Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Conformational Investigation of the Structure – Activity Relationship of GdFFD and Its Analogs on an Achatin-like Neuropeptide Receptor of *Aplysia californica* Involved in the Feeding Circuit

Thanh D. Do,[‡,] James W. Checco,[‡] Michael Tro,[±] Joan-Emma Shea,[¥,±] Michael T. Bowers[±] and Jonathan V. Sweedler[‡,*]*

[‡]Department of Chemistry and the Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States, and [±]Department of Chemistry and Biochemistry, [¥] Department of Physics, University of California at Santa Barbara, Santa Barbara, Californica 93106, United States.

***** Corresponding authors:
Jonathan V. Sweedler. Email: jsweedle@illinois.edu
Thanh D. Do. Email: tdo5@utk.edu

SUPPORTING INFORMATION

Table of Contents

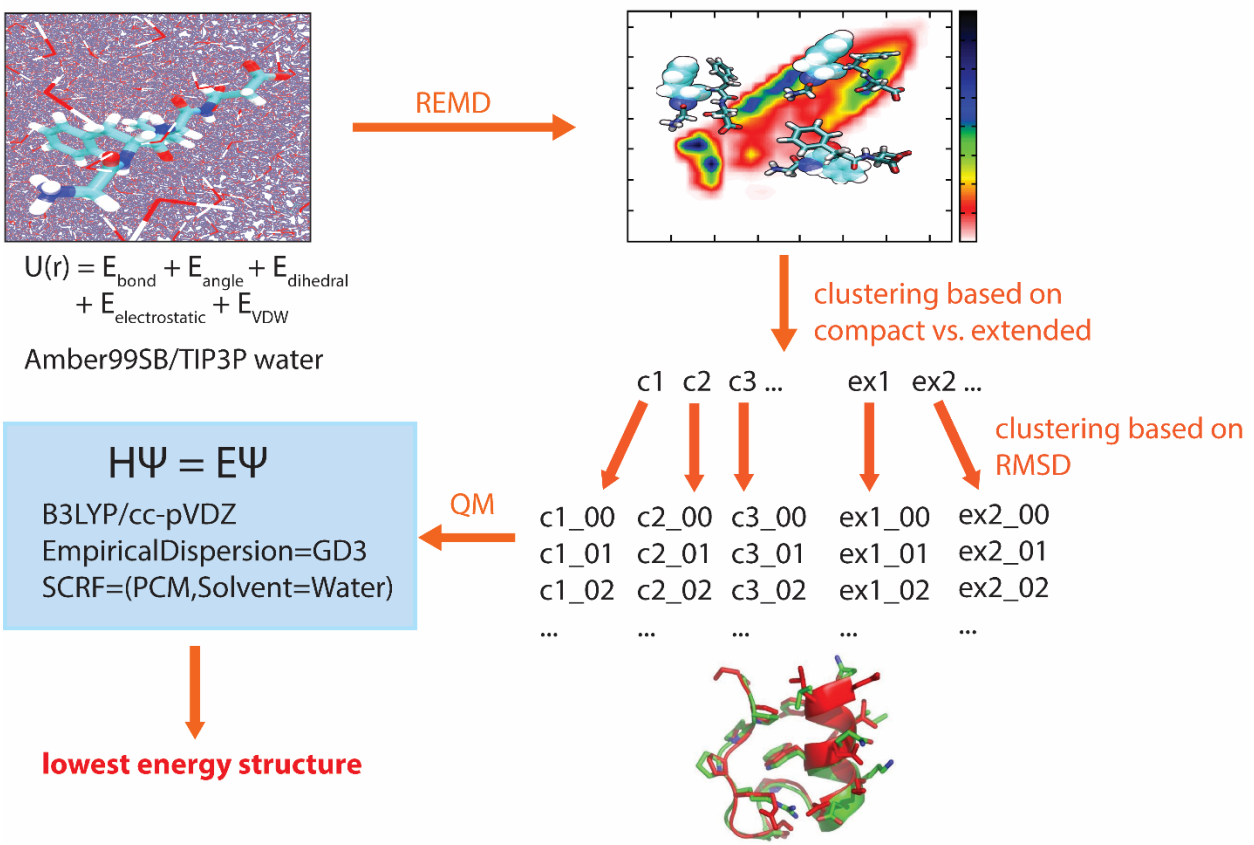


Figure S1. The computational modeling workflow to obtain global energy minimum structures of GdFFD and its analogs. For each peptide, an ensemble of structures was generated through the means of the enhanced sampling T-REMD method. The peptide was simulated using the combination of the Amber ff99SB force field and TIP3P explicit solvent water model. After the REMD simulation reached 200 ns, the trajectory at 300K was clustered into compact and extended families of structures. Each family was then clustered into smaller sub-populations. A representative structure of each subpopulation was subjected to geometry optimization and relative energy calculations using the B3LYP level of theory, the cc-pVDZ basis set, the Grimme's dispersion correction GD3, and the implicit water model PCM. The resulting structures were sorted based on energy to determine the global minimum structure.

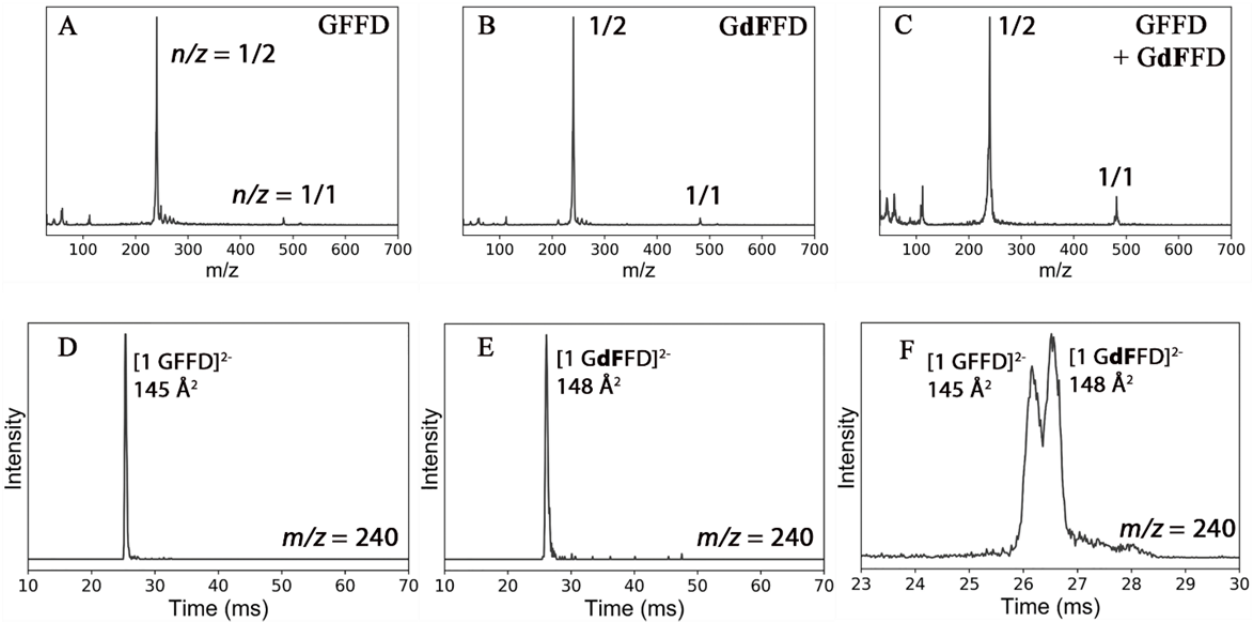


Figure S2. Representative mass spectra of (A) GFFD, (B) G**dF**FD, and (C) equal molar mixture of GFFD and G**dF**FD in water at the concentration of 50 µM. Mass spectral peaks are annotated with *n/z* where *n* is the number of peptide chain and *z* in the number of (negative) charge. (D–F) Representative ATDs of doubly charged species (*z* = -2, *m/z* 240) collected from pure standards of GFFD and G**dF**FD, and the mixture. All data were collected in negative polarity.

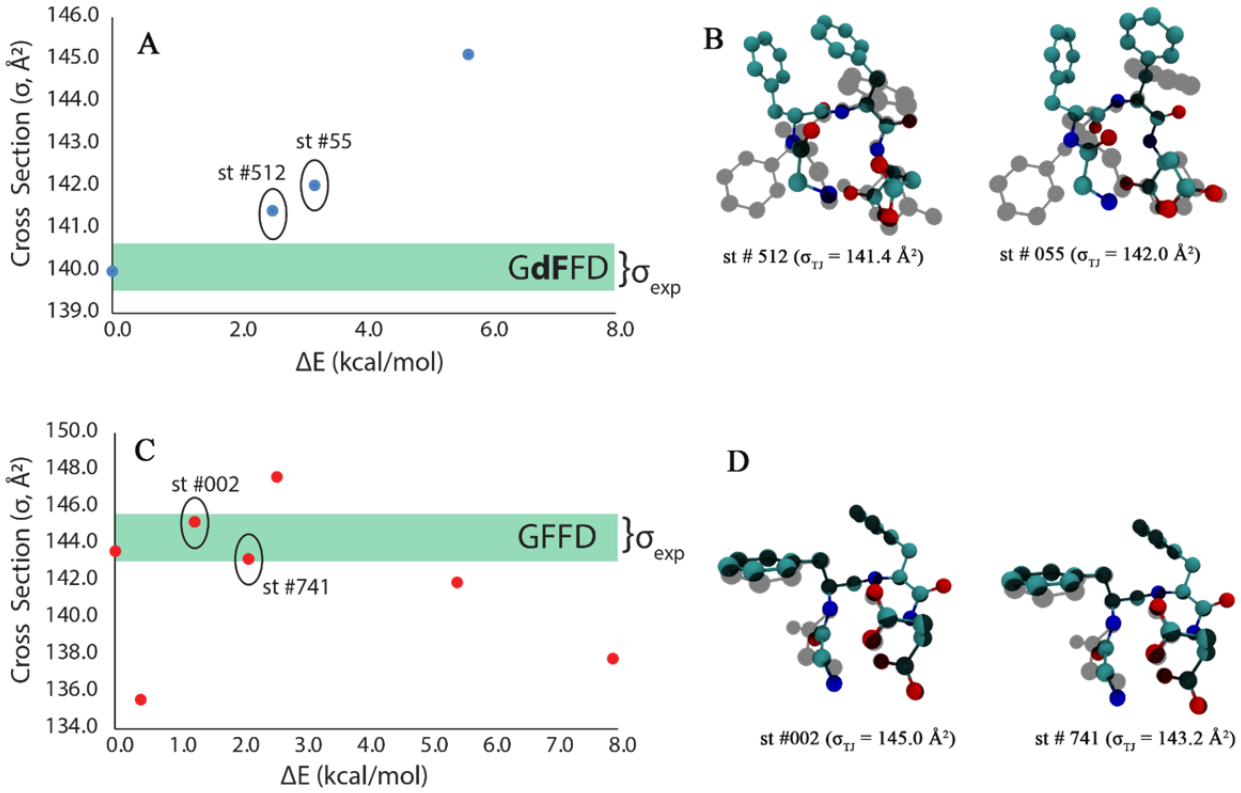


Figure S3. Theoretical CCS vs. ΔE for (A) G**dF**FD and (C) GFFD. (B, D) Overlaid images of structures (shown in colors) with similar CCSs onto the lowest energy structures of GdFFD/GFFD (shown as shadow).

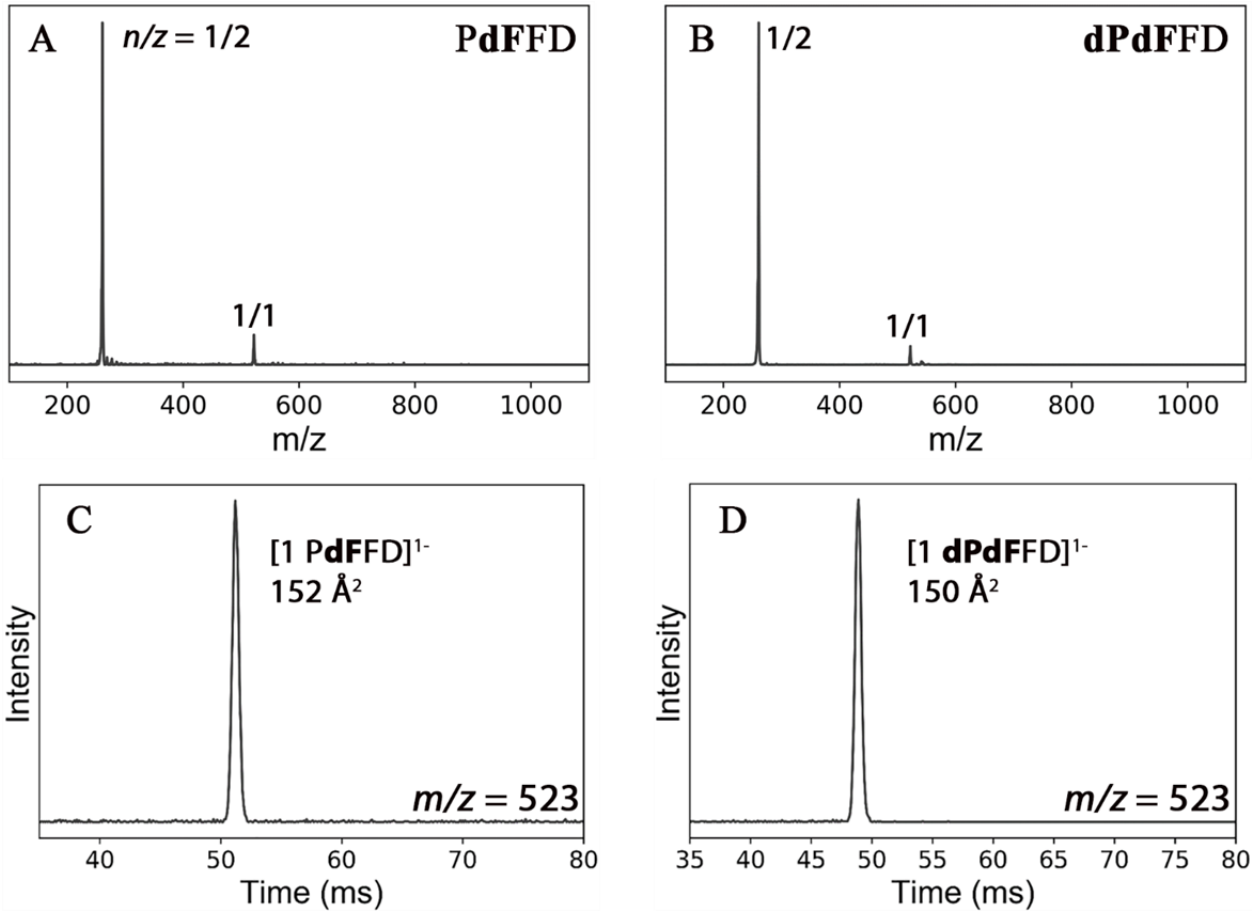


Figure S4. (A, B) Representative mass spectra of P**dF**FD and **dPdF**FD in water. Mass spectral peaks were annotated with *n/z* where *n* the number of peptide chain and *z* is the charge. (C, D) Representative ATDs of singly charged P**dF**FD and **dPdF**FD. The CCSs are also shown.

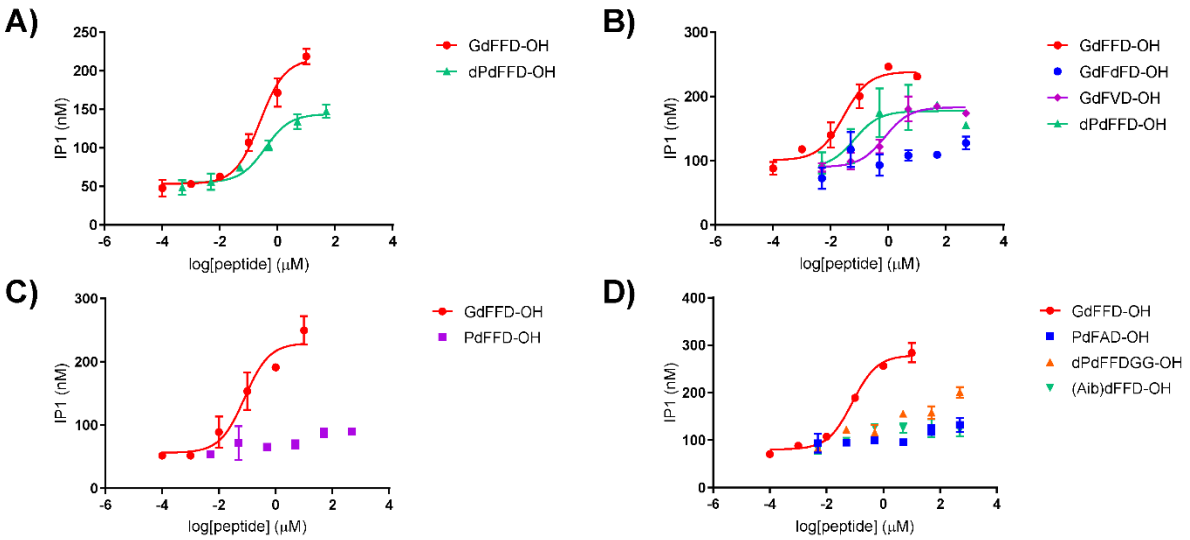


Figure S5. Dose-response curves of peptides tested for activation of apALNR, as determined by IPOne assay. CHO-K1 cells were co-transfected with plasmids for both *apALNR* and *Gα-16*. Each point represents the mean ± SEM from duplicate wells on the plate. Calculated EC_{50} values for each experiment are: (A) GdFFD-OH = 252 nM, dPdFFD-OH = 373 nM; (B) GdFFD-OH = 28 nM, dPdFFD-OH = 66 nM, GdFVD-OH = 748 nM; (C) GdFFD-OH = 81 nM; (D) GdFFD-OH = 82 nM. The EC_{50} value for dPdFFD-OH given in the main text (200 nM) is the average of the value obtained from panels (A) and (B), rounded to 1 significant figure.

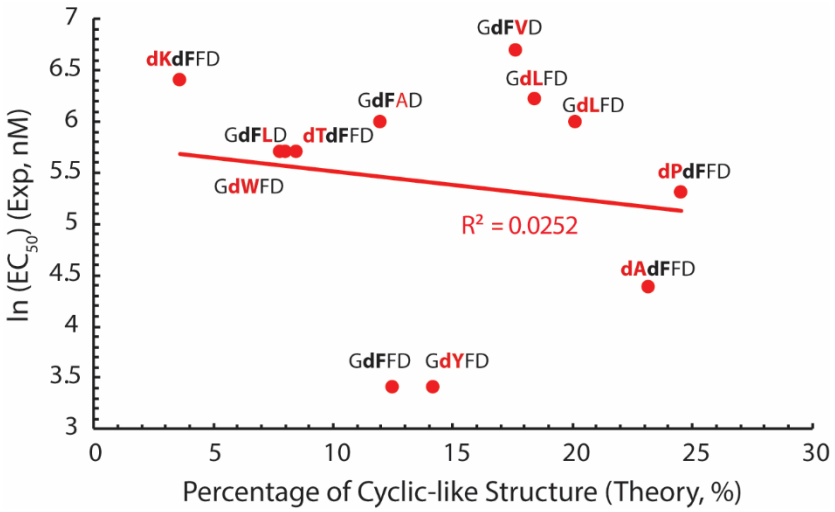


Figure S6. Correlation between the percentage of cyclic-like structures; i.e., end-to-end distance < 0.42 nm, to natural log of experimental EC_{50} for active peptides discussed in this work. In general, peptides with a low percentage of cyclic-like structures (< 10%) have weak activities. On the other hand, those with a high percentage of cyclic-like structures are not necessarily strongly active.

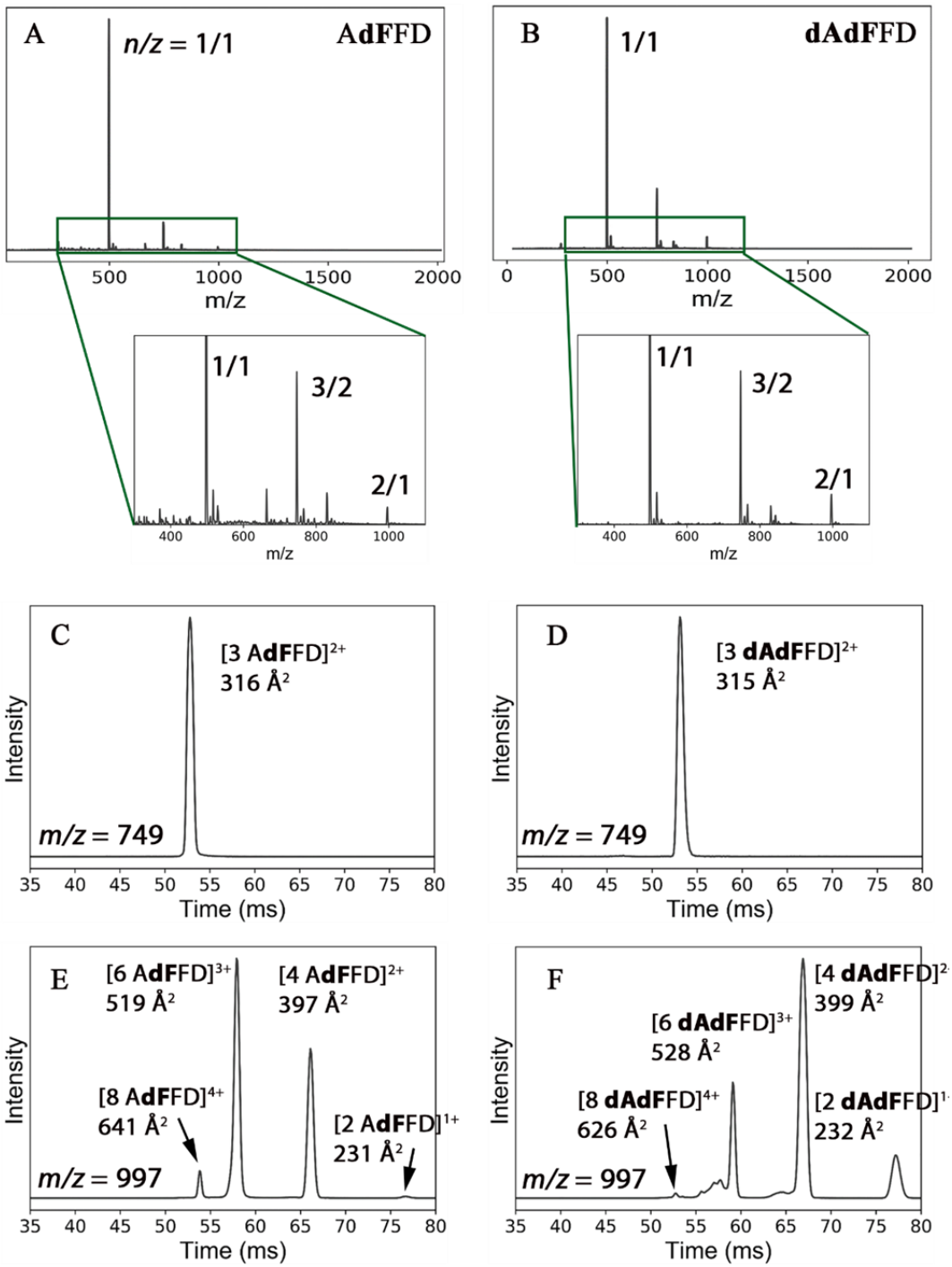


Figure S7. Representative mass spectra of (A) A**d**FFD and (B) **dAdF**FD in water. The *m/z* regions between 300 and 1100 are also shown. Mass spectral peaks are annotated with *n/z* where *n* is the oligomer size and *z* is the number of (positive) charge. (C, D) Representative ATDs of *n/z* = 3/2 at *m/z* = 749. (E, F) Representative ATDs of *n/z* = 2/1 at *m/z* = 997 show the formation of large oligomers up to octamer (*n* = 8). Each feature is labeled with the oligomer species and its corresponding CCS. For [8 dAdFFD]$^{4+}$, the intensity of the feature is low, thus the CCS may not be accurate. All data were collected in positive polarity.

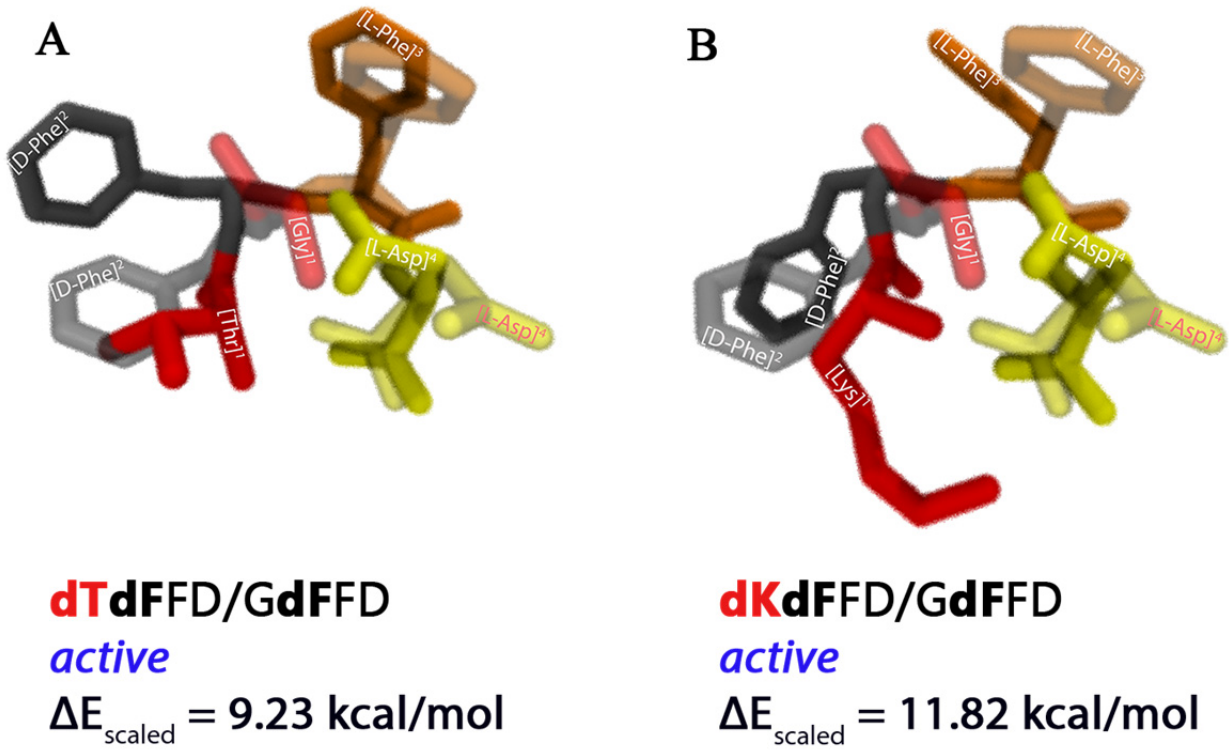


Figure S8. Overlaid structures of (A) **dTdF**FD and (B) **dKdF**FD onto G**dF**FD (shown in lighter colors). In both cases, dT and dK residues do not overlap in coordinate with $[Gly]^1$ of GdFFD.

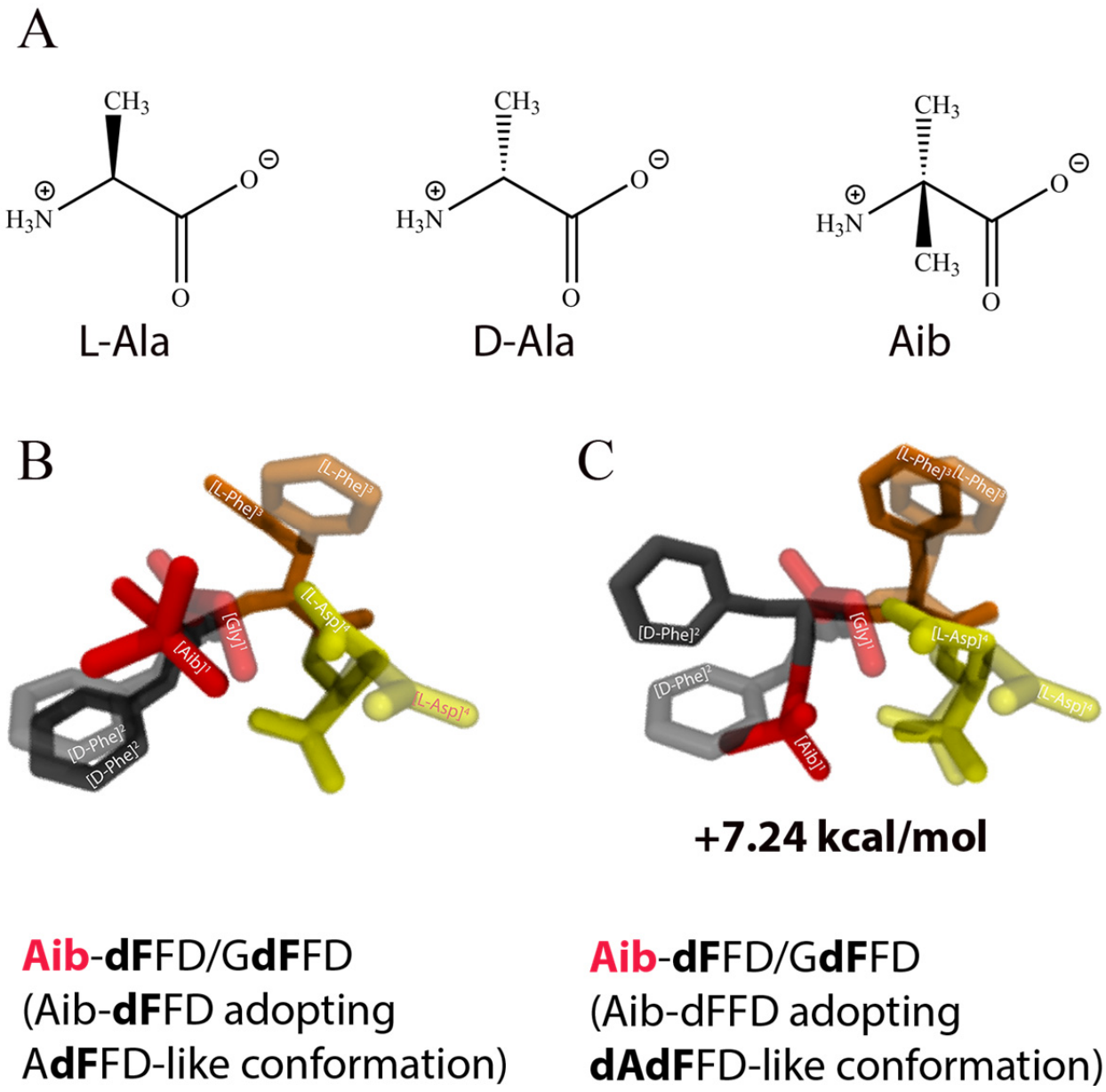


Figure S9. (A) Chemical structures of L-Ala, D-Ala and Aib amino acids. (B, C) Overlaid structures of two Aib-**dF**FD conformations onto G**dF**FD. These two structures were obtained by replacing Hα atoms in A**dF**FD and **dAdF**FD with a methyl (CH_3-group). The resulting structures were geometry optimized following the same protocol as discussed in the main text. The relative positions of Aib residues (in colors) to $[Gly]^1$ in G**dF**FD (in grey) are shown. G**dF**FD is shown in lighter colors in the overlaid images.

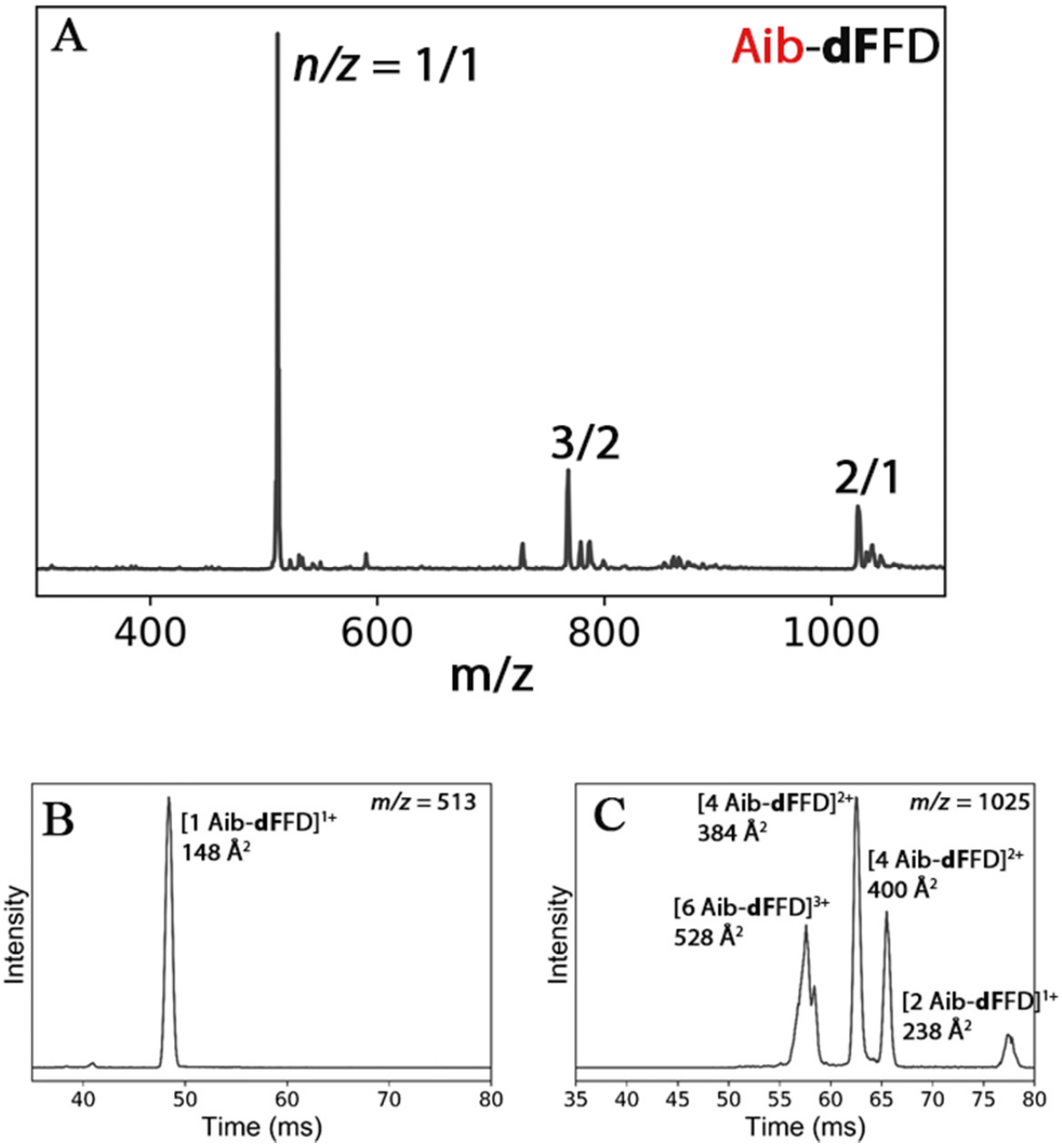


Figure S10. (A) Representative mass spectrum of Aib-**dF**FD in water. Mass spectral peaks were annotated with *n/z* where *n* the number of peptide chain and *z* is the charge. (B, C) Representative ATDs of *n/z* = 1/1 and 2/1. The CCSs are also shown. The data were obtained in positive polarity.

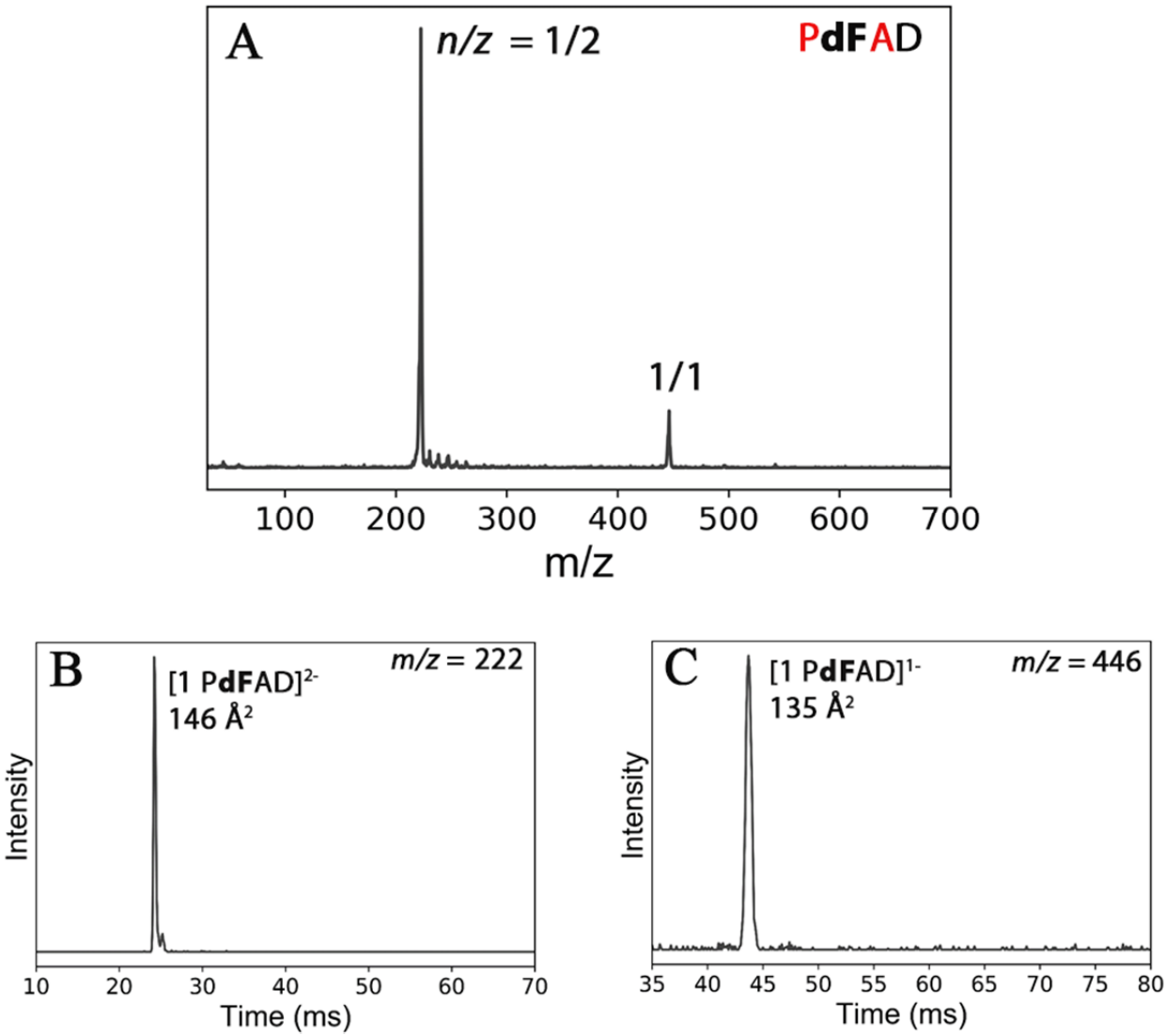


Figure S11. (A) Representative mass spectrum of P**dF**AD in water. Mass spectral peaks were annotated with *n/z* where *n* the number of peptide chain and *z* is the charge. (B, C) Representative ATDs of *n/z* = 1/1 and 1/2. The CCSs are also shown. The data were obtained in negative mode.

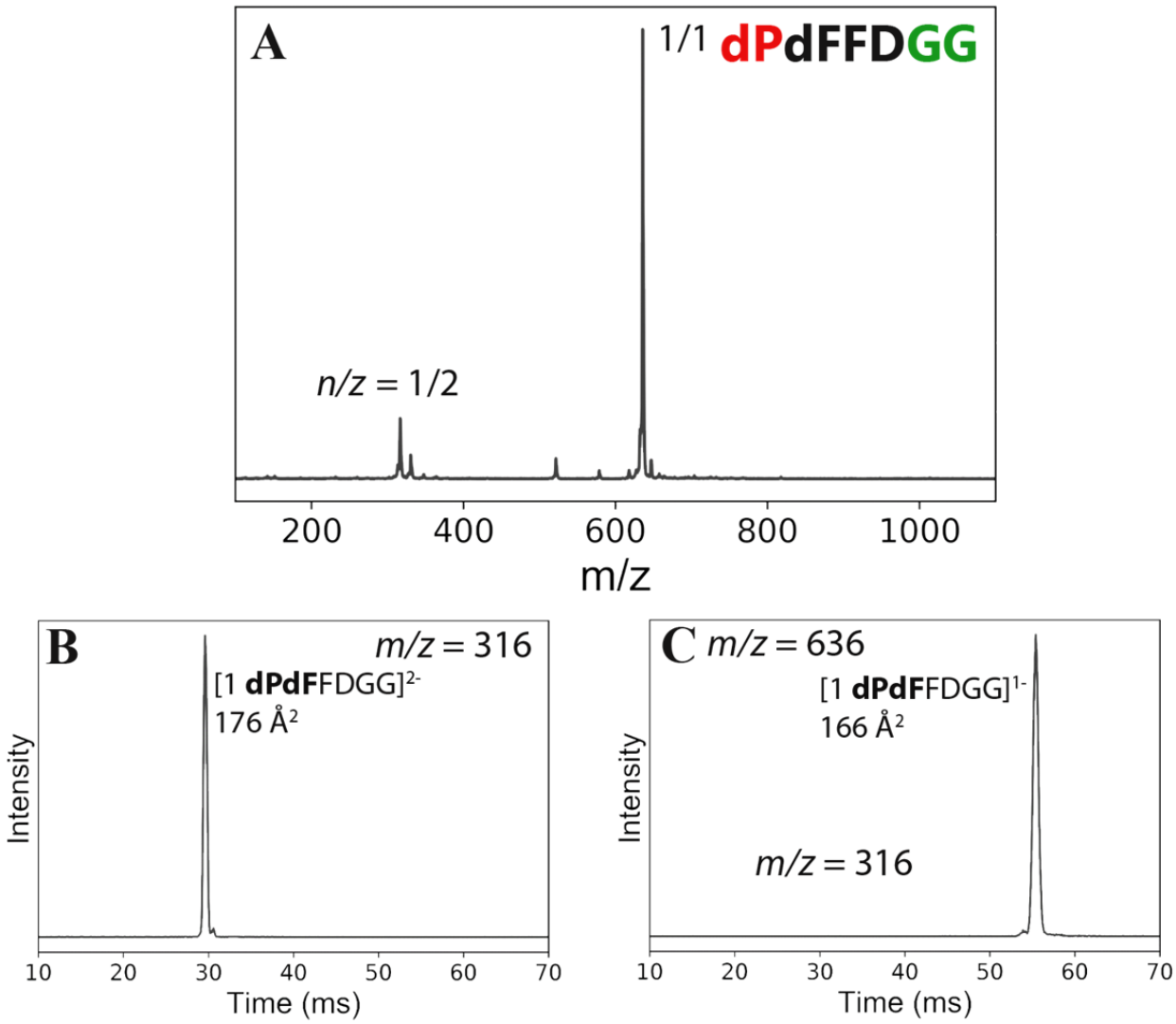
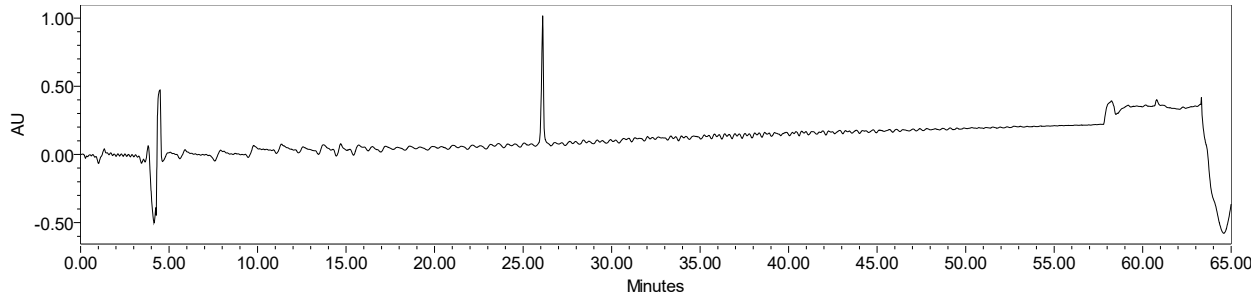
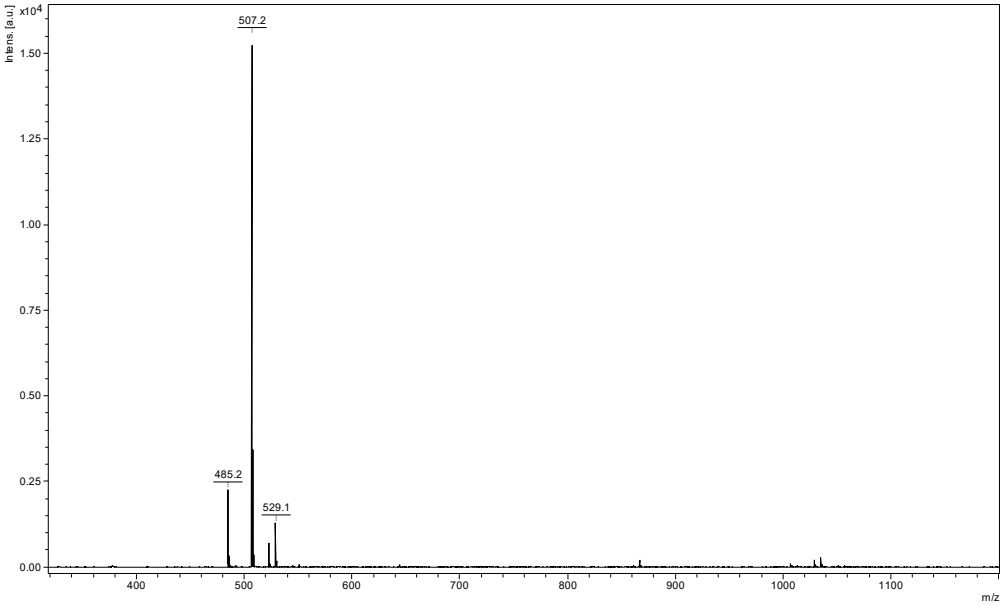


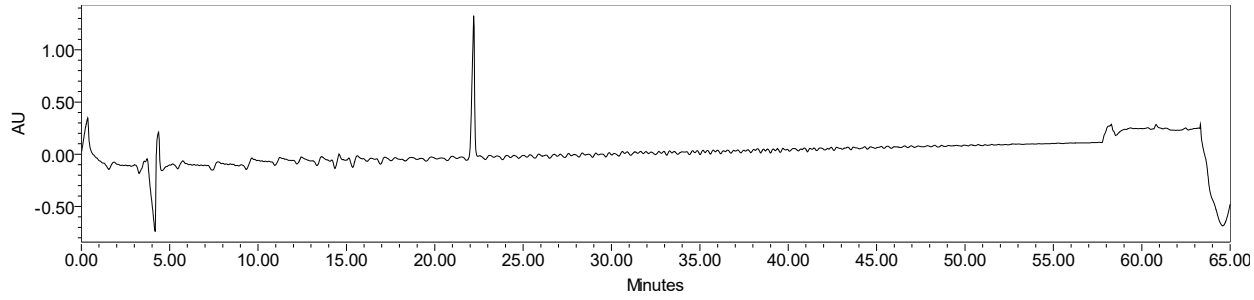
Figure S12. (A) Representative mass spectrum of **dPdFF**DGG in water. Mass spectral peaks were annotated with *n/z* where *n* the number of peptide chain and *z* is the charge. (B, C) Representative ATDs of *n/z* = 1/1 and 1/2. The CCSs are also shown. The data were obtained in negative mode.

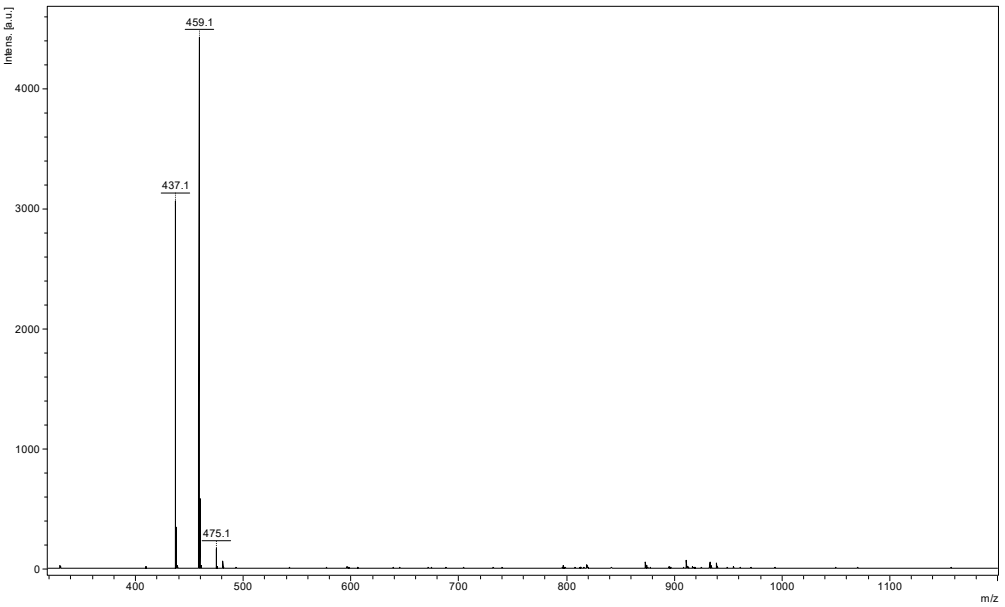
A. G**dFdF**D. Expected $[M+H]^+$ = 485.2, $[M+Na]^+$ = 507.2



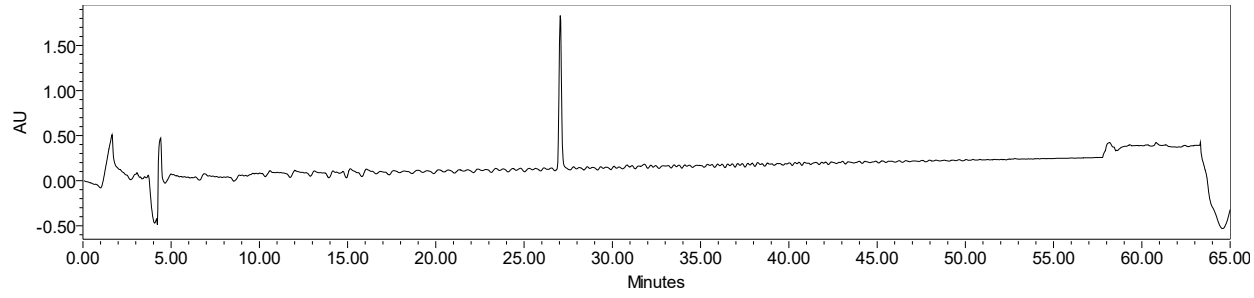


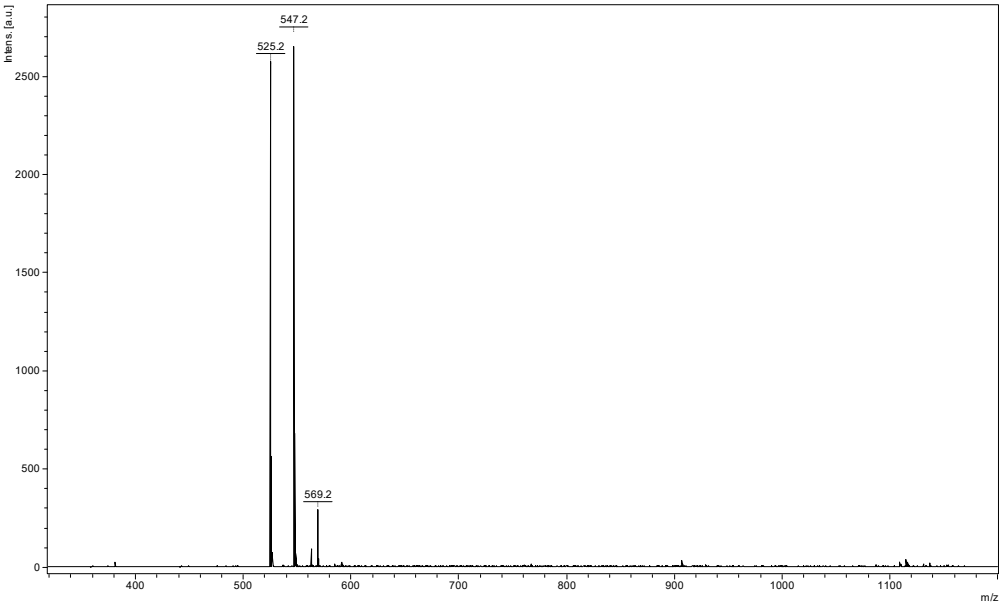
B. G**dF**VD. Expected $[M+H]^+$ = 437.2, $[M+Na]^+$ = 459.2



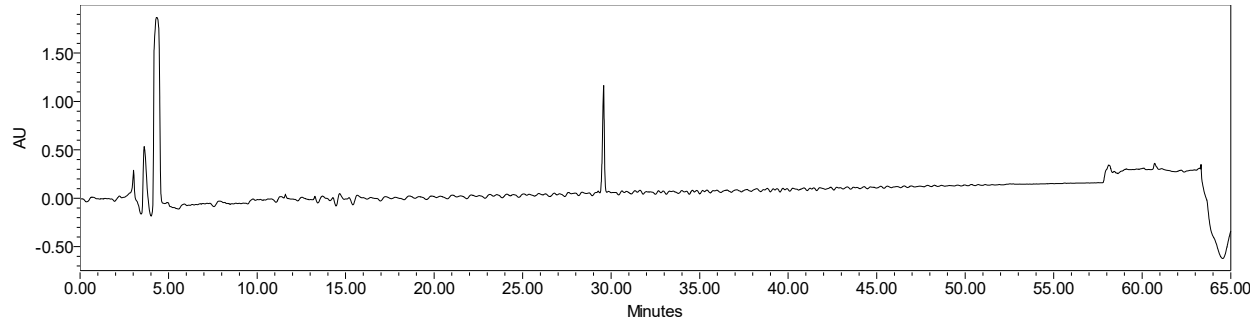


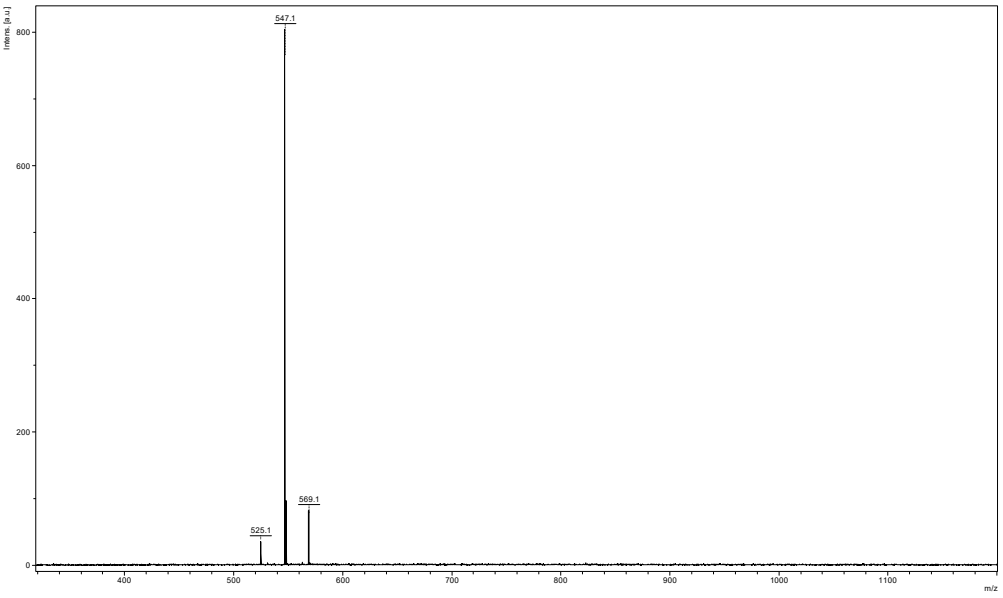
C. dPdFFD. Expected $[M+H]^+$ = 525.2, $[M+Na]^+$ = 547.2



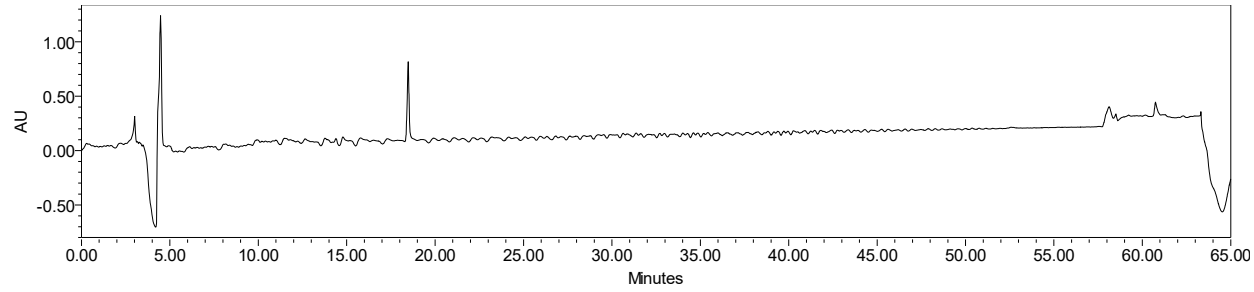


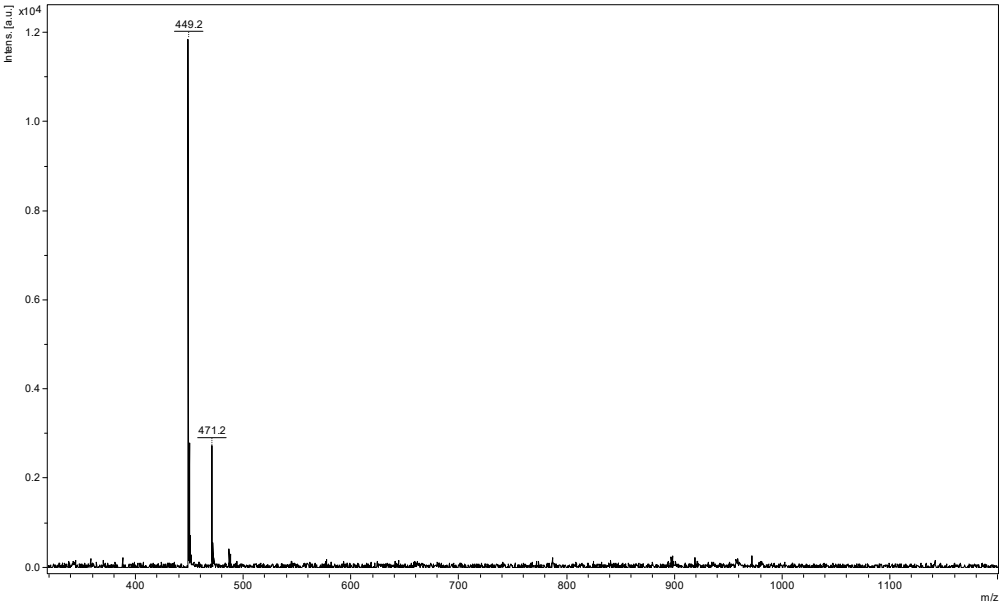
D. P**dF**FD. Expected $[M+H]^+$ = 525.2, $[M+Na]^+$ = 547.2



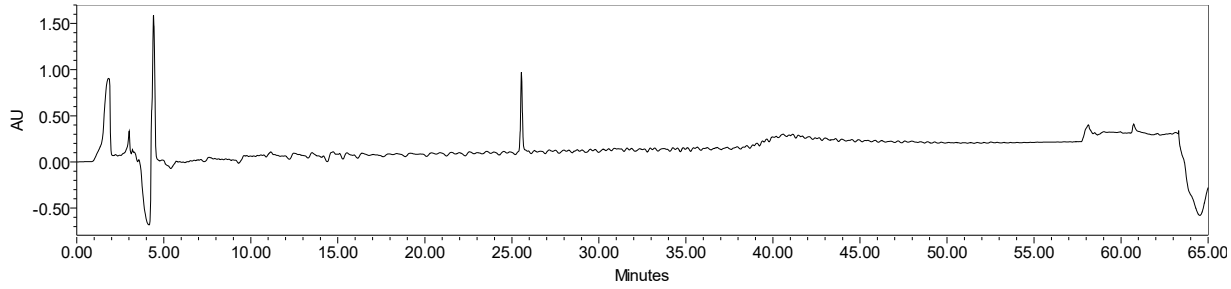


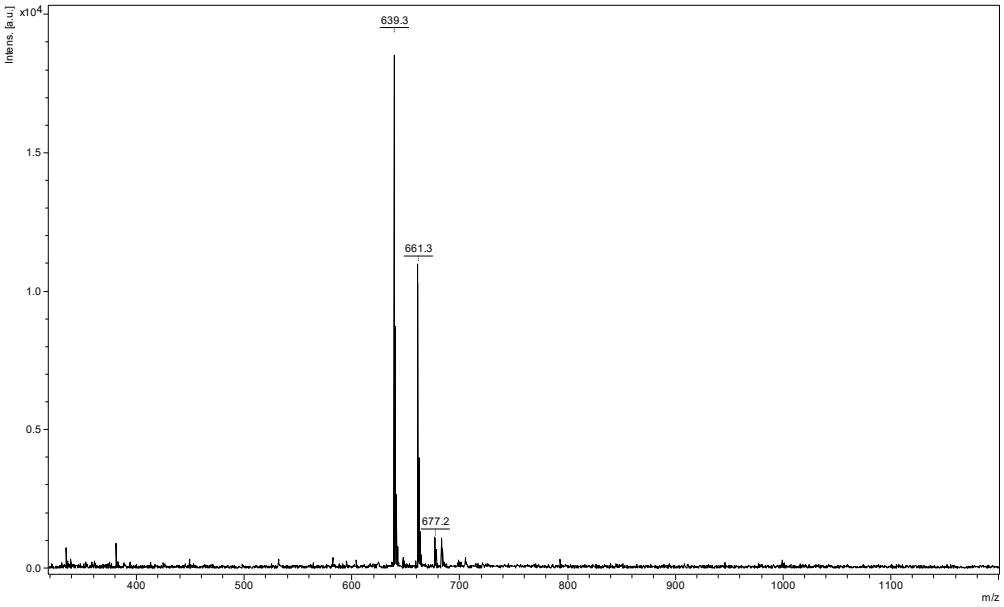
E. P**dF**AD. Expected $[M+H]^+$ = 449.2, $[M+Na]^+$ = 471.2





F. dPdFFDGG. Expected $[M+H]^+$ = 639.3, $[M+Na]^+$ = 661.3





G. Aib-**dF**FD. Expected $[M+H]^+$ = 513.2, $[M+Na]^+$ = 535.2

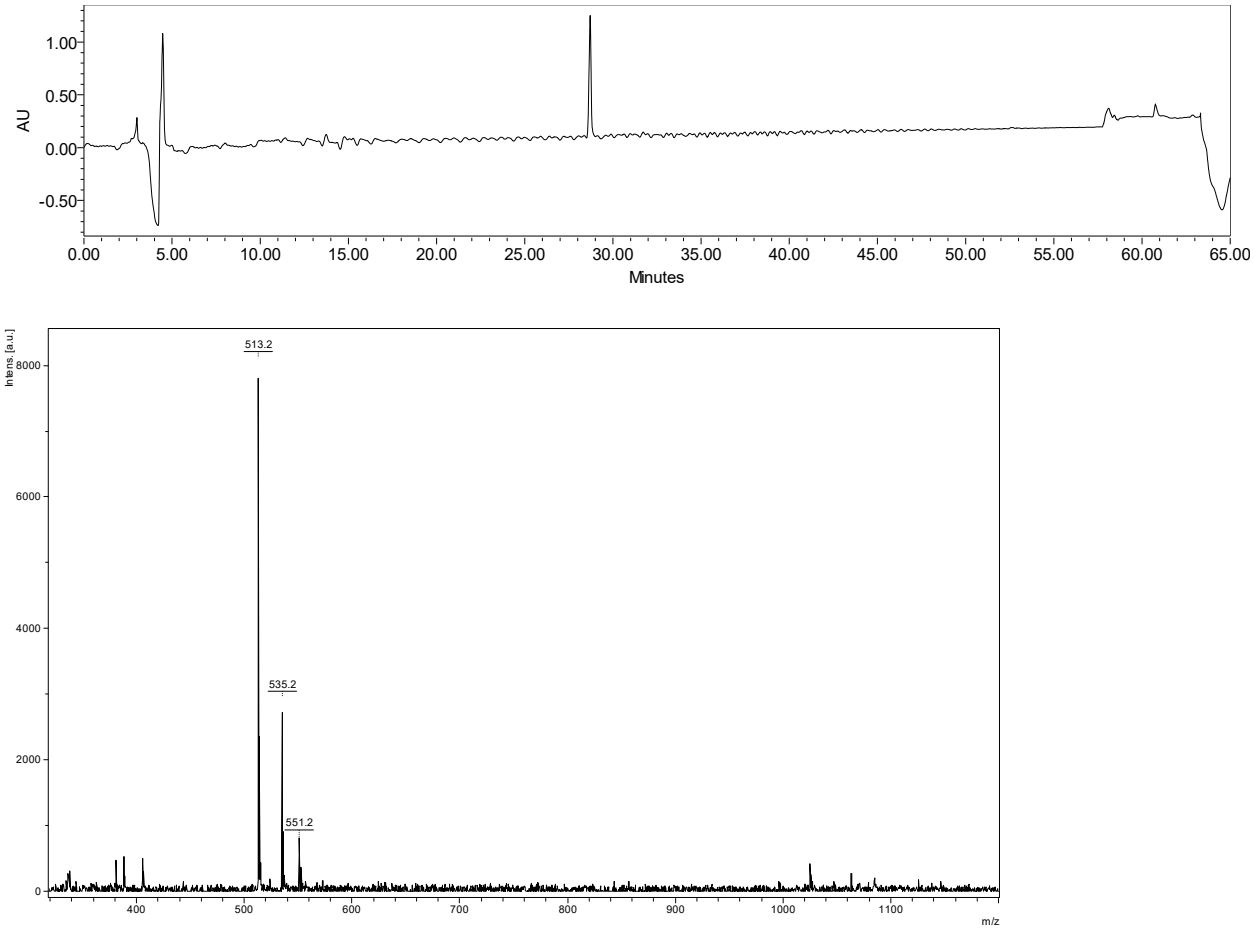


Figure S13. Peptide characterization. The MS-based characterization was performed for each new peptide synthesized and evaluated in this report. After purification by reversed-phase HPLC, peptide purity was assessed by analytical-scale reversed-phase HPLC (top panel) and peptide identity was confirmed by MALDI-TOF MS (bottom panel). Purity analysis by HPLC was monitored based on UV absorbance at 220 nm, using the following conditions: solvent A = 0.1% formic acid (FA) in H^2O, solvent B = 0.1% FA in acetonitrile, flow rate = 500 µL/min, gradient = 5–55% solvent B over 50 min, temperature = 35 °C. For MALDI-TOF MS analysis, the monoisotopic $[M+H]^+$ is labeled, along with the $[M+Na]^+$ (+23), $[M+K]^+$ (+39), or $[M+2Na-H]^+$ (+45) in some spectra.