Supporting Information
High Serum Stability of Collagen Hybridizing Peptides

Lucas L. Bennink[a], Daniel J. Smith[a], Catherine A. Foss[b], Martin G. Pomper[b], Yang Li[a,*], and S. Michael Yu[a,c,*]

[a]Department of Bioengineering, University of Utah, Salt Lake City, UT 84112, USA.
[b]The Russel H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, School of Medicine, Baltimore, MD 21228, USA
[c]Department of Pharmaceutics and Pharmaceutical Chemistry, University of Utah, Salt Lake City, UT 84112, USA.

*Corresponding authors: yang.d.li@utah.edu (Yang Li), michael.yu@utah.edu (S. Michael Yu)

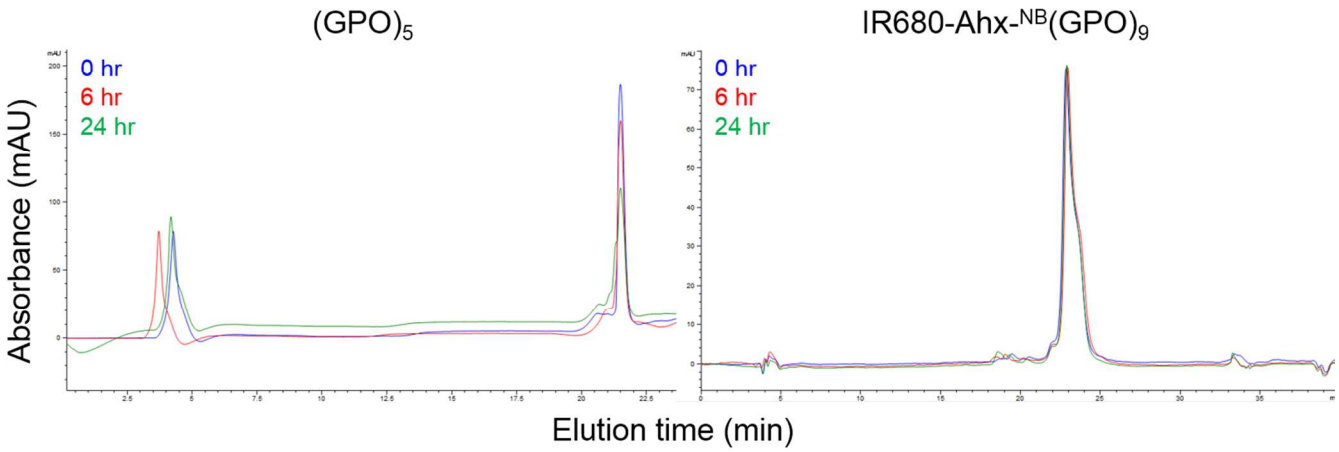


Figure S1. Representative raw HPLC profiles of peptides $(GPO)_5$ (left) and IR680-Ahx-$^{NB}(GPO)_9$ (right) after 0 hr (blue), 6 hr (red), and 24 hr (green) of incubation in 25% mouse serum at 37 °C. The $(GPO)_5$ target peak eluted at 21.5 min and the IR680-Ahx-$^{NB}(GPO)_9$ target peak eluted at 23 min. Size of the target peak at each time point was integrated, normalized, and presented in Figures 1-3. The IR680-Ahx-$^{NB}(GPO)_9$ was used for *in vivo* imaging applications.

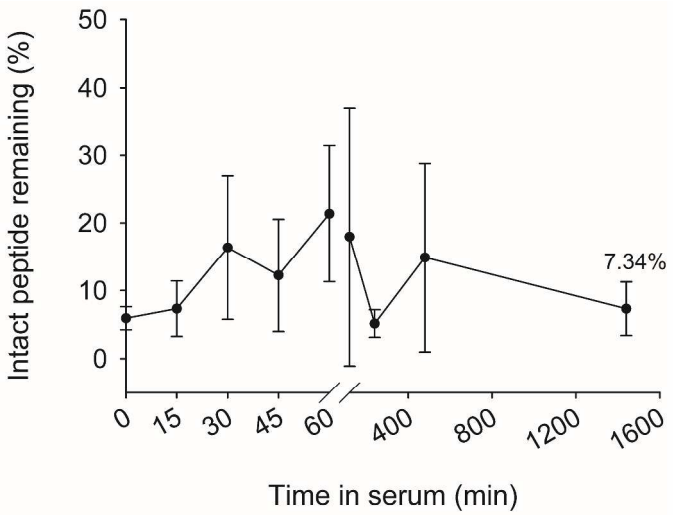


Figure S2. Serum stability profile of penetratin (sequence: RQIKIWFQNRRMKWKKGG). After 24 hr in serum most of penetratin is degraded or bound to serum proteins.

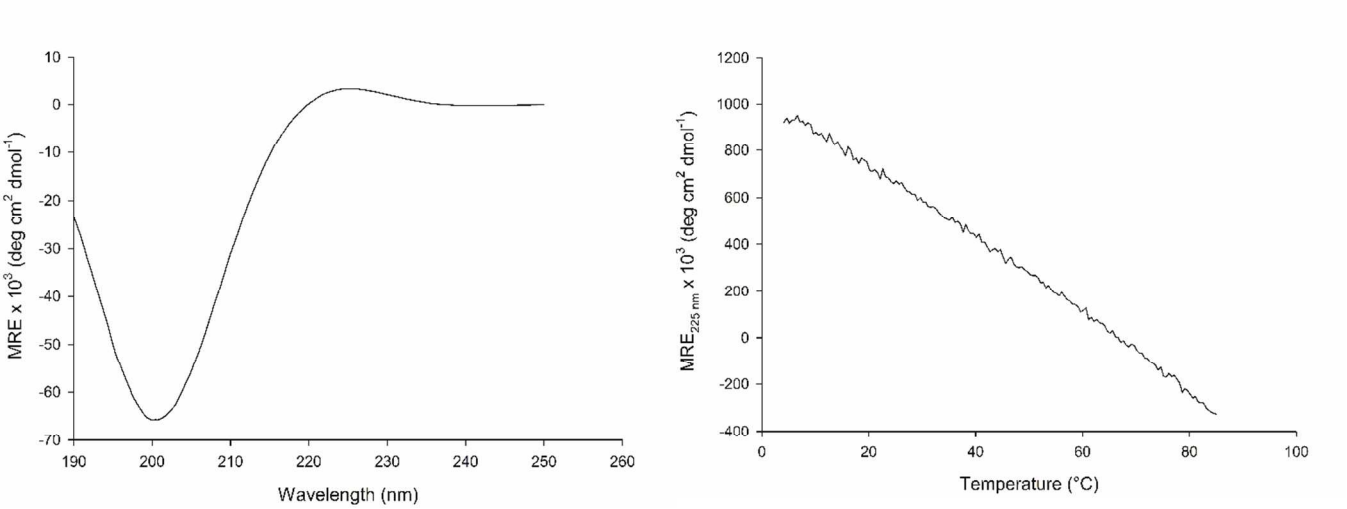


Figure S3. The CD spectrum (left) and melting curve (right) of $(GPO)_5$ (150 μM in 1×PBS solution). The low ellipticity value at 225 nm at 4 °C and the linear decrease of signal at 225 nm during thermal melting indicate lack of the triple helical structure. Samples were incubated at 4 °C for at least 24 hr before CD measurement to ensure peptide folding.

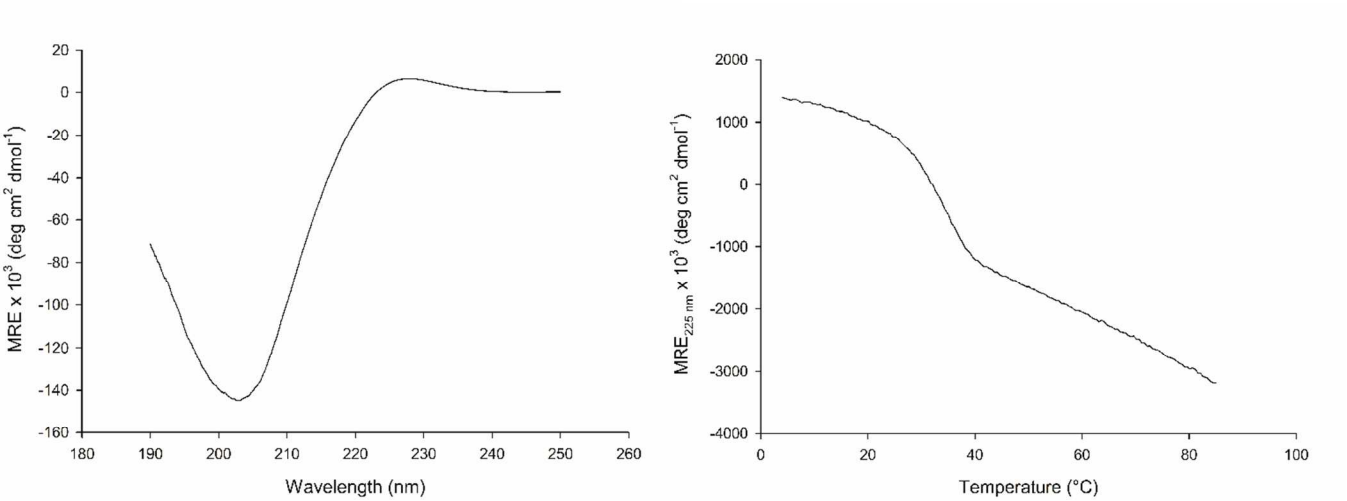


Figure S4. The CD spectrum (left) of $(GPP)_9$ (150 μM in 1×PBS solution) at 4 °C and its melting curve (right) showing very little triple helical content at 37 °C, and complete peptide unfolding above 39 °C. Samples were incubated at 4 °C for at least 24 hr before CD measurement to ensure folding.

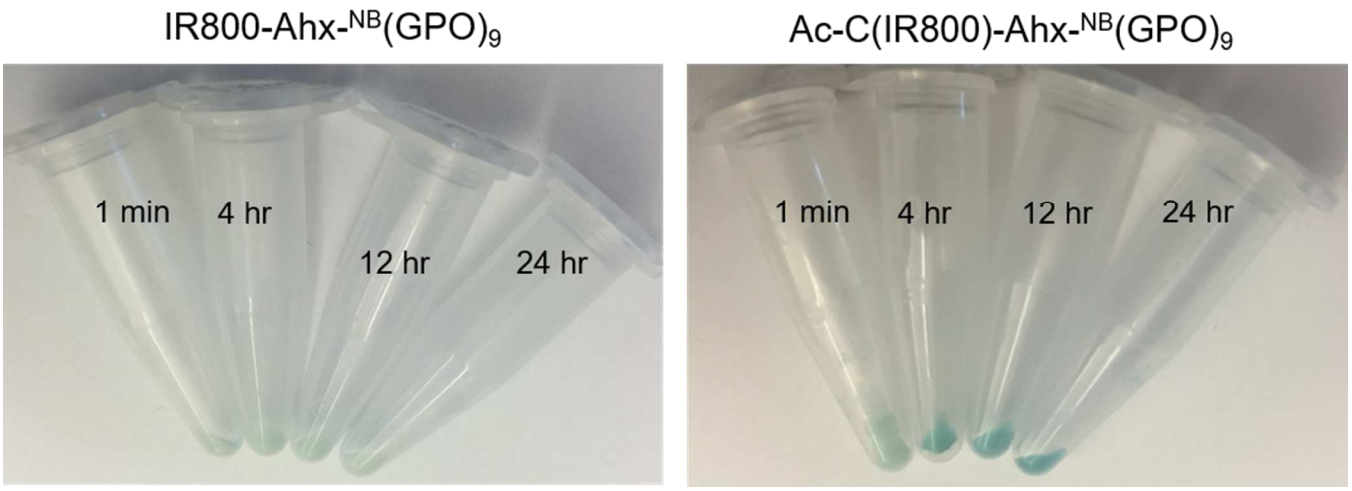


Figure S5. Photographs of protein pellets precipitated from mouse serum samples incubated with IR800-Ahx-$^{NB}(GPO)_9$ or Ac-C(IR800)-Ahx-$^{NB}(GPO)_9$ for 1 min, 4 hr, 12 hr, and 24 hr. The pellets from serum samples incubated with Ac-C(IR800)-Ahx-$^{NB}(GPO)_9$ showed a clear dark green color indicating the presence of IR800 dye within the pellets. The interaction between Ac-C(IR800)-Ahx-$^{NB}(GPO)_9$ and the serum proteins seemed to happen immediately and reach a steady state within 4 hr, as visualized by the color intensities of the pellets. In contrast, the pellets from serum incubated with IR800-Ahx-$^{NB}(GPO)_9$ showed minimal color change up to 24 hr of incubation.

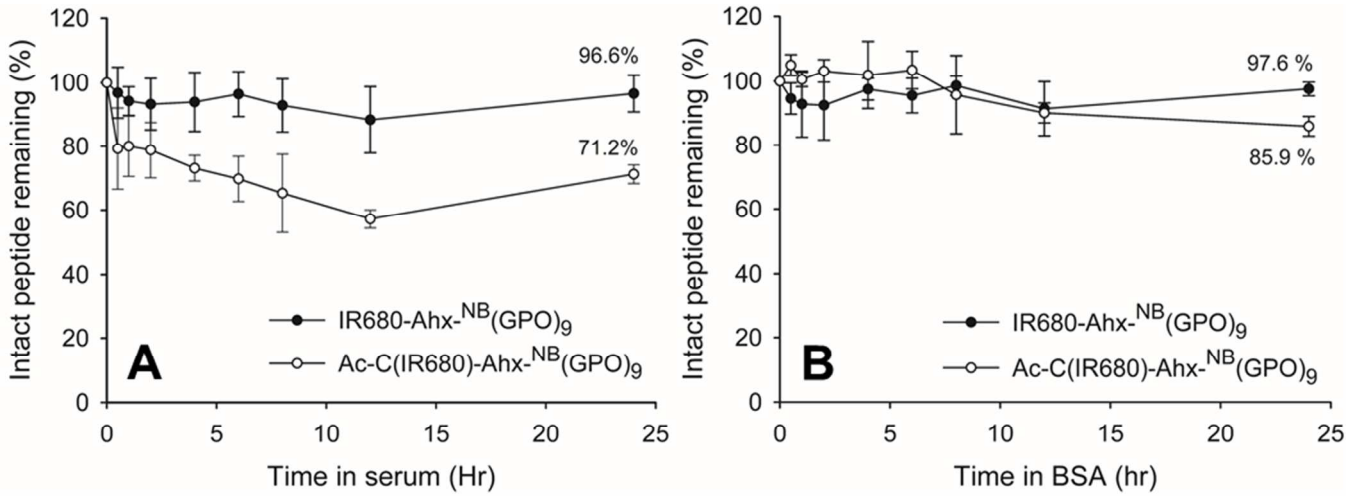


Figure S6. Stability of IR680-labeled CHPs. (A) Stability profiles of IR680-Ahx-$^{NB}(GPO)_9$ and Ac-C(IR680)-Ahx-$^{NB}(GPO)_9$ in mouse serum for 24 hr. (B) Percentages of intact IR680-Ahx-$^{NB}(GPO)_9$ and Ac-C(IR680)-Ahx-$^{NB}(GPO)_9$ incubated with BSA for 24 hr. Overall, less interaction with serum proteins occurred with the IR680 dye than the IR800 dye (Figure 3B). Ac: acetyl.

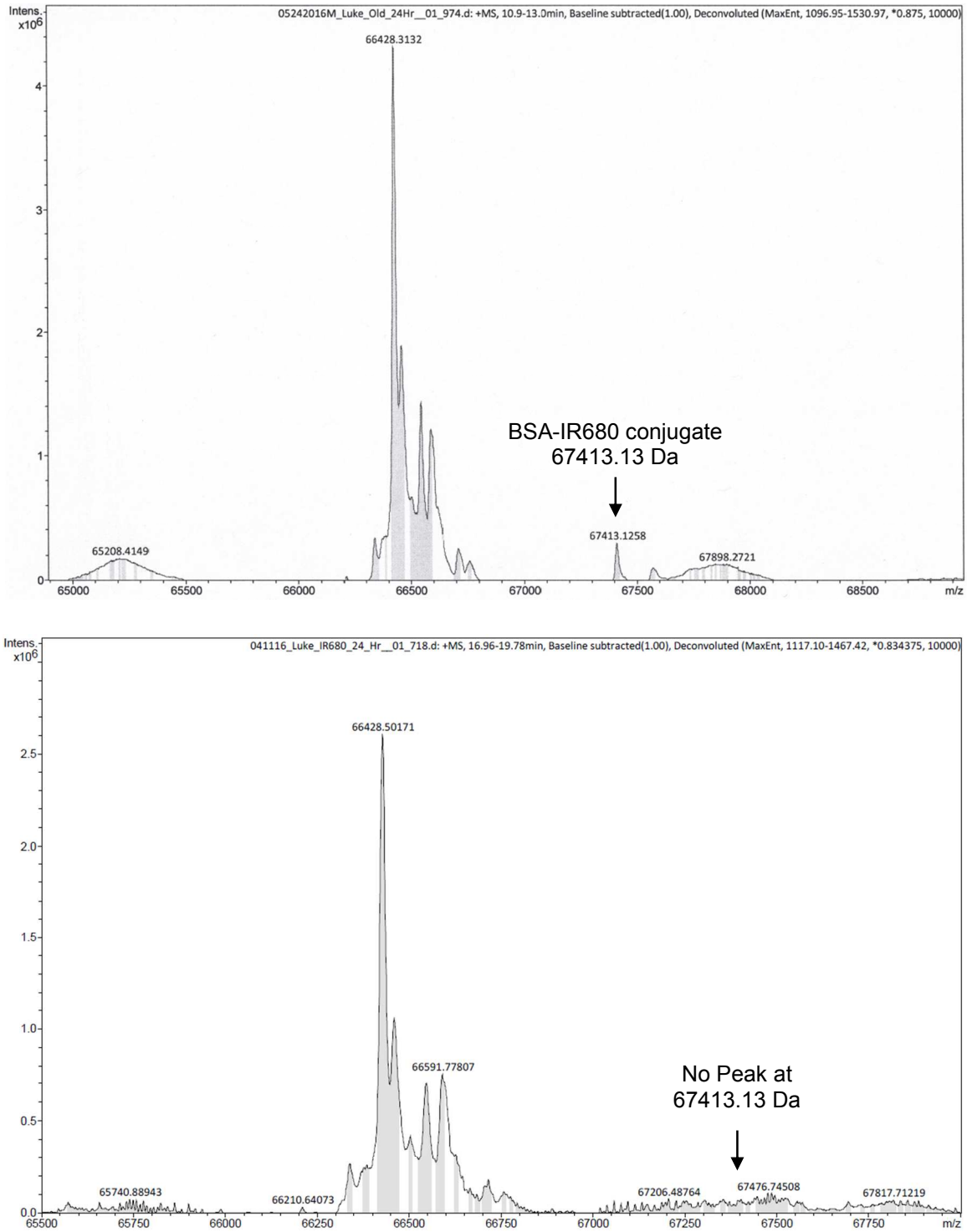


Figure S7. The ESI-MS spectrum of BSA incubated with Ac-C(IR680)-Ahx-$^{NB}(GPO)_9$ for 24 hr (top). The large peak at 66428.62 m/z corresponds to the BSA protein. The peak at 67413.13 m/z corresponds to BSA (MW: 66428.3 g/mol) covalently conjugated with IR680-maleimide (MW: 982.5g/mol). In the ESI-MS spectrum of BSA incubated with IR680-Ahx-$^{NB}(GPO)_9$ for 24 hr (bottom), the BSA-IR680 conjugate peak is absent.

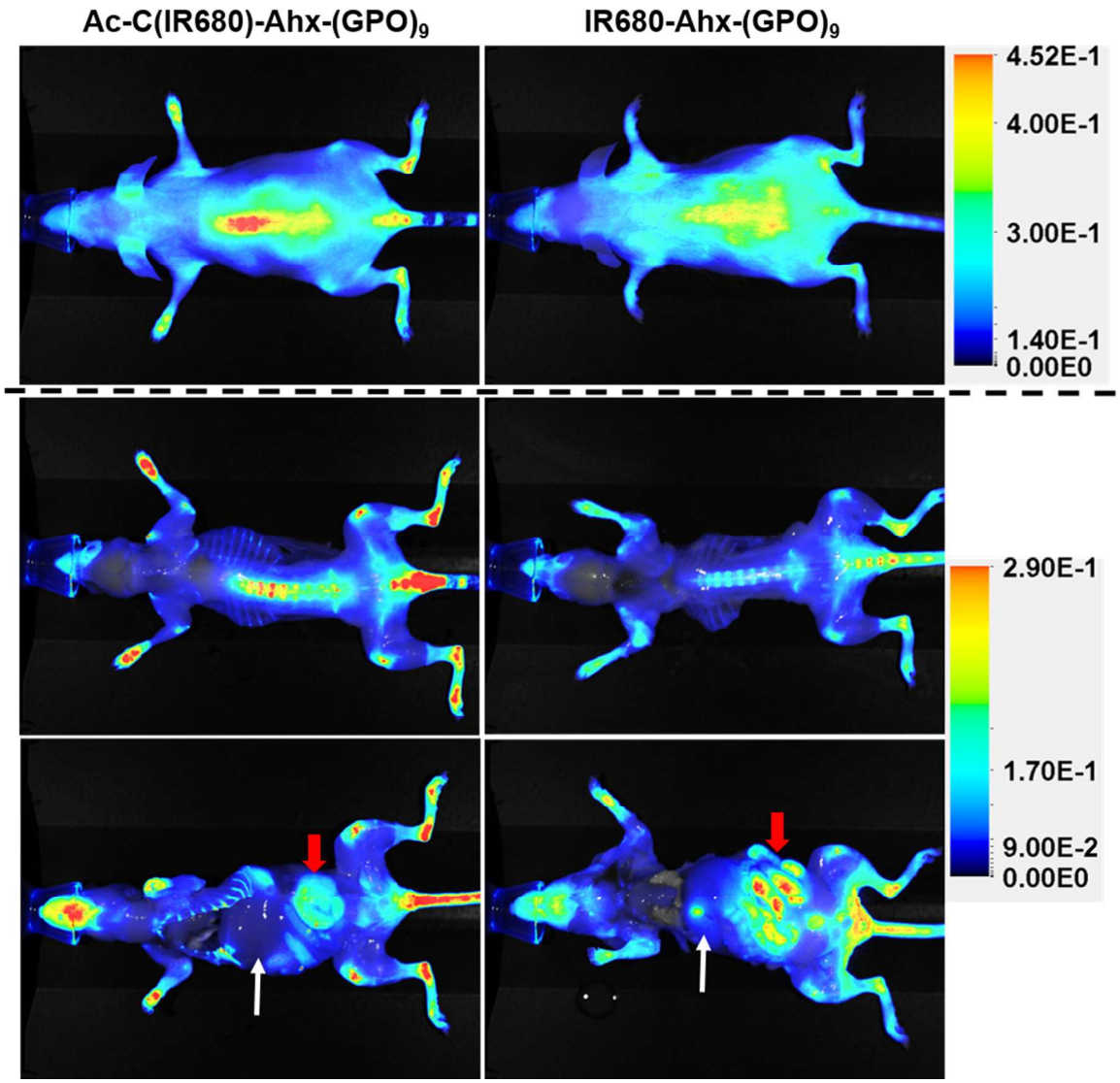
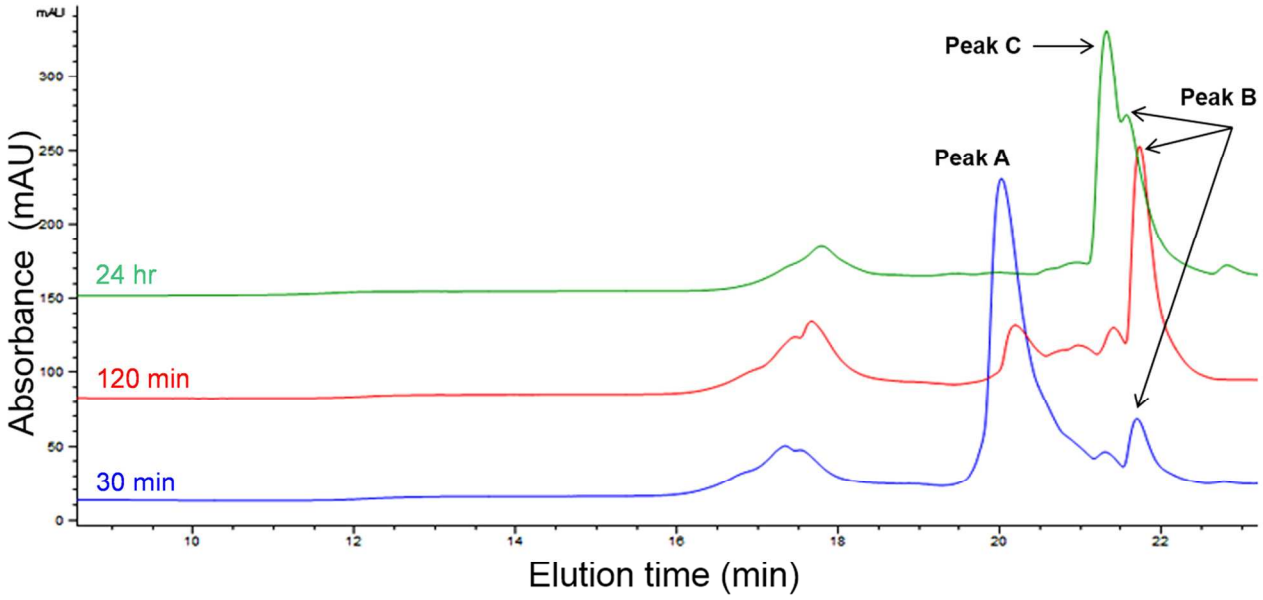


Figure S8. NIR fluorescence imaging showing *in vivo* distribution of Ac-C(IR680)-Ahx-$(GPO)_9$ and IR680-Ahx-$(GPO)_9$ in mice 48 hr post tail vein injection. Both conjugates showed very similar biodistribution pattern with low uptake in the liver (white arrows). Red arrows indicate false positive signal in the digestive system caused by the fluorescence of the chlorophyll from the diet of the mice, which is significant in the 680 channel. The images below the dashed line were taken after skin removal for clear display of signals in the skeleton and internal organs, and their fluorescence intensities were adjusted to the same scale for direct comparison. All *in vivo* experiments were performed three times with similar results.



Sample	Sequence or Fragment		m/z Calculated	m/z Found
Peak A	CF-Ahx-GPOGPKGPOGPKGPOGPKGPOGPKGPO-NH_2	[M+H$^+$]	2952.0	2952.6
Peak B	CF-Ahx-GPOGPKGPOGPKGPOG	[M+H$^+$]	1912.7	1904.3
	CF-Ahx-GPOGPK	[M+H$^+$]	1039.4	1039.2
Peak C	CF-Ahx-GPOGP	[M+H$^+$]	911.3	911.1

Figure S9. The HPLC chromatograms showing the degradation profile of CF-$(GPO/K)_9$. The size of the intact peptide peak (Peak A) decreased rapidly as the degradation peaks (Peaks B and C) increased over the course of the 24 hr incubation in mouse serum. At the end of the assay, the original peptide was no longer present, leaving only degraded fragments. Peptide fragments were collected and their mass was verified with MALDI-TOF MS (included above) to determine the possible sequences and cleavage sites. The MS results suggest that the degradation occurred in the middle of the $(GPO/K)_9$ sequence, not from the CF labeled N-terminus. The C-termini of fragments in peaks B and C are assumed as COOH.

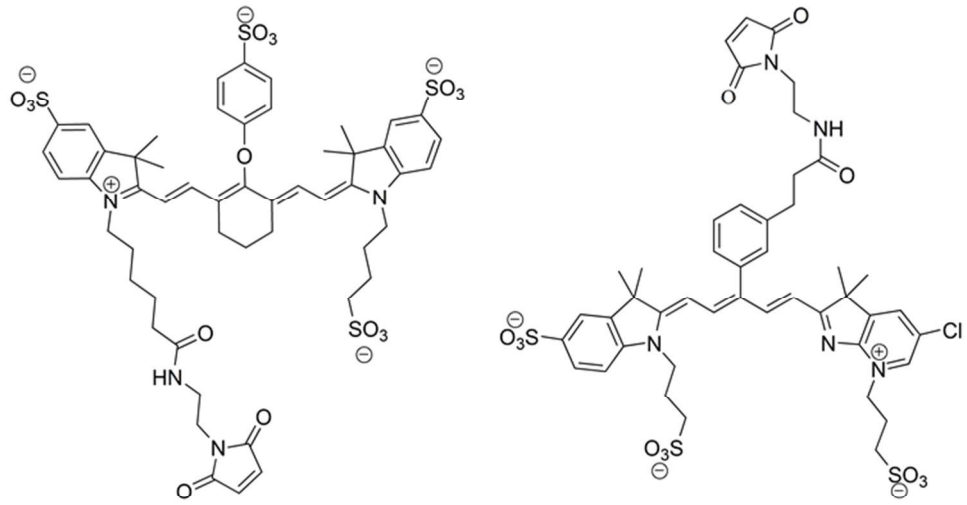


Figure S10. Structures of IR800-maleimide and IR680-maleimide. The two dyes are different in backbone structure and maleimide position. In addition, at physiological conditions, the net charge of IR800-maleimide is -3 and that of IR680-maleimide is -2.

Table S1. Sequences and MALDI-TOF MS of all CHPs.

Peptide	Sequence		*m/z* Calculated	*m/z* Found
$(GPO)_9$	NH_2-GPOGPOGPOGPOGPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	2422.86	2423.72
$^{NB}(GPO)_9$	NH_2-GPOGPOGPOGPO^{NB}GPOGPOGPOGPOGPO-$CONH_2$	$[M+Na^+]$	2577.98	2578.34
$(GPO)_5$	NH_2-GPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	1354.05	1354.23
$^{S}(G_9P_9O_9)$	NH_2-PGOGPGPOPOGOGOPPGOOPGGOOPPG-$CONH_2$	$[M+H^+]$	2422.86	2422.32
$(GPP)_9$	NH_2-GGG-GPPGPPGPPGPPGPPGPPGPPGPPGPP-$CONH_2$	$[M+Na^+]$	2471.23	2471.41
CF-$(GPO)_9$	CF-GGG-GPOGPOGPOGPOGPOGPOGPOGPOGPO-$CONH_2$	$[M+Na^+]$	2973.24	2972.63
CF-$^{NB}(GPO)_9$	CF-GGG-GPOGPOGPOGPO^{NB}GPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	3084.36	3085.51
CF-$^{S}(G_9P_9O_9)$	CF-GGG-PGOGPGPOPOGOGOPPGOOPGGOOPPG-$CONH_2$	$[M+Na^+]$	2973.24	2972.73
CF-$^{NB}(GPP)_9$	CF-GGG-GPPGPPGPPGPP^{NB}GPPGPPGPPGPPGPP-$CONH_2$	$[M+Na^+]$	2961.67	2964.4
CF-$(GPO)_5$	CF-GGG-GPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	1882.42	1883.15
CF-$(GPO/K)_9$	CF-Ahx-GPOGPKGPOGPKGPOGPKGPOGPKGPO-$CONH_2$	$[M+H^+]$	2952.8	2952.66
Ac-C(IR800)-Ahx-$^{NB}(GPO)_9$	Ac-C(IR800)-Ahx-GPOGPOGPOGPONBGPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	3939.86	3919.77
IR800-Ahx-$^{NB}(GPO)_9$	IR800-Ahx-GPOGPOGPOGPONBGPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	3654.32	3654.41
Ac-C(IR680)-Ahx-$^{NB}(GPO)_9$	Ac-C(IR680)-Ahx-GPOGPOGPOGPONBGPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	3798.72	3798.73
IR680-Ahx-$^{NB}(GPO)_9$	IR680-Ahx-GPOGPOGPOGPONBGPOGPOGPOGPOGPO-$CONH_2$	$[M+H^+]$	3513.57	3513.49