Construction and analyses of hybrid *Azotobacter vinelandii* mannuronan C-5 epimerases with new epimerization pattern characteristics

Tonje M. BJERKAN, Bjørn E. LILLEHOV, Wenche I. STRAND, Gudmund SKJÅK-BRÆK, Svein VALLA and Helga ERTESVÅG¹

Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

The secreted mannuronan C-5 epimerases from *Azotobacter vinelandii* form a family of seven homologous modular type enzymes, which appear to have evolved through duplications and point mutations in the individual modules. The catalytic A modules of these enzymes are responsible for generating the characteristic sequence distribution patterns of G residues in the industrially important polymer alginate by epimerizing M (β -D-mannuronic acid) moieties to G (α -L-guluronic acid). Forty-six different hybrid enzymes were constructed by exchanging parts of the sequences encoding the A modules of AlgE2 (generates consecutive stretches of G residues) and AlgE4 (generates alternating structures). These hybrid enzymes introduce a variety of new mono-

mer-sequence patterns into their substrates, and some regions important for the subsite specificity or processivity of the enzymes were identified. By using time-resolved NMR spectroscopy, it became clear that the rates for introducing alternating structures and consecutive stretches of G residues are different for each enzyme, and that it is the ratio between these rates that determines the overall epimerization pattern. These findings open up new possibilities in biotechnology and in studies of the many biological functions of alginates.

Key words: alginate, *Azotobacter vinelandii*, enzyme mechanism, mannuronan C-5 epimerase, time-resolved NMR.

INTRODUCTION

Alginate is a widely used polysaccharide with many industrial applications. It is composed of two monomers, M (β -D-mannuronic acid) and its C-5 epimer G (α -L-guluronic acid), linked to each other by 1,4 bonds [1]. The polymer is produced by brown algae and by some of the bacteria belonging to the genera *Pseudomonas* and *Azotobacter*. Probably, in all alginate-producing organisms, the polymer is first made as mannuronan, and the G residues originate from mannuronan C-5 epimerases acting at the polymer level [2]. The C-5 epimerases involved in the biosynthesis of heparin/heparan sulphate and dermatan sulphate also act at the polymer level and share some sequence similarity with the mannuronan C-5 epimerases [2].

In contrast with most other polysaccharides, alginates are not composed of monotonously repeating units. Specifically, the relative content of M and G residues and their sequence distributions vary widely among alginates produced by different species, and even within different parts of one organism [1,3]. This variability has a profound effect on the physicochemical properties of the polymer, and one particularly important parameter is the frequency and length of contiguous stretches of G residues referred to as G blocks. These blocks form the basis for the reversible and bivalent cation-based (typically Ca^{2+}) formation of gels [1]. In our laboratory, we have used *Azotobacter vinelandii* to study the origin of the variability in alginate structures. It has been found that the genome of this organism encodes a seven-member family (AlgE1–7) of Ca^{2+} -dependent mannuronan C-5 epimerases, which are secreted to the surface and extracellular environment [4–6]. Interestingly, after expression in *Escherichia coli*, each epimerase has been found to introduce its own characteristic sequence distribution of G residues in the substrate, explaining the extensive variability in the monomer sequences of the alginates produced by *A. vinelandii* [7]. It is probable that this ability is biologically

significant for the formation of a particular resting stage, designated cyst, characteristic of this genus [8–10].

Each enzyme is composed of a varying number of two types of modules: A (approx. 385 amino acids) and R (approx. 150 amino acids). By using AlgE1 as a model, it has been shown that the A module alone is sufficient both for epimerization and for the determination of the sequence distribution [11]. As the amino acid sequences of all A modules are very similar to each other, it follows that a limited number of residues must be responsible for each of the specific epimerization patterns. The molecular explanation for this is of both biotechnological and biological significance and, in the present study, we describe the construction of hybrid epimerases by exchanging, at the DNA level, parts of the A modules between AlgE2 and AlgE4. These enzymes were chosen as models because AlgE2 is a G-block-forming enzyme, whereas AlgE4 almost exclusively generates alternating sequences known as MG blocks or $(\text{MG})_n$ in the polymer chains [12,13]. The resulting hybrids were expressed in *E. coli* and studied biochemically. It was found that all the hybrids were catalytically active and many of them generated alginates different from those of their parent enzymes. We believe that these discoveries are of both fundamental and applied interest and should stimulate a variety of further studies.

EXPERIMENTAL

Growth of bacteria

Bacteria were grown at 37 °C in L broth (10 g/l tryptone, 5 g/l yeast extract and 5 g/l NaCl) with shaking or on L agar. The media were supplemented with 200 $\mu\text{g/ml}$ ampicillin or 12.5 $\mu\text{g/ml}$ tetracycline when appropriate. For enzyme purification, *E. coli* JM109, containing the plasmid of interest, was grown in a medium containing 30 g/l tryptone, 15 g/l yeast extract and 5 g/l NaCl.

Abbreviations used: G, α -L-guluronic acid; M, β -D-mannuronic acid.

¹ To whom correspondence should be addressed (e-mail helga.ertesvag@biotech.ntnu.no).

Construction of plasmids

Standard recombinant DNA procedures were performed as described by Sambrook and Russell [14], whereas chemical transformations were performed as described by Chung et al. [15]. *E. coli* JM109 [16] was used in most of the study and *E. coli* JM110 [16] was used for producing unmethylated vector DNA. *E. coli* ES1301-mutS (Promega, Madison, WI, U.S.A.) was used for transforming mutagenized plasmids. The expression vector pTrc99A [17] was used for the production of all enzymes. When needed for mutagenesis, parts of the genes were transferred to pAlter-1 (Promega). Restriction sites in the coding regions were then inserted, using the pAlter site-specific mutagenesis system (Promega) and the helper phages R408 and M13K07. In two cases, mutagenesis resulted in the introduction of the amino acid of AlgE4 into the AlgE2 sequence. The different constructs (see Figure 3) were made by exchanging homologous parts of the genes using earlier constructs as donors and recipients. Plasmid pHH2 [18] was used as a source for *algE2* and plasmids pHH4 [12] and pHH43 [12] as sources for *algE4*. The oligonucleotides used for mutagenesis were (changed nucleotides are underlined, and the introduced restriction site and any amino acid changes are shown in parentheses):

- (1) 5'-CAACTCCACCTACGGGATCCAGGAGCGC-3' (*Bam*HI, AlgE2 and AlgE4),
 - (2) 5'-ACGGCTTCAACGTGGTCACCAGCACC-3' (*Bst*EII, AlgE2, Ile²⁰⁰ → Val²⁰⁰),
 - (3) 5'-ACCTGCCCCGGGACCGGCCAGCA-3' (*Sma*I, AlgE2),
 - (4) 5'-ACTTCGGCATGCGCGACCTGAC-3' (*Sph*I, AlgE2, Ser¹⁰³ → Arg¹⁰³),
 - (5) 5'-GCGAACCCGGGGCGACACCGCA-3' (*Sma*I, AlgE4),
 - (6) 5'-CAACAACGTCGCGTACGGCAACGGC-3' (*Bsi*WI, AlgE2 and AlgE4),
 - (7) 5'-TGAAATGTCCGGATACGGTTTCGATC-3' (*Bsp*EI, AlgE2),
 - (8) 5'-CACCTGCAGAATGCCGAGATCTA-3' (*Bsm*I, AlgE2),
 - (9) 5'-CACCTGCAGAATGCCGATATCCA-3' (*Bsm*I, AlgE4)
- and
- (10) 5'-CGAGATGTCCGGATACGGCTTCGAC-3' (*Bsp*EI, AlgE4).

Measurement of epimerase activities by radioisotope assay and NMR

Enzyme extracts were prepared as described by Ertesvåg et al. [7], except that 20 mM Mops (pH 6.9), containing 3 mM CaCl₂, was used both in the cell disruption and in the purification steps. For some experiments, clarified, but not FPLC-purified, enzyme extracts were used. Epimerase activities were quantified by measuring the liberation of tritium from [5-³H]alginate to water as described previously [19]. For NMR spectroscopy, unlabelled alginate containing less than 5% G was used as substrate, except for the enzyme KA1, for which 38 of the spectra were obtained using pure mannuronan as substrate. These spectra do not differ significantly in epimerization pattern from those obtained using alginate containing less than 5% G as substrate. The epimerizations were performed in a total volume of 6 ml containing 20 mM Mops (pH 6.9), 3 mM CaCl₂ and 7.5 mg of alginate. The desired degrees of epimerization were achieved by varying the amount of enzyme or the incubation time. NMR spectra were recorded on a Bruker DPX 400 (400 MHz) spectrometer, and the fractions F_G, F_M, F_{GG}, F_{MM} and F_{GM, MG} were calculated from the integrated spectra as described by Grasdalen [20]. Labelled and unlabelled M-rich alginate were prepared

from *Pseudomonas aeruginosa* [21] and contained less than 5% G. Mannuronan (100% M) was prepared from a *P. fluorescens* mutant harbouring a defective epimerase [22].

Time-resolved ¹H-NMR spectroscopy

Before NMR recording, hybrid epimerases were partially purified by FPLC, dialysed against Mops [10 mM BL6, KA1 and TB5 (pH 6.8) or 5 mM TB14 (pH 6.9), all containing 1 mM CaCl₂] and freeze-dried. The amounts of active enzyme in the samples were then estimated by using the Bio-Rad Coomassie Brilliant Blue-based protein assay and estimating the purity of the proteins on a Coomassie Blue-stained SDS/polyacrylamide gel.

Individual solutions in ²H₂O of the different components, mannuronan (9 mg), CaCl₂ (2.5 mM for TB5 and TB14, and 5 mM for BL6 and KA1) and epimerase [BL6 (0.15 mg), KA1 (0.18 mg), TB5 (0.05 mg) and TB14 (0.23 mg), these are the estimated amounts] were prepared separately. The calculated volumes, the amounts given above refer to the mixed solution, were then transferred to an NMR sample tube (total volume, 0.5 ml).

Spectra were recorded on a Bruker DPX 400 (400 MHz) spectrometer. To monitor the progress of a single epimerization, a series of several hundred successive ¹H-NMR spectra were recorded as described by Hartmann et al. [13]. Each spectrum was calculated from 64 scans and represents an average of 8 min. We chose to set the time for each result as the end-time for each uptake. Between each uptake, there was a pause of 1 min. The scanning for the first spectrum was started 10 min after the addition of the enzyme. The reactions were continued until the fraction of G no longer increased or for a maximum of 250 spectra.

RESULTS

Construction of hybrid epimerases by DNA-level exchanges of segments within the A modules of AlgE2 and AlgE4

To facilitate construction of hybrid enzymes and to simplify interpretations of results, we first synthesized a derivative of *algE2* in which the DNA region encoding the four R modules was substituted with the single R-module-encoding region of *algE4*. The new protein, BL6, is thus similar to AlgE4 in terms of its modular structure (Figure 1A), and the two proteins differ only in the amino acid sequences of their A modules (Figure 1B). Consequently, any differences in enzyme properties between BL6 and AlgE4 must originate from the sequences of their A modules. The genes for these two enzymes and for AlgE2 were expressed in *E. coli* and the corresponding proteins were used to epimerize mannuronan. Since we had shown previously, using AlgE1 as a model, that the epimerization patterns are determined mainly by the A module [11], we expected BL6 to generate a pattern similar to that of AlgE2. This also turned out to be the case (Figure 2). BL6 and AlgE4 could therefore be used as tools for a more detailed study of the structure–function relationships.

As it seemed most probable that more than one amino acid is responsible for the differences in the epimerization patterns generated by BL6 and AlgE4, we decided to exchange A-module-encoding segments between these enzymes. The new hybrids were then expressed and analysed with respect to their catalytic activities and epimerization patterns. To simplify the construction work, we identified unique restriction sites in one or both of the A-module-encoding DNA sequences. In several cases, new restriction sites were made by site-specific mutagenesis in one or both of the genes, such that the same sites were present in both genes at the corresponding position, according to the alignment

Figure 1 Comparison of the primary structures of AlgE2, AlgE4 and the hybrid enzyme BL6

(A) The modular structures of the mannuronan C-5 epimerases AlgE2, AlgE4 and the hybrid enzyme BL6. Identical patterns indicate identical modules. The Ca²⁺-binding motifs are indicated as vertical lines above the R modules. (B) Amino acid sequences of the A modules of AlgE2 and AlgE4. Identical amino acids are shown as dots in the AlgE2 sequence. The sequences are divided into nine regions bordered by restriction sites.

shown in Figure 1(B). In no cases were amino acids introduced, which are not already present in AlgE4. This enabled us to divide the A modules into nine regions (numbered 1–9 in Figure 1B). With these new constructs available, we could easily make the desired hybrid genes, and 46 such hybrids were constructed (Figure 3B). All the enzymes were expressed in *E. coli* and, interestingly, activity measurements (radioisotope assays) showed that they were all sufficiently active to allow standard incubation conditions in the subsequent experiments.

Analyses of the epimerization patterns generated by the hybrid enzymes

The best way to analyse the epimerization pattern generated by an epimerase is to use NMR spectroscopy. However, since the epimerases are found to respond differently to the changes in the substrate, which take place as it becomes epimerized, spectra at several different degrees of epimerization need to be obtained for each enzyme [7]. This procedure is too laborious to be used for a detailed characterization of as many as 47 enzymes (the hybrids plus BL6). As a preliminary screening, we obtained one or more NMR spectra for all the enzymes at various degrees of epimerization of a substrate initially containing less than 5% G. The epimerization patterns varied widely among the enzymes, indicating that the hybrids in many cases generate patterns that are different from both of their parent enzymes (Figure 3A). These observations are important, because they open new possibilities both in terms of biotechnological applications and for basic studies of the enzymes both *in vitro* and *in vivo*.

To assist in the inspection of the results shown in Figure 3(A), we have included lines that indicate maximum and minimum values that can be theoretically obtained for the values displayed. For F_{GG} (the fraction of GG dyads), the upper broken line indicates the maximum value that can be obtained for any given F_G , since F_{GG} can never exceed the corresponding F_G . Similarly, the lower broken line shows the minimum value that F_{GG} can have for any

Figure 2 Anomeric region of the ¹H-NMR spectra (400 MHz) of an M-rich alginate after being epimerized by AlgE2, BL6 and AlgE4

G and M denote GulA and ManA residues, and the underlined letters and the numbers denote the residue and proton causing the signal.

given F_G . A line showing the theoretical F_{GG} to be expected for an enzyme, which at all stages attacks all M residues with the same probability (random attack mechanism), is included. As can be seen, there are data points covering most of the area between the theoretical minimum F_{GG} line and the data for BL6, whereas only very few data points are found in the region between those of BL6 and the theoretical maximum F_{GG} line. Thus most or all of the hybrid enzymes display properties similar to or intermediate between those of the parent enzymes with respect to their G-block-forming capabilities. It has been shown earlier that AlgE4 can introduce some G blocks in M-rich alginate, but at a very low rate [12]. This can also be seen in Figure 3(A) for BL5, which

Figure 3 Structure and epimerization patterns of the hybrid enzymes

(A) Fraction of GG (F_{GG}) versus degree of conversion (F_G) of M alginate epimerized by different epimerases. $\diamond$, KA1 (46). All the other hybrid enzymes are represented with their number (B) and $\times$, for even numbers; $\blacklozenge$, for odd numbers. The data obtained using clarified cell extracts are shown as $+$. —, a theoretical randomly epimerized alginate; - - - and - - - - -, the statistical maximum and minimum limits respectively, for F_{GG} for any enzyme. (B) Structure and analyses of the hybrid enzymes. Each enzyme is assigned both a name and a number for reference to (A). Regions originating from AlgE2 are coloured grey, whereas regions originating from AlgE4 are coloured white. Both members of the same hybrid pair are shown on the same line. The enzymes are divided into five groups based on the results in (A) as described in the text. The Figure is divided into three parts, I–III, according to the structure of the hybrids.

has an amino acid sequence identical with AlgE4. Consequently, enzymes generating products with F_{GG} values < 0.05 when F_G is > 0.30 should not be considered significantly different from AlgE4.

Correlations between epimerization patterns and structural compositions of the enzymes

The epimerization products of the different hybrids shown in Figure 3(A) seem to represent a continuum of properties between

those of the two parent enzymes. As a first step in the analyses of these data, the enzymes were divided into five groups: GG, GGm, R, MGg and MG enzymes (Figure 3B). GG enzymes are similar to AlgE2/BL6, whereas MG enzymes are similar to AlgE4/BL5. R enzymes are those for which the formation of GG dyads as a function of the G fraction develops very similarly to the predictions of a random attack model. GGm enzymes introduce more G blocks than expected from a random attack model, but seem to introduce less than AlgE2 at least at some degree of epimerization. The MGg enzymes are those that are capable of

introducing GG above a frequency of 0.05, but such that the data points are below the random attack line. All the AlgE4-type enzymes (MG enzymes) identified in the initial screen were reanalysed using excessive amounts of enzyme to distinguish them from the MGg enzymes.

By using this classification, some correlation between structure and epimerization patterns could be observed. Part I of Figure 3(B) shows how the properties of the enzymes were affected when the site of exchange was moved stepwise from the N-terminal end. A major change in the epimerization pattern is seen when the exchange includes region 6, indicating an important function for this region. This hypothesis is also consistent with part II of Figure 3(B), where for all enzymes the epimerization patterns are similar to the enzyme from which region 6 originates.

Part III of Figure 3(B) summarizes the results for the enzymes, where region 6, alone or together with regions 4, 5 or 8 (or combinations of them), is exchanged. Again, no enzyme-containing region 6 from AlgE4 generates more G blocks than that expected from a random epimerization pattern; most of them generate less. However, region 6 from AlgE2 alone (BL43) is not sufficient for the hybrid to generate more G blocks than expected for a random enzyme. The corresponding hybrid enzyme (BL42) still retains a significant ability to form G blocks. These results show that other regions also have an impact on the epimerization pattern. However, whereas exchange of region 6 causes a transition between mostly GG- and MG-forming enzymes, the exchange of some of the other regions causes more subtle modulations, for instance, between GG and GGm enzymes or between MG and MGg enzymes. Part III of Figure 3(B) clearly shows an effect of exchanging modules 4, 5, 8 or combinations of them. In the other parts of the Figure, an influence of these three regions, and also of regions 3 and 7, on the epimerization pattern are seen from the enzymes belonging to the GGm and MGg classes. In several cases, the effects seem to be dependent on the combination of several regions.

The enzyme KA1 was of particular interest since the spectra of its reaction products were similar to what one would expect if the enzyme were attacking its substrate according to a random attack mechanism. Since this could have implications for the understanding of how the enzymes act, we collected many more spectra for KA1 (Figure 3A). These additional experiments confirmed the initial hypothesis and also helped to assess the level of accuracy in the measurements shown in Figure 3(A). None of the naturally occurring epimerases studied previously, displayed properties consistent with a random attack mechanism [12,23]. Furthermore, for all *in vivo* produced alginates that have had their dyad frequencies determined, it can be shown that the G residues have been introduced by a non-random mechanism [3].

Improved method for the calculation of GG dyads in time-resolved NMR spectroscopy

The experiments described above can provide only limited information about the way a particular enzyme generates its characteristic epimerization pattern. The epimerization reactions may be run directly in the NMR tube, allowing continuous recording of spectra as the catalytic reactions progress [13]. To follow the epimerization reactions quantitatively, the spectra have to be integrated and the fraction of the different dyads calculated. A technical obstacle in these procedures arises from the fact that ^1H -NMR spectroscopy of alginates are run in deuterated water, and when this solvent is present during the catalytic reaction, epimerization leads to an exchange of H-5 on G residues with deuterium. The $\underline{\text{G}}\text{-5-G}$ signal, which is normally used to estimate the G-block formation, then becomes undetectable by ^1H -NMR.

Figure 4 Details of the ^1H -NMR spectra (400 MHz) of G-block alginate (top) and MG-block alginate showing that the $\underline{\text{G}}\text{-1-G}$ signal is shifted higher than the $\underline{\text{G}}\text{-1-M}$ signal

The fraction of GG sequences may be calculated indirectly from the difference $F_{\text{G}} - F_{\text{MG}}$, but, as described by Hartmann et al. [13], it is difficult to obtain an accurate estimate of F_{MG} from the $\underline{\text{M}}\text{-1-G}$ peak, because it is positioned on top of the $\underline{\text{M}}\text{-1}$ end signal. As shown in Figure 5, the anomeric signal from guluronic acids shifts depending on the degree of epimerization, suggesting that the G-1 signal is sensitive to neighbouring residues. To verify this hypothesis, NMR spectra were run on homopolymeric G blocks and on alternating blocks. The G blocks ($F_{\text{G}} = 0.92$) were prepared from *Laminaria hyperborea* outer cortex after fractionation [24,25], whereas the alternating blocks ($F_{\text{G}} = 0.47$ and $F_{\text{GG}} = 0$) were made by epimerization of mannuronan with AlgE4. The NMR spectra shown in Figure 4 demonstrate that $\underline{\text{G}}\text{-1-G}$ shift slightly up field of $\underline{\text{G}}\text{-1-M}$. The splitting of the G-1 signal can therefore be used to calculate the relative amounts of MG and GG.

Analyses of the catalytic reactions by time-resolved NMR spectroscopy

To achieve a better understanding of how the different epimerization patterns are generated, three of the hybrids (KA1, TB5 and TB14) were studied in more detail by time-resolved NMR spectroscopy. KA1 was chosen because of its apparently random mode of action (R group). TB5 and TB14 were both classified as MGg enzymes; however, from the data points available (Figure 3A), it appeared that TB5 generated G blocks somewhat more efficiently than TB14. BL6 was included as a reference (GG type).

Simple visual inspection of the spectra (Figure 5) immediately confirmed that the relative rate by which the GG signals ($\underline{\text{G}}\text{-1-G}$) develop compared with that of GM ($\underline{\text{G}}\text{-1-M}$) varied strongly between the enzymes. BL6 clearly introduced relatively more GG at early stages of the reaction than the three hybrids, which appeared mostly to generate alternating structures in the same period. An interesting feature is the diminishing of the $\underline{\text{M}}\text{-1-G}$ signal towards the end of the reaction for all of the hybrid enzymes. The only way to achieve this is to epimerize the M residue in the triplet -GMG-. This was not observed for BL6; since F_{MG} remained lower for this enzyme throughout the experiment, it cannot be excluded that BL6 is also capable of using -GMG- as a substrate. The $\underline{\text{G}}_{\text{red}}\text{-}\beta$ signal seen for BL6 in Figure 6 shows that this enzyme, similar to AlgE2 [18], has a weak lyase activity. This is not seen for any of the three hybrid enzymes, even though their final degrees of epimerization are approximately equal to that of BL6.

Figure 5 Stacked plot of the anomeric region of ^1H -NMR spectra (400 MHz) during epimerization of mannanuronan with the C-5 epimerases

The enzymic reaction was carried out in the spectrometer at 50 °C. Each spectrum shows an average of 8 min scans. The direction of the arrow indicates increasing reaction time. Since the reaction proceeds fastest in the beginning, the time between each spectrum shown here is less at the start of the reaction than towards the end to provide some detail. The notations are as described in the legend to Figure 2.

Integration of the spectra (Figure 6) confirmed the results obtained by the single-point experiments (Figure 3). BL6 preferably introduces G in GG dyads as shown by the steeper slope of F_{GG} as a function of F_{G} compared with that of F_{MG} . In contrast, the G blocks first accumulated at a very low rate relative to total G for TB5 and TB14, and then started to accumulate more rapidly at $F_{\text{G}} > 0.5$. These enzymes are therefore similar to AlgE4 in the sense that they preferentially avoid epimerizing a G next to another G; however, in contrast with AlgE4 [13], they have a significant capacity to do so. For KA1 the dyad frequencies, as a function of F_{G} , developed according to a random attack mode.

Kinetic analyses of the reactions

The multiple NMR spectra shown in Figure 5 were recorded at fixed time intervals, allowing also kinetic analyses of the

epimerization reactions. To do this, it had to be taken into account that the amount of epimerase per weight unit of freeze-dried total protein varied somewhat in the four preparations. Additionally, it has been shown previously that for AlgE4 the active form migrates as a 79 kDa protein in an SDS/polyacrylamide gel, while other forms (apparently inactive) are also present [12]. Similar observations were made in the present study, and the relative amount of the active form (79 kDa) was not constant. To compensate for these differences, the amounts of assumed active protein were estimated from a Coomassie Blue-stained SDS/polyacrylamide gel. The results could then be used to visualize rates for introducing G and the dyads GG and MG (Figure 7). In this Figure, the slopes of the curves are proportional to the specific activity of each enzyme. TB5 clearly has a much higher epimerization rate than the three other enzymes. For BL6, it is seen that the initial rate for introducing G blocks is lower than for TB5, indicating a

Figure 6 Fraction of GG and MG versus degree of conversion (F_G) by the BL6 ($\square$), KA1 ($\triangle$), TB5 ($\blacklozenge$) and TB14 ($\times$) epimerases

The thin line represents AlgE4, whereas the thick line represents a randomly epimerized alginate. The parameters were calculated by integration of the spectra in Figure 5.

need for preformed G residues to obtain the maximum rate. At higher F_G , the rate of forming GG dyads is similar for TB5 and BL6, and it is at all times significantly lower for KA1 and TB14. TB5 introduced MG dyads much faster than the other three, which display approximately the same rates for increasing F_{MG} . Note also that although TB5 and TB14 have similar epimerization patterns, TB5 has a significantly higher epimerization rate for both dyads. These results indicate that the epimerization pattern made by any enzyme is determined by the relative rates of introducing MG and GG displayed by that enzyme.

DISCUSSION

The results obtained in the present study showed that one can generate new active epimerases by randomly exchanging corresponding parts of the AlgE4 and AlgE2 A modules. Secondly, many of the new hybrid enzymes generated epimerization patterns different from those of the parent enzymes. Among these was an enzyme that appeared to epimerize according to a random attack model (KA1).

In recent studies regarding the reaction mechanisms of AlgE2 and AlgE4, it was found that AlgE4 probably epimerizes the polymer in a processive manner, whereas AlgE2 probably acts by a preferred attack mechanism where the affinity for the substrates depends on preexisting G residues [23,26,27]. From a structural point of view, the repeating unit in a segment of mannuronic acid residues can be considered as being composed of two residues (Figure 8). AlgE4, making alternating structures, can then be processive by moving two residues (one repeating unit) along the strand in each catalytic step. In contrast, if two consecutive

Figure 7 Fraction of G, GG and MG introduced by the BL6 ($\square$), KA1 ($\triangle$), TB5 ($\blacklozenge$) and TB14 ($\times$) epimerases as a function of time (min) and enzyme amounts (mg of active enzyme)

The parameters were calculated by integration of the spectra in Figure 5. The specific epimerization rates are proportional to the slopes of the graphs.

residues were to be epimerized (to make G blocks), a processive enzyme would need to rotate 180° relative to the polymer strand. It seems that AlgE2 may have achieved its ability to make G blocks by losing the processivity and developing a preference for a G residue next to the residue bound at the catalytic site. This is also consistent with the significantly lower reaction rate observed for G-block-forming epimerases relative to MG-block-forming enzymes [28].

Time-resolved NMR of the epimerization reactions of TB5, TB14, KA1 and BL6 showed that only TB5 is capable of generating MG dyads at a high rate, whereas the rate of introducing GG dyads is similar to that of AlgE2. TB5 is the only enzyme among the four that contains region 6 from AlgE4, whereas both TB5 and TB14 contain region 5 from AlgE4. The single-point experiments summarized in Figure 3 show that region 6 from AlgE4 alone is sufficient to alter the epimerization pattern of AlgE2. The effect of region 5 on the other hand seems to be less prominent, the difference between BL42 containing only region 6 from AlgE4 and TB5 is fairly weak. These results indicate that amino acids in region 6 could be important for the processivity of AlgE4,

Figure 8 Schematic drawing showing the minimum amount of discriminating subsites needed for substrate binding in (A) AlgE2 and (B) MG-forming enzyme, assuming a preferred attack mechanism

The preferred residue is shown in boldface; if the epimer is accepted, it is written in parentheses. The alginate molecule is drawn as a stretched polymer; the actual angles of the 1–4 bonds of the polymer when bound to the enzyme are unknown. Note that the smallest repeated unit in an alginate chain is a dimer.

although the possibility of a co-operation between regions 5 and 6 cannot be excluded.

Enzymes that act according to a processive or preferred attack mechanism must recognize more than one residue in the polymer substrate. The epimerization pattern of AlgE2 might thus be explained by assuming that the enzyme can recognize a minimum of four consecutive residues [23] as illustrated in Figure 8(A). The catalytic site (+1) must bind an M, whereas the epimerase will prefer a G residue at site –1. Site –2 preferably will recognize a G residue to elongate existing G blocks, but an M residue is also readily accepted. AlgE2 seems to prefer the substrate -G/MGMM- [23], thus site +2 preferably binds an M residue. For an MG-block-forming enzyme using a preferred attack mechanism, subsites –1, +1 and +2 each must bind an M residue, whereas –2 preferably would bind a G residue (Figure 8B). Considering the difference in structure between oligomers containing G and M (Figure 8), it is quite conceivable that a subsite may differentiate between them. Not only will the angles be different, but also more important is the fact that the bulky, negatively charged carboxy group will be in a different position.

The present study has demonstrated that region 6 is important for the epimerization pattern, and the sequence alignment in Figure 1 shows 15 differing amino acids in this region. Yet, the family of secreted AlgE epimerases comprises three MG- and six GG-forming A modules [7]. When all these sequences are compared, only two amino acid changes in region 6 are found to correlate with the epimerization pattern: amino acid 233 is a His in the GG formers (Phe in AlgE7) and a Leu or Gln in the MG formers. Similarly, amino acid 235 is a Tyr in the GG formers (Glu in AlgE6) and Ser or Thr in the MG formers. Region 8 also seemed to have an impact on the epimerization pattern; in this region all MG formers have a Phe as amino acid 307, whereas all GG formers have a Tyr residue. The other differences in amino

acids in this region do not seem to correlate with the epimerization pattern. In region 4, the sequence LVDS (amino acids 180–183) in the MG-forming enzymes, is changed to QI/VGG/A in the GG-forming ones. No amino acid changes correlating with the epimerization pattern are found in region 5.

What emerges from the present study is that the epimerization pattern created by any epimerase is determined by the relative rates of introducing GG and MG. The GG-forming rate is determined by the affinity for G at subsite –1, and by the ability to accept a G at subsite +2. A high MG-forming rate is dependent on the processivity of the enzyme and a specificity for M at subsite –1. The exact mechanisms by which the different patterns are generated are difficult to unravel in the absence of a detailed understanding of the molecular structure of the epimerases and their interactions with the substrate. Yet, the biochemical studies reported here identified some regions important for the subsite specificity or processivity of the enzymes. Since the epimerization patterns are of crucial importance both in the biotechnological uses of alginates and for their biological functions, these findings should stimulate many further studies.

This work was funded by a grant from the Norwegian Research Council, and by FMC Biopolymers AS. We thank Hilde Foros, Elisabeth Kalvik, Erik Brodin Eide and Kim Andreassen for constructing some of the plasmids used in this study and Aslak Einbu for providing 33 samples of KA1-epimerized alginate.

REFERENCES

- Smidsrød, O. and Draget, K. I. (1996) Chemistry and physical properties of alginates. *Carbohydr. Europe* **14**, 6–13.
- Valla, S., Li, J., Ertesvåg, H., Barbeyron, T. and Lindahl, U. (2001) Hexuronyl C5-epimerases in alginate and glycosaminoglycan biosynthesis. *Biochimie* **83**, 819–830.
- Skjåk-Bræk, G., Grasdalen, H. and Larsen, B. (1986) Monomer sequence and acetylation pattern in some bacterial alginates. *Carbohydr. Res.* **154**, 239–250.

- 4 Ertesvåg, H., Doseth, B., Larsen, B., Skjåk-Bræk, G. and Valla, S. (1994) Cloning and expression of an *Azotobacter vinelandii* mannuronan C-5-epimerase gene. *J. Bacteriol.* **176**, 2846–2853
- 5 Ertesvåg, H., Høidal, H. K., Hals, I. K., Rian, A., Doseth, B. and Valla, S. (1995) A family of modular type mannuronan C-5-epimerase genes controls alginate structure in *Azotobacter vinelandii*. *Mol. Microbiol.* **16**, 719–731
- 6 Svanem, B. I., Skjåk-Bræk, G., Ertesvåg, H. and Valla, S. (1999) Cloning and expression of three new *Azotobacter vinelandii* genes closely related to a previously described gene family encoding mannuronan C-5-epimerases. *J. Bacteriol.* **181**, 68–77
- 7 Ertesvåg, H., Høidal, H. K., Schjerven, H., Svanem, B. I. and Valla, S. (1999) Mannuronan C-5-epimerases and their application for *in vitro* and *in vivo* design of new alginates useful in biotechnology. *Metab. Eng.* **1**, 262–269
- 8 Sadoff, H. L. (1975) Encystment and germination in *Azotobacter vinelandii*. *Bacteriol. Rev.* **39**, 516–539
- 9 Campos, M., Martinez-Salazar, J. M., Lloret, L., Moreno, S., Nunez, C., Espin, G. and Soberon-Chavez, G. (1996) Characterization of the gene coding for GDP-mannose dehydrogenase (algD) from *Azotobacter vinelandii*. *J. Bacteriol.* **178**, 1793–1799
- 10 Høidal, H. K., Svanem, B. I. G., Gimmedstad, M. and Valla, S. (2000) Mannuronan C-5 epimerases and cellular differentiation of *Azotobacter vinelandii*. *Environ. Microbiol.* **2**, 27–38
- 11 Ertesvåg, H. and Valla, S. (1999) The A modules of the *Azotobacter vinelandii* mannuronan-C-5-epimerase AlgE1 are sufficient for both epimerization and binding of Ca^{2+} . *J. Bacteriol.* **181**, 3033–3038
- 12 Høidal, H. K., Ertesvåg, H., Skjåk-Bræk, G., Stokke, B. T. and Valla, S. (1999) The recombinant *Azotobacter vinelandii* mannuronan C-5-epimerase AlgE4 epimerizes alginate by a nonrandom attack mechanism. *J. Biol. Chem.* **274**, 12316–12322
- 13 Hartmann, M., Duun, A. S., Markussen, S., Grasdalen, H., Valla, S. and Skjåk-Bræk, G. (2002) Time-resolved ^1H and ^{13}C NMR spectroscopy for detailed analyses of the *Azotobacter vinelandii* mannuronan C-5 epimerase reaction. *Biochim. Biophys. Acta* **1570**, 104–112
- 14 Sambrook, J. and Russell, D. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn, Cold Spring Harbor Laboratory Press, Plainview, NY
- 15 Chung, C. T., Niemela, S. L. and Miller, R. H. (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2172–2175
- 16 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119
- 17 Amann, E., Ochs, B. and Abel, K. J. (1988) Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**, 301–315
- 18 Ramstad, M. V., Ellingsen, T. E., Josefsen, K. D., Høidal, H. K., Valla, S., Skjåk-Bræk, G. and Levine, D. (1999) Properties and action pattern of the recombinant mannuronan C-5-epimerase AlgE2. *Enzyme Microbial Technol.* **24**, 636–646
- 19 Ertesvåg, H. and Skjåk-Bræk, G. (1999) Modification of alginate using mannuronan C-5-epimerases. In *Methods in Biotechnology 10, Carbohydrate Biotechnology Protocols* (Bucke, C., ed.), pp. 71–78, Humana Press, Totowa, NJ
- 20 Grasdalen, H. (1983) High-field, ^1H -n.m.r. spectroscopy of alginate. Sequential structure and linkage-conformation. *Carbohydr. Res.* **118**, 255–260
- 21 Rehm, B. H., Ertesvåg, H. and Valla, S. (1996) 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.* **178**, 5884–5889
- 22 Gimmedstad, M., Sletta, H., Ertesvåg, H., Bakkevig, K., Jain, S., Suh, S.-J., Skjåk-Bræk, G., Ellingsen, T. E., Ohman, D. E. and Valla, S. (2003) The *Pseudomonas fluorescens* algG protein, but not its mannuronan C5-epimerase activity, is needed for alginate polymer formation. *J. Bacteriol.* **185**, 3515–3523
- 23 Hartmann, M., Holm, O. B., Johansen, G. A. B., Skjåk-Bræk, G. and Stokke, B. T. (2002) Mode of action of recombinant *Azotobacter vinelandii* mannuronan C-5 epimerases AlgE2 and AlgE4. *Biopolymers* **63**, 77–88
- 24 Haug, A. (1959) Fractionation of alginic acid. *Acta Chem. Scand.* **13**, 601–603
- 25 Haug, A., Larsen, B. and Smidsrød, O. (1967) Studies on the sequence of uronic acid residues in alginic acid. *Acta Chem. Scand.* **21**, 691–704
- 26 Campa, C., Holtan, S., Nilsen, N., Bjerkan, T. M., Stokke, B. T. and Skjåk-Bræk, G. (2004) Biochemical analysis of the processive mechanism for epimerisation of alginate by mannuronan C-5 epimerase AlgE4. *Biochem. J.*, DOI 10.1042/BJ20031265
- 27 Sletmoen, M., Skjåk-Bræk, G. and Stokke, B. T. (2004) Indications of a processive mode of action of a non-degrading polysaccharide modifying enzyme obtained by single-molecular pair unbinding studies. *Biomacromolecules*, in the press
- 28 Ertesvåg, H., Høidal, H. K., Skjåk-Bræk, G. and Valla, S. (1998) The *Azotobacter vinelandii* mannuronan C-5-epimerase AlgE1 consists of two separate catalytic domains. *J. Biol. Chem.* **273**, 30927–30932

Received 16 October 2003/14 April 2004; accepted 19 April 2004

Published as BJ Immediate Publication 19 April 2004, DOI 10.1042/BJ20031580