Protein signatures distinctive of chlamydial species: horizontal transfers of cell wall biosynthesis genes *glmU* from archaea to chlamydiae and *murA* between chlamydiae and *Streptomyces*

Emma Griffiths and Radhey S. Gupta

Author for correspondence: Radhey S. Gupta. Tel: +1 905 525 9140 ext. 22639. Fax: +1 905 522 9033.
e-mail: gupta@mcmaster.ca

Department of Biochemistry,
McMaster University,
Hamilton, Ontario,
Canada L8N 3Z5

Chlamydiae are major human and animal pathogens. Based on alignments of different protein sequences, a number of conserved indels (insertion/deletions) were identified that appear to be unique and distinctive characteristics of the chlamydial species. The identified signatures include one 16 aa and two single aa inserts in the enzyme UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase (MurA), a 1 aa insert in protein synthesis elongation factor P (EF-P), a 1 aa insert in the Mg²⁺ transport protein (MgtE), a 1 aa insert in the carboxy-terminal protease and a 1 aa deletion in the tRNA (guanine-*N*¹)-methyltransferase (TrmD) protein. The homologues of these proteins are found in all major groups of bacteria and the observed indels are present in all available chlamydial sequences but not in any other species (except for the large insert in MurA in *Streptomyces*). The validity of three of these signatures (MurA, EF-P and MgtE) was tested by PCR amplifying the signature regions from several chlamydial species for which no sequence information was available. All *Chlamydiaceae* species for which specific fragments could be amplified (*Chlamydia suis*, *Chlamydophila abortus*, *Chlamydophila psittaci*, *Chlamydophila felis*) contained the expected signatures. Additionally, a fragment of the *murA* gene from *Waddlia chondrophila* and the *efp* gene from *Simkania negevensis*, two chlamydia-like species, were also cloned and sequenced. The presence of respective indels in these species provides strong evidence that they are specifically related to the traditional chlamydial species, and that these signatures may be distinctive of the entire *Chlamydiales* order. A 17 aa conserved indel was also identified in the cell wall biosynthesis enzyme UDP-*N*-acetylglucosamine pyrophosphorylase (GlmU), which is shared by all archaeal and chlamydial homologues. The gene for this protein is indicated to have been horizontally transferred from an archaeon to a common ancestor of the chlamydiae. The results also support a lateral transfer of the *murA* gene between chlamydiae and *Streptomyces*. The large inserts in these peptidoglycan synthesis related genes in chlamydiae could account for their unusual cell-wall characteristics. These signatures are also potentially useful for screening of the chlamydiae species.

Keywords: chlamydia-like organisms, lateral gene transfer, *Archaea*, cell-wall biosynthesis

Abbreviations: indel, insertion/deletion.

The GenBank accession numbers for the sequences reported in this paper are indicated in the text.

INTRODUCTION

Because of their widespread pathogenicity to human and animals, chlamydial species are a major health concern. *Chlamydia trachomatis* in humans is the major causative agent of sexually transmitted genitourinary infections, ocular trachoma, neonatal pneumonitis and the systemic disease lymphogranuloma venereum (Fields & Barnes, 1992; Schachter, 1999). *Chlamydophila pneumoniae* infections are responsible for several respiratory diseases including bronchitis and sinusitis, about 10–15% cases of community-acquired pneumonia (Pettersson *et al.*, 1997), and it has also been implicated in many other diseases including atherosclerosis and increased susceptibility to HIV infection (Moulder *et al.*, 1984; Saikku, 2000; Daian *et al.*, 2000; Laga *et al.*, 1991; Kuo *et al.*, 1995).

We have recently described a new approach based on shared conserved insertion and deletions (indels) in various proteins that has proven very useful in identifying main groups within the domain *Bacteria* and in understanding their relationship to each other (Gupta, 1998, 2000a, b). By tracking the presence or absence of specific indels in various proteins in different phyla, this approach allows the logical deduction of the relative branching order of different groups from a common ancestor (Gupta, 1998, 2000b, 2001; Griffiths & Gupta, 2001). The use of this approach indicates that the *Chlamydiaceae* branch is in a similar position to the large, diverse *Cytophaga*–*Flavobacterium*–*Bacteroides* group which has been placed between the *Spirochaetes* and the δ,ϵ -*Proteobacteria* (Gupta, 2000b, 2001). The present work describes a number of conserved indels in various proteins which provide specific molecular markers for the chlamydial group of species, and can unambiguously define and identify this group from all other groups of bacteria. It is significant that two of the largest indels identified in the present work on chlamydiae (in *MurA* and *GlmU*) are in enzymes involved in cell wall peptidoglycan biosynthesis, which could provide insights into the unusual cell wall characteristics of these organisms.

METHODS

Identification of signature sequences. Multiple sequence alignments for homologues of different proteins were constructed using the ALIGN PLUS 4 program (Scientific and Educational Software) as described in earlier work (Gupta & Johari, 1998). Chlamydial signatures were identified in these alignments by visual inspection, where a conserved indel was uniquely present in this group of species. To qualify as a useful signature, any identified indel was required to be flanked on both sides by regions of high sequence conservation, ensuring that the observed indel was not due to sequencing errors or alignment artefacts (Gupta, 1998, 2000b).

PCR amplification and sequencing. DNA from *Chlamydiaceae* strains (*Chlamydophila felis* FP Cello; *Chlamydophila abortus* EBA (EP12); *Chlamydophila psittaci* MN (ATCC VR 122); and *Chlamydia suis* R24) was generously made available to us by Dr K. Everett (University of Georgia, Athens, USA), and genomic DNA from *Simkania negevensis* (ATCC VR1471) and *Waddlia chondrophila* (ATCC 1470) was kindly

provided by Dr A. Petrich (St. Joseph's Hospital, Hamilton, Canada).

Oligonucleotide primers, in opposite orientations, were designed for regions that flanked three of the signatures, based on the sequences of the genes/proteins from available species of *Chlamydiaceae*. Degeneracy was incorporated into the primers to account for nucleotide variability at different sites in the alignment. The primers were synthesized at MOBIX, McMaster University.

PCR was performed in a Techne Techgene thermocycler. Each reaction had a final volume of 10 μ l and each primer set was optimized for Mg^{2+} concentration (in the range of 1.5–4 mM) for each DNA strain tested. PCR amplification was carried out over 30 cycles (15 s at 94 °C, 15 s at 55 °C, 1 min at 72 °C) with an initial 1 min hot start at 94 °C and a final extension step (15 s at 94 °C, 15 s at 55 °C, 7 min at 72 °C). DNA fragments of the expected size were purified from 0.8% (w/v) agarose gels (using a GeneClean kit), and subcloned into the plasmid pCR2.1-TOPO using a TA cloning kit (Invitrogen). *Escherichia coli* JM109 cells were transformed with the ligated plasmid, and the inserts from a number of positive clones were sequenced by the dideoxy chain termination method using a T7 sequencing kit (Pharmacia). Sequences of all cloned fragments were run through a BLAST search to ensure that they were from a novel source. Attempts were made to generate *mgtE*, *efp* and *murA* fragments corresponding to signature regions in *Chl. suis*, *Chlam. psittaci*, *Chlam. abortus*, *Chlam. felis*, *S. negevensis* and *W. chondrophila* genes for which sequence data were not known. However, amplification was not successful in all cases. Due to the small quantity of available genomic DNA, different primer sets for PCR amplification could not be attempted, though they may prove useful in future studies. The primer sequences used for amplification of different genes were as follows.

UDP-N-acetylglucosamine 1-carboxyvinyltransferase (*MurA*). The forward and reverse primers used to amplify this gene were: 5'-CCAGATAARATYGARGCDGCGYGGWATGGCYGCRGTWGT-3' and 5'-GCAATVAGWGCKGCCATRACATARGCAAAYCCYGCDGCGYAAATC-3', R represents A or G; Y represents C or T; D represents G, A or T; W represents A or T; V represents G, C or A; K represents G or T; M represents A or C; S represents G or C. This primer set was successfully used to amplify approximately 0.5 kb fragments of the *murA* gene from *Chlam. abortus*, *Chlam. psittaci* and *Chlam. felis* genomic DNA. A 690 bp fragment from *W. chondrophila* was generated using a forward primer for the conserved amino acid sequence VGATEN (5'-GTNGGNGCNACNGARAA-3') and a reverse primer for the amino acid sequence AGFAYVMA (5'-GCCATNACRTANGCR-AANCCNGC-3'), where N is A, C, G or T. *murA* sequence data were submitted to GenBank under the accession numbers AY038586, AY038585, AY038587 and AF468694 for *Chlam. abortus*, *Chlam. psittaci*, *Chlam. felis* and *W. chondrophila* respectively.

Translation elongation factor P (EF-P) protein. The forward and reverse primers for this gene were: 5'-TTRMGMRTWRARATYATGG-3' and 5'-WACDCGRGASTCRTARCTYC-3'. These primers successfully amplified an approximately 0.5 kb fragment of the *efp* gene from *Chl. suis* and *Chlam. abortus*. The primers used to amplify an approximately 0.4 kb fragment from *S. negevensis* genomic DNA were based on the conserved amino acid sequences VKPGKG (forward primer: 5'-GTNARCCNGGNAARGG-3') and TGAKIMVP (reverse primer: 5'-GGNACCATDATYTTNGCNCCNGT-3'). Sequence information was submitted to GenBank under the accession

numbers AY038589, AY038588 and AF468693 for *Chl. suis*, *Chlam. abortus* and *S. negevensis* respectively.

Mg²⁺ transport (MgtE) protein. A 0.5 kb fragment from *Chl. suis* genomic DNA was amplified using the following forward and reverse primers: 5'-GTTKCBRCYTGYATWCGMARTAA-TCCTGGSRTTGA-3' and 5'-GTWGCCATACTMCGVAC-WARAATHGTGCT-3', where B is G, C or T; and H is A, C or T. Sequence data for this gene fragment were submitted to GenBank under the accession number AY038584.

Phylogenetic analysis. Phylogenetic analyses based on MurA or GlmU protein sequences were carried out as described in our earlier work (Gupta & Singh, 1994; Gupta *et al.*, 1997). The aligned sequences were analysed using the programs SEQBOOT, PROTDIST, NEIGHBOR and CONSENSE from the PHYLIP 3.5 program software (Felsenstein, 1994). The sequences for these proteins were analysed both with and without the large inserts found in these proteins to exclude the effects of these regions on the observed topology.

RESULTS

Description of chlamydia-specific signatures in different proteins

Our work on alignment of different protein sequences has led to the identification of a number of conserved indels (referred to as signatures) that appear to be specific for the chlamydial group of species. Several proteins [elongation factor P (EF-P), UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA), a Mg²⁺ transport protein (MgtE) and a carboxyl terminal protease (Ctp)], encoded by the genomes of the four different chlamydia species (*Chl. trachomatis*, *Chlam. pneumoniae*, *Chlamydomophila caviae* and *Chlamydia muridarum*) whose sequences were available in the NCBI microbial database were found to contain single amino acid inserts within conserved regions that were not found in other bacteria. Using the PCR primers described in Methods, we confirmed that these inserts were also present in the EF-P of *Chlam. abortus* and *Chl. suis* and in MgtE of *Chl. suis* (accession numbers AY038589, AY038588 and AY038584). We also noted a 1 aa chlamydia-specific deletion in a conserved region of the tRNA (guanine-N¹)-methyltransferase (TrmD) protein (not shown). The sequence alignments showing these signatures are included as supplementary material available at <http://mic.sgmjournals.org>. In addition to these single amino acid indels we have also identified a 16 aa insertion signature in the MurA of *Chl. trachomatis*, *Chlam. pneumoniae*, *Chlam. caviae* and *Chl. muridarum* (Fig. 1). Using the PCR primers described in Methods, we have cloned this sequence region from *Chlam. abortus*, *Chlam. psittaci* and *Chlam. felis*; all of these sequences contained this signature (Fig. 1). Interestingly, MurA of two *Streptomyces* species also possess a similar but not identical 16 aa insertion in the same location (Fig. 1).

Presence of the identified signatures in chlamydia-like organisms

In addition to the traditional chlamydial species, which are placed in the family *Chlamydiaceae*, a number of

other organisms that show chlamydia-like replication cycles have been identified (Everett *et al.*, 1999; Kahane *et al.*, 1999; Ossewaarde & Meijer, 1999; Rurangirwa *et al.*, 1999). Based on the 16S and 23S rRNA phylogenies, these organisms have been placed in three separate families: *Parachlamydiaceae*, *Waddliaceae* and *Simkaniaceae*, within the order *Chlamydiales* (Everett *et al.*, 1999; Rurangirwa *et al.*, 1999). Besides 16S and 23S rRNA sequence data, very little or no sequence information is available for these species. Hence, it was of much interest to determine whether any of the identified signatures were present in these organisms. Using DNA for two of the chlamydia-related organisms, *W. chondrophila* and *S. negevensis*, we initially attempted to amplify the gene fragments for MurA, EF-P and MgtE, employing oligonucleotide primers that were successful in amplifying fragments from some of the *Chlamydiaceae* species. However, such attempts did not lead to amplification of any specific fragment. We reasoned that the failure of these primers was perhaps due to the fact that these species are more distantly related to the *Chlamydiaceae* species. We therefore designed degenerate primers based on amino acid sequence regions that were conserved among distantly related bacteria. Using these new sets of primers, we were successful in amplifying a 691 bp fragment of the *murA* gene from *W. chondrophila* and a 414 bp fragment of the *efp* gene from *S. negevensis*. These fragments were cloned, and their nucleotide and translated amino acid sequences deposited in GenBank under the accession numbers AY038587 and AF468693, respectively. The partial sequence for MurA from *W. chondrophila* is shown in Fig. 1 and that for EF-P from *S. negevensis* is included in the supplementary material available at <http://mic.sgmjournals.org>. Both these sequences from chlamydia-like organisms contained the expected signatures. These results provide evidence that these organisms are specifically related to the *Chlamydiaceae* and that the signatures described are very likely distinctive of the entire *Chlamydiales* order.

We have also performed phylogenetic analyses based on an alignment of MurA protein sequences. A consensus neighbour-joining phylogenetic tree based on 100 bootstrap replicates of the MurA protein sequences is presented in Fig. 2. All *Chlamydiaceae* species grouped together in this tree with a bootstrap confidence score of 95 out of 100. Interestingly, and as expected based on 16S and 23S rRNA trees (Everett *et al.*, 1999; Bush & Everett, 2001), *W. chondrophila* formed an outgroup of the *Chlamydiaceae* family and the clade consisting of these species was recovered 99 times out of 100 in bootstrap replicates. The association of *W. chondrophila* with the *Chlamydiaceae* was not dependent upon the large common insert shared by these species, as omission of this region from the alignment did not affect the bootstrap score of the node leading to these species. These results provide strong evidence that *W. chondrophila* is specifically related to the *Chlamydiaceae*. The phylogenetic trees based on MurA sequences also supported a strong and specific relationship of the

Proteobacteria	H. influenzae Rd	1171069	327	ETIFENRFMHIPELIRMGKAEI	374	EGNTAVCHGVEQLSGTEVIATDLRA
	Enter. cloacae	417329	-----V-----AH---	-----V-----AH---	-----S-----K-----AQ-M-----	-----S-----K-----AQ-M-----
	E. coli	4972324	--V-----V-----S---AH---	--V-----V-----S---AH---	--S-----VI-----K-----AQ-M-----	--S-----VI-----K-----AQ-M-----
	Sal. enterica	CAD07824	--V-----V-----S---AR---	--V-----V-----S---AR---	--S-----VI-----T-----AQ-M-----	--S-----VI-----T-----AQ-M-----
	Sal. typhimurium	AAL22176	--V-----V-----S---AR---	--V-----V-----S---AR---	--S-----VI-----T-----AQ-M-----	--S-----VI-----T-----AQ-M-----
	V. cholerae	AE003852	-----V-----K-----A----	-----V-----K-----A----	-----VI-GD-DR-----AQ-M-----	-----VI-GD-DR-----AQ-M-----
	Y. pestis	CAC92798	-----V-----AH-----	-----V-----AH-----	--S-----VI-Y-----AQ-M-----	--S-----VI-Y-----AQ-M-----
	X. fastidiosa	AAF84224	-----VD---L-L-A-IQ---	-----VD---L-L-A-IQ---	--H-IVQ---R-----AP-M-----	--H-IVQ---R-----AP-M-----
	Buch. sp. APS	BAB13089	-----IYTS-----A-IK---	-----IYTS-----A-IK---	KN-II-Y-IPK-ISSN-FSS---	KN-II-Y-IPK-ISSN-FSS---
	Bru. melitensis	AAL52847	-----VQ---A-L-A-ISL---	-----VQ---A-L-A-ISL---	S-Q-TVE---R-K-AQ-M-----	S-Q-TVE---R-K-AQ-M-----
	Pas. multocida	P57821	-----V-----AH-----	-----V-----AH-----	-----I-----DH-----A-M-----	-----I-----DH-----A-M-----
	Pse. aeruginosa	AE004859	--V-----VY-MN---AQILV	--V-----VY-MN---AQILV	-----IVT---PK-K-AP-M-----	-----IVT---PK-K-AP-M-----
	Pse. putida	4336804	-----VY-MH---AQIQV	-----VY-MH---AQIQV	-----IVT---K-KALKVPGNGHLR	-----IVT---K-KALKVPGNGHLR
	Acin. calcoaceticus	462661	-----V-----A-----ANIQV	-----V-----A-----ANIQV	--D-VT---K-----AAP-M-----	--D-VT---K-----AAP-M-----
	Nei. meningitidis	AL162759	-----V-----N-----ANITT	-----V-----N-----ANITT	-----FVQ---R-----AV-K-----	-----FVQ---R-----AV-K-----
	Ral. solanacearum	CAD16660	-----VQ---N-L-ANIIT	-----VQ---N-L-ANIIT	-----AVT-----AT-M-----	-----AVT-----AT-M-----
	Meso. loti	BA000012	-----V-----M-L-ANIKL	-----V-----M-L-ANIKL	Q-TM-LVR-G-K-H-AQ-M-----	Q-TM-LVR-G-K-H-AQ-M-----
	Ri. prowazekii	3861126	-N-----V-----C-----ADIVV	-N-----V-----C-----ADIVV	R--K--VR---M-K-A-M-S---	R--K--VR---M-K-A-M-S---
	Ri. conorii	AAL03421	-N-----V-----C-----ADIVV	-N-----V-----C-----ADIVV	R--K--VR---M-K-A-M-S---	R--K--VR---M-K-A-M-S---
	A. tumefaciens	AAL41557	-----VQ---A-L-A-ISL---	-----VQ---A-L-A-ISL---	S-QM-RIE---TR-K-AP-M-----	S-QM-RIE---TR-K-AP-M-----
	Ca. crescentus	AAK24321	-----A-----M-L-ADISV	-----A-----M-L-ADISV	S-GE-RVR--D--E-AQ-M-----	S-GE-RVR--D--E-AQ-M-----
	Sin. meliloti	CAC45181	-----VQ---A-L-A-ISL---	-----VQ---A-L-A-ISL---	S-Q-KVE---SR-K-AP-M-----	S-Q-KVE---SR-K-AP-M-----
	Myx. xanthus	AAK64441	-NI-----V-----H-L-ADIT-	-NI-----V-----H-L-ADIT-	Q-P--VK--KG---AP-M-----	Q-P--VK--KG---AP-M-----
	Camp. jejuni	CAB73123	-RL-----VS---L-----ADIKL	-RL-----VS---L-----ADIKL	N-HI-TIV-GKE-NAAD-M-----	N-HI-TIV-GKE-NAAD-M-----
	Hel. pylori 26695	2829458	-L-----AS---Q-L-ANISL	-L-----AS---Q-L-ANISL	KT-V-TIS-STE-T-SD-M-----	KT-V-TIS-STE-T-SD-M-----
Chlamydiaceae and related species	Chl. trachomatis	4887172	--VH---LGYLKG-VK--AHCDL	--VH---LGYLKG-VK--AHCDL	FHECLSAKSCRYSTGN	FHECLSAKSCRYSTGN
	Chl. muridarum	Q9PJ77	--VH---LGYLSG-AK--AHCDL	--VH---LGYLSG-AK--AHCDL	FHECLSAKSCRYSTGN	FHECLSAKSCRYSTGN
	Chlam. caviae	*	--VH---LGYLRG-QQ--ASC-L	--VH---LGYLRG-QQ--ASC-L	FYQCLSSKACRYATGN	FYQCLSSKACRYATGN
	Chlam. pneumoniae	4376864	--VH---LGYLHG-QH--AECOL	--VH---LGYLHG-QH--AECOL	FHQCLSTKACRYAIGN	FHQCLSTKACRYAIGN
	Chlam. abortus	AY038586	--VH---LGYLRG-QK--ANC-L	--VH---LGYLRG-QK--ANC-L	FYQCLSSKACRYATGN	FYQCLSSKACRYATGN
	Chlam. psittaci	AY038585	--VH---LGYLRG-QQ--ANC-L	--VH---LGYLRG-QQ--ANC-L	FYQCLSSKACRYATGN	FYQCLSSKACRYATGN
Streptomyces	Chlam. felis	AY038587	--VH---LGYLRG-QK--ANC-L	--VH---LGYLRG-QK--ANC-L	FYQCLSSKACRYATGN	FYQCLSSKACRYATGN
	Wad. chondrophila	AF468694	--VY---GYTDT-KE--AEITL	--VY---GYTDT-KE--AEITL	FRQCLGGKECRFSSQA	FRQCLGGKECRFSSQA
	Str. lividans	BAA85335	--VY-S-LGFTSA-NQ--AHQL	--VY-S-LGFTSA-NQ--AHQL	YRECLGGSDCRFGQRN	YRECLGGSDCRFGQRN
Other Gram (-)ve Bacteria	Str. coelicolor	CAB72195	--VY-S-LGFTSA-NQ--AHQL	--VY-S-LGFTSA-NQ--AHQL	YRECLGGSDCRFGQRN	YRECLGGSDCRFGQRN
	Aqu. aeolicus	2983705	-N---H--H-AQ--N-L-ANITV	-N---H--H-AQ--N-L-ANITV		
	Cb. tepidum	*	DR-YLE--N-----N-L-AHI--	DR-YLE--N-----N-L-AHI--		
	Fibro. succinogenes	*	D-VYND--K-VA---E-L-ASIQ	D-VYND--K-VA---E-L-ASIQ		
	Nostoc sp. PCC 7120	BAB77689	-SV---LRHAS--N-L-ADIRV	-SV---LRHAS--N-L-ADIRV		
	Sy. sp. PCC6803	2833465	--V---LQ-VA--Q---ANIKL	--V---LQ-VA--Q---ANIKL		
	Tre. pallidum	2780178	-KM--S-MFFVDK--T--ARIILC	-KM--S-MFFVDK--T--ARIILC		
	Bor. burgdorferi	2688372	-KM--S-MFFVDK--K--ARIVLC	-KM--S-MFFVDK--K--ARIVLC		
	D. radiodurans	AAF10696	DPVYPD-LT-VA--H---ATITV	DPVYPD-LT-VA--H---ATITV		
	T. maritima	4980599	-NV-KT--L-VD--K---ADI-V	-NV-KT--L-VD--K---ADI-V		
Gram (+)ve Bacteria	Bac. subtilis(1)	1648861	--V---AE-FR--N-DIK-	--V---AE-FR--N-DIK-		
	Bac. subtilis(2)	CAB15738	D--YSA--K--D--R---ANMKV	D--YSA--K--D--R---ANMKV		
	Bac. halodurans(1)	BAB07503	D--YNA--K--D--R---ADIKV	D--YNA--K--D--R---ADIKV		
	Bac. halodurans(2)	BAB07468	--V---VE-FR--N-NIK-	--V---VE-FR--N-NIK-		
	Des. thermocisternum	AAC96104	-----KVAD--K---ARIKV	-----KVAD--K---ARIKV		
	Clo. acetobutylicum(1)	AAK80805	-----V---K-F-ANIK-	-----V---K-F-ANIK-		
	Clo. acetobutylicum(2)	AAK81464	-S-W-S-LK-VD--KK--ANIKV	-S-W-S-LK-VD--KK--ANIKV		
	Sta. aureus N315(1)	BAB43186	--V---VA-FK--NANINV	--V---VA-FK--NANINV		
	Sta. aureus(2)	BAB43210	D--YPE--K-VE--K---ANI-V	D--YPE--K-VE--K---ANI-V		
	L. lactis(1)	AE005176	--V---Q-LK-MR-I-LEVD-	--V---Q-LK-MR-I-LEVD-		
	L. lactis(2)	AAK04410	D--YQK-VN-V---A---ANISV	D--YQK-VN-V---A---ANISV		
	Lis. monocytogenes(1)	CAD00604	--V-----VE-MR--NADMK-	--V-----VE-MR--NADMK-		
	Lis. monocytogenes(2)	AAF86297	D--YPS--K--A-IE-----FKL	D--YPS--K--A-IE-----FKL		
	Lis. innocua(1)	CAC97896	--V-----VE-MR--NADMK-	--V-----VE-MR--NADMK-		
	Lis. innocua(2)	CAC97923	D--YPS--K--A-IE-----FKL	D--YPS--K--A-IE-----FKL		
	Strep. pneumoniae(1)	AAL00584	--V-----Q-LE-MR---LHS-	--V-----Q-LE-MR---LHS-		
	Strep. pneumoniae(2)	AAK75194	D--Y-K-VN-VF--AK-DADIST	D--Y-K-VN-VF--AK-DADIST		
	Strep. pyogenes(1)	AAK33706	--V---Q-LE-MR---LQS--	--V---Q-LE-MR---LQS--		
	Strep. pyogenes(2)	AAK34186	D--Y-K-IN-V---M---ADISV	D--Y-K-IN-V---M---ADISV		
	Cor. diphtheriae	*	-N---A--RFVD--V-L-AD-TV	-N---A--RFVD--V-L-AD-TV		
	Myc. smegmatis	1361415	-NV--A--RFVE-M--L-AD-RT	-NV--A--RFVE-M--L-AD-RT		
	Myc. leprae	1171078	-NV--A--RFVE-M--L-AD-RT	-NV--A--RFVE-M--L-AD-RT		
	Myc. tuberculosis	1361416	-NV--A--RFVE-M--L-AD-RT	-NV--A--RFVE-M--L-AD-RT		

Fig. 1. Excerpt from an alignment for the UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) protein showing a 16 aa insert in a conserved region (boxed) in the seven tested *Chlamydiaceae* species and *W. chondrophila*. A similar insert in this position is also present in the two *Streptomyces* species. The sequences for MurA from *Chlam. abortus*, *Chlam. psittaci*, *Chlam. felis* and *W. chondrophila* were determined in the present work. Dashes in the sequence alignments indicate identity with the amino acid shown on the top line. Accession numbers of the sequences are shown in the second column. Numbers (1) and (2) beside sequences represent the two different homologues found in some Gram-positive bacteria. An asterisk (*) denotes a sequence retrieved from the NCBI incomplete (www.ncbi.nlm.nih.gov/BLAST/unfinishedgenome.html) Microbial Genome Sequence database. Abbreviations in the species names are as follows: A., *Agrobacterium*; Acin., *Acinetobacter*; Aqu., *Aquifex*; Bac., *Bacillus*; Bor., *Borrelia*; Bru., *Brucella*; Buch., *Buchnera*; Ca., *Caulobacter*; Camp., *Campylobacter*; Cb., *Chlorobium*; Chl., *Chlamydia*; Chlam., *Chlamydophila*; Clo., *Clostridium*; Cor., *Corynebacterium*; D., *Deinococcus*; Des., *Desulfotomaculum*; E., *Escherichia*;

Fig. 2. An unrooted consensus neighbour joining phylogenetic tree based on MurA protein sequences. The tree shown here is based on 390 aligned positions and was bootstrapped 100 times. The observed bootstrap scores for all nodes that were higher than 50 are shown.

chlamydiae to the *Streptomyces* species. These two groups of species formed a clade, exclusive of any other bacteria, which was recovered 100% of the time in phylogenetic trees constructed either with or without the large insert (Fig. 2). The unusual branching of *Streptomyces* with chlamydiae, instead of with other Gram-positive bacteria, strongly suggests that *murA* gene has been laterally transferred between these two groups of species.

Horizontal transfer of a cell wall biosynthesis gene (*glmU*) between *Archaea* and chlamydiae

We have also identified a prominent signature that is uniquely shared by the various archaeal homologues and those from the chlamydial species. In the protein UDP-N-acetylglucosamine pyrophosphorylase (GlmU), which catalyses both the acetylation of glucosamine 1-phosphate and the uridylation of N-acetylglucosamine 1-phosphate to produce UDP-N-acetylglucosamine (Gehring *et al.*, 1996; Pompeo *et al.*, 2001), a 17 aa indel in a conserved region has been identified that is commonly shared by various archaea and chlamydial species (Fig. 3). Although this protein is found in most other bacteria, they do not contain this insert. The presence of this uniquely shared indel between *Chlamydiaceae* and archaea suggests a specific relationship between these groups, exclusive of all other bacterial phyla. However, such a relationship is inconsistent with all of the other phylogenies and signatures in different

proteins (Woese, 1987; Brown & Doolittle, 1997; Gupta, 1998; Ludwig & Klenk, 2001), as well as various characteristics distinguishing *Archaea* and *Bacteria* (Woese, 1987). The most likely explanation for this shared signature is that the *glmU* gene was laterally transferred from an archaeon to a common ancestor of the chlamydial species. In phylogenetic trees based on GlmU protein sequences, the chlamydial and archaeal homologues exhibit strong affinity for each other, grouping together in 94% of the bootstrap replicates (results not shown). These results support our contention that the gene for this protein was laterally transferred from an archaeon to an ancestor of the chlamydial group. The conserved nature of this indel and its presence in all known chlamydial and archaeal homologues provide arguments against the possibility that this indel was independently introduced in these two groups of prokaryotes.

DISCUSSION

In the present work, we have identified a number of conserved indels or signatures in different proteins that appear to be unique and distinctive characteristics of chlamydial species. The proteins in which these indels are present are highly conserved and found in most known groups of bacteria. The identified signatures are found in conserved regions and they are flanked on both sides by conserved residues that ensure that the observed changes are not due to sequence misalignment or other

Enter., *Enterobacter*; *Fibro.*, *Fibrobacter*; *H.*, *Haemophilus*; *Hel.*, *Helicobacter*; *L.*, *Lactococcus*; *Lis.*, *Listeria*; *Meso.*, *Mesorhizobium*; *Myc.*, *Mycobacterium*; *Myx.*, *Myxococcus*; *Nei.*, *Neisseria*; *Pas.*, *Pasteurella*; *Pse.*, *Pseudomonas*; *Ral.*, *Ralstonia*; *Ri.*, *Rickettsia*; *Sal.*, *Salmonella*; *Sin.*, *Sinorhizobium*; *Sta.*, *Staphylococcus*; *Str.*, *Streptomyces*; *Strep.*, *Streptococcus*; *Sy.*, *Synechocystis*; *T.*, *Thermotoga*; *Tre.*, *Treponema*; *V.*, *Vibrio*; *X.*, *Xylella*; *Wad.*, *Waddlia*; *Y.*, *Yersinia*.

Archaeobacteria	Pyrococcus horikoshii	3258369	HLNVYVGSIIIGENTNLGAGTITANL	RHDNKTVKVEIKGNLED	SGRRKLGAIGHNVKGT
	Pyrococcus abyssi	5459119	-----V-----	---R-I-----K---	-----
	Halobacterium sp. NRC-1	AAG18651	--HA-----VL-ADV-F---KV---	---DAP-Q-QVN-E-V---	T----F-VVV-DDT---
	Sulfolobus solfataricus	AAK41040	--S-----V-C-GV-F-----	-F-EEE---N- NERVS	---K-----V-AH-R--
	Sulfolobus tokodaii	BAB65442	--S-----V-A-DV-F---LI---	-F-E-E---NV--KRIS	-----F--GH-R--
Chlamydiae	Aeropyrum pernix	5104874	-----VV--HV-----V---	-F-GRS-RMTV--ERV-	T-L-----V--GYAQ--
	Methanococcus jannaschii	2127932	--S-----C-F-CN-----	-F-D-P---N- SKRVE	-V---V-M-D---
	Meth. thermautotrophicus	2622712	--S-----V-M-C-IA---NI---	-F-DGP-RMVV-DDVVE	T-----VF-DG---
Gram(+)ve Bacteria	Chlamydia trachomatis	3329076	-FA----VF-SRV----VRC--F	-L-G--IFFHHS-ERCN	TK-K----FL-RG-SV-
	Chlamydia muridarum	NP_297290	-FA----VL-SRV----VRC--F	-L-GGNIFLRY-S-DRY-	TQ-K----FL-KG-SI-
	Chlamydophila psittaci	*	-FA----VL-AEV----VRC--F	-L-G-NIS-TC-EGKV-	TQL--V--FL-TK-SI-
	Chlamydophila pneumoniae	4377055	-FA-L---VLSSEV----VRC--F	-L-GRNII-RSTSDKSK	T-----FL-KG-AI-
	Listeria monocytogenes	532203	-FI-M--AE--K-V-V-C-S-AV-Y		D-KN-AKT---D-FV-
	Listeria innocua	NP_469582	-FI-M--AE--K-V-V-C-S-AV-Y		D-KN-AKT---D-FV-
	Clostridium acetobutylicum	NP_349818	--T-I--AEV--GC-F-C--VVV-Y		D-KD-HKTV--NKSFI-
	Lactococcus lactis	NP_268040	--T-I-NATV--KV-F-----F		D-KN-FNTE-DDFAFI-
	Corynebacterium ammoniagenes	2289092	--S-I--AQV--DV--C-S-V-Y		D-KN-FKTV--NG-FI-
	Bacillus halodurans	NP_240931	--S-I--AEV--RV-FSC-S--V-Y		D-KN-FLTK-EDDAFI-
	Bacillus subtilis	135927	-S-----AEV--TDV--C-S-V-Y		D-KN-YLTK-EDGAFI-
	Bacillus caldolyticus	1169872	--S-I--AEV-ADV--C-S-V-Y		D-VN-HMTK-ENGAFI-
	Mycobacterium leprae	NP_301308	--T---AD---HS-I--SSVFV-Y		D-TA-QRTT--SH-R--
	Mycobacterium tuberculosis	1870010	--T---AD---YS-I--SSVFV-Y		D-TS-RRTTV-SH-R--
	Streptococcus pneumoniae	NP_345467	--T-I-NCEV-S-V-F-----V-Y		D-KN-YKTV--N--FV-
Other Gram(-)ve Bacteria	Streptococcus pyogenes	NP_268741	-T-I-NAQV-SSV-V-----V-Y		D-QN-YETV--DHAFI-
	Staphylococcus epidermidis	AAL11408	--S-I--AE--R-I-C-S--V-Y		D-AN-FKT-V-KDAFI-
	Staphylococcus aureus	BAB56661	--S-I--AV---R--I-C---V-Y		D-EN-FKT-V-KDSFV-
	Thermotoga maritima	4982202	--S-I--AFV-K-V-V-----C-Y		D-KK-NPTF-EDGAFI-
	Deinococcus radiodurans	AAF10386	--A-L--VE--AE--V-----V-F		D-LN-HQSKV-AG-FI-
	Nostoc sp PCC 7120	BAB75620	--S-L--ATA-TGV-I-----Y		D-VK-HRTK--DRT---
	Synechocystis sp. PCC6803	1001368	--S-L--AML-QRV-V-----Y		D-VS-HATV--SGT---
	Aquifex aeolicus	2983227	--A-I--ATV-----I---VF--Y		D-K--YESYV-KSAFI-
	Helicobacter pylori J99	4155177	--S-L--CE--K--V--V--C-Y		D-KK-HQT---E--FI-
	Campylobacter jejuni	NP_281982	--S-L--CE-DSG--I-C---C-Y		D-VK-HKT---K--FV-
	Caulobacter crescentus	NP_421107	--S-L--GSV--KA-I-----FC-Y		D-FE-FETHV-KGAFI-
	Mesorhizobium loti	BAB48341	--T-I--ARV-AGA-I-----C-Y		D-FS-FFTD--EGAFV-
	Agrobacterium tumefaciens	NP_354772	--T-I--AFV-AGS-I---A--C-Y		D-YN-SETR--A-SFI-
	Ralstonia solanacearum	CAD13705	--A---ATV-SRV-I-----C-Y		D-AN-FRT--EDDAFI-
	Sinorhizobium meliloti	CAC46236	--T-I--AFV-AGS-I-----C-Y		D-VN-HVTR--A-AFV-
	Brucella melitensis	AAL53926	--A-I--AV--ASS-I-----C-Y		D-YN-FKT---D-AF--
	Neisseria meningitidis	NP_273104	--T-I--AEV-SK--F-----I--Y		D-VH-HKTV--DE-RI-
	Escherichia coli	290578	--T-L--AE--D-V-I-----C-Y		D-AN-FKT---DD-FV-
	Haemophilus influenzae Rd	1169921	--T-----E--S-C-I--V--C-Y		D-AN-FKT---DD-FV-
	Vibrio cholerae	NP_232388	--T-L--AE--KGV-V--V--C-Y		D-AN-HKTV--DD-FV-
	Pseudomonas aeruginosa	AAG08937	--S-L--AEL-A-C-I-----C-Y		D-AN-FRTEL--DD-FI-
	Yersinia Pestis	CAC93568	--S-L--AE--AGV-I-----C-Y		D-AN-FKT---DD-FV-
	Buchnera sp. APS	NP_239868	--S-L-N-E--SKV-I---S--C-Y		D-AN-FKT---D--LV-
	Xylella fastidiosa	NP_298430	--T-L--AQ--TKV-I-----C-Y		D-IN-SITL--DGAFI-
	Pasteurella multocida	15603671	--S---AEV-MQC-I---V--C-Y		D-AN-FKT---D--FV-
	Salmonella typhi	6762464	--T-L--AE--D-V-I-----C-Y		D-AN-FKTV--DD-FV-
	Salmonella typhimurium	AAL22720	--T-L--AE--D-V-I-----C-Y		D-AN-FKTV--DD-FV-

Fig. 3. Partial sequence alignment of GlmU protein sequences showing the presence of a 17 aa indel that is shared by various archaeal and chlamydial homologues. The insert in *Chlam. pneumoniae* contains an extra 3 aa (KID) at the C-terminal end which are not shown. The gene for this protein is postulated to have been laterally transferred from an archaeon to a common ancestor of the chlamydiae species.

artefacts. For three of these signatures (in the MurA, EF-P and MgtE proteins), their specificity for chlamydial species was tested by cloning and sequencing the signature regions from a number of species for which no sequence information was available. In all cases where we were successful in amplifying the corresponding genes, the expected indels were present in these species confirming and validating the usefulness of these molecular signatures.

The chlamydial group of species until very recently consisted of a single family, the *Chlamydiaceae*, which contained one genus and four species then named *Chlamydia trachomatis*, *Chlamydia pecorum*, *Chlamydia pneumoniae* and *Chlamydia psittaci* (Fields & Barnes, 1992; Bush & Everett, 2001; Meijer *et al.*, 1999; Pettersson *et al.*, 1997; Fukushima & Hirai, 1992;

Grayston *et al.*, 1989; Moulder *et al.*, 1984; Page, 1968). However, the taxonomy of this group has undergone a major revision within the last three years. Based on DNA-DNA hybridization, and 16S rRNA and 23S rRNA sequence analyses, Everett *et al.* (1999) have proposed a new taxonomic classification for this group. The proposal recognizes nine species within the *Chlamydiaceae*, which are placed into two different genera, *Chlamydia* and *Chlamydophila*. *Chlamydia* consists of three species, *Chl. trachomatis*, *Chl. suis* and *Chl. muridarum*, whereas the genus *Chlamydophila* is made up of six species, *Chlam. abortus*, *Chlam. psittaci*, *Chlam. felis*, *Chlam. caviae*, *Chlam. pneumoniae* and *Chlamydophila pecorum* (Bush & Everett, 2001; Herrmann *et al.*, 2000). In addition to *Chlamydiaceae*, the proposal recognizes two new families of chlamydial-like organisms, *Parachlamydiaceae* and *Simkaniaceae*

(Bush & Everett, 2001; Everett *et al.*, 1999; Kahane *et al.*, 1999; Ossewaarde & Meijer, 1999). A fourth family of chlamydia-like organisms, *Waddliaceae*, was recognized in a separate proposal (Rurangirwa *et al.*, 1999). The assignment of distinct family status to these chlamydia-like organisms is presently based on very limited sequence information, consisting primarily of differences in their 16S and 23S rRNA sequences (Everett *et al.*, 1999). The proposal to identify new families and genera within chlamydiae based mainly on arbitrary degree of differences in the 16S and 23S rRNA sequences has met with strong opposition from many scientists (Schachter *et al.*, 2001). It has been argued that since all known chlamydiae species share a unique and highly conserved biological replication cycle, they should be retained within a single genus unless there is a compelling reason to do otherwise.

The signatures described here should prove helpful in clarifying the classification of the chlamydial species. For all eight signatures described here, sequence information is known for at least two species from each of the two proposed genera within *Chlamydiaceae*. For some signatures, sequence data are known for six to seven out of a possible nine species. Since the two genera, *Chlamydia* and *Chlamydophila*, are indicated to be monophyletic in their 16S and 23S rRNA trees (with 100% bootstrap scores) (Everett *et al.*, 1999; Bush & Everett, 2001), the presence of these signatures in at least two species from each genus strongly suggests that they will also be found in other species of *Chlamydiaceae* for which sequence information is lacking at present. Therefore, all of the signatures described here are at least distinctive of the *Chlamydiaceae* group of species. For the MurA and EF-P proteins, we were also successful in amplifying a gene fragment from the two chlamydia-like organisms, *W. chondrophila* and *S. negevensis*, respectively. The sequences from both of these species contained the indicated indels. In phylogenetic trees based on MurA sequences, *Waddlia* formed an outgroup of the *Chlamydiaceae* species and the clade consisting of these species was supported at a very high degree of bootstrap confidence level (99 out of 100). These results provide strong molecular and phylogenetic evidence that these species are specifically related to the traditional chlamydiae (i.e., *Chlamydiaceae*) species. The presence of these signatures in these organisms, which have been assigned to different families within the order *Chlamydiales*, is strongly suggestive that these signatures may prove distinctive for the entire *Chlamydiales* order. It is unclear at present whether the other identified signatures are also present in these chlamydia-like organisms. It is quite possible that of these signatures, some may prove specific for the *Chlamydiaceae*, while others will be commonly shared by the entire *Chlamydiales* order. In future studies, it should then be possible to unambiguously identify species belonging to the *Chlamydiaceae* family or the *Chlamydiales* order from all other bacteria, simply by determining the presence or absence of these molecular signatures. Further studies should also reveal whether some of these signatures are

commonly shared by the *Chlamydiaceae* and one or more of the *Parachlamydiaceae*, *Simkaniaceae* and *Waddliaceae* families. This should provide insight as to how the various proposed families within *Chlamydiales* have evolved and are related to each other.

An important question that needs to be understood in future work concerns the functional significance of the identified indels on the biochemical and physiological characteristics of the chlamydial species. Since the identified signatures are present in all available chlamydial sequences but not in any other bacteria, it strongly suggests that they were introduced in a common ancestor of this group of bacteria at the time of its evolution. Because these indels have not been lost in any of the chlamydial species examined to date, they likely play an important role in the biology of these organisms and could have been important in their evolution. Thus, it is of much interest to understand the functional significance of these site-specific alterations in these conserved and widely distributed genes. Although all of these signatures are potentially interesting, the two containing the most prominent indels are both in proteins involved in cell-wall-peptidoglycan biosynthesis. The MurA protein, which contains the 16 aa insert, is responsible for carrying out the first essential and committed step in peptidoglycan biosynthesis in various bacteria (Brown *et al.*, 1995; Du *et al.*, 2000). The large insert seen in MurA is also present in the two streptomyces species *Streptomyces coelicolor* and *Streptomyces lividans*, but not any other Gram-positive or Gram-negative bacteria. The shared presence of this common insert as well as the phylogenetic studies based on MurA protein sequences strongly suggests that a lateral transfer of this gene has occurred between these two groups of species. The biochemical significance of this lateral gene transfer event and whether the exchange occurred from chlamydiae to *Streptomyces* or *vice versa* is unclear at present.

Similar to the MurA protein, the GlmU protein that contains the 17 aa indel is also involved in peptidoglycan biosynthesis. Interestingly, the large insert in this protein is also present in all archaeal homologues and phylogenetic studies of GlmU protein sequences indicate that the gene containing this indel has been laterally acquired by chlamydiae from an archaeon. The above observations are of much interest because although the genomes of chlamydial species contain a full complement of the genes involved in peptidoglycan synthesis (Stephens *et al.*, 1998; Hatch, 1998; Kalman *et al.*, 1999; Read *et al.*, 2000), these species are generally believed to lack, or be deficient in, peptidoglycan (Hua *et al.*, 1985; Fox *et al.*, 1990; Hatch, 1996, 1998; Ghuyssen & Goffin, 1999). It has been suggested that chlamydiae synthesize a defective peptidoglycan or an atypical cell wall differing from other bacteria (Hatch, 1996; Ghuyssen & Goffin, 1999). However, the biochemical basis of this phenomenon remains to be understood. In this context, the presence of large inserts in essential enzymes involved in cell wall synthesis has the potential to inactivate or modify the cellular functions of such proteins. It should

be pointed out in this regard, that GlmU is a bifunctional protein which catalyses two of the essential steps leading to the synthesis of UDP-N-acetylglucosamine, a fundamental precursor for bacterial cell wall synthesis (Gehring *et al.*, 1996; Pompeo *et al.*, 2001). It has been shown that the N-terminal domain of this protein (residues 3–227 in *E. coli*) catalyses the uridyltransferase activity whereas the C-terminal domain is responsible for the acetyltransferase activity (Brown *et al.*, 1999; Pompeo *et al.*, 2001). The insert in GlmU is in the C-terminal region and hence its main effect should be on the acetyltransferase activity. The acetyltransferase domain of GlmU is made up of 10 regular coils, with an inter-coil distance of generally 17 aa (Brown *et al.*, 1999; Sulzenbacher *et al.*, 2001). The observed 17 aa insert in GlmU is expected to add an extra coil in the protein structure. How this may affect the function of the protein remains to be determined. It is also of much interest that the GlmU homologues from chlamydiae are much shorter than those from other species (about 205 aa compared to 450 aa from other bacteria) and they seem to lack the N-terminal region. Thus, the enzyme from chlamydial species should be lacking the uridyltransferase activity. Similar to chlamydiae, the cell wall composition/structure in the *Archaea* also differs from the *Bacteria* (Woese, 1987; Kandler & König, 1993). If there have been lateral gene transfers between *Archaea* and chlamydiae, or chlamydiae and *Streptomyces*, as strongly suggested by the results presented here, then it becomes of much interest to determine the functional significance of these lateral gene-transfer events in these divergent prokaryotes.

Chlamydial infections are responsible or have been implicated in a wide variety of diseases affecting many different systems and organs (Moulder *et al.*, 1984). Hence, identification of the known chlamydial species and other novel species related to this group in different pathological conditions is of importance. The dependence of the *Chlamydiaceae* on host cells for their growth, and the presence of other contaminating bacteria in clinical specimens, have presented problems in their definitive identification in clinical situations. Some of the large chlamydial-specific signatures described here could prove useful in this regard. For example, the MurA protein contain a 16 aa or 48 nt insert that is distinctive for the chlamydial species and not found in any other bacteria. The homologues of MurA are also not found in eukaryotes. Based on alignment of the *murA* gene sequences from various chlamydial species, several conserved regions both within this large insert, and in regions flanking it, can be identified. PCR amplification utilizing such sequences should exhibit a high degree of specificity in detecting chlamydial species. Likewise, the large 17 aa or 51 nt insert in the GlmU protein could also prove useful for similar purposes.

ACKNOWLEDGEMENTS

We gratefully acknowledge the generous gifts of DNA from various species by Dr Karin D. E. Everett (University of Georgia, Athens, USA) and Dr A. Petrich (St Joseph's

Hospital, Hamilton, Ontario, Canada). The present work was supported by funds from the Committee on Scientific Development at McMaster University. We also thank and acknowledge the use of some sequence data from the NCBI Microbial Unpublished Genomes Sequence Database, which is contributed to by different sequencing centres.

REFERENCES

- Brown, J. R. & Doolittle, W. F. (1997). *Archaea* and the prokaryote-to-eukaryote transition. *Microbiol Rev* **61**, 456–502.
- Brown, E. D., Vivas, E. I., Walsh, C. T. & Kolter, R. (1995). MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. *J Bacteriol* **177**, 4194–4197.
- Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreux, D., Cambillau, C. & Bourne, Y. (1999). Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily. *EMBO J* **18**, 4096–4107.
- Bush, R. M. & Everett, K. D. E. (2001). Molecular evolution of the *Chlamydiaceae*. *Int J Syst Evol Microbiol* **51**, 203–220.
- Daian, C. M., Wolff, A. H. & Bielory, L. (2000). The role of atypical organisms in asthma. *Allergy Asthma Proc* **21**, 107–111.
- Du, W., Brown, J. R., Sylvester, D. R., Huang, J., Chalker, A. F., So, C. Y., Holmes, D. J. & Wallis, N. G. (2000). Two active forms of UDP-N-acetylglucosamine enolpyruvyltransferase in gram-positive bacteria. *J Bacteriol* **182**, 4146–4152.
- Everett, K. D., Bush, R. M. & Andersen, A. A. (1999). Emended description of the order *Chlamydiales*, proposal of *Parachlamydiaceae* fam. nov. and *Simikaniaceae* fam. nov., each containing one monotypic genus, revised taxonomy of the family *Chlamydiaceae*, including a new genus and five new species, and standards for the identification of organisms. *Int J Syst Bacteriol* **49**, 415–440.
- Felsenstein, J. (1994). PHYLIP version 3.5. Seattle, WA: University of Washington.
- Fields, P. I. & Barnes, R. C. (1992). The genus *Chlamydia*. In *The Prokaryotes*, pp. 3691–3709. Edited by A. Balows and others. New York: Springer.
- Fox, A., Rogers, J. C., Gilbert, J., Morgan, S., Davis, C. H., Knight, S. & Wyrick, P. B. (1990). Muramic acid is not detectable in *Chlamydia psittaci* or *Chlamydia trachomatis* by gas chromatography-mass spectrometry. *Infect Immun* **58**, 835–837.
- Fukushi, H. & Hirai, K. (1992). Proposal of *Chlamydia pecorum* sp. nov. for *Chlamydia* strains derived from ruminants. *Int J Syst Bacteriol* **42**, 306–308.
- Gehring, A. M., Lees, W. J., Mindiola, D. J., Walsh, C. T. & Brown, E. D. (1996). Acetyltransferase precedes uridyl transfer in the formation of UDP-N-acetylglucosamine in separable active sites of the bifunctional GlmU proteins of *Escherichia coli*. *Biochemistry* **35**, 579–585.
- Ghuysen, J.-M. & Goffin, C. (1999). Lack of cell wall peptidoglycan versus penicillin sensitivity: new insights into the chlamydial anomaly. *Antimicrob Agents Chemother* **43**, 2339–2344.
- Grayston, J. T., Kuo, C. C., Campbell, A. & Wang, S. P. (1989). *Chlamydia pneumoniae* sp. nov. for *Chlamydia* sp. strain TWAR. *Int J Syst Bacteriol* **39**, 88–90.
- Griffiths, E. & Gupta, R. S. (2001). The use of signature sequences in different proteins to determine the relative branching order of bacterial divisions: evidence that *Fibrobacter* diverged at a similar time to *Chlamydia* and the *Cytophaga-Flavobacterium-Bacteroides* division. *Microbiology* **147**, 2611–2622.

- Gupta, R. S. (1998).** Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol Mol Biol Rev* **62**, 1435–1491.
- Gupta, R. S. (2000a).** The *natural* evolutionary relationships among prokaryotes. *CRC Crit Rev Microbiol* **26**, 111–131.
- Gupta, R. S. (2000b).** The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. *FEMS Microbiol Rev* **24**, 367–402.
- Gupta, R. S. (2001).** The branching order and phylogenetic placements of species from completed bacterial genomes based on conserved indels found in various proteins. *Int Microbiol* **4**, 187–202.
- Gupta, R. S. & Johari, V. (1998).** Signature sequences in diverse proteins provide evidence of a close evolutionary relationship between the *Deinococcus/Thermus* group and Cyanobacteria. *J Mol Evol* **46**, 716–720.
- Gupta, R. S. & Singh, B. (1994).** Phylogenetic analysis of 70 kDa protein sequences suggests a chimeric origin of the eukaryotic cell nucleus. *Curr Biol* **4**, 1104–1114.
- Gupta, R. S., Bustard, K., Falah, M. & Singh, D. (1997).** Sequencing of heat shock protein 70 (DnaK) homologs from *Deinococcus proteolyticus* and *Thermomicrobium roseum* and their integration in a protein based phylogeny of prokaryotes. *J Bacteriol* **179**, 345–357.
- Hatch, T. P. (1996).** Disulfide cross-linked envelope proteins: the functional equivalent of peptidoglycan in chlamydiae? *J Bacteriol* **178**, 1–5.
- Hatch, T. P. (1998).** Chlamydia: old ideas crushed, new mysteries bared. *Science* **282**, 638–639.
- Herrmann, B., Pettersson, B., Everett, K. D. E., Mikkelsen, N. E. & Kirsebom, L. A. (2000).** Characterization of the *rnpB* gene and RNase P RNA in the order *Chlamydiales*. *Int J Syst Evol Microbiol* **50**, 149–158.
- Hua, S., Youxum, Z. & Rungte, L. (1985).** Presence of muramic acid in *Chlamydia trachomatis* proved by liquid chromatography-mass spectrometry. *Kexue Tongbao (Chinese Science Bulletin)*, **30**, 695–699.
- Kahane, S., Everett, K. D. E., Kimmel, N. & Friedman, M. G. (1999).** *Simkania negevensis* strain Z^T: growth, antigenic and genome characteristics. *Int J Syst Bacteriol* **49**, 815–820.
- Kalman, S., Mitchell, W., Marathe, R. & 7 other authors (1999).** Comparative genomes of *Chlamydia pneumonia* and *C. trachomatis*. *Nat Genet* **21**, 385–389.
- Kandler, O. & König, H. (1993).** Cell envelopes of archaea: structure and chemistry. In *The Biochemistry of Archaea (Archaeobacteria)*, pp. 223–259. Edited by M. Kates, D. J. Kushner & A. T. Matheson. New York: Elsevier.
- Kuo, C. C., Jackson, L. A., Campbell, L. A. & Grayston, J. T. (1995).** *Chlamydia pneumoniae* (TWAR). *Clin Microbiol Rev* **8**, 451–461.
- Laga, M., Nzila, N. & Goeman, J. (1991).** The interrelationship of sexually transmitted diseases and HIV infection: implications for the control of both epidemics in Africa. *AIDS* **5**, S55–S63.
- Ludwig, W. & Klenk, H.-P. (2001).** Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, *The Archaea and the deeply branching and phototrophic Bacteria*, pp. 49–65. Edited by D. R. Boone and R. W. Castenholz. Berlin: Springer.
- Meijer, A., Morre, S. A., van den Brule, A. J. C., Savelkoul, P. H. M. & Ossewaarde, J. M. (1999).** Genomic relatedness of chlamydia isolates determined by amplified fragment length polymorphism analysis. *J Bacteriol* **181**, 4469–4475.
- Moulder, J. W., Hatch, T., Kuo, C. C., Schachter, J. & Storz, J. (1984).** Genus *Chlamydia*. In *Bergey's Manual of Systematic Bacteriology*, pp. 729–739. Edited by N. R. Krieg and J. G. Holt. Baltimore, MD: Williams & Wilkins.
- Ossewaarde, J. M. & Meijer, A. (1999).** Molecular evidence for the existence of additional members of the order *Chlamydiales*. *Microbiology* **145**, 411–417.
- Page, L. A. (1968).** Proposal for the recognition of two species in the genus *Chlamydia* Jones, Rake and Stearns, 1945. *Int J Syst Bacteriol* **18**, 51–66.
- Pettersson, B., Andersson, A., Leitner, T., Olsvik, O., Uhlen, M., Storey, C. & Black, C. M. (1997).** Evolutionary relationships among members of the genus *Chlamydia* based on 16S ribosomal DNA analysis. *J Bacteriol* **179**, 4195–4205.
- Pompeo, F., Bourne, Y., Heijenoer, J. V., Fassy, F. & Mengin-Lecreux, D. (2001).** Dissection of the bifunctional *Escherichia coli* N-acetylglucosamine-1-phosphate uridylyltransferase enzyme into autonomously functional domains and evidence that trimerization is absolutely required for glucosamine-1-phosphate acetyltransferase activity and cell growth. *J Biol Chem* **276**, 3883–3889.
- Read, T. D., Brunham, R. C., Shen, C. & 22 other authors (2000).** Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. *Nucleic Acids Res* **28**, 1397–1406.
- Rurangirwa, F. R., Dilbeck, P. M., Crawford, T. B., McGuire, T. C. & McElwain, T. F. (1999).** Analysis of the 16S rRNA gene of microorganism WSU 86-1044 from an aborted bovine foetus reveals that it is a member of the order *Chlamydiales*: proposal of *Waddliaceae* fam. nov., *Waddlia chondrophila* gen. nov., sp. nov. *Int J Syst Bacteriol* **49**, 577–581.
- Saikku, P. (2000).** *Chlamydia pneumoniae* in atherosclerosis. *J Intern Med* **247**, 396.
- Schachter, J. (1999).** Infection and disease epidemiology. In *Chlamydia: Intracellular Biology, Pathogenesis, and Immunity*, pp. 139–169. Washington, DC: American Society for Microbiology.
- Schachter, J., Stephens, R. S., Timms, P. & 29 other authors (2001).** Radical changes to chlamydial taxonomy are not necessary just yet. *Int J Syst Evol Microbiol* **51**, 249.
- Stephens, R. S., Kalman, S., Lammel, C. & 9 other authors (1998).** Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**, 754–759.
- Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F. & Bourne, Y. (2001).** Crystal structure of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridylyltransferase bound to acetyl-coenzyme A reveals a novel active site architecture. *J Biol Chem*, **276**, 11844–11851.
- Woese, C. R. (1987).** Bacterial evolution. *Microbiol Rev* **51**, 221–271.

Received 28 September 2001; revised 26 March 2002; accepted 3 April 2002.