View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Immunity, Vol. 10, 681–690, June, 1999, Copyright ©1999 by Cell Press

Presentation of Out-of-Frame Peptide/MHC Class I Complexes by a Novel Translation Initiation Mechanism

Subramaniam Malarkannan,* Tiffany Horng,* Patty P. Shih,* Susan Schwab,† and Nilabh Shastri*‡
*Division of Immunology
Department of Molecular and Cell Biology
†Energy and Resources Group
University of California
Berkeley, California 94720

Summary

Immune surveillance by CD8 T cells requires that peptides derived from intracellular proteins be presented by MHC class I molecules on the target cell surface. Interestingly, MHC molecules can also present peptides encoded in alternate translational reading frames, some even without conventional AUG initiation codons. Using T cells to measure expression of MHC bound peptides, we identified the non-AUG translation initiation codons and established that their activity was at the level of translational rather than DNA replication or transcription mechanisms. This translation mechanism decoded the CUG initiation codon not as the canonical methionine but as the leucine residue, and its activity was independent of upstream translation initiation events. Naturally processed peptide/MHC complexes can thus arise from "noncoding" mRNAs via a novel translation initiation mechanism.

Introduction

On the cell surface, the peptide/MHC class I complexes serve as ligands for the antigen receptors of $CD8^+$ T cells. The existence of a large and diverse set of peptide/MHC complexes allows the antigen receptors of $CD8^+$ T cells to detect novel intracellular proteins that arise due to viral or bacterial infection, tumorigenic transformation, or polymorphic histocompatibility loci. Many peptides presented by the MHC molecules are derived from endogenously synthesized proteins and are generated as by-products of protein turnover (Rock et al., 1994; Pamer and Cresswell, 1998). Yet, considerable evidence has accumulated that MHC molecules can also present peptides that were not expected to be even translated in the antigen-presenting cells (APC) (Gooding et al., 1988; Boon et al., 1989; Fetten et al., 1991; Hahn et al., 1991; Shastri and Gonzalez, 1993; Bullock and Eisenlohr, 1996; Elliott et al., 1996; Bullock et al., 1997). In addition, identification of unknown antigenic peptides presented by MHC molecules to CD8 T cells specific for tumors (Uenaka et al., 1994; Wang et al., 1996,1998; Shichijo et al., 1998; Dolstra et al., 1999), virus infected (Mayrand et al., 1998), or allogeneic cells (Malarkannan et al., 1995a) has shown that the target peptides were encoded in cryptic translational reading frames, some even without upstream AUG initiation codons. These examples indicate that cryptic peptide/MHC complexes can be generated by the APC and that these complexes are used to elicit cytotoxic T cell responses.

How APCs generate cryptic peptide/MHC complexes is obscure. A decade ago, Boon and Van Pel had formulated the Pepton hypothesis to explain how peptide/MHC complexes could arise from DNA fragments lacking obvious features that allow transcription and translation (Boon and Van Pel, 1989). The Pepton hypothesis suggested that a novel RNA polymerase and/or a translational mechanism generated short peptides for presentation by MHC molecules. However, our earlier studies failed to reveal a requirement for a novel transcription mechanism and instead showed that expression of the cryptic model peptide/MHC class I complexes correlated with translation initiation at non-AUG codons in alternate translational reading frames (Shastri and Gonzalez, 1993; Shastri et al., 1995). Independently, the Eisenlohr group showed that expression of out-of-frame peptide/MHC complexes in vaccinia-infected cells was consistent with aberrant translation initiation events at the AUG codons (Bullock and Eisenlohr, 1996; Bullock et al., 1997). As another possibility, Yewdell and colleagues have suggested that antigenic peptides could be derived from defective ribosomal products (DRiPs) due to premature termination events (Yewdell et al., 1996). Taken together, the key features of the Pepton and the DRiPs hypotheses as well as these experimental observations suggest that abnormal protein translation events can serve as a potential source of antigenic peptides, but the characteristics of this translational mechanism and its relationship to normal protein translation is not clear (Lindahl, 1991; Mayrand and Green, 1998).

Using an expression cloning strategy, we had earlier identified the SVVEFSSL (JAL8) peptide/K^b MHC class I complex as a ligand for the alloreactive T cell bm1BZ19.4 (Malarkannan et al., 1995a). Intriguingly, this JAL8 peptide was encoded in an alternate translational reading frame without an upstream AUG initiation codon. The ability to detect expression of this out-of-frame JAL8 peptide as well as other model peptides as MHC bound ligands in T cell activation assays provided an opportunity to systematically characterize the mechanisms that allow cryptic peptide/MHC expression in APC. We first defined the set of non-AUG initiation codons that allow expression of peptide/MHC complexes on the cell surface. Second, we established that the activity of the non-AUG codons is at the level of translation itself rather than DNA replication or transcription errors. Third, analysis of the naturally processed peptides in cell extracts showed that the CUG initiation codon was decoded not as the canonical methionine but as the leucine residue. This translational mechanism was capable of generating peptides from "noncoding" regions and was independent of other conventional translation initiation events. These findings strongly support the notion that a novel

‡To whom correspondence should be addressed (e-mail: nshastri@socrates.berkeley.edu).

```
STOP-44SC1         ...ctatagacccccgccagtgtggtggaattctccagcctgatggccagatgcccaagt...
                        *  T  P  A  S  V  V  E  F  S  S  L  M  A  R  C  P  S  .
                               ▽
STOP-START-44SC1   ...ctctagaccNNNgccagtgtggtggaattctccagcctgatggccagatgcccaagt...
                        *  T  ?  A  S  V  V  E  F  S  S  L  M  A  R  C  P  S  .
```

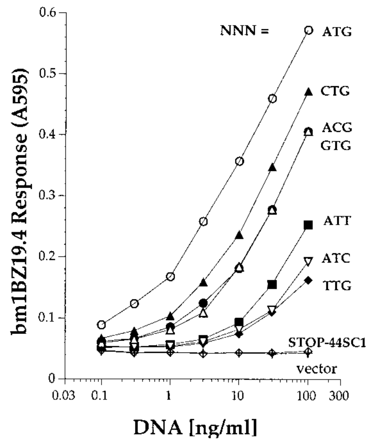

Figure 1. Non-AUG Codons Allow Generation of SVVEFSSL/K^b MHC Complex

The plasmid DNA construct "STOP-44SC1" contains the nucleotide sequence encoding the antigenic peptide (bold, underlined) with an upstream in-frame stop codon (tag). The "STOP-START-44SC1" constructs were generated using oligonucleotides with a degenerate "NNN" codon indicated with an arrowhead. Varying concentrations of purified plasmid DNAs were transfected into K^b-COS cells. Two days later, the lacZ inducible, bm1BZ19.4 T cells were added, and after overnight culture their response to the peptide/K^b MHC ligand was measured as the lacZ activity by the conversion of chlorophenol red β D-galactopyranoside. Data are shown as the average absorbance of replicate wells for each point with a representative construct. The nucleotide triplets shown next to each curve indicate the NNN sequence in the STOP-START-44SC1 sequence above.

translational mechanism can supply cryptic peptides for presentation by MHC molecules.

Results

Non-AUG Initiation Codons Allow Expression of Cryptic JAL8/K^b Complexes

To elucidate the mechanism by which cryptic peptide/MHC complexes are generated from alternate translational reading frames, we first identified the initiation codons that allowed their expression on the cell surface. The bm1BZ19.4 T cell was used to detect surface expression of the JAL8/K^b complex in K^b-COS cells transfected with DNA constructs. First, the STOP-44SC1 construct was derived from the earlier 44SC1 construct by inserting a termination codon nine nucleotides upstream of the SVVEFSSL (JAL8) codons (Malarkannan et al., 1995a) (Figure 1). This construct was completely inactive in stimulating the T cells. To identify potential initiation codons, a set of STOP-START-44SC1 plasmid constructs were prepared by inserting a degenerate "NNN" codon, representing all 64 possibilities between the termination and the JAL8 codons (Figure 1). Among a total of 252 random recombinant plasmids tested, 33 (13%) stimulated the bm1BZ19.4 T cells. Nucleotide sequences of 29 of these plasmids showed that the NNN codons were specified by ATG (N = 4), ACG (N = 7), CTG (N = 2), GTG (N = 2), ATT (N = 5), ATC (N = 4), and TTG (N = 5). As expected from its established role in translation initiation, the ATG as the NNN codon elicited the strongest T cell response. However, the activities of the CTG, ACG, and GTG codons were within 10- to 30-fold, while the activities of the ATT, ATC, and TTG were within 100-fold of the activity of the ATG codon. This analysis confirms and extends our previous findings with minimal precursors (Shastri et al., 1995), that generation of peptide/MHC complexes was not a random event but depended strictly upon the presence of either an ATG or an appropriate non-ATG codon. In further confirmation of this conclusion, presentation of the JAL8/K^b complex

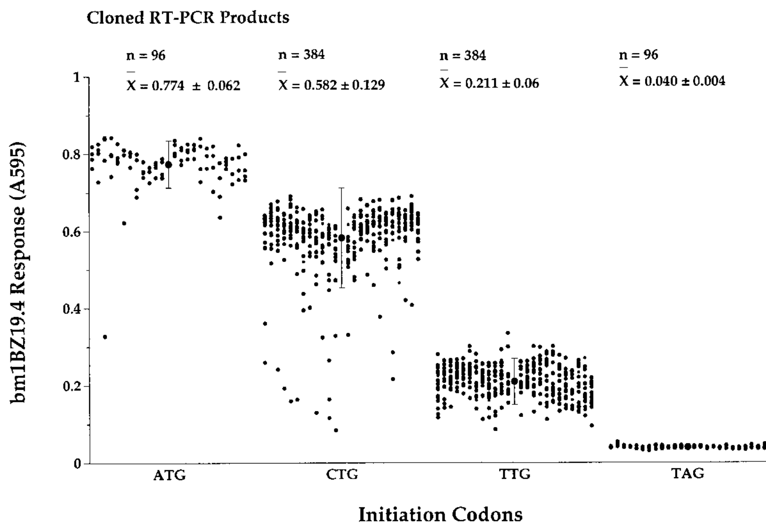

Figure 2. Analysis of Potential Non-AUG to AUG Mutations in the Transcripts

K^b-COS cells were transfected with plasmid DNA constructs encoding the JAL8 peptide and the indicated initiation codons (see Figure 1). Two days later, mRNA was prepared from the cells, converted into cDNA, and subjected to PCR using primers flanking the JAL8 coding sequence as described in the Experimental Procedures. Plasmid DNAs isolated from individual recombinant bacteria expressing the cloned PCR products were transfected into Kb-COS cells. Two days later, the cells were tested for expression of the JAL8/K^b complex using the bm1BZ19.4 T cells as in the legend to Figure 1. Each point represents the T cell response to an individual plasmid. The mean ± standard deviation of the T cell response to each set of plasmids is indicated in the figure above each codon set. n, number of recloned RT-PCR products tested.

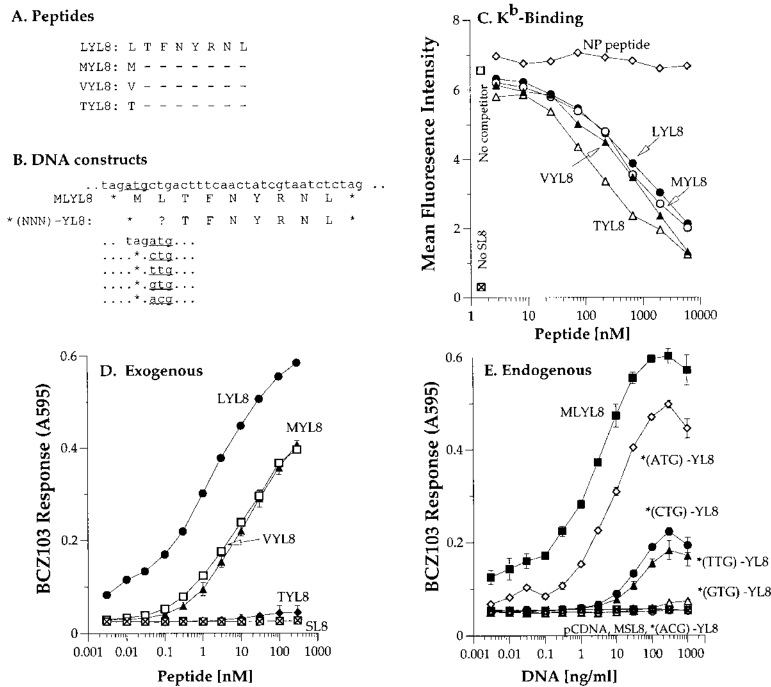

Figure 3. The H60 Model for Identifying the Amino Acid Residue Incorporated at the Initiation Codon of Cryptic Translation Products

(A) Amino acid sequence and abbreviations of the LYL8 peptide analogs in single letter code. (B) Sequence of DNA constructs encoding the MLYL8 or the LYL8 analogs with the indicated initiation codons (NNN) following a termination codon (*). (C) K^b MHC binding capacity of the LYL8 analogs is comparable. The peptide binding to K^b MHC on the surface of TAP-RMA/S cells was measured in a competitive flow cytometric assay using the SIINFEKL(SL8)/K^b-specific mAb as described earlier (Malarkannan et al., 1998). (D) The BCZ103 T cell recognizes the LYL8, MYL8, and VYL8 but not the TYL8 or the irrelevant SL8 peptides presented by K^b-L cells as APC. (E) The BCZ103 stimulating ligand is generated in cells transfected with the MLYL8 construct or the *(ATG)-, *(CTG), *(TTG)-YL8 constructs but not with the pcDNAI vector alone, the *(ACG)-, *(GTG)-YL8, or the irrelevant MSL8 constructs. Data points represent average ± standard deviation of triplicate measurements and are representative of at least three different experiments.

was abrogated when the upstream non-ATG initiation codons in the 44SC1 construct were mutated to those without any initiation activity (data not shown).

DNA Replication or Transcription Errors Do Not Account for Presentation of Out-of-Frame Translation Products

The peptide/MHC presentation activity was profoundly influenced by a subset of codons that have also been found as natural translation initiation codons in several cellular genes as well (Kozak, 1991; Boeck and Kolakofsky, 1994). However, given the exquisite sensitivity of T cells that can respond to even one peptide/MHC complex (Sykulev et al., 1996) and the high DNA replication and transcription levels normally obtained in transient transfections tested here and in tumors and virally infected cells used by others, we considered the alternate hypothesis that a single base mutation during DNA replication or transcription could have converted the non-AUG codons to the AUG codon. Because of the high intrinsic activity of the canonical AUG codon in translation initiation (Figure 1), the mutant transcript could then be translated in a conventional manner and could allow generation of the JAL8 peptide/K^b complex.

To detect the fraction of mutant transcripts with high translational activity, mRNA was isolated from cells transfected with the STOP-START-44SC1 constructs with either the ATG, CTG, TTG, or the TAG (STOP) codons at the NNN position shown in Figure 1. The coding sequences of these transcripts were then recovered using RT-PCR with a flanking forward and a reverse primer, and the DNA fragments were recloned into the pcDNAI vector using appropriate restriction enzymes. The plasmid DNA from individual recombinant bacteria was prepared in 96-well plates and transfected into K^b-COS cells. Two days later, the transfected cells were assayed for their ability to generate the JAL8/K^b complex using bm1BZ19.4 T cells (Figure 2). As expected, no T cell responses were observed to any of the 96 RT-PCR products of mRNA from cells transfected with the TAG (STOP) construct. In contrast, each of the 96 RT-PCR products from cells transfected with the ATG construct were active, and 95/96 had high activity. Interestingly, the activity of the 384 RT-PCR products tested from cells transfected with either the CTG or the TTG constructs, at the level of each plasmid and the overall average was remarkably similar to that of the original constructs (compare Figures 2 and 1). From this failure to detect 1%–10% of RT-PCR products of mutant transcripts with high "ATG-like" activity, we inferred that the activity of the non-ATG codons in allowing expression of cryptic peptide/MHC complexes was not due to low frequency mutations to the ATG codon but was rather at the level of the translation mechanism itself.

Model for Analysis of the Decoding of the Translation Initiation Codons Using the Peptide/MHC Expression Assay

To further characterize the translational mechanism for generating cryptic peptides, we determined how the non-AUG translation initiation codons were decoded. The initiation codons identified in the antigen presentation screen specify five different amino acids (ATG, Met; CTG and TTG, Leu; ACG, Thr; GTG, Val; ATT and ATC, Ile) according to the universal genetic code. But, as initiation codons they were all expected to be decoded as Met because the 40S ribosome is presumed to be loaded exclusively with the $tRNA_i^{met}$ before it binds to the 5′ end of the mRNA (Kozak, 1992; Rajbhandary and Chow, 1995). To distinguish whether the initiation mechanism for non-AUG codons was the same or distinct from this classical model, we wished to identify the actual amino acid incorporated in the antigenic precursors

translated using the non-AUG initiation codons. Toward this end, we developed a model system in which the initiation codons could be included within the coding sequence of the processed antigenic peptide itself. This was an absolute requirement because all N- and C-terminal residues flanking the antigenic peptide are precisely removed from the endogenous precursor by proteolysis, leaving only the MHC bound peptide available for analysis in cell extracts (Falk et al., 1990; Malarkannan et al., 1995b).

The first N-terminal Leu residue of the LTFNYRNL (LYL8) peptide, which is the naturally processed product of the *H60* minor histocompatibility locus (Malarkannan et al., 1998), could be substituted with either Met (MYL8), Val (VYL8), or Thr (TYL8) without compromising its K^b MHC binding capacity (Figures 3A and 3C). In an exogenous antigen presentation assay using K^b-L cells as APC, the LYL8/K^b-specific T cell hybrid, BCZ103, recognized the LYL8 peptide at picomolar concentrations and the MYL8 and VYL8 peptides at about 10-fold higher concentration (Figure 3D). Despite their excellent K^b binding capacities (Figure 3C), the TYL8 and the irrelevant ovalbumin derived SIINFEKL (SL8) peptides failed to stimulate the BCZ103 T cell even at 10^4-fold higher concentration. Most importantly, K^b cells transfected with the *(NNN)-YL8 DNA constructs encoding the "?-YL8" peptide sequence with ATG, CTG, and TTG, but not GTG and ACG as the initiation codons generated the BCZ103 stimulating peptide/K^b ligand (Figure 3E). The activity of the non-AUG initiation codon constructs was readily detectable and was within 10- to 100-fold of the maximal activity obtained with the MLYL8 construct with the canonical AUG initiation codon. Thus, as shown above for the JAL8 peptide, K^b MHC can also present the LYL8 peptide encoded in a cryptic context. The lack of BCZ103 T cell recognition of the Thr substituted analog, TYL8, and the poor activity of the GTG initiated construct precluded analysis of the ACG and the GUG initiated translation products in this model system. Nevertheless, these results showed that BCZ103 T cells could detect the Met or the Leu as the first residue of the LYL8 peptide and could therefore be used to determine how the CUG initiation codon was decoded.

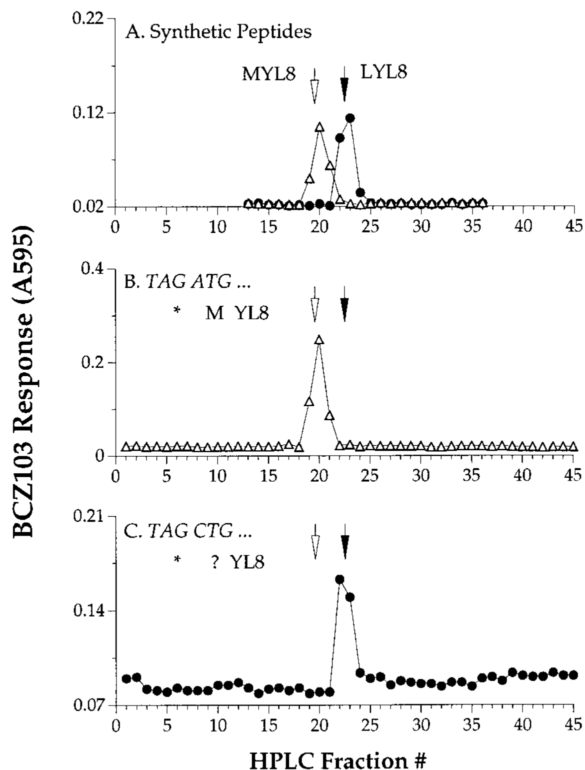

Figure 4. The AUG and CUG Initiation Codons Are Decoded as Met and Leu, Respectively

HPLC analysis of (A) synthetic MYL8 and LYL8 peptides or TFA extracts of COS cells expressing the K^b MHC and (B) the *(ATG)-YL8 or (C) the *(CTG)-YL8 constructs. Extracts were passed through a 10 kDa cut-off filter and the filtrate was run through a narrow bore reverse phase C18 HPLC column. The fractions were tested for BCZ103 T cell stimulating activity with K^b-L cells as APC. The vertical arrows mark the peak elution time for synthetic MYL8 and LYL8 peptides. The data are from one of four experiments with similar results.

The CUG Initiation Codon Is Decoded as the Leu Rather Than the Canonical Met Residue

The MHC bound naturally processed peptides generated in cells expressing the *(ATG)-YL8 and *(CTG)-YL8 constructs were analyzed by fractionating cell extracts by HPLC and testing the fractions for presence of BCZ103 T cell stimulating activity. Synthetic MYL8 and LYL8 peptides were run and assayed under identical conditions to compare their elution profiles with those detected in the extracts (Figure 4A). In extracts from the *(ATG)-YL8 transfected cells a single peak of activity was detected that coeluted with the MYL8 peptide, establishing that the AUG initiation codon was, as expected, decoded as the Met residue (Figure 4B). Surprisingly, the activity peak in extracts from the *(CTG)-YL8 transfected cells did not coelute with the MYL8 peptide but instead reproducibly coeluted with the synthetic LYL8 peptide (Figure 4C). Thus, the CUG initiation codon as well as the UUG initiation codon (data not shown) were decoded as the Leu rather than the canonical Met residue.

Cryptic Translation of the CUG Codon Is Independent of Upstream Translation Initiation Events

The observation that the CUG initiation codon was decoded as the Leu rather than the canonical Met residue was unexpected. Although the above analysis with the STOP-START 44SC1 constructs had clearly shown that peptide/MHC expression was absolutely dependent upon the presence of an appropriate initiation codon (Figure 1), we considered the formal possibility that for LYL8/K^b expression, the ribosomes could have initiated translation at some upstream site, but via leaky read-through of the TAG termination codon (Figure 3B), decoded the CUG codon as Leu during the elongation rather than the initiation step of translation. This hypothesis predicted that the amount of the cryptic LYL8 peptide would be proportional to the efficiency of translation initiation, not at its first CUG codon but at some other initiation codon upstream of the 5′ termination codon. This prediction was tested by placing the SIINFEKL (SL8) peptide upstream of the termination codon (*) preceding the

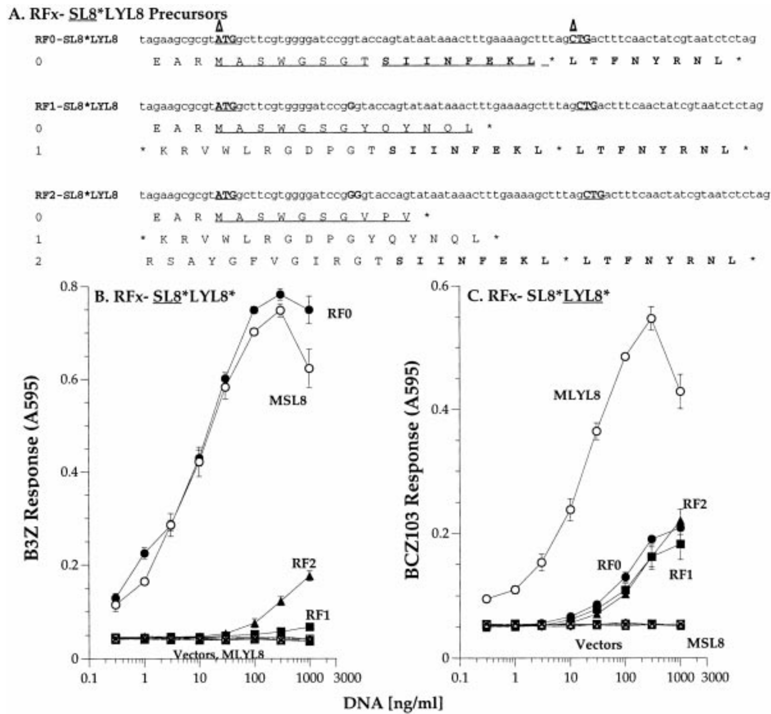

Figure 5. Expression of the Cryptic LYL8/K^b Complex Is Independent of the Efficiency of Upstream Translation Initiation

(A) Nucleotide sequences of the RFx-SL8*LYL8 DNA constructs. RF0, RF1, and RF2 constructs differ in the translational reading frame for the SL8*LYL8 cassette relative to the ATG initiation codon (bold, underlined). The ATG and the CTG initiation codons are indicated by arrowheads. Expression of (B) the B3Z stimulating SL8/K^b or (C) the BCZ103 stimulating LYL8/K^b complexes in K^b-COS cells transfected with the RFx-SL8*LYL8 constructs as described in the legend to Figure 1. Each of the three RFx vectors was used as a negative control, and the MSL8 and MLYL8 minigene constructs are reciprocal specificity control for B3Z and BCZ103 T cells. Data points represent average ± standard deviation of triplicate measurements and are representative of at least four different experiments.

CUG initiated, LYL8 peptide in the SL8*LYL8 cassette. The SL8/K^b complex is recognized by the B3Z T cell (Karttunen et al., 1992). The SL8*LYL8 cassette was then inserted into the RFx vectors in each of the three translational reading frames relative to the canonical ATG initiation codon to vary the translational efficiency (Figure 5A) (Shastri and Gonzalez, 1993). In cells transfected with the RFx-SL8*LYL8 constructs, the strongest SL8/K^b-specific B3Z T cell response was obtained with the RF0 construct encoding the SL8 peptide in-frame with the ATG codon and was superimposible with that to the minimal Met-SL8 (MSL8) construct (Figure 5B). In contrast, the B3Z T cell response was 2–3 orders of magnitude lower to the RF1 and RF2 constructs where the SL8 coding sequence was in the +1 or +2 translational reading frames relative to the ATG codon (Figures 5A and 5B). The B3Z T cell response was specific only for the SL8/K^b complex with no detectable activity to the vectors alone or to the MLYL8 construct, ruling out the possibility that differences in B3Z responses were due to spurious cross-reactions. The presentation efficiency of the SL8/K^b complex (RF0>>RF2>RF1) was thus profoundly influenced by the relative potency of AUG (RF0) versus non-AUG initiation (RF2, RF1) codons.

If the cryptic LYL8 peptide was generated by read-through of the termination codon by ribosomes that had initiated translation upstream, we would expect the same relative hierarchy (RF0>>RF2>RF1) for presentation of the LYL8 peptide as was observed for the upstream SL8 peptide. However, in the same RFx-SL8*LYL8 transfected cells, the peptide/K^b complex recognized by BCZ103 T cells was generated at comparable levels (RF0~RF2~RF1) (Figure 5C). The BCZ103 response was also specific for cells transfected only with the LYL8-encoding constructs, with no detectable response to any of the vectors alone or to the potent B3Z stimulating MSL8 construct. Further, the ~30-fold stronger BCZ103 response to the AUG-initiated Met-LYL8 construct showed that the same APCs could generate and BCZ103 cells could respond to far higher levels of the peptide/K^b complexes, ruling out saturation of the T cell response or the antigen presentation capacity of APC as a potential explanation for the comparable BCZ103 response to the cryptic LYL8 peptide. In other experiments (data not shown), we determined that comparable expression of the LYL8/K^b complex also occurred in cells transfected with the RFx-HY*LYL8 constructs in which the SL8 peptide was replaced by the HY peptide (WMHHNMDLI) presented by the D^b MHC. Expression of the cryptic LYL8/K^b complex was also obtained in constructs where instead of the cytomegalovirus (CMV) promoter, the MHC class I promoter controlled the transcription of M-SL8*LYL8 and M-HY*LYL8 coding sequences ruling out fortuitous requirements for a specific sequence preceding the termination codon or for the CMV promoter in the RFx constructs to generate the cryptic LYL8/K^b complex. We conclude that not only could a peptide/MHC complex arise from 3′ noncoding region but that its expression was independent of other upstream translation initiation events and was thus not due to a read-through mechanism.

To directly establish the differences in the relative abundance of these naturally processed peptides as well as to confirm the identity of the amino acid incorporated at the CUG initiation codon we analyzed TFA extracts of cells expressing the RFx-SL8*LYL8 constructs. The B3Z stimulating activity was readily detected in the extract of RF0-SL8*LYL8 transfectants and comigrated with the synthetic SL8 peptide (Figures 6A and 6B). In accord with the functional assay (Figure 5B), the concentration of recovered naturally processed SL8 peptide was diminished over 200-fold (3124 fM versus 15 fM) in

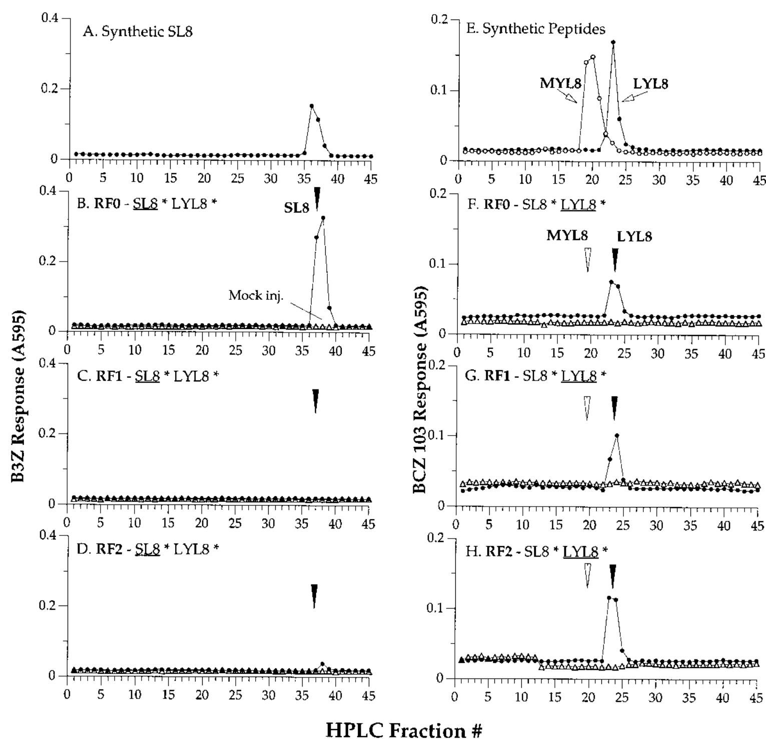

Figure 6. Relative Abundance of Naturally Processed Peptides in TFA Extracts of Cells Expressing K^b MHC and RFx-SL8*LYL8 Constructs

The extracts were fractionated by HPLC as in Figure 3. The fractions were dried, resuspended, divided into two equal parts, and tested for stimulating (A–D) SL8/K^b-specific B3Z and (E–H) LYL8/K^b-specific BCZ103 T cells using K^b L cells as APC. The synthetic peptides SL8 (A) and MYL8 or LYL8 (E) were run under identical conditions to determine their peak elution times indicated by vertical arrows. Each sample was preceded by mock injection of sample buffer alone, and fractions were collected and assayed in the same experiment to assess potential cross-contamination among the injected samples.

extracts of cells expressing the RF2-SL8*LYL8 construct (Figure 6D). No SL8 peptide was detected in RF1-SL8*LYL8 transfected cells (Figure 6C). In contrast, the concentration of BCZ103 stimulating LYL8 peptide in the same extracts varied by only 4-fold (190, 423, and 823 fM in RF0-, RF1- and RF2- respectively), directly confirming the results of the functional assay (Figures 6F–6H). Interestingly, these concentrations represent less than 1% of the LYL8 activity recovered from cells transfected with the Met-LYL8 construct shown in Figure 3E (93pM). Notably, the naturally processed peptide in each extract coeluted with the LYL8 rather than the MYL8 peptide, establishing that again Leu rather than the Met residue was incorporated at the CUG initiation codon (Figures 4E–4H). These results confirmed that presence of strong AUG (RF0) or weak non-AUG (RF2, RF1) upstream initiation site(s) had little quantitative and no qualitative influence on cryptic translation of the LYL8 peptide encoded in the 3′ noncoding region. They also established that the translational activity of the non-AUG initiation codons is less than 1% of that of the canonical AUG initiation codon.

Discussion

We have shown here that MHC class I molecules can present cryptic peptides arising from noncoding mRNA sequences. Peptide/MHC expression was absolutely dependent upon the presence of a subset of translation initiation codons and could not be accounted for by DNA replication or transcription errors. Surprisingly, we found that the CUG initiation codon was decoded as the Leu rather than the canonical Met residue and that its translational activity was independent of upstream translation initiation events. These findings indicate that a novel translation initiation mechanism can supply antigenic peptides for presentation by MHC molecules. This mechanism can significantly enlarge the potential pool of antigenic peptides that become available for display as peptide/MHC complexes on the cell surface (Figure 7) and has implications for protein translation as well as immune surveillance of endogenously synthesized proteins by the $CD8^+$ T cell repertoire.

Translational Mechanism for Generating Cryptic Peptides

How do ribosomes read apparently noncoding mRNA sequences and initiate translation at non-AUG codons? The scanning model of translation in eukaryotes suggests that the 43S ribosomal complex initiates translation at the first 5′ AUG codon. Departures from this conventional model have been observed mostly during viral infections and occasionally for some normal cellular genes (Engelberg-Kulka and Schoulaker-Schwarz, 1994). Translational frame shifts occur frequently in retroviruses (Jacks et al., 1988), and the polio virus allows 5′CAP-independent translation via internal ribosomal entry sites (IRES) (Sonenberg, 1990; Chen and Sarnow, 1995). These aberrations in the translational mechanism are essential for viral growth but depend upon the sequence-specific secondary structure of the mRNA. The CUG-initiated translation of the 3′ noncoding LYL8 peptide from the RFx-constructs (Figure 5) was independent of the upstream translational events and is therefore conceptually similar to an IRES but without the sequences that define natural IRES elements. Also, compared to relatively efficient IRES mediated translation initiation, the efficiency of expression of the cryptic peptide/MHC complexes was $<1\%$ of that when the

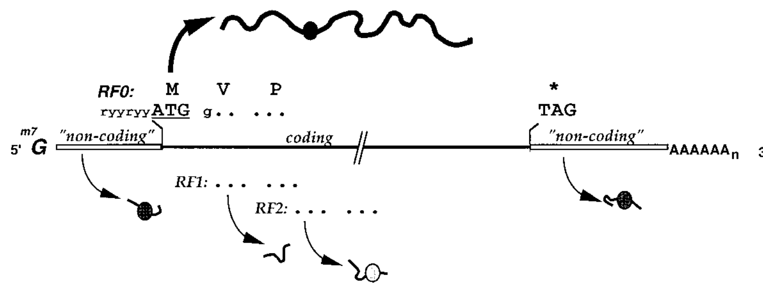

Figure 7. Schematic Representation of the Origin of Antigenic Peptides Displayed by MHC I Molecules

MHC bound processed peptides (filled circles) can be encoded in the open reading frame (RF0, e.g., MVP....) whose translation is initiated at the first 5′ AUG that is in the appropriate sequence context. MHC can also present peptides that are encoded in alternate translational reading frames (RF1 or RF2) as well as in the 5′ or 3′ untranslated regions.

same peptides were placed in a conventional AUG initiation context (Figures 1, 3E, and 5B). Our results are therefore consistent with the notion that a small fraction of ribosomes deviate from the "first-AUG" rule and can initiate translation wherever appropriate initiation codons are found. Given that these unusual translation products are present only in small quantities, they are unlikely to cause disruptions in normal cell physiology but can be immunologically relevant for detection by T cells. A potential mechanism for generating cryptic translation products may be related to the recent discovery that ribosomal recognition of the conventional 5′ AUG initiation codon is not simply a default consequence of 5′→3′ scanning of mRNA by the 43S preinitiation ribosomal complex but is regulated by the eIF1 and eIF1A factors (Pestova et al., 1998). The regulated recognition of the conventional 5′ AUG initiation codon allows for the possibility that recognition of cryptic initiation codons may also be regulated, as was actually observed when translation was initiated at the UUG codon in yeast with mutations in the eIF1 homolog (Yoon and Donahue, 1992). The potential role of these initiation factors in the recognition of cryptic initiation codons in mammalian cells however remains to be tested.

The decoding of the CUG initiation codon as Leu rather than the canonical Met residue is unprecedented. While the translation of synthetic RNAs encoding a variety of amino acids by bacterial ribosomes in vitro was the key to deciphering the genetic code (Nirenberg and Leder, 1964), translation in vivo is by far more stringent and initiates almost exclusively at the first 5′ AUG codon in a suitable context (Cigan et al., 1988; Kozak, 1991). Nevertheless, for a growing list of prokaryotic and eukaryotic genes emerging from genome sequencing projects, translation initiates at non-AUG codons (http://www.ncbi.nlm.nih.gov/dbGSS/), but the context rules and the amino acids incorporated at these codons in vivo are not clear (Boeck and Kolakofsky, 1994; Reynolds et al., 1995; Kozak, 1997). Synthesis of dihydrofolate reductase initiated in vitro with ACG, CUG, GUG, UUG, AUA, AUC, and AUU codons was shown to incorporate predominantly the Met residue, but other residues were not tested (Peabody, 1989). Using mutant initiator t-RNAs, it has been recently reported that ribosomes can not only utilize non-AUG initiation codons but also incorporate nonmethionine amino acids, thus establishing that the mammalian translational machinery is not restricted to the exclusive use of only the methionine residue (Drabkin and RajBhandary, 1998). But the extent to which the unmanipulated mammalian translational machinery incorporates non-Met residues in the non-AUG-initiated ORFs is not clear. Note that conventional assays are unlikely to allow detection of cryptic translation products such as those described here because the peptides presented by MHC molecules are small, relatively rare, and extraordinarily labile in the absence of appropriate MHC molecules (Falk et al., 1990; Malarkannan et al., 1995b). Because the Thr or Val residues could not be tested in this model system (Figures 3D and 3E), it is not yet clear whether translation initiated at the other relatively strong ACG and GUG initiation codons incorporates the Thr and Val residues as well. The use of a non-Met residue by this translational mechanism could possibly be due to novel t-RNAs, synthetases, or other factors that regulate translational reinitiation in prokaryotic cells (Janosi et al., 1998). Regardless of the mechanism that incorporates the natural residue rather than the canonical methionine residue, this is a teleological necessity when these peptides are used to elicit CD8 T cell responses. This ensures that the same peptide/MHC ligand will be available for T cell recognition regardless of whether it was derived from a conventional or cryptic translational reading frame.

Implications for Immune Surveillance

What fraction of the thousands of peptides presented by the MHC molecules at a level of <10 copies on the cell surface are cryptic translation products is not known (Hunt et al., 1992). But T cell recognition of many cryptic peptide/MHC complexes in experimental, tumor, and virus models raises the question of why APCs utilize noncoding information in the transcripts in addition to the conventional source derived from open reading frames. Note that from the perspective of the $CD8^+$ T cell surveillance of novel intracellular events such as a viral infection or tumorigenic transformation it is irrelevant whether the peptides presented by the MHC molecules originate from the predominant open or cryptic translational reading frames. The ability to translate noncoding sequences by the APCs can significantly increase peptide supply. This could be an evolutionary solution to ease the bottlenecks that limit the diversity of the peptide/MHC display and as a consequence, increase the number of intracellular transcripts detectable by $CD8^+$ T cells. Indeed, rate-limiting steps in the antigen presentation pathway have been documented for proteolysis (Ossendorp et al., 1996), TAP transport (Neisig et al., 1995) and peptide/MHC binding (Falk et al., 1991). Furthermore, cryptic translation products because of their small size require only minimal processing to yield the precisely cleaved 8–11 mers and could therefore be excellent precursors for MHC binding peptides.

In conclusion, we have shown here that cryptic peptide/MHC complexes can arise from noncoding regions of mRNA by a novel translation initiation mechanism. The existence of this mechanism provides a potential explanation for the numerous observations that cryptic translation products are presented by MHC class I molecules. Whether this translational mechanism is available in unmanipulated normal cells and whether it is regulated in different cell types are intriguing possibilities that could allow cryptic translation products to play a role in the development and maintenance of the T cell repertoire as well.

Experimental Procedures

Cell Lines

Cell lines were maintained in RPMI 1640 medium (GIBCO-BRL) supplemented with 2 mM glutamine, 1 mM pyruvate, 50 μM β-mercaptoethanol, 100 U/ml penicillin, 100 μg/ml streptomycin, and 10% FBS (Hyclone). COS7, Lmtk-, K^b-L cells, RMA (H-2^b), and its TAP-derivative RMA/S cell lines have been described (Karttunen et al., 1992; Malarkannan et al., 1998). T cell hybridomas used to detect peptide/MHC expression were bm1BZ19.4 (anti-JAL8/K^b) (Malarkannan et al., 1995a), BCZ103 (anti-H60 (LYL8)/K^b) (Malarkannan et al., 1998), and B3Z (anti-OVA (SL8)/K^b) (Karttunen et al., 1992).

DNA Constructs and Peptides

All DNA constructs were prepared using complementary oligonucleotides with the sequences shown in Figures 1, 3, and 5 in either the 44SC1 (Malarkannan et al., 1995a), pcDNAI (Invitrogen), or the pEVRFx (RFx) vectors (Matthias et al., 1989). Constructs for identifying alternate initiation codons were prepared with the NNN degenerate (N = 25% A,C,G,T) codon (Shastri et al., 1995). Constructs were sequenced to identify the degenerate NNN codon and to confirm their integrity. The synthetic peptides, LTFNYRNL (LYL8), MTFNYRNL (MYL8), VTFNYRNL (VYL8), TTFNYRNL (TYL8), and SIINFEKL (SL8) were prepared using solid phase F-Moc chemistry on the ABI Model 421 synthesizer. The peptides were purified by HPLC and their sequence confirmed by mass spectrometry.

Transient Transfections and T Cell Activation Assays

DNA constructs were transiently transfected into 3×10^4 COS7 cells together with the K^b MHC class I cDNA (10 ng/ml) and B7-1 cDNA (5 ng/ml) (Malarkannan et al., 1995a, 1998). Two days later, 10×10^4 T cells were added per well and cocultured overnight. Peptide titrations were carried out using K^b-L cells as APC and the respective T hybrids. The lacZ expression in the T cells was measured in culture lysates by the conversion of the CPRG substrate to chlorophenol red at 595 and 655 nm as the reference wavelength (Sanderson and Shastri, 1994). Data show the mean absorbance of replicate cultures and are representative of at least three independent experiments.

RT-PCR Analysis

COS cells were transfected with constructs coding for the JAL8 peptide, preceded by conventional ATG, the CTG, TTG initiation codons or the TAG stop codon (Figure 1). After 48 hr, the mRNA from the transfected cells was purified using Qiagen Oligotex Direct mRNA purification kit. Purified mRNA was treated with DNAase to eliminate contaminating plasmid DNA confirmed by the absence of amp^r+tet^r transformants of MC1061/p3 *E. coli*. First and second strand cDNA synthesis was carried out with the 5′ T7 and 3′ SP6 flanking primers using the Super Script II reverse transcriptase as per the manufacturer's instructions. The cDNA was cut with BamHI and EcoRV, cloned into the pcDNAI vector, and transformed into competent MC1061/p3 *E. coli* strain. Individual bacterial colonies were selected on ampicillin and tetracycline plates and grown in 96-well plates. Plasmid DNA were prepared in the 96-well plates and used in transient transfection assays together with the K^b cDNA using recipient COS cells to elucidate the functional activity of the each plasmid in for stimulating bm1BZ19.4 T cell hybrid as described above.

HPLC Analysis of Naturally Processed Peptides in Cell Extracts

Total acid soluble peptide pool from transiently transfected cells was extracted and analyzed as described earlier (Rotzschke et al., 1990; Malarkannan et al., 1995b). In brief, 4×10^7 COS7 cells were transfected with 1 μg/ml of ATG-YL8, CTG-YL8, or the RFx-SL8*CTG-YL8 constructs using the DEAE-Dextran/chloroquin method in 100 mm petri dishes. After 48 hr, the cells were harvested, washed once with PBS, and extracted with 4 ml of 1% Trifluoroacetic acid (TFA) in water. Cellular debris and >10 kDa material was removed by centrifugation and passing the supernatant through a 10 kDa filter. The <10 kDa filtrate was analyzed by HPLC using a reverse phase C18 column (Vydac, 2.1 × 250 mm, 5 μm) in 0.1% TFA in water (solvent A) and 0.1% TFA in acetonitrile (solvent B). Three drop fractions were collected between 24% to 32% solvent B at a flow rate of 0.35 ml/min in 96-well flat-bottom plates. After drying the plates in Speed Vac Plus (Savant), each fraction was resuspended in 30 μl PBS + 15%DMSO and assayed for T cell stimulating activity using K^b-L cells as APC as described above. Synthetic peptides (100 fm) and mock injections were run and assayed under identical conditions in parallel to determine their elution profiles and to confirm the absence of sample carry over between experimental runs. The minimal peptide concentrations required to activate T cells were typically in the range of 100–1000 fM. Processed peptide amounts in cell extracts were estimated by comparison with synthetic peptide standard curves.

Acknowledgments

We thank Federico Gonzalez, Vu Nguyen, Minh Nguyen, Van Nguyen, and Hussein Hadeiba for their excellent technical support. This study was supported by National Institutes of Health grants (N. S.). S. M. was supported in part by the University of California Cancer Research Coordination Committee and T. H. by the Beckman undergraduate scholarship. T. H. and P. S. were also supported by the Undergraduate Biology Fellows Program of the Howard Hughes Medical Institute. S. S. is a predoctoral fellow of the National Science Foundation. We are grateful to U. Rajbhandary (MIT), A. Kaji (U. Pennsylvania), D. Peabody (Univ. New Mexico), R. Boeck (Univ. California), and other members of our laboratory for discussions.

Received March 19, 1999.

References

Boeck, R., and Kolakofsky, D. (1994). Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. EMBO J. *13*, 3608–3617.

Boon, T., and Van Pel, A. (1989). T cell-recognized antigenic peptides derived from the cellular genome are not protein degradation products but can be generated directly by transcription and translation of short subgenic regions. A hypothesis. Immunogenetics *29*, 75–79.

Boon, T., Van Pel, A., De Plaen, E., Chomez, P., Lurquin, C., Szikora, J.-P., Sibille, C., Mariame, B., Van Den Eynde, B., Lethe, B., and Brichard, V. (1989). Genes coding for T-cell-defined tum transplantation antigens: point mutations, antigenic peptides, and subgenic expression. Cold Spring Harbor Symp. Quant. Biol. *54*, 587–596.

Bullock, T.N.J., and Eisenlohr, L.C. (1996). Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. J. Exp. Med. *184*, 1319–1329.

Bullock, T.N.J., Patterson, A.E., Franlin, L.L., Notidis, E., and Eisenlohr, L.C. (1997). Initiation codon scanthrough versus termination codon readthrough demonstrates strong potential for major histocompatibility complex class I-restricted cryptic epitope expression. J. Exp. Med. *186*, 1051–1058.

Chen, C.Y., and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. Science *268*, 415–417.

Cigan, A.M., Feng, L., and Donahue, T.F. (1988). $tRNA_i^{Met}$ functions in directing the scanning ribosome to the start site of translation. Science *242*, 93–97.

Dolstra, H., Fredrix, H., Maas, F., Coulie, P.G., Brasseur, F., Mensink, E., Adema, G.J., de Witte, T.M., Figdor, C.G., Wiel-Van, V.D., and Kemenade, E. (1999). A human minor histocompatibility antigen specific for B cell acute lymphoblastic leukemia. J. Exp. Med. *189*, 301–308.

Drabkin, H.J., and RajBhandary, U.L. (1998). Initiation of protein synthesis in mammalian cells with codons other than AUG and amino acids other than methionine. Mol. Cell. Biol. *18*, 5140–5147.

Elliott, T., Bodmer, H., and Townsend, A. (1996). Recognition of out-of-frame major histocompatibility complex class I-restricted epitopes in vivo. Eur. J. Immunol. *26*, 1175–1179.

Engelberg-Kulka, H., and Schoulaker-Schwarz, R. (1994). Regulatory implications of translational frameshifting in cellular gene expression. Mol. Microbiol. *11*, 3–8.

Falk, K., Rötzschke, O., and Rammensee, H.-G. (1990). Cellular peptide composition governed by major histocompatibility complex class I molecules. Nature *348*, 248–251.

Falk, K., Rotzschke, O., Stevanovic, S., Jung, G., and Rammensee, H.-G. (1991). Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature *351*, 290–296.

Fetten, J.V., Roy, N., and Gilboa, E. (1991). A frameshift mutation at the NH2 terminus of the nucleoprotein gene does not affect generation of cytotoxic T lymphocyte epitopes. J. Immunol. *147*, 2697–2705.

Gooding, L.R., O'Connell, K.A., Geib, R.W., and Pipas, J.M. (1988). Cytotoxic T cell recognition of the SV40 tumor antigen: a note of caution. In Processing and Presentation of Antigens, B. Pernis, S. Silverstein, and H. Vogel, eds. (New York: Academic Press), pp. 87–95.

Hahn, Y.S., Braciale, V.L., and Braciale, T.J. (1991). Presentation of viral antigen to class I major histocompatibility complex-restricted cytotoxic T lymphocyte. Recognition of an immunodominant influenza hemagglutinin site by cytotoxic T lymphocyte is independent of the position of the site in the hemagglutinin translation product. J. Exp. Med. *174*, 733–736.

Hunt, D.F., Henderson, R.A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A.L., Appella, E., and Engelhard, V.H. (1992). Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science *255*, 1261–1263.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., and Varmus, H.E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature *331*, 280–283.

Janosi, L., Mottagui-Tabar, S., Isaksson, L.A., Sekine, Y., Ohtsubo, E., Zhang, S., Goon, S., Nelken, S., Shuda, M., and Kaji, A. (1998). Evidence for in vivo ribosome recycling, the fourth step in protein biosynthesis. EMBO J. *17*, 1141–1151.

Karttunen, J., Sanderson, S., and Shastri, N. (1992). Detection of rare antigen presenting cells by the lacZ T-cell activation assay suggests an expression cloning strategy for T-cell antigens. Proc. Natl. Acad. Sci. USA *89*, 6020–6024.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. *266*, 19867–19870.

Kozak, M. (1992). Regulation of translation in eukaryotic systems. Annu. Rev. Cell Biol. *8*, 197–225.

Kozak, M. (1997). Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. EMBO J. *16*, 2482–2492.

Lindahl, K.F. (1991). Do we need a pepton hypothesis? Immunogenetics *34*, 1–4.

Malarkannan, S., Afkarian, M., and Shastri, N. (1995a). A rare cryptic translation product is presented by K^b MHC class I molecule to alloreactive T-cells. J. Exp. Med. *182*, 1739–1750.

Malarkannan, S., Goth, S., Buchholz, D.R., and Shastri, N. (1995b). The role of MHC class I molecules in the generation of endogenous peptide/MHC complexes. J. Immunol. *154*, 585–598.

Malarkannan, S., Shih, P., Eden, P., Horng, T., Zuberi, A.R., Christianson, G., Roopenian, D., and Shastri, N. (1998). The molecular and functional characterization of a dominant minor H antigen, *H60*. J. Immunol. *161*, 3501–3509.

Matthias, P., Muller, M.M., Schreiber, E., Rusconi, S., and Schaffner, W. (1989). Eukaryotic expression vectors for the analysis of mutant proteins. Nucleic Acids Res. *17*, 6418.

Mayrand, S.-M., and Green, W.R. (1998). Non-traditionally derived CTL epitopes: exceptions that prove the rules? Immunol. Today *19*, 551–555.

Mayrand, S.M., Schwarz, D.A., and Green, W.R. (1998). An alternative translational reading frame encodes an immunodominant retroviral CTL determinant expressed by an immunodeficiency-causing retrovirus. J. Immunol. *160*, 39–50.

Neisig, A., Roelse, J., Sijts, A.J.A.M., Ossendorp, F., Feltkamp, M.C.W., Kast, W.M., Melief, C.J.M., and Neefjes, J.J. (1995). Major differences in transporter associated with antigen presentation (TAP)-dependent translocation of MHC class I-presentable peptides and the effect of flanking sequences. J. Immunol. *154*, 1273–1279.

Nirenberg, M., and Leder, P. (1964). RNA code words and protein synthesis. Science *145*, 1399–1407.

Ossendorp, F., Eggers, M., Neisig, A., Ruppert, T., Groettrup, M., Sijts, A., Mengede, E., Kloetzel, P.M., Neefjes, J., Koszinowski, U., and Melief, C. (1996). A single residue exchange within a viral CTL epitope alters proteasome-mediated degradation resulting in lack of antigen presentation. Immunity *5*, 115–124.

Peabody, D.S. (1989). Translation initiation at non-AUG triplets in mammalian cells. J. Biol. Chem. *264*, 5031–5035.

Pestova, T.V., Borukhov, S.I., and Hellen, C.U.T. (1998). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. Nature *394*, 854–859.

Rajbhandary, U.L., and Chow, C.M. (1995). Initiator tRNAs and initiation of protein synthesis. In tRNA: Structure, biosynthesis, and function, D. Soll and U.L. Rajbhandary, eds. (Washington, D.C.: American Society of Microbiology), pp. 511–528.

Reynolds, J.E., Kaminski, A., Kettinen, H.J., Grace, K., Clarke, B.E., Carroll, A.R., Rowlands, D.J., and Jackson, R.J. (1995). Unique features of internal initiation of hepatitis C virus RNA translation. EMBO J. *14*, 6010–6020.

Rotzschke, O., Falk, K., Wallny, H.-J., Faath, S., and Rammensee, H.-G. (1990). Characterization of naturally occurring minor histocompatibility peptides including H-4 and H-Y. Science *249*, 283–287.

Sanderson, S., and Shastri, N. (1994). LacZ inducible peptide/MHC specific T-hybrids. Int. Immunol. *6*, 369–376.

Shastri, N., and Gonzalez, F. (1993). Endogenous generation and presentation of the OVA peptide/K^b complex to T-cells. J. Immunol. *150*, 2724–2736.

Shastri, N., Nguyen, V., and Gonzalez, F. (1995). Major histocompatibility class I molecules can present cryptic translation products to T-cells. J. Biol. Chem. *270*, 1088–1091.

Shichijo, S., Nakao, M., Imai, Y., Takasu, H., Kawamoto, M., Niiya, F., Yang, D., Toh, Y., Yamma, H., and Itoh, K. (1998). A gene encoding antigenic peptides of human squamous cell carcinoma recognized by cytotoxic T lymphocytes. J. Exp. Med. *187*, 277–288.

Sonenberg, N. (1990). Poliovirus translation. Curr. Top. Microbiol. Immunol. *161*, 23–47.

Sykulev, Y., Joo, M., Vturina, I., Tsomides, T.J., and Eisen, H.N. (1996). Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response. Immunity *4*, 565–571.

Uenaka, A., Ono, T., Akisawa, T., Wada, H., Yasuda, T., and Nakayama, E. (1994). Identification of a unique antigen peptide pRL1 on BALB-c RL 1 leukemia recognized by cytotoxic T lymphocytes and its relation to the Akt oncogene. J. Exp. Med. *180*, 1599–1607.

Wang, R.-F., Parkhurst, M.R., Kawakami, Y., Robbins, P.F., and Rosenberg, S.A. (1996). Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. J. Exp. Med. *183*, 1131–1140.

Wang, R.-F., Johnston, S.L., Zeng, G., Topalian, S.L., Schwartzentruber, D.J., and Rosenberg, S. (1998). A breast and melanoma-shared tumor antigen: T cell responses to antigenic peptides translated from different open reading frames. J. Immunol. *161*, 3596–3606.

Yewdell, J.W., Anton, L.C., and Bennink, J.R. (1996). Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? J. Immunol. *157*, 1823–1826.

Yoon, H.J., and Donahue, T.F. (1992). The suil suppressor locus in Saccharomyces cerevisiae encodes a translation factor that functions during tRNA(iMet) recognition of the start codon. Mol. Cell. Biol. *12*, 248–260.