Mapping the 'Farnesylome'. Structure based predictions of Farnesyl-Transferase targets

Nir London

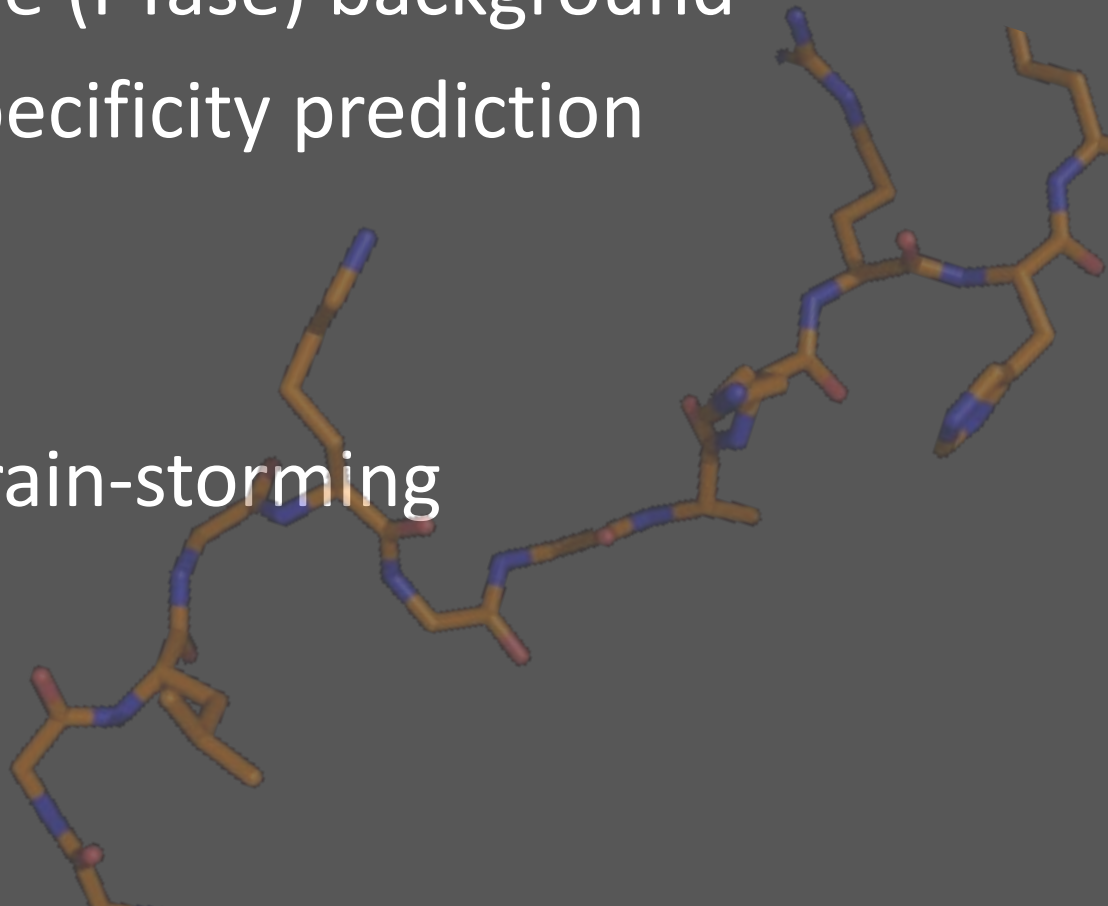
Ora Schueler-Furman

RosettaCON 2010

* This is NOT a renegade talk! Slides will be shared!
Data is unpublished but not confidential…

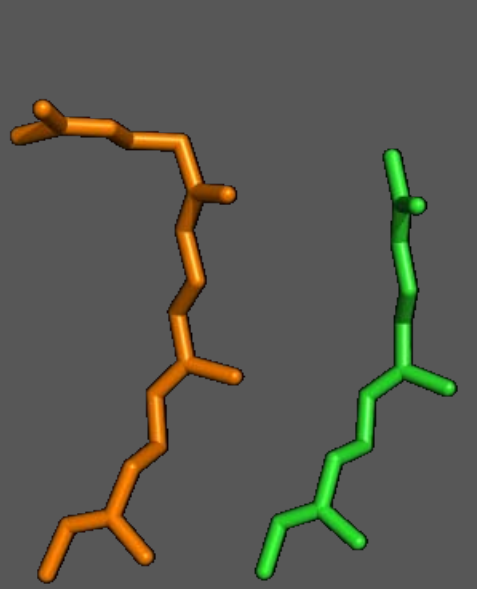
Outline

- This talk is about peptide binding specificity
- Farnesyl Transferase (FTase) background
- Our protocol for specificity prediction
- Good results
- Bad results
- Conclusions and brain-storming



Protein Prenylation

- The addition of a Prenyl group to a CYS residue on the protein
- Usually as means to direct it the membrane
- CaaX box enzymes:
  - Franesyl-Transferase (FTase)
  - Geranyl-Geranyl-Transferase (GGTase)



Geranyl
Geranyl

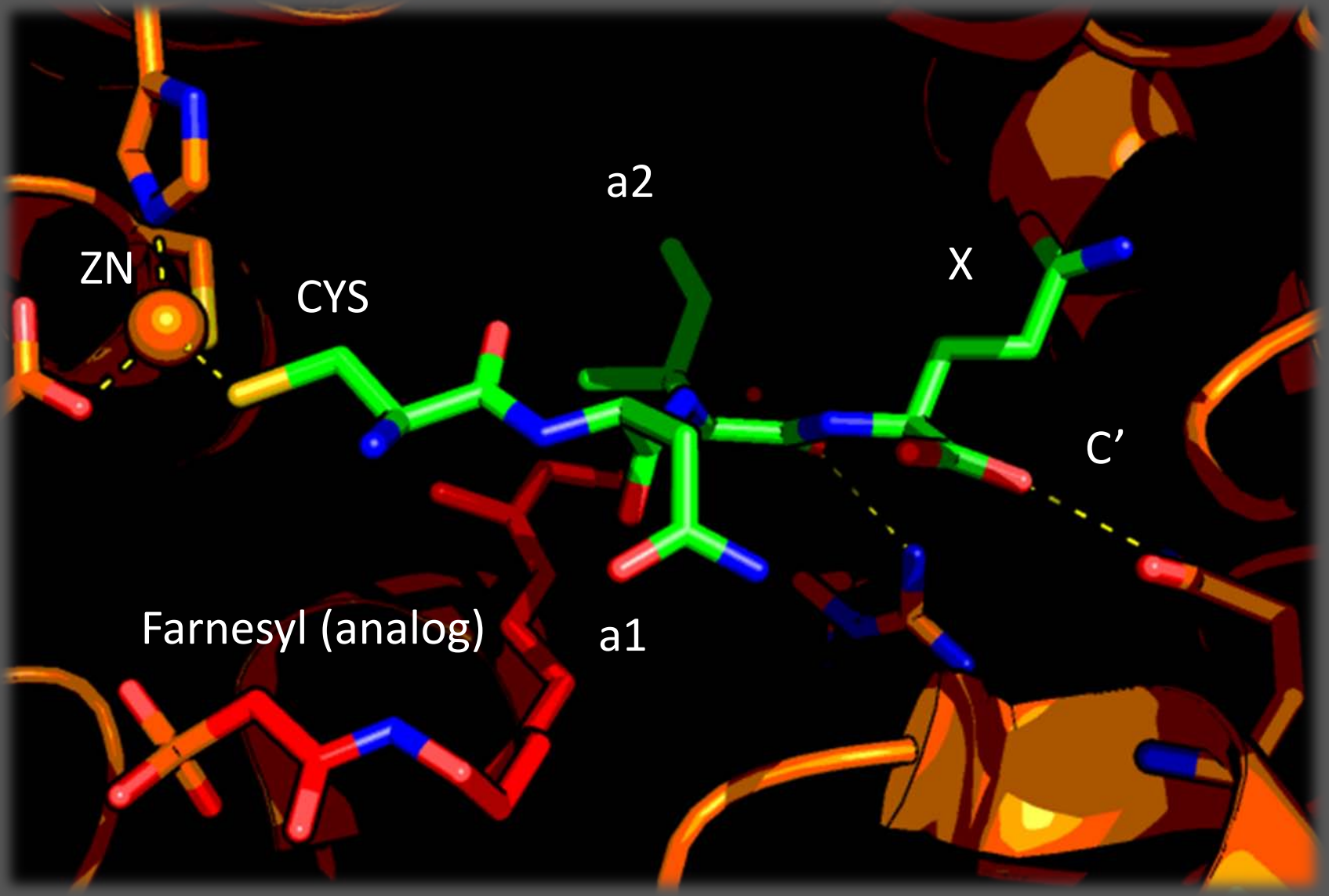
Farnesyl

CaaX Box Proteins

- Motif: Ca_1a_2X
  - Found at the target protein C'
  - a_1 & a_2 – usually aliphatic residues
  - X determines FTase/GGTase specificity

- Targets: many signal transduction related proteins (Ras-like, Rab, GTP-binding …)

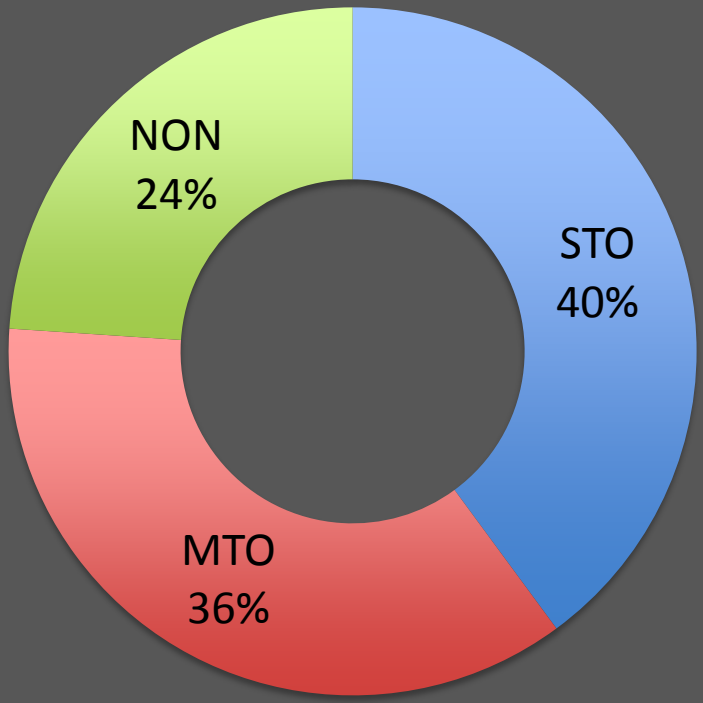
Overview of structure



Novel substrate peptides revealed

- Hougland *et al.* synthesized and characterized 213 hexapeptides of the form TKCxxx(C')

Do not undergo the reaction (51)



Undergoes reaction only under [E]>>[S] (85)

Undergoes reaction (77)

*Hougland et al., JMB (2009)

Can we distinguish MTO & NON?

Slow protocol:

Thread C' sequence onto structure

FlexPepDock

Generate 200 decoys in a standard run

Select top scoring

Using Score 12

Rescore

Sum over the peptide energy contribution (less the sequence reference energy term)

Can we distinguish MTO & NON?

Fast protocol:

Thread C' sequence onto structure

Pack peptide side chains using extra $\chi 1, \chi 2$ rotamers. FTase rotamers are kept fixed

Repack

Short minimization of all interface side-chains and peptide's backbone.

Minimize

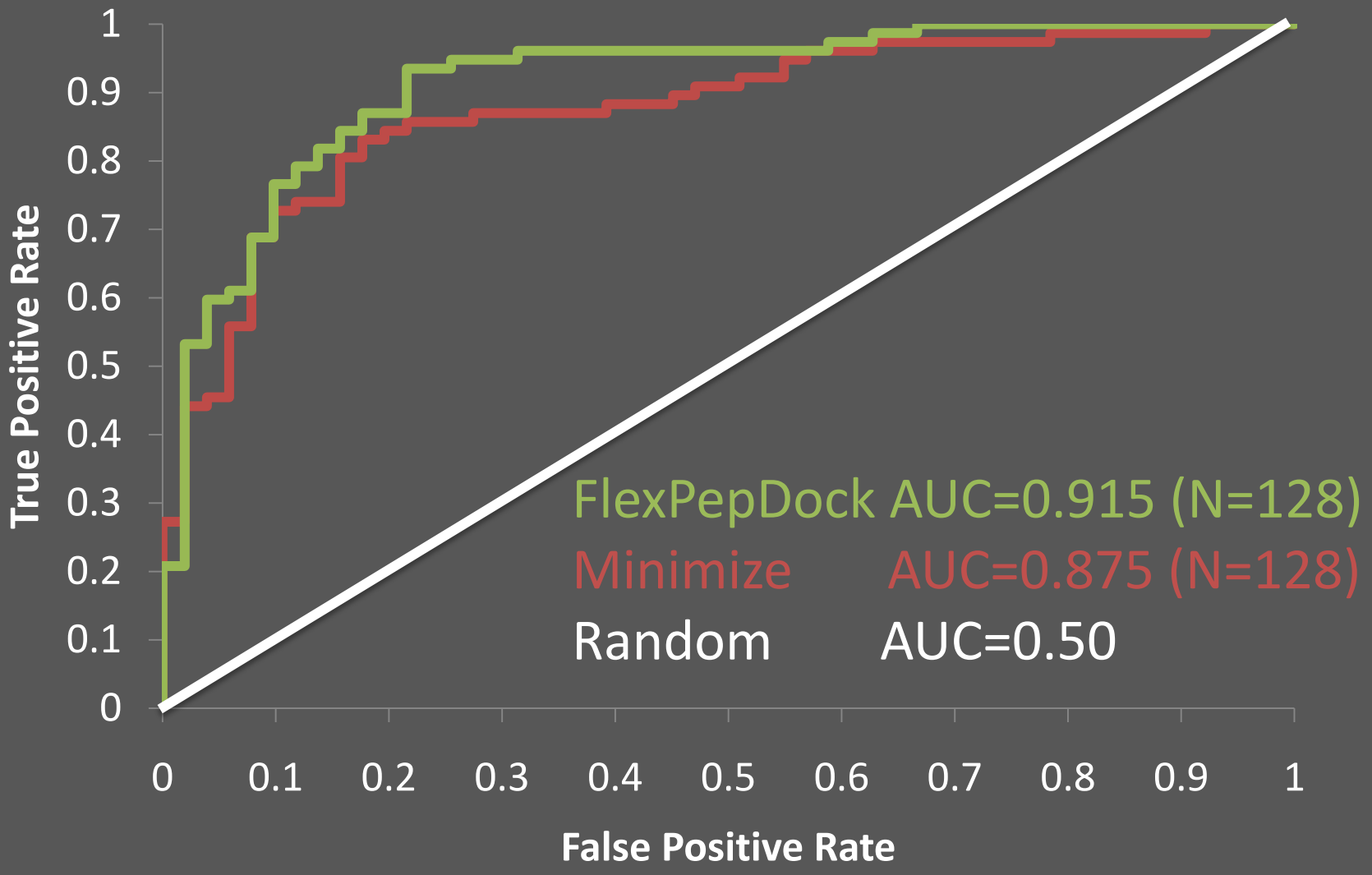
Rescore

Sum over the peptide energy contribution (less the sequence reference energy term)

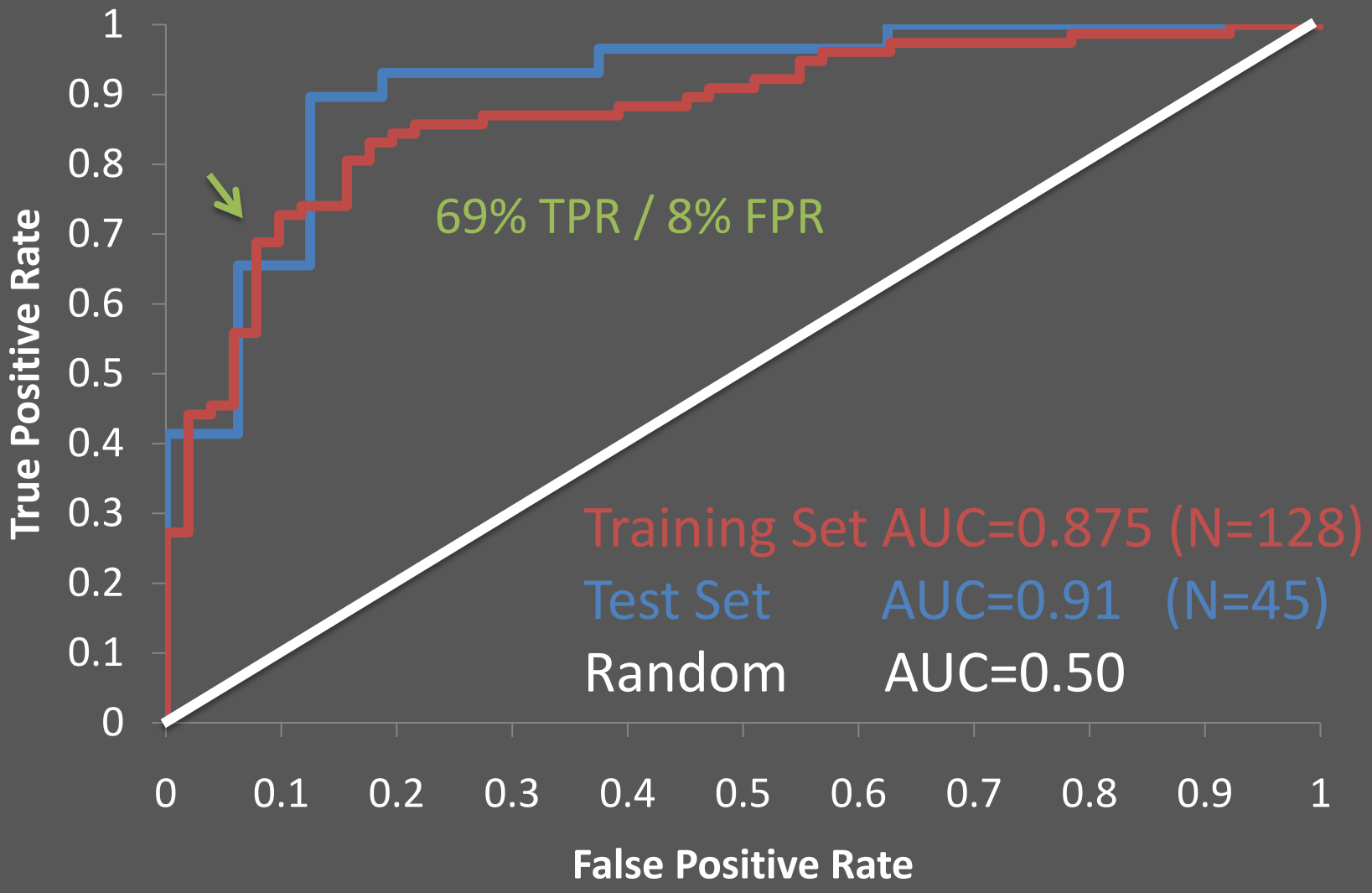
Important points (or: lessons learnt)

- Three constraints were enforced during the simulations
- Farnesyl (analog) molecule was included
- Re-scoring function was selected manually and not computationally optimized
- All these were learnt on a training set.

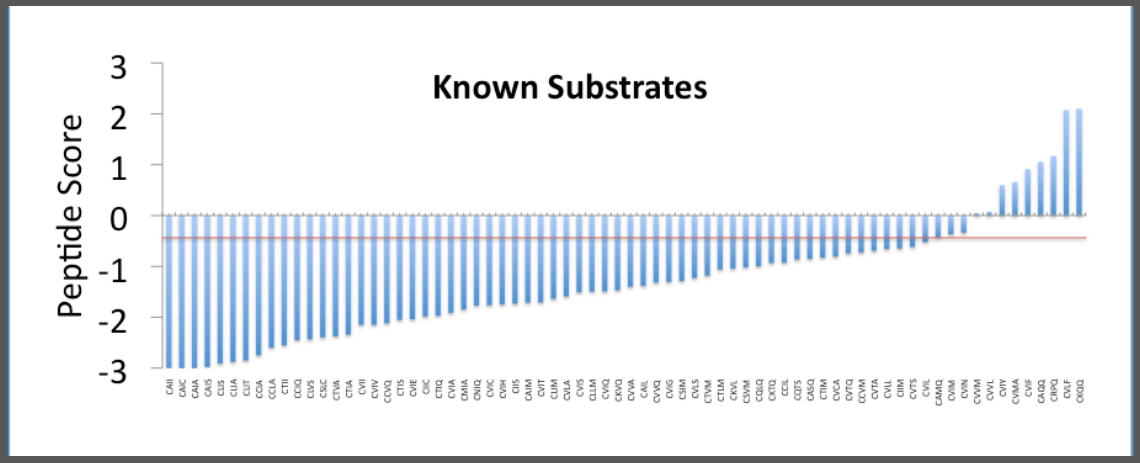
Good discrimination MTO/NON

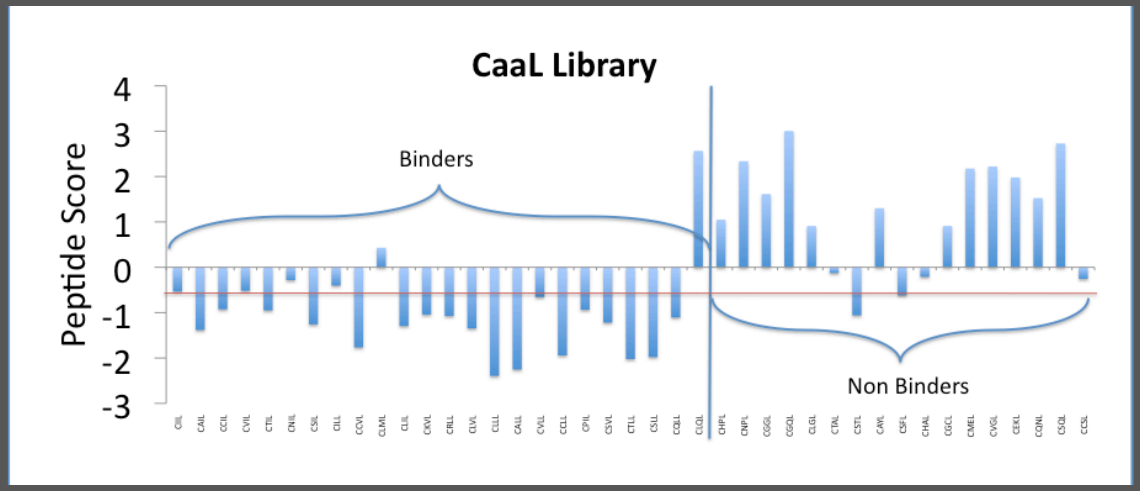


No over-fitting



Validation

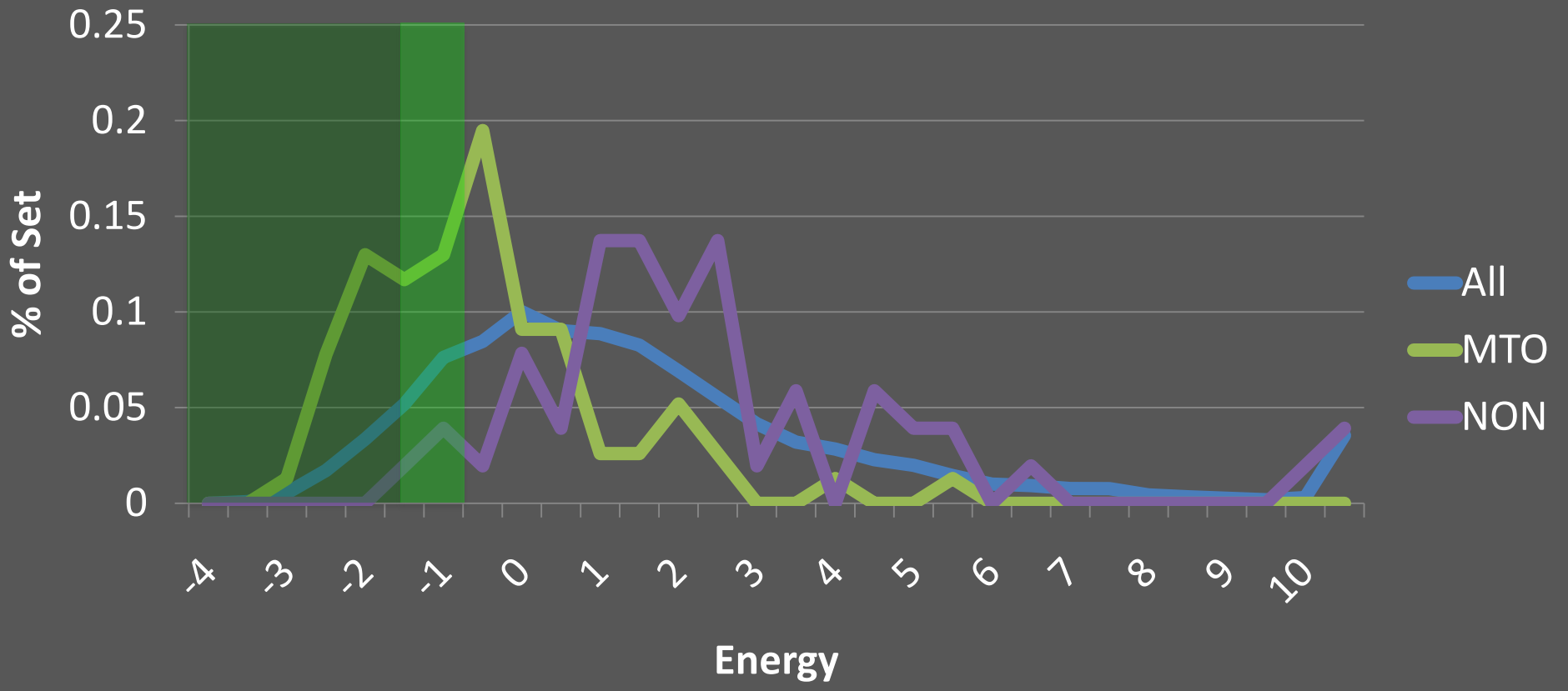




- 85% of known substrates are recovered
- 87.5% of synthetic CaaL library binders are recovered. 12% False Positive Rate.

FTase – sequence mapping

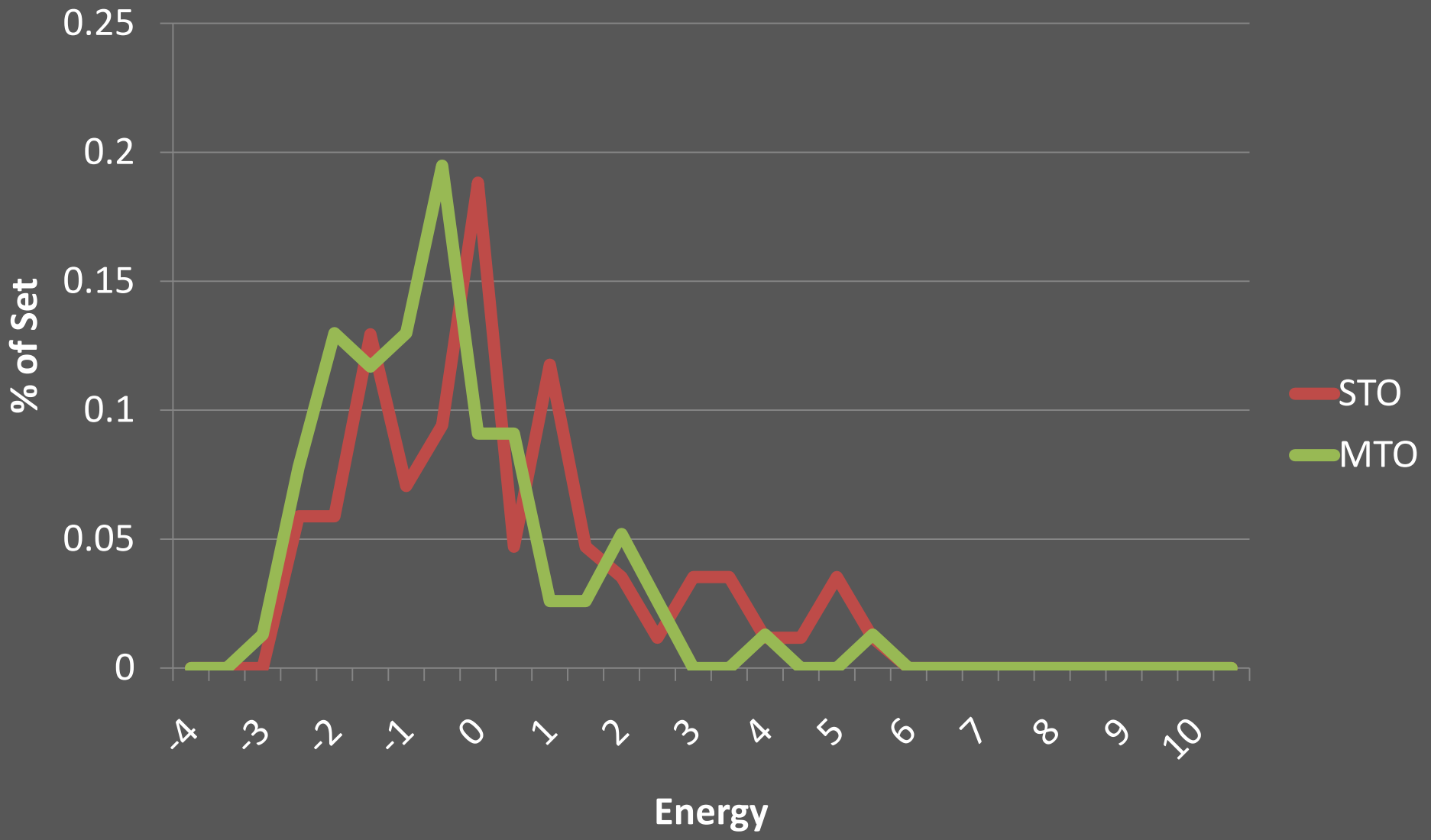
- Peptide scores were calculated for all possible 8000 combinations



Genomic Scan -> Putative novel targets

- Scan the Human genome for proteins with Cxxx at the C’.
- Rank according to their scores
- Top ranking proteins were either known FTase targets or looks like promising novel substrates

MTO vs. STO



Outperform sequence

- PrePS* – prediction of prenylation targets using sequence information. Not restricted to motif. Shows good results.
- Of the 167 protein sequences that passed the stringent threshold. 77 were undetected by PrePS
- Our protocol recovers 47% of STO peptides vs. 14% with PrePS
- Experimental validation of top binders is underway**

** Maurer-Stroh et al., Genome Biology (2005) ** Carol Fierke Lab U. Michigan*

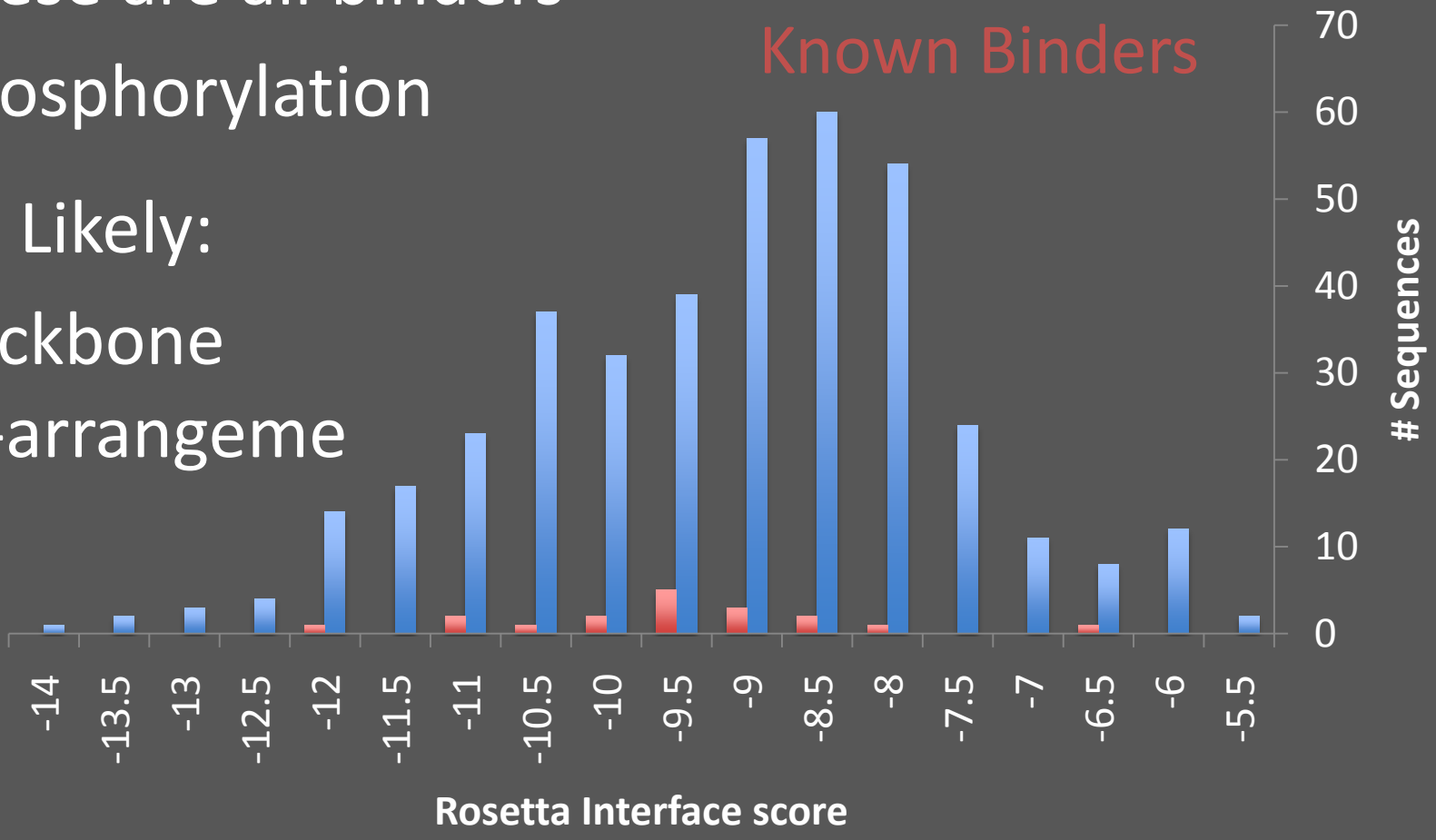
Modulating Protein Degradation

- Apply the same methodology for other "motif reading" proteins
- DS_pGxxS_p Motif

Cold Spring Harbor Perspectives in Biology (2010) Kanarek, London, Schueler-Furman and Ben-Neriah

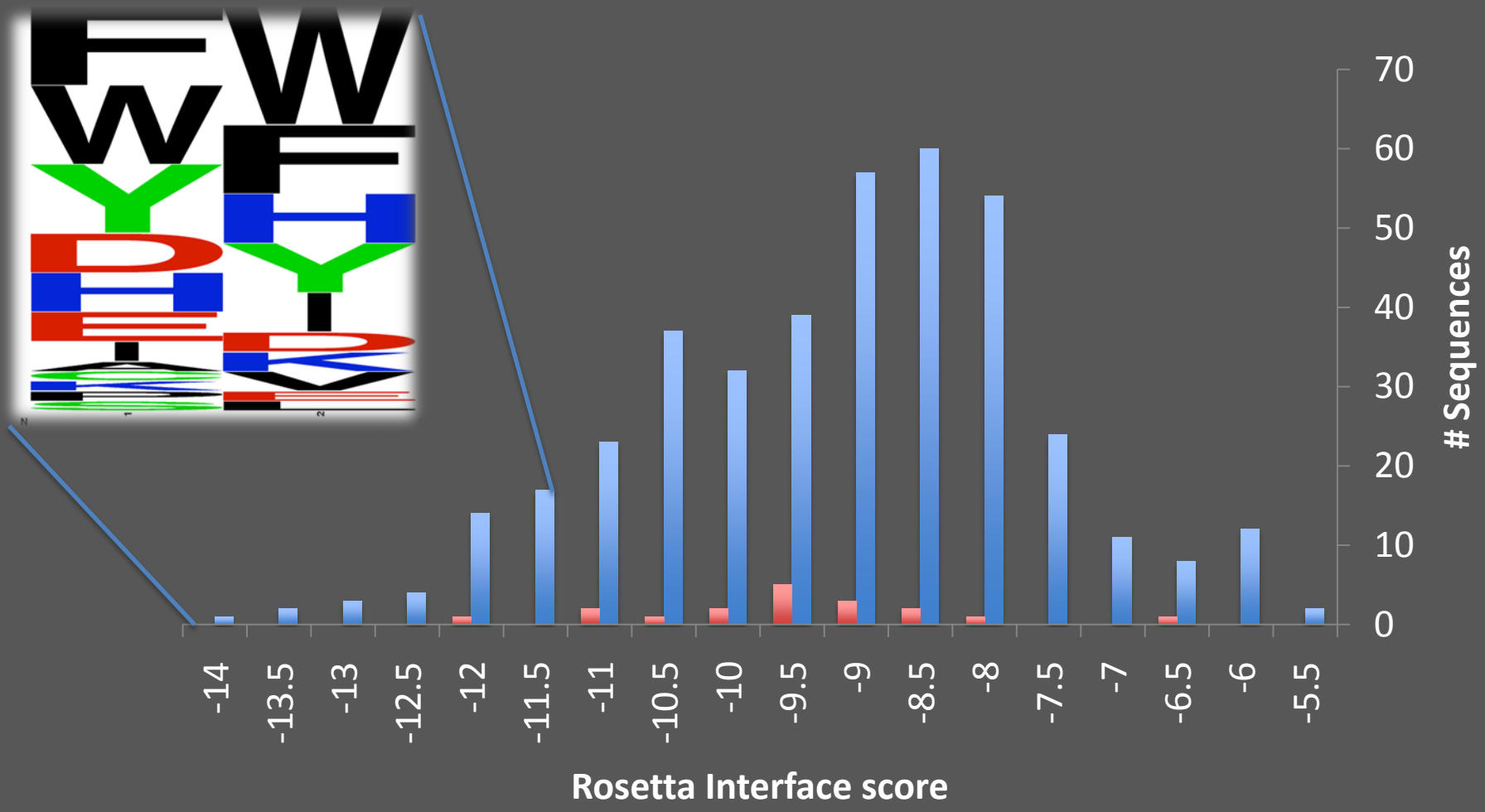
Failure 1

- Possible Excuses:
  - These are all binders
  - Phosphorylation
- More Likely:
  - Backbone re-arrangement



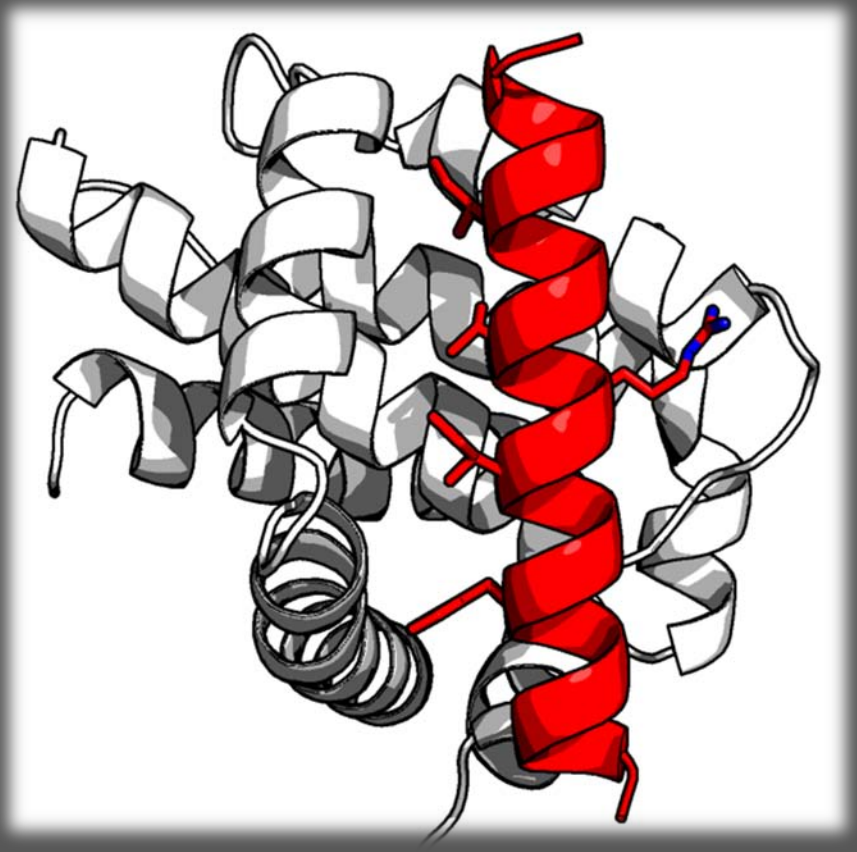
Failure 1

- Energy function bias



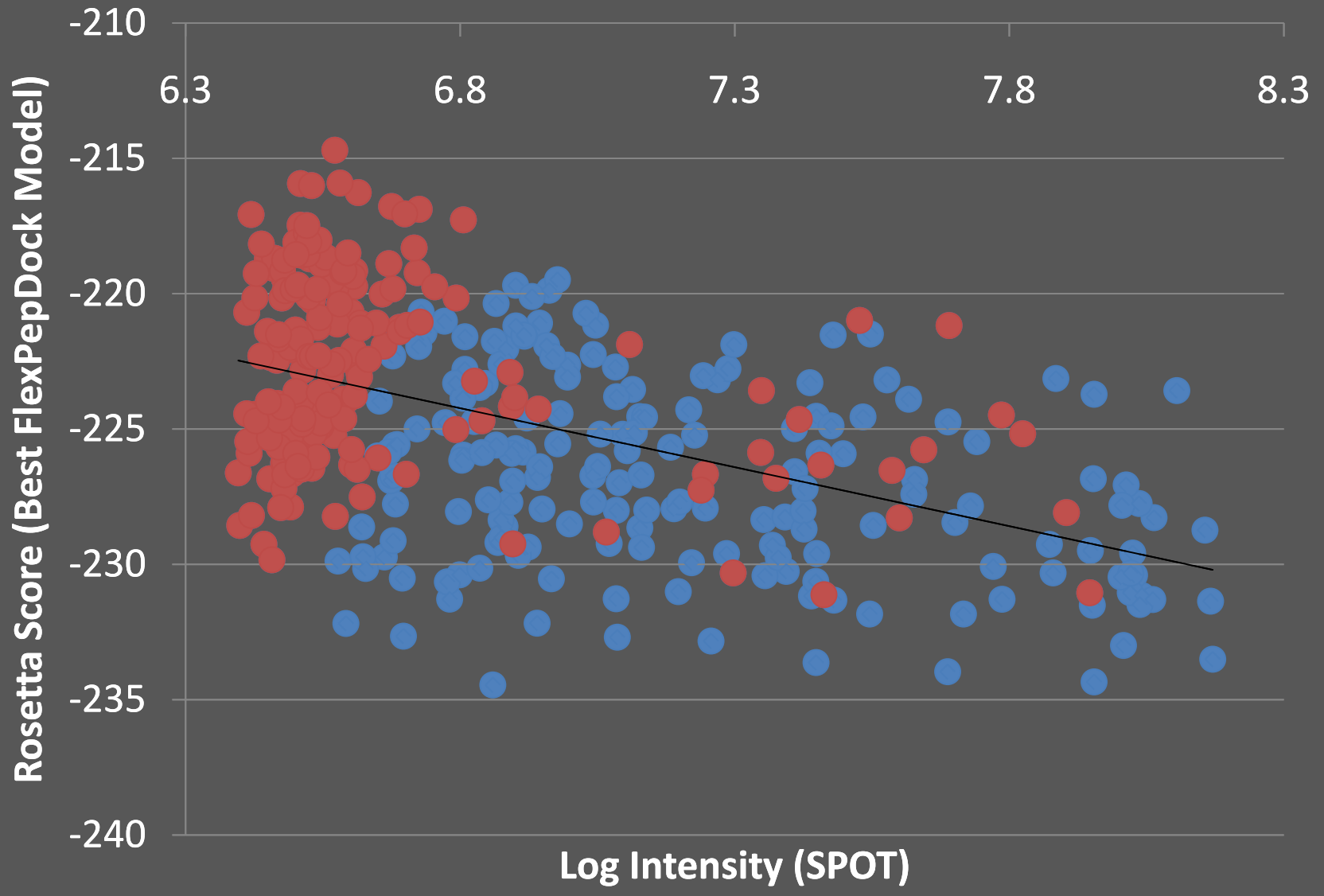
Failure 2

- MCL1/BH3 binding
  - Major role in apoptosis regulation
  - Peptide binds as a helix
  - SPOT binding measurements available



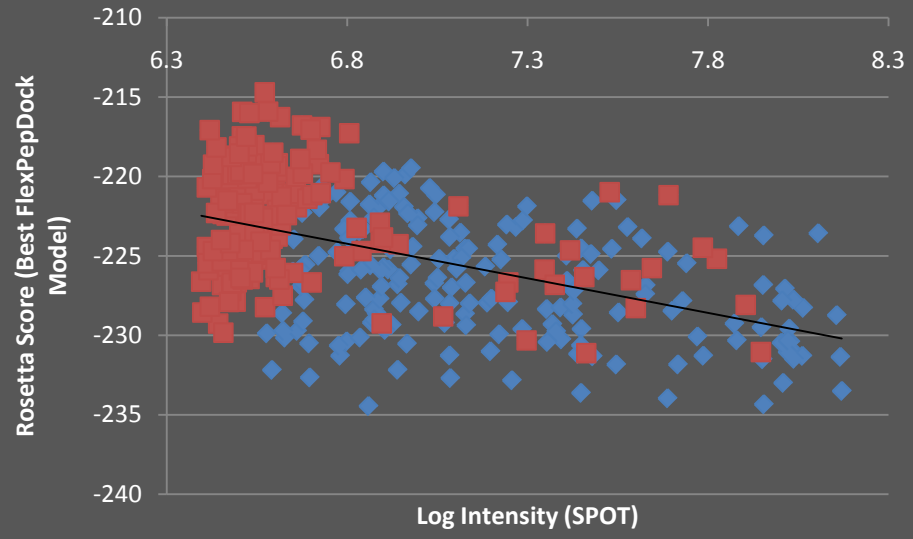
** Dutta et al., JMB (2010)*

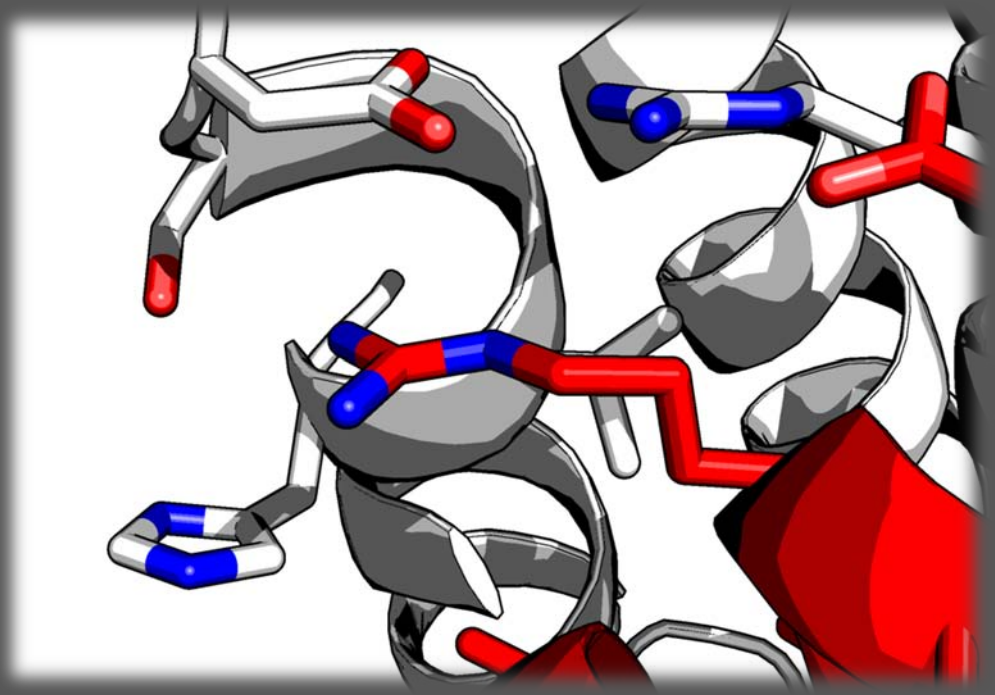
Poor Correlation



Arg-Stacking

- Arg->Asp deleterious mutation undetected by energy function.
- Even when this is taken into account there is still poor correlation in predicting “regular” hydrophobic mutations
- Backbone re-arrangement isn’t likely in this case.





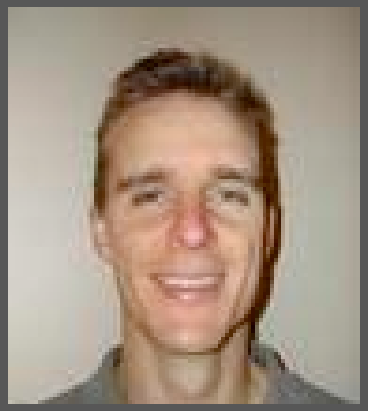
Conclusions

- FTase was a good system to work on
- For specific systems one can learn a lot using this methodology
- FlexPepDock can improve sampling (ranking)
  - Aligns with the fact that modeling bb flexibility improves design

- Energy function needs work
- What makes specific systems work ? Constraints ?

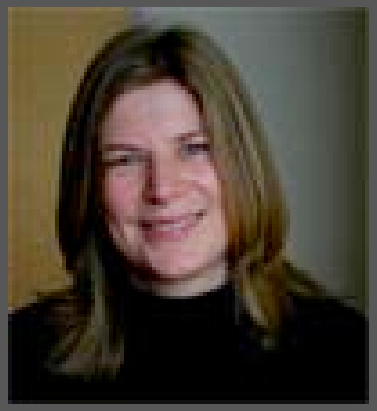
Peptide Specificity in Rosetta World

System: PDZ
Performance:
R=0.66 for $\Delta\Delta G$ calculations
Kaufmann et al.,
J. Mol Model (2010)



System: Any
Performance: GOOD!
King & Bradley. Proteins
2010(accepted)

System: HIV protease
Performance:
Identifying drug resistance mutations and recovering specificity of cleavable peptides
Chaudhury & Gray,
Structure (2009)



System: PDZ
Performance: Recovering Phage display data.
Smith & Kortemme
JMB (2010)

Acknowledgements

- Barak Raveh
- Dana Movshovitz-Attias
- Ora

Rest of the lab:
Assaf Faragy,
Dan Reshef, Lior Zimmerman

Rosetta Community

Funding:
Converging Technologies Scholarship
ISF, GIF Young investigator, NIH

Our Lab