Multi-State Design of Antibody-Antigen Interactions Confers Conformational Flexibility Hypothesis

Jordan Willis
Crowe Lab
Meiler Lab
RosettaCon 2011

Antibody Structure is Constructed through 3 Genes

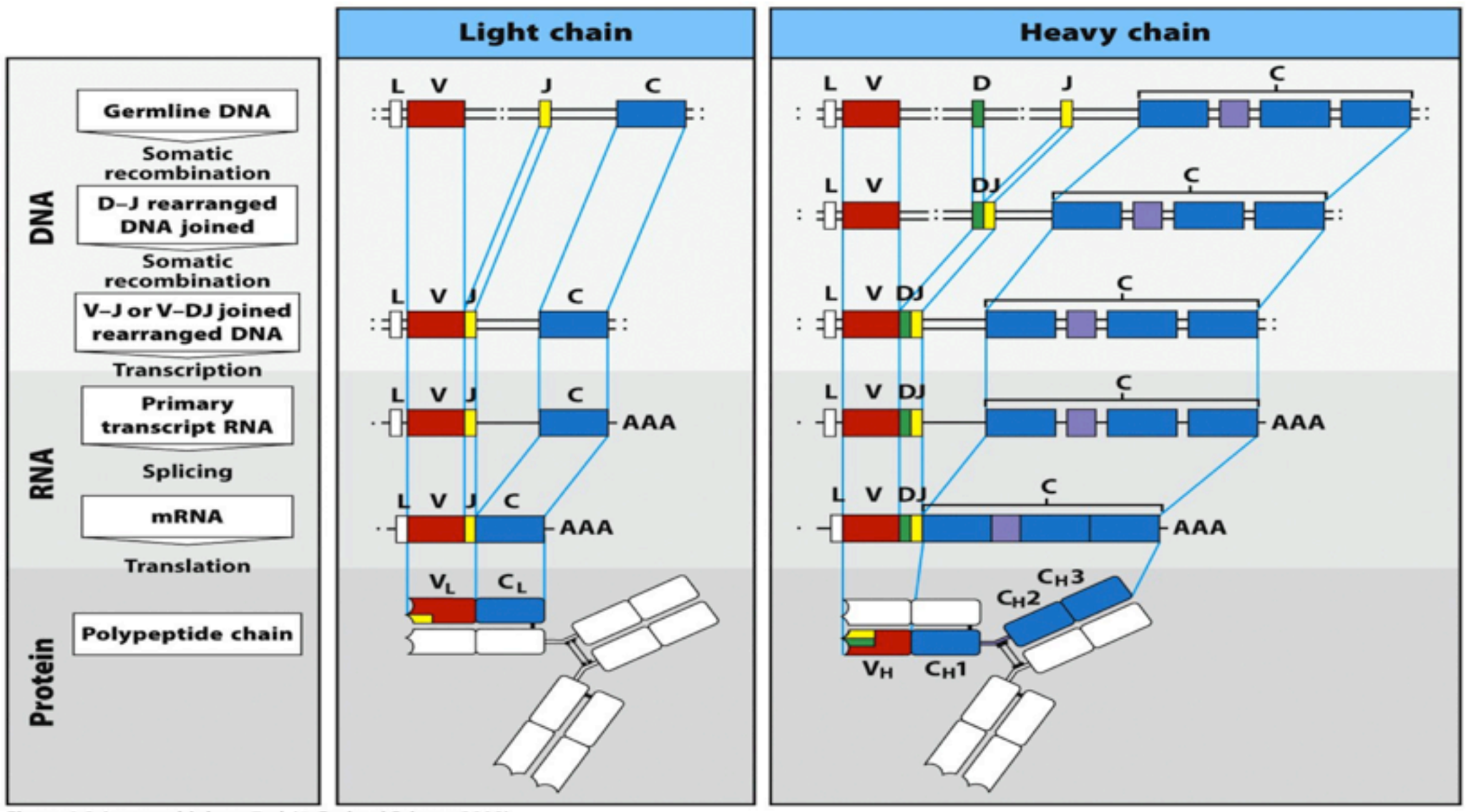
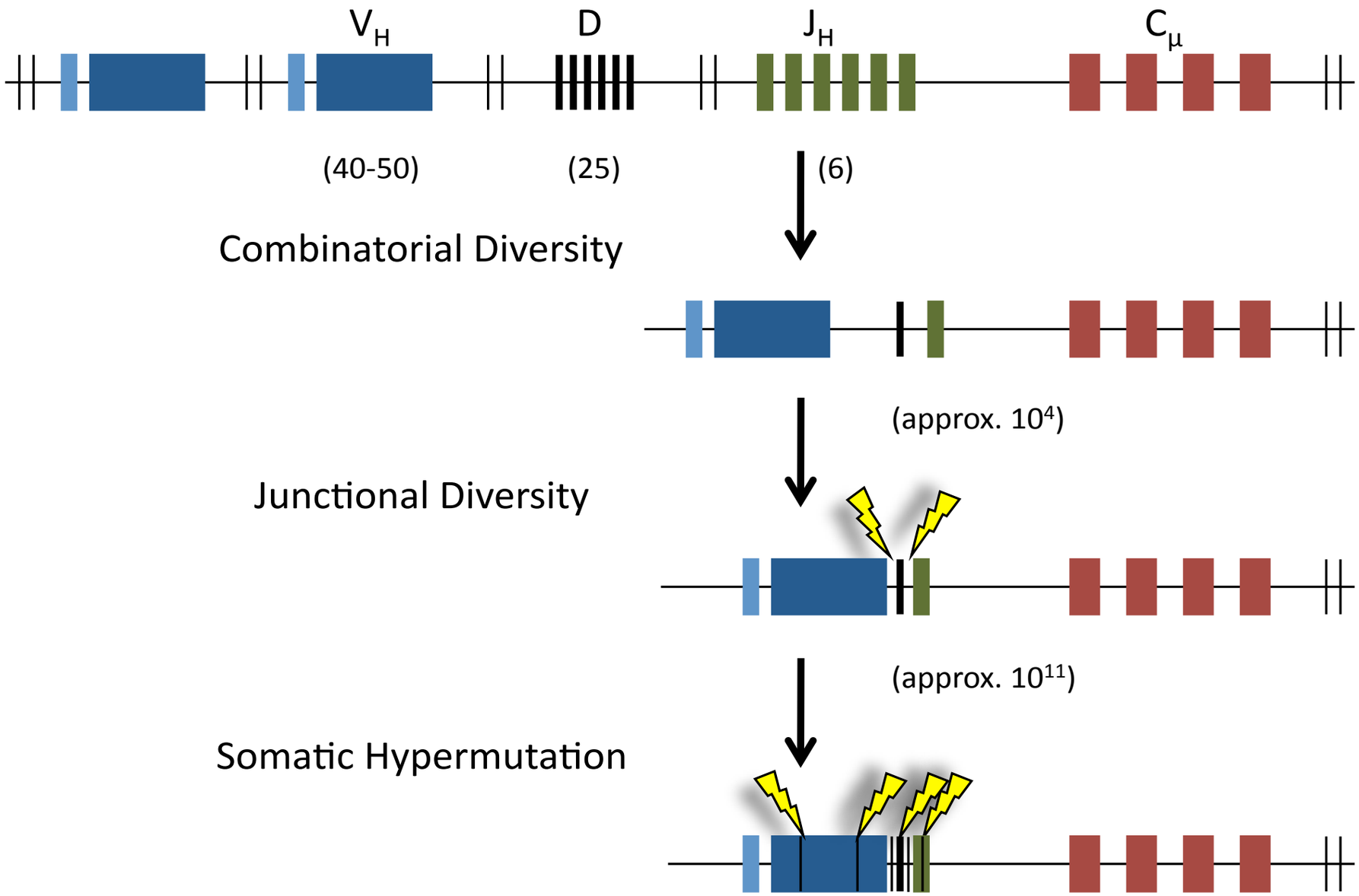


Figure 4-2 Immunobiology, 7ed. (© Garland Science 2008)

Heavy chain and light chain joins 3 and 2 gene segments respectively to form combinatorial diversity. Junctions form complimentary determining regions

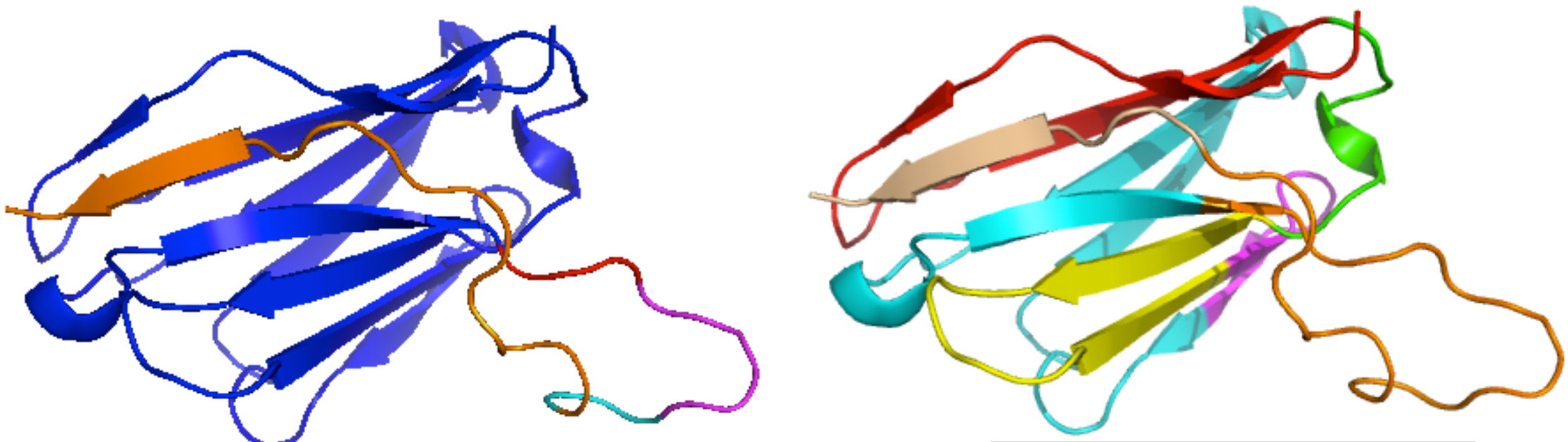
Antibody Diversity



Segment Count: *Immunobiology* (Janeway)
Junctional Diversity

V-Gene codes for a majority of antibody variable region

Heavy chain variable region



Blue - V Gene
Red - N-additon
Pink - D Gene
Cyan - n addition
Orange - J Gene

V gene codes for

Red - Framework I
Green - CDR1
Yellow - Framework II
Pink - CDR2
Cyan - Framework III

Orange - CDR3
Wheat - Framework IV

Motivation - HT sequencing reveals progenitor genes

- Crowe lab uses 454 pyro-sequencing to access antibody repertoire of healthy and viral infected patients.
- Antibody repertoire is the same for all healthy patients

Total Reads: 1,965,037

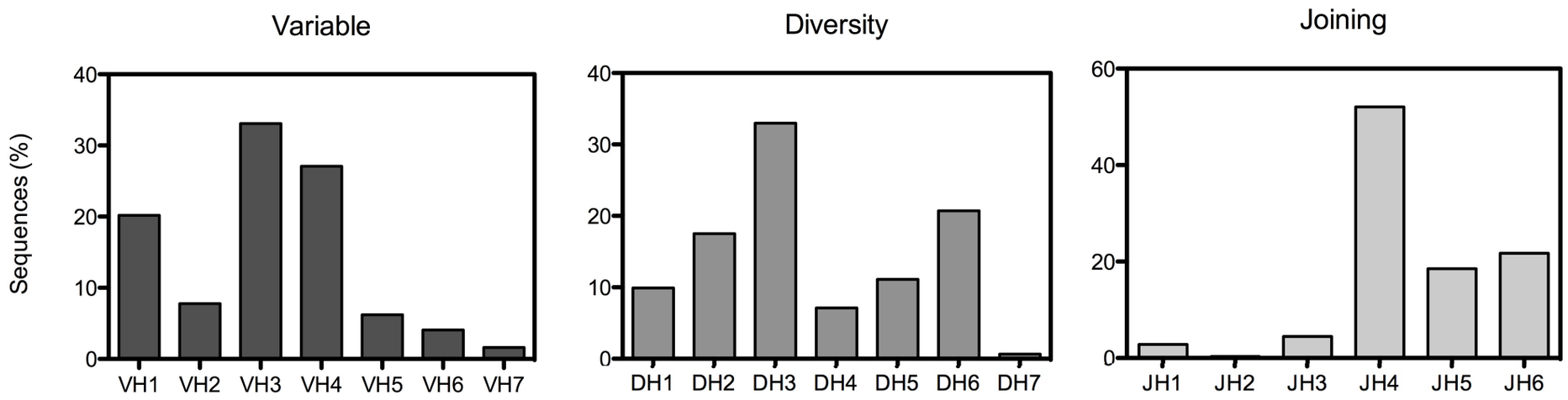
High Quality Reads: 1,868,183

95.1% of all reads are high quality antibody sequences

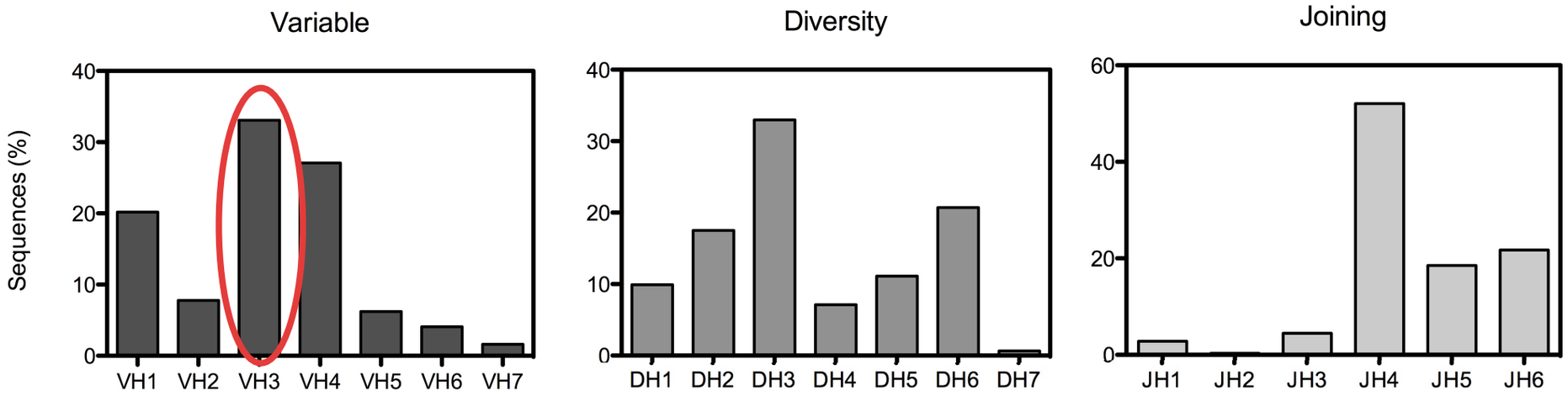
	Reads	High-Quality Reads
Peripheral Blood	149896	132248
Bone Marrow	171111	156177
Small Intestine	118044	108279
Lung	198660	181299
Lymph Node	165091	152175
Tonsil	197846	180319
TOTAL	**1000648**	**910497**

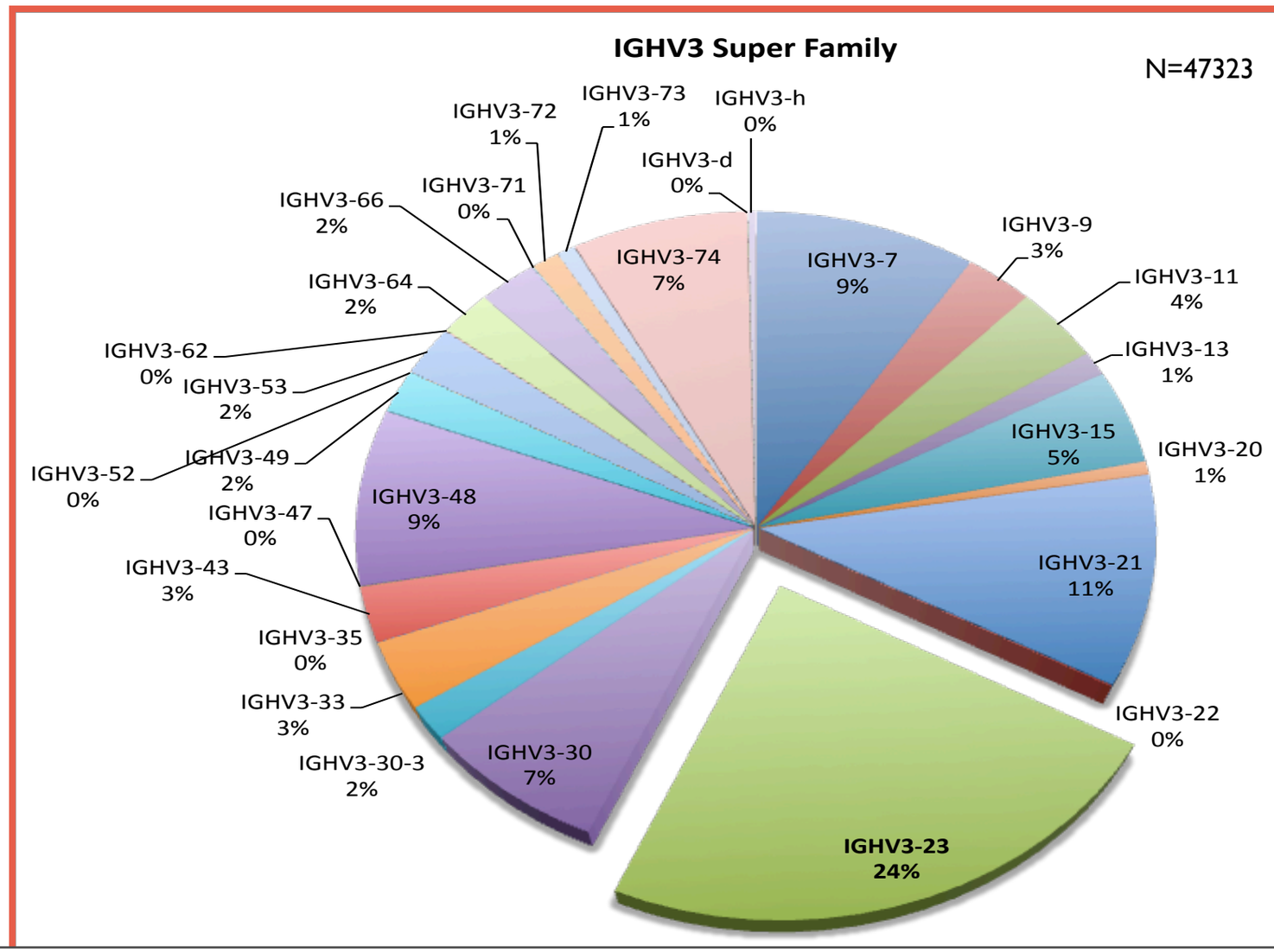
Briney, Willis, Crowe *Blood* 2011

Antibody Repertoire - VH3-23 dominates

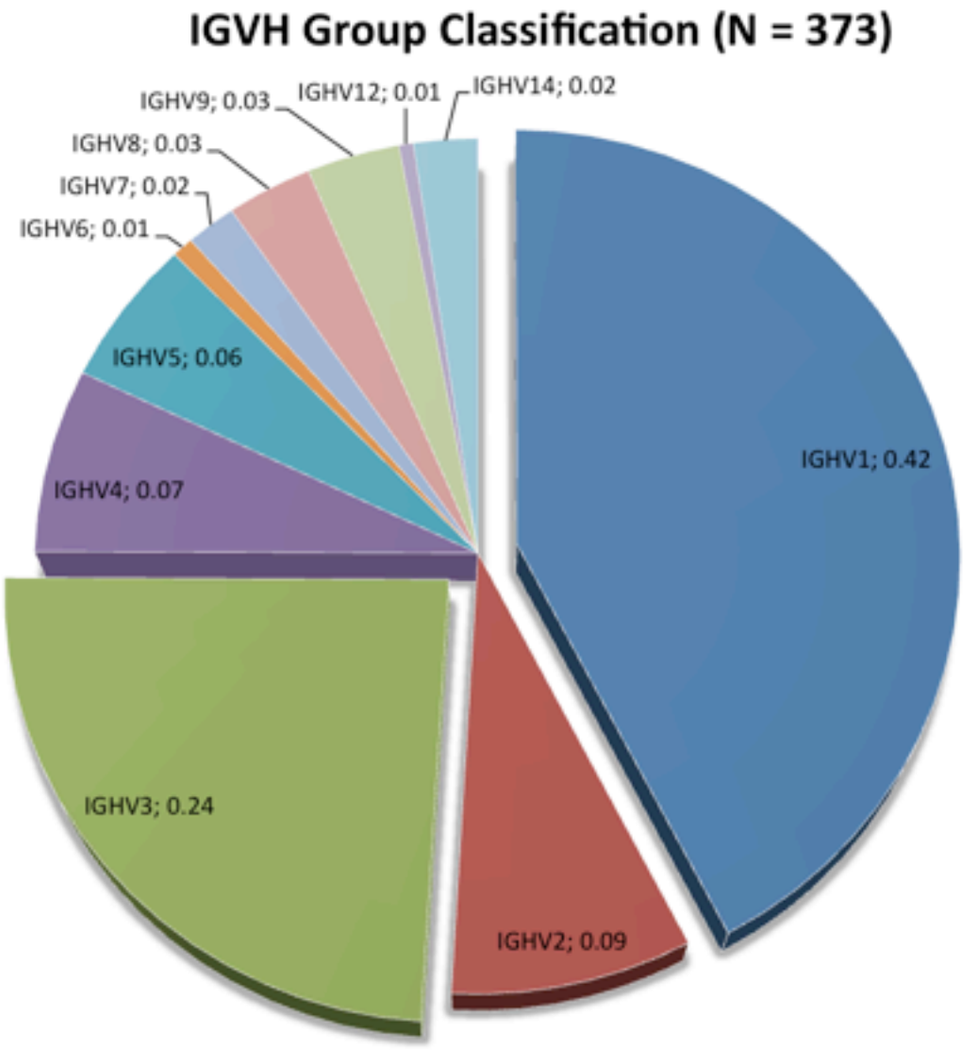


Antibody Repertoire - VH3-23 dominates

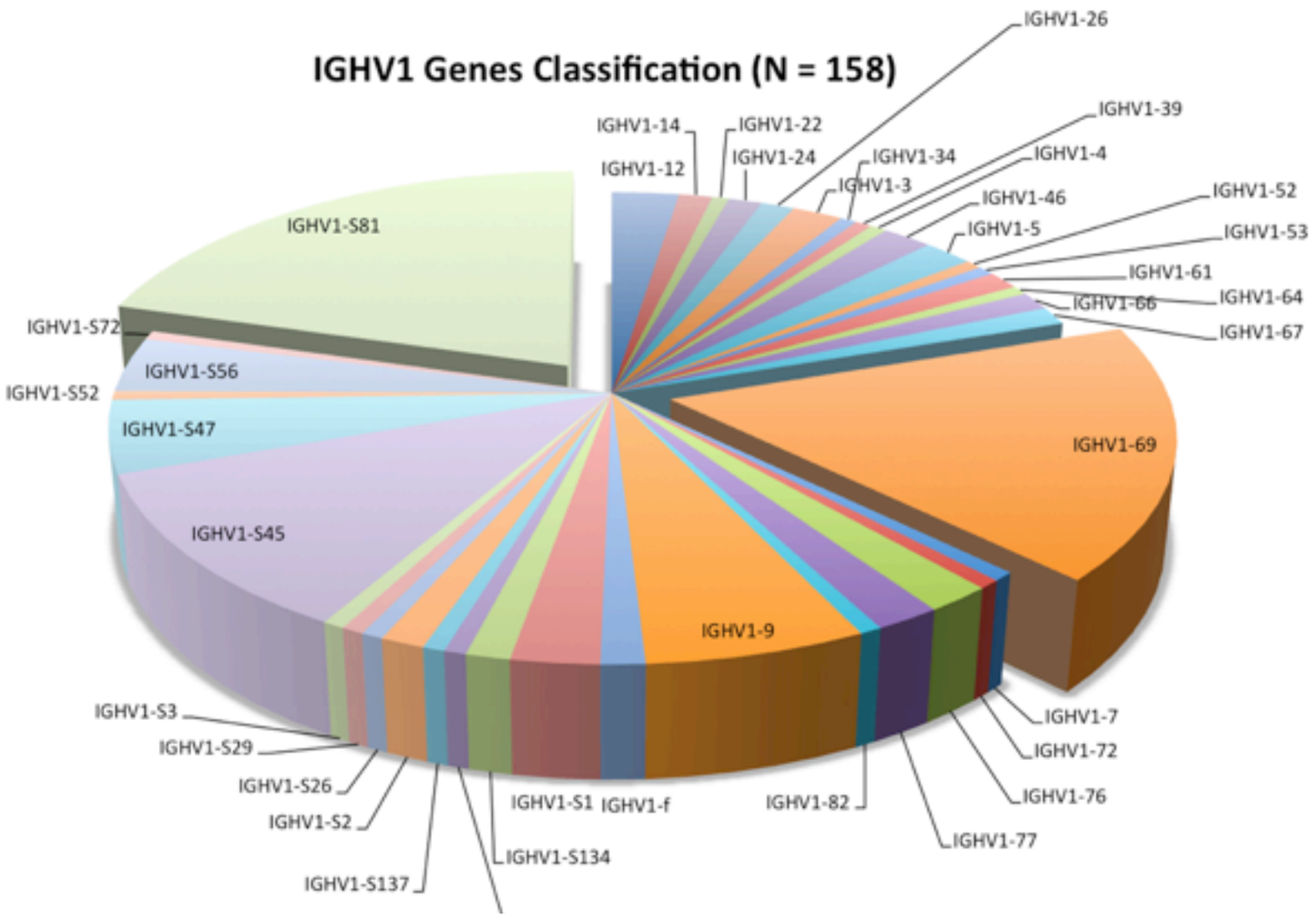




PDB Antibody Repertoire - Recapitulates Sequencing Repertoire

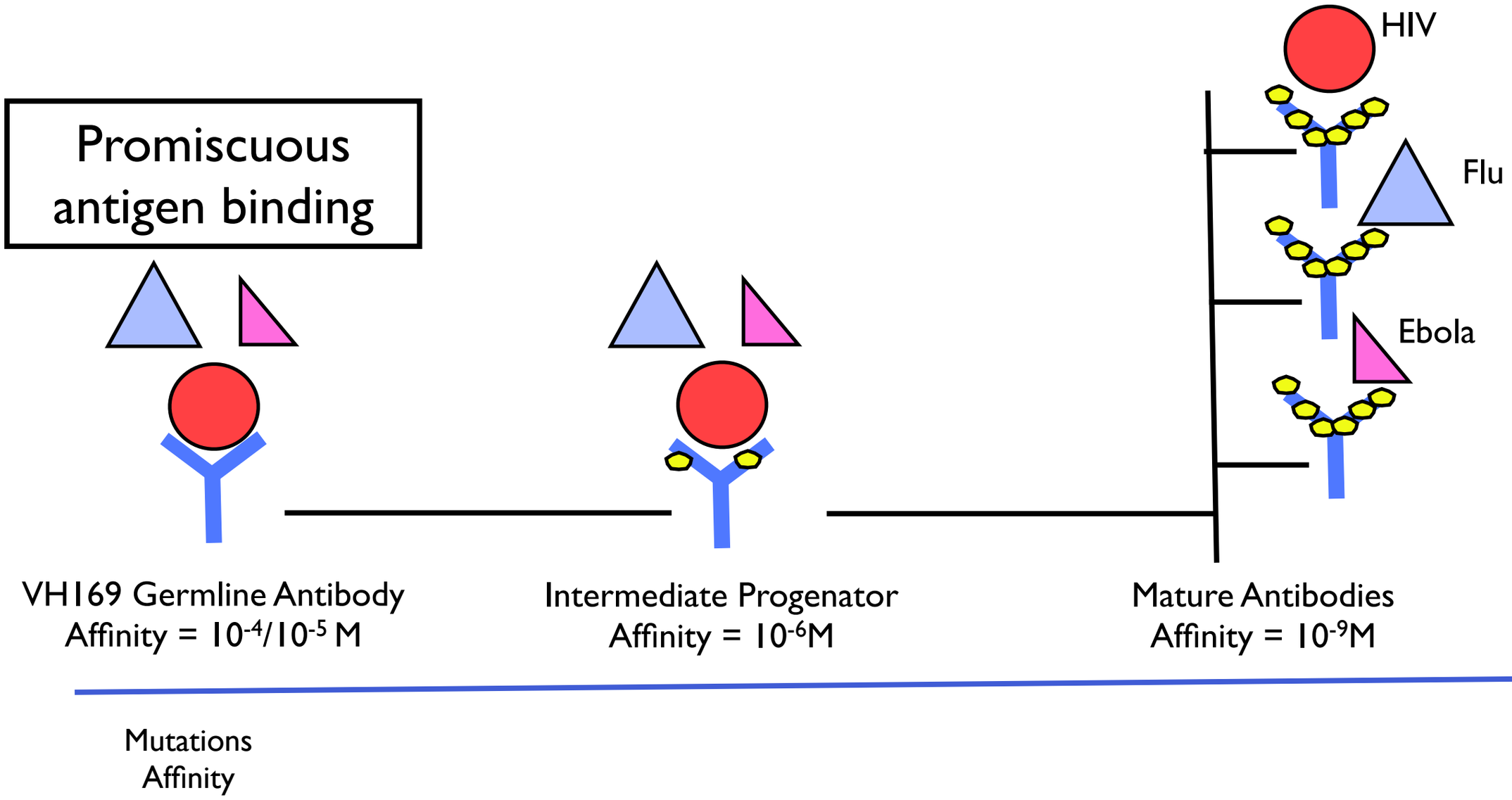


Search by antibody-antigen protein complexes



Hypothesis

There exists conformational flexibility on commonly used germline genes that accommodates a variety of antigenic structures. Using multi-state design we can test if germline sequences are optimal to bind a set of native complexes.



Multi-State Design

Multi-constraint computational design suggests that native sequences of germline antibody H3 loops are nearly optimal for conformational flexibility

Mariana Babor[1,2] and Tanja Kortemme[1,2]*

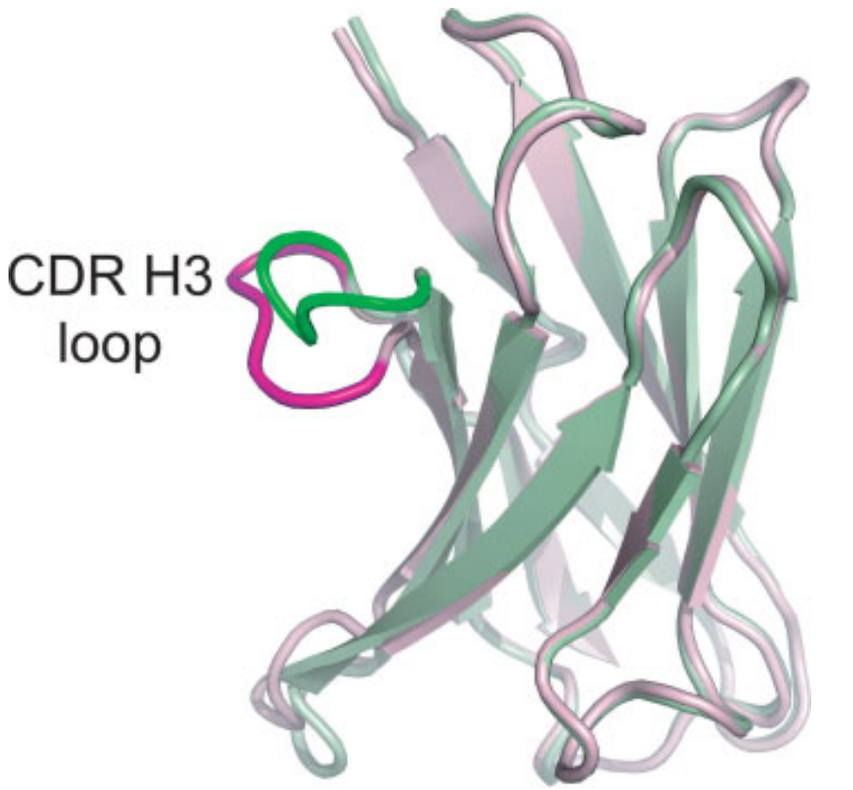
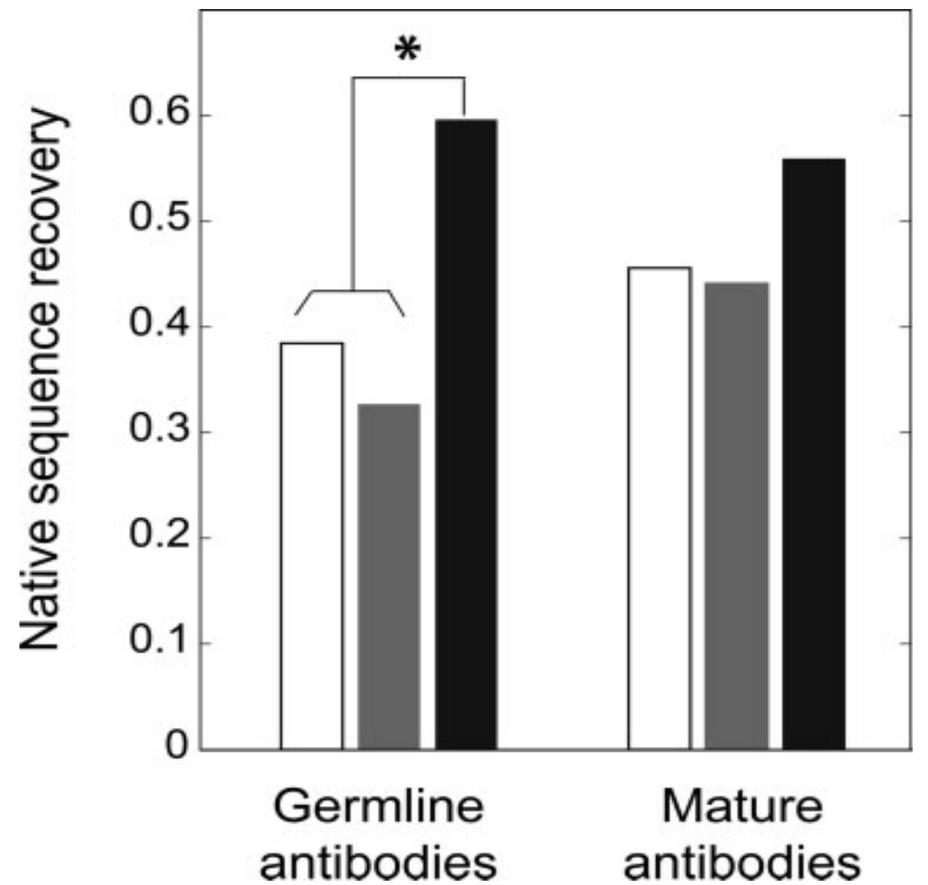


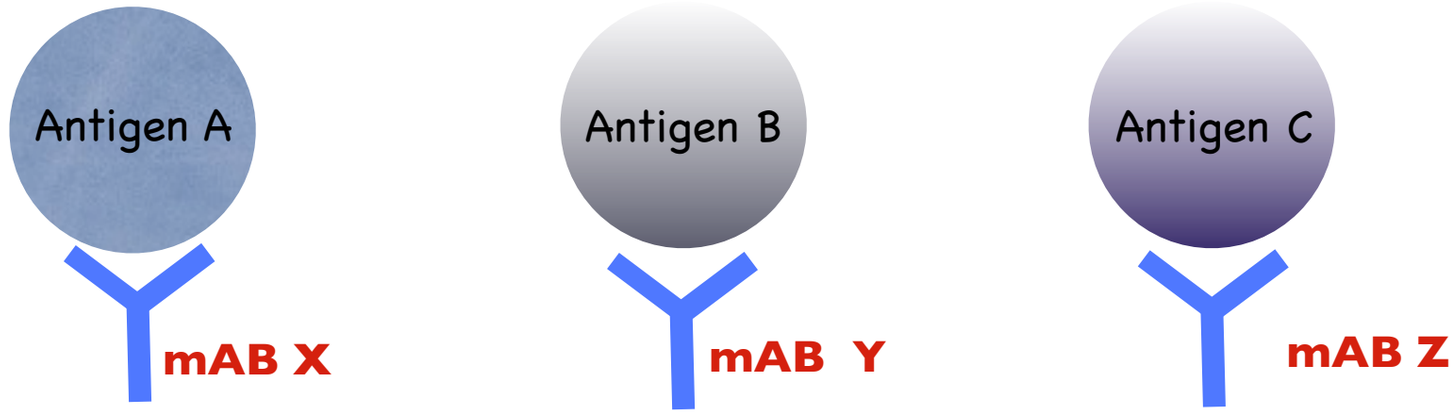
Figure 2

Superimposition of the V_H domain of the germline 7g12 antibody in its bound (pdb: 1n7m) and free (pdb: 1ngz) forms (green and magenta, respectively).



Multi-State Design may reveal promiscuous sequences

ex. VH1-69 mAb's



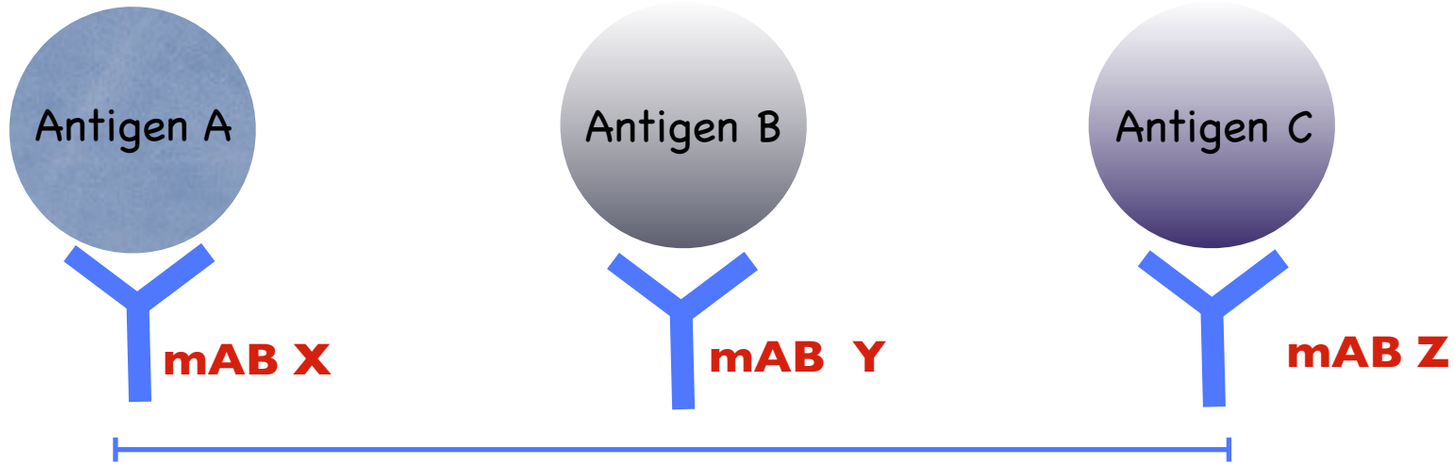
mpi_msd.linuxrelease - design in mutations that give a lower energy structure for each state

mAB X, Y, and Z all use the same germline, but bind separate, and structurally unique antigens. The germline sequence must be flexible to accommodate all these positions.

Conformational Flexibility Hypothesis

Multi-State Design may reveal promiscuous sequences

ex. VH1-69 mAb's



mpi_msd.linuxrelease - design in mutations that give a lower energy structure for each state

mAB X, Y, and Z all use the same germline, but bind separate, and structurally unique antigens. The germline sequence must be flexible to accommodate all these positions.

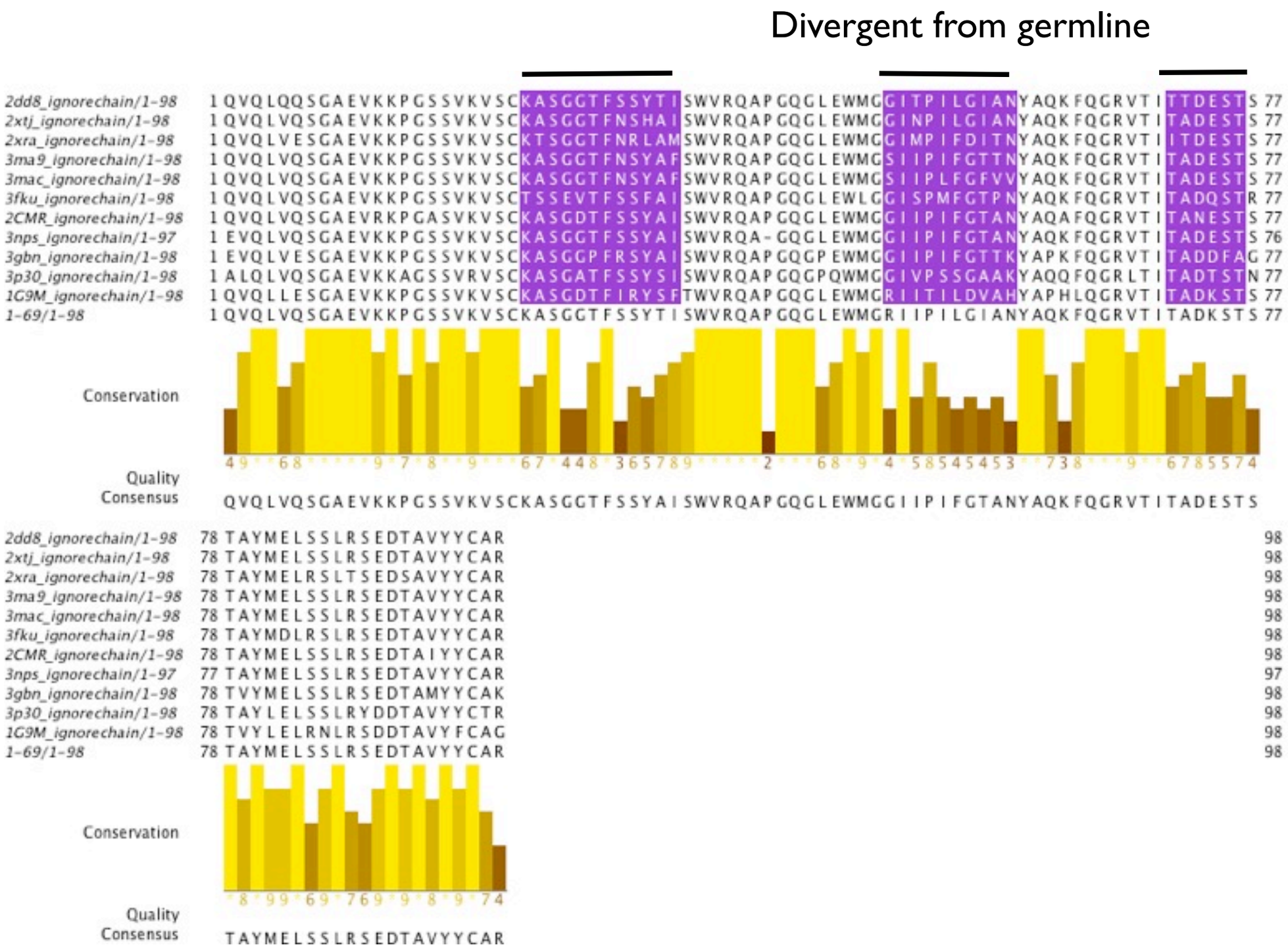
Conformational Flexibility Hypothesis

VH1-69 Mature Antibody Complexes

Entry (PDB ID)	Antibody Name	Type	Ligand	Resolution
1g9m	17b	FAB Kappa	Envelope Glycoprotein gp120 (HXBC2)	2.20
2b4c	X5	FAB Kappa	Envelope Glycoprotein gp120 (JRFL)	3.30
2cmr	D5	FAB Kappa	Gp41 Fusion Intermediate	2.0
2dd8	m396	FAB Lambda	SARS Spike	2.30
2xra	HK20	FAB Kappa	Transmembrane protein (synthetic)	2.30
2xtj	1D05	FAB Kappa	Proprotein convertase substilin	2.70
3fku	F10	ScFV Kappa	Hemmaglutanin	3.20
3gbn	CR6261	FAB Lamda	Hemmaglutanin Peptide	2.20
3ma9	8066	FAB Lamda	Transmembrane Glycoprotein	2.05
3mac	8062	FAB Lambda	Transmembrane Glycoprotein	2.50
3nps	S4	FAB Kappa	Suppressor of tumorgenicity protein	1.50
3p30	1281	FAB Lamda	Gp41 Fusion Intermediate	3.30

12 candidate test complexes using VH1-69

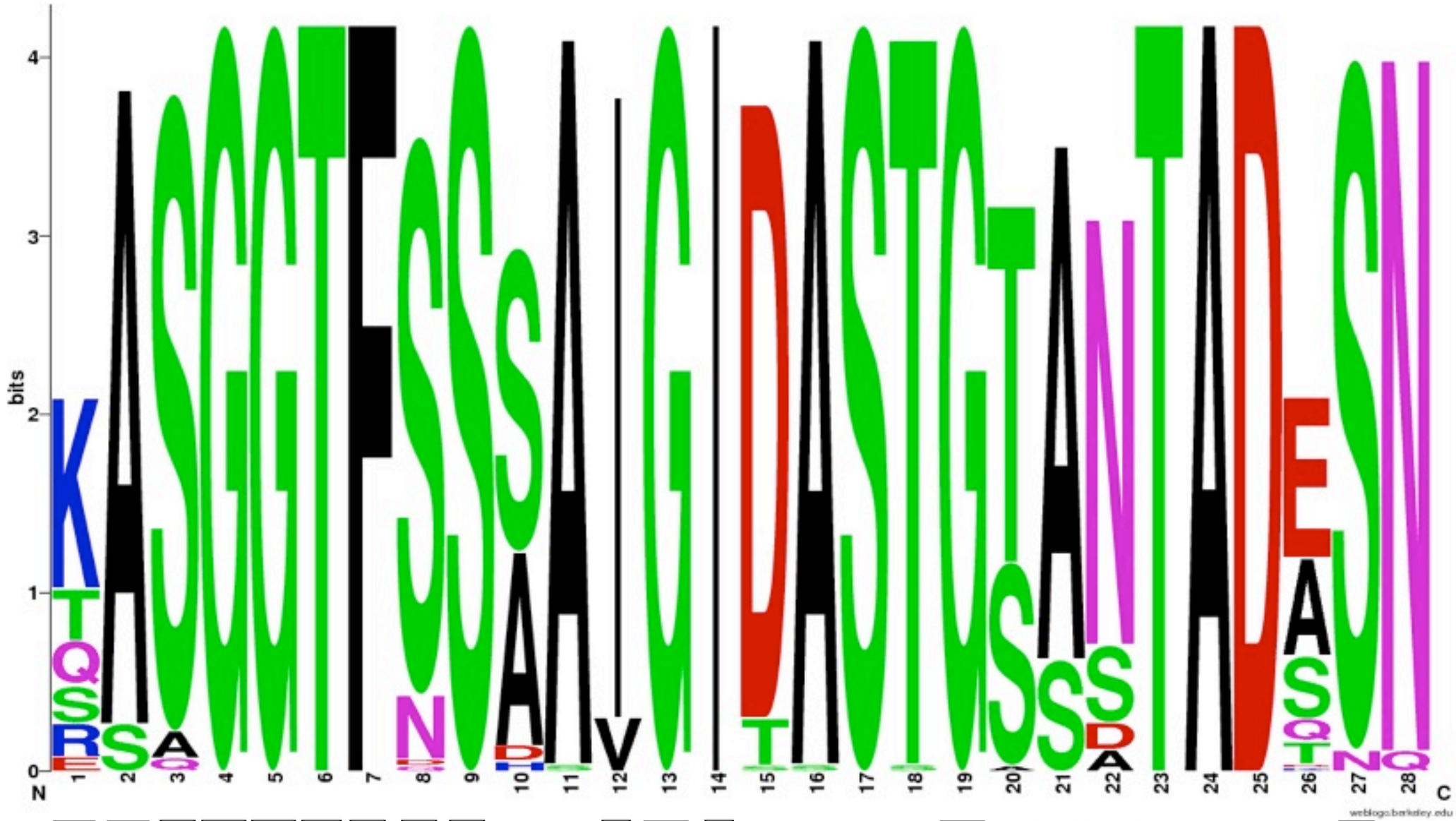
VH1-69 Mature Antibody Complexes



VH1-69 Multi-State Design

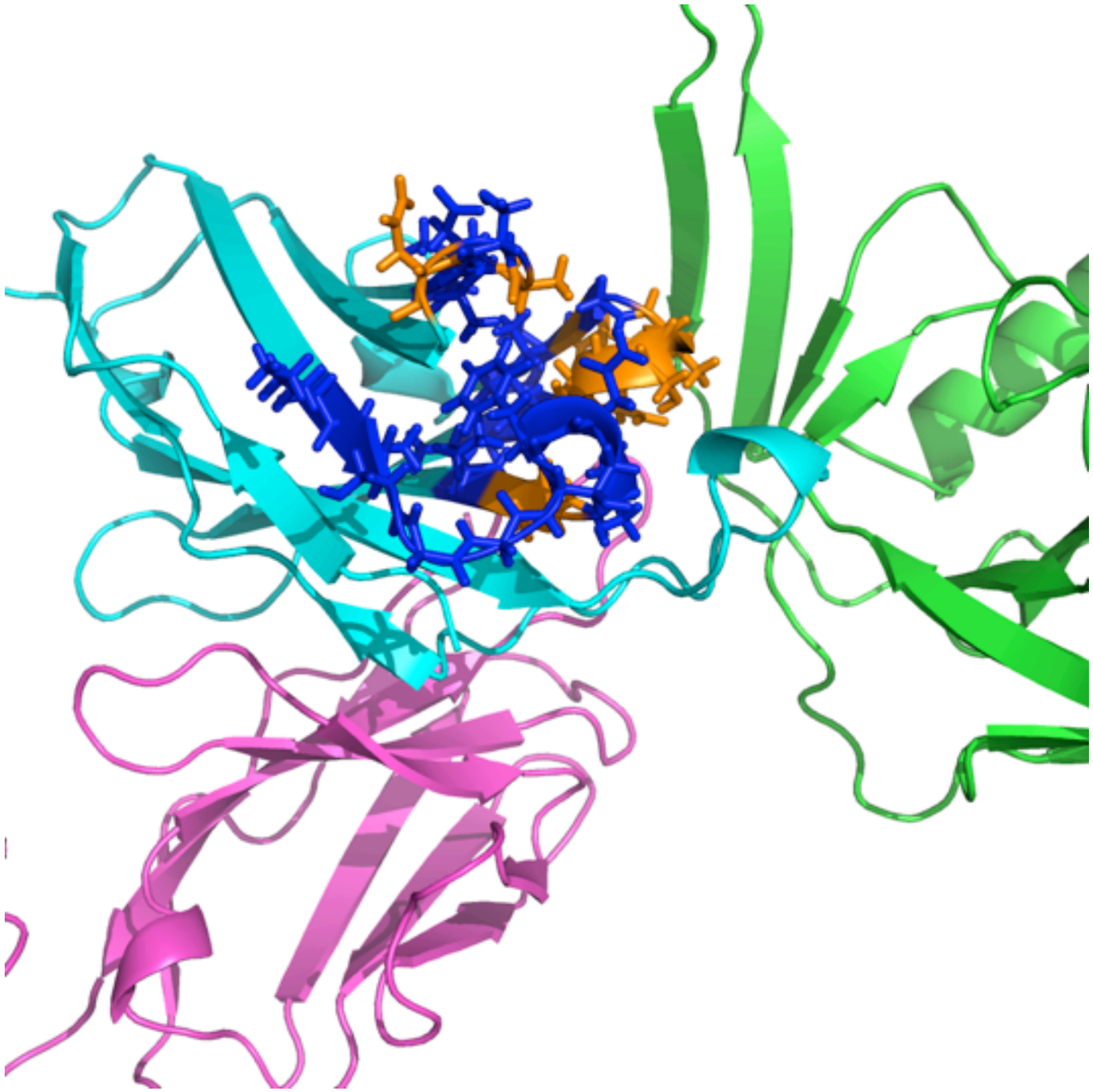
VH1-69 MSD (11 States)

19/28 recovered to VH1-69



VH1-69 K A S G G T F S S Y T I G I I P I L G I A N T A D K S T

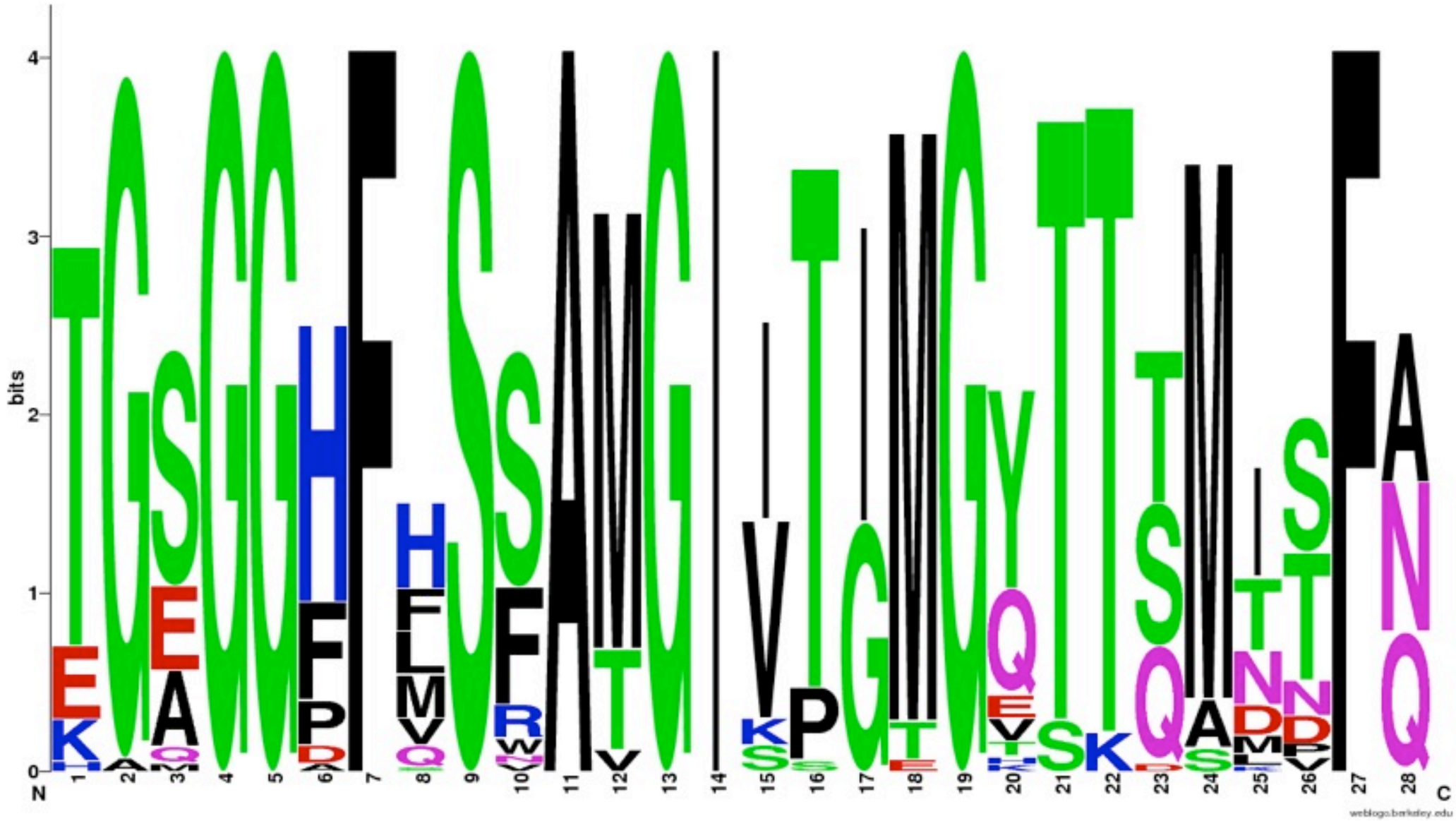
VH1-69 Multi-State Design



Antibody 17b designs towards germline in MSD with correctly designed amino acids shown in dark blue. Incorrect designed are shown in orange

VH1-69 Single-State Design (3GBN)

16/28 recovered to VH1-69
24/28 Native

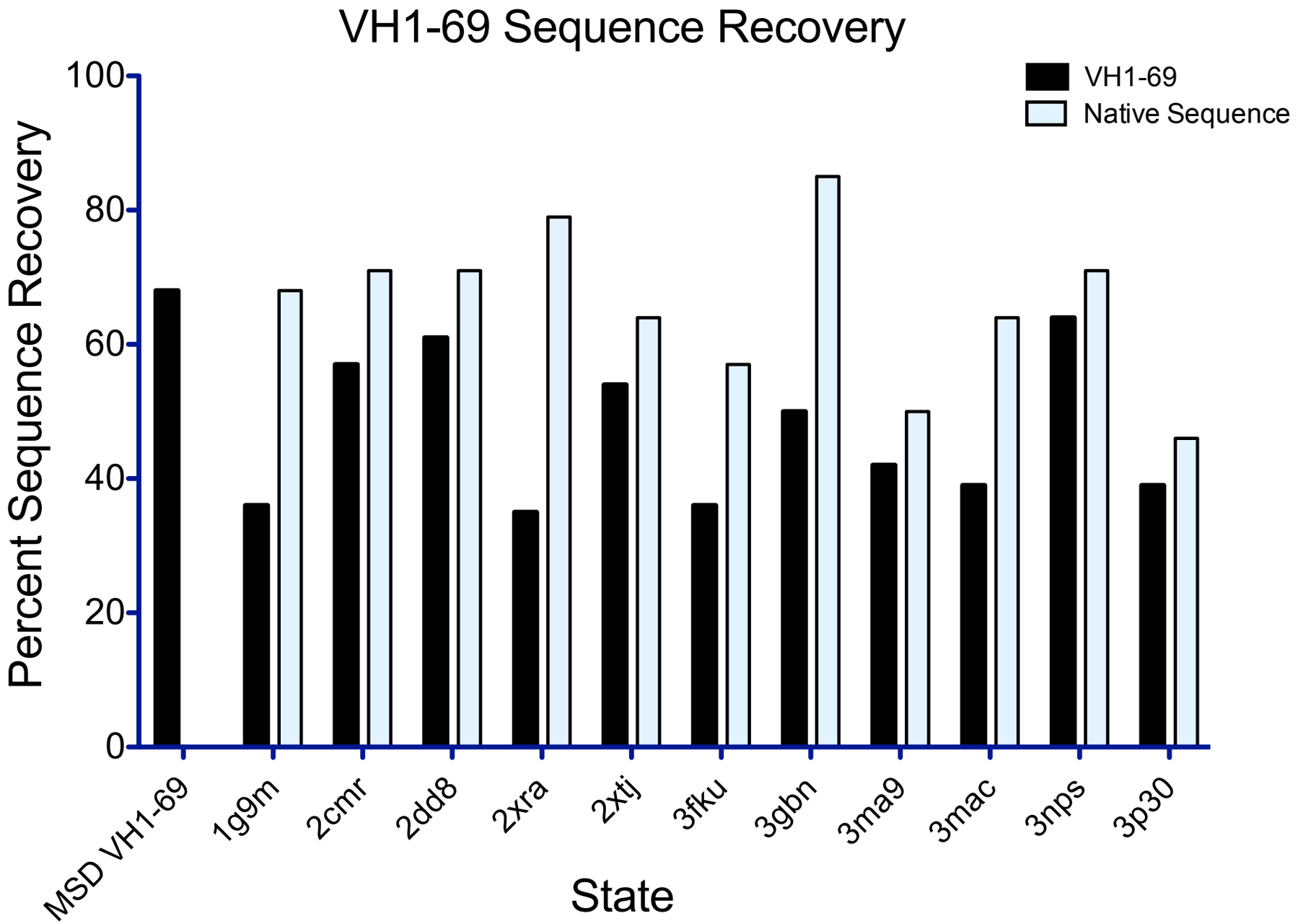


VH1-69	K	A	S	G	G	T	F	S	S	Y	T	I	G	I	I	P	I	L	G	I	A	N	T	A	D	K	S	T
3gbn	K	A	S	G	G	P	F	R	S	Y	A	I	G	I	I	P	I	F	G	T	T	K	T	A	D	D	F	A

VH1-69 MSD/SSD Design

Design	Percentage recovered to native	Percentage to VH1-69
MSD of 11 VH1-69 States	-	68
1g9m	68	36
2cmr	71	57
2dd8	71	61
2xra	***79***	***35***
2xtj	64	54
3fku	57	36
3gbn	85	50
3ma9	50	42
3mac	64	39
3nps	71	64
3p30	46	39

VH1-69 MSD/SSD Design



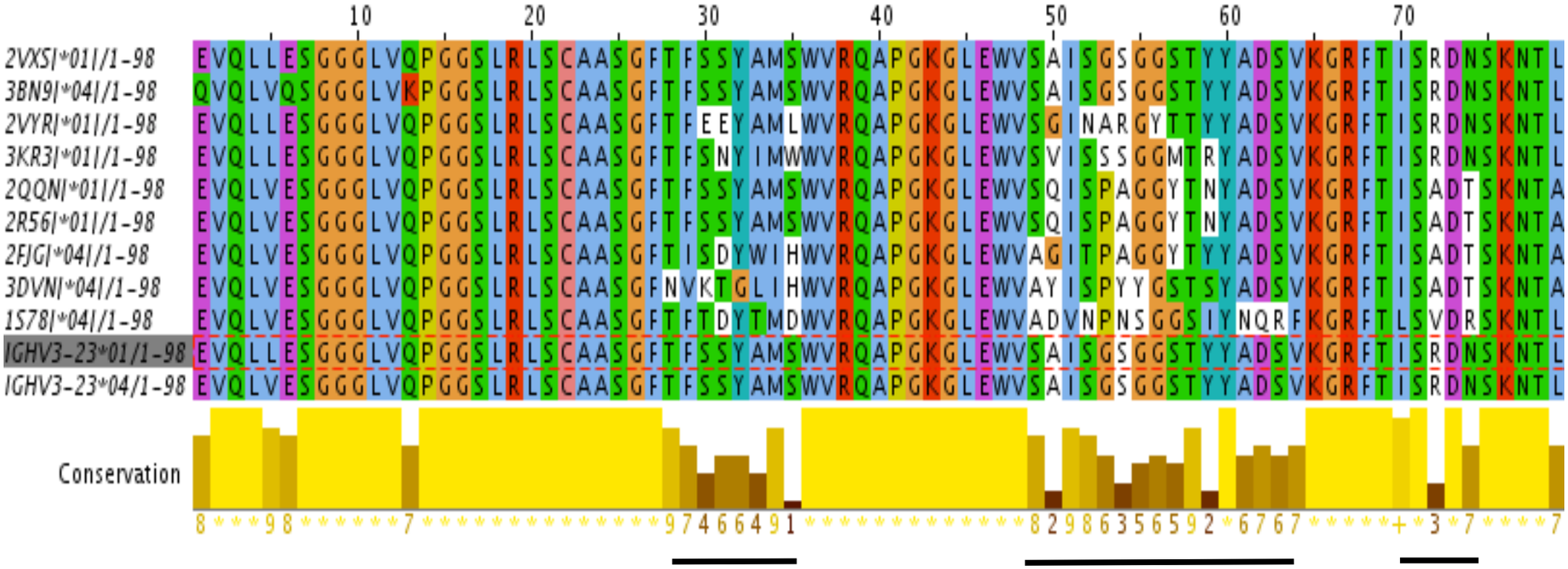
Native Sequence Recovery Average = 66%
VH1-69 Recovery Average = 46%

VH3-23 Mature Antibody Complexes

Entry (PDB ID)	Antibody Name	Type	Ligand	Resolution
1s78	Pertuzumab	FAB Kappa	ErbB-2	3.25
2fjg	G6	FAB Kappa	Vascular endothelial growth factor 1	2.80
2qqn	Semaphorin Blocking	FAB Lamda	Neurophilin-1	2.20
2r56	IgE Fab Fragment	FAB Kappa	Beta-lactoglublin allergen	2.80
2vxs	Unnamed	FAB Lamda	Interleukin-17A	2.63
2vyr	Unnamed Single VH chain	Single Chain	MDM4 Protein	2.00
3bn9	E2	FAB Kappa	Supressor of tumorigenicity protein 14	2.17
3dvn	Apu2.16	FAB Kappa	Ubiquitin	2.70
3kr3	DX-2647	FAB Kappa	Insulin-like growth factor II	2.20

9 candidate test complexes using VH3-23

VH3-23 Mature Antibody Complexes

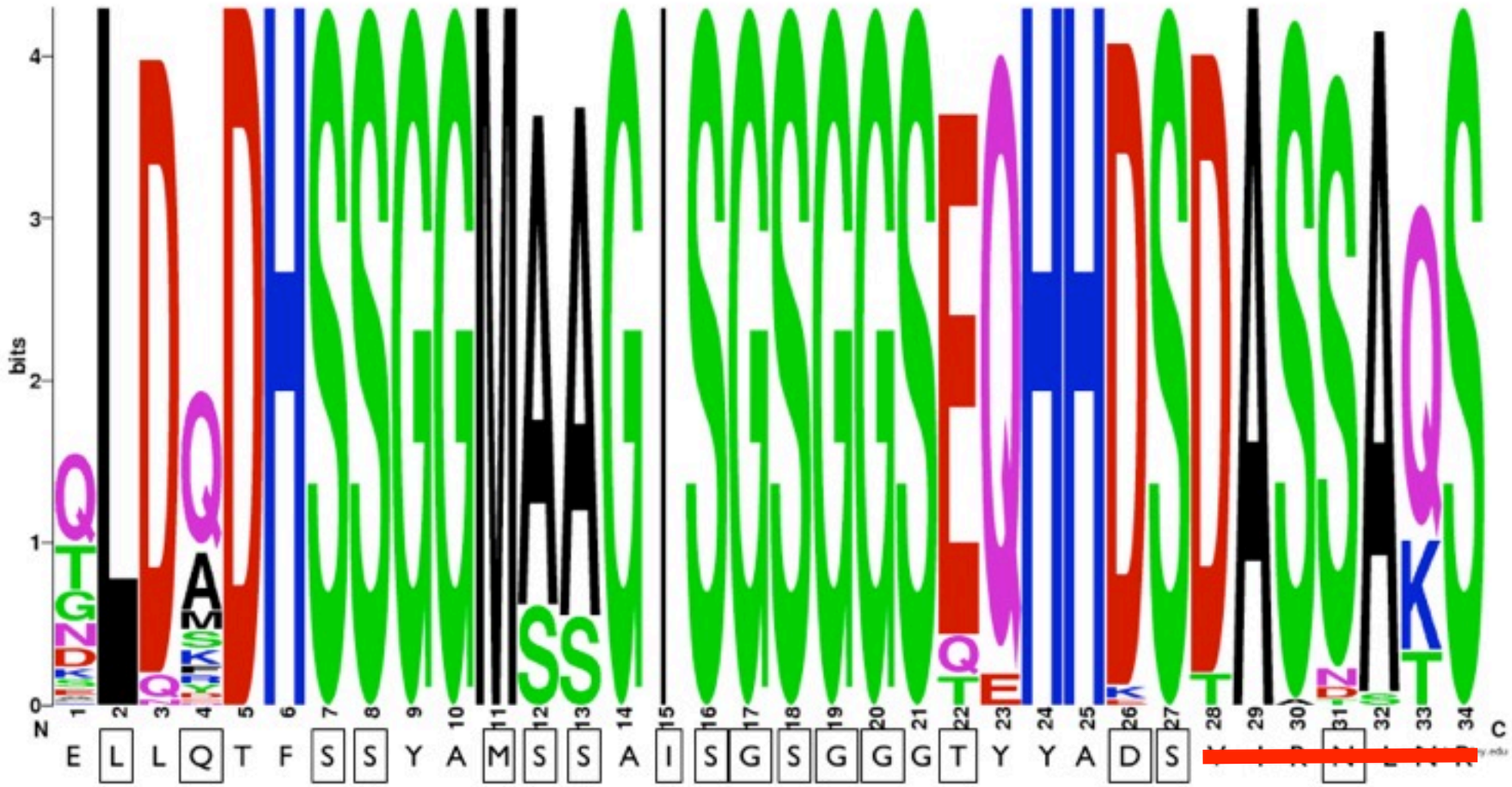


Divergent from germline

9 candidate test complexes using VH3-23

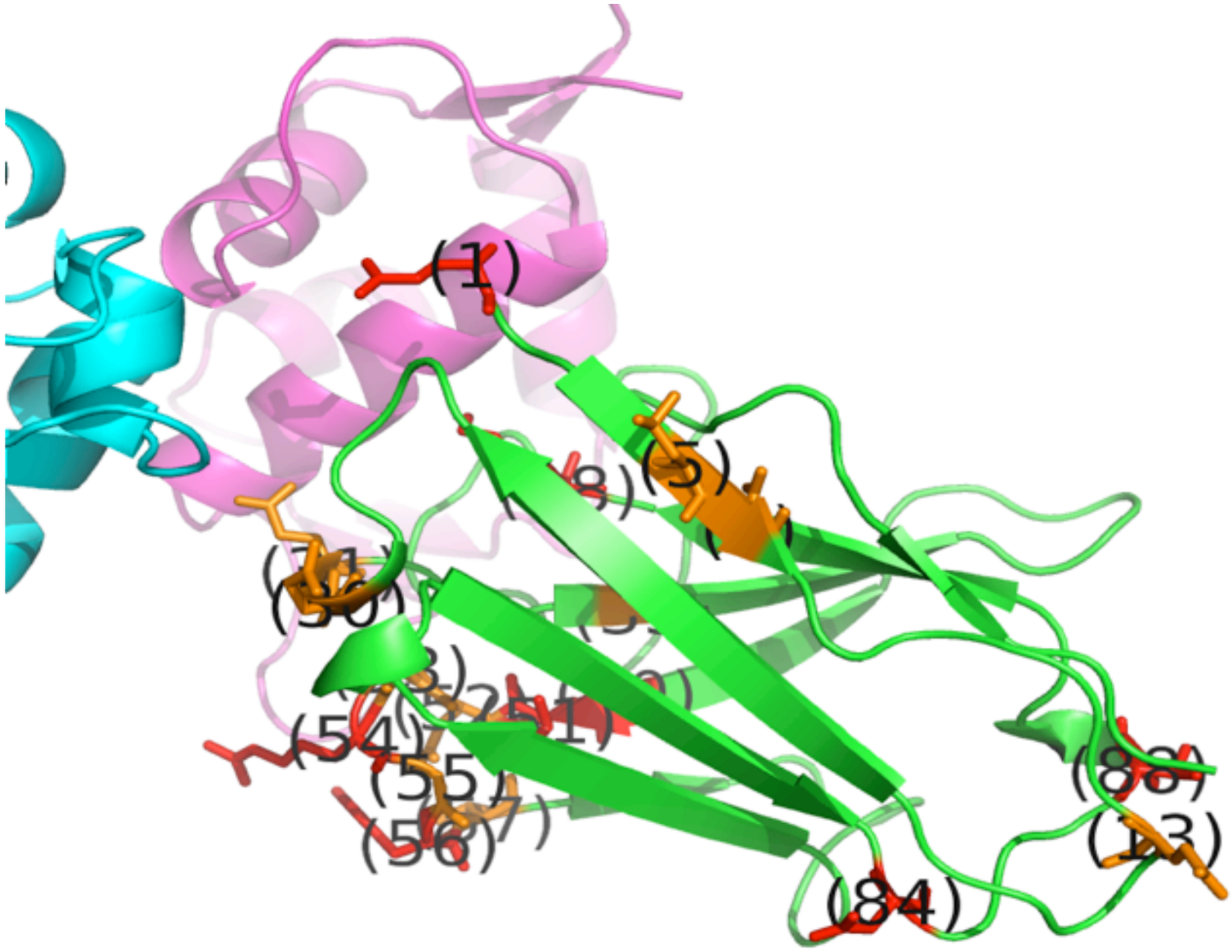
VH3-23 Multi-State Design

9 States - VH3-23 fixed backbone



16/27 ~~17/33~~ Amino Acids Recovered

VH3-23 Multi-State Design

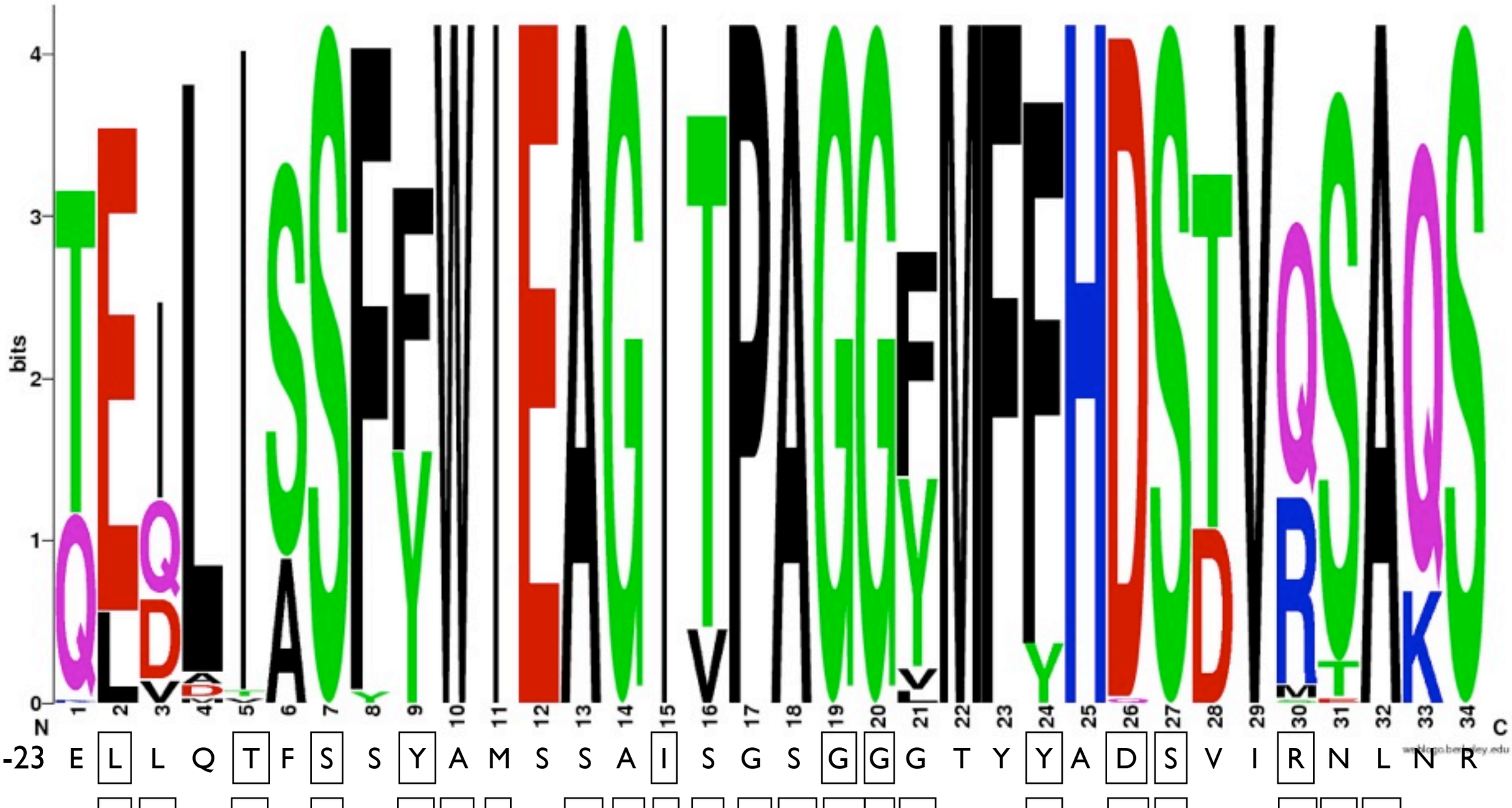


Orange - Correct, Red - Incorrect

VH3-23 Single-State Design

2FJG SSD

11/34 recovered to VH3-23
22/34 recovered to Native
11 more recovered to single state

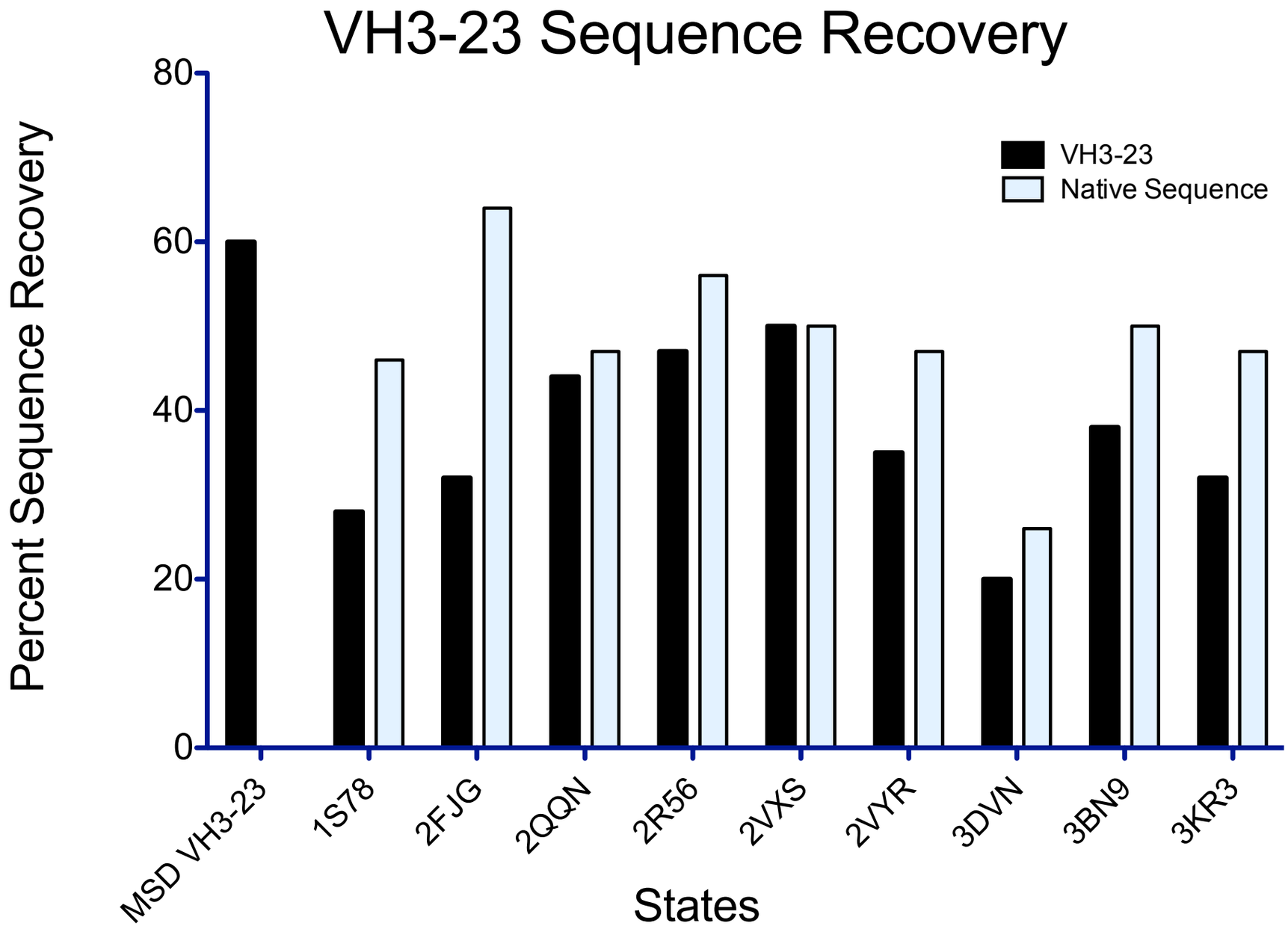


VH3-23	E	L	L	Q	T	F	S	S	Y	A	M	S	S	A	I	S	G	S	G	G	G	T	Y	Y	A	D	S	V	I	R	N	L	N	R
2FJG	E	L	V	Q	T	I	S	D	Y	W	I	H	A	G	I	T	P	A	G	G	Y	T	Y	Y	A	D	S	V	I	A	T	A	N	R

VH3-23 MSD/SSD Design

Design	Percentage recovered to native	Percentage recovered to VH3-23
MSD of VH3-23 States	-	60
1S78	46	28
2FJG	64	32
2QQN	47	44
2R56	56	47
2VXS	50	50
2VYR	47	35
3DVN	26	20
3BN9	50	38
3KR3	47	32

VH3-23 MSD/SSD Design



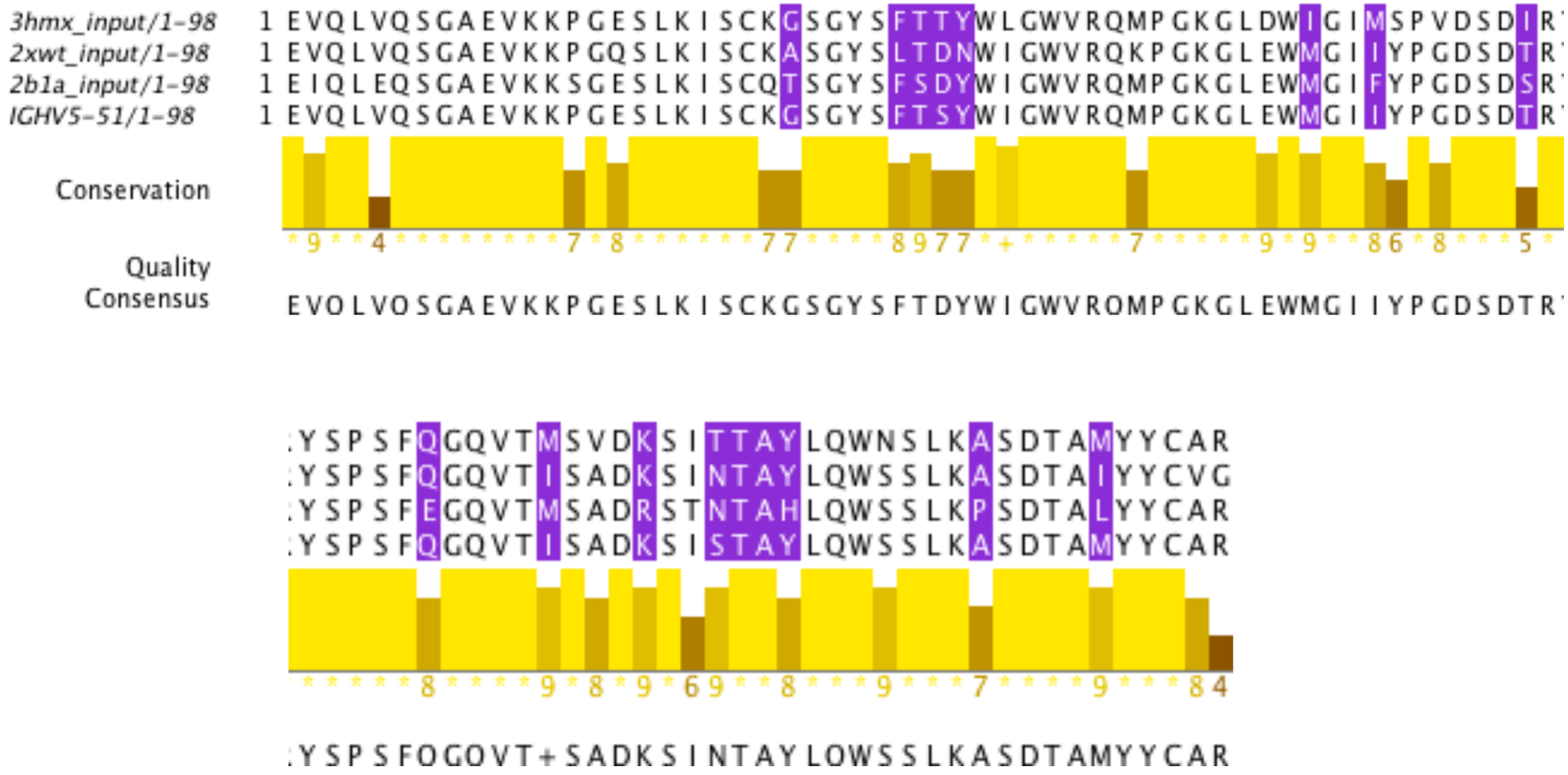
Native Sequence Recovery Average = 48%
VH1-69 Recovery Average = 36%

VH5-51 Mature Antibody Complexes

Entry (PDB ID)	Antibody Name	Antibody Description	Ligand	Resolution
2b1a	2219	FAB Lamda	UG1033 Peptide	2.35
2xwt	K1-70	FAB Lamda	TSH-R	1.90
3hmx	ustekinumab Fab	FAB Lamda	IL-12	3.00
2dd8	m396	FAB Lambda	SARS Spike	2.30

4 candidate test complexes using VH5-51

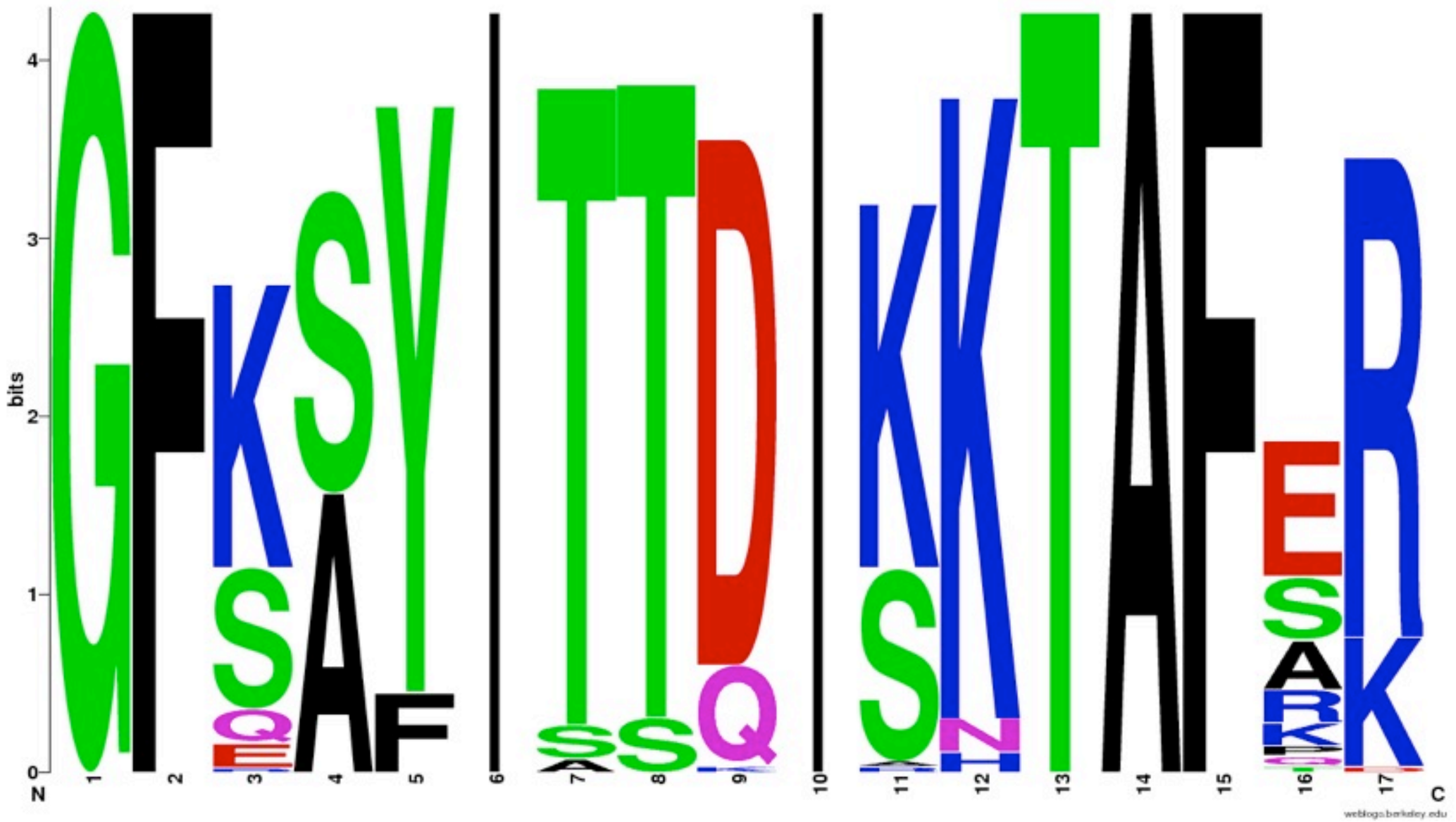
VH5-51 Mature Antibody Complexes



VH5-51 Multi-State Design

VH5-51 MSD

11/17 recovered to VH5-51

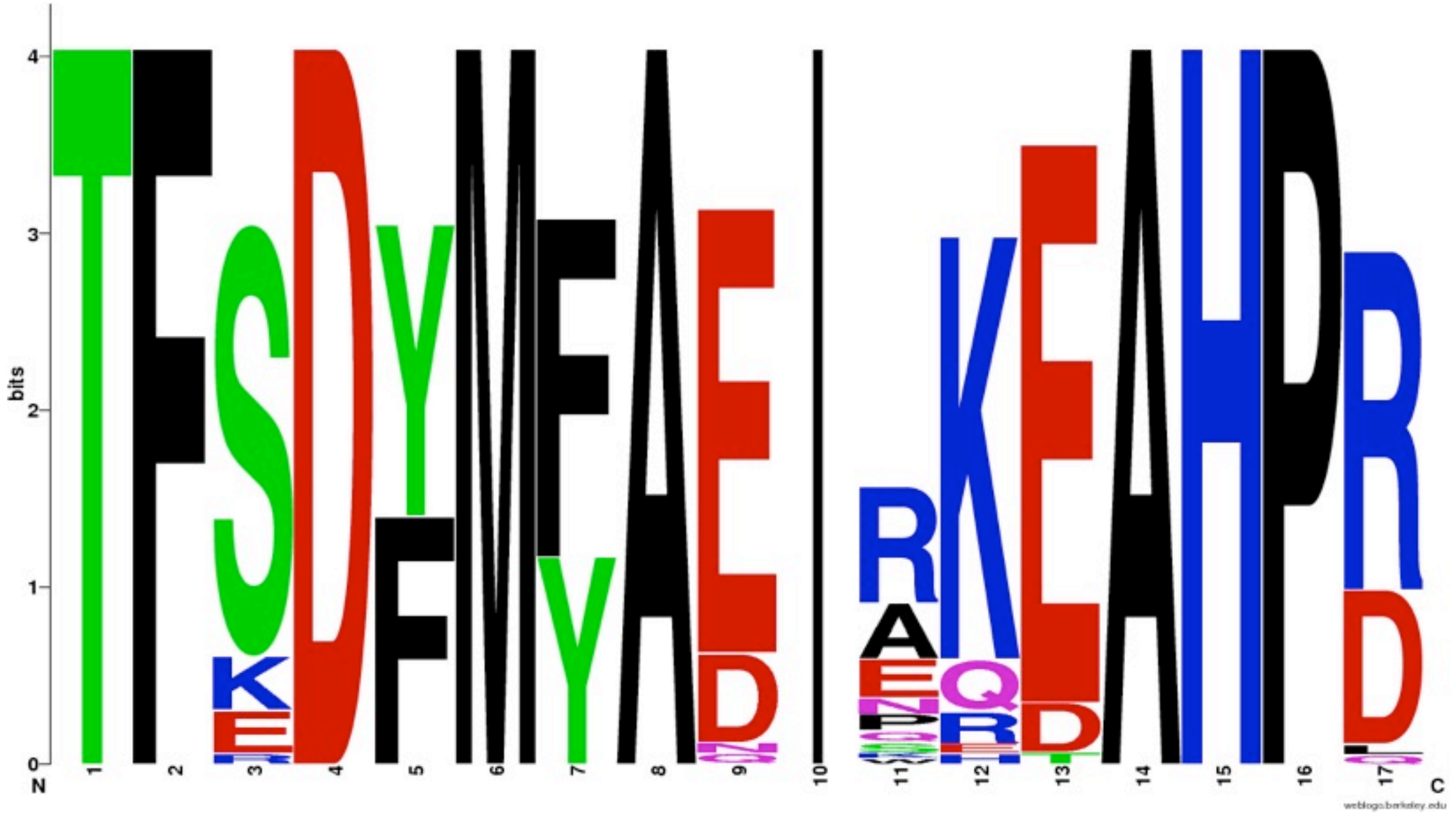


VH5-51	G	F	T	S	Y	M	I	T	Q	I	K	S	T	A	Y	A	M

VH5-51 Single-State Design

2B1A MSD

6/17 recovered to VH5-51
14/17 Native

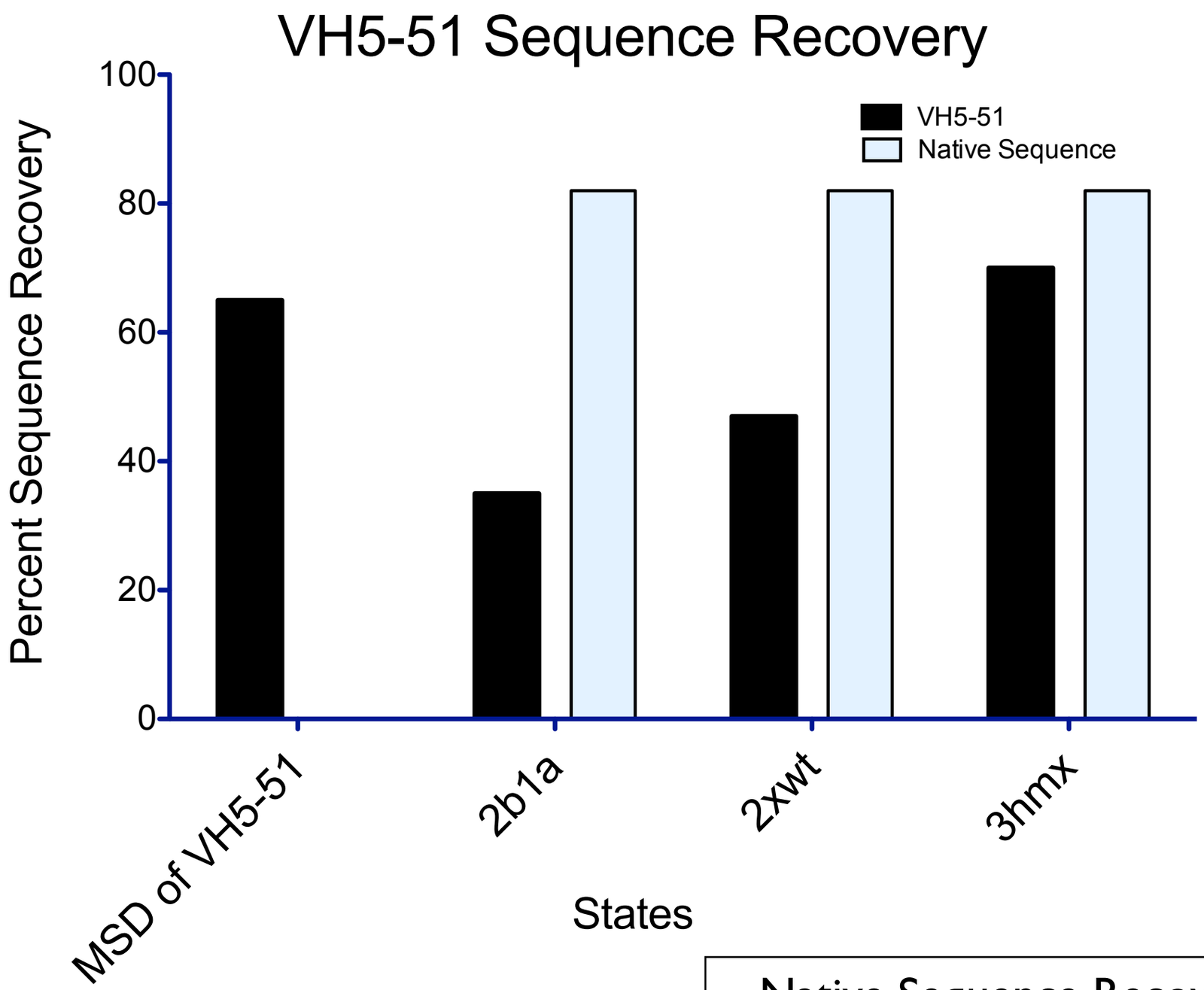


VH5-51	G	F	T	S	Y	M	I	T	Q	I	K	S	T	A	Y	A	M	R
2B1A	T	F	S	D	Y	M	F	S	E	M	R	N	T	A	H	P	L	R

VH5-51 MSD/SSD Design

Design	Percentage recovered to native	Percentage to VH5-51
MSD of VH5-51 States	-	65
2b1a	82	35
2xwt	82	47
3hmx	82	70

VH5-51 MSD/SSD Design



Native Sequence Recovery Average = 82%
VH1-69 Recovery Average = 50.6%

Conclusions

- Multi-state design recovers sequences closer to germline progenitor.

- Single state design recovers sequences closer to native (mature) antibody sequences, showing an *in silico* maturation.

- Germline sequences are optimally flexible in frequently used germline genes to accommodate binding of many antigens.

Future Directions

- Combine states of frequent and infrequently used germline genes to see which sequences are recovered.

- Full quanitative workup to find frequently used amino acids (PSSM)

- Iterative relax and MSD to accommodate clashing rotamers and improve sequence recovery.

- Apply MSD to HIV antibodies to bind a diverse panel of antigens

Acknowledgements

James Crowe, MD

Bryan Briney

Mark Hicar MD, PhD

Gordon Lemmon

Jens Meiler, PhD

David Nannemann, PhD

Sam Deluca

Steven Combs

Acknowledgments.

Crowe Lab

- Natalie Thornburg PhD
- Bryan Briney
- Gopal Sapparapu
- Mohammed Aiyegbo
- Frances House, MS
- Fyza Shaikh
- John Bates, PhD
- Mark Hicar, MD PhD
- Lidiya Pundor
- Valentine Chukwuma
- Scott Smith MD PhD
- Chelsey Huffman

Meiler Lab

- Will Lowe, PhD
- Brian Weiner PhD
- Carrie Fortenberry MS
- Jeff Mendenhall
- Caitlin Heuberger
- Mert Karakas
- Nils Woetzal
- Mariusz Butkiewicz
- Steffen Lindert
- Julia Koehler
- Sten Heinze
- Elizabeth Dong
- Nathan Alexander
- Gordon Lemmon
- David Nannemann
- Steven Combs
- Sam DeLuca
- Stephanie Hirst
- Gregory Sliwoski
- Britney Allison

Spearman Lab, Emory

- Paul Spearman MD
- Xuemin Chen

Williams Lab

- John Williams MD
- Reagan Cox

Kortemme Lab

- Colin Smith
- Mariana Babor
- Tanja Kortemme

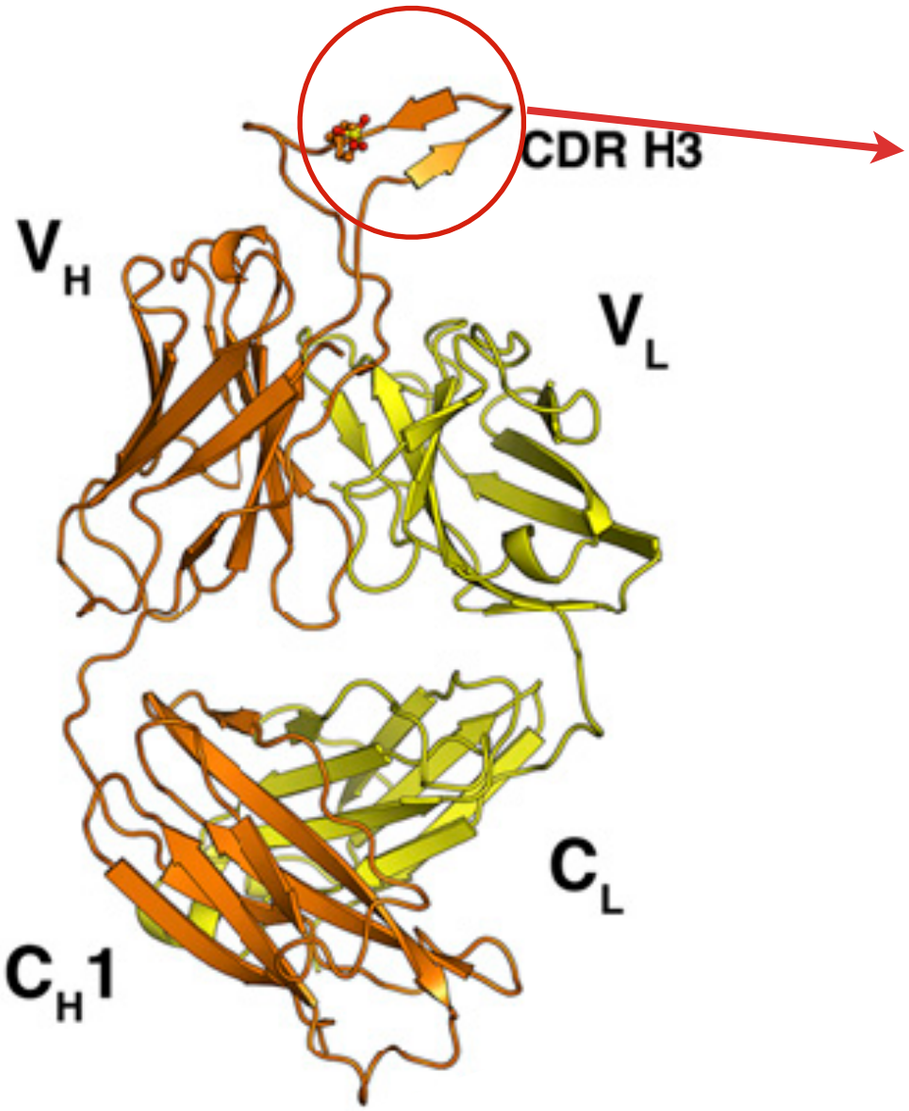
Kuhlman Lab

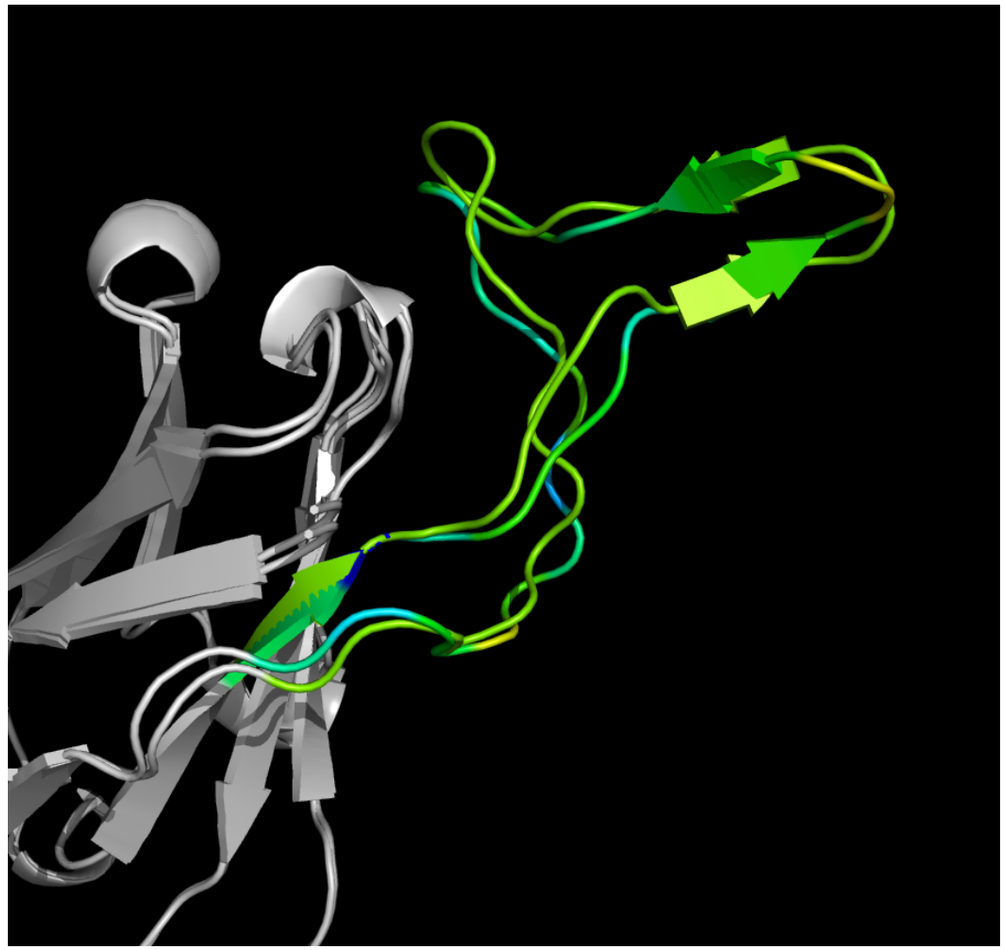
- Brian Kuhlman
- **Andrew Leaver-Fay**

Funding

- NIH U01 - Clonal Analysis of the Human B-Cell Response
- NIH U01 - Broadly Neutralizing Monoclonal Antibodies to HIV-1
- HIV Training Grant - M&IM (Chris Aiken)

Antibody Diversity

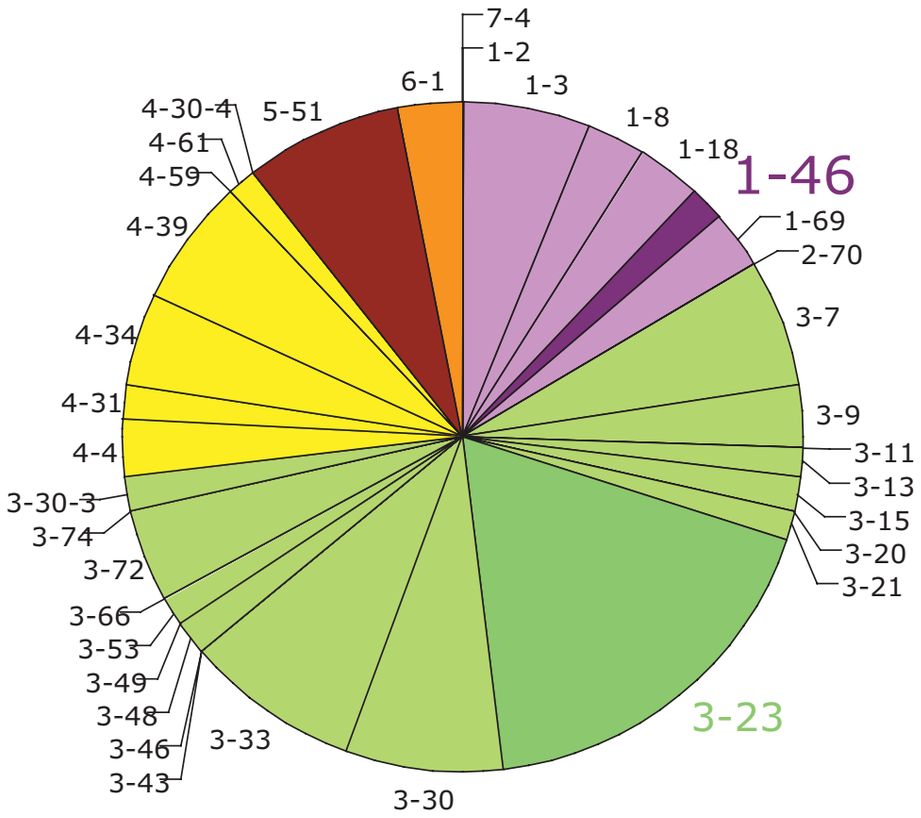




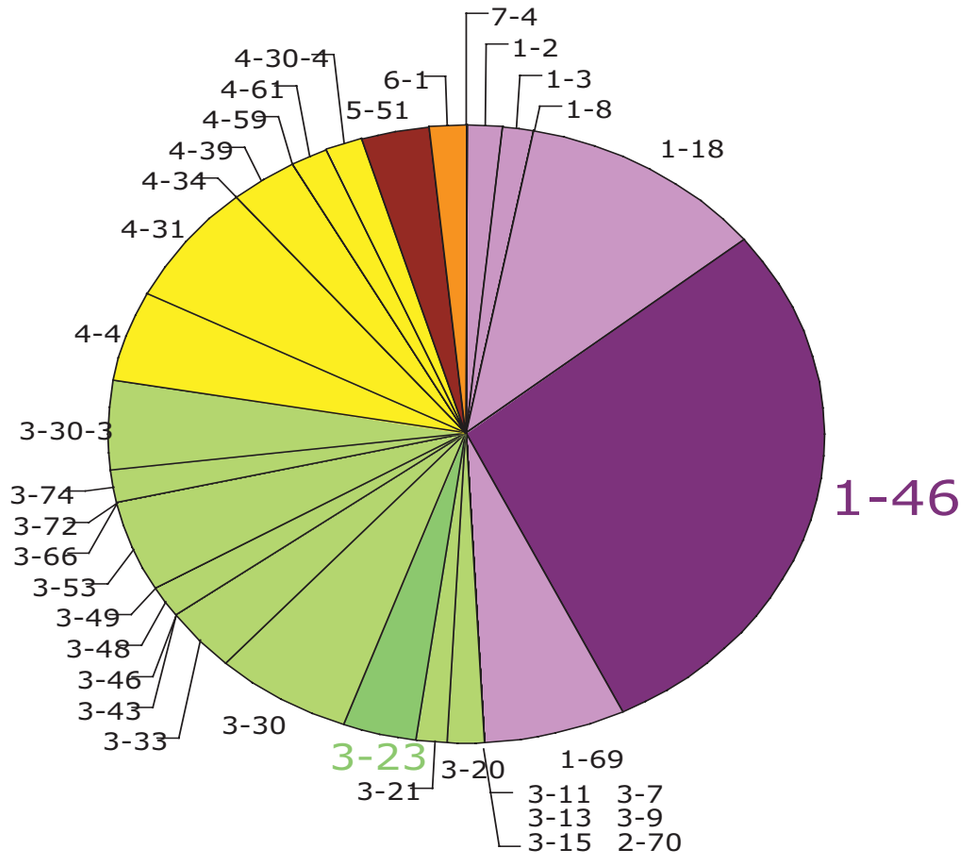
Pejchal et al., *PNAS* 2010

Gene Usage is Driven by Structure

- Tian et. al reported on healthy and diseased repertoire using Sanger sequencing



Healthy Donor



RSV Infected

Tian et. al Immunologoy 2007