Identification and Biosynthesis of Cyclic Enterobacterial Common Antigen in Escherichia coli

Paul J. A. Erbel,1 Kathleen Barr,2 Ninguo Gao,3 Gerrit J. Gerwig,4 Paul D. Rick,2,* and Kevin H. Gardner1*

Department of Biochemistry1 and Department of Pharmacology,3 University of Texas Southwestern Medical Center, Dallas, Texas 75390-9038; Department of Microbiology and Immunology, F. Edward Hebert School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, Maryland 208142; and Department of Bio-Organic Chemistry, Bijvoet Center, Utrecht University, 3508 TB Utrecht, The Netherlands4

Received 8 October 2002/Accepted 13 December 2002

Phosphoglyceride-linked enterobacterial common antigen (ECAPG) is a cell surface glycolipid that is synthesized by all gram-negative enteric bacteria. The carbohydrate portion of ECAPG consists of linear heteropolysaccharide chains comprised of the trisaccharide repeat unit Fuc4NAc-ManNAcA-GlcNAc, where Fuc4NAc is 4-acetamido-4,6-dideoxy-D-galactose, ManNAcA is N-acetyl-d-mannosaminouronic acid, and GlcNAc is N-acetyl-d-glucosamine. The potential reducing terminal GlcNAc residue of each polysaccharide chain is linked via phosphodiester linkage to a phosphoglyceride aglycone. We demonstrate here the occurrence of a water-soluble cyclic form of enterobacterial common antigen, ECACYC, purified from Escherichia coli strains B and K-12 with solution nuclear magnetic resonance (NMR) spectroscopy, electrospray ionization mass spectrometry (ESI-MS), and additional biochemical methods. The ECACYC molecules lacked an aglycone and contained four trisaccharide repeat units that were nonstoichiometrically substituted with up to four O-acetyl groups. ECACYC was not detected in mutant strains that possessed null mutations in the wecA, wecF, and wecG genes of the wec gene cluster. These observations corroborate the structural data obtained by NMR and ESI-MS analyses and show for the first time that the trisaccharide repeat units of ECAcyC and ECApg are assembled by a common biosynthetic pathway.

Lipopolysaccharide (LPS) is the major cell surface glycolipid of gram-negative bacteria. However, the cell surface of all gram-negative enteric bacteria contains an additional glycolipid, the phosphoglyceride-linked enterobacterial common antigen (ECApg) (16, 21, 24, 32). The carbohydrate portion of ECA is a linear heteropolysaccharide comprised of the trisaccharide repeat unit \(\alpha\)-d-Fuc4NAc-(1→3)-\(\beta\)-d-ManNAcA-(1→4)-\(\alpha\)-d-GlcNAc-(1→) where Fuc4NAc, ManNAcA, and GlcNAc denote 4-acetamido-4,6-dideoxy-D-galactose, N-acetyl-d-mannosaminouronic acid, and N-acetyl-d-glucosamine, respectively (Fig. 1A) (19, 22, 32). Individual ECA polysaccharide chains are covalently linked to diacylglycerolphosphate via a glycosidic linkage between the potential reducing terminal GlcNAc residue and the phosphate residue of the algycone (16, 17, 30); the phosphoglyceride aglycone anchors the ECA chains to the outer membrane (2, 33). Accordingly, phosphoglyceride-linked ECA chains are referred to as ECApg. While the function of ECAPG remains to be established, recent studies suggest that Wzy-catalyzed polymerization of repeat units is assembled as an undecaprenylpyrophosphate-linked intermediate (lipid III) (Fig. 1B) (4, 5, 31, 32). Accordingly, synthesis of the repeat unit is initiated by the transfer of GlcNAc 1-P from UDP-GlcNAc to undecaprenylphosphate to yield undecaprenylpyrophosphate-GlcNAc (lipid I) catalyzed by WecA (5, 31, 32). Subsequent reactions involve the successive transfer of ManNAcA and Fuc4NAc from the donors UDP-ManNAcA and TDP-Fuc4NAc, catalyzed by WecG and WecF, respectively.

Although synthesis of lipid III occurs on the cytosolic face of the cytoplasmic membrane, currently available evidence suggests that Wzy-catalyzed polymerization of repeat units to form linear polysaccharide chains occurs on the periplasmic face of the membrane. This requires the transbilayer movement of lipid III to the periplasmic face of the membrane, and it has been suggested that this step is mediated by a “flippase” encoded by the wexE gene (o416) (20). Finally, polymerization is followed by the transfer of polysaccharide chains from the lipid carrier to an as yet unidentified acceptor to yield phosphoglyceride-linked chains, and the completed ECApg molecules are then translocated to the outer membrane. However, essentially nothing is known regarding the genes and mechanisms involved in the latter two steps.

ECApg is regarded as the major form of ECA, and it is present in all gram-negative enteric bacteria (16). ECApg accounts for approximately 0.2% of the cellular dry weight of E. coli K-12 (18, 22). Two related forms, ECA_LPS and ECA_mur.
have also been identified in certain organisms. ECA_{LPS} molecules possess the same linear ECA polysaccharide chains found in ECA_{PG}, but in the case of ECA_{LPS} these chains are covalently linked to the core region of LPS (15, 16) instead of a phosphoglyceride aglycone. In contrast, ECA_{Cyc} is a water-soluble polymer that contains only ECA trisaccharide repeat units (16). In addition, the degree of polymerization of ECA_{Cyc} molecules is quite different from that observed for linear ECA polysaccharide chains. For example, the polysaccharide chains of ECA_{PG} synthesized by E. coli K-12 exhibit a population that ranges from 1 to 14 repeat units in length, with a modal value of 5 to 7 repeat units (3). In contrast, ECA_{Cyc} molecules isolated from Shigella sonnei contain only four to six trisaccharide repeat units (11).

Although the structure of ECA_{Cyc} has been characterized, nothing is known about its function, and there is no information available regarding the genetics and biosynthesis of this novel molecule. This is due, in large part, to the general belief that the occurrence of ECA_{Cyc} within members of the Enterobacteriaceae is rather restricted, since it has only been found in cell extracts of Shigellosis phase I (11, 19), Yersinia pestis (39), and Plesiomonas shigelloides (7, 37); the last organism has now been included in the Enterobacteriaceae.

The results presented in this communication describe the occurrence and characterization of ECA_{Cyc} in E. coli strain B as determined by a variety of methods, including nuclear magnetic resonance (NMR) spectroscopy and electrospray ionization mass spectrometry (ESI-MS). ECA_{Cyc} was initially found to copurify with the C-terminal PAS (Per-Arnt-Sim) domain of the human hypoxia-inducible factor 2 (HIFd) (38) following its overexpression as a recombinant protein in E. coli B. However, the detection of ECA_{Cyc} in these preparations was fortuitous because it was not found to be associated with HIFd and its synthesis was independent of the overexpression of this protein. ECA_{Cyc} was also found in cell extracts of E. coli K-12, and similar to the results obtained with E. coli B, its synthesis was independent of the overexpression of HIFd. Finally, the results of genetic and biochemical analyses show for the first time that the trisaccharide repeat units of ECA_{Cyc} and ECA_{PG} are assembled by a common biosynthetic pathway.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and plasmids. E. coli strains used in this study are listed in Table 1. Transformations were carried out with phage P1 Puri according to Silhavy et al. (36). Introduction of plasmid pKG31 into recipient strains was carried out either by transformation or electroporation with standard procedures. For the routine propagation of bacteria were grown at 37°C in Luria-Bertani (LB) broth or on LB agar containing 0.2% glucose (28). Where indicated, 15N-labeled protein and 13C-labeled ECA were isolated from cells grown in M9 minimal medium (28) containing 0.1% 15NH4Cl and 0.3% glucose (either natural abundance or 99% enriched and uniformly labeled), and expression was allowed to proceed overnight at 20°C. The cells were harvested by centrifugation and handled at 4°C for all remaining purification steps. The pellet was resuspended in 25 mL of 50 mM sodium phosphate buffer (pH 7.6)–15 mM NaCl–5 mM dithiothreitol, lysed by high-pressure extrusion, centrifuged, and filtered (0.22 μm), and the supernatant was purified with a Source 15Q anion-exchange column (Amersham Biosciences) pre-equilibrated with the above buffer. GB1-HIFd eluted from the column with the column with 2 volumes of the same buffer. Protein-containing fractions were pooled and concentrated in an Amicon pressure-driven ultrafiltration cell with YM10 10-kDa filters.

![Diagram](https://example.com/diagram.png)
TABLE 1. Bacterial strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype and information</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21(DE3)</td>
<td>F' ompT hsdS, (rB mB) gal dcm met (DE3)</td>
<td>Novagen</td>
</tr>
<tr>
<td>BL21(DE3)pLysS</td>
<td>F' ompT hsdS, (rB mB) gal dcm met (DE3) pLysS (Cm')</td>
<td>Novagen</td>
</tr>
<tr>
<td>HMS174(DE3)</td>
<td>F recA1 hsdR1K12, mK12 (II) Rif (DE3)</td>
<td>Novagen</td>
</tr>
<tr>
<td>AB1133</td>
<td>thr-1 leuB6 Xgal-proA66 hisG4 argF3 rbi-1 rbd1 lacY1 ara-14 galK2 ysl-5-nit-mgl-5 rpsL31 kdgK31 supE44</td>
<td>Laboratory Collection, received as CGSC 11339</td>
</tr>
<tr>
<td>MC4100</td>
<td>F' araD139 Δ(arg-lac)169 X c14 fliD5301 fnuA25 relA1 rpsL150 rbsR22 deoC1</td>
<td>Laboratory Collection, received as CGSC 6152a</td>
</tr>
<tr>
<td>21548</td>
<td>As AB1133 but wacA: Tn10</td>
<td>36</td>
</tr>
<tr>
<td>21568</td>
<td>As AB1133 but wgcG: Tn10</td>
<td>36</td>
</tr>
<tr>
<td>PND788</td>
<td>As MC4100 but ompR:: Tn10(Tet) wecA: Tn10(Cam) 8R8S885 [degP-lac-Z]</td>
<td>This study</td>
</tr>
<tr>
<td>PR4185</td>
<td>BL2(DE3)pLysSpKGM3</td>
<td>This study</td>
</tr>
<tr>
<td>PR4186</td>
<td>BL2(DE3)pKGM3</td>
<td>This study</td>
</tr>
<tr>
<td>PR4153</td>
<td>PR4153 as wacA:: Tn10 [P1(21548) × BL2(DE3)] pLysS, then pKGM3 (transformation)</td>
<td>This study</td>
</tr>
<tr>
<td>PR4164</td>
<td>PR4164 as wgcG: Tn10 [P1(21568) × BL21(DE3), then pKGM3] (transformation)</td>
<td>This study</td>
</tr>
<tr>
<td>PR4167</td>
<td>PR4167 as wacA:: Tn10 [P1(PND788) × BL21(DE3), then pKGM3] (transformation)</td>
<td>This study</td>
</tr>
</tbody>
</table>

a E. coli Genetic Stock Center; M. Berlyn, Biology Department, Yale University, New Haven, CT 06520.

The concentrated GB1-HIFd was digested with tobacco etch virus protease (13), followed by removal of the cleaved GB1 fragment by passage of the digest through an immunoglobulin G-Sepharose affinity column (Amersham Biosciences). HIFd, which eluted in the flowthrough volume of this column, was concentrated in an Amicon ultrafiltration system with a YM3 3-kDa filter and then loaded onto a HiLoad 26/60 Superdex 75 column (Amersham Biosciences) equilibrated in 50 mM sodium phosphate buffer (pH 7.2–7.5) with a linear gradient of 0 to 5% acetonitrile in H2O. The chromatogram was high-pressure liquid chromatography (HPLC) (Vydac C18 column, 0.5 by 25 cm) introduced into the source at a rate of 5 μl/min by an infusion pump (Harvard Apparatus), and mass spectra were acquired over a range of 300 to 1,700 per s.

**FACE analyses.** ECA samples were hydrolyzed at 100°C with 0.1, 0.25, 0.5, and 1 N HCl for 30 min and then dried under reduced pressure with a Speed Vac apparatus (Savant Instruments). The hydrolyzed samples were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE) with a FACE apparatus, oligosaccharide profiling kit, and reagents according to the directions of the manufacturer (Glyko). Accordingly, reducing termini were labeled with the neutral fluorophore 2-aminoacridone and then resolved on an oligosaccharide profiling gel. Labeled oligosaccharides were detected on the gel, and electronic images of the gel were generated with a Bio-Rad Fluor-S Multi-Imager equipped with a 500- and 600-MHz spectrometers, generally at 27°C. NMR spectra of the ECA\_\text{score}, which copurified with HIFd were recorded in samples containing 0.8 mM protein in 50 mM sodium phosphate (pH 7.2)–15 mM NaCl-5 mM dithiothreitol-10% D2O. NMR spectroscopic analysis of HPLC-purified ECA\_\text{score} was performed on a sample that contained the polysaccharide at a concentration of approximately 100 μM. ECA samples were dissolved in 550 μl of either 99.96% D2O or H2O:D2O (90:10, by volume) mixtures. One-dimensional 1H-NMR spectra were recorded with presaturation of the water resonance during the 2-s relaxation delay. Two-dimensional total correlation spectroscopy spectra were recorded with an MLEV-17 spin-lock pulse sequence (7 and 100 ms) and two-dimensional nuclear Overhauser enhancement spectroscopy spectra were recorded with a mixing time of 300 and 500 ms. Carbon and nitrogen chemical shift assignments were based on 13C–1H and 15N–1H heteronuclear single-quantum coherence (HSQC) spectra, respectively. Additional chemical shift data were obtained from standard three-dimensional NMR experiments generally used to assign protein backbone and side chain atoms, including HNCA, CBCA(CO)NH, HNCO (34), and (HBCBCA)COCAHA (14). Chemical shifts were referenced to the methyl H signals of sodium 2,2-dimethyl-2-silapentane 5-sulfonate, with direct referencing for all H shifts and indirect referencing for 13C and 15N shifts (23).

Crude cell lysate NMR samples were prepared from 200 ml cultures of BL21(DE3) or HMS174(DE3) cells transformed with pKG31 that were allowed to grow overnight at 30°C in M9 minimal medium containing 15NH4Cl, with and without induction of GB1-HIFd fusion protein expression. Cells were harvested, lysed, centrifuged, and filtered as described before. The supernatant was concentrated to 0.5 ml and used to record standard 1H–15N HSQC spectra (12).

**RESULTS**

Copurification of ECA with HIFd. HIFd is believed to play an integral role in the function of human hypoxia-inducible factor 2, a eukaryotic transcription factor that responds to reduced intracellular oxygen levels (38). To conduct structural studies of this domain, we expressed the 13.2-kDa HIFd protein fragment in E. coli PR4186 (Table 1), a derivative of E. coli B. Initial 1H–15N HSQC spectra of HIFd displayed a well-dispersed resonance pattern indicative of a well-folded protein. Interestingly, three sets of 15N/1H correlations in these spectra exhibited particularly narrow line widths and peak doubling (Fig. 2A). These resonances gave unusual correlations in standard triple-resonance experiments commonly used for protein backbone chemical shift assignment, including HNCA and CBCA(CO)NH spectra.

When used on uniformly 15N/13C protein samples, these methods link 15N/1H chemical shifts to the 13C\_n and 15N\_n shifts of carbons on either side of the amide linkage. However, the
signals in these spectra generated by the three \( ^{15}\)N/\(^{1}H \) pairs clearly indicated that they were from amides linked to groups not normally found within protein samples. This is demonstrated in segments of CBCA(CO)NH and HNCACB spectra corresponding to these amides, which show that each \(^{15}\)N/\(^{1}H \) pair has only a single carbon at approximately 25 ppm present on the distal side of the amide linkage rather than the two expected \( ^{13}\)C/\(^{1}H \) signals (Fig. 2B). \(^{1}H\)H-\(^{13}\)C HSQC spectra show that this carbon is directly attached to protons at approximately 2 ppm, establishing that it is an \( N\)-acetyl group. The triple resonance spectra (Fig. 2B) also indicated that each amide had three signals to the proximal side, one from a \( ^{13}\)C/\(^{1}H \)-like site (\( ^{1}H \) at 55 ppm) and two from \( ^{13}\)C/\(^{1}H \)-like sites (\( ^{1}H \) at 70 and 100 ppm). All of these data are clearly inconsistent with standard amino acid structure, strongly suggesting that this sample contained nonprotein material.

To identify the source of these signals, we purified this material by ethanol extraction and reverse-phase HPLC to obtain a protein-free sample. A one-dimensional \(^{1}H\)NMR spectrum of this material showed a typical carbohydrate pattern, including very intense signals at approximately 2 ppm corresponding to various \( N\)-acetyl groups (Fig. 3).

FIG. 2. Identification of nonprotein amide resonances. (A) Expansion of the \(^{1}H\)-\(^{15}\)N HSQC spectrum of HIFd with copurified ECA (black) and HPLC purified ECA (red). Signals from the amides of the \( N\)-acetyl groups of ECA are characterized by notably narrow line widths and signal doubling compared to protein signals. (B) CBCA(CO)NH and HNCACB strips for each \( N\)-acetyl group of ECA. Black and red indicate positive and negative cross peaks, respectively. These spectra correlate carbon resonances with the amide \(^{15}\)N and \(^{1}H \) shifts. In particular, the CBCA(CO)NH experiment gives a positive cross peak between the \( C_{\alpha}\)-like site to the carbonyl (\( C_{\alpha}\)[H\_2]). The HNCACB experiment shows positive cross peaks for the same \( C_{\alpha}\)-like site (\( C_{\alpha}\)[H\_3]) and the \( C_{\beta}\)-like site with respect to the HN (\( C_{\beta}\)) with the same amide. In addition, negative cross peaks are observed in the HNCACB experiment for the \( C_{\beta}\)-like signals (see header). Note that the \( C_{\beta}\)-like anomeric signals (\( \sim\)100 ppm) occur outside of the chemical shift range of the \(^{13}\)C dimension (15 to 75 ppm) and are aliased into this range near 40 ppm. Arrows indicate the characteristic signal doubling of ECA\(_{\text{Cyc}}\) signals, the origin of which may be sample heterogeneity or slow time scale dynamics (see Discussion).
corresponding to anomic protons indicated that this compound contained three monosaccharides. These signals appeared with equal intensities at 4.86, 4.97, and 5.12 ppm, indicating that the monosaccharides were present in equal amounts. The identities of these were provided by preliminary monosaccharide composition analyses conducted on this NMR sample. These revealed that the carbohydrate contained three N-acetylated amino sugars in equal amounts: N-acetylgalactosamine, 4-acetamido-4,6-dideoxyhexosamine, and N-acetylhexosamine uronic acid. Glucose was also detected in lesser amounts (data not shown).

To more completely assign the NMR chemical shifts of this carbohydrate, we used a combination of two-dimensional homonuclear and heteronuclear NMR methods. The majority of the signals in the $^1$H-NMR spectrum were easily assigned with total correlation spectroscopy data, starting at the anomic proton resonances. Carbon resonances were assigned from $^1$H,$^1$C HSQC spectra, and nitrogen resonances were assigned by $^1$H,$^15$N HSQC of $^15$N-labeled ECA. These assignments are provided in Table 2. The program SUGABASE (http://boc.chem.uu.nl/sugabase/sugabase.html) was used to search for carbohydrate structures that correlated with these NMR data and the monosaccharide composition. Our $^1$H and $^13$C chemical shift assignments generally agree with those previously reported for ECA$_{cyc}$ from Plesiomonas shigelloides (37) and Yersina pestis (39) and also with the $^13$C assignments reported for Plesiomonas shigelloides ECA$_{pg}$ (7) (Table 2). However, a comparison of these chemical shift assignments of ECA$_{cyc}$ revealed small variations that are likely due to heterogeneity occurring from natural modifications, possible degradation during purification, and variations in experimental conditions.

To confirm the glycosidic linkages for this polysaccharide, we acquired several nuclear Overhauser enhancement spectroscopy cross peaks, and nitrogen resonances were assigned by $^1$H-15N HSQC of 15N-labeled ECA. These assignments revealed molecular ions of ECA$_{cyc}$ obtained from Plesiomonas shigelloides (37), were not found in the mass spectra of ECA$_{cyc}$ from E. coli. In this regard, it should be noted that the ECA$_{cyc}$ obtained from Plesiomonas shigelloides was also found to contain only four repeat units (37). However, the basis for the apparent organism-dependent variation in the degree of polymerization is not understood.

Additionally, NMR analysis of the ECA isolated by copurification with HIFd confirmed that this material is indeed ECA$_{cyc}$. Aside from ECA$_{cyc}$, all other forms of ECA (including ECA$_{pg}$, ECA$_{lps}$, and various biosynthetic intermediates) are extremely hydrophobic due to the chemical nature of the lipid molecules to which they are linked. These forms of ECA are poorly soluble, and they form large micelles in aqueous solution that would dramatically increase the line width and decrease the signal-to-noise ratio of NMR signals. As evidence of this, Basu et al. used high temperature (70°C) and large sample tubes (10 mm diameter) to record a one-dimensional $^13$C spectrum of ECA$_{pg}$ (7). In contrast, HIFd-associated ECA gave highly resolved NMR spectra at 27°C in a standard 5-mm NMR tube (Fig. 3A). Moreover, the $^1$H one-dimensional and $^1$H,$^13$C HSQC spectra of the ECA material did not reveal $^1$H and $^13$C signals of aliphatic groups, which are typically found between 1 and 2 ppm for $^1$H and around 30 ppm for $^13$C, that would be indicative of the lipid components of ECA$_{pg}$ and ECA$_{lps}$ (7). In addition, $^31$P-NMR analysis of HIFd-associated ECA did not indicate the presence of phosphodiester groups, which would be expected to occur in both ECA$_{pg}$ and ECA$_{lps}$, or phosphomonomesters, which would be expected to occur in ECA$_{lps}$ (data not shown).

Finally, it is important to note that ECA$_{pg}$, ECA$_{lps}$, and ECA$_{cyc}$ are characterized by the absence of a free terminal reducing sugar. Indeed, no signals for terminal reducing sugar residues were detected in the $^1$H- and $^13$C-NMR spectra obtained from the ECA associated with purified preparations of HIFd. The lack of a free reducing terminus was confirmed by FACE analyses. In this method, saccharides and oligosaccharides are labeled at the reducing terminus with the appropriate fluorescein probe to yield a derivative with a net negative charge, and the fluorescein derivative is then analyzed by gel electrophoresis. The electrophoretic mobility of the derivative is dependent on its charge-mass ratio as well as its hydrodynamic volume. Fluorescent labeling of HIFd-associated ECA with 2-aminoacridone proved to be unsuccessful (Fig. 4, lane 6), consistent with the lack of a free reducing terminus. In contrast, mild acid treatment of HIFd-associated ECA generated several oligosaccharide fragments with free reducing termini, as indicated by the 2-aminoacridone-labeled products shown in Fig. 4, lanes 2 to 5.
<table>
<thead>
<tr>
<th>Atom</th>
<th>GlcNAc</th>
<th>ManNAcA</th>
<th>Fuc4NAc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>H2</td>
<td>3.98</td>
<td>—</td>
<td>3.98</td>
</tr>
<tr>
<td>H3</td>
<td>3.84</td>
<td>—</td>
<td>3.82</td>
</tr>
<tr>
<td>H4</td>
<td>3.77</td>
<td>—</td>
<td>3.76</td>
</tr>
<tr>
<td>H5</td>
<td>3.84</td>
<td>—</td>
<td>3.86</td>
</tr>
<tr>
<td>H6</td>
<td>3.84, 3.79</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CH3</td>
<td>2.08</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>NAcH</td>
<td>2.02</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>OAcH</td>
<td>2.08</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>C1</td>
<td>94.1</td>
<td>94.3</td>
<td>94.2</td>
</tr>
<tr>
<td>C2</td>
<td>54.1</td>
<td>—</td>
<td>54.2</td>
</tr>
<tr>
<td>C3</td>
<td>70.4 (68.8)</td>
<td>—</td>
<td>70.6</td>
</tr>
<tr>
<td>C4</td>
<td>79.7</td>
<td>—</td>
<td>79.0</td>
</tr>
<tr>
<td>C5</td>
<td>71.3</td>
<td>—</td>
<td>71.4</td>
</tr>
<tr>
<td>C6</td>
<td>61.1</td>
<td>—</td>
<td>61.2</td>
</tr>
<tr>
<td>CH4</td>
<td>16.6</td>
<td>—</td>
<td>16.6</td>
</tr>
<tr>
<td>NHAc</td>
<td>23.2</td>
<td>—</td>
<td>23.0</td>
</tr>
<tr>
<td>OHAc</td>
<td>17.7</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Atom</th>
<th>1H(N)*</th>
<th>13N(H)*</th>
<th>13C(O)**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7.99</td>
<td>122 (122.4)</td>
<td>175.2 (175.6, 174.8)</td>
</tr>
<tr>
<td></td>
<td>8.30 (8.36)</td>
<td>114.8</td>
<td>176.4</td>
</tr>
<tr>
<td></td>
<td>7.87 (7.91, 7.98)</td>
<td>115.2 (115.8, 116.0)</td>
<td>175.2 (175.6, 174.8)</td>
</tr>
</tbody>
</table>

A, assignments of ECA from E. coli at 27°C (1H reference DSS = 0.007 ppm; 13C reference DSS = 1.84 ppm). B, partial assignment of ECA from P. shigelloides at 25°C (37) (1H reference TMS = 0.00 ppm; 13C reference dioxane = 67.40 ppm). C, Assignments of ECA from Y. pestis at 70°C (19) (1H reference acetone = 2.23 ppm; 13C reference acetone = 31.45 ppm). D, 13C assignments of ECA from P. shigelloides at 70°C (39) (13C reference TMS = 1.31 ppm). E, assignments of linear and lipid-free ECA from E. coli at 25°C (9) (1H reference TMS = 0.00 ppm; 13C reference dioxane = 67.40 ppm). !, assignment interchanged with respect to original paper (19); ? assignment may be reversed according to original paper (19, 39); *, measured in a 15N-labeled sample in H2O; **, measured in a 13C-labeled sample in the presence of HIFd in H2O at 30°C. Bold type indicates data obtained in the present study. Chemical shift values in parentheses are from minor species, as discussed in the text. —, chemical shift not reported.
We note that the FACE results here could possibly be produced by a new form of linear ECA which has its reducing end blocked by a novel modification. However, such modification would lead to unique NMR signals from the terminal Fuc4NAc, particularly for the anomeric proton. No such signals were observed (Fig. 3). Furthermore, the molecular mass of such a molecule would be significantly above 2,447 Da (molecular ion of unmodified linear polymer), which is not supported by our ESI-MS data.

In summary, these data establish that ECA purified in this manner does not consist of a linear ECA polysaccharide that was generated by some uncharacterized degradative process. Rather, we conclude that the HIFd-associated polysaccharide consists of ECA_CYC molecules, each of which contains four trisaccharide repeat units and an average of approximately two O-acetyl groups.

**Synthesis of ECA CYC in E. coli B and K-12 strains is independent of HIFd overexpression.** ECA_CYC has not been previously identified in *E. coli*. Accordingly, the data presented thus far raise questions as to whether ECA_CYC biosynthesis is limited to *E. coli* B strains or is induced by HIFd overexpression. To address these questions, the soluble fraction of cell lysates of *E. coli* B strain BL21(DE3) and *E. coli* K-12 strain HMS174(DE3) cultures transformed with pKG31 were analyzed for ECA_CYC with and without induction of HIFd overexpression. NMR analysis of 15N-labeled cell lysates showed intense amide signals characteristic of soluble ECA_CYC in 15N-1H HSQC spectra of HMS174(DE3) lysates independent of HIFd expression (Fig. 5). Similar results were obtained for BL21(DE3) lysates (data not shown). The chemical shift values of the amide 15N and 1H resonances of ECA_CYC overlaid very closely and were unaffected by HIFd (Fig. 2A and 5), suggesting that there is no interaction between ECA_CYC and HIFd. In these spectra, additional signals from other nitrogen-containing metabolites or small proteins can be observed. These data demonstrate that HIFd overexpression has no effect on ECA_CYC biosynthesis and that ECA_CYC is present in both *E. coli* B and K-12 strains.

**FIG. 4.** FACE analysis of acid hydrolyzed ECA. Lane 1, 2-aminoacridone control; lanes 2, 3, 4, and 5, 2-aminoacridone-derivatized products resulting from the treatment of HPLC-purified ECA with 1 N, 0.5 N, 0.25 N, and 0.1 N HCl for 30 min at 100°C, respectively; lane 6, unhydrolyzed HPLC-purified ECA control that was incubated with 2-aminoacridone under derivatizing conditions; lane 7, oligosaccharide standards containing four to nine glucose residues labeled with 8-aminonaphthalene-1,3,6-trisulfonic acid (three negative charges).

**FIG. 5.** NMR signals of ECA CYC in crude cell lysate. (A) Expansion of the 1H-15N HSQC spectrum of 15N-labeled crude cell lysate of HMS174(DE3) cells without induction of GB1-HIFd expression. (B) Overlay of 1H-15N HSQC spectra of 15N-labeled crude cell lysate of HMS174(DE3) cells with (red) and without (black) induction of GB1-HIFd expression. The characteristic strong amide signals of soluble ECA are indicated. The crude cell lysate spectrum of noninduced *E. coli* cells showed only a few 1H-15N resonances, indicating that ECA_CYC is present along with only a few other nitrogen-containing metabolites or small proteins present at high concentrations. In contrast, overexpression of a small protein like GB1-HIFd gives rise to many additional signals readily observed in this experiment.
Genetic loci involved in synthesis of ECA\textsubscript{CYC}. A considerable amount is known about the genes involved in synthesis of the linear ECA chains of ECA\textsubscript{PG} and ECA\textsubscript{LPS} (3, 8, 29, 32). In contrast, essentially nothing is known about the genetic determinants of ECA\textsubscript{CYC}. The \textit{wecA}, \textit{wecF}, and \textit{wecG} genes of \textit{E. coli} K-12 encode the GlcNac 1-P, Fuc4NAc, and ManNAcA transferases, respectively, that are involved in the assembly of the ECA trisaccharide repeat unit of linear ECA polysaccharide chains (8, 26, 27). Accordingly, null mutations in these genes completely abolish the synthesis of ECA\textsubscript{CYC} and ECA\textsubscript{LPS} (26, 27, 29), as determined by a variety of assays, including colony immunoblot assay (25), passive hemagglutination assay (31), and SDS-PAGE and fluorography analyses of cell envelopes prepared from cells grown in the presence of radiolabeled GlcNac (31). However, it is not known whether these genes also function for the synthesis of ECA\textsubscript{CYC}.

In this regard, it should be noted that the relatively small size of ECA\textsubscript{CYC} and the lack of a hydrophobic aglycone component of ECA\textsubscript{CYC} preclude its detection by many of the methods used to study the biosynthesis of the other ECA forms. In addition, it is not yet known whether the antibodies used for the colony immunoblot and passive hemagglutination assays recognize ECA\textsubscript{CYC}. Accordingly, HIF\textsubscript{d} was overexpressed in strains PR4153 (\textit{wecA::Tn10/pKG31}), PR4164 (\textit{wecG::Tn10/pKG31}), and PR4161 (\textit{wecF::Tn10/pKG31}) grown in the presence of \textit{15}N\textsubscript{H}\textsubscript{4}Cl, and attempts were made to detect the presence of \textit{15}N-ECA\textsubscript{CYC} in cell extracts of each of these strains as revealed in \textit{1H-15N} HSQC spectra. No N-acetyl signals attributable to ECA\textsubscript{CYC} were detected in any of these preparations (data not shown). These observations corroborate the conclusion stemming from NMR studies described earlier that the carbohydrate material present in purified preparations of HIF\textsubscript{d} is indeed an ECA polysaccharide. Furthermore, these findings provide the first evidence that the trisaccharide repeat unit of both ECA\textsubscript{CYC} and linear ECA polysaccharide chains is assembled by a common pathway.

**DISCUSSION**

ECA\textsubscript{PG} is a cell surface component of all gram-negative enteric bacteria (16), and it accounts for approximately 0.2% of the cellular dry weight of \textit{E. coli} K-12 (18, 22). In contrast, ECA\textsubscript{CYC} has thus far been found to occur in only a few members of the \textit{Enterobacteriaceae} (11, 16, 19, 39). The data presented in this study now demonstrate that ECA\textsubscript{CYC} is also synthesized by \textit{E. coli}, and it is entirely possible that future studies will reveal that ECA\textsubscript{CYC} is a constituent of many other gram-negative enteric bacteria. The apparent limited occurrence of ECA\textsubscript{CYC} among members of the \textit{Enterobacteriaceae} may simply reflect the fact that only a few organisms have been examined for the presence of ECA\textsubscript{CYC} because of the lack of readily available assays for its detection.

ECA\textsubscript{CYC} was initially found to copurify with the C-terminal PAS domain of the human hypoxia-inducible factor 2 (HIF\textsubscript{d}) (38) following its overexpression as a recombinant protein in \textit{E. coli} B. However, detailed analysis of the NMR spectra of HIF\textsubscript{d}, including three-dimensional \textit{15}N- and \textit{13}C-edited nuclear Overhauser enhancement spectroscopy spectra, did not reveal interactions between ECA\textsubscript{CYC} and HIF\textsubscript{d}. Furthermore, NMR analyses revealed strong signals for ECA\textsubscript{CYC} in the soluble fraction obtained from crude cell lysates of both \textit{E. coli} strains B and K-12 in the absence of HIF\textsubscript{d} synthesis. These data indicate that the synthesis of ECA\textsubscript{CYC} is independent of the overexpression of HIF\textsubscript{d}, and they also suggest that its occurrence in \textit{E. coli} is not strain specific.

Subsequent work has revealed that the methods used here to purify HIF\textsubscript{d} were chiefly responsible for the fortuitous discovery of ECA\textsubscript{CYC} in \textit{E. coli} K-12. Accordingly, the initial anion-exchange step for the isolation of the fusion protein, GB1-HIF\textsubscript{d}, is not very efficient because the protein binds weakly to the Source 15Q anion-exchange resin with the described phosphate buffer. Subsequent refinements of the isolation protocol employed a lower-ionic-strength buffer (50 mM Tris, pH 7.6), which resulted in increased binding of the protein to the resin without concomitant binding of ECA\textsubscript{CYC}, allowing the separation of these molecules from one another (P. Erbel, unpublished results). In contrast, attempts to employ gel filtration chromatography and molecular cutoff filters to separate the highly negatively charged and unusually shaped ECA\textsubscript{CYC} (2.4 kDa to 2.6 kDa) from HIF\textsubscript{d} (13.2 kDa) were unsuccessful.

Bruix et al. (10) were also unsuccessful in their initial attempts to use size exclusion chromatography to separate ECA from the comparably sized chemotactic protein CheY (14.0 kDa) from \textit{E. coli}. Che\textsubscript{Y}-associated ECA was ultimately isolated by repeated phenol extraction of the protein (9), generating material that was identified as a linear and lipid-free ECA polysaccharide. It is significant to note that these investigators identified a free reducing terminal Fuc4NAc residue (α-anomer at 5.23 ppm and β-anomer at 4.77 ppm) in their purified preparations. Since GlcNAc is the potential reducing terminal amino sugar of ECA trisaccharides, this observation suggests the possibility that these polysaccharides resulted from degradation of ECA\textsubscript{CYC} that was originally present in the fractions containing Che\textsubscript{Y}. All of these data suggest that ECA\textsubscript{CYC} could be a more general contaminant in samples of other recombinant proteins expressed in \textit{E. coli}.

ECA\textsubscript{CYC} is readily identified by characteristic signal doubling in a \textit{1H-15N} HSQC spectrum (Fig. 2), possibly caused by chemical heterogeneity (e.g., differential O-acetylation) and restrained rotational motion around the C-N bonds of the N-acetyl groups. In this context, it is interesting that Staaf and colleagues (7, 37) suggested that ECA\textsubscript{CYC} undergoes slow conformational changes based on NMR and molecular dynamics studies on unlabeled ECA\textsubscript{CYC} isolated from \textit{Plesiomonas shigelloides}. Accordingly, such slow conformational exchange processes might also account for the signal doubling observed in this study.

Based on the molar extinction coefficient of HIF\textsubscript{d} at 280 nm, the amount of purified HIF\textsubscript{d} from 1 g (wet weight) of cells was calculated to be approximately 15 mg. Consequently, the amount of ECA\textsubscript{CYC} can be estimated from the peak volumes of the N-acetyl signals of ECA\textsubscript{CYC} relative to the protein backbone amide signals (ratio is 1.2) in \textit{1H-15N} HSQC spectra (Fig. 2A). With this ratio, and taking into account both the average molecular mass of ECA\textsubscript{CYC} and the fourfold redundancy of the N-acetyl signals, it was estimated that approximately 0.4 mg of ECA\textsubscript{CYC} was isolated from 200 mg (dry weight) of \textit{E. coli} cells. Surprisingly, this suggests that ECA\textsubscript{CYC} and ECA\textsubscript{PG} are present in similar amounts in \textit{E. coli}. It is important to stress that nothing is known about possible fac-
tors that may affect the amounts of ECA CYC and ECA PG synthesized by cells. Thus, a more accurate determination of the cellular quantity of these molecules will require direct assays as well as additional information about the possible regulation of their synthesis. This investigation revealed that the trisaccharide repeat units of ECA CYC and linear ECA polysaccharide chains are assembled as a lipid-linked intermediate (lipid III) by a common biosynthetic pathway that involves enzymes encoded by the wecA, wectG, and wecF genes of the wec gene cluster of E. coli K-12. Although it has previously been assumed that these genes play a role in ECA CYC synthesis, there in fact exists no direct evidence to validate this assumption. Accordingly, the data presented here constitute the first information regarding genetic loci involved in the synthesis of this molecule.

It is highly likely that the pathways for the assembly of water-soluble ECA CYC and the linear ECA chains of ECA PG and ECA LPSP diverge following synthesis of lipid III. Indeed, it has been suggested that ECA CYC may be a component of the cytoplasm (1, 16). In this event, it seems likely that the assembly of ECA CYC would most likely occur on the inner leaflet of the cytoplasmic membrane by a WzyE-independent mechanism, and it would not require WzxE-mediated translocation of lipid III across the cytoplasmic membrane. It also seems reasonable to assume that the enzyme that catalyzes the cyclization reaction is specifically involved in the synthesis of ECA CYC. In this regard, the functions of essentially all of the genes in the wec gene cluster have been defined, and none of these genes appear to be specifically involved in the assembly of ECA CYC. Therefore, genetic determinants specifically involved in the synthesis of this polymer must be located at sites on the chromosome outside of this gene cluster; however, these genetic loci have not yet been identified.

The functions of ECA CYC and ECA PG are not known, and attempts to identify their functions would be greatly facilitated by the availability of mutants specifically defective in the synthesis of either of these molecules. However, as stated above, the identification of genetic determinants specifically involved in the assembly of ECA CYC has yet to be accomplished. In addition, the isolation of mutants specifically defective in the synthesis of ECA PG has also proven to be problematic. Accordingly, the wzyE and wzxE genes encode the polymerase and putative flippase involved in the assembly of linear ECA polysaccharide chains, respectively. Although mutations in these genes specifically abolish the synthesis of ECA PG, recent experiments have revealed that the use of such mutants to investigate the functions of ECA CYC and ECA PG is not feasible because mutations in these genes are deleterious to the cell; this appears to be due to toxicity resulting from the accumulation of lipid III (P. D. Rick, unpublished results). Furthermore, attempts to isolate mutants specifically defective in the synthesis of ECA PG due to the inability to transfer ECA polysaccharide chains to a diacylglyceride or phosphoglyceride acceptor have not yet been successful.

Despite these obstacles, the discovery that ECA CYC is synthesized by E. coli K-12 now affords a tractable experimental system that will greatly facilitate efforts to identify the function of this novel molecule as well as to define the genes and enzymes involved in its assembly.

ACKNOWLEDGMENTS

This research was supported by NIGMS grant GM52882 to P.D.R. and grants from the NIH (CA09061 and CA95471), Searle Scholars Program, Robert A. Welch Foundation (I-1424), and UT Southwestern Endowed Scholars Program to K.H.G. N.G. was supported by grants from the NIH (GM38545) and Robert A. Welch Foundation (I-1168) to Mark Lehrman (UT Southwestern Medical Center).

REFERENCES


ERRATUM

Identification and Biosynthesis of Cyclic Enterobacterial Common Antigen in *Escherichia coli*

Paul J. A. Erbel, Kathleen Barr, Ninguo Gao, Gerrit J. Gerwig, Paul D. Rick, and Kevin H. Gardner

*Department of Biochemistry and Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9038; Department of Microbiology and Immunology, F. Edward Hebert School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20814; and Department of Bio-Organic Chemistry, Bijvoet Center, Utrecht University, 3508 TB Utrecht, The Netherlands*


Page 1997, Table 1, strains PR4153, PR4164, and PR4161: The relevant genotypes should read “As PR4185. . .,” “As PR4186. . .,” and “As PR4186. . .,” respectively.