



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) **PI 1003646-6 A2**

(22) Data de Depósito: 08/09/2010
(43) Data da Publicação: 08/01/2013
(RPI 2192)



(51) *Int.Cl.:*
C40B 30/02
C40B 30/04
A01N 63/00

(54) Título: IDENTIFICAÇÃO DE ALVOS TERAPÊUTICOS PARA DESENHO COMPUTACIONAL DE DROGAS CONTRA BACTÉRIAS DOTADAS DA PROTEÍNA PILT

(73) Titular(es): Embrapa - Empresa Brasileira de Pesquisa Agropecuária.

(72) Inventor(es): Goran Neshich, Ivan Mazoni, Izabella Agostinho Pena Neshich, José Gilberto Jardine, José Salim, Leticia Nishimura

(57) Resumo: IDENTIFICAÇÃO DE ALVOS TERAPÊUTICOS PARA DESENHO COMPUTACIONAL DE DROGAS CONTRA BACTÉRIAS DOTADAS DA PROTEÍNA PILT. A presente invenção se refere a um método para identificar regiões alvo existentes na interface de monômeros constituintes da proteína PilT com o objetivo de desenhar moléculas potencialmente aplicáveis no comprometimento da atividade desta proteína, controlando, assim, processos infecciosos. O método é caracterizado pela (i) seleção de pelo menos uma sequência de aminoácidos constituidora de monômero de PilT; (ii) desenvolvimento de um modelo computacional tridimensional da estrutura homo-hexamérica de PilT; (iii) análise computacional para determinar os resíduos de aminoácidos formadores de interface (IFR) e suas características físico-químicas e estruturais para todas as cadeias dos modelos de complexos hexaméricos gerados; (iv) seleção das regiões a serem usadas como alvos terapêuticos (e alvos terapêuticos preferenciais) na interface entre os monômeros baseada na intensidade de determinados parâmetros; (v) modelagem computacional para o desenho de moléculas potencialmente capazes de efetuar ligações e/ou interações entre regiões alvo dos monômeros. Notadamente, algumas das principais aplicações desta tecnologia consistem no combate à bactéria *Xylella fastidiosa*, agente etiológico da Clorose Variegada dos Citrus (CVC ou Amarelinho), e a doença de Pierce.

Relatório Descritivo de Patente de Invenção: “IDENTIFICAÇÃO DE ALVOS TERAPÊUTICOS PARA DESENHO COMPUTACIONAL DE DROGAS CONTRA BACTÉRIAS DOTADAS DA PROTEÍNA PILT”.

CAMPO DA INVENÇÃO

5 O presente documento de patente refere-se a um método para identificar regiões alvo existentes na interface de monômeros constituintes da proteína PilT com o objetivo de desenhar moléculas potencialmente aplicáveis no comprometimento da atividade desta proteína, controlando, assim, processos infecciosos. Notadamente, uma das principais aplicações desta tecnologia consiste no combate à bactéria *Xylella fastidiosa*,
10 agente etiológico da Clorose Variegada dos Citrus (CVC ou Amarelinho) e a doença de Pierce.

ESTADO DA TÉCNICA

A *Xylella fastidiosa* é uma bactéria gram-negativa e não-flagelada que provoca várias doenças em plantas, tais como a Clorose Variegada dos Citrus (CVC),
15 popularmente conhecida como “amarelinho” e a doença de Pierce, que afeta a citricultura e vinicultura, respectivamente. O microrganismo *Xylella fastidiosa* é limitado a persistir apenas colonizando os vasos do xilema, os vasos condutores de água e sais em plantas, e no trato digestório anterior de alguns insetos, como os popularmente conhecidos como cigarrinhas (Hemiptera: Cicadellidae), que se alimentam da seiva das
20 plantas e servem como vetores para carregamento e inserção da bactéria nas plantas (Hopkins, D. L., e A. H. Purcell. 2002. *Xylella fastidiosa*: cause of Pierce's disease of grapevine and other emergent diseases. Plant Dis. 86:1056-1066). Há muitos anos tem-se testado o uso de inseticida contra os insetos-vetores desta bactéria, porém não houve suficiente eficácia, e os prejuízos somados para o ano de 2000, decorrentes da doença de
25 Pierce, no estado da Califórnia, Estados Unidos da América, atingiram os 14 bilhões de dólares (ver documento US 2005/0257285). Os mecanismos de virulência da *Xylella fastidiosa* e a forma como esta interage com plantas hospedeiras não são totalmente compreendidos. No entanto, a explicação mais plausível é a formação de agregados como biofilmes que, ao colonizar vasos do xilema, causa um bloqueio ao fluxo de seiva
30 e culmina nos sintomas da doença (Hopkins, D. L. 1989. Xylem-limited bacterial pathogen of plants. Annu. Rev. Phytopathol. 27:271-290; Newman, K. L., R. P. P. Almeida, A. H. Purcell, and S. E. Lindow. 2003. Use of a green fluorescent strain for

analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl. Environ. Microbiol. 69:7319-7327).

- Hopkins e colaboradores sugeriram que a colonização e a patogenicidade da bactéria *Xylella fastidiosa* que levam à doença de Pierce estão estritamente relacionadas com a sua capacidade de se movimentar dentro dos elementos de vaso do xilema, o que permite a colonização de outras regiões da planta. Além disso, é sabido que as cepas não virulentas ou atenuadas raramente movem-se do ponto de inoculação (Hopkins, D. L. 1989. Xylem-limited bacterial pathogen of plants. Annu. Rev. Phytopathol. 27:271-290).
- O projeto genoma da *Xylella fastidiosa* revelou a presença de genes que codificam proteínas envolvidas na biogênese e função do type IV pili (T4P) (Simpson, A. J. G., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. C. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, E. D. Bonaccorsi, S. Bordin, J. M. Bove, M. R. S. Briones, M. R. P. Bueno, A. A. Camargo, L. E. A. Camargo, D. M. Carraro, H. Carrer, N. B. Colauto, C. Colombo, F. F. Costa, M. C. R. Costa, C. M. Costa-Neto, L. L. Coutinho, M. Cristofani, E. Dias-Neto, C. Docena, H. El-Dorry, A. P. Facincani, A. J. S. Ferreira, V. C. A. Ferreira, J. A. Ferro, J. S. Fraga, S. C. Franca, M. C. Franco, M. Frohme, L. R. Furlan, M. Garnier, G. H. Goldman, M. H. S. Goldman, S. L. Gomes, A. Gruber, P. L. Ho, J. D. Hoheisel, M. L. Junqueira, E. L. Kemper, J. P. Kitajima, J. E. Krieger, E. E. Kuramae, F. Laigret, M. R. Lambais, L. C. Leite, E. G. M. Lemos, M. V. F. Lemos, S. A. Lopes, C. R. Lopes, J. A. Machado, M. A. Machado, A. Madeira, H. M. F. Madeira, C. L. Marino, M. V. Marques, E. A. L. Martins, E. M. F. Martins, A. Y. Matsukuma, C. F. M. Menck, E. C. Miracca, C. Y. Miyaki, C. B. Monteiro-Vitorello, D. H. Moon, M. A. Nagai, A. Nascimento, L. E. S. Netto, A. Nhani, F. G. Nobrega, L. R. Nunes, M. A. Oliveira, M. C. de Oliveira, R. C. de Oliveira, D. A. Palmieri, A. Paris, B. R. Peixoto, G. A. G. Pereira, H. A. Pereira, J. B. Pesquero, R. B. Quaggio, P. G. Roberto, V. Rodrigues, A. J. D. Rosa, V. E. de Rosa, R. G. de Sa, R. V. Santelli, H. E. Sawasaki, A. C. R. da Silva, A. M. da Silva, F. R. da Silva, W. A. Silva, J. F. da Silveira, 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-157), que podem gerar mobilidade do tipo "Twitching motility" (Meng, Y., Li, Y., Galvani, C. D., Hao, G., Turner, J. N., Burr, T. J. & Hoch, H. C. 2005. Upstream migration of *Xylella fastidiosa* via pilus-driven

- twitching motility. J Bacteriol 187, 5560–5567). O “Twitching motility” é uma forma de movimento associado à superfície pelo qual as bactérias puxam-se rapidamente ao longo das superfícies através de ciclos de polimerização e despolimerização do Pilus (T4P) (Skerker, J. M. & Berg, H. C. 2001. Direct observation of extension and retraction of type IV pili. Proc Natl Acad Sci USA 98, 6901–6904). A energia necessária para o movimento é fornecida por meio de hidrólise do ATP por proteínas chamadas PilB e PilT para montagem e desmontagem, respectivamente, do Pilus. A perda de função da proteína PilT ou da PilB resulta na ausência deste tipo de motilidade, associada à privação da extensão ou retração do pilus (Whitchurch, C. B., Hobbs, M., Livingston, S. P., Krishnapillai, V. & Mattick, J. S. 1991. Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. Gene 101, 33–44).
- Satyshur *et al.* 2007 (K. A. Satyshur, G. A. Worzalla, L. S. Meyer, E. K. Heiniger, K. G. Aukema, A. M. Masic, and K. T. Fore. Crystal Structures of the Pilus Retraction Motor PilT Suggest Large Domain Movements and Subunit Cooperation Drive Motility. Structure 15, 363–376. 03/2007) realizaram a caracterização estrutural do motor de retração do pilus, PilT. Estes autores resolveram, por difração de raios X, quatro estruturas tridimensionais de proteínas PilT (códigos PDB: 2GSZ, 2EWV, 2EWW e 2EYU) de um organismo hipertermófilo, *Aquifex aeolicus*. PilT é uma ATPase hexamérica de um subgrupo do “bacterial type II/type IV secretion systems”, e tem dois grandes domínios estruturais: o domínio N-terminal (NTD) e o domínio C-Terminal (CTD), que contém o núcleo ATPásico. Adicionalmente, foi demonstrada a notável importância de interações polares e carregadas nas interfaces CTD_n: NTD_{n+1} (1157Å², dos 1782 Å² da área total da interface é fornecida pelos resíduos polares e carregados) para a função da PilT. Através de mutagênese sítio-dirigida na PilT de *Pseudomonas aeruginosa*, demonstraram que alguns resíduos são cruciais para a função da proteína, sendo que metade deles estão localizados na região da interface (D29, R95 e R207) e ressaltaram a importância da mesma para o correto funcionamento da proteína. O trabalho ainda menciona que esta elucidação é de suma importância para o desenvolvimento de ferramentas que possam prevenir a disseminação de doenças infecciosas em animais e plantas causadas pelos patógenos: *Ralstonia solanacearum*, agente causal do moko ou murcha bacteriana em plantas solanáceas (Kang, Y., Liu, H.,

- Genin, S., Schell, M.A., and Denny, T.P. 2002). *Ralstonia solanacearum* requires type 4 pili to adhere to multiple surfaces and for natural transformation and virulence. Mol. Microbiol. 46, 427–437.), *Pseudomonas syringae* (existem mais de 50 patovares que atuam como agentes patogênicos em diversas plantas) (Hirano, S.S. e C.D. Upper.
- 5 Bacteria in the Leaf Ecosystem with Emphasis on *Pseudomonas syringae* — a Pathogen, Ice Nucleus, and Epiphyte. Microbiology and Molecular Biology Reviews 64 624-653. 2000), *Pseudomonas aeruginosa* (patógeno oportunista em seres humanos e também causa infecções em plantas como manjerição (*Ocimum basilicum*) e alface (*Lactuca sativa*) (Rahme, L. G., Stevens, E. J., Wolfort, S.F., Shao, J., Tompkins, R. G.,
- 10 Calderwood, S.B., Ausube,l F.M. 1995. Common virulence factors for bacterial pathogenicity in plants and animals. Science 268: 1899-1902), *Neisseria meningitidis* e *Neisseria gonorrhoeae* (Importantes patógenos obrigatórios de mucosas humanas, sendo o primeiro agente causal da meningite e o segundo da gonorréia) (Merz, A.J., e So, M. Interactions of pathogenic Neisseriae with epithelial cell membranes. Annu.
- 15 Rev. Cell Dev. Biol. 16, 423–457. 2000.; Pujol, C., Eugene, E., Marceau, M., e Nassif, X. 1999. The meningococcal PilT protein is required for induction of intimate attachment to epithelial cells following pilus-mediated adhesion. Proc. Natl. Acad. Sci. USA 96, 4017–4022). Outros importantes patógenos também se utilizam de motilidade do tipo “twitching” através de pilus tipo IV, como *Dichelobacter nodosus*, causador de
- 20 pododermatite em bovinos e *Vibrio cholerae*, importante patógeno humano, causador da cólera (J. S. Mattick. Type IV pili and twitching motility. Annu Rev Microbiol. 2002; 56:289-314. Epub 2002 Jan 30. Review). Além destes, *Xanthomonas axonopodis* pv *citri*, outra bactéria muito importante para a citricultura por ser a causadora do cancro cítrico, também possui pilus tipo IV (Yang, Y. C., Chou, C. P., Kuo, T. T., Lin , S. H.,
- 25 Yang, M. K. PilR enhances the sensitivity of *Xanthomonas axonopodis* pv. citri to the infection of filamentous bacteriophage Cf. Curr Microbiol. 2004 Apr;48(4):251-61).

Recentemente outro trabalho, do mesmo grupo que publicou a estrutura da PilT de *A. Aeolicus*, resolveu outras estruturas tridimensionais da proteína PilT, porém de *Pseudomonas aeruginosa*, uma bactéria mais próxima filogeneticamente da *X.*

30 *fastidiosa* em relação ao *A. aeolicus*. As novas estruturas são de complexos parciais, um ligado a um análogo do ATP, AMP-PCP (código PDB: 3jvv) e um não-ligado (código PDB: 3jvu), e sugeriram mecanismos de funcionamento da proteína-motor que, é

considerada o motor-biológico mais forte conhecido na natureza (A. M. Misić, K. A. Satyshur, K. T., Forest. *P. aeruginosa* PilT structures with and without nucleotide reveal a dynamic type IV pilus retraction motor. J Mol Biol. 2010 Jul 30;400(5):1011-21. Epub 2010 Jun). Estas estruturas mostram complexos que seriam os mais próximos

5 estruturalmente do que se hipotetizaria serem os complexos da PilT de *X. fastidiosa* e ressalta que esta é uma proteína extremamente dinâmica e possui certas diferenças estruturais quando ligadas a ADP, ATP ou sem ligantes.

Atualmente, avanços importantes na produção de fármacos foram obtidos por meio de abordagens computacionais. Os altos prejuízos originados com as

10 fitopatologias que a *Xylella fastidiosa* e a ausência de soluções eficazes contra este mal nos estimulam a buscar novas alternativas e alvos terapêuticos distintos dos trabalhados atualmente para o combate a este microrganismo. Portanto, este trabalho descreve uma tentativa de inovar em termos de realizar uma caracterização de novos alvos para

desenho computacional de fármacos baseado na estrutura ("Structure-Based drug design"), em uma proteína ainda não usada para este fim, a PilT.

15

Anderson (A.C. Anderson. The process of structure-based drug design. Chem Biol. 2003 Sep;10(9):787-97. 2003) publicou uma descrição de como deve ser realizado o procedimento de desenho computacional de fármacos baseado em estruturas protéicas, desde os critérios que devem ser adotados para a escolha do alvo terapêutico para tal

20 desenho até os processos de desenvolvimento, "docking" e "virtual screening". O trabalho discute que o desenho de drogas antimicrobianas deve se basear em alvos que sejam essenciais, sejam principalmente encontrados nos patógenos (em relação a organismos não-patogênicos), tenha uma função única no patógeno e que sejam passíveis de sofrerem inibição por pequenas moléculas.

25 A escolha da PilT e, mais especificamente, dos resíduos a serem usados como alvos, seguiu estas recomendações. É uma proteína essencial à patogenicidade de *Xylella fastidiosa*, é sabido que a movimentação *via* pilus tipo IV é a principal responsável pela dispersão do patógeno desde o ponto de inoculação e desenvolvimento da doença, é importante para outros patógenos em termos de movimentação, formação

30 de agregados, adesão e evasão de sistema imunológico hospedeiro (J. S. Mattick. Type IV pili and twitching motility. Annu Rev Microbiol. 2002;56:289-314. Epub 2002 Jan 30. Review) e os resíduos descritos a seguir, escolhidos como os alvos preferenciais,

têm ocorrência restrita às PilTs de organismos patogênicos estudados em relação a organismos de vida livre como cianobactérias. Como já foi mencionado, é sabido que ao mutar alguns aminoácidos localizados na interface, a PilT deixou de exercer sua função e o organismo mutado não foi capaz de mover-se (“Twitching motility”) (K. A. Satyshur, G. A. Worzalla, L. S. Meyer, E. K. Heiniger, K. G. Aukema, A. M. Mistic, and K. T. Fore. Crystal Structures of the Pilus Retraction Motor PilT Suggest Large Domain Movements and Subunit Cooperation Drive Motility. *Structure* 15, 363–376. 03/2007), portanto partimos do pressuposto de que as interfaces são essenciais à manutenção da função e motilidade, e que a criação de um composto que seja capaz de se ligar a esta região impedirá este correto funcionamento e/ou impedirá a correta associação dos monômeros no momento da formação do hexâmero.

Existem vários trabalhos reconhecidos cientificamente que usaram técnicas similares às usadas neste trabalho, com o intuito de compreender melhor as estruturas de proteínas e/ou objetivando a identificação de alvos terapêuticos para desenvolvimento de fármacos. A seguir expõe-se uma série de citações de alguns destes trabalhos.

Li *et al.* 1996 (R. Li, X. Chen.; B. Gong, P. M. Seizer, Z. Li, E. Davidson, G. Kurzban, R. E. Miller, E. O. Nuzum, J. H. McKerrow, R. J. Fletterick, S. A. Gillmor C. S. Craik, I. D. Kuntz, F. E. Cohen e G. L. Kenyon. Structure-Based Design of Parasitic Protease Inhibitors. *Bioorganic & Medicinal Chemistry* 1996, 4, No. 9, 1421-1427) exploraram a utilização de estruturas protéicas similares conhecidas a fim de se desenvolver possíveis candidatos a drogas a serem utilizados, não apenas contra parasitas, mas também contra outras doenças infecciosas e processos de multiplicação celular descontrolada. Os autores defendem ainda que o design de drogas baseados em estruturas identifica interações favoráveis e não favoráveis entre um potencial inibidor e um determinado alvo.

Sheng *et al.* 2004 (Sheng, C., Zhang, W., Zhang, M., Song, Y., Ji, H., Zhu, J., Yao, J., Yu, J., Yang, S., Zhou, Y., Zhu, J., Lu, J. Homology Modeling of Lanosterol 14 α -Demethylase of *Candida albicans* and *Aspergillus fumigatus* and Insights into the Enzyme-Substrate Interactions. *J Biomol Struct Dyn.* 2004; Aug;22(1):91-99) apresentaram a utilização da estrutura de uma proteína de *Mycobacterium tuberculosis* (14 α -sterol demethylase) como template para a modelagem da estrutura

tridimensional da proteína lanosterol 14 α -demethylase de outros microrganismos, no caso: *Candida albicans* e *Aspergillus fumigatus*.

Marrone *et al.* 1997 (T. J. Marrone, J. M. Briggs, and J. A. McCammon. STRUCTURE-BASED DRUG DESIGN: Computational Advances. Annu. Rev. Pharmacol. Toxicol. 1997 37:71–90) apresentaram uma revisão com os métodos computacionais baseados em estruturas mostrando a utilização destes para a descoberta, refinamento ou alteração de agentes terapêuticos. A revisão cita vários campos como a visualização molecular, modelagem molecular, técnicas de modelagem utilizando bancos de dados 3-D, métodos de fragmentação e perturbação da energia livre.

No documento WO200135316 foi descrito um método computacional para a seleção de drogas baseado em polimorfismo, compreendendo a obtenção de mais de uma sequência de aminoácido das proteínas alvos que são produto de um gene exibindo polimorfismo, onde as sequências representem diferentes polimorfismos genéticos e a geração de variações de modelos estruturais protéicos em 3D; ensaios computacionais de “Docking drug” moleculares com os modelos protéicos alvos; refinamento energético dos complexos de ancoragem (docked complexes); determinação da interação de ligação entre a droga ou potenciais 15 novas moléculas candidatas a droga e os seus respectivos modelos; seleção de terapias de drogas baseada na droga ou drogas com maior potencial de ligação/interação com os modelos estruturais variantes.

O documento WO9425860 descreveu um sistema computacional para modelagem da estrutura tridimensional de uma proteína modelo sendo esta modelagem baseada na estrutura tridimensional de uma proteína template, na sequência de aminoácidos da proteína modelo, nas proteínas contendo a pluralidade dos aminoácidos, em cada aminoácido possuindo uma cadeia principal de átomos e cadeias laterais de átomos, em cada átomo em uma estrutura tridimensional contendo um determinado posicionamento.

No documento US5884230 foi descrito um sistema computacional para a geração de uma coleção de informações relativas à posição entre pares de aminoácidos para uso na modelagem de uma estrutura tridimensional de uma região variável de uma proteína modelo. O sistema computacional possui informações sobre as posições relativas entre pares de aminoácidos de regiões variáveis de uma coleção de protéica.

No documento WO2006110064 foi descrito um método para a seleção de potenciais compostos medicinais e design de drogas, que compreende a predição de valores da afinidade de ligação ou da energia da interação liberada da interação proteína-ligante a partir de um score calculado.

5 No documento US2001000807 foi descrito um método computacional para a identificação de alvos de ligação em proteínas e outras macromoléculas. A invenção inclui um algoritmo para predição de alvos de ligação em proteínas. Apesar de requerer o conhecimento da estrutura da proteína, o algoritmo não necessita da localização, da natureza dos sítios de ligação e nem dos ligantes. Os alvos de ligação na proteína são
10 identificados e classificados de acordo com a afinidade ótima esperada. Está invenção possui uma significativa aplicação no desenho de drogas baseados na estrutura.

No documento US5580723 foi descrito um método para a identificação de, pelo menos, um domínio ativo desconhecido em uma região de sequência conhecida de aminoácidos de um polipeptídeo correlacionado, contendo similaridade, de ocorrência
15 natural. O polipeptídeo contendo a relação de parentesco deve ter sua atividade biológica predefinida. E o domínio ativo em identificação deve ser capaz de interagir com o mesmo alvo do polipeptídeo correlacionado, em sua forma ativa, cuja interação é responsável por sua atividade biológica.

No documento US2005123995 foi descrito um método de um código de
20 simulação computacional para avaliar a liberação de energia de ligações entre resíduos de aminoácidos de polipeptídios. Ao fazer uso das propriedades de amostragem canônica, base para a tecnologia, a afinidade para a ligação de fragmentos nas proximidades de cada resíduo de proteína pode ser eficientemente calculada. O volume de ligação associado a cada par de fragmentos-resíduo é estimado com base em critérios
25 de proximidade simples e um mapeamento de afinidade da superfície protéica pode ser obtido desta forma. A análise desses dados para vários tipos de fragmento fornece informações valiosas para auxiliar a identificação de sítios de ligação da proteína, bem como para identificar fragmentos chave que podem ser utilizados para a construção de potenciais drogas.

30 No documento WO0135316 foi descrito um método baseado em computação para design de drogas baseado em polimorfismo genético. Tal método compreende: obtenção de mais de uma sequência de aminoácido da proteína alvo, geração de

modelos tridimensionais da estrutura protéica e suas variantes, desenho de candidatos a drogas, modificação de drogas existentes baseada na predição da interação molecular entre os modelos e as drogas candidatas, com suas respectivas variações.

No documento US2005257285 foi descrito uma proteína quimérica antimicrobiana para o tratamento terapêutico e profilático de doenças de plantas causadas pela *Xylella fastidiosa*. A proteína antimicrobiana é composta por um domínio polipeptídico de reconhecimento de superfície, capaz de se ligar a componentes da membrana bacteriana, fundido a um domínio polipeptídico de lise bacteriana, capaz de afetar a lise ou a ruptura da membrana bacteriana. Particularmente, o método e composição para o tratamento da Doença de Pierce são fornecidos. O método para a geração da planta *Vitis vinefera* transgênica, expressando a proteína quimérica no xilema, também é apresentado.

No documento US2005053584 foi descrito a introdução de uma cepa com menor virulência, definida como “benigna”, de *X. fastidiosa* nas plantas de interesse que normalmente são suscetíveis à doença infecciosa causada por este patógeno. As plantas inoculadas são, geralmente, resistentes à Doença de Pierce, pelo menos, cerca de 3-4 anos. A invenção apresenta ainda a cepa bacteriana “benigna” adequada para o controle da Doença além de prover injeções de reforços (Booster) para cada 3-4 anos.

No documento WO2006069160 foi descrito uma construção compreendendo uma molécula de ácido nucléico que codifica para uma hemaglutinina tipo HecA recombinante, ou um fragmento da mesma, o qual confere resistência à infecção pela *Xylella fastidiosa* quando expresso em plantas.

No documento MXNL06000034 é descrito uma substância Anti-*Xylella fastidiosa* desenvolvida através da junção de um csFy aviário e um peptídeo lítico Cecropin P1 (CP1) que é expresso em um bacteriófago M13. Este bacteriófago é capaz de remover 100% da bactéria em cultivos in-vitro em uma concentração de $3,6 \times 10^5$ fagos/célula após 28 horas de incubação.

No documento US6548265 foi descrito um método para tratamento e/ou profilaxia de doenças infecciosas causadas por bactérias que se aderem aos tecidos através do pilus. O método se baseia na interação com chaperonas moleculares periplasmáticas interferindo assim na montagem do pilus prevenindo ou diminuindo a infecção bacteriana. Porém, este método não inclui o pilus tipo IV (usado para o

twitching motility dos patógenos como *Xylella fastidiosa*), já que este sistema não requer chaperonas para correta montagem e funcionamento.

No documento US2003/0176662 foi descrito um procedimento de desenho de drogas na base estrutural, a partir de uma estrutura proteica de hepatitis C. Os autores
5 não chegaram a apresentar nenhuma estrutura de fármaco, porém sugeriram os alvos e, com base nesta identificação, sugeriram o desenho “de novo” dos fármacos que obedecem aos determinantes de farmacóforo encontrado no alvo centrado em determinado aminoácido da estrutura proteica revelada.

No documento US2005/0004766 foi descrito um procedimento de uso de
10 estruturas obtidas através de modelagem por homologia, usando as seqüências protéicas polimórficas, para desenho de drogas afinadas ao paciente específico. Nesta patente, verifica-se que os autores não relatam nenhum fármaco para as estruturas modeladas, porém descrevem em detalhes como obter os modelos estruturais e como se deve proceder a identificação dos alvos para o desenho racional de drogas.

15 O presente pedido de patente refere-se a uma nova alternativa aos atuais processos e pesquisas que objetivam erradicação/inativação das cepas fitopatogênicas de *Xylella fastidiosa*. Existem vários casos de sucesso de desenho computacional de fármacos, dirigidos a alvos terapêuticos estruturalmente delimitados, como pode ser visto no artigo de Villoutreix e colaboradores (Villoutreix, B.O., Eudes, R., Miteva,
20 M.A. Structure-based virtual screening: recent success stories. Comb Chem High Throughput Screen. 2009 Dec; 12(10):1000-16). A presente descrição de novos alvos terapêuticos e correspondentes fármacos, realizada para regiões de interface da proteína hexamérica PilT é distinta das técnicas já apresentadas por se tratar de um novo alvo terapêutico até então não explorado com este fim, a proteína PilT, em uma região
25 específica e pouco estudada na ciência do desenho computacional de fármacos (aminoácidos de interface) e que objetivará o impedimento da formação do complexo hexamérico e/ou sua desestabilização com concomitante inibição da motilidade da bactéria.

SUMÁRIO DA INVENÇÃO

30 A invenção refere-se a identificações de regiões específicas na proteína PilT para que a mesma possa ser inativada ou ter sua atividade reduzida, reduzindo ou evitando assim o desenvolvimento da infecção bacteriana. A proteína em questão está

relacionada à motilidade do patógeno e é descrita na literatura como essencial para o desenvolvimento das fitopatogenias correlacionadas ao patógeno supracitado. As principais fitopatogenias correlacionadas a esta bactéria são a Clorose Variegada dos Citrus (CVC ou Amarelinho) e a doença de Pierce e podem acometer as cultivares de citrus de uma forma geral e de videiras respectivamente, além de outras várias doenças que afetam outras plantas de interesse econômico. A importância de combate a estas doenças é ressaltada por sua rápida capacidade de disseminação zoocórica, tendo como principal vetor a cigarrinha (Hemiptera: Cicadellidae), e do seu alto impacto negativo sobre os cultivares.

Foi desenvolvido um método de identificação de alvos terapêuticos em resíduos formadores de interface (IFR) entre os polipeptídeos monoméricos que constituem os hexâmeros da proteína PilT em três possíveis conformações (ligado a ATP, ADP e sem ligantes), através de modelagem de três complexos hexaméricos de moldes em diferentes conformações. Os moldes usados para tal modelagem foram a estrutura da PilTs de *A. aeolicus* (aqui chamada AaPilT) ligada a ADP e duas estruturas de *P. aeruginosa* (PaPilT) ligadas, respectivamente, a ATP e nenhum ligante e depositadas no PDB ("Protein Data Bank"). Posteriormente, realizou-se a predição de resíduos localizados na interface, suas características, e análises de Biologia Computacional para indicação de locais-alvo para o desenho de fármacos baseado nas estruturas geradas da PilT de *Xylella fastidiosa* (aqui chamada XfPilT). Em seguida, foi proposta, em detalhes, uma descrição de como se procederá o desenho computacional de fármacos que se ligarão nos sítios-alvo preditos. Como complemento a esta descrição, foi exposto um exemplo prático com fármacos que desenhemos na base de determinantes de ligação ao alvo identificado, no caso centrado na aminoácido Glu89, possuindo um respaldo teórico muito forte para poder projetar um ótimo "lead", uma vez que encontramos o alto grau de index de atracamento (é importante ressaltar que o escopo da invenção não se restringe a estes fármacos desenhados).

BREVE DESCRIÇÃO DAS FIGURAS

Figura 1. Representação de alinhamento estrutural dos modelos de hexâmeros criados e os templates usados como base para sua modelagem por homologia. Figura 1a: Sobreposição estrutural do modelo da PilT de *X. fastidiosa* (Xf) nomeado como XfAa1 (em cinza) criado a partir do hexâmero da PilT da bactéria *A. aeolicus* (Aa) (em

sua conformação ligada a ADP) cujo código PDB é 2gsz (em preto). Figura 1b: Sobreposição estrutural do modelo da PilT de Xf nomeado como XfPa1 (em cinza) criado com base no hexâmero gerado a partir do Biological unit do PDB (3jvu) da PilT de *P. Aeruginosa* (Pa) em sua conformação sem ligantes (em preto). Figura 1c: 5 Sobreposição estrutural do modelo da PilT de Xf nomeado como XfPa2 (em cinza) criado com base no hexâmero gerado a partir do Biological unit do PDB (3jvv) da PilT de Pa em sua conformação ligada a um análogo do ATP (o AMP-PCP) (em preto). Os valores de RMSD estão representados.

Figura 2. Valores percentuais de ocupação média de área na interface e 10 superfície livre dos hexâmeros por duas categorias de aminoácidos (hidrofóbicos e polares+carregados). As estruturas representadas nesta figura são: as usadas como molde (da esquerda para a direita: 2gsz (AaPilT em conformação de ligação a ADP), 3jvu (PaPilT, conformação não-ligada), 3jvv (PaPilT, conformação ligação a ATP), XfAa1 (modelo da XfPilT baseado em 2gsz), XfPa1 (modelo da XfPilT baseado em 15 3jvu) e XfPa2 (modelo da XfPilT baseado em 3jvv).

Figura 3. 3a: Densidade de Contatos nas Interfaces (ICD) por complexo: o número total de contatos de todas as interfaces estabelecidas em cada complexo hexamérico dividida pela soma da área da interface de todo o hexâmero para AaPilT 2gsz, PaPilT 3jvu, PaPilT 3jvv, XfAa1, XfPa1 e XfPa2. 3b: Densidade de Energia de 20 Contatos de Interface (ICD) por complexo: soma das energias de contatos de interface estabelecidos em cada complexo dividida pela soma das áreas de interfaces de todo o hexâmero para AaPilT 2gsz, PaPilT 3jvu, PaPilT 3jvv, XfAa1, XfPa1 e XfPa2.

Figura 4: Diagrama de Venn demonstrando a ocorrência dos resíduos sugeridos como alvos (ver tabela 2) nos três diferentes complexos. (Neste caso, é considerada a 25 ocorrência se o resíduo está presente em pelo menos uma cadeia dos modelos hexaméricos de XfPilTs, com pelo menos 10 kcal/mol de energia de contatos de interface, mais que $1 \text{ \AA}^2$ de área exposta ao solvente quando em complexo, que tenham a característica de serem capazes de estabelecer pontes de hidrogênio e/ou contatos de cunho eletrostático e, preferencialmente, que se localizem em bolsos (pockets) em 30 complexo e/ou isolamento. Os resíduos destacados em branco se localizam em bolsos em complexo (ou seja, na conformação hexamérica há algum bolso adjacente) em pelo

menos uma cadeia. O resíduo E89 está sublinhado pois será posteriormente usado como exemplo de representação estrutural de um alvo.

Figura 5. Alinhamento da seqüência primária, realizado por meio do *software* ClustalW (Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J., D., Gibson, T. J., Higgins, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948), de PilTs de bactérias patogênicas de Proteobacteria que se utilizam de movimentação guiada por pilus tipo IV (ou como no caso de *Xanthomonas axonopodis*, possui o pilus tipo IV) como citado em “Estado da Técnica” (*Xylella fastidiosa* (XfPilT_Pt), *Xanthomonas axonopodis* pv *citri* (XaPilT_Pt), *Pseudomonas syringae* pv *tabaci* (PsPilT_Pt), *Pseudomonas aeruginosa* (PaPilT_Pt), *Ralstonia solanacearum* (RsPilT_Pt), *Vibrio cholerae* (VcPilT_Pt), *Dichelobacter nodosus* (DnPilT_Pt), *Neisseria gonorrhoeae* (NgPilT_Pt) e *Neisseria meningitidis* (NmPilT_Pt), bactérias de vida livre essenciais à manutenção do equilíbrio de inúmeros ecossistemas como Cianobactérias (*Nostoc* sp. NsPilT_Cy e *Synechocystis* sp. (SsPilT_Cy) e a bactéria do filo Aquificales, *Aquifex aeolicus* (AaPilT). Os identificadores GI das seqüências estão descritos no parágrafo anterior. Códigos das PilTs de Proteobactérias patogênicas são destacados em preto, Cianobactérias em cinza *A. aeolicus* marcado em branco. Os resíduos mais importantes na interface, sugeridos como alvos (que sejam presentes em pelo menos uma cadeia dos modelos hexaméricos de XfPilTs, com pelo menos 10 kcal/mol de energia de contatos de interface, mais que 1Å² de área exposta ao solvente quando em complexo, que tenham a característica de serem capazes de estabelecer pontes de hidrogênio e/ou contatos de cunho eletrostático e presentes ou não em pockets) são destacados da seguinte maneira: os resíduos destacados com barras longas em cinza são possíveis alvos que ocorrem em todas as PilTs de todos organismos usadas no alinhamento de sequência primária; os resíduos destacados com barras em cinza escuro com comprimento igual ao da barra de identificadores de proteobactérias patogênicas, são os possíveis alvos específicos desta categoria de microrganismos e que podem ser interpretados como alvos preferenciais ao desenho de fármacos para combate a estes patógenos em comum; A barra pequena em cinza escuro representa os possíveis alvos que ocorrem unicamente em PilT da *Xylella fastidiosa*, sendo estes interpretados como alvos únicos à Xf; A barra pequena

transparente mostra alvos que variam em termos de aminoácido (ex: D33 para E33) mas o aminoácido trocado tem propriedades similares, considerado como um “Positivo” por indicar substituições nas quais a matriz BLOSUM-62 matrix pontua positivamente segundo Altschul e colaboradores (S. F. Altschul, T. L. Madden, A. A. Schäffer, Z. Zhang, W. Miller, D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res.* 1997 Sep 1;25(17):3389-402.); Os círculos pretos representam os resíduos que têm distribuição variada entre os organismos, que são os resíduos cujos campos estão em branco na tabela 2.

10 **Figura 6:** Representação estrutural de um exemplo de alvo terapêutico, o resíduo E89 na cadeia A do complexo XfAa1, mostrando os resíduos adjacentes que formam o nanoambiente no qual o fármaco a ser desenhado se ligará. Este alvo foi escolhido para representar visualmente, pois é de extrema importância por existir unicamente na XfPilT e existir em todos os três modelos (XfAa1, XfPa1 e XfPa2) 15 contendo “pocket adjacente”, sendo essa característica muito importante ao desenho de fármaco.

Figura 7: Representação em 2D (obtida através de uma montagem de imagens extraídas do software Ligand Scout) dos fármacos desenhados e modelados que obtiveram os melhores valores de MolDock Score em cada uma das rodadas de docking 20 realizadas. Foram realizadas cinco rodadas de docking, sendo a primeira realizada com compostos desenhados diretamente das características do sítio alvo. As rodadas seguintes foram realizadas com moléculas que desenhadas com base nos melhores compostos das rodadas anteriores, sendo o E89C4m4 obtido como melhor candidato a ligação na primeira rodada de docking; o E89Cm29 obtido na segunda rodada, 25 E89C4m30 na terceira, E89C4m32 na quarta e E89C4m34 obtido na quinta rodada e classificado como o melhor candidato a se ligar no sítio-alvo E89. Estão representados com setas pontilhadas as pontes de hidrogênio que o composto estabelece com resíduos do sítio-alvo; três triângulos cinza representam os átomos ionizáveis e, adjacentes a essa representação, estão as identificações de resíduos dos sítios-alvo que interagem 30 eletrostaticamente com o átomo em questão; a ocorrência de interações hidrofóbicas é representada com duas linhas curvas em cinza claro. Há mais detalhes na porção da figura que possui a imagem 2D do composto E89C4m34 que pode ser sugerido (dentre

os compostos desenhados) como o que mais possivelmente se ligará ao sítio-alvo. Está representada uma distância (12,22 Å) do átomo 18 ao 8, que mostra a correspondência com o tamanho da cavidade (que tem aproximadamente 14 Å de comprimento em sua porção mais inferior), os átomos estão numerados (segundo a numeração do ChemDraw) e estão presentes as formas isoméricas do composto.

Figura 8: Representação das superfícies que delimitam o “pocket” e que compõem o sítio-alvo de ligação está mostrada (8a), mostrando como o ligante possui formato similar ao da região. Em 8b está representado o E89C4m34 e as posições em 3D dos resíduos que interagem com esta molécula no formato “stick”.

10

DESCRIÇÃO DETALHADA DA INVENÇÃO

A presente invenção refere-se a um método para identificar regiões passíveis de serem usadas como alvos terapêuticos para o desenho computacional de fármacos baseado em estrutura existentes na interface de monômeros constituintes da proteína PilT, bem como a listagem e caracterização dos alvos preditos, com o objetivo futuro de desenhar moléculas que se liguem nestas regiões e que sejam potencialmente aplicáveis no comprometimento da atividade desta proteína, controlando, assim, processos infecciosos.

O método proposto consiste das seguintes etapas:

1 – Caso o complexo da proteína de interesse, por qualquer razão, não esteja descrita ou tenha sua estrutura indeterminada, deve se proceder a geração de um modelo a partir de outra proteína com alta similaridade em sua sequência primária utilizando-se a modelagem por homologia através do programa Swiss-Model (Guex, N. e Peitsch, M. C. 1997. SWISS-MODEL and the Swiss-Pdb-Viewer: An environment for comparative protein modelling. Electrophoresis 18:2714-2723) ou o programa Modeller (Eswar, N., Marti-Renom, M. A. , Webb, B. , Madhusudhan, M. S., Eramian, D., Shen, M., Pieper, U., Sali., A. 2006. Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30). A validação da modelagem é em seguida realizada através da análise dos gráficos de Ramachandran (Ramachandran, G. N., Ramakrishnan, C., Sasisekharan, V. 1963. Stereochemistry of polypeptide chain conformations. J. Mol Biol 7:95-99) e a utilização do ProSA web-service (Whiederstein, M., Sippl, M. J. 2007. ProSA-web: interactive

web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acid Research, Web Server issue: W407-10).

2 – Utilização do programa Deep-View para a geração do arquivo PDB do complexo da proteína de interesse, por meio de sobreposição estrutural com o complexo
5 usado como template pela modelagem e realização de minimização de energia usando o programa Gromacs (Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., Berendsen, H. J. 2005. "GROMACS: fast, flexible, and free". J Comput Chem 26 (16): 1701–18. doi:10.1002/jcc.20291.);

3 – Utilização do servidor STING (Neshich, G., Togawa, R., Mancini, A. L.,
10 Kuser, P. R., Yamagishi, M. E. B., Pappas Jr., G., Torres, W. V., Campos, T. F., Ferreira, L. L., Luna, F. M., Oliveira, A. G., Miura, R. T., Inoue, M. K., Horita, L. G., de Souza, D. F., Dominiquini, F., Álvaro, A., Lima, C. S., Ogawa, F. O., Gomes, B. G., Palandrani, J. C. F., dos Santos, G. F., de Freitas, E. M., Mattiuz, A. R., Costa, I. C., de Almeida, C. L., Souza, S., Baudet, C. and Higa, R. H. 2003. STING Millennium: a Web
15 based suite of programs for comprehensive and simultaneous analysis of protein structure and sequence. Nucleic Acids Research, 31:13, 3386-3392) para a geração do arquivo TGZ contendo todos os parâmetros físico-químicos e estruturais do STING_DB.

4 – Todos os parâmetros gerados pelo STING são carregados no Star STING
20 (Neshich, G., Mazoni, I., Oliveira, S.R., Yamagishi, M.E., Kuser-Falcão, P.R., Borro, L.C., Morita, D.U., Souza, K.R., Almeida, G.V., Rodrigues, D.N., Jardine, J.G., Togawa, R.C., Mancini, A.L., Higa, R.H., Cruz, S.A., Vieira, F.D., Santos, E.H., Melo, R.C., Santoro, M.M. The Star STING server: a multiplatform environment for protein structure analysis. Genet Mol Res. 2006 Dec 1;5(4):717-22) e no Java Protein Dossier
25 (Neshich, G., Rocchia, W., Mancini, A.L., Yamagishi, M.E., Kuser, P.R., Fileto, R., Baudet, C., Pinto, I.P., Montagner, A.J., Palandrani, J.F., Krauchenco, J.N., Torres, R.C., Souza, S., Togawa, R.C., Higa, R.H. 2004. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. Nucleic Acids Res. 2004 Jul 1;32 (Web Server issue):W595-601), uma plataforma para detalhar
30 e integrar as análises de função estrutural.

5 – Identificação dos resíduos formadores de interface (IRFs) e obtenção dos dados das respectivas áreas através da utilização do programa SurfV (Sridharan, S.,

Nicholls, A. e Honig, B. 1992. A new vertex algorithm to calculate solvent accessible surface areas. *Biophys. J.*, 61, A174) para calcular a área de acessibilidade ao solvente por resíduos em dois cenários: para a cadeia isolada e para a cadeia no complexo com outras cadeias.

5 6 – Estudar a ocorrência de categorias de aminoácidos em Resíduos Formadores de Interface (IFR) e Superfície livre (que é dada pela diferença entre a superfície molecular total e a interface), no que se refere aos aminoácidos que a formam, e também no que se refere às categorias de aminoácidos: Polares: Cys, Ser, Thr, Tyr, Asn, Gln, His and Trp; Carregados: Asp, Glu, Arg, Lys; Hidrofóbicos: Ala, Ile, Leu, Val, Met, Phe and Pro; e Glicina: que pode ser considerada como uma quarta categoria
10 (Branden, C. e Tooze, J. 1991. *Introduction to protein structure*. Garland Publishing, New York, ISBN 0-815-30270-3). Esta fase é importante para indicar o quão polar e carregada é a interface em questão. A análise é feita usando bancos de dados MySQL (<http://www.mysql.com/>) e tabelas contendo valores de áreas obtidos com o programa
15 SurfV.

7 – Utilização do STING (Neshich, G., Togawa, R., Mancini, A. L., Kuser, P. R., Yamagishi, M. E. B., Pappas Jr., G., Torres, W. V., Campos, T. F., Ferreira, L. L., Luna, F. M., Oliveira, A. G., Miura, R. T., Inoue, M. K., Horita, L. G., de Souza, D. F., Dominiquini, F., Álvaro, A., Lima, C. S., Ogawa, F. O., Gomes, B. G., Palandrani, J. C.
20 F., dos Santos, G. F., de Freitas, E. M., Mattiuz, A. R., Costa, I. C., de Almeida, C. L., Souza, S., Baudet, C. e Higa, R. H. 2003. STING Millennium: a Web based suite of programs for comprehensive and simultaneous analysis of protein structure and sequence. *Nucleic Acids Research*, 31:13, 3386-3392) e do Java Protein Dossier (Neshich G., Rocchia W., Mancini A.L., Yamagishi M.E., Kuser P.R., Fileto R., Baudet
25 C., Pinto I.P., Montagner A.J., Palandrani J.F., Krauchenco J.N., Torres R.C., Souza S., Togawa R.C., Higa R.H. 2004. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic Acids Res.* 2004 Jul 1;32 (Web Server issue):W595-601) para calcular o número, o tipo e, portanto, a energia total dos contatos estabelecidos entre a interface dos aminoácidos de cadeia
30 selecionadas e as cadeias adjacentes.

8 – Uma vez possuindo os dados descrevendo a área ocupada por resíduos específicos em uma determinada superfície, realizar a criação de dois índices:

Densidade de Energia nas Interfaces de Contato (ICED – “Interface Contacts Energy Density”), que é dado pela soma das energias para todos os contatos IFRs dividido pela soma da área ocupada na interface por todos os IFRs; e Densidade de Contatos de Interface (ICD – “Interface Contacts Density”) que é dado pelo número total de contatos estabelecidos através de cada contato dividido pela área total. Adicionalmente, o programa Java Protein Dossier (Neshich, G., Rocchia, W., Mancini, A.L., Yamagishi, M.E., Kuser, P.R., Fileto, R., Baudet, C., Pinto, I.P., Montagner, A.J., Palandrani, J.F., Krauchenco, J.N., Torres, R.C., Souza, S., Togawa, R.C., Higa, R.H. 2004. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic Acids Res.* 2004 Jul 1;32 (Web Server issue):W595-601) é utilizado para a análise estrutural geral e o programa PyMol (Delano, W. L. 2002. The PyMOL Molecular Graphics System Delano Scientific, San Carlos, CA, USA. <http://www.pymol.org>) é utilizado para gerar as imagens moleculares.

9 – Realização da seleção dos alvos terapêuticos com base em características físico-químicas (como: valores altos de energia de contatos, polaridade) e estruturais (área exposta ao solvente, presença em “pocket” (bolso), dentre outras) de interesse usando o módulo Select do STING Java Protein Dossier (Neshich, G., Rocchia W., Mancini, A.L., Yamagishi, M.E., Kuser, P.R., Fileto, R., Baudet, C., Pinto, I.P., Montagner, A.J., Palandrani, J.F., Krauchenco, J.N., Torres, R.C., Souza, S., Togawa, R.C., Higa, R.H. 2004. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic Acids Res.* 2004 Jul 1;32 (Web Server issue):W595-601).

10 – Alinhamento entre a estrutura primária de proteínas homólogas à proteína de interesse através do programa ClustalW 2.0 (Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J., D., Gibson, T. J., Higgins, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948) evidenciando as similaridades e diferenças entre estes dois conjuntos de proteínas e buscando as correspondências no alinhamento de estrutura primária dos resíduos encontrados em “9”. Seguindo certos critérios como: presença exclusivamente nas seqüências de bactérias patogênicas, escolher os alvos terapêuticos preferenciais.

11 – Proceder com métodos e protocolos de desenho estrutural de fármacos baseado em estrutura, seja este um método *de novo* ou baseado em estruturas já conhecidas por meio de virtual screening.

EXEMPLOS

5 A invenção será agora descrita em maiores detalhes por meio dos exemplos a seguir, os quais não devem ser interpretados como limitadores do escopo da invenção.

Para entender mais sobre as características físico-químicas das interfaces da PilT de *Xylella fastidiosa*, as quais acredita ser a chave para o correto funcionamento da proteína e como a PilT de *Xylella fastidiosa* possui uma similaridade de seqüência primária elevada em relação a PilT de *A. aeolicus* (cerca de 68% de similaridade e 49% de aminoácidos idênticos) e mais elevada ainda em relação a PilT de *P. aeruginosa* (cerca de 87% de similaridade e 74% de aminoácidos idênticos, como indica o programa BLASTp (S. F. Altschul, T. L. Madden, A. A. Schäffer, Z. Zhang, W. Miller, D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res. 1997 Sep 1;25(17):3389-402), foram utilizadas as estruturas 3D dos hexâmeros PilT de *Aquifex aeolicus* (Aa) (código PDB: 2gsz, em uma conformação ligada a ADP obtido em <ftp://ftp.wwpdb.org/pub/pdb/data/structures/all/pdb/pdb2gsz.ent.gz>) e os complexos hexaméricos de PilTs de *Pseudomonas aeruginosa* (Pa) (códigos PDB: 3jvu e 3jvv, em conformações não-ligada e ligada a ATP, respectivamente, obtidos nos endereços <ftp://ftp.wwpdb.org/pub/pdb/data/biounit/coordinates/all/3jvu.pdb1.gz> e <ftp://ftp.wwpdb.org/pub/pdb/data/biounit/coordinates/all/3jvv.pdb1.gz>) para a modelagem por homologia de cada uma das 6 cadeias distintas da PilT da *X. fastidiosa* (XfPilT) com base nas cadeias correspondentes nos moldes, e nas três conformações, baseadas em 2gsz, 3jvu e 3jvv, utilizando o programa Modeller (Eswar, N., Marti-Renom, M. A. , Webb, B. , Madhusudhan, M. S., Eramian, D., Shen, M., Pieper, U., Sali, A. 2006. Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30), gerando ao todo 18 cadeias, de três complexos.

30 Posteriormente, usando o programa Deep-View (Guex, N. e Peitsch, M. C. 1997. SWISS-MODEL and the Swiss-Pdb-Viewer: An environment for comparative protein modelling. Electrophoresis 18:2714-2723), foi realizada a montagem dos complexos

com base em sobreposição estrutural das cadeias de XfPilT em seus respectivos “templates” (moldes). Desta maneira, foram criados três arquivos PDB para os hexâmeros XfPilT: um baseado em 2gsz, outro em 3jvu e o último em 3jvv. A confecção destes três modelos de complexos é de grande importância, pois visa simular os vários estados da proteína PilT, que é extremamente dinâmica e provavelmente muda de conformação em conformação para o correto funcionamento e quebra do ATP em ADP. A existência destas diferentes conformações nos permite sugerir alvos que estejam expostos em uma conformação e/ou cadeia e interno à interface em outra como alvos interessantes para impedir o correto funcionamento da proteína quando esta já está montada na base do pilus. Portanto, se faz necessário o estudo das três conformações e busca por similaridades e diferenças entre elas.

Após a montagem dos complexos, foi realizada minimização de energia usando o programa Gromacs (Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., Berendsen, H. J. 2005. "GROMACS: fast, flexible, and free". J Comput Chem 26 (16): 1701–18. doi:10.1002/jcc.20291). Os três complexos minimizados gerados foram nomeados como XfAa1, XfPa1 e XfPa1, e estão listados na tabela 1. Foi feita a validação dos complexos modelados usando a análise de gráficos de Ramachandran (Ramachandran, G. N., Ramakrishnan, C., Sasisekharan, V. 1963. Stereochemistry of polypeptide chain conformations. J. Mol Biol 7:95-99), que indicou que os três modelos tiveram mais de 98% dos resíduos nas regiões permitidas. A validação através do ProsaWeb indicou que os valores de z-score variaram entre -8.96 e -9.67, valores que estão contidos no intervalo de resultados normalmente encontrados para as proteínas nativas de tamanho similar (Whiederstein, M., Sippl, M. J. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acid Research, Web Server issue: W407-10). Portanto, os três modelos de complexos de XfPilT foram considerados aceitáveis para a continuação das análises.

Tabela 1. Identificação dos três modelos criados, templates nos quais se baseou a modelagem por homologia, descrição e valores obtidos para validação dos modelos.

Modelos de hexâmero de Xf PilT gerados	Templates usados na modelagem por homologia	Descrição

Xfa1.pdb	2gsz.pdb	Modelo criado usando o Modeller com base na AaPilT hexamérica ligada a ADP (PDB: 2gsz), com adicional minimização de energia realizada com o programa Gromacs
Xfa1.pdb	2gsz.pdb	Modelo criado usando o Modeller com base no hexâmero de PaPilTsem ligantes (PDB: 3jvu), com adicional minimização de energia realizada com o programa Gromacs
Xfa1.pdb	2gsz.pdb	Modelo criado usando o Modeller com base na PaPilT ligado a um análogo do ATP (AMP-PCP) (PDB: 3jvv), com adicional minimização de energia realizada com o programa Gromacs

As estruturas modeladas dos hexâmeros da PilT de *X. fastidiosa* (XfAa1, XfPa1 e XfPa2) foram estruturalmente alinhadas contra seus templates (AaPilT com ADP: 2gsz, PaPilT sem ligantes: 3jvu e PaPilT com ATP: 3jvv), como mostrado na figura 1, o que resultou em sobreposições com desvios de 0.388, 0.297 e 0.231 para os três alinhamentos respectivamente mostrados em 1a, 1b e 1c, o que indica que as estruturas dos complexos são muito similares entre si (template vs. o modelo gerado).

Utilizando a plataforma Blue Star STING (Neshich, G., Togawa, R., Mancini, A. L., Kuser, P. R., Yamagishi, M. E. B., Pappas Jr., G., Torres, W. V., Campos, T. F., Ferreira, L. L., Luna, F. M., Oliveira, A. G., Miura, R. T., Inoue, M. K., Horita, L. G., de Souza, D. F., Dominiquini, F., Álvaro, A., Lima, C. S., Ogawa, F. O., Gomes, B. G., Palandrani, J. C. F., dos Santos, G. F., de Freitas, E. M., Mattiuz, A. R., Costa, I. C., de Almeida, C. L., Souza, S., Baudet, C. e Higa, R. H. 2003. STING Millennium: a Web based suite of programs for comprehensive and simultaneous analysis of protein structure and sequence. *Nucleic Acids Research*, 31:13, 3386-3392) e o programa Java Protein Dossier (JPD) (Neshich, G., Rocchia, W., Mancini, A.L., Yamagishi, M.E., Kuser, P.R., Fileto, R., Baudet, C., Pinto, I.P., Montagner, A.J., Palandrani J.F., Krauchenco, J.N., Torres, R.C., Souza, S., Togawa, R.C., Higa, R.H. 2004. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic Acids Res.* 2004 Jul 1;32 (Web Server issue):W595-601), principalmente no módulo chamado “Formiga” (Higa, R. H. ; Neshich, G. Defining 3D residue environment in protein structures using SCORPION and FORMIGA. *Bioinformatics* (Oxford), Oxford, v. 20, n. 12, p. 1989-1991, 2004), foi possível selecionar e analisar os resíduos formadores de interface bem como desvendar os valores de áreas que estes ocupam na Interface e na Superfície livre, dados estes que são

calculados através do programa SurfV (Sridharan, S., Nicholls, A. and Honig, B. 1992. A new vertex algorithm to calculate solvent accessible surface areas. *Biophys. J.*, 61, A174).

Com esta análise, foi possível observar que a natureza dos Resíduos Formadores de interface (IFR) dos complexos da PilT de *Xylella fastidiosa* é predominantemente de
 5 de interface (IFR) dos complexos da PilT de *Xylella fastidiosa* é predominantemente de resíduos carregados e polares, como pode ser visto na figura 2, similar aos resultados obtidos por Satyshur e colaboradores (Satyshur K. A., Worzalla G. A., Meyer L. S., Heiniger E. K., Aukema K. G., Misis A. M., Forest K. T. 2007. Crystal structures of the pilus retraction motor PilT suggest large domain movements and subunit cooperation
 10 drive motility. *Structure*. 2007 Mar;15(3):363-76) que indicaram que as interfaces seriam predominantemente carregadas e polares. A soma de todas as áreas de interface de todos os monômeros do hexâmero da PilT de *Aquifex aeolicus* é 18890.20 Å² dos quais 13455.55 Å² (71,23%) são formados por resíduos das categorias polares e carregados. Para os hexâmeros da PilT de *X. fastidiosa*, tem-se 18173.88 Å², 18446.96
 15 Å² e 18882.558 Å² de áreas totais de interface e 13592.32 Å² (72,5%), 14020.77 Å² (70,88%) e 14208.91 Å² (70,88%) de áreas ocupadas por resíduos polares e carregados, respectivamente para XfAa1, XfPa1 e XfPa2, o que demonstra de maneira clara a importância que tais resíduos exercem na constituição das IFRs. As interfaces dos hexâmeros de Pa têm valores de ocupação de área por aminoácidos polares e carregados
 20 similar à encontrada para Xf: 71,42% e 70,93% para 3jvu e 3jvv, respectivamente.

Foram usadas abordagens estatísticas para testar se os valores das médias dos percentuais de áreas ocupadas por resíduos polares e carregados eram
 significativamente diferentes entre os encontrados para as cadeias dos hexâmeros de PilTs de Pa e Xf em relação ao hexâmero de Aa. Foi observado que estes dados tem alta
 25 probabilidade de seguirem uma distribuição normal (P-value do teste de normalidade D'Agostino = 0.092) e possuem desvios-padrão de valores similares. Sendo assim, foi utilizado o Teste T de Student que resultou em um valor de P (5.25×10^{-6}) que demonstrou que as interfaces da PilT de Aa (2gsz) tem alta probabilidade de serem
 30 PilTs das proteobactérias patogênicas Pa e Xf sendo, portanto, sugerido que as interfaces de 2gsz têm menor área ocupada por tais resíduos se comparada às interfaces das PilTs de Pa e Xf. Foi testado também, através do teste T, se os valores de áreas

ocupadas por resíduos polares e carregados eram similares entre os encontrados nas interfaces dos modelos em relação aos templates usados para sua modelagem. Na comparação de 3jvu (hexamero da Pa PilT sem ligantes) contra XfPa1 e de 3jvv (hexamero da Pa PilT ligada a AMP-PCP) contra XfPa2 resultaram em valores de P equivalentes a 0.96 e 0.89, respectivamente, indicando que existe uma alta probabilidade de se referirem a dados semelhantes que, neste caso, sugerem que as interfaces das PilTs de Pa e Xf são igualmente ocupadas em sua maioria por resíduos polares e carregados. Na comparação de XfAa1 contra o template usado em sua modelagem, 2gsz (hexamero da PilT de Aa), foi obtido um valor de P de 0.019, o que indica que se tratam de amostras diferentes, sugerindo que Aa PilT tem uma distinta ocupação de área por tais resíduos em relação a Xf PilTs. Estes resultados podem sugerir que as interfaces das proteobactérias patogênicas são mais polares e carregadas que as mesmas de outras bactérias distantes como Aa. Esta característica é de suma importância pois a presença de vários resíduos capazes de estabelecer contatos de cunho eletrostático e outros contatos polares como pontes de hidrogênio indica que vários sítios dentre estes podem ser usados para a identificação de um alvo para o desenho de drogas baseados em estrutura (A. C. Anderson. The process of structure-based drug design. Chem Biol. 2003 Sep;10(9):787-97).

Estes resultados podem indicar uma característica importante das PilTs de Pa e Xf: suas interfaces parecem ter área ocupada ligeiramente maior por resíduos que são capazes de formar interações ou contatos fortes (por exemplo: de cunho eletrostático atrativo/repulsivo e pontes de hidrogênio), em relação ao complexo da PilT de Aa, 2gsz. Complementarmente, as interfaces da PilT de *A. aeolicus* tem área ligeiramente mais hidrofóbica que interfaces da PilT de *X. fastidiosa*, o que pode indicar uma menor propensão dos IFRs de 2gsz de *A. aeolicus* para formar contatos de alta energia. Uma maneira de testar essa última hipótese é através da análise do número, tipo e energia de contatos estabelecidos entre as IFRs de ambos hexâmeros. Utilizando o programa STING JPD pode-se identificar quais IFRs estavam estabelecendo contatos através das dadas interfaces. Foi notado que muitos aminoácidos hidrofílicos são identificados estabelecendo contatos entre as interfaces da PilT de *X. fastidiosa*.

Utilizando os programas Blue Star Sting e JPD foram coletados dados sobre o número, tipo e as energias de contatos estabelecidos na interface. Para valores de área

aproximadamente similares entre Xf e AaPilT, a média da soma de todas as energias de contato de IFR para os três complexos das PilT de *X. fastidiosa* (7.008,53 Kcal.mol⁻¹, com erro padrão da média (SEM) de 190.55) é superior ao valor correspondente para a soma das energias de contato nas interfaces da PilT de *A. aeolicus* (5.203,60 Kcal.mol⁻¹). Em relação ao número de contatos estabelecidos por IFRs, a interface da PilT de *A. aeolicus* estabelece um total de 884 contatos, enquanto que as interfaces dos modelos da PilT de *X. fastidiosa* faz, em média um número surpreendentemente superior: 1.218 contatos (SEM de 12.86). O mesmo é observado para as interfaces das PilTs de Pa: 1023 contatos (SEM de 47) que estabelecem cerca de 6.487,40 Kcal.mol⁻¹ (SEM de 175,80) de energias de interação em média, para todo o complexo.

Quanto aos contatos carregados de força atrativa, carregados de força repulsiva e ligações de hidrogênio (contatos não-hidrofóbicos/energeticamente mais fortes) estabelecidos entre os IFRs, as interfaces de *X. fastidiosa* e *P. aeruginosa* demonstram ter contatos em maior número e com maior energia do que os estabelecidos nas interfaces da PilT de *A. aeolicus*. Enquanto AaPilT (2gsz) tem, em média, 88 desses contatos por cadeia e uma média de 824,87 Kcal.mol⁻¹ de energia de contatos por cadeia, as PilTs de Xf (XfAa1, XfPa1 e XfPa2) tem, respectivamente: média de 115, 133 e 134 contatos não-hidrofóbicos por cadeia, que geram uma energia de interações na interface de 1.044,8, 1.143,4 e 1.157,47 Kcal.mol⁻¹. Os mesmos valores para as PilTs de Pa são similares aos das PilTs de Xf: número de contatos não-hidrofóbicos de 126 e 119 e energias de 1.074,13 e 1.018,93 Kcal.mol⁻¹ em média, por cadeia dos hexâmeros 3jvu e 3jvv. Isso indica que as interfaces dos hexâmeros das PilTs das proteobactérias patogênicas Xf e Pa, independentemente da conformação, estabelecem maior número de contatos de alta energia que as interfaces de 2gsz.

Os dados obtidos foram utilizados para calcular a Densidade de Energia de Contatos nas Interfaces (ICED), que é um índice que pode ilustrar os índices de energia de contatos estabelecidos entre todas as cadeias dos hexâmeros normalizados pela soma da total área de interface ocupada pelos resíduos que estabelecem estes contatos. As estruturas das PilTs de *X. fastidiosa* e *P. aeruginosa* tem maiores ICEDs do que a PilT de *A. aeolicus*, o que sugere que as interfaces das PilTs destes organismos estabelecem mais contatos e com maior energia por área em relação ao que é visto para Aa PilT. Os valores para Xf PilTs são, ainda, superiores aos de Pa PilT, como pode ser visto na

figura 3. A média dos ICEDs para Aa PilT, 2gsz é de $0,28 \text{ Kcal.mol}^{-1}.\text{\AA}^{-2}$, enquanto para as PilTs de Xf e Pa são, respectivamente, 0,38 e 0,34 $\text{Kcal.mol}^{-1}.\text{\AA}^{-2}$.

Para responder a questão se a diferença entre os ICEDs entre as PilTs das bactérias Xf, Pa com relação à Aa é estatisticamente significativa, novamente foi usado o Teste T de Student. Foi observado que os dados de ICED têm alta probabilidade de seguirem uma distribuição normal (P-value do teste de normalidade D'Agostino = 0.061) e possuem desvios-padrão de valores similares. Sendo assim, foi obtido um valor de P equivalente a $3,48 \times 10^{-7}$ para o teste que comparou as médias dos valores de ICED da PilT de Aa (2gsz) contra os mesmos das PilTs de Pa e Xf. Este valor indica que existe uma alta probabilidade de se tratarem de amostras estatisticamente diferentes, e o que se pode sugerir é que as interfaces dos hexâmeros das PilTs das proteobactérias patogênicas tem maiores energias de contatos por área em relação ao hexâmero da PilT de Aa.

Finalmente, outras análises dos contatos de interface foram realizadas: o programa Blue Star STING JPD foi utilizado para visualizar e selecionar os aminoácidos que estabelecem os contatos mais ricos energeticamente (aqui colocamos como valor de corte para sua seleção como estabelecendo no mínimo 10 Kcal.mol^{-1} de energia de contatos na interface), que possuíssem um mínimo de área de exposição ao solvente quando em complexo (nestes casos, na forma hexamérica), que tenham a característica de serem capazes de estabelecer pontes de hidrogênio e/ou contatos de cunho eletrostático e que, não necessariamente, se localizem em bolsos (pockets) em complexo e/ou isolamento. Esta busca resultou em 54 resíduos distintos (sendo que, destes, 42 fazem parte dos bolsos (pockets) adjacentes em complexo ou em isolamento) que provavelmente são os mais prováveis de serem passíveis de uso como alvos terapêuticos para o desenho de fármacos orientado pela estrutura, como é discutido no trabalho de Anderson (A.C. Anderson. The process of structure-based drug design. Chem Biol. 2003 Sep;10(9):787-97. 2003). A listagem destes resíduos, bem como algumas de suas características como: número de cadeias de Xf PilTs que os contenham seguindo tais parâmetros, valores máximos e médios de energia de contatos que o resíduo estabelece na interface e presença em "pockets", estão mostradas na tabela 2.

Tabela 2. Listagem dos 54 resíduos identificados como possíveis alvos terapêuticos para o processo de desenho de fármacos baseados na estrutura e informações

importantes sobre estes, como o número de cadeias de modelos de XfPilTs hexaméricos nas quais o resíduo aparece com alvo, o valor máximo e valor médio de energia de contatos de interface encontrado para o mesmo, presença em “pocket” (bolso) em isolamento (quando a cadeia está isolada do restante do complexo, ou seja: está em seu estado monomérico) e em complexo (ou seja, se na conformação hexamérica há algum bolso adjacente). A última coluna se refere ao alinhamento de sequência primária mostrado na figura 5, em relação a presença/ausência em três grupos de organismos: Proteobactérias patogênicas para animais e plantas (Pt), Aquificae e Cianobactérias (Cy). As seqüências usadas na confecção deste alinhamento foram extraídas do Protein database do NCBI (<http://www.ncbi.nlm.nih.gov/protein/>). As seqüências de Proteobactérias patogênicas são as seguintes: PilT de *Xylella fastidiosa* 9a5c com GI:15838234 (Xf), *Xanthomonas axonodis pv citri* str. 306 GI:21243651 (Xa), *Pseudomonas syringae pv tabaci* ATCC 11528 GI:257483590 (Ps), *Pseudomonas aeruginosa* GI: 301015820 (Pa), *Ralstonia solanacearum* UW551 GI: 83747263 (Rs), *Vibrio cholerae* V52 GI: 121728590 (Vc), *Dichelobacter nodosus* VCS1703A GI: 146329541 (Dn), *Neisseria gonorrhoeae* DGI18 GI: 240013043 (Ng), *Neisseria meningitidis* Z2491 GI: 218767246 (Nm). As seqüências de PilTs de Cianobactérias são: *Nostoc* sp. PCC 7120 GI: 17229935 (Ns) e *Synechocystis* sp. PCC 6803 GI: 16331158 (Ss). A seqüência usada de *Aquifex aeolicus* VF5 (Aa) foi a GI: 119389376.

Identificação do resíduo (aminoácido e posição)	Número de cadeias de complexos de XfPilTs na qual o resíduo aparece como alvo	Valor máximo de energias de contatos de interface para o resíduo	Valor médio de energias de contatos de interface para o resíduo	Presença em pockets (bolso) de isolamento	Presença em pockets em complexo (na conformação hexamérica)	Organismo (dos usados no alinhamento da figura 5) que possuem, em sua PilT, este resíduo em posição correspondente em alinhamentos de estruturas primárias
D160	5	30.00	16.00	sim	-	todas
D17	17	40.60	38.39	sim	sim	todas
D196	18	153.80	107.10	-	sim	todas
D198	16	40.00	36.25	sim	sim	todas
D207	2	20.00	15.00	sim	sim	todas
D242	4	41.20	17.95	-	sim	todas
D70	18	41.82	37.73	sim	sim	todas
E159	5	10.00	10.00	sim	sim	todas

E163	2	20.00	15.00	sim	-	todas
E204	1	20.00	20.00	sim	-	todas
E209	2	40.00	25.00	-	-	todas
E219	17	45.30	29.52	sim	sim	todas
H154	5	51.50	44.10	-	-	todas
H19	8	62.70	55.59	-	-	todas
H222	1	11.50	11.50	-	-	todas
H229	2	15.00	13.25	sim	-	todas
R176	2	42.10	41.80	-	-	todas
R194	18	101.80	79.91	-	sim	todas
R206	6	40.00	18.33	sim	-	todas
R239	16	60.00	34.87	sim	sim	todas
R29	18	104.10	86.07	sim	sim	todas
R294	7	40.00	14.46	-	sim	todas
R80	12	40.00	25.00	sim	sim	todas
R82	9	40.60	29.26	sim	sim	todas
R97	9	50.46	19.31	sim	sim	todas
E258	18	70.00	29.85	sim	sim	Todas as Pt
E74	2	30.00	25.00	-	sim	Todas as Pt
K235	13	40.60	22.54	sim	sim	Todas as Pt
K249	2	10.00	10.00	-	sim	Todas as Pt
R35	7	42.70	28.04	-	sim	Todas as Pt
R90	6	46.40	19.53	sim	sim	Todas as Pt
D33	16	41.20	28.21	sim	sim	Todas as Pt
E248	2	20.00	15.00	sim	-	Todas as Pt
R36	5	53.10	28.26	-	-	Todas as Pt
D181	2	30.00	35.00	-	sim	Somente a Pila de A. testidosa
E89	16	145.80	92.38	-	sim	Somente a Pila de A. testidosa
K157	30	30.00	21.00	-	-	Somente a Pila de A. testidosa
H152	17	84.10	47.41	sim	sim	Xf, Ps, Dn
E336	5	40.00	28.30	sim	sim	Xf, Xa, Ps, Pa, Rs, Dn
K58	2	72.60	46.30	-	sim	Xf, Xa, Ps, Pa, Rs, Ng, Nm
R212	2	33.20	23.20	-	-	Xf, Xa, Ps, Pa, Rs, Dn

R335	15	51.20	34.28	sim	sim	Xf, Xa, Ps, Pa, Vc, Dn
E65	9	36.10	20.49	sim	sim	Xf, Xa, Rs, Vc
E177	4	60.00	42.12	-	sim	Xf, Xa, Ps, Pa, Rs, Aa
K170	1	10.00	10.00	sim	-	Xf, Xa, Ps, Pa, Ng+, Nm+, Aa++
E68	11	40.00	22.73	sim	sim	Xf, Xa, Ps, Pa, Rs, Vc, Dn, Aa
E64	3	20.00	15.27	-	-	Xf, Xa, Ps, Pa, Rs, Vc, Dn, Ng Nm, Aa, Ns, Ss+
R290	18	108.80	78.48	sim	-	Xf, Xa, Ps, Pa, Rs, Vc, Dn, Ng Nm, Ns, Ss
R180	1	10.00	10.00	sim	-	Xf, Xa, Ps, Pa, Vc, Ns, Ss
D181	3	10.00	10.00	sim	-	Xf, Xa, Ps, Pa, Rs, Vc, Dn+, Ng, Nm, Aa
H179	10	36.80	24.00	-	sim	Xf, Xa, Ps, Pa, Vc, Dn, Aa, Ns, Ss
D62	1	21.50	21.50	-	-	Xf, Xa, Ps, Vc, Dn, Ng, Nm
H183	1	10.00	10.00	-	-	Xf, Xa, Ps, Pa, Vc+, Ss+
E120	1	10.00	10.00	-	-	Xf, Xa+, Ps+, Pa+, Dn+, Ng, Nm, Ns+, Ss

Outra importante característica que será considerada no momento do desenho dos fármacos é a frequência de ocorrência do resíduo há ser considerado como um alvo em todas as 18 cadeias possíveis em três complexos modelados. Um determinado resíduo pode ocorrer com estas características em apenas uma conformação da PilT de Xf (como exemplo, o resíduo E74 no modelo XfAa1) como também pode ocorrer com estas características em todas as três conformações (como exemplo, o resíduo E89, que possui tais características nos três hexâmeros: XfAa1, XfPa1 e XfPa2). A importância desta descrição reside no fato que um determinado fármaco pode ter como alvo a ligação em um resíduo em um determinado momento da proteína PilT (ex: se em seu estado ligado ATP ela apresentar estrutura do hexâmero como prevemos no modelo XfPa2), ou em vários. Os futuros processos de “docking” de compostos químicos terão como alvo uma ou mais cadeias de um ou mais hexâmeros, conforme preditas as características de cada alvo em cada cadeia de cada hexâmero. Portanto, foi gerado um diagrama de Venn contendo a distribuição dos 54 resíduos selecionados nos três hexâmeros modelados, como mostrado na figura 4.

Em relação a definição dos alvos terapêuticos, como já foi citado anteriormente, o desenho de drogas antimicrobianas deve se basear em alvos que sejam principalmente encontrados nos patógenos (e não em organismos não-patogênicos). Para averiguar se,

dentre os 54 resíduos selecionados, existiriam resíduos de ocorrência restrita a organismos patogênicos como Xf, Pa e outros descritos anteriormente, em relação a organismos não-patogênicos (como Aa) e mesmo a organismos essenciais ao equilíbrio de ecossistemas como as cianobactérias (bactérias fotossintetizantes e fixadoras de nitrogênio atmosférico), foi realizado um alinhamento de estrutura primária incluindo PilT de vários organismos que se incluem nestas categorias e que se tem registro na literatura sobre presença da PilT, Pilus tipo IV e motilidade do tipo “twitching”. Este alinhamento, representado na figura 5, indica que existe uma alta conservação dos IFRs, descritos acima como potenciais alvos, neste alinhamento de seqüências PilTs.

Esta é mais uma possível evidência da importância dos IFRs para organismos patogênicos que fazem uso do PilT IV. Pode-se especular que as seqüências das PilT de Proteobacteria, por serem extremamente conservadas (Wall, D. e Kaiser, D. Type IV pili and cell motility. Mol Microbiol. 1999 Apr;32(1):1-10. Review), teriam suas estruturas igualmente semelhantes e, em caso afirmativo, as semelhanças e as diferenças observadas neste alinhamento também poderiam refletir as reais características da estrutura.

Neste alinhamento (figura 5) foram destacados os 54 resíduos sendo que, destes, 25 ocorrem em todos os organismos, co-localizando com a posição do resíduo em XfPilT (D160, D17, D196, D198, D207, D242, D70, E159, E163, E204, E209, E219, H154, H19, H222, H229, R176, R194, R206, R239, R29, R294, R80, R82 e R97), que constituiriam alvos inespecíficos dos patógenos; 9 ocorrem em todas as Proteobactérias patogênicas (E258, E74, K235, K249, R35, R90, D33, E248 e R36) constituindo bons alvos para o desenvolvimento de fármacos que atinjam esta gama de patógenos; 3 resíduos são exclusivos da PilT de *Xylella fastidiosa* (D184, E89 e K187) que podem constituir alvos terapêuticos para desenvolvimento de fármacos mais específicos à PilT da Xf; outros 6 resíduos ocorrem em distribuição variada mas ainda somente em Xf e outras proteobactérias patogênicas, porém não em todas (H152, E336, K58, R212, R335 e E65); e, por fim, 17 resíduos que ocorrem em distribuição variada entre Xf, outras proteobactérias e sempre incluindo Aa ou alguma das cianobactérias, constituindo também alvos mais inespecíficos (H152, E336, K58, R212, R335, E65, E177, K170, E68, E64, R290, R180, D181, H179, D62, H183 e E120).

Além da indicação dos possíveis alvos terapêuticos para o desenho de fármacos contra XfPilT, esta análise nos permite selecioná-los quanto à sua possível especificidade, já que se objetiva alvos que ocorram preferencialmente em Xf e/ou outras bactérias patogênicas de interesse médico/econômico. Portanto, os 18 alvos aqui chamados de “preferenciais” por ocorrerem somente em Xf ou em proteobactérias patogênicas são: D184, E89, K187, E258, E74, K235, K249, R35, R90, D33, E248, R36, H152, E336, K58, R212, R335 e E65 (porém isso não restringe que os outros 36 resíduos não sejam usados, mas sim que os 18 resíduos citados serão preferenciais ao desenho dos fármacos). Outro fato interessante é que vários resíduos carregados que possuem a característica de contribuírem muito em termos energéticos para as interfaces de XfPilT não ocorrem em AaPilT, o que pode ser interpretado como uma possível explicação para as diferenças estatisticamente comprovadas entre os valores de ICED e ocorrência de áreas de resíduos polares e carregados entre os hexâmeros das PilTs de Pa e Xf e o hexâmero da PilT de Aa.

Várias hipóteses podem ser formuladas para explicar como estas observações refletem sobre a função da proteína, desempenho e estabilidade, e estes podem ser discutidas em estudos futuros.

Assim, os aminoácidos mais importantes dentre os IFRs, ou seja: os IFRs que estabelecem contatos mais energéticos, são conservados dentre um grupo de Proteobacterias patogênicas que se utilizam do “twitching motility” dependente de PilT como mecanismo de movimentação. Estes aminoácidos conservados (somados aos específicos para Xf) são os apontados como os alvos terapêuticos preferenciais para o desenvolvimento de fármacos por complementaridade físico-química-estrutural, não invalidando a possibilidade de uso dos outros resíduos (por exemplo, os que ocorrem em todos os organismos).

O principal objetivo será a produção de um fármaco que se ligue na futura região de interface de certo monômero e impeça sua oligomerização e/ou que esse fármaco se ligue já no hexâmero oligomerizado, em alguma de suas conformações, e impeça o correto funcionamento da interface e atividade ATPásica.

Exemplo prático de Desenho de fármacos baseado na estrutura (“Structure based drug design (SBDS))

Técnicas de desenho de drogas baseado em estrutura protéica podem ser aplicadas com base em possíveis alvos terapêuticos preditos. A seguir será feita uma exemplificação de um dos vários caminhos que podem ser adotados para o desenho *de novo* de fármacos, que se inicia no mapeamento de posições favoráveis de interações para grupos funcionais (ex: em que posições podem ser desenhados grupos hidroxila, amina, hidrofóbicos, cíclicos, dentre outros) ou mesmo pequenos fragmentos de moléculas. Novos compostos podem, então, ser desenhados *de novo* de modo que relevantes grupos funcionais se localizem em posições que determinem uma correta relação espacial com o sítio-alvo. Após o desenho deve ocorrer a modelagem de sua estrutura tridimensional, ensaios de “docking”, escolha dos melhores ligantes, predição das bases moleculares/estruturais de sua ligação (ex: tipo de interações que este é capaz de realizar). Existe uma série de “softwares” úteis para o desenho de fármacos *de novo* e abordagens de “screening”, como o SPROUT (<http://www.simbiosys.ca/sprout/>), muito usado para o desenho *de novo* baseado em fragmentos. Este programa inclui módulos para identificar e selecionar grupos funcionais e posições nos sítios-alvo para formar fragmentos iniciais de compostos para geração de estrutura (módulo EleFAnT) e, conforme estes são selecionados, são gerados esqueletos que satisfaçam as restrições estéricas do pocket-alvo através de crescimento de fragmentos espaçadores e conectando-os aos fragmentos iniciais (módulo SPIDeR). Por fim, é feita a substituição de átomos no esqueleto até gerar moléculas que sejam compatíveis com as propriedades eletrostáticas do sítio-alvo (módulo MARABOU). As soluções podem ser clusterizadas e terem scores de ligação calculados usando o módulo ALLigaTOR.

Além do desenho *de novo* de compostos, objetivar-se-á a busca por compostos que se liguem nos sítios-alvo preditos por meio de “virtual screening” em larga escala, através do uso de bancos de dados de estruturas tridimensionais de pequenas moléculas, como exemplo, o PubChem Compound (ftp://ftp.ncbi.nlm.nih.gov/pubchem/Compound_3D/) como fonte de ligantes a serem usados em screening contra um sítio alvo. A simulação da interação proteína-ligante com o intuito de avaliar a capacidade de um ligante em formar interações fortes e de encaixar-se a um sítio alvo é chamado “docking”. Algoritmos de “docking”, que incluem os seguintes (porém não restritos a estes): Dock (Kuntz, I., Blaney, J., Oatley, S., Langridge, R., e Ferrin, T. 1982. A geometric approach to macromolecular-ligand

interactions. J. Mol. Biol. 161, 269–288), FlexX (Kramer, B., Metz, G., Rarey, M., e Lengauer, T. 1999. Ligand docking and screening with FlexX. Med. Chem. Res. 9, 463–478) AUTODOCK (Goodsell, D., Morris, G., and Olson, A. 1996. Automated docking of flexible ligands: applications of AutoDock. J. Mol. Recognit. 9, 1–5). MOL-DOCK (Thomsen, R., Christensen, M. H. MolDock: a new technique for high-accuracy molecular docking. J. Med Chem. 2006 Jun 1;49(11):3315-21.) e GOLD (Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. J. Mol. Biol. 1997, 267, 727–748) podem ser usados para realizar o “docking” na região de interesse para encontrar as moléculas que estabelecem interações mais favoráveis e que se adaptem melhor ao sítio-alvo, sendo escolhidas por meio do ranqueamento do “score” de “docking”.

No exemplo a ser descrito a seguir, objetivou-se a demonstração prática de um, dos vários possíveis, “pipeline” de como poderá ser feita a predição de estruturas de compostos que possivelmente se ligarão aos alvos preditos através de um desenho *de novo*. A metodologia não se restringe às moléculas a serem descritas neste texto, porém a todas as moléculas que serão desenhadas com base nas características físico-químicas e estruturais dos sítios escolhidos como alvos terapêuticos, seja por meio de desenho *de novo* de fármacos usando fragmentos, softwares de predição, “virtual screening”, “screening” usando farmacóforos ou por outros métodos.

Foram desenhados *de novo* cerca de 50 compostos, baseados no sítio que abriga o resíduo-alvo E89, da cadeia A do modelo XfAa1, de *Xylella fastidiosa*, sendo este um resíduo considerado exclusivo de Xf, o que sugere que o fármaco que será desenhado com base neste alvo, possivelmente terá especificidade à PilT de Xf. Primeiramente é indispensável a descrição físico-química e estrutural do sítio-alvo em questão, que pode ser visto na figura 6. Nesta estrutura, o E89 está localizado entre as interfaces CTDⁿ:NTDⁿ⁺¹, sendo localizado no NTD da cadeia A e interage com vários resíduos em interface (Ponte de hidrogênio com Q175, duas interações eletrostáticas de repulsão com E74 e uma interação hidrofóbica com I162, todos da cadeia F), resultando em um total de 23,8 Kcal/mol de energia de interação. Usando os programas STING ¹PD e Molegro Virtual Docker (MVD), pode-se observar que este resíduo faz parte dos “pockets adjacentes”. O programa MVD realizou a predição de dois “pockets” que circundam o resíduo E89 (“pocket” 1 que possui 55,29 Å³ de volume e “pocket” 2 com

45,05Å³ de volume). A fusão destes bolsos originou um novo “pocket” de volume igual a 100,35Å³ e área de superfície equivalente a 381,44Å², como mostrada na figura 6. Os resíduos adjacentes a este “pocket” (à distâncias de menos de 6Å) incluem, além do E89, mais dois resíduos carregados negativamente (Glu 177 e Glu 74 da cadeia F), três
 5 resíduos carregados positivamente (Arg90 e Lys45 da cadeia A e Arg180 e Arg80), oito resíduos apolares (Ile162, Val165, Leu172 e Ile173 da cadeia F, Ala22, Gly23, leu24 e Ala92 da cadeia A) e três resíduos polares (Gln88 e Asn87 da cadeia A e Gln 175 da cadeia F).

Com isso, seguindo a estrutura do “pocket” no que diz respeito aos limites deste
 10 e tamanhos de átomos, foram desenhadas cerca de 40 estruturas químicas de fármacos que possivelmente interagiriam com a região mencionada, principalmente com os resíduos-alvo preditos e localizados próximos ao “pocket” (E89 (resíduo exclusivo de Xf), R90 e E74 (resíduos exclusivos de Proteobactérias patogênicas)). O trabalho se iniciou com o desenho de quatro estruturas-base usando o programa ChemDraw
 15 (<http://www.cambridgesoft.com/>) e modelagem da estrutura tridimensional no módulo Chem3Dpro. Este módulo permitiu, ainda, que as pequenas moléculas modeladas em 3D passassem por minimização de energia e dinâmica molecular para atingir uma conformação estrutural aceitável. Estas estruturas desenhadas têm em comum certas características: presença de duas extremidades cíclicas, que podem ser: anel de cinco carbonos e anel de seis carbônicos (aromático ou não), para interação com resíduos
 20 apolares e ocupar os espaços largos do “pocket”, uma região carbônica espaçadora entre os anéis (entre quatro e cinco carbonos), um grupo carboxila ionizável ligado a um dos anéis para interação (preferencialmente de ordem eletrostática) com a Arg90, mais próxima do “pocket”, um ou dois grupos amina ionizáveis para interação com os
 25 resíduos polares e negativos e presença ou ausência de grupo hidroxila.

Foi então realizada uma simulação de “docking” (com ligante flexível) usando-se o programa MVD, centrado no “pocket” mencionado e com um raio de espaço de busca de 15Å. O algoritmo de cálculo de “score” usado foi o MolDock Score [GRID], usando uma resolução de Grid de 0,3Å (avaliando o ligante por meio de Interações
 30 internas de pontes de hidrogênio, interações eletrostáticas e torções Sp²-Sp²). O algoritmo de busca e “docking” usado foi o MolDock Optimizer (que é uma implementação de uma variação do algoritmo evolucionário), com 10 rodadas, tamanho

de população igual a 100 e com 3000 iterações. Após o “docking”, o MVD realizou minimização de energia e otimização de pontes de Hidrogênio.

Das quatro estruturas-base desenhadas e modeladas, o composto que obteve o melhor “dock score” destas (-105,62) foi nomeado como E89C4m4. Posteriormente, avaliando-se que átomos poderiam ser alterados/retirados/adicionados para otimizar a ligação ao sítio alvo e, especificamente ao resíduo-alvo (E89), foram feitas mudanças através do ChemDraw e nova modelagem usando Chem3Dpro. Desta forma, foram realizadas, ao todo adicionais quatro rodadas de “docking” com moléculas que foram desenhadas com base nos melhores compostos das rodadas anteriores. Os compostos com melhores MolDock Scores de cada uma das cinco rodadas (E89Cm4, E89Cm29, E89C4m30, E89C4m32 e E89C4m34), bem como as porções envolvidas nas interações e os resíduos que interagiram com a molécula em questão estão representados na figura 7 e os dados referentes ao “docking” destes ligantes estão mostrados na tabela 3. Foi obtida, ao final, uma estrutura com melhor valor de MolDock Score (-131,118), que estabelece -142,72 Kcal/mol de energia de interação com a proteína.

Tabela 3. Dados referente ao “docking” das moléculas citadas e os valores obtidos de MolDock Score, energia de interação com a proteína, energias referentes à interações eletrostáticas de curta distância (menor que 4,5Å) e longa distância (maior que 4,5Å), energia associada ao estabelecimento de pontes de hidrogênio (sendo que todos os valores de energias são dados em Kcal/mol), número de átomos pesados e peso molecular de cada composto e a eficiência do ligante (LE1) dada como a razão entre o MolDock Score e o número de átomos pesados.

Ligando	MolDock Score	Interação	Interação eletrostática (curta distância)	Interação eletrostática (longa distância)	Pontes de Hidrogênio	Resíduo Alvo	Peso Molecular	LE1
E89C4M34	-131.118	-142.772	-22.8288	-2.55371	-6.36607	24	333.468	-5.46237
E89C4M30	-124.067	-128.484	0	0	-11.3605	24	333.422	-5.16945
E89C4M29	-120.767	-127,069	0	0	-6.8006	24	333.4222	-5.03195
E89C4M32	-111.706	-133.849	-21.9648	-1.95334	-7.15845	24	332.48	-4.6544
E89C4M4	-105.62	-108.811	-17.8373	-0.42324	-5.55281	20	283.386	-5.281

O composto aqui nomeado como E89C4m34 (Figura 7) tem fórmula $C_{19}H_{32}N_3O_2$, notação SMILES [O-]C(=O)C2=CC=C(C(CCCC\CCCCC)=C/[NH3+])CC[NH3+])CN2, peso molecular de

334,484 Da, 56 átomos (sendo 24 átomos pesados), 11 ligações rotáveis, 2 anéis, sendo um aromático. Foi usado o programa Ligand Scout para predição de algumas características polares e químicas, como o coeficiente de partição em octanol/água (LogP), que influencia nas propriedades ADME (Absorção, Distribuição, 5 Metabolização e Excreção) da droga por indicar a lipofilicidade do composto. O E89C4m34 possui cLogP de -0,923, o que indica que é solúvel em água, e com TPSA (Área de Superfície Polar Total) de 108Å². Possui, ainda 3 átomos doadores de pontes de hidrogênio (átomo N 23 que estabelece pontes de hidrogênio com os resíduos Glu89A e Glu74F, átomo N 17 que estabelece ponte de hidrogênio com Gln175F, e 10 átomo N 10), dois átomos aceptores de ponte de hidrogênio (O19 e O20), um negativamente ionizável (O19, que interage fortemente através de interação eletrostática atrativa de curta distância (menor que 4,5Å) com Arg90A) e dois positivamente ionizáveis (átomos N 17 e N 23, sendo este último responsável por duas fortes interações eletrostáticas atrativas de curta distância com os resíduos E89A e E74F). Os 15 anéis são importantes para o estabelecimento de interações hidrofóbicas com resíduos apolares do sítio-alvo, como Ala22A e Ile173F, tabela 04.

Tabela 4: Dados referentes aos alvos de interação do composto E89C4m34 com o sítio-alvo, mostrando a que cadeia este pertence, seu nome, identificação, energia total estabelecida em contatos e energia proveniente de interação eletrostática de curta e 20 longa distância.

Cadeia alvo	Aminoácido	Identificação do resíduo	Energia total estabelecida com o resíduo	Energia proveniente de interação eletrostática de curta distância	Energia proveniente de interação eletrostática de longa distância
XfAa1[A]	Glu	89	-41.3234	-1.02331	-14.3845
XfAa1[A]	Arg	90	-18.0581	0.457393	-1.18198
XfAa1[F]	Glu	74	-13.1333	-0.00599243	-7.26196
XfAa1[F]	Gln	175	-9.68925		
XfAa1[A]	Lys	45	-8.52262	-1.34031	
XfAa1[A]	Gly	23	-7.64588		
XfAa1[A]	Ala	22	-7.18962		
XfAa1[F]	Ile	173	-7.06743		

XfAa1[A]	Gln	88	-5.28542		
XfAa1[A]	Gly	91	-5.21324		
XfAa1[F]	His	166	-3.54791	0.877045	
XfAa1[F]	Thr	167	-3.28522		
XfAa1[F]	Val	165	-2.68033		
XfAa1[F]	Asn	174	-2.24359		
XfAa1[F]	Ser	168	-1.96354		
XfAa1[F]	Glu	159	-1.70918	-1.70918	
XfAa1[F]	Glu	177	-1.44564	-1.44564	
XfAa1[F]	Asp	181	-1.38837	-1.38837	
XfAa1[A]	Leu	24	-1.37025		
XfAa1[A]	Ser	21	-1.25635		
XfAa1[A]	Ala	92	-0.65466		
XfAa1[A]	Val	47	-0.45698		
XfAa1[F]	Asp	145	-0.38244	-0.382442	
XfAa1[F]	Glu	163	-0.34223	-0.342231	
XfAa1[A]	His	44	-0.30635		
XfAa1[F]	Lys	149	0.308608	0.308608	
XfAa1[F]	Lys	102	0.442583	0.442583	
XfAa1[F]	Arg	180	1.21.20099	1.20099	
XfAa1[F]	Arg	80	1.73437	2.22489	

A listagem completa dos resíduos que interagem com o composto, bem como a energia total de interações estabelecidas entre eles e os átomos do E89C4m34 e a energia estabelecida através de interações eletrostáticas de curta e longa distância. Os resíduos que estabelecem as interações mais fortes são: E89 e R90 da cadeia A (-41,32 e -18,05 Kcal/mol, respectivamente) e com E74 da cadeia F (-13,13 Kcal/mol), que são resíduos incluídos na lista de alvos terapêuticos exclusivos de Proteobactérias patogênicas, sendo a E89 exclusiva de Xf. A representação das superfícies que delimitam o “pocket” que compõem o sítio-alvo de ligação está mostrada na figura 8a, mostrando como o ligante possui formato similar ao da região. Na figura 8b está representado o E89C4m34 e as posições dos resíduos que interagem com esta molécula.

Demonstramos, neste pedido de patente, a metodologia para identificação de alvos terapêuticos específicos para o desenho de novos fármacos em estruturas modeladas da proteína PilT de *Xylella fastidiosa* e que tem possível aplicação para organismos patogênicos que também possuem esta proteína com alto grau de

5 similaridade de sequência e correspondência dos resíduos indicados como alvos em alinhamento de estrutura primária. Em seguida, propomos uma lista de 54 resíduos alvos a serem usados em processos de desenho de drogas baseado em estrutura sendo que, desses, 18 serão usados preferencialmente. Finalmente, foram propostas metodologias de desenho computacional de fármacos e a proposição de algumas

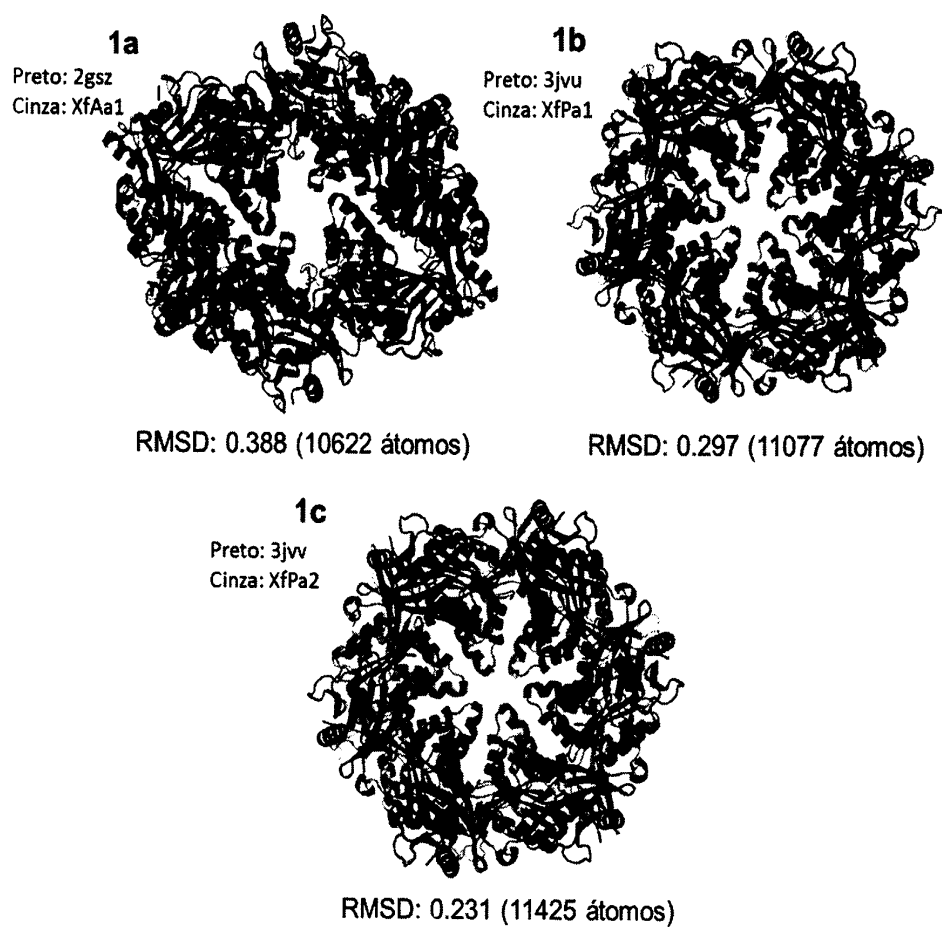
10 estruturas de compostos químicos para ligação no sítio-alvo E89 da PilT de Xf. A caracterização dos determinantes da interação entre o composto desenhado que possuiu as maiores energias de interação com o correspondente alvo na PilT da Xf (composto este aqui nomeado como E89C4m34) foi descrita em detalhes, englobando características físico-químicas e estruturais. É importante ressaltar que o presente

15 pedido de patente não se resume aos fármacos citados, mas sim à metodologia para identificação dos resíduos-alvo na proteína PilT, os resíduos-alvo sugeridos para o desenho de drogas e a proposição de métodos que podem culminar no desenho/identificação de diversos fármacos possíveis de serem usados para impedir o correto funcionamento do motor PilT (seja por impedir oligomerização ou por impedir a

20 dinâmica da interface no hexâmero formado). O fármaco sugerido é apenas uma exemplificação do trabalho que está sendo realizado.

REIVINDICAÇÃO

1. Método de identificação de alvos protéicos caracterizado pelas seguintes etapas: (i) seleção de pelo menos uma seqüência de aminoácidos constituidora de monômero de PilT; (ii) desenvolvimento de um modelo computacional tridimensional da estrutura
- 5 homo-hexamérica de PilT; (iii) análise computacional para determinar os resíduos de aminoácidos formadores de interface (IFR) e suas características físico-químicas e estruturais para todas as cadeias dos modelos de complexos hexaméricos gerados; (iv) seleção das regiões a serem usadas como alvos terapêuticos (e alvos terapêuticos preferenciais) na interface entre os monômeros baseada na intensidade dos seguintes
- 10 parâmetros: (a) energia de contatos interfaciais, (b) exposição, em área em complexo, (c) propensidade do aminoácido-alvo em estabelecer pontes de hidrogênio (como doador ou aceptor), (d) propensidade do aminoácido-alvo em estabelecer contatos de cunho eletrostático, (e) preferencialmente, presença em “pockets” (em isolamento ou em complexo), (f) preferencialmente, que estejam presentes somente em organismos-alvo
- 15 em relação a organismos não patogênicos; (v) desenho computacional de fármacos ou agroquímicos baseado em estrutura-alvo, seja através dos métodos *de novo* descritos, ou com base em estruturas conhecidas através de “*virtual screening*”.

**Fig. 1**

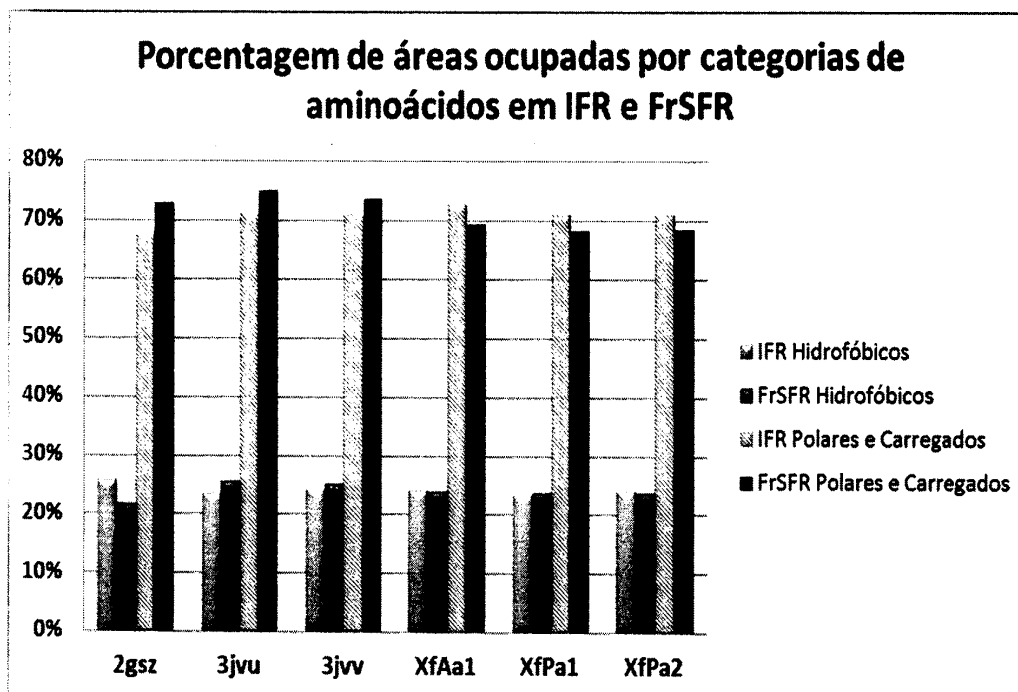


Fig. 2

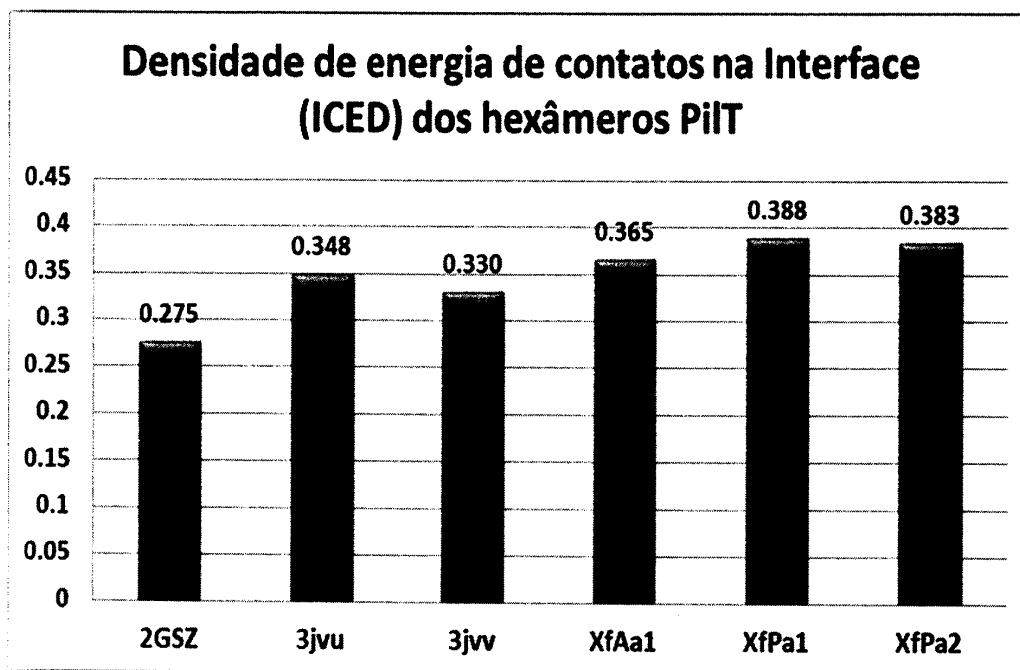
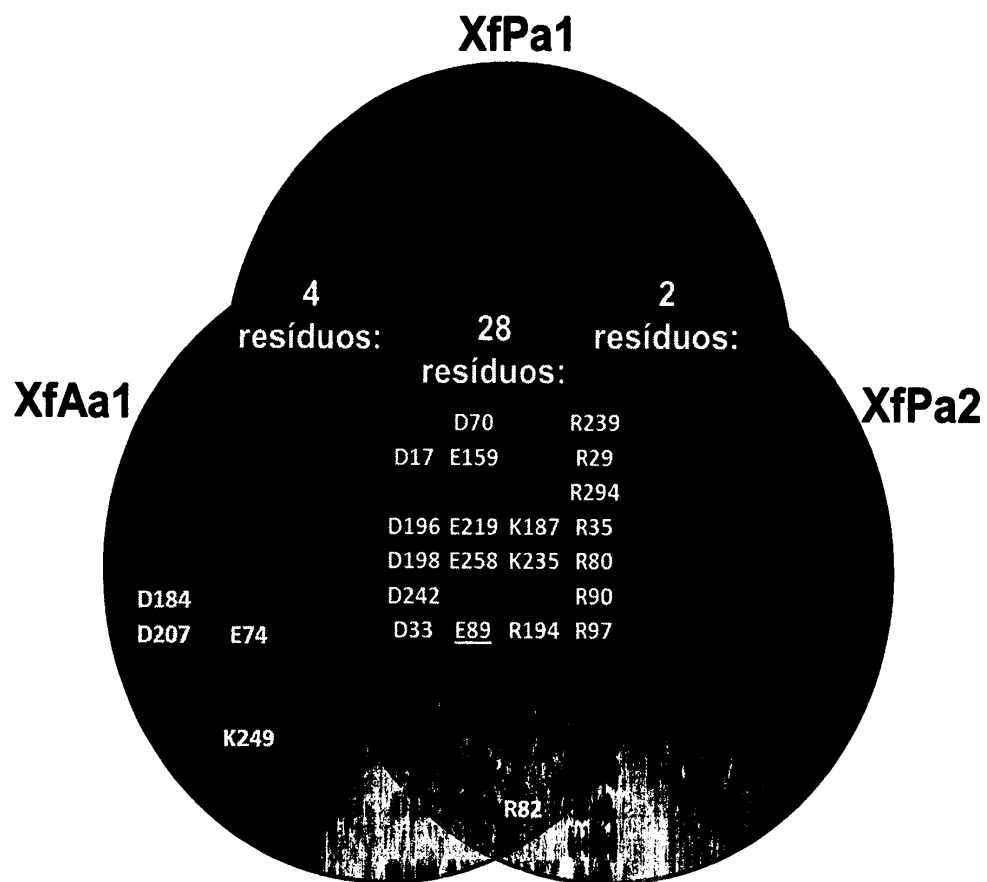


Fig. 3

**Fig. 4**

[illegible]

Fig. 5a

X: RKLIEEPQGLILVTGPTGSGKSTTLAAMIDHINKNAHGHIILTIIDPIEFVHTSQKCLINQ 175
 X: RQLIDQPQGLILVTGPTGSGKSTTLAGMIDYINKNEYGHILTVEDPIEFVHTSQKCLINQ 175
 P: RKITDVARGLILVTGPTGSGKSTTLAAMIDYLNCKNHHIILTIIDPIEFVHESKKCLVNQ 175
 L: KRVS DVPRGLVLTGPTGSGKSTTLAAML DYLNNTKYH IILTIIDPIEFVHESKKCLVNQ 175
 N: VDIIDVPRGLVLTGPTGSGKSTTLAAMIDHINNRRHEIILTVEDPIEFVHESKKCLVNQ 178
 N: QKIAESPRGMVLVTGPTGSGKSTTLAAMINYINETQPAHILTIIDPIEFVHQS KKS LINQ 175
 N: QKIAESPRGMVLVTGPTGSGKSTTLAAMINYINETQPAHILTIIDPIEFVHQS KKS LINQ 175
 R: SDLAMKPRGLVLTGPTGSGKSTTLAAMVNHNRNESDLGHILTVEDPIEFVHESKKSLINQ 175
 V: SKIANYEKGLVLTGPTGSGKSTTLAAMVNVNAHNNKH IILTIIDPIEFVHSNNKCLINQ 175
 N: REMCDKPRGLILVTGPTGSGKSTTLAAMIDLINRTKAEHILTVEDPIEFVYEPKSLVHQ 177
 S: REMAERPRGLILVTGPTGSGKSTTLAAILDLINRTAEHILTIIDPIEFVFPNVRSLFHQ 179
 AaPilt LELCHRKXGLILVTGPTGSGKSTTIASXIDYINQTKSYH IITIDPIEFVFKHKSIVNQ 177
 *: *****:***.::* *:*****:*.:::.*
 178 184 187 194 196 198 204 206 209 219 222 229 235
 X: REVHRDTHFNALSSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 X: REVHRDTHGFNEALRSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 S: REVHRDTHGFSEALRSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 P: REVHRDTHGFSEALRSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 D: REVHRDTQSFSNALRAALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 238
 N: RELHQHTLSFANALSSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 N: RELHQHTLSFANALSSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 R: RELGPHTHSFANALKSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 V: REVHRDTHSFKNALRSALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 235
 N: RQLGEDTKSFANALKAALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 237
 S: RQRGEDTKSFANALRAALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 239
 AaPilt BEVGEDTKSFADALRAALREDHIIILVGEIRDLTIRLALTANETGHLVFGTLETSSAA 237
 : . .* .** :*****:::*** ** *: *:*****. *

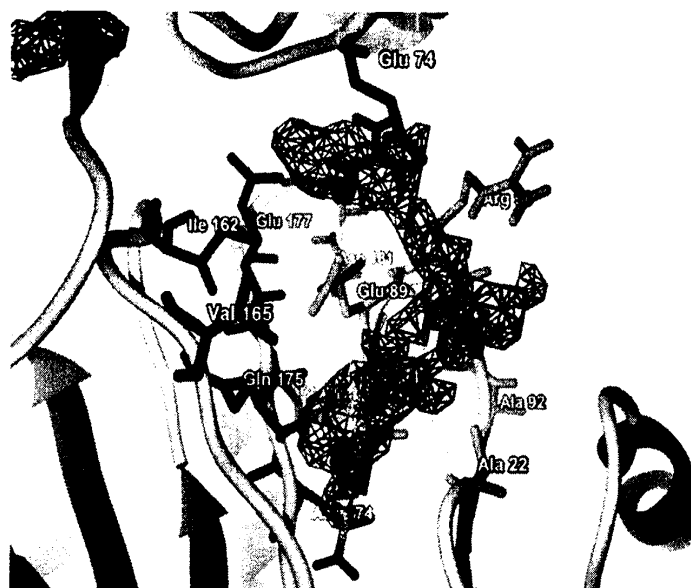
Fig. 5b

```

      239 242      248      249      258
X: SINRIIVFPAAEMVRSMLS SLCGIISQMLLKK----VGG-GRTAWEIMVGTPAIR 290
X: TIDRIIDVFPAGPMVRSMLS SLRAVISQALLKK----VGG-GRTAWEIMVGTPAIR 290
P: TIDRIIDVFPAGSMIRSMLS SLHAVVSQALLKK----VGG-GRVAAHEIMMGTPAIR 290
P: TIDRIIDVFPAGAMVRSMLS SLQSVISQTLIKK----IGG-GRVAAHEIMIGTPAIR 290
D: TIDRIIDVFPGEQLVRSMLS SLRAVIAQTLKK----IGG-GRVAAHEVLVGTSVAVK 293
N: TVDRIIDVFPAGPMVRSMLS SLTAVISQNLKKT----HDGNGRVASHEILIANPAVR 291
N: TVDRIIDVFPAGPMVRSMLS SLTAVISQNLKKT----HDGNGRVASHEILIANPAVR 291
E: TIDRIIDVFPSEDMVRTMLS SLEAVISQTLKKT----RDGSGRVAHEIMICTPAIR 291
E: TIDRIIDVFPGSEDMVRSMLS SLRAVIAQKLLKR----VGG-GRVACHEIMLATPAIR 290
S: TVDRIIDVFPHEKQTQVRVQLSNSLVAVFSQTLVPKKNPKPGEYGRVMAQEIMIITPAIS 297
AaPilt TIHRIIDVIFPLNQEQVRIVLSFILQGIISQRLLPK----IGG-GRVLAYELLIPNTAIR 292
      ...*:*:* * : * * * * : * : * : * : * : * : * : * : * : * :
      294
X: NLIREDKVAQMYSSIQTG-QQYGMQTLDDQHLQDLIKRNLITRQQAREYAK-DKANF---- 344
X: NLIREDKVAQMYSSIQTG-QQYGMQTLDDQHLQDLVKRSLITRNQAREYAK-DKRIFE--- 345
P: NLIREDKVAQMYSSIQTG-GSLGMQTLDMCLADLVKKGLITRESARERAK-VPDNF---- 344
P: NLIREDKVAQMYSAIQTG-GSLGMQTLDMCLKGLVAKGLISRENAREKAK-IPENFGAAA 348
D: NLIREDKVAQIYSTIQTG-SQYGMQTLDDQALSALVKEGKVDRMLAASKAH-DKDNFM--- 348
N: NLIRENKITQINSVLQTG-QASGMQTMDSLSQSLVRQGLIAPEAARRRAQ-NSESMSF-- 347
N: NLIRENKITQINSVLQTG-QASGMQTMDSLSQSLVRQGLIAPEVARRRRAQ-NSESMSF-- 347
E: HLIRENKISQMYSMQTS-SGLGMQTLDDQCLAEIKRSAINYADARAIK-NPDAFAN-- 347
E: NLIREDKVAQMYSSIQTG-AAHGMQTMENAKQLIARGVVDAQEVQSKIETDLKAF---- 345
N: NLIREGKTSQIYSAIQTG-GKLGMTLEKVLADYYKSGTISFEAAMSKTS-KPDEIQRLI 355
S: NLIREGKAAQIYSAIQTG-ARLGMQTMEEGLATLVVSGVISLEEGLAESG-KPDELQRLI 357
AaPilt NLIRENKLQQVYSLXQSGQAETGXQTXNQTLTKLYKQGLITLEDAXEASP-DPKELERXI 351
      :***** * : * * : * * : * * : * * : * * : * * : * * :

```

Fig. 5c

**Fig. 6**

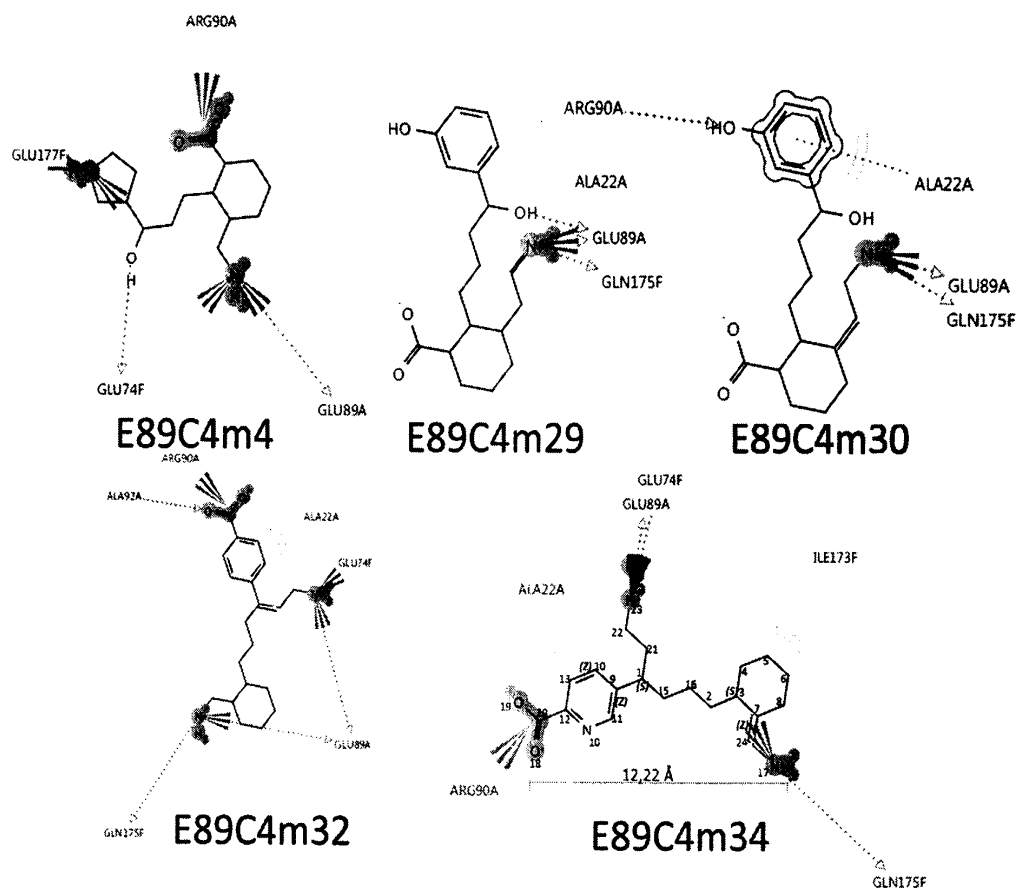


Fig. 7

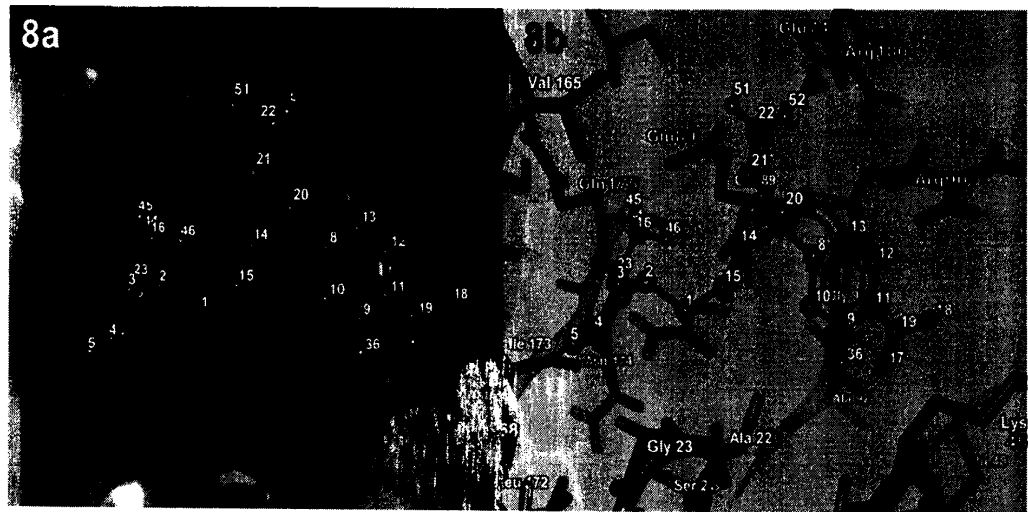


Fig. 8

RESUMO

“IDENTIFICAÇÃO DE ALVOS TERAPÊUTICOS PARA DESENHO COMPUTACIONAL DE DROGAS CONTRA BACTÉRIAS DOTADAS DA PROTEÍNA PILT”.

A presente invenção se refere a um método para identificar regiões alvo existentes na interface de monômeros constituintes da proteína Pilt com o objetivo de desenhar moléculas potencialmente aplicáveis no comprometimento da atividade desta proteína, controlando, assim, processos infecciosos. O método é caracterizado pela (i) seleção de pelo menos uma seqüência de aminoácidos constituidora de monômero de Pilt; (ii) desenvolvimento de um modelo computacional tridimensional da estrutura homo-hexamérica de Pilt; (iii) análise computacional para determinar os resíduos de aminoácidos formadores de interface (IFR) e suas características físico-químicas e estruturais para todas as cadeias dos modelos de complexos hexaméricos gerados; (iv) seleção das regiões a serem usadas como alvos terapêuticos (e alvos terapêuticos preferenciais) na interface entre os monômeros baseada na intensidade de determinados parâmetros; (v) modelagem computacional para o desenho de moléculas potencialmente capazes de efetuar ligações e/ou interações entre regiões alvo dos monômeros. Notadamente, algumas das principais aplicações desta tecnologia consistem no combate à bactéria *Xylella fastidiosa*, agente etiológico da Clorose Variegada dos Citrus (CVC ou Amarelinho), e a doença de Pierce.