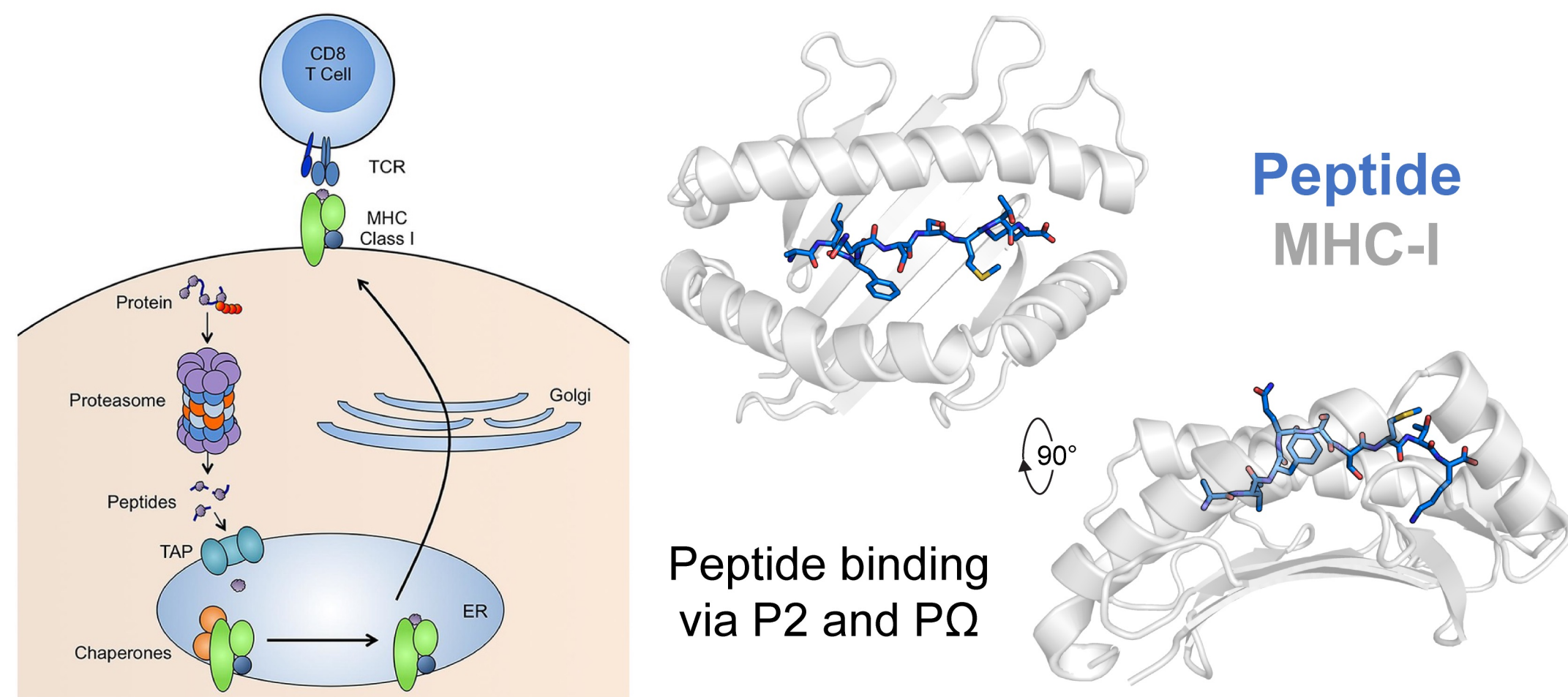


## Introduction

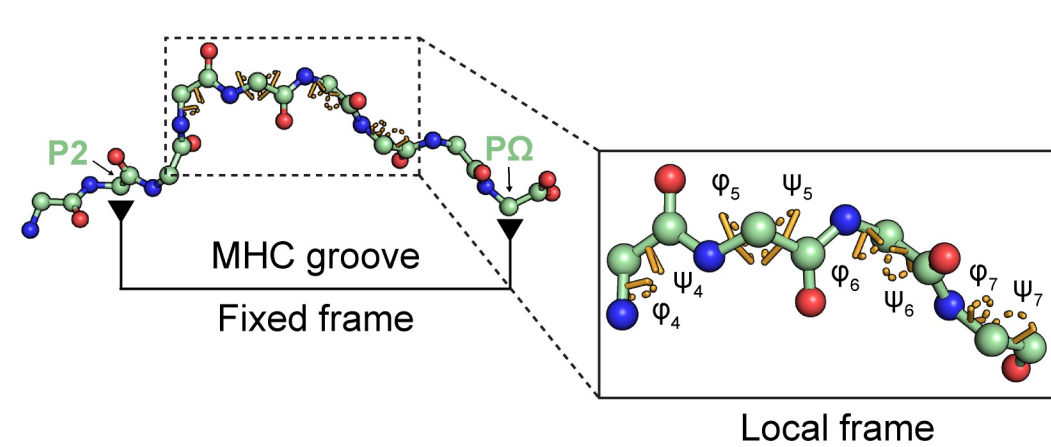


Human leukocyte antigens (HLAs), the human version of major histocompatibility complex molecules (MHC-I), display self, foreign, or aberrant peptides on the cell surface for immune surveillance<sup>1</sup>.

The sequence and conformational diversity offered by the center of the peptide defines the immunologically important area of peptide/HLA-I (pHLA) structures<sup>2</sup>.

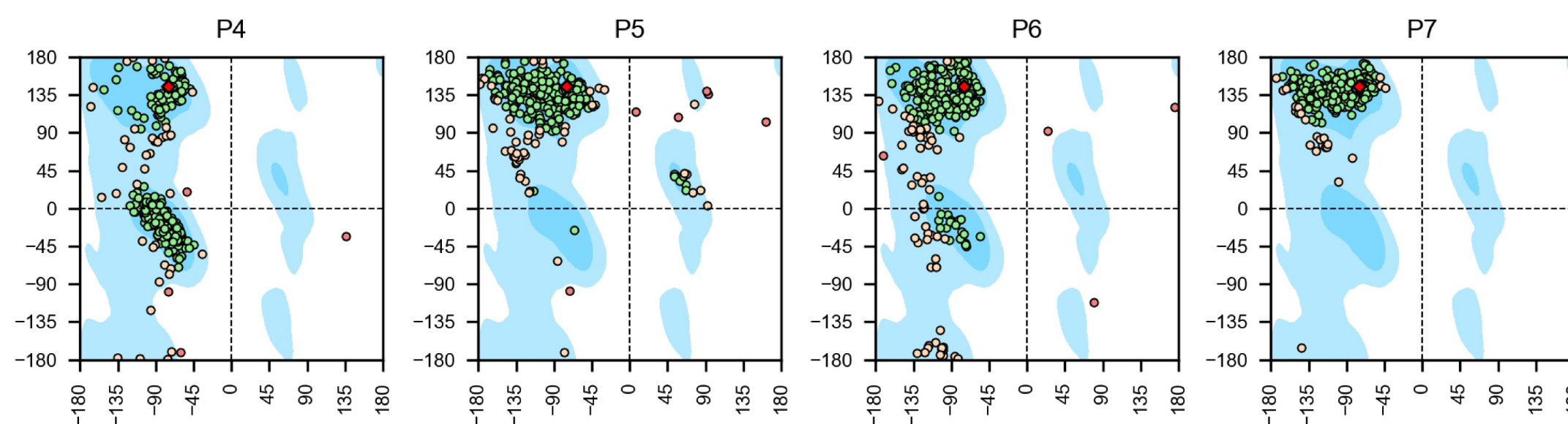
Thus, accurate modeling of the center of the peptide is necessary to understand and predict the molecular basis for immunogenicity.

## Peptide conformational diversity



A structural framework, termed the fixed-local frame, captures peptide conformational diversity.

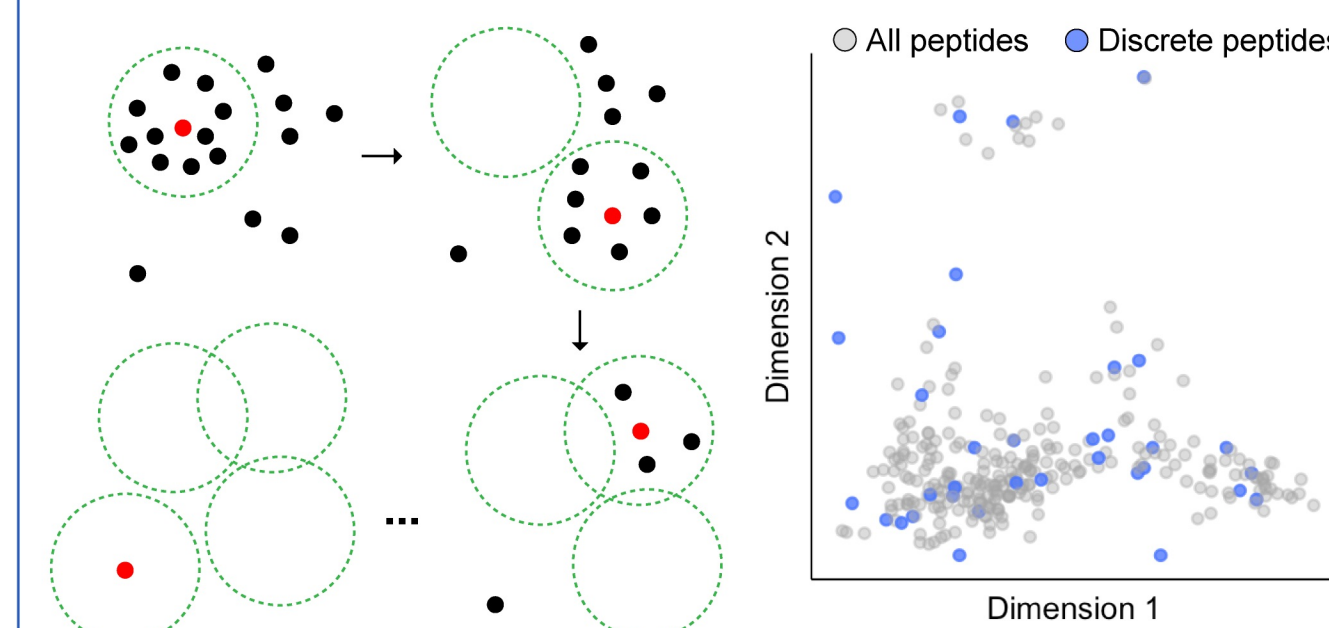
Primary anchor residues work as a fixed frame while the center of the peptide enables backbone changes.



A comprehensive structural analysis of 9mer pHLA X-ray crystal structures (n = 293) from our database (HLA3DB) was conducted to determine the distributions of φ and ψ dihedral angles for each peptide position.

The central residues of the peptide (P4 to P7) showed frequent deviations towards more extended (β-strand) or more condensed (α-helix) structure.

## Unbiased classification of peptide backbones

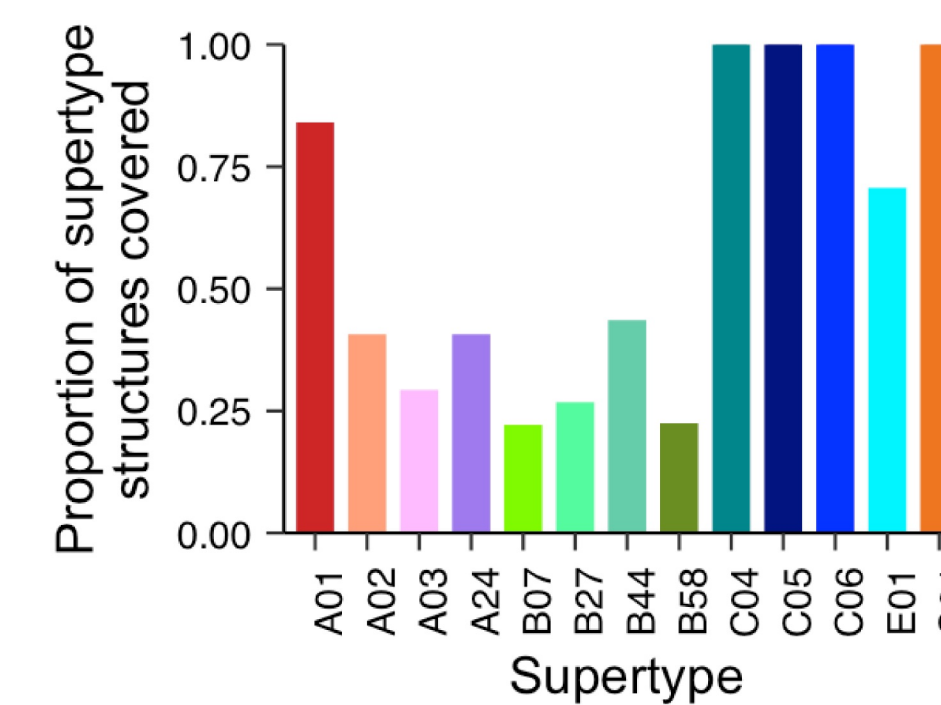


A minimal set of 34 peptide backbones can describe the entire conformational landscape.

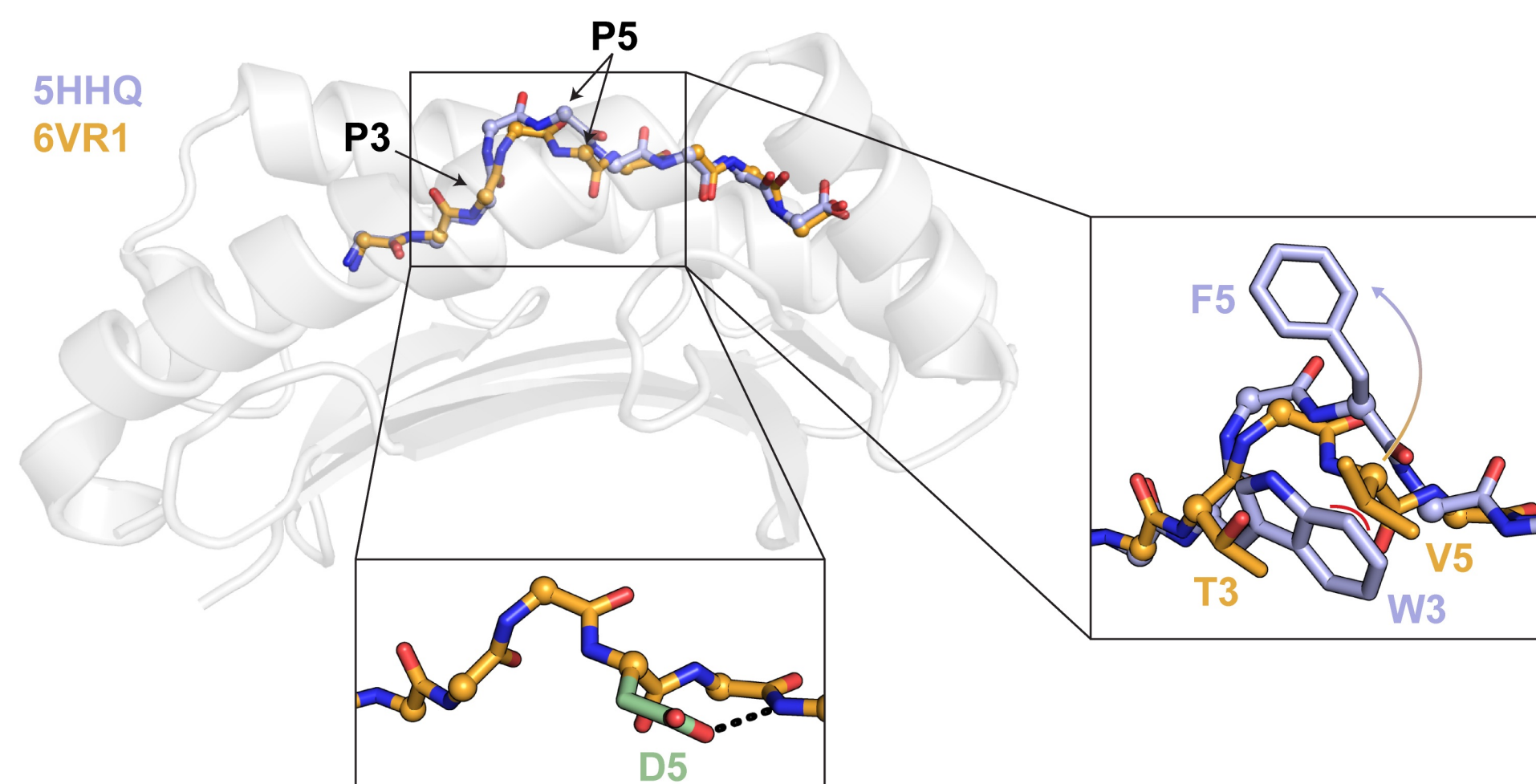
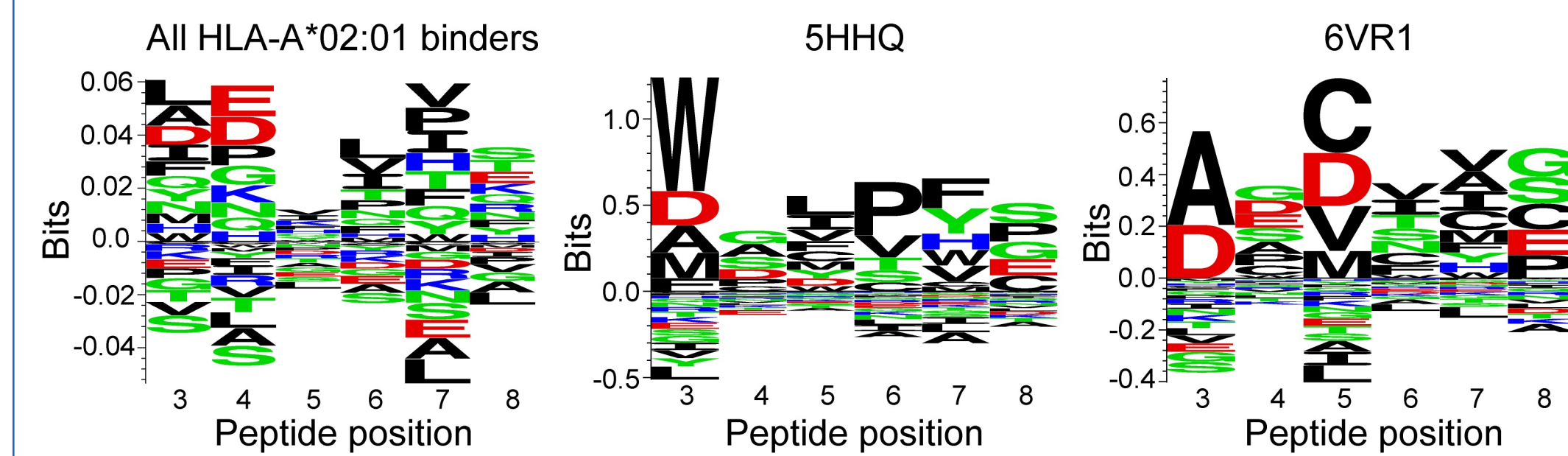
A two-dimensional PCA plot shows that rare configurations are represented.

The most common peptide backbone accounted for 35% of all configurations in our dataset.

The pHLA complexes represented by this peptide backbone originated from diverse HLA supertypes, indicating that similar antigen conformations are not solely defined by the identity of HLA groove residues.

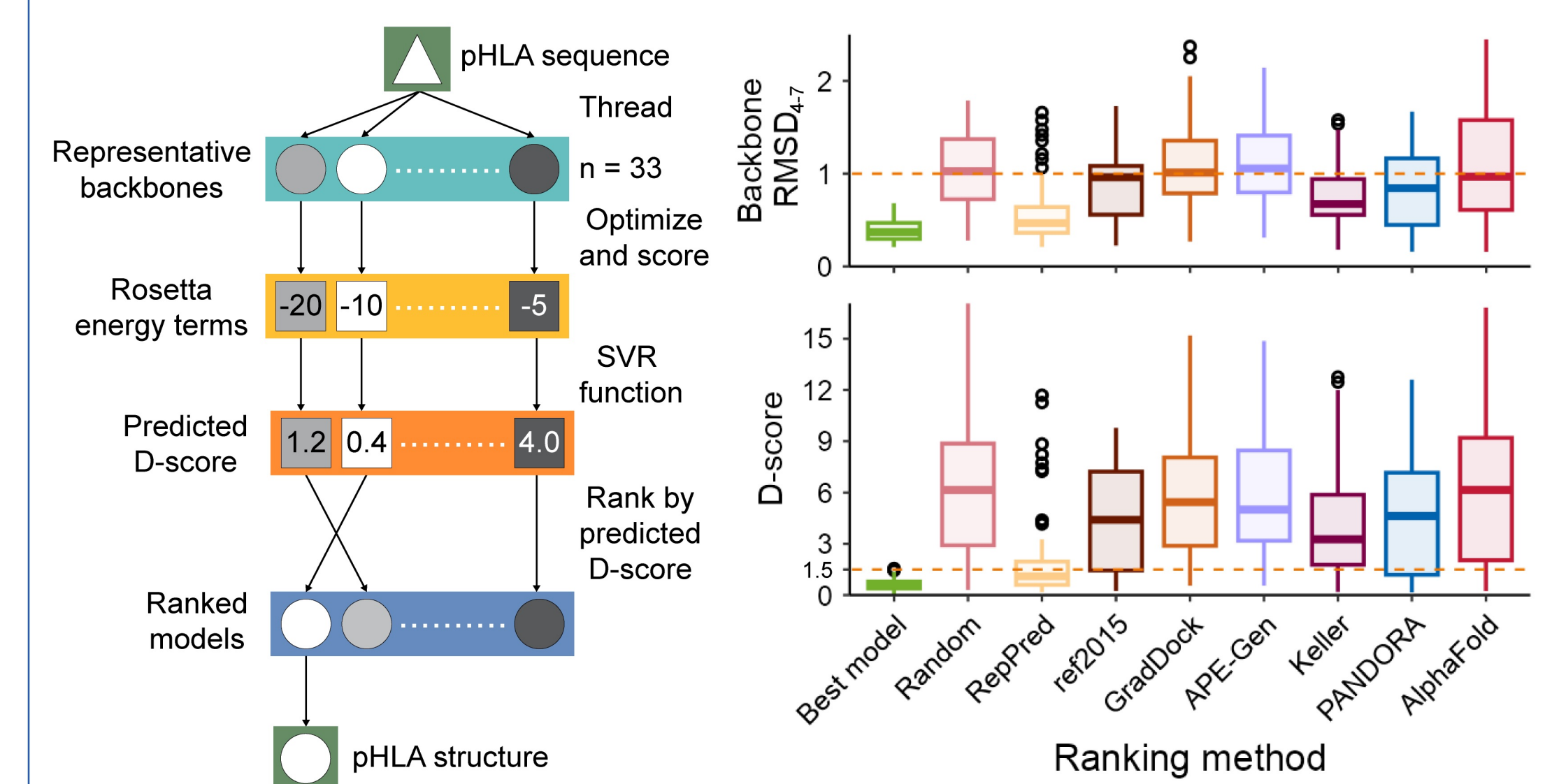


## Sequence biases across discrete conformations



Augmentation of the peptide sequence space by Rosetta demonstrate how backbone conformation can impose steric constraints or introduce favorable interactions via crosstalk between backbone and side chain features.

## Accurate structural modeling



Leveraging the discrete backbones to enable structural modeling of new peptides on all HLA allotypes, RepPred outperforms five existing methods by at least 32% with respect to D-score.

RepPred models an immunodominant HIV epitope with near-native side chain placement as a result of high-fidelity generation of the peptide backbone.

QATQEVKNW/HLA-B\*57:01 (target)  
FEDLRVLSF/HLA-B\*37:01 (template)  
D-score: 0.39

## Conclusions

Using HLA3DB, our database of pHLA structures, we characterize peptide backbone diversity for 9mer pHLA complexes.

We find that peptide backbone similarity is allotype-independent and identify 34 discrete configurations which cover the conformational landscape.

We discover strong peptide sequence trends influenced by distinct backbone features.

RepPred, an accurate, pan-allelic structural modeling approach for pHLA complexes, demonstrates improved accuracy over existing methods.

## References and Acknowledgments

<sup>1</sup>Wieczorek, M. *et al.* Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. *Front Immunol* **8**, 292 (2017).  
<sup>2</sup>Rudolph, M. G., Stanfield, R. L. & Wilson, I. A. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* **24**, 419–466 (2006).

This research was supported through grants by NIAID (5R01AI143997), NIGMS (5R35GM125034), and NIDDK (5U01DK112217).

HLA3DB can be accessed via <https://hla3db.research.chop.edu>.