Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*

Richard A. Young*, Vijay Mehra*||, Douglas Sweetser*,
Thomas Buchanan†, Josephine Clark-Curtiss‡,
Ronald W. Davis§ & Barry R. Bloom||

* Whitehead Institute for Biomedical Research, Nine Cambridge
Center, Cambridge, Massachusetts 02142, and Department of
Biology, Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA
† Immunology Research Laboratory, Pacific Medical Center, and
Departments of Pathobiology and Medicine, University of
Washington, Seattle, 98144, USA
‡ Department of Biology, Washington University, St Louis,
Missouri 63130, USA
§ Department of Biochemistry, Stanford University School of
Medicine, Stanford, California 94305, USA
|| Department of Microbiology and Immunology, and Irvington
House Institute for Medical Research, Albert Einstein College of
Medicine, Bronx, New York 10461, USA

Leprosy, a chronic infectious disease afflicting between 10 and 15 million people, is caused by the obligate intracellular parasite *Mycobacterium leprae*[1]. Although *M. leprae* was the first identified bacterial pathogen of man[2], basic biochemical, immunological, diagnostic and therapeutic investigations have been severely limited because it remains one of the few human pathogens that have not been cultured *in vitro*. An *M. leprae* recombinant DNA expression library was constructed to provide a source of genes encoding proteins relevant for such studies. Monoclonal antibodies directed against *M. leprae* specific antigens[3-7] have been used to isolate the genes encoding the five most immunogenic protein antigens of the leprosy bacillus. We report here that *M. leprae* specific epitopes recognized by all of 13 monoclonal antibodies tested were produced by recombinant phage in *Escherichia coli*.

A recombinant DNA expression library of *M. leprae* DNA was constructed using λgt11, a bacteriophage vector that is capable of driving the expression of foreign insert DNA with *E. coli* transcription and translation signals[8-10]. An attempt was made to increase the likelihood that all possible foreign coding sequences would be expressed in *E. coli* by using an approach that proved successful in the isolation of *Mycobacterium tuberculosis* genes[11]. *M. leprae* was purified from an armadillo that had been inoculated with bacillus from a single human patient[12]; DNA was purified from the bacillus[13] and was mechanically sheared to produce fragments of 1–7 kilobases (kb). *Eco*RI linkers were added to the ends of the DNA fragments to allow insertion at the unique *Eco*RI site of λgt11. The ligated recombinant DNA was packaged into phage heads and this material was used to infect *E. coli* cells[14]. The aim of this approach was to generate DNA fragments with random endpoints throughout the foreign genome and to then produce recombinant phage in sufficient numbers that insert endpoints occurred at each base pair in the pathogen genome. This strategy should ensure that all coding sequences are inserted in the correct transcriptional orientation and translational frame to be expressed as a fusion protein with the β-galactosidase encoded in λgt11.

The *M. leprae* DNA library constructed in this manner contained 2.5×10^6 individual recombinant phage. This library was amplified in *E. coli* Y1088 (ref. 9) by producing a plate stock[14] whose titre was 2×10^{11} PFU (plaque-forming units) ml^{-1}. The amplified library consisted of 25% recombinants whose foreign DNA insert lengths averaged 2 kb, as determined by DNA restriction endonuclease analysis of 25 independent phage clones. As the *M. leprae* genome consists of $\sim 10^6$ base pairs (bp)[13], it is likely that this library comprehensively represents the DNA of the bacillus.
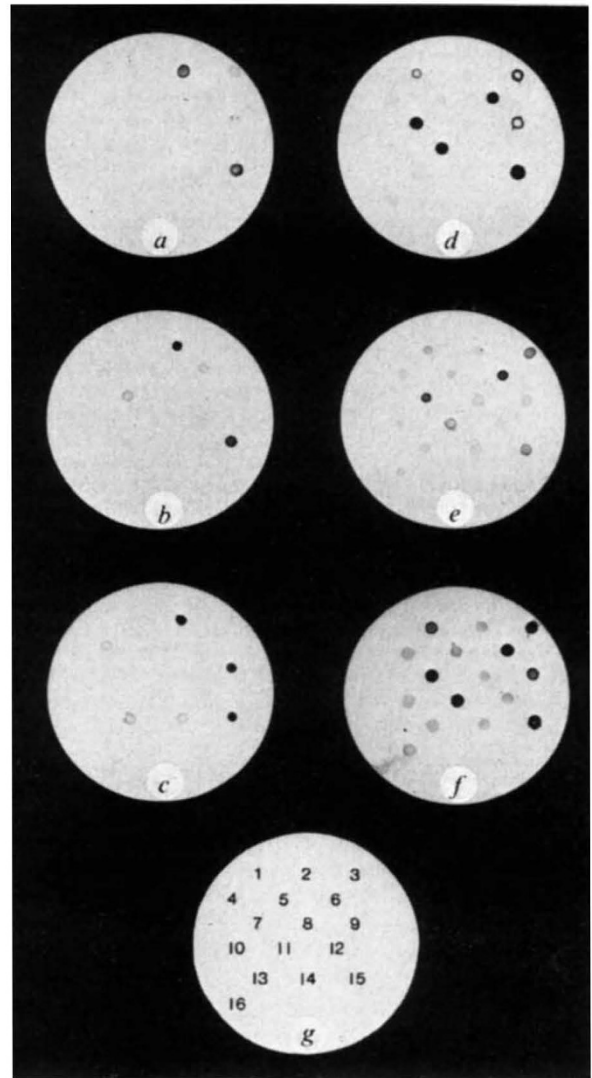
Fig. 1 Arrays of antigen from *M. leprae* recombinant DNA clones probed with individual monoclonal antibodies. Drops containing $\sim 10^4$ PFU each of 15 cloned λgt11 recombinants were arrayed on lawns of *E. coli* Y1090. The phage were grown and the antigens blotted and probed with individual monoclonal antibodies at $\sim 1:200$ dilution as described in ref. 11. The monoclonal antibodies used here are: *a*, MLIIC8; *b*, MLIIIC8; *c*, MLIIIE9; *d*, MLIIH9; *e*, Y1-2; and *f*, Cl-1. The recombinant DNA clones are coded in *g*: 1, Y3159; 2, Y3160; 3, Y3161; 4, Y3162; 5, Y3165; 6, Y3166; 7, Y3170; 8, Y3171; 9, Y3172; 10, Y3173; 11, Y3174; 12, Y3175; 13, Y3176; 14, Y3177; 15, Y3178; 16, λgt11 (this phage plaque produced the background signal for each monoclonal antibody).

The monoclonal antibodies were produced in mice immunized with intact or crude extracts of armadillo-derived, purified *M. leprae*[3-7], and the sizes of the antigens to which they bind are listed in Table 1. All of the antibodies directed against the antigen of relative molecular mass (M_r) 65,000 (65K) were pooled at approximately 1:200 dilution and used to probe 10^6 plaques (0.25×10^6 recombinant plaques) according to protocols described previously[10,11]. Seventeen plaques produced signals; 15 of these were successfully purified to homogeneity in one or two successive rescreens with the antibody pool.

Recombinant DNA clones isolated in this manner were then arrayed and probed with each of the monoclonal antibodies individually. Figure 1 shows the results obtained with six of the seven antibodies directed against the 65K antigen. All seven monoclonal antibodies, each directed against a different epitope (T. P. Gillis, R. A. Miller, D. B. Young, S. R. Khanolkar and T.B., personal communication), were capable of recognizing antigen produced in *E. coli*. At least four different signal patterns

65K

36K

28K

18K

12K
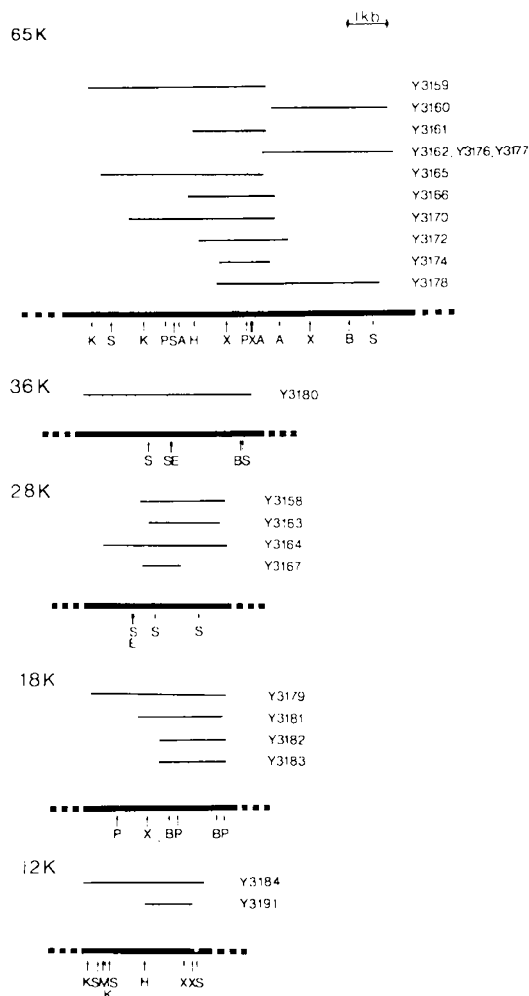
Fig. 2 Restriction maps of *M. leprae* DNA. Recombinant DNA clones isolated by using antibodies directed against the 65K, 36K, 28K, 18K and 12K antigens were subjected to restriction endonuclease mapping. λ DNA was prepared from phage plate stocks according to Davis *et al.*[16]. A, *Sac*I; B, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; M, *Bam*HI; P, *Pvu*I; S, *Sal*I; X, *Xho*I.

	M. leprae	
Antibody	antigen	Ref.
MLIIC8	65K	4
MLIIIC8	65K	4
Y1-2	65K	7
MLIIH9	65K	4
MLIIIE9	65K	4
CI-1	65K	7
ML-30	65K	3
F47 CL9.1	36K	7
SA1.D2D	28K	6
SA1.B11H	28K	6
L7-15	18K	7
ML-06	12K	3

Table 1 Monoclonal antibodies used to isolate *M. leprae* genes

Monoclonal antibodies directed against *M. leprae* protein antigens. The mouse monoclonal antibodies used in this study are tabulated according to the size of the *M. leprae* antigen that they bind and a reference describing their source. All of the antibodies are IgG1.

were observed in the array of clones. Antibodies MLIIC8, MLIIIC8 and MLIIIE9 (Fig. 1*a*, *b* and *c*, respectively) produced three distinct patterns. A fourth pattern was generated by antibodies MLIIH9, Y1-2 and C1-1 (Fig. 1*d*, *e* and *f*, respectively). Antibody ML-30, directed against an epitope shared with *E. coli*, produced strong signals with all clones, generating a poorly discernible pattern (not shown). There was considerable variation in the number of different epitopes produced by each clone, as evidenced by their reaction with different monoclonal antibodies. While some clones produced antigen that was recognized by only one antibody (for example, clone 13 in Fig. 1*c*), and one of the clones produced antigen recognized by all seven antibodies (clone 15), most clones produced antigen that was bound by some intermediate fraction of the seven antibodies.

Recombinant DNA clones were also isolated and characterized with antibodies directed against the 36K, 28K, 18K and 12K antigens. Approximately 10^6 λgt11 plaques were screened with a pool of these monoclonal antibodies. Eleven plaques that produced signals were purified to homogeneity. Clones were arrayed as in Fig. 1 and probed with each of the individual antibodies that comprised the pool. The anti-36K, 28K, 18K and 12K antibodies produced signals with 1, 4, 4 and 2 recombinant clones, respectively.

The insert DNAs of all the recombinant DNA clones isolated using these monoclonal antibodies were mapped with restriction endonucleases. Figure 2 shows the genomic DNA restriction

maps deduced for genes encoding each of the five antigens of interest and illustrates how each of the cloned DNAs aligns with that map. All clones isolated with monoclonal antibodies directed against any single antigen appear to align with a single genomic DNA segment. For example, the insert DNAs of the 12 recombinatnt clones isolated with the seven different anti-65K monoclonal antibodies overlap sufficiently to allow the construction of a unique genomic DNA restriction map with which all clones align unambiguously. This result indicates that all the anti-65K antibodies recognize epitopes encoded in a single DNA locus, presumably in a single gene. Similarly, the multiple clones isolated with the anti-28K, 18K or 12K antibodies produced overlapping restriction maps.

One concern with the approach used here is that some recombinant clones may be isolated not because they express the protein of interest but because they express an unrelated polypeptide containing a similar or identical immunological determinant. However, when multiple recombinant DNA clones were isolated by using a single monoclonal antibody, all contained overlapping DNA; this suggests that each of the epitopes of interest is encoded by a single genomic DNA segment.

These results attest to the power of the approach used here to detect and isolate specific antigen-coding sequences from λgt11 recombinant DNA expression libraries. It is striking that all of the monoclonal antibodies used in the present study could recognize antigen produced by recombinant phage in *E. coli*. The nature of the epitopes recognized by these antibodies is presently under investigation. Some of the monoclonal antibodies used here bind small linear peptide determinants. For example, at least two of the anti-65K monoclonal antibodies bind antigen from λgt11 subclones of Y3178 (Fig. 2) which contain less than 100 bp of insert DNA, encoding at most 33 amino acids (R.A.Y., unpublished data).

In summary, the recombinant DNA strategy used here to express the coding capacity of the *M. leprae* genome in *E. coli* and to detect specific antigenic determinants with monoclonal antibodies, appears to be effective. Each of the genes of interest was isolated by screening no more than 10% of the recombinant DNA library. These results suggest that a similar approach is feasible with genomic DNA from parasites having more complex genomes, such as schistosomes, filaria, leishmania and plasmodia, parasites that cause disease in hundreds of millions of people.

The availability of antigens produced by recombinant DNA technology may make it possible to address some problems in leprosy that cannot be approached by other means. It should be possible to develop simple and specific seroepidemiological assays using these protein antigens, in order to screen populations for individuals producing antibodies to *M. leprae*-specific antigenic determinants. This could make early diagnosis of

leprosy feasible, permitting early treatment to reduce its transmission and prevent the development of nerve damage and deformity.

As resistance to leprosy is provided by cell-mediated immunity[1], the availability of antigens produced by recombinant DNA technology should allow experiments to be performed to address the basic immunological question of whether epitopes recognized by monoclonal antibodies are also recognized by T cells. In addition, a mixture of peptides recognized by helper T cells could, for the first time, provide a specific skin test antigen for use in assessing the immunological status of patients and their contacts—this is urgently needed to permit rapid evaluation of the immunological efficacy of the candidate vaccines against leprosy that are being developed[15]. Finally, the identification of antigens involved in cellular immunity could lead to the development of a new generation of vaccines. Because of the intrinsic adjuvant properties of mycobacteria, one attractive approach is to introduce genes whose products provide protection into cultivatable mycobacteria, to produce a vaccine capable of engendering long-lasting cell-mediated immunity.

1. Bloom, B. R. & Godal, T. *Rev. Infect. Dis.* **5**, 765-780 (1983).
2. Hansen, G. A. *Norsk Mag. Laegezitensk.* **4**, 1-88, I-L III (1874).
3. Ivanyi, J., Morris, J. A. & Keen, M. in *Monoclonal Antibodies Against Bacteria* (eds Macario, A. J. L. & Macario, E. C.) (Academic, New York, 1984).
4. Gillis, T. P. & Buchanan, T. M. *Infect. Immunity* **37**, 172 (1982).
5. Coates, A. R. M., Allen, B. W., Hewitt, J., Mitchison, D. A. & Ivanyi, J. *Lancet* ii, 167 (1981).
6. Young, D. B., Fohn, M. J., Khanolkar, S. R. & Buchanan, T. M. *Clin. exp. Immun.* (in the press).
7. Engers, H. *et al. Infect. Immunity* **48**, 603-605 (1985).
8. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194-1198 (1983).
9. Young, R. A. & Davis, R. W. *Science* **222**, 778-782 (1983).
10. Young, R. A. & Davis, R. W. in *Genetic Engineering: Principles and Techniques* Vol. 7 (eds Setlow, J. & Hollaender, A.) 29-41 (Plenum, New York, 1985).
11. Young, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 2583-2587 (1985).
12. Draper, P. *Int. J. Lepr.* **44**, 95-98 (1976).
13. Clark-Curtiss, J. E., Jacobs, W. R., Docherty, M. A., Ritchie, L. R. & Curtiss, R. *J. Bact.* **161**, 1093-1102 (1985).
14. Huynh, T., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D.) (IRL, Eynsham, in the press).
15. Bloom, B. R. & Mehra, V. in *New Approaches to Vaccine Development* (eds Bell, R. & Torrigiani, G.) 368-389 (Schwabe & Co., Basel, 1984).
16. Davis, R. W., Botstein, D. & Roth, J. R. *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, New York, 1980).

Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG

R. B. Parekh*, R. A. Dwek*, B. J. Sutton†,
D. L. Fernandes*, A. Leung‡, D. Stanworth‡
& T. W. Rademacher*

Departments of *Biochemistry and †Molecular Biophysics, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
‡ Department of Immunology, Rheumatoid and Allergy Research Unit, University of Birmingham, The Medical School, Vincent Drive, Birmingham B15 2TJ, UK

T. Mizuochi§, T. Taniguchi§, K. Matsuta||,
F. Takeuchi||, Y. Nagano||, T. Miyamoto||
& A. Kobata§

§ Department of Biochemistry, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108, Japan
|| Department of Medicine and Physical Therapy, Faculty of Medicine, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Rheumatoid arthritis (RA) is a widely prevalent (1–3%) chronic systemic disease thought to have an autoimmune component[1]; both humoral[1-4] and cellular[5,6] mechanisms have been implicated. Primary osteoarthritis (OA) is considered to be distinct from rheumatoid arthritis, and here damage is thought to be secondary to cartilage degeneration. In rheumatoid arthritis, immune complexes are present that consist exclusively of immunoglobulin[7], implying that this is both the 'antibody' (rheumatoid factor [RF]) and the 'antigen' (most commonly IgG). Autoantigenic reactivity has been localized to the constant-region ($C_\gamma 2$) domains of IgG[8,9]. There is no evidence for a polypeptide determinant but carbohydrate changes have been reported[10]. We have therefore conducted a study, simultaneously in Oxford and Tokyo, to compare in detail the N-glycosylation pattern of serum IgG (Fig. 1) isolated from normal individuals and from patients with either primary osteoarthritis or rheumatoid arthritis. The results, which required an evaluation of the primary sequences of ~1,400 oligosaccharides from 46 IgG samples, indicate that: (1) IgG isolated from normal individuals, patients with RA and patients with OA contains different distributions of asparagine-linked bi-antennary complex-type oligosaccharide structures, (2) in neither disease is the IgG associated with novel oligosaccharide structures, but the observed differences are due to changes in the relative extent of galactosylation compared with normal individuals. This change results in a 'shift' in the population of IgG molecules towards those carrying complex oligosaccharides, one or both of whose arms terminate in N-acetylglucosamine. These two arthritides may therefore be glycosylation diseases, reflecting changes in the intracellular processing, or post-secretory degradation of N-linked oligosaccharides.

At least 30 different complex-type bi-antennary oligosaccharides are associated with human serum IgG (Fig. 2). To compare the molar proportions of each of these structures, each serum IgG sample was subjected to controlled hydrazinolysis to release intact its associated oligosaccharide moieties[11]. Reduction of the reducing terminal N-acetylglucosamine residues using NaB^3H_4 was then performed to label radioactively each carbohydrate chain. Each labelled oligosaccharide mixture was subjected to exhaustive neuraminidase digestion in order to analyse the distribution of neutral structures. The resulting 'asialo' oligosaccharide mixtures were fractionated by Bio-Gel P-4 (~400 mesh) gel filtration chromatography, a technique that separates neutral oligosaccharides on the basis of their effective hydrodynamic volumes[12] (Figs 2, 3).

A comparison of the individual profiles obtained reveals several interesting points. First, all P-4 chromatograms of asialo mixtures from control individuals (Fig. 3) were essentially identical to one another (data not shown) and similar to those reported previously[13]. As any individual IgG molecule contains, on average, 2.8 N-linked oligosaccharides[14], many different IgG subpopulations must exist, each unique with respect to the sequence of its associated oligosaccharides. Further, the overall relative molar contribution of each of the 30 or so structures in the analysis of polyclonal IgG is remarkably constant. The occurrence of this large number of different oligosaccharides is not the result of performing the analysis on polyclonal IgG, because this same 'set' of structures is found on human myeloma proteins[13] and an analogous situation exists for mouse monoclonal antibodies[14]. Second, the P-4 chromatograms of asialo mixtures from patients with rheumatoid arthritis are also essentially the same between patients (data not shown), but differ from control profiles (Fig. 3). Third, the asialo oligosaccharide