
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Fedorov, Dmitrii; Batys, Piotr; Hayes, David B.; Sammalkorpi, Maria; Linder, Markus B.
Analyzing the weak dimerization of a cellulose binding module by analytical ultracentrifugation

Published in:
International Journal of Biological Macromolecules

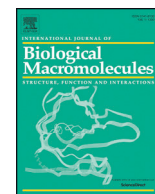
DOI:
[10.1016/j.ijbiomac.2020.09.054](https://doi.org/10.1016/j.ijbiomac.2020.09.054)

Published: 15/11/2020

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Fedorov, D., Batys, P., Hayes, D. B., Sammalkorpi, M., & Linder, M. B. (2020). Analyzing the weak dimerization of a cellulose binding module by analytical ultracentrifugation. *International Journal of Biological Macromolecules*, 163, 1995-2004. <https://doi.org/10.1016/j.ijbiomac.2020.09.054>



Analyzing the weak dimerization of a cellulose binding module by analytical ultracentrifugation

Dmitrii Fedorov^a, Piotr Batys^b, David B. Hayes^c, Maria Sammakorpi^{a,d}, Markus B. Linder^{a,*}

^a Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University, Box 16100, 00076-Aalto Espoo, Finland

^b Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Niezapominajek 8, PL-30239 Krakow, Poland

^c International Solidarity of Scientists, LLC, Gorham, NH, USA

^d Department of Chemistry and Materials Science, School of Chemical Engineering, Aalto University, Box 16100, 00076-Aalto Espoo, Finland

ARTICLE INFO

Article history:

Received 23 June 2020

Received in revised form 1 September 2020

Accepted 8 September 2020

Available online 13 September 2020

Keywords:

Analytical ultracentrifugation

Dimerization

Protein engineering

ABSTRACT

Cellulose binding modules (CBMs) are found widely in different proteins that act on cellulose. Because they allow a very easy way of binding recombinant proteins to cellulose, they have become widespread in many biotechnological applications involving cellulose. One commonly used variant is the CBM_{CipA} from *Clostridium thermocellum*. Here we studied the oligomerization behavior of CBM_{CipA}, as such solution association may have an impact on its use. As the principal approach, we used sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. To enhance our understanding of the possible interactions, we used molecular dynamics simulations. By analysis of the sedimentation velocity data by a discrete model genetic algorithm and by building a binding isotherm based on weight average sedimentation coefficient and by global fitting of sedimentation equilibrium data we found that the CBM_{CipA} shows a weak dimerization interaction with a dissociation constant K_D of $90 \pm 30 \mu\text{M}$. As the K_D of CBM_{CipA} binding to cellulose is below $1 \mu\text{M}$, we conclude that the dimerization is unlikely to affect cellulose binding. However, at high concentrations used in some applications of the CBM_{CipA}, its dimerization is likely to have a marked effect on its solution behavior.

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cellulose binding modules (CBMs) are proteins that are found as parts of many different enzymes and other proteins that interact with cellulose. Their role is to bind to cellulose and typically they do so by surface exposed aromatic side chains. It is well known that CBMs play a crucial role in cellulose degradation and the function of enzymes [1]. There are over 80 different families of functionally related, but structurally different proteins that bind to carbohydrates. They form an example of convergent evolution with different protein scaffolds achieving a similar binding function [2]. The details of these binding mechanisms differ between the individual families, although they share some common elements such as aromatic ring-pyranose stacking as part of the binding interface.

CBMs have been utilized in a large number of biotechnical applications because they provide an easy way to control the binding of proteins to cellulose. Especially the ability to use recombinant DNA technology to design new structurally engineered proteins with CBMs attached, has led to a wide range of applications. For example, the activity of enzymes can be enhanced by increasing their binding to cellulose by CBMs. In one study, different types of CBMs were used to either

enhance or inhibit the activity of lytic polysaccharides monooxygenases (LPMO) by swapping between different types of CBMs [3]. The activity of hydrolases have also been increased by fusing enzymes to CBMs [4]. Immobilization to cellulose is also an application that has been explored. CBMs can be used as a high-capacity purification tag, a targeting molecule or an affinity tag for enzyme immobilization and processing [5]. In addition, CBMs can be used to increase the rate of catalysis by creating association between biocatalysts and substrates [6,7]. Other applications include cell immobilization [8], diagnostics, and even the construction of protective devices against nerve gas [9].

Because CBMs attach to cellulose surfaces, they can also be used to modify the properties of cellulose based materials. An example of this is improving the mechanical properties of paper. A protein constructed to contain two CBMs at each end of a linker resulted in paper with higher folding endurance [10]. In another example, a recombinant silk-protein based adhesive for cellulose depended on the binding of CBMs [11].

Here we focus on the CBM_{CipA} from the *Cellulomonas fimi* cellulosome anchoring protein CipA. CBM_{CipA} belongs to family III according to the Cazy-classification [2]. CBM_{CipA} has a nine-stranded β sandwich like structure with a jellyroll topology and belongs to surface binding type of CBMs (type A). It has a face that contains a planar linear strip of aromatic and polar residues, which participates in interaction with crystalline cellulose [12].

* Corresponding author.

E-mail address: markus.linder@aalto.fi (M.B. Linder).

CBM_{CipA} shows a very strong binding to crystalline cellulose, with a dissociation constant (K_D) of 0.6 μM [13–15]. This is an order of magnitude higher affinity than, for example the fungal Family I type of CBMs have [16]. The CBM_{CipA} has a big practical value in applications due to its high affinity to cellulose. Because of its wide use and because a potential multimerization behavior could influence the use of the protein in applications, we decided to investigate the multimerization characteristics of CBM_{CipA} in more detail. Furthermore, a possibility of dimerization was indicated in our previous work using CBM_{CipA} as a fusion partner in silk-material forming proteins. Fusion proteins containing the CBM_{CipA} seemed to associate with each other stronger than other proteins that did not contain the CBM_{CipA}. This stronger self-association effected the way the assembled into different protein-based materials [11,17,18]. For CBM_{CipA}, no detailed information of multimerization is available, but the formation of dimers has qualitatively been reported for the closely related CBD_{Cex} from the enzyme Cex, a β -1,4-glycanase from *C. fimi* [19]. Moreover, a cellobiose-mediated dimerization of a CBM belonging to class 29 has been shown [20]. As a method to study the solution interactions of the CBM_{CipA} we chose analytical ultracentrifugation (AUC) which is one of the most powerful techniques available for quantifying weak interactions [21]. Experiments performed by AUC are either based on analyzing sedimentation velocity or sedimentation equilibrium. We analyzed the CBM_{CipA} by multiple approaches to obtain the K_D , as well as diffusion coefficients and information on the anisotropy of the protein in solution [22]. Molecular dynamics (MD) simulations were done to widen the understanding of the system.

2. Materials and methods

2.1. Protein expression

The cloning, purification, and characterization of CBM_{CipA} has been described previously [15]. It contained a 6xHis-tag in the C-terminus. The sequence of the CBM_{CipA} was:

E. coli strain BL21 (DE3) (ThermoFisher Scientific) was used for CBM_{CipA} production. CBM_{CipA} was expressed using MagicMedia (ThermoFisher Scientific) expression medium in accordance to the manufacturers protocol. Cell lysis was performed with a high-pressure homogenizer EmulsiFlex. The purification process consisted of two parts, first precipitation of contaminating bacterial proteins and debris by heating and then affinity purification of CBM_{CipA} by immobilized metal affinity chromatography (IMAC). The sample was heated up to 70 °C for 30 min, and then clarified by centrifugation at 5000 rpm for 20 min. An ÄKTA-Pure liquid chromatography system (GE Healthcare Life Sciences) with a 5 mL HisTrap column was used for IMAC of the supernatant. Gel filtration was used after the IMAC to transfer CBM_{CipA} into suitable buffers for subsequent experiments.

2.2. Analytical ultracentrifugation

Analytical ultracentrifugation experiments were performed using a Beckman Coulter Optima AUC. Sample cells with 12 mm Epon centerpieces and sapphire windows were used. Sample cells were cleaned with deionized water (13.0 $\text{M}\Omega \cdot \text{cm}$ resistivity), 20% ethanol, and 1% Hellmanex detergent. All measurements were made in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 , and 1.8 mM KH_2PO_4 , pH 7.35). The sample volumes were 400 μL . All experiments were performed at 20 °C and prior to each experiment the system was thermally stabilized for 1 h 30 min (except sedimentation equilibrium experiments).

In sedimentation velocity experiments, UV–Vis absorbance at 280 nm and interference detection were used. The experiments were carried out at 50000 rpm using an An-50 rotor. CBM_{CipA} samples were measured in the concentration range from 0.05 mg/mL to 0.7 mg/mL (3–38 μM) that corresponds to 0.1–1.3 OD range (280 nm absorbance)

and from 1 to 5 mg/mL (55–270 μM) measured by interference. To get complete sedimentation and sufficient amount of data points, 400 scans per each sample were collected with a frequency of one scan every 80 s.

Sedimentation equilibrium experiments were performed using UV–Vis absorbance at 280 nm and an An-60 rotor. Measurement were conducted at three loading concentrations: 0.1, 0.25, and 0.4 mg/mL (5, 14, 22 μM) and at four rotor speeds: 27,000, 35,000, 40,000, and 45,000 rpm.

Speeds were chosen to have the reduced buoyant molecular weight of the sedimenting species in the range of 1–5 to provide optimal steepness in the exponent graph. The time until equilibrium and the reduced buoyant molecular weight were estimated according to [23].

2.3. Data analysis by Ultrascan

Sedimentation velocity absorbance data analysis were performed using the software UltraScan III (v. 4.0 revision 2466) [24]. Rotor-stretching calibration and chromatic aberration corrections were applied during data import. A partial specific volume of 0.7148 mL/g was used. It was based on the amino acid sequence and calculations by the estimation tool implemented in Ultrascan. A density of 1.0056 g/mL and viscosity of 1.0120 mPa·s for the buffer were obtained from the tabulated values in Ultrascan. During data editing, the meniscus position and end cutoff distance were entered manually. The plateau position corresponding to a uniform concentration prior to acceleration and the baseline buffer absorbance were determined automatically.

Concentration profiles obtained in the sedimentation velocity experiments were treated using 2-dimensional spectrum analysis (2DSA) [25]. We used 2DSA to perform primary analysis coupled with simultaneous time-invariant and radial-invariant noise reduction to obtain values for frictional ratios (f/f_0), and sedimentation coefficients (S). The grid parameters used were 1 to 10 for S and 1 to 4 for f/f_0 with 64 grid points in each direction. The choice of parameters was based on time-derivate analysis and was selected to cover the distribution of parameters for all species in solution. The 2DSA was done in three steps. First the time-invariant noise was removed after which the meniscus was fitted. The meniscus fit range was 0.03 cm around the previously chosen position. In the meniscus fitting procedure, 10 positions were chosen, and corresponding root mean square deviation (RMSD) values were approximated by second order polynomial to determine the position with lowest RMSD. The position with lowest RMSD was used for further analysis. The third step was an iterative 2DSA refinement consisting of repetitions with improved meniscus position and fitting of the noise profile. The number of iterations was 10, which was sufficient for convergence of the fit. The parameters of the grid were the same in all steps.

The 2DSA results and noise reduction profile obtained by the analysis described above were used for further analysis. A genetic algorithm (GA) and a genetic algorithm Monte Carlo (GAMC) [26] optimization were used to extract the main populations of frictional ratio and sedimentation coefficient combinations present in the data. The existence of reversible self-association was analyzed using the van Holde-Weischet (vHW) [27] method. The dissociation constant (K_D) was determined by a discrete model genetic algorithm (DMGA). The specific parameters used in GA, GAMC, vHW and DMGA are provided in the supplemental information. Hydrodynamic parameters of monomer and dimer were estimated using the UltraScan Solution Modeler (US-SOMO) [28]. Calculations were performed using the ZENO hydrodynamic computation algorithm [29,30] and the van der Waals overlap bead model [31].

2.4. Data analysis by Sedfit

$C(s)$ [32] and $l_s-g^*(s)$ [33] analysis of sedimentation velocity absorbance and interference data were performed using Sedfit version

16.1c [34]. We loaded every fifth of the first 350 scans, in total 70 scans per loading concentration for both $c(s)$ and $ls-g^*(s)$ type of analysis. The values for partial specific volume, buffer density, and viscosity were obtained from Ultrascan, as described above. The resolution was 100. Analysis and fitting of isotherms of weight-average sedimentation coefficients was carried out according to the modified Hill equation [35]:

(1)

where $S(c)$ is the weight-average sedimentation coefficient at concentration c , S_{mon} is the sedimentation coefficient of the monomer, and S_{dim} is the sedimentation coefficient of dimer.

2.5. Data analysis of sedimentation equilibrium data

Analysis of sedimentation equilibrium data was carried out by Sedphat version 15.2b [34]. Fitting parameters included local concentrations and dissociation constants, while the molecular weight was fixed. Sedfit was used for initial data inspection.

2.6. Molecular modelling

All-atom molecular dynamics (MD) simulations were carried out by using the Gromacs software (version 5.1.4) [36,37]. We used the particle Mesh Ewald (PME) electrostatics calculation scheme [38], and a non-bonded interactions cut-off of 1 nm. The modelling was performed with the Amber03ws force field [39] and TIP4P2005 water model [40]. Two systems were studied, one with a single CBM_{CipA} molecule and one with two CBM_{CipA} molecules in the simulation box. A cubic box of 8 nm × 8 nm × 8 nm in initial size was used for the single CBM_{CipA} system and for the two molecule-system we used a box size of 12.5 nm × 12.5 nm × 12.5 nm (Fig. 1).

We used the crystal structure PDB ID: 1NBC as the initial structure for CBM_{CipA}. The data were modified to leave only protein itself. The structural model differed from the actual protein by the H₆-tag and two amino acids in the beginning and at the end of the sequence. The

structure for the two-molecule system was obtained from the final frame of the single CBM_{CipA} simulation. The structure was cut with a 1 nm water shell. The initial arrangement of the two molecules relative to each other was chosen based on the location of the β -sheets and the geometric shape. After this, the proteins were solvated with explicit water molecules. Neutralization of the net charge in the system was performed by replacement of water molecules by Na⁺ ions. Both systems were energy-minimized. After initial equilibration, the simulations were performed in NPT (isothermal-isobaric) ensemble. The temperature and pressure control were carried out by the V-rescale thermostat [41] with a time constant of 0.1 ps and the Parrinello-Rahman barostat [42] with a time constant of 2 ps. The temperature and pressure reference levels were set at 300 K and 1 bar. The visualization was done with the VMD software (v. 1.9.4a9) [43]. The main simulation was conducted for 160 ns for the single CBM_{CipA} and for 450 ns for the two CBM_{CipA} system. In the dimer simulation the first 150 ns were disregarded in the analysis as dimer formation period.

Diffusion coefficients were calculated using the mean square displacement (MSD). The stable linear region on the MSD plot was approximated by a straight line according to the equation

(2)

where $\langle |r(t) - r(0)|^2 \rangle$ is the MSD, D is the diffusion coefficient and t is the time.

3. Results and discussion

3.1. Analysis of reversibility

Our initial question was to determine the presence of self-association for CBM_{CipA} in the chosen concentration range. For this we used, two ways of treating the data, the vHW and the $ls-g^*(s)$ methods. The methods are implemented in Ultrascan and Sedfit, respectively. The vHW method is based on a graphical transformation of the sedimentation velocity experimental data and allows obtaining a distribution of sedimentation coefficients that is independent of diffusion. By the vHW method, one can differentiate between non-interacting and self-associating species and one can also identify boundary spreading due to diffusion and heterogeneity in the sedimentation coefficient. If the kinetics of interactions are slow compared to the time of the experiment or if the system is non-interacting, the interacting components will be separated by the centrifugal force and the boundaries of the concentration profiles will be separated. However, in the case of relatively fast kinetics, the components have sufficient time to re-equilibrate and a single boundary would be observed. Increasing the sample concentration results in changes in the ratios between species, and therefore in differences in the boundary shape. These differences represent a shift toward higher sedimentation coefficients and higher partial concentrations of oligomers. For CBM_{CipA}, vHW distributions demonstrated self-association behavior (Fig. 2a), showing a shift of the distribution to higher values of sedimentation coefficients with an increase of the loading concentrations. The presence of a shift indicated clearly that the CBM_{CipA} undergoes reversible self-association [22].

The least-squares apparent sedimentation coefficient distribution, i.e. the $ls-g^*(s)$, method was used as an alternative approach to analyzing reversible association. The method is based on calculating the apparent sedimentation coefficient distribution based on direct least-squares boundary modelling. A shift in the distributions as a result of concentration changes serves as a robust indication of protein self-association [44]. Moreover, in the case of reversible association, the distributions should follow a certain pattern: the height of the peak should decrease, and width should increase when approaching the K_D concentration. However, at

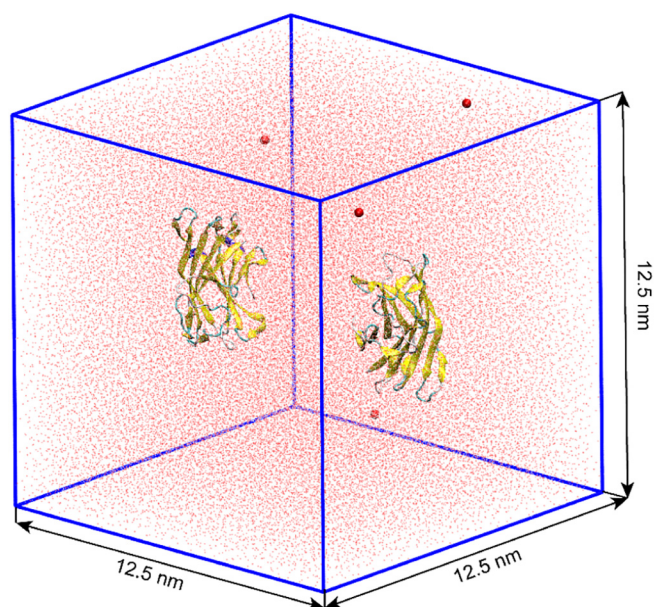


Fig. 1. The initial all-atom molecular dynamics simulation configuration of two CBM_{CipA} molecules in the simulation box. The explicit water molecules are shown as red dots and the Na⁺ ions that are required for charge neutrality as red spheres. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

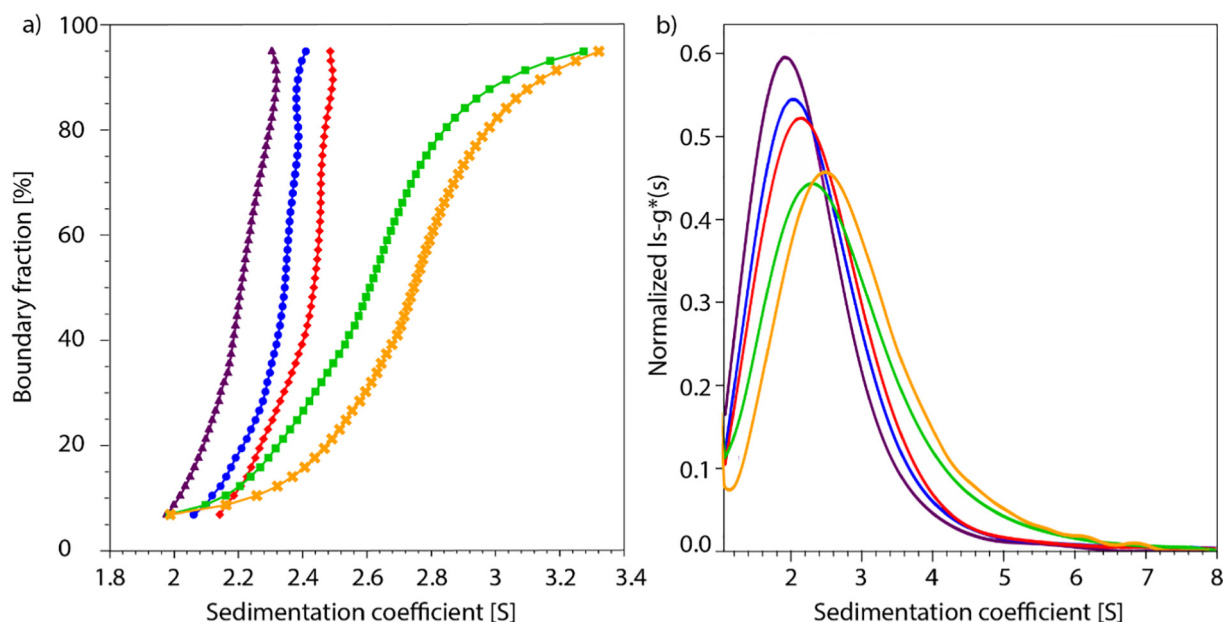


Fig. 2. a) A van Holde-Weischet integral distribution plot for different loading concentrations: 0.05 mg/mL (3 μM) – violet triangles, 0.3 mg/mL (16 μM) – blue circles, 0.6 mg/mL (32 μM) – red diamonds, 1.85 mg/mL (100 μM) – green squares and 4.93 mg/mL (267 μM) – yellow crosses. The shift toward higher sedimentation coefficients with an increase of loading concentration and the half-parabolic shape of distributions clearly indicate the presence of reversible self-association in the system. b) ls-g*(s) distribution: 0.05 mg/mL (3 μM) – violet line, 0.3 mg/mL (16 μM) – blue line, 0.6 mg/mL (32 μM) – red line, 1.85 mg/mL (100 μM) – green line, 4.93 mg/mL (267 μM) – yellow. A clear shift of the curves toward higher S values correspond to a reversible self-association of the components. At the highest concentration the system has switched to a state where over half of the protein is as dimers. Therefore, we can predict that the K_D value is between 32 and 267 μM.

concentrations higher than the K_D concentration the height of the peak should increase and its width decrease. This pattern represents the successive changes in composition of the system with concentration. There is a shift from a state with mostly monomers to states with increasing fractions of oligomers. All mentioned features were found in ls-g*(s) distributions shown in Fig. 2b.

3.2. Initial analysis of data and estimation of parameters

We continued to analyze data in order to establish parameters for possible monomers and oligomers, and to further choose the appropriate interaction model. We used the two programs, Ultrascan and Sedfit, to obtain starting parameters and to perform qualitative analysis.

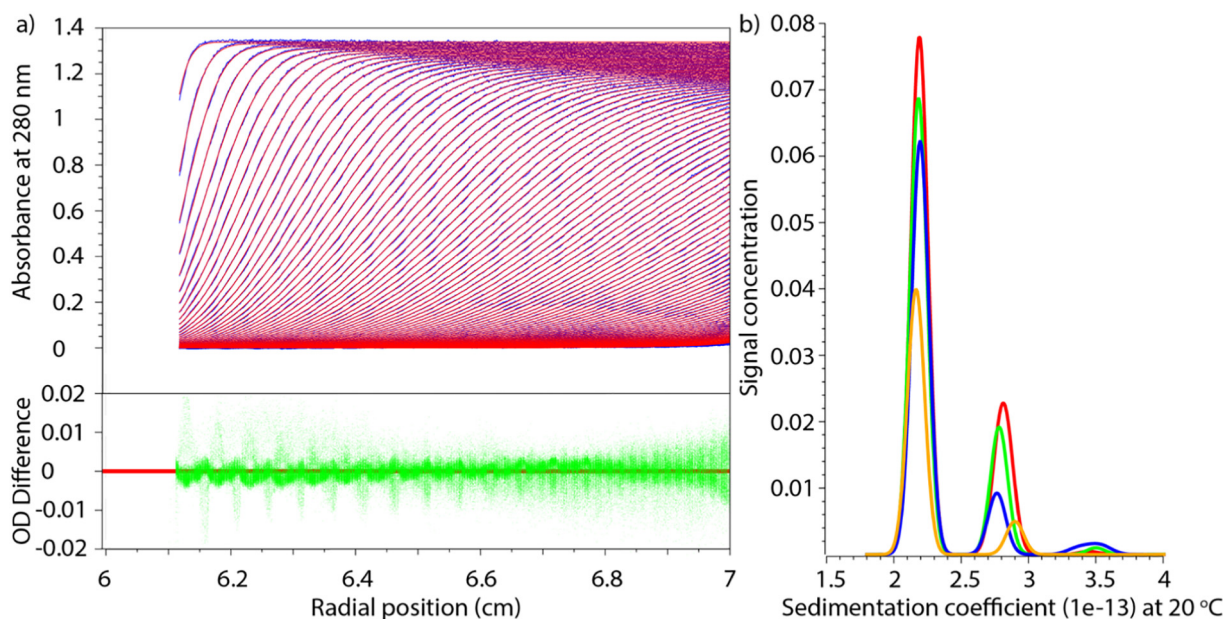


Fig. 3. a) Data from sedimentation velocity experiments (blue lines) and simulated data based on the GA analysis (red lines). Residuals are shown in green in the panel below. Experimental and simulated data are right on top of each and residuals are random and not intense, indicating a high precision of the fit. b) GA analysis results shown as sedimentation distribution plots at four concentrations: 0.3 mg/mL (16 μM) – yellow, 0.5 mg/mL (27 μM) – blue, 0.6 mg/mL (32 μM) – green and 0.7 mg/mL (38 μM) – red lines. The signal concentration relates to the optical concentration of each component.

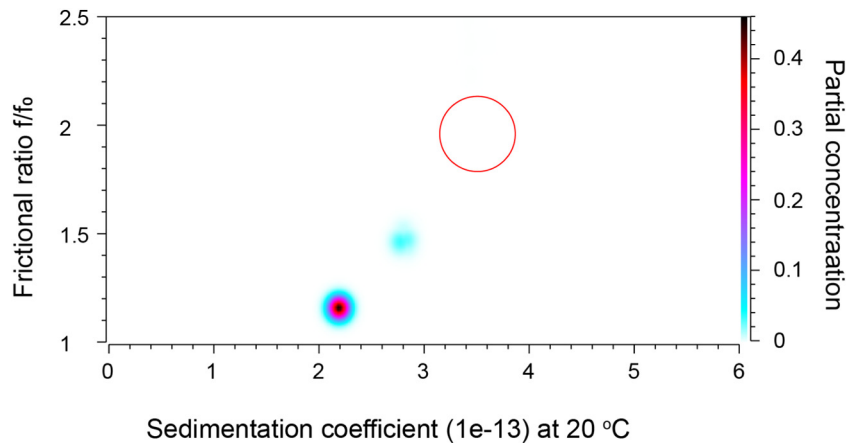


Fig. 4. Pseudo 3D plot of the CBM_{CipA} solution composition at a concentration of 0.7 mg/mL (38 μM) after 64 Monte Carlo iterations by GAMC. The pseudo 3D distribution connects sedimentation coefficient *S*, frictional ratio *f/f*₀, and partial concentrations of the solution. The graph shows two peaks, a monomer-like at 2.15S and *f/f*₀ ≈ 1.15 and a dimer-like at 2.8 S and *f/f*₀ ≈ 1.5. The red circle marks the area where 3.5S peak would be expected to be found.

Starting parameters were then taken for further analysis using methods that have been designed to take reversible interactions between molecules into account and allow to collect quantitative information about interacting system. Ultrascan has multiple methods for data analysis. We used two of them: 2DSA and GA and their Monte-Carlo versions. Both methods are intended for non-interacting or non-reversible interacting systems and therefore in the case of reversible system can be used only for estimation of parameters. After initial data treatment and noise reduction by 2DSA, GA was used for further data refinement. The sedimentation profile, fitted data, and residuals at 0.7 mg/mL (38 μM) concentration are shown in Fig. 3a. The GA analysis of datasets at concentrations higher than 0.2 mg/mL (11 μM) revealed two clearly identifiable peaks at 2.15 S and 2.8 S (Fig. 3b). These peaks correspond to molecular weights of 18 and 37 kDa, respectively.

Next, we conducted a GAMC analysis to evaluate the quality of data fitting and to eliminate peak artefacts due to random noise. In particular, the peak at 3.5S did not show good convergence, indicating that it was a result of background noise. No significant changes in the two main peaks at 2.15 S and 2.8 S were observed in the pseudo 3D distribution after 64 Monte Carlo iterations (Fig. 4).

In addition, GAMC was used to estimate sedimentation coefficients, diffusion coefficients, frictional ratios, and partial concentration values of the monomers and oligomers in the range of concentrations from 0.05 to 0.7 mg/mL (3–38 μM) (Table 1). The sedimentation velocity fitting showed low RMSD values. However, all parameters showed strong concentration dependency and variation. This variation is likely to be due to the system being interacting while the model is based on

a non-interacting model. Moreover, it should be noted that the values listed should be seen as fitting (or apparent) parameters. In general, the best agreement is achieved at low and at high concentration, when mostly one of the states prevail: monomer or dimer.

We obtained similar results using the Sedfit *c(s)* distribution. Results are shown in Figs. 5 and 6.

Both data fitting with the Sedfit *c(s)* non-interacting model and with Ultrascan GA showed separation into two peaks, but these analyses are suitable only for irreversible systems or systems with slow kinetics. The analysis is thus potentially insufficient since above we obtained *vHW* and *ls-g*(s)* distributions that clearly showed that the system was reversibly interacting.

Since both GAMC and *c(s)* showed two peaks in the range between 2 and 3 S, we concluded that the simplest and most appropriate model to investigate further is a monomer-dimer interaction model. Hydrodynamic parameters of the monomer were taken from the dataset with the lowest concentration to minimize the influence of the dimer fraction and were: *S* = 2.1S, *f/f*₀ = 1.26.

3.3. Molecular dynamics modelling

Molecular dynamics simulations were conducted as a complementary approach to study molecular parameters. Two types of systems were investigated: one with a single CBM_{CipA} molecule and one system with two CBM_{CipA} molecules in the simulation box. The first system was used to estimate the hydrodynamic properties of the CBM_{CipA} molecule as a monomer.

Table 1
Sedimentation velocity fitting results by Ultrascan for 7 different loading concentrations. Parameters of monomer-like and dimer-like peaks and RMSD of the fitting are shown. Data for the dimer-like peak is shown only for the concentrations where the peak was clearly identifiable. Parameters are normalized to 20 °C and water as solvent.

Conc. mg/mL	Monomer fraction				Dimer fraction				
	<i>S</i> * 10 ^{−13} s	<i>D</i> cm ² /s	<i>f/f</i> ₀	Fraction	<i>S</i> * 10 ^{−13} s	<i>D</i> cm ² /s	<i>f/f</i> ₀	Fraction	RMSD
0.05	2.09	9.8	1.26	95.4%					0.00086
0.1	2.13	10.4	1.21	95.6%					0.00102
0.2	2.15	10.4	1.19	91.4%					0.00119
0.3	2.17	10.6	1.16	88.5%	2.90	5.5	1.68	11.0%	0.00144
0.5	2.19	11.0	1.16	82.9%	2.76	5.3	1.74	12.6%	0.00209
0.6	2.18	10.9	1.16	76.8%	2.78	6.9	1.46	21.6%	0.00248
0.7	2.19	10.9	1.16	74.7%	2.81	6.8	1.47	24.0%	0.00269
Average	2.16	10.5	1.2		2.8	6.1	1.6		

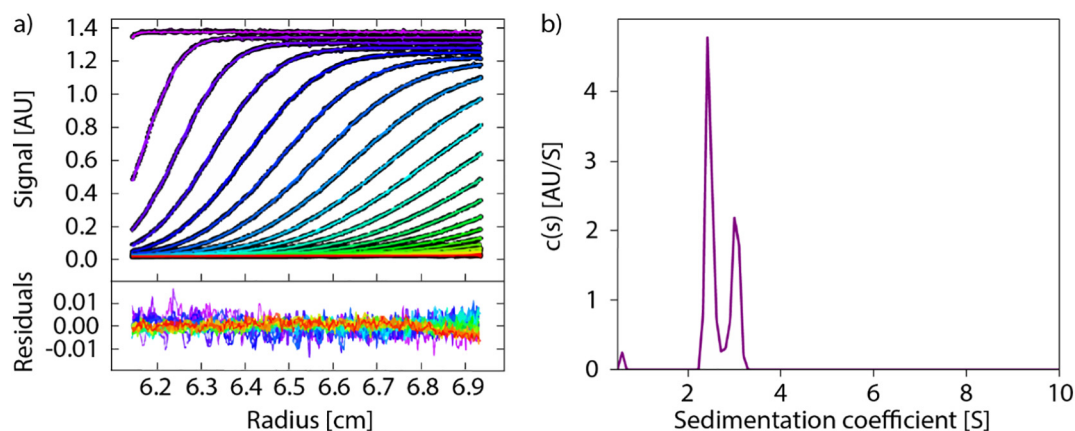


Fig. 5. a) Absorbance sedimentation boundaries of 0.6 mg/mL (32 μ M) CBM_{CipA} in BSA buffer pH 7.35 at 50000 rpm (circles) and best-fit solution by $c(s)$ model (lines). Color gradient represents the time from beginning of the experiment (violet) to the end (red). b) $c(s)$ distribution showing two peaks at approximately monomer and dimer positions.

From this, a diffusion coefficient of $7.9 \cdot 10^{-7} \pm 0.5 \cdot 10^{-7} \text{ cm}^2/\text{s}$ was calculated. Compared to the diffusion coefficient at lowest concentration of $9.8 \cdot 10^{-7} \pm 0.1 \cdot 10^{-7} \text{ cm}^2/\text{s}$, deviation is about 20%. We consider the agreement as sufficiently close, taking into account the widely different approaches used.

The second system allowed us to study the process of CBM_{CipA} dimerization. Initially, the CBM_{CipA} molecules were apart from each other in the simulation box (Fig. 1) but a stable dimer form was formed after 150 ns of MD simulations. The dimer form did not change until the end of the simulation, i.e. the next 300 ns (Fig. 7). The CBM_{CipA} dimer diffusion coefficient (D) was determined to be $6.2 \cdot 10^{-7} \pm 1.6 \cdot 10^{-7} \text{ cm}^2/\text{s}$, which is in very good agreement with the measured one of $6.1 \cdot 10^{-7} \pm 0.8 \cdot 10^{-7} \text{ cm}^2/\text{s}$. The single monomer and dimer in the simulation box corresponds to the concentration 6 mg/mL and 3.1 mg/mL, respectively. The latter concentration is significantly above the concentration range studied in the experiments. This suggests, that the D determined from the MD simulations should be rather compared with the 0.7 mg/mL diffusion coefficient value, i.e. $6.8 \cdot 10^{-7} \text{ cm}^2/\text{s}$, and not the one for dilute solution.

For both, the monomer and the dimer, the D calculated from the MD simulations is underestimated by 12–20%. This might originate from the oversimplifications in the modelling methodology. For example, the D of the TIP4P/2005 water which is used in this work, is $\sim 10\%$ smaller than experimental values [40]. Nonetheless, the MD method clearly reflects the decrease in D caused by dimerization.

A residue contact map (RCM) (Fig. 8) was calculated from the two CBM_{CipA} simulation trajectory to understand which residues are involved in the dimer formation in the simulated system. The RCM shows the mean distance between the CBM_{CipA} residues, that is, each point on the map shows the distance between two residues, as specified by the residue index. We plotted the two CBM_{CipA} molecules with consecutive numbering, meaning that the lower left quadrant shows interactions within molecule 1, the upper right quadrant shows interactions within molecule 2, and both the lower right and upper left show interactions between molecules. Notably, an RCM is symmetric with respect to the diagonal from lower left to upper right. Residues on this line have a distance equal to zero, since each residue on this line is compared to itself. Square areas through which the zero-line passes display the distances between residues inside the molecule. The lines parallel and perpendicular to the zero line inside these areas identify parallel and anti-parallel β -sheets. The distance between residues belonging to different molecules is shown in square areas perpendicular to the zero line. The RCM highlights at least two areas which can participate in the interactions between the CBM_{CipA} molecules. The first area represents the interaction between residues Met-81 and Ser-82 of one CBM_{CipA} and Gly-65 of the other CBM_{CipA}. The second area shows interaction between Gln-48 of first molecule and group of residues Ser-63, Glu-102, Pro-103, Ala-105 belonging to the second molecule.

We used the program SOMO to estimate the results of the sedimentation velocity experiments. The SOMO-estimations were based on the

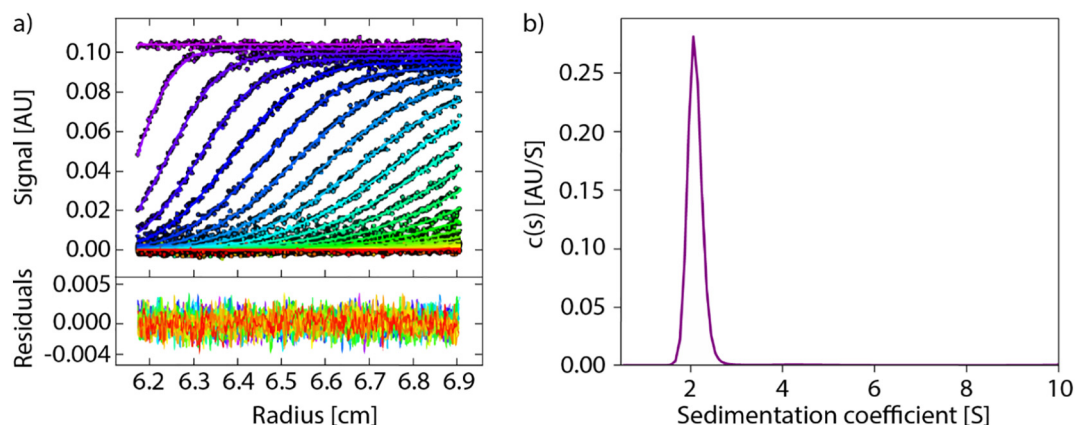


Fig. 6. Left: Absorbance sedimentation boundaries of 0.05 mg/mL (3 μ M) CBM_{CipA} in BSA buffer pH 7.35 at 50000 rpm (circles) and best-fit solution by $c(s)$ model (lines). Color gradient represents the time from beginning of the experiment (violet) to the end (red). Right: $c(s)$ distribution, that shows only one peak at $S \approx 2.1 \text{ S}$.

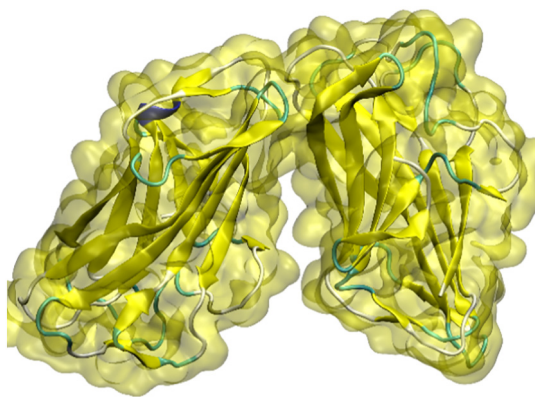


Fig. 7. CBM_{CIPa} dimer structure as obtained by the MD simulations. The snapshot corresponds to the final configuration at 450 ns.

crystal structure for the monomer and the molecular dynamic simulations for the dimer. Results are presented in Table 2 and show that experimental and predicted values are in good agreement for the monomer but were slightly off for the dimer. Both sets of SOMO data show overestimation of the sedimentation and diffusion coefficients and underestimation of the frictional ratio. The difference between experimental and SOMO results could be explained by the absence of a poly-His tag and a few other amino acids in the crystal structure or by tight packing of proteins in the crystal, that led to increase in diffusion coefficient. In the case of the dimer, an inexact partial specific volume (could also contribute).

3.4. K_D determination

AUC allows different routes toward analyzing interactions and estimating K_D . Here, we compared three different approaches: 1) global fitting in Sedphat, 2) DMGA method in Ultrascan and 3) building of isotherm on basis of Sedfit fitting. Together these methods allow obtaining quantitative information about the system.

Table 2

Comparison between experimental results at concentration 0.05 mg/mL (3 μ M) for monomer and at concentration 0.7 mg/mL (38 μ M) for dimer and SOMO modelling.

	$S \cdot 10^{-13} \text{ s}$	$D, \text{ cm}^2/\text{s}$	f/f_0	
Monomer				
Experimental	2.1	9.8	1.26	0.7148
SOMO	2.23	11.1	1.13	0.7180
Dimer				
Experimental	2.81	6.8	1.47	0.7148
SOMO	3.08	8.0	1.26	0.71

Global fitting of sedimentation equilibrium data on sedimentation equilibrium experiments were conducted to study the oligomerization process in absence of kinetics effects (Fig. 9a, b, and c). In the experiments, we used three loading concentrations, 0.1 mg/mL (5 μ M), 0.25 mg/mL (13 μ M), and 0.4 mg/mL (20 μ M) in combination with four speeds: 27000 rpm, 35,000 rpm, 40,000 rpm, and 45,000 rpm. Equilibrium data was collected after 46, 24, 20, and 20 h respectively for each speed (Fig. 9d). The dimerization model in Sedphat was the simplest model showing the best fit for the data. Adding more complexity did not improve fitting results. A global fit of the obtained equilibrium data at three loading concentrations and four speeds to monomer-dimer self-association model provided the dissociation constant K_D of 99 μ M.

An analysis using the DMGA implemented in Ultrascan gave K_D values that differed slightly with loading concentrations (Table 3). The analysis showed overall lower values than described above, and had also a larger variability. The shift in values and variability in results probably reflect the effect of noise in the data. Nonetheless, the analysis reassures the overall order of magnitude for the K_D .

We further analyzed the data through isotherms of weight-average sedimentation coefficients. The weight-average sedimentation coefficients were calculated by integration of sedimentation coefficient distribution obtained by the $c(s)$ method in Sedfit. Eq. (1) was then used for isotherm fitting. We fixed the value of S at 2.1 S for the monomer and fitted the values for S for the dimer and K_D . Fitting of the binding isotherm based on the weight-average sedimentation coefficients gave $S = 2.8$ S for dimer and the value of $K_D = 71 \mu\text{M}$ (Fig. 10).

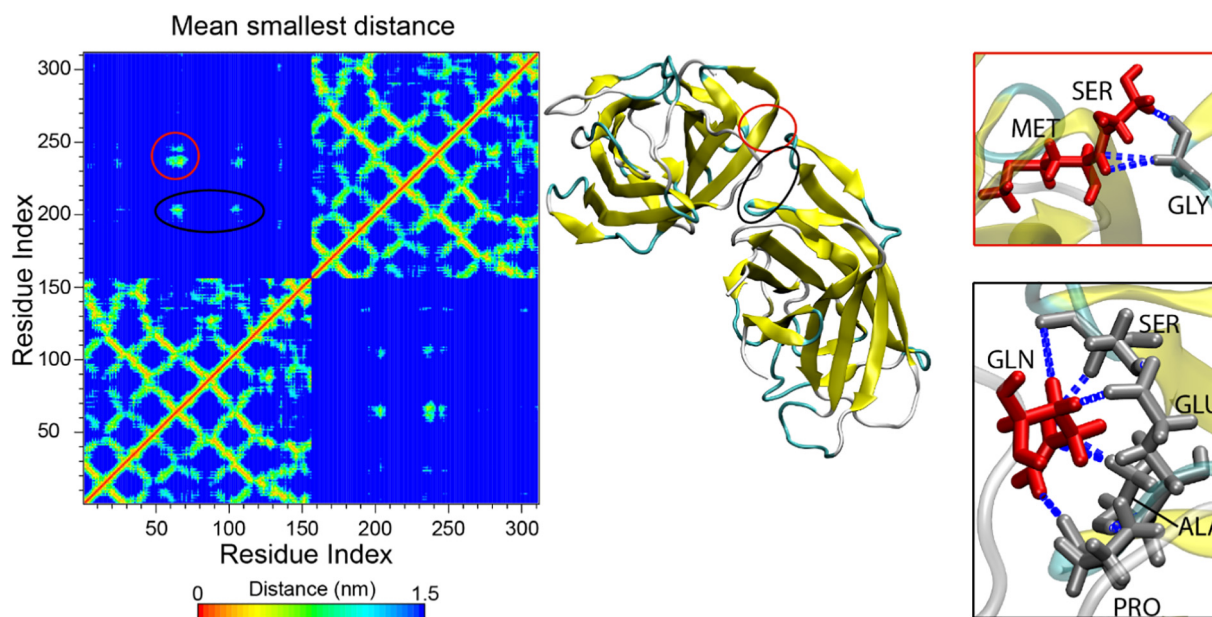


Fig. 8. Residue contact map calculated for the two CBM_{CIPa} molecular dynamics simulations and the corresponding dimer configuration with contact regions zoomed in. Red and black circles show areas of contact between molecules. Hydrogen bonds are shown by blue dot lines.

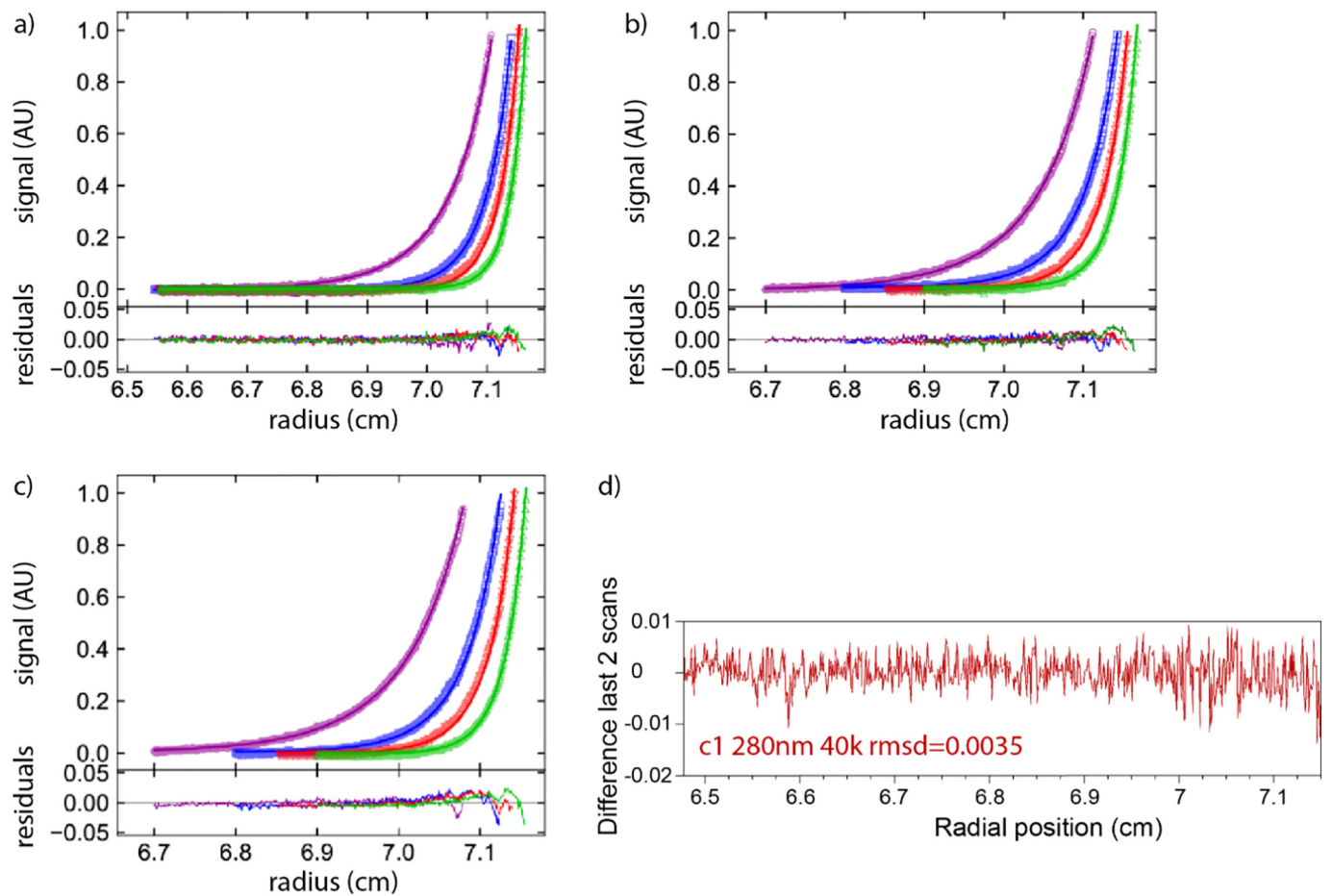


Fig. 9. Sedimentation equilibrium distributions of three loading concentrations: a) 0.1 mg/mL (5 μ M), b) 0.25 mg/mL (13 μ M) and c) 0.4 mg/mL (20 μ M) and at four speeds: 27000 rpm (violet circles), 35,000 rpm (blue squares), 40,000 rpm (red down triangles), 45,000 rpm (green up triangles). Fitting curves shown by lines of corresponding color. d) Difference between 2 last scans for 0.1 mg/mL (5 μ M) loading concentration at 40000 rpm. The difference is on level of random noise shows that equilibrium was achieved.

4. Conclusions

We showed that the CBM_{CipA} form dimers with a weak K_D of $90 \pm 30 \mu\text{M}$. In a highly conservative interpretation of the combined data, the CBM_{CipA} K_D range could be 50–200 μM . A K_D lower or higher than this range would cause significant changes in AUC data, which would have clearly been observed. Each of three methods: building of isotherms on basis of Sedfit fitting, DMGA method in Ultrascan, and global fitting in Sedphat have their own drawbacks, and solely looking at only one of them would not have given an equally wide understanding of both dynamics and affinity of the interaction. However, together the results obtained by these methods are not contradictory and their K_D values are close, allowing us to be confident in the stated K_D range for CBM_{CipA}.

Table 3
The value of the dissociation constant K_D as determined at different concentrations and using DMGA for analysis of data.

Conc., mg/mL (μ M)	K_D , μ M	95% confidence interval	
		lower	upper
0.20 (10.83)	81.3	68.8	93.8
0.30 (16.24)	86.7	76.9	94.5
0.50 (27.07)	93.9	81.3	106.6
0.60 (32.49)	92.3	83.4	101.1
0.70 (37.90)	80.5	68.8	92.2
Average	87.0		

The interaction with K_D about 90 μM can be classified as weak, on the edge of ultra-weak. Typically interactions that have a K_D over 100 μM are classified as ultra-weak [21]. The K_D corresponds to a half of the

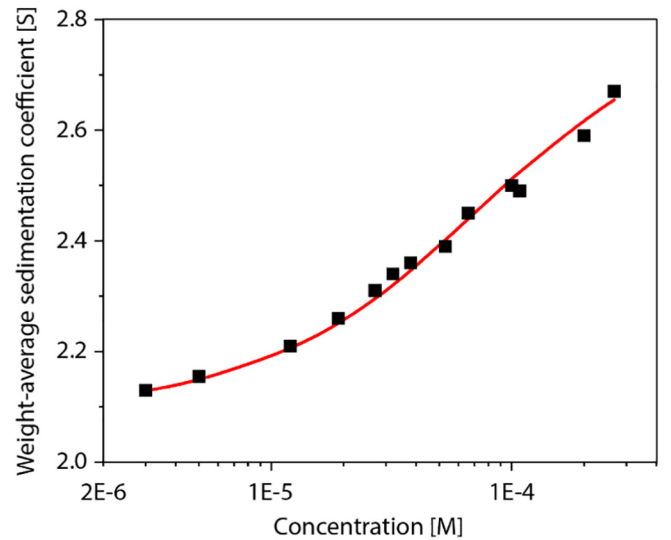


Fig. 10. Weight-average sedimentation coefficients in 0.05–5 mg/mL (3–270 μM) concentration range (black squares) and expected binding isotherm for $K_D = 71 \mu\text{M}$ (red line).

CBM_{CipA} molecules being as dimers at a concentration of 1.7 mg/mL. As the K_D for the binding of CBM_{CipA} is 0.6 μ M [13–15], which is around two orders of magnitude lower we conclude that its dimerization has a negligible effect on cellulose binding (dimer fraction at such concentration is about 1% or lower). However, in other uses of the CBM_{CipA} such as using them for molecular adhesives, much higher concentrations were used, reaching over 100 mg/mL [11,17]. At these concentrations the majority (over 98%) of protein molecules will be in the dimer form. The dimerization could therefore affect the proteins solution behavior, such as coacervation [18]. Modelling indicated a possible mechanism of interaction between CBM_{CipA} molecules, and the modelled structures showed calculated properties that can be used for estimation of hydrodynamic parameters of oligomers. In general AUC is suitable for studying weak and ultraweak interactions, and several examples are found. Approaches include both sedimentation velocity and sedimentation equilibrium setups and use of different software-packages for analysis [45–48]. Despite different approaches implemented in the different software-packages, each software has its own benefits. The approaches did not contradict other, and they and all supported the same overall conclusion.

CRedit authorship contribution statement

Dmitrii Fedorov did the AUC measurements, modelling, and data analysis, and wrote the manuscript, Piotr Batys supervised and assisted in the modelling and wrote the manuscript, David B Hayes critically evaluated data, supervised and assisted in AUC data analysis, commented and corrected the manuscript, Maria Sammalkorpi supervised modelling and wrote the manuscript, Markus B Linder supervised the work, and wrote the manuscript.

Acknowledgements

Professor Borries Demeler is thanked for help with setting up UltraScan. We thank the Aalto School of Chemical Engineering graduate school for funding (DF). This work was also supported by the Academy of Finland through its Centres of Excellence Programme (2014–2019, HYBER) and under Projects No. 307474 and No. and 309324 (MS), and National Science Centre Grant No. 2018/31/D/ST5/01866 (PB). We are grateful for the support by the FinnCERES Materials Bioeconomy Ecosystem, and use of the Bioeconomy Infrastructure at Aalto. Computational resources by CSC IT Centre for Science, Finland, and RAMI – RawMatTERS Finland Infrastructure are also gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijbiomac.2020.09.054>.

References

- [1] O. Shoseyov, Z. Shani, I. Levy, Carbohydrate binding modules: biochemical properties and novel applications, *Microbiol. Mol. Biol. Rev.* 70 (2006) 283–295, <https://doi.org/10.1128/MMBR.00028-05>.
- [2] The CAZypedia Consortium, Ten years of CAZypedia: a living encyclopedia of carbohydrate-active enzymes, *Glycobiology*. 28 (2017) 3–8.
- [3] L.I. Crouch, A. Labourel, P.H. Walton, G.J. Davies, H.J. Gilbert, The contribution of non-catalytic carbohydrate binding modules to the activity of lytic polysaccharide monoxygenases, *J. Biol. Chem.* 291 (2016) 7439–7449, <https://doi.org/10.1074/jbc.M115.702365>.
- [4] H. Kim, M. Goto, H. Jeong, K.H.W.A. Jung, I.I. Kwon, K. Furukawa, Functional analysis of a hybrid endoglucanase of bacterial origin having a cellulose binding domain from a fungal exoglucanase, *Appl. Biochem. Biotechnol.* (1998) <https://doi.org/10.1007/BF02787774>.
- [5] C.A. Boraston, A. B. B.W. McLean, J.M. Kormos, M. Alam, N.R. Gilkes, R.A.J.W. Haynes, P. Tomme, D.G. Kilburn, Carbohydrate-binding modules: diversity of structure and function, *Recent Adv. Carbohydr. Bioeng.* 27 (1999) 202–211, [https://doi.org/10.1016/S0168-1656\(97\)00088-6](https://doi.org/10.1016/S0168-1656(97)00088-6).
- [6] M. Srisodsuk, J. Lehtiö, M. Linder, E. Margolles-Clark, T. Reinikainen, T.T. Teeri, Trichoderma reesei cellobiohydrolase I with an endoglucanase cellulose-binding domain: action on bacterial microcrystalline cellulose, *J. Biotechnol.* 57 (1997) [https://doi.org/10.1016/S0168-1656\(97\)00088-6](https://doi.org/10.1016/S0168-1656(97)00088-6).
- [7] M.C. Lim, M. Penttilä, Addition of substrate-binding domains increases substrate-binding capacity and specific activity of a chitinase from Trichoderma harzianum, *FEMS Microbiol. Lett.* 198 (2001) 57–63.
- [8] A.A. Wang, A. Mulchandani, W. Chen, Whole-cell immobilization using cell surface-exposed cellulose-binding domain, *Biotechnol. Prog.* (2001) 407–411.
- [9] A.A. Wang, A. Mulchandani, W. Chen, Specific adhesion to cellulose and hydrolysis of organophosphate nerve agents by a genetically engineered Escherichia coli strain with a surface-expressed cellulose-binding domain and organophosphorus hydrolase, *Appl. Environ. Microbiol.* 68 (2002) 1684–1689, <https://doi.org/10.1128/AEM.68.4.1684>.
- [10] X. Shi, F. Zheng, R. Pan, J. Wang, S. Ding, Engineering and comparative characteristics of double carbohydrate binding modules as a strength additive for papermaking applications, *BioResources* 9 (2014) 3117–3131, <https://doi.org/10.15376/biores.9.2.3117-3131>.
- [11] P. Mohammadi, G. Beaune, B.T. Stokke, J.V.I. Timonen, M.B. Linder, Self-coacervation of a silk-like protein and its use as an adhesive for cellulosic materials, *ACS Macro Lett.* 7 (2018) 1120–1125, <https://doi.org/10.1021/acsmacrolett.8b00527>.
- [12] J. Tormo, R. Lamed, A.J. Chirino, E. Morag, E.A. Bayer, Y. Shoham, T.A. Steitz, Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose, *EMBO J.* 15 (1996) 5739–5751, <https://doi.org/10.1002/j.1460-2075.1996.tb00960.x>.
- [13] N. Georgelis, N.H. Yennawar, D.J. Cosgrove, Structural basis for entropy-driven cellulose binding by a type-A cellulose-binding module (CBM) and bacterial expansin, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 14830–14835, <https://doi.org/10.1073/pnas.1213200109>.
- [14] E. Morag, A. Lapidot, D. Govorko, R. Lamed, M. Wilchek, E.A. Bayer, Y. Shoham, Expression, purification, and characterization of the cellulose-binding domain of the scaffoldin subunit from the cellulosome of Clostridium thermocellum, *Appl. Environ. Microbiol.* 61 (1995) 1980–1986.
- [15] S. Voutilainen, A. Paananen, M. Lille, M. Linder, Modular protein architectures for pH-dependent interactions and switchable assembly of nanocellulose, *Int. J. Biol. Macromol.* 137 (2019) 270–276, <https://doi.org/10.26434/chemrxiv.8010872.v1>.
- [16] S. Arola, M.B. Linder, Binding properties of single and double cellulose binding modules reveal differences between cellulose substrate, *Sci. Rep.* 6 (2016) 35358, <https://doi.org/10.1038/srep35358>.
- [17] P. Mohammadi, A.S. Aranko, L. Lemetti, Z. Cenev, Q. Zhou, S. Virtanen, C.P. Landowski, M. Penttilä, W.J. Fischer, W. Wagermaier, M.B. Linder, Phase transitions as intermediate steps in the formation of molecularly engineered protein fibers, *Commun. Biol.* 1 (2018) 86, <https://doi.org/10.1038/s42003-018-0090-y>.
- [18] L. Lemetti, S. Hirvonen, D. Fedorov, P. Batys, M. Sammalkorpi, H. Tenhu, M.B. Linder, A.S. Aranko, Molecular crowding facilitates assembly of spider-like proteins through phase separation, *Eur. Polym. J.* 112 (2019) 539–546, <https://doi.org/10.1016/j.eurpolymj.2018.10.010>.
- [19] G.Y. Xu, E. Ong, N.R. Gilkes, D.G. Kilburn, D.R. Muhandiram, M. Harris-Brandts, J.P. Carver, L.E. Kay, T.S. Harvey, Solution structure of a cellulose-binding domain from Cellulomonas fimi by nuclear magnetic resonance spectroscopy, *Biochemistry* 34 (1995) 6993–7009, <https://doi.org/10.1021/bi00021a011>.
- [20] J. Flint, D. Nurizzo, S.E. Harding, E. Longman, G.J. Davies, H.J. Gilbert, D.N. Bolam, Ligand-mediated dimerization of a carbohydrate-binding module reveals a novel mechanism for protein-carbohydrate recognition, *J. Mol. Biol.* 337 (2004) 417–426, <https://doi.org/10.1016/j.jmb.2003.12.081>.
- [21] A.J. Rowe, Ultra-weak reversible protein – protein interactions, *Methods* 54 (2011) 157–166, <https://doi.org/10.1016/j.jymeth.2011.02.006>.
- [22] B. Demeler, E. Brookes, R. Wang, V. Schirf, C.A. Kim, Characterization of reversible associations by sedimentation velocity with ultrascan, *Macromol. Biosci.* 10 (2010) 775–782, <https://doi.org/10.1002/mabi.200900481>.
- [23] J.W. Williams, R.L. Baldwin, K.E. Van Holde, H. Fujita, The theory of sedimentation analysis, *Chem. Rev.* 58 (1958) 715–744, <https://doi.org/10.1021/cr50022a005>.
- [24] B. Demeler, A comprehensive data analysis software package for analytical ultracentrifugation experiments, *Mod. Anal. Ultracentrifugation Tech. R. Soc. Chem* 2005, pp. 210–229.
- [25] E. Brookes, B. Demeler, Genetic algorithm optimization for obtaining accurate molecular weight distributions from sedimentation velocity experiments, *Anal. Ultracentrifugation VIII*. 131 (2006) 33–40.
- [26] E.H. Brookes, B. Demeler, Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments, *Proc. 9th Annu. Conf. Genet. Evol. Comput* 2007, pp. 361–368, <https://doi.org/10.1145/1276958.1277035>.
- [27] B. Demeler, K.E. Van Holde, Sedimentation velocity analysis of highly heterogeneous systems, *Anal. Biochem.* 335 (2004) 279–288, <https://doi.org/10.1016/j.ab.2004.08.039>.
- [28] E. Brookes, B. Demeler, C. Rosano, M. Rocco, The implementation of SOMO (SOLUTION MOdeller) in the UltraScan analytical ultracentrifugation data analysis suite: enhanced capabilities allow the reliable hydrodynamic modeling of virtually any kind of biomacromolecule, *Eur. Biophys. J.* (2010) 423–435, <https://doi.org/10.1007/s00249-009-0418-0>.
- [29] E. Kang, M.L. Mansfield, J.F. Douglas, Numerical path integration technique for the calculation of transport properties of proteins, *Phys. Rev. E* (2004) 1–11, <https://doi.org/10.1103/PhysRevE.69.031918>.
- [30] M.L. Mansfield, J.F. Douglas, Improved path integration method for estimating the intrinsic viscosity of arbitrarily shaped particles, *Phys. Rev. E* (2008) 1–11, <https://doi.org/10.1103/PhysRevE.78.046712>.

- [31] E. Brookes, M. Rocco, Recent advances in the UltraScan SOLUTION MOdeller (US - SOMO) hydrodynamic and small - angle scattering data analysis and simulation suite, *Eur. Biophys. J.* 47 (2018) 855–864, <https://doi.org/10.1007/s00249-018-1296-0>.
- [32] P. Schuck, Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling, *Biophys. J.* 78 (2000) 1606–1619, [https://doi.org/10.1016/S0006-3495\(00\)76713-0](https://doi.org/10.1016/S0006-3495(00)76713-0).
- [33] P. Schuck, P. Rossmanith, Determination of the sedimentation coefficient distribution by least-squares boundary modeling, *Biopolymers*. 54 (2000) 328–341, [https://doi.org/10.1002/1097-0282\(20001015\)54:5<328::AID-BIP40>3.0.CO;2-P](https://doi.org/10.1002/1097-0282(20001015)54:5<328::AID-BIP40>3.0.CO;2-P).
- [34] P. Schuck, On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation, *Anal. Biochem.* 320 (2003) 104–124, [https://doi.org/10.1016/S0003-2697\(03\)00289-6](https://doi.org/10.1016/S0003-2697(03)00289-6).
- [35] H.P. van Holde KE, W.C. Johnson, Chemical equilibria involving macromolecules, *Princ. Phys. Biochem.* (2006) 605–658.
- [36] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, *Comput. Phys. Commun.* 91 (1995) 43–56, [https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/10.1016/0010-4655(95)00042-E).
- [37] A. Juréus, Ü. Langel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, *Acta Chim. Slov.* 43 (1996) 51–60, <https://doi.org/10.1007/s008940100045>.
- [38] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* 103 (1995) 8577–8593, <https://doi.org/10.1063/1.470117>.
- [39] R.B. Best, W. Zheng, J. Mittal, Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association, *J. Chem. Theory Comput.* 10 (2014) 5113–5124, <https://doi.org/10.1021/ct500569b>.
- [40] J.L.F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, *J. Chem. Phys.* 234505 (2005) <https://doi.org/10.1063/1.2121687>.
- [41] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.* 126 (2007) <https://doi.org/10.1063/1.2408420>.
- [42] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, *J. Appl. Phys.* 52 (1981) 7182–7190, <https://doi.org/10.1063/1.328693>.
- [43] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* 7855 (1996) 33–38.
- [44] Y. Li, W.F. Stafford, M. Hesselberg, D. Hayes, Z. Wu, M. Byrne, Characterization of the self-association of human interferon- α 2b, albinterferon- α 2b, and pegasys, *J. Pharm. Sci.* 101 (2012) 68–80, <https://doi.org/10.1002/jps.22751>.
- [45] X. Mao, Z. Ren, G.N. Parker, H. Sondermann, M.A. Pastorello, W. Wang, J.S. McMurray, B. Demeler, J.E. Darnell, X. Chen, Structural bases of unphosphorylated STAT1 association and receptor binding, *Mol. Cell* 17 (2005) 761–771, <https://doi.org/10.1016/j.molcel.2005.02.021>.
- [46] J. Zheng, C.D. Fage, B. Demeler, D.W. Ho, A.T. Keatinge-clay, The missing linker: a dimerization motif located within polyketide synthase modules, *ACS Chem. Biol.* 8 (2013) 1263–1270, <https://doi.org/10.1021/cb400047s>.
- [47] H. Silkowski, S.J. Davis, A.N. Barclay, A.J. Rowe, S.E. Harding, O. Byron, Characterisation of the low affinity interaction between rat cell adhesion molecules CD2 and CD48 by analytical ultracentrifugation, *Eur. Biophys. J.* 58 (1997) 455–462.
- [48] T.R. Patel, S.E. Harding, A. Ebringerova, M. Deszczynski, Z. Hromadkova, A. Togola, S. Paulsen, G.A. Morris, A.J. Rowe, Weak self-association in a carbohydrate system, *Biophys. J.* 93 (2007) 741–749, <https://doi.org/10.1529/biophysj.106.100891>.