Case Study

LENS^{ai}™ Epitope Mapping

BioStrand
a subsidiary of IPA

Comparing LENS^{ai} *in silico* Epitope Mapping with x-ray crystallography

The Challenge:

Epitope mapping is an essential process for characterizing lead candidates. Identifying the epitope region on the target of a pool of antibodies (Abs) provides critical information on both diversity and potential function. As the demand for expedited workflows in antibody discovery and development increases, *in silico* epitope mapping emerges as a highly attractive, rapid and scalable alternative to traditional methods such as x-ray crystallography and HDX mass spectrometry. Given that the binding information obtained through crystallographic methods is considered the gold standard, it is crucial to strive for similar accuracy levels with *in silico* procedures.

The Approach:

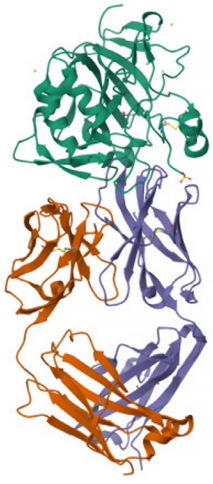
LENS^{ai} Epitope Mapping utilizes unbound structures as input to determine the epitope region. In the absence of available x-ray crystallography structure, predicted models can be employed. This case study compared the accuracy of LENS^{ai} Epitope Mapping for a selected antibody (Ab) and a target using contact information derived from x-ray crystallography. A contact is defined when an atom in target's residue is within 5 Å of an atom in an antibody's residue. Target residues are weighted by the number of antibody residues they interact with. For evaluation of prediction accuracy three measures are used: receiver operating characteristic (ROC), precision-recall and contact map overlap.

Ground truth

Contacts are derived from crystal structure 6RPS. These contacts are used as "ground truth".

CAXII

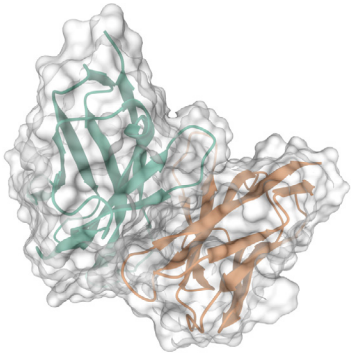
Fab 6A10



LENS^{ai} Epitope Mapping input

The input information required is the unbound structures of Ab and target. Both predicted models can be used, as well as crystal models from PDB. In this case study the PDB model 1JCZ (EC domain of CAXII) was used and the model of Fv 6A10 obtained by LENS^{ai} antibody-specific *in silico* structure prediction pipeline.

Fv 6A10:



Results: contact map overview of both methods

LENS[ai] sequence view: CAXII monomer

X-ray blue bars: the height indicates the number of antibody residues that are in contact with the target residue.

LENS[ai] blue bars: The height indicates the LENS[ai] Epitope Mapping confidence score.

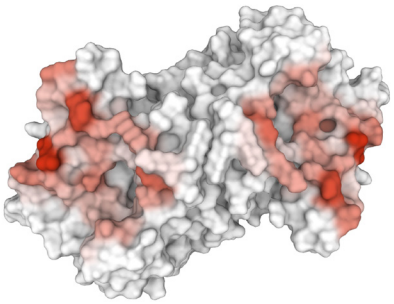
The LENS[ai] Epitope Mapping confidence score (ranging from 0-1) predicts the probability of a target residue interacting with the Ab, and thus, being part of the epitope. The height of the bar above each residue reflects the likelihood of interaction with the paratope, where tall bars indicate a high likelihood and no to low bars indicate no or low likelihood.

3D surface visualization: CAXII homodimer (extracellular domain)

The Ab can bind to either of the two chains due to symmetry

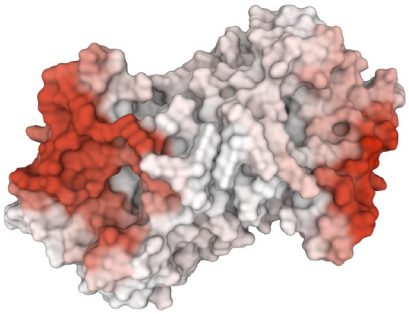
6RPS x-ray crystallography

Dark red = most contacts with Ab
White = no contacts



LENS[ai] Epitope Mapping

Dark red = highest confidence score
White = no predicted interaction



Case Study Summary

Goal:

Compare accuracy of LENS[ai] epitope mapping prediction with x-ray crystallography

- Crystal complex = 6RPS
- Input LENS[ai] = Fv 6A10 (predicted model) and unbound CAXII (1JCZ)

Method:

Prediction Reliability measures:
- ROC AUC
- Precision recall F1 score
- Contact map overlap

Outcome:

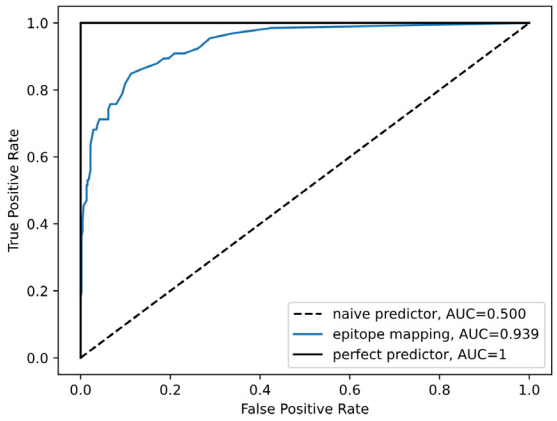
- LENS[ai] Epitope Mapping overlaps 71.2% with epitope size determined by x-ray crystallography
- AUC = 0.939 (outstanding)
- F1 = 0.726 (good)

Prediction performance

Evaluating LENS[ai] Epitope Mapping: classification performance metrics

1. ROC and AUC

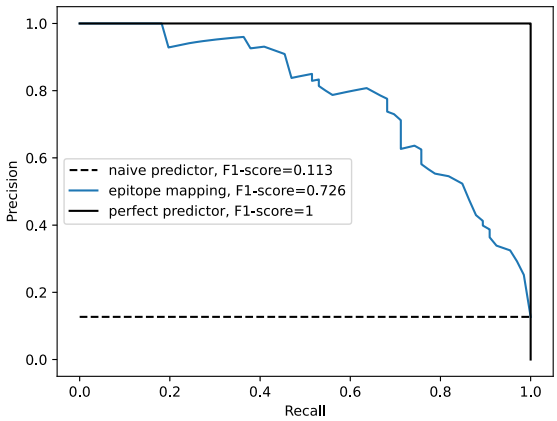
Fig. 1



- TPR (True Positive Rate or Recall) = TP/(TP+FN)
- FPR (False Positive Rate) = FP/(FP+TN)
- LENS[ai] Epitope Mapping AUC = 0.939
  In general AUC > 0.9 is considered outstanding discrimination*

2. Precision Recall

Fig. 2



- Precision = TP/(TP+FP)
- F1 = harmonic mean of precision and recall
- LENS[ai] Epitope Mapping F1 = 0.726
  In general F1 > 0.7 is considered good*

LENS[ai] Epitope Mapping predicts the likelihood of each target residue being part of the epitope. If the confidence score of an residue exceeds a set threshold, the residue is classified as part of the predicted epitope.

- **True Positive Rate (TPR or Recall):** measures the proportion of residues accurately predicted to be part of the "true" epitope (as per x-ray crystallography) to the total number of residues in the true epitope
- **False Positive Rate (FPR):** Measures the proportion of residues incorrectly predicted as part of the true epitope to the total number of residues not present in the true epitope determined by x-ray crystallography.
- **Precision:** Indicates how many of the residues predicted to be part of the epitope are correctly predicted.

The Receiver Operating Characteristic (ROC) curve (Fig 1), which plots TPR against FPR, demonstrates the prediction's ability to distinguish between epitope and non-epitope residues. An AUC of 1 indicates perfect prediction, while an AUC of 0.5 indicates performance no better than random guessing.

The precision-recall curve (Fig 2) visualizes the trade-off between precision and recall. A well-performing classifier trends towards the top-right of the graph, indicating it can accurately identify true epitope residues (high recall) while minimizing incorrect predictions (high precision). The F1 score, ranging from 0 to 1, combines precision and recall into a single measure, with a higher score indicating better performance.

*Applied Logistic Regression, 3rd edition, ISBN: 978-0-470-58247-3

Prediction performance

Evaluating LENS[ai] Epitope Mapping: contact map overlap

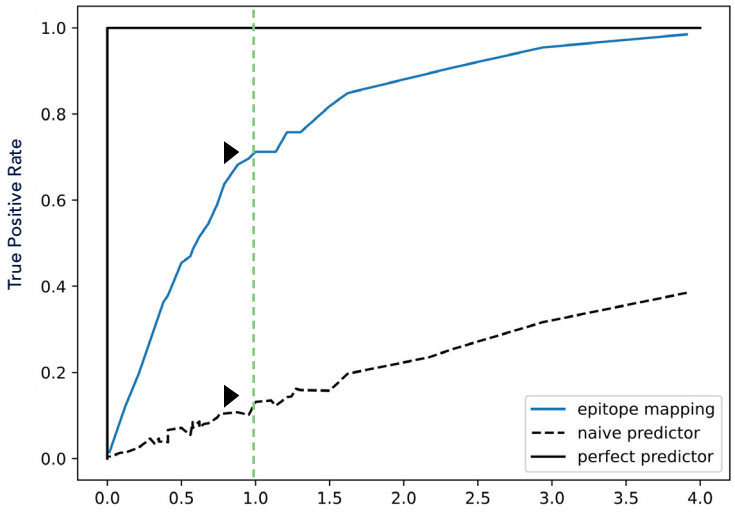
Applying a threshold on the LENS[ai] Epitope Mapping confidence score influences the size of the predicted epitope.

Figure 3 shows the true positive rate at varying sizes of the predicted epitope. This measure ranges from 0 for very small predicted epitopes (high threshold, predicted epitope doesn't contain any residues) to 1 for very large predicted epitopes (low threshold, predicted epitope completely encompasses the true epitope).

- At the threshold that yields a predicted epitope similar in size to the true epitope (value of 1 on the x-axis), LENS[ai] Epitope Mapping retrieves 71.2% (upper arrow) of the true epitope.
- In comparison, a random classifier predicting an epitope of the same size as the true epitope achieves, on average, an overlap of only 10% (lower arrow) with the true epitope on this target.

Figure 3

Taking the ground truth epitope size as a reference, LENS[ai] Epitope Mapping achieves an 71.2% overlap (- - - - - -).



Size of predicted epitope (number of predicted contact residues) divided by size of the ground truth (number of contacts identified by x-ray crystallography)

Conclusion

LENS[ai] Epitope Mapping combines a physics-based and machine learning approach. Coarse, global information is iteratively refined into atomic-level local contact information. The method is fast and applicable to various target types, including transmembrane proteins.

The LENS[ai] epitope mapping algorithm achieves high accuracy in determining the residues constituting the epitope. Using a machine learning predicted model of the Ab and an unbound PDB model of the target, the accuracy of the predicted epitope region on this dimeric target is notable. By taking the ground truth epitope size as a reference, LENS[ai] Epitope Mapping retrieves 71.2% of the ground truth residues in the epitope.

BioStrand
a subsidiary of IPA

For more information, contact: info@biostrand.ai