Supplementary Information

TAPBPR Promotes Antigen Loading on MHC-I Molecules Using a Peptide Trap

Andrew C. McShan[1,2†], Christine A. Devlin[3†], Giora I. Morozov[1,2], Sarah A. Overall[1], Danai Moschidi[1], Neha Akella[3], Erik Procko[3*] and Nikolaos G. Sgourakis[1,2,4*]

[1]Department of Chemistry and Biochemistry, University of California Santa Cruz, Santa Cruz, CA 95064, USA.

[2]Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA, 19104, USA

[3]Department of Biochemistry and Cancer Center at Illinois, University of Illinois, Urbana, IL 61801, USA.

[4]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd, Philadelphia, PA, 19104, USA

† These authors contributed equally to this work.

* To whom correspondence should be addressed: Erik Procko (procko@illinois.edu) or Nikolaos G. Sgourakis (Nikolaos.Sgourakis@Pennmedicine.upenn.edu).

Lead contact: Nikolaos.Sgourakis@Pennmedicine.upenn.edu

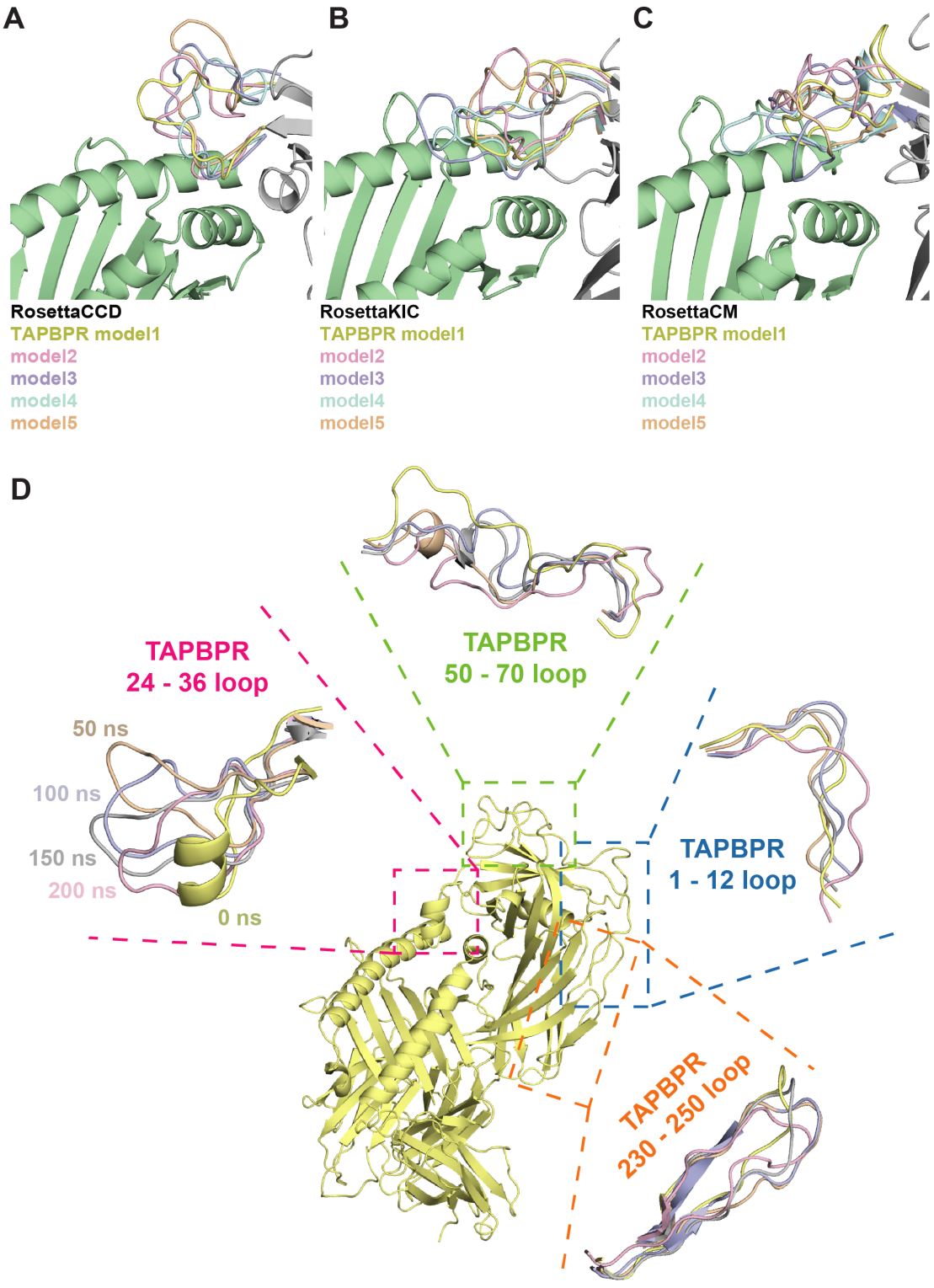

Figure S1. Conformational variation of the TAPBPR G24-R36 loop. (A-C) Different *Rosetta* protocols are used to model the TAPBPR G24-R36 loop: **(A)** Cyclic coordinate descent (CCD), **(B)** kinematic closure (KIC), or **(C)** comparative modeling (CM). The template used was PDB ID 5WER where HLA-A*02:01 sequence was threaded onto H2-D^{d} using the partial_thread

application in *Rosetta*. HLA-A*02:01 is green; TAPBPR is gray. The five lowest energy models of the TAPBPR G24-R36 loop from a total of 1,000 calculated are shown in different colors. **(D)** Model of the peptide-deficient HLA-A*02:01/hβ2m/TAPBPR complex at 0 nsec (start of the MD simulation) is shown in yellow. The dotted boxes highlight different loops present in TAPBPR. The range of conformations sampled by various loops of TAPBPR captured at different times during the MD simulation are shown within each box. Only the TAPBPR loops are highlighted for clarity but the entire peptide-deficient HLA-A*02:01/hβ2m/TAPBPR complex was simulated.

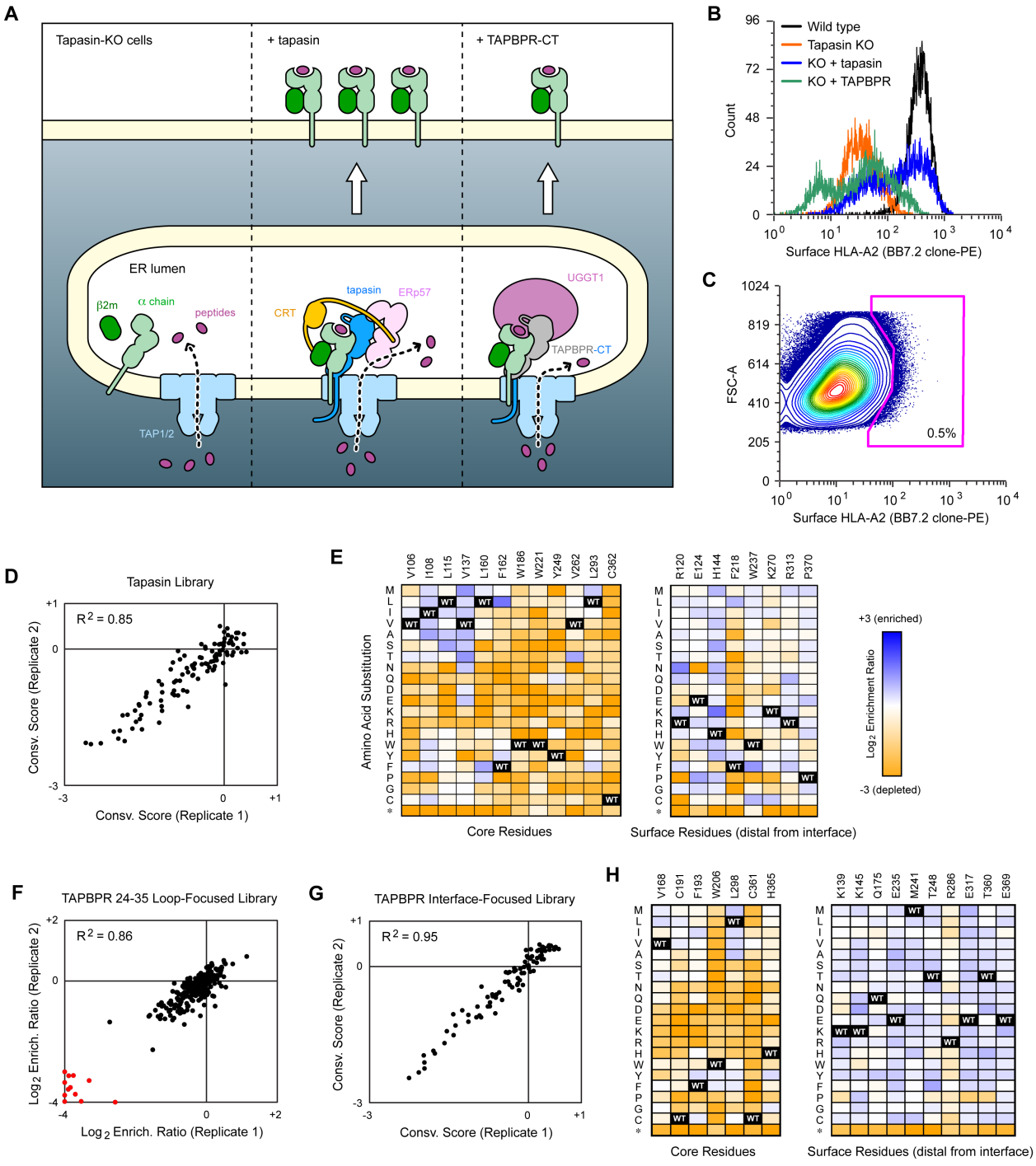

Figure S2. Selections based on functional replacement of endogenous tapasin for HLA-A2 processing and surface trafficking. (A) (*Left*) In tapasin-KO cells, HLA-A2 is poorly processed and loaded, and is retained intracellularly. (*Center*) Expression of tapasin from a plasmid rescues formation of the peptide-loading complex (PLC). HLA-A2 is now efficiently loaded with peptides, folds, and traffics to the plasma membrane. (*Right*) Expression of a chimera between the TAPBPR extracellular domains and the tapasin transmembrane and cytosolic domains partially rescues HLA-A2 processing. We hypothesize that the tapasin C-terminal tail might help bring TAPBPR

and its associated glucosyltransferase UGGT1 in to a PLC-like complex with the TAP1/2 transporter. **(B)** Tapasin was knocked out by targeting indel mutations to exon 2 in Expi293F cells using Cas9/CRISPR-based editing (confirmed by deep sequencing genomic DNA; data deposited in GEO). Surface HLA-A*02:01 expression was substantially reduced in tapasin-KO cells (red) compared to wild type cells (black). Surface trafficking of HLA-A*02:01 was rescued by transfection with a plasmid encoding tapasin (blue), or partially rescued by over-expression of TAPBPR (green). When TAPBPR is over-expressed, a subset of cells also shows increased intracellular retention of HLA-A*02:01 (represented by the left-most peak in green). **(C)** Libraries of tapasin or TAPBPR-CT variants were expressed in tapasin-KO Expi293F cells, and after first gating by scattering and viability, the 0.5 % of cells with the highest levels of endogenous HLA-A2 at the cell surface were collected (magenta gate). Under these transfection conditions, cells typically express no more than a single sequence variant, and most cells are negative. The example shown here is following transfection of a TAPBPR library. **(D)** Conservation scores (calculated from the mean of the $\log_2$ enrichment ratios for each mutation at a residue position) following two independent selections of the tapasin library show close agreement. Tapasin residues with negative scores are tightly conserved for functional rescue of surface HLA-A*02:01. **(E)** Following selection of the tapasin library and deep sequencing, $\log_2$ enrichment ratios for mutations at control positions (in the core or at surface sites distal from the MHC-I interface) are plotted as heatmaps, colored from orange ($\leq$ -3, i.e. depleted/deleterious) to pale/white (0, i.e. neutral) to dark blue ($\geqslant$ +3, i.e. enriched). Due to read length limits for Illumina sequencing, enrichment of the wild type sequence in the experiment is unknown and shown in black. Mutated tapasin positions are on the horizontal axis, while amino acid substitutions are on the vertical axis (*, stop codon). **(F)** Following selection of a 24-35 loop-focused TAPBPR-CT library, $\log_2$ enrichment ratios for nonsynonymous mutations (black) in the 24-35 loop cluster near the origin. Nonsense mutations (red) are depleted. Data from two independent experiments show close agreement. **(G)** Conservation scores from the selection of a larger TAPBPR-CT library focused on the entire MHC-I interface are closely correlated between independent sorting experiments. **(H)** Heatmaps colored as in E, showing that core TAPBPR residues are generally restricted to hyrophobics (*left*) while surface TAPBPR residues distal from the MHC-I interface are mutationally tolerant (*right*).

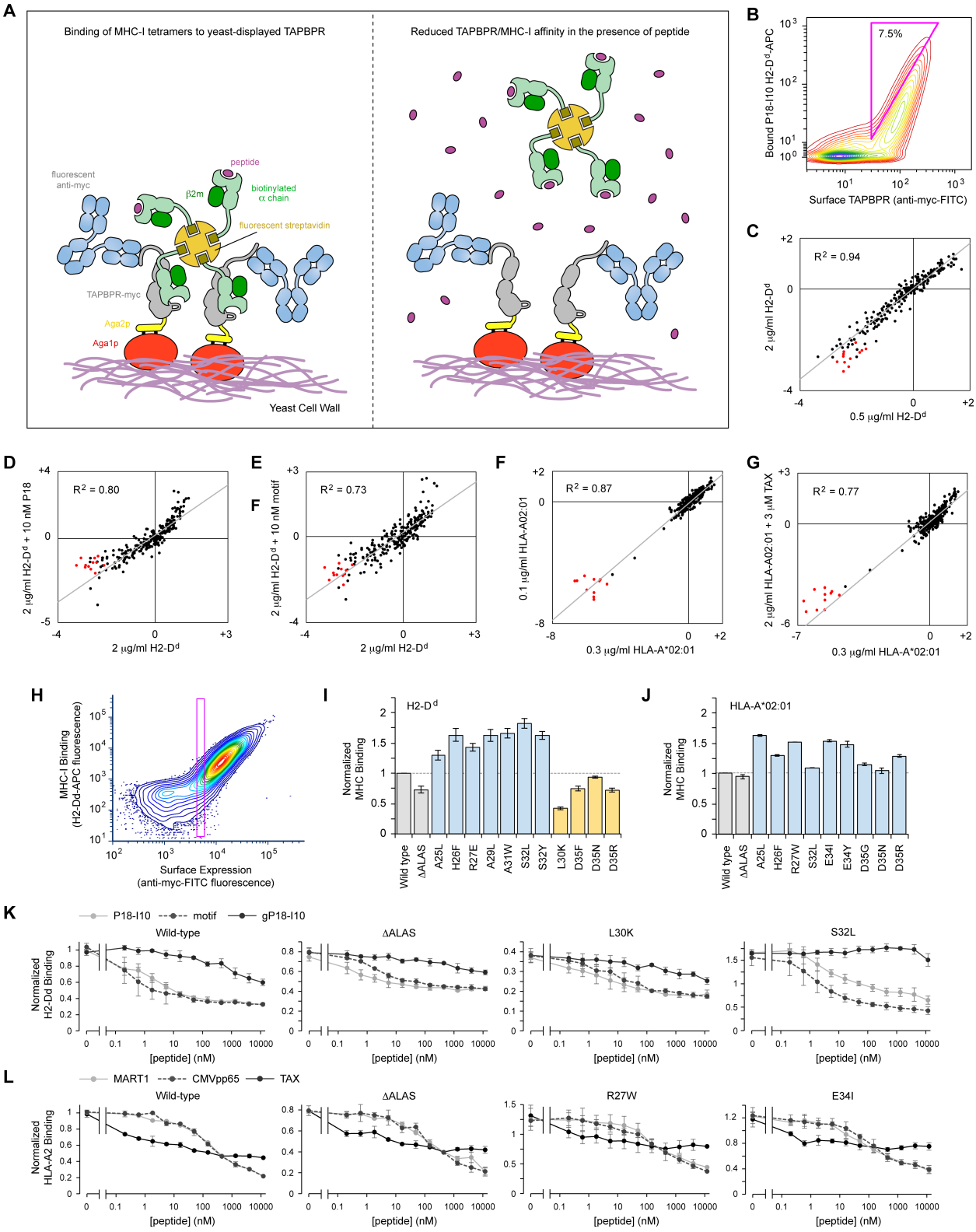

Figure S3. Yeast display for assessing how the sequence of the TAPBPR 24-35 loop influences binding to folded MHC-I. (A) The extracellular domains of TAPBPR are expressed on the surface of the yeast cell wall via an N-terminal fusion with Aga2p. A c-myc epitope tag at the C-

terminus is used for fluorescence detection of expression levels with anti-myc antibodies. MHC-I binding is measured by flow cytometry after incubation of the yeast with fluorescent pMHC-I tetramers. TAPBPR binds MHC-I with highest affinity when it is peptide free, and the addition of peptide reduces MHC-I binding to TAPBPR-expressing yeast. **(B)** The 24-35 loop of TAPBPR was diversified by saturation mutagenesis. The yeast-displayed library was sorted under a variety of conditions for binding to tetramers of P18-I10-loaded H2-D^d or TAX8-loaded HLA-A*02:01, in the presence or absence of free peptides P18-I10 or motif (for H2-D^d) or HTLV-1 TAX (for HLA-A*02:01). The gating strategy for one of the sorting experiments is shown. In this example, yeast were incubated with 0.5 μg/ml H2-D^d tetramer, and yeast expressing TAPBPR variants with the highest levels of binding were collected (magenta gate). MHC-I tetramer staining was below saturation (staining remained unsaturated up to at least 10 μg/ml tetramer). **(C-G)** Yeast-displayed TAPBPR libraries were sorted for binding to MHC-I tetramers and the enrichment or depletion of mutations were calculated following Illumina sequencing of the naive and sorted populations. Comparisons are shown between different sorting experiments for the enrichment ratios of nonsynonymous (black) and nonsense (red) mutations in the TAPBPR 24-35 loop. Data are highly correlated between replicate experiments using different concentrations of H2-D^d (C) or HLA-A*02:01 (F). Enrichment ratios are qualitatively similar when free peptides P18-I10 (D), motif (E), or TAX (G) are co-incubated with H2-D^d (D, E) or HLA-A*02:01 (G). **(H)** Yeast displaying surface TAPBPR were dual stained with P18-I10-loaded H2-D^d-APC tetramers and anti-myc-FITC for simultaneous detection by flow cytometry of bound MHC-I versus TAPBPR expression. Due to the use of avid MHC-I tetramers, binding to TAPBPR was found to be highly dependent on surface expression levels. To control for this, yeast were gated (magenta box) for low TAPBPR expression to minimize avidity effects and control for any differences in expression among TAPBPR mutants. Bound MHC-I (based on mean APC fluorescence) was measured within the magenta gate. **(I, J)** Validation by targeted mutagenesis of TAPBPR mutants predicted from the deep mutational scans to have reduced (orange) or increased (blue) binding to H2-D^d (I; 1.0 μg/ml, mean ± SD from n = 4) or HLA-A*02:01 (J; 2.0 μg/ml, mean ± range from n = 2). TAPBPR ΔALAS (grey) has slightly reduced MHC-I binding. Mean fluorescence for MHC-I binding is normalized to wild type TAPBPR (grey). **(K, L)** Binding of H2-D^d (K; 2.0 μg/ml, mean ± SD from n = 4) or HLA-A*02:01 (L; 2.0 μg/ml, mean ± range from n = 2) to TAPBPR-expressing yeast is competed by peptides. Shown are data from yeast expressing wild type TAPBPR, TAPBPR ΔALAS, and representative mutants. gP18-I10 is a low affinity peptide for H2-D^d missing an N-terminal anchoring residue for the A pocket. Peptide sequences are: RGPGRAFVTI (P18-I10), GPGRAFVTI (gP18-I10), AGPARAAAL (motif), LLFGYPVYV (HTLV-1 TAX9), ELAGIGILTV (MART1), NLVPMVATV (CMVpp65).

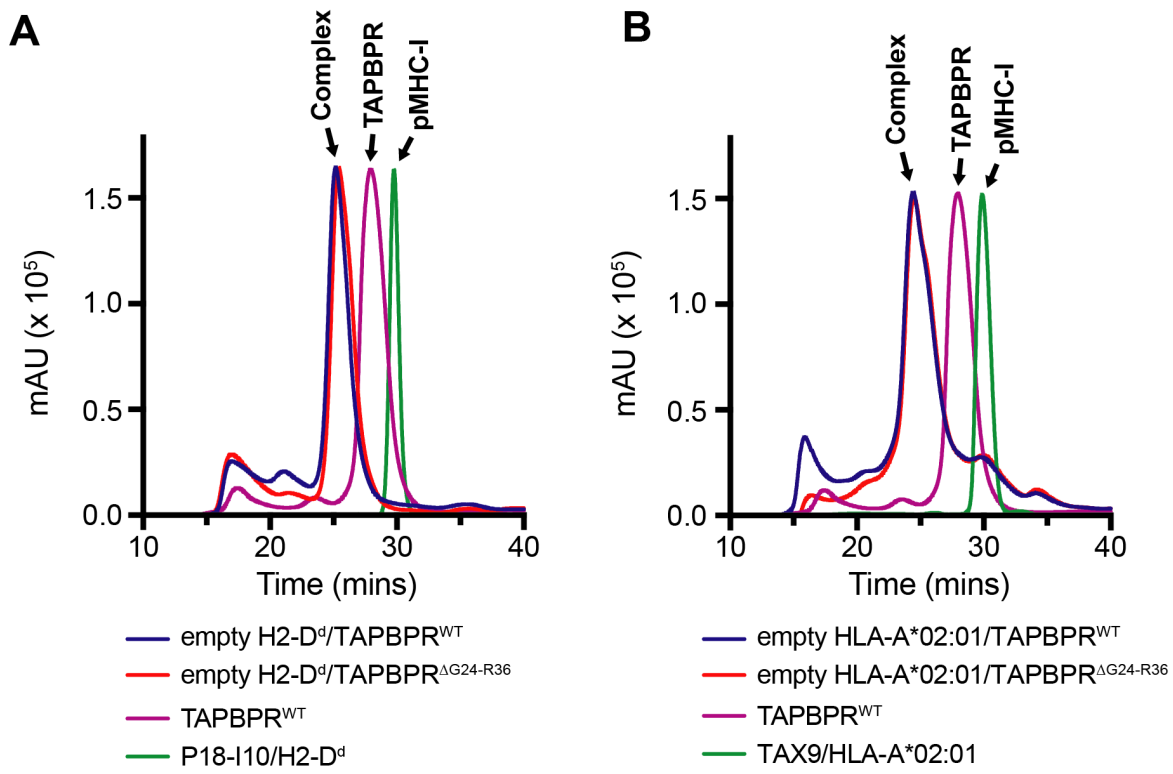

Figure S4. Purification of peptide-deficient MHC-I/TAPBPR complexes.
Size exclusion chromatographs (0.5 mL/min flow rate) of peptide-deficient MHC-I/TAPBPR complexes, relative to the free pMHC-I and TAPBPR species. photoP18-I10/H2-D^d/hβ_2m **(A)** or photoFluM1/HLA-A*02:01/hβ_2m **(B)** were mixed with TAPBPR at 1:1 molar ratio and incubated for 1 hour at 25°C. The mixture was then UV-irradiated at 365 nm for 1 hour followed by centrifugation at 13,000 r.p.m. for 10 min to remove precipitates. Pure 1:1 stoichiometric H2-D^d/hβ_2m/TAPBPR (A) or HLA-A*02:01/hβ_2m/TAPBPR (B) complexes were isolated by SEC with a Superdex 200 Increase 10/300 GL column at flow rate of 0.5 mL/min in 50 mM NaCl, 20 mM sodium phosphate pH 7.2. The exact procedure was followed for complexes containing either TAPBPR^{WT} or TAPBPR$^{\Delta G24\text{-}R36}$, with chromatograms shown as different colors. Validation of peptide removal from the peaks at approx. 26 min corresponding to the MHC-I/TAPBPR complexes was achieved using liquid chromatography-mass spectroscopy (LC-MS) and carried out with passage through a Higgins PROTO300 C4 column (5 µm, 100 mm × 2.1 mm) followed by electron ion spray mass spectroscopy performed on a Thermo Finnigan LC-MS/MS (LTQ).

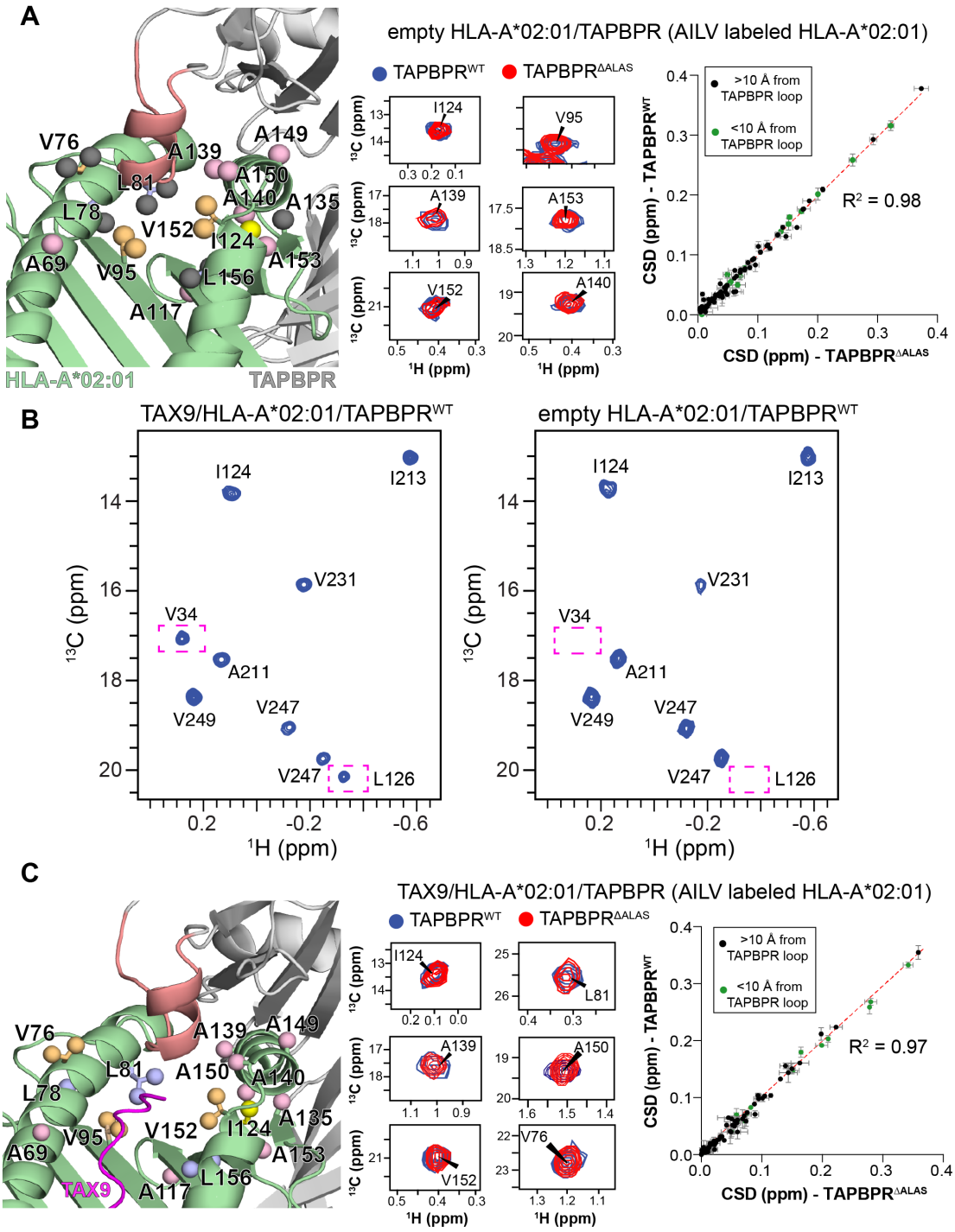

Figure S5. The TAPBPR G24-R36 loop does not form a stable interaction with the HLA-A*02:01 groove. (A) (Left) View of the peptide-deficient HLA-A*02:01/TAPBPR model (template PDB ID 5OPI) showing AILV methyl probes on HLA-A*02:01 (as spheres) within 10 Å angstroms of the TAPBPR G24-R36 loop (salmon). Methyl resonances of HLA-A*02:01 residues shown in black are missing in 2D ^{1}H-^{13}C methyl HMQC spectra of peptide-deficient HLA-A*02:01/hβ_2m/TAPBPR complex due to conformational exchange induced line broadening. (Middle) Zoom in of representative peaks from 2D ^{1}H-^{13}C methyl HMQC spectra of 80 μM peptide-deficient HLA-A*02:01 (AILV labeled)/hβ_2m in complex with TAPBPR^{WT} (red) and

TAPBPR$^{\Delta ALAS}$ (blue) recorded at 25°C at a ^{1}H field of 800 MHz. (Right) Comparison of chemical shift deviation (CSD) values measured between free TAX9/HLA-A*02:01/hβ_2m and peptide-deficient HLA-A*02:01/hβ_2m in complex with TAPBPR^{WT} or TAPBPR$^{\Delta ALAS}$. CSD values for HLA-A*02:01 methyl probes within 10 Å of the TAPBPR G24-R36 loop are shown in green, while those further than 10 Å are shown in black. The dotted red line is a linear fit of the data with R^2 value noted. **(B)** Comparison of 2D ^{1}H-^{13}C HMQC spectra of 80 μM TAX9/HLA-A*02:01 (AILV labeled)/hβ_2m/TAPBPR^{WT} in the presence of 200 μM excess TAPBPR^{WT} (left) with 80 μM peptide-deficient HLA-A*02:01 (AILV labeled)/hβ_2m/TAPBPR^{WT} (right). Dotted pink boxes highlight methyl resonances that become exchange broadened in the wild-type empty HLA-A*02:01/TAPBPR complex but are present in the wild-type TAX9/HLA-A*02:01/ complex. **(C)** (Left) View of the TAX9/HLA-A*02:01/TAPBPR model (template PDB ID 5OPI) showing AILV methyl probes on HLA-A*02:01 (as spheres) within 10 Å angstroms of the TAPBPR G24-R36 loop (salmon). (Middle) Zoom in of representative peaks from 2D ^{1}H-^{13}C methyl HMQC spectra of 80 μM TAX9/HLA-A*02:01 (AILV labeled)/hβ_2m in complex with TAPBPR^{WT} (red) and TAPBPR$^{\Delta ALAS}$ (blue) recorded at 25°C at a ^{1}H field of 800 MHz. (Right) Comparison of chemical shift deviation (CSD) values measured between free TAX9/HLA-A*02:01/hβ_2m and TAX9/HLA-A*02:01/hβ_2m in complex with TAPBPR^{WT} or TAPBPR$^{\Delta ALAS}$. CSD values for HLA-A*02:01 methyl probes within 10 Å of the TAPBPR G24-R36 loop are shown in green, while those further than 10 Å are shown in black. The dotted red line is a linear fit of the data with R^2 value noted.

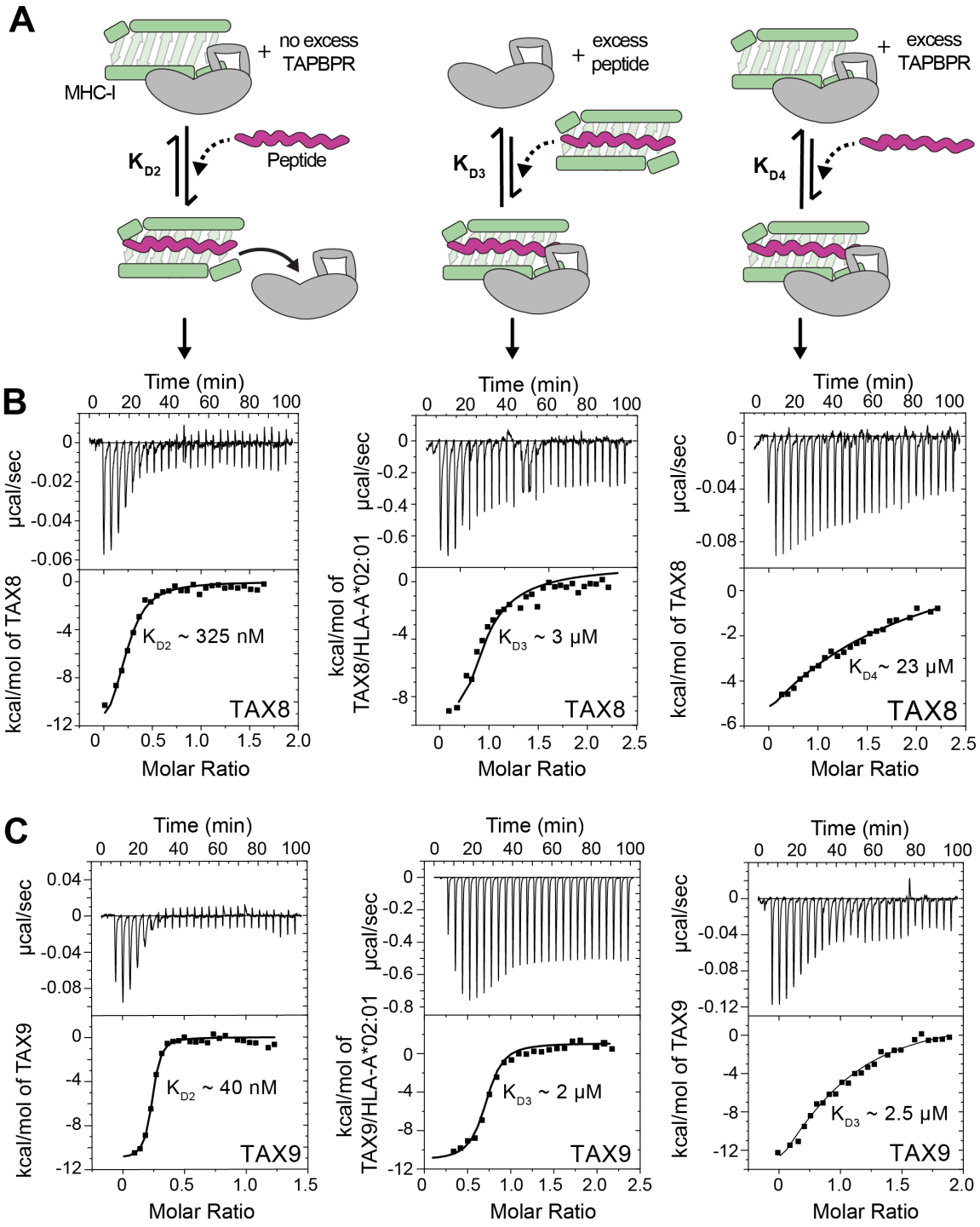

Figure S6. Design of ITC experiments to examine specific steps of the thermodynamic cycle of TAPBPR mediated peptide exchange. (A) Schematic of conditions in the calorimetry cell (top) and the syringe (titration represented by the dotted arrow) during ITC experiments for each step of the peptide exchange cycle. **(B and C)** Representative examples of raw ITC data for each step of the peptide exchange cycle for TAX8 and TAX9 peptide. The line represents the best fit of the data using a 1:1 model in Origin. Fitted apparent K_D values are reported. Details of the experiments are outline in the *Materials and Methods*.

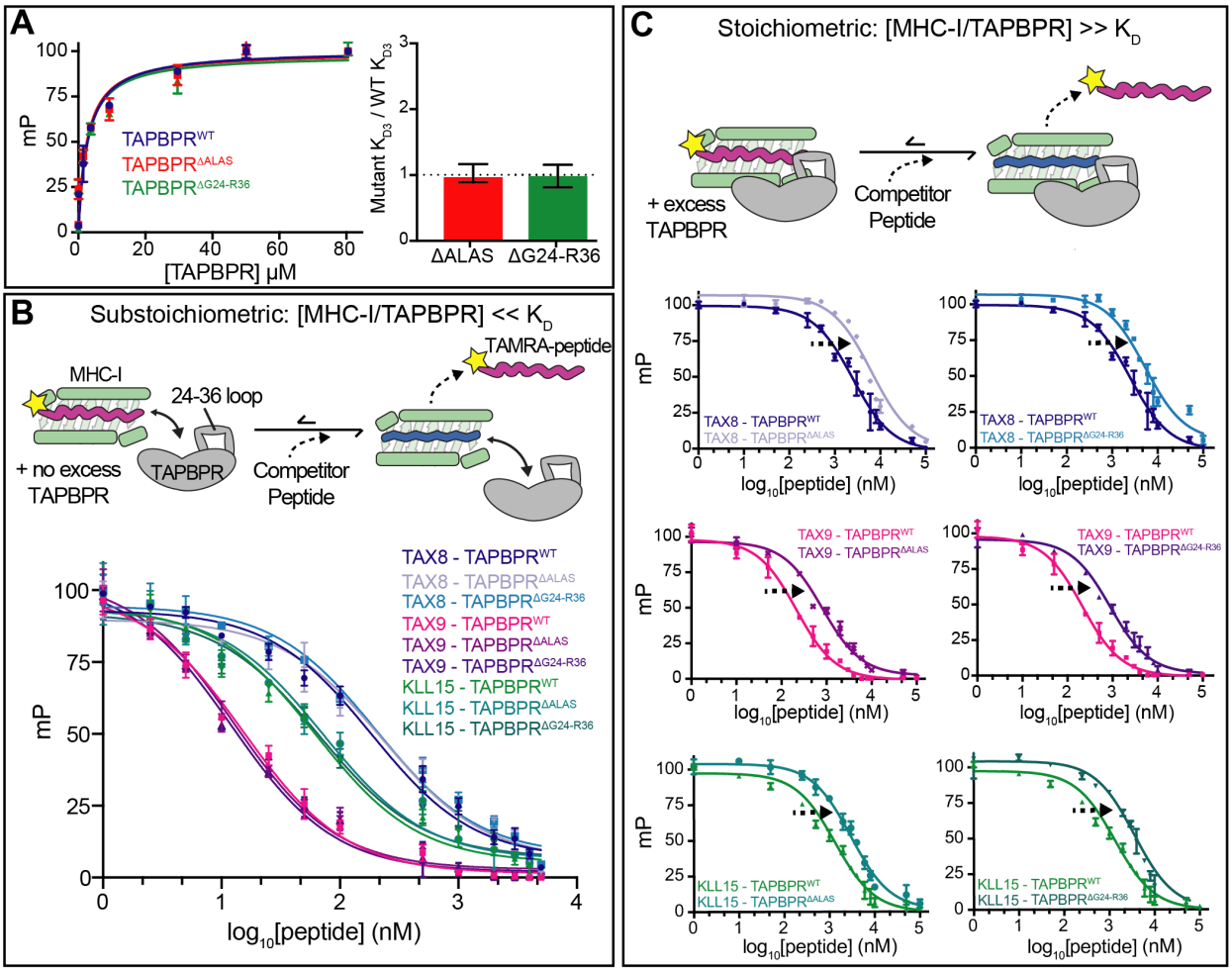

Figure S7. Design of FP assays to examine specific steps of the thermodynamic cycle of TAPBPR mediated peptide exchange. (A) FP experiments to probe the K_{D3} step of the peptide exchange cycle. (Left) Titration of graded concentrations of $TAPBPR^{WT}$, $TAPBPR^{\Delta ALAS}$ or $TAPBPR^{\Delta G24\text{-}R36}$ into 1 nM TAMRA-TAX9 and 50 nM TAX8/HLA-A*02:01/hβ_2m. (Right) Ratio of FP-determined K_{D3} values for mutant / wild-type TAPBPR. **(B)** Schematic of the design of FP experiments performed under substoichiometric conditions (no excess TAPBPR). Under substoichiometric conditions, 50 nM of wild-type or G24-R36 loop mutant peptide-deficient HLA-A*02:01/hβ_2m/TAPBPR complex is incubated with 1 nM TAMRA-TAX9 peptide and a range of concentrations of competitor peptide. millipolarization (mP) values are plotted as a function of the $\log_{10}$ peptide concentration for TAX8, TAX9 and KLL15 peptides. **(C)** Schematic of the design of FP experiments performed under stoichiometric conditions (excess TAPBPR). Under stoichiometric conditions, 50 nM of wild-type or G24-R36 loop mutant peptide-deficient HLA-A*02:01/hβ_2m/TAPBPR complex is incubated with 1 nM TAMRA-TAX9 peptide, 2 μM TAPBPR (or $TAPBPR^{\Delta ALAS}$ or $TAPBPR^{\Delta G4\text{-}R36}$) and a range of concentrations of competitor peptide. mP values are plotted as a function of the $\log_{10}$ peptide concentration for TAX8, TAX9 and KLL15 peptides. Error bars were obtained from three technical replicates.

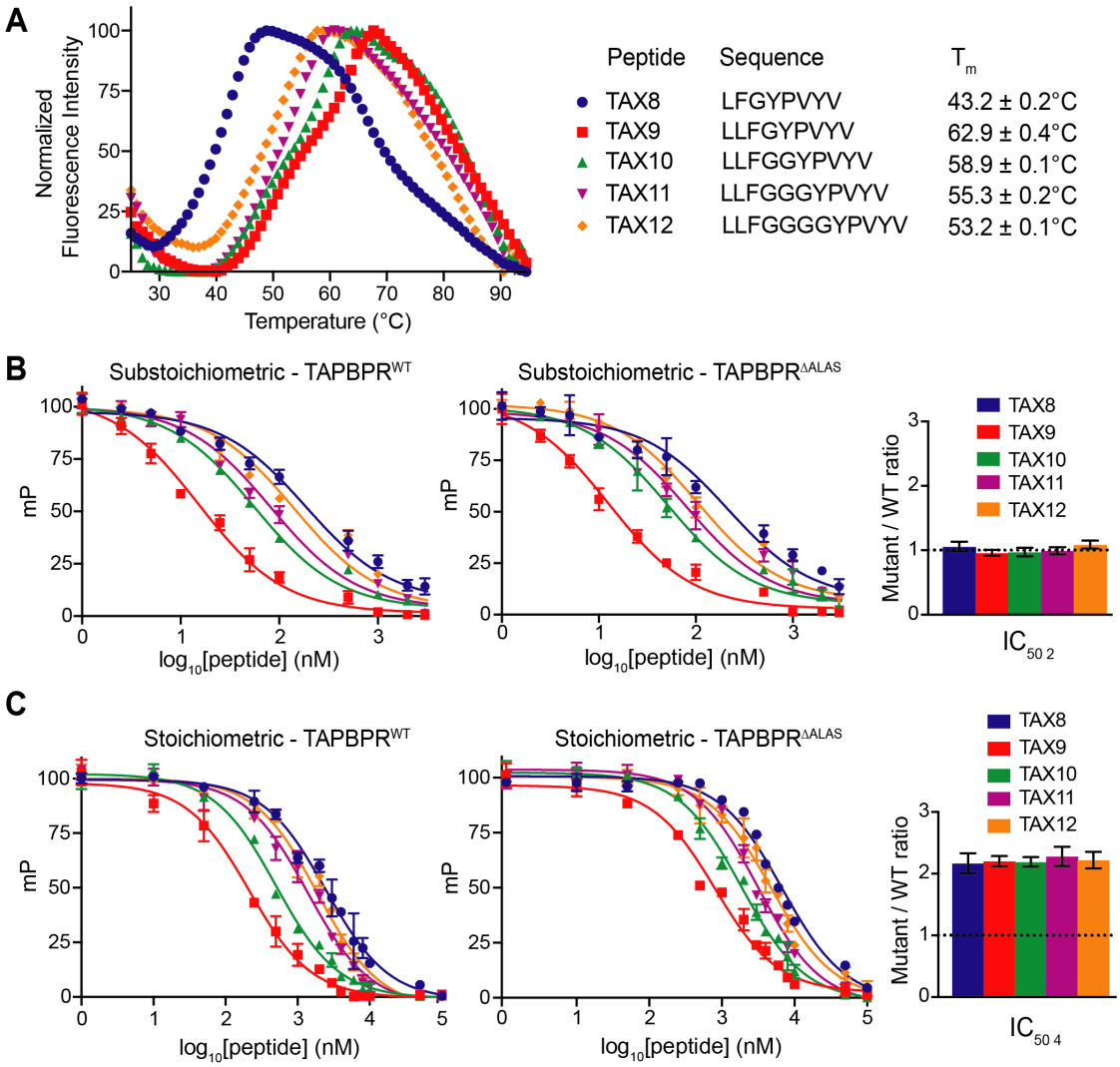

Figure S8. Evaluation of peptide length dependence for TAPBPR mediated exchange. (A) DSF experiments performed on 7 µM HLA-A*02:01/hβ$_2$m in complex with different TAX length variants. The peptide sequence and determined T_m values are shown. Errors were determined from three replicates. Fluorescence intensities were normalized for comparison. **(B)** (Left, Middle) FP performed under substoichiometric conditions, 50 nM of peptide-deficient HLA-A*02:01/hβ$_2$m/TAPBPR^{WT} or HLA-A*02:01/hβ$_2$m/TAPBPR$^{\Delta ALAS}$ complex is incubated with 1 nM TAMRA-TAX9 peptide and a range of concentrations of competitor peptide. millipolarization (mP) values are plotted as a function of the $\log_{10}$ peptide concentration for the TAX length variant peptides. (Right) Comparison of the ratio of FP determined $^2IC_{50}$ values for TAPBPR$^{\Delta ALAS}$ versus TAPBPR^{WT}. The dotted line represents no effect. **C.** (Left, Middle) FP performed under stoichiometric conditions, 50 nM of peptide-deficient HLA-A*02:01/hβ$_2$m/TAPBPR^{WT} or HLA-A*02:01/hβ$_2$m/TAPBPR$^{\Delta ALAS}$ complex is incubated with 1 nM TAMRA-TAX9 peptide, 2 µM TAPBPR (or TAPBPR$^{\Delta ALAS}$) and a range of concentrations of competitor peptide. mP values are plotted as a function of the $\log_{10}$ peptide concentration for the TAX length variant peptides. (Right) Comparison of the ratio of FP determined $^4IC_{50}$ values for TAPBPR$^{\Delta ALAS}$ versus TAPBPR^{WT}. The dotted line represents no effect.

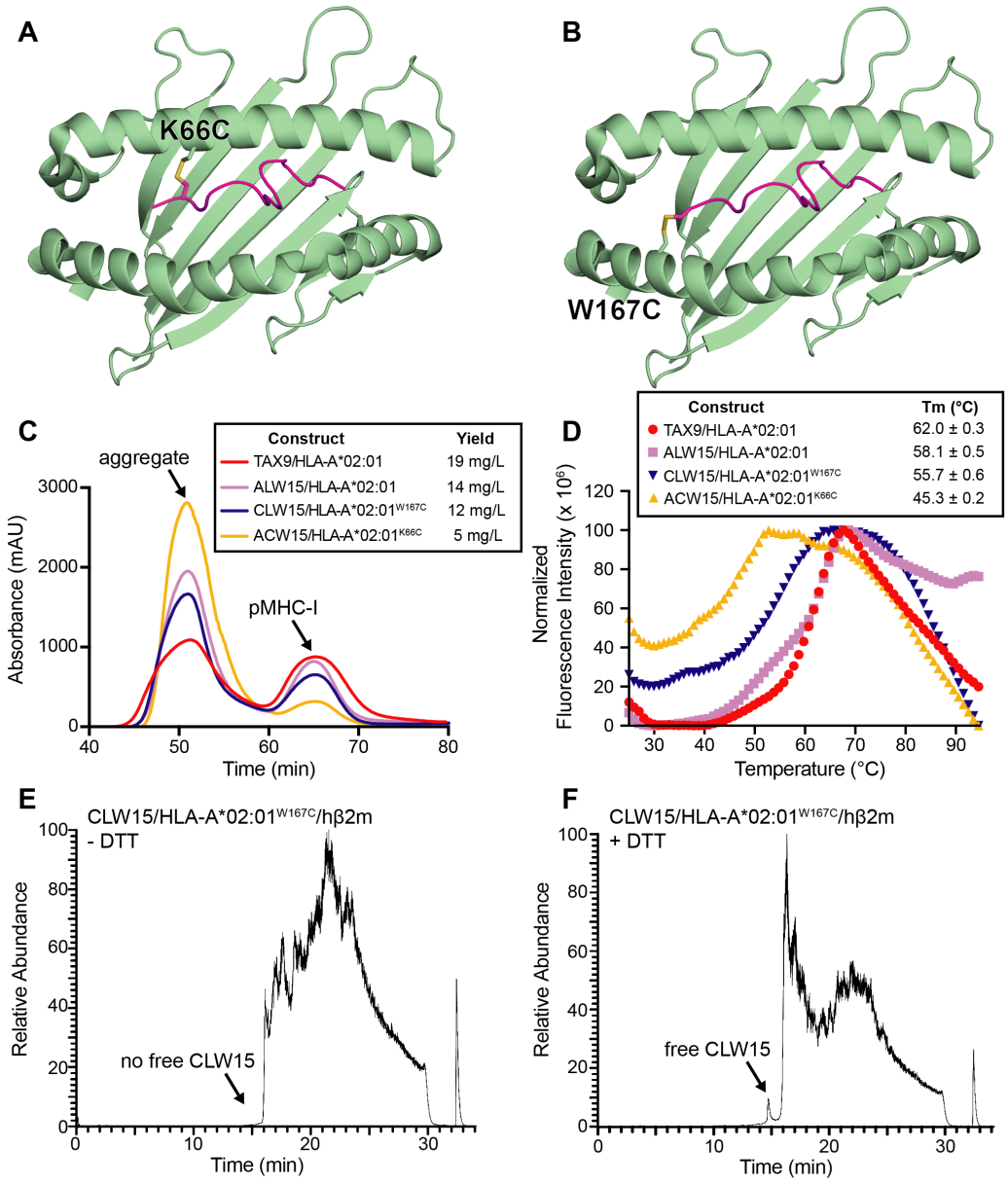

Figure S9. Design of disulfide mutants linking peptide to the MHC-I groove. Structural models of disulfides between ALW15 peptide (ALWDIETGQQKTVFV) and the HLA-A*02:01 groove based on Disulfide by Design v2. **(A)** Predicted disulfide between L2C of ALW15 (herein called ACW15) and K66C of HLA-A*02:01. **(B)** Predicted disulfide between A1C of ALW15 (herein called CLW15) and W167C of HLA-A*02:01. **(C)** Purification of *in vitro* refolded pMHC-I constructs by size exclusion chromatography. TAX9 is a high affinity peptide reference. ALW15 is the reference for the wild-type 15mer peptide. ACW15 and CLW15 peptides are the disulfide mutant designs shown in A. and B., respectively. The final refolded protein yield after purification is noted. **(D)** Differential scanning fluorimetry of purified pMHC-I constructs. The fitted melting temperature (Tm) is noted. **(E and F)** LC-MS of CLW15/HLA-A*02:01/hβ2m without and with 1 mM DTT.

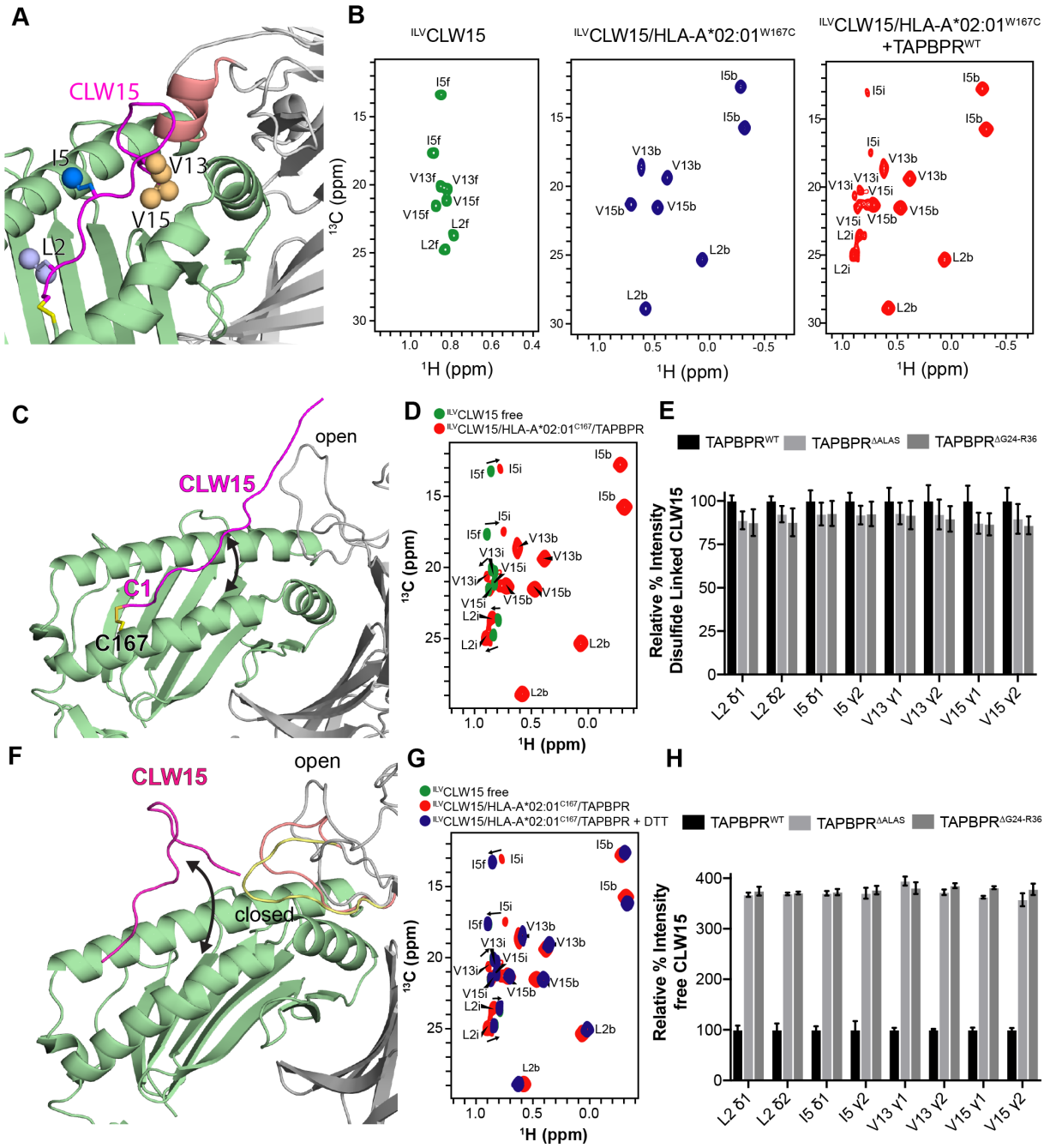

Figure S10. NMR characterization of chaperone mediated peptide exchange on MHC-I for a 15mer peptide. (A) Model of the CLW15/HLA-A*02:01/TAPBPR complex designed using Disulfide by Design v2. The disulfide is formed between C1 of CLW15 (CLWDIETGQQKTVFV) and W167C of HLA-A*02:01. The open conformation of the TAPBPR G24-R36 loop obtained from MD simulations is shown in grey. **(B)** 2D ^{1}H-^{13}C methyl HMQC spectra of ^{15}N/^{13}C ILV labeled CLW15 in the free state (green) and in complex with HLA-A*02:01 in the presence of 8-fold excess TAPBPR (red) recorded at 25 °C at a ^{1}H field of 800 MHz. Methyl resonances in the pMHC-I state are denoted "b" for bound. Methyl resonances in the pMHC-I/TAPBPR state are denoted "i" for intermediate. Methyl resonances for free CLW15 are denoted "f" for free. **(C)** Comparison of NMR signal intensity for each methyl resonance of disulfide linked CLW15

intermediate in the pMHC-I/TAPBPR state with and without the TAPBPR G24-R36 loop. **(D)** Model of the CLW15/HLA-A*02:01/TAPBPR complex is shown when the disulfide between CLW15 and HLA-A*02:01 is reduced with DTT. The open and closed conformations of the TAPBPR G24-R36 loop obtained from MD simulations are shown in grey, salmon and yellow. **(E)** 2D ^{1}H-^{13}C methyl HMQC spectra of ^{15}N/^{13}C ILV labeled CLW15 in the free state (green), in complex with HLA-A*02:01 in the presence of 8-fold excess TAPBPR (red) and in complex with HLA-A*02:01 in the presence of 8-fold excess TAPBPR and 1 mM DTT (blue) recorded at 25 °C at a ^{1}H field of 800 MHz. Methyl resonances corresponding to the pMHC-I bound state are denoted "b" for bound. Methyl resonances corresponding to the pMHC-I/TAPBPR state are denoted "i" for intermediate. Methyl resonances corresponding to the free CLW15 state are denoted "f" for free. **(F)** Comparison of NMR signal intensity for each methyl resonance of free CLW15 peptide when released from the MHC-I/TAPBPR complex with and without the TAPBPR G24-R36 loop.