

# SCIENTIFIC REPORTS



OPEN

## Diversity between mammalian tolloid proteinases: Oligomerisation and non-catalytic domains influence activity and specificity

Received: 02 December 2015

Accepted: 25 January 2016

Published: 23 February 2016

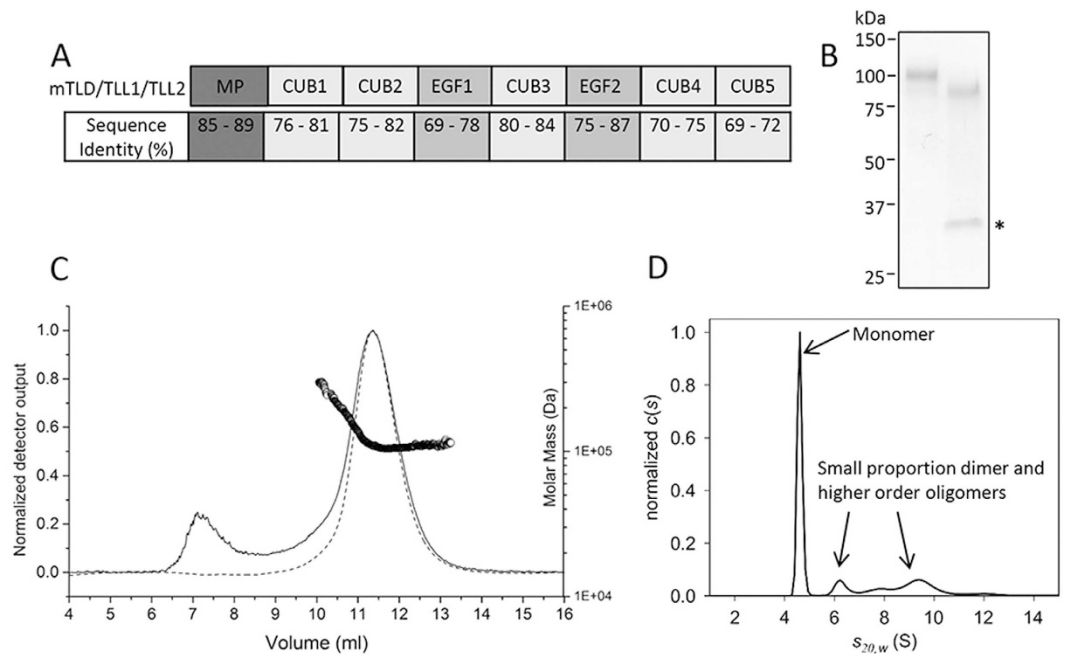
Christopher P. Bayley<sup>1</sup>, Hilda D. Ruiz Nivia<sup>1</sup>, Rana Dajani<sup>1</sup>, Thomas A. Jowitt<sup>1</sup>, Richard F. Collins<sup>2</sup>, Heather Rada<sup>3</sup>, Louise E. Bird<sup>3</sup> & Clair Baldock<sup>1</sup>

The mammalian tolloid family of metalloproteinases is essential for tissue patterning and extracellular matrix assembly. The four members of the family: bone morphogenetic protein-1 (BMP-1), mammalian tolloid (mTLD), tolloid-like (TLL)-1 and TLL-2 differ in their substrate specificity and activity levels, despite sharing similar domain organization. We have previously described a model of substrate exclusion by dimerisation to explain differences in the activities of monomeric BMP-1 and dimers of mTLD and TLL-1. Here we show that TLL-2, the least active member of the tolloid family, is predominantly monomeric in solution, therefore it appears unlikely that substrate exclusion via dimerisation is a mechanism for regulating TLL-2 activity. X-ray scattering and electron microscopy structural and biophysical analyses reveal an elongated shape for the monomer and flexibility in the absence of calcium. Furthermore, we show that TLL-2 can cleave chordin *in vitro*, similar to other mammalian tolloids, but truncated forms of TLL-2 mimicking BMP-1 are unable to cleave chordin. However, both the N- and C-terminal non-catalytic domains from all mammalian tolloids bind chordin with high affinity. The mechanisms underlying substrate specificity and activity in the tolloid family are complex with variation between family members and depend on both multimerisation and substrate interaction.

BMP-1 (procollagen C-proteinase-1) and mammalian tolloid (mTLD) are alternatively spliced products of the *Bmp1* gene<sup>1</sup>. Together with mammalian tolloid like-1 (TLL-1) and -2, they comprise a small group of zinc and calcium dependent proteinases, fundamental to morphogenesis and extracellular matrix (ECM) assembly. The BMP-1/TLD family is conserved in species ranging from *Drosophila* to humans and their importance is highlighted by the embryonic lethal phenotype of *Bmp1/Tll1* homozygous null mice, which display heart malformations and abnormal procollagen processing<sup>2</sup>. In vertebrates, BMP-1/TLD proteinases are involved in the biosynthetic processing of a wide range of ECM precursors including collagens, the crosslinking enzymes lysyl oxidase (LOX) and LOX-like, laminin basement membrane proteins and small leucine-rich proteoglycans osteoglycin and probiglycan (for reviews see<sup>3,4</sup>). BMP-1/TLD proteinases also release a number of TGF $\beta$  superfamily members from their corresponding latent complexes, including BMP-2 and -4, growth and differentiation factors 8 (promyostatin) and 11 and TGF $\beta$ 1. This modulates dorsal ventral patterning, growth of skeletal muscle and neural tissue, and cellular behaviour respectively<sup>5-8</sup>. These dual roles have fuelled speculation that BMP-1/TLD proteinases orchestrate ECM assembly via signalling by TGF- $\beta$ -like proteins<sup>9</sup>.

BMP-1/TLD proteinases have an N-terminal protease domain related to the digestive enzyme astacin<sup>10</sup> followed by non-catalytic CUB (Complement/Uegf/BMP-1) and EGF-like domains (Fig. 1A). mTLD, TLL-1 and TLL-2 have 5 CUB and 2 EGF-like domains whereas BMP-1 lacks two CUB and one EGF C-terminal domains.

<sup>1</sup>Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, Manchester, M13 9PT, UK. <sup>2</sup>Faculty of Life Sciences, University of Manchester, Manchester, M13 9PT, UK. <sup>3</sup>OPPF-UK, Research Complex at Harwell, Rutherford Appleton Laboratory, Oxford, OX11 0FA, UK. Correspondence and requests for materials should be addressed to C.B. (email: clair.baldock@manchester.ac.uk)



**Figure 1. Expression and characterization of human TLL-2 which is predominantly a monomer in the presence of  $\text{Ca}^{2+}$ .** (A) Schematic diagram of the domain organisation of mTLD, TLL-1 and TLL-2 (MP = metalloproteinase domain) and sequence identities between domains in mTLD, TLL-1 and TLL-2. (B) Coomassie stained SDS-PAGE of purified TLL-2, in the absence and presence of PNGaseF. TLL-2 runs as a single band of >100 kDa and has a shift in mobility after treatment with PNGaseF. The PNGaseF can be seen on the gel as a single band at ~30 kDa (highlighted with an asterisk). (C) Multi-angle light scattering (MALS) analysis of TLL-2, the graph shows the normalised differential refractive index (dashed line), light scattering at 90 degrees (solid line) and molecular mass of the proteins as they elute from the gel filtration column. The majority of the protein elutes as a monomeric protein with an apparent molecular mass of  $106.7 \pm 0.97$  kDa (experimental errors from polydispersity). A small amount of higher molecular weight species co-elutes at the beginning of the protein peak which could be a mixture of dimer and higher molecular weight species. (D) C(s) analysis of TLL-2 as derived from sedimentation velocity AUC analysis, indicating the majority of the sample is comprised of a monomer, with a small proportion of higher molecular weight species. Biophysical analyses were carried out in the presence of 2 mM calcium chloride.

BMP-1 has the highest activity whereas its splice variant mTLD has much lower activity and TLL-2 is generally accepted to have the lowest activity of the mammalian family members. BMP-1, mTLD and TLL-1 are highly expressed in developing skeletal tissue, whereas, TLL-2 has specific expression profiles in developing skeletal muscle and in the central nervous system<sup>11</sup>. All members of the mammalian tolloid family are capable of cleaving pro-myostatin but TLL-2 efficiently cleaves pro-myostatin *in vitro*, despite having low activity for other tolloid substrates (i.e. procollagen and chordin)<sup>8</sup> and muscle mass is increased in TLL-2 deficient mice consistent with decreased levels of processed myostatin<sup>12</sup>. TLL-2 has also recently been shown to have a genetic association with bipolar disorder, although the mechanism is not yet understood<sup>13</sup>.

There is evidence that the non-catalytic domains have a negative regulatory role<sup>14,15</sup> and we have previously provided an explanation as to why mTLD is a less efficient proteinase relative to BMP-1<sup>16</sup>. Our findings illustrated that BMP-1 is a monomer whereas mTLD and TLL-1 form dimers in the presence of calcium-ions, which restricts their activities by a substrate exclusion mechanism<sup>16,17</sup>. The C-terminal non-catalytic domains mediate dimerisation and removal of these domains increases the activity of mTLD and TLL-1. However, we recently showed that *Drosophila* tolloid is a monomer<sup>18</sup>, despite having high activity for its substrate Sog and removal of the C-terminal domains resulted in loss of activity, suggesting that there may be differences in the requirement of the non-catalytic domains and dimerisation between tolloids.

In the present report, we provide evidence that TLL-2 is predominantly monomeric in the presence of calcium. We show that truncated forms of TLL-2, having the same domain structure as BMP-1, are unable to cleave the tolloid substrate, chordin. We also demonstrate that all mammalian tolloids bind chordin with high affinity via the N-terminal and C-terminal non-catalytic domains.

## Results

**TLL-2 is preferentially monomeric in solution unlike mTLD and TLL-1.** Human TLL-2 was expressed in a mammalian expression system and purified as a secreted protein (Fig. 1B). Tryptic peptide analysis by MS validated the identity of the purified protein. The size and oligomeric status of TLL-2 was first analysed by multi-angle light scattering (MALS) in conjunction with size exclusion chromatography. In the presence of 2 mM calcium chloride, the majority of the protein elutes as a monomeric species with molecular mass of

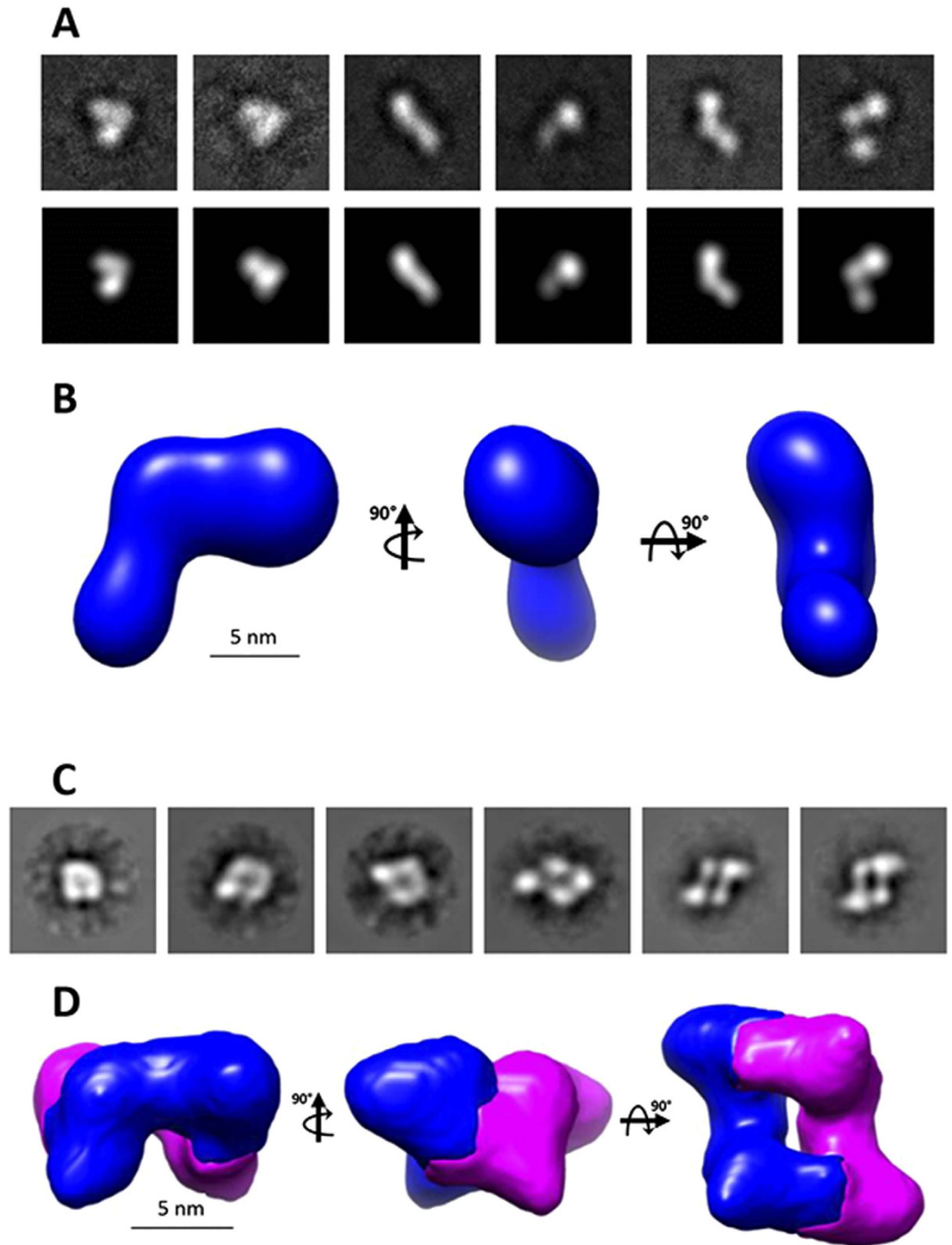
106.7 ± 0.97 kDa (Fig. 1C), which is slightly larger than the predicted mass of a monomer (98.58 kDa). A small amount of higher molecular weight species co-elutes at the beginning of the protein peak which could be a mixture of dimer or higher molecular weight species. TLL-2 has 5 N-linked glycans predicted from primary sequence and digestion with PnGaseF, an amidase that cleaves between the innermost GlnAc and Asp residues of N-linked glycoproteins, yielded a shift in mobility on SDS-PAGE consistent with glycosylation (Fig. 1B). In-line quasi-elastic light scattering, which measures the diffusion of particles and therefore the hydrodynamic radius ( $R_H$ ) and sedimentation velocity analytical ultracentrifugation (AUC) which measures the sedimentation coefficient, were used to help determine the solution properties of monomeric mTLL2. A  $R_H$  of  $4.8 \pm 0.21$  nm and a sedimentation coefficient of 4.78 S (Fig. 1D) are consistent with a slightly elongated monomeric protein. This is also consistent with the frictional ratio ( $f/f_0$ ) determined by AUC of 1.62 which is characteristic of a folded, but slightly extended molecule. Although the main species was a monomer both the MALS and AUC showed evidence of small amounts of higher order oligomers.

**The structure of TLL-2 is consistent with the other mammalian tolloids and stabilised by  $\text{Ca}^{2+}$ .** The structure of TLL-2 was investigated using single particle EM with negative staining (Fig. 2). There were two oligomeric states of TLL-2 observed on the EM grid, a monomeric form and a dimer, which was consistent with the shape seen for mTLD and TLL-1 previously<sup>16,17</sup>. Using reference-free methods, 3D reconstructions were calculated using angular reconstitution. The 3D model generated for the monomer was an elongated asymmetric shape with dimensions 17.2 nm × 6.9 nm × 6.9 nm (Fig. 2B). The dimer had a similar overall shape to the models produced for mTLD and TLL-1 and it was possible to segment the dimer into two units with consistent shape to the monomer (Fig. 2D). To determine the 3D shape of the protein in solution, SAXS measurements were made at the PETRAIII synchrotron radiation source (Fig. 3). In the presence of calcium, TLL-2 had a tendency to aggregate at high concentrations therefore SAXS data were collected in the presence of EDTA (Fig. 3A). The data quality was assessed using Guinier plots, to check for aggregation in the sample (Fig. 3B). The radius of gyration ( $R_g$ ) obtained from the Guinier plot was 5.7 nm. The maximum particle dimension was estimated as 19 nm using indirect Fourier transform with GNOM (Fig. 3C)<sup>19</sup>. The dimensionless Kratky plot (Fig. 3D) shows that the protein is folded but not globular as there is a peak with maxima ~2.3, 1.4 rather than 1.73, 1.1 as expected for a globular protein<sup>20</sup>. Rigid body modeling was also performed using homology models of the eight TLL-2 domains in SASREF<sup>21</sup>. The models fitted the experimental data with a mean discrepancy factor  $\chi$  of 1.1 ( $N = 8$ ;  $q < 2.5 \text{ nm}^{-1}$ ). Each simulation produced fairly similar models with an asymmetric elongated shape (Fig. 3E).

As these SAXS data were collected in the absence of  $\text{Ca}^{2+}$  which is known to stabilize cbEGF-CUB domain interfaces<sup>22</sup>, analysis of the flexibility of the protein was performed. For a non-flexible protein, the Porod-Debye plot ( $q^4$ ) should contain a plateau within the low resolution region of the SAXS data<sup>23</sup>, however no plateau was observed in the Porod-Debye plot (not shown). Furthermore, a plateau was present in the SIBYLS plot ( $q^3$ ) (Fig. 4A)<sup>23</sup> which indicates a flexible protein. The Porod exponent, used to assess the compactness of the protein, was 2.4 which suggests that the protein is fairly elongated and is consistent with the AUC sedimentation parameters. The cross-sectional radius is 2.5 nm, significantly smaller than the  $R_g$  and  $R_h$ , which again is consistent with an elongated protein. To gain further insight into the apparent flexibility of TLL-2 in the absence of  $\text{Ca}^{2+}$  the conformational flexibility was evaluated using the ensemble optimization method (EOM)<sup>24</sup>. Since the structures of homologous domains are known, we were able to generate a pool of models that could be compared to the experimental scattering data (Fig. 4B). From a pool of 10,000 models an ensemble of three structures was predicted based on the SAXS data, two compact collapsed conformations with  $R_g$  of ~4.6 nm (62.5% of the ensemble) and a more elongated extended conformation with  $R_g$  6.1 nm (37.5% of the ensemble) (Fig. 4C). The average  $R_g$  of 5.1 nm and  $D_{\text{max}}$  of 16.8 nm of the ensemble are similar to the estimated values by GNOM. The more elongated conformation bears most similarity to the EM structure, suggesting this most resembles the conformation in the presence of  $\text{Ca}^{2+}$ .

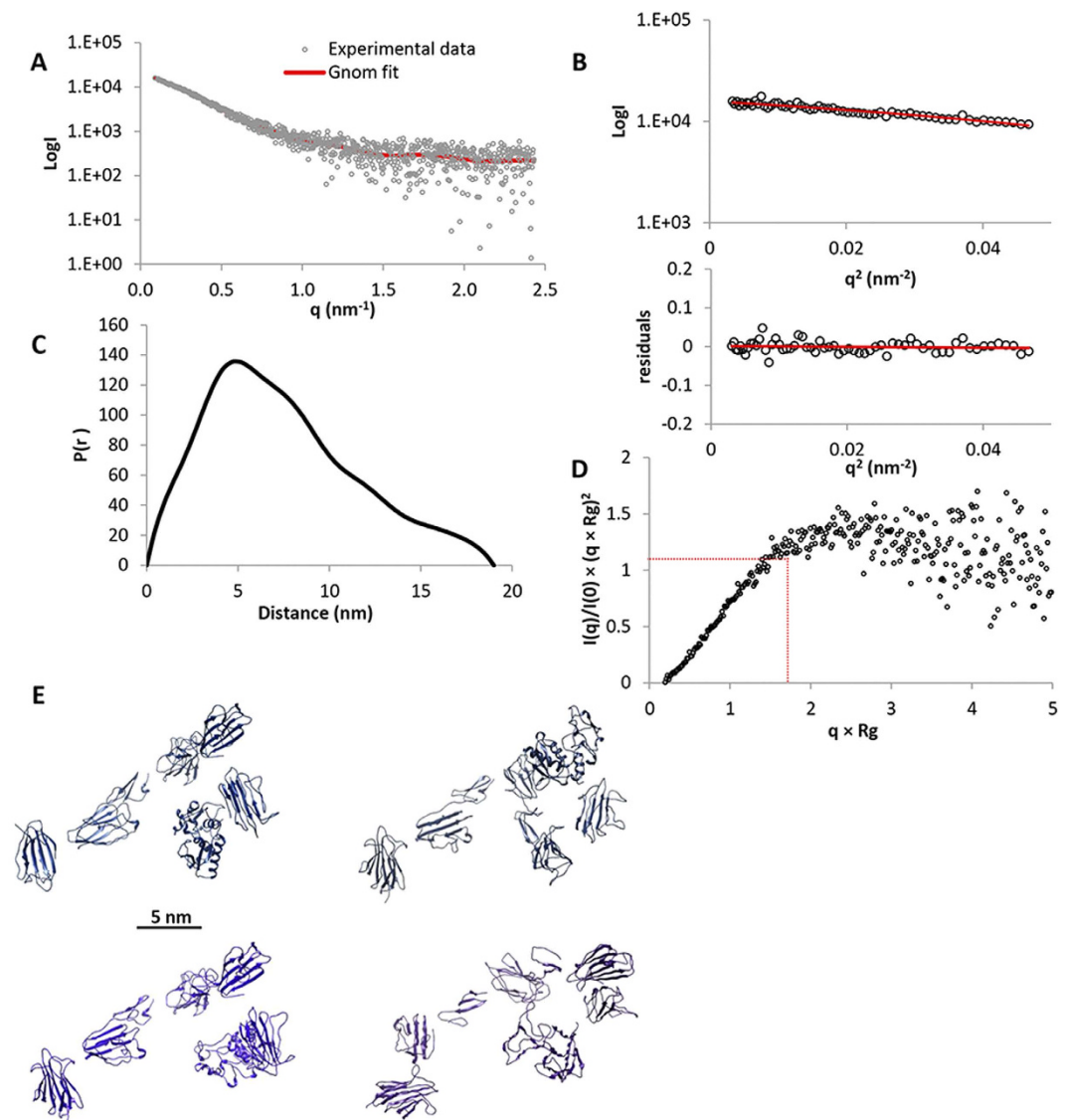
**Hydrodynamic analysis of TLL-2 also supports an extended conformation in the presence of  $\text{Ca}^{2+}$ .** To further compare the predicted conformations with the experimental hydrodynamic data measured by MALS and AUC, a different approach was undertaken. First, the program Ranch<sup>24</sup> was used to generate 5,000 random models using the eight individual domains of TLL-2. Following this, the program SoMo<sup>25</sup> was used to calculate the hydrodynamic parameters ( $s$ ,  $R_h$ ,  $f/f_0$  and  $R_g$ ) of each model (Fig. 4D). A subset of eight models that fitted the hydrodynamic data ( $R_h$ ,  $s$ ) were identified (Fig. 4E). This procedure showed molecular shapes that are similar and an average model was generated rendered to 20 Å resolution (Fig. 4F) which fitted well with the EM data (Fig. 4G). These models also suggest that a significant degree of interdomain flexibility is possible in TLL-2, even in the presence of bound calcium (Fig. 4E).

**Deletion of CUB4 and CUB5 domains ablates TLL-2 chordinase activity.** To investigate the role of the non-catalytic domains and determine whether TLL-2 is restricted by a substrate exclusion mechanism, TLL-2 variants lacking one or both C-terminal CUB domains (TLL2TC4 and TLL2TE2), and a construct with the same domain structure as BMP-1 (TLL2TC3) were made (Fig. 5A). The variants were analysed by SDS-PAGE and size exclusion chromatography to confirm in each case that the constructs were monomeric and intrinsic fluorescence spectroscopy to confirm that the proteins were folded (Fig. 5B). To determine the proteolytic activity of TLL-2 and variants, we compared their chordinase activity to that of mTLD (Fig. 5C). Chordinase activity was detected for TLL-2, this was similar to mTLD. However, the truncated forms of TLL-2 had no activity against chordin (Fig. 5C), unlike truncated TLL-1 and BMP-1 which are more active than their longer forms<sup>16,17</sup> (i.e. full-length TLL-1 and mTLD). These assays suggest that CUB5 in TLL-2 is essential for chordin cleavage but is dispensable in TLL-1 and mTLD (i.e. in BMP1). This finding is consistent with studies on *Drosophila* TLD which also requires CUB5 for Sog cleavage<sup>18</sup>.



**Figure 2. Monomeric and dimeric forms of TLL-2 can be resolved by electron microscopy.** Low dose negatively stained EM images were recorded on a Tecnai T12 TEM using a CCD detector. (A) Selected representative class sum images (top row) and reprojections (bottom row) of different views of TLL-2 monomer (box size =  $36 \times 36$  nm) and dimer (box size =  $39 \times 39$  nm) (C). (B) Orthogonal views of the 3D reconstruction of TLL-2 monomer generated by angular reconstitution of the class sums are shown. (D) The dimeric TLL-2 can be compared and shows a similar conformation to the monomer. Images of the 3D reconstruction represented as a solid surface are shown.

**Binding of non-catalytic domains from mammalian tolloid family to chordin.** Following from this finding, to determine if different non-catalytic domains are responsible for the interaction with chordin or if the affinity of the interaction differs between tolloids, we used SPR binding assays. The N-terminal non-catalytic domains (C1C2E1) from TLL-2, TLL-1 and mTLD/BMP1 were analysed for binding to  $\Delta$ N-chordin (Fig. 6A). The  $K_D$  of each interaction was calculated by fitting a 1:1 Langmuir binding model to the data. All constructs bound with high affinity (e.g. TLL-1:  $K_D = 0.057$  nM) (Table 1). The same assay was performed with the

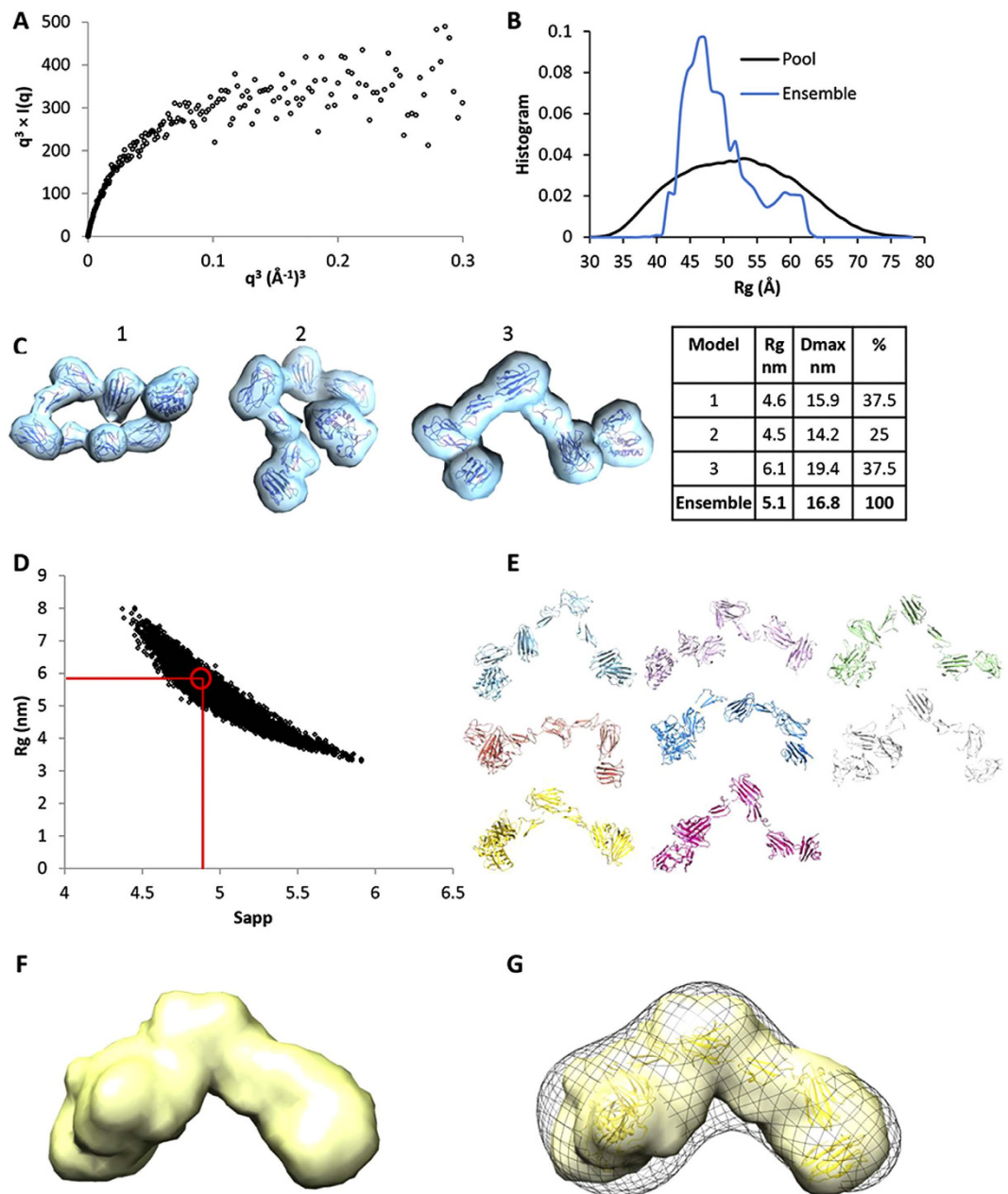


**Figure 3. Solution small angle X-ray scattering data for TLL-2 in the absence of  $\text{Ca}^{2+}$ .** (A) The experimental SAXS data for TLL-2 in the presence of EDTA are plotted as a function of  $q$ . (B) The low-angle region of the X-ray scattering data is shown in the form of a Guinier plot, which is linear. (C) The distance distribution function shows a maximum dimension of 19 nm. (D) The dimensionless Kratky plot is shown with peak at  $\sim 2.3$   $q \times R_g$ , the expected peak maxima for a globular protein is indicated in red. (E) Four representative rigid body models generated with SASREF are shown.

C-terminal non-catalytic domains, C4C5 from TLL-2, TLL-1 and mTLD (Fig. 6B). Again the binding of all three proteinases to chordin was high affinity, with  $K_D$ s from 0.25 nM (TLL-2) to 0.056 nM (TLL-1). Together these data suggest a slightly stronger interaction between chordin and TLL-1 and weaker to mTLD/BMP1 and TLL-2 but in all cases binding was tight and in the sub-nanomolar range. Although CUB5 is required for TLL-2 cleavage of chordin, but dispensable in mTLD/TLL-1, the N-terminal non-catalytic domains also bind to chordin with high affinity.

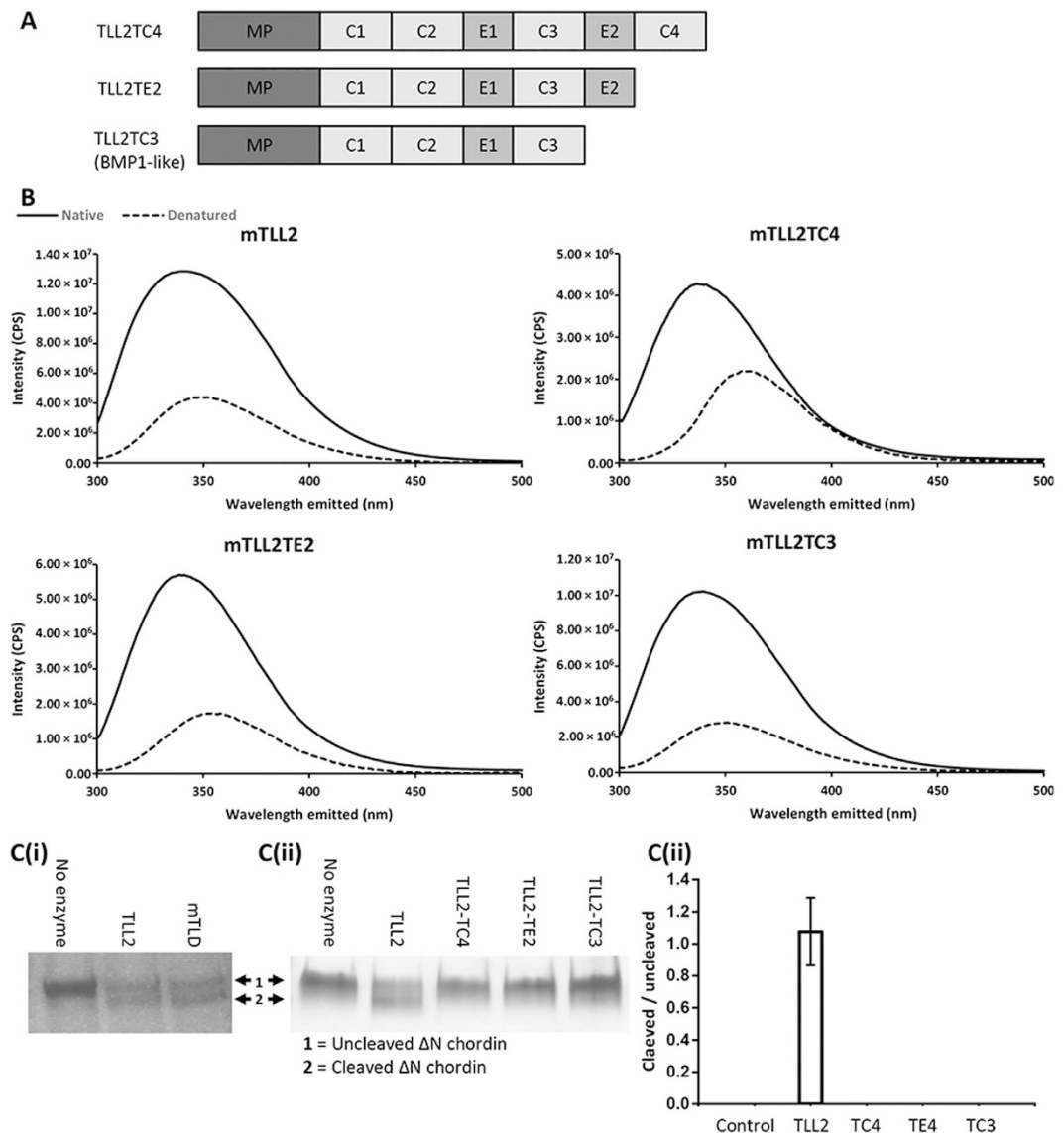
## Discussion

Our previous studies had demonstrated that a substrate exclusion mechanism controls the activity of mTLD and TLL-1<sup>16,17</sup>. Calcium-dependent dimerisation of mTLD and TLL-1 was postulated to lower tolloid activity by either reducing the number of monomers available to bind substrate or reducing the substrate affinity of each protease molecule within the dimer (due to steric hindrance caused by the C-terminal non-catalytic domains partially occluding the adjacent protease domains). However, here we show that the third member of the long mammalian tolloid proteinase family, TLL-2 does not follow the same pattern as shown for mTLD and TLL-1. We report here that although some dimerisation does occur, mTLL-2 is predominantly a monomer in the presence of calcium (Figs 1 and 2). Consequently, the activity of mTLL-2 cannot be controlled by the same substrate exclusion mechanism as for mTLD and TLL-1.



**Figure 4. Flexibility in the absence of  $\text{Ca}^{2+}$  and conformation with  $\text{Ca}^{2+}$ -bound.** (A) SIBYLS plot shows a flexible protein in the absence of  $\text{Ca}^{2+}$ . (B) Plot showing the distribution of the pool of models generated by EOM plotted as a function of  $R_g$ . The distribution of the ensemble selected from this pool is shown. (C) An ensemble of three structures generated EOM that represent the SAXS data. (D) Plot of the  $R_g$  vs  $S$  parameters of the Ranch models as calculated with SoMo. (E) Eight models that satisfy the hydrodynamic properties. (F) The average of these 8 models rendered to 20 Å resolution. (G) The average in (F) with one model fitted, superimposed with the EM map.

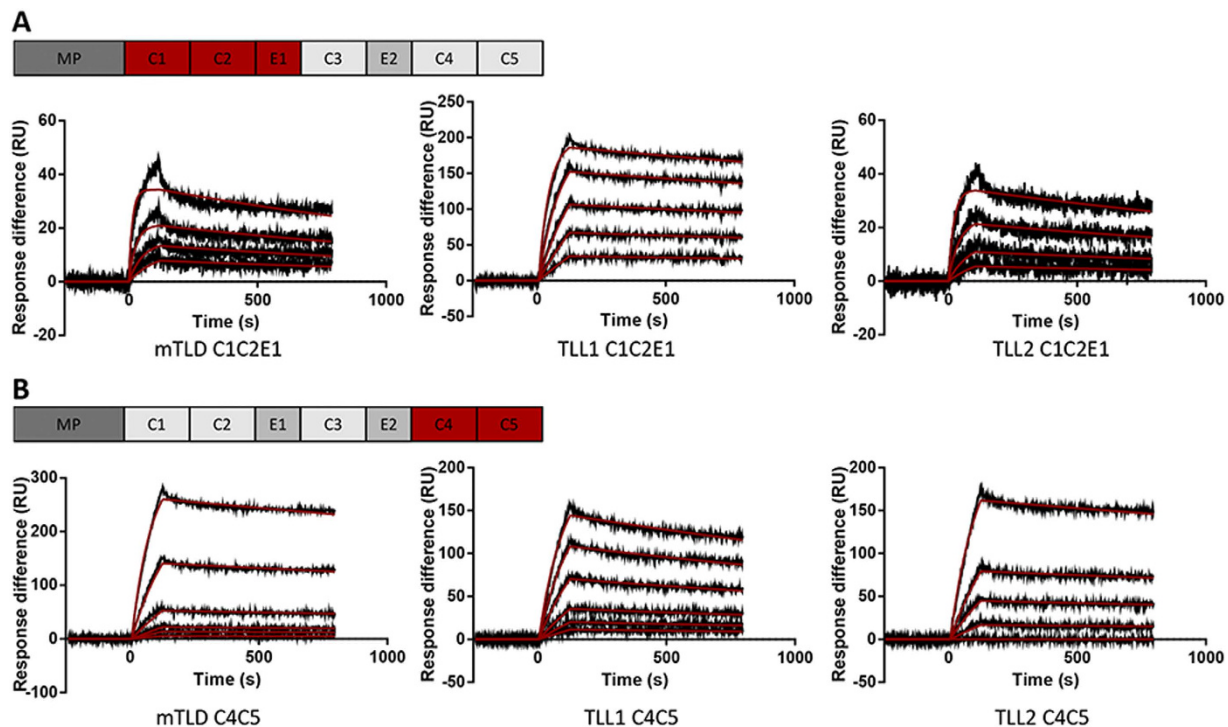
We have previously shown that removing CUB and EGF domains from the C-terminus of either mTLD or TLL-1 increases the chordinase activity of these proteinases<sup>16,17</sup>. This finding is consistent with other reports that the protease domain alone of BMP-1 cleaved procollagen VII with greater efficiency than full-length BMP-1<sup>26</sup> and that the protease domains of both TLL-1 and BMP1 cleaved procollagen VII with greater efficiency than the corresponding full-length tolloids<sup>14</sup>. The present study demonstrates that removing the corresponding domains from the C-terminus of TLL-2 does not alter the fold of the truncated proteinase, yet ablates chordinase activity (Fig. 5), indicating that the deleted non-catalytic domains form critical interactions with the substrate. Consequently, our data suggest that substrate recognition is an important mechanism of TLL-2 activity. Similarly, *Drosophila* tolloid (dTLD) has also been found to have a requirement for CUB5 in order to cleave Sog, the chordin analogue in flies<sup>18</sup>.



**Figure 5. Unlike mTLD and TLL-1, CUB5 is necessary for TLL-2 chordinase activity.** (A) Domain organisation of truncated forms of TLL-2 terminating after CUB4 (TLL2-TC4), EGF2 (TLL2-TE2) and CUB3 (TLL2-TC3) is an analogue of BMP1. TLL2-TC3 is a BMP1-like. (B) Intrinsic fluorescence spectroscopy on TLL-2, TLL2-TC4, TLL2-TE2, TLL2-TC3 in native and denatured (8M Urea) conditions shows a shift in emittance indicating all proteins were folded. (C) SDS-PAGE analysis showing the cleavage of  $\Delta$ N-chordin by TLL-2 and mTLD and (ii) TLL-2 compared to the truncated forms (TLL2-TC4, TLL2-TE2 and TLL2-TC3). This cleavage assay was carried out in triplicate and a representative gel is shown. (iii) Quantification of the cleavage assay by densitometry using the ratio of cleaved to uncleaved chordin. Errors from standard deviation (N = 3).

The non-catalytic domains are certainly important in the activity and substrate specificity of tolloids. For example, Kadler and co-workers demonstrated that CUB2 was necessary for maximum cleavage of procollagen I by BMP-1<sup>27</sup>. Stocker and co-workers analysed the binding of CUB and EGF domain doublets to procollagen I by SPR and showed that binding affinity to procollagen increases for domains towards the C-terminal end of mTLD<sup>28</sup>. They also show that fragments containing EGF domains bind procollagen more strongly than those containing only CUB domains. However, we show that the CUB4-CUB5 region binds very tightly to chordin so EGF domains appear not to be a requirement for tolloid-chordin binding. A previous study using BMP1 and TLL-2 domain swaps showed that the protease domain of TLL-2 fused to the non-catalytic domains of BMP1 was able to cleave chordin as efficiently as BMP1 demonstrating that the non-catalytic domains are essential for activity. Indeed the protease domain of BMP1 fused to the non-catalytic domains of TLL-2 had no chordinase activity and in addition, the protease domain of TLL2 was unable to cleave procollagen even when fused to the CUB domains of BMP1 showing that the non-catalytic domains also strongly influence substrate specificity<sup>15</sup>.

In this study, we have shown that protein constructs from all mammalian tolloids, comprising either CUB1-CUB2-EGF1 or CUB4-CUB5, bind to chordin (Fig. 6). This observation indicates that when a tolloid



**Figure 6.** N- and C-terminal non-catalytic tolloid domains bind substrate tightly. SPR analysis of  $\Delta$ N-chordin binding to CUB and EGF domains from mammalian tolloids. Real-time binding curves in black. Models of analysis in red. (A) Langmuir kinetic analysis of  $\Delta$ N-chordin binding to immobilised protein fragments comprising CUB1, CUB2 and EGF1 (C1C2E1) from mTLD, TLL-1 or TLL-2. (B) Langmuir kinetic analysis of  $\Delta$ N-chordin binding to immobilised protein fragments comprising CUB4 and CUB5 (C4C5) from mTLD, TLL-1 or TLL-2. Experiments performed in triplicate, representative curves shown.

Immobilised tolloid fragment	$K_D$ (nM) Langmuir binding
mTLD CUB1-CUB2-EGF1	$0.278 \pm 0.85$
TLL1 CUB1-CUB2-EGF1	$0.057 \pm 0.003$
TLL2 CUB1-CUB2-EGF1	$0.247 \pm 0.26$
mTLD CUB4-CUB5	$0.19 \pm 0.02$
TLL1 CUB4-CUB5	$0.056 \pm 0.01$
TLL2 CUB4-CUB5	$0.25 \pm 0.02$

**Table 1.** Kinetic analysis of  $\Delta$ N-chordin binding to different fragments of mammalian tolloids, obtained by SPR. Experiments performed in triplicate.

proteinase interacts with chordin, it makes protein-protein interactions that span the majority of the non-catalytic domains. These interactions may be necessary in order to bring both proteinase and substrate into a favourable conformation for cleavage to occur. With respect to TLL-2, since CUB5 is required for chordin cleavage, a specific interaction between CUB5 and chordin may be necessary to enable cleavage by the proteinase. The extended, flexible conformation of TLL-2, indicated by the biophysical data (Figs 3 and 4), suggests that TLL-2 has the potential to change conformation upon substrate binding which may be mediated by the non-catalytic CUB and EGF domains.

TLL-2 must interact with the substrate chordin in a different way to the other members of the tolloid family. This difference is conferred by two factors. Firstly, TLL-2 is predominantly a monomer in the presence of calcium, whereas mTLD and TLL-1 are calcium dependent dimers. The TLL-1 dimer has a strong intermolecular interaction ( $< \text{nM Kd}$ ) which is a similar affinity to the tolloid-chordin interaction we observe, whereas the mTLD intermolecular dimer interaction is weaker ( $\mu\text{M Kd}$ ). Secondly, the non-catalytic domains of TLL-2 show differences to the corresponding non-catalytic domains of mTLD and TLL-1. CUB5 of TLL-2 is necessary for chordinase activity, whereas removal of C-terminal CUB domains from mTLD and TLL-1 increases chordinase activity. Similarly for dTLD, in the absence of CUB4 or CUB5 dTLD was unable to bind Sog (as assessed by immunoprecipitation)<sup>18</sup>. However, dTLD also has further differences to the mammalian tolloids, for example the requirement of dpp (BMP) ligand for Sog cleavage<sup>29</sup>. Collagen IV also provides a scaffold for the Sog-tolloid interaction<sup>18</sup> whereas the mammalian tolloid family do not appear to require these other contributors. Differences between



mTLD/TLL-1 and TLL-2/dTLD may point to a discrete range of substrates processed by TLL-2/dTLD, where the requirement could be for specialised substrate cleavage rather than the broad specificity seen for TLL-1/BMP1. Analysis of TLL-2, the third member of the mammalian tollid proteinases, has provided an insight into how substrate recognition and substrate cleavage by tollids could be mediated by the non-catalytic domains of these proteinases.

## Materials and Methods

**Expression and purification of recombinant proteins.** Full length human TLL-2<sup>15</sup> was used to generate constructs encoding TLL2-TC4, TLL2-TE2 and TLL2-TC3 by PCR. The constructs were ligated into a modified pCEP-Pu vector<sup>30</sup> and transfected into HEK 293-EBNA cells cultured as described previously<sup>16</sup>. Constructs encoding mTLD-C1C2E1, TLL1-C1C2E1, TLL2-C1C2E1, mTLD-C4C5, TLL1-C4C5 and TLL2-C4C5 were generated at the Oxford Protein Production Facility (OPPF-UK), Harwell, UK in a pOPINTTgneo expression vector<sup>31</sup> provided by OPF and transfected into HEK 293S cells.  $\Delta$ N-chordin was generated as previously described<sup>32</sup>. For all constructs a 6x histidine tag was incorporated at the C-terminus. Conditioned media was concentrated and buffer exchanged into 10 mM HEPES, 500 mM NaCl, 10 mM imidazole, 2 mM CaCl<sub>2</sub>, pH 7.4 using tangential flow ultrafiltration (Pall Life Sciences). All recombinant proteins were then purified by nickel affinity chromatography followed by size-exclusion chromatography on an AKTA purifier HPLC using a Superdex200 10/300GL column (GE Healthcare) in 10 mM HEPES, 500 mM NaCl, 2 mM CaCl<sub>2</sub>, pH 7.4. Where needed, proteins were concentrated using Vivaspin centrifugal concentrators (Sartorius). Protein identities were confirmed by in-gel trypsin digestion and liquid chromatography tandem mass spectrometry (LC-MS/MS) using a NanoAcquity LC (Waters) coupled to a LTQ Velos (Thermo Fisher Scientific).

**Intrinsic fluorescence spectroscopy (IFS).** Samples (0.25  $\mu$ M) in a 16.100-F quartz cuvette (Starna Scientific, Essex, UK) were analysed using a JASCO FP-750 Spectrometer (JASCO UK, Essex, UK) under the control of the Spectra Manager software. The excitation wavelength was 290 nm and an emission scan between 300–500 nm was recorded. Analyses were repeated following addition of 8 M urea (final concentration) to each sample. Spectrums generated by buffers alone were subtracted from the protein measurements.

**Multi-angle light scattering (MALS) Analysis.** Samples (0.5 ml at approximately 0.5 mg/ml) were loaded onto a Superdex200 10/300GL column running at a flow rate of 0.75 ml/min in 10 mM Tris, 500 mM NaCl, 2 mM CaCl<sub>2</sub>, pH 7.4. Samples eluting from the column passed through a DAWN Wyatt HeliosII 18-angle laser photometer. One of the detectors on the photometer was replaced with a Wyatt QELS detector. This was coupled to a Wyatt Optilab rEX refractive index detector and the molecular mass moments, polydispersity, hydrodynamic radii and concentrations of the resulting peaks were analysed using Astra 6.1 (Wyatt, Santa Barbara, USA).

**Analytical Ultracentrifugation (AUC).** Sedimentation velocity AUC on TLL-2 (0.4 mg/ml) was carried out at 45,000 rpm, 20 °C using an XL-A ultracentrifuge with an An50Ti-4-hole rotor (Beckman Coulter, High Wycombe, UK). Sedimentation was scanned every 90 seconds for 200 scans. Data were analysed using Sedfit<sup>33</sup>.

**Solution Small angle X-ray scattering (SAXS).** SAXS intensity data on TLL-2 in 10 mM Tris, 500 mM NaCl, 2 mM EDTA pH 7.4 at 2 mg/ml were collected from protein samples and matched buffer blanks at the EMBL-P12 beamline at PETRAIII (DESY, Hamburg) employing automated data acquisition and radial averaging protocols<sup>34</sup>. The forward scattering intensity,  $R_g$  and distance distribution function  $p(r)$  were evaluated with GNOM<sup>19</sup>. Rigid body modelling against the experimental SAXS data was performed with SASREF<sup>21</sup> using the eight individual domains of TLL-2 with a distance range of 5–10 Å between each domain. EOM<sup>24</sup> was used to evaluate the flexibility of TLL-2 and to select the best fitting ensemble to the SAXS data.

**Hydrodynamic analysis.** A library of 5,000 conformers of TLL-2 were generated by Ranch<sup>24</sup> then SoMo<sup>25</sup> was used for calculation of hydrodynamic parameters for each model including ( $R_g$ ), hydrodynamic radius ( $R_h$ ), sedimentation coefficient ( $s$ ), maximal linear distance ( $D_{max}$ ) and frictional ratios. Experimental hydrodynamic data from MALS and AUC were used to select a pool of models that fit the experimental values.

**EM and Single-Particle Analysis.** TLL-2 (~10  $\mu$ g/ml) in 10 mM HEPES, 500 mM NaCl, 2 mM CaCl<sub>2</sub>, pH 7.4 was adsorbed onto glow-discharged carbon-coated grids and stained with 4% (w/v) uranyl acetate (pH 4.7). Grids were observed using either a FEI Tecnai Twin (120 keV) or Biotwin (100 keV) transmission EM for the monomer and dimer, respectively. Images were recorded under low dose conditions ( $< 10 e^-/\text{Å}^2$ ) on 2,048  $\times$  2,048 pixel CCD cameras at 30,000  $\times$  (2.8 Å/pixel monomer) and 28,000  $\times$  magnification (3.5 Å/pixel dimer) between –0.2 and –2.0  $\mu$ m defocus. Eman2<sup>35</sup> was used for particle picking and image processing. Images were CTF corrected and the total number of particles in each the dataset was 1,260 for TLL-2 monomers and 5,524 for TLL-2 dimers. Characteristic class-sum images were used as references to align the dataset. Angular reconstruction produced unique projection classes, enabling calculation of an initial 3D reconstruction which was then subjected to 5 rounds of iterative refinement. C2 symmetry was applied to the TLL-2 dimer and a 20 Å low-pass Gaussian filter was used at each stage of the refinement.

**Activity Assays.** Purified  $\Delta$ N-chordin (1.8  $\mu$ g) was incubated in the presence or absence of 110 ng enzyme in 50 mM Tris-HCl (pH 7.4) containing 150 mM NaCl and 5 mM CaCl<sub>2</sub> at 37 °C for 16 hours<sup>36</sup>. Reactions were stopped by adding LDS sample buffer (Life Technologies) and 2.5 %  $\beta$ -mercaptoethanol and heating to 95 °C for 5 min. Reaction products were separated by SDS-PAGE and visualised by silver staining. Chordinase assay products were quantified by densitometry using GeneTools software (Syngene UK, Cambridge, UK).

**Surface plasmon resonance (SPR).** Protein-protein interactions were measured with a ProteOn XPR36 (Bio-Rad Laboratories) in 10 mM HEPES, 150 mM NaCl, 2 mM CaCl<sub>2</sub>, 0.05% Tween 20, pH 7.4 (running buffer) at 25 °C. Proteins were immobilised on a GLC chip (Bio-Rad Laboratories) via amine coupling and a reference lane blocked with ethanolamine. Analytes diluted in running buffer were injected onto the surface of the chip at increasing concentrations for kinetic analysis. Response curves were analysed using the ProteOn Manager Software (Bio-Rad Laboratories) fitting a 1:1 Langmuir model to each interaction.

## References

1. Takahara, K., Lyons, G. E. & Greenspan, D. S. Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encoded by alternatively spliced transcripts which are differentially expressed in some tissues. *J Biol Chem* **269**, 32572–8 (1994).
2. Pappano, W. N., Steiglit, B. M., Scott, I. C., Keene, D. R. & Greenspan, D. S. Use of Bmp1/Tll1 doubly homozygous null mice and proteomics to identify and validate *in vivo* substrates of bone morphogenetic protein 1/tolloid-like metalloproteinases. *Mol Cell Biol* **23**, 4428–38 (2003).
3. Muir, A. & Greenspan, D. S. Metalloproteinases in Drosophila to humans that are central players in developmental processes. *J Biol Chem* **286**, 41905–11 (2011).
4. Vadon-Le Goff, S., Hulmes, D. J. & Moali, C. BMP-1/tolloid-like proteinases synchronize matrix assembly with growth factor activation to promote morphogenesis and tissue remodeling. *Matrix Biol* **44–46C**, 14–23 (2015).
5. Scott, I. C. *et al.* Homologues of Twisted gastrulation are extracellular cofactors in antagonism of BMP signalling. *Nature* **410**, 475–8 (2001).
6. Ge, G. & Greenspan, D. S. BMP1 controls TGFbeta1 activation via cleavage of latent TGFbeta-binding protein. *J Cell Biol* **175**, 111–20 (2006).
7. Ge, G., Hopkins, D. R., Ho, W. B. & Greenspan, D. S. GDF11 forms a bone morphogenetic protein 1-activated latent complex that can modulate nerve growth factor-induced differentiation of PC12 cells. *Mol Cell Biol* **25**, 5846–58 (2005).
8. Wolfman, N. M. *et al.* Activation of latent myostatin by the BMP-1/tolloid family of metalloproteinases. *Proc Natl Acad Sci USA* **100**, 15842–6 (2003).
9. Hopkins, D. R., Keles, S. & Greenspan, D. S. The bone morphogenetic protein 1/Tolloid-like metalloproteinases. *Matrix Biol* **26**, 508–23 (2007).
10. Bond, J. S. & Beynon, R. J. The astacin family of metalloendopeptidases. *Protein Sci* **4**, 1247–61 (1995).
11. Scott, I. C. *et al.* Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. *Dev Biol* **213**, 283–300 (1999).
12. Lee, S. J. Genetic analysis of the role of proteolysis in the activation of latent myostatin. *PLoS One* **3**, e1628 (2008).
13. de Mooij-van Malsen, J. G. *et al.* Cross-species genetics converge to TLL2 for mouse avoidance behavior and human bipolar disorder. *Genes Brain Behav* **12**, 653–7 (2013).
14. Ge, G., Zhang, Y., Steiglit, B. M. & Greenspan, D. S. Mammalian tolloid-like 1 binds procollagen C-proteinase enhancer protein 1 and differs from bone morphogenetic protein 1 in the functional roles of homologous protein domains. *J Biol Chem* **281**, 10786–98 (2006).
15. Petropoulou, V., Garrigue-Antar, L. & Kadler, K. E. Identification of the minimal domain structure of bone morphogenetic protein-1 (BMP-1) for chordinase activity: chordinase activity is not enhanced by procollagen C-proteinase enhancer-1 (PCPE-1). *J Biol Chem* **280**, 22616–23 (2005).
16. Berry, R. *et al.* Role of dimerization and substrate exclusion in the regulation of bone morphogenetic protein-1 and mammalian tolloid. *Proc Natl Acad Sci USA* **106**, 8561–8566 (2009).
17. Berry, R., Jowitt, T. A., Garrigue-Antar, L., Kadler, K. E. & Baldock, C. Structural and functional evidence for a substrate exclusion mechanism in mammalian tolloid like-1 (TLL-1) proteinase. *FEBS Lett* **584**, 657–61 (2010).
18. Winstanley, J., Sawala, A., Baldock, C. & Ashe, H. L. Synthetic enzyme-substrate tethering obviates the Tolloid-ECM interaction during Drosophila BMP gradient formation. *Elife* **4** (2015).
19. Semenyuk, A. V. & Svergun, D. V. GNOM - a program package for small-angle scattering data processing *Journal of Applied Crystallography* **24**, 537–540 (1991).
20. Receveur-Brechot, V. & Durand, D. How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr Protein Pept Sci* **13**, 55–75 (2012).
21. Petoukhov, M. V. & Svergun, D. I. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* **89**, 1237–50 (2005).
22. Thielens, N. M. *et al.* The N-terminal CUB-epidermal growth factor module pair of human complement protease C1r binds Ca<sup>2+</sup> with high affinity and mediates Ca<sup>2+</sup>-dependent interaction with C1s. *J Biol Chem* **274**, 9149–59 (1999).
23. Rambo, R. P. & Tainer, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559–71 (2011).
24. Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656–64 (2007).
25. Rai, N. *et al.* SOMO (SOLUTION MOdeler) differences between X-Ray- and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics. *Structure* **13**, 723–34 (2005).
26. Wermter, C. *et al.* The protease domain of procollagen C-proteinase (BMP1) lacks substrate selectivity, which is conferred by non-proteolytic domains. *Biol Chem* **388**, 513–21 (2007).
27. Hartigan, N., Garrigue-Antar, L. & Kadler, K. E. Bone morphogenetic protein-1 (BMP-1). Identification of the minimal domain structure for procollagen C-proteinase activity. *J Biol Chem* **278**, 18045–9 (2003).
28. Hintze, V. *et al.* The interaction of recombinant subdomains of the procollagen C-proteinase with procollagen I provides a quantitative explanation for functional differences between the two splice variants, mammalian tolloid and bone morphogenetic protein 1. *Biochemistry* **45**, 6741–8 (2006).
29. Marques, G. *et al.* Production of a DPP activity gradient in the early Drosophila embryo through the opposing actions of the SOG and TLD proteins. *Cell* **91**, 417–26 (1997).
30. Poschl, E., Fox, J. W., Block, D., Mayer, U. & Timpl, R. Two non-contiguous regions contribute to nidogen binding to a single EGF-like motif of the laminin gamma 1 chain. *EMBO J* **13**, 3741–7 (1994).
31. Nettleship, J. E. *et al.* Transient expression in HEK 293 cells: an alternative to E. coli for the production of secreted and intracellular mammalian proteins. *Methods Mol Biol* **1258**, 209–22 (2015).
32. Troilo, H. *et al.* Nanoscale structure of the BMP antagonist chordin supports cooperative BMP binding. *Proc Natl Acad Sci USA* **111**, 13063–8 (2014).
33. Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys J* **78**, 1606–19 (2000).
34. Franke, D., Kikhney, A. G. & Svergun, D. I. Automated acquisition and analysis of small angle X-ray scattering data. *Nucl. Instr. Meth. Phys. Res. A* **689**, 52–59 (2012).
35. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **157**, 38–46 (2007).
36. Garrigue-Antar, L., Francois, V. & Kadler, K. E. Deletion of epidermal growth factor-like domains converts mammalian tolloid into a chordinase and effective procollagen C-proteinase. *J Biol Chem* **279**, 49835–41 (2004).

## Acknowledgements

CPB was supported by a BBSRC grant (Ref: BB/I012265/1) to CB. We would like to thank Dr Paul Mould in the Biomolecular analysis facility and the EM facility in the Faculty of Life Sciences, University of Manchester. We would also like to thank PETRAIII, Hamburg for SAXS beamtime on beamline P12 and Dr Cy Jefferies for assistance during data collection. The OPPF-UK is gratefully acknowledged for assistance in generating tolloid constructs.

## Author Contributions

C.B. conceived and coordinated the study and wrote the paper. C.P.B. performed and analyzed the experiments and contributed to the preparation of the figures. R.D. and H.D.R.N. expressed and purified proteins used in Figures 5 and 6. R.F.C. performed and analyzed the data shown in Figure 2C and D. T.A.J. performed and analyzed the experiments shown in Figure 4D. H.R. and L.E.B. performed high-throughput cloning of tolloid constructs at the OPPF. All authors reviewed the results and approved the final version of the manuscript.

## Additional Information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Bayley, C. P. *et al.* Diversity between mammalian tolloid proteinases: Oligomerisation and non-catalytic domains influence activity and specificity. *Sci. Rep.* **6**, 21456; doi: 10.1038/srep21456 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>