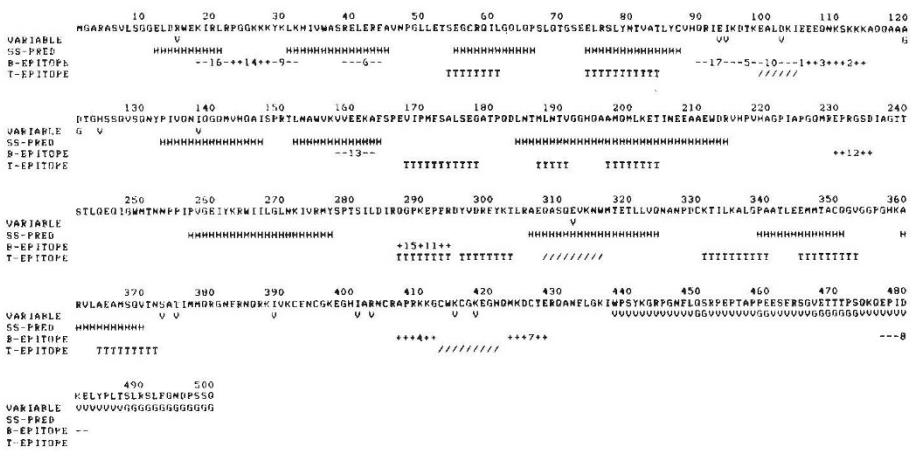

First line, numbered amino-acid (single letter code) sequence of a LAV strain[7] of HIV-1. Second line (VARIABLE), sequence variations among four HIV-1 strains, where V is an amino-acid variation and G is a gap (sequences from Protein Information Resource[8] file numbers FOVWLV, FOVWH3, FOVWA2 and FOVWVL, corresponding to refs 7, 9, 10 and 11, respectively). Third line (SS-PRED), predicted secondary structure: H, α-helix (no β-strands predicted). Fourth line (B-EPITOPE), predicted B-cell epitopes, where numbers denote the rank order of Hopp and Woods hydrophilic peaks[4], + denotes a peak conserved in sequence and not in a secondary structure, and − denotes other peaks. Fifth line (T-EPITOPE), predicted T-cell epitope[5], where T denotes a conserved sequence and / a non-conserved sequence.

have performed a prediction of the B and T epitopes of the gag proteins of one HIV-1 strain[7]. The results of standard algorithms to predict B epitopes[4] and T epitopes[5] are refined by considering inter-strain sequence variability and secondary structure prediction to yield suggestions for experimental verification. Applied to the env protein, this approach has identified seven B epitopes (to be reported elsewhere) of which three correspond to sections 1,6 and 7 of Robson *et al.*[2].

Three proteins are derived from the *gag* gene product. The first 132 amino-acid residues form p17 which is predicted to belong to the α/α class of protein with a carboxy-terminal tail of 40 residues that does not have any regular secondary structure and may well be flexible. This section would be an ideal candidate for stimulating a B-cell antigenic response. The main capsid protein, p24, is formed from residues 133 to about 373. This again is predicted to be an α/α class of protein, with the region between 220 and 260 probably being a long exposed loop that will stimulate a B-cell response. By contrast, the carboxy terminus of gag consists of two structurally quite different regions. The region from residues 375 to 438 consists of the twofold repeat (392–405 with 413–426) typical of a retroviral nucleic-acid binding protein[7]. But from residue 439 to the terminus there is a proline-rich section having a high degree of sequence variation between the four isolates and predicted not to adopt an α- or β-structure.

This analysis predicts likely B and T epitopes on gag proteins which would be suitable candidates for vaccines. We are synthesizing these for *in vitro* testing. A peptide of particular interest comprising residues 288–304 (GPKEPFRDYVDRFYKTL) which is the only 15–30 residue section containing both B and T epitopes in a sequence conserved region.

ANTHONY R.M. COATES
Department of Medical Microbiology,
JOHN COOKSON
Hill Computer Centre,
London Hospital Medical College,
Turner Street, London, E1 2AD, UK
GEOFFREY J. BARTON
MARKETA J. ZVELEBIL
MICHAEL J E STERNBERG
Department of Crystallography,
Birkbeck College, Malet Street,
London, WC1E 7HX, UK

1. Newmark, P. *Nature* **324**, 304 (1986).
2. Robson, B., Fishleigh, R.V. & Morrison, C.A. *Nature* **325**, 395 (1986).
3. Schupbach, J. *et al. New Engl. J. Med.* **312**, 265-270 (1985).
4. Hopp, T.P. & Woods, K.R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824-3828 (1981).
5. DeLisi, C. & Berzotsky, J. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7048–7052 (1985).
6. Weber, J.N. *et al. Lancet* **i**, 119-122 (1987).
7. Wain-Hobson, S. *et al. Cell* **40**, 9-17 (1985).
8. George, D.G. *et al. Nucleic Acids Res.* **14**, 11-17 (1986).
9. Ratner, L. *et al. Nature* **313**, 277-284 (1985).
10. Sanchez-Pescador, R. *et al. Science* **227**, 484-492 (1985).
11. Muesing, M.A. *et al. Nature* **313**, 450-458 (1985).

Knob-associated histidine-rich protein gene

SIR—Pologe and Ravetch have described the cloning of the complementary DNA for knob-associated histidine-rich protein (KAHRP)[1], also knownf as the knob protein of *Plasmodium falciparum*. We have also isolated cDNA for this protein from the same human malaria parasite (FCR-3 Gambia)[2] and agree with the location of *Bam* HI and *Xmn*-I restriction sites. But we disagree that there is no difference between the relative location of these sites in cDNA and genomic DNA (in preparation). The *Bam*HI and *Xmn*-I sites in the genomic DNA are much further apart because of the presence of a 444-base-pair intron between them. This intron separates exon 1, which codes for a signal peptide, from the main coding region, exon 2. We did not find any *Eco*RI site in the gene, though we sequenced the entire intron, exon 1 and about 375 nucleotides of 5′ untranslated region.

The laboratory-derived knob⁻ variants which lose their virulence[3] do not synthesize the knob protein[4] because they do not transcribe the gene[1,2]. Pologe and Ravetch concluded that this lack of messenger RNA and protein in knob⁻ variants is associated with deletions in the 3′ coding sequence of the knob protein gene. But we have found that not all knob⁻ variants have such deletions. Southern blot analysis of the knob protein gene showed no difference in the sizes of restriction fragments from knob⁺ (wild-type) and knob⁻ (FCR-3) using cDNA or genomic clones of the knob protein gene as probes, indicating that this gene is intact in these knob⁻ variants (ref. 2 and unpublished data). Thus the factors involved in the regulation of expression of the knob protein gene in knob⁻ variants remain unknown.

YAGYA D. SHARMA
Malaria Research Centre,
22-Sham Nath Marg, Delhi-110054, India

1. Pologe, L. G. & Ravetch, J. V. *Nature* **322**, 474–477 (1986).
2. Kilekian, A., Sharma, Y. D., Karoui, M. & Naslund, L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7938–7941 (1986).
3. Langreth, S. G. & Peterson, E. *Infec. Immun.* **47**, 760–766 (1985).
4. Kilejian, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4650–4653 (1979).

POLOGE AND RAVETCH REPLY—The map in our Fig. 2[1] to which Sharma refers does not claim that the complementary DNA and genomic organization of KAHRP gene are identical. Rather, it is a composite of cDNA data (Fig. 1) and genomic mapping data as described in the text. We recently described an intron in the region noted by Sharma from restriction mapping and sequence analysis of cDNA and genomic clones of the KAHRP gene[2]. In addition, an *Eco*RI site is found 500 bases 5′ of the intron–exon border in the 5′ untranslated sequence[2], placing it 15 bases 5′ of the region sequenced by Sharma.

Our conclusions on the mechanism of loss of KAHRP expression in knobless mutants were based on our determination of the structure and chromosome location of the KAHRP gene in three independently cloned, knobless parasites from Gambian and Vietnamese isolates. Analysis of other clonal isolates would be of interest as not all knobless parasites may arise by the same mechanism. But Sharma's conclusion, based on a non-clonal (FcR-3) isolate, is subject to uncertainty.

LAURA G. POLOGE
JEFFREY V RAVETCH
Sloan-Kettering Institute,
New York, New York 10021, USA

1. Pologe, L.G. & Ravetch, J.V. *Nature* **322**, 474-477 (1986).
2. Pologe, L.G., Pavlovec, A., Shio, H. & Ravetch, J.V. *Proc. natn. Acad. Sci. U.S.A.* (in the press).