Supporting information for: Quantifying Co-Oligomer Formation by Alpha-Synuclein

Marija Iljina*, Alexander J. Dear*, Gonzalo A. Garcia, Suman De, Laura Tosatto, Patrick Flagmeier, Daniel R. Whiten, Thomas C. T. Michaels, Daan Frenkel, Christopher M. Dobson, Tuomas P. J. Knowles**, and David Klenerman**

*Equally contributing authors.

** Correspondence may be addressed to: tpjk2@cam.ac.uk, dk10012@cam.ac.uk*

Supporting Methods

TEM Imaging

We have previously shown that the fluorescently-labelled proteins used in this study, αS isoforms,[S1–S3] tau k18[S4] and Aβ40 or Aβ42[S5] can assemble into amyloid fibrils. Because in this study we aimed to characterise oligomer formation, we chose the conditions that would minimise fibril formation such as low total protein concentrations and absence of agitation. We carried out transmission electron microscopy (TEM) imaging to visualise protein solutions following the sample preparation and incubation protocols as described above, at the highest protein concentrations, 3 μM. 10-15 μL volumes of the samples were applied onto carbon-coated 400-mesh copper grids (EM Resolutions) for 5 min and washed with distilled

water. The application was repeated in the cases when the absorbed aggregate density was too low. Negative staining was carried out by using 2% (w/v) uranyl acetate. TEM images were acquired using Tecnai G2 microscope (13218, EDAX, AMETEK) operating at an excitation voltage of 200 kV. Representative images are shown in Supplementary Fig. 1, and amyloid fibril formation was observed only in the samples containing Aβ isoforms, and absent in all other protein solutions.

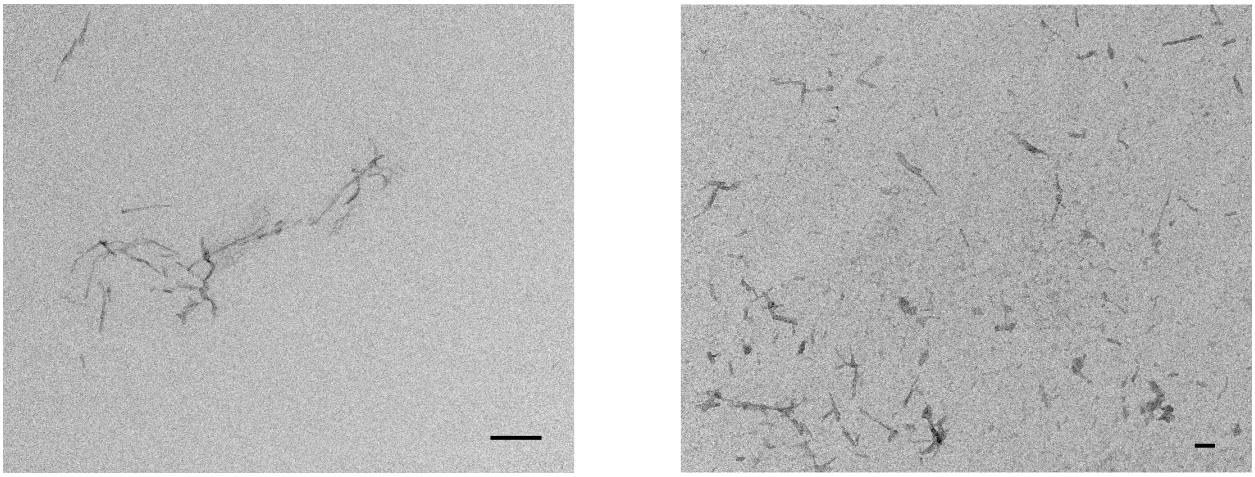


Figure S1: Representative TEM images of fluorescently labelled protein samples following 72-h incubation under quiescent conditions at 3 μM total protein concentration. Left: WT:Aβ40 1:1, right: WT:Aβ42 (scale bar 100 nm in both images).

TCCD controls using free dyes in solution

To confirm that the TCCD experiments were measuring the interactions between protein molecules rather than a random association of the fluorescent dyes, control measurements were performed using pairs of free fluorescent dyes in solution. The free unbound dyes were prepared by reacting 5 mM of AF488 NHS ester (Succinimidyl ester) or 5 mM of AF594 NHS ester with Tris.HCl (pH 7.2, 250 mM) at room temperature for 3 h. The solutions were further diluted with PBS buffer, defined previously, aliquoted and stored at $-80°C$. For the experiments, 1:1 molar ratio of AF488 and AF594 were combined up to the total concentration of 500 nM, recorded immediately after the preparation and then following

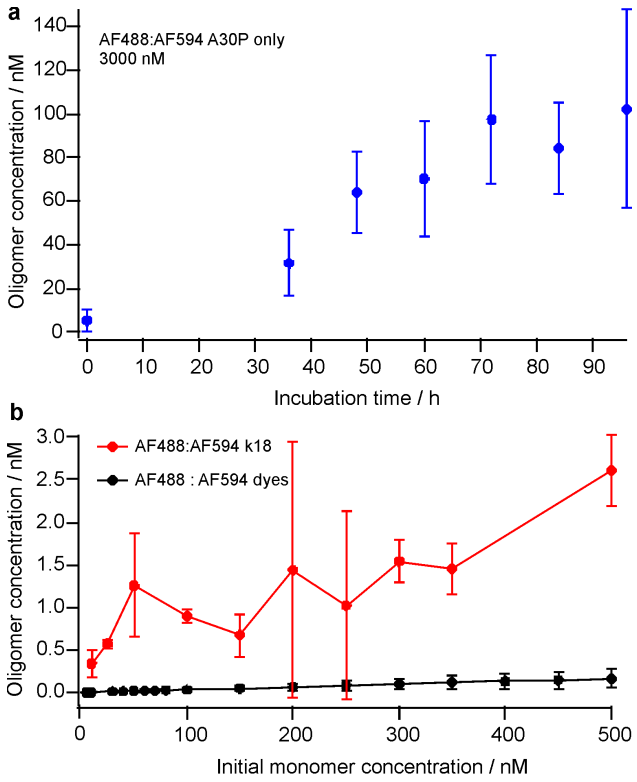


Figure S2: (a) Timecourse of oligomer formation by A30P point-mutant at the starting total protein concentration of 3 μM, recorded by sm-TCCD (n=3, SEM). (b) Result of control sm-TCCD measurement of free 1:1 AF488 and AF594 dye mixture, derived from a triplicated measurement of Q (eq. 1, main text) multiplied by a range of numbers between 0-500 nM, using the same analysis as for the derivation of apparent oligomer concentrations for the measurements of protein solutions in sm-TCCD experiments. The results obtained for k18 self-oligomerising system are shown in a range of nM concentrations for comparison (full range is in main text, Fig. 1), where samples were measured in triplicate for every starting protein concentration. Most of the data-points for the protein measurement are above the control values derived from the measurement of free dyes in solution, confirming that sm-TCCD experiments were quantifying the protein aggregates.

a 72-h incubation at the same incubation conditions as for the protein samples. For the detection using TCCD, the samples were diluted to 100 pM, and recorded under flow using the same procedure and settings as for the detection of protein samples, and the resulting Q (eq. 1, main text) was multiplied by a range of values corresponding to the starting total protein concentrations (i.e. following the same analysis as for the protein datasets). The result is presented in Supplementary Fig. 2 and compared to the data obtained for k18 protein solutions, which yielded the lowest apparent oligomer concentrations across all analysed self-oligomerising systems.

The control measurement of free dyes was repeated at the total concentration of 3 μM. Following the same analysis as for the protein samples, it yielded 0.64±0.12 nM (n=3, SD) from the free dyes. In comparison, the lowest oligomer concentration measured at the same starting monomer concentration (3 μM) is from tau k18 solutions and is 4.9±0.88 nM (n=3, SD), followed by 18.18±5.93 nM (n=3, SD) from WT:k18 samples and much higher concentrations from all other samples that were investigated over this monomer concentration range. Therefore, the results at higher concentration of the free dye solution are consistent with the results presented in Fig. S2(b).

Measurements of critical aggregation concentration of Aβ isoforms in the presence of WT αS

Having observed fibril formation in the 1:1 samples of αS and Aβ40 or Aβ42, we sought to determine whether these species were self-fibrils of Aβ, or mixed fibrils containing Aβ and αS. This question cannot be straightforwardly answered using high-resolution fluorescence imaging methods, by monitoring the degree of overlap of fluorescently-labelled αS and Aβ within fibrillar aggregates, because the overlap may be due to either mixed fibrillar structures, or due to intertwined self-fibrils of the two proteins, which is plausible due to the propensity of fibrils to clump. Therefore, we employed an alternative measurement, by determining whether the critical aggregation concentration (CAC) of Aβ is altered by the presence of αS. In our previous work, we used a single-molecule technique to measure the CAC of fluorescently labelled Aβ isoforms Aβ40 and Aβ42 in aqueous buffer, by measuring the concentration of monomeric peptide released from pre-formed fibrils into aqueous solution at equilibrium, and reported values of 94 ± 37 nM for Aβ40, and 28 ± 4 nM for Aβ42.[S5] The experimental protocols for these measurements are fully described in our previous study.[S5] Following the same protocols, we measured the CAC of fluorescently labelled Aβ isoforms in the presence of 500 nM of unlabelled WT αS. The obtained values in this experiment were 101 ± 27 nM and 32 ± 8 nM for Aβ40 and Aβ42, respectively, which was in close agreement with the

previously measured values in the absence of αS. This result suggested the absence of the effect of soluble αS on Aβ fibrillization, favouring the explanation that Aβ assembled into self-fibrils under our experimental conditions.

Table S1: ΔG° values of oligomerisation, derived from the fitting of sm-TCCD data as described in the main text. The summary of these values is in Fig. 3 (main text).

Protein combination	ΔG° **/kJmol^{-1}**	**Error in** ΔG°
A30P-A30P	-26.9	1.37
A53T-A53T	-25.2	0.44
E46K-E46K	-28.7	0.59
k18-k18	-19.4	0.49
WT-WT	-24.0	0.31
WT-A30P	-24.4	0.85
WT-A53T	-27.1	0.46
WT-E46K	-28.7	0.64
WT-k18	-22.7	0.75
WT-Aβ40	-29.4	0.62
WT-Aβ42	-30.8	1.03

Self-Oligomer Modelling

In our previous work,[S5] a streamlined statistical mechanical model was developed to describe equilibrium oligomer size distributions and to extract the free energy of oligomer growth by single-species monomer addition. Filamentous growth was assumed, such that oligomers are treated as one-dimensional chains. A single equilibrium constant K then describes oligomer formation and growth at temperature T in a reaction volume V. We summarize here the key equations describing this model (for more detail, see our previous paper[S5]).

Statistical Mechanical Linear Oligomer Model

The grand canonical partition function for a system containing linear oligomers up to size M, assuming no interaction between oligomers, is found to be

$$\Xi(T,V,\mu) = \exp\left(\sum_{j=1}^{M} q_j(T,V)\, e^{\beta j \mu}\right). \tag{1}$$

where $q_j(T,V)$ is the canonical partition function of a j-mer. The number of clusters of size j is then given by

$$N_j(T,V) = q_j(T,V)\, e^{\beta j \mu}. \tag{2}$$

The partition function $q_j(T,V)$, or $q(j)$, can be factorized into translational and internal components: $q(j) = q_{\text{trans}}(j)\, q_{\text{int}}(j)$. The translational partition function $q_{\text{trans}}(j)$ is proportional to the system volume, and so can be written as $q_{\text{trans}} = V/v_0(j)$, where $v_0(j)$ is a fundamental volume (in the gaseous phase, it is given by the cube of the thermal wavelength). The concentration of j-mers is given by $f(j) \equiv N_j(T,V)/(N_A\, V)$, where N_A is Avogadro's number:

$$f(j) = \frac{1}{N_A\, v_0(j)}\, q_{\text{int}}(j)\, e^{\beta j \mu}. \tag{3}$$

The chemical potential μ is set implicitly by conservation of the initial total monomeric protein concentration m_{tot}:

$$\sum_{j=1}^{M} j f(j) = -\frac{1}{N_A\, V}\frac{k_B\, T}{\Xi}\frac{\partial \Xi}{\partial \mu} = m_{\text{tot}}. \tag{4}$$

Oligomer partition function

We assume size-independent oligomer growth and shrinkage rates, such that $q_{\text{int}}(j) \approx e^{-\beta\, \varepsilon\, (j-1)}$ and $v_0(j) \equiv v_0$. The resulting expression for the oligomer size distribution is:

$$f(j) = \frac{1}{N_A\, v_0}\, e^{-\beta\, \varepsilon\, (j-1)}\, e^{\beta\, j\, \mu}. \tag{5}$$

The now size-independent standard free energy change ΔG° upon addition of a monomer to an oligomer is:

$$\Delta G^\circ = -RT \ln K = -RT \ln \frac{\frac{f(j+1)}{c_0}}{\frac{f(1)}{c_0} \cdot \frac{f(j)}{c_0}} = \varepsilon - R\,T \ln\,(N_A\,c_0\,v_0)\,. \tag{6}$$

This shows that ϵ is the free energy measured at standard concentration $c_0 = 1/(N_A v_0)$. Hence, we choose $v_0 = 1/(N_A\,c_0)$ for convenience, such that $\Delta G^\circ = \epsilon$, where standard conditions are chosen as $T = 25$ °C and $c_0 = 1$ M.

Correcting for experimental observations and fitting

In each experiment, half of the monomer species are labeled "red", and the other half "blue". Only oligomers which contain at least one red monomer and at least one blue monomer are accounted for in our apparent oligomer concentration measurements; all single-color species detected are branded free monomer molecules by the experimental procedure, and the two labels are assumed to have no differing effects besides introducing a degree of distinguishability between molecules.

Eq. (3) is thus modified to take into account the two colors:

$$f(j_1, j - j_1) = \frac{1}{N_A\,v_0}\,e^{-\beta\,\varepsilon(j-1)} e^{\beta\,j\,\mu} \binom{j}{j_1} \tag{7}$$

where $f(j_1, j_2)$ denotes the concentration of oligomers of size $j = j_1 + j_2$ monomers, of which j_1 are red and j_2 are blue. The apparent size distribution $f_{\exp}(j)$ is then:

$$f_{\exp}(j) = \begin{cases} \sum_{j_1=1}^{M} f(j_1, 0) + \sum_{j_2=1}^{M} f(0, j_2) & j = 1 \\ \sum_{j_1=1}^{j-1} f(j_1, j - j_1) & j > 1 \end{cases}. \tag{8}$$

Combining Eqs. (7) and (8), the total apparent oligomer concentration F_{exp} is given by:

$$F_{\mathrm{exp}} = \sum_{j=2}^{M} \frac{2^{j-1}}{N_A\, v_0}\, e^{-\beta\, \varepsilon (j-1)} e^{\beta\, j\, \mu} \tag{9}$$

The conservation-of-mass condition is also modified, and becomes:

$$\sum_{j=1}^{M} \sum_{j_1=0}^{j} j_1 \cdot f(j_1, j-j_1) = -\frac{1}{2} \frac{1}{N_A\, V} \frac{k_B\, T}{\Xi} \frac{\partial \Xi}{\partial \mu} = \frac{1}{2} m_{\mathrm{tot}}. \tag{10}$$

Eq. (9), with M chosen to be large enough to encompass virtually all oligomers, can be fitted directly via least-squares methods to experimental measurements of equilibrium apparent oligomer concentrations, with ε as the sole fitting parameter. At each step, μ is evaluated numerically from the implicit analytical conservation-of-mass condition Eq. (10).

Oligomer sizes

Having obtained fitted values for ε, it is possible to simulate oligomer size distributions formed by different αS mutants at different initial monomer concentrations. It may be shown that, at the concentrations visited by our study, oligomers formed by the mutants under investigation are predominantly dimeric. For example, at 1 μM, 99% of oligomers formed by WT αS are dimeric. The steep decline in the size distributions implies that possible deviations from linear geometry at larger sizes will not significantly perturb our analysis, justifying our choice of a linear model.

Co-oligomer Modelling

The approach detailed above was generalized in our previous work[S5] to consider the co-oligomerization of two monomeric species to form dimers. Here, we go further and develop a model for co-oligomerization of two species to form linear oligomers of any length.

Exact solution

The grand canonical partition function for a co-oligomerizing ensemble of 2 different monomer types is given as a direct generalization of Eq. (1):

$$\Xi(T,V,\vec{\mu}) = \exp\left(\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} q_{\vec{j}}(T,V)\, e^{\beta\,\vec{j}\cdot\vec{\mu}}\right), \tag{11}$$

where $\vec{j} = (j_1, j_2)$ is a composition vector whose components j_i describe the number of monomeric residues of type i. Similarly, $\vec{\mu} = (\mu_1, \mu_2)$ with μ_i being the chemical potential associated with species i, $q_{\vec{j}}$ gives the canonical partition function for an oligomer with composition described by $\vec{j}$. If nearest-neighbor interactions dominate, $q_{\vec{j}}(T,V)$ is given by:

$$q_{\vec{j}}(T,V) \approx \frac{V}{v_0}\sum_{o(\vec{j})} e^{-\beta\,U(\vec{j},o(\vec{j}))}, \tag{12}$$

where $U(\vec{j},o(\vec{j}))$ denotes the internal energy of an oligomer with composition $\vec{j}$ and with monomer units arranged in an order specified by $o(\vec{j})$; a sum must be carried out over all possible arrangements. The concentration of oligomers with composition $\vec{j}$ is then given by:

$$f(\vec{j}) = \frac{1}{N_A v_0}\sum_{o(\vec{j})} e^{-\beta\,U(\vec{j},o(\vec{j}))}\, e^{\beta\,\vec{j}\cdot\vec{\mu}}. \tag{13}$$

The chemical potential μ_i of monomer type i is defined for corresponding initial type i total monomeric protein concentration m^i_{tot} by

$$\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} j_i\, f(\vec{j}) = -\frac{1}{N_A\,V}\frac{k_B\,T}{\Xi}\frac{\partial\,\Xi}{\partial\,\mu_i} = m^i_{\mathrm{tot}}. \tag{14}$$

The sum over arrangements $o(\vec{j})$ can be carried out by noting that if the number of cross-bonds x present in a J-mer is known, then $U(\vec{j},o(\vec{j}))$ is almost uniquely defined. There are three cases to consider: x is odd (contributions to sum must be multiplied by a factor of 2 as chains have directionality); x is even and the chain terminates with type 1 monomers on

both ends; x is even and the chain terminates with type 2 monomers on both ends. These three cases all give different energies, and so the Boltzmann factors for each of these energies must be multiplied by the relevant degeneracy and added together to give the contribution to $q(\vec{j})$ for each value of x. A sum can then be carried out over all possible values of x to arrive at a final expression, setting $x = 2r$ or $x = 2r - 1$ in the relevant sums to ensure only odd or even values are included as required. We ultimately arrive at the expression:

$$\frac{q(\vec{j})}{V/v_0} = \begin{cases} e^{-\beta\varepsilon_{2,2}(J-1)}, & j_1 = 0 \\ e^{-\beta\varepsilon_{1,1}(J-1)}, & j_2 = 0 \\ 2 \displaystyle\sum_{r=1}^{\mathrm{Min}(j_1,j_2)} \binom{j_1-1}{r-1}\binom{j_2-1}{r-1} e^{-\beta E(j_1,j_2,r,1)} & \\ + \mathcal{H}[j_1-2] \displaystyle\sum_{r=1}^{\mathrm{AltMin}(j_1,j_2)} \binom{j_1-1}{r}\binom{j_2-1}{r-1} e^{-\beta E(j_1,j_2,r,2)} & \\ + \mathcal{H}[j_2-2] \displaystyle\sum_{r=1}^{\mathrm{AltMin}(j_1,j_2)} \binom{j_2-1}{r}\binom{j_1-1}{r-1} e^{-\beta E(j_1,j_2,r,3)}, & j_1 > 0, j_2 > 0, \end{cases} \quad , \tag{15}$$

where $J = j_1 + j_2$, and the smaller of the two numbers j_1 and j_2 is given by $\mathrm{Min}(j_1, j_2)$. The discrete Heaviside step function is denoted $\mathcal{H}[z]$, defined such that $H[0] = 1$, and the remaining functions are defined as follows:

$$\mathrm{AltMin}(j_1, j_2) = \begin{cases} j_1 - 1 & j_1 = j_2 \\ \mathrm{Min}(j_1, j_2) & j_1 \neq j_2 \end{cases}, \tag{16}$$

$$E(j_1, j_2, r, 1) = (j_1 - r)\,\varepsilon_{1,1} + (j_2 - r)\,\varepsilon_{2,2} + (2r - 1)\,\varepsilon_{1,2}, \tag{17}$$

$$E(j_1, j_2, r, 2) = (j_1 - r - 1)\,\varepsilon_{1,1} + (j_2 - r)\,\varepsilon_{2,2} + 2r\,\varepsilon_{1,2}, \tag{18}$$

$$E(j_1, j_2, r, 3) = (j_1 - r)\,\varepsilon_{1,1} + (j_2 - r - 1)\,\varepsilon_{2,2} + 2r\,\varepsilon_{1,2}. \tag{19}$$

Correcting for experimental observations and fitting

Where one monomer type is labelled with AF488 dye and one with AF594 dye, then only mixed oligomers will be recorded. The total mixed oligomer concentration $F(\vec{m_{\rm tot}})$ is given from combining Eqs. (13) and (15):

$$\begin{aligned} F_{\rm exp}(\vec{m_{\rm tot}}) = \sum_{j=2}^{M}\sum_{j_1=1}^{j-1} \quad & 2 \sum_{r=1}^{\mathrm{Min}(j_1,j-j_1)} \binom{j_1-1}{r-1}\binom{j-j_1-1}{r-1} e^{-\beta E(j_1,j-j_1,r,1)} \\ & + \mathcal{H}[j_1-2] \sum_{r=1}^{\mathrm{AltMin}(j_1,j-j_1)} \binom{j_1-1}{r}\binom{j-j_1-1}{r-1} e^{-\beta E(j_1,j-j_1,r,2)} \\ & + \mathcal{H}[j-j_1-2] \sum_{r=1}^{\mathrm{AltMin}(j_1,j-j_1)} \binom{j-j_1-1}{r}\binom{j_1-1}{r-1} e^{-\beta E(j_1,j-j_1,r,3)} \end{aligned} \quad (20)$$

This expression can then be fitted directly to experimental measurements.

The fitting procedure now involves evaluating both chemical potentials at each stage numerically from implicit algebraic expressions, obtained via Eq. (14) and Eq. (15). The sole fitting parameter is $\varepsilon_{1,2}$, once $\varepsilon_{1,1}$ and $\varepsilon_{2,2}$ are established from fits to self-oligomerization data sets. The same standard value for v_0 is used for all fits.

Justifying the choice of model

Our statistical mechanical modelling approach relies on a set of basic assumptions that are easily justified considering the experimental conditions and previous work. It has been established (see fig. S2) that chemical equilibrium is attained during the long equilibration times (72 hours) of our experiments (see also ref. S2). Furthermore, fibrils are known not to form at the low protein concentrations investigated here (see ref. S3). An equilibrium statistical mechanical model, considering only monomers and soluble oligomers, is thus clearly appropriate. At the low concentrations used (nM protein and pM oligomers), activities can be well-approximated by concentrations, and we can define equilibrium constants in the usual way. Furthermore, since oligomers are at such low concentrations, any possible bimolecu-

lar reactions involving oligomers are effectively suppressed, and we need only to consider monomer addition and subtraction reaction steps in a statistical mechanical model. For the same reason the equilibrium monomer concentration is well-approximated by the initial monomer concentration.

These assumptions are not enough to arrive at a unique model; we have further assumed that the populations of oligomers formed can be well-represented by a model that considers linear oligomers of any length. This model was chosen firstly because a wide range of amyloidogenic proteins form oligomers regardless of their primary sequence, implying that oligomers are held together by nonspecific interactions and can exist in a range of sizes. Secondly, we have little knowledge of the geometry of amyloid oligomers, and a linear geometry gives rise to the simplest model possible that allows a range of sizes.

Since there is no direct evidence that oligomers of a range of sizes form in these experiments, we now test the possibility that instead a specific kind of self-oligomer forms, with a specific size. These can be easily modeled in the same grand canonical framework as the existing model, by simply allowing only oligomers of size n to form. We test this model for $n = 2,\ 3$ and 4. We find that $n = 3$ and $n = 4$ give substantially worse fits, both in terms of sum-of-square-residuals and visually (see Figure S3 and Table S2). There is no need to consider higher values of n; the oligomer concentration scales as the n^{th} power of the monomer concentration and the curvature of the data are clearly insufficient to support any higher values. As expected, since we have already calculated the best-fit linear size distribution to be dimer-dominated at the concentrations investigated, we find that $n = 2$ gives fits similar in quality to the linear model.

We conclude that there is no especial preference for formation of a specific oligomer species of size $n > 2$. Although it is possible that the oligomers are formed via specific interactions that lead to dimers only, the marginal improvement in fit quality is insufficient to justify making this additional assumption of specificity. In any case, changing to this model would have no effect on the key results of our study, since bonding free energies

obtained under the dimer model are very close to those obtained from the linear model, as expected (see Table S3).

Additionally, it is now apparent that the curvature of the data in all experiments precludes the presence of significant concentrations of oligomers of size $n > 2$. The best fit to the data that is theoretically possible to obtain will thus be given by any model that permits the size distribution to be dimer-dominated. All such models will fit the data equally well and will give similar fitted bonding free energy values. There is therefore no reason to consider a model more complicated than the 1-parameter linear model. We do *not* conclude that alpha-synuclein oligomers are always linear in morphology, merely that there is no evidence of non-linear oligomers in the experiments performed in this study, and that no other physically reasonable model can therefore give a better description of the system under the present experimental conditions.

Finally, a visual inspection of the co-oligomer data reveals that again the curvature is insufficient to justify significant concentrations of $n > 2$ oligomers, and the co-oligomer modelling approach can therefore be justified on the same grounds as the self-oligomer modelling approach.

Table S2: Residual sum-of-squares values for self-oligomer models (units: M)

Oligomer type	WT-WT	A30P-A30P	A53T-A53T	E46K-E46K	K18-K18
Linear model	$2.1 \cdot 10^{-16}$	$4.4 \cdot 10^{-15}$	$5.6 \cdot 10^{-16}$	$6.2 \cdot 10^{-15}$	$4.1 \cdot 10^{-17}$
Trimer model	$4.8 \cdot 10^{-16}$	$5.1 \cdot 10^{-15}$	$1.0 \cdot 10^{-15}$	$9.5 \cdot 10^{-15}$	$6.0 \cdot 10^{-17}$
Tetramer model	$7.2 \cdot 10^{-16}$	$5.9 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$	$1.2 \cdot 10^{-14}$	$7.5 \cdot 10^{-17}$
Dimer model	$2.0 \cdot 10^{-16}$	$4.3 \cdot 10^{-15}$	$5.4 \cdot 10^{-16}$	$5.5 \cdot 10^{-15}$	$4.0 \cdot 10^{-17}$

Table S3: Bond energy values for self-oligomer models (units: kJ mol^{-1})

Oligomer type	WT-WT	A30P-A30P	A53T-A53T	E46K-E46K	K18-K18
Linear model	−24.0	−26.9	−25.2	−28.7	−19.4
Dimer model	−24.1	−27.5	−25.4	−29.3	−19.4

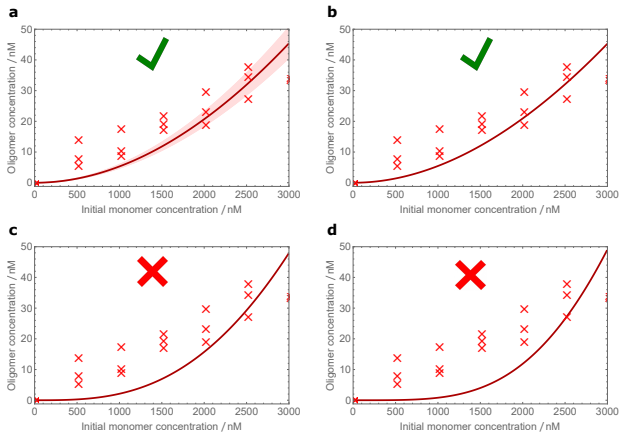


Figure S3: Fits of alternative models to WT-WT self-oligomer data. **a**: Fit to a linear model (the same as in Fig. 1e, main text). **b**: A dimer-only model gives an equally good fit. **c**: A trimer-only model gives a worse fit, having curvature significantly greater than that of the data. **d**: A tetramer-only model has even greater curvature and thus gives an even worse fit. Since the minimum possible curvature for a physically-reasonable model is given by the dimer-only model, and it is the low data curvature that is the primary source of fitting error, the linear and the dimer model give the best fit to the data that is theoretically possible.

References

(S1) Cremades, N.; Cohen, S. I. A.; Deas, E.; Abramov, A. Y.; Chen, A. Y.; Orte, A.; Sandal, M.; Clarke, R. W.; Dunne, P.; Aprile, F. A.; Bertoncini, C. W.; Wood, N. W.; Knowles, T. P. J.; Dobson, C. M.; Klenerman, D. Direct Observation of the Interconversion of Normal and Toxic Forms of α-Synuclein. *Cell* **2012**, *149*, 1048–1059.

(S2) Tosatto, L.; Horrocks, M. H.; Dear, A. J.; Knowles, T. P. J.; Dalla Serra, M.; Cremades, N.; Dobson, C. M.; Klenerman, D. Single-Molecule FRET Studies on Alpha-Synuclein Oligomerization of Parkinson's Disease Genetically Related Mutants. *Sci. Rep.* **2015**, *5*, 16696.

(S3) Iljina, M.; Garcia, G. A.; Horrocks, M. H.; Tosatto, L.; Choi, M. L.; Ganzinger, K. A.; Abramov, A. Y.; Gandhi, S.; Wood, N. W.; Cremades, N.; Dobson, C. M.; Knowles, T. P. J.; Klenerman, D. Kinetic Model of the Aggregation of Alpha-Synuclein Provides Insights Into Prion-Like Spreading. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E1206–

E1215.

(S4) Shammas, S. L.; Garcia, G. A.; Kumar, S.; Kjaergaard, M.; Horrocks, M. H.; Shivji, N.; Mandelkow, E.; Knowles, T. P. J.; Mandelkow, E.; Klenerman, D. A Mechanistic Model of Tau Amyloid Aggregation Based on Direct Observation of Oligomers. *Nat. Commun.* **2015**, *6*, 7025.

(S5) Iljina, M.; Garcia, G. A.; Dear, A. J.; Flint, J.; Narayan, P.; Michaels, T. C. T.; Dobson, C. M.; Frenkel, D.; Knowles, T. P. J.; Klenerman, D. Quantitative Analysis of Co-Oligomer Formation by Amyloid-Beta Peptide Isoforms. *Sci. Rep.* **2016**, *6*, 28658.

APPENDIX

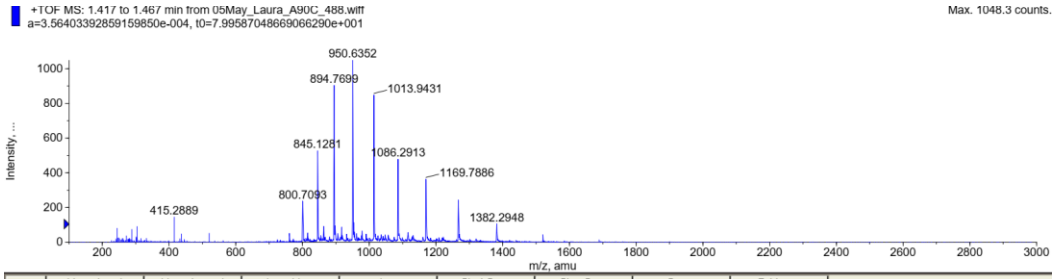
Mass spectrometry results to control α-synuclein labeling efficiency

Table A

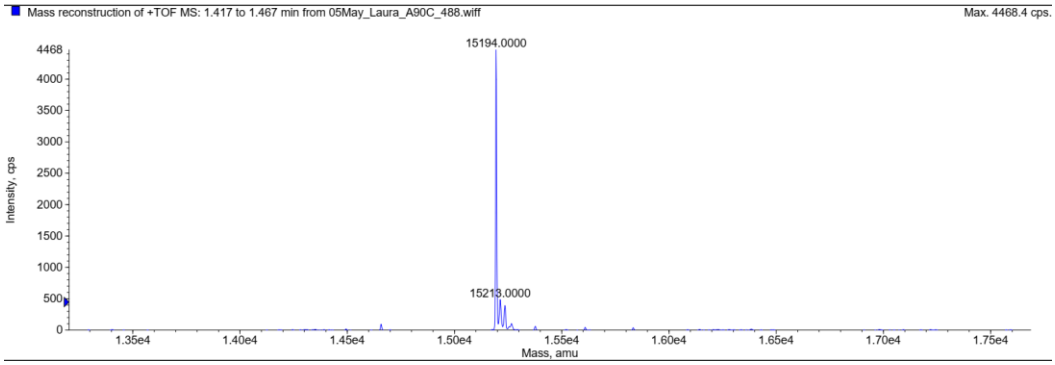
Protein mass values (theoretical)			
αSyn variant		*+Alexa488*	*+Alexa594*
	Unlabeled protein mass	+700 Da	+887 Da
wt	14460.1		
A90C	14492.2	15192.2	15379.2
A30P	14486.1		
A30P/A90C	14518.2	15218.2	15405.2
A53T	14490.1		
A53T/A90C	14522.2	15222.2	15409.2
E46K	14459.2		
E46K/A90C	14491.2	15191.2	15378.2

Variations from mass spec values of <5 Da are due to mass calibration methodologies (optimized for small molecules)

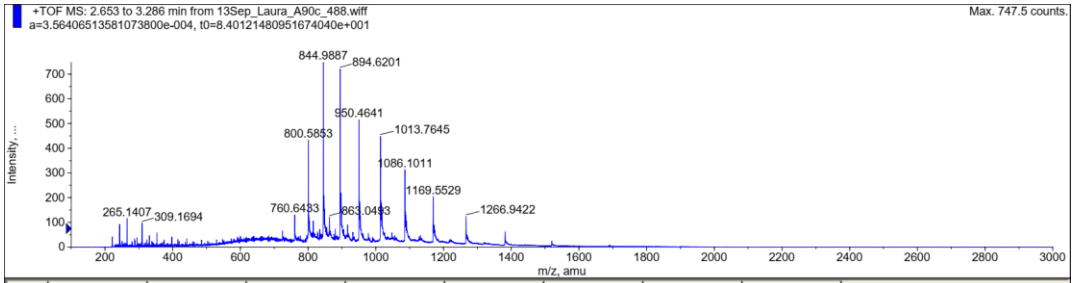
A90C + AlexaFluor488



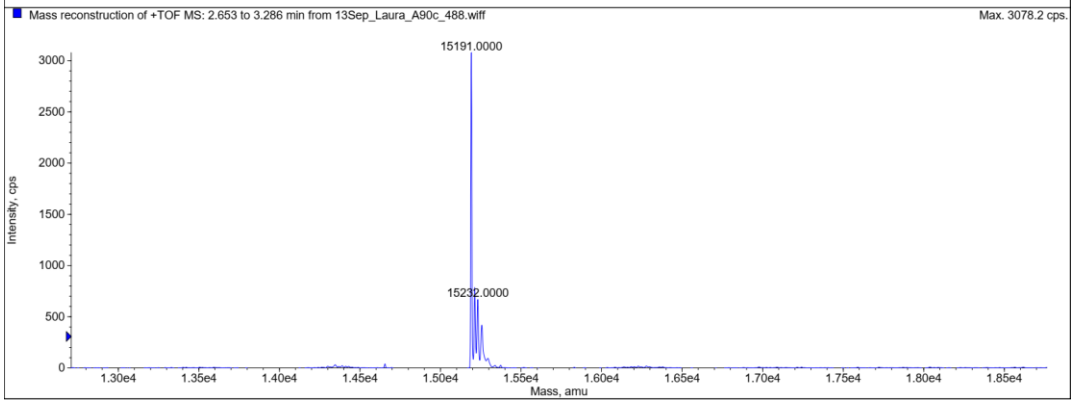
	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15193.9680		15193.8626	23738.4419	85	88	1.0000	C
2	15212.9850		15212.4163	4707.9401	85	88	1.0000	C
3	15235.7390		15235.9505	2874.6825	85	88	1.0000	C
4	15198.9003		15198.8870	2324.6210	85	88	1.0000	C
5	15265.9439		15265.9972	1145.8105	85	88	1.0000	C
6	15228.0814		15228.0745	1063.2306	85	88	1.0000	C



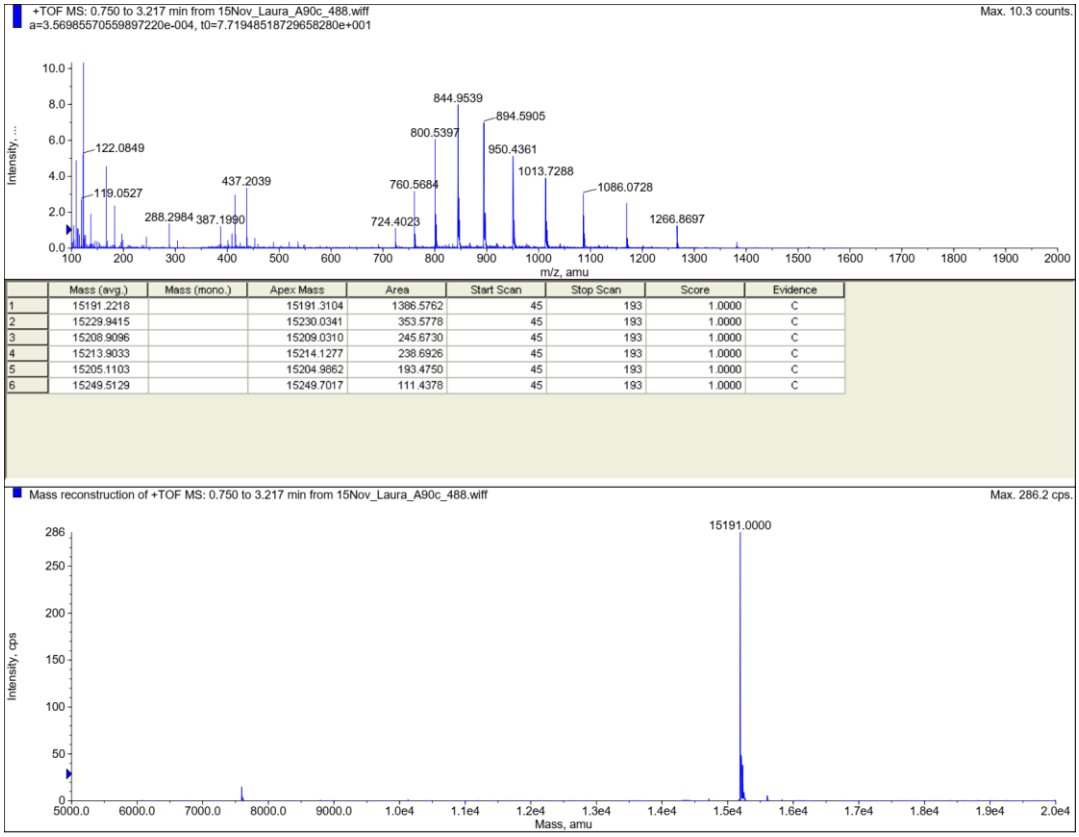
Undetected unlabeled protein



	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15191.4622		15191.5060	20077.1112	159	197	1.0000	C
2	15231.6610		15231.5178	7393.1476	159	197	1.0000	C
3	15212.7157		15212.7840	7303.7548	159	197	1.0000	C
4	15256.5891		15257.4111	5735.3003	159	197	1.0000	C

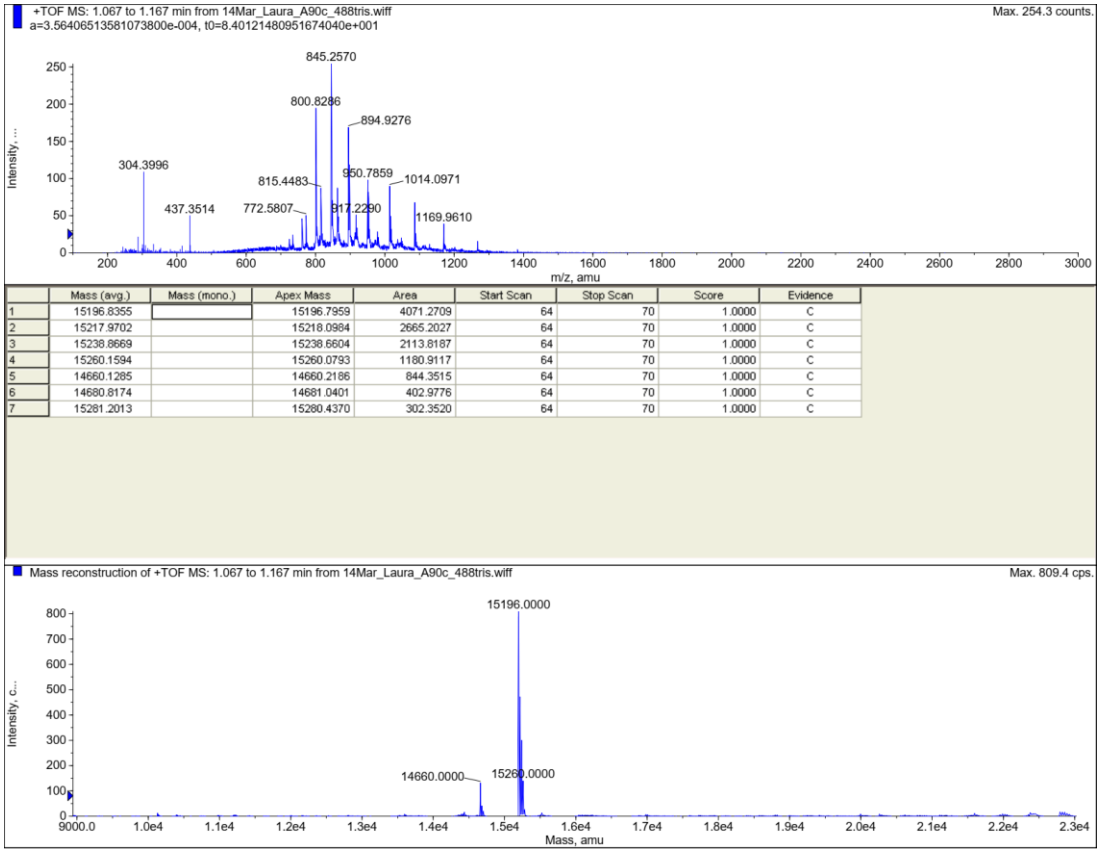


Undetected unlabeled protein



	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15191.2218		15191.3104	1386.5762	45	193	1.0000	C
2	15229.9415		15230.0341	353.5778	45	193	1.0000	C
3	15208.9096		15209.0310	245.6730	45	193	1.0000	C
4	15213.9033		15214.1277	238.6926	45	193	1.0000	C
5	15205.1103		15204.9862	193.4750	45	193	1.0000	C
6	15249.5129		15249.7017	111.4378	45	193	1.0000	C

Undetected unlabeled protein

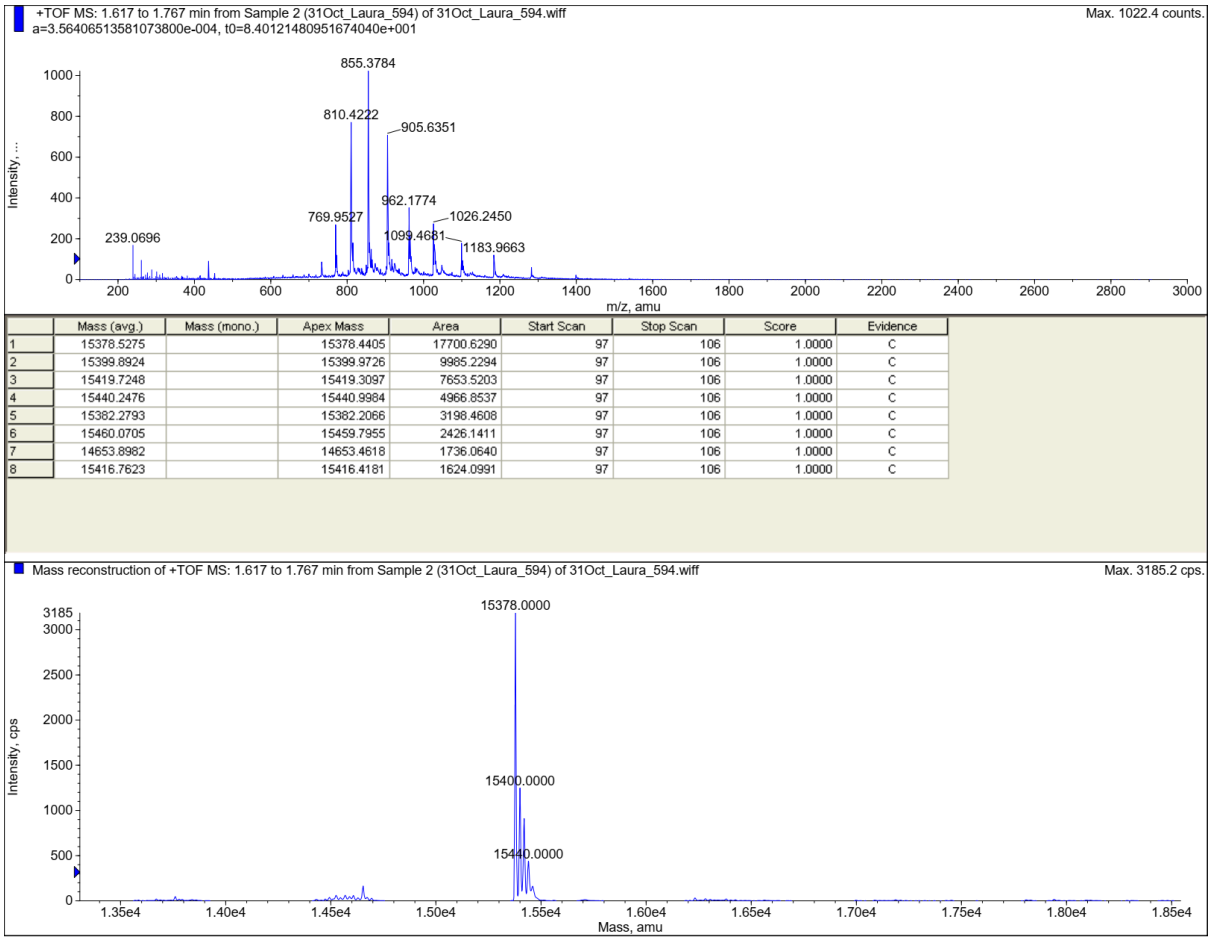


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15196.8355		15196.7959	4071.2709	64	70	1.0000	C
2	15217.9702		15218.0984	2665.2027	64	70	1.0000	C
3	15238.8669		15238.8604	2113.8187	64	70	1.0000	C
4	15260.1594		15260.0793	1180.9117	64	70	1.0000	C
5	14660.1285		14660.2186	844.3515	64	70	1.0000	C
6	14680.8174		14681.0401	402.9776	64	70	1.0000	C
7	15281.2013		15280.4370	302.3520	64	70	1.0000	C

Detectable unlabeled protein - Sample discarded (see notes below).

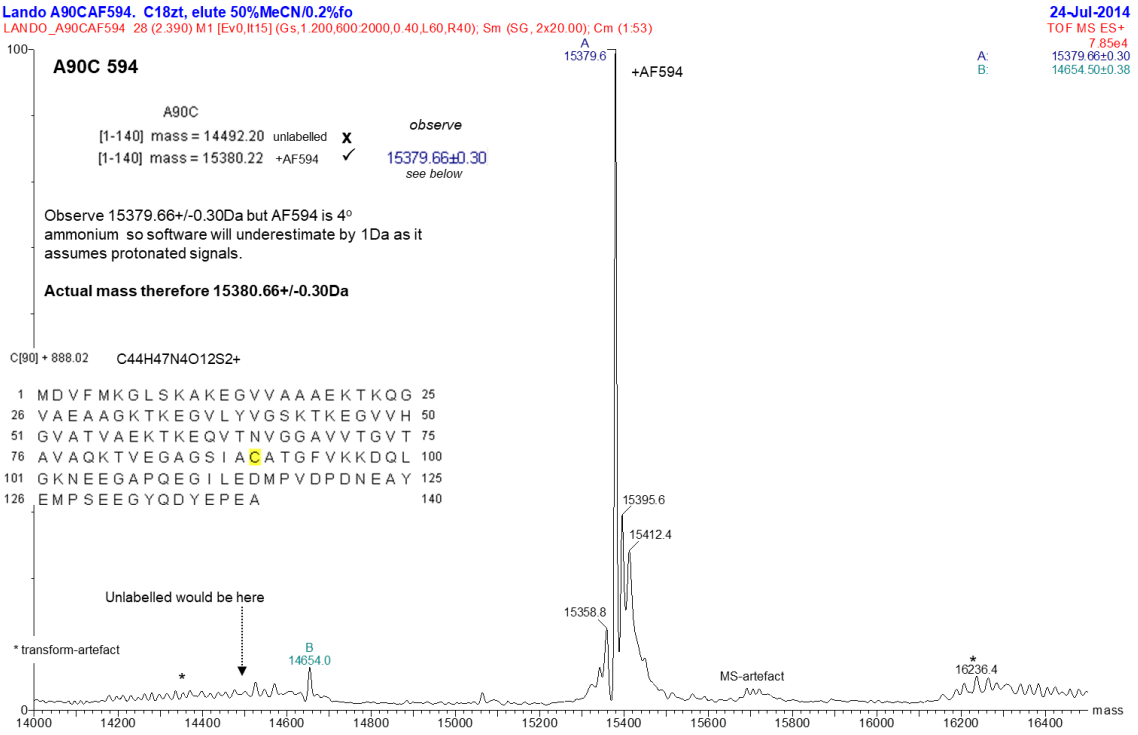
Notes on the previous spectrum (bottom of p. 3): In the presence of unlabeled protein/protein batches containing protein with adducts (mass 14660, for example), the areas of the masses corresponding to labeling protein /other species have been integrated. In the case of adduct corresponding to more than the 10% compared to labeled protein (mass 15196), the sample has been discarded.
Here we report this example: species corresponding to A90C-488 labeled protein and various salt complexes are mass the ones above 15 kDa (lines 1, 2, 3, 4, 7), while unlabeled protein masses are reported in lines 5 and 6. In this case, this batch has been discarded as the unlabeled protein adduct is around 12%.

A90C + AlexaFluor594

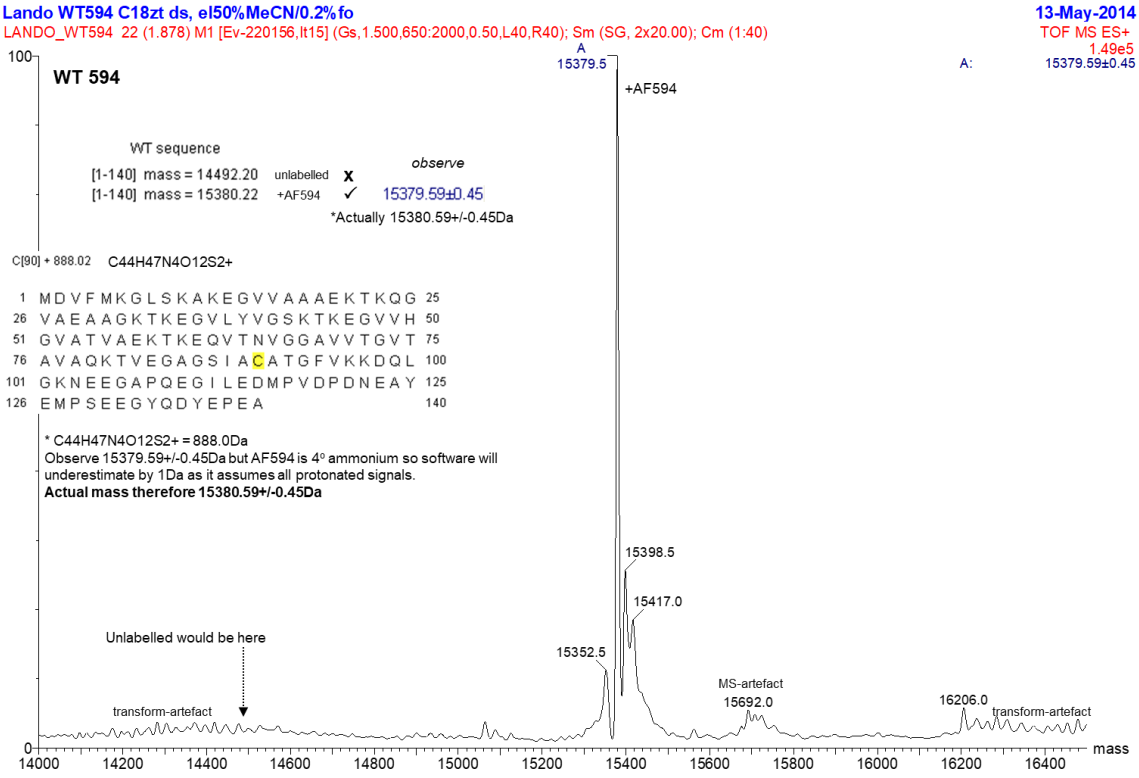


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15378.5275		15378.4405	17700.6290	97	106	1.0000	C
2	15399.8924		15399.9726	9985.2294	97	106	1.0000	C
3	15419.7248		15419.3097	7653.5203	97	106	1.0000	C
4	15440.2476		15440.9984	4966.8537	97	106	1.0000	C
5	15382.2793		15382.2066	3198.4608	97	106	1.0000	C
6	15460.0705		15459.7955	2426.1411	97	106	1.0000	C
7	14653.8982		14653.4618	1736.0640	97	106	1.0000	C
8	15416.7623		15416.4181	1624.0991	97	106	1.0000	C

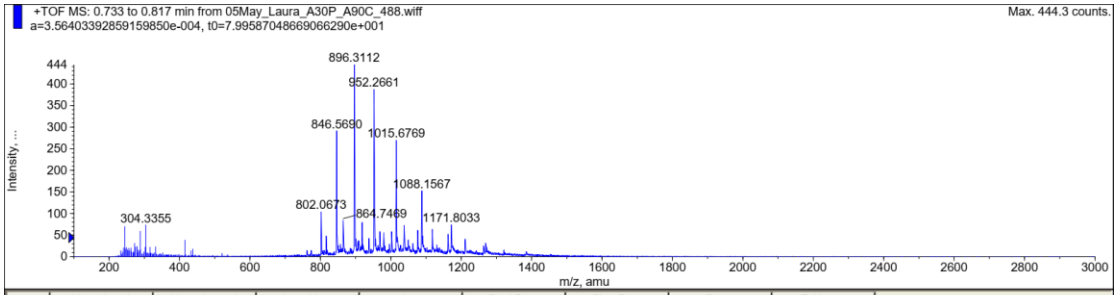
ESI analysis cap 3000V, cone 70V, calib CytC same file, 70%MeOH/0.2%formic mobile phase



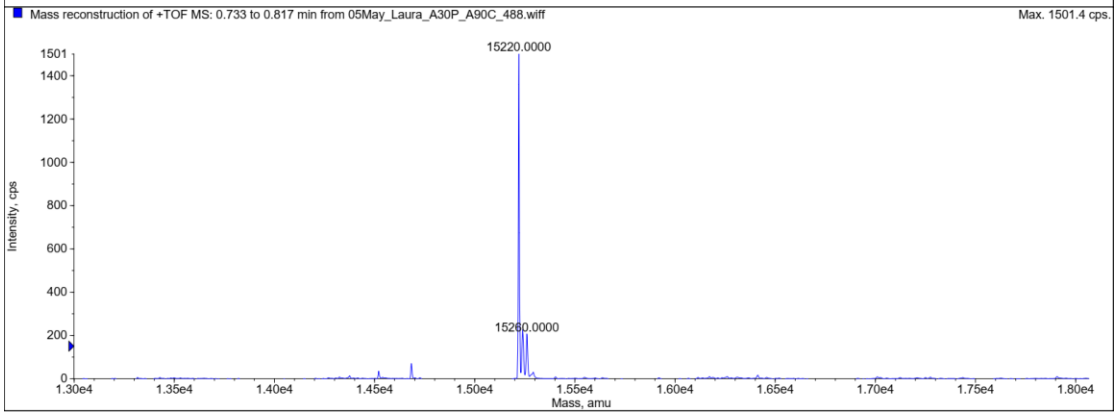
ESI analysis cap 3000V, cone 70V, calib ubiquitin same file, 70%MeOH/0.2%formic mobile phase

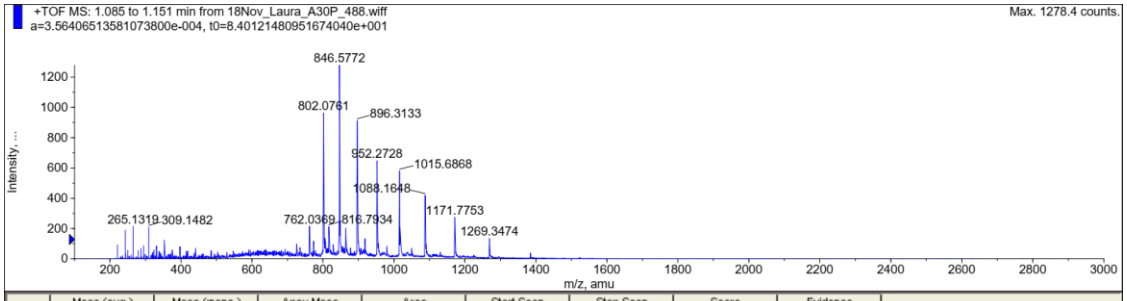


A30P A90C + AlexaFluor488

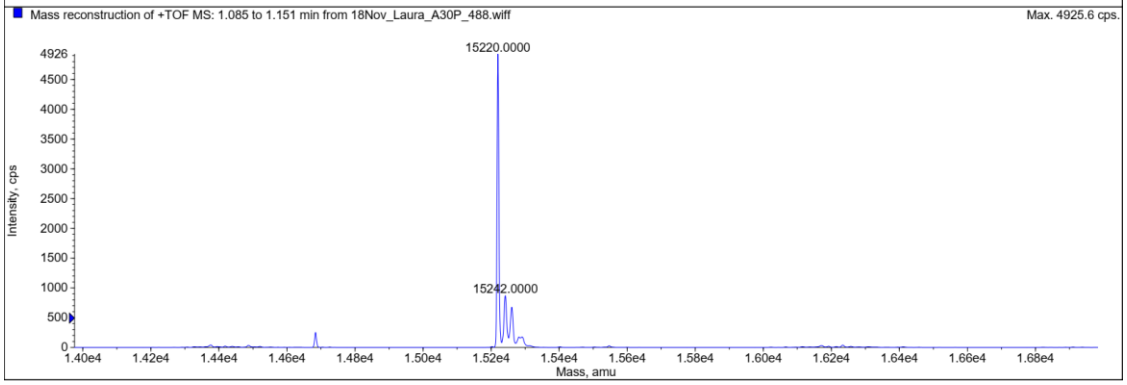


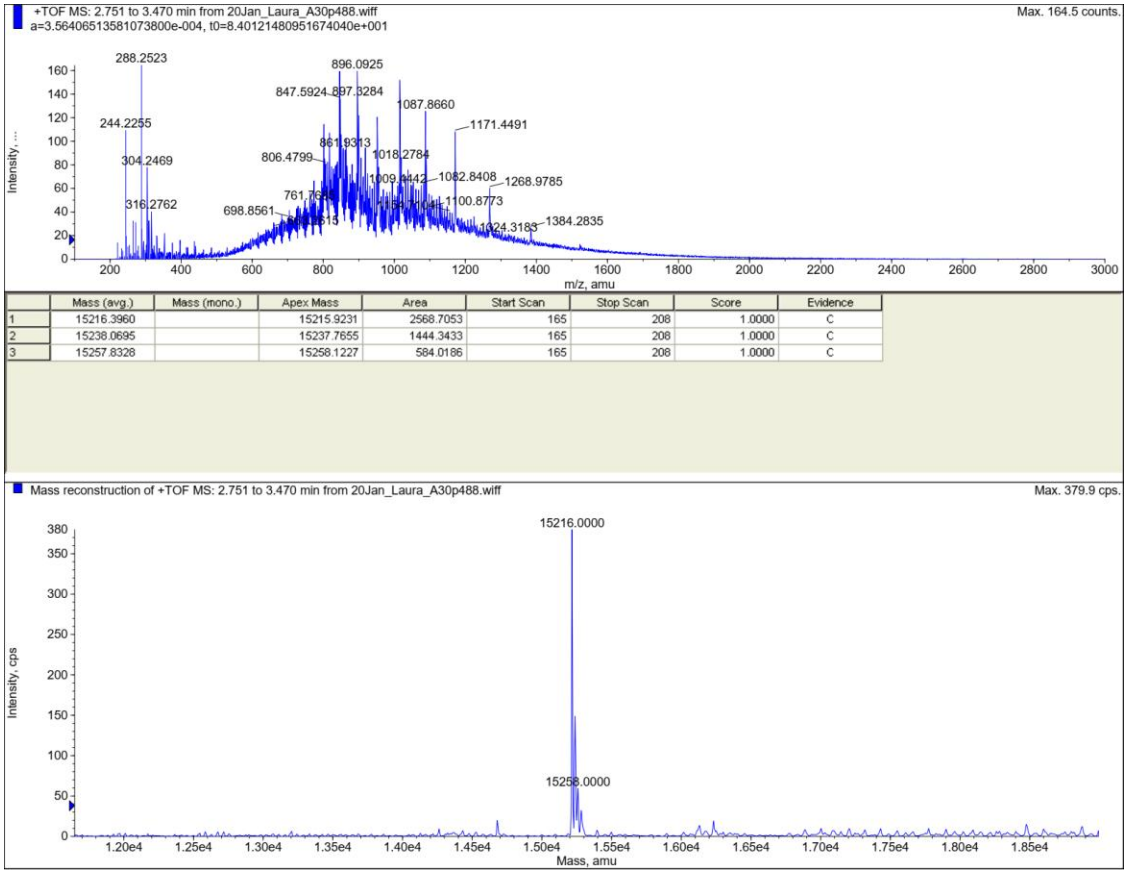
	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15220.1865		15220.1904	8090.8354	44	49	1.0000	C
2	15239.0661		15238.6380	2047.1615	44	49	1.0000	C
3	15260.4048		15260.6214	1716.6494	44	49	1.0000	C



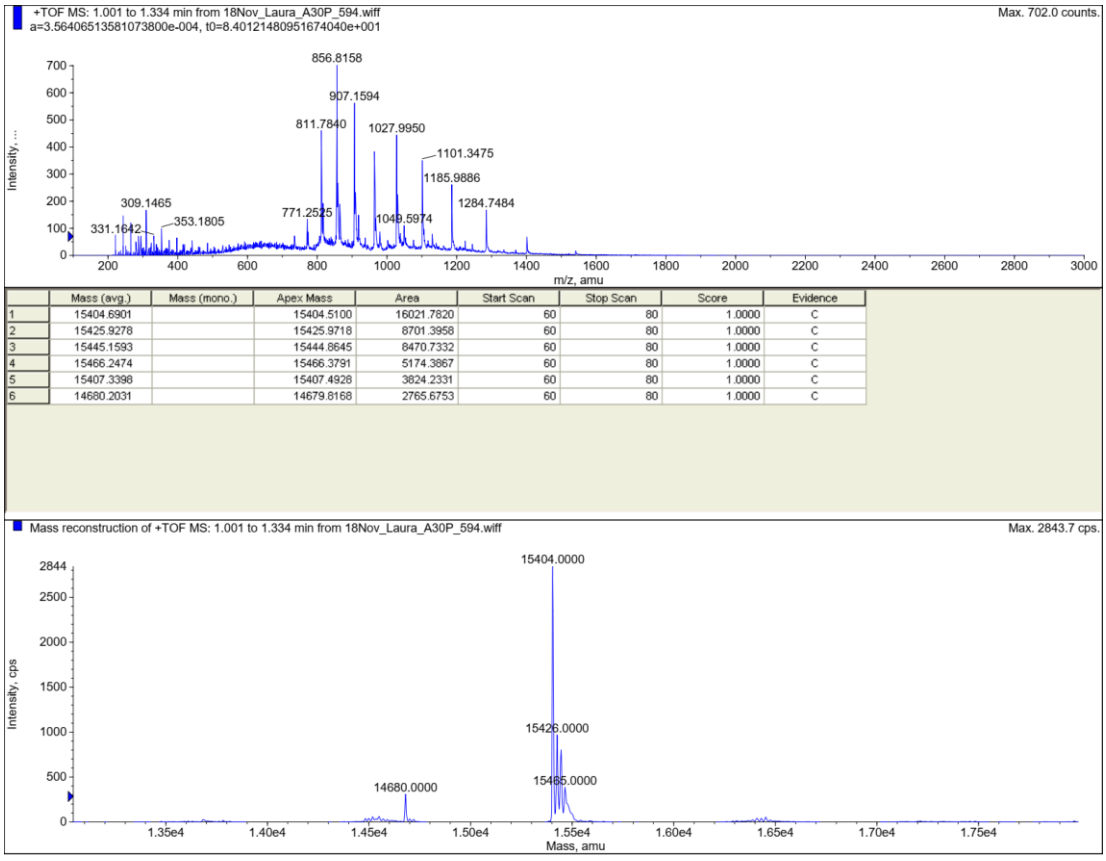


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15220.4483		15220.3473	29077.9182	65	69	1.0000	C
2	15241.5234		15241.2257	7071.7664	65	69	1.0000	C
3	15260.8614		15260.8890	5637.1457	65	69	1.0000	C
4	15280.7550		15279.9514	2501.9496	65	69	1.0000	C
5	15237.6030		15237.2267	1938.8473	65	69	1.0000	C



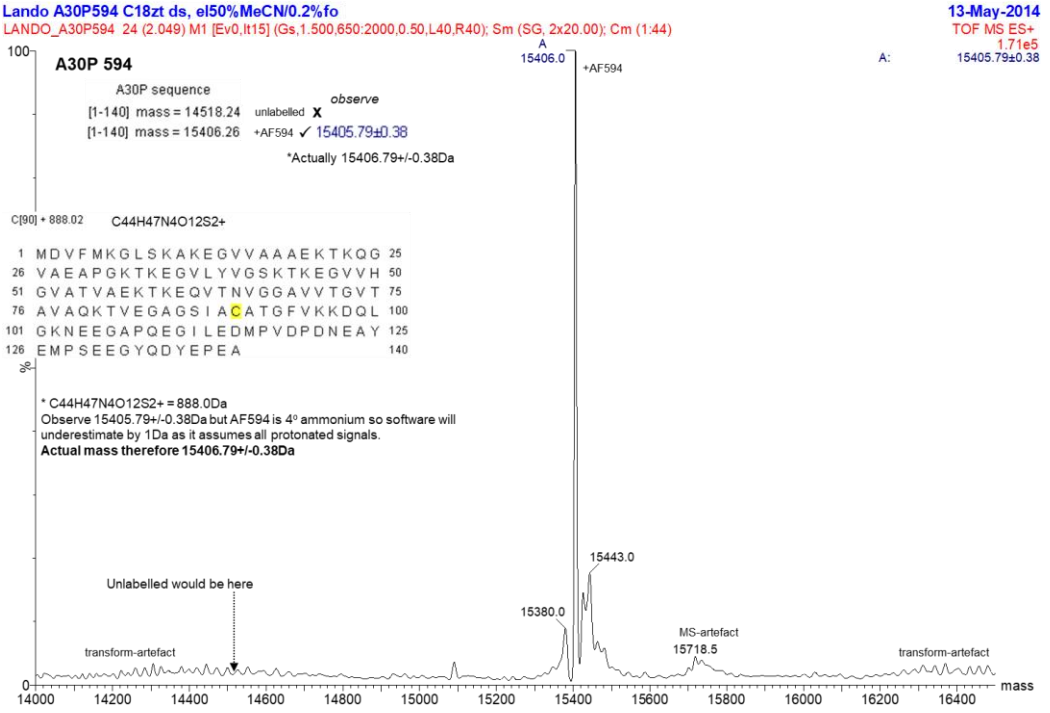


A30P A90C + AlexaFluor594

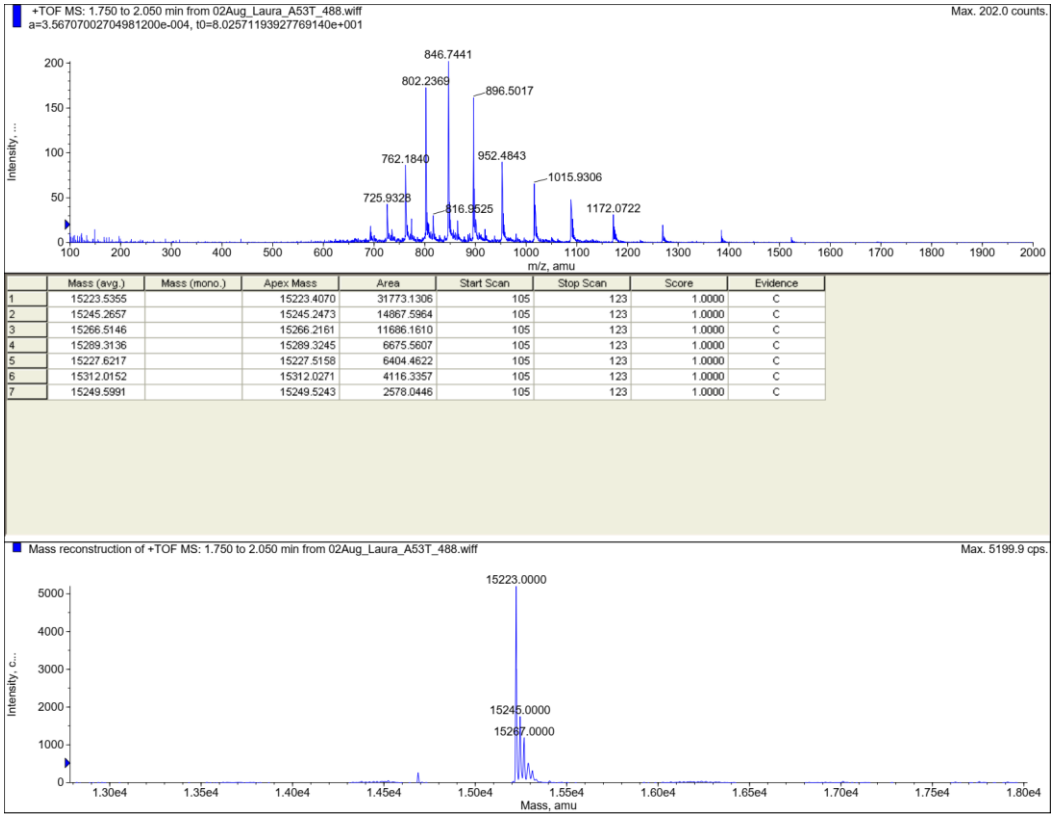


(protein product at 14680 Da is around 6.5% of the total protein species: sample kept for analysis)

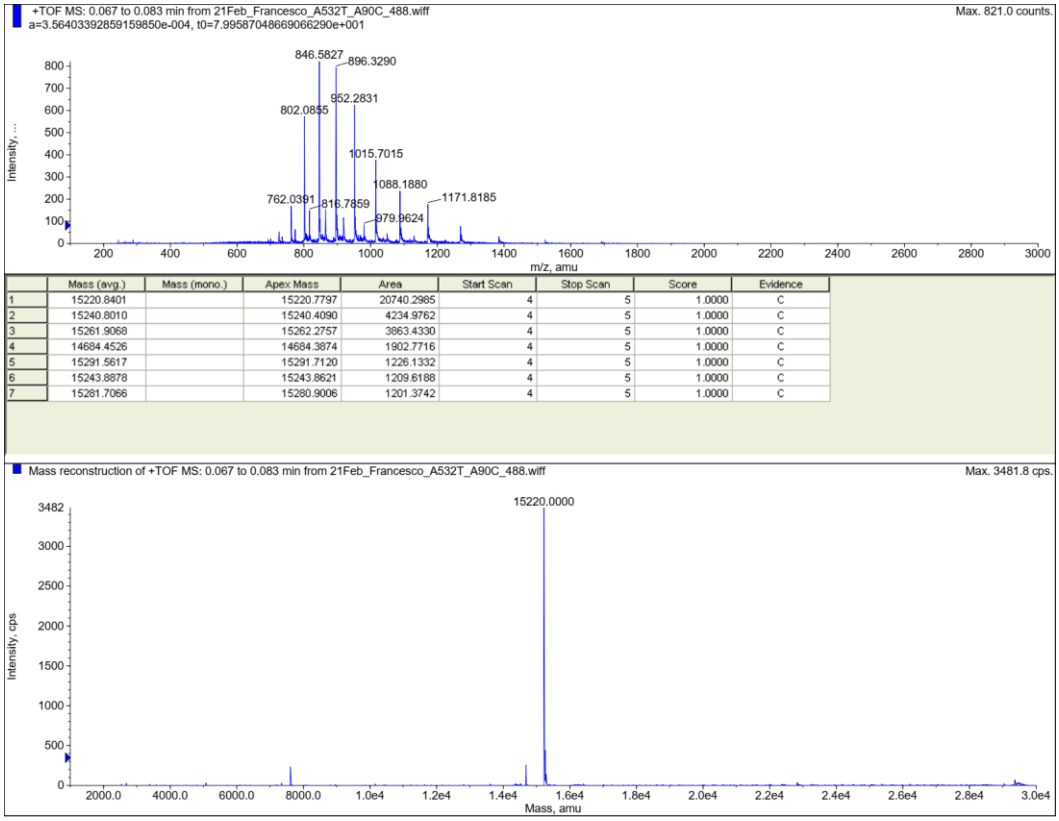
ESI analysis cap 3000V, cone 70V, calib ubiquitin same file, 70%MeOH/0.2%formic mobile phase



A53T A90C + AlexaFluor488



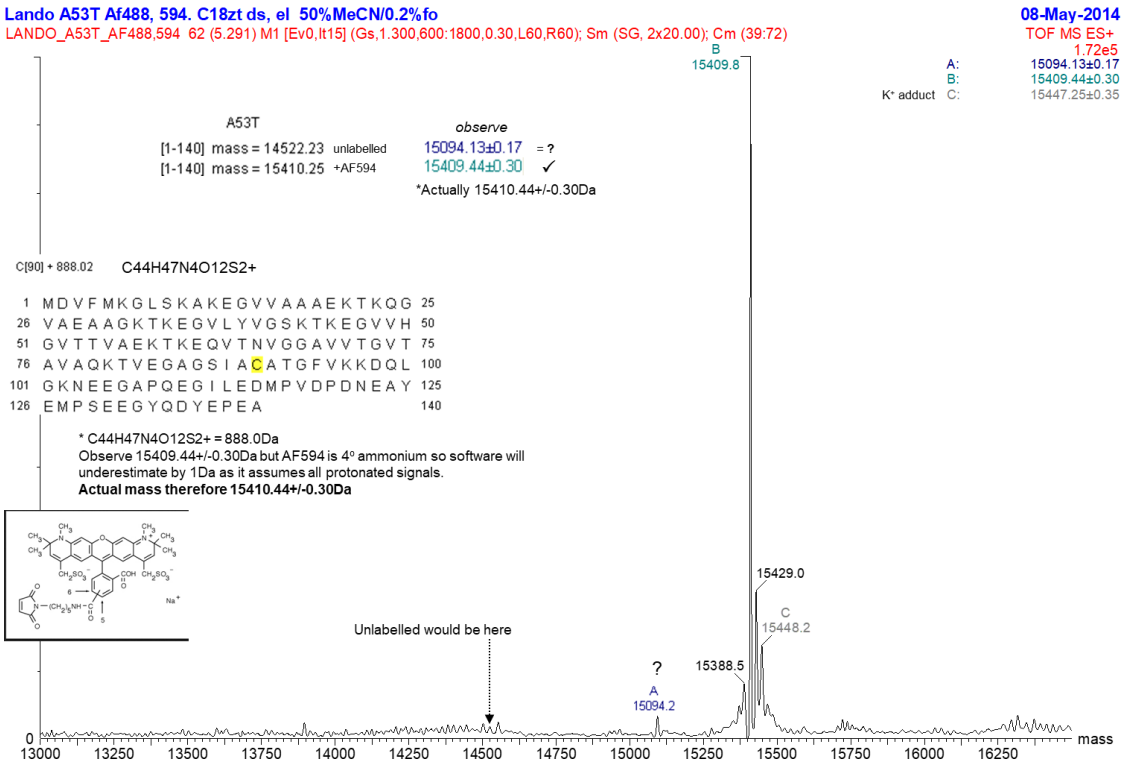
	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15223.5355		15223.4070	31773.1306	105	123	1.0000	C
2	15245.2657		15245.2473	14867.5964	105	123	1.0000	C
3	15266.5146		15266.2161	11686.1610	105	123	1.0000	C
4	15289.3136		15289.3245	6675.5607	105	123	1.0000	C
5	15227.6217		15227.5158	6404.4622	105	123	1.0000	C
6	15312.0152		15312.0271	4116.3357	105	123	1.0000	C
7	15249.5991		15249.5243	2578.0446	105	123	1.0000	C

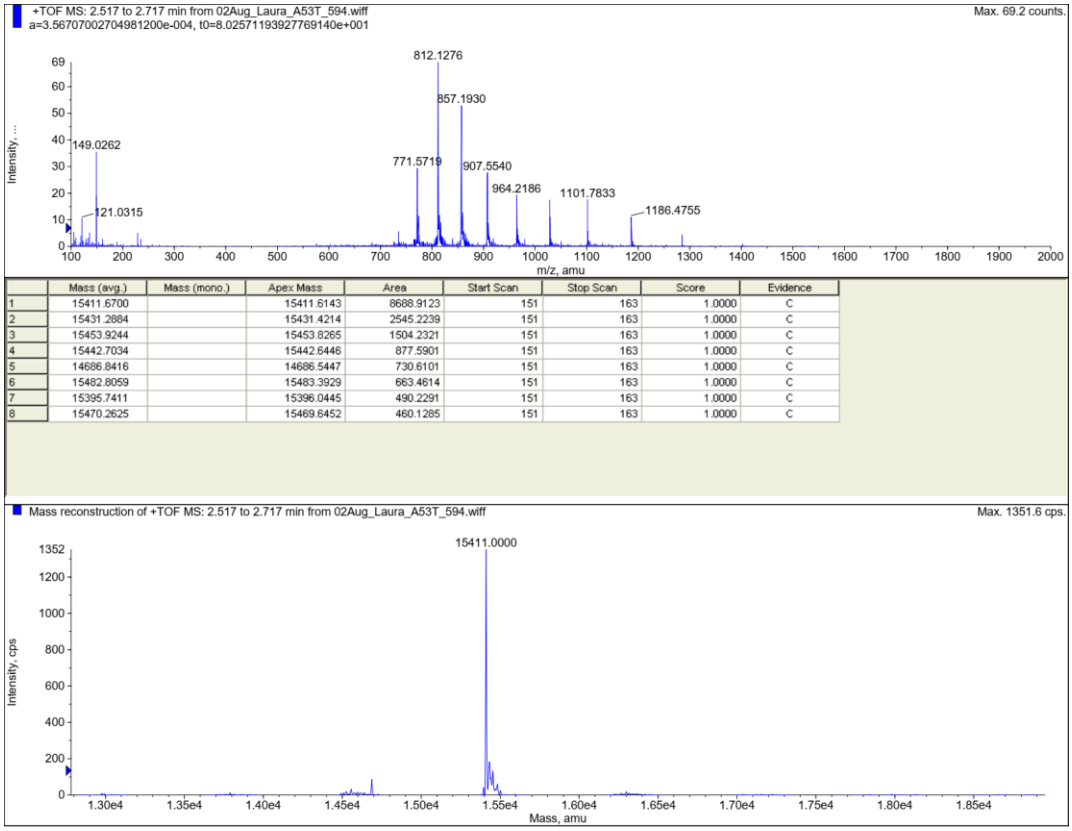


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15220.8401		15220.7797	20740.2985	4	5	1.0000	C
2	15240.8010		15240.4090	4234.9762	4	5	1.0000	C
3	15261.9068		15262.2757	3863.4330	4	5	1.0000	C
4	14684.4526		14684.3874	1902.7716	4	5	1.0000	C
5	15291.5617		15291.7120	1226.1332	4	5	1.0000	C
6	15243.8878		15243.8621	1209.6188	4	5	1.0000	C
7	15281.7066		15280.9006	1201.3742	4	5	1.0000	C

A53T A90C + AlexaFluor594

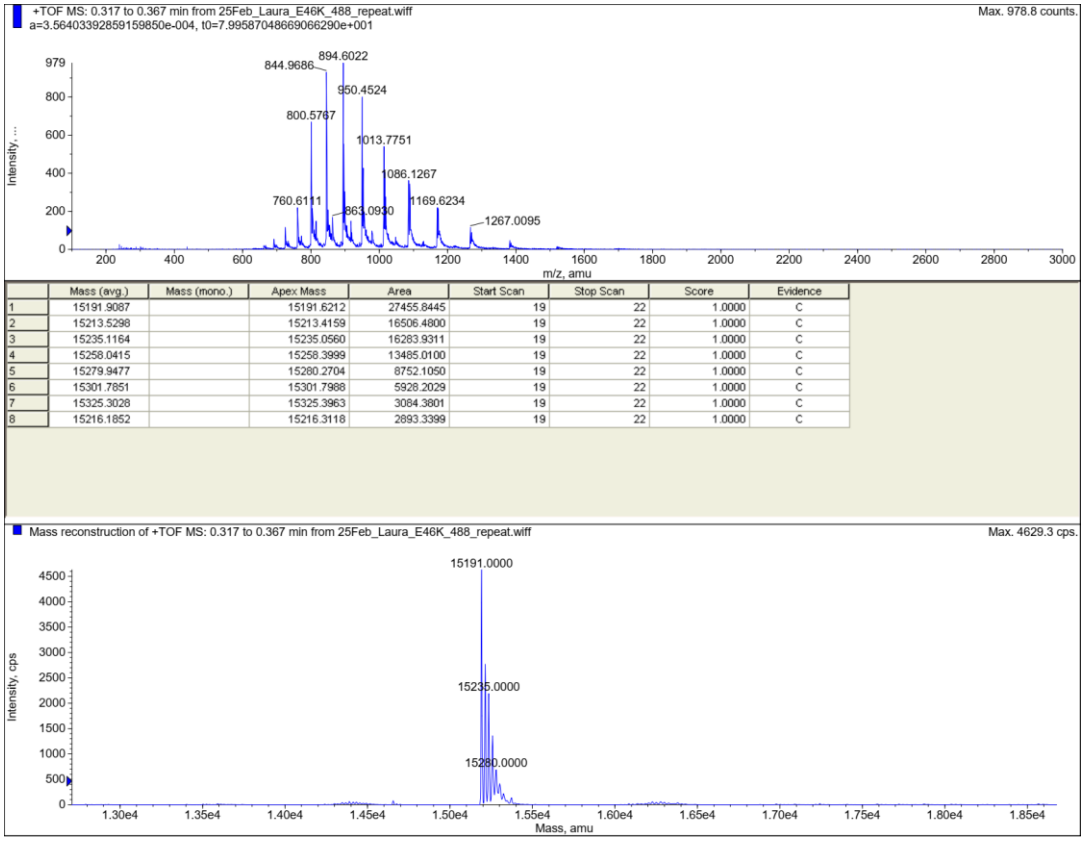
ESI analysis cap 3000V, cone 70V, calib ubiquitin same file, 70%MeOH/0.2%formic mobile phase



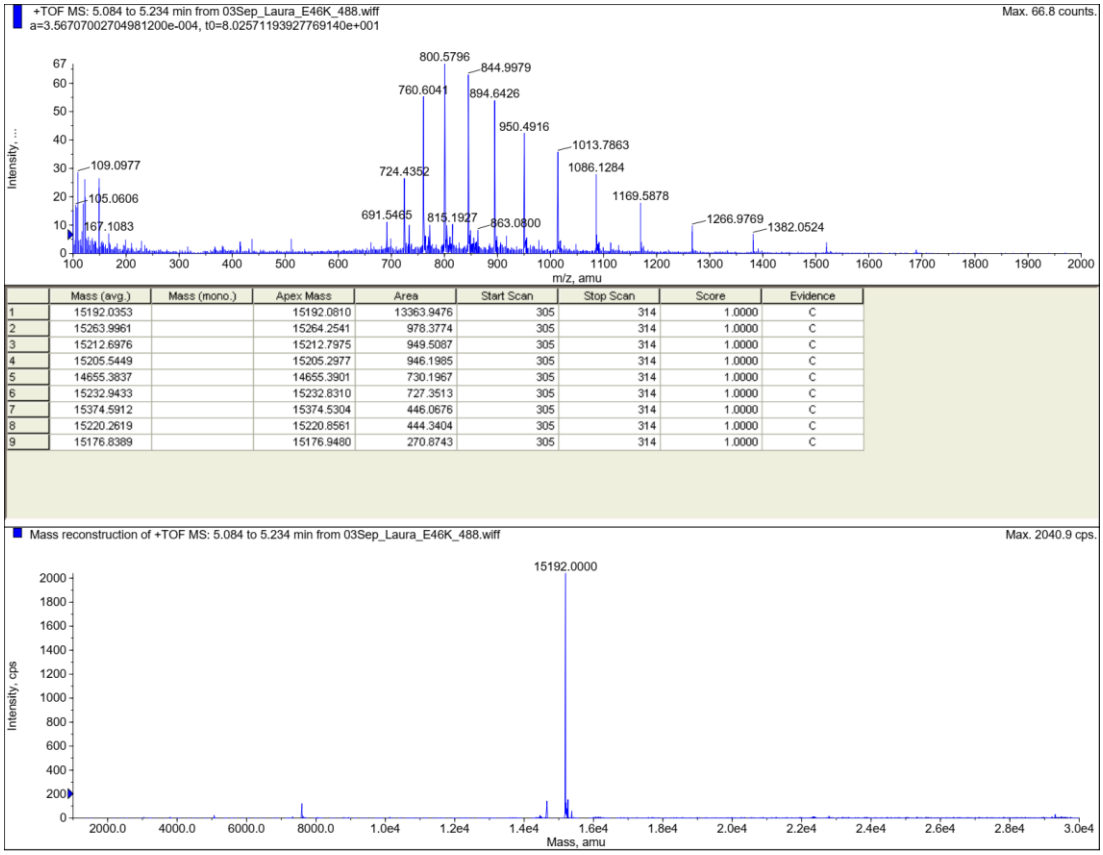


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15411.6700		15411.6143	8688.9123	151	163	1.0000	C
2	15431.2884		15431.4214	2545.2239	151	163	1.0000	C
3	15453.9244		15453.8265	1504.2321	151	163	1.0000	C
4	15442.7034		15442.6446	877.5901	151	163	1.0000	C
5	14686.8416		14686.5447	730.6101	151	163	1.0000	C
6	15482.8059		15483.3929	663.4614	151	163	1.0000	C
7	15395.7411		15396.0445	490.2291	151	163	1.0000	C
8	15470.2625		15469.6452	460.1285	151	163	1.0000	C

E46K A90C + AlexaFluor488

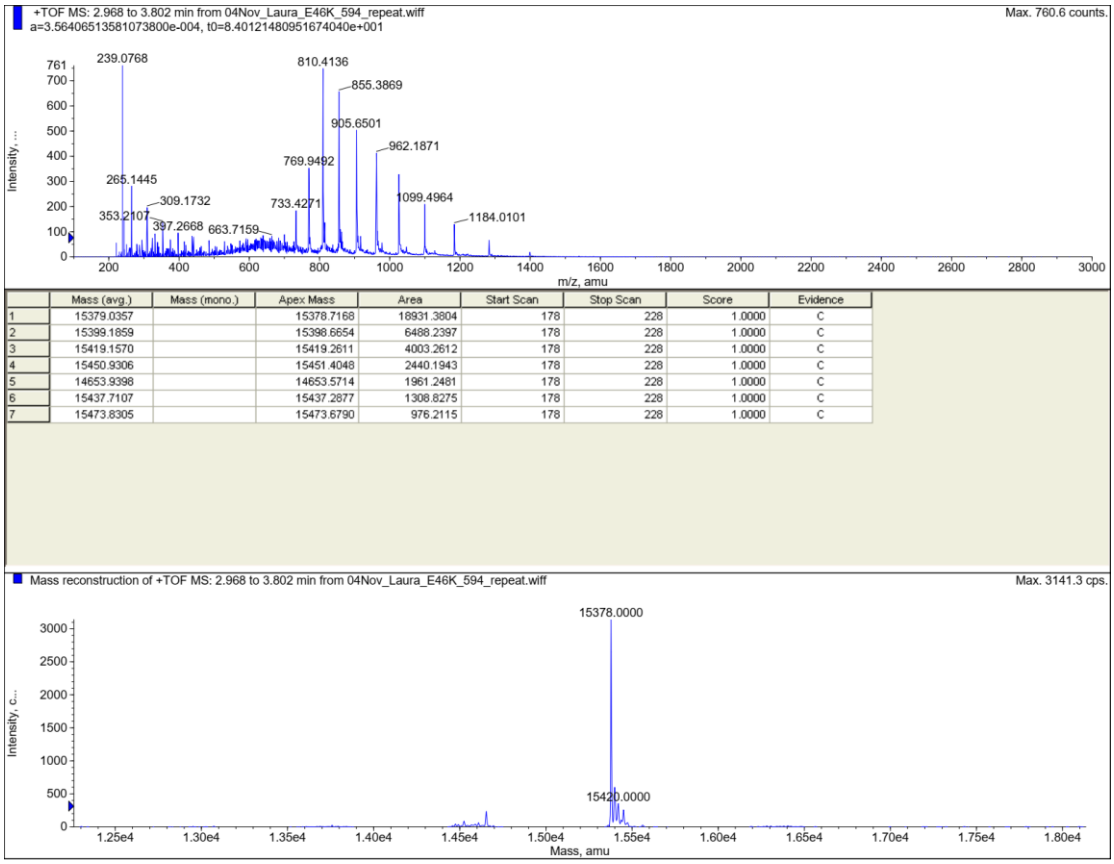


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15191.9087		15191.6212	27455.8445	19	22	1.0000	C
2	15213.5298		15213.4159	16506.4800	19	22	1.0000	C
3	15235.1164		15235.0560	16283.9311	19	22	1.0000	C
4	15258.0415		15258.3999	13485.0100	19	22	1.0000	C
5	15279.9477		15280.2704	8752.1050	19	22	1.0000	C
6	15301.7851		15301.7988	5928.2029	19	22	1.0000	C
7	15325.3028		15325.3963	3084.3801	19	22	1.0000	C
8	15216.1852		15216.3118	2893.3399	19	22	1.0000	C

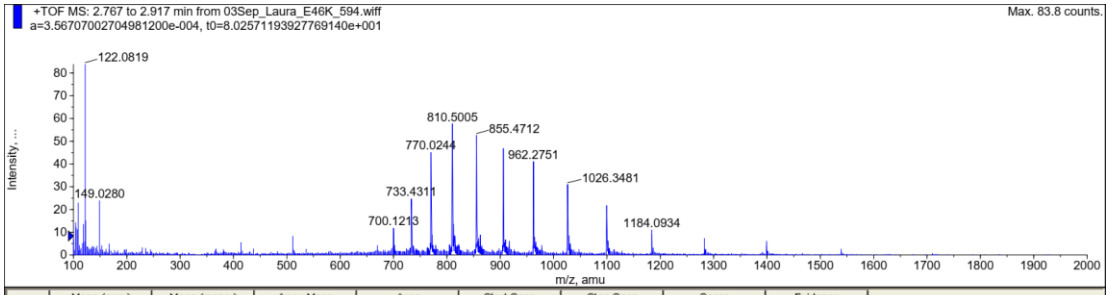


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15192.0353		15192.0810	13363.9476	305	314	1.0000	C
2	15263.9961		15264.2541	978.3774	305	314	1.0000	C
3	15212.6976		15212.7975	949.5087	305	314	1.0000	C
4	15205.5449		15205.2977	946.1985	305	314	1.0000	C
5	14655.3837		14655.3901	730.1967	305	314	1.0000	C
6	15232.9433		15232.8310	727.3513	305	314	1.0000	C
7	15374.5912		15374.5304	446.0676	305	314	1.0000	C
8	15220.2619		15220.8561	444.3404	305	314	1.0000	C
9	15176.8389		15176.9480	270.8743	305	314	1.0000	C

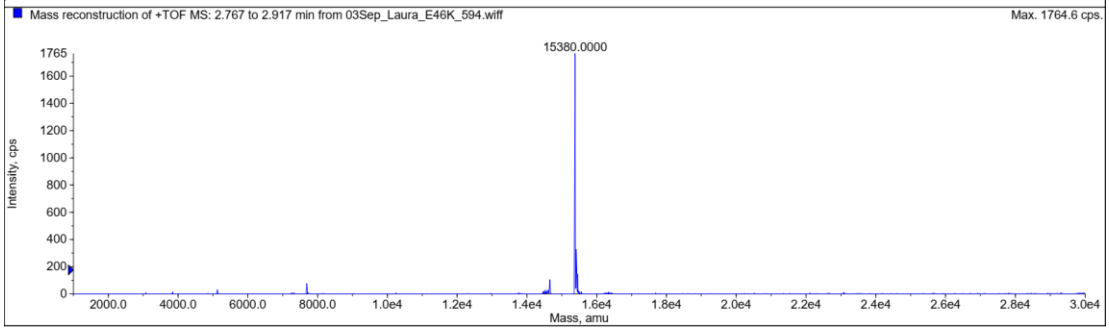
E46K A90C + AlexaFluor594

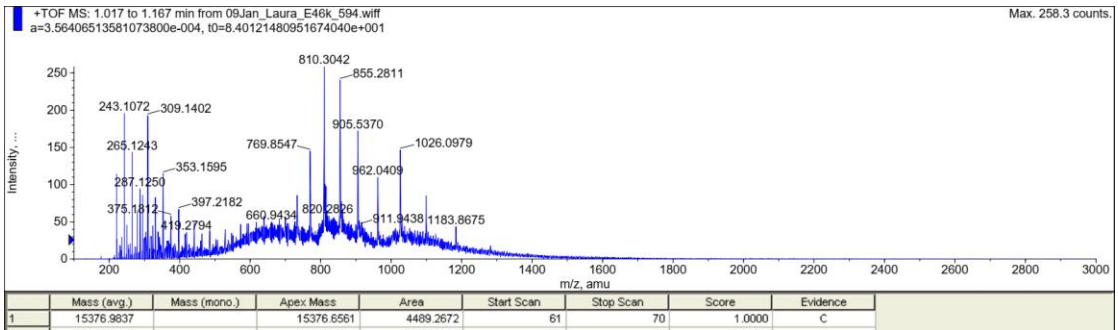


	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15379.0357		15378.7168	18931.3804	178	228	1.0000	C
2	15399.1859		15398.6654	6488.2397	178	228	1.0000	C
3	15419.1570		15419.2611	4003.2612	178	228	1.0000	C
4	15450.9306		15451.4048	2440.1943	178	228	1.0000	C
5	14653.9398		14653.5714	1961.2481	178	228	1.0000	C
6	15437.7107		15437.2877	1308.8275	178	228	1.0000	C
7	15473.8305		15473.6790	976.2115	178	228	1.0000	C



	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15380.5494		15380.5073	11911.0992	166	175	1.0000	C
2	15412.0301		15412.1732	2769.8532	166	175	1.0000	C
3	15395.4483		15395.3441	2589.8950	166	175	1.0000	C
4	15452.3984		15452.6537	1193.0292	166	175	1.0000	C
5	15424.6903		15425.2612	896.4905	166	175	1.0000	C
6	14655.3959		14655.1611	825.6120	166	175	1.0000	C
7	15365.3737		15365.3667	739.9532	166	175	1.0000	C
8	15398.7985		15398.6306	667.9696	166	175	1.0000	C
9	15438.1588		15437.8309	404.4049	166	175	1.0000	C
10	15481.8194		15481.8797	308.0934	166	175	1.0000	C





	Mass (avg.)	Mass (mono.)	Apex Mass	Area	Start Scan	Stop Scan	Score	Evidence
1	15376.9837		15376.6561	4489.2672	61	70	1.0000	C
2	15396.7630		15397.4559	1080.1214	61	70	1.0000	C
3	15416.6555		15415.8666	420.7365	61	70	1.0000	C