Supplemental Information for

Peptides Displayed as High Density Brush Polymers Resist Proteolysis and Retain Bioactivity

Angela P. Blum, Jacquelin K. Kammeyer, Jian Yin, Dustin T. Crystal, Anthony M. Rush, Michael K. Gilson, Nathan C. Gianneschi*

I. PREPARATION OF MATERIALS FOR CELL PENETRATION STUDIES

Ia. Characterization of control peptides and monomers

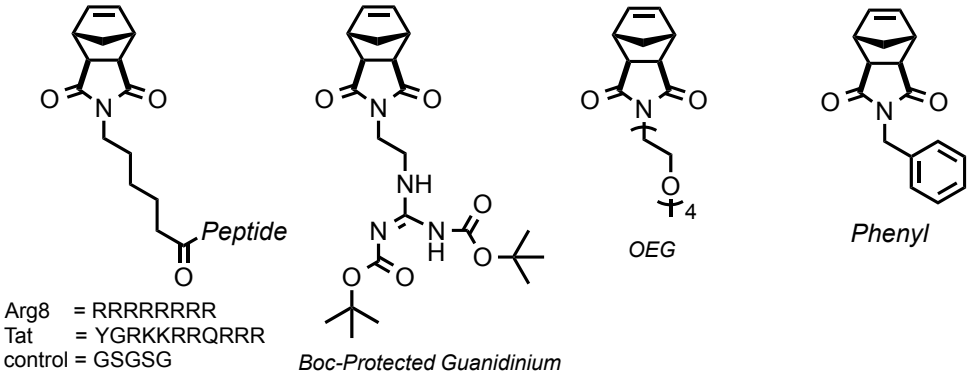

Figure S1. Chemical structures of monomers used in cell penetration studies. Note that the side-chains on the Arg8 monomer are protected with a Pbf group.

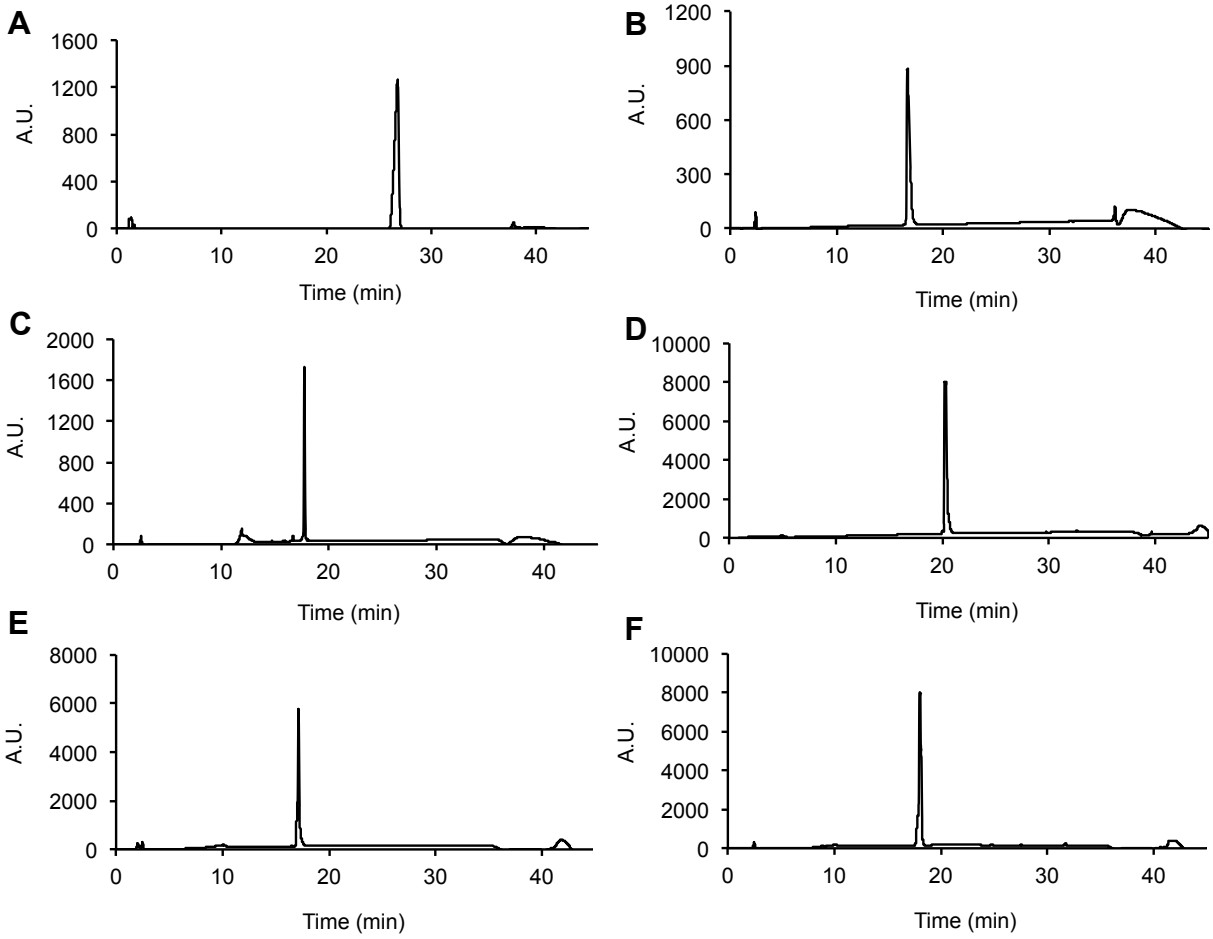

Figure S2. HPLC chromatograms of purified control peptides and peptide monomers. A) NorAha-RRRRRRRR-NH_2 (NorAha-Arg8) at 80-100% B gradient; B) NorAhaYGRKKRRQRRR-NH_2 (NorAha-Tat) at 10-40% B gradient; C) NorAha-GSGSG-NH_2 (NorAha-GSGSG) at 0 to 67% B gradient. D) K(Fluorescein)-RRRRRRRR-NH_2 (Flu-Arg8) at 0-67% B gradient; E) K(Fluorescein)-YGRKKRRQRRR-NH_2 (Flu-Tat) at 0-67% B gradient; F) K(Fluorescein)-GSGSG-NH_2 (Flu-GSGSG) at 0-67% B gradient.

Peptide	Sequence	Mass Calculated	Mass Obtained
Flu-Tat Peptide	K(5/6-Flu)YGRKKRRQRRR-NH_2	2045.4	2044.8
Flu-Arg8 Peptide	K(5/6-Flu)RRRRRRRR-NH_2	1753.02	1752.1
Flu-GSGSG Peptide	K(5/6-Flu)GSGSG-NH_2	848.82	849.1
NorAha-TAT Monomer	NorAha-YGRKKRRQRRR-NH_2	1818.9	1818.5
NorAha-Arg8 Monomer	NorAha-R(Pbf)R(Pbf)R(Pbf)R(Pbf)R(Pbf)R(Pbf)R(Pbf)R(Pbf)-NH_2	3541.6	3542.4
NorAha-GSGSG Monomer	NorAha-GSGSG-NH_2	622.3	623.3

Table S1. Calculated and obtained molecular weight values of peptides prepared for cell penetration studies. Tat and Arg8 masses were obtained via MALDI-TOF MS and GSGSG masses were taken on ESI (+) MS.

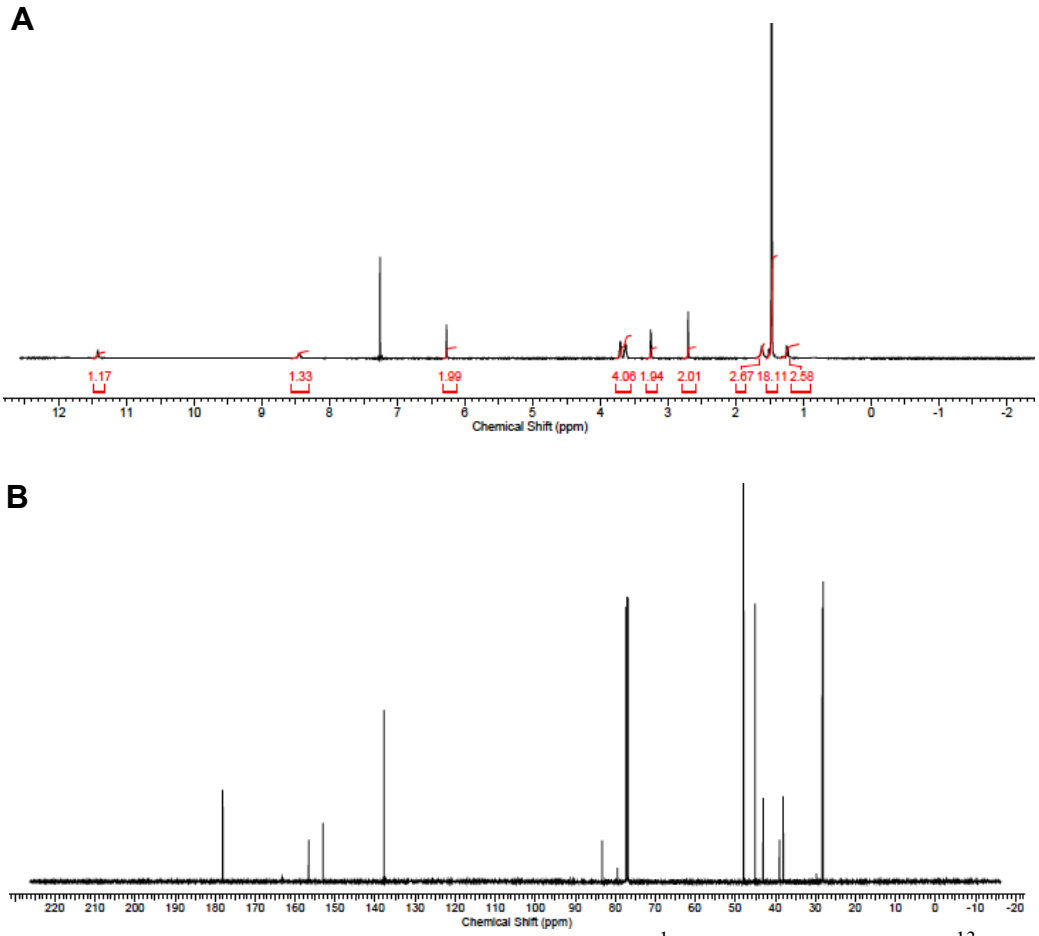

Figure S3. NMR spectra of guanidinium monomer. A) ^{1}H NMR spectrum; B) ^{13}C NMR spectrum.

Ib. Polymerization of functional monomers and controls

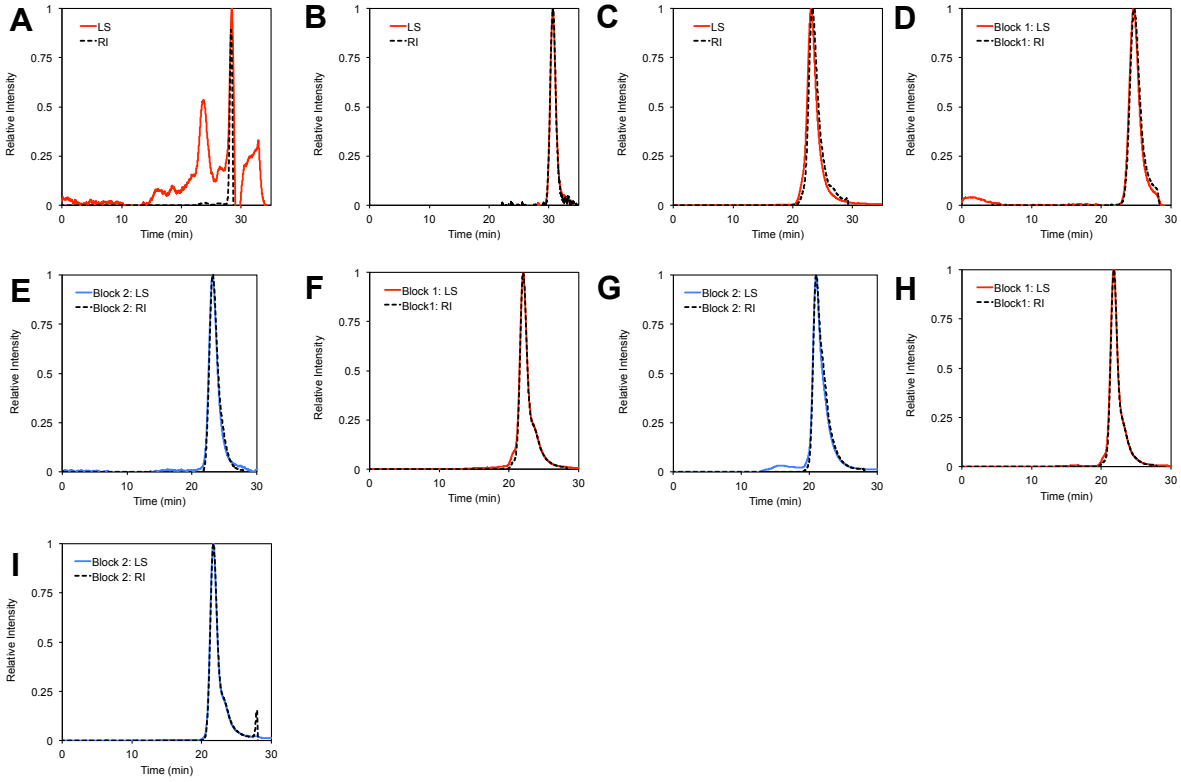

Figure S4. Characterization of polymers by SEC-MALS. Chromatograms for SLS (solid colored line) and refractive index (RI, black dotted line) are given for A) OEG polymer; B) GSGSG polymer; C) Arg8 polymer; D) OEG-*b*-Guanidinium polymer, first "m" block (OEG); E) OEG-*b*-Guanidinium polymer, second "n" block (guanidinium); F) Phenyl-*b*-GSGSG polymer, first "m" block (Phenyl); G) Phenyl-*b*-GSGSG polymer, second "n" block (GSGSG); H) Phenyl-*b*-Tat polymer, first "m" block (Phenyl); I) Phenyl-*b*-Tat polymer, second "n" block (Tat). Red traces represent homopolymers or the first block of a block copolymer while the second block of block copolymers are shown in blue. SLS data is only tabulated for peaks that have a corresponding RI signal. Note that chromatogram A appears noisy in the SLS component because the polymer runs coincident with the solvent front due to its low molecular weight (~3500).

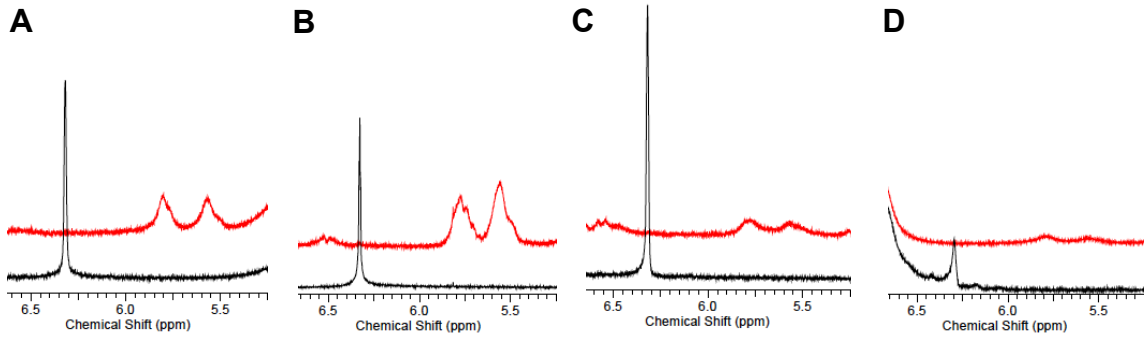

Figure S5. ^{1}H NMR spectra of polymerizations of monomers that have not been reported previously. Black spectrum is taken prior to addition of the initiator. Resonance at δ 6.32 ppm corresponds to the norbornene olefin protons of the monomer. The red spectrum is recorded at the end of the polymerization (t = 3 hr) and verifies complete consumption of the monomer (no resonance at δ 6.32 ppm). The new resonance at ~ δ 5.5-6 ppm corresponds to the cis-trans olefin protons of the polymerized material. A) NorAha-GSGSG polymerization; B) NorGuanidinium polymerization; C) NorAha-Tat polymerization; D) NorAha-Arg8 polymerization.

II. DETERMINATION OF CELLULAR UPTAKE OF MATERIALS

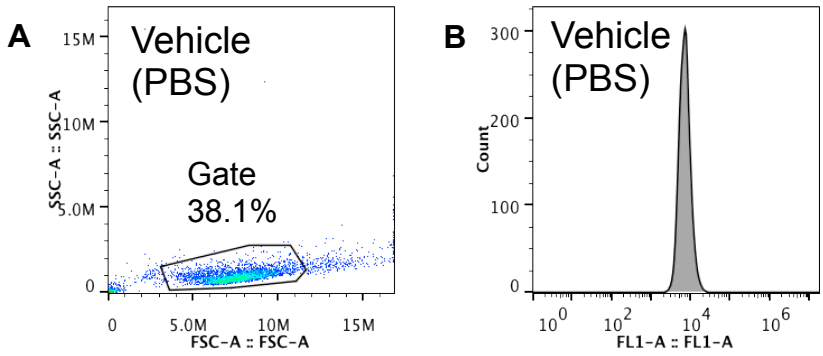

Figure S6. Representative flow cytometry data for the vehicle control used in cellular uptake experiments. Each experiment was performed in triplicate on separate cultures of HeLa cells, and data from one of these recordings is given. A) Plot of side and forward scattering showing the region gated for each flow cytometry experiment. The percentage given is the percentage of cells that fall within the gated region (10,000 events total). B) Histogram showing fluorescence intensity of the gated region for the vehicle control. Y-axis is number of cells and x-axis is fluorescence.

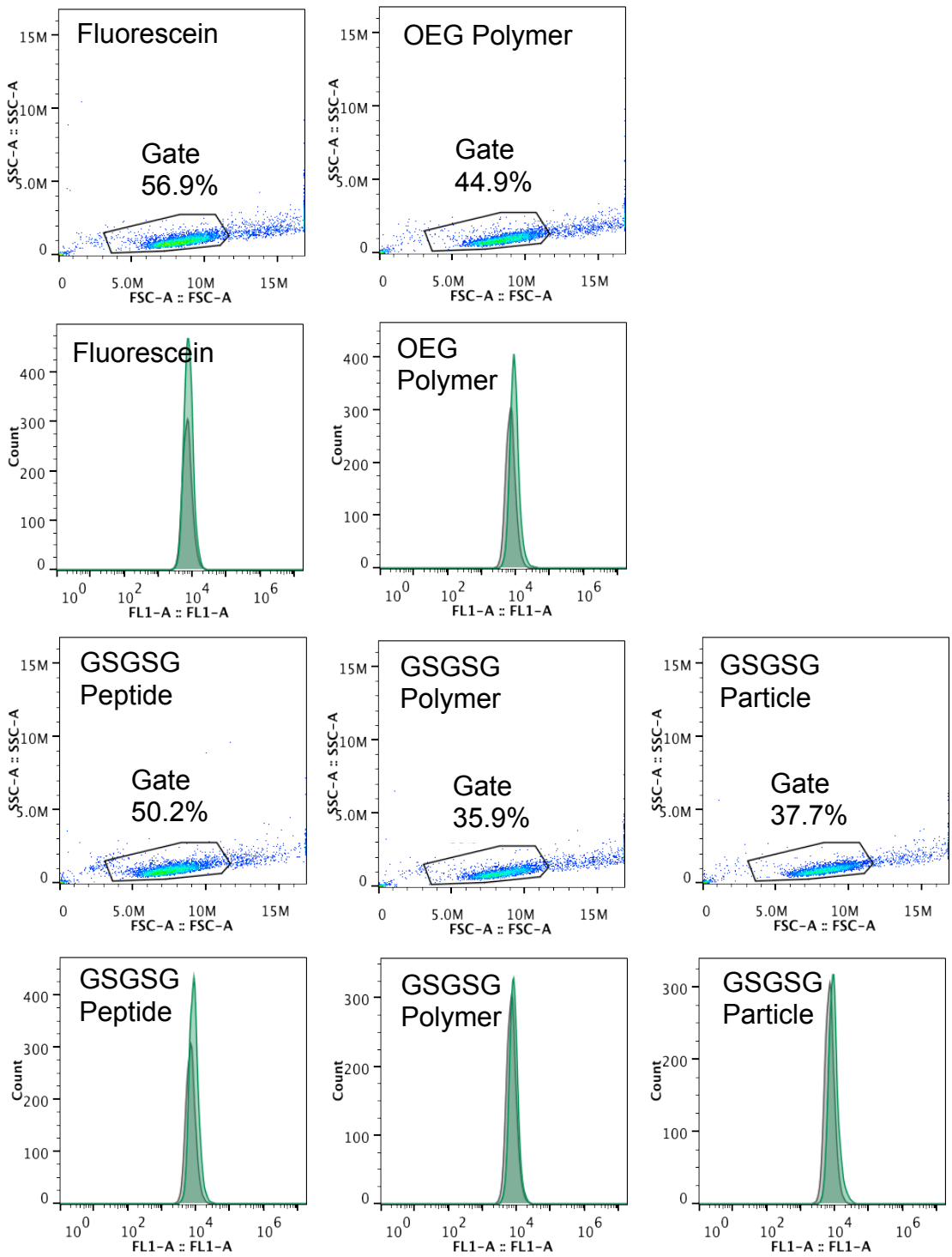

Figure S7. Representative flow cytometry data for all control peptides, polymers, particles, and small molecules. All recordings were gated to the vehicle control (PBS) given in Figure S6. The percentage given is the percentage of cells within the gated region for each sample (10,000 events total). The histograms display any shifts in fluorescence of the material (in green) relative to the vehicle control (in grey). Y-axis is number of cells and x-axis is fluorescence.

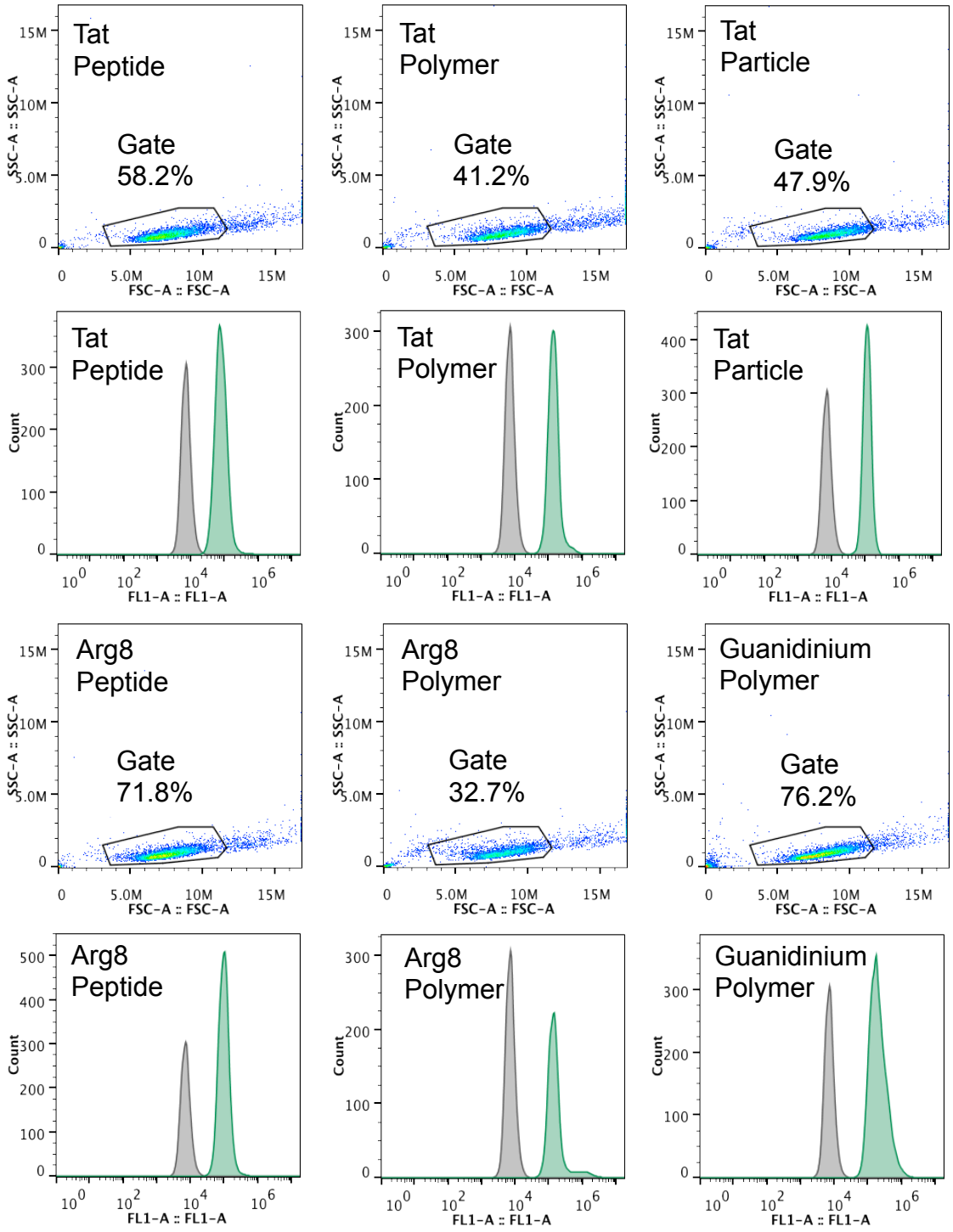

Figure S8. Representative flow cytometry data for cell penetrating peptides, polymers and particles. All recordings were gated to the vehicle control (PBS) given in Figure S6. The percentage given is the percentage of cells within the gated region for each sample (10,000 total events). The histograms display shifts in fluorescence of the material (in green) relative to the vehicle control (in grey). Y-axis is number of cells and x-axis is fluorescence.

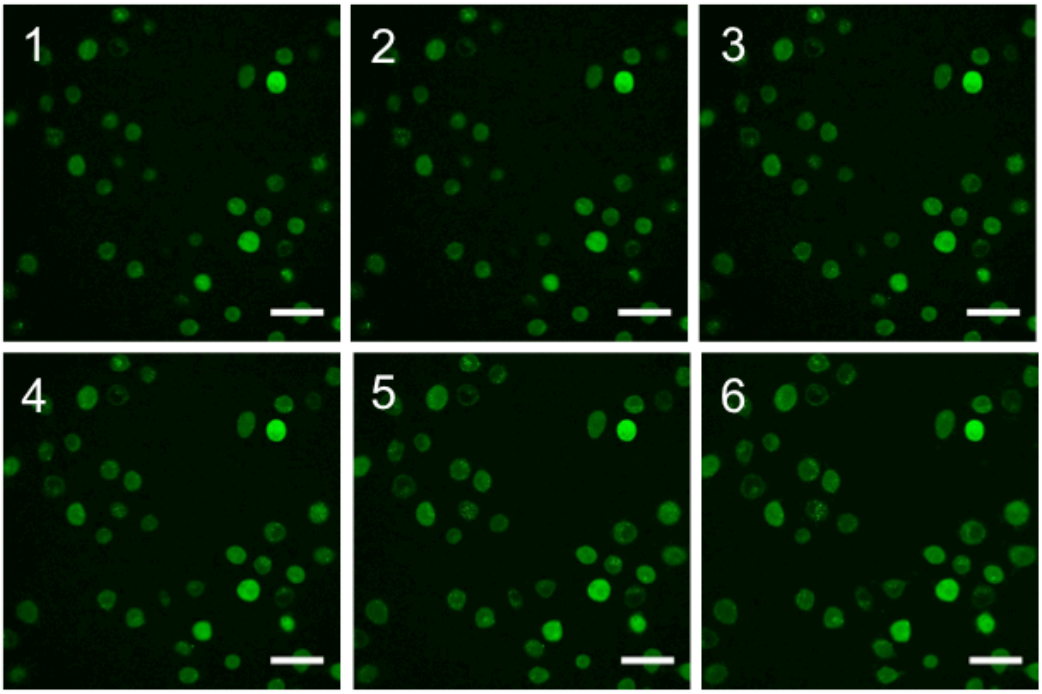

Figure S9. The six consecutive Z-stack slices that were averaged together to yield the averaged image of the Tat peptide in Figure 5 of the main text. Scale bar is 50 μm. Slices are acquired every 1 μm. Punctate and diffuse fluorescence are seen in each slice, suggesting that the Tat peptide has permeated the cell and does not just reside at the cell membrane.

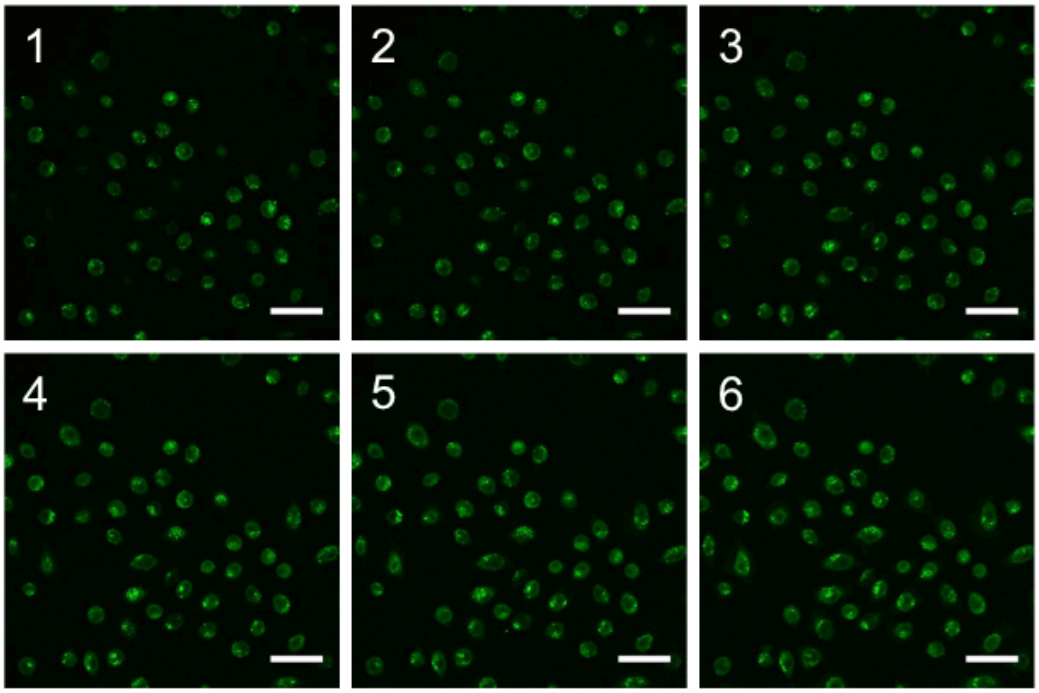

Figure S10. The six consecutive Z-stack slices that were averaged together to yield the averaged image of the Tat polymer in Figure 5 of the main text. Scale bar is 50 μm. Slices are acquired every 1 μm. Punctate and diffuse fluorescence are seen in each slice, suggesting that the Tat polymer has permeated the cell and does not just reside at the cell membrane.

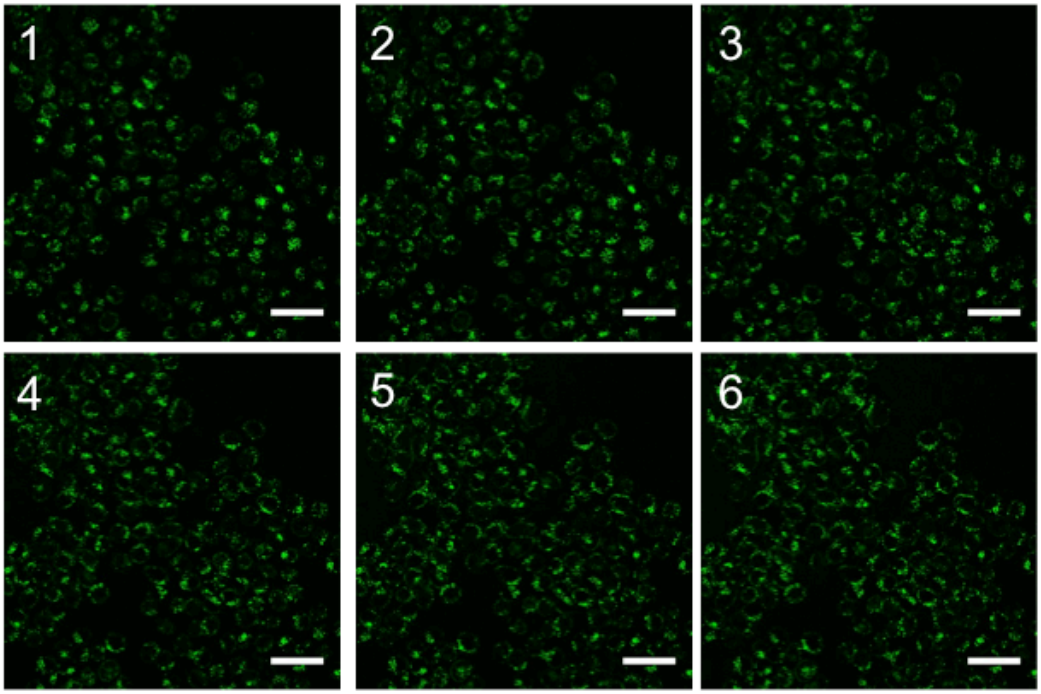

Figure S11. The six consecutive Z-stack slices that were averaged together to yield the averaged image of the Tat particle in Figure 5 of the main text. Scale bar is 50 μm. Slices are acquired every 1 μm. Fluorescence is seen in each slice, suggesting that the Tat particle has permeated the cell and does not just reside at the cell membrane. We also note that fluorescence appears more punctate here than for the Tat peptide or polymer.

III. CELL VIABILITY ASSAYS OF TAT AND GSGSG MATERIALS

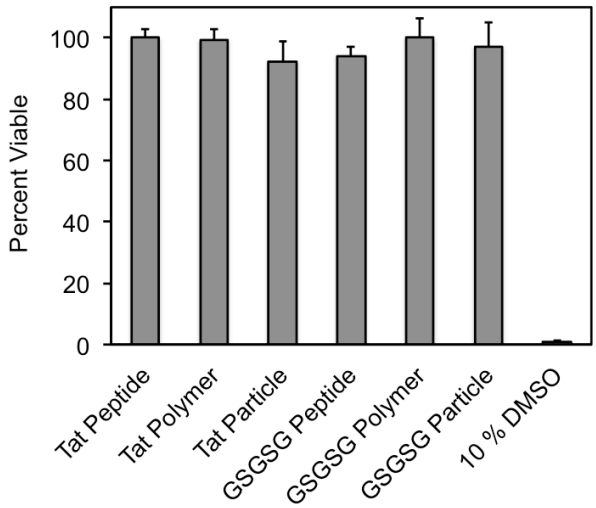

Figure S12. Cell cytotoxicity studies using the CellTiter Blue assay. Percentages are relative to vehicle control. Cells were treated for 48 hrs with the Tat and GSGSG peptide, polymer and particle at 5 μM. The positive control is 10 % DMSO. Cells treated with any peptide-based material remained at least 92 % viable.

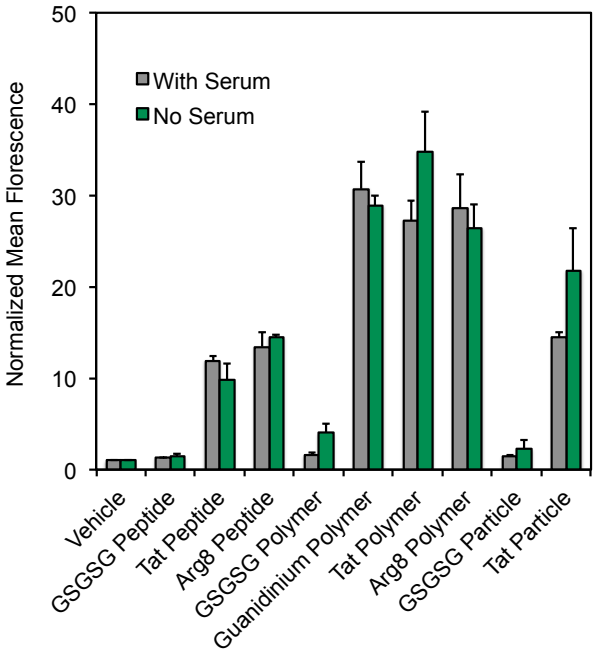

Figure S13. Flow cytometry data for materials incubated with and without fetal bovine serum. No significant differences in values were obtained, indicating that serum components do not affect cellular uptake of the materials.

IV. CONCENTRATION DETERMINATION OF CELL PENETRATING MATERIALS AND CONTROLS

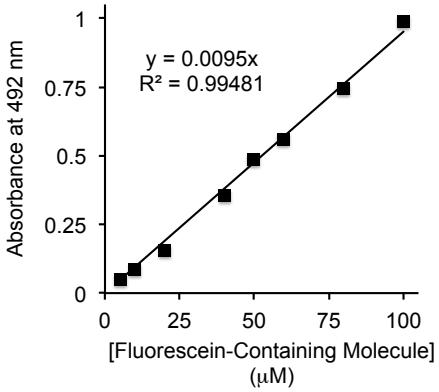

Figure S14. Standard curve used to determine concentration of fluorescein-containing materials. Data was recorded and fit for the small molecule 5/6-carboxyfluorescein and confirmed for use with the fluorescein-Tat polymers and particles. Absorbance is at 492 nm.

V. PROTEOLYSIS STUDIES OF TAT PEPTIDE, POLYMER AND PARTICLE

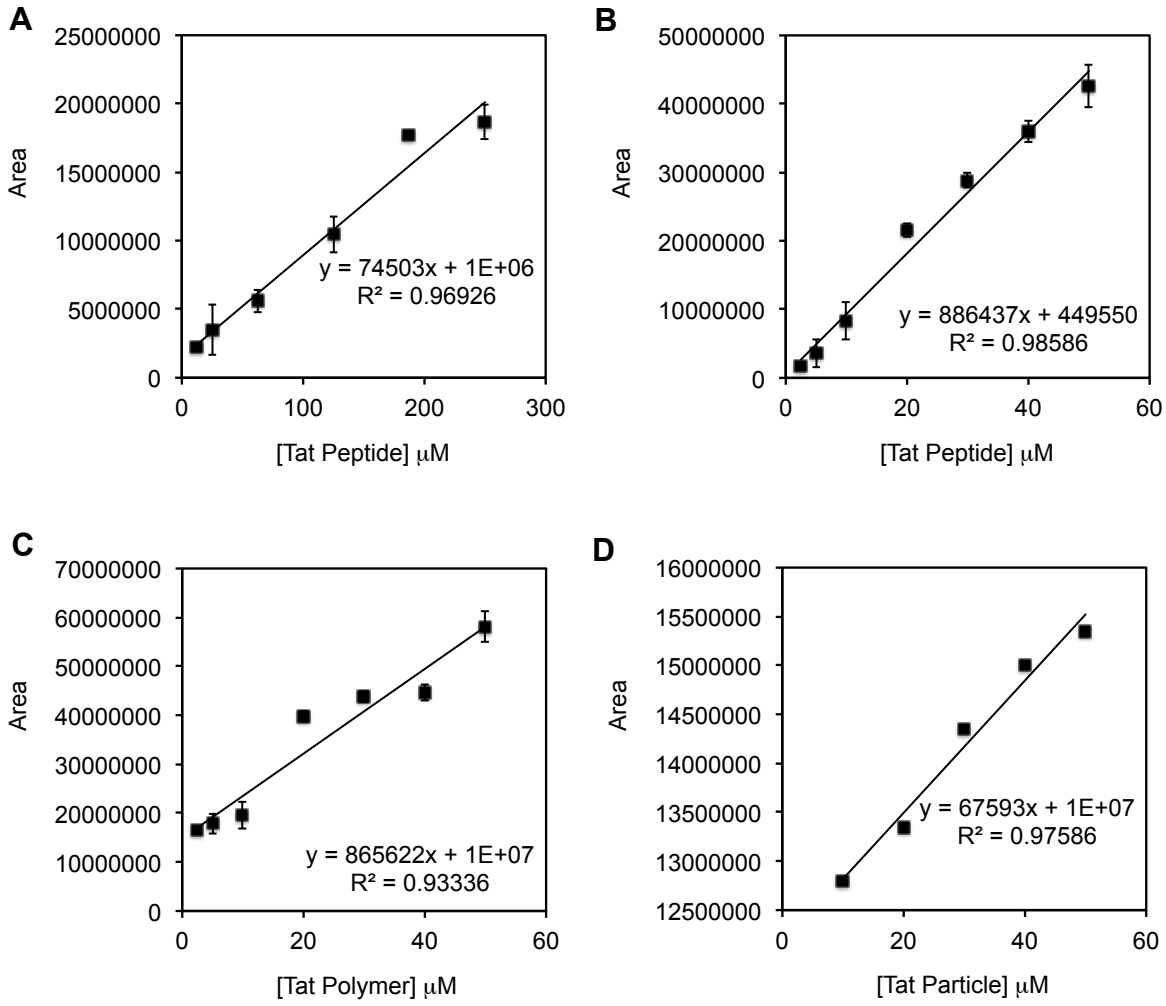

Figure S15. Standard curves for peak area of untreated A) Tat peptide- 99 μL injection; B) Tat peptide- 20 μL injection; C) Tat polymer- 99 μL injection; D) Tat particle- 15 μL injection.

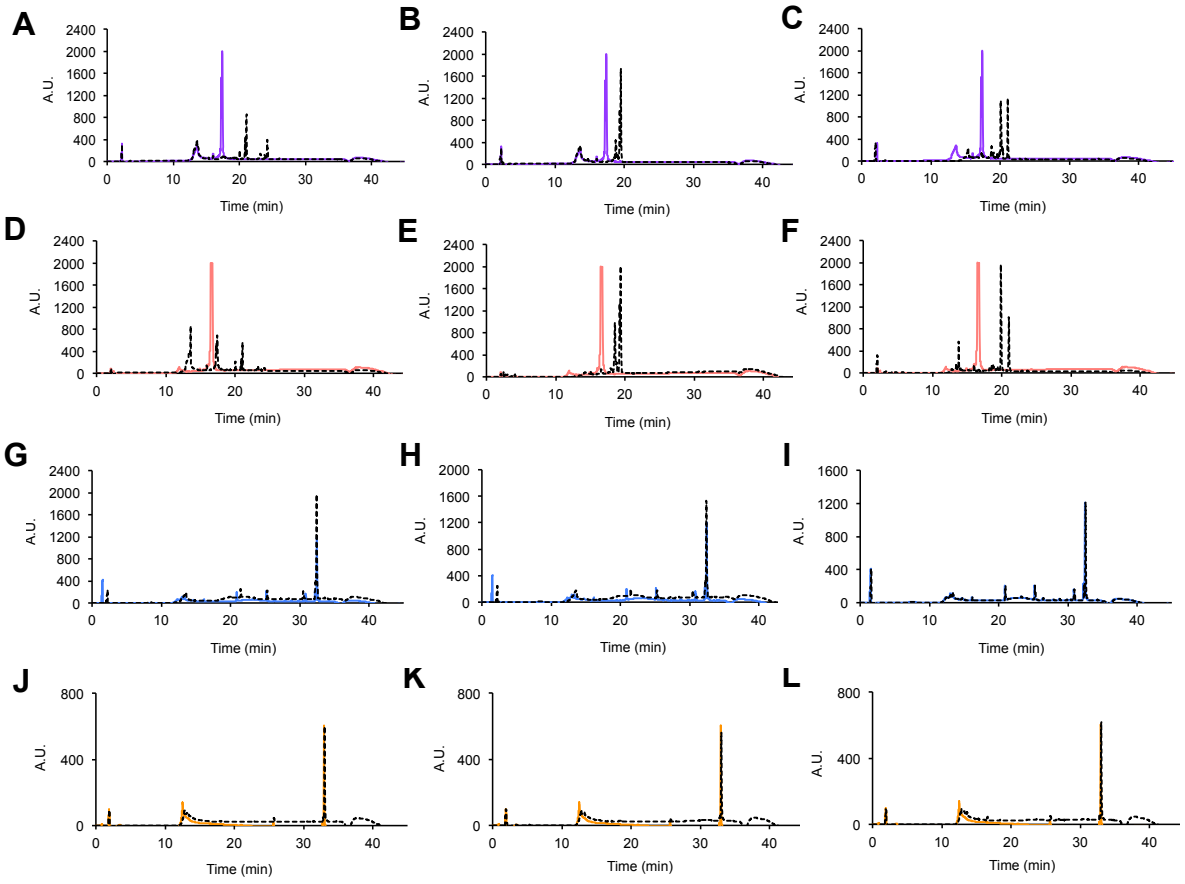

Figure S16. HPLC chromatograms of Tat-containing materials before (solid colored line) and after (dotted black line) treatment with 1 μM protease. A) Tat peptide at 2.5 μM with chymotrypsin; B) Tat peptide at 2.5 μM with trypsin; C) Tat peptide at 2.5 μM with pronase; D) Tat peptide at 12.5 μM with chymotrypsin; E) Tat peptide at 12.5 μM with trypsin; F) Tat peptide at 12.5 μM with pronase; G) Tat polymer at 2.5 μM with chymotrypsin; H) Tat polymer at 2.5 μM with trypsin; I) Tat polymer at 2.5 μM with pronase; J) Tat particle at 2.5 μM with chymotrypsin; K) Tat particle at 2.5 μM with trypsin; L) Tat particle at 2.5 μM with pronase.

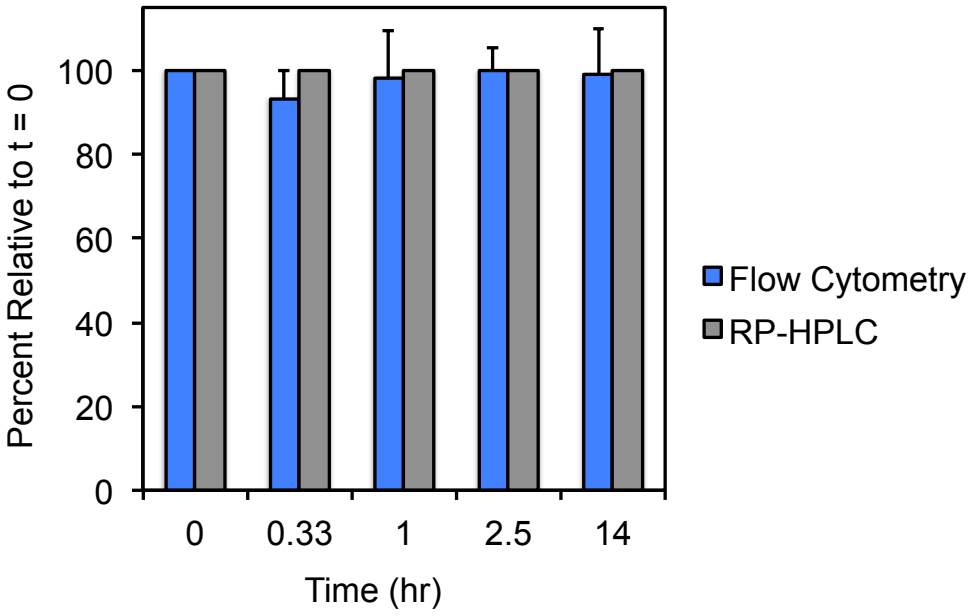

Figure S17. Time course of the treatment of the Tat polymer (2.5 μM) with chymotrypsin (1 μM). Note that the peptide is cleaved to less than 10% of the starting concentration of material after 20 min of incubation with chymotrypsin (See Figure 7 in main text). The functional impact of proteolytic treatment is measured by flow cytometry while the percentage of intact peptide remaining after incubation was assayed by RP-HPLC. Note that all values are the percentage remaining after incubation relative to the value seen at t = 0. Therefore a value of 100% for either measurement is indicative of no cleavage or no functional effect.

VI. TESTING THE GENERALITY OF THE APPROACH BY ASSESSING PROTEOLYSIS OF A THROMBIN SUBSTRATE

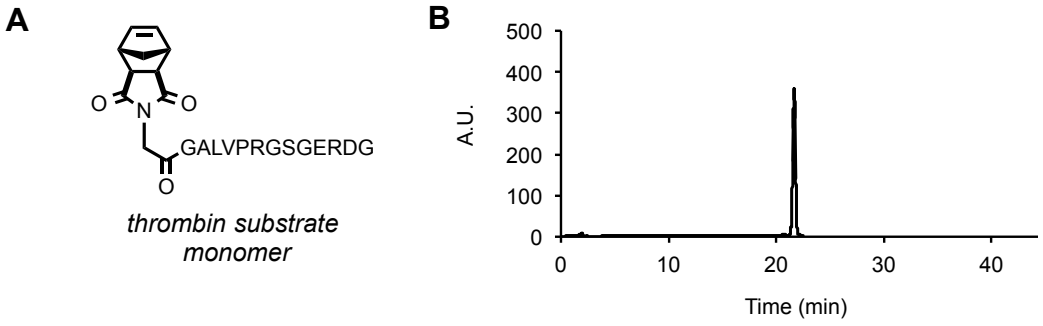

Figure S18. Chemical structure and characterization of thrombin substrate monomer. A) Chemical structure of thrombin substrate monomer. B) RP-HPLC chromatogram of purified monomer at gradient 15-30 % B. ESI-MS confirms identity- calculated: 1473.5 *m/z*, obtained: 1471.6 *m/z*.

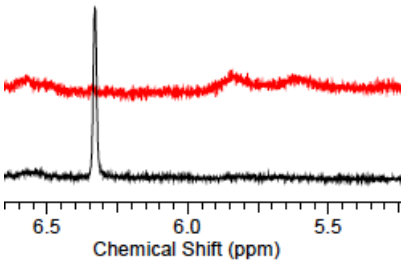

Figure S19. Example of a successful polymerization of thrombin substrate monomer as determined by ^{1}H NMR. The monomer at t = 0 is given in black and the polymerization reaction at t = 3 hr is in red.

Polymer	M_n	DP
Thrombin Homopolymer (10 mer)	16,800	11 (10)
Thrombin Homopolymer (20 mer)	33,300	23 (20)
Thrombin Homopolymer (30 mer)	43,900	30 (30)

Table S2. Characterization data for each homopolymer containing the thrombin peptide substrate as determined by batch mode SLS in a cuvette.

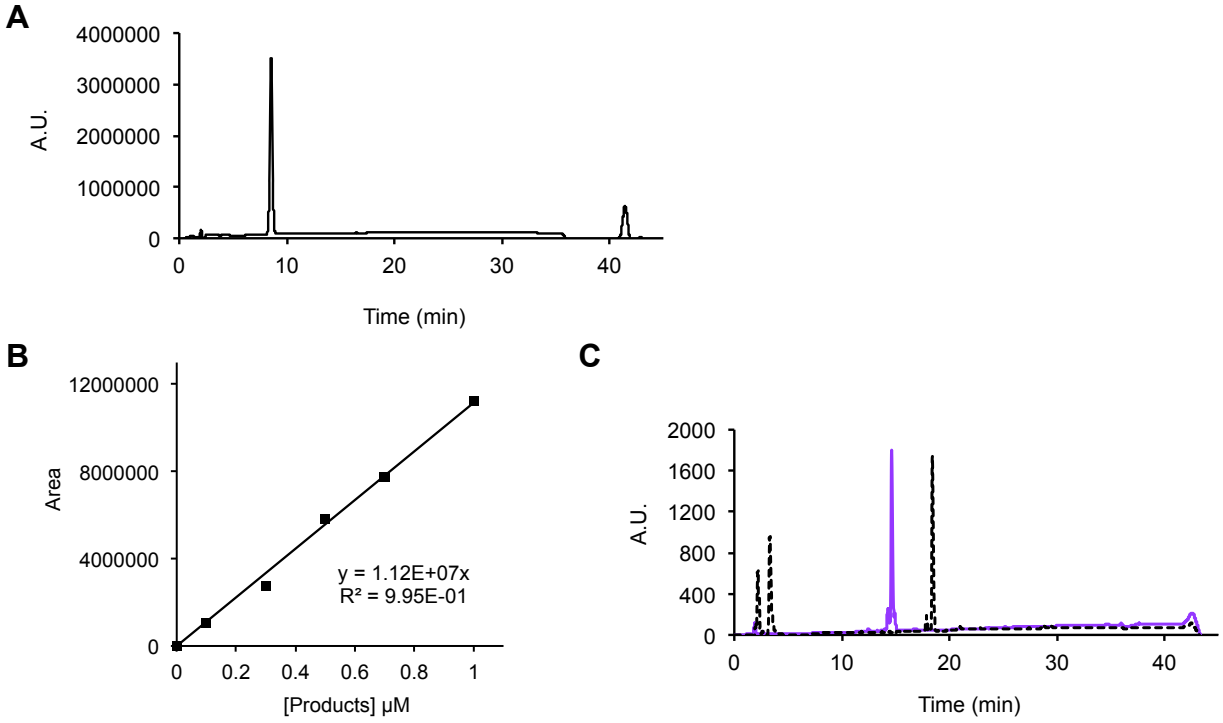

Figure S20. Characterization of the cleavage of the thrombin substrate monomer and polymer. A) RP-HPLC chromatogram of authentic product resulting from cleavage reaction (GSGERDG-NH_2), gradient is 0 – 20% Buffer B. ESI-MS confirms identity- calculated: 675.7 *m/z*, obtained: 674.4 *m/z*. B) Standard curve used for determining the concentration of products, which compares the concentration of the authentic product to the corresponding peak area in RP-HPLC chromatograms with gradient 0 – 80% Buffer B. C) An example proteolytic digestion of the monomer at t =12 hr. The chromatogram of the polymer at time t = 0 min is purple and the reaction at t = 12 hr is black with gradient at 0 – 80% Buffer B.

VII. TESTING THE GENERALITY OF THE APPROACH WITH A FLUOROGENIC SUBSTRATE FOR MT1-MMP

The design of the fluorogenic substrate (given in Figure S21) is as follows: EDANS, the fluorescent donor, was incorporated as a conjugate to a Glu side chain at the *N*-terminus of the substrate sequence, while a fluorescence quencher, DABCYL was conjugated to the ε-amino group of a lysine residue toward the *C*-terminal end of the peptide substrate. The EDANS/DABCYL pair is an established donor/quencher pair with appropriate spectral overlap and has been used in a variety of biochemical assays. An additional five amino acid sequence, GSGSG, was added to the *C*-terminal end of the peptide to render the resulting polymers soluble in aqueous media. Upon cleavage of the peptide substrate, the *C*-terminal quencher is separated from the *N*-terminal donor, and a fluorescent signal can be detected (Figure S25-S27). In this case, the theoretical *C*-terminal cleavage products for both the monomer and polymer substrates are identical.

To ensure that our kinetic data, shown in Figure 8d of the main text, were not biased by incomplete digestion of the homopolymer, which could lead to persistent quenching of the EDANS donor by any remaining DABCYL that was not cleaved from the polymer, we designed and polymerized a monomer in which the locations of the fluorescent donor and quencher are flipped (Figure S21). In this scenario, the donor is liberated from the polymer upon enzymatic digestion and would no longer be susceptible to quenching by any DABCYL remaining on the polymer. Again, very little cleavage of the peptide in the homopolymer was observed in our fluorescence assays (Figure S30). Finally, we prepared and polymerized a monomer where the location of the GSGSG sequence was moved from the *C*-terminus of the peptide to the *N*-terminus (the point of attachment to the norbornene units) in order to further distance the peptide substrate from the polymer backbone. Again, little cleavage was observed for homopolymers bearing this peptide (Figure S29). These data establish that the fluorogenic substrate cannot be degraded appreciably by proteases as a homopolymer.

VIIA. Preparation of fluorogenic monomers

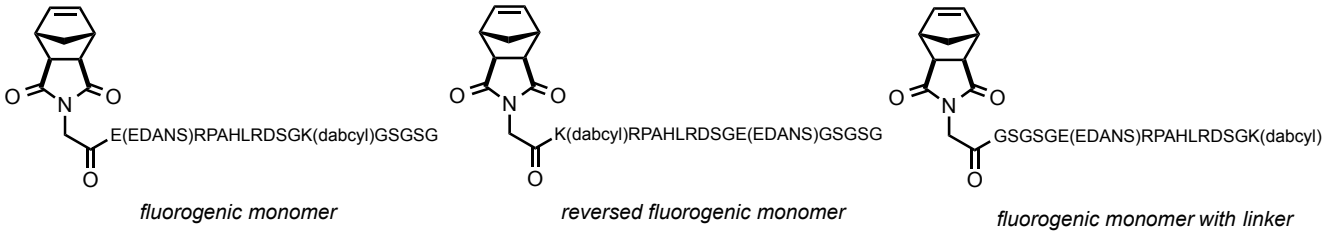

Figure S21. Chemical structures of fluorogenic monomers prepared.

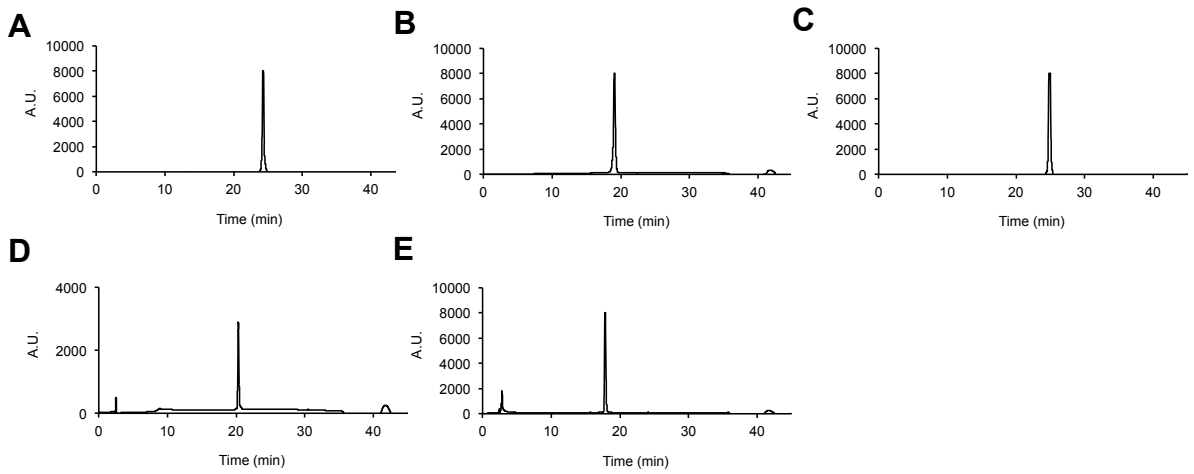

Figure S22. RP-HPLC chromatograms of fluorogenic monomers and authentic products. A) NorGly-E(EDANS)RPAHLRDSGK(dabcyl)GSGSG-NH_2 at 0-67% B. B) NorGly-K(dabcyl)RPAHLRDSGE (EDANS)GSGSG-NH_2, gradient at 0-67 % B. C) NorGly-GSGSGE(EDANS)RPAHLRDSGK(dabcyl)-NH_2, gradient at 10-67% B. D) The C-terminal fragment H-RPAHLRDSGK(dabcyl)GSGSG-NH_2, gradient at 0-67% B. E) The N-terminal fragment NorGly-E(EDANS)RPAH-H, gradient at 10-30 %B. Other enzymes used in this study cleave the substrate at different locations, but for simplicity, D and E are used to generate standard curves for all kinetic studies.

Peptide	Sequence	Mass Calculated	Mass Obtained
Fluorogenic Monomer	NorGly-E(EDANS)RPAHLRDSGK(dabcyl)GSGSG-NH_2	2311.03	2312.9
Reversed Fluorogenic Monomer	NorGly-K(dabcyl)RPAHLRDSGE(EDANS)GSGSG-NH_2	2311.03	2312.8
Fluorogenic Monomer with Linker	NorGly-GSGSGE(EDANS)RPAHLRDSGK(dabcyl)-NH_2	2311.03	2313.7
N-terminal Fragment	NorGly-E(EDANS)RPAH-H	1059.42	1060.6
C-terminal Fragment	H-LRDSGK(dabcyl)GSGSG-NH_2	1269.5	1270.6

Table S3. Amino acid sequences and molecular weights calculated and obtained for fluorogenic monomers and authentic products. Molecular weights were obtained by ESI (+) MS.

VIIb. Polymerization of fluorogenic monomers

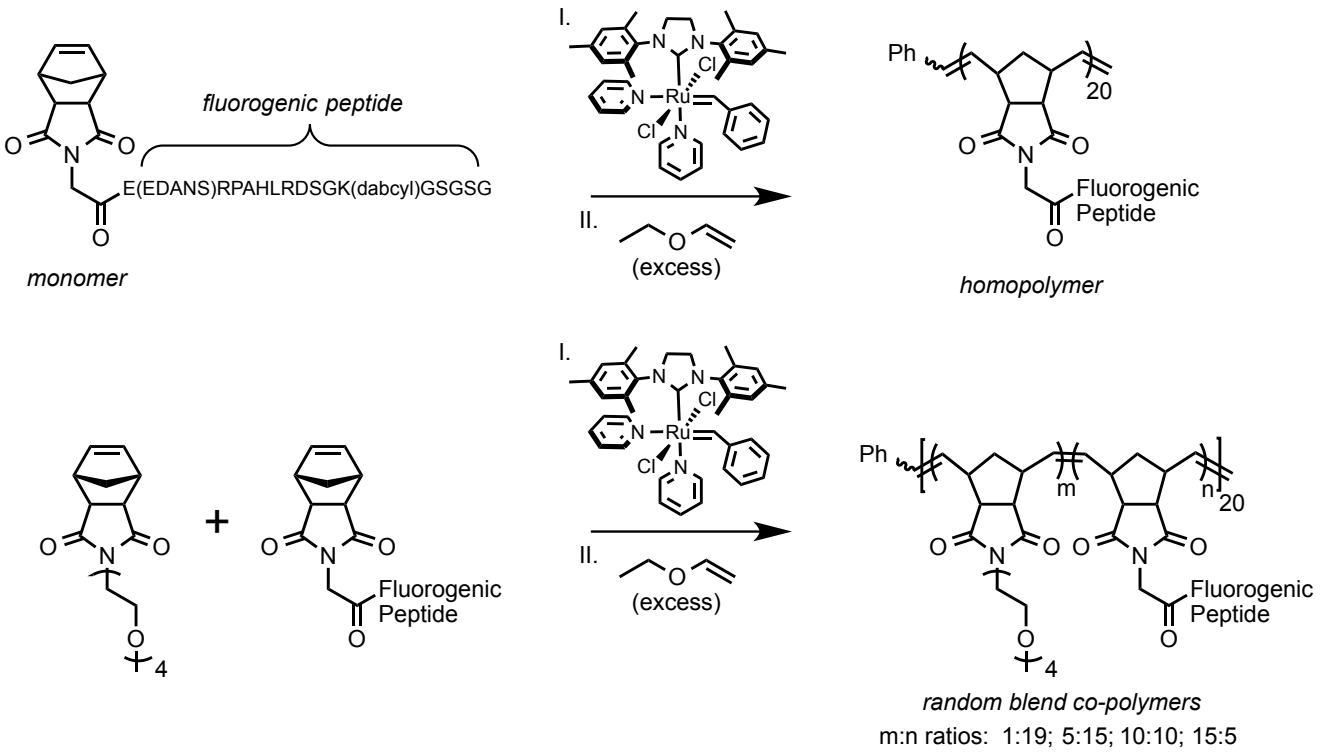

Figure S23. Polymerization schemes of fluorogenic monomer into a A) homopolymer and B) random blend co-polymer with an OEG monomer.

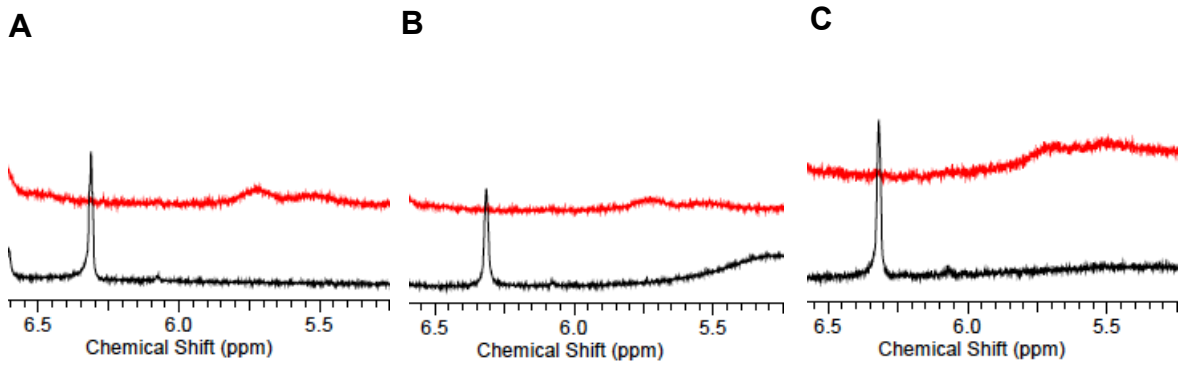

Figure S24. Monitoring of polymerization by ^{1}H NMR of A) NorGly-E(EDANS)RPAHLRDSGK (dabcyl)GSGSG-NH_2; B) NorGly-K(dabcyl)RPAHLRDSGE(EDANS)GSGSG-NH_2; C) NorGly-GSGSGE(EDANS)RPAHLRDSGK(dabcyl)-NH_2. Note the absence of the olefenic proton from the monomer (seen at ~ δ 6.32 ppm in the black spectrum of the monomer) after 3 hrs of polymerization (red spectrum) and the corresponding appearance of two broad peaks from ~ δ 5.4-6 ppm signifying the cis- and trans- olefin protons of the norbornyl polymer backbone.

Polymer	Theoretical M_n	M_n	DP
Fluorogenic Homopolymer	46,000	43,000	19
Reversed Fluorogenic Homopolymer	46,000	43,000	19
Fluorogenic Homopolymer with Linker	46,000	45,000	20

Table S4. Theoretical and experimentally determined M_n values for fluorogenic homopolymers. M_n is obtained by batch mode SLS with a cuvette.

VIIc. Fluorescence assays for quantifying cleavage of fluorogenic substrates

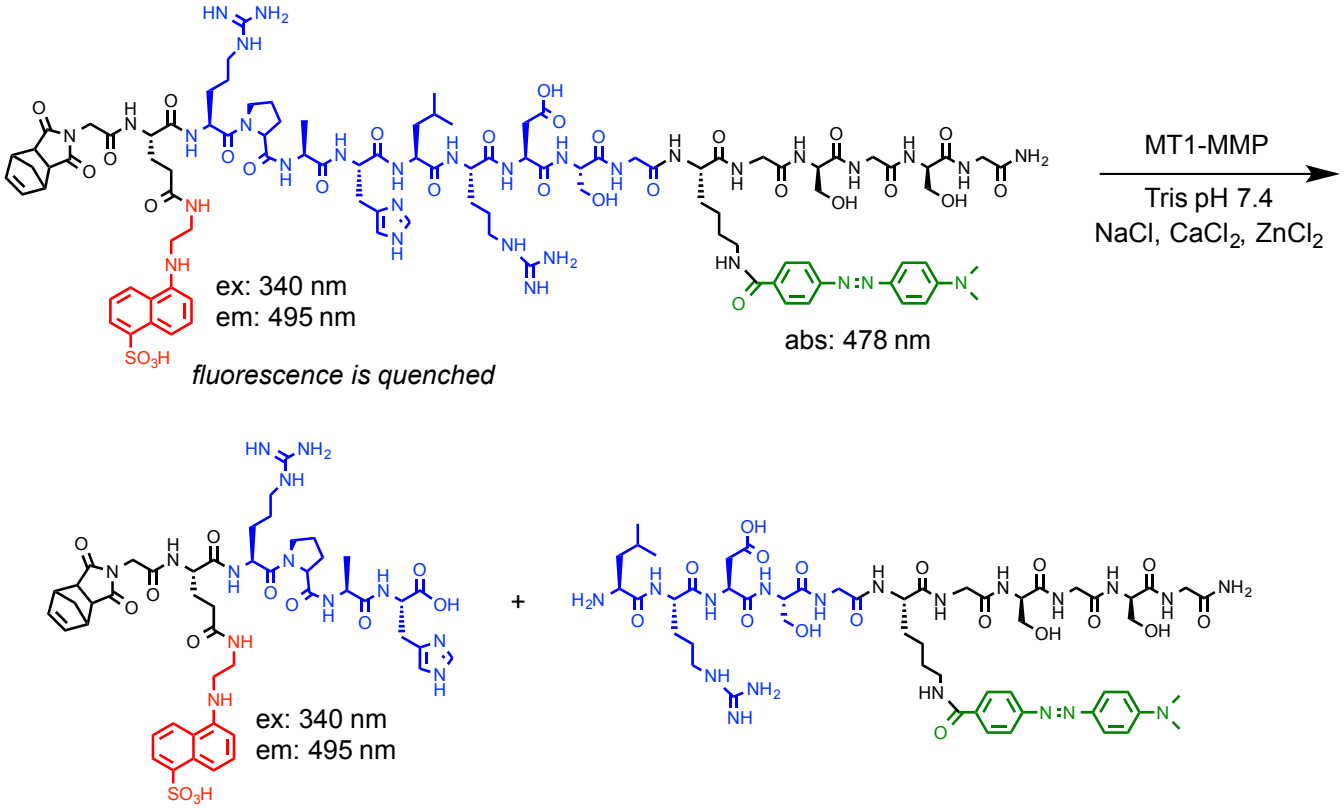

Figure S25. Fluorescent assay to detect cleavage of the fluorogenic substrate by MT1-MMP. In the intact monomer (NorGly-E(EDANS)RPAHLRDSGK(dabcyl)GSGSG-NH_2), the fluorescence of the donor, EDANS (red), is quenched by the acceptor, dabcyl (green) until cleavage of the substrate (blue) by the protease. Note that other proteases used in this study will cleave at other locations along the peptide substrate, but scission at any amide bond of the substrate sequence will result in liberation of the quencher and the onset of fluorescence.

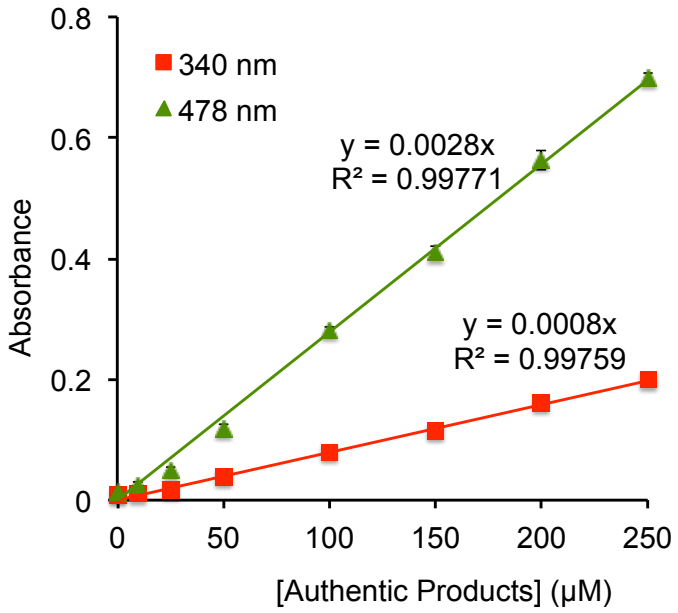

Figure S26. Determination of concentration of fluorogenic materials in water at 340 nm (EDANS absorbance) and at 378 (dabcyl absorbance). The concentrations used in proteolysis studies is the average concentration obtained at the two wavelengths.

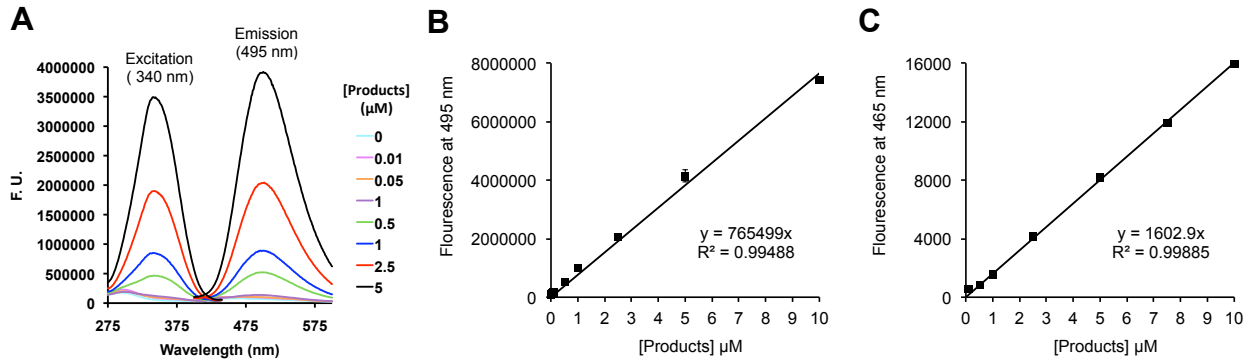

Figure S27. Standard curves for assessing degree of proteolytic digestion using authentic products described in Figure S22D-E and Table S3. Note that only example standard curves are given because new standard curves were generated each day at an appropriate concentration range for the given experiment. A) Excitation and emission spectra for authentic products on a fluorimeter. Max excitation and emission were determined to be 340 nm and 495 nm, respectively. B) Example standard curve on a fluorimeter (used for monitoring cleavage kinetics of homopolymers and all cleavage reactions by MMP-9 due to higher sensitivity than the plate reader). C) Example standard curve on a plate reader (used for assessing cleavage kinetics of monomers for enzymes other than MMP-9).

VIId. Additional kinetic data on fluorogenic substrate proteolysis

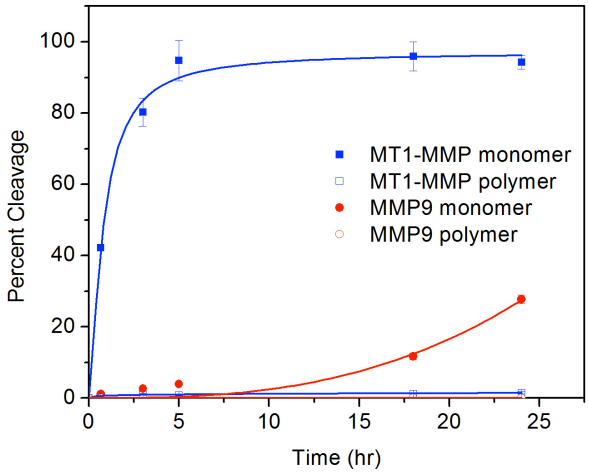

Figure S28. Time course of the proteolysis of the fluorogenic monomer and homopolymer by MT1-MMP and MMP-9 (for which the amino acid sequence should not be a substrate). While the monomer is cleaved readily by MT1-MMP, MMP-9 shows little activity. Neither enzyme was able to proteolytically digest the homopolymer.

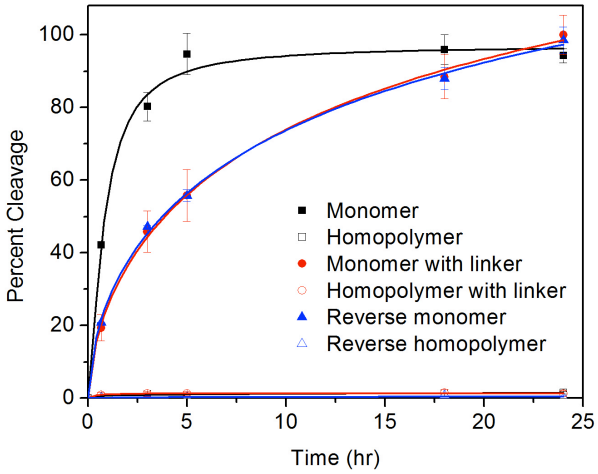

Figure S29. Time course of proteolysis of fluorogenic homopolymers and their monomers building blocks (40 μM) by MT1-MMP (25 nM). Chemical structures of the materials are given in Figure S21. Little proteolysis of the homopolymers was observed regardless of the location of the quencher or donor or whether a five-amino acid linker (GSGSG) was used to space the substrate from the polymer backbone.

VIIe. Preparation and characterization of random blend copolymers of the fluorogenic substrate

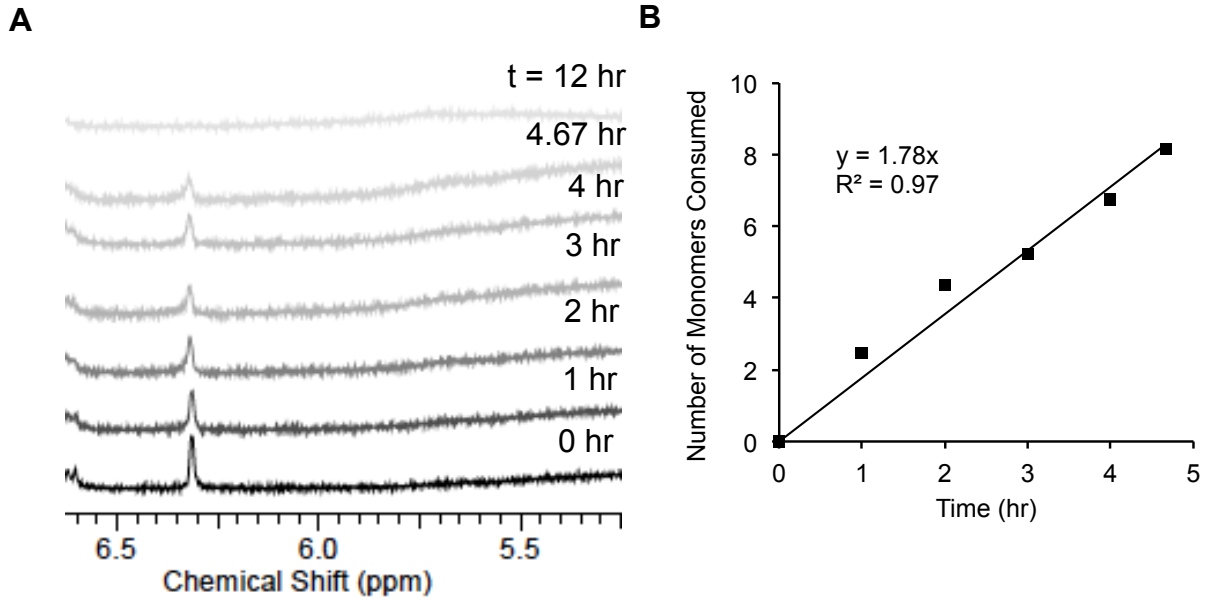

Figure S30. Strategy for preparing random blend copolymers of the fluorogenic substrate. A) Time course of the fluorogenic substrate monomer as monitored by ^{1}H NMR. At each time point, integration of the olefinic proton of the monomer δ 6.32 is tabulated and normalized relative to a time point at which the reaction is complete (final spectrum; t = 12 hr). This value is converted into the number of monomers consumed by assuming that the integration value at t = 12 hr represents the consumption of 20 monomers and the value at t = 0 hr represents that of no monomers consumed. B) Plot of the number of monomers consumed as calculated in A vs. time. The slope of the line is the rate of monomers consumed per hr.

OEG: Peptide Ratio (m:n)	Peptide Volume (µL)	OEG Solution Volume (µL)	Initiator Solution Volume (µL)	Total Reaction Volume (µL)	Calculated Reaction Time (hr)	Rate of OEG Addition (µL/hr)	Theoretical M_n	M_n
0:20	368.4	0	20	388.4	11.2	-	46,200	43,000
5:15	276.3	92.1	20	388.4	8.4	10.9	36,500	36,000
10:10	184.2	184.2	20	388.4	5.6	32.8	26,700	26,000
15:5	92.1	276.3	20	388.4	2.8	98.4	16,900	16,300
19:1	18.4	350	20	388.4	0.56	-	8,670	8,700

Table S5. Tabulated values for the preparation of reasonably interdigitated random copolymers. Given that the OEG monomer is fast to polymerize (at the concentration used, the OEG monomer polymerizes to DP = 20 in less than 15 m), this monomer is doped into a stirring solution of the substrate monomer and the initiator via a syringe pump in a glove box. For the 19:1 blend copolymer, the OEG monomer was added in one portion to the reaction vessel after ~2 hr. The initial concentration of the OEG monomer, substrate monomer and catalyst were 0.012 M, 0.012 M and 0.6 mM, respectively. The bulk M_n obtained for the homopolymer was 43,000 (DP = 19). Given that the monomers and initiators were taken from the same pots, it is assumed that the ratios of m:n are as indicated. Nevertheless, M_n values obtained by batch mode SLS are listed and are in good agreement with the theoretical values calculated form the intended m:n ratios.

VIII. MOLECULAR DYNAMICS SIMULATIONS

In order to gain insight into the structures and behavior of these novel molecular systems, we used explicit solvent, all-atom molecular dynamics (MD) to simulate peptide-norbornene constructs representative of those studied experimentally.

As shown in Figure 11 of the main text, polymers of the following structures were simulated:

Simulated polymer **1**: The homopolymer with ten fluorogenic peptides (as shown in main text, Figure 11A): (Monomer = NorG-E(EDANS)RPAHLRDSGK(dabcyl)GSGSG), MW=23.2 kDa, N=200

Simulated polymer **2**: random blend polymer, where five norbornene moieties are linked to the fluorogenic peptide and the other five are linked to OEG-4, MW=13.4 kDa, N=125 (Figure 11D)

Simulated polymer **3**: blend polymer, where nine of the norbornene moieties are linked to OEG-4 and the tenth, at position 10 of the norbornene chain, is linked to the fluorogenic peptide, M= 5.60 kDa, N=65 (Figure 11B)

Simulated polymer **4**: same as **3** except that the peptide is linked to the fifth norbornene of the chain, MW=5.60 kDa, N=65 (Figure 11C)

Simulated polymer **5**: same as **3** except without the dye components, MW=5.11 kDa, N=63

Simulated polymer **6**: same as **4** except without the dye components, MW=5.11 kDa, N=63

Here, MW is molecular weight (Da); and N, the number of constituent monomers, is computed as the number of norbornene monomers (10) plus the number of amino acid residues (number of peptides × 17) plus the number of ethylene oxide units (number of OEG chains × 4) plus the number of dye moieties (either 2 or 0). Each simulation began with the molecule in an artificial extended conformation, with a straight norbornene chain and with straight peptides and OEGs arrayed at right angles to the norbornene chain. After an initial 20 ns equilibration phase at 300K, each simulation was split and continued in two ways. One simulation was simply continued for another 100 ns at 300K. The other was randomized further by briefly heating it to 500K, cooling it back to 300K, and then continuing it at 300 K for the rest of the 100 ns production simulation.

In every simulation, the initial extended structure collapsed quickly into a much more compact conformational ensemble. The collapse is evident from Figure S31, which shows rapid early drops in the radii of gyration, R_g. The values of R_g then changed little for the runs that continued at 300K (blue and green). For simulated polymers **1** and **2**, which contain ten and five peptides, respectively, the heat-cool cycle (red and orange) generated a spike in R_g, followed by a further reduction in R_g below the value seen for the standard 300K runs. Thus, for these two molecules, the heat-cool cycle allows the polymer construct to anneal into an even more compact conformation. For simulated polymers **3-6**, which have one peptide and nine PEGs, the heat-cool cycle changed R_g but did not consistently lead to a lower value. Thus, the heat-cool cycle did not seem to lead to greater compaction of these one-peptide constructs.

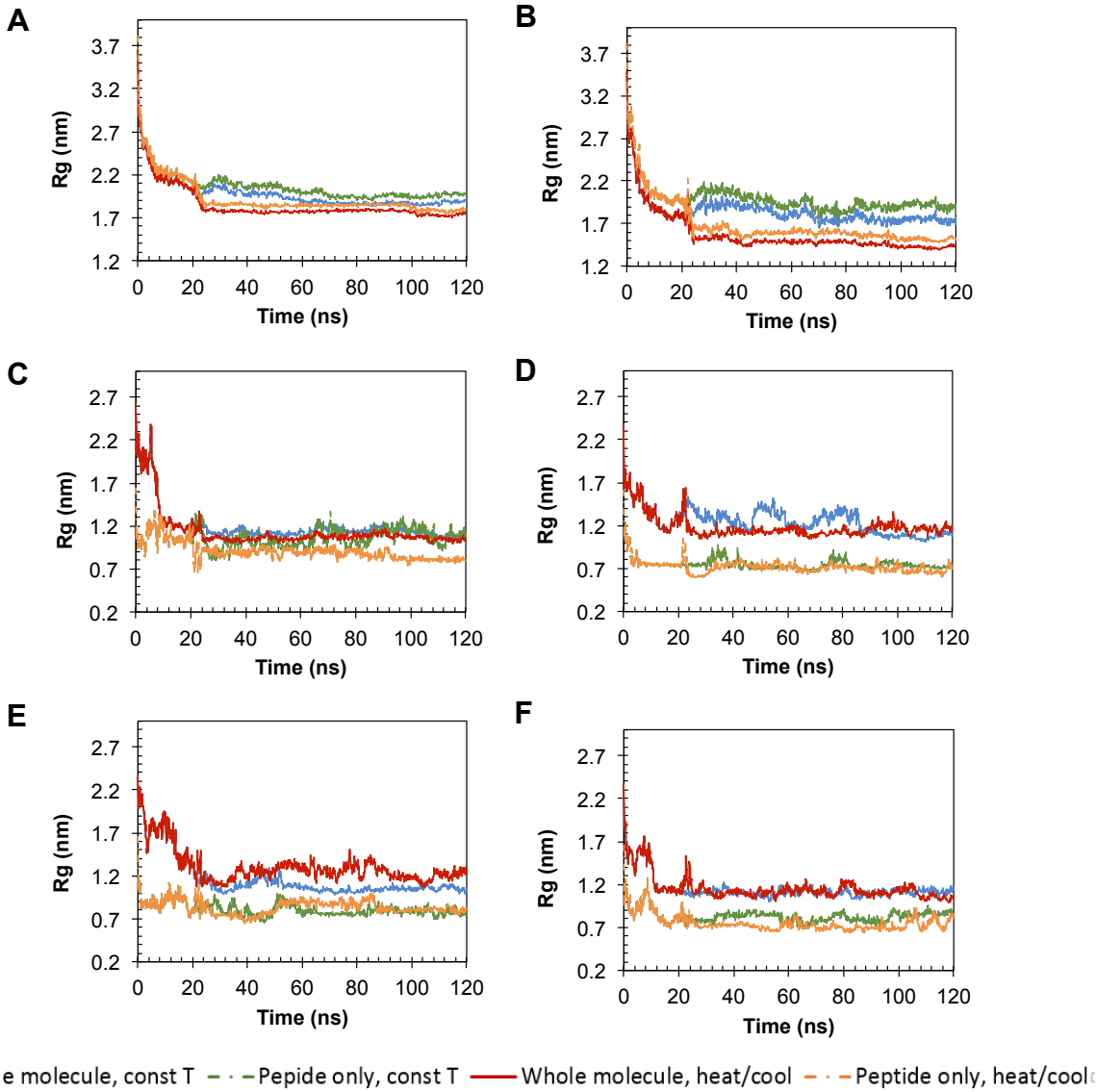

Figure S31. Radii of gyration (Rg) of simulated polymers under constant temperature condition (300 K) and using simulated annealing procedure. Graphs describe A) simulated polymer **1**, B) simulated polymer **2**, C) simulated polymer **3**, D) simulated polymer **4**, E) simulated polymer **5**, F) simulated polymer **6**. Note that the results split at 20 ns, because the runs are continued at 300K (blue and green) or subjected to a brief heat/cool cycle and then continued at 300 K (red and orange).

The values of R_g observed here may be put into context by comparing them with the values associated with folded proteins having a similar number of residues or molecular weight. Thus, simulated polymer **1**, with 200 monomers and molecular weight of 23.2 kDa, has final R_g values of 1.8-2.0 nm. For comparison, folded proteins with 151-200 residues have R_g values ranging approximately from 1.6 – 1.8 nm,[1] and the proteins ribonuclease and chymotrypsinogen alpha, with molecular weights of 17 and 38 kDa, have R_g values of 1.5 and 1.8 nm, respectively.[2] Although simulated polymers **3**-**6**, have molecular weights about four times smaller (5-6 kDa), their R_g values of ~1.2 nm are not even two-fold smaller, suggesting that these single-peptide constructs do not compact as tightly as the more peptide-rich **1** and **2**.

We quantified conformational fluctuations of the simulated molecules by using overall rotations and translations to optimize the overlay of each MD snapshot on the final snapshot (120 ns) of the corresponding MD trajectory, and computing the root-mean-square deviations (RMSDs) of the overlaid

snapshots relative to the final snapshots, as a function of time. As further detailed in Figure S32 and Table S6, the RMSD values in all cases remained less than 1.5 nm in the final 100 ns of each simulation, and in many cases, they remained below 1.0 or even 0.4 nm, particularly for the trajectories which include the heat-cool cycle. Presumably, the heat-cool cycle in these cases allowed the structures to find a particularly stable conformational state. The structural fluctuations observed here are larger than those typically observed in similar MD simulations of stably folded, naturally occurring proteins, but the lower end of the present range is commensurate. It should be noted that the final conformations generated from the simple 300K simulations and from the corresponding heat-cool simulations are quite different from each other: for simulated polymers **1** and **2**, the RMSDs between these conformer are 1.35 and 1.32 nm respectively. We anticipate that further conformational sampling, especially if facilitated by further heat-cool cycles, would lead to additional different, collapsed conformations.

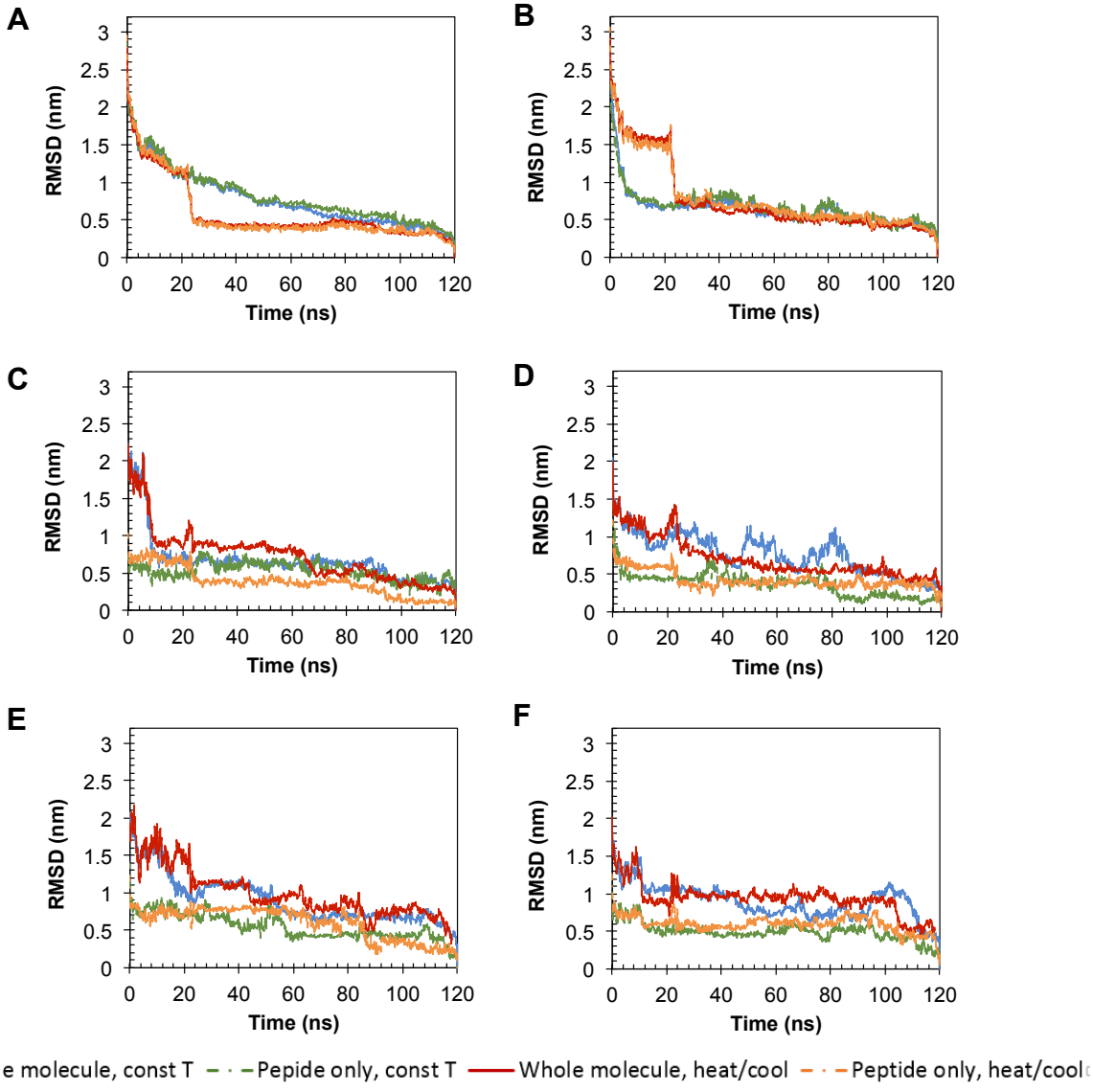

Figure S32. RMSD of simulated polymers. Graphs describe A) simulated polymer **1**, B) simulated polymer **2**, C) simulated polymer **3**, D) simulated polymer **4**, E) simulated polymer **5**, F) simulated polymer **6**. RMSD values were calculated with respect to the structure in the last frame in each run. The program g_rms in Gromacs[3] was used to generate the rotational and translational overlays and compute the RMSD values.

	Polymer 1	Polymer 2	Polymer 3	Polymer 4	Polymer 5	Polymer 6
AA-AA	88	33	7	6	8	5
EDANS-AA	24	9	0	1	-	-
DABCYL-AA	7	3	0	0	-	-
Peptide-peptide	119	45	7	7	8	5
PEG-AA	0	2	1	1	2	0
NOR-AA	2	1	2	0	0	2
EDANS-OEG	0	0	0	0	0	0
DABCYL-OEG	0	0	0	0	0	0
EDNAS-NOR	1	1	0	0	0	0
DABCYL-NOR	0	0	0	0	0	0
Peptide-nonpeptide	3	4	3	1	2	2

Table S6. Hydrogen bond counts of simulated polymers **1-6**, averaged over the last 40 ns of the heat-cool trajectories. Here "peptide" refers to the entire peptide or peptide-dye component of each construct, AA refers to only the amino acid residues of these peptide chains, and "nonpeptide" refers to the OEG and norbornene (NOR) components.

Representative conformations of simulated polymers **1-4** (Figure S33) were obtained by applying an RMSD-based clustering algorithm to the last 40 ns of the respective heat-cool simulations. Homopolymer **1** and the 5:5 copolymer **2** collapsed into an elongated globule, with their peptide chains tangled around the polymer backbone. The single peptide chains in the 1:9 copolymers, **3** and **4**, are visible as relatively isolated components at the surface of these peptide-polymer constructs. Salt-bridges and occasional segments of alpha-helix were observed (Figure S33).

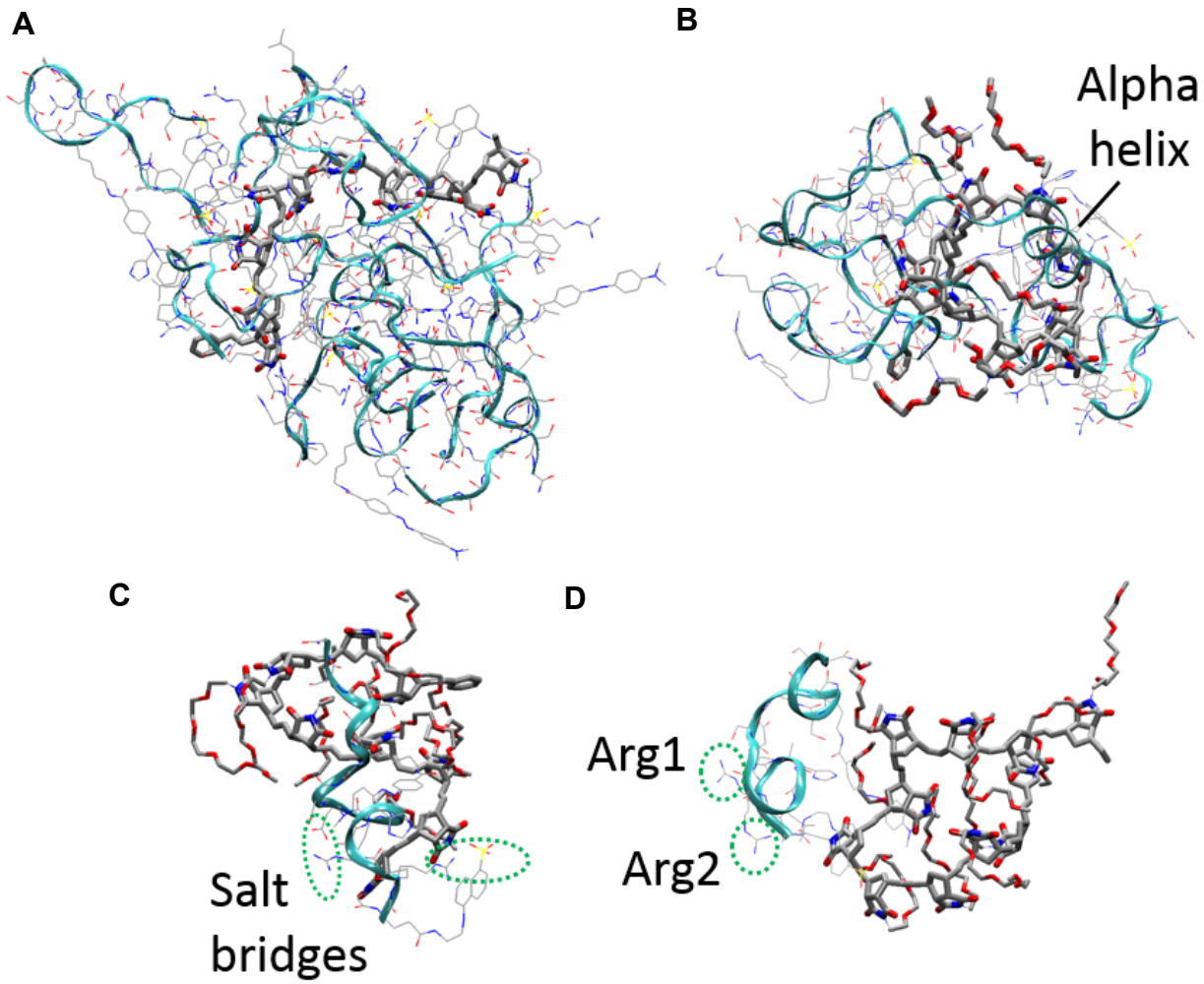

Figure S33. Representative conformation of peptide-polymer models. Representative conformations for A) simulated polymer **1**; B) simulated polymer **2**; C) simulated polymer **3**; D) simulated polymer **4**. The peptide chain is shown in ribbon representation. The norbornyl polymer backbone and OEG units are given in a licorice representation. Peptide residues and dye molecules (EDANS and DABCYL) are depicted by lines. Carbon atoms are colored in silver, nitrogen in blue, oxygen in red and sulfur in yellow. Hydrogen atoms are neglected for clarity.

REFERENCES

(1) Lobanov, M.; Bogatyreva, N. S.; Galzitskaia, O. V. *Molekuliarnaia Biologiia* **2008**, *42*, 701.

(2) He, L.; Niemeyer, B. *Biotechnology Progress* **2003**, *19*, 544.

(3) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845.