Supplementary Material

Application of Derandomisation of Peptide Circular Dichroism Spectra to Determine Its Secondary Structure Content

Adewale Olamoyesan[1], Dale Ang[1], Lisa Martin[2], Alison Rodger[1]

[1]Macquarie University, Faculty of Science and Engineering, Department of Molecular Sciences, Macquarie Park, NSW 2109, Australia
[2]Monash University, School of Chemistry, Clayton, VIC 3800, Australia

Table of Contents

Figure S10 NRMSDs for the fitting of the U3.5R7A model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 195 nm.

Figure S11 NRMSDs for the fitting of the U3.6wt model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 195 nm.

Figure S12 NRMSDs for the fitting of the U3.6 K7A model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 195 nm.

Figure S13 NRMSDs for the fitting of the U3.4 wt model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 200 nm.

Figure S14 NRMSDs for the fitting of the U3.4 R7A model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 200 nm.

Figure S15 NRMSDs for the fitting of the U3.5 wt model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h. while using data down to 200 nm.

Figure S16 NRMSDs for the fitting of the U3.5 R7A model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 200 nm.

Figure S17 NRMSDs for the fitting of the U3.6 wt model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 200 nm.

Figure S18 NRMSDs for the fitting of the U3.6 K7A model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h while using data down to 200 nm.

Figure S19 Secondary structure predictions for regenerated U3.4 wt in water only and after addition buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 wt predictions.

Figure S20 Secondary structure predictions for regenerated U3.4 R7A in water only and after addition buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 R7A predictions.

Figure S21 Secondary structure predictions for regenerated U3.5 wt in water only and after addition buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 wt predictions.

Figure S22 Secondary structure predictions for regenerated U3.5 R7A in water only and after addition buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 R7A predictions.

Figure S23 Secondary structure predictions for regenerated U3.6 wt in water only and after additon of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

Figure S24 Secondary structure predictions for regenerated U3.6 K7A in water only and after additon of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 K7A predictions.

Figure S25 Secondary structure predictions for regenerated U3.4 wt in water only and after additon of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 wt predictions.

Figure S26 Secondary structure predictions for regenerated U3.4 R7A in water only and after additon of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 R7A predictions.

Figure S27 Secondary structure predictions for regenerated U3.5 wt in water only and after additon of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 wt predictions.

F**igure S28** Secondary structure predictions for regenerated U3.5 R7A in water only and after addition of buffer over time, created by removing 5 nm from the original spectra.. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 R7A predictions.

Figure S29 Secondary structure predictions for regenerated U3.6 wt in water only and after addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

F**igure S30** Secondary structure predictions for regenerated U3.6 K7A in water only and after addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 K7A predictions.

Figure S31 Secondary structure predictions for regenerated U3.4 wt in the absence and presence of created by removing 10 nm from the original spectra. The SOMSpec predictions for the derandomised U3.4 wt spectra were scaled to account for the removed RC percentage and then then RC percentage added to the other

category. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 wt predictions.

Figure S32 Secondary structure predictions for regenerated U3.4 R7A in the absence and presence of created by removing 10 nm from the original spectra. The SOMSpec predictions for the derandomised U3.4 R7A spectra were scaled to account for the removed RC percentage and then then RC percentage added to the other category. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 R7A predictions.

Figure S33 Secondary structure predictions for regenerated U3.5 R7A in the absence and presence of created by removing 10 nm from the original spectra. The SOMSpec predictions for the derandomised U3.5 R7A spectra were scaled to account for the removed RC percentage and then then RC percentage added to the other category. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 R7A predictions.

Figure S34 Secondary structure predictions for regenerated U3.6 wt in water only and after addtion of buffer over time, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

Figure S35 Secondary structure predictions for regenerated U3.6 K7A in water only and after addtion of buffer over time, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

Table S1 Spectral NRMSDs for U3.4 wt and its seventh position variant before and after addition of buffer with 0-90%RC while using reference set and test file down to 190 nm, 195 nm and 200 nm.

Table S2 Spectral NRMSDs for U3.5 wt and its seventh position variant before and after addition of buffer with 0-90%RC while using reference set and test file down to 190 nm, 195 nm and 200 nm.

Table S3 Spectral NRMSDs for U3.6 wt and its seventh position variant before and after addition of buffer with 0-90%RC while using reference set and test file down to 190 nm, 195 nm and 200 nm.

This section contains the plots of the spectral NRMSDs, secondary structure predictions, comparison of spectral NRMSDs for wavelength range down to 190 nm, 195 nm and 200 nm for the each of six uperin peptides

Figures S1 shows the spectral NRMSDs of the best model spectrum from U3.4wt experimental spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set with data down to 190 nm for SOMSpec training and prediction respectively.

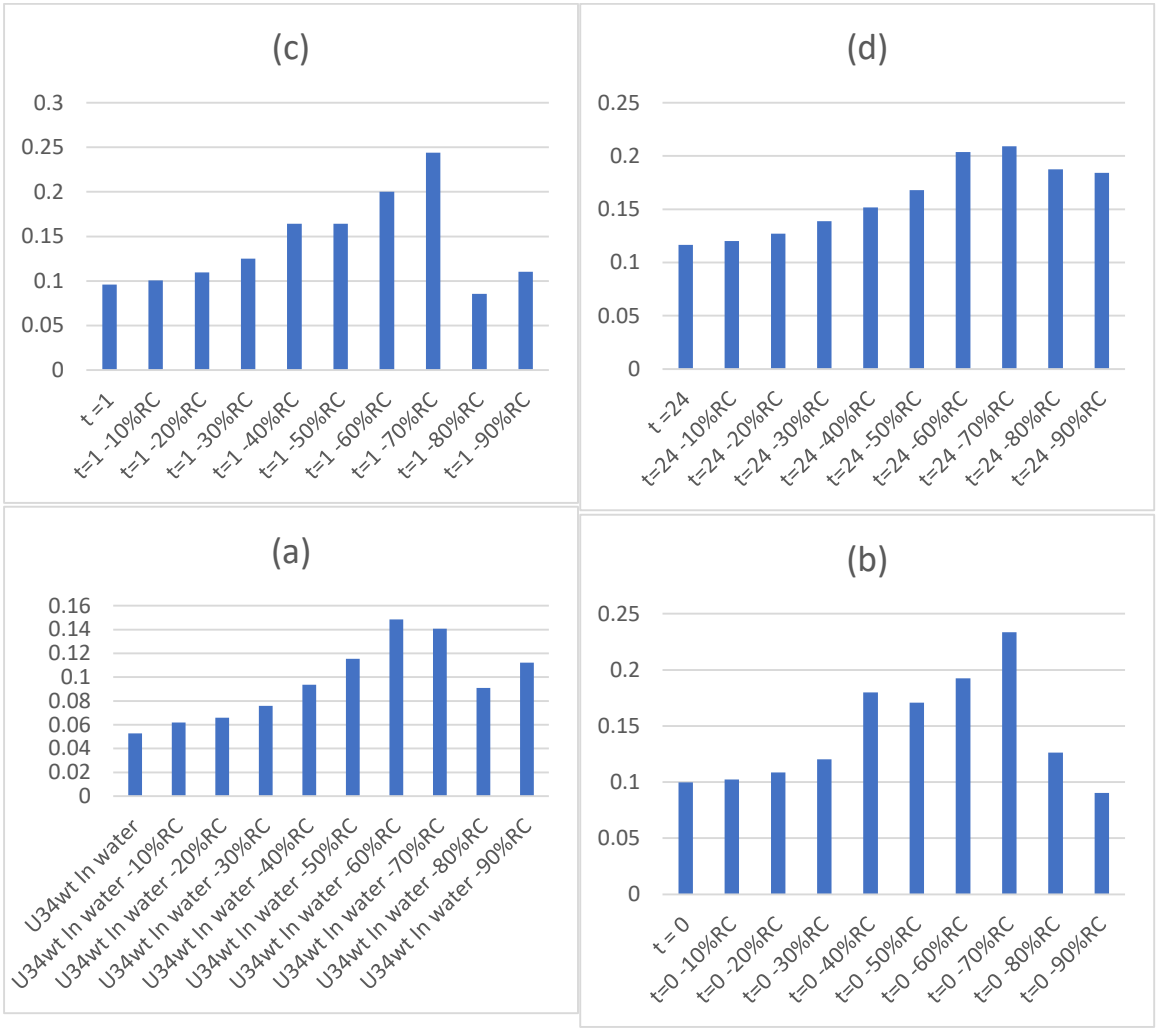


Figure S1 NRMSDs for the fitting of the U3.4wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S2 shows the spectral NRMSDs of the best model spectrum from the experimental U3.4R7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set with data down to 190 nm for SOMSpec training and prediction respectively.

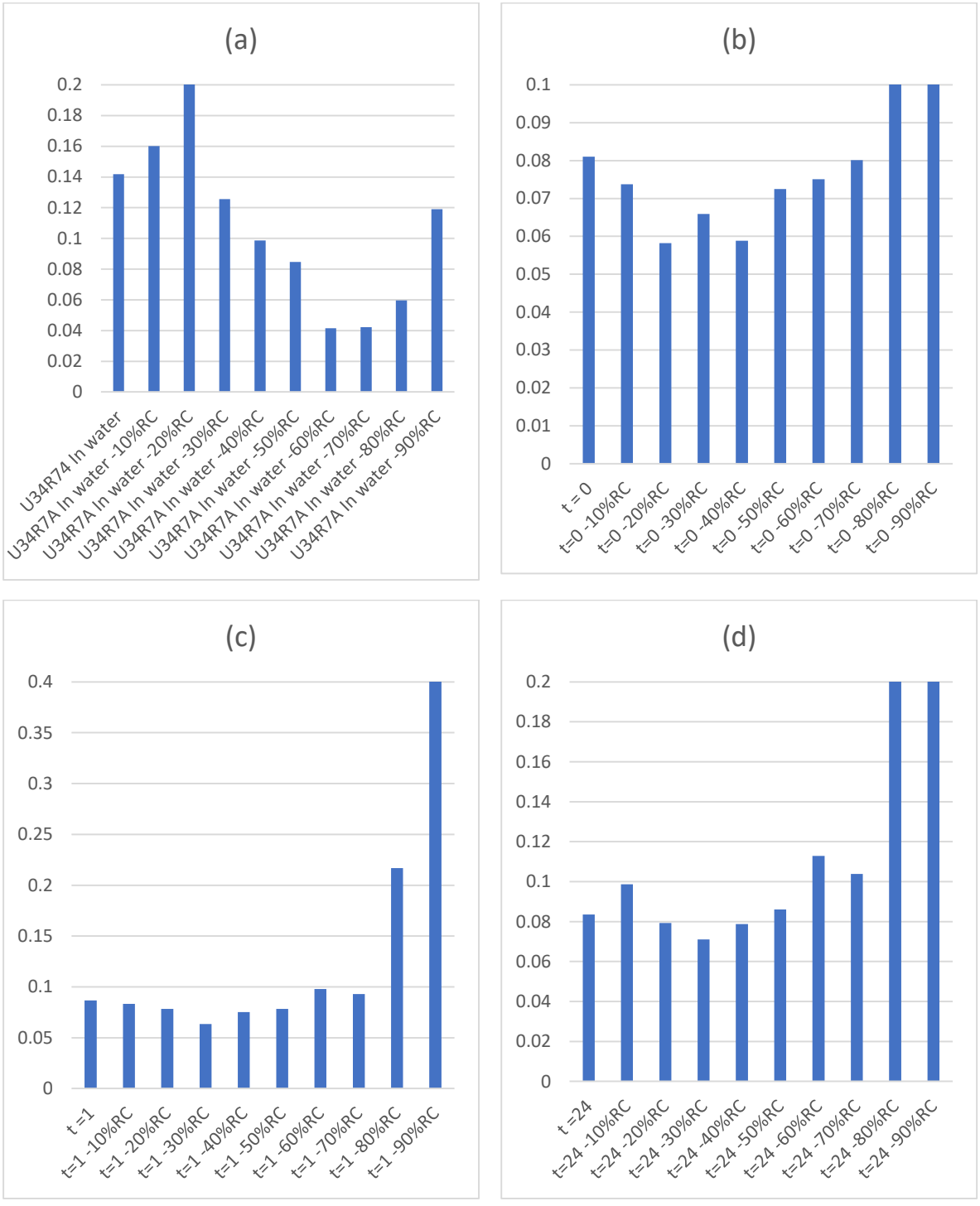


Figure S2 NRMSDs for the fitting of the U3.4R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h

Figures S3 shows the spectral NRMSDs of the best model spectrum from the experimental U3.5 wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set with data down to 190 nm for SOMSpec training and prediction respectively.

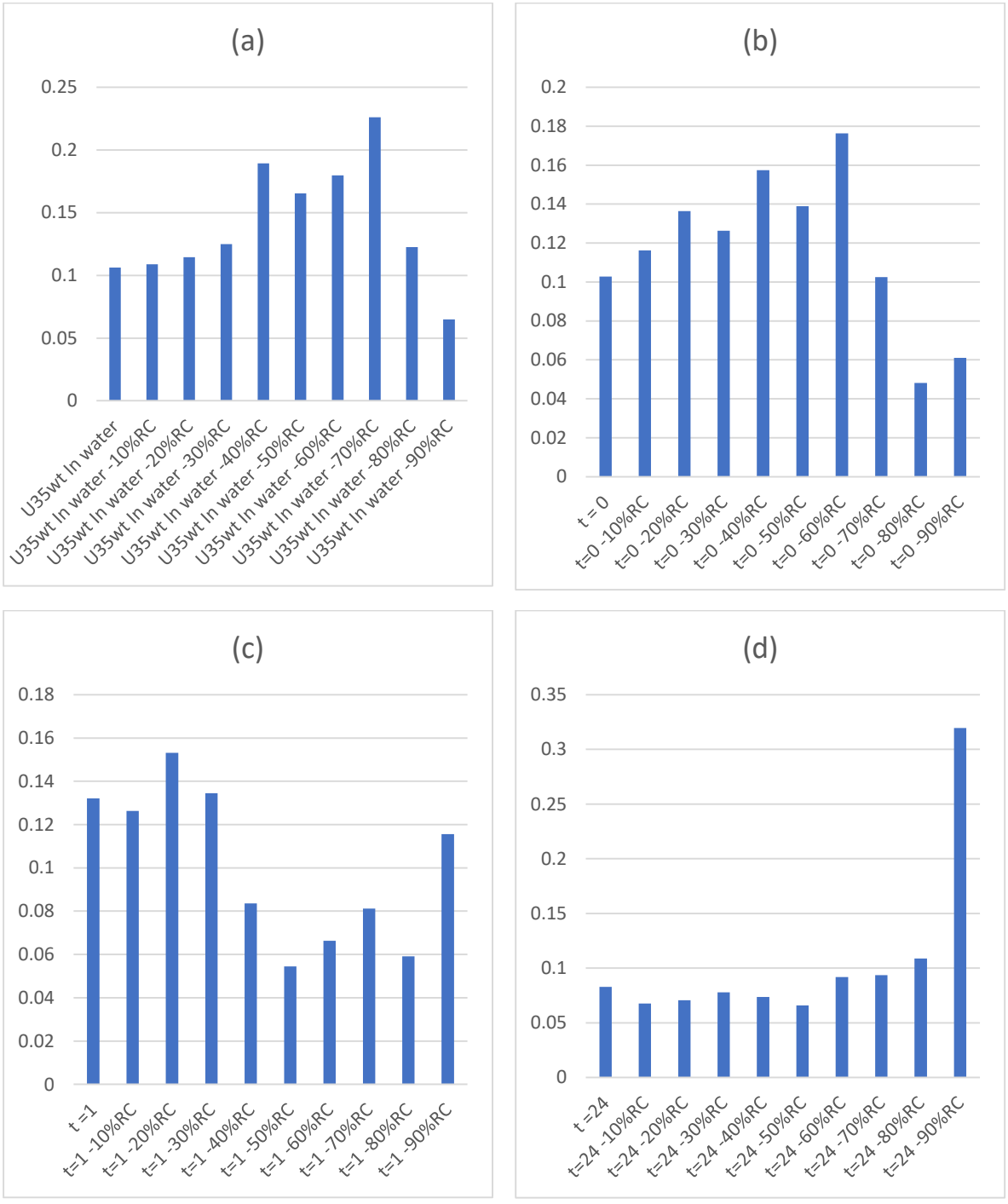


Figure S3 NRMSDs for the fitting of the U3.5 wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S4 shows the spectral NRMSDs of the best model spectrum from the experimental U3.5 R7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set with data down to 190 nm for SOMSpec training and prediction respectively.

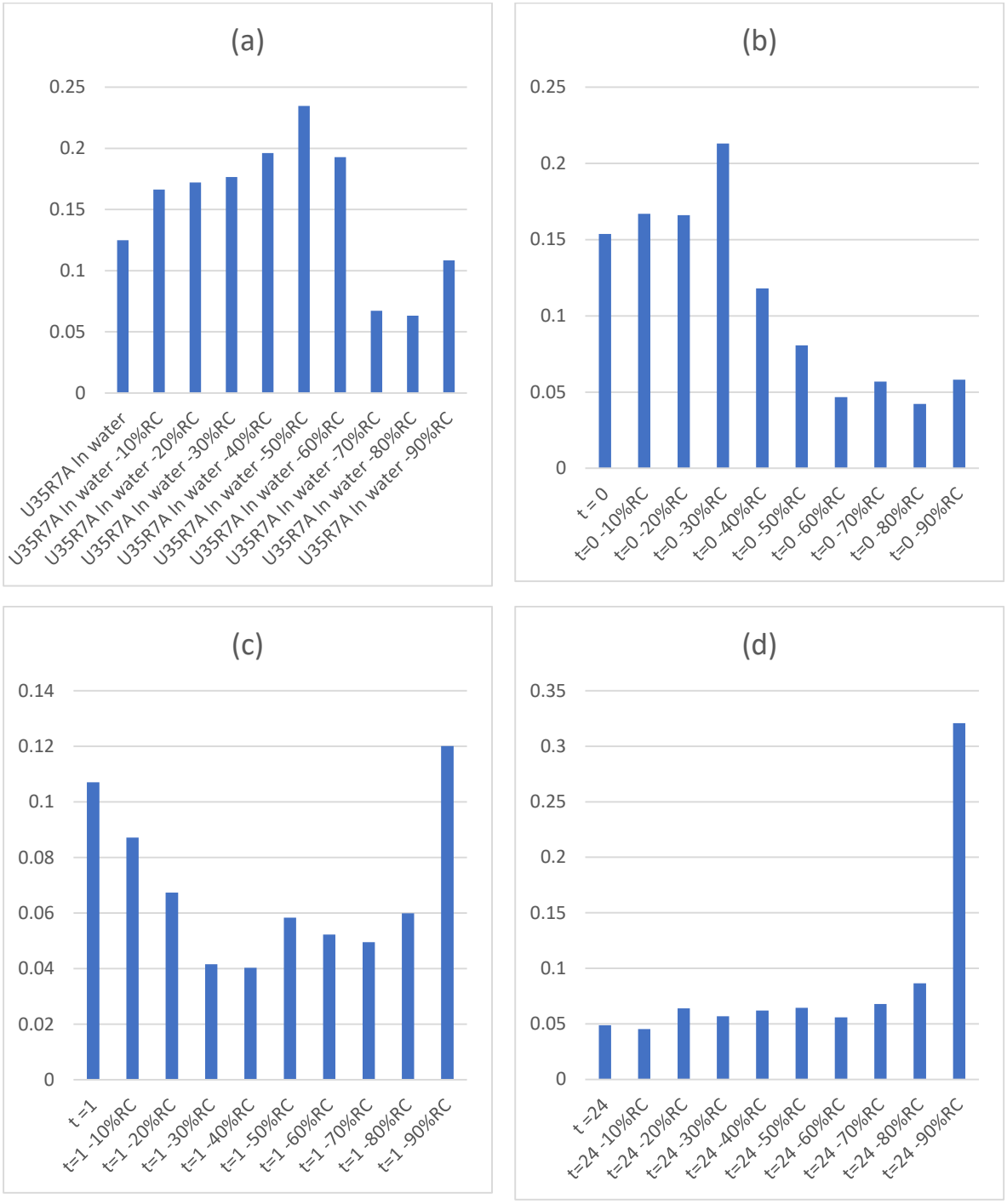


Figure S4 NRMSDs for the fitting of the U3.5R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S5 shows the spectral NRMSDs of the best model spectrum from the experimental U3.6wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set with data down to 190 nm for SOMSpec training and prediction respectively.

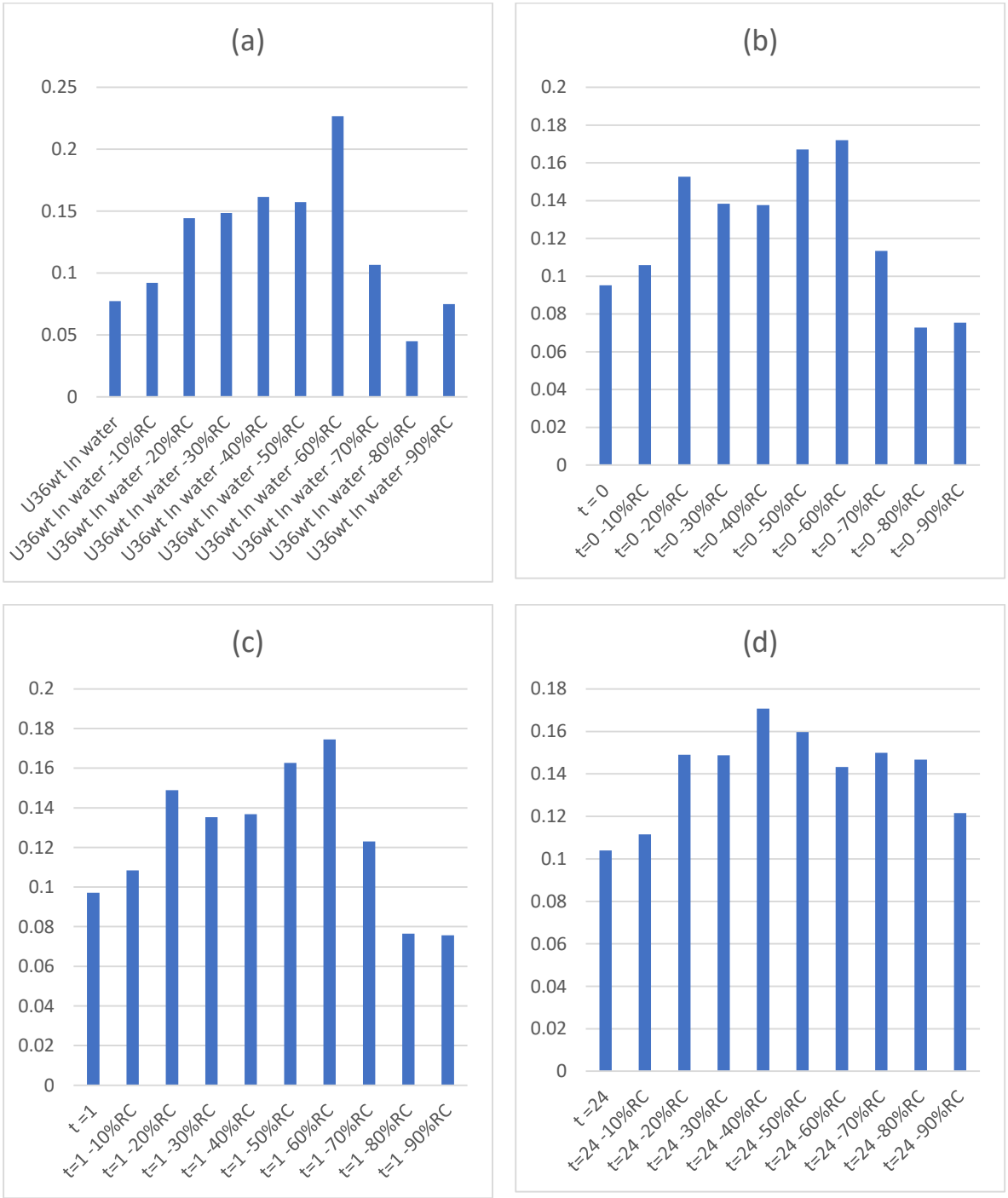


Figure S5 NRMSDs for the fitting of the U3.6wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S6 shows the spectral NRMSDs of the best model spectrum from the experimental U3.6R7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set with data down to 190 nm for SOMSpec training and prediction respectively.

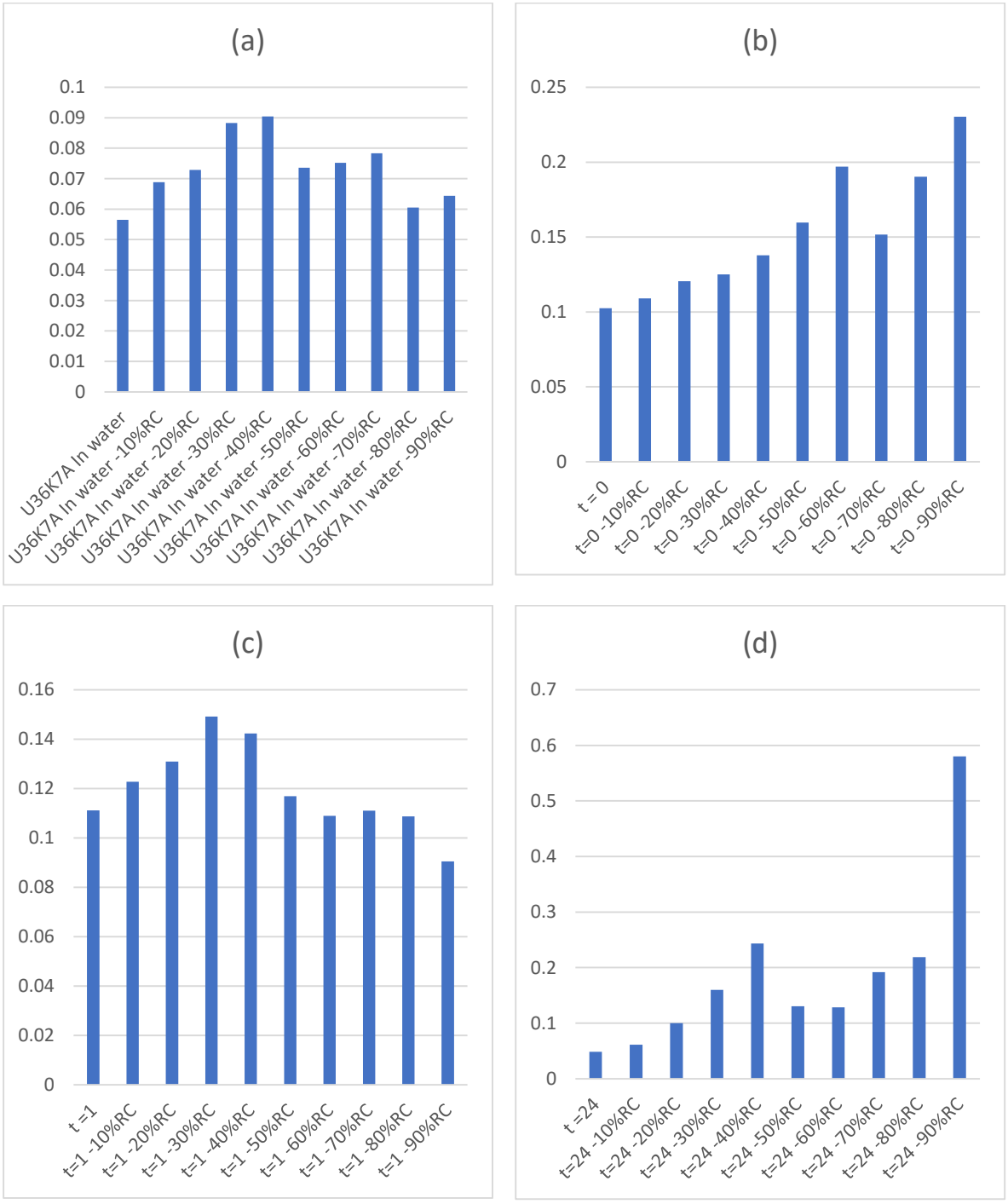


Figure S6 NRMSDs for the fitting of the U3.6R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S7 shows the spectral NRMSDs of the best model spectrum from the experimental U3.4wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 195 nm for SOMSpec training and prediction respectively

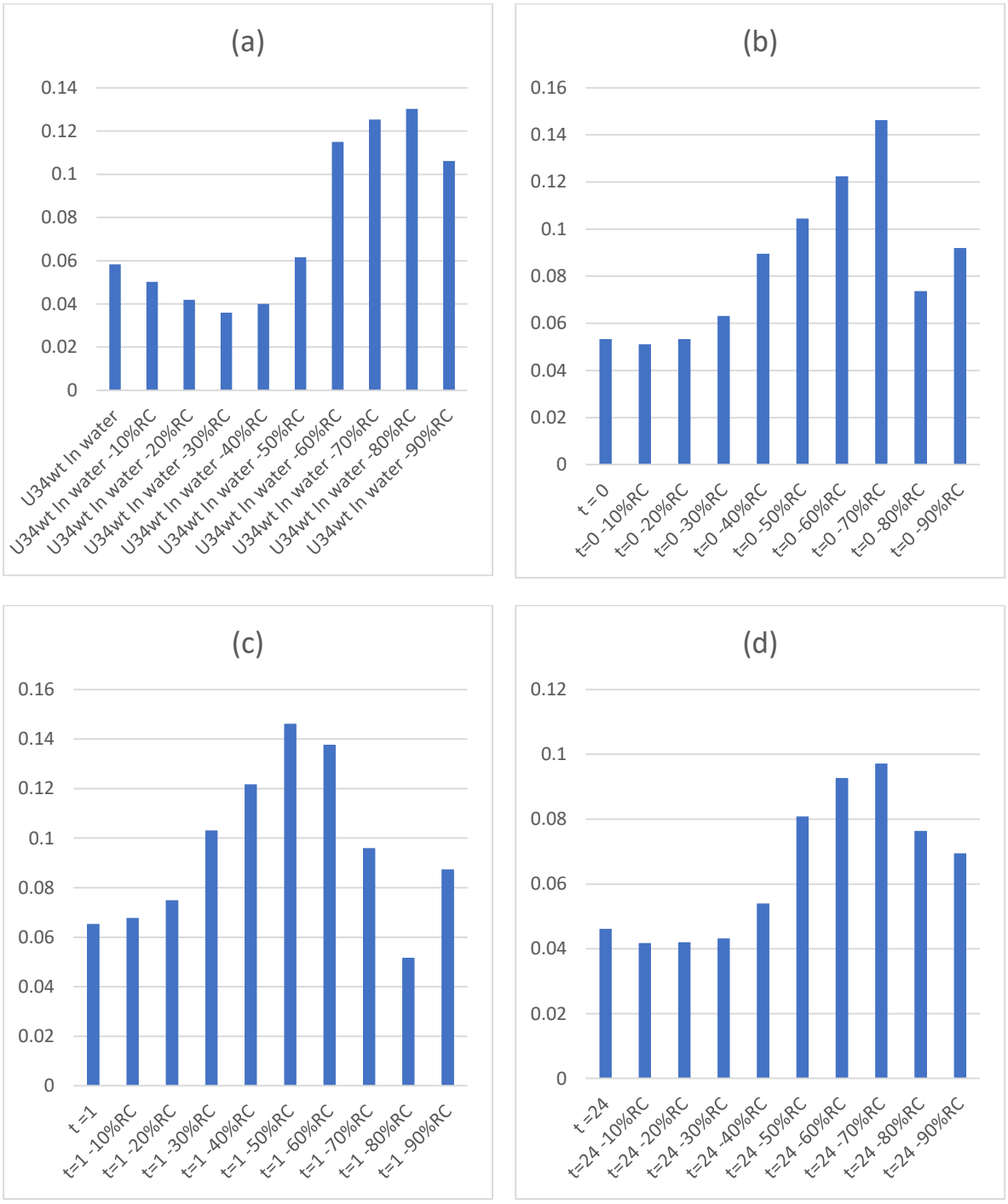


Figure S7 NRMSDs for the fitting of the U3.4 wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S8 shows the spectral NRMSDs of the best model spectrum from the experimental U3.4R7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 195 nm for SOMSpec training and prediction respectively

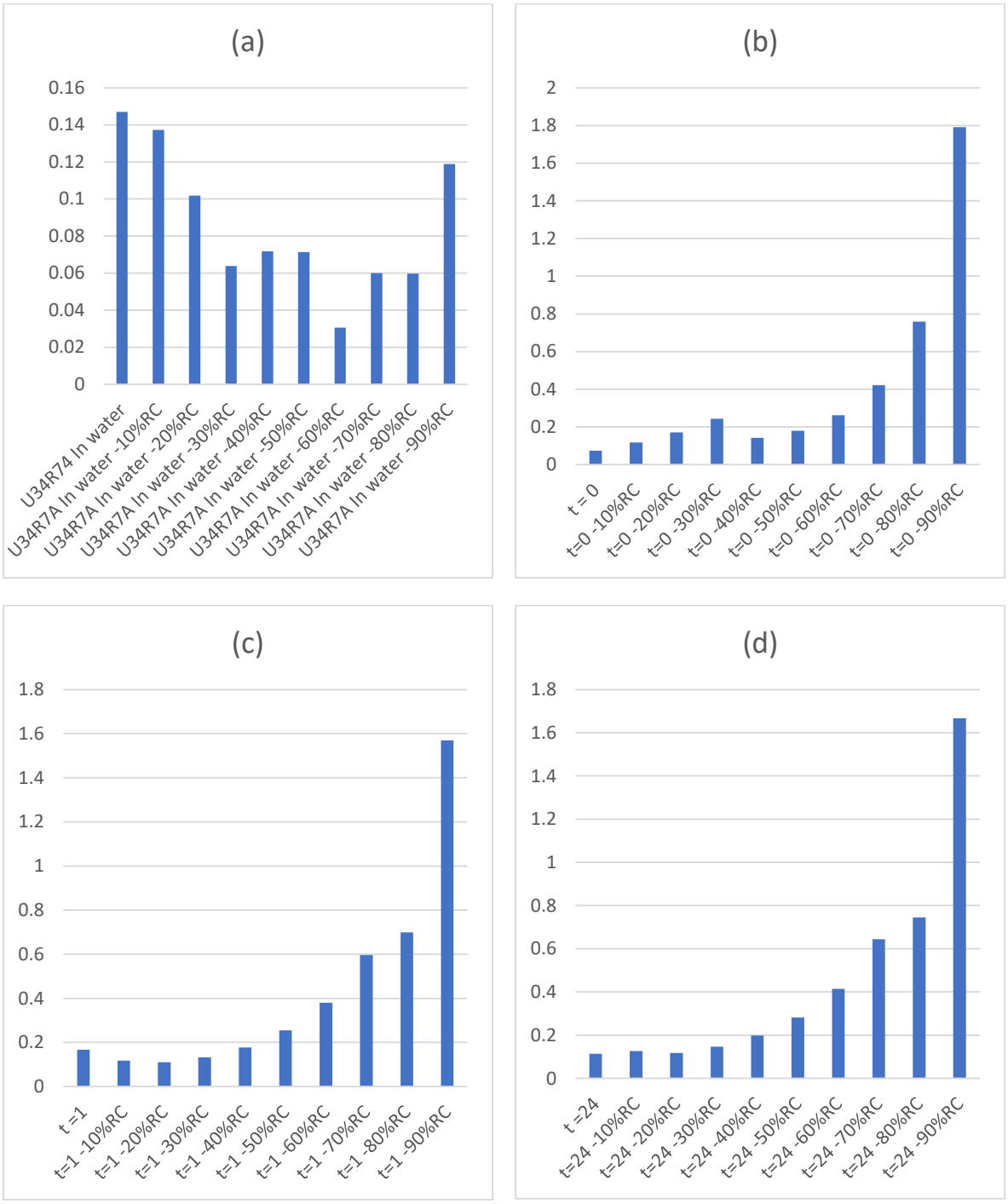


Figure S8 NRMSDs for the fitting of the U3.4R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S9 shows the spectral NRMSDs of the best model spectrum from the experimental U3.5wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 195 nm for SOMSpec training and prediction respectively

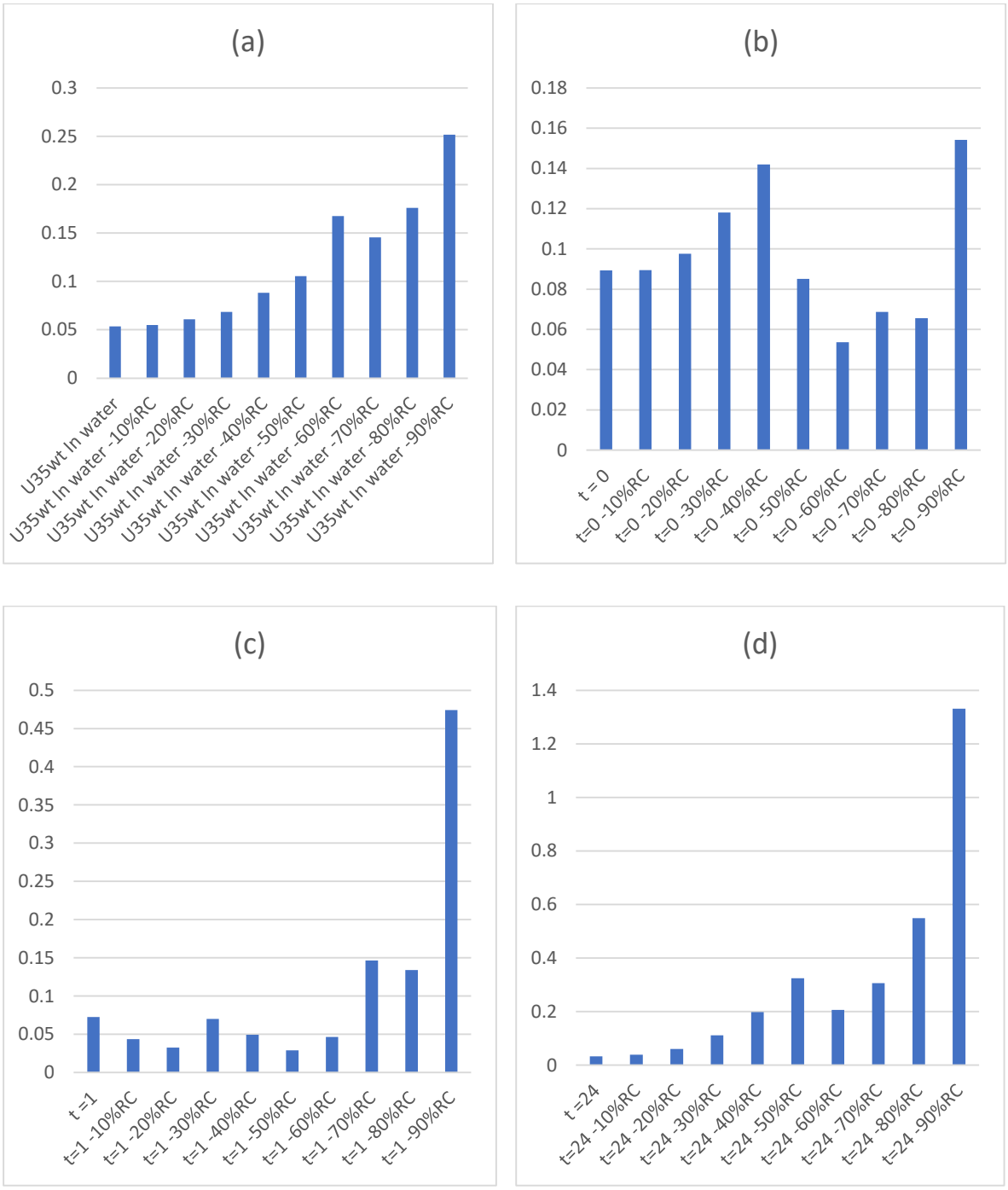


Figure **S9** NRMSDs for the fitting of the U3.5wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S10 shows the spectral NRMSDs of the best model spectrum from the experimental U3.5R7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 195 nm for SOMSpec training and prediction respectively

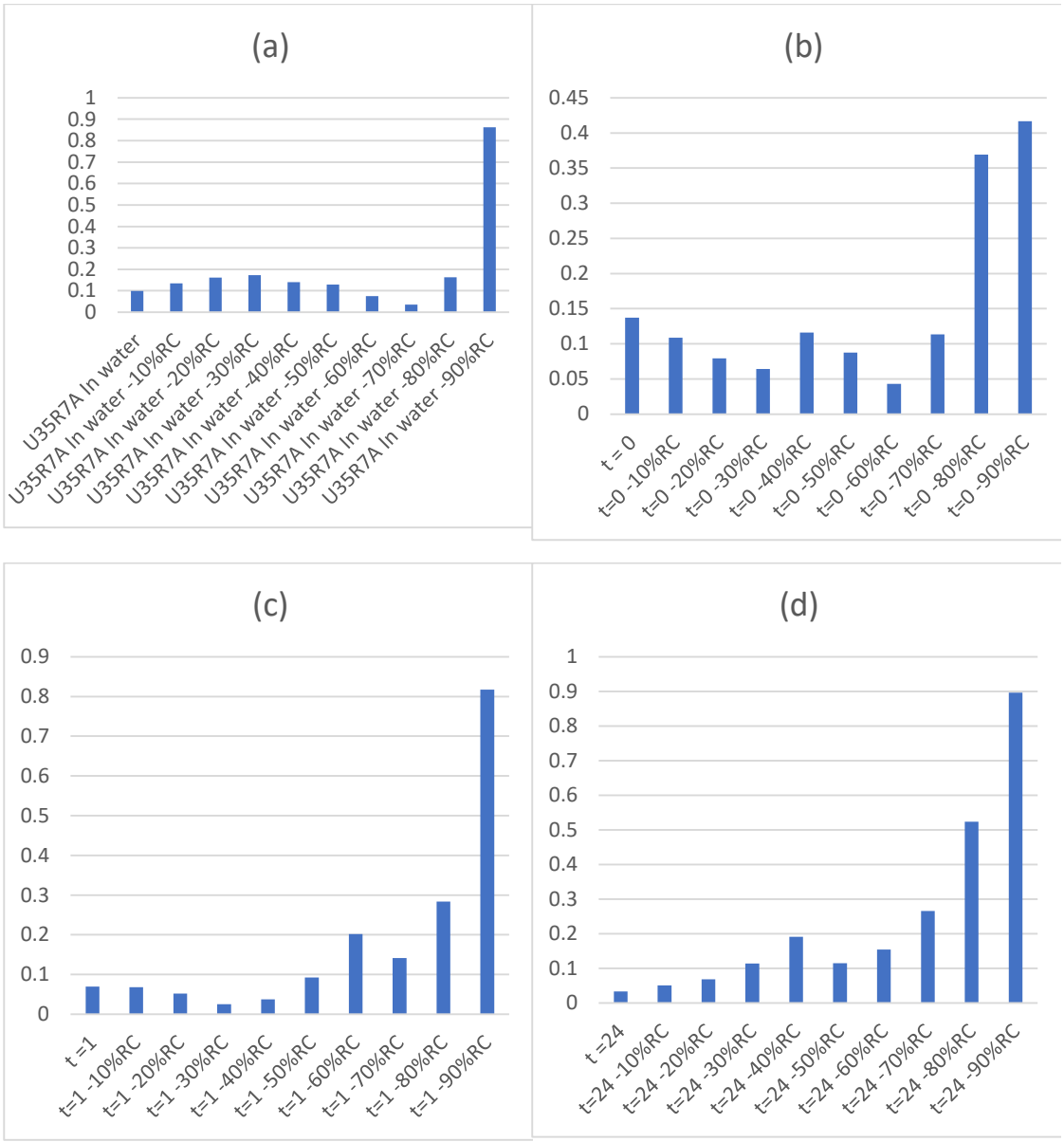


Figure S10 NRMSDs for the fitting of the U3.5R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S11 shows the spectral NRMSDs of the best model spectrum from the experimental U3.6wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 195 nm for SOMSpec training and prediction respectively

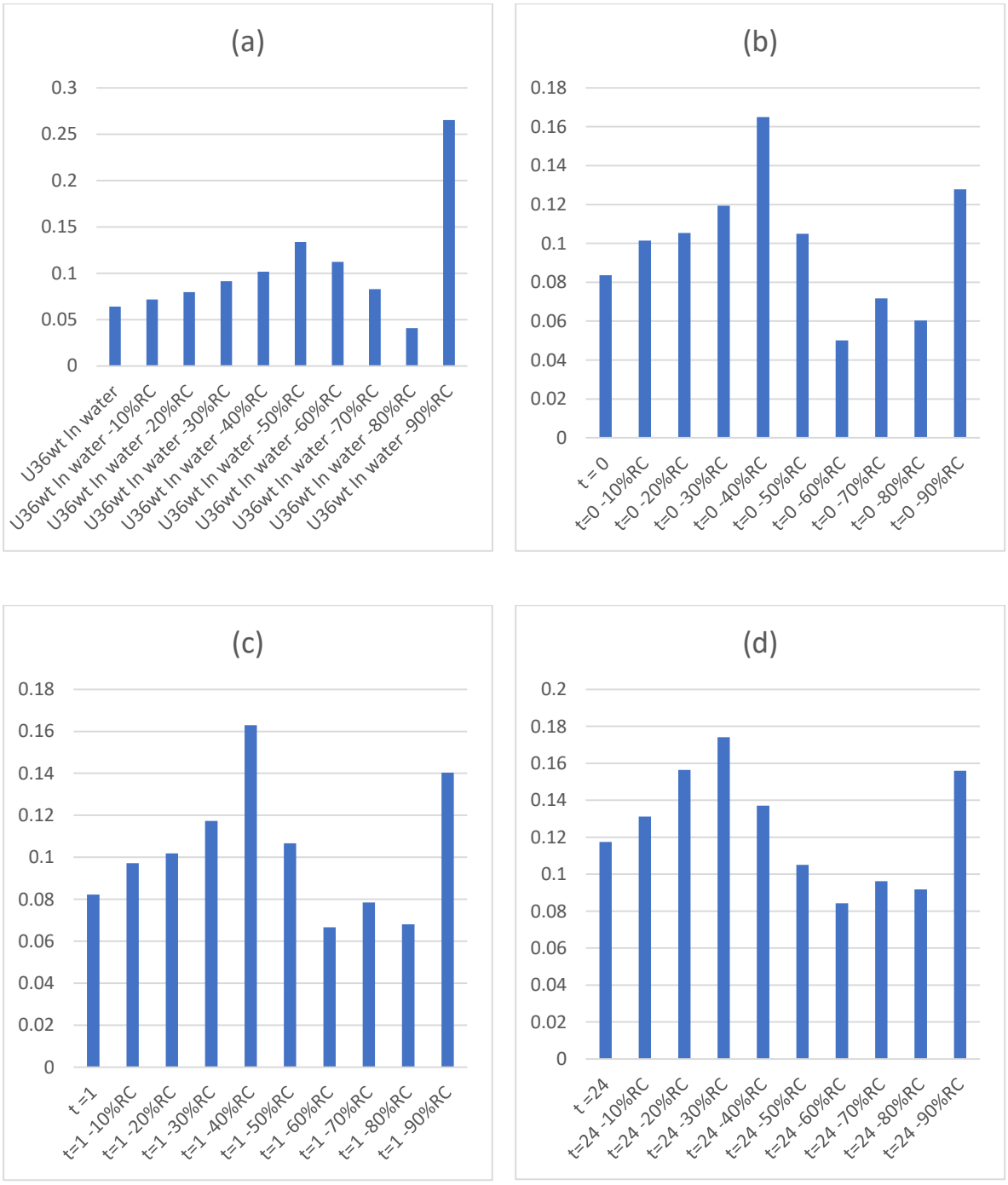


Figure S11 NRMSDs for the fitting of the U3.6wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S12 shows the spectral NRMSDs of the best model spectrum from the experimental U3.6 K7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 195 nm for SOMSpec training and prediction respectively

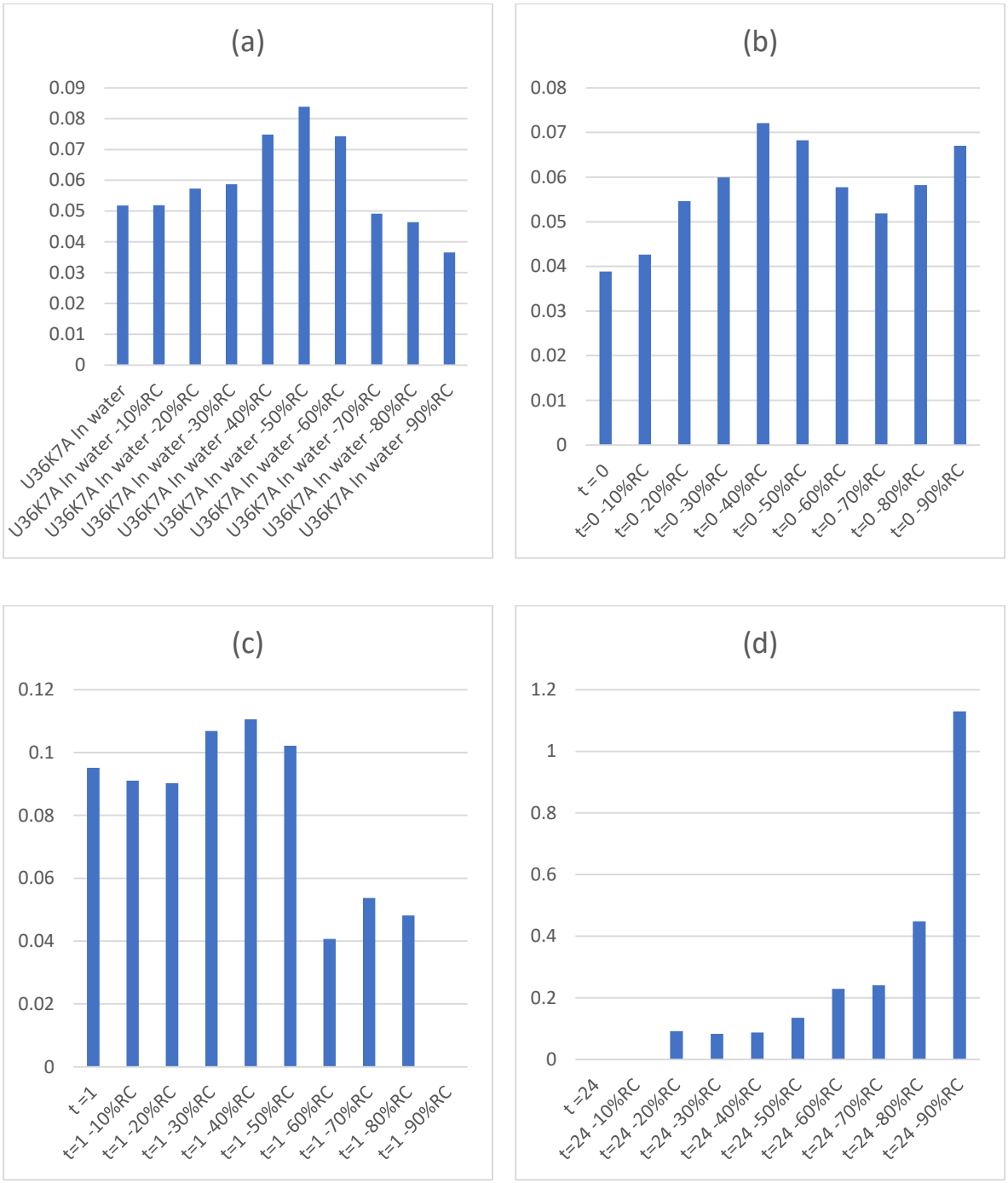


Figure S12 NRMSDs for the fitting of the U3.6 K7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figure S13 shows the spectral NRMSDs of the best model spectrum from the experimental U3.4 wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 200 nm for SOMSpec training and prediction respectively

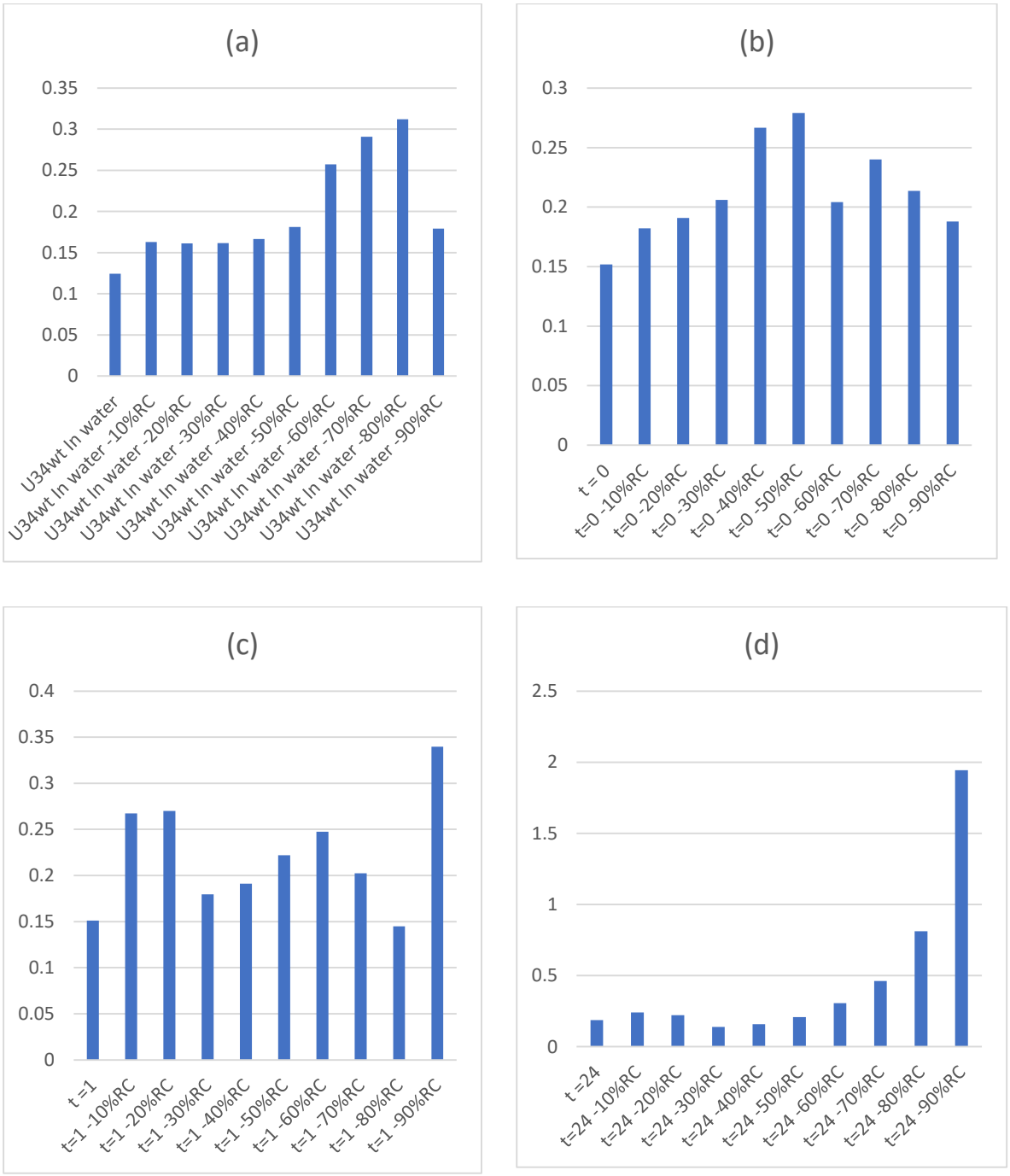


Figure S13 NRMSDs for the fitting of the U3.4 wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figure S14 shows the spectral NRMSDs of the best model spectrum from the experimental U3.4 R7A spectra with varying percentage RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 200 nm for SOMSpec training and prediction respectively

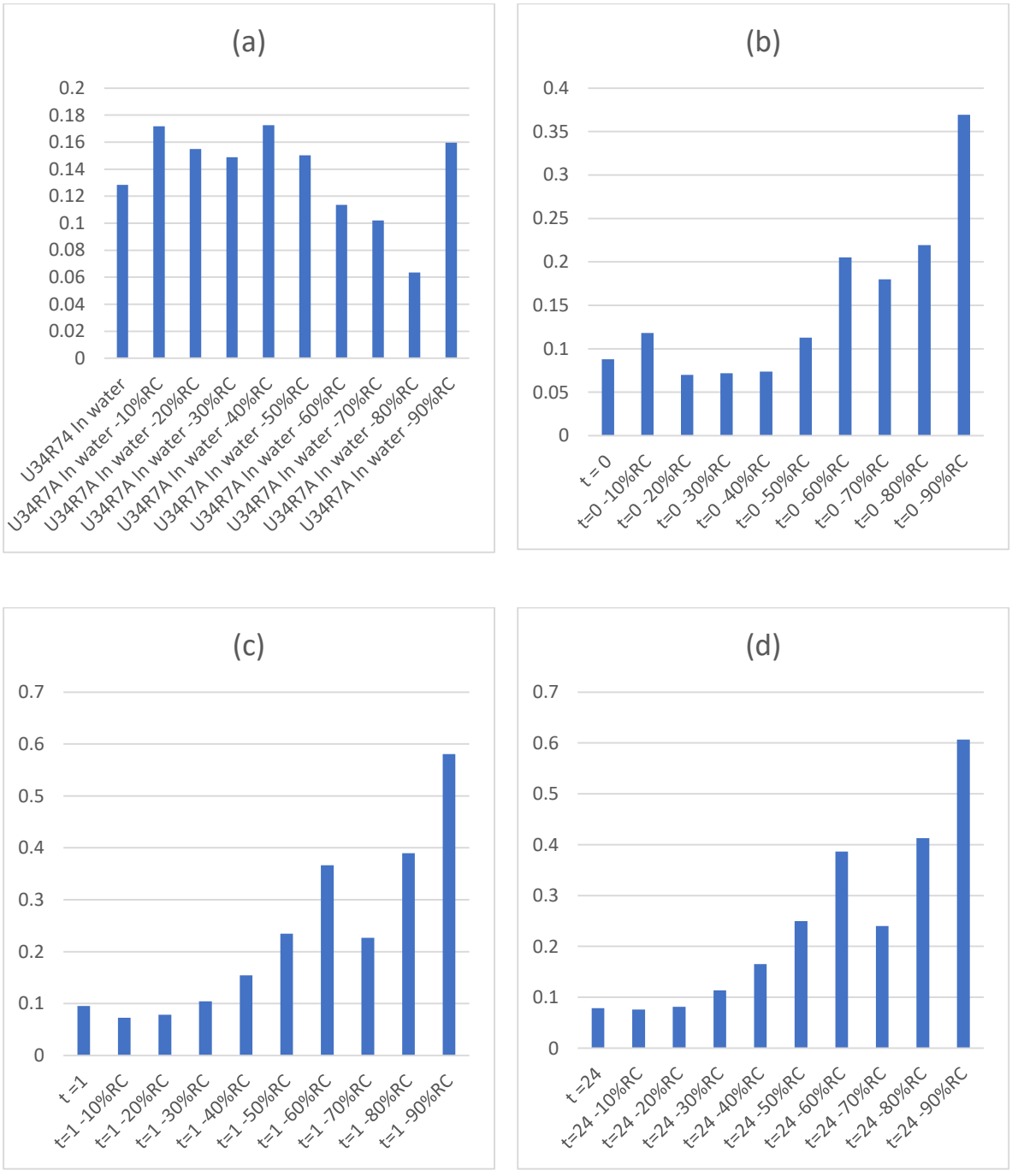


Figure S14 NRMSDs for the fitting of the U3.4 R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figure S15 shows the spectral NRMSDs of the best model spectrum from the experimental U3.5 wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 200 nm for SOMSpec training and prediction respectively

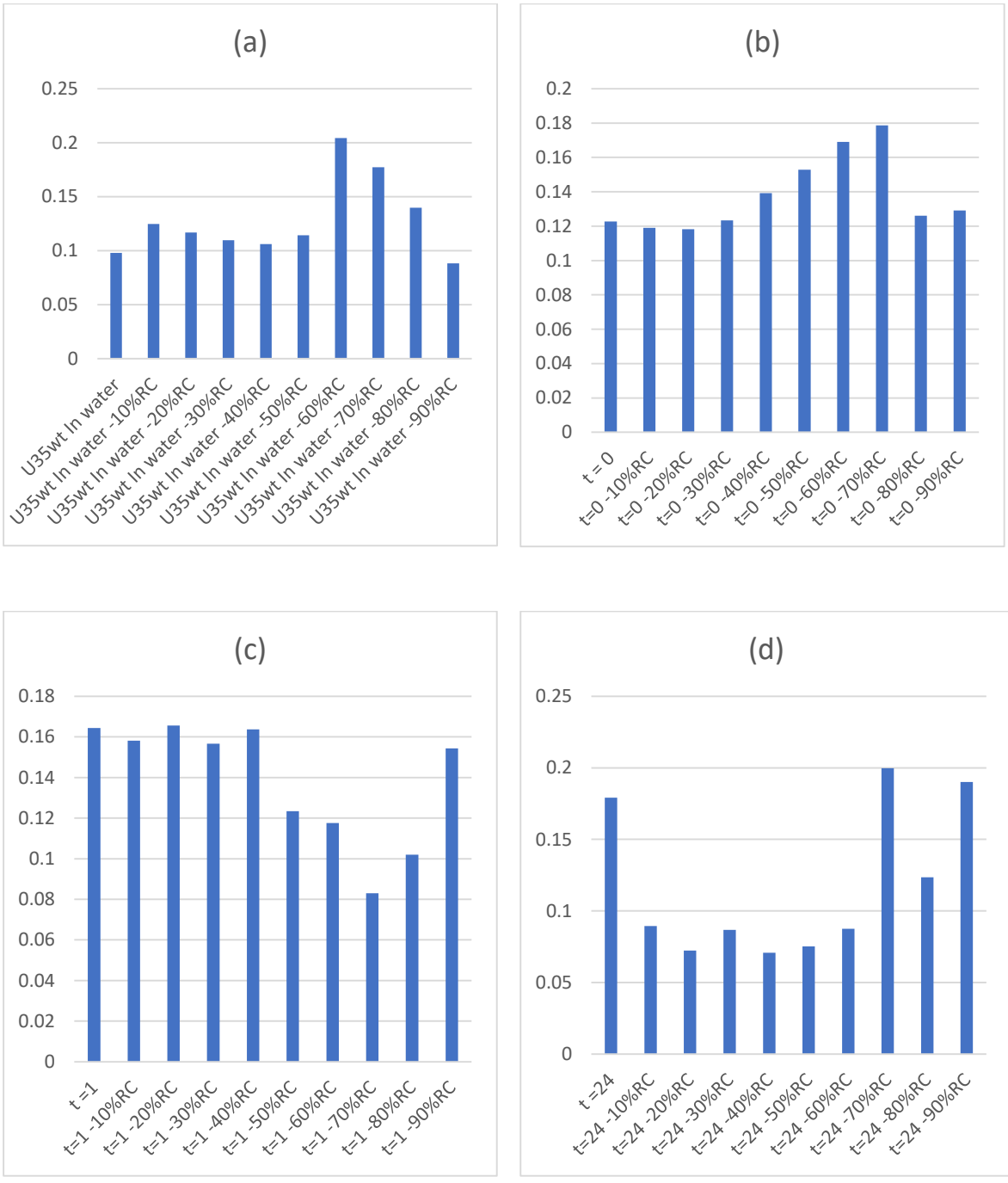


Figure S15 NRMSDs for the fitting of the U3.5 wt model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S16 shows the spectral NRMSDs of the best model spectrum from the experimental U3.5 R7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 200 nm for SOMSpec training and prediction respectively

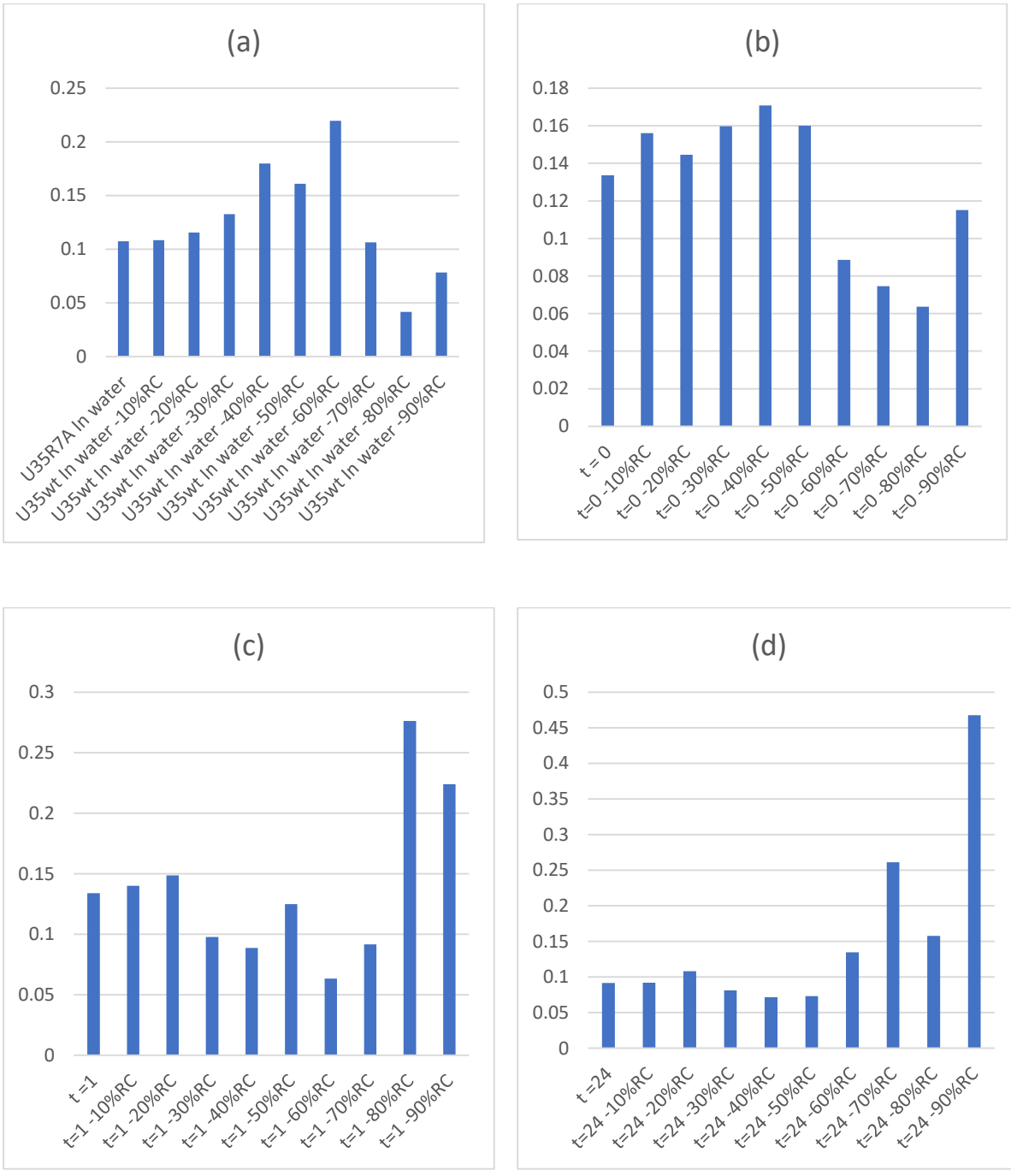


Figure S16 NRMSDs for the fitting of the U3.5 R7A model spectrum to experimental data as a function of fractions of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figure S17 shows the spectral NRMSDs of the best model spectrum from the experimental U3.6 wt spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 200 nm for SOMSpec training and prediction respectively

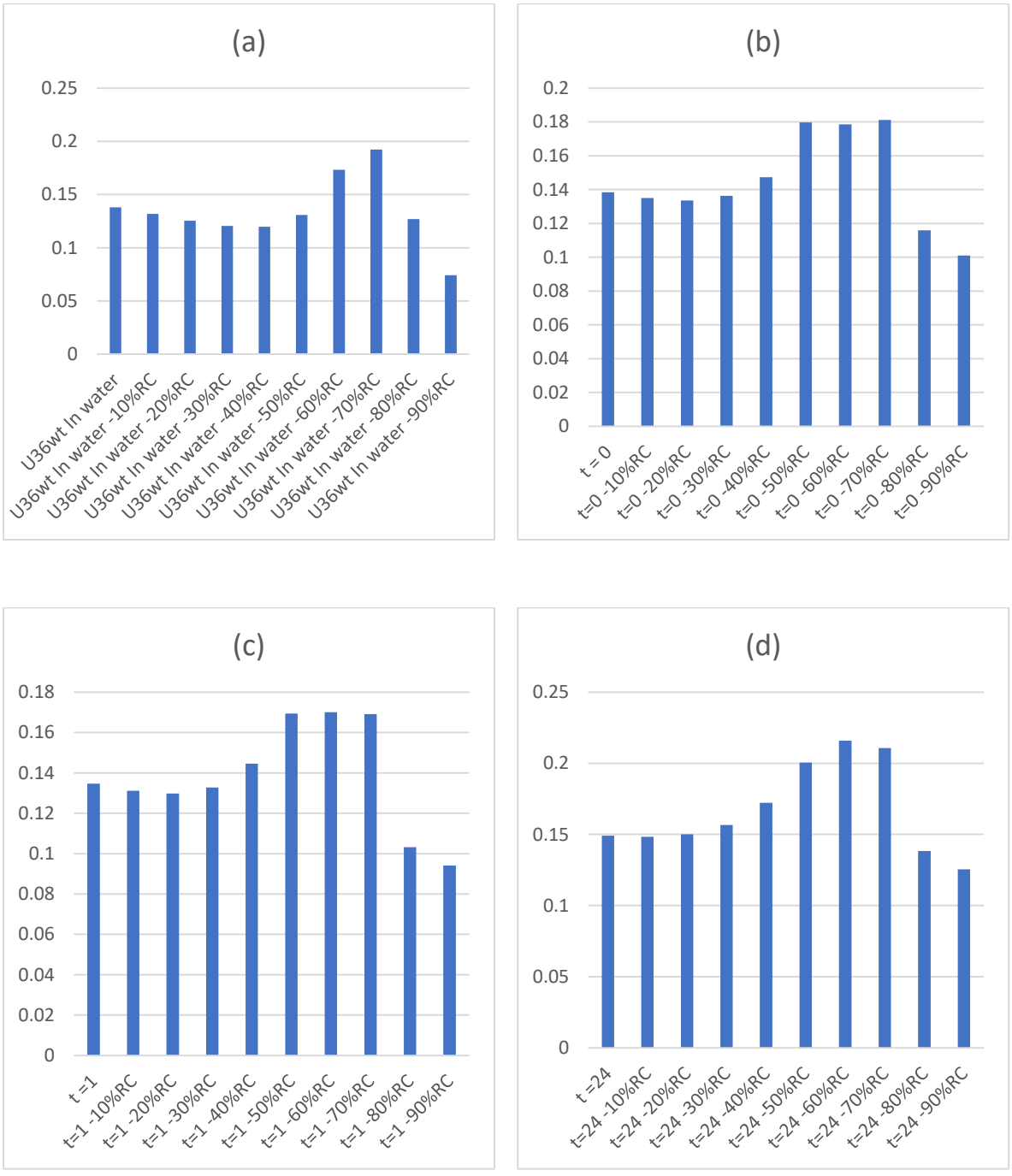


Figure S17 NRMSDs for the fitting of the U3.6 wt model spectrum to experimental data as a function of fraction of RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

Figures S18 shows the spectral NRMSDs of the best model spectrum from the experimental U3.6 K7A spectra with varying percentage of RC removed before and after addition of buffer at t = 0 to 24 h while using reference set and test with data down to 200 nm for SOMSpec training and prediction respectively

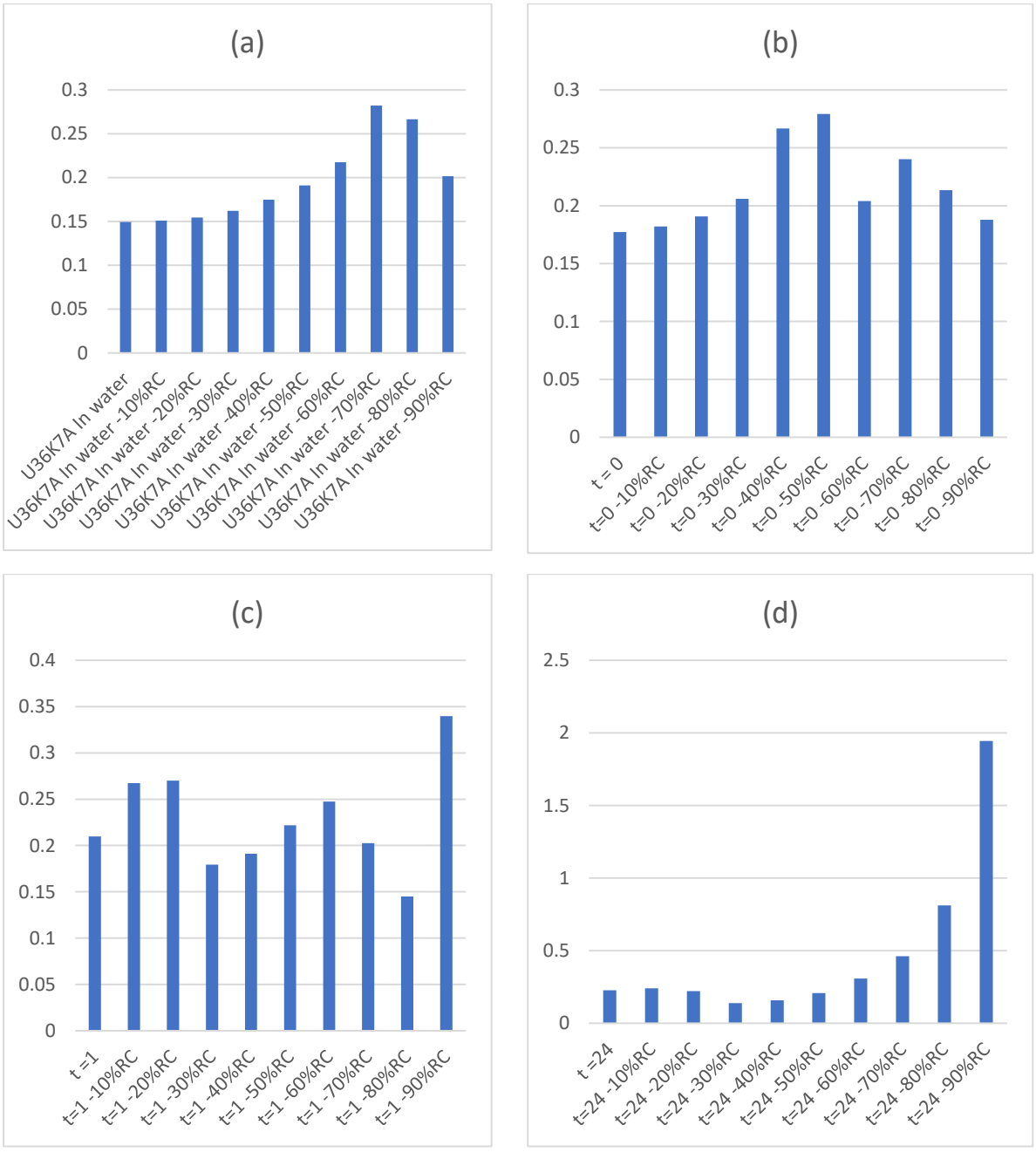
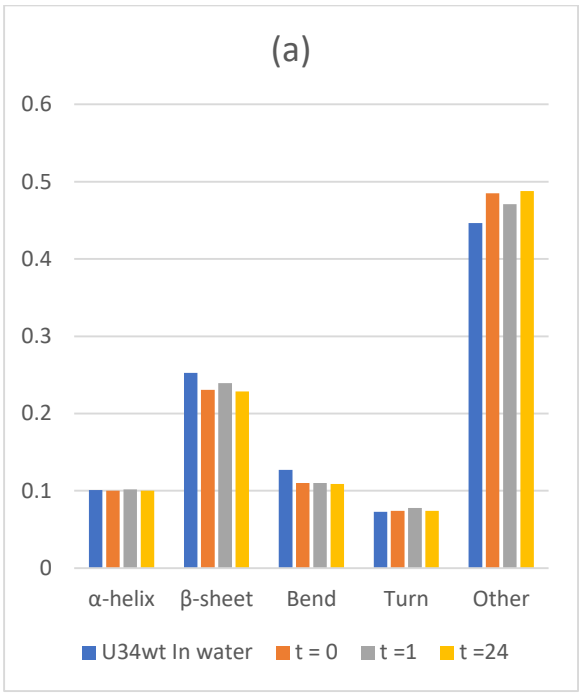
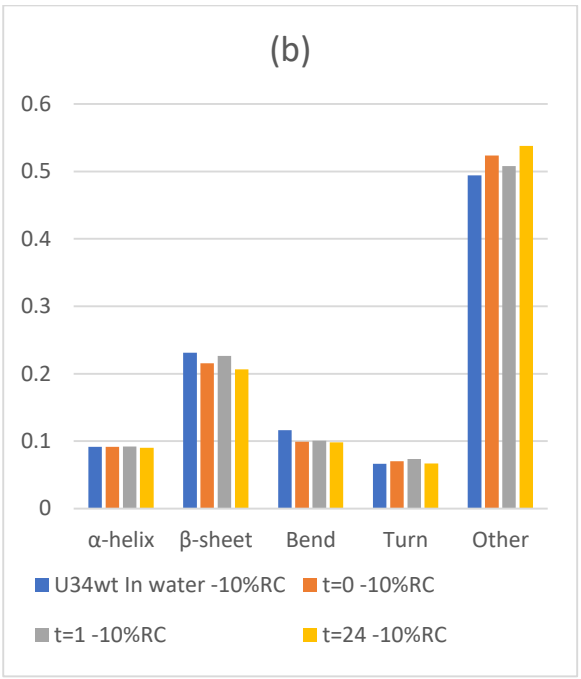
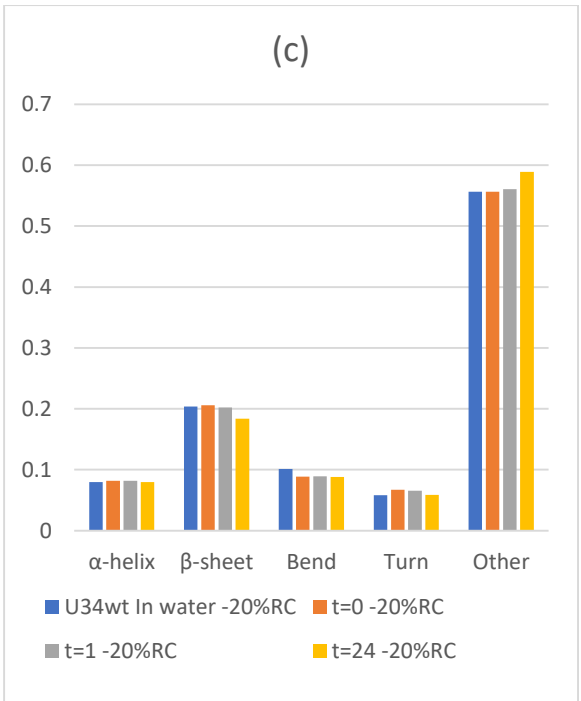


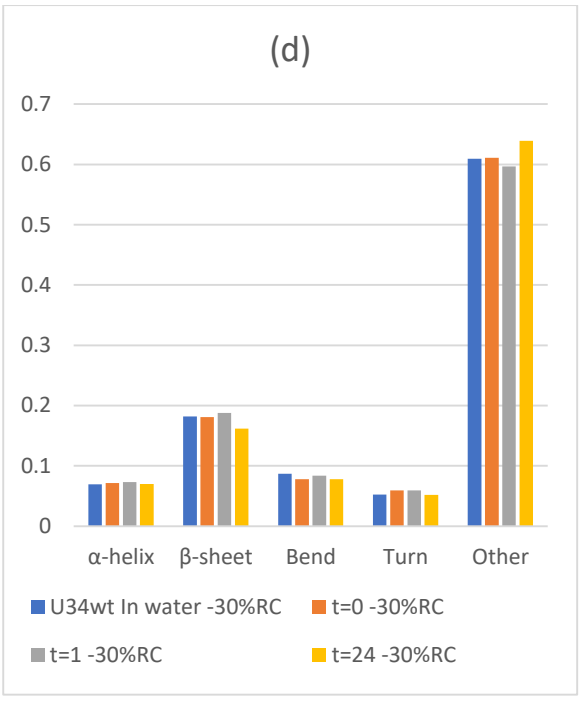
Figure S18 NRMSDs for the fitting of the U3.6 K7A model spectrum to experimental data as a function of percentage RC removed from the original spectrum in (a) water, and after buffer addition at (b) t = 0, (c) t = 1, (d) and t = 24 h.

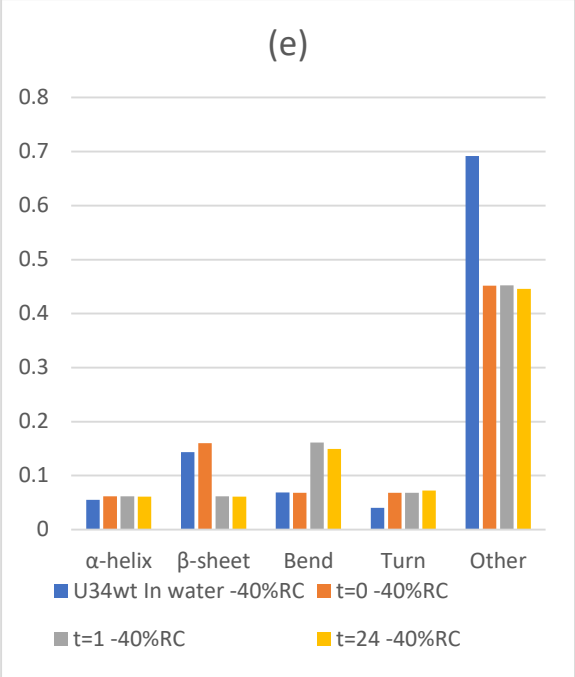
The SOMSpec secondary structure predictions for original U3.4wt data in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted and tested against reference data down to 190 nm are shown in Figures S19. Figure S19a is for the original protein, whereas Figures S19 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

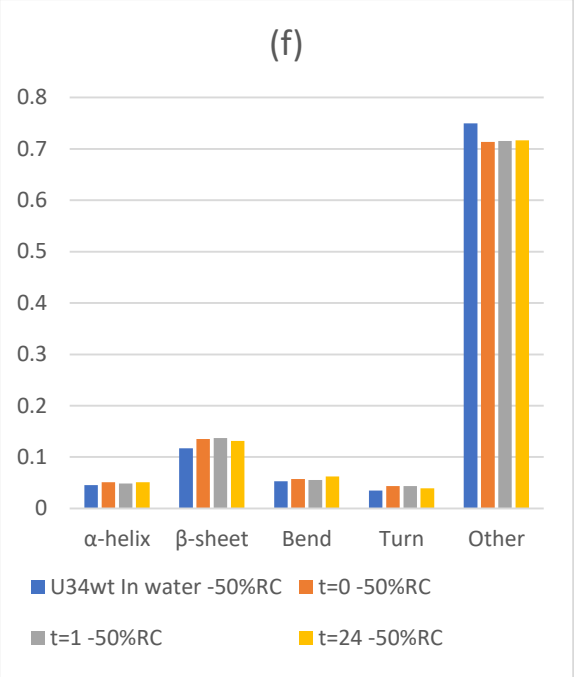


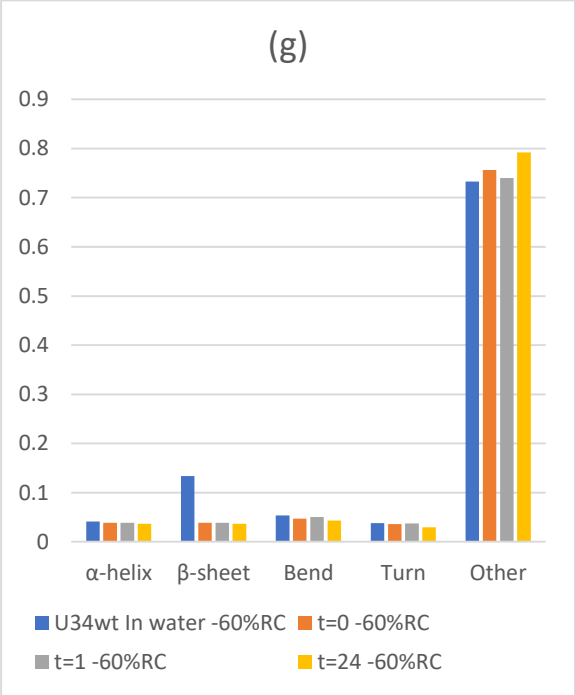


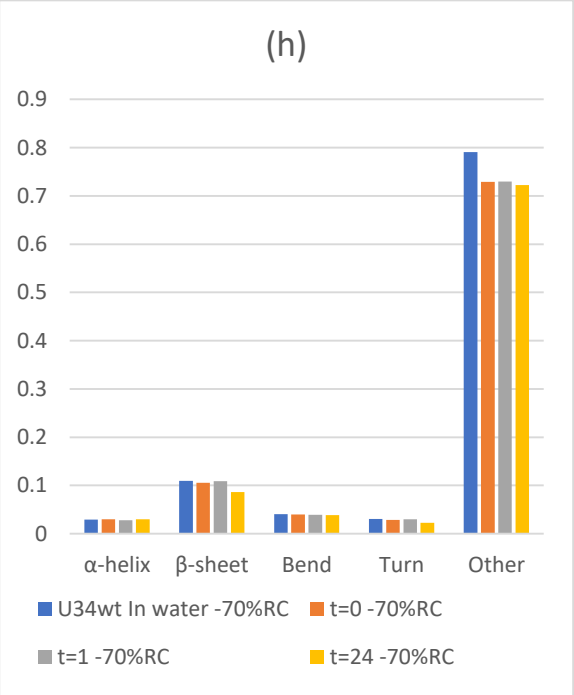





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

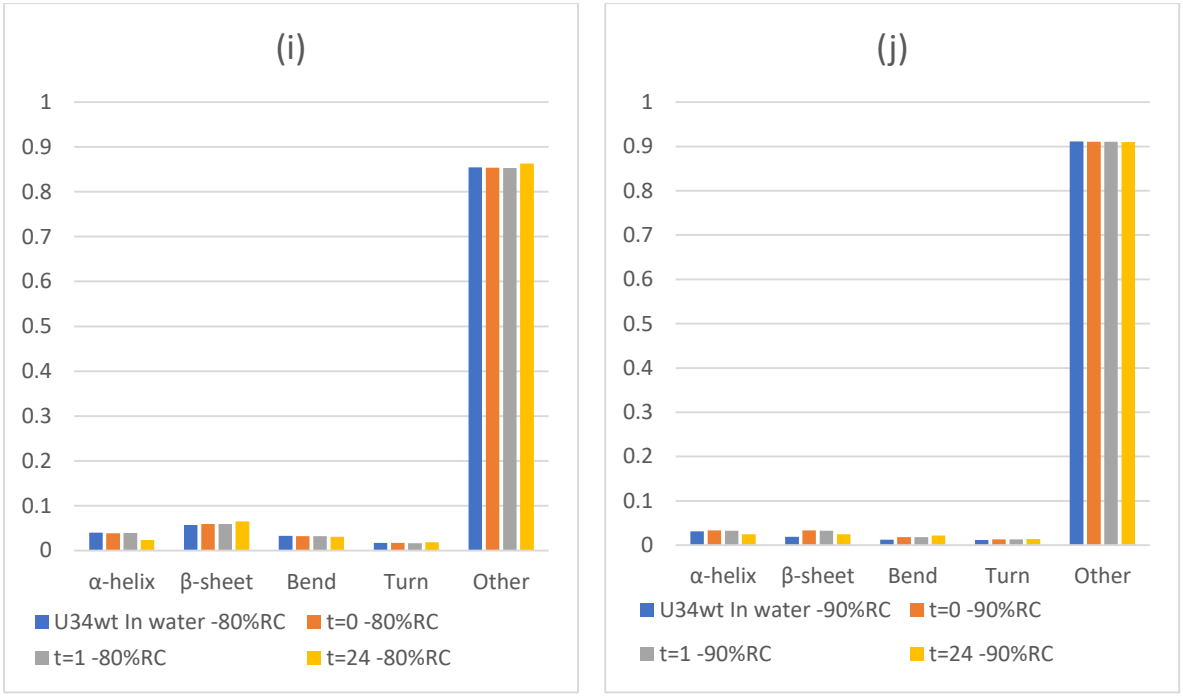
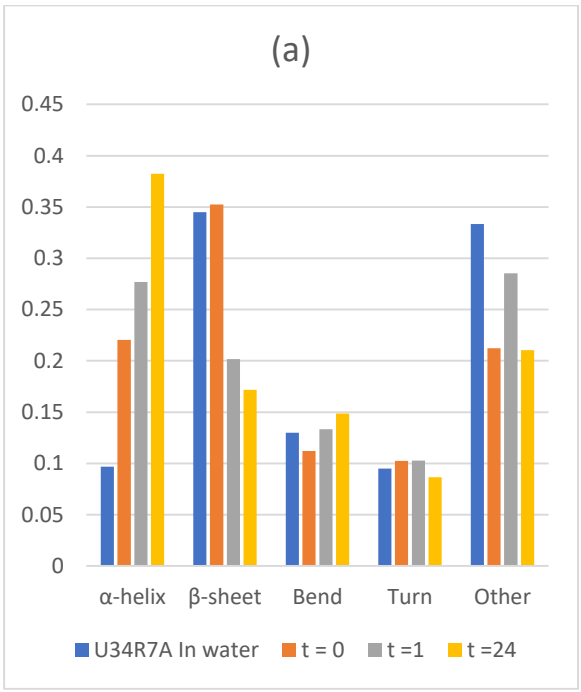
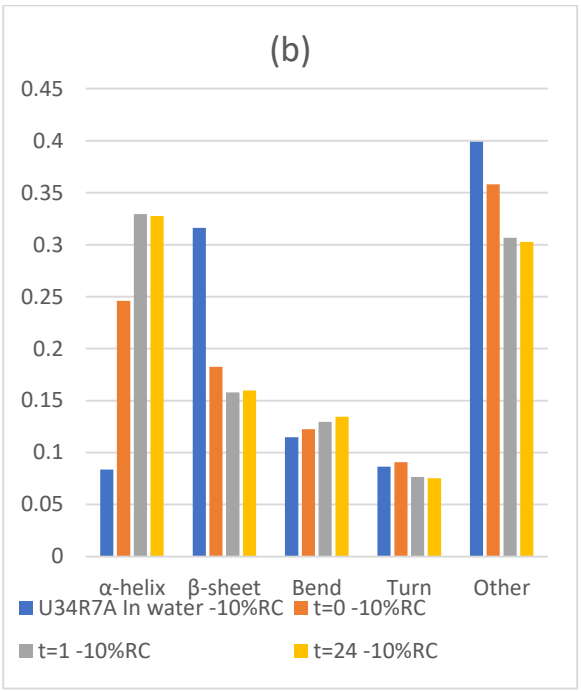
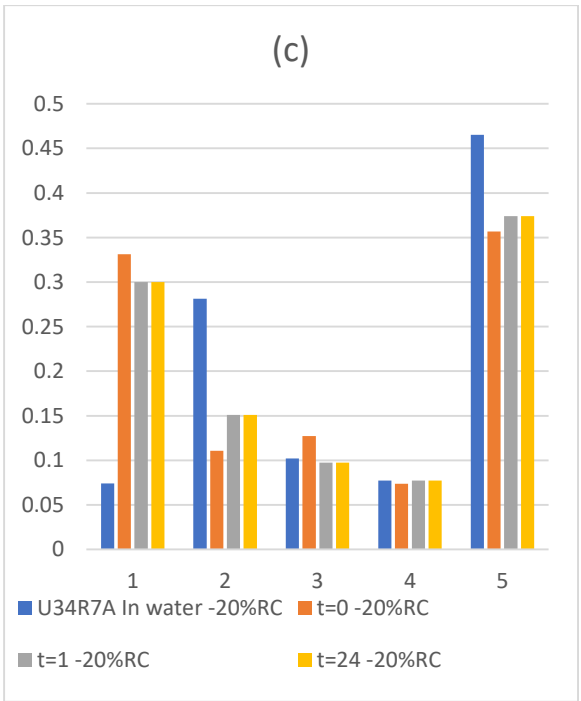


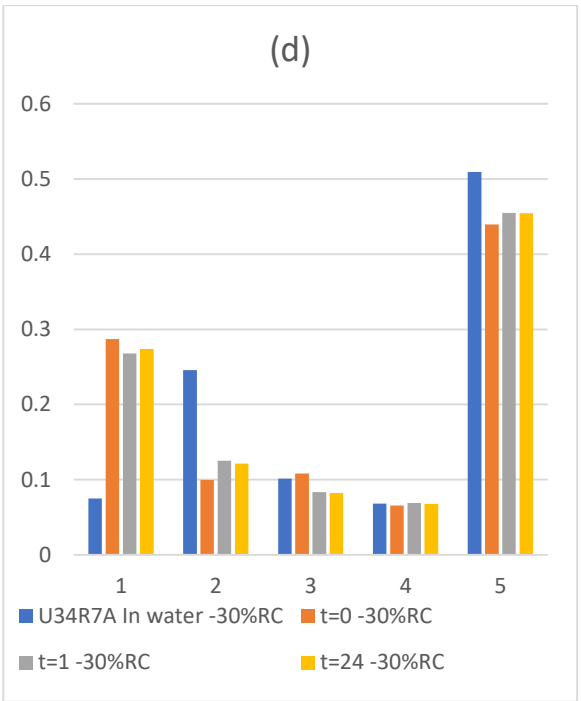
Figure S19 Secondary structure predictions for regenerated U3.4wt in water only and after addition of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4wt predictions.

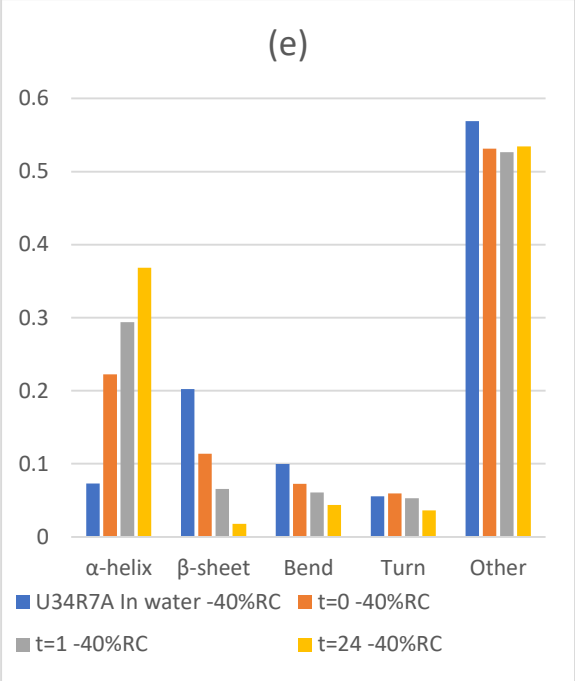
The SOMSpec secondary structure predictions for original U3.4 R7A data in the absence and/or presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference data down to 190 nm are shown in Figures S20. Figure S20a is for the original protein, whereas Figures S20 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

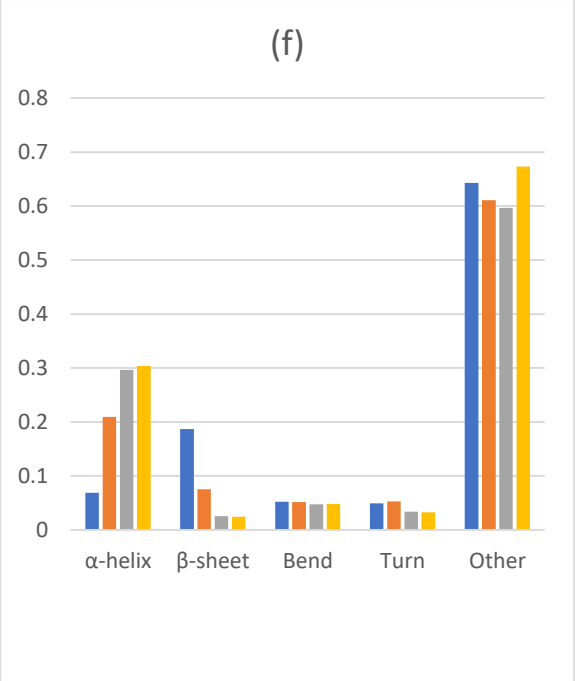


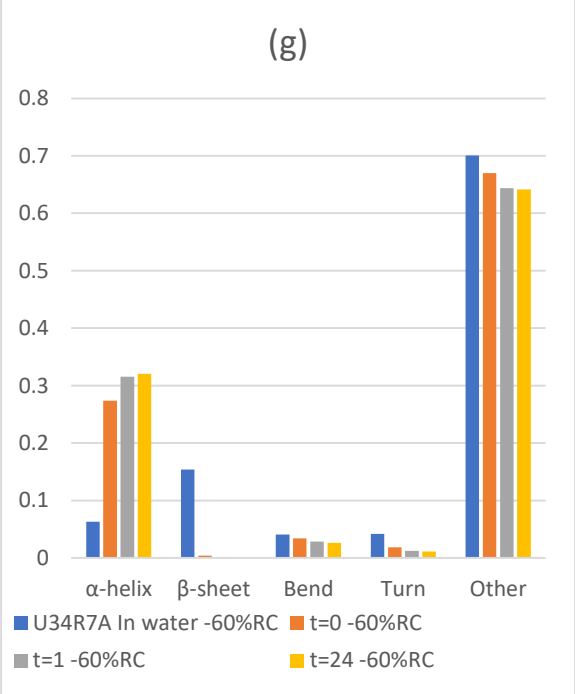


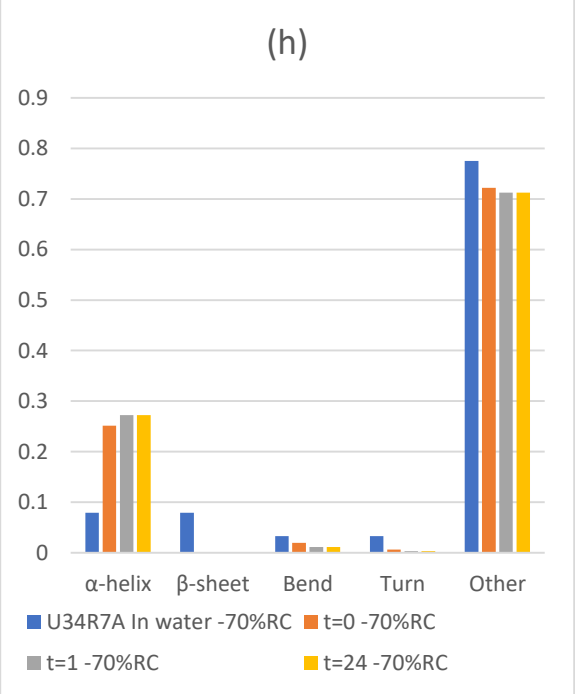





(e)
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

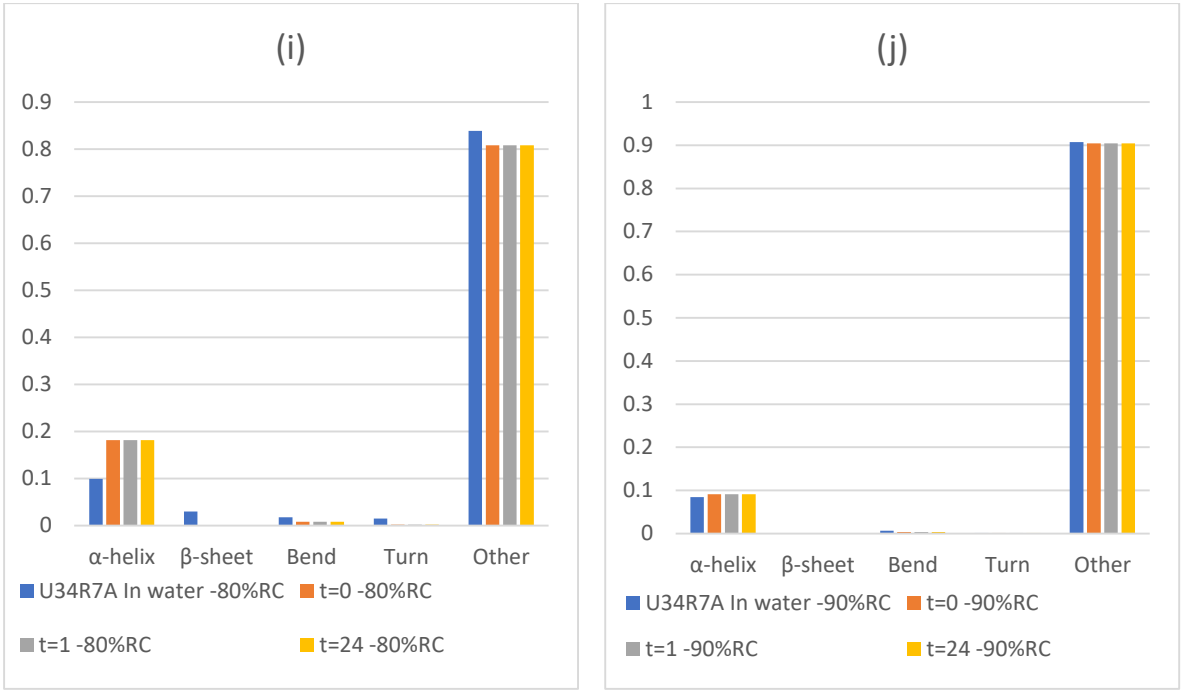
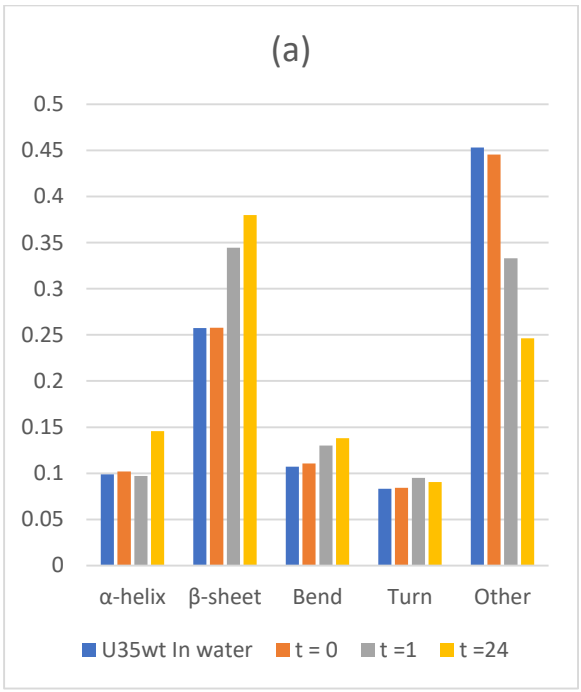
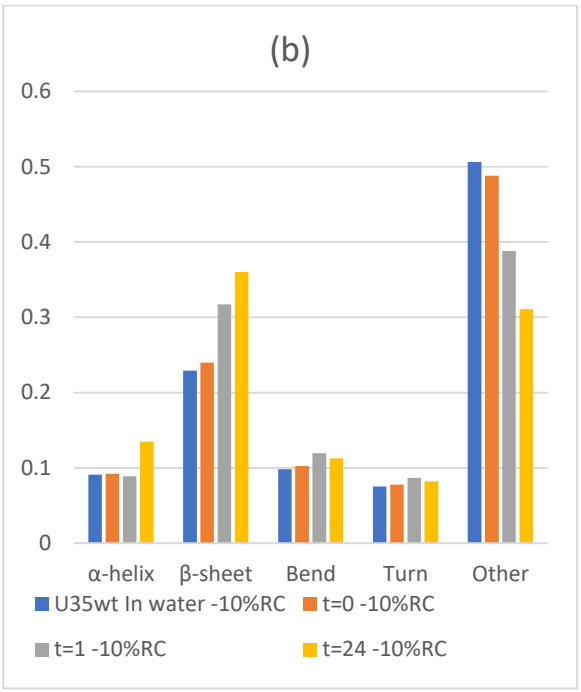
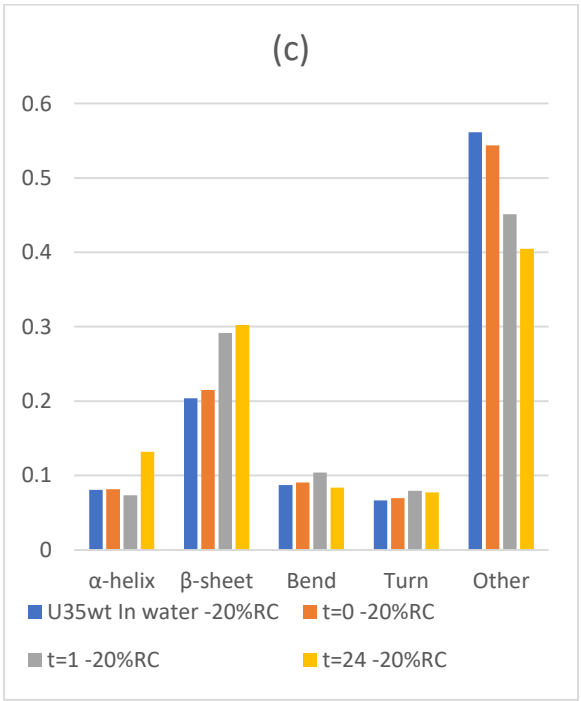


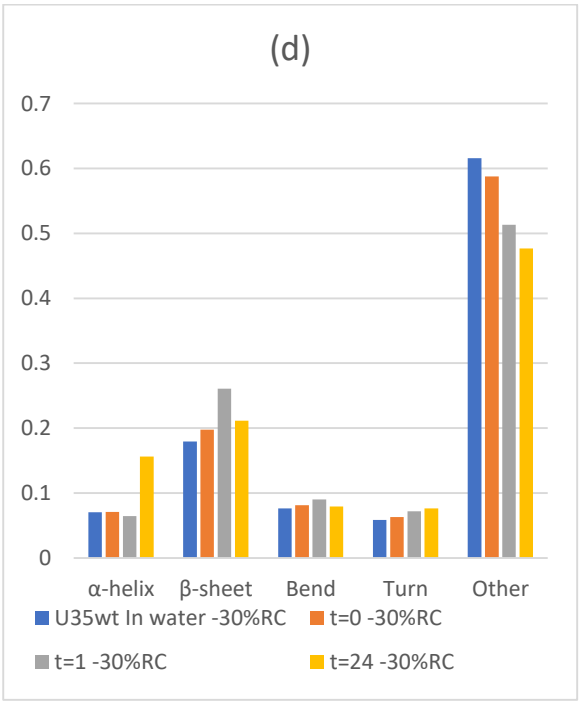
Figure S20 Secondary structure predictions for regenerated U3.4 R7A in water only and after the addition of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 R7A predictions.

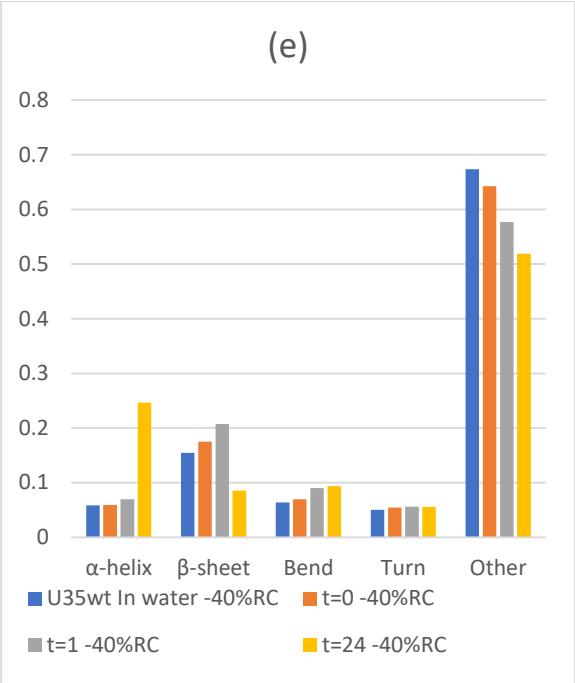
The SOMSpec secondary structure predictions for original U3.5 wt data in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference data down to 190 nm are shown in Figures S21. Figure S21a is for the original protein, whereas Figures S21 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

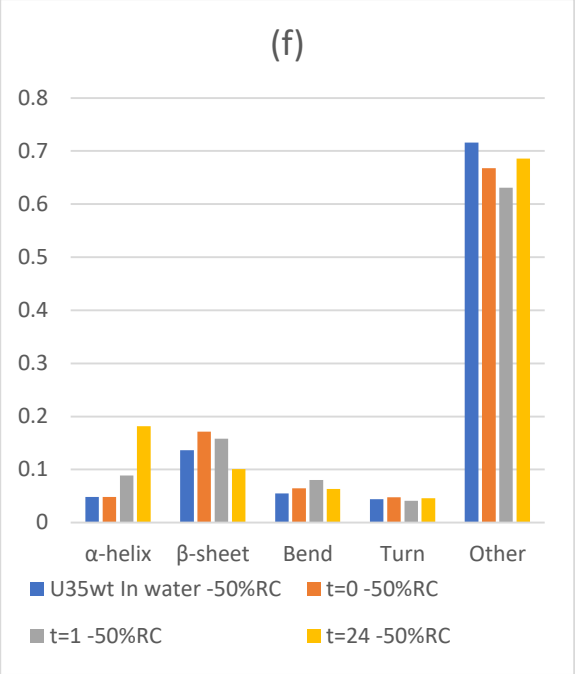


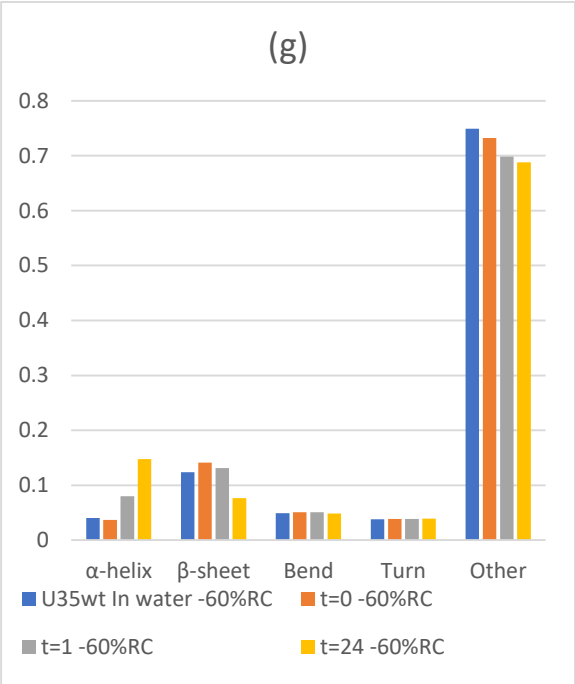


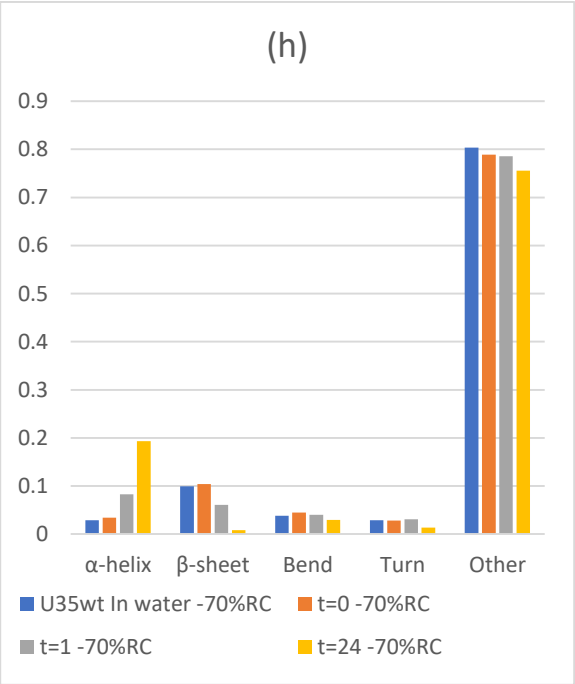





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

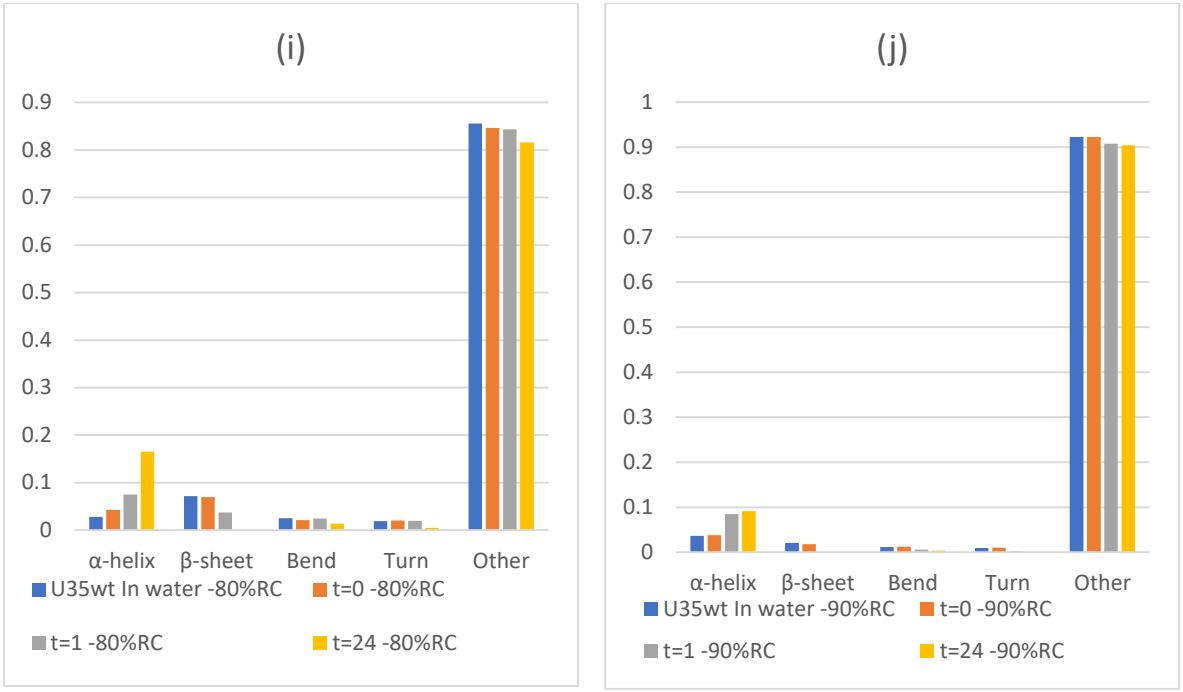
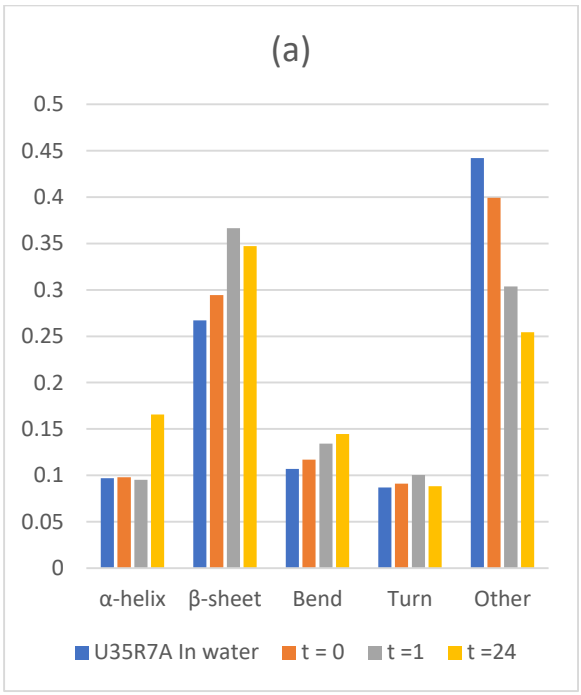
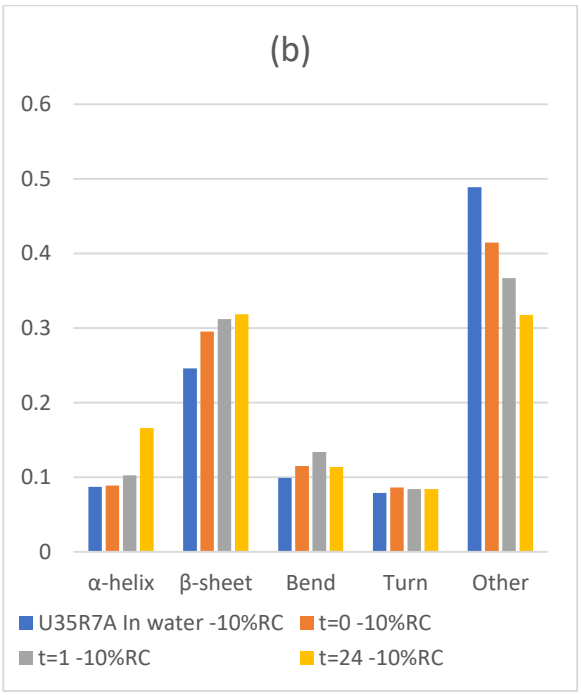
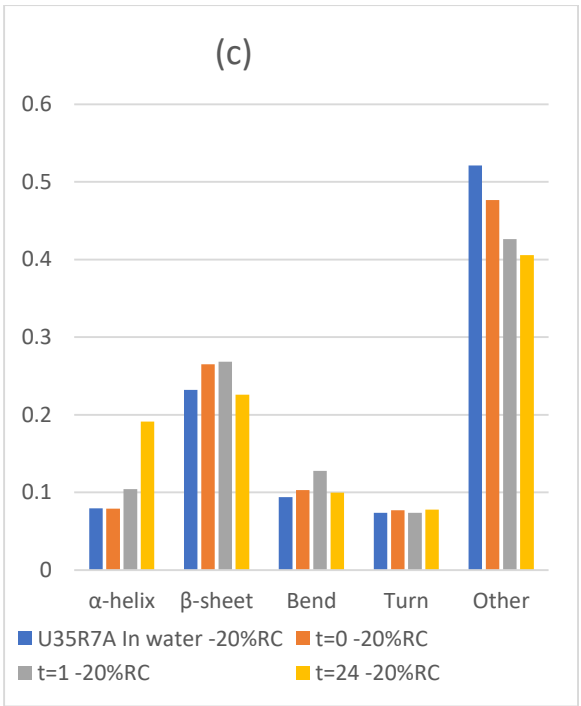


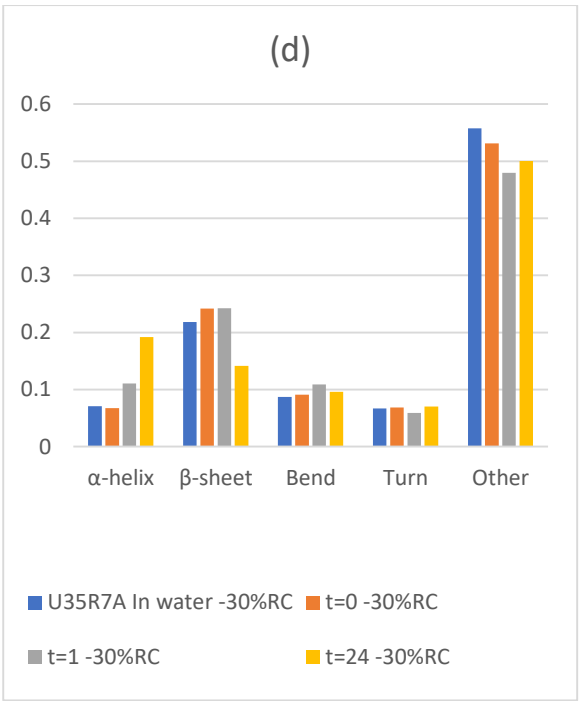
Figure S21. Secondary structure predictions for regenerated U3.5 wt in water only and after addition of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 wt predictions.

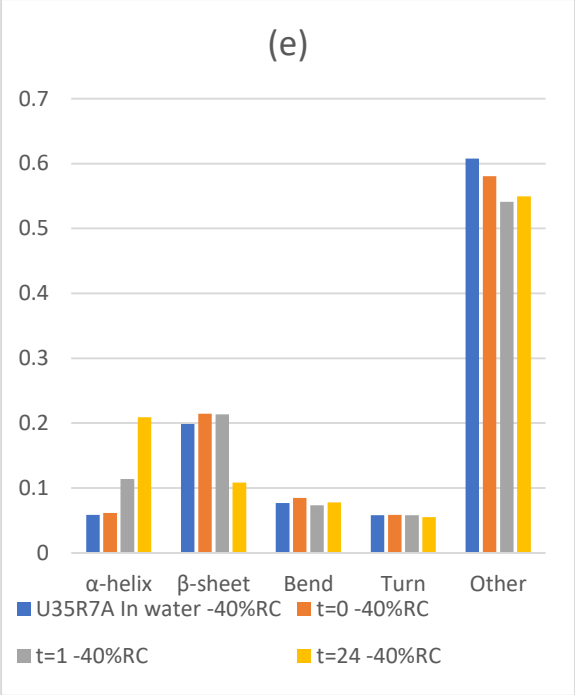
The SOMSpec secondary structure predictions for original U3.5 R7A data in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference data down to 190 nm are shown in Figures S22. Figure S22a is for the original protein, whereas Figure S22 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

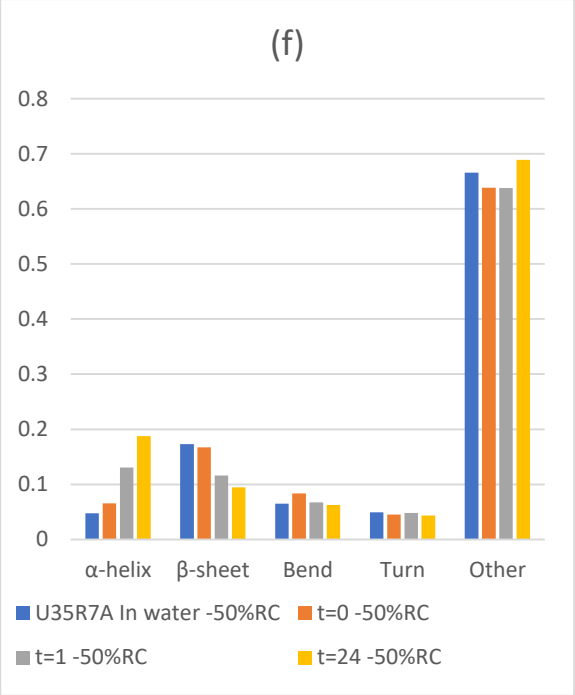


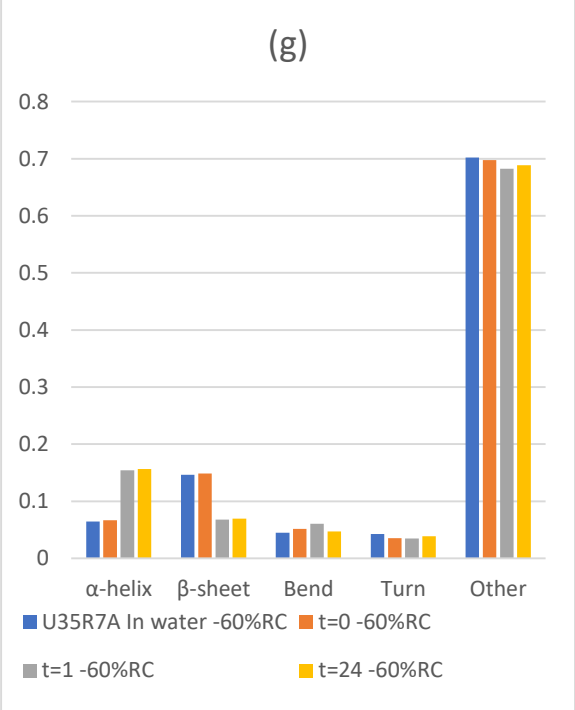


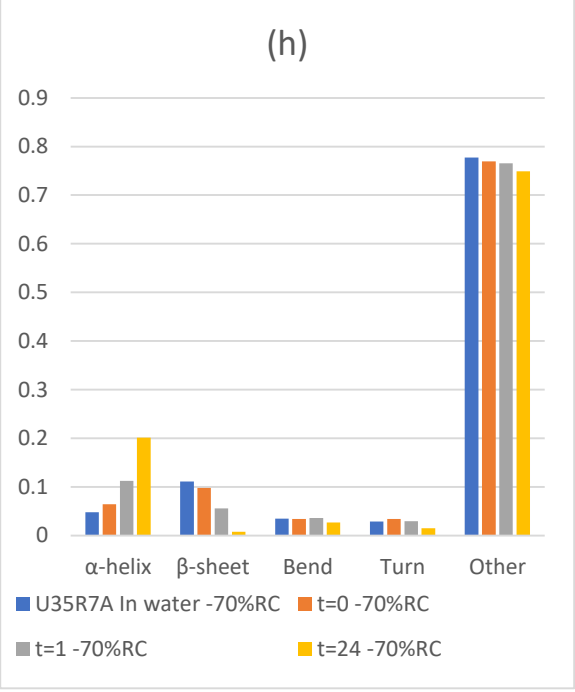





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

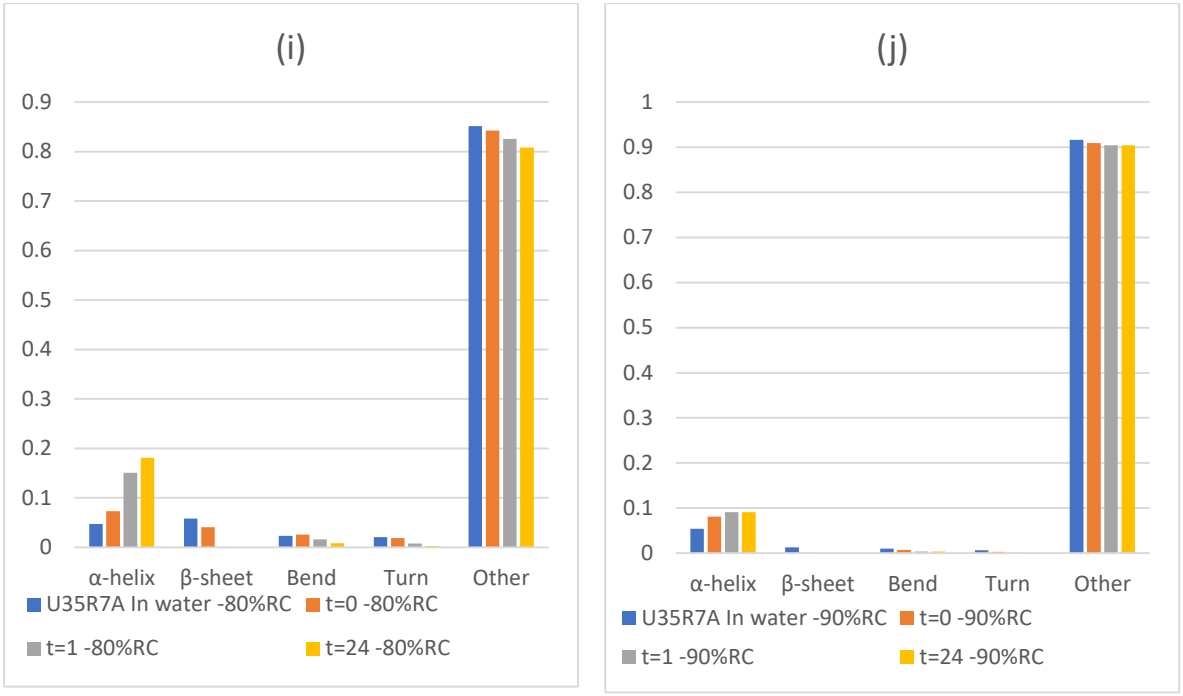
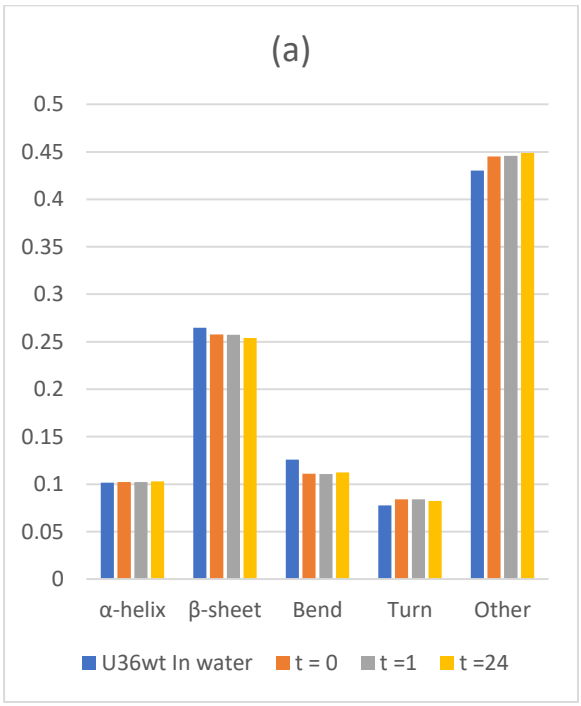
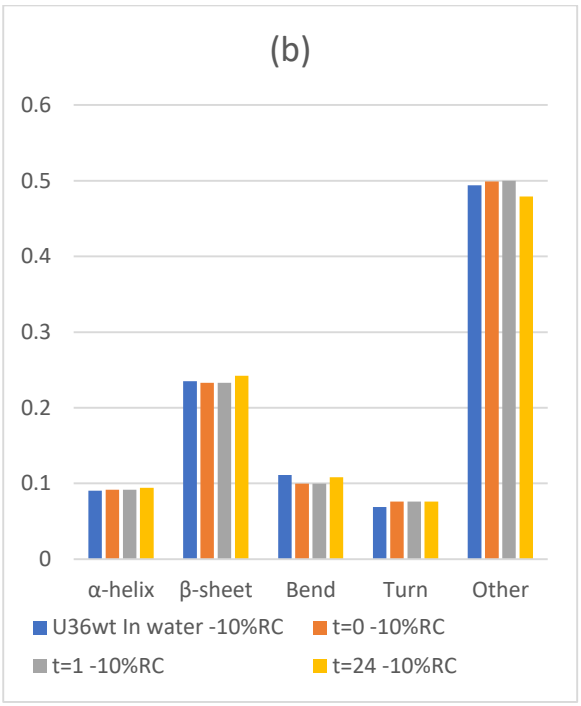
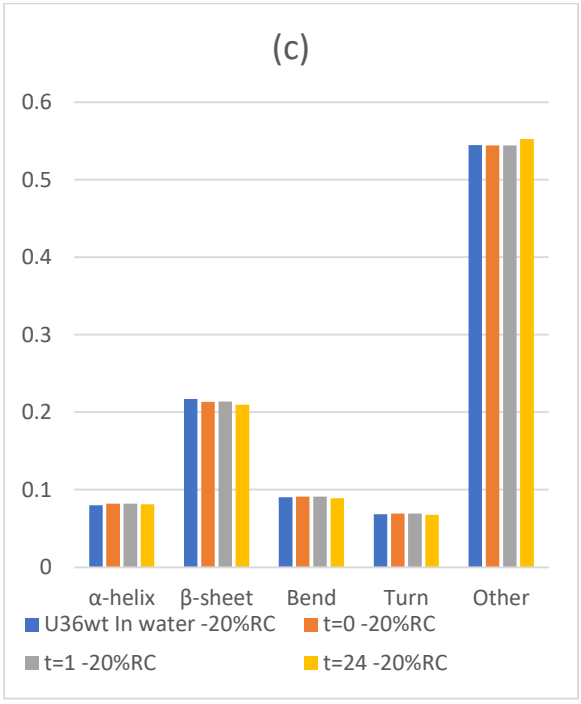


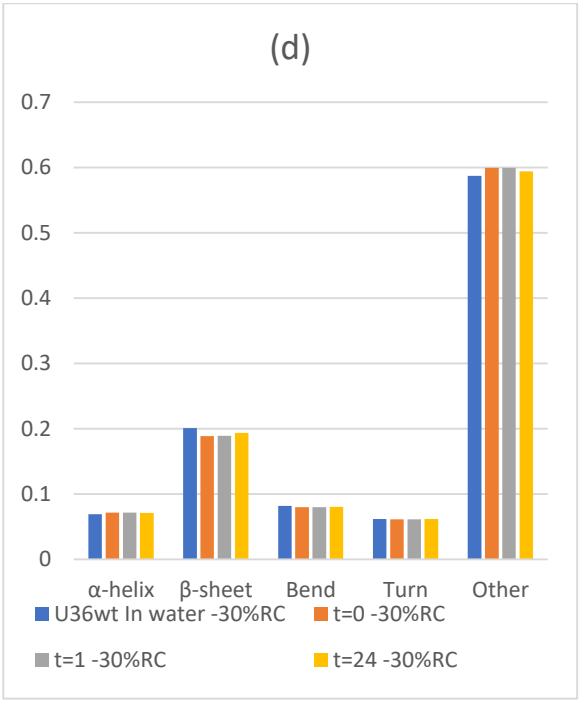
Figure S22 Secondary structure predictions for regenerated U3.5 R7A in water only and after addition of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 R7A predictions.

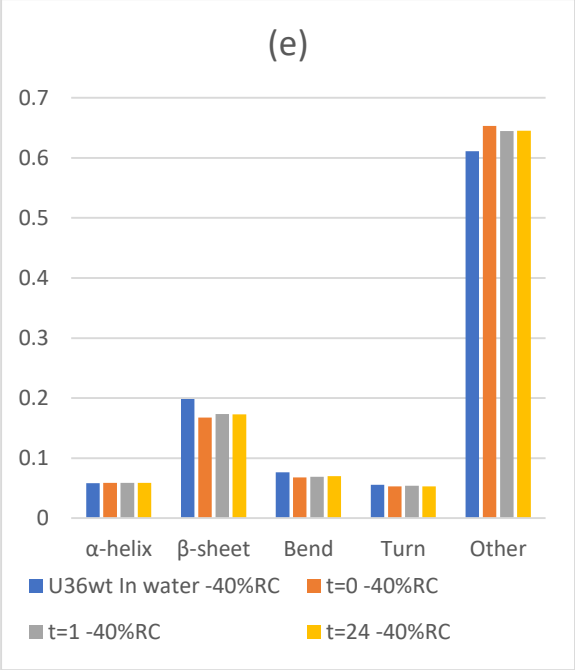
The SOMSpec secondary structure predictions for original U3.6 wt data in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference data down to 190 nm are shown in Figures S23. Figure S23a is for the original protein, whereas Figures S23 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

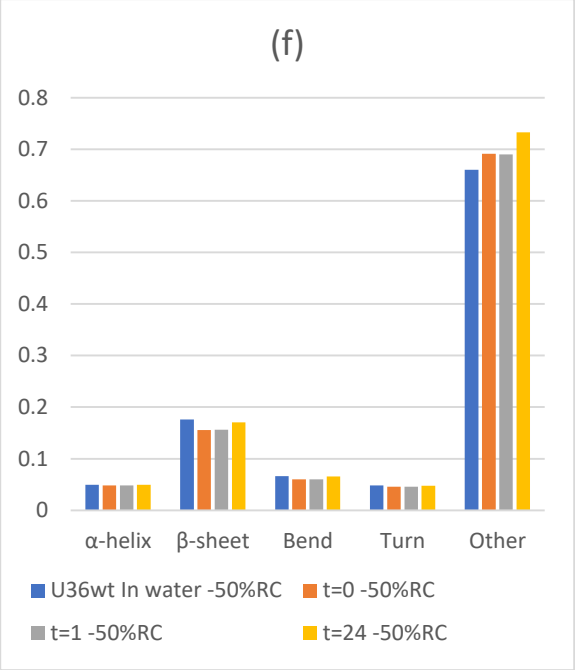


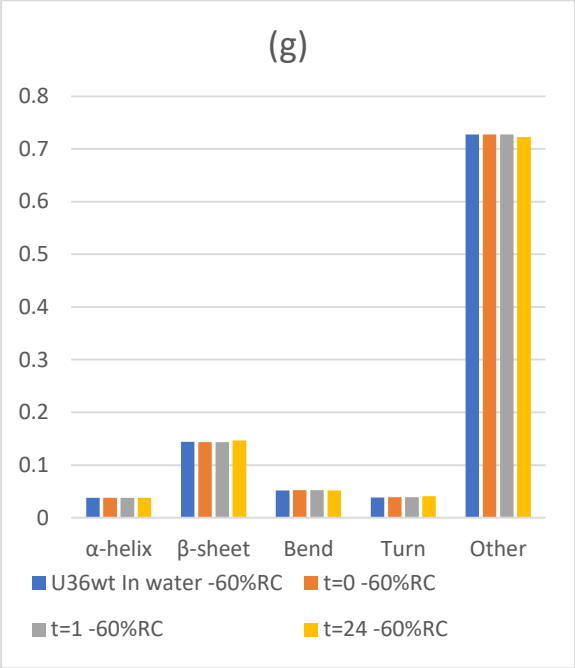


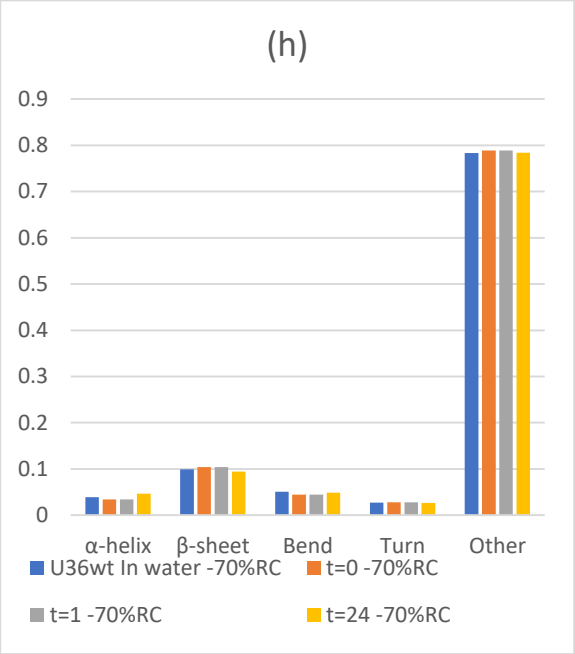





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

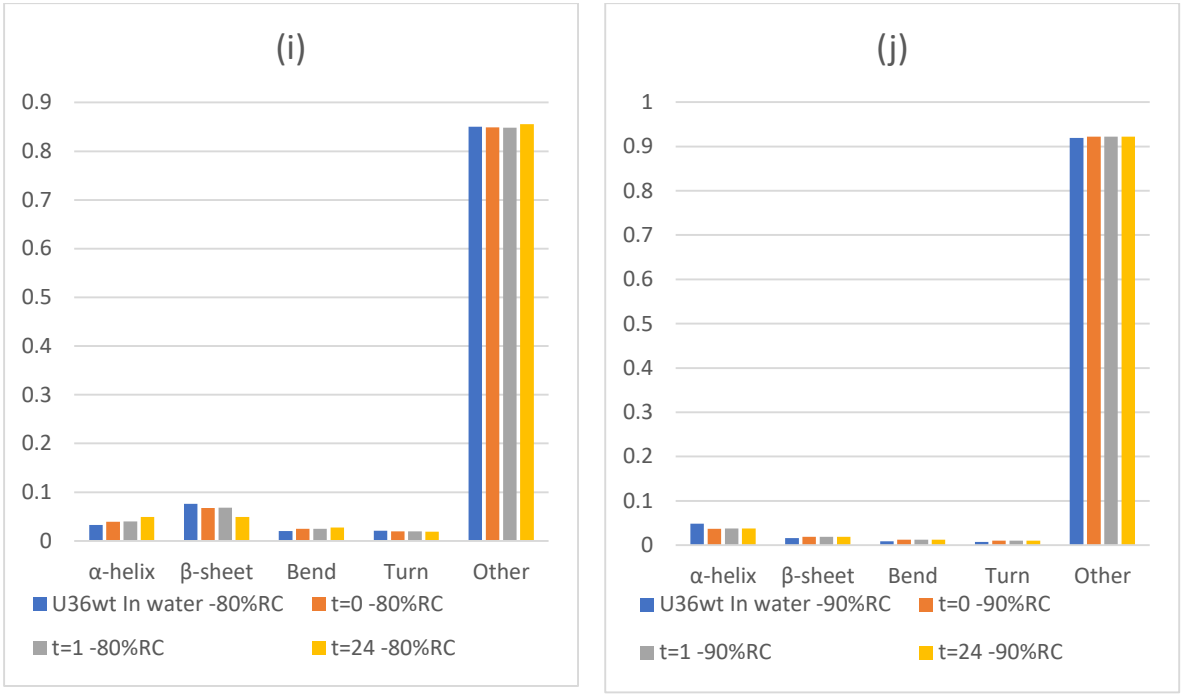
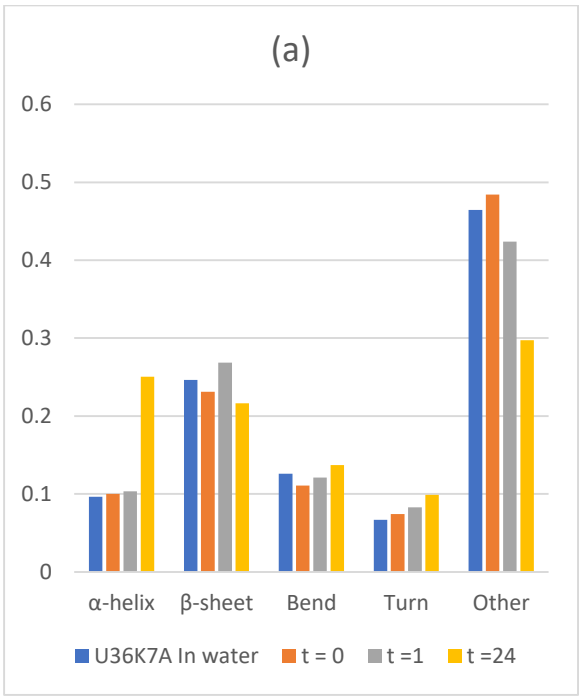
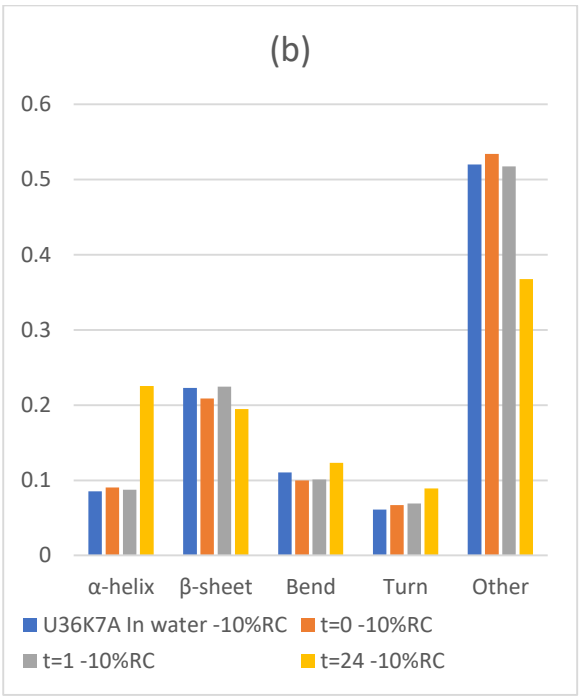
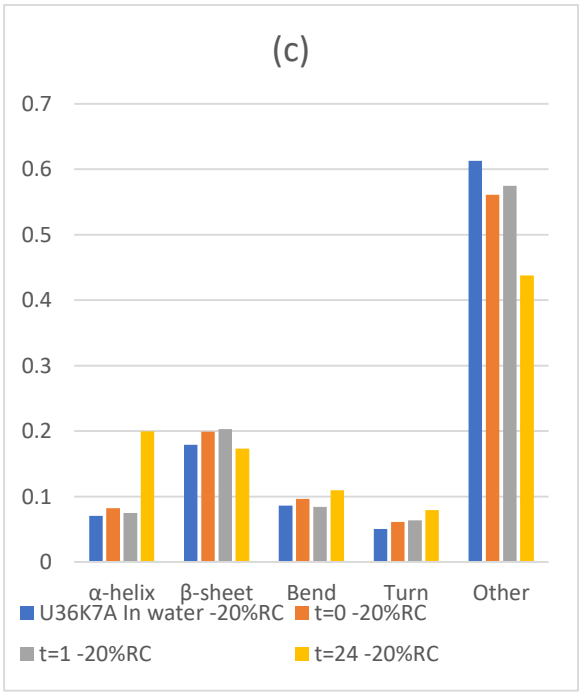


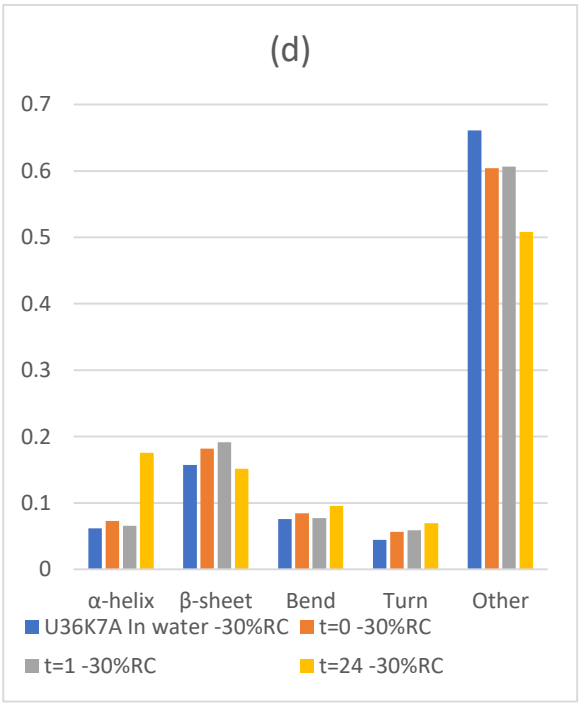
Figure S23 Secondary structure predictions for regenerated U3.6 wt in water only and after addition of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

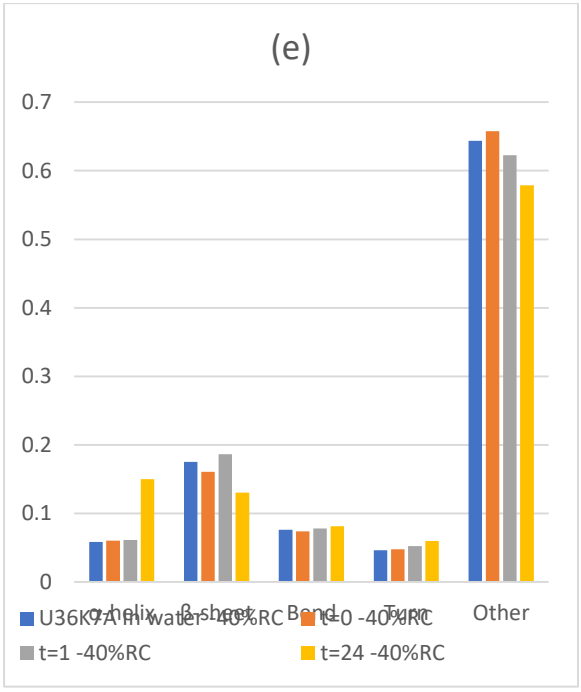
The SOMSpec secondary structure predictions for original U3.6 K7A data in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference data down to 190 nm are shown in Figures S24. Figure S24a is for the original protein, whereas Figures S24 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

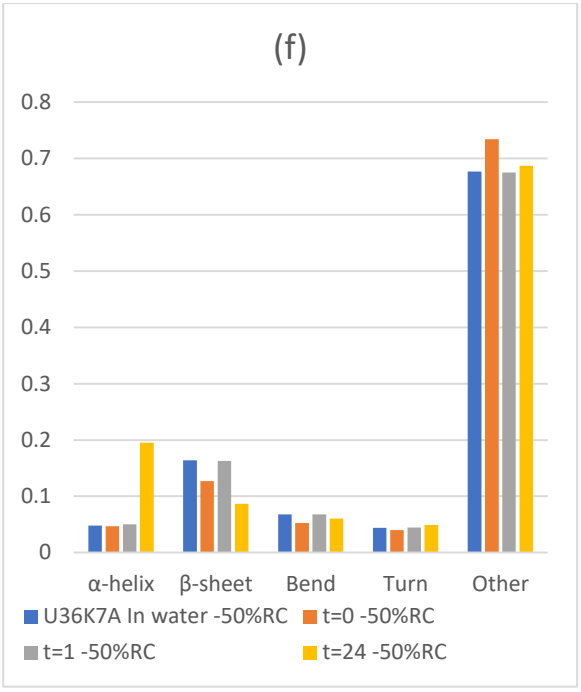


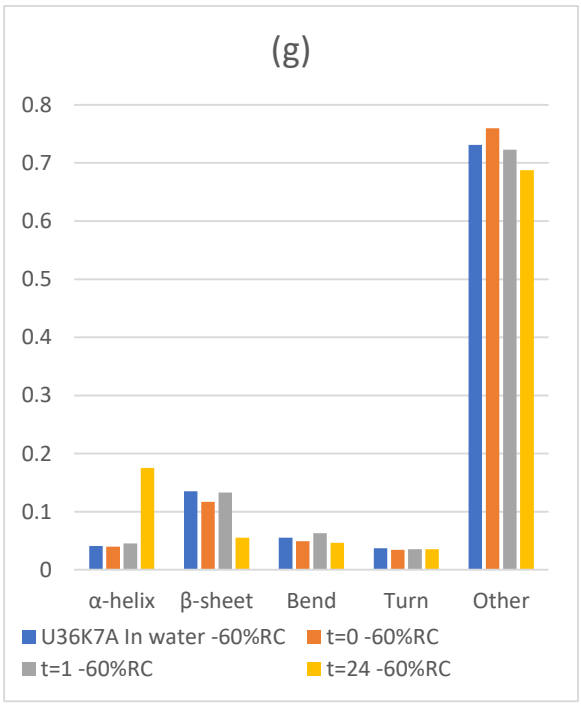


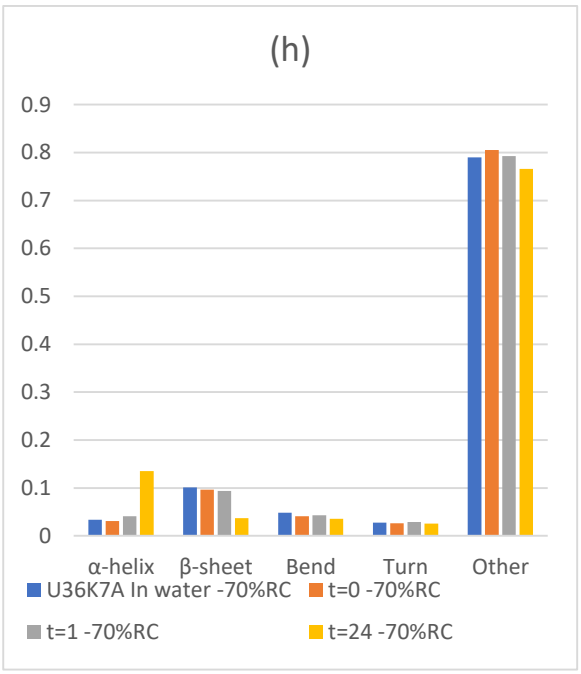





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

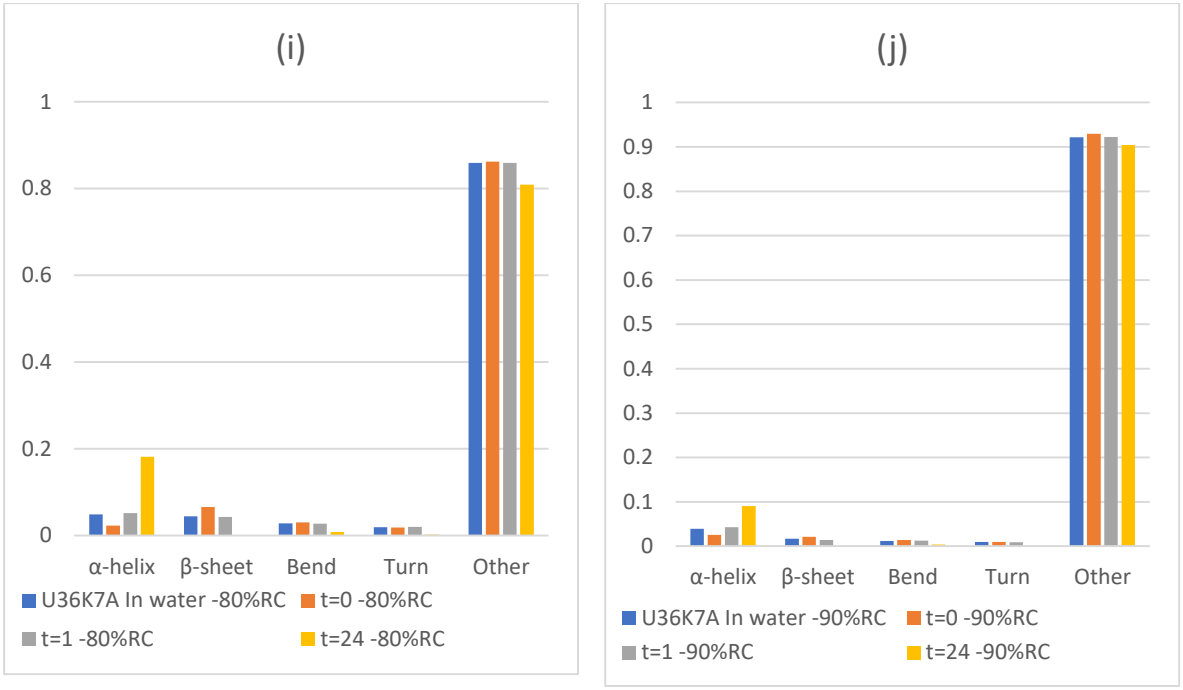
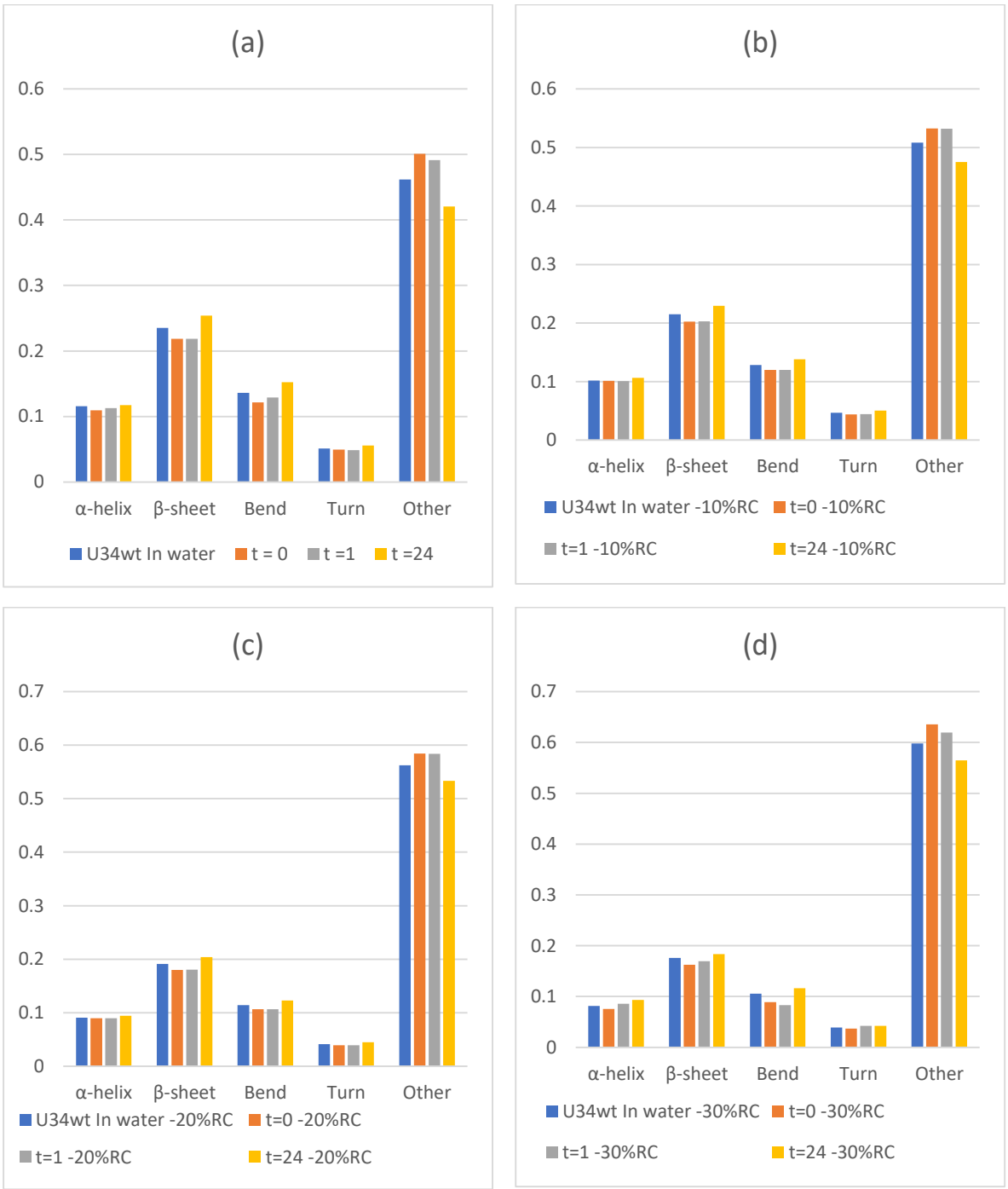
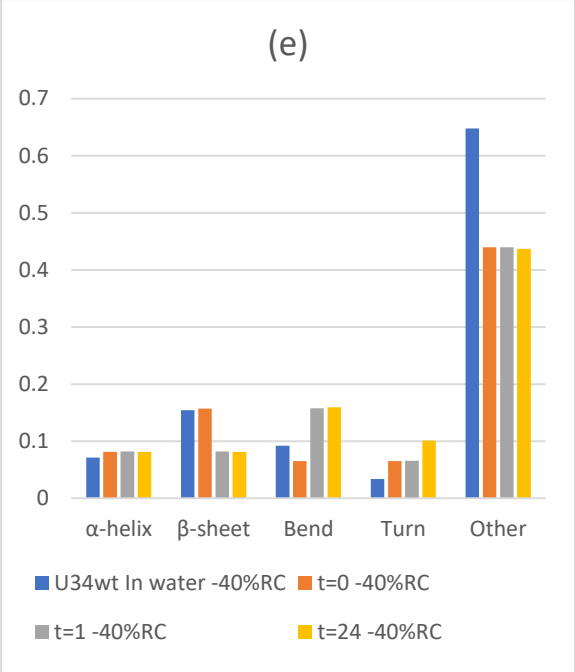
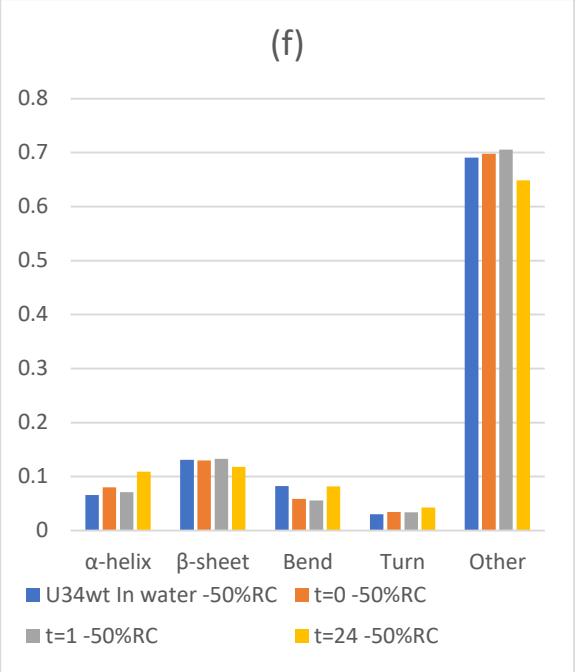


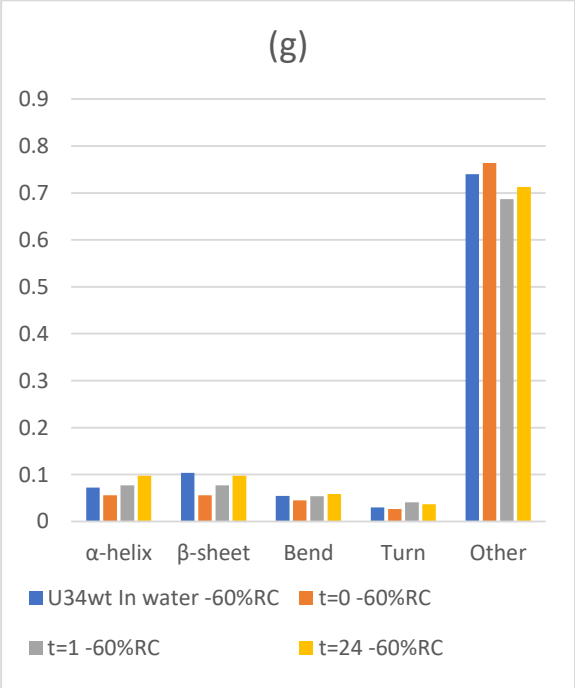
Figure S24 Secondary structure predictions for regenerated U3.6 K7A in water only and after of buffer over time. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 K7A predictions.

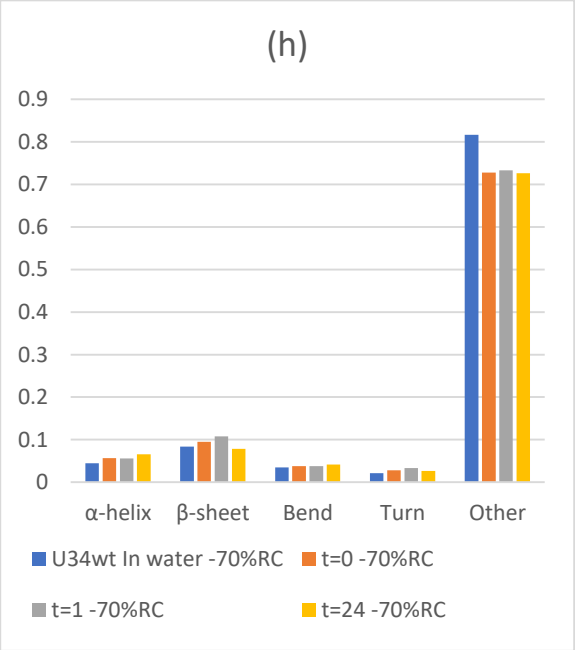
The SOMSpec secondary structure predictions for U3.4 wt data down to 195 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S25. Figure S25a is for the original protein, whereas Figures S25 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.




(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

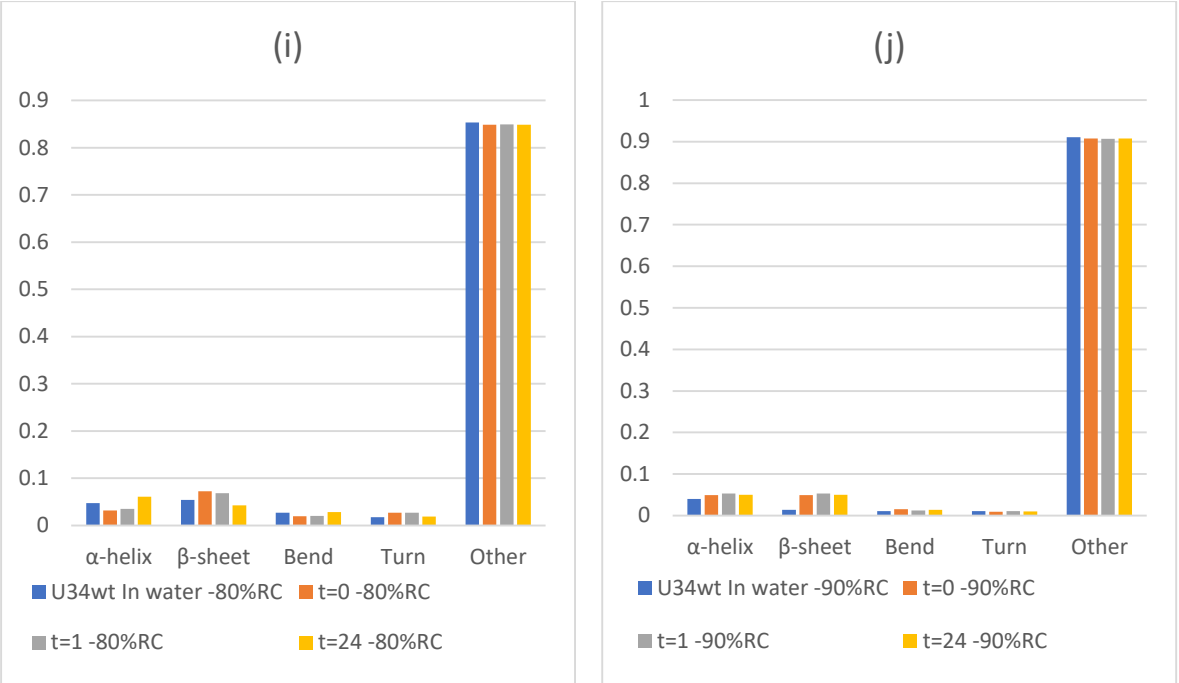
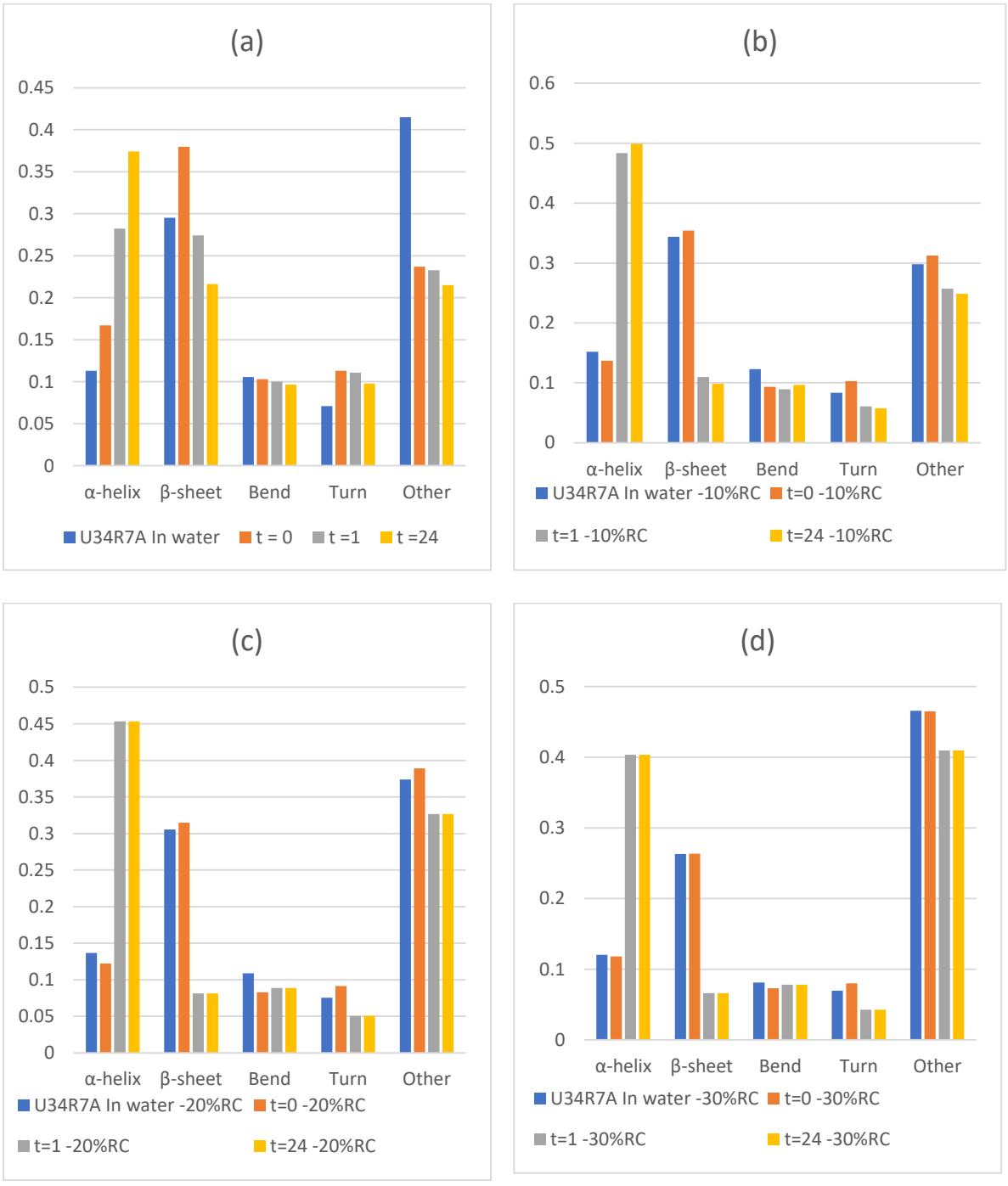
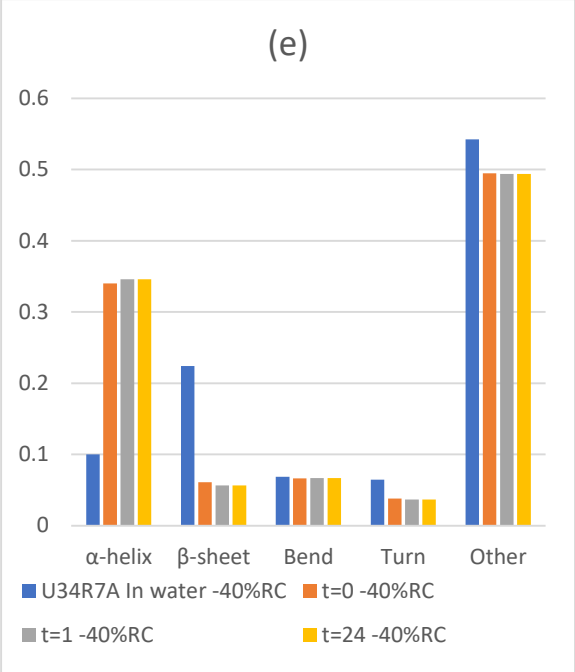
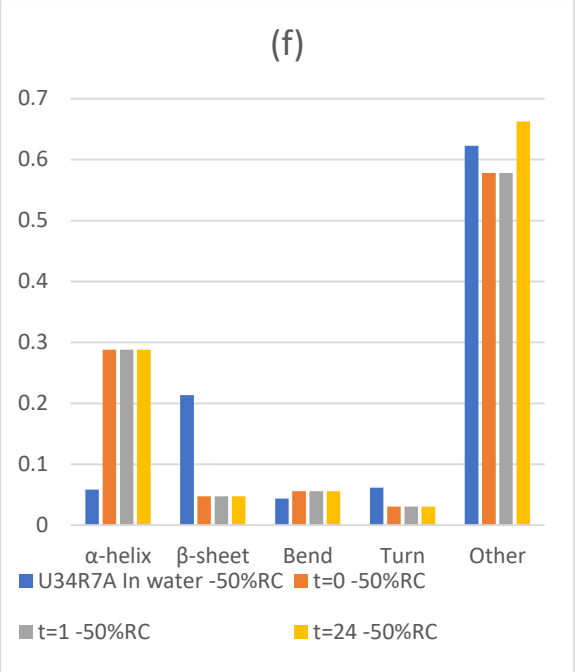


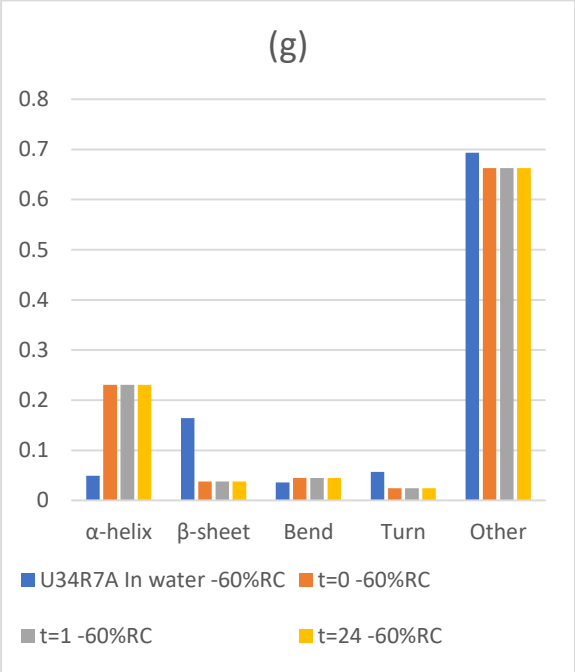
Figure S25 Secondary structure predictions for regenerated U3.4 wt in water only and after addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 wt predictions.

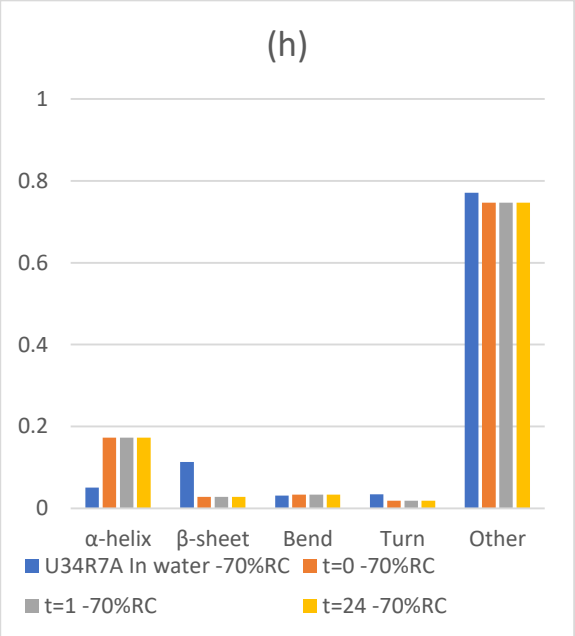
The SOMSpec secondary structure predictions for U3.4 R7A data down to 195 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S26. Figure S26a is for the original protein, whereas Figures S26 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.




(e)
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

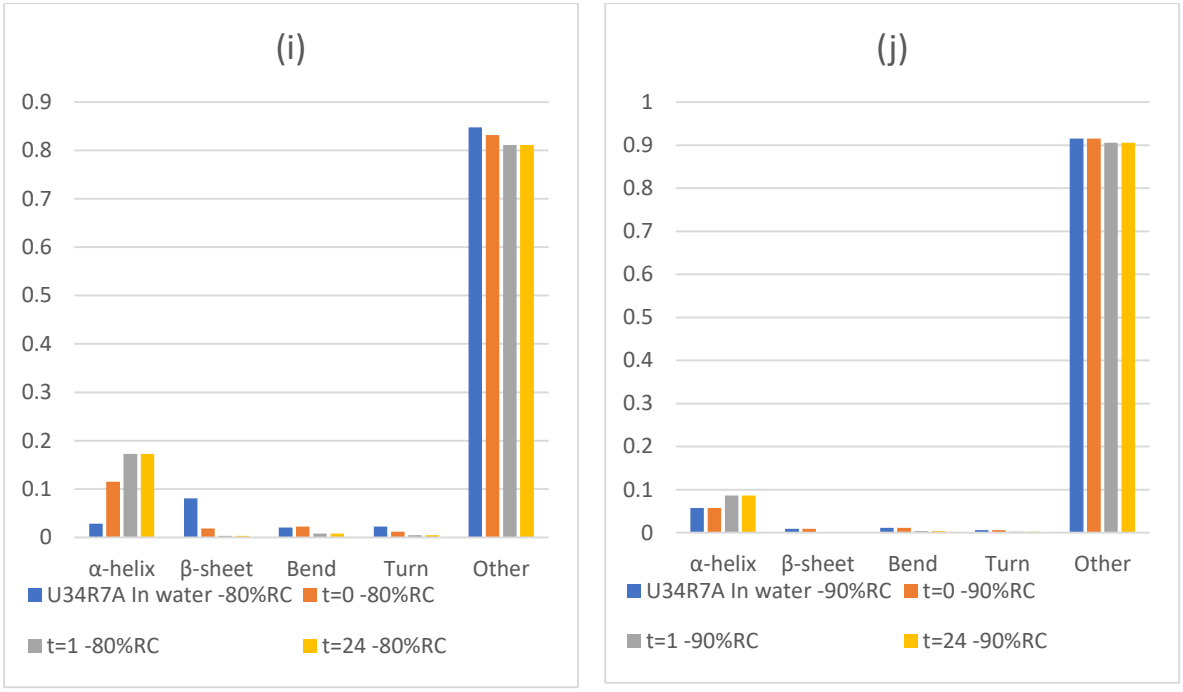
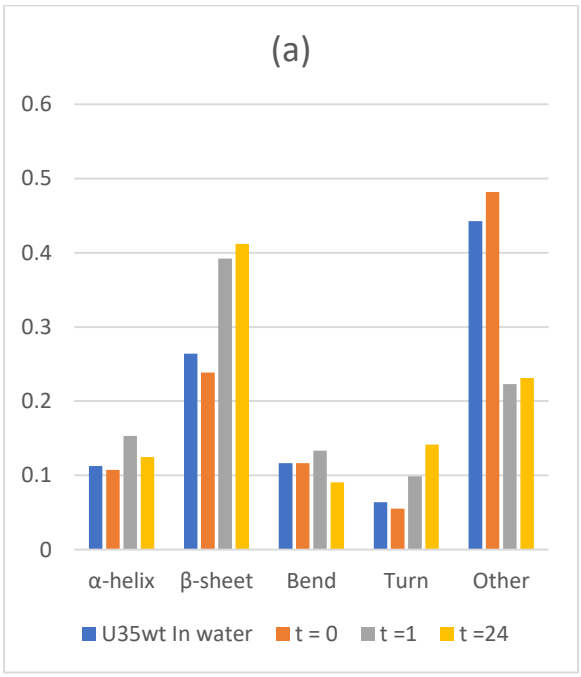
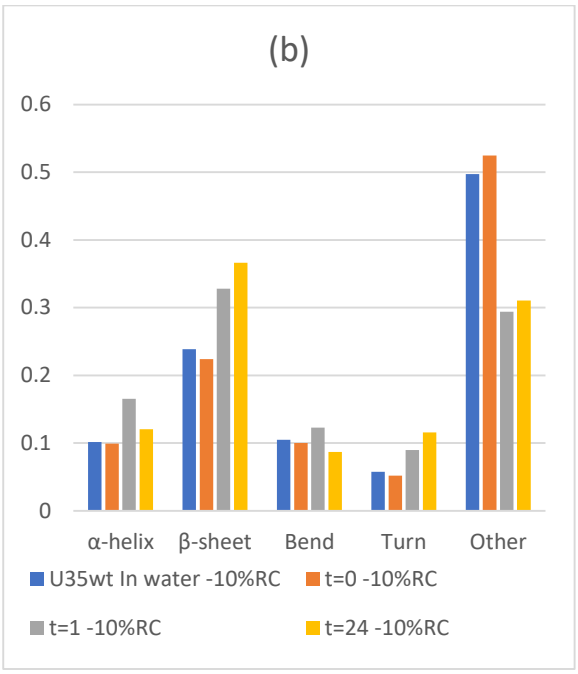
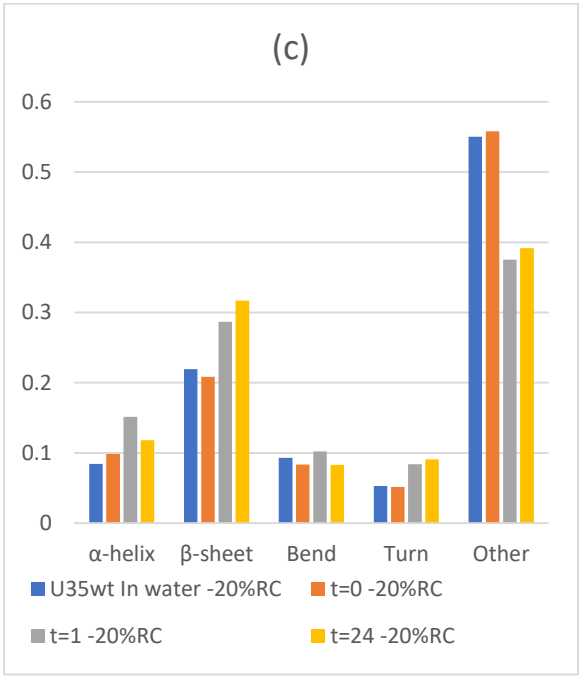


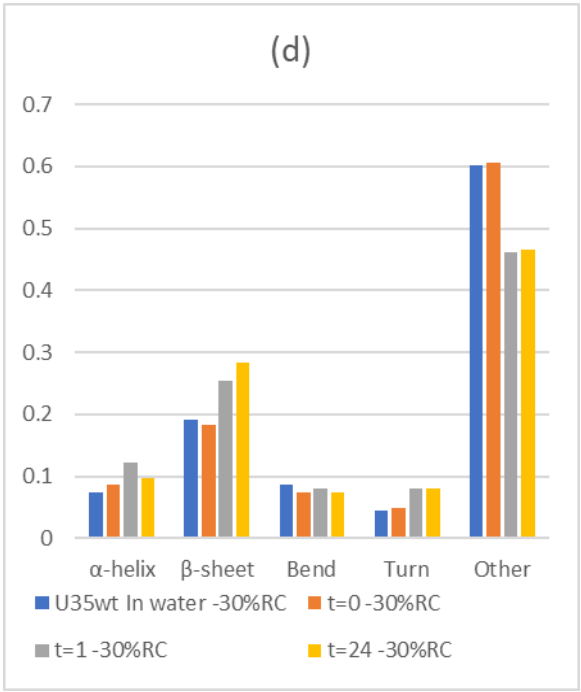
Figure S26 Secondary structure predictions for regenerated U3.4 R7A in water only and after the addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 R7A predictions.

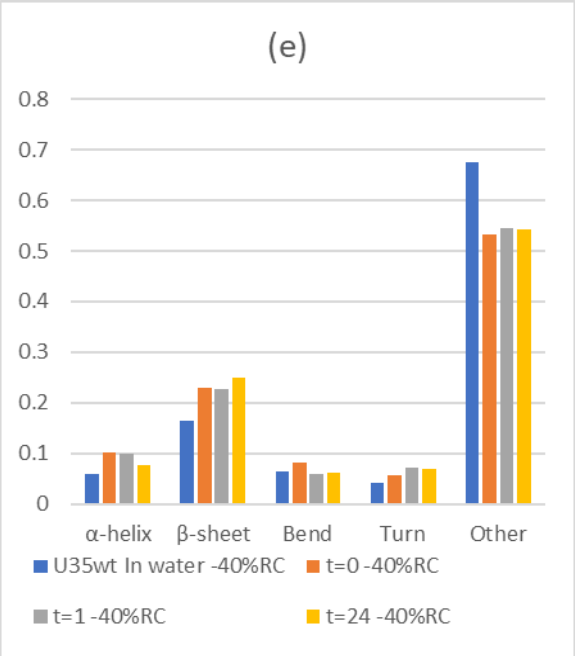
The SOMSpec secondary structure predictions for U3.5 wt data down to 195 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S27. Figure S27a is for the original protein, whereas Figures S27 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

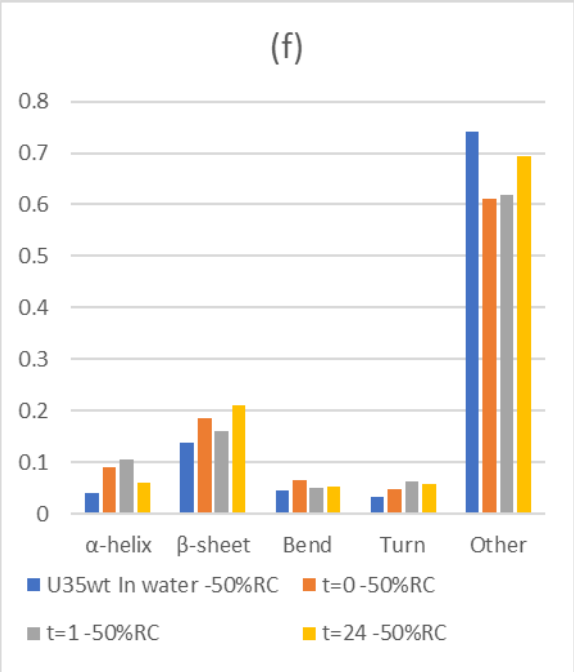


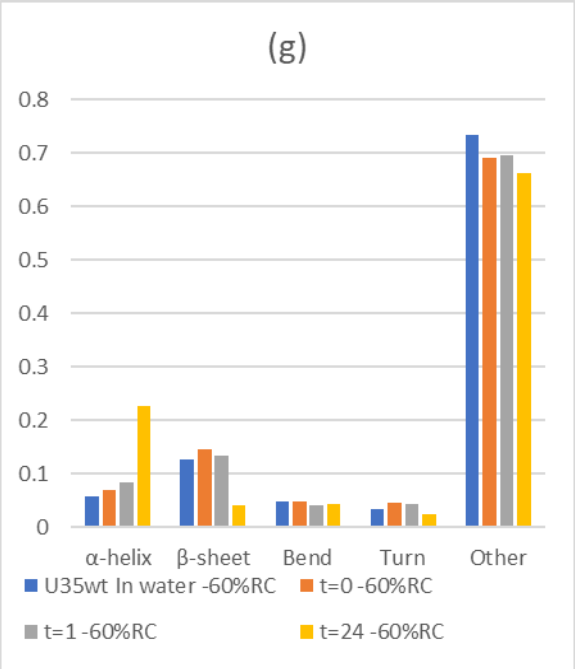


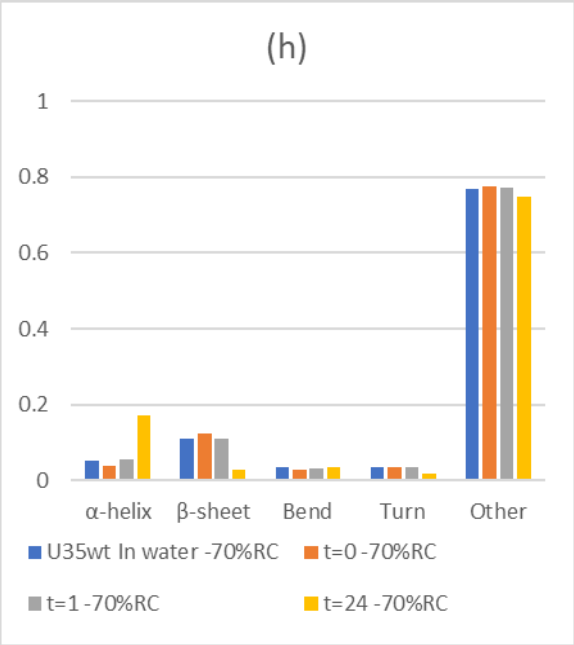





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U35wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

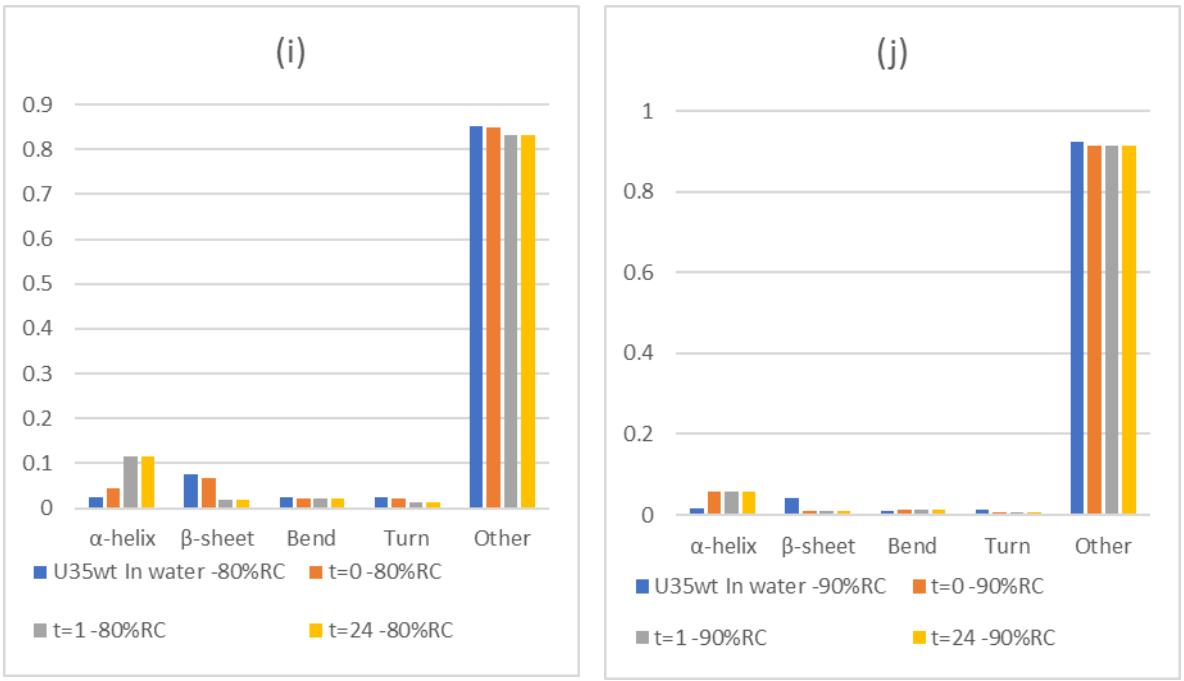
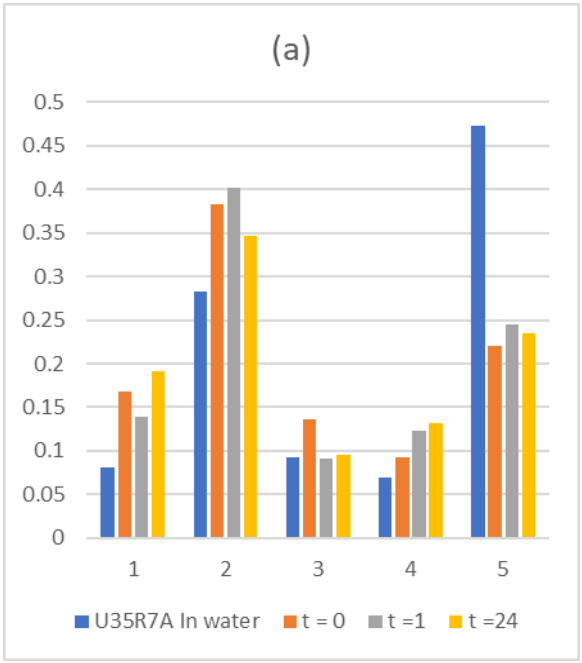
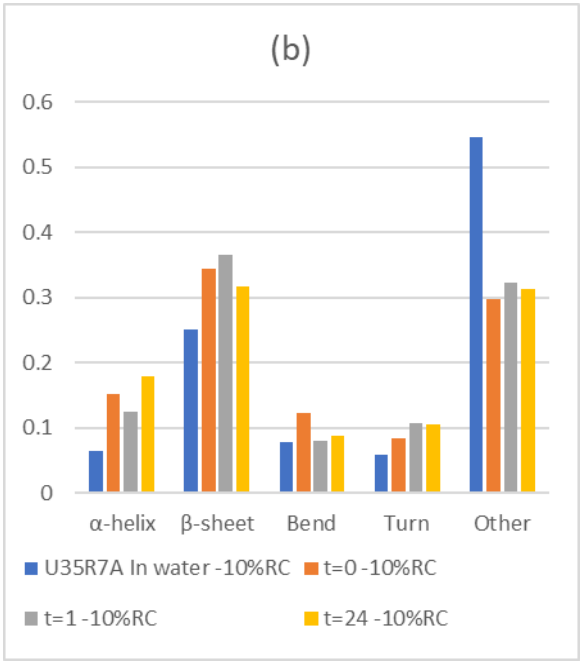
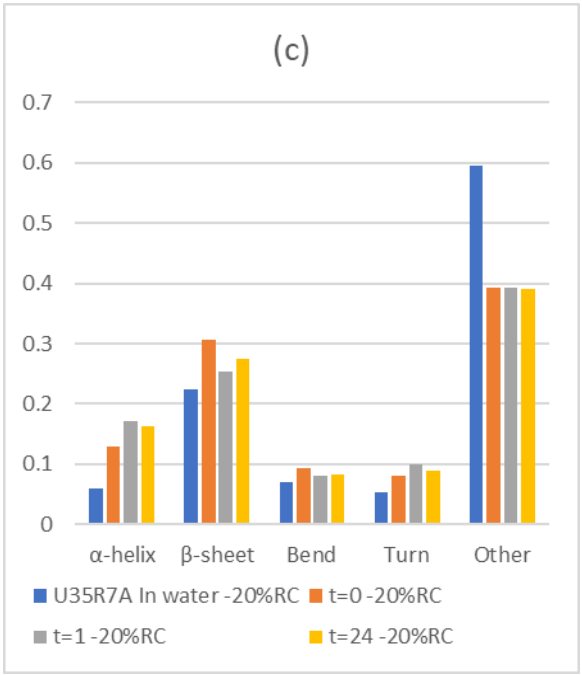


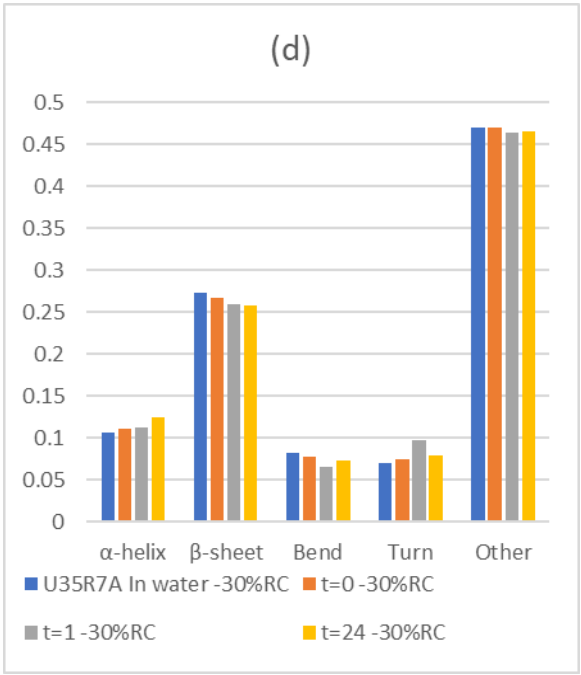
Figure S27 Secondary structure predictions for regenerated U3.5 wt in water only and after the addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 wt predictions.

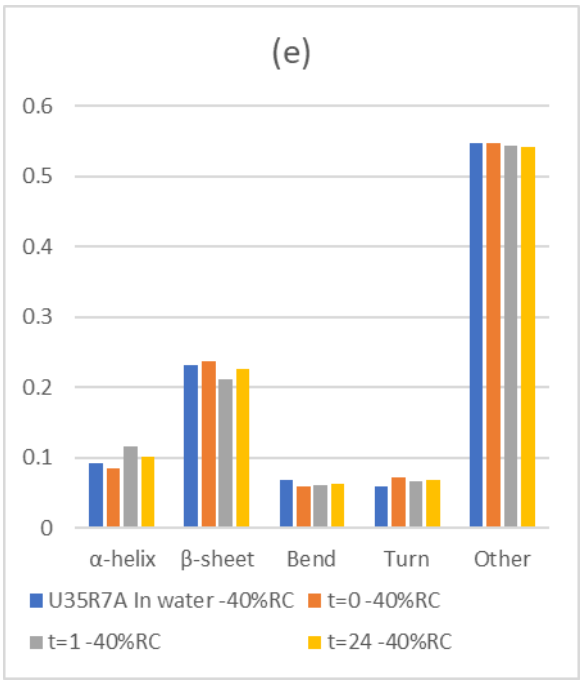
The SOMSpec secondary structure predictions for U3.5 R7A data down to 195 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S28. Figure S28a is for the original protein, whereas Figures S28 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

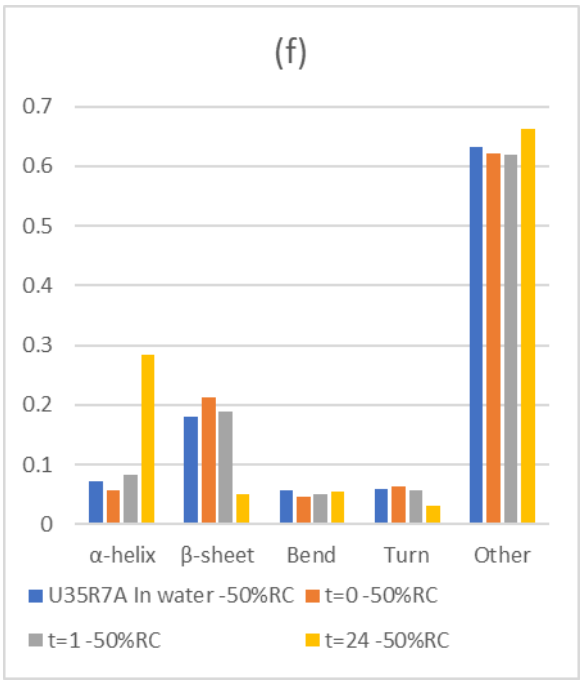


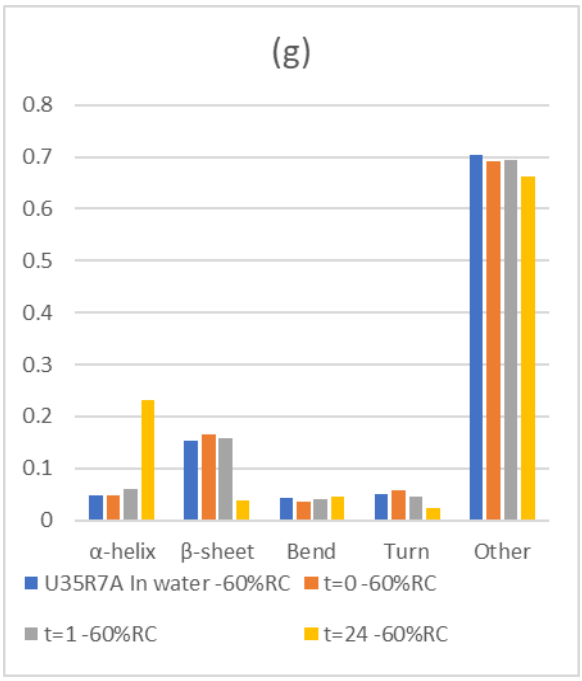


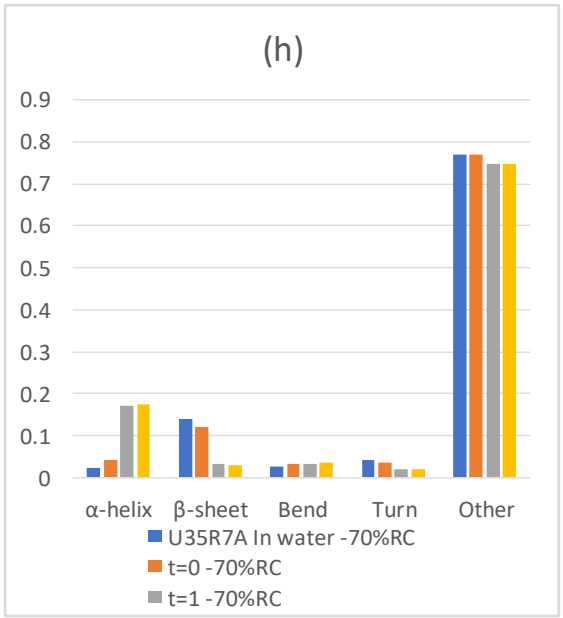





(e)
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC

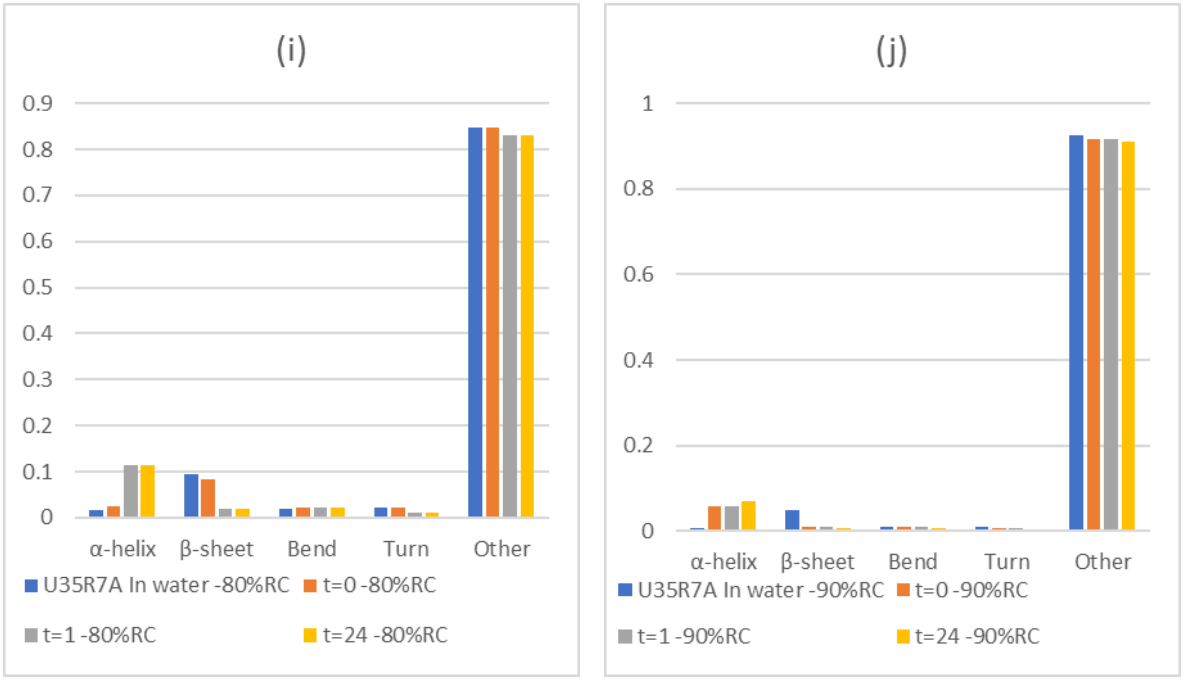
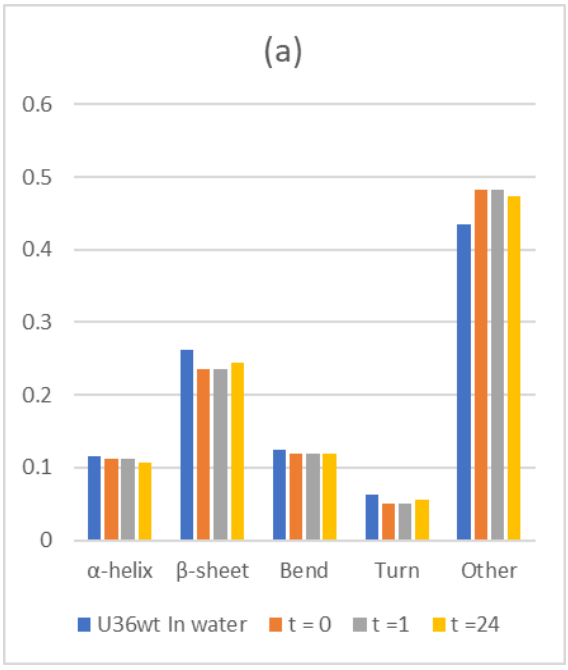
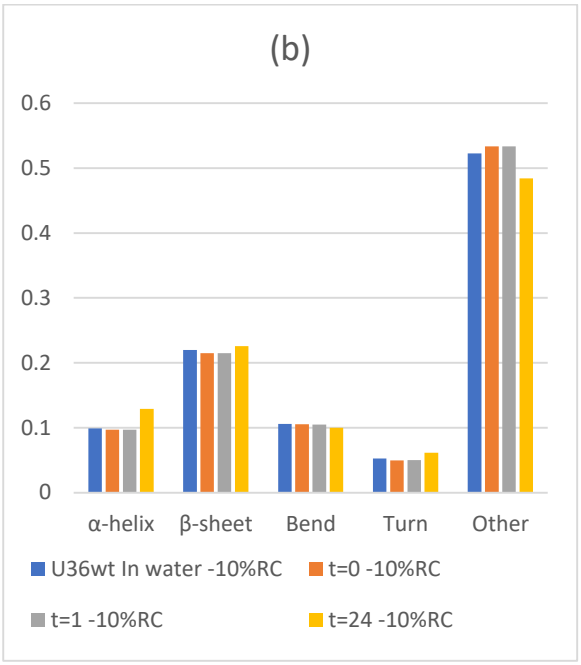
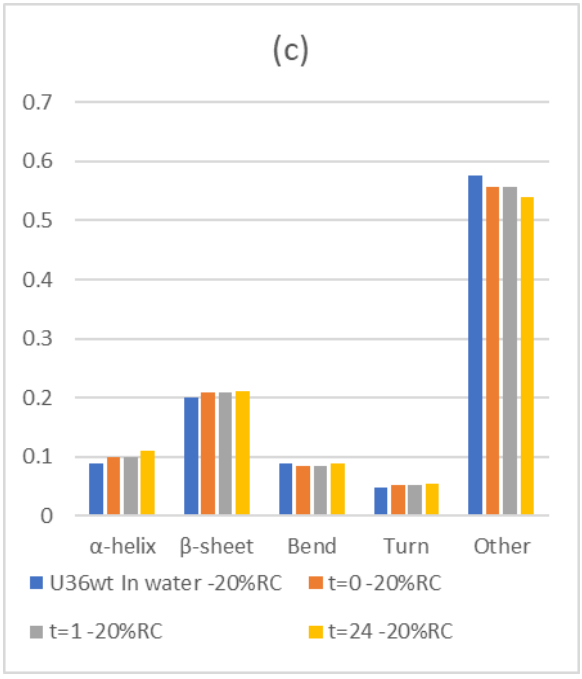


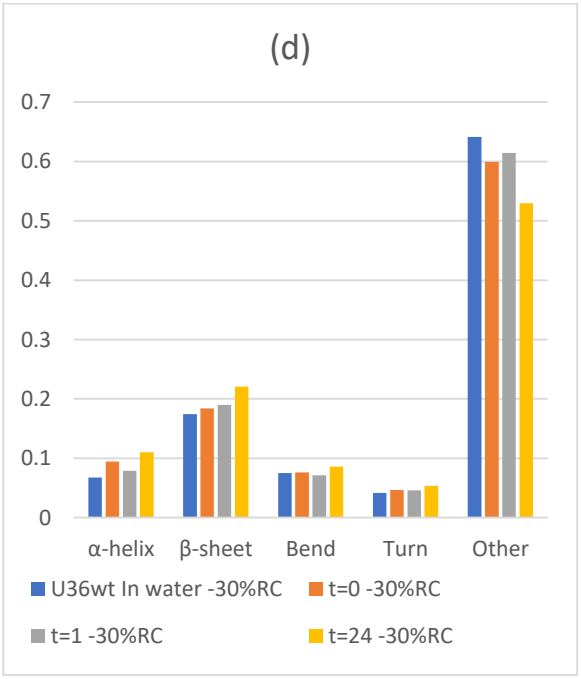
Figure S28 Secondary structure predictions for regenerated U3.5 R7A in water only and after addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 R7A predictions.

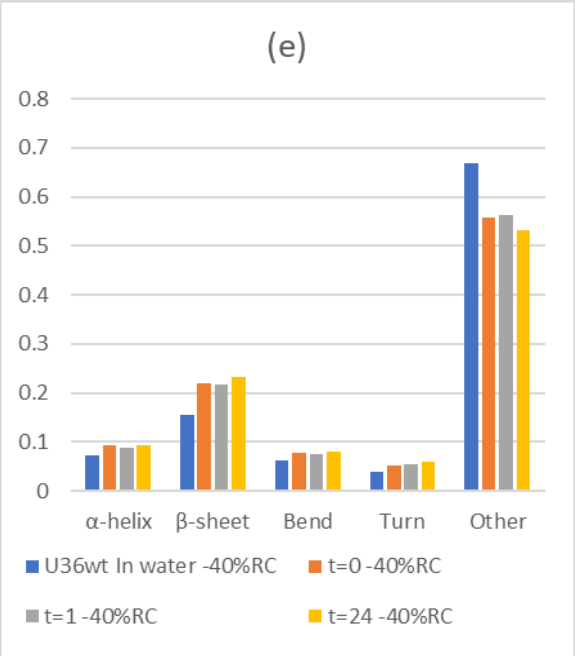
The SOMSpec secondary structure predictions for U3.6 wt data down to 195 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S29. Figure S29a is for the original protein, whereas Figures S29 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

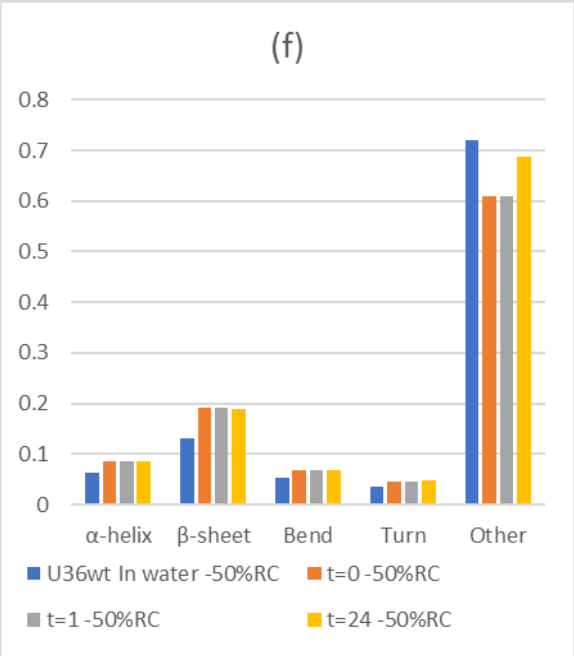


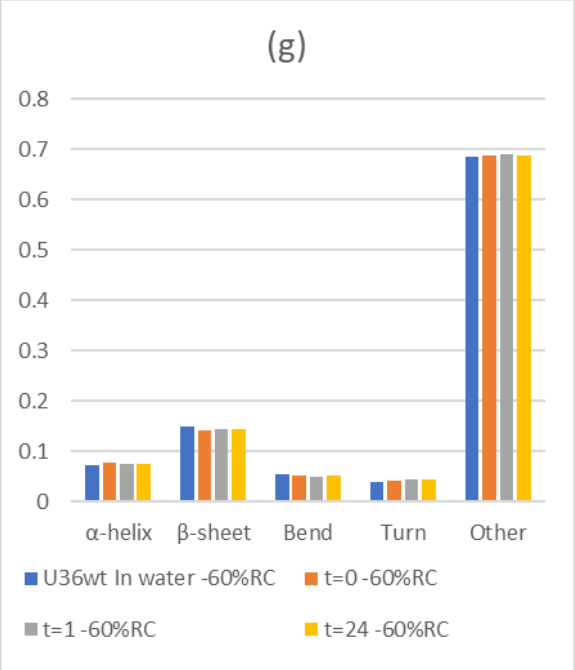


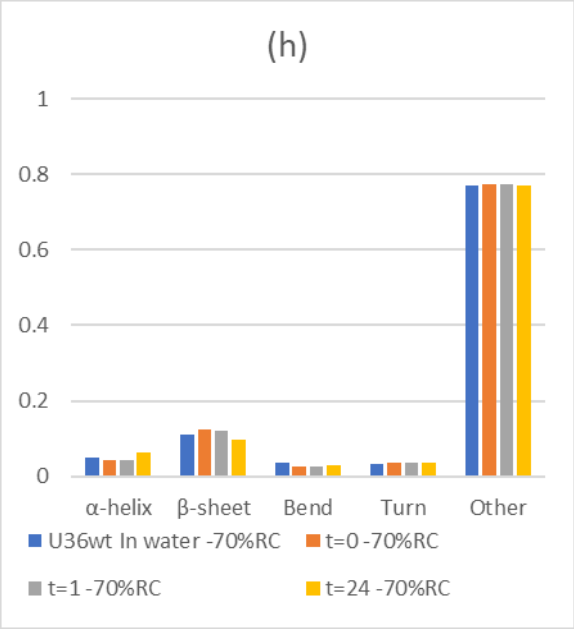





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

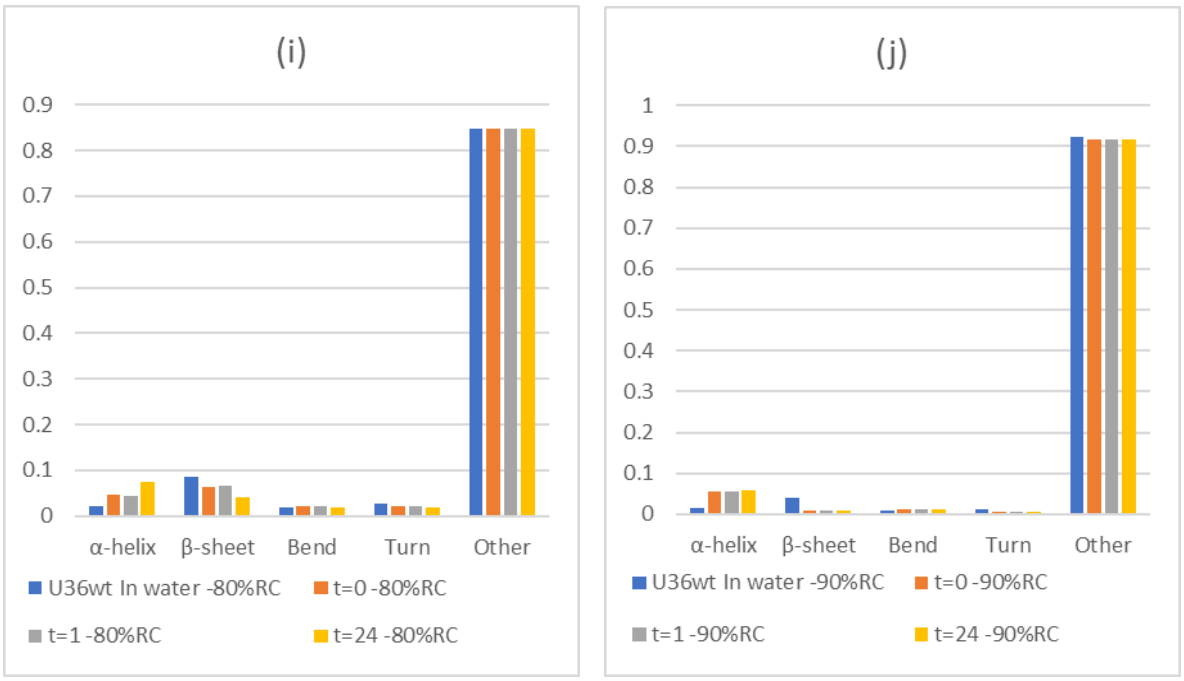
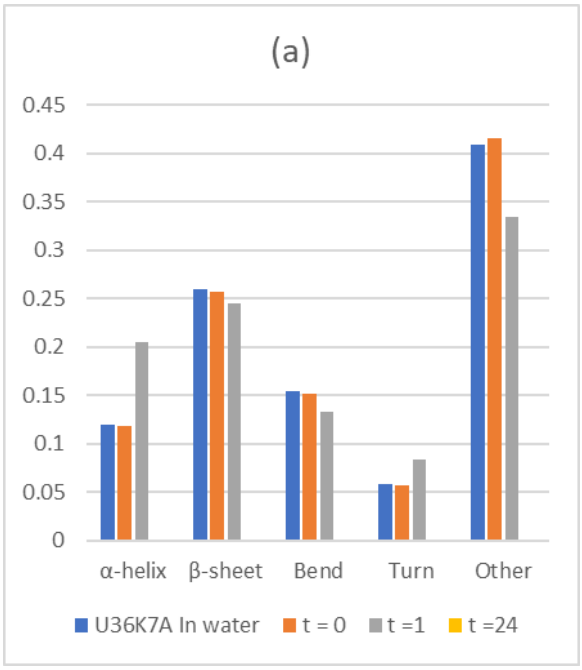
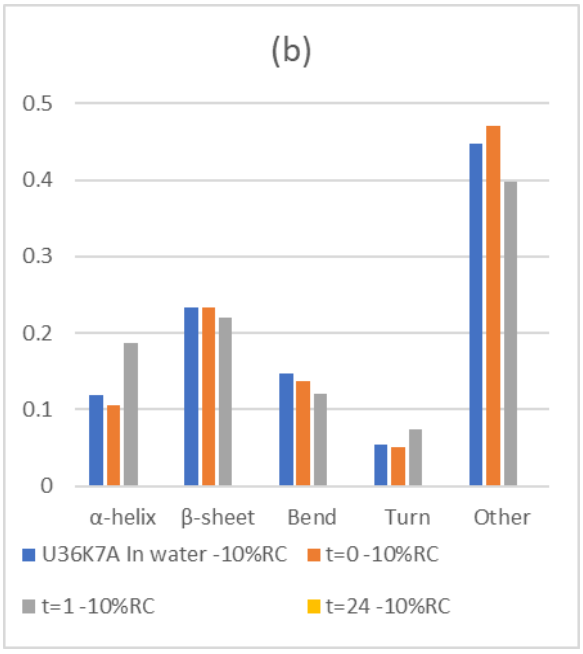
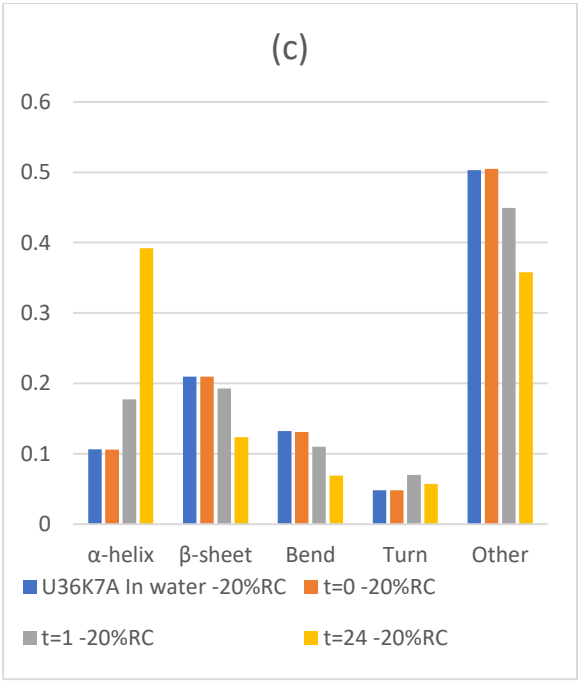


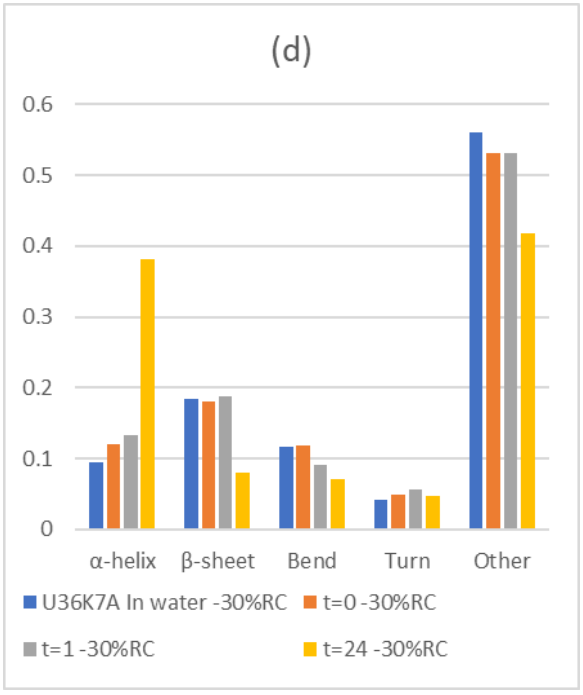
Figure S29 Secondary structure predictions for regenerated U3.6 wt in water only and after addition of buffer over time, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

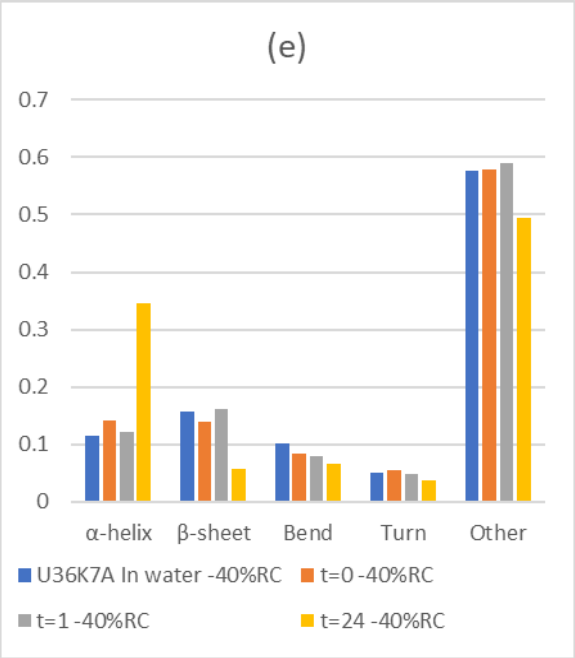
The SOMSpec secondary structure predictions for U3.6 K7A data down to 195 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S30. Figure S30a is for the original protein, whereas Figures S30 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

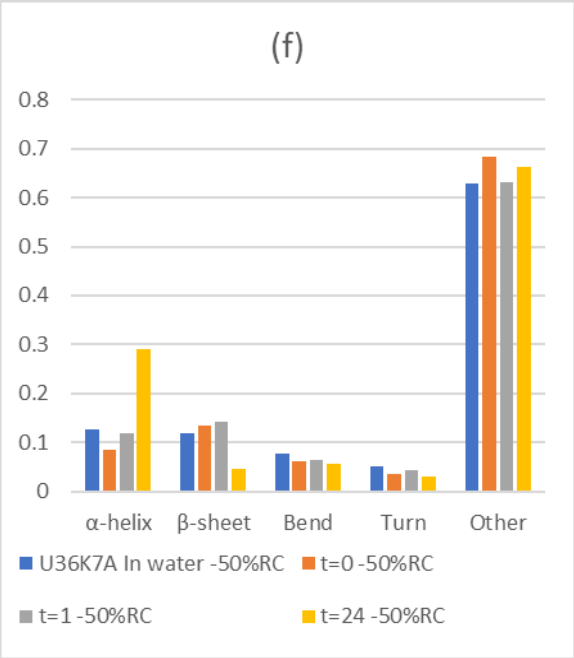


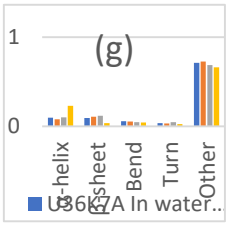


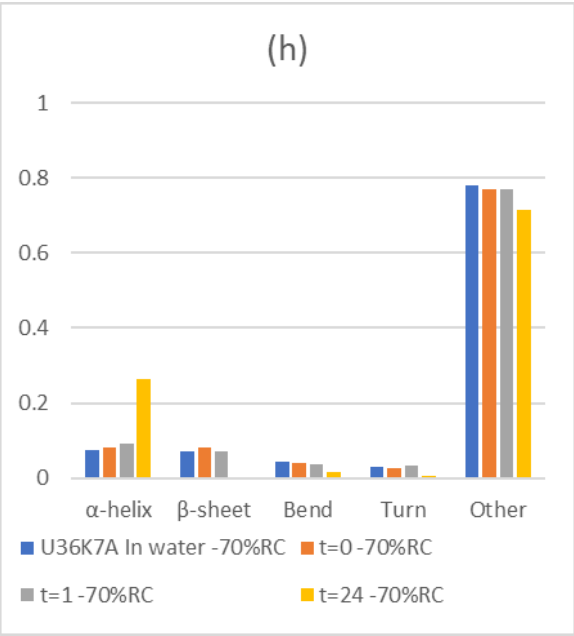





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water..


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

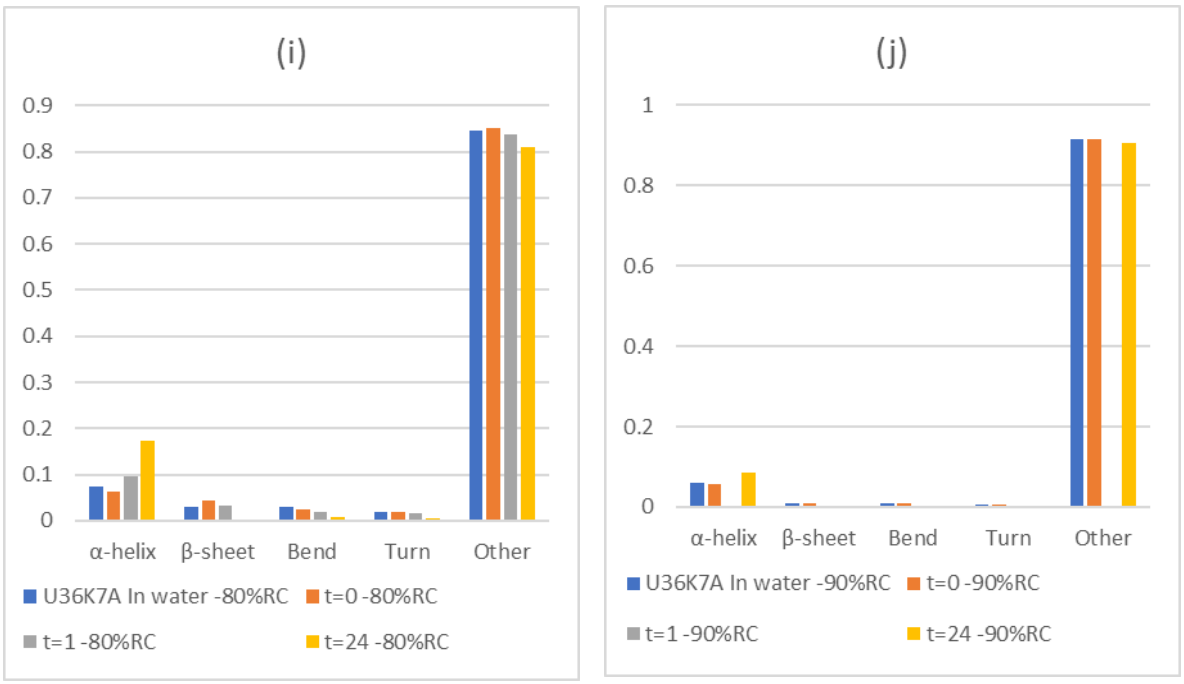
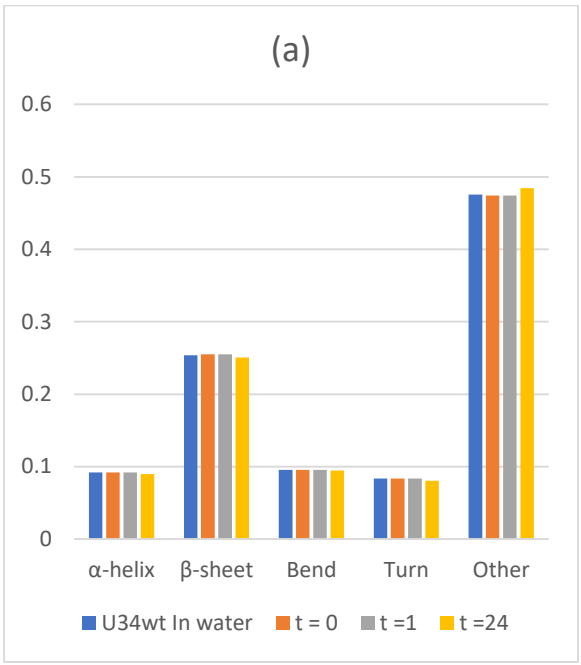
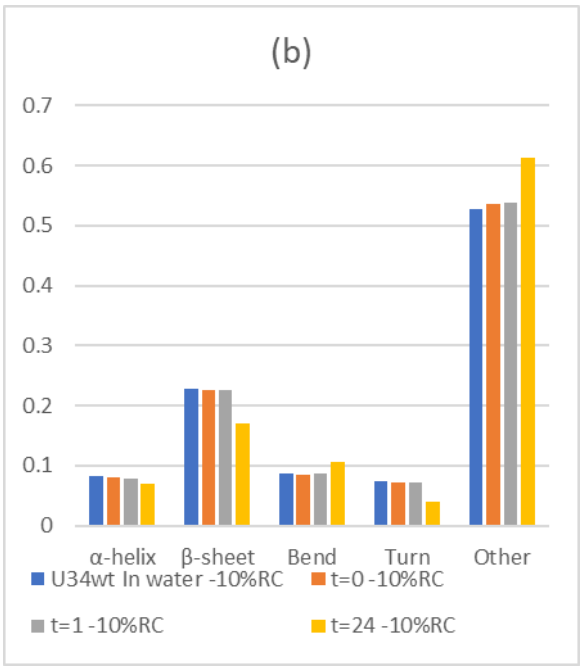
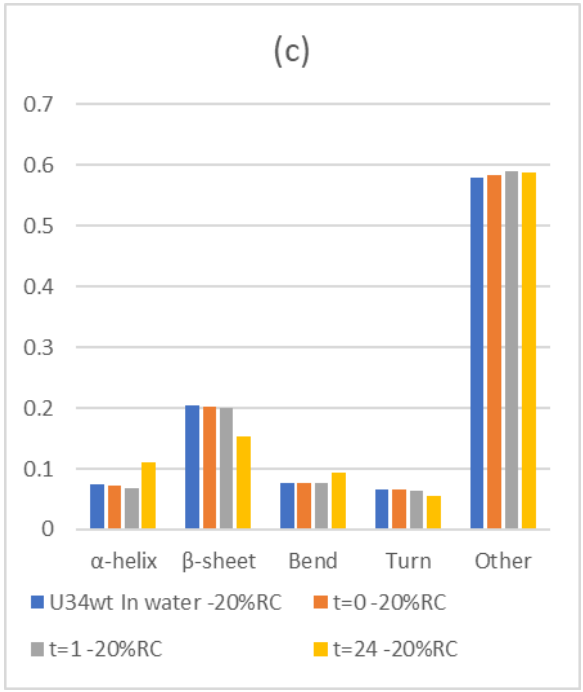


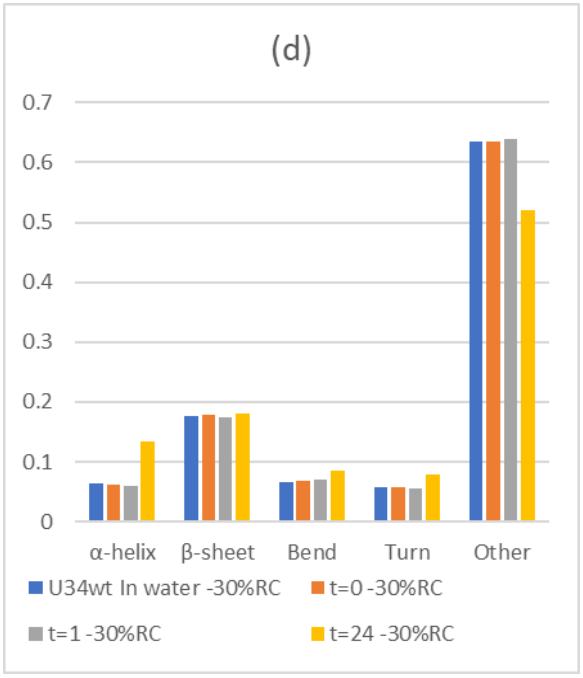
Figure S30 Secondary structure predictions for regenerated U3.6 K7A in water only and after the addition of buffer, created by removing 5 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 K7A predictions.

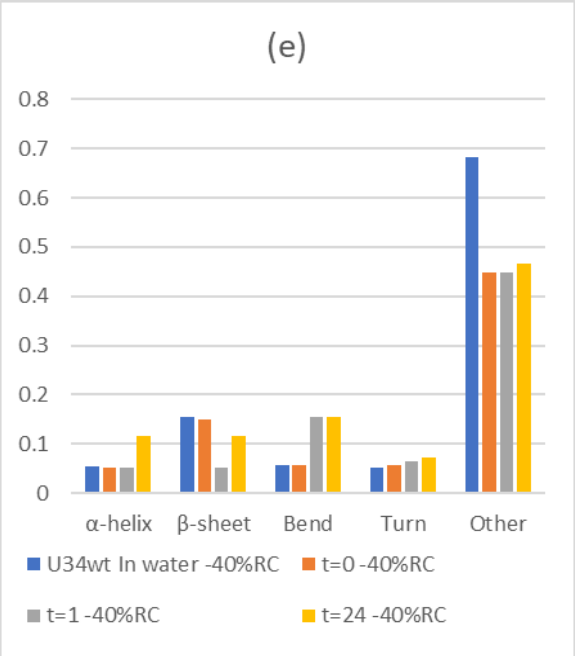
The SOMSpec secondary structure predictions for U3.4 wt data down to 200 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S31. Figure S31a is for the original protein, whereas Figures S31 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

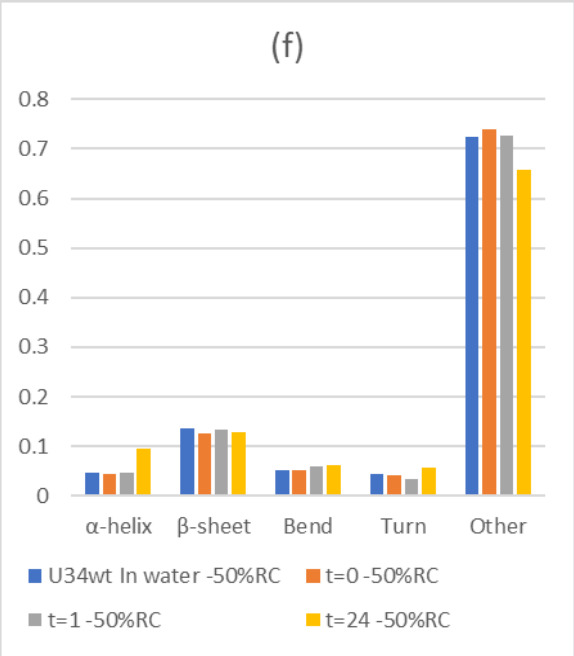


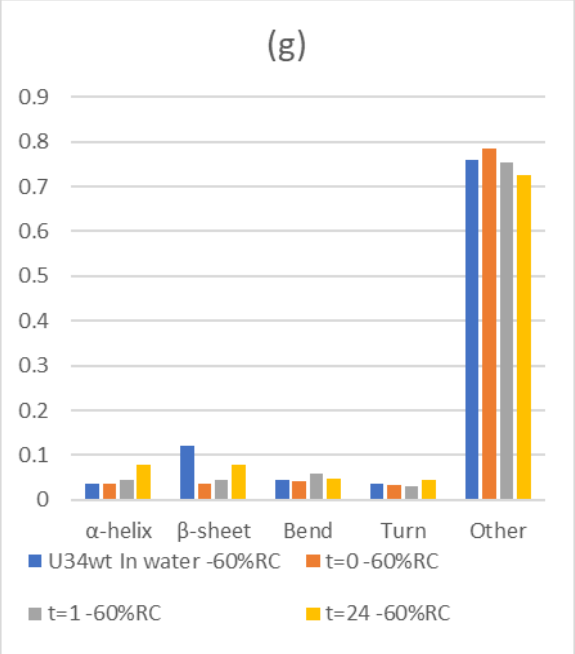


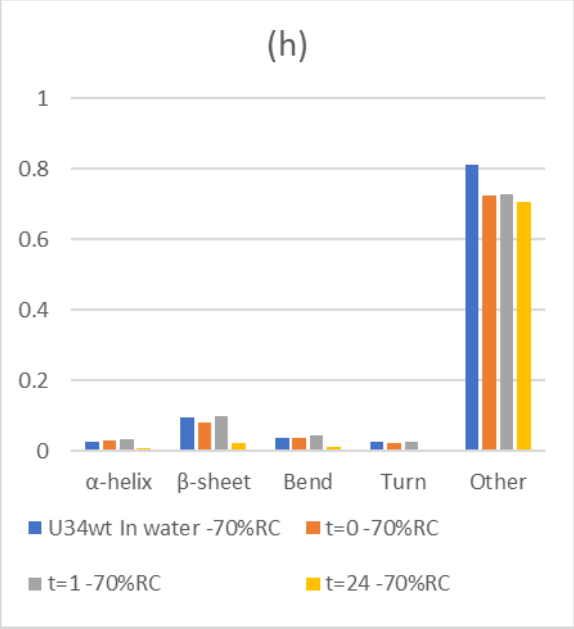





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U34wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

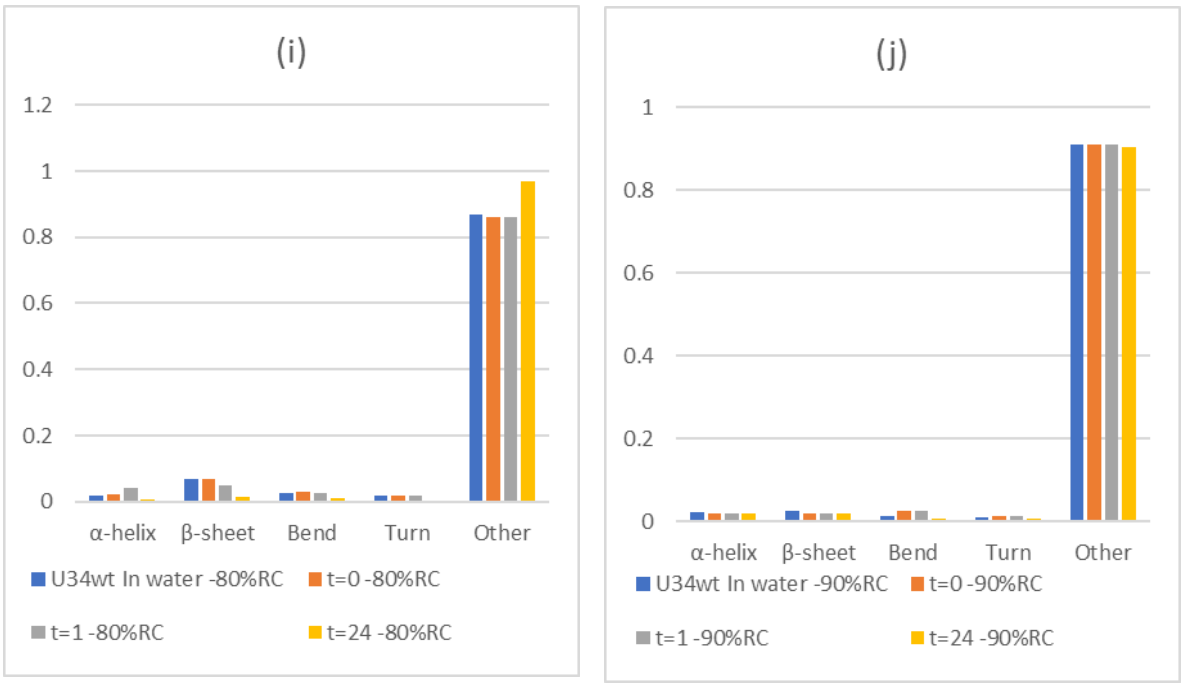
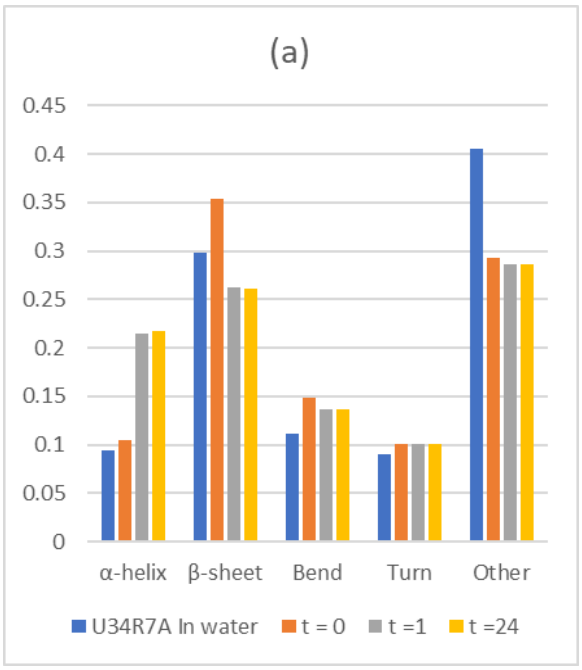
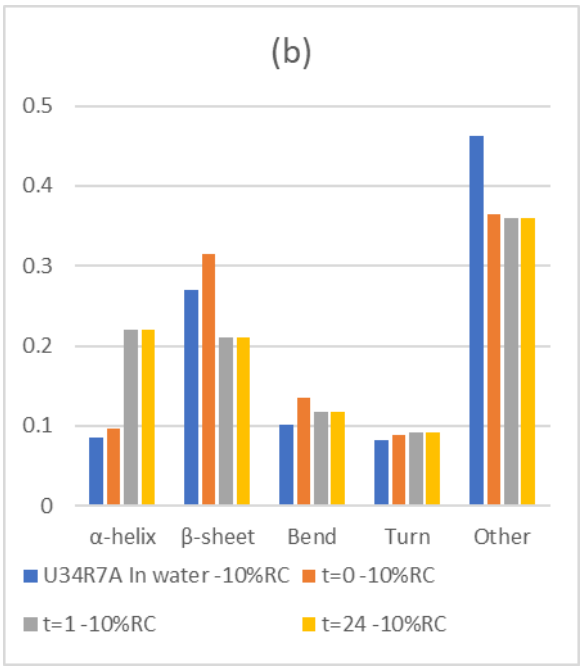
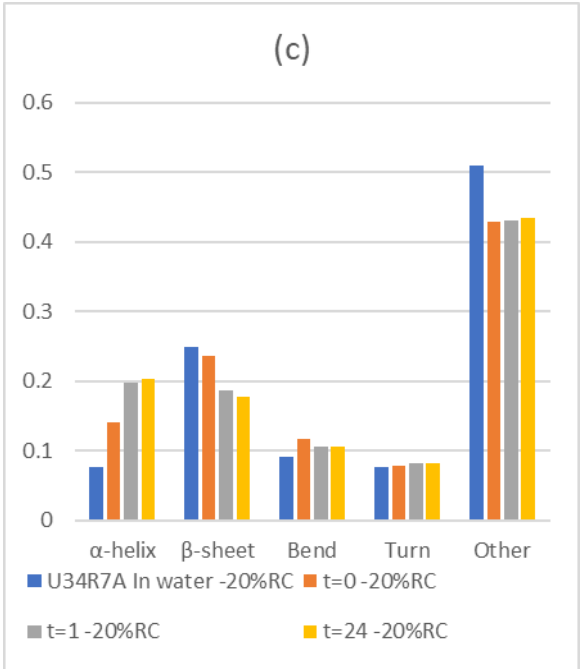


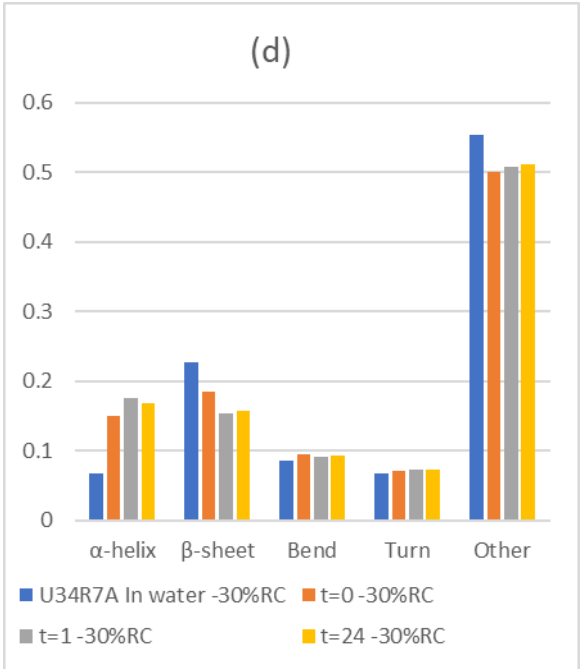
Figure S31 Secondary structure predictions for regenerated U3.4 wt in water only and after addition of buffer, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 wt predictions.

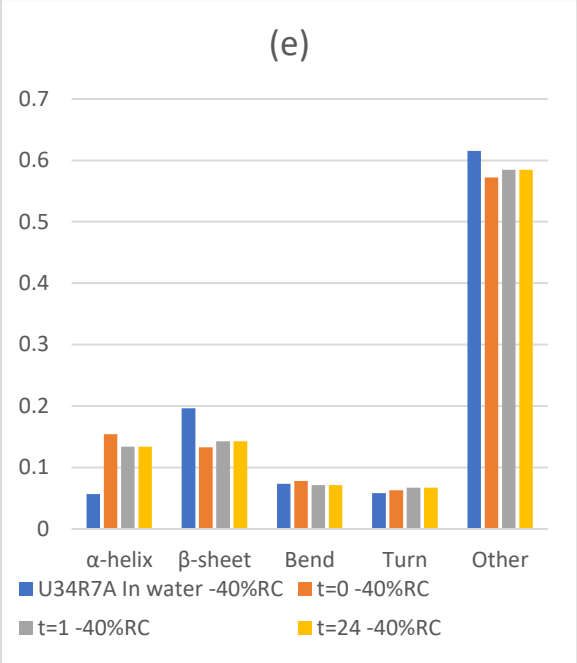
The SOMSpec secondary structure predictions for U3.4 R7A data down to 200 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S32. Figure S32a is for the original protein, whereas Figures S32 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

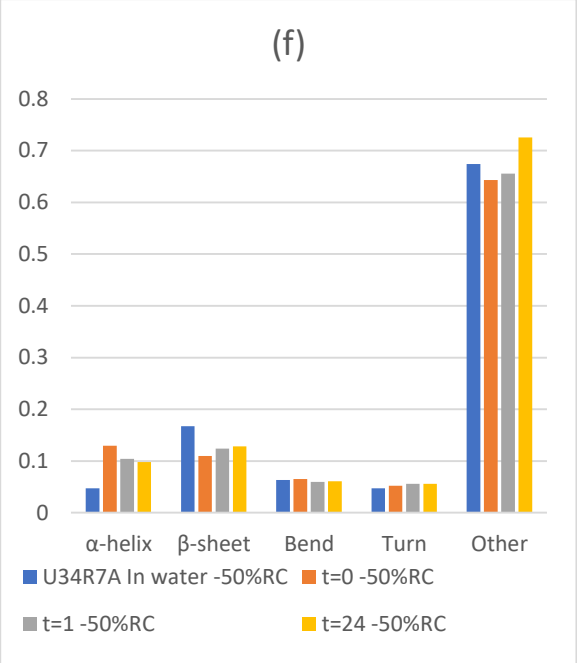


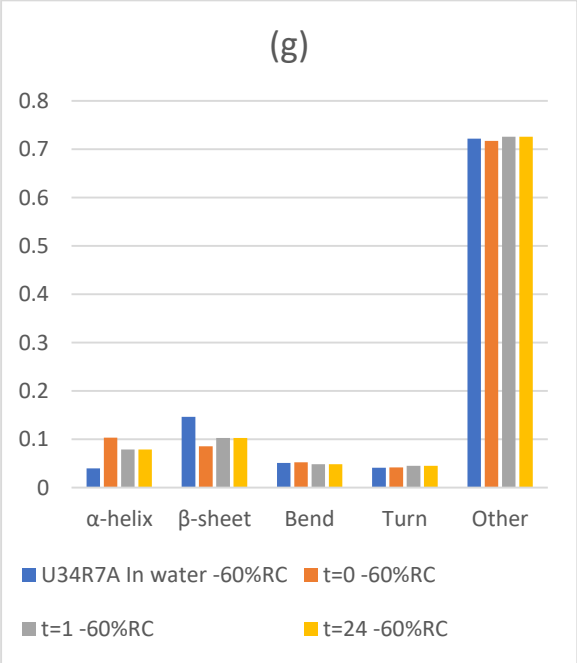


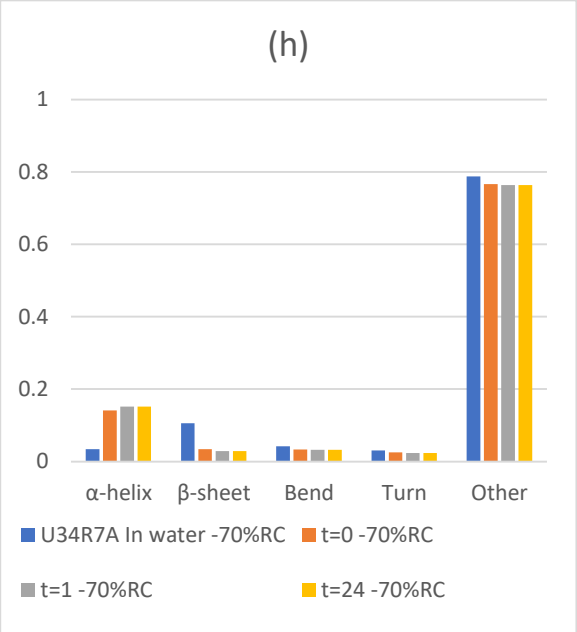





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U34R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

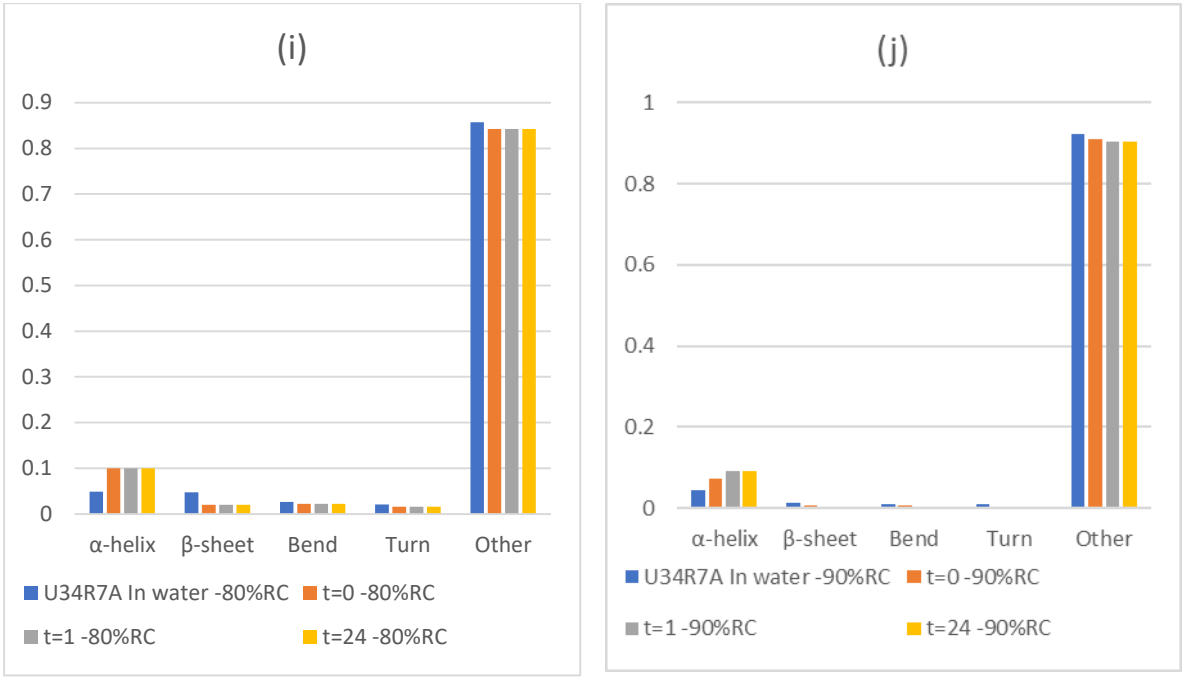
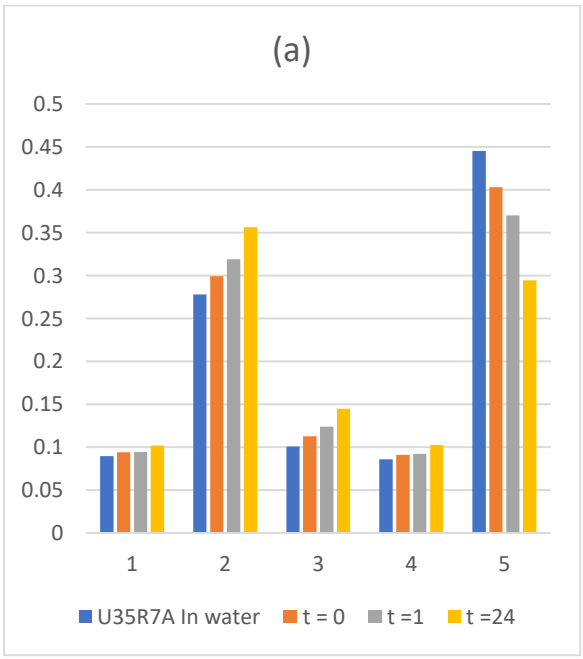
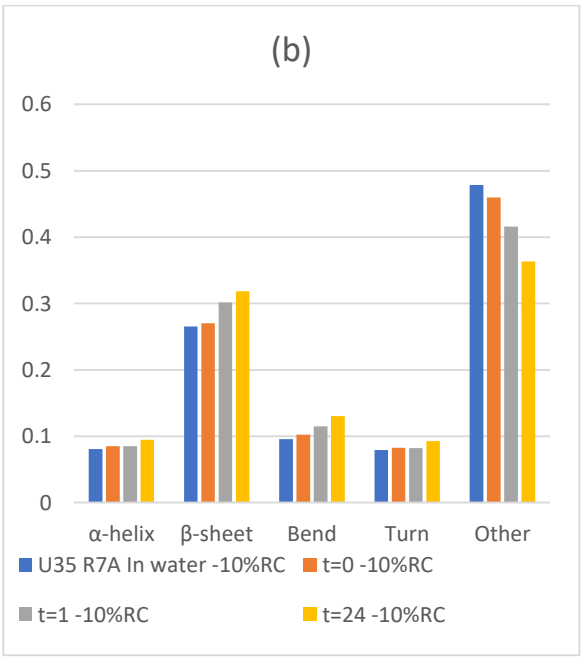
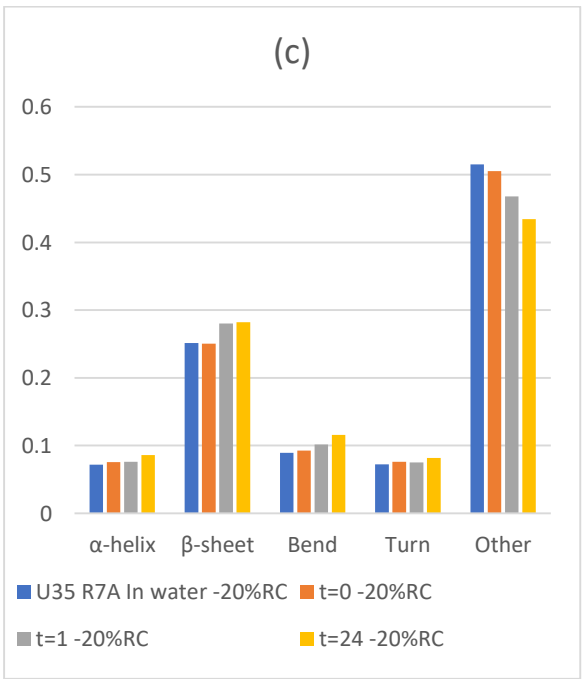


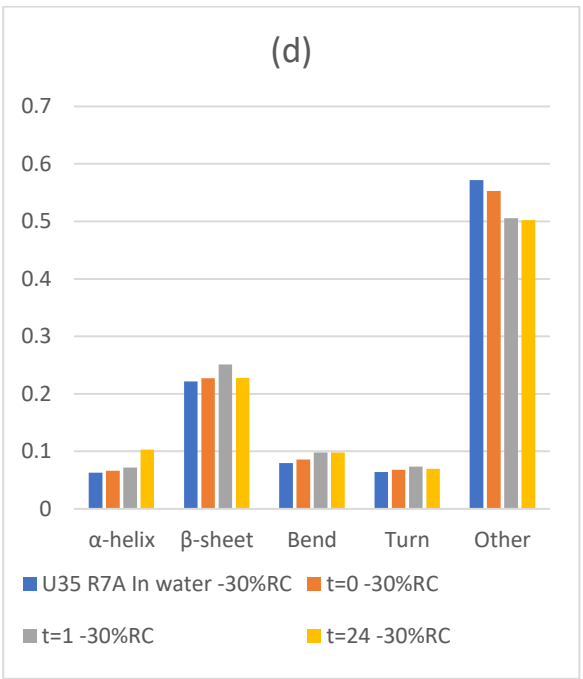
Figure S32 Secondary structure predictions for regenerated U3.4 R7A in water only and after addition of buffer over time, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.4 R7A predictions

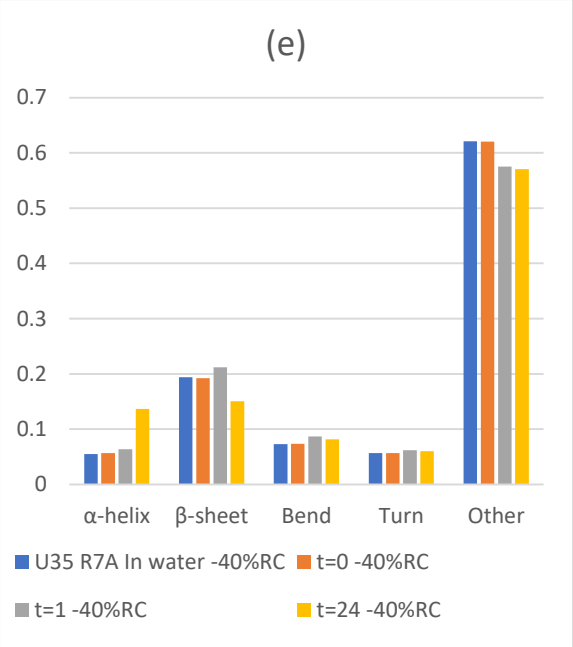
The SOMSpec secondary structure predictions for U3.5 R7A data down to 200 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S33. Figure S33a is for the original protein, whereas Figures S33 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

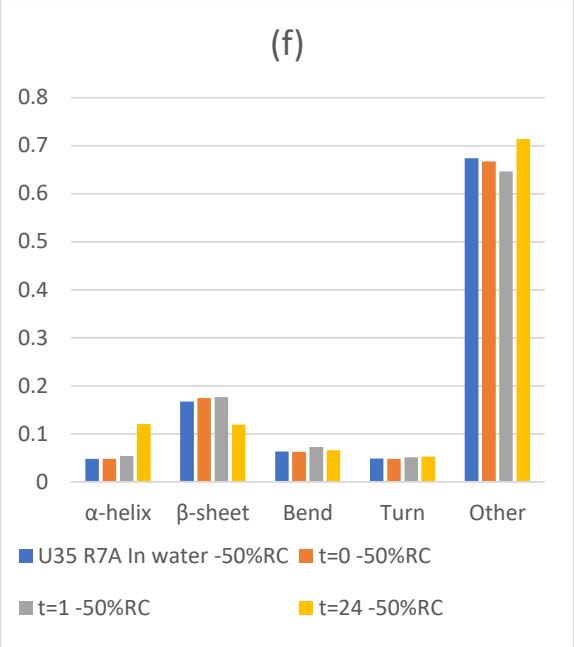


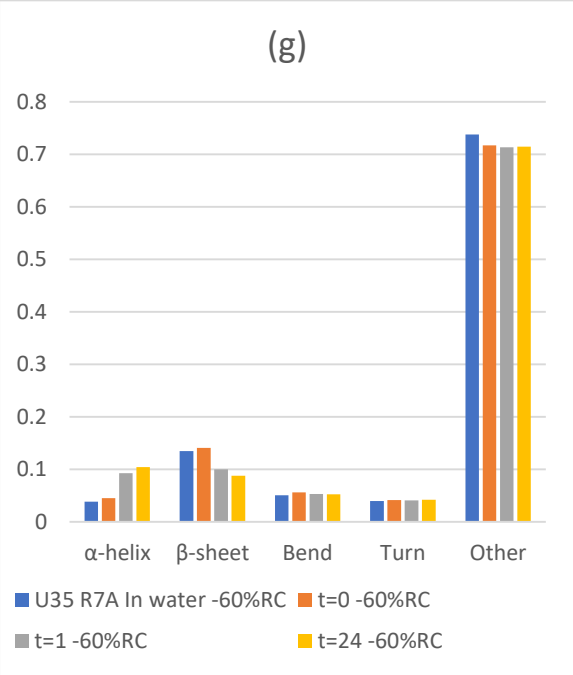


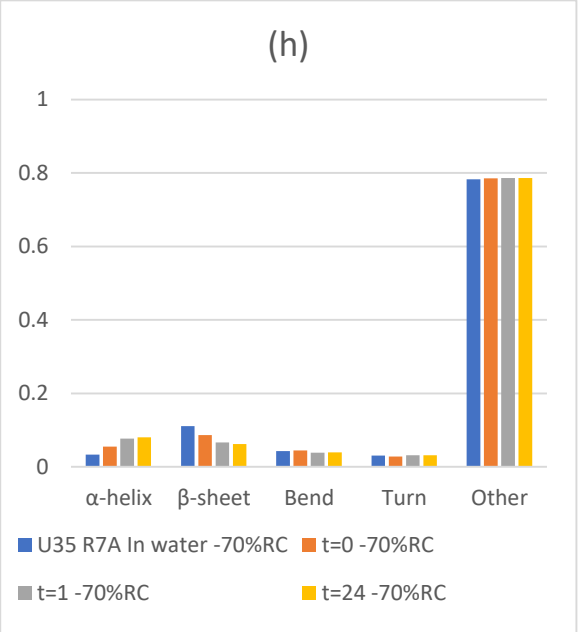





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

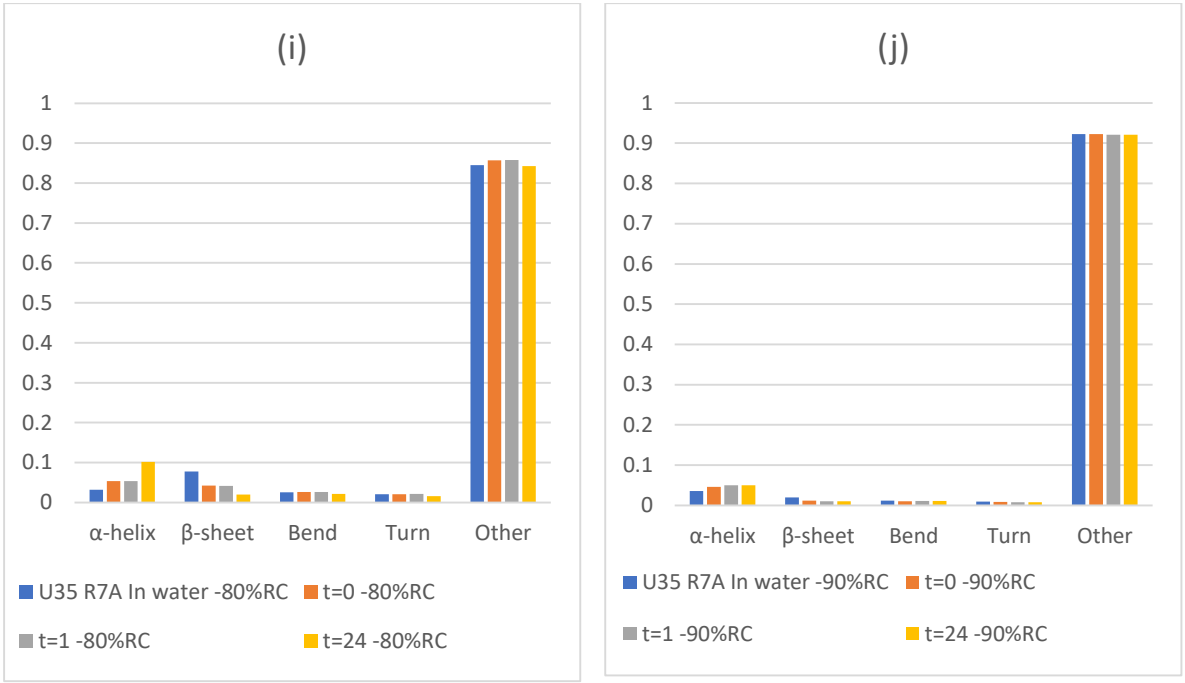
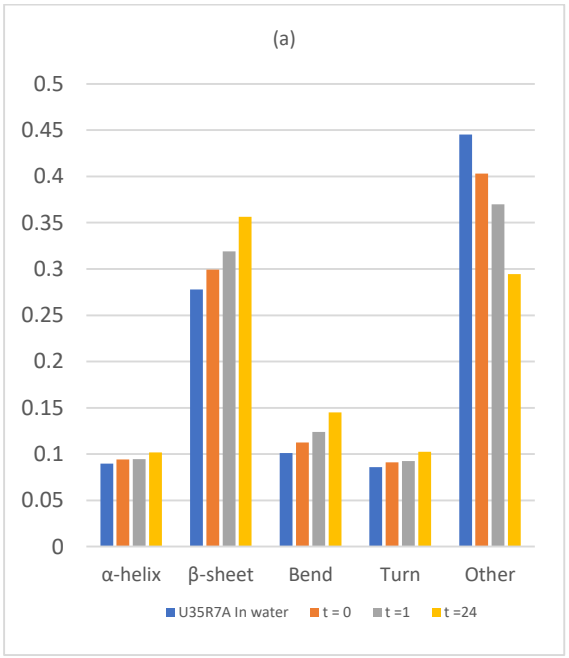
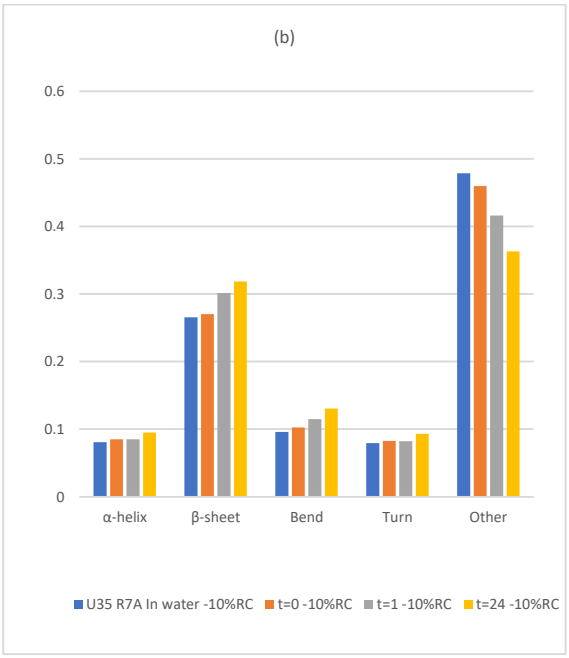
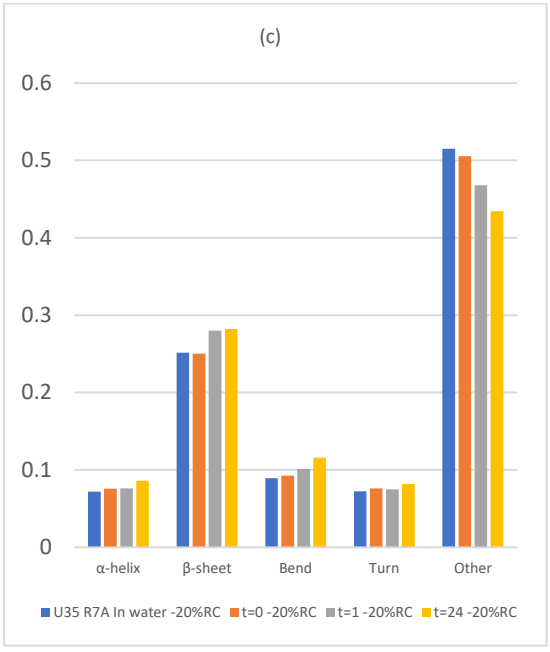


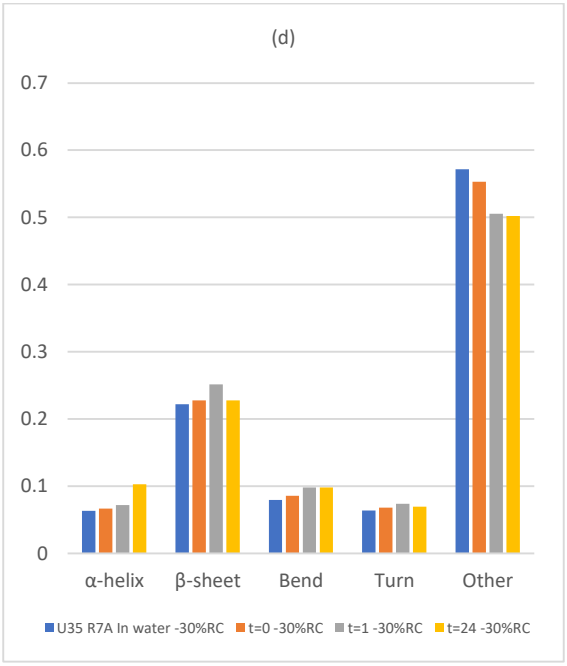
Figure S33 Secondary structure predictions for regenerated U3.5 wt in water only and after the addition of buffer over time, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.5 R7A predictions.

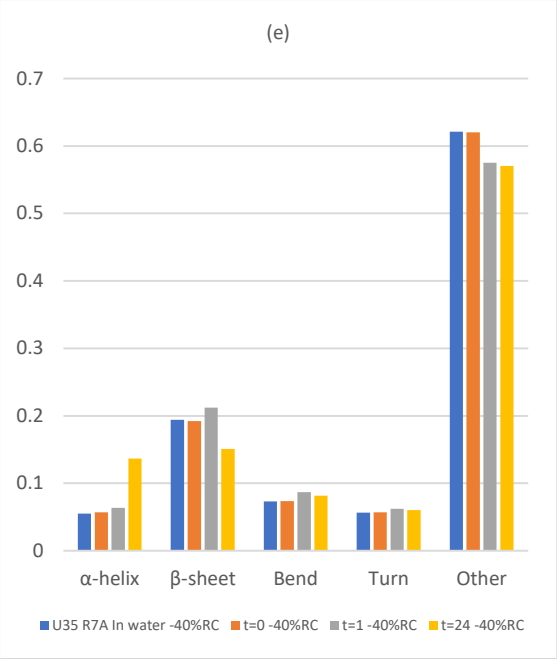
The SOMSpec secondary structure predictions for U3.5 R7A data down to 200 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S34. Figure S34a is for the original protein, whereas Figures S34 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

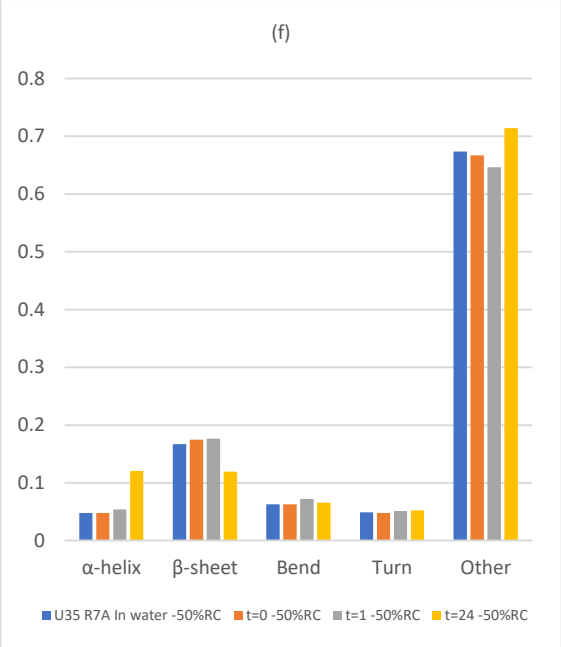


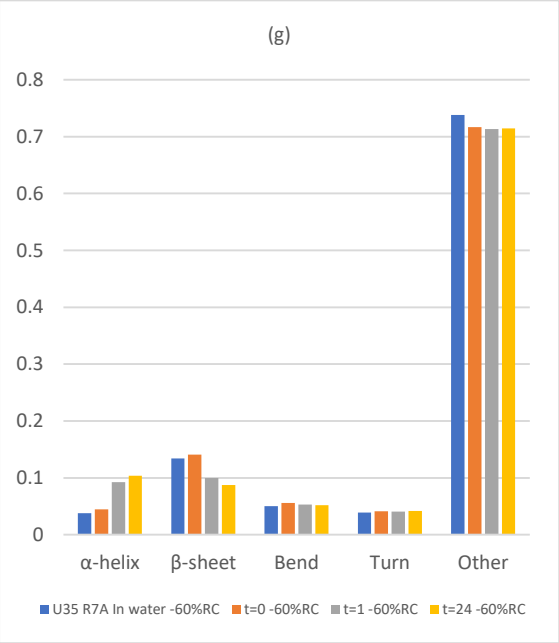


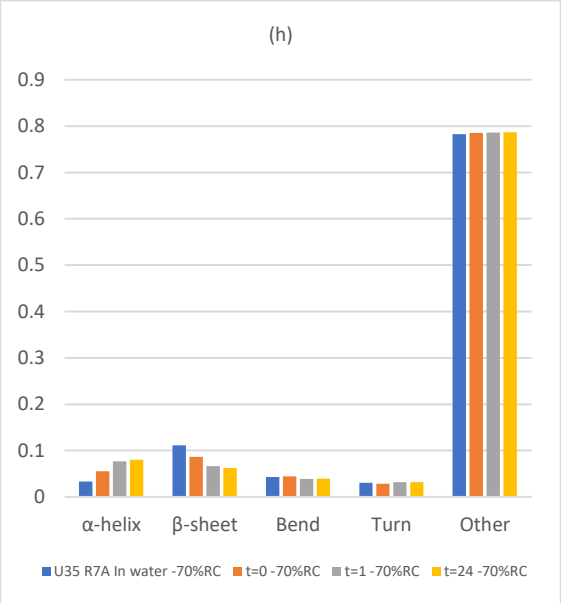





(e)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U35 R7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

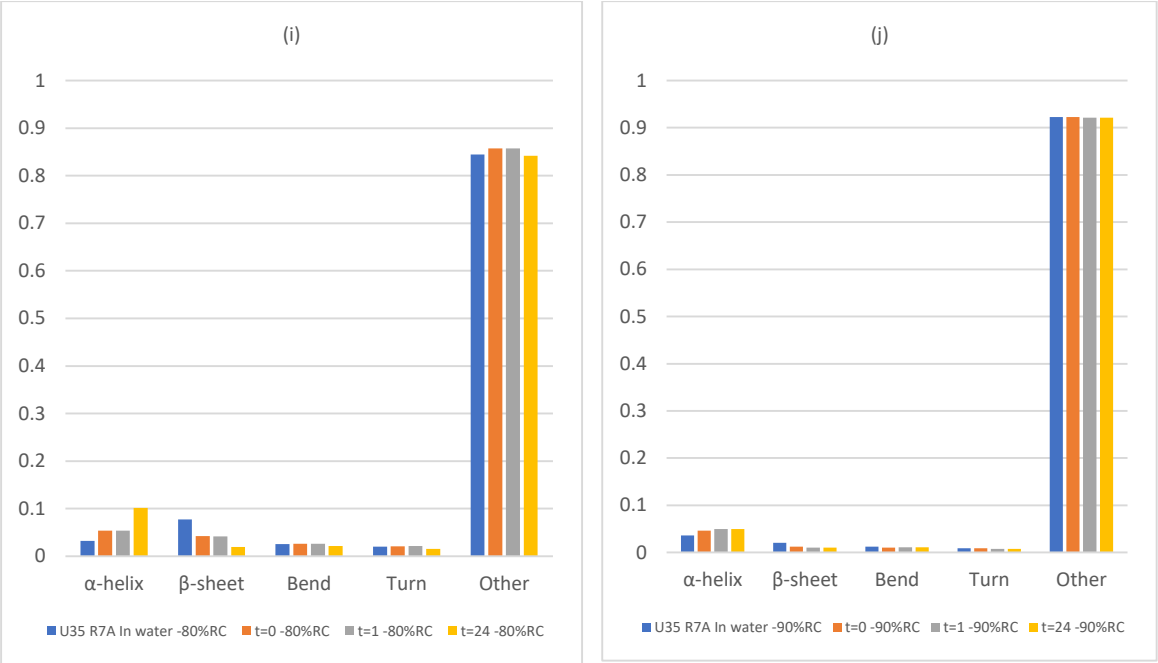
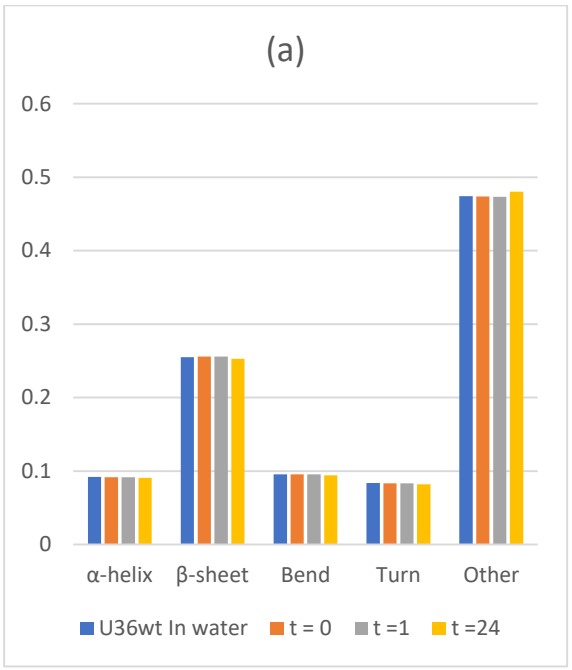
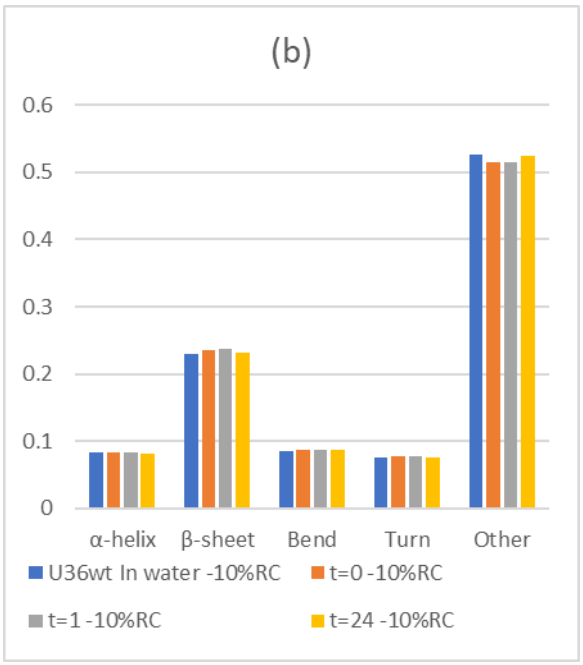
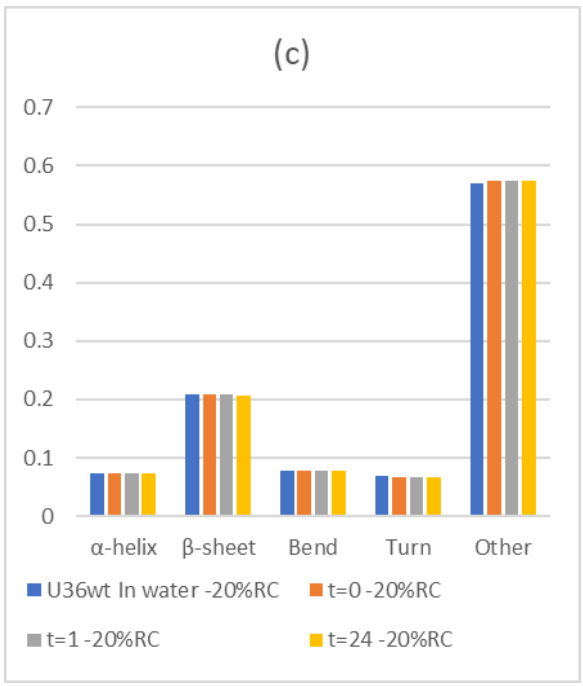


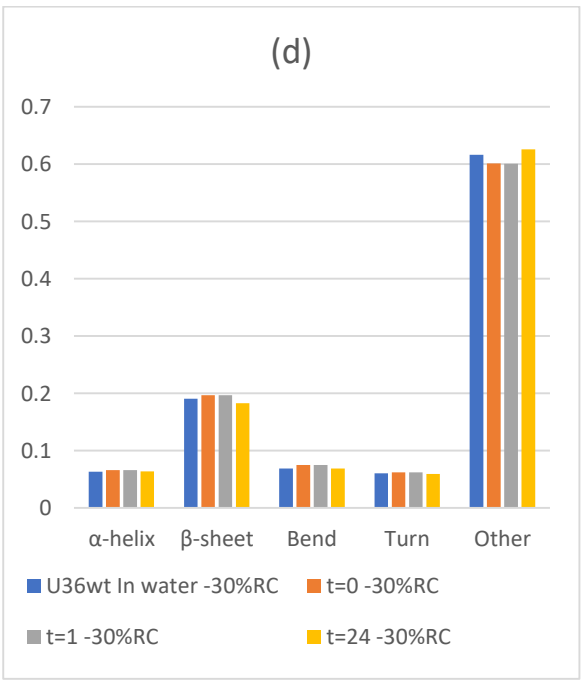
Figure S34 Secondary structure predictions for regenerated U3.5 R7A in water only and after addition of buffer over time, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

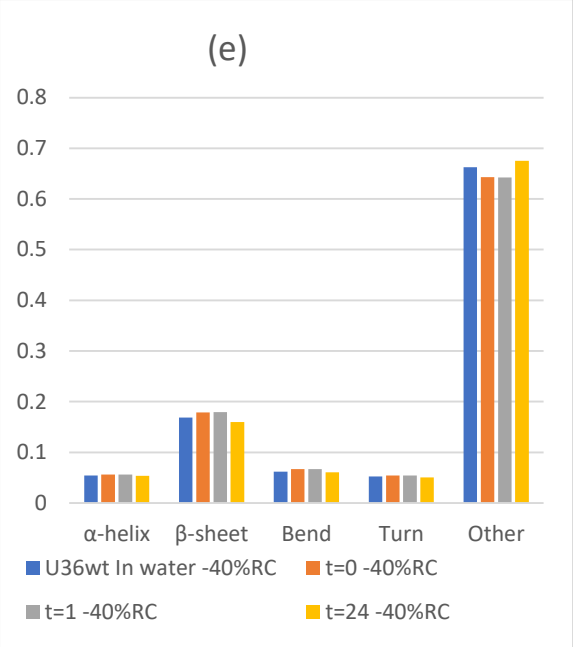
The SOMSpec secondary structure predictions for U3.6 wt data down to 200 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S35. Figure S35a is for the original protein, whereas Figures S35 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

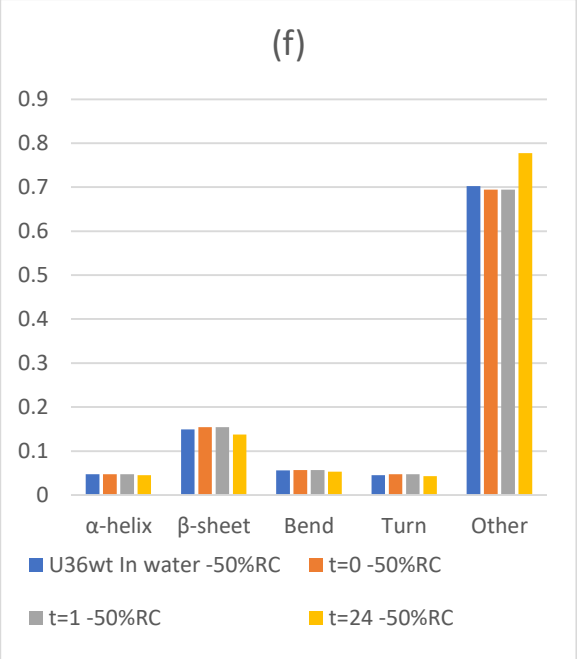


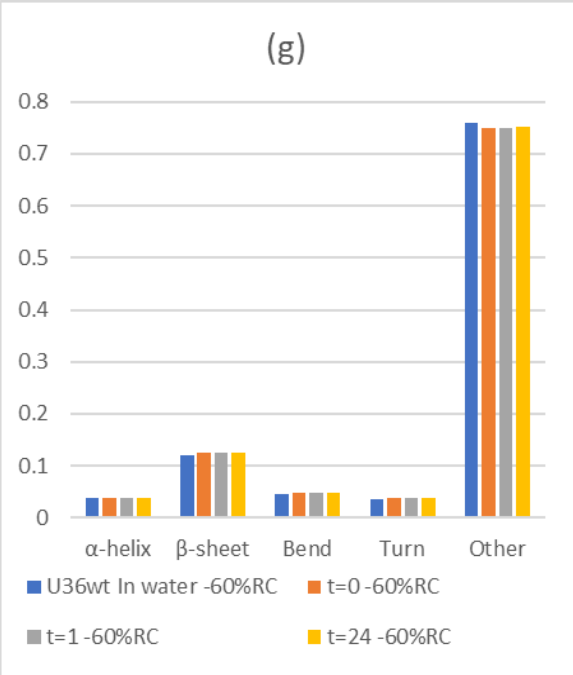


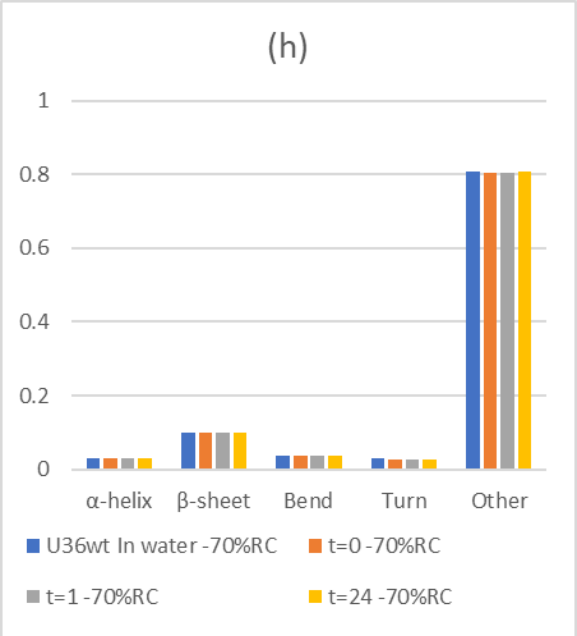





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U36wt In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

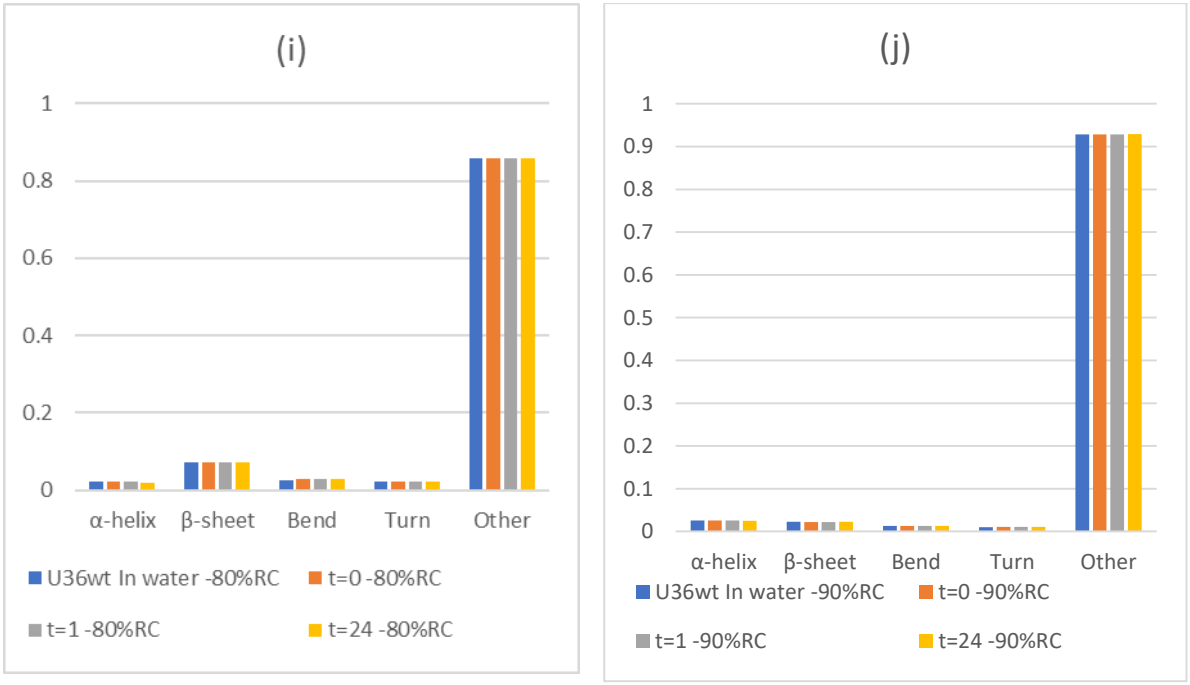
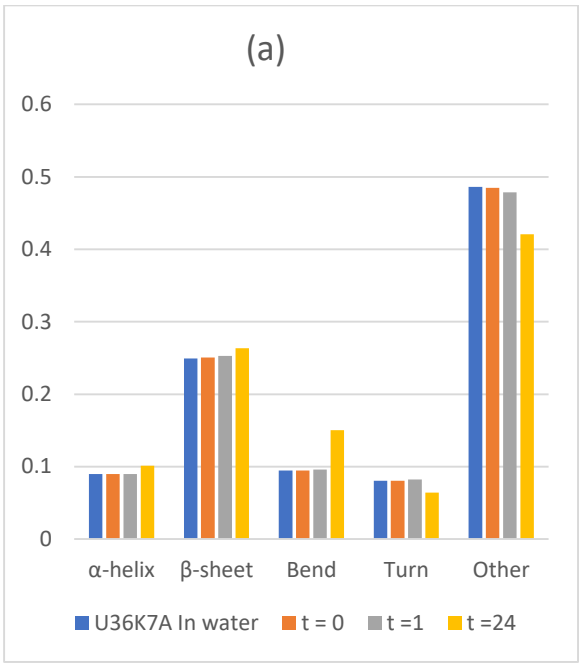
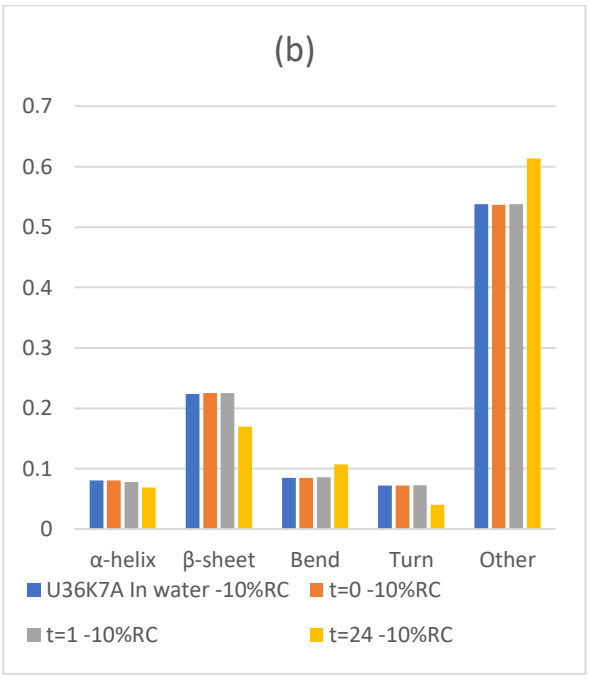
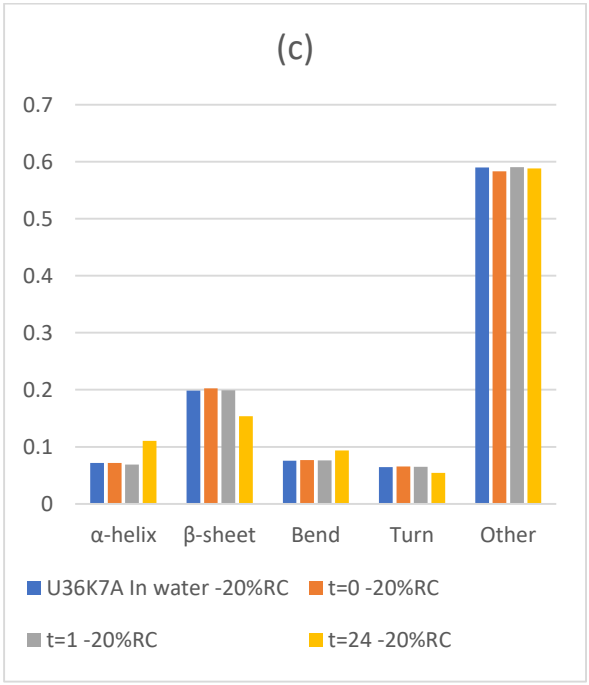


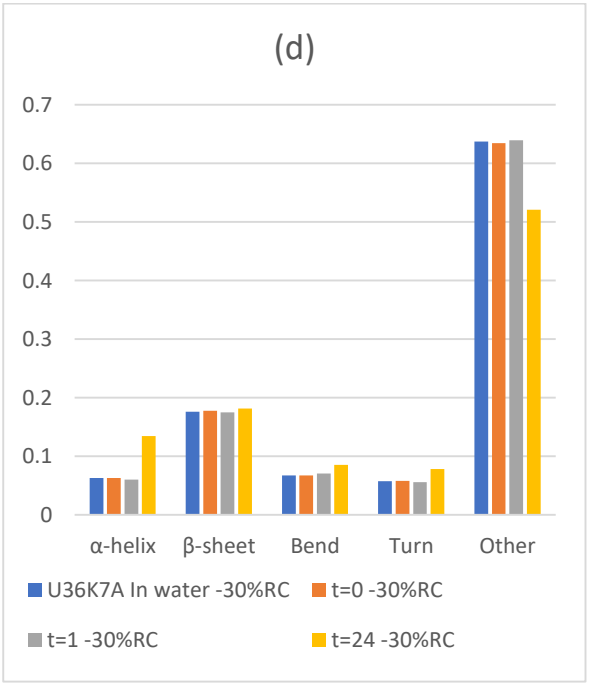
Figure S35 Secondary structure predictions for regenerated U3.6 wt in water only and after the addition of buffer over time, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 wt predictions.

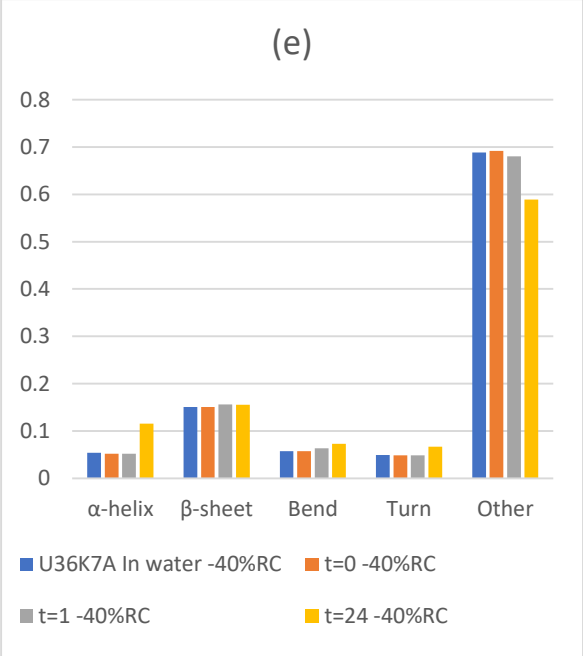
The SOMSpec secondary structure predictions for U3.6 K7A data down to 200 nm in the absence and presence of buffer (readings on the effect of buffer addition at three different time intervals were taken at t = 0, t = 1, and t = 24 h) for each percentage RC content subtracted tested against reference set down to the same wavelength range are shown in Figures S36. Figure S36a is for the original protein, whereas Figures S36 (b–j) are the regenerated proteins that are obtained by adding back the fraction of RC removed during derandomisation.

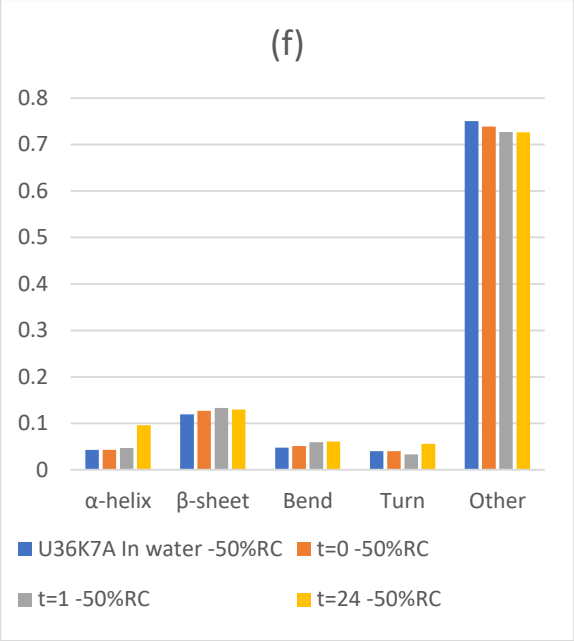


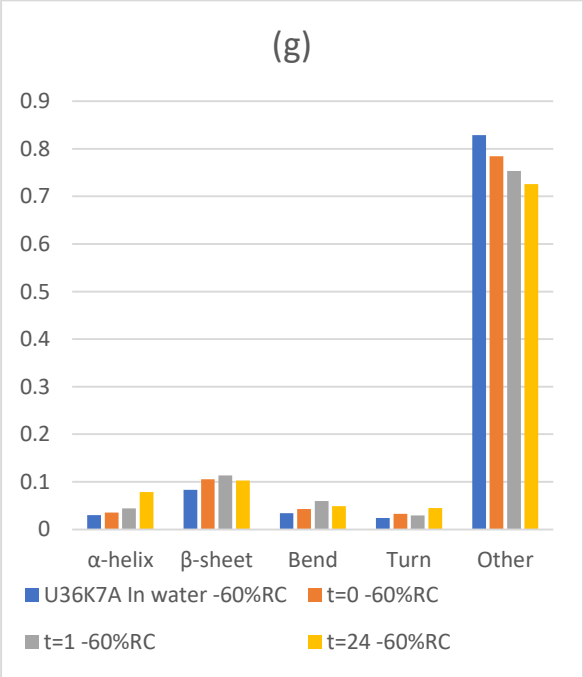


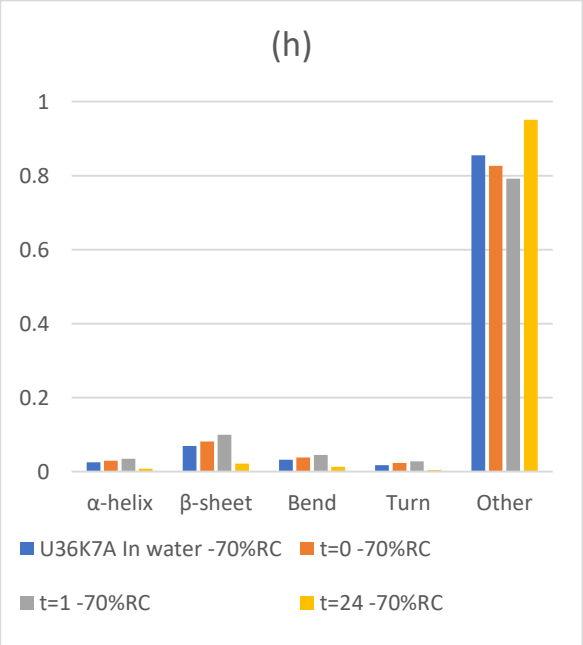





(e)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -40%RC
t=0 -40%RC
t=1 -40%RC
t=24 -40%RC


(f)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -50%RC
t=0 -50%RC
t=1 -50%RC
t=24 -50%RC


(g)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -60%RC
t=0 -60%RC
t=1 -60%RC
t=24 -60%RC


(h)
1
0.8
0.6
0.4
0.2
0
α-helix
β-sheet
Bend
Turn
Other
U36K7A In water -70%RC
t=0 -70%RC
t=1 -70%RC
t=24 -70%RC

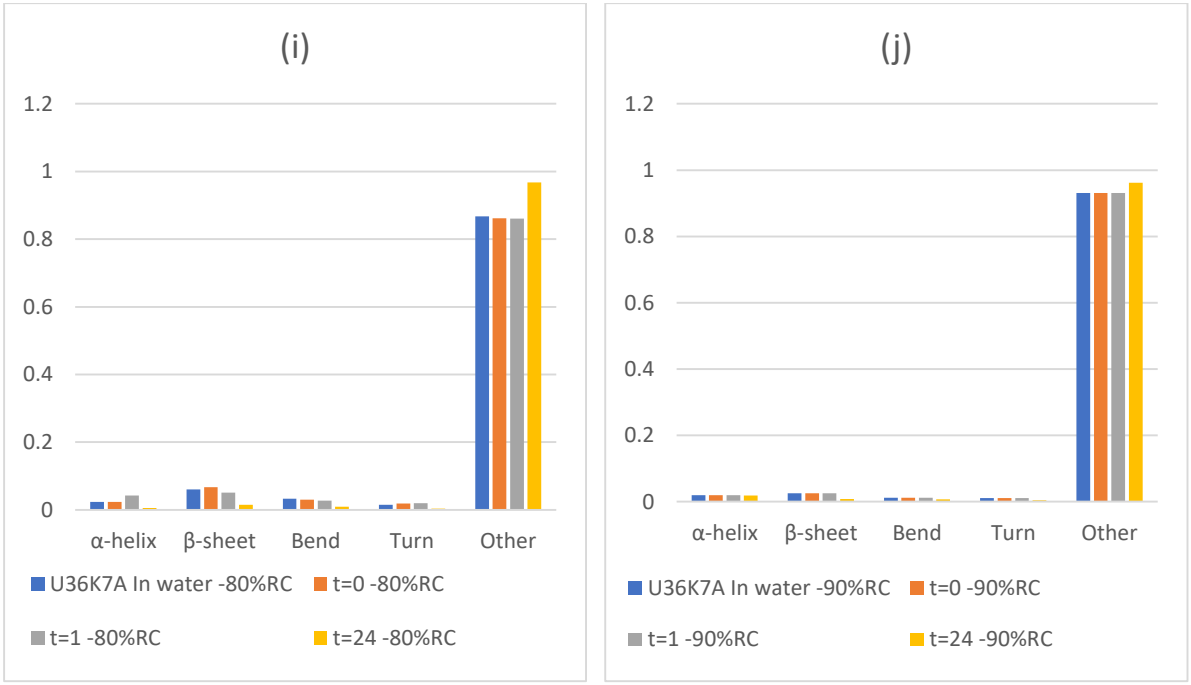


Figure S36 Secondary structure predictions for regenerated U3.6 K7A in water only and after addition of buffer over tme, created by removing 10 nm from the original spectra. (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 40%, (f) 50%, (g) 60%, (h) 70%, (i) 80%, (j) 90% RC removed for the prediction process and added back to the derandomised U3.6 K7A predictions

Table S1 Spectral NRMSDs for U3.4 wt and its seventh position variant before and after addition of buffer with 0-90%RC while using reference set and test file down to 190 nm, 195 nm and 200 nm.

Peptide	190 nm	195 nm	200 nm	Peptide	190 nm	195 nm	200 nm
U34wt In water	0.053	0.058	0.124	U34R74 In water	0.142	0.147	0.128
t = 0	0.100	0.053	0.152	t = 0	0.081	0.074	0.088
t =1	0.096	0.065	0.151	t =1	0.087	0.167	0.095
t =24	0.117	0.046	0.187	t =24	0.083	0.113	0.079
U34wt In water -10%RC	0.062	0.050	0.163	U34R7A In water -10%RC	0.160	0.137	0.172
t=0 -10%RC	0.103	0.051	0.182	t=0 -10%RC	0.074	0.118	0.118
t=1 -10%RC	0.101	0.068	0.267	t=1 -10%RC	0.083	0.118	0.073
t=24 -10%RC	0.120	0.042	0.240	t=24 -10%RC	0.099	0.127	0.076
U34wt In water -20%RC	0.066	0.042	0.161	U34R7A In water -20%RC	0.204	0.102	0.155
t=0 -20%RC	0.109	0.053	0.191	t=0 -20%RC	0.058	0.171	0.070
t=1 -20%RC	0.110	0.075	0.270	t=1 -20%RC	0.078	0.110	0.078
t=24 -20%RC	0.127	0.042	0.220	t=24 -20%RC	0.079	0.117	0.081
U34wt In water -30%RC	0.076	0.036	0.162	U34R7A In water -30%RC	0.126	0.064	0.149
t=0 -30%RC	0.120	0.063	0.206	t=0 -30%RC	0.066	0.243	0.072
t=1 -30%RC	0.125	0.103	0.179	t=1 -30%RC	0.064	0.132	0.104
t=24 -30%RC	0.139	0.043	0.139	t=24 -30%RC	0.071	0.146	0.114
U34wt In water -40%RC	0.094	0.040	0.167	U34R7A In water -40%RC	0.099	0.072	0.173
t=0 -40%RC	0.180	0.090	0.267	t=0 -40%RC	0.059	0.142	0.074
t=1 -40%RC	0.164	0.122	0.191	t=1 -40%RC	0.075	0.178	0.154
t=24 -40%RC	0.152	0.054	0.158	t=24 -40%RC	0.079	0.198	0.165
U34wt In water -50%RC	0.115	0.062	0.181	U34R7A In water -50%RC	0.085	0.071	0.150
t=0 -50%RC	0.171	0.104	0.279	t=0 -50%RC	0.073	0.180	0.113
t=1 -50%RC	0.164	0.146	0.222	t=1 -50%RC	0.078	0.255	0.235

t=24 -50%RC	0.168	0.081	0.208	t=24 -50%RC	0.086	0.281	0.250
U34wt In water -60%RC	0.149	0.115	0.257	U34R7A In water -60%RC	0.042	0.031	0.114
t=0 -60%RC	0.192	0.122	0.204	t=0 -60%RC	0.075	0.263	0.205
t=1 -60%RC	0.200	0.138	0.247	t=1 -60%RC	0.098	0.380	0.366
t=24 -60%RC	0.204	0.093	0.307	t=24 -60%RC	0.113	0.415	0.387
U34wt In water -70%RC	0.141	0.125	0.291	U34R7A In water -70%RC	0.042	0.060	0.102
t=0 -70%RC	0.234	0.146	0.240	t=0 -70%RC	0.080	0.422	0.180
t=1 -70%RC	0.244	0.096	0.203	t=1 -70%RC	0.093	0.596	0.227
t=24 -70%RC	0.209	0.097	0.461	t=24 -70%RC	0.104	0.644	0.240
U34wt In water -80%RC	0.091	0.130	0.312	U34R7A In water -80%RC	0.060	0.060	0.064
t=0 -80%RC	0.126	0.074	0.214	t=0 -80%RC	0.110	0.759	0.219
t=1 -80%RC	0.085	0.052	0.145	t=1 -80%RC	0.217	0.700	0.390
t=24 -80%RC	0.188	0.076	0.812	t=24 -80%RC	0.248	0.745	0.413
U34wt In water -90%RC	0.112	0.106	0.179	U34R7A In water -90%RC	0.119	0.119	0.160
t=0 -90%RC	0.090	0.092	0.188	t=0 -90%RC	10.477	1.791	0.369
t=1 -90%RC	0.111	0.087	0.340	t=1 -90%RC	1.551	1.570	0.581
t=24 -90%RC	0.184	0.069	1.944	t=24 -90%RC	1.105	1.667	0.607

Table S2 Spectral NRMSDs for U3.5 wt and its seventh position variant before and after addition of buffer with 0-90%RC while using reference set and test file down to 190 nm, 195 nm and 200 nm.

Peptide	190 nm	195 nm	200 nm	Peptide	190 nm	195 nm	200 nm
U35wt In water	0.106	0.053	0.098	U35R7A In water	0.125	0.100	0.108
t = 0	0.103	0.089	0.123	t = 0	0.154	0.137	0.134
t =1	0.132	0.072	0.164	t =1	0.107	0.070	0.134
t =24	0.083	0.033	0.179	t =24	0.049	0.033	0.091
U35wt In water -10%RC	0.109	0.055	0.125	U35R7A In water -10%RC	0.166	0.135	0.109
t=0 -10%RC	0.116	0.089	0.119	t=0 -10%RC	0.167	0.109	0.156
t=1 -10%RC	0.126	0.044	0.158	t=1 -10%RC	0.087	0.068	0.140
t=24 -10%RC	0.068	0.039	0.090	t=24 -10%RC	0.045	0.051	0.092
U35wt In water -20%RC	0.114	0.061	0.117	U35R7A In water -20%RC	0.172	0.162	0.116
t=0 -20%RC	0.136	0.098	0.118	t=0 -20%RC	0.166	0.079	0.145
t=1 -20%RC	0.153	0.033	0.166	t=1 -20%RC	0.067	0.052	0.149
t=24 -20%RC	0.071	0.061	0.072	t=24 -20%RC	0.064	0.068	0.108
U35wt In water -30%RC	0.125	0.068	0.110	U35R7A In water -30%RC	0.177	0.172	0.132
t=0 -30%RC	0.126	0.118	0.123	t=0 -30%RC	0.213	0.064	0.160
t=1 -30%RC	0.134	0.070	0.157	t=1 -30%RC	0.042	0.025	0.098
t=24 -30%RC	0.078	0.111	0.087	t=24 -30%RC	0.057	0.113	0.081
U35wt In water -40%RC	0.189	0.088	0.106	U35R7A In water -40%RC	0.196	0.140	0.180
t=0 -40%RC	0.157	0.142	0.139	t=0 -40%RC	0.118	0.116	0.171
t=1 -40%RC	0.084	0.049	0.164	t=1 -40%RC	0.040	0.038	0.089
t=24 -40%RC	0.074	0.198	0.071	t=24 -40%RC	0.062	0.191	0.072
U35wt In water -50%RC	0.166	0.106	0.114	U35R7A In water -50%RC	0.235	0.129	0.161
t=0 -50%RC	0.139	0.085	0.153	t=0 -50%RC	0.081	0.088	0.160
t=1 -50%RC	0.054	0.029	0.123	t=1 -50%RC	0.058	0.093	0.125
t=24 -50%RC	0.066	0.324	0.075	t=24 -50%RC	0.065	0.115	0.073

U35wt In water -60%RC	0.180	0.168	0.204	U35R7A In water -60%RC	0.193	0.075	0.220
t=0 -60%RC	0.176	0.054	0.169	t=0 -60%RC	0.047	0.043	0.089
t=1 -60%RC	0.066	0.046	0.118	t=1 -60%RC	0.052	0.202	0.063
t=24 -60%RC	0.092	0.206	0.088	t=24 -60%RC	0.056	0.155	0.135
U35wt In water -70%RC	0.226	0.145	0.177	U35R7A In water -70%RC	0.067	0.035	0.107
t=0 -70%RC	0.102	0.069	0.179	t=0 -70%RC	0.057	0.113	0.075
t=1 -70%RC	0.081	0.147	0.083	t=1 -70%RC	0.050	0.141	0.092
t=24 -70%RC	0.094	0.306	0.200	t=24 -70%RC	0.069	0.266	0.261
U35wt In water -80%RC	0.123	0.176	0.140	U35R7A In water -80%RC	0.063	0.163	0.042
t=0 -80%RC	0.048	0.066	0.126	t=0 -80%RC	0.042	0.369	0.064
t=1 -80%RC	0.059	0.134	0.102	t=1 -80%RC	0.060	0.284	0.276
t=24 -80%RC	0.109	0.549	0.123	t=24 -80%RC	0.087	0.523	0.158
U35wt In water -90%RC	0.065	0.252	0.088	U35R7A In water -90%RC	0.108	0.863	0.078
t=0 -90%RC	0.061	0.154	0.129	t=0 -90%RC	0.058	0.417	0.115
t=1 -90%RC	0.116	0.474	0.154	t=1 -90%RC	0.120	0.818	0.224
t=24 -90%RC	0.319	1.331	0.190	t=24 -90%RC	0.321	0.896	0.468

Table S1 Spectral NRMSDs for U3.6 wt and its seventh position variant before and after addition of buffer with 0-90%RC while using reference set and test file down to 190 nm, 195 nm and 200 nm.

Peptide	190 nm	195 nm	200 nm	Peptide	190 nm	195 nm	200 nm
U36wt In water	0.077	0.064	0.138	U36K7A In water	0.056	0.052	0.149
t = 0	0.095	0.084	0.138	t = 0	0.102	0.039	0.177
t =1	0.097	0.082	0.135	t =1	0.111	0.095	0.210
t =24	0.104	0.117	0.149	t =24	0.048		0.227
U36wt In water -10%RC	0.092	0.072	0.132	U36K7A In water -10%RC	0.069	0.052	0.151
t=0 -10%RC	0.106	0.102	0.135	t=0 -10%RC	0.109	0.043	0.182
t=1 -10%RC	0.109	0.097	0.131	t=1 -10%RC	0.123	0.091	0.267
t=24 -10%RC	0.112	0.131	0.148	t=24 -10%RC	0.061		0.240
U36wt In water -20%RC	0.144	0.080	0.125	U36K7A In water -20%RC	0.073	0.057	0.155
t=0 -20%RC	0.153	0.105	0.134	t=0 -20%RC	0.120	0.055	0.191
t=1 -20%RC	0.149	0.102	0.130	t=1 -20%RC	0.131	0.090	0.270
t=24 -20%RC	0.149	0.156	0.150	t=24 -20%RC	0.100	0.092	0.220
U36wt In water -30%RC	0.148	0.091	0.120	U36K7A In water -30%RC	0.088	0.059	0.162
t=0 -30%RC	0.138	0.119	0.136	t=0 -30%RC	0.125	0.060	0.206
t=1 -30%RC	0.135	0.117	0.133	t=1 -30%RC	0.149	0.107	0.179
t=24 -30%RC	0.150	0.174	0.157	t=24 -30%RC	0.160	0.083	0.139
U36wt In water -40%RC	0.162	0.102	0.120	U36K7A In water -40%RC	0.090	0.075	0.175
t=0 -40%RC	0.138	0.165	0.147	t=0 -40%RC	0.138	0.072	0.267
t=1 -40%RC	0.137	0.163	0.144	t=1 -40%RC	0.142	0.111	0.191
t=24 -40%RC	0.171	0.137	0.172	t=24 -40%RC	0.243	0.088	0.158
U36wt In water -50%RC	0.157	0.134	0.131	U36K7A In water -50%RC	0.074	0.084	0.191
t=0 -50%RC	0.167	0.105	0.180	t=0 -50%RC	0.160	0.068	0.279
t=1 -50%RC	0.163	0.107	0.169	t=1 -50%RC	0.117	0.102	0.222
t=24 -50%RC	0.150	0.105	0.200	t=24 -50%RC	0.130	0.135	0.208

U36wt In water -60%RC	0.227	0.112	0.173	U36K7A In water -60%RC	0.075	0.074	0.218
t=0 -60%RC	0.172	0.050	0.179	t=0 -60%RC	0.197	0.058	0.204
t=1 -60%RC	0.175	0.067	0.170	t=1 -60%RC	0.109	0.041	0.247
t=24 -60%RC	0.143	0.084	0.216	t=24 -60%RC	0.128	0.230	0.307
U36wt In water -70%RC	0.107	0.083	0.192	U36K7A In water -70%RC	0.078	0.049	0.282
t=0 -70%RC	0.113	0.072	0.181	t=0 -70%RC	0.152	0.052	0.240
t=1 -70%RC	0.123	0.078	0.169	t=1 -70%RC	0.111	0.054	0.203
t=24 -70%RC	0.150	0.096	0.211	t=24 -70%RC	0.192	0.241	0.461
U36wt In water -80%RC	0.045	0.041	0.127	U36K7A In water -80%RC	0.060	0.046	0.267
t=0 -80%RC	0.073	0.060	0.116	t=0 -80%RC	0.190	0.058	0.214
t=1 -80%RC	0.076	0.068	0.103	t=1 -80%RC	0.109	0.048	0.145
t=24 -80%RC	0.147	0.092	0.138	t=24 -80%RC	0.219	0.448	0.812
U36wt In water -90%RC	0.075	0.265	0.074	U36K7A In water -90%RC	0.064	0.037	0.202
t=0 -90%RC	0.075	0.128	0.101	t=0 -90%RC	0.230	0.067	0.188
t=1 -90%RC	0.076	0.140	0.094	t=1 -90%RC	0.090		0.340
t=24 -90%RC	0.122	0.156	0.125	t=24 -90%RC	0.580	1.130	1.944