Supporting information

The Kinetics of Amyloid Fibrillar Aggregation of Uperin 3.5 Is Directed by the Peptide's Secondary Structure

Torsten John,[†,‡,§] Tiara J. A. Dealey,[†] Nicholas P. Gray,[†] Nitin A. Patil,[∥,#] Mohammed A. Hossain,[∥] Bernd Abel,[‡,§] John A. Carver,[⊥] Yuning Hong,[∇] and Lisandra L. Martin[*,†]

†. School of Chemistry, Monash University, Clayton, Victoria 3800, Australia. E-mail: Lisa.Martin@monash.edu

‡. Leibniz Institute of Surface Engineering (IOM), Permoserstraße 15, 04318 Leipzig, Germany

§. Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University, Linnéstraße 3, 04103 Leipzig, Germany

∥. Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Parkville, Victoria 3010, Australia

⊥. Research School of Chemistry, The Australian National University, Acton, Australian Capital Territory 2601, Australia

∇. Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia

#. Current address: Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia.

Table of Contents

1 Supporting materials and methods
2 Supporting results
3 Supporting references

1 Supporting materials and methods

1.1 Peptide synthesis, purification and characterization

Table S1. Overview of uperin 3.5 variants, their source and purity determined by HPLC-MS.

Abbreviation	Sequence	Source	Purity by Analytical HPLC-MS [%]
U3.5 wt	GVGDLIRKAVSVIKNIV-NH_2	Authors & Peptide 2.0	≥ 99 ≥ 98
U3.5 G1L	LVGDLIRKAVSVIKNIV-NH_2	GenicBio	≥ 96
U3.5 D4A	GVGALIRKAVSVIKNIV-NH_2	GenicBio	≥ 90
U3.5 R7A	GVGDLIAKAVSVIKNIV-NH_2	Authors & Peptide 2.0	≥ 90 ≥ 98
U3.5 G1L-D4A	LVGALIRKAVSVIKNIV-NH_2	GenicBio	≥ 93
U3.5 G1L-R7A	LVGDLIAKAVSVIKNIV-NH_2	GenicBio	≥ 97
U3.5 G1L-D4A-R7A	LVGALIAKAVSVIKNIV-NH_2	Authors	≥ 99

The purity values of the peptides were determined from the integrated HPLC peak areas.

1.2 Secondary structure prediction models

The secondary structure of the U3.5 peptide variants was predicted based on their sequences using multiple algorithms on webservers. The Scratch Protein Predictor with SSpro server is based on sequence homology.[1–3] PredictProtein[4] provides access to the PROFsec algorithm.[5] The Network Protein Sequence analysis (NPS@)[6] provides access to the PHD algorithm.[5,7] The PSIPRED[8,9] and PSSpred (Protein Secondary Structure prediction)[10] are neural network algorithms.

2 Supporting results

2.1 Thioflavin T fluorescence experiments

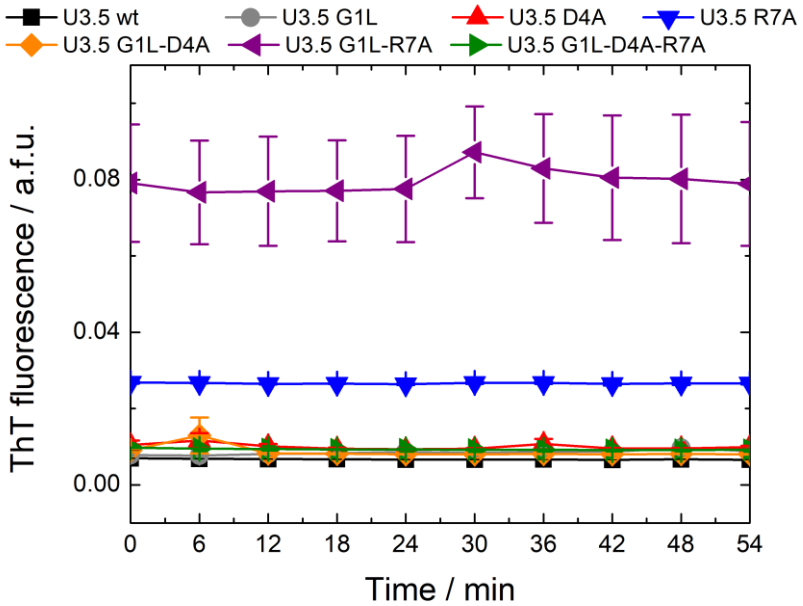

Figure S1. ThT fluorescence data for U3.5 variants at 111 µM (100 µM after buffer addition) during the initial 54 minutes in ultrapure water. Only the peptide variants U3.5 R7A and U3.5 G1L-R7A present some fluorescence.

Table S2. Kinetic analysis of ThT fluorescence assay. For all U3.5 variants, the net charge (with partial charges in parentheses), time at half-maximum fluorescence ($t_{1/2}$), aggregation rate at $t_{1/2}$ (k) and maximum reached fluorescence ($I_{Fluorescence}$) are listed. The aggregation rate and amount of formed fibrils for U3.5 G1L is negligibly small.

U3.5 variant	Net charge, pH 7	$t_{1/2}$ [h]	k at $t_{1/2}$ [h^{-1}]	$I_{Fluorescence}$ [a.f.u.]
U3.5 G1L-R7A	+2 (+3,-1)	1.0	2.10	1.00
U3.5 R7A	+2 (+3,-1)	1.1	1.08	0.58
U3.5 G1L-D4A-R7A	+3 (+3)	1.6	0.21	0.39
U3.5 wt	+3 (+4,-1)	5.3	0.11	0.31
U3.5 G1L	+3 (+4,-1)	-	-	0.05
U3.5 D4A	+4 (+4)	5.2	0.05	0.28
U3.5 G1L-D4A	+4 (+4)	2.9	0.02	0.15

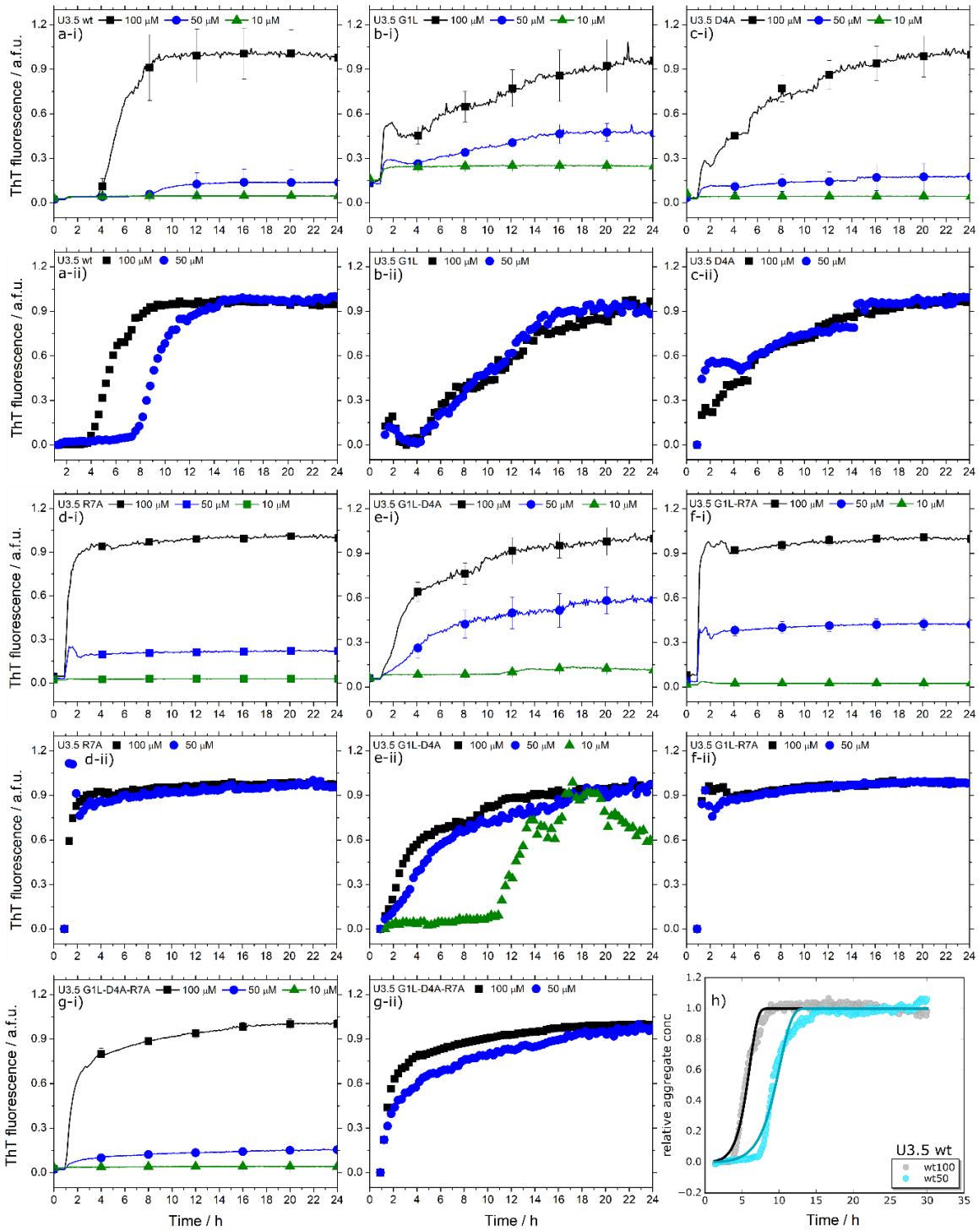

Figure S2. ThT fluorescence data for U3.5 variants at 100 µM (black squares), 50 µM (blue circles) and 10 µM (green triangles) with buffer addition after 1 hour at 25°C. In i), the data represent the fluorescence intensities with the highest fluorescence at 100 µM set to 1. In ii), only concentrations at which the U3.5 variant aggregated are shown, and the fluorescence intensities for each concentration are (0-1) normalized. While U3.5 wt starts to aggregate after a lag time of more than three hours after buffer addition, all other U3.5 variants at 50 and 100 µM rapidly started to form amyloid fibrils without apparent lag time. Only the U3.5 G1L-D4A variant at 10 µM formed fibrils after a lag time of around ten hours. The ThT fluorescence data for the U3.5 variants were fitted to existing fibrillar aggregation models using the AmyloFit online tool.[11] h) For U3.5 wt, a fragmentation and secondary nucleation dominated mechanism was identified. The other U3.5 peptides likely follow the same mechanism, but are seeded with fibril nuclei that formed in water prior to buffer addition.

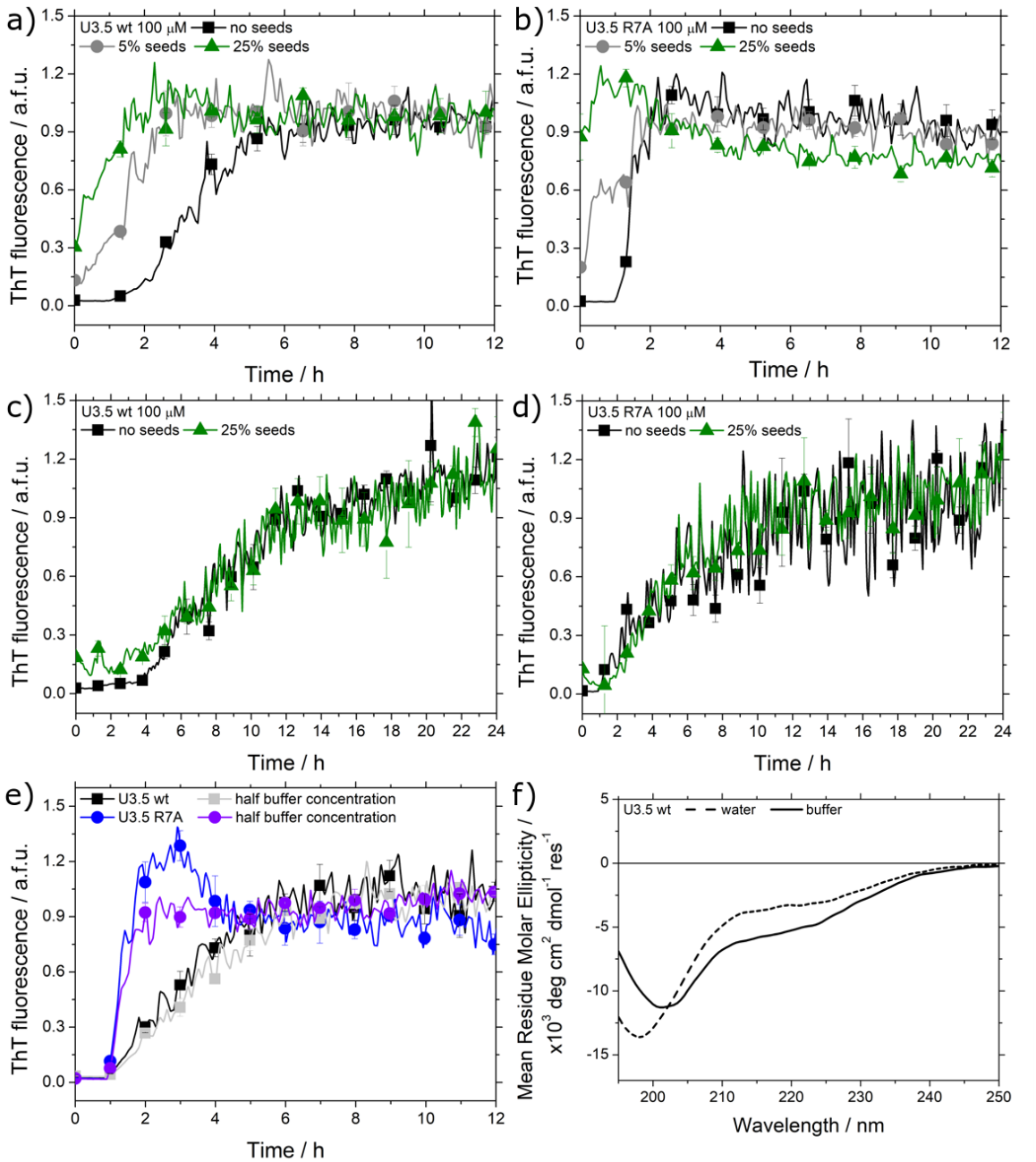

Figure S3. (a-e) ThT fluorescence data for U3.5 wt and U3.5 R7A at 100 µM with buffer addition after 1 hour (final fluorescence set to 1). (a-b) Seed experiments at 37°C with 0% (black squares), 5% (grey circles) or 25% (green triangles) seeds added at the start. The seed stock solutions contained buffer and thus initiated the aggregation process. (c-d) Seed experiments at 25°C with seeds that are free of buffer by prior centrifugation and subsequent resuspension of the aggregates into ultrapure water. The mature fibrils seem to not accelerate peptide aggregation. (e) Experiments at 25°C with varying buffer concentration. Half of the original PBS buffer concentration was sufficient to initiate peptide aggregation. U3.5 wt also aggregated without lag time as U3.5 R7A does. This is likely due to the formation of nuclei in water prior to buffer addition. (f) Samples that were prepared from older (repeatedly) frozen or higher concentrated stock solutions had higher α-helical content in water and/or upon addition of buffer. The CD spectra were recorded at 111 µM (in water) / 100 µM (upon buffer addition).

2.2 Circular dichroism experiments and peptide secondary structure estimation

Table S3. Summary of quantitative secondary structure estimation for the **U3.5 wt peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 wt in water					
BeStSel	0.02	0.34	0.16	0.48	0.011
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.10	0.32		0.58	0.32
DICROWEB K2D	0.05	0.38		0.57	0.119
DICROWEB CDSSTR Ref 7	0.02	0.13	0.09	0.76	0.011
DICROWEB CDSSTR Ref 4	0.08	0.31	0.25	0.36	0.010
DICROWEB CONTIN-LL Ref 7	0.04	0.11	0.07	0.78	0.048
DICROWEB CONTIN-LL Ref 4	0.08	0.28	0.23	0.41	0.113
DICROWEB SELCON3 Ref 7	0.04	0.06	0.05	0.82	0.062
DICROWEB SELCON3 Ref 4	poor	fit			
U3.5 wt in buffer					
BeStSel	0.02	0.24	0.18	0.56	0.016
CAPITO basis spectra	0.07	0.28		0.71	0.18
CAPITO lowest area difference	0.10	0.32		0.58	0.22
DICROWEB K2D	0.06	0.37		0.58	0.106
U3.5 wt in buffer, 1d					
BeStSel	0.05	0.34	0.10	0.51	0.010
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.13	0.34		0.53	0.34
DICROWEB K2D	poor	fit			
U3.5 wt in buffer, 5d					
BeStSel	0.04	0.25	0.16	0.60	0.012
CAPITO basis spectra	0.30	0.29		0.55	0.36
CAPITO lowest area difference	0.07	0.37		0.56	0.33
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S4. Summary of quantitative secondary structure estimation for the **U3.5 G1L peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 G1L in water					
BeStSel	0.04	0.32	0.16	0.48	0.017
CAPITO basis spectra	0.02	0.25		0.70	0.21
CAPITO lowest area difference	0.10	0.32		0.58	0.28
DICROWEB K2D	0.06	0.37		0.57	0.108
DICROWEB CDSSTR Ref 7	0.04	0.12	0.09	0.76	0.006
DICROWEB CDSSTR Ref 4	0.08	0.31	0.26	0.35	0.008
DICROWEB CONTIN-LL Ref 7	0.04	0.11	0.07	0.78	0.042
DICROWEB CONTIN-LL Ref 4	0.09	0.29	0.23	0.39	0.114
DICROWEB SELCON3 Ref 7	0.04	0.07	0.06	0.83	0.052
DICROWEB SELCON3 Ref 4	poor	fit			
U3.5 G1L in buffer					
BeStSel	0.03	0.24	0.15	0.58	0.014
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.10	0.32		0.58	0.32
DICROWEB K2D	0.08	0.35		0.57	0.142
U3.5 G1L in buffer, 1d					
BeStSel	0.10	0.20	0.09	0.61	0.015
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.21	0.26		0.53	0.23
DICROWEB K2D	0.28	0.32		0.40	0.159
U3.5 G1L in buffer, 5d					
BeStSel	0.04	0.21	0.11	0.64	0.016
CAPITO basis spectra	0.21	0.16		0.64	0.22
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	0.27	0.26		0.47	0.107

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S5. Summary of quantitative secondary structure estimation for the **U3.5 D4A peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 D4A in water					
BeStSel	0.04	0.29	0.17	0.50	0.016
CAPITO basis spectra	0	0.31		0.72	0.20
CAPITO lowest area difference	0.10	0.32		0.58	0.34
DICROWEB K2D	0.06	0.37		0.57	0.118
DICROWEB CDSSTR Ref 7	0.03	0.12	0.09	0.76	0.006
DICROWEB CDSSTR Ref 4	0.09	0.29	0.25	0.37	0.007
DICROWEB CONTIN-LL Ref 7	0.04	0.09	0.07	0.80	0.057
DICROWEB CONTIN-LL Ref 4	0.09	0.27	0.23	0.41	0.123
DICROWEB SELCON3 Ref 7	poor	fit			
DICROWEB SELCON3 Ref 4	poor	fit			
U3.5 D4A in buffer					
BeStSel	0.04	0.24	0.16	0.56	0.013
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.13	0.35		0.51	0.26
DICROWEB K2D	0.09	0.36		0.55	0.100
U3.5 D4A in buffer, 1d					
BeStSel	0.04	0.26	0.13	0.57	0.018
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.13	0.30		0.57	0.27
DICROWEB K2D	poor	fit			
U3.5 D4A in buffer, 5d					
BeStSel	0.03	0.29	0.13	0.55	0.017
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.09	0.41		0.50	0.40
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S6. Summary of quantitative secondary structure estimation for the **U3.5 R7A peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 R7A in water					
BeStSel	0.06	0.32	0.15	0.47	0.014
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.10	0.32		0.58	0.27
DICROWEB K2D	0.08	0.41		0.50	0.143
DICROWEB CDSSTR Ref 7	0.05	0.19	0.13	0.62	0.014
DICROWEB CDSSTR Ref 4	0.09	0.31	0.25	0.35	0.017
DICROWEB CONTIN-LL Ref 7	0.07	0.19	0.13	0.61	0.053
DICROWEB CONTIN-LL Ref 4	0.09	0.31	0.23	0.37	0.046
DICROWEB SELCON3 Ref 7	0.06	0.13	0.12	0.67	0.060
DICROWEB SELCON3 Ref 4	0.12	0.25	0.22	0.41	0.061
U3.5 R7A in buffer					
BeStSel	0.28	0	0.12	0.60	0.010
CAPITO basis spectra	0.39	0.25		0.41	0.27
CAPITO lowest area difference	0.39	0.14		0.46	0.14
DICROWEB K2D	poor	fit			
U3.5 R7A in buffer, 1d					
BeStSel	0.06	0.16	0.11	0.67	0.024
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.07	0.37		0.56	0.26
DICROWEB K2D	poor	fit			
U3.5 R7A in buffer, 7d					
BeStSel	0.03	0.29	0.13	0.55	0.022
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.09	0.41		0.50	0.47
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S7. Summary of quantitative secondary structure estimation for the **U3.5 G1L-D4A peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 G1L-D4A in water					
BeStSel	0.04	0.31	0.16	0.49	0.019
CAPITO basis spectra	0.04	0.26		0.73	0.17
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	0.06	0.37		0.57	0.123
DICROWEB CDSSTR Ref 7	0.03	0.12	0.09	0.76	0.007
DICROWEB CDSSTR Ref 4	0.09	0.29	0.25	0.36	0.010
DICROWEB CONTIN-LL Ref 7	0.04	0.09	0.07	0.80	0.045
DICROWEB CONTIN-LL Ref 4	0.09	0.27	0.23	0.41	0.115
DICROWEB SELCON3 Ref 7	0.04	0.10	0.08	0.77	0.139
DICROWEB SELCON3 Ref 4	poor	fit			
U3.5 G1L-D4A in buffer					
BeStSel	0.04	0.27	0.13	0.56	0.014
CAPITO basis spectra	0.20	0.21		0.68	0.20
CAPITO lowest area difference	0.12	0.35		0.54	0.14
DICROWEB K2D	0.18	0.26		0.56	0.077
U3.5 G1L-D4A in buffer, 1d					
BeStSel	0.06	0.17	0.09	0.68	0.015
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	poor	fit			
U3.5 G1L-D4A in buffer, 5d					
BeStSel	0.08	0.11	0.11	0.70	0.018
CAPITO basis spectra	0.27	0.17		0.55	0.14
CAPITO lowest area difference	0.26	0.28		0.46	0.15
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S8. Summary of quantitative secondary structure estimation for the **U3.5 G1L-R7A peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 G1L-R7A in water					
BeStSel	0.08	0.29	0.14	0.49	0.018
CAPITO basis spectra	0.05	0.31		0.61	0.33
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	0.09	0.36		0.55	0.199
DICROWEB CDSSTR Ref 7	0.08	0.22	0.16	0.54	0.015
DICROWEB CDSSTR Ref 4	0.11	0.31	0.24	0.33	0.024
DICROWEB CONTIN-LL Ref 7	0.03	0.18	0.16	0.63	0.082
DICROWEB CONTIN-LL Ref 4	0.13	0.31	0.23	0.33	0.048
DICROWEB SELCON3 Ref 7	0.11	0.20	0.12	0.54	0.115
DICROWEB SELCON3 Ref 4	0.15	0.28	0.21	0.35	0.155
U3.5 G1L-R7A in buffer					
BeStSel	0.19	0.08	0.11	0.62	0.013
CAPITO basis spectra	0.15	0.37		0.42	0.57
CAPITO lowest area difference	0.26	0.28		0.46	0.18
DICROWEB K2D	poor	fit			
U3.5 G1L-R7A in buffer, 1d					
BeStSel	0.06	0.15	0.11	0.68	0.017
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.27	0.24		0.49	0.19
DICROWEB K2D	poor	fit			
U3.5 G1L-R7A in buffer, 5d					
BeStSel	0.14	0.11	0.10	0.65	0.018
CAPITO basis spectra	0.14	0.44		0.49	0.38
CAPITO lowest area difference	0.12	0.36		0.53	0.38
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S9. Summary of quantitative secondary structure estimation for the **U3.5 G1L-D4A-R7A peptide** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 G1L-D4A-R7A in water					
BeStSel	0.11	0.24	0.16	0.49	0.016
CAPITO basis spectra	0.13	0.13		0.78	0.15
CAPITO lowest area difference	0.13	0.18		0.69	0.22
DICROWEB K2D	0.09	0.23		0.68	0.165
DICROWEB CDSSTR Ref 7	0.06	0.07	0.07	0.80	0.012
DICROWEB CDSSTR Ref 4	0.10	0.27	0.24	0.37	0.010
DICROWEB CONTIN-LL Ref 7	0.06	0.07	0.05	0.82	0.014
DICROWEB CONTIN-LL Ref 4	0.11	0.24	0.25	0.39	0.079
DICROWEB SELCON3 Ref 7	poor	fit			
DICROWEB SELCON3 Ref 4	poor	fit			
U3.5 G1L-D4A-R7A in buffer					
BeStSel	0.27	0.16	0.05	0.52	0.014
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.33	0.19		0.48	0.39
DICROWEB K2D	poor	fit			
U3.5 G1L-D4A-R7A in buffer, 1d					
BeStSel	0.01	0.31	0.06	0.62	0.023
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	poor	fit			
U3.5 G1L-D4A-R7A in buffer, 5d					
BeStSel	0.03	0.26	0.08	0.63	0.015
CAPITO basis spectra	0.39	0.25		0.55	0.42
CAPITO lowest area difference	0.38	0.13		0.48	0.20
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S10. Summary of quantitative secondary structure estimation for the **U3.5 R7A peptide and influence of shaking** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 R7A in water					
BeStSel	0.06	0.32	0.15	0.47	0.014
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.10	0.32		0.58	0.27
DICROWEB K2D	0.08	0.41		0.50	0.143
DICROWEB CDSSTR Ref 7	0.05	0.19	0.13	0.62	0.014
DICROWEB CDSSTR Ref 4	0.09	0.31	0.25	0.35	0.017
DICROWEB CONTIN-LL Ref 7	0.07	0.19	0.13	0.61	0.053
DICROWEB CONTIN-LL Ref 4	0.09	0.31	0.23	0.37	0.046
DICROWEB SELCON3 Ref 7	0.06	0.13	0.12	0.67	0.060
DICROWEB SELCON3 Ref 4	0.12	0.25	0.22	0.41	0.061
U3.5 R7A in buffer					
BeStSel	0.23	0.02	0.13	0.62	0.013
CAPITO basis spectra	0.20	0.23		0.46	0.31
CAPITO lowest area difference	0.31	0.21		0.48	0.17
DICROWEB K2D	poor	fit			
U3.5 R7A in buffer, 60 min, idle					
BeStSel	0.19	0.03	0.11	0.67	0.011
CAPITO basis spectra	0.28	0.26		0.45	0.17
CAPITO lowest area difference	0.29	0.28		0.43	0.17
DICROWEB K2D	poor	fit			
U3.5 R7A in buffer, 4 min, shaken					
BeStSel	0.06	0.14	0.15	0.65	0.033
CAPITO basis spectra	0.14	0.34		0.55	0.19
CAPITO lowest area difference	0.09	0.41		0.50	0.24
DICROWEB K2D	0.18	0.27		0.55	0.274
U3.5 R7A in buffer, 242 min, shaken					
BeStSel	0.05	0.13	0.15	0.67	0.037
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.09	0.41		0.50	0.26
DICROWEB K2D	0.20	0.25		0.55	0.236

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S11. Summary of quantitative secondary structure estimation for the **U3.5 wt peptide in 8 % (v/v) TFE** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 wt in 8 % TFE/water					
BeStSel	0.16	0.21	0.13	0.50	0.018
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	0.28	0.33		0.39	0.127
DICROWEB CDSSTR Ref 7	0.23	0.14	0.15	0.49	0.021
DICROWEB CDSSTR Ref 4	0.21	0.21	0.23	0.35	0.015
DICROWEB CONTIN-LL Ref 7	0.06	0.11	0.03	0.81	0.066
DICROWEB CONTIN-LL Ref 4	0.17	0.22	0.29	0.32	0.059
DICROWEB SELCON3 Ref 7	poor	fit			
DICROWEB SELCON3 Ref 4	poor	fit			
U3.5 wt in 8 % TFE/buffer					
BeStSel	0.44	0	0.07	0.49	0.018
CAPITO basis spectra	0.39	0.01		0.49	0.53
CAPITO lowest area difference	0.49	0.15		0.35	0.23
DICROWEB K2D	poor	fit			
U3.5 wt in 8 % TFE/buffer, 1d					
BeStSel	0.03	0.28	0.13	0.56	0.021
CAPITO basis spectra	0.10	0.39		0.53	0.21
CAPITO lowest area difference	0.14	0.41		0.34	0.25
DICROWEB K2D	poor	fit			
U3.5 wt in 8 % TFE/buffer, 5d					
BeStSel	0.03	0.30	0.14	0.53	0.017
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S12. Summary of quantitative secondary structure estimation for the **U3.5 wt peptide in 20 % (v/v) TFE** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 wt in 20 % TFE/water					
BeStSel	0.27	0.37	0.12	0.24	0.011
CAPITO basis spectra	0.26	0.22		0.41	0.46
CAPITO lowest area difference	0.39	0.14		0.46	0.20
DICROWEB K2D	0.42	0.16		0.42	0.146
DICROWEB CDSSTR Ref 7	0.41	0.21	0.15	0.23	0.019
DICROWEB CDSSTR Ref 4	0.32	0.24	0.17	0.28	0.028
DICROWEB CONTIN-LL Ref 7	0.56	0.01	0	0.43	0.081
DICROWEB CONTIN-LL Ref 4	0.57	0.01	0	0.42	0.081
DICROWEB SELCON3 Ref 7	0.40	0.16	0.15	0.27	0.139
DICROWEB SELCON3 Ref 4	0.40	0.16	0.17	0.28	0.153
U3.5 wt in 20 % TFE/buffer					
BeStSel	0.30	0.33	0.19	0.18	0.010
CAPITO basis spectra	0.22	0.35		0.38	0.28
CAPITO lowest area difference	0.27	0.29		0.44	0.14
DICROWEB K2D	poor	fit			
U3.5 wt in 20 % TFE/buffer, 1d					
BeStSel	0.23	0.36	0.21	0.20	0.011
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	poor	fit			
U3.5 wt in 20 % TFE/buffer, 5d					
BeStSel	0.09	0.45	0.22	0.24	0.012
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.02	0.68		0.31	0.20
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S13. Summary of quantitative secondary structure estimation for the **U3.5 wt peptide in 40 % (v/v) TFE** using multiple approaches.[12,13,22–25,14–21] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 wt in 40 % TFE/water					
BeStSel	0.58	0	0.02	0.40	0.023
CAPITO basis spectra	0.77	0.01		0.63	0.43
CAPITO lowest area difference	0.72	0		0.28	0.36
DICROWEB K2D	poor	fit			
DICROWEB CDSSTR Ref 7	0.82	0.04	0.06	0.08	0.003
DICROWEB CDSSTR Ref 4	0.76	0.06	0.05	0.11	0.003
DICROWEB CONTIN-LL Ref 7	0.74	0.03	0.06	0.17	0.061
DICROWEB CONTIN-LL Ref 4	0.88	0.05	0.07	0	0.061
DICROWEB SELCON3 Ref 7	0.57	0	0.12	0.36	0.027
DICROWEB SELCON3 Ref 4	0.57	0	0.12	0.36	0.028
U3.5 wt in 40 % TFE/buffer					
BeStSel	0.42	0.26	0	0.32	0.050
CAPITO basis spectra	0.78	0.01		0.62	0.45
CAPITO lowest area difference	0.61	0.01		0.38	0.36
DICROWEB K2D	poor	fit			
U3.5 wt in 40 % TFE/buffer, 1d					
BeStSel	0.44	0.24	0	0.32	0.047
CAPITO basis spectra	0.78	0.01		0.59	0.45
CAPITO lowest area difference	0.61	0.01		0.38	0.35
DICROWEB K2D	poor	fit			
U3.5 wt in 40 % TFE/buffer, 5d					
BeStSel	0.47	0.21	0	0.32	0.043
CAPITO basis spectra	0.71	0.01		0.57	0.45
CAPITO lowest area difference	0.61	0.01		0.38	0.33
DICROWEB K2D	poor	fit			

The methods CDSSTR, CONTIN-LL and SELCON3 with reference sets 4 or 7 require CD data from 190 nm to 240 nm; thus, these methods are only applied for the peptide before the addition of buffer (buffer (chloride ions) interference below 195 nm).

Table S14. Summary of quantitative secondary structure estimation for **U3.5 wt in the presence of lipids** (peptide:lipid ratio was 1:9) using multiple approaches.[12–18] Only fits that roughly represented the experimental CD spectrum (shape) are shown, independent of the NRMSD value (poor fits are not shown). A low NRMSD value (normalized root mean square deviation) indicates a good fit.

Method	Helix	β-sheet/strand	Turn	Unordered/random	NRMSD
U3.5 wt in buffer					
BeStSel	0.20	0.05	0.10	0.65	0.013
CAPITO basis spectra	0.13	0.38		0.34	0.46
CAPITO lowest area difference	0.27	0.29		0.44	0.10
DICROWEB K2D	poor	fit			
U3.5 wt in buffer, 15 h					
BeStSel	0.03	0.23	0.12	0.62	0.014
CAPITO basis spectra	poor	fit			
CAPITO lowest area difference	0.14	0.42		0.44	0.20
DICROWEB K2D	poor	fit			
U3.5 wt in cholesterol					
BeStSel	0.18	0.09	0.11	0.62	0.012
CAPITO basis spectra	0.14	0.43		0.36	0.27
CAPITO lowest area difference	0.27	0.29		0.44	0.16
DICROWEB K2D	poor	fit			
U3.5 wt in cholesterol, 15 h					
BeStSel	0.05	0.23	0.12	0.60	0.012
CAPITO basis spectra	0.13	0.45		0.45	0.30
CAPITO lowest area difference	0.14	0.42		0.44	0.24
DICROWEB K2D	poor	fit			
U3.5 wt in DMPC					
BeStSel	0.37	0	0.10	0.53	0.009
CAPITO basis spectra	0.31	0.36		0.36	0.09
CAPITO lowest area difference	poor	fit			
DICROWEB K2D	0.40	0.06		0.54	0.155
U3.5 wt in DMPC, 15 h					
BeStSel	0.01	0.33	0.14	0.52	0.015
CAPITO basis spectra	0.12	0.49		0.44	0.31
CAPITO lowest area difference	0.07	0.56		0.37	0.40
DICROWEB K2D	0.14	0.38		0.48	0.208
U3.5 wt in DMPG					
BeStSel	0.32	0.19	0.12	0.37	0.007
CAPITO basis spectra	0.30	0.16		0.44	0.34
CAPITO lowest area difference	0.29	0.17		0.53	0.10
DICROWEB K2D	0.51	0.17		0.32	0.106
U3.5 wt in DMPG, 15 h					
BeStSel	0.51	0.08	0.11	0.30	0.004
CAPITO basis spectra	0.42	0.03		0.45	0.40
CAPITO lowest area difference	0.57	0.02		0.41	0.09
DICROWEB K2D	0.62	0.06		0.31	0.124

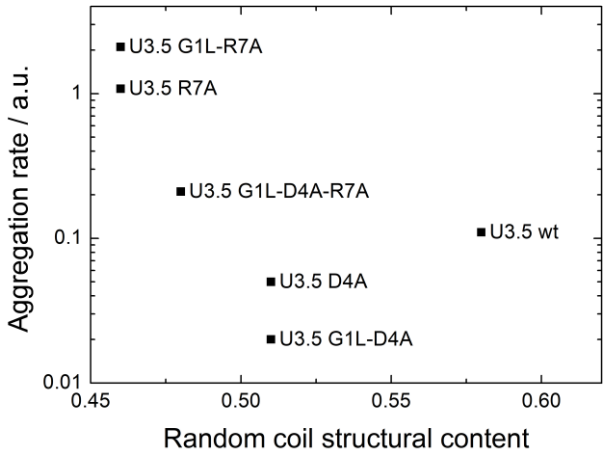

Figure S4. Correlation between the U3.5 amyloid fibril formation rates (from ThT assay, see Table S2) and the "random coil" structural content of the U3.5 variants in water (predicted using the CAPITO lowest area difference[14]). The amyloid fibril formation rate is higher for a lower "random coil" and thus higher α-helical or β-sheet structural content for most U3.5 variants.

2.3 Secondary structure prediction models

Table S15. Summary of sequence-based secondary structure prediction results for the U3.5 variants using several approaches. Most methods predict a helical structure with unstructured termini, except the NPS@/PHD algorithm with higher predicted strand-content for some variants. (C = Coil, H = Helix, E = Strand, L = Loop)

	U3.5 wt	**U3.5 G1L**	**U3.5 D4A**	**U3.5 R7A**
Scratch Protein Predictor/SSpro[1–3]	CCHHHHHHHHHHHHHCC	CCHHHHHHHHHHHHHCC	CHHHHHHHHHHHHHHCC	CCHHHHHHHHHHHHHCC
Predict-Protein/ PROFsec[4,5]	LLHHHHHHHHHHHHLLL	LLHHHHHHHHHHHHHLL	LLHHHHHHHHHHHHHLL	LLHHHHHHHHHHHHHLL
NPS@/PHD[5–7]	CCCCECEEEEEEEECCC	CCCCHHHHHHHHHHHHC	CCEEEEEEHHHHHHHHC	CCCHHHEEEEEEHHHCC
PSIPRED[8,9]	CHHHHHHHHHHHHHHHC	CHHHHHHHHHHHHHHHC	CHHHHHHHHHHHHHHHC	CHHHHHHHHHHHHHHHC
PSSpred[10]	CHHHHHHHHHHHHHHCC	CHHHHHHHHHHHHHHCC	CCHHHHHHHHHHHHHCC	CHHHHHHHHHHHHHHCC
	U3.5 G1L-D4A	**U3.5 G1L-R7A**	**U3.5 G1L-D4A-R7A**	
Scratch Protein Predictor/SSpro[1–3]	CHHHHHHHHHHHHHHCC	CCHHHHHHHHHHHHHCC	CHHHHHHHHHHHHHHCC	
Predict-Protein/ PROFsec[4,5]	LHHHHHHHHHHHHHHLL	LLHHHHHHHHHHHHHLL	LHHHHHHHHHHHHHHHL	
NPS@/PHD[5–7]	CCCCHHHHHHHHHHHHC	CCCHHHHHHHHHHHHHC	CCCHHHHHHHHHHHHHC	
PSIPRED[8,9]	CHHHHHHHHHHHHHHHC	CHHHHHHHHHHHHHHHC	CHHHHHHHHHHHHHHHC	
PSSpred[10]	CCHHHHHHHHHHHHHCC	CHHHHHHHHHHHHHHHC	CCHHHHHHHHHHHHHHC	

3 Supporting references

(1) Cheng, J., Randall, A. Z., Sweredoski, M. J., and Baldi, P. (2005) SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res. 33*, W72–W76.

(2) Magnan, C. N., and Baldi, P. (2014) SSpro/ACCpro 5: Almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles, machine learning and structural similarity. *Bioinformatics 30*, 2592–2597.

(3) Pollastri, G., Przybylski, D., Rost, B., and Baldi, P. (2002) Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins Struct. Funct. Genet. 47*, 228–235.

(4) Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., Hönigschmid, P., Schafferhans, A., Roos, M., Bernhofer, M., Richter, L., Ashkenazy, H., Punta, M., Schlessinger, A., Bromberg, Y., Schneider, R., Vriend, G., Sander, C., Ben-Tal, N., and Rost, B. (2014) PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res. 42*, W337–W343.

(5) Rost, B., and Sander, C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins Struct. Funct. Genet. 19*, 55–72.

(6) Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000) NPS@: Network Protein Sequence Analysis. *Trends Biochem. Sci. 25*, 147–150.

(7) Rost, B., and Sander, C. (1993) Prediction of Protein Secondary Structure at Better than 70% Accuracy. *J. Mol. Biol. 232*, 584–599.

(8) Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res. 41*, W349–W357.

(9) Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol. 292*, 195–202.

(10) Yan, R., Xu, D., Yang, J., Walker, S., and Zhang, Y. (2013) A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Sci. Rep. 3*, 2619.

(11) Meisl, G., Kirkegaard, J. B., Arosio, P., Michaels, T. C. T., Vendruscolo, M., Dobson, C. M., Linse, S., and Knowles, T. P. J. (2016) Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat. Protoc. 11*, 252–272.

(12) Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y. H., Goto, Y., Réfrégiers, M., and Kardos, J. (2018) BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res. 46*, W315–W322.

(13) Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., and Kardos, J. (2015) Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. 112*, E3095–E3103.

(14) Wiedemann, C., Bellstedt, P., and Görlach, M. (2013) CAPITO - A web server-based analysis and plotting tool for circular dichroism data. *Bioinformatics 29*, 1750–1757.

(15) Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers 89*, 392–400.

(16) Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res. 32*, W668–W673.

(17) Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: An interactive website for the

analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics 18*, 211–212.

(18) Andrade, M. A., Chacón, P., Merelo, J. J., and Morán, F. (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng. Des. Sel. 6*, 383–390.

(19) Compton, L. A., and Johnson, W. C. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem. 155*, 155–167.

(20) Manavalan, P., and Johnson Jr., W. C. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem. 167*, 76–85.

(21) Sreerama, N., and Woody, R. W. (2000) Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochem. 287*, 252–260.

(22) van Stokkum, I. H. M., Spoelder, H. J. W., Bloemendal, M., van Grondelle, R., and Groen, F. C. A. (1990) Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal. Biochem. 191*, 110–118.

(23) Provencher, S. W., and Glöckner, J. (1981) Estimation of Globular Protein Secondary Structure from Circular Dichroism. *Biochemistry 20*, 33–37.

(24) Sreerama, N., and Woody, R. W. (1993) A Self-Consistent Method for the Analysis of Protein Secondary Structure from Circular Dichroism. *Anal. Biochem. 209*, 32–44.

(25) Sreerama, N., Venyaminov, S. Y. U., and Woody, R. W. (1999) Estimation of the number of α-helical and β-strand segments in proteins using circular dichroism spectroscopy. *Protein Sci. 8*, 370–380.