Antibody Structure: Prediction and Design

Rocco Moretti In the lab of Jens Meiler





UNIVERSITY

Center for Structural Biology and Institute of Chemical Biology Departments of Chemistry, Pharmacology, and Biomedical Informatics

10 June 2018

Rosetta

- Rosetta consists of multiple modules: protein folding, comparative modeling, ligand docking, protein design, antibody/antigen interactions, etc.
- Rosetta is developed in a consortium of twelve laboratories by around 50 developers
- Rosetta is free for academic us; user guide and tutorials are available
- PyRosetta is a python interface that allows integration with Pymol
- FoldIt is the better video game for you and your kids
- Rosetta@home uses your computer for our research

Bender, B. J., et al. "Protocols for Molecular Modeling with Rosetta3 and RosettaScripts."" Biochemistry 2016.







RosettaCon 2017, Leavenworth, WA, USA

What is an antibody? (Structurally)





What is an antibody? (Structurally)





Antibody terminology: Ig



"Immunoglobulin" or "full length antibody"

Antibody terminology: Fab





Antibody terminology: Fv





What is an antibody? (Structurally)





The immunoglobulin domain fold





Anthony S. Serianni

Complementarity determining regions





Recognition of antigens



Protein-protein interactions are determined by shape and physicochemical properties of solvent exposed loops, such as CDR loops.



Antibody diversification is achieved primarily through four mechanisms



... and affinity maturation.

V(D)J recombination





Junctional diversity





V(D)J gene sequences mapped to CDRs





Sequence/structure relationship





IMGT DomainGapAlign

Prediction of Antibody Structures: Match Sequences with Known Structures





Canonical Structures:

Morea, Tramontano, Rustici, Chothia and Lesk, J Mol Bio 1998 North, Lehmann and Dunbrack, J Mol Bio 2011 Adolf-Bryfogle, Xu, North, Lehmann and Dunbrack, Nuc Acids Res 2015

Very Long HCDR3 Sequences are Rare but Present





HCDR3 torso domain conformations





Finn JA, et al. PLOS ONE. 2016

Weitzner et al. Structure 2015; 23:302-311.

Bulged and non-bulged dihedral angle measurements differ





Bulged (n = 218)
 Non-bulged (n = 38)

Finn JA, et al. PLOS ONE. 2016

Bulged torso structures share similar sequences





The bulged torso sequence motif is germline encoded







Germline Consensus

Antibody Structure Prediction

- RosettaAntibody
 - http://rosie.rosettacommons.org/antibody/
- Kotai Antibody Builder
 - https://sysimm.ifrec.osaka-u.ac.jp/kotaiab/
- PIGS (Prediction of ImmunoGlobulin Structures)
 - http://circe.med.uniroma1.it/pigs/
- ABodyBuilder
 - http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/Modelling.php

Antibody/Antigen Docking

Coarse Grained Docking

Docking Refinement

Binding Types

Lock and Key:

$$(Ab + Ag) \Rightarrow (Ab - Ag)$$

Conformational Isomerism:

Induced-Fit:

Bax et al. Front Immunol. 2012 3:229

Antibody Specific Issues: V_H-V_L

Marze et al, Protein Eng Des Sel. 2016;29(10):409-418

Caveats

 Glycosylation can be an issue.

- Rigid body docking may be highly loop dependent
 - Working with ensembles of models may help

Antibody Docking

- Rosetta SnugDock
 - http://rosie.rosettacommons.org/snug_dock/
- PIPER with the ADARS potential
 - https://www.schrodinger.com/piper
- ZDOCK with Antibody i-Patch potential
 - http://zdock.umassmed.edu/

(Inverse) Protein Folding Problem Holy Grail of Comp. Struct. Biology

- Given a protein's AA sequence, what is its 3-dimensional fold , and how does it get there?
- Assume 100 conformations for each amino acid in a 100 amino acid protein ⇒ 10²⁰⁰ possible conformations!
- Cyrus Levinthal's paradox of protein folding,1968.

- Given a protein fold, which primary sequence(s) fold into it?
- Assume a total of 100 conformations for all 20 natural occurring amino acids side chains in a 100 amino acid protein ⇒ 10²⁰⁰ possible conformations!
- Earth is less than 10¹⁰ years old.

Sidechain Degrees of Freedom

All Likely Side Chain Conformations are Present in the Protein Databank

Sampling and Scoring for Side Chain Repacking and Design

Simulated Annealing Monte Carlo energy minimization

© Jens Meiler

Designed Model (blue) and X-ray Structure (red) of Top7

10 June 2018

Combinatorial Optimization Assumes a *Fixed Backbone*.

Enhancing Polyspecificity: Multistate Design

Multi-State design of three antibody-antigen complexes

REstrained CONvergence multi-specificity design (RECON)

Willis, J. R. *et al. PLoS Comput Biol*; **2013**; 9 (4):e1003045.

AA AB Encouraged convergence Forced convergence Forced convergence BB BB

Sevy, A. *et al. PLoS Comput Biol.* **2015**; 11(7):e1004300

RosettaAntibodyDesign

- Monte Carlo Algorithm.
- Whole CDR Sampling (GraftDesign)
 - Sample structures from database
 - Insert and optimize peptide bond without CDR disruption
 - New Graft Algorithm
 - 100% Closure
 - Roughly mimics Somatic Recombination
- Sequence Sampling (SeqDesign)
 - Sample amino acid types
 - According to CDR Cluster Profiles
 - Roughly mimics somatic hypermutation

Adolf-Bryfogle, J. et al. PLoS Comput Biol. 2018;14(4):e1006112

RosettaAntibodyDesign Top Decoy H3 Variability

2dqc

Antigen Design Example

Nature (2014), 507, 201-206.

Scaffold Grafting

Silva, D-A. et al. Methods Mol Biol. 2016;1414:285-304.

Other Structural Modeling Examples

Single Particle EM Indicative of Binding over Receptor Binding Site

10 June 2018

© Jens Meiler

Relative Orientation of H5.3 Fab is Ambiguous by EM Density Fitting

Mutational Analysis Indicates H5.3 Binding in Orientation B

© Jens Meiler

N. J. Thornburg, et al., *J Clin Invest*; **2013**; Vol. 123 (10): p. 4405-9.

PG9-like is a Broadly Neutralizing HIV Antibody

- 30 residue long HCDR3 loop accounts for neutralization functionality
- Variable chain has few mutations relative to other broadly neutralizing Abs
- Germline reversion of the variable chain retains neutralization

Do HIV-Naïve Humans have PG9like Antibodies?

Willis JR, et al. Proc Natl Acad Sci U S A. 2016;113(16):4446-51.

10 June 2018

"Thread" Long HCDR3 Sequences over PG9 co-crystal Structure

Rapid Screening of 25,000 HCDR3s using Rosetta-Inspired PSSMs

© Jens Meiler

Threading of Top 1000 HCDR3s Discriminates Unlikely Matches

Threading of Top 1000 HCDR3s Discriminates Unlikely Matches

Unfavorable

10 June 2018

Threading of Top 1000 HCDR3s Discriminates Unlikely Matches

Energetic Gap Between PG9 and Top Naïve Donor HCDR3 Sequences

Today's tutorial materials:

http://csb.vanderbilt.edu/~morettr/AntibodySummerSchool.html

Plug:

More in-depth protocols covered in RosettaWorkshop

- http://structbio.vanderbilt.edu/comp/workshops/rosetta/
 Workshop content at
- http://meilerlab.org/index.php/rosetta-tutorials

Email: rocco.moretti@vanderbilt.edu