PART II: HELPER T-CELL EPITOPES

SUMMARY

Part II includes tables and maps of HIV-specific helper T-cell (Th) epitopes arranged sequentially according to the location of the proteins in the HIV-1 genome. We attempted to make this section as comprehensive as possible, requiring that the epitope be contained within a region of approximately 30 amino acids maximum, but not that the precise boundaries be defined. The HLA specificity is usually not determined for Th epitopes. For more recent updates, epitope sequence alignments, and useful searching capabilities, please see our web site: http://hiv-web.lanl.gov/immunology. The same epitope can have multiple entries, as each entry represents a single publication.

A. TABLES:

Each Th epitope has a six-part basic entry:

- **HXB2 Location:** The viral strain HXB2 is used as a reference strain throughout this publication. The position of the defined epitope location on the sequence of the HXB2 protein is indicated. Obviously HXB2 may not be identical to a given defined reactive sequence, so we simply indicate the location of the aligned positions. The HXB2 numbering is used in to the protein maps of this database.
- **Author Location:** The amino acid positions of the epitope boundaries and the reference sequence are listed as given in the primary publication. Frequently, these positions as published are imprecise, and do not truly correspond to the numbering of the sequence, but they provide a reasonable guide to the peptide's approximate location in the protein. Also, in many cases the reference sequence identification was not provided, and in such cases it is not possible to use these numbers to specify precise locations. If you are interested in finding the precise positions of epitopes you are studying, please try using the interactive position locator at our web site: http://hiv-web.lanl.gov/NUM-HXB2/HXB2.MAIN.html.
- **Epitope Sequence:** The amino acid sequence of the epitope of interest as defined in the reference, based on the reference strain used in the study defining the epitope. On rare occasions, when only the epitope location and not the actual epitope was specified in the original publication, and the sequences were numbered inaccurately by the primary authors, we may have misrepresented the epitope's amino acid sequence. Therefore, epitopes that were not explicitly written out in the text in the primary publication, those that we determined by looking up the reference strain and the numbered location, are followed by a question mark in the table.
- **Immunogen:** The antigenic stimulus of the Th response to the defined epitope.
- **Species(HLA):** The species responding and HLA specificity of the epitope, when known.
- **Reference:** The primary reference.

Following each entry for a given Th epitope is a brief comment explaining the context of the study that defined or studied the epitope. If the same epitope was studied in several labs, each study is cited in its own entry.

B. HIV PROTEIN EPITOPE MAPS:

All human and primate Th epitopes defined to within 21 amino acids or less are indicated on the HIV protein epitope maps. HLA restriction information is included when known.

The location and HLA restriction elements (when known) of Th epitopes are indicated on protein sequences of the HXB2. These maps are meant to provide the relative location of epitopes on a given protein, but the HXB2 sequence may not actually carry the epitope of interest, as it may vary relative to the sequence for which the epitope was defined.

ALIGNMENTS:

Because of space limitations, alignments that correspond to the epitopes are only available from the web site, not in the hard copy of the compendium. All epitopes are aligned to the HXB2 sequence, with the sequence used to define the epitope indicated directly above it. In consensus sequences an upper case letter indicates the amino acid was present in all sequences, a lower case letter indicates the amino acid was present in most sequences in a given position, and a question mark indicates two or more amino acids were represented with equal frequency. The master alignment files from which the epitope alignments were created are available from our Web site at (http://hiv-web.lanl.gov/ALIGN_CURRENT/ALIGN-INDEX.html), and we restricted ourselves to full gene region sequences for these alignments, excluding short frag-

ments of sequences. The subtype designation and the country of isolation are indicated along with the common name of the sequence. The alignments were modified in some cases to optimize the alignment relative to the defined epitope and minimize insertions and deletions. A dash indicates identity to the consensus sequence, and a period indicates an insertion made to maintain the alignment. Stop codons are indicated with a $, and frameshifts by a #; they are inserted to maintain the alignments.

C. REFERENCES AND NOTES