Structural and functional characterization of PL28 family ulvan lyase NLR48 from Nonlabens ulvanivorans

ThirumalaiSelvi Ulaganathan¹, Ehud Banin², William Helbert³ and Miroslaw Cygler¹,4,*

From the ¹Department of Biochemistry, University of Saskatchewan, Saskatoon, SK S7N 5E5, Canada; ²Faculty of Life Sciences, Bar Ilan University; ³CERMAV, CNRS and Grenoble Alpes Université, BP53, 38000 Grenoble Cedex 9, France; ⁴Department of Biochemistry, McGill University, Montreal, Quebec H3G 0B1. Canada

Running title: Crystal structure of PL28 family polysaccharide lyase

*To whom the correspondence should be addressed: Miroslaw Cygler, ¹Department of Biochemistry, University of Saskatchewan, Saskatoon, SK S7N 5E5, Canada; miroslaw.cygler@usask.ca; Tel: +1 (306) 966-4361.

Keywords: ulvan lyase, catalytic mechanism, β-jelly roll fold, PL28 family, substrate binding

ABSTRACT

Ulvan is a complex sulfated polysaccharide present in the cell wall of green algae of genus Ulva (Chlorophyta). The first ulvan-degrading polysaccharide lyases have been identified several years ago and more were discovered through genome sequencing of marine bacteria. Ulvan lyases are now grouped in three polysaccharide lyase (PL) families in the CAZy database, PL24, PL25 and PL28. The structures of the representative lyases from families PL24 and PL25 have been recently determined, showing that they adopt a seven-bladed β-propeller fold and utilize the His/Tyr catalytic mechanism. No structural information is yet available for PL28 ulvan lyases. NLR48 from Nonlabens ulvanivorans belongs to PL28 together with its close paralog, NLR42. The biochemical studies of NLR42 elucidated that it can cleave ulvan next to both uronic acid epimers. We report the crystal structure of ulvan lyase NLR48 at 1.9 Å resolution. It has a β-jelly roll fold with an extended, deep and positively charged substrate binding cleft. Putative active site residues were identified from the sequence conservation pattern and their role confirmed by site-directed mutagenesis. The structure of an inactive K162M mutant with a tetrasaccharide substrate shows the substrate occupying the ‘-’ subsites. Comparison with lyases from other PL families with β-jelly roll fold supported assignment of the active site and explained the ability to degrade ulvan next to either epimer of uronic acid. NLR48 contains the His/Tyr catalytic machinery, with Lys162 and Tyr281 playing the catalytic base/acid roles.

Red, brown and green algae are major members of the marine ecosystem and contribute to the huge aquatic biomass (1). Polysaccharides present in the algal cell wall play a structural role providing rigidity and represent a major component of their biomass. They are also a rich source of carbon for cohabiting microorganisms. Carrageenan is the
main cell wall polysaccharide of many red algae, alginate is a cell wall constituent of brown algae and ulvan is present in the green algae belonging to the genus *Ulva* (Chlorophyra). Carrageenan and alginate have been studied extensively and reported to have many industrial applications (2) and the enzymes degrading these polysaccharides have been well characterized (3,4). Ulvan is biosynthesized by the marine green algae genus *Ulva* and *Enteromorpha* spp. and contributes to 8-30% of the dry weight of the algae. This sulfated polysaccharide has recently gained attention for a variety of industrial applications in agriculture, food, pharmaceutical, chemical, and biomaterials industries (1,5,6).

Major components of the ulvan polysaccharide are 3-O-sulfated α-L-rhamnose, α-L-iduronic acid, β-D-glucuronic acid and xylose. Ulvan is made of two main repetitive disaccharide units. Type A ulvanobiouronic-3-sulfate (A3S) has \((\alpha1\rightarrow4)\beta-D-GlcA-(1\rightarrow4)\alpha-L-Rha3S-(1\rightarrow)\) and type B ulvanobiouronic-3-sulfate (B3S) has \((\alpha1\rightarrow4)\alpha-L-IdoA-(1\rightarrow4)\alpha-L-Rha3S(1\rightarrow)\), disaccharide units (Schema 1) (7). The occurrence of a sulfate group in ulvan resembles the glycosaminoglycans present in vertebrates and the presence of uronic acid makes it subject to depolymerization by a lytic mechanism.

Ulvan lyases degrade ulvan through a β-elimination mechanism, cleaving the 1-4 glycosidic linkage on the non-reducing end of a uronic acid residue with the formation of an unsaturated product at the newly formed non-reducing end of the polysaccharide. The first identified ulvan lyase and the first representatives of the PL28 family was NLR42 (NCBI accession number: AEN28574, 46 KDa with two domains) isolated from *Nonlabens ulvanivors* (syn. *Persicivirga ulvanivorans*) (8). The enzyme was able to cleave ulvan next to both GlcA and IdoA. In addition to the catalytic domain, NLR42 also harbors carbohydrate-binding and sorting domains (9).

Genome sequencing of marine microorganisms undertaken in the last few years led to identification of more genes exhibiting ulvan lyase activity (10-12). Polysaccharide degrading machineries occur often as co-regulated gene clusters called Polysaccharide Utilization Loci (PUL). The PUL regions include genes coding for various transport proteins, highly specific carbohydrate acting enzyme (glycoside hydrolases, polysaccharide lyases and sulfatases) and recognition proteins. The complete degradation of the polysaccharides requires the coordination of this different components. We had previously characterized the ulvan PUL from *Alteromonas* and *Pseudoalteromonas* species (13), which led to addition of the family PL24 (LOR107) and PL25 (PLSV3936) to the CAZy database (http://www.cazy.org) (14). Both LOR107 and PLSV3936 display <14% sequence identity, nevertheless they both adopt a β-propeller scaffold. LOR107 is reported to degrade ulvan specifically after the IdoA site (15) whereas PLSV3936 can cleave after both GlcA and IdoA site (16). Recently, new ulvan polysaccharide family PL28 was added to CAZy database.

Here we report first crystal structure and catalytic mechanism of ulvan lyase NLR48 from *Nonlabens ulvanivors* grouped in the PL28 family. NLR48 is a 30 KDa protein that was first identified as an ulvan lyase together with its larger homolog NLR42 (8). NLR48 shows 37% sequence identity with the catalytic module of NLR42 suggesting structural and functional similarity of these enzymes. NLR42 can cleave the ulvan polysaccharide next to both GlcA and IdoA (8). We have determined the crystal structure of NLR48 determined at 1.9 Å resolution. Unlike the previously
characterized ulvan lyases LOR107 and PLSV3936, NLR48 has a β-jelly roll fold. The structure of the NLR active site mutant K162M in complex with a tetrasaccharide substrate facilitated the identification of active site residues and allowed us to propose a catalytic mechanism for ulvan degradation.

RESULTS AND DISCUSSION

We have determined the three-dimensional structure of NLR48, the first representative of the polysaccharide lyase family PL28, which has recently been added to the CAZy database. Several constructs were cloned and tested for expression. Expression yields of His-NLR48 and a shorter construct starting at Asp53, called His-NLR48(53D), were high but the proteins aggregated in solution and were not suitable for crystallization. Therefore, the MBP and GST fusions were tested. The expression level of His-MBP-NLR48 was low in BL21(DE3), but that of the GST-NLR48 fusion was moderate and no aggregation was observed. It was essential to purify the latter construct at 4°C to prevent aggregation. This protein yielded well diffracting crystals.

Crystal structure of NLR48

The three-dimensional crystal structure of NLR48 was solved by a single anomalous dispersion method using SeMet-incorporated protein crystals. There are two molecules in the asymmetric unit, which are related by a non-crystallographic 2-fold rotation. Both molecules are nearly identical, with a root-mean-square deviation (RMSD) of 0.2 Å for all Ca atoms. NLR48 displays a β-sandwich jelly roll fold with two concave β-sheets stacking on each other (Fig. 1A). Each β-sheet has seven antiparallel, strongly bent β-strands. The inner β-sheet is formed from strands β1-β4-β13-β6-β7-β8-β9 while the outer β-sheet contains strands β2-β3-β14-β5-β10-β11-β12 (Fig. 1B). There are seven crossovers between the two sheets. The N-terminal segment, aa 21-50, is devoid of secondary structure elements and is fixed to the loop connecting strands β1 (inner sheet) and β2 (outer sheet) by a disulfide bond between Cys32 and Cys59. The inner β-sheet forms a deep cleft extending across the entire β-sheet in a direction nearly perpendicular to the strands. The bottom of the cleft is formed by the centers of the β-stands, one wall is formed from the loops connecting strands β3 to β4, β5 to β6 (inter sheet), while the other wall is made from the loops connecting β6 to β7, and β8 to β9 (intra inner sheet). The cleft is lined predominantly with basic and polar sidechains and displays a positive electrostatic surface potential as calculated with APBS (17) (Fig. 1C). Two phosphate molecules acquired from the crystallization solution are found in equivalent positions in both molecules. This cleft is likely the substrate binding site of NLR48. A large positive difference map peak was identified in each molecule near the disulfide bond described above. Based on the environment and the coordination geometry, these peaks are assigned as Ca^{2+} ions. Each ion is hepta-coordinated in a pentagonal bipyramid arrangement by OD1 Asp^{79}, OD2 Asp^{79}, OD1 Asn^{85}, O=C Gly^{59} and O=C Ser^{81} in equatorial plane and OD1 Asn^{61} and O=C Ala^{84} in axial positions. The Ca—O distances are in the 2.25-2.50 Å range (Fig. 2A). This Ca^{2+} ion provides additional stability to the N-terminal region of NLR48. This Ca^{2+} binding site is separated from the positively charged cleft, suggesting that Ca^{2+} plays a structural rather than catalytic role in NLR48.

Two phosphate (PO_4^{3-}) ions are located within the positively charged cleft and make many contacts with the surrounding residues (Fig. 2B). PO_4^{3-} at site 1 binds in the middle of the cleft and forms hydrogen bond with N2 Lys^{162}, N2 Gln^{160} and O=C^1 Gly^{160}, NH1 Arg^{17}, O^1 Tyr^{281}. The PO_4^{3-} at site 2 makes H-bonds with NH2 Arg^{216}, O^5 Ser^{171} and backbone
The phosphate in site 1 has lower B-factors suggesting a somewhat stronger binding than in site 2.

**Putative substrate binding and catalytic site**

The β-sandwich jelly roll fold is not unique within the polysaccharide lyases. It is also observed in alginate lyase families PL7, PL14, and PL18, in heparin lyase family PL13 and in exo-glucuronase PL20 (18). Representative structures from these PL families were compared with NLR48. Representative structures for these families were: PL7 – alginate lyase (A1-II; Aly) complexed with tetrasccharide, PDB code 2ZAA (19); PL13 – heparin lyase I with octasaccharide, PDB code 5INA (20); PL18 – alginate lyase aly-SJ02, PDB code 4Q8K (21); PL20 – glucuronan lyase, PDB code 2ZZJ (22); PL14 – *Chlorella* virus enzyme vAL-1, PDB code 3A0N (23). The number of β-strands in the inner and outer sheets varies between these enzymes but the core structures (~100 residues, Fig. 3B) superpose well with that of NLR48 (Table 2). They differ mainly in the loop regions (Fig. 3A). The PL14 lyase is the least similar and is excluded from further analysis. Despite sharing the same fold, the sequence identity between NLR48 and these other lyases is less than 14%. Nevertheless, despite low sequence identity, these enzymes have similar active sites and shared catalytic mechanism belonging to the His/Tyr class (18). The active site is located in the cleft extending along the inner β-sheet.

To identify the putative catalytic site, we first utilized the ConSurf software (24), which showed the highly conserved residues among ~30 NLR48 homologs (Fig. 1D). The solvent accessible and conserved residues located within the cleft are: Glu203, Leu178, Gln160, Arg277 on the bottom of the cleft, Lys162, Asp173 and Arg216 along one wall and Tyr157, Tyr281, Arg117, Arg80 and Asp78 along the other wall. Next, we compared the structure of NLR48 with the lyases from the aforementioned PL families (excluding PL14). The structural alignment of these four lyases showed that within the cleft only Glu119, Gln160 and Tyr281 had counterparts in four other lyase families while Arg117 is replaced by a histidine in PL20, Lys162 is a histidine in PL7, PL13 and PL18 but a leucine in PL20, and Arg277 has lysine counterparts in PL7, PL13 and PL18 (Fig. S1). PL14 lyases differ in this respect and their putative active site residues are located in different region of the cleft (23).

**Site directed mutagenesis of putative active site and substrate binding residues**

Based on these observations the cleft region was considered as a substrate binding site. As such, Arg117, Gln160, Tyr281 and Lys162 were deemed potential candidates for active site residues, with Arg216 being likely involved in substrate binding (H-bonds to phosphates in the native structure). These residues were mutated as follows: R117N, Q160A, K162M, R216N and Y281F. The single mutants were expressed, purified and their activity on the ulvan substrate was measured. The enzymatic activity was monitored as a function of time by following the appearance of unsaturated sugars in the reaction product, with a characteristic absorbance at 232 nm. While the native NLR48 showed easily measurable activity, no measurable activity was observed for the mutants (Fig. 4). This confirmed their involvement in either catalysis or substrate binding.

**Structure of NLR48 mutant with a tetrasccharide substrate**

Three of the aforementioned mutants, K162M, Q160A and Y281F, were screened for crystallization. Well-diffracting crystals could only be obtained for K162M mutant and these
crystals required high concentration of phosphate ions. These crystals were soaked for 2 hours in a solution containing the tetrasaccharide substrate and diffracted to 2.2 Å resolution. They belong to the space group \( P2_12_12_1 \) and are isomorphous to the native NLR48 crystals. The K162M mutation did not introduce any significant changes into the protein structure, the root-mean-square deviation (RMSD) between the native and mutant structures for all Ca atoms is 0.3 Å. The electron density map calculated for the complex structure showed a blob of positive density associated with each molecule. The entire tetrasaccharide could be built into this density (Fig. 5A). The tetrasaccharide does not fully enter the cleft but is docked in the entrance and does not reach into the part where the putative active site is located (residues conserved between the three lyases). We adopt the nomenclature of Davies et al. (25) where the sugars are marked as +1, +2, etc. starting from the first sugar on the reducing end of the cleaved bond and -1, -2, etc. on the non-reducing end. Correspondingly, the sugars binding sites on the protein are referred to as ± subsites. The tetrasaccharide is oriented with its reducing end directed toward the center of the cleft, indicating that it occupies the ‘-‘ subsites in the substrate binding sites.

In contrast to the native structure, the NL48 mutant harbors only one phosphate at site 1. The second phosphate, observed in site 2 of the native structure, is replaced by Rha3S of the substrate in the mutant. The phosphate in site 1 was refined to an occupancy of ~0.7. Nevertheless, high concentration of phosphate in the crystallization conditions might effectively block the center of the cleft from the incoming tetrasaccharide and prevent it from occupying the +1 subsite. Indeed, a comparison with the lyase-oligosaccharide complexes from PL7 and PL13 confirms that the tetrasaccharide in NLR48 occupies ‘-‘ subsites, with the Rha3S on the reducing end occupying the -1 subsite. The first three sugars make multiple contacts with the lyase. Rha3S sugar at the -1 subsite interacts with Tyr281 (at the beginning of the crossover loop following β13), Arg216 (beginning of β9) and Tyr157 (beginning of β6) while its sulfate H-bonds with Arg282 (loop following β13). GlcA sugar at the -2 subsite interacts with Lys284 (crossover loop after β13) and its carboxylic group H-bonds with Ser211 (loop connecting β8 and β9). Finally, Rha2S at the -3 subsite interacts through its sulfate group with Ser213 (loop connecting β8 and β9) (Fig. 5B). The last sugar, AUA, makes no contacts with this molecule of NLR48 nor with symmetry related molecules but its rotational freedom is restricted through the constraints imposed by the presence of bulky substituents at C3 and C5 positions of the preceding Rha3S sugar. Consequently, this sugar is more mobile than the others and has higher B-factors. No aromatic stacking interaction are observed in the lyase-tetrasaccharide contacts. Another factor that may restrict passage of the tetrasaccharide into the cleft is the dynamics of walls motion: they may have to move out to allow the substrate to slide in. In particular, the loop 164-174 with the Asp173 protruding into the cleft could play a gating function in solution. However, in the crystal structure this loop butts against its counterpart in the neighboring molecule related by the non-crystallographic two-fold axis (Fig. 6) and has no ability the cleft expand. Since the complex was obtained by soaking the substrate into a pre-formed crystal, this could be an additional obstacle to the presence of a phosphate in the center of the cleft.

Interestingly, an additional phosphate ion was observed in this structure, near Arg124 and Arg277 that define the ‘+‘ subsites. Although these sidechains have the same conformation in the native structure, there are only water molecules in this area. This new phosphate
binding site might reflect a subtle change of the electrostatic potential when the substrate in occupies the ‘-’ sites.

Catalytic mechanism

The β-elimination catalytic mechanism requires neutralization of the negative charge on the C5 carboxylic group. This reduces the pKa of the C5 proton and makes it vulnerable to abstraction by a base. Concomitant donation of a proton by a protonated amino acid or water molecule to the leaving group results in the cleavage of the O-C4 bond. The end product carries an unsaturated molecule at the newly formed non-reducing end (26). Two general mechanisms have been identified among polysaccharide lyases: (i) metal-assisted catalysis or (ii) His/Tyr dependent catalysis (18,27). Lyases adopting the β-jelly roll fold employ the latter mechanisms. Based on residue conservation, it is likely that NLR48 utilizes the same mechanism.

To define the roles of residues within the cleft we superimposed the three lyase-oligosaccharide complexes, NLR48, heparinase I and alginate lyase (Al-I;Aly), based only on the backbone atoms of the five conserved residues representing the local environment of the active site (Fig. 7A). In this superposition the only sugar that has similar position and orientation is the +1 uronate (Fig. 7B). The +2 and -1 sugars are oriented differently in these structures, reflecting different local shapes of the clefts. Nevertheless, the O1 of Rha3S in the NLR48 complex is in nearly the same place as the corresponding anomeric O atoms in -1 sugars of the other two complexes (Fig. 7B). It is therefore very likely that the +1 GlcA/IdoA sugar is placed in the same way as in the other two lyases. In such a case, Gln160 would serve to neutralize the acidic group on the +1 sugar promoting protonation through formation of two hydrogen bonds (Fig. 7A). Additional stabilization of the acidic group of the +1 sugar is provided by Arg117, which forms hydrogen bond to the acidic group of +1 sugar. The Arg117 - Gln160 arrangement is further stabilized by the formation of hydrogen bonds with a bridging Glu119 (Fig. 7A). The same arrangement is present in the other two lyases. The hydroxyl group of Tyr281 would point toward the bridging oxygen. With GlcA in the +1 site the Tyr281 hydroxyl would be also in close proximity to the C5 proton. This residue could therefore perform the proton abstraction and donate it to the leaving group – the oxygen bridging -1 and +1 sugars (Schema 2, upper left). In the case of IdoA, the C5 proton is on the opposite side of the sugar ring to Tyr281 (Schema 2, bottom left). The only sidechain close to this proton is Lys162, which was mutated to methionine in the NLR48-tetrasaccharide complex (Fig. 5A). There is a histidine in an equivalent position in PL13 lyases that break the bond next to IdoA and in PL7 lyases that break the bond next to guluronate, with same configuration of the C5 carboxylic group as IdoA. In PL20 lyases, which break the bond next to GlcA, the C5 proton is facing the tyrosine and the residue in a position equivalent to Lys162 does not participate in the reaction. Indeed, an isoleucine is found in PL20 family. Since NRL48 can likely cleave the bond next to either epimer of the uronic acid, like its homolog NRL42 (8), Lys162 must be able to act as a base for proton abstraction. This would require an unusual pKa for this sidechain in the context of the enzyme. The pKa could be influenced by the proximity to 3-sulfate of +2 Rha3S. In heparinase I the equivalent residue is a histidine. The role of a lysine residue as a proton acceptor is not unusual; it fulfills the same role in lyases from several other families with different folds, namely PL3, PL9, PL6 and PL11 (18). Overall, both the fold and the catalytic machinery are conserved between
families PL28, PL7, PL13, PL18 and PL20 (Fig. 7).

Comparison with other ulvan lyases

Lyases that degrade ulvan polysaccharide have been identified and characterized only in the last several years. They have led to the addition of three new families to the CAZy classification, namely PL24, PL25 and PL28 family in the CAZy database. We have previously characterized the structure and catalytic mechanism of LOR107 (PL24) and PLSV3936 (PL25). Both ulvan lyases adopt seven-bladed β-propeller fold. PLSV3936 utilizes the His/Tyr catalytic machinery with Arg204, Tyr188 and His123 as neutralizer, catalytic base and acid, respectively (16). LOR107 uses Arg259 to neutralize the C5 carboxylic group and His146 appears to play both catalytic base/acid role aided by His167 and Tyr243 (28). Moreover, LOR107 binds the substrate in two stages, with Arg320 initiating substrate binding at the top of the cleft, followed by substrate sliding deeper into the cleft toward the active site (28).

Despite its different fold, NLR48 shares similarities with the PL24 and PL25 lyases in their catalytic mechanisms. All of them utilize the His/Tyr mechanism, however, the neutralization of the C5 acidic group is achieved by different means. NLR48 uses glutamine that faces the carboxyl and forms a two h-bonds and is supported by a glutamate and an arginine, while PL24/25 utilize an arginine to form two H-bonds with the carboxyl.

CONCLUSIONS

The marine genome sequencing project resulted in the identification many ulvan degrading enzymes. These are classified under PL24, PL25 and PL28 families in the CAZy database. While PL24 and PL25 ulvan lyases have a seven-bladed β-propeller fold, PL28 lyases have a β-jelly roll fold, adding another family to the list of PL enzymes with this architecture. The catalytic mechanism of the PL28 representative lyase, NLR48 from Nonlaben ulvanivorans, is very similar to that of other lyase families with β-jelly roll fold despite very low sequence identity, and is a modification of the His/Tyr catalytic mechanism. The fold similarity of PL28 to several other families of polysaccharide lyases indicates their common evolutionary origin with adaptation to different substrates. Finally, the three families of ulvan lyases utilize a different mode of neutralization of the carboxylic group of the uronic acid.

MATERIALS AND METHODS

Cloning

The gene encoding NLR48 (NCBI accession number: KEZ94336) lacking the first 25 residues as they are predicted to be a signal peptide was cloned into three expression vectors with different tags to improve solubility and to aid in purification. NLR48(26-303) was inserted into pET28 vector using the restriction cloning method, while the ligation independent cloning protocol (29) was used to insert NLR48(26-303) into pRL652 (modified version of pMAL with N-terminal GST tag followed by TEV cleavage site). All mutants used in this study were based on GST-NLR48(26-303) as a template by following the Quickchange site directed mutagenesis protocol. All constructs were verified by DNA sequencing.

Expression and, purification of NLR48

GST-NLR48(26-303) was expressed in BL21(DE3) strain. Overnight inoculum was sub-cultured into 6 L of Terrific Broth media. The cells were induced at ~1.5 OD$_{600}$ with 0.5 mM isopropyl β-D-1-thiogalactopyranoside
(IPTG). After 16 hours of growth at 18°C, cells were spun down at 6200 x g for 20 min. Cell pellet was resuspended in the lysis buffer (20 mM Tris pH 8, 400 mM NaCl and 5% glycerol) and disrupted by sonication. The supernatant was separated from cell debris by centrifugation at 39,000 x g for 55 min. The clarified supernatant was incubated for 2 hours with pre-equilibrated GST resin. Unbound proteins were removed by washing with high salt buffer (lysis buffer supplemented with 200 mM NaCl). All purification steps were performed at 4°C. The GST tag was cleaved on the column by incubation for ~10 h with His-TEV protease, added at a mass ratio of 1:50. The protein eluted from the column was loaded on a Ni-NTA column to remove the TEV protease. The flow-through contained highly pure NLR48 was loaded on size exclusion column for final purification.

**Crystallization of NLR48**

NLR48 was concentrated to 16 mg/ml and screened for crystallization by sitting drop vapor diffusion method in 96-well plate format by mixing 300 nL of the protein mixed with 300 nL of well solution using the Gryphon crystallization robot (ArtRobbins Instruments, Sunnyvale, CA). The duplicate plates were maintained at temperatures of 9°C and 15°C. The best crystals grew from solution containing 0.056 M sodium phosphate monobasic, 1.344 M potassium phosphate dibasic, pH 8.2. Crystals grown at 9°C were somewhat bigger than crystals from 15°C.

**Expression, purification and crystallization of selenomethionine labeled NLR48**

Expression plasmid carrying GST-NLR48 fusion construct was transformed into methionine auxotroph cells B834(DE3) (Novagen). Overnight inoculum grown in 50 ml M9 medium supplemented with methionine was sub-cultured into 1 L of M9 medium. For selenomethionine (SeMet) labeling, the cells were grown at 37°C until OD600 reaches 1.0. The culture was spun gently at 3000 x g for 20 min, the cell pellet was resuspended in 1 L of M9 medium and the culture continued to grow at 37°C for another 4 h to deplete the remaining methionine. Subsequently, 50 mg/L of SeMet was added to the cell culture 30 min prior to the addition of IPTG. The cells were further grown at 18°C. After ~16 hours of growth the cells were collected by centrifugation and SeMet-incorporated protein was purified following the same protocol as used for the native enzyme purification. SeMet-incorporated NLR48 was concentrated to 16 mg/ml for crystallization experiments by hanging drop vapor diffusion method. The crystals grown from 0.49 M sodium phosphate monobasic monohydrate, 0.91 M potassium phosphate dibasic, pH 6.9 at 9°C were cryoprotected with 30% ethylene glycol and diffraction data were collected at the 08BM beamline at the Canadian Light Source. Data was processed with the XDS software package (30) to 1.9 Å resolution and the data collection statistics are shown in Table 1. These crystals diffracted significantly better and showed better spot shape then the native crystals and were used for structure determination and refinement.

**Structure solution and refinement of wildtype NLR48**

The structure of NLR48 was solved by the single anomalous dispersion method with the SeMet dataset using autosol procedure in Phenix software (31). There are two molecules in the asymmetric unit. All six expected Se sites were identified. The initial model contained ~90% of the expected residues. The remaining residues were built manually using Coot (32). Since SeMet-containing crystal diffracted better than wild type NLR48 this dataset was used for not only for solving the structure but also for the refinement. The refinement was
performed with Phenix software interspersed with manual rebuilding using Coot. The final $R_{\text{work}} = 0.169$ and $R_{\text{free}} = 0.201$. Several N-terminal residues are not visible in the electron density. The model contains residues 44-303 in both chains, two Ca$^{2+}$ ions, four phosphates and 710 solvent molecules. The validation by Molprobity (33) identified Arg282 and Val283 as Ramachandran outliers. However, the inspection of electron density map confirms the correct modeling of these residues. Data collection and statistics are shown in Table 1. The coordinates and structure factors have been deposited to the Protein Data Bank with accession number 6D2C.

**Site directed mutagenesis**

Selected residues were mutated to assess their role in catalysis. The mutant R117A, Q160A, K162M, R217N and Y281F were made using the GST-NLR48 fusion construct as a template by following the Quickchange site directed mutagenesis protocol. All the constructs were confirmed by sequencing. The mutants were expressed in BL21(DE3) cells and purified using the same protocol as for the wildtype protein.

**Enzyme activity**

The activity of NLR48 and its various mutants was measured by monitoring the absorption at 232 nm, indicative of the formation of an unsaturated double bond in the sugar on the non-reducing end of one of the two reaction products. The assay solution contained 0.5 mg/ml ulvan polysaccharide in 20 mM Hepes, pH 7.5, 150 mM NaCl. 1μg of the enzyme was added directly to the assay solution with constant stirring and the absorbance was measured as a function of time. The absorbance of the substrate alone was used as a control.

**Expression, purification and crystallization of NLR48 mutants**

Mutants showing no activity, namely K162M, Y281F and R117N were expressed and purified. The GST fusion constructs of the mutants were expressed and purified using the same protocol as for wild type NLR48. Of these, only K162M expressed in high yield and was screened for crystallization at 17 mg/ml. Crystals were obtained in the same conditions as wild type NLR48 (0.49 M sodium phosphate monobasic monohydrate, 0.91 M potassium phosphate dibasic, pH 6.9 at 9 °C).

**Structure determination of NLR48-tetrasaccharide complex**

The K162M mutant crystals were used for soaking experiments with an ulvan tetrasaccharide, which contained a mixture of ΔUA-Rha3S-GlcA-Rha3Sα/β and ΔUA-Rha3S-IdoA-Rha3Sα/β. Briefly, a drop containing 3 μl of substrate solution (~50 mM tetrasaccharide solution in water) mixed with 3 μl of crystallization well solution on a cover slip was equilibrated against 1 ml of well solution. The tetrasaccharide was prepared as described before (28). The crystals were soaked in the substrate solution for ~2 hours, cryoprotected in well solution supplemented with 30% ethylene glycol and flash frozen in liquid N$_2$. Diffraction data were collected at 08ID beamline, Canadian Light Source. These crystals diffracted to 2.2 Å resolution and were isomorphous with the wild type NLR48 crystals. One round of rigid body refinement was done on the complex dataset with the Phenix refinement program. The tetrasaccharide was built into positive difference electron density with Coot. Further refinement was done using Phenix software. The complex structure was evaluated for stereochemistry using MolProbity. The final model has 44-303 residues in each molecule, 212 water molecules, two calcium ions, two...
tetrasaccharides and four phosphate ions. Data collection and statistics are shown in Table 1. The coordinates and structure factors have been deposited to the Protein Data Bank with accession number 6D3U.

ACKNOWLEDGEMENTS: We would like to thank Mr. Kevin Voth for his comments. This research was supported by grant 155375-2012-RGPIN from the Natural Science and Engineering Council (NSERC) to MC. We acknowledge the Protein Characterization and Crystallization Facility, College of Medicine, University of Saskatchewan for access to the crystallization robot. Research described in this paper was performed using beamline 08ID-1 at the Canadian Light Source, which is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research.

CONFLICT OF INTEREST: The authors declare that they have no conflicts of interest with the contents of this article.

AUTHOR CONTRIBUTIONS: T.U. expressed and purified the protein and mutants, performed crystallographic studies and contributed to writing the manuscript. W.H. prepared and purified tetrasaccharides, contributed biochemical characterization, data analysis and manuscript writing. E.B. provided the gene and contributed to the experimental design and data analysis and interpretation. M.C. contributed to the experimental design, data analysis, interpretation and manuscript writing.

REFERENCES
1. Shalaby, E. A. (2011) Algae as promising organisms for environment and health. Plant Signaling & Behavior 6, 1338-1350
2. Sudha, P. N., Aisverya, S., Nithya, R., and Vijayalakshmi, K. (2014) Chapter Eight - Industrial Applications of Marine Carbohydrates. in Adv. Food Nutr. Res. (Kim, S.-K. ed.), Academic Press. pp 145-181
3. Hehemann, J.-H., Boraston, A. B., and Czjzek, M. (2014) A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. Curr. Opin. Struct. Biol. 28, 77-86
4. Michel, G., and Czjzek, M. (2013) Polysaccharide-degrading enzymes from marine bacteria. in Marinwe Enzymes for Biocatalysis (Trincone, A. ed.), Woodhead Publishing Ltd, Cambridge. pp 429-464
5. Cunha, L., and Grenha, A. (2016) Sulfated Seaweed Polysaccharides as Multifunctional Materials in Drug Delivery Applications. Marine Drugs 14
6. Popa, E. G., Reis, R. L., and Gomes, M. E. (2015) Seaweed polysaccharide-based hydrogels used for the regeneration of articular cartilage. Crit. Rev. Biotechnol. 35, 410-424
7. Lahaye, M., and Robic, A. (2007) Structure and Functional Properties of Ulvan, a Polysaccharide from Green Seaweeds. Biomacromolecules 8, 1765-1774
8. Nyvall Collen, P., Sassi, J. F., Rogniaux, H., Marfaing, H., and Helbert, W. (2011) Ulvan lyases isolated from the Flavobacteria Persicivirga ulvanivorans are the first members of a new polysaccharide lyase family. *J. Biol. Chem.* **286**, 42063-42071

9. Melcher, R. L. J., Neumann, M., Fuenzalida Werner, J. P., Gröhn, F., and Moerschbacher, B. M. (2017) Revised domain structure of ulvan lyase and characterization of the first ulvan binding domain. *Scientific Reports* **7**, 44115

10. Kopel, M., Helbert, W., Henrissat, B., Doniger, T., and Banin, E. (2014) Draft Genome Sequence of *Pseudoalteromonas* sp. Strain PLSV, an Ulvan-Degrading Bacterium. *Genome announcements* **2**(6)

11. Kopel, M., Helbert, W., Henrissat, B., Doniger, T., and Banin, E. (2014) Draft Genome Sequences of Two Ulvan-Degrading Isolates, Strains LTR and LOR, That Belong to the *Alteromonas* Genus. *Genome announcements* **2**(5)

12. Kopel, M., Helbert, W., Henrissat, B., Doniger, T., and Banin, E. (2014) Draft Genome Sequence of Nonlabens ulvanivorans, an Ulvan-Degrading Bacterium. *Genome announcements* **2**(4)

13. Foran, E., Buravenkov, V., Kopel, M., Mizrahi, N., Shoshani, S., Helbert, W., and Banin, E. (2017) Functional characterization of a novel “ulvan utilization loci” found in Alteromonas sp. LOR genome. *Algal Research* **25**, 39-46

14. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-D495

15. Kopel, M., Helbert, W., Belnik, Y., Buravenkov, V., Herman, A., and Banin, E. (2016) New Family of Ulvan Lyases Identified in Three Isolates from the Alteromonadales Order. *J. Biol. Chem.* **291**, 5871-5878

16. Ulaganathan, T., Boniecki, M. T., Foran, E., Buravenkov, V., Mizrachi, N., Banin, E., Helbert, W., and Cygler, M. (2017) New Ulvan-Degrading Polysaccharide Lyase Family: Structure and Catalytic Mechanism Suggests Convergent Evolution of Active Site Architecture. *ACS chemical biology* **12**, 1269-1280

17. Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G. W., Holst, M. J., McCammon, J. A., and Baker, N. A. (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112-128

18. Garron, M. L., and Cygler, M. (2010) Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology* **20**, 1547-1573

19. Ogura, K., Yamasaki, M., Mikami, B., Hashimoto, W., and Murata, K. (2008) Substrate recognition by family 7 alginate lyase from *Sphingomonas* sp. A1. *J Mol.Biol.* **380**, 373-385
20. Han, Y. H., Garron, M. L., Kim, H. Y., Kim, W. S., Zhang, Z., Ryu, K. S., Shaya, D., Xiao, Z., Cheong, C., Kim, Y. S., Linhardt, R. J., Jeon, Y. H., and Cygler, M. (2009) Structural Snapshots of Heparin Depolymerization by Heparin Lyase I. *J. Biol. Chem.* **284**, 34019-34027

21. Dong, S., Wei, T.-D., Chen, X.-L., Li, C.-Y., Wang, P., Xie, B.-B., Qin, Q.-L., Zhang, X.-Y., Pang, X.-H., Zhou, B.-C., and Zhang, Y.-Z. (2014) Molecular Insight into the Role of the N-terminal Extension in the Maturation, Substrate Recognition, and Catalysis of a Bacterial Alginate Lyase from Polysaccharide Lyase Family 18. *J. Biol. Chem.* **289**, 29558-29569

22. Konno, N., Ishida, T., Igarashi, K., Fushinobu, S., Habu, N., Samejima, M., and Isogai, A. (2009) Crystal structure of polysaccharide lyase family 20 endo-beta-1,4-glucuronan lyase from the filamentous fungus Trichoderma reesei. *FEBS Lett.* **583**, 1323-1326

23. Ogura, K., Yamasaki, M., Yamada, T., Mikami, B., Hashimoto, W., and Murata, K. (2009) Crystal Structure of Family 14 Polysaccharide Lyase with pH-dependent Modes of Action. *J. Biol. Chem.* **284**, 35572-35579

24. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **33**, W299-302

25. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559

26. Gacesa, P. (1987) Alginate-modifying enzymes. A proposed unified mechanism for the lyases and epimerases. *FEBS Lett.* **212**, 199-202

27. Garron, M. L., and Cygler, M. (2014) Uronic polysaccharide degrading enzymes. *Curr. Opin. Struct. Biol.* **28C**, 87-95

28. Ulaganathan, T., Helbert, W., Kopel, M., Banin, E., and Cygler, M. (2018) Structure–function analyses of a PL24 family ulvan lyase reveal key features and suggest its catalytic mechanism. *J. Biol. Chem.*

29. Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**, 6069-6074

30. Kabsch, W. (2010) Xds. *Acta crystallogr. D* **66**, 125-132

31. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.D* **66**, 213-221

32. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta crystallogr. D* **66**, 486-501

33. Chen, V. B., Arendall Iii, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.D* **66**, 12-21
34. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714-2723
Table 1. Data collection and refinement statistics.

|                          | NLR48(SeMet)          | NLR48(K162M)-tetrasaccharide |
|--------------------------|-----------------------|-----------------------------|
| Crystal form             |                       |                             |
| Space group              | $P 2_1 2_1 2_1$       | $P 2_1 2_1 2_1$             |
| a, b, c (Å), β (°)       | 82.5, 102.1, 103.7    | 82.5, 102.7, 103.0          |
| Wavelength (Å)           | 0.98013*              | 0.97949                     |
| Resolution (Å)           | 50-1.9                | 50-2.2                      |
| Observed hkl             | 1,001,093             | 432,911                     |
| # Unique hkl             | 131,854               | 44,398                      |
| Completeness (%)         | 99.8 (97.7)           | 99.8 (97)                   |
| Redundancy               | 7.6                   | 9.8                         |
| $R_{\text{meas}}$        | 0.143 (1.25)          | 0.100 (1.16)                |
| $I/(\sigma I)$           | 10.01 (1.82)          | 14.86 (2.0)                 |
|                          |                       |                             |
| Refinement statistics    |                       |                             |
| $R_{\text{work}}$        | 0.169                 | 0.189                       |
| $R_{\text{free}}$        | 0.201                 | 0.220                       |
| Wilson B (Å$^2$)         | 26.28                 | 49.57                       |
| Ramachandran plot (%)    |                       |                             |
| Favored                  | 96.57                 | 95.35                       |
| Allowed                  | 2.67                  | 3.88                        |
| Outliers                 | 0.76                  | 0.78                        |
| R.m.s. deviation         |                       |                             |
| Bonds (Å)                | 0.015                 | 0.008                       |
| Angles (°)               | 1.214                 | 1.091                       |
| PDB code                 | 6D2C                  | 6D3U                        |

The dataset for structure determination was collected at the wavelength of 0.97864. Dataset used for refinement was collected at the wavelength 0.98013 as shown in the table.
Table 2. Superposition of polysaccharide lyase with a β-sandwich jelly roll fold on NLR48.

| Name  | PL family | No. of overlapping Ca | RMSD (Å) |
|-------|-----------|------------------------|----------|
| 4Q8K  | PL18      | 100                    | 1.47     |
| 2ZZJ  | PL20      | 97                     | 1.55     |
| 3INA  | PL13      | 93                     | 1.60     |
| 2ZAA  | PL17      | 125                    | 1.49     |
| 3A0N  | PL14      | 83                     | 1.84     |
Schema 1

[\rightarrow 4]-%D-GlcA-(\rightarrow 4)-{\alpha}-L-Rha3S-(\rightarrow ]

-2 GlcA
-1 Rha3S
+1 GlcA
+2 Rha3S

-1 Rha3S
+1 IdoA
+2 Rha3S

[\rightarrow 4]-{\alpha}-L-IdoA-(\rightarrow 4)-{\alpha}-L-Rha3S-(\rightarrow ]

-2 IdoA

Schema 1. Diagram of a tetrasaccharide fragment of ulvan containing either GlcA or IdoA. The product of the cleavage by lyase is shown on the right.
Schema 2. Proposed mechanism of catalysis. Tyr281 acts as a base and acid for the cleavage near GlcA, while for IdoA the base is Lys162 and the acid is Tyr281.
Figure 1. Structural details of NLR48. A) Stereo view of the cartoon representation of the structure. The molecule is rainbow colored from blue at the N-terminus to red at the C-terminus. The β-strands are numbered sequentially along the protein sequence. This and other figures were made with PyMol (Schrödinger, pymol.org); B) Topology diagram showing the connectivity of the strands. The strands are similarly rainbow colored as in the cartoon figure. The inner sheet is on light blue background, the outside sheet is on pink background; C) View of the electrostatic potential projected on the protein surface within the cleft in the concave side of the protein. The cleft displays strongly positive potential matching the negatively charged polysaccharide substrate; D) Residue conservation among sequence homologs of NLR48 as calculated by Consurf (24). The highly conserved residues are colored burgundy. They surface representation shows their concentration within the canyon.
Figure 2. Ion binding sites in the native structure of NLR48. A) The coordination sphere of the Ca$^{2+}$ ion, which plays a structural role in this enzyme; B) the environment of the PO$_4^{3-}$ ions.
Figure 3. Stereo view of the superposition of polysaccharide lyases that display the β-sandwich jelly roll fold. **A)** Representative structures from families PL7, PL13, PL14, PL18 and PL20 were superimposed on NLR48 using the sPDBViewer (34). NLR48 is colored green, HepI (PL13) colored magenta, alginate lyase (PL7) colored cyan, PL18 colored forest green and PL20 colored gray; **B)** The core ~100 residues of NLR48 that superimpose well among the PL lyases colored red, the less structurally conserved segments are colored green. Residues Arg117, Gln160, Lys162, Arg216 and Tyr281 that were mutagenized are shown explicitly in ball-and-stick representation.
Figure 4. The activity measurements of the wild type NLR48 and several mutants on ulvan polysaccharide. Only the native NLR48 showed measurable activity.
Figure 5. Tetrasaccharide substrate bound to NLR48. A) The final 2Fo-DFc map contoured at 1.5 σ with the superimposed model of the tetrasaccharide; B) Close-up of the tetrasaccharide and its contacts with NLR48.
Figure 6. Packing of a neighboring molecule against the loop forming one wall of the substrate binding site. This tight packing restricts the movement of this loop and partially block the entry into the active site.
Figure 7. Stereo view of the superposition of enzyme-oligosaccharide complexes based on the backbone of five conserved residues in the cleft (Arg117, Glu119, Gln160, Arg277 and Tyr281 and their equivalences in other lyases). **A)** The key residues are displayed in a stick representation. NLR48 – green carbons, PL7 (2ZAA) – cyan carbons, PL13 (2INA) – yellow carbons. Only the +1 sugar in PL7 and PL13 is shown for clarity. The Rha3S in -1 subsite of NLR48 is shown. The hydrogen bond network between the acidic group of +1 sugar, Gln, Arg and Glu is shown with cyan dashed lines; **B)** Overlay of the oligosaccharides from the three lyases from the above superposition. Only -2 to +2 sugars are shown. NLR48 – green carbons, PL7 (2ZAA) – pink carbons, PL13 (2INA) – yellow carbons. Only the +1 sugar has similar position and orientation as is required for the catalysis. The orientation of the ‘-’ and ‘+’ sugars differ to accommodate well into the clefts of individual enzymes. The O1 of Rha3S in NLR48 is very close to the oxygen bridging -1 and +1 sugars suggesting that the +1 sugar would take similar position to that observed in the other two lyases.
