מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Anti-tumour immunity induces aberrant peptide presentation in melanoma

Anti-tumor immunity induces aberrant peptide presentation in melanoma

Osnat Bartok[1,14], Abhijeet Pataskar[2,14], Remco Nagel[2,14], Maarja Laos[3,#], Eden Goldfarb[1,#], Deborah Hayoun[1], Ronen Levy[1], Pierre-Rene Körner[2], Inger Z M Kreuger[2], Julien Champagne[2], Esther Zaal[4], Onno B Bleijerveld[5], Xinyao Huang[6], Juliana Kenski[6], Jennifer Wargo[7], Alexander Brandis[8], Yishai Levin[9], Orel Mizrahi[10], Michal Alon[1], Sacha Lebon[11], Weiwen Yang[3], Morten M Nielsen[3], Noam Stern-Ginossar[10], Maarten Altelaar[4,5], Celia Berkers[4], Tamar Geiger[12], Daniel S Peeper[6], Johanna Olweus[3], Yardena Samuels[1]*, Reuven Agami[2,13,*]

[1]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel

[2]Division of Oncogenomics, Oncode institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX, Amsterdam, the Netherlands

[3]Department of Cancer Immunology, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, and Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[4]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research, and Department of Biochemistry and Cell Biology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

[5]Proteomics facility The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX, Amsterdam, the Netherlands

[6]Division of Molecular Oncology & Immunology, Oncode institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX, Amsterdam, the Netherlands

[7]Departments of Surgical Oncology and Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas

[8]Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel

[9]The Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science, Rehovot, Israel

[10]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel.

[11]Department of Biological Regulation, Weizmann Institute of Science, Rehovot, Israel.

[12] Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Israel.

[13] Erasmus MC, Rotterdam University, Rotterdam, the Netherlands

[14] These authors contributed equally

#These authors contributed equally

* Corresponding author

Abstract

Extensive tumor inflammation, reflected by high levels of infiltrating T-cells and Interferon gamma (IFNγ) signaling, improves checkpoint immunotherapy response[1,2]. Many tumors, however, escape by activating cellular pathways that induce immunosuppression. One such mechanism is the production of tryptophan metabolites along the kynurenine pathway by the IFNγ-induced enzyme IDO1[3-5]. Clinical trials using IDO1-inhibition in combination with PD1-pathway-blockade, however, failed to improve melanoma treatment[6,7], pointing to an incomplete understanding of the role of IDO1 and the consequent tryptophan degradation on mRNA translation and cancer progression. Here, we investigated the effects of prolonged IFNγ treatment on mRNA translation in melanoma cells by ribosome profiling. Surprisingly, we observed an accumulation of ribosomes downstream of tryptophan codons (W-Bumps) along with their expected stalling at the tryptophan codon. This indicated ribosomal-bypass of tryptophan codons in the absence of tryptophan. Detailed examination of W-Bumps pinpointed ribosomal frameshifting events. Indeed, reporter assays combined with proteomic and immunopeptidomic analyses demonstrated the induction of ribosomal frameshifting, and the generation and presentation of aberrant *trans-frame* peptides at the cell surface after IFNγ treatment. Priming of naïve T cells from healthy donors with aberrant peptides induced peptide-specific T cells. Altogether, our results suggest that IFNγ-induced IDO1-mediated tryptophan depletion plays a role in the immune recognition of melanoma cells by contributing to diversification of the peptidome landscape.

IFNγ-induced IDO1-mediated tryptophan deprivation stimulates the uncharged-tRNA-sensing machinery whose main components are EIF2AK4 (GCN2), eIF2α, ATF4[8-10]. When active, the EIF2AK4-ATF4 cascade suppresses protein synthesis initiation (Fig. 1a). Conversely, cancers compensate for IFNγ-induced tryptophan deprivation by upregulating both the expression of several amino acid transporters and WARS, the tryptophanyl-tRNA synthetase. This improves survival and quick recovery once tryptophan is replenished[3,11]. However, the long-term impact of sustained IFNγ-mediated tryptophan depletion on melanoma cells remains largely unknown.

Effects of IFNγ on translating ribosomes

To address this issue, we first confirmed IDO1 induction, tryptophan depletion, and kynurenine accumulation in IFNγ-treated 12T melanoma cells (Extended Data. Fig. 1a,b). Then, we performed ribosome profiling and analyzed the ribosome-protected fragments (RPFs) by diricore (DIfferential RIbosome COdon REading) to detect differential ribosome occupancy patterns at the codon level[12] (Extended Data Fig. 1c). Metagene examination of RPF distribution revealed IFNγ-induced accumulation at the translation start site (Fig. 1b). This corresponds to the expected reduction in global protein synthesis, measured by OP-Puro (O-propargyl-puromycin) incorporation assays (Fig. 1c), in line with global reduction in translation initiation. Subsequence (codon occupancy bias) and 5'-RPF density[12] analyses uncovered reduced ribosome occupancy at RPF position 12 (corresponding to ribosomal P-site) at the initiator methionine codon (ATG[start]) (Fig. 1d). Similar results were obtained with the melanoma cell lines MD55A3 and 108T (Extended Data Fig. 1a,b,d,e).

Further analysis of RPFs containing tryptophan codons at position 15 (ribosomal A-site) indicated stalling of ribosomes on the tryptophan codon following IFNγ treatment (Fig. 1e, Extended Data Fig. 1f). 5'-RPF measurements confirmed increase in the tryptophan codon density at the A-site of ribosomes, whereas the codons of other amino acids showed no such pattern (Fig. 1d, Extended Data Fig. 1e). These observations are consistent with the expected suppressed translation initiation and stalling of ribosomes at the tryptophan codon due to tryptophan shortage[12].

Surprisingly, we observed a massive accumulation of RPFs downstream of tryptophan codons (herein W-Bumps, Fig. 1d, Extended Data Fig. 1e; shadowed grey). We verified the

presence of W-Bumps on individual genes (ATF4 and CDC6) as opposed to the tryptophan-less ATP5G1 gene (Extended Data Fig. 1g; shadowed grey). The existence of W-Bumps downstream of tryptophan codons suggests that ribosomes, unexpectedly, bypass these sites in IFNγ-treated cells. To examine specificity, we added IDO inhibitor 1-Methyl-L-tryptophan (IDOi) to IFNγ-treated cells (Extended Data Fig. 1h,i). As expected, IDOi negated the global redistribution of RPFs towards the translation initiation sites and the enrichment of the tryptophan codon signal at A-site (Fig. 1f,g, Extended Data Fig. 1j,k). More importantly, IDOi rescued the accumulation of RPFs at the W-Bumps region (Fig. 1f). Additionally, diricore analysis of tryptophan-depleted cells revealed that tryptophan depletion phenocopied IFNγ treatment (Extended data Fig. 1l-o): It inhibited global initiation of translation, reduced RPF density at the ATG^{start} and, most importantly, generated W-Bumps. Altogether, tryptophan exhaustion not only causes ribosome-stalling at tryptophan codons but also induces their accumulation downstream thereof.

Characterization of W-Bumps

We constructed a bioinformatics pipeline (Bump-finder) that unbiasedly identifies regional RPF accumulations in ribosome profiling data (Extended Data Fig. 2a). The vicinity of detected bump regions was then scanned for the frequencies of codons for each individual amino acid. Interestingly, while none of the codons were enriched in the vicinity of bumps in control cells, IFNγ treatment induced a tryptophan signal ~20 amino acid upstream of bumps (Fig. 2a, Extended Data Fig. 2b,c). Next, we examined the abundance-ratio of each codon, 30 triplets upstream and downstream of bumps, and observed an IFNγ-induced tryptophan codon enrichment upstream of the identified bumps (Fig. 2b, Extended Data Fig. 2d). We then analyzed the occurrence of bumps right after each tryptophan codon. While in control cells no bump signal appeared, IFNγ treatment induced bumps downstream of tryptophan codons (Fig. 2c, Extended Data Fig. 2e). A scaled-up analysis of all tryptophans indicated that tryptophan-associated bumps are a widespread phenomenon after IFNγ treatment (Extended Data Fig. 2f). The analysis of IDOi and tryptophan-depletion datasets further strengthened the connection of bumps to tryptophan (Fig. 2d,e). To assess whether this is a global phenomenon that relates to amino acid shortages, we performed diricore analysis of tyrosine-deprived melanoma cells. Interestingly, we identified the induction of Y-Bumps, with similar characteristics to W-Bumps, only associated with tyrosine (Fig. 2f).

5

We sought to identify coding elements associated with W-Bumps by selecting tryptophan codons strongly or weakly associated with W-Bumps (Extended Data Fig. 2g). The only delineating signal-associated with strong W-Bumps was the presence of multiple *in-frame* tryptophan codons within a region of 8 codons (Fig. 2g, Extended Data Fig. 2h-j). We used this feature to assess impact of W-Bumps on protein expression by differential protein expression analysis of IFNγ-treated versus control cells (Extended Data Fig. 3a). This revealed an inverse association between tryptophan content and protein expression (Fig. 2h). Interestingly, this association was not observed when the proteasome was inhibited, and was not observed with other amino acids (Fig. 2i, Extended Data Fig. 3b-e). Further analysis excluded protein length as a contributing factor (Extended Data Fig. 3f). To link W-Bumps to the detected decrease in protein synthesis, we analyzed two groups of proteins, either containing two tryptophans encoded within eight amino acid (<8), or separated by more than eight amino acid (>8). The <8 group showed a stronger W-Bumps signal (Extended Data Fig. 3g,h), and a greater IFNγ-mediated reduction in expression (p-value 2×10^{-6}, Extended Data Fig. 3i), compared to >8. In contrast, for asparagine no significant effect was observed (Extended Data Fig. 3j). Altogether, our results pinpoint the biological importance of W-Bumps in restraining protein synthesis upon IFNγ signaling.

W-Bump connection with ribosomal frameshifts

Given the average distance of ~20 codons between W-Bumps and tryptophan codons, and the periodicity of tryptophan codons within the Bumps, we hypothesized that W-Bumps might be connected to the secondary structure of the nascent peptide in the ribosomal exit tunnel. Though little is known about the influence of secondary structure in this exit tunnel on ribosomal stalling, the formation of an α-helical structure in the tunnel zone is thought to be a major determinant for ribosomal progression [13,14]. Indeed, peptide sequences surrounding W-Bumps form α-helical structures more frequently than other regions in the proteome (Fig. 3a). Possibly, the bypass of ribosome stalling sites after amino acid starvations by frameshifting events induces the loss of these α-helical structures (Extended Data Fig. 4a) [15-21]. To examine this *in-silico*, we scored for disorderedness of *in-frame* and *out-of-frame* peptides by computationally introducing frameshifts at the site of all tryptophans. In general, the level of disorderedness of newly formed peptides after frameshifts greatly increased (Fig. 3b, green line). Nevertheless, a selected outlier group showed highly ordered *out-of-frame* peptides downstream of tryptophan (Fig. 3b, red line). While in general, the *out-of-frame* regions downstream of tryptophan were associated with W-Bumps, the selected group with ordered

regions were not (Fig. 3c, Extended Data Fig. 4b). Therefore, W-Bumps could in part be the result of ribosomes that bypassed tryptophan codons by frameshifting, but then paused with *out-of-frame* aberrant polypeptides in their lower exit tunnel (Extended Data Fig. 4a).

To examine occurrence of frameshifting, we used V5-ATF4[(1-63)]-His lentiviral reporter constructs containing the first 63 amino acid of ATF4 (tryptophan at position 60, preceding a W-Bump, Extended Data Fig. 1g) flanked with V5- and His- tags at the N- and C-termini, respectively. We generated one *in-frame* and two *out-of-frame* His-tag constructs (+1, +2, Fig. 3d), of which the His-tag would only be expressed upon frameshifting events surrounding the tryptophan (Extended Data Fig. 4c). We stably expressed the constructs and examined reporter expression in either mock or IFNγ-treated cells using His-tag pulldowns and V5-tag immunoblotting. Figure 3e shows efficient His-tag pulldown of the *in-frame* reporter in control and IFNγ-treated conditions. In contrast, both *out-of-frame* reporter proteins were retained in the supernatant in control conditions but partially pulled down following IFNγ-treatment, indicating frameshifting events (Fig. 3e). Interestingly, when the supernatants of *in-frame* reporter lysates were subjected to anti-V5 immunoprecipitation to enrich for residual V5-tagged proteins lacking a His-tag at their C-terminus, only IFNγ-treatment induced such proteins (Extended Data Fig. 4d), supporting *frameshifting* also in the *in-frame* reporter situation.

Confirming the causative role of tryptophan shortage, IDOi treatment abolished IFNγ-induced frameshifting and tryptophan depletion induced frameshifting (Extended Data Fig. 4e-h). To examine the importance of tryptophan, we substituted it with tyrosine and observed cessation of frameshifting (Extended Data Fig. 4i). Instead, tyrosine depletion (associated with Y-Bumps formation (Fig. 2f) now led to frameshifting (Extended data Figs. 4j,k). We further confirmed IFNγ-induced frameshifting using different reporters at a single-cell level using the turboGFP (tGFP) gene, which contains no tryptophans. Indeed, V5-ATF4[(1-63)]-tGFP showed IFNγ-induced frameshifting of *out-of-frame* constructs by an increase in *fluorescent* and protein signals, while *in-frame* signals remained largely unchanged (Extended Data Fig. 4l-n).

Next, we examined frameshifting in the presence of T cells, where IFNγ is locally secreted upon recognition of antigens on target cells. We co-cultured anti-melan A (MART-1) T-cells[22] with either IFNγ-sensitive (D10) or the more resistant (888-Mel) melanoma cells, both

7

expressing the MART-1 antigen and the V5-ATF4[(1-63)]-His tagged reporter gene either *in-* or *out-of-frame*. V5-tag immunoblot analysis of His-tag pulldowns demonstrated frameshifting also in this native context (Extended Data Fig. 4o). Frameshifting was only apparent in the D10 cells, in line with their magnitude of IDO1 protein induction (Extended Data Fig. 4p). This effect was recapitulated by IFNγ-treatment, while tryptophan depletion induced frameshifting in both cell lines (Extended Data Fig. 4q-t), indicating that weaker IFNγ-mediated IDO1 induction in 888-Mel is likely the cause of lower frameshifting rate. Altogether, these results confirm the causal role of tryptophan-depletion in the induction of chimeric *trans-frame* proteins by IFNγ.

Induction of endogenous aberrant peptides

We next searched for aberrant peptides in the full proteome of IFNγ-treated and control MD55A3 cells expressing the +1-*out-of-frame*-tGFP reporter. Cells were subjected to 2D-LC MS/MS and a differential expression analysis on this data confirmed upregulation of an IFNγ signature (Extended Data Fig. 5a,b) and the production of tGFP only following IFNγ treatment (Fig. 4a). To identify endogenous tryptophan-associated *trans-frame* proteins, we predicted *out-of-frame* -1 and +1 polypeptides created by frameshifting at endogenously-expressed tryptophans (resulting in 66,728 *trans-frame* polypeptides, Extended Data Fig. 5c). This library was supplemented with the entire proteome and used for scanning of 2D-LC MS/MS data. This led to the detection of 124 *out-of-frame* and *trans-frame* peptides not present in any of the *in-frame* polypeptides (including pseudogenes, alternative mRNA isoforms or upstream open reading frames, Extended Data Fig. 5d), 41 of them were reproduced in two biological replicates (Fig. 4b, Extended Data Fig. 5e). Remarkably, whereas IFNγ treatment led to reduced intensity of proteomic peptides in general (Extended Data Fig. 5f), likely due to reduced mRNA translation (Fig. 1c), most aberrant peptides only appeared in this condition (34 out of 41, p=4.42e-10, Fig. 4b and Extended Data Fig. 5g). In contrast, no general induction was observed in the corresponding *in-frame* genes of the aberrant peptides, and a control reverse-oriented *trans-frame* polypeptide library yielded only 3 peptides, none induced by IFNγ (Fig. 4b and Extended Data Fig. 5h,i). This data demonstrates the occurrence of endogenous frameshifting events at tryptophan residues following IFNγ treatment.

Presentation of aberrant peptides

In accordance to the observed IFNγ protein signature in treated cells (Extended Data Fig. 5b), a strong induction of the immunoproteasome and antigen presentation via Human Leukocyte Antigen (HLA) molecules was observed [23] (Extended Data Fig. 3a,5j-k). As a significant source of HLA-presented peptides are derived from newly-synthesized, rapidly degraded proteins[24-26], as well as from cryptic, non-canonical translation[27-36], we assessed whether IFNγ-induced aberrant peptides may be presented on the cell surface. We first used the model peptide SIINFEKL from mouse ovalbumin that binds to H2-K^b [37]. A375 melanoma cells, which frameshift following IFNγ treatment (Extended Data Fig. 5l), were engineered to express H2-K^b and either *in-frame* or +1-*out-of*-frame-tGFP constructs extended with a SIINFEKL sequence. Figure 4c and Extended Data Fig. 5n show that *in-frame* SIINFEKL is well expressed and presented, and its presentation is mildly induced by IFNγ-treatment in an IDO1-independent manner, presumably by an enhanced antigen processing machinery (Extended Data Fig. 3a[38,39]). In contrast, the presentation of the *out-of-frame* SIINFEKL was undetectable in mock treated cells, but strongly stimulated upon IFNγ-treatment in an IDO1-dependent manner (Fig. 4c, Extended Data Fig. 5m). These results indicate that IFNγ-induced IDO1-mediated tryptophan depletion results in ribosomal frameshifting leading to aberrant peptide presentation.

Next, we examined the presentation of endogenous aberrant peptides using immuno-peptidomics[40] of MD55A3 cells, either untreated, treated with IFNγ, or grown in tryptophan-free media (Extended Data Fig. 6). To detect tryptophan-associated *out-of-frame* peptides, we generated a database containing -1 and +1 polypeptides sequences starting 12 amino acid before tryptophans until the nearest stop codon, or the next Trp codon. We then searched for these aberrant peptides in immunopeptidomics data derived from MD55A3 cells as well as fresh metastases derived from the same patient (Extended Data Fig. 7a, b, patient 55 [40]). We detected 94 HLA-I-bound aberrant peptides, containing both -1 and +1 *out-of-frame* and *trans-frame* sequences, of which 13 were exclusively presented in the metastatic samples, and 81 in the different MD55A3 cell samples (Extended Data Fig. 7a,c,d). Remarkably, comparison between the relative intensities of the altered peptides presented in the treated versus untreated cells, revealed an enrichment in the treated cells (Extended data Fig. 7e). Importantly, out of the fifteen aberrant peptides that were found presented exclusively in the treated samples, six were also detected in the metastases (Extended Data Fig. 7d, Supplementary information).

Overall, the identified aberrant peptides found in the treated MD55A3 cells and in the fresh metastases (n=28) shared similar properties with the *in-frame* HLA-I bound peptides (Extended Data Fig. 7f, g). To validate the identification of those peptides, we generated synthetic peptides, and compared their resulting MS/MS spectra with those of the native, endogenous aberrant peptides. This derived 20 peptides that had a good correlation between their native and synthetic counterparts (Fig 4d, Extended Data Fig. 7h, Supplementary information). In addition, two peptides derived from -1 and +1 frameshifts were further validated by spiking stable isotopically labeled peptides that co-eluted with them (supplementary information).

Interestingly, analysis of a metastatic sample derived from another patient (patient 42 [40]), who shares three of six HLA-I alleles with patient 55, revealed three identical aberrant peptides (Extended Data Fig. 8), suggesting recurrent aberrant peptide presentation across patients. These results confirmed the endogenous production of aberrant peptides and their presentation on the cell surface.

Immunogenicity of aberrant peptides

We next asked whether the HLA complexes containing aberrant peptides can be immunogenic. Previous studies have shown that T cells from healthy donors can recognize tumor-specific peptides ignored by tumor-infiltrating T cells[41]. Aberrant peptides identified in our immunopeptidomic analysis were hence tested for ability to prime naïve CD8+ T cells from healthy donors[42]. Monocyte-derived dendritic cells (MoDCs) isolated from healthy donor peripheral blood mononuclear cells (PBMCs) were pulsed with aberrant peptides and co-cultured with autologous naive CD8+ T cells. Following co-culture, combinatorial tetramer staining analyzed by flow cytometry revealed CD8+ T cells from different donors reactive to the KCNK6- and ZNF513-derived aberrant peptides (Fig. 4e, Extended Data Fig. 9a, b). T cells staining positively with pMHC-multimers were sorted as single cells to generate T-cell clones. Sixteen T cell clones reactive to the KCNK6 peptide expanded sufficiently for subsequent analysis. Among these, thirteen stained positively with multimers and were strongly activated by target cells loaded with relevant peptide (Fig. 4f). Sorted T cells reactive to the ZNF513 peptide did not expand sufficiently for further analysis. These data demonstrate that IFNγ-induced aberrant peptides can be presented on HLA molecules to T cells and induce an immune response.

Conclusions

We show that melanoma cells exposed to prolonged IFNγ exposure derived from their interaction with T cells, induce IDO1-mediated tryptophan depletion. Despite this depletion and despite the stalling of ribosomes on tryptophan codons, mRNA translation proceeds via ribosomal frameshifting. This leads to ribosome stalling downstream of the tryptophan codon due to *out-of-frame* and *trans-frame* aberrant peptide production and loss of secondary structure within the ribosome exit tunnel. Most notably, these aberrant peptides can be detected in the full proteome, can be found presented on HLA-I molecules on melanoma cells, and can prime T cells (Fig. 4g). This novel translational mechanism by which cancer cells cope with amino acid shortages are of particular relevance as they provide a new layer to the complex landscape of melanoma-presented HLA-peptides.

Figure Legends

Figure 1: IFNγ induces IDO1-mediated ribosome pausing on tryptophan codons, and form W-Bumps downstream thereof.

(a) Schematic model depicting the effect of IFNγ (IFN) signaling on IDO1 positive cells. IFNγ induction leads to increased IDO1 expression, an enzyme catalyzing the conversion of tryptophan to kynurenine. This leads to an increase in uncharged tRNAs, which negatively affects protein translation process. But, on the other hand, the production of kynurenine inhibits T cell function. (b) Metagene density profiles depicting global shifts of Ribosomal Protected Fragments (RPFs) to the start of the coding sequence in 12T cells upon IFNγ treatment (red and yellow) as compared to control (black and grey). Y-axis: intra-gene normalized RPF density. (c) Quantification of flow cytometry analysis of OP-Puro incorporation assays as a readout for nascent protein synthesis in mock treated (Ctrl) or IFNγ-treated 12T cells. Bars represent averages plus standard deviation of three independent experiments. (d) Diricore analysis line-plots depicting differential ribosome occupancy in 12T cells (5'-RPF) at -30 to +60 codons surrounding the initiator ATG (ATG^{Start}, green), tryptophan (red) and cysteine codons (grey). Y-axis: Ratio between the number of reads in IFNγ versus control conditions. (e, f) Diricore analysis bar-plots depicting differential codon usage (at position 15 of the RPFs) in IFNγ versus control (panel e), and IFNγ versus IFNγ+IDOi (panel f), in 12T cells. (g) Diricore bar-plot depicting differential codon usage in IDOi versus IFNγ+IDOi treatment of 12T cells.

Figure 2: IFNγ-induced W-Bump formation is associated with the presence of multiple tryptophan codons within a region of eight codons, and is indicative of a reduction in protein synthesis.

(a) Density of codons for alanine (A), serine (S) and tryptophan (W), 60 codons upstream and downstream of bumps identified with bump-finder after IFNγ-treatment of 12T cells. (b) Ratio of reads upstream/downstream 30 codons of bumps identified in control and IFNγ-treated conditions. (c-f) RPF-density plots 200 nucleotides surrounding the tryptophan (W) codon closest to bumps in control (Ctrl), IFNγ-treated (IFN), IFNγ+IDOi, tryptophan (W) depletion, and tyrosine (Y) depletion in 12T cells. (g) Line plot depicting mean tryptophan codon enrichment in the 'Bumps'-group over the 'no-Bumps'-group. Red arrows indicate the enrichment of multiple alternating tryptophan codons within a region of eight codons in the 'Bumps'-group over the 'No-Bumps'-group. (h) Boxplots depicting log fold change in protein levels (red; average of three replicates) and mRNA (black; average of two replicates)

in IFNγ versus control treated cells. Proteins were grouped according to the number of tryptophans in the protein sequence. Boxes depict first, second and third quartile and the whiskers depict the range excluding the outliers. Test: Wilcoxon Test (two-tailed); ns = not significant, ** p<0.005 and *** p<0.0005. Actual p-values are 0.14, $1.2e^{-3}$, $1.5e^{-3}$, $4.5e^{-11}$ for the protein quantification (left) and 0.8, 0.6, 0.7, 0.9 for the RNA quantifications (right) in the order of depiction. (i) Boxplots depicting log fold change in protein levels (red; average of three replicates) in IFNγ versus control treated cells with inclusion of proteasomal inhibition (MG132). Boxes depict first, second and third quartile and the whiskers depict the range excluding the outliers. Test: Wilcoxon Test (two-tailed); ns = not significant, *** p<0.0005. Actual p-values are 0.13, $3.2e^{-4}$, $9.8e^{-6}$, $8.9e^{-12}$ in the order of depiction.

Figure 3: W-Bumps are associated with disordered out-of-frame peptides, and shortage in tryptophan leads to frameshifting events.

(a) Bar-plots depicting total occurrences of structural elements for tryptophan-containing peptide sequences corresponding to W-Bumps, all other sequences, and the difference between these. (b) Line-plots depicting the probability of disorderedness of the *in-frame* peptide up until the tryptophan (dashed line), and the possible frameshifted (both +1 and -1) peptides downstream of that. The average probability of disorderedness of all peptides ('All tryptophan codons', n=36508, green) and a selected group with relatively ordered *out-of-frame* peptide sequences ('Selected tryptophan codons', n=94, red) are shown. (c) Line-plots depicting mean frequency of RPFs at single nucleotide resolution across all tryptophans (green) and the selected outlier set (red). (d) Schematic representation of the reporter constructs (V5-ATF4$^{(1-63)}$-His) used for detection of frameshifting events with the His-tag placed in all three frames. (e) Western blots showing V5-stainings on the supernatants (S) and pulled-down (PD) reporter proteins in MD55A3 cells after control and IFN treatment (n=3).

Figure 4: Aberrant frameshifted peptides are detected in full proteomes, are presented on HLA class I molecules, and induce peptide-specific T cell responses.

(a) Volcano plot showing differentially expressed proteins (IFN/Ctrl) in MD55A3 cells expressing the V5-ATF4$^{(1-63)}$-tGFP +1 *out-of-frame* construct. The arrowhead indicates a peptide that corresponds to tGFP originating from a frameshifting event. (b) Heatmap showing frameshifted peptides identified in the full proteome of IFN-treated MD55A3 cells. (c) Bar graph depicting the median fluorescence intensity (MFI) of H2-K^b-bound SIINFEKL

peptides in A375 cells expressing H2-K^b in combination with *in-frame* (Frame) or +1 out-of-frame (+1) V5-ATF4$^{(1-63)}$-tGFP-SIINFEKL reporters. Bars represent the average of three independent experiments, +/- SD. (d) Panel displaying frameshifted, HLA class I-presented peptides identified by immunopeptidomics. *In-frame* (black), +1 frame (green) and -1 frame (red) peptide sequences are indicated. (e) Flow cytometric analysis of CD8+ T cells reactive to APC and PE-labeled pMHC multimers complexed with the KCNK6-derived transpeptide in co-cultures of naïve CD8+ T cells and autologous MoDCs pulsed with peptide (right) or DMSO vehicle (left). Cells were from an HLA-B*07:02^{pos} healthy donor. (f) Up-regulation of the activation marker CD137 on T cell clones after co-incubation with K562-B*07:02^{pos} cells pulsed with indicated concentrations of peptides (left). Gating strategy is outlined in Extended Data Fig. 9c. Percentage of T cell clones derived from sorted KCNK6 pMHC$^+$ cells staining positively with APC and PE-labeled KCNK6 pMHC multimers (right). Representative results are shown for five reactive and one non-reactive T cell clone. T cells transduced with an HSV-2 HLA-B*07:02-restricted TCR [43] served as a positive control in the activation assay and for pMHC multimer labelling using the relevant HSV-2 peptide. (g) Schematic representation of the effects of IFNγ signaling. The conversion of tryptophan to kynurenine by IDO1 leads to inhibition of T cell function. The current study indicates that the IFNγ-induced depletion of tryptophan leads to frameshifting events and the production of aberrant peptides that are presented on HLA class I molecules and that potentially can activate T cells.

Uncategorized References

1   Ayers, M. *et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* **127**, 2930-2940, doi:10.1172/JCI91190 (2017).

2   Ji, R. R. *et al.* An immune-active tumor microenvironment favors clinical response to ipilimumab. *Cancer Immunol Immunother* **61**, 1019-1031, doi:10.1007/s00262-011-1172-6 (2012).

3   Timosenko, E. *et al.* Nutritional Stress Induced by Tryptophan-Degrading Enzymes Results in ATF4-Dependent Reprogramming of the Amino Acid Transporter Profile in Tumor Cells. *Cancer Res* **76**, 6193-6204, doi:10.1158/0008-5472.CAN-15-3502 (2016).

4   Zhai, L. *et al.* Molecular Pathways: Targeting IDO1 and Other Tryptophan Dioxygenases for Cancer Immunotherapy. *Clin Cancer Res* **21**, 5427-5433, doi:10.1158/1078-0432.CCR-15-0420 (2015).

5   Amobi, A., Qian, F., Lugade, A. A. & Odunsi, K. Tryptophan Catabolism and Cancer Immunotherapy Targeting IDO Mediated Immune Suppression. *Adv Exp Med Biol* **1036**, 129-144, doi:10.1007/978-3-319-67577-0_9 (2017).

6   Labadie, B. W., Bao, R. & Luke, J. J. Reimagining IDO Pathway Inhibition in Cancer Immunotherapy via Downstream Focus on the Tryptophan-Kynurenine-Aryl Hydrocarbon Axis. *Clin Cancer Res* **25**, 1462-1471, doi:10.1158/1078-0432.CCR-18-2882 (2019).

7   Gunther, J., Dabritz, J. & Wirthgen, E. Limitations and Off-Target Effects of Tryptophan-Related IDO Inhibitors in Cancer Treatment. *Front Immunol* **10**, 1801, doi:10.3389/fimmu.2019.01801 (2019).

8   Battu, S., Minhas, G., Mishra, A. & Khan, N. Amino Acid Sensing via General Control Nonderepressible-2 Kinase and Immunological Programming. *Front Immunol* **8**, 1719, doi:10.3389/fimmu.2017.01719 (2017).

9   Wek, R. C. & Staschke, K. A. How do tumours adapt to nutrient stress? *EMBO J* **29**, 1946-1947, doi:10.1038/emboj.2010.110 (2010).

10   Ye, J. *et al.* The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *EMBO J* **29**, 2082-2096, doi:10.1038/emboj.2010.81 (2010).

11   Adam, I. *et al.* Upregulation of tryptophanyl-tRNA synthethase adapts human cancer cells to nutritional stress caused by tryptophan degradation. *Oncoimmunology* **7**, e1486353, doi:10.1080/2162402X.2018.1486353 (2018).

12   Loayza-Puch, F. *et al.* Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. *Nature* **530**, 490-494, doi:10.1038/nature16982 (2016).

13   Bhushan, S. *et al.* alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat Struct Mol Biol* **17**, 313-317, doi:10.1038/nsmb.1756 (2010).

14   Woolhead, C. A., McCormick, P. J. & Johnson, A. E. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **116**, 725-736, doi:10.1016/s0092-8674(04)00169-2 (2004).

15   Caliskan, N. *et al.* Conditional Switch between Frameshifting Regimes upon Translation of dnaX mRNA. *Mol Cell* **66**, 558-567 e554, doi:10.1016/j.molcel.2017.04.023 (2017).

16   Yelverton, E., Lindsley, D., Yamauchi, P. & Gallant, J. A. The function of a ribosomal frameshifting signal from human immunodeficiency virus-1 in Escherichia coli. *Mol Microbiol* **11**, 303-313, doi:10.1111/j.1365-2958.1994.tb00310.x (1994).

17   Gurvich, O. L., Baranov, P. V., Gesteland, R. F. & Atkins, J. F. Expression levels influence ribosomal frameshifting at the tandem rare arginine codons AGG_AGG and AGA_AGA in Escherichia coli. *J Bacteriol* **187**, 4023-4032, doi:10.1128/JB.187.12.4023-4032.2005 (2005).

18   Olubajo, B. & Taylor, E. W. A -1 frameshift in the HIV-1 env gene is enhanced by arginine deficiency via a hungry codon mechanism. *Mutat Res* **579**, 125-132, doi:10.1016/j.mrfmmm.2005.02.018 (2005).

19   Barak, Z., Lindsley, D. & Gallant, J. On the mechanism of leftward frameshifting at several hungry codons. *J Mol Biol* **256**, 676-684, doi:10.1006/jmbi.1996.0117 (1996).

20   Laine, S., Thouard, A., Komar, A. A. & Rossignol, J. M. Ribosome can resume the translation in both +1 or -1 frames after encountering an AGA cluster in Escherichia coli. *Gene* **412**, 95-101, doi:10.1016/j.gene.2008.01.018 (2008).

21   Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R. N. & Chrzanowska-Lightowlers, Z. M. Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* **327**, 301, doi:10.1126/science.1180674 (2010).

22   Vredevoogd, D. W. *et al.* Augmenting Immunotherapy Impact by Lowering Tumor TNF Cytotoxicity Threshold. *Cell* **178**, 585-599 e515, doi:10.1016/j.cell.2019.06.014 (2019).

23   Zhou, F. Molecular mechanisms of IFN-gamma to up-regulate MHC class I antigen processing and presentation. *Int Rev Immunol* **28**, 239-260, doi:10.1080/08830180902978120 (2009).

24   Bourdetsky, D., Schmelzer, C. E. & Admon, A. The nature and extent of contributions by defective ribosome products to the HLA peptidome. *Proc Natl Acad Sci U S A* **111**, E1591-1599, doi:10.1073/pnas.1321902111 (2014).

25   Yewdell, J. W. DRiPs solidify: progress in understanding endogenous MHC class I antigen processing. *Trends Immunol* **32**, 548-558, doi:10.1016/j.it.2011.08.001 (2011).

26   Trentini, D. B. *et al.* Role for ribosome-associated quality control in sampling proteins for MHC class I mediated antigen presentation. *Proc Natl Acad Sci U S A* **117**, 4099-4108, doi:10.1073/pnas.1914401117 (2020).

27   Erhard, F. *et al.* Improved Ribo-seq enables identification of cryptic translation events. *Nat Methods* **15**, 363-366, doi:10.1038/nmeth.4631 (2018).

28   Prasad, S., Starck, S. R. & Shastri, N. Presentation of Cryptic Peptides by MHC Class I Is Enhanced by Inflammatory Stimuli. *J Immunol* **197**, 2981-2991, doi:10.4049/jimmunol.1502045 (2016).

29   Starck, S. R. & Shastri, N. Nowhere to hide: unconventional translation yields cryptic peptides for immune surveillance. *Immunol Rev* **272**, 8-16, doi:10.1111/imr.12434 (2016).

30   Chen, J. *et al.* Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140-1146, doi:10.1126/science.aay0262 (2020).

31   Chong, C. *et al.* Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun* **11**, 1293, doi:10.1038/s41467-020-14968-9 (2020).

32   Laumont, C. M. & Perreault, C. Exploiting non-canonical translation to identify new targets for T cell-based cancer immunotherapy. *Cell Mol Life Sci* **75**, 607-621, doi:10.1007/s00018-017-2628-4 (2018).

33   Laumont, C. M. *et al.* Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med* **10**, doi:10.1126/scitranslmed.aau5516 (2018).

34   Pearson, H. *et al.* MHC class I-associated peptides derive from selective regions of the human genome. *J Clin Invest* **126**, 4690-4701, doi:10.1172/JCI88590 (2016).

35   Laumont, C. M. *et al.* Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun* **7**, 10238, doi:10.1038/ncomms10238 (2016).

36   Saulquin, X. *et al.* +1 Frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J Exp Med* **195**, 353-358, doi:10.1084/jem.20011399 (2002).

37   Dersh, D., Yewdell, J. W. & Wei, J. A SIINFEKL Based System to Measure MHC Class I Antigen Presentation Efficiency and Kinetics. *Methods Mol Biol* **1988**, 109-122, doi:10.1007/978-1-4939-9450-2_9 (2019).

38   McCarthy, M. K. & Weinberg, J. B. The immunoproteasome and viral infection: a complex regulator of inflammation. *Front Microbiol* **6**, 21, doi:10.3389/fmicb.2015.00021 (2015).

39   Goldberg, A. L., Cascio, P., Saric, T. & Rock, K. L. The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides. *Mol Immunol* **39**, 147-164, doi:10.1016/s0161-5890(02)00098-6 (2002).

40   Kalaora, S. *et al.* Combined Analysis of Antigen Presentation and T-cell Recognition Reveals Restricted Immune Responses in Melanoma. *Cancer Discov* **8**, 1366-1375, doi:10.1158/2159-8290.CD-17-1418 (2018).

41   Stronen, E. *et al.* Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337-1341, doi:10.1126/science.aaf2288 (2016).

42   Ali, M. *et al.* Induction of neoantigen-reactive T cells from healthy donors. *Nat Protoc* **14**, 1926-1943, doi:10.1038/s41596-019-0170-6 (2019).

METHODS:

Alignment of Ribosome profiling data:

Preprocessing of FASTQ files (Extended Data Table 2) consisted of adapter removal using cutadapt (https://doi.org/10.14806/ej.17.1.200) with the parameters (--quality-base=33 -O 7 -e 0.15 -m 20 -q 5) and removal of rRNA and tRNA contaminants by means of alignment against a reference (rRNA reference data from GENCODE [44]v19: rRNA,MT_RNA,rRNA_pseudogene, and tRNA reference data from GtRNAdb) using bowtie2 [45] with parameters (--seed 42 -p1 –local).

The actual alignment of pre-processed FASTQ files was done with TopHat2[46] and Bowtie2 [45] against GRCh37/hg19 and GENCODE v19/BASIC transcript with Ensembl coordinates using parameters (seed 42 -n 2 -m 1 --no-novel-juncs --no-novel-indels --no-coverage-search --segment-length 25). In a subsequent step the primary aligned reads filtered for a minimum mapping quality of 10.

Quality check of the FASTQ files were undertaken using FASTQC package (ref), and the quality analysis of frames and periodicity of RPFs was undertaken using RiboWaltz [47](Extended Data Fig. 11).

Diricore Analysis:

For the subsequence analysis RPF codon occupancy frequency between two conditions (e.g. plus IFNγ versus minus IFNγ) are compared as described in [12]. RPF density analysis is performed by the comparison of normalized 5' RPF density analysis per codon between conditions [12].

RNA-seq data analysis:

RNA-seq data (Extended Data Table 2), as FASTQ file, were aligned to human hg19 genome using TopHat [46] SAMTOOLS [48] was used for file format conversions. HTSeq [49] was used to count reads at exons of protein coding genes. Library size normalization of read counts was done using DESEQ [50].

Finding Bumps in the Ribosome profiling data and associating them with amino-acids:

The reads (FASTQ) from Riboseq experiments were aligned to human transcript assembly (gencode v19) after removal of low quality reads and the reads that align to tRNA and rRNA (refer methods section- Diricore). Transcript alignment was performed using Bowtie [45] with the default parameters.

BAM files were converted to BED using BEDTOOLS[51] and later file formats were edited using PERL scripts. Each gene was divided into 100 windows of equal length, and read (separately for every sample) at each window were quantified using BEDTOOLS. The array of reads, logged (base 2), were normalized between 0 to 1. For this, the replicates for both the conditions (untreated and

IFNγ treated conditions) were taken collectively as average. Thereafter, PEAKS were identified in an array of reads per window using findpeaks (pracma v1.9.9) function in R with (nups =1 , ndown=1, minpeakheight=10) parameters.

Only peaks called in treatment condition (merged in two replicates) were identified as TREATMENT (IFNγ-treated) peaks, while peaks called in minus condition were identified as CONTROL peaks. The highest point (window with the greatest number of reads) per peak, was marked as reference point.

The transcript position of the reference point, was converted to protein coordinate using ensembldb 2.8.0 [52] in R. The amino-acids were mapped 30 codons at each side of the reference point and was quantified as sum at every individual position in PERL. Line plots were then plotted in R. Upstream/Downstream ratio for every amino-acid was quantified as the ratio of average presence of a particular amino acid upstream (30 codons) upstream versus 30 codons downstream. The scripts are available with additional detailing in the github package; https://github.com/apataskar/bump_finder_example2.

Transcript density plots, are plotted as the function of density (R function) of read across the nearest "tryptophan" to the reference-points identified in the respective cell-lines. For global Bump-analysis, transcript positions of TGG codon encoding tryptophan amino-acid *in-frame* were obtained using PERL script. The frequency of occurrence of P-sites (12th position from offset of the read) from Ribosome profiling samples across 30 codons (upstream and downstream) were plotted as density function in R and as heatmap using pheatmap in R.

Proteomics analysis (relates to Fig. 2h,i, and Extended Data Figs. 3 a,b,e-h)

I) Sample preparation for proteomics
Frozen MD55A3 cell pellets were lysed, reduced and alkylated in heated guanidine (GuHCl) lysis buffer as described by Jersie-Christensen et al [53]. Proteins were digested with Lys-C (Wako) for 2h at 37°C, enzyme/substrate ratio 1:100. The mixture was then diluted to 2 M GuHCl and digested overnight at 37°C with trypsin (Sigma) in enzyme/substrate ratio 1:50. Digestion was quenched by the addition of TFA (final concentration 1%), after which the peptides were desalted on a Sep-Pak C18 cartridge (Waters, Massachusetts, USA). Samples were vacuum dried and stored at -80°C until LC-MS/MS analysis.

Peptides were reconstituted in 2% formic acid and analyzed by nano LC-MS/MS on an Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer equipped with an EASY-NLC 1200 system (Thermo Scientific). Samples were directly loaded onto the analytical column (ReproSil-Pur 120 C18-

AQ, 2.4 μm, 75 μm × 500 mm, packed in-house). Solvent A was 0.1% formic acid/water and solvent B was 0.1% formic acid/80% acetonitrile. Peptides were eluted from the analytical column at a constant flow of 250 nl/min. For single-run proteome analysis, a 3h gradient was employed containing a linear increase from 4% to 26% solvent B, followed by a 15-min wash. MS settings were as follows: full MS scans (375-1500 m/z) were acquired at 60,000 resolution with an AGC target of 3×10^6 charges and max injection time of 45 ms. Loop count was set to 20 and only precursors with charge state 2-7 were sampled for MS2 using 15,000 resolution, MS2 isolation window of 1.4 m/z, 1×10^5 AGC target, a max injection time of 22 ms and a normalized collision energy of 26.

II) Data analysis

RAW files were analyzed by Proteome Discoverer (version 2.3.0.523, Thermo Scientific) using standard settings. MS/MS data were searched in Sequest HT against the the human Swissprot database (20,381 entries, release 2018_08). The maximum allowed precursor mass tolerance was 50 ppm and 0.06 Da for fragment ion masses. False discovery rates for peptide and protein identification were set to 1%. Trypsin was chosen as cleavage specificity allowing two missed cleavages. Carbamidomethylation (C) was set as fixed modification, whereas oxidation (M) and protein N-terminal acetylation were set as variable modifications. Peptide spectrum matches (PSM) were filtered for Sequest HT Xcorr score ≥ 1. The Proteome Discoverer output file containing the LFQ abundances was loaded into Perseus (version 1.6.5.0) [54]. Abundances were Log2-transformed and the proteins were filtered for at least two out of three valid values in one condition. Missing values were replaced by imputation based on the standard settings of Perseus, i.e. a normal distribution using a width of 0.3 and a downshift of 1.8. Differentially expressed proteins were determined using a t-test (threshold: FDR 1% or FDR 5% and S0: 0.13).

GENCODE annotations (gencode v19) were used to calculate number of amino acids per protein as well as the least distance between a particular amino-acid (W,Y and N) Boxplots for every group were plotted in R. Statistical tests were down using Wilcoxon test in R.

Proteome analysis for the detection of frameshifted polypeptides (relates to Fig. 4 a-b, and Extended Data Fig. 5 a-i, Supplementary ~~XXX~~information)

I) Sample preparation

For deeper proteome coverage in search of IFNγ-induced frameshifts, IFNγ- or mock treated melanoma cells were lysed and digested as described above, after which dried digests were subjected to basic reversed-phase (HpH-RP) high-performance liquid chromatography for offline peptide fractionation. 250 μg peptides were reconstituted in 95% 10 mM ammonium

hydroxide (NH$_4$OH, solvent A)/5% (90% acetonitrile (ACN)/10mM NH$_4$OH, solvent B) and loaded onto a Phenomenex Kinetex EVO C18 analytical column (150 mm x 2.1 mm, particle size 5 μm, 100 Å pores) coupled to an Agilent 1260 HPLC system equipped with a fraction collector. Peptides were eluted at a constant flow of 100 μL/min in a 90-minute gradient containing a nonlinear increase from 5-30% solvent B.   Fractions were collected and concatenated to 24 fractions per sample replicate. All fractions were analyzed by nanoLC-MS/MS on an Orbitrap Fusion Tribrid mass spectrometer equipped with an Easy-nLC1000 system (Thermo Scientific) as described previously[55]. Peptides were directly loaded onto the analytical column (ReproSil-Pur 120 C18-AQ, 1.9μm, 75 μm × 500 mm, packed in-house). Solvent A was 0.1% formic acid/water and solvent B was 0.1% formic acid/80% acetonitrile. Samples were eluted from the analytical column at a constant flow of 250 nl/min in a 2h-gradient containing a linear increase from 8-32% solvent B.

II) In silico tryptophan- associated frameshift database construction

For an overview see Extended Data Fig. 12. The CDS sequences of GRCh38 were downloaded from Ensembl [56]. Prime transcripts (annotated as -001), which contain tryptophan codon and are highly expressed (log2(Normalized Read Counts) > 5) in the ribosome profiling data were included for further analysis. Transcripts with less than 50bp were discarded. Only CDS starting with ATG were kept. In cases of multiple *in-frame* TGG-codons per transcript, each TGG along the sequence was frameshifted separately. Both +1 and -1 frameshifts at the TGG codon position were implemented. The CDS *out-of-frame* was *in silico* translated until the first stop codon. Finally, we generated a database of chimeric polypeptides, starting at the first tryptic cleavage start site, upstream of tryptophan (33628 instances), frame-shifted at TGG codon (Extended Data Figure 5f) via both -1 and +1 frame-shifts, until out-of-frame stop codons. No further filtering was implemented at the pre-scanning stage.

III) Data search, filtering and analysis

Further, the MS data was analyzed by MaxQuant (version 1.6.0.16)[57]. The WT human proteome to run against was obtained from ~~Uniprot~~ UniProt database with SWISSPROT protein evidence level of 1 [58] . The in-frame protein expression data was further normalized and analyzed using DEP[59], For the frame-shift proteome analysis, the MS data was analyzed by MaxQuant (version 1.6.0.16)[57] with LFQ normalization, and then was scanned against the trans-frame polypeptide database together with the SWISSPROT proteins with Evidence

22

level of 1. After scanning, a total of 124 peptides from the trans-frame polypeptide database were retained for further quantitative analysis, after subjecting to filtering for mapping to any proteomic and non-coding sequences (Extended Data Fig. 5c). Only the peptides reproducibly detected across replicates were retained for further analysis, and the list included reverse peptide hits (n=41) (Fig. 4c).

Immunopeptidomics analysis (relates to Fig. 4d and Extended Data Figs. 6, 7, Supplementary ~~XXX~~information)

I) Sample preparation

Untreated MD55A3 (n=4), IFNγ treated (n=4), and Trp depleted (n=4) cell pellets were subjected to HLA-purification as described previously[40,60], with slight modifications; Briefly, cell pellets were lysed with lysis buffer containing 0.25% sodium deoxycholate, 0.2mM iodoacetamide, 1mM EDTA, 1:200 protease inhibitors cocktail (Sigma-Aldrich), 1mM PMSF and 1%octyl-b-D glucopyranoside in PBS, and then incubated at 4˚C for 1 hour. The lysates were cleared by centrifugation at 4˚C and 48,000g for 60 minutes, and then passed through a pre-clearing column containing Protein-A Sepharose beads (GenScript). HLA-I molecules were immunoaffinity purified from cleared lysate with the pan-HLA-I antibody (W6/32 antibody purified from HB95 hybridoma cells) covalently bound to Protein-A Sepharose beads). Affinity column was washed first with 10 column volumes of 400mM NaCl, 20mM Tris–HCl, pH 8.0 and then with 10 volumes of 20mM Tris–HCl, pH 8.0. The HLA peptides and HLA molecules were eluted with 1% TFA followed by separation of the peptides from the proteins by binding the eluted fraction to disposable reversed-phase Sep-Pak tC18 (Waters). Elution of the peptides was done with 30% acetonitrile (ACN) in 0.1% trifluoracetic acid (TFA).

II) Liquid chromatography MS-analysis

The HLA peptides were dried by vacuum centrifugation, re-solubilized with 0.1% formic acid and separated using reversed phase chromatography using the nanoAquity system (Waters Corp., USA), with a Symmetry trap column (180x20mm) and HSS T3 analytical column, 0.75x250mm (Waters Corp. USA). The chromatography system was coupled by electrospray to tandem mass spectrometry to Q-Exactive-Plus (Thermo Fisher Scientific).

The HLA peptides were eluted with a linear gradient over 2 h from 5 to 28% acetonitrile with 0.1% formic acid at a flow rate of 0.35μl/min.

Data was acquired using a data-dependent "top 10" method, fragmenting the peptides by higher-energy collisional dissociation (HCD). Full scan MS spectra was acquired at a resolution of 70,000 at 200 m/z with a target value of 3x106 ions. Ions were accumulated to an AGC target value of 105 with a maximum injection time of generally 100 msec. The peptide match option was set to Preferred. Normalized collision energy was set to 25% and MS/MS resolution was 17,500 at 200 m/z. Fragmented m/z values were dynamically excluded from further selection for 20 sec.

Data analysis

III) In-silico tryptophan-associated frameshift database construction

For overview ~~see Extended Data Fig. 12~~see Supplementary information.

The CDS sequences of GRCh38 were downloaded from Ensembl [56]. All transcript variants were included. Transcripts with less than 50bp were discarded. Sequences in which no *in-frame* TGG codon (corresponding to tryptophan) exists were excluded. Only CDSs starting with ATG were kept. In cases where there were multiple *in-frame* TGG-codons per transcript, each TGG along the sequence was frameshifted separately. Both +1 and -1 frameshifts at the TGG codon position were implemented. The CDS *out-of-frame* was in silico translated until the first stop codon, or the next tryptophan obtained. The *in-frame* portion of the sequence was trimmed at the N', such that it contained 12 amino acids upstream to the frameshift for the peptidomics database (as a 12 amino acid window upstream to the slippery tryptophan consists of all possible HLA I bound altered peptides, derived from this ribosomal slippage). At the last step, sequence redundancy was removed in cases of 100% sequence identity, and the longest sequence was kept using CD-HIT[61].

IV) Database search and filtration

The RAW MS data files were analyzed by MaxQuant (version 1.6.0.16). Files were searched against the frame-shifted database and the full canonical human proteome. The canonical human proteome was obtained from Ensembl GRCh38 and Uniprot database [58] following removal of 100% sequence redundancy using CD-HIT[61]. The maximum allowed precursor mass tolerance was 20 ppm. N-terminal acetylation and methionine oxidation were set as variable modifications. A PSM false discovery rate (FDR) of 0.05 was used, and no protein

FDR was set. Enzyme specificity was set as "unspecific" and "match between runs" option was set with default settings and LFQ was set to "minimum ratio count" of 1. The obtained peptides were filtered, by multiple criteria (Extended Data Fig. 7a). Only peptides obtained by the frameshifted database and not the canonical database were kept. Peptides with Maxquant scores less than 80 or PEP larger than 0.1 were discarded. Any peptide not predicted by NetMHCpan (version 4.0)[62] to bind the cell line HLA-alleles (either as strong-threshold of % rank 0.5 or weak binders threshold of % rank 5) were removed. In addition, in order to avoid false positive hits derived from poorly fragmented spectra and ambiguous sequence, we further filtered the detected peptides based on "Fragmentation coverage"(FC), defined as matched ions (a, b or y) divided by the total theoretical ions of the matched peptide sequence (peptide length-1). FC was calculated and only peptides showing MS/MS FC greater than 60% were kept). Peptides derived from source protein with expression in at least one dataset (transcriptome /translatome) were kept. Furthermore, peptides that were obtained in one or more control (non-treated) samples were not further investigated. Lastly, we confirmed that none of the corresponding identified aberrant peptides were generated from pseudogenes by aligning them to gencode protein-coding transcript sequences version 34, containing polymorphic pseudogenes entries. Similarly, we confirmed that the identified altered peptides were not derived from INDELS (using GATK4 version 4.1.4.1[63] haplotype caller for variant calling) or from intron retention (IR) events [64].

Gibbs clustering

Quality assessment of the identified peptides was done using the GibbsCluster2.0 server [65] by clustering to 1-6 groups. These groups were compared to the expected motifs identified, using curated IEDB database[66]. The expected motifs were derived using http://hlathena.tools/ with peptide length set to 9 [67,68].

V) Hydrophobicity index (HI) prediction

Sequence specific HI was calculated with an online available tool SSRCalc[31],[69]: http://hs2.proteome.ca/SSRCalc/SSRCalcQ.html. HI prediction obtained from the SSRCalc based on the 100Å C18 column, 0.1% formic acid separation system and without cysteine protection.

VI) Synthetic peptide validation

Light synthetic peptides for spectra validation were ordered from GenScript, as HPLC grade (≥85% purity). These were analyzed using the same LC-MSMS system and acquisition parameters as indicated above for the endogenous peptides, with the following changes: the gradient was from 4% to 30% acetonitrile in 20min; NCE was set to 27. The data was processed with MaxQuant using the following parameters, all FDRs were set to 1, the individual peptide mass tolerance was set to false.

To compare the endogenous and synthetic spectra, we utilized the MSnbase R package [70] to calculate the Pearson correlations of fragment ions including a, b and y ions without neutral losses, detected in spectra of both endogenous and synthetic peptides, and to plot the head to tail graph. In all cases, we selected only the spectra that harbored the same precursor charge as the endogenous peptides and that were not post translationally modified. We then selected the synthetic spectra that had the best score in MaxQuant.

Selected peptides were ordered from JPT as synthetic peptides with one stable isotope-labeled amino acid, at ≥ 95 % purity. The mass spectrometer was operated at a resolution of 70,000 (at m/z = 200) for the MS1 full scan, scanning a mass range from 300 to 1650 m/z with an ion injection time of 120 ms and an AGC of 3e6. Then each peptide was isolated with an isolation window of 1.7 m/z prior to ion activation by high-energy collision dissociation (HCD, NCE = 27). Targeted MS/MS spectra were acquired at a resolution of 35,000 (at m/z = 200) with an ion injection time of 100 ms and an AGC of 2e5.

The PRM data were processed and analyzed by Skyline (v20.1.0.76)[71], and an ion mass tolerance of 0.02 m/z was used to extract fragment ion chromatograms. Data was smoothed by the Savitzky Golay algorithm.

Prediction of disorderedness for variant peptides

The frame-shifted library for proteomics was subject to disorderedness prediction, albeit from the in-frame start codon. Only those peptides were retained for which *in-frame* part is longer than 25 amino-acids, while the *out-of-frame* part is longer than 30 amino-acids (stop codon occurs later). Disorderedness probability was obtained using IUPRED2A [72]. Outlier groups were selected by using following cutoffs. Average *out-of-frame* (both -1 and +1) disordered score < 0.15, difference between disorderedness prediction *in-frame* and *out-of-frame* part is less than 0.95 (FoldChange).

Cells and reagents

26

Cell lines 12T and 108T were derived from pathology-confirmed metastatic melanoma tumor resections collected from patients enrolled in institutional review board (IRB)-approved clinical trials at the Surgery Branch of the National Cancer Institute. The MD55A3 cell line was derived from metastatic melanoma tumor resections collected with informed patient consent under a protocol approved by the NIH Institutional Review Board (IRB) Ethics Committee and approved by the MD Anderson IRB (protocol numbers 2012-0846, LAB00-063, and 2004-0069; NCT00338377). All cell lines were tested regularly and were found negative for mycoplasma contamination (EZ-PCR Mycoplasma Kit, Biological Industries). 12T and 108T cells were authenticated by Finger printing with STR profiling (Panel: PowerPlex_16_5Nov142UAGC, Size: GS500 x35 x50 x250, Analysis Type: Fragment (Animal), Software Package: SoftGenetics GeneMarker 1.85). 12T, 108T, MD55A3, 888-Mel and D10 cells were cultured in Roswell Park Memorial Institute 1640 Medium (RPMI 1640, Gibco) supplemented with heat-inactivated 10% fetal bovine serum (Sigma), 25 mM HEPES (Gibco) and 100 U/mL penicillin/streptomycin (Gibco). HEK293T and A375 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Gibco), supplemented with 10% fetal bovine serum and 100 U/mL penicillin/streptomycin. All cell lines were maintained in a humidified atmosphere containing 5% of CO_2 at 37°C. PBMCs were isolated from the blood of healthy donors provided by the Norwegian Blood Bank. In-house donor PBMCs were isolated from the healthy donors under informed consent and HLA-typed. Tryptophan-free DMEM/F12 media was purchased from US Biologicals, custum-made tyrosine-free medium was custom purchased from Cell Culture Technologies, IFNγ (PeproTech) was used at 250 U/mL for 48hrs. MG-132 (Selleckchem), dissolved in DMSO, was used at a final concentration of 10 μM. IDO inhibitor, 1-Methyl-L-tryptophan (Sigma) was dissolved in 0.1N NaOH at a 20mM concentration, adjusted to pH=7.5, filter sterilized and used at a final concentration of 300μM for 48h. Polyethylenimine (PEI, Polysciences) was dissolved in water at a concentration of 1 mg/mL, after which it was filter sterilized, aliquoted and stored at -20°C

Ribosome profiling

The construction of ribosome protected fragments (RPF) libraries were done as previously described [73]. For the generation of total RNA libraries, total RNA was extracted using TRI Reagent (Sigma), and mRNA was purified using Dynabeads mRNA DIRECT Purification Kit (Invitrogen), according to the manufacturers' protocol. Libraries were constructed using

SENSE Total RNA-Seq Library Prep Kit for Illunima (Lexogen). RPF and Total RNA Libraries were loaded onto Illumina NextSeq 500 sequencer (Illumina)

Lentiviral production and transduction

For lentivirus production, 4 x 10⁶ HEK293T cells were seeded per 100 mm dish, one day prior to transfection. For each transfection, 10 µg of the pCDH reporter, 5 µg of pMDL RRE, 3.5 µg pVSV-G AND 2.5 µg of pRSV-REV plasmids were mixed in 500 µL of serum-free DMEM. Next, 500 µL of serum-free DMEM containing 63 µL of a 1 mg/mL PEI solution was added. The entire mix was vortexed and left for 15 minutes at room temperature, after which it was added to the HEK293T cells to be transfected. The next day, the medium was replaced by RPMI. The lentivirus-containing supernatants were collected 48 and 72 hours post transfection, and snap frozen in liquid nitrogen. Target cells were transduced on two consecutive days by supplementation of the lentiviral supernatant with 8 µg/mL polybrene (Sigma). One day after the last transduction, transduced cells were selected by addition of 5 µg/mL blasticidin (Invivogen) to the medium.

His-tag pulldown and Western blotting

At the end of each experiment intended for His-tag pulldown, cells were treated with 10 µM MG-132 for 4 hours and subsequently collected by trypsinization and centrifugation. Next, cells were lysed in 400 µL 1x binding/washing buffer containing 1% Triton X-100. Next, his-tag pulldowns were performed with Dynabeads His-tag isolation and pulldown (Thermo Fisher Scientific) according to manufacturer's protocol. All pulled down protein was eluted in 100 µL elution buffer, after which the samples pulled down from *+1* and *+2* reporter-expressing cells were precipitated using acetone. Four volumes of ice-cold acetone were added to the samples, after which they were stored for 30 minutes at -20°C. The samples were then centrifuged at maximum speed at 4°C for 10 minutes, after which the supernatant was removed and the samples were resuspended in 40 µL of 1x Laemmli buffer. All V5-ATF4-His proteins were visualized by first running the samples on 20% SDS-PAGE gels and blotting on 22 µm pore size nitrocellulose membranes (Pall Corporation). V5-stainings were performed using V5 tag monoclonal antibodies (Thermo Fisher Scientific, #R960-25; 1:1000), tGFP stainings with Rabbit anti TurboGFP (Thermo Scientific, cat. Nr. PA5-22688; 1:1000) and IRDye 680RD Donkey anti-Mouse (LI-COR, #926-68072, 1:10,000) and IRDye 800CW Goat anti-Rabbit (LI-COR, #926-32211, 1:10,000) secondary antibodies. Visualization was performed by use of an Odyssey infrared scanning device (LI-COR). IDO1

was visualized with anti-IDO D5J4E ™ Rabbit mAb (Cell Signaling, 1:1000) followed by Peroxidase-conjugated Goat anti mouse antibody (Jackson ImmunoResearch #115-035-003, 1:10000) secondary antibody blot, and tubulin- via anti-Tubulin (DM1A, SIGMA, 1:10000), followed by Peroxidase-conjugated Goat anti rabbit antibody (Jackson ImmunoResearch #111-035-003, 1:10000) secondary antibody blot.

Flow cytometry analyses OPP measurments
Cells expressing the V5-ATF4-tGFP reporters were seeded and mock treated or treated with IFNγ the next day. 48 hours after the start of the experiment, cells were treated with 10 µM MG-132 for 4 hours and subsequently collected by trypsinization and centrifugation. Next, the cells were analyzed on a Attune NxT machine (Thermo Fisher Scientific). The data were analyzed using FlowJo V10 software (FlowJo).

For OPP measurements, cells were seeded and mock treated or treated with IFNγ the next day. 48 hours after the start of treatment, 10 µM of OPP (O-propargyl-puromycin Life Technologies) was added and incorporation was allowed for 60 minutes at 37 degrees. Next, the cells were harvested by trypsinization and centrifugation and fixed in 70% ethanol overnight at 4 degrees Celsius. The next day, the cells were washed with PBS, permeabilized with 0.1% Triton X-100 (Sigma) and blocked with 3% bovine serum albumin (Sigma) in PBS. Subsequently the click-it reaction was performed using click-it reagents and picol azide AF488 (all from Thermo Fisher Scientific). The cells were analyzed on a BD LSR Fortessa (BD Biosciences). The data were analyzed using FlowJo V10 software (FlowJo).

Amino Acid Mass Spectrometry
Cells were washed with cold PBS and lysed with lysis buffer composed of methanol/acetonitrile/H$_2$O (2:2:1). The lysates were collected and centrifuged at 16,000 g (4°C) for 15 minutes and the supernatant was transferred to a new tube for liquid-chromatography mass spectrometry (LC-MS) analysis. For media samples, 10 µl of medium was mixed with 1 ml lysis buffer and processed as above.

LC-MS analysis was performed on an Exactive mass spectrometer (Thermo Scientific) coupled to a Dionex Ultimate 3000 autosampler and pump (Thermo Scientific). Metabolites were separated using a Sequant ZIC-pHILIC column (2.1 x 150 mm, 5 µm, guard column 2.1 x 20 mm, 5 µm; Merck) using a linear gradient of acetonitrile (A) and eluent B (20 mM

29

(NH4)2CO3, 0.1% NH4OH in ULC/MS grade water (Biosolve)), with a flow rate of 150 µL/min. The MS operated in polarity-switching mode with spray voltages of 4.5 kV and -3.5 kV. Metabolites were identified on the basis of exact mass within 5 ppm and further validated by concordance with retention times of standards. Quantification was based on peak area using LCquan software (Thermo Scientific).

T cell co-culture

Melan-A/MART-1 expressing 888-Mel and D10 cells were transduced with V5-ATF4-His *in frame* and *+1* constructs. One day after seeding these 888-Mel and D10 cells in 10 cm dishes, MART-1-specific T cells were added overnight at three different dilutions along with a control that did not receive any T cells. The next day, the plates with the dilution showing around 40% efficient T cell-mediated killing of D10 cells were harvested for both lines, along with the control. His-tag pulldown and Western blotting was performed as indicated above.

SIINFEKL based peptide display

A DNA sequence coding for the amino acids LEQLESIINFEKL was cloned immediately downstream of the tGFP sequence in the pCDH-V5-ATF4$^{(1-63)}$-tGFP reporter constructs (in-frame and +1). This was done by PCR on the V5-ATF4$^{(1-63)}$-tGFP *in-frame* and +1 constructs as templates and with the primers listed in Supplementary Methods (Supplementary information).. The resulting PCR products were then inserted by restriction/ligation cloning in the XbaI and NotI sites in the pCDH-Blast vector. Resulting plasmids were sequence verified.

A375 cells were first transfected with pCMV(CAT)T7-SB100 (a gift from Zsuzsanna Izsvak, Addgene plasmid # 34879) and pSBbi-pur H-2K^b (a gift from Jon Yewdell, Addgene plasmid # 111623) in a 1:10 ratio, using PEI as a transfection reagent. Five days after transfection, cells with stable H2-K^b expression were selected with puromycin. Next, the H2-K^b expressing A375 were transduced with lentiviruses generated from the pCDH-V5-ATF4$^{(1-63)}$-tGFP-SIINFEKL constructs, after which they were selected with blasticidin.

For the detection of presented SIINFEKL peptides, cells were treated for two days with the indicated treatments. Then the cells were washed with PBS and detached using Versene solution (Gibco). Next, cells were washed in PBS-BSA (0.1%) and incubated with APC anti-mouse H-2K^b bound to SIINFEKL Antibody (Biolegend, clone 25-D1.16, #141606; 1:200)

for 30 minutes. Next, the cells were washed three times and analyzed on a BD LSR Fortessa (BD Biosciences). The data were analyzed using FlowJo V10 software (FlowJo).

Aberrant peptides screened for immunogenicity

Gene	Aberrant peptide sequence	HLA restriction	# of screened healthy donors
HSP90AB1	MVSPLAGVPK	A*03:01	5
STK25	SPALRTLTL	B*07:02, C*07:02	4, 4
MAPK12	SPAAGGQGL	B*07:02	4
RPL7A	VAAAESHPL	B*07:02	4
KCNK6	VPLSLAEHAL	B*07:02	4
GBA	APASLKASA	B*07:02	4
FCGBP	APSGVAAGL	B*07:02	4
ZNF513	VGQEGLVSL	C*07:02	4
ESPNL	LFLSHLEEI	C*07:02	4

Induction of T cells reactive to aberrant peptides

Peripheral blood mononuclear cells (PBMCs) were isolated either from buffy-coats from healthy anonymous blood donors provided by Oslo University Hospital Blood Bank, or from blood processed from HLA-typed in-house healthy donors. The study was approved by the Regional Ethics Committee (REC) and informed consent was obtained from healthy donors in accordance with the declaration of Helsinki and institutional guidelines (REC # 2018/2006 and # 2018/879). Isolation of T cells reactive to aberrant peptides was performed as previously described [42], with modifications. Briefly, on day -4 monocytes were isolated from PBMCs of HLA-A*03:01[pos], HLA-B*07:02[pos] and HLA-C*07:02[pos] healthy donors using CD14-reactive microbeads and AutoMACS Pro Separator (Miltenyi Biotec), and cultured for three days in CellGro GMP DC medium (CellGenix) supplemented with 1% (vol/vol) human serum (HS, Trina biotech) and 1% (vol/vol) P/S containing 10 ng/ml Interleukin (IL)-4 (PeproTec) and 800 IU/ml GM-CSF (Genzyme). Subsequently, monocyte-derived-dendritic cells (MoDCs) were matured for 14-16 h by supplementing cultures with 800 IU/ml GM-CSF, 10 ng/ml IL-4, 10 ng/ml Lipopolysaccharide (LPS; Sigma-Aldrich) and 5 ng/mL IFN-γ (PeproTech). On day -1 autologous naïve CD8+ T cells were isolated using CD8+ T cell isolation kit and AutoMACS Pro Separator (Miltenyi Biotec). Naïve CD8+ T cells were

cultured overnight in CellGro GMP DC medium supplemented with 5% human serum (CC medium) and 5 ng/ml IL-7 (PeproTech). On day 0, MoDCs were peptide-pulsed (list of peptides found in supplementary information) for 2 h at a concentration of 1 μg/mL for individual transpeptides, or incubated with DMSO vehicle. A total of 9 different transpeptides with different HLA restrictions were screened. After harvesting, MoDCs were co-cultured with naïve T cells in CellGro GMP DC medium supplemented with 5% human serum and 30 ng/ml IL-21 (PeproTech) at a DC:T cell ratio of 1:2. In parallel control cultures, naïve T cells were co-cultured with DMSO vehicle-treated MoDCs. On days 3, 5 and 7 half of the medium was removed and replaced with fresh CC medium supplemented with 10 ng/ml of both IL-7 and IL-15 (PeproTech). On day 10, co-cultures were screened for the presence of transpeptide pMHC multimer-reactive CD8+ T cells. pMHC multimers conjugated to four different streptavidin (SA)-fluorochrome conjugates were prepared in-house as previously described [74,75]: SA-phycoerythrin (SA-PE), SA-phycoerythrin-Cy7 (SA-PE-Cy7), SA-Allophycocyanin (SA-APC) and SA-Brilliant Violet 421 (SA-BV421). Each pMHC multimer was labeled with two different fluorochromes for increased specificity and a total of 70 fluorochrome-labeled pMHC multimers were prepared for analysis. Positive T cells were identified by Boolean gating strategy in FlowJo (TreeStar) V.10.6.2 software as live CD8+ T cells staining positively in two pMHC multimer channels and negatively in three other pMHC multimer channels, as previously described [76]. Double pMHC multimer positive T cells were sorted by fluorescence activated cell sorting (FACS) using PE- and APC-conjugated pMHC multimers.

Cloning of CD8+ T cells reactive to aberrant peptides
Single cell cloning of CD8+ T cells reactive to MHC multimers complexed with KCNK6-derived or ZNF513-derived aberrant peptides was performed as previously described [42]. Briefly, for feeder preparation, PBMCs from three different donors were mixed in 1:1:1 ratio, irradiated with 35 Gy, washed and re-suspended in X-vivo 20 medium (BioNordika) supplemented with 5% (vol/vol) HS and 1% (vol/vol) P/S (T-cell cloning medium). Feeder cells were added to 96-well tissue culture treated plates (0.2 x 10^6 cells/well in a volume of 100 μL) and incubated overnight at 37 °C with 5% CO_2. On the day of sorting, 100 μL of T-cell cloning medium containing 4 ng/ml IL-15 (PeproTech), 2 μg/ml phytohaemagglutinin (PHA; Remel Thermo Scientific) and 200 U/ml IL-2 (R&D Systems) was added to the feeder cells and double pMHC multimer+ live cells were sorted as single cells onto the feeder cells using SH800 cell sorter (Sony Biotechnology). A total of 180 pMHC multimer+ single cells

were sorted for both KCNK6 and ZNF513-reactive CD8+ T cells. Every 7 days, cultures were supplied with fresh T-cell cloning medium containing IL-15 and IL-2, and expanding clones were identified by microscopic observation. Day 14 after sorting, growing clones were harvested and re-stimulated with a freshly prepared feeder mix in T-cell cloning medium, as described above. Double pMHC multimer staining and T cell activation assay (below) were performed to confirm presence of specific and functional T cells.

Assessment of T cell specificity

Reactivity of T cell clones was investigated by measuring up-regulation of the activation marker CD137 upon stimulation of 50,000 T cells/well with 100,000 target cells /well, using 1-3 parallels/condition. Target cells (K562 transduced to express relevant HLA allele) were pulsed with indicated concentrations of peptide for 2 h, washed and co-cultured with effector cells. Following 16 h of co-incubation cells were stained to measure up-regulation of CD137 on live CD8+ T cells, measured by flow cytometry. T cells transduced with an HSV-2 HLA-B*07:02-restricted TCR [43] (TCR epitope: RPRGEVRFL) were used as a positive control in functionality assay with target cells loaded with relevant HSV-2 peptide and for pMHC multimer labelling when stained with relevant pMHC multimer. Results are shown as the percentage of live CD137+/CD8+ cells.

Antibodies and flow cytometry – T cell assays

Flow cytometry was performed on a BD LSR II flow cytometer (BD Biosciences), and data were analyzed using FlowJo (TreeStar) V.10.6.2 software. For surface staining, cells were incubated with antibodies for 15-20 min on ice followed by washing steps. The following antibodies were used: anti-human CD8-FITC (Biolegend, clone RPA-T8, cat no. 301050, 1:200), anti-human CD8a-BV421 (Biolegend, clone RPA-T8, cat no. 301036, 1:200) and anti-human CD137-Alexa Fluor 647 (Thermo Fisher Scientific, clone 4B4-1, cat no. A51019, 1:100). Live/Dead Fixable Near-IR Dead Cell Stain kit (Life Technologies) was used to exclude dead cells in all flow cytometry experiments.

Data availability

Data were deposited in GEO with accession code GSE142822. The genomic data relevant to the MD55A3 are found in BioProject ID PRJNA316754, identified 1M1. Proteomics and peptidomics data were deposited in PRIDE [77] with accession code PXD020224. The codes used in the study are available on GITHUB with these links;

https://github.com/apataskar/bump_finder_example2,
https://github.com/apataskar/accessory_scripts_manuscript

References for the Method section

43   Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47**, D766-D773, doi:10.1093/nar/gky955 (2019).

44   Langmead, B. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* **Chapter 11**, Unit 11 17, doi:10.1002/0471250953.bi1107s32 (2010).

45   Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).

46   Lauria, F. *et al.* riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* **14**, e1006169, doi:10.1371/journal.pcbi.1006169 (2018).

47   Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

48   Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).

49   Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).

50   Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

51   Rainer, J., Gatto, L. & Weichenberger, C. X. ensembldb: an R package to create and use Ensembl-based annotation resources. *Bioinformatics* **35**, 3151-3153, doi:10.1093/bioinformatics/btz031 (2019).

52   Jersie-Christensen, R. R., Sultan, A. & Olsen, J. V. Simple and Reproducible Sample Preparation for Single-Shot Phosphoproteomics with High Sensitivity. *Methods Mol Biol* **1355**, 251-260, doi:10.1007/978-1-4939-3049-4_17 (2016).

53   Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

54   Ameziane, N. *et al.* A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. *Nat Commun* **6**, 8829, doi:10.1038/ncomms9829 (2015).

55   Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res* **46**, D754-D761, doi:10.1093/nar/gkx1098 (2018).

56   Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).

57   UniProt, C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506-D515, doi:10.1093/nar/gky1049 (2019).

58   Zhang, X. *et al.* Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc* **13**, 530-550, doi:10.1038/nprot.2017.147 (2018).

59   Kalaora, S. *et al.* Use of HLA peptidomics and whole exome sequencing to identify human immunogenic neo-antigens. *Oncotarget* **7**, 5110-5117, doi:10.18632/oncotarget.6960 (2016).

60   Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152, doi:10.1093/bioinformatics/bts565 (2012).

61   Jurtz, V. *et al.* NetMHCpan 4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J Immunol* **199**, 3360-3368, doi:10.4049/jimmunol.1700893 (2017).

62   Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-11 10 33, doi:10.1002/0471250953.bi1110s43 (2013).

63   Li, H. D., Funk, C. C. & Price, N. D. iREAD: a tool for intron retention detection from RNA-seq data. *BMC Genomics* **21**, 128, doi:10.1186/s12864-020-6541-0 (2020).

64   Andreatta, M., Alvarez, B. & Nielsen, M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res* **45**, W458-W463, doi:10.1093/nar/gkx248 (2017).

65   Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* **47**, D339-D343, doi:10.1093/nar/gky1006 (2019).

66   Krokhin, O. V. & Spicer, V. Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides. *Anal Chem* **81**, 9522-9530, doi:10.1021/ac9016693 (2009).

67   Gatto, L. & Lilley, K. S. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288-289, doi:10.1093/bioinformatics/btr645 (2012).

68   MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968, doi:10.1093/bioinformatics/btq054 (2010).

69   Meszaros, B., Erdos, G. & Dosztanyi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **46**, W329-W337, doi:10.1093/nar/gky384 (2018).

70   Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* **7**, 1534-1550, doi:10.1038/nprot.2012.086 (2012).

71   Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nat Med* **12**, 246-251, doi:10.1038/nm1360 (2006).

72   Hadrup, S. R. *et al.* Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods* **6**, 520-526, doi:10.1038/nmeth.1345 (2009).

73   Philips, D., van den Braber, M., Schumacher, T. N. & Kvistborg, P. pMHC Multiplexing Strategy to Detect High Numbers of T Cell Responses in Parallel. *Methods Mol Biol* **1514**, 93-101, doi:10.1007/978-1-4939-6548-9_7 (2017).

74   Dong, L., Li, P., Oenema, T., McClurkan, C. L. & Koelle, D. M. Public TCR use by herpes simplex virus-2-specific human CD8 CTLs. *J Immunol* **184**, 3063-3071, doi:10.4049/jimmunol.0903622 (2010).

75   Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450, doi:10.1093/nar/gky1106 (2019).

Author Contributions

OB, AP, RN, conceived the project, designed and performed experiments, analyzed data, wrote the manuscript; RA and YS conceived the project, designed experiments, wrote the manuscript, and supervised the project; JO and ML designed experiments for aberrant peptide immunogenicity testing and wrote the manuscript, and ML performed immunogenicity testing of aberrant peptides; WY produced monomers; MMN optimized combinatorial tetramer staining and contributed to HLA typing and blood processing of in-house healthy PBMC donors; RL generated the frame-shift sequence database for peptidomics data. MA, DH and EG performed the bioinformatics analyses of the peptidomics data; TG and DH performed the correlation analysis between the synthetic and endogenous spectra; SL assisted in peptide validation; PRK performed the Diricore analysis of ribosome profiling data; IK, SL performed cloning of reporters, western blot and flow cytometry analyses; JC performed the tGFP western blot; EZ, CB and AB, performed the metabolomics analyses; OBB, MA,

37

YL, mass spectrometry; XH, JK, DSP provided reagents and technical assistance for the T cell co-culture assays; JW helped in generating melanoma cell lines; OM and NSG assisted in performing ribosome profiling. All authors read and approved the manuscript.

Conflict of interest

The authors declare no potential conflicts of interest.

Corresponding authors

Yardena Samuels

Email Yardena.Samuels@weizmann.ac.il

Reuven Agami

Email r.agami@nki.nl

Extended Figure legends

Extended Data Figure 1: Diricore analysis characterizes ribosome occupancy changes following IFNγ and tryptophan depletion treatments of melanoma cells

(a) Western blot analysis for IDO1 expression in control and IFNγ treated conditions in three different melanoma cell lines as indicated (n=3). (b) Tryptophan (left) and Kynurenine (right) levels as determined by mass spectrometry of the indicated melanoma cell lines in control conditions or 48h after the start of IFNγ treatment. Data represent averages of three independent experiments +/- SD. (c) Schematic depicting the principle behind Diricore analysis: E, P and A-site occupancy of the ribosome can be mapped on a typical Ribosomal Protected Fragment (RPF) at 9th, 12th and 15th position, respectively. This allows to perform a position specific subsequence analysis (bottom left panel) to probe for the codon enriched differentially between two conditions at E, P or A sites (presented as bar-plots). Additionally, Diricore analysis depicts 5'RPF densities across a codon of interest (bottom right panel, presented as line-plots). (d) Metagene density profiles depicting global shifts of RPFs to the start of the coding sequence upon IFNγ treatment (red, yellow) as compared to control (grey). Y-axis: intra-gene normalized RPF density. Ribosome profiling data of two independent biological replicates for MD55A3 (upper panel) and 108T (lower panel) are represented. (e) Diricore analysis line plots depicting differential ribosome occupancy (5'RPF) at -30 to +60 codons in MD55A3 (left) and 108T (right) cells. Plots depict 5'RPF densities after the start ATG codon (upper panels), tryptophan codon (middle panels) and cysteine codon (lower panels). Y-axis: Ratio between number of reads in IFNγ versus control conditions. (f) Diricore analysis bar-plots depicting differential codon usage (at position 15 of the RPFs) in IFNγ versus the control condition for MD55A3 (top panel) and 108T cells (lower panel). Data represent the average of two independent biological replicates. (g) Diricore line-plots depicting cumulative signal of RPFs across the coding region normalized into percentiles for ATF4, CDC6 and ATP5G1 in 12T control (grey tones) and IFNγ-treated cells (yellow, red). W-Bumps (greyed areas) indicate an increased number of reads downstream of the tryptophan codons (dashed lines). (h) Tryptophan levels as measured by mass spectrometry from 12T cells subjected to the indicated treatments. Bars represent the average of three independent experiments, +/- SD. (i) UCSC tracks representing IDO1 mRNA reads in MD55A3 cells in the indicated conditions as a measure for the induction of transcription of this gene. (j) Metagene RPF density profiles for control (black), IFNγ (red), IDOi (grey) and IFNγ+IDOi (green) treated 12T cells (k) Diricore analysis bar-plots depicting differential

codon usage (at position 15 of the RPFs) in IFNγ versus control-treated 12T cells. The graph represents the average of two independent biological replicates. This experiment is the control of 1f. (l) Tryptophan levels as measured by mass spectrometry in MD55A3 cells following 48h growth in control or tryptophan-depleted media. Values represent the average of three independent replicates +/- SD. (m) Metagene density profiles for 12T control (dark grey) and tryptophan depletion cells (red). The lines represent the average of two independent biological replicates. (n) Diricore analysis line-plots depicting differential (IFNγ/Control) ribosome occupancy (5'RPF) at -30 to +30 codons of the indicated amino acid in control versus tryptophan depletion 12T cells. (o) Diricore analysis bar-plots depicting differential codon usage (at position 15 of the RPFs) in control and tryptophan-depleted 12T cells. All Diricore plots represent the average of at least two biological replicates.

Extended Data Figure 2: Bump-finder identifies tryptophan-associated bumps

(a) Computational approach for unbiased detection of Bumps in comparative ribosome profiling experiments. The algorithm scans transcripts in 100 windows of equal length for peaks in ribosome occupancy and filters for differential peaks in treated versus untreated samples. (b-c) Density of codons per amino acid in the region of 60 codons upstream and downstream of the peak of bumps identified with bump-finder in ribosome profiling data of 12T cells. Data shown is derived from two independent biological replicates treated with either IFNγ (panel b) or control (panel c). (d) Ratio between upstream and downstream reads 30 codons from the peak of bumps identified in control conditions when treated with IFNγ in MD55A3 (top) and 108T cells (bottom). (e) RPF density in control (grey line) and IFNγ (red line) treated conditions in MD55A3 and 108T cells. The area marked in grey indicates the W-Bump region. (f) Densities (upper lines) and heatmaps (lower panel) of ribosomal P-sites 100 nucleotides surrounding every tryptophan codon in control and IFNγ conditions in 12T, MD55A3 and 108T cells. (g) Classification of all transcripts containing tryptophan codons, one group associated with W-Bumps ('Bumps'), another which is not associated with bumps ('No-bumps'). Graphs represent RPF density in the region of 300 nucleotides surrounding the tryptophan codon. Bumps indicated in grey shading. (h) Bar plot depicting the enrichment of presence of two tryptophan codons within a region of eight codons in the 'Bumps'-group over the 'No-Bumps'-group. W indicates a codon for tryptophan, whereas X indicates all remaining amino acids. (i) Heatmap depicting frequencies of codons for each amino-acid at every position 25 codons upstream and downstream, with respect to tryptophan. (j) Line plots

depicting RPF density at tryptophan stratified by distance between two tryptophan (black). In each graph the control (the occurrence of a single tryptophan) is presented in red.

Extended Data Figure 3: Proteomics analyses following IFNγ treatment of MD55A3 melanoma cells.

(a) Volcano plot depicting overall changes in the proteome upon IFNγ treatment as observed by analysis of quantitative mass-spectrometry data. X-axis: Log fold change between IFNγ versus control conditions in MD55A3 cells. Y-axis: Corresponding log transformed adjusted p-value calculated from three independent biological replicates. Highlighted in blue, are proteins that are significantly differentially expressed. Both IFNγ-mediated induction of IDO1 and WARS and the immunoproteasome components are indicated. (b) Left: Boxplots depicting log fold change in the levels of protein (red; average of three replicates) and mRNA (brown; average of two replicates) in IFNγ versus control treated cells. Proteins were grouped according to the number of asparagines (N), tyrosine (Y) or phenylalanine (F) in the protein sequence. Boxes depict first, second and third quartiles while the whiskers depict the range excluding the outliers. Test: Wilcoxon Test (two-tailed); ns = not significant, * $p < 0.05$. For asparagine box plots the actual p-values are 0.7, 0.8, 0.055 for protein quantification (left plot) and 0.8, 0.2, 0.013 for RNA quantifications (right plot) in the order of depiction. Mid: Same as Left but for number of Tyrosine (Y). Actual p-values are 0.7, 0.86 and 0.02. for protein quantification (left plot) and 0.98, 0.13 and 0.14 for RNA quantifications (right plot) in the order of depiction. Right: Same as mid but for number of Phenylalanine (F). Actual p-values are 0.7, 0.97 and 0.23 for protein quantification (left plot) and 0.3, 0.015 and 0.0038 for RNA quantifications (right plot) in the order of depiction. (c) Western blot analysis of ubiquitinylated proteins in total cell lysates of MD55A3 cells mock- or IFNγ-treated, which were additionally incubated minus or plus MG132 (n=1). (d) A panel showing proteins with increased abundance in total cell lysates of MD55A3 cells treated with MG132 versus controls. Log2 fold changes were calculated on data of three independent replicates. (e) Boxplots depicting log fold change in protein levels (average of three replicates) in IFNγ-treated versus control conditions in MD55A3 cells treated with proteasome inhibitor, for proteins stratified for different numbers of asparagines (N), tyrosines (Y) and phenylalanine (F) in their sequence. Boxes depict first, second and third quartiles while the whiskers depict the range excluding the outliers. Test: Wilcoxon Test (two-tailed); ns = not significant and * $p < 0.05$. Actual p-values are 0.34, 0.44, 0.77 and 0.96, 0.84, 0.024 and 0.29, 0.84, 0.34 in the

order of depiction. (f) Boxplots depicting protein length (number of amino-acids) for stratified group of genes with increasing number of tryptophan (left) and asparagine (mid), with their ratios (right). (g) Density of RPFs (average of two replicates) 300 codons across individual tryptophan codons (black line), two tryptophans that are present within a distance of 8 codons (green line) and two tryptophans that are present at a distance greater than 8 codons (red line). (h) A boxplot depicting bump scores (average of two replicates) for instances of two tryptophan separated by fewer than 8 codons (green) and more than 8 codons (red). Bump scores are calculated in MD55A3 cells. Boxes depict first, second and third quartiles while the whiskers depict the range excluding the outliers. Test: Wilcoxon Test (two-tailed); *** $p < 0.0005$. Actual p-value is $< 2.2e^{-16}$. (i) A boxplot depicting protein level changes (log fold change, average of three replicates) between IFNγ and control conditions. The graph represents genes which have two tryptophan codons within a distance of 8 codons (green) or genes having a distance of more than 8 codons between two tryptophans (red). Test: Wilcoxon Test (two-tailed); *** $p < 0.0005$. Actual p-value is $< 2.e^{-16}$ (j) Same as (i), but for asparagine (N), in MD55A3 cells. Test: Wilcoxon Test (two-tailed); ns = not significant. Actual p-value is 0.83.

Extended Data Figure 4: Reporter assays for the detection of tryptophan-associated *out-of-frame* events.

(a) A hypothetical model suggesting a possible mechanism causative for W-Bumps. In the normal scenario, ribosomes do not encounter problems when translating a tryptophan and progress translation at regular speed (top panels). Tryptophan shortage, on the other hand, can lead to stalling on the tryptophan codon (bottom left panel), or could in theory induce frameshifting events, leading to aberrant peptide production (lower right panel). Since secondary structure of growing polypeptide chains is attained in the lower tunnel of the ribosome, the loss of an α-helical secondary structure in this tunnel could hamper ribosomal progression. (b) A boxplot depicting bump-score (from two replicates) in MD55A3 cells from the group of selected peptides with ordered *out-of-frame* peptides ('Selected', from Fig. 3b) and every tryptophan in the proteome ('All' from Fig. 3b). Boxes depict first, second and third quartiles while the whiskers depict the range excluding the outliers. Test: Two-sample t-test; * $p = 0.056$. (c) Schematic representation of the final protein sequences that would form due to frameshifting events. The *in-frame* construct (upper panel) contains a His-tag and would end up in the pulldown (PD) fraction. Whenever a frameshift occurs at the position of the tryptophan in ATF4, this protein would lose its His-tag, and consequently would end up

in the supernatant fraction (S). The +1 and +2 *out-of-frame* constructs (bottom panel) do not contain a His-tag, whereby the resulting proteins always end up in the supernatant fraction (S) in a His-tag pulldown assay. When frameshifting events take place, however, the His-tag is incorporated into the peptide, whereby the resulting protein ends up in the pulldown fraction (PD). (d) Western blot analysis of a sequential V5-tag immunoprecipitation on the supernatant samples of the Frame reporter expressing cells from Fig. 3e (n=1). The image indicates the presence of V5-tag containing peptides that do not contain a His-tag generated from the *in-frame* reporter. (e) Tryptophan levels in MD55A3 cells expressing the V5-ATF4-His reporter constructs that were used in Fig. 3e and Extended Data Fig. 4d. Tryptophan levels were analyzed by mass spectrometry after 48 hours of treatment. Bars represent the average of three independent replicates +/- SD. (f) Tryptophan levels in MD55A3 cells expressing the V5-ATF4-His reporter constructs that were used in Extended Data Fig. 4h. Tryptophan levels were analyzed by mass spectrometry after 24 and 48 hours of treatment in triplicate. (g) Western blot analysis showing V5-tagged peptides in pulldown samples of MD55A3 reporter cells, either mock, IFNγ, or IFNγ+IDOi treated (n=2). (h) Western blot analysis of V5-tagged proteins in pulldown samples of MD55A3 reporter cells that were either mock-treated or cultured in tryptophan-depleted medium (n=1). (i) Western blot analyses showing pulldown assays followed by V5-staining on western blot of MD55A3 cells expressing the original reporter constructs as depicted in Fig. 3d (Wt), or MD55A3 cells expressing the same reporters where the tryptophan codon was mutated to a codon for tyrosine (Y mut, n=1). (j) Amino acid levels as determined by mass spectrometry from lysates of Tyr-depleted cells (48 hours) versus control cells. Bars represent the average of three independent replicates +/- SD. (k) Western blot analysis showing V5-tagged peptides in pulldown samples of MD55A3 reporter cells that were mock-treated or depleted of tyrosine for 48 hours (n=1). (l) Flow cytometry analyses showing the histograms obtained of MD55A3 cells expressing the V5-ATF4-tGFP reporters in all 3 frames in mock and IFNγ-treated conditions. Plots are a representative graph out of a triplicate biological experiment. (m) Flow cytometry analyses showing the quantification of histograms obtained in (l). Bars represent the average of three independent replicates +/- SD, * p-values in order from left to right: 4.0×10^6 and 9.5×10^4 as determined by a two-sided T-Test (n) Anti-V5-tag and anti-tGFP western blot analysis of whole cell lysates from MD55A3 cells expressing the indicated reporters, which were subjected to mock or IFNγ treatments. In each blot the position of the full-length *in-frame* protein and the shorter *out-of-frame* protein are marked by the arrowheads (n=2). (o) Western blot analyses with anti-V5 antibody of His-tag pulldown

samples of 888-Mel and D10 cells expressing the *in-frame* and +1 reporters (n=2). The cells were either grown in isolation (-), or co-cultured with MART-1 specific T cells for 16 hours (+) before the pulldown was performed. (p) Western blot analysis showing IDO1 upregulation in 888-Mel and D10 cells in control and T-cell co-culture conditions (n=2). The same cells were used for pulldown experiments in Fig. 3m. (q) His-tag pull-down was executed on the lysates of mock and IFNγ-treated 888-Mel and D10 cells expressing the *in-frame* and *+1* reporters (n=2). Both supernatant (S) and His-tag pull down samples (PD) of these cells were stained with V5 antibodies. (r) The lysates of cells used in panel q were used for a western blot analysis to show the level of IDO1 induction (n=1). (s) Tryptophan levels as determined by mass spectrometry analysis of 888-Mel and D10 lysates in mock, IFNγ-treated and tryptophan-depleted conditions. Bars represent the average of three independent replicates +/- SD, except for the first bar, which is an average of two independent replicates. (t) Western blot analyses with anti-V5 antibody of His-tag pulldown samples of 888-Mel and D10 that were mock-treated (+), or grown in tryptophan-less medium for 48 hours (-) before the pulldown was performed (n=2).

Extended Data Figure 5: Proteomics analysis of 2D LC-MS/MS data reveals the induction of endogenous aberrant peptides following IFNγ treatment of MD55A3 melanoma cells.

(a) PCA plot illustrating the clustering of 2D LC-MS/MS data from MD55A3 cells that were treated with IFNγ for 48 hours as compared to mock. Both conditions were treated with MG-132 for the last 4 hours. (b) Gene set enrichment analysis (GSEA) based depiction of the proteomics data showing the induction of the IFN response genes following IFNγ treatment of MD55A3 cells. (c) Schematic representation of the *in silico* generation of +1 and -1 frameshifted protein database (see methods). (d) Bar-plot depicting number of hits for the detected aberrant (frame-shifted) polypeptides in the Ensemble translations database of coding and non-coding genes. Only the aberrant peptides without any matches (n=0) were retained as true hits (red bar). (e) Heatmap showing frameshifted peptides identified in the full proteome of IFN-treated MD55A3 cells along with sequence information. (f). Line plot depicting log scaled fold change in proteomic intensity values of IFNγ versus mock-treated MD55A3 cells. (g) Same as panel f, but for LFQ normalized intensity values of frame-shifted peptides hits along with reverse peptide hits. (h) Same as panel f, but only for reverse peptide hits. (i) Same as panel f, but only for frame-shifted peptide hits. (j) Gene set enrichment analysis (GSEA) based depiction of the proteomics data showing the induction of the HLA

44

genes following IFNγ treatment of MD55A3 cells. (k) GO term analysis on the differentially expressed proteins detected in panel j. (l) Anti-V5-tag and anti-tGFP western blot analysis of whole cell lysates from A375 cells expressing the indicated reporters, which were subjected to mock treatment or tryptophan depletion (n=2). In each blot the position of the full-length *in-frame* protein and the shorter *out-of-frame* protein are marked by the arrowheads. (m) FACS plots representing the signal for H2-K^b-bound SIINFEKL peptides in A375 cells expressing H2-K^b in combination with *in-frame* (Frame) or +1 out-of-frame (+1) V5-ATF4[(1-63)]-tGFP-SIINFEKL reporters. Graphs shown are a representative curve from one of three independent experiments.

Extended Data Figure 6: Quality control analysis for the immunopeptidomics of MD55A3 cells. (a) Length distribution of the identified peptides. (b) HLA binding prediction using netMHCpan shows a high percentage of eluted peptides predicted to bind MD55A3 haplotypes. (c) HLA-I peptides were clustered by Gibbs clustering to assign the peptides to the different HLA alleles of the patient. Logos were created to identify the HLA alleles' motif using all peptides matching to this allele in the IEDB database (top panel). Peptides that were identified in the cells treated with IFNγ, tryptophan derived (mTRP) and non-treated (NT) were clustered to 1-6 clusters, and only the best fit (highest Kullbach Leiber Distance) was used for comparison with the IEDB. The clusters show the motifs expected of the cells' HLA alleles. The last column indicates the outlier peptides that were not clustered.

Extended Data Figure 7: Immunopeptidomics analysis identifies tryptophan-associated aberrant peptides.

(a) Scheme summarizing the filtration steps that led to the detection of HLA-bound aberrant peptides, containing W-associated frame-shift sequences (derived from 16 samples: 4 NT, 4 +IFNγ, 4 mTpr, 4 metastases from patient 55); Immunopeptidomics spectra were searched against W-specific frame-shifted human proteome. MS-identified peptides were filtered according to described steps, including HLA binding prediction, spectral quality and source gene expression validation in transcriptome/translatome datasets. Following the above steps, 28 aberrant peptides were detected either in MD55 metastasis or treated MD55A3 samples, and not in the untreated cells. (b) Predicted hydrophobicity index (HI) and observed retention time (RT) of canonical peptides (black) and aberrant peptides at two filtration steps described

in a, step 3 (blue) and step 7 (red). The observed correlation between the RT versus HI, (using the SSRCalc tool [31,69]) supports the identification of the aberrant peptides presented in a. (c) Schematic representation of aberrant peptides KCNK6 and PCNX2 detected via immunopeptidomics, derived from a -1 and +1 frameshift at a Trp codon (marked in red), respectively. The in-frame and out of frame amino acid sequences are indicated above (black) and below (red/ green) the mRNA sequence, respectively. (d) Venn diagram including the 28 aberrant peptides detected following the filtration steps presented in a (were not identified in the untreated MD55A3 cells), grouped according to the samples they were detected in. (e) Mean LFQ normalized intensities. Canonical and aberrant peptides that exhibit above three-fold change in their intensities, are marked in black and blue/red, respectively. (f-g) Similarities between the aberrant (n=28) and canonical (n=6,304) bound peptides. f: Andromeda score. g: Binding affinity prediction with NetMHCpan showing a similar binding distribution of the canonical and aberrant peptides. The aberrant peptides are highlighted in red. Dashed line mark strong binders rank cutoff (<0.5). (h) a list of eight out of the thirteen -1 and +1 OOF and trans-frame peptides found in the metastasis samples of patient MD55. The colors of the peptide sequences represent in-frame (black), +1 frame (green) and -1 frame (red).

Extended Data Figure 8: Tandem mass spectra of endogenous aberrant peptides identified in patient MD55 match to the spectra identified in patient 42 metastasis.

Left – experimental spectra of peptides identified in patient 42 metastasis samples who shares three of six HLA class I alleles with patient 55 (A03:01; B07:02; C07:02); right - head-to-tail plots show similarities between overlapping ions of the two spectra, Upper panel (blue) – patient 42; lower plane (red)– MD55.

Extended Data Figure 9: Identification of T cells reactive to tryptophan-derived aberrant peptides.

(a) Flow cytometric analysis of CD8+ T cells reactive to BV421 and PE-labeled pMHC multimers complexed with the ZNF513-derived aberrant peptide in co-cultures of naïve CD8+ T cells and autologous MoDCs pulsed with peptide (right) or DMSO vehicle (left). Cells were from an HLA-C*07:02[pos] healthy donor. (b) Boolean gating strategy for analysis of combinatorial pMHC multimer staining limits false positive signals arising due to background fluorescence when combining in one staining multiple different pMHC multimers labeled with different fluorochromes. In total, 4 different fluorochromes were

used in pMHC multimer preparation and combined into 4 dual-color pMHC multimer pairs, each pair complexed with a different peptide. For simplicity, gating for only one dual-color pMHC multimer population is shown. Gating strategy: 1) Live CD8+ T cell singlets were identified with the use of FSC and SSC gates, live/dead fixable near-IR dead cell and CD8 staining → 2) Separate gates were used to define positive events in each pMHC channel → 3) NOT gates were used to select CD8+/pMHC multimer negative events for each pMHC channel → 4) AND gates for two pMHC multimer positive populations and NOT gates for each of the remaining pMHC channels were employed to select for cells that are positive only in two channels. Events staining positively for one or more than two pMHC channels were gated out → 5) AND gate was used to combine all pMHC NOT gates → 6) OR gate was made with gates from step 4 and 5. The final plot shows only CD8+ T cells that are positive for two pMHC channels and excludes cells positive for only one or positive for more than two channels. (c) For analysis of T cells up-regulating the activation marker CD137, viable CD8+ T cell singlets were identified by FSC and SSC gates, live/dead fixable near-IR dead cell staining and CD8 staining. Subsequently, activated CD8+ T cells were identified as CD137+ events, where cut-off was set based on staining of T cells incubated without target cells. Plots depicting CD137+ events show the KCNK6 T cell clone #1 incubated with target cells pulsed with the KCNK6 aberrant peptide (1 μg/mL, upper plot), or not (lower plot).

Uncategorized References

1   Ayers, M. *et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* **127**, 2930-2940, doi:10.1172/JCI91190 (2017).
2   Ji, R. R. *et al.* An immune-active tumor microenvironment favors clinical response to ipilimumab. *Cancer Immunol Immunother* **61**, 1019-1031, doi:10.1007/s00262-011-1172-6 (2012).
3   Timosenko, E. *et al.* Nutritional Stress Induced by Tryptophan-Degrading Enzymes Results in ATF4-Dependent Reprogramming of the Amino Acid Transporter Profile in Tumor Cells. *Cancer Res* **76**, 6193-6204, doi:10.1158/0008-5472.CAN-15-3502 (2016).
4   Zhai, L. *et al.* Molecular Pathways: Targeting IDO1 and Other Tryptophan Dioxygenases for Cancer Immunotherapy. *Clin Cancer Res* **21**, 5427-5433, doi:10.1158/1078-0432.CCR-15-0420 (2015).
5   Amobi, A., Qian, F., Lugade, A. A. & Odunsi, K. Tryptophan Catabolism and Cancer Immunotherapy Targeting IDO Mediated Immune Suppression. *Adv Exp Med Biol* **1036**, 129-144, doi:10.1007/978-3-319-67577-0_9 (2017).
6   Labadie, B. W., Bao, R. & Luke, J. J. Reimagining IDO Pathway Inhibition in Cancer Immunotherapy via Downstream Focus on the Tryptophan-Kynurenine-Aryl

Hydrocarbon Axis. *Clin Cancer Res* **25**, 1462-1471, doi:10.1158/1078-0432.CCR-18-2882 (2019).

7   Gunther, J., Dabritz, J. & Wirthgen, E. Limitations and Off-Target Effects of Tryptophan-Related IDO Inhibitors in Cancer Treatment. *Front Immunol* **10**, 1801, doi:10.3389/fimmu.2019.01801 (2019).

8   Battu, S., Minhas, G., Mishra, A. & Khan, N. Amino Acid Sensing via General Control Nonderepressible-2 Kinase and Immunological Programming. *Front Immunol* **8**, 1719, doi:10.3389/fimmu.2017.01719 (2017).

9   Wek, R. C. & Staschke, K. A. How do tumours adapt to nutrient stress? *EMBO J* **29**, 1946-1947, doi:10.1038/emboj.2010.110 (2010).

10   Ye, J. *et al.* The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *EMBO J* **29**, 2082-2096, doi:10.1038/emboj.2010.81 (2010).

11   Adam, I. *et al.* Upregulation of tryptophanyl-tRNA synthethase adapts human cancer cells to nutritional stress caused by tryptophan degradation. *Oncoimmunology* **7**, e1486353, doi:10.1080/2162402X.2018.1486353 (2018).

12   Loayza-Puch, F. *et al.* Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. *Nature* **530**, 490-494, doi:10.1038/nature16982 (2016).

13   Bhushan, S. *et al.* alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat Struct Mol Biol* **17**, 313-317, doi:10.1038/nsmb.1756 (2010).

14   Woolhead, C. A., McCormick, P. J. & Johnson, A. E. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* **116**, 725-736, doi:10.1016/s0092-8674(04)00169-2 (2004).

15   Caliskan, N. *et al.* Conditional Switch between Frameshifting Regimes upon Translation of dnaX mRNA. *Mol Cell* **66**, 558-567 e554, doi:10.1016/j.molcel.2017.04.023 (2017).

16   Yelverton, E., Lindsley, D., Yamauchi, P. & Gallant, J. A. The function of a ribosomal frameshifting signal from human immunodeficiency virus-1 in Escherichia coli. *Mol Microbiol* **11**, 303-313, doi:10.1111/j.1365-2958.1994.tb00310.x (1994).

17   Gurvich, O. L., Baranov, P. V., Gesteland, R. F. & Atkins, J. F. Expression levels influence ribosomal frameshifting at the tandem rare arginine codons AGG_AGG and AGA_AGA in Escherichia coli. *J Bacteriol* **187**, 4023-4032, doi:10.1128/JB.187.12.4023-4032.2005 (2005).

18   Olubajo, B. & Taylor, E. W. A -1 frameshift in the HIV-1 env gene is enhanced by arginine deficiency via a hungry codon mechanism. *Mutat Res* **579**, 125-132, doi:10.1016/j.mrfmmm.2005.02.018 (2005).

19   Barak, Z., Lindsley, D. & Gallant, J. On the mechanism of leftward frameshifting at several hungry codons. *J Mol Biol* **256**, 676-684, doi:10.1006/jmbi.1996.0117 (1996).

20   Laine, S., Thouard, A., Komar, A. A. & Rossignol, J. M. Ribosome can resume the translation in both +1 or -1 frames after encountering an AGA cluster in Escherichia coli. *Gene* **412**, 95-101, doi:10.1016/j.gene.2008.01.018 (2008).

21   Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R. N. & Chrzanowska-Lightowlers, Z. M. Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* **327**, 301, doi:10.1126/science.1180674 (2010).

22   Vredevoogd, D. W. *et al.* Augmenting Immunotherapy Impact by Lowering Tumor TNF Cytotoxicity Threshold. *Cell* **178**, 585-599 e515, doi:10.1016/j.cell.2019.06.014 (2019).

23   Zhou, F. Molecular mechanisms of IFN-gamma to up-regulate MHC class I antigen processing and presentation. *Int Rev Immunol* **28**, 239-260, doi:10.1080/08830180902978120 (2009).

24   Bourdetsky, D., Schmelzer, C. E. & Admon, A. The nature and extent of contributions by defective ribosome products to the HLA peptidome. *Proc Natl Acad Sci U S A* **111**, E1591-1599, doi:10.1073/pnas.1321902111 (2014).

25   Yewdell, J. W. DRiPs solidify: progress in understanding endogenous MHC class I antigen processing. *Trends Immunol* **32**, 548-558, doi:10.1016/j.it.2011.08.001 (2011).

26   Trentini, D. B. *et al.* Role for ribosome-associated quality control in sampling proteins for MHC class I-mediated antigen presentation. *Proc Natl Acad Sci U S A* **117**, 4099-4108, doi:10.1073/pnas.1914401117 (2020).

27   Erhard, F. *et al.* Improved Ribo-seq enables identification of cryptic translation events. *Nat Methods* **15**, 363-366, doi:10.1038/nmeth.4631 (2018).

28   Prasad, S., Starck, S. R. & Shastri, N. Presentation of Cryptic Peptides by MHC Class I Is Enhanced by Inflammatory Stimuli. *J Immunol* **197**, 2981-2991, doi:10.4049/jimmunol.1502045 (2016).

29   Starck, S. R. & Shastri, N. Nowhere to hide: unconventional translation yields cryptic peptides for immune surveillance. *Immunol Rev* **272**, 8-16, doi:10.1111/imr.12434 (2016).

30   Chen, J. *et al.* Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140-1146, doi:10.1126/science.aay0262 (2020).

31   Chong, C. *et al.* Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun* **11**, 1293, doi:10.1038/s41467-020-14968-9 (2020).

32   Laumont, C. M. & Perreault, C. Exploiting non-canonical translation to identify new targets for T cell-based cancer immunotherapy. *Cell Mol Life Sci* **75**, 607-621, doi:10.1007/s00018-017-2628-4 (2018).

33   Laumont, C. M. *et al.* Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med* **10**, doi:10.1126/scitranslmed.aau5516 (2018).

34   Pearson, H. *et al.* MHC class I-associated peptides derive from selective regions of the human genome. *J Clin Invest* **126**, 4690-4701, doi:10.1172/JCI88590 (2016).

35   Laumont, C. M. *et al.* Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun* **7**, 10238, doi:10.1038/ncomms10238 (2016).

36   Saulquin, X. *et al.* +1 Frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J Exp Med* **195**, 353-358, doi:10.1084/jem.20011399 (2002).

37   Dersh, D., Yewdell, J. W. & Wei, J. A SIINFEKL-Based System to Measure MHC Class I Antigen Presentation Efficiency and Kinetics. *Methods Mol Biol* **1988**, 109-122, doi:10.1007/978-1-4939-9450-2_9 (2019).

38   McCarthy, M. K. & Weinberg, J. B. The immunoproteasome and viral infection: a complex regulator of inflammation. *Front Microbiol* **6**, 21, doi:10.3389/fmicb.2015.00021 (2015).

39   Goldberg, A. L., Cascio, P., Saric, T. & Rock, K. L. The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides. *Mol Immunol* **39**, 147-164, doi:10.1016/s0161-5890(02)00098-6 (2002).

40   Kalaora, S. *et al.* Combined Analysis of Antigen Presentation and T-cell Recognition Reveals Restricted Immune Responses in Melanoma. *Cancer Discov* **8**, 1366-1375, doi:10.1158/2159-8290.CD-17-1418 (2018).

41   Stronen, E. *et al.* Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337-1341, doi:10.1126/science.aaf2288 (2016).

42   Ali, M. *et al.* Induction of neoantigen-reactive T cells from healthy donors. *Nat Protoc* **14**, 1926-1943, doi:10.1038/s41596-019-0170-6 (2019).

43   Dong, L., Li, P., Oenema, T., McClurkan, C. L. & Koelle, D. M. Public TCR use by herpes simplex virus-2-specific human CD8 CTLs. *J Immunol* **184**, 3063-3071, doi:10.4049/jimmunol.0903622 (2010).

44   Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47**, D766-D773, doi:10.1093/nar/gky955 (2019).

45   Langmead, B. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* **Chapter 11**, Unit 11 17, doi:10.1002/0471250953.bi1107s32 (2010).

46   Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).

47   Lauria, F. *et al.* riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* **14**, e1006169, doi:10.1371/journal.pcbi.1006169 (2018).

48   Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

49   Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).

50   Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).

51   Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

52   Rainer, J., Gatto, L. & Weichenberger, C. X. ensembldb: an R package to create and use Ensembl-based annotation resources. *Bioinformatics* **35**, 3151-3153, doi:10.1093/bioinformatics/btz031 (2019).

53   Jersie-Christensen, R. R., Sultan, A. & Olsen, J. V. Simple and Reproducible Sample Preparation for Single-Shot Phosphoproteomics with High Sensitivity. *Methods Mol Biol* **1355**, 251-260, doi:10.1007/978-1-4939-3049-4_17 (2016).

54   Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

55   Ameziane, N. *et al.* A novel Fanconi anaemia subtype associated with a dominant-negative mutation in RAD51. *Nat Commun* **6**, 8829, doi:10.1038/ncomms9829 (2015).

56   Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res* **46**, D754-D761, doi:10.1093/nar/gkx1098 (2018).

57   Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).

58   UniProt, C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506-D515, doi:10.1093/nar/gky1049 (2019).

59   Zhang, X. *et al.* Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc* **13**, 530-550, doi:10.1038/nprot.2017.147 (2018).

60   Kalaora, S. *et al.* Use of HLA peptidomics and whole exome sequencing to identify human immunogenic neo-antigens. *Oncotarget* **7**, 5110-5117, doi:10.18632/oncotarget.6960 (2016).

61 Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152, doi:10.1093/bioinformatics/bts565 (2012).

62 Jurtz, V. *et al.* NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J Immunol* **199**, 3360-3368, doi:10.4049/jimmunol.1700893 (2017).

63 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-11 10 33, doi:10.1002/0471250953.bi1110s43 (2013).

64 Li, H. D., Funk, C. C. & Price, N. D. iREAD: a tool for intron retention detection from RNA-seq data. *BMC Genomics* **21**, 128, doi:10.1186/s12864-020-6541-0 (2020).

65 Andreatta, M., Alvarez, B. & Nielsen, M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res* **45**, W458-W463, doi:10.1093/nar/gkx248 (2017).

66 Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* **47**, D339-D343, doi:10.1093/nar/gky1006 (2019).

67 Abelin, J. G. *et al.* Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. *Immunity* **46**, 315-326, doi:10.1016/j.immuni.2017.02.007 (2017).

68 Sarkizova, S. *et al.* A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol* **38**, 199-209, doi:10.1038/s41587-019-0322-9 (2020).

69 Krokhin, O. V. & Spicer, V. Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides. *Anal Chem* **81**, 9522-9530, doi:10.1021/ac9016693 (2009).

70 Gatto, L. & Lilley, K. S. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288-289, doi:10.1093/bioinformatics/btr645 (2012).

71 MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968, doi:10.1093/bioinformatics/btq054 (2010).

72 Meszaros, B., Erdos, G. & Dosztanyi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **46**, W329-W337, doi:10.1093/nar/gky384 (2018).

73 Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* **7**, 1534-1550, doi:10.1038/nprot.2012.086 (2012).

74 Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nat Med* **12**, 246-251, doi:10.1038/nm1360 (2006).

75 Hadrup, S. R. *et al.* Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods* **6**, 520-526, doi:10.1038/nmeth.1345 (2009).

76 Philips, D., van den Braber, M., Schumacher, T. N. & Kvistborg, P. pMHC Multiplexing Strategy to Detect High Numbers of T Cell Responses in Parallel. *Methods Mol Biol* **1514**, 93-101, doi:10.1007/978-1-4939-6548-9_7 (2017).

77 Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450, doi:10.1093/nar/gky1106 (2019).