Supplementary information

Picking the tyrosine-lock: chemical synthesis of the tyrosyl-DNA phosphodiesterase I inhibitor recifin A and analogues

Taylor B. Smallwood[a], Lauren R. H. Krumpe[b], Colton D. Payne[a], Victoria G. Klein[c,1], Barry R. O'Keefe[b,d], Richard J. Clark[a], Christina I. Schroeder[c,e] and K. Johan Rosengren[a]

[a] The University of Queensland, School of Biomedical Science, Brisbane QLD 4072, Australia
[b] Molecular Targets Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD, 21702, USA
[c] Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD, 21702, USA
[d] Natural Products Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Frederick, MD, 21702, USA
[e] Peptide Therapeutics, Genentech Inc. 1 DNA Way, South San Francisco, CA, 94080, CA, USA
[1] Current address: Laboratory of Bioorganic Chemistry, National Institute for Diabetes and Digestive Kidney Diseases, National Institutes of Health, Bethesda, MD, 20892

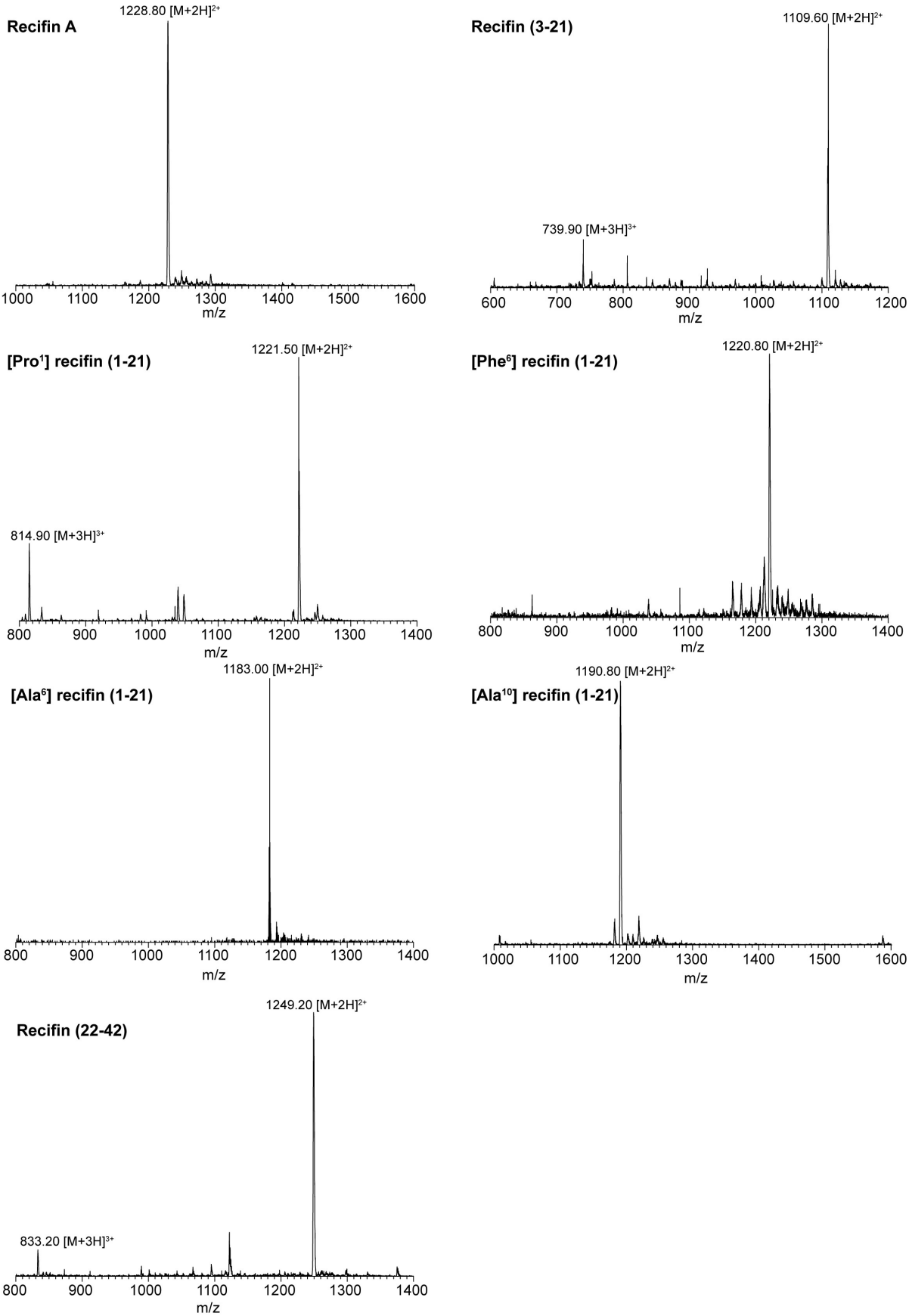

Figure S1: ESI-MS spectra of the N-terminal hydrazide fragments (fragment 1) of recifin A and its analogues and the C-terminal cysteine fragment (fragment 2) used for native chemical ligation.

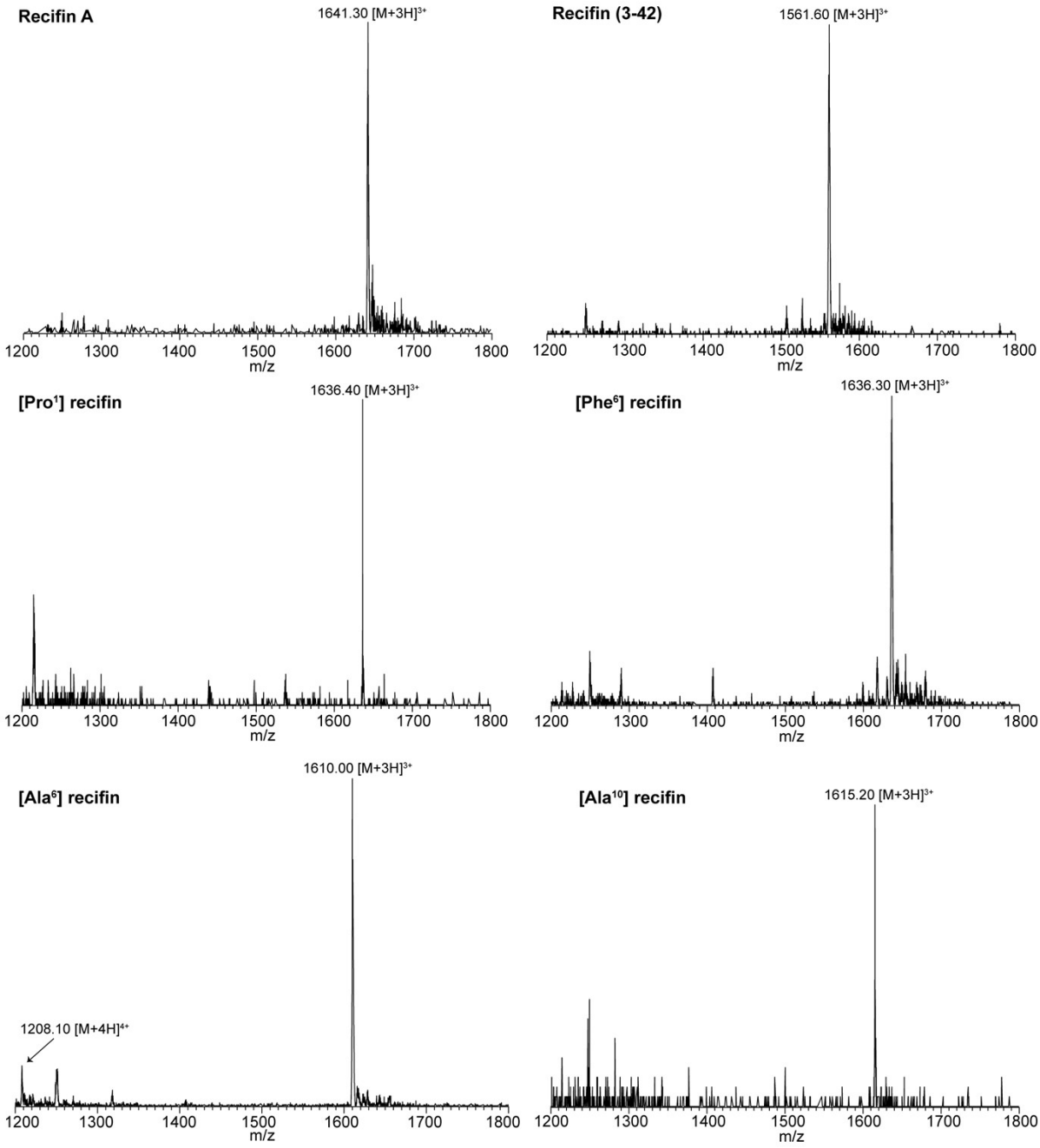

Figure S2: ESI-MS spectra of full length linear ligated recifin A and analogues.

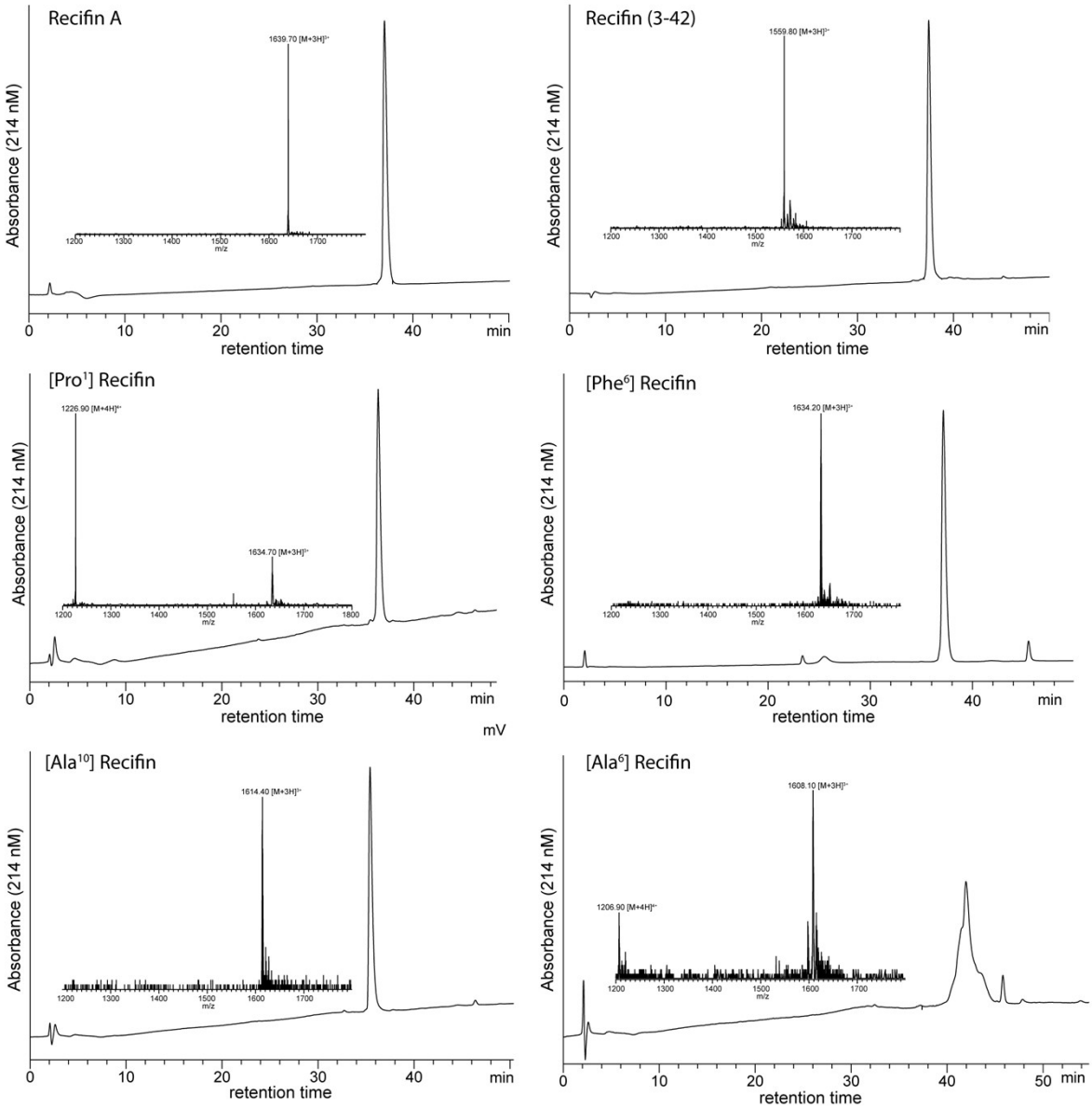

Figure S3: Final analytical traces and ESI-MS spectra of oxidized recifin A and analogues.

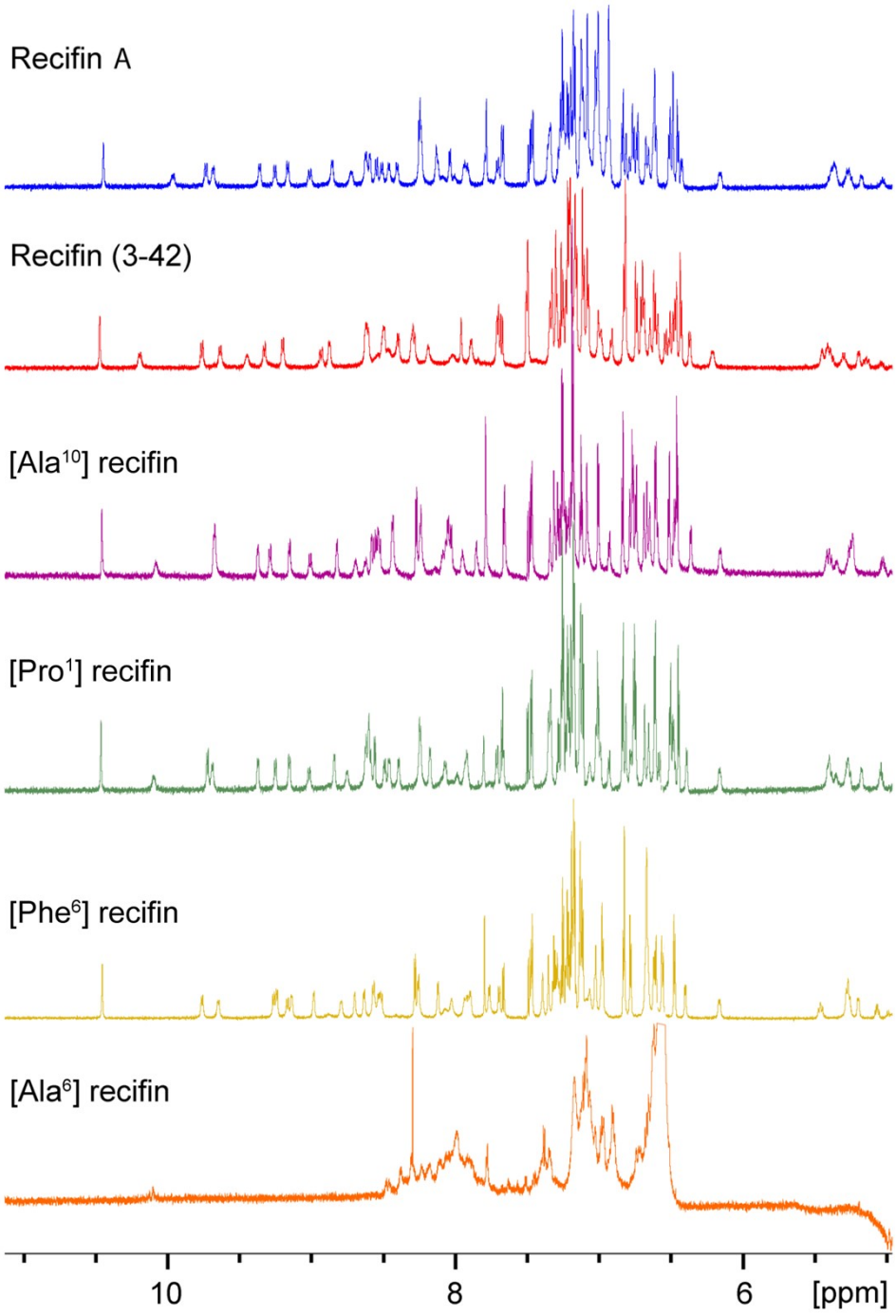

Figure S4: 1D ^{1}H Nuclear Magnetic Resonance spectra of recifin A and analogues in 90/10% H_2O/D_2O at 298 K acquired on a Bruker Avance III 900 MHz spectrometer equipped with a cryoprobe. The majority of the purified peptides gave dispersed ^{1}H NMR spectra with sharp lines, implying that they adopt ordered structures in solution. However, the [Ala^{6}] recifin analogue spectra appeared broad and lacked dispersion of the HN signals indicating that the peptide is misfolded. Thus, while substitution of Tyr6 with Phe is well tolerated, incorporating an alanine at position 6 prevents folding of the peptide.

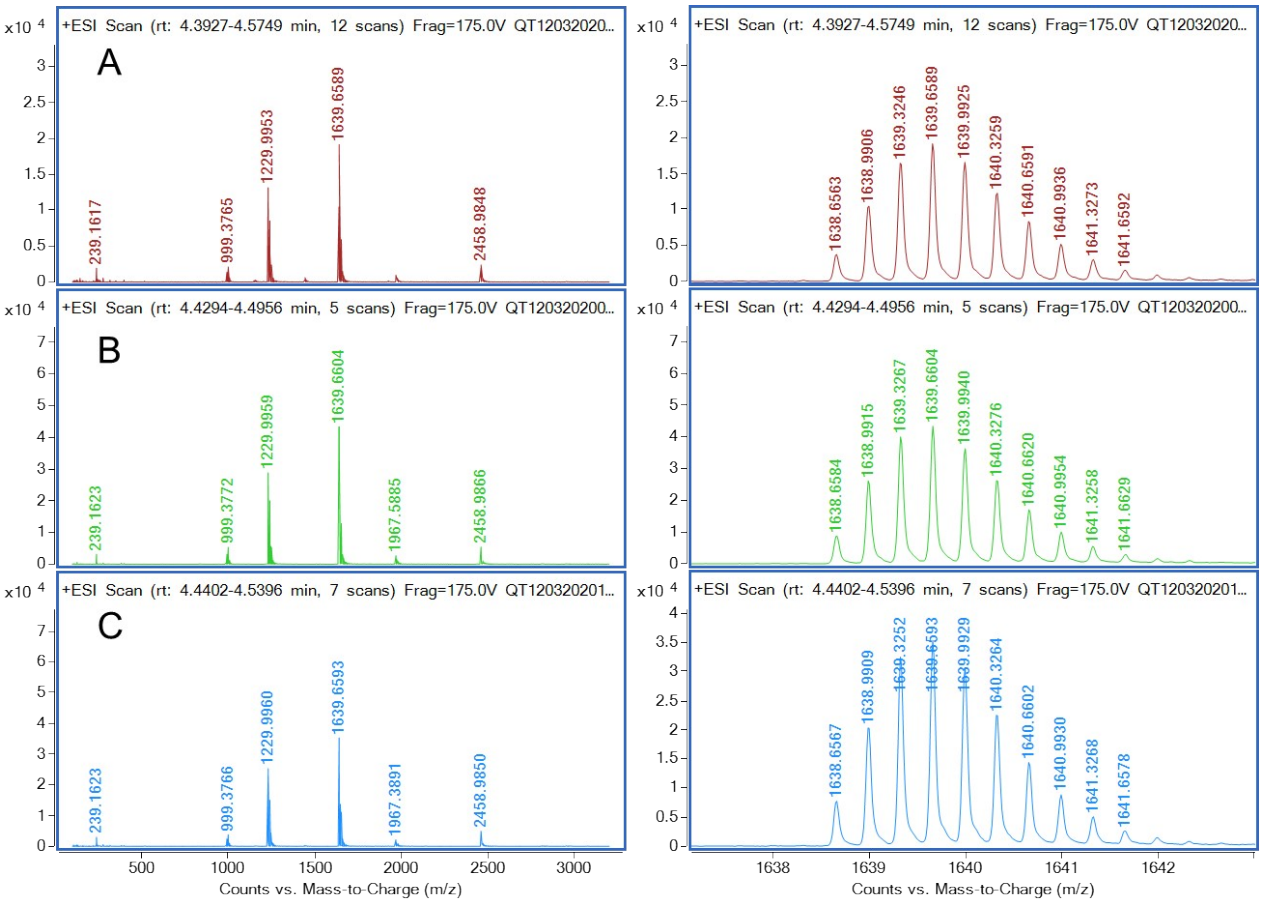

Figure S5: Molecular charge envelopes for **A)** native and **B)** synthetic recifin A and **C)** native and synthetic recifin A co-elution experiments showing identical ionization patterns and distribution of charge states with near identical isotopic distribution $[M+3H]^{3+}$.

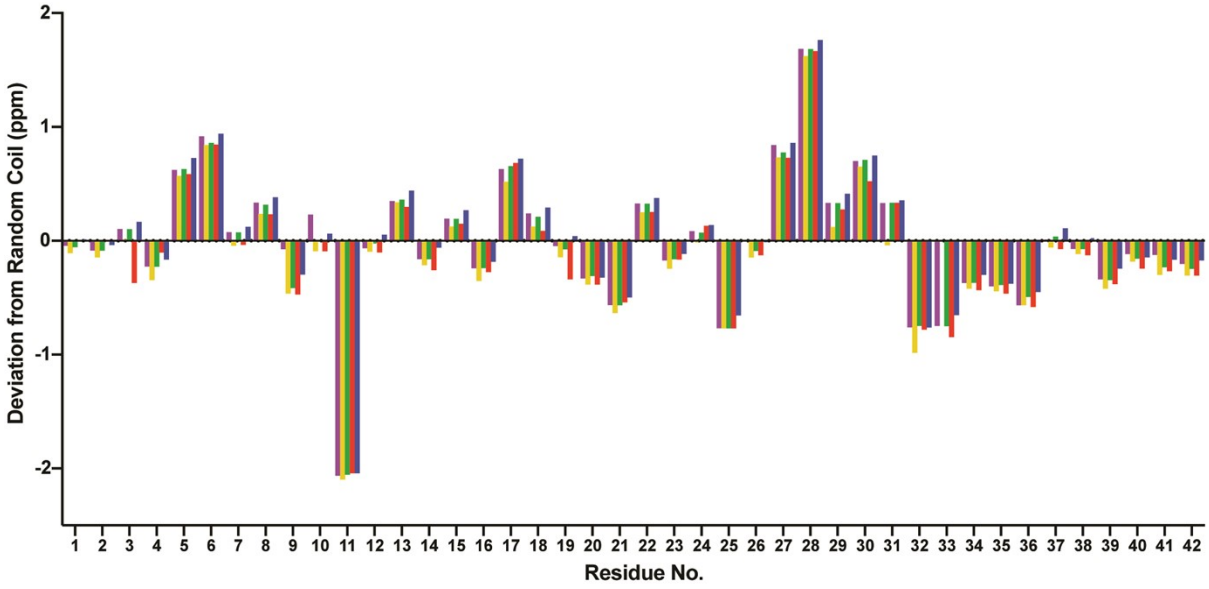

Figure S6: Secondary Hα chemical shifts compared to random coil values[15] highlighting positive stretches of secondary chemical shifts indicative of β-sheets combined with negative stretches suggesting α-helices. Recifin A in blue, recifin (3–42) in red, [Pro^{1}] recifin in purple, [Phe^{6}] recifin in green and [Ala^{10}] recifin in yellow. Secondary Hα chemical shifts for recifin (3–42) at residue number 1 and 2 was set to 0 for the truncated version.

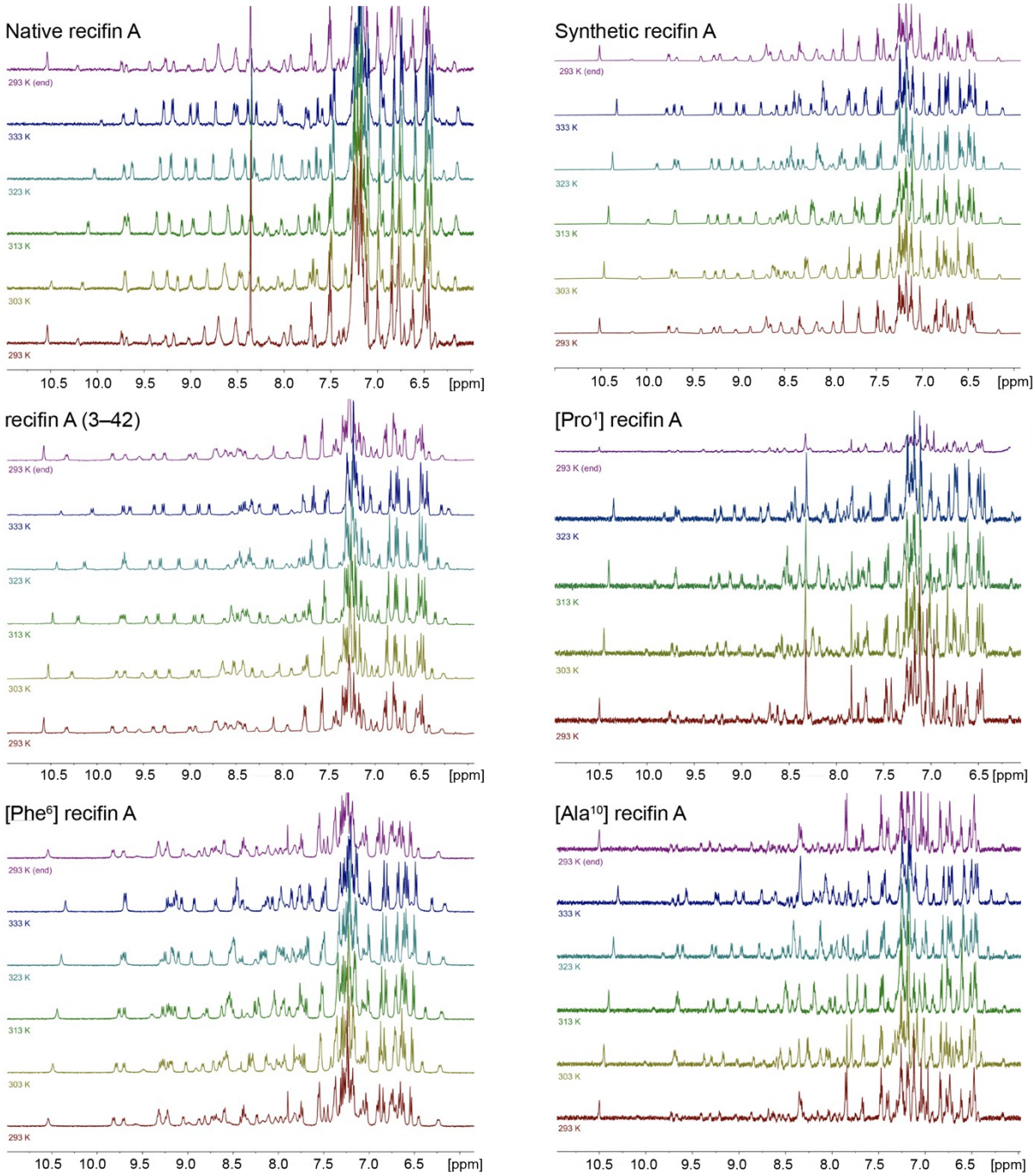

Fig S7: Thermal stability (293–333 K) of native recifin A, synthetic recifin A and synthetic recifin A analogues carried out using nuclear magnetic resonance in 90/10% H_2O/D_2O, pH5.8–7.9, on a 500 or 700 MHz Bruker Avance III equipped with a cryo probe.

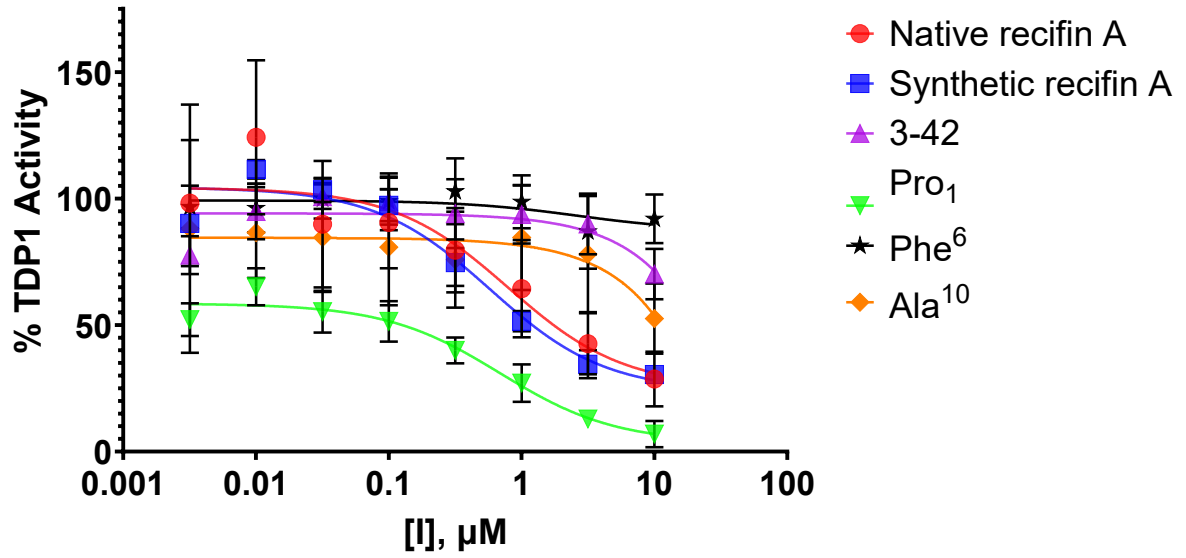

Figure S8: Biochemical activity of native recifin A, synthetic recifin A and structural analogues. Peptides were tested for inhibition of TDP1 activity in a FRET-based assay system. Reactions were conducted in triplicate, and data were normalized to no enzyme (0% activity) and non-inhibited enzyme (100% activity) assay controls. Concentration-response curves and IC_{50} values (see main text, Table 1) were produced with GraphPad Prism software.

Table S1: Expected and observed molecular mass (Da) of synthesized recifin A and analogues

Peptide	Hydrazide fragment		Ligated peptide		Oxidized peptide	
	Expected $[M+2H]^{2+}$	Observed $[M+2H]^{2+}$	Expected $[M+3H]^{3+}$	Observed $[M+3H]^{3+}$	Expected $[M+3H]^{3+}$	Observed $[M+3H]^{3+}$
Recifin A	1229.34	1228.80	1641.74	1641.30	1639.74	1639.70
Recifin (3-42)	1109.22	1109.60	1561.74	1561.60	1559.74	1559.80
[Pro^{1}] recifin	1222.32	1221.50	1637.15	1636.40	1635.15	1634.70
[Phe^{6}] recifin	1221.33	1220.80	1636.47	1636.30	1634.47	1634.20
[Ala^{6}] recifin	1183.28	1183.00	1611.11	1610.00	1609.11	1608.10
[Ala^{10}] recifin	1191.28	1190.80	1616.44	1615.20	1614.44	1613.80

Table S2: NMR distance and dihedral statistics for [Phe^6] recifin

	[Phe^6] recifin
Distance restraints	
Intra residual (|i-j| = 0)	151
Sequential (|i-j| = 1)	170
Medium range (1 < |i-j| < 5)	117
Long range (|i-j| ≥ 5)	180
Hydrogen bonds	24
Total	642
Dihedral angle restraints	
Φ	18
Ψ	15
Total	33
Energies (kcal/mol, mean ± SD)	
Overall	-1374.73 ± 36.81
Bonds	23.26 ± 0.93
Angles	65.42 ± 4.63
Improper	31.43 ± 3.52
Dihedral	189.78 ± 1.67
Van der Waals	-189.80 ± 5.84
Electrostatic	-1496.15 ± 35.99
NOE (experimental)	0.36 ± 0.04
Constrained dihedrals (experimental)	0.96 ± 0.36
Atomic RMSD (Å)	
Mean global backbone (1-42)	0.92 ± 026
Mean global heavy atoms (1-42)	1.52 ± 0.26
Mean global backbone (3-17,27-42)	0.39 ± 0.09
Mean global heavy atoms (3-17,27-42)	0.92 ± 0.10
MolProbity	
Clash Score, all atoms	11.74 ± 2.94
Poor rotamers	0 ± 0
Favored rotamers (%)	94.05 ± 3.23
Ramachandran Outliers (%)	0 ± 0
Ramachandran Favored (%)	92.05 ± 3.32
MolProbity score	2.02 ± 0.14
MolProbity score percentile	74.26 ± 7.20
Violations from experimental restraints	
NOE violations exceeding 0.2 Å	0
Dihedral violations exceeding 2.0°	0