

1 Introdução

Hello (Kendrew et al., [1958](#))

2 Revisão de literatura

Hello (Kendrew et al., 1958)

2.1 Métodos de atribuição de estruturas secundárias

2.1.1 DSSP

As ligações de hidrogênio são definidas utilizando um modelo eletrostático. Por esse modelo, uma ligação de hidrogênio HB ocorrerá se, e somente se, a energia E for menor que -0.5 kcal/mol. Para o cálculo são utilizadas as cargas parciais $+q_1, -q_1$ nos átomos C e O, e $-q_2, +q_2$ nos átomos N e H, onde $q_1 = 0.42e$ e $q_2 = 0.20e$.

$$E < -0.5 \text{ kcal/mol} \implies \text{HB} = \text{Verdade} \quad (2.1)$$

onde

$$E = q_1 q_2 (1/r(ON) + 1/r(CH) - 1/r(OH) - 1/r(CN)) * f \quad (2.2)$$

Na equação (2.2), $r(AB)$ é a distância interatômica entre A e B em ângstroms e o fator dimensional $f = 332$.

Os autores afirmam que, por este modelo, uma boa ligação de hidrogênio teria aproximadamente -3 kcal/mol. Assim, a escolha de um limiar em -0.5 kcal/mol torna o modelo mais tolerante à erros nas coordenadas atômicas e à ligações de hidrogênios bifurcadas (Kabsch and Sander, 1983).

2.1.2 Stride

stride

2.1.3 KAKSI

kaksi

2.1.4 PROSS

pross

2.2 Métodos de predição de estruturas secundárias

2.2.1 Primeira geração (1957-1978)

primeira geração - chou e fasman - gor

2.2.2 Segunda geração (1983-1992)

segunda geração - GORIII

2.2.3 Terceira geração

terceira geração - aprendizado de máquina (NN) - PHDsec

2.2.4 Quarta geração

quarta geração - aprendizado de máquina (NN) + dados evolutivos (PSSM) - PSIPred

3 Resultados

3.1 Análise de métodos de atribuição de estrutura secundária

Diferentemente de vários métodos publicados de predição de estrutura secundária que utilizaram os resultados de atribuição provenientes de apenas um método, comumente o DSSP, nós optamos por utilizar os resultados de um consenso entre quatro métodos. Consequentemente, é importante analisar como esses métodos diferem entre si.

As diferenças entre as metodologias de atribuição foram previamente discutidas na seção 2.1. Nesta seção, analisaremos as diferenças nas estruturas secundárias atribuídas aos resíduos.

As análises a seguir foram feitas entre os resíduos de 6749 proteínas do banco de dados Top8000-HOM50 que tiveram estrutura secundária atribuída por todos os métodos. Isso exclui 5,19% dos resíduos que não tiveram a estrutura secundária atribuída por não terem sido visualizados na estrutura experimental e 0,05% dos resíduos que não tiveram sua estrutura atribuída por um ou mais programas devido à possíveis falhas dos programas.

Notamos que, entre os quatro métodos de atribuição analisados, há a ocorrência de uma variação considerável entre as estruturas secundárias atribuídas para cada resíduo. A figura 3.1 demonstra a similaridade observada entre os métodos. Como esperado, a maior similaridade é observada entre os métodos DSSP e STRIDE, uma vez que o último é inspirado no primeiro. Entretanto, a similaridade entre os demais métodos encontra-se na faixa entre 83% e 85%. A porcentagem de resíduos que apresentaram consenso entre os quatro métodos foi de apenas 74,43%.

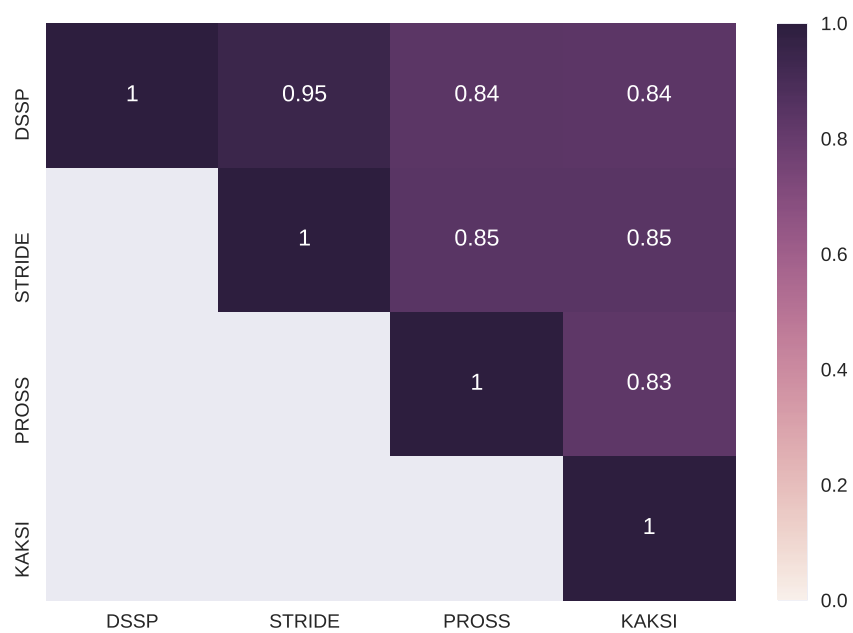


Figure 3.1: Similaridade entre as estruturas secundárias atribuídas para cada resíduo entre os quatro métodos de atribuição de estruturas secundárias: DSSP, STRIDE, KAKSI e PROSS.

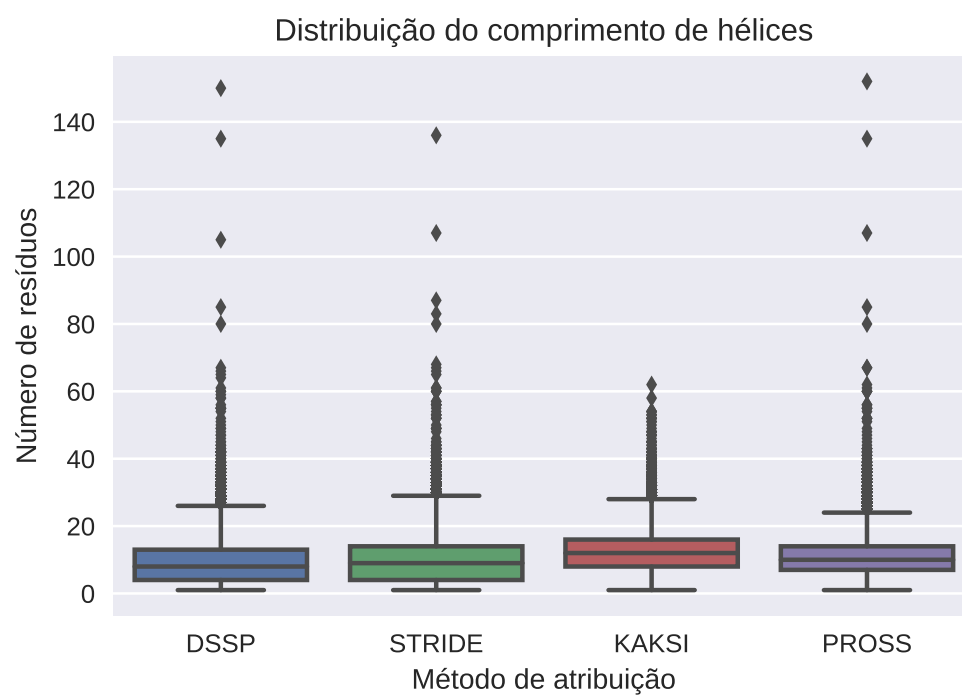


Figure 3.2: Distribuição do comprimento de hélices por método de atribuição de estrutura secundária.

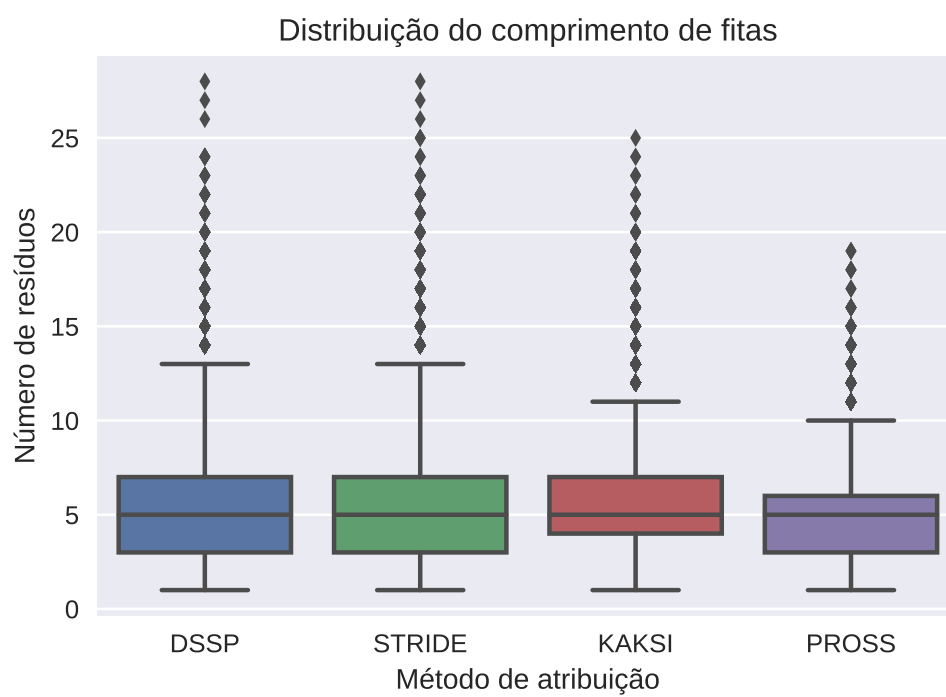


Figure 3.3: Distribuição do comprimento de fitas por método de atribuição de estrutura secundária.

Bibliography

- Kabsch, Wolfgang and Christian Sander (1983). “Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features”. In: *Biopolymers* 22.12, pp. 2577–2637.
- Kendrew, John C et al. (1958). “A three-dimensional model of the myoglobin molecule obtained by x-ray analysis”. In: *Nature* 181.4610, pp. 662–666. DOI: [10 . 1038 / 181662a0](https://doi.org/10.1038/181662a0).