STRUCTURAL BIOLOGY

ISSN 2059-7983

Structure, substrate recognition and therapeutic targeting of the human ADAMTS-5 spacer domain

Mario Milani,[a,b]* Michela Visintin,[c] Ivet Krastanova,[c]‡ Marco Visentini,[c]‡ Elisa Margotti,[c]‡ Gabriele Ugolini,[c] Martino Bolognesi,[b] Lucio C. Rovati[c] and Eloise Mastrangelo[a,b]*

[a]Biophysics Institute, CNR–IBF, Via Corti 12, 20133 Milan, Italy, [b]Department of Bioscience, Università degli Studi di Milano, Via Celoria 26, 20133 Milan, Italy, and [c]Rottapharm Biotech Srl, Via Valosa di Sopra 9, 20900 Monza, Italy. *Correspondence e-mail: mario.milani@cnr.it, eloise.mastrangelo@cnr.it

Received 26 August 2025
Accepted 16 November 2025

Edited by K. Djinovic-Carugo, EMBL Grenoble, France

‡ Present address: Alifax R&D, SS 14 Km 163, 34149 Basovizza, Trieste, Italy.

Keywords: spacer domain; aggrecanase; osteoarthritis; cysteine-rich domain; selenomethionine labeling.

PDB reference: human ADAMTS-5 C_B and spacer domains, 9rwm

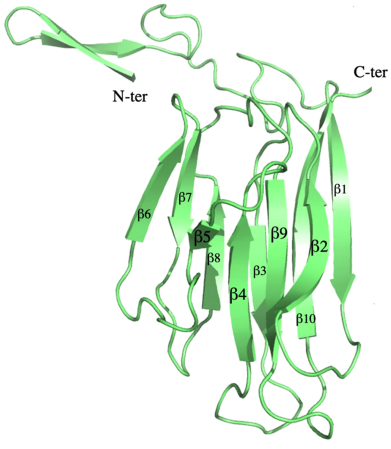
Supporting information: this article has supporting information at journals.iucr.org/d

The ADAMTS (a disintegrin-like and metalloproteinase domain with thrombospondin type 1 motifs) family of secreted metalloproteinases plays essential roles in extracellular matrix remodeling. ADAMTS-5 contributes to cartilage degradation, cleaving proteoglycans such as aggrecan and versican, and being involved in both physiological tissue turnover and pathological processes such as osteoarthritis and atherosclerosis. Although structural insights into its catalytic domain have informed inhibitor development, the role of ancillary domains, particularly the spacer domain, in substrate recognition and specificity remains underexplored. Here, we report the crystal structure of a segment of human ADAMTS-5 encompassing the C-terminal portion of the cysteine-rich domain and the spacer domain (residues 694–876). This structure reveals critical features of the spacer domain, including the hypervariable loops that function as exosites essential for the binding of aggrecan and versican. Our findings provide new structural insights into the molecular determinants of the substrate specificity of ADAMTS-5 and underscore the spacer domain as a promising target for the development of selective inhibitors.

1. Introduction

The ADAMTS proteins (a disintegrin-like and metalloproteinase domain with thrombospondin type 1 motifs) constitute a family of 19 secreted proteases that are active in extracellular matrix (ECM) remodeling. ADAMTS proteins play key roles in embryonic development and tissue homeostasis, and their dysregulation contributes to various diseases, including cancer, fibrosis, arthritis and cardiovascular disorders (Kelwick *et al.*, 2015; Mead & Apte, 2018). Traditionally, ADAMTS enzymes have been classified based on their primary substrates: ADAMTS-2, ADAMTS-3 and ADAMTS-14 are procollagenases and ADAMTS-1, ADAMTS-4, ADAMTS-5, ADAMTS-8, ADAMTS-9, ADAMTS-15 and ADAMTS-20 are hyalectanases, while ADAMTS-13 specifically cleaves von Willebrand factor (VWF). The remaining members, including ADAMTS-6, ADAMTS-7, ADAMTS-10, ADAMTS-12, ADAMTS-16, ADAMTS-17, ADAMTS-18 and ADAMTS-19, following the identification of multiple new substrates, have been grouped as fibrillin- or fibronectin-associated ADAMTS proteases (Rose *et al.*, 2021).

Encoded by the *ADAMTS5* gene on chromosome 21, human ADAMTS-5 (also known as aggrecanase-2; UniProt Q9UNA0; 930 amino acids) is a zinc-dependent metalloproteinase with a critical role in cartilage biology. It is composed of a signal peptide (16 amino acids), a pro-domain, a catalytic metalloproteinase domain (containing the zinc-binding active



site), a disintegrin-like domain, a thrombospondin type 1 domain, a cysteine-rich domain, a spacer domain and another thrombospondin type 1 domain (Supplementary Fig. S1). ADAMTS-5 belongs to the hyalectanase subgroup of the ADAMTS family, collectively known as proteoglycanases due to their ability to cleave large aggregating proteoglycans (Tortorella *et al.*, 2002; Apte, 2009). ADAMTS-5 specifically cleaves aggrecan (a protein heavily modified with negatively charged polysaccharide chains, *i.e.* chondroitin sulfate and keratan sulfate; Tortorella *et al.*, 2002), the principal proteoglycan in cartilage responsible for maintaining tissue hydration and providing resistance to compressive forces. While this proteolytic activity is crucial for normal ECM turnover, it becomes detrimental in osteoarthritis (OA), where excessive aggrecan degradation drives cartilage breakdown and joint degeneration (Glasson *et al.*, 2005; Stanton *et al.*, 2005). These insights have spurred the development of ADAMTS-5 inhibitors as potential therapeutics for OA (Cuffaro *et al.*, 2022). Beyond aggrecan, ADAMTS-5 also cleaves versican, a large chondroitin sulfate proteoglycan prominently expressed in vascular tissues, skin, developing organs and tumors. Versican plays essential roles in embryonic development, cell proliferation, adhesion and migration (Wight, 2017). Notably, ADAMTS-5-mediated cleavage of versican regulates lipoprotein retention in atherosclerosis. ADAMTS-5 knockout (AKO) mice models exhibit reduced ADAMTS-5 activity, increased versican (and biglycan) accumulation and enhanced low-density lipoprotein (LDL) binding in arterial walls (Didangelos *et al.*, 2012). This discovery highlights a pathophysiological role for ADAMTS-5 beyond OA, linking its activity on versican to cardiovascular disease mechanisms (Santamaria & De Groot, 2020; Kemberi *et al.*, 2023; Santamaria, 2020). More recently, ADAMTS-5 has also been proposed as a potential therapeutic target in other pathological contexts, including ovarian cancer (Yuan *et al.*, 2025) and Duchenne muscular dystrophy (Dulos *et al.*, 2025), further underscoring its relevance across diverse disease mechanisms.

Despite the critical roles of ADAMTS-5 in both physiological and pathological processes, the molecular mechanisms underlying its substrate specificity and proteolytic activity remain only partially understood. Although the full-length structure of ADAMTS-5 has not yet been solved, several high-resolution crystal structures of its catalytic domain have provided key insights. For example, the structure of the catalytic domain bound to an inhibitor at 1.4 Å resolution (PDB entry 3b8z; Shieh *et al.*, 2008) revealed a unique substrate-binding cleft and a novel calcium-binding site. Other structures (PDB entry 3ljt; Shieh *et al.*, 2011) highlighted the flexibility of the S1′ pocket using an amino-2-indanol inhibitor, which is essential for the design of selective inhibitors. Additional insights have come from structures of the metalloprotease and disintegrin-like domains bound to inhibitors (PDB entry 2rjq; Mosyak *et al.*, 2008) and the catalytic domain in complex with broad-spectrum and selective inhibitors, such as marimastat (PDB entry 3hy7; Tortorella *et al.*, 2009) and GLPG1972 (PDB entry 6yjm; Brebion *et al.*, 2021), respectively. These structures collectively underscore the distinct architectural features of ADAMTS-5 and support the development of highly specific therapeutic agents that are capable of modulating its activity while sparing related metalloproteinases.

1.1. Relevance of the spacer structure for inhibitor design

While the metalloprotease domain determines cleavage-site specificity, the C-terminal ancillary domains provide essential substrate recognition (Apte, 2009). In particular, the spacer domain (Sp) is critical for substrate recognition and protease activity (Zheng *et al.*, 2003; Gendron *et al.*, 2007; Fushimi *et al.*, 2008). Gao and coworkers demonstrated that deletion of the ADAMTS-5 Sp domain severely impairs aggrecan degradation, and swapping this region with analogous segments from ADAMTS-13 abolishes substrate recognition (Gao *et al.*, 2012). Thus, the Sp domain is critical for the aggrecanase activity and substrate specificity of ADAMTS-5, although the precise molecular mechanisms remain unclear (Gao *et al.*, 2012). Further, Santamaria *et al.* (2019) showed that Sp-domain deletion reduces versicanase activity by 19-fold, with a combined deletion of the cysteine-rich domain causing a 153-fold reduction. Two hypervariable loops within the spacer (residues 739–743 between β1 and β2 and 837–844 between β9 and β10) serve as key exosites for substrate interaction. Replacing these loops with the corresponding regions from ADAMTS-13, which lacks proteoglycanase activity, drastically impairs function, underscoring their essential role in substrate specificity.

A similar requirement for the Sp domain has been observed in ADAMTS-13, which relies on this region for the effective recognition and cleavage of VWF (Akiyama *et al.*, 2009). These findings highlight the spacer domain as a promising target for the design of selective ADAMTS-5 inhibitors.

In this study, we determined the crystal structure of a portion of human ADAMTS-5 (residues 694–876) comprising the C-terminal segment of the cysteine-rich domain (residues 694–731, *i.e.* the C_B domain; Akiyama *et al.*, 2009) and the spacer domain (residues 732–852). Despite good diffraction data, all attempts to solve the structure by molecular replacement (MR) were unsuccessful. Therefore, we employed selenomethionine (SeMet) labeling and determined the initial phases by using single isomorphous replacement with anomalous scattering (SIRAS). By expanding the structural data beyond the catalytic site of ADAMTS-5, this work offers new opportunities for targeted therapeutic strategies, particularly for diseases such as osteoarthritis in which ADAMTS-5 plays a pathogenic role.

2. Materials and methods

2.1. Expression and purification of the native fragment of ADAMTS-5

The native fragment of ADAMTS-5 (residues 694–876, *i.e.* the C_B and Sp domains with a C-terminal His-tag; molecular weight 21.4 kDa) was expressed in *Escherichia coli* SHuffle cells transformed via heat-shock with 5 µl plasmid DNA

(100 ng µl^{-1}; pDXV4 vector; Domainex, United Kingdom; Reich *et al.*, 2006), including a T7 promoter for high-level expression, a C-terminal His_6-tag for purification and an ampicillin-resistance gene for selection, and containing the nucleotide sequence of interest inserted via the TOPO cloning method (Invitrogen). Transformed cells were cultured in four 750 ml flasks containing Super Broth (26.25 g per flask) and incubated for 20 h at 18°C. After harvesting by centrifugation, the cell pellets were resuspended in lysis buffer consisting of 50 m*M* HEPES pH 7.2, 300 m*M* NaCl, 100 m*M* sodium sulfate, 10 m*M* imidazole, 5%(*v*/*v*) glycerol supplemented with 0.9 mg ml^{-1} lysozyme and 100 µg ml^{-1} DNase I. The cells were lysed by sonication and the lysates were clarified by centrifugation. The supernatant was subjected to Ni–NTA-based affinity chromatography with buffer *A* consisting of 50 m*M* HEPES pH 7.2, 300 m*M* NaCl, 100 m*M* sodium sulfate, 10 m*M* imidazole, 5% glycerol. The protein, which was eluted with buffer *A* supplemented with 250 m*M* imidazole, was directly applied onto a Superdex 75 16/600 pg size-exclusion chromatography column equilibrated with a buffer consisting of 300 m*M* NaCl, 100 m*M* sodium sulfate, 50 m*M* HEPES pH 7.2, 5% glycerol. After elution, the protein was concentrated to 10 mg ml^{-1}. The total yield of purified protein was approximately 3.3 mg per litre of culture.

SDS–PAGE analysis confirmed high purity, with a single dominant band at the expected molecular weight (∼21 kDa; Supplementary Fig. S2). The activity of the purified protein was tested with an ELISA assay, demonstrating its capability to bind bovine aggrecan in a concentration-dependent manner, with an EC_{50} of 14 ± 2.1 n*M* (Supplementary Fig. S3).

The purified protein (at a concentration of 10 mg ml^{-1}) was centrifuged at 13 000*g* for 10 min prior to DLS analysis; all measurements were carried out at 10°C in a DynaPro instrument (Protein Solutions, Charlottesville, USA). The DLS data revealed a low polydispersity index (<20%) and a hydrodynamic radius (R_h) of 4.05 nm, indicative of a likely tetrameric assembly (∼88 kDa).

Thermal stability was assessed via thermofluorimetric analysis conducted in a MiniOpticon Real-Time PCR Detection System (Bio-Rad). The fluorescent dye SYPRO Orange (which binds efficiently to hydrophobic residues, displaying a higher quantum yield in a lower dielectric environment) was used to monitor protein unfolding. Solutions of 2.5 µl of the protein were mixed with 3.5 µl SYPRO Orange (Sigma) diluted 1:500 and 19 µl protein buffer. The sample plates were heated from 4 to 99°C at a heating rate of 0.2°C per 5 s. Fluorescence intensity was measured within the excitation and emission ranges 470–505 and 540–700 nm, respectively. The protein exhibited a melting temperature (T_m) of approximately 41°C.

2.2. Crystallization and X-ray data collection

Crystallization screening was performed using an Oryx8 liquid-handling robot (Douglas Instruments) in sitting-drop vapor-diffusion format with Crystal Screen and Crystal Screen 2 from Hampton Research. Crystals were obtained in two different conditions: (i) 30% polyethylene glycol (PEG) 4000, 0.2 *M* ammonium acetate, 0.1 *M* sodium citrate pH 5.6 (Crystal Screen condition 9; A9) and (ii) 2.0 *M* sodium sulfate, 2% PEG 400 (Crystal Screen condition 39; D3).

X-ray diffraction data were collected on the ID23-2 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, adding 25% glycerol to the crystals as a cryoprotectant before flash-cooling. The data were indexed and scaled using *iMOSFLM* (Battye *et al.*, 2011) and *SCALA* (Evans, 2006). Diffraction statistics are reported in Table 1. Despite promising diffraction data, all attempts to solve the structure using molecular replacement (MR) were unsuccessful.

2.3. Expression and purification of selenomethionine (SeMet)-labeled protein

SeMet-substituted protein was expressed in *E. coli* SHuffle cells using standard methionine-suppression protocols. A single colony was grown overnight in lysogeny broth at 37°C, harvested (3000*g*, 10 min), washed in 30 ml minimal medium and transferred into 3 l methionine-deficient medium (Molecular Dimensions). Cultures were grown at 37°C to an OD_{600} of 0.6–0.8 and were then supplemented with 100 mg l^{-1} each of lysine, phenylalanine and threonine, 50 mg l^{-1} each of isoleucine, leucine and valine, and 60 mg l^{-1} L-selenomethionine.

Following 15–20 min of incubation, protein expression was induced with 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside and the cultures were grown overnight at 20°C. The cells were collected and lysed in buffer *A* (0.3 *M* NaCl, 0.1 *M* sodium sulfate, 40 m*M* imidazole, 50 m*M* HEPES pH 7.2, 5% glycerol) via sonication. The lysates were clarified by centrifugation and the combined supernatants from 3 l of culture were purified using the protocol previously described for the native protein, yielding approximately 3 mg SeMet-substituted protein. Protein purity was confirmed by SDS–PAGE (Supplementary Fig. S2). The DLS profile at 10°C revealed a low polydispersity index (<20%) and an R_h of 4.15 nm, consistent with the value obtained for the native protein.

2.4. Crystallization of SeMet-substituted protein and X-ray data collection

Crystals of SeMet-labeled protein were obtained using an Oryx8 robot with 1.6 *M* ammonium sulfate, 0.1 *M* NaCl, 0.1 *M* HEPES pH 7.5 (Crystal Screen 2 condition 32; G8). A SAD data set was collected on the ID23-1 beamline at ESRF (λ = 0.9686 Å), indexed (trigonal crystal, $P3_121$) and scaled using *MOSFLM*/*SCALA* (Table 1). This data set (resolution 15–3.5 Å, AnoCorr/SigAno = 25/1.03) exhibited anomalous signal up to ∼4.7 Å resolution. For SIRAS (single isomorphous replacement anomalous scattering) phasing, the data set was combined with the native trigonal data set from D3 (Table 1).

2.5. Molecular-dynamics simulation

A molecular-dynamics (MD) simulation was performed with *GROMACS* starting from the coordinates of subunit *B* of

Table 1
X-ray data collection and refinement.

Values in parentheses are for the highest resolution shell.

Crystal	Native, monoclinic, A9	SeMet, trigonal, G8	Native, trigonal, D3
Crystallization conditions	0.2 *M* ammonium acetate, 30% PEG 4000, 0.1 *M* sodium citrate pH 5.6	1.6 *M* ammonium sulfate, 0.1 *M* NaCl, 0.1 *M* HEPES pH 7.5	2%(*v*/*v*) PEG 400, 2.0 *M* ammonium sulfate
Beamline	ID23-2, ESRF	ID23-1, ESRF	ID23-2, ESRF
Wavelength (Å)	0.8729	0.9686	0.8729
Space group	$P2_1$	$P3_121$	$P3_121$
a, *b*, *c* (Å)	49.4, 78.3, 77.4	87.1, 87.1, 79.6	86.4, 86.4, 80.8
α, β, γ (°)	90, 105.1, 90	90, 90, 90	90, 90, 90
Resolution (Å)	36.14–2.60 (2.74–2.60)	15.65–3.50 (3.59–3.50)	40.42–3.80 (4.01–3.80)
Unique reflections	17475 (2542)	8518 (638)	3662 (511)
Completeness (%)	99.3 (99.6)	99.9 (100.0)	99.9 (100.0)
Multiplicity	2.8 (2.9)	21.7 (22.2)	7.2 (6.9)
R_{meas}† (%)	18.8 (94.4)	21.7 (137.7)	22.7 (86.7)
$CC_{1/2}$ (%)	98.4 (37.4)	99.9 (96.9)	99.5 (80.3)
Average $I/\sigma(I)$	7.2 (1.9)	13.85 (2.35)	9.4 (2.9)
Molecules in asymmetric unit	3	1	1
R factor‡/R_{free}§ (%)	19.3/26.0		
R.m.s.d., bond lengths (Å)	0.007		
R.m.s.d., angles (°)	0.52		
Ramachandran statistics (%)			
Residues in favored regions	98		
Residues in additionally allowed regions	2		
Average *B* factors (Å^2)			
Chain *A*	36.6		
Chain *B*	42.7		
Chain *C*	36.3		
PDB code	9rwm		

† $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl) \times 100$, where $I_i(hkl)$ is the intensity of the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is its average intensity. ‡ *R* factor $= \sum_{hkl} \big||F_{obs}| - |F_{calc}|\big|/\sum_{hkl}|F_{obs}| \times 100$. § R_{free} is calculated on a randomly selected 5% of reflections for cross-validation.

the A9 crystal (keeping amino acids 695–852) using the AMBER99SB-ILDN force field (Lindorff-Larsen *et al.*, 2010). The parallelepipedal box for the simulation was filled with 10 887 water molecules, and the system charge (+9) was equilibrated with 20 and 29 atoms of Na^+ and Cl^-, respectively. After minimization with the steepest-descent algorithm (final potential energy = -5.70×10^5 kJ mol^{-1}), we performed two successive equilibrations, each lasting for 2 ns, in the NVT ensemble at T = 100 and 300 K. The productive run (1 µs; coordinates saved every 1.6 ns) was performed in the NPT ensemble at T = 300 K and atmospheric pressure with the particle mesh Ewald method for long-range electrostatics, with a 2 fs time step (leap-frog integrator), T-coupling with a modified Berendsen thermostat and Parrinello–Rahman P-coupling, with periodic boundary conditions.

2.6. *AlphaFold* model of ADAMTS-5–aggrecan

Models of the interaction of the ADAMTS5 C_B–Sp domains with aggrecan (without glycosaminoglycans; UniProt P16112) were generated by the *AlphaFold* server, without using any data from the crystal structure (https://alphafoldserver.com/; Jumper *et al.*, 2021).

3. Results

3.1. Production and crystallization of ADAMTS-5 C_B–Sp domains

We expressed and purified a highly charged fragment of human ADAMTS-5 encompassing the cysteine-rich subdomain B (C_B, amino acids 694–731) and the Sp domain (amino acids 732–876). This region contains 20 lysines (∼10% of the sequence) and has a calculated isoelectric point of 9.55, suggesting the possibility of electrostatic interactions with negatively charged moieties including glycosaminoglycans (for example chondroitin sulfate and keratan sulfate). After the purification process (see Section 2 for details), both the native and SeMet-labeled proteins showed a hydrodynamic radius of ∼4 nm (indicative of a likely tetrameric assembly) and low polydispersity (<20%) favorable for crystallization.

3.2. X-ray data analysis and structure solution

Diffracting crystals of native and SeMet-labeled protein were obtained using the crystallization conditions described in Table 1 (see also Section 2).

Although high-quality diffraction data were collected for the native protein, specifically the A9 data set (space group $P2_1$, unit-cell parameters a = 49.4, b = 78.3, c = 77.4 Å, β = 105.1, resolution 2.6 Å), all attempts to determine the native structure via MR proved unsuccessful. The use of homology models based on the exosite-containing fragment of human ADAMTS-13 (PDB entries 3ghm and 3vn4; Akiyama *et al.*, 2013) and sharing approximately 35% sequence identity with the N-terminal region of the ADAMTS-5 spacer construct as search templates failed to yield a viable solution.

We therefore turned to experimental phasing of the SeMet-labeled protein using single-wavelength anomalous dispersion (SAD). A SAD data set (λ = 0.9686 Å; space group $P3_121$, 43.6–3.5 Å resolution; AnoCorr/SigAno = 25/1.03) showed

anomalous signal up to ∼4.7 Å resolution (AnoCorr/SigAno = 21/0.94). For improved phasing, we performed SIRAS phasing by combining the SAD data set with the native data set D3. Matthews analysis suggested the presence of one or two molecules in the crystal asymmetric unit with four (including the N-terminal methionine) or eight Se atoms. Heavy-atom positions were identified with *SHELXD* (Schneider & Sheldrick, 2002), revealing three sites with isomorphous differences [CC all/weak = 30.70/10.40; CFOM = 41.10; CC(E) = 0.178] and seven with anomalous differences [CC all/weak = 60.80/27.10; CFOM = 87.90; CC(E) = 0.461]. Refinement of the heavy-atom positions and phasing, performed with *SHARP* (Bricogne *et al.*, 2003), yielded figures of merit [FOM (acentric/centric)] of 0.223/0.093, an ISO phasing power of 0.229/0.191 and an anomalous phasing power (ANO) of 0.704.

The initial electron-density map was uninterpretable (Supplementary Fig. S4). After iterative residual map interpretation and rephasing in *SHARP*, we retained only three heavy-atom sites [FOM (acentric/centric) = 0.212/0.094; ISO phasing power (acentric/centric) = 0.234/0.202; ANO phasing power = 0.711]. Density modification (DM) improved the initial phases, with the original mutual disposition of the haevy atoms giving a better DM score than the inverted (enantiomorphic) one (0.43 versus 0.27). Final DM cycles (one molecule in the asymmetric unit, 71.6% solvent content) yielded a map suitable for model building (DM map shown in Supplementary Fig. S5; final DM score = 3.68).

After additional DM (Cowtan, 2010), automated model building was performed using *Buccaneer* (Cowtan, 2006), with manual interpretation (*Coot*; Emsley *et al.*, 2010) guided by the crystal structure of ADAMTS-13 (PDB entry 3ghm) to ensure correct chain-tracing topology and side-chain positioning. This process produced an initial model with R and R_{free} values of 30.1% and 34.8%, respectively. The resulting model was then employed for molecular replacement into the native monoclinic data set (space group $P2_1$, three molecules per asymmetric unit, 2.6 Å resolution), followed by rigid-body refinement (R/R_{free} = 46.1/45.8%). After several rounds of manual rebuilding (in *Coot*) and refinement [using *CCP*4 (Agirre *et al.*, 2023), *REFMAC*5 (Murshudov *et al.*, 2011), *BUSTER* (Blanc *et al.*, 2004) and *Phenix* (Liebschner *et al.*, 2019)], the final model converged to R and R_{free} values of 19.3% and 26.0%, respectively, with overall good stereochemical quality (Table 1 and Supplementary Fig. S6).

3.3. Structural features of the ADAMTS-5 C_B–Sp domains

The three molecules in the asymmetric unit were fully modeled from Asn695 (or Thr706 in subunit *A*) to Lys853 (or Asn860 for subunit *B*), except for amino acids 765–767 (belonging to the β3–β4 loop; see below) of chain *C*, which remained disordered (Fig. 1*a*).

The C_B domain (amino acids 694–731) consists of two subdomains: (i) a β-hairpin, spanning residues 696–705, featuring an Arg700–Gly701 turn, stabilized by the disulfide bond between Cys698 and Cys703 (Fig. 1*b*), and (ii) a left-handed superhelix (residue 706–730) forming two curved segments (Thr706–Ser714 and Asp719–Gly726), resembling the rims of spectacles, stabilized by two disulfide bonds (Cys708–Cys724 and Cys721–Cys731) as well as by several electrostatic interactions such as Asp709–Ser714 and Asp727–Ser729 (Fig. 1*b*). In the crystal, the two C_B subdomains exhibit variable mutual orientations, with Thr706 acting as a hinge point (as observed in chains *B* and *C*; Supplementary Fig. S7).

The Sp domain (amino acids 731–852) adopts a β-sandwich with jelly-roll topology composed of two antiparallel β-sheets: the first with six β-strands [β1 (amino acids 731–738), β10 (845–852), β3 (756–762), β8 (817–820), β7 (804–808) and β6 (797–801)] and the second with four [β2 (744–751), β9 (830–836), β4 (773–779) and β5 (784–786)] (Fig. 1*a*). As observed in ADAMTS-13, in the Sp domain the classical jelly-roll topology is extended by an additional β-hairpin (β6–β7) located on the same side of the C_B subdomain (Fig. 1*a*). Structural comparison of all four crystallographic copies (the three molecules in the asymmetric unit and one from the trigonal crystal) revealed a high degree of overall structural conservation (Supplementary Fig. S7). Notable differences were confined to the extended β3–β4 loop (Lys764–Arg770), which is not conserved in the ADAMTS family (Supplementary Fig. S7), and the first C_B subdomain (residues 696–705). The β3–β4 loop (amino acids 763–773) protrudes towards the solvent from the Sp-domain core with considerable conformational freedom, especially between Ala765 and Arg770, only being restrained by the hydrogen bond between Lys764 and the main chain of His813 (in the β7–β8 loop; Fig. 1*c*). The mobility of the β3–β4 loop, as shown by higher *B* factors and poor electron density for chain *C* (amino acids 765–767 not modeled), suggested a potential role in exosite-mediated substrate recognition, together with its two flanking loops: β7–β8 (residues 809–817) and β9–β10 (residues 837–844) (Akiyama *et al.*, 2009; Jin *et al.*, 2010; Fig. 1*c*).

The structure of the β7–β8 loop is stabilized by hydrogen bonds between Ser809 and the main chain of Phe817, as well as between Ser812 and the main chains of amino acids 814 and 815. Additionally, a salt bridge between Asp815 and Arg761 (located in β3) further restrains this region (Fig. 1*c*).

The conformation of the β9–β10 loop (residues 837–844) is constrained by multiple stabilizing factors: the presence of two prolines (Pro839 and Pro842) and the internal electrostatic interaction among Asp838, Thr840 and Lys841, together with hydrogen bonds between Thr837 and the main chain of Thr772 (in the β3–β4 loop) and between the main chains of Thr840 and Lys741 (in the β1–β2 loop), and finally the insertion of the Leu843 side chain into the hydrophobic core of the protein (Fig. 1*c*).

In the modified jelly-roll topology, the presence of the additional β-hairpin β6–β7 enables the formation of a short β7–β8 loop that connects two strands within the same β-sheet. This configuration introduces a unique structural feature that may facilitate substrate interaction in a manner distinct from the other surface-exposed loops. Moreover, in the crystal asymmetric unit, β-sheet β6 (residues 797–801) is in contact

with β1 (residues 731–738) of a neighboring molecule. This crystallographic intermolecular contact suggests the possibility of an assembly interface for the Sp domain, which is consistent with the oligomeric state observed in DLS analysis (Supplementary Fig. S8). While analysis with *PISA* (Krissinel & Henrick, 2007) does not predict any stable quaternary assembly based on this interaction, such an arrangement could increase the effective substrate-binding surface by clustering multiple Sp domains, thereby enhancing the potential for multivalent or cooperative interactions with diverse substrates (Kosasih *et al.*, 2016).

3.4. Molecular-dynamics simulation

Molecular-dynamics simulations were carried out to investigate the structural flexibility and internal dynamics of the C_B and Sp domains of ADAMTS-5 (amino acids 695–852). During the simulation (1 µs) the selected region remained quite stable with respect to the initial conformation (r.m.s.d. of ∼1.5 Å over 130–140 amino acids; Fig. 2*a*), with a transition occurring around 467 ns (with an r.m.s.d. varying from ∼1.5 to ∼3.0 Å; Fig. 2*a*) due to the different conformation of the C_B domain (Fig. 2*b*). Notably, the β-hairpin in the first C_B subdomain (residues 696–705) remained structurally stable throughout the simulation, suggesting its role as a rigid element that may contribute to maintaining the local domain architecture. In contrast, analysis of the second C_B subdomain revealed notable conformational variability (as shown by the root-mean-square fluctuations in Fig. 2*c*), with the two 'rims of spectacles' motifs adopting a flexible mutual orientation, pivoting around a hinge centered at Gln717 (Fig. 2*b*). This hinge-like behavior suggests that Gln717 acts as the bridge between the two 'rims of spectacles', allowing the coordinated movement of adjacent structural elements and potentially facilitating interactions with substrates.

To explore correlated motions within the domain, dynamic cross-correlation maps (DCCMs) were computed (Fig. 2*d*). The covariance DCCM revealed a positive correlation between the β3–β4 loop and both the β7–β8 and β9–β10 loops (Fig. 2*d*), consistent with their spatial proximity. In addition, the β1–β2 loop showed positive correlation with the nearby β9–β10 loop (Fig. 2*d*), showing the functional significance of the former as a potential regulator of the dynamics of the whole loop region, contributing to the overall plasticity and adaptability of the spacer domain during substrate recognition and binding. Collectively, these findings enhance our understanding of the structural determinants underlying the substrate specificity and catalytic efficiency of ADAMTS-5.

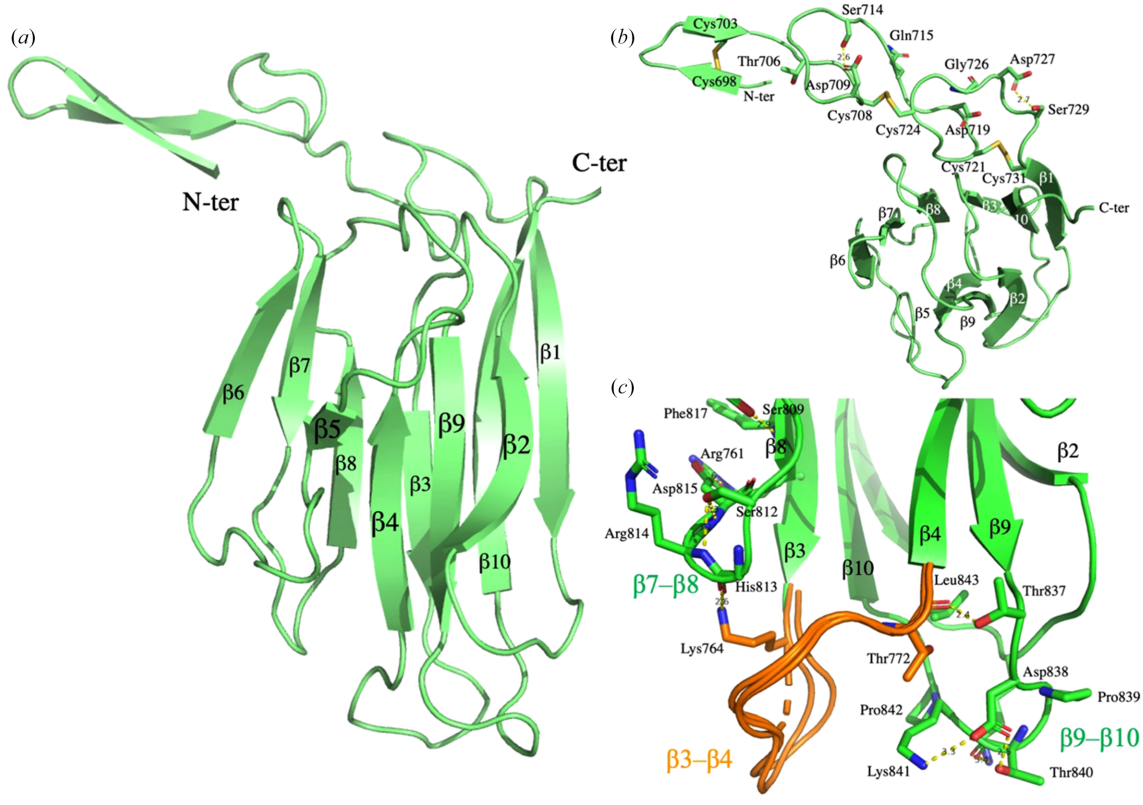


Figure 1
Crystal structure of the C_B and Sp domains. (*a*) Overall structure of the ADAMTS-5 construct comprising the N-terminal C_B domain (β-hairpin and two-rim-like superhelix) and the Sp domain (β-sandwich formed by two antiparallel β-sheets: β1–β10–β3–β8–β7–β6 and β2–β9–β4–β5). (*b*) Different orientation of the overall structure of the ADAMTS-5 construct, highlighting the C_B domain, with the three disulfide bridges and selected amino acids represented as sticks. (*c*) Close-up view of the flexible loops β7–β8, β3–β4 (orange) and β9–β10, showing some of the electrostatic interactions stabilizing the loop structures. Structural variability of the β3–β4 loop (orange) is shown by the superposition of the three molecules in the crystal asymmetric unit (the dotted segment represents the unmodelled amino-acid residues 765–767 missing in chain *C*). Figures were generated using *PyMOL* (Schrödinger).

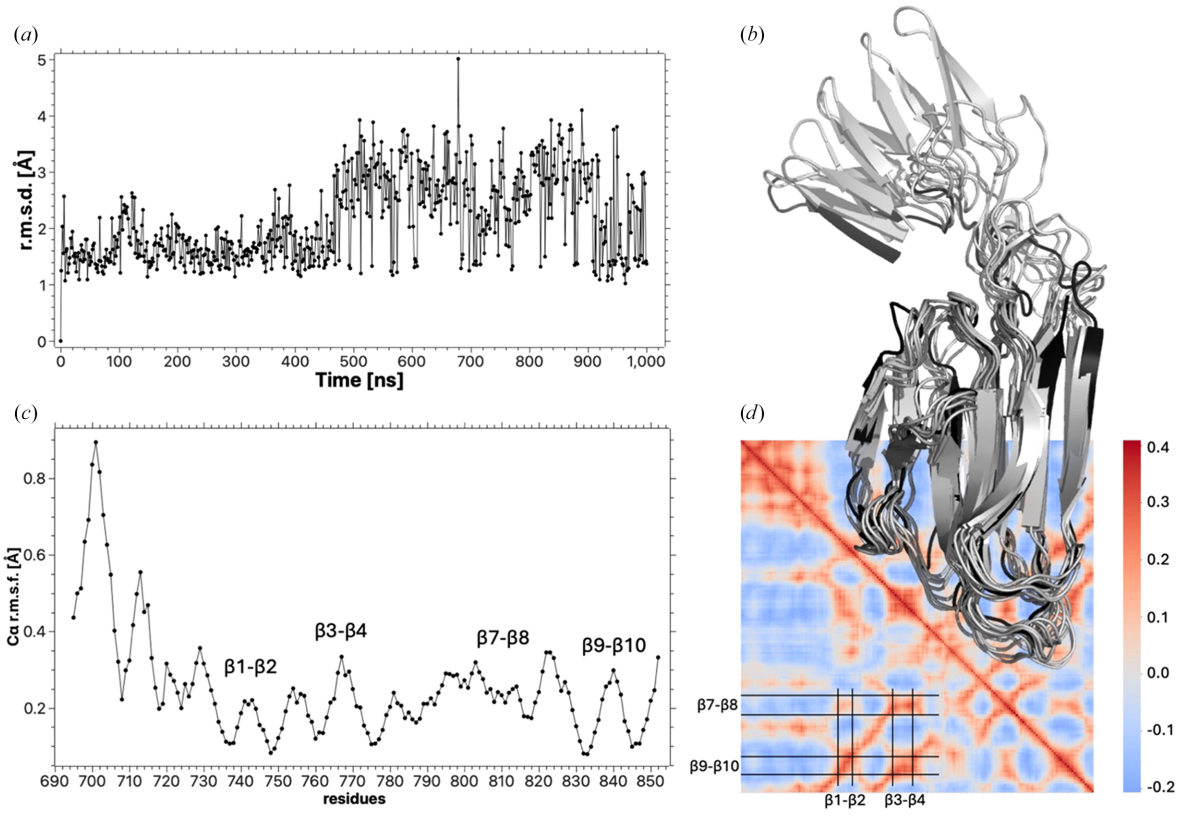


Figure 2
Molecular-dynamics simulation of the C_B and Sp domains (amino acids 695–852). (*a*) R.m.s.d. values with respect to the starting conformation during 1 µs of simulation. (*b*) Superposition of eight conformations during the simulation: the conformation at time 0 (black) and seven contiguous snapshots from 465.6 to 475.2 ns (Δt = 1.6 ns). (*c*) Root-mean-square fluctuations (r.m.s.f.s) evidencing the mobility of C^{α} during the MD simulation (the positions of selected loops are reported). (*d*) Dynamic cross-correlation maps (DCCMs) evidencing the positive covariance (dark red color) of the β3–β4 loop with β7–β8 and β9–β10 and of the β1–β2 loop with β9–β10.

3.5. Interaction model(s) of the ADAMTS-5 C_B–Sp domains with aggrecan

To investigate the possible interactions between the C_B–Sp domains of ADAMTS-5 and its proteoglycan substrate aggrecan, we used *AlphaFold* modeling. It is known that the interaction surface between ADAMTS-5 and aggrecan is primarily mediated by the loop region projecting from the β-sandwich core of the Sp domain (Santamaria *et al.*, 2019). These loops, which are prominently solvent-exposed, are likely to function as a flexible, surface-accessible exosite that mediates substrate recognition through shape complementarity and electrostatic interactions. The five *AlphaFold* models predicted different poses of the C_B–-Sp domains on aggrecan, none of which involved the loop regions (Supplementary Fig. S8). This evidence underscores the limitation of *AlphaFold*'s covariance-based predictions in capturing the dynamic interactions of a domain whose structural plasticity has evolved to transiently recognize multiple binding sites.

4. Discussion

In this work, we analyzed the structure of the human ADAMTS-5 C_B and Sp domains to understand how their architecture contributes to substrate recognition and enzymatic specificity. The crystal structure revealed a jelly-roll fold centered on a rigid β-sandwich scaffold that can maintain domain stability under physiological stress and proteolytic conditions. Surrounding this core, a set of flexible loop regions, including β3–β4, β7–β8 and β9–β10, introduces conformational adaptability suitable for substrate engagement. This architectural duality, coupling core rigidity with peripheral plasticity, enables the Sp domain to interact selectively and efficiently with structurally diverse proteoglycans such as aggrecan and versican, and reflects a conserved structural strategy within the ADAMTS family.

Structural analysis highlights the β3–β4, β7–β8 and β9–β10 loops as candidate exosites that support multisite, cooperative substrate binding, enhancing enzymatic specificity. Molecular-dynamics simulations analyzed with dynamic cross-correlation reveal communication between the selected distal loops, including coupling between β1–β2 and β9–β10, underscoring the presence of a coordinated dynamic network that may fine-tune substrate recognition. These results are consistent with the findings of Santamaria *et al.* (2019), who identified the β1–β2 and β9–β10 loops as critical exosites for substrate recognition, underscoring their essential role in defining spacer-domain specificity.

Crystal structures and molecular-dynamics simulations further reinforce this model by demonstrating that the N-terminal β-hairpin (residues 696–705) of the C_B domain remains stable, while the C-terminal C_B subdomain displays hinge-like flexibility around Thr706 (in the crystal) or Gln717 (in the MD simulation). These hinges allow variable orientation of the 'rims of spectacles' motifs, suggesting a mechanism for conformational adaptation.

Together, these structural and dynamic insights reveal an evolved binding platform capable of balancing specificity with flexibility, principles that are essential for the physiological function of ADAMTS-5. Understanding this duality provides not only mechanistic insight into substrate recognition and enzymatic efficiency, but also a foundation for the design of domain-targeted inhibitors with potential therapeutic value in osteoarthritis, atherosclerosis and other proteoglycan-related pathologies.

Acknowledgements

We thank the beamline scientists of ESRF beamlines ID23-1 and ID23-2 for support during data collection. We thank the European Union–NextGenerationEU–PNRR, MUR code IR0000011, CUP B51E22000150006, project EBRAINS-Italy for the support of computational studies and the Italian Consortium 'Cineca' (https://www.hpc.cineca.it/) for supercomputing resources allocated for the projects EIRI_E_UNIMIL (ARES-MILANI-20241213111827) and IsCc8_cea (2025–2026).

Conflict of interest

The authors declare no conflicts of interest.

Funding information

Open access publishing facilitated by Consiglio Nazionale delle Ricerche, as part of the Wiley – CRUI-CARE agreement.

References

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. É., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. & Yamashita, K. (2023). *Acta Cryst.* D**79**, 449–461.

Akiyama, M., Nakayama, D., Takeda, S., Kokame, K., Takagi, J. & Miyata, T. (2013). *J. Thromb. Haemost.* **11**, 1399–1406.

Akiyama, M., Takeda, S., Kokame, K., Takagi, J. & Miyata, T. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 19274–19279.

Apte, S. S. (2009). *J. Biol. Chem.* **284**, 31493–31497.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). *Acta Cryst.* D**60**, 2210–2221.

Brebion, F., Gosmini, R., Deprez, P., Varin, M., Peixoto, C., Alvey, L., Jary, H., Bienvenu, N., Triballeau, N., Blanque, R., Cottereaux, C., Christophe, T., Vandervoort, N., Mollat, P., Touitou, R., Leonard, P., Ceuninck, F. D., Botez, I., Monjardet, A., van der Aar, E. & Amantini, D. (2021). *J. Med. Chem.* **64**, 2937–2952.

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.

Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.

Cowtan, K. (2010). *Acta Cryst.* D**66**, 470–478.

Cuffaro, D., Ciccone, L., Rossello, A., Nuti, E. & Santamaria, S. (2022). *J. Med. Chem.* **65**, 13505–13532.

Didangelos, A., Mayr, U., Monaco, C. & Mayr, M. (2012). *J. Biol. Chem.* **287**, 19341–19345.

Dulos, J., Debruin, D. A., van der Aar, E., Chaitanya, V. B. S., Lecru, L., Campelj, D., Addinsall, A. B. & Stupka, N. (2025). *Sci. Transl. Med.* **17**, eado2112.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Fushimi, K., Troeberg, L., Nakamura, H., Lim, N. H. & Nagase, H. (2008). *J. Biol. Chem.* **283**, 6706–6716.

Gao, W., Zhu, J., Westfield, L. A., Tuley, E. A., Anderson, P. J. & Sadler, J. E. (2012). *J. Biol. Chem.* **287**, 26944–26952.

Gendron, C., Kashiwagi, M., Lim, N. H., Enghild, J. J., Thøgersen, I. B., Hughes, C., Caterson, B. & Nagase, H. (2007). *J. Biol. Chem.* **282**, 18294–18306.

Glasson, S. S., Askew, R., Sheppard, B., Carito, B., Blanchet, T., Ma, H. L., Flannery, C. R., Peluso, D., Kanki, K., Yang, Z., Majumdar, M. K. & Morris, E. A. (2005). *Nature*, **434**, 644–648.

Jin, S. Y., Skipwith, C. G. & Zheng, X. L. (2010). *Blood*, **115**, 2300–2310.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021). *Nature*, **596**, 583–589.

Kelwick, R., Desanlis, I., Wheeler, G. N. & Edwards, D. R. (2015). *Genome Biol.* **16**, 1–16.

Kemberi, M., Salmasi, Y. & Santamaria, S. (2023). *Int. J. Mol. Sci.* **24**, 12135.

Kosasih, H. J., Last, K., Rogerson, F. M., Golub, S. B., Gauci, S. J., Russo, V. C., Stanton, H., Wilson, R., Lamande, S. R., Holden, P. & Fosang, A. J. (2016). *J. Biol. Chem.* **291**, 3197–3208.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. (2010). *Proteins*, **78**, 1950–1958.
Mead, T. J. & Apte, S. S. (2018). *Matrix Biol.* **71–72**, 225–239.
Mosyak, L., Georgiadis, K., Shane, T., Svenson, K., Hebert, T., McDonagh, T., Mackie, S., Olland, S., Lin, L., Zhong, X., Kriz, R., Reifenberg, E. L., Collins-Racie, L. A., Corcoran, C., Freeman, B., Zollner, R., Marvell, T., Vera, M., Sum, P., Lavallie, E. R., Stahl, M. & Somers, W. (2008). *Protein Sci.* **17**, 16–21.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Reich, S., Puckey, L. H., Cheetham, C. L., Harris, R., Ali, A. A. E., Bhattacharyya, U., Maclagan, K., Powell, K. A., Prodromou, C., Pearl, L. H., Driscoll, P. C. & Savva, R. (2006). *Protein Sci.* **15**, 2356–2365.
Rose, K. W. J., Taye, N., Karoulias, S. Z. & Hubmacher, D. (2021). *Front. Mol. Biosci.* **8**, 701959.
Santamaria, S. (2020). *Int. J. Exp. Pathol.* **101**, 4–20.
Santamaria, S. & de Groot, R. (2020). *Open Biol.* **10**, 200333.
Santamaria, S., Yamamoto, K., Teraz-Orosz, A., Koch, C., Apte, S. S., de Groot, R., Lane, D. A. & Ahnström, J. (2019). *Sci. Rep.* **9**, 10914.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Shieh, H., Mathis, K. J., Williams, J. M., Hills, R. L., Wiese, J. F., Benson, T. E., Kiefer, J. R., Marino, M. H., Carroll, J. N., Leone, J. W., Malfait, A., Arner, E. C., Tortorella, M. D. & Tomasselli, A. (2008). *J. Biol. Chem.* **283**, 1501–1507.
Shieh, H., Tomasselli, A. G., Mathis, K. J., Schnute, M. E., Woodard, S. S., Caspers, N., Williams, J. M., Kiefer, J. R., Munie, G., Wittwer, A., Malfait, A. & Tortorella, M. D. (2011). *Protein Sci.* **20**, 735–744.
Stanton, H., Rogerson, F. M., East, C. J., Golub, S. B., Lawlor, K. E., Meeker, C. T., Little, C. B., Last, K., Farmer, P. J., Campbell, I. K., Fourie, A. M. & Fosang, A. J. (2005). *Nature*, **434**, 648–652.
Tortorella, M. D., Liu, R. Q., Burn, T., Newton, R. C. & Arner, E. (2002). *Matrix Biol.* **21**, 499–511.
Tortorella, M. D., Tomasselli, A. G., Mathis, K. J., Schnute, M. E., Woodard, S. S., Munie, G., Williams, J. M., Caspers, N., Wittwer, A. J., Malfait, A. & Shieh, H. (2009). *J. Biol. Chem.* **284**, 24185–24191.
Wight, T. N. (2017). *Matrix Biol.* **60–61**, 38–56.
Yuan, S., Bacchetti, R., Adams, J., Cuffaro, D., Rossello, A., Nuti, E., Santamaria, S. & Rainero, E. (2025). *FEBS J.* **292**, 4491–4515.
Zheng, X., Nishio, K., Majerus, E. M. & Sadler, J. E. (2003). *J. Biol. Chem.* **278**, 30136–30141.