Amino Acid Metaclusters: Implications of Growth Trends on Peptide Self-Assembly Structure

*Thanh D. Do, Natália E. C. de Almeida, Nichole E. LaPointe,† Ali Chamas, Stuart C. Feinstein,† Michael T. Bowers**

Department of Chemistry and Biochemistry, and †Neuroscience Research Institute and Department of Molecular Cellular and Developmental Biology, University of California, Santa Barbara, California 93106, United States.

*M. T. Bowers. Email: bowers@chem.ucsb.edu. Tel: +1-805-893-2673

SUPPORTING INFORMATION

S1. Materials and Methods

S1.1 Sample Preparation

The amino acids (L-serine, L-isoleucine and L-asparagine) were purchased from Sigma Aldrich (St. Louis, MO). The spray samples used in IM-MS experiments were prepared at concentrations of 300 μM to 6 mM in pure water (JT Baker). The hexa-peptides were custom synthesized by Genscript Corp. (Piscataway, NJ) with purity >90%. For IM-MS experiments, the peptide powders were dissolved in water adjusted to pH = 8 with ammonium hydroxide. The concentration of the peptide in the spray sample was 200 μM. For TEM, peptide powers were dissolved in H_2O/DMSO to a final concentration of 200 μM.

S1.2 Ion-mobility Mass Spectrometry

Samples were loaded into gold coated borosilicate capillaries that had been pulled to a fine point on an in-house micropipette puller (Sutter Instrument Co., Novato, CA). A positive voltage was applied to the tip and ions were generated through the means of nano-electrospray ionization and entered the instrument capillary through a small orifice (0.010" id) and stored in an ion funnel. In ion-mobility experiments, the ions were pulsed into the drift cell so that the initial time t_0 could be set. The ions travelled through the drift cell under the influence of a weak electrical field and the drag force due to multiple collisions with helium buffer gas. At the end of the drift cell, the ions were collected into an exit funnel and steered into a quadrupole mass filter, at which the arrival time t_A was determined. The arrival time distributions (ATDs) were measured at different drift voltages for each mass-selected peak. Each ATD could contain multiple features corresponding to oligomers at different sizes or distinct conformers of the same oligomer. The

reduced mobility of each species could be determined from the ATDs by plotting t_A as a function of P/V where P is pressure and V is drift voltage.[1]

$$t_A = \frac{L^2 T_0}{K_0 P_0 T}\frac{P}{V} + t_0 \ (Eq.\,\mathrm{S1})$$

where L is the length of the drift cell (200 cm), P_0 is 760 torr and T_0 is 273.15 K. From the reduced mobility, the experimental cross section which reflects the size and shape of the ion can be approximated using Eq. 2.

$$\sigma \approx \frac{3q}{16N}\left(\frac{2\pi}{\mu k_B T}\right)^{\frac{1}{2}}\frac{1}{K_0} \ (Eq.\,S2)$$

Here, N is the buffer gas number density, μ is the reduced mass of the collision system (ion + helium), k_B is Boltzmann's factor and T is the drift cell temperature. The experimental cross section σ is often reported in the unit of Å^2. The flux of ions exiting the drift tube can be calculated. It is assumed that the ion packet takes the form of a periodic delta function and the flux is given by Eq. 3.[2]

$$\varphi(0,z,t) = \frac{s\cdot a\cdot e^{-\alpha t}}{4(\pi D_L t)^{1/2}}\cdot\left(v_d + \frac{z}{t}\right)\cdot\left[1 - e^{\left(-r_0{}^2/4D_T t\right)}\right]\cdot e^{\left(-(z-v_d t)^2/4D_{Lt}\right)} \ (Eq.\,S3)$$

Where z is the distance the ions travel, r_0 is the radius of the initial ion packet, a is the area of the exit aperture, D_L and D_T are the longitudinal and transverse diffusion coefficients, s is the initial ion density and α is the loss of ions due to the reactions in the drift tube.[2] The line shape generated from Eq. 3 would correspond to that of an ion packet composed of a single ion conformation shown in the ATDs. Experiment line shapes broader than this limit indicate more than one conformation/oligomeric species is generating the experimental peak.

S1.3 Tables of Cross Sections

Table S1: L-Serine metaclusters obtained from experiment 1 at P = 13 torr

m/z	N	Z	Metacluster	Cross section σ, Å^2
106	1	1	$[1(Ser)+H]^{+}$	52
210	2	1	$[2(Ser)+H]^{+}$	82
421	8	2	$[8(Ser)+2H]^{2+}$	196
421	4	1	$[4(Ser)+1H]^{2+}$	110
473	9	2	$[9(Ser)+2H]^{2+}$	222
578	11	2	$[11(Ser)+2H]^{2+}$	238
689	13	2	$[13(Ser)+2H]^{2+}$	263
736	21	3	$[21(Ser)+3H]^{3+}$	356
736	14	2	$[14(Ser)+2H]^{2+}$	270
771	22	3	$[22(Ser)+3H]^{3+}$	366
788	15	2	$[15(Ser)+2H]^{2+}$	277
806	23	3	$[23(Ser)+3H]^{3+}$	377
841	24	3	$[24(Ser)+3H]^{3+}$	385
877	25	3	$[25(Ser)+3H]^{3+}$	394
911	26	3	$[26(Ser)+3H]^{3+}$	406
946	18	2	$[18(Ser)+2H]^{2+}$	306
946	27	3	$[27(Ser)+3H]^{3+}$	402
981	28	3	$[28(Ser)+3H]^{3+}$	416
1016	29	3	$[29(Ser)+3H]^{3+}$	424
1051	20	2	$[20(Ser)+2H]^{2+}$	328
1051	30	3	$[30(Ser)+3H]^{3+}$	429
1051	41	4	$[40(Ser)+4H]^{4+}$	522
1086	31	3	$[31(Ser)+3H]^{3+}$	439

1104	21	2	$[21(Ser)+2H]^{2+}$	340
1104	42	4	$[42(Ser)+4H]^{4+}$	539
1121	32	3	$[32(Ser)+3H]^{3+}$	446
1129	43	4	$[43(Ser)+4H]^{4+}$	538
1156	22	2	$[22(Ser)+2H]^{2+}$	345
1156	33	3	$[33(Ser)+3H]^{3+}$	452
1156	44	4	$[44(Ser)+4H]^{4+}$	550
1182	45	4	$[45(Ser)+4H]^{4+}$	555
1191	34	3	$[34(Ser)+3H]^{3+}$	462
1207	46	4	$[46(Ser)+4H]^{4+}$	561
1226	35	3	$[35(Ser)+3H]^{3+}$	471
1233	47	4	$[47(Ser)+4H]^{4+}$	565
1261	36	3	$[36(Ser)+3H]^{3+}$	476
1261	48	4	$[48(Ser)+4H]^{4+}$	584
1286	49	4	$[49(Ser)+4H]^{4+}$	582
1314	50	4	$[50(Ser)+4H]^{4+}$	590
1366	39	3	$[39(Ser)+3H]^{3+}$	501
1366	52	4	$[52(Ser)+3H]^{4+}$	612
1366	65	5	$[65(Ser)+5H]^{5+}$	707

Table S2: L-Asparagine metaclusters obtained at P = 12-13 torr

m/z	**N**	**Z**	**Metacluster**	**Cross section σ, Å^2**
133	1	1	$[1(Asn)+H]^{+}$	59
264	2	1	$[2(Asn)+H]^{+}$	94
528	8	2	$[8(Asn)+H]^{2+}$	247
594	9	2	$[9(Asn)+H]^{2+}$	261

660	10	2	$[10(Asn)+2H]^{2+}$	270
660	15	3	$[15(Asn)+3H]^{3+}$	360
704	16	3	$[16(Asn)+3H]^{3+}$	372
748	17	3	$[17(Asn)+3H]^{3+}$	381
794	12	2	$[12(Asn)+2H]^{2+}$	303
794	18	3	$[18(Asn)+3H]^{3+}$	393
839	19	3	$[19(Asn)+3H]^{3+}$	407
884	20	3	$[20(Asn)+3H]^{3+}$	417
929	14	2	$[14(Asn)+2H]^{2+}$	309
929	21	3	$[21(Asn)+3H]^{3+}$	412
972	22	3	$[22(Asn)+3H]^{3+}$	422
990	30	4	$[30(Asn)+4H]^{4+}$	533
1023	31	4	$[31(Asn)+4H]^{3+}$	545
1016	23	3	$[23(Asn)+3H]^{3+}$	444
1056	24	3	$[24(Asn)+3H]^{3+}$	454
1056	32	4	$[32(Asn)+4H]^{4+}$	553
1056	40	5	$[40(Asn)+5H]^{5+}$	640
1089	33	4	$[33(Asn)+4H]^{4+}$	562
1122	34	4	$[34(Asn)+4H]^{3+}$	572
1155	35	4	$[35(Asn)+4H]^{4+}$	585
1189	27	3	$[27(Asn)+3H]^{3+}$	507
1189	36	4	$[36(Asn)+4H]^{4+}$	613
1189	45	5	$[45(Asn)+5H]^{5+}$	716
1212	37	4	$[37(Asn)+4H]^{4+}$	592
1254	38	4	$[38(Asn)+4H]^{4+}$	608
1287	39	4	$[39(Asn)+4H]^{4+}$	619

1320	40	4	$[40(Asn)+4H]^{4+}$	621
1320	50	5	$[50(Asn)+5H]^{5+}$	732
1320	60	6	$[60(Asn)+6H]^{6+}$	844
1354	41	4	$[41(Asn)+4H]^{4+}$	628
1453	44	4	$[44(Asn)+4H]^{4+}$	649
1453	55	5	$[55(Asn)+5H]^{5+}$	761
1453	66	6	$[66(Asn)+6H]^{6+}$	865

Table S3: L-Tryptophan metaclusters obtained at P = 12-13 torr

m/z	**N**	**Z**	**Metacluster**	**Cross section σ, Å^2**
205	1	1	$[1(Trp)+H]^{+}$	80
409	2	1	$[2(Trp)+H]^{+}$	122
612	3	1	$[3(Trp)+1H]^{+}$	164
612	6	2	$[6(Trp)+2H]^{2+}$	264
715	7	2	$[7(Trp)+2H]^{2+}$	283
817	4	1	$[4(Trp)+1H]^{+}$	194
817	8	2	$[8(Trp)+2H]^{2+}$	314
817	12	3	$[12(Trp)+3H]^{3+}$	424
884	13	3	$[13(Trp)+3H]^{3+}$	445
919	9	2	$[9(Trp)+2H]^{2+}$	341
953	14	3	$[14(Trp)+3H]^{3+}$	465
1021	10	2	$[10(Trp)+2H]^{2+}$	368
1021	15	3	$[15(Trp)+3H]^{3+}$	484
1021	15	3	$[15(Trp)+3H]^{3+}$	466
1089	16	3	$[16(Trp)+3H]^{3+}$	506

1123	11	2	$[11(Trp)+2H]^{2+}$	388
1157	17	3	$[17(Trp)+3H]^{3+}$	526
1225	12	2	$[12(Trp)+2H]^{2+}$	415
1225	18	3	$[18(Trp)+3H]^{3+}$	546
1225	24	4	$[24(Trp)+4H]^{4+}$	672
1277	25	4	$[25(Trp)+4H]^{4+}$	687
1294	19	3	$[19(Trp)+3H]^{3+}$	565
1328	13	2	$[13(Trp)+2H]^{2+}$	430
1328	26	4	$[26(Trp)+4H]^{4+}$	703
1362	20	3	$[20(Trp)+3H]^{3+}$	584
1379	27	4	$[27(Trp)+4H]^{4+}$	720
1430	14	2	$[14(Trp)+2H]^{2+}$	445
1430	21	3	$[21(Trp)+3H]^{3+}$	600
1430	28	4	$[28(Trp)+4H]^{4+}$	737

Table S4: L-Isoleucine metaclusters obtained at P = 12-13 torr

m/z	**N**	**Z**	**Metacluster**	**Cross section σ, Å^2**
132	1	1	$[1(Ile)+H]^{+}$	65
263	2	1	$[2(Ile)+H]^{+}$	107
590	9	2	$[9(Ile)+2H]^{2+}$	303
590	9	2	$[9(Ile)+2H]^{2+}$	316
656	5	1	$[5(Ile)+H]^{+}$	180
656	10	2	$[10(Ile)+2H]^{2+}$	313
656	15	3	$[15(Ile)+3H]^{3+}$	405
721	11	2	$[11(Ile)+H]^{2+}$	327
787	12	2	$[12(Ile)+2H]^{2+}$	357

787	18	3	$[18(Ile)+3H]^{3+}$	484
831	19	3	$[19(Ile)+3H]^{3+}$	507
853	13	2	$[13(Ile)+2H]^{2+}$	376
853	26	4	$[26(Ile)+4H]^{4+}$	670
874	20	3	$[20(Ile)+3H]^{3+}$	525
918	14	2	$[14(Ile)+2H]^{2+}$	405
918	21	3	$[21(Ile)+3H]^{3+}$	550
961	22	3	$[22(Ile)+3H]^{3+}$	568
1012	31	4	$[31(Ile)+4H]^{4+}$	757
1049	16	2	$[16(Ile)+2H]^{2+}$	442
1049	24	3	$[24(Ile)+3H]^{3+}$	612
1049	32	4	$[32(Ile)+4H]^{4+}$	787
1049	40	5	$[40(Ile)+5H]^{5+}$	927
1081	33	4	$[33(Ile)+4H]^{4+}$	800
1081	33	4	$[33(Ile)+4H]^{4+}$	769
1115	34	4	$[34(Ile)+4H]^{4+}$	822
1115	34	4	$[34(Ile)+4H]^{4+}$	777
1137	26	3	$[26(Ile)+3H]^{3+}$	649
1132	52	6	$[52(Ile)+6H]^{6+}$	1199
1181	18	3	$[18(Ile)+3H]^{3+}$	483
1181	36	4	$[36(Ile)+4H]^{4+}$	815
1181	45	5	$[45(Ile)+5H]^{5+}$	982
1246	38	4	$[38(Ile)+4H]^{4+}$	902
1246	57	6	$[57(Ile)+6H]^{6+}$	1287
1246	76	8	$[76(Ile)+8H]^{8+}$	1654
1246	95	10	$[95(Ile)+10H]^{10+}$	1902

S1.4 Representative ATDs

S1.4.1 Serine

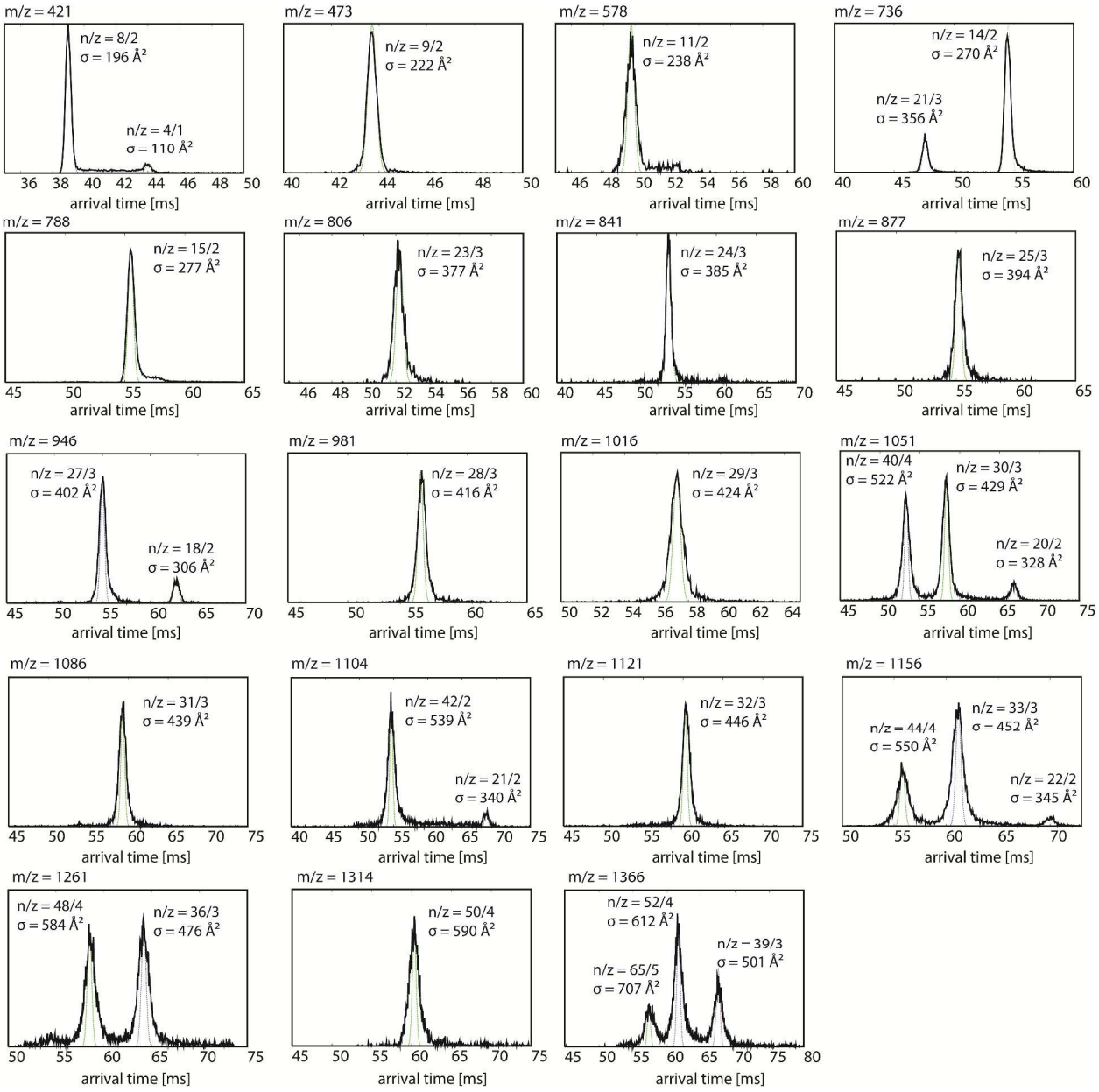

Figure S1. Representative ATDs of L-serine metaclusters. Each ATD feature is annotated with *n/z* where *n* is the cluster number and *z* is the charge, and experimental cross sections. The data were collected at P = 13 torr.

S1.4.2 Tryptophan

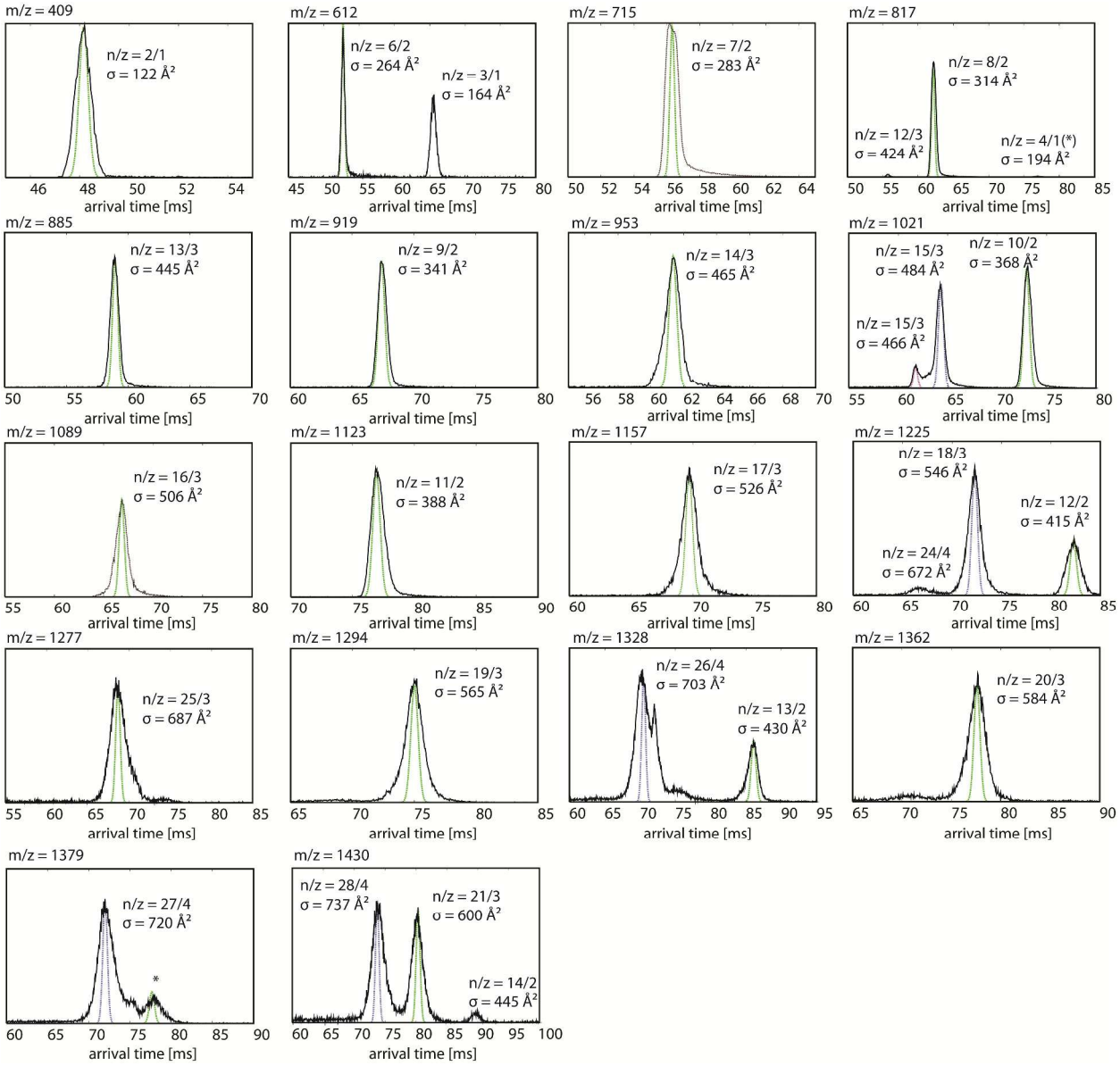

Figure S2. Representative ATDs of L-tryptophan mass spectral peaks. Each ATD feature is annotated with *n/z* where *n* is the cluster number and *z* is the charge, and experimental cross sections. The data were collected at P = 12-13 torr.

S1.4.3 Isoleucine

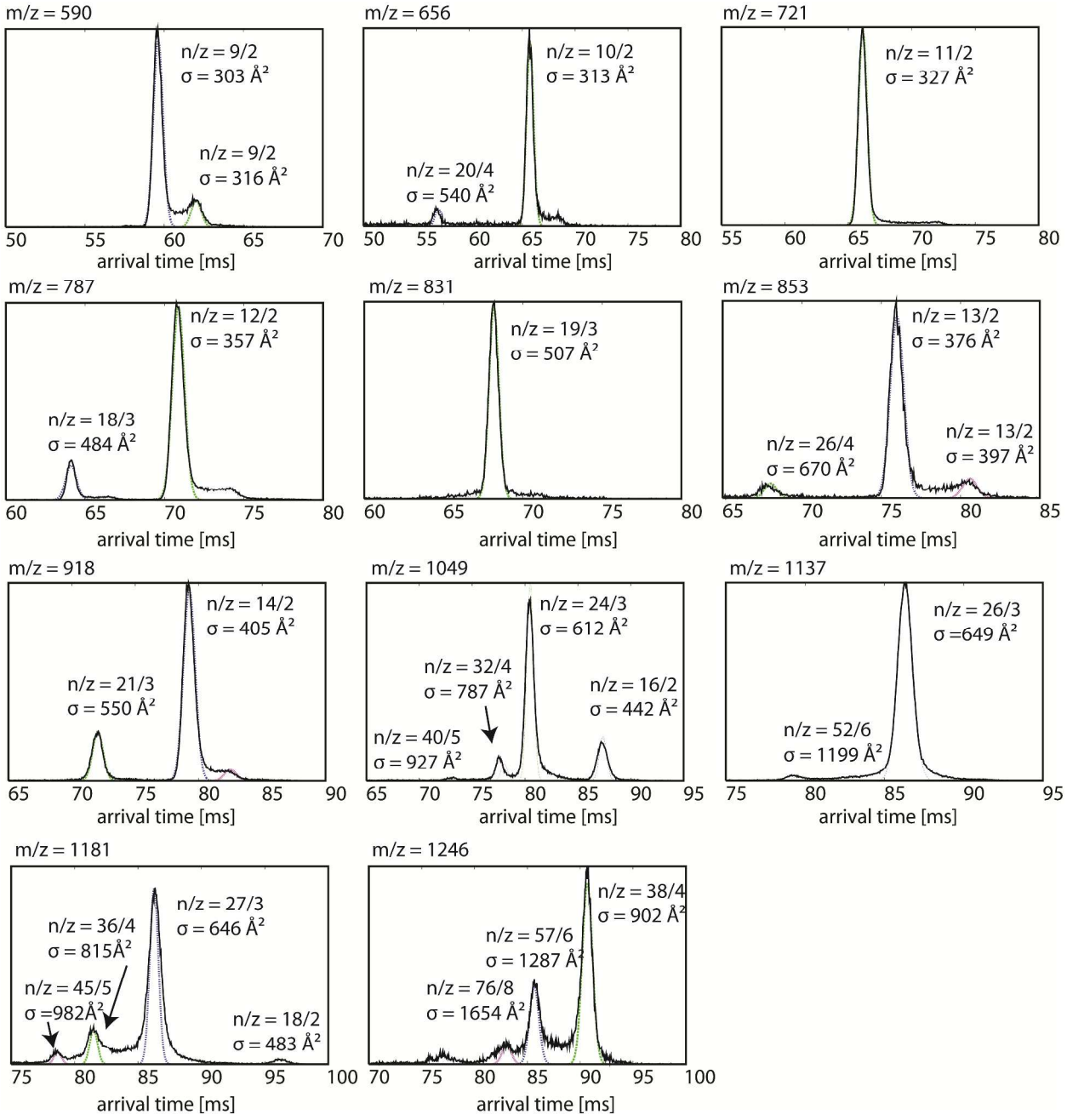

Figure S3. Representative ATDs of L-isoleucine mass spectral peaks. Each ATD feature is annotated with *n/z* where *n* is the cluster number and *z* is the charge, and experimental cross sections. The data were collected at P = 12-13 torr.

S1.4.4 Asparagine

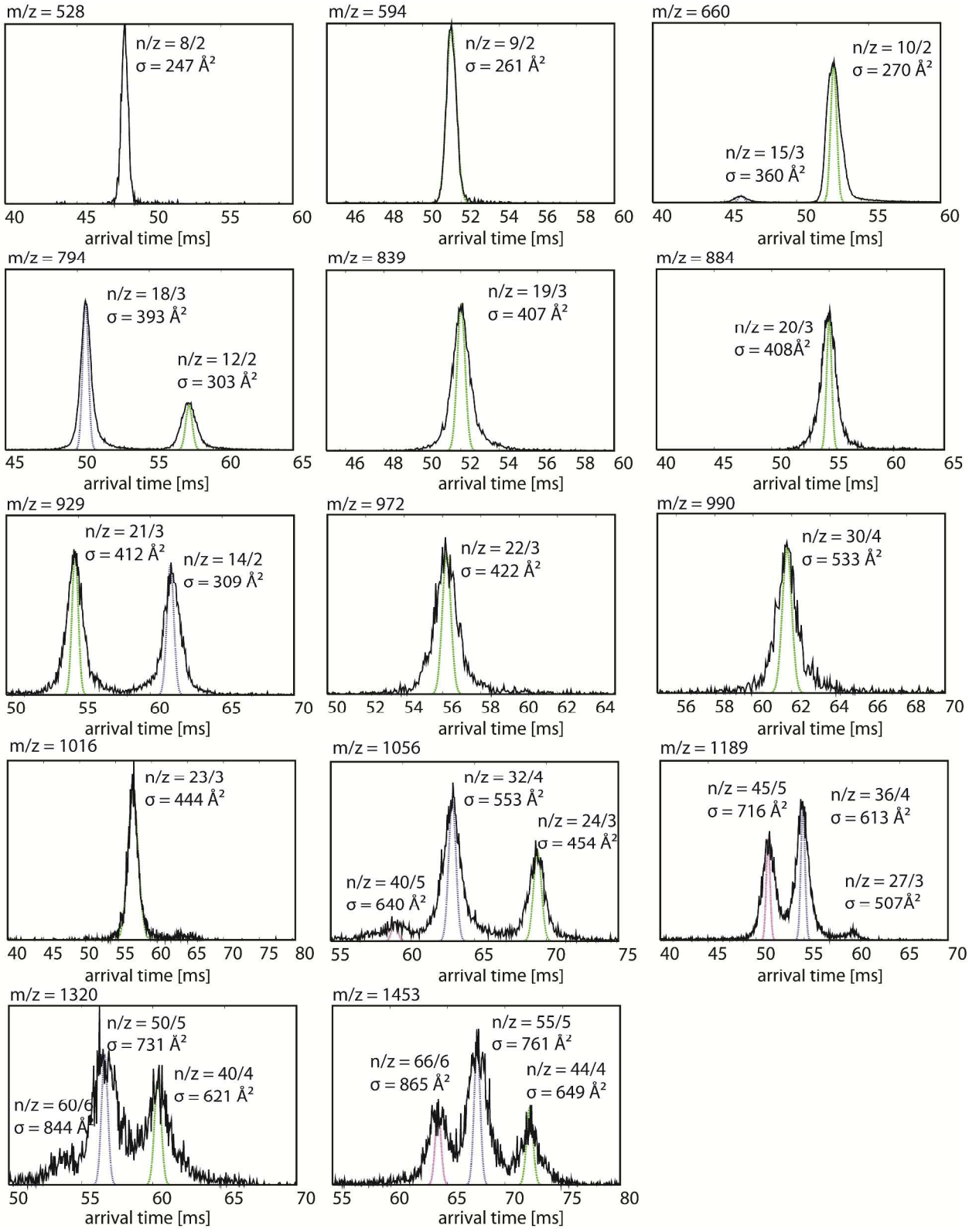

Figure S4. Representative ATDs of L-asparagine mass spectral peaks. Each ATD feature is annotated with *n/z* where *n* is the cluster number and *z* is the charge, and experimental cross sections. The data were collected at P = 12-13 torr.

S1.4.5 NININI

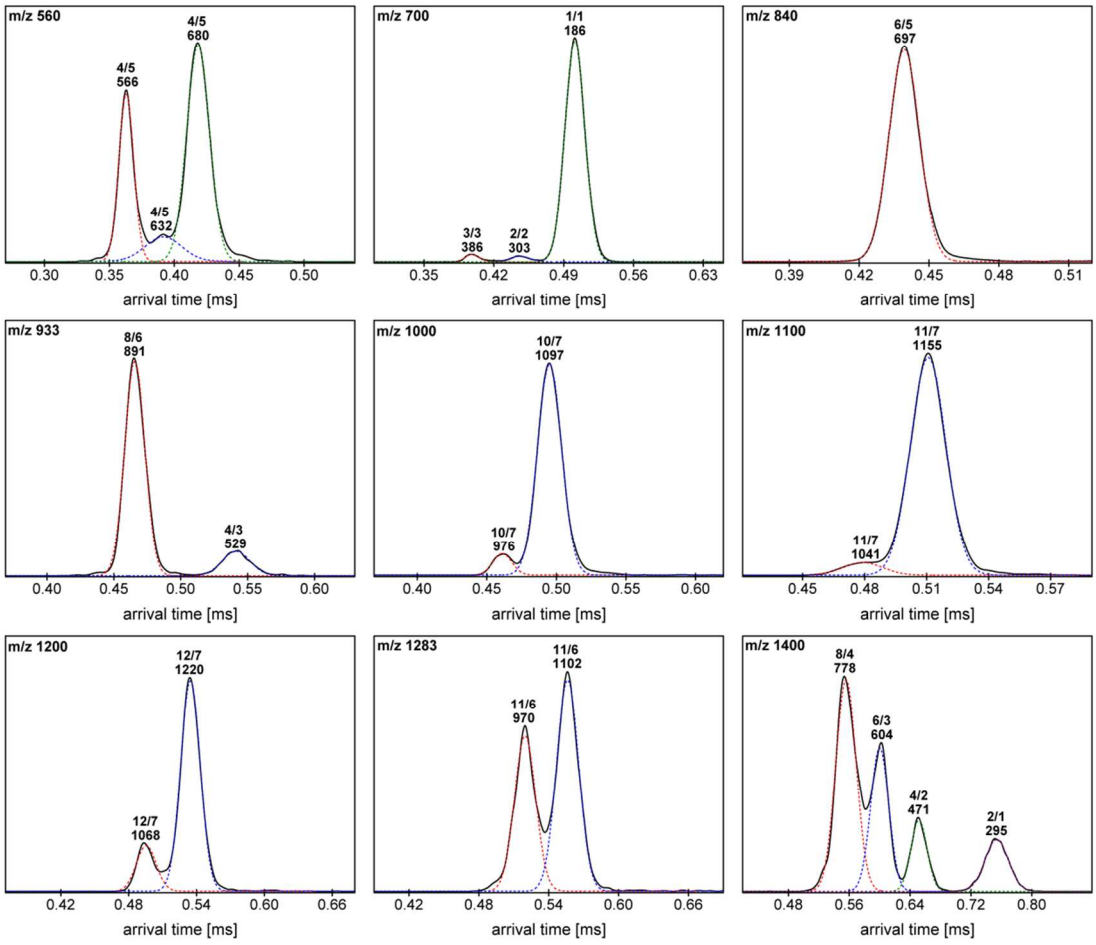

Figure S5. Representative ATDs of NININI (200 μM). The peak at *m/z* 560 contains tetramer species with collisional cross sections (CCS) of 566, 632 and 680 Å^2, whereas the peak at *m/z* 700 displayed three different oligomeric species comprising monomer (186 Å^2), dimer (303 Å^2) and trimer (386 Å^2). The peak at *m/z* 840 contains a single hexamer with a CCS of 697 Å^2 and the other abundant peaks are *m/z* 933 (tetramer and octamer, 529 and 891 Å^2, respectively), 1000

(decamers, 976 and 1097 Å^2), 1100 (undecamers, 1041 and 1155 Å^2), 1200 (dodecamers, 1068 and 1220 Å^2), 1283 (undecamers, 970 and 1102 Å^2). The minor peak at *m/z* 1400 contains dimer (295 Å^2), tetramer (471 Å^2), hexamer (604 Å^2) and octamer (778 Å^2).

S1.4.6 SISISI

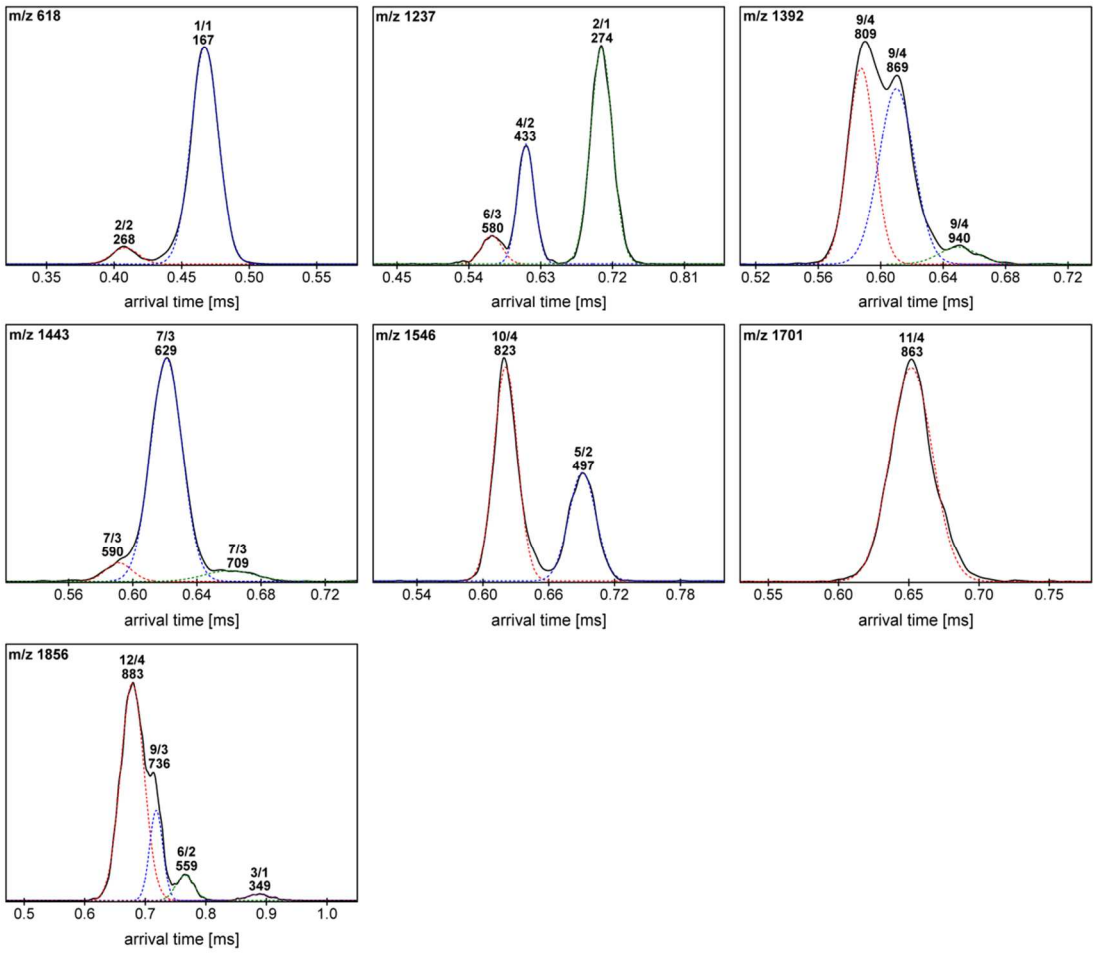

Figure S6. Representative ATDs of SISISI (200 μM). The most abundant peaks are m/z 618 (monomer and dimer, 167 and 268 Å^2, respectively), 1237 (dimer, tetramer and hexamer, 274, 433 and 580 Å^2, respectively), 1392 (nonamers, 809, 869 and 940 Å^2), 1443 (heptamers, 590, 629 and 709 Å^2), 1546 (pentamer and decamer, 497 and 823 Å^2, respectively), 1701 (undecamer,

863 Å^2) and 1856 (trimer, hexamer, nonamer and dodecamer, 349, 559, 736 and 883 Å^2, respectively).

S1.4.7 SFSFSF

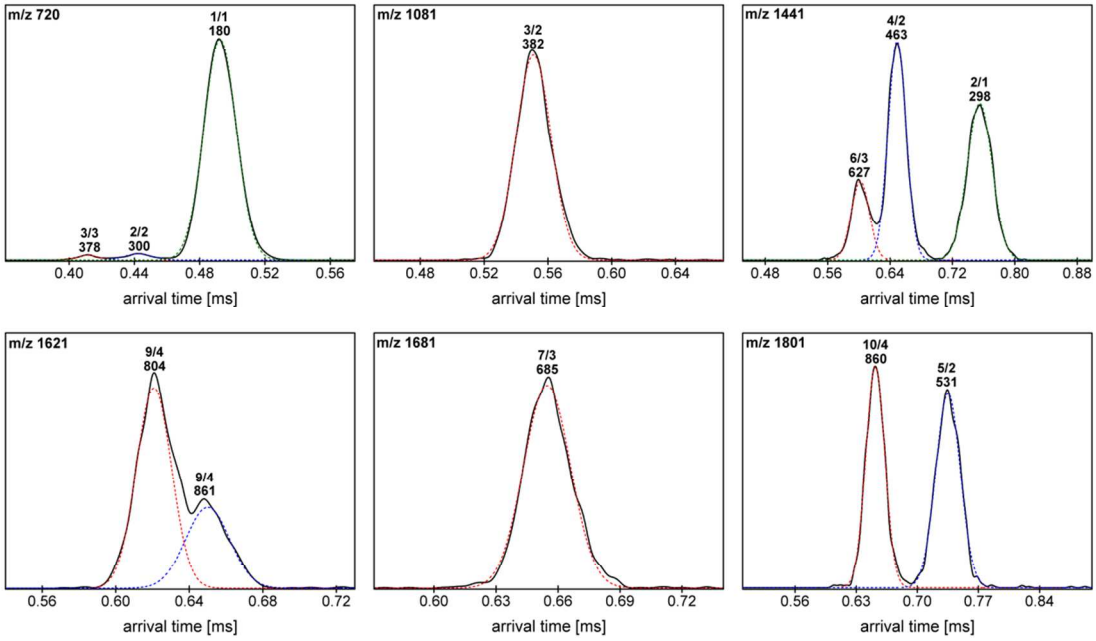

Figure S7. Representative ATDs of SFSFSF (200 μM). The peak at m/z 720 was the most abundant and displayed three different oligomeric species, monomer (180 Å^2), dimer (300 Å^2) and trimer (378 Å^2), whereas the peak at m/z 1081 containso a single trimer with a CCS of 382 Å^2. Other minor peaks are m/z 1441 (dimer, tetramer and hexamer, 298, 463 and 627 Å^2, respectively), 1621 (nonamers, 804 and 861 Å^2), 1681 (heptamer, 685 Å^2) and 1801 (pentamer and decamer, 531 and 860 Å^2, respectively).

S1.4.8 Additional IM-MS data

Serine metacluster data obtained in negative mode ESI polarity and at the concentration of 250 μM.

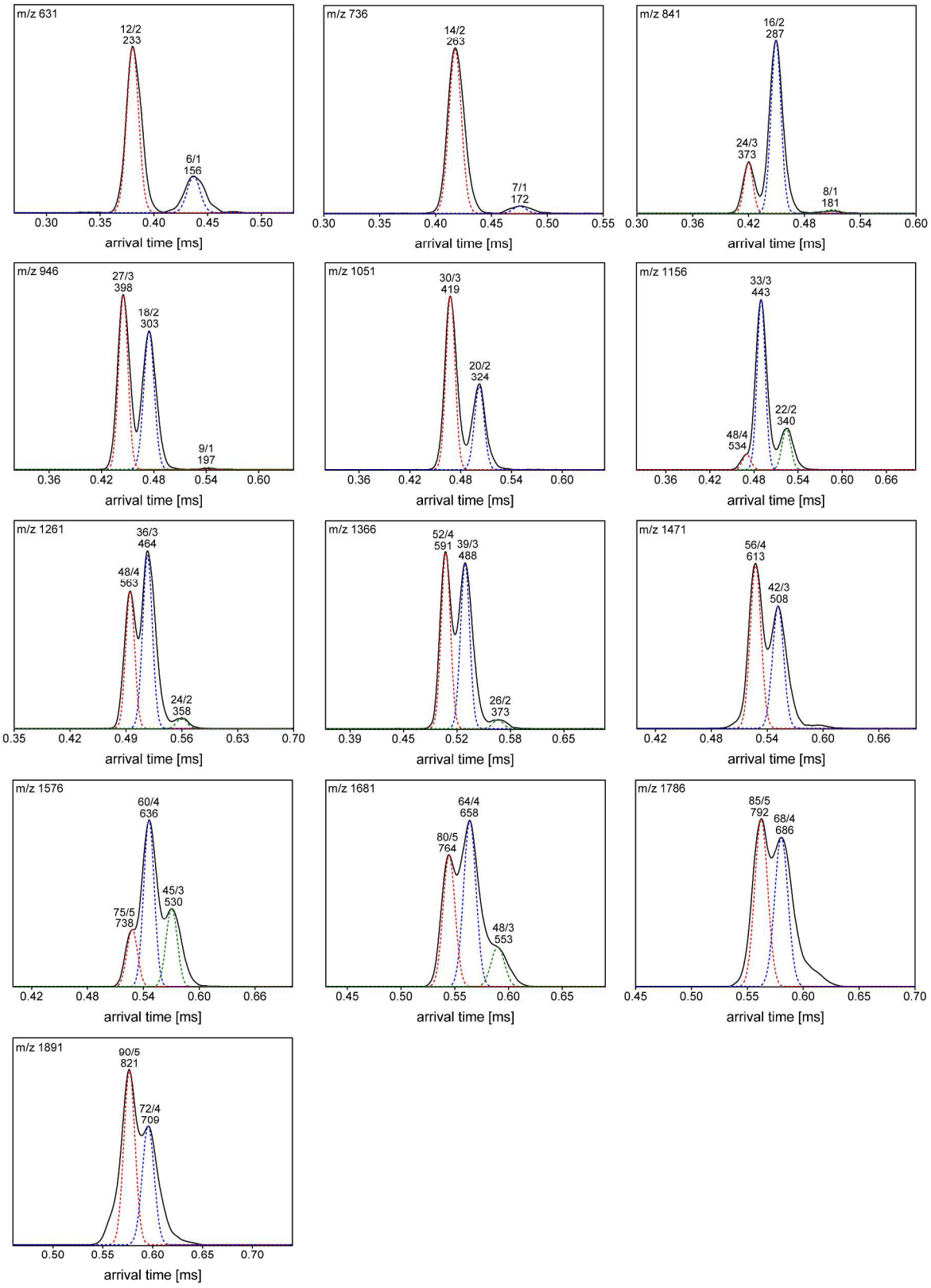
m/z 631
12/2
233
6/1
156
0.30 0.35 0.40 0.45 0.50
arrival time [ms]
m/z 736
14/2
263
7/1
172
0.30 0.35 0.40 0.45 0.50 0.55
arrival time [ms]
m/z 841
16/2
287
24/3
373
8/1
181
0.30 0.36 0.42 0.48 0.54 0.60
arrival time [ms]
m/z 946
27/3
398
18/2
303
9/1
197
0.36 0.42 0.48 0.54 0.60
arrival time [ms]
m/z 1051
30/3
419
20/2
324
0.36 0.42 0.48 0.54 0.60
arrival time [ms]
m/z 1156
33/3
443
22/2
340
48/4
534
0.36 0.42 0.48 0.54 0.60 0.66
arrival time [ms]
m/z 1261
36/3
464
48/4
563
24/2
358
0.35 0.42 0.49 0.56 0.63 0.70
arrival time [ms]
m/z 1366
52/4
591
39/3
488
26/2
373
0.39 0.45 0.52 0.58 0.65
arrival time [ms]
m/z 1471
56/4
613
42/3
508
0.42 0.48 0.54 0.60 0.66
arrival time [ms]
m/z 1576
60/4
636
45/3
530
75/5
738
0.42 0.48 0.54 0.60 0.66
arrival time [ms]
m/z 1681
64/4
658
80/5
764
48/3
553
0.45 0.50 0.55 0.60 0.65
arrival time [ms]
m/z 1786
85/5
792
68/4
686
0.45 0.50 0.55 0.60 0.65 0.70
arrival time [ms]
m/z 1891
90/5
821
72/4
709
0.50 0.55 0.60 0.65 0.70
arrival time [ms]

Figure S8. Representative ATDs of Ser metaclusters obtained under negative mode ESI condition. The sample concentration is 250 μM in water. The data show that the metacluster formation is independent of ESI process and concentration.

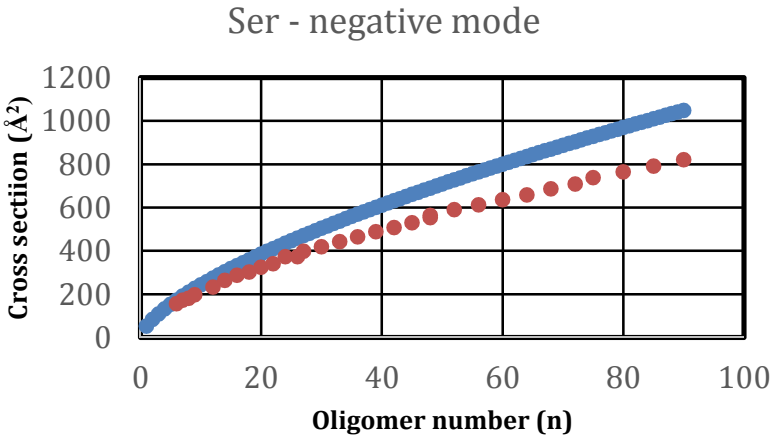

Figure S9. Growth trend of Ser metaclusters obtained from negative mode IMS experiments. The amino acid concentration is 250 μM in water. The isotropic curve is shown in blue and experimental cross sections are shown in red. The data are consistent with the cross section data obtained in positive mode.

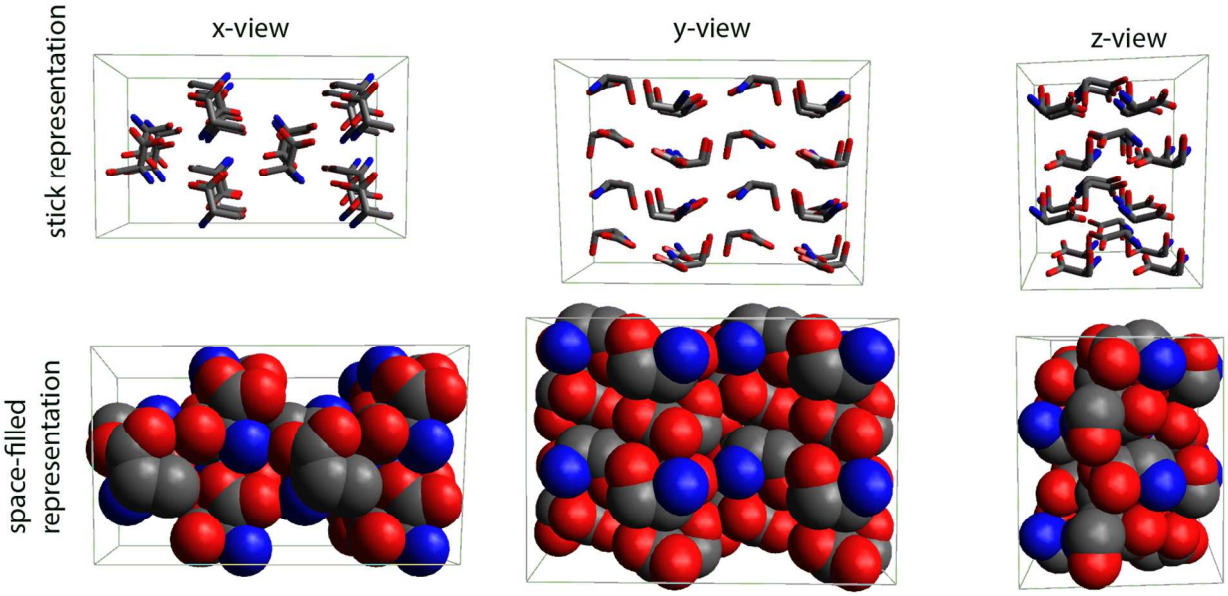

Figure S10. Representative structure of Serine 24-mer constructed from the crysal structure of L-serine. This model was constructed from the crystal structure obtained at 5.4 GPa by Moggach et al.

Table S6. Theoretical cross sections of model structures. The model structures are Serine 24-mer.

Structure	Cross section (σ, Å^2)			
	TJ[a]	PSA	Experiment	Isotropic
L-Serine I at 0.3 GPa	426	374		
L-Serine I at 1.4 GPa	419	387		
L-Serine I at 2.9 GPa	412	381	385[b]	435
L-Serine I at 4.1 GPa	408	374	373[c]	
L-Serine I at 4.8 GPa	405	367		
L-Serine II at 5.4 GPa	393	345		

[a] It has been shown that the TJ method can overapproximate the cross sections of small systems.[3,4]

[b] in positive mode ESI and at 6 mM.

[c] in negative mode ESI and at 250 μM.

S1.5 Transmission Electron Microscopy

TEM – Samples were fixed for 15 min in glutaraldehyde (1.6% final) and then adsorbed onto formvar/carbon-coated copper grids and subsequently stained with 2% uranyl acetate. Images were collected using a JEOL 123 microscope operating at 80 kV.

S2. Aggregation Propensity Prediction

S2.1 PASTA

PASTA score of each peptide was obtained from http://protein.bio.unipd.it/pasta2/.[5] Table S7 tabulates the PASTA 2.0 predictions for the four hexapeptides. PASTA 2.0 evaluates the aggregation propensity by calculating the stability of putative cross-β pairing between different sequence stretches. The method relies on a large dataset of globular proteins having known native structures.[5-7]

Table S7. PASTA 2.0 self-aggregation predictions for the four hexapeptides.

Peptide	Best energy	% disorder	% α-helix	% β-strand	% coil
NININI	-5.873696	100.0	0.0	0.0	100.0
NFNFNF	-2.623743	100.0	0.0	0.0	100.0
SISISI	-4.585828	100.0	0.0	0.0	100.0
SFSFSF	-2.349734	100.0	0.0	0.0	100.0

S2.2 3D Profile Method

The 3D Profile Method computes the composite score (C-Score) of a given peptide and uses it to predict aggregation propensity. C-Score takes into account the Rosetta energy of one layer composed of two β-strands in kcal/mol (i.e. the more negative score is preffered), the shape

complementary (from 0.0 to 1.0; the higher value indicates more contact between the two mating sheets), the area of interface, contact area and solvent-accessible surface area (SASA). Table S8 lists the statistics of the four hexapeptides calculated by the 3D Profile Method. It is noted here that the current 3D Profile method uses the steric zipper structure of NNQQNY as the only model, although 8 different classes of steric zippers are known to exist.[8,9]

Table S8. 3D Profile self-aggregation predictions for the four hexapeptides.

Peptide	Rosseta Energy	Shape Compl.	Area of Interface	Contact Area	SASA	C-Score
NININI	-27.6	0.8640	0.0	0.0	0.0	0.0
NFNFNF	-20.3	0.8589	48	0.0	196	-37.983
SISISI	-29.1	0.8141	70	0.0	200	-48.261
SFSFSF	-21.7	0.8141	70	0.0	193	-40.861

S2.4 Zyggregation Method

Zyggregator is an algorithm that predicts aggrgegation propensity based on optimized equations that are parameterized to take into account the overall hydrophobicity of the peptide, the overall charge, secondary structure propensity and hydrophobic/hydrophilic pattern. The score for each peptide was obtained from http://www-mvsoftware.ch.cam.ac.uk/index.php/zyggregator and at pH = 8.[10,11]

S2.5 Empirical Scoring Method based on Metacluster Data

As mentioned in the main text, the peptide aggregation propensity is defined by

$$SC_{agg} = -\left[\frac{1}{l}\sum_{i}^{all\ amino\ acids} n_i \times \left(\frac{\sigma_{1,Trp}}{\sigma_{1,i}}\right) \times (\Delta i - 1) \times 100\right] (Eq.4)$$

Isoleucine: ΔIle = 1.19 and $\left(\frac{\sigma_{1,Trp}}{\sigma_{1,Ile}}\right)$ = 1.23.

Asparagine: ΔAsn = 0.93 and $\left(\frac{\sigma_{1,Trp}}{\sigma_{1,Asn}}\right)$ = 1.36.

Serine: ΔSer = 0.86 and $\left(\frac{\sigma_{1,Trp}}{\sigma_{1,Ser}}\right)$ = 1.54.

Phenylalanine: ΔPhe = 1.05 and $\left(\frac{\sigma_{1,Trp}}{\sigma_{1,Phe}}\right)$ = 1.10.

Tryptophan: ΔTrp is set to be equal 1.

REFERENCES

(1) Mason, E. A. *Transport Properties of Ions in Gases*; 99 ed.; John Wiley & Sons, 1988.

(2) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 15132.

(3) Anderson, S. E.; Bleiholder, C.; Brocker, E. R.; Stang, P. J.; Bowers, M. T. *Int. J. Mass Spectrom.* **2012**, *330*, 78.

(4) Bleiholder, C.; Contreras, S.; Bowers, M. T. *Int. J. Mass Spectrom.* **2013**, *354*, 275.

(5) Walsh, I.; Seno, F.; Tosatto, S. C.; Trovato, A. *Nucleic Acids Res.* **2014**, *42*, W301.

(6) Trovato, A.; Chiti, F.; Maritan, A.; Seno, F. *Plos Computational Biology* **2006**, *2*, 1608.

(7) Trovato, A.; Seno, F.; Tosatto, S. C. E. *Protein Eng Des Sel.* **2007**, *20*, 521.

(8) Goldschmidt, L.; Teng, P. K.; Riek, R.; Eisenberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 3487.

(9) Thompson, M. J.; Sievers, S. A.; Karanicolas, J.; Ivanova, M. I.; Baker, D.; Eisenberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4074.

(10) Pawar, A. P.; Dubay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. *J. Mol. Biol.* **2005**, *350*, 379.

(11) Tartaglia, G. G.; Pawar, A. P.; Campioni, S.; Dobson, C. M.; Chiti, F.; Vendruscolo, M. *J. Mol. Biol.* **2008**, *380*, 425.