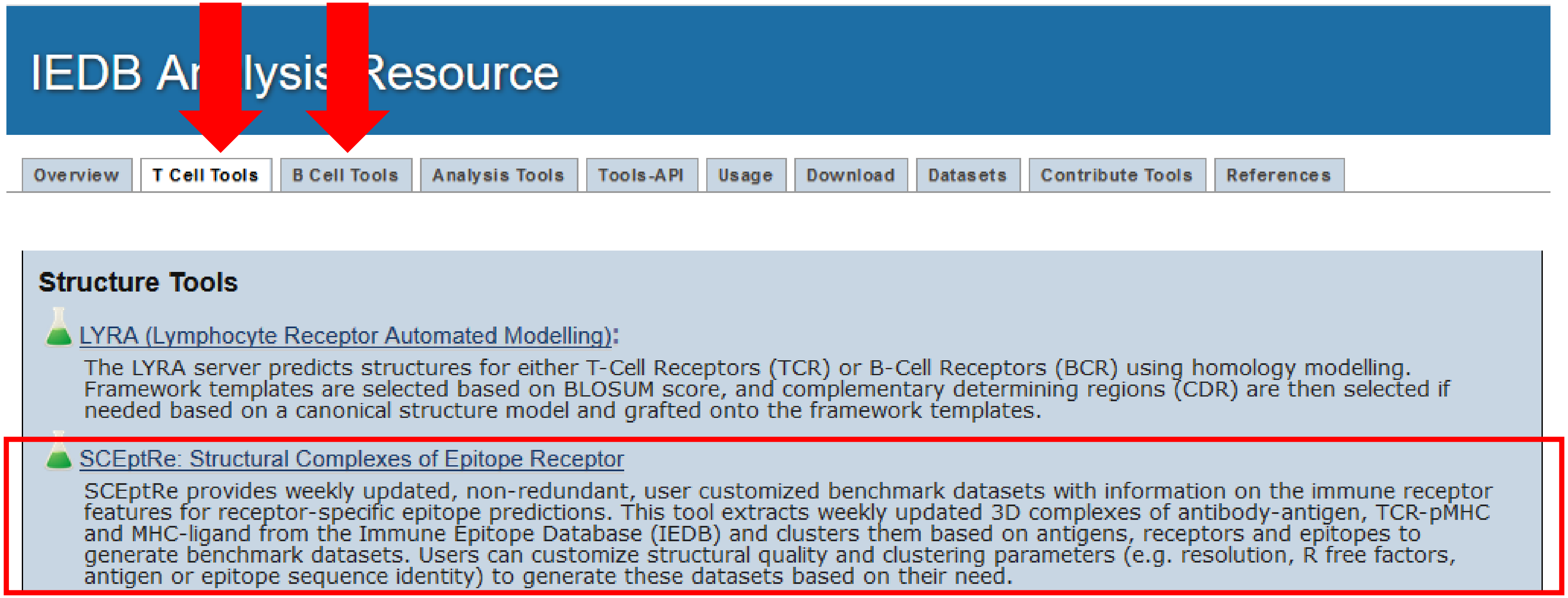
IMMUNE EPITOPE DATABASE
AND ANALYSIS RESOURCE

LYRA & SCEptRe

tools.iedb.org

Presented by: Morten Nielsen, Professor

Tools accessible from both T Cell & B Cell tabs



Structure Tools

LYRA (Lymphocyte Receptor Automated Modelling):

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.

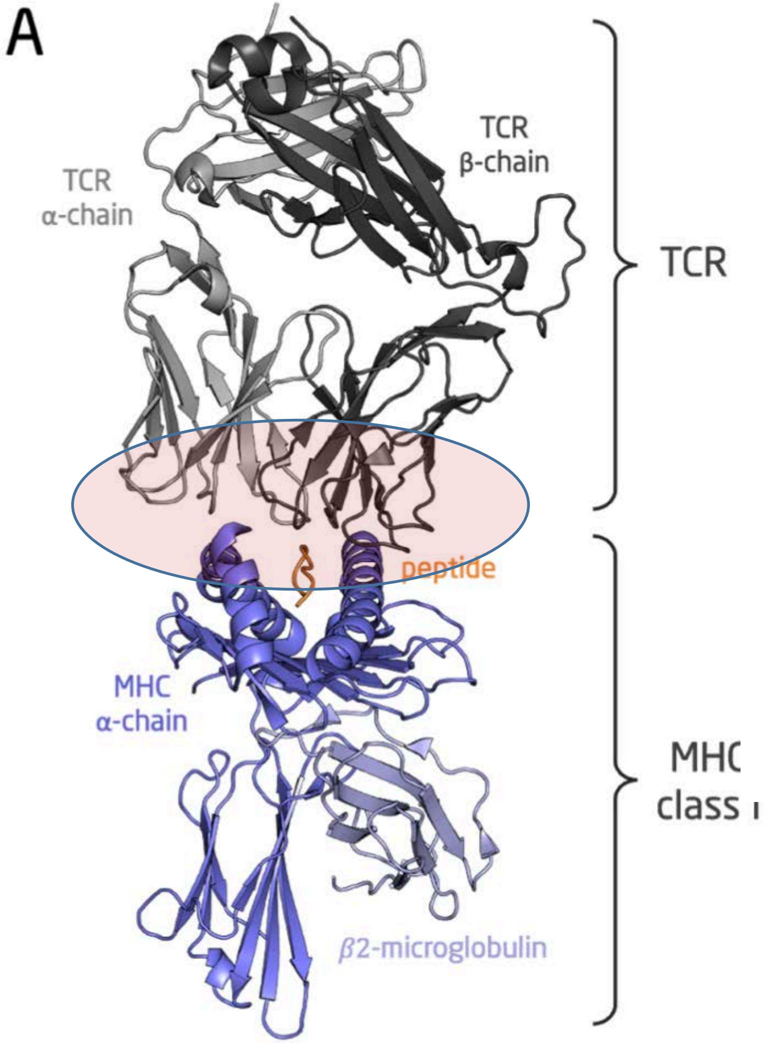
SCEptRe: Structural Complexes of Epitope Receptor

SCEptRe provides weekly updated, non-redundant, user customized benchmark datasets with information on the immune receptor features for receptor-specific epitope predictions. This tool extracts weekly updated 3D complexes of antibody-antigen, TCR-pMHC and MHC-ligand from the Immune Epitope Database (IEDB) and clusters them based on antigens, receptors and epitopes to generate benchmark datasets. Users can customize structural quality and clustering parameters (e.g. resolution, R free factors, antigen or epitope sequence identity) to generate these datasets based on their need.

SCEptRe: Structural Complexes of Epitope-Receptor

- Provides automatically updated benchmark datasets based on curated structural IEDB data
- User defines criteria to select set of quality and clustering parameters
- Clustering to address redundancy (and overfitting) in the resulting dataset when used for epitope predictions.
- Antibody-antigen, TCR-pMHC and MHC-ligand complexes

What are BCR and TCR structures



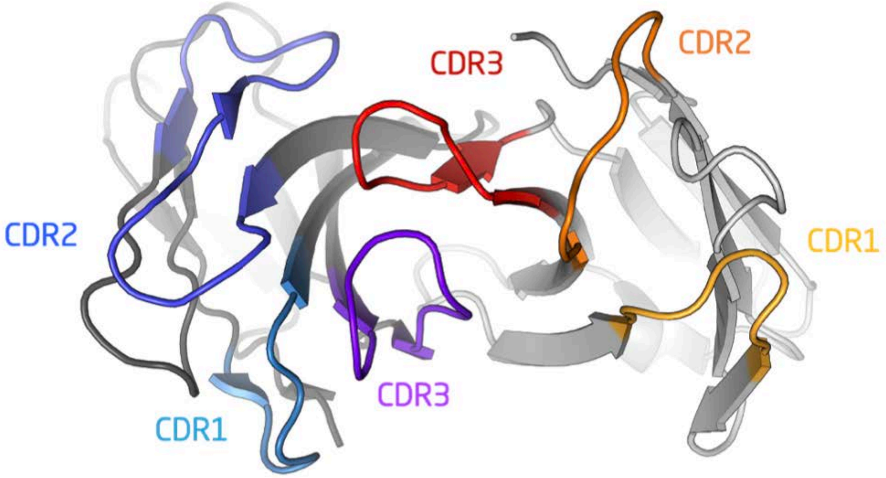


Figure curtesy KK Munk.

Why is this important?

- Most current methods for epitope prediction are BCR/TCR agnostic
- Advances in T and B cell profiling by sequencing T and B cell receptors (TCR/BCR)
- Limited ability to interpret these data, since we most often do not know the cognate target of the differentially profiled TCR/BCRs
- The next major computational challenge is to develop tools to predict this
- TCR/BCR data sets are however very redundant, with multiple highly similar sequences

Why is this important?

- Example from the IEDB. TCRb targeting GLCTLVAML

ASSQSPGGTQY
ASSQSPGGTQYS
ASSQSPGGTQYL

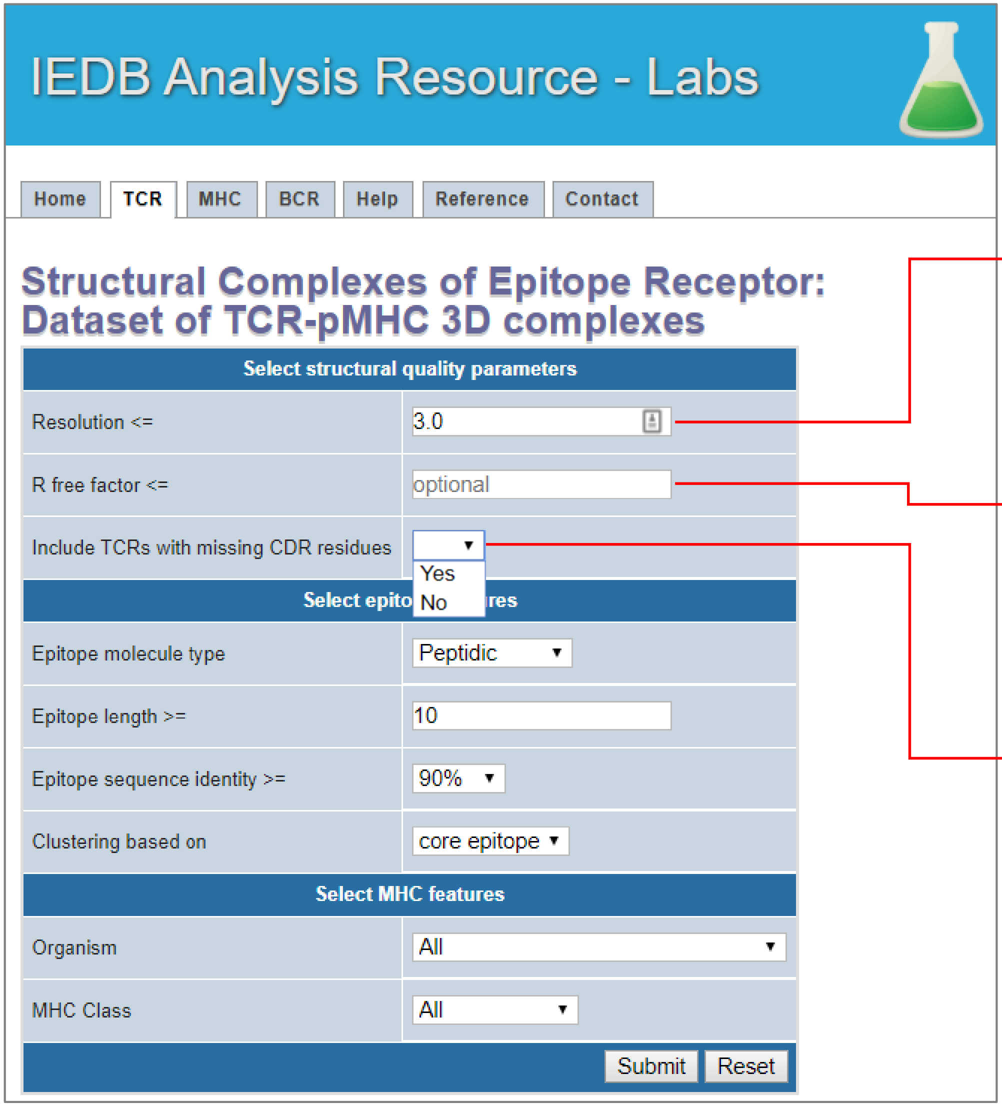
ASSQSPGGTQH
ASSQSPGGTQN
ASSQSPGGTQS

- Likewise do many distinct BCR/TCRs share identical CDR sequence
- and often a TCR binds several very similar peptides

ASSLDQGAQDNEQF YLYDRLLRI
ASSLDQGAQDNEQF YLYDRLLRV

- If one seeks to make extrapolatable predictions/rules (and not just reproduce the measurements), one needs to deal with this redundancy

SCEptRe: TCR-pMHC 3D complexes



Structural Complexes of Epitope Receptor: Dataset of TCR-pMHC 3D complexes

Select structural quality parameters	
Resolution <=	3.0
R free factor <=	optional
Include TCRs with missing CDR residues	Yes No

Select epitope features	
Epitope molecule type	Peptidic
Epitope length >=	10
Epitope sequence identity >=	90%
Clustering based on	core epitope

Select MHC features	
Organism	All
MHC Class	All

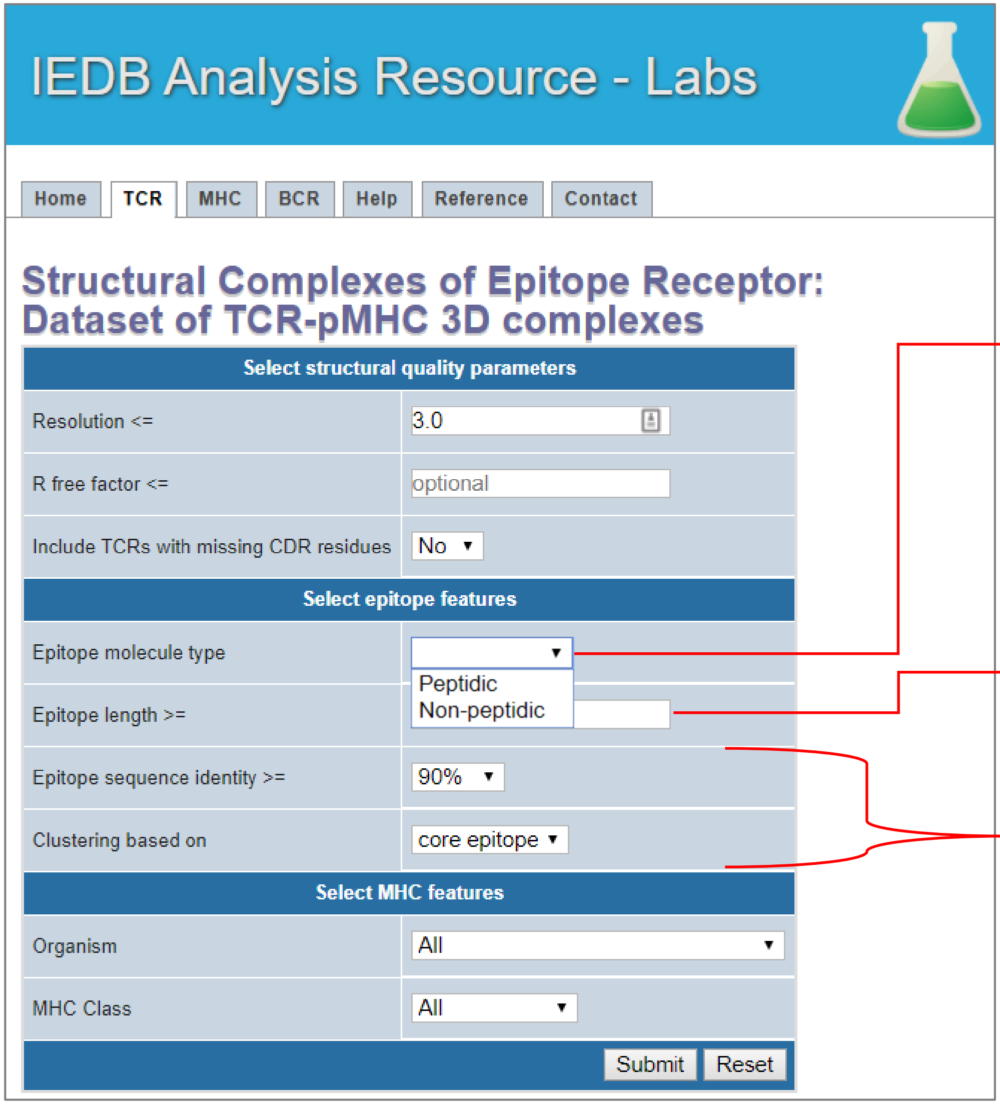
http://tools.iedb.org/sceptre/tcr/

Resolution is a measure of the quality of the crystal structures

R-free factor is a measure of the quality of the atomic model obtained from the crystallographic data

Decide whether to include complexes with missing residues in the CDR regions in the final dataset

SCEptRe: TCR-pMHC 3D complexes



Structural Complexes of Epitope Receptor: Dataset of TCR-pMHC 3D complexes

Select structural quality parameters	
Resolution <=	3.0
R free factor <=	optional
Include TCRs with missing CDR residues	No
Select epitope features	
Epitope molecule type	Peptidic Non-peptidic
Epitope length >=	
Epitope sequence identity >=	90%
Clustering based on	core epitope
Select MHC features	
Organism	All
MHC Class	All

http://tools.iedb.org/sceptre/tcr/

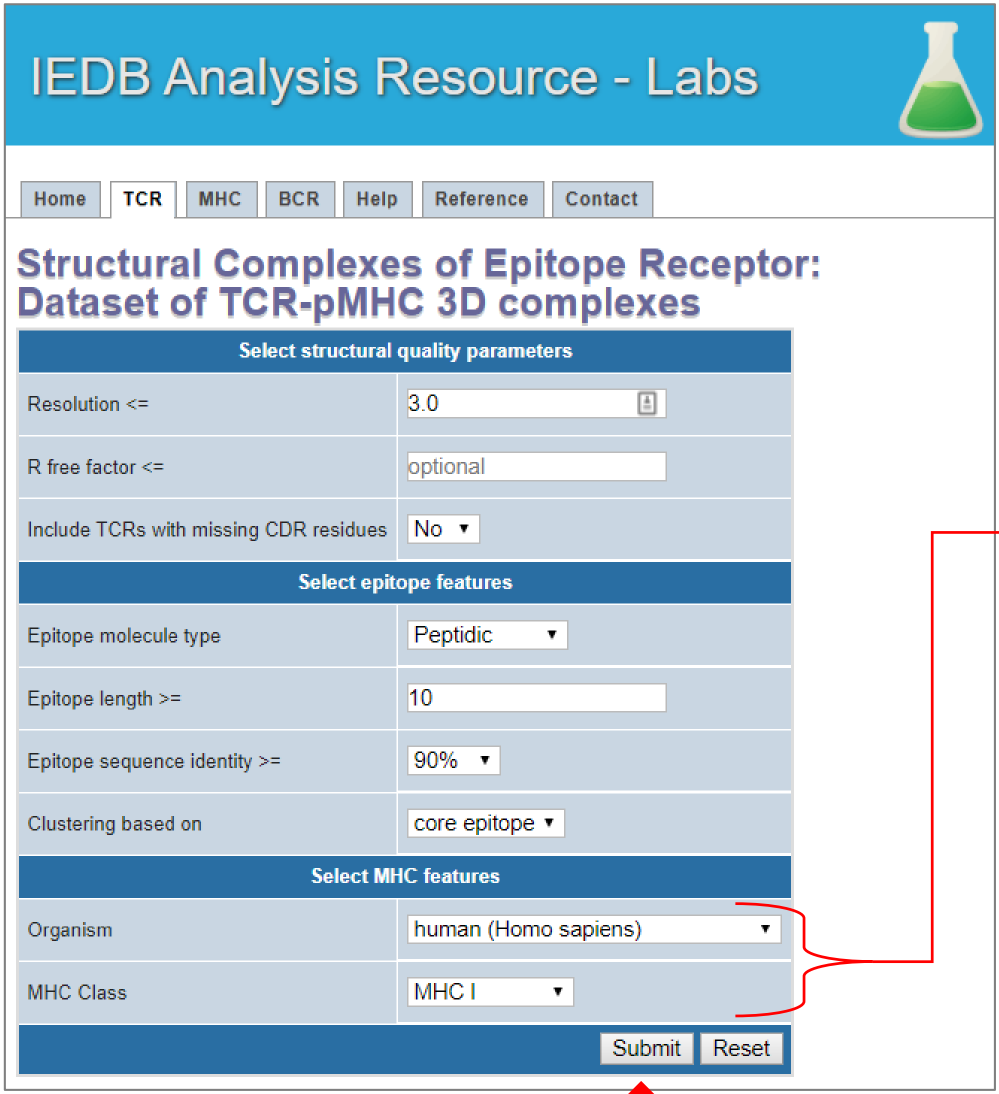
Peptidic - clustered based on epitope sequence, TCR CDR sequences and MHC G-domains
Non-peptidic - filtered based on specified structural quality parameters & clustered based on TCR CDR sequences and MHC G-domains

Filter out complexes below the specified epitope length threshold

Clustering Parameters

- Specify sequence identity threshold
- Choose core-epitopes (peptide residues in direct contact) or whole epitope sequences

SCEptRe: TCR-pMHC 3D complexes



IEDB Analysis Resource - Labs

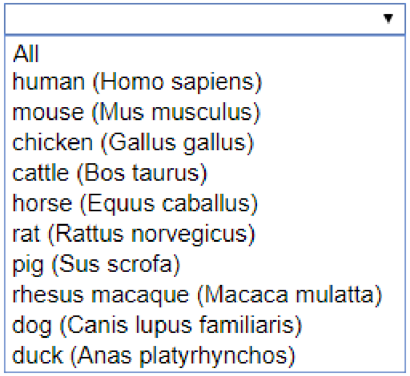
Structural Complexes of Epitope Receptor: Dataset of TCR-pMHC 3D complexes

Select structural quality parameters	
Resolution <=	3.0
R free factor <=	optional
Include TCRs with missing CDR residues	No
Select epitope features	
Epitope molecule type	Peptidic
Epitope length >=	10
Epitope sequence identity >=	90%
Clustering based on	core epitope
Select MHC features	
Organism	human (Homo sapiens)
MHC Class	MHC I

http://tools.iedb.org/sceptre/tcr/

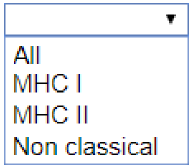
Filter complexes based on:

- MHC source organism



All
human (Homo sapiens)
mouse (Mus musculus)
chicken (Gallus gallus)
cattle (Bos taurus)
horse (Equus caballus)
rat (Rattus norvegicus)
pig (Sus scrofa)
rhesus macaque (Macaca mulatta)
dog (Canis lupus familiaris)
duck (Anas platyrhynchos)

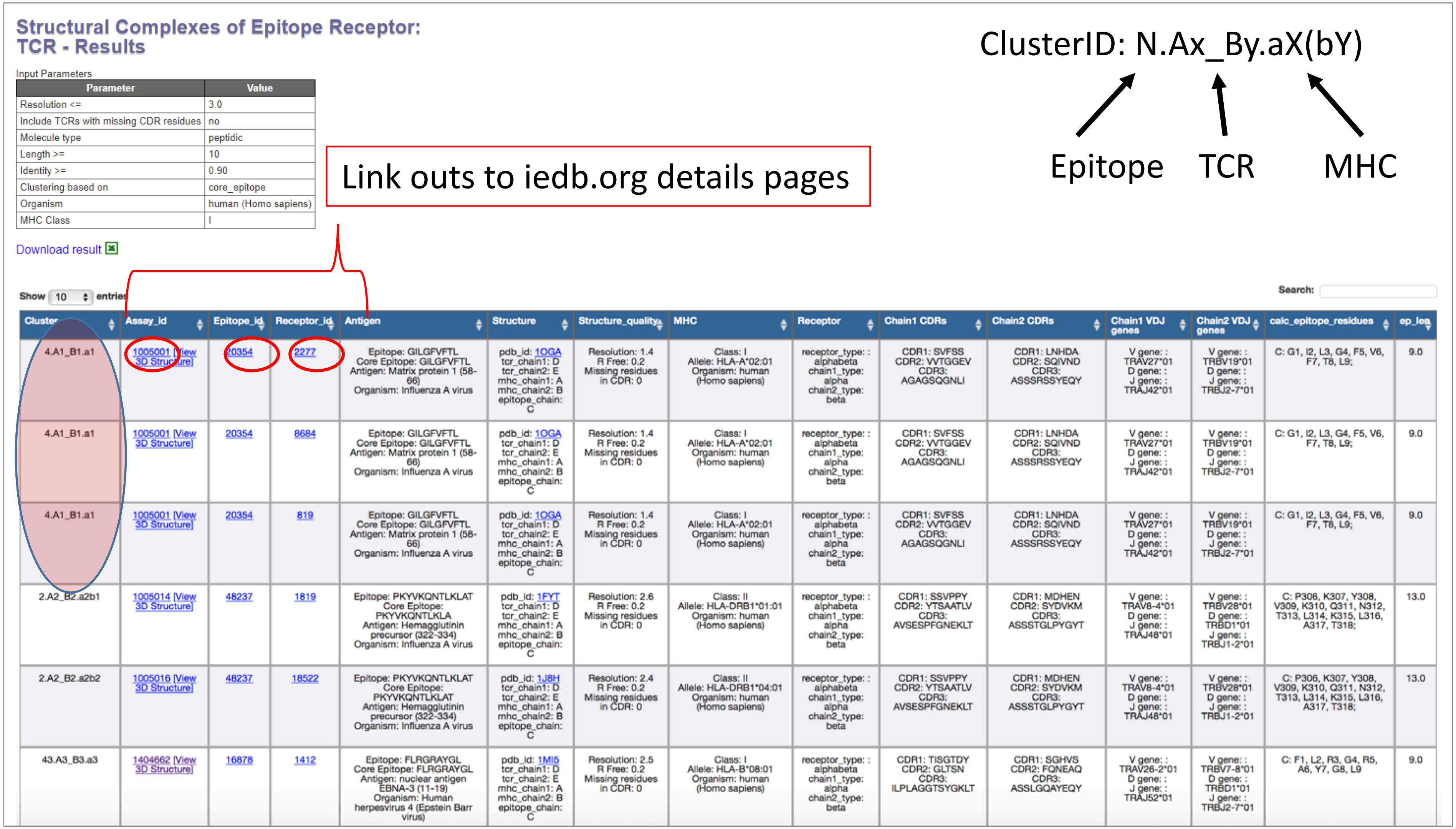
- MHC class



All
MHC I
MHC II
Non classical

SCEptRe: TCR-pMHC 3D complexes -results

http://tools.iedb.org/sceptre/tcr/



Structural Complexes of Epitope Receptor: TCR - Results

Input Parameters

Parameter	Value
Resolution <=	3.0
Include TCRs with missing CDR residues	no
Molecule type	peptidic
Length >=	10
Identity >=	0.90
Clustering based on	core_epitope
Organism	human (Homo sapiens)
MHC Class	I

Download result

Show 10 entries

Search:

Cluster	Assay_id	Epitope_id	Receptor_id	Antigen	Structure	Structure_quality	MHC	Receptor	Chain1 CDRs	Chain2 CDRs	Chain1 VDJ genes	Chain2 VDJ genes	calc_epitope_residues	ep_len
4.A1_B1.a1	1005001 [View 3D Structure]	20354	2277	Epitope: GILGFVFTL Core Epitope: GILGFVFTL Antigen: Matrix protein 1 (58-66) Organism: Influenza A virus	pdb_id: 1OGA tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 1.4 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-A*02:01 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: SVFSS CDR2: VVTGGEV CDR3: AGAGSQGNLI	CDR1: LNHDA CDR2: SQIVND CDR3: ASSSRSSYEQY	V gene: : TRAV27*01 D gene: : J gene: : TRAJ42*01	V gene: : TRBV19*01 D gene: : J gene: : TRBJ2-7*01	C: G1, I2, L3, G4, F5, V6, F7, T8, L9;	9.0
4.A1_B1.a1	1005001 [View 3D Structure]	20354	8684	Epitope: GILGFVFTL Core Epitope: GILGFVFTL Antigen: Matrix protein 1 (58-66) Organism: Influenza A virus	pdb_id: 1OGA tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 1.4 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-A*02:01 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: SVFSS CDR2: VVTGGEV CDR3: AGAGSQGNLI	CDR1: LNHDA CDR2: SQIVND CDR3: ASSSRSSYEQY	V gene: : TRAV27*01 D gene: : J gene: : TRAJ42*01	V gene: : TRBV19*01 D gene: : J gene: : TRBJ2-7*01	C: G1, I2, L3, G4, F5, V6, F7, T8, L9;	9.0
4.A1_B1.a1	1005001 [View 3D Structure]	20354	819	Epitope: GILGFVFTL Core Epitope: GILGFVFTL Antigen: Matrix protein 1 (58-66) Organism: Influenza A virus	pdb_id: 1OGA tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 1.4 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-A*02:01 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: SVFSS CDR2: VVTGGEV CDR3: AGAGSQGNLI	CDR1: LNHDA CDR2: SQIVND CDR3: ASSSRSSYEQY	V gene: : TRAV27*01 D gene: : J gene: : TRAJ42*01	V gene: : TRBV19*01 D gene: : J gene: : TRBJ2-7*01	C: G1, I2, L3, G4, F5, V6, F7, T8, L9;	9.0
2.A2_B2.a2b1	1005014 [View 3D Structure]	48237	1819	Epitope: PKYVKQNTLKLAT Core Epitope: PKYVKQNTLKLA Antigen: Hemagglutinin precursor (322-334) Organism: Influenza A virus	pdb_id: 1FYT tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 2.6 R Free: 0.2 Missing residues in CDR: 0	Class: II Allele: HLA-DRB1*01:01 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: SSVPPY CDR2: YTSAATLV CDR3: AVSESPFGNEKLT	CDR1: MDHEN CDR2: SYDVKM CDR3: ASSSTGLPYGYT	V gene: : TRAV8-4*01 D gene: : J gene: : TRAJ48*01	V gene: : TRBV28*01 D gene: : TRBD1*01 J gene: : TRBJ1-2*01	C: P306, K307, Y308, V309, K310, Q311, N312, T313, L314, K315, L316, A317, T318;	13.0
2.A2_B2.a2b2	1005016 [View 3D Structure]	48237	18522	Epitope: PKYVKQNTLKLAT Core Epitope: PKYVKQNTLKLAT Antigen: Hemagglutinin precursor (322-334) Organism: Influenza A virus	pdb_id: 1J8H tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 2.4 R Free: 0.2 Missing residues in CDR: 0	Class: II Allele: HLA-DRB1*04:01 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: SSVPPY CDR2: YTSAATLV CDR3: AVSESPFGNEKLT	CDR1: MDHEN CDR2: SYDVKM CDR3: ASSSTGLPYGYT	V gene: : TRAV8-4*01 D gene: : J gene: : TRAJ48*01	V gene: : TRBV28*01 D gene: : J gene: : TRBJ1-2*01	C: P306, K307, Y308, V309, K310, Q311, N312, T313, L314, K315, L316, A317, T318;	13.0
43.A3_B3.a3	1404662 [View 3D Structure]	16878	1412	Epitope: FLRGRAYGL Core Epitope: FLRGRAYGL Antigen: nuclear antigen EBNA-3 (11-19) Organism: Human herpesvirus 4 (Epstein Barr virus)	pdb_id: 1MI5 tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 2.5 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-B*08:01 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: TISGTDY CDR2: GLTSN CDR3: ILPLAGGTSYGKLT	CDR1: SGHVS CDR2: FQNEAQ CDR3: ASSLGQAYEQY	V gene: : TRAV26-2*01 D gene: : J gene: : TRAJ52*01	V gene: : TRBV7-8*01 D gene: : TRBD1*01 J gene: : TRBJ2-7*01	C: F1, L2, R3, G4, R5, A6, Y7, G8, L9	9.0

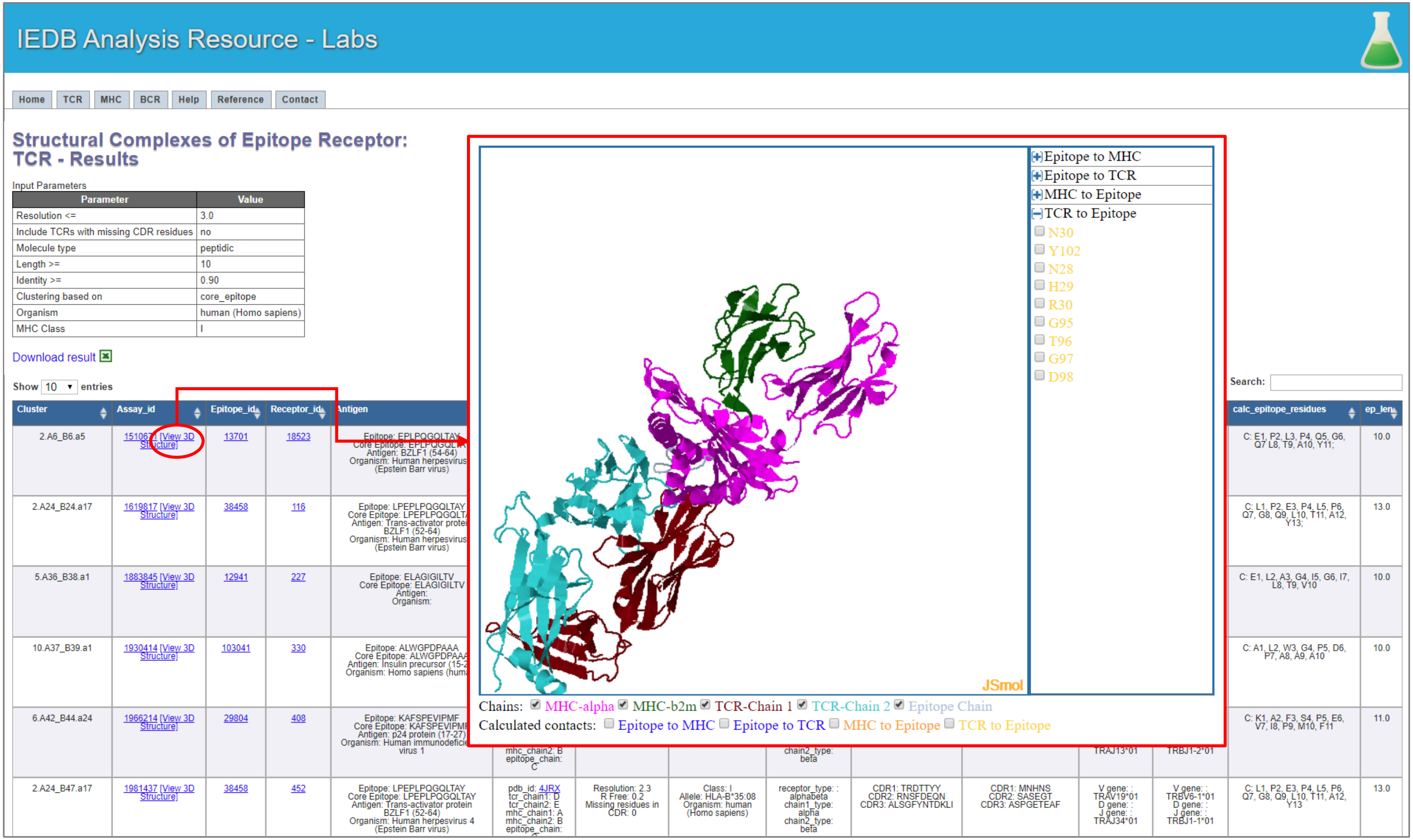
ClusterID: N.Ax_By.aX(bY)

Epitope TCR MHC

Link outs to iedb.org details pages

SCEptRe: TCR-pMHC 3D complexes -results

http://tools.iedb.org/sceptre/tcr/



Structural Complexes of Epitope Receptor: TCR - Results

Input Parameters

Parameter	Value
Resolution <=	3.0
Include TCRs with missing CDR residues	no
Molecule type	peptidic
Length >=	10
Identity >=	0.90
Clustering based on	core_epitope
Organism	human (Homo sapiens)
MHC Class	I

Download result

Show 10 entries

Search:

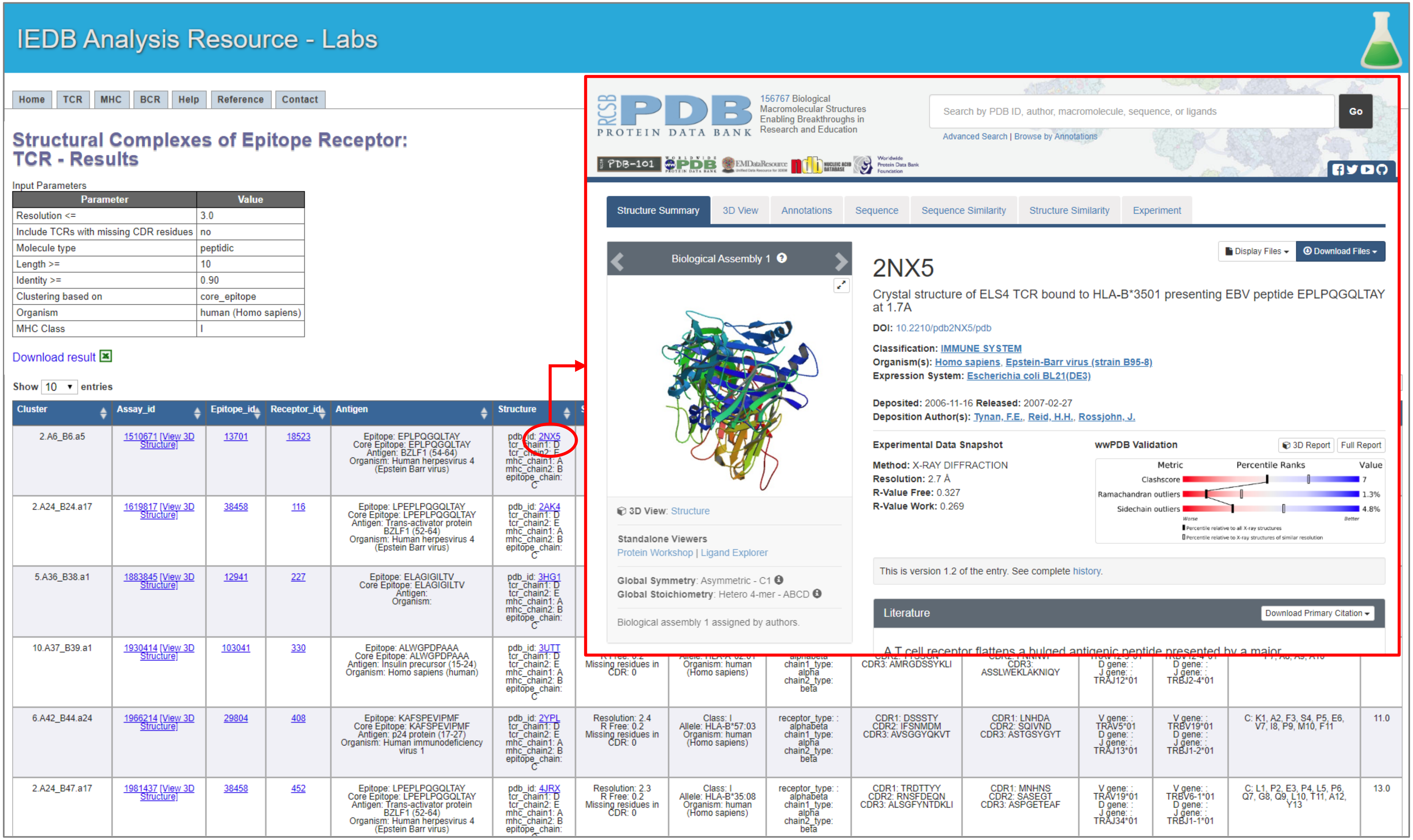
Cluster	Assay_id	Epitope_id	Receptor_id	Antigen								calc_epitope_residues	ep_len
2.A6_B6.a5	1510671 [View 3D Structure]	13701	18523	Epitope: EPLPQGQLTAY Core Epitope: EPLPQGQLTA Antigen: BZLF1 (54-64) Organism: Human herpesvirus (Epstein Barr virus)								C: E1, P2, L3, P4, Q5, G6, Q7 L8, T9, A10, Y11;	10.0
2.A24_B24.a17	1619817 [View 3D Structure]	38458	116	Epitope: LPEPLPQGQLTAY Core Epitope: LPEPLPQGQLTA Antigen: Trans-activator protei BZLF1 (52-64) Organism: Human herpesvirus (Epstein Barr virus)								C: L1, P2, E3, P4, L5, P6, Q7, G8, Q9, L10, T11, A12, Y13;	13.0
5.A36_B38.a1	1883845 [View 3D Structure]	12941	227	Epitope: ELAGIGILTV Core Epitope: ELAGIGILTV Antigen: Organism:								C: E1, L2, A3, G4, I5, G6, I7, L8, T9, V10	10.0
10.A37_B39.a1	1930414 [View 3D Structure]	103041	330	Epitope: ALWGPDPAAA Core Epitope: ALWGPDPAAA Antigen: Insulin precursor (15-2 Organism: Homo sapiens (huma								C: A1, L2, W3, G4, P5, D6, P7, A8, A9, A10	10.0
6.A42_B44.a24	1966214 [View 3D Structure]	29804	408	Epitope: KAFSPEVIPMF Core Epitope: KAFSPEVIPMF Antigen: p24 protein (17-27) Organism: Human immunodeficie virus 1	mhc_chain2: B epitope_chain: C			chain2_type: beta			TRAJ13*01	TRBJ1-2*01	C: K1, A2, F3, S4, P5, E6, V7, I8, P9, M10, F11	11.0
2.A24_B47.a17	1981437 [View 3D Structure]	38458	452	Epitope: LPEPLPQGQLTAY Core Epitope: LPEPLPQGQLTAY Antigen: Trans-activator protein BZLF1 (52-64) Organism: Human herpesvirus 4 (Epstein Barr virus)	pdb_id: 4JRX tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 2.3 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-B*35:08 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: TRDTTYY CDR2: RNSFDEQN CDR3: ALSGFYNTDKLI	CDR1: MNHNS CDR2: SASEGT CDR3: ASPGETEAF	V gene: : TRAV19*01 D gene: : J gene: : TRAJ34*01	V gene: : TRBV6-1*01 D gene: : J gene: : TRBJ1-1*01	C: L1, P2, E3, P4, L5, P6, Q7, G8, Q9, L10, T11, A12, Y13	13.0

Chains: ☑ MHC-alpha ☑ MHC-b2m ☑ TCR-Chain 1 ☑ TCR-Chain 2 ☑ Epitope Chain

Calculated contacts: ☐ Epitope to MHC ☐ Epitope to TCR ☐ MHC to Epitope ☐ TCR to Epitope

SCEptRe: TCR-pMHC 3D complexes -results

http://tools.iedb.org/sceptre/tcr/



Structural Complexes of Epitope Receptor: TCR - Results

Input Parameters

Parameter	Value
Resolution <=	3.0
Include TCRs with missing CDR residues	no
Molecule type	peptidic
Length >=	10
Identity >=	0.90
Clustering based on	core_epitope
Organism	human (Homo sapiens)
MHC Class	I

Download result

Show 10 entries

Cluster	Assay_id	Epitope_id	Receptor_id	Antigen	Structure									
2.A6_B6.a5	1510671 [View 3D Structure]	13701	18523	Epitope: EPLPQGQLTAY Core Epitope: EPLPQGQLTAY Antigen: BZLF1 (54-64) Organism: Human herpesvirus 4 (Epstein Barr virus)	pdb_id: 2NX5 tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C									
2.A24_B24.a17	1619817 [View 3D Structure]	38458	116	Epitope: LPEPLPQGQLTAY Core Epitope: LPEPLPQGQLTAY Antigen: Trans-activator protein BZLF1 (52-64) Organism: Human herpesvirus 4 (Epstein Barr virus)	pdb_id: 2AK4 tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C									
5.A36_B38.a1	1883845 [View 3D Structure]	12941	227	Epitope: ELAGIGILTV Core Epitope: ELAGIGILTV Antigen: Organism:	pdb_id: 3HG1 tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C									
10.A37_B39.a1	1930414 [View 3D Structure]	103041	330	Epitope: ALWGPDPAAA Core Epitope: ALWGPDPAAA Antigen: Insulin precursor (15-24) Organism: Homo sapiens (human)	pdb_id: 3UTT tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Missing residues in CDR: 0	Organism: human (Homo sapiens)	chain1_type: alpha chain2_type: beta	CDR3: AMRGDSSYKLI	CDR3: ASSLWEKLAKNIQY	D gene: : J gene: : TRAJ12*01	D gene: : J gene: : TRBJ2-4*01		
6.A42_B44.a24	1966214 [View 3D Structure]	29804	408	Epitope: KAFSPEVIPMF Core Epitope: KAFSPEVIPMF Antigen: p24 protein (17-27) Organism: Human immunodeficiency virus 1	pdb_id: 2YPL tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain: C	Resolution: 2.4 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-B*57:03 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: DSSSTY CDR2: IFSNMDM CDR3: AVSGGYQKVT	CDR1: LNHDA CDR2: SQIVND CDR3: ASTGSYGYT	V gene: : TRAV5*01 D gene: : J gene: : TRAJ13*01	V gene: : TRBV19*01 D gene: : J gene: : TRBJ1-2*01	C: K1, A2, F3, S4, P5, E6, V7, I8, P9, M10, F11	11.0
2.A24_B47.a17	1981437 [View 3D Structure]	38458	452	Epitope: LPEPLPQGQLTAY Core Epitope: LPEPLPQGQLTAY Antigen: Trans-activator protein BZLF1 (52-64) Organism: Human herpesvirus 4 (Epstein Barr virus)	pdb_id: 4JRX tcr_chain1: D tcr_chain2: E mhc_chain1: A mhc_chain2: B epitope_chain:	Resolution: 2.3 R Free: 0.2 Missing residues in CDR: 0	Class: I Allele: HLA-B*35:08 Organism: human (Homo sapiens)	receptor_type: : alphabeta chain1_type: alpha chain2_type: beta	CDR1: TRDTTYY CDR2: RNSFDEQN CDR3: ALSGFYNTDKLI	CDR1: MNHNS CDR2: SASEGT CDR3: ASPGETEAF	V gene: : TRAV19*01 D gene: : J gene: : TRAJ34*01	V gene: : TRBV6-1*01 D gene: : J gene: : TRBJ1-1*01	C: L1, P2, E3, P4, L5, P6, Q7, G8, Q9, L10, T11, A12, Y13	13.0

2NX5

Crystal structure of ELS4 TCR bound to HLA-B*3501 presenting EBV peptide EPLPQGQLTAY at 1.7A

DOI: 10.2210/pdb2NX5/pdb

Classification: IMMUNE SYSTEM
Organism(s): Homo sapiens, Epstein-Barr virus (strain B95-8)
Expression System: Escherichia coli BL21(DE3)

Deposited: 2006-11-16 **Released:** 2007-02-27
Deposition Author(s): Tynan, F.E., Reid, H.H., Rossjohn, J.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 2.7 Å
R-Value Free: 0.327
R-Value Work: 0.269

wwPDB Validation

Metric	Value
Clashscore	7
Ramachandran outliers	1.3%
Sidechain outliers	4.8%

3D View: Structure

Standalone Viewers
Protein Workshop | Ligand Explorer

Global Symmetry: Asymmetric - C1
Global Stoichiometry: Hetero 4-mer - ABCD

Biological assembly 1 assigned by authors.

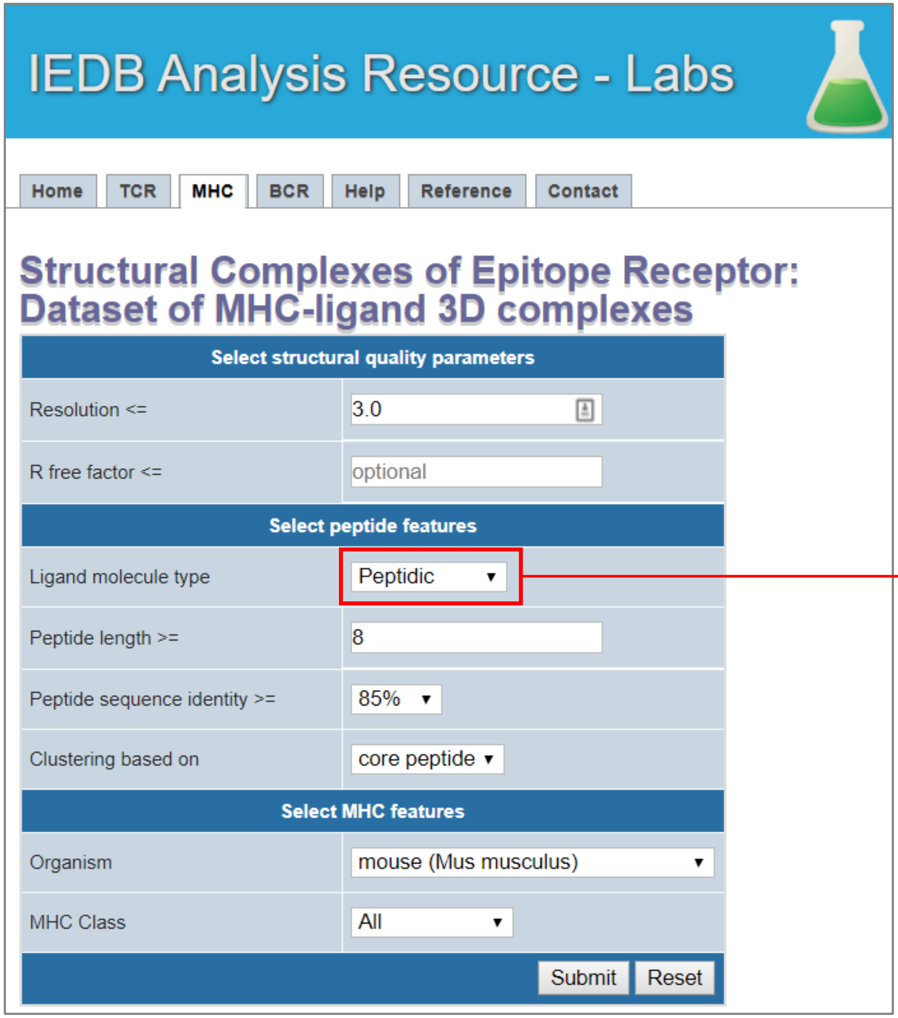
This is version 1.2 of the entry. See complete history.

Literature

A T cell receptor flattens a bulged antigenic peptide presented by a major

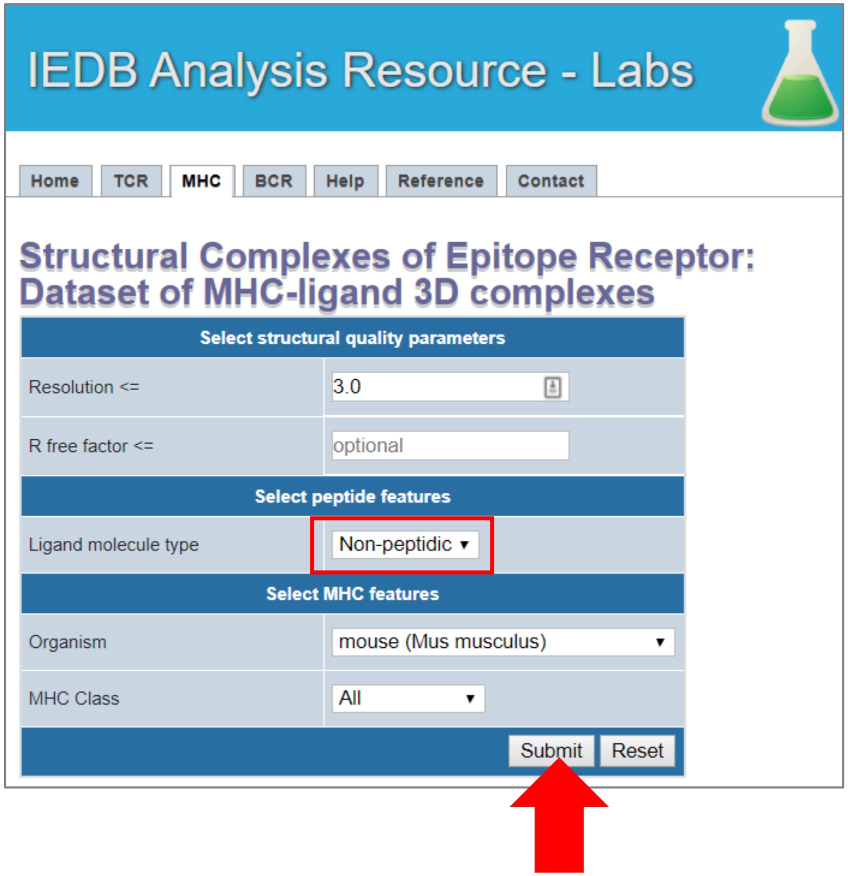
SCEptRe: MHC-ligand 3D complexes

http://tools.iedb.org/sceptre/mhc/



Structural Complexes of Epitope Receptor: Dataset of MHC-ligand 3D complexes

Select structural quality parameters	
Resolution <=	3.0
R free factor <=	optional
Select peptide features	
Ligand molecule type	Peptidic ▾
Peptide length >=	8
Peptide sequence identity >=	85% ▾
Clustering based on	core peptide ▾
Select MHC features	
Organism	mouse (Mus musculus) ▾
MHC Class	All ▾

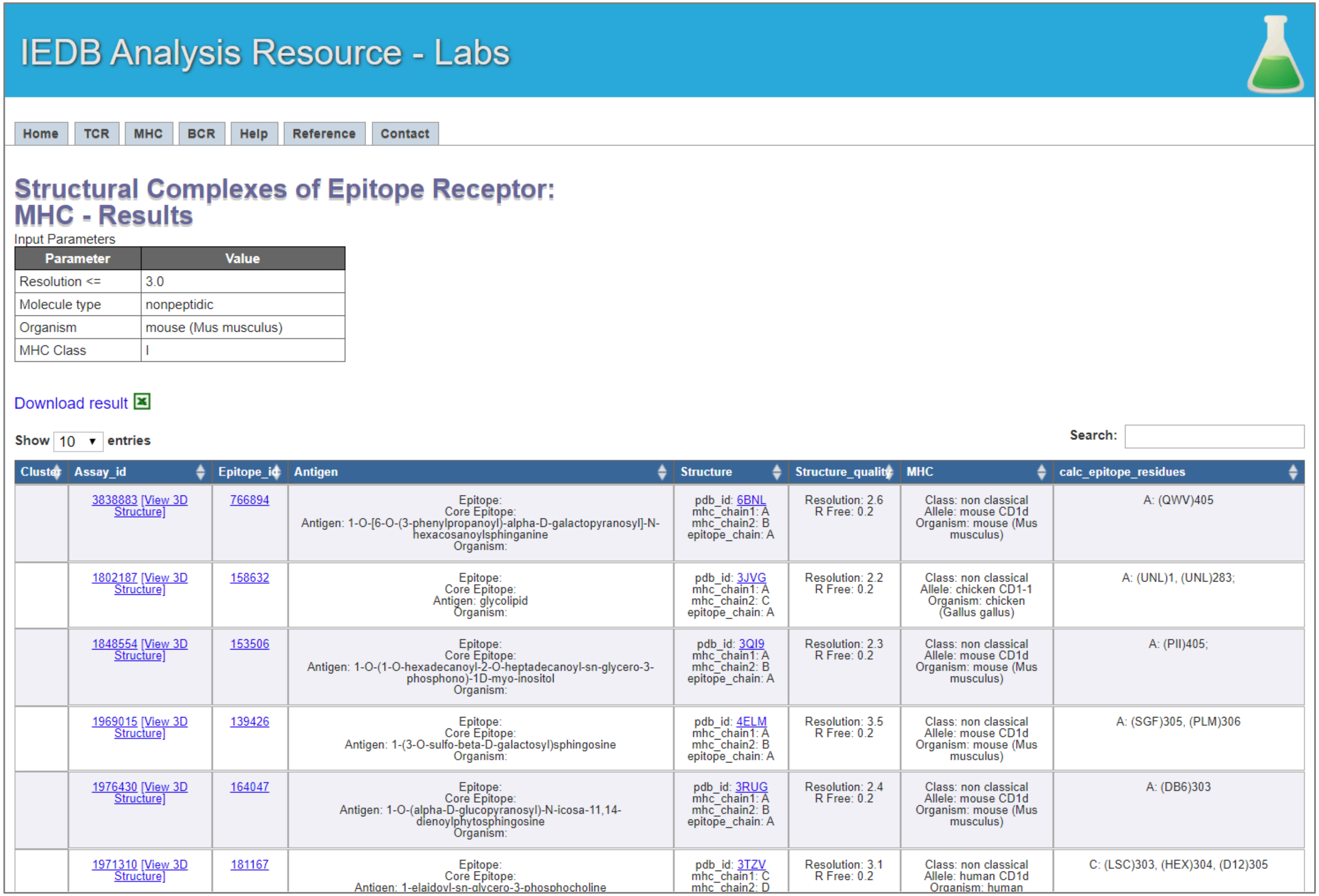


Structural Complexes of Epitope Receptor: Dataset of MHC-ligand 3D complexes

Select structural quality parameters	
Resolution <=	3.0
R free factor <=	optional
Select peptide features	
Ligand molecule type	Non-peptidic ▾
Select MHC features	
Organism	mouse (Mus musculus) ▾
MHC Class	All ▾

SCEptRe: MHC-ligand 3D complexes -results

http://tools.iedb.org/sceptre/mhc/



Structural Complexes of Epitope Receptor: MHC - Results

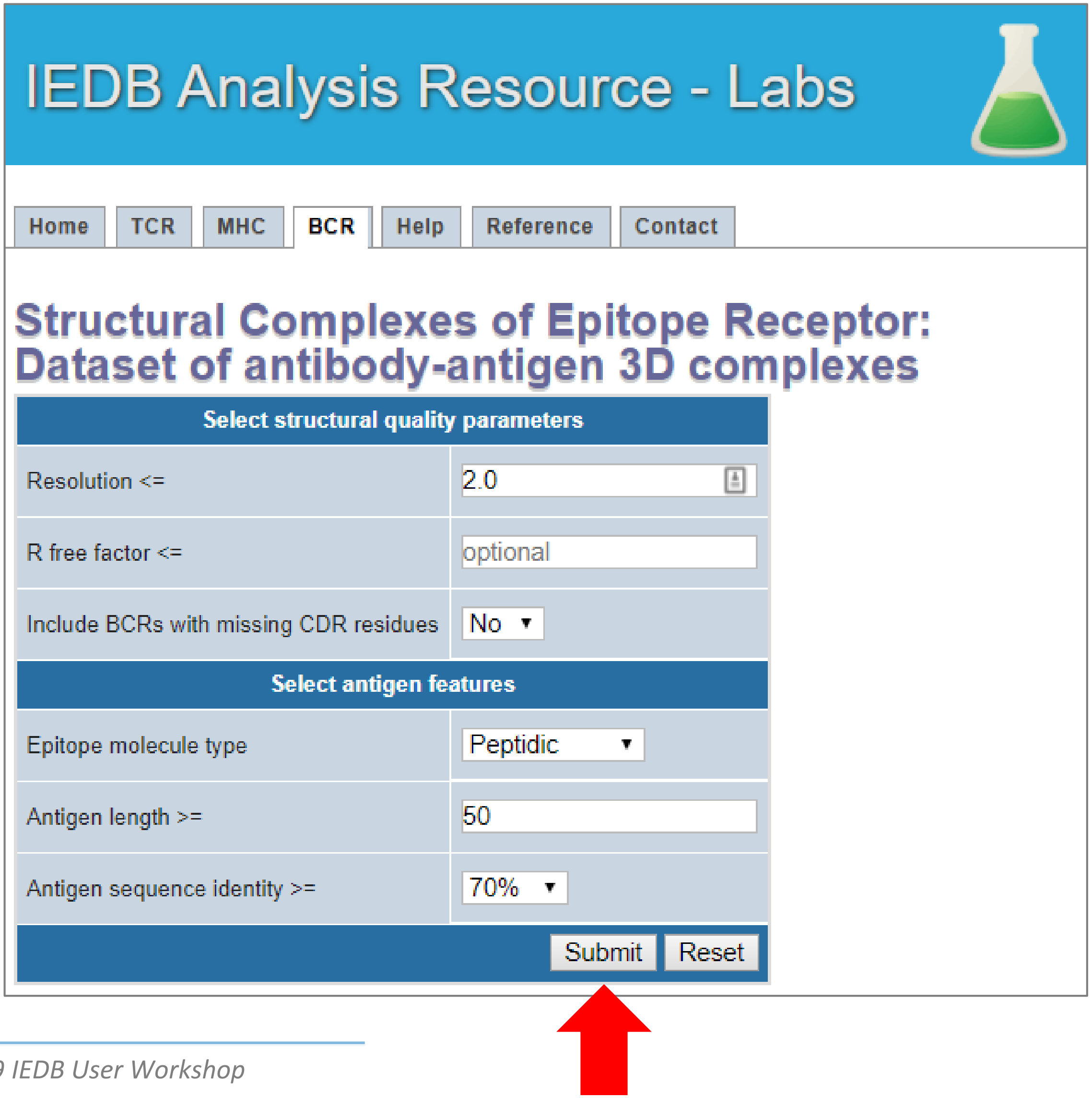
Input Parameters

Parameter	Value
Resolution <=	3.0
Molecule type	nonpeptidic
Organism	mouse (Mus musculus)
MHC Class	I

Cluster	Assay_id	Epitope_id	Antigen	Structure	Structure_quality	MHC	calc_epitope_residues
	3838883 [View 3D Structure]	766894	Epitope: Core Epitope: Antigen: 1-O-[6-O-(3-phenylpropanoyl)-alpha-D-galactopyranosyl]-N-hexacosanoylsphinganine Organism:	pdb_id: 6BNL mhc_chain1: A mhc_chain2: B epitope_chain: A	Resolution: 2.6 R Free: 0.2	Class: non classical Allele: mouse CD1d Organism: mouse (Mus musculus)	A: (QWV)405
	1802187 [View 3D Structure]	158632	Epitope: Core Epitope: Antigen: glycolipid Organism:	pdb_id: 3JVG mhc_chain1: A mhc_chain2: C epitope_chain: A	Resolution: 2.2 R Free: 0.2	Class: non classical Allele: chicken CD1-1 Organism: chicken (Gallus gallus)	A: (UNL)1, (UNL)283;
	1848554 [View 3D Structure]	153506	Epitope: Core Epitope: Antigen: 1-O-(1-O-hexadecanoyl-2-O-heptadecanoyl-sn-glycero-3-phosphono)-1D-myo-inositol Organism:	pdb_id: 3QI9 mhc_chain1: A mhc_chain2: B epitope_chain: A	Resolution: 2.3 R Free: 0.2	Class: non classical Allele: mouse CD1d Organism: mouse (Mus musculus)	A: (PII)405;
	1969015 [View 3D Structure]	139426	Epitope: Core Epitope: Antigen: 1-(3-O-sulfo-beta-D-galactosyl)sphingosine Organism:	pdb_id: 4ELM mhc_chain1: A mhc_chain2: B epitope_chain: A	Resolution: 3.5 R Free: 0.2	Class: non classical Allele: mouse CD1d Organism: mouse (Mus musculus)	A: (SGF)305, (PLM)306
	1976430 [View 3D Structure]	164047	Epitope: Core Epitope: Antigen: 1-O-(alpha-D-glucopyranosyl)-N-icosa-11,14-dienoylphytosphingosine Organism:	pdb_id: 3RUG mhc_chain1: A mhc_chain2: B epitope_chain: A	Resolution: 2.4 R Free: 0.2	Class: non classical Allele: mouse CD1d Organism: mouse (Mus musculus)	A: (DB6)303
	1971310 [View 3D Structure]	181167	Epitope: Core Epitope: Antigen: 1-elaidoyl-sn-glycero-3-phosphocholine	pdb_id: 3TZV mhc_chain1: C mhc_chain2: D	Resolution: 3.1 R Free: 0.2	Class: non classical Allele: human CD1d Organism: human	C: (LSC)303, (HEX)304, (D12)305

SCEptRe: antibody-antigen 3D complexes

http://tools.iedb.org/sceptre/bcr/

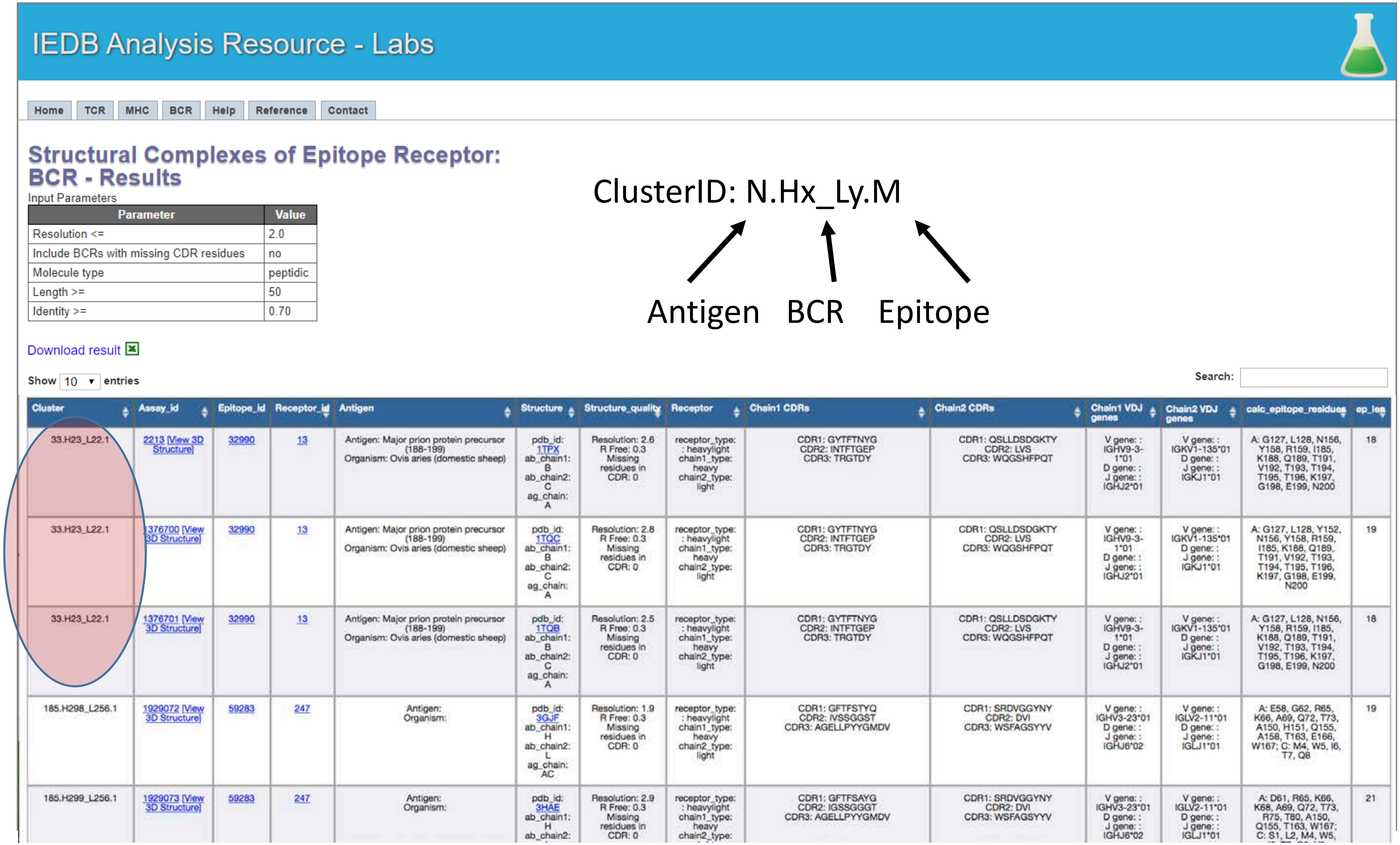


Structural Complexes of Epitope Receptor: Dataset of antibody-antigen 3D complexes

Select structural quality parameters	
Resolution <=	2.0
R free factor <=	optional
Include BCRs with missing CDR residues	No
Select antigen features	
Epitope molecule type	Peptidic
Antigen length >=	50
Antigen sequence identity >=	70%

SCEptRe: antibody-antigen 3D complexes - results

http://tools.iedb.org/sceptre/bcr/



Structural Complexes of Epitope Receptor: BCR - Results

Input Parameters

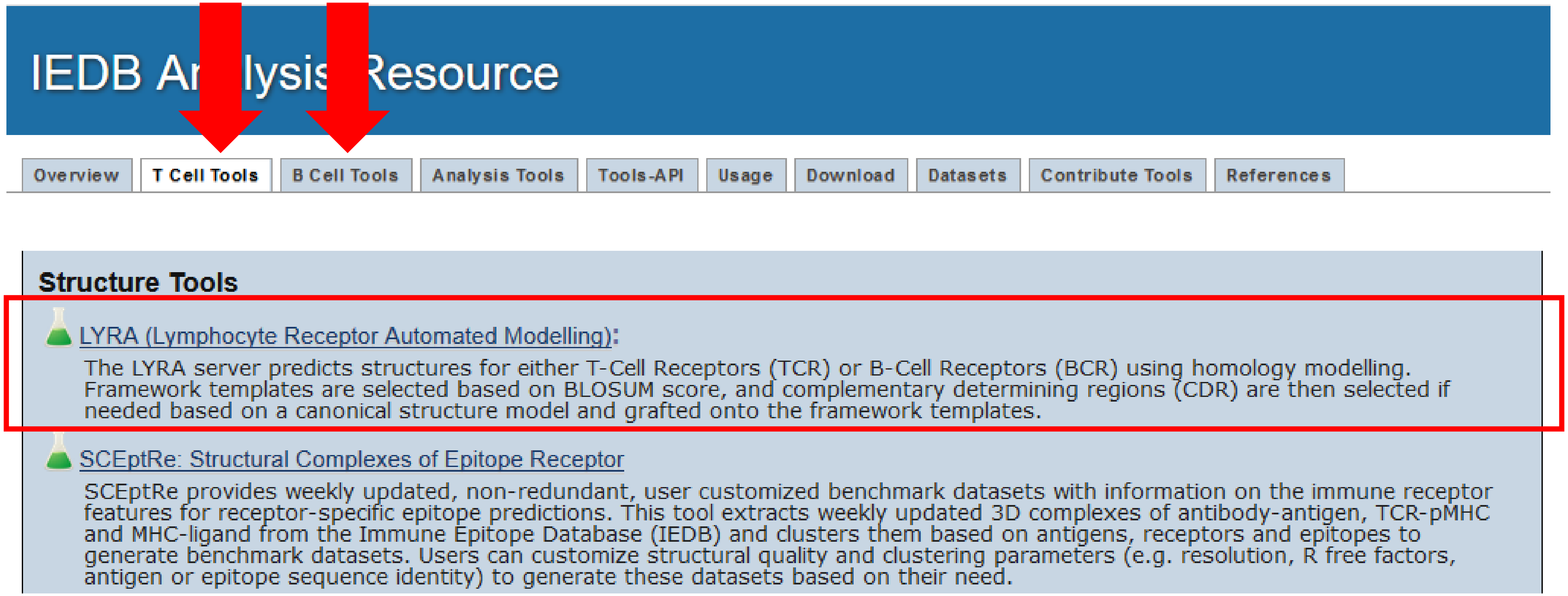
Parameter	Value
Resolution <=	2.0
Include BCRs with missing CDR residues	no
Molecule type	peptidic
Length >=	50
Identity >=	0.70

ClusterID: N.Hx_Ly.M

Antigen BCR Epitope

Cluster	Assay_id	Epitope_id	Receptor_id	Antigen	Structure	Structure_quality	Receptor	Chain1 CDRs	Chain2 CDRs	Chain1 VDJ genes	Chain2 VDJ genes	calc_epitope_residues	ep_len
33.H23_L22.1	2213 [View 3D Structure]	32990	13	Antigen: Major prion protein precursor (188-199) Organism: Ovis aries (domestic sheep)	pdb_id: 1TPX ab_chain1: B ab_chain2: C ag_chain: A	Resolution: 2.6 R Free: 0.3 Missing residues in CDR: 0	receptor_type: : heavylight chain1_type: heavy chain2_type: light	CDR1: GYTFTNYG CDR2: INTFTGEP CDR3: TRGTDY	CDR1: QSLLDSDGKTY CDR2: LVS CDR3: WQGSHFPQT	V gene: : IGHV9-3-1*01 D gene: : J gene: : IGHJ2*01	V gene: : IGKV1-135*01 D gene: : J gene: : IGKJ1*01	A: G127, L128, N156, Y158, R159, I185, K188, Q189, T191, V192, T193, T194, T195, T196, K197, G198, E199, N200	18
33.H23_L22.1	1376700 [View 3D Structure]	32990	13	Antigen: Major prion protein precursor (188-199) Organism: Ovis aries (domestic sheep)	pdb_id: 1TQC ab_chain1: B ab_chain2: C ag_chain: A	Resolution: 2.8 R Free: 0.3 Missing residues in CDR: 0	receptor_type: : heavylight chain1_type: heavy chain2_type: light	CDR1: GYTFTNYG CDR2: INTFTGEP CDR3: TRGTDY	CDR1: QSLLDSDGKTY CDR2: LVS CDR3: WQGSHFPQT	V gene: : IGHV9-3-1*01 D gene: : J gene: : IGHJ2*01	V gene: : IGKV1-135*01 D gene: : J gene: : IGKJ1*01	A: G127, L128, Y152, N156, Y158, R159, I185, K188, Q189, T191, V192, T193, T194, T195, T196, K197, G198, E199, N200	19
33.H23_L22.1	1376701 [View 3D Structure]	32990	13	Antigen: Major prion protein precursor (188-199) Organism: Ovis aries (domestic sheep)	pdb_id: 1TQB ab_chain1: B ab_chain2: C ag_chain: A	Resolution: 2.5 R Free: 0.3 Missing residues in CDR: 0	receptor_type: : heavylight chain1_type: heavy chain2_type: light	CDR1: GYTFTNYG CDR2: INTFTGEP CDR3: TRGTDY	CDR1: QSLLDSDGKTY CDR2: LVS CDR3: WQGSHFPQT	V gene: : IGHV9-3-1*01 D gene: : J gene: : IGHJ2*01	V gene: : IGKV1-135*01 D gene: : J gene: : IGKJ1*01	A: G127, L128, N156, Y158, R159, I185, K188, Q189, T191, V192, T193, T194, T195, T196, K197, G198, E199, N200	18
185.H298_L256.1	1929072 [View 3D Structure]	59283	247	Antigen: Organism:	pdb_id: 3GJF ab_chain1: H ab_chain2: L ag_chain: AC	Resolution: 1.9 R Free: 0.3 Missing residues in CDR: 0	receptor_type: : heavylight chain1_type: heavy chain2_type: light	CDR1: GFTFSTYQ CDR2: IVSSGGST CDR3: AGELLPYYGMDV	CDR1: SRDVGGYNY CDR2: DVI CDR3: WSFAGSYYV	V gene: : IGHV3-23*01 D gene: : J gene: : IGHJ6*02	V gene: : IGLV2-11*01 D gene: : J gene: : IGLJ1*01	A: E58, G62, R65, K66, A69, Q72, T73, A150, H151, Q155, A158, T163, E166, W167; C: M4, W5, I6, T7, Q8	19
185.H299_L256.1	1929073 [View 3D Structure]	59283	247	Antigen: Organism:	pdb_id: 3HAE ab_chain1: H ab_chain2:	Resolution: 2.9 R Free: 0.3 Missing residues in CDR: 0	receptor_type: : heavylight chain1_type: heavy chain2_type:	CDR1: GFTFSAYG CDR2: IGSSGGGT CDR3: AGELLPYYGMDV	CDR1: SRDVGGYNY CDR2: DVI CDR3: WSFAGSYYV	V gene: : IGHV3-23*01 D gene: : J gene: : IGHJ6*02	V gene: : IGLV2-11*01 D gene: : J gene: : IGLJ1*01	A: D61, R65, K66, K68, A69, Q72, T73, R75, T80, A150, Q155, T163, W167; C: S1, L2, M4, W5,	21

Tools accessible from both T Cell & B Cell tabs



Structure Tools

LYRA (Lymphocyte Receptor Automated Modelling):

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.

SCEptRe: Structural Complexes of Epitope Receptor

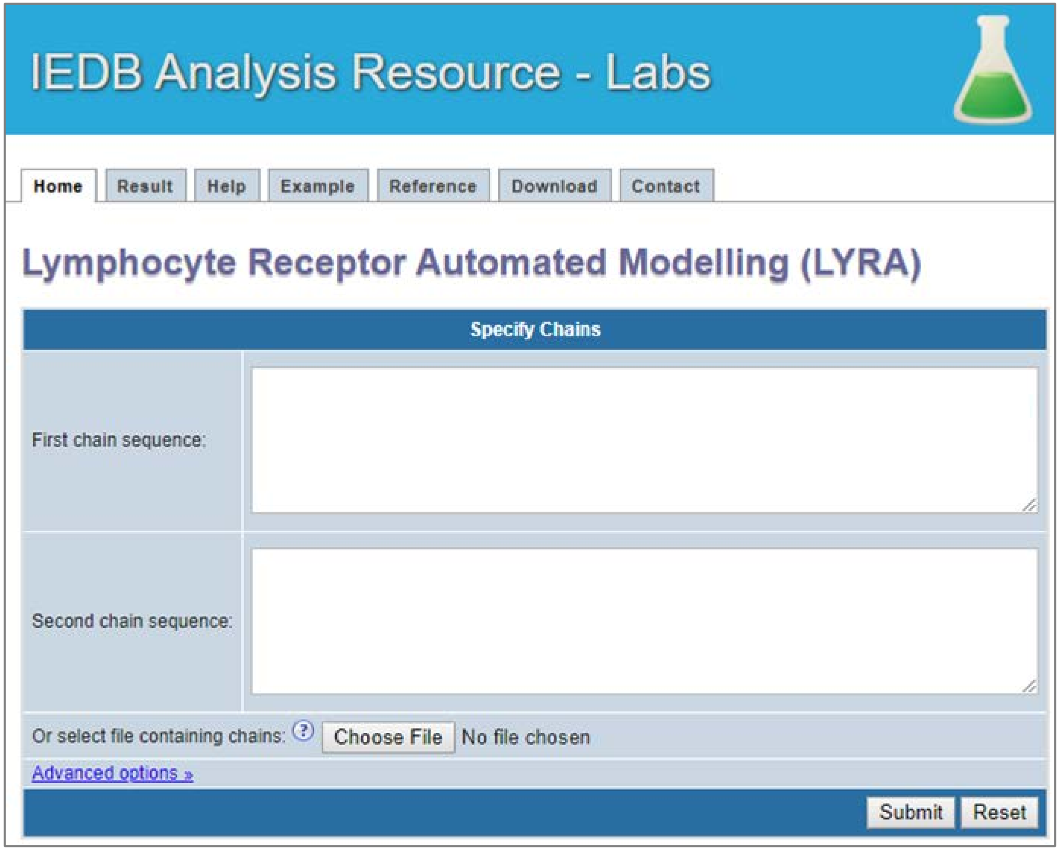
SCEptRe provides weekly updated, non-redundant, user customized benchmark datasets with information on the immune receptor features for receptor-specific epitope predictions. This tool extracts weekly updated 3D complexes of antibody-antigen, TCR-pMHC and MHC-ligand from the Immune Epitope Database (IEDB) and clusters them based on antigens, receptors and epitopes to generate benchmark datasets. Users can customize structural quality and clustering parameters (e.g. resolution, R free factors, antigen or epitope sequence identity) to generate these datasets based on their need.

Why yet another specialized tool?

- Structure provides detailed information essential to access function of a protein
- Homology modeling is still the most reliable approach for accurate protein structure prediction
  - Works accurately if similar templates exist
  - Low accuracy of variable loops
- The fold of BCR and TCR is highly conserved (similar templates exits 👍), but the function is dictated by the highly sequence variable CDR loops. 👎
- However, the CDR loops share relative low structural diversity (canonical structures)

LYRA: Lymphocyte Receptor Automated Modeling

http://tools.iedb.org/lyra/



Lymphocyte Receptor Automated Modelling (LYRA)

Specify Chains

First chain sequence:

Second chain sequence:

Or select file containing chains: ? Choose File No file chosen

Nucleic Acids Res. 2015 Jul 1;43(W1):W349-55. doi: 10.1093/nar/gkv535. Epub 2015 May 24.

LYRA, a webserver for lymphocyte receptor structural modeling.

Klausen MS[1], Anderson MV[1], Jespersen MC[1], Nielsen M[2], Marcatili P[3].

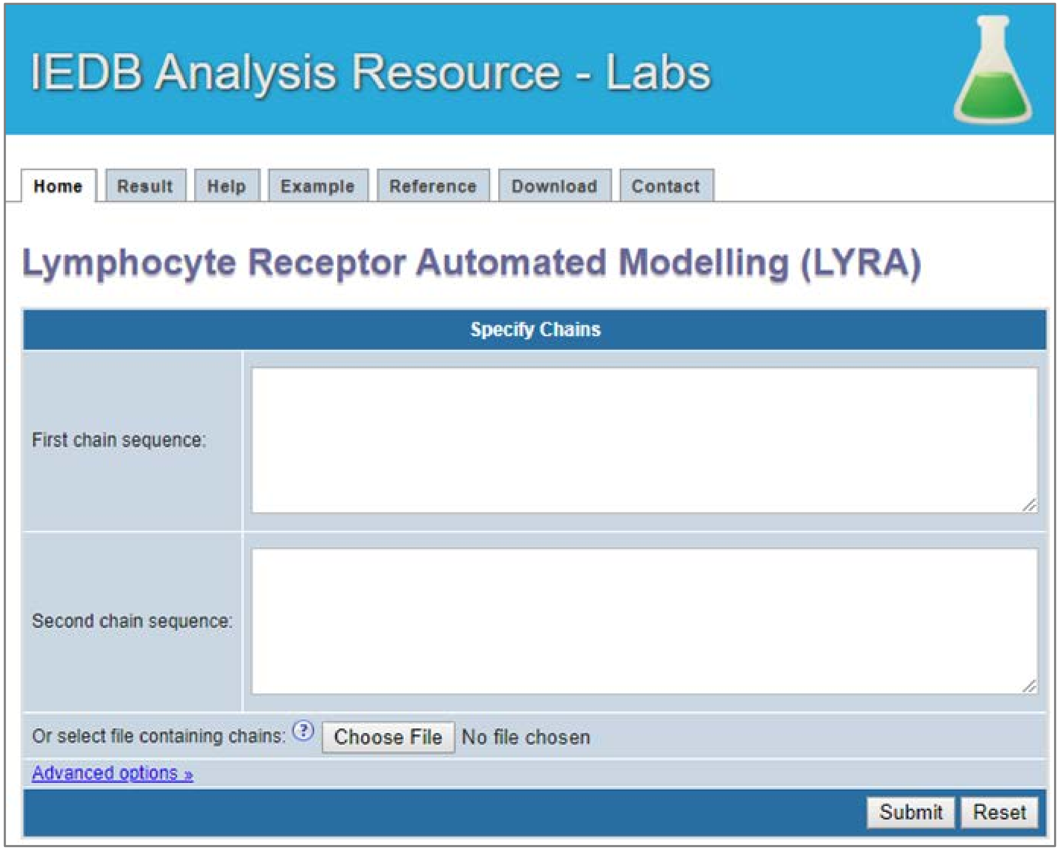
PMID: 26007650 PMCID: PMC4489227 DOI: 10.1093/nar/gkv535

- B- and T-cell receptor structure modeling

- Canonical structures (CS):
  - Hypervariable CDR loops only assume a limited number of conformations *(Chothia & Lesk, JMB 1987)*
  - Usually identified by specific sequence features

LYRA: Lymphocyte Receptor Automated Modeling

http://tools.iedb.org/lyra/



IEDB Analysis Resource - Labs

Lymphocyte Receptor Automated Modelling (LYRA)

Specify Chains

First chain sequence:

Second chain sequence:

Or select file containing chains: ? Choose File No file chosen

Advanced options »

Nucleic Acids Res. 2015 Jul 1;43(W1):W349-55. doi: 10.1093/nar/gkv535. Epub 2015 May 24.

LYRA, a webserver for lymphocyte receptor structural modeling.

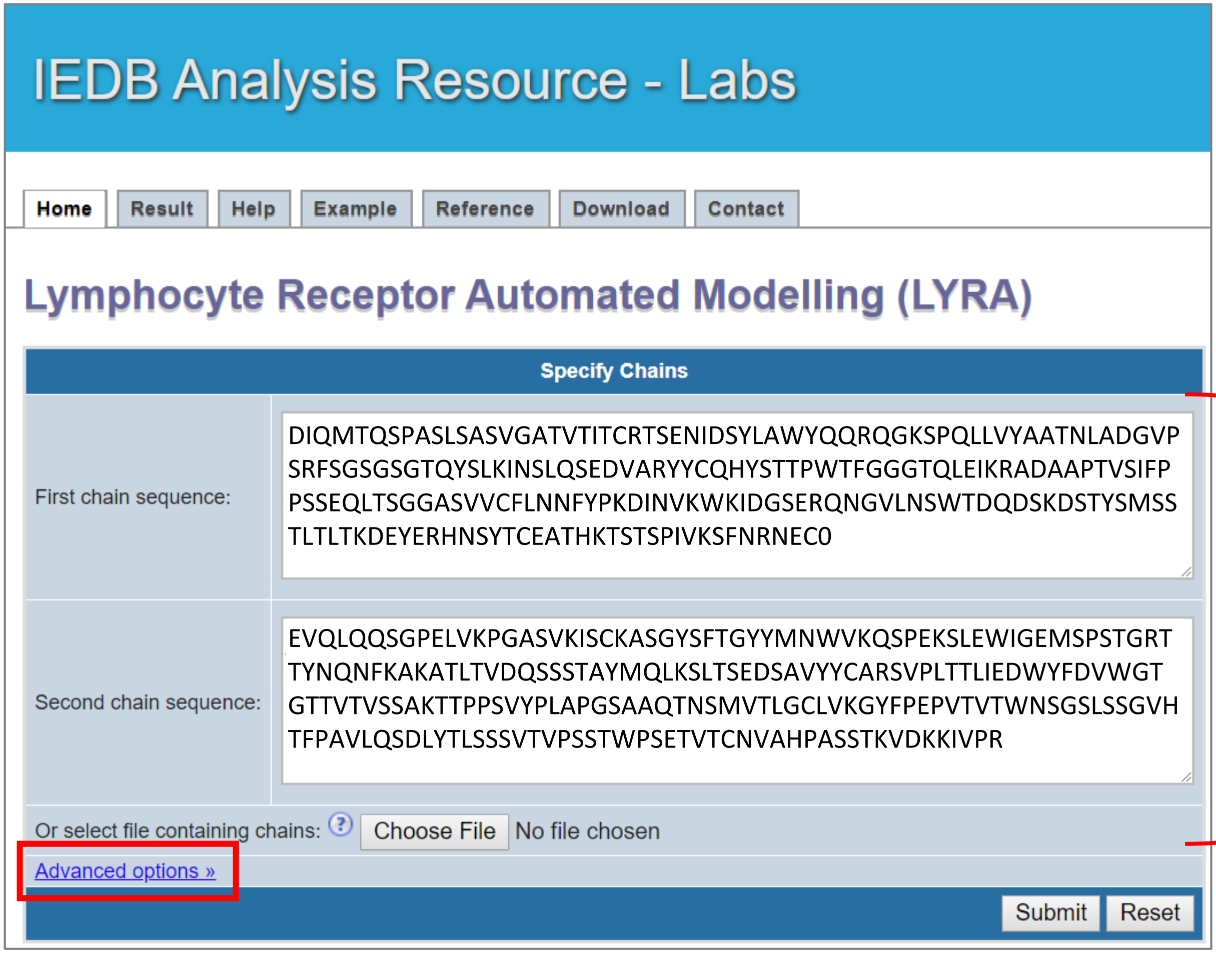
Klausen MS[1], Anderson MV[1], Jespersen MC[1], Nielsen M[2], Marcatili P[3].

PMID: 26007650 PMCID: PMC4489227 DOI: 10.1093/nar/gkv535

- Templates are identified using BLOSUM62 scores
- Loop modeling:
  If the CS of target and template loops do not match, then the highest scoring identical canonical structure loops from other structures are selected

LYRA – example (BCRs)

http://tools.iedb.org/lyra/



Lymphocyte Receptor Automated Modelling (LYRA)

First chain sequence:

DIQMTQSPASLSASVGATVTITCRTSENIDSYLAWYQQRQGKSPQLLVYAATNLADGVP
SRFSGSGSGTQYSLKINSLQSEDVARYYCQHYSTTPWTFGGGTQLEIKRADAAPTVSIFP
PSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSS
TLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC0

Second chain sequence:

EVQLQQSGPELVKPGASVKISCKASGYSFTGYYMNWVKQSPEKSLEWIGEMSPSTGRT
TYNQNFKAKATLTVDQSSSTAYMQLKSLTSEDSAVYYCARSVPLTTLIEDWYFDVWGT
GTTVTVSSAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVH
TFPAVLQSDLYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPR

Or select file containing chains: No file chosen

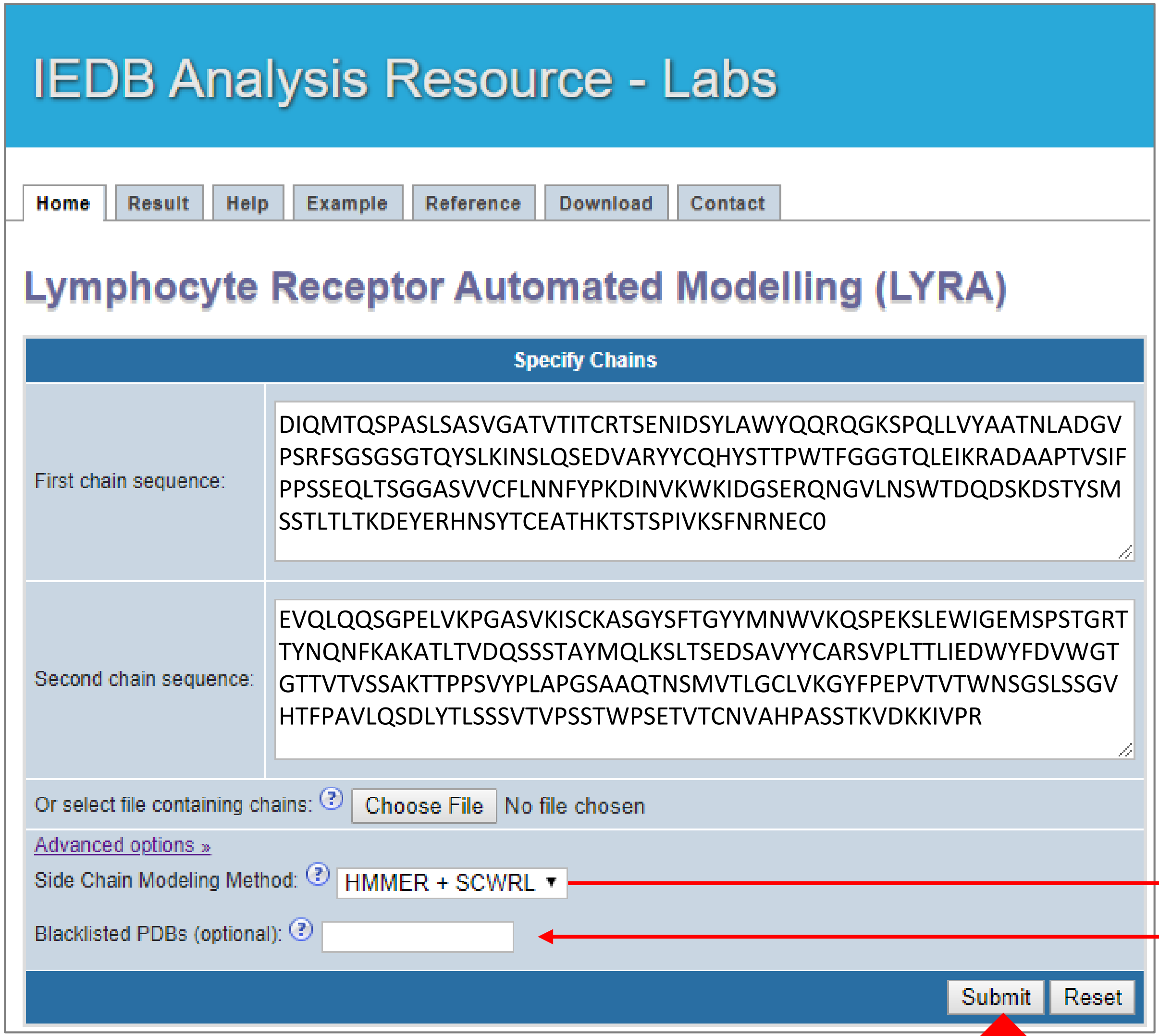
Chain sequences
(input or upload)

FASTA upload format:
>light chain header
sequence for light chain here...

>heavy chain header
sequence for heavy chain here...

LYRA – example (BCRs)

http://tools.iedb.org/lyra/



Lymphocyte Receptor Automated Modelling (LYRA)

First chain sequence:

DIQMTQSPASLSASVGATVTITCRTSENIDSYLAWYQQRQGKSPQLLVYAATNLADGV
PSRFSGSGSGTQYSLKINSLQSEDVARYYCQHYSTTPWTFGGGTQLEIKRADAAPTVSIF
PPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSM
SSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC0

Second chain sequence:

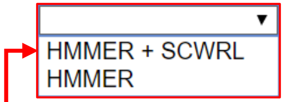
EVQLQQSGPELVKPGASVKISCKASGYSFTGYYMNWVKQSPEKSLEWIGEMSPSTGRT
TYNQNFKAKATLTVDQSSSTAYMQLKSLTSEDSAVYYCARSVPLTTLIEDWYFDVWGT
GTTVTVSSAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGV
HTFPAVLQSDLYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPR

Or select file containing chains: No file chosen

Side Chain Modeling Method: HMMER + SCWRL

Blacklisted PDBs (optional):

Side Chain Modeling Method



- HMMER – searches sequence databases for sequence homologs & makes sequence alignments
- SCWRL 4.0 models non conserved residues (uses conserved residue side chains as constraints)

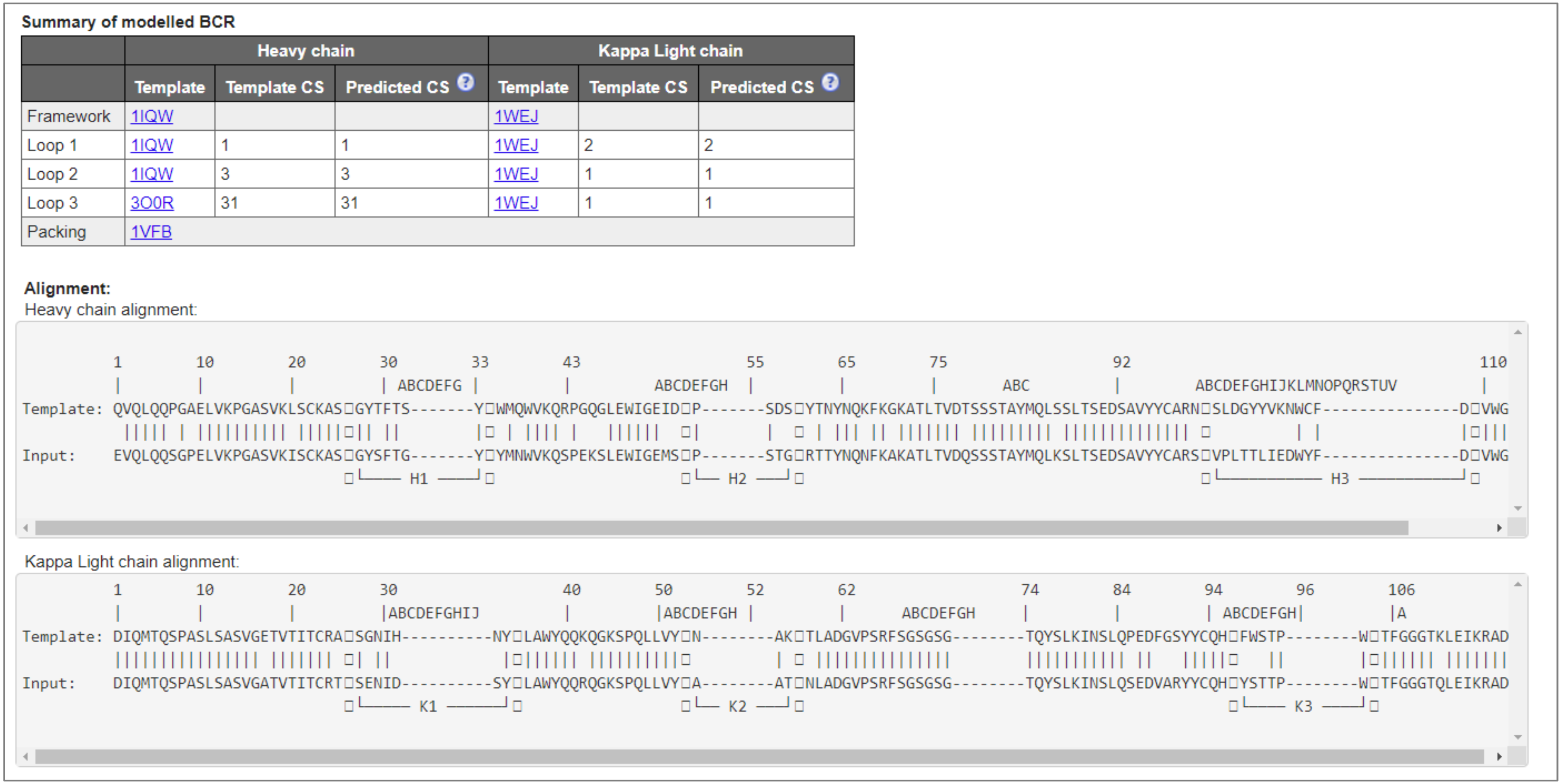
Prevents use as framework or loop template for both chains

LYRA – example (BCRs)

http://tools.iedb.org/lyra/

LYRA – example (BCRs)

http://tools.iedb.org/lyra/



Summary of modelled BCR

	Heavy chain			Kappa Light chain		
	Template	Template CS	Predicted CS ?	Template	Template CS	Predicted CS ?
Framework	1IQW			1WEJ		
Loop 1	1IQW	1	1	1WEJ	2	2
Loop 2	1IQW	3	3	1WEJ	1	1
Loop 3	3O0R	31	31	1WEJ	1	1
Packing	1VFB					

Alignment:

Heavy chain alignment:

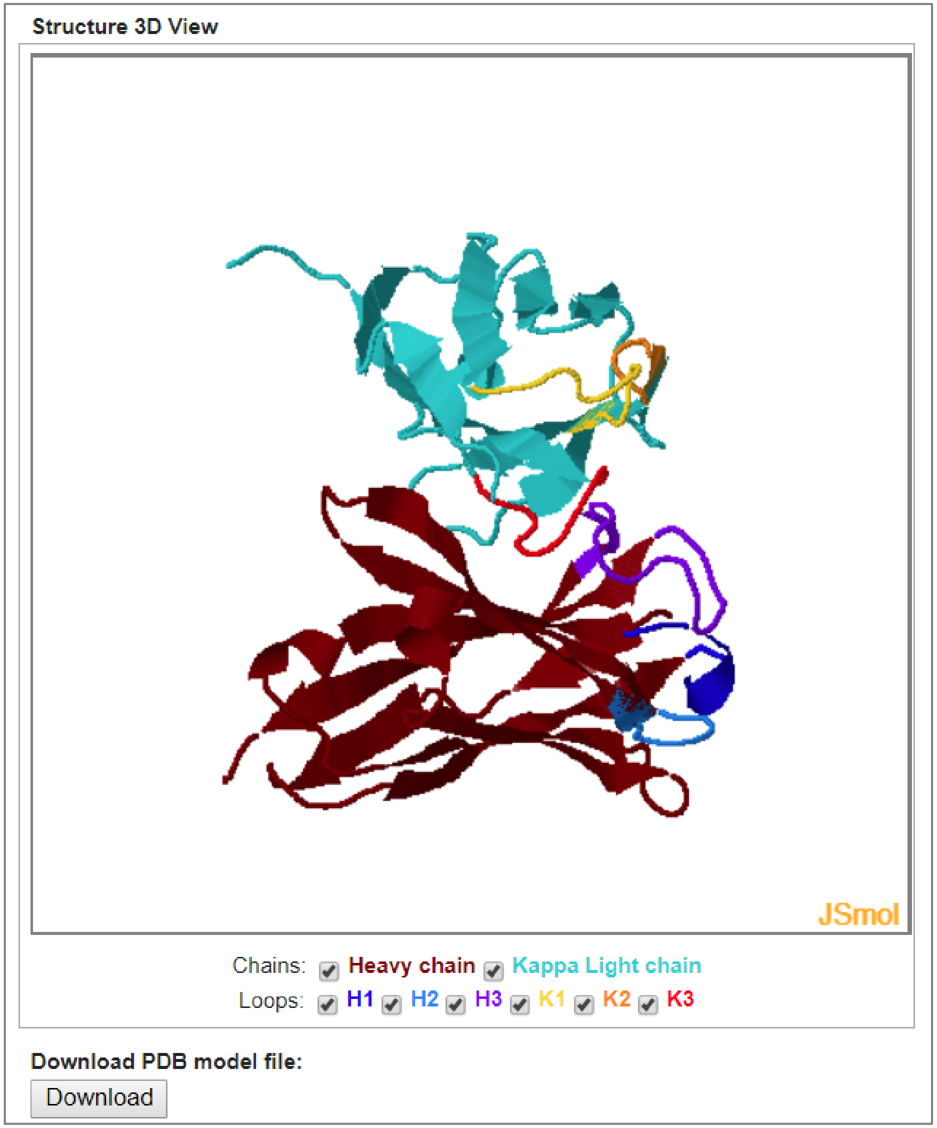
```
          1        10        20        30        33        43                  55        65        75                  92                                      110
          |        |         |         | ABCDEFG |         |         ABCDEFGH  |         |         |       ABC         |        ABCDEFGHIJKLMNOPQRSTUV         |
Template: QVQLQQPGAELVKPGASVKLSCKAS□GYTFTS-------Y□WMQWVKQRPGQGLEWIGEID□P-------SDS□YTNYNQKFKGKATLTVDTSSSTAYMQLSSLTSEDSAVYYCARN□SLDGYYVKNWCF---------------D□VWG
           ||||| | |||||||||| |||||□|| ||        |□ | |||| |   ||||||  □|       |  □ | ||| || ||||||| ||||||||| |||||||||||||| □         | |               |□|||
Input:    EVQLQQSGPELVKPGASVKISCKAS□GYSFTG-------Y□YMNWVKQSPEKSLEWIGEMS□P-------STG□RTTYNQNFKAKATLTVDQSSSTAYMQLKSLTSEDSAVYYCARS□VPLTTLIEDWYF---------------D□VWG
                                   □└──── H1 ────┘□                    □└─ H2 ────┘□                                           □└────────── H3 ────────────┘□
```

Kappa Light chain alignment:

```
          1        10        20        30                  40        50        52        62                  74        84        94        96        106
          |        |         |         |ABCDEFGHIJ         |         |ABCDEFGH |         |      ABCDEFGH     |         |         | ABCDEFGH|         |A
Template: DIQMTQSPASLSASVGETVTITCRA□SGNIH----------NY□LAWYQQKQGKSPQLLVY□N--------AK□TLADGVPSRFSGSGSG--------TQYSLKINSLQPEDFGSYYCQH□FWSTP--------W□TFGGGTKLEIKRAD
          |||||||||||||||| ||||||| □| ||            |□|||||| ||||||||||□         | □ |||||||||||||||        ||||||||||| ||   |||||□   ||        |□|||||| |||||||
Input:    DIQMTQSPASLSASVGATVTITCRT□SENID----------SY□LAWYQQRQGKSPQLLVY□A--------AT□NLADGVPSRFSGSGSG--------TQYSLKINSLQSEDVARYYCQH□YSTTP--------W□TFGGGTQLEIKRAD
                                   □└──── K1 ───────┘□                 □└─ K2 ────┘□                                              □└─── K3 ─────┘□
```

LYRA – example (BCRs)

http://tools.iedb.org/lyra/



Structure 3D View

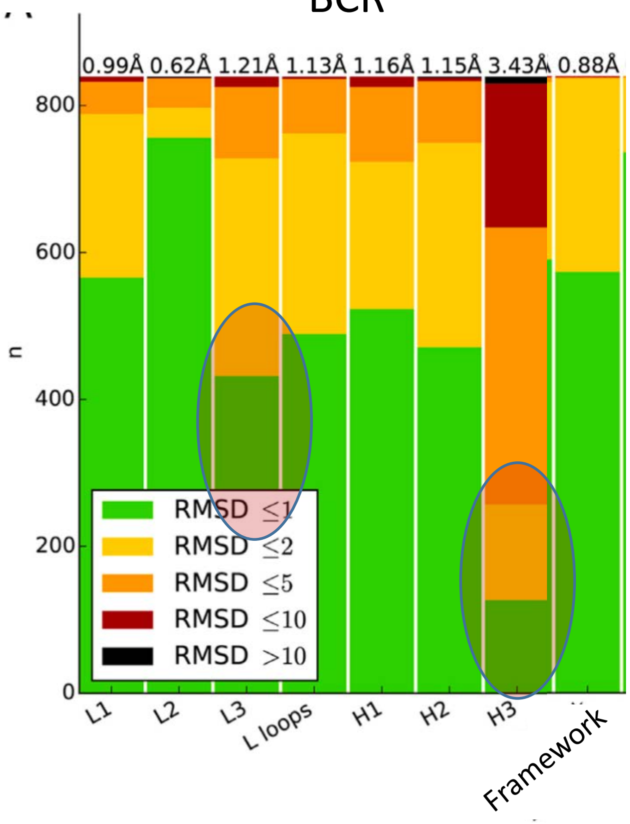
Chains: ☑ Heavy chain ☑ Kappa Light chain

Loops: ☑ H1 ☑ H2 ☑ H3 ☑ K1 ☑ K2 ☑ K3

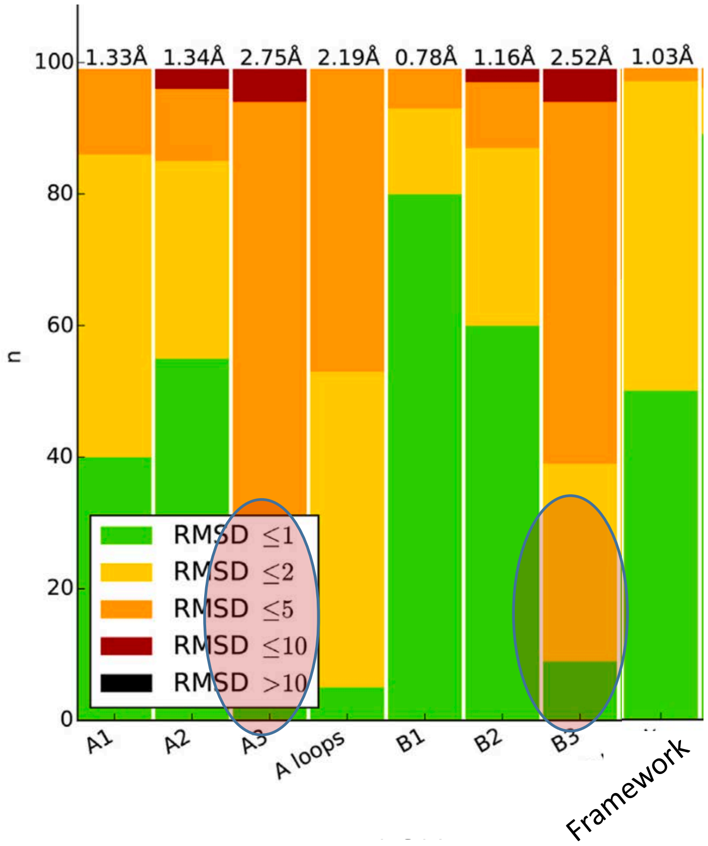
Download PDB model file:

Benchmark evaluation

BCR

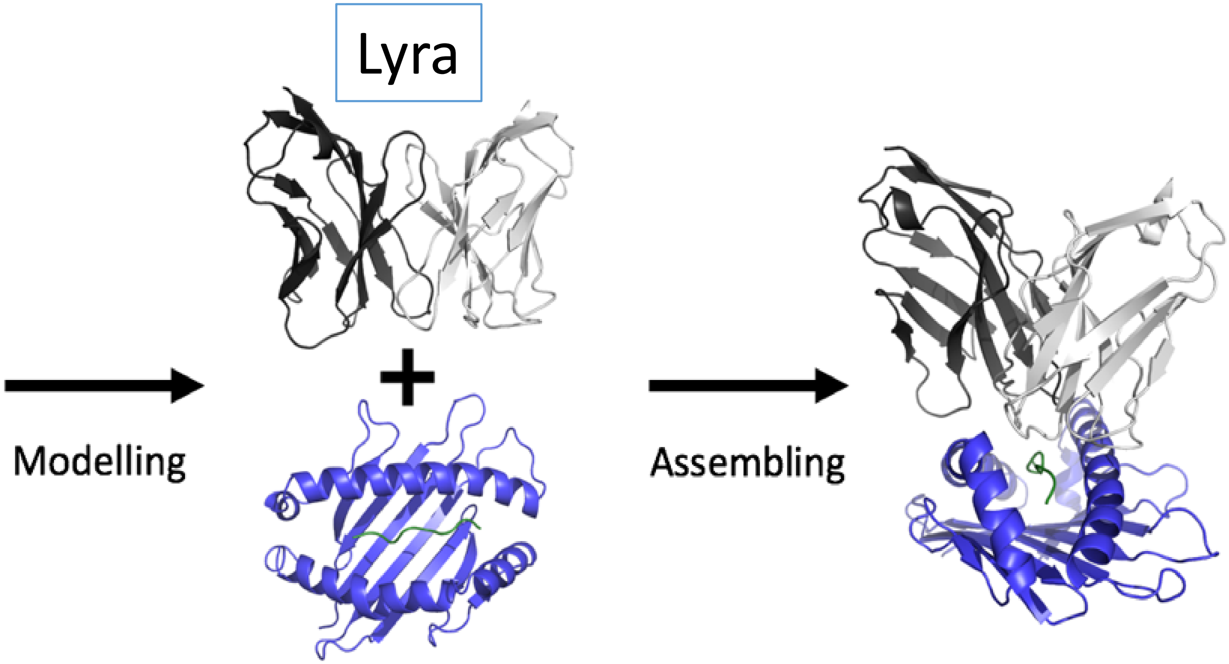


TCR



Application: Modeling TCR – peptide:MHC interactions

```
>MHC
GPHSLRYFVTAVSRPGLGEPRYMEVGYVDDT
EFVRFDSDAENPRYEPRARWMEQEGPEYWER
ETQKAKGNEQSFRVDLRTLLGYYNQSKGGSH
TIQVISGCEVGSDGRLLRGYQQYAYDGCDYI
ALNEDLKTWTAADMAALITKHKWEQAGEAER
LRAYLEGTCVEWLRRYL
>Peptide
SIYRYYGL
>TCR_alpha
QSVTQPDARVTVSEGASLQLRCKYSYSATPY
LFWYVQYPRQGLQLLLKYYSGDPVVQGVNGF
EAEFSKSNSSFHLRKASVHWSDSAVYFCAVS
GFASALTFGSGTKVIVLP
>TCR_beta
VTQSPRNKVAVTGGKVTLSCNQTNNHNNMYW
YRQDTGHGLRLIHYSYGAGSTEKGDIPDGYK
ASRPSQENFSLILELATPSQTSVYFCASGGG
GTLYFGAGTRLSVL
```



Lyra

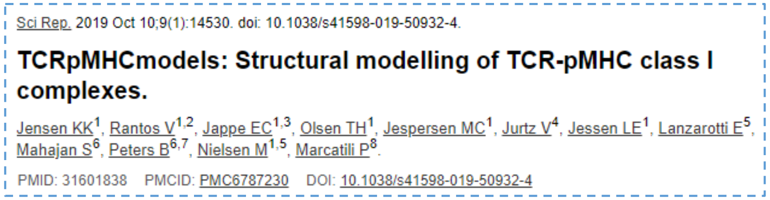
Modelling

Assembling

Input sequence

TCR and pMHC models

Final model

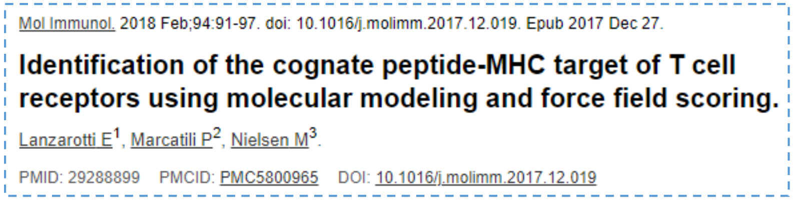


Sci Rep. 2019 Oct 10;9(1):14530. doi: 10.1038/s41598-019-50932-4.

TCRpMHCmodels: Structural modelling of TCR-pMHC class I complexes.

Jensen KK[1], Rantos V[1,2], Jappe EC[1,3], Olsen TH[1], Jespersen MC[1], Jurtz V[4], Jessen LE[1], Lanzarotti E[5], Mahajan S[6], Peters B[6,7], Nielsen M[1,5], Marcatili P[8].

PMID: 31601838 PMCID: PMC6787230 DOI: 10.1038/s41598-019-50932-4



Mol Immunol. 2018 Feb;94:91-97. doi: 10.1016/j.molimm.2017.12.019. Epub 2017 Dec 27.

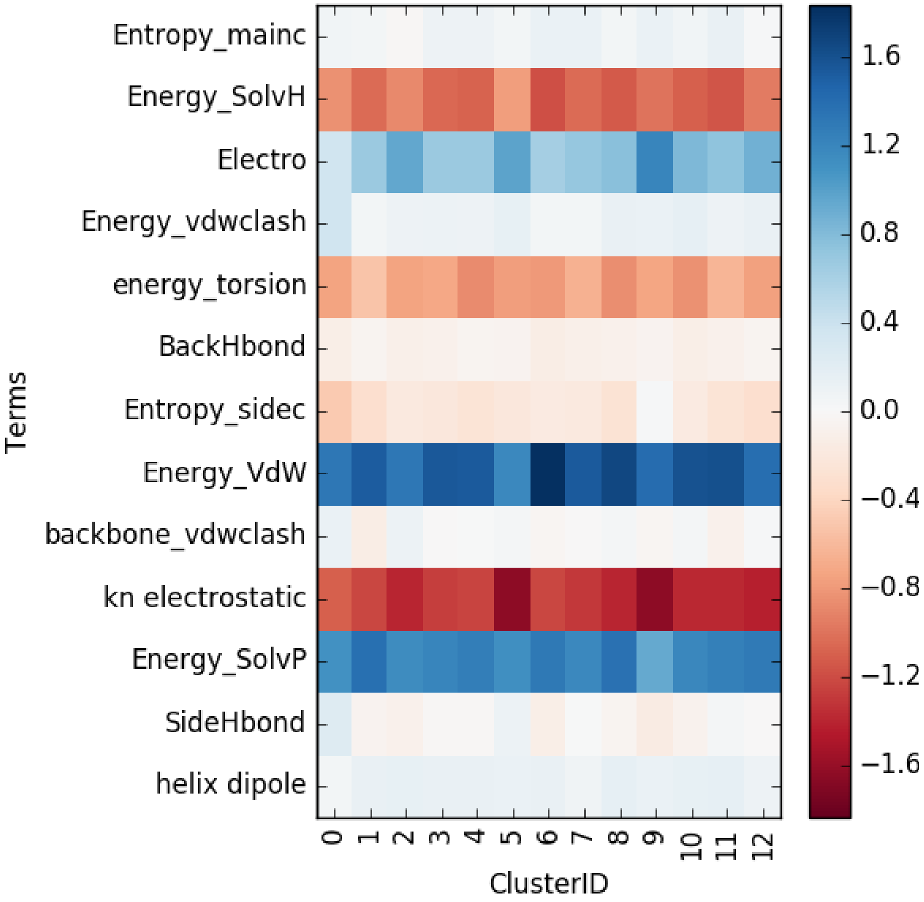
Identification of the cognate peptide-MHC target of T cell receptors using molecular modeling and force field scoring.

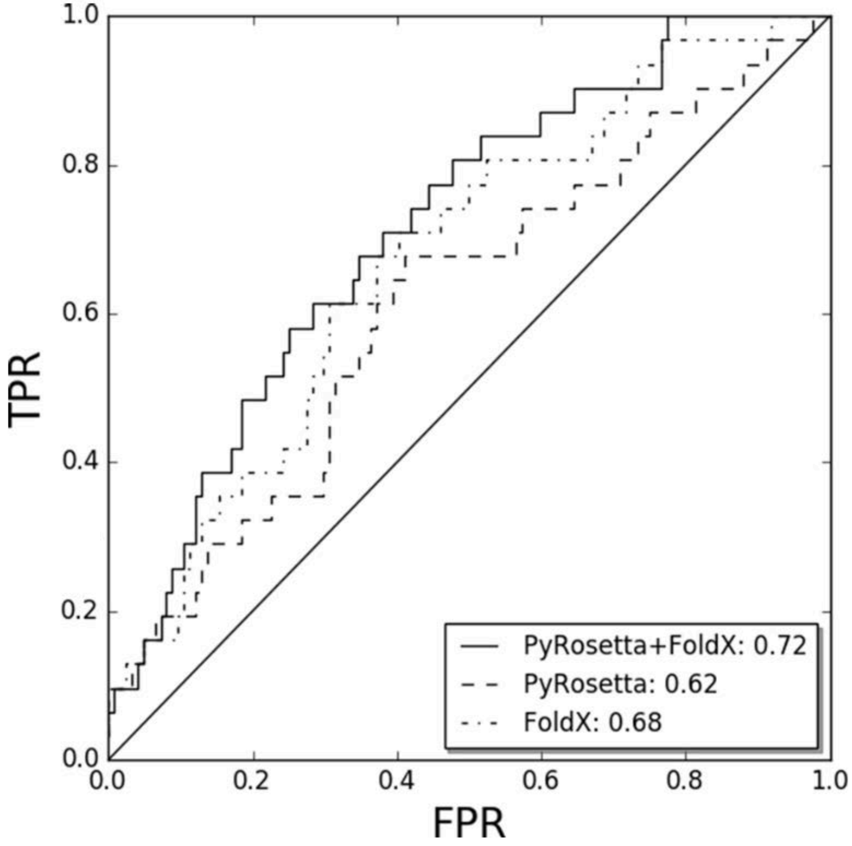
Lanzarotti E[1], Marcatili P[2], Nielsen M[3].

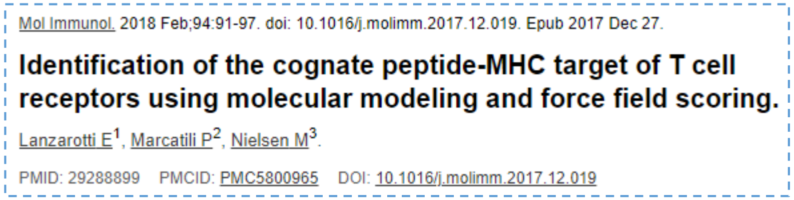
PMID: 29288899 PMCID: PMC5800965 DOI: 10.1016/j.molimm.2017.12.019

Modeling TCR – peptide:MHC interactions

Optimized force field terms







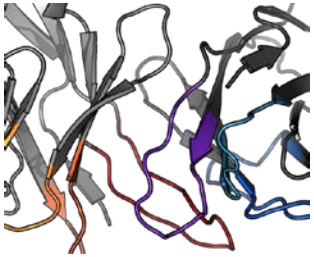
Mol Immunol. 2018 Feb;94:91-97. doi: 10.1016/j.molimm.2017.12.019. Epub 2017 Dec 27.

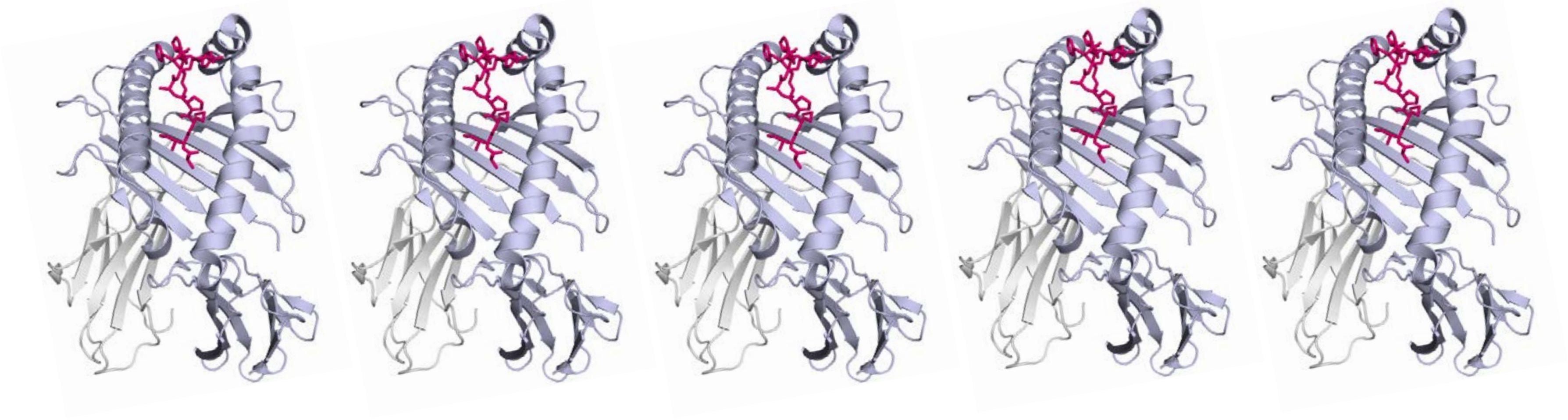
Identification of the cognate peptide-MHC target of T cell receptors using molecular modeling and force field scoring.

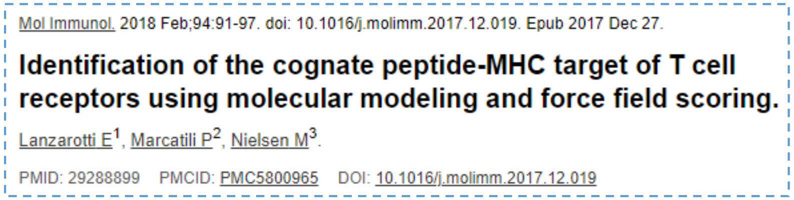
Lanzarotti E[1], Marcatili P[2], Nielsen M[3].

PMID: 29288899 PMCID: PMC5800965 DOI: 10.1016/j.molimm.2017.12.019

Modeling TCR – peptide:MHC interactions







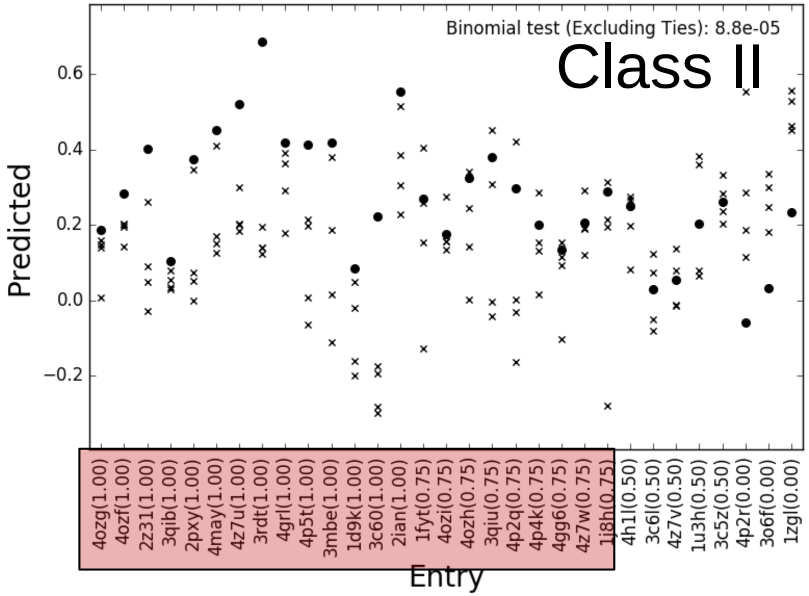
Mol Immunol. 2018 Feb;94:91-97. doi: 10.1016/j.molimm.2017.12.019. Epub 2017 Dec 27.

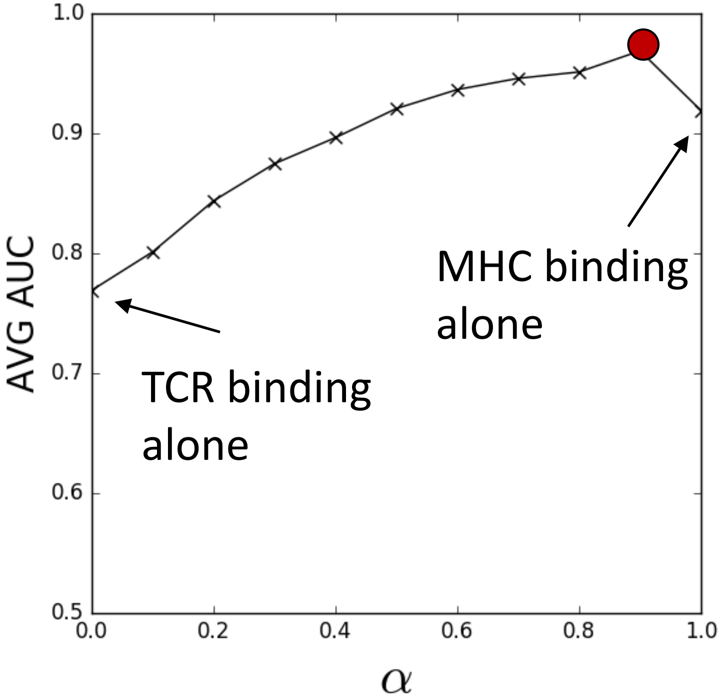
Identification of the cognate peptide-MHC target of T cell receptors using molecular modeling and force field scoring.

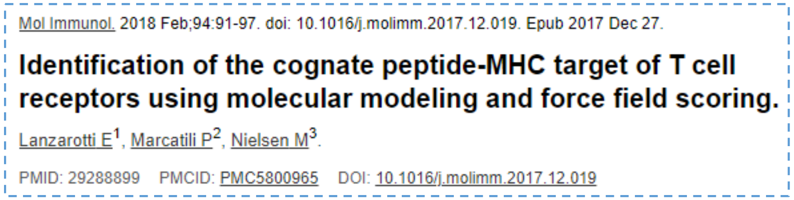
Lanzarotti E[1], Marcatili P[2], Nielsen M[3].

PMID: 29288899 PMCID: PMC5800965 DOI: 10.1016/j.molimm.2017.12.019

Modeling TCR – peptide:MHC interactions







Mol Immunol. 2018 Feb;94:91-97. doi: 10.1016/j.molimm.2017.12.019. Epub 2017 Dec 27.

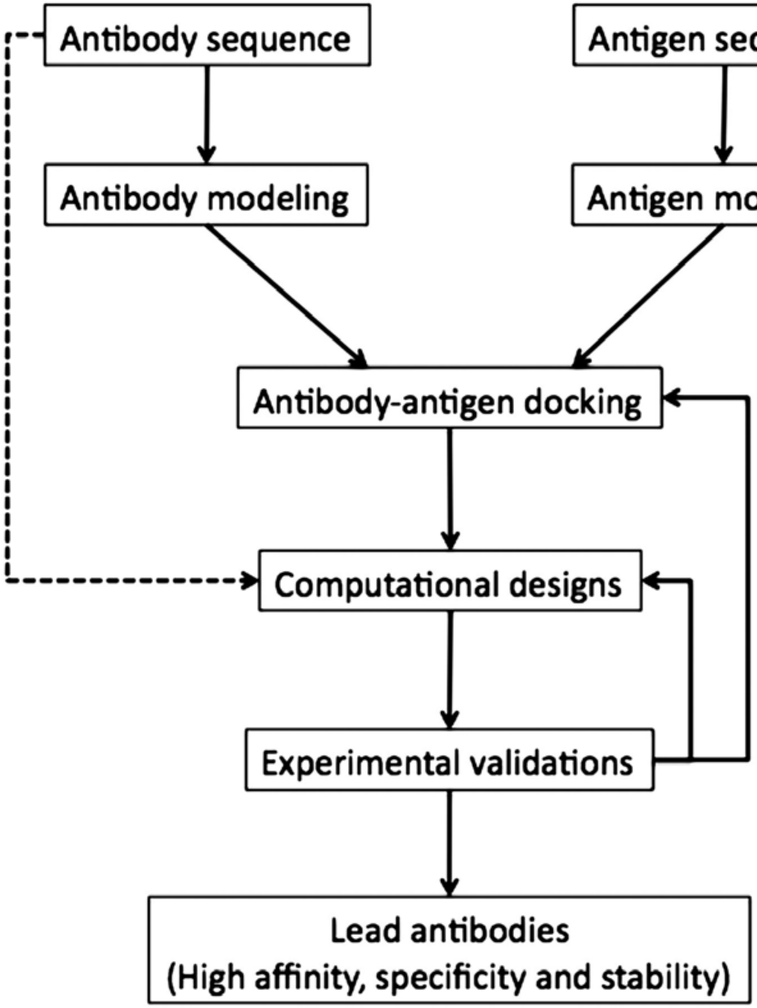
Identification of the cognate peptide-MHC target of T cell receptors using molecular modeling and force field scoring.

Lanzarotti E[1], Marcatili P[2], Nielsen M[3].

PMID: 29288899 PMCID: PMC5800965 DOI: 10.1016/j.molimm.2017.12.019

Application of LYRA in Computational Epitope discovery

LYRA
RosettaAntibody



Antibody sequence

Antigen sequence

Antibody modeling

Antigen modeling

Antibody-antigen docking

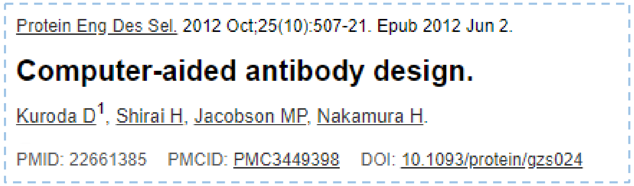
Computational designs

Experimental validations

Lead antibodies
(High affinity, specificity and stability)

PDB
PMP
I-TASSER

ClusPro
SnugDock
PocketMatch



Protein Eng Des Sel. 2012 Oct;25(10):507-21. Epub 2012 Jun 2.

Computer-aided antibody design.

Kuroda D[1], Shirai H, Jacobson MP, Nakamura H.

PMID: 22661385 PMCID: PMC3449398 DOI: 10.1093/protein/gzs024

Summary

- SCEptRE
  - Provides easy access to all BCR and TCR structures available from the IEDB
  - The tools allows customized clustering of the data based on CDR and epitope information
  - Next, step will be to develop a similar tool for BCR/TCR sequence data
- LYRA
  - A tool for accurate modeling of BCR and TCR structures
  - CDR loops are modeled based on canonical structures
  - Accuracy high with the exception of CDR3b for BCR and CBR3a and b for TCR
  - For these loops, further development of *ab-initio* modeling is needed