

# **Rational design of antibodies to target amyloid fibril structure in Alzheimer's disease**

**Kath Nicholls**

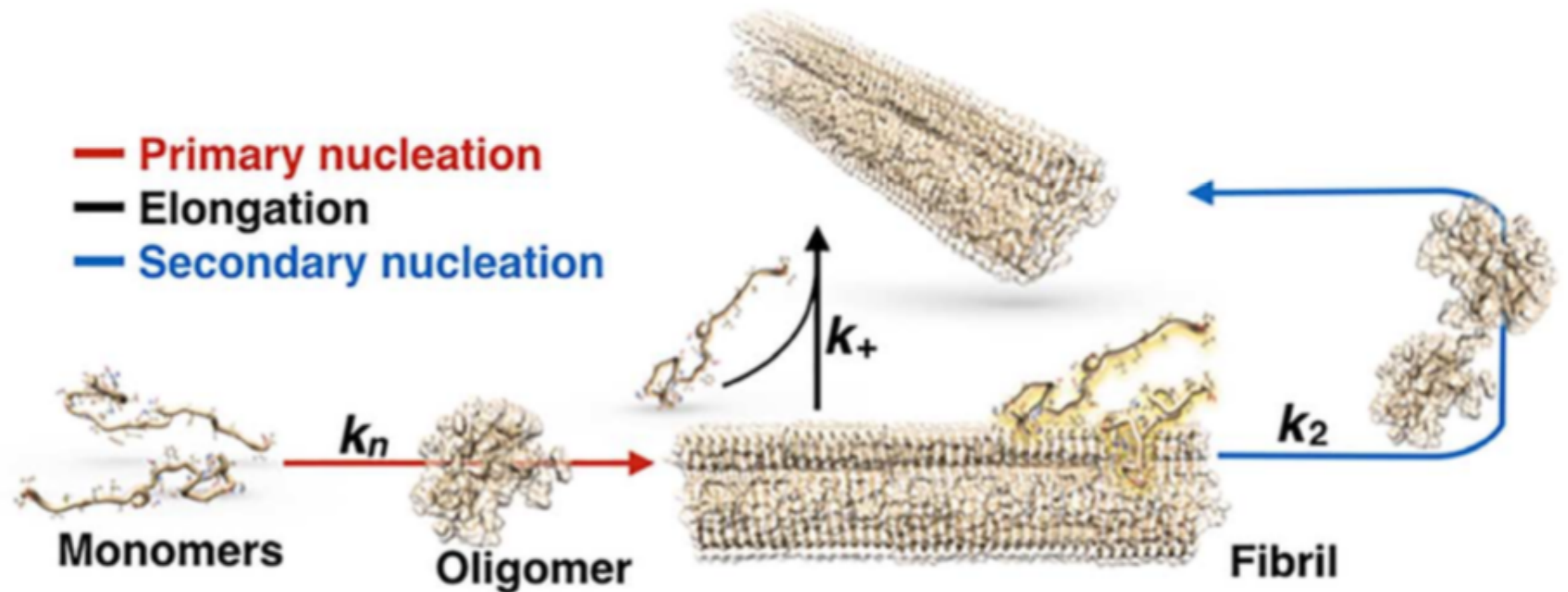
**Supervised by Michele Vendruscolo, Paul Kirk and Pietro Lio'**

# Overview

- Motivation - Alzheimer's disease, therapeutics
- Method - dataset, features, statistical models
- Results and further work

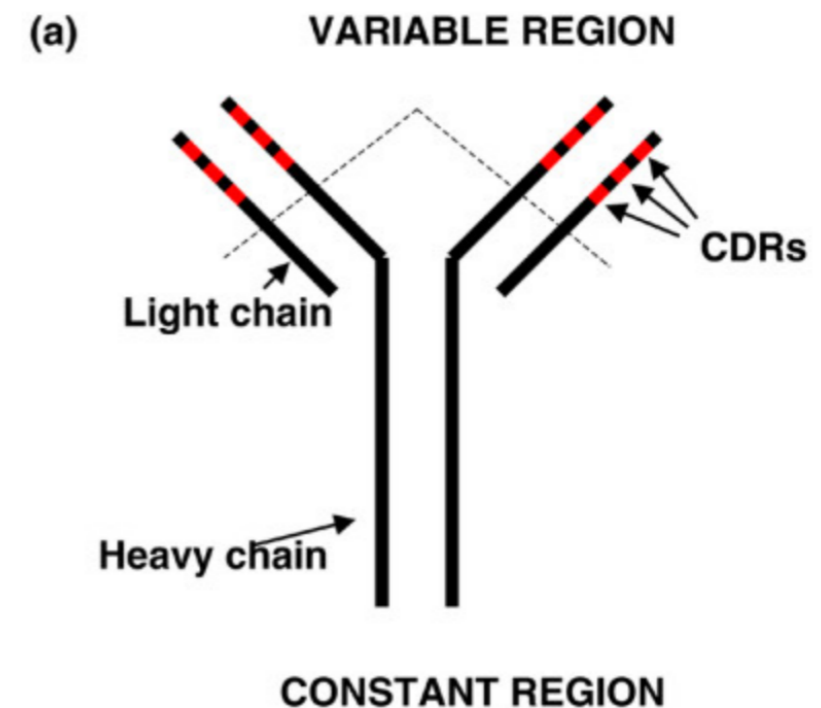
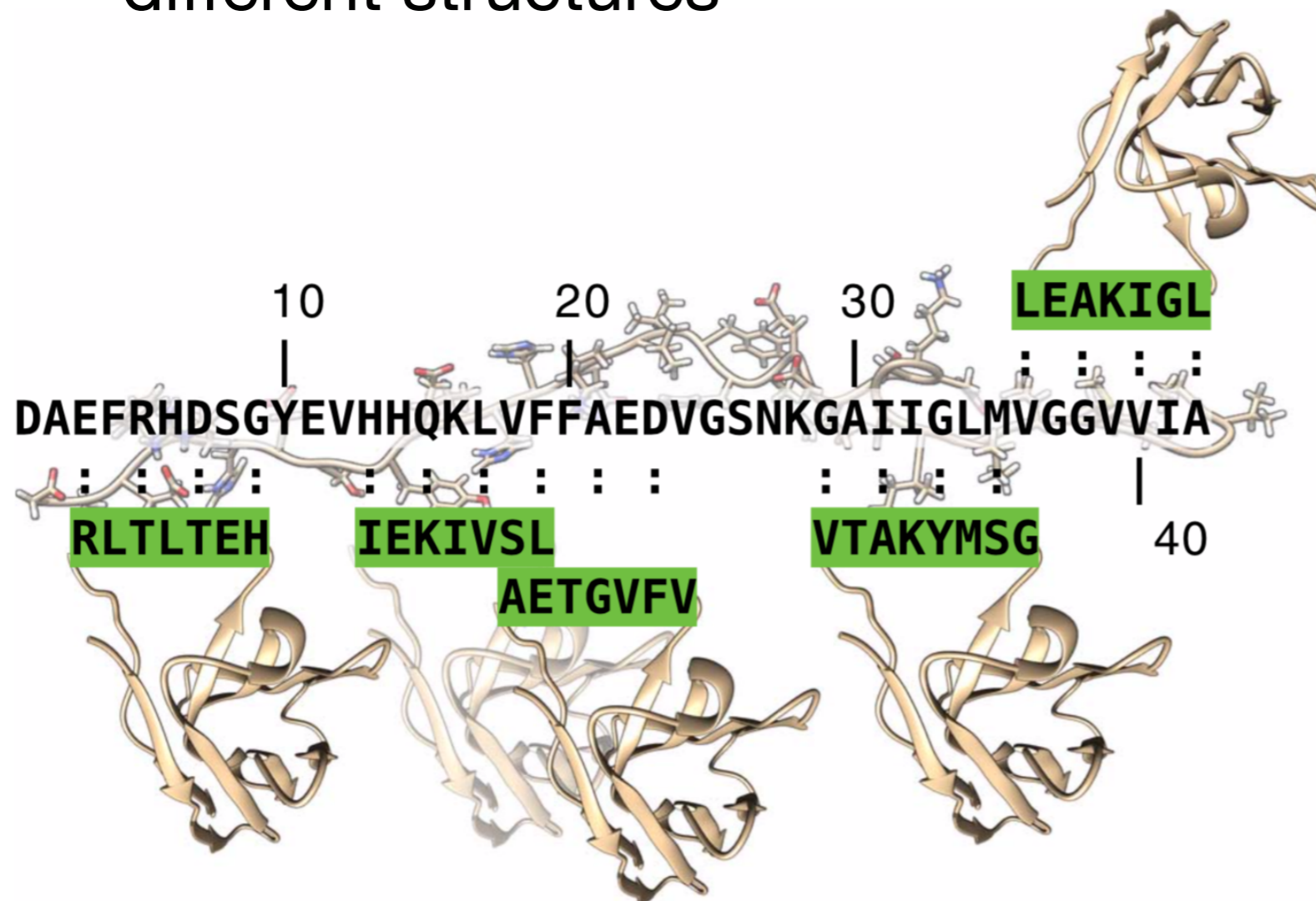
# Aggregation cascade in Alzheimer's Disease

**A**



# Existing methods can target sequences

- Modular approach targets linear fragments of AB42
- Want to target structures - different forms will have different structures



# Other applications

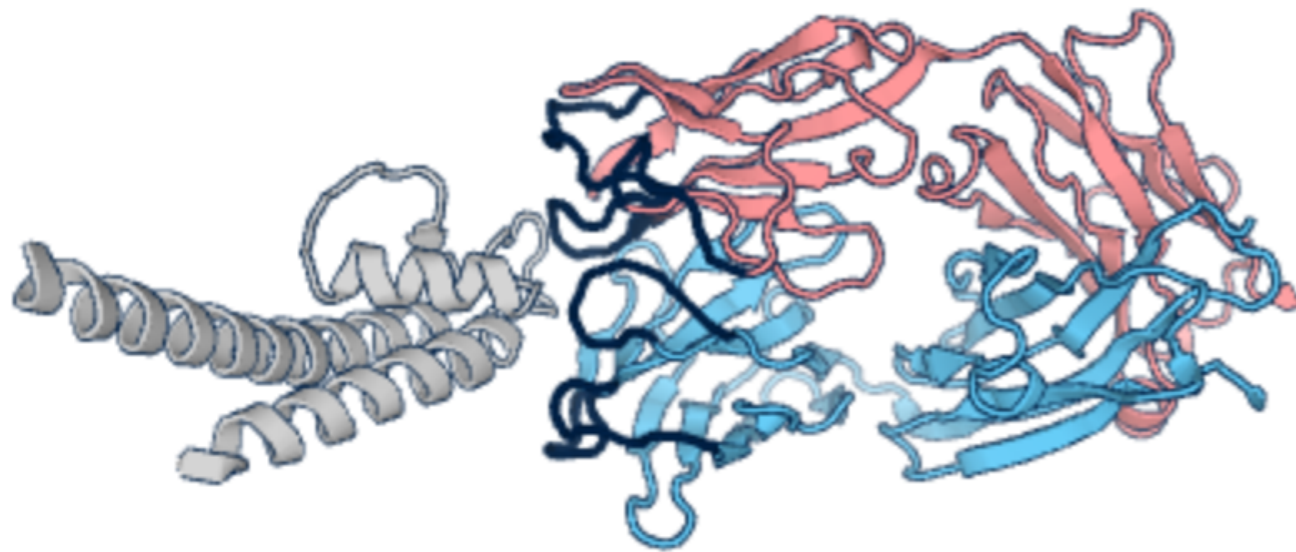
- Antibodies used as drugs
- Many other applications in misfolding diseases and elsewhere

# Overview

- Motivation - Alzheimer's disease, therapeutics
- Method - dataset, features, statistical models
- Results and further work

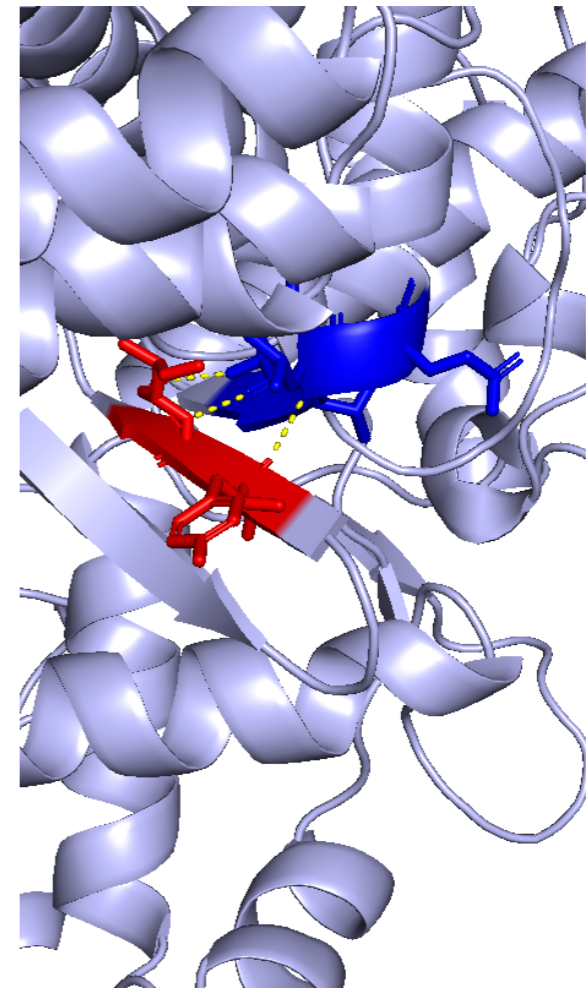
# Dataset construction

**Real aim:  
predicting antibody binding**



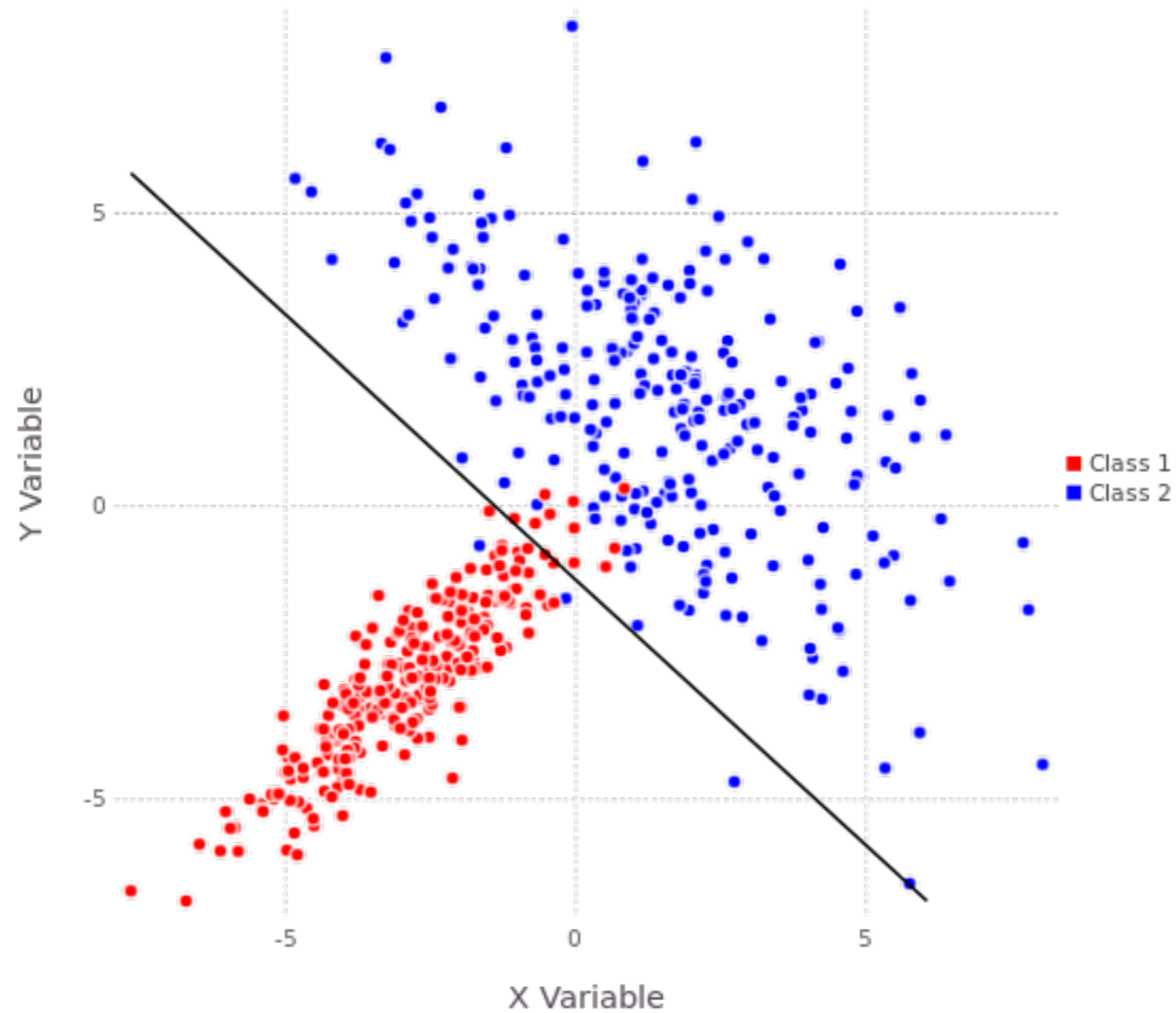
**~ 3000 antibodies bound to antigens**

**Larger dataset:  
general binding of fragments**



**~ 40,000 proteins  
each with ~ 1,000 internal interactions**

# Logistic regression





# Key challenge: examples of 'no binding'

- PDB gives only examples of binding (positive examples)
- Require positive and negative to train a model

# Generating examples of 'no binding'

- Generated by swapping CDR-like fragments
- Only if CDR-like fragments dissimilar and targets dissimilar



$< 0$ , so swap CDRs to create a negative example

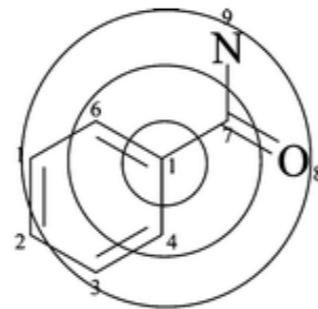


# Features

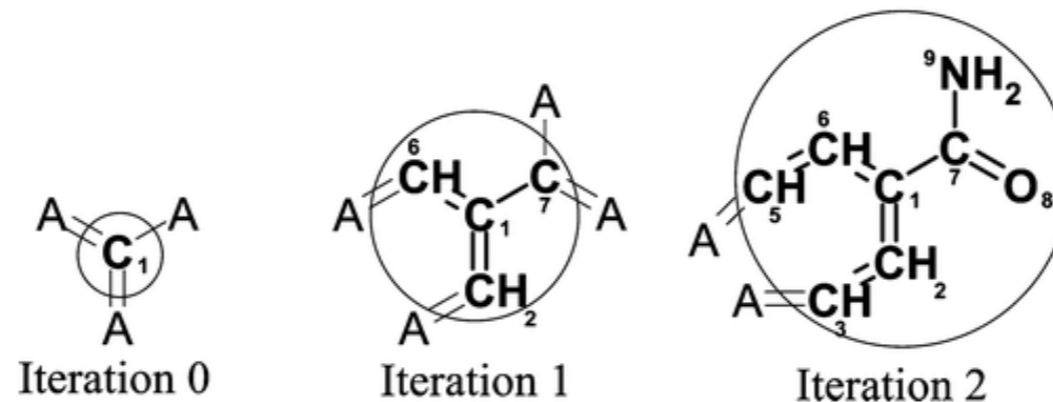
- Sequence-based

CDR	Target	C	S	P	G	...	L	V	F	C	S	P	G	...	L	V	F
VCLL	SLSS	1	0	0	0	...	2	1	0	0	3	0	0	...	1	0	0
GCLV	PSVFPL	1	0	0	1	...	1	1	0	0	1	2	0	...	1	1	1

- Structural



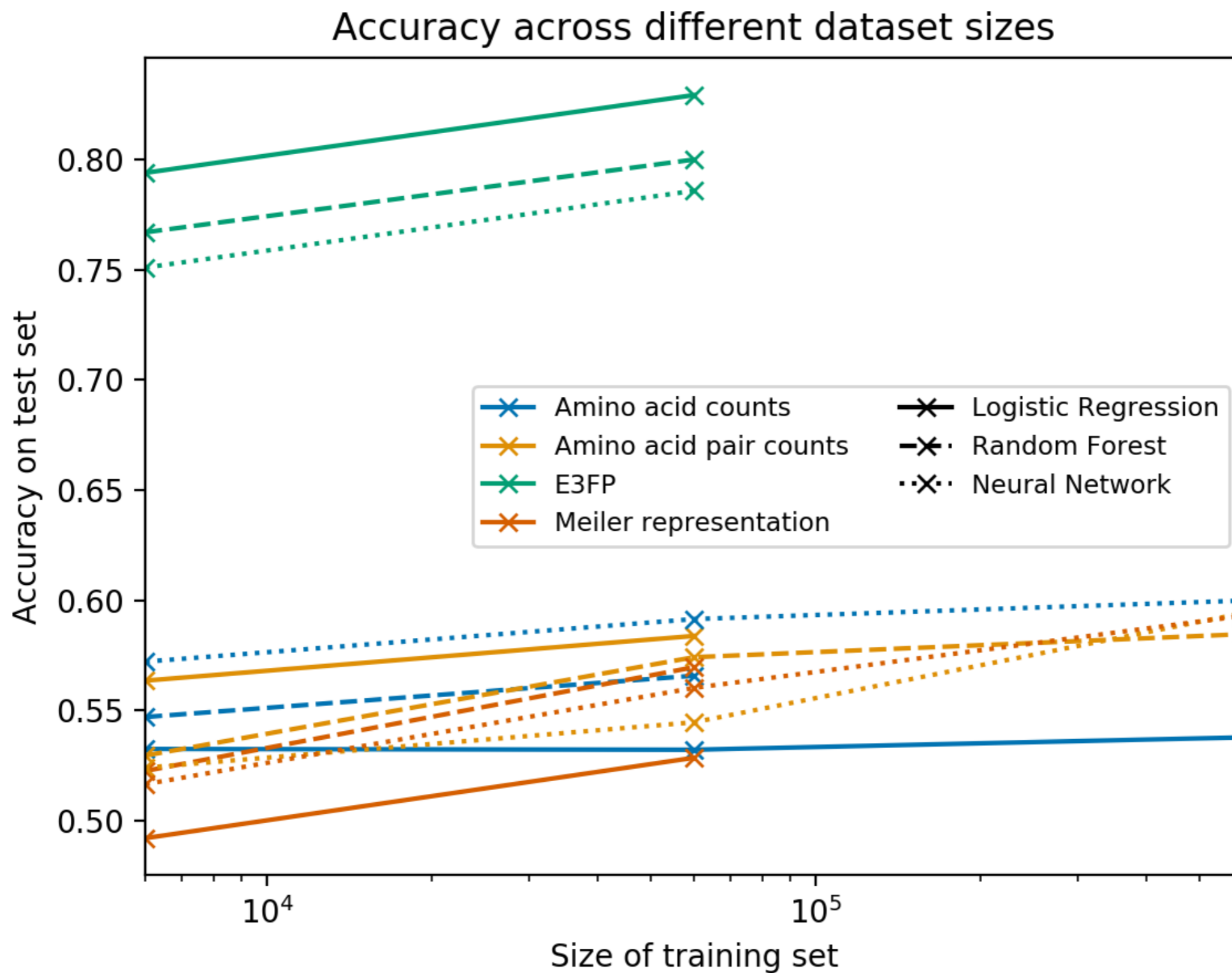
Considering atom 1 in benzoic acid amide



# Overview

- Motivation - Alzheimer's disease, therapeutics
- Method - dataset, features, statistical models
- Results and further work

# Comparing representations



# Generating examples of 'no binding'

VCLL  
SLSS

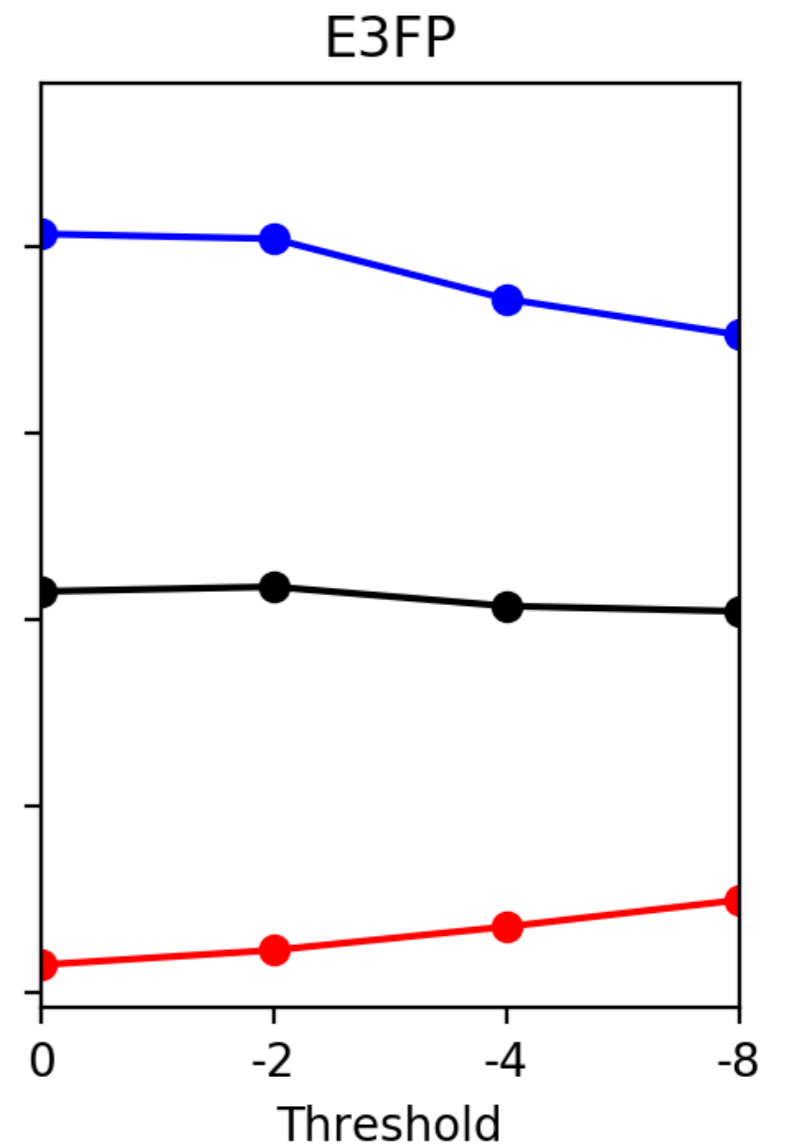
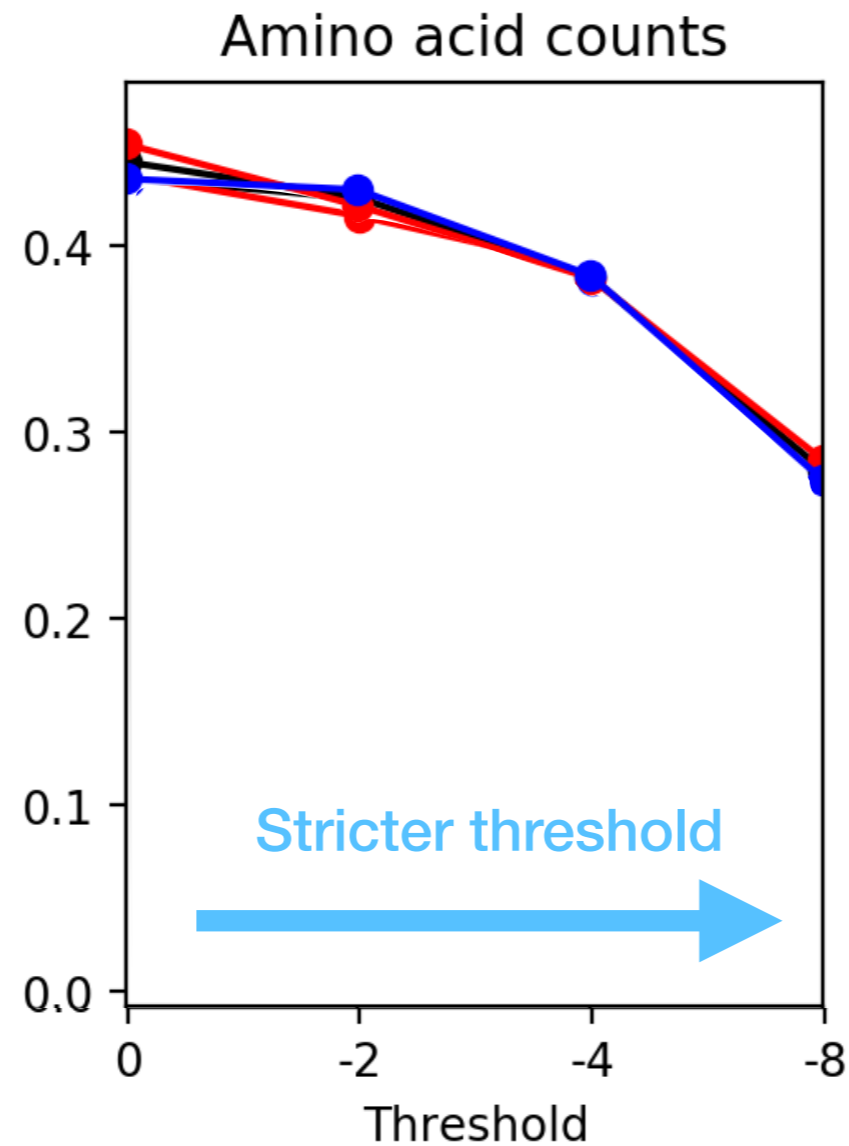
--VCLL  
NNAS--  
score: -5

SLSS-  
A-TAR  
score: 1

$-5 + 1 = -4$

VCLL  
ATAR

NNAS  
ATAR

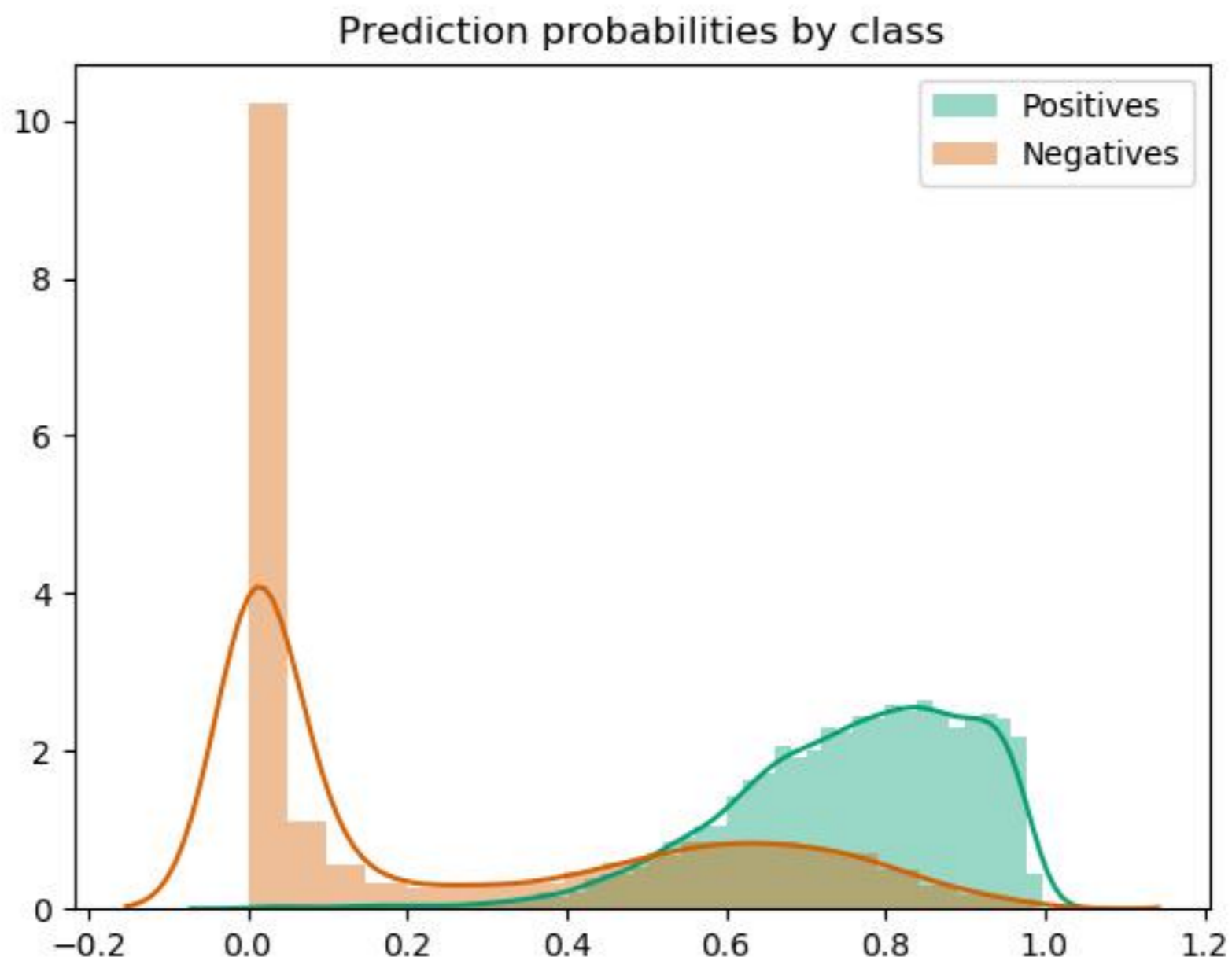


● Error    ● False Negative Rate    ● False Positive Rate

# Further work

- Interpreting the model
- Validate predictions in 'real' problem of antibody binding
- Validate experimentally

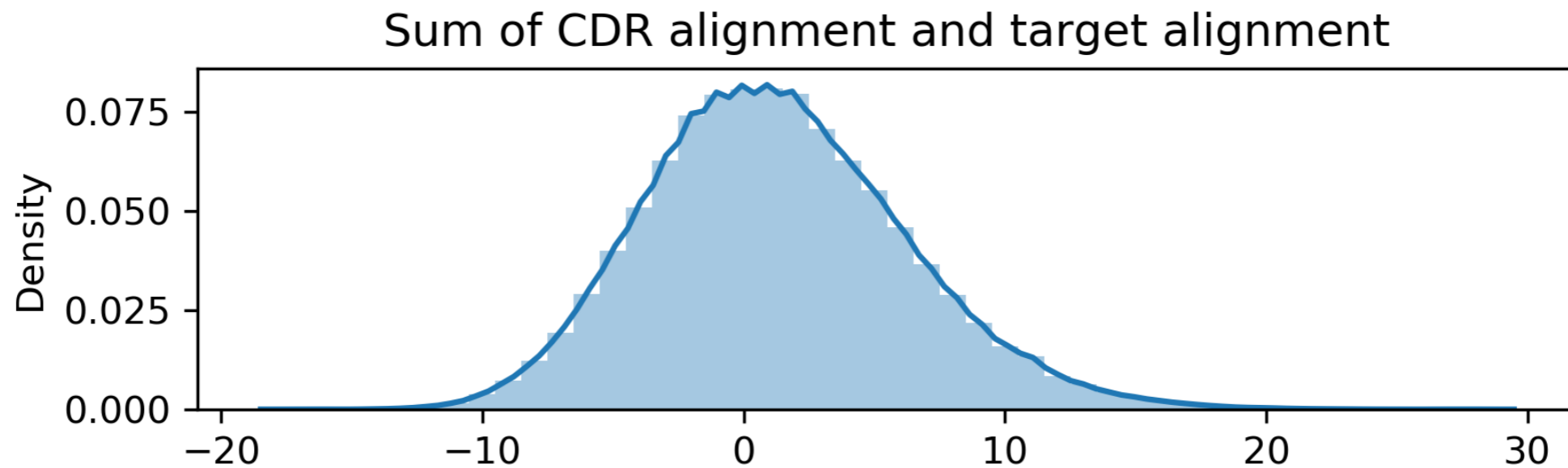
# Classification threshold



Structural fingerprint representation, logistic regression model

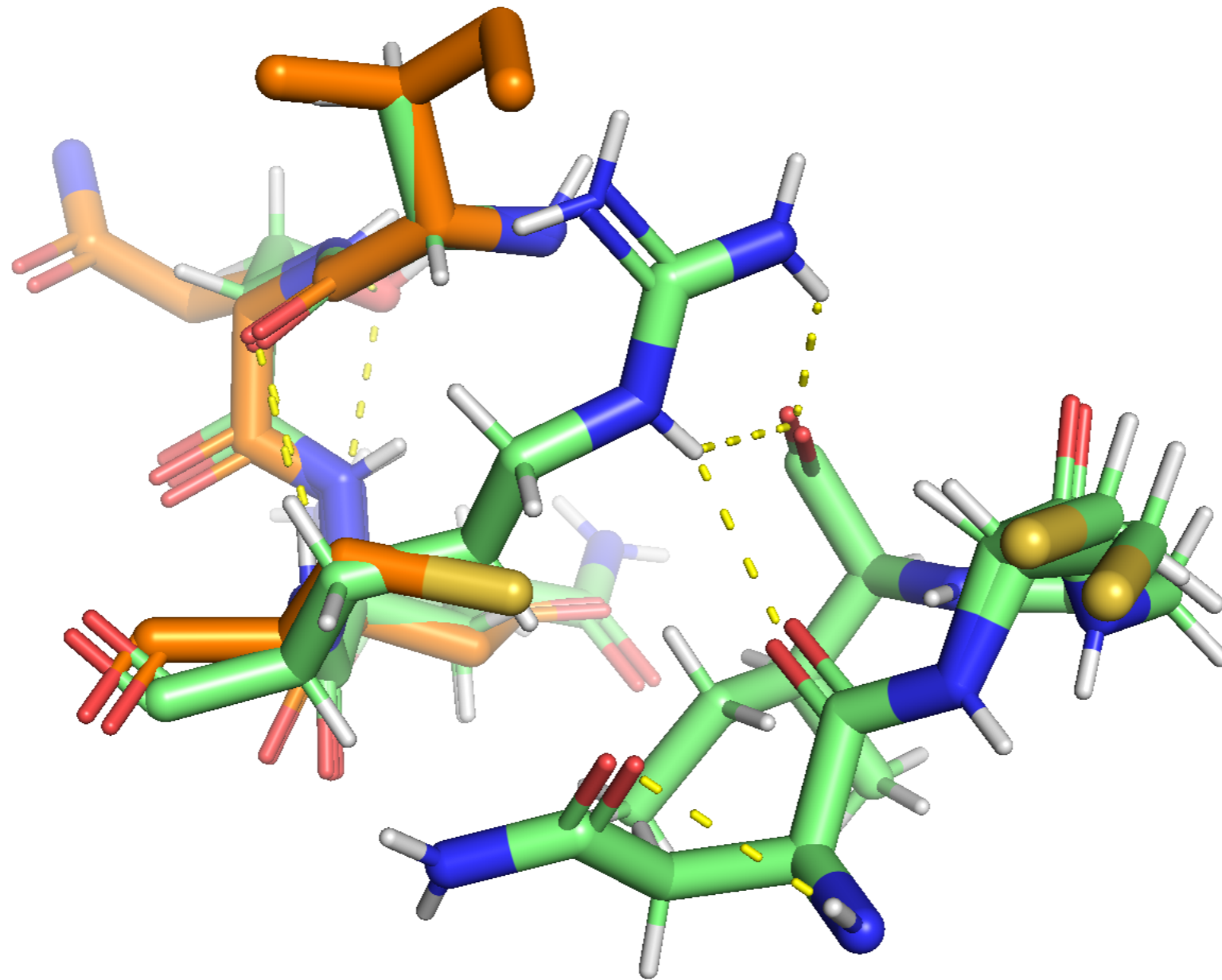


# Negative generation threshold



Threshold, $m$	Retention rate
0	38.1%
-2	22.8%
-4	11.4%
-8	1.3%

# Superimpose CDRs to create negatives



*Created using Pymol*