INTRODUCTION

The term intein refers to unusual protein sequences that are excised from the precursor protein during maturation. Its coding sequence is always inserted in-frame with a protein-coding sequence (Perler et al., 1994). Protein maturation involves excision of the central protein from the precursor molecule and ligation of the N- and C-terminal domains to form the mature protein. Inteins probably possess endonuclease activity since they show homology to endonucleases of eukaryotes which mediate homing of the encoding sequence during maturation of the mature protein. Inteins have so far been described in the yeast vacuolar proton-pump ATPase subunit (Saccharomyces cerevisiae and Candida tropicalis) and in archaeobacterial DNA polymerases (Thermococcus litoralis and Pyrococcus spp.) (for a review see Colston & Davis, 1994). Recently, as many as 18 putative inteins have been identified in the genome of Methanococcus jannaschii, an archaeon, from which the complete genome sequence has been determined (Bult et al., 1996).

Inteins have so far been described in the yeast vacuolar proton-pump ATPase subunit (Saccharomyces cerevisiae and Candida tropicalis) and in archaeobacterial DNA polymerases (Thermococcus litoralis and Pyrococcus spp.) (for a review see Colston & Davis, 1994). Recently, as many as 18 putative inteins have been identified in the genome of Methanococcus jannaschii, an archaeon, from which the complete genome sequence has been determined (Bult et al., 1996).

In mycobacteria, inteins have been found only occasionally in the corresponding proteins of Mycobacterium kansasii (Colston & Davis, 1994). Among mycobacteria, the presence of GyrA inteins apparently is not uniform. While in gyrA of M. leprae intein coding sequences have been constantly observed, inteins have been found only occasionally in the corresponding proteins of Mycobacterium kansasii, Mycobacterium flavescens and Mycobacterium gordona. Other mycobacterial species were investigated at a single strain level, thus excluding conclusions concerning the spreading of this characteristic feature (Fsihi et al., 1996).

Among mycobacteria, the presence of GyrA inteins apparently is not uniform. While in gyrA of M. leprae intein coding sequences have been constantly observed, inteins have been found only occasionally in the corresponding proteins of Mycobacterium kansasii, Mycobacterium flavescens and Mycobacterium gordona. Other mycobacterial species were investigated at a single strain level, thus excluding conclusions concerning the spreading of this characteristic feature (Fsihi et al., 1996). The finding that some strains of a given species have GyrA inteins, while others have not, prompted us to investigate the nature of this phenomenon.

METHODS

Strains. Several strains, isolated at Medizinische Hochschule Hannover (Germany), Zürich (Switzerland), Berne (Switzerland), Antwerpen (Belgium) and Firenze (Italy), each of Mycobacterium kansasii (n = 23), Mycobacterium malmoense (n = 8), Mycobacterium marinum (n = 8), Mycobacterium ulcerans (n = 4) and Mycobacterium xenopi (n = 10) were included in this analysis.

Identification. Isolates were identified by conventional methods as well as by 16S rRNA sequencing (Rogall et al., 1990). Strains of M. kansasii were also investigated by
hybridization with a commercially available probe for \textit{M. kansasii} (Gen-Probe; Tortoli et al., 1994a,b).

**Molecular biological methods and protein alignment.** PCR amplification of the non-coding 16S–23S rRNA intergenic spacer region was performed with primers 248 (5′-GTGGTTTCTTCCCTTGG-3′) and 42 (5′-CCACACGGGTTAACCTCGC-3′) using standard conditions (Rogall et al., 1990). Nucleic acid sequencing was done manually using 32P-labelled dCTP and sequenase (USB).

A previously devised PCR strategy using primers H49 (5′-AGGTGTCGCGCGGATATGTT-3′) and H50 (5′-TCCGCCCGAGCCAGCCACGC-3′) was used to investigate the presence of intein coding sequences in \textit{gyrA} (Fsihi et al., 1996). Sequencing of \textit{gyrA} intein coding sequence of \textit{M. malmoense} was performed by PCR-mediated Taq cycle sequencing using an ABI373 sequencer. Protein alignment of known GyrA inteins (\textit{M. leprae}, accession no. Z68206; \textit{M. flavescens}, accession no. Z68209; \textit{M. gordonae}, accession no. Z68208; \textit{M. kansasii}, accession no. Z68207; and \textit{M. xenopi}, accession no. U67876) was performed with the program CLUSTAL (PC/Gene, IntelliGenetics, release 6.85) as was the 16S rRNA sequence alignment of these species (Rogall et al., 1990).

**RESULTS AND DISCUSSION**

Mycobacteria with intein coding sequences in \textit{gyrA} yield a PCR product of 1.6 kbp or 0.9 kbp, whereas those with inteinless GyrA generate smaller fragments of around 350 bp.

Representative PCR results are presented in Fig. 1. The different isolates of \textit{M. malmoense}, \textit{M. marinum}, \textit{M. ulcerans} and \textit{M. xenopi} investigated showed a constant pattern: (i) isolates of \textit{M. malmoense} were characterized by a gene fragment of 1.6 kbp, indicative of a GyrA intein; (ii) isolates of \textit{M. marinum} and \textit{M. ulcerans} showed an amplified gene fragment of 350 bp character-

![Fig. 1. PCR amplification of a gyrA gene fragment from different mycobacteria. Mycobacteria with intein coding sequences in gyrA yield a PCR product of 1.6 kbp or 0.9 kbp, whereas those with inteinless GyrA generate a smaller fragment of around 350 bp. Lanes: 1–4, \textit{M. malmoense}; 5–8, \textit{M. marinum}; 9 and 10, \textit{M. kansasii} type I; 11 and 12, \textit{M. kansasii} type II; 13–16, \textit{M. xenopi}.](image-url)
malmoense, M. marium and M. ulcerans, which appear as genetically homogeneous taxons, M. kansasii genetically is more diverse and consists of two genetically distinct subspecies. Analysis of the PCR amplification data (see Fig. 1) demonstrated that the presence of an intein coding sequence in gyrA was invariably associated with M. kansasii type I (13 of 13 investigated), whereas none of the type II M. kansasii isolates showed the presence of an intein coding sequence (10 of 10 investigated).

Sequence determination of PCR products obtained with primers H49 and H50 (Fsihi et al., 1996) and M. malmoense genomic DNA as template revealed an insertion following codon 130 of gyrA (Perler et al., 1997); the insertion comprises 1260 bp, encoding a putative intein of 420 aa. Comparison with gyrA genes of other mycobacterial species indicates a conservation of size (420 aa in M. leprae, M. kansasii type I and M. gordonae and 421 aa in M. flavescens), the N-terminal amino acid (Cys) and the C-terminal splice junction (His Asn/Thr). These GyrA inteins are much longer than the M. xenopi GyrA intein (198 aa), which lacks homing endonuclease activity and has undergone a complex series of recombination events (Telenti et al., 1997).

Phylogenetic trees based on 16S rRNA sequences indicate a closer relationship of M. malmoense, M. gordonae and M. kansasii to one another than to M. flavescens or M. leprae. A similar relationship was found when the GyrA intein sequences were compared (data not shown).

From our analysis, we draw the conclusion that the presence of a GyrA intein is not random but is a taxonomic character. This character defines a mycobacterial species either at a species or at a subspecies level depending on the degree of genetic homogeneity within a given taxon. Genetically homogeneous species such as M. malmoense, M. marium, M. ulcerans, M. tuberculosis and M. leprae uniformly by the presence or absence of a GyrA intein. M. xenopi is characterized by its unique short gyrA element. In genetically more heterogeneous species, such as M. kansasii – and probably M. flavescens and M. gordonae (Fsihi et al., 1996) – the presence or absence of an intein coding sequence in gyrA corresponds to a subspecies characteristic. Our findings do not imply that the intein itself – although a valid taxonomic marker – is useful for defining phylogenetic relationships as the underlying mechanism of acquisition or loss of that marker remains to be determined.

ACKNOWLEDGEMENTS

We are indebted to Francoise Portaels and Gaby Pfyffer for providing strains. We thank Kerstin Teschner for excellent technical assistance.

This work was supported in part by grants from the Bundesministerium für Bildung, Wissenschaft, Forschung und Technik (Verbund Mykobakterielle Infektionen). E.C.B. is a Hermann- und Lilly-Schilling Professor of the Stifterverband für die Deutsche Wissenschaft. A. T. is supported by the Swiss National Science Foundation.

REFERENCES


Received 1 April 1997; revised 23 September 1997; accepted 1 October 1997.