- 12. D. J. Reynolds, C. W. Penn, *Microbiology* **140**, 2649 (1994).
- 13. Materials and Methods are available as supporting material on *Science* Online.
- 14. J. W. Olson, R. J. Maier, data not shown.
- M. M. Zukowski et al., Proc. Natl. Acad. Sci. U.S.A. 80, 1101 (1983).
- 16. j.-F. Tomb et al., Nature 388, 539 (1997)
- R. E. McCrae, J. Hanus, H. J. Evans, *Biochem. Biophys. Res. Commun.* 80, 384 (1978).
- A. Salyers, D. D. Whitt, *Bacterial Pathogenesis: A Molecular Approach* (ASM Press, Washington, DC, 2002).
- 19. P. Doig et al., Microbiol. Mol. Biol. Rev. 63, 675 (1999).
- J. H. Bond, M. D. Levitt, J. Lab. Clin. Med. 85, 546 (1975).
- 21. G. E. Noether, Introduction to Statistics: A Fresh Approach (Houghton Mifflin, Boston, 1971).
- 22. This study was supported by the Georgia Research Alliance. We thank R. Seyler for cloning of the pro-

# N-Linked Glycosylation in *Campylobacter jejuni* and Its Functional Transfer into *E. coli*

Michael Wacker,<sup>1\*</sup> Dennis Linton,<sup>2\*</sup> Paul G. Hitchen,<sup>3</sup> Mihai Nita-Lazar,<sup>1</sup> Stuart M. Haslam,<sup>3</sup> Simon J. North,<sup>3</sup> Maria Panico,<sup>3</sup> Howard R. Morris,<sup>3,4</sup> Anne Dell,<sup>3</sup> Brendan W. Wren,<sup>2</sup> Markus Aebi<sup>1</sup>†

N-linked protein glycosylation is the most abundant posttranslation modification of secretory proteins in eukaryotes. A wide range of functions are attributed to glycan structures covalently linked to asparagine residues within the asparagine-X-serine/threonine consensus sequence (Asn-Xaa-Ser/Thr). We found an N-linked glycosylation system in the bacterium *Campylobacter jejuni* and demonstrate that a functional N-linked glycosylation pathway could be transferred into *Escherichia coli*. Although the bacterial N-glycan differs structurally from its eukaryotic counterparts, the cloning of a universal N-linked glycosylation cassette in *E. coli* opens up the possibility of engineering permutations of recombinant glycan structures for research and industrial applications.

Glycosylation has generally been considered to be restricted to eukaryotes where the attachment of glycan structures to proteins usually occurs at an Asn-Xaa-Ser/Thr consensus (Nlinked) or at Ser/Thr residues (O-linked). It is now evident that protein glycosylation is also abundant in prokaryotes (1, 2). N-linked protein glycosylation of S-layer proteins seems to be restricted to the archaeal domain, whereas serine-, threonine-, or tyrosine-linked (Olinked) glycosylation is predominantly found in bacteria. However, specific N-glycoproteins in bacteria have been reported (3).

N-linked glycosylation is the most frequent protein modification in eukaryotes. In the central step of the process that takes place at the luminal side of the endoplasmic reticulum

\*These authors contributed equally to this work. †To whom correspondence should be addressed. Email: aebi@micro.biol.ethz.ch (ER) membrane, a preassembled oligosaccharide (Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub>, where Glc is glucose, Man is mannose, and GlcNAc is N-acetylglucosamine) is transferred from the lipid carrier dolichyl pyrophosphate to asparagine residues of nascent polypeptide chains by the oligosaccharyltransferase (4). Such general glycosylation systems have not been described in prokaryotes, with the exception of the enteropathogenic bacterium Campylobacter jejuni, where the pgl gene cluster (Fig. 1A) seems to encode a general protein glycosylation system (5, 6). A number of pgl genes encode proteins with sequence similarity to glycosyltransferases and to enzymes required for sugar biosynthesis, a fact that supports this hypothesis. The pglB gene encodes a protein with strong similarity to Stt3p found exclusively in eukaryotes and in archaea but not in any other bacterial species (Fig. 1B). Genetic and biochemical studies in yeast have demonstrated that the Stt3 protein is an essential component of the oligosaccharyltransferase complex (7-9), the central enzyme in the process of N-linked protein glycosylation. In eukaryotes, the oligosaccharyltransferase complex consists of at least eight different subunits (8), yet, the precise catalytic mechanism of this enzyme is unknown (10-12).

moter-less xylE gene, M. Johnson for technical assistance in developing the XylE assay, S. Maier for assistance with the mouse colonization experiments, and A. Olzcak for assistance with mouse colonization and Southern blot experiments.

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/298/5599/1788/ DC1

Materials and Methods

8 August 2002; accepted 24 September 2002

we constructed a C. jejuni pglB mutant strain and probed the membrane protein extract with a polyclonal antiserum raised against C. jejuni whole-cell extracts (Fig. 2A) (13). The reactivity of this antiserum is markedly reduced by prior treatment of C. jejuni proteins with a deglycosylating agent, indicating that the glycan portion of C. jejuni glycoproteins is highly immunogenic (5). Mutation of the pglB gene also resulted in changes to the immunoreactivity of C. jejuni membrane proteins (Fig. 2A) suggesting that PglB was involved in the biosynthesis of these putative glycoproteins. This mutant phenotype was partially complemented by introducing the 16-kilobase pgl locus on a plasmid. We attributed the partial complementation of the pglB mutant phenotype to the instability of the large plasmid in C. jejuni. Our experiments confirmed a direct involvement of PglB in the generation of glycoproteins in C. jejuni (5). One of the immunoreactive proteins (arrow in Fig. 2A) was purified from the C. jejuni extract and identified by matrixassisted laser desorption/ionization (MALDI) mass mapping as the periplasmic AcrA (Cj0367c). An acrA deletion mutant of C. *jejuni* was analyzed by immunoblot analysis using the glycoprotein-specific antiserum. The 47-kD immunoreactive protein was absent in *acrA* mutant cells, and this phenotype was complemented by the expression of plasmid-borne acrA gene (Fig. 2A). Antiserum raised against recombinant C. jejuni AcrA produced in Escherichia coli revealed mature protein migrating at an apparent molecular size of 47 kD (Fig. 2B). In contrast to the glycoprotein-specific serum (Fig. 2A), an unglycosylated AcrA with increased mobility was recognized by the AcrA-specific antibodies in a *pglB* mutant strain (Fig. 2B). Partial complementation of the pglB phenotype was achieved with the pgl locus, as visualized by three distinct AcrA-specific bands. Thus, mature AcrA carries two oligosaccharide modifications, and partial complementation of the pglB mutation resulted in diglycosylated, monoglycosylated, and nonglycosylated AcrA protein. Only the two former forms were recognized by the glycosylation-specific serum. Therefore, AcrA was a target for the general glycosylation system and glycosylation required the pgl gene cluster, in particular PglB activity.

<sup>&</sup>lt;sup>1</sup>Institute of Microbiology, Department of Biology, Swiss Federal Institute of Technology, Zürich, CH-8092 Zürich, Switzerland. <sup>2</sup>Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. <sup>3</sup>Department of Biological Sciences, Imperial College of Science, Technology and Medicine, London SW7 2AY, UK. <sup>4</sup>M-SCAN Mass Spectrometry Research and Training Centre, Silwood Park, Ascot SL5 7PZ, UK.

### REPORTS

We reconstituted the C. jejuni glycosylation machinery in E. coli. acrA was overexpressed in E. coli in the presence of the complete C. jejuni pgl gene cluster (Fig. 3A). When the E. coli extracts were analyzed with the C. jejuni glycosylation-specific antiserum, reactive proteins were not detected in E. coli strains lacking the acrA expression construct or in strains containing the acrA expression construct but lacking the pgl gene cluster (Fig. 3B). However, expression of acrA in the presence of the pgl gene cluster led to the production of two immunoreactive proteins. Their production in E. coli was PglB dependent, because the C. jejuni pgl gene cluster expressing the  $Trp^{458} \rightarrow Ala^{458}$ ,  $Asp^{459} \rightarrow Ala^{459}$  (W458A, D459A) PglB mutant was unable to direct production of immunoreactive AcrA. The highly conserved amino acid motif  $[_{457}WWDYGY_{462}$  (Fig. 1B) (14)] is located on the hydrophilic COOH-terminal part of PglB oriented toward the periplasmic space (15). Thus, the *pgl* gene cluster was sufficient to direct AcrA glycosylation in E. coli, and PglB was essential for this process. To demonstrate that the pglmediated modification was indeed glycosylation, we tested the E. coli proteins for their affinity toward the lectin soybean agglutinin

(SBA) (6). We observed weak cross-reactivity of the lectin toward the strongly overproduced AcrA protein (Fig. 3C); however, the lectin also detected two AcrA proteins with a higher molecular weight, corresponding to the two modified forms visualized by the glycosylation-specific antiserum. Thus, the *pgl* gene cluster directed the synthesis of glycosylated AcrA in *E. coli* and was required for a general protein glycosylation system in *C. jejuni*. The involvement of the Stt3-like PglB protein indicated that this process was N-linked protein glycosylation involving the Asn-Xaa-Ser/Thr consensus acceptor sequence.

To verify this hypothesis, *C. jejuni*–derived PEB3 glycoprotein (6) and recombinant AcrA glycoprotein produced in *E. coli* were purified by affinity chromatography and digested with trypsin. The resulting glycopeptides were analyzed by mass spectrometry (MS). Doubly charged species in the MS spectrum of PEB3 not corresponding to trypsin autodigest products or unmodified PEB3 tryptic peptides were subjected to collisionally activated dissociation (CAD) MS-MS (*16*). Data indicative of a glycosylated peptide were obtained from analysis of the doubly charged species at a mass-to-charge ratio (m/z) of 1057.5<sup>2+</sup> (Fig. 4A). The



**Fig. 1.** The *C. jejuni* protein glycosylation locus and its role in protein glycosylation. (**A**) Schematic representation of the *C. jejuni* 81116 protein glycosylation locus. This locus is described in (22). Putative glycosyltransferases and enzymes involved in sugar biosynthesis are shaded dark and light, respectively. The *C. jejuni* STT3 homolog (*pglB*) is shaded black. For comparison, a cluster of *N. meningitidis* genes involved in pilin glycosylation (19) is presented below the *C. jejuni pgl* gene locus. Connecting lines indicate genes encoding proteins with strong amino acid sequence identity, as indicated. (**B**) Alignment of conserved residues from putative oligosaccharyltransferase homologs (14). Invariant residues are boxed. Arrows indicate residues altered in the *C. jejuni* PglB protein by site-directed mutagenesis (described herein). Bracketed letters A, B, or E preceding species names indicate that the corresponding species belongs to the archaeal, bacterial, or eukaryal domains, respectively.

series of singly charged fragment ions beginning at m/z 1911 and ending at m/z 709 was consistent with glycosidic cleavage products (17) from the tryptic peptide DFNVSK (14) carrying a glycan moiety of composition 2,4diacetamido-2,4,6-trideoxyhexose (DATDH), N-acetylhexosamine (HexNAc)<sub>5</sub>, hexose (Hex). Sufficient information was obtained to putatively assign the sequence and branching pattern of the glycan (Fig. 4A, inset). The peptide sequence was confirmed by CAD MS-MS using higher collision energies. The unusual sugar DATDH has been detected on the pili of Neisseria men-



Fig. 2. Analysis of membrane proteins of the wild type (WT), pglB and acrA mutants, and complementation in trans. Membrane proteins were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunodetected with (A) antiserum raised against C. jejuni or with (B) the antiserum raised against recombinant AcrA produced in E. coli. Lane 1, WT; lane 2, pglB mutant; lane 3, pglB mutant containing a plasmid that expressed PglB; lane 4, acrA mutant; lane 5, acrA mutant containing a plasmid expressing AcrA. Molecular size markers are indicated on the left. The arrow indicates the position of the AcrA glycoprotein which is absent in the mutant strains but present when complemented in trans by introduction of the specific plasmids.

*ingitidis* (18), and its putative assignment in *C. jejuni* was based on its mass increment in the CAD MS-MS spectrum.

When recombinant AcrA in the presence of the *pgl* locus was analyzed in the same way as PEB3, two doubly charged molecular ions (m/z 978.5<sup>2+</sup> and 966.0<sup>2+</sup>) were identified as putative glycopeptides. When subjected to CAD

pET24b

Α

coli BL21 membranes

pET24(AcrA)

o(W458A, D459A) p(W458A, D459A pACYC(pgf) pACYC184 pACYC184 175 83 62 47.5 32.5 kD 2 3 4 5 6 **B** <sub>175</sub> 83 62 47 5 32.5 kD 2 3 4 5 6 1 С 175 83 62 47.5 32.5 kD 2 1 3 4 5

Fig. 3. Analysis of membrane proteins of E. coli BL21(DE3) overexpressing AcrA in the presence or absence of the C. jejuni pgl locus. Membrane proteins were separated by SDS-PAGE and either stained with (A) Coomassie Blue or, after transfer to nitrocellulose, immunodetected with (B) the C. *jejuni* antiserum or probed with (C) SBA lectin. Lanes 1 to 3, cells containing the plasmid backbone (pET24b) only; lanes 4 to 6, cells overexpressing AcrA of C. jejuni. As a second plasmid, these cells contain vector pACYC184 (lanes 1 and 4), pACYC184 with the pgl locus (lanes 2 and 5), or pACYC184 with the pgl locus expressing PglB with amino acid substitutions (W458A and Y459A) in the highly conserved domain (lanes 3 and 6). Molecular size markers are indicated on the left.

MS-MS, the doubly charged species at m/z 978.5<sup>2+</sup> (Fig. 4B) gave a fragmentation pattern that was remarkably similar to the PEB3 gly-copeptide (Fig. 4A), suggesting that the glyco-sylation cassette was capable of synthesizing the full glycan structure as seen on native *C. jejuni* protein. Furthermore, the mass and fragmentation of the peptide component of the AcrA glycopeptide was consistent with the sequence DFNR, which is the predicted

product of tryptic cleavage at  $Arg^{124}$  within the Asn-Arg-Ser consensus sequence at residues 123 to 125. CAD MS-MS analysis of m/z 966.0<sup>2+</sup> (Fig. 4C) showed a very similar pattern to that for m/z 978.5<sup>2+</sup>, except that all major fragment ions were shifted to lower mass by 25 daltons, which represented the difference in mass of the molecular ions. The 25-dalton increment was also the difference between a DATDH



**Fig. 4.** Tandem mass spectrometry of glycosylated peptides derived from tryptic cleavage of PEB3 and AcrA. (**A**) CAD MS-MS spectrum of  $[M+2H]^{2+} m/z$  1057.5 derived from PEB3. (**B**) CAD MS-MS spectrum of  $[M+2H]^{2+} m/z$  978.5 derived from AcrA. (**C**) CAD MS-MS spectrum of  $[M+2H]^{2+} m/z$  966.0 derived from AcrA. Fragmentation patterns for each glycan are shown in the insets (top right of each panel) and are labeled using the nomenclature as described (23). The signals at (A) m/z 1749, (B) 1591, and (C) 1566 are attributable to loss of HexNAc plus Hex from the corresponding molecular ions.

and a HexNAc residue, suggesting that the glycan in the  $m/2 966.0^{2+}$  glycopeptide was linked to the peptide through HexNAc rather than DATDH (Fig. 4C, inset). In view of the DATDH synthesis pathway (19), it is likely that the HexNAc-linked oligosaccharide represents a precursor of the DATDH oligosaccharide (fig. S1).

A number of C. jejuni N-linked glycoproteins, whose glycosylation is dependent on the pgl gene cluster, have been identified (6). These glycoproteins have diverse functions in the cell but are all located extracytoplasmically. This is in agreement with the periplasmic space being the proposed site of PglB-mediated glycosylation in C. jejuni. The composition of the oligosaccharide did not show similarities to the highly conserved Glc<sub>3</sub>Man<sub>o</sub>GlcNAc<sub>2</sub> oligosaccharide that is transferred to protein in the ER of eukaryotes. However, as in eukaryotes, the protein-bound oligosaccharide in C. jejuni was linked to the amide nitrogen of asparagine within the consensus sequence Asn-Xaa-Ser/Thr. This sequence is essential for N-linked protein glycosylation in eukaryotes, and it has been proposed that it forms a specific secondary structure, the Asn-turn, recognized by the oligosaccharyltransferase (10). The presence of this consensus sequence on glycoproteins of C. jejuni suggests that the reaction mechanisms in eukaryotes and in C. jejuni are very similar. It is proposed that PglB fulfills the oligosaccharyltransferase function in the prokaryote and that Stt3 protein represents the catalytic subunit of the eukaryotic oligosaccharyltransferase complex. With the bipartite nature of the eukaryotic N-glycosylation pathway in mind, we speculate that in C. jejuni the heptasaccharide is assembled in the cytoplasm on a lipid carrier, most likely bactoprenyl pyrophosphate, translocated to the periplasmic side of the membrane, and transferred to protein (fig. S1).

To our knowledge, a general N-glycosylation system very similar to the one found in eukaryotes has not been described in other bacteria, and the C. jejuni genome is the only bacterial genome sequenced to date that harbors a gene that encodes a protein with strong sequence homology to a eukaryotic oligosaccharyltansferase component. The unique biosynthetic potential of C. jejuni and the isolated appearance of such an oligosaccharyltransferase ortholog in this organism led us to conclude that lateral gene transfer (20) of the oligosaccharyltransferase gene either from the archaeal or the eukaryal domain is responsible for this unique biosynthetic property of C. *jejuni*.

*Note added in proof*: The structure of the N-linked glycan of *C. jejuni* has recently been published (21).

### **References and Notes**

- 1. C. Schäffer, M. Graninger, P. Messner, Proteomics 1, 248 (2001).
- 2. I. Benz, M. A. Schmidt, Mol. Microbiol. 45, 267 (2002).

- P. R. Erickson, M. C. Herzberg, J. Biol. Chem. 268, 23780 (1993).
- 4. R. Kornfeld, S. Kornfeld, Annu. Rev. Biochem. 54, 631 (1985).
- 5. C. M. Szymanski, R. Yao, C. P. Ewing, T. J. Trust, P. Guerry, *Mol. Microbiol.* **32**, 1022 (1999).
- D. Linton, E. Allan, A. V. Karlyshev, A. D. Cronshaw, B. W. Wren, *Mol. Microbiol.* 43, 497 (2002).
- 7. R. Zufferey et al., EMBO J. 14, 4949 (1995).
- R. Knauer, L. Lehle, Biochim. Biophys. Acta 1426, 259 (1999).
- 9. S. Silberstein, R. Gilmore, *FASEB J.* **10**, 849 (1996). 10. B. Imperiali, K. L. Shannon, M. Unno, K. W. Rickert,
- J. Am. Chem. Soc. 114, 7944 (1992).
- E. Bause, G. Legler, *Biochem. J.* **195**, 639 (1981).
  D. Karaoglu, D. J. Kelleher, R. Gilmore, *Biochemistry* **40**, 12193 (2001).
- 13. See Supporting Online Material on Science Online.
- 14. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. X indicates any residue.
- 15. P. Kotyczka, M. Wacker, M. Aebi, unpublished data.
- H. R. Morris et al., Rapid Commun. Mass Spectrom. 10, 889 (1996).

- 17. P. Teng-umnuay et al., J. Biol. Chem. 273, 18242 (1998).
- 18. E. Stimson et al., Mol. Microbiol. 17, 1201 (1995).
- 19. P. M. Power et al., Microbiology 146, 967 (2000).
- 20. W. F. Doolittle, Trends Cell Biol. 9, M5 (1999).
- 21. N. M. Young et al., J. Biol. Chem. 227, 42530 (2002).
- B. N. Fry et al., Microbiology 144, 2049 (1998).
  B. Domon, C. E. Costello, *Glycoconjugate J.* 5, 397
- (1988). 24 We there 9. Come for an idiation with C initia
- 24. We thank P. Guerry for providing us with C. *jejuni* 81-176 strain and plasmid containing AcrA, B. N. Fry for plasmid pBTLPS, and D. G. Newell for providing antibodies raised against C. *jejuni* whole-cell extracts. Supported by grants from the Swiss National Science Foundation to M.A. and by grants from the UK Biotechnology and Biological Sciences Research Council and the Wellcome Trust to B.W.W., H.R.M., and A.D.

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/298/5599/1790/ DC1

Materials and Methods

Fig. S1 References and Notes

24 July 2002; accepted 17 September 2002

# Structural Adaptations in a Membrane Enzyme That Terminates Endocannabinoid Signaling

Michael H. Bracey,<sup>1\*</sup> Michael A. Hanson,<sup>2\*</sup> Kim R. Masuda,<sup>1</sup> Raymond C. Stevens,<sup>2,3</sup>† Benjamin F. Cravatt<sup>1,2,4</sup>†

Cellular communication in the nervous system is mediated by chemical messengers that include amino acids, monoamines, peptide hormones, and lipids. An interesting question is how neurons regulate signals that are transmitted by membrane-embedded lipids. Here, we report the 2.8 angstrom crystal structure of the integral membrane protein fatty acid amide hydrolase (FAAH), an enzyme that degrades members of the endocannabinoid class of signaling lipids and terminates their activity. The structure of FAAH complexed with an arachidonyl inhibitor reveals how a set of discrete structural alterations allows this enzyme, in contrast to soluble hydrolases of the same family, to integrate into cell membranes and establish direct access to the bilayer from its active site.

Several amidated lipids serve as endogenous signaling molecules in the mammalian central nervous system, modulating behaviors such as pain, cognition, feeding, sleep, and locomotor activity (1–5). A subset of these, including anandamide (1), bind to the central cannabinoid receptor CB1, a heterotrimeric GTP-binding protein (G protein)-coupled receptor activated by the psychotropic component of marijuana,  $\Delta^9$ -THC (6). Additional classes of lipid transmitters have also emerged as regulators of nervous system function (7). Nonetheless, the molecular mechanisms for regulating lipid-based signal-

ing events remain largely unknown. Although classical neurotransmitters can be confined spatially by membrane-delineated compartments for controlled storage, release, and uptake, lipid messengers are not easily contained by such physical boundaries. Because lipids can readily partition to and diffuse throughout cell membranes, termination of their signaling capacity may rely primarily on chemical transformation, possibly within the bilayer itself, instead of on cellular uptake and compartmentalization.

The amplitude and duration of fatty acid amide signals are regulated in vivo primarily by a single degradative enzyme, the integral membrane protein fatty acid amide hydrolase (FAAH) (8). Transgenic mice lacking FAAH possess high endogenous concentrations of anandamide and related fatty acid amides in the brain that correlate with increased CB1dependent analgesia in these animals (9), suggesting that FAAH may represent an at-

Departments of <sup>1</sup>Cell Biology, <sup>2</sup>Chemistry, and <sup>3</sup>Molecular Biology, <sup>4</sup>Skaggs Institute for Chemical Biology, Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.

<sup>\*</sup>These authors contributed equally to this work. †To whom correspondence should be addressed. Email: cravatt@scripps.edu, stevens@scripps.edu

www.sciencemag.org SCIENCE VOL 298 29 NOVEMBER 2002