Volume 240, number 1,2, 115–117 FEB 06511 November 1988

Immunogenic protein MPB57 from *Mycobacterium bovis* BCG: molecular cloning, nucleotide sequence and expression

Ryuji Yamaguchi, Kazuhiro Matsuo, Akihiro Yamazaki, Sadamu Nagai[+], Kunihiro Terasaka[°] and Takeshi Yamada[°]

Central Research Laboratories, Ajinomoto Co. Inc., Suzuki-cho 1-1, Kawasaki-ku, Kawasaki 210, [+]Toneyama Institute for Tuberculosis Research, Osaka City University Medical School, Toneyama 5-1-1, Toyonaka 560 and *[°]Research Institute for Microbial Diseases, Osaka University, Yamada-oka 3-1, Suita, Osaka 565, Japan*

Received 30 August 1988; revised version received 5 October 1988

Using a single-probe method, we have cloned the gene for an immunogenic MPB57 protein of *Mycobacterium bovis* BCG. The nucleotide sequence includes an ORF of 300 base pairs encoding a protein of 99 amino acids with an M_r of 10 818. This cloned gene was expressed in an *Escherichia coli* expression vector to give a mature protein which reacted with a polyclonal antibody raised against MPB57.

Mycobacterial antigen; Single-probe method; Amino acid sequence; DNA sequence

1. INTRODUCTION

Interest in mycobacterial proteins has increased in recent years because it has been reported that the epitope of 65 kDa *Mycobacterium bovis* BCG protein seems responsible for the immunopathology of arthritis [1], and because some extracellular proteins from mycobacteria may offer an attractive possibility for use in serological assays for the detection of mycobacterial diseases.

Using a single-probe method [2] which has great potential for the isolation of mycobacterial protein genes, we have recently cloned two genes encoding α antigen [2] and MPB64 [3] derived from *M. bovis* BCG, both of which were expressed in *E. coli*.

Correspondence address: R. Yamaguchi, Central Research Laboratories, Ajinomoto Co. Inc., Suzuki-cho 1-1, Kawasaki-ku, Kawasaki 210, Japan

Abbreviations: aa, amino acid(s); PPD, purified protein derivatives; ORF, open reading frame; SD sequence, Shine-Dalgarno sequence; SDS-PAGE, SDS-polyacryl amide gel electrophoresis; IPTG, isopropyl-β-D-thiogalactopyranoside; Amp, ampicillin

The MPB57 protein was isolated by one of the authors from the culture filtrates of *M. bovis* BCG-Tokyo and *M. tuberculosis* H37Rv and found in the cell-sonicates of these species. A similar component can be found in the PPD prepared after heat sterilization. In this communication we report the cloning, sequencing, and expression in *E. coli* of an approx. 12 kDa MPB57 protein gene. It was also shown that this protein reacted with the monoclonal antibody IT3 [4].

2. MATERIALS AND METHODS

2.1. *Synthesis and design of oligonucleotide probes*

About 15 μg of purified MPB57 was analyzed by a 470A gas-phase protein sequencer (Applied Biosystems Inc.) and the sequence of 20 N-terminal aa was determined (fig.1). Oligonucleotide probes (probe I: 21-mer; probe II: 53-mer) corresponding to the N-terminal aa sequence were chemically synthesized by a 370A automated DNA synthesizer (Applied Biosystems Inc.), using the most frequently used codons which were selected on the basis of the reported mycobacterial protein sequences (65 kDa [5], α antigen [2], and MPB64 [3]). The codon usage of these proteins is extremely biased to the use of G+C (~80–90%) in the third position. About 100 pmol of purified oligonucleotide probes were labeled with ^{32}P by T_4 polynucleotide kinase. The probes were purified with NEN

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

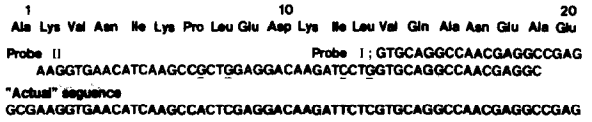

Fig.1. N-terminal aa sequence of MPB57 and oligonucleotide probes. The short probe (21-mer) gave a completely matched (21/21 = 100%) base pair in a synthetic oligonucleotide-DNA duplex by selecting a region of biased codons, and the longer one (53-mer) also gave a high matching rate (49/53 = 92%). The unmatched bases are underlined. Both were extremely efficient in the Southern and the colony hybridizations.

SORB (Du Pont Co., Ltd) and then used for the hybridization tests.

2.2. *DNA cloning*

M. bovis BCG DNA was digested with restriction enzyme *Pst*I and size-fractionated on a 0.8% agarose gel by electrophoresis. Since a *Pst*I fragment with a length of 9.0 kbp gave a relatively clear hybridized band with probe I, it was harvested, cloned into pUC18, and transformed into *E. coli* JM109. One positive colony was obtained by colony hybridization with probes I and II.

2.3. *DNA sequencing*

The *Pst*I fragment from the mycobacterial DNA insert was digested with *Bam*HI and *Xho*I to get a *Bam*HI fragment and an *Xho*I fragment, respectively. Then the fragments were cloned into pUC18 digested with *Bam*HI and *Sal*I. The nucleotide sequence of the subcloned DNA fragments was determined by the dideoxy chain termination method [6], using intact pUC plasmid [7] and *E. coli* bacteriophage M13 mp19 [8].

3. RESULTS AND DISCUSSION

The N-terminal aa sequence of this protein was identical to that of BCG-a of Minden et al. [10] except for one aa; Glu-15 of BCG-a was replaced by Gln-15 in MPB57. The DNA sequence contained an ORF beginning with GTG at position 1–3 and ending with the double stop codons TAGTAG at position 301–306 (fig.2). Its theoretical M_r was calculated to be 10818. Examination of the DNA sequence revealed that this protein lacked a signal peptide sequence, suggesting that it might be originally an intracellular protein but released into the medium by lysis during cultivation; the possibility of excretion without the signal peptide remains to be elucidated. Sequences closely resembling the −10 and −35 regions for *E. coli* promoters were at positions −173 to −168 (TATAGA) and −194 to −189 (TTGAGT) upstream of the ORF [11]. There are 15 bases between these putative promoter elements, which is in agreement with the 15 to 20 base spacing described for *E. coli* promoters [11]. The SD sequence GGAGG can be seen 9 bases upstream of the start codon GUG. Besides a preference for the use of G and C in the third codon position, there is a similar strong bias in codon usage patterns (table 1), compared with those of α antigen [2] and

Fig.2. Nucleotide and deduced aa sequences of MPB57. The sequence is numbered beginning with the first nucleotide of the start codon GTG. The −35 and −10 regions of the proposed promoters are underlined. A possible SD sequence is boxed.

Table 1

Codon usage for an MPB57 gene

TTT	Phe	0	TCT	Ser	0	TAT	Tyr	0	TGT	Cys	0
TTC		1	TCC		1	TAC		4	TGC		0
TTA	Leu	0	TCA		0	TAA	***	0	TGA	***	0
TTG		0	TCG		1	TAG		1	TGG	Trp	1
CTT		0	CCT	Pro	2	CAT	His	0	CGT	Arg	2
CTC		2	CCC		0	CAC		0	CGC		1
CTA		0	CCA		1	CAA	Gln	0	CGA		0
CTG		4	CCG		2	CAG		2	CGG		2
ATT	Ile	2	ACT	Thr	0	AAT	Asn	0	AGT	Ser	0
ATC		5	ACC		6	AAC		3	AGC		1
ATA		0	ACA		0	AAA	Lys	0	AGA	Arg	0
ATG	Met	0	ACG		1	AAG		9	AGG		0
GTT	Val	3	GCT	Ala	0	GAT	Asp	0	GGT	Gly	2
GTC		5	GCC		4	GAC		7	GGC		8
GTA		0	GCA		1	GAA	Glu	1	GGA		0
GTG		2	GCG		3	GAG		10	GGG		0

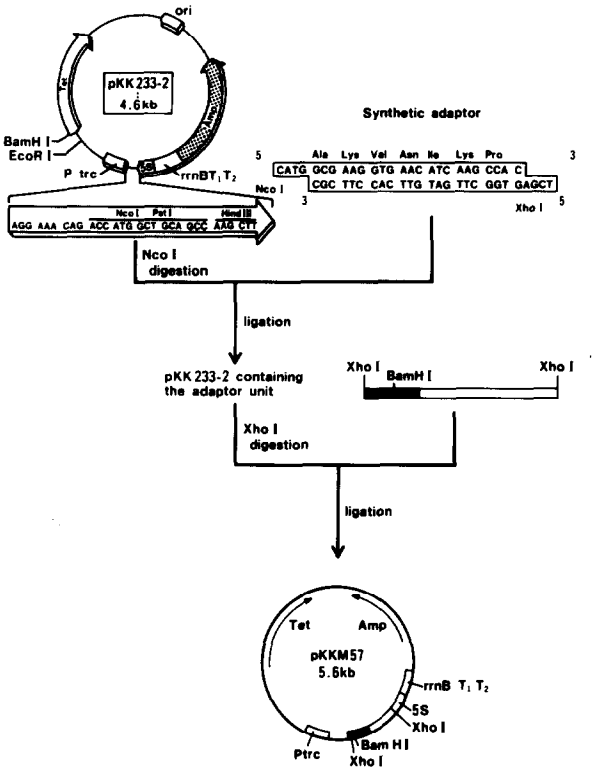

Fig.3. Construction of *E. coli* expression vector pKKM57. The 22-bp *Nco*I-*Xho*I synthetic adaptor encoding the N-terminus (7 aa) of MPB57 was ligated to a pKK233-2 plasmid vector [9] digested with *Nco*I. After digesting with *Xho*I, the linearized plasmid was linked to an *Xho*I fragment containing all but the terminal 7 aa of the MPB57 gene, creating pKKM57.

MPB64 [3]. The G + C content of the third letter is 85%. Almost all codons which were not used possessed A or T at the third position. Interestingly in the case of Leu, only CTC and CTG were used among synonymous codons. It seems likely that CTT and CTA were converted by the so-called GC pressure [12] during evolution to the codons having G or C, since the mutation of T to C or A to G is silent.

The protein expressed in *E. coli* had a mobility in SDS-PAGE identical to that of authentic MPB57, corresponding to a molecular mass of 12 kDa (fig.4). An interesting point is that the MPB57 protein was produced in much larger quantity (~5% of the total proteins) than α antigen [2] or MPB64 [3]. The hydropathy profile (not shown) showed three hydrophilic regions around aa 30–40, 50–60, and 75–85, which might be regarded as exposed epitopes. The mature MPB57

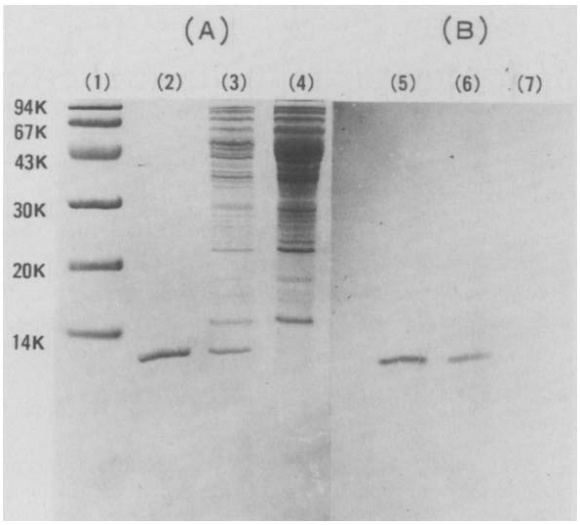

Fig.4. SDS-PAGE and immunoblot analysis of an MPB57 protein expressed in *E. coli*. Protein bands are visualized by Coomassie brilliant blue staining (A) or by Western blot analysis (B) using a rabbit antiserum raised against MPB57. Lanes: (1) molecular mass markers; (2) and (5), the authentic MPB57 protein; (3) and (6), *E. coli* JM109 [pKKM57]; (4) and (7), no plasmid.

protein and some poly eric epitopes would be good candidates for ser liagnosis [13].

REFERENCES

[1] Van Eden, W., Thole, J.E.R., Van der Zee, R., Noordzi, A., Van Embden, J.D.A., Hensen, E.J. and Cohen, I.R. (1988) Nature 331, 171–173.
[2] Matsuo, K., Yamaguchi, R., Yamazaki, A., Tasaka, H. and Yamada, T. (1988) J. Bacteriol., in press.
[3] Yamaguchi, R., Matsuo, K., Yamazaki, A., Abe, C., Nagai, S., Terasaka, K. and Yamada, T. (1988) Infect. Immun., submitted.
[4] Engers, H.D. et al. (1986) Infect. Immun. 51, 718–720.
[5] Shinnick, T.M. (1987) J. Bacteriol. 169, 1080–1088.
[6] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[7] Hattori, M. and Sakaki, Y. (1986) Analytical Biochem. 152, 232–238.
[8] Messing, J. (1983) Methods Enzymol. 101, 20–78.
[9] Amann, E. and Brosius, J. (1985) Gene 40, 183–190.
[10] Minden, P., Kelleher, P.J., Freed, J.H., Nielsen, L.D., Brennan, P.J., McPheron, L. and McClatchy, J.K. (1984) Infect. Immun. 46, 519–525.
[11] Hawley, D.K. and McClure, W.R. (1983) Nucleic Acids Res. 11, 2237–2255.
[12] Osawa, S., Ohama, T., Yamao, F., Muto, A., Jukes, T.H., Ozeki, H. and Umesono, K. (1988) Proc. Natl. Acad. Sci. USA 85, 1124–1128.
[13] Minden, P., Voegtline, M.S. and Houghten, R.A. (1987) J. Clin. Lab. Anal. 1, 287–292.