Cloning and Overexpression in Soluble Form of Functional Shikimate Kinase and 5-Enolpyruvylshikimate 3-Phosphate Synthase Enzymes from Mycobacterium tuberculosis

Jaim S. Oliveira, Clotilde A. Pinto, Luiz A. Basso, and Diógenes S. Santos
Departamento de Biologia Molecular e Biotecnologia, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves 9500, Porto Alegre-RS 91501-970, Brazil

Received January 26, 2001, and in revised form April 3, 2001; published online July 20, 2001

In 1997, a total of 7.96 million new cases of tuberculosis (TB) were estimated, there were 16.2 million existing cases of the disease, and an estimated 1.87 million people died of TB (1). The reemergence of TB as a public health threat, the high susceptibility of human immunodeficiency virus-infected persons to the disease, and the proliferation of multidrug-resistant (MDR) strains have created much scientific interest in developing new antimycobacterial agents to both treat Mycobacterium tuberculosis strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance. Moreover, treatment of patients infected with MDR M. tuberculosis must rely on second-line drugs, which are less effective and more expensive, can cost up to $250,000 per person and take 2 years (2).

The shikimate pathway is an attractive target for herbicides and antimicrobial agents development because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals (3). Moreover, biochemical, genetic, and chemotherapeutic evidence presented for the existence of a functional shikimate pathway in apicomplexan parasites should provide attractive targets for the development of new antiparasite agents (4, 5). In mycobacteria, the shikimate pathway leads to the biosynthesis of chorismic acid, which is a precursor for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactin (6).

Evidence that the shikimate pathway is essential in M. tuberculosis is lacking, although the disruption of the aroD gene was used successfully to generate attenuated oral vaccine strains of Salmonella typhi and other bacteria (7). Furthermore, the salicylate-derived mycobactin siderophores were recently shown to be essential for M. tuberculosis growth in macrophages (8).

Homologues to the shikimate pathway enzymes were identified in the complete genome sequence of M. tuberculosis H37Rv strain (9). Among them, the shikimate
kinase I (SK, EC 2.7.1.71) encoding gene (aroK, Rv2539c) and 5-enolpyruvylshikimate 3-phosphate synthase (EPSPS, EC 2.5.1.19) encoding gene (aroA, Rv3227) were proposed to be present by sequence homology. SK catalyzes a phosphate transfer from ATP to the carbon-3 hydroxyl group of shikimate forming shikimate-3-phosphate (S3P) and EPSPS catalyzes an unusual transfer of an enolpyruvyl moiety from phosphoenolpyruvate (PEP) to S3P with the elimination of inorganic phosphate forming 5-enolpyruvylshikimate-3-phosphate (EPSP) (10). Incidentally, EPSPS is the target of glyphosate [N-(phosphonomethyl)glycine], which is a widely used broad-spectrum herbicide (11).

In order to pave the way for structural and functional efforts currently under way in our laboratory, and assess the feasibility of utilizing SK and EPSPS as potential targets for antimycobacterial agents development, the aroK and aroA genes from M. tuberculosis H37Rv strain were PCR amplified, cloned, sequenced, and overexpressed in Escherichia coli BL21(DE3) host cells. Protein overexpression could be achieved by a low-cost and simple protocol. The activity of the soluble enzymes was measured by coupled assays in crude extracts, confirming the correct assignment to the structural genes encoding SK and EPSPS in M. tuberculosis. We hope that the work here presented may contribute to current efforts toward the development of new antimycobacterial agents.

MATERIALS AND METHODS

PCR Amplification and Cloning of aroK and aroA genes

Synthetic oligonucleotide primers (5'-attccatatggaagatctctggccagcccc-3' and 5'-gcggatccatcaaacgccgaccccctgcc-3', for aroA gene; 5'-attccatatgccaccaaggggtctc-3' and 5'-ggcggatctctatgtgcccgcctgcgtgg-3', for aroK gene) were designed based on the complete genome sequence of M. tuberculosis H37Rv (9). These two pairs of primers were complementary to the amino-terminal coding and carboxyl-terminal noncoding strands of their respective structural genes containing 5' NdeI and 3' BamHI restriction sites, which are underlined, in noncomplementary overhangs. These primers were used to amplify the M. tuberculosis aroA and aroK structural genes from genomic DNA using standard PCR conditions (Perkin-Elmer). The PCR products were purified by electrophoresis on low-melting agarose, digested with NdeI and BamHI and ligated into a pET23a(+) expression vector (Novagen) which had previously been digested with the same restriction enzymes. The DNA sequences of the amplified M. tuberculosis aroA and aroK genes were determined, using the Thermo Sequenase radiolabeled terminator cycle sequencing kit (Amersham Pharmacia Biotech), to both confirm the identity of the cloned genes and assess that no mutations were introduced by the PCR amplification step.

Overexpression of SK and EPSPS

The recombinant plasmids, either pET-23a(+):aroK or pET-23a(+):aroA, were transformed into electrocompeent Escherichia coli BL21(DE3) cells and selected on LB agar plates containing 50 μg mL⁻¹ carbenicillin. Single colonies were used to inoculate 1.5 L of LB medium containing 50 μg mL⁻¹ carbenicillin, grown at 37°C until OD₆₀₀ ranged between 0.4 and 0.6, and either induced by addition of 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) or not induced. SK and EPSPS culture cells were grown for an additional 23 h at 37°C in shaker flasks. Cells were harvested by centrifugation at 6000 g for 20 min at 4°C and stored at −80°C.

For protein overexpression analysis, stored cells were suspended in 700 μL of 50 mM Tris–HCl, pH 7.8 (40 mg of cells/mL of buffer) and lysed by sonication, and the cell debris was removed by centrifugation at 20,800 g (30 min at 4°C). The subunit molecular weights of soluble protein extracts were analyzed by SDS–PAGE (Laemmli method) (12). EPSPS extract for enzyme activity measurements was prepared with 50 mM Tris–HCl, pH 7.8. Cell extract for SK activity analysis was suspended in 50 mM Tris–HCl, 50 mM KCl, 5 mM MgCl₂, pH 7.5. Control experiments were performed under the same experimental conditions except that transformed E. coli cells harbored the expression vector lacking the target genes. The proportions of SK and EPSPS to total soluble proteins in the SDS–PAGE gels were estimated using a GS-700 imaging densitometer (Bio-Rad).

In order to evaluate EPSP synthase protein expression as a function of cell growth phase, E. coli BL21(DE3) cells harboring pET-23a(+:aroK recombinant plasmid were grown in the absence of IPTG and samples were removed at various times for OD₆₀₀ measurements and for SDS–PAGE analysis.

Protein Determination

Protein concentration was determined by the method of Bradford et al. (13) using the Bio-Rad protein assay kit (Bio-Rad) and bovine serum albumin as standard.

Shikimate Kinase Activity Assay

Shikimate kinase enzyme activity was assayed in the forward direction by coupling the ADP product formation to the pyruvate kinase (PK; EC 2.7.1.40) and lactate dehydrogenase (LDH; EC 1.1.1.27) reactions following the protocol described by Millar et al. (14). Shikimate-dependent oxidation of NADH was continuously monitored at 340 nm (ε = 6.22 × 10³ M⁻¹ cm⁻¹).
All reactions were carried out at 25°C and initiated with addition of SK extracts. The assay mixture contained 100 mM Tris-HCl buffer, pH 7.5, 50 mM KCl, 5 mM MgCl₂, 1.6 mM shikimic acid, 2.5 mM ATP, 1 mM phosphoenolpyruvate, 0.1 mM NADH, 3 U mL⁻¹ pyruvate kinase, and 2.5 U mL⁻¹ units of lactate dehydrogenase. Initial steady-state rates were calculated from the linear portion of the reaction curve relative to rates obtained likewise with extracts of E. coli BL21(DE3) cells harboring pET-23a(+) plasmid. PEP, (-)-shikimate, LDH and PK enzymes, ATP, and NADH were all purchased from Sigma. One unit of enzyme activity (U) is defined as the amount of enzyme catalyzing the conversion of 1 μmol of substrate per minute.

EPSP Synthase Activity Assay

Recombinant M. tuberculosis EPSPS was assayed in the forward direction with 50 mM Tris-HCl, 50 mM KCl, 2.5 mM MgCl₂, pH 7.6, at 25°C. Enzyme activity was measured by estimating the rate of inorganic phosphate (P_i) release using a continuous spectrophotometric coupled assay with purine nucleoside phosphorylase (PNP; EC 2.4.2.1) and 2-amino-6-mercapto-7-methylpurinenucleoside (MESG) (15). The PNP enzyme and MESG are commercially available as the EnzChek phosphate assay kit (Molecular Probes). This assay is based on the difference in absorbance between MESG and the assay kit (Molecular Probes). This assay is based on are commercially available as the EnzChek phosphate assay kit (Molecular Probes). This assay is based on the difference in absorbance between MESG and the

RESULTS AND DISCUSSION

The nucleotide sequence analysis using the dideoxy-chain termination method of the PCR-amplified M. tuberculosis aroK (531 bp) and aroA (1353 bp) genes confirmed the identity of the coding DNA sequence of cloned genes and demonstrated that no mutation was introduced by the PCR amplification step.

Shikimate kinase and EPSP synthase from M. tuberculosis H37Rv were overexpressed in E. coli BL21(DE3) cells carrying pET-23a(+)::aroK and pET-23a(+)::aroA recombinant plasmids, respectively. Analysis by SDS-PAGE with Coomassie blue staining indicated that the cell extracts contained a significant amount of protein with subunit molecular weight in agreement with the predicted MWs for shikimate kinase (Fig. 1A, 18.58 kDa) and EPSP synthase (Fig. 1B, 46.43 kDa). Densitometric quantification of the SDS-PAGE protein bands showed that both SK and EPSPS constituted approximately 30% of total protein present in the soluble cell extract under the experimental conditions utilized. The value of 30% is an underestimate for EPSPS protein overexpression, since a lower load of protein on SDS-PAGE indicated a proportion of ~40% EPSPS to total soluble protein (data not shown).

It should be pointed out that overexpression of both proteins was achieved without addition of IPTG and by

![FIG. 1. SDS-PAGE analysis of protein-soluble crude extracts. Overexpression of shikimate kinase and EPSP synthase from M. tuberculosis. All samples were grown in LB liquid medium containing 50 μg mL⁻¹ carbenicillin at 37°C to an OD₆₀₀ between 0.4 and 0.6 and either induced by addition of 0.5 mM IPTG or not induced and then grown for an additional 23 h. (A) 15% SDS-PAGE. Lanes 1 and 2, E. coli BL21(DE3) [pET-23a(+)::aroK] cells uninduced (control); lanes 3 and 4, E. coli BL21(DE3) [pET-23a(+)::aroK] cells induced; lane 5, MW markers; lanes 6 and 7, E. coli BL21(DE3) [pET-23a(+)::aroK] cells induced; lanes 8 and 9, E. coli BL21(DE3) [pET-23a(+)::aroK] cells induced.]
growing cells to high saturation density, as described under Materials and Methods and in the legend to Fig. 1. M. tuberculosis EPSP synthase protein expression as a function of cell growth phase in the absence of IPTG could be detected at all time intervals tested (Fig. 2). However, EPSPS expression reached a maximum in the stationary phase (Fig. 2B; lane 6, 14.7 h of cell growth) and high levels of expression remained throughout the time intervals tested in the stationary phase (Fig. 2B).

In the pET vector system (Novagen), target genes are positioned downstream of the bacteriophage T7 late promoter on medium copy number plasmids. Typically, production hosts contain a prophage (λDE3) encoding the highly processive T7 RNA polymerase under control of the IPTG-inducible lacUV5 promoter. Although it is often argued that the cost of IPTG limits the usefulness of the lac promoter to high-added-value products, we showed here that high levels of expression of SK and EPSPS could be obtained in this system without addition of IPTG at midlog phase in LB medium. Protein overexpression was obtained at higher levels for SK without IPTG induction compared to 0.5 mM IPTG induction and at an equal level for EPSPS (data not shown). It should also be noted that only fresh transformants containing pET-23a(+)::aroA produced EPSPS protein at high levels. EPSPS protein production levels declined considerably when transformants were stored either on agar plates at 4°C or as glycerol stocks at −20°C. Similar experimental observations were reported by Kelley et al. (16) using the pET system. Furthermore, high levels of uninduced protein production could be achieved by cell growth to stationary phase (OD$_{600}$ = 1.6–1.7) at 37°C for both the M. tuberculosis EPSP synthase (Fig. 2) and SK (data not shown) constructs. It was recently proposed that leaky protein expression is a property of the lac-controlled system as cells approach stationary phase in complex medium and

![Graph showing growth curve of uninduced E. coli BL21(DE3) pET-23a(+)::aroA cells.](image)

**FIG. 2.** Recombinant EPSP synthase protein expression increases at the stationary phase. E. coli BL21(DE3) cells harboring pET-23a(+)::aroA recombinant plasmid were grown at 37°C in LB liquid medium containing 50 μg mL$^{-1}$ carbenicillin and samples were removed at indicated time intervals for OD$_{600}$ measurements and SDS–PAGE analysis. (A) Growth curve of uninduced E. coli BL21(DE3) [pET-23a(+)::aroA] cells indicating time intervals of sample removal for OD$_{600}$ measurements (□) and SDS–PAGE analysis of total soluble protein (■). (B) 12.5% SDS–PAGE of total soluble protein (25 μg was loaded on each lane). Lane 1, 6 h; lane 2, 6.8 h; lane 3, 8.3 h; lane MW, protein markers; lane 4, 9.4 h; lane 5, 10.5 h; lane 6, 14.7 h; lane 7, 18.4 h; lane 8, 24.5 h; lane 9, 29 h.
that cyclic AMP, acetate, and low pH are required to effect high-level expression in the absence of IPTG induction (17). The authors also suggested that derepression of the lac operon in the absence of IPTG may be part of a general cellular response to nutrient limitation.

The soluble protein cellular extracts were assayed for SK and EPSPS enzyme activities. EPSPS-specific enzyme activity from cells expressing aroA gene product was determined to be 101-fold higher than the specific activity obtained for E. coli BL21(DE3) cells transformed with pET-23a(+) (control), as shown in Table 1. Analysis of SK-specific enzyme activity showed a 328-fold increase (see Table 1) compared to control cell extracts. Moreover, SK and EPSPS enzyme activities were linearly dependent on cell extract volume added to the reaction mixture. The SK enzyme activity was stable, since storage of crude extracts for 5 days at 4°C showed no decline in enzyme activity (data not shown).

In this report, we presented the PCR amplification, cloning, sequencing, and overexpression of SK and EPSPS from M. tuberculosis. The target proteins reached a native conformation as could be verified by enzyme activity measurements. Moreover, SK and EPSPS proteins comprised at least 30% of total soluble cell protein. To the best of our knowledge, this is the first report of cloning and overexpression both in soluble and in active forms of M. tuberculosis SK and EPSPS enzymes. Although cloning of the aroA gene from M. tuberculosis has been reported by Garbe et al. (18), only 345 amino acids from the aroA gene product were produced as an insoluble fusion protein to glutathione S-transferase in order to raise antibodies. Therefore, M. tuberculosis EPSPS was only characterized by immunological analysis and not by functional studies. The three-dimensional structure of E. coli EPSPS has been determined to 3.0 Å in the absence of any ligands (19).

As previously pointed out, EPSPS is the target of glyphosate, which is a broad-spectrum herbicide widely used (11). However, more complete structural studies of the EPSPS–S3P–glyphosate and EPSPS–EPSP-glyphosate ternary complexes will be needed to assist inhibitor design for the proposed allosteric glyphosate-binding domain (20). Accordingly, the results presented here should provide protein in quantities necessary for structural and enzymological studies. Our goal is to develop antimycobacterial agents; however, we also hope to contribute to the development of new potent herbicides.

ACKNOWLEDGMENTS

This work was supported by PADCT, CNPq, and FAPERGS (Brazil) grants awarded to L.A.B. and D.S.S. We thank Professor John S. Blanchard for generously supporting this work and Professor William R. Jacobs for his generous gift of genomic DNA from M. tuberculosis H37Rv.

REFERENCES


