

# Uncovering the Early Assembly Mechanism for Amyloidogenic $\beta_2$ -Microglobulin Using Cross-linking and Native Mass Spectrometry\*

Received for publication, September 12, 2015, and in revised form, November 22, 2015. Published, JBC Papers in Press, December 10, 2015, DOI 10.1074/jbc.M115.691063

Zoe Hall<sup>†§</sup>, Carla Schmidt<sup>§</sup>, and Argyris Politis<sup>¶1</sup>

From the <sup>†</sup>Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom,

<sup>§</sup>Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom, and <sup>¶</sup>Department of Chemistry, King's College London, 7 Trinity Street, London SE1 1DB, United Kingdom

$\beta_2$ -Microglobulin ( $\beta_2m$ ), a key component of the major histocompatibility class I complex, can aggregate into fibrils with severe clinical consequences. As such, investigating the structural aspects of the formation of oligomeric intermediates of  $\beta_2m$  and their subsequent progression toward fibrillar aggregates is of great importance. However,  $\beta_2m$  aggregates are challenging targets in structural biology, primarily due to their inherent transient and heterogeneous nature. Here we study the oligomeric distributions and structures of the early intermediates of amyloidogenic  $\beta_2m$  and its truncated variant  $\Delta N6$ - $\beta_2m$ . We established compact oligomers for both variants by integrating advanced mass spectrometric techniques with available electron microscopy maps and atomic level structures from NMR spectroscopy and x-ray crystallography. Our results revealed a stepwise assembly mechanism by monomer addition and domain swapping for the oligomeric species of  $\Delta N6$ - $\beta_2m$ . The observed structural similarity and common oligomerization pathway between the two variants is likely to enable  $\Delta N6$ - $\beta_2m$  to cross-seed  $\beta_2m$  fibrillation and allow the formation of mixed fibrils. We further determined the key subunit interactions in  $\Delta N6$ - $\beta_2m$  tetramer, revealing the importance of a domain-swapped hinge region for formation of higher order oligomers. Overall, we deliver new mechanistic insights into  $\beta_2m$  aggregation, paving the way for future studies on the mechanisms and cause of amyloid fibrillation.

The major histocompatibility class I complex (MHC I) is found on the cell surface of all nucleated cells and is responsible for antigen presentation (1, 2).  $\beta_2$ -Microglobulin ( $\beta_2m$ )<sup>2</sup> is a key component of this complex. After its dissociation from MHC I, serum  $\beta_2m$  is broken down in the kidney. The buildup of circulating  $\beta_2m$  can result as a consequence of renal dysfunction and long term hemodialysis. This leads to amyloid fibril deposition in osteoarticular tissues and joint destruction in a condition known as dialysis-related amyloidosis.

\* The authors declare that they have no conflicts of interest with the contents of this article.

Author's Choice—Final version free via Creative Commons CC-BY license.

<sup>1</sup> To whom correspondence should be addressed: Dept. of Chemistry, King's College London, 7 Trinity St., London, SE1 1DB, UK. Tel.: 44-207-848-7514; E-mail: argyris.politis@kcl.ac.uk.

<sup>2</sup> The abbreviations used are:  $\beta_2m$ ,  $\beta_2$ -microglobulin; BS3, bis[sulfosuccinimidyl] suberate; CCS, collision cross-section; CX, cross-linking;  $\Delta N6$ - $\beta_2m$ , truncated  $\beta_2$ -microglobulin variant lacking six N-terminal residues; IM, ion mobility; PA, projection approximation; MD, molecular dynamics.

Amyloidogenic proteins such as  $\beta_2m$  tend to self-assemble into higher order oligomeric species through a complex aggregation process, which can ultimately lead to the formation of fibrils (3–6). Despite the clinical significance of amyloid formation, the principles governing the mechanisms for the aggregation process of  $\beta_2m$  and related proteins remains largely unknown, primarily due to the transient nature of intermediates on-pathway to fibril formation (7). Structural elucidation of  $\beta_2m$  oligomeric intermediates is, therefore, a challenging task that is further complicated by the uncertainty in differentiating between specific and nonspecific protein aggregates.

Despite these obstacles progress has been made most prominently from techniques such as nuclear magnetic resonance spectroscopy (NMR) (8), x-ray crystallography (9), atomic force microscopy (10), cryo-electron microscopy (EM) (11), and hydrogen/deuterium exchange (12) and by combining NMR with mass spectrometry (MS) (13). MS in particular is well suited for studying heterogeneous assembly intermediates, including proteins populating multiple oligomeric states (14). When coupled with ion mobility (IM), IM-MS allows the separation of different conformational states of co-populated oligomers (5, 15). IM-MS has been successfully employed to investigate the structure and dynamics of amyloid assembly intermediates revealing information on the aggregation process of A $\beta$ 40 and A $\beta$ 42 complexes (15, 16) and more recently human amylin (17). Structural insights have also been gained for  $\beta_2m$  (18, 19) where IM-MS experiments suggested either an elongated or more compact assembly mechanism for full-length  $\beta_2m$  under different solution conditions (20).

The integration of different experimental techniques with modeling can provide powerful means to interrogate candidate models of protein assemblies (21–25). In particular so-called hybrid approaches, which combine information from complementary experiments, have shed light on complexes intractable by single techniques, exemplified by the structural elucidation of the nuclear pore complex (26), the 26S proteasome (27, 28), and the eukaryotic translation initiation factor 3 (29, 30).

Here we used an integrative MS-based strategy for generating three-dimensional structural models of the oligomeric assembly intermediates of  $\Delta N6$ - $\beta_2m$ , a truncated  $\beta_2m$  isoform (11.1-kDa monomer).  $\Delta N6$ - $\beta_2m$  makes up to 30% of amyloid deposits extracted from dialysis-related amyloidosis patients and can act as a seed for full-length  $\beta_2m$  fibrillogenesis *in vitro* (1). Contrary to full-length  $\beta_2m$ ,  $\Delta N6$ - $\beta_2m$  is highly amyloido-

genic at neutral pH, making it a convenient model for studying  $\beta_2$ m aggregation under laboratory conditions (2). Furthermore, the x-ray crystal structure of the dimeric intermediate built by the self-association of two  $\Delta N6$ - $\beta_2$ m monomers was recently solved, and proposed as a building block for growing oligomers on-pathway to fibril formation (1). In this structure domain-swapping occurs through the so-called hinge region, which corresponds to two NHVTLSSQ heptapeptides interacting in an antiparallel fashion (1).

Using a combination of experimental and computational techniques, we predict the structures and an early assembly mechanism for  $\Delta N6$ - $\beta_2$ m oligomers. We further compare oligomers of the truncated variant with those of the full-length protein, highlighting similar oligomeric distributions and compact topologies as well as inter- and intraprotein distances. This points to a common assembly mechanism in the early stages of their aggregation and may facilitate the ability of the truncated variant to cross-seed and form mixed fibrils (31) with full-length  $\beta_2$ m *in vivo*. The data and the structural models generated from the integrative strategy further suggest an elongation mechanism of monomer addition consistent with domain swapping and self-templated growth. Furthermore, our model for  $\Delta N6$ - $\beta_2$ m tetramer shows that the domain-swapped hinge region found in  $\Delta N6$ - $\beta_2$ m dimer is key to both intra- and inter-dimer interactions.

## Experimental Procedures

**Protein Preparation**— $\Delta N6$ - $\beta_2$ m and  $\beta_2$ m were expressed in *Escherichia coli* and purified using ion exchange chromatography and size exclusion chromatography as previously described (1). Lyophilized protein was dissolved in 100 mM ammonium acetate, pH 5, before MS analysis.

**Ion Mobility-Mass Spectrometry**—IM-MS experiments were performed on a quadrupole ion mobility time-of-flight mass spectrometer (Synapt HDMS, Waters Corp., Manchester, UK) modified such that the traveling-wave IM cell is replaced with an 18-cm drift cell with radial RF confinement (RF amplitude 200 V) and a linear voltage gradient along the axis of ion transmission, as described in detail previously (32). The following parameters were used: source pressure 4–6 mbar, capillary voltage 1.0–1.5 kV, sample cone voltage 20 V, bias voltage 20 V, IM entrance DC 5 V, trap gas 6 ml min<sup>-1</sup>, trap collision energy 5 V. Helium (2 torr) was used as the buffer gas, and the drift voltage varied from 50 to 200 V. All spectra were mass-calibrated using cesium iodide (100 mg ml<sup>-1</sup>).

**Chemical Cross-linking MS**—50  $\mu$ l of  $\Delta N6$ - $\beta_2$ m or  $\beta_2$ m were cross-linked with 10  $\mu$ l of a 25 mM 1:1 mixture of deuterated (d4) and non-deuterated (d0) bis[sulfosuccinimidyl] suberate (BS3). The reaction mixture was incubated at 25 °C and 400 rpm for 1 h. 10  $\mu$ l of the cross-linked proteins were analyzed by gel electrophoresis (NuPAGE system, Invitrogen) according to the manufacturer's protocol. The proteins were digested in-gel as described elsewhere (33).

The mixture of cross-linked and non-cross-linked peptides was analyzed by liquid chromatography-coupled tandem-mass spectrometry (LC-MS/MS) employing an LTQ-Orbitrap XL hybrid mass spectrometer (Thermo Scientific) coupled with a Dionex UltiMate 3000 RSLC nano System (Thermo Scientific).

The peptides were separated by nanoLC (mobile phase A, 0.1% (v/v) formic acid; mobile phase B, 80% (v/v) acetonitrile/0.1% (v/v) formic acid). The peptides were loaded onto a trap column (2 cm, HPLC column Acclaim PepMap 100, C18, 100- $\mu$ m inner diameter, particle size 5  $\mu$ m; Thermo Scientific) and separated on an analytical C18 capillary column (50 cm, HPLC column Acclaim PepMap 100, C18, 75- $\mu$ m inner diameter particle size 3  $\mu$ m; Thermo Scientific) at a flow rate of 300 nl/min and a gradient of 5–80% (v/v) mobile phase B over 74 min.

The peptides were directly eluted into the mass spectrometer. Mass spectrometric conditions were: spray voltage of 1.8 kV, capillary temperature 180 °C, normalized collision energy 35% at an activation of  $q = 0.25$ , and an activation time of 30 ms. The LTQ-Orbitrap XL was operated in data-dependent mode. Survey full scan MS spectra were acquired in the orbitrap ( $m/z$  300–2,000) with a resolution of 30,000 at  $m/z$  400 and an automatic gain control target at 10<sup>6</sup>. The five most intense ions were selected for collision-induced dissociation in the linear ion trap at an automatic gain control target of 30,000. Selection of previously selected precursor ions was dynamically excluded for 30 s. Singly charged ions as well as ions with unrecognized charge state were also excluded. Internal calibration of the Orbitrap was performed using the lock mass option (lock mass:  $m/z$  445.120025) (34).

mzXML files were generated from raw data using the MassMatrix file conversion tool. Potential cross-links were identified by searching against a reduced database containing  $\Delta N6$ - $\beta_2$ m and  $\beta_2$ m protein sequences using the MassMatrix Database Search Engine (35). Search parameters were: enzyme, trypsin; missed cleavage sites, two; variable modifications, carbamidomethylation of cysteines and oxidation of methionine; mass accuracy filter, 10 ppm for precursor ions and 0.8 Da for fragment ions; minimum pp and pp2 values, 5.0; minimum ptag, 1.3; maximum number of cross-links per peptide, 1. All searches were performed twice, including the deuterated and the non-deuterated BS3 cross-linker, respectively.

**Modeling Restraints from Cross-linking**—To test if the identified cross-links in the  $\Delta N6$ - $\beta_2$ m tetramer (or trimer) were arising from interdimer or intradimer interactions, we projected the cross-links confirmed by MS/MS quality onto the dimer x-ray crystal structure (PDB ID 2X89). We measured the physical C $\alpha$ -C $\alpha$  distances (36) to check if these were within the upper-bound interresidue distance threshold (35 Å) (37). The cross-links, which do not satisfy our distance threshold, are more likely to generate from interdimer interactions in the tetramer (or trimer) and were, therefore, assigned as such in our modeling analysis. Those that were within the distance threshold were most likely to arise from intradimer interactions.

**MS-based Hybrid Approach**—We employed a hybrid approach for structural determination of oligomeric intermediates of  $\Delta N6$ - $\beta_2$ m, primarily based on native MS, IM-MS and cross-linking MS (CX-MS) (27, 29) (Fig. 1). From native MS, we established the oligomeric state of the identified complexes (38). By combining MS with IM, topological information in the form of an orientation-averaged collision cross-section (CCS) was derived (39). The measured CCS from IM-MS was used as shape restraint for interrogating candidate structural models

## Assembling Oligomers of $\Delta N6$ - $\beta_2$ -Microglobulin

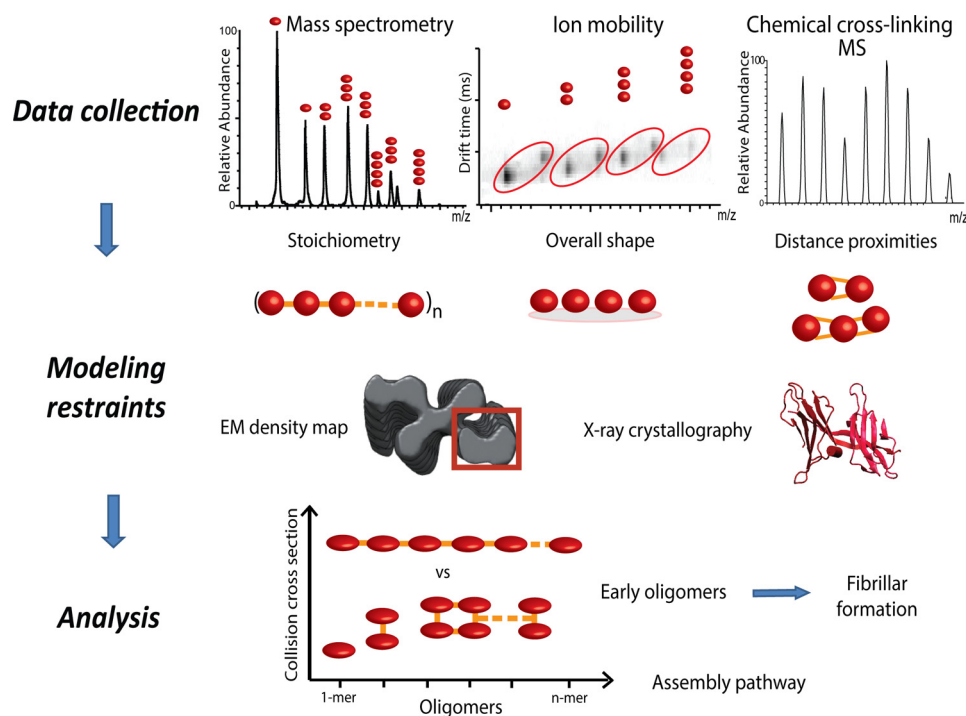


FIGURE 1. **Integrative modeling workflow for structure characterization of oligomeric protein assemblies.** To predict the structures of higher oligomeric states of proteins, we implement an integrative modeling strategy. Here we use information derived from native MS, IM-MS, and CX-MS. The acquired data and other available information are converted into spatial restraints, which are subsequently utilized by a scoring function to guide the search for candidate model structures. Finally, an analysis step (e.g. clustering) of the top-scoring models determined the most likely structures of the oligomeric assembly pathways.

(22). CX-MS identified lysines in close proximity and was used as a distance restraint (27). In addition to MS data, we made use of available structures from x-ray crystallography and NMR as well as EM density maps. The atomic level structures were used as starting points in our modeling strategy, whereas a segment of the EM map of  $\beta_2$ m assembled into fibrils (EM Database ID 1613; type A) was used as a volume restraint, as there is no available EM density map for  $\Delta N6$ - $\beta_2$ m fibrils (11). Structural information obtained from these methods was encoded into restraints and exploited by a scoring function for subsequent modeling analysis.

**Integrative Modeling and Scoring Function**—We generated structural models of the assemblies by employing a Monte Carlo search algorithm. The building process was guided by a scoring function that encoded the experimental data as a sum of individual restraints. This scoring function ( $S$ ) evaluates the ensemble of candidate models generated against their quality-of-fit with the input data.

$$S = W_{\text{IM-MS}}S_{\text{IM-MS}} + W_{\text{CX-MS}}S_{\text{CX-MS}} + W_{\text{EM}}S_{\text{EM}} \quad (\text{Eq. 1})$$

where  $S_{\text{IM-MS}}$  and  $S_{\text{CX-MS}}$  refer to IM-MS and CX-MS restraints, respectively.  $S_{\text{IM-MS}}$  was implemented as a harmonic potential function (22, 23), whereas  $S_{\text{CX-MS}}$  was applied as a distance restraint between two interacting residues (27, 29). The EM ( $S_{\text{EM}}$ ) restraint assessed the quality-of-fit between the model and the corresponding molecular volume of an appropriate section of the EM map, as defined by the cross-correlation coefficient. IM-MS, CX-MS, and EM were given 1:2:2 weightings ( $W$ ), respectively, consistent with previous benchmark studies (27).

To assess the uniqueness of the ensemble of generated mod-

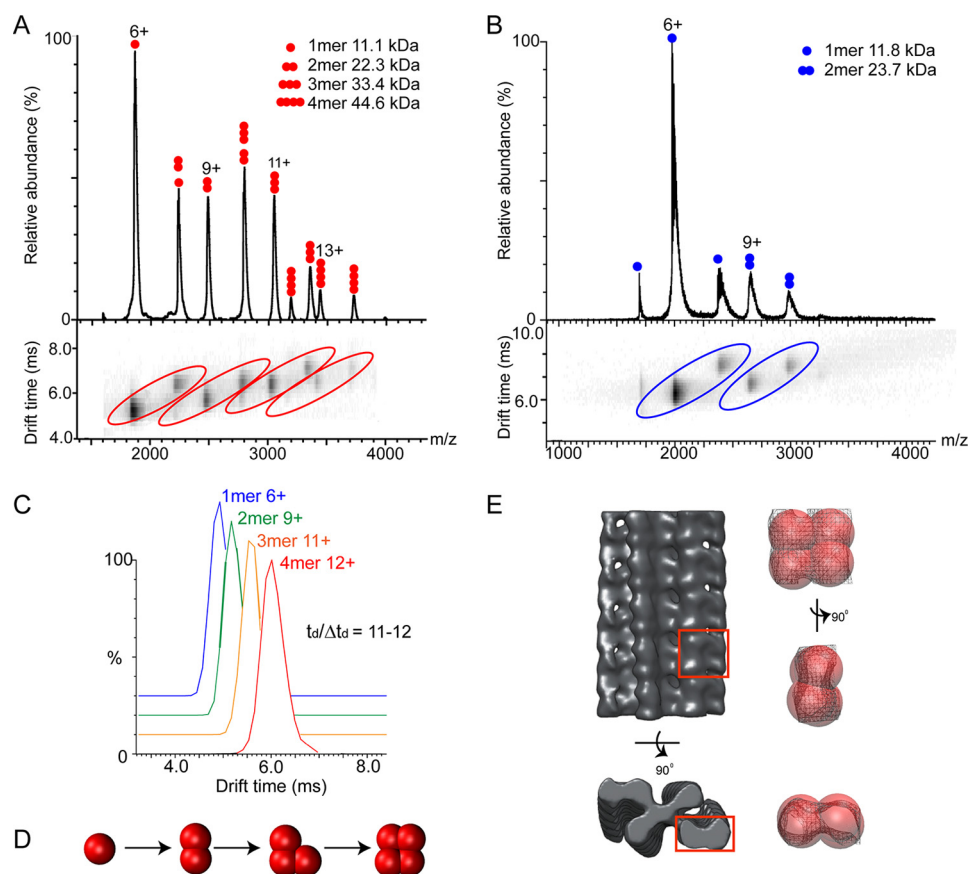
els we performed ensemble analysis (e.g. clustering of top-scoring solutions), and the final solution was selected from the major cluster. The visual molecular dynamics (VMD) and the UCSF Chimera packages were used for visualization of the structures (40).

**Collision Cross-section Calculations**—To interpret the experimentally obtained CCS values, we compared them to theoretically calculated CCSs (41). Theoretical CCSs were obtained with the open source MOBCAL code using the projection approximation (PA) algorithm (42, 43). The PA method is known to underestimate the experimental CCS of proteins by neglecting multiple collisions between ions and buffer gas (7, 43). However, it has been shown that it is correlated with the experimental CCS for protein complexes ( $R^2 > 0.99$ ) (44). We use the scaled PA CCS as previously described, where the experimental CCS can be typically predicted ( $\pm 3\%$ ) by multiplying the PA CCS by a factor of 1.14 (44). All CCS calculations include hydrogen atoms.

**Molecular Dynamics Simulations**—All simulations were performed in single (solution phase) or double floating-point precision (gas phase) with GROMACS 4.5.3 using the OPLS-AA/L forcefield (45).

Gas-phase simulations (10 ns) were performed at 300 K as described previously (46). The MS observed charge state was distributed evenly over solvent-accessible basic residues ( $< 5 \text{ \AA}$  from the surface) (47, 48). As such, we assigned the charged residues for the 6+ monomer, 9+ dimer, 11+ trimer, and 13+ tetramer charge states of  $\Delta N6$ - $\beta_2$ m. CCSs were calculated for structures every 25 ps using the scaled PA method implemented in MOBCAL.





**FIGURE 2. IM-MS data and coarse-grained modeling reveal a compact tetrameric intermediate in  $\Delta N6$ - $\beta_2$ m aggregation.** *A*, IM-MS of  $\Delta N6$ - $\beta_2$ m ( $15 \mu\text{M}$ , pH 5) revealed four charge state distributions, corresponding to the 1-mer, 2-mer, 3-mer, and 4-mer. *B*, IM-MS of full-length  $\beta_2$ m ( $15 \mu\text{M}$ , pH 5) reveals charge state distributions corresponding to 1-mer and 2-mer. Higher order oligomers up to tetramer could be observed at higher protein concentrations ( $30 \mu\text{M}$ , not shown). *C*, representative arrival time distributions for the 4-mer (12+), 3-mer (11+), 2-mer (9+), and 1-mer (6+) of  $\Delta N6$ - $\beta_2$ m. *D*, the  $\Delta N6$ - $\beta_2$ m tetramer was assembled from the monomer by the stepwise addition of subunits using a coarse-grained modeling strategy. The model with lowest score at each stage was chosen as the starting point for the subsequent addition of further monomer units. *E*, manual fitting of the low resolution  $\Delta N6$ - $\beta_2$ m tetramer (best scored model after clustering) into the cryo-EM density map of  $\beta_2$ m fibrils (EM Database ID 1613). Segments (red box) of the EM map are shown as surface-and-mesh representations in three different orientations.

Solution phase simulations (10 ns) were carried out similarly, except periodicity and a cutoff of 0.9 and 1.4 nm were used for electrostatic and van der Waals forces respectively; an integration step of 2 fs was used. Acidic and basic residues were charged as appropriate for solution. Total charge of the system was neutralized by the addition of an appropriate number of sodium ions.

**Software and Scripts**—Our integrative protocol was implemented within the open source Integrative Modeling Platform (IMP) software package.

## Results

**$\Delta N6$ - $\beta_2$ m Assembles into Compact Oligomers**—We carried out IM-MS on  $\Delta N6$ - $\beta_2$ m at various monomer concentrations ( $10$ – $30 \mu\text{M}$ , pH 5) revealing multiple oligomeric species in equilibrium (Fig. 2*A*). Our experiments demonstrated that the formation of specific oligomers in solution is highly concentration-dependent (data not shown). At  $10 \mu\text{M}$ , the predominant species observed were monomers and dimers, with a low amount of trimers formed. At  $15 \mu\text{M}$ , we could clearly observe four charge state distributions corresponding to monomers, dimers, trimers, and tetramers of  $\Delta N6$ - $\beta_2$ m (Fig. 2*A*). Higher concentrations ( $30 \mu\text{M}$ ) revealed higher order oligomers

(>pentamer). As a control we carried out similar experiments using cytochrome *c*, a monomeric protein of similar molecular mass (12 kDa). At  $15 \mu\text{M}$  and below, only monomeric species were detected, whereas at higher concentrations (*e.g.*  $30 \mu\text{M}$ ), we could observe low intensity peaks for dimeric and trimeric species (data not shown). Therefore, we carried out experiments at a protein concentration of  $15 \mu\text{M}$  to minimize any contribution from nonspecific aggregation (49). To correctly assign the oligomeric species of  $\Delta N6$ - $\beta_2$ m, separation in mobility space was critical, as peaks at certain  $m/z$  values were co-populated with multiple species. For instance,  $m/z$  2800 is composed of both dimer (8+ charge state) and trimer (12+ charge state) (Fig. 2*A*). Separation of different species with the same  $m/z$  but different CCS is a major strength of IM-MS.

CCS measurements were performed for all charge states of the four oligomers observed for  $\Delta N6$ - $\beta_2$ m ( $15 \mu\text{M}$ ). Arrival time distributions (ATD) had Gaussian peak shapes with  $t_D/\Delta t_D \sim 11$ – $12$ , consistent with a single conformer or group of closely related conformations (representative ATDs are shown in Fig. 2*C*) (39). The CCS deviation between different charge states of a particular oligomer was typically <5% (considered to be within experimental error), and we therefore averaged the CCS

## Assembling Oligomers of $\Delta N6$ - $\beta_2$ -Microglobulin

across all measured charge states. The average CCS for monomer ( $CCS_{exp}$ ,  $1200 \pm 36 \text{ \AA}^2$ ), dimer ( $CCS_{exp}$ ,  $1900 \pm 57 \text{ \AA}^2$ ), trimer ( $CCS_{exp}$ ,  $2530 \pm 76 \text{ \AA}^2$ ), and tetramer ( $CCS_{exp}$ ,  $3057 \pm 92 \text{ \AA}^2$ ) were measured (Table 1).

We applied a similar approach to interrogate full-length  $\beta_{2m}$ , which is expected to oligomerize to a lesser extent than its truncated variant at pH 5 (19). We observed monomer and dimer at a protein concentration of  $15 \mu\text{M}$  (Fig. 2B), with trimers and tetramers only observed at higher concentrations ( $30 \mu\text{M}$ , data not shown). We measured CCSs for the oligomers of  $\beta_{2m}$  at  $15 \mu\text{M}$  (monomer and dimer) and at  $30 \mu\text{M}$  (trimer and tetramer) (Table 1) revealing values  $<5\%$  greater than those measured for  $\Delta N6$ - $\beta_{2m}$ , consistent with the higher molecular weight. This suggests that the oligomers of full-length  $\beta_{2m}$  and

its truncated variant adopt similar conformations for their early oligomeric intermediates.

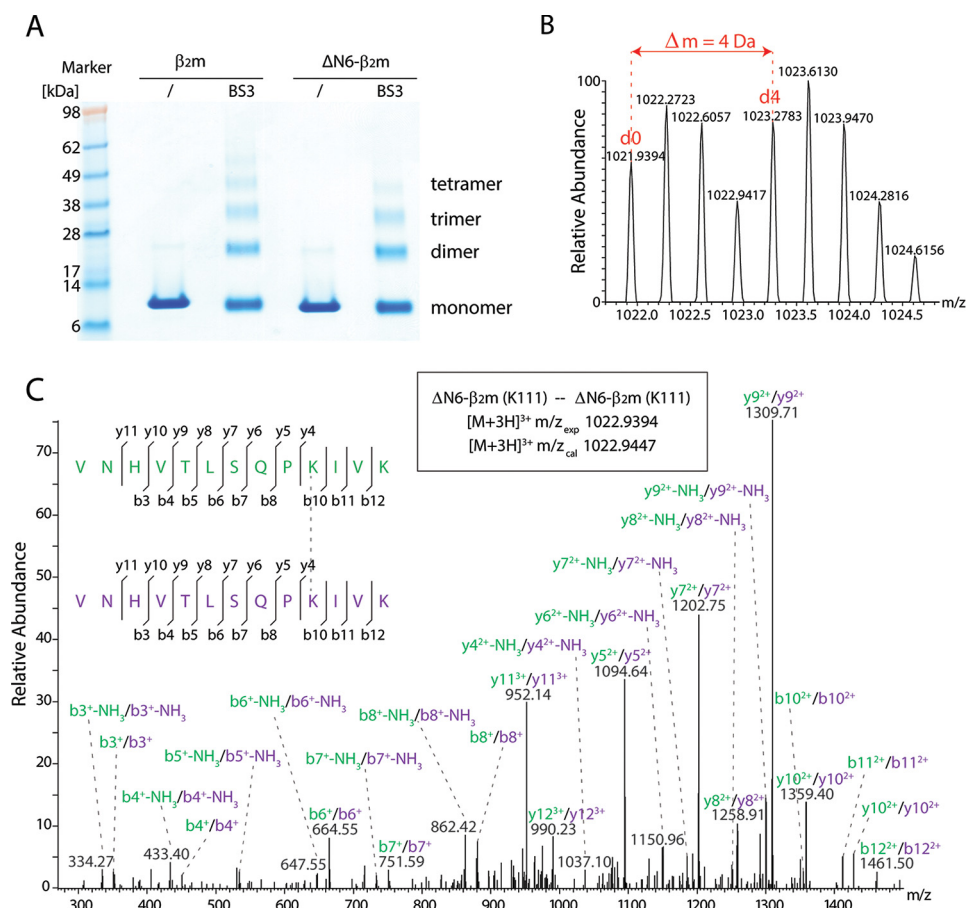
To establish the overall topology (*i.e.* compact or elongated) of  $\Delta N6$ - $\beta_{2m}$  oligomers, we used the IM-MS data to inform a coarse-grained modeling strategy reported previously (22). Monomers were represented as spheres, with radii defined by the monomer CCS. Models were then built for the dimer and trimer by varying the intersubunit distances and angles. These models were then scored, and the one with the lowest total score at each stage was taken to the next step, to form the  $(n+1)$  oligomer (Fig. 2D). To sample conformational space for the tetramer, a Monte Carlo approach was used keeping the relative position of the three other subunits fixed. The 1% lowest scoring models were clustered, revealing three distinct clusters. The largest cluster (84.4%) is represented by a compact topological arrangement of subunits. This low resolution model of tetrameric  $\Delta N6$ - $\beta_{2m}$  from IM-MS is consistent with a recently published EM map of  $\beta_{2m}$  fibrils (Fig. 2E).

**Protein Interactions in  $\Delta N6$ - $\beta_{2m}$  Oligomers**—Next, we employed CX-MS to gain insights into the interprotein interactions at atomic level.  $\Delta N6$ - $\beta_{2m}$  was incubated with a 1:1 mixture of non-deuterated ( $d_0$ ) and deuterated ( $d_4$ ) BS3. The cross-linked proteins were separated by gel electrophoresis (Fig. 3A). The bands corresponding to monomer, dimer, trimer, and tetramer of  $\Delta N6$ - $\beta_{2m}$  were cut, and the proteins were digested

**TABLE 1**  
IM-MS experiments showed similar measured CCS for oligomeric species of  $\Delta N6$ - $\beta_{2m}$  ( $15 \mu\text{M}$ , pH 5) and full-length  $\beta_{2m}$  ( $15$ – $30 \mu\text{M}$ , pH 5)

Oligomeric species	$\Delta N6$ - $\beta_{2m}$		$\beta_{2m}$	
	CCS	Error	CCS	Error
Monomer	$1200$	$36$	$1250$	$13$
Dimer	$1900$	$57$	$1994$	$20$
Trimer	$2530$	$76$	$2622^a$	$52$
Tetramer	$3057$	$92$	$3213^a$	$64$

<sup>a</sup> CCS measurement at  $30 \mu\text{M}$ .

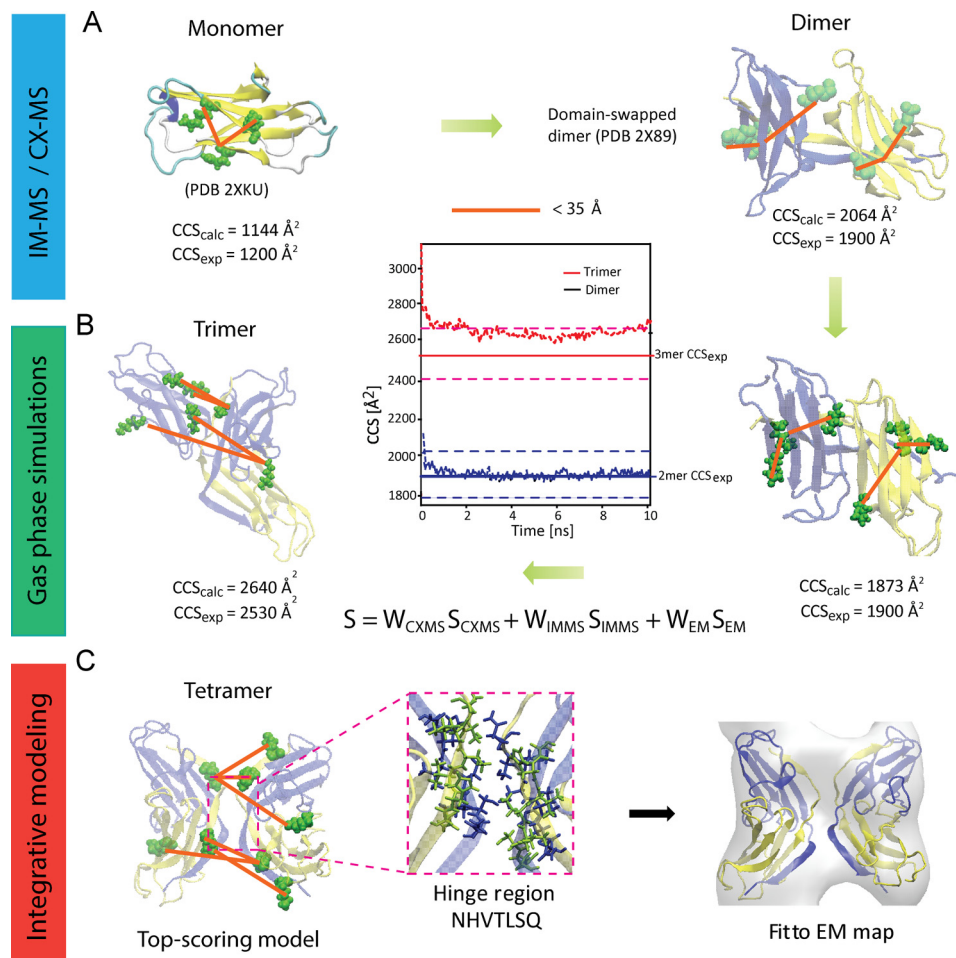


**FIGURE 3. Cross-linking of  $\Delta N6$ - $\beta_{2m}$  and  $\beta_{2m}$ .** A,  $\Delta N6$ - $\beta_{2m}$  and  $\beta_{2m}$  were cross-linked with a 1:1 mixture of BS3- $d_0/d_4$ . The gels show the presence of monomers, dimers, trimers, and tetramers. B, cross-links were verified by the presence of peak pairs in the mass spectra corresponding to light ( $d_0$ ) and heavy ( $d_4$ ) cross-linked peptides. C, cross-linked dipeptides containing the same or overlapping peptide sequences are unambiguously assigned to interprotein cross-links. An example spectrum is shown.

**TABLE 2**  
Identified cross-links for  $\Delta$ N6- $\beta_2$ m and  $\beta_2$ m oligomers

Peptide sequence 1	Peptide sequence 2	Residue 1	Residue 2	No. spectra
<b><math>\Delta</math>N6-<math>\beta_2</math>m monomer</b>				
VNHVTLSPKIVK	<u>H<sub>2N</sub></u> -IQVYSR	110	27	6
VNHVTLSPKIVKWRDMD		110/113		64
VNHVTLSPKIVKWR		110/113		29
<b><math>\Delta</math>N6-<math>\beta_2</math>m dimer</b>				
VNHVTLSPKIVK	IVKWRD	110	113	2
VNHVTLSPKIVK	IVKWRDMD	110	113	4
VNHVTLSPKIVK	<u>H<sub>2N</sub></u> -IQVYSR	110	27	4
VNHVTLSPKIVK	VNHVTLSPKIVK	110	110	13
IEKVEHSDLSFSK	IVKWRDMD	67	113	1
VNHVTLSPKIVKWRD		110/113		40
VNHVTLSPKIVKWRDMD		110/113		44
IEKVEHSDLSFSKDWFSYLLYYTEFTPEKDEYACR		67/77		1
<b><math>\Delta</math>N6-<math>\beta_2</math>m trimer</b>				
VNHVTLSPKIVKWRD	IVKWRD	110	113	1
VNHVTLSPKIVKWRD	IVKWRDMD	110	113	1
VNHVTLSPKIVK	IVKWRD	110	113	2
VNHVTLSPKIVK	IVKWRDMD	110	113	4
VNHVTLSPKIVK	<u>H<sub>2N</sub></u> -IQVYSR	110	27	6
IEKVEHSDLSFSK	VNHVTLSPKIVK	67	110	2
IEKVEHSDLSFSK	IEKVEHSDLSFSK	67	67	1
VNHVTLSPKIVK	VNHVTLSPKIVK	110	110	9
VNHVTLSPKIVKWRD		110/113		23
VNHVTLSPKIVKWRDMD		110/113		24
<b><math>\Delta</math>N6-<math>\beta_2</math>m tetramer</b>				
VNHVTLSPKIVK	IVKWRD	110	113	2
VNHVTLSPKIVK	IVKWRDMD	110	113	2
VNHVTLSPKIVK	<u>H<sub>2N</sub></u> -IQVYSR	110	27	2
IEKVEHSDLSFSK	VNHVTLSPKIVK	67	110	2
VNHVTLSPKIVK	VNHVTLSPKIVK	110	110	7
VNHVTLSPKIVKWRD		110/113		15
VNHVTLSPKIVKWRDMD		110/113		15
<b><math>\beta_2</math>m monomer</b>				
VNHVTLSPKIVK	TPKIQVYSR	110	26	84
VNHVTLSPKIVKWRDMD	TPKIQVYSR	110	26	8
VNHVTLSPKIVKWRD	TPKIQVYSR	110	26	3
VNHVTLSPKIVKWRD	TPKIQVYSR	113	26	2
VNHVTLSPKIVKWRD		110/113		40
VNHVTLSPKIVKWRDMD		110/113		62
<b><math>\beta_2</math>m dimer</b>				
<u>H<sub>2N</sub></u> -IQRTPKIQVYSR	TPKIQVYSR	21	26	3
IQRTPKIQVYSR	TPKIQVYSR	26	26	3
TPKIQVYSR	TPKIQVYSR	26	26	20
IEKVEHSDLSFSK	TPKIQVYSR	67	26	2
IEKVEHSDLSFSK	VNHVTLSPKIVK	67	110	1
VNHVTLSPKIVK	TPKIQVYSR	110	26	71
VNHVTLSPKIVK	VNHVTLSPKIVK	110	110	6
VNHVTLSPKIVK	IVKWRDMD	110	113	1
VNHVTLSPKIVKWRDMD	TPKIQVYSR	110	26	4
IVKWRDMD	TPKIQVYSR	113	26	1
VNHVTLSPKIVKWRD	TPKIQVYSR	113	26	1
VNHVTLSPKIVKWRD	TPKIQVYSR	113	26	3
VNHVTLSPKIVKWRD		110/113		30
VNHVTLSPKIVKWRDMD		110/113		41
<b><math>\beta_2</math>m trimer</b>				
VNHVTLSPKIVK	TPKIQVYSR	110	26	50
VNHVTLSPKIVKWRD	TPKIQVYSR	113	26	3
VNHVTLSPKIVKWRDMD	TPKIQVYSR	110	26	2
TPKIQVYSR	TPKIQVYSR	26	26	9
IQRTPKIQVYSR	TPKIQVYSR	26	26	1
IEKVEHSDLSFSK	VNHVTLSPKIVK	67	110	2
IEKVEHSDLSFSK	TPKIQVYSR	67	26	2
VNHVTLSPKIVK	VNHVTLSPKIVK	110	110	7
VNHVTLSPKIVKWRD		110/113		16
VNHVTLSPKIVKWRDMD		110/113		26
<b><math>\beta_2</math>m tetramer</b>				
VNHVTLSPKIVK	TPKIQVYSR	110	26	30
TPKIQVYSR	TPKIQVYSR	26	26	8
IEKVEHSDLSFSK	TPKIQVYSR	67	26	4
VNHVTLSPKIVK	VNHVTLSPKIVK	110	110	5
VNHVTLSPKIVKWRD		110/113		19
VNHVTLSPKIVKWRDMD		110/113		25

## Assembling Oligomers of $\Delta N6$ - $\beta_2$ -Microglobulin



**FIGURE 4. Workflow for structural characterization of  $\Delta N6$ - $\beta_2$ m tetramer.** *A*, structural data from IM-MS and CX-MS were used to select a suitable starting structure for building the tetramer. The cross-linking data (orange) was consistent with  $\Delta N6$ - $\beta_2$ m monomer NMR (PDB ID 2XKU) and dimer x-ray (PDB ID 2X89) crystal structures. Calculated CCS for the energy-minimized monomer was in good agreement with experimental CCS from IM-MS; however, calculated dimer CCS was larger than observed experimentally. *B*, gas-phase MD simulations were performed on the dimer, and the subsequently calculated CCS was in good agreement with experimental CCS, suggesting subtle compaction in the gas phase. Information from IM-MS and cross-linking MS were combined using a scoring function in an integrative approach to suggest model structures for trimeric  $\Delta N6$ - $\beta_2$ m, starting from the validated monomer and dimer structures from NMR and x-ray crystallography, respectively. Gas-phase MD simulations were performed on the best-scoring trimeric model structure and CCS calculated, showing good agreement with experimental CCS. *C*, similarly, using restraints from IM-MS, cross-linking MS and EM model structures were suggested for tetrameric  $\Delta N6$ - $\beta_2$ m, starting from the validated dimeric structure (1). Docking of the best-scoring tetramer model into the  $\beta_2$ m fibril EM density map (EM Database ID 1613) showed excellent agreement, with a cross-correlation coefficient of 0.77.

with trypsin. The mixture of peptides and cross-linked dipeptides for each band were analyzed by LC-MS/MS. In total, we obtained 138 potential hits after database searching, of which we validated 92 spectra by (i) the presence of peak pairs in the MS spectra (corresponding to  $BS3-d_0$  and  $d_4$ ; Fig. 3B) and (ii) by the quality of tandem-MS spectra, resulting in a false discovery rate of 33.33% (Table 2).

Importantly we obtained dipeptides with the same or overlapping peptide sequences in the dimer, trimer, and tetramer bands, allowing their unambiguous classification as interprotein cross-links (Fig. 3C). The remaining cross-links are difficult to assign to intra- or interprotein cross-links, as the oligomers are formed by monomers with the same amino acid sequence, and their peptide sequences are not unique. We have, therefore, excluded these cross-links from our modeling and only implemented cross-links that were unique and unambiguously related to interprotein interactions.

To assess the structural differences between full-length  $\beta_2$ m and its truncated form, we performed cross-linking experi-

ments on  $\beta_2$ m following the same strategy as above. Bands corresponding to monomer, dimer, trimer, and tetramer of  $\beta_2$ m were cut, proteins were digested, and peptides were analyzed by LC-MS/MS. We obtained 529 potential hits after database searching and validated 322 of these manually (false discovery rate 39.13%). We obtained two unique cross-links from the monomer band and up to nine unique cross-links for the dimer, trimer, or tetramer bands (Table 2). These cross-links were in good agreement with those obtained for  $\Delta N6$ - $\beta_2$ m (Table 2). Due to the longer amino acid sequence in  $\beta_2$ m we identified additional interactions in the N-terminal regions of the protein. The high similarity between the observed cross-links in the oligomers of the two variants suggests highly conserved solution structures and initial assembly mechanism.

**Building Atomic Models of  $\Delta N6$ - $\beta_2$ m Oligomers**—Having established  $\Delta N6$ - $\beta_2$ m interresidue proximities from CX-MS and the overall compact assembly using IM-MS, we turned our attention to identify suitable atomic resolution structures from the protein data bank (PDB) for building further  $\Delta N6$ - $\beta_2$ m olig-



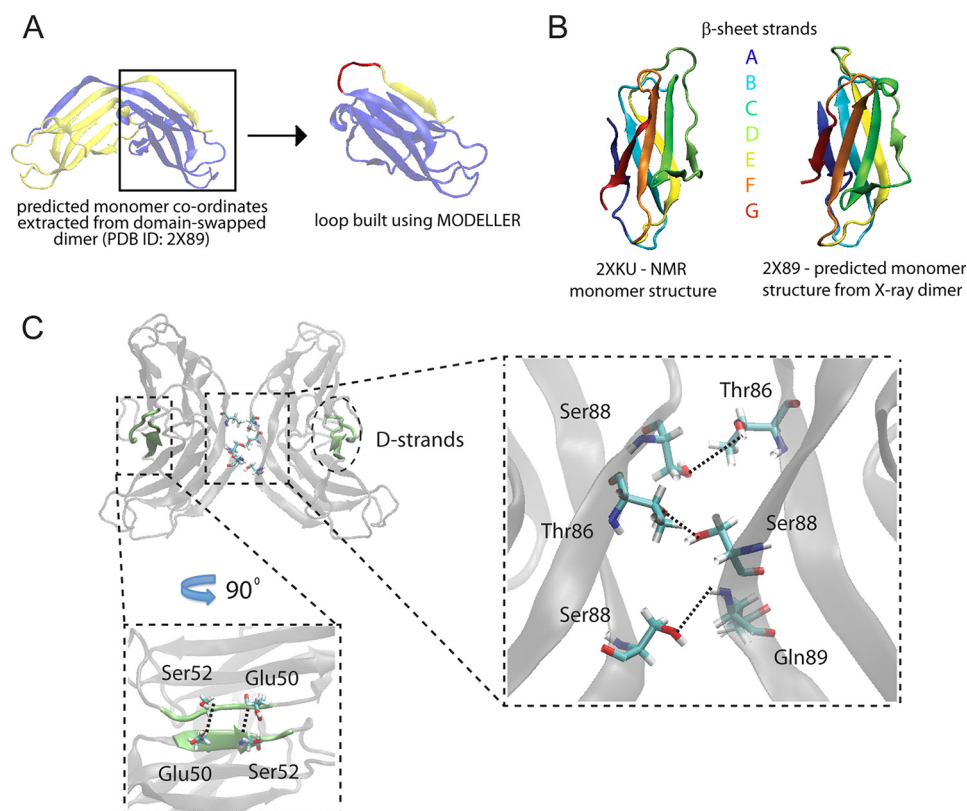


FIGURE 5. **Starting structures and interaction interfaces for tetrameric  $\Delta N6$ - $\beta_2$ m.** *A*,  $\Delta N6$ - $\beta_2$ m monomer was reconstructed from the domain-swapped dimer x-ray crystal structure (PDB ID 2X89) by extracting the coordinates manually and subsequently rebuilding a missing loop using MODELLER. *B*, the reconstructed monomer shows striking similarity with the monomeric  $\Delta N6$ - $\beta_2$ m structure determined by NMR (PDB ID 2XKU). Colored regions show the corresponding  $\beta$ -sheets A to G in the two structures. *C*, model structure of  $\Delta N6$ - $\beta_2$ m tetramer reveals that intradimeric interfaces are formed by the interaction of the domain-swapped hinge regions and on the opposite side, the D strands interact through hydrogen bonding between Glu-50 and Ser-52. The inter-dimeric interface is formed from the interaction of the domain-swapped hinge regions by hydrogen bonding of Ser-88 (chain D), Thr-86, and Ser-88 (chain C) with Thr-86 (chain B), Ser-88, and Gln-89 (chain A). Hydrogen bonds were identified using the HBonds tool in VMD (visual molecular dynamics).

omers. NMR structure (PDB ID 2XKU) for  $\Delta N6$ - $\beta_2$ m monomer (50) was used to compare the calculated monomer CCS with that from IM-MS. These were found to differ by 6.3% ( $CCS_{exp}$   $1200 \pm 36 \text{ \AA}^2$ ;  $CCS_{calc}$   $1276 \text{ \AA}^2$ ). Rearrangements in the gas phase are thought to be responsible for overall compaction of the structure in the absence of solution leading to lower experimental CCS values than anticipated from the crystal structure (51). To account for this possibility, gas-phase molecular dynamics (MD) simulations were performed before CCS calculations (46). The calculated CCS of the simulated structure was in good agreement with the measured CCS ( $CCS_{exp}$   $1200 \pm 36 \text{ \AA}^2$ ;  $CCS_{calc}$   $1144 \text{ \AA}^2$ ; 4.5% deviation). Next, we projected the experimentally identified cross-links (Table 2) onto the atomic structure and measured the  $C\alpha$ - $C\alpha$  distances (Fig. 4A). The calculated distances revealed that the experimentally determined cross-links were in good agreement with the NMR structure using an upper bound interresidue distance threshold of  $35 \text{ \AA}$  (27).

X-ray crystal structure (PDB ID 2X89) is composed of two  $\Delta N6$ - $\beta_2$ m monomers associated through domain-swapping. We compared the calculated CCS from this structure with the measured dimer CCS from IM-MS and found them to differ by 8% ( $CCS_{exp}$   $1900 \pm 57 \text{ \AA}^2$ ;  $CCS_{calc}$   $2064 \text{ \AA}^2$ ). Similar to the monomer, we carried out gas-phase MD simulations and found that the calculated CCS of the simulated dimer was in good agreement with the measured CCS from IM-MS ( $CCS_{exp}$

$1900 \pm 57 \text{ \AA}^2$ ;  $CCS_{calc}$   $1873 \text{ \AA}^2$ ; 1.4% deviation) (Fig. 4B). Structural rearrangement in the gas phase primarily occurred through the compaction of the gross structure as measured by center of mass distances between the two interacting monomers, which decreased from 2.93 to 1.97 nm during the MD simulation. Structural agreement with the experimentally identified cross-links (Table 2) was confirmed by projecting them onto the x-ray crystal structure before and after the MD simulations.

Overall we conclude that both the NMR structure for the  $\Delta N6$ - $\beta_2$ m monomer and the x-ray structure for the dimer are consistent with our experimental data and are themselves structurally similar (Fig. 5, A and B). We, therefore, used these structures as a starting point to build higher order oligomers. Of particular interest is the dimeric structure, as domain-swapping has been proposed as a plausible assembly mechanism for amyloidogenesis (1, 9).

*Integrative Modeling Predicts  $\Delta N6$ - $\beta_2$ m Oligomers at Atomic Resolution*—Having validated starting structures for building higher oligomers, we used an integrative approach to model trimeric and tetrameric  $\Delta N6$ - $\beta_2$ m. To achieve this we computationally integrated our MS data (trimer and tetramer) with available information from EM (tetramer only) using a suitable scoring function (“Experimental Procedures”).

We began by building a model for trimeric  $\Delta N6$ - $\beta_2$ m at atomic resolution. We used as starting structures the domain-



## Assembling Oligomers of $\Delta N6$ - $\beta_2$ -Microglobulin

swapped dimer and the monomer from NMR. We generated 10,000 atomic models using a Monte Carlo sampling of conformational space, and the models generated were evaluated using the scoring function described above. In particular, we scored all models using the experimentally measured CCS ( $CCS_{exp}$  2530  $\text{\AA}^2$ ) and four interdimer cross-links, K110:K113 (19.5  $\text{\AA}$ ) K110:K110 (13.4  $\text{\AA}$ ), K110:K67 (29.3  $\text{\AA}$ ), and K110:K27 (18.0  $\text{\AA}$ ), identified in the trimer band (Table 2). Clustering analysis of the top-scoring 1% models revealed two main clusters (threshold 5  $\text{\AA}$ ). We chose a representative structure of the largest cluster (60%). We finally performed gas phase MD simulations (46) showing a compaction (10%;  $CCS_{calc}$  2640  $\text{\AA}^2$ ) of the trimeric structure in vacuum (Fig. 4B), in good agreement with the measured CCS (4% deviation).

Next, we employed a similar strategy to predict the most likely architecture of the tetrameric  $\Delta N6$ - $\beta_2m$ , starting with two dimers (Fig. 4C). We scored all models using the experimentally measured CCS ( $CCS_{exp}$  3057  $\text{\AA}^2$ ), the three identified interdimer cross-links, K110:K113 (22.3  $\text{\AA}$ ), K110:K110 (20.9  $\text{\AA}$ ) and K110:K67 (32.4  $\text{\AA}$ ), for the tetramer (Table 2), and the section of density map corresponding to the globular tetramer of  $\beta_2m$  (EM Database ID 1613) (11). To reflect the molecular envelope of tetrameric  $\Delta N6$ - $\beta_2m$ , we chose a globular rather than elongated section from the EM map (Fig. 4C), consistent with our IM-MS experiments and coarse-grained modeling (Fig. 2E). As performed on the trimer, the top 1% scoring models for the tetramer were clustered, and the top scoring model in the largest cluster was chosen as the representative model structure (Fig. 4C). Interestingly, in this model the hinge region consisting of heptapeptides NHVTLSSQ, which readily form amyloids in isolation, are stacked together (Fig. 5C). These interact through dimer-dimer interfaces by hydrogen bonding between Ser-88 (D chain), Thr-86 and Ser-88 (C chain) from one dimer with Thr-86, Ser-88, and Gln-89 from a second dimer, respectively (Fig. 5C). Of further interest are the D strands, dynamic regions which are thought to play a role in amyloidogenicity (2, 52). Here, these form the dimer interface on the opposite side to the domain-swapped hinge region (Fig. 5C).

Finally, to assess the stability of the predicted tetramer, we performed solution and gas phase MD simulations (46). These simulations revealed that the tetramer underwent only subtle changes over the simulation timeframe, suggesting a stable conformation (Fig. 6, A and B).

### Discussion

$\Delta N6$ - $\beta_2m$  oligomers have been proposed as intermediate assemblies leading to fibrillogenesis either through nucleation and elongation of their own fibrils or through cross-seeding with full-length  $\beta_2m$  (2).  $\beta_2m$  fibrillogenesis may be seeded by preformed  $\Delta N6$ - $\beta_2m$  filaments or fibrils (53). Alternatively,  $\Delta N6$ - $\beta_2m$  monomers may interact with full-length  $\beta_2m$ , enabling a transition into an aggregation-prone conformation (50). Evidence in support of this, from NMR (50) and IM-MS (18), shows that  $\beta_2m$  monomers undergo a conformational change before oligomerization. This intermediate state is thought to have enhanced amyloidogenic potential (2). A similar intermediate state may be more readily accessible for  $\Delta N6$ - $\beta_2m$ , giving

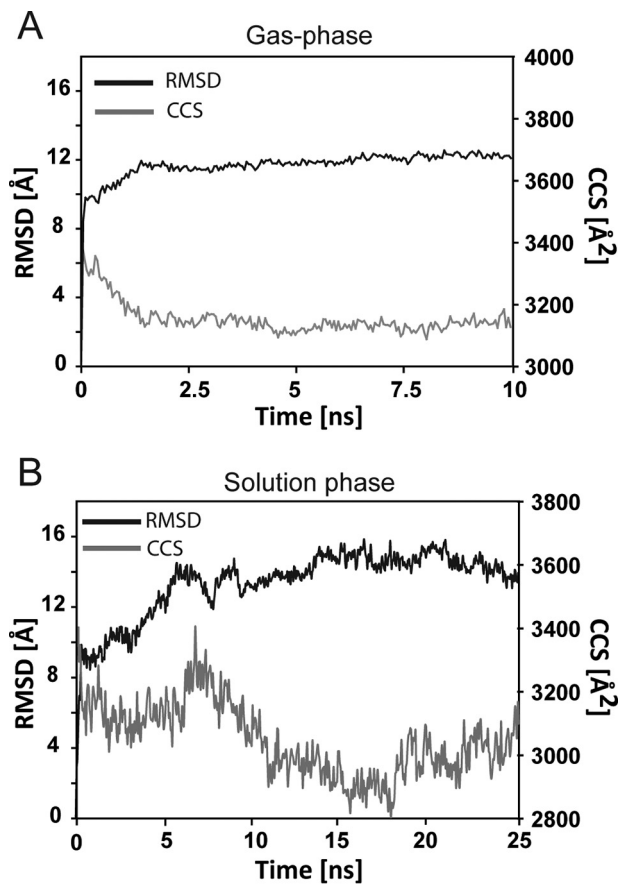
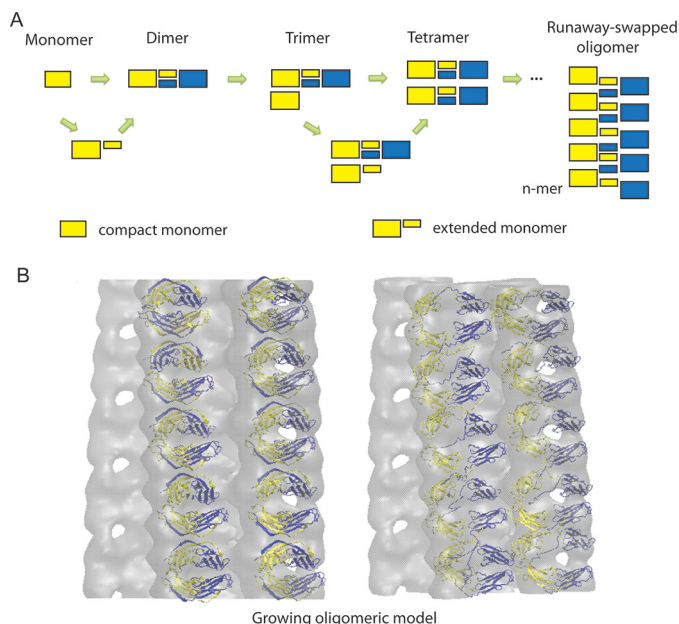


FIGURE 6. MD simulations on  $\Delta N6$ - $\beta_2m$  tetramer. A, gas-phase MD simulations of  $\Delta N6$ - $\beta_2m$  tetramer revealed subtle structural rearrangement measured by a decrease in CCS ( $\sim 5.9\%$ ; gray line) over the course of the simulation (10 ns). B, solution phase MD simulations of  $\Delta N6$ - $\beta_2m$  tetramer revealed an overall stable conformation as measured by the moderate root mean square deviation (RMSD, black line) over 25 ns.

rise to its increased amyloidogenicity compared with full-length  $\beta_2m$  (2, 54, 55). Another possibility is that a proteolytic step may precede aggregation, in which the N-terminal hexapeptide is removed from  $\beta_2m$  and that  $\Delta N6$ - $\beta_2m$  is, therefore, itself an on-pathway intermediate of  $\beta_2m$  fibrillation (54). Although the specific interaction between  $\beta_2m$  and  $\Delta N6$ - $\beta_2m$  remains a topic of intense debate, it is becoming increasingly clear that studying the early oligomeric pathway of  $\Delta N6$ - $\beta_2m$  at atomic level may be essential to understand the assembly mechanism for full-length  $\beta_2m$ .

Here we predict structural models and an early assembly mechanism for oligomers of  $\Delta N6$ - $\beta_2m$  up to tetramer, highlighting their compact topologies and key intersubunit interactions and comparing oligomers of the full-length protein. Because fibrillogenesis has been observed for  $\Delta N6$ - $\beta_2m$  under similar solution conditions to those used in our MS experiments (1), we believe that the observed oligomers serve as pre-amyloid intermediates en route to higher order oligomeric species and fibril formation.

From our comparative IM-MS and CX-MS results, we established that  $\Delta N6$ - $\beta_2m$  and full-length  $\beta_2m$  had similar oligomer profiles, CCS, and identified cross-links. This suggests a high structural similarity for the lower order oligomers of these two variants and points to a conserved initial assembly mechanism



**FIGURE 7. Mechanistic models for growing  $\Delta N6$ - $\beta_2 m$  oligomers.** *A*, proposed assembly mechanism of early oligomers preceding amyloidosis. A stepwise monomeric addition is suggested for oligomerization that likely involves an extended “open” intermediate state before domain swapping. *B*, Multiple copies of tetrameric  $\Delta N6$ - $\beta_2 m$  model were fitted into the  $\beta_2 m$  fibrillar EM map (EM Database ID 1613). The growing oligomer models were built from repeating tetramers, where the domain-swapped hinge region was perpendicular to the fibrillar growth axis (*left*), consistent with the transition of stacked domain-swapped dimers into runaway oligomers (*right*).

before fibrillogenesis. The specific experimental conditions, however, are important for determining the *ex vivo* aggregation pathway followed and resulting fibril morphology (18, 19, 56). Previous studies of full-length  $\beta_2 m$  have shown that at low pH and low ionic strength, long straight fibrils are formed via a nucleated mechanism. On the other hand, at higher pH and ionic strength, worm-like fibrils are formed, with elongation proceeding via a non-nucleated, monomer addition pathway, and a conformational change in monomeric  $\beta_2 m$  thought to be responsible for initiating oligomerization. Interestingly, assembly of  $\beta_2 m$  in the presence of Cu(II) ions proceeds through dimer addition (52), with domain-swapping also implicated (57).

Although the precise assembly mechanism(s) followed *in vivo* are not fully understood, the evidence is growing for a domain-swapped oligomeric intermediate preceding fibril formation (1, 9). By exploiting the power of MS, we put forward a model at the atomic level of resolution, describing the early formation of  $\Delta N6$ - $\beta_2 m$  oligomers, indicating a stepwise assembly mechanism through the addition of monomeric subunits (Fig. 7A). We further mapped the oligomeric growth into the fibrillar EM map to show the conversion of stacked domain- to runaway-swapped oligomers in which the ends of the growing oligomers are capable of binding open monomers by self-templated growth (Fig. 7B). Overall, this study provides novel insights into the early mechanism of pre-amyloid assemblies and further makes inroads toward an atomic level description of  $\beta_2 m$  that entails detailed structural models of intermediate states and their associated transient interactions.

**Author Contributions**—Z. H. and A. P. conceived and designed the study. Z. H. conducted all native MS experiments. C. S. conducted all cross-linking experiments. Z. H. and A. P. performed the modeling. Z. H., C. S., and A. P. analyzed the data and wrote the paper.

**Acknowledgments**—We thank Professor Jan Steyaert and Katarzyna Domanska (Vrije Universiteit Brussels) for the purified samples of  $\Delta N6$ - $\beta_2 m$  and  $\beta_2 m$ . We thank Professor Dame Carol Robinson for generous support. We acknowledge Drs. Antoni Borysik and Eamonn Reading (both King’s College London) for critical reading of the manuscript.

## References

- Domanska, K., Vanderhaegen, S., Srinivasan, V., Pardon, E., Dupeux, F., Marquez, J. A., Giorgetti, S., Stoppini, M., Wyns, L., Bellotti, V., and Steyaert, J. (2011) Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic  $\beta_2$ -microglobulin variant. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1314–1319
- Eichner, T., and Radford, S. E. (2011) Understanding the complex mechanisms of  $\beta_2$ -microglobulin amyloid assembly. *FEBS J.* **278**, 3868–3883
- Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366
- Ashcroft, A. E. (2010) Mass spectrometry and the amyloid problem: how far can we go in the gas phase? *J. Am. Soc. Mass Spectrom.* **21**, 1087–1096
- Williams, D. M., and Pukala, T. L. (2013) Novel insights into protein misfolding diseases revealed by ion mobility-mass spectrometry. *Mass Spectrom. Rev.* **32**, 169–187
- Tipping, K. W., Karamanos, T. K., Jakhria, T., Iadanza, M. G., Goodchild, S. C., Tuma, R., Ranson, N. A., Hewitt, E. W., and Radford, S. E. (2015) pH-induced molecular shedding drives the formation of amyloid fibril-derived oligomers. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5691–5696
- Bleiholder, C., Dupuis, N. F., Wytttenbach, T., and Bowers, M. T. (2011) Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to  $\beta$ -sheet in amyloid fibril formation. *Nat. Chem.* **3**, 172–177
- Carulla, N., Zhou, M., Arimon, M., Gairi, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2009) Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7828–7833
- Liu, C., Sawaya, M. R., and Eisenberg, D. (2011)  $\beta_2$ -microglobulin forms three-dimensional domain-swapped amyloid fibrils with disulfide linkages. *Nat. Struct. Mol. Biol.* **18**, 49–55
- Kad, N. M., Myers, S. L., Smith, D. P., Smith, D. A., Radford, S. E., and Thomson, N. H. (2003) Hierarchical assembly of  $\beta_2$ -microglobulin amyloid *in vitro* revealed by atomic force microscopy. *J. Mol. Biol.* **330**, 785–797
- White, H. E., Hodgkinson, J. L., Jahn, T. R., Cohen-Krausz, S., Gosal, W. S., Müller, S., Orlova, E. V., Radford, S. E., and Saibil, H. R. (2009) Globular tetramers of  $\beta_2$ -microglobulin assemble into elaborate amyloid fibrils. *J. Mol. Biol.* **389**, 48–57
- Carulla, N., Zhou, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2010) Structure and intermolecular dynamics of aggregates populated during amyloid fibril formation studied by hydrogen/deuterium exchange. *Acc. Chem. Res.* **43**, 1072–1079
- Carulla, N., Caddy, G. L., Hall, D. R., Zurdo, J., Gairi, M., Feliz, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2005) Molecular recycling within amyloid fibrils. *Nature* **436**, 554–558
- Stengel, F., Baldwin, A. J., Painter, A. J., Jaya, N., Basha, E., Kay, L. E., Vierling, E., Robinson, C. V., and Benesch, J. L. (2010) Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2007–2012
- Bernstein, S. L., Dupuis, N. F., Lazo, N. D., Wytttenbach, T., Condron, M. M., Bitan, G., Teplow, D. B., Shea, J.-E., Ruotolo, B. T., Robinson, C. V., and Bowers, M. T. (2009) Amyloid- $\beta$  protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer’s

## Assembling Oligomers of $\Delta N6$ - $\beta_2$ -Microglobulin

- disease. *Nat. Chem.* **1**, 326–331
16. Bleiholder, C., Do, T. D., Wu, C., Economou, N. J., Bernstein, S. S., Buratto, S. K., Shea, J. E., and Bowers, M. T. (2013) Ion mobility spectrometry reveals the mechanism of amyloid formation of A $\beta$ (25–35) and its modulation by inhibitors at the molecular level: epigallocatechin gallate and scyllo-inositol. *J. Am. Chem. Soc.* **135**, 16926–16937
  17. Young, L. M., Saunders, J. C., Mahood, R. A., Revill, C. H., Foster, R. J., Tu, L. H., Raleigh, D. P., Radford, S. E., and Ashcroft, A. E. (2015) Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nat. Chem.* **7**, 73–81
  18. Smith, D. P., Radford, S. E., and Ashcroft, A. E. (2010) Elongated oligomers in  $\beta_2$ -microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6794–6798
  19. Smith, D. P., Woods, L. A., Radford, S. E., and Ashcroft, A. E. (2011) Structure and dynamics of oligomeric intermediates in  $\beta_2$ -microglobulin self-assembly. *Biophys. J.* **101**, 1238–1247
  20. Woods, L. A., Radford, S. E., and Ashcroft, A. E. (2013) Advances in ion mobility spectrometry-mass spectrometry reveal key insights into amyloid assembly. *Biochim. Biophys. Acta* **1834**, 1257–1268
  21. Baldwin, A. J., Lioe, H., Hilton, G. R., Baker, L. A., Rubinstein, J. L., Kay, L. E., and Benesch, J. L. (2011) The polydispersity of  $\alpha B$ -crystallin is rationalized by an interconverting polyhedral architecture. *Structure* **19**, 1855–1863
  22. Hall, Z., Politis, A., and Robinson, C. V. (2012) Structural modeling of heteromeric protein complexes from disassembly pathways and ion mobility-mass spectrometry. *Structure* **20**, 1596–1609
  23. Politis, A., Park, A. Y., Hall, Z., Ruotolo, B. T., and Robinson, C. V. (2013) Integrative modelling coupled with ion mobility mass spectrometry reveals structural features of the clamp loader in complex with single-stranded DNA binding protein. *J. Mol. Biol.* **425**, 4790–4801
  24. Politis, A., Park, A. Y., Hyung, S. J., Barsky, D., Ruotolo, B. T., and Robinson, C. V. (2010) Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes. *PLoS ONE* **5**, e12080
  25. van Duijn, E., Barbu, I. M., Barendregt, A., Jore, M. M., Wiedenhof, B., Lundgren, M., Westra, E. R., Brouns, S. J., Doudna, J. A., van der Oost, J., and Heck, A. J. (2012) Native tandem and ion mobility mass spectrometry highlight structural and modular similarities in clustered-regularly interspaced shot-palindromic-repeats (CRISPR)-associated protein complexes from *Escherichia coli* and *Pseudomonas aeruginosa*. *Mol. Cell Proteomics* **11**, 1430–1441
  26. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2007) The molecular architecture of the nuclear pore complex. *Nature* **450**, 695–701
  27. Politis, A., Stengel, F., Hall, Z., Hernández, H., Leitner, A., Walzthoeni, T., Robinson, C. V., and Aebersold, R. (2014) A mass spectrometry-based hybrid method for structural modelling of protein complexes. *Nat. Methods* **11**, 403–406
  28. Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1380–1387
  29. Politis, A., Schmidt, C., Tjioe, E., Sandercock, A. M., Lasker, K., Gordiyenko, Y., Russel, D., Sali, A., and Robinson, C. V. (2015) Topological models of heteromeric protein assemblies from mass spectrometry: application to the yeast eIF3:eIF5 complex. *Chem. Biol.* **22**, 117–128
  30. Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H., Cimermančić, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014) Molecular architecture of the 40SeIF1eIF3 translation initiation complex. *Cell* **158**, 1123–1135
  31. Su, Y., Sarell, C. J., Eddy, M. T., Debelouchina, G. T., Andreas, L. B., Pashley, C. L., Radford, S. E., and Griffin, R. G. (2014) Secondary structure in the core of amyloid fibrils formed from human  $\beta_2m$  and its truncated variant  $\Delta N6$ . *J. Am. Chem. Soc.* **136**, 6313–6325
  32. Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V., and Ruotolo, B. T. (2010) Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Anal. Chem.* **82**, 9557–9565
  33. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858
  34. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021
  35. Xu, H., Hsu, P. H., Zhang, L., Tsai, M. D., and Freitas, M. A. (2010) Database search algorithm for identification of intact cross-links in proteins and peptides using tandem mass spectrometry. *J. Proteome Res.* **9**, 3384–3393
  36. Kahraman, A., Malmström, L., and Aebersold, R. (2011) Xwalk: computing and visualizing distances in cross-linking experiments. *Bioinformatics* **27**, 2163–2164
  37. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* **9**, 1634–1649
  38. Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2**, 715–726
  39. Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S. J., and Robinson, C. V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protoc.* **3**, 1139–1152
  40. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38
  41. Zemla, A. (2003) LGA: A method for finding 3D similarities in protein structures. *Nucleic Acids Res.* **31**, 3370–3374
  42. Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C., and Jarrold, M. F. (1996) Structural information from ion mobility measurements: effects of the long-range potential. *J. Phys. Chem.* **100**, 16082–16086
  43. Shvartsburg, A. A., and Jarrold, M. F. (1996) An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86–91
  44. Benesch, J. L., and Ruotolo, B. T. (2011) Mass spectrometry: come of age for structural and dynamical biology. *Curr. Opin. Struct. Biol.* **21**, 641–649
  45. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447
  46. Hall, Z., Politis, A., Bush, M. F., Smith, L. J., and Robinson, C. V. (2012) Charge-state dependent compaction and dissociation of protein complexes: insights from ion mobility and molecular dynamics. *J. Am. Chem. Soc.* **134**, 3429–3438
  47. Benesch, J. L., Aquilina, J. A., Ruotolo, B. T., Sobott, F., and Robinson, C. V. (2006) Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chem. Biol.* **13**, 597–605
  48. Chakravarty, S., and Varadarajan, R. (1999) Residue depth: a novel parameter for the analysis of protein structure and stability. *Structure Fold. Des.* **7**, 723–732
  49. Lane, L. A., Ruotolo, B. T., Robinson, C. V., Favrin, G., and Benesch, J. L. P. (2009) A Monte Carlo approach for assessing the specificity of protein oligomers observed in nano-electrospray mass spectra. *Int. J. Mass Spectrom.* **283**, 169–177
  50. Eichner, T., Kalverda, A. P., Thompson, G. S., Homans, S. W., and Radford, S. E. (2011) Conformational conversion during amyloid formation at atomic resolution. *Mol. Cell* **41**, 161–172
  51. Jurneczko, E., and Barran, P. E. (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst* **136**, 20–28
  52. Mendoza, V. L., Barón-Rodríguez, M. A., Blanco, C., and Vachet, R. W. (2011) Structural insights into the pre-amyloid tetramer of  $\beta_2$ -microglobulin from covalent labeling and mass spectrometry. *Biochemistry* **50**, 6711–6722
  53. Mangione, P. P., Esposito, G., Relini, A., Raimondi, S., Porcari, R., Giorgetti, S., Corazza, A., Fogolari, F., Penco, A., Goto, Y., Lee, Y. H., Yagi, H., Ceccconi, C., Naqvi, M. M., Gillmore, J. D., Hawkins, P. N., Chiti, F., Ro-



- landi, R., Taylor, G. W., Pepys, M. B., Stoppini, M., and Bellotti, V. (2013) Structure, folding dynamics, and amyloidogenesis of D76N  $\beta_2$ -microglobulin: roles of shear flow, hydrophobic surfaces, and  $\alpha$ -crystallin. *J. Biol. Chem.* **288**, 30917–30930
54. Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernández, H., Robinson, C. V., Amoresano, A., Dal Piaz, F., Monti, M., Pucci, P., Mangione, P., Stoppini, M., Merlini, G., Ferri, G., and Bellotti, V. (2000) Removal of the N-terminal hexapeptide from human  $\beta_2$ -microglobulin facilitates protein aggregation and fibril formation. *Protein Sci.* **9**, 831–845
55. Estácio, S. G., Krobath, H., Vila-Viçosa, D., Machuqueiro, M., Shakhnovich, E. I., and Faisca, P. F. (2014) A simulated intermediate state for folding and aggregation provides insights into  $\Delta$ N6  $\beta_2$ -microglobulin amyloidogenic behavior. *PLoS Comput. Biol.* **10**, e1003606
56. Gosal, W. S., Morten, I. J., Hewitt, E. W., Smith, D. A., Thomson, N. H., and Radford, S. E. (2005) Competing pathways determine fibril morphology in the self-assembly of  $\beta_2$ -microglobulin into amyloid. *J. Mol. Biol.* **351**, 850–864
57. Eakin, C. M., Attenello, F. J., Morgan, C. J., and Miranker, A. D. (2004) Oligomeric assembly of native-like precursors precedes amyloid formation by  $\beta_2$  microglobulin. *Biochemistry* **43**, 7808–7815