

CALIXTO, T. M. R. **Análises de propriedades eletrostáticas e estruturais de complexos de proteínas para o desenvolvimento de preditores de complexação em larga escala.** 2010. 228f. Dissertação (Mestrado) - Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2010.

ERRATA

| Página | Linha | Onde se lê | Leia-se |
|--------|-------|--|--|
| xxvii | 15 | NT Nozaki e Tanford. PDB Protein data bank (banco de dados de proteínas). | NT Nozaki e Tanford. PB Poisson-Boltzmann. PDB Protein data bank (banco de dados de proteínas). |
| xxii | 11 | IN SILÍCIO | IN SILICO |
| xxiv | 17 | D.2 Arquivos no formato <i>st</i> criados com base no campo de força AMBER99 198 | D.2 Arquivos no formato <i>st</i> criados com base no campo de força AMBER99 198 Material suplementar disponível em: http://glu.fcfrp.usp.br/tulio/supp |
| 2 | 4 | inimaginário | inimaginável |
| 2 | 22 | DNA (50), nucleotídeos (51) | DNA (50, 51) |
| 4 | 2 | identificar. | identificar (208). |
| 4 | 15 | <i>in silício</i> | <i>in silico</i> |
| 7 | 11 | de forma combinatória | combinando diferentes quantidades de aminoácidos e fazendo mutações entre os resíduos |
| 9 | 24 | A situação é equivalente a se conhecer os “ingredientes” (eletrostática, <i>van der Waals</i>, hidrofobicidade, etc.) para o preparo de um delicioso bolo, porém, não se tem a quantidade necessária de cada um deles (falta a “receita!”). | A situação é equivalente a se conhecer os elementos constituintes do processo de complexação, porém, não é possível quantificar particularmente cada elemento diretamente, ou seja, obter a <i>Hamiltoniana</i> exata do sistema. |

| | | | |
|----|----|---|---|
| 10 | 3 | . Em outras palavras: procura-se também pela “receita” (27). | , visando desvendar as interações físicas e manipular racionalmente os eventos biológicos. |
| 12 | 16 | Porém, o estado protonado de uma proteína não depende somente do pH, mas também de moléculas vizinhas; o potencial eletrostático de moléculas vizinhas perturbam os grupos ionizáveis presentes na proteína, o mesmo efeito também é observado internamente na proteína (46, 106, 120). | Porém, o estado protonado de uma proteína não depende somente do pH, mas também da influência de moléculas vizinhas. O potencial eletrostático gerado por moléculas vizinhas carregadas afeta o equilíbrio ácido-base dos grupos ionizáveis (46, 106, 120). |
| 12 | 24 | <i>docking</i> | predição do sítio ativo da proteína e afinidade com o ligante (<i>docking</i>) |
| 13 | 17 | Uma das características dos potenciais estatísticos é avaliar a densidade de uma determinada distribuição e, assim, propor funções para descrever a energia de contato entre os pares, comumente expressa através de potencial de força média. | A partir dos potenciais estatísticos e com o auxílio dos métodos inversos, podemos determinar funções que descrevem e quantifiquem as energias de interações do complexo, comumente expressa através de potencial de força média. |
| 14 | 29 | Esses portais são caracterizados por sua simplicidade de uso, rápido retorno ao usuário e com resultados satisfatórios. | Esses portais são caracterizados por sua simplicidade de uso e rápido retorno ao usuário, apresentando uma estimativa de comportamento do sistema. |
| 15 | 2 | Elementos em cinza indicam as ferramentas que serão implementadas no futuro. | O nível de detalhamento do sistema, precisão e custo computacional aumenta conforme passamos do nível 0 para o 2. O arquivo PQR contém a carga e o raio de cada átomo presente na proteína, o qual é utilizado em simulações PB (MEAD) e MC. Elementos em cinza indicam as ferramentas que serão implementadas no futuro. |

| | | | |
|----|----|---|--|
| 18 | 11 | predição complexos | predição de complexos |
| 18 | 19 | Compreensão | Melhorar a compreensão |
| 20 | 11 | Todas as proteínas | As proteínas |
| 25 | 3 | <i>IN SILÍCIO</i> | <i>IN SILICO</i> |
| 25 | 22 | Os átomos da biomolécula e da solução eletrolítica são as coordenadas do núcleo e dos elétrons de cada componente do sistema. | Os átomos da biomolécula e da solução eletrolítica são representados pelas coordenadas do núcleo e dos elétrons de cada um dos componentes do sistema. |
| 26 | 8 | isto é, as variáveis são somente as coordenadas e momentos da molécula do soluto. O soluto participa apenas com a média de suas coordenadas e momentos. | isto é, as variáveis tratadas explicitamente são somente as coordenadas e momentos da molécula do soluto. Apenas o comportamento médio do solvente é considerado. |
| 27 | 23 | $\nabla \cdot [\varepsilon(r) \nabla \phi(r)] = -4\pi\rho(r)$ <p>A equação de <i>Poisson</i> permite relacionar a variação espacial do potencial ϕ na posição r com a distribuição da densidade de carga $\rho(r)$ em um meio de constante dielétrica ε. A carga parcial de cada átomo na molécula pode ser descrita como cargas fixas, enquanto que as cargas dos eletrólitos dissolvidos no solvente podem ser descritas como cargas móveis, cuja densidade é determinada pela distribuição de <i>Boltzmann</i>:</p> | $\nabla \cdot [\varepsilon(r) \nabla \phi(r)] = -4\pi\rho(r)$ <p>Supondo $\rho(r)$ a distribuição de cargas em r e $\phi(r)$ o potencial elétrico na posição r provido por cargas fixas e móveis, haverá uma maior concentração de íons negativos em r quando $\phi(r)$ for positivo e vice-versa (167). A concentração de íons móveis (eletrólitos dissolvidos em solução) pode ser obtida pela distribuição de <i>Boltzmann</i> (139, 167, 212), definida na Equação 4.2, e as cargas fixas são as cargas parciais de cada átomo que constitui a molécula.</p> |

| | | | |
|----|----|---|---|
| 28 | 6 | $\nabla \cdot \varepsilon(r) \nabla \phi(r) =$ $-4 \pi \{ \rho(r) + \lambda(r) \sum_i q_i n_i \exp[\frac{-q_i \phi(r)}{k_B T}] \}$ onde λ é igual 1 em regiões acessíveis aos íons móveis e 0 nas demais regiões. | $\nabla \cdot [\varepsilon(r) \nabla \phi(r)] =$ $-4 \pi \{ \rho(r) + \lambda(r) \sum_i q_i n_i \exp[\frac{-q_i \phi(r)}{k_B T}] \}$ assumimos λ igual a 0 em regiões impenetráveis e 1 nas regiões acessíveis aos íons móveis. |
| 28 | 10 | linearizada (EPBL) | linearizada (EPBL)(139, 143, 167) |
| 28 | 11 | $-\nabla \cdot \varepsilon \nabla \phi(r) + \varepsilon k^2 \phi = \rho$ | $\nabla \cdot [\varepsilon(r) \nabla \phi(r)] - k^2 \varepsilon(r) \phi(r) = -4 \pi \rho(r)$ |
| 28 | 13 | $k^2 = \frac{2Ie^2}{k_B T \varepsilon}$ | $k^2 = \frac{8 \pi e^2 N_A I}{k_B T \varepsilon}$ |
| 29 | 20 | adotando uma política de focagem | focando em uma região de interesse |
| 31 | 10 | Pode ser vista com uma | Pode ser vista como uma |
| 39 | 8 | expressões: | expressões (105): |
| 39 | 19 | Para o cálculo da variação da energia livre eletrostática com o mecanismo de regulação de cargas, utilizamos a Equação (5): $\Delta G_{cap}^{ele} = \frac{l_B Z_i Z_j}{r} - \frac{l_B^2}{2r^2} (C_i C_j + C_i Z_j^2 + C_j Z_i^2)$ | onde Z_i e Z_j são as cargas das proteínas i e j, ε_0, a constante dielétrica do vácuo, ε_s, a constante dielétrica do solvente ⁷, r, a distância de separação (em Ångströms) entre o centro geométrico das duas proteínas, K_B, a constante de <i>Boltzmann</i> e T, a temperatura em Kelvin. |
| | | A Equação 5.12 exibe o cálculo da variação da energia livre eletrostática com a incorporação da força iônica do meio. Note que, a blindagem eletrostática provocada pela força iônica afeta termo puramente eletrostático e o mecanismo de regulação de cargas de maneiras diferentes (106): | Para o cálculo da variação da energia livre eletrostática com o mecanismo de regulação de cargas (189), utilizamos a Equação (5): $\Delta G_{cap}^{ele} = \frac{l_B Z_i Z_j}{r} - \frac{l_B^2}{2r^2} (C_i C_j + C_i Z_j^2 + C_j Z_i^2) \quad (5.11)$ onde C_i e C_j são as capacitâncias das proteínas i e j e l_B é o comprimento de <i>Bjerrum</i> (5): $l_B = \frac{e^2}{4 \pi \varepsilon_0 \varepsilon_s k_B T}$ (5.12) A Equação 5.13 exibe o cálculo da variação da energia livre eletrostática com a incorporação da força iônica do meio. |

$$\Delta G_{cap}^{ele} = \frac{l_B Z_i Z_j \exp(-kr)}{r} - \frac{l_B^2 \exp(-2kr)}{2r^2} (C_i C_j + C_i Z_j^2 + C_j Z_i^2) \quad (5.12)$$

onde l_B é o comprimento de Bjerrum (5): $l_B = \frac{e^2}{4\pi\epsilon_0\epsilon_s k_B T}$ (5.13) Z_i , Z_j e C_i , C_j são, respectivamente, as cargas e capacitâncias das proteínas i e j , ϵ_0 é a constante dielétrica do vácuo, ϵ_s é a constante dielétrica do solvente⁷ e r e a distância de separação (em Ångströms) entre o centro geométrico das duas proteínas.

(em g/mol).

referências (15, 18, 98).

$$\Delta G_{cap}^{ele} = \frac{l_B Z_i Z_j \exp(-kr)}{r} - \frac{l_B^2 \exp(-2kr)}{2r^2} (C_i C_j + C_i Z_j^2 + C_j Z_i^2) \quad (5.13)$$

onde k é o inverso do comprimento de Debye. Note que, a blindagem eletrostática provocada pela força iônica afeta o termo puramente eletrostático e o mecanismo de regulação de cargas de maneiras diferentes (106).

(em Å³).

referências (15, 18, 98). Outras abordagens podem ser utilizadas para prever a complexação a partir da seqüência primária das proteínas (209-211).

Ferramentas computacionais desenvolvidas

assim os modelos propostos são melhor aplicados em complexos protéicos para os quais as interações eletrostáticas são as predominantes.

de peptídeos com uma seqüência de resíduos do tipo a em diferentes concentrações de sal

simulação MC (5)

referência (13) em força iônica igual a 0,1M de cloreto de potássio (KCl) a 25°C.

experimentais (99, 168)

[Ref. (13)]

[Ref. (14)]

[Ref. (168)]

[Ref.(168)]

[Ref. (168)]

42 2

42 24

48 24

50 4

95 19

97 12

97 16

98 18

101 2

101 3

101 4

101 5

101 6

Resultados

limitando sua aplicabilidade a qualquer sistema real.

de uma seqüência de peptídeos da mesma espécie a e em diferentes concentrações de sal

simulação MC

referência (13).

experimentais

(13)

(14)

(168)

(168)

(168)

| | | | |
|-----|----|--|---|
| 101 | 7 | (168) | [Ref. (168)] |
| 101 | 8 | (168) | [Ref. (168)] |
| 101 | 9 | (168) | [Ref. (168)] |
| 108 | 5 | Juffer ¹⁹ | Juffer ²⁰ |
| 109 | 4 | Kesvatera (MC) ²⁰ | Kesvatera (MC) ²¹ |
| 121 | 3 | (13) | [Ref. (13)] |
| 121 | 4 | (14) | [Ref. (14)] |
| 121 | 5 | (168) | [Ref. (168)] |
| 121 | 6 | (168) | [Ref. (168)] |
| 123 | 11 | moléculas | moléculas (148) |
| 124 | 5 | proteico | protéico |
| 131 | 13 | Coulombicas | Coulombianas |
| 136 | 8 | <i>pI</i> das proteínas. | <i>pI</i> das proteínas. Predições de outros complexos protéicos podem ser vistas em material complementar disponível em: http://glu.fcfrp.usp.br/proteinInteraction.pdf |
| 139 | 10 | com base na literatura | com base em trabalhos que seguem esta mesma abordagem |
| 139 | 12 | O conjunto controle foi selecionado para ser o de referência para esses cálculos. Tal conjunto é constituído por complexos de proteínas que foram resolvidos por cristalografia de raios X com resolução de 2,5 Å ou melhor. | O conjunto controle foi selecionado para ser o de referência para esses cálculos, pois tal conjunto é constituído por complexos de proteínas que foram resolvidos por cristalografia de raios X com resolução de 2,5 Å ou melhor, proporcionando mais confiabilidade na determinação da estrutura protéica. |
| 140 | 1 | Conjunto controle | Conjunto referência |
| 147 | 16 | $X_{Total}^2 = \sum_{i=1; j=1}^{i=20; j=20} X_{ij}^2$ | $X_{Total}^2 = \sum_{i=1}^{i=20} \sum_{j=1}^{j=20} X_{ij}^2$ |
| 166 | 13 | seqüência primária | seqüência primária ¹ |
| 166 | 28 | | ¹ Em sistemas onde as interações eletrostáticas são as predominantes, modelos simplificados com base na seqüência primária das proteínas apresentam |

resultados tão bons quanto outros mais complexos como, por exemplo, as simulações computacionais (PB e MC) que utilizam a estrutura 3D da proteína. Além disso, a quantidade de seqüências primárias disponíveis é muito maior do que as estruturas 3D (veja em <http://www.ncbi.nlm.nih.gov/>), destacando assim a importância em se ter um preditor inicial de complexos protéicos com base apenas na seqüência primária das proteínas.

| | | | |
|-----|----|---|---|
| 169 | 5 | Monte carlo move for proton titration of proteins and macromolecules in salt solutions. | Fast Proton Titration Scheme for Multi-scale Modeling of Protein Solutions. <i>Journal of chemical theory and computation</i> , 6, p. 3259–3266, AUG 13 2010. |
| 170 | 21 | <i>Computer Sciences</i> , AMBER CHEMICAL SOC, | <i>Computer Sciences</i> , |
| 172 | 1 | NESHICH, G.; CALIRI, A.; DA SILVA, F. L. B. Biocomputacao estrutural: dos sinais em sequencia primária de proteías ao folding e a função biológica. Projeto de pesquisa, Universidade de São Paulo, 2003. | REDFERN, OLIVER C. et al. Exploring the structure and function paradigm. <i>Current Opinion in Structural Biology</i> , 18, n. 3, p. 394-402, JUN 2008. |
| 173 | 34 | proteinprotein | protein-protein |
| 174 | 30 | Nature, Nature Publishing Group, v. 426, | Nature, v. 426, |
| 174 | 32 | Nature, Nature Publishing Group, v. 426, | Nature, v. 426, |
| 175 | 25 | <i>B-Biological Sciences</i> , ROYAL SOC LONDON, 300, | <i>B-Biological Sciences</i> , |
| 178 | 26 | JNSSON | JÖNSSON |
| 181 | 31 | <i>Proteins-Structure Function And Bioinformatics</i> , 65, n. 2, p. 424–437, NOV 1 2006. | <i>Proteins-Structure Function And Bioinformatics</i> , 65, n. 2, p. 424–437, NOV 1 2006. |

208 SACCENTI, E.; ROSATO, A. The war of tools: how can NMR spectroscopists detect errors in their structures?. *Nature*, 40, n. 4, p. 251-261 MAR 5 2008.

209 PAN, X. Y. et al.; Large-Scale Prediction of Human Protein-Protein Interactions from Amino Acid Sequence Based on Latent Topic Features. *Journal of Proteome Research*, 9, p. 4992 – 5001, SEP 8 2010.

210 YU, C. Y. et al.; Predicting protein-protein interactions in unbalanced data using the primary structure of proteins. *BMC Bioinformatics*, 11, p. 167-177, 2010.

211 CHEN, P.; Li, J. Sequence-based identification of interface residues by an integrative profile combining hydrophobic and evolutionary information. *BMC Bioinformatics*, 11, p. 402, 2010.

212 SOARES, T. A. Aplicação da Equação de Poisson-Boltzmann ao cálculo de propriedades dependentes do pH em proteínas. *Química Nova*, 27, p.640-647, 2004.