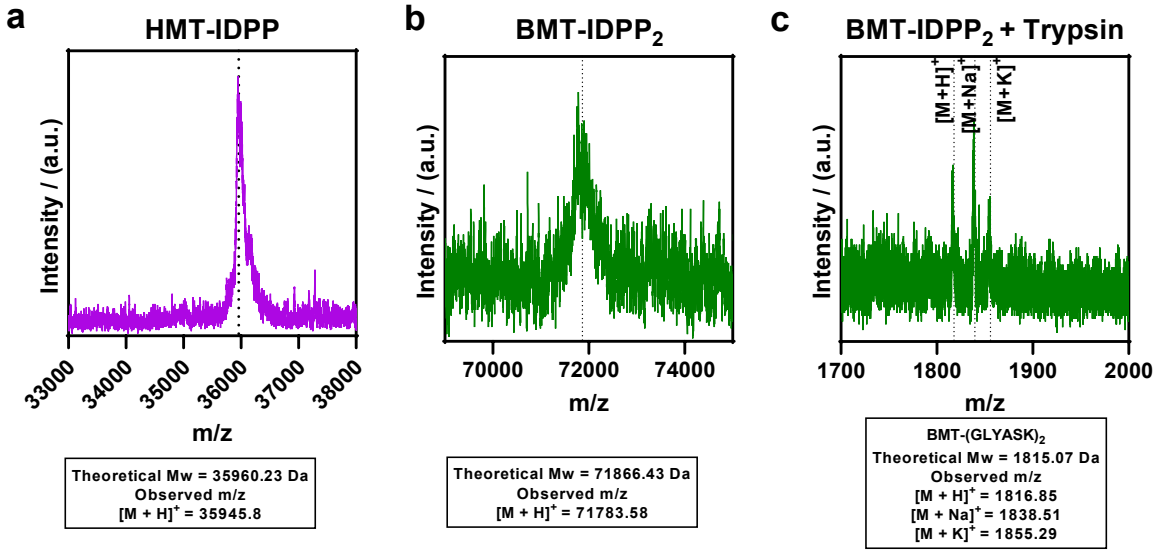

Fig. S13 MALDI-TOF-MS of HMT-IDPP (a) and BMT-IDPP$_2$ (b). MALDI-TOF-MS of N-terminal peptide fragment of BMT-IDPP$_2$ digested with trypsin (c). Vertical dashed lines denote theoretical molecular weight.

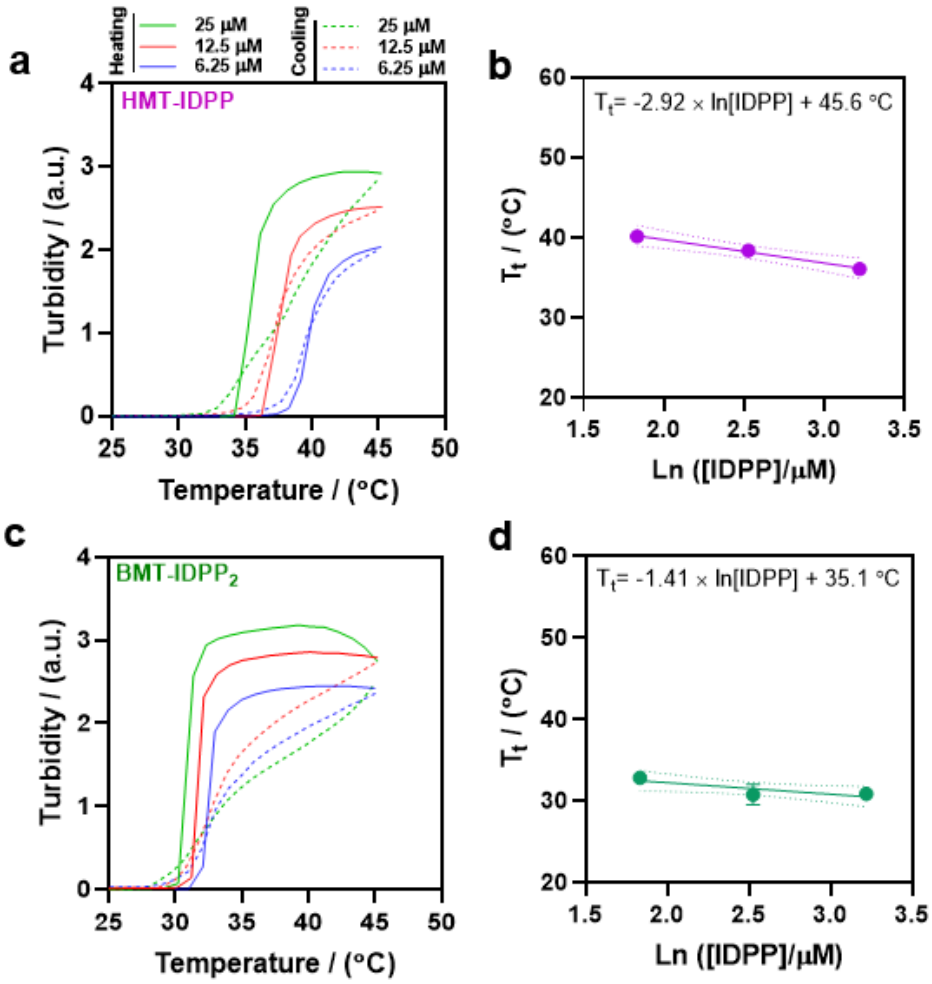

Fig. S14 Temperature-programmed turbidimetry was used to monitor the LCST phase-transition of HMT-IDPP (a,b) and BMT-IDPP$_2$ (c,d). Both HMT-IDPP and BMT-IDPP$_2$ exhibits reversible LCST phase behavior. Since the concentration of IDPP influences the LCST, we chose to use the concentration of the IDPP domain instead of the nominal concentration of each (i.e., [HMT-IDPP] = [IDPP]; 2 x [BMT-IDPP2] = [IDPP]). The phase-diagram of HMT-IDPP (b) and BMT-IDPP$_2$ (d), showing the boundaries between the single phase (below each line) and the two-phase region (above each line). Both lipoproteins exhibited lower transition temperatures compared to unmodified IDPP at identical concentrations (Fig. 2a). However, the concentration dependence of T_t was different for each construct. The dashed lines represent the 90% confidence interval of the line fitted to the experimentally observed transition temperatures, see the summary of derived parameters in Table S3.

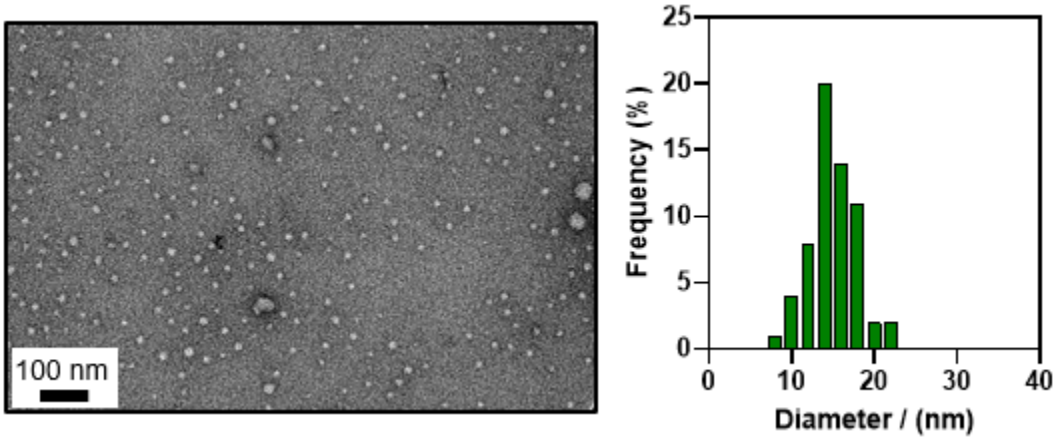

Fig. S15 TEM of BMT-IDPP$_2$ below LCST which forms spherical aggregates with an average diameter of 14.1 ± 3.0 (N= 62), and the distribution of measured diameters.

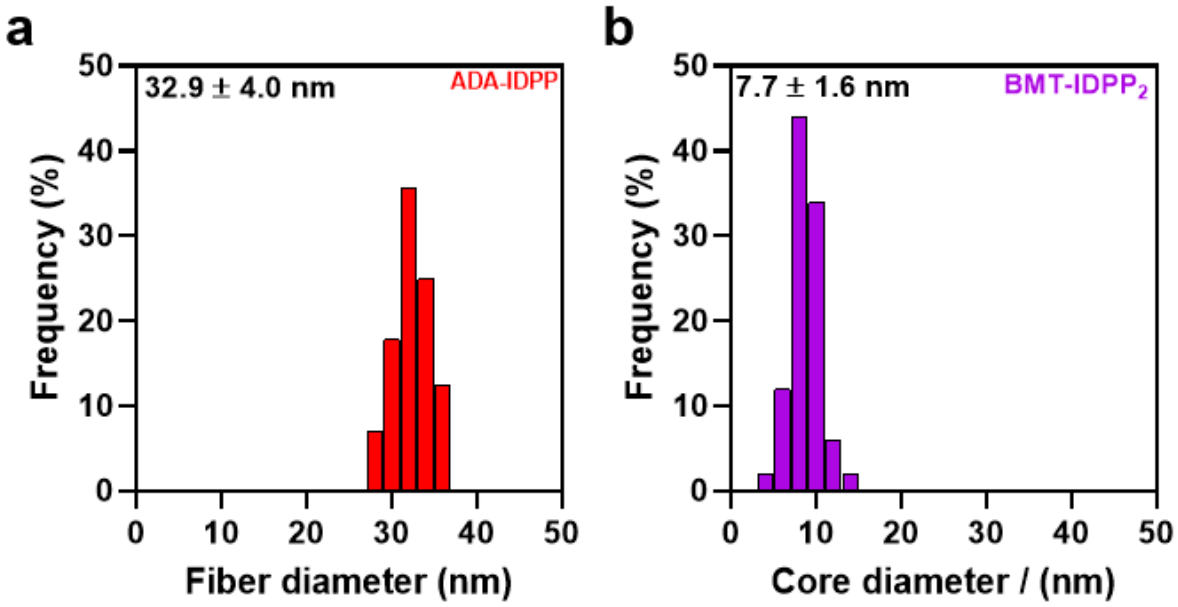

Fig. S16 (a) The distribution of measured diameter of fibers formed by ADA-IDPP after thermal annealing (N= 80), corresponding to Fig. 4a. (b) The distribution of core diameters in bottlebrush structures formed by BMT-IDPP2 (N=50), corresponding to Fig. 4b.

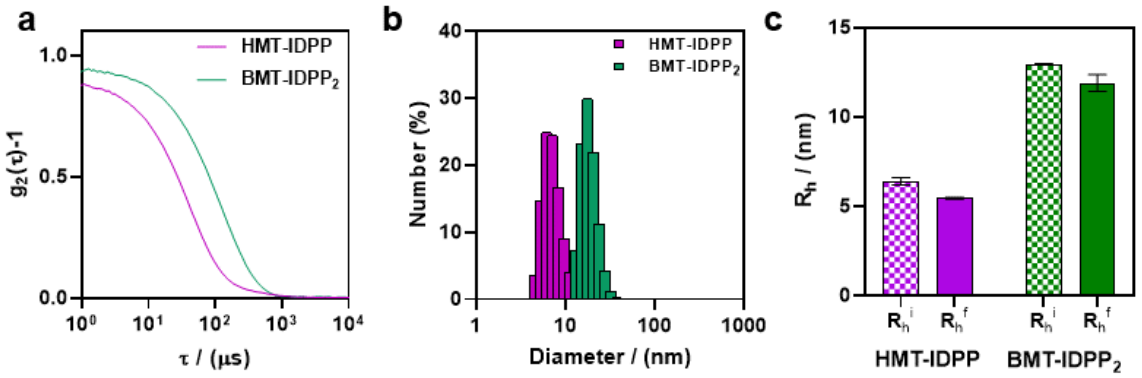

Fig. S17 (a) The DLS auto-correlation functions of HMT-IDPP and BMT-IDPP$_2$ below LCST. (b) Number-size distributions of HMT-IDPP and BMT-IDPP$_2$. HMT-IDPP does not self-assemble as it formed small particles with an average (peak mean) diameter of 7.3 ± 0.5 (which is consistent with the size of unassembled IDPP with comparable size 5.2 nm[20]). BMT-IDPP$_2$, on the other hand, formed nanostructures with an average (peak mean) diameters of 13.2 ± 0.4 nm, which is significantly larger than the expected hydrodynamic radii of a coiled polypeptide with this length, i.e., for IDPP-IDPP, Rh = ~7.4 nm[20]). (c) Cumulants analysis (hydrodynamic radius) of the size of HMT-IDPP and BMT-IDPP$_2$ before and after thermal annealing (checkered versus solid fill). No irreversible increase in the size of the aggregates is observed for these two constructs, consistent with the reversible LCST phase-transition observed in the turbidimetry (Fig. S14).

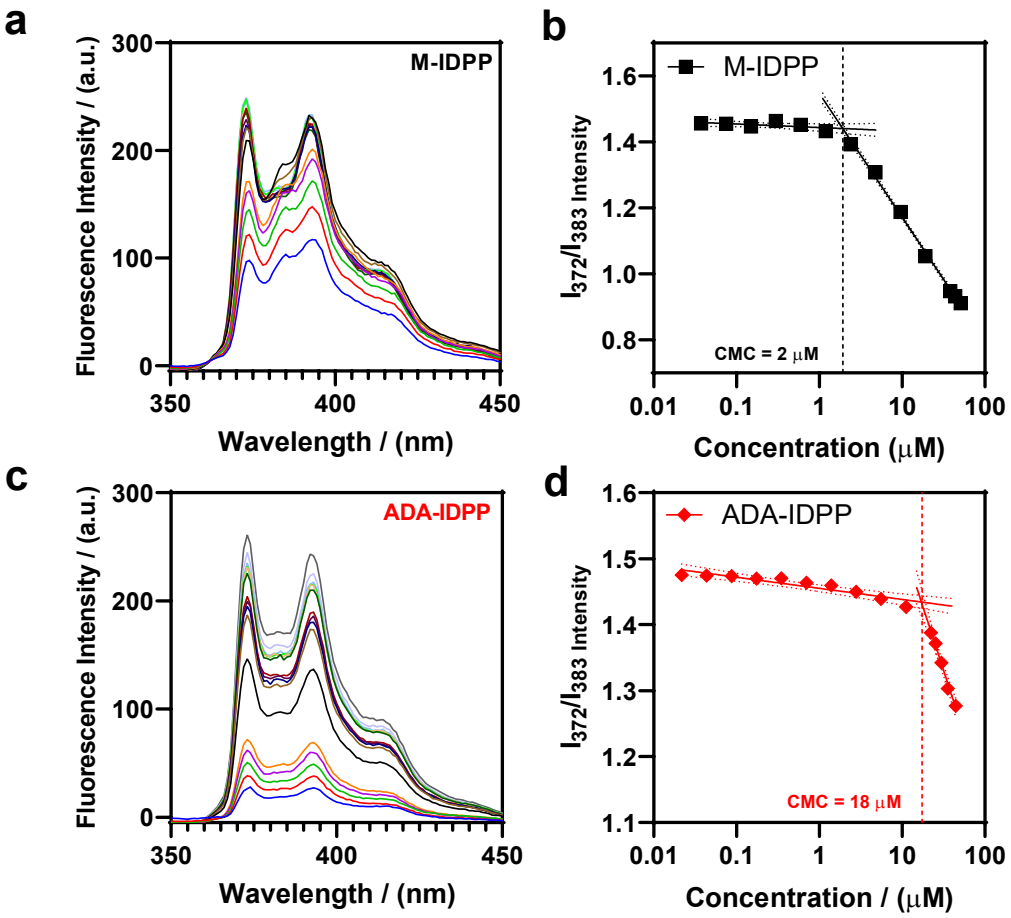

Fig. S18 Pyrene fluorescence assay to determine the CMC of M-IDPP (a,b) and ADA-IDP (c,d). Raw fluorescence data is plotted in (a) and (c). Calculated peak 1 to peak 3 absorbance ratios for pyrene fluorescence are plotted against M-IDPP and ADA-IDPP concentrations, made in serial dilutions. The calculated CAC is shown as a dashed, vertical line for M-IDPP (b) and ADA-IDPP (d).

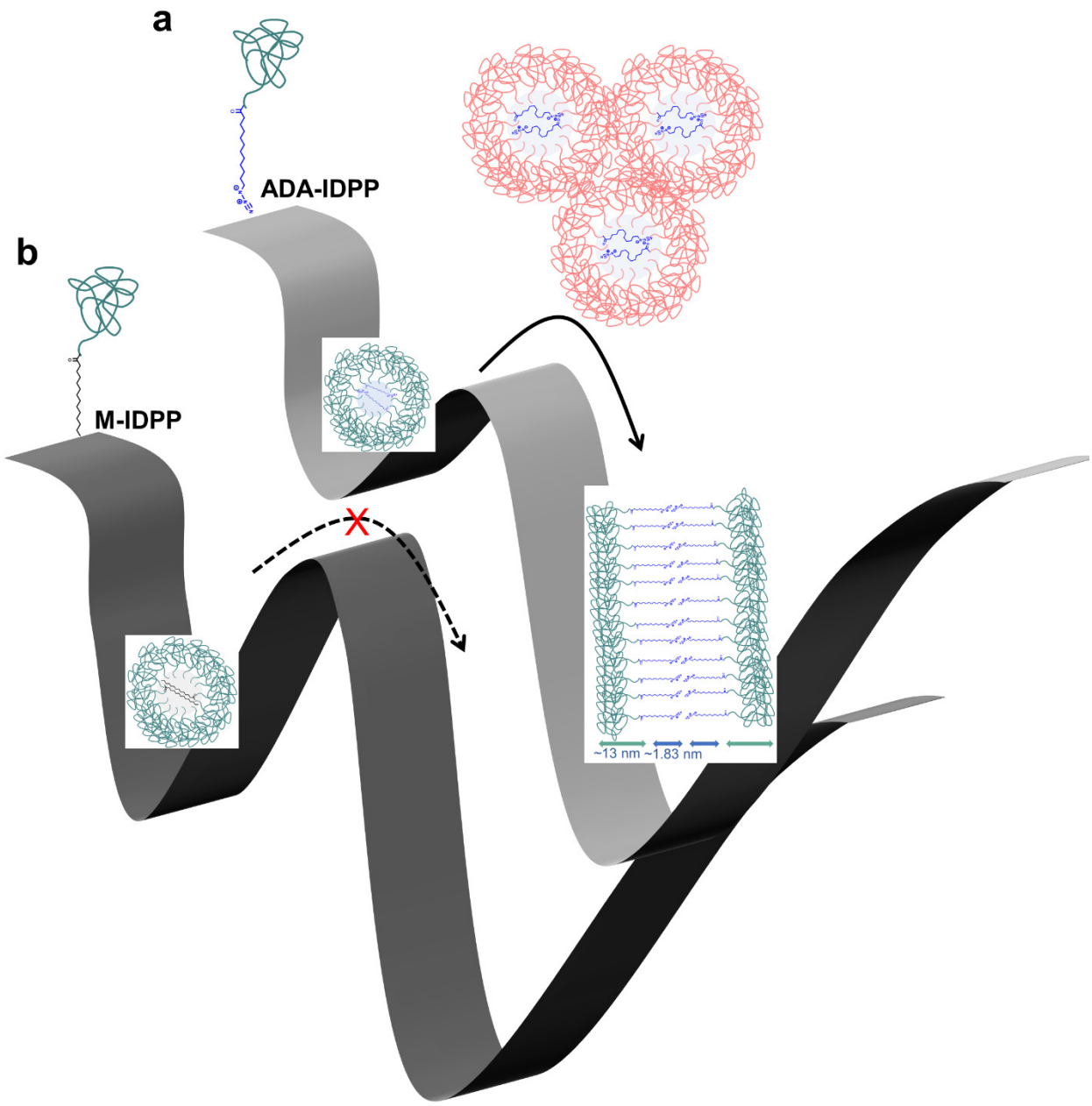

Fig. S19 Proposed mechanism to explain the unique temperature-triggered self-assembly observed by ADA-IDPP. a) ADA-IDPP forms micelles below the LCST but the cores of these micelles are more dynamic than the M-IDPP micelles due to the lower melting point of the ADA. Heating of the sample above the LCST and T_m of ADA can lead to partial melting of the core and rearrangement of the lipoproteins into fibers after thermal annealing. b) M-IDPP micelles form coacervates above the LCST but the rearrangement of the lipoprotein from micelles to other nanostructures is hindered due to the stability of the micellar core. Once the temperature is lowered below T_t, the coacervates will dissolve to form the original micelles.

We propose the following model (Fig. S19) to explain the observed self-assembly of ADA-IDPP and its unique behavior (hysteresis and fiber formation) compared to its natural analogue of myristic acid. The modification of IDPPs with lipids (myristic acid or ADA) results in the self-assembly of proteins into spherical micelles driven by the difference in the hydrophobicity of the lipid tail and the polypeptide chain, consistent with the results of turbidimetry and DLS. The lipid group is likely positioned at the core of these micelles while the hydrated IDPP chain forms the corona of these structures. As the temperature is increased above the transition temperature of the IDPP domain, the polypeptide chains in the corona undergo the LCST transition, thus becoming more hydrophobic as the bound water molecules are released from the peptide chains. This LCST transition results in the coalescence of micelles with each other as the polypeptides undergoes liquid-liquid phase separation to form coacervates (red chains in Fig. S19a).

We suggest that the difference between the assembly of M-IDPP and ADA-IDPP can be explained by considering the packing efficiency (or mobility) of the hydrophobic cores of these structurally-related lipoproteins during the LCST phase transition. Unlike myristic acid, the ADA chain cannot efficiently pack inside the hydrophobic core due to the linear arrangement of the terminal azide group ($-N_3$) compared to the terminal propyl group in myristic acid. This difference in packing efficiency is supported by differences in the observed critical micelle concentration for ADA-IDPP (18 μM) and M-IDPP (2 μM), Fig. S19. Additionally, the melting point (T_m) of ADA is lower than myristoyl group (~30 °C *vs*. 54 °C), suggesting that part of the hydrophobic core in ADA-IDPP is melted as the sample is heated. This increased mobility can facilitate the rearrangement of the ADA-IDPP chains in the coacervate phase during the heating cycle, leading to the conversion of the spherical micelles into rod-like aggregates due to the molecular anisotropy of the amphiphile (*i.e.*, only one terminus is lipidated). On the other hand, the M-IDPP micelles are likely to be stable at 30–40 °C (lower than the myristic acid melting temperature) as the lower mobility of the hydrophobic cores precludes rearrangement of the lipidated polypeptides. As the temperature is reduced below the T_t, the polypeptide chains are rehydrated. In the case of M-IDPP, this will result in the dissolution of the coacervates back into the original micelles. In contrast, for ADA-IDPP, the barrier to convert the rod-like aggregates and extended fibers into the original micelles is likely to be high, and

therefore these structures are other nanostructures is hindered due to the stability of the micellar core. Once the temperature is lowered below T_t, the coacervates will dissolve to form the original micelles.

9. References

1 W. J. Rocque, C. A. McWherter, D. C. Wood and J. I. Gordon, *J. Biol. Chem.*, 1993, **268**, 9964–9971.

2 B. Devadas, T. Lu, A. Katoh, N. S. Kishore, A. C. Wade, P. P. Mehta, D. A. Rudnick, M. L. Bryant, S. P. Adams, Q. Li, G. W. Gokel and J. I. Gordon, *J. Biol. Chem.*, 1992, **267**, 7224–7239.

3 L. Liu, J. A. Shadish, C. K. Arakawa, K. Shi, J. Davis and C. A. DeForest, *Adv. Biosyst.*, 2018, **2**, 1800240.

4 S. H. Ho and D. A. Tirrell, *J. Am. Chem. Soc.*, 2016, **138**, 15098–15101.

5 Y. Fujita, H. Matsuoka and K. Hirooka, *Mol. Microbiol.*, 2007, **66**, 829–839.

6 D. E. Meyer and A. Chilkoti, *Nat. Biotechnol.*, 1999, **17**, 1112–1115.

7 J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson and Y. Zhang, *Nat. Methods*, 2015, **12**, 7–8.

8 S. Jo, T. Kim, V. G. Iyer and W. Im, *J. Comput. Chem.*, 2008, **29**, 1859–1865.

9 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**–**2**, 19–25.

10 J. Huang and A. D. MacKerell, *J. Comput. Chem.*, 2013, **34**, 2135–2145.

11 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

12 G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.

13 C. Braga and K. P. Travis, *J. Chem. Phys.*, 2005, **123**, 134101.

14 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

15 S. Nosé and M. L. Klein, *Mol. Phys.*, 1983, **50**, 1055–1076.

16 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

17 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

18 H. Hofmann, A. Soranno, A. Borgia, K. Gast, D. Nettels and B. Schuler, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 16155–16160.

19 P. J. Flory, *J. Chem. Phys.*, 1949, **17**, 303–310.

20 S. Fluegel, K. Fischer, J. R. McDaniel, A. Chilkoti and M. Schmidt, *Biomacromolecules*, 2010, **11**, 3216–3218.

21 W. Hassouneh, E. B. Zhulina, A. Chilkoti and M. Rubinstein, *Macromolecules*, 2015, **48**, 4183–4195.

22 K. Kalyanasundaram and J. K. Thomas, *J. Am. Chem. Soc.*, 1977, **99**, 2039–2044.

23 Y. Li, J. N. Hoskins, S. G. Sreerama and S. M. Grayson, *Macromolecules*, 2010, **43**, 6225–6228.

24 J. R. Simon, N. J. Carroll, M. Rubinstein, A. Chilkoti and G. P. López, *Nat. Chem.*, 2017, **9**, 509–515.

25 J. F. Arenas, J. I. Marcos, J. C. Otero, I. L. Tocón and J. Soto, *Int. J. Quantum Chem.*, 2001, **84**, 241–248.