PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07H 21/04, 21/00, C07K 1/00, C12N 1/19, 9/00, 9/02, 15/00, C12P 21/00, C12Q 1/00, 1/26, 1/68

A1

(11) International Publication Number: **WO 98/29432**

(43) International Publication Date: 9 July 1998 (09.07.98)

(21) International Application Number: PCT/US97/23557

(22) International Filing Date: 19 December 1997 (19.12.97)

(30) Priority Data:

60/034,725	3 January 1997 (03.01.97)	US
60/039,737	14 February 1997 (14.02.97)	US
08/990,791	15 December 1997 (15.12.97)	US

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant: SOUTHERN RESEARCH INSTITUTE [US/US]; 2000 North Avenue, South, Birmingham, AL 35205 (US).

(72) Inventors: BARROW, William, W.; 460 Russet Hill Road, Hoover, AL 35244 (US). VAN GINKEL, Sabrina, Z.; 504 Morris Boulevard, Homewood, AL 35209 (US). DOOLEY, Thomas, P.; 839 Mountain Branch Lane, Vestavia Hills, AL 35226 (US). SULING, William, J.; 537 Canterbury Road, Pehlam, AL 35124 (US).

(74) Agents: REMENICK, James et al.; Baker & Botts, L.L.P., The Warner, 1299 Pennsylvania Avenue, N.W., Washington, DC 20004 (US).

(54) Title: MYCOBACTERIUM FOL A GENE THAT ENCODES FOR THE ENZYME DIHYDROFOLATE REDUCTASE

(57) Abstract

The invention relates to the nucleic acid sequence and amino acid sequence of dihydrofolate reductase (DHFR) from mycobacteria and to expression of recombinant DHFR protein. Utilizing the recombinant protein, novel therapies and diagnostic strategies can be developed and selective antimycobacterial compositions can be designed and utilized to treat mycobacterial infections in patients. This invention includes all or portions of novel recombinant nucleic acids encoding DHFR for mycobacteria such as *M. avium*, to novel recombinant DHFR peptides produced by such sequences, and to vaccines, diagnostic kits, cells and therapies utilizing these peptides and nucleic acid sequences. The present invention relates to methods for using the sequences of the present invention to develop drugs specific to *M. avium* and other mycobacteria, to identify and sequence corresponding sequences in species other than *M. avium*, as well as diagnostic and treatment methods incorporating the disclosed sequences and peptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

MYCOBACTERIUM FOL A GENE THAT ENCODES FOR THE ENZYME DIHYDROFOLATE REDUCTASE

Background of the Invention

1. Field of the Invention

This invention relates to novel recombinant nucleic acids encoding the enzyme dihydrofolate reductase (DHFR) from mycobacteria, to novel recombinant DHFR peptides produced by such sequences, and to vaccines, diagnostic kits, cells and therapies utilizing these peptides and nucleic acid sequences. The invention is also directed to methods for using the sequences and peptides to develop drugs specific to *M. avium* and other species of mycobacteria, to identifying other DHFR sequences and peptides, as well as diagnostic and treatment methods incorporating the disclosed sequences and peptides.

2. Description of Background.

The *Mycobacterium avium* complex represents one of the most serious opportunistic infections and is often associated with advanced stages of autoimmune deficiency syndrome or AIDS (J.J. Ellner et al., J. Infect. Dis. 163:1326-35, 1991; J.A. Havlok, Jr. et al., J. Infec. Dis. 165:577-80, 1992; C.C. Hawkins et al., Ann. Intern. Med. 105:184-88, 1986; D.S. O'Brien et al., Am. Rev. Respir. Dis. 135:1007-14, 1989; N. Rastogi et al., Res. Microbiol. 145:167-261, 1994). Unlike *Mycobacterium tuberculosis,* which can be successfully treated with two or three drug combinations (except for multidrug resistant *M. tuberculosis;* MDR-TB), the *M. avium* complex is resistant to many antimycobacterial agents (B.D. Agins et al., J. Infect. Dis. 159:784-87, 1989; C. Benson et al., Sixth International Conference on AIDS, San Francisco, 1990; J. Chiu et al., Ann. Intern. Med. 113:358-61, 1990; F. De Lalla et al., Antimicrob. Agents Chemother. 36:1567-69, 1992; L. Heifets et al., Antimicrob. Agents Chemother. 37:2364-70, 1993; D.Y. Rosenzweig, Amer. Rev. Resp. Dis. 113(Suppl.):55,

1976). Drug resistance in *M. avium* is still considered an inherent property of the wild type organism (S.L. Morris et al., Complex. Res. Microbiol. 147:68-73, 1996), resulting in large part from the refractory nature of the organism's cell envelope (H.L. David, Rev. Infect. Dis. 3:878-84, 1981; N. Rastogi et al., Res. Microbiol. 145:243-52, 1994; N. Rastogi et al., Antimicrob. Agents Chemother. 20:666-77, 1981). Although *M. avium* infections in AIDS patients are treated with 3-6 different drugs, the long term prognosis is still poor (B.D. Agins et al., J. Infect. Dis. 159:784-87, 1989; C. Benson et al., Sixth International Conference on AIDS, San Francisco, 1990; J. Chiu et al., Ann. Intern. Med. 113:358-61, 1990; F. De Lalla et al., Antimicrob. Agents Chemother. 36:1567-69, 1992; J.J. Ellner et al., J. Infect. Dis. 163:1326-35, 1991).

Tuberculosis is a disease of worldwide significance and notoriety. At any one time, about one-third of the world is infected with *M. tuberculosis* resulting in eight million new cases of tuberculosis and 2.9 million deaths annually (A. Arachi, Tubercle. 72:1-6, 1991). It is estimated that about 0.3% of U.S. residents are infected and at risk to develop active disease (CDC 1996, CDC Revises HIV Infection Estimates. HIV/AIDS Prevention. August:2). This risk becomes even greater if the person is co-infected with the human immunodeficiency virus (HIV). If so, estimates indicate that progression to tuberculosis will occur in about 30% of those cases and the risk for developing tuberculosis becomes 113 times greater (tuberculosis., N.a.p.t.c.m.-r., MMWR. 41 RR-11:1-71, 1992).

As the projected figure for HIV infections is more than 20 million by the year 2000, it is probable that the number of tuberculosis cases worldwide will also increase. Even in 1991, the figure for people co-infected with HIV and *M. tuberculosis* was estimated to be 3.1 million (J.F. Murray, Bull. Int. Union Tuberc. Lung Dis. 66:21-15, 1991). In addition,

life-threatening strains of MDR-TB are appearing. Some of these strains can result in a high mortality rate (e.g. 72-89%), with death occurring in a short period (e.g. 4-16 weeks) (CDC, Mortal. Morbid. Weekly Rep. 39:718-22, 1990; CDC, Mortal. Morbid. Weekly Rep. 40:649-652, 1991; B.R. Edlin et al., New Engl. J. Med. 326:1514-21, 1992). In summary, the impact of tuberculosis on the world today can best be appreciated by the fact that the World Health Organization declared tuberculosis a global public health emergency, a distinction never before given to any other disease (WHO., Soz Praventivmed. 38:251-52, 1993). Consequently, the development of new antimycobacterial drugs is an important research endeavor.

The dihydrofolate reductase (DHFR) enzyme is an important target for medicinal chemistry (K. Bowden et al., J. Chemother. 5:377-88, 1993) and DHFR inhibitors have been used in anticancer therapy (e.g. methotrexate (W.A. Bleyer, Cancer Treat. Rev. 41:36-51, 1978)), antibacterial therapy (e.g. trimethoprim (M. Finland et al., J. Infect. Dis. 128:S425-816, 1973)), and antimalarial therapy (e.g. pyrimethamine (A.K. Saxena, Prog. Drug Res. 30:221-80, 1986)). Dihydrofolate reductase is present in all cells and is necessary for the maintenance of intracellular folate pools in a biochemically active reduced state (M. McCourt et al., J. Am. Chem. Soc. 113:6634-39, 1991). Inhibition of the enzyme is effective because binding affinities for substrate analogs are so great that such analogs are not readily displaced by the natural substrates. Enzyme inhibition results in the depletion of intracellular reduced folates that are required for one carbon transfer reactions, which in turn are important for the biosynthesis of thymidylate, purine nucleotides, methionine, serine, glycine and many other compounds needed for RNA, DNA, and protein synthesis. Figure 1 depicts DHFR's role in the biosynthesis of tetrahydrofolate and cell metabolism. (P.G. Hartman, J. Chemother. 5:369-76, 1993). Some bacteria have an

uptake system for folates, but most have to synthesize folates *de novo* by reduction of dihydrofolate to tetrahydrofolates.

Although DHFR is not a new drug target, enthusiasm in the development of improved derivatives to inhibit DHFR is very intense, (D.P. Baccanari et al., J. Chemother. 5:393-99, 1993; K. Bowden et al., J. Chemother. 5:377-88, 1993; M. McCourt et al., J. Am. Chem. Soc. 113:6634-39, 1991; J.R. Piper et al., J. Med. Chem. 39:1271-80, 1996; B.I. Schweitzer et al., FASEB. 4:2441-52, 1990; J.K. Seydel, J. Chemother. 5:422-29, 1993), and particularly with regard to mycobacteria (K.H. Czaplinski et al., Eur. J. Med. Chem. 30:779-87, 1995; M. Kansy et al., Eur. J. Med. Chem. 27:237-44, 1992; H.H. Locher et al., Antimicrob. Agents Chemother. 40:1376-81, 1996; S.C.C. Meyer et al., Antimicro. Agents Chemother. 39:1862-63, 1995; R.L. Then, J. Chemother. 5:361-68, 1993). A unique feature of DHFR is the selectivity possible in the design of inhibitors for this target, thus making it an ideal target for antimycobacterial agents using rational and effective drug design. Although genes for DHFR (*fol* A) have been identified in other bacteria, they are not equivalent to the *fol* A gene from *M. avium* or other mycobacteria. The enzyme product of the *M. avium* gene (*i.e.* DHFR) is structurally different from other known DHFRs. For these reasons and others having to do with, for example, toxicity, selective drugs can be designed for individual species. Thus, a selective drug for *M. avium* can be designed in such a way that it will not affect DHFRs in other species, such as humans. One example of the species specific nature of the enzyme is demonstrated by the fact that the diaminopyrimidines, when properly substituted, can be several thousand times more active against bacterial than mammalian DHFR (P.G. Hartman, J. Chemother. 5:369-76, 1993). In addition, there are many possible inhibitors of this enzyme that have not been synthesized or studied (K.

Bowden et al., J. Chemother. 5:377-88, 1993). DHFR also represents an enzyme that has been extensively used in the development of site directed inhibitors based upon X-ray crystallographic and molecular graphic studies (K. Bowden et al., J. Chemother. 5:377-388, 1993; M.P. Bradley, J. Med. Chem. 36:3171-77, 1993; B.J. Denny et al., J. Med. Chem. 35:2315-20, 1992; M. McCourt et al., J. Am. Chem. Soc. 113:6634-39, 1991; B. Roth, FASEB. 45:2765-72, 1986; W.M. Southerland, J. Computer-Aided Molecular Design. 8:113-22, 1994).

An important objective in the future development of DHFR inhibitors will be improving delivery of antimycobacterial drugs. One research program, designed to develop sustained and targeted delivery of first-line antituberculosis drugs using micro-encapsulation techniques has successfully formulated a micro-encapsulated form of rifampicin that shows good release characteristics (W.W. Barrow et al., European Society for Mycobacteriology, Institute Pasteur, Paris, France: 1996:50). Use of this formulation has resulted in reduction in colony forming units (CFUs) in both the *M. tuberculosis* H37Rv infected macrophage and mouse models, suggesting that this technology can also be used for other antimycobacterial drugs including the lipophilic DHFR inhibitors.

The history of the development of antifolates is long and includes significant contributions to that area of anticancer and anti-infective drug discovery. Improved agents against opportunistic infections have recently been developed including a synthetic process to prepare 5-alkyl-5-deaza analogs of antifolates, both classical and the lipophilic types. These analogs have been used in several studies (J.R. Piper et al., J. Med. Chem. 39:1271-80, 1996).

Mycobacterial DHFR has been a target for drug design for about three decades. Two groups in particular have synthesized lipophilic

antifolates targeting this enzyme in mycobacteria. One group published activity results on a small number of 2,4-diamino-6-substituted pteridines, 6-substituted 8-deazapteridine, 5,6-substituted-5-deazapteridines and 5-methyl-6-substituted-5,8-di-deazapteridines (quinazolines) against *M. species* 607 (W.T. Colwell et al., Chemistry and Biology of Pteridines. vol. Elsevier North Holland, Inc., Amsterdam. 215-18, 1979; J.I. DeGraw et al., J. Medicinal Chem. 17:144-46, 1974; J.I. DeGraw et al., J. Medicinal Chem. 17:762-64, 1974). Another group published several papers describing the design and synthesis of substituted 2,4-diamino-5-benzyl pyrimidines (Trimethoprim analogs) active against mycobacterial DHFR from *M. lufu* and screened *in vitro* against *M. lufu, M. tuberculosis and M. marinum* (K.H. Czaplinski et al., Eur. J. Med. Chem. 30:779-87, 1995; M. Kansy et al., Eur. J. Med. Chem. 27:237-44, 1992; J.K. Seydel et al., Chemother. 29:249-61, 1983).

These data support the activity of folate analogs against mycobacteria. Active compounds were obtained in each series of experiments. The quinazoline analogs were targeted as lead compounds. Although quite active against isolated DHFR, these compounds demonstrated poor selectivity. However, it was apparent that selectivity was enhanced by a 5-methyl substitution for both the 5-deaza- and 5,8-di-deazapteridines (quinazolines). Although low toxicity was noted for the quinazolines in mice, equivocal results against *M. leprae* in the mouse foot pad model were noted (W.T. Colwell et al., *Chemistry and Biology of Pteridines*, Elsevier North Holland, Inc., Amsterdam. 215-18, 1979).

One research group has been pursuing trimethoprim analogs such as 2,4-diamino-5-benzylpyrimidine derivatives (M. Kansy et al., Eur. J. Med. Chem. 27:237-44, 1992; J.K. Seydel, J. Chemother. 5:422-29, 1993; J.K. Seydel et al., Chemother. 29:249-61, 1983). Several 4'-modified

derivatives have been synthesized to extend into the glutamate binding region of the enzyme (L.F. Kuyper et al., J. Med. Chem. 28:303-11, 1985). A small number of these compounds showed very good selective activity against *M. lufu in vitro* and against the isolated bacterial DHFR (K.H. Czaplinski, Eur. J. Med. Chem. 30:779-87, 1995). The compounds showed an activity profile against *M. tuberculosis in vitro* (M. Kansy et al., Eur. J. Med. Chem. 27:237-44, 1992).

The efficacy of using antifolates against *M. avium* and other mycobacteria has not yet been fully determined. In addition, as new antifolates are developed, their efficacy against *M. avium* and other mycobacteria will need to be evaluated. Such studies could be greatly facilitated through the use of purified recombinant mycobacterial DHFR.

Thus far, there has not been complete or accurate identification, sequencing and cloning of the mycobacterial DHFR gene of any species. Two papers have been published about mycobacterial DHFR. Al-Rubeai et al. (M. Al-Rubeai et al., Biochem. J. 235:301-3, 1986), reported on the purification and characterization of DHFR from *M. phlei* and Sirawaraporn et al. (W. Sirawaraporn et al., Exper. Parisitol. 72:184-90, 1991), reported on the purification and characterization of DHFR from a strain of *M. smegmatis*. The reported molecular weights were 15 and 23 kDa for *M. phlei*, and *M. smegmatis* DHFR, respectively. In Sirawaraporn et al., the authors reported on the amino terminal sequencing of the protein. Of the fifteen assignments reported, twelve were stated to be clear, one ambiguous and two could not be determined. Other than this incomplete and partial sequencing of *M. segmatis* DHFR protein, no mycobacterial DHFR DNA sequences have previously been reported.

Summary of the Invention

The invention is directed to all or portions of novel recombinant nucleic acids encoding, in whole or in part, a dihydrofolate reductase (DHFR) peptide from mycobacteria such as *M. avium* or *M. tuberculosis*, to novel recombinant DHFR peptides produced by such sequences, and to vaccines, diagnostic kits, cells and therapies utilizing these peptides and nucleic acid sequences. The invention is also directed to methods for using the sequences to develop drugs specific to *M. avium* and other mycobacteria, to identify and sequence corresponding sequences in other species of *Mycobacterium*, as well as to diagnostic and treatment methods incorporating the disclosed sequences and peptides.

One embodiment of the invention is directed to recombinant nucleic acids comprising all or a portion of the nucleic acid sequence that encodes a mycobacterial DHFR protein, such as the DHFR protein of *Mycobacterium avium, Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium leprae*. The nucleic acid may comprise DNA, RNA or PNA, and may include additional sequences to direct transcription or translation, such as a promoter, a polymerase binding site, an enhancer, or a transcription or translation termination site. The nucleic acid may encode portions of the DHFR protein, such as an enzymatically active portion or antigenically active portion. Alternatively, the sequence may encode the entire amino acid sequence of the DHFR protein.

Another embodiment of the invention is directed to vectors comprising one of these recombinant nucleic acids, or a recombinant cell containing one of these nucleic acids. The nucleic acid may be integrated into the cell's genome, or it may be episomal. The cell may be prokaryotic or eukaryotic.

Another embodiment of the invention is directed to recombinant peptides comprising an amino acid sequence containing all or portion of a mycobacterial DHFR protein, such as *M. avium, M. bovis, M. tuberculosis* or *M. leprae* protein. The recombinant peptide may encode only an enzymatically or antigenically active portion of the peptide, or the entire amino acid sequence of the protein.

Another embodiment of the invention is directed to methods for screening for an agent which inhibits the activity of recombinant DHFR. This method comprises determining the activity of DHFR protein upon incubation with a plurality of agents and selecting the agent that inhibits the activity. The DHFR protein may be derived from *M. avium, M. bovis, M. leprae, M. tuberculosis* or other mycobacteria. The DHFR protein may be an entire DHFR protein or comprise only selected enzymatically active or antigenically active portions thereof. The plurality of agents may be, for example, over 10^2 different agents, and may be drawn from a collection of related chemical compounds, including chemical modifications of folate, methotrexate, trimethoprin or combinations thereof. The incubation may comprise mixing the DHFR protein with the agents under conditions allowing for molecular interaction, such as binding, inhibition of enzymatic activity or inhibition of immunogenicity. A preferred embodiment of this method includes the steps of selecting a plurality of agents that inhibit the *Mycobacterium* DHFR protein, determining the molecular conformation of each agent, and identifying a common inhibitory molecular conformation.

Another embodiment of the invention relates to methods for assessing the ability of an agent to inhibit the activity of a DHFR protein comprising the steps of incubating a recombinant mycobacterial DHFR protein with the agent, determining the activity of the incubated DHFR protein, and comparing the activity with the wild-type activity of the protein.

In this method, the activity may be enzymatic activity or immunogenic activity. The agent may be incubated by simply mixing the agent with the protein under, for example, physiological conditions. The agent preferably is useful for treatment of a mycobacterial infection. The DHFR protein may be a protein produced by a species selected from the group of *M. avium, M. bovis, M. tuberculosis, M. leprae*, or another mycobacteria, such that the agent is specific for treatment of one of these same species.

Another embodiment of the invention is directed to methods for selecting an antimycobacterial agent specific against a mycobacterial infection comprising crystallizing a recombinant *Mycobacterium* DHFR protein and determining the molecular conformation of the protein, identifying a binding site within the molecular conformation, and selecting the agent with the molecular structure that fits the binding site. The binding site is preferably a substrate binding site.

Another embodiment of the invention is directed to methods for identifying the sequence of a mycobacterial DHFR gene comprising amplifying nucleic acid in a biological sample containing *Mycobacterium* by a polymerase chain reaction with two probes which span all or part of the *M. avium* DHFR gene, and determining the sequence of the amplified nucleic acid to identify the sequence of the mycobacterial DHFR gene. Preferably, the sequence identified encodes the DHFR protein of *M. avium, M. bovis, M. tuberculosis* or *M. leprae*. In a preferred embodiment, the sequence of the first probe contains a sequence from the 5' terminus of the *M. avium* gene and the sequence of the second probe contains a sequence from the 3' terminus of the *M. avium* gene.

Another embodiment is directed to methods for detecting *Mycobacterium* infection such as infection with *M. avium, M. bovis, M. tuberculosis* or *M. leprae*, by contacting a biological sample obtained

from a patient with an antibody specific to a recombinant *Mycobacterium* protein, and detecting bound antibody in the sample. The biological sample may be a sample of bodily fluid from a human. The antibody may be a monoclonal or polyclonal antibody.

Another embodiment of the invention is directed to methods for detecting a *Mycobacterium* infection by contacting a biological sample obtained from a patient with a recombinant mycobacterial protein or an active portion thereof, and detecting protein bound with antibody from the patient in the sample. The infection may be an infection of *M. avium, M. bovis, M. tuberculosis* or *M. leprae*. In this embodiment, the protein may be labeled with a detectable label, such as a radioisotope, stable isotope, fluorescent chemical moiety, enzyme, metal or combinations thereof.

Another embodiment of the invention is directed to methods for detecting a mycobacterial infection comprising amplifying nucleic acid in a biological sample containing *Mycobacterium* by a polymerase chain reaction comprising two probes, wherein the two probes span all or part of the *M. avium* DHFR gene, and detecting amplified nucleic acid that corresponds to an amplification of the nucleic acid between the two probes. In a preferred embodiment, the sequence of the first probe contains a sequence from a terminus of the *M. avium* gene and the sequence of the second probe contains a sequence from the opposite terminus of the *M. avium* gene. The two probes may be labeled with detectable labels, such as radioisotopes, stable isotopes, fluorescent chemical moieties, enzymes, metals and combinations thereof. In one embodiment, the step of detecting comprises determining the size of the nucleic acid amplified.

Another embodiment of the invention is directed to methods for screening for an agent that interacts with *M. avium* DHFR protein by immunizing an animal with a protein containing at least a portion of the

M. avium DHFR protein to generate anti-protein antibodies, immunizing another animal with the anti-protein antibodies to generate a collection of anti-idiotypic antibodies, selecting a anti-idiotypic antibody of the collection that binds to dihydrofolate, and identifying an agent that binds to said anti-idiotypic antibody. In this embodiment, the portion of the protein may be an enzymatically active or antigenically active portion, or a conserved region of the DHFR protein. The anti-idiotypic antibody may have an affinity for dihydrofolate that is comparable to the affinity of the catalytic site of *M. avium* DHFR for dihydrofolate. Preferably, the portion is a peptide corresponding to a region of mycobacterial DHFR protein that is not present in mammalian DHFR protein.

Another embodiment of the invention is directed to methods for detecting *M. avium* in a sample comprising immunizing an animal with a protein containing the DHFR sequence of the present invention to generate antibodies specific to the sequence, immunizing another animal with the antibodies to generate anti-idiotypic antibodies, and detecting *M. avium* DHFR protein in an immunoassay containing said anti-idiotypic antibodies. In this embodiment, the immunoassay may be a competitive immunoassay, an indirect immunofluorescence assay, an ELISA assay, an immunoprecipitation assay or other well-know or useful assay.

Another embodiment of the invention is directed to methods of detecting *M. avium* DHFR in a biological sample comprising the steps of combining a portion of the sample with an idiotypic antibody to *M. avium* DHFR protein, an anti-idiotypic monoclonal antibody to the idiotypic antibody such that the anti-idiotypic monoclonal antibody exhibits structural congruence with at least one epitope of the protein to form an assay mixture in which there is competition between the protein and the anti-idiotypic monoclonal antibody for binding to the anti-idiotypic antibody, and

detecting *M. avium* DHFR protein in the sample by determining the amount of bound labeled antibodies disposed within the anti-idiotypic antibody pairs. In one embodiment, determination of the amount of bound labeled antibody disposed within the anti-idiotypic antibody pairs follows a separation of the anti-idiotypic pairs from unbound antibody. Such separation may be achieved by precipitation. In one embodiment, at least one component of the mixture may be labeled with a detectable label, such as a fluorophore, radioactive compound, chemiluminescent compound, latex bead, enzyme, enzyme cofactor or enzyme inhibitor. The idiotypic antibody may be attached to a substrate or, alternatively, the anti-idiotypic antibody may be attached to a substrate.

Another embodiment of the invention is directed to antibodies specifically reactive against DHFR peptides of the invention. Antibodies may be polyclonal or monoclonal and expressed from a population of hybridoma cells. Such antibodies may be reactive against specific epitopes of the peptide such as the substrate binding site.

Another embodiment of the invention is directed to vaccines and diagnostic kits incorporating recombinant *Mycobacterium* DHFR peptides or antibodies specifically reactive against these recombinant peptides and to methods for using such vaccines and diagnostic kits.

Other embodiments and advantages of the invention are set forth, in part, in the description which follows, and, in part, will be obvious from this description and may be learned from the practice of the invention.

Description of the Drawings

Figure 1 Block diagram of DHFR's role in biosynthesis of tetrahydrofolate and cell metabolism.

Figure 2 Autoradiogram of Southern blot of genomic DNA from (A) *M. tuberculosis* H37Ra and (B) *M. avium* serovar 4.

Figure 3 Flow diagram depicting PCR procedure to search for the DHFR gene in *M. avium*.

Figure 4 Flow diagram depicting procedures used for screening genomic libraries for the DHFR gene *M. avium*.

Figure 5 Flow diagram depicting procedures used for complementation studies and cloning by deduced amino acid sequences.

Figure 6 Comparison of deduced amino acid translation of the p502 DHFR clone from *M. avium* (Consensus SEQ ID NO 2) with the deduced amino acid translation DHFR sequences from other prokaryotic DHFR sequences from the GenBank (Stepi SEQ ID NO 5; Ecoli SEQ ID NO 6; Citrob SEQ ID NO 7; Hinf SEQ ID NO 8; Bacsub SEQ ID NO 9; Llactis SEQ ID NO 10; Lactob SEQ ID NO 11; P502dhfr).

Figure 7 Sequence of *M. avium* DHFR gene (SEQ ID NO 1) and protein (SEQ ID NO 2).

Figure 8 SDS PAGE of *M. avium* DHFR expressed from BL 21 (DE3) plyS/p807.

Figure 9 Alignment of *M. avium* deduced N-terminal peptide sequence (SEQ ID NO 4) with *M. smegmatis* N-terminal peptide sequence (SEQ ID NO 4) previously reported by Sirawaraporn et al.

Description of the Invention

As embodied and broadly described herein, the present invention is directed to novel recombinant nucleic acids encoding DHFR for mycobacteria such as, for example, *M. avium,* to novel recombinant DHFR peptides produced by such sequences, and to vaccines, diagnostic kits, cells and therapies utilizing these peptides and nucleic acid sequences. The present invention is also directed to methods for using the sequences of the

present invention to develop drugs specific to *M. avium* and other mycobacteria, to identify and sequence corresponding sequences in species other than *M. avium,* as well as diagnostic and treatment methods incorporating the disclosed sequences and peptides.

In addition to uses relating to *M. avium, M. tuberculosis, M. bovis* and *M. leprae,* the nucleic acid sequences and peptides of the present invention may be useful in sequencing, treatments and diagnostic strategies relating to all species of mycobacteria. These include, but are not limited to, *M. fortuitum* (associated with mastitis in cows, and pulmonary infections, lymph node and cutaneous lesions in animals), *M. chelonei* (associated with contaminated wounds and injection abscesses), *M. marinum* (a human and cold-blooded animal pathogen), *M. scrofulaceum, M. xenopi,* and *M. lepraemurium* (leprosy-like pathogen of cats and rats).

Prior to the present invention, a DHFR gene had not been identified or cloned from any mycobacterial species. However, DHFR genes from other organisms are known. Figure 9 depicts alignment of the *M. avium* deduced N-terminal peptide sequence of the present invention with the *M. smegmatis* N-terminal peptide sequence reported by Sirawaraporn et al. (W. Sirawaraporn et al., Exper. Parisitol. 72:184-190, 1991). Letters connected with vertical line (|) indicate identical amino acid residues; question marks (?) could not be determined, and letters connected by a point (•) indicate conserved residues. Letters underlined and in *italics* indicate ambiguous residues. As depicted in Figure 9, comparison of the deduced N-terminal peptide sequence of the *M. avium* DHFR of the present invention with that of the *M. smegmatis* DHFR reported by Sirawaraporn, *et al.*, revealed that six of the fifteen residues were identical and three were conserved amino acids.

A *fol* A locus is listed in the GenBank as a putative dihydrofolate reductase gene for *M. tuberculosis* (GenBank Accession No. X59271). However, this gene sequence was not accurately identified. The listed sequence was compared with sixteen other known DHFR gene sequences in the GenBank. The reported sequence for *fol* A was not significantly homologous with any of the DHFR sequences and it lacked homology with important binding sites for the cofactor NADPH, and the inhibitors trimethoprim and methotrexate. Using primers homologous to the 5' and 3' ends of the putative gene sequence, this presumptive *fol* A gene was recloned by PCR and expressed in an *E. coli* expression system. The recombinant protein was purified by using a His/tag fusion protein and found to have a molecular weight of about 22 kDa by SDS PAGE. This is close to other DHFR proteins, examples being that from *M. smegmatis* (23 kDa) (W. Sirawaraporn et al., Exper. Parasitol. 72:184-90, 1991) and *Staphylococcus haemolyticus* (20 kDa) (G.E. Dale et al., Antimicrob. Agents Chemother. 39:1920-24, 1995). However, the expression vector with the presumptive *M. tuberculosis* GenBank *fol* A did not complement the DHFR deficient *E. coli* strain D3-157. As a result of these findings, the source for Accession No. X59271 was contacted. It was determined that indeed the sequence of Accession No. X59271 was not the *fol* A gene and that subsequent work had not been performed to identify the sequence further. Therefore, it appears that a complete DHFR gene has not previously been accurately identified in mycobacteria, cloned in any system or expressed.

Next, both the *M. avium* and *M. tuberculosis* genomic libraries were screened with a probe made from the DHFR gene of *Pneumocystis carinii.* No clones that would hybridize with this probe were identified from either library. Two optimized oligo probes were made from the binding sites of trimethoprim, methotrexate, and the cofactor NADPH on the *P. carinii*

DHFR gene using the aligned sequences of 16 species as a guide to the homologous regions (G.E. Dale et al., Antimicrob. Agents Chemother. 39:1920-24, 1995). These probes were used in a Southern blot reaction with genomic DNA from *M. avium* (serovar 4) and *M. tuberculosis* H37Ra, digested with several restriction enzymes. Figure 2 depicts an Autoradiogram of a Southern blot of genomic DNA from (A) *M. tuberculosis* H37Ra and (B) *M. avium* serovar 4. As indicated by Figure 2, both probes hybridized to the genomic DNA of both species, indicating homologous regions.

The functional recombinant *M. avium* DHFR gene has now been successfully cloned via genetic complementation and expressed as an active protein in *E. coli*. Figure 7 depicts the sequence of this cloned *M. avium* gene (SEQ ID NO 1) with flanking 5' and 3' regions obtained by automated DNA sequencing. The open reading frame is in bold (nucleotide 57 to 559). The polypeptide sequence (single letter code) of the open reading frame is indicated below the DNA sequence in capital letters. The methionine initiation codon used by *M. avium* is indicated by 'M' at position 57 (ATCC Accession No. AF006616).

As shown in Figure 7, the *M. avium* DHFR gene has an open reading frame of 543 bp and contains a guanosine plus cytosine content of 73%, consistent with other mycobacterial DNA. The translated polypeptide sequence of the *M. avium* gene compares favorably with that of other bacterial DHFRs, showing 58% identity to the consensus sequence of the conserved regions from eight other bacterial DHFRs. The recombinant *M. avium* DHFR was expressed actively in *Escherichia coli*. SDS PAGE analysis of recombinant DHFR protein revealed a 20 kDa species, agreeable with that predicted from the polypeptide sequence.

Figure 6 depicts a comparison of the deduced amino acid translation of the p502 DHFR clone from *M. avium* with the deduced amino acid translation DHFR sequences from other prokaryotic DHFR sequences from the GenBank, namely: *Staphylococcus aureus* (Saureus); *Staphylococcus epidermidis* (Stepi); *Escherichia coli* (Ecoli); *Citrobacter freundii* (Citrob); *Haemophilus influenzae* (Hinf); *Bacillus subtilis* (Bacsub); *Lactococcus lactis* (Llactis); *Lactobacillus casei* (Lactob). Analysis of predicted amino acid sequences were performed using Wisconsin Package Version 9.0, Genetics Computer Group (GCG) (Madison, WI). Alignments were carried out using TRANSLATE, PILEUP, and PRETTYBOX programs. Black boxes indicate amino acid identity, gray boxes indicate similar amino acids, and dots indicate gaps in the alignment. Following genetic complementation cloning, the open reading frame was found for which the peptide sequence aligns with that of other DHFR genes.

Although the DHFR gene had not previously been cloned from any mycobacterial species, the nucleotide sequence for the DHFR gene has been known for a number of bacterial species. Of the known DHFR genes, the most related species to *Mycobacterium* are *Staphylococcus aureus, Staphylococcus epidermidis, Bacillus subtilis and Lactobacillus casei.* These sequences were used to choose regions of homology between the genes, particularly in the regions involved in the binding of trimethoprim, methotrexate and the cofactor NADPH. These sequences were also used to make probes and PCR primers based on the known nucleotide codon preference of *M. tuberculosis.* Identification and cloning of the DHFR gene in *E. coli* by genetic complementation allows for expression of the gene and gene segments in *M. smegmatis* MC^2155, to produce functional enzymes and enzyme fragments. Both *M. avium* and *M. tuberculosis* genomic libraries, genomic DNA from both species *M. tuberculosis* H37Ra and *M. avium*

serovars 4 and 8, *Escherichia coli* D3-157 deficient in DHFR are available and can be used in complementation experiments, and *M. smegmatis* MC^2155 for expression of recombinant DNA.

A comparison of the specific activity of the soluble DHFR protein expressed by p807 with that previously reported for purified mycobacterial DHFR (M. Al-Rubeai et al., Biochem. J. 235:301-3, 1986, W. Sirawaraporn et al., Exper. Parisitol. 72:184-90, 1991) indicates that the p807-expressed enzyme represented about 2% of the total soluble protein. In contrast, DHFR, present in endogenous *M. avium* extracts represents about 0.01% of the total soluble protein.

In addition to genetic complementation, strategies for screening and cloning the DHFR gene from *M. avium* and other mycobacteria such as *M. tuberculosis*, include, for example, use of PCR for cloning from genomic DNA (Figure 3, Example 9), bacteriophage genomic library screening (Figure 4, Example 10), and gene cloning by deduced amino acid sequence (Figure 5, Example 12). The present invention represents the first report of the cloning, expression, purification and demonstration of biochemical activity of a recombinant mycobacterial DHFR protein.

Based upon experiences with the development of DHFR inhibitors and inhibitor-drug combinations, and experience with *Mycobacterium* species, the successful cloning of the DHFR gene from *M. avium* and other species now allows for the recombinant proteins to be used, for example, for kinetic analyses and binding studies with drug compounds. DHFR proteins are also useful for obtaining information on molecular structure, including x-ray crystallographic structure and molecular graphics, for drug design and modeling studies.

One embodiment of the invention is directed to the recombinant DHFR nucleic acid of *Mycobacterium* and, specifically, the *M. avium* DHFR gene sequence (SEQ ID NO 1), as depicted in Figure 7. This sequence comprises the *M. avium* DHFR gene with flanking 5' and 3' regions obtained by automated DNA sequencing. The nucleic acid sequence may comprise DNA, RNA or PNA, and may further contain one or more additional sequences to direct transcription or translation of the protein product. These additional sequences may be origins of replication, promoters, polymerase binding sites, enhancers, or transcription or translation termination sites. Such sequences may be homologous to mycobacterial cells or heterologous, selected for a level of expression in a heterologous host cell.

Another embodiment of the invention is directed to a vector, such as prokaryotic, eukaryotic or combination shuttle vectors, or an other suitable vector which contains the nucleic acid sequence of SEQ ID NO 1. Alternatively, the nucleic acid sequence may be integrated into the genome of a recombinant cell, or be contained within a recombinant cell episomally. The recombinant cell may be prokaryotic such as, for example, *E. coli, B. subtilis* or *P. aeruginosa*, or eukaryotic such as, for example, human, primate, rabbit or another mammalian cell. Integration may be performed by transfection, transformation, lipofection or any of the well-known methods of integrating a nucleic acid into a cellular genome. Integrated cells can be maintained in tissue culture or *in vivo* as desired.

Another embodiment of the invention is directed to a recombinant nucleic acid which contains a nucleic acid sequence encoding at least a portion of the DHFR protein of a *Mycobacterium*. The portion may be an enzymatically active portion which may be more useful in certain drug-development procedures. Alternatively, the portion may be an

antigenically active portion. The sequence may encode for the entire amino acid sequence of a mycobacterial DHFR protein. The DHFR protein may be derived from *M. avium, M. bovis, M. tuberculosis* or *M. leprae,* or another *Mycobacterium* species.

Another embodiment of the invention is directed to a recombinant DHFR peptide comprising an amino acid sequence that contains a sequence of mycobacterial DHFR such as, for example, SEQ ID NO 2, or at least a portion of the amino acid sequence of a *Mycobacterium* DHFR protein. The portion of the recombinant peptide may be an enzymatically active portion, or may be an antigenically active portion. The sequence may also contain the entire amino acid sequence of a DHFR protein of *Mycobacterium.* This DHFR protein may be derived from any *Mycobacterium* species, including *M. avium, M. bovis, M. tuberculosis,* or *M. leprae.* The DHFR protein may be a purified protein from *M. avium, M. bovis* or *M. tuberculosis.*

Recombinant DHFR peptides of the present invention are useful in the development of anti-infectives. For example, recombinant peptide may be used to screen potentially effective inhibitors of that drug target, determine enzyme kinetics, and to develop more effective and better drugs through molecular modeling and other drug-development techniques, which are well-known to those of ordinary skill in the art.

Recombinant DHFR peptides may also be used as vaccines against or to prevent mycobacterial infections, or diagnostically in a kit. For example, DHFR peptide may be mixed with an appropriate diluent and injected into an animal, person or other host to stimulate a desired immune response. Alternatively, antibodies specifically reactive against the recombinant peptide, such as monoclonal or polyclonal antibodies may be generated, collected and, as necessary or desired, purified by techniques all

of which are well-known to those of ordinary skill in the art. Monoclonal antibodies can be purified from cell cultures of hybridoma cells from human, primate or other mammalian cells. The antibody may be, for example, a recombinant such as a humanized antibody. Such antibodies may be used, for example, as a vaccine or in a diagnostic kit.

Another embodiment of the invention is directed to a method for assessing the ability of an agent to inhibit the activity of a DHFR protein. Recombinant DHFR of, for example, *M. avium* or *M. tuberculosis* can be isolated, characterized and used to screen for antimycobacterial agents. It can be used in studies to improve the binding potential of the lipophilic DHFR inhibitors and predict new second generation analogs with improved antimycobacterial activity. The sequence of the DHFR gene of *M. avium, M. tuberculosis* or related species, can be subcloned into appropriate vectors and the DHFR protein expressed to develop site directed lipophilic DHFR inhibitors by means of X-ray crystallographic and other molecular graphic techniques. Specifically, potential lipophilic inhibitors can be synthesized to screen against mycobacterial DHFR *in vitro*. In the absence of a 3-dimensional enzyme structure from X-ray diffraction, the *in vitro* mycobacterial DHFR data for these lipophilic antifolates can be used to carry out a comparative molecular field analysis (CoMFA) (K.H. Czaplinski, Eur. J. Med. Chem. 30:779-87, 1995). This analysis can be used to relate activity to specific molecular properties and shapes in order to predict newer, second generation analogs with improved activity as part of an iterative drug-design process.

In addition to using recombinant *M. avium* DHFR to develop new drugs, existing anti-DHFR analogs can be screened using the DHFR protein for inhibition activity. Further, the disclosed sequence of the present invention can be used to identify corresponding genes in other mycobacterial

species, and to identify other genes coding for enzymes important in folic acid synthesis. This information can then be used to develop drugs that inhibit enzymes in the same pathway as DHFR. Such drugs may be used in combination with anti-DHFR drugs to increase the therapeutic advantage of such a treatment strategy.

Another method comprises incubating a recombinant mycobacterial DHFR protein with the agent, determining the activity of the incubated DHFR protein, and comparing the activity determined with the wild-type activity of the protein. The activity inhibited may be enzymatic activity, immunogenicity or another activity. The step of incubating may comprise mixing the agent with the protein under physiological conditions such as physiological temperature or lower (room temperature; about 22°C), physiological pH (*e.g.* about 7.2), and physiological ionic conditions (*e.g.* 0.9% NaCl or PBS). In a preferred embodiment, the agent selected is useful for the treatment of any mycobacterial infection. The *Mycobacterium* DHFR protein may be of a specific species selected from the group consisting of *M. avium, M. bovis, M. tuberculosis, M. leprae* or another *Mycobacterium* species, such that the agent is specific for treatment caused by only that species of a selected group of species.

Other embodiments of the invention are directed to methods of identifying and screening agents suitable for inhibiting the activity of mycobacterial DHFR. One method of screening for an agent that inhibits the activity of recombinant DHFR protein of a *Mycobacterium* comprises the steps of determining the activity of the DHFR protein upon incubation with each of a plurality of agents, and selecting the agent that inhibits DHFR protein activity. The agent may inhibit all mycobacterial species, or may selectively inhibit *M. avium, M. bovis, M. tuberculosis, M. leprae*, or others. The DHFR protein may contain the sequence of SEQ ID NO 2, or an active

portion thereof. This active portion may be an enzymatically active portion or an antigenically active portion. Through the use of high through-put screening, a large number of agents can be screened, such as greater than 10^2 different agents, preferably greater than 10^4, more preferably greater than 10^6, or greater than about 10^8, or more. These agents may comprise a collection of related chemical compounds such as, for example, chemical modifications of folate, methotrexate, trimethoprin, quinazoline or combinations of these agents, or combinantions of these agents with other anti-mycobacterial agents such as, for example, rifampicin, streptomycin and isoniazid.

The step of incubation may comprise mixing the DHFR protein with the agent under conditions well-known to those of ordinary skill in the art that allow for molecular interaction to occur. The desired molecular interaction may be binding, or inhibition of activity such as enzymatic activity. Alternatively, the activity inhibited may be immunoginecity. Additionally, the method may further comprise the step of determining the molecular conformation of said agent again using techniques well-known to those of ordinary skill in the art such as, for example, a comparative molecular field analysis (CoMFA).

This method may further comprise the steps of selecting a plurality of agents that inhibit the activity of the *Mycobacterium* DHFR protein, determining the molecular conformation of each agent selected, and identifying a common inhibitory molecular conformation. For example, in many cases, agents tested may have a known or easily determinable molecular conformation. Identification of a common structure or sub-structure within that conformation usinf, for example, well-known crystallographic techniques or structure-activity relationships learned from

folate gene products, allows one of ordinary skill in the art to design and construct more effective agents.

Use of *Mycobacterium* DHFR protein to identify useful antimycobacterial drugs has several advantages. Looking at inhibition of *enzyme* rather than inhibition of growth of viable microorganisms is faster, less hazardous, less expensive and poses fewer complications. In addition, using the DHFR protein itself to identify inhibitors of *Mycobacterium* DHFR allows for more direct evaluation of efficacy and facilitates additional drug discovery.

To discover antimycobacterial drugs, in the absence of an actual X-ray structure, computer modeling may be used to obtain a 3-dimensional picture of the recombinant *M. avium* DHFR active site and correlate structural alterations with differences in pharmacologic activity. The success of this approach can depend on how robustly a CoMFA performs. The quality of the analysis and subsequent predictions are directly correlated to the quality and number of components that are analyzed. A CoMFA has been published using 14 Trimethoprim analogs against DHFR from various species including *M. lufu* (K.H. Czaplinski, Eur. J. Med. Chem. 30:779-87, 1995). However, very poor correlation coefficients were obtained for the *M. lufu* analysis. Derivation of structure activity relationships for small data sets non-homogeneous in structure and/or conformation can lead to erroneous results (K.H. Czaplinski, Eur. J. Med. Chem. 30:779-87, 1995). In other CoMFA studies involving enzyme receptors, the use of at least 50 inhibitors has resulted in much better predictive capability (M.A. El-Bermawy et al., Med. Chem. Res. 2:290-97, 1992; C.L. Waller et al., J. Med. Chem. 36:2390-403, 1993; C.L. Waller et al., Chem. Res. Toxic. 8:847-58, 1995). Thus, by using at least 50 structurally similar derivatives, a high quality first model from which to

make predictions for second generation compounds to be synthesized can be produced. Further refinement of the model with data from new inhibitors could validate and increase its predictive capability. The recombinant protein could serve as a valuable reagent for these studies to identify new drugs.

In one embodiment, a CoMFA may also be used in connection with a method for selecting an antimycobacterial agent specific against a *Mycobacterium* infection. This method comprises the steps of crystallizing a recombinant mycobacterial DHFR protein and determining the molecular conformation of the protein, identifying a binding site within the molecular conformation, and selecting the agent with a molecular structure that fits within the binding site. In this method, the binding site may be a substrate binding site. Positron emission topography (PET) can also be used to analyze molecular structure. PET techniques are well know to those of ordinary skill in the art.

Another embodiment of the invention is directed to methods for identifying sequences of mycobacterial DHFR genes. A preferred method comprises amplifying nucleic acid in a biological sample containing *Mycobacterium* by a polymerase chain reaction with two probes which span the *M. avium* DHFR gene. The sequence of the first probe preferably contains a sequence from the 5' terminus of the *M. avium* gene and the second probe preferably contains a sequence from a 3' terminus of the *M. avium* gene. The method comprises a second step of determining the sequence of the amplified nucleic acid to identify the sequence of the mycobacterial DHFR gene. Sequences identified may encode the DHFR protein of *M. avium, M. bovis, M. tuberculosis, M. leprae* or other mycobacterial species.

In another embodiment, the mycobacterial DHFR gene can be sequenced by contacting genomic DNA of a target mycobacterial species with a probe containing all or part of SEQ ID NO 1 and identifying hybridized regions of homology between the genomic DNA and the sequence. The *M. avium* sequence utilized may comprise one or more binding sequences, such as the portion of the DHFR protein that binds with trimethoprin, methotrexate, or NADPH. The target DNA may then be isolated and sequenced.

Another embodiment of the invention is directed to a method for detecting a *Mycobacterium* infection. In this method, a biological sample from a patient is contacted with an antibody specific to a recombinant *Mycobacterium* protein. Bound antibody is then detected in the sample. The mycobacterial infection may be due to *M. avium, M. bovis, M. tuberculosis, M. leprae*, or other mycobacterial infection. The biological sample may be a sample of bodily fluid or tissue. The patient may be an animal or a human suspected of harboring a mycobacterial infection. The antibody may be a monoclonal antibody, a polyclonal antibody or another type of antibody. Detection can be performed using a wide variety of techniques, all well-known and widely used in the art.

Another embodiment is directed to a method of detecting *Mycobacterium* infection by contacting a biological sample obtained from a patient with a recombinant mycobacterial protein or active portion thereof, and detecting protein bound with antibody from the patient in the sample. The infection may be due to *M. avium, M. bovis, M. tuberculosis, M. leprae*, or other *Mycobacterium.* The protein may be labeled with a detectable label such as a radioisotope, stable isotope, fluorescent chemical moiety, enzyme, metal or combination thereof.

Another embodiment of the invention is directed to a method for detecting mycobacterial infection by a first step of amplifying nucleic acid in a biological sample containing *Mycobacterium* by polymerase chain reaction. This can be accomplished through the use of two probes, wherein the sequences of the two probes span the *M. avium* DHFR gene. The sequence of the first probe preferably contains a sequence from a terminus of DHFR gene and the sequence of the second probe preferably contains a sequence from the opposite terminus of the *M. avium* DHFR gene. The method includes the further step of detecting amplified nucleic acid that corresponds to an amplification of the nucleic acid between the two probes. The two probes may be labeled with detectable labels, such as radioisotopes, stable isotopes, fluorescent chemical moieties, enzymes, metals and combinations thereof. In the step of detecting, the size of the nucleic acid amplified may be determined.

Another embodiment of the invention is directed to an immunoassay for the detection of *M. avium* DHFR. The assay detects *M. avium* DHFR by its inhibition of the reaction between the combining site of an idiotypic antibody to DHFR and an anti-idiotypic antibody. The assay provides a method for measuring the existence and concentrations of single epitopes without requiring purified DHFR or mycobacteria. One advantage of this method is that it can detect unpurified *M. avium* DHFR even in a mixture of antigens. Another advantage of the method is that it provides a consistent, specific reagent for DHFR measurement which is not dependent on the catalytic function of DHFR. A third advantage of the system is that it is more resistant to enzyme inhibitors, detergents and ions than assays based on DHFR enzyme activity.

Another embodiment of the invention is directed to a method to screen for an agent or group of agent which interact with *M. avium* DHFR

protein. In this method, an animal is immunized with a protein comprising a sequence that contains, for example, at least a portion of the protein identified in SEQ ID NO 2 to generate anti-protein antibodies, immunizing another animal with the anti-protein antibodies to generate a collection of anti-idiotypic antibodies, selecting an anti-idiotypic antibody of the collection that binds to dihydrofolate, and identifying an agent that binds to the anti-idiotypic antibody. The portion of SEQ ID NO 2 may be an enzymatically active portion, an antigenically active portion, or a conserved region of the DHFR protein. The anti-idiotypic antibody may have an affinity for dihydrofolate that is comparable to the affinity of the catalytic site of *M. avium* DHFR for dihydrofolate. In a preferred embodiment, the portion is a peptide that corresponds to a region of mycobacterial DHFR protein that is not present in mammalian DHFR protein. Such an embodiment allows for treatment of a mammalian patient with an agent which would have an adverse effect on mycobacterial DHFR while sparing the host's DHFR protein.

Another embodiment of the invention is directed to a method for detecting *M. avium* in a sample by immunizing an animal with a protein containing, for example, SEQ ID NO 2 or a portion thereof to generate antibodies specific to the sequence, immunizing another animal with the antibodies to generate anti-idiotypic antibodies, and detecting a *M. avium* DHFR protein in an immunoassay containing the anti-idiotypic antibodies. In this embodiment, the immunoassay may be a competitive immunoassay, an indirect immunofluorescence assay, an ELISA assay, an immunoprecipitation assay, or other assay known in the art.

Another embodiment of the invention is directed to a method for detecting *M. avium* DHFR in a biological sample comprising the steps of combining a portion of the sample with an idiotypic antibody to *M. avium*

DHFR protein, an anti-idiotypic monoclonal antibody to the idiotypic antibody wherein the anti-idiotypic monoclonal antibody exhibits structural congruence with at least one epitope of the protein to form an assay mixture in which there is competition between the protein and the anti-idiotypic monoclonal antibody for binding to the anti-idiotypic antibody, and then detecting *M. avium* DHFR protein in the sample by determining the amount of bound labeled antibodies disposed within the anti-idiotypic antibody pairs. In a preferred embodiment, determining the amount of bound labeled antibodies disposed within the anti-idiotypic antibody pairs follows a separation of the anti-idiotypic antibody pairs from unbound antibody. Separation may be by precipitation and at least one component of the mixture may be labeled with a detectable label such as, for example, a fluorophore, radioactive compound, chemiluminescent compound, latex beads, enzyme, enzyme cofactor or enzyme inhibitor. The idiotypic antibody may be attached to a substrate. Alternatively, the anti-idiotypic antibody may be attached to a substrate.

The following examples are offered to illustrate embodiments of the invention, and should not be viewed as limiting the scope of the invention.

Examples

Example 1 Cloning of DHFR gene from mycobacterial species by genetic complementation.

At least four independent strategies can be used to screen or clone the DHFR gene from *M. avium* and *M. tuberculosis* H37Ra and other mycobacteria. The four approaches include PCR (Figure 3), bacteriophage genomic library screening (Figure 4), genetic complementation (Figure 5), and gene cloning by deduced amino acid sequence (Figure 5) As discussed,

the *M. avium fol* A gene (DHFR) was first successfully cloned via genetic complementation.

Bacterial Strains. For genetic complementation, a DHFR-deficient *Escherichia coli* strain D3-157 (S. Singer et al., J. Bacteriol. 164:470-72, 1985) was purchased from the American Type Culture Collection (Rockville, Maryland). This strain contains a DHFR mutation that maps at or near the *fol* A gene and was given the designation *fol*- 200 (S. Singer et al., J. Bacteriol. 164:470-72, 1985). Other *E. coli* strains which may be used are JM109 (Promega, Madison, WI), for cloning of plasmid constructs, and BL21(DE3)plysS (Novagen, Madison, WI), a lysogen of bacteriophage lambda DE3 containing the plasmid plysS, for controlling expression of proteins from the pET-15b vector. Useful bacterial strains and plasmids are listed in Table 1.

Table 1

List of *Escherichia coli* strains and plasmids.

Strain or plasmid	Relevant genotype or properties	Origin
D3-157	F-, *guaB22, xyl-7, rpsL125, fol - 200, Strep^r*	ATCC
JM109	*endA*1, *recA*1, *gyrA96, thi, hsdR17,* (r_k^-, m_k^+), *relA 1, supE^{44}*, λ-, Δ (*lac-pro AB*), [F-, *traD36, proAB, lac* I^{q}ZΔM15]	Promega
BL21(DE3)plysS	F-, *ompT, hsd* S_B (r_B-m_B-), *gal, dcm,* (DE3)plysS	Novagen
pBS+	*Amp^r, lac Z* promoter	Stratagene
pGEM®-7Zf+	*Amp^r*, T7 and SP6 promoters	Promega
pGEM®-T Easy	*Amp^r*, T7 and SP6 promoters	Promega
pET-15b	*Amp^r*, T7 *lac* promoter, His•Tag fusion protein	Novagen
p502	pBS$^+$ with *M. avium* genomic DNA insert	This study
p807	pET-15b with *M. avium folA* gene inserted in frame with *T7 lac* promoter	This study

Screening of genomic DNA library by complementation. An *M. avium* genomic DNA library in the λ ZAPII vector (Stratagene, La Jolla, CA) was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH (Catalog Number 1786, contributed by BioTechnology General). The vector was excised with helper phage and recircularized according to the manufacturer's protocol to make subclones in the pBluescript® (pBS+) vector which represented the genomic DNA fragments from 10^6 pfu of the phage library. The plasmid genomic library (P502) was then purified twice with a CsCl gradient. A 100 ng aliquot of the plasmid genomic library was transformed into the DHFR deficient *E. coli* strain D3-157 and plated onto M9 minimal salts medium lacking thymidine (J. Sambrook et al., *Molecular Cloning: a Laboratory Manual*, second ed., New York: Cold Spring Harbor Laboratory Press, 1989) supplemented with tryptophan, tyrosine, and histidine (50 μg/ml), guanine (20 μg/ml), thiamine (10μg/ml), streptomycin and ampicillin (100 μg/ml), and IPTG (1 mM).

The *M. avium* genomic library successfully complemented the *E. coli* DHFR deficient strain, D3-157, using described methods (G. Vasanthakumar et al., Gene 147:153-54, 1994). One concern with this approach was that the *Mycobacterium* promotor might not be recognized by the *E. coli* host cell; however, it was expected that some of the genomic DNA constructs would contain the DHFR gene in close enough proximity to the *E. coli* plasmid promotor to utilize it in making DHFR to complement the deficient strain. The deficient *E. coli* strain D3-157 was transformed with the *M. avium* genomic library, as described above, and plated onto minimal media supplemented with D3-157 strain nutritional requirements except thymidine. After the plates were incubated at 37°C for five days, four

colonies were seen, indicating successful complementation and functional expression of the *M. avium* gene insert.

The deficient strain should be complemented by any *M. avium* plasmid from the genomic library that expresses the DHFR gene, such as the above described plasmid clone P502.

Example 2 Southern Blot Analysis.

Plasmids from the *M. avium* library clones of Example 1 were purified by CsCl gradient centrifugation. Positive clones were verified by Southern blotting and subjected to DNA sequencing. Restriction enzyme digests and Southern blots were done according to standard protocol (J. Sambrook et al., *Molecular Cloning: a Laboratory Manual*, second ed., New York: Cold Spring Harbor Laboratory Press, 1989). The oligonucleotide probe (5'-AATGCCGTGATGATGGGCCGGAAGACCTGGGAAT CGAT-3' SEQ ID NO 12), which was prepared by Genosys (The Woodlands, TX), was 3'-end labeled with digoxigen-11-ddUTP and hybridization was performed with the Genius™ system according to the manufacturers' specifications (Boehringer Mannheim, Indianapolis, IN).

Restriction enzyme digestion of p502, a plasmid isolated from one of the complementation colonies, revealed a 6.4 kb insert of genomic DNA. To identify the DHFR gene, the 38 base oligonucleotide probe described above was made based on the aligned sequences of the methotrexate binding sites of known DHFR genes from other species and the codon preferences of *Mycobacterium* (S.G.E. Anderson et al., Microbiol. 142:915-25, 1996). The complementation plasmid insert was cut with *Bam*, HI, *Xho I*, *Apa I*, and *Hin* dIII, and then hybridized in a Southern blot. The probe hybridized to the 1.6 kb *Apa* I fragment which was subsequently cloned into the *Apa* I site of pGEM®-7Zf+ (Promega, Madison, WI).

Example 3 DNA Sequencing and Analysis.

The DNA sequencing was determined by dideoxy chain termination method using the ABI PRISM™ Dye Terminator Cycle Sequencing kit (Perkin-Elmer; Foster City, CA) and the ABI PRISM™ 377 automated DNA Sequencer, either in house or through the DNA Sequencing Facility at Iowa State University, Ames, Iowa. Primers used for sequencing were made by Gibco BRL (Gaithersburg, MD). The DNA sequence encoding the *M. avium fol* A gene was submitted to GenBank and assigned to the Accession No. AF006616.

Analysis of nucleotide and predicted peptide sequences was performed using the Wisconsin Package Version 9.0, Genetics Computer Group (GCG) Madison, WI) (J. Devereux et al., Nucleic Acids Res. 12:387-95, 1984). Alignments were carried out using TRANSLATE, PILEUP and PRETTYBOX programs.

The 1.6 kb subclone of the p502 complementation plasmid was sequenced and an open reading frame (ORF) was found (Figure 7). A deduced polypeptide sequence comparison with other DHFR genes indicated that a portion of the 5' end of the gene was not in the 1.6 kb insert. Therefore, the sequence of the entire gene was determined from both strands of p502. The nucleotide sequence of the *M. avium fol* A gene is given in Figure 7, along with the translated amino acid sequence. The open reading frame is 543 bp in length and has a guanosine plus cytosine (G/C) content of 73%, in agreement with that found for most mycobacterial DNA (S.G.E. Anderson et al., Microbiol. 142:915-25, 1996; J.E. Clark-Curtiss, *Molecular Biology of the Mycobacteria*, 77-96, 1990).

The translated polypeptide sequence from the *M. avium fol* A gene was aligned with those of several other known bacterial DHFR genes, including *Staphylococcus aureus, Staphylococcus epidermidis, E. coli,*

Citrobacter freundii, Haemophilus influenzae, Bacillus subtilis, Lactoccus lactis, and *Lactobacillus casei* (Figure 6). Comparison of the polypeptide sequence of the *M. avium fol* A gene with that from eight other bacterial DHFR genes revealed a 58% identity to the consensus sequence derived from the conserved regions of the bacterial genes.

In a Southern hybridization, the probe BSDHFR was found to hybridize strongly with p502 DNA. A 1.6 kb restriction enzyme digested fragment from the 6.4 kb *Mycobacterium avium* genomic DNA insert from p502 that also hybridized with the BSDHFR probe was subcloned, p502-7, and partially sequenced. Within this clone, p502-7, a region was found that contained an open reading frame that was homologous to DHFR genes from other species and contained the binding sites for the cofactor NADPH, and the inhibitors trimethoprim and methotrexate.

Example 4 Construction of expression plasmid.

PCR may performed with oligonucleotide primers (GIBCO BRL; Gaithersburg, MD), using Taq DNA polymerase (Fisher Scientific; Pittsburgh, PA) according to manufacturer specifications except that dimethyl sulfoxide may be added to the reaction at a final concentration of 5%. Primers, are 8DHFR, 5'-CATATGACCC GTGCCGAGGTG-3' (SEQ ID NO 13) and 7DHFR, 5'-GGATCCTCAGCTCGGGCGT GAGG-3' (SEQ ID NO 14), include *Nde* I and *Bam* HI restriction enzymes sites to the 5' and 3' ends of the *fol* A gene respectively. The template DNA is denatured at 95°C for 3 minutes and PCR is performed for 35 cycles of 95°C for 1 minute, 60°C for 1 minute, and 72°C for 1 minute with a 7 minute extension cycle at 72°C. PCR product is isolated in a 0.8% agarose gel and the DNA eluted from the gel slice by centrifugation with an Ultrafree™-MC

spin column (Millipore; Bedford, MA). PCR products are cloned into the pGEM™-T Easy vector system (Promega; Madison, WI).

To make the DHFR plasmid construct, p807, the pET-15b vector (Novagen, Madison, WI), is digested with *Nde* I and *Bam* HI, then treated with calf intestinal alkaline phosphatase (Promega; Madison, WI). The PCR subclone of the *M. avium fol* A gene in pGEM®-T Easy vector is digested with *Nde* I and *Bam* HI, then the restriction enzyme digested-fragments of the *M. avium* DHFR gene are isolated in a 0.8% agarose gel and the DNA eluted from the gel slice. The DNA is directionally cloned into the pET-15b expression system using the Ligation Express™ kit (Clontech; Palo Alto, CA) with the start codon in close proximity to the T7 promotor.

Example 5 Expression of recombinant *M. avium* DHFR.

The host strain BL21(DE3) plysS, containing the DHFR recombinant gene in p807, is grown in LB broth (J. Sambrook et al., *Molecular Cloning: a Laboratory Manual* (second ed.), 1989) with 100 µg/ml ampicillin at 28°C to an A_{600} of 1.0 and expression induced with 0.1 mM IPTG overnight. The complementation plasmid p502, in the deficient strain D3-157 is grown under the same conditions in LB broth with 50 µg/ml thymidine, 100 µg/ml ampicillin and 100 µg/ml streptomycin. Thirty ml samples of cultures are washed twice with cold standard DHFR assay buffer and the pelleted cells are stored overnight at -20°C. Cells are resuspended in one ml of cold assay buffer and lysed by sonication. The insoluble proteins are pelleted at 16,000 *g* in a microcentrifuge at 4°C and the supernatant assayed for DHFR activity, as described below.

Example 6 Purification and thrombin cleavage of recombinant DHFR.

Recombinant *M. avium* DHFR may be expressed and purified under denaturing conditions using His•Bind resin (Novagen; Madison, WI), taking advantage of the His•Tag fusion protein. The *M. avium* DHFR gene

in p807 was expressed in BL21(DE3)plysS as described above, except that it was grown at 37°C and induced with 1 mM IPTG for three hours. Cells were washed once with ice cold PBS and stored at -20°C. Recombinant fusion protein was purified from the inclusion bodies following the protocol in the pET System Manual, 4th Ed. (1994), and eluted from the His•Bind resin column with 300 mM imidazole. His•Tag fusion protein was then cleaved from 140μg of recombinant protein with 0.82 Units of Thrombin (Novagen; Madison, WI) in 100 μl of 1x Thrombin cleavage buffer (20 mM Tris, pH 8.4; 0.15 M NaCl; 2.5 mM $CaCl_2$) at RT overnight. SDS PAGE analysis was performed with a 4-20% polyacrylamide gel (ICN Pharmaceuticals, Inc.; Costa Mesa, CA) with Perfect Protein™ molecular weight markers (Novagen; Madison, WI) and stained with Colloidal Coomassie Solution (ICN Pharmaceuticals, Inc.; Costa Mesa, CA) according to the manufacturer's protocol.

Recombinant protein was expressed as a 22 kDa band that was not present before induction with IPTG. Figure 8 depicts SDS PAGE analysis of the expression of *M. avium* DHFR in BL21(DE3)plysS/p807. Proteins were analyzed by SDS PAGE, noninduced culture (40 μg) (Lane 1), culture induced with 1 mM IPTG (40 μg) (Lane 2), *M. avium* DHFR His•Tag fusion protein purified on His•Bind resin (7 μg) (Lane 3), and Thrombin cleaved (indicated as "cp" for cleaved protein) and uncleaved *M. avium* DHFR (10 μg) (indicated as "fp" for fusion protein) (Lane 4). Molecular weight markers are indicated in kilodaltons (kDa) (Lane 5). As indicated in Figure 8, when the recombinant protein was purified on a His•Bind resin column and cleaved with thrombin to remove the His•Tag fusion protein, a 20 kDa band of cleaved protein could be seen which is in close agreement with the predicted size of the protein (19.9 kDa), based upon the polypeptide sequence (Figure 8).

Example 7 Activity of recombinant DHFR.

Both recombinant *M. avium* DHFR plasmid constructs (p807 and p502) were expressed under slow growth conditions at 28°C to maximize the amount of DHFR expressed as soluble protein. Enzyme assays were performed on crude cell lysates from the plasmid constructs expressed in both the deficient strain D3-157 and the expression strain BL21(DE3)plysS. As indicated in Table 2, DHFR activity was present in the deficient strain *E. coli* D3-157, containing the p502 complementation plasmid, but as expected, not in the D3-157 strain containing only pBS+ without the *M. avium* DHFR gene insert. Also, in the DHFR-proficient host strain BL21(DE3)plysS, DHFR specific activity was 1000-fold greater in the presence of the recombinant DHFR plasmid p807 than with the plasmid pET-15b without the recombinant DHFR gene.

Table 2

Expression of the DHFR Recombinant Gene and Enzyme Activity in *Escherichia coli*†

Plasmid	***E. coli* Host Strain**	**Protein (mg/ml)**	**DHFR 10^3 Units/ml**	**Specific Activity 10^3 Units/mg**
p502	D3-157	34.1 ± 3.1	290 ± 5.8	8.5 ± 0.95
pBS+	D3-157	37.8 ± 2.6	< 1.6	< 0.042
p807	BL21(DE3)plysS	21.7	25,000	1,200
pET-15b	BL21(DE3)plysS	13.5	16	1.2

† Data with p502 and pBS^+ are presented as the mean and standard deviation of three samples. Values listed for pET-15b are from one sample and those for p807 are the mean of two samples.

Example 8 Dihydrofolate reductase assay.

Dihydrofolic acid (FAH_2), NaDPH (tetrasodium salt), 2-mercaptoethanol and EDTA were obtained from Sigma (St. Louis, MO).

FAH_2 (20 mM) was suspended in 5 mM HCl containing 50 mM 2-mercaptoethanol and stored at -20°C until used for assay. On the day of assay, the FAH_2 suspension was dissolved in 50 mM potassium phosphate buffer (pH 7) and kept on ice.

DHFR activity was measured at 30°C in a Spectronic Genesis 5 Spectrophotometer as the decrease in the A_{340}. The reaction mixture was modified from that described by Al-Rubeai and Dale (M. Al-Rubeai, Biochem. 235:301-303, 1986) for the assay of DHFR from *Mycobacterium phlei* and consisted of 10 mM 2-mercaptoethanol, 0.1 mM NADPH, 0.1 mM FAH_2 and 0.01-0.05 ml of enzyme in a standard buffer of 50 mM potassium phosphate-1 mM EDTA, pH 7. The total assay volume was one ml. The reaction was initiated by the addition of FAH_2 after preincubation of the other components for 3 min. Dihydrofolate reductase activity was corrected for NADPH oxidase activity which was measured as a decrease in A_{340} in the absence of FAH_2 during the preincubation period. One unit of enzyme is defined as the amount which reduces 1 μmole of FAH_2 per min using a molar extinction coefficient at A_{340} of 12,300 $M^{-1} \cdot cm^{-1}$ (B.L. Hillcoat et al., Anal. Biochem. 21:178-89, 1967).

Example 9 Use of PCR for cloning from genomic DNA.

Figure 3 is a flow diagram outlining a general procedure to search for the DHFR gene in *M. avium* using PCR to clone the DHFR gene from genomic DNA. This may be accomplished by the following steps:

1. Sequences of DHFR genes from related species in question may be obtained from the GenBank and aligned by means of computer programs to yield areas of consensus. For instance, analysis of nucleotide and predicted amino acid sequences may be performed using Genetics Computer Group Sequence Analysis Software for the VAX (GCG; Madison, WI). Predicted peptide sequences may be translated with the PEPDATA

program, nucleotide and peptide sequence alignments may be carried out using GAP, PILEUP and PRETTYBOX programs in the GCG package. The known DHFR genes, *S. aureus*, *S. epidermidis*, *B. subtilis*, and *L. casei* are the most closely related to *Mycobacterium* areas of homology, particularly those known from the literature to be involved in binding of trimethoprim, methotrexate, and the cofactor NADPH. These may be used to make primers for PCR. The choice for an ambiguous nucleotide that may occur within the consensus region may be made based on the known codon preference of *M. tuberculosis*. Primers may be 20-30 bases in length and may be synthesized commercially (Genosys).

2. Primers may be used in a PCR reaction to amplify respective areas from genomic DNA.

3. PCR products may be separated on a gel and examined for bands corresponding to segments containing the number of bases expected in the gene product.

4. A commercial Cloning Kit may be used to clone the PCR product into a plasmid.

5. Products may then be sequenced.

6. The sequences of the products may be compared with gene sequences in the GenBank to determine degree of homology.

Example 10 Bacteriophage genomic library screening.

Figure 4 is a flow diagram of procedures which may be used for screening genomic libraries for the DHFR gene in species such as *M. avium*. As indicated in Figure 4, the following steps may be used for Genomic library screening:

1. Genomic Library (phage stock) may be used to infect *E. coli* (*e.g.* XL-1 Blue mrf' for λZAPII; Stratagene).

2. Lysis of *E. coli* takes place.

3. Using nylon membrane, DNA blots are made of plaques.

4. A probe will be made using the aligned sequences. Regions of homology will be chosen for use as a probe and a 50 base pair oligo will be synthesized commercially. The probe is nonradioactively labeled with digoxigenin.

5. Membranes are processed, leaving single-stranded DNA fixed. Membranes are incubated with digoxigenin-labeled probe to identify the original sites of plaques carrying the transformed gene. Positive sites are detected by fluorescent tag, using the Genius™ system (Boehringer Mannheim).

6. Insert is sequenced and compared with gene sequences in GenBank for homology.

7. Alternatively, a sublibrary may be screened with a DNA probe: Genomic DNA is digested with a restriction enzyme, run on an agarose gel, transferred to a nylon membrane and hybridized with a probe made to a homologous region in a Southern blot hybridization. The DNA in the region of the gel that hybridizes to the probe is eluted from the gel and used to make a plasmid library that will potentially contain the gene. This sublibrary can be plated in *E. coli* and screened with the same Southern blot conditions as with the genomic blot to identify the clones containing the homologous sequence. These clones may then be sequenced and compared to the sequences from the GenBank.

Example 11 Genetic complementation cloning.

Figure 5 is a flow diagram depicting a procedure which may be used for genetic complementation cloning, and gene cloning by deduced amino acid sequences. Genetic complementation cloning can be performed by the following steps:

1. Excision of lambda ZAP genomic library from phage to plasmid form.

2. Complementation of deficient strain of *E. coli*.

3. Growth of gene complemented mutant on minimal medium containing appropriate substrate.

4. Pick colonies, culture, and isolate plasmids. Perform restriction enzyme digestion and Southern Blot. Probe blot with oligonucleotide with sequence analogous to binding site.

5. Subclone DNA fragments that hybridize with probe and sequence.

Example 12 Genetic cloning by deduced amino acid sequence.

Gene cloning by deduced amino acid sequence as depicted in Figure 5 may be accomplished by the following steps:

1. The native DHFR protein from *M. avium* is purified, isolated on an SDS PAGE gel and transferred to polyvinylidene difluoride membrane. The protein is excised from the membrane and subjected to N-terminal sequencing.

2. The nucleotide sequence of the peptide can then be deduced and the sense strand used as one of the primers for PCR.

3. Internal peptides may be obtained by cleavage of the native DHFR protein into smaller peptides that may be purified and sequenced as above. The antisense nucleotide sequence from these peptides may be used as the opposing primer in a PCR reaction with genomic *M. avium* DNA.

4. The PCR product from this reaction may be cloned into *E. coli* and sequenced. This partial sequence of the DHFR gene may then be used to probe the Genomic library to obtain a clone containing the entire DHFR gene sequence.

Example 13 Expression and purification of recombinant DHFR.

The *Mycobacterium* DHFR gene may be cloned into the *pTrcHis* System (Invitrogen; San Diego, CA) for expression in *E. coli*. Restriction enzyme sites can be added to the 5' and 3' ends of the cDNA respectively by polymerase chain reaction (PCR) using primers containing these sites. PCR can be performed with *Taq* DNA polymerase (Perkin-Elmer Cetus; Norwalk, CO) under conditions recommended by the manufacturer. The PCR product can be cloned into the *pTrcHis* vector in close proximity to the *E. coli* promotor and expressed in TOPIO™ cells (Invitrogen; San Diego, CA) according to the manufacturer's protocol. *E. coli* containing recombinant plasmid can be grown in the presence of 0.1 mg ampicillin/ml and expression can be induced by the addition of 1mM IPTG. The cell pellet can be disrupted in a French Press (3x at 15,000 lb/inch^2) in ice cold lysis buffer and the recombinant protein purified by using a His/tag fusion protein with the Invitrogen affinity purification system. This procedure was used successfully to express the putative (and incorrect) DHFR gene listed in GenBank as Accession No. X59271.

Example 14 Use of *M. avium* DHFR sequence to identify other sequences.

The disclosed *M. avium* DHFR sequence may be used to identify corresponding DHFR genes in mycobacteria other than *M. avium* as well as in other species. The *M. avium* DHFR nucleic acid sequence, or portions thereof, may be used to make probes and PCR primers. These probes and primers may then be used to identify regions of homology on other genes. Regions particularly useful to identify corresponding DHFR genes include sequences encoding the portions of the DHFR protein involved in the binding of trimethoprim, methotrexate and NADPH.

For example, target DHFR sequences may be screened with a probe incorporating all or part of the mycobacterial DHFR nucleic acid

sequence. Oligo probes may be made from the binding sites of trimethoprim, methotrexate and NADPH on the DHFR gene, and used in a Southern Blot reaction with genomic DNA from the target sequence. The probe hybridizes to genomic DNA of the target having homologous regions. To facilitate this process, target sequences may be hybridized with ^{32}P-radiolabeled oligo probes homologous to binding site regions on the *M. avium* DHFR gene. Specific procedures involving the use of probes to identify homologous regions on target sequences are outlined in Figure 4. The nucleic acid sequence of the invention can be substituted and used in probes in these same procedures, to identify corresponding DHFR genes in species other than *M. avium*.

Example 15 Cell-free enzyme system and purification of DHFR from mycobacteria.

A crude cell-free DHFR enzyme system may be prepared directly from mycobacteria. Procedures for cell disruption and isolation of subcellular fractions may include probe sonication for disruption of mycobacteria. Alternatively, the Bead Beater (commercially available), which results in far better breakage of mycobacteria, may be used. The Bead Beater is used to obtain crude enzyme preparations. Before use, the Bead Beater is washed, rinsed with ethanol, and allowed to dry. The chamber is filled with sterilized glass beads and the breaking buffer added to wet beads. Breaking buffer consists of a 10 mM Tris-HCI solution (pH 7.4) with 150 mM NaCl, 10 mM EDTA, and 0.1% Tween 80. Following sterilization by autoclaving, the following are added before filter sterilization (per 100 ml of buffer): 1) DNase and RNase, 0.06%, 2) 23.3, µl of a 3 mg/ml stock of Pepstatin in ethanol(-20°C) 3) 50 µl of a 1 mg/ml stock of Leupeptin in ethanol (-20°C) 4) 200 µl of a 100 mM stock of PMSF in isopropanol (-20°C). mycobacteria are added to chamber in the breaking buffer solution

and pulsed for 30 seconds (x's 8) with cooling. 20-24 g of wet weight (from 6 liters of mycobacteria) was used in 25-ml of buffer. Following breakage, the amount of protein is quantitated by the Bradford procedure (M.M. Bradford, Anal. Biochem. 72:248-254, 1976) to determine optimum breakage (N. Ramasesh et al., Infect. Immun. 60:308-311, 1992). Disrupted cell suspension is transferred to 50-ml conical centrifuge tubes and centrifuged at 4,000 x g to pellet unbroken cells and glass beads. Supernatant is centrifuged at 40,000 x g for 2 hr to obtain a crude supernatant. This is filtered through a 0.45 μm filter (low-protein binding), followed by a 0.2 μm filter (low-protein binding) to remove any unbroken mycobacteria. Proteins are precipitated from the resulting supernatant by slow addition of ammonium sulfate, on ice with stirring, to 80% saturation. The resulting suspension is then stored at -20°C until used in assays. Prior to DHFR assay, a measured volume of the crude enzyme preparation is centrifuged at 2-8°C, the supernatant is discarded, and the pellet is dissolved in an equal volume of cold 50 mM potassium phosphate containing 1 mM EDTA, pH 7.

Example 16 Isolation and purification of DHFR.

The crude cell-free enzyme preparation of Example 15 (80% ammonium sulfate precipitate) can be centrifuged at 0-4°C, the supernatant discarded, and the pellet dissolved in cold standard buffer. The proteins can then be fractionated at 0-4°C using solid ammonium sulfate at successive saturations of 25, 40, 50, 60 and 80%. Enzyme activity can be measured in each fraction, after dissolution in standard buffer, and the active fractions pooled. The protein concentration of pooled fractions can be measured to determine specific activity (units/mg of protein). Further purification can be accomplished by gel filtration chromatography (A.M. Albrecht et al., Biochem. 8:960-967, 1969; A.M. Albrecht et al., Ag. Biochem. Biophys.

153:16-25, 1972), followed by methotrexate affinity chromatography (B.T. Kaufman, Meth. Enzymol. 34:272-81, 1974; T.D. Meek et al., Biochemist. 24:678-686, 1985; W. Sirawaraporn et al., Exper. Prasitol. 72:184-90, 1991). A 4% beaded agarose matrix is available from Sigma. Gel filtration chromatography can be done on a Sephadex G-75 column equilibrated with standard buffer. Fractions can be monitored for protein by measuring their absorbance at 280 nm and for DHFR activity. Active fractions can be pooled and total and specific enzyme activity determined. Gel filtration has been used successfully for the partial purification of *M. phlei* DHFR (M. Al-Rubeai et al., Biochem. J. 235:301-3, 1986). However, it may be necessary to add bovine serum albumin (1 mg/ml) to stabilize the enzyme. This step is initially performed on a small scale to determine the stability of the enzyme before proceeding further with the purification. Purified DHFR can be characterized by SDS PAGE to determine molecular weight.

Once purified enzyme preparations are obtained, the apparent K_m for dihydrofolate and NADPH may be determined at saturating concentrations of each substrate and the V_{max} may be determined using double reciprocal Lineweaver-Burk plots (I.H. Segal, *Biochemical Calculations*, 366-96, 1968). Double reciprocal plots with and without inhibitor can also be used to establish the nature of the inhibition (competitive or noncompetitive). The inhibition constant (K_i) of an inhibitor can be determined using Dixon plots of the reciprocal of the velocity vs. inhibitor concentration for increasing concentrations of DHFR.

Example 17 *In vitro* testing of drugs using a wild-type or recombinant enzyme system.

Table 3 presents the results of an experiment designed to establish that the substrate requirements for enzyme activity are those expected for DHFR and also that a linear relationship exists between rate of

reaction and amount of enzyme. The upper portion of Table 3 presents results which demonstrate the requirement for both NADPH and dihydrofolate for enzyme activity, measured as the decrease in A_{340nm} over time (30°C). In the absence of either substrate, the change in absorbance at 340 nm/min is 10 to 20-fold less than when both substrates are present. The lower portion of the table demonstrates a linear relationship between reaction rate and enzyme amount. The conclusion from these results is that the assay is selectively measuring DHFR activity in the crude cell-free extract. The small amount of change in absorbance in the presence of either NADPH or dihydrofolate alone was due to oxidase activity or instability of dihydrofolate, respectively.

Table 3
Requirements of *M. avium* DHFR
for NADPH and Dihydrofolate (FAH_2)
and Relationship of Rate to Enzyme Amount

µl Enzyme Preparation	NADPH	FAH_2	Δ A340nm/min.
50	+	-	-0.0070
50	+	+	-0.0740
50	-	+	-0.0041
50	+	+	-0.0930
15	+	+	-0.0230
30	+	+	-0.0440
30	+	+	-0.0470
60	+	+	-0.0950

Because of the demonstrated linear reaction rate with endogenous or wild type DHFR, recombinant enzyme of the present invention should demonstrate the same activity and thus may be used to screen and identify drugs and other agents with activity against DHFR. In one experiment to screen for potential antimycobacterial agents, the drugs

may be dissolved in DMSO followed by 10-fold serial dilutions in a final concentration of 10% DMSO. Ten µl of each drug dilution will be added to the reaction mixture having a total volume of 1 ml. Both enzyme and drug will be preincubated together for 3 minutes before the reaction is started by the addition of dihydrofolate. The reaction rate will be monitored both before and after dihydrofolate addition and the change in absorbance corrected for endogenous NADPH oxidase activity. The reaction rate, percent inhibition and approximate IC_{50} will be calculated to determine the effect, if any, of the agent on the DHFR.

The enzyme reaction rate may be calculated as follows:

Δ A340nm/minute corrected for oxidase activity (mean $\pm$ 0.0019, $n = 11$)

The % inhibition may be calculated as follows:

$$\frac{\Delta \text{ A340nm (0.1\% DMSO)} - \Delta \text{ A340nm (inhibitor)}}{\Delta \text{ A340nm (0.1\% DMSO)}} \times 100$$

The IC_{50} may be calculated as follows:

The amount of inhibitor required to inhibit the reaction rate by 50%.

Example 18 Determining Inhibitory Activity of DHFR Inhibitors in Various Systems.

As noted in Example 17, potential DHFR inhibitors can be tested in cell-free enzyme systems containing wild-type DHFR and/or enzyme systems containing the recombinant enzyme of the present invention. This is useful to obtain data for CoMFA, discussed below. In one cell-free enzyme system, DHFR activity may be measured at 30°C as the decrease in absorbance at 340 nm. The reaction mixture is adapted from that described by Al-Rubeai and Dale (M. Al-Rubeai et al., Biochem. J. 235:301-

3, 1986) for DHFR from *Mycobacterium phlei* and consists of 10 mM 2-mercaptoethanol, 0.1 mM NADPH, 0.1 mM dihydrofolate and 0.01-0.05 ml of enzyme in a standard buffer of 50 mM potassium phosphate- 1 mM EDTA, pH 7. The total assay volume is 1 ml. Dihydrofolate is added to initiate the reaction after the other components are preincubated for 3 minutes. Activity is corrected for NADPH oxidase activity that is measured in the absence of dihydrofolate during the preincubation period. For inhibition assays, the inhibitor and enzyme are added before the 3 minute preincubation period. One unit of enzyme is defined as that amount that reduces 1 μmole of DHFR per minute using a molar extinction coefficient at 340 nm, of 12,300 M^{-1cm-1} (B.L. Hillcoat et al., Anal. Biochem. 21:178-89, 1967). The IC_{50} can be determined as the amount of inhibitor required to inhibit the reaction rate by 50% under a defined set of conditions.

In addition to cell-free systems, intracellular activity of candidate compounds may also be evaluated. Two established cell lines available for determining intracellular activity of antimycobacterial drugs are the murine J774 macrophage cell line and the human Mono Mac 6 monocytic cell line (E.L. Wright et al., J. Clin. Microbiol. 34, 1996). Both cell lines can be infected with *M. tuberculosis* (E.L. Wright et al., J. Clin. Microbiol. 34, 1996) and *M. avium*. Essentially three concentrations of drug can be tested, the MIC value (predetermined in TAACF or broth microdilution assay at SRI), a concentration 3-fold above the MIC and a concentration 3-fold below the MIC concentration.

The testing procedure may consist of two stages. In the first stage, the toxicity of each drug may be tested using a MTT Cytotoxicity Assay Kit (Advanced Tissue Sciences) (E.L. Wright et al., J. Clin. Microbiol. 34, 1996). This can be done with both cell lines. The percent of untreated control for each dilution of a given test compound is plotted on the

y-axis vs. the concentration of the test compound on the x-axis. LD_{50} (*i.e.* IC_{50}) endpoints are then determined from the graph by reading from where the 50% point intercepts the Dose Response Curve to the concentration along the x-axis. That concentration is the LD_{50} value. Drugs that are cytotoxic at relevant concentrations will not be tested further, but information gained from this is useful in developing more selective and less toxic derivatives.

In a second stage of testing, macrophage cell lines may be used to examine the effectiveness of each drug to inhibit intracellular growth of *M. tuberculosis* and *M. avium* (*M. tuberculosis* H37Ra or H37Rv and *M. avium* serovar 4) and several *M. avium* and M. *tuberculosis* strains. Briefly, the J774 cell line is maintained in RPMI, containing 10% fetal bovine serum (FBS). At the time of infection with mycobacteria, the FBS is decreased to 1% to reduce division of J774 cells (N. Rastogi, Curr. Microb. 16:79-92, 1987). J774 cells are infected with mycobacteria for four hours, after which time cells are washed and fresh media added with appropriate concentrations of drugs. Cells are exposed to 3 different concentrations of each drug to be tested with the drugs being replenished daily, by topping off. The infection is allowed to continue for 5 days. J774 cells are lysed with sodium dodecyl sulfate at time zero (immediately after infection), and at five days to release the mycobacteria. Viable mycobacteria are enumerated by plating on 7h10 agar. In parallel with *in vitro* combination drug studies, lipophilic DHFR inhibitors may be evaluated in the macrophage cell lines for synergy in combination with the following drugs: isoniazid, ethambutol, and rifampin. Evaluation of the *in vitro* drug combination studies may be helpful before this phase is initiated and other antimycobacterial agents can be added to this scheme or replace one of the three listed here.

The Mono Mac 6 cell line may be obtained from the German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. This cell line is a human acute monocytic leukemia. Cells are grown as a suspension culture and are maintained in RPMI 1640, containing 10% (v/v) FCS, 2 mM L-glutamine, nonessential amino acids, 1 mM Na-pyruvic acid, and 9 μg/ml of bovine insulin (Sigma Chem. Co.; St. Louis, MO). The Mono Mac 6 cell line is the only human cell line to constitutively express phenotypic and functional features of mature monocytes (H.W.L. Ziegler-Heitbrock et al., Int. J. Cancer. 41:456-61, 1988). Unlike the U937 and THP-1 human monocytic cell lines that must be induced to develop phagocytic properties (F. Herrmann, J. Exp. Med. 162:1111-16, 1985), the MM6 cell line can constitutively phagocytize antibody-coated erythrocytes (H.W.L. Ziegler-Heitbrock et al., Int. J. Cancer. 41:456-61, 1988) and mycobacteria (J.S. Friedland et al., Cytokine 5:150-56, 1993; R.J. Shattock et al., J. Gen. Virology 75:849-56, 1994). For experiments involving drug inhibition studies, the J774 cell line may be infected with five mycobacteria per macrophage for four hours. For the Mono Mac 6, ten mycobacteria per macrophage are required, and overnight infection is necessary. The rest of the method is the same.

Activity of select compounds may also be evaluated in a mouse model. *M. tuberculosis* H37Rv infected mouse model may be used for single and combination drug evaluation involving 1-3 drugs, tested individually, and in combination with rifampin and in combination with isoniazid and in the *M. avium* infected Beige model.

In addition to the foregoing, other models known in the art for evaluating inhibition of DHFR of various species, such as mammalian systems, may be used, and the data from these experiments correlated with data obtained using the recombinant enzyme of the present invention.

Example 19 Use of CoMFA analysis.

3-D QSAR CoMFA analysis may be used to determine if there is any correlation between an agent's structure and its ability to inhibit DHFR. The IC_{50} values of existing compounds or newly synthesized analogs can be determined against isolated mycobacterial DHFR enzyme. The CoMFA analysis should encompass the areas critical to enzyme binding in order to obtain the most information from the analysis.

A CoMFA analysis of target drugs using inhibitory data from mammalian systems may also be done. Comparison of a mammalian binding model with the analysis derived from the mycobacterial enzyme inhibition data should increase the ability to predict alterations in structure that can not only enhance binding to the mycobacterial DHFR, but should also increase selectivity versus the mammalian enzyme. CoMFA analysis may also be done on the *M. avium* DHFR enzyme itself.

Actual CoMFA can be performed using the TRIPOS software (SYBYL 6.2 version). The analysis consists of the following: 1) establishing the conformation of each molecule; 2) superimposing each of the conformations from step 1; 3) calculating for each of the interaction energies with suitable probes at many points on a lattice; 4) performing a statistical analysis of the relationship between the interaction energies and the property of interest (IC_{50} for DHFR inhibition); and 5) displaying the 3-D QSAR coefficient contour map. The final contour map gives a real sense in space where the QSAR terms have high or low values and can be obtained for both the steric and electrostatic terms. A contour map represents in real space a picture of what portions of the molecule positively and negatively affect the biological property of interest either electrostatically or sterically. From these maps, one can predict what modifications can enhance biological activity. In fact, further computer analysis using the model for second

generation targets can predict biological activity of these new targets. One method of predicting new structures for synthesis and screening is based on the SYBYL program LEAPFROG. One can obtain a hypothetical binding cavity based on the CoMFA. This program can be used to predict new structures using molecular probes and binding energies to the cavity can be calculated for these structures.

Once the CoMFA is performed on a suitable number of agents, a rational approach based upon the inhibition data, the CoMFA results, and previous experience with antifolate drug design may be used to produce more selective second generation DHFR inhibitors. Computer modeling programs such as LEAPFROG can be used for second generation compounds and to test the validity of the CoMFA model by comparing predicted IC_{50} values with those obtained experimentally. As more compounds with data are generated, the CoMFA model can be updated to increase its predictive power and the quality of the hypothetical binding cavity. This work dovetails into and complements additional experiments once the crystal structures of inhibitors in the mycobacterial DHFR active site are actually obtained. Structures of compounds
in the active site can be compared to the minimized structures derived from CoMFA model, and the actual enzyme active site can be grossly compared to the putative DHFR cavity generated by LEAPFROG.

Other embodiments and uses of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. All documents, including U.S. patents and applications disclosed herein, are specifically incorporated herein by reference. U.S. provisional patent application serial number 60/034,725, filed January 3, 1997, and U.S. provisional patent application serial number 60/039,737, filed February 14, 1997, both of which are entitled

Mycobacterium avium fol A gene that encodes for the enzyme, Dihydrofolate Reductase (DHFR), are both specifically incorporated by reference. The specification and examples should be considered exemplary only with the true scope and spirit of the invention indicated by the following claims.

We Claim:

1. A recombinant nucleic acid comprising a sequence that encodes the DHFR protein of *Mycobacterium*.

2. The nucleic acid of claim 1 comprising SEQ ID NO 1.

3. The nucleic acid of claim 1 wherein the sequence comprises DNA, RNA or PNA.

4. The nucleic acid of claim 1 wherein the sequence further contains one or more additional sequences that direct transcription or translation.

5. The nucleic acid of claim 4 wherein the one or more additional sequences is a promoter, a polymerase binding site, an enhancer or a transcription or translation termination site.

6. A vector comprising the nucleic acid of claim 1.

7. A recombinant cell containing the nucleic acid of claim 1.

8. The cell of claim 7 wherein the nucleic acid is episomal or integrated into the genome.

9. The cell of claim 7 which is prokaryotic or eukaryotic.

10. A recombinant nucleic acid comprising a sequence that encodes at least a portion of the DHFR protein of *Mycobacterium*.

11. The nucleic acid of claim 10 wherein the portion is an enzymatically active portion.

12. The nucleic acid of claim 10 wherein the portion is an antigenically active portion.

13. The nucleic acid of claim 10 wherein the sequence encodes the entire amino acid sequence of a DHFR protein of *Mycobacterium*

14. The nucleic acid of claim 10 wherein the DHFR protein is derived from *M. avium*, *M. bovis*, *M. tuberculosis*, or *M. leprae*.

15. A recombinant peptide comprising the amino acid sequence of a *Mycobacterium* DHFR protein.

16. The peptide of claim 15 comprising SEQ ID NO 2.

17. The peptide of claim 15 that comprises a portion of the amino acid sequence of a *Mycobacterium* DHFR protein.

18. The peptide of claim 17 wherein the portion is an enzymatically active portion.

19. The peptide of claim 17 wherein the portion is an antigenically active portion.

20. The peptide of claim 17 wherein the sequence contains the entire amino acid sequence of a DHFR protein of *Mycobacterium*

21. The peptide of claim 17 wherein the DHFR protein is derived from *M. avium*, *M. bovis*, *M. tuberculosis*, or *M. leprae*.

22. The peptide of claim 17 wherein the DHFR protein is a purified DHFR protein derived from *M. avium*, *M. bovis* or *M. tuberculosis*.

23. The peptide of claim 20, wherein the DHFR protein is a purified DHFR protein derived from *M. avium*, *M. bovis* or *M. tuberculosis*.

24. A diagnostic kit comprising the peptide of claim 17.

25. A method of screening for an agent that inhibits the activity of a recombinant DHFR protein of a *Mycobacterium* comprising the steps of:

a. determining the activity of the DHFR protein upon incubation with each of a plurality of agents; and

b. selecting the agent that inhibits DHFR protein activity.

26. The method of claim 25 wherein the *Mycobacterium* is *M. avium*, *M. bovis*, *M. tuberculosis*, or *M. leprae*.

27. The method of claim 25 wherein the DHFR protein contains the sequence of SEQ ID NO 2 or an active portion thereof.

28. The method of claim 25 wherein the active portion is an enzymatically active portion.

29. The method of claim 25 wherein the active portion is an antigenically active portion.

30. The method of claim 25 wherein the plurality comprises greater than 10^2 different agents.

31. The method of claim 25 wherein the plurality of agents comprises a collection of related chemical compounds.

32. The method of claim 31 wherein the collection comprises chemical modifications of folate, methotrexate, trimethoprim, quinazoline or combinations thereof.

33. The method of claim 25 wherein incubating comprises mixing the DHFR protein with the agent under conditions that allow for a molecular interaction.

34. The method of claim 33 wherein the molecular interaction is binding.

35. The method of claim 25 wherein the activity inhibited is enzymatic activity.

36. The method of claim 25 wherein the activity inhibited is immunogenicity.

37. The method of claim 25 further comprising the step of determining the molecular conformation of said agent.

38. The method of claim 25 further comprising the steps of selecting a plurality of agents that inhibit the activity of the *Mycobacterium* DHFR protein, determining the molecular conformation of each agent selected and identifying a common inhibitory molecular conformation.

39. A method for assessing the ability of an agent to inhibit the activity of a DHFR protein comprising:

a. incubating a recombinant mycobacterial DHFR protein with said agent;

b. determining the activity of the incubated DHFR protein; and

c. comparing the activity determined with the wild-type activity of said protein.

40. The method of claim 39 wherein the activity is enzymatic activity.

41. The method of claim 39 wherein the activity is immunogenicity.

42. The method of claim 39 wherein the step of incubating comprises mixing the agent with the protein under physiological conditions.

43. The method of claim 39 wherein the agent selected is useful for the treatment of a mycobacterial infection.

44. The method of claim 43 wherein the *Mycobacterium* DHFR protein is of a specific species selected from the group consisting of *M. avium*, *M. bovis*, *M. tuberculosis*, and *M. leprae*, and the agent is specific for the treatment of the same species.

45. A method for selecting an antimycobacterial agent specific against a mycobacterial infection comprising:

a. analyzing a molecular conformation of a recombinant mycobacterial DHFR protein;

b. identifying a binding site within the molecular conformation; and

c. selecting the agent with a molecular structure that fits within the binding site.

46. The method of claim 45 wherein the binding site is a substrate binding site.

47. The method of claim 45 wherein the molecular conformation is analyzed from crystallographic analysis.

48. The method of claim 45 wherein the molecular conformation is analyzed from structure-activity relationships of the DHFR protein with a plurality DHFR protein substrates.

49. A method for identifying the sequence of a mycobacterial DHFR gene comprising:

a. amplifying nucleic acid in a biological sample containing *Mycobacterium* by a polymerase chain reaction with two probes wherein the sequences of the two probes span the *M. avium* DHFR gene; and

b. determining the sequence of the amplified nucleic acid to identify the sequence of the mycobacterial DHFR gene.

50. The method of claim 49 wherein the first probe contains a sequence from the 5' terminus of the *Mycobacterium avium* gene and the sequence of the second probe contains a sequence from the 3' terminus of the *Mycobacterium avium* DHFR gene.

51. The method of claim 49 wherein the sequence identified encodes the DHFR protein of *M. avium, M. bovis* or *M. tuberculosis* or *M. leprae.*

52. A method for detecting a mycobacterial infection comprising:

a. amplifying nucleic acid in a biological sample containing *Mycobacterium* by a polymerase chain reaction comprising two probes wherein the sequences of the two probes span the *M. avium* DHFR gene; and

b. detecting amplified nucleic acid that corresponds to an amplification of the nucleic acid between the two probes.

53. The method of claim 52 wherein the first probe contains a sequence from a terminus of the *Mycobacterium avium* DHFR gene and the sequence of the second probe contains a sequence from the opposite terminus of the *Mycobacterium avium* DHFR gene.

54. The method of claim 52 wherein the two probes are labeled with detectable labels.

55. The method of claim 54 wherein the detectable labels are selected from the group consisting of radioisotopes, stable isotopes, fluorescent chemical moieties, enzymes, metals and combinations thereof.

56. The method of claim 54 wherein detecting comprises determining the size or sequence of the nucleic acid amplified.

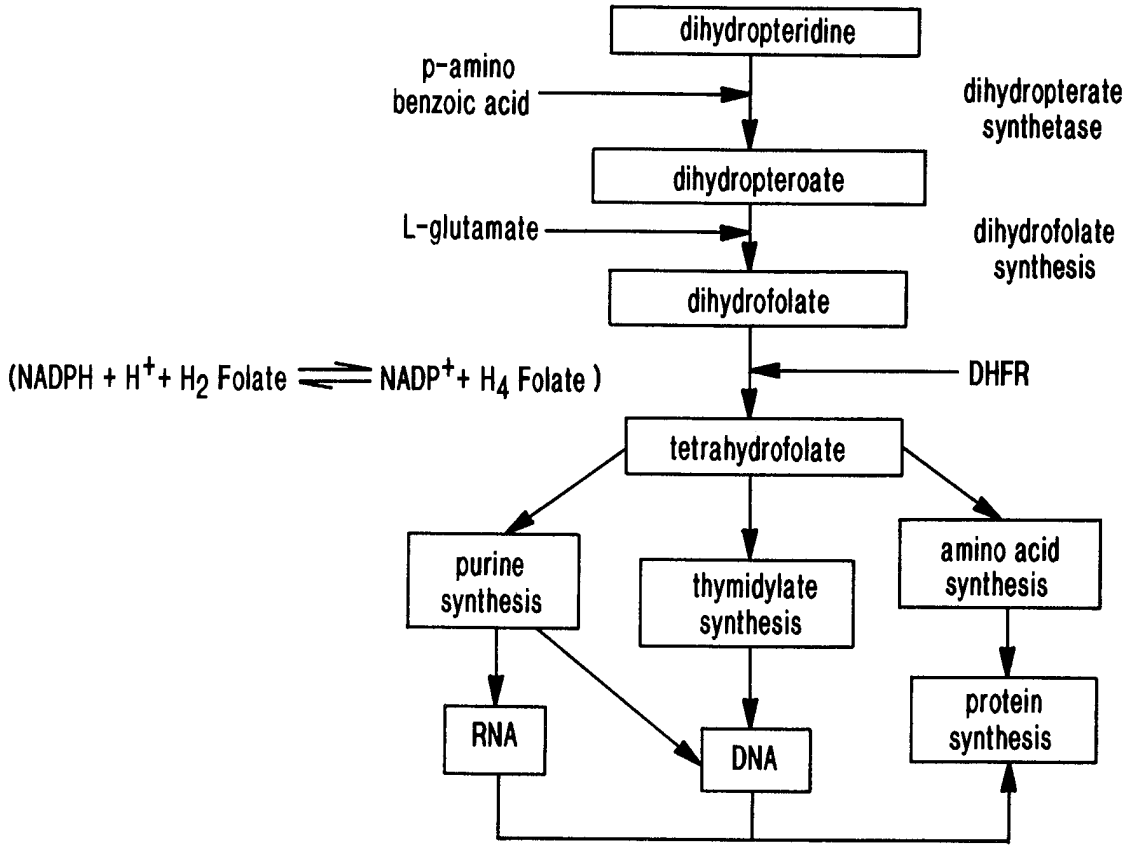

dihydropteridine
p-amino benzoic acid
dihydropterate synthetase
dihydropteroate
L-glutamate
dihydrofolate synthesis
dihydrofolate
(NADPH + H+ + H2 Folate ⇌ NADP+ + H4 Folate)
DHFR
tetrahydrofolate
purine synthesis
thymidylate synthesis
amino acid synthesis
RNA
DNA
protein synthesis

FIG. 1

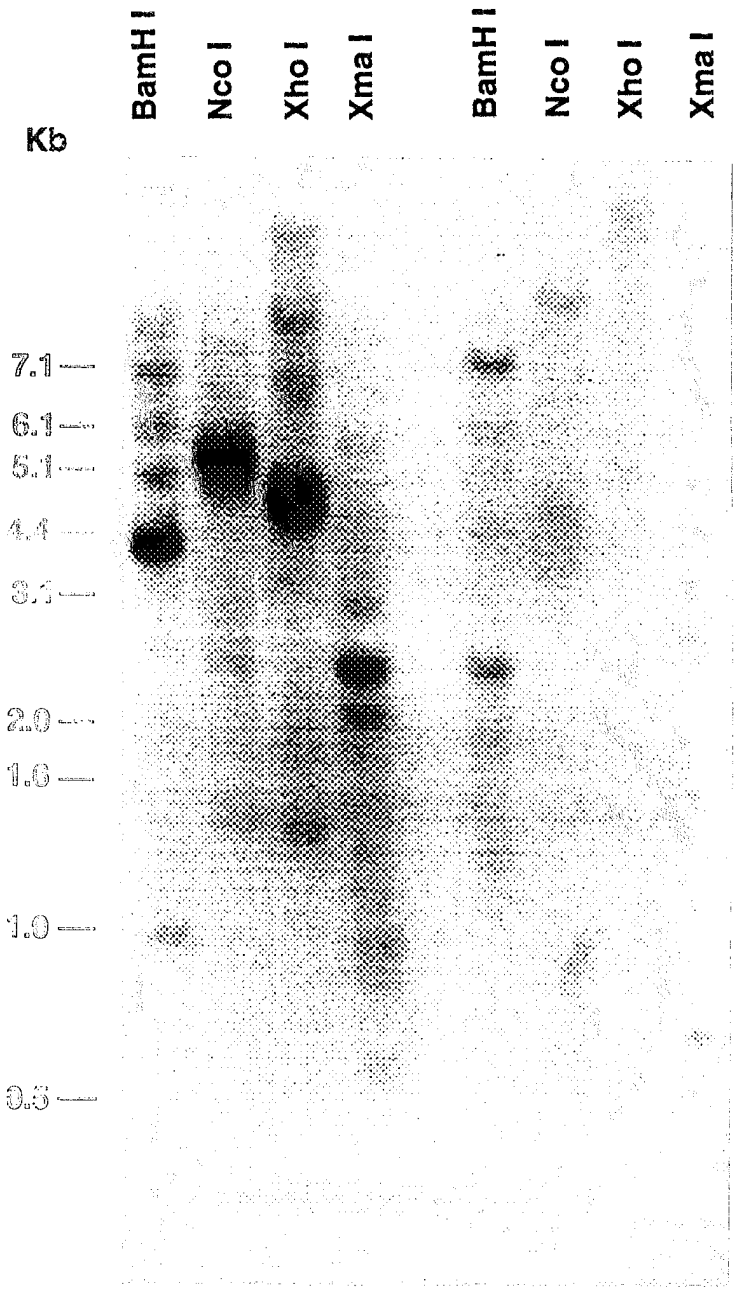

BamH I
Nco I
Xho I
Xma I
BamH I
Nco I
Xho I
Xma I
Kb
7.1
6.1
5.1
4.1
3.1
2.0
1.6
1.0
0.5

FIG. 2

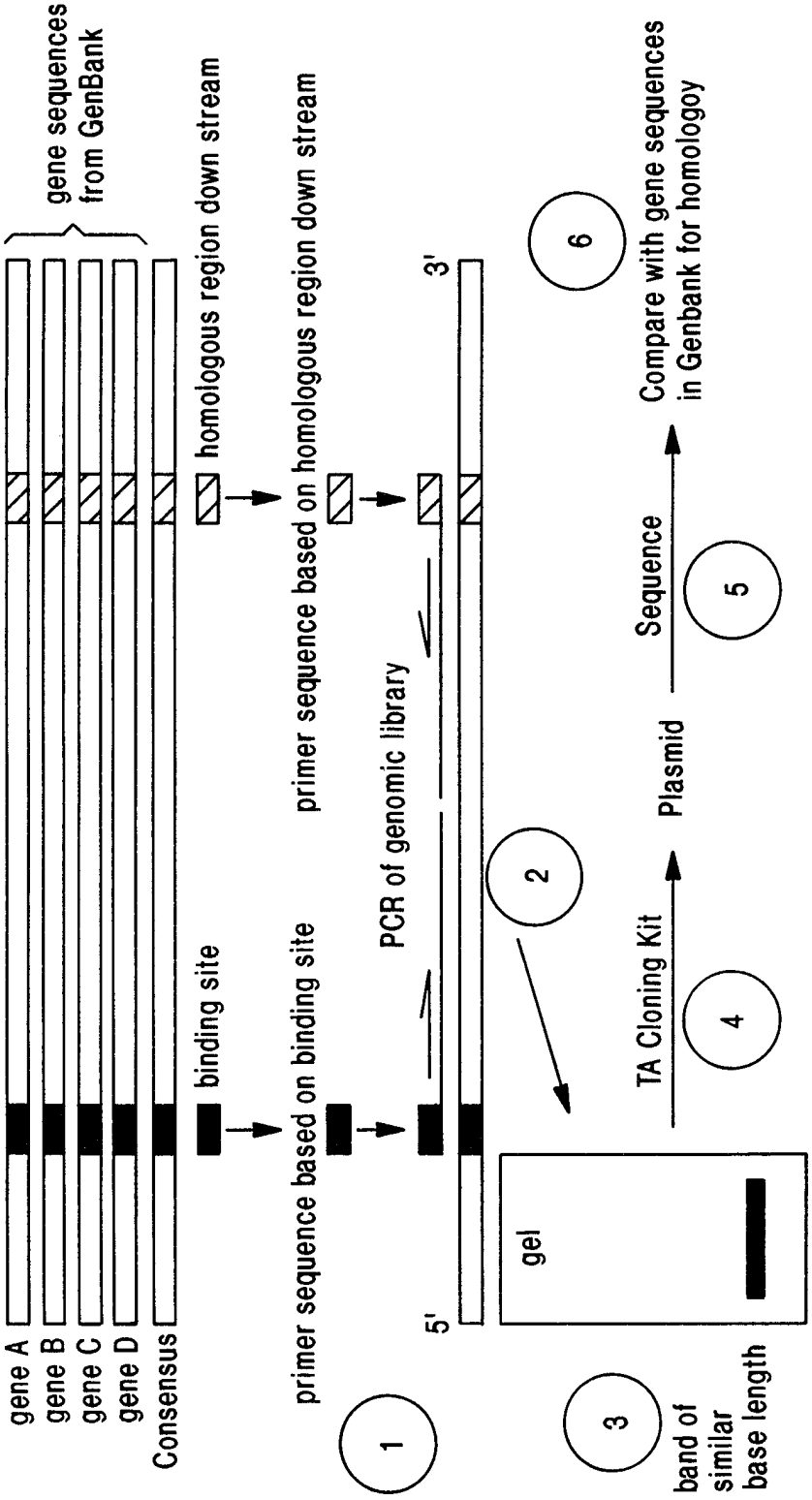

gene sequences from GenBank
gene A
gene B
gene C
gene D
Consensus
binding site
homologous region down stream
primer sequence based on binding site
primer sequence based on homologous region down stream
PCR of genomic library
5'
3'
1
2
3
4
5
6
gel
band of similar base length
TA Cloning Kit
Plasmid
Sequence
Compare with gene sequences in Genbank for homologoy

FIG. 3

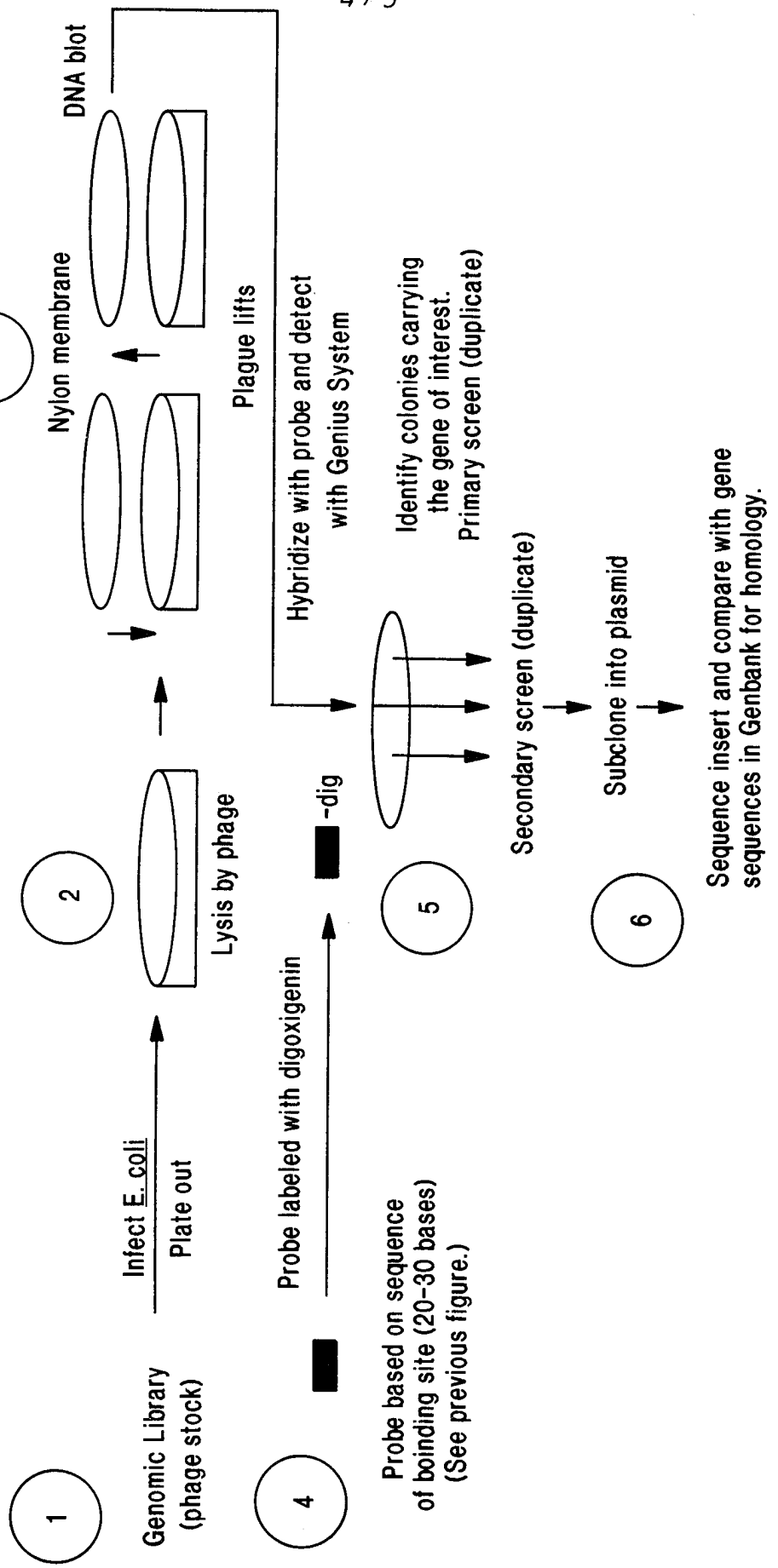

1
Genomic Library
(phage stock)
Infect E. coli
Plate out
2
Lysis by phage
3
Nylon membrane
DNA blot
Plague lifts
4
Probe based on sequence
of boinding site (20-30 bases)
(See previous figure.)
Probe labeled with digoxigenin
-dig
Hybridize with probe and detect
with Genius System
5
Identify colonies carrying
the gene of interest.
Primary screen (duplicate)
Secondary screen (duplicate)
6
Subclone into plasmid
Sequence insert and compare with gene
sequences in Genbank for homology.

FIG. 4

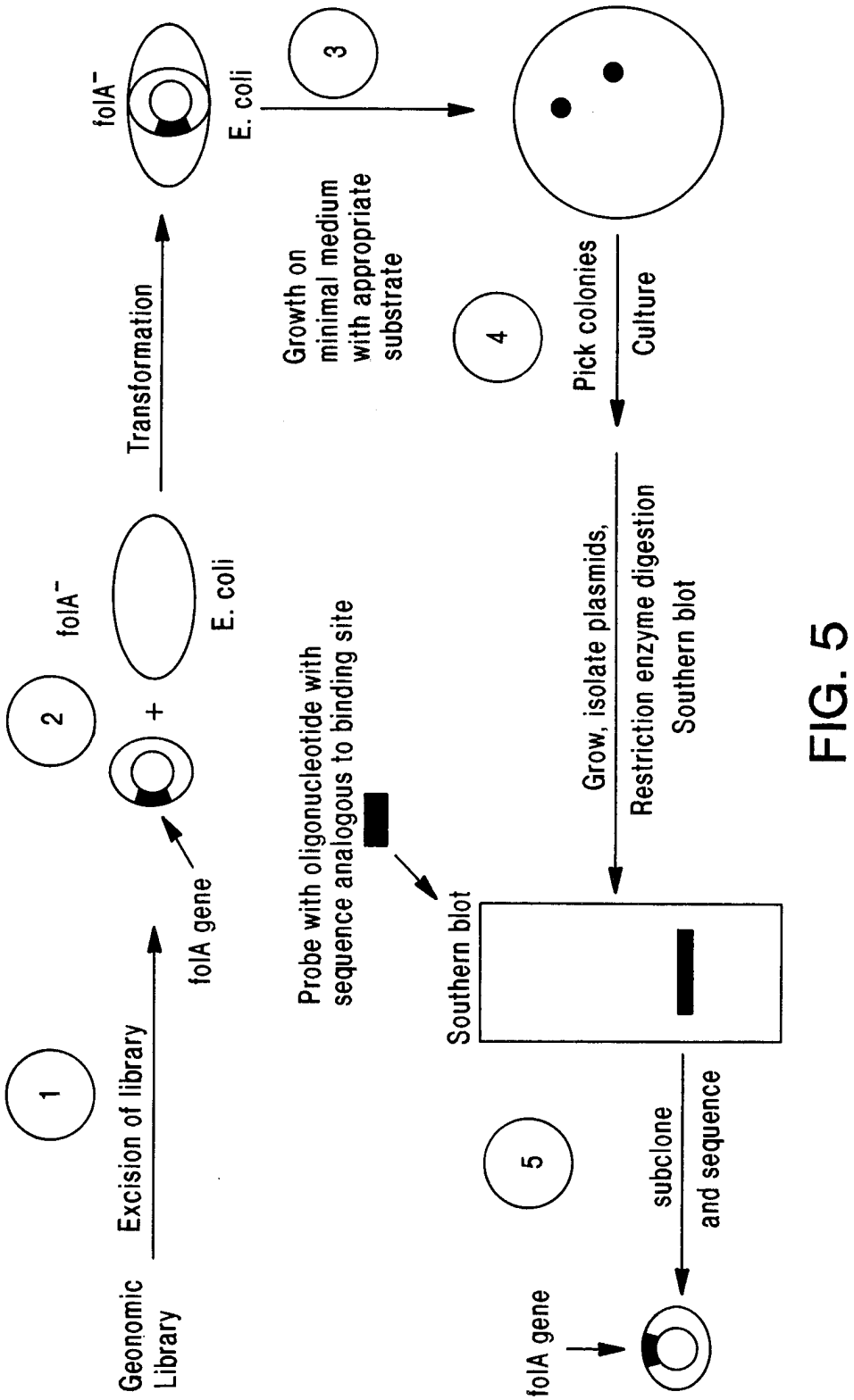

1
Excision of library
Geonomic Library
folA gene
2
folA⁻
E. coli
Transformation
folA⁻
E. coli
3
Growth on minimal medium with appropriate substrate
4
Pick colonies
Culture
Grow, isolate plasmids,
Restriction enzyme digestion
Southern blot
Probe with oligonucleotide with sequence analogous to binding site
Southern blot
5
subclone
and sequence
folA gene

FIG. 5

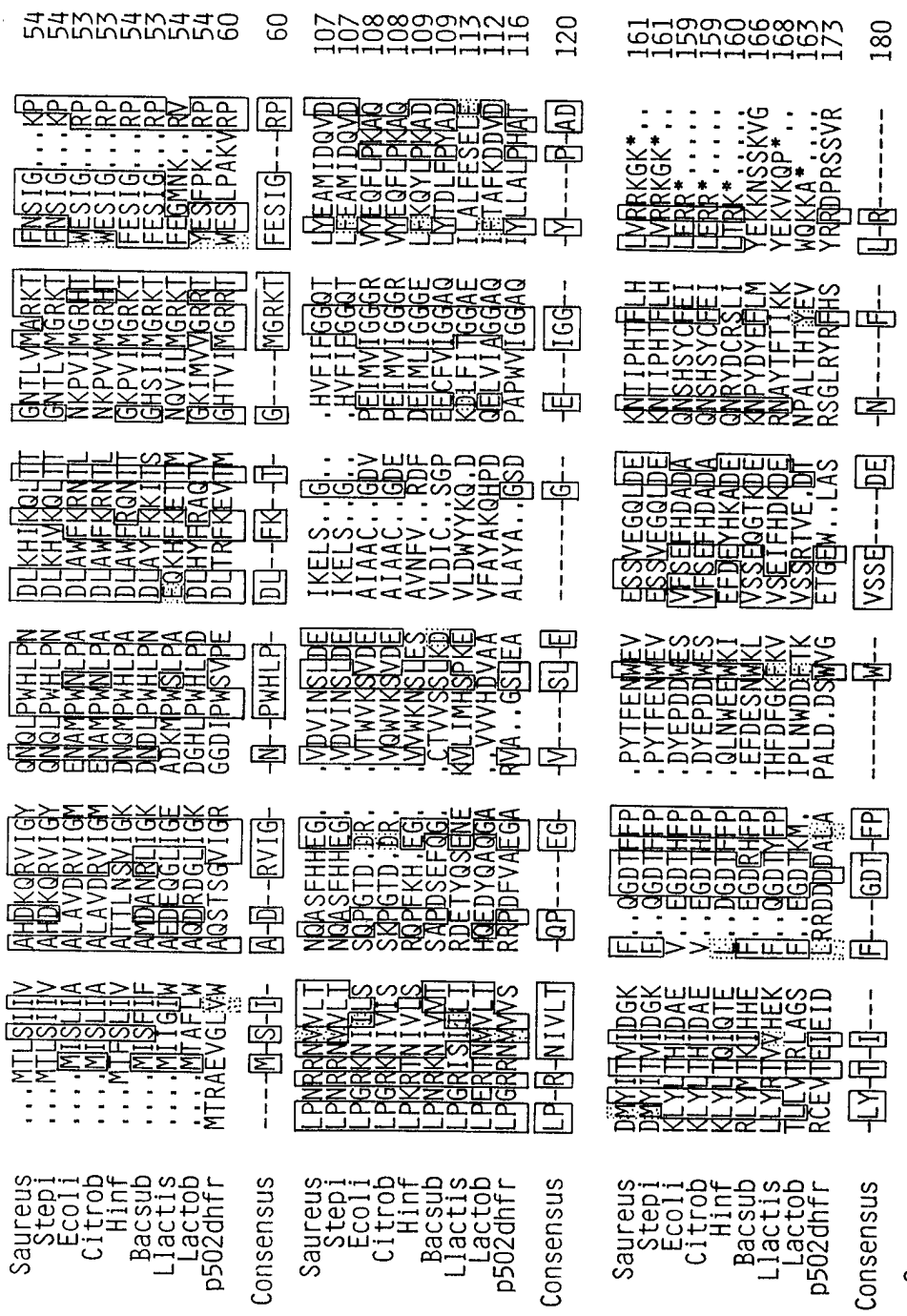


FIG. 6

```
1   GACGTCGTGGTGCACAACTACGATCCGCACCCGGCCATCAAGGCCCCCGTCGCGGTATGA  60
                                                        M   T

61  CCCGTGCCGAGGTGGGCCTGGTGTGGGCCCAGTCGACGTCTGGCGTCATCGGCCGCGGCG 120
      R   A   E   V   G   L   V   W   A   Q   S   T   S   G   V   I   G   R   G   G

121 GTGACATCCCGTGGAGCGTGCCCGAGGACCTCACCCGGTTCAAAGAGGTGACCATGGGC 180
      D   I   P   W   S   V   P   E   E   D   L   T   R   F   K   E   V   T   M   G   H

181 ACACCGTGATCATGGGCCGGCGGACCTGGGAGTCGTTGCCGGCCAAGGTGCGGCCGCTGC 240
      T   V   I   M   G   R   R   T   W   E   S   L   P   A   K   V   R   P   L   P

241 CCGGCCGGCGCAACGTGGTGGTGTCCCGCCGGCCCGACTTCGTCGCCGAGGGGGCGCGGG 300
      G   R   R   N   V   V   V   S   R   R   P   D   D   F   V   A   E   G   A   R   V

301 TGGCCGGGTCGCTGGAGGCGGCCCTCGCGTACGCCGGGAGCGACCCGGCGCCGTGGGTGA 360
      A   G   S   L   E   A   A   L   A   Y   A   G   S   D   P   A   P   W   V   I

361 TCGGCGGCGCGCAGATCTATCTGCTGGCGCTGCCCCATGCCACCCGCTGCGAGGTCACCG 420
      G   G   A   Q   I   Y   L   L   A   L   P   H   A   T   R   C   E   V   T   E

421 AAATCGAGATCGACCTGCGCCGCGACGACGACGCGCTGGCGCCGGCGCTGGACGACA 480
      I   E   I   D   L   R   R   D   D   D   D   A   L   A   P   A   L   D   D   S

481 GCTGGGTAGGCGAGACGGGCGAGTGGCTGGCCAGCCGCTCCGGGCTGCGGTACCGGTTCC 540
      W   V   G   E   T   G   E   W   L   A   S   R   S   G   L   R   Y   R   F   H

541 ACAGCTACCGTCGGGACCCGCGCTCTTCCGTTCGGCGCTGTTCGCCCTCACGCCCGAGCT 600
      S   Y   R   R   D   P   R   S   S   V   R   G   C   S   P   S   R   P   S   *

601 GACATACTCGGACGCGGGGGTCGGTCACACACCGTCTACCAGCCGCTGTTCGGGAAAAGG 660
```

FIG. 7

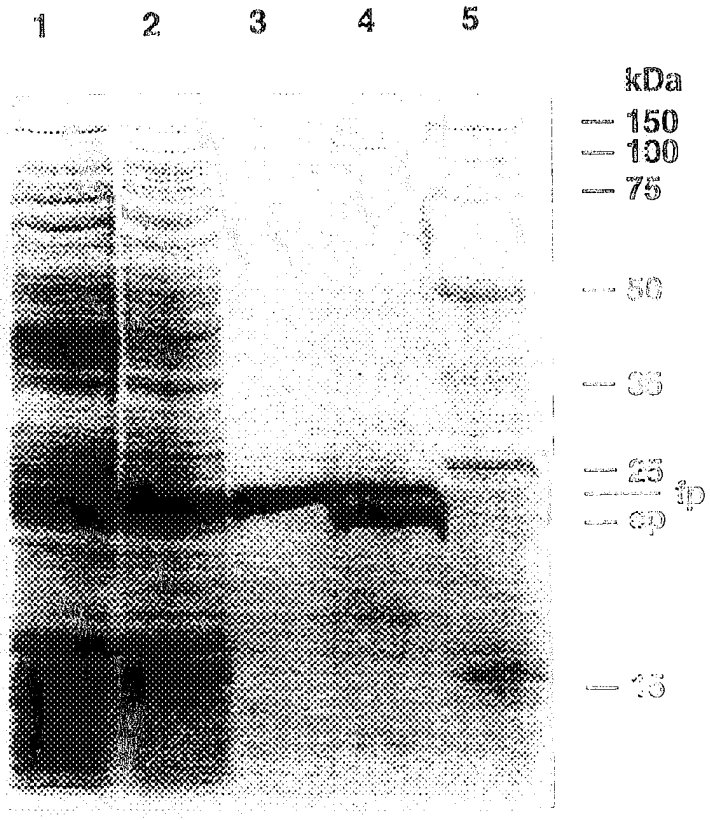


FIG. 8

M. avium M T R A E V G L V W A Q S T S G V I G

M. smegmatis S M S L I ? A Q ? T G G I I S

FIG. 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23557

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1, 23.2, 23.7, 24.3, 24.33; 530/350, 825; 435/4, 69.1, 70.1, 71.1, 91.4, 183, 189, 252.3, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, GENBANK, STN CAplus, MEDLINE, DISSERTATION ABSTRACTS
mycobacteria, dihydrofolate reductase, sequence, inhibit, inhibitor.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	Database on GENBANK. DALE, J. 'Mycobacterium tuberculosis folA gene (putative) for dihydrofolate reductase.' Genbank Accession Number X59271, 9 June 1991. See both amino acid and nucleotide sequence	1, 3, 10-14 --- 15, 17-24
X, P ---- Y, P	Database on GENBANK. PARKHILL, J. Mycobacterium tuberculosis sequence v002. GENBANK Accession Number AL008967. 18 November 1997. See Keywords, Amino acid sequence "complement 24753-25232 /gene dhfr", and nucleic acid sequence.	1, 3-7, 9-14 ---- 8, 15, 17-24

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
13 MARCH 1998	15 JUN 1998

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	MARY K ZEMAN
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23557

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	PHILIPP, W.J. An Integrated Map of the Genome of the Tubercle Bacillus, Mycobacterium tuberculosis H37Rv, and Comparison with Mycobacterium leprae. Proc. Nat'l. Acad. Sci. April 1996, Vol. 93 No. 7, pages 3132-3137. See Table 2 for cosmid and locus of dhfr gene, and Figure 3 for location on genomic map.	1, 3-7, 9-14 --------- 8, 15, 17-23
X --- Y	SCUDERI, J.D. Molecular Genetic and Biochemical Analysis of Thymidylate Synthase and Dihydrofolate Reductase from Mycobacteria. Dissertation Abstracts International. 1996, Vol. 57 No. 8B page 4874B-4875B. See Entire Abstract.	1, 3-15, 17-23, ---------- 16, 49-56
Y	CZAPLINKSI, K-H et al. New Benzylpyrimidines: Inhibition of DHFR from Various Species. QSAR, CoMFA, and PC analysis. Eur. J. Med. Chem. 1995, Vol. 30, pages 779-787. See Summary and Results sections.	25, 26, 28-48
Y	KANSY, M. et al. Syntehsis of New 2,4-diamino-5-benzylpyrimidines Active Against Various Bacterial Species. Eur. J. Med. Chem. 1992, Vol. 27, pages 237-244. See Summary and Results sections.	25, 26, 28-48

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23557

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 2, 16, 27
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

No Sequence disk or searchable sequence information was provided.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23557

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6): C07H 21/04, 21/00; C07K 1/00; C12N 1/19, 9/00, 9/02, 15/00; C12P 21/00; C12Q 1/00, 1/26, 1/68

A. CLASSIFICATION OF SUBJECT MATTER:
US CL : 536/23.1, 23.2, 23.7, 24.3, 24.33; 530/350, 825; 435/4, 69.1, 70.1, 71.1, 91.4, 183, 189, 252.3, 320.1