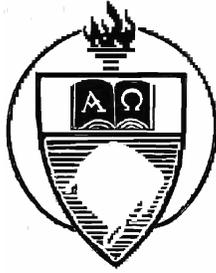


UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
GRADUAÇÃO EM BIOMEDICINA

Utilização de Ferramentas de Bioinformática para
a Análise do Potencial de
Reatividade Cruzada entre Epitopos Virais.

Dinler Amaral Antunes

Porto Alegre, 2008



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
GRADUAÇÃO EM BIOMEDICINA

**Utilização de Ferramentas de Bioinformática para
a Análise do Potencial de
Reatividade Cruzada entre Epitopos Virais**

*Trabalho de Conclusão de Curso apresentado ao curso
de Biomedicina, do Instituto de Ciências Básicas da
Saúde, da Universidade Federal do Rio Grande do Sul,
como um dos pré-requisitos para a obtenção do grau
de Biomédico.*

Dinler Amaral Antunes

Orientador: Prof. Dr. José Artur Bogo Chies

Co-orientador: BSc Gustavo Fioravanti Vieira

Curso de Biomedicina

2008

Agradecimentos

Agradeço, sobretudo, a meus pais, Aristeu e Ana Cristina. Por terem sido o exemplo perfeito, no qual busquei inspiração para definir meu caráter; meu empenho, minha perseverança e meu comprometimento. Agradeço ainda, além do esforço para manter meus estudos, pela educação – em todos os aspectos – e por todo o amor, carinho e confiança depositados em mim. Serei eternamente grato.

Agradeço a meus padrinhos, Éder e Dina, por terem me recebido como um filho em sua casa, durante o cursinho. Especialmente ao tio Éder, com quem dividi apartamento por mais um semestre antes de entrar na UFRGS.

Ao tio Sílvio, pelas dicas em Porto Alegre, pelo carinho, pelos jantares e corridas de Táxi que me presenteou, bem como à tia Adriana, pelos inúmeros telefonemas e todo o envolvimento para garantir a continuidade de meus estudos. Neste sentido, agradeço também a Fernando José Marroni de Abreu pela ajuda fundamental, sem a qual eu não estaria concluindo meu curso neste ano.

Aos demais familiares, pelo afeto, o constante desejo de “boa sorte”, os presentes, os telefonemas...O porto seguro sempre disponível, ainda que distante. Em especial, dedico um muito obrigado a meus avós, pela preocupação e carinho incondicional. A Vó Vanda, pelas discussões filosóficas, confidências e “aulas” de história da família. Aos “mi-mis”, Valmi e Wolmy, pelos telefonemas (sempre que possível), pelas palavras de apoio (sempre que necessário) e inclusive pelos “troquinhos” que me mandaram.

Aos amigos, tanto os antigos quanto os novos. Em um mundo onde a preocupação com o sucesso e os interesses próprios inviabiliza cada vez mais a existência de amizades verdadeiras e duradouras, considero-me uma pessoa abençoada por ter vários grandes amigos.

Aos colegas de curso, por terem sido uma verdadeira família durante a faculdade. Pelo convívio, pelas impagáveis horas de divertimento, pelas festas, pelas jantares, passeios, estudos em grupo e pelo apoio total no momento em que mais precisei. Certamente sem o apoio da turma eu não estaria me formando neste ano – Muito Obrigado!

Aos Confirmados, pela incrível afinidade, pelos vários ótimos encontros e pela amizade que, espero, será mantida pela vida toda.

Aos meus antigos colegas de apartamento, pelo divertido convívio durante mais da metade do curso. Pelas risadas, filmes, jogos, idas em conjunto ao mercado, pelas várias horas de profundas discussões em plena madrugada – Viva ao AP da Biomed!

Ao meu irmão, Jonier. Pela amizade, pelas risadas e por todos os momentos de conversa e troca de conhecimentos.

À Laura, minha namorada. Pelo companheirismo e apoio em todos os momentos, desde o início do curso. Pela atenção, pelo carinho, pelo abraço sempre disponível. Tens sido o meu mais intenso vínculo, a pessoa que mais me conhece. Simplesmente, Te Amo!

Aos Stertz. Por terem me recebido com um carinho e atenção impressionantes. Apesar de nos conhecermos a poucos anos, neste período vocês foram sem dúvida minha segunda família. Agradeço em especial a Suzi, por ter me recebido sem objeções em seu apartamento, onde durante um bom tempo fui praticamente um morador extra.

Aos meus orientadores durante a Iniciação Científica, Zéca e Gustavo, pelos ensinamentos, pela amizade e por acreditarem no meu trabalho.

Aos demais colegas de trabalho, pelo ambiente agradável de amizade e colaboração. Pela não discriminação aos “sem jaleco” da bioinformática. Em especial ao Maurício, meu primeiro grande amigo em Porto Alegre, meu colega de curso e de estágio.

À Duda, pela paciência em responder meus e-mails “quilométricos” e por toda a ajuda técnica, sem a qual, este trabalho não seria possível.

A Deus, por ter me possibilitado conhecer todas estas pessoas a quem estou agradecendo agora e por estar me permitindo alcançar, passo a passo, meus sonhos.

“Toda a nossa ciência, comparada com a realidade, é primitiva e infantil - e, no entanto, é a coisa mais preciosa que temos”.

Albert Einstein (1879 - 1955)

Sumário

Resumo.....	6
Introdução.....	7
Trabalho experimental (artigo).....	15
Histórico.....	16
Resultados.....	18
Análise de geração do epitopo HBsAg ₂₈₋₃₉	18
Identificação de um padrão estrutural alelo específico.....	18
Estabelecimento de uma estratégia para a construção de complexos pMHC.....	19
Estudo <i>in silico</i> do potencial de reatividade cruzada entre os epitopos virais PA ₂₂₄₋₂₃₃ (Influenza) e HBsAg ₂₈₋₃₉ (HBV).....	19
Discussão.....	20
Conclusões.....	25
Métodos.....	26
Estudos de geração de epitopos.....	26
Construção de complexos pMHC.....	26
Construção do complexo D ^b HBsAg ₃₀₋₃₉	27
Construção do APL PA ₂₂₄₋₂₃₃ R7W.....	28
Análise da topologia e da distribuição de cargas.....	28
Obtenção das imagens para o artigo.....	28
Contribuição dos Autores.....	28
Agradecimentos.....	29
Referências.....	30
Figuras.....	33
Figura 1 - Sobreposição de epitopos restritos a H-2D ^b e a H-2K ^b	33
Figura 2 - Padrão conformacional dos epitopos apresentados por H-2D ^b	34
Figura 3 - Organização estrutural dos epitopos apresentados em H-2D ^b	35
Figura 4 - Comparação dos complexos pMHC quanto a topologia e carga.....	36
Figura 5 - Fluxograma da construção de um novo complexo pMHC.....	37
Tabelas.....	38
Tabela 1- Relação dos ligantes de H-2D ^b disponíveis no PDB.....	38
Arquivos Adicionais.....	39
Arquivo Adicional 1 - Exceções ao padrão de H-2D ^b	39
Arquivo Adicional 2 - Relação dos ligantes de H-2K ^b disponíveis no PDB.....	40
Arquivo Adicional 3 - Padrão estrutural dos ligantes de H-2K ^b	41
Arquivo Adicional 4 - Resultados do <i>docking</i> indicam variação em p7 no APL.....	42
Arquivo Adicional 5 - Diferença estrutural do epitopo 1CE6.....	43
Conclusões e Perspectivas.....	44
Referências Adicionais.....	50

Resumo

A reatividade cruzada é definida como a capacidade de um linfócito T em reconhecer, no contexto do Complexo Principal de Histocompatibilidade (MHC), peptídeos não relacionados, provenientes de um mesmo organismo ou mesmo de organismos heterólogos. Este fenômeno já foi descrito em inúmeros trabalhos, embora os mecanismos que permitam este reconhecimento cruzado ainda não tenham sido completamente estabelecidos. O reconhecimento do complexo MHC:peptídeo (pMHC) pelo Receptor de Célula T (TCR) leva à lise da célula apresentadora, o que torna a reatividade cruzada alvo de profundo interesse em estudos que envolvem os mecanismos citotóxicos da resposta imune. Neste trabalho realizamos um estudo *in silico* da possível ocorrência de reatividade cruzada entre os epitopos virais PA₂₂₄₋₂₃₃ (Influenza) e HBsAg₂₈₋₃₉ (HBV) no contexto do MHC murino H-2D^b. O complexo H-2D^b:PA₂₂₄₋₂₃₃ (D^bPA₂₂₄₋₂₃₃) possui estrutura depositada no *Protein Data Bank* (PDB) sob o código de acesso 1WBY, enquanto a estrutura do complexo D^bHBsAg₂₈₋₃₉ ainda não foi determinada. Utilizando o programa AutoDock 4 para realizar o *docking* do epitopo ao MHC e o pacote GROMACS para realizar a minimização de energia (EM) das estruturas geradas, conseguimos construir o complexo D^bHBsAg₃₀₋₃₉. O programa GRASP2 foi utilizado para as análises de topologia e distribuição de cargas do complexo. A comparação entre a estrutura 1WBY e o complexo gerado indicou forte correlação estrutural, contudo identificamos uma diferença de cargas em uma posição crítica para o reconhecimento pelo TCR. Para contornar esta diferença, construímos um epitopo de Influenza mutado (R7W) e repetimos as análises. Este novo peptídeo apresentou alta afinidade pelo MHC e maior semelhança com o epitopo de HBV, possivelmente induzindo reatividade cruzada. A estratégia desenvolvida para a realização deste trabalho pode ser utilizada para simular complexos pMHC formados com qualquer epitopo e o estudo de topologia/potencial do complexo formado nos permite comparar dois diferentes complexos pMHC sob o ponto de vista do TCR, possibilitando discutir reatividade cruzada *in silico*. Em conjunto, estas técnicas apresentam grande potencial de aplicação no estudo de patologias autoimunes e no desenvolvimento de vacinas antivirais de amplo espectro.

Introdução

O sistema imune realiza uma constante vigilância no organismo e respostas imunológicas estão sendo montadas a todo o momento, independentemente da existência de infecções ou processos patológicos. As interações químicas e celulares que formam a base dos processos neste sistema nunca deixam de ocorrer, sendo contudo, reguladas em intensidade de acordo com o estado imunológico do indivíduo e a interação com o ambiente (Coombes & Powrie, 2008). Um dos exemplos mais claros desta constante vigilância é a rota de apresentação de antígenos pelo Complexo Principal de Histocompatibilidade (MHC) de classe I.

O MHC foi descoberto como um *locus*¹ extenso com genes altamente polimórficos, que determinavam o resultado de transplantes de órgãos entre pessoas. Apesar do nome equivocado, herança do contexto em que foi descoberta, a função fisiológica desta molécula não guarda qualquer relação com hiscompatibilidade, sendo responsável pela apresentação de peptídeos na superfície celular. Os genes MHC podem ainda ser subdivididos em pelo menos dois grupos principais – MHC de classe I e MHC de classe II – que codificam proteínas estruturalmente distintas mas homólogas. As moléculas do MHC de classe II estão envolvidas na apresentação de peptídeos exógenos² para reconhecimento pelos linfócitos T CD4+. Elas estão presentes em um grupo restrito de células (p.e. macrófagos, células dendríticas e linfócitos B) conhecidas como apresentadoras profissionais de antígenos (APCs). Em contraste, as moléculas do MHC de classe I estão presentes em todas as células nucleadas do organismo, inclusive nas APCs.

O MHC de classe I é formado por duas cadeias polipeptídicas ligadas de forma não covalente, sendo uma cadeia pesada (cadeia α) e uma subunidade invariável não codificada no mesmo *locus*, chamada β_2 -microglobulina. Os segmentos amino-terminais (N-term) da cadeia α ($\alpha 1$ e $\alpha 2$), interagem formando uma plataforma de oito β -folhas pregueadas na qual se apóiam duas α -hélices paralelas. Este conjunto forma a chamada “fenda” do MHC-I, um espaço com tamanho suficiente para abrigar peptídeos com

¹ Localização física, ou posição, de um gene específico em um cromossomo.

² Peptídeos longos (30 aminoácidos ou mais) oriundos de proteínas extra-celulares que foram endocitadas por fagócitos e degradadas por hidrolases em endossomos.

comprimento entre oito e onze aminoácidos. A molécula de MHC-I totalmente montada é um heterodímero que consiste nas duas cadeias (α e β_2 -microglobulina) e um peptídeo endógeno³ ligado a ela, sendo que a expressão estável deste complexo na superfície celular requer a presença dos três componentes (Hammer e cols, 2008).

O *locus* do MHC abriga ainda outros genes envolvidos na geração desses segmentos protéicos - epitopos - que serão apresentados aos linfócitos T citotóxicos (CD8+) no contexto do MHC-I. Entre eles destacam-se as seqüências codificadoras de um complexo enzimático com atividade proteásica - o proteossomo – e do Transportador Associado ao Processamento de Antígenos (TAP). Uma amostra de todas as proteínas celulares é marcada pelo processo de ubiquitinação⁴ o que as direciona para o proteossomo. Este complexo é formado por um *core* central (20S) altamente conservado e por subunidades regulatórias (19S ou 11S). A subunidade 19S tem alta afinidade pela ubiquitina, sendo ainda capaz de realizar o desdobramento, ATP dependente, da proteína ubiquitinada, direcionando a seqüência de aminoácidos semi-desnaturada para a região central do complexo (Wang & Maldonado, 2006). No *core* a proteína será digerida, liberando ubiquitinas que serão reaproveitadas e peptídeos menores que serão transportados pela TAP para o Retículo Endoplasmático (ER). Os cortes do proteossomo irão gerar com exatidão a extremidade carbóxi-terminal (C-term) dos epitopos apresentados pelo MHC-I, existindo uma “sintonia” entre proteossomo, TAP e MHC em termos de afinidade em C-term por um aminoácido (ou um grupo físico-químico), sendo esta semelhança provável fruto da co-evolução destes que são passos decisivos dentro de uma mesma rota (Paulsson, 2004). A extremidade N-term, no entanto, não é gerada pelo proteossomo, e os peptídeos transportados pela TAP sofrerão ainda um desbastamento amino-terminal – por proteases do ER, como ERAP - antes de serem enfim carregados na cadeia nascente do MHC-I (Falk & Röttschke, 2002).

O encaixe do epitopo na fenda do MHC é realizado por um Complexo de Carregamento do Peptídeo (PLC) que, além das duas cadeias do MHC-I recém

³ Peptídeos curtos (8-11 aa) oriundos de proteínas citosólicas que foram degradadas por proteases da célula.

⁴ Uma cascata de eventos envolvendo as enzimas E1, E2 e E3 adiciona uma cauda de ubiquitinas à seqüência alvo. A ubiquitina é uma proteína com 76 aminoácidos, altamente conservada nos eucariotos (Wang & Maldonado, 2006).

sintetizado, envolve uma série de chaperonas⁵ do ER e moléculas acessórias, destacando-se as proteínas calnexina, Bap31, ERp57, tapasina e a própria TAP. Sabe-se que Bap31 está envolvida na agregação de complexos pMHC em sítios de saída para a exportação do ER (Zhang & Williams, 2006). A cadeia nascente do MHC-I (α) é direcionada ao ER onde é “recebida” e estabilizada pelas chaperonas Grp78 e calnexina. A ligação da cadeia β_2 -microglobulina induz a substituição da calnexina por uma chaperona ortóloga, a calreticulina (CRT) e neste ERp57 e tapasina são também recrutadas, formando o PLC (Fig 1). A ERp57 - uma tiol oxidoreductase - é conhecida por catalizar a formação de pontes dissulfeto, estando, em conjunto com a CRT, envolvida no dobramento correto e na maturação de uma série de glicoproteínas do ER, bem como do próprio MHC-I. O prolongado tempo de interação da ERp57 com o MHC-I e com os demais integrantes do PLC indica uma provável função desta chaperona na formação da fenda e na ligação do peptídeo (Purcell & Elliott, 2008), conforme será posteriormente abordado neste trabalho. A tapasina é uma proteína responsável pela colocação entre o PLC e a TAP. Esta proteína liga-se covalentemente à ERp57 e possui diversas funções já conhecidas, tanto sobre o MHC quanto sobre a TAP. Acredita-se também que este heterodímero, ERp57-tapasina, seja responsável pelo processo de edição do peptídeo após a dissociação entre a TAP e o complexo pMHC recém construído (Wearsch & Cresswell, 2007).

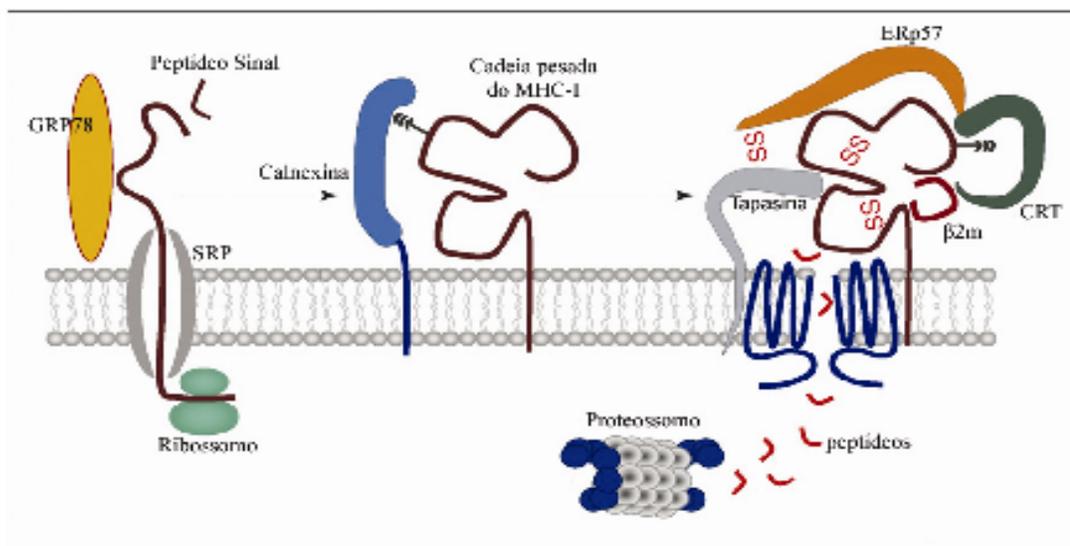


Figura 1. Os componentes do Complexo de Carregamento do Peptídeo (PLC). A cadeia nascente do MHC-I

⁵ Proteína que auxilia o enovelamento de outras proteínas.

(α) é direcionada para o ER por um peptídeo sinalizador (clivado posteriormente) sendo translocada ao lúmen desta organela pela partícula reconhecedora de sinal (SRP). Dentro do ER, um grupo de chaperonas e moléculas acessórias é recrutado para dar estabilidade ao MHC recém sintetizado e ainda realizar o carregamento, em sua fenda de ligação, dos peptídeos oriundos do processamento de proteínas citosólicas. Modificado de Purcell & Elliott, 2008. (continuação legenda fig.1.)

Após a completa montagem do pMHC, este é encaminhado para a superfície celular onde, conforme já apresentado, sofrerá reconhecimento pelos linfócitos T citotóxicos através da interação com o Receptor de Célula T (TCR), processo este estabilizado pela molécula acessória CD8. O receptor de células T é 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 realizam o contato direto com sítios do peptídeo e da molécula de MHC (Brehm et al., 2004). Os linfócitos T sofrem um rígido processo de seleção no timo onde, mediante competição por estímulos, linfócitos não responsivos são negligenciados e morrem, linfócitos altamente auto-reativos são negativamente selecionados, restando potencialmente uma população de linfócitos cujo TCR possui apenas baixa afinidade por peptídeos próprios (Sohn & cols, 2007).

Através da rota de apresentação de peptídeos endógenos, que envolve desde o proteossomo até a apresentação ao TCR, peptídeos próprios estão sendo constantemente apresentados aos linfócitos CD8⁺ por todas as células nucleadas do organismo. De acordo com o processo de seleção tímica, resumidamente apresentado neste texto, complexos pMHC apresentando epitopos próprios, na periferia, não serão reconhecidos pelo TCR. Esta constante vigilância é de suma importância na prevenção e na resposta contra patologias celulares, especialmente infecções virais e oncogênese (Hammer e cols, 2008).

Sendo parasitas intracelulares obrigatórios, os vírus infectam sua célula alvo e induzem a realização dos processos necessários a sua replicação. Independentemente da estratégia utilizada, que varia muito de acordo com o vírus e seu tipo de genoma, proteínas virais serão produzidas em grandes quantidades e liberadas no citoplasma da célula infectada. Inevitavelmente, algumas moléculas destas proteínas sofrerão o processo de ubiquitinação, sendo então encaminhadas para a via do MHC-I. Os epitopos

virais no contexto do MHC, normalmente distintos dos pMHC próprios, serão reconhecidos por TCRs levando à ativação do linfócito T CD8+, liberação de seu conteúdo citotóxico e conseqüente lise da célula infectada.

A liberação da citotoxicidade celular não é, no entanto, a única conseqüência possível após o contato entre pMHC e TCR. A intensidade desta interação pode variar dentro de um amplo espectro de possibilidades, desencadeando diferentes comportamentos no linfócito, entre os quais destacam-se a ativação, a expansão clonal e o processo de anergia (Nel, 2002). Não obstante, uma conseqüência deste contato pode ser a geração de linfócitos T de memória, mecanismo central na resposta imune adaptativa e com especial importância para o desenvolvimento de vacinas profiláticas (Seder e cols, 2008). Após sofrer infecção com um determinado vírus, ou ser vacinado, inicia-se no organismo uma resposta imunológica que deverá eliminar o agente e gerar células de memória. O advento de um segundo contato, com o mesmo agente, reativa estas células de memória que, sendo específicas para o vírus infectante, guiarão uma resposta rápida e eficaz, impedindo que o vírus desencadeie seu efeito patológico sobre o organismo infectado. No entanto, o contato com um segundo agente – diferente daquele presente no primeiro contato – pode também recrutar uma parte destas células de memórias, as quais serão ativadas, podendo proteger o indivíduo contra este agente desconhecido. Este fenômeno pelo qual um contato prévio com um agente induz um estado de imunidade que pode conferir proteção parcial contra outros agentes (relacionados ou não), ou afetar de maneira significativa a resposta ao segundo agente (causando alterações no curso da patologia) foi denominado imunidade heteróloga. Este fenômeno é dependente da história imunológica⁶ do hospedeiro, dos linfócitos de memória que ele possui e está envolvido em uma série de eventos como resposta efetiva a um patógeno (*clearance*), infecção viral crônica e indução de autoimunidade (Welsh & Selin, 2002).

Conforme apresentado anteriormente, o reconhecimento de um complexo pMHC por um TCR específico permite a geração da resposta e a indução de memória. As células de memória não alteram a especificidade de seus TCRs, logo, a reativação deste grupo de células é dependente do reconhecimento dos novos pMHCs, pelos mesmos

⁶ o conjunto de infecções ao longo da vida do hospedeiro altera sua resposta a infecções futuras.

TCRs das células de memória. Considerando que infecções subseqüentes com vírus heterólogos gerarão complexos pMHC apresentando diferentes peptídeos virais, a imunidade heteróloga só é possível mediante um reconhecimento degenerado por parte do linfócito T. O termo reatividade cruzada foi utilizado para definir esta capacidade de um linfócito T em reconhecer peptídeos não relacionados, provenientes de um mesmo organismo ou mesmo de organismos heterólogos (Welsh & Fujinami, 2007). Contrariando a visão clássica da imunologia, acredita-se que um mesmo TCR possa reconhecer mais de um milhão de diferentes complexos pMHC (Mason, 1998) e há cada vez mais dados corroborando a ocorrência de reconhecimento degenerado pelo TCR (Godfrey e cols, 2008). Um dos principais motivos para a ocorrência deste fenômeno parece ser a limitação de espaço imposta ao sistema imune. Cada vez que uma população de células expande, uma outra população de células deverá ser eliminada, mantendo o número total de linfócitos relativamente estável. Não será possível, portanto, guardar células de memória contra todos os antígenos com os quais se teve contato sendo, neste contexto, de especial importância a existência de células de memória “*cross reativas*”, as quais poderão maximizar a capacidade de resposta do *pool* de memória enquanto minimizam o espaço físico necessário a seu armazenamento (Welsh & Selin, 2002; Selin e cols, 2004). A reatividade cruzada está possivelmente envolvida em muitos outros mecanismos do sistema imune, como na seleção positiva de timócitos imaturos⁷, na indução de autoimunidade por mimetismo molecular, na diferenciação de linfócitos e na manutenção da memória de longo prazo (Reali e cols, 1999; Regner, 2001).

A reatividade cruzada é citada em muitos trabalhos, sendo discutida há vários anos. Os trabalhos mais antigos se referiam a este fenômeno como um achado, não fazendo maiores referências aos mecanismos envolvidos. Trabalhos posteriores conseguiram relacionar a ocorrência deste fenômeno a semelhanças na seqüência de aminoácidos dos epítopos envolvidos, o que determinaria uma semelhança entre os pMHC formados e explicaria o reconhecimento cruzado. Nessa linha foram descritos vários casos de reatividade cruzada, entre eles a presença de células T CD8+ específicas para um epítipo 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

⁷ Linfócitos T imaturos em processo de seleção no Timo.

hepatite B (HBV), mas que já haviam sido infectadas pelo vírus Influenza (IV), agente causador da gripe verdadeira. Neste trabalho, publicado em 2001, Wedemeyer e colaboradores identificaram os epitopos NA₂₃₁₋₂₃₉ (IV) e NS₃₁₀₇₃₋₁₀₈₁ (HCV) como alvos da reatividade cruzada, sendo que estes compartilhavam 7 de seus 9 aminoácidos. Embora a semelhança entre as seqüências dos epitopos visivelmente colabore com o reconhecimento cruzado, muitos estudos recentes passaram a relatar a ocorrência de reatividade cruzada entre epitopos que divergiam em mais de 50% de sua seqüência simples de aminoácidos. Em um estudo publicado em 2005, Sandalova e colaboradores demonstraram que dois epitopos (um viral e um próprio) compartilhando apenas 44% de sua seqüência, quando apresentados pelo mesmo alelo de MHC, apresentavam estrutura praticamente idêntica. A superfície dos complexos pMHC formados compartilhava ainda topologia e distribuição de cargas, sendo esta semelhança estrutural sugerida como responsável pelo reconhecimento cruzado e, neste caso, desenvolvimento de uma patologia autoimune.

Apesar de todos os dados gerados, e da visível importância da reatividade cruzada em inúmeros processos imunológicos, incluindo potencial uso no desenvolvimento de vacinas (Vieira & Chies, 2005), mais estudos são necessários para esclarecer as dúvidas que ainda restam sobre as bases físicas e químicas envolvidas no reconhecimento do pMHC pelo TCR. O desenvolvimento de novas técnicas capazes de facilitar, agilizar e reduzir custos nos processos de indentificação e testagem de epitopos “*cross reativos*” também seria de extrema importância, tanto para a pesquisa básica quanto para aplicações voltadas à clínica.

Algumas técnicas computacionais já estabelecidas poderiam ser empregadas neste contexto, desde que observados os parâmetros e restrições relativos ao sistema imune. O *docking* molecular é uma importante ferramenta computacional que permite “encaixar” um ligante a seu receptor ou explorar os possíveis sítios de interação entre um composto e uma proteína de interesse (Morris e cols, 1998). Esta ferramenta vem sendo utilizada no *design* de fármacos e estudos de interação sítio-específica, colaborando para a melhor compreensão do funcionamento de certos compostos e para o desenvolvimento de novas drogas (Cavasotto & Orry, 2007). Para refinar os modelos gerados por *docking* utiliza-se o processo de Minimização de Energia (EM), técnica que

simula *in silico* o comportamento do complexo de interesse em solução aquosa⁸, o que induz a proteína e o ligante a adotarem uma conformação mais estável, provavelmente mais próxima da que ocorreria *in vivo* (Kmiecik e cols, 2007).

Neste trabalho utilizamos dois epitopos virais como modelos para um estudo de reatividade cruzada *in silico*. Um dos epitopos já possuía estrutura cristalografada – complexado ao alelo de MHC murino H-2D^b - depositada no Protein Data Bank (Meijers e cols, 2005). O outro epitopo selecionado para o estudo, no entanto, não possuía estrutura determinada. Através do uso de *docking* molecular e EM, estabelecemos um protocolo que permite construir complexos pMHC com qualquer epitopo de interesse, partindo de sua seqüência simples de aminoácidos. De posse de ambos os complexos pMHC, partimos para um estudo da superfície destes complexos, analisando sobretudo a topologia e a distribuição de cargas. Nossos dados sugerem ser esta uma boa estratégia para estudos de reatividade cruzada, permitindo a rápida e barata (todos os programas são disponíveis sem custo) identificação e análise dos epitopos com potencial para induzir autoimunidade ou imunidade heteróloga.

⁸ O processo de Minimização de Energia também pode ser realizado no vácuo.

Trabalho experimental (artigo)

Utilização de Ferramentas de Bioinformática para a
Análise do Potencial de Reatividade Cruzada entre
Epitopos Virais.

Dinler A Antunes, Maurício M Rigo, Gustavo F Vieira, José AB Chies[§]

Núcleo de Bioinformática do Laboratório de Imunogenética, Departamento de
Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

[§]Autor para contato.

Endereços de E-mail:

DAA: dinler@gmail.com

MMR: mauriciomr1985@gmail.com

GFV: gusfioravanti@yahoo.com.br

JABC: jabchies@terra.com.br

Histórico

Os linfócitos T CD8⁺ reconhecem peptídeos processados, que são apresentados na superfície celular no contexto do Complexo Principal de Histocompatibilidade (MHC) de classe I [1]. Os peptídeos apresentados são seqüências curtas contendo de oito a doze aminoácidos, chamadas de epitopos. Estas seqüências são derivadas de proteínas endógenas à célula, ou seja, podem ser tanto provenientes de proteínas citosólicas próprias quanto de proteínas pertencentes a parasitas intracelulares, o que torna esta rota um importante mecanismo de defesa antiviral [2]. A interação entre o Receptor de Células T (TCR) e o complexo MHC:peptídeo (pMHC) pode ser bastante degenerada e alguns estudos demonstram que um mesmo TCR pode reconhecer até 10⁶ diferentes combinações de pMHC [3]. Esta capacidade de um linfócito T em reconhecer peptídeos não relacionados, provenientes de um mesmo organismo ou mesmo de organismos heterólogos, é denominada reatividade cruzada [4].

Imunidade heteróloga é um termo utilizado para descrever uma imunidade parcial que ocorre em resposta a um patógeno quando o hospedeiro foi previamente infectado ou imunizado com outro patógeno não relacionado. Este fenômeno, diretamente relacionado ao mecanismo de reatividade cruzada, é dependente da história imunológica do hospedeiro, dos linfócitos de memória que ele possui e está envolvido em uma série de eventos como resposta efetiva a um patógeno (*clearance*), infecção viral crônica e indução de autoimunidade [5].

Existem muitos relatos de imunidade heteróloga entre infecções por vírus não relacionados e de autoimunidade induzida por mimetismo (ou mímica) molecular [6, 7], mas os mecanismos que possibilitam esta imunização cruzada ainda não foram completamente esclarecidos [4]. A maioria dos trabalhos que se propõe a estudar reatividade cruzada utiliza como modelo epitopos que compartilham entre si a maior parte da seqüência de aminoácidos [8], no entanto, já foram demonstrados *in vitro* alguns casos de reatividade cruzada entre epitopos que compartilhavam menos de 40% da seqüência [9].

O antígeno de superfície da hepatite B apresenta alto poder de imunização, sendo utilizado para o desenvolvimento de vacinas e para monitoração da imunidade

induzida pela vacina anti-HBV. Um dos epitopos desta proteína é o peptídeo imunodominante HBsAg₂₈₋₃₉ (IPQSLDSWWTSL) [10]. O epitopo PA₂₂₄₋₂₃₃ (SSLENFRAYV) da proteína polimerase 2 do vírus Influenza A [11] compartilha apenas dois aminoácidos com o HBsAg₂₈₋₃₉, contudo, sete de seus dez aminoácidos pertencem ao mesmo grupo físico-químico dos aminoácidos correspondentes no epitopo de HBV. Considerando que a topologia e a distribuição de cargas do complexo pMHC são críticos para a interação com o TCR e que a estimulação do linfócito depende da interação com apenas alguns aminoácidos do epitopo [5] e não necessariamente com toda a seqüência, acreditamos que os epitopos acima descritos sejam bons candidatos a um estudo de reatividade cruzada.

O *docking* molecular vem sendo utilizado com sucesso para realizar a complexação entre um ligante e seu receptor ou explorar os possíveis sítios de interação (caso existentes) entre um composto e uma proteína de interesse [12]. Vários trabalhos descrevem o uso desta ferramenta no *design* de fármacos e estudos de interação sítio-específica, colaborando para a melhor compreensão do funcionamento de certos compostos e para o desenvolvimento de novas drogas [13]. Para refinar os modelos gerados por *docking* utiliza-se a Minimização de Energia (EM). Esta técnica, que pode ser realizada em solução aquosa ou no vácuo, induz a proteína e o ligante a adotarem uma conformação mais estável, provavelmente mais próxima da que ocorreria *in vivo* [14, 15].

A estrutura do epitopo PA₂₂₄₋₂₃₃ cristalografado no contexto do alelo H-2D^b, está disponível no *Protein Data Bank* (PDB) [16] sob o código de acesso 1WBY. Neste trabalho, através do uso combinado de *docking* e EM, construímos a estrutura do epitopo HBsAg₃₀₋₃₉ no contexto do alelo H-2D^b (D^bHBsAg₃₀₋₃₉), a fim de utilizá-lo – em conjunto com a estrutura 1WBY - no estudo da superfície destes complexos, comparando sobretudo a topologia e a distribuição de cargas. Acreditamos que a estratégia utilizada para a construção do complexo D^bHBsAg₃₀₋₃₉, seguida do uso de programas de visualização que permitem analisar a topologia do complexo gerado, a distribuição de cargas e a conformação espacial dos aminoácidos responsáveis pela interação, permite estudar novos complexos pMHC sob o “ponto de vista” do TCR e inferir um possível mecanismo de reatividade cruzada entre dois epitopos distintos [7]. Já foram determinados quais aminoácidos do epitopo são os mais importantes para a ligação ao

MHC [17], assim como foi descrito que os aminoácidos centrais do epitopo são decisivos para o contato com o TCR e a indução do linfócito T [5], mas as bases moleculares dessa interação e da própria hierarquia de imunodominância dos epitopos permanecem pouco conhecidas [4]. Através de estudos que verifiquem a estrutura dos epitopos no contexto do MHC poderemos potencialmente chegar a um conhecimento mais claro sobre as características físico-químicas e estruturais que determinam fenômenos como imunodominância e mimica molecular, utilizando então tal conhecimento no desenvolvimento de vacinas mais abrangentes [18] e na prevenção de autoimunidade.

Resultados

Análise de geração do epitopo HBsAg₂₈₋₃₉

A fim de confirmar a capacidade de geração do epitopo HBsAg₂₈₋₃₉ pela via do MHC-I e avaliar seu potencial em ser apresentado pelo alelo H-2D^b, foi realizado um estudo da geração deste epitopo. A compilação dos dados obtidos indicou que o processo de clivagem do HBsAg gera o peptídeo TIPQLDSWWTSL, o qual será transportado pelo Transportador Associado ao Processamento de Antígenos (TAP) para o retículo endoplasmático onde sofrerá um desbastamento N-terminal, tornando-se o decâmero (10-mer) QSLDSWWTSL, que será apresentado no contexto do alelo de MHC-I murino H-2D^b.

Identificação de um padrão estrutural alelo específico

Realizamos uma busca por todas as estruturas de MHC murinos depositadas no PDB e analisamos a conformação dos diferentes epitopos aos quais estes MHCs estão ligados. Encontramos um padrão conformacional alelo-específico, demonstrando que a estrutura tridimensional do epitopo apresentado não é intrínseca à sua seqüência de aminoácidos, mas ativamente determinada pela fenda do MHC-I, a qual varia de acordo com o alelo (Figuras 1 e 2). Os dados estruturais dos ligantes de H-2D^b e H-2K^b (alelos murinos com maior número de estruturas depositadas no PDB) foram cruzados com informações sobre suas posições de ancoragem e estrutura da fenda, demonstrando claramente quais regiões eram mantidas rígidas e em quais havia possibilidade de

variação das cadeias laterais, para dar especificidade ao contato com o TCR (Figura 3).

Estabelecimento de uma estratégia para a construção de complexos pMHC

A existência de um padrão estrutural alelo-específico nos permitiu desenvolver um procedimento através do qual um epitopo já cristalografado - no contexto do MHC de interesse - é utilizado como molde para a construção da estrutura de um outro epitopo que se deseja estudar (vide métodos). Esta estratégia conseguiu reproduzir a estrutura cristalina 1WBY, com um valor de *Root Mean Square Deviation* (RMSD) igual a 1,3 Å (calculado sobre todos os átomos do epitopo), sendo que, de acordo com a literatura, médias de desvio inferiores a 2,2 Å são consideradas reproduções válidas [19]. A energia de ligação (*BE*) apresentada pelo *docking* flexível deste epitopo “modelado” foi de -12,57 kcal/mol, enquanto a *BE* do *docking* rígido do epitopo original apresentou valor de -12,82 kcal/mol (valor utilizado como referência para as demais análises). Esta abordagem também mostrou-se capaz de reproduzir a estrutura de um pMHC envolvendo o alelo H-2K^b, utilizando um arquivo PDB como molde para o epitopo e outro arquivo PDB como doador da estrutura do MHC, embora este procedimento não tenha sido tão eficaz, em função de diferenças entre as estruturas cristalografadas.

A reprodutibilidade da estratégia foi testada pela reprodução do epitopo SSIEFARL (Glicoproteína B de HSV) no contexto do alelo H-2K^b (1RJY), utilizando como molde um epitopo do precursor do receptor de insulina de camundongo - seqüência GNYSFYAL (1LK2) - e como doador do MHC uma estrutura de H-2K^b disponível sob o código 3C8K (cristalografada apresentando um epitopo de ovoalbumina). O melhor resultado do Dock 2 apresentou *BE* = -11,47 kcal/mol. Como o complexo gerado e a estrutura original não possuem o mesmo número de átomos, o RMSD foi calculado apenas sobre a cadeia principal do pMHC, apresentando valor de 1,13 Å. No entanto, a análise da superfície apresentou algumas diferenças de carga e topologia, inclusive em regiões do MHC.

Estudo *in silico* do potencial de reatividade cruzada entre os epitopos virais PA₂₂₄₋₂₃₃ (Influenza) e HBsAg₂₈₋₃₉ (HBV)

O resultado do *docking* indicou alta energia livre de ligação entre o epitopo HBsAg₃₀₋₃₉ e o alelo H-2D^b (*BE* = -13,13 kcal/mol) e a topologia do complexo

apresentava forte semelhança com a estrutura pMHC do epitopo PA₂₂₄₋₂₃₃ (Figuras 4A e 4B). O epitopo de Influenza apresentava, no entanto, uma crista com carga positiva que não possuía correspondente na estrutura do epitopo de HBV. Esta crista, conferida por uma arginina na posição 7 do epitopo (p7) - justamente uma das posições onde havia diferença físico-química entre as seqüências - encontra-se em uma região crítica de contato com o TCR, o que talvez dificulte uma possível reação cruzada. No intuito de melhor explorarmos o modelo utilizado, desenhamos um ligante alternativo (APL) de Influenza, com uma mutação R7W, tornando-o mais semelhante ao epitopo de HBV. A análise de geração deste APL não apresentou alterações com relação ao resultado do epitopo original e ambos apresentaram o mesmo escore de ligação no SYFPEITHI (Banco de ligantes de MHC) [20]. As estruturas, do APL e do complexo com o MHC murino, foram construídas seguindo os mesmos passos já descritos. O APL apresentou no *docking* um valor de ligação ao alelo H-2D^b maior do que o epitopo original ($BE = -17,5$ kcal/mol) e, conforme o esperado, apresentou uma maior semelhança com a estrutura construída do epitopo de HBV (Figura 4C).

Discussão

O complexo de carregamento do peptídeo (PLC) é formado por uma série de chaperonas do ER e moléculas acessórias. O heterodímero ERp57-tapasina parece exercer papel central neste complexo, mantendo-se durante muito tempo associado ao pMHC, mesmo após a dissociação entre este e a TAP. A ERp57 é uma tiol oxidoreductase conhecida por catalizar a formação de pontes dissulfeto, estando envolvida no dobramento correto e na maturação da cadeia pesada do MHC-I, possuindo ainda provável função na formação da fenda de ligação ao peptídeo [21]. A tapasina é responsável pela colocalização entre o PLC e a TAP, além de possuir outras muitas funções como aumento da expressão de TAP, aumento da ligação de peptídeos a TAP e manutenção do estado receptivo do MHC aos peptídeos [22]. Esta proteína liga-se covalentemente à ERp57 para formar o heterodímero ERp57-tapasina, que parece ser responsável pelo processo de edição do peptídeo após a dissociação entre a TAP e o complexo pMHC recém construído, favorecendo o carregamento dos melhores ligantes, em detrimento de outros [23].

Neste trabalho, verificamos que a conformação da cadeia principal era compartilhada entre diferentes epitopos apresentados por um mesmo alelo de MHC. Pode-se argumentar que alguns dos epitopos incluídos na análise são variantes que diferem em apenas um aminoácido, o que poderia ser um viés na identificação de um padrão geral do alelo. Contudo, epitopos com seqüência muito diversa mantiveram o mesmo padrão. Além disso, as variantes foram incluídas com base na idéia de que a troca de um aminoácido pode induzir uma alteração nas torções do esqueleto de carbono e inclusive cancelar o reconhecimento por um TCR específico para a seqüência original [24]. Analisando as estruturas, verificamos que os epitopos que fugiam a regra não o faziam por serem oriundos de proteínas distintas, e sim, pela necessidade de acomodarem mais aminoácidos que os outros (seqüências maiores). Percebemos ainda que este padrão estrutural é específico do alelo de MHC, de modo que epitopos apresentados por diferentes alelos, apresentam conformações diferenciadas (Figura 1A). Mais do que isso, estudos anteriores já haviam sinalizado para o fato de que um mesmo epitopo apresenta conformações distintas quando apresentado no contexto de diferentes alelos de MHC (Figura 1C) [25]. Em conjunto, estes dados indicam que a conformação adotada pelo epitopo na fenda do MHC não é determinada unicamente por sua origem ou seqüência, mas também, pelo formato da fenda do alelo no qual será apresentado. O epitopo deve ser ativamente induzido a adotar a conformação adequada durante o processo de carregamento. O PLC possui a maquinaria necessária para esta função e o heterodímero ERp57-tapasina, pelas funções já descritas, se torna um forte candidato a exercer esta tarefa. Evidentemente nem todos os peptídeos conseguirão adotar a conformação adequada e ainda manter uma ligação estável com o MHC, sendo neste caso, substituídos por melhores ligantes.

O fato de o padrão encontrado não se aplicar às cadeias laterais, não implica uma distribuição aleatória destas na fenda. Na verdade, um estudo mais detalhado dos 18 diferentes ligantes de H-2D^b encontrados (Tabela 1), demonstrou claramente a existência de regiões de maior e menor controle sobre as cadeias laterais dos epitopos. O epitopo na fenda do H-2D^b encontra-se envolto de tal modo, que uma parte de sua extremidade amino-terminal (N-term) está oculta sob cadeias laterais do MHC (p2-3 do epitopo) enquanto outra parte (p6-7) encontra-se externalizada à fenda do MHC (Figura 3A). Não

por acaso, existe maior identidade entre as cadeias laterais em p2-3 e maior diversidade de cadeias laterais na região p6-7 (Figura 3B). Esta região, onde as cadeias laterais estão mais livres e existe grande diversidade, será responsável pela especificidade do contato com o TCR. Observamos ainda grande identidade entre cadeias laterais em p5 e p9, justamente as posições de ancoragem à este alelo. Um alinhamento das seqüências dos ligantes, de acordo com suas características físico-químicas, indicou a presença preferencial de uma asparagina na posição 5, substituída em 3 casos (1WBZ, 1BZ9 e 1JUF) por um aminoácido não polar. A última posição também mostrou conservação, apenas uma seqüência apresentou uma cisteína, enquanto todas as demais possuíam aminoácidos apolares. A maior parte dos ligantes de H-2D^b possuía 9 aminoácidos, contudo, também encontramos 10-mers e 11-mers. Apesar das diferenças no comprimento, a posição dos âncoras é conservada e o ajuste do comprimento da cadeia ocorre justamente na região mais livre (p6-7), posições em que verificamos maior divergência entre as estruturas, mesmo considerando apenas o esqueleto de carbono (Figuras 2A e 2C). Um dos ligantes, o peptídeo sintético FAPGVFPYM (1BZ9) apresentou acentuado desvio na cadeia principal, sobretudo em p5, elevando-se para fora da fenda (Arquivo Adicional 1). Este epitopo não possui o resíduo de ancoragem convencional na posição 5, sendo que os autores que o descreveram discutiram brevemente sua conformação incomum, afirmando que eventualmente um epitopo sem os âncoras convencionais poderia ligar-se ao MHC, desde que outros aminoácidos em sua seqüência compensem esta deficiência [26].

A análise dos ligantes de H-2K^b também demonstrou conservação das cadeias laterais nos aminoácidos de ancoragem. Das 18 seqüências analisadas (AA2), 14 possuíam aminoácidos aromáticos (F ou Y) na quinta posição, enquanto as quatro exceções apresentavam aminoácidos polares (N, S ou T). Todas as seqüências apresentaram aminoácidos não polares na última posição (C-term). No entanto, quando comparado ao H-2D^b, o alelo H-2K^b parece envolver menos os seus ligantes, permitindo uma maior divergência já em p4 (AA3 - Figura A). Além disso, o ajuste de cadeia neste alelo também ocorre em uma posição diferente, aparentemente entre p3 e p4, sendo muito mais conservados em p6-7 (AA3 - Figura B). Embora tenhamos estudado em maiores detalhes apenas dois alelos de MHC murino, acreditamos que a existência deste padrão

possa ser extrapolada para outros alelos, inclusive humanos. Alguns trabalhos, sobretudo aqueles em que se fez cristalografia de complexos pMHC, já haviam descrito semelhanças estruturais entre as cadeias principais de alguns epitopos, mas não encontramos trabalhos que relatassem padrões estruturais tão abrangentes como os aqui descritos.

Sendo a estrutura da cadeia principal determinada pela fenda, podemos adiantar que, mesmo epitopos cuja estrutura não foi determinada, adotarão – salvo raras exceções - esta mesma conformação no contexto do MHC. Esta foi a premissa para o desenvolvimento de uma nova abordagem, que nos permite construir complexos pMHC cuja a estrutura não foi determinada. Cabe salientar que nem todos os peptídeos conseguem ser apresentados pelo MHC-I (ou por todos os alelos). Acreditamos que o *docking* seja capaz de identificar “maus ligantes” através da variação na energia de ligação, contudo, esta estratégia torna-se ainda mais confiável se precedida de um estudo de geração do epitopo ou, no mínimo, da verificação da afinidade da seqüência de interesse pelo alelo de MHC-I, em um banco de ligantes de MHC.

Baseados em um epitopo descrito no contexto do alelo H-2L^d, identificamos qual a seqüência com maior probabilidade de ser apresentada em H-2D^b e verificamos o escore de afinidade desta seqüência no SYFPEITHI [20]. Como o escore indicou capacidade de ligação, construímos a estrutura do complexo D^bHBsAg₃₀₋₃₉, a qual foi utilizada em um estudo de reatividade cruzada com o complexo D^bPA₂₂₄₋₂₃₃. Sobre os parâmetros utilizados para a construção do pMHC, cabe salientar que o número máximo de avaliações (*Evals*) utilizadas no *docking* foi escolhido por uma questão de restrição computacional, estando muito abaixo do que seria um valor ótimo para este estudo, tendo-se em vista o número de ligações flexíveis utilizado. Contudo, existem trabalhos na literatura utilizando com sucesso valores semelhantes de avaliações e, em nosso estudo, o valor escolhido demonstrou-se capaz de reproduzir uma estrutura cristalografada [27]. O uso de valores de avaliação maiores certamente traria resultados ainda melhores, sendo indicado em estudos que disponham de maiores recursos computacionais.

Apesar de não termos confirmado nossa expectativa de que ambos os complexos – com o epitopo de HBV e de IV - fossem idênticos em topologia e carga, predissemos um APL, baseado no epitopo de Influenza, que possivelmente induz imunidade heteróloga ao

HBV, ou até mesmo a ambos. A mutação R7W solucionou a diferença de cargas na região de contato com o TCR, mas a posição das cadeias laterais do triptofano (p7) não ficaram exatamente na mesma orientação do correspondente no epitopo de HBV. Esta diferença de topologia é conferida pelo anel indólico do triptofano, que possui uma certa liberdade rotacional conferida pela ligação simples que o une ao grupamento metileno, conforme demonstrado pelo próprio *docking*, que apresentou outras possíveis conformações da cadeia lateral do triptofano sem comprometer a possibilidade de apresentação pelo H-2D^b (AA4). Além disso, diferenças de cargas no pMHC são apresentadas pela literatura como sendo mais críticas para o reconhecimento pelo TCR do que pequenas variações na topologia [28, 29]. Assim sendo, acreditamos que esta pequena variação estrutural (em p7) não impeça, *in vivo*, o reconhecimento do APL e do epitopo de HBV pelo mesmo TCR (no contexto do H-2D^b) e testes *in vitro* seriam necessários para a confirmação da possível reatividade cruzada entre estes complexos. Caso confirmada, este APL poderia ser facilmente incorporado na formulação de vacinas da gripe, induzindo proteção também contra o HBV. Sendo uma estratégia muito mais barata do que a atual vacina anti-HBV, esta nova vacina teria potencial aplicação para pacientes HIV+ e em populações carentes, como em várias regiões da América Latina, onde a hepatite B é considerada uma doença endêmica [30].

Cabe ainda salientar, no tocante ao procedimento para a construção de pMHCs e, sobretudo, ao experimento em que reproduzimos a estrutura 1RJY, o fato de que a possibilidade de se utilizar um terceiro arquivo PDB como doador do MHC (3C8K) e ainda assim construir um complexo semelhante ao original, demonstra que o resultado do processo não é dependente do uso de um arquivo PDB específico, e que o resultado do *docking* não está sendo direcionado pelo *Fit*. No entanto, existe uma margem de erro no processo de cristalografia, acarretando diferenças estruturais entre diferentes arquivos PDB de um mesmo alelo de MHC, sendo assim esperada alguma variação de topologia/potencial entre o complexo gerado e o original. Para o uso desta estratégia na construção de novos complexos pMHC, recomendamos que se utilize como “doador do MHC” um arquivo PDB com a melhor resolução possível. Após identificado este arquivo, ele deve ser utilizado como a estrutura 3D daquele alelo na construção de todos os demais pMHCs que se deseja estudar. Deste modo se reduz o erro oriundo do processo

de obtenção da estrutura (cristalografia ou ressonância magnética nuclear) e pode se concentrar na variabilidade oriunda do processo de construção deste novo complexo (fruto da EM e do *docking*).

Outra observação importante neste trabalho é a marcante contribuição das características físico-químicas dos aminoácidos, tanto no processo de ligação do epitopo ao MHC, quanto na atribuição de especificidade ao contato com o TCR. Conforme apresentado, os aminoácidos âncoras para os alelos estudados pertenciam ao mesmo, ou aos mesmos, grupos físico químicos. No estudo de reatividade cruzada realizado, os epitopos compartilhavam apenas 2 aminoácidos, mas 70% da seqüência pertencia ao mesmo grupo. Quando realizamos o estudo da superfície dos pMHC destes epitopos, encontramos diferença de topologia e de cargas justamente em uma das posições que divergia quanto ao grupo, enquanto nas demais posições, a despeito da divergência na seqüência, havia forte correlação estrutural/potencial.

Conclusões

Concluimos que existe um padrão conformacional alelo-específico compartilhado pelos epitopos apresentados no contexto de alelos de MHC-I, provavelmente induzido pela maquinaria de carregamento do peptídeo. Este padrão atua somente sobre a cadeia principal, embora as cadeias laterais também sofram ajustes impostos pela fenda do MHC-I. Epitopos que não conseguirem adotar a conformação adequada, ou cujas cadeias laterais posicionem-se em regiões não permitidas, serão provavelmente incapazes de ser apresentados por este alelo.

O conhecimento de tal padrão nos permitiu desenvolver uma estratégia computacional para a construção da estrutura de complexos MHC-I:peptídeo, partindo-se da seqüência simples de aminoácidos do epitopo de interesse. Sendo o complexo pMHC um ponto chave na resposta imune adaptativa, esta abordagem rápida e barata torna-se útil para emprego nos mais diversos campos de estudo dos mecanismos imunológicos. Aliada a ferramentas que permitem analisar topologia e carga na superfície dos complexos, esta estratégia pode ser ainda utilizada na seleção de candidatos com potencial para induzir reatividade cruzada, tanto em estudos de autoimunidade quanto no desenvolvimento de vacinas mais abrangentes.

Métodos

Estudos de geração de epitopos

Os dados analisados foram compilados a partir do SYFPEITHI (banco de ligantes de MHC) [20], do simulador de proteossomo NetChop 3.0 (C-Term) [31] e do preditor de epitopos (MHC-I) do *Immune Epitope Database and Analysis Resource* (IEDB) [32]. No SYFPEITHI, utilizamos a sessão *Epitope Prediction* para determinar o escore de ligação dos epitopos. A seqüência FASTA da proteína de origem do epitopo foi utilizada como entrada no NetChop e no IEDB. Em ambos escolheu-se a opção referente ao imunoproteossomo. No IEDB foram feitas simulações com *Maximum Precursor Extension* (MPE) da TAP nos valores 1, 2, 3 e 4. Ambos os métodos de predição (SMM e ARB) foram utilizados, com os quatro valores de MPE. Os resultados foram cruzados entre si e com os dados de clivagem do NetChop 3.0.

Construção de complexos pMHC

A construção de complexos pMHC é realizada de acordo com as seguintes etapas (Figura 5): Primeiro identifica-se uma estrutura depositada no PDB (complexo “A”, por exemplo), que possua o alelo de MHC de interesse apresentando um epitopo com o mesmo número de aminoácidos do epitopo que se deseja modelar (epitopo c, por exemplo). Em seguida, separa-se as estruturas do MHC e do epitopo (do complexo A), salvando o epitopo em um arquivo independente que será aberto com o SPDBV 3.7.[33]. Utilizando o painel de modelagem do programa (aba “SwissModel”) realiza-se o carregamento da seqüência FASTA do epitopo que se deseja modelar (epitopo c) e, com o comando “*Magic Fit*”, realiza-se a sobreposição desta seqüência no epitopo molde (retirado do complexo “A”). A estrutura “fitada” do epitopo c pode conter átomos em posições energeticamente desfavoráveis e aminoácidos com torções impróprias. Esta situação é corrigida realizando-se uma curta EM (nsteps = 100). A estrutura resultante é então submetida ao primeiro *docking* (*Dock1*) com uma molécula de MHC, que pode ser a mesma presente no complexo “A” ou pode ser outra estrutura de MHC do mesmo alelo. Esta estrutura doadora do MHC (complexo B, na Figura 5) pode não ter sido obtida apresentando um epitopo com o mesmo número de aminoácidos do “epitopo c”, no

entanto, é importante que se escolha para esta função uma estrutura com a melhor resolução possível. O MHC é mantido rígido durante o *docking*, assim como a maior parte das ligações na cadeia principal do epitopo, considerando que esta já se encontra na conformação desejada. As cadeias laterais do epitopo permanecem completamente flexíveis, o que permite grandes variações conformacionais. O pMHC formado com a conformação do epitopo que apresentou melhor resultado do *Dock1*, de acordo com a energia de ligação (BE), é submetido a uma EM um pouco mais longa ($nsteps = 1000$), a fim de ajustar as cadeias laterais do MHC ao novo ligante, reduzindo interações desfavoráveis. Por último, os componentes deste complexo minimizado são separados (epitopo e MHC) e utilizados como entrada para o segundo *docking* (*Dock2*), que gerará o pMHC de interesse.

Construção do complexo D^bHBsAg₃₀₋₃₉

A estrutura do 10-mer QSLDSWWTSL foi construída a partir de sua seqüência simples de aminoácidos, utilizando o programa SPDBV 3.7 e tendo como molde o epitopo restrito a H-2D^b, HA₄₆₈₋₄₇₇ (SQLKNNAKEI), disponível no PDB sob o código 1WBX. Apesar do epitopo molde também ser de Influenza, pertence a uma proteína de envelope Hemaglutinina, não compartilhando qualquer relação estrutural ou funcional com o epitopo PA₂₂₄₋₂₃₃ e tendo sido escolhido como molde apenas por ser um 10-mer apresentado por H-2D^b (enquanto a maioria dos ligantes deste alelo são 9-mer). Utilizamos o pacote GROMACS (versão 3.3.3) [34] para realizar a EM da estrutura do epitopo, utilizando o campo de força GROMOS 96 43a1, $nsteps = 100$ e $emstep = 0,01$. O programa AutoDock 4 [12] para realizar o *docking* molecular do epitopo modelado à estrutura do alelo H-2D^b. Durante o *docking*, 21 ligações foram mantidas flexíveis (*fewest atoms*) mantendo rígida a maior parte da cadeia principal.

A caixa (*Grid Box*) foi centrada na fenda do H-2D^b e ajustada para 42 X 42 X 68 pontos nos eixos x, y e z, respectivamente, com espaçamento (*grid-point spacing*) de 0,375 Å. Utilizamos o Algoritmo Genético de Lamarck, com *Maximum Number of Evals* = 250000, *Maximum Number of Generations* = 270000, população de 150 indivíduos e demais valores de acordo com o padrão do AutoDock4.

Realizamos ainda um passo adicional de EM ($nsteps = 1000$) do pMHC para

ajuste das cadeias laterais do MHC ao novo epítopo e então repetimos o *docking*, mantendo o MHC rígido e as cadeias laterais do epítopo flexíveis. A fim de comparar o complexo pMHC cristalografado do epítopo PA₂₂₄₋₂₃₃ (1WBY) com o pMHC do epítopo HBsAg₃₀₋₃₉, construído para esta análise, realizamos uma EM na estrutura 1WBY (mesmos parâmetros da minimização do D^bHBsAg₃₀₋₃₉) e, a seguir, utilizamos o programa GRASP2 [35] para análise da topologia e da distribuição de cargas dos complexos em estudo.

Construção do APL PA₂₂₄₋₂₃₃R7W

A seqüência FASTA do APL (SSLENFWAYV) foi carregada no SPDBV para o Fit sobre o epítopo HA₄₆₈₋₄₇₇ (1WBX) e a estrutura resultante foi utilizada para construção do pMHC, utilizando os mesmos parâmetros anteriormente descritos. O MHC utilizado para o *docking* também foi obtido a partir da estrutura 1WBX.

Análise da topologia e da distribuição de cargas

As análises da superfície dos pMHC formados foram realizadas com o programa GRASP2 [35], na plataforma Windows XP. O potencial na superfície foi calculado, conforme padrão do programa, em um intervalo de -10 a +10 KTeslas. Para comparação entre as imagens dos complexos, utilizamos uma grade com 270 células (18 linhas e 15 colunas), fazendo-se uma estimativa do número de células positivas (compartilhavam carga e topologia) e negativas (diferiam em pelo menos um dos aspectos).

Obtenção das imagens para o artigo

A sobreposição dos epítopos foi realizada no programa SPDBV 3.7., no entanto, as imagens destas sobreposições foram obtidas no programa PyMOL 1.0 [36]. As imagens dos complexos pMHC foram obtidas no programa GRASP2. Todas as imagens foram editadas no programa Adobe Photoshop CS2 v.9.0.

Contribuição dos Autores

DAA realizou todas as análises e escreveu o artigo. MMR colaborou na discussão dos resultados. GFV desenhou o estudo, coordenou as atividades e contribuiu tecnicamente com o trabalho. JABC supervisionou o grupo, colaborou tecnicamente nos

aspectos imunológicos e revisou o artigo. Todos os autores leram e aprovaram a versão final do manuscrito.

Agradecimentos

Agradecemos a colaboração do Grupo de Bioinformática Estrutural do Centro de Biotecnologia da UFRGS, no compartilhamento de informações sobre a instalação dos programas. Em especial, a aluna Eduarda Schultze, com quem mantivemos contato ao longo do trabalho discutindo parâmetros dos programas e a interpretação de alguns dados. Agradecemos o auxílio financeiro CNPq/CAPES.

Referências

1. Yewdell JW, Bennink JR: **Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses.** *Annu Rev Immunol* 1999, **17**:51-88.
2. Welsh RM, Selin LK, Szomolanyi-Tsuda E: **Immunological memory to viral infections.** *Annu Rev Immunol* 2004, **22**:711-743.
3. Wilson DB, Wilson DH, Schroder K, Pinilla C, Blondelle S, Houghten RA, Garcia KC: **Specificity and degeneracy of T cells.** *Mol Immunol* 2004, **40**(14-15):1047-1055.
4. Welsh RM, Fujinami RS: **Pathogenic epitopes, heterologous immunity and vaccine design.** *Nat Rev Microbiol* 2007, **5**(7):555-563.
5. Welsh RM, Selin LK: **No one is naive: the significance of heterologous T-cell immunity.** *Nat Rev Immunol* 2002, **2**(6):417-426.
6. Ang CW, Jacobs BC, Laman JD: **The Guillain-Barre syndrome: a true case of molecular mimicry.** *Trends Immunol* 2004, **25**(2):61-66.
7. Sandalova T, Michaelsson J, Harris RA, Odeberg J, Schneider G, Karre 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**(29):27069-27075.
8. 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**(23):11392-11400.
9. Clute SC, Watkin LB, Cornberg M, Naumov YN, Sullivan JL, Luzuriaga K, Welsh RM, Selin LK: **Cross-reactive influenza virus-specific CD8+ T cells contribute to lymphoproliferation in Epstein-Barr virus-associated infectious mononucleosis.** *J Clin Invest* 2005, **115**(12):3602-3612.
10. Ishikawa T, Kono D, Chung J, Fowler P, Theofilopoulos A, Kakumu S, Chisari FV: **Polyclonality and multispecificity of the CTL response to a single viral epitope.** *J Immunol* 1998, **161**(11):5842-5850.
11. Belz GT, Xie W, Altman JD, Doherty PC: **A previously unrecognized H-2D(b)-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**(8):3486-3493.
12. Morris GM, Goodsell, D.S., Halliday R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J.: **Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.** *J Computational Chemistry*, 1998, **19**(14).
13. Cavasotto CN, Orry AJ: **Ligand docking and structure-based virtual screening in drug discovery.** *Curr Top Med Chem* 2007, **7**(10):1006-1014.
14. Kmiecik S, Gront D, Kolinski A: **Towards the high-resolution protein structure prediction. Fast refinement of reduced models with all-atom force field.** *BMC Struct Biol* 2007, **7**:43.
15. Chen J, Brooks CL, 3rd: **Can molecular dynamics simulations provide high-resolution refinement of protein structure?** *Proteins* 2007, **67**(4):922-930.

16. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nucleic Acids Res* 2000, **28**(1):235-242.
17. Davies MN, Hattotuwigama CK, Moss DS, Drew MG, Flower DR: **Statistical deconvolution of enthalpic energetic contributions to MHC-peptide binding affinity.** *BMC Struct Biol* 2006, **6**:5.
18. 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**(5):873-879.
19. Madurga S, Belda I, Llorca X, Giralt E: **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* 2005, **14**(8):2069-2079.
20. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S: **SYFPEITHI: database for MHC ligands and peptide motifs.** *Immunogenetics* 1999, **50**(3-4):213-219.
21. Purcell AW, Elliott T: **Molecular machinations of the MHC-I peptide loading complex.** *Curr Opin Immunol* 2008, **20**(1):75-81.
22. Lehner PJ, Surman MJ, Cresswell P: **Soluble tapasin restores MHC class I expression and function in the tapasin-negative cell line .220.** *Immunity* 1998, **8**(2):221-231.
23. Wearsch PA, Cresswell P: **Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer.** *Nat Immunol* 2007, **8**(8):873-881.
24. 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**(1):106-113.
25. Velloso LM, Michaelsson J, Ljunggren HG, Schneider G, Achour A: **Determination of structural principles underlying three different modes of lymphocytic choriomeningitis virus escape from CTL recognition.** *J Immunol* 2004, **172**(9):5504-5511.
26. Zhao R, Loftus DJ, Appella E, Collins EJ: **Structural evidence of T cell xeno-reactivity in the absence of molecular mimicry.** *J Exp Med* 1999, **189**(2):359-370.
27. Becker CF, Guimarães JA, Mourão PAS, Verli H: **Conformation of sulfated galactan and sulfated fucan in aqueous solutions: Implications to their anticoagulant activities.** *J Mol Graph Model* 2007, **26**(1):391-399.
28. Jorgensen JL, Esser U, Fazekas de St Groth B, Reay PA, Davis MM: **Mapping T-cell receptor-peptide contacts by variant peptide immunization of single-chain transgenics.** *Nature* 1992, **355**(6357):224-230.
29. Kessels HW, de Visser KE, Tirion FH, Coccoris M, Kruisbeek AM, Schumacher TN: **The impact of self-tolerance on the polyclonal CD8+ T cell repertoire.** *J Immunol* 2004, **172**(4):2324-2331.
30. Parana R, Almeida D: **HBV epidemiology in Latin America.** *J Clin Virol* 2005, **34 Suppl 1**:S130-133.
31. 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**(1-2):33-41.
32. **Immune Epitope Database and Analysis Resource (IEDB)** [<http://www.immuneepitope.org/home.do>].
 33. Guex N, Peitsch MC: **SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling.** *Electrophoresis* 1997, **18**(15):2714-2723.
 34. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ: **GROMACS: fast, flexible, and free.** *J Comput Chem* 2005, **26**(16):1701-1718.
 35. Petrey D, Honig B: **GRASP2: visualization, surface properties, and electrostatics of macromolecular structures and sequences.** *Methods Enzymol* 2003, **374**:492-509.
 36. **PyMOL** [<http://pymol.sourceforge.net/>].

Figuras

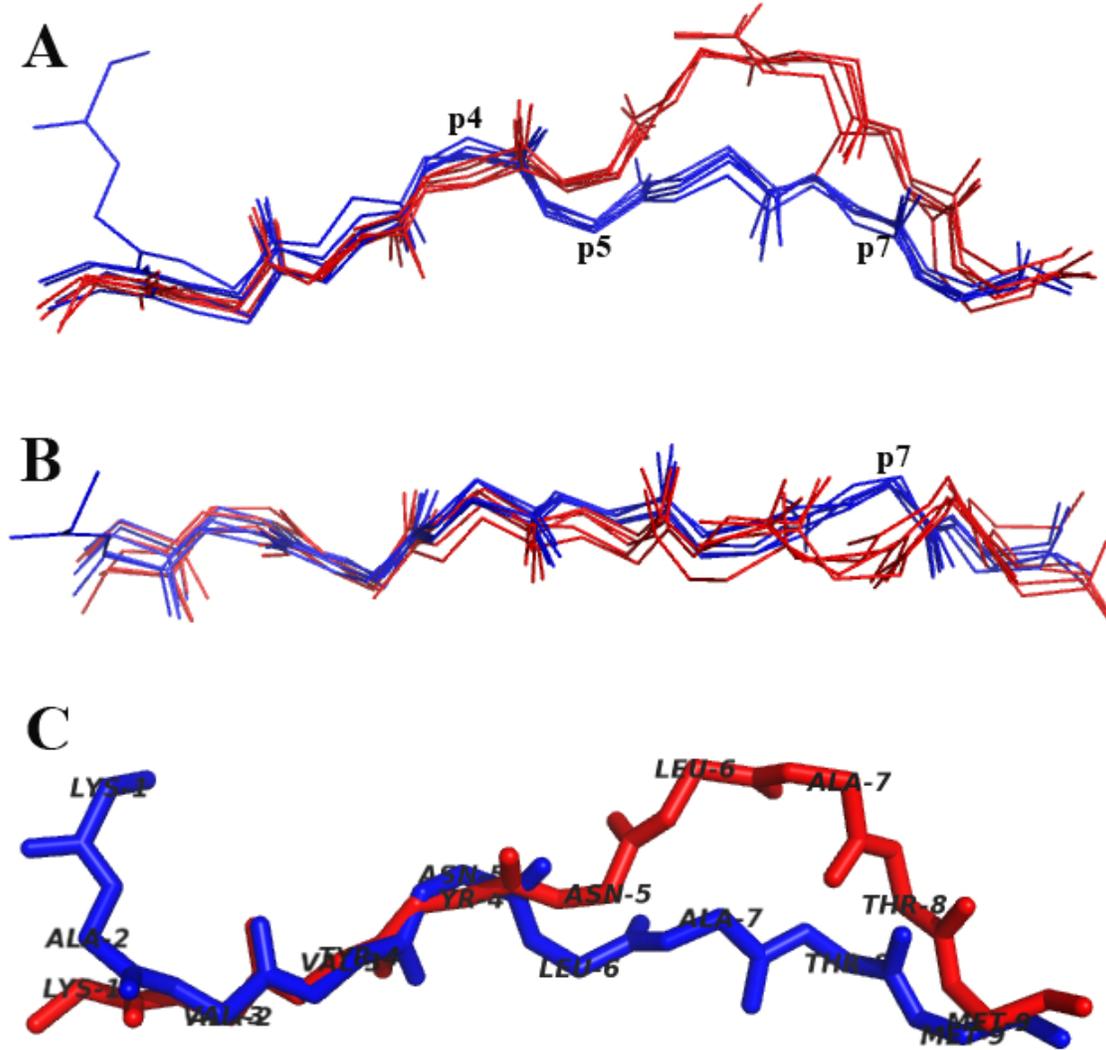


Figura 1 – Sobreposição de epitopos restritos a H-2D^b e a H-2K^b.

As estruturas de cinco epitopos restritos a H-2D^b (1CE6, 1S7V, 1WBZ, 1ZHB e 3BUY) e de cinco epitopos restritos a H-2K^b (1FO0, 1FZJ, 1LK2, 1RJY e 1S7R) foram sobrepostas. A: a vista lateral salienta as diferenças conformacionais entre os epitopos de H-2D^b, representados em vermelho, e os epitopos de H-2K^b, representados em azul, sobretudo entre as posições 5 e 7 (p5-7) dos peptídeos. O padrão de cores foi mantido nas imagens B e C. B: vista superior dos 10 peptídeos sobrepostos. C: As estruturas 1S7V e 1S7R apresentam o mesmo epitopo (KAVYNLATM) cristalografado no contexto dos dois diferentes alelos.

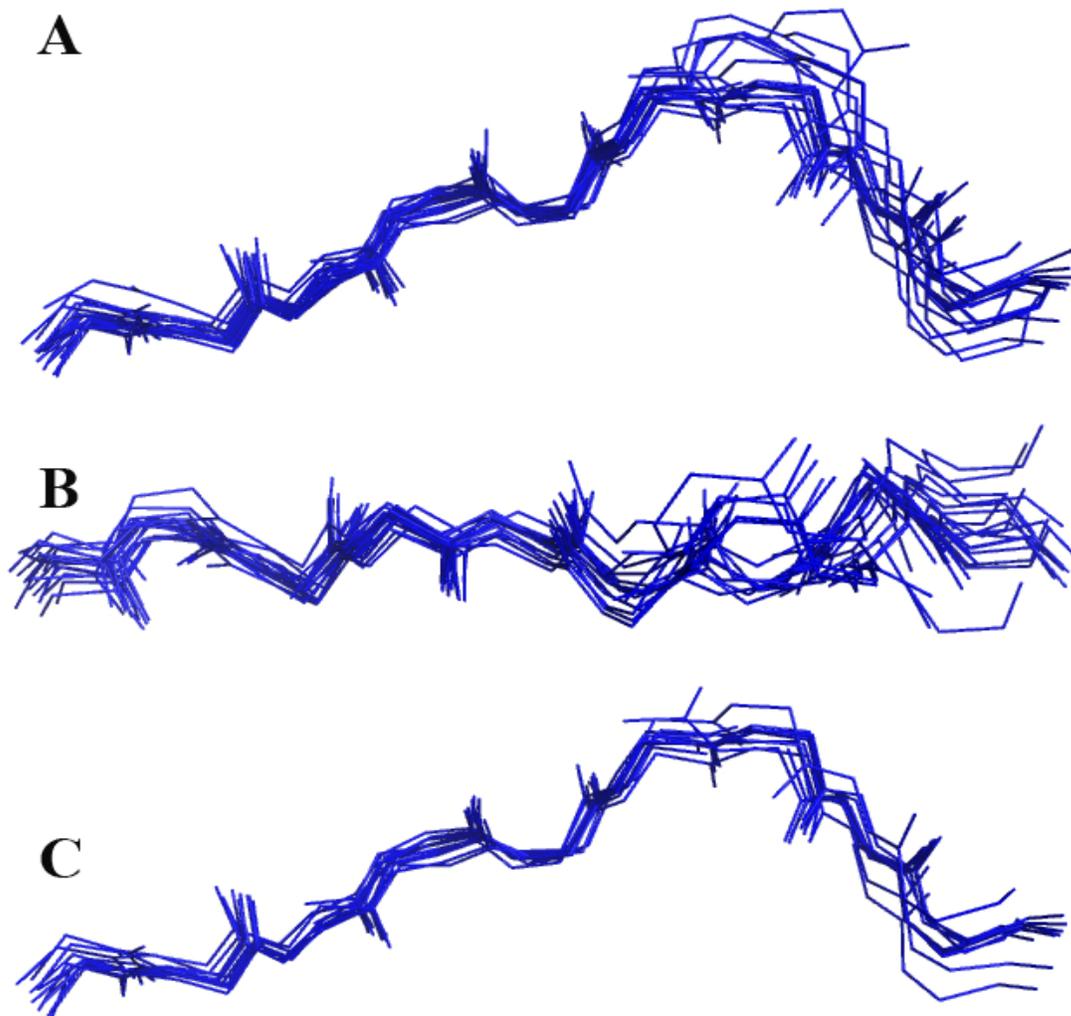


Figura 2 - Padrão conformacional dos epitopos apresentados por H-2D^b.

A: Dezesete epitopos (ver tabela 1), contendo entre 8 e 11 aminoácidos de extensão, cristalografados no contexto deste alelo, foram abertos e sobrepostos mantendo-se a mesma orientação de amino (à esquerda) pra carbóxi (à direita). Observou-se claramente a existência de um padrão compartilhado entre as cadeias principais dos epitopos. B: Os dezesete epitopos em visão panorâmica. C: Foram sobrepostos apenas os epitopos com nove aminoácidos, demonstrando que a maior variabilidade estrutural entre os epitopos era determinada pela diferença no comprimento das seqüências, no sentido de acomodar epitopos maiores (10-mer e 11-mer). As cadeias laterais dos epitopos não foram representadas nestas imagens, bem como um aminoácido (p7) do epitopo 2VE6, excluído por ter sido alterado e conter um anel fixo à cadeia principal (para maiores detalhes verifique o Arquivo Adicional 5).

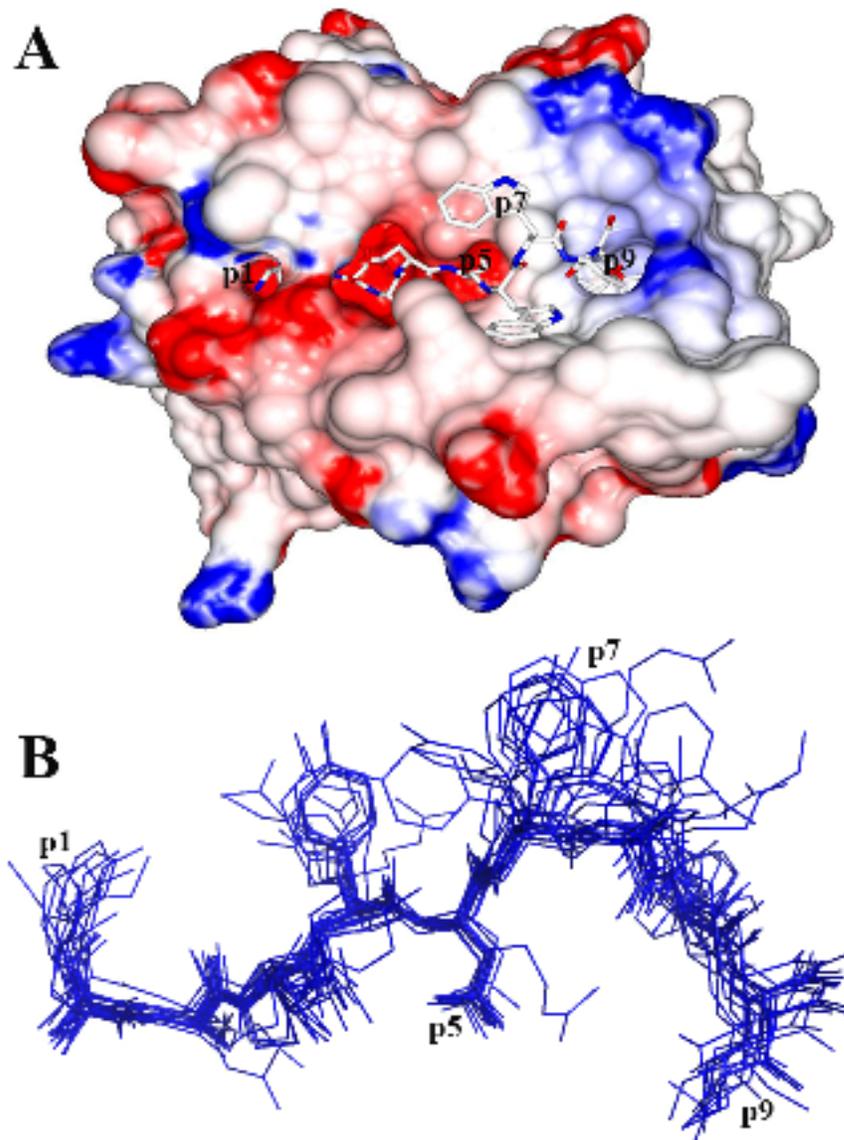


Figura 3 - Organização estrutural dos epitopos apresentados em H-2D^b.

A. Imagem do epitopo HBsAg₃₀₋₃₉ (apresentado em *Ball and Stick_CPK*) na fenda do alelo H-2D^b (representado com a superfície molecular colorida conforme a distribuição de cargas). Parte da extremidade N-term do epitopo está oculta sob cadeias laterais do MHC (p2-3) enquanto outra parte (p6-7) encontra-se externalizada à fenda do MHC. B. Sobreposição da estrutura de 17 epitopos restritos a H-2D^b (tabela 1), incluindo cadeias laterais. É possível observar maior variabilidade entre os epitopos na região em que este se externaliza ao MHC e pouca variabilidade na região oculta (N-term) e nos aminoácidos âncoras (p5 e p9). As posições das cadeias laterais dos aminoácidos 1,5,7 e 9 são indicadas em ambas as imagens.

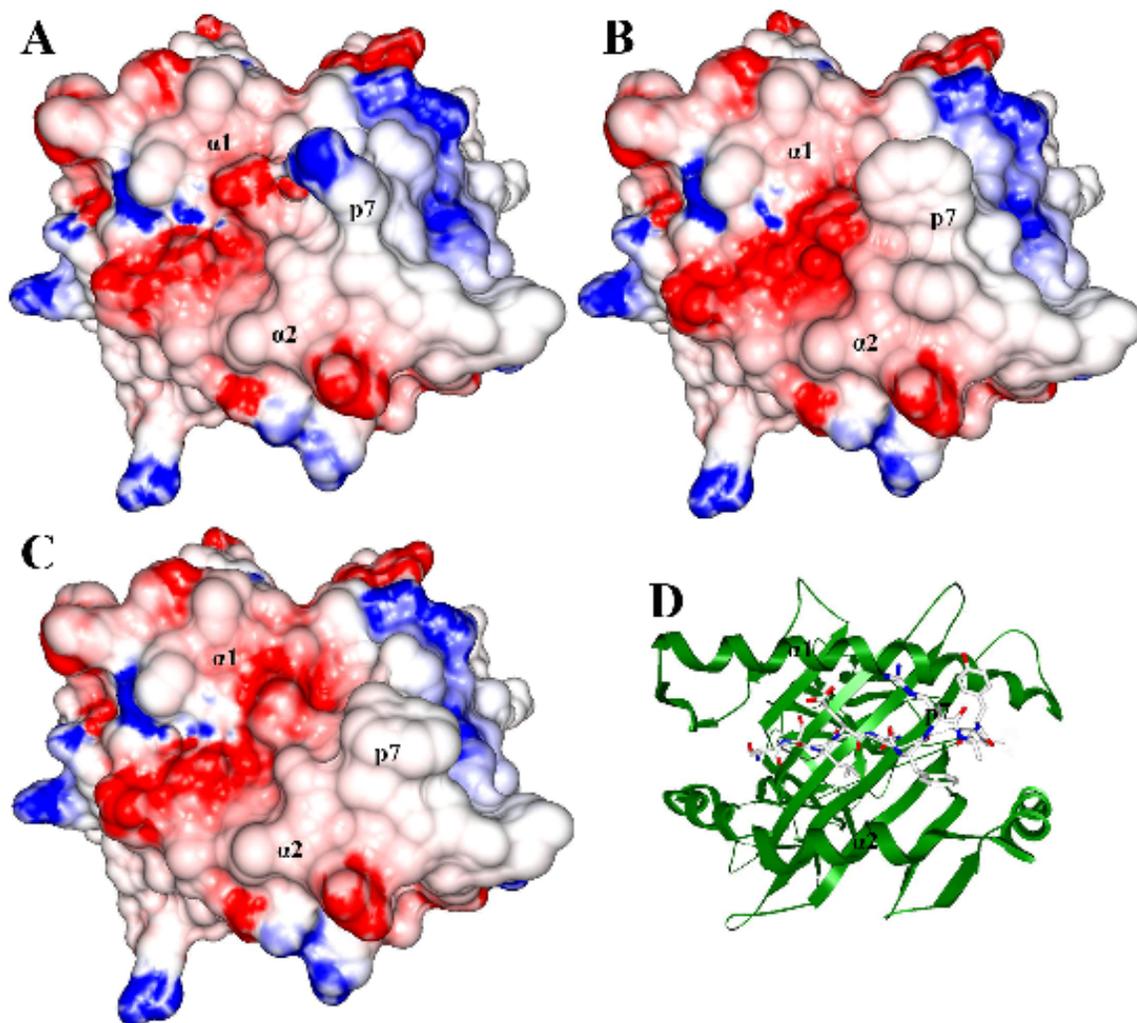


Figura 4 - Comparação dos complexos pMHC quanto a topologia e carga.

As imagens apresentam uma vista panorâmica da superfície de interação com o TCR dos diferentes complexos pMHC. Regiões com carga positiva (azul) e negativa (vermelho) são representadas em uma escala de -10 a +10 kiloteslas. A: 1WBY (SSLENFRAYV). B: Complexo D^b HBsAg₃₀₋₃₉ (QSLDSWWTSL). C: APL (SSLENFWAYV). D: Imagem da estrutura 1WBY com o MHC representado em *Ribon* e o epitopo em *Ball and Stick*, no mesmo tamanho e posição dos demais complexos. São indicados ainda os domínios $\alpha 1$ e $\alpha 2$ do MHC, bem como a posição 7 dos epitopos.

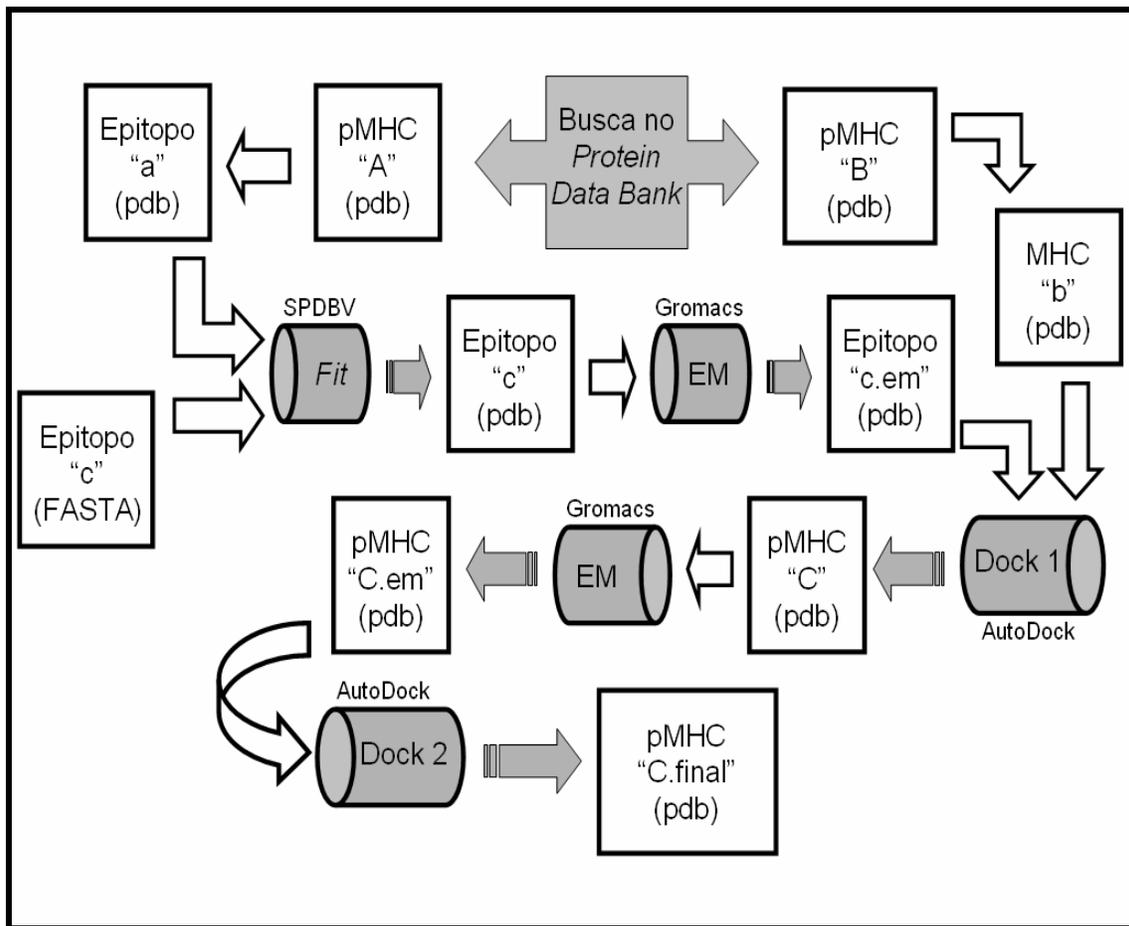


Figura 5 - Fluxograma da construção de um novo complexo pMHC.

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).

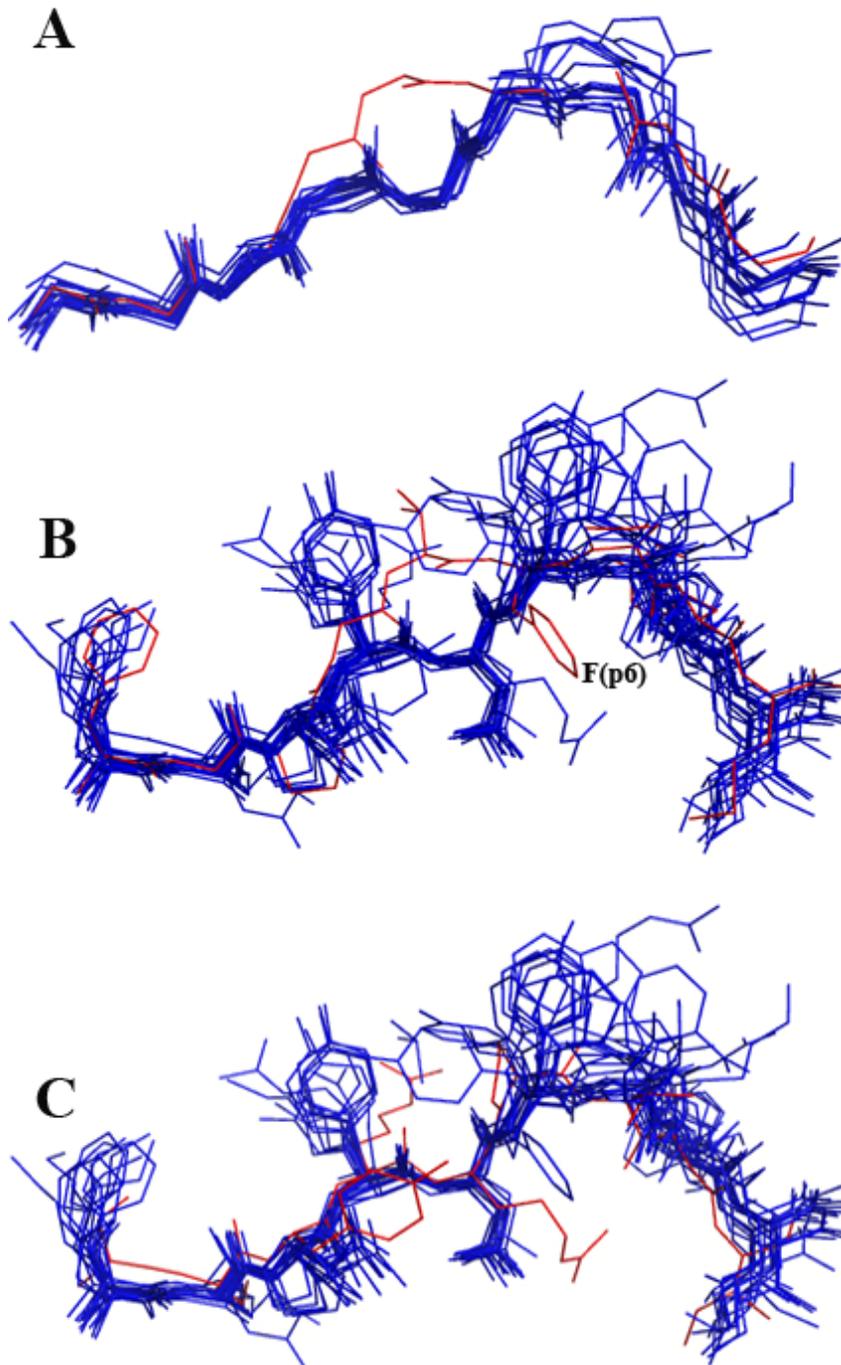
Tabelas

Tabela 1- Relação dos ligantes de H-2D^b disponíveis no PDB.

Código	Descrição do epitopo	Sequência (aa)	Nº de aa
1BZ9*	Peptídeo sintético p1027	FAPGVFPYM	9
1CE6*	Epitopo da Nucleoproteína do vírus Sendai (NP ₃₂₄₋₃₃₂)	FAPGNYPAL	9
1FG2*	Epitopo de LCMV (gp33)	LAVYNFATC	9
1JPF*	Epitopo de LCMV (gp276)	SGVENPGGYCL	11
1JPG*	Epitopo de LCMV (np396)	FQPQNGQFI	9
1JUF*	Peptídeo de antígeno secundário de histocompatibilidade (h13b)	SSVIGVWYL	9
1N3N*	Epitopo de hsp60	SALQNAASIA	10
1QLF*	Sendai NP 324-332 (<i>gly327 replaced by o-glcnae serine</i>)	FAPSNYPAL	9
1S7U/2F74*	Epitopo imunodominante de LCMV (gp33)	KAVYNFATM	9
1S7V*	Epitopo de LCMV (gp33) - variante de escape viral	KAVYNLATM	9
1S7W*	Epitopo de LCMV (gp33) - variante de escape viral	KALYNFATM	9
1S7X*	Epitopo de LCMV (gp33) - variante de escape viral	KAVFNFATM	9
1WBX*	Epitopo de Influenza A (pr8) HA ₄₆₈₋₄₇₇	SQLKNNAKEI	10
1WBY/1YN6*	Epitopo de Influenza A (pr8) PA ₂₂₄₋₂₃₃	SSLENFRAYV	10
1YN7*	Peptídeo mutado (R7A) de RNA polimerase (P2)	SSLENFAAYV	10
1WBZ*	Epitopo de Influenza A (pr8) PB1 ₇₀₃	SSYRPVGI	8
1ZHB*	Peptídeo 9-mer de Dopamine beta-monooxygenase	KALYNYAPI	9
2VE6	Epitopo alterado do vírus Sendai (<i>Photocleavable peptide</i>)	FAPGNYPAL	9
3BUY*	Epitopo de PB1-F2	LSLRNPILV	9

Foram encontrados 21 arquivos contendo o alelo H-2D^b no PDB, totalizando 18 (*) estruturas deste alelo com diferentes epitopos. Estes 18 epitopos foram utilizados nas análises de estrutura e sequência. O epitopo 1BZ9 foi incluído em todas as análises, mas foi excluído na Figuras 1 e na figura 3B.

Arquivos Adicionais



Arquivo Adicional 1 – Exceções ao padrão de H-2D^b.

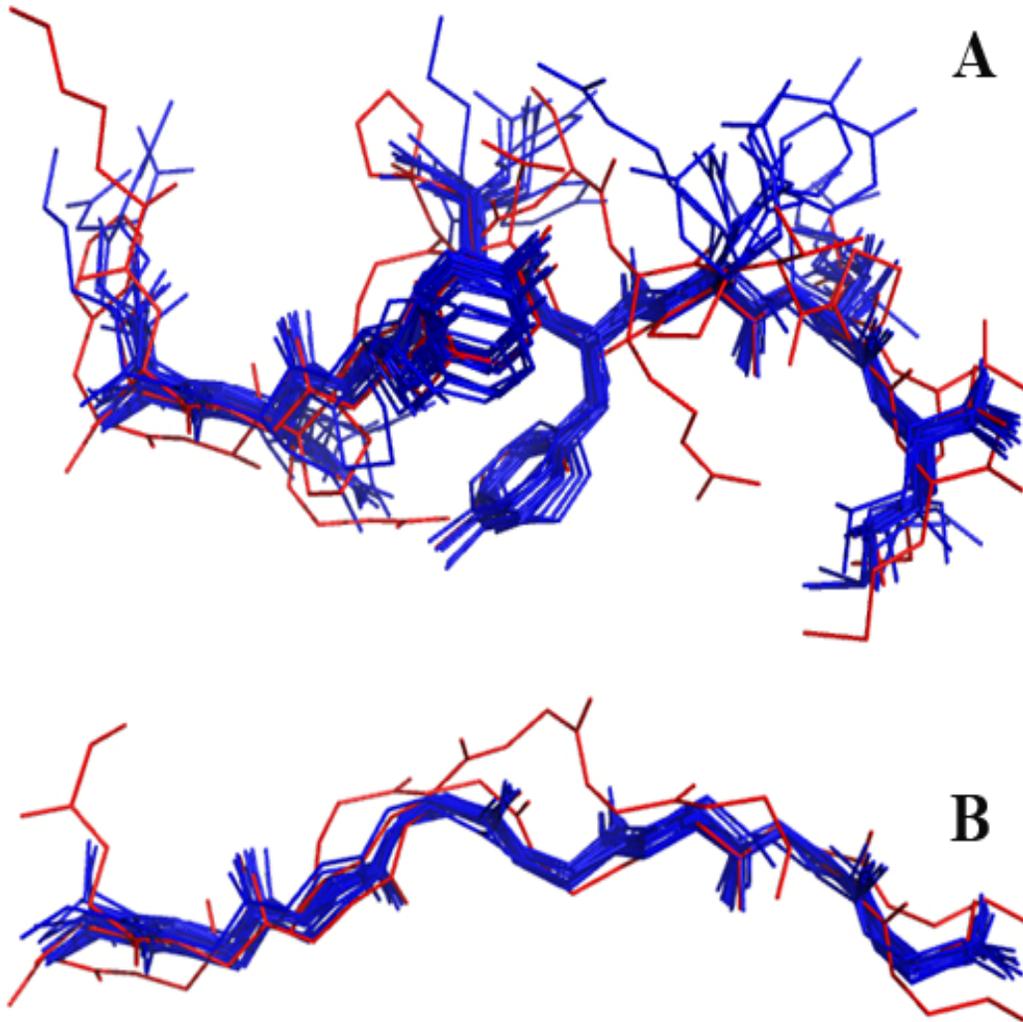
A: O epitopo 1BZ9 (em vermelho) não possui o aminoácido convencional na posição de âncora (p5) e apresentou um importante desvio na cadeia principal quando comparado aos demais epitopos. B: as cadeias laterais do epitopo 1BZ9 (também em vermelho) comportam-se de forma semelhante a dos demais epitopos deste alelo, excetuando-se

uma fenilalanina na posição 6 do epitopo, que talvez seja utilizada com uma “âncora” alternativa. A presença do aminoácido âncora em C-term e a extremidade N-term “presa” sob as cadeias laterais do MHC devem colaborar para que mesmo com esta conformação incomum, este epitopo consiga ser apresentado. C: O epitopo 1WBZ (em vermelho) apresenta um desvio na cadeia lateral do aminoácido âncora (p5), quando comparado aos demais epitopos. Este epitopo foi o único 8-mer encontrado no contexto do H-D^b (com estrutura depositada no PDB) e possui uma prolina (não polar) como âncora em p5.

Arquivo Adicional 2 - Relação dos ligantes de H-2K^b disponíveis no PDB.

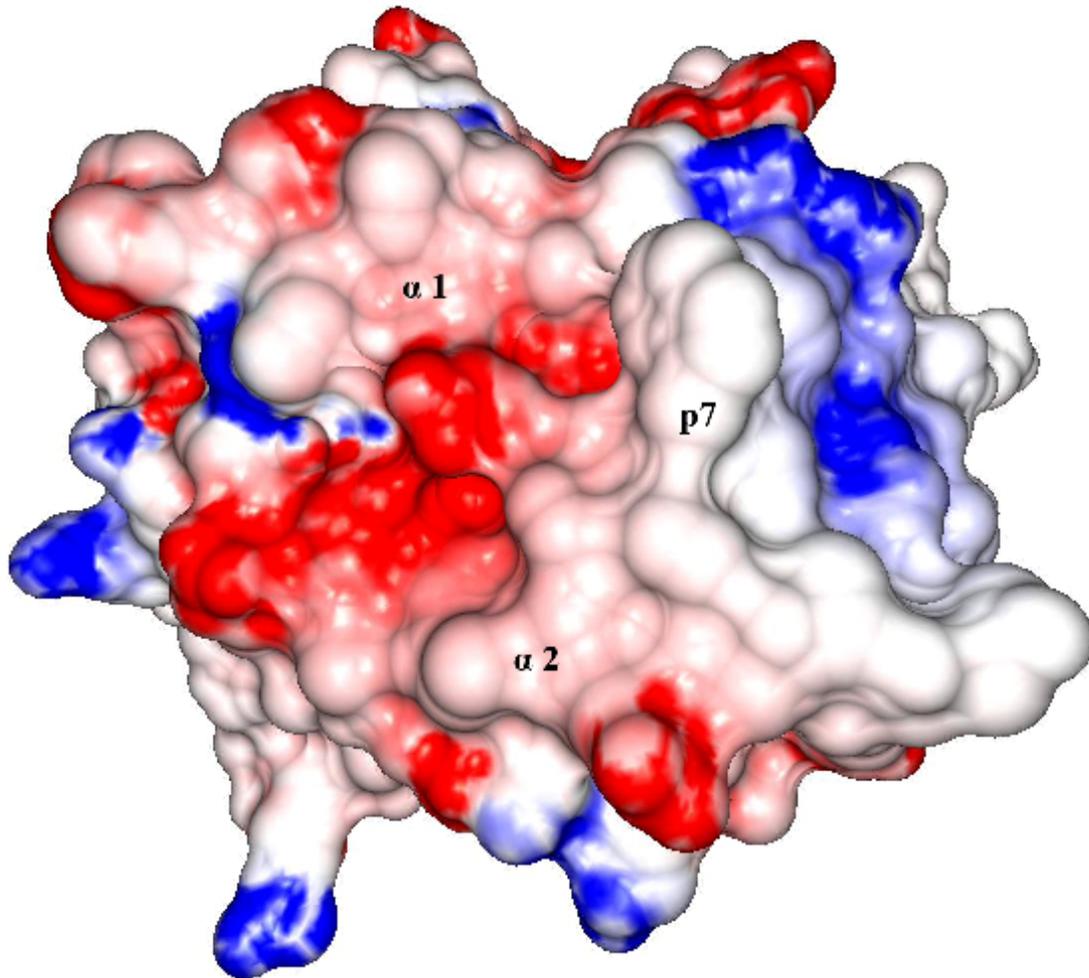
Código:	Descrição do epitopo	Sequência	Nº de aa
1FO0*/2CLZ 1NAN	Epitopo derivado de Pbm1	INFDFNIT	8
1FZJ/1KPU/1NAM 2VAA/1FZM*	Epitopo de VSV NP ₅₂₋₅₉	RGYVYQGL	8
1FZK*/1FZO/ 1KPV/2VAB	Epitopo do vírus Sendai NP ₃₂₄₋₃₃₂	FAPGNYPAL	9
1G7P*	Peptídeo de Alfa-Glicosidase de levedura (438-446)	SRDHSRTPM	9
1G7Q*	Peptídeo de Muc1 (180-187)	SAPDTRPA	8
1KJ3*	Epitopo derivado de Pkb1	KVITFIDL	8
1LEG/1LEK*	Dev _{8,54-61} (NADH-Ubiquinone Oxidoreductase MLRQ)	EQYKFYSV	8
1LK2*	Peptídeo derivado de receptor de insulina (423-430)	GNYSFYAL	8
1N59*/1S7Q	Epitopo de LCMV (Gp33)	AVYNFATM	8
1OSZ*	Epitopo mutado (L4V) da Np de VSV	AGYLYQGL	8
1P1Z*/1P4L/1VAC 3C8K/2QRI	Peptídeo derivado de Ovoalbumina (257-264)	SIINFEKL	8
2QRS/2QRT			
1RJY*/1RKO/ 1T0M/1TON	Epitopo da Glicoproteína B de Herpes (HSV)	SSIEFARL	8
1RJZ*/1RK1	Epitopo mutado S2E da Glicoproteína B de HSV	SEIEFARL	8
1S7R*	Epitopo de LCMV (Gp33) – variante de escape	KAVYNLATM	9
1S7S*	Epitopo de LCMV (Gp33) – variante de escape	ALYNFATM	8
1S7T*	Epitopo de LCMV (Gp33) – variante de escape	AVFNFATM	8
2CLV*	Peptídeo derivado de Rbm5 (189-196)	SGYYYNSL	8
2OL3*	Epitopo derivado de Pbm8	SQYYYNSL	8

Esta tabela apresenta os 39 arquivos contendo o alelo H-2K^b disponíveis no PDB. Destes, 18 estruturas (*) apresentavam epitopos com sequência diferindo em pelo menos um aminoácido. Estes 18 epitopos foram utilizados nas análises.



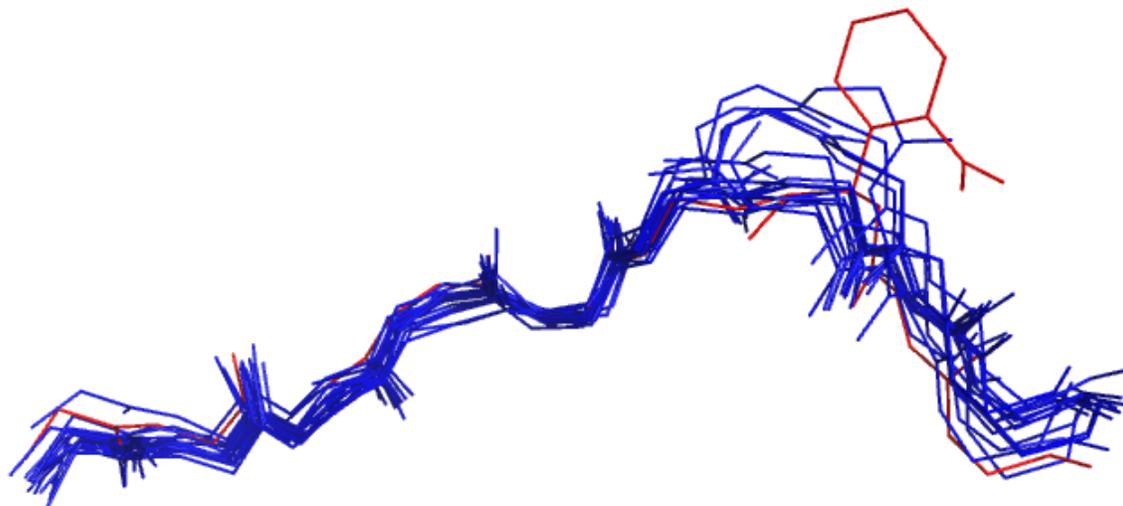
Arquivo Adicional 3 – Padrão estrutural dos ligantes de H-2K^b.

A: 18 epitopos (vide AA2) restritos a este alelo foram sobrepostos, sendo três 9-mers (vermelho) e 15 8-mers (azul). Assim como os epitopos de H-2D^b, os ligantes de H-2K^b apresentam uma maior identidade de cadeias laterais nos aminoácidos ancoras, enquanto uma maior divergência é visível entre as cadeias laterais voltadas para fora da fenda. B: A sobreposição da cadeia principal dos 18 ligantes revela uma conformação compartilhada entre as cadeias principais dos epitopos do mesmo comprimento (8-mer). O ajuste de cadeia neste alelo parece ocorrer em uma posição diferente do descrito para H-2D^b, aparentemente entre p3 e p4, sendo muito mais conservados em p6-7.



Arquivo Adicional 4 – Resultados do *docking* indicam variação em p7 no APL.

Estrutura alternativa apresentada pelo *docking* como resultado para a ligação do APL (PA₂₂₄₋₂₃₃R7W) ao H-2D^b ($BE = -15,93$ kcal/mol). A diferença de topologia mais intensa entre este complexo, o complexo com melhor resultado neste *docking* ($BE = -17,5$ kcal/mol) e o complexo D^bHBsAg₃₀₋₃₉ é conferida pela rotação do anel indólico do triptofano, na posição 7 do epitopo (p7). Como tanto o APL quanto o epitopo de HBV possuem este mesmo anel, na mesma posição, provavelmente *in vivo* esta diferença topológica será minimizada. Esta imagem foi obtida com o programa GRASP2 e a superfície do complexo esta colorida de acordo com a distribuição de cargas em um gradiente de -10 (azul) a +10 (azul) kiloteslas.



Arquivo Adicional 5 – Diferença estrutural do epitopo 1CE6.

Este arquivo do PDB (1CE6) possui o alelo H-2D^b complexado a um epitopo derivado da proteína NP do vírus Sendai. Este epitopo, no entanto, foi alterado na posição 7 para tornar-se um peptídeo fotoclivável (*3-amino-3-(2-nitro)phenyl-propionic acid*). Esta estrutura acrescentada à cadeia principal do epitopo faz com que este desvie do padrão proposto para a cadeia principal de epitopos apresentados por este alelo de MHC. No entanto, esta estrutura se localiza exatamente na região onde foi evidenciada maior variação de estruturas entre as cadeias laterais, conforme representado na Figura 3B do artigo, a qual inclui o epitopo 2VE6.

Conclusões e Perspectivas

Conforme apresentado, o MHC-I representa um passo vital na resposta imune celular. Sua interação com o TCR está envolvida em uma série de eventos responsáveis não só pelo desfecho – lise de uma célula alvo, por exemplo – mas também pela própria estruturação do sistema imune adaptativo, exercendo, por exemplo, importante papel na seleção de linfócitos e na manutenção de memória imunológica. Além disso, esta rota que envolve o MHC-I, é a principal forma de defesa contra duas das mais importantes patologias celulares: infecções virais e neoplasias.

Não por acaso, entender o funcionamento desta molécula e sua interação com o TCR, tem sido o objeto de estudo de muitos imunologistas desde sua descoberta em camundongos, nos anos 40, pelo grupo de George Snell (Fischer Lindahl, 1997). As características dos epitopos capazes de se ligar a cada alelo de MHC, na maior parte obtida por estudos *in vitro*, foram repetidamente analisadas até que padrões de afinidade emergissem destes estudos. Foram identificados os aminoácidos “âncoras” de inúmeros alelos e a predição da capacidade de ligação de peptídeos ao MHC tornou-se um novo e promissor campo de estudo. Neste contexto, a agilidade e o poder de processamento da informática foram recrutados possibilitando a busca rápida em bancos de ligantes de MHC - cada vez maiores – e a construção de preditores *on-line*, capazes de fornecer em segundos um escore de ligação de uma seqüência qualquer de aminoácidos ao alelo de interesse.

Também por motivos óbvios, a análise de ligação ao MHC passou a ser empregada na identificação de alvos para resposta imune a tumores e a infecções virais. Em ambos os casos, a forte pressão seletiva do sistema imune sobre células apresentando peptídeos não-próprios e o fraco controle sobre mutações (nos vírus e nas células tumorais) leva ao estabelecimento de uma série de mecanismos de evasão ao sistema imune. O estudo destes mecanismos de escape - sem os quais não haveria infecção viral produtiva ou evolução de um quadro tumoral - forneceu mais pistas sobre as características de um peptídeo imunógeno. A ligação de um peptídeo ao MHC era o ponto mais importante, mas não era o suficiente quando se desejava construir uma vacina. As outras etapas da rota de apresentação de peptídeos endógenos também eram essenciais para a correta

apresentação deste alvo e além disso, a correta apresentação não garantia a desejada estimulação dos linfócitos T CD8+. Na tentativa de reproduzir *in silico* etapas do sistema imune, bem como desenvolver novas ferramentas capazes de identificar alvos com potencial para gerar uma resposta – tanto celular quanto humoral – desenvolveu-se uma nova área de pesquisa, denominada imunoinformática (ou bioinformática imunológica). A ambição maior neste campo é reproduzir todo o sistema imune *in silico*, combinando a informação gerada com dados experimentais, e visando dar suporte a diversas aplicações clínicas. A União Européia tem liderado estes estudos com a criação do ImmunoGrid, um consórcio multicêntrico que pretende reproduzir o sistema imune humano (*The European Virtual Human Immune System Project*).

Como parte deste trabalho, em fevereiro deste ano Todman e colaboradores publicaram o desenvolvimento de um servidor *on-line* - MHCsim - capaz de construir complexos pMHC de alelos de MHC humanos (HLAs) com qualquer epitopo de interesse (Todman e cols, 2008). A estratégia utilizada consiste em buscar em um banco de estruturas de pMHCs aquela cuja sequência do epitopo apresentado mais se assemelhe com a sequência fornecida pelo usuário. Depois de identificado, este “molde” é mutado nas posições em que há divergência até que se obtenha o complexo pMHC de interesse. Antes de fornecer a estrutura o servidor realiza ainda uma curta Dinâmica Molecular (MD), processo semelhante à EM descrita neste trabalho, mas que gera mais dados sobre a afinidade do peptídeo ao MHC.

Apesar de fazer parte de um grande projeto de cunho imunológico, o artigo que descreve a ferramenta não faz nenhum comentário a respeito do embasamento imunológico de tal estratégia. Será que simplesmente alterar os aminoácidos de uma estrutura cristalografada é um procedimento confiável para se gerar um complexo pMHC completamente novo? Qual o número máximo de mutações que uma estrutura pode sofrer sem perder a validade como molde?

A estratégia descrita no presente trabalho está fundamentada na evidência de que epitopos apresentados por um mesmo alelo de MHC compartilham uma conformação da cadeia principal. O novo epitopo é então construído sobre este molde – válido apenas para a cadeia principal – e é realizado um processo de encaixe (*docking*) deste epitopo na estrutura do MHC, onde as cadeias laterais estarão completamente livres para

rotacionar até encontrarem uma conformação mais favorável. O *docking* é uma ferramenta já estabelecida na literatura, empregado inclusive para abordagens imunológicas. Após o *docking* realizamos uma minimização de energia, que ajustará cadeias laterais do MHC, favorecendo a interação com este novo ligante. A segunda rodada de *docking* irá permitir um refinamento da busca pela conformação mais adequada do epitopo na fenda do MHC.

Embora baseados em estratégias distintas, os resultados de ambos os procedimentos – o desenvolvido por nosso grupo e o oferecido pelo MHCsim – devem ser semelhantes, uma vez que as mutações induzidas pelo MHCsim devem afetar sobretudo as cadeias laterais e o passo de solvatação (MD) irá acomodar as cadeias laterais tanto do peptídeo quanto do MHC. De certo modo, a identificação – por nosso grupo - de um padrão conformacional alelo-específico fornece a base imunológica da qual o MHCsim necessitava.

Um ponto importante, brevemente abordado neste trabalho, é o fato de que usando nossa estratégia se pode construir complexos pMHCs com quaisquer epitopo e alelo de MHC desejado. Sempre se obterá um pMHC. Contudo existe uma importante implicação imunológica nesta abordagem, pois, nem todos os peptídeos serão capazes de se ligar a todos os MHCs. Em nosso procedimento, o peptídeo será submetido ao *docking* e, caso seja incompatível com a fenda do MHC ou possua aminoácidos desfavoráveis à ligação, este peptídeo apresentará baixa energia livre de ligação, e o pesquisador saberá que o pMHC formado não é confiável. O mesmo problema se aplica ao MHCsim e, neste caso, o processo de solvatação trará os dados de afinidade do ligante pelo MHC. Em ambos os casos é, portanto, necessário estabelecer-se um “ponto de corte”, baseado em estruturas cristalografadas ou em ensaios *in vitro*, que defina quais valores indicam que o ligante em estudo provavelmente não conseguirá ser apresentado pelo alelo desejado. Uma perspectiva de nosso grupo é realizar um detalhado experimento de comparação entre ambos os procedimentos, tanto com um bom ligante, quanto com um epitopo descrito na literatura como sendo uma variante de escape viral, incapaz de ser apresentada no contexto de determinado MHC. Assim teremos uma visão mais clara acerca das semelhanças e diferenças entre os procedimentos e também, acerca da sensibilidade de ambos na detecção de maus

ligantes.

Considerando que ambos os procedimentos sejam confiáveis, a grande vantagem do MHCsim é a automação do serviço. O usuário não precisa entender os procedimentos, basta fornecer os dados ao servidor e aguardar o resultado. Contudo, o MHCsim só consegue construir complexos pMHC com HLAs, enquanto o nosso procedimento permite o estudo de qualquer alelo de MHC que já tenha sido cristalografado. Para contornar a questão da complexidade do processo, nosso grupo também pretende automatizar a construção de pMHCs, facilitando sua aplicação por outros grupos e possibilitando sua disponibilização *on line*.

Esta automatização independe da determinação do padrão estrutural de outros alelos, desde que o usuário forneça o arquivo PDB que considera ser um molde válido para seu trabalho. Apesar disso, nosso grupo pretende ainda estabelecer o padrão de outros alelos de MHC, pertencentes a outros organismos (incluindo alguns alelos humanos), a fim de consolidar a existência deste padrão alelo-específico, além de fornecer informações sobre as características da interação de peptídeos com outros alelos de MHC utilizados em pesquisa.

No presente trabalho, desenvolvemos um estudo de reatividade cruzada entre os epitopos virais PA₂₂₄₋₂₃₃ (IV) e HBsAg₂₈₋₃₉ (HBV). Propusemos ainda a construção de um APL, baseado no epitopo de IV, com grande possibilidade de induzir resposta cruzada contra o epitopo de HBV. Cabe ressaltar que, neste caso, ambos os epitopos são imunógenos conhecidos, e que o epitopo PA₂₂₄₋₂₃₃ é um epitopo imunodominante, com populações de linfócitos específicos frequentemente presentes em indivíduos infectados pelo agente. Embora menos provável que a reação cruzada contra HBV, existe ainda, em função desta grande capacidade do epitopo de IV em estimular linfócitos, a possibilidade de ocorrência de resposta cruzada entre o APL e o epitopo original (PA₂₂₄₋₂₃₃), o que tornaria este APL um forte candidato a inclusão nas vacinas contra gripe, atuando na geração de imunidade e memória contra ambos os alvos (IV e HBV).

Este estudo de reatividade cruzada é apenas um exemplo da possível aplicação do uso combinado de um procedimento para construção de pMHCs e da análise da superfície dos complexos gerados, sobretudo quanto a distribuição de cargas. Esta estratégia *in silico* pode ser usada na busca por seqüências alvos em vírus, bactérias

intracelulares e proteínas tumorais, reduzindo etapas e custos de projetos que visem desenvolver novas vacinas contra estas patologias celulares, ou simplesmente estudar a atuação do sistema imune nestas situações.

A possibilidade de aplicação na identificação de alvos para futuras vacinas é certamente o que mais chama atenção. Neste sentido, o segundo *docking* (Dock) pode ser melhor explorado, fornecendo dados sobre quais átomos do epitopo estão interagindo com quais átomos do MHC e ainda o tipo de ligação que sustenta esta interação (Ligações de Hidrogênio ou *Van Der Waals*). Além disso, aqueles que forem identificados como sendo os melhores ligantes, podem ser submetidos a um passo posterior de Dinâmica Molecular (MD). Apesar de requerer um gasto computacional um pouco maior, esta técnica permite ter dados mais confiáveis de afinidade do epitopo pelo MHC e de estabilidade do complexo pMHC em solução. Deste modo, os epitopos que chegarem ao final do estudo certamente serão capazes de ser apresentados no contexto do MHC de interesse, com uma margem de erro muito pequena. De forma complementar, os dados de semelhança estrutural/potencial com estruturas descritas na literatura como sendo bons imunógenos já podem fornecer a pista da capacidade deste complexo em ativar populações de linfócitos.

De fato, os dados gerados por este trabalho estão sendo utilizados na construção de uma proposta completamente nova no desenvolvimento de vacinas. Esperamos em breve construir uma verdadeira linha de montagem envolvendo atividades *in silico*, *in vitro* e *in vivo*, as quais poderiam ser direcionadas para qualquer alvo viral de interesse. Uma vez definido o alvo, seria realizada uma busca em larga escala por seqüências deste agente que contenham os pré-requisitos para a correta clivagem no proteossomo, transporte pela TAP e desbastamento no retículo, de modo que o peptídeo final ainda possua resíduos capazes de oferecer a ancoragem a algum dos alelos de MHC do organismo hospedeiro. Esta busca, combinando dados da literatura e de preditores *on-line*, levaria ao levantamento de seqüências candidatas, as quais seriam então submetidas a um estudo estrutural da capacidade de ligação destes epitopos aos MHCs do hospedeiro, identificação das conformações mais estáveis na fenda, e inferência da possível ocorrência de resposta cruzada contra proteínas próprias e contra outros vírus de interesse. Os epitopos com os melhores resultados em todas as análises poderiam ser

então sintetizados e testados quanto a real capacidade de estimular a resposta citotóxica de linfócitos (*in vitro*) e memória imunológica (*in vivo*).

Referências Adicionais

- Brehm, M. A., L. K. Selin, et al. (2004). "CD8 T cell responses to viral infections in sequence." Cell Microbiol **6**(5): 411-21.
- Cavasotto, C. N. and A. J. Orry (2007). "Ligand docking and structure-based virtual screening in drug discovery." Curr Top Med Chem **7**(10): 1006-14.
- Coomes, J. L. and F. Powrie (2008). "Dendritic cells in intestinal immune regulation." Nat Rev Immunol **8**(6): 435-46.
- Falk, K. and O. Rotzschke (2002). "The final cut: how ERAP1 trims MHC ligands to size." Nat Immunol **3**(12): 1121-2.
- Fischer Lindahl, K. (1997). "On naming H2 haplotypes: functional significance of MHC class Ib alleles." Immunogenetics **46**(1): 53-62.
- Godfrey, D. I., J. Rossjohn, et al. (2008). "The fidelity, occasional promiscuity, and versatility of T cell receptor recognition." Immunity **28**(3): 304-14.
- Hammer, G. E., T. Kanaseki, et al. (2007). "The final touches make perfect the peptide-MHC class I repertoire." Immunity **26**(4): 397-406.
- ImmunoGrid - The European Virtual Human Immune System Project.
<http://www.immunogrid.org/immunogrid>.
- Kmiecik, S., D. Gront, et al. (2007). "Towards the high-resolution protein structure prediction. Fast refinement of reduced models with all-atom force field." BMC Struct Biol **7**: 43.
- Mason, D. (1998). "A very high level of crossreactivity is an essential feature of the T-cell receptor." Immunol Today **19**(9): 395-404.
- Meijers, R., C. C. Lai, et al. (2005). "Crystal structures of murine MHC Class I H-2 D(b) and K(b) molecules in complex with CTL epitopes from influenza A virus: implications for TCR repertoire selection and immunodominance." J Mol Biol **345**(5): 1099-110.
- Morris, G. M., Goodsell, D.S., Halliday R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J. (1998). "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function." J. Computational Chemistry, **19**(14).
- Nel, A. E. (2002). "T-cell activation through the antigen receptor. Part 1: signaling components, signaling pathways, and signal integration at the T-cell antigen receptor synapse." J Allergy Clin Immunol **109**(5): 758-70.
- Paulsson, K. M. (2004). "Evolutionary and functional perspectives of the major histocompatibility complex class I antigen-processing machinery." Cell Mol Life Sci **61**(19-20): 2446-60.
- Purcell, A. W. and T. Elliott (2008). "Molecular machinations of the MHC-I peptide loading complex." Curr Opin Immunol **20**(1): 75-81.
- Reali, E., R. Guerrini, et al. (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**(1): 106-13.
- Regner, M. (2001). "Cross-reactivity in T-cell antigen recognition." Immunol Cell Biol **79**(2): 91-100.
- Sandalova, T., J. Michaelsson, et al. (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**(29): 27069-75.

- Seder, R. A., P. A. Darrah, et al. (2008). "T-cell quality in memory and protection: implications for vaccine design." Nat Rev Immunol **8**(4): 247-58.
- Selin, L. K., M. Cornberg, et al. (2004). "CD8 memory T cells: cross-reactivity and heterologous immunity." Semin Immunol **16**(5): 335-47.
- Sohn, S. J., J. Thompson, et al. (2007). "Apoptosis during negative selection of autoreactive thymocytes." Curr Opin Immunol **19**(5): 510-5.
- Todman, S. J., M. D. Halling-Brown, et al. (2008). "Toward the atomistic simulation of T cell epitopes automated construction of MHC: peptide structures for free energy calculations." J Mol Graph Model **26**(6): 957-61.
- Vieira, G. F. and J. A. Chies (2005). "Immunodominant viral peptides as determinants of cross-reactivity in the immune system--Can we develop wide spectrum viral vaccines?" Med Hypotheses **65**(5): 873-9.
- Wang, J. and M. A. Maldonado (2006). "The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases." Cell Mol Immunol **3**(4): 255-61.
- Wearsch, P. A. and P. Cresswell (2007). "Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer." Nat Immunol **8**(8): 873-81.
- Wedemeyer, H., E. Mizukoshi, et al. (2001). "Cross-reactivity between hepatitis C virus and Influenza A virus determinant-specific cytotoxic T cells." J Virol **75**(23): 11392-400.
- Welsh, R. M. and R. S. Fujinami (2007). "Pathogenic epitopes, heterologous immunity and vaccine design." Nat Rev Microbiol **5**(7): 555-63.
- Welsh, R. M. and L. K. Selin (2002). "No one is naive: the significance of heterologous T-cell immunity." Nat Rev Immunol **2**(6): 417-26.
- Zhang, Y. and D. B. Williams (2006). "Assembly of MHC class I molecules within the endoplasmic reticulum." Immunol Res **35**(1-2): 151-62.