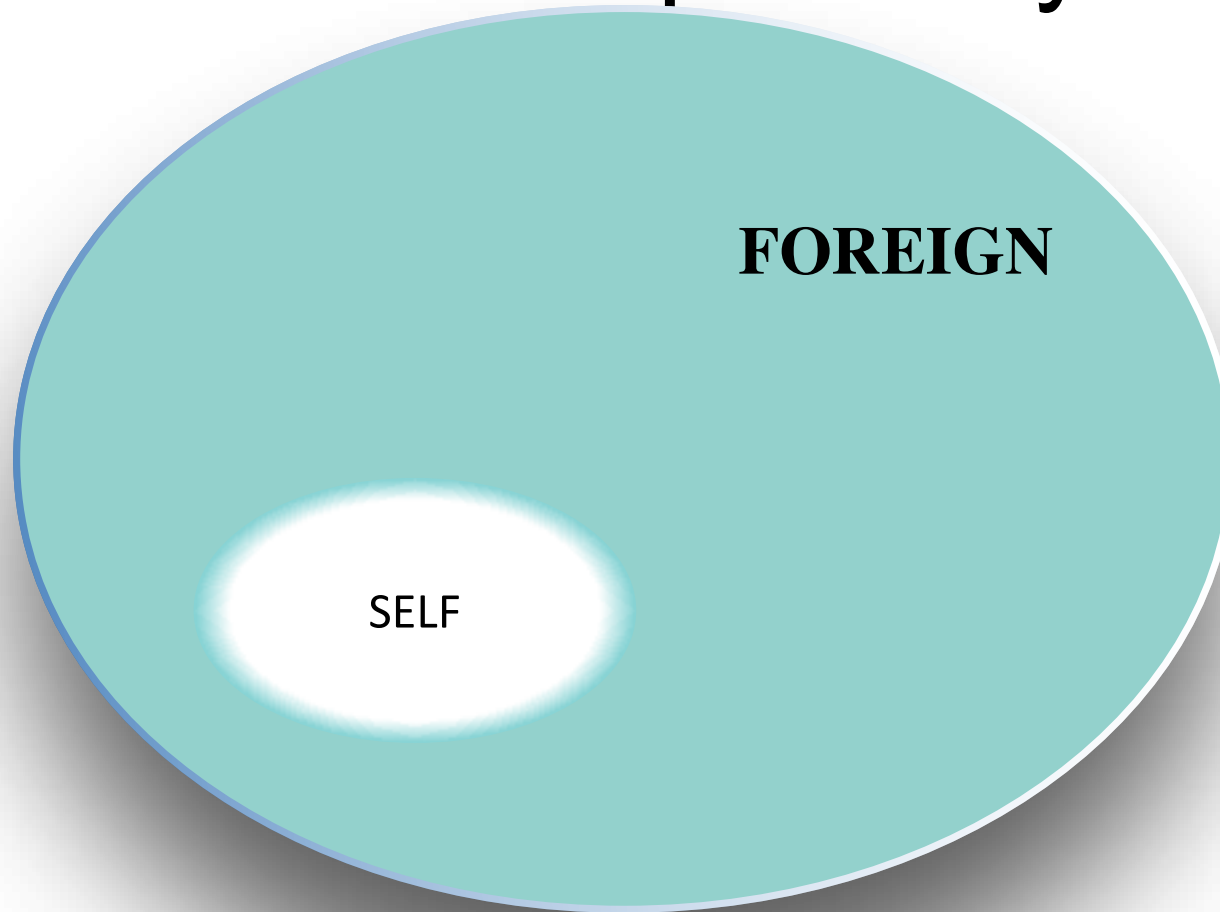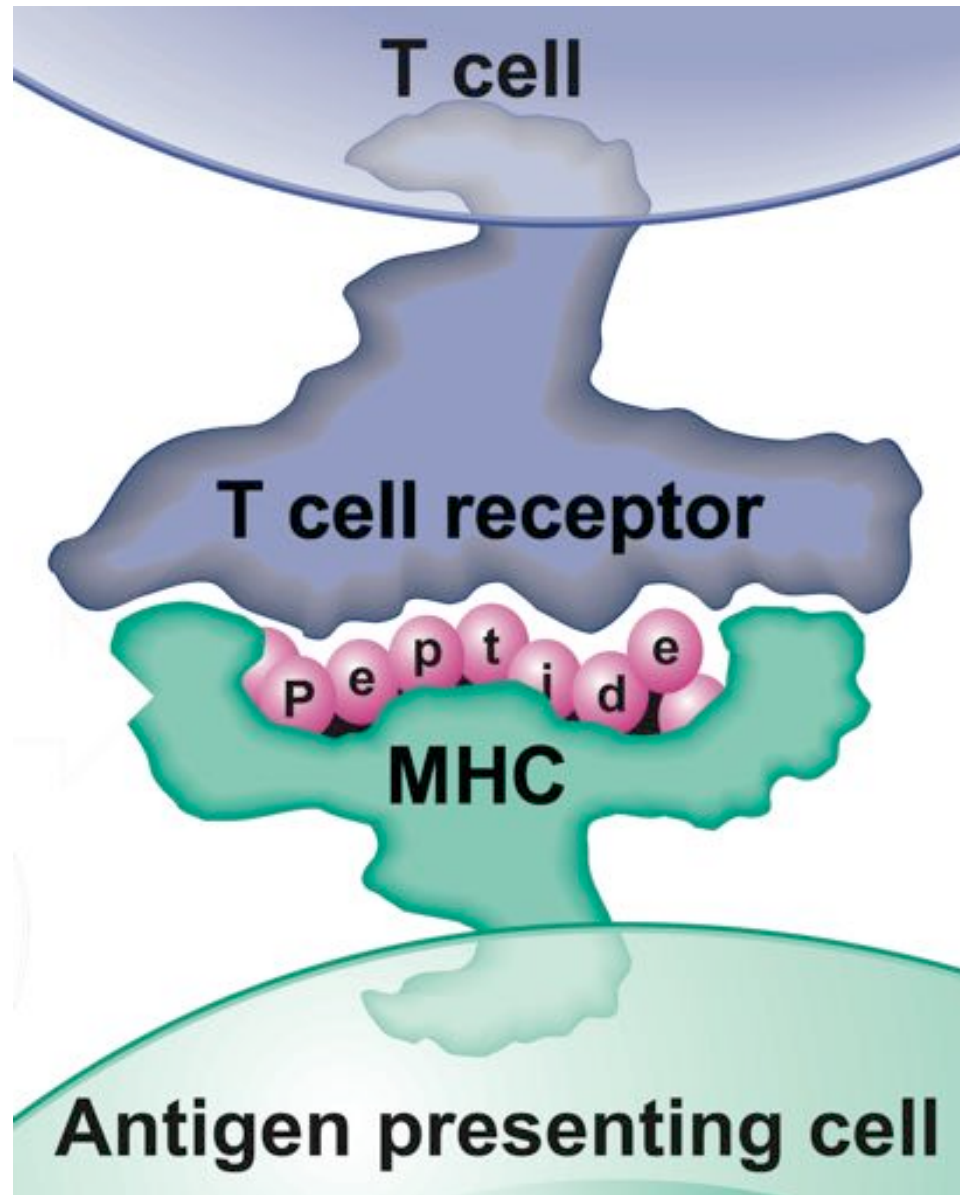# The impact of Immunological Bioinformatics on immunological research

Soren Buus
University of Copenhagen

Controlling lymphocytes is controlling immune specificity
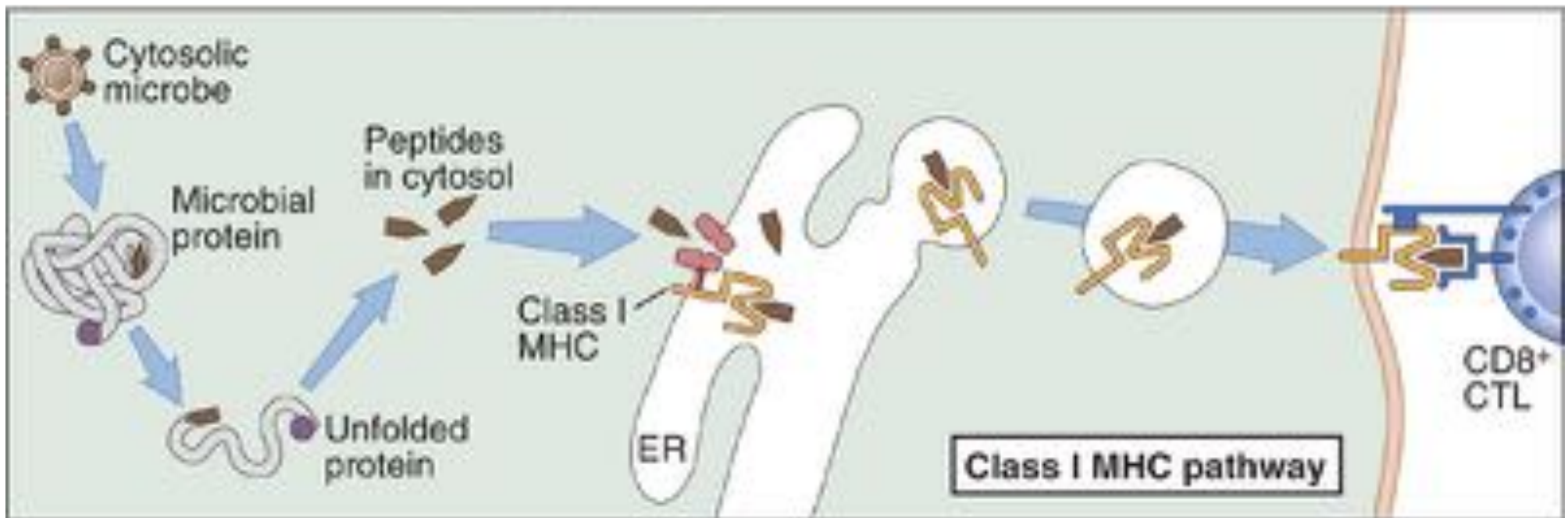

FOREIGN

SELF

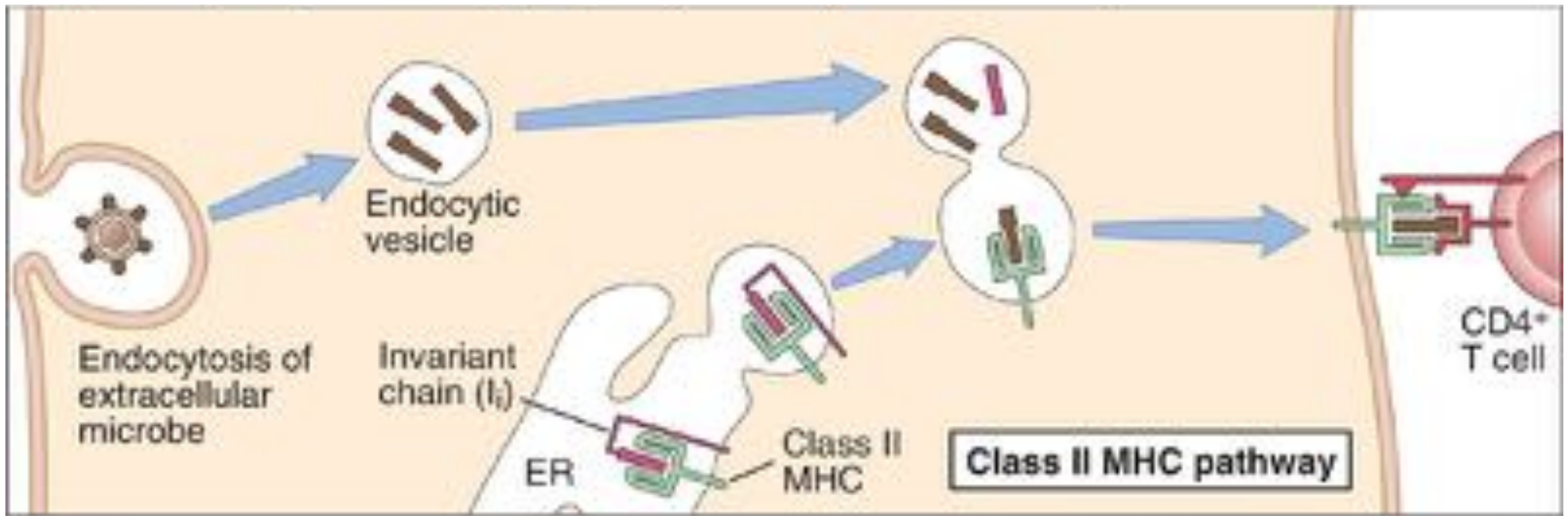# T cells recognize peptide presented by MHC / HLA

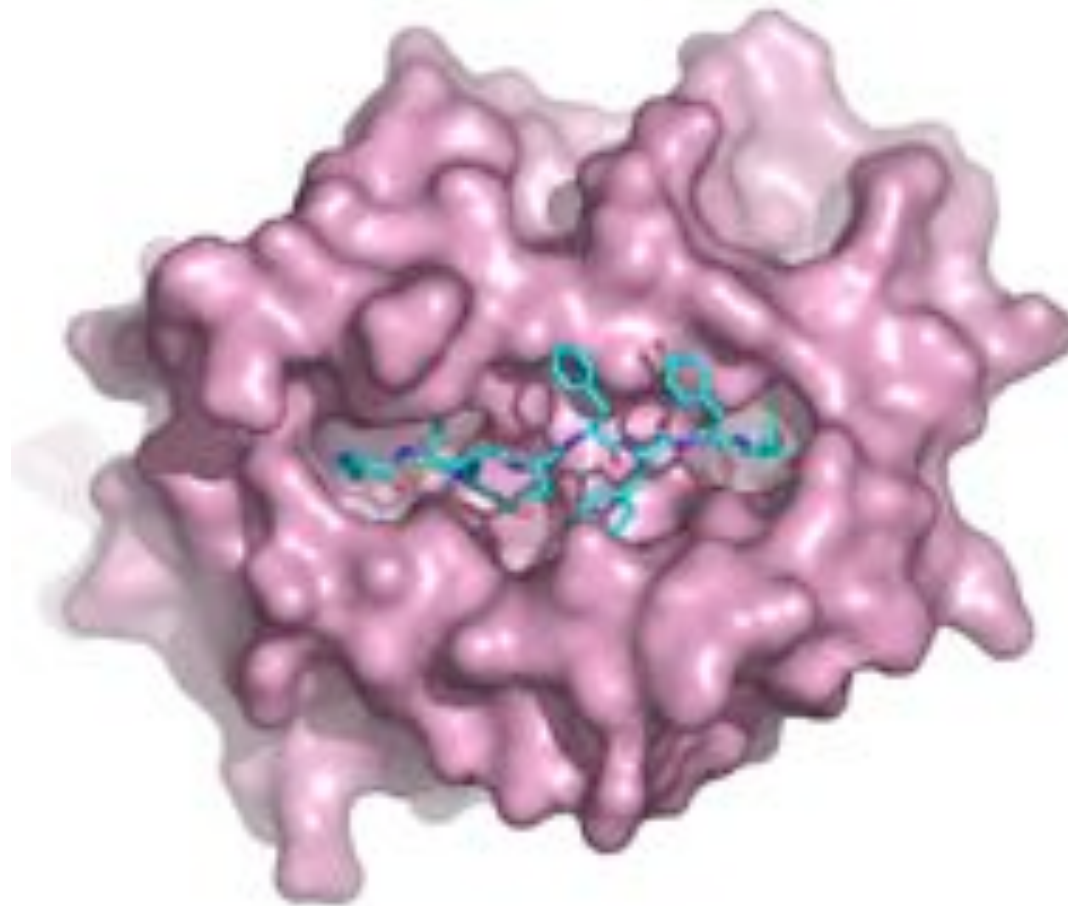# T cell recognition of proteins of intracellular origin
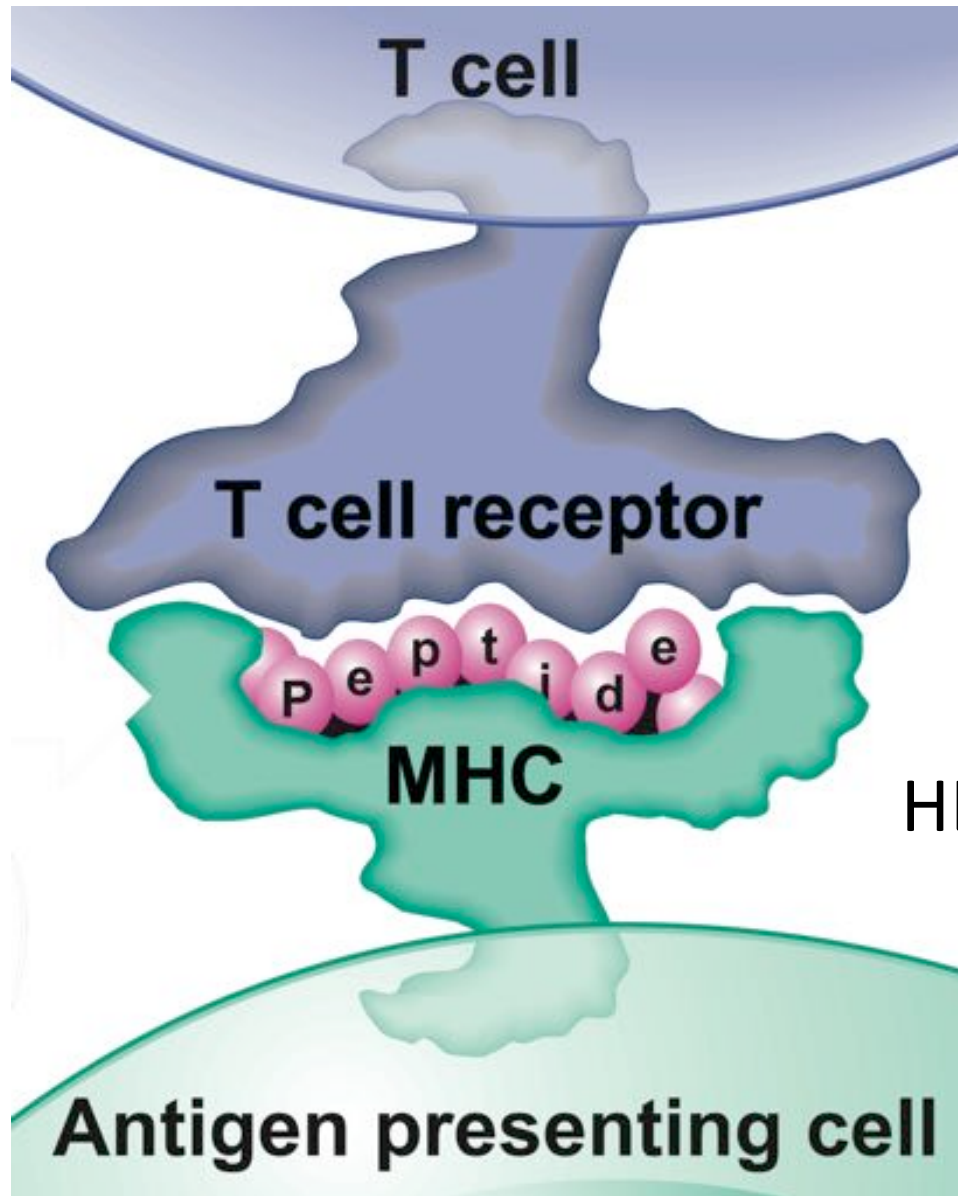
MHC class I presents cytosolic peptides to CD8+ T cells

T cell recognition of proteins of extracellular origin

MHC class II presents endosomal peptides to CD4+ T cells

Immunological Bioinformatics @ Center for Biological Sequence Analysis

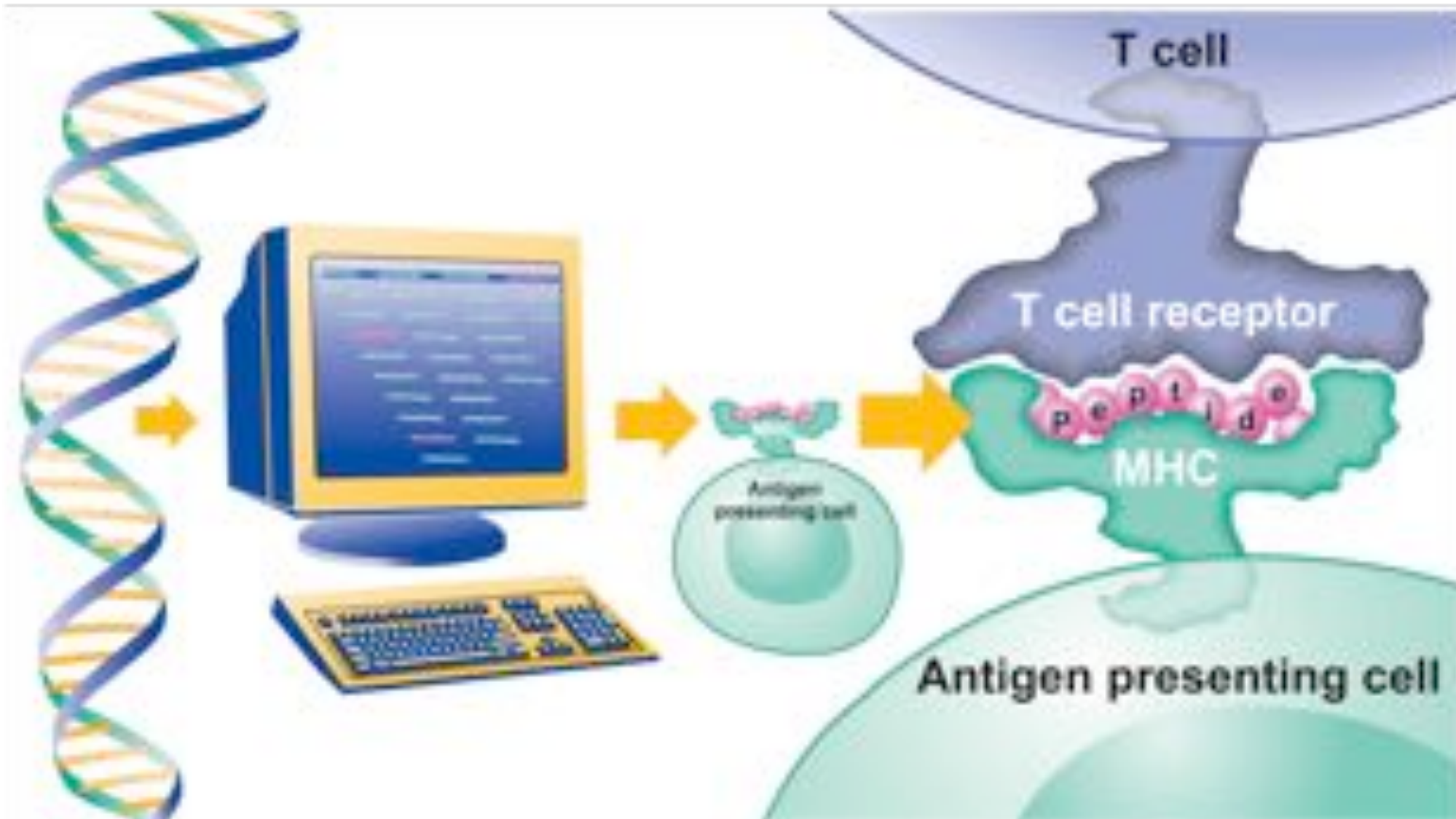# The number of possible interactions is staggering

TcR $> 10^9$

Peptide $> 10^{12}$

HLA I and II each > 7000

quence
Analysis

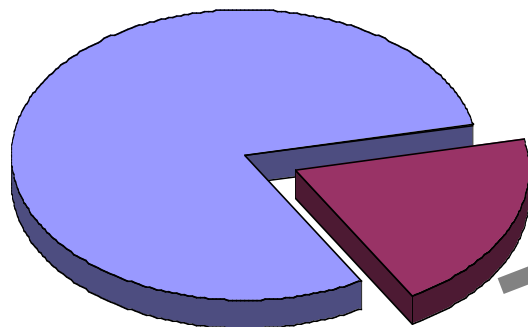# A bioinformatics approach is needed

Relative contribution of processing events

T cells responses depend on
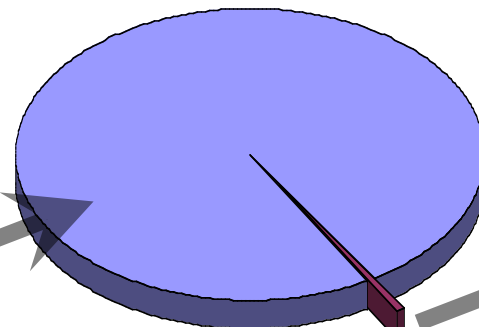
- Antigen processing i.e. peptide generation
- Peptide selection i.e. binding to MHC
- Antigen presentation i.e. peptide display

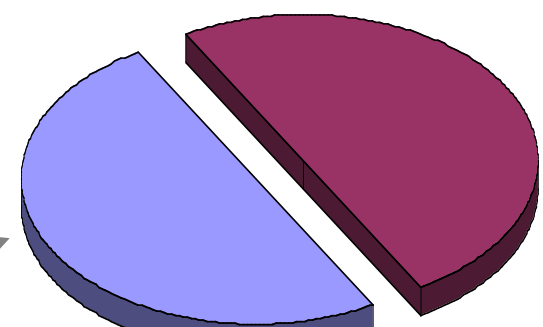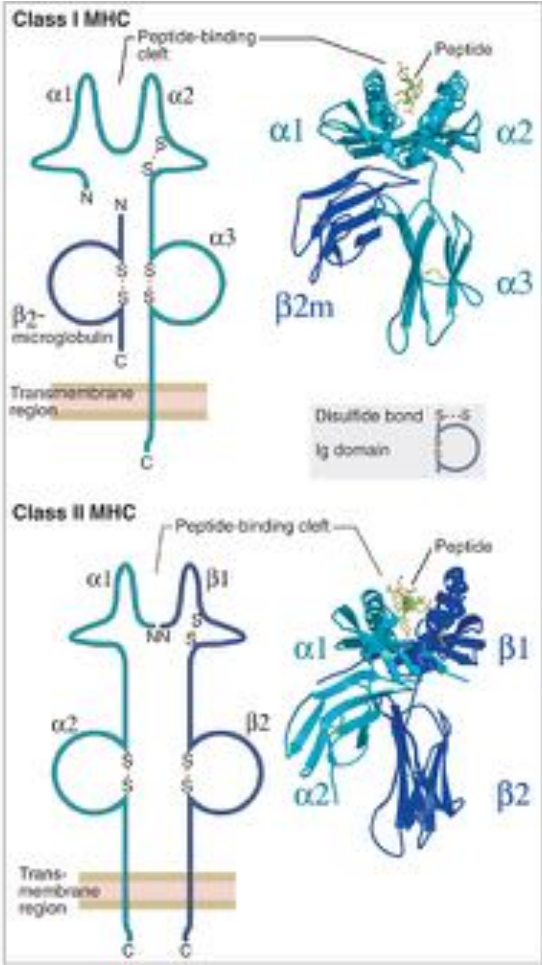Antigen processing   MHC binding   T cell repertoire

1 / 5   1 / 200   1 / 2

Structure of MHC / HLA molecules

Class I

Class II

Immunological Bioinformatics @ Center for Biological Sequence
Analysis

Binding of Peptides to MHC

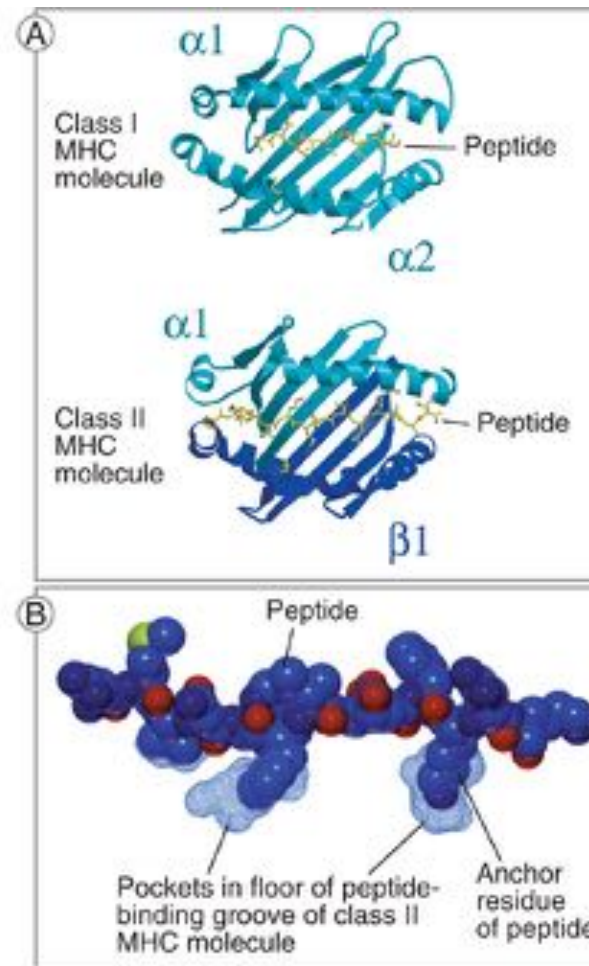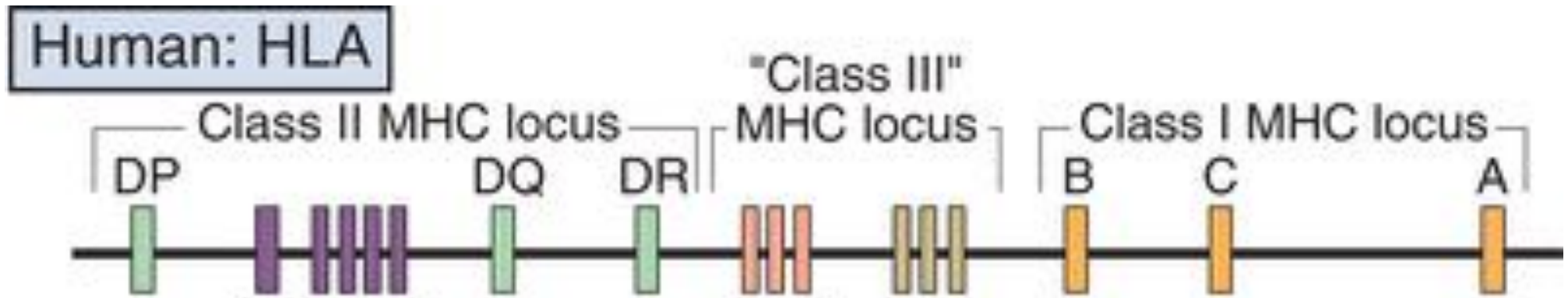MHC class I closed
Peptide short

MHC class II open
Peptide longer

Abbas & Lichtman: Basic Immunology 3e, Updated Edition.
Copyright © 2010 by Saunders, an imprint of Elsevier, Inc. All rights reserved.

MHC gene region

MHC / HLA polymorphism

- The most polymorphic gene region known
  - \> 7000 different HLA class I registered
  - \> 7000 different HLA class II registered

Gene complexity at the MHC locus in man

Class I			Class II			
gene	alleles		gene	A alleles	B alleles	A x B
HLA-A	1,519		DR	3	966	2,898
HLA-B	2,069		DQ	35	144	5,040
HLA-C	1,016		DP	28	145	4,060
HLA-E	10		DM	4	7	28
HLA-F	22		DO	12	9	108
HLA-G	46					

data from the European Bioinformatics Institute (EBI) server (http://www.ebi.ac.uk/imgt/hla/stats.html)

MHC polymorphism individualizes immune specificity
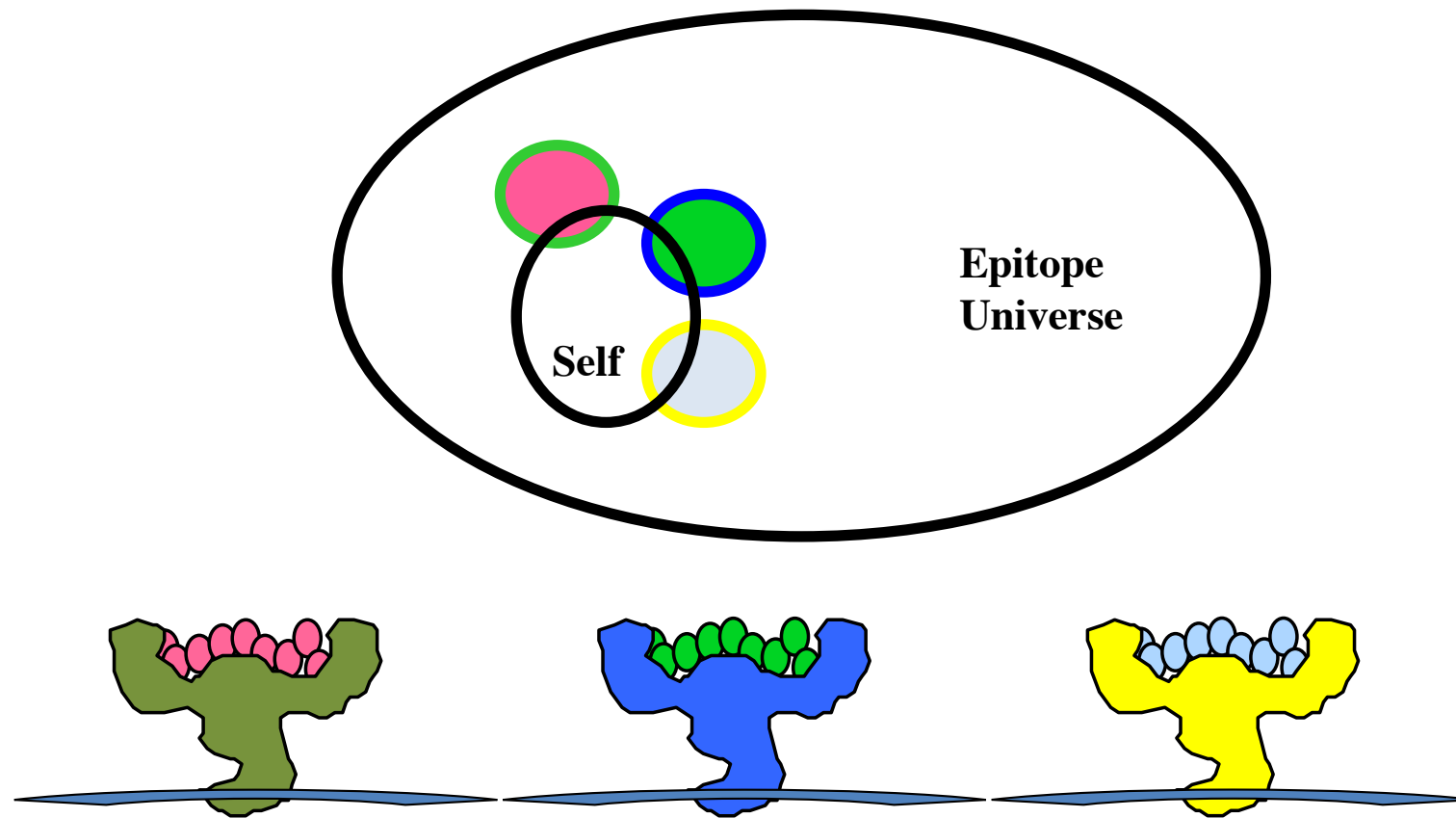
"The human MHC project"

> **Description and prediction of peptide–MHC binding: the 'human MHC project'**
> Søren Buus
>
> Current Opinion in Immunology 1999, 11:209–213

Complete mapping of all human MHC specificities

- Generate MHC molecules, peptides, assays and data

- Develop and validate binding predictors

- Disseminate and apply predictors

Strategy

- HLA molecules

HLA-A	Phenotype (%)	HLA-B	Phenotype (%)	HLA-C	Phenotype (%)	HLA-DRB1	Phenotype (%)	HLA-DRB3/4/5	Phenotype (%)	HLA-DQB1	Phenotype (%)	HLA-DQA1	Phenotype (%)	HLA-DPB1	Phenotype (%)	HLA-DPA1	Phenotype (%)
02:01	46.8	08:01	27.8	07:01	29.8	03:01	28.1	4*01:03	53.3	03:01	29.4	01:02	33.8	04:01	64.8	01:03	95.9
01:01	34.6	07:02	23.2	07:02	28.4	04:01	23.5	3*01:01	38.6	02:01	28.1	05:01	29.1	02:01	28.4	02:01	25.4
03:01	22.5	15:01	17.0	03:04	24.6	15:01	22.3	3*02:02	28.7	03:02	23.5	03:01	22.6	04:02	25.4	02:02	7.0
24:02	18.6	44:02	16.4	04:01	17.6	07:01	19.7	5*01:01	24.9	05:01	22.3	01:01	21.3	01:01	14.5		
11:01	13.2	40:01	15.5	06:02	16.0	01:01	19.1	3*03:01	15.0	06:02	21.6	02:01	20.0	03:01	13.8		
68:01	7.8	35:01	10.6	05:01	13.8	13:01	12.8	4*01:01	13.9	02:02	14.2	03:03	15.8	05:01	4.6		
32:01	7.5	51:01	8.4	03:03	12.5	13:02	12.5	5*01:02	2.6	06:03	12.8	01:03	14.8	10:01	3.9		
26:01	6.2	44:03	6.2	12:03	10.6	11:01	9.4			06:04	11.8	05:05	14.2	06:01	3.2		
31:01	5.6	57:01	5.5	02:02	6.7	04:04	7.4			03:03	8.8	04:01	5.7	09:01	2.8		
29:02	4.9	18:01	5.5	01:02	4.7	08:01	5.7			04:02	5.7	01:04	3.9	17:01	2.5		
25:01	3.9	27:05	5.2	15:02	4.4	14:54	3.9			05:03	4.6	03:02	3.2	13:01	2.5		
23:01	2.6	38:01	4.6	07:04	4.1	09:01	3.2			06:01	2.1	06:01	2.1				
		55:01	4.6	16:01	3.0	04:07	2.5										
		35:03	3.6	08:02	3.0	04:02	2.1										
		39:01	3.3	12:02	2.0	15:02	2.1										
		50:01	2.6			04:08	2.1										
		39:06	2.6														
		45:01	2.6														
		13:02	2.3														
		49:01	2.3														
		52:01	2.3														
		37:01	2.3														
		14:01	2.0														
		56:01	2.0														
		14:02	2.0														
		40:02	2.0														

Positional Scanning Combinatorial Peptide Library

Strategy

- MHC molecules

- Peptides

Position 1	Position 2	....	Position 9
AXXXXXXX (AX7)	XAXXXXXX (XAX6)	....	XXXXXXXA (X8A)
CXXXXXXX (CX7)	XCXXXXXX (XCX6)	....	XXXXXXXC (X8C)
DXXXXXXX (DX7)	XDXXXXXX (XDX6)	....	XXXXXXXD (X8D)
...	....	....	....
YXXXXXXX (YX7)	XYXXXXXX (XYX6)	....	XXXXXXXY (X8Y)

HLA-A02:01
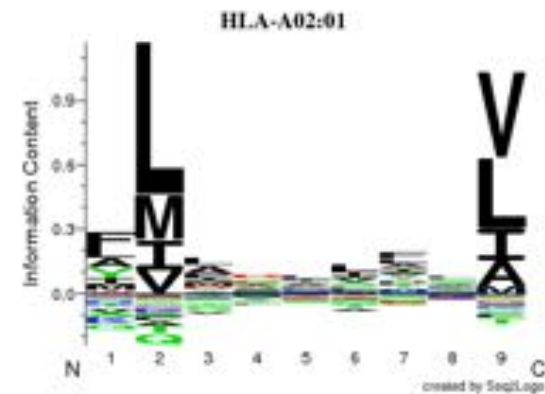
Overlapping Peptide Library

Strategy

- MHC molecules

- Peptides

Yellow Fever Virus RNA genome

3411 aa long YF precursor protein

MSGRKAQGKTLGVNMVRRGVRSLSNKIKQKTKQIGNRPGPSRGVQGFIFFFLFNILTGKKITAHLKRLWKMLDPRQGLAVLRKVKRVVASLMRGLSSRKRRSHDVLTVQFLILGMLLMT

MSGRKAQGKTLGVNM
   KAQGKTLGVNMVRRG
      KTLGVNMVRRGVRSL
         VNMVRRGVRSLSNKI
            RRGVRSLSNKIKQKT
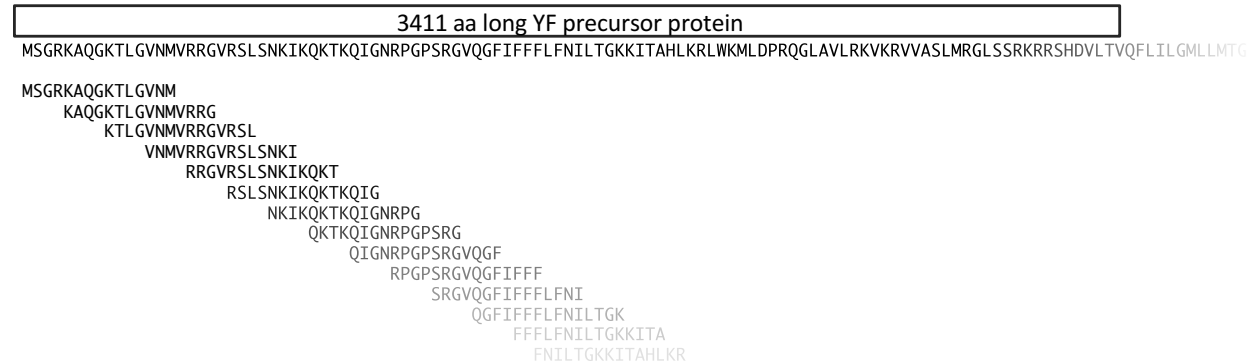   RSLSNKIKQKTKQIG
                  NKIKQKTKQIGNRPG
                     QKTKQIGNRPGPSRG
                        QIGNRPGPSRGVQGF
                           RPGPSRGVQGFIFFF
                              SRGVQGFIFFFLFNI
                                 QGFIFFFLFNILTGK
                                    FFFLFNILTGKKITA
                                      FNILTGKKITAHLKR

Proximity-based high-throughput assays (affinity)

Strategy

- MHC molecules

- Peptides

- Assays

Excitation

Peptide-MHC

Emission

Donor bead

Acceptor bead

O_2^*

Proximity-based high-throughput assays (affinity)

Strategy

- MHC molecules

- Peptides

- Assays

Immunological Bioinformatics @ Center for Biological Sequence Analysis

Proximity-based high-throughput assays (stability)
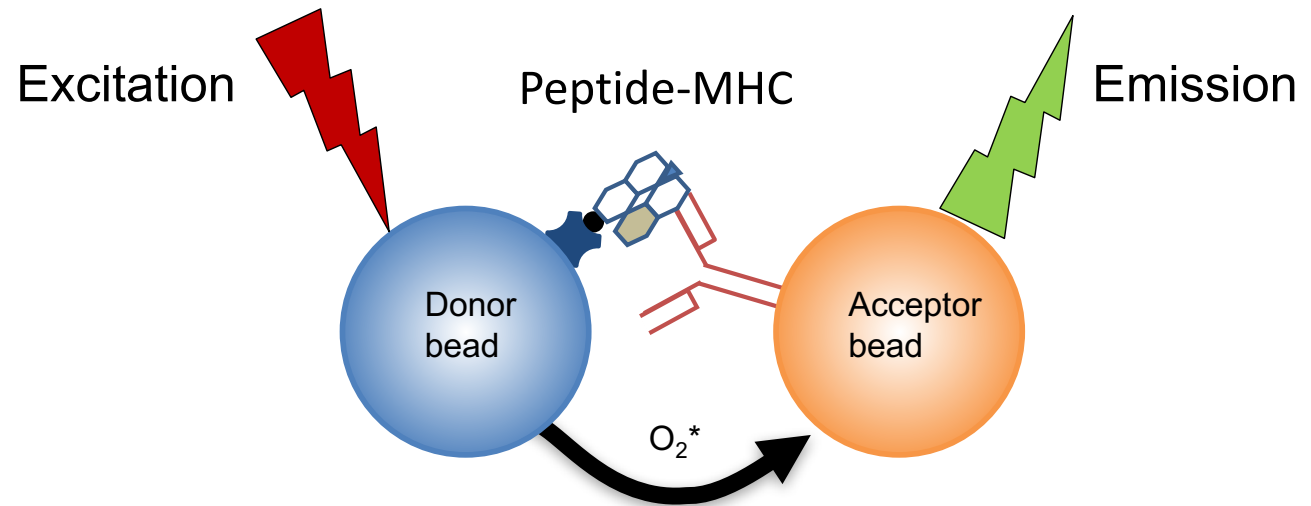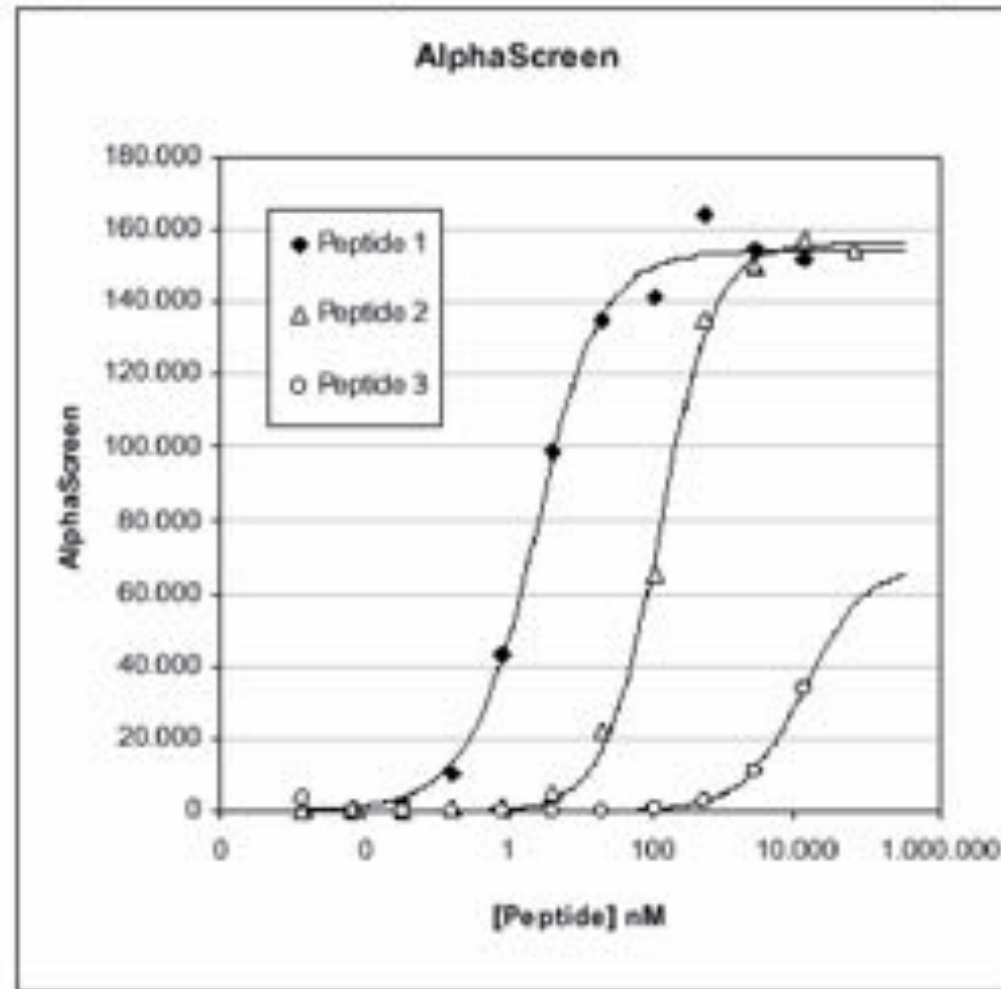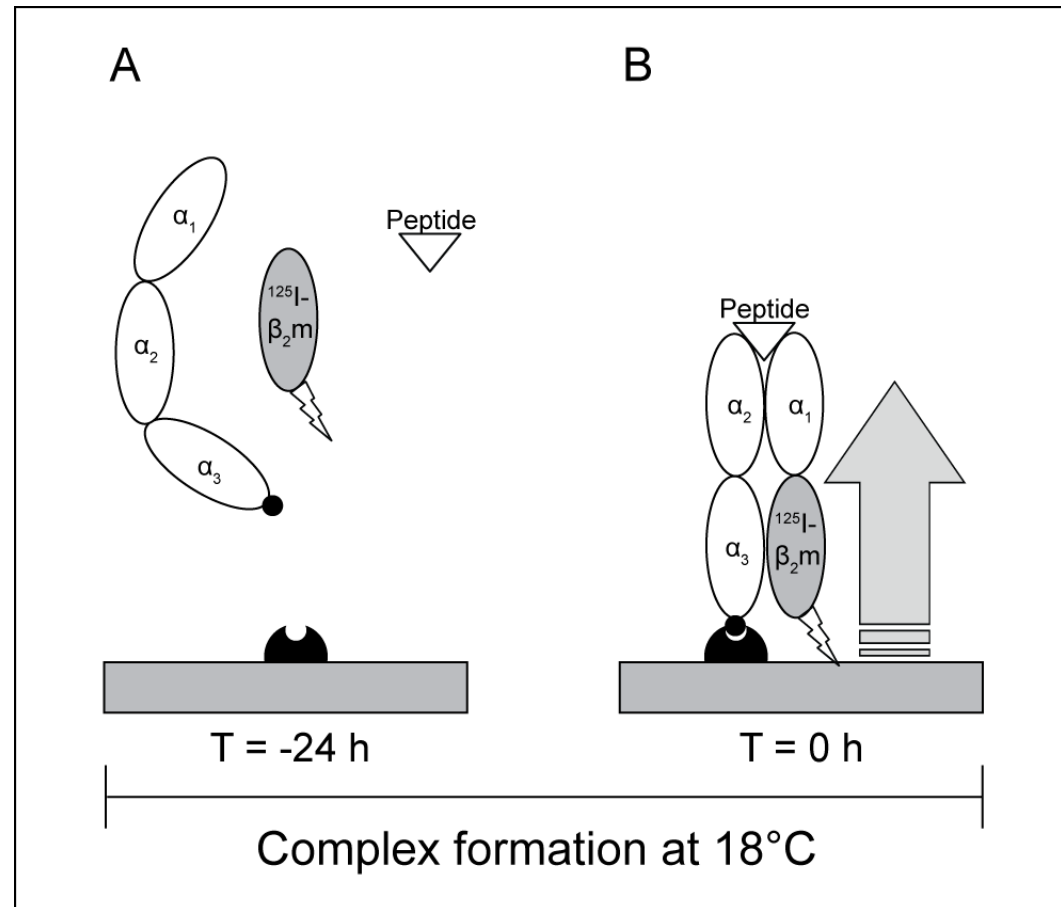
Strategy

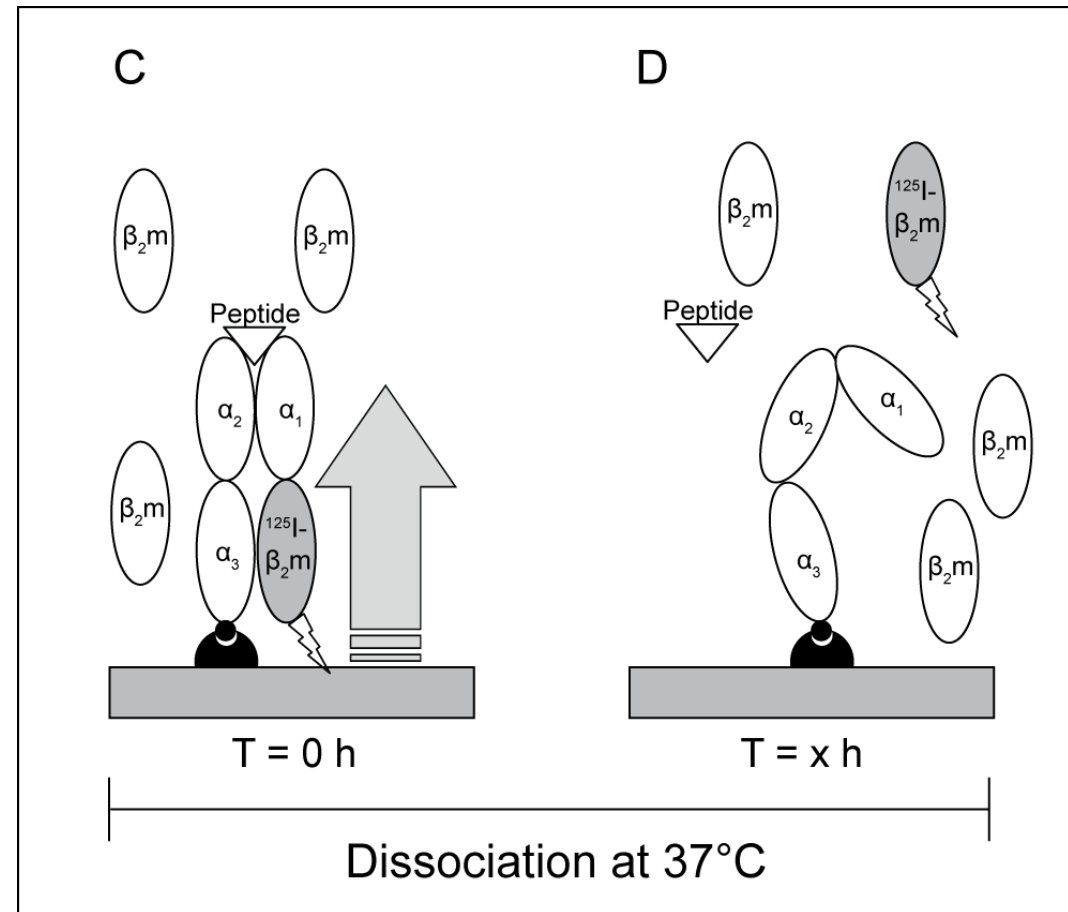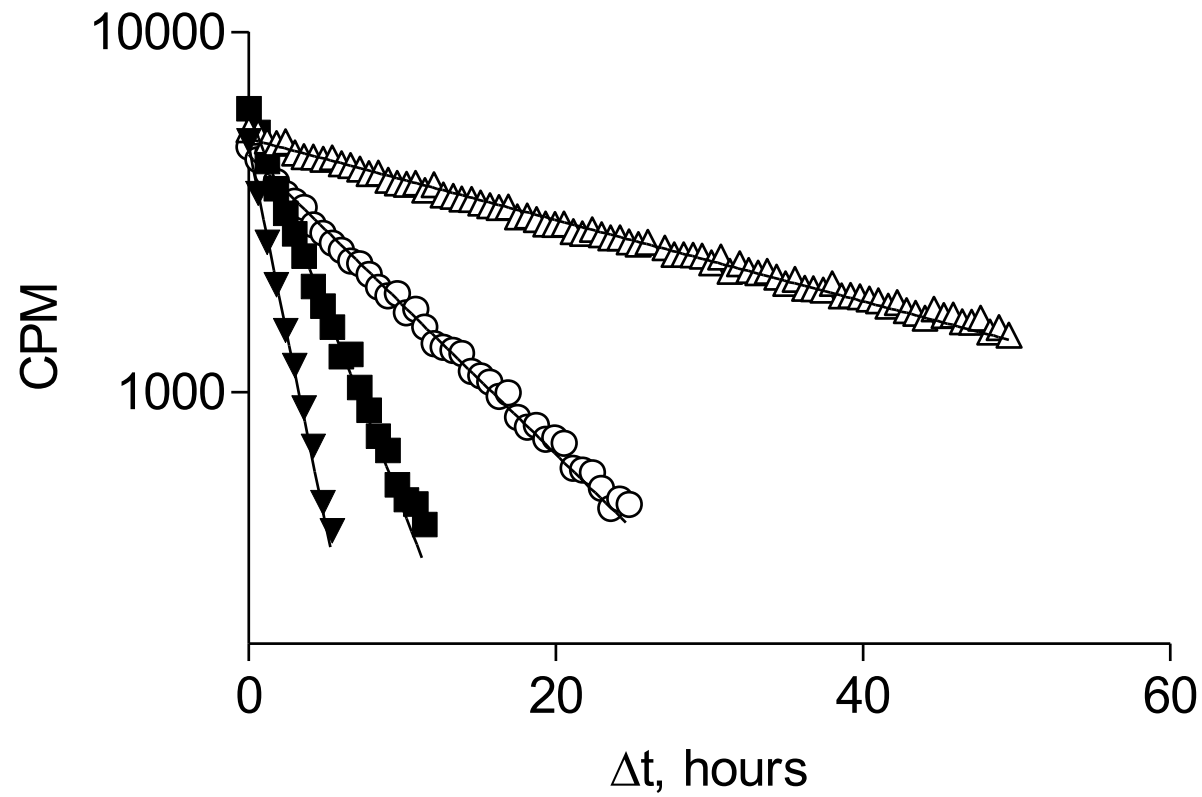- MHC molecules

- Peptides

- Assays

Proximity-based high-throughput assays (stability)
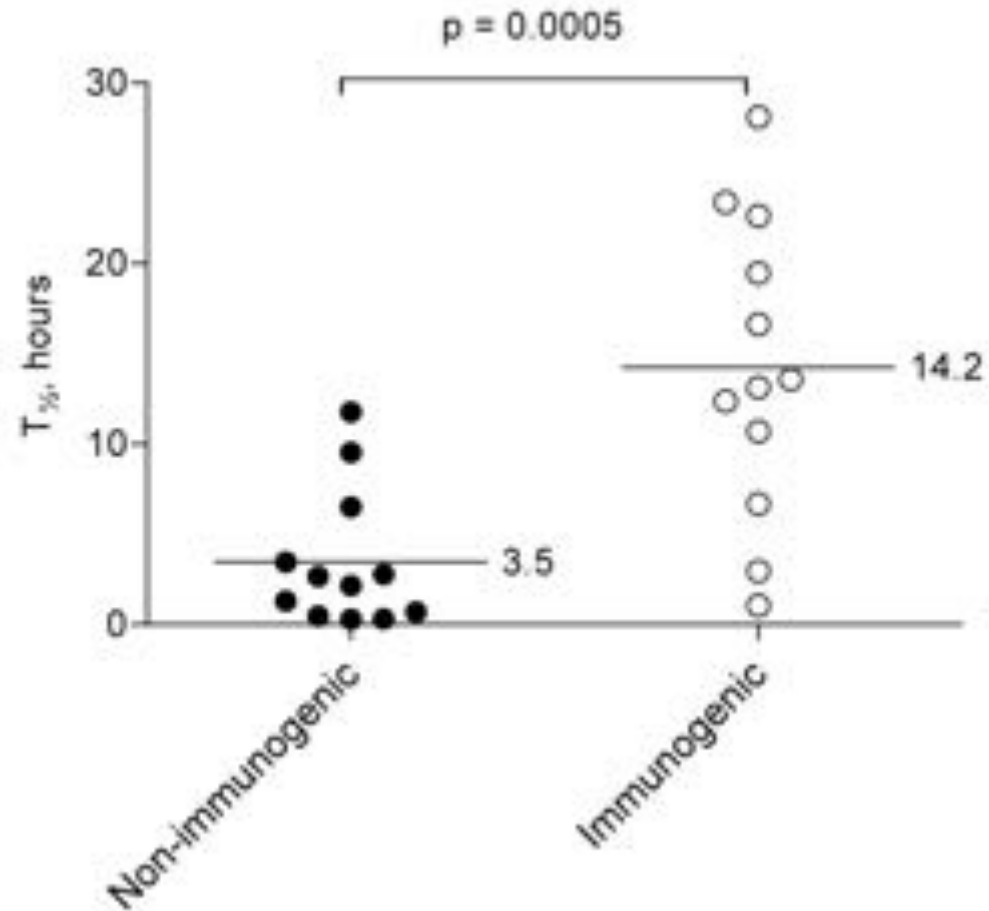
Strategy

- MHC molecules

- Peptides

- Assays

Immunological Bioinformatics @ Center for Biological Sequence Analysis

Proximity-based high-throughput assays (stability)

Stability ≈ more likely to be an immunogen
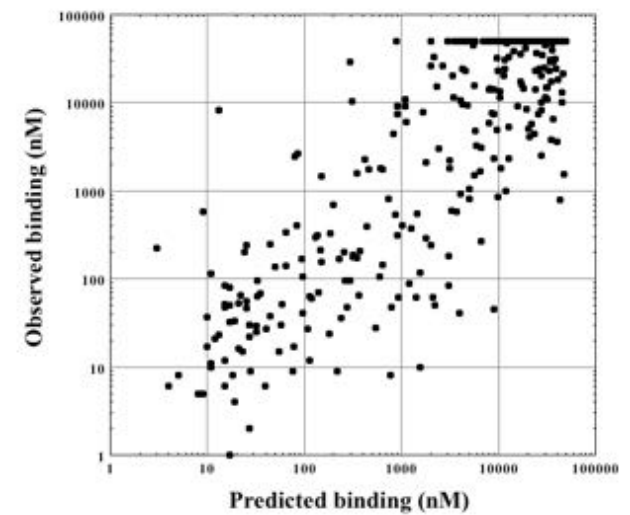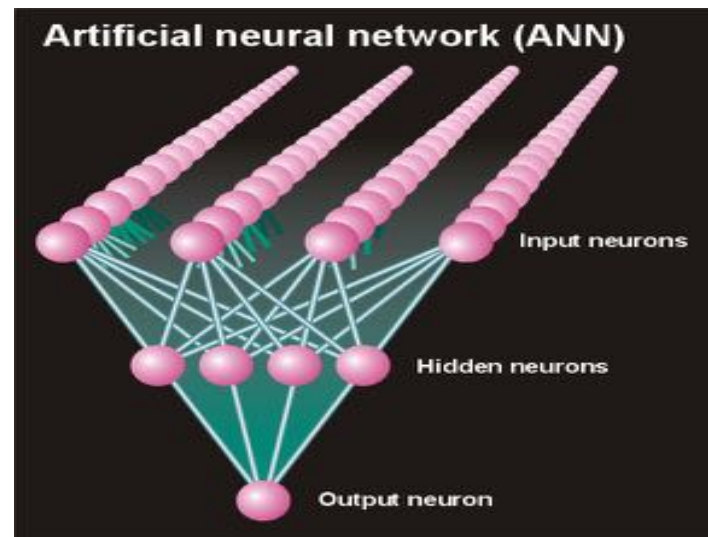
Collecting HLA-specific data

Strategy

- MHC molecules

- Peptides

- Assays

- **Binding data**

MHC Alpha Chain	MHC Beta Chain	Peptide Sequence	Peptide Batchnumbers	Affinity [nM]	Stability [hrs]
HLA-A*02:01	hB2m	VTTEVAFGL	10274	68,6	1,8
HLA-A*02:01	hB2m	FVRQCFNPM	10299	20000	0,1
HLA-A*02:01	hB2m	FVRTLFQQM	10301	20000	0,1
HLA-A*02:01	hB2m	IPKRNRSIL	10292	20000	0,1
HLA-A*02:01	hB2m	QARQMVQAM	10298	20000	0,1
HLA-A*02:01	hB2m	VERLKHGTF	10259, 10308	20000	0,1
HLA-A*02:01	hB2m	YYLEKANKI	10182, 10237, 10319	20000	0,1
HLA-A*02:01	hB2m	FMIDWILDA	8701	3	14,1
HLA-A*02:01	hB2m	LMQWWSDYV	8693	5,7	11,45
HLA-A*02:01	hB2m	GQWDGWVWL	9209	10,4	8,4
HLA-A*02:01	hB2m	TMPELAWAV	8692	15,8	2,2
HLA-A*02:01	hB2m	WMLGTGVYL	8716	17,3	9,8
HLA-A*02:01	hB2m	FVIGGMTGV	8726	19,1	4,5
HLA-A*02:01	hB2m	ALYWALMES	8724	21	4,1
HLA-A*02:01	hB2m	YQYGDNLIL	9295	22,1	2
HLA-A*02:01	hB2m	HIMPNSFRV	8708	24	9,6
HLA-A*02:01	hB2m	FIAEIDHWI	8697	35	2,7
HLA-A*02:01	hB2m	YLHRDIFDI	8695	36,5	5,4
HLA-A*02:01	hB2m	KLADMSIYC	8713	42,7	22,3
HLA-A*02:01	hB2m	LMAEDLANV	8696	47,2	23,4
HLA-A*02:01	hB2m	YMLWNSWLS	8706	48,4	5,9

Immunological Bioinformatics @ Center for Biological Sequence Analysis

Strategy

- MHC molecules

- Peptides

- Assays

- Binding data

- **Predictors**

Immunological Bioinformatics @ Center for Biological Sequence Analysis
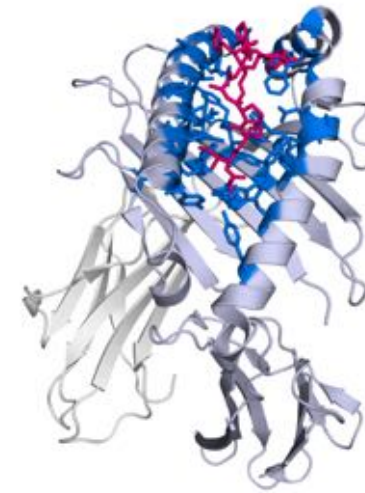
Strategy

- MHC molecules

- Peptides

- Assays

- Binding data

- Predictors

- Publish

NetMHC   # NetMHCpan

- NetMHC
- NetMHCII
- NetMHCstab
- NetChop
- NetCTL

- NetMHCpan
- NetMHCIIpan
- NetMHCstabpan
- NetCTLpan

http://www.cbs.dtu.dk/services or http://tools.iedb.org/main/

Peptide-HLA class I and II tetramers