Improvement of the Thermostability and Activity of a Pectate Lyase by Single Amino Acid Substitutions, Using a Strategy Based on Melting-Temperature-Guided Sequence Alignment††

Zhizhuang Xiao, Hélène Bergeron, Stephan Grosse, Manon Beauchemin, Marie-Line Garron, David Shaya, Traian Sulea, Miroslaw Cygler, and Peter C. K. Lau*

Biotechnology Research Institute, National Research Council Canada, Montreal, Quebec H4P 2R2

Received 28 September 2007/Accepted 8 December 2007

In the vast number of random mutagenesis experiments that have targeted protein thermostability, single amino acid substitutions that increase the apparent melting temperature ($T_m$) of the enzyme more than 1 to 2°C are rare and often require the creation of a large library of mutated genes. Here we present a case where a single beneficial mutation (R236F) of a hemp fiber-processing pectate lyase of Xanthomonas campestris origin (PL$_{Xc}$) produced a 6°C increase in $T_m$ and a 23-fold increase in the half-life at 45°C without compromising the enzyme’s catalytic efficiency. This success was based on a variation of sequence alignment strategy where a mesophilic amino acid sequence is matched with the sequences of its thermophilic counterparts that have established $T_m$ values. Altogether, two-thirds of the nine targeted single amino acid substitutions were found to have effects either on the thermostability or on the catalytic activity of the enzyme, evidence of a high success rate of mutation without the creation of a large gene library and subsequent screening of clones. Combination of R236F with another beneficial mutation (A31G) resulted in at least a twofold increase in specific activity while preserving the improved $T_m$ value. To understand the structural basis for the increased thermal stability or activity, the variant R236F and A31G R236F proteins and wild-type PL$_{Xc}$ were purified and crystallized. By structure analysis and computational methods, hydrophobic desolvation was found to be the driving force for the increased stability with R236F.

Improving the thermostability of a protein or enzyme is desirable for commercially viable bioprocesses that prolong product shelf life, increase energy efficiency, and save costs (30). Unfortunately, there are no simple rules for improving thermostability, the mechanisms of which include the formation of disulfide bridges, hydrophobic or aromatic interactions, contact order, hydrogen bonding, ion pairing, dimer-dimer interaction, and a preponderance of glutamine usage (32, 36, 41). A critical and challenging task in thermostability engineering is, therefore, choosing a method whereby promising results can be obtained quickly and efficiently. Both rational (structure-based, including molecular modeling) and randomized or irrational (directed evolution or random and combinatorial engineering) approaches have been applied with varying degrees of success (11, 19, 30).

Notable single amino acid substitutions that have led to dramatic increases in melting temperatures ($T_m$; the temperature at which 50% of the protein is unfolded), by 20°C and 24°C, respectively, were identified in the Drosophila protein drk SH3 domain (24), and in a bacterial malate dehydrogenase (4). However, these successes were made possible only through a good understanding of the thermodynamics and kinetics of folding, or through knowledge of the electrostatic interactions in the dimer-dimer interface of the tetrameric protein. Otherwise, single amino acid substitutions that increase the apparent $T_m$ of a given enzyme more than 1 to 2°C are rare (19).

In this study, we selected a bacterial-genome-mined pectate lyase sequence as a template in order to evolve added thermostability and/or catalytic activity using a newly improvised strategy that does not require the construction of a large library of clones, as in directed evolution experiments, or specialized computational skills, as in rational protein engineering approaches. We further characterized the beneficial variants at the biochemical and structural levels. Pectate lyases (EC 4.2.2.2) are secreted enzymes produced by a variety of plant-pathogenic bacteria and therefore are best studied as virulence factors (for a review, see reference 15). Biotechnologically, these pectinolytic enzymes, as well as various endopolygalacturonases, are useful retting or bioscouring reagents for the processing of natural bast fibers (e.g., hemp and flax), and cotton fabric (1, 2, 28, 35).

MATERIALS AND METHODS

Molecular cloning and site-directed mutagenesis. Standard methods were used for isolation of plasmid DNA, cloning, and transformation (33). The Xan-thomonas campestris pectate lyase II (PL$_{Xc}$) gene (NP_638163) was amplified from its genomic DNA (ATCC 33913) and cloned into pDS80, an isopropyl-$\beta$-thiogalactopyranoside (IPTG)-inducible expression vector, at the EcoRI and HindIII restriction sites (34). This cloning is part of a larger genome-mining effort for pectinase-encoding sequences (Z. Xiao, J. Boyd, S. Grosse, E. Coupe, and P. C. K. Lau, unpublished data). The recombinant plasmid (pDS80PLII) and resultant derivatives were transformed into Escherichia coli strain Rosetta 2 (Novagen). Site-directed mutagenesis was performed using the QuikChange mutagenesis kit (Stratagene, La Jolla, CA) according to its instruction manual. The mutagenic oligonucleotide primers are listed in Table S1 in the supplemental material. The expected mutations were confirmed by gene sequencing using
MgSO₄ (20 g/liter) was started. Culture was induced a few minutes later by a batch phase of about 9 h, a feed with 0.22 ml/min of glucose (50% solution) and Thermotoga maritima TS-47 (37), and AAD35518 from BAA96478 from their established thermostability in reported literature. The thermostable candidate selected by screening of more than a thousand candidates in the NCBI database for sequence (9) with the four available thermostable pectate lyase sequences, sequencer (model 377; ABI Prism). A BigDye DNA sequencing kit (Applied Biosystems) with an automated DNA X. campestris /clustalw/) was used to align the cells were harvested subsequently. Cells were centrifuged (at 6,000 g/ml ampicillin and 30 g/ml chloramphenicol) and grown overnight at 30°C in LB medium. Glucose (0.7%) was used as the preculture (giving a starting optical density at 600 nm of about 0.3), which had been grown overnight at 30°C on LB medium. The pH was maintained automatically using NH₄OH (29%; sole carbon source. The pH was maintained automatically using NH₄OH (29%)

FIG. 1. Multiple alignment of protein sequences of thermostable pectate lyases and thermolabile PLXc, in family PL1. The signal peptide sequences of the respective proteins are not shown. The numbering of each sequence starts from the initiation codon. Conserved catalytic sites (K199, R235, R236), conserved calcium binding sites (D137, D175, D179), the core structure of the parallel β-helices (vWiDH region), and sites conserved in all thermostable PL1 pectate lyases (CAD65882 from Bacillus licheniformis, BAA96478 from Bacillus sp. strain P-4-N, BABA036 from Bacillus sp. strain TS-47, AAD35518 from Thermotoga maritima MSB8) but variant in PLXc (NP_638163) are shaded.


Protein expression and fermentation. Single colonies of E. coli Rosetta 2 harboring the X. campestris pectate lyase II gene and the desired mutations (V26A, A31G, L64I, Y66V, K69T, F70I, Q123R, V187I, R236F, and the A31G substitution in the parallel β-helices (vWiDH region)) but variants in PLXc but variant in PLXc (NP_638163) are shaded.
points, followed by the pectate lyase assay. Purified proteins were adjusted to about 0.5 mg/ml in 10 mM phosphate buffer, pH 7.0, for determining the circular dichroism (CD) spectrum and \( T_m \). Thermal denaturation experiments were performed at a temperature increase of 4°C per h. The data were collected with a JASCO J-710 CD spectrometer at a wavelength of 222 nm, where maximal signal difference was observed.

Protein crystallization and structure determination. Initial crystallization conditions were identified using the PEGs (Nextal, Montreal, Canada), Classic Suite I and II (Qiagen), screens, with the sitting drop vapor diffusion set up at 20°C. Numerous chemical conditions yielded triclinic crystals overnight. The conditions were reproduced and optimized using the hanging drop vapor diffusion method in 24-well Linbro plates (Hampton Research). The best conditions found were 20 to 30% (wt/vol) polyethylene glycol 3500 and various buffers with pHs ranging from 5.5 to 8.5. The crystals belong to the primitive triclinic space group \( P1 \), with unit cell dimensions of 47.2 Å for \( a \), 53.2 Å for \( b \), 73.0 Å for \( c \), 71° for \( \alpha \), 68° for \( \beta \), and 69° for \( \gamma \), and two protein molecules in the asymmetric unit. The Matthews coefficient \((23) V_m = 2.20 \text{ Å}^3\text{Da}^{-1}\) corresponds to a solvent content of 44%. The cryoprotectant solution used consisted of mother liquor supplemented with 10 to 12% (wt/vol) glycerol, and data were collected at 100 K. Diffraction data extended to 2.0, 2.1, and 1.9 Å resolution for wild-type, R236F, and A31G R236F proteins, respectively. The images were processed with the HKL2000 program package (27). The structure of wild-type PL\(_{Xc}\) was solved by molecular replacement using the MolRep program (38) and the structure of \( X.\) campestris pectate lyase II and its variants on an SDS-PAGE gel (left) and a ruthenium red-stained zymogram gel containing 0.1% PGA (right). Lanes: M, molecular standard; 1, parent enzyme; 2, A31G variant; 3, R236F variant; 4, A31G R236F variant.

Results and Discussion

Sequence characteristics of a new pectate lyase in comparison to its thermophilic counterparts. The predicted pectate lyase NP_638163 entry of Xanthomomas campestris pv. campes- tris ATCC 33913 in the CAZY database (PL\(_{Xc}\)) is a biochemically uncharacterized entity except for its annotated sequence classification as a member of the polysaccharide lyase family 1 of proteins (http://www.cazy.org/fam/PL1.html). We established the \( T_m \) of PL\(_{Xc}\) to be 48°C, certifying it as a mesophilic protein (see below). Figure 1 shows a multiple alignment of the predicted amino acid sequence of PL\(_{Xc}\) with four available sequences of thermostable PLs, whose optimal temperatures have been established as ranging from 65 to 90°C. Interestingly, only nine amino acid positions were invariant in all four thermostable PLs but not in PL\(_{Xc}\). Pairwise comparison of the subject and database PL\(_{Xc}\) sequences ranging from 212 to 301 amino acids gave only 25 to 36% identity, indicating low sequence homology among this group of proteins.

![Image](http://aem.asm.org/Downloaded from http://aem.asm.org)
Characterization of PL<sub>Xc</sub> and its variants. We set out to test whether any of the nine variable positions could contribute to added thermostability and/or catalytic activity in PL<sub>Xc</sub> when they were replaced by the corresponding residue of a thermophilic counterpart (Table 1). A PGA assay was used to determine the activity and thermal inactivation of PL<sub>Xc</sub> and its variants in crude cell extracts. One mutation, R236F, increased the T<sub>m</sub> by 6°C and extended the half-life at 45°C 23-fold. Three other mutations, F70I, Q123R, and V187I, reduced thermostabilities by 18 to 53% (Fig. 2). Thus, four out of the nine positions appeared to be crucial for the thermostability of PL<sub>Xc</sub>. Of the other mutations, the Y66V substitution resulted in a loss of activity, while the A31G variant was fivefold more active than its parent enzyme, although its thermostability was not substantially changed. Combining the best attributes of A31G and R236F in a double variant, the A31G R236F variant, resulted in a 5-fold improvement in activity and a 12-fold increase in thermostability over those of the wild-type enzyme (Table 1). All in all, two-thirds of the amino acid substitutions (A31G, Y66V, F70I, Q123R, V187I, and R236F) substantially affected either the thermostability or the catalytic activity of the PL<sub>Xc</sub> enzyme, indicating a high success rate of identifying susceptible residues without having to generate a large library.

Enzyme purification and kinetics. The wild-type enzyme and the A31G, R236F, and A31G R236F variants were characterized further by purification to apparent homogeneity with a single cation-exchange SP-Sepharose column yielding 40% recovery of the respective proteins. These proteins appeared as a single band on an SDS-PAGE gel with an apparent molecular mass of 35 kDa, in good agreement with the theoretical mass of 35,239 Da. Zymogram analyses of the purified enzymes indicated that they are active pectate lyases (Fig. 3). The intensities of the clear zones of the respective enzymes were consistent with the measured specific activities. The K<sub>m</sub> values of the purified A31G, R236F, and A31G R236F proteins with PGA as a substrate were determined and found not to differ substantially from each other and from that of the parent enzyme (Table 2). However, the catalytic efficiencies for both the A31G and A31G R236F variants were about twofold greater than that for the parent enzyme. The activity of the most thermostable variant, the R236F variant, was the same as that of the parent. Overall, the specific activities of the purified enzymes are consistent with their k<sub>cat</sub>/K<sub>m</sub> values. Similar results

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>K&lt;sub&gt;m&lt;/sub&gt; (g liter&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>k&lt;sub&gt;cat&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>k&lt;sub&gt;cat&lt;/sub&gt;/K&lt;sub&gt;m&lt;/sub&gt; (liter g&lt;sup&gt;-1&lt;/sup&gt; s&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Sp act (U/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
<td>0.98 ± 0.20</td>
<td>114 ± 11</td>
<td>116 ± 17</td>
<td>196 ± 20</td>
</tr>
<tr>
<td>A31G</td>
<td>0.73 ± 0.11</td>
<td>194 ± 29</td>
<td>266 ± 40</td>
<td>333 ± 53</td>
</tr>
<tr>
<td>R236F</td>
<td>0.83 ± 0.15</td>
<td>121 ± 7</td>
<td>146 ± 16</td>
<td>208 ± 27</td>
</tr>
<tr>
<td>A31G R236F</td>
<td>0.93 ± 0.19</td>
<td>216 ± 12</td>
<td>232 ± 19</td>
<td>370 ± 22</td>
</tr>
</tbody>
</table>

FIG. 4. Amounts of pectin released from natural hemp fiber by PL<sub>Xc</sub> and its variants.

FIG. 5. T<sub>m</sub>s of the wild-type and variant enzymes determined by CD. (A) Original CD data; (B) derivative of CD signal.
were observed when citrus pectin was used as a substrate (Table 3). Additionally, these variants were found to be 20 to 40% more effective than the parent enzyme in releasing pectin from natural hemp fiber (Fig. 4).

**CD study.** The structures of the wild-type enzyme and the A31G, R236F, and A31G R236F variant enzymes showed no discernible differences in their α-helical structure contents by CD spectroscopy (see Fig. S1 in the supplemental material). However, the $T_m$s of the four proteins were established to be 48, 47.5, 54, and 53°C, respectively (Fig. 5). The 5 to 6°C increase in the $T_m$s of the R236F variant and the double variant are consistent with the increases in the half-lives of the proteins.

**Structures of wild-type PL$_{Xc}$ and variant proteins.** To understand the structural basis for the thermal stability of the variant proteins, R236F, A31G R236F, and wild-type PL$_{Xc}$ were crystallized and their structures solved by molecular replacements. These structures have the same right-handed parallel β-helix architecture as that described for Erwinia chrysanthemi (presently classified as Dickeya dadantii) pectate lyase PelC (43) (Fig. 6A; see also Fig. S2, the supplementary note, and Table S3 in the supplemental material). Residue R236 is partially solvent exposed and is located in the T1.6 loop (the 6th loop of turn T1), which stacks against the T1.5 loop, containing a helical structure previously shown to enhance the catalytic activity of related pectate lyases (10). Its guanidinium
The double mutation A31G R236F also does not introduce major variations relative to PLXG wild-type and R236F variant structures (Fig. 6B). A noticeable change is a peptide flip at the A31G mutation site, bringing the backbone carbonyl of A30 closer to the side chain of K151, which undergoes a concerted conformational change (Fig. 6C). The A31G mutation resides in the N-terminal region of the enzyme on the opposite side of the β-helix relative to the catalytic site. Thus, the twofold increase in catalytic efficiency upon the A31G mutation cannot be accounted for by the present structural data. Mutations distant from catalytic sites are documented to be as effective as close mutations in improving enzymatic activity (25), but the underlying factors are often subtle and elude structural interpretations. Computational predictions yielded a marginal destabilizing effect (0.2 kcal/mol) for the A31G mutation, underscoring the gain afforded by hydrophobic desolvation (−2.8 kcal/mol) as driving the increase in stability. This more than offsets the major destabilizing contribution, i.e., the loss of hydrogen bonding energy upon mutation (1.4 kcal/mol), between the carbonyl group of D207 and the guanidinium group of R236.

ACKNOWLEDGMENTS

This research was largely supported by the Climate Change Technology and Innovation Biotechnology Program of the Canadian Biomas Innovation Network of Natural Resources Canada.

We thank A. Matte for critical reading of the manuscript. This article is issued as NRCC publication 49095.

REFERENCES

highly thermophilic and alkaline pectate lyase Pel-4R from alkalophilic Bacillus sp. strain P-4N and the entire nucleotide and amino acid sequences. Extremophiles 5:127–133.


