Case Study

MindWalk™

LensAI™ Epitope Mapping

LensAI Epitope Mapping retains near x-ray precision on 'unseen' diverse clinically relevant complexes

Overview:

Epitope mapping remains a cornerstone of therapeutic antibody development. In a previous case study LensAI Epitope Mapping was evaluated against various wet-lab methods, using epitopes determined by x-ray crystallography as the ground truth. LensAI demonstrated superior performance, achieving near x-ray accuracy. In this study we expand the benchmark. LensAI Epitope Mapping retains near x-ray precision—validated on an extended benchmark of 30 clinically relevant complexes, 17 entirely novel to the system.

Challenge:

LensAI *in silico* Epitope Mapping is based on a predictive model. The challenge of any predictive model is to demonstrate its ability to generalize beyond its training data.

Background: a head-to-head comparison with x-ray crystallography

On 30 known complexes the epitope predicted by LensAI is compared with the epitope identified by x-ray crystallography, which is considered the ground truth. Epitope residues are identified as those of the antigen whose heavy atoms are less than 5 Angströms away from the antibody atoms.

13 complexes have been part of the training dataset of our models. 17 complexes are out of set and thus 'unseen', in other words not used to train the model. The 17 out of set complexes were carefully chosen to be substantially different, e.g. protein families distant from those in the training set.

Benchmarking 30 complexes:

Target	Training set	Non-training set
Number of AAs in target	173–1114	200–1312
Monomeric	8	17
Dimeric	5	0
Soluble proteins	3	3
(EC part of) transmembrane proteins	7	5
Virus, toxin, allergen, ...	3	9

Antibody	Training set	Non-training set
VH	2	0
Fv	4	3
Fab	5	11
VHH	2	3

Method:

To quantify LensAI epitope mapping prediction accuracy the following standard metrics are used:

The True Positive Rate | Recall

$$TPR = \frac{TP}{(TP+FN)}$$

The proportion of residues being part of the true epitope that are correctly identified.

The False Positive Rate

$$FPR = \frac{FP}{(FP+TN)}$$

The proportion of residues not being part of the true epitope that are incorrectly predicted as part of it.

AUC (Area Under the Curve)

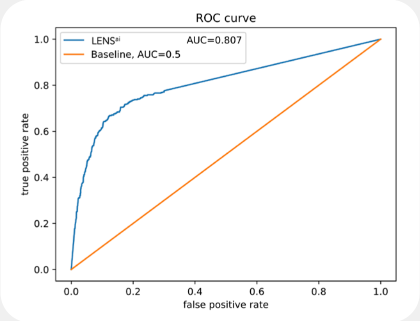
AUC is a single value derived from the ROC curve, which plots TPR against the FPR. AUC quantifies the model's ability to distinguish epitopes from non-epitopes.

AUC = 1 perfect prediction.
AUC >0.8 excellent
AUC = 0.5 no better than random guessing.

Outcome:

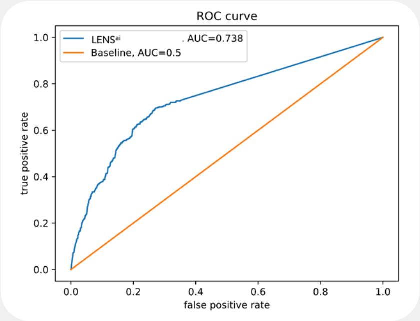
A nearly similar AUC is achieved for both datasets. This shows the robustness of LensAI predictions on 'unseen-novel' input. LensAI Epitope Mapping retains x-ray like precision on models that were not used for ML training of the algorithm.

IN training dataset **AUC = 0.81**



The AUC score of the ROC curve aggregated over the training set

OUT of set dataset **AUC = 0.74**



The AUC score of the ROC curve aggregated over the out of set complexes

Conclusion:

LensAI Epitope Mapping continues demonstrating x-ray like precision in this extended benchmark study of 30 clinically relevant targets. Average AUC score of the complexes that were not part of the training set (n=17) was only slightly lower than the AUC score obtained for the training set, indicating the robustness of the prediction. Unlike traditional methods, LensAI requires only the target and antibody sequences as input and delivers results within hours, enabling high-throughput application. Early epitope mapping with LensAI helps accelerate decision-making and manage potential risks in discovery and development.

MindWalk™

For more information, contact:
info@mindwalkAI.com

mindwalkAI.com