Epitope-based vaccines and the educated search for epitopes

Dr. Pedro A Reche, PhD.
Facultad de Medicina,
Dpto. Microbiology I – Immunology,
Universidad Complutense de Madrid, Madrid, Spain

Content

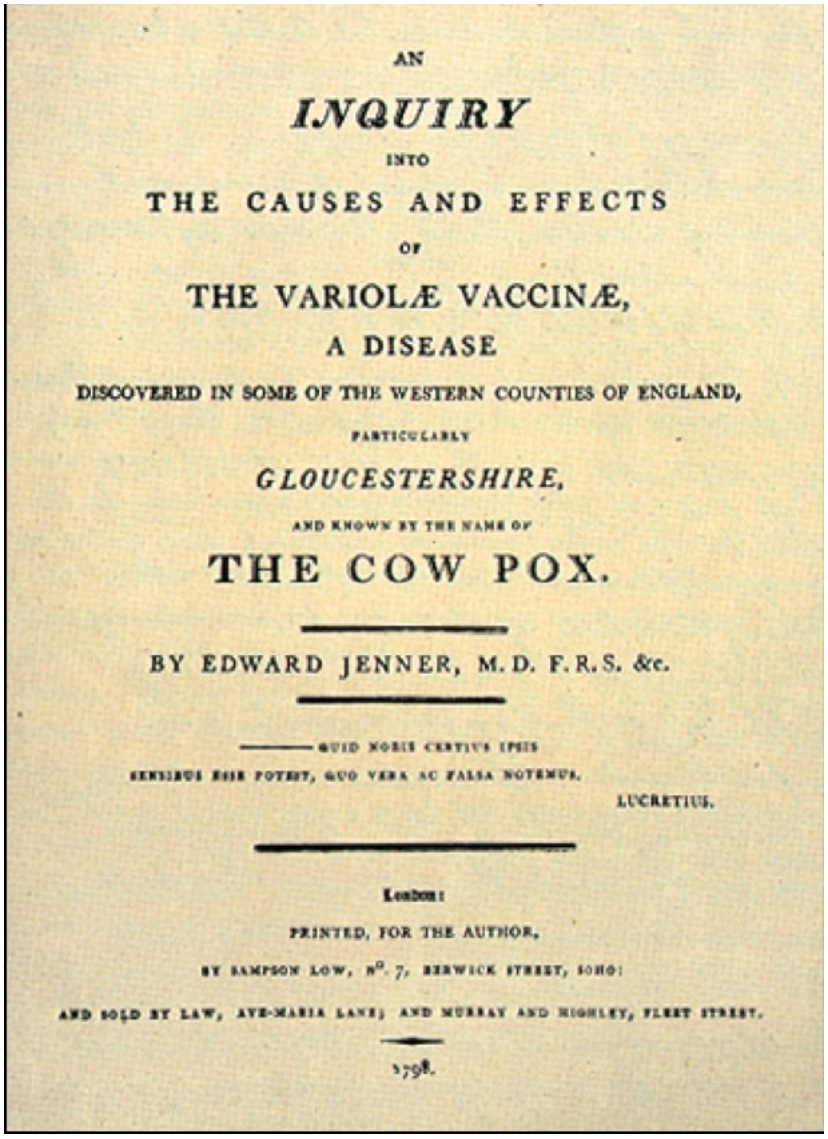
1. Something about vaccines & epitopes

1. Prediction of B cell epitopes

1. Prediction of T cell epitopes

Vaccines development: the science of inducing protective adaptive immunity

- 1600 AD
  - Turks uses variolation to prevent smallpox
- 1700AD
  - Introduction of variolation in England and later in the US
- 1780AD
  - Edward Jenner introduces cowpox (*variolae vaccinae*) inoculation to prevent smallpox
- 1860-80AD
  - Pasteur introduced various treatment for infectious diseases using with attenuated pathogens (vaccines)



AN

INQUIRY

INTO

THE CAUSES AND EFFECTS

OF

THE VARIOLÆ VACCINÆ,

A DISEASE

DISCOVERED IN SOME OF THE WESTERN COUNTIES OF ENGLAND,

PARTICULARLY

GLOUCESTERSHIRE,

AND KNOWN BY THE NAME OF

THE COW POX.

BY EDWARD JENNER, M.D. F.R.S. &c.

——— QUID NOBIS CERTIUS IPSIS
SENSIBUS HAEC POTEST, QUO VERA AC FALSA NOTEMUS.

LUCRETIUS.

London:

PRINTED, FOR THE AUTHOR,

BY SAMPSON LOW, N°. 7, BERWICK STREET, SOHO:

AND SOLD BY LAW, AVE-MARIA LANE; AND MURRAY AND HIGHLEY, FLEET STREET.

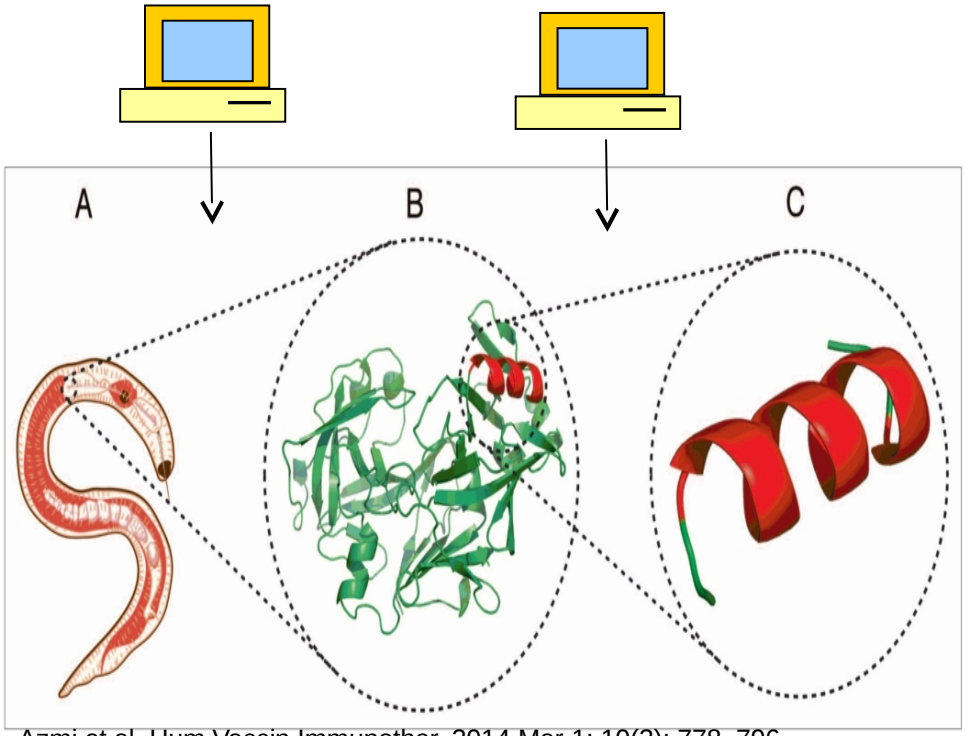
1798.

Vaccines: past & future

Current vaccines

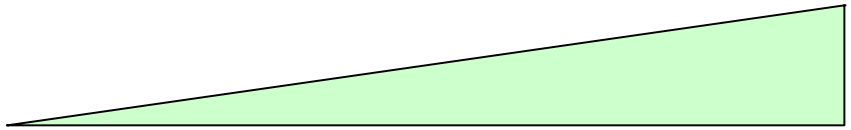
Organism	Type	Vaccine Type	Year
Variola virus	Virus	Live	1798
Rabies virus	Virus	Inactivated	1885
Salmonella typhi	Bacteria	Live	1896
Vibrio cholerae	Bacteria	Inactivated	1896
Yersinia pestis	Bacteria	Inactivated	1897
Corynebacterium diphtheriae	Bacteria	Toxoid	1923
Bordetella pertussis	Bacteria	Acellular	1926
Clostridium tetani	Bacteria	Toxoid	1927
Mycobacterium tuberculosis	Bacteria	Live	1927
Yellow fever virus	Virus	Live	1935
Influenza virus type A	Virus	Inactivated	1936
Influenza virus type B	Virus	Inactivated	1936
Coxiella burnetii	Bacteria	Inactivated	1938
Rickettsia prowazekii	Bacteria	Inactivated	1938
Rickettsia rickettsii	Bacteria	Inactivated	1938
Central European encephalitis virus	Virus	Inactivated	1939
Poliovirus types 1, 2, and 3	Virus	Inactivated/Live	1962
Measles virus	Virus	Live	1963
Mumps virus	Virus	Live	1967
Rubivirus	Virus	Live	1969
Staphylococcus aureus	Bacteria	Staphage lysate	1976
Streptococcus pneumoniae	Bacteria	Polysaccharide	1977
Human adenovirus types 4 and 7	Virus	Live	1980
Neisseria meningitidis	Bacteria	Polysaccharide	1981
Hepatitis B	Virus	Recombinant	1986
Haemophilus influenzae	Bacteria	Conjugate	1987
Hantaan virus	Virus	Inactivated	1989
Japanese encephalitis virus	Virus	Inactivated	1992
Varicella-zoster virus	Virus	Live	1994
Hepatitis A	Virus	Inactivated	1995
Escherichia coli	Bacteria	Inactivated	1995
Junin virus	Virus	Live	1996
Bacillus anthracis	Bacteria	Adsorbed	1998
Borrelia burgdorferi	Bacteria	Recombinant	1998

Evolution in vaccine development



Azmi et al. Hum Vaccin Immunother. 2014 Mar 1; 10(3): 778–796

Pathogen Antigen Epitope

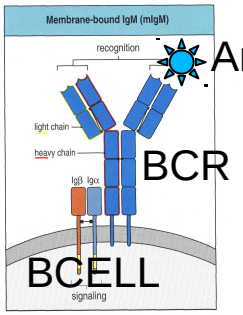


knowledge & technology

Why bother with new epitope based vaccines?

- Epitope-based vaccines have been shown to confer protection in animal models (Rodriguez et al. [1998] and Sette and Sidney [1999])
- No conventional vaccines are available for highly pathogenic organisms (HIV-1, HCV, Plasmodium etc).
- They have important advantages over pathogen or protein-based vaccines
  - Controlled specificity: Force immune response towards conserved subdominant epitopes when dominant are variable)
  - Less undesired responses
  - Epitope vaccines can be formulated in many ways including as DNA vaccines
  - Fast and safe to develop: We have knowledge and technology to predict epitopes from primary sequences. No need for pathogen cultures

B cell and T cell epitopes



Antigen

BCR

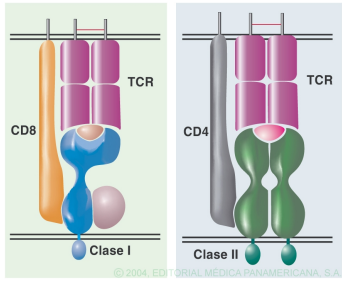
BCELL

Antigen

B cell epitope

Antibody

T cell



APC

TCR

T cell epitope

MHC

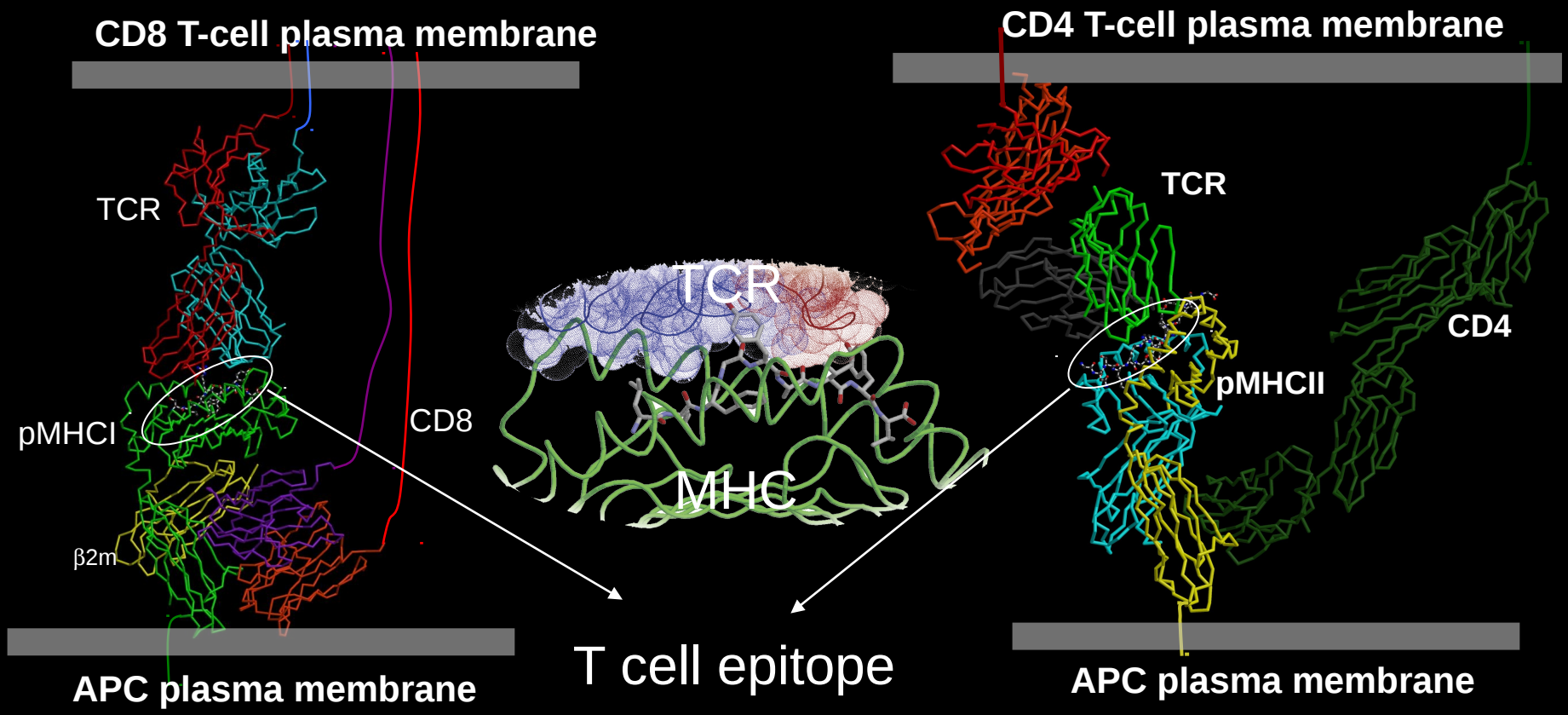
B cell epitopes

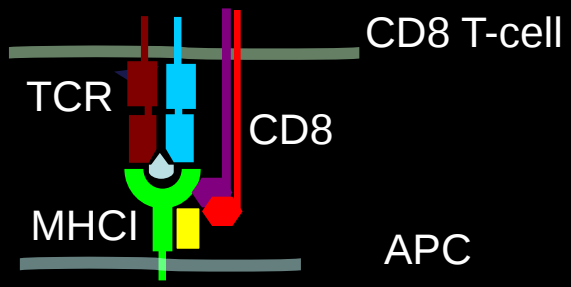
Solvent exposed antigen regions recognized by BCR & Antibodies

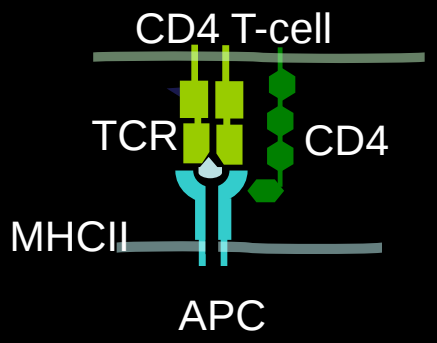
T cell epitopes

Peptides recognized by TCR when presented by APC bound to MHC molecules

Antigen recognition by T cells

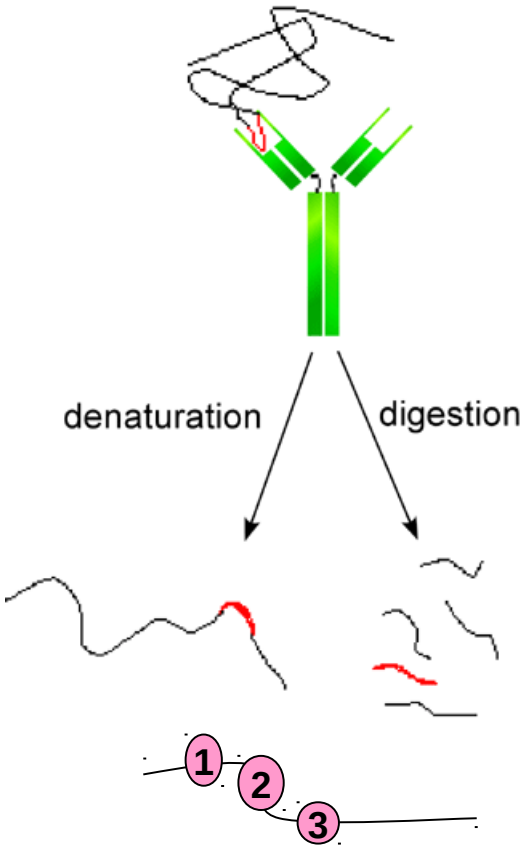






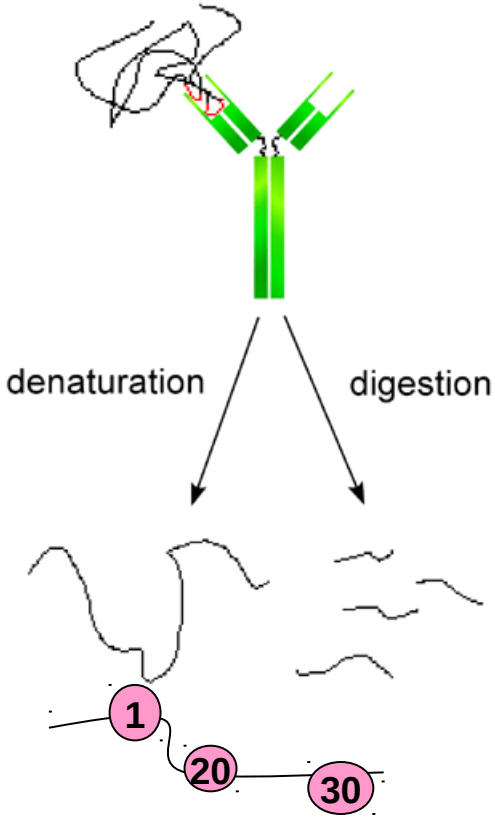
B cell epitopes

Linear/continuous



linear B epitope.
sequential residues

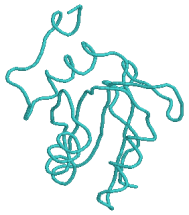
Conformational/discontinous



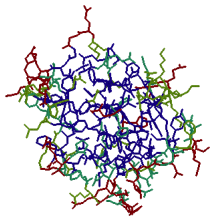
Conformational B epitope
non-sequential residues

Prediction of discontinuous B cell epitopes

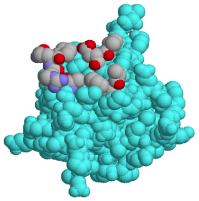
1. Get or model 3D-structure of antigen



1. Determine Residue Solvent Accessible (RAS)



1. B cell epitopes: clusters of solvent of accessible residues + features

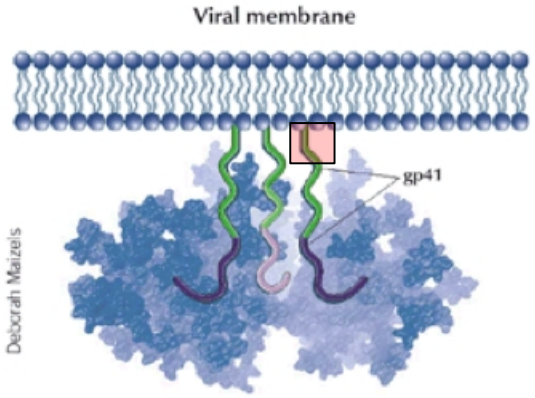


1. Examples:
    1. CEP: http://196.1.114.49/cgi-bin/cep.pl
    2. DISCOTOPE: http://www.cbs.dtu.dk/services/DiscoTope/
    3. PEPITO: www.igb.uci.edu

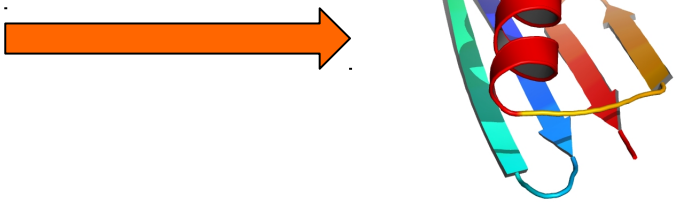
Conformational epitopes and epitope vaccines

- Conformational epitopes can not be isolated from their context
- Solutions to deliver isolated conformational epitopes
  - Mimetopes: Linear peptides obtained by phage display that binds to the same antibodies than conformational epitopes
  - Epitope grafting: Transfer conformational epitope to proteins

HIV-1 gp41 peptide



GB1 grafted with gp41 peptide



He L, Cheng Y, Kong L, Azadnia P, Giang E, Kim J, Wood MR, Wilson IA, Law M, Zhu J. Approaching rational epitope vaccine design for hepatitis C virus with meta-server and multivalent scaffolding. Sci Rep. 2015;5:12501. doi: 10.1038/srep12501.

Prediction of linear B cell epitopes

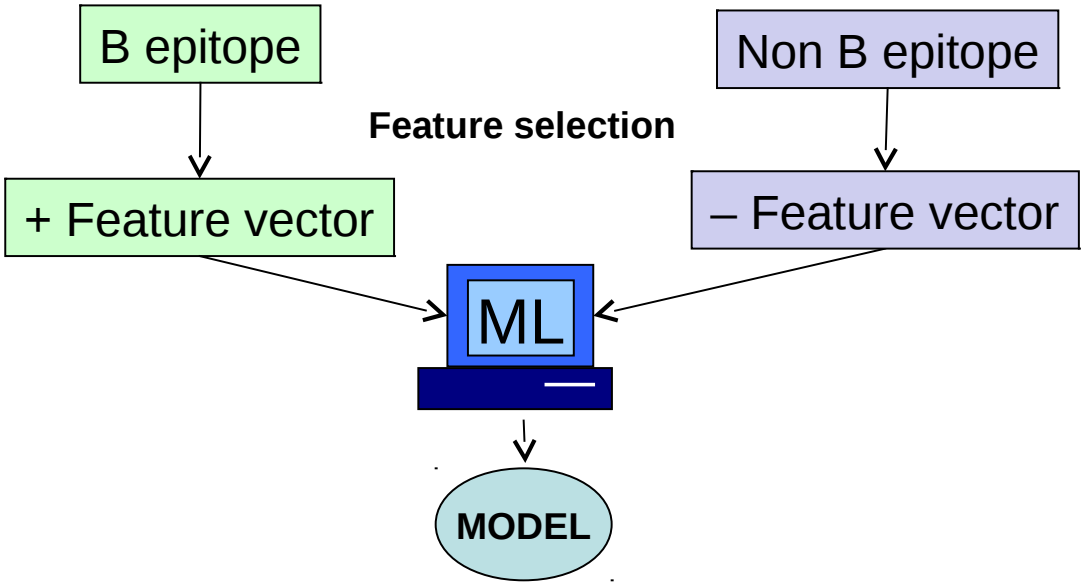
Amino Acid Propensity Scales

- Hopp and Woods, 1981. *Hydrophilicity scale*
- Welling et al, 1985. *Relative occurrence of amino acids*
- Parker & Hodges, 1986. *Hydrophilicity, Accessibility & Flexibility*
- Kolaskar & Tongaonkar, 1990. *Relative occurrence and amino acid properties*
    - **antigenic:** http://imed.ucm.es/Tools/antigenic.pl

Blythe and Flower, 2005, Protein Sci 14:246: “single-scale amino acid propensity profiles cannot be used to predict epitope location reliably

Prediction of linear B cell epitopes

- **Machine Learning approaches:** Classification of B cell epitopes vs non B cell epitopes



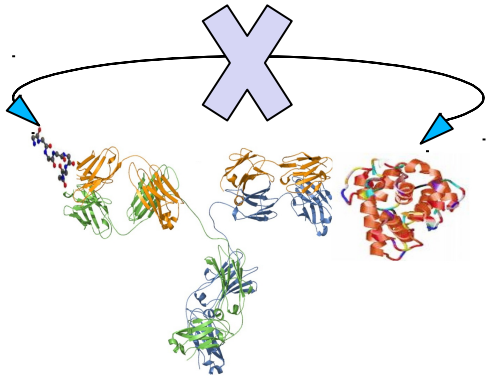
- Examples:
  - ABCPRED: http://www.imtech.res.in/raghava/abcpred/
  - BEPIPRED: http://www.cbs.dtu.dk/services/BepiPred/
  - BCPRED: http://ailab.ist.psu.edu/bcpred/

Greenbaum et al., 2009 J Mol Recognit 20:75–82. : The combination of scales with machine learning algorithms showed little improvement over single scale-based methods, which were considered to perform inadequately''

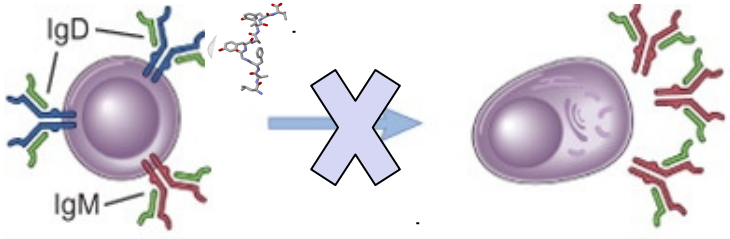
Prediction of linear B cell epitopes: good and bad news

- Any peptide equal or longer than 15 aa is likely to be antigenic: there is always a BCR up to the job

- Antibodies elicited by peptides do not always recognize native antigens

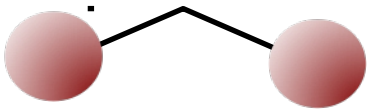


- Absence of antibody specific response can be due to lack of immunogenicity (think of haptens)

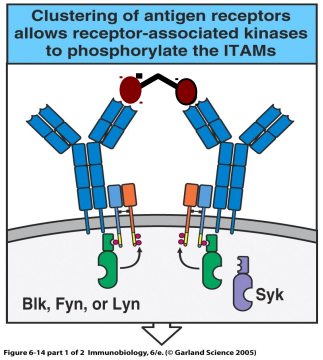


Increasing immunogenicity of B cell epitopes

1. Improve BCR clustering: Concatenate B cell epitopes

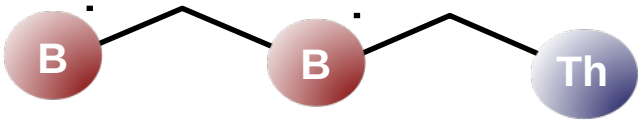


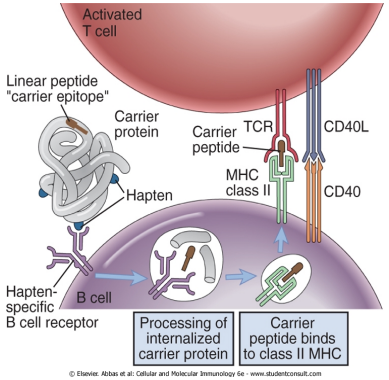
B-linker-B



2. Improve Th recruitment by enhancing MHC II presentation

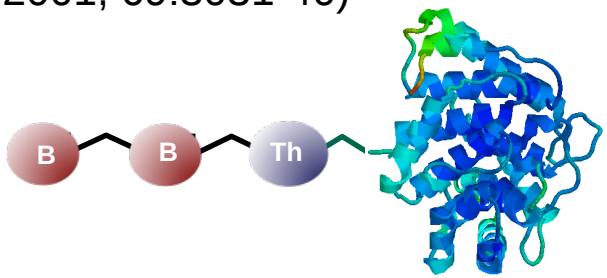
2.1. Conjugate B cell epitopes with Th epitopes (Brumeanu et al J Virol. 1997, 71:5473-80

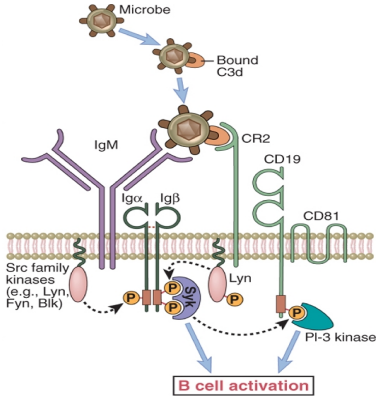




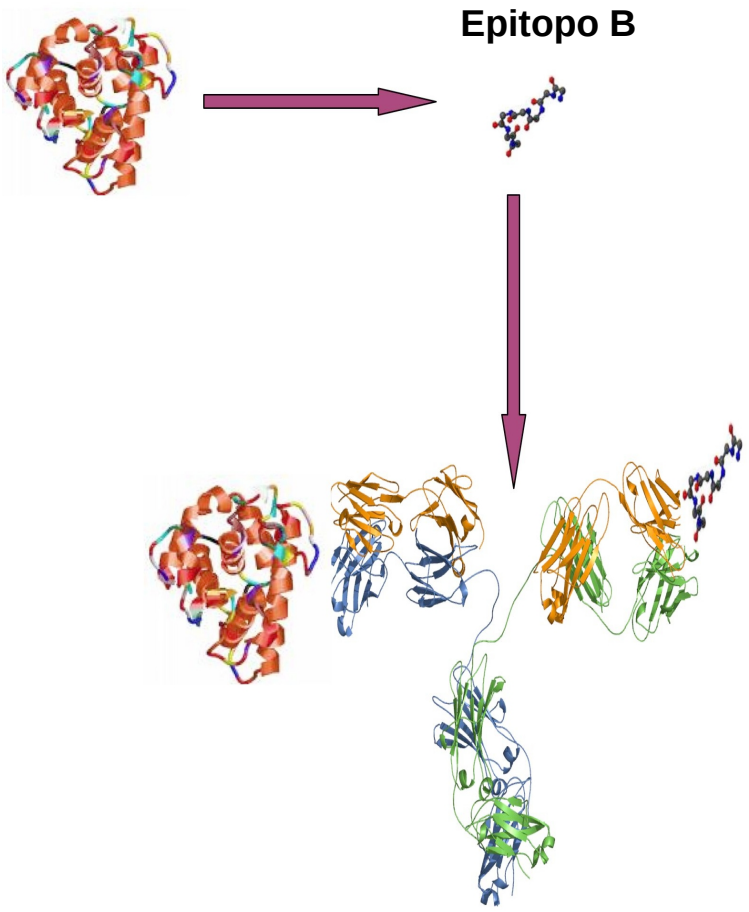
2.3. Conjugate with carrier protein

3. Engage BCR co-receptor: conjugate with C3d (Test et al., Infect Immun. 2001, 69:3031-40)





Bceps: A software that predicts linear B cell epitopes eliciting cros-reactive antibodies with native antigen



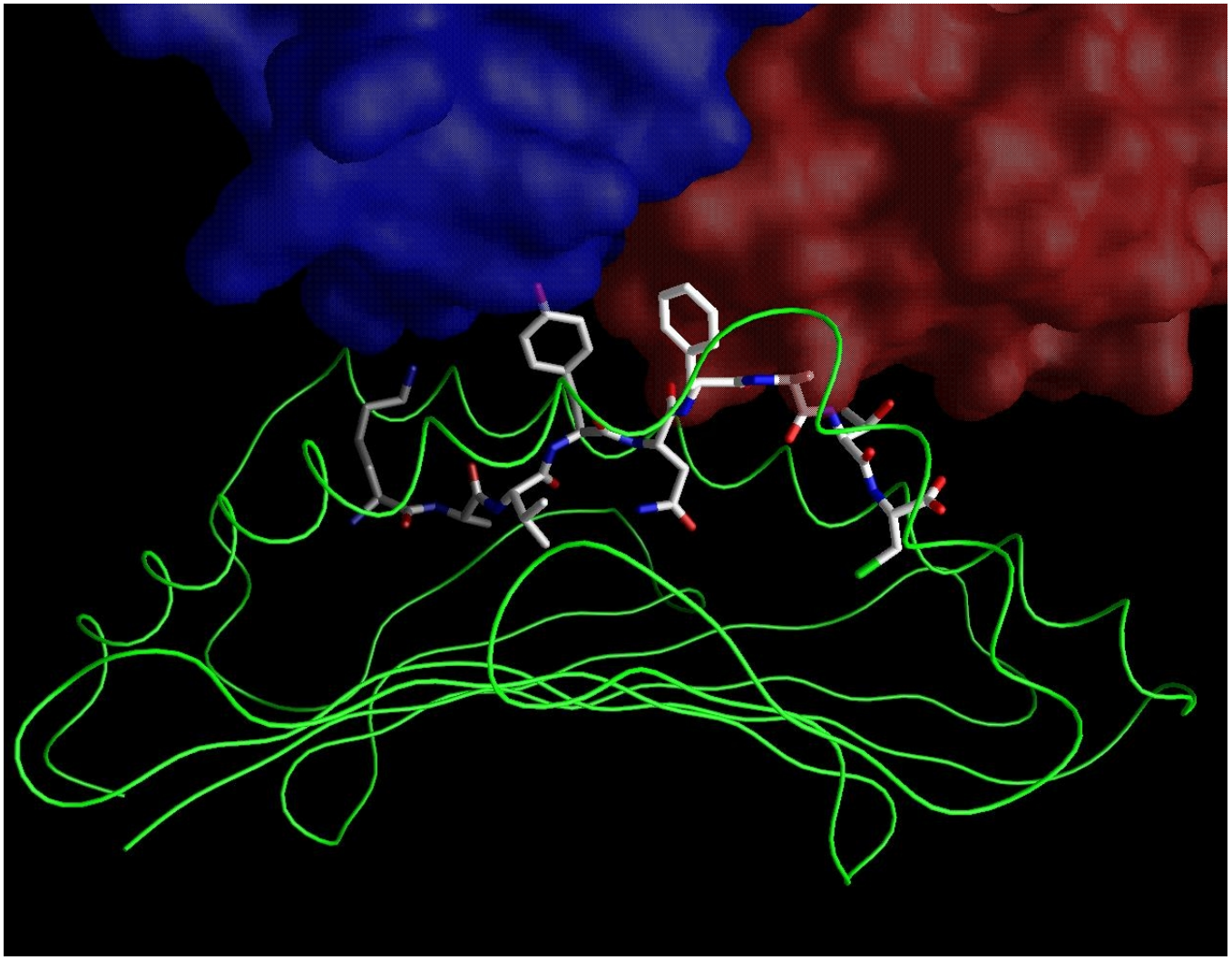
- **1. Predict B cell epitopes using support vector machine model**
- **2. Select epitopes mapping in solvent accessible loops as determined by flexibility measures (F):**

  • 2.1. Known 3D structures: Use normalized B factors

  • 2.2. Unknown 3D structure: Predict flexibility using ProfBal

  _

T cell epitopes

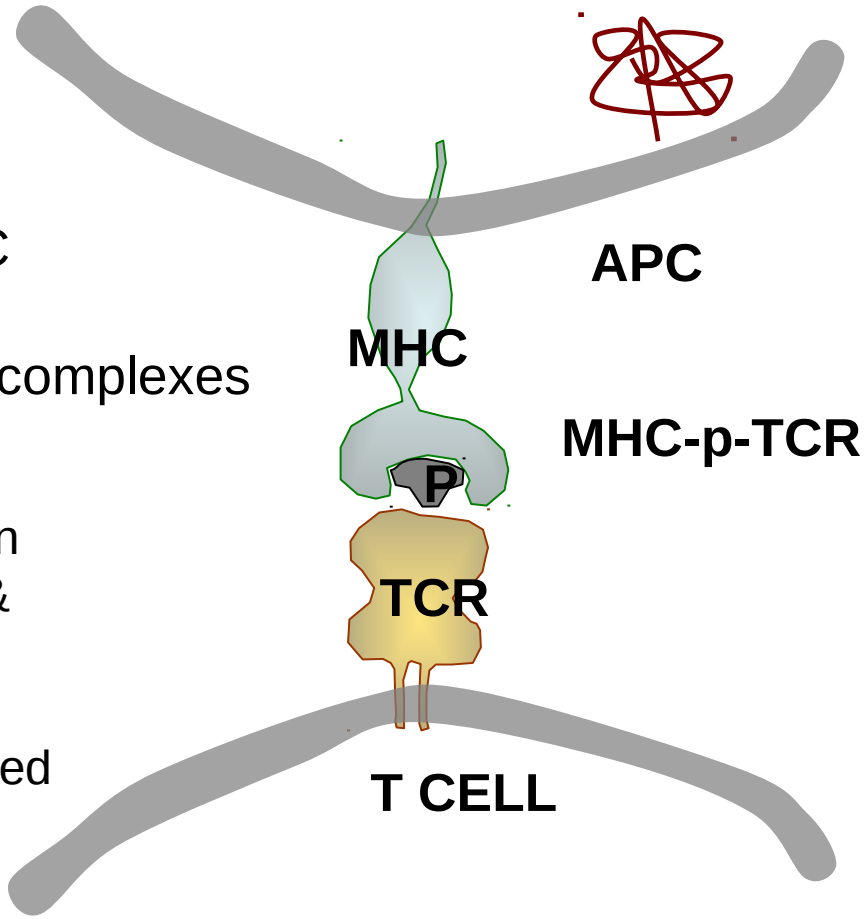


Peptides presented by MHC molecules capable of eliciting a T cell response

T cell epitope prediction

T cell epitope immunogenicity is contingent on:

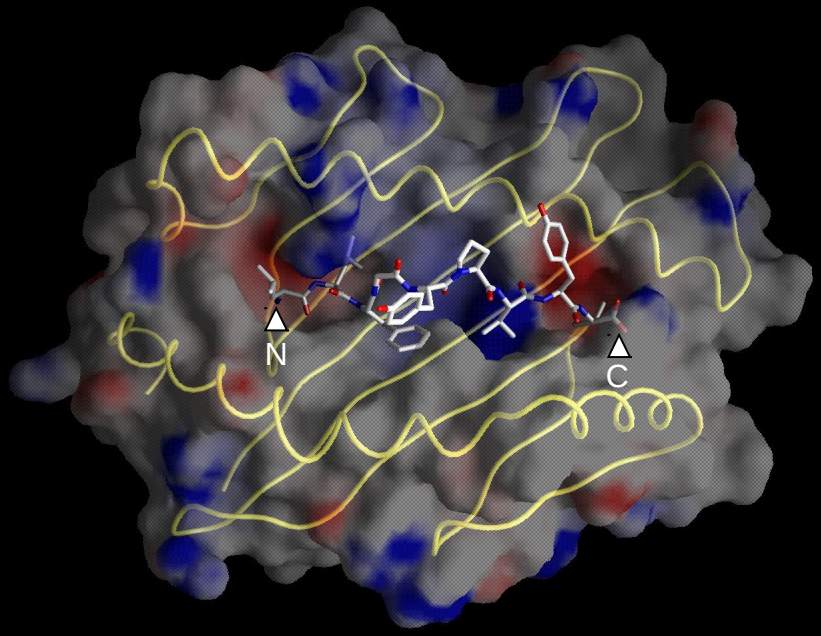
- Appropriated antigen processing
- Binding of processed peptides to MHC
- Availability of cognate TCR for pMHC complexes

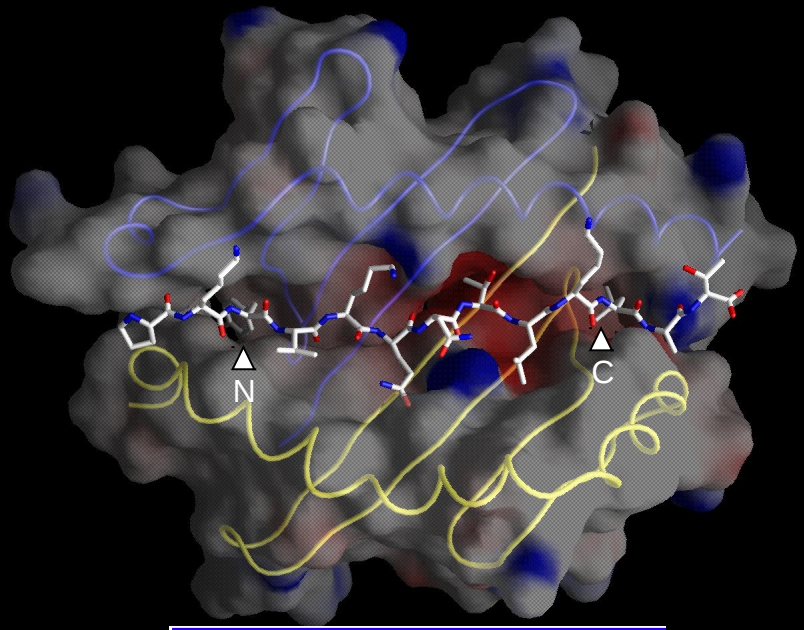


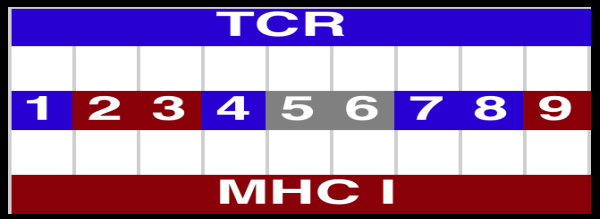
Prediction of peptide-MHC binding is the main basis to anticipate T cell epitopes (Lafuente & Reche, Curr Pharm Des. 2009;15:3209)

- T cells only recognize peptides presented by MHC
- Binding of peptides to MHC is the most selective step determining epitope immunogenicity

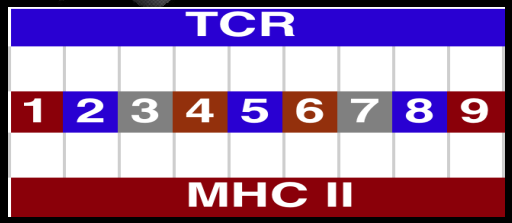
Peptide binding to MHC molecules







anchor



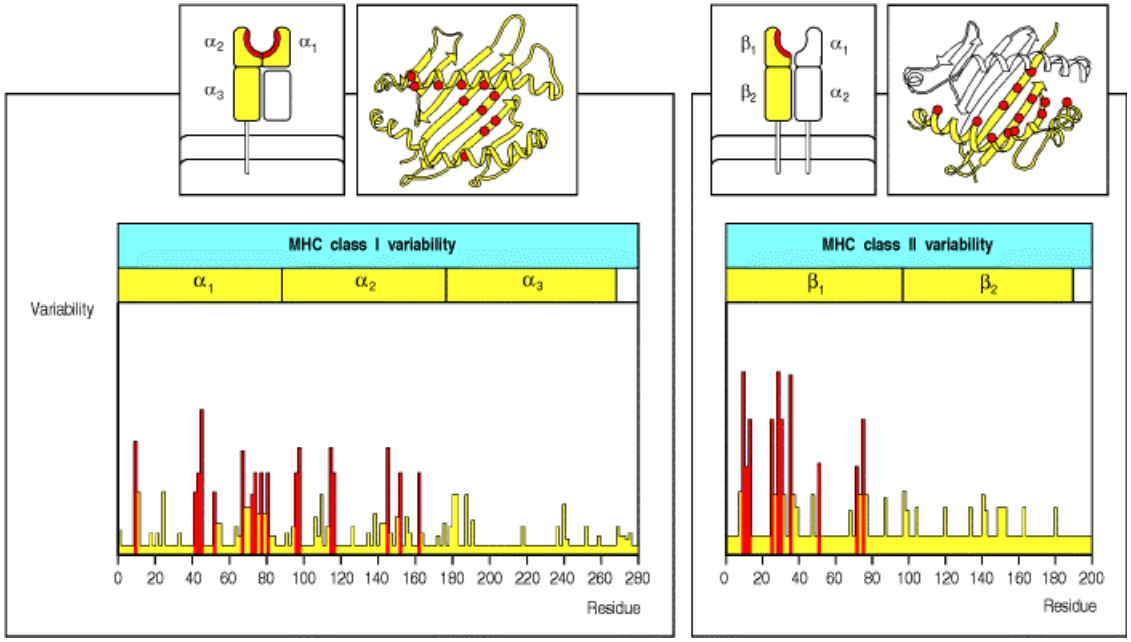
- MHC I molecules bind short peptides (9-11) and are very selective

- MHC II molecules bind long peptides (9-22), a core of 9 residues sits in the binding groove and ends protrude, and display less peptide binding selectivity than MHC I molecules

Human MHC molecules aka HLA molecules are highly polymorphic

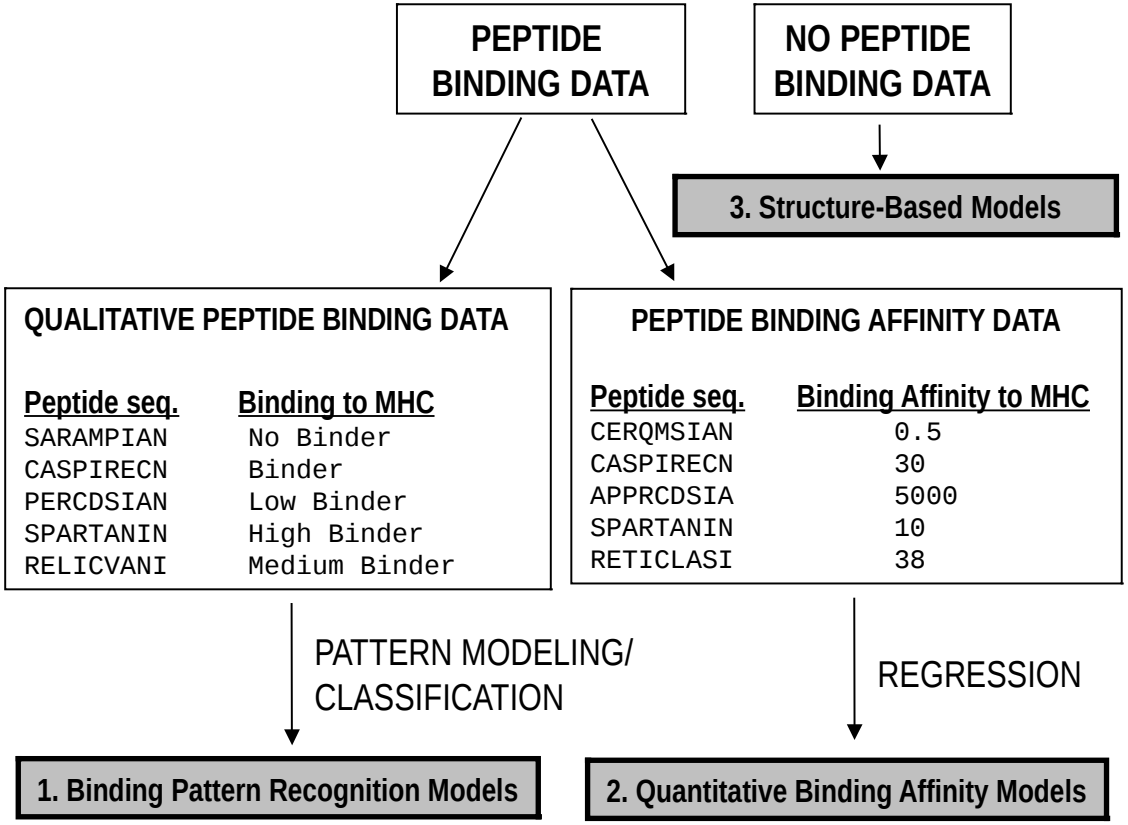
- There are hundred of allelic variants that are expressed at different frequencies in distinct ethnic groups

- HLA allelic variants bind distinct sets of peptides (Reche and Reinherz, 2005, J Mol Biol.; 331:623)



- T cell epitope presentation and recognition varies between individuals and each MHC molecule requires a specific predictor of peptide binding

Classification of peptide-MHC binding prediction methods

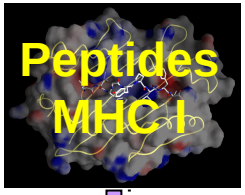


RANKPEP: http://imed.med.ucm.es/Tools/rankpep.html
SVMHC: http://abi.inf.uni-tuebingen.de/Services/SVMHC

PROPED: http://www.imtech.res.in/raghava/propred/
PREPEDII: http://www.imtech.res.in/raghava/propredII/
NETMHCI: http://www.cbs.dtu.dk/services/NetMHCI/
NETMHCI: http://www.cbs.dtu.dk/services/NetMHCII/

Lafuente & Reche, Curr Pharm Des. 2009;15:3209

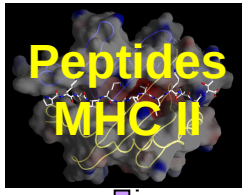
Prediction of peptide binding to MHC molecules using motif profiles



class

DATABASE OF PEPTIDES KNOWN TO BIND TO MHC

class



Size, allele

allele

pMHCI$_a$,8 | pMHCI$_a$,9 | pMHCI$_a$, 10

pMHCII$_a$,9-22

Ungapped peptides alignement

G	A	P	V	P	I	Q	M	L
A	I	R	R	Q	L	A	A	L
V	L	V	Q	N	Y	T	V	I
D	L	N	N	S	A	D	D	V
R	V	R	R	K	S	R	R	V
Y	M	S	A	E	F	Y	F	I
S	V	F	D	S	Y	S	S	A
T	L	D	G	V	L	G	K	M

Meme motif finder, 9aa

$w_{ij} = \ln (f_{obs} / f_{exp})$

Matrix Profile

Prediction

RANK	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
1	91	GCA	SLAKEMFAM	KTK	1009.25	13.527	32.14 %
2	4	MSA	LLILALVGA	AVA	864.14	12.512	29.73 %
3	6	ALL	ILALVGAAV	ADY	808.03	12.133	28.83 %
4	23	DDD	KLAAAYDFT	NCD	981.12	11.958	28.41 %
5	30	AYD	FTNCDFEKI	KAA	1098.24	11.0	26.13 %
6	2	M	SALLILALV	GAA	894.17	9.468	22.49 %
7	49	ISK	DLITYMSGT	KST	982.11	8.163	19.39 %
107	31	YDF	TNCDFEKIK	AAY	1079.23	-35.202	-83.63 %
108	110	IWC	PGYSETQIN	AT	990.04	-37.247	-88.49 %
109	34	TNC	DFEKIKAAY	LST	1066.23	-38.476	-91.41 %
110	20	YKD	DDDKLAAAY	DFT	963.02	-39.348	-93.48 %
111	72	SNR	PHCLTEIQS	LTF	1009.15	-40.417	-96.02 %
112	21	KDD	DDKLAAAYD	FTN	963.02	-42.127	-100.09 %

Protein/Antigen

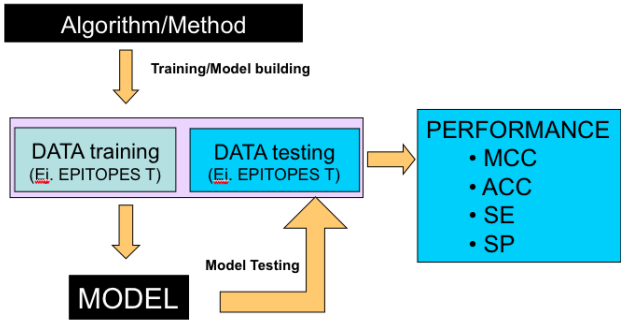
```
MSALLILALVGAAVADYKDDDDKLAAAYDF
TNCDFEKIKAAYLSTISKDLITYMSGTKST
EFNNTVSCSNRPHCLTEIQSLTFNPTAGCA
SLAKEMFAMKTKAALAIWCPGYSETQINAT
```

Reche, P.A, Glutting, J.P., and Reinherz, E.L. (2002) Prediction of peptide binding to class I MHC molecules using profile motifs. *Hum Immunol.,* **63**, 701-709. PMID 12175724.

Reche, P.A.*, Glutting, J.P., Zhang, H. and Reinherz, E.L. (2004) Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. *Immunogenetics*, **56**, 405-419. PMID 15349703.

Profile validations for peptide-MHC binding prediction

- Prediction power of RANKPEP profiles in cross-validation



	Sensitivity (3%)	Specificity (3%)	AUC
MHCI	93 % ± 6	98% ± 5	0.95 ± 0.06
MHCII	87 % ± 8	89% ± 8	0.89 ± 0.07

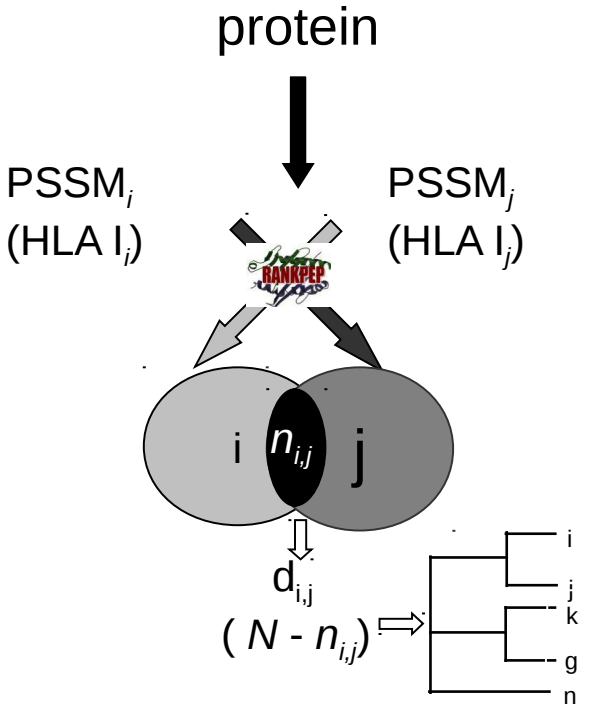
AUC: Area under ROC (Se vs 1-SP). Capacity of models to discriminate between the peptides that bind to MHC from those that do not bind

- Most T cell epitopes rank among the 2% top scoring peptides in their protein sources:

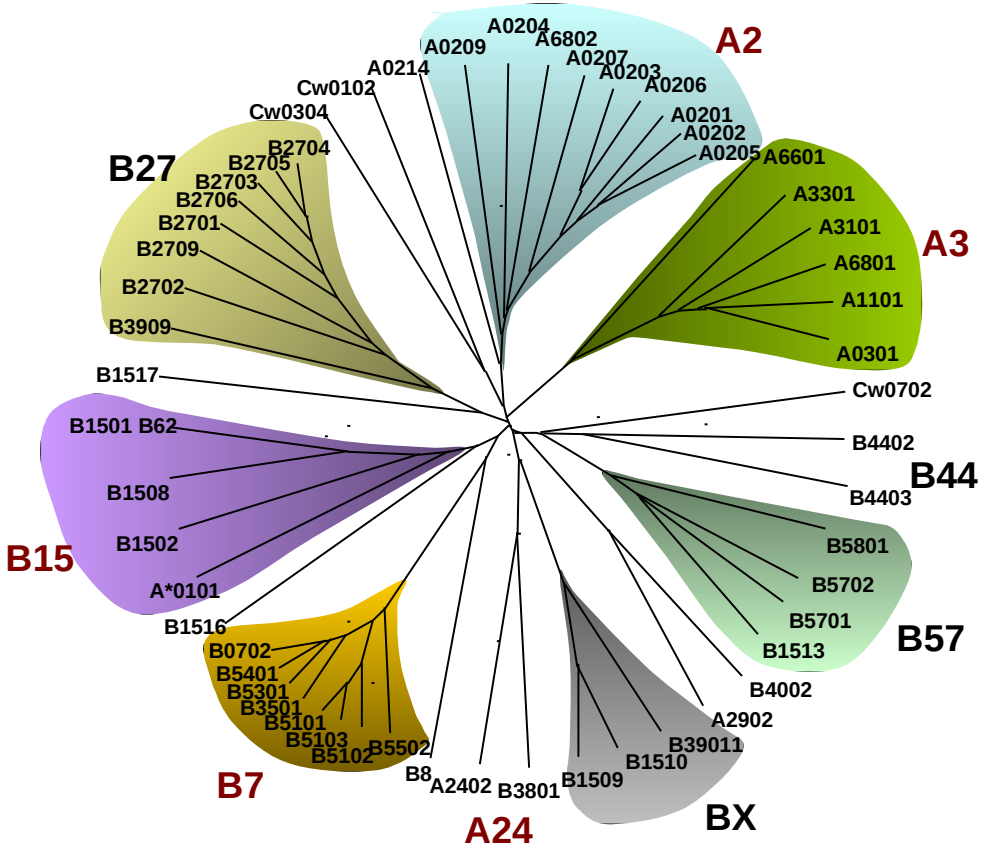
  95 CD8 T cell epitopes

  85 CD4 T cell epitopes

Peptide binding repertoires of MHC molecules are largely overlapping



Reche and Reinherz, *Nucleic Acids Res.2005, 33:W138*



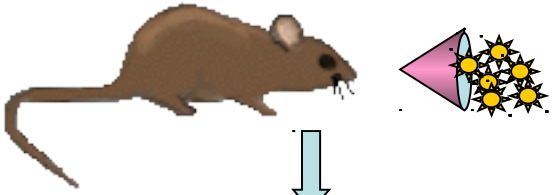
Supertype	Alleles	Blacks	Caucasians	Hispanics	*N.A.Natives	Asians	Promiscuous peptides
A2	A*0201-7, A*6802	43.7%	49.9%	51.8%	52.4%	44.7%	0.8%
A3	A*0301, A*1101, A*3101, A*3301, A*6801, A*6601	35.4%	46.9%	41.5%	40.7%	47.9%	0.6%
B7	B*0702, B*3501, B*5101-02, B*5301, B*5401	45.9%	42.2%	40.5%	52.0%	31.3%	1.2%
B15	A*0101, B*1501_B62, B1502	13.06%	37.80%	16.75%	27.26%	21.04%	0.6%
A24	A*2402, B*3801	15.5%	17.28%	25.85%	41.94%	35.0%	0.5%

Prediction of promiscuous peptides binding to supertypes facilitate development of broadly protective vaccine

T cell epitope discovery efficiency

Genome-wide characterization of a viral cytotoxic T lymphocyte epitope repertoire.
Zhong W, Reche PA, Lai CC, Reinhold B, Reinherz EL. J Biol Chem. 2003, 278:45135

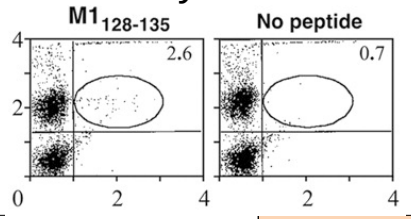
B6 mice (Kb, Db)



Influenza A Virus (4500 aa)

BAL cells + peptides

Intra-cellular cytokine staining (FACS)



Peptides	Kb + Db
Predicted	101
Bind	82 (82 %)
CTL	12 (13 %)

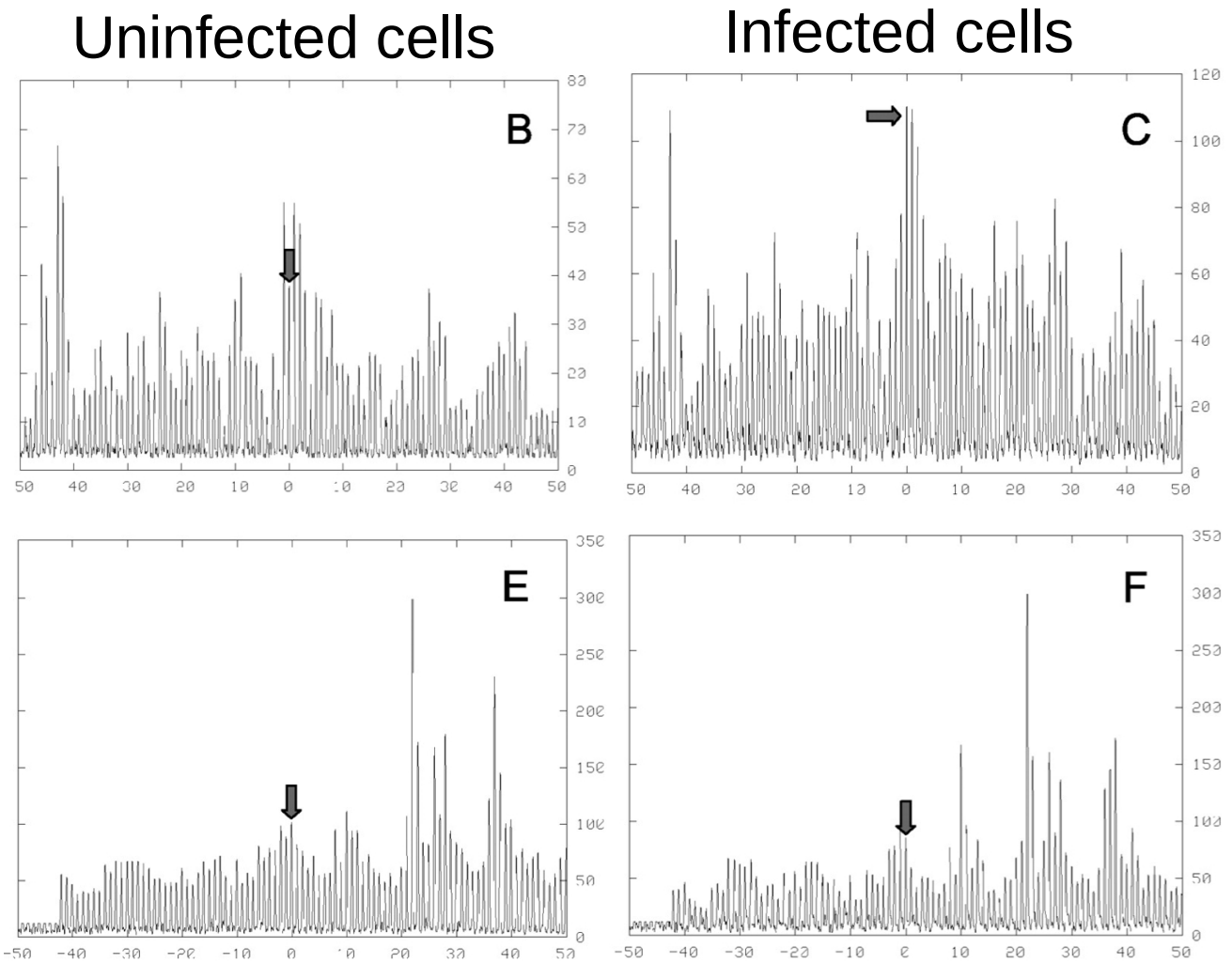
~ 12 % of predicted T cell epitopes are immunogenic

Failure in antigen processing impairs immunogenicity of peptides binding with high affinity to MHC

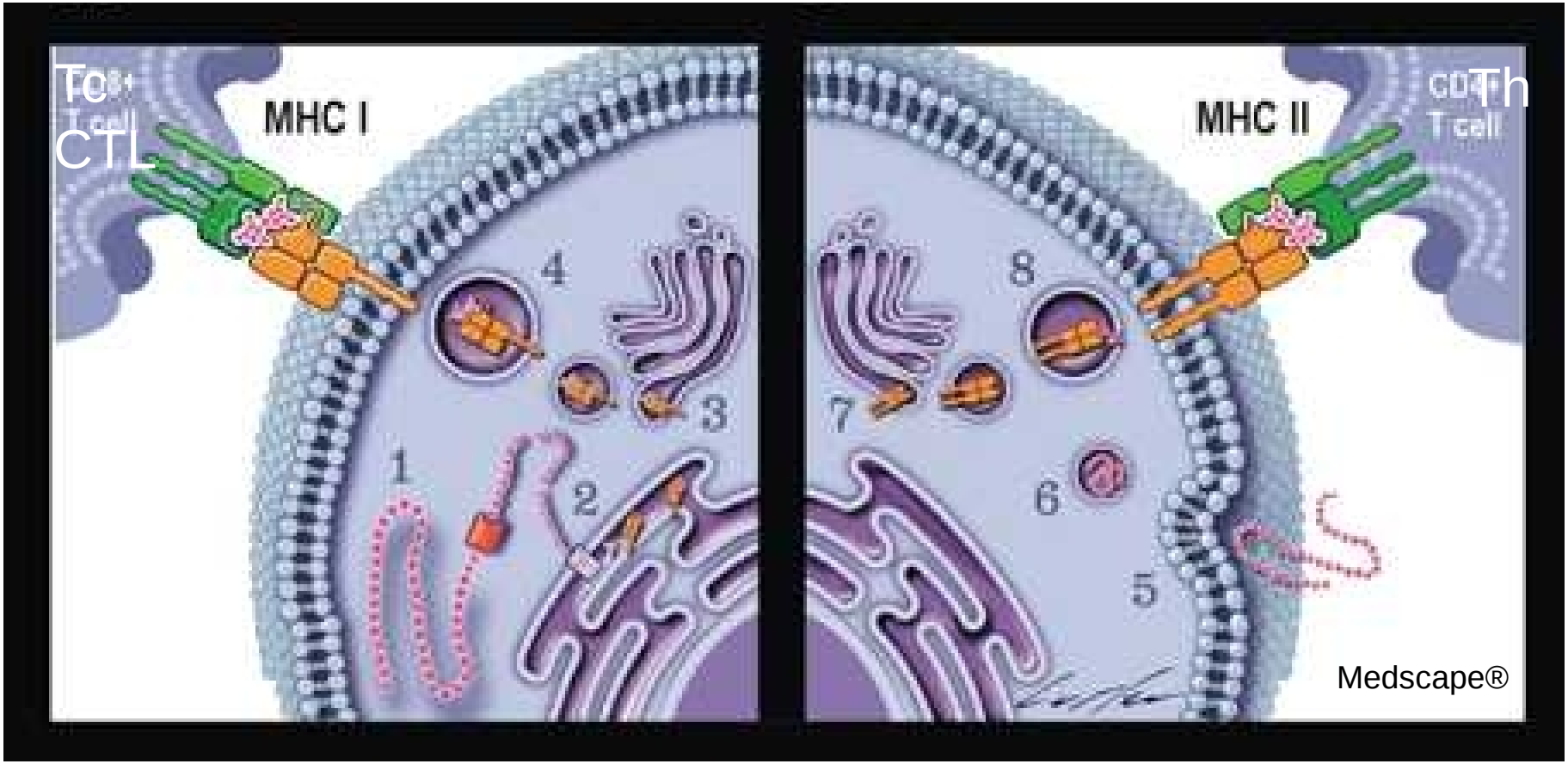
MS^n analysis of peptides eluted from MHC molecules of EL4 cells infected with IAV

Immunogenic
PA224-233
SSLENFRAYV

Non-immunogenic
HA332-340
TGLRNIPSI



Class I and II antigen processing pathways

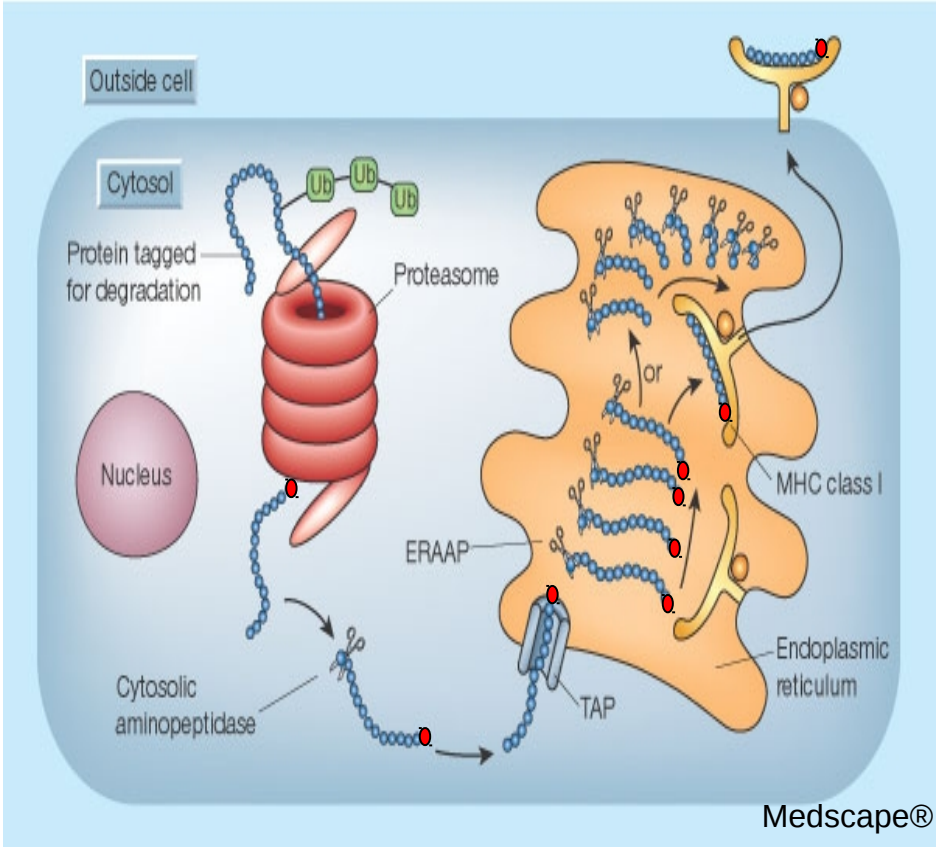


Class I: Antigens degraded in the cytosol and loaded in ER (Models available for proteasome cleavage and TAP binding)

Class II: Endocytosed antigens degraded and loaded in endosomal compartments (No models available)

C-terminal end of peptides presented by MHCI molecules is the result of protesomal cleavage

The C-terminal end of peptides presented by MHCI molecules reflects the cleavage by the protesaome. Cleavage site can be obtained by mapping peptides onto the source proteins.



Peptide presented by MHC I

S-D-V-D-V-S-S-V-L

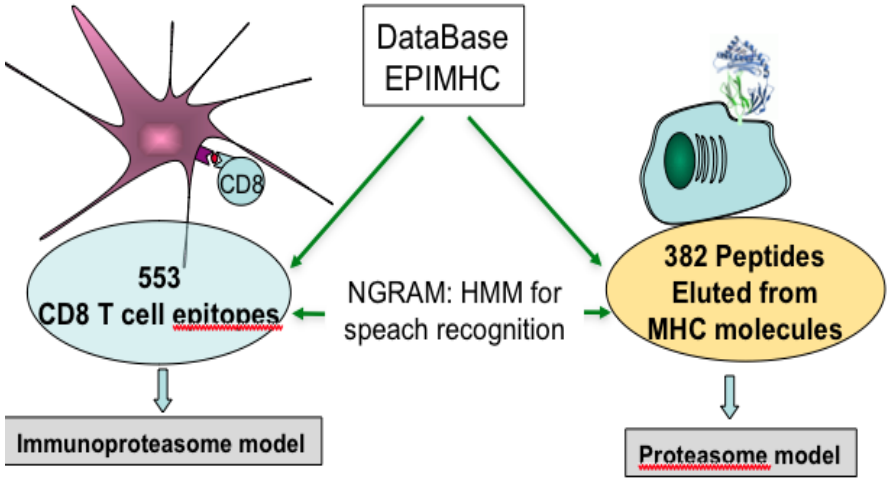
3HN-A-C-V-**S-D-V-D-V-S-S-V-L**-S-V-S-H-COO-

P1 **P1'**

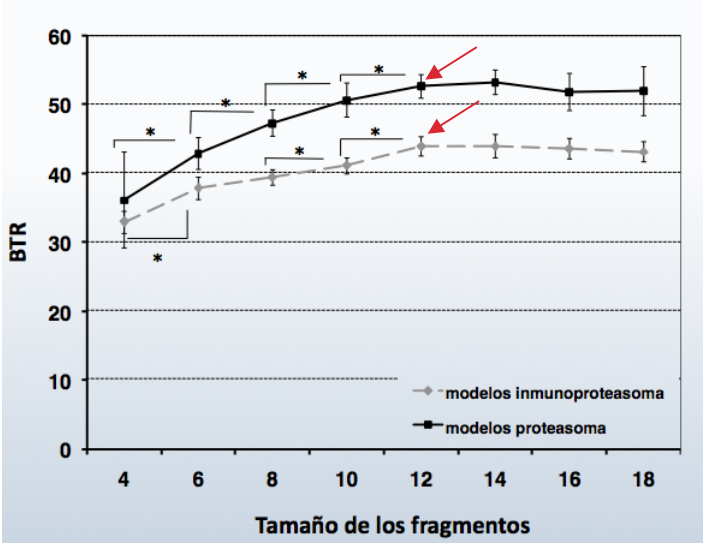
Cleavage site

Computational modeling of cleavage by the proteasome and the immunoproteasome

A. Datasets for model building



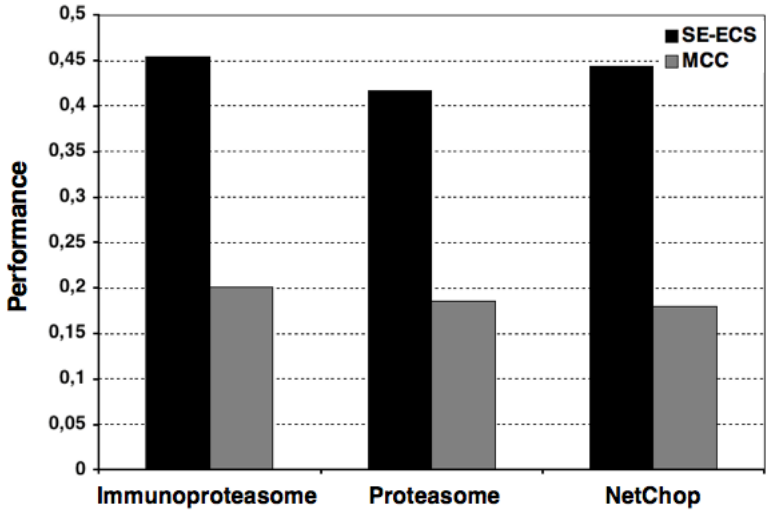
B. Model performance with regard to fragment size



Data For Model building:

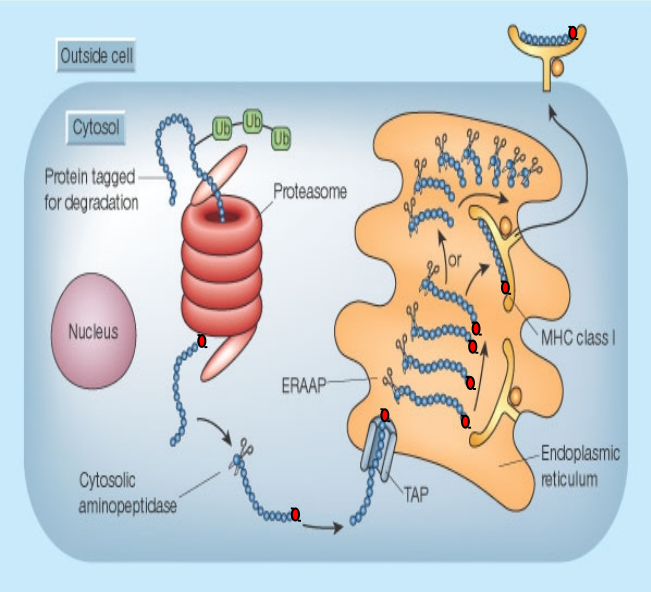
	MHC I PEPTIDE	FLANCKING RESIDUES
Mode 4 →	I Y	K R
Mode 6 →	E I Y	K R W
Mode 10 →	V G E I Y	K R W I I
Mode 12 →	P V G E I Y	K R W I I L
Mode 18 →	P P I P V G E I Y	K R W I I L G L N

C. Performance on HIV-1 epitope set



Computational modeling of peptide transport to ER by TAP

A) TAP transport peptides to ER



B) Model building

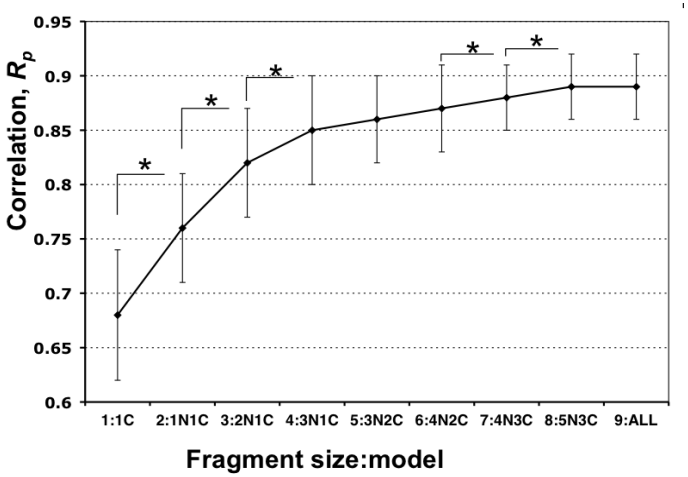
QB TAP DATASET (613 9 mer)

```
PLAKAAAAV   3.63
.
RRFGDKLNF   -2.0
```

SVR

TAPREG

C) Performance of TAP models

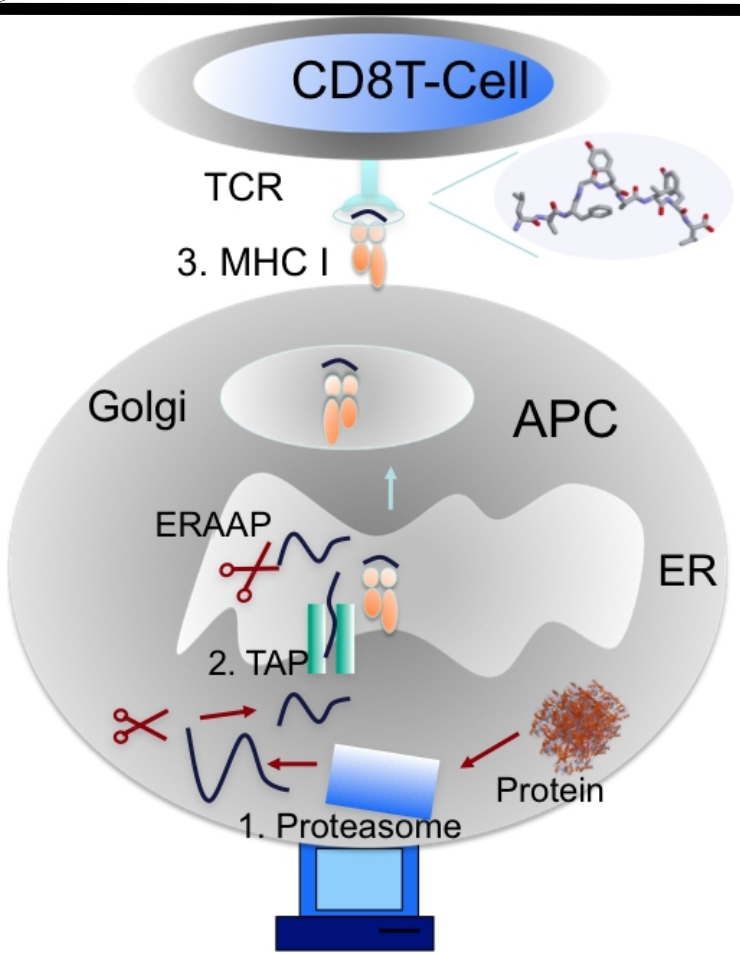


D) Algorithm to estimate TAP transport of MHCI-restricted ligands

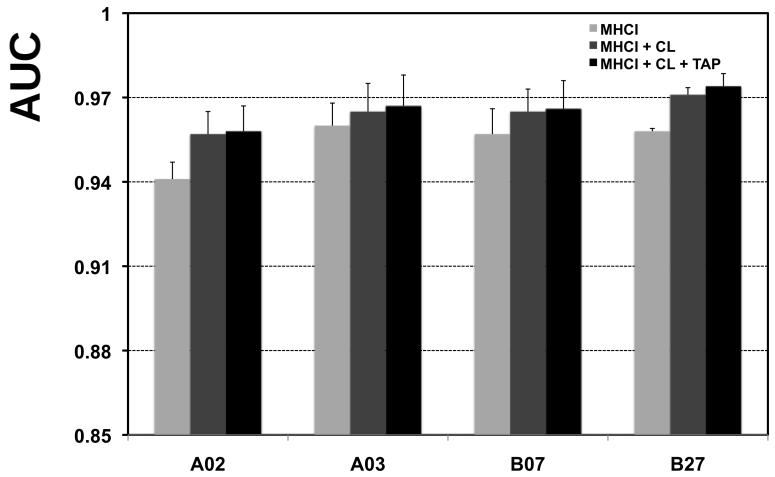
PEPTIDE	TAPREG AFFINITY (EC 50)
ACPDDDSFV	4000
SACPDDDSFV	2500
VSACPDDDSFV	1000
IVSACPDDDSFV	300*
DIVSACPDDDSFV	600
ADIVSACPDDDSFV	800
LADIVSACPDDDSFV	2000
GLADIVSACPDDDSFV	3000

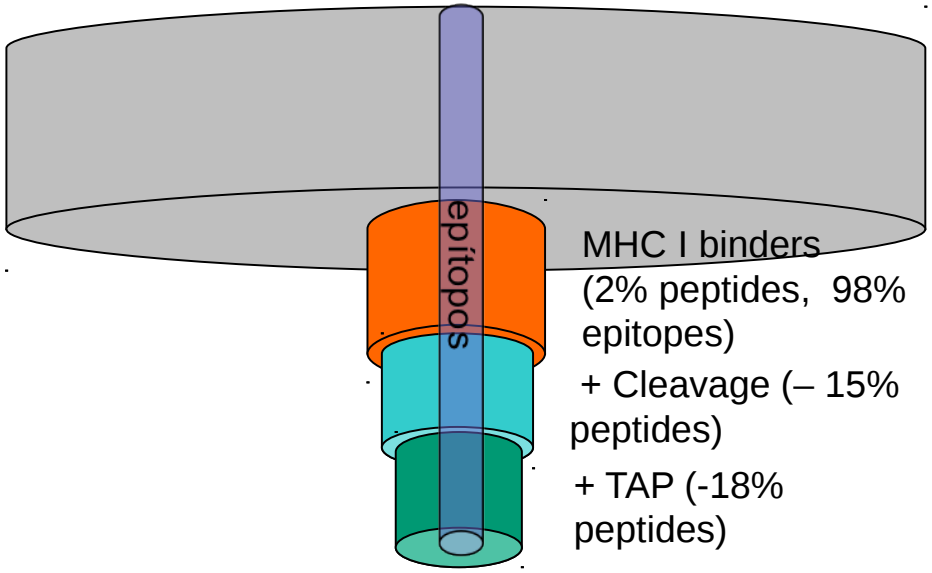
Best TAP affinity of all possible elongated peptides up to 16

Integrated CD8 T cell epitope prediction



B) CD8 T cell epitope discrimination





1. Proteasome cleavage: Diez-Rivero, .. & Reche (2010) Computational analysis and modeling of cleavage by the immunoproteasome and the constitutive proteasome, *BMC Bioinformatics*, **11**:479.

1. TAP transport: Diez-Rivero, .. & Reche (2010) Quantitative modeling of peptide binding to TAP using support vector machine.Proteins, 78:63

T cell Epitope prediction is not a precise science

Antigen

Epitope prediction

Epitope

Potential candidates

We cannot predict antigen processing accurately

Dealing with low T cell epitope discovery rate for epitope vaccine design

- Rely in technology: peptide synthesis is getting cheaper and there are high-throughput assays capable of evaluating thousands of peptides

- Prioritize antigens: e.g. highly expressed proteins are prime source of T cell epitopes (Diez-Rivero & Reche, 2012, PLoS One, 7: e43674)

- **Use legacy experimentation to overcome limitations with antigen processing predictions**

Legacy experimentation & epitope vaccine design

- There are thousands of pathogen specific T cell epitopes known to be processed that are available on specialized databases

Epitope	Source Organism	Source Protein	Restriction	PMID
LPFDKTTIM	Influenza A	Nucleocapsid	B35:01	14764717

- 1. We can predict binding to other MHC molecules and compute epitope vaccine coverage

    - T cell epitope vaccine coverage: Percentage of people capable of eliciting a response ~ cumulative phenotypic frequency of HLA molecules presenting the epitope

- 2. Select epitopes reaching a determined coverage

T cell epitope based vaccine for HIV1

HIV-1 DATABASE

1. Select CD8 T cell epitopes processed in vivo

195 CD8 T cell epitopes

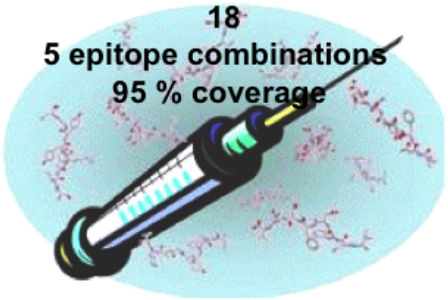
2. Discard epitopes with variables sites (H ≥ 1)

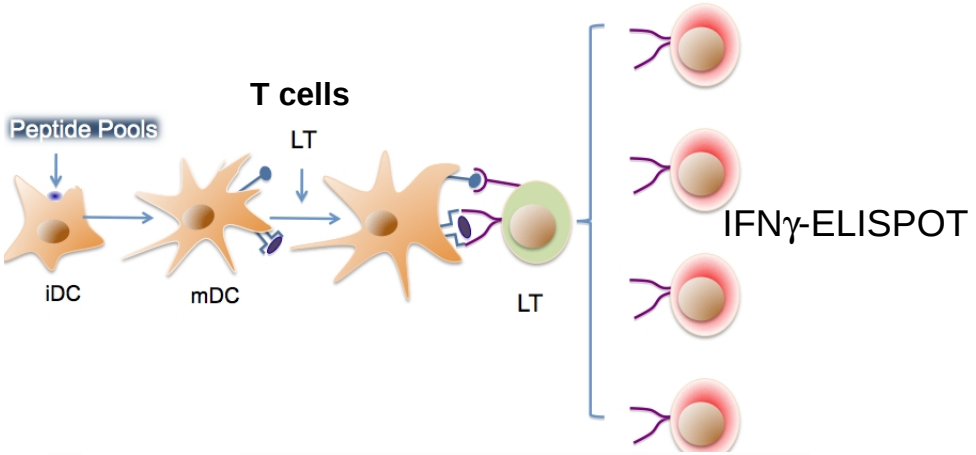
25 CD8 T cell epitopes

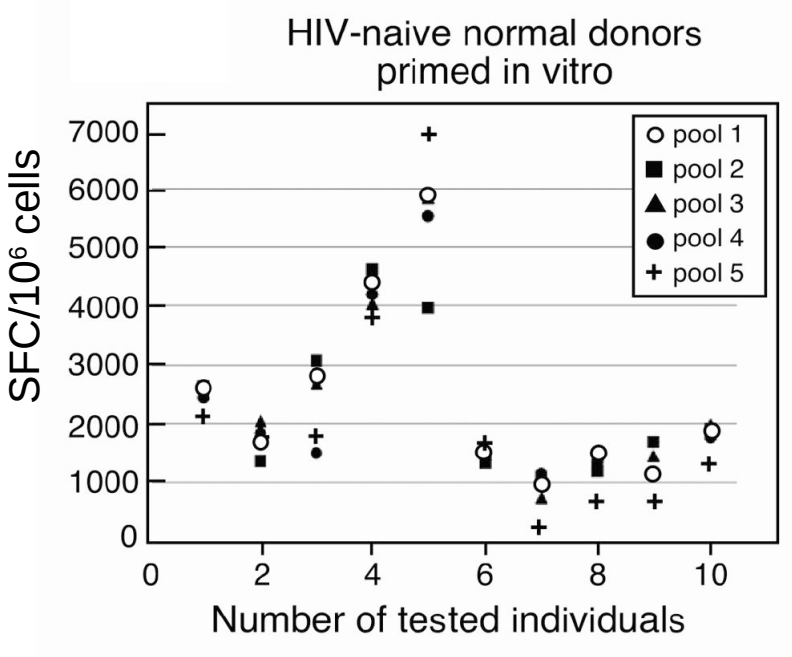
3. Prediction of HLA I binding profiles

PEPTIDE	BINDING PROFILE	Coverage
TLFCASDAK	A0301 A1101 A3101 A3301 A6801	0.32
LTFGWCFKL	A0201	0.18
...	...	...
WYIKIFIMI	A0203 A0206 A2402	0.05

4. Algorithm select epitopes reaching 95 % coverage



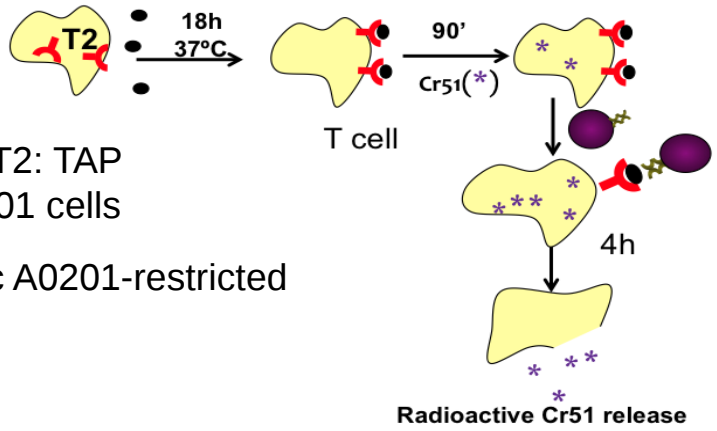




1: Reche PA et al. Med Immunol. 2006 May 18;5:1. PMID: 16674822

Primed T cells with DC loaded peptides are functional

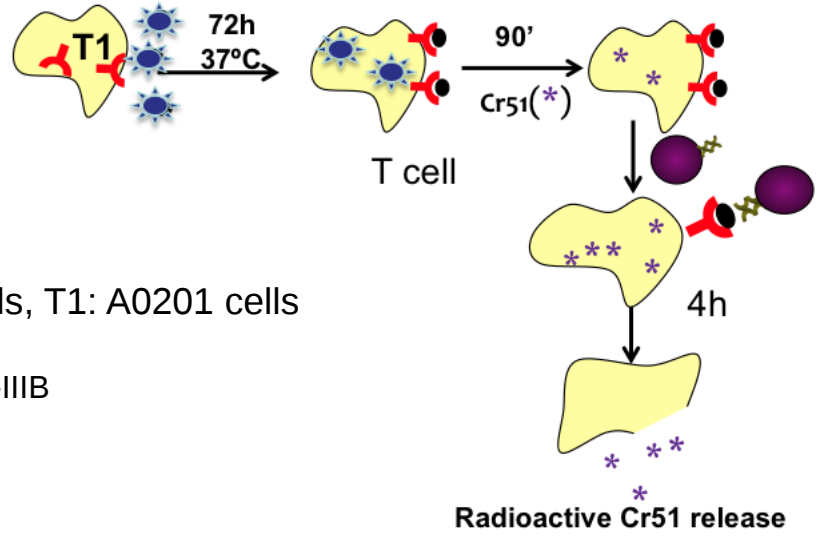
A



Target cells, T2: TAP deficient A0201 cells

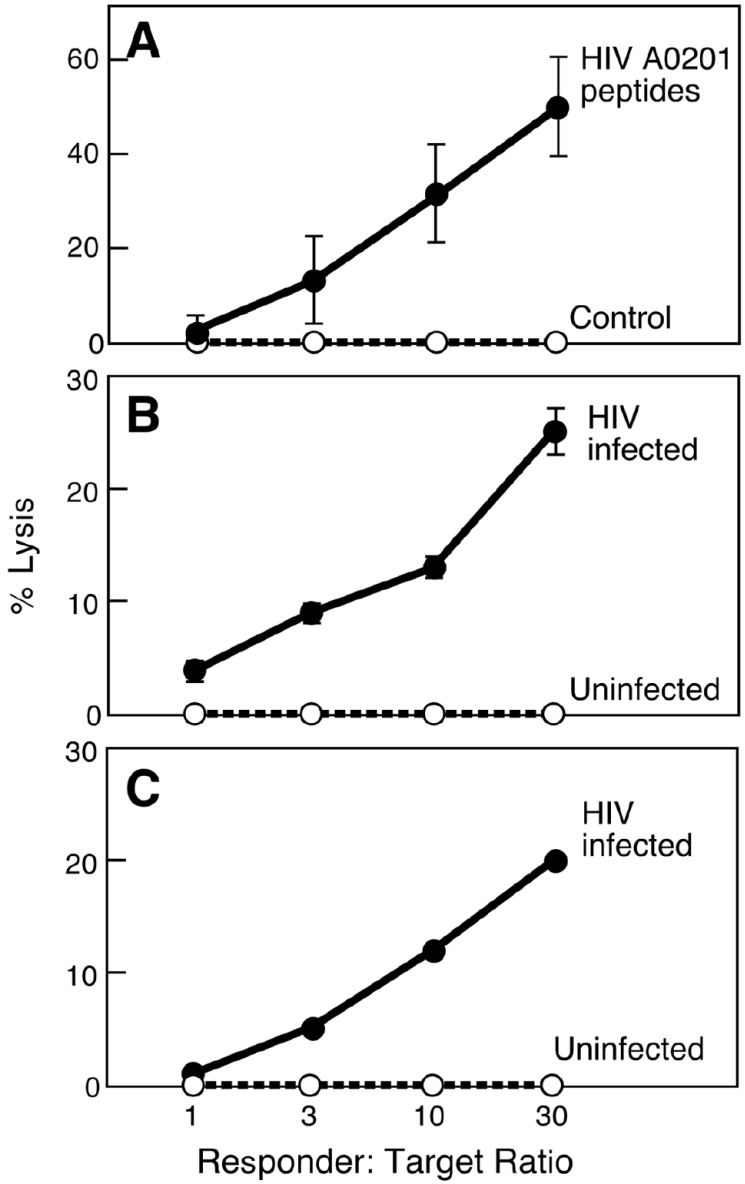
- HIV-specific A0201-restricted epitopes

B,C



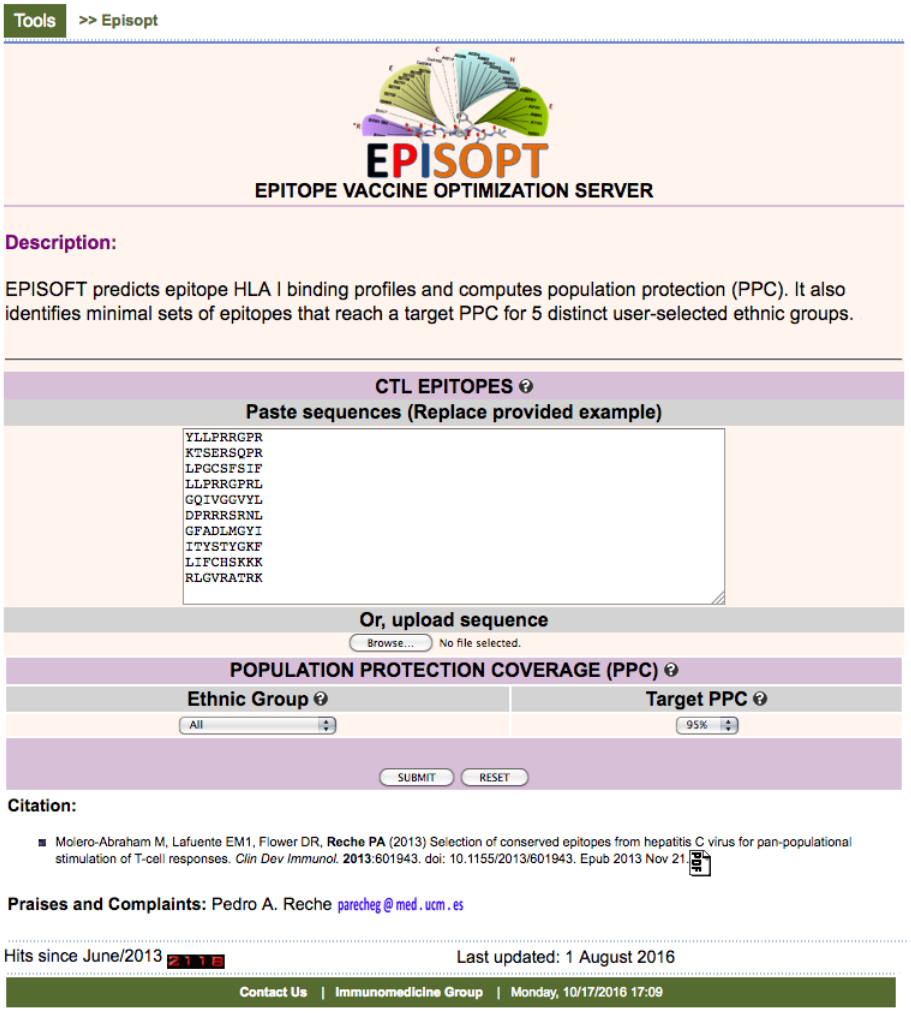
Target cells, T1: A0201 cells

HIV1-IIIB



EPISOPT: T cell Epitope vaccine optimization

http://bio.med.ucm.es/episopt.html



EPISOPT RESULTS

HLAI BINDING PROFILES AND POPULATION PROTECTION COVERAGE

PEPTIDE	PPC	HLA I binding profile
LLPRRGPRL	0.3578	A0201 A0203 A0205 B0801 C0102
KTSERSQPR	0.3218	A0301 A1101 A3101 A3301 A6801
LIFCHSKKK	0.3218	A0301 A1101 A3101 A3301 A6801
LPGCSFSIF	0.2782	B0702 B1502 B1508 B3501 B5301 B5401
DPRRRSRNL	0.2180	B0702 B0801 B4402 B5101 B5102 B5103
ITYSTYGKF	0.1486	A2402 B1516 B5801
GFADLMGYI	0.0527	A0202 A0203 A6802 B3801
RLGVRATRK	0.0193	A0301
YLLPRRGPR	0.0140	A3301 A6601 A6801
GQIVGGVYL	0.0040	A0206 A0214 B1510 B4002

MINIMUN TOP EPITOPE-HLAI SET

Coverage: 97.14	Epitopes: 5	HLA I-restiction alleles: 22
Sequences:	LLPRRGPRL GFADLMGYI KTSERSQPR/LIFCHSKKK ITYSTYGKF LPGCSFSIF 5	
HLA I-restriction alleles:	A0201 A0202 A0203 A0205 A0301 A1101 A2402 A3101 A3301 A6801 A6802 B0702 B0801 B1502 B1508 B1516 B3501 B3801 B5301 B5401 B5801 C0102 22	

1: Molero-Abraham M, Lafuente EM, Flower DR, Reche PA. Selection of conserved epitopes from hepatitis C virus for pan-populational stimulation of T-cell responses. Clin Dev Immunol. 2013; PMID: 24348677

2: Sheikh QM, Gatherer D, Reche PA, Flower DR. Towards the knowledge-based design of universal influenza epitope ensemble vaccines. Bioinformatics. 2016 PMID: 27402904.

Conclusions

1. Epitope-based vaccines are next step in vaccine formulations
2. We have immunoinformatics tools to predict and select epitopes of interest that save time and resources
3. Epitope prediction is not a precise science: B cell epitope prediction is somewhat less useful than T cell epitope prediction
    - Most B cell epitopes are conformational
    - We do not have methods to convert 3D B cell epitopes in 1D B cell epitopes
4. The Achilles heal in T cell epitope prediction is antigen processing: we need to develop new and better tools to predict antigen processing
5. Combination of legacy experimentation and bioinformatics can speed epitope vaccine development

Спасибо за внимание !