

OPEN

# Enhanced oligomerization of full-length RAGE by synergy of the interaction of its domains

Alexander Moysa <sup>1\*</sup>, Dietmar Hammerschmid<sup>2</sup>, Roman H. Szczepanowski<sup>3</sup>, Frank Sobott<sup>2,4</sup> & Michal Dadlez<sup>1</sup>

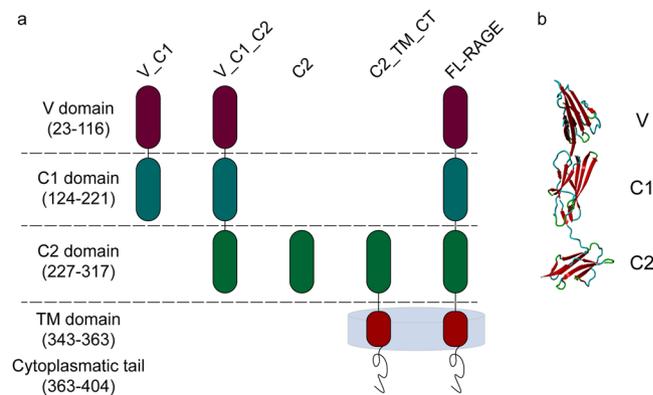
The pattern recognition receptor RAGE (receptor for advanced glycation end-products) transmits proinflammatory signals in several inflammation-related pathological states, including vascular diseases, cancer, neurodegeneration and diabetes. Its oligomerization is believed to be important in signal transduction, but RAGE oligomeric structures and stoichiometries remain unclear. Different oligomerization modes have been proposed in studies involving different truncated versions of the extracellular parts of RAGE. Here, we provide basic characterization of the oligomerization patterns of full-length RAGE (including the transmembrane (TM) and cytosolic regions) and compare the results with oligomerization modes of its four truncated fragments. For this purpose, we used native mass spectrometry, analytical ultracentrifugation, and size-exclusion chromatography coupled with multi-angle light scattering. Our results confirm known oligomerization tendencies of separate domains and highlight the enhanced oligomerization properties of full-length RAGE. Mutational analyses within the GxxxG motif of the TM region show sensitivity of oligomeric distributions to the TM sequence. Using hydrogen–deuterium exchange, we mapped regions involved in TM-dependent RAGE oligomerization. Our data provide experimental evidence for the major role of the C2 and TM domains in oligomerization, underscoring synergy among different oligomerization contact regions along the RAGE sequence. These results also explain the variability of obtained oligomerization modes in RAGE fragments.

Oligomerization is common among proteins, and two thirds of cellular proteins form oligomers<sup>1</sup>. Oligomer formation often seems crucial for membrane protein interaction with ligands and signal transduction<sup>2–6</sup>. Oligomerization is important for different types of receptors, such as immunoreceptors<sup>7</sup> and T cell<sup>8</sup>, B cell<sup>9</sup>, and hormone receptors<sup>10</sup>. The receptor for advanced glycation end-products (RAGE) is a member of the immunoglobulin (Ig) superfamily of cell-surface receptors<sup>11</sup>. Apart from advanced glycation end-products, RAGE ligands include a variety of proteins, peptides, and nucleic acid chains<sup>12–16</sup>. RAGE signalling plays an important role in the adaptive immune response. As a crucial mediator of proinflammatory signalling, however, RAGE activation may also enhance a variety of pathophysiological states, such as arteriosclerosis, diabetes, and Alzheimer's disease<sup>17</sup>.

The role of RAGE in a variety of pathologies makes it a potential target for therapeutic intervention. However, like other membrane proteins, RAGE presents a significant challenge in structural studies. Its inherent oligomerization tendencies, the dynamic nature of its domains and the need for detergents or lipids to maintain protein solubility render RAGE intractable to many biophysical techniques. In contrast, native mass spectrometry (Native MS) and hydrogen–deuterium exchange monitored by MS (HDX-MS) allow acquisition of structural information for oligomerizing proteins and protein–protein/ligand interactions in the presence of membrane-mimicking detergents or lipids<sup>18</sup>.

RAGE contains three extracellular domains – a V-type Ig domain (residues 23–116) and two C-type Ig domains (C1 and C2) (residues 124–221 and residues 227–317, respectively), a transmembrane (TM) helix (residues 343–363), and a short cytoplasmic tail (residues 363–404) (Fig. 1a). Each extracellular domain of RAGE contains two cysteines forming intramolecular disulphide bridges within a domain<sup>19–21</sup>. Cysteines of the C2 region may also form intermolecular bridges, stabilizing a RAGE dimer<sup>14,19</sup>. Structural studies of the extracellular part of

<sup>1</sup>Institute of Biochemistry and Biophysics, PAN, Pawinskiego 5a, 02-109, Warsaw, Poland. <sup>2</sup>Biomolecular & Analytical Mass Spectrometry, University of Antwerp, Groenenborgerlaan 171, 2020, Antwerp, Belgium. <sup>3</sup>International Institute of Molecular and Cell Biology, Trojdena 4, 02-109, Warsaw, Poland. <sup>4</sup>Astbury Centre for Structural Molecular Biology and School of Molecular and Cellular Biology, University of Leeds, Woodhouse Lane, LS2 9JT, Leeds, UK. \*email: [a.alexandrmoya@gmail.com](mailto:a.alexandrmoya@gmail.com)



**Figure 1.** (a) Domain composition of full-length RAGE and different truncated variants used in this work. RAGE consists of the N-terminal V domain followed by two C-domains (C1, C2) of immunoglobulin-like fold, a transmembrane region (TM), and a cytoplasmic tail (CT). (b) X-ray structure of V\_C1\_C2 4YBH<sup>16</sup>.

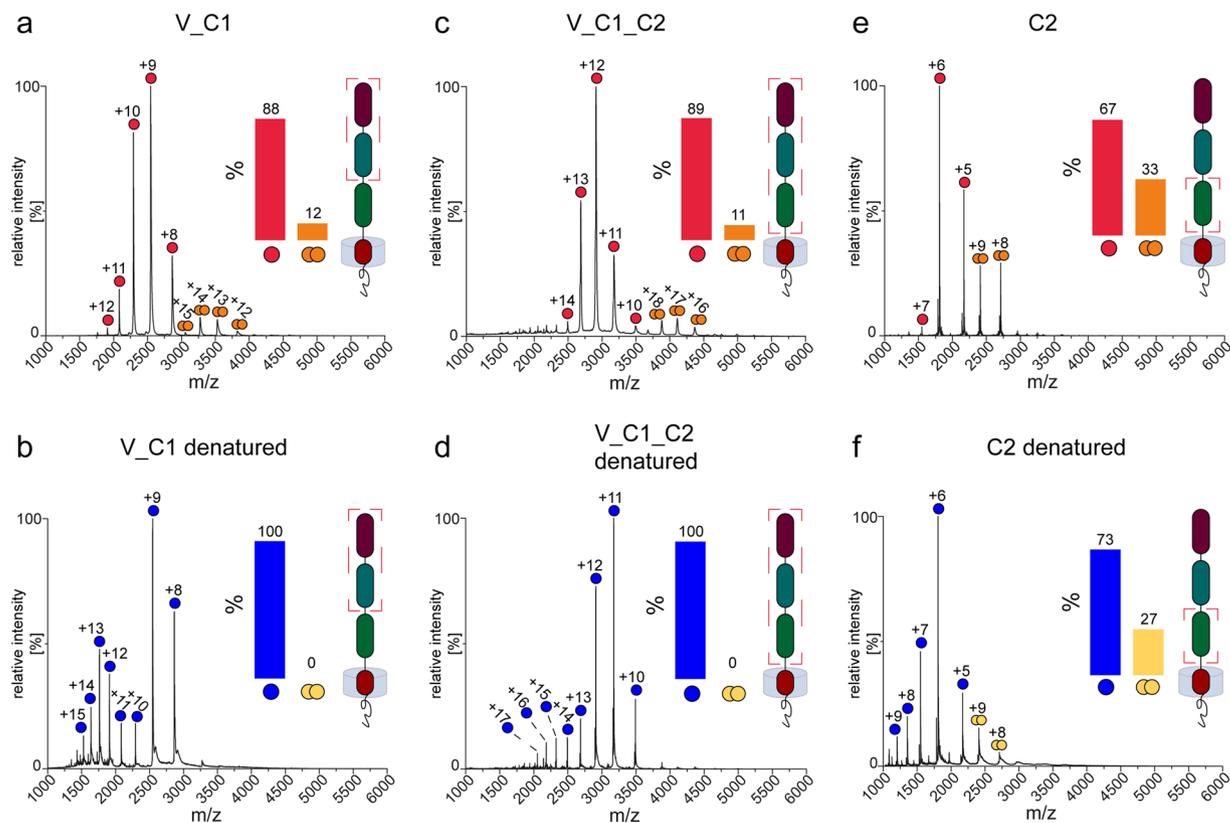
RAGE (V\_C1\_C2) have shown that the V and C1 domains form an elongated structural unit connected with the C2 domain by a flexible linker (Fig. 1b)<sup>16,22</sup>.

Intrinsic oligomerization tendencies have hampered the classic approaches to elucidating the structure of this membrane-embedded protein complex in its full-length version. Instead, truncated variants, covering different regions of the extracellular part of RAGE and missing the TM region have been used. For instance, X-ray structures of isolated domains V\_C1<sup>13</sup> and V\_C1\_C2<sup>22</sup> have been used as a basis for molecular modelling of the complex. Covalently stabilized oligomeric forms of subdomains also have been employed, cross-linked either by an engineered di-tyrosine bond<sup>23</sup> or by standard chemical cross-linking in the V\_C1\_C2 construct, followed by molecular modelling<sup>20</sup>. In result of this work different modes of oligomerization were proposed, without consensus. Because electrostatic charges are unevenly distributed between the domains, pI values in different constructs are widely disparate (V\_C1, pI = 9.9; V\_C1\_C2, pI = 7.8; C2, pI = 4.2). These differences may lead to a variety of arrangements when domains are studied in separation. Therefore, attempts to verify the obtained models using full-length RAGE that includes all potential interacting regions are crucial for revealing the native pairing and stoichiometry of the RAGE oligomers.

## Results

**RAGE constructs.** To examine the role of different human RAGE domains in oligomerization, we compared the oligomerization patterns in full-length RAGE (FL\_RAGE), covering the three extracellular domains, the TM region, and the intracellular tail with its four, differently truncated variants (V\_C1, V\_C1\_C2, C2, and C2\_TM\_CT) (Fig. 1a). The two C-terminally truncated variants, V\_C1 and V\_C1\_C2, contained the extracellular domains of RAGE. C2\_TM\_CT represented the N-terminally truncated variant, containing the C2, TM, and cytoplasmic tail, while C2 contained only the extracellular domain C2. Proteins were obtained by overexpression in *E. coli* and subjected to *in vitro* studies of their oligomerization status and structural dynamics by native MS, analytical ultracentrifugation (AUC), size-exclusion chromatography (SEC) coupled with multi-angle light scattering (MALS) and HDX-MS.

**Disulphide bonds in FL\_RAGE.** Disulphide linkages significantly influence protein structure and function. Incorrectly paired disulphide bonds trigger alteration in the three-dimensional structure followed by a resultant change in protein properties. Therefore, obtaining native disulphide bonds is a crucial moment in protein production (see “Protein production” section in Supplementary information). It has been established before in numerous studies<sup>13,14,19–21</sup> that cysteines within V and C1 domains form only intradomain disulphide bonds while cysteines within the C2 domain can form intra- and interprotein disulphide bridges. Both variants of linkages between cysteines of C2 domains were observed in our study for C2 and C2\_TM\_CT samples (Figs. 2f and S4). In contrast, for the V\_C1 protein, where the C2 domain was omitted, our results revealed only non-covalent dimers (Fig. 2a,b) which is in agreement with the previous results<sup>13,21,24</sup>. Apparently, the tendency of cysteine linkages between C2 domains caused the formation of high-order disulphide mediated oligomers of the full-length RAGE and V\_C1\_C2 during *E. coli* cytoplasmic expression (Fig. S2c). In order to overcome this problem, we used a periplasmic variant of *E. coli* expression, where intraprotein disulphide bonds were formed predominantly. For confirmation of the correct FL\_RAGE disulphide bond formation, we used a MS-based strategy (Supplementary Fig. S1). Based on previous structural studies of extracellular domains, three intradomain disulphide bonds (linking residues 38–99, 144–208, and 259–301) were expected<sup>14,20,21</sup>. Our results (Supplementary Table S1) confirmed the correct disulphide bond pairing in FL\_RAGE. For V\_C1\_C2 protein, we used the same production protocol as for the FL\_RAGE. To further confirm the structural identity of the V\_C1\_C2 used here and obtained before, we have verified that the hydrogen-deuterium exchange pattern for V\_C1\_C2 protein was the same (Fig. S6a) as demonstrated in a previous study<sup>23</sup>. Therefore, we conclude a proper formation of the expected structure of V\_C1\_C2 protein necessitating the correct formation of disulphide bridges.



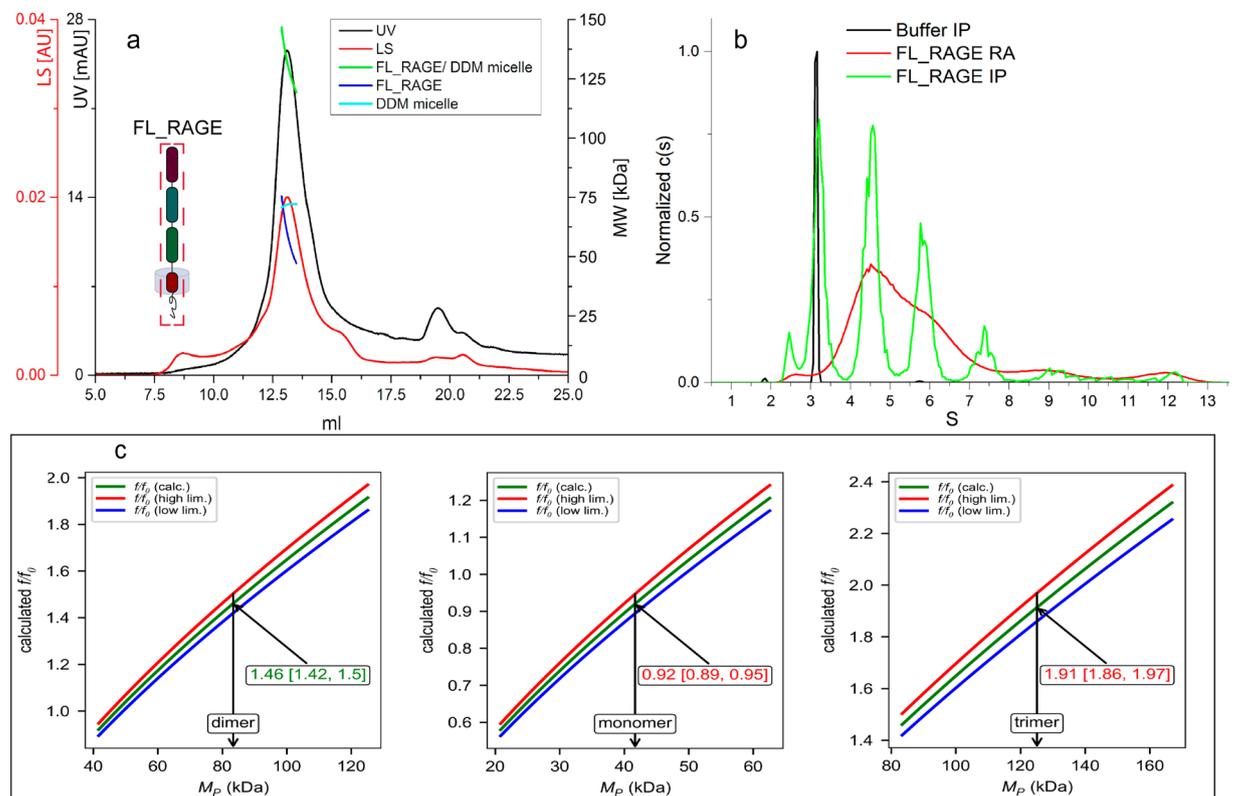
**Figure 2.** Mass spectra of V\_C1 (a,b), V\_C1\_C2 (c,d), and C2 (e,f) introduced into the mass spectrometer under non-denaturing conditions (a,c,e) and denaturing conditions (b,d,f). Monomeric and dimeric species including their corresponding charge state are assigned accordingly for the different protein variants. The bar diagrams in the inset show the relative amount of monomers and dimers, respectively. The illustration highlights the composition of each construct in relation to the full-length protein by a red frame.

## Oligomerization States of FL\_RAGE

**SEC-MALS.** In order to maintain FL\_RAGE in solution, we applied a detergent environment capable of mimicking the native lipid environment in which the TM region is normally embedded. Therefore, FL\_RAGE was purified and characterized in the presence of the non-ionic detergent n-Dodecyl  $\beta$ -D-maltoside (DDM). The SEC-MALS analysis of FL\_RAGE in the presence of DDM revealed a well-defined peak (Fig. 3a). In the case of the protein–detergent complex, however, the interpretation of the corresponding masses is not straightforward. The amount of detergent bound to the protein may differ depending on both the protein structure and the nature of the detergent. Therefore, the mass ratio between the protein and detergent in the micelle can vary dramatically. A dedicated approach is required in the data analysis step, which considers the presence of the detergent<sup>25</sup>. The “protein conjugate” module of ASTRA software (Wyatt Technologies) enables the calculation of the molecular weight of the whole complex as well as its protein component and the detergent micelle mass associated with the protein.

The SEC-MALS chromatogram of FL\_RAGE (Fig. 3a) showed a main protein peak eluting at ~13 ml. The molar mass of the complex, estimated by ASTRA, was heterogeneous across the peak, spanning the range of masses of the protein–detergent micelle complex (~120–140 kDa). The calculated molar mass of the protein was also heterogeneous (~47–75 kDa), indicating a substantial non-monomeric fraction in the samples of FL\_RAGE (the expected molecular mass of FL\_RAGE is 41.8 kDa).

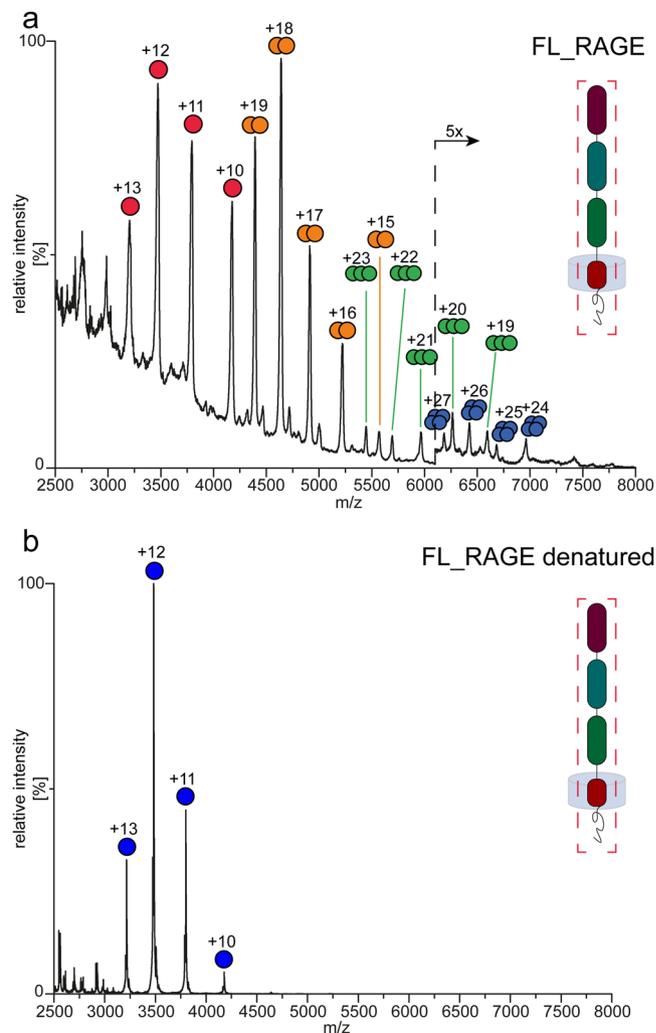
**AUC.** The sedimentation velocity of FL\_RAGE (Fig. 3b) was investigated by monitoring the absorbance and interference data. The absorbance at 280 nm displayed a broad signal between 3.5S and 8.0S, with the maximum around 4.4S and a trace amount of higher molecular weight species. At the same time, the interference, characterized by a better spatial resolution as compared to the absorption, showed three well-defined peaks (4.4S, 5.8S, and 7.4S) located in the same range of 3.5–8.0S. The additional peak at 3.2S clearly corresponded to the interference boundary observed for the buffer. This peak was not detected when the buffer alone was investigated with absorbance (data not shown) indicating the presence of free DDM. Similar signal was also observed by other authors<sup>26</sup>. Interpreting these findings in terms of the oligomerization status of FL\_RAGE required a more detailed analysis. For this purpose, we used the  $f/f_0$  analysis procedure described in the GUSSE Membrane Protein Module package by Brautigam<sup>27</sup>. Following this, the results indicated that the 4.4S, 5.8S, and 7.4S peaks corresponded to monomer, dimer, and trimer, respectively, as only in this case the  $f/f_0$  values were within the allowed range of 1.2–1.7. The assumption of another degree of oligomerization led to  $f/f_0$  values deviated from the permissible range: >1.9 and



**Figure 3.** SEC-MALS and AUC analysis of full-length RAGE. **(a)** SEC-MALS of FL\_RAGE. Chromatograms show the detector readings of the LS and UV detectors in red and black, respectively. Scales for LS and UV detectors are shown on the left-hand axis; scale on right-hand axis represents the molecular mass. The green, blue, and cyan lines across the peaks indicate calculated molecular masses of the protein/detergent complex (green), the protein (blue), and the attached detergent micelle (cyan). **(b)** Sedimentation velocity experiment of DDM-solubilized FL\_RAGE (absorption at 280 nm in red and interference in green). The black trace represents the interference data of the buffer including DDM micelles. **(c)** Example of the analysis of the dataset of FL\_RAGE (panel b) using type  $f/f_0$  GUSSE Membrane Protein Calculation module. Red, green, and blue lines represent the high, optimal, and low limits of  $f/f_0$  for the protein/micelle fraction, respectively (see text). The pictures show the analysis of the 5.8S peak and indicate the presence of a dimer. The 4.4S and 7.4S peaks represent monomers and trimers, respectively (see Fig. S5 in Supplementary information).

$<1$  for higher order and smaller oligomers, respectively, which is not possible. Figure 2c illustrates this type of analysis based on peak number 2 (5.8S). These results of the AUC experiment revealed the presence of a mixture of monomers, dimers, trimers and even a small fraction of higher order oligomers, which is in good agreement with the SEC-MALS result. Obviously, a rapid exchange between monomeric and dimeric species resulted in a broad peak at 280 nm.

**Native MS.** For native MS analysis, FL\_RAGE was solubilized in DDM or in Triton X-100 at a concentration of two times the critical micelle concentration (CMC). Despite the stability of FL\_RAGE in the presence of DDM during protein purification, the buffer exchange to an MS-friendly ammonium acetate (AmAc) solution, which also contained 2x CMC of DDM, led to protein precipitation, even at a quite high concentration of AmAc (1M). The precipitation of FL\_RAGE was not observed when Triton X-100 (2x CMC) was used instead of DDM. Following this, the native MS spectrum acquired at 600 mM AmAc revealed monomeric, dimeric, trimeric, and even tetrameric signals (Fig. 4a). Both monomeric and dimeric species are close to equilibrium and accounted for the vast majority of the signal. Interestingly, the purification of FL\_RAGE using SEC instead of ion exchange chromatography enabled a better preservation of higher-order oligomers, e.g. trimers and tetramers (Fig. S3), with however some impurities (unidentified contaminants) and a more challenging membrane protein liberation in the gas phase (poor resolution). The addition of 0.1% formic acid and 50% acetonitrile to FL\_RAGE led to the disappearance of signals stemming from oligomeric species (Fig. 4b), indicating a non-covalent character of the oligomerization. As native MS of membrane proteins depends on a collision-activated liberation of the protein from the detergent micelle cluster, we required a minimum of (15V) 30V collision energy along with an activation of 200V in the sampling cone in order to observe (monomeric) oligomeric species of FL-RAGE in the spectrum.



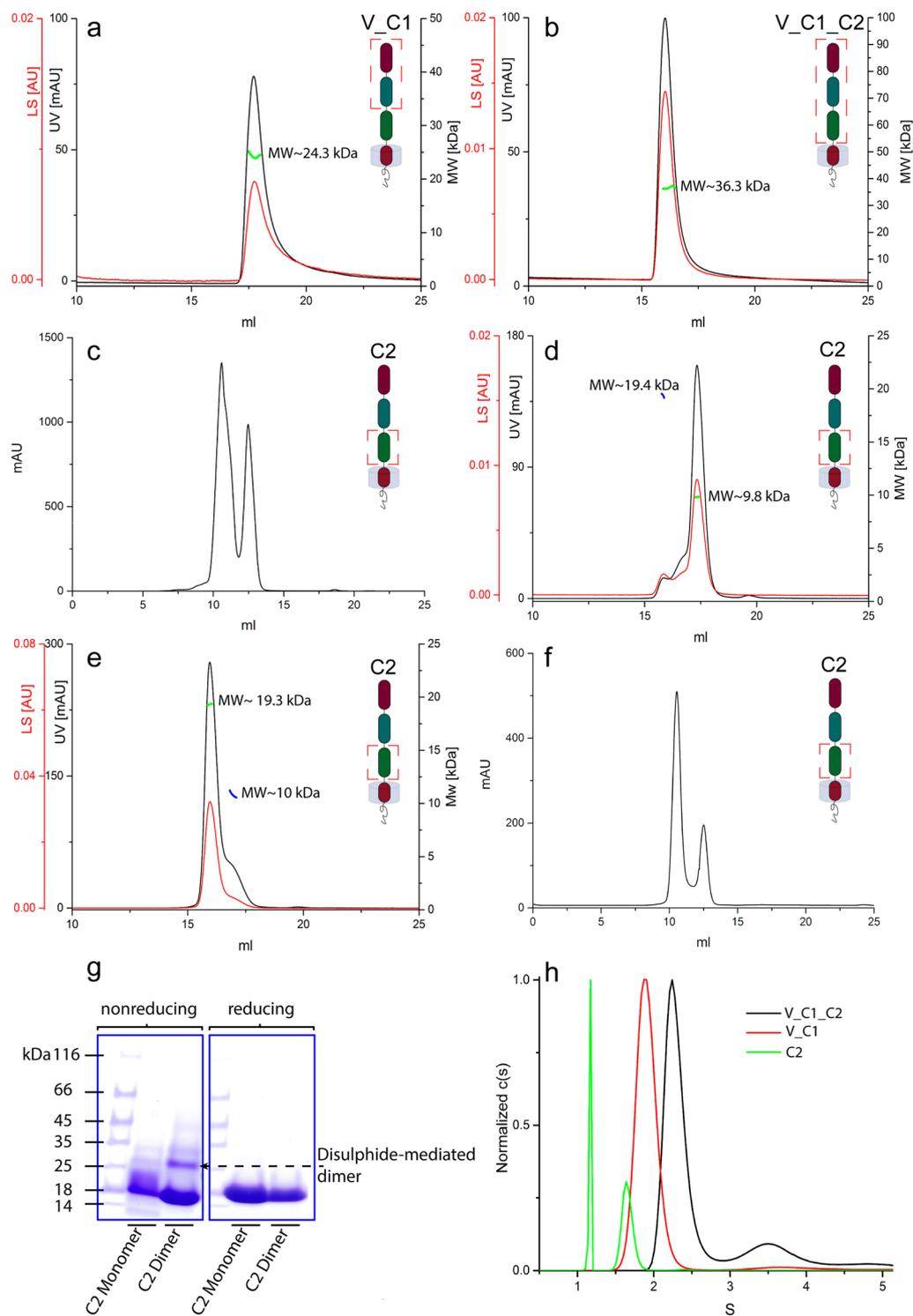
**Figure 4.** Mass spectra of FL\_RAGE under native (a) and (b) denaturing conditions. Peaks are marked by colored dots and the corresponding charge state for monomers and multimers of FL\_RAGE accordingly.

### Oligomerization of truncated variants of RAGE: V\_C1, V\_C1\_C2, and C2

**SEC-MALS.** The elution profiles of V\_C1 and V\_C1\_C2 in SEC-MALS showed homogeneous peaks (Fig. 5a,b). The calculated molecular masses for V\_C1 and V\_C1\_C2 revealed values of 24.3 and 36.3 kDa, respectively. These values are in close proximity to the molecular weight of monomeric V\_C1 ( $MW_{\text{theoret}} = 22.8$  kDa) and V\_C1\_C2 ( $MW_{\text{theoret}} = 34.9$  kDa). The elution profile of C2 displayed two peaks during SEC purification (Fig. 5c). Both fractions were collected and re-chromatographed separately for the determination of the molecular mass by MALS. The calculated Mw of ~10 and ~19 kDa corresponded to the monomer ( $MW_{\text{theoret}} = 10.7$  kDa) and dimer ( $MW_{\text{theoret}} = 21.4$  kDa) of C2, respectively (Fig. 5d,e). A rechromatography of the dimer fraction showed a limited monomeric content (Fig. 5e), indicating that the dimer of C2 was stable.

As the C2 variant contains two cysteines, we considered the possibility of a covalently formed dimer. In order to verify the type of dimerization, we performed SEC in the presence of a reducing agent (Fig. 5f), which still showed a reasonable fraction of dimers. The SDS-PAGE analysis of the dimeric fraction (Fig. 5g) under non-reducing conditions showed a minor dimeric fraction, confirming the disulphide-mediated, non-covalent character of the dimer. These results underscore the strong intrinsic oligomerization properties of the C2 domain.

**AUC.** The sedimentation velocity of V\_C1, V\_C1\_C2, and C2 was investigated by monitoring the absorbance at 280 nm. In the case of V\_C1\_C2, the  $c(s)$  distribution showed two peaks: a main peak (72% of the signal) at 2.3S and a minor peak (17% of the signal) at 3.58S corresponding to the monomer and dimer, respectively (Fig. 5h). The experimental-derived molecular masses of 32.4 kDa and 62.7 kDa were in good agreement with the theoretical sequence-based masses, namely 34.9 kDa for the monomer and 69.8 kDa for the dimer. Trace amounts of even higher molecular weights were observed with however less than 1.5% of the total signal area. This result is in excellent agreement with previous AUC results<sup>28</sup>. The contradicting SEC-MALS result – only a monomeric fraction was present – may be the result of a dilution effect during SEC. V\_C1 only displayed a monomeric peak at 1.91S (93% of the total signal) (Fig. 5h), with excellent agreement of the experimental and theoretical masses: 22.5 kDa versus 22.9 kDa, respectively.



**Figure 5.** SEC, SEC-MALS, and AUC analysis of the constructs V\_C1, V\_C1\_C2, and C2. (a–f) SEC/SEC-MALS chromatograms show the detector readings of the LS and UV in red and black, respectively. Scales for LS and UV detectors are shown in the left-hand axis; and the calculated molecular mass on the right-hand axis. (a) SEC-MALS of V\_C1. (b) SEC-MALS of V\_C1\_C2. (c) SEC of C2. (d) SEC-MALS of the monomeric fraction of C2. (e) SEC-MALS of the dimeric fraction of C2. (f) SEC of C2 upon the addition of a reducing agent (1 mM TCEP). (g) SDS-PAGE of the monomeric and dimeric fractions of C2 under reducing and non-reducing conditions (full-length gels see Fig. S2 in Supplementary information). (h) AUC (sedimentation velocity profiles) of V\_C1, V\_C1\_C2, and C2.

The standard  $c(s)$  distribution of C2 revealed two main peaks (Fig. 5b). In contrast to previous experiments, however, the resulting masses did not agree with the theoretical values calculated from the amino acid sequence. In order to solve this problem, a Sedfit “Continuous  $c(s)$  with bimodal  $f/f_0$ ” procedure was used. During the calculations, we considered a change in the coefficient  $f/f_0$  from 1.39 in the case of the monomer to 1.65 in the case of the dimer, which enabled a good match between experimental and theoretical masses: 9.7 kDa for the monomeric peak ( $MW_{\text{theoret}} = 10.7$  kDa) and 21.1 kDa for the dimeric peak ( $MW_{\text{theoret}} = 21.4$  kDa). This finding suggests that dimerization resulted in the elongation of protein shape. The AUC of C2 showed the prevalence of a dimer (over 65% of the total signal), which is in accordance with the SEC result (Fig. 5c).

**Native MS.** We also conducted native MS experiments for V\_C1, V\_C1\_C2, and C2. The spectrum of V\_C1 (Fig. 2a) revealed the prevalence of monomeric species accompanied by prominent dimeric signals (~12%). V\_C1\_C2 showed approximately the same fraction of dimers (~11%) (Fig. 2c), with well-resolved charge state series corresponding to both monomeric (predominant) and dimeric states (Fig. 2c). The spectrum of C2 (Fig. 2e) was also dominated by monomeric signals, but the dimeric signals was much higher in intensity (~33%) compared to V\_C1 and V\_C1\_C2. This pattern suggested a relatively higher tendency for the formation of C2 dimers, which is in accordance with the results obtained from AUC and SEC experiments (Fig. 5).

As all variants contain cysteine residues, we also conducted MS experiments under denaturing conditions to verify the non-covalent character of the oligomerization. Upon the addition of denaturants (50% acetonitrile and 0.1% formic acid), which causes the disruption of non-covalent interactions, the dimeric signals in V\_C1 (Fig. 2b) and V\_C1\_C2 (Fig. 2d) disappeared entirely. Moreover, the intensity of peaks corresponding to highly charged monomers increased, indicating the denaturation of the proteins and highlighting the non-covalent character of the dimerization in these constructs. The spectrum of C2 (Fig. 2f) under denaturing conditions however still exhibited a significant fraction (~27%) of dimeric peaks. A covalent dimerization in the C2 domain, as observed in Fig. 5g, is possible, although the MS data do not allow a clear statement in this respect. Overall, native MS experiments confirmed increasing oligomerization tendencies for the C2 domain in comparison with V\_C1 or V\_C1\_C2 as well as for FL\_RAGE in comparison with V\_C1 or V\_C1\_C2. Hence, the C2 and TM domains seems to play an important role in terms of the oligomerization of RAGE. Native oligomerization should thus be assessed in the context of the full-length protein, as truncated constructs tend to distort native oligomeric patterns.

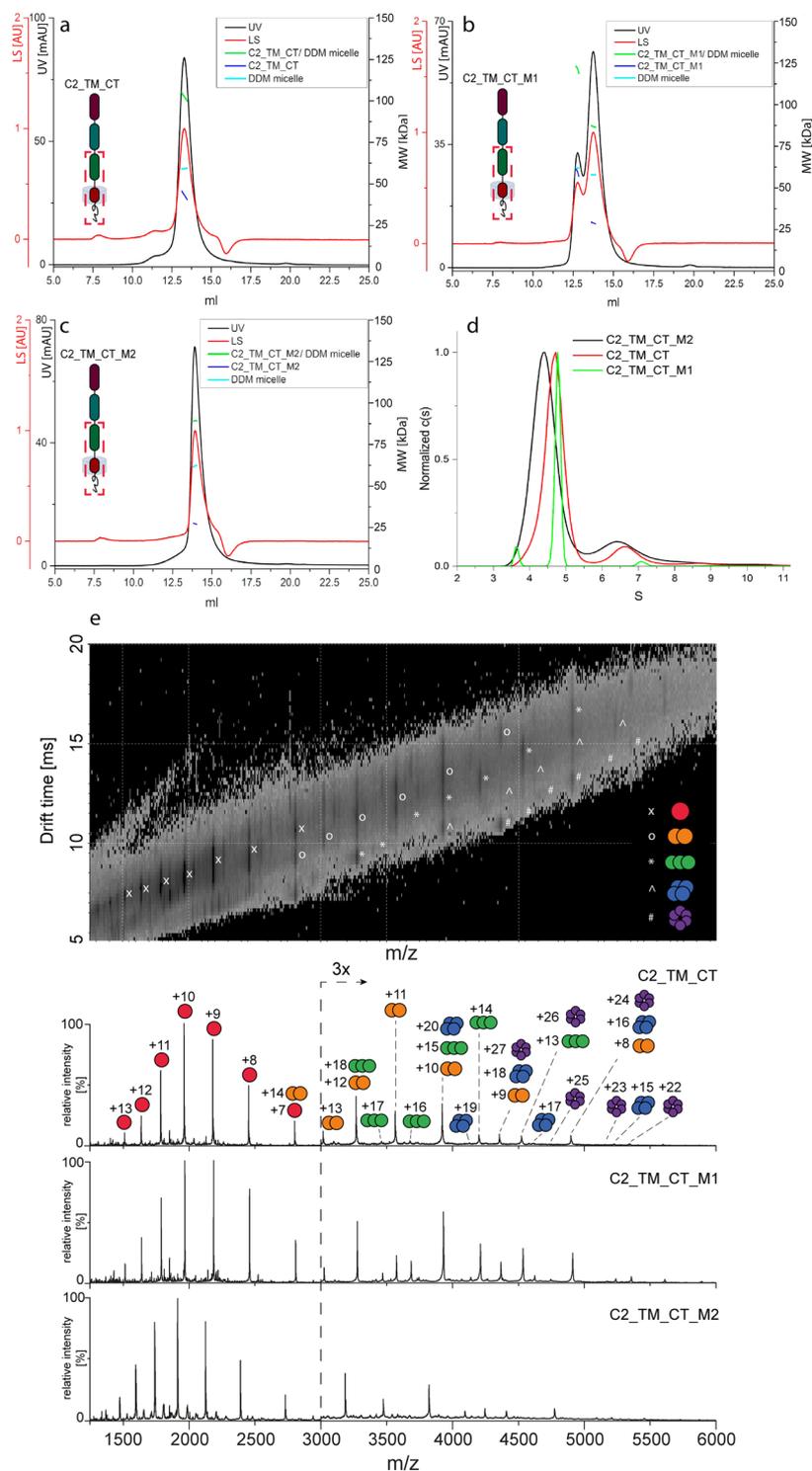
### Oligomerization of the N-truncated variant of FL\_RAGE (C2\_TM\_CT)

**SEC-MALS.** For the entire C-terminal part of RAGE (C2\_TM\_CT), SEC-MALS was performed in the presence of DDM and the data were analyzed according to the “protein conjugate” method (Fig. 6a). The chromatogram displayed a protein peak at ~13.3 ml. The ASTRA analysis revealed an average molecular mass of the protein of 43 kDa, indicating to the formation of a dimer (monomeric mass: 19.6 kDa).

**Native MS.** Native MS analysis of C2\_TM\_CT was performed in the presence of DDM at 2x CMC and 200 mM AmAc (Fig. 6e). The spectrum highlighted monomers, dimers, trimers, and even small amounts of tetramers and hexamers. Since mass spectrometry measures the  $m/z$  ratio of the analyte, different oligomers can result in the same  $m/z$  values which may cause partial overlap in the spectrum. In the case of overlapping oligomers, ion mobility measurements can help to identify the underlying species (Fig. 6e), though the analysis of peak intensities remains difficult. Nevertheless, the fact that higher-order oligomers (i.e. trimers, tetramers, and hexamers) were only observed for C2\_TM\_CT but not for V\_C1, V\_C1\_C2, and C2 - all species were acquired at the same concentration - suggests an increased oligomerization tendency in the TM domain of the protein. However, based on intensities of the oligomeric signals in native MS of FL-RAGE and C2\_TM\_CT, it is difficult to compare these results directly, as the collision-activated liberation depends on the size of the complex as well as the detergent used for solubilization, and both varied in these constructs. Therefore, the higher intensity of FL\_RAGE oligomeric signals may demonstrate an enhanced oligomerization of FL\_RAGE. On the other hand, lower intensity of the oligomeric signals in C2\_TM\_CT may originate from the dissociation of a substantial fraction of oligomers, indicated by a relatively broader charge state distribution for monomers of C2\_TM\_CT (Fig. 6e) compared to FL\_RAGE (Fig. 4a). The spectrum of C2\_TM\_CT (Fig. S4) under denaturing conditions exhibited not only dimeric but also tetrameric and hexameric species. These higher-order oligomers (tetramer and hexamer) might originate from strong non-covalent interactions - membrane proteins are generally known as more challenging to denature<sup>29</sup> - or from interprotein disulphide bonds.

**Mutational analysis of the TM domain.** As previous experiments indicated a direct involvement of the TM region in the oligomerization of RAGE, we attempted to confirm this hypothesis by performing mutational analyses. The presence of a GxxxG motif (two glycines separated by any three amino acids) in the TM region (LGTLALALGILGGLGTAALLIGVILW) is an indicator of possible TM helix-helix interactions<sup>30,31</sup>. The GxxxG motif mediates the packing of TM alpha helices by permitting close association and thus allowing interactions between the helix backbones as well as the side chains of surrounding residues<sup>32</sup>. Therefore, we directed our attention to this region and designed two variants of C2\_TM\_CT with mutations affecting the GxxxG motif. In the first variant, four glycines in the region of the GxxxG motif were mutated to alanine (Mutant 1). In the second variant, six amino acids that cover the GxxxG motif were deleted (Mutant 2) (Supplementary Table S2).

The SEC-MALS experiments indicated the influence of mutations on the oligomerization pattern of both mutants. The chromatograms showed an increased retention volume (13.8 ml) of the major peaks of C2\_TM\_CT\_M1 and C2\_TM\_CT\_M2 as compared to C2\_TM\_CT (13.3 ml), indicating lower molecular masses of the corresponding peaks (Fig. 6a-c). The calculated average molecular masses of the protein, detergent micelle, and protein-detergent micelle complex were at 27, 62, and 89 kDa, respectively, and in close proximity between both



**Figure 6.** SEC-MALS, AUC and native MS analysis of C2\_TM\_CT wild-type and mutants. **(a)** SEC-MALS of C2\_TM\_CT. **(b)** SEC-MALS of C2\_TM\_CT\_M1. **(c)** SEC-MALS of C2\_TM\_CT\_M2. Chromatograms show the detector readings of the LS and UV in red and black, respectively. Scales for LS and UV detectors are shown in the left-hand axis; and calculated molecular mass on the right-hand axis. **(d)** AUC (sedimentation velocity profiles) of C2\_TM\_CT and mutants. **(e)** Native mass spectra of C2\_TM\_CT wild-type and the mutant variants (1 & 2) solubilized in DDM micelles with the 2D drift plot of the wild-type in the top panel. In the case of overlapping oligomeric species, the drift plot can help to identify the correct oligomeric states that are represented by the same m/z peak. Monomeric and oligomeric species including their charge states are assigned accordingly. Corresponding peaks of C2\_TM\_CT\_M2 are slightly shifted to lower m/z values due to the deletion of six amino acids.

mutants. The obtained molecular mass of the protein fits neither to the monomer ( $MW_{\text{theoret}} = 19.6$  kDa, Mutant 1; 19.1 kDa, Mutant 2) nor to the dimer. Therefore, it is plausible that this peak contained a mixture between a monomer and dimer in fast equilibrium, resulting in an experimental molecular mass of 27 kDa. In addition to the main peak (13.8 ml), the glycine-to-alanine substitution in Mutant 1 resulted in an additional peak at 12.7 ml, with a calculated molecular mass of 59 kDa, which corresponds to the trimer ( $MW_{\text{prot/det}} = 121$  kDa;  $MW_{\text{det}} = 61$  kDa).

AUC (Fig. 6d) also showed differences in the sedimentation of C2\_TM\_CT and its variants. The highest value of  $c(s) = 4.8S$  was observed for the main peak of C2\_TM\_CT\_M1. This peak contained 87.4% of the total signal (280 nm) and was surrounded by two other minor peaks, each with less than 7% of the signal. The C2\_TM\_CT was represented by a slightly slower boundary at 4.7S (85.0% of the signal) and the second one at 6.7S (11.9% of the signal). The  $c(s)$  distribution of C2\_TM\_CT\_M2 revealed a main peak with the lowest  $c(s)$  at 4.4S (80.1% of the total signal) and a second, much smaller peak at 6.8S (18.1% of the signal). The lower  $c(s)$  of C2\_TM\_CT\_M2 indicated, on average, a lower oligomerization compared to C2\_TM\_CT and C2\_TM\_CT\_M1. Apparently, the G/A substitution in C2\_TM\_CT\_M1 resulted in a slightly enhanced oligomerization, as the main peak was also narrower than those of C2\_TM\_CT and C2\_TM\_CT\_M2.

The two C2\_TM\_CT mutants were also subjected to native MS experiments. In comparison with wild-type C2\_TM\_CT, the native MS measurements displayed an increased intensity of oligomeric signals for C2\_TM\_CT\_M1, while wild-type and C2\_TM\_CT\_M2 did not differ significantly (Fig. 6e). Overall, AUC, SEC, and native MS analysis of the two mutants indicated that the G/A substitutions might enhance the oligomerization, while the deletion of the GxxxG motif decreases it. Oligomeric distributions proved to be sensitive to substitutions in the TM region, confirming its role in oligomerization.

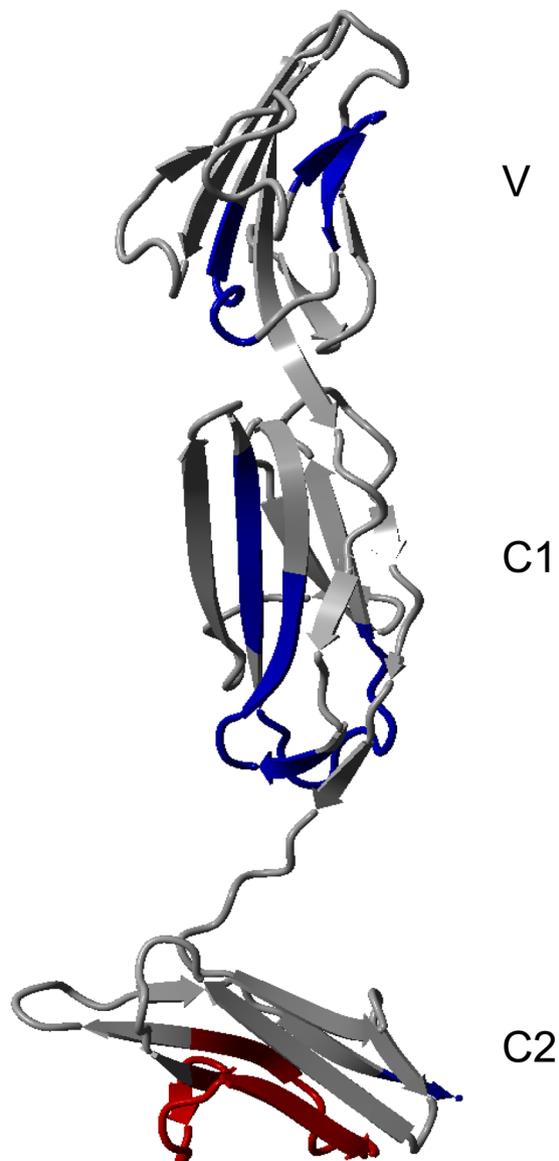
**Amide H/D exchange and differences in dynamics between V\_C1\_C2 and FL\_RAGE.** In order to investigate structural differences between V\_C1\_C2 and FL\_RAGE, we measured the H/D uptake in both proteins. As of the enhanced oligomerization in FL\_RAGE, we expected changes in structural dynamics upon the presence of the TM domain, revealing protein regions involved in the TM-mediated oligomerization. Therefore, we compared the exchange in a set of peptic peptides derived from V\_C1\_C2 and FL\_RAGE. Overall, 71 peptides with satisfactory signal-to-noise ratio and a 78% coverage of the V\_C1\_C2 sequence were reproducibly identified after brief pepsin digestion (Supplementary Table S3). The intertwined pattern of protected and exposed structural elements of V\_C1\_C2 – as observed somewhere else<sup>23,33</sup> and also in this work (Fig. S6a, blue symbols) – was fully retained in FL\_RAGE (Fig. S6a, red symbols). The observed differences between V\_C1\_C2 and FL\_RAGE are visualized in the differential plot, covering the extracellular part of the protein (Fig. S6b). Of interest, only local and no global changes in the protection levels were observed.

The characterization of FL\_RAGE indicated three regions in the V and C1 domain, which became stabilized (Fig. 7). Of note, these regions form the structural core of the protein and showed already strong protection in V\_C1\_C2, and in the case of FL\_RAGE, became even more protected. The first region (R<sup>77</sup>VLPNGSLFLPAVG<sup>90</sup>) in the V domain contains six hydrophobic key residues (marked bold), which has been indicated as the oligomerization interface of the V-V domain (see Fig. 3a in<sup>34</sup>). Similarly, Yatime and Andersen<sup>22</sup> designated the amino acids Leu79, Pro80, and Pro87 as important for the formation of a hydrophobic pocket that takes part in the stabilization of V\_C1\_C2 oligomers. Thus, V-domain oligomerization seems to be enhanced in higher order oligomers. The other two regions, which showed increased protection in FL\_RAGE over V\_C1\_C2, were identified in the C1 domain, namely L<sup>133</sup>TAGVPNKVGTC<sup>144</sup> and F<sup>186</sup>TLQSEL<sup>192</sup>. Although these two peptides are distant in protein sequence, they are in close proximity to each other in the crystal structure of V\_C1\_C2 (Fig. 7), forming the  $\beta$ -strands S11, S12, and S17a on the surface of a  $\beta$ -sheet, distal to the V domain and near to the flexible C1-C2 linker. Of interest, S11 has previously been found to undergo stabilization during dimer formation, obtained by C-terminal tyrosine crosslinking in exRAGE<sup>23</sup>. This evidence indicates that the C1 region participates in the formation of both dimers and higher-order oligomers. This three strands also face each other in the C1-C1 interface of the structural model of the heparin sulphate-stabilized hexamer (see Fig. 4a,b in<sup>34</sup>).

Moreover, we observed regions deprotected in FL\_RAGE relative to V\_C1\_C2. In the C2 domain, we observed an increased rate of deuterium uptake for the proline-rich region W<sup>271</sup>MKDGVPPLPLPSPVL<sup>286</sup>, in which the residues LPL form a  $\gamma$ -turn. This finding indicates a significant destabilization in a region of the C2 domain, possibly resulting from an allosteric effect caused by the oligomerization of FL\_RAGE. Moreover, this proline-rich region has been shown to form a loop on the external surface of the hexameric model mentioned above (Fig. 5d in<sup>34</sup>).

Changes in deuterium uptake were spread along the entire sequence, yet predominantly observed in the V-C1 region. Intermolecular-mediated contacts in the TM domain seemed to shift the oligomeric equilibria towards higher-order oligomers, resulting in a stabilization of regions that have previously been indicated as oligomerization contact sites. Their synergistic action seems to stabilize higher-order oligomers, as observed, for instance, in the heparan-stabilized hexamer structure<sup>34</sup>. The TM domain itself was also highly protected in FL\_RAGE, unsurprisingly, as its amide hydrogens form a system of interacting intramembrane helices which are surrounded by detergent.

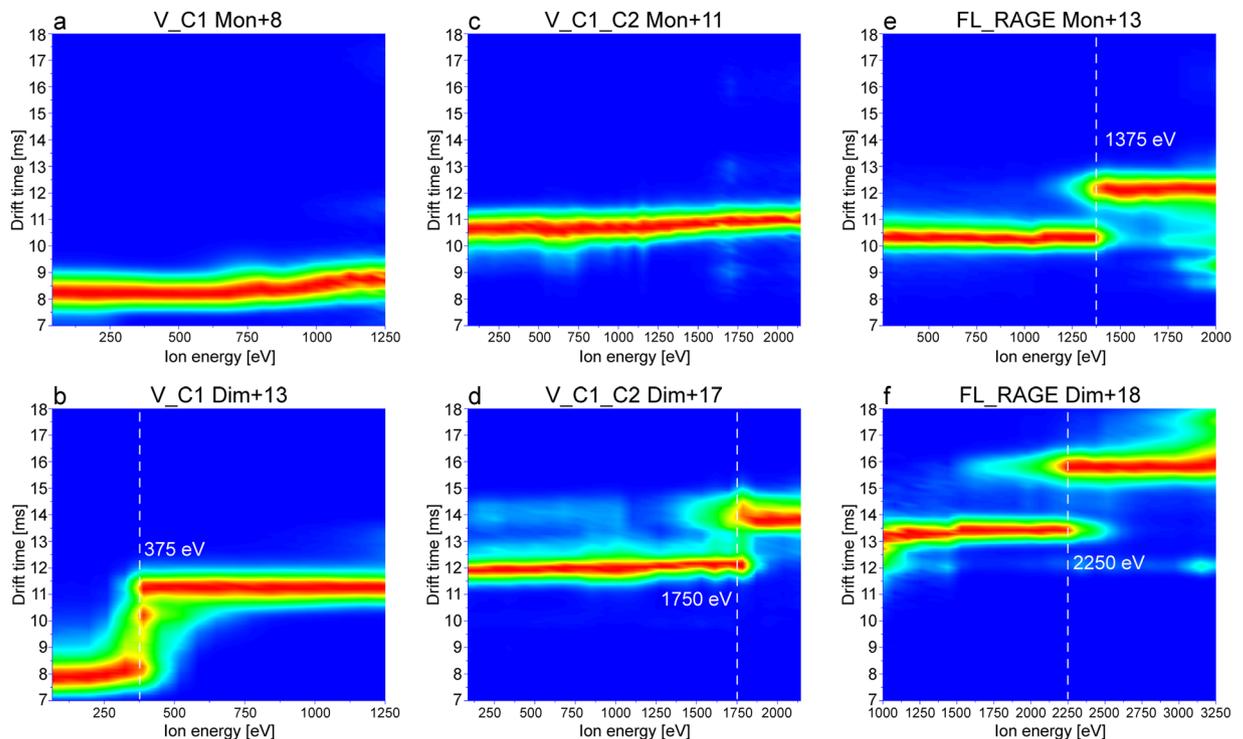
**Collision-induced unfolding (CIU) of V\_C1, V\_C1\_C2, and FL\_RAGE.** Recording ion mobility (IM)-MS along with a stepwise increase of collision energy, known as collision-induced unfolding (CIU)<sup>35</sup>, provides a useful tool for the characterization of the fold stability and domain architecture of proteins, as well as the stability of their dimerization interface. For CIU, protein ions are subjected to increasing activation energies in the collision cell prior to IM-MS separation, which enables the observation of gas-phase unfolding events. We performed such experiments for the constructs V\_C1, V\_C1\_C2, and FL\_RAGE in order to investigate the effect of different domains on the conformational stability and dimerization interface of the proteins. Figure 8 summarizes



**Figure 7.** H/D exchange differences between FL\_RAGE and V\_C1\_C2 in blue (protection increased in FL\_RAGE) and red (protection decreased in FL\_RAGE) are depicted on the X-ray structure of V\_C1\_C2 4YBH<sup>16</sup>. The differential plot of deuterium uptake degrees of peptides is in Supplementary materials Fig. S6.

the obtained CIU fingerprints for monomeric and dimeric species of V\_C1, V\_C1\_C2, and FL\_RAGE, where the ion energy, defined as the ion's charge state multiplied by the collision energy, is plotted against the drift time of the ions. Of interest, the monomers of V\_C1 (Fig. 8a) and V\_C1\_C2 (Fig. 8c) did not show any noticeable unfolding, suggesting a quite stable conformation. The dimers of V\_C1 (Fig. 8b) and V\_C1\_C2 (Fig. 8d), however, exhibited unfolding but with a much higher activation energy required for V\_C1\_C2 (1,750 eV) compared to V\_C1 (375 eV).

Investigating micelle-embedded membrane protein complexes (e.g. FL\_RAGE) with native MS requires part of the activation energy for the stripping of detergent molecules in order to release the membrane protein from this hydrophobic micellar environment<sup>36</sup>. Thus, only the remaining part of the collision energy is available for CIU. In the case of FL\_RAGE, the release of the dimer required ca. 30 V, i.e. 540 eV based on the +18 dimer, of collision energy in addition to 200 V of activation using the sampling cone. Taking these conditions for the liberation of FL\_RAGE into account, we can conclude that the thresholds for unfolding of both V\_C1\_C2 dimer (1,750 eV; Fig. 8d) and FL\_RAGE dimer (2,250 eV; Fig. 8f) are relatively similar and much higher compared to the dimeric V\_C1 (375 eV; Fig. 8b), indicating a considerable effect of the C-terminal C2 and TM domain on the stability of the dimer. The comparison of CIU plots of monomeric V\_C1, V\_C1\_C2 and FL\_RAGE highlights unfolding exclusively in the case of FL\_RAGE at 1,375 eV (Fig. 8e), which suggests structural changes in the TM and CT regions that may weaken the stability of the fold.



**Figure 8.** CIU plots of monomeric (a,c,e) and dimeric (b,d,f) species of the constructs V\_C1, V\_C1\_C2, and FL\_RAGE. The monomeric species of V\_C1 (+8; m/z range 2290–2291) and V\_C1\_C2 (+11; m/z range 3174.3–3175.3) did not show unfolding in this energy range, while the dimeric species showed unfolding at 375 eV (+13; m/z range 3522–3523) and 1750 eV (+17; m/z range 4113–4115), respectively. FL\_RAGE exhibited unfolding at 1375 eV for the monomer (+13; m/z range 3215–3217) and 2250 eV for the dimer (+18; m/z range 4642–4644). The relatively higher ion energy necessary for the unfolding of V\_C1\_C2 and FL\_RAGE as compared to V\_C1 indicates a stabilizing effect of the C2 and TM domain in the protein.

## Discussion

Oligomerization has been implicated in numerous studies as a RAGE receptor signal transduction mechanism<sup>19,20,22,34</sup>. *In vitro*, structural aspects of oligomerization have been studied mainly for different variants of the extracellular parts of RAGE<sup>13,16,20,22,23,34,37</sup>, whereas the impact of TM and cytoplasmic tail have not been assessed. However, in many cases, oligomerization tendencies may be encoded in the TM parts. Mimicking the native properties of the target protein complex requires including this region. Here, we report a basic analysis of full-length RAGE oligomerization in comparison with its different truncated variants. We show major differences in oligomerization properties between full-length RAGE and separated extracellular RAGE domains. The analysis of truncated RAGE variants indicated that the C-terminal parts of the protein are more prone of forming oligomers as compared to the N-terminal domains. The results indicate that both the C2 domain and the TM region are crucial for the formation of integrated, intrinsically oligomeric native structural units of the receptor, building the flexible scaffold necessary for interaction with a variety ligands, and triggering ligand-dependent signal transduction.

Oligomerization of membrane receptors may be induced by ligands but can be constitutive. Apparently, RAGE belongs to the second group<sup>38,39</sup>, although the formation of the protein–ligand complex may further affect the structure and stoichiometry of the receptor complex. Structural studies of different truncated extracellular RAGE variants in the absence of ligands have shown different oligomerization modes in which different domains in different arrangements mediate self-association. Yatime and Andersen<sup>22</sup> proposed V-V domain dimerization based on analysis of the X-ray structure of the extracellular part of RAGE (V\_C1\_C2). A similar type of oligomerization through the V domain has also been observed in solution<sup>40</sup>. A different mode of dimerization, with side-by-side contacts between V\_C1 molecules and mainly involving residues in the C1 domain, was found in the X-ray structure of V\_C1 protein<sup>13</sup>. In contrast to these V-V and C1-C1 modes of oligomerization, an NMR-based study of the soluble form of exRAGE (V\_C1\_C2) demonstrated an asymmetric type of dimer formation, mediated by both V\_C1 and C2 interaction<sup>20</sup>. The evaluation of the oligomeric status of V\_C1\_C2 expressed in mammalian cells was carried out by native PAGE, which showed a prevalence of oligomers for V\_C1\_C2<sup>19</sup>. This observation contrasts with results obtained for V\_C1\_C2 *in vitro* that predominantly formed a monomer in solution<sup>23,28,34,40</sup>.

In the presence of ligands, diverse arrangements of the domains have also been found. Results of Small Angle X-ray Scattering in combination with crystallographic studies revealed, for instance, that in the presence of heparan sulphate, V\_C1\_C2 exists as a hexamer<sup>34</sup>. In this model of the configuration of the receptor complex, positively charged V domains interacting with negatively charged heparan sulphate form the central part of a wheel-like structure, in which the C1 and C2 domains form the spokes and rim of the wheel, respectively.

Therefore, all domains can participate in the interaction network, which is in agreement with current results. Studies of complexes between RAGE and proteins of the S100 family have shown different models of RAGE oligomerization, mediated by ligand binding. X-ray structure analysis of the interaction between the V\_C1\_C2 and S100A6<sup>16</sup> revealed that the homodimer of S100A6 binds two molecules of V\_C1\_C2, mainly through the C1 domain and a linker between C1 and C2 domains. Of interest, a completely different model of the RAGE S100 complex was proposed for the RAGE:S100B interaction<sup>20</sup> in which the S100B homodimer bridged two preformed, asymmetrical RAGE homodimers. Taken together, these results indicate a multimodal type of RAGE oligomerization, required for recognition of different ligands<sup>22</sup>. Ligand-induced transitions between different oligomeric modes may be important in the regulation of ligand-specific signal transduction.

The role of the TM domain upon oligomerization of RAGE has not previously been characterized using high-resolution techniques. Yatime and Andersen<sup>22</sup> suggested an involvement of the TM region in RAGE oligomerization. Thus, the TM region likely supplements or modifies the oligomerization of the extracellular parts of the protein<sup>41</sup>. Mimicking the native state requires this region to be taken into account. Studies of the oligomerization of RAGE using cross-linking of overexpressed FL\_RAGE in mammalian cells have shown that the protein may exist as monomers or higher-order oligomers, although the dimeric state appears to be the prevalent oligomeric form<sup>19,42</sup>. These results agree with the observation that RAGE can form a disulphide-mediated dimer in mammalian cells<sup>19</sup>. In such a case, the structural unit of the RAGE receptor would be the disulphide-mediated dimer, where cysteines of the C2 domain form intermolecular bridges. Indeed, our data confirm that such a scenario is possible for C2 and C2\_TM\_CT (observed by non-reducing SDS-PAGE), in which we have observed a fraction of disulphide-linked dimers (Figs. 5g, 2f and S2b). In contrast, V\_C1, V\_C1\_C2 (Fig. 2b,d), and FL\_RAGE (Fig. 4b) did not reveal covalent dimers after non-reducing denaturation. C2 and C2\_TM\_CT were produced by cytoplasmic expression in *E. coli* where numerous reductases and reducing agents, such as reduced glutathione, prevent oxidation of cysteines and therefore the formation of disulphide bridges occurs later during protein purification in neutral/alkaline buffers. The C2 domain has the propensity to form a stable non-covalent dimer, in which intermolecular disulphides may be thermodynamically favored over intramolecular ones. V\_C1 was also obtained by cytoplasmic expression in *E. coli*, however, in contrast to C2 and C2\_TM\_CT, this variant did not form a disulphide-mediated dimer. This lack of dimerization may be explained by its weaker oligomerization tendencies promoting intra- rather than inter-domain disulphides. Cytoplasmic expression of V\_C1\_C2 and FL\_RAGE resulted in disulphide-mediated aggregates (Supplementary Fig. S2). Using periplasmic *E. coli* expression, described earlier for V\_C1\_C2<sup>28</sup>, allowed us to overcome this issue, leading to the formation of only intramolecular disulphide bonds<sup>43</sup> (Supplementary Table S1).

It has become possible, for the first time, to observe the oligomerization of the full-length protein. We obtained a native MS spectrum of FL\_RAGE, acquired at a relatively high AmAc concentration (600 mM), that displays abundant oligomeric signals up to tetramers. In general, high ionic strength buffers are expected to weaken electrostatically driven oligomer formation<sup>44,45</sup>. Thus, the fact that oligomers of FL\_RAGE remain despite the relatively high AmAc concentration suggests a significant hydrophobically driven oligomerization in the TM domain of the protein. Besides that, interactions between extracellular domains, also of hydrophobic nature, may help to further stabilize the oligomeric complexes, which is indicated by native MS (Figs. 4a and 6), HDX (Fig. 7) and CIU (Fig. 8). In addition to hydrophobic effects, synergistic electrostatic interactions influence the oligomerization of RAGE.

## Conclusion

Our data showed differences in oligomerization tendencies for different variants of RAGE and revealed that oligomerization interfaces are spread along the RAGE sequence. The strongest oligomerization was observed in the C-terminal part of the protein, with significant involvement of the TM region. This finding suggested that both the C2 and TM domain are major factors for the stabilization of RAGE oligomers. At the same time, the increased protection in the V and C1 domain, in the context of full-length RAGE, as well as the dimer formation observed for V\_C1 and V\_C1\_C2, indicate that the V and C1 domain also participate, at different levels, in oligomerization and that each domain plays a specific role.

## Materials and Methods

**Protein expression and purification.** The production of recombinant proteins and mutagenesis are described in Supplementary information.

**Disulphide bond identification.** Disulphide bond identification is described in Supplementary information.

**SEC-MALS analysis.** The SEC-MALS setup consisted of an Agilent HPLC 1260 (including degasser, quaternary pump, autosampler, column holder, and UV-VIS diode array detector) in line with a DAWN-HELEOS multi-angle laser light-scattering detector and Optilab-T-rEX relative differential interferometer (Wyatt Technology). Samples were monitored at wavelengths of 280, 254, and 215 nm. One hundred microliters of protein samples were loaded onto a Superdex 200 Increase 10/300 column (GE Healthcare) equilibrated with buffer (25 mM Tris-Cl, 0.2 M NaCl, pH 8.0). Samples were run at room temperature at a flow rate of 0.5 ml/min. Agilent software was used to control the HPLC, and Wyatt Astra software was used for data collection and analysis. The results were analysed with the ASTRA software (Wyatt Technology) in accordance with the manufacturer's instructions. For FL\_RAGE and C2\_TM\_CT proteins, a buffer containing two times the critical micellar concentration (CMC) DDM was used. Because the TM domain of FL\_RAGE and C2\_TM\_CT was surrounded by a DDM micelle, for molecular weight calculation, we used the "protein conjugate" module of the ASTRA software, which allows calculation of protein and detergent micelle molecular weights<sup>25</sup>. This method requires values of dn/

dc for proteins and detergent and A280 for detergent and proteins. The value of  $dn/dc_{\text{protein}}$  is approximately the same (around 0.187 ml/g) for most soluble proteins. The values of  $dn/dc_{\text{DDM}} = 0.1435$  ml/g and  $A280_{\text{DDM}} = 0.04$  were taken from the (<https://www.anatrace.com/Products/Detergents/MALTOSIDES/D310A>). A280 for proteins was calculated using the ProtParam Tool (<http://us.expasy.org>).

**HDX-MS.** Analysis of the H/D exchange was performed as described previously<sup>23</sup> with modifications of the data analysis procedures. Peptides were identified using ProteinLynx Global SERVER software (PLGS, Waters). The list of identified peptides generated from the on-line pepsin digestion is presented in the supplementary data (Supplementary Table S3). Identified peptides after exchange were analysed by the DynamX 3.0 program (Waters) with the following acceptance criteria: minimum intensity threshold of 3000, minimum products per amino acid of 0.2, RT deviation  $\pm 10$ S, and m/z deviation  $\pm 10$  ppm. The isotopic envelopes of the peptides after exchange were analysed using DynamX 3.0 with manual corrections wherever necessary. Final data containing molecular weights of all peptides at a given incubation time ( $M_{\text{ex}}$ ) were exported to Excel (Microsoft Office) for calculations. Percent of peptide deuteration was calculated with a formula that takes into consideration the molecular weights from the minimum ( $M_{\text{ex}}^0$ ) and maximum exchange ( $M_{\text{ex}}^{100}$ ) values of a given peptide. These values were obtained in control experiments as described previously<sup>23</sup>.

$$D(\%) = \frac{M_{\text{ex}} - M_{\text{ex}}^0}{M_{\text{ex}}^{100} - M_{\text{ex}}^0} * 100\%$$

Error bars for percent deuteration D(%) were calculated as the square root of the sum of the standard deviation values of the compared states of three independent experiments.

To calculate the overall degree of protection of a given peptide (aggregated protection, show in Fig. S6a), we used kinetic plots showing the fraction of exchange at different incubation times. These plots allow monitoring of the increase in deuteration fraction over time in each peptide. For fully flexible regions, the exchange is instantaneous in the timeframe of the experiment, and the kinetic curve, linking the datapoints at different incubation times (1 minute, 30 minutes, 24 hours), is parallel to the horizontal axis at 100% of exchange. For fully protected regions, the kinetic curve is parallel to the horizontal axis at 0% exchange. To characterize the overall protection by a single number, we calculated the area over the kinetic curve, obtaining a degree of protection in % (aggregated protection), usually a number between 0% of protection in fully flexible regions and 100% of protection of fully protected (non-exchanging in the time frame of the experiment) regions. These numbers were further used to calculate the differential aggregated protection ( $\Delta\text{HDex}$ ) by subtraction of the aggregated protection in common peptides of FL\_RAGE and V\_C1\_C2. Estimation of the error in the difference in exchange ( $\Delta\text{HDex}$ ), obtained by subtracting the aggregated protection in the FL\_RAGE and V\_C1\_C2, was calculated as the square root of the sum of the standard deviation values of the compared states. Final graphs were plotted using OriginLab 9.

**Native MS.** The samples were desalted and buffer-exchanged using P-6 Micro BioSpin columns (Bio-Rad) pre-equilibrated with 100 mM AmAc for V\_C1 and C2, 500 mM (AmAc) for V\_C1\_C2, or 200 mM (AmAc) and 0.02% (w/v, 2 CMC) DDM for the C2\_TM\_CT wild-type and mutant proteins. The AmAc concentration for FL-RAGE was ranged from 200 to 1,000 mM containing either 0.04% (w/v, 2 CMC) Triton X-100 or 0.02% (w/v, 2 CMC) DDM. The concentration of all proteins was approximately 10  $\mu\text{M}$ . MS measurements of all protein variants were also conducted under denaturing conditions achieved by the addition of 50% acetonitrile and 0.1% formic acid. Samples were introduced into the mass spectrometer by nano-electrospray ionization using in-house produced gold-coated glass capillaries at a voltage of +1.6 kV. MS experiments were performed either on a travelling-wave ion mobility Q-TOF (Synapt G2 HDMS, Waters, Manchester, UK)<sup>46,47</sup> or a high-mass modified Q-TOF (MS Vision, Almere, NL)<sup>48</sup> mass spectrometer tuned to maintain the native structure of soluble proteins and preserve the micelle-embedded membrane protein complexes upon entry into the gas phase. MS settings were adjusted to obtain an optimal ion transmission as follows: 1.4–1.6 kV capillary voltage, 25–75 V sampling cone, 3 V extractor cone, 30 °C source temperature, 15–30 V trap collision energy, and 5 V transfer collision energy. Gentle release of C2\_TM\_CT and FL\_RAGE membrane protein complexes was achieved by the following instrument settings: 1.6–1.8 kV capillary voltage, 200 V sampling cone, 5 V extractor cone, 30 °C source temperature, 50–200 V trap collision energy, and 5 V transfer collision energy. Pressures throughout the instrument were 6–9 mbar backing, 3.1 mbar in the ion mobility cell, and 2.5  $10^{-2}$  mbar for the trap and transfer cell, respectively. IM-MS was performed with a constant wave height of 40 V and a wave velocity of 800 m/s. CIU experiments were performed by stepwise (5 V) ramping the collision energy between 5–175 V for V\_C1 and V\_C1\_C2 or between 85–175 V for FL\_RAGE. Deconvolution of native mass spectra was performed using the online tool ([http://www.bioprocess.org/esiprot/esiprot\\_form.php](http://www.bioprocess.org/esiprot/esiprot_form.php)) and UniDec software<sup>49</sup>.

**AUC.** Sedimentation velocity AUC experiments were performed using a ProteomeLab XL-I analytical ultracentrifuge (Beckman-Coulter, Indianapolis, USA), equipped with an An-50, 8-hole analytical rotor and double-sector charcoal-Epon cells (12-mm path length). The experiments were carried out at 20 °C, at 40,000 rpm (extracellular truncated variants V\_C1\_C2, V\_C1, and C2) and 42,000 rpm (FL\_RAGE, C2\_TM\_CT, and mutants the containing TM and cytosolic domains), using continuous scan mode and radial spacing of 0.003 cm. Scans were collected with no intervals between scans, both in absorbance (280 nm) and interference mode. Cells were loaded with 400  $\mu\text{L}$  of sample and 410  $\mu\text{L}$  of buffer. In case of TM-containing proteins, standard sample buffer (25 mM Tris, pH 7.5, 200 mM NaCl; but not reference buffer) was supplemented by the addition of 0.03% of DDM. Solvent densities and viscosities were measured at 20 °C using an Anton Paar (Graz, Austria) DMA 5000 densitometer and Lovis 2000 M rolling-ball viscometer. Partial specific volume and extinction coefficients

for proteins were calculated using SEDNTERP software ([http://bitwiki.sr.unh.edu/index.php/Main\\_Page](http://bitwiki.sr.unh.edu/index.php/Main_Page)). Extracellular truncated variants: buffer density = 1.00716 g/cm<sup>3</sup>, and viscosity = 1.017 mPas. Partial specific volume ( $\bar{V}$ ) of V\_C1\_C2 = 0.73366 cm<sup>3</sup>/g; V\_C1 = 0.73208 cm<sup>3</sup>/g; and C2 = 0.73204 cm<sup>3</sup>/g. Full-length and N-terminal truncated variants: buffer with DDM (0.02%), density = 1.00772 g/cm<sup>3</sup>, and viscosity = 1.012 mPas; reference buffer without DDM, FL\_RANGE  $\bar{V}$  = 0.7324. Data were analysed using mostly the “Continuous c(s) distribution” model of the SEDFIT (version 16.1) program<sup>50</sup>, but in case of the C2 fragment, “Continuous c(s) with bimodal  $f/f_0$ ” procedure was useful. In all cases, confidence level (F-ratio) was specified to 0.68. The proteins before loading were diluted to absorbance 0.9 at 280 nm. After loading into the centrifuge and reaching a temperature of 20 °C, the cells were equilibrated for 1 hour. Centrifugation was monitored by both absorbance (280 nm) and interference optics.

## Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on request.

Received: 24 June 2019; Accepted: 11 December 2019;

Published online: 30 December 2019

## References

1. Ali, M. H. & Imperiali, B. Protein oligomerization: how and why. *Bioorg. Med. Chem.* **13**, 5013–5020 (2005).
2. Sakaguchi, T., Tu, Q., Pinto, L. H. & Lamb, R. A. The active oligomeric state of the minimalistic influenza virus M2 ion channel is a tetramer. *Proc. Natl. Acad. Sci. USA* **94**, 5000–5005 (1997).
3. Arthur, C. P. & Stowell, M. H. B. Structure of synaptophysin: a hexameric MARVEL-domain channel protein. *Structure* **15**, 707–714 (2007).
4. Dean, M., Hamon, Y. & Chimini, G. The human ATP-binding cassette (ABC) transporter superfamily. *J. Lipid Res.* **42**, 1007–1017 (2001).
5. Guo, W., Shi, L. & Javitch, J. A. The fourth transmembrane segment forms the interface of the dopamine D2 receptor homodimer. *J. Biol. Chem.* **278**, 4385–4388 (2003).
6. Gupta, K. *et al.* The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **541**, 421–424 (2017).
7. Nobre, L., Wise, D., Ron, D. & Volmer, R. Modulation of innate immune signalling by lipid-mediated MAVS transmembrane domain oligomerization. *PLoS One* **10**, e0136883 (2015).
8. Futosi, K., Fodor, S. & Mocsai, A. Reprint of Neutrophil cell surface receptors and their intracellular signal transduction pathways. *Int. Immunopharmacol.* **17**, 1185–1197 (2013).
9. Treanor, B. B-cell receptor: from resting state to activate. *Immunology* **136**, 21–27 (2012).
10. Ng, S. Y. L., Lee, L. T. O. & Chow, B. K. C. Receptor oligomerization: from early evidence to current understanding in class B GPCRs. *Front. Endocrinol. (Lausanne)*. **3**, 175 (2012).
11. Schmidt, A. M. *et al.* Isolation and characterization of two binding proteins for advanced glycosylation end products from bovine lung which are present on the endothelial cell surface. *J. Biol. Chem.* **267**, 14987–14997 (1992).
12. Huttunen, H. J., Fages, C., Kuja-Panula, J., Ridley, A. J. & Rauvala, H. Receptor for advanced glycation end products-binding COOH-terminal motif of amphoterin inhibits invasive migration and metastasis. *Cancer Res.* **62**, 4805–4811 (2002).
13. Koch, M. *et al.* Structural basis for ligand recognition and activation of RAGE. *Structure* **18**, 1342–1352 (2010).
14. Xue, J. *et al.* Advanced glycation end product recognition by the receptor for AGEs. *Structure* **19**, 722–732 (2011).
15. Yatime, L. & Andersen, G. R. The specificity of DNA recognition by the RAGE receptor. *The Journal of experimental medicine* **211**, 749–750 (2014).
16. Yatime, L. *et al.* The Structure of the RAGE:S100A6 Complex Reveals a Unique Mode of Homodimerization for S100 Proteins. *Structure* **24**, 2043–2052 (2016).
17. Cai, Z. *et al.* Role of RAGE in Alzheimer’s Disease. *Cell. Mol. Neurobiol.* **36**, 483–495 (2016).
18. *Hydrogen Exchange Mass Spectrometry of Proteins.*, <https://doi.org/10.1002/9781118703748> (John Wiley & Sons, Ltd, 2016).
19. Wei, W. *et al.* Disulfide bonds within the C2 domain of RAGE play key roles in its dimerization and biogenesis. *PLoS One* **7**, e50736 (2012).
20. Xue, J. *et al.* Change in the Molecular Dimension of a RAGE-Ligand Complex Triggers RAGE Signaling. *Structure* **24**, 1509–1522 (2016).
21. Park, H., Adsit, F. G. & Boyington, J. C. The 1.5 Å crystal structure of human receptor for advanced glycation endproducts (RAGE) ectodomains reveals unique features determining ligand binding. *J. Biol. Chem.* **285**, 40762–40770 (2010).
22. Yatime, L. & Andersen, G. R. Structural insights into the oligomerization mode of the human receptor for advanced glycation end-products. *FEBS J.* **280**, 6556–6568 (2013).
23. Sitkiewicz, E., Tarnowski, K., Poznanski, J., Kulma, M. & Dadlez, M. Oligomerization interface of RAGE receptor revealed by MS-monitored hydrogen deuterium exchange. *PLoS One* **8**, e76353 (2013).
24. Dattilo, B. M. *et al.* The extracellular region of the receptor for advanced glycation end products is composed of two independent structural units. *Biochemistry* **46**, 6957–6970 (2007).
25. Slotboom, D. J., Duurkens, R. H., Olieman, K. & Erkens, G. B. Static light scattering to characterize membrane proteins in detergent solution. *Methods* **46**, 73–82 (2008).
26. Salvay, A. G., Santamaria, M., le Maire, M. & Ebel, C. Analytical Ultracentrifugation Sedimentation Velocity for the Characterization of Detergent-Solubilized Membrane Proteins Ca<sup>++</sup>-ATPase and ExbB. *J. Biol. Phys.* **33**, 399–419 (2007).
27. Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods in enzymology* **562**, 109–133 (2015).
28. Wilton, R., Yousef, M. A., Saxena, P., Szpunar, M. & Stevens, F. J. Expression and purification of recombinant human receptor for advanced glycation endproducts in *Escherichia coli*. *Protein Expr. Purif.* **47**, 25–35 (2006).
29. Roman, E. & González Flecha, F. L. Kinetics and Thermodynamics of Membrane Protein Folding. *Biomolecules* **4**, 354–373 (2014).
30. Kleiger, G., Grothe, R., Mallick, P. & Eisenberg, D. GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry* **41**, 5990–7 (2002).
31. Cymer, F., Veerappan, A. & Schneider, D. Transmembrane helix–helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochim. Biophys. Acta - Biomembr.* **1818**, 963–973 (2012).
32. Mueller, B. K., Subramaniam, S. & Senes, A. A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical Calpha-H hydrogen bonds. *Proc. Natl. Acad. Sci. USA* **111**, E888–95 (2014).
33. Kupniewska-Kozak, A., Gospodarska, E. & Dadlez, M. Intertwined structured and unstructured regions of exRAGE identified by monitoring hydrogen-deuterium exchange. *J. Mol. Biol.* **403**, 52–65 (2010).
34. Xu, D. *et al.* Stable RAGE-heparan sulfate complexes are essential for signal transduction. *ACS Chem. Biol.* **8**, 1611–1620 (2013).

35. Dixit, S. M., Polasky, D. A. & Ruotolo, B. T. Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Curr. Opin. Chem. Biol.* **42**, 93–100 (2018).
36. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).
37. Xie, J. *et al.* Hexameric calgranulin C (S100A12) binds to the receptor for advanced glycosylated end products (RAGE) using symmetric hydrophobic target-binding patches. *J. Biol. Chem.* **282**, 4218–4231 (2007).
38. Xie, J. *et al.* Structural basis for pattern recognition by the receptor for advanced glycation end products (RAGE). *J. Biol. Chem.* **283**, 27255–27269 (2008).
39. Popa, I., Ganea, E. & Petrescu, S. M. Expression and subcellular localization of RAGE in melanoma cells. *Biochem. Cell Biol.* **92**, 127–136 (2014).
40. Sarkany, Z. *et al.* Solution structure of the soluble receptor for advanced glycation end products (sRAGE). *J. Biol. Chem.* **286**, 37525–37534 (2011).
41. Fink, A., Sal-Man, N., Gerber, D. & Shai, Y. Transmembrane domains within the membrane milieu: principles, advances and challenges. *Biochim. Biophys. Acta* **1818**, 974–983 (2012).
42. Zong, H. *et al.* Homodimerization is essential for the receptor for advanced glycation end products (RAGE)-mediated signal transduction. *J. Biol. Chem.* **285**, 23137–23146 (2010).
43. Berkmen, M. Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expr. Purif.* **82**, 240–251 (2012).
44. Gavriilidou, A. F. M., Gulbakan, B. & Zenobi, R. Influence of Ammonium Acetate Concentration on Receptor-Ligand Binding Affinities Measured by Native Nano ESI-MS: A Systematic Study. *Anal. Chem.* **87**, 10378–10384 (2015).
45. Seetoh, W.-G., Chan, D. S.-H., Matak-Vinkovic, D. & Abell, C. Mass Spectrometry Reveals Protein Kinase CK2 High-Order Oligomerization via the Circular and Linear Assembly. *ACS Chem. Biol.* **11**, 1511–1517 (2016).
46. Sobott, F., McCammon, M. G., Hernandez, H. & Robinson, C. V. The flight of macromolecular complexes in a mass spectrometer. *Philos. Trans. A. Math. Phys. Eng. Sci.* **363**, 379–391 (2005).
47. Konijnenberg, A., Butterer, A. & Sobott, F. Native ion mobility-mass spectrometry and related methods in structural biology. *Biochim. Biophys. Acta* **1834**, 1239–1256 (2013).
48. van den Heuvel, R. H. H. *et al.* Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal. Chem.* **78**, 7473–7483 (2006).
49. Marty, M. T. *et al.* Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
50. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–19 (2000).

## Acknowledgements

We thank Lilia Zhukova for help with protein expression and purification. We thank Dr. Zbigniew Pietras for help with SEC-MALS experiments. We acknowledge financial support from EU infrastructural projects CEPT (POIG.02.02.00-14-024/08-00). The equipment used in this study was sponsored by the National Multidisciplinary Laboratory of Functional Nanomaterials (POIGT.02.02.00-00-025/09-00). The work was supported by the National Science Centre, Poland: MAESTRO grant (2014/14/A/NZ1/00306) to M.D. and A.M. European Cooperation in Science and Technology (COST BM1403 - Native Mass Spectrometry and Related Methods for Structural Biology) Short-Term Scientific Mission to A.M. D.H. and F.S. acknowledge the Antwerp University Research Fund for the Concerted Research Actions grant (BOF-GOA 4D protein structure).

## Author contributions

A.M., D.H., F.S., R.S. and M.D. designed the study and wrote the manuscript. A.M. designed and constructed vectors for expression of recombinant full-length RAGE, truncated variants and mutant proteins. A.M. designed, performed, and analysed HDX-MS and disulphide x-linking data. A.M., F.S. and D.H., performed native MS experiments and analysed the data. R.S. performed AUC experiments and analysed the data. A.M. designed SEC-MALS experiments and analysed the data. All authors approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-019-56993-9>.

**Correspondence** and requests for materials should be addressed to A.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019