Computational and Structural Biotechnology Journal 21 (2023) 2188–2196

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

Computational modeling of the molecular basis for the calcium-dependence of the mannuronan C-5 epimerase AvAlgE6 from *Azotobacter vinelandii*

Margrethe Gaardløs [a,*], Anders Lervik [b], Sergey A. Samsonov [a,**]

[a] Faculty of Chemistry, University of Gdańsk, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland
[b] Department of Chemistry, NTNU Norwegian University of Science and Technology, N-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 15 December 2022
Received in revised form 13 March 2023
Accepted 13 March 2023
Available online 15 March 2023

Keywords:
Alginate
Mannuronan C-5 epimerases
Calcium dependence
Reaction mechanism
Molecular dynamics
Umbrella sampling

ABSTRACT

The mannuronan C-5 epimerases catalyze epimerization of β-D-mannuronic acid to α-L-guluronic acid in alginate polymers. The seven extracellular *Azotobacter vinelandii* epimerases (AvAlgE1–7) are calcium-dependent, and calcium is essential for the structural integrity of their carbohydrate binding R-modules. Ca^{2+} is also found in the crystal structures of the A-modules, where it is suggested to play a structural role. In this study, the structure of the catalytic A-module of the *A. vinelandii* mannuronan C-5 epimerase AvAlgE6 is used to investigate the role of this Ca^{2+}. Molecular dynamics (MD) simulations with and without calcium reveal the possible importance of the bound Ca^{2+} in the hydrophobic packing of β-sheets. In addition, a putative calcium binding site is found in the active site, indicating a potential direct role of this calcium in the catalysis. According to the literature, two of the residues coordinating calcium in this site are essential for the activity. MD simulations of the interaction with bound substrate indicate that the presence of a calcium ion in this binding site increases the binding strength. Further, explicit calculations of the substrate dissociation pathways with umbrella sampling simulations show and energetically higher dissociation barrier when calcium is present. The present study eludes to a putative catalytic role of calcium in the charge neutralizing first step of the enzymatic reaction. In addition to the importance for understanding these enzymes' molecular mechanisms, this could have implications for engineering strategies of the epimerases in industrial alginate processing.

1. Introduction

Mannuronan C-5 epimerases (alginate epimerases) catalyze the C-5 epimerization of β-D-mannuronic acid (M) to α-L-guluronic acid (G) residues in alginate polymers, creating block patterns of un-epimerised residues (M-blocks), alternating epimerizations (MG-blocks) and subsequent epimerizations (G-blocks) [1–3]. Alginate is a linear copolymer of (1→4) linked, negatively charged M- and G-residues, synthesised by brown algae and soil bacteria in the genera *Azotobacter* and *Pseudomonas* [4–6]. The proposed function of alginate in the *Azotobacter* genera is related to survival during cyst formation. In opportunistic pathogenic species of *Pseudomonas*, alginate is used in biofilm formation and protection from host reactions [7–10]. The organisms use alginate epimerases to fine-tune the block patterns of MG- and G-blocks in the polymer, resulting in unique hydrogels with specific properties depending on the block composition [11]. Alginate is a valuable biopolymer in industrial and medical applications, and in addition to its material properties, it is non-toxic, biocompatible, biodegradable, and immunogenic. In order to create the industrially high-valued G-blocks, the epimerases can be used in vitro for alginate modification [12]. Decades of research have elucidated the mode of action of ten different extracellular alginate epimerases called AlgE, produced by the soil bacteria *A. vinelandii* (AvAlgE1–7) and *A. chroococcum* (AcAlgE1–3) [13–15]. These block-forming enzymes require calcium for their activities [2,16], but calcium is also an inhibitor of G-block formation [17]. If the engineered AlgE variants could be made calcium-independent, it could aid in vitro modifications of alginate to produce a high G-

* Corresponding author at: Faculty of Chemistry, University of Gdańsk, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland.
** Corresponding author.
E-mail addresses: margrethe.gaardlos@gmail.com (M. Gaardløs), sergey.samsonov@ug.edu.pl (S.A. Samsonov).

https://doi.org/10.1016/j.csbj.2023.03.021

content. Such modifications require a thorough understanding of the molecular mechanism behind the calcium dependency in the AlgEs.

The AlgEs consist of two modules, the active A-module and the carbohydrate binding module (CBM)-like R-modules. Calcium ions are bound in the β-roll folds of the R-modules, a known requirement for the structural stability of this type of fold [18,19]. The A-modules of AvAlgE4 and AvAlgE6 (PDB IDs 2PYH and 5LW3, respectively) have a parallel β-helix fold with an N-terminal α-helix cap, known as carbohydrate-binding proteins and sugar hydrolases domain (CASH) [20,21]. In both crystal structures, a calcium ion is coordinated in a buried site between two loops close to the active site. A similar metal binding site is found in the dermatan sulfate epimerases, a group of polysaccharide epimerases that epimerise β-D-glucuronic acid to α-L-iduronic acid in glycosaminoglycan chains. They coordinate a Mn^{2+}-ion in a loop close to the active site, equivalent to Ca^{2+} in the AlgE [22]. Mutations in the residues coordinating the Mn^{2+}-ion are deleterious to enzymatic activity, and the manganese binding site was assumed to interact with the negatively charged substrate and affect the catalytic reaction [22].

It is not known if calcium has a role in the enzymatic reaction performed by the A-modules, or if its role is purely structural. In addition to possible electrostatic interactions with the C-6 carboxylates of the bound substrate, a catalytic role could constitute a neutralization of the carboxylate of the sugar ring in the active site during catalysis. Four residues, all located in the enzyme subsite +1, have been found to be essential for activity: Y149, D152, H154, and D178 [21]. As there are acidic residues in and around the active site, a putative calcium coordination has been proposed to be established by the catalytic residues E155 and D178 [23]. Neutralization of the substrate charge is the first step in the suggested epimerase reaction, and it activates C-5 of the mannuronate residue for subsequent proton abstraction of H-5 [24]. Donation of a proton to the opposite face of the sugar ring leads to formation of the C-5 epimer guluronate. It is suggested that the catalytic H154 is the proton abstractor, and the catalytic Y149 is the proton donor [25].

There are mechanistic similarities between the reactions of uronic acid lyases and epimerases. Three of the AlgE epimerases exhibit a dual lyase and epimerase activity, presumably in the same active site [14,24,26]. Some alginate lyases use Ca^{2+} in a metal-assisted β-elimination where the ion neutralises the substrate charge, whereas other alginate lyases use protein residues for the neutralization. Usually, polysaccharide lyases with the parallel β-helix fold use Ca^{2+} for substrate charge neutralization [27]. Calcium is, however, not a universal requirement for alginate epimerases with this fold, because the periplasmic alginate epimerase AlgG is calcium independent [28,29]. This epimerase is found in alginate producing species of *Azotobacter* and *Pseudomonas*, including *A. vinelandii*. It has a larger α-helical region at the N-terminus than the AlgEs, and an arginine positioned in the active site instead of D178 found in the AlgEs. This could contribute to charge neutralization and putatively compensate for the absence of a calcium ion in the AlgE active site. Although they likely follow equivalent catalytic routes, AlgE and AlgG yield different alginate polymers, as AlgE epimerises residues block-wise and AlgG creates single G-residues in the chain [29,30].

In this study, we investigate calcium's role in the AlgEs through a computational analysis of the deposited structure of AvAlgE6′s A-module (AvAlgE6A, PDB ID 5LW3). Molecular dynamics (MD) simulations indicate that the structurally bound calcium ion has a role in maintaining the structural integrity of the enzyme. We identify a possible calcium coordination in the active site, and perform MD simulations of enzyme-substrate complexes with and without calcium in the active site. In addition to conventional MD simulations, we perform umbrella sampling MD where the substrate is pulled away from a stable binding pose with and without calcium, and compare the resulting potential of mean force pathways. Calcium can indeed be coordinated stably in the active site, and we obtain indications that the presence of calcium is favourable for the interaction energies of the enzyme-substrate complex. In addition, we observe stable binding poses where calcium is coordinated by the substrate carboxylate, and the proton at C-5 points towards the putative proton abstractor H154. This supports the hypothesis that the ion has a role in the first charge neutralizing step in the reaction mechanism. Our results inidicate that calcium might be essential in the catalytic mechanism. In our discussion, we comment on the possible implications of this for efforts to engineer calcium-independent enzymes, as calcium also inhibits the production of valuable G-rich alginate.

2. Methods

2.1. Electrostatic potential calculations of AvAlgE6A and AlgG

Electrostatic surface potential calculations are relevant in systems where proteins interact with negatively charged carbohydrates [31]. The A-module of AvAlgE6 (PDB ID 5LW3) and AlgG from *P. aeruginosa* (PDB ID 4NK6) were subjected to Poisson-Boltzmann surface area (PBSA) electrostatic potential calculations with the PBSA-program from AmberTools in AMBER16 [32] with a grid spacing of 1 Å. The resulting protein surfaces were visualised in VMD [33].

2.2. Molecular dynamics simulations

Three different molecular systems analysed by MD simulations were set up: i) wild type AvAlgE6A with and without the calcium ion from the crystal structure; ii) AvAlgE6A wild type with calcium ion initially placed near the active site; and iii) AvAlgE6A wild type with mannuronate substrates, with and without calcium in the active site.

i) The AvAlgE6A crystal structure solved at 1.2 Å resolution (PDB ID 5LW3) with and without coordinated calcium was studied with MD simulations using the ff14SBonlysc force field in AMBER 20 [34,35]. In order to determine the effect of salt concentration, we also investigated the system with and without NaCl, resulting in a total of 4 different simulation setups which were all repeated three times. In all setups, Na^+ counterions were added to create an electroneutral system. A concentration corresponding to 100 mM NaCl was obtained in the other setup by adding 35 Na^+ and 15 Cl^- ions according to the calculation of Machado and Pantano [36]. Histidine residues were protonated on the ε-nitrogens (HIE residue library), except for the active site histidine H154 which was protonated on the δ-nitrogen (HID residue library) based on the assumption that it is the catalytic base and its relative orientation in the active site [25]. The structures were first solvated in TIP3P water with minimal distances of 8 Å to the periodic boundary. A two-step energy minimization was performed, first with restraints on the solute during 0.5×10^3 steepest descent cycles and 10^3 conjugate gradient cycles, and second with no restraints during 3×10^3 steepest descent cycles and 3×10^3 conjugate gradient cycles. The systems were heated to 300 K during 10 ps using the Langevin thermostat. Finally, they were equilibrated for 100 ps at 300 K and 10^5 Pa in the isobaric isothermic ensemble (NPT) with the Langevin thermostat and the Berendsen barostat, before a final production run of 200 ns under the same conditions. Harmonic force restraints of 100 kcal/mol/Å^2 were put on the protein and calcium ion during the first minimization step and heating. All covalent bonds containing hydrogens were under the SHAKE algorithm, and Particle Mesh Ewald method treated electrostatics. Trajectories were visualised with VMD [33], and the RMSD relative to the energy minimised wild type structure was obtained using the AMBER module CPPTRAJ [37].

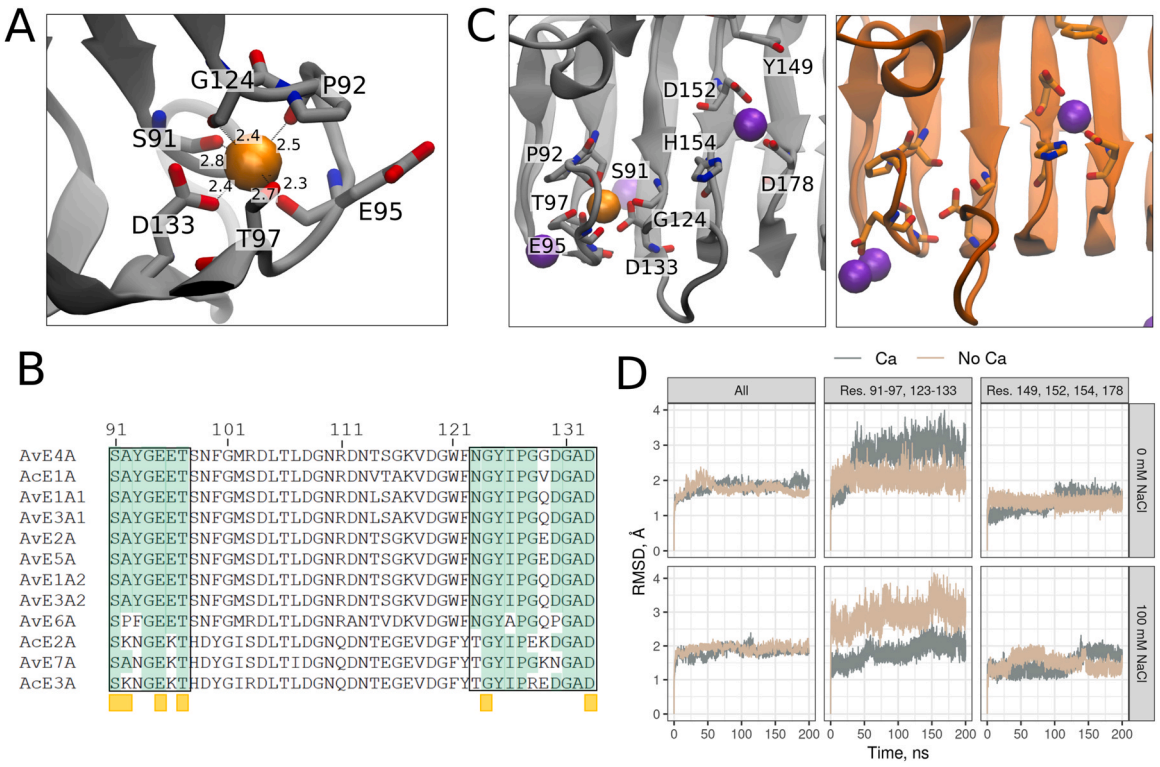

Fig. 1. Structure analysis and MD simulation results of AvAlgE6A with and without calcium in the calcium binding site. A. The calcium coordination site in the AvAlgE6A crystal structure (PDB ID 5LW3). The calcium ion is shown as an orange sphere, and the six coordination partners are shown as sticks. B. Multiple sequence alignment produced using an exhaustive pairwise multi-way alignment and the BLOSUM 62 scoring matrix with Clone Manager Professional Suite software 9.2 (Sci-Ed). Black boxes show the alignment of the two loops coordinating calcium. Orange squares are placed under the calcium coordinating residues. C. Cartoon representations of the last frames from 200 ns MD simulations of AvAlgE6A wild type with (left) and without (right) calcium present, in the presence of 100 mM NaCl. The calcium ion is shown as an orange sphere, and the sodium ions associating with the proteins as purple spheres. The residues coordinating calcium in the two loops (S91, P92, E95, T97, G124 and D133) and the active site residues (Y149, D152, H154 and D178) are shown in sticks. Figures were made using VMD [33]. D. RMSD values with reference to the starting structure averaged over three simulation repeats of 200 ns each, with calcium (grey), without calcium (brown), and with 0 (top) or 100 (bottom) mM NaCl. Every 10th frame is plotted. RMSD is compared for all residues, for the calcium binding site residues (residues 91–97, 123–133) and the catalytic residues (residues 149, 152, 154, and 178).

ii) A calcium ion was placed in three different locations (called startposes 1–3, where a pose denotes a position) close to the active site, and simulations were performed in the same manner as described above for AMBER 20 with the ff14SBonlysc force field, but without the addition of 100 mM NaCl. Startposes 1 and 2 were simulated 10 times in total, 3 repetitions of 100 ns and 7 repetitions of 50 ns. Startpose 3 was simulated once for 100 ns. In total, 1.4 μs of MD simulations were performed in this step. Contacts between protein atoms (excluding hydrogens) and calcium were defined as residues within 3 Å of the ion, and measured throughout the trajectories with the *nativecontacts* command in CPPTRAJ.

iii) Two alginate substrates (mannuronate tetramer and octamer, M4 and M8) were modeled and docked to the A-module of AvAlgE4 in our previous work [38]. The resulting poses were used as starting structures in this study by aligning AvAlgE4A with AvAlgE6A and copying the binding poses from stable AvAlgE4A MD-trajectories to AvAlgE6A. Then, the complexes with or without calcium in the active site were subjected to MD simulations as described above for AMBER 20 with the ff14SBonlysc and GLYCAM06 [39] force field parameters used for protein and alginate, respectively. This was done without the addition of 100 mM NaCl. For each complex, 3 repetitions were performed for 100 ns, and 7 repetitions were performed for 50 ns, resulting in a total of 2.6 μs of MD simulations in this step. The same starting structure of the protein complexed with the substrate was used with and without calcium. In the simulations with calcium, the ion coordinates were taken from the last frame of one of the trajectories with startpose 1 from the previous experiments (ii). The resulting trajectories of the complexes were analysed using CPPTRAJ.

2.3. Umbrella sampling

The strength of binding between AvAlgE6A and the two alginate substrates in the presence and absence of a calcium ion in the active site was studied using the umbrella sampling (US) approach [40]. For each substrate, two of the US experiments had starting poses from the last frames from stable MD simulations with calcium described in the previous section: one where the calcium was kept and one where it was removed (denoted MD1). In addition, an experiment was set up without calcium where the starting pose corresponded to the frame of a stable MD simulation without calcium (denoted MD2). The dissociation processes were modeled by subjecting the distance of a backbone nitrogen in the protein and a glycosidic linkage oxygen in the sugar to a harmonic restraint with a force constant of 8 (M4) or 4 (M8) kcal/mol/Å^2. The restraint distance was increased by 1 Å in each window (simulation step) through 40 (M4) or 30 (M8) subsequent windows. Each window was simulated for 10 ns, following equilibration steps of 0.1 ns. The overlap of probability distributions of the distances between windows was analysed visually. Potential of mean force (PMF) plots were calculated from the probability distributions using the weighted histogram analysis method (WHAM) with the Grossfield WHAM software [41,42]. Force convergence tolerance was set to 0.001 kcal/mol/Å, and 1000 fake data sets were created in the Monte Carlo bootstrap error analysis.

In total, 6.4 μs of MD simulations and 2.1 μs of umbrella sampling MD simulations were performed.

3. Results and discussion

3.1. Investigating the role of the bound calcium

This study started with a visual analysis of the crystal structure of AvAlgE6A solved at 1.2 Å resolution and deposited in the PDB (PDB ID 5LW3, [23]). A calcium ion with an octahedral coordination is found in a site spanning two hairpin loops at the N-terminal end of the alginate binding groove (residues 91–97 and 123–133, Fig. 1 A). Bound calcium ions in proteins commonly have six to eight coordination partners [43], and in the crystal structures of AvAlgE6A and AvAlgE4A there are six: the side chains of S91, T97, and D133, and the main chains of E95, G124, and the non-conserved P92 in AvAlgE6A and A92 in AvAlgE4A [21]. Their coordinating atoms are placed from 2.3 to 2.8 Å from the calcium ion in the crystal structure. The site is partially conserved in the AlgEs, and of the six coordination partners, only residue 92 is not conserved (Fig. 1B). The calcium binding site could be important for the structural integrity of the protein, similar to calcium's role in the R-modules where it is vital for the structural stability [18].

In order to investigate this metal site, we performed MD simulations of AvAlgE6A with and without the bound calcium ion (see Methods for details). Interestingly, the side chain of E95 rotates and participates in calcium coordination, establishing an additional seventh coordinating group. This rotation is not observed in the simulations where calcium is removed. Moreover, we found an effect of calcium on the structural integrity of the protein, shown in Fig. 1C-D. The tight packing of the parallel β-sheets around the metal site is affected upon removal of the calcium. In the calcium site, the electronegative hydroxyl groups of S91 and T97 and carboxylate groups of E95 and D133 start to repel each other, and this causes a loosening of the packing around this site that extends to the active site. In both systems, sodium ions associate with negatively charged and polar residues in the two sites. This indicates a putative location for a binding site for positively charged ions in the active site, which is investigated in subsequent sections. The RMSDs for all residues, the calcium binding site residues, and the catalytic residues are relatively similar regardless of the presence of calcium (Fig. 1D). However, in the presence of 100 mM NaCl, the simulation with calcium has a decrease in the RMSD of the calcium binding loop atoms, and the simulation without calcium has an increase in RMSD of these atoms compared to the systems without NaCl. The enzyme has a binding groove that accommodates a negatively charged substrate, and by associating with these residues, the sodium ions appear to stabilise the structure overall. When calcium is removed, sodium ions are seen to associate with the calcium binding site, causing the increase of RMSD relative to the starting structure. Previous experimental results have shown a complex interplay between calcium and NaCl concentration for the activity of the bifunctional lyase and epimerase AlgE7 [25]. An increase in NaCl concentration decreases the lyase activity and increases epimerase activity and G-block formation, whereas an increase in calcium concentration increases both lyase and epimerase activity but decreases G-block formation. This complex relationship is probably related to the substrate interaction, but here we also see that the ions have effects on the protein structure flexibility.

If the bound calcium is in contact with substrate residues, it might be indirectly necessary for catalysis even without a direct participation in the catalytic mechanism. Keeping this in mind, the postulated role of calcium as the initial charge neutraliser before proton abstraction in the reaction mechanism and the observed coordination of sodium ions in the active site motivated us to proceed with computational studies of calcium binding in the active site.

In the following simulations we did not add 100 mM NaCl, in order to restrict our analyses to only concern the effects of an active site calcium ion. However, we know from previous experimental results that in this highly electronegative enzyme-substrate system, the NaCl concentration is important, and in the future, it would be of great interest to extend our computational investigations of the effect of NaCl on substrate interactions.

3.2. Structural analysis of AvAlgE6A

Polysaccharide epimerases constitute a relatively small group in the carbohydrate-active enzymes (CAZymes) framework: in addition to the alginate epimerases, only two other enzyme families are known [44]. However, they share structural and functional similarities with the polysaccharide lyases, and the two enzyme activities have similar reaction mechanisms, as discussed in the introduction. In the polysaccharide lyase families PL1, PL3, PL6, and PL9, the carboxylate group of the uronic acid is neutralised by a positive charge from either calcium or arginine [27]. To propose a similar role of calcium in the active site of the AlgEs, a possible first step is to compare them to the calcium independent AlgG. AlgG has a similar active site, but the AlgEs have no equivalent residue to R345, which is probably neutralising the carboxylate group in AlgG (*P. aeruginosa* AlgG numbering from PDB ID 4NK6, Fig. 2) [23]. Instead, E155 and D178 occupy the equivalent spacial position in the AlgEs, where Wolfram et al. proposed they coordinate a calcium ion. Previous mutation studies point to these residues being essential for activity [21]. The surface potentials obtained from PBSA analysis of AvAlgE6A and AlgG shown in Fig. 2 reveal that these residues contribute to the establishment of the negative electrostatic potential area in the active site in AvAlgE6a, whereas AlgG has more positive electrostatic potential in the active site. How this negative potential could affect the positioning of a sugar ring in the active site is described further below.

3.3. Is there a calcium binding site in the active site?

We proceeded to investigate if calcium could be responsible for the charge neutralizing first step in the proposed mechanism for the epimerases [24]. We started by placing a calcium ion close to the active site in AvAlgE6A without any substrate, to determine if it could be coordinated stably for neutralization of the substrate carboxyl upon substrate binding. Simulations were initiated from slightly different positions of the ion around the active site (Fig. 3 A). The ion in one of the positions (startpose 3) dissociated, and the other two moved into stable coordination configurations in the active site. The simulations were repeated 10 times each for startposes 1 and 2, and in these 20 simulations, a stable end pose with calcium in the active site was found in 15 of the simulations (Fig. S1).

Two different coordination sites were sampled in the active site, called endpose 1 and 2 (see Fig. 3B). Endpose 1 is coordinated by E155, D178, and E229, and represents 11 of the 15 stable poses. It is close to startpose 1, but is also sampled from startpose 2 (Fig. 3 C). Endpose 2 is coordinated by D152 and D178, and represents 4 of 15 stable coordinations. It is closer to startpose 2, and yet is only found in 3 of 7 stable poses from startpose 2. Based on this, we deem endpose 1 as the more likely position for the calcium ion. Visual inspection shows that a calcium ion in endpose 2 would be placed between the active tyrosine and histidine residues, which might cause a steric hindrance for a bound substrate. Previous mutational studies of AvAlgE4, which has 81% sequence identity with AvAlgE6A, show that D152 and D178 are essential for the activity, whereas E229 has been identified as a potential calcium binding site through crystal soaking with $CaCl_2$.

We find that calcium can sample two distinct coordination configurations in the active site, and it stays bound once they are found. The existence of these sites suggests the possibility that calcium participates in the reaction mechanism, as there would

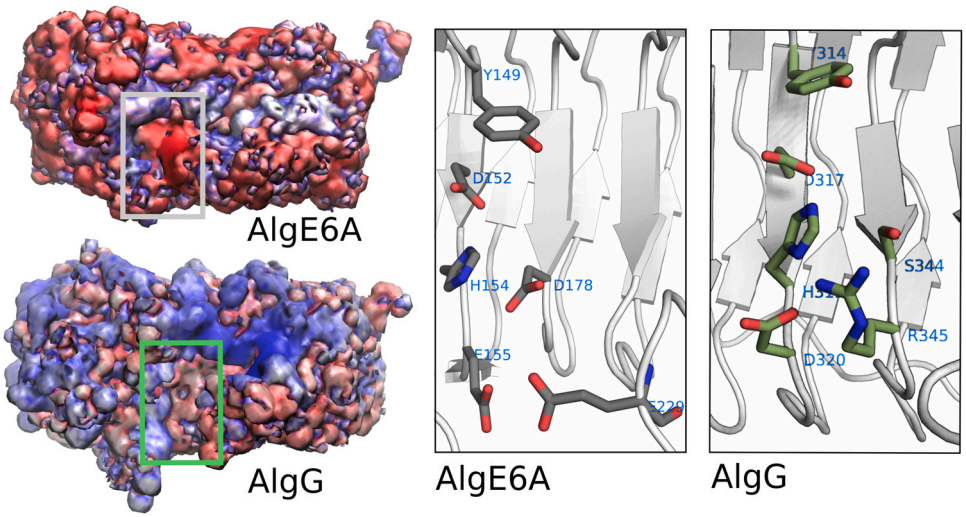

Fig. 2. Comparison of AvAlgE6A and AlgG electrostatic surface potentials calculated by AMBER PBSA (illustrated as isosurfaces coloured from −5 kcal/mol/e (red) to + 5 kcal/mol/e (blue) and visualised with VMD [33]). The negative patches present on the surface of AvAlgE6A might function as cation storage that aids in structural stability and donate ions to the substrate for gel formation after epimerization, driving the reaction by removing the product [18]. The insets show the two active sites, where Y149, D152, H154 and E155 in AvAlgE6A occupy equivalent positions to Y314, D317, H319, and D320 in AlgG, whereas D178 and E229 in AvAlgE6A occupy similar positions as S344 and R345 in AlgG. The active sites were visualised using PyMOL [45].

naturally be an abundance of calcium when the enzymes are active. As we find endpose 1 to be the most likely position, it is used as the starting position in the following simulations with alginate substrates.

3.4. How does the bound calcium affect substrate binding?

We followed by investigating how the coordinated calcium ion would affect a substrate placed in the binding groove. Two alginate

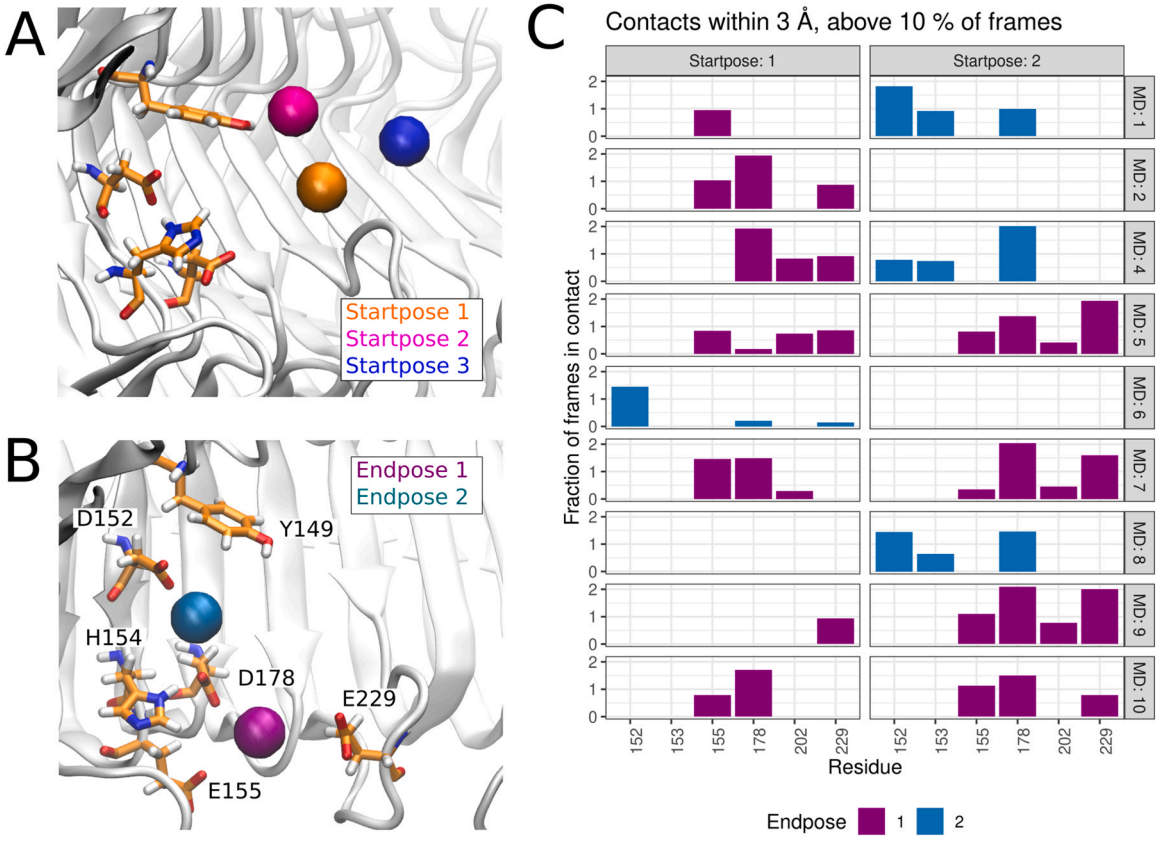

Fig. 3. MD-based calcium coordination analysis. A. Start positions of the calcium ion. B. The two stable end poses, after 50–100 ns of simulation (MD run 1–3: 100 ns, MD run 4–10: 50 ns). Figures created with VMD [33]. C. Contacts (defined by the cut-off distance of 3 Å) between the calcium ion and protein residues. In total, the complex established stable contacts for more than 10% of the frames in 15 simulation runs. Positions defined as endpose 1 from visual analysis correspond to purple bars, and poses defined as endpose 2 correspond to blue bars.

substrates, mannuronate tetramer and octamer (M4 and M8), were docked to AvAlgE4A in a previous work and the complexes were subjected to MD simulations [37]. In our current study, starting structures were selected from converged MD trajectories where the active site sugar is in the middle of the chain and close to the active site residues, and the alginate poses were placed in the AvAlgE6A corresponding binding site superimposed to the AvAlgE4A one. 10 simulation repeats starting from the same positions were then performed for the enzyme bound to M4 and M8 with and without calcium in endpose 1 (Fig. 3B) in the active site.

The resulting 40 MD trajectories are shown in Fig. S2. Of these trajectories, the substrate dissociated 3 times with calcium in the active site and 8 times without calcium in the active site. Calcium remained in endpose 1 and did not dissociate in any of the 20 runs, even when the substrate dissociated.

In order to quantitatively assess the differences in substrate binding with and without calcium, we measured the distance between the H-5 of the active site sugar (defined as the sugar residue closest to the active residues in each frame) and the active site atoms Y149-OH and H154-Nε. During catalysis, this distance has to be short, especially to the assumed proton abstractor H154. Only frames where the sugar was bound (the total RMSD of all the sugar residues relative to the starting position below 15 Å) were considered. From the resulting plots in Fig. 4 A, we observe that the median distance between the active site sugar H-5 and H154-Nε is lower for simulations with calcium (5.8 Å with calcium versus 8.4 Å without calcium for M4, and 5.2 Å with calcium versus 6.7 Å without calcium for M8). Y149 is closer to H-5 when calcium is absent, although the differences are smaller (with medians of 6.3 Å with calcium versus 5.6 Å without calcium for M4, and 5.7 Å with calcium versus 5.3 Å without calcium for M8). Visual inspection reveals that in poses without calcium, the substrate is repelled from the active site due to the proximity of negatively charged E152, E155 and D178. When calcium is present, the sugar is able to move in between the active site residues. In some of the trajectories, the carboxylate of the substrate in the active site takes part in calcium coordination, and the Nε of H154 is pointing towards H-5 (Fig. 4B). Such a positioning supports the hypotheses of calcium being the charge neutraliser in the catalysis, activating the H-5 for abstraction by H154.

To determine if the calcium ion affects the substrate binding strength, we modeled the dissociation process with umbrella sampling (US) in the presence and absence of the calcium ion. When calcium was present, the starting pose was the last frame of a stable MD simulation with calcium (MD7 for M4 and MD3 for M8, see Fig. S2). Without calcium, the dissociation process either started from the same initial structure as in the simulation with calcium, or it started from the last frame of a converged MD trajectory without calcium (MD10 for M4 and MD8 for M8). In the former pose, the dissociation with calcium is directly comparable to that without calcium as the only difference is the presence or absence of the ion. In the latter, the substrate in the active site is allowed to adapt to an energetically more favourable conformation corresponding to the scenario in the absence of calcium.

The restraint was in all runs applied for the distance between a backbone nitrogen in the stable α-helix cap of the protein and a glycosidic linkage oxygen in the sugar as shown in Fig. S3A. Studying the trajectories of the substrate dissociation (Fig. S3B), we observe that when calcium is present, the active site sugar is coordinated by the calcium ion and stays attached to the active site longer. When the distance increases, the sugar moves in the binding site so that the oligosaccharide unit previously located in subsite +2 (towards the N-terminal of AvAlgE6A) coordinates calcium instead, until at last, the substrate dissociates. Without calcium, the sugar dissociates more evenly from the binding groove, and there is no preference to stay attached in the active site. Density plots of the distances in each US window are shown in Fig. S3C. The PMF pathway along the reaction coordinate shows that the presence of calcium leads to a higher dissociation barrier for the ligand in all systems (Fig. 4 C). Larger differences are observed for M8 than for M4. For the dissociation processes of M8 without calcium, when the starting pose is obtained from a trajectory without calcium (MD2, blue curve in Fig. 4 C), the energy barrier is lower than when the starting pose is taken from a trajectory with calcium (MD1, orange curve in Fig. 4 C). This can signify that the binding is energetically more favourable when calcium is present.

To conclude, our results show that a bound calcium ion in the active site is beneficial for substrate binding. Stable poses are found in 18 of 20 MD runs with calcium present, whereas without calcium this number is 12 of 20. The average and median distance between H-5 and H154 are shorter when calcium is present. The stable pose with the active histidine pointing towards H-5 is not found without calcium, due to the repulsion of the carboxylate group by the acidic residues in the active site. This pose is not sampled in every simulation run with calcium, but when it is sampled, it is stable for the rest of the run. Assuming H154 is the catalytic base, it is positioned well for proton abstraction, and calcium acts as the charge neutralizing residue. The results presented here are in other words supporting the proposed reaction mechanism by Gacesa [24], with calcium as the charge neutraliser, H154 as the catalytic base, and Y149 as the catalytic acid. From this follows that the calcium dependence of the AlgEs could be catalytic, in addition to the structural role of calcium in the R-modules and the proposed structurally important calcium binding site in the A-modules.

The function of the calcium binding site found in the crystal structure is more elusive. The dermatan epimerases also have a similar metal binding site thought to affect substrate binding. In the dermatan epimerases, a bound substrate is positioned close to the metal site and might be affected by it, at least indirectly. In our simulations, we do not observe direct interaction between the substrate and the coordinate calcium, even for M8 substrates positioned above the loops, as the calcium is not exposed to the solvent. An indirect effect of this calcium on the substrate binding is still feasible, by allowing for proper movement of the two flexible loops coordinating it. We observed the largest effects of the removal of the calcium ion in these loops (Fig. 1D). The longest of the two loops in this site was previously suggested to be involved in the interaction with the substrate and processive action through clamping of the substrate [38,46]. Movement of the loop to perform such an action might be inhibited by changes in the calcium binding site, and this could affect substrate interaction, processive action, and proper substrate positioning in the active site for catalysis. In addition to calcium's putative involvement in the reaction mechanism, this would represent a different, indirect impact on catalysis.

Assuming that the buried calcium site is involved in processive action and substrate interactions through the flexible loops, there are biological implications when comparing various carbohydrate epimerases. Like the AlgEs, the dermatan sulfate epimerase is processive and requires divalent cations (Mn^{2+} in the case of the dermatan sulfate epimerases). AlgG lacks the metal dependence and is not processive, and perhaps there is a functional connection between the processivity of certain epimerases and their calcium binding sites. In its natural environment, AlgG has different requirements for its substrate interaction than the AlgEs, as it is placed in a periplasmic complex which the alginate chains are threaded through on their way out of the bacterial cells [47]. The AlgEs are, on the other hand, extracellular and modify alginate that is secreted from the bacteria through the processive creation of blocks of MG-residues and G-residues. This tailors the functional properties of the protective coats produced by the bacteria during cyst formation. The calcium dependence of the AlgEs might regulate the enzyme activity and prevent too high G-block concentrations for the organisms, as G-blocks bind calcium and make it unavailable for the enzymes [48].

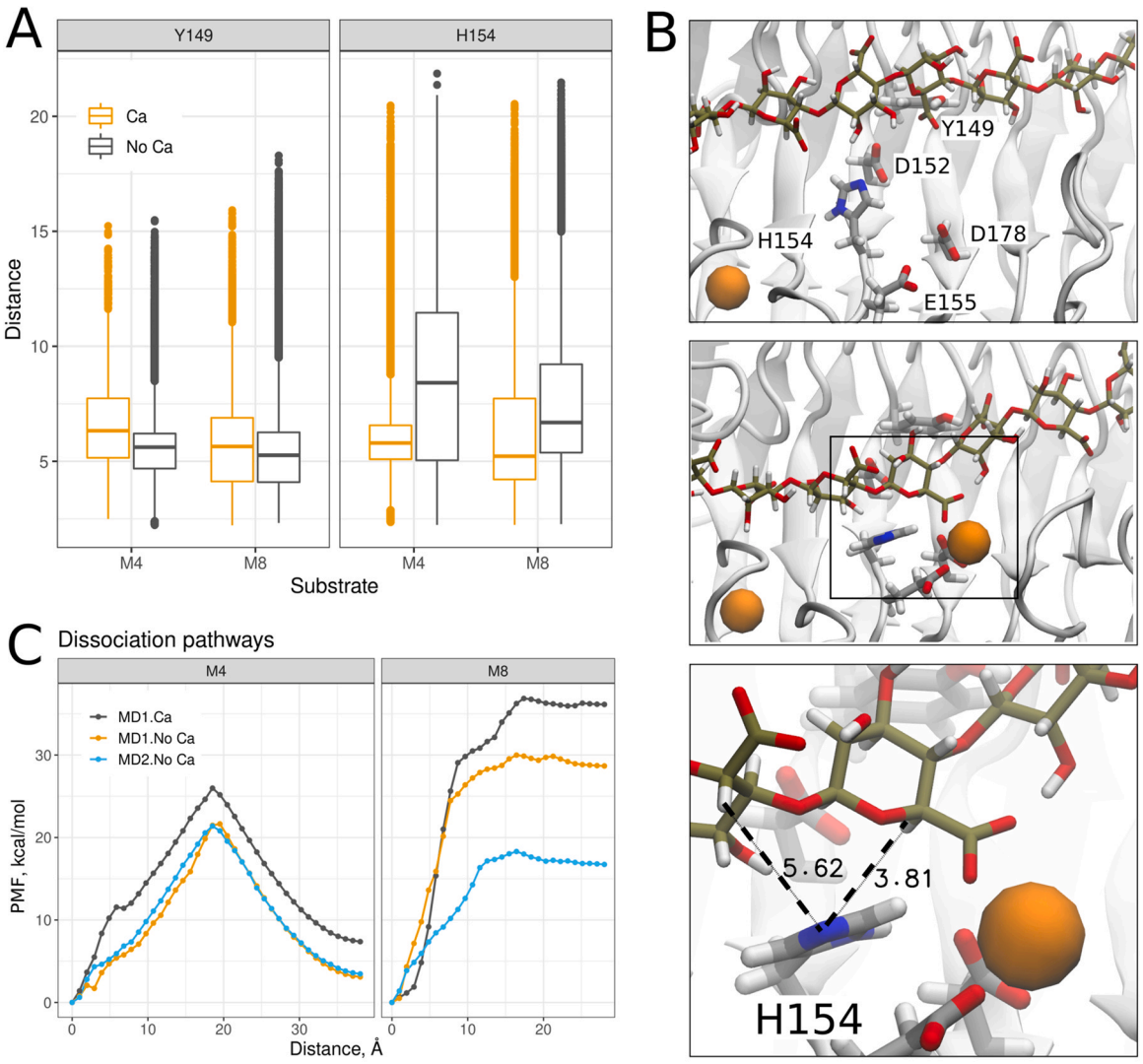

Fig. 4. The substrate interaction with and without calcium. A. Distances between H-5 of the sugar residue in the active site and Y149-OH (left) or H154-Nε (right), with (orange) or without (black) calcium, for all frames from 10 simulation runs where the sugar is bound to the protein. B. Binding poses from stable trajectories without (top) and with (middle) calcium. The inset (bottom) is a close-up showing the distance between the Nε of H154 and two different H-5 of the substrate. A substrate carboxyl group takes part in the calcium coordination. C. The PMF of dissociation of M4 and M8 in the presence (black curves) and absence (orange and blue curves) of calcium. The orange curves are from experiments where the starting pose is taken from an MD simulation with calcium, whereas the blue curves are from experiments where the starting pose is taken from an MD simulation without calcium.

On the other hand, very high G-content alginate is desired for industrial purposes, and engineered AlgEs are able to create alginate with a higher G-content than the native enzymes [46,49,50]. As epimerisation is less effective because of the concomitant production of calcium gels, the possibility to design calcium-independent AlgEs could be worth investigating [51]. Our indications that calcium is coordinated in the active site and takes part in the reaction mechanism implicate that such an enzyme design route could require mutations in the active site of the active A-module. If verified, this also adds to our fundamental knowledge of polysaccharide epimerases.

4. Conclusion

Our overall goal was to use computational methods to elucidate the role of calcium in the A-modules of the AlgEs, including its potential as a charge neutraliser in the catalytic mechanism. US simulations of dissociation showed that AvAlgE6A bound substrate stronger when calcium was present. Through MD simulations we obtained stable coordination configurations for calcium in the active site, where it interacts with the active site sugar carboxyl and where H154 is positioned closer to the H-5 proton than in the absence of calcium. Such a pose is in line with calcium being responsible for the charge neutralizing first step, and we propose that it has a catalytic role in addition to being important for the structural integrity of the AlgEs. This is supported by previously published experimental results where the calcium coordinating residues are found to be essential for catalytic activity, as in the case of E229, which was identified as a direct participant in a calcium binding site. Stable poses obtained through MD trajectories could in the future be used as starting points for quantum mechanical calculations of the reaction mechanism, or inspire experimental validation setups. The implications obtained from this study contribute to a deeper understanding of the role of calcium in the epimerization activity of the AlgEs. This could have importance for industrial tailoring of alginate and form a basis for future enzyme engineering, where it would be

desirable to avoid the need for calcium which inhibits production of alginate with a high G-content.

Funding

The study was funded by the National Science Centre of Poland (Narodowe Centrum Nauki, grant number UMO-2018/31/G/ST4/00246). Computational resources were provided by the local cluster "Piasek" at the Faculty of Chemistry, University of Gdansk.

CRediT authorship contribution statement

All authors have seen and approved submission of the final version of the revised manuscript and, have contributed significantly to its preparation. MG, AL and SAS designed the research; MG and, AL performed the research; MG analyzed data; MG, AL and SAS wrote the paper; and SAS wrote, the grant applications. The research has been performed with full responsibility, and there is no, duplicate publication, fraud or plagiarism. The research meets all applicable standards regarding the, ethics of experimentation and research integrity. All contributors have been acknowledged. None of, the authors of this paper has financial or personal relationships with other people or organizations that could inappropriately influence or bias the content of the paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.03.021.

References

[1] Haug A, Larsen B. Biosynthesis of alginate: Part II. Polymannuronic acid C-5-epimerase from *Azotobacter vinelandii* (Lipman). Carbohydr Res 1971;17:297–308. https://doi.org/10.1016/S0008-6215(00)82537-9
[2] Haug A, Larsen B. Biosynthesis of alginate. Epimerisation of D-mannuronic to L-guluronic acid residues in the polymer chain. Biochim Et Biophys Acta (BBA) - Gen Subj 1969;192:557–9. https://doi.org/10.1016/0304-4165(69)90414-0
[3] Larsen B, Haug A. Biosynthesis of alginate: part I. Composition and structure of alginate produced by *Azotobacter vinelandii* (Lipman). Carbohydr Res 1971;17:287–96. https://doi.org/10.1016/S0008-6215(00)82536-7
[4] Gorin PAJ, Spencer JFT. Exocellular alginic acid from *Azotobacter vinelandii*. Can J Chem 1966;44:993–8. https://doi.org/10.1139/v66-147
[5] Linker A, Jones RS. A new polysaccharide resembling alginic acid isolated from Pseudomonads. J Biol Chem 1966;241:3845–51.
[6] Stanford ECC. On algin: a new substance obtained from some of the commoner species of marine algae. Chem N 1883;47:254–69.
[7] Bisset KA. Evidence from the cytology of *Azotobacter chroococcum* of a relationship with *Rhizobium* and the Bacillaceae. Microbiology 1955;13:442–5. https://doi.org/10.1099/00221287-13-3-442
[8] Campos M, Martínez-Salazar JM, Lloret L, Moreno S, Núñez C, Espín G, et al. Characterization of the gene coding for GDP-mannose dehydrogenase (algD) from *Azotobacter vinelandii*. J Bacteriol 1996;178:1793–9. https://doi.org/10.1128/jb.178.7.1793-1799.1996
[9] Harmsen M, Yang L, Pamp SJ, Tolker-Nielsen T. An update on *Pseudomonas aeruginosa* biofilm formation, tolerance, and dispersal. FEMS Immunol Med Microbiol 2010;59:253–68. https://doi.org/10.1111/j.1574-695X.2010.00690.x
[10] Pier GB, Coleman F, Grout M, Franklin M, Ohman DE. Role of alginate O acetylation in resistance of mucoid *Pseudomonas aeruginosa* to opsonic phagocytosis. Infect Immun 2001;69:1895–901. https://doi.org/10.1128/IAI.69.3.1895-1901.2001
[11] Ertesvåg H, Høidal HK, Schjerven H, Svanem BIG, Valla S. Mannuronan C-5-epimerases and their application for in vitro and in vivo design of new alginates useful in biotechnology. Metab Eng 1999;1:262–9. https://doi.org/10.1006/mben.1999.0130
[12] Tøndervik A, Aarstad OA, Aune R, Maleki S, Rye PD, Dessen A, et al. Exploiting mannuronan C-5 epimerases in commercial alginate production. Mar Drugs 2020;18:565. https://doi.org/10.3390/md18110565
[13] Ertesvåg H, Høidal KH, Hals KI, Rian A, Doseth B, Valla S. A family of modular type mannuronan C-5-epimerase genes controls alginate structure in *Azotobacter vinelandii*. Mol Microbiol 1995;16:719–31. https://doi.org/10.1111/j.1365-2958.1995.tb02433.x
[14] Gawin A, Tietze L, Aarstad OA, Aachmann FL, Brautaset T, Ertesvåg H. Functional characterization of three *Azotobacter chroococcum* alginate-modifying enzymes related to the *Azotobacter vinelandii* AlgE mannuronan C-5-epimerase family. Sci Rep 2020;10:12470. https://doi.org/10.1038/s41598-020-68789-3
[15] Svanem BIG, Skjåk-Bræk G, Ertesvåg H, Valla S. Cloning and expression of three new *Azotobacter vinelandii* genes closely related to a previously described gene family encoding mannuronan C-5-epimerases. J Bacteriol 1999;181:68–77.
[16] Ertesvåg H, Valla S. The A modules of the *Azotobacter vinelandii* mannuronan-C-5-epimerase AlgE1 are sufficient for both epimerization and binding of. J Bacteriol 1999:Ca^{2+}.
[17] Ofstad R, Larsen B. The effect of calcium ion concentration on poly-D-mannuronate C-5 epimerase. Tore Levring (Ed.), Proceedings of the 10th International Seaweed Symposium. Berlin: Walter de Gruyter; 1981. p. 485–93.
[18] Aachmann FL, Svanem BIG, Güntert P, Petersen SB, Valla S, Wimmer R. NMR Structure of the R-module. A parallel β-roll subunit from an *Azotobacter vinelandii* mannuronan C-5 epimerase. J Biol Chem 2006;281:7350–6. https://doi.org/10.1074/jbc.M510069200
[19] Jenkins J, Pickersgill R. The architecture of parallel β-helices and related folds. Prog Biophys Mol Biol 2001;77:111–75. https://doi.org/10.1016/S0079-6107(01)00013-X
[20] Ciccarelli FD, Copley RR, Doerks T, Russell RB, Bork P. CASH – a β-helix domain widespread among carbohydrate-binding proteins. Trends Biochem Sci 2002;27:59–62. https://doi.org/10.1016/S0968-0004(01)02046-1
[21] Rozeboom HJ, Bjerkan TM, Kalk KH, Ertesvåg H, Holtan S, Aachmann FL, et al. Structural and mutational characterization of the catalytic A-module of the mannuronan C-5-epimerase AlgE4 from *Azotobacter vinelandii*. J Biol Chem 2008;283:23819–28.
[22] Hasan M, Khakzad H, Happonen L, Sundin A, Unge J, Mueller U, et al. The structure of human dermatan sulfate epimerase 1 emphasizes the importance of C5-epimerization of glucuronic acid in higher organisms. Chem Sci 2021;12:1869–85. https://doi.org/10.1039/D0SC05971D
[23] Wolfram F, Kitova EN, Robinson H, Walvoort MTC, Codée JDC, Klassen JS, et al. Catalytic mechanism and mode of action of the periplasmic alginate epimerase AlgG. J Biol Chem 2014;289:6006–19.
[24] Gacesa P. Alginate-modifying enzymes: a proposed unified mechanism of action for the lyases and epimerases. FEBS Lett 1987;212:199–202.
[25] Gaardløs M, Heggeset TMB, Tøndervik A, Tezé D, Svensson B, Ertesvåg H, et al. Mechanistic basis for understanding the dual activities of the bifunctional azotobacter vinelandii mannuronan C-5-epimerase and alginate lyase AlgE7. e01836-21 Appl Environ Microbiol 2022;88. https://doi.org/10.1128/aem.01836-21
[26] Svanem BIG, Strand WI, Ertesvåg H, Skjåk-Bræk G, Hartmann M, Barbeyron T, et al. The catalytic activities of the bifunctional *Azotobacter vinelandii* mannuronan C-5-epimerase and alginate lyase AlgE7 probably originate from the same active site in the enzyme. J Biol Chem 2001;276:31542–50.
[27] Garron M-L, Cygler M. Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. Glycobiology 2010;20:1547–73. https://doi.org/10.1093/glycob/cwq122
[28] Franklin MJ, Chitnis CE, Gacesa P, Sonesson A, White DC, Ohman DE. *Pseudomonas aeruginosa* AlgG is a polymer level alginate C5-mannuronan epimerase. J Bacteriol 1994;176:1821–30.
[29] Rehm BHA, Ertesvåg H, Valla S. A new *Azotobacter vinelandii* mannuronan C-5-epimerase gene (algG) is part of an alg gene cluster physically organized in a manner similar to that in *Pseudomonas aeruginosa*. J Bacteriol 1996;178:5884–9.
[30] Gimmestad M, Sletta H, Ertesvåg H, Bakkevig K, Jain S, Suh S, et al. The *Pseudomonas fluorescens* AlgG protein, but not its mannuronan C-5-epimerase activity, is needed for alginate polymer formation. J Bacteriol 2003;185:3515–23.
[31] Samsonov SA, Pisabarro MT. Computational analysis of interactions in structurally available protein–glycosaminoglycan complexes. Glycobiology 2016;26:850–61. https://doi.org/10.1093/glycob/cww055
[32] D.A. Case , R.M. Betz , D.S. Cerutti , T.E.I. Cheatham , T.A. Darden , R.E. Duke , et al. , AMBER 2016, 2016.
[33] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(33–38):27–8. https://doi.org/10.1016/0263-7855(96)00018-5
[34] D.A. Case , H.M. Aktulga , K. Belfon , I.Y. Ben-Shalom , S.R. Brozell , D.S. Cerutti et al. 2021, University of California, San Francisco, 2021.
[35] Nguyen H, Maier J, Huang H, Perrone V, Simmerling C. Folding simulations for proteins with diverse topologies are accessible in days with a physics-based force field and implicit solvent. J Am Chem Soc 2014;136:13959–62. https://doi.org/10.1021/ja5032776
[36] Machado MR, Pantano S. Split the charge difference in two! a rule of thumb for adding proper amounts of ions in MD simulations. J Chem Theory Comput 2020;16:1367–72. https://doi.org/10.1021/acs.jctc.9b00953
[37] Roe DR, Cheatham TEI. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 2013;9:3084–95. https://doi.org/10.1021/ct400341p
[38] Gaardløs M, Samsonov SA, Sletmoen M, Hjørnevik M, Sætrom GI, Tøndervik A, et al. Insights into the roles of charged residues in substrate binding and mode of action of mannuronan C-5 epimerase AlgE4. Glycobiology 2021;31:1616–35. https://doi.org/10.1093/glycob/cwab025

[39] Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, et al. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem 2008;29:622–55. https://doi.org/10.1002/jcc.20820
[40] Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. J Comput Phys 1977;23:187–99. https://doi.org/10.1016/0021-9991(77)90121-8
[41] A. Grossfield, WHAM: the weighted histogram analysis method. Version 2.0, n.d.
[42] Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J Comput Chem 1992;13:1011–21. https://doi.org/10.1002/jcc.540130812
[43] Katz AK, Glusker JP, Beebe SA, Bock CW. Calcium ion coordination: a comparison with that of beryllium, magnesium, and zinc. J Am Chem Soc 1996;118:5752–63. https://doi.org/10.1021/ja953943i
[44] Valla S, Li J, Ertesvåg H, Barbeyron T, Lindahl U. Hexuronyl C5-epimerases in alginate and glycosaminoglycan biosynthesis. Biochimie 2001;83:819–30. https://doi.org/10.1016/S0300-9084(01)01313-X
[45] L. Schrödinger, The PyMOL Molecular Graphics System, Version 2.2.3, 2018.
[46] Stanisci A, Tøndervik A, Gaardløs M, Lervik A, Skjåk-Bræk G, Sletta H, et al. Identification of a pivotal residue for determining the block structure-forming properties of alginate C-5 epimerases. ACS Omega 2020;5:4352–61. https://doi.org/10.1021/acsomega.9b04490
[47] Urtuvia V, Maturana N, Acevedo F, Peña C, Díaz-Barrera A. Bacterial alginate production: an overview of its biosynthesis and potential industrial production. World J Microbiol Biotechnol 2017;33:198. https://doi.org/10.1007/s11274-017-2363-x
[48] Stanisci A, Aarstad OA, Tøndervik A, Sletta H, Dypås LB, Skjåk-Bræk G, et al. Overall size of mannuronan C5-Epimerases influences their ability to epimerize modified alginates and alginate gels. Carbohydr Polym 2018;180:256–63. https://doi.org/10.1016/j.carbpol.2017.09.094
[49] Tøndervik A, Klinkenberg G, Aachmann FL, Svanem BIG, Ertesvåg H, Ellingsen TE, et al. Mannuronan C-5 epimerases suited for tailoring of specific alginate structures obtained by high-throughput screening of an epimerase mutant library. Biomacromolecules 2013;14:2657–66. https://doi.org/10.1021/bm4005194
[50] Buchinger E, Knudsen DH, Behrens MA, Pedersen JS, Aarstad OA, Tøndervik A, et al. Structural and functional characterization of the R-modules in alginate C-5 epimerases AlgE4 and AlgE6 from *Azotobacter vinelandii*. J Biol Chem 2014;289:31382–96.
[51] Petersen AB, Tøndervik A, Gaardløs M, Ertesvåg H, Sletta H, Aachmann FL. Mannuronate C-5 epimerases and their use in alginate modification. EBC20220151 Essays Biochem2023. https://doi.org/10.1042/EBC20220151