Cell-associated α-amylases of butyrate-producing Firmicute bacteria from the human colon

Alan G. Ramsay,† Karen P. Scott, Jenny C. Martin, Marco T. Rincon and Harry J. Flint

Rowett Research Institute, Greenburn Road, Bucksburn, Aberdeen AB21 9SB, UK

Selected butyrate-producing bacteria from the human colon that are related to Roseburia spp. and Butyrivibrio fibrisolvens showed a good ability to utilize a variety of starches for growth when compared with the Gram-negative amylolytic anaerobe Bacteroides thetaiotaomicron. A major cell-associated amylase of high molecular mass (140–210 kDa) was detected in each strain by SDS-PAGE zymogram analysis, and genes corresponding to these enzymes were analysed for two representative strains. Amy13B from But. fibrisolvens 16/4 is a multi-domain enzyme of 144-6 kDa that includes a family 13 glycoside hydrolase domain, and duplicated family 26 carbohydrate-binding modules. Amy13A (182-4 kDa), from Roseburia inulinivorans A2-194, also includes a family 13 domain, which is preceded by two repeat units of ~116 aa rich in aromatic residues, an isoamylase N-terminal domain, a pullulanase-associated domain, and an additional unidentified domain. Both Amy13A and Amy13B have N-terminal signal peptides and C-terminal cell-wall sorting signals, including a modified LPXTG motif similar to that involved in interactions with the cell surface in other Gram-positive bacteria, a hydrophobic transmembrane segment, and a basic C terminus. The overexpressed family 13 domains showed an absolute requirement for Mg\(^{2+}\) or Ca\(^{2+}\) for activity, and functioned as 1,4-α-glucanohydrolases (α-amylases; EC 3.2.1.1). These major starch-degrading enzymes thus appear to be anchored to the cell wall in this important group of human gut bacteria.

INTRODUCTION

Starch that escapes digestion in the small intestine goes on to provide a major source of energy for microbial growth in the large intestine (van Munster et al., 1994). Dietary starch escapes breakdown by small intestinal amylases via several mechanisms, including protection by cell wall polymers that are not degradable by host enzymes, high amylase content, and the state of gelatinization or retrogradation, which is largely determined by cooking and processing (Englyst et al., 1992). Starch is a complex polysaccharide consisting of a mixture of amylase (1,4-α-linked glucose residues) and amylopectin, a branched polymer composed of amylase chains linked to an amylase backbone by 1,6-α linkages. The relative proportions of amylase and amylopectin have a considerable effect on the availability of different types of starch for bacterial growth. A range of catalytic specificities is involved in starch breakdown, including α-amylases that hydrolyse 1,4-α linkages, type I pullulanases that specifically cleave 1,6-α bonds, and amylopullulanases that possess both 1,4-α and 1,6-α activities (Erra-Pujada et al., 1999). Enzymes involved in starch degradation belong to families 13, 14 and 57 (http://afmb.cnrs-mrs.fr/CAZY/), with by far the greatest number of starch-degrading enzymes, including α-amylases, pullulanases and amylopullulanases, falling into family 13 glycoside hydrolases (MacGregor et al., 2001).

The addition of dietary starch has been shown to affect the composition of the gut microbiota in many studies (e.g. Macfarlane & Englyst, 1986; Silvi et al., 1999), yet we know little about which bacterial groups are most successful in competing for starch as an energy source in vivo, or about the variety of mechanisms involved in starch utilization. Indeed, the only human colonic bacterium for which the organization of the enzyme systems responsible for starch utilization has been established is Bacteroides thetaiotaomicron (D’Elia & Salyers, 1996; Reeves et al., 1997); it appears that an elaborate system of starch-binding proteins and periplasmic hydrolases enables this bacterium to sequester and degrade starch molecules (Reeves et al., 1997), presumably allowing it to compete more effectively for the available substrate. Little is known about the ability of
human colonic bacteria possessing a Gram-positive cell-wall ultrastructure to compete for, and process, polysaccharides such as starch. This study attempts to elucidate the organization of \( \alpha \)-amylases in amylolytic Gram-positive anaerobes belonging to clostridial cluster XIVa, isolated from human faeces. Evidence from 16S rRNA diversity studies indicates that cluster XIVa is numerically one of the most abundant bacterial groups in the human colon (Franks et al., 1998; Suau et al., 1999; Hold et al., 2002; Eckburg et al., 2005).

Bacteria that represent the most abundant groups of butyrate producers have been isolated previously from human faeces (Barcenilla et al., 2000). Many of these isolates, particularly clostridial cluster XIVa bacteria related to \textit{Roseburia intestinalis}, \textit{Roseburia inulinivorans}, \textit{Eubacterium rectale} and \textit{Butyrivibrio fibrisolvens}, are able to degrade starch (Duncan et al., 2002, 2003, 2006). There is evidence that starch fermentation by colonic bacteria favours the production of butyric acid (Le Blay et al., 1999; Jenkins et al., 1998), and butyrate is believed to help protect against colitis and colorectal cancer (McIntyre et al., 1993; Archer et al., 1998; Wachtershauser & Stein, 2000). The mildly acidic pH resulting from the production of short-chain fatty acid fermentation products improves \( \text{Ca}^{2+} \) reabsorption from the colon (Abrams et al., 2005), while also selecting for the \textit{Roseburial/E. rectale} group (Walker et al. 2005). Thus, it was of particular interest to understand the ability of these bacteria to compete for starch as an energy source, and to investigate the mechanisms used for starch degradation.

### METHODS

#### Bacterial strains and growth conditions.

Strains (listed in Table 1) were chosen as representatives of the low DNA mol\% G+C butyrate-producing Gram-positive bacteria in cluster XIVa (Rumney et al., 1995; Barcenilla et al., 2000; Duncan et al., 2002). \textit{Bacteroides thetaiotaomicron} 5482, a gift from A. Salyers (University of Illinois, Urbana-Champaign, USA), was included for comparison.

Routine culturing of bacteria was in M2GSC medium (Miyazaki et al., 1997). BYCFA medium (Duncan et al., 2003), supplemented with 0-2 \% w/v (unless otherwise stated) of a range of autoclaved 10 \% starch polysaccharide preparations with varying contents of amylose and amylopectin (Table 2), was used for specific growth experiments. The disaccharide maltose was included in the study as a comparison for bacterial growth on the starch substrates. Growth rate experiments were carried out in triplicate, and bacterial growth was determined spectrophotometrically by monitoring changes in \( \text{OD}_{600} \) every hour until stationary phase was reached. For slow-growing cultures, readings were taken up to 72 h after initial inoculation. The turbidity of the basal media containing the starch solutions varied depending on the specific starch substrate; therefore, each set of growth data was blanked against a tube containing that specific substrate. The extent of growth of selected amylolytic strains was confirmed by measuring total bacterial protein concentration using the Lowry method. The protein present in 1 ml culture was measured at \( t_0 \) (background), and subtracted from the value at \( t_{24} \) (after 24 h growth), to get a corrected figure for the final protein concentration.

#### Enzyme activity.

Bacterial strains were grown to exponential or stationary phase in BYCFA medium containing 0-5 \% glucose, maltose or amylpectin corn starch (Table 2), and cell pellets and supernatants were separated from 1 ml culture by centrifugation (14 000 \( \times \) g for 10 min at 4 \( ^\circ \)C). Pellets were washed and resuspended in 50 \( \mu \)l 50 mM sodium phosphate buffer (pH 6-5), and pellets and supernatants were stored frozen, and thawed once before assaying. The enzyme activity was determined by measuring the release of reducing sugars (Lever, 1977), as described previously (Flint et al., 1991). Cell pellet and supernatant culture fractions were incubated aerobically for up to 2 h at 37 \( ^\circ \)C with 1 \% amylpectin corn starch substrate. Three independent experiments were carried out in triplicate, and the protein concentration in each sample was measured by using the method of Lowry. One unit of enzyme activity is equivalent to the release of 1 nmol glucose \( \text{min}^{-1} \) (mg protein \( \text{g}^{-1} \)).

#### Zymogram analysis.

Proteins exhibiting amylase activities were identified using a modification of the method of Saul et al. (1989). Fresh total cell extracts were prepared from 1 ml of cell culture during exponential growth. Cells were pelleted by centrifugation (14 000 \( \times \) g for 10 min at 4 \( ^\circ \)C), and the cell pellets were washed and resuspended in 50 \( \mu \)l 50 mM sodium phosphate buffer (pH 6-5). Extracts were incubated with 5 \( \times \) SDS loading buffer (Sambrook et al., 1989) at 60 \( ^\circ \)C for 20 min, and then loaded on a 7-5 \% SDS-PAGE gel containing 0-2 \% amylpectin corn starch. Following electrophoretic separation, the gel was washed (2 \( \times \)25 min) in 200 ml solution 1 [10 mM Tris/HCl, pH 7-5, 5 mM \( \beta \)-mercaptoethanol, 20 \% (v/v) 2-propanol], and enzymes were renatured overnight at 4 \( ^\circ \)C in 200 ml solution 2 (50 mM Tris/HCl, pH 6-8, 1 mM EDTA, 5 mM \( \beta \)-mercaptoethanol), with gentle shaking. Finally, the gel was incubated for 1 h at 4 \( ^\circ \)C in 200 ml 50 mM sodium phosphate buffer (pH 6-5), and transferred to a glass plate, wrapped in cling film, and incubated at 37 \( ^\circ \)C for 4 h. Protein bands were visualized by staining with Coomassie blue, and starch hydrolysis was detected by the formation of clear zones following staining with Gram’s iodine solution.

### Table 1. Bacterial strains used in this study

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>16S rRNA sequence accession no.</th>
<th>Reference/source</th>
<th>Catalogue no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butyrivibrio fibrisolvens</td>
<td>16/4</td>
<td>AJ250365</td>
<td>Rumney et al. (1995)</td>
<td>RRI collection*</td>
</tr>
<tr>
<td>R. inulinivorans</td>
<td>A2-194</td>
<td>AJ270473</td>
<td>Duncan et al. (2006)</td>
<td>NCIMB 14030/DSM 16841</td>
</tr>
<tr>
<td>R. intestinalis</td>
<td>L1-952</td>
<td>AJ270479</td>
<td>Duncan et al. (2002)</td>
<td>RRI collection*</td>
</tr>
<tr>
<td>R. intestinalis</td>
<td>L1-82</td>
<td>AJ312385</td>
<td>Duncan et al. (2002)</td>
<td>NCIMB 13810/DSM 14610</td>
</tr>
<tr>
<td>Bacteroides thetaiotaomicron</td>
<td>5482</td>
<td>L16489</td>
<td>Gift from A. Salyers</td>
<td>ATCC 29148/DSM 2079</td>
</tr>
</tbody>
</table>

*Strains held in a collection at the Rowett Research Institute.
Table 2. Maximum $\text{OD}_{600}$ and $\mu_{\text{max}}$ values for cultures on BYCFA medium containing different starch substrates

Values are the means values of triplicate readings (± SD). All reagents were obtained from Sigma Aldrich, except maize starch 30261 (from VWR). The amylpectin content (%) of each starch is given; the remainder is amyllose. HA, High amyllose; ND, not determined (growth was too poor to calculate $\mu_{\text{max}}$).

<table>
<thead>
<tr>
<th>Strain</th>
<th>Measure-</th>
<th>Added starch substrate and catalogue no. (amylopectin content)</th>
<th>No addition</th>
<th>Malrose</th>
<th>A potato A8515 (100%)</th>
<th>A corn A7870 (100%)</th>
<th>Maize 30261 (74%)</th>
<th>Wheat 55127 (75%)</th>
<th>Rice S7260 (75–87%)</th>
<th>HA corn S4180 (30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butyribivibrio</td>
<td>$\mu_{\text{max}}$</td>
<td></td>
<td>0.12 ± 0.01</td>
<td>0.97 ± 0.02</td>
<td>0.99 ± 0.06</td>
<td>0.86 ± 0.02</td>
<td>0.68 ± 0.08</td>
<td>0.49 ± 0.04</td>
<td>0.69 ± 0.02</td>
<td>0.12 ± 0.02</td>
</tr>
<tr>
<td>fibrisolvens</td>
<td></td>
<td></td>
<td>16/4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roseburia</td>
<td>$\mu_{\text{max}}$</td>
<td></td>
<td>0.10 ± 0.05</td>
<td>1.22 ± 0.04</td>
<td>1.25 ± 0.01</td>
<td>1.07 ± 0.01</td>
<td>0.80 ± 0.00</td>
<td>0.63 ± 0.01</td>
<td>0.80 ± 0.03</td>
<td>0.06 ± 0.01</td>
</tr>
<tr>
<td>inulinivorans</td>
<td>A2-194</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roseburia</td>
<td>$\mu_{\text{max}}$</td>
<td></td>
<td>0.04 ± 0.00</td>
<td>1.23 ± 0.06</td>
<td>1.28 ± 0.03</td>
<td>1.17 ± 0.03</td>
<td>0.65 ± 0.01</td>
<td>0.58 ± 0.01</td>
<td>0.61 ± 0.01</td>
<td>0.06 ± 0.00</td>
</tr>
<tr>
<td>intestinalis</td>
<td>L1-952</td>
<td></td>
<td>0.12 ± 0.01</td>
<td>1.24 ± 0.08</td>
<td>1.31 ± 0.02</td>
<td>1.12 ± 0.01</td>
<td>0.60 ± 0.03</td>
<td>0.80 ± 0.01</td>
<td>0.80 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>Roseburia</td>
<td>$\mu_{\text{max}}$</td>
<td></td>
<td>0.08 ± 0.01</td>
<td>1.60 ± 0.06</td>
<td>0.92 ± 0.13</td>
<td>0.63 ± 0.00</td>
<td>0.65 ± 0.05</td>
<td>0.38 ± 0.03</td>
<td>0.87 ± 0.00</td>
<td>ND</td>
</tr>
<tr>
<td>intestinalis</td>
<td>L1-82</td>
<td></td>
<td>0.41 ± 0.01</td>
<td>0.89 ± 0.06</td>
<td>0.84 ± 0.02</td>
<td>0.92 ± 0.00</td>
<td>0.55 ± 0.04</td>
<td>0.42 ± 0.05</td>
<td>0.59 ± 0.04</td>
<td>0.45 ± 0.02</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>$\mu_{\text{max}}$</td>
<td></td>
<td>0.19 ± 0.02</td>
<td>0.31 ± 0.03</td>
<td>0.87 ± 0.01</td>
<td>0.40 ± 0.05</td>
<td>0.27 ± 0.04</td>
<td>0.23 ± 0.00</td>
<td>0.26 ± 0.01</td>
<td>0.19 ± 0.02</td>
</tr>
<tr>
<td>thetaliotammonicron</td>
<td></td>
<td></td>
<td>5482</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PCR amplification. All PCR amplifications were carried out using standard conditions (initial cycle 5 min at 94 °C, 30 cycles with a denaturation step of 1 min at 94 °C, a 1 min annealing step at the appropriate temperature, and an extension step of 3 min at 72 °C; followed by a final cycle of 10 min at 72 °C) on a DNA thermal cycler 480 (Perkin Elmer). Degenerate PCR primers (synthesized by MWG Biotech) were designed using CLUSTALW (Thompson et al. 1997) alignments of published $\alpha$-amyrase sequences from Lactobacillus amylovorus (AAAC5781; Giraud & Cuny, 1997), Streptococcus bovis 148 (BA24177; Satoh et al., 2000), Bacillus subtilis SUH4-2 (AAF14358; Cho et al., 2000) and B. fibrisolvens H17c (AAA32305; Rumbak et al., 1991) to target conserved regions in the catalytic domain of glycoside hydrolase family 13 $\alpha$-amylases, including conserved domain 4 (Rumbak et al., 1991). The resulting primers [AmyPfor, GA(T/C)GCGGIGTA(T/C)AC; and AmyPrev, TC(A/G)TGGI(G/C)(A/T)(T/C)TCIACCCAIAGT] were used to amplify a 750 bp fragment from B. fibrisolvens 16/4 chromosomal DNA, using an annealing temperature of 50 °C.

Completion of amylase gene sequences. The amylase coding sequences were completed by PCR walking. B. fibrisolvens 16/4 chromosomal DNA was digested using one of several restriction enzymes (EcoRV, EcoRI, HindIII, Pst or ClaI), and ligated with pUC18 cut with a compatible restriction enzyme. Direct amplification of the ligation mix with the M13 forward or reverse primer (recognizing the vector sequence), and a specific primer based on a known sequence (recognizing the Bf-amy13B sequence), yielded DNA bands that were purified from agarose gels (Qiagen gel extraction kit) by following the manufacturer’s instructions, and sequenced. Bands caused by single primer amplification were eliminated by comparison with control reactions using single primers. Repeated use of this PCR amplification and sequencing approach allowed the construction of a contiguous DNA sequence. The Clontech Universal GenomeWalker kit was used, according to the manufacturer’s instructions, to complete the sequence of the R. inulinivorans $\alpha$-amyrase. Chromosomal DNA was digested using blunt-cutting restriction enzymes (PvuII, EcoRV, Stul or DraI), adaptor molecules were ligated, and PCR amplification done using a gene-specific primer (~26mer), which was designed based on previous sequence, and adaptor-specific primers provided in the kit.

DNA sequence analysis. Samples were sequenced using a Taq ABI PRISM dye terminator cycle sequencing ready reaction kit (Perkin-Elmer), and separated on an ABI 377 automated sequencer. Prior to sequencing, PCR products were purified using a Wizard purification column (Promega). Sequences were assembled into a contiguous sequence using either the Contig Assembly Program or the GCG package (Devereux et al., 1984) available previously through the Human Genome Mapping Project (HGMP). Database homology searching utilized the BLAST search tool (www.ncbi.nlm.nih.gov/blast/). Multiple sequence alignments were performed using the CLUSTALW program (www.ebi.ac.uk/clustalw/). Sequences used for computer-assisted analysis and comparison were retrieved from ProDom (http://protein.toulouse.inra.fr/prodom/current/html/home.php), Pfam (www.sanger.ac.uk/Software/Pfam/), and the carbohydrate-active enzyme server (http://afmb.cnrs-mrs.fr/CAZY/).

Expression cloning and enzyme assays. The coding region for the family 13 catalytic domains of both Bf-Amy13B and Ri-Amy13A were cloned in the pET30 EK/LIC expression vector (Novagen). The DNA sequences encoding the catalytic domains of each enzyme were amplified from chromosomal DNA using gene-specific primers designed to include 14–15 bp of sequence specific for the LIC cloning site of the vector. After amplification, PCR products (50 μl) were incubated at 37°C for 1 h with 5 units T4 DNA polymerase (Promega), in a total volume of 60 μl, which contained 2·5 mM dATP, 1 mM DTT, 33 mM Tris/acetate (pH 7·9), 66 mM potassium acetate, 10 mM magnesium acetate, and 0·1 mg BSA ml⁻¹. After heat-inactivating the enzyme at 75°C for 15 min, the resulting blunt-ended products were gel purified using the QiAquick gel extraction kit (Qiagen), according to the manufacturer’s instructions. The
eluted DNA was concentrated to a volume of 5–10 µl, and the full volume was ligated into 1 µl pET-30 Ek/LIC expression vector, at 16 °C overnight. The resulting recombinant plasmid was transformed into *Escherichia coli* SoloPack Gold XL10 chemically competent cells (Invitrogen), plated on LB agar supplemented with 30 µg kanamycin ml⁻¹, and incubated at 37 °C for 16 h. Positive clones were verified by PCR and sequencing across the cloning junction.

Double His₆-S-tagged protein from positive pET30 Ek/LIC constructs were overexpressed following transformation in *E. coli* BL21(DE3) cells (Novagen). Transformed *E. coli* BL21 DE3 cells were grown in 1 l volumes of LB containing 1-2 % (v/v) glycerol, 1 % glucose and 30 µg kanamycin ml⁻¹, with shaking for 4–5 h at 37 °C. Cells at an OD₆₀₀ of 0.8–1 were placed on ice for 1 h, before adding IPTG to 0.5 mM (final concentration). Incubation was continued at 16 °C for 1 h (no shaking), and for 16 h (with shaking) in a refrigerated incubator. Protein constructs were purified by nickel affinity chromatography, as described previously (Ding et al., 2001). The purified proteins were then concentrated using a Vivaspin column (Vivasience 10 000 MWCO), with a final wash using 50 mM sodium phosphate buffer, pH 6.5. Purified proteins were analysed by SDS-PAGE (10 % polyacrylamide gel), blotted onto a PVDF membrane, and hybridized to a conjugated Anti-His-C-term antibody, following the manufacturer’s instructions (Invitrogen). Additionally, purified proteins were tested for activity in wells on 0.4 % agarose plates prepared with 25 mM potassium phosphate buffer, pH 6.8, containing 0.25 % starch. Activity was visualized as a clear halo after iodine staining.

The protein concentration of each extract, and the activity against starch. Activity was visualized as a clear halo after iodine staining.

**RESULTS AND DISCUSSION**

**Starch utilization by selected human colonic anaerobes**

Several abundant species of butyrate-producing bacteria from the human colon, including *R. intestinalis* and *E. rectale*, are reported to utilize starch (D’Elia & Salyers, 1996; Duncan et al., 2002; Barcenilla et al., 2000). The growth of four selected butyrate-producing strains, all of which belong to the Gram-positive clostridial cluster XIVa, on six different types of starch is shown in Table 2. Starch stimulated the growth of the four butyrate-producing strains to a greater extent than the amylolytic Gram-negative bacterium *Bacteroides theta-totaeomicron* 5482. There was a general preference for high amylpectin starches, and the maximum specific growth rate (μmax) values decreased as the amyllose content of the starch increased (Table 2). Because OD₆₀₀ values were required to be corrected for the turbidity of high-amyllose (HA) substrates (see Methods), additional experiments were conducted to monitor the production of bacterial protein. These confirmed the growth of *B. fibrisolvens* 16/4 and *R. inulinivorans* A2-194 on amylpectin corn starch (4·95 and 9·23 mg protein ml⁻¹, respectively, in 24 h), and on waxy corn starch (4·12 and 8·66 mg protein ml⁻¹, respectively); the two strains grew poorly on HA corn starch (1·43 and 2·18 mg protein ml⁻¹, respectively).

**Activity, location and size of amylases from human *Roseburia* and *Butyrivibrio* strains**

At least 85 % of the amylase activity detected in *B. fibrisolvens* 16/4, *R. inulinivorans* A2-194 and *Bacteroides theta-totaeomicron* 5482 cultures was associated with the cellular fraction rather than the supernatant (Fig. 1). In the case of *Bacteroides theta-totaeomicron*, most amylolytic activity is known to be periplasmic or outer-membrane associated (Anderson & Salyers 1989; Shipman et al., 1999). For the two Gram-positive strains, however, this result implies that the amylases are cell associated. Amylase activities were induced in cells grown on maltose or amylpectin corn starch, compared with glucose-grown cells. *B. fibrisolvens* 16/4 showed 10-fold higher amylase activity than *R. inulinivorans* A2-194 when grown on amylpectin (Fig. 1). Amylase activity was detected following SDS-PAGE zymogram analysis of total cell proteins from one *B. fibrisolvens*, and three *Roseburia* strains, as shown in Fig. 2. A major high molecular mass (>150 kDa) band showing starch hydrolisys activity was detected in each strain, whether grown on maltose or amylpectin, indicating the production of a large active amylase enzyme.
Identification of two genes encoding large α-amylases from *Roseburia* and *Butyrivibrio* strains

Consensus PCR primers designed to recognize family 13 amylases (AmyPfor and AmyPrev) were used to amplify the predicted 750 bp fragment from *B. fibrisolvens* 16/4, and this was then sequenced. The gene sequence was completed by genome walking. The full-length ORF was 4002 bp, and encoded a protein of 1333 aa, with a predicted molecular mass of 145 kDa, named Bf-Amy13B, and this was comparable with the size of the major band estimated by zymogram analysis (Fig. 2). The protein is most similar (50 % amino acid identity) to a previously characterized extracellular α-amylase (Bf-Amy13A) from the ruminal *B. fibrisolvens* H17c (Rumbak *et al.*, 1991). The enzyme from *B. fibrisolvens* H17c (976 aa) is significantly smaller than *B. fibrisolvens* 16/4 Amy13B, and the greatest amino acid sequence homology is found in the N-terminal catalytic region.

Separately, a number of clones were identified from *R. inulinivorans* A2-194 that were specifically expressed following growth on starch (K. P. Scott, unpublished results). The sequence of one of these clones, after extension by genome walking using a succession of gene-specific primers, was found to encode a large 1674 aa polypeptide, which also contained a family 13 domain. The predicted molecular mass of the protein (184 kDa) was again consistent with the size of the major amylase observed in

---

**Fig. 2.** Zymogram showing activity of cell-associated amylases against amylpectin corn starch. Cells were pregrown to exponential phase on amylpectin (A) or maltose (M). Values on the left correspond to the molecular masses determined by staining the gel with Coomassie blue, prior to staining the gel with iodine to visualize clear zones of amylase activity. Lanes: 1, molecular mass markers; 2, *R. inulinivorans* A2-194 (A); 3, *R. inulinivorans* A2-194 (M); 4, *B. fibrisolvens* 16/4 (A); 5, *B. fibrisolvens* 16/4 (M); 6, *R. intestinalis* L1-952 (A); 7, *R. intestinalis* L1-952 (M); 8, *R. intestinalis* L1-82 (A); 9, *R. intestinalis* L1-82 (M).

---

**Fig. 3.** Multi-domain structure of the (a) *B. fibrisolvens* 16/4 Bf-Amy13B, and (b) *R. inulinivorans* A2-194 Ri-Amy13A α-amylase enzymes. Specific domains are marked with different shading. R1 and R2, major amino acid repeat units; SG, serine/glycine-rich domain; HD, hydrophobic C-terminal domain; B, cluster of basic residues; PUD, pullulanase-associated domain. The locations of the alternative XXXTG motifs are shown. Domain predictions are based on ProDom, Pfam and PDB database searching. This figure is drawn approximately to scale. Double-headed arrows below each sequence indicate the regions cloned into the pET30 Ek/LIC expression vector.
this strain by zymogram analysis (Fig. 2). BLASTP analysis indicated that the complete enzyme sequence had greatest identity (39%) to an amylpolullanase from Bacillus sp. (accession no. BAA11322.1).

**Multi-domain organization of the α-amylase enzymes**

The two putative amylase gene products identified here have complex multi-domain structures (Fig. 3), starting with characteristic Gram-positive N-terminal signal peptides (SPs), predicted using the SignalP 3.0 program and neural-network prediction (Nielsen et al., 1997; Bendtsen et al., 2004). In *B. fibrisolvens* 16/4, the SP cleavage site is predicted between A33 and T34, while in *R. inulinivorans* A2-194, it is between A34 and L35. The *B. fibrisolvens* 16/4 Bf-Amy13B includes a family 13 glycoside hydrolase domain of 462 aa (V46–L507) that shares 62% amino acid identity with the catalytic domain of Amy13A from the rumen *B. fibrisolvens* strain H17c (AMY BUTFI in Fig. 4; Rumbak et al., 1991). The *R. inulinivorans* A2-194 enzyme, which we will designate Ri-Amy13A, also contains a family 13 catalytic domain of 483 aa (A924–K1407), which shares only 20% amino acid identity with that from *Bf*-Amy13B, and has closest identity (45%) to the α-amylase from *Micrococcus* sp. strain 207 (AMY MICSP in Fig. 4). The three key catalytic residues within the four conserved regions of α-amylases (Rumbak et al., 1991; MacGregor et al., 2001) can be identified in both enzymes (*Bf*-Amy13B, D233,E282,D354; *Ri*-Amy13A, D1127,E1156,D1222). The catalytic specificity of the two family 13 catalytic domains was determined following expression in *E. coli* (discussed later; Table 3).

The *B. fibrisolvens* 16/4 Bf-Amy13B enzyme contains two short repeated sequences of 88 aa, which share 46% amino acid identity and resemble family 26 carbohydrate-binding modules (CBM). Similar repeat units have been identified in a number of other α-amylases from Gram-positive bacteria, including *B. fibrisolvens* H17c (Rumbak et al., 1991) and *L. amylovorus* (Giraudeau & Cuny, 1997). In the latter bacterium, the repeat units are responsible for binding substrates, including starch (Rodriguez et al., 2000). The 3D crystal structure of CBM26 from *Bacillus halodurans* has been recently resolved (BhCBM26, Boraston et al., 2006), and a computer model of *Bf*-Amy13B-CBM26 was generated using BhCBM26 as a template (Fig. 5). The predicted model illustrates the accessibility of the conserved aromatic residues, proven to interact with sugar molecules in BhCBM26 (Boraston et al., 2006). A further two repeats were present adjacent to CBM26 in *Bf*-Amy13B, with significant homology to CBM25 (21% identity), also found in *Bacillus halodurans* amylase (Boraston et al., 2006). Attempts were made to create a computer-generated model using BhCBM25 as a template, but the resulting model could not mirror the binding pocket of BhCBM25. Nevertheless, the role of these repeats in carbohydrate recognition cannot be ruled out, and further physico-chemical characterization is necessary.

The *R. inulinivorans* A2-194 Ri-Amy13A protein contains three large and three small domains, of undetermined
function (Fig. 3). Only two of these have significant similarity with sequences in the ProDom and Pfam databases: an amylopullulanase \( \alpha \)-amylase alkaline domain, and a pullulanase-associated carbohydrate-binding domain. Two short repeat domains of 132 aa each, which were rich in aromatic residues, and followed the SP, were identified in \( R_i \)-Amy13A using Radar (www.ebi.ac.uk/Radar/), and found to share 37% amino acid identity with each other. Comparison of these two repeats with the CBM26 repeats identified in \( B_f \)-Amy13B revealed a low, but significant, similarity, specifically in the alignment of most of the aromatic amino acids. Secondary structure prediction analysis of these repeats also revealed a predominance of \( \beta \)-sheet folding, which is characteristic of carbohydrate-binding domains (Boraston et al., 2004). These results indicate that the two repeats located at the N-terminus of \( R_i \)-Amy13A are probably involved in carbohydrate-binding activity, despite not belonging to any known family of CBM.

**Table 3.** Activity of the purified catalytic domains of the \( \alpha \)-amylase enzymes of \( B. \) fibrisolvens 16/4 and \( R. \) inulinivorans A2-194

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Enzyme activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>16/4</td>
<td>12-144</td>
</tr>
<tr>
<td>Maize</td>
<td>12-144</td>
</tr>
<tr>
<td>Potato</td>
<td>12-144</td>
</tr>
<tr>
<td>Pulullan</td>
<td>12-144</td>
</tr>
<tr>
<td>Rice</td>
<td>12-144</td>
</tr>
<tr>
<td>Wheat</td>
<td>12-144</td>
</tr>
<tr>
<td>pnp-Maltopentaoside</td>
<td>12-144</td>
</tr>
</tbody>
</table>

Enzyme activities are expressed as nmol reducing sugars released min\(^{-1}\) (mg protein\(^{-1}\)). Values are the mean of triplicate readings from three independent experiments (±SD), except the readings for \( p \)-nitrophenyl \( \alpha \)-D-maltopentaoside (pnp-maltopentaoside), which are the means of triplicate results from one experiment.

**Putative C-terminal cell-surface anchoring sequences**

The C-terminus of the \( B_f \)-Amy13B and \( R_i \)-Amy13A enzymes contains a hydrophobic region (16/4, L\(_{1307}\)–V\(_{1329}\) and A2-194, L\(_{1660}\)–Y\(_{1688}\)) capable of forming a helical transmembrane region, and a short basic C-terminus (\( B_f \)-Amy13B, RKRK\(_{1333}\); \( R_i \)-Amy13A, RKNKHF\(_{1674}\); Fig. 3). This structure is characteristic of proteins that are anchored to the cell wall in Gram-positive bacteria (Navarre &
Schneewind, 1999), and it is typically preceded by a signal motif that is recognized by a sortase enzyme. Such sortase signal motifs (LPXTG) can be quite variable among the first three amino acid residues of the motif (Pallen et al., 2001; Ton-That et al., 2004), but the threonine and glycine are highly conserved since they form the cleavage site for the sortase enzyme, and the subsequent covalent linkage to the cell-wall peptidoglycan (Marraffini et al., 2004). The sequences AVD TG in Bf-Amy13B, and SQTTG in Ri-Amy13A, fit the criteria for cell-wall sorting signals. Furthermore, hydrophobicity plots confirm the putative transmembrane nature of this region (data not shown). The Bf-Amy13B enzyme also contains a 27 aa sequence rich in serine and glycine residues (117GASGNGGSSGSSSAGSSAGS1198), including the SGNG consensus sequence for glycosaminoglycan attachment. Serine/glycine-rich regions have been associated with binding sensus sequence for glycosaminoglycan attachment. Serine/glycine-rich regions have been associated with binding proteins to the Gram-positive cell wall (e.g. Sara et al., 1998), but they can also act as flexible poly-linker sequences (Hogg et al., 2003). Attachment of these two proteins to the peptidoglycan of the cell wall may account for additional bands of activity in zymograms that appear higher than the estimated molecular mass for the mature protein (Fig. 2).

Catalytic activity of the cloned family 13 domains from B. fibrisolvens 16/4 and R. inulinivorans A2-194

The family 13 putative catalytic domains of each enzyme (Fig. 3) were cloned into the pET-30 vector for expression in E. coli. Expressed proteins were purified from bulk preparations of active clones (see Methods), and analysed by SDS-PAGE. The observed sizes for the recombinant B. fibrisolvens 16/4 (His6-Amy13B) and R. inulinivorans A2-194 (His6-Amy13A) enzymes corresponded to those expected for the cloned products (~66 and 94 kDa, respectively), and the sequence across the cloning junction was confirmed. The purified enzymes formed clear zones on agarose plates containing amylopectin, but no activity was observed in reducing sugar assays (data not shown). Further investigation revealed a requirement for Mg\(^{2+}\) or Ca\(^{2+}\): optimal activity was obtained for Bf-His6-Amy13B with at least 20 mM Mg\(^{2+}\) or 2 mM Ca\(^{2+}\), while Ri-His6-Amy13A required at least 2 mM Mg\(^{2+}\) or 2 mM Ca\(^{2+}\) (data not shown). Enzyme assays performed in the presence of 20 mM Mg\(_{2}\text{Cl}_2\) and 20 mM Ca\(_{2}\text{Cl}_2\) (Table 3) showed that both enzymes were active against a range of starch substrates, with greatest activities against the high-amylopectin-containing starches (rice and potato; Table 3). Bf-His6-Amy13B showed 50-fold greater activity against p-nitrophenyl \(\alpha\)-D-maltopentaoside than Ri-His6-Amy13A. TLC was used to analyse the degradation products of maltotriose, maltose and panose, following incubation with the purified recombinant enzymes. Neither Bf-His6-Amy13B nor Ri-His6-Amy13A was able to degrade panose, but both enzymes degraded maltotriose into maltose and glucose. These results are consistent with 1,4-\(\alpha\)-glucan glucanohydrolase (\(\alpha\)-amylase; EC 3.2.1.1) activity, and the observed activity against amylopectin and pullulan is probably due to cleavage of 1,4-\(\alpha\) linkages in the structure, but not hydrolysis of 1,6-\(\alpha\) linkages.

Conclusions

Most of the actively amylolytic bacteria that have been identified from the human colon are Bacteroides and Bifidobacterium spp., but this may partly reflect the relative oxygen tolerance and cultivability of these groups. This study establishes that abundant Gram-positive anaerobes from the human colon that are related to Roseburia and Butyribivrio also have the ability to degrade and utilize a variety of starches. These bacteria may therefore play an important role in the formation of short-chain fatty acids, especially butyrate, from dietary starch that reaches the large intestine, and this may explain the butyrogenic effect reported for resistant starch in vivo (Silvi et al., 1999; Le Blay et al., 1999; Schwertz et al., 2002).

Amylase activity was largely cell associated in the Roseburia and Butyribivrio strains studied here. A major amylase, > 150 kDa in size, was detected in each strain by zymogram analysis, although we do not exclude the existence of other amylases that are less active, or that do not renature following SDS-PAGE. Amylase genes whose products correspond in size to the major amylases produced by these strains were analysed from two species: R. inulinivorans and B. fibrisolvens. Both enzymes were multi-modular, including an SP, a family 13 \(\alpha\)-amylase catalytic domain, and a C-terminus indicative of sortase-mediated attachment to the bacterial cell wall. This reliance on cell-wall anchored amylases in Gram-positive anaerobes provides a contrast with the largely periplasmic multi-enzyme starch-degrading system of Bacteroides thetaiotaomicron described by others (Shipman et al., 1999; Cho et al., 2001). The family 13 catalytic domains of the enzymes from B. fibrisolvens and R. inulinivorans appear to be \(\alpha\)-amylases capable of attacking 1,4-\(\alpha\) linkages in both starch and pullulan (Mathupala et al., 1990). The roles of the additional unidentified domains, and their possible effects on substrate binding and catalytic activity, merit further investigation. In the highly competitive community of the large intestine, there is a clear ecological benefit from retaining such enzymes on the bacterial surface, thus ensuring that the site of hydrolysis is close to the systems for transporting hydrolysis products into the bacterial cell.

ACKNOWLEDGEMENTS

The authors would like to thank Pauline Young for the automated DNA sequencing, and Sylvia Duncan for critical reading of the manuscript. This work was supported in part by SEERAD (Scottish Executive Environment and Rural Affairs Department), A.G.R. received a BBSRC/CASE PhD studentship sponsored by Nestlé UK Ltd. M. T. R. is supported by EU Grant GEMINI (QLRT-2001-02056).
REFERENCES


