

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**Peptídeos Virais Imunogênicos como Determinantes de Reatividade
Cruzada no Sistema Imune**

Gustavo Fioravanti Vieira

Tese submetida ao Programa
de Pós-Graduação em
Genética e Biologia Molecular
da UFRGS como requisito
parcial para a obtenção do
grau de Doutor em Ciências.

Orientador:

Prof Dr José Artur Bogo Chies

Porto Alegre, Agosto de 2008

Pesquisa desenvolvida no Núcleo de Bioinformática do Laboratório de Imunogenética da UFRGS e financiada com recursos da Capes - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

“Dedico essa tese aos meus pais por terem me ensinado que a melhor profissão é aquela que nos traz felicidade e que a vida somente faz sentido quando lutamos por nossos sonhos.”

Gustavo *et al* 2008

Como todo projeto, torna-se impossível sua realização sem o apoio (sentimental, fraternal, amoroso, técnico, intelectual, financeiro, futebolístico, musical, festivo. etc) de incontáveis colaboradores, cada um com suas expertises específicas. Todas elas fundamentais para que ele fosse concluído com sucesso.

De acordo com as regras do nosso programa de pós-graduação, em referências com mais de dez autores apenas os dez primeiros devem ser listados, seguidos por *et al*. Aqui, serei obrigado a citar mais de dez colaboradores, no entanto, aqueles que não encontrarem seu nome especificado (a maioria provavelmente por problemas de memória acarretados pela síndrome pré-defesa de tese), considerem-se no grupo do *et al*.

Aos meus pais que tanto amo, os quais tenho a honra de chamar de amigos, por terem dedicado suas vidas para que eu pudesse construir a minha. Obrigado por serem meu porto seguro, por me apoiarem e, além disso, por serem cúmplices em todas as minhas maluquices. Vocês são os “caras”!

Ao Zéca (Prof. Dr. José Artur Bogo Chies) que viu uma oportunidade onde todo mundo via um oportunista (inclusive eu, hehehe). Espero que não o tenha decepcionado e muito obrigado por acreditar em mim!

À Maricota, por ter devolvido a paz e o equilíbrio à minha vida. Teu companheirismo e o teu jeito sereno foram fundamentais nessa reta final tão conturbada. Fostes minha doce calma em meio à tempestade. Te amo muito!

À minha família, por serem os alicerces da minha vida. E olha que eu tenho vários. Entre tios, tias, primos e primas vai pra lá de sessenta entre os Fioravanti e os Vieira (dá pra construir um condomínio inteiro com esses alicerces). Um agradecimento especial à minha avózinha Neuza pelo carinho e pelas boas risadas que dou quando nos encontramos.

Aos meu co-orientados (Dinler e Maurício) por ajudarem a transformar minhas idéias em realizações. Destaque para a criação do nosso Núcleo de Bioinformática e para a formação do grupo musical “Los Epitopos”.

Aos amigos, parceiros de toda hora e de todas indíadas (desculpem, nessa hora vai ser difícil lembrar de todos, vou tentar dividir em categorias, hehehe).

Aos colegas da Pós da turma de 2004 (à Tati G pela grande amizade e pelas boas vindas ao lab, ao Rafa Diehl, à Júlia, ao Graxa, e a toda aquela turma das aulas de Molecular, Estatística e Evolução). Aos amigos dos laboratórios de Drosófila (Fabiano, Ronaldo, Marícia) da Hemostasia (Ana, Bandi) da Genética Humana (Verônica, Fabi, Vinícius) e da Evolução (Ernesto, Lucas, Adriano) e aos parcerias do departamento todo.

Aos colegas e professores do Meu lab (Imunopueblo). Especialmente ao Dani Garcia, à Paula Rohr, ao Andrés Cañedo, à Déia Wick, ao Felipe Narigudo (imunoagregado) à Lari, ao Tiago Dalberto, à Fernanda Malinsky, à Mel, ao Tiago Veit, à Elvira, à Pati e a todos os outros integrantes desse grupo tão agradável de trabalhar. Vocês todos tornam a rotina de trabalho mais leve.

Aos colegas do futebol, meu passatempo preferido e meu único vício. Menção especial ao Guano (Juliano, Lucas, Nariz, Dani) e aos multicampeões amarelinhos (Nico, Maurício, Alemão, Ezequiel, Merenda). Foi uma honra defender suas cores.

Aos colegas de caminhada, aqueles que surgem nos momentos cruciais da nossa vida. Glauquito, valeu por ser meu colega de apartamento, de composições, de discursos políticos e filosóficos e por todos nosso papos sobre assuntos aleatórios. Ao Maurício Almerão, por ter sido um irmão num momento que eu mais precisava de um amigo.

Ao Elmo e à Ellen por serem os “grandes facilitadores” do nosso trabalho e, principalmente, pelo convívio alegre no dia-a-dia do departamento.

À meu Deus, luz da minha vida e fonte inesgotável de inspiração para os meus projetos. Obrigado por todas as dádivas e pela força para enfrentar meus desafios. Que eu consiga retribuir com o meu trabalho todas as bençãos que tenho recebido.

“Uma pessoa com uma crença é um poder social igual a noventa e nove que possuem apenas interesses.”

John Stuart Mill (1861)

SUMÁRIO

LISTA DE ABREVIATURAS	8
RESUMO	9
ABSTRACT	11
CAPÍTULO I - INTRODUÇÃO	13
Peptídeos Virais Imunogênicos como Determinantes de Reatividade Cruzada no Sistema Imune	13
Reatividade Cruzada	19
Reconhecimento Promíscuo pelo Receptor de Célula T (TCR)	28
Como o Sistema Imune Reconhece os Vírus?	34
Implicações e Aplicações do Mimetismo Molecular	40
OBJETIVOS	45
CAPÍTULO II - IMMUNODOMINANT VIRAL PEPTIDES AS DETERMINANTS OF CROSS- REACTIVITY IN THE IMMUNE SYSTEM - CAN WE DEVELOP WIDE SPECTRUM VIRAL VACCINES?	46
CAPÍTULO III - CONSTRAINS ON THE VARIATION OF VIRAL PROTEINS AND VACCINE DEVELOPMENT	67
CAPÍTULO IV - DISCUSSÃO	108
REFERÊNCIAS	120
ANEXO I	128
ANEXO II	129

LISTA DE ABREVIATURAS

APC:	célula apresentadora de antígeno
APL:	ligante peptídico alterado
CCR7:	receptor de quimiocina 7
CDRs:	regiões determinantes de complementaridade
CTL:	linfócito T citotóxico
EBV:	Epstein-Barr vírus
ERV:	retrovírus endógeno
HBV:	vírus da hepatite B
HCMV:	citomegalovírus humano
HCV:	vírus da hepatite C
HIV:	vírus da imunodeficiência humana
HIV-1:	vírus da imunodeficiência humana tipo I
HIV-2:	vírus da imunodeficiência humana tipo II
HLA:	antígeno leucocitário humano - MHC de classe I humano
HPV:	papilomavírus humano
HPV16:	papilomavírus humano tipo 16
HTN:	haantan vírus
IL-2:	interleucina 2
IV:	vírus influenza
IVA:	vírus influenza A
JEV:	vírus da encefalite japonesa
LCMV:	vírus da coriomeningite linfocítica
LES:	Lúpus Eritematoso Sistêmico
MCMV:	citomegalovírus murino
MHC-I:	complexo principal de histocompatibilidade de classe I
MHC-II:	complexo principal de histocompatibilidade de classe II
MHC-pep:	complexo MHC-peptídeo
PKR:	proteína quinase R
PUU:	puumala vírus
PV:	vírus pichinde
SNV:	sin nombre vírus
TAP:	proteína do transporte associada ao processamento de antígenos
TCR:	receptor de célula T
TLR:	receptor do tipo "Toll"
VV:	vírus da vaccinia

RESUMO

A identificação de epitopos e motivos virais para serem utilizados na imunização de humanos e animais é um objetivo importante e essencial em pesquisa imunológica. No momento, com novas ferramentas de bioinformática, diferentes abordagens são possíveis. A importância da bioinformática reside na possibilidade de trabalhar com grandes quantidades de dados de forma que, várias etapas experimentais do desenvolvimento de vacinas podem ser abreviadas. Quando buscamos por motivos virais (objetivando a vacinação) é necessário estar ciente de todos os passos envolvidos na seleção e apresentação de peptídeos ao sistema imune. Devemos considerar que muitos peptídeos podem ser gerados a partir de uma única proteína, mas apenas uma fração destes peptídeos é realmente apresentada ao sistema imune, pois os peptídeos devem ser capazes de atravessar diferentes “bottlenecks” e apenas aqueles apresentando características específicas serão capazes de estimular o sistema imune. As abordagens apresentadas nesta tese incluem o estabelecimento de um banco de dados de epitopos virais descritos na literatura e a comparação desses epitopos buscando identificar características similares entre eles. Os epitopos selecionados foram classificados de acordo com suas propriedades físico-químicas (polaridade e carga dos grupos R de seus aminoácidos). Das 69 sequências de epitopos incluídas em nossa base de dados, 31 (44,93%) apresentaram, em sítios potenciais de ancoragem ao MHC, aminoácidos com resíduos não-polares. A partir desses resultados é possível inferir o seguinte motivo consenso: X [AGPVLIM] X(6) [AGPVLIM] para epitopos virais. As sequências virais foram então comparadas àquelas de outras proteínas buscando verificar se elas são exclusivamente representadas em vírus: 1) primeiro os epitopos foram comparados a todas as sequências depositadas no GenBank (independente da origem); 2) a seguir, as comparações foram direcionadas a sequências de origem humana. A segunda abordagem foi usada para verificar o potencial de indução de reações autoimunes. As sequências de saída foram classificadas de acordo com seu organismo de origem. Das 31 sequências

alinhas de acordo com a similaridade dos resíduos de ancoragem, 29 (93,54%) apresentaram similaridade significativa (estabelecida como 80% ou mais) com outras sequências virais. Destas, 12 (38,71%) apresentavam similaridade apenas com outras sequências de origem viral, nove (29,03%) apresentavam similaridade com sequências de origem bacteriana e duas (6,45%) apresentavam similaridade com sequências humanas, sugerindo que a grande maioria dos epitopos virais pode ser utilizada no desenvolvimento de vacinas. A habilidade dos epitopos serem gerados pela via de processamento de antígenos foi também testada. Uma parte das proteínas citosólicas sofre o processo de ubiquitinação que as dirige para o complexo enzimático proteolítico, denominado proteossomo. Do total de epitopos, cinquenta (73,53%) apresentaram uma sequência que permitia um corte exato na extremidade carbóxi terminal. Este número alcançou 86,67% (26 epitopos), quando restringimos a análise aos epitopos apresentando os resíduos de ancoragem compartilhados, sugerindo que a maioria dos epitopos apresentava os requerimentos clássicos para o processamento antigênico. Das estruturas do Protein Data Bank e de dados de modelagem, foi possível observar que os sítios de clivagem preditos na região amino terminal dos epitopos eram estruturalmente relacionados a alças na estrutura da proteína original (66,7%). Estes dados sugerem que há uma clivagem preferencial em alças ($\chi^2=6.09$ $p=0.047$). O banco de dados de ligantes peptídicos do EpiJen foi utilizado para avaliar a capacidade dos epitopos em serem carregados pelo complexo da TAP. A partir dessa comparação observamos que o motivo predito estava mais representado do que todas outras possíveis sequências entre as saídas, sugerindo novamente que estas características são necessárias à seleção e apresentação de epitopos. Concluindo, sugerimos que é possível identificar padrões entre epitopos derivados de vírus e que a predição de motivos virais conservados pode ser aplicada ao desenvolvimento de vacinas. Além disso, considerando a existência de reatividade cruzada, sugerimos que é possível imunizar contra uma quantidade consideravelmente grande de alvos virais utilizando um número de epitopos reduzido. Estudos sobre os aspectos e características dos epitopos virais são o primeiro passo rumo a uma nova geração de vacinas.

ABSTRACT

The identification of epitopes and viral motifs to be used in the immunization of both humans and other animals is an important and essential objective in immunology research. At present, with the new tools of the bioinformatics different approaches are possible. The importance of the bioinformatics is exemplified by its capacity to handle a large amount of data in order to bypass several methodological steps in vaccine development. When searching for conserved viral motifs it is necessary to be aware of all the steps involved in peptide selection and presentation. In this way, we should consider that many different peptides can be generated from a given protein, but only a fraction of these peptides will actually be presented to the immune system. The approaches presented in this thesis include the establishment of a viral epitope databank from sequences described in the literature and the comparison of these epitopes in order to identify similar features among them. The selected epitopes were classified according to their physicochemical properties (i.e. polarity and charge of their amino acid-R groups). From the 69 sequences of epitopes included in our database, 31 (44.93%) presented, in potential MHC anchor sites, amino acids with non-polar residues. From this, it is possible to infer the following consensus motif: X [AGPVLIM] X(6) [AGPVLIM]. The viral sequences were then compared to those of other proteins in order to verify if they are exclusively represented in viruses: 1) first, the epitopes were compared to all sequences stored in GenBank (disregarding their origin); 2) then, the comparisons were directed to the sequences from human origin. The output sequences were classified according to the organism of origin. From the 31 sequences aligned according to their anchor residues, 29 (93.54%) presented significant similarity (established as 80% or above) with other viral sequences. From these, 12 (38.71%) presented similarity only with other sequences from viral origin, nine (29.03%) presented similarity with sequences from bacterial origin and two (6.45%) presented significant similarity to human sequences, suggesting that a great majority of these viral epitopes could be used in vaccine development. The ability of the epitopes to be generated by antigen processing pathway was also

tested. A fraction of all cytosolic proteins suffers the ubiquitination process that directs them to the proteolytic enzymatic complex, called proteasome. From the whole databank, fifty epitopes (73.53%) presented a sequence that allowed a precise cut at the carboxy terminal region. This number reached 86.67% (26 epitopes) when we restricted this analysis to the epitopes presenting the shared anchor residues, suggesting that the majority of the epitopes presented the classical requirements to antigenic processing. From the Protein Data Bank structures and from the modelling data, we could observe that the predicted cleavage sites on the amino-terminal region of the epitopes were structurally related to loops on the structure of the original protein (66.7%), suggesting a preferential cleavage at loops ($\chi^2=6.09$ $p=0.047$). The TAP ligands peptide database of EpiJen was used in order to evaluate the epitopes capacity to be carried by the TAP complex. We were able to observe that our predicted motif was present more frequently than every other possible sequence among the outputs, again suggesting that this feature is necessary to epitope selection and presentation. In conclusion, we suggest that it is possible to identify patterns among virus-derived epitopes and that the prediction of viral conserved motifs would allow the development of vaccines. Also, considering the existence of cross reactivity, we suggest that it will be possible to cover a considerably large amount of targets using a limited number of viral peptides. Equally important is the immunogenicity of viral peptides, since only a fragment and not the whole viral particle will challenge the immune system, therefore reducing risks of undesired immune responses. Studies on viral epitopes features and characteristics are the first step towards a new generation of vaccines.

“Não me inspiro nas citações; valho-me delas para corroborar o que digo e que não sei tão bem expressar, ou por insuficiência da língua ou por fraqueza do intelecto.”

Montaigne (1592)

CAPÍTULO I

Introdução



Peptídeos Virais Imunogênicos como Determinantes de Reatividade Cruzada no Sistema Imune

O sistema imune humano tem desenvolvido, ao longo de sua evolução, diversas formas de eliminar organismos patógenos. Em contrapartida, estes também buscam aprimorar sua evasão dos mecanismos de defesa. Alguns dos mecanismos de interação sistema imune:patógeno já se encontram em avançado estágio de esclarecimento. Em sistemas virais, entretanto, o assunto não está totalmente resolvido, principalmente devido ao alto poder de mutabilidade destes organismos, o que pressupõe a necessidade de um sistema de defesa com certo grau de especialização e complexidade direcionado contra esses organismos.

Muitas células do sistema imune estão envolvidas nos mecanismos voltados para a erradicação de infecções virais. Daremos atenção especial à interação entre as células que expressam MHC de classe I (potencialmente todas as células nucleadas) e os linfócitos T citotóxicos (CTLs) CD8+, embora se saiba da importância dos linfócitos B na resposta antígeno-específica a infecções virais. Outro mecanismo que será também abordado, mesmo que com menor ênfase é o que envolve os linfócitos T auxiliares CD4+ que têm importante participação na estimulação da imunidade celular e humoral contra os microorganismos. O foco dessa discussão se localizará nos mecanismos envolvidos na resposta elicitada logo após a infecção, mais especificamente na resposta celular.

Após a entrada do vírus no hospedeiro, a qual pode se dar por diferentes vias (mucosas, sistema respiratório, digestivo ou diretamente na circulação

sanguínea), o sistema imune imediatamente se prepara para combater o invasor. Quando essa entrada se dá pela corrente sanguínea, os vírus são provavelmente encaminhados para o baço onde encontram um grande repertório de células B e T, além de células APCs (células apresentadoras de antígenos) direcionados à montagem de uma resposta imune adequada. Normalmente as APCs migram para tecidos periféricos onde aguardam ativação mediada por “sinais de perigo” ou “sinais de infecção”. Estes sinais podem ser desencadeados pela infecção da própria célula ou pela liberação de substâncias de células vizinhas que sofreram morte celular induzida por vírus. Outras formas de ativação são mediadas por citocinas liberadas pela interação de proteínas virais com receptores do tipo Toll (TLRs) ou ativação de proteína Quinase R (PKR) pela presença de RNA de dupla fita. Estas APCs ativadas que estão apresentando o antígeno infeccioso, começam a expressar o receptor de quimiocina CCR7, migram para os linfonodos, onde entrarão em contato com as células T e B necessárias para iniciar a resposta imune apropriada (Welsh *et al.*, 2004).

Como o foco do nosso trabalho se dá na resposta citotóxica a vírus, nosso interesse se volta para a forma como as células que expressam MHC-I sinalizam para as células T CD8+ a ocorrência de uma infecção viral. Além do contexto infeccioso, o único elemento que elas dispõem para tal tarefa são os peptídeos que elas carregam. Elas precisam “demonstrar” que alguns daqueles antígenos (epitopos) que elas estão portando são derivados de antígenos virais, lembrando-se que esses epitopos correspondem a uma ínfima amostra de todo o conjunto de proteínas que compõem uma partícula viral. Estima-se que em vaccinia vírus (VV), dos 175.000 peptídeos potenciais, apenas 75 chegarão a entrar em contato

e ser reconhecidos pelo receptor de célula T (TCR) específico, após passarem pelas diferentes fases do processamento antigênico (Yewdell, 2006). Dessa forma, temos que levar em conta quais os mecanismos que estão agindo concomitantemente para que ocorra a produção, e fundamentalmente, a seleção de tais alvos, para o reconhecimento pelos receptores das células.

Quando as células que expressam MHC de classe I são infectadas por vírus, estas começam a produzir as proteínas virais necessárias à montagem da partícula viral. Parte dessas proteínas é, no entanto, degradada proteoliticamente nos proteossomos. Também serão degradadas nos proteossomos proteínas citosólicas próprias, proteínas específicas produzidas em células tumorais, ou mesmo proteínas derivadas de bactérias intracelulares, que tenham sido ubiquitinadas. Esta degradação gera peptídeos com tamanho entre 6 e 30 aminoácidos. No entanto, para que estes peptídeos se liguem à fenda peptídica do MHC de classe I no retículo endoplasmático e migrem para a superfície celular existe uma restrição de tamanho e apenas aqueles com comprimento entre 8 e 12 aminoácidos serão capazes de tal ligação. Este ajuste de tamanho é realizado após esta primeira clivagem, pelo encurtamento da extremidade amino terminal por outras peptidases citosólicas (Levy *et al.*, 2002; Reits *et al.*, 2003). Essa restrição em relação ao tamanho dos peptídeos se deve principalmente à natureza da fenda peptídica do MHC de classe I que se apresenta fechada em ambas extremidades, ao contrário da encontrada no MHC de classe II que pode acomodar seqüências maiores (até 30 aa).

Imagine que determinada proteína viral possua 200 aminoácidos e que a proteólise gere, em média, peptídeos com 10 aminoácidos (lembrando que a

fissão não precisa, necessariamente, ocorrer a 10 aa da extremidade inicial, mas que pode ocorrer um corte randômico em qualquer parte da proteína gerando peptídeos de tamanhos variáveis). O número de possíveis seqüências formadas, neste caso, atingiria 191 seqüências. Obviamente devemos considerar que nem todas as seqüências potenciais serão geradas. O proteossomo irá gerar apenas alguns desses epitopos e nem todos conseguirão ser translocados pela proteína do transporte associada ao processamento de antígenos (TAP) para dentro do retículo endoplasmático (RE) (Yellen-Shaw *et al.*, 1997; Beismann-Driemeyer and Tampé, 2004). Além disso, dentre os epitopos selecionados nem todos conterão os resíduos necessários para a correta ligação à molécula de MHC no RE. Cálculos indicam que a freqüência de peptídeos capazes de se ligar com suficiente afinidade ao MHC de classe I é de aproximadamente 1% de todos os possíveis peptídeos de uma proteína (Istrail *et al.*, 2004). Portanto, o número de epitopos resultantes ao final do processo de geração de antígenos é sensivelmente menor do que o esperado inicialmente.

Considerando que o tamanho da proteína de nosso exemplo é relativamente pequeno, e considerando a idéia clássica monoespecífica em que cada célula T reconhece apenas um simples peptídeo no contexto do MHC, o repertório de células T deveria ser extremamente maior do que o tamanho estimado (entre 10^7 e 10^8 células T, em humanos) (Sospedra and Martin, 2006). Isso fica mais evidente quando se considera o número teórico de antígenos possíveis que podem ser apresentados no contexto do MHC, o qual atinge valores superiores a 10^{15} (considerando-se peptídeos com tamanhos entre 8 e 14 aminoácidos, incluindo epitopos restritos a MHC de classe I e II). Se cada célula T

fosse monoespecífica, calcula-se que o baço de um camundongo deveria ser centenas de vezes maior do que o tamanho real desse órgão linfóide para que o animal pudesse conter infecções de forma eficiente (Mason, 1998).

Por este e outros motivos, o mecanismo de reconhecimento de seqüências peptídicas virais pelo sistema imune deve ser extremamente adaptável, ou seja, ter uma certa plasticidade, porém sem perder a especificidade, permitindo que o repertório de células imunes seja selecionado em função das experiências prévias, mas, ao mesmo tempo, deve possuir um estoque de células intocadas (virgens) capazes de reconhecer antígenos que ainda sejam desconhecidos dos organismos.

Vamos discutir alguns mecanismos que podem ser os primeiros indícios de eventos que contornam o problema da limitação numérica do sistema imune e, ao mesmo tempo, conferem a vertebrados um poderoso sistema de reconhecimento viral. A maior parte destes mecanismos envolve o fenômeno de reatividade cruzada.

Reatividade Cruzada

A capacidade de um linfócito T em reconhecer dois ou mais peptídeos não relacionados derivados de um mesmo vírus, ou mesmo peptídeos de vírus heterólogos, é denominada de reatividade cruzada (Selin *et al.*, 2006). Esse fenômeno é observado principalmente em linfócitos T citotóxicos (Selin *et al.*, 1994), embora também ocorra em linfócitos T auxiliares (Moran *et al.*, 2008) e em linfócitos B (Tischler *et al.*, 2008). Diversos exemplos de reatividade cruzada entre vírus heterólogos podem ser citados e serão apresentados ao longo do texto.

O trabalho de Brehm *et al.* (2002) demonstrou que a sensibilização de um camundongo com um epítipo subdominante do vírus da coriomeningite linfocítica (LCMV) induz também forte resposta contra um epítipo heterólogo subdominante do vírus pichinde. É importante salientar que o peptídeo responsável pela reatividade cruzada apresentava seis aminoácidos compartilhados de um total de oito, em comparação ao peptídeo cross-reativo. No caso, a infecção por um vírus heterólogo estimulou uma forte resposta imunodominante de células T a um epítipo que outrora era fraco e subdominante, indicando não somente que a hierarquia de células T vírus específica pode ser mais variável do que previamente se imaginava, mas que a hierarquia de imunodominância pode ser marcadamente afetada por um encontro prévio com um patógeno heterólogo.

Outro trabalho importante envolveu um estudo de quatro vírus heterólogos: o vírus da coriomeningite linfocítica (LCMV), o vírus pichinde (PV), o vírus da vaccinia (VV) e o citomegalovírus murino (MCMV). O estudo evidenciou que a

presença de células T de memória para vírus não relacionados pode alterar a resposta imune primária do hospedeiro a um segundo vírus. Foi observada reatividade cruzada entre os vírus estudados, principalmente entre LCMV e PV, e entre LCMV e VV. Também foi demonstrado que a imunização prévia com um desses vírus, em alguns casos, potencializava a eliminação da infecção precoce de um segundo vírus não relacionado. Interessantemente, tais efeitos protetivos dependiam da seqüência da infecção viral e a proteção não era necessariamente recíproca (Selin *et al.*, 1998).

No clássico trabalho de Wedemeyer *et al.* (2001), a presença de células T CD8+ específicas para um epitopo imunodominante do vírus da hepatite C (HCV) entre doadores de sangue que não possuíam uma história de infecção nem por HCV nem pelo vírus da hepatite B (HBV), levou a uma investigação que acabou revelando que estes pacientes haviam sido sensibilizados por uma infecção prévia com vírus influenza A (IVA). A reatividade cruzada ocorria entre um epitopo endogenamente processado da neuraminidase (IVNA-231) do vírus influenza (IV), o qual corresponde a uma seqüência conservada entre diferentes variantes de vírus influenza, sendo normalmente incluída em vacinas, e um determinante imunodominante de HCV denominado HCVNS3-1073 o qual é um dos determinantes HLA-A2 (MHC-I humano) mais freqüentemente reconhecido durante infecção aguda de HCV.

Também chama a atenção neste exemplo o grau de similaridade entre as duas seqüências, as quais compartilham sete aminoácidos, diferindo em apenas dois dos nove resíduos. Além disso, ambos peptídeos apresentam aminoácidos conservados nos resíduos 2 e 9, os resíduos críticos para a ligação do peptídeo à

molécula de HLA-A2. Embora dois aminoácidos sejam distintos na comparação direta dos peptídeos, os aminoácidos que ocupam as mesmas posições nos dois peptídeos pertencem ao mesmo grupo e compartilham características físico-químicas. Este trabalho sugere que um dos fatores que influencia o número de células T com reatividade cruzada existente em um determinado organismo é a frequência de exposição a um dado vírus e a variabilidade de seqüências deste vírus.

Diversos exemplos de reatividade cruzada que envolvem o vírus influenza já foram observados. Isso, de certa forma, é esperado, visto que o sistema imune ao longo da vida de um indivíduo é confrontado diversas vezes com o IV e, conseqüentemente, células de memória direcionadas contra diversos tipos de epitopos de diferentes linhagens de influenza devem ser abundantes.

Já em 1989, Shimojo *et al.* haviam observado que um peptídeo derivado de uma proteína de um rotavírus era capaz de sensibilizar células HLA-A2.1+ para lise por CTLs específicas de M1 55-73 (peptídeo da matriz de IV). Uma mesma variante desse peptídeo, o FLU-M1: 58-66, apresenta reatividade cruzada com o epitopo de HIV-I p17 GAG: 77-85. Células estimuladas *in vitro* com o peptídeo FLU-M1 são capazes de lisar não apenas células marcadas com o epitopo de IV, mas também células-alvo apresentando o epitopo GAG: 77-85 do HIV.

É interessante salientar que esta observação foi realizada em células mononucleares de sangue periférico tanto em indivíduos infectados quanto em indivíduos não infectados pelo HIV, sugerindo que a imunidade à proteína de matriz do vírus influenza pode direcionar uma resposta imune específica a um

epitopo GAG HIV restrito a HLA-A2 em indivíduos vacinados contra influenza. Os dados ainda sugerem que a reatividade cruzada pode ser melhor observada em indivíduos que apresentam uma forte resposta reincidente a este vírus (Acierno *et al.*, 2003). Já se sabia que a proteína de matriz do vírus influenza A e as proteínas do capsídeo e matriz do HIV apresentavam notáveis similaridades estruturais, principalmente referentes à estrutura tridimensional da proteína mais que em relação à identidade na seqüência de aminoácidos. As duas proteínas, M1 de IV e GAG de HIV têm uma função similar na montagem do vírus, elas medeiam a encapsidação do complexo ribonucleoprotéico pela membrana viral (Harris *et al.*, 1999).

Outro trabalho recente especula que a presença de imunidade ao vírus da encefalite japonesa (JEV) modularia a resposta imune induzida pela infecção pelo vírus da dengue. Os ensaios apresentados no trabalho demonstraram que havia reatividade cruzada na resposta imune direcionada a estes dois vírus (Yamada *et al.*, 2003).

O número de exemplos que poderiam ser citados é bem maior, e inclui diferentes níveis de relação entre as seqüências que estimulam reatividade cruzada. Podendo ocorrer reatividade cruzada entre seqüências variantes de um mesmo epitopo de HIV-I restritos a uma família de alelos de MHC de classe I humano (HLA-B57) (Currier *et al.*, 2005). Existe também a ocorrência de reconhecimento cruzado a diferentes peptídeos oriundos do mesmo subtipo do vírus influenza A (IVA) (Belz *et al.*, 2001), ou a epitopos de diferentes subtipos de IVA (Haanen *et al.*, 1999). Outro exemplo de reatividade cruzada, que é relativamente comum, ocorre entre membros de uma mesma família, como entre

membros da família Bunyaviridae (Nicacio *et al.*, 2002; Maeda *et al.*, 2004) ou entre vírus da família polyomaviridae (Tagaram *et al.*, 2008). Um fato que parece bastante interessante é que também são relativamente comuns exemplos de reatividade cruzada entre vírus amplamente divergentes, tais como o vírus da imunodeficiência humana tipo 1 (HIV-1) e o vírus da imunodeficiência humana tipo 2 (HIV-2), os quais apresentam várias diferenças biológicas e estruturais, inclusive em relação às suas seqüências primárias de aminoácidos (Bertoletti *et al.*, 1998).

Diversas observações a respeito da reatividade cruzada são comuns nos trabalhos comentados e parecem estar bem fundamentadas. Primeiro, parece que a ordem hierárquica de imunodominância parece ser influenciada pela reatividade de células T a vírus heterólogos através de dois mecanismos:

- células T que apresentam reatividade cruzada são dominantes na resposta aguda a infecções virais;

- dentro do reservatório de células T de memória, células que apresentam reatividade cruzada são mantidas, enquanto que células específicas para um único epítipo são seletivamente perdidas, sugerindo que toda e qualquer infecção viral modula o repertório de células T disponíveis (Wedemeyer *et al.*, 2001; Brehm *et al.*, 2002; Selin *et al.*, 2004).

As assertivas supracitadas não apenas são corroboradas como também estão exemplificadas no trabalho de Welsh & Selin (2002) que demonstra a modulação do repertório de células T durante as infecções virais, e pode ser ilustrado na figura 1.

A história imunológica de um indivíduo durante seu ciclo vital irá selecionar constantemente o repertório de células T e, conseqüentemente, afetar a indução de respostas mediadas por células CD8+ contra futuras infecções virais, possibilitando que o sistema imune responda convenientemente dentro do contexto no qual está inserido e ao mesmo tempo capacitando-o a responder a possíveis mudanças ambientais representadas pela presença de outros patógenos (Brehm *et al.*, 2002).

Uma visão importante é que as células de memória não são um grupo de unidades isoladas, ao contrário, são parte de uma rede interativa, a qual está continuamente evoluindo, tanto que a resposta imune de um grupo de células altera a frequência, distribuição e atividade das outras (Selin & Welsh, 2004).

Anteriormente pensava-se que células T de memória estavam presentes em frequência relativamente baixa e eram essencialmente células em repouso, mas estudos recentes têm demonstrado que subsequente a infecções de camundongos com (LCMV), (PV) ou (VV), CTLs específicas são mantidas em alta frequência durante toda a vida do animal. Além disso, uma subpopulação dessas células de memória, em um dado momento, consiste de células circulantes citotolicamente ativas, expressando receptores de IL-2 e altos níveis de moléculas de adesão, o que parece estar mais de acordo com um modelo de sistema imune plástico e interconectado (Selin *et al.*, 1998). Dada esta alta frequência, estado de ativação e facilidade de ativação posterior, bem como a propensão à reatividade com agentes heterólogos, pode-se questionar o quanto células de memória geradas em resposta a um agente infeccioso podem interferir no estabelecimento e progressão de uma doença elicitada por um agente heterólogo.

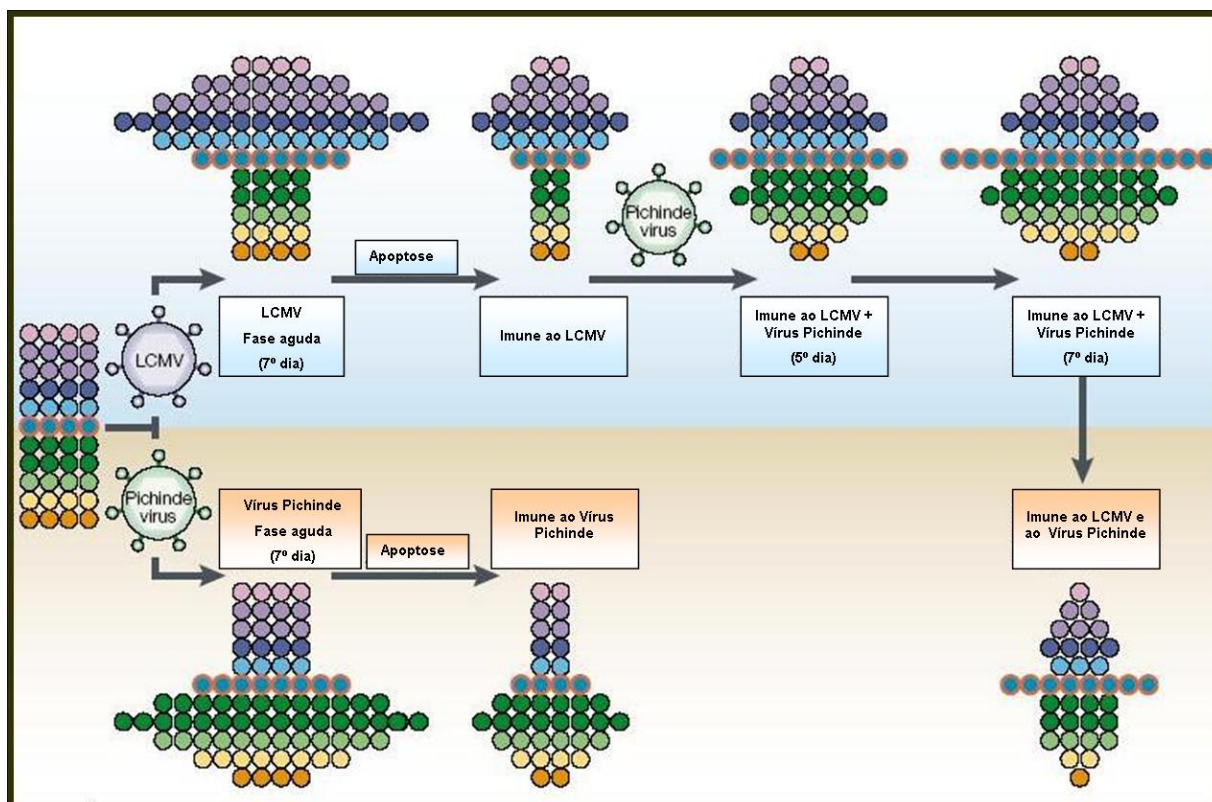


Figura 1: **Modulação do repertório de células T durante uma infecção viral.** Os círculos coloridos representam populações de células T com diferentes especificidades (em um sistema imune “virgem”) as quais são desafiadas com dois vírus heterólogos: Vírus da Coriomeningite Linfocítica (LCMV) e Vírus Pichinde. Algumas populações de células T específicas sofrem expansão para combater a infecção (dependendo do vírus ao qual elas foram desafiadas), e posteriormente entram em apoptose, no entanto deixando uma “prole” de células que resultará em um hospedeiro com um reservatório de células T de memória direcionada. Se um sistema imune que foi condicionado por um vírus (LCMV), for posteriormente exposto a um segundo vírus distinto (Vírus Pichinde), as populações de células que apresentam reatividade cruzada aos dois vírus (azuis com uma linha vermelha circundante) vão sofrer expansão preferencial e dominarão a resposta. Além disso, estas populações serão preservadas e abundantes no reservatório de memória resultante (Modificado de Welsh & Selin, 2002).

A relação do fenômeno da reatividade cruzada com a tolerância a transplantes é bem ilustrada no trabalho de Brehm *et al.* (2003), o qual conseguiu induzir CTLs específicos para os peptídeos derivados do LCMV, através da

infecção de LCMV em camundongos C57Bl/6 (B6), mas que também reconheciam antígenos alogênicos. Estas células aloespecíficas geradas em resposta à infecção viral foram mantidas em altas frequências no reservatório de células de memória, indicando que populações de células T CD8+ de memória aloespecíficas podem surgir como consequência de infecções virais. De fato, camundongos previamente infectados com LCMV eram refratários à indução de tolerância a enxertos alogênicos de pele, diferentemente do que ocorria com camundongos controle que não haviam sofrido infecção viral. Assim, sugere-se que a susceptibilidade à indução de tolerância seja influenciada pela história imunológica do indivíduo: indivíduos que sofreram múltiplas infecções virais tendem a ser mais refratários à indução de tolerância do que indivíduos que tiveram um menor número de infecções. Tal fenômeno tem implicações diretas na pesquisa de transplantes (Adams *et al.*, 2003).

A imunidade prévia a um vírus pode significativamente potencializar a eliminação de um segundo vírus não relacionado no estágio inicial da infecção. Esta eliminação, ou pelo menos uma tentativa de eliminação, ocorre antes do desenvolvimento de células T vírus específicas de alta afinidade, ou seja, antes da estimulação de um repertório distinto de células T virgens. Assim, é delegada às células capazes de reatividade cruzada a tarefa de diminuição da dispersão da infecção, mais que a realização de uma resposta imune efetiva, e o organismo ganha tempo para a montagem de uma resposta mais adequada de células T e B antígeno-específicas de alta afinidade.

A suposição que células T com reatividade cruzada levam a uma resposta de baixa afinidade a epitopos heterólogos levou à argumentação que isto poderia

ocasionar efeitos deletérios ao hospedeiro, pois este careceria de uma resposta imune vírus específica aguda em alguns tipos de infecções virais. Realmente, alguns trabalhos verificaram que tais situações são possíveis, embora pouco freqüentes, e podem se constituir em mais um mecanismo de evasão utilizado pelos vírus. O que se observou é que indivíduos infectados com papilomavírus humano tipo 16 (HPV16) que haviam desenvolvido câncer cervical apresentavam uma quase que total ausência de linfócitos T específicos para o epitopo da oncoproteína (HPV16)E711-19(20). Foi identificada reatividade cruzada entre este epitopo e um epitopo de coronavírus (vírus causadores de doenças respiratórias leves e moderadas em humanos) denominado NS252-60. Os autores argumentaram que um contato freqüente com um vírus intimamente relacionado poderia levar a uma “exaustão” do repertório de células T. Isto poderia provocar uma resposta CTL ineficiente. Além disso, o mesmo epitopo de HPV apresentava considerável homologia de seqüência com peptídeos de diversos organismos, inclusive algumas seqüências provindas de peptídeos próprios, reforçando a sugestão de existência de uma tolerização com intuito de diminuição de chance de desenvolvimento de auto-imunidade (Nilges *et al.*, 2003).

Como podemos ver, a reatividade cruzada está envolvida com a ocorrência de diversos fenômenos, como a imunidade a diferentes patógenos, a manutenção da memória, a autoimunidade e a imunologia de transplantes. Tudo isso graças à capacidade de um único TCR em interagir com diferentes antígenos. Com isso, torna-se fundamental compreender a estrutura do TCR e sua interação com o complexo MHC:peptídeo.

Reconhecimento Promíscuo pelo Receptor de Célula T (TCR)

Muito do fenômeno da reatividade cruzada envolve a capacidade de reconhecimento degenerado de antígenos pelo receptor de células T (TCR). O receptor de células T constitui-se de uma estrutura heterodimérica compreendida por uma cadeia α e uma cadeia β ou, alternativamente, uma cadeia γ e uma δ . Dentro de cada uma dessas cadeias existem três sítios hipervariáveis, conhecidos como regiões determinantes de complementariedade (CDRs), os quais se protraem como alças do TCR e se contatam diretamente com sítios do peptídeo e da molécula de MHC (Brehm *et al.*, 2004).

A visão clássica de uma relação monogâmica entre o TCR de determinado linfócito e seu complexo MHC-peptídeo (MHC-pep) correspondente tem se desgastado. Considerações teóricas recentes têm sugerido, como já comentado anteriormente, que o número de complexos MHC-pep que poderiam ser gerados a partir de antígenos ambientais excederia enormemente o número de células T maduras inclusas no repertório de um indivíduo. Mais especificamente, o cálculo se baseia no número limitado de células T capazes de ocorrerem em um organismo em um determinado momento e não na limitação dos rearranjos do TCR. Só para se ter uma idéia, o tamanho do repertório de diferentes $\alpha\beta$ TCR passível de ser formado é estimado em 10^{15} combinações diferentes, considerando todos os rearranjos dos segmentos gênicos, combinados às junções imprecisas devido a adições e deleções na região N (junções dos segmentos VDJ em que são adicionados nucleotídeos, ao acaso, pela

desoxirribonucleotidil transferase). Esse número é bem superior ao número teórico de possíveis peptídeos nãoâmeros diferentes que podem ser gerados, que é igual a 20^9 (Regner, 2001). É claro que o potencial recombinatório não pode ser exaurido em humanos, mesmo considerando-se um tempo de vida relativamente longo do organismo, e muito menos em pequenos animais, mesmo que existisse um mecanismo garantindo que um dado TCR fosse produzido apenas uma vez no indivíduo. Acredita-se que o repertório de diferentes TCRs de um ser humano varie entre 10^7 e 10^8 diferentes receptores (Mason, 1998; Arstila e cols, 1999).

Por esses e outros motivos, a reatividade cruzada se faz necessária nos receptores de células T. Tais células podem ser ativadas não só pelo peptídeo imunogênico primário, mas também por peptídeos intimamente relacionados, ou até mesmo por peptídeos que apresentam um mínimo de homologia com a seqüência requerida. A reatividade cruzada de receptores de células T se manifesta em diferentes níveis, influenciando desde a seleção positiva de timócitos imaturos até a resposta imune a antígenos heterólogos, em células T periféricas maduras (Regner, 2001).

Considerando que no mesmo indivíduo espera-se que múltiplos clones de células T reconheçam o mesmo complexo MHC-pep, e que existem mais antígenos potenciais que células T em humanos, pode-se pensar que um alto grau de reatividade cruzada é uma propriedade intrínseca e necessária ao reconhecimento de antígenos por essas células. Uma análise por modelo matemático apóia esta idéia predizendo que um simples TCR pode ser capaz de interagir com mais de um milhão de diferentes, apesar de similares, peptídeos (Mason, 1998). As alças do CDR3 do TCR sofrem significantes alterações

conformacionais significantes para acomodar a estrutura tridimensional da superfície do complexo MHC-pep. A flexibilidade conferida ao TCR permite que este seja promíscuo no reconhecimento de peptídeos, fazendo com que ele possua, inerentemente, capacidade de reconhecimento cruzado (Brehm *et al.*, 2004). Um panorama simplificado das possíveis interações teóricas entre o TCR e os complexos MHC:peptídeo pode ser visualizado na figura 2.

A interação das alças do CDR3 do TCR com o complexo MHC:peptídeo apresentado explica mais um nível na degeneração do processo de reconhecimento. Esse contato pode se dar em apenas uma das alças (CDR3 α ou CDR3 β), ou pode se dar em conjunto, com as duas alças contribuindo diferentemente para as interações, dependendo da conformação que o peptídeo adota na fenda do MHC (figura 3). É importante salientar que quando as funções efetoras de um determinado linfócito T são ativadas, o fenômeno de reconhecimento degenerado é reduzido, tornando-se menos sensível a ligantes de baixa afinidade (Mazza *et al.*, 2007).

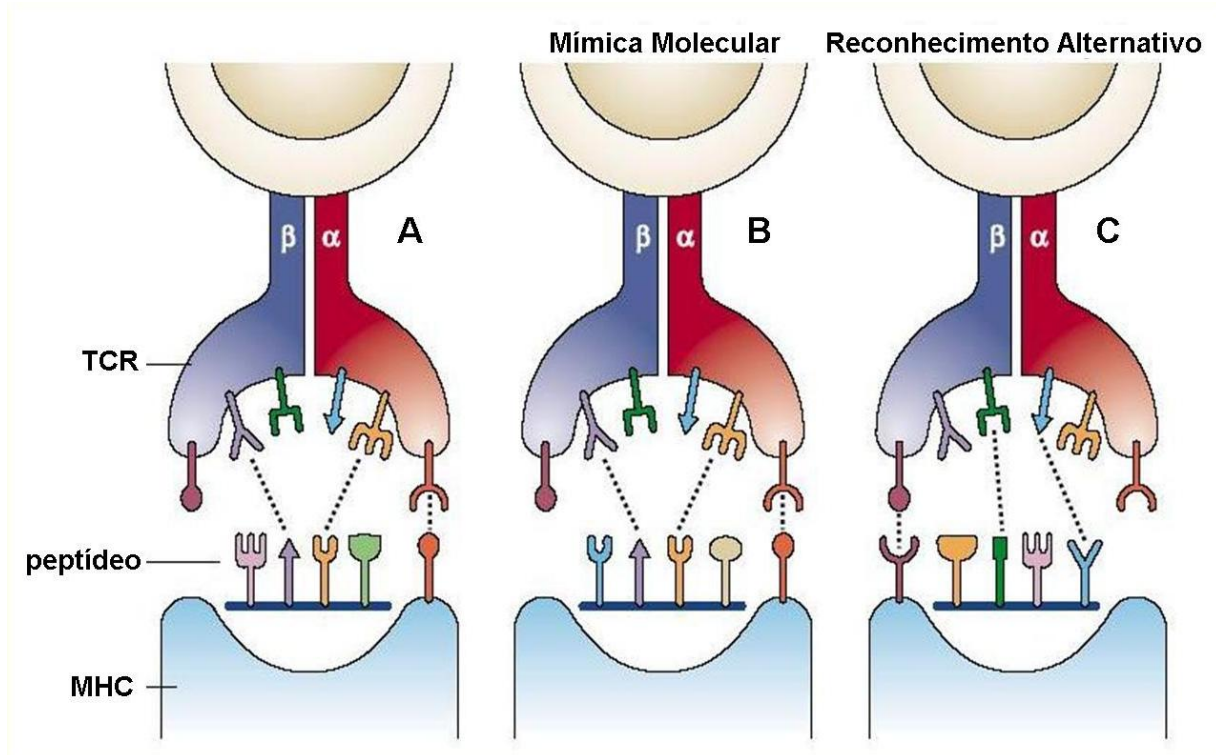


Figura 2: **Mecanismos potenciais de reatividade cruzada em nível de célula T.** Em (A) ocorre a interação das cadeias α e β de um TCR interagindo com um peptídeo no contexto do MHC de classe I. A figura central (B), apresenta outro peptídeo com determinantes similares ao primeiro e que interage de maneira semelhante com o TCR na situação A. Neste esquema o fenômeno apresentado é referido como mimetismo molecular, mas essa situação também ocorre em exemplos de reatividade cruzada usuais. Na terceira situação, os determinantes do TCR que interagem com o peptídeo apresentado são diferentes, e o fenômeno é referido como reconhecimento alternativo (Modificado de Welsh & Selin, 2002).

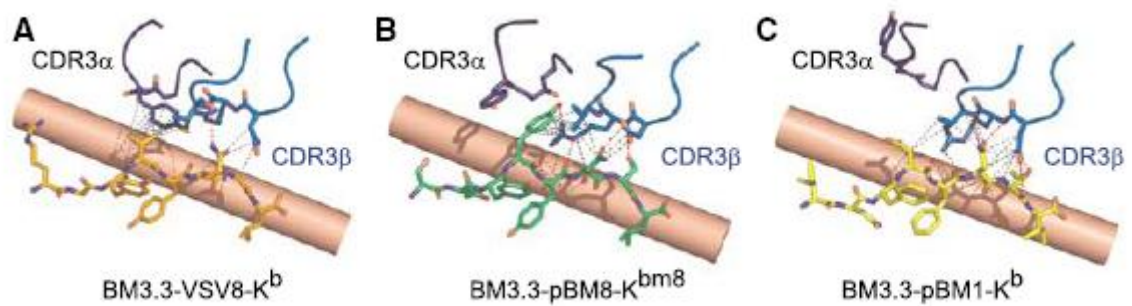


Figura 3: **Três complexos MHC:peptídeo interagindo de diferentes formas com as alças do CDR3 do receptor de célula T.** Em A o complexo interage com ambas alças do CDR3 (α , representada em azul escuro, ou β , representada em azul claro). Em B e C, a interação ocorre com apenas uma das alças, no caso, com as da cadeia β . A barra inclinada marrom representa a hélice $\alpha 1$ do MHC. As linhas pontilhadas entre os átomos dos peptídeos e do TCR, representam pontes de hidrogênio e forças de Van der Waals (Modificado de Mazza *et al.*, 2007).

Um outro nível de flexibilidade no reconhecimento de antígenos ocorre com o MHC. Também nesse ponto, pensava-se que a maioria dos peptídeos antigênicos poderia ser reconhecida no contexto de apenas um ou poucos alelos do MHC. Atualmente foi observado que determinantes de células T podem ser reconhecidos no contexto de vários alelos de classe II (Joshi *et al.*, 2001), permitindo diferentes formas de apresentação do antígeno em questão, embora a base estrutural para este reconhecimento promíscuo não esteja ainda totalmente estabelecida.

É claro que se pode argumentar que nem todos os peptídeos possíveis teoricamente ocorrerão naturalmente e destes, nem todos irão conseguir se associar a moléculas do MHC, mas mesmo assim o número de possíveis complexos MHC-pep efetivamente formado é muito grande. No entanto, devemos considerar que nos complexos MHC-pep formados, muitos não irão desencadear

expansão clonal, reativação dos CTLs de memória e a ação dos mecanismos de lise da célula infectada, sugerindo que um linfócito T possui capacidade em responder diferentemente a distintos complexos apresentados (Reali *et al.*, 1999; Mazza *et al.*, 2007).

Como o Sistema Imune Reconhece os Vírus?

A ocorrência de reatividade cruzada tanto ao nível do reconhecimento de células T (TCR), quanto ao nível da ligação do peptídeo ao MHC dá idéia de um sistema imune dinâmico, que apesar de apresentar um grau de degeneração no reconhecimento de antígenos, possui evidente eficácia na eliminação de infecções virais.

A maioria dos trabalhos anteriormente citados observa a reatividade cruzada a vírus heterólogos. Normalmente considerando os epitopos analisados como não relacionados. Será mesmo que estas seqüências peptídicas não apresentam algum padrão de homologia? Como será que o sistema imune identifica um epitopo viral caracterizando-o como um antígeno viral e não como um peptídeo próprio? Que tipo de particularidades estas seqüências apresentam que as diferenciam, ou melhor, que as caracterizam como peptídeos virais?

Certamente, a indução de uma resposta imune específica leva em consideração não somente o reconhecimento direto do peptídeo apresentado pelas células apresentadoras de antígeno como também o contexto no qual esta apresentação está ocorrendo. Ou seja, em um contexto infeccioso e/ou inflamatório ocorrerá indução de resposta imune específica dirigida contra um determinado peptídeo, ao passo que este mesmo peptídeo apresentado fora do contexto infeccioso e/ou inflamatório não induzirá desenvolvimento de resposta imune, desencadeando uma condição de anergia celular ou tolerância adaptativa (Ryan *et al.*, 2007). O termo tolerância adaptativa é mais atual e descreve um

estado de linfócitos T não-responsivos generalizado (sem proliferação, produção de IL-2 e produção de citocinas efetoras). Entretanto, ao contrário da condição anérgica clássica, este estado não-responsivo pode ser revertido na ausência do antígeno, sendo, portanto, mais plástico. Esta necessidade de co-estimulação para indução de resposta imune específica não invalida, no entanto, as perguntas anteriormente formuladas.

A limitação no tamanho do repertório de células T, já comentada anteriormente, pressupõe que, além dos linfócitos apresentarem reatividade cruzada a diferentes antígenos, nem todas as possíveis seqüências que teoricamente possam ser processadas, contêm os elementos que as caracterizem como proteínas de origem viral.

Verdadeiramente, nem todos estes peptídeos processados são capazes de estimular uma resposta imune propriamente dita. Parece que há uma hierarquia de imunodominância entre os peptídeos formados. Os que estimulam forte resposta imune são considerados imunodominantes, seguidos pelos subdominantes que estimulam resposta fraca ou pouco detectável, e por outros que apresentam resposta detectável apenas na ausência de epitopos dominantes (Yewdell & Bennink, 1999).

Os fatores que influenciam a imunodominância não são ainda bem compreendidos, mas parece que esta característica resulta da combinação de uma série de fatores, entre eles: a eficiência do processamento e apresentação do peptídeo; a afinidade do peptídeo em se ligar à(s) molécula(s) do MHC; o número total de complexos MHC-pep formados; o repertório de TCR disponível e

o fenômeno da imunodominância (células T específicas para peptídeos imunodominantes suprimem a resposta a outros epitopos virais) (Yewdell & Bennink, 1999).

É importante salientar que, principalmente em humanos, esta hierarquia de imunodominância não é rígida, podendo ser alterada por encontros prévios com vírus heterólogos (Brehm *et al.*, 2002).

Em que nível porém reside o padrão que sinaliza que determinada seqüência é originária de uma proteína viral? Os epitopos imunodominantes são os candidatos mais prováveis a responderem esta pergunta, talvez por serem os possuidores de tais características.

Estarão estas características contidas na seqüência primária de aminoácidos (com a necessidade de identidade total para que ocorra reconhecimento)? Ou poderão envolver seqüências de aminoácidos diferentes mas que levem a estruturas tridimensionais semelhantes do peptídeo devido a determinantes estruturalmente parecidos? Ou seja, serão as características de cadeias laterais dos aminoácidos de contato com o TCR ou o padrão de hidrofobicidade dos resíduos as responsáveis pelo “sinal viral”?

Alguns trabalhos dão algumas pistas referentes às potenciais similaridades entre os “sinais virais”. Joshi *et al.*, (2001) estudaram a reatividade cruzada em três epitopos promíscuos de célula T auxiliar (ligados a MHC classe II), os quais não apresentavam forte homologia de seqüência e eram de origens distintas. As seqüências pareciam ser totalmente não relacionadas. Um olhar mais cuidadoso,

porém revelou algumas similaridades estruturais, como a presença de resíduos positivamente carregados no interior de suas seqüências.

O mesmo trabalho comenta que o foco de ação do TCR se concentra sobre poucas cadeias laterais dos aminoácidos do peptídeo. Para se ter uma idéia, o trabalho de Garcia *et al.* (1998) comenta que o contato é realizado basicamente por um a quatro resíduos do TCR com aproximadamente três a quatro resíduos do peptídeo ligado ao MHC. Além disso, o mesmo TCR pode acomodar seqüências peptídicas com diferentes cadeias laterais, dependendo do tamanho e superfície química de contato do mesmo.

Outro indício da existência de seqüências que caracterizam epitopos virais veio do estudo de um epitopo imunodominante de Epstein-Barr vírus (EBV) que representa o principal alvo de reconhecimento por CTLs. No epitopo analisado, verificou-se que apenas 3 dos 9 aminoácidos do peptídeo possuíam função crucial para o reconhecimento por células T. Dois destes aminoácidos eram resíduos de ancoragem e ocupavam as posições 2 e 9 do peptídeo, um outro resíduo na posição 4 realizava o contato com o TCR. Análogos de polialanina sintetizados que continham estes 3 aminoácidos eram capazes de induzir reativação e expansão clonal de CTLs específicos para o epitopo selvagem, indicando que um simples, mas específico, resíduo de aminoácido é suficiente para interagir produtivamente com o TCR. Outro dado interessante deste trabalho é que não apenas o ácido aspártico na posição 4 era capaz de induzir reativação de CTLs específicos, mas também o ácido glutâmico e a glutamina, sendo que todos apresentavam um grupo carbonílico na cadeia lateral. Tais peptídeos, no entanto, falharam em desencadear os mecanismos líticos, demonstrando que

existem diferentes requerimentos para o reconhecimento e para a estimulação do TCR. Assim, diferentes interações devem induzir reativação de CTLs de memória ou desencadear os mecanismos citotóxicos (Reali *et al.*, 1999).

Os autores sugerem que o limiar requerido para ativação dos mecanismos líticos ocorre através de interações íntimas mais rigorosas entre o TCR e as cadeias laterais dos resíduos de contato, reforçando a idéia de que o sistema imune pode ficar “alerta” em resposta a determinados sinais, mas que a total ativação dos mecanismos efetores depende de uma interação cuidadosa e específica.

A partir de trabalhos como estes podemos sugerir que, mesmo considerando-se a necessidade de existência de co-estimulação, realmente alguns fatores caracterizam os epitopos derivados de seqüências virais. Serão estas seqüências tão distintas de outras seqüências também processadas nos proteossomos, principalmente daquelas oriundas de proteínas próprias? Ao que tudo indica, a resposta é sim, pois caso contrário, haveria uma alta incidência de doenças autoimunes. Essa suposição é plausível mesmo que as diferenças suscitadas nos pareçam sutis. O repertório de células T consegue diferenciar e desenvolver resposta apropriada entre antígenos próprios de não-próprios mesmo se a diferença entre eles for de um único aminoácido, desde que esse seja o principal resíduo de contato com o TCR (Kessels *et al.*, 2004).

Entretanto, em alguns casos, epitopos virais apresentam semelhanças consideráveis a peptídeos de origem própria e podem estar envolvidos na patogênese de doenças autoimunes. Já foram descritas evidências

epidemiológicas que infecções freqüentemente precedem reações autoimunes. Além disso, epítopos variantes de origem viral poderiam atuar como antagonistas de seus próprios CTLs, por possuírem os resíduos de ancoragem, mas não o(s) resíduo(s) de contato (Reali *et al.*, 1999).

Um indício do desenvolvimento de autoimunidade a partir de infecções virais foi observado em pacientes com Lúpus Eritematoso Sistêmico (LES), os quais produziam autoanticorpos contra HRES-1/p28, uma proteína nuclear codificada por um retrovírus endógeno humano (ERV). Esta proteína apresentava reatividade cruzada com um peptídeo codificado pela ORF2a de TTV (vírus transmitido por transfusão). A elevada prevalência deste vírus em pacientes com LES (>50%) e o mimetismo molecular com HRES-1/p28 podem estar contribuindo para a geração de autoanticorpos nucleares e patogênese de Lúpus Eritematoso Sistêmico (Gergely *et al.*, 2005).

O mecanismo de reconhecimento cruzado parece ser de extrema importância para a plasticidade e resiliência do sistema imune. A possibilidade de prever o padrão peculiar dessas seqüências virais é de extrema relevância e o número de aplicações possíveis é amplo e será discutido a seguir.

Implicações e Aplicações do Mimetismo Molecular

É difícil dissociar o conceito de reatividade cruzada do mimetismo molecular. Podemos dizer que este último se restringe, fundamentalmente, à capacidade que os vírus “desenvolveram” (selecionaram), ao longo da evolução, de imitar proteínas de seus hospedeiros, tentando controlar seus sistemas de defesa. O termo controlar é mais adequado do que combater, pois não é interessante ao parasita destruir sua maquinaria de replicação. Um bom exemplo disso é o que observamos no HCMV, o qual evoluiu formas de combater tanto a morte da célula do hospedeiro quanto a rejeição pelo mesmo. O HCMV possui genes de proteínas que imitam proteínas antiapoptóticas, evitando a morte celular, e proteínas que induzem a célula infectada a manter um baixo perfil imunológico (análogos de receptores de quimiocinas) (Michelson, 2004).

O envolvimento do mimetismo molecular na patogênese de distúrbios autoimunes baseia-se principalmente em dados epidemiológicos e dados de homologia de seqüências, embora nenhuma evidência direta desse mecanismo, como um iniciador ou acelerador do fenômeno em humanos tenha sido demonstrada até hoje, indicando que mais de um fator deve agir concomitantemente para desencadear a autoimunidade. Mimetopos de baixa avidéz provavelmente requerem eventos inflamatórios locais adicionais para que células T com reatividade cruzada a peptídeos de origem própria possam desencadear uma doença.

Um modelo para diabetes do tipo I já foi desenvolvido em que ratos transgênicos que expressam nucleoproteína (NP) e glicoproteína (GP) de LCMV sob a ação de um promotor de insulina de rato em células β -pancreáticas, quando infectados com o vírus (LCMV) desenvolvem a doença em 95% dos casos. Além disso, infecções seqüenciais com vírus que apresentam epitopos extremamente relacionados (mimotopos) a antígenos próprios podem potencializar um processo autoimune em desenvolvimento, mas não são capazes de quebrar a tolerância, em animais virgens (Christen & von Herrath, 2004).

O exemplo acima descrito, embora tenha sido importante para demonstrar o possível envolvimento de infecções virais com o desenvolvimento de autoimunidade é melhor definido como exemplo de identidade molecular, pois os genes que estavam sendo expressos nos ratos eram idênticos aos do vírus. Em outro trabalho, no entanto, uma situação de mimetismo molecular é melhor evidenciado: um epitopo próprio de rato derivado da dopamina- β mono-oxigenase (KALYNYAPI) apresenta 56% de identidade com a proteína gp33 de LCMV (KAVYNFATC). Afora a evidente disparidade de seqüência, os dois peptídeos são apresentados de forma aproximadamente idêntica no MHC I ao TCR, com uma justaposição da região central de ambos epitopos entre os resíduos 3 e 7. Os complexos MHC-peptídeo formados são muito similares com relação à forma, topologia e distribuição de cargas (figura 4), provendo uma base estrutural para a ativação de células CD8+ por mimetismo molecular e potencial desenvolvimento subsequente de autoreatividade (Sandalova *et al.*, 2005).

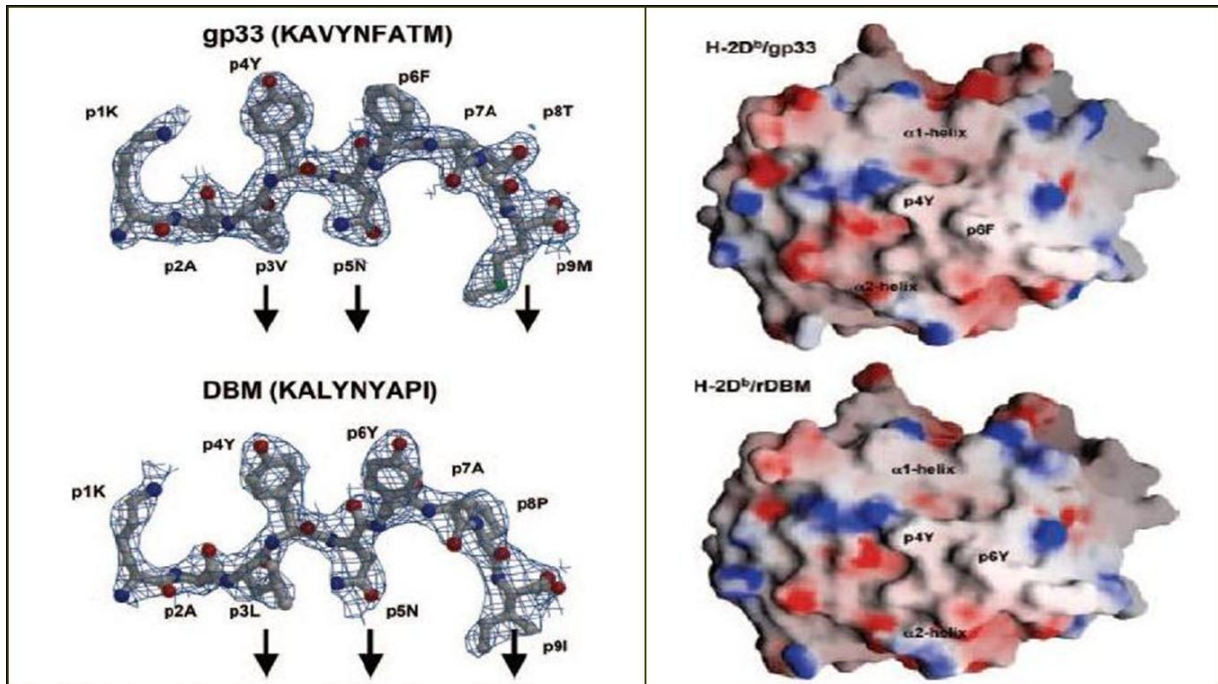


Figura 4: **Exemplo de mimetismo molecular entre um antígeno murino e um viral (LCMV).** No painel à esquerda são apresentados os dois nonâmeros que, a despeito de sua divergência de seqüência (56% de identidade), apresentam conformações estruturais similares. À direita, visão superior de alelos de MHC murino (H2Db) complexados aos antígenos respectivos (na parte superior o epítopo murino e na parte inferior do painel o epítopo viral), apresentando uma topologia de cargas dos dois complexos. Chama a atenção a quase total sobreposição das imagens, evidenciando um potencial reconhecimento cruzado pelo TCR. (Modificado de Sandalova *et al.*, 2005).

Quando se pensa em mimetismo molecular e reatividade cruzada, duas principais aplicações podem ser apontadas: O desenvolvimento de vacinas peptídicas e a indução de tolerância em distúrbios autoimunes. Uma questão importante a ser abordada quando se pensa em vacinas virais é o contexto inflamatório no desenvolvimento da resposta imune. A administração de peptídeos deve ser acompanhada de fortes adjuvantes para capacitá-los com suficiente imunogenicidade, fazendo com que as vacinas sejam protetivas contra desafios virais. Dentre as modificações que permitem a passagem dos epítopos

pela membrana plasmática das APCs, o uso de lipídeos, além de óleos naturais e componentes bacterianos que estimulam respostas inflamatórias, é bastante comum.

Uma abordagem interessante foi realizada por Daftarian *et al.* (2005), os quais testaram misturas de antígenos protéicos com DNA sintético com motivos CpG não metilados (CpG). Esta proposta procurou reduzir a quantidade de vacina e criar uma plataforma molecular capaz de incorporar múltiplos epitopos e entregá-los eficientemente, assim, a entrega do componente antigênico associada ao adjuvante foi pertinente, pois co-localizou ambos estimulantes imunes. Foi teorizado que essa plataforma de oligodeoxinucleotídeos seria capaz de anexar 12 epitopos independentes em suas duas extremidades (5' e 3'). O uso de oligodeoxinucleotídeos contendo motivos CpG em estratégias de vacinação como forte indutor de resposta celular já foi confirmado em outro trabalho (Liu *et al.*, 2006), sendo que neste trabalho ele foi utilizado como emulsificante da proteína de nucleocapsídeo de coronavírus (CoV).

Com a predição de motivos virais e/ou o reconhecimento dos elementos desencadeadores da resposta imune, aliados ao desenvolvimento de técnicas que permitam entregar esses alvos em um contexto infeccioso apropriado, a expectativa do desenvolvimento de vacinas virais de amplo espectro torna-se um objetivo factível e extremamente atraente.

A seguir apresentaremos os capítulos II e III, os quais serão apresentados sob a forma de artigos científicos, tratando respectivamente sobre a elaboração da nossa hipótese de trabalho além da revisão dos mecanismos que envolvem o

fenômeno da reatividade cruzada (capítulo II) e sobre os resultados obtidos através da montagem de um banco de epitopos imunodominantes virais e a comparação dos mesmos em busca de caracteres compartilhados (capítulo III).

OBJETIVOS

O presente trabalho possui como objetivo determinar se existem epitopos virais imunodominantes capazes de sinalizar infecção viral ao sistema imune, de modo similar ao processo existente em bactérias.

Através da análise dos epitopos virais imunodominantes pretendemos estabelecer correlações, buscando um padrão de homologia, para verificar se este padrão se repete em proteínas de diferentes organismos. Temos como objetivo, verificar se tais características se repetem em peptídeos de origem humana e se estas características, conforme discutido anteriormente, podem estar envolvidas no desenvolvimento de doenças autoimunes.

Além disso, procurar-se-á avaliar, através da criação de um banco de epitopos e das diversas abordagens que serão utilizadas, se as sequências de origem viral apresentam identidade preferencial com outras sequências de mesma origem e se existem motivos virais grupo-específicos, além de realizar uma busca contínua por novas sequências consenso.

“Digo o mesmo sobre a filosofia. Ela tem tantas faces, tanta variedade e tem sido tão falastrona que todos os nossos delírios e sonhos podem ser encontrados nela. A fantasia humana nada mais pode conceber de bom ou de mau que lá não se possa achar. ‘Nada se pode dizer de tão absurdo que já não tenha sido dito por algum filósofo.’ Isso me dá maior liberdade ainda para tornar públicos os meus devaneios. Embora eu saiba que eles emanaram de mim, e não de modelos alheios, sempre se poderá encontrar um ou outro autor antigo cujas fantasias são semelhantes às minhas. E sempre haverá alguém para dizer: ‘Vejam aqui, eis de onde ele as tirou’.”

Montaigne (1592)

CAPÍTULO II

Immunodominant Viral Peptides as Determinants of Cross-Reactivity in the Immune System - Can we develop wide spectrum viral vaccines?

Med Hypotheses (2005) 65:873-9

Immunodominant Viral Peptides as Determinants of Cross-Reactivity in the Immune System - Can we develop wide spectrum viral vaccines ?

GF Vieira and JAB Chies

Immunogenetics Laboratory, Department of Genetics, Universidade Federal do Rio Grande do Sul - UFRGS, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil

* Corresponding author: José A. B. Chies, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil. Tel.: +55-51-33-16-67-40; fax: +55-51-33-16-73-11 E-mail: jabchies@terra.com.br.

Summary

When we look back to Edward Jenner vaccination of a young man in 1796, we cannot help thinking that he was both lucky and crazy. Crazy because he decided to test in a human being a hypothesis based mainly in the traditional belief that people who had acquired cowpox from the udders of a cow were thereafter resistant to smallpox, a quite devastating disease, and lucky because (even considering that he did not know this at that time) he succeeded to induce protection against a pathogen through the induction of an immune response directed against a different agent. Not only was he able to protect the young man but he took the first step towards the development of a vast new field, vaccination. It is acceptable to say that Jenner was lucky because he succeeded in promoting protection against smallpox using a cowpox virus, and this induction of protection in a cross-reactive way is believed to be quite rare. Nevertheless, more and more examples of cross-reactive immune responses are being described and we are starting to admit that cross-reactivity is far more common and important than we thought before. Here we review cross-reactivity in the immune system and we discuss the potential use of this feature to the development of wide spectrum vaccines based in conserved immunodominant viral peptides.

Introduction

The human immune system has developed, throughout its evolution, different mechanisms to eliminate pathogenic organisms. Pathogens, in turn, have developed expedients to evade our defense mechanisms. Some of the interaction mechanisms between the immune system and pathogens are already well understood. For viral systems, however, the subject is not fully elucidated, mainly due to the highly mutable nature of viruses, which requires the development of specialized and complex defense mechanisms.

Many cells of the immune system are involved in the eradication of viral infections. We will give special attention to the interactions between cells that express major histocompatibility complex (MHC) class I molecules (all the nucleated cells) and cytotoxic CD8⁺ T lymphocytes (CTLs).

When an MHCI-expressing cell is infected by a virus, it starts producing proteins necessary to viral assembly. Part of these proteins is, however, proteolytically degraded in the proteasomes, along with other cytosolic self-proteins or even proteins provenient from intracellular bacteria that had been also ubiquitinated. This degradation process generates peptides 6 to 30 amino acids (aa) long that could, potentially, bind to the MHC class I molecule peptidic cleft, in the endoplasmic reticulum, being transported to the cell surface. These peptides have little variation in size, presenting normally 8 to 12 amino acids. The peptide size restriction is due to the nature of the MHC I peptidic cleft, which presents closed extremities, contrasting to the peptidic cleft of MHC class II molecules which can accommodate longer sequences (up to 30 aa).

Imagine that a given viral protein possesses 200 amino acids and that its proteolysis generates, in average, peptides with 10 aa (it must be emphasized that fission does not occur, necessarily, at 10 aa from the initial extremity, but that random cuts in any part of the protein generate peptides of different sizes). The number of sequences yielded, in such case, would reach some hundreds. Considering that the exemplified protein size is relatively small, and considering the classical idea that each CD8+ T cell recognizes only one peptide, the repertoire of memory T cells should have to be bigger than 10^{12} , the total number of lymphocytes in human beings, to be efficient (1). Therefore, the recognition of peptidic sequences must involve mechanisms with a certain plasticity and adaptability, without, however, lose specificity. This would allow the immune cell repertoire to be shaped according to ongoing experiences leading to the establishment of memory, but at the same time, it would allow the maintenance of a source of naive, inexperienced cells.

We will discuss here some mechanisms that could indicate events that skirt the problem of the size limitation of the immune system and, at the same time, could confer it a powerful resource to viral recognition. Most of these mechanisms involve cross-reactivity.

Cross-reactivity

The capacity of a T lymphocyte to recognize non-related peptides derived from the same virus, or even peptides from heterologous viruses, will be defined here as cross-reactivity (2,3). This phenomenon is mainly observed in cytotoxic T

lymphocytes, even though it also occurs in T helper cells (4). Several examples of cross-reactivity between heterologous viruses have already been described and some of them will be presented throughout the text.

Brehm *et al.* (5) demonstrated that the sensitization with a subdominant epitope of the lymphocytic choriomeningitis virus (LCMV) elicits strong immune response against an heterologous subdominant epitope of the pichinde virus (PV). It is important to note that the PV peptide which elicited the cross-reactivity shared six out of eight amino acids with the one used to sensitize the cells. In that case, infection with an heterologous virus stimulated a strong T cell immunodominant response to an epitope that was previously weak and subdominant, indicating that not only the hierarchy of virus-specific T cells can be more malleable than formerly imagined, but also that the hierarchy of immunodominance can be deeply affected by previous encounters with heterologous pathogens.

Another interesting work involved a panel of four heterologous viruses: LCMV, PV, vaccinia virus (VV) and murine cytomegalovirus (mCMV). This study evidenced that memory T cells primed by a given virus may alter the host immune response to a second unrelated virus. Thus, it was observed cross-reactivity between those viruses, mainly between LCMV and PV, and between LCMV and VV. It was also demonstrated that previous immunization with one of those viruses, in some cases, enhanced the clearance of a second unrelated virus, early in infection, although the sequence of viral infection was important and cross-protection was not necessarily reciprocal (6).

In the work of Wedemeyer *et al.* (1), the presence of specific CD8⁺ T cells directed to one immunodominant epitope of the hepatitis C virus (HCV) among

blood donors that did not present any history of infection by HCV or hepatitis B virus (HBV), led to an inquiry that revealed that those patients had been cross-sensitized by a previous encounter with influenza A virus (IVA). The cross-response occurred against one endogenously processed epitope of the IV neuraminidase (IV_{NA-231}), which is conserved among influenza viruses, being normally included in vaccines, and an HCV immunodominant viral determinant called HCV_{NS3-1073}, which is an HLA-A2 determinant frequently recognized during acute HCV infection. In the previous example, the degree of similarity between the sequences of both epitopes, which share seven out of nine amino acids, seems to be quite important. Additionally, both peptides present conserved aa in residues 2 and 9, the critical residues for binding to the HLA-A2 molecule. Although they differ in two aa, these non-identical amino acids belong to the same chemical group and share certain physicochemical characteristics.

Several examples of cross-reactivity involving the influenza virus have already been reported. This is, to a certain extent, an expected fact since any given human immune system is challenged by this virus several times throughout life. Consequently, memory cells for many epitopes, from different influenza strains, must be abundant in the T cell repertoire of any human being.

In 1989, Shimojo *et al.* (7) observed that a rotavirus-derived peptide could sensitize HLA-A2.1+ targets, inducing their lysis by CTLs specific to a IV-derived matrix peptide. One variant of this IV peptide, the FLU-M1:58-66, presented cross-reactivity with an HIV-I epitope. Cells stimulated *in vitro* with FLU-M1 were capable of lyse not only cells marked with the IV epitope but also cells presenting the HIV epitope (8). It is interesting to point out that this observation was carried out in

peripheral blood mononuclear cells from both HIV-infected and uninfected individuals, suggesting that, in individuals vaccinated against influenza, the response generated against the IV matrix protein could direct a specific immune response to an HIV epitope. It is known that the IV matrix protein and the HIV capsid and matrix proteins present notable structural similarities, mainly in the three-dimensional structure of the proteins rather than homology in the amino acid sequence. Interestingly, both virus proteins have similar functions in its assembly, mediating the encapsidation of the ribonucleoprotein complex by the viral membrane (9). Another work demonstrated the existence of cross-reactivity in the immune responses directed to japanese encephalitis and dengue viruses (10).

Among phylogenetically-related viruses there are several examples of cross-recognition, such as the occurrence of cross-recognition to different peptides from the same subtype of influenza virus A (IVA) (11), or among epitopes from different subtypes of IVA (12), although cross-reaction between viruses with widely divergent primary amino acid sequences (HIV-1 and HIV-2, for instance) seems also to be common (13).

T cell receptor promiscuous recognition

Much of the cross-reactivity phenomenon involves the degenerated capacity for antigen recognition of the T cell receptor (TCR). The TCR consists of an heterodimeric structure formed by an alpha chain and a beta chain or, alternatively, a gamma and a delta chain. Within each one of these chains, there are three hypervariable sites known as complementary determining regions

(CDRs), which protrude as loops from the TCR and directly contact sites on the peptide and MHC molecule (14).

The classical view of a monogamous relationship between the TCR of a given lymphocyte and the corresponding MHC-peptide (MHC-pep) complex has been wearing. As previously mentioned, the number of MHC-pep complexes that could be generated exceeds that of T cells in the repertoire of an individual. Specifically, this limitation is due to spatial and/or numerical restrictions on the mature T cell pool, rather than the number of possible TCR rearrangements. The size of the potential TCR repertoire is estimated in 10^{15} , considering all gene segment rearrangements and the imprecise junctions generated by insertions and deletions in the N terminal regions. This number is greater than the theoretical number for different nonamer peptides that could be generated, which equals 20^9 (15). It is logical that the repertoire full combinatorial potential cannot be exhausted in humans or other animals, even if there were a mechanism ensuring that any TCR is produced only once in a given individual.

Thus, cross-reactivity is necessary for T cell receptors. Such wide spectrum cells can be activated by a primary peptide and also by closely-related peptides, or even by peptides that present a certain homology with the required sequence. The TCR flexibility is exhibited at different levels, influencing the positive selection of immature thymocytes as well as the immune response to heterologous antigens, in peripheral mature T cells (15).

Considering that, during an immune response, it is expected that multiple T cell clones recognize the same MHC-pep complex, and considering that the number of potential antigens exceeds that of available T cells, it can be suggested

that a high degree of cross-reactivity is an intrinsic and necessary property for an efficient immune system. A mathematical modeling supports this idea, predicting that a single TCR may be able to interact with over one million different peptides (16). The CDR3 loops of the TCR are subject to significant conformational alterations to accommodate the three-dimensional structure of the MHC-pep complex surface. The natural TCR flexibility allows it to be promiscuous in peptide recognition, making it inherently able to cross-recognition (14).

It can be argued that not all possible peptides will naturally occur and that, among them, some will associate with MHC molecules while others will not. Even so, the number of MHC-pep complexes effectively formed is extremely high. Nevertheless, we must consider that not all the MHC-pep complexes formed will trigger clonal expansion, memory CTLs reactivation and infected cell lysis.

Another level in the antigen recognition concerns the MHC. Although the classical view dictates that the majority of antigenic peptides could be recognized in the context of only one or a few alleles of the MHC, nowadays it has been observed that T cell determinants can be recognized in the context of a variety of class II alleles (17), even though the structural basis for this promiscuous recognition is not yet understood.

Implications of cross-reactivity in the available T cell repertoire

Brehm *et al.* (18), through the infection of mice, obtained specific CTLs for LCMV-derived peptides which also recognized allogeneic antigens. These allospecific cells generated in response to the viral infection have been kept in

high frequencies in the memory cell compartment, indicating that allospecific CD8+ memory T cells can develop as a consequence of viral infection. In fact, mice previously infected with LCMV were refractory to the tolerance induction to skin allogeneic graft. The susceptibility to the tolerance induction is influenced by the immunological history of the individual, ie, individuals that have suffered multiple viral infections tend to be more refractory to the tolerance induction than those individuals that have had less infections. Such observation can have some implications in the transplant research field (19).

Also, previous immunity to a given virus can significantly improve the clearance of a second, non related virus, in initial stages of the infection, when specific high affinity T cells directed against this second virus (generated by the stimulation of naïve T cells) were not yet available. Thus, the task of reducing infection spreading can be delegated to cross-reactive cells and the organism will gain time to assemble a more specific artillery. The idea that cross-reactive T cells lead to a low affinity response to heterologous epitopes generated the argument that this could cause deleterious effects to the host, who would lack an acute, virus-specific immune response for some types of viral infections. In fact, some works have verified that such situation is possible and can be used by viruses as an extra evasion mechanism. It was observed that human papillomavirus type 16 (HPV16)-infected individuals that had developed cervical cancer presented an almost complete absence of specific T lymphocytes directed to the oncoprotein (HPV16) E7₁₁₋₁₉₍₂₀₎ epitope. It was identified cross-reactivity between this epitope and a coronavirus epitope called NS2₅₂₋₆₀. It was proposed that frequent contacts with a virus could lead to an “exhaustion” of the T cell repertoire, which could lead

to an inefficient CTL response. Moreover, this same HPV epitope presented considerable sequence homology with peptides from diverse organisms, including some sequences from self-peptides, providing support for the suggestion of toleration of the immune system (20). Another strategy used by viruses to evade the immune system is exemplified by the murine cytomegalovirus (MCMV), which codes for an evasion immune protein (m152/gp40). This protein prevents the presentation of an immunodominant peptide, provenient from an antiapoptotic protein (M45), in pathologically relevant tissues. Thus, this peptide stimulates a CD8+ long-term response, but the target is not present in the tissues of interest (21).

A possible role for the plasticity of T cell recognition in the development of human immunity is supported by De Silva-Udawatta *et al.* (22). Human T cell clones were generated against two autoantigens which are frequent targets of autoantibodies and T cells in the same lupus patient. Interestingly, the TCRs from all isolated clones had substantial sequence homology in their CDR3 region 3. Cloning of the TCR α e β chains from these cells, in a TCR-negative human cell line, and a subsequent analysis of interaction between the TCR and the antigenic peptides revealed TCR stimulation by both peptides, evidencing a degeneration or plasticity in the recognition of these lupus autoantigens by the T cell receptor.

Memory, in the immune system, is also affected by cross-reactivity. The immunological history of an individual will constantly shape the T cell repertoire and, consequently, affect the induction of cells in future viral infections. Memory cells are not isolated in the immune system but participate in an interactive network, which is continuously evolving in such a way that the any immune

response modifies the frequency, distribution and activity of all other components of the immune system (23). Finally, if it was previously thought that memory T cells were present in relatively low frequency and were essentially resting cells, recent studies have demonstrated that virus-specific CTLs were kept in high frequencies during the lifespan of a mice that had been infected by LCMV, PV or VV. In addition, a subpopulation of those memory T cells consisted of cytolytically active circulating cells, expressing IL-2 receptors and high levels of adhesion molecules, which is in agreement with a plastic and interconnected immune system (6).

How does a virus reveal us its identity?

Considering the previous discussion, and focusing specifically in viral immunodeterminants, some questions arise. If the same peptide can, potentially, trigger either a strong immune response or a barely detectable response, how can we talk about immunodominant peptides? In other words, how can the immune system detect a virus? Beyond the need for costimulation, it seems that viral immunodominant peptides share some characteristics.

Joshi *et al* (17). studied cross-reactivity of three promiscuous epitopes of T helper cells (linked to MHC class II), which did not show strong sequence homology and were of distinct origins. The sequences seemed totally not related. A more careful analysis, however, revealed some structural similarities, such as the presence of positively charged residues in the middle of their sequences.

The same work comments that few lateral chains from the amino acids of the peptide concentrate the focus of the TCR action. Moreover, the same TCR can

accommodate peptide sequences with different lateral chains, depending on the size and chemical properties of the TCR contact surface.

In addition, a study with an Epstein-Barr virus immunodominant epitope that represents the main target for CTLs identified that only 3 out of 9 amino acids of the peptide possessed crucial function for T cell recognition. Two of those amino acids were anchorage residues and occupied positions 2 and 9, and the residue in position 4 was responsible for making contact with the TCR. Polyalanine analogues that share these 3 amino acids were capable of inducing reactivation and clonal expansion of specific CTLs for the wild-type epitope, indicating that a simple but specific amino acid residue is sufficient to productively interact with the TCR. Not only was the original amino acid (aspartic acid) at position 4 capable of inducing reactivation of specific CTLs, but also the glutamic acid or glutamine, all of them presenting a carbonylic group on the side chain. Such peptides, however, have failed to trigger the cytolytic mechanisms, demonstrating the existence of different requirements in TCR stimulation. Thus, different interactions must induce reactivation of memory CTLs and unchain the cytotoxic mechanisms (24).

We can suggest that, even though there is a need for co-stimulation, some factors really characterize some epitopes as viral sequences and that these “sequences” will probably not be represented in higher organisms, otherwise we would observe a high incidence of autoimmune diseases. However, in some cases, viral epitopes present considerable similarities to self peptides and can be involved in pathogenesis of autoimmune diseases. Epidemiologic evidences alert that infections frequently precede autoimmune reactions, moreover, variant

epitopes of viral origin may act as antagonists of their own CTLs, by possessing the anchor residues but not the contact residue.

Can we manipulate cross-reactivity?

Cross-reactivity in the TCR-MHC-pep interaction illustrates the dynamic of the immune system. Certainly, the induction of a specific immune response not only takes into account the peptide presented by antigen-presenting cells (APCs), but also the context in which this presentation occurs. In other words, in an infectious and/or inflammatory context the induction of specific immune response against a determined peptide will occur, while the same peptide, if not presented under the same conditions, will not induce an immune response. This need for costimulation does not invalidate, however, the possible use of immunodominant viral epitopes to trigger a protective immune response.

The T cell repertoire size limitation presupposes that several MHC-pep should never be recognized if cross-reactivity inexists. Actually, not all of the processed peptides are capable of stimulate a truly immune response. It seems to exist an hierarchy of immunodominance between the formed peptides (23).The factors that influence the immunodominance are still poorly understood, but it seems that this feature results from the combination of different factors: the efficacy of the peptide processing and presentation; the affinity with which the peptide binds to the MHC molecule; the total number of MHC-pep complexes formed; the available TCR repertoire and the phenomenon of immunodomination (specific T cells for immunodominant peptides suppress the response to other viral

epitopes) (25). Nevertheless, we must have in mind that this immunodominancy hierarchy is not rigid and can be altered by previous immunizations or encounters with heterologous viruses (5).

The cross-recognition mechanism seems to be of extreme importance for the plasticity and resilience of the immune system. The possibility to predict the peculiar pattern of these viral sequences is of extreme relevance and the number of possible applications is wide. Among those, the use of synthetic peptides as vaccines that can stimulate memory cells for a diversity of heterologous viruses and the use of variant peptides that can induce tolerance in autoimmune diseases.

Acknowledgments: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Apoio à Pesquisa do Estado do Rio Grande do Sul (FAPERGS). We thank Andréia Escostesguy Vargas for critical review of this manuscript.

References

1. Wedemeyer H., Mizukoshi E., Davis A.R., Bennink J.R., Rehermann B. Cross-reactivity between hepatitis C virus and influenza A virus determinant-specific cytotoxic T cells. *J Virol* 2001; 75: 11392-400.
2. Williams O.M., Hart K.W., Wang E.C., Gelder C.M. Analysis of CD4(+) T-cell responses to human papillomavirus (HPV) type 11 L1 in healthy adults reveals a high degree of responsiveness and cross-reactivity with other HPV types. *J Virol* 2002; 76: 7418-29.
3. Snyder J.T., Belyakov I.M., Dzutsev A., Lemonnier F., Berzofsky J.A. Protection against lethal vaccinia virus challenge in HLA-A2 transgenic mice by immunization with a single CD8+ T-cell peptide epitope of vaccinia and variola viruses. *J Virol* 2004; 78: 7052-60.
4. Selin L.K., Nahill S.R., Welsh R.M. Cross-reactivities in memory cytotoxic T lymphocyte recognition of heterologous viruses. *J Exp Med* 1994; 179: 1933-43.
5. Brehm M.A., Pinto A.K., Daniels K.A., Schneck J.P., Welsh R.M., Selin L.K. T cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens. *Nat Immunol* 2002; 3: 627-34.
6. Selin L.K., Varga S.M., Wong I.C., Welsh R.M. Protective heterologous antiviral immunity and enhanced immunopathogenesis mediated by memory T cell populations. *J Exp Med* 1998; 188: 1705-15.
7. Shimojo N., Maloy W.L., Anderson R.W., Biddison W.E., Coligan J.E. Specificity of peptide binding by the HLA-A2.1 molecule. *J Immunology* 1989; 143: 2939-47.

8. Acierno P.M., Newton D.A., Brown E.A., Maes L.A., Baatz J.E., Gattoni-Celli S. Cross-reactivity between HLA-A2-restricted FLU-MI: 58-66 and HIV p17 GAG: 77-85 epitopes in HIV-infected and uninfected individuals. *J Transl Med* 2003; 1: 3.
9. Harris A., Sha B., Luo M. Structural similarities between influenza virus matrix protein M1 and human immunodeficiency virus matrix and capsid proteins: an evolutionary link between negative-stranded RNA viruses and retroviruses. *J Gen Virol* 1999; 80: 863-69.
10. Yamada K., Takasaki T., Nawa M., Yabe S., Kurane I. Antibody responses determined for Japanese dengue fever patients by neutralization and hemagglutination inhibition assays demonstrate cross-reactivity between dengue and Japanese encephalitis viruses. *Clin Diagn Lab Immunol* 2003; 10: 725-28.
11. Belz G.T., Xie W., Doherty P.C. Diversity of epitope and cytokine profiles for primary and secondary influenza A virus-specific CD8⁺ T cell responses. *J Immunology* 2001; 166: 4627-33.
12. Haanen J.B.A.G., Wolkers M.C., Kruisbeek A.M., Schumacher T.N.M. Selective expansion of cross-reactive CD8⁺ memory T cells by viral variants. *J Exp Med* 1999; 190: 1319-28.
13. Bertoletti A., Cham F., McAdam S. et al. Cytotoxic T cells from Human Immunodeficiency virus type 2- Infected patients frequently cross-react with different human immunodeficiency type I clades. *J Virol* 1998; 72: 2439-48.
14. Brehm M.A., Selin L.K., Welsh R.M. CD8⁺ T cell responses to viral infections in sequence. *Cell Microbiol* 2004; 6: 411-21.

15. Regner M. Cross-reactivity in T-cell antigen recognition. *Immunol Cell Biol* 2001; 79: 91-100.
16. Mason D. A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol Today* 1998; 19: 395-404.
17. Joshi S.K., Suresh P.R., Chauhan V.S. Flexibility in MHC and TCR recognition: degenerate specificity at the T cell level in the recognition of promiscuous Th epitopes exhibiting no primary sequence homology. *J Immunology* 2001; 166: 6693-03.
18. Brehm M.A., Markees T.G., Daniels K.A., Greiner D.L., Rossini A.A., Welsh R.M. Direct visualization of cross-reactive effector and memory allo-specific CD8 T cells generated in response to viral infections. *J Immunol* 2003; 170: 4077-86.
19. Adams A.B., Williams M.A., Jones T.R. et al. Heterologous immunity provides a potent barrier to transplantation tolerance. *J Clin Invest* 2003; 111: 1887-95.
20. Nilges K., Höhn H., Pilch H. et al. Human papillomavirus type 16 E7 peptide-directed CD8⁺ T cells from patients with cervical cancer are cross-reactive with the coronavirus NS2 protein. *J Virol* 2003; 77: 5464-74.
21. Holtappels R., Podlech J., Pahl-Seibert M.F. et al. Cytomegalovirus misleads its host by priming of CD8 T cells specific for an epitope not presented in infected tissues. *J Exp Med* 2004; 199: 131-36.
22. De Silva-Udawatta M., Kumar S.R., Greidinger E.L., Hoffman R.W. Cloned human TCR from patients with autoimmune disease can respond to two structurally distinct autoantigens. *J Immunol* 2004; 172: 3940-47.

23. Selin L.K., Welsh R.M. Plasticity of T cell memory responses to viruses. *Immunity* 2004; 20: 5-16.
24. Reali E., Guerrini R., Marastoni M. et al. A single specific amino acid residue in peptide antigens is sufficient to activate memory CTL: potential role of cross-reactive peptides in memory T cell maintenance. *J Immunol* 1999; 162: 106-13.
25. Yewdell J.W., Bennink J.R. Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu Rev Immunol* 1999; 17: 51-88.

“A ignorância produz atrevimento; a reflexão, vagar.”

Péricles (Século V a.C.)

CAPÍTULO III

Constrains on the variation of viral proteins and vaccine development

BMC Microbiology (submetido)

Constrains on the variation of viral proteins and vaccine development

Gustavo Fioravanti Vieira^{1,2}, Andrés Delgado-Cañedo³, Dinler Amaral Antunes¹, Maurício Menegatti Rigo¹, José Artur Bogo Chies^{1,2§}

¹Genetics Department, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

²Post-Graduation Program in Genetics and Molecular Biology, UFRGS, Porto Alegre, Brazil

³Laboratory of Molecular and Cellular Cardiology, Cardiology Institute, Porto Alegre, Brazil

§Corresponding author

Email addresses:

GFV: gusfioravanti@yahoo.com.br

ADC: andres.lcmc@cardiologia.org.br

DAA: dinler@gmail.com

MMR: mauriciomr1985@gmail.com

JABC: jabchies@terra.com.br

ABSTRACT

Background

An important concern involving the ability to discriminate between self and non-self peptides is that not all peptides generated by antigen processing will be efficiently presented to T-cells in the MHC and, conversely, there will not necessarily be T-cells in a given organism capable of recognizing each specific peptide. In order to bypass this situation, the immune system developed mechanisms that enable a T-Cell Receptor to recognize more than one peptide (cross-reactivity) as well as evolved in such a way that prioritize the presentation of selected peptides that induce efficient immune responses.

Results

Here we analysed a group of 69 viral derived peptides predominantly described as immunodominant, focusing on how they pass through the peptide processing pathway, in order to identify shared features that made them ideal targets to the immune system.

Conclusions

We describe a common virus-derived epitope motif **X [AGPVLIM]X(6)[AGPVLIM]** and discuss the possibility to develop wide spectrum viral vaccines.

Background

The vertebrate immune system is characterized by its ability to develop and direct immune responses in a specific manner. Therefore, the ability to mount a response is strictly dependent on antigen recognition. T-cell antigen binding is always achieved through the recognition of small peptides presented in the context of the Major Histocompatibility Complex (MHC). The recognition of immunogenic signals is the first step to the occurrence of the immune system activation. There are different molecular recognition pathways both in the innate as well as in specific immune system. In an innate immune response, Toll Like Receptors (TLRs), among other molecules, are capable of identifying molecules and molecular patterns typical from pathogens (pathogen-associated molecular patterns – PAMPs), and quickly activate response mechanisms [1].

Besides the TLR targets, it was suggested that peptidic sequences derived from viruses, mainly those that are referred as immunodominant epitopes, should share features that would provide markers of viral infection to the immune system [2]. This idea is based on the principle that, if each peptide originated from the proteolysis should be recognized by a unique T-lymphocyte, the T-cell repertoire would obligatorily be extremely vast. Moreover, it is known that not all peptidic sequences can induce an immune response, what implies the existence of elements, both inherent to the epitopes themselves as well as related to the environment/context of the immune presentation, that allows a proper identification of the peptides against which an immune response should be directed.

One evidence pointing to the existence of shared elements on viral epitopes allies the inherent cross-reactivity of T-cells, meaning the recognition of more than

one antigenic determinant by the same T-cell [3], with the fact that this phenomenon seems to be quite common when viruses are the targets. The recognition of different antigens by the same T-cell allows the development of an immune response against elements that were not, necessarily, the first targets of the response. Cross-reactivity can be observed against different peptides from the same virus [4], against viruses belonging to the same family [5,6] or even against highly divergent viruses [7].

The relation between the number of peptides that can potentially be produced and the number of peptides that can actually be presented to the immune system also suggests a need for conserved features among immunodominant viral epitopes. Although, theoretically, hundreds of peptides can be generated from a single protein, the number of peptides able to pass through the different bottlenecks that lead to the presentation in the MHC is quite reduced since a series of special requirements are needed at each step (Figure 1). Therefore, a significant number of epitopes can theoretically be produced, but only a small proportion of these peptides will be presented in the context of the MHC [8,9] and will be able to stimulate a T-cell receptor and induce an immune response [10].

Through the use of bioinformatics tools, it is possible to predict putative immunogenic epitopes derived from protein sequences on the basis of the identification of “typical” epitope features on the original protein (such as patterns of proteasomic cleavage, the presence of MHC anchor residues or Transporter Associated with Antigen Processing (TAP) affinity) [11,12]. Nevertheless, the simple presence of all these features in a given protein region does not assure that

this region will elicit an efficient immune response *in vivo* [13,14]. Therefore, in the field of viral vaccine development, more important than simply predict putative immunogenic epitopes, it is necessary to identify epitopes that can actually be used as targets to an immune response.

In the present work, we approached the identification of similarities among viral peptides through the comparison of immunogenic epitopes already described in the literature. Therefore, not only the features considered to be important to epitope generation should be present in these sequences, but also the elements that contribute to the induction of an efficient immune response. Most of these epitopes were described as immunodominants, but subdominant epitopes were also included in our analysis since the hierarchy of immunodominance is affected by previous encounters with other epitopes [10]. We also included epitopes described as cross-reactive, what allows us to analyze if this cross-reactivity can be identified at amino acid level, considering both the amino acid sequence and biochemical characteristics.

Results

Selection of epitopes and establishment of a Databank

A total of 69 viral epitopes were included in the analysis: 23 were described as immunodominant or putative immunodominants, 19 were described as presenting some degree of cross-reactivity, 03 presented both immunodominance and cross-reactivity, 03 were subdominant, and 17 were immunogenic viral epitopes (references in additional file 3). We also included modified or artificial peptides, specifically a SV40 synthetic peptide, containing an optimal binding motif

to H-2D^b; a Cytomegalovirus (CMV) peptide with an Human Leukocyte Antigen A2 (HLA-A2) consensus binding motif; an antigenic variant of the immunodominant Influenza Virus (IV) peptide NP₃₆₆₋₃₇₄ and an enhanced epitope from a highly conserved sequence RT₁₇₉₋₁₈₇ de HIV-1 “Supplementary Epitope Databank”.

Identification of a Common Viral Motif

The first step of our experimental approach was the alignment of the epitopes from this databank in order to identify potential viral motifs. This alignment took into consideration the physicochemical features of all residues that composed the epitope sequence. When each amino acid was classified according to the biochemical properties of their R groups, the first evidence of the existence of a viral motif started appearing. Thirty one (44.93%) out of the 69 sequences from our databank presented an alignment generating a putative motif that included anchor residues **X [AGPVLIM] X(6) [AGPVLIM]**, in which **X** represents any of the essential amino acids (Table 1). These anchor residues (the amino acids alanine, glycine, proline, valine, leucine, isoleucine and methionine) possess aliphatic non-polar R groups as side chains. This value (44.93%) is almost twice the expected value to the random occurrence of such a motif in this group of sequences (meaning 25.7%, $\chi^2 = 13.365$; $p < 0,001$), according to a chi-square goodness of fit test.

BLAST Analysis and identification of similarity with other sequences

From the original 31 sequences aligned according to the presence of consensus residues on their anchor sites and submitted to a BLAST [15] search, 29 (93.54%) presented significant similarity (established as more than 80%) to other viral sequences. It is important to point out that 12 out of them (38.71%) presented significant similarity exclusively with sequences derived from viruses. Fifteen sequences (48.39%) also presented significant similarity with bacterial-derived sequences, and only two of the tested epitopes (6.45%) presented a significant similarity with sequences from human origin (Table 2).

Bottlenecks in the pathway of epitope generation and presentation

Proteolysis

From our databank, 50 epitopes (73.53%) presented a sequence at the COOH-term cut point needed for a good proteasomic processing. Considering only the epitopes aligned on Table 1 (31 sequences), 26 presented a sequence indicating a precise cut at the carboxy terminal region, representing 86.67% of the sequences.

Twenty one proteins related to the epitopes from our databank possessed structure on the Protein Data Bank – PDB [16] and 5 proteins could be modeled by homology “Supplementary Models”(anexo II). As a result, we could observe that the amino terminal end of 67% epitopes from the total databank and 75% from de

aligned epitopes from Table 1 were structurally related to loops on the structure of the protein (Table 3). Detailed results of the structural and proteolytic cleavage analyses are shown in Table 3.

TAP transport and MHC Anchoring

From 1822 TAPligands in the AntiJen [17] at the time of the analysis (December, 2007), 839 (46.04%) presented non-polar amino acids at the putative anchor residues. All outputs obtained from this analysis are represented on “Supplementary Table 5”. From data obtained on the literature, we were able to identify the MHC affinity in 29 out of the 31 epitopes grouped in our alignment. Twelve out of the 29 epitopes were described as ligands to human MHC-I and other 15 to murine MHC-I. Interestingly, 2 of the epitopes were described as able to bind to both human and murine MHC-I. All alleles able to bind to human MHC-I were HLA-A restricted, suggesting that the peptide motif described in the presenting work could be allele specific. The complete analysis is presented in Table 4.

Discussion

Although it is already known that anchor residues are needed to the binding of peptides to the MHC molecule and that the presence of specific amino acid residues on the anchor sites were already reported in other studies involving viral cross-reactivity [14], this is the first work where it was possible to predict a broader preferential viral motif. Although it was possible to identify conserved amino acid

residues in the anchor regions, the same was not true to the central region of the motif. A higher amino acid diversity on the central region of the peptides was expected since these central amino acids are not anchor to the MHC but responsible for the contact with T-Cell Receptors (TCRs). It was suggested that the amino acids on the central region of a peptide determine both the intensity of a CD8+ T-cell response as well as the hierarchy of dominance of this peptide [14]. As our databank was comprised of a sample of viral epitopes described in the literature as immunodominant or cross-reactive, independently of origin (type of virus) or protein function, a complete shared pattern, common to such diverse organisms would be quite improbable. Nevertheless, the existence of motifs and patterns more restricted to specific viral subgroups is a reliable hypothesis. In fact, it was observed when individual viral-derived sequences were submitted to a BLAST [15] search (see the subsequent discussion on the nucleocapsid protein of 46 viruses of the hantavirus genus).

As a rule, we can observe that the output sequences of each analysis were consistently related to a given viral group. For example, the BLAST of epitope 24 from Table 2 (EBV – BMLF1₂₈₀₋₂₈₈) against the whole GenBank generated as output a series of sequences, belonging to different viruses from the Herpesviridae family. More interestingly is the fact that these output sequences were restricted to herpesviruses that infect primates, although the Herpesviridae family contains members that infect a wide range of hosts. This strongly suggests that viral conserved motifs that are used by the immune system as immunodominant targets can be group specific, although examples of immunity induced by cross-reactivity against non-related virus were already reported [18]. As a control, from the entire

sequence of the protein corresponding to each peptide, a randomly selected sequence, ten amino acids long, was also submitted to a BLAST search. The results indicated, differently from what was observed for the immunodominant peptides, a broader presence among different organisms, suggesting that good viral targets to the immune systems should be more restricted [see Table 6 (SI)].

Other examples suggesting the existence of conserved motifs in viral families that help the establishment of a family-specific immune response were also observed. In fact, 21 out of the 31 epitopes generated a BLAST output with sequences derived predominantly from viral organisms phylogenetically related to the original query. The epitope 07 (Variola – HRP2₇₄₋₈₂), for example, showed similarity with proteins derived from other viruses from the Poxviridae family; and the output of four other epitopes (10, 16, 23 and 25) resulted in sequences predominantly derived from other subtypes or lineages from the original viruses. For instance, the epitope 10 (Measles – MV-C₈₄₋₉₂) showed similarity only with other lineages of measles viruses (Table 2).

Similarity among viral sequences and human derived sequences was rarely observed. If this was not the case, we should expect a high incidence of autoimmune reactions. Development of autoimmunity in response to vaccination is indeed an uncommon phenomenon, and the selection of targets which direct immunodominant responses should act to maintain low levels of such a situation [19].

It should be pointed out that, to be presented in the MHC, a peptide must be processed and, consequently, several pathways and enzymatic machineries will be involved in this process. In order to observe the action of these mechanisms,

we analyzed the epitopes that share the putative anchor residues in what concerns their ability to pass through the several bottlenecks of peptide processing and presentation.

Initially, the protein should be cleaved through the recognition of specific flanking regions by a proteasome. We retrieved the original protein sequences from the epitopes (at the GenBank), and from the whole databank, fifty epitopes (73.53%) presented a sequence that allowed a precise cut at the carboxy terminal region. This number reached 26 (86.67%), when we restricted this analysis to the epitopes presenting the conserved anchor residues, suggesting that the majority of the epitopes presented the classical described requirements to antigenic processing. Interestingly, some of the epitopes, that according to the literature actually induce immune responses *in vivo*, do not possess the required flanking sequences needed to proteasome cleavage. Therefore, the programs of proteasome cleavage simulation are not as complete as the real proteasome and approaches that use such programs to identify relevant sequences to be used in vaccine development are possibly neglecting important targets. It is also known that cytosolic peptidases, such as Tripeptidyl peptidase II (TPPII), are involved in the generation of Cytotoxic T-lymphocytes (CTL) epitopes, either in combination with or in a proteasome independent way [20] and therefore, the epitopes lacking sequences needed to proteasome cleavage could be processed through different pathways.

It is interesting to speculate that protein regions on the molecule periphery are more exposed and are easier to be accessed by whichever enzymatic machinery that could be involved in epitope generation. An approach that can

suggest accessibility to cleavage includes the analysis of the tertiary structure of a protein. From the PDB and from the modelling data, we could observe that the predicted cleavage sites on the amino-terminal region of the epitopes were structurally related to loops on the structure of the original protein (66.7%). In order to test if this do not simply represent a higher occurrence of amino acids in loops in the analyzed proteins, the frequency of such structures was determined in the whole molecules. The frequency of amino acids in loops was 47%, considering the whole proteins. This data suggests that there is preferential cleavage at loops ($\chi^2=6.09$ $p=0.047$). It is not yet established if proteins, to be enzymatically processed by a proteasome, should obligatorily be previously completely denatured. If denaturation occurs only partially, it is possible to imagine that some three-dimensional domains (such as alpha-helices or beta-sheets) could interfere in the action of enzymes and, conversely, that a loop would confer a more accessible region to proteasome digestion. Therefore, a loop at the amino terminal end would be a point of choice for a first cut. After that, the protein would suffer a partial denaturation and the proteasome would gain access to the precise point of cut at the carboxy terminus.

The transport of the peptide through the TAP complex is also an important step in epitope processing. In the TAPligand databank, 46.04% of all ligands deposited shared the presence of non polar amino acids on the putative anchor residues, suggesting that the motif inferred by the alignment of the epitopes from our databank represents a good and common motif to TAP binding. Again, it points to the need of specific features to an epitope go through this bottleneck (peptides able to be carried by TAP) and proceed to immune stimulation.

In the present work, we were able to identify the occurrence of preferential patterns on immunodominant viral epitopes, probably due to the selection of characteristics that enable the peptide to go through all the bottlenecks of antigen processing and presentation. Initially, we had postulated that the identification of a motif common to all viral-derived peptides would be quite improbable since it is known that secondary characteristics of amino acids, such as charge, polarity and hydrophobicity, are as important to the specific immunogenicity of the epitope as the linear amino acid sequence is [2,21,22]. Nevertheless, we were able to identify motifs restricted to specific viral groups. For instance, from the output sequences obtained from the Blast search with the epitopes 27 and 31 (which represent cross-reactive epitopes), an alignment was performed with the nucleocapsid protein of 46 viruses of the hantavirus genus of the Bunyaviridae family. A motif was identified from this alignment (NC₁₈₀₋₁₈₈) S[LM][PA][STN][AS]QS[ST][MI] and was submitted to a pattern search in the site Max Planck Institute for Developmental Biology [23] (<http://toolkit.tuebingen.mpg.de/patsearch>). This search returned 359 sequences, all but one belonging to proteins of hantaviruses [the only discordant sequence belonged to *Aspergillus oryzae* - (SI)]. Besides, it is important to point out this exclusive hantavirus motif presented no similarity with sequences from higher organisms, a quite important feature in order to avoid risks of autoimmunity if we consider vaccine development.

Then, considering the different steps of antigen processing and presentation, we can imagine the following picture: from a theoretical 100 amino acid long viral protein, the number of potential nonamers that can hypothetically be generated by cleavage would be 92 (Figure 1A); nevertheless, the number of

peptides actually generated by proteolysis (generally occurring at the proteasomes) will never reach this theoretical value due to, among other factors, the absence of sites to the action of proteases, unfavorable conditions of ubiquitination, or even because of their inaccessibility to the enzymatic machinery (Figure 1B). Besides, not all of these peptides will be efficiently translocated by TAP to the ER (Figure 1C), and, from the translocated peptides, only those with amino acids able to anchor to the MHC will effectively fit on the peptide groove and will be able to be presented to lymphocytes, eventually inducing an effective immune response (Figure 1D). Thus, these “selected” peptides should possess the requirements needed for cleavage, transport and binding to MHC and, in addition, they need to contain elements capable of inducing an immune response. Considering that the phenomenon of cross-reactivity is a relatively common event, immunodominant epitopes should share some features and, consequently, specific patterns could be identified among such peptides.

Conclusions

The existence of motifs shared by different viruses brings a question concerning the commonness of such motifs. If they are used as favorite targets by the immune system, how can one expect that such a situation of vulnerability can be accepted by organisms with high protein malleability as viruses? A possibility is that these epitopes are localized in regions, or domains, on proteins that could not suffer alterations due to the involvement in essential functions [24]. Therefore,

mutations in certain gene regions are not allowed since any alteration in such structures would put in risk the organism viability. Although considered as organisms with extreme flexibility, viruses present constrains on their potential of diversification and therefore it is expected that natural selection should favor MHC alleles able to bind to these restricted sequences. In the evolutionary tug-of-war between viruses and the immune system, these constrains on viral diversification would favor the immune system, allowing the development of cross-reactivity as a strategy in the defense against viral infections.

In conclusion, we suggest that it is possible to identify shared patterns specific to viruses among immunogenic epitopes and therefore to draw strategies to develop wide spectrum viral vaccines. We can postulate the use of a reduced number of epitopes in order to sensitize an organism against several pathogens. The use of epitopes instead of attenuated pathogens will be safer and will help to prevent unwanted or exacerbated immune reactions against other components of the vaccine.

METHODS

Establishment of the Epitope Databank

An exhaustive search on the literature lead to the identification of 69 viral epitopes described as immunodominant or putative immunodominants. We annotated in the databank the following informations: ID and sequence from the protein that gives origin to the epitope; natural or experimental host where the original tests were performed; MHC restriction and abbreviated bibliographical reference.

Alignment

In order to verify if the selected epitopes share physicochemical features important to the induction of an immune response, they were compared, and their amino acids were classified according to their R groups (aliphatic non-polar R, aromatic R, polar non-charged R, negatively charged R and positively charged R) [25], using the program BioEdit *version 7.0.4.1* [26]. All epitopes with size equal or inferior to 10 amino acids were aligned. We limited the analysis to epitopes up to 10 amino acids in length due to a probabilistic limitation. The probability of occurrence of a given amino acid group in two independent positions increases progressively as peptide length increases. Thus, epitopes longer than eleven amino acids restricted to MHC-I as well as epitopes restricted to MHC-II, which are naturally longer, were excluded from our databank, since they would affect the expected values of occurrence of a given motif.

BLAST protein-protein

To compare the epitope sequences from our databank with sequences from other organisms to determine the degree of exclusivity of such epitopes, two different approaches were used: first, a comparison between each epitope and all sequences from the GenBank was performed. Second, each sequence was compared to sequences from human origin, in order to verify the possibility of cross-reactivity and potential to induce autoimmune reactions through molecular mimicry. In these analyses, identity values upper than 80% were considered significant in terms of similarity (this value was arbitrarily chosen, although it is known that cross-reactivity can occur between proteins with lower identity values). Outputs that contained gaps were excluded. Both analyses used the BLAST *protein-protein* section from NCBI. Sequences considered similar to our input sequences were identified according to their origin (organism), and the origin of viral sequences was classified using the ICTVdB [27].

Proteolytic Simulation

A simulation of proteasomic cleavage was performed with NetChop 3.0 [28] (<http://www.cbs.dtu.dk/services/NetChop/>) and with Prediction Algorithm for

Proteasomal Cleavages – PAProC [29] (<http://www.paproc.de/>). The proteolysis simulation was performed on the whole protein sequences, which originated the immunodominant peptides and, consequently, the amino acids flanking the peptide regions were known. As positivity criteria in processing, we considered the exact cut in COOH-term of epitope and a putative cut up to seven amino acids upstream in amino-terminal extremity, taking into account that this extremity is degraded later by other cytosolic and endoplasmic reticulum (ER) proteases [28,30,31]. All the 69 epitopes from our databank (except epitope 48 since it is a synthetic epitope and therefore it is impossible to define flanking sequences) were submitted to the proteasome simulation.

Molecular Modeling

Five **proteins**, corresponding to the epitopes (012, 043, 045, 053, 055 and 056) from our databank, do not possessed structure on the PDB and were modeled by homology from other proteins (Additional file 4) (anexo II). This approach was accomplished using the program Geno 3D [32]. The models generated were analysed and visualized with the program Swiss Protein Data Bank Viewer – SPDBV [33] (SI).

TAP binding

The TAPligand section on AntiJen [17] (<http://www.jenner.ac.uk/AntiJen/>) that comprises peptides ligands to TAP was used to analyze potential binding motifs, classified according to the physicochemical characteristics of amino acids present on the anchor sites. Twenty five putative anchor motifs were tested, including the predicted non-polar motif determined in our alignment; and the number of output sequences with the same characteristics was annotated.

Statistical Analysis

In the analysis of the results obtained in the alignment (table 1), we used a chi-square goodness of fit test since we had a theoretical expected value. In order to obtain a probability at random of the motif found in a specific epitope group, we

calculated the average among the expected values of all sequences from the alignment of Table 1. In a sequence of 8 amino acids, for instance, the expected value, needed to occur 2 non polar amino acids (A, G, P, V, L, I or M) interspersed by 6 random amino acids, which correspond to the probable anchor residues, is equal to 12.25% ($7/20 \times 7/20$). These values correspond to the number of non polar amino acids divided by the total number of amino acids. According to the increase in length of epitopes, the probability of occurrence of the motif increases proportionally. For example, a nonamer presents two possibilities of occurrence of the motif (in first and eighth positions, and in second and ninth positions), resulting an expected value of 25%. Because of this limitation, epitopes longer than 10 amino acids were excluded from our databank in order to avoid illegitimate associations.

Author Contributions: J.A.B.C. analysed the data and wrote the paper; G.F.V. designed the study, established the epitope databank, analysed data and wrote the paper; D.A.A. performed and analysed the proteolytical data, performed the BLAST search; M.M.R. performed the analysis on the hantavirus genus; A.D.C. was involved in study design. All authors discussed the results and commented on the manuscript.

Acknowledgments: G.F.V received a scholarship from CAPES, (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) Brazil. D.A.A. received a scholarship from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), Brazil.

References:

1. Bowie AG, Hago IR: **The role of Toll-like receptors in the host response to viruses.** *Mol Immunol* 2005, **42**:859-867.
2. Vieira GF, Chies JA: **Immunodominant viral peptides as determinants of cross-reactivity in the immune system - Can we develop wide spectrum viral vaccines?** *Med Hypotheses* 2005, **65**:873-879.
3. Felix NJ, Donermeyer DL, Horvath S, Walters JJ, Gross ML, Suri A, Allen PM: **Alloreactive T cells respond specifically to multiple distinct peptide-MHC complexes.** *Nat Immunol* 2007, **8**:335-337.
4. Belz GT, Xie W, Doherty PC: **Diversity of epitope and cytokine profiles for primary and secondary influenza A virus-specific CD8+ T cell responses.** *J Immunol* 2001, **166**:4627-4633.
5. Krymskaya L, Sharma MC, Martinez J, Haq W, Huang EC, Limaye AP, Diamond DJ, Lacey SF: **Cross-reactivity of T lymphocytes recognizing a human cytotoxic T-lymphocyte epitope within BK and JC virus VP1 polypeptides.** *J Virol* 2005, **79**:11170-11178.
6. Tesoro-Cruz E, Hernández-González R, Kretschmer-Schmid R, Aguilar-Setién A: **Cross-reactivity between caprine arthritis-encephalitis virus and type 1 human immunodeficiency virus.** *Arch Med Res* 2003, **34**:362-366.
7. Wedemeyer H, Mizukoshi E, Davis AR, Bennink JR, Rehermann B: **Cross-reactivity between hepatitis C virus and Influenza A virus determinant-specific cytotoxic T cells.** *J Virol* 2001, **75**:11392-11400.

8. Stevanovic S: **Antigen processing is predictable: From genes to T cell epitopes.** *Transpl Immunol* 2005, **14**:171-174.
9. Lundegaard C, Nielsen M, Lund O: **The validity of predicted T-cell epitopes.** *Trends Biotechnol* 2006, **24**:537-538.
10. Yewdell JW, Bennink JR: **Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses.** *Annu Rev Immunol* 1999, 51-88.
11. Larsen MV, Lundegaard C, Lamberth K, Buus S, Brunak S, Lund O, Nielsen M: **An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions.** *Eur J Immunol* 2005, **35**:2295-2303.
12. Doytchinova IA, Guan P, Flower DR: **EpiJen: a server for multistep T cell epitope prediction.** *BMC Bioinformatics* 2006, **7**:131.
13. Micheletti F, Canella A, Vertuani S, Marastoni M, Tosi L, Volinia S, Traniello S, Gavioli R: **Supra-agonist peptides enhance the reactivation of memory CTL responses.** *J Immunol* 2000, **165**:4264-4271.
14. Reali E, Guerrini R, Marastoni M, Tomatis R, Masucci MG, Traniello S, Gavioli R: **A single specific amino acid residue in peptide antigens is sufficient to activate memory CTL: potential role of cross-reactive peptides in memory T cell maintenance.** *J Immunol* 1999, **162**:106-113.
15. **Basic Local Alignment Search Tool** [<http://blast.ncbi.nlm.nih.gov/Blast.cgi>]

16. **The Protein Data Bank** [<http://www.rcsb.org/pdb/home/home.do>]
17. McSparron H, Blythe MJ, Zygouri C, Doytchinova IA, Flower DR: **JenPep: a novel computational information resource for immunobiology and vaccinology.** *J Chem Inf Comput Sci* 2003, **43**:1276-1287.
18. Brehm MA, Pinto AK, Daniels KA, Schneck JP, Welsh RM, Selin LK: **T cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens.** *Nat Immunol* 2002, **3**:627-634.
19. Schattner, A: **Consequence or coincidence? The occurrence, pathogenesis and significance of autoimmune manifestations after viral vaccines.** *Vaccine* 2005, **23**:3876-3886.
20. Seifert U, Marañón C, Shmueli A, Desoutter JF, Wesoloski L, Janek K, Henklein P, Diescher S, Andrieu M, de la Salle H, Weinschenk T, Schild H, Laderach D, Galy A, Haas G, Kloetzel PM, Reiss Y, Hosmalin A: **An essential role for tripeptidyl peptidase in the generation of an MHC class I epitope.** *Nat Immunol* 2003, **4**:375-379.
21. Joshi SK, Suresh PR, Chauhan VS: **Flexibility in MHC and TCR recognition: degenerate specificity at the T cell level in the recognition of promiscuous Th epitopes exhibiting no primary sequence homology.** *Immunol* 2001, **166**:6693-6703.
22. Sandalova T, Michaëlsson J, Harris RA, Odeberg J, Schneider G, Kärre K, Achour A: **A structural basis for CD8+ T cell-dependent recognition of non-**

homologous peptide ligands: implications for molecular mimicry in autoreactivity. *J Biol Chem* 2005, **280**:27069-27075.

23. **Bioinformatics Toolkit – Max Plank Institute for Developmental Biology**
[<http://toolkit.tuebingen.mpg.de/patsearch>]

24. Zhou T, Xu L, Dey B, Hessel AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, Burton DR, Dimitrov DS, Sodroski J, Wyatt R, Nabel GJ, Kwong PD: **Structural definition of a conserved neutralization epitope on HIV-1 gp120.** *Nature* 2007, **445**:732-737.

25. Nelson DL, Cox MM: *Lehninger - Principles of Biochemistry*. Fourth Edition. H. Freeman & Company. New York 2004: 78.

26. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp Ser* 1999, **41**:95-98.

27. **The Universal Virus Database of the International Committee on Taxonomy of Viruses** [<http://www.ncbi.nlm.nih.gov/ICTVdb>]

28. Nielsen M, Lundegaard C, Lund O, Kesmir, C: **The role of the proteasome in generating cytotoxic T cell epitopes: Insights obtained from improved predictions of proteasomal cleavage.** *Immunogenetics* 2005, **57**:33-41.

29. **Prediction Algorithm for Proteasomal Cleavages**
[<http://www.paproc2.de/paproc1/paproc1.html>].

30. York IA, Brehm MA, Zendzian S, Towne CF, Rock KL: **Endoplasmic reticulum aminopeptidase 1 (ERAP1) trims MHC class I-presented peptides**

in vivo and plays an important role in immunodominance. *Proc Natl Acad Sci U S A* 2006, **103**:9202-9207.

31. Reits E, Neijssen J, Herberts C, Benckhuijsen W, Janssen L, Drijfhout JW, Neefjes J: **A major role for TPPII in trimming proteasomal degradation products for MHC class I antigen presentation.** *Immunity* 2004, **20**:495-506.

32. **Geno3D** [http://geno3d-pbil.ibcp.fr/cgi-bin/geno3d_automat.pl?page=/GENO3D/geno3d_home.html]

33. Guex N, Peitsch MC: **Swiss Model and the Swiss-PdbViewer: an environment for comparative protein modeling.** *Electrophoresis* 1997, **18**:2714-2723.

34. Belz GT, Xie W, Altman, JD, Doherty, PC: **A Previously Unrecognized H-2Db-Restricted Peptide Prominent in the Primary Influenza A Virus-Specific CD8+ T-Cell Response Is Much Less Apparent following Secondary Challenge.** *J Virol* 2000, **74**:3486-3493.

35. Brehm MA, Pinto AK, Daniels KA, Schneck JP, Welsh RM, Selin LK: **T cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens.** *Nat Immunol* 2002, **3**:627-634.

36. Mylin LM, Bonneau RH, Lippolis JD, Tevethia SS: **Hierarchy among multiple H-2b-restricted cytotoxic T-lymphocyte epitopes within simian virus 40 T antigen.** *J Virol* 1995, **69**:6665-6677.

37. Oldstone MB, Lewicki H, Homann D, Nguyen C, Julien S, Gairin JE: **Common Antiviral Cytotoxic T-Lymphocyte Epitope for Diverse Arenaviruses.** *J Virol* 2001, **75**:6273-6278.
38. Snyder JT, Belyakov IM, Dzutsev A, Lemonnier F, Berzofsky JA: **Protection against Lethal Vaccinia Virus Challenge in HLA-A2 Transgenic Mice by Immunization with a Single CD8+ T-Cell Peptide Epitope of Vaccinia and Variola Viruses.** *J Virol* 2004, **78**:7052-7060.
39. Gold MC, Munks MW, Wagner M, Koszinowski UH, Hill AB, Fling SP: **The Murine Cytomegalovirus Immunomodulatory Gene m152 Prevents Recognition of Infected Cells by M45-Specific CTL But Does Not Alter the Immunodominance of the M45-Specific CD8 T Cell Response In Vivo.** *J Immunol* 2002, **169**:359-365.
40. Jaye A, Herberts CA, Jallow S, Atabani S, Klein MR, Hoogerhout P, Kidd M, van Els CA, Whittle HC: **Vigorous but Short-Term Gamma Interferon T-Cell Responses against a Dominant HLA-A*02-Restricted Measles Virus Epitope in Patients with Measles.** *J Virol* 2003, **77**:5014-5016.
41. Wills MR, Carmichael AJ, Mynard K, Jin X, Weekes MP, Plachter B, Sissons JG: **The human cytotoxic T-lymphocyte (CTL) response to cytomegalovirus is dominated by structural protein pp65: frequency, specificity, and T- cell receptor usage of pp65-specific CTL.** *J Virol* 1996, **70**:7569-7579.
42. Tourdot S, Herath S, Gould KG: **Characterization of a new H-2Dk-restricted epitope prominent in primary influenza A virus infection.** *J Gen Virol* 2001, **82**:1749-1755.

43. van der Most RG, Murali-Krishna K, Whitton JL, Oseroff C, Alexander J, Southwood S, Sidney J, Chesnut RW, Sette A, Ahmed R: **Identification of Db- and Kb-Restricted Subdominant Cytotoxic T-Cell Responses in Lymphocytic Choriomeningitis Virus-Infected Mice.** *Virology* 1998, **240**:158-167.
44. Acierno PM, Newton DA, Brown EA, Maes LA, Baatz JE, Gattoni-Celli S: **Cross-reactivity between HLA-A2-restricted FLU-M1:58–66 and HIV p17 GAG:77–85 epitopes in HIV-infected and uninfected individuals.** *J Transl Med.* 2003, **1**:3.
45. Welsh RM, Selin LK: **No one is naive: the significance of heterologous T-cell immunity.** *Nat Rev Immunol* 2002, **2**:417-426.
46. Okazaki T, Pendleton CD, Lemonnier F, Berzofsky JA: **Epitope-Enhanced Conserved HIV-1 Peptide Protects HLA-A2-Transgenic Mice Against Virus Expressing HIV-1 Antigen.** *J Immunol* 2003, **171**:2548-2555.
47. Sette A, Sidney J, Bui HH, del Guercio MF, Alexander J, Loffredo J, Watkins DI, Mothé BR: **Characterization of the peptide-binding specificity of Mamu-A*11 results in the identification of SIV-derived epitopes and interspecies cross-reactivity.** *Immunogenetics* 2005, **57**:53-68.
48. Maeda K, West K, Toyosaki-Maeda T, Rothman AL, Ennis FA, Terajima M: **Identification and analysis for cross-reactivity among hantaviruses of H-2b-restricted cytotoxic T-lymphocyte epitopes in Sin Nombre virus nucleocapsid protein.** *J Gen Virol* 2004, **85**:1909-1919.

49. Lee KY, Chun E, Kim NY, Seong BL: **Characterization of HLA-A2.1-restricted epitopes, conserved in both Hantaan and Sin Nombre viruses, in Hantaan virus-infected patients.** *J Gen Virol* 2002, **83**:1131-1136.

Figure Legend:

Figure 1: mechanisms of viral epitope selection under the bottleneck effect. **A** native protein; **B** proteolysis; **C** TAP binding and translocation; **D** anchoring in the MHC cleft.

Table 1: Epitope alignment considering their putative anchor residues

VIRUS – (viral sequence)	Peptide				MHC restriction	EP	References
	AR		AR				
01) IV – (<i>PA</i> ₂₂₄₋₂₃₃)	SS	L	ENFRAY	V	H-2D	0.3675	[34]
02) PV – (<i>NP</i> ₁₆₋₂₅)	RG	L	SNWTHP	V	H-2D	0.3675	[35]
03) SV40 – (<i>T antigen</i> ₂₀₆₋₂₁₅)	SA	I	NNYAQK	L	H-2D	0.3675	[36]
04) LCMV – (<i>NP</i> ₁₁₈₋₁₂₆)	R	P	QASGVY	M	H-2L	0.2450	[37]
05) Mopeia Virus – (<i>NP</i> ₁₁₈₋₁₂₆)	R	P	LAAGVY	M	H-2L	0.2450	[37]
06) Lassa Fever Virus – (<i>NP</i> ₁₁₈₋₁₂₆)	R	P	LSSGVY	M	H-2L	0.2450	[37]
07) Variola – (<i>HRP2</i> ₇₄₋₈₂)	K	V	DDTFYY	V	HLA-A	0.2450	[38]
08) Vaccinia – (<i>HRP2</i> ₇₄₋₈₂)	K	V	DDTFYY	V	HLA-A	0.2450	[38]
09) mCMV – (<i>M45</i> ₉₈₅₋₉₉₃)	H	G	IRNASF	I	H-2D	0.2450	[39]
10) Measles – (<i>MV-C</i> ₈₄₋₉₂)	K	L	WESPQE	I	HLA-A	0.2450	[40]
11) IV – (<i>M1</i> ₁₂₈₋₁₃₅)		M	GLIYNR	M	H-2K	0.1225	[34]
12) Vaccinia – (<i>VETFsm</i> ₄₉₈₋₅₀₆)	V	L	PFDIKK	L	HLA-A and H-2D	0.2450	[38]
13) Variola – (<i>VETFsm</i> ₄₉₈₋₅₀₆)	V	L	PFDIKK	L	HLA-A and H-2D	0.2450	[38]
14) Variola – (<i>A26L</i> ₆₋₁₄)	N	L	WNGIVP	M	HLA-A	0.2450	[38]
15) HCV – (<i>NS3</i> ₁₀₇₃₋₁₀₈₁)	C	V	NGVCWT	V	HLA-A	0.2450	[7]
16) IV – (<i>NA</i> ₂₂₃₋₂₃₁)	C	V	NGSCFT	V	HLA-A	0.2450	[7]
17) CMV – (<i>pp65</i> ₄₉₅₋₅₀₄)	N	L	VPMVAT	V	HLA-A	0.3675	[41]
18) IV – (<i>PB1</i> ₃₄₉₋₃₅₇)		A	RLKGY	M	H-2D	0.2450	[42]
19) LCMV – (<i>GP</i> ₉₂₋₁₀₁)	CS	A	NNSHHY	I	H-2D	0.3675	[43]
20) SV40 - Synthetic peptide	S	M	IKNLEY	M	H-2D	0.2450	[36]
21) SV40 – (<i>T antigen</i> ₄₈₉₋₄₉₇)	Q	G	INNLDN	L	H-2D	0.2450	[36]
22) HIV-1 – (<i>GAG</i> ₇₇₋₈₅)	S	L	YNTIAV	L	HLA-A	0.2450	[44]
23) IV – (<i>M1</i> ₅₈₋₆₆)	G	I	LGFVFT	L	HLA-A	0.2450	[44]
24) EBV – (<i>BMLF1</i> ₂₈₀₋₂₈₈)	G	L	CTLVAM	L	HLA-A	0.2450	[45]
25) HIV-1 – (<i>RT</i> ₁₇₉₋₁₈₇)	V	L	YQYMDD	V	HLA-A	0.2450	[46]
26) SIV – (<i>Env</i> ₄₉₅₋₅₀₂)		G	DYKLVE	I	Mamu-A	0.1225	[47]
27) SNV – (<i>NC</i> ₁₈₀₋₁₈₈)	S	M	PTAQST	M	H-2D	0.2450	[48]
28) SNV – (<i>NC</i> ₃₃₂₋₃₄₁)	AI	L	QDMRNT	I	H-2D	0.3675	[48]
29) HTNV – (<i>NC</i> ₃₂₈₋₃₃₅)		L	GAFFSI	L	H-2K	0.1225	[48]
30) HTNV – (<i>NC</i> ₃₃₄₋₃₄₂)	I	L	QDMRNT	I	HLA-A	0.2450	[49]
31) PUU – (<i>NC</i> ₁₈₀₋₁₈₈)	S	M	PTAQST	M	NC	0.2450	[48]

AR: Anchor Residues
EP: Expected Probability

μ= 25.7%

Table 2: Sequence comparison using the BLAST *protein-protein* program

Organism	Bacterial sequences	Human sequences	Non-related sequences	Viral Related Sequences
01) IV – (<i>PA</i> ₂₂₄₋₂₃₃)	80%		80% (fungus)	Different strains of IVA
02) PV – (<i>NP</i> ₁₆₋₂₅)				Arenaviridae Family
03) SV40 – (<i>T antigen</i> ₂₀₆₋₂₁₅)	80%	88%	80% (helminth)	Polyomaviridae Family
04) LCMV – (<i>NP</i> ₁₁₈₋₁₂₆)				Arenaviridae Family
05) Mopeia virus – (<i>NP</i> ₁₁₈₋₁₂₆)				Arenaviridae Family
06) Lassa Fever Virus – (<i>NP</i> ₁₁₈₋₁₂₆)			80% (plant)	Arenaviridae Family
07) Variola – (<i>HRP</i> ₂₇₄₋₈₂)				Poxviridae Family
08) Vaccinia – (<i>HRP</i> ₂₇₄₋₈₂)				Poxviridae Family
09) mCMV – (<i>M45</i> ₉₈₅₋₉₉₃)	88%			
10) Measles – (<i>MV-C</i> ₈₄₋₉₂)				Different strains of measles virus
11) IV – (<i>M1</i> ₁₂₈₋₁₃₅)	87%		87% (fish)	Orthomyxoviridae Family
12) Vaccinia – (<i>VETFsm</i> ₄₉₈₋₅₀₆)	88%			Poxviridae Family
13) Variola – (<i>VETFsm</i> ₄₉₈₋₅₀₆)	88%			Poxviridae Family
14) Variola – (<i>A26L</i> ₆₋₁₄)			88% (fungus)	Poxviridae Family
15) HCV – (<i>NS3</i> ₁₀₇₃₋₁₀₈₁)				Flaviviridae Family
16) IV – (<i>NA</i> ₂₂₃₋₂₃₁)	87%		87% (insect and cnidarian)	Different strains of IVA
17) CMV – (<i>pp65</i> ₄₉₅₋₅₀₄)				Herpesviridae Family
18) IV – (<i>PB1</i> ₃₄₉₋₃₅₇)	88%			Orthomyxoviridae e Retroviridae Family
19) LCMV – (<i>GP</i> ₉₂₋₁₀₁)				Arenaviridae Family
20) SV40 - <i>Synthetic peptide</i>				
21) SV40 – (<i>T antigen</i> ₄₈₉₋₄₉₇)			88% (yeast and fungus)	Polyomaviridae Family
22) HIV-1 – (<i>GAG</i> ₇₇₋₈₅)	88%		88% (fungus)	Herpesviridae e Retroviridae Family
23) IV – (<i>M1</i> ₅₈₋₆₆)	88%			Different strains of IVA
24) EBV – (<i>BMLF1</i> ₂₈₀₋₂₈₈)	88%			Herpesviridae that infect primates
25) HIV-1 – (<i>RT</i> ₁₇₉₋₁₈₇)				Differents proteins of HIV
26) SIV – (<i>Env</i> ₄₉₅₋₅₀₂)	87%		87% (fungus)	Retroviridae e Poxviridae
27) SNV – (<i>NC</i> ₁₈₀₋₁₈₈)				Bunyaviridae
28) SNV – (<i>NC</i> ₃₃₂₋₃₄₁)	80%		80% (moss)	Bunyaviridae
29) HTNV – (<i>NC</i> ₃₂₈₋₃₃₅)	87%	87%	87% (mammals, mollusca, protozoan and fungus)	Bunyaviridae
30) HTNV – (<i>NC</i> ₃₃₄₋₃₄₂)	88%			Bunyaviridae
31) PUU – (<i>NC</i> ₁₈₀₋₁₈₈)				Bunyaviridae
Total = 31	15 = 48.39%	2 = 6.45%	11 = 35.48%	

*Significant sequence similarity was defined as an identity value above 80%
 Updated in March, 2008.

Table 3: Proteolytic simulation and structural data of epitopes

Epitopes from Table 1							
Exact cut in carboxy terminal extremity		Structure					
Yes	No	carboxy			amino		
		α -helix	β -sheet	loop	α -helix	β -sheet	loop
26/30	4/30	1/8	4/8	3/8	2/8	-	6/8
86.67%	13.33%	12.5%	50%	37.5%	25%	0%	75%

Epitopes from databank							
Exact cut in carboxy terminal extremity		Structure					
Yes	No	carboxy			amino		
		α -helix	β -sheet	loop	α -helix	β -sheet	loop
50/68	18/68	4/18	3/18	11/18	8/27	1/27	18/27
73.53%	26.47%	22.22%	16.66%	61.11%	29.62%	3.7%	66.7%

Epitope 48 was excluded since it is a synthetic epitope and therefore it is impossible to define flanking sequences, which define cleavage sites.

Table 4: Peptide MHC restriction

MHC – I			
HUMAN		MURINE	
HLA-A	H-2K	H-2D	H-2L
14/14	2/17	12/17	3/17
100%	11.76%	70.59%	17.65%

Peptides 12 and 13 are capable of binding as in HLA-A as in H-2D and were included from the allele counting in human beings as well as from murines.

Peptide 26 was not included because it is the only specific of simian.

Peptide 31 did not present data about allele specificity in the original article.

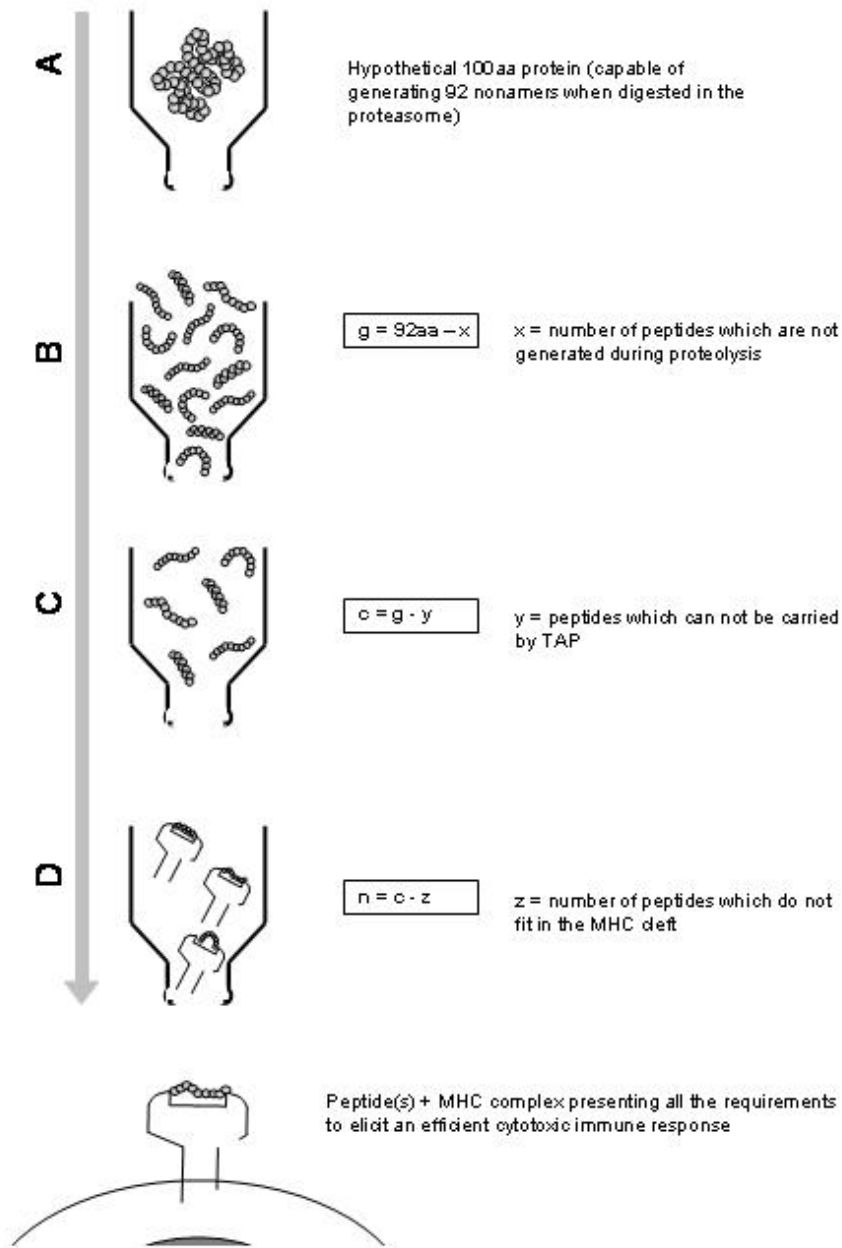


Figure 1

Additional file 1: Supplementary Table 5

Table 5: Outputs from the AntiJen Data Bank (2007, December), containing the number of sequences that showed each query motif.			
Peptide		TAP ligands	
Putative AR [R groups]		Putative AR [R groups]	
[Non polar-aliphatic]	X(6)	[Non polar-aliphatic]	839
[Aromatic]		[Aromatic]	28
[Polar Uncharged]		[Polar Uncharged]	07
[Positively Charged]		[Positively Charged]	32
[Negatively Charged]		[Negatively Charged]	00
[Non polar-aliphatic]		[Aromatic]	609
[Non polar-aliphatic]		[Polar Uncharged]	34
[Non polar-aliphatic]		[Positively Charged]	43
[Non polar-aliphatic]		[Negatively Charged]	15
[Aromatic]		[Polar Uncharged]	16
[Aromatic]		[Positively Charged]	08
[Aromatic]		[Negatively Charged]	00
[Polar Uncharged]		[Positively Charged]	03
[Polar Uncharged]		[Negatively Charged]	01
[Positively Charged]		[Negatively Charged]	30
[Negatively Charged]		[Non polar-aliphatic]	10
[Negatively Charged]		[Polar Uncharged]	00
[Negatively Charged]		[Positively Charged]	02
[Negatively Charged]		[Aromatic]	08
[Polar Uncharged]		[Non polar-aliphatic]	36
[Aromatic]		[Non polar-aliphatic]	26
[Positively Charged]		[Non polar-aliphatic]	62
[Positively Charged]		[Aromatic]	18
[Polar Uncharged]		[Aromatic]	25
[Positively Charged]		[Polar Uncharged]	00

AR= anchor residue

Additional file 2: Supplementary Table 6

Table 6: Flanking Sequence comparison using the BLAST *protein-protein* program

Organism	Sequence	Bacterial sequences	Human sequences	Non-related sequences
01) IV – (<i>PA</i> ₁₉₁₋₂₀₀)	frqsergeet	80%		80% (insect and plant)
02) PV – (<i>NP</i> ₃₄₁₋₃₅₀)	iqqpvrrnggs		80%	80% (mammal, plant and fungus)
03) SV40 – (<i>T antigen</i> ₅₅₁₋₅₆₀)	dylkhclers			80% (insect and helminth)
04) LCMV – (<i>NP</i> ₉₁₋₁₀₀)	vgrlsaeelm	80%		80-90% (mammal, aves, fish, plant and protozoan)
05) Mopeia virus – (<i>NP</i> ₅₅₁₋₅₆₀)	prpvlsrdlv	80%		80-90% (mammal, insect and plant)
06) Lassa Fever Virus – (<i>NP</i> ₃₃₋₄₂)	kdaqallhgl	80%		80-90% (plants and protozoan)
07) Variola – (<i>HRP2</i> ₁₁₋₂₀)	ingdialrnl			80% (insect and protozoan)
08) Vaccinia – (<i>HRP2</i> ₁₂₁₋₁₃₀)	skkykvkeen			80% (simian, aves, insect, plant, protozoan and fungus)
09) mCMV – (<i>M45</i> ₆₁₋₇₀)	qtattaagga	80%		80% (mammal, insect, arachnid, helminth, plant, protozoan, alga, fungus, yeast)
10) Measles – (<i>MV-C</i> ₆₁₋₇₀)	ldqkavhla	80%		80% (insect, plant and fungus)
11) IV – (<i>M1</i> ₂₃₁₋₂₄₀)	dnllenlqay	80%		80% (mammal, insect, protozoan and yeast)
12) Vaccinia – (<i>VETFsm</i> ₆₅₋₇₄)	fkkyvilvnp	80%		80% (mammal and yeast)
13) Variola – (<i>VETFsm</i> ₁₈₀₋₁₈₉)	nlmseetidf	80%		80% (amphibian)
14) Variola – (<i>A26L</i> ₁₉₉₋₂₀₈)	yheiydgnvg	80-90%		80% (mammal and insect)
15) HCV – (<i>NS3</i> ₇₇₅₋₇₈₄)	awyikgrlvp		80%	80-90% (insect and fungus)
16) IV – (<i>NA</i> ₆₁₋₇₀)	ntnflaeqav	80%		80% (mammal and fungus)
17) CMV – (<i>pp65</i> ₄₃₀₋₄₃₉)	stsagrkrks			80% (mammal, insect, helminth, plant and fungus)
18) IV – (<i>PB1</i> ₉₀₋₉₉)	eamafleesh			80% (mammal)
19) LCMV – (<i>GP</i> ₄₃₁₋₄₄₀)	stplalmdll	80%		80% (fish, amphibian, cnidarian and fungus)
20) SV40 - Synthetic peptide				
21) SV40 – (<i>T antigen</i> ₅₅₈₋₅₆₇)	ersefllekr	80%	80%	80% (mammal and fungus)
22) HIV-1 – (<i>GAG</i> ₆₁₋₇₀)	lgqlqpalht	80%		80% (mammal and yeast)
23) IV – (<i>M1</i> ₂₃₁₋₂₄₀)	dnllenlqay	80%		80-90% (mammal, insect, protozoan and yeast)
24) EBV – (<i>BMLF1</i> ₁₁₋₂₀)	ssissnedpa	80%		80-90% (mammal, helminth and fungus)
25) HIV-1 – (<i>RT</i> ₄₆₃₋₄₇₂)	rqkivsltet		80%	80-90% (mammal, amphibian, insect and protozoan)
26) SIV – (<i>Env</i> ₆₄₆₋₆₅₅)	fleenitall			80% (mammal, amphibian, plant, protozoan and fungus)
27) SNV – (<i>NC</i> ₅₄₋₆₃)	ledkladfkr	80%		80% (mammal and fungus)
28) SNV – (<i>NC</i> ₁₃₁₋₁₄₀)	lpiilkalym	80%		80% (plant, moss, protozoan and fungus)
29) HTNV – (<i>NC</i> ₅₁₋₆₀)	aasiqakide	80-90%		80% (insect, helminth, plant, soil amoeba, protozoan and fungus)
30) HTNV – (<i>NC</i> ₆₁₋₇₀)	fkrrmadavs	80%		80-90% (plant and protozoan)
31) PUU – (<i>NC</i> ₁₁₋₂₀)	itrheqqlvv	80%		80% (insect, protozoan and fungus)

Total = 31

Additional file 3: Supplementary Epitope Databank

Virus	Viral sequence	Host	CD4/CD8	MHC	Reference	Epitope	N°
Influenza	NP ₃₆₆₋₃₇₄	<i>Homo sapiens</i>	CD8	H2-D ^b	Kedzierska <i>et al.</i> , PNAS April 6, (2004) vol. 101 no. 14 (001)	ASNENMETM immunodominant	001
Measles	MV-C ₈₄₋₉₂	<i>Homo sapiens</i>	CD8	HLA-A*O2	Jaye <i>et al.</i> , J Virol Apr. (2003), p. 5014–5016	KLWESPQEI immunodominant	002
Vaccinia	VETFsm ₄₉₈₋₅₀₆	<i>Homo sapiens</i>	CD8	HLA-A2 e H2 Db	Snyder <i>et al.</i> J VIROL July (2004), p. 7052–7060	VLPFDIKKL Cross-reactive	003
Vaccinia	A26L ₆₋₁₄	<i>Homo sapiens</i>	CD8	HLA-A2 e H2 Db	Snyder <i>et al.</i> J VIROL July (2004), p. 7052–7060	NLWNGIVPT Cross-reactive	004
Vaccinia*	HRP2 ₇₄₋₈₂	<i>Homo sapiens</i>	CD8	HLA-A2	Snyder <i>et al.</i> J VIROL July (2004), p. 7052–7060	KVDDTFYYV Cross-reactive	005
Variola	VETFsm ₄₉₈₋₅₀₆	<i>Homo sapiens</i>	CD8	HLA-A2	Snyder <i>et al.</i> J VIROL July (2004), p. 7052–7060	Idêntica a de Vaccinia (004) Cross-reactive	006
Variola	A26L ₆₋₁₄	<i>Homo sapiens</i>	CD8	HLA-A2	Snyder <i>et al.</i> J VIROL July (2004), p. 7052–7060	Homóloga (NLWNGIVPM) Cross-reactive	007
Variola	HRP2 ₇₄₋₈₂	<i>Homo sapiens</i>	CD8	HLA-A2	Snyder <i>et al.</i> J VIROL July (2004), p. 7052–7060	Idêntica à de Vaccinia (006) Cross-reactive	008
mCMV	M45 ₉₈₅₋₉₉₃	<i>Mus musculus (C57Bl/6)</i>	CD8	H2-D ^b	Gold <i>et al.</i> J Immunol (2002) 169:359–365	HGIRNASFI immunodominant	009
HCV	NS3 ₁₀₇₃	<i>Homo sapiens</i>	CD8	HLA-A2	Wedemeyer <i>et al.</i> J Virol Dec. (2001), p. 11392–11400	CVNGVCWTV Cross-reactive immunodominant	010
IV	NA ₂₃₁	<i>Homo sapiens</i>	CD8	HLA-A2	Wedemeyer <i>et al.</i> J Virol Dec. (2001), p. 11392–11400	CVNGSCFTV Cross-reactive	011
HPV44	L1 ₃₁₅	<i>Homo sapiens</i>	CD8	HLA-A2	Wedemeyer <i>et al.</i> J Virol Dec. (2001), p. 11392–11400	HNNGICWGN Epitope	012
CMV	pp65 ₄₉₅₋₅₀₄	<i>Homo sapiens</i>	CD8	HLA-A2	Wills <i>et al.</i> J Virol, Nov. (1996), p. 7569–7579	NLVPMVATVQ consensus HLA A2 binding motif Peptide dominant (pp65)	013

PV	NP ₁₆₋₂₅	<i>Mus musculus</i>	CD8	H-2D ^b	Brehm <i>et al</i> , Nat Immun Vol3 n7 July (2002) P627-634	RGLSNWTHPV Immunodominant	014
PV	NP ₃₈₋₄₅	<i>Mus musculus</i>	CD8	H-2K ^b	Brehm <i>et al</i> , Nat Immun Vol3 n7 July (2002) P627-634	SALDFHKV immunodominant	015
PV	NP ₂₀₅₋₂₁₂	<i>Mus musculus</i>	CD8	H-2K ^b	Brehm <i>et al</i> , Nat Immun Vol3 n7 July (2002) P627-634	YTVKFPNM weak subdominant	016
VSV	N ₅₃₋₅₉	<i>Mus musculus</i>	CD8	H-2K ^b	Bleek <i>et al</i> , Nature Vol 348 (1990), p213- 216	GYVYQGL immunodominant	017
IV	NP ₃₉₋₄₇	<i>Mus musculus</i>	CD8	H-2K ^d	Chen <i>et al</i> , Immunity 12 (2000), p83-93	FYIQMCTEL Putative immunodominant	018
IV	NP ₁₄₇₋₁₅₅	<i>Mus musculus</i>	CD8	H-2K ^d	Chen <i>et al</i> , Immunity 12 (2000), p83-93	TYQRTRALV “mais forte dos 3” Putative immunodominant	019
IV	HA ₅₁₈₋₅₂₆	<i>Mus musculus</i>	CD8	H-2K ^d	Chen <i>et al</i> , Immunity 12 (2000), p83-93	IYSTVASSL Putative immunodominant	020
IV	PA ₂₂₄₋₂₃₃	<i>Mus musculus</i>	CD8	H-2D ^b	Belz <i>et al</i> , J Virol, Vol 74 (2000), p3486- 3493	SSLENFRAYV Epitope	021
IV	NS ₂₁₁₄₋₁₂₁	<i>Mus musculus</i>	CD8	H-2K ^b	Belz <i>et al</i> , J Virol, Vol 74 (2000), p3486- 3493	RTFSFQLI Epitope	022
IV	M1 ₁₂₈₋₁₃₅	<i>Mus musculus</i>	CD8	H-2K ^b	Belz <i>et al</i> , J Virol, Vol 74 (2000), p3486- 3493	MGLIYNRM (tentar S do Blast em humans) Epitope	023
IV	HA ₁₂₅₉₋₂₆₆	<i>Mus musculus</i>	CD8	H-2K ^k	Gould <i>et al</i> , J Virol, Vol 65 (1991), p5401- 5409	FEANGNLI Epitope (putative immunodominant)	024
IV	HA ₂₁₀₋₁₈	<i>Mus musculus</i>	CD8	H-2K ^k	Gould <i>et al</i> , J Virol, Vol 65 (1991), p5401- 5409	IEGGWTGMI Epitope (putative immunodominant)	025
IV	NP ₅₀₋₅₇	<i>Mus musculus</i>	CD8	H-2K ^k	Gould <i>et al</i> , J Virol, Vol 65 (1991), p5401- 5409	SDYEGRLI Epitope	026
IV	NS ₁₁₅₂₋₁₆₀	<i>Mus musculus</i>	CD8	H-2 ^b	Belz <i>et al</i> , J Immunol, Vol 165 (2000), p2404-2409	EEGAIVGEI Epitope	027
IV	PB ₁₃₄₉₋₃₅₇	<i>Mus musculus</i>	CD8	H-2D ^K	Tourdot <i>et al</i> , J Gen Virol, Vol 82 (2001), P1749-1755	ARLGKGYMF Immunodominant a D ^K -restricted peptide motif, consisting of arginine at position two, arginine or lysine at position five and a hydrophobic residue at the carboxy terminus.	028
Polyoma	MT ₃₈₉₋₃₉₇	<i>Mus</i>	CD8	H-2D ^K	Lukacher <i>et al</i> , J Immunol, Vol	RRLGRTLLL immunodominant	029

Virus		<i>musculus</i>			160 (1998), p1724-1734		
LCMV	NP ₃₉₆₋₄₀₄	<i>Mus musculus</i>	CD8	H-2D ^b	Gairin <i>et al</i> , J Virol, Vol 69 (1995), p 2297-2305	FQPQNGQFI Immunodominant	030
LCMV	GP1 ₃₃₋₄₁	<i>Mus musculus</i>	CD8	H-2D ^b and H-2K ^b	Gairin <i>et al</i> , J Virol, Vol 69 (1995), p 2297-2305. Hudrisier <i>et al</i> , Virology, Vol 234 (1997), p62-73. ("see picture") Sandalova <i>et al</i> , JBC, Vol 280 (2005), p27069-27075.	KAVYNFATC Immunodominant or GP1 ₃₃₋₄₃ KAVYNFATCGI depending on the cell line used to test binding and lysis. KAVYNFATC/M Cross reactive with self epitope murine KALYDYAPI (dopamine beta mono oxygenase)	031
LCMV	NP ₂₀₅₋₂₁₂	<i>Mus musculus</i>	CD8	H-2K ^b	Van der Most <i>et al</i> , Virology, Vol 240 (1998), P158-167	YTVKYPNL Subdominant	032
LCMV	GP ₉₂₋₁₀₁	<i>Mus musculus</i>	CD8	H-2D ^b	Van der Most <i>et al</i> , Virology, Vol 240 (1998), P158-167	CSANNSHHYI Subdominant	033
SV40	Synthetic peptide	<i>Mus musculus</i>	CD8	H-2D ^b	Mylin <i>et al</i> , J Virol, Vol 69 (1995), p 6665-6677	SMIKNLEYM optimal H-2D ^b binding motif	034
SV40	T antigen 206-215	<i>Mus musculus</i>	CD8	H-2D ^b	Mylin <i>et al</i> , J Virol, Vol 69 (1995), p 6665-6677	SAINNYAQKL Epitope	035
SV40	T antigen 223-231	<i>Mus musculus</i>	CD8	H-2D ^b	Mylin <i>et al</i> , J Virol, Vol 69 (1995), p 6665-6677	CKGVNKEYL Epitope	036
SV40	T antigen 404-411	<i>Mus musculus</i>	CD8	H-2K ^b	Mylin <i>et al</i> , J Virol, Vol 69 (1995), p 6665-6677	VVYDFLKC Epitope	037
SV40	T antigen 489-497	<i>Mus musculus</i>	CD8	H-2D ^b	Mylin <i>et al</i> , J Virol, Vol 69 (1995), p 6665-6677	QGINLNDNL Epitope	038
IV	M1 ₅₈₋₆₆	<i>Homo sapiens</i>	CD8	HLA-A2	Acierno <i>et al</i> , J Transl Med, Vol 1 (2003), p3	GILGFVFTL Cross-reactive with 054 Strecht of 023	039
HIV-1	GAG ₇₇₋₈₅	<i>Homo sapiens</i>	CD8	HLA-A2	Acierno <i>et al</i> , J Transl Med, Vol 1 (2003), p3	SLYNTI AVL Cross-reactive with 053	040
IV	PB1 ₇₀₃₋₇₁₁	<i>Mus musculus</i>	CD8	H-2D ^b	Belz <i>et al</i> , J Immunol, Vol 166 (2001), p 4627-33	SSYRRPVGI Immunodominant	041
IV	PB2 ₁₉₈₋₂₀₆	<i>Mus musculus</i>	CD8	H-2D ^b	Belz <i>et al</i> , J Immunol, Vol 166 (2001), p 4627-33	ISPLMVAYM Cross-reactive mimotope with 055	042
IV	NP ₃₆₆₋₃₇₄	<i>Mus musculus</i>	CD8	H-2D ^b	Haanen <i>et al</i> , J Exp Med, Vol 190 (1999), p 1319-1328	ASNENMDAM Antigenic variant of 001	043

HIV-1y	GAG ₂₄₀₋₂₄₉	<i>Homo sapiens</i>	CD8	HLA-B57 and HLA-B5801	Bertoletti <i>et al</i> , J Virol, Vol 72 (1998), p2439-2448	TSTLQEIQGW Immunodominant and cross-reactive with 059	044
HIV-2	GAG ₂₄₁₋₂₅₀	<i>Homo sapiens</i>	CD8	HLA-B5801	Bertoletti <i>et al</i> , J Virol, Vol 72 (1998), p2439-2448	TSTVEEQIQW Immunodominant and cross-reactive with 058	045
HPV16	E7 ₁₁₋₁₉	<i>Homo sapiens</i>	CD8	HLA-A2	Nilges <i>et al</i> , J Virol, Vol 77 (2003), p5464-5474	YMLDLQPET Cross-reactive with 061	046
Coronavirus	NS ₂₅₂₋₆₀	<i>Homo sapiens</i>	CD8	HLA-A2	Nilges <i>et al</i> , J Virol, Vol 77 (2003), p5464-5474	TMLDIQPED Cross-reactive with 060	047
LCMV	NP ₁₁₈₋₁₂₆	<i>Homo sapiens</i>	CD8	H-2L ^d	Oldstone <i>et al</i> , J Virol, Vol 75 (2001), p6273-6278	RPQASGVYM Cross-reactive with 063 and 064	048
Mopeia virus	NP ₁₁₈₋₁₂₆	<i>Homo sapiens</i>	CD8	H-2L ^d	Oldstone <i>et al</i> , J Virol, Vol 75 (2001), p6273-6278	RPLAAGVYM Cross-reactive with 062 and 064	049
Lassa Fever Virus(LFV)	NP ₁₁₈₋₁₂₆	<i>Homo sapiens</i>	CD8	H-2L ^d	Oldstone <i>et al</i> , J Virol, Vol 75 (2001), p6273-6278	Strain (G) RPLSSGVYM Strain (J) RPLSAGVYM Cross-reactive with 062 and 063	050
Sabia Virus	NP ₁₁₈₋₁₂₆	<i>Homo sapiens</i>	CD8	H-2L ^d	Oldstone <i>et al</i> , J Virol, Vol 75 (2001), p6273-6278	RSASGGYYL No cross-reactive	051
Dengue Virus	NS ₃₂₉₈₋₃₀₆	<i>Homo sapiens</i>	CD8	H-2K ^d	Spaulding <i>et al</i> , J Virol, Vol 73 (1999), p398-403	GYISTRVEM type 4 GYISTRVGM type 1 NS₃₂₉₉₋₃₀₇ GYISTRVEM type 2 GYISTRVGM type 3 NS₃₂₉₉₋₃₀₇	052
HIV-1	GAG ₂₄₀₋₂₄₉ “TW10”	<i>Homo sapiens</i>	CD8	HLA-B57 HLA-B58	Stephens <i>et al</i> , Trends Immunol, Vol 26 (2005), p41-47	TSTLQEQIAW Immunodominant	053
HIV-1	Nef ₇₃₋₈₂	<i>Homo sapiens</i>	CD8	HLA-A11	Stephens <i>et al</i> , Trends Immunol, Vol 26 (2005), p41-47	QVPLRPMTYK Immunodominant See mutations article	054
MVE (Murray Valley Encephal)	NS ₃₁₇₈₅	<i>Mus musculus</i>	CD8	H-2K ^K	Regner <i>et al</i> , J Immunol, Vol 166 (2001), p3820-3828	REHSGNEI Immunodominant	055
MVE (Murray Valley Encephal)	NS ₃₁₉₇₁	<i>Mus musculus</i>	CD8	H-2K ^K	Regner <i>et al</i> , J Immunol, Vol 166 (2001), p3820-3828	DEGEGRVI Immunodominant	056
EBV	BMLF1 ₂₈₀₋₂₈₈	<i>Homo sapiens</i>	CD8	HLA-A2	Welsh <i>et al</i> , Nature Immunol, Vol 2 (2002), p417-426	GLCTLVAML Immunodominant	057
HIV-1	RT ₁₇₉₋₁₈₇	<i>Homo sapiens</i>	CD8	HLA-A2.1	Okazaki <i>et al</i> , J Immunol, Vol 171 (2003), p 2548-55	VIYQYMDDL Highly conserved Low affinity to the HLA-A2.1 VLYQYMDDV	058

						Enhanced epitope	
SIV	Env ₄₉₅₋₅₀₂	<i>Macaca mulatta</i>	CD8	Mamu-A*11	Mothé <i>et al</i> , Immunogen, Vol 57 (2005), p53-68	GDYKLVEI Epitope	059
SNV	NC _{94-101*}	<i>Mus musculus</i>	CD8	H-2K ^b	Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	SSLRYGNV Epitope	060
SNV	NC ₁₈₀₋₁₈₈	<i>Mus musculus</i>	CD8	H-2D ^b	Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	SMPTAQSTM Epitope	061
SNV	NC ₂₁₈₋₂₂₆	<i>Mus musculus</i>	CD8	H-2K ^b	Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	PVMGVIGFS Epitope	062
SNV	NC ₃₃₁₋₃₄₀ NC ₃₃₂₋₃₄₁	<i>Mus musculus</i>	CD8	H-2D ^b	Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	FAILQDMRNT AILQDMRNTI Epitope	063
HTNV	NC ₃₂₈₋₃₃₅	<i>Mus musculus</i>	CD8	H-2K ^b	Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	LGAFFSIL Epitope	064
HTNV	NC ₃₃₄₋₃₄₂	<i>Homo sapiens</i>	CD8	HLA-A2	Lee <i>et al</i> , J of General Virology, Vol 83 (2002), p 1131-1136	ILQDMRNTI Epitope (elicit strong CTL activity)	065
HTNV	NC ₂₂₁₋₂₂₈	<i>Mus musculus</i>	CD8	H-2K ^b	Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	SVIGFLAL Immunodominant	066
PUU	NC ₉₄₋₁₀₁		CD8		Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	SSLRYGNV Cross-reactive with 075	067
PUU	NC ₁₈₀₋₁₈₈		CD8		Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	SMPTAQSTM Cross-reactive with 076	068
PUU	NC ₂₁₈₋₂₂₆		CD8		Maeda K. <i>et al</i> , J of General Virology, Vol 85 (2004), p 1909-1919	PVMGVIGFS Cross-reactive with 077	069

Additional file 4: PDB Identifier from used proteins

<i>N° on databank:</i>	<i>PDB Identifier or modeled structures:</i>
001	2IQH – PDB
010	1DXP – PDB
011	2HTY – PDB
012	L1– Geno 3D (model 1)
017	2GIC – PDB
018	2IQH – PDB
019	2IQH – PDB
023	1EA3 - PDB
024	1RU7 – PDB
025	1RUZ – PDB
026	2IQH – PDB
027	2GX9 – PDB
035	2FUF – PDB
036	2FUF – PDB
037	1SVL – PDB
038	1N25 – PDB
039	1EA3 - PDB
040	1L6N – PDB
043	NP–Geno 3D (model 2)
044	1L6N – PDB
045	GAG(HIV2) – Geno 3D (model 3)
052	2BMF – PDB
053	GAG(HIV) – Geno3D (model 4)
054	2NEF – PDB
055	NS3(MVE) – Geno3D (model 5)
056	NS3(MVE) – Geno3D
058	1DLO – PDB (model 5)

“Nós somente possuímos convicções sob a condição de nada termos estudado plenamente.”

E.M. Cioran (1973)

CAPÍTULO IV

Discussão



Desde a elaboração da nossa idéia (descrita no capítulo II) até a montagem do banco e a análise dos epitopos (apresentados no capítulo III), algumas observações importantes foram feitas sobre as relações entre epitopos virais e destes com o sistema imune. A seguir estas observações serão sintetizadas:

- a análise inicial, através da comparação das características físico-químicas dos aminoácidos das seqüências dos peptídeos, demonstrou que os epitopos realmente compartilham algumas particularidades (motivo encontrado entre os resíduos de ancoragem ao MHC), sendo esse o primeiro indício para a construção de um perfil completo de identificação de seqüências de origem viral;

- os epitopos reconhecidamente imunogênicos possuem os requerimentos para um bom processamento antigênico e, portanto, são positivamente selecionados ao longo desse processo;

- como existe uma co-evolução entre os mecanismos que geram epitopos (ou seja, os mecanismos de clivagem pelo proteossomo e imunoproteossomo, translocação por TAP, ligação ao MHC) e o histórico de infecções virais, é particularmente plausível que os epitopos imunodominantes contêm as características comentadas anteriormente;

- não só as seqüências dos epitopos (os fragmentos de 8-12 aminoácidos que efetivamente estão ligados à fenda peptídica), mas também as seqüências que flanqueiam essas regiões são importantes em estudos que visam a compreensão dos mecanismos de geração dos epitopos e, por exemplo, em estudos que visam a identificação de alvos que possam ser utilizados em vacinas;

- o sistema imune seleciona como alvos, seqüências que são tipicamente de origem viral, e que não apresentam (ou apresentam poucas) relações de identidade com seqüências filogeneticamente distintas, principalmente seqüências oriundas de organismos superiores (as quais poderão estar envolvidas no desenvolvimento de autoimunidades).

Embora essas conclusões sejam de extrema relevância e corroborem nossas suposições iniciais, o que obtivemos foi apenas um esboço inicial das relações de similaridade entre os alvos selecionados pelo sistema imune. Essa limitação das análises se deu principalmente porque as ferramentas até então disponíveis (ferramentas de alinhamento) permitiu-nos no máximo, uma categorização grosseira das seqüências lineares.

Seria de extrema valia se pudéssemos acessar toda a topologia dos complexos MHC:peptídeos formados, para que mais níveis de similaridade pudessem ser avaliados (inclusive de peptídeos que são aparentemente não relacionados). Com isso, esse tipo de abordagem pode auxiliar, inclusive, na inferência de reatividade cruzada entre dois epitopos virais, para serem utilizados no desenvolvimento de vacinas de amplo espectro.

Como discutimos anteriormente, esse tipo de estudo é possível desde que existam as estruturas cristalografadas tanto de moléculas do MHC como de peptídeos depositadas em bancos de proteínas. Para o nosso objetivo, o ideal seria que pudéssemos realizar as análises em qualquer epitopo desejado, associado ao alelo de MHC de interesse.

Em parte, essa limitação já está sendo suplantada por nosso grupo. Estamos construindo uma abordagem de bioinformática em que é possível inferir a estrutura tridimensional, e conseqüentemente a topologia de cargas, de qualquer epitopo, a partir de sua seqüência de aminoácidos. Esse tipo de ferramenta também está sendo disponibilizado, ao menos em carácter inicial e experimental, por outros grupos, com algumas diferenças experimentais (Todman *et al.*, 2008; Madurga *et al.*, 2005).

A abordagem supracitada compreende algumas técnicas reconhecidas de bioinformática [combinações de docagem molecular (molecular docking) e minimizações de energia, entre outros passos]. A única limitação é que para poder montar o complexo desejado, é necessário que o alelo do MHC em questão possua sua estrutura cristalografada depositada nos bancos de proteínas (PDB, por exemplo). A partir desse ponto é possível obter a estrutura tridimensional de qualquer epitopo complexado ao MHC desejado (metodologia no anexo I). Além disso é possível inferir se algum peptídeo apresenta afinidade considerável ao MHC que está sendo estudado, uma questão fundamental na apresentação de antígenos. É claro que devemos salientar que as ferramentas de bioinformática são criadas para abreviar o número de candidatos que devem ser testados em experimentos *in vitro* e *in vivo* e não para substituir esses passos fundamentais do método científico.

Para demonstrar a importância do desenvolvimento desse tipo de técnica em um estudo de compartilhamento de características, ou inferência de reatividade cruzada, citaremos um exemplo envolvendo dois epitopos, sendo um derivado do vírus influenza (PA224-233) e outro do vírus da hepatite B (HbsAg30-

39) (figura 5). Quando realizamos o alinhamento desses dois epitopos levando em consideração apenas as suas seqüências de aminoácidos, o valor de identidade encontrado é de apenas 20% (figura 5A). Quando classificamos os aminoácidos das seqüências de acordo com as características físico-químicas das suas cadeias laterais (tipo de categorização que utilizamos no alinhamento dos peptídeos do banco de epitopos), a identidade sobe para 70% (figura 5B). Com isso aumenta-se a sensibilidade em detectar seqüências similares que poderiam estar sendo ignoradas. Esse tipo de categorização das cadeias laterais de acordo com as características físico-químicas, buscando inferir reatividade cruzada entre epitopos aparentemente não relacionados foi recentemente corroborada no trabalho de Frankild *et al.* (2008), o qual demonstrou que substituições que preservam essas características parecem manter a capacidade desses epitopos em elicitar respostas imunes.

Ainda assim, quando desejamos descobrir todas as nuances compartilhadas entre alvos virais, e além disso, quando um de nossos objetivos é encontrar padrões que possam ser utilizados como alvos para sensibilizar o sistema imune contra diferentes vírus, a simples seqüência linear dos aminoácidos ainda é um parâmetro grosseiro. Assim, a possibilidade de comparar a estrutura tridimensional de dois epitopos complexados a um mesmo alelo de MHC, permite-nos que acessemos diferentes parâmetros simultaneamente (a distribuição de cargas na região do epitopo, a forma dos resíduos que realizam o contato com o TCR, etc) e possamos assim realizar inferências mais precisas.

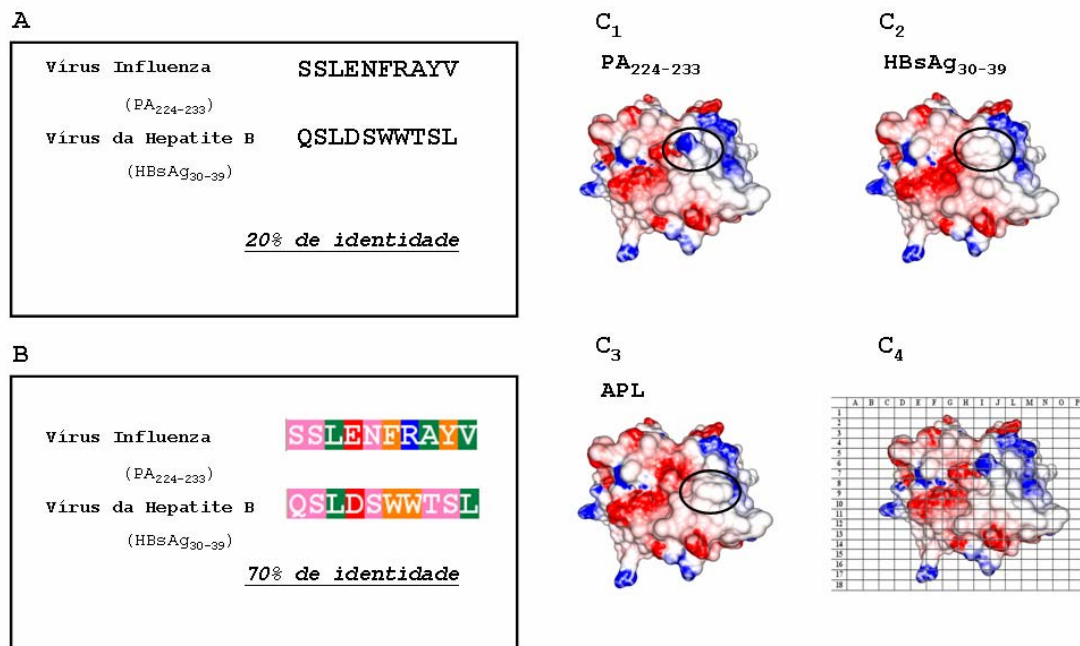


Figura 5: Diferentes níveis de acesso à verificação do compartilhamento de identidade entre os epitopos. Em A observamos dois epitopos aparentemente não relacionados, apresentando apenas 20% de identidade de seqüência. Na parte inferior, em B, os aminoácidos das duas seqüências foram classificados de acordo com suas características físico-químicas (não-polares na cor verde, aromáticos na cor laranja, polares não-carregados na cor rosa, carregados positivamente na cor azul e carregados negativamente na cor vermelha). Em C são apresentadas às topologias de cargas de estruturas de MHCs murinos (H2D^b) ligados a diferentes peptídeos. Azul, carga positiva; vermelho, carga negativa; branco, neutro. C₁ é a estrutura cristalografada 1WBY de H2D^b com epitopo PA₂₂₄₋₂₃₃ do vírus Influenza. C₂ e C₃ são estruturas tridimensionais obtidas através de nosso processo de inferência de complexos MHC-peptídeo e referem-se respectivamente ao complexo H2D^b ligado ao epitopo do vírus da hepatite B (HbsAg₃₀₋₃₉) e H2D^b ligado ao APL (ligante peptídico alterado). A comparação das estruturas C₁ e C₂ demonstra estruturas muito parecidas, com exceção da área das elipses, onde ocorre uma diferença de cargas, a qual é corrigida na construção do APL (C₃). C₄ é a grade de comparação de cores que foi montada para comparar duas estruturas diferentes para facilitar a co-localização de suas topologias (dados não publicados).

No caso dos epitopos do vírus influenza e da hepatite B, a comparação estrutural seria inviável se dependêssemos apenas de dados cristalográficos.

Entretanto, como um dos epitopos (PA224-233 do vírus Influenza) possuía sua estrutura ancorada a um MHC de classe I murino (H2Db), depositada no Protein Data Bank sob o código 1WBY, conseguimos modelar o epitopo do vírus da hepatite B (HbsAg30-39) ligado a esse mesmo alelo do MHC. O resultado desse experimento permitiu-nos vislumbrar toda a área de contato com o TCR dessa região do complexo MHC: peptídeo (figura 5C). As topologias são muito similares em toda a extensão das duas estruturas, à exceção de uma diferença de cargas conferida pela presença de aminoácidos pertencentes a grupos diferentes na posição sete do epitopo (elipses em C1 e C2). O epitopo do vírus influenza apresenta uma arginina nessa posição, a qual lhe confere uma carga positiva (C1), enquanto que no epitopo do vírus da hepatite B, o triptofano tem uma carga neutra, porém um volume mais proeminente (C2).

Não se pode dizer se essa diferença comprometeria o reconhecimento de um destes complexos MHC:peptídeo por um linfócito T que tivesse sido sensibilizado pelo outro complexo MHC:peptídeo. No entanto, sabe-se que diferenças que envolvem troca de cargas parecem ser mais importantes para a abolição de reatividade cruzada do que alterações na forma dos epitopos (Kessels *et al.*, 2004). Se objetivássemos utilizar a possível reatividade cruzada entre esses dois epitopos para o desenvolvimento de vacinas, uma estratégia que poderíamos adotar seria a de utilização de um ligante peptídico alterado (APL) que fosse capaz de induzir uma imunização contra os dois alvos simultaneamente.

Os APLs são ligantes que apresentam variações em relação à seqüência original de determinado peptídeo e são capazes de induzir diferentes níveis de

resposta em um linfócito T (desde o desencadeamento de todas funções citotóxicas até a indução de um estado anérgico total) (Sloan-Lancaster & Allen, 1996). O APL que desenhamos (SSLENFWAIV) tem como base a seqüência do epitopo de influenza trocando-se a arginina da posição 7 por um triptofano. Esta configuração anula a carga que estava sendo o diferencial entre os dois epitopos (figura 5C3), mas mantém uma topologia conservada. Não só o problema da carga é resolvido através do uso deste APL como é também possível verificar que esse APL tem afinidade pelo alelo do MHC de interesse (H2Db), através da comparação dos valores de docking das três estruturas (dados não apresentados). Estes dados foram obtidos em colaboração com o estudante de Biomedicina da UFRGS Dinler Amaral Antunes.

Como podemos observar na figura 5, a capacidade de comparação entre os dois epitopos apresenta um crescente ao longo das três abordagens exemplificadas (figura 5A, 5B e 5C). No entanto, na técnica que permite a comparação das estruturas tridimensionais, apesar do aumento na riqueza dos detalhes observados ser evidente, a quantificação da identidade entre duas estruturas comparadas não é precisa. Para realizar as comparações usamos um sistema de grade com coordenadas (figura 5C4). Esse sistema pode ser refinado permitindo a comparação das imagens pixel a pixel.

Com o estabelecimento dessa técnica, temos à disposição diferentes metodologias de comparação de peptídeos. Assim, podemos voltar nossa atenção para outros fatores que podem estar envolvidos na seleção de epitopos virais. Uma idéia interessante é a de analisar proteínas virais completas, para que possamos investigar quais pontos dessas proteínas estão sendo

preferencialmente “escolhidos” como alvos pelo sistema imune. Com isso, pretendemos entender quais fatores estão influenciando para que essas regiões sejam escolhidas.

Para testar essa hipótese escolhemos estudar um grupo viral no qual já havíamos observado a ocorrência de epitopos com seqüências de aminoácidos conservadas entre diferentes representantes de um mesmo gênero. O gênero escolhido foi o Hantavírus pertencente à família Bunyaviridae e a proteína alvo foi a proteína do nucleocapsídeo (NC). Entre as características interessantes dos representantes do gênero Hantavírus podemos citar sua distribuição geográfica ampla (global) e o fato de já haver descrição de reatividade cruzada entre representantes desse grupo (Maeda *et al.*, 2004).

Resumidamente, foi feito um alinhamento das seqüências completas da proteína de NC de 34 diferentes Hantavirus e assinaladas as regiões de ocorrência de epitopos descritos para três integrantes desse gênero (vírus haantan - HTN, vírus puumala - PUU e vírus sin nombre - SNV) (figura 6). Interessantemente, observamos que os epitopos localizavam-se em regiões conservadas da proteína. Essa “colocalização” corrobora uma de nossas hipóteses iniciais de que o sistema imune pode estar direcionando suas respostas contra alvos que possuem uma restrição mutacional evidente. De fato, duas dessas regiões que compreendem epitopos têm função reconhecida na viabilidade desses organismos. A região que vai do aminoácido 1-125 está envolvida com a oligomerização da proteína NC enquanto que a região entre 175-217 é um local de interação com o RNA (Kaukinen *et al.*, 2005).

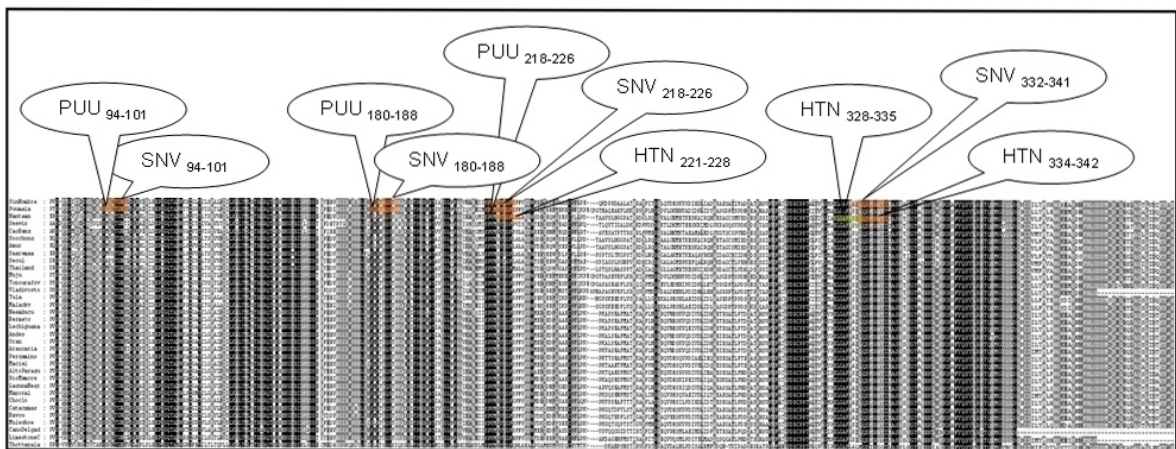


Figura 6: **Alinhamento da proteína de nucleocapsídeo de 34 diferentes Hantavírus.** Os balões assinalam os locais de ocorrência de epitopos descritos para vírus haantan - HTN, vírus puumala - PUU e vírus sin nombre – SNV. A figura foi propositalmente reduzida para permitir uma visão esquemática da proteína, apresentando blocos de alta conservação (colunas mais escuras) e regiões de menor similaridade de seqüências (colunas mais claras). Observe que os epitopos co-localizam-se com regiões de alta conservação (dados não publicados).

Outra questão que sempre abordamos é de que o sistema imune deve direcionar suas respostas para regiões que além de ser típicas de vírus, sejam relativamente diferentes de proteínas próprias (do hospedeiro). Para testar esta hipótese utilizamos uma técnica, que envolve a ferramenta BLAST, denominada “janela deslizante”. Esta estratégia permitiu-nos verificar similaridades entre todos os possíveis decâmeros da proteína NC de SNV e proteínas de roedores (hospedeiro natural de Hantavirus). Escolhemos a proteína de SNV por ser a que possui mais epitopos descritos dentre os Hantavirus, incluídos no nosso trabalho. Mais uma vez os epitopos localizaram-se em regiões que possuíam baixa similaridade com proteínas do hospedeiro (dados não apresentados). Estes dados

foram obtidos em colaboração com o estudante de Biomedicina da UFRGS Maurício Menegatti Rigo.

Para finalizar, os locais preditos como pontos de clivagem pelo proteassomo encontravam-se em posições altamente conservadas (em ambas as extremidades amino e carbóxi terminal), ou sobre resíduos que apresentavam substituições por outros aminoácidos pertencentes ao mesmo grupo físico-químico do aminoácido predominante, evidenciando que as regiões adjacentes aos epitopos são também fundamentais na seleção de epitopos. Outra evidência relativa à necessidade da conservação das seqüências flanqueadoras dos fragmentos epitopos é o fato de mutações nesses locais poderem abortar a existência de um determinado epitopo. Mutações nestes pontos são, portanto, mais uma estratégia de escape viral (Yeh *et al.*, 2006).

Podemos deduzir que, aparentemente, tanto regiões conservadas (incluindo o epitopo e suas seqüências adjacentes) como regiões que possuem baixa similaridade com proteínas próprias são preferencialmente selecionadas para a geração de epitopos imunodominantes. Isso favorece a ocorrência de reatividade cruzada contra diferentes vírus além de diminuir o risco do surgimento de eventos autoimunes.

De posse das diferentes ferramentas de comparação de epitopos que possuímos e de outras ferramentas que simulam passos do processamento de antígenos, associadas a outros conhecimentos que temos incorporado sobre as forças que atuam na seleção de alvos virais pelo sistema imune ao longo desse trabalho, pretendemos aumentar cada vez mais nosso banco de epitopos, em

busca de bons marcadores de infecção viral. A partir daí, pretendemos selecionar os melhores candidatos desses testes in silico (candidatos altamente imunogênicos, possuidores de características para o bom processamento antigênico e com grande potencial para indução de reatividade cruzada), os quais poderão passar para uma segunda fase de testagem: in vitro e in vivo. Nossa perspectiva é a de criar um sistema de fluxo contínuo de eleição e testagem de epitopos candidatos a serem utilizados no desenvolvimento de vacinas de DNA. A etapa in vivo envolverá a construção de plasmídeos contendo politopos (grupos de epitopos virais) para estimular a resposta celular e humoral, além de memória de longo prazo, em um modelo experimental de camundongos. Desta forma postulamos que será possível induzir imunidade simultaneamente contra diferentes proteínas de um mesmo vírus ou mesmo contra diferentes patógenos.

“O que quer que um outro disser bem, é meu.”

Sêneca (Século I d.C.)

Referências



Acierno PM, Newton DA, Brown EA, Maes LA, Baatz JE and Gattoni-Celli S (2003) Cross-reactivity between HLA-A2-restricted FLU-MI: 58-66 and HIV p17 GAG: 77-85 epitopes in HIV-infected and uninfected individuals. *J Transl Med* 1:3.

Adams AB, Williams MA, Jones TR, Shirasugi N, Durham MM, Kaech SM, Wherry EJ, Onami T, Lanier JG, Kokko KE, Pearson TC, Ahmed R and Larsen CP (2003) Heterologous immunity provides a potent barrier to transplantation tolerance. *J Clin Invest* 111:1887-1895.

Arstila TP, Casrouge A, Baron V, Even J, Kanellopoulos J and Kourilsky P (1999) A direct estimate of the human alphabeta T cell receptor diversity. *Science* 286:958-961.

Beismann-Driemeyer S and Tampé R (2004) Function of the transporter associated with antigen processing (TAP) in cellular immunity, tumor escape, and virus persistence. In: Broer S and Wagner CA (eds) *Membrane Transporter Diseases*. 1st edition. Springer-Verlag, New York, pp320-337.

Belz GT, Xie W and Doherty PC (2001) Diversity of epitope e cytokine profiles for primary and secondary influenza A virus-specific CD8+ T cell responses. *J Immunology* 166:4627-4633.

Bertoletti A, Cham F, McAdam S, Rostrom T, Rowland-Jones S, Sabally S, Corrah T, Ariyoshi K and Whittle H (1998) Cytotoxic T cells from Human Immunodeficiency virus type 2- Infected patients frequently cross-react with different human immunodeficiency type I clades. *J Virol* 72:2439-2448.

Brehm MA, Markees TG, Daniels KA, Greiner DL, Rossini AA and Welsh RM (2003) Direct visualization of cross-reactive effector and memory allo-specific CD8 T cells generated in response to viral infections. *J Immunol* 170:4077-4086.

Brehm MA, Pinto AK, Daniels KA, Schneck JP, Welsh RM and Selin LK (2002) T cell immunodominance and maintenance of memory regulated by unexpectedly cross-reactive pathogens. *Nat Immunol* 3:627-634.

Brehm MA, Selin LK and Welsh RM (2004) CD8+ T cell responses to viral infections in sequence. *Cell Microbiol* 6:411-421.

Christen U and von Herrath MG (2004) Induction, acceleration or prevention of autoimmunity by molecular mimicry. *Mol Immunol* 40:1113-1120.

Currier JR, Harris ME, Cox JH, McCutchan FE, Birx DL, Maayan S and Ferrari G (2005) Immunodominance and cross-reactivity of B5703-restricted CD8 T lymphocytes from HIV type 1 subtype C-infected Ethiopians. *AIDS Res Hum Retroviruses* 21:239-245.

Daftarian P, Sharan R, Haq W, Ali S, Longmate J, Termini J and Diamond DJ (2005) Novel conjugates of epitope fusion peptides with CpG-ODN display enhanced immunogenicity and HIV recognition. *Vaccine* 23:3453-3468.

de Carvalho Nicacio C, Gonzalez Della Valle M, Padula P, Björling E, Plyusnin A and Lundkvist A (2002) Cross-protection against challenge with Puumala virus after immunization with nucleocapsid proteins from different hantaviruses. *J Virol* 76:6669-6677.

Frankild S, de Boer RJ, Lund O, Nielsen M and Kesmir C (2008) Amino acid similarity accounts for T cell cross-reactivity and for "holes" in the T cell repertoire. *PLoS ONE* 3:e1831.

Garcia KC, Degano M, Pease LR, Huang M, Peterson PA, Teyton L and Wilson IA (1998) Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* 279:1166-1172.

Gergely P Jr, Pullmann R, Stancato C, Otvos L Jr, Koncz A, Blazsek A, Poor G, Brown KE, Phillips PE and Perl A (2005) Increased prevalence of transfusion-transmitted virus and cross-reactivity with immunodominant epitopes of the HRES-1/p28 endogenous retroviral autoantigen in patients with systemic lupus erythematosus. *Clin Immunol* 116:124-134.

Haanen JBAG, Wolkers MC, Kruisbeek AM and Schumacher TNM (1999) Selective expansion of cross-reactive CD8⁺ memory T cells by viral variants. *J Exp Med* 190:1319-1328.

Harris A, Sha B and Luo M (1999) Structural similarities between influenza virus matrix protein M1 and human immunodeficiency virus matrix and capsid proteins: an evolutionary link between negative-stranded RNA viruses and retroviruses. *J Gen Virol* 80:863-869.

Istrail S, Florea L, Halldórsson BV, Kohlbacher O, Schwartz RS, Yap VB, Yewdell JW and Hoffman SL (2004) Comparative immunopeptidomics of humans and their pathogens. *Proc Natl Acad Sci U S A* 101:13268-13272.

Joshi SK, Suresh PR and Chauhan VS (2001) Flexibility in MHC and TCR recognition: degenerate specificity at the T cell level in the recognition of promiscuous Th epitopes exhibiting no primary sequence homology. *J Immunology* 166:6693-6703.

Kaukinen P, Vaheri A and Plyusnin A (2005) Hantavirus nucleocapsid protein: a multifunctional molecule with both housekeeping and ambassadorial duties. *Arch Virol* 150:1693-1713.

Kessels HW, de Visser KE, Tirion FH, Coccoris M, Kruisbeek AM and Schumacher TN (2004) The impact of self-tolerance on the polyclonal CD8⁺ T cell repertoire. *J Immunol* 172:2324-2331.

Lévy F, Burri L, Morel S, Peitrequin AL, Lévy N, Bachi A, Hellman U, Van den Eynde BJ and Servis C (2002) The final N-terminal trimming of a subaminoterminal proline-containing HLA class I-restricted antigenic peptide in the cytosol is mediated by two peptidases. *J Immunol* 169:4161-4171.

Liu SJ, Leng CH, Lien SP, Chi HY, Huang CY, Lin CL, Lian WC, Chen CJ, Hsieh SL and Chong P (2006) Immunological characterizations of the nucleocapsid protein based SARS vaccine candidates. *Vaccine* 24:3100-3108.

Madurga S, Belda I, Llorà X and Giralt E (2005) Design of enhanced agonists through the use of a new virtual screening method: application to peptides that bind class I major histocompatibility complex (MHC) molecules. *Protein Sci* 14:2069-2079.

Maeda K, West K, Toyosaki-Maeda T, Rothman AL, Ennis FA and Terajima M (2004) Identification and analysis for cross-reactivity among hantaviruses of H-2b-restricted cytotoxic T-lymphocyte epitopes in Sin Nombre virus nucleocapsid protein. *J Gen Virol* 85:1909-1919.

Mason D (1998) A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol Today* 19:395-404.

Mazza C, Auphan-Anezin N, Gregoire C, Guimezanes A, Kellenberger C, Roussel A, Kearney A, van der Merwe PA, Schmitt-Verhulst AM and Malissen B (2007) How much can a T-cell antigen receptor adapt to structurally distinct antigenic peptides? *EMBO J* 26:1972-1983.

Michelson S (2004) Consequences of human cytomegalovirus mimicry. *Hum Immunol* 65:465-475.

Moran E, Simmons C, Vinh Chau N, Luhn K, Wills B, Dung NP, Thao le TT, Hien TT, Farrar J, Rowland-Jones S and Dong T (2008) Preservation of a critical epitope core region is associated with the high degree of flaviviral cross-reactivity exhibited by a dengue-specific CD4+ T cell clone. *Eur J Immunol* 38:1050-1057.

Nilges K, Höhn H, Pilch H, Neukirch C, Freitag K, Talbot PJ and Maeurer J (2003) Human papillomavirus type 16 E7 peptide-directed CD8+ T cells from patients with cervical cancer are cross-reactive with the coronavirus NS2 protein. *J Virol* 77:5464-5474.

Reali E, Guerrini R, Marastoni M, Tomatis R, Masucci MG, Traniello S and Gavioli R (1999) A single specific amino acid residue in peptide antigens is sufficient to activate memory CTL: potential role of cross-reactive peptides in memory T cell maintenance. *J Immunol.* 162:106-113.

Regner M (2001) Cross-reactivity in T-cell antigen recognition. *Immunol Cell Biol* 79:91-100.

Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, Janssen H, Calafat J, Drijfhout JW and Neefjes J (2003) Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. *Immunity* 18:97-108.

Ryan KR, Patel SD, Stephens LA and Anderton SM (2007) Death, adaptation and regulation: the three pillars of immune tolerance restrict the risk of autoimmune disease caused by molecular mimicry. *J Autoimmun* 29:262-271.

Sandalova T, Michaelsson J, Harris RA, Odeberg J, Schneider G, Karre K and Achour A (2005) A structural basis for CD8+ T cell-dependent recognition of non-homologous peptide ligands: Implications for molecular mimicry in autoreactivity. *J Biol Chem* 280:27069-27075.

Selin LK and Welsh RM (2004) Plasticity of T cell memory responses to viruses. *Immunity* 20:5-16.

Selin LK, Brehm MA, Naumov YN, Cornberg M, Kim SK, Clute SC and Welsh RM (2006) Memory of mice and men: CD8+ T-cell cross-reactivity and heterologous immunity. *Immunol Rev* 211:164-181.

Selin LK, Cornberg M, Brehm MA, Kim SK, Calcagno C, Ghersi D, Puzone R, Celada F and Welsh RM (2004) CD8 memory T cells: cross-reactivity and heterologous immunity. *Semin Immunol* 16:335-347.

Selin LK, Nahill SR and Welsh RM (1994) Cross-reactivities in memory cytotoxic T lymphocyte recognition of heterologous viruses. *J Exp Med* 179:1933-1943.

Selin LK, Varga SM, Wong IC and Welsh RM (1998) Protective heterologous antiviral immunity and enhanced immunopathogenesis mediated by memory T cell populations. *J Exp Med* 188:1705-1715.

Shimojo N, Maloy WL, Anderson RW, Biddison WE and Coligan JE (1989) Specificity of peptide binding by the HLA-A2.1 molecule. *J Immunology* 143:2939-2947.

Sloan-Lancaster J and Allen PM (1996) Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. *Annu Rev Immunol* 14:1-27.

Sospedra M and Martin R (2006) When T cells recognize a pattern, they might cause trouble. *Curr Opin Immunol* 18:697-703.

Tagaram HR, Watson AM, Lemonnier FA, Staveley-O'Carroll K, Tevethia SS and Schell TD (2008) An SV40 VP1-derived epitope recognized by CD8⁺ T cells is naturally processed and presented by HLA-A*0201 and cross-reactive with human polyomavirus determinants. *Virology* 376:183-190.

Tischler ND, Roseblatt M and Valenzuela PD (2008) Characterization of cross-reactive and serotype-specific epitopes on the nucleocapsid proteins of hantaviruses. *Virus Res* 135:1-9.

Todman SJ, Halling-Brown MD, Davies MN, Flower DR, Kayikci M and Moss DS (2008) Toward the atomistic simulation of T cell epitopes automated construction of MHC: peptide structures for free energy calculations. *J Mol Graph Model* 26:957-961.

Wedemeyer H, Mizukoshi E, Davis AR, Bennink JR and Rehermann B (2001) Cross-reactivity between hepatitis C virus and influenza A virus determinant-specific cytotoxic T cells. *J Virol* 75:11392-11400.

Welsh RM and Selin LK (2002) No one is naive: the significance of heterologous T-cell immunity. *Nat Rev Immunol* 2:417-426.

Welsh RM, Selin LK and Szomolanyi-Tsuda E (2004) Immunological memory to viral infections. *Annu Rev Immunol* 22:711-743.

Yamada K, Takasaki T, Nawa M, Yabe S and Kurane I (2003) Antibody responses determined for Japanese dengue fever patients by neutralization and

hemagglutination inhibition assays demonstrate cross-reactivity between dengue and Japanese encephalitis viruses. *Clin Diagn Lab Immunol* 10:725-728.

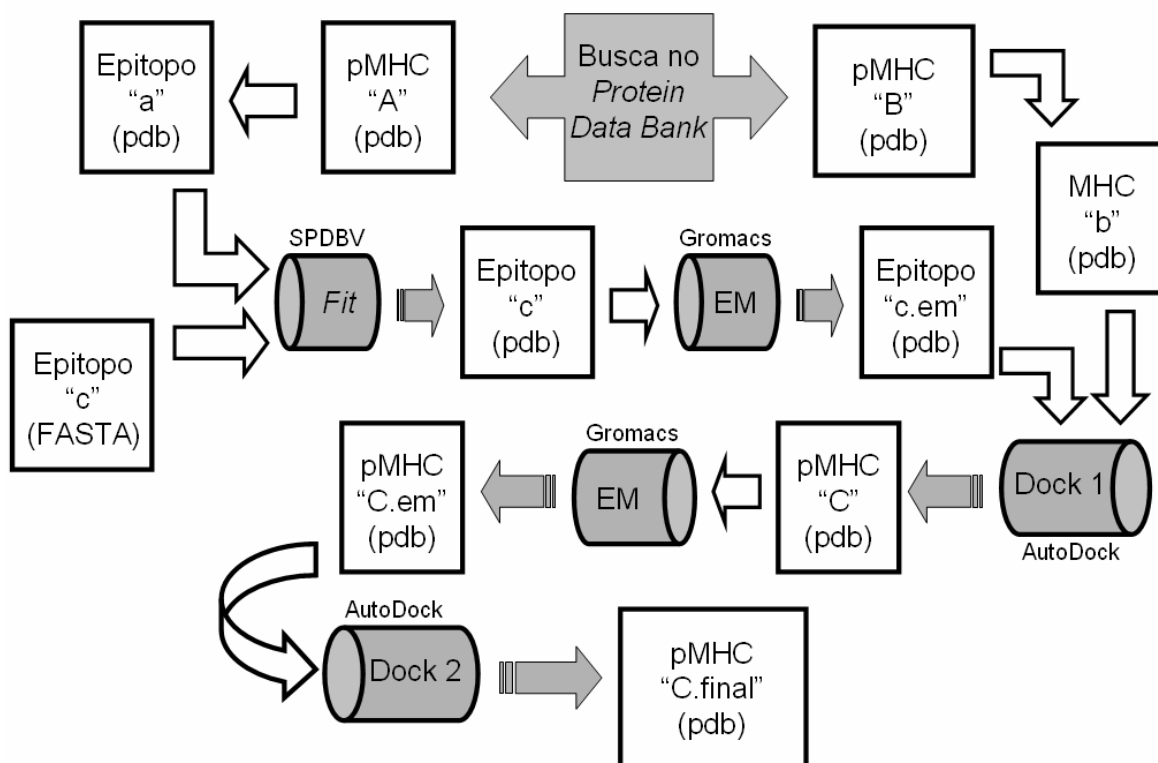
Yeh WW, Cale EM, Jaru-Ampornpan P, Lord CI, Peyerl FW and Letvin NL (2006) Compensatory substitutions restore normal core assembly in simian immunodeficiency virus isolates with Gag epitope cytotoxic T-lymphocyte escape mutations. *J Virol* 80:8168-8177.

Yellen-Shaw AJ, Laughlin CE, Metrione RM and Eisenlohr LC (1997) Murine transporter associated with antigen presentation (TAP) preferences influence class I-restricted T cell responses. *J Exp Med* 186:1655-1662.

Yewdell JW (2006) Confronting complexity: real world immunodominance in antiviral CD8+ t cell response. *Immunity* 25:533-543.

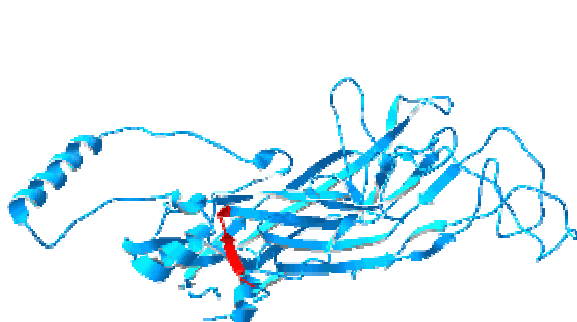
Yewdell JW and Bennink JR (1999) Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu Rev Immunol* 17:51-88.

ANEXO I

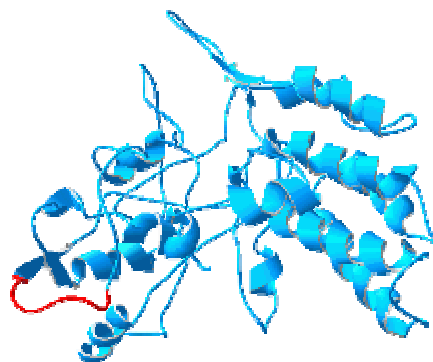


Fluxograma da construção de um novo complexo pMHC. Considere um epitopo "c" cuja estrutura no contexto de um dado alelo de MHC não foi determinada. Busca-se no PDB um arquivo contendo o alelo de interesse apresentando um epitopo com o mesmo número de aminoácidos do epitopo "c", nesse exemplo, encontramos o complexo "A" e o complexo "B". No programa SPDBV realiza-se o *Fit* da sequência de aminoácidos do epitopo "c" sobre a estrutura 3D do epitopo presente no complexo "A". A estrutura gerada do epitopo "c" sofre uma minimização de energia (EM), sendo então utilizada como entrada para o *docking* com a estrutura do MHC de interesse (neste caso retirada do complexo "B"). A fim de ajustar o MHC a este novo epitopo realiza-se uma EM do complexo C. Após a minimização, epitopo e MHC são separados e utilizados como entrada para um segundo *docking*, o qual irá gerar o complexo pMHC desejado (maiores informações vide métodos).

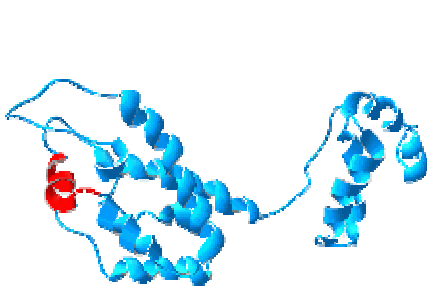
ANEXO II



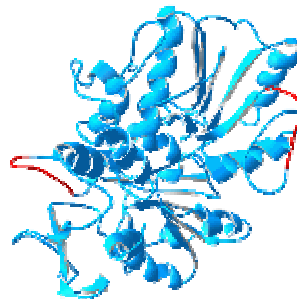
Modelo 1(contendo o epitopo 012)



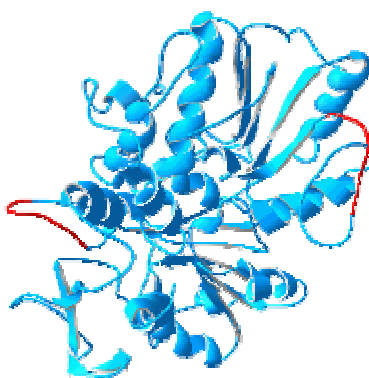
Modelo 2 (contendo o epitopo 043),



Modelo 3 (contendo o epitopo 045)



Modelo 4 (contendo o epitopo 053)



Modelo 5 (contendo os epitopos 055 e 056)

Estruturas obtidas através de modelagem por homologia.