

O desenvolvimento de ferramentas computacionais para análise das interações moleculares nas estruturas protéicas vem sendo o alvo fundamental de respeitados centros de pesquisa como *Protein Data Bank* (PDB) (Berman et al., 2000; Research Collaboratory for Structural Bioinformatics, 2003), SwissModel Repository (Expasy, 2003), National Center of Biotechnology Information (NCBI) (National Center of Biotechnology Information, 2003; Wang et al., 2000), The Brutlag Bioinformatics Group (2003), Laboratory of Molecular Biophysics – Andrej Sali Lab (The Rockefeller University, 2003) e de multinacionais da área farmacológica e biotecnológica como a Glaxo Welcome, Pfizer e Roche entre outras.

O *Sting Millennium Suite* (SMS) é um conjunto de programas desenvolvidos no Structural Bioinformatics Group (2003), da Embrapa Informática Agropecuária, para a análise das propriedades físico-químicas, interação molecular e estrutura tridimensional de proteínas. A possibilidade de combinar a capacidade de análise da estrutura protéica no SMS com algoritmos de alinhamento estrutural de proteínas permite a construção “*in silico*” de complexos virtuais proteína-inibidor e consequentemente de matrizes de contato protéico com inibidores específicos. Esta abordagem permite o desenvolvimento de descritores estruturais de especificidade enzimática, e o estudo da evolução molecular de famílias protéicas.

ISSN 1677-8464

## Utilização de Alinhamento Estrutural de Proteínas no Estudo de Interface Forming Residues de Proteína-quinases com Inibidores Protéicos

Jorge Hernandez Fernandez <sup>1</sup>

Goran Neshich <sup>2</sup>

### **Alinhamento estrutural de proteínas e análise de interação proteína-inibidor**

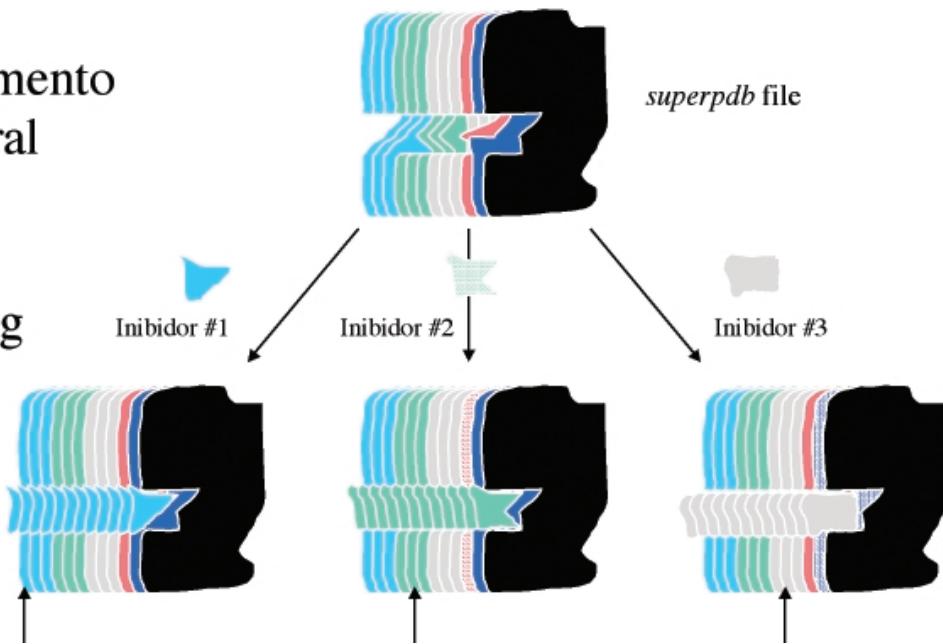
O alinhamento múltiplo de estruturas protéicas pode disponibilizar para todas as proteínas estudadas, a análise da interação protéica com certo inibidor (Fig. 1). Basta um complexo enzima-inibidor ter sido cristalizado. Este complexo é usado no *hard docking* com todas as outras estruturas alinhadas, na superposição das coordenadas do inibidor estudado com cada uma das outras proteínas em estudo. Esta superposição de coordenadas cria “*in silico*” uma grande quantidade de complexos enzima-inibidor (inclusive aqueles que não foram cristalizados ainda, e aqueles que não é possível cristalizar pela pouca afinidade Enzima-inibidor). Isto possibilita a criação de matrizes de contato de grande quantidade de proteínas com um inibidor, partindo de um único complexo proteína-inibidor cristalizado (Fig. 1). Como exemplo do possível resultado desta abordagem, pode ser utilizado a descrição da matriz de contatos de 73 SER-proteases com 3 inibidores diferentes, disponível no aplicativo SPIDER (Serine Protease Interface Data Resource, 2003) desenvolvido no Núcleo de Bioinformática Estrutural (NBI) da Embrapa Informática Agropecuária.

<sup>1</sup> Dr. em Biologia Molecular, Colaborador Científico do Núcleo de Bioinformática Estrutural da Embrapa Informática Agropecuária, Caixa Postal 6041, Barão Geraldo - 13083-970 - Campinas, SP. (e-mail: jorgehf@cnptia.embrapa.br)

<sup>2</sup> Ph.D. em Biofísica, Pesquisador da Embrapa Informática Agropecuária, Caixa Postal 6041, Barão Geraldo - 13083-970 - Campinas, SP. (e-mail: neshich@cnptia.embrapa.br)

## Alinhamento estrutural

### Docking



**Fig. 1.** Esquema demonstrativo dos passos seguidos na construção da Interface Forming Residues (IFR) de famílias protéicas. Inicialmente são alinhadas estruturalmente as estruturas pertencentes à família protéica a ser estudada, obtendo-se o *superpdb*. O *superpdb* é superposto às coordenadas de proteínas cristalizadas com inibidores no processo de *hard-docking*, e nos pdbs resultantes, são estudadas as interações proteína-inibidor. As setas mostram os complexos enzima-inibidor conhecidos antes desta abordagem, que são usados na superposição de coordenadas dos inibidores com o *superpdb*.

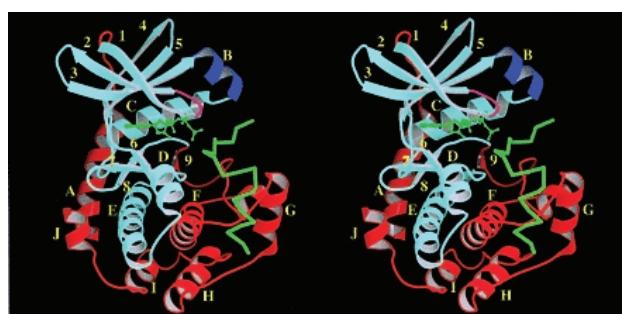
Um passo importante na construção do *Interface Forming Residues* (IFR) de uma família protéica é a construção do alinhamento estrutural de todas as proteínas a serem estudadas. O **Protein Informatics System for Modeling** (PrISM) (Yang & Honig, 1999) foi concebido inicialmente como um sistema para modelagem e análise de proteínas, e desenvolvido para o uso automático e repetitivo. Este sistema de algoritmos apresenta um poderoso módulo de alinhamento estrutural, especificamente desenvolvido para obter um ótimo resultado com seqüências de baixa identidade (~30%) (Yang & Honig, 2000a, 2000b). O algoritmo de alinhamento estrutural múltiplo é dividido em duas partes: no primeiro passo, são comparadas a distância e a orientação de elementos da estrutura secundária (SSEs) e no segundo passo, a superposição interativa e otimização do *rmsd* (*root mean square deviation*) do alinhamento global. A combinação da otimização dos SSEs e do *rmsd* no alinhamento global é calculado pelo PrISM como parâmetro **Protein Structural Distance** (PSD). Testes em diferentes bases de dados serviram para determinar o PSD = 2.0 como o ponto de separação entre diferentes grupos protéicos (Yang & Honig, 2000a). Este algoritmo mostrou-se ágil e efetivo tanto no alinhamento de estruturas similares como distantes (Yang & Honig, 2000a, 2000b) e por isto é usado no projeto como ferramenta principal na etapa inicial: alinhamento estrutural de proteínas-quinase (PK).

### As proteínas-quinases (PK) como alvo biológico

A fosforilação catalisada pelas PK é um evento chave na mediação de respostas celulares aos estímulos do ambiente. Esta mediação é feita por dois mecanismos diferentes. O primeiro envolve SER/THR quinases ativadas por mensageiros secundários (como cAMP) gerados por executores conectados via proteína G a receptores de membrana (review em Nishizuka, 1992). No segundo mecanismo os sinais extracelulares são transmitidos por receptores transmembrânicos com domínios intercelulares que apresentam atividade enzimática, como o grupo das TYR quinases (Van der Geer et al., 1994). Entre as SER-THR proteínas-quinases mais estudadas encontram-se as Map PK (Mitogen activated protein kinases) (Widmann et al., 1999; Ligterink et al., 1997; Wilson et al., 1996), CK I e II (Caseine kinase) (Niefind et al., 1998; Mashhoon et al., 2000), cAMP-PK (cAMP dependent protein kinase) (Zheng et al., 1993; Engh et al., 1996), Cd PK2 (cyclin-dependent protein kinase) (Brown et al., 1999; Maccioni et al., 2001). Entre as TYR proteínas quinases mais estudadas encontram-se a Insulin receptor tyrosine kinase (IRK) (Hubbard, 1997), e o FGFRK (fibroblasts growth factor receptor kinase) (Mohammadi et al., 1997, 1998). O interesse pelas PK vem crescendo depois de ser descoberto o importante papel que jogam estas enzimas em patologias como câncer (CK II, Battistutta et al., 2000), Alzheimer (CK I, Mashhoon et al., 2000),

angiogenese (FGFRK, Mohammadi et al., 1998), entre outras. A maioria das TYR quinases (E.C. 2.7.1.112) são consideradas genes envolvidos no desenvolvimento de câncer (BRENDA Enzyme Database) (Cologne University Bioinformatics Center, 2003) o que incentiva os trabalhos de obtenção de drogas terapêuticas. Outro campo importante de pesquisas é o envolvimento de PK no mecanismo de defesa de plantas (Romeis et al., 2001).

A estrutura do domínio catalítico destas enzimas é evolutivamente conservado embora apresentem seqüências com identidade abaixo dos 35%, dado que justifica uma abordagem estrutural geral neste grupo tão heterogêneo. O domínio catalítico das PK consiste basicamente num polipeptídio de aproximadamente 30-32 kDa (Bossemeyer et al., 1993; Mohammadi et al., 1997, Kruse et al., 1997) com estrutura  $\alpha$ - $\beta$  evolutivamente conservada (Kruse et al., 1997). Nesta estrutura estão bem definidos os sítios de contato com o co-substrato (ATP e/ou GTP) e com a proteína-substrato (Fig. 2). Para maior comodidade, o sítio de contato com o co-substrato é chamado de sítio **C** e o sítio de contato com a proteína substrato é chamado de sítio **S**.



**Fig. 2.** Representação estereométrica da estrutura do complexo cAPK com ATP e o peptídeo inibidor PKI (PDB-ID 1APM). As  $\alpha$ -hélices estão assinaladas com letras (A-J) e as folhas- $\beta$  com números (1-9) de forma incremental no sentido N-C terminal da proteína. O ATP e o peptídeo PKI estão na cor verde. Em azul-claro, as estruturas mais conservadas do domínio catalítico das PK, e em vermelho as mais variáveis. Em rosa a “glycine-rich loop”. Fonte: Steinbacher et al. (1999), adaptada pelo autor.

Nas estruturas das proteínas-quinases, as estruturas em folha- $\beta$  1-5 formam um plano- $\beta$  (o chapéu) que se sobrepõe às  $\alpha$ -hélices D-J (que formam uma espécie de barril). As  $\alpha$ -hélices A e C conectam estas duas

estruturas bem definidas. O sítio C é formado na zona de contato destas duas estruturas e compreende as folhas- $\beta$  1 e 6, e a  $\alpha$ -hélice D (Fig. 2). É nesta região que aloja-se o trinucleotídeo (ATP-GTP) doador do grupo fosfato. O “loop” de passagem da folha- $\beta$  1 à 2 conhecido como “glycine-rich loop” por conter o motivo altamente conservado Gly X Gly XX Gly (Hemmer et al., 1997). O sítio S encontra-se definido na zona de contato entre as  $\alpha$ -hélices D-G-H (Steinbacher et al., 1999) (Fig. 2). Esta zona é de estrutura mais variável pois é a responsável pela especificidade catalítica da enzima (Steinbacher et al., 1999; Lowe et al., 1997; Hubbard, 1997). O importante papel em processos celulares de sinalização, diferenciação e ciclo celular das CDPK (Engh et al., 1996), cAPK (Brown et al., 1999), FGF (Mohammadi et al., 1998) e MAPK (Widmann et al., 1999), entre outras PK, tem impulsionado a procura de inibidores para estas enzimas. No PDB podem ser encontradas 68 estruturas de proteínas-quinases cristalizadas com diferentes inibidores. A maioria são proteínas humanas, fato facilmente explicável devido à concentração de esforços no desenvolvimento de drogas de interesse terapêutico.

## Material e Métodos

Para o alinhamento estrutural da família das PK foram escolhidas 45 estruturas depositadas no PDB (Berman et al., 2000) que representam seqüências não redundantes (<95% de identidade) de proteínas classificadas como proteínas-quinases e no SCOP (Murzin et al., 1995; 2003) como *SER-THR* e *TYR* quinases. No alinhamento seqüencial múltiplo das 45 PK foi usado o ClustalW 1.8 (Thompson et al., 1994). Na análise filogenética foi usado o programa Treeconw (Van de Peer & De Wachter, 1994). O alinhamento seqüencial das PK produzido no ClustalW foi salvo em formato PHYLIP (arquivo .phy). A correção de *Poison* (Dickerson, 1971) foi usada no cálculo das distâncias. No cálculo da significância estatística da similaridade entre as seqüências foi usada a análise de “bootstrap” (Swofford et al., 1996) com 100 repetições. O método de “Neighbor-joining” (Saitou & Nei, 1987) foi usado na construção da árvore filogenética.

Na construção do alinhamento estrutural das 45 PK foi implementado um *Prism-script* (Fig. 3-A) para separar as estruturas de PK escolhidas nos domínios N-terminal e catalítico, calcular o alinhamento de todas as estruturas e agrupar estas estruturas em alinhamentos de proteínas com estrutura semelhante (*P.S.D.- cutoff* = 2.0, Yang & Honig, 2000a).

**A:**

```

main>>yes
main>>load_structure
ldst>>PKselected
.
.## Loading structural
.## Data
.
ldst>>/pdb/name.pdb
ldst>>name
ldst>>chain
ldst>>do_cut
ldst>>save
.
.## Repeat for all
.## structural data
.
ldst>>done
main>>alignment
ali2>>structure_alignment
ali2>>data
ali2>>yes
ali2>>rmscut = 100
ali2>>scocut = 0
ali2>>done
ali2>>load
ali2>>batch/PKselected.dat
ali2>>@all_records
ali2>>batch/null.dat
ali2>>done
ali2>*>** main>>exit
prof>>no
prof>>./3rd/ali2nd.log
prof>>done

```

**B:**

```

Group:3
CHEK:: creating superimpositions....
@ 103ljwh A114 337 :@ 1101j91 A115 333 e-pairs= 219 rmsd= 0.714
@ 521k2p A538 654 :@ 751jpa A771 900 e-pairs= 98 rmsd= 0.876
@ 882ptk 403 528 :@ 931ad5 A403 531 e-pairs= 109 rmsd= 0.952
@ 1081qpd A318 501 :@ 751jpa A771 900 e-pairs= 102 rmsd= 1.012
@ 641m14 A768 996 :@ 521k2p A538 654 e-pairs= 116 rmsd= 1.100
@ 832src 339 533 :@ 882ptk 403 528 e-pairs= 108 rmsd= 1.201
@ 771fvr A903 121 :@ 521k2p A538 654 e-pairs= 112 rmsd= 1.192
@ 882ptk 403 528 :@ 521k2p A538 654 e-pairs= 101 rmsd= 1.188
@ 431m52 A317 502 :@ 521k2p A538 654 e-pairs= 116 rmsd= 1.277
@ 521k2p A538 654 :@ 461byg A331 448 e-pairs= 100 rmsd= 1.357
@ 791fgi A562 762 :@ 461byg A331 448 e-pairs= 105 rmsd= 1.635
@ 671h41 A143 287 :@ 381hf8 B199 384 e-pairs= 141 rmsd= 1.892
@ 1101j91 A115 333 :@ 671h41 A143 287 e-pairs= 141 rmsd= 2.059
@ 751jpa A771 900 :@ 671h41 A143 287 e-pairs= 98 rmsd= 2.588
@ 1061ias A367 500 :@ 521k2p A538 654 e-pairs= 103 rmsd= 2.629
@ 721how A247 737 :@ 671h41 A143 287 e-pairs= 104 rmsd= 2.392
@ 671h41 A143 287 :@ 261ia8 A147 276 e-pairs= 81 rmsd= 2.686

Group:4
CHEK:: creating superimpositions....
@ 211f3m C406 542 :@ 111a06 96 316 e-pairs= 80 rmsd= 1.646
@ 991luf A656 860 :@ 951irk 078 283 e-pairs= 179 rmsd= 1.368
@ 291koia 017 266 :@ 211f3m C406 542 e-pairs= 80 rmsd= 1.660
@ 951irk 078 283 :@ 211f3m C406 542 e-pairs= 81 rmsd= 1.731
@ 211f3m C406 542 :@ 1011jkh A081 286 e-pairs= 81 rmsd= 1.714
@ 291koia 017 266 :@ 351tki A101 338 e-pairs= 238 rmsd= 2.004
@ 991luf A656 860 :@ 701vr2 A045 168 e-pairs= 78 rmsd= 1.242
@ 311p38 108 354 :@ 211f3m C406 542 e-pairs= 83 rmsd= 3.445

```

**Fig. 3.** Separação das 45 PK segundo similaridade das suas estruturas do domínio catalítico usando o PrISM. A: *Prism-script* usado na separação das estruturas em alinhamentos estruturais de domínios catalíticos semelhantes. O script carrega todas as estruturas a serem estudadas, alinha estruturalmente e separa as estruturas alinhadas em grupos semelhantes, conforme parâmetros pré-estabelecidos. B: Resultados obtidos no alinhamento estrutural de PK nos grupos 3 e 4. Os diferentes grupos de proteínas foram alinhados usando como parâmetro de similaridade o *PSD-cutoff.= 2.0*.

## Resultados e Discussão

### Três grupos de estruturas diferentes no domínio catalítico de PK

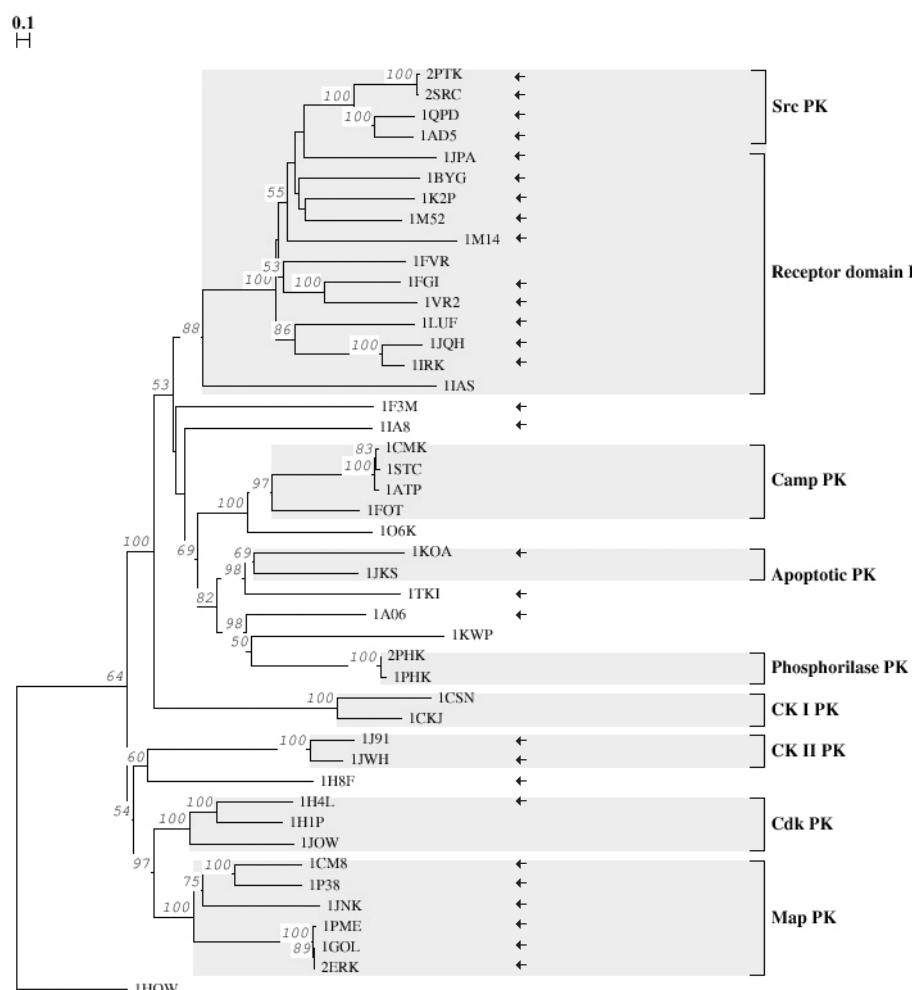
Como produto do *Prism-script* utilizado, as 45 estruturas de PK foram separadas em 5 grupos diferentes (Fig. 3-B). No Grupo 1 foi obtido o alinhamento dos 85-100 a.a do N-terminal das 45 PK. Este resultado era esperado pois a estrutura deste domínio N-terminal é evolutivamente muito conservada. Já nos Grupos 3, 4 e 5 foram separados os alinhamentos dos domínios catalíticos de 16, 9 e 5 PKs diferentes (Fig. 3B e 4). Além do cálculo do *rmsd* em cada um dos alinhamentos, o PrISM fornece os arquivos no formato PDB de todas as estruturas alinhadas em cada grupo. Estes arquivos, chamados de *superpdb* serão utilizados na superposição das coordenadas dos complexos PK-inibidor usados no processo de cálculo das IFR de PK.

### Superposição dos dados obtidos no alinhamento seqüencial e estrutural

Os resultados do alinhamento seqüencial e análise filogenética das 45 PKs podem ser observados na Fig. 4. Os dados obtidos anteriormente no alinhamento estrutural foram comparados com a filogenia proposta pelo alinhamento seqüencial e observou-se uma superposição não ordenada de PKs de grupos diferentes. Embora a filogenia calculada pelo alinhamento seqüencial demonstre grande coerência com os grupos funcionais das proteínas estudadas, os resultados obtidos no alinhamento estrutural indicam uma relação estrutura-função diferente. Estes dados aparentemente contraditórios podem estar indicando que os aminoácidos que interagem com o substrato influenciam mais na especificidade enzimática do que a estrutura do domínio catalítico da enzima.

Este resultado pode variar conforme a família protéica, e será confirmado com o estudo dos aminoácidos que compõem o IFR das PKs. Os

resultados obtidos serão comparados aos já disponíveis no SPIDER (Serine Protease Interface Data Resource, 2003).



**Fig. 4.** Árvore filogenética construída no Treecrown com resultados do alinhamento seqüencial de 45 estruturas de proteína-quinase escolhidas no PDB. Na cor cinza estão assinaladas as famílias de proteína-quinases mais encontradas na literatura atual. As setas marcam as estruturas agrupadas pelo PrISM no Grupo 3 (↔), Grupo 4 (↔) e Grupo 5 (↔), usando PSD = 2.0 no alinhamento estrutural destas quinases. Em cada grupo encontra-se o domínio catalítico das quinases alinhado estruturalmente. O alinhamento estrutural dos 80-100 a.a. no N-terminal das 45 quinases foi agrupado no Grupo 1 (não mostrado), indicando que esta parte da estrutura é evolutivamente bem conservada.

## Conclusões

O algoritmo do programa PrISM (Yang & Honig, 1999) mostrou-se suficientemente eficiente no alinhamento e separação por grupos homólogos de proteínas com níveis de ID<30%. Esta separação semi-automática de estruturas semelhantes permite um melhor posicionamento dos inibidores na transferência de coordenadas estruturais no processo de *hard docking*. Como resultado, o processo de alinhamento estrutural de proteínas de uma mesma família acaba aumentando a confiabilidade dos resultados obtidos na descrição do IFR destas proteínas, e facilitando a descrição de parâmetros estruturais de especificidade enzimática.

## Referências Bibliográficas

- BATTISTUTTA, R.; SARNO, S.; DE MOLINER, E.; MARIN, O.; ISSINGER, O.-G.; ZANOTTI, G.; PINNA, L. A. The crystal structure of the complex of *Zea mays* {alpha} subunit with a fragment of human {beta} subunit provides the clue to the architecture of protein kinase CK2 holoenzyme. **Eur. J. Biochem.**, v. 267, p. 5184-5190, 2000.
- BERMAN, H. M.; BHAT, T. N.; BOURNE, P. E.; FENG, Z.; GILLILAND, G.; WEISSIG, H.; WESTBROOK, J. The Protein Data Bank and the challenge of structural genomics. **Nat. Struct. Biol.**, v. 7, p. 957-9, 2000.

- BOSSEMEYER, D.; ENGH, R. A.; KINZEL, V.; PONSTINGL, H.; HUBER, R. Phosphotransferase and substrate binding mechanism of the cAMP-dependent - protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn<sup>2+</sup> adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). **EMBO J.**, v. 12, p. 849-859, 1993.
- BROWN, N. R.; NOBLE, M. E. M.; LAWRIE, A. M.; MORRIS, M. C.; TUNNAH, P.; DIVITA, G.; JOHNSON, L. N.; ENDICOTT, J. A. Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity. **J. Biol. Chem.**, v. 274, p. 8746-8756, 1999.
- THE BRUTLAG BIOINFORMATICS GROUP. **The Brutlag Bioinformatics Group [home page]**. Disponível em: <<http://dna.stanford.edu/>>. Acesso em: fev. 2003.
- COLOGNE UNIVERSITY BIOINFORMATICS CENTER. **Enzyme Database – BRENDA**: the comprehensive enzyme information system. Disponível em: <<http://www.brenda.uni-koeln.de/>>. Acesso em: fev. 2003.
- DICKERSON, R. E. The structure of cytochrome c and the rates of molecular evolution. **J. Mol. Evol.**, v. 1, p. 26-45, 1971.
- ENGH, R. A.; GIROD, A.; KINZEL, V.; HUBER, R.; BOSSEMYER, D. Crystal structures of catalytic subunit of cAMP-dependent protein kinase in complex with isoquinolinesulfonyl protein kinase inhibitors H7, H8, and H89. Structural implications for selectivity. **J. Biol. Chem.**, v. 271, n. 42, p. 26157-26164, 1996.
- EXPASY. **SWISS-MODEL repository**. Disponível em: <[http://www.expasy.org/swissmod/SM\\_3DCrunch\\_Search.html](http://www.expasy.org/swissmod/SM_3DCrunch_Search.html)>. Acesso em: fev. 2003.
- HEMMER, W.; MCGLONE, M.; TSIGELNY, I.; TAYLOR, S. S. Role of the Glycine triad in the ATP-binding site of cAMP-dependent protein kinase. **J. Biol. Chem.**, v. 272, p. 16946-16954, 1997.
- HUBBARD, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. **EMBO J.**, v. 16, p. 5572-5581, 1997.
- KRUSE, M.; MULLER, I. M.; MULLER, W. E. Early evolution of metazoan serine/threonine and tyrosine kinases: identification of selected kinases in marine sponges. **Mol. Biol. Evol.**, v. 12, p. 1326-1334, 1997.
- LIGTERINK, W.; KROJ, T.; ZUR NIEDEN, U.; HIRT, H.; SCHEEL, D. Receptor-mediated activation of a MAP kinase in pathogen defense of plants. **Science**, v. 276, p. 2054-2057, 1997.
- LOWE, E. D.; NOBLE, M. E. M.; SKAMNAKI, V. T.; OIKONOMAKOS, N. G.; OWEN, D. J.; JOHNSON, L. N. The crystal structure of a phosphorylase kinase peptide substrate complex: kinase substrate recognition. **EMBO J.**, v. 16, p. 6646-6658, 1997.
- MACCIONI, R. B.; OTTH, C.; CONCHA, I. I.; MUÑOZ, J. P. The protein kinase Cdk5: Structural aspects, roles in neurogenesis and involvement in Alzheimer's pathology. **Eur. J. Biochem.**, v. 268, p. 1518-1527, 2001.
- MASHHOON, N.; DEMAGGIO, A. J.; TERESHKO, V.; BERGMAYER, S. C.; EGLI, M.; HOEKSTRA, M. F.; KURET, J. Crystal structure of a conformation-selective casein kinase-1 inhibitor. **J. Biol. Chem.**, v. 275, p. 20052-20060, 2000.
- MOHAMMADI, M.; FROUM, S.; HAMBY, J. M.; SCHROEDER, M. C.; PANEK, R. L.; LU, G. H.; ELISEENKOVA, A. V.; GREEN, D.; SCHLESSINGER, J.; HUBBARD, S. R. Crystal structure of an angiogenesis inhibitor bound to the Fgf receptor tyrosine kinase domain. **EMBO J.**, v. 17, p. 5896, 1998.
- MOHAMMADI, M.; MCMAHON, G.; SUN, L.; TANG, C.; HIRTH, P.; YEH, B. K.; HUBBARD, S. R. SCHLESSINGER, J. Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. **Science**, v. 276, p. 955, 1997.
- MURZIN, A. G.; BRENNER, S. E.; HUBBARD, T.; CHOTHIA, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. **J. Mol. Biol.**, v. 247, p. 536-540, 1995.
- MURZIN, A. G.; CONTE, L.; AILEY, B. G.; BRENNER, S. E.; HUBBARD, T. J. P.; CHOTHIA, C. **SCOP - structural classification of proteins**. Disponível em: <<http://scop.mrc-lmb.cam.ac.uk/scop/index.html>>. Acesso em: fev. 2003.
- NATIONAL CENTER OF BIOTECHNOLOGY INFORMATION (Estados Unidos). **Entrez-Structure**. Disponível em: <<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Structure>>. Acesso em: fev. 2003.
- NIEFIND, K.; GUERRA, B.; PINNA, L. A.; ISSINGER, O. G.; SCHOMBURG, D. Crystal structure of the catalytic subunit of protein kinase CK2 from *Zea mays* at 2.1 Å resolution. **EMBO J.**, v. 17, p. 2451, 1998.
- NISHIZUKA, Y. Intracellular signaling by hydrolysis of phospholipids and activation of protein kinase C. **Science**, v. 258, p. 607-614, 1992.
- RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS. Protein Data Bank. Disponível em: <<http://www.rcsb.org/>>. Acesso em: fev. 2003.
- THE ROCKEFELLER UNIVERSITY. **Laboratory of Molecular Biophysics - Andrej Sali Lab**. Disponível em: <[http://www.salilab.org/our\\_resources.shtml](http://www.salilab.org/our_resources.shtml)>. Acesso em: fev. 2003.
- ROMEIS, T.; LUDWIG, A. A.; MARTIN, R.; JONES, J. D. Calcium-dependent protein kinases play an essential role in a plant defense response. **EMBO J.**, v. 20, n. 20, p. 5556-5567, 2001.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.**, v. 4, p. 406-425, 1987.

SERINE Protease Interface Data Resource. Disponível em: <<http://www.cbi.cnptia.embrapa.br/SMS/SPIDER/>>. Acesso em: fev. 2003.

STEINBACHER, S.; HOF, P.; EICHINGER, L.; SCHLEICHER, M.; GETTEMANS, J.; VANDEKERCKHOVE, J.; HUBER, R.; BENZ, J. The crystal structure of the *Physarum polycephalum* actin-fragmin kinase: an atypical protein kinase with a specialized substrate-binding domain. **EMBO J.**, v. 18, p. 2923-2929, 1999.

STRUCTURAL BIOINFORMATICS GROUP. **Sting Millennium Suite**. Disponível em: <<http://www.ncbi.cnptia.embrapa.br/>>. Acesso em: fev. 2003.

SWOFFORD, D. L.; OLSEN, G. J.; WADDELL, P. J.; HILLIS, D. M. Phylogenetic inference. In: HILLIS, D. M.; MORITZ, C.; MABLE, B. K. (Ed.). **Molecular systematics**. 2<sup>nd</sup> ed. Sunderland: Sinauer Associates, 1996. p. 407-514.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p. 4673-4680, 1994.

VAN DE PEER, Y.; DE WACHTER, R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. **Computer Applications in the Biosciences**, v. 10, n. 5, p. 569-570, 1994.

VAN DER GEER, P.; HUNTER, T.; LINDBERG, R. A. Receptor protein-tyrosine kinases and their signal transduction pathways. **Annu. Rev. Cell. Biol.**, v. 10, p. 251-337, 1994.

WANG, Y.; BRYANT, S.; TATUSOV, R.; TATUSOVA, T. Links from genome proteins to known 3-D structures. **Genome Res.**, v. 10, n. 10, p. 1643-1647, 2000.

WIDMANN, C.; GIGSON, S.; JARPE, M. B.; JOHNSON, G. L. Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. **Physiol Rev.**, v. 79, p. 143-180, 1999.

WILSON, K. P.; FITZGIBBON, M. J.; CARON, P. R.; GRIFFITH, J. P.; CHEN, W.; MCCAFFREY, P. G.; CHAMBERS, S. P.; SU, M. S.-S. Crystal structure of p38 mitogen-activated protein kinase. **J. Biol. Chem.**, v. 27, p. 27696-27700, 1996.

YANG, A.-S.; HONIG, B. Sequence to structure alignment in comparative modeling using PRISM. **Proteins: Struc., Func. and Genet.**, v. 3, p. 66-72, 1999.

YANG, A.-S.; HONIG, B. An integrated approach to the analysis of sequence and structure. I. Protein structural alignment and a quantitative measure for protein structural distance. **J. Mol. Biol.**, v. 301, p. 665-678, 2000a.

YANG, A.-S.; HONIG, B. An integrated approach to the analysis of sequence and structure. III. A comparative study of sequence conservation in protein structural families using multiple structural alignments. **J. Mol. Biol.**, v. 301, p. 691-711, 2000b.

ZHENG, J. H.; TRAFNY, E. A.; KNIGHTON, D. R.; XUONG, N. H.; TAYLOR, S. S.; TENEYCK, L. F.; SOWADSKI, J. M. 2.2-Angstrom refined crystal-structure of the catalytic subunit of cAMP-dependent protein-kinase complexed with MNATP and a peptide inhibitor. **Acta Crystallogr. D Biol. Crystallogr.**, v. 49, p. 362, 1993.

**Comunicado Técnico, 44** Embrapa Informática Agropecuária  
Área de Comunicação e Negócios (ACN)  
Av. André Tosello, 209  
Cidade Universitária - "Zeferino Vaz"  
Barão Geraldo - Caixa Postal 6041  
13083-970 - Campinas, SP  
Telefone (19) 3789-5743 - Fax (19) 3289-9594  
e-mail: sac@cnptia.embrapa.br

1<sup>a</sup> edição  
2002 - on-line  
Todos os direitos reservados

**Comitê de Publicações** Presidente: José Ruy Porto de Carvalho  
Membros efetivos: Amarindo Fausto Soares, Ivanilde Dispato, Luciana Alvim Santos Romani, Marcia Izabel Fugisawa Souza, Suzilei Almeida Carneiro  
Suplentes: Adriana Delfino dos Santos, Fábio Cesar da Silva, João Francisco Gonçalves Antunes, Maria Angélica de Andrade Leite, Moacir Pedroso Júnior

**Expediente** Supervisor editorial: Ivanilde Dispato  
Normalização bibliográfica: Marcia Izabel Fugisawa Souza  
Capa: Intermídia Publicações Científicas  
Editoração Eletrônica: Intermídia Publicações Científicas