

# PROTEIN BUILDER

Leonardo Santos<sup>1</sup>

<sup>1</sup>Biotecnologia - Bioinformática  
Universidade Federal do Rio Grande do Sul

Bioinformática Estrutural, 2017

# ROTEIRO

## PDB FILE

Leitura

## BUILDER

Dictionary

Peptide Bond

Dihedral Angles

## PROBLEMAS ENCONTRADOS

# FRAGMENTAÇÃO

## LEITURA DE ÁTOMOS:

Para que apenas os átomos de interesse sejam utilizados:

- ▶ Apenas linhas que comecem com 'ATOM'
- ▶ Em caso de backbone e carbonos alfa:
  - ▶ Átomos: N, CA, C, O para backbone
  - ▶ CA para carbono alfa
  - ▶ N para carbono amina
  - ▶ C para carbono carboxila
- ▶ Retorna uma lista de vetores, onde cada vetor representa um átomo

# MONTA DICIONÁRIO DE AMINOÁCIDOS

- ▶ Input: 20 arquivos PDB, um para cada aminoácido
- ▶ Output: Dic[AA]:" aa
  - ▶ AA: código de letra correspondente. (Ex: A,C,P,M.
  - ▶ aa: átomos correspondentes correspondente.

# PEPTIDE BOND

- ▶ Input: sequência de aminoácidos
- ▶ Output: arquivo .pbd com a cadeia de aminoácidos correspondentes
  - ▶ Retorna os aminoácidos apenas alinhados, sem rotação

# ENCADEAMENTO

- ▶ Primeiro aminoácido: é definido como o ponto de referência
- ▶ Próximos aminoácidos:
  - ▶ O átomo N do aminoácido  $n$  é posicionado no mesmo lugar que o átomo C da carboxila do aminoácido  $n-1$
  - ▶ Os átomos C e H da carboxila são removidos do aminoácido  $n-1$
  - ▶ O deslocamento do átomo N do aminoácido  $n$  é calculado e aplicado aos outros átomos

└─ PEPTIDE BOND

VSCEDCPEHCSTQKAQAKCDNDKCVCEPI.PDB

1	ATOM	1	N	VAL	1	0.095	-0.181	0.000	1.00	0.00
2	ATOM	2	CA	VAL	1	1.547	-0.026	0.000	1.00	0.00
3	ATOM	3	C	VAL	1	1.935	1.434	0.000	1.00	0.00
4	ATOM	4	O	VAL	1	1.090	2.337	0.000	1.00	0.00
5	ATOM	5	CB	VAL	1	2.166	-0.772	-1.248	1.00	0.00
6	ATOM	6	CG1	VAL	1	1.977	-2.308	-1.217	1.00	0.00
7	ATOM	7	CG2	VAL	1	3.683	-0.551	-1.501	1.00	0.00
8	ATOM	8	1H	VAL	1	-0.571	0.669	0.000	1.00	0.00
9	ATOM	9	2H	VAL	1	-0.377	-1.152	0.000	1.00	0.00
10	ATOM	10	HA	VAL	1	1.952	-0.468	0.928	1.00	0.00
11	ATOM	11	HB	VAL	1	1.638	-0.395	-2.150	1.00	0.00
12	ATOM	12	1HG1	VAL	1	2.529	-2.782	-0.384	1.00	0.00
13	ATOM	13	2HG1	VAL	1	0.914	-2.590	-1.107	1.00	0.00
14	ATOM	14	3HG1	VAL	1	2.323	-2.783	-2.155	1.00	0.00
15	ATOM	15	1HG2	VAL	1	4.307	-0.876	-0.647	1.00	0.00
16	ATOM	16	2HG2	VAL	1	4.043	-1.094	-2.396	1.00	0.00
17	ATOM	17	3HG2	VAL	1	3.928	0.511	-1.696	1.00	0.00
18	ATOM	18	N	SER	2	3.338	1.708	0.000	1.00	0.00
19	ATOM	19	CA	SER	2	4.797	1.744	0.000	1.00	0.00
20	ATOM	20	C	SER	2	5.304	3.166	0.000	1.00	0.00
21	ATOM	21	O	SER	2	4.539	4.131	0.000	1.00	0.00
22	ATOM	22	CB	SER	2	5.368	0.935	-1.192	1.00	0.00
23	ATOM	23	OG	SER	2	5.119	1.557	-2.457	1.00	0.00
24	ATOM	24	2H	SER	2	2.788	0.778	0.000	1.00	0.00
25	ATOM	25	HA	SER	2	5.149	1.278	0.939	1.00	0.00
26	ATOM	26	1HB	SER	2	6.463	0.817	-1.072	1.00	0.00
27	ATOM	27	2HB	SER	2	4.970	-0.098	-1.203	1.00	0.00
28	ATOM	28	HG	SER	2	4.166	1.586	-2.580	1.00	0.00
29	ATOM	29	N	CYS	3	6.725	3.326	0.000	1.00	0.00
30	ATOM	30	CA	CYS	3	8.181	3.440	0.000	1.00	0.00
31	ATOM	31	C	CYS	3	8.609	4.888	0.000	1.00	0.00
32	ATOM	32	O	CYS	3	7.788	5.817	0.000	1.00	0.00
33	ATOM	33	CB	CYS	3	8.734	2.662	-1.210	1.00	0.00
34	ATOM	34	SG	CYS	3	10.539	2.613	-1.168	1.00	0.00
35	ATOM	35	2H	CYS	3	6.226	2.368	0.000	1.00	0.00
36	ATOM	36	HA	CYS	3	8.565	2.984	0.931	1.00	0.00
37	ATOM	37	1HB	CYS	3	8.366	1.620	-1.218	1.00	0.00
38	ATOM	38	2HB	CYS	3	8.401	3.112	-2.168	1.00	0.00
39	ATOM	39	HG	CYS	3	10.746	3.505	-2.134	1.00	0.00

FIGURA: Arquivo PDB

VSCEDCPEHCSTQKAQAKCDNDKCVCEPI

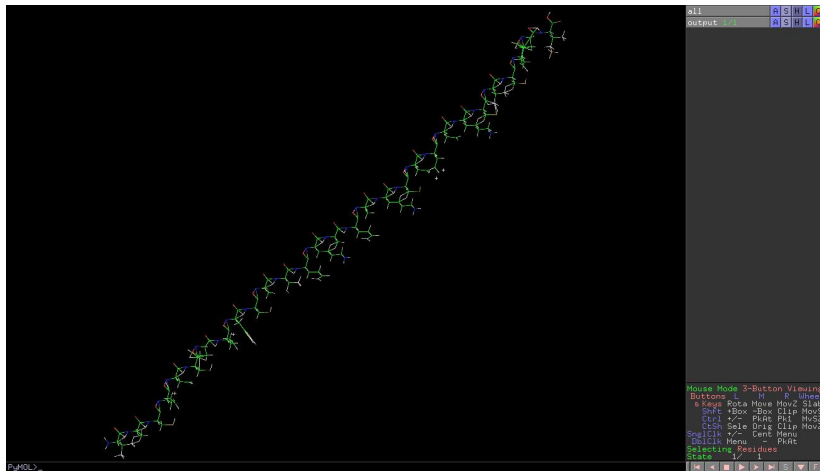


FIGURA: Criada no PyMol



# INICIALIZAÇÃO

- ▶ Input: arquivo .pdb
- ▶ Output1: arquivo .txt com o aminoácido, ângulo Phi e Psi
- ▶ Output2: mapa de Ramachandran

## CÁLCULO DE ÂNGULO

- ▶ Input PSI:  $[N_n, CA_n, C_n, N_{n+1}]$
- ▶ Input PHI:  $[C_{n-1}, N_n, CA_n, C_n]$ 
  - ▶ Cada átomo é representado por um vetor  $[P_x, P_y, P_z]$
- ▶ Calculo de distancia entre átomos
  - ▶ Para cada dupla de átomos é calculado um vetor  $b_{1-2}, b_{2-3}, b_{3-4}$

$$n1 = \frac{b1 \times b2}{\| b1 \times b2 \|}, n2 = \frac{b2 \times b3}{\| b2 \times b3 \|} \rightarrow m1 = n1 \times \frac{b2}{\| b2 \|} \quad (1)$$

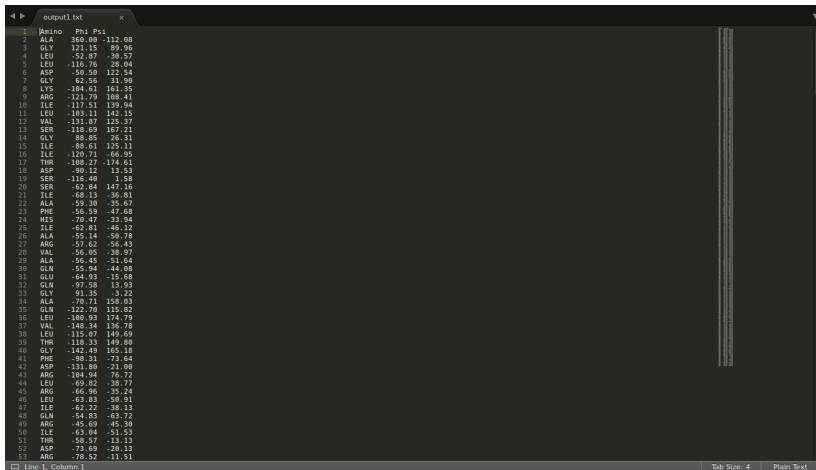
# CÁLCULO DE ÂNGULO

$$x = n1 \times n2, y = m1 \times n2 \rightarrow \phi/\psi = \arctan(y, x) \quad (2)$$

# CÁLCULO DE ÂNGULO

- ▶ Primeiro átomo:
  - ▶ PHI: 360
  - ▶ PSI: cálculo de ângulo
- ▶ Último átomo:
  - ▶ PHI: cálculo de ângulo
  - ▶ PSI: 360

1ENY



	Amino	Phi	Psi
1	ALA	360.00	-112.88
2	GLY	121.15	89.96
3	LEU	-52.87	-30.57
4	LEU	-116.76	28.04
5	ASP	-50.50	122.54
6	GLY	62.56	31.90
7	LYS	-104.61	161.35
8	ARG	-121.79	188.41
9	ILE	-117.51	139.94
10	LEU	-103.11	142.15
11	VAL	-131.07	125.37
12	SER	-118.69	167.21
13	GLY	88.85	26.31
14	ILE	-88.61	125.11
15	ILE	-120.71	-66.95
16	THR	-108.27	-174.61
17	ASP	-90.12	13.53
18	SER	-116.40	1.58
19	SER	-62.84	147.16
20	ILE	-68.13	-36.81
21	ALA	-59.30	-35.67
22	PHE	-56.59	-47.60
23	MET	-70.47	-33.94
24	ILE	-62.81	-46.12
25	ALA	-55.14	-50.78
26	ARG	-57.62	-56.43
27	VAL	-56.05	-38.97
28	ALA	-56.45	-51.64
29	GLN	-55.94	-44.08
30	GLU	-64.93	-15.68
31	GLN	-97.58	13.93
32	GLY	91.35	-3.22
33	ALA	-70.71	158.03
34	GLN	-122.70	115.82
35	LEU	-100.93	174.79
36	VAL	-148.34	136.78
37	LEU	-115.07	149.69
38	THR	-110.33	149.00
39	GLY	-142.49	165.18
40	PHE	-98.31	-73.64
41	ASP	-131.80	-21.00
42	ARG	-104.94	76.72
43	LEU	-69.82	-38.77
44	ARG	-66.96	-35.24
45	LEU	-63.83	-50.91
46	ILE	-62.22	-38.13
47	GLN	-54.83	-63.72
48	ARG	-45.69	-45.30
49	ILE	-63.04	-51.53
50	THR	-58.57	-13.13
51	ASP	-73.69	-20.13
52	ARG	-78.52	-11.51

FIGURA: Output AA-PHI-PSI

# 1ENY

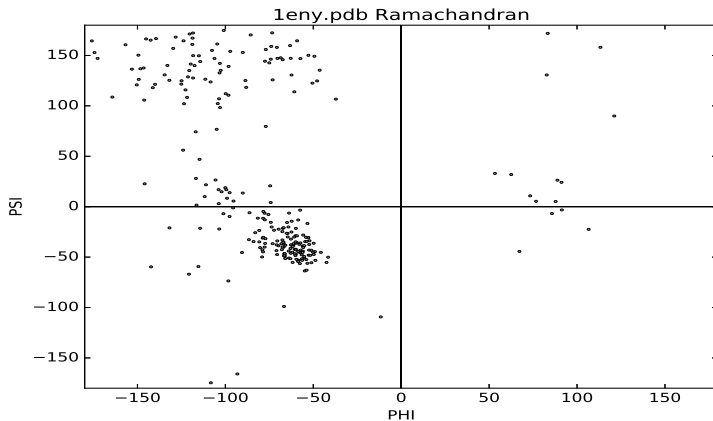


FIGURA: 1ENY Ramachandran's map

# ROTAÇÃO PHI E PSI

- ▶ Input:
  - ▶ PDB: arquivo no formato PDB
- ▶ Output:
  - ▶ PDB: arquivo no formato PDB com a estrutura modificada

# ALGORITMO:

- ▶ Primeiro átomo:
  - ▶ PHI: 360
  - ▶ PSI: cálculo de ângulo
- ▶ Último átomo:
  - ▶ PHI: cálculo de ângulo
  - ▶ PSI: 360
- ▶ Cálculo de ângulo:
  - ▶ PHI: Para o primeiro aminoácido → todos os átomos, menos N e Ca
  - ▶ PSI: Para o primeiro aminoácido → nenhum átomo, exceto O da carbonila



## ALGORITMO:

- Cálculo de ângulo:
  - PHI: Para o primeiro aminoácido → todos os átomos, menos N e Ca
  - PSI: Para o primeiro aminoácido → nenhum átomo, exceto O da carbonila

$$\mathbf{R}_{x,\alpha} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}$$

(A) Eixo X

$$\mathbf{R}_{y,\phi} = \begin{bmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{bmatrix}$$

(B) Eixo Y

$$\mathbf{R}_{z,\theta} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(C) Eixo Z

FIGURA: Rotação quando os eixos são fixados

## SAÍDA:

- Arquivo PDB com as novas posições

1	ATOM	1	N	TYR	1	0.115	-0.105	0.000	1.00	0.00
2	ATOM	2	CA	TYR	1	1.574	-0.047	0.000	1.00	0.00
3	ATOM	3	C	TYR	1	1.234	2.147	-0.146	1.00	0.00
4	ATOM	4	O	TYR	1	1.279	2.337	0.000	1.00	0.00
5	ATOM	5	CB	TYR	1	2.157	-0.753	-1.259	1.00	0.00
6	ATOM	6	CG	TYR	1	1.856	-2.248	-1.412	1.00	0.00
7	ATOM	7	CD1	TYR	1	2.567	-3.198	-0.673	1.00	0.00
8	ATOM	8	CD2	TYR	1	0.854	-2.668	-2.292	1.00	0.00
9	ATOM	9	CE1	TYR	1	2.273	-4.553	-0.810	1.00	0.00
10	ATOM	10	CE2	TYR	1	0.562	-4.022	-2.428	1.00	0.00
11	ATOM	11	CZ	TYR	1	1.272	-4.963	-1.686	1.00	0.00
12	ATOM	12	OH	TYR	1	0.985	-6.293	-1.815	1.00	0.00
13	ATOM	13	1H	TYR	1	-0.493	0.788	0.000	1.00	0.00
14	ATOM	14	2H	TYR	1	-0.421	-1.042	0.000	1.00	0.00
15	ATOM	15	HA	TYR	1	1.952	-0.534	0.918	1.00	0.00
16	ATOM	16	1HB	TYR	1	1.824	-0.213	-2.170	1.00	0.00
17	ATOM	17	2HB	TYR	1	3.258	-0.632	-1.277	1.00	0.00
18	ATOM	18	HD1	TYR	1	3.342	-2.888	0.014	1.00	0.00
19	ATOM	19	HD2	TYR	1	0.290	-1.942	-2.863	1.00	0.00
20	ATOM	20	HE1	TYR	1	2.816	-5.287	-0.234	1.00	0.00
21	ATOM	21	HE2	TYR	1	-0.219	-4.332	-3.106	1.00	0.00
22	ATOM	22	HH	TYR	1	0.276	-6.389	-2.453	1.00	0.00

FIGURA: PHI e PSI rotados

# MONTAGEM DA ESTRUTURA:

- ▶ Distância: 1.32 Å
- ▶ Cálculo de ângulo: Sobreposição do  $N_n$  com o  $OH_{n-1}$

FIGURA: Ligação tripla de hidrogenios e ângulo de ligação

## DIFERENÇAS DE ARQUIVOS PDB:

- ▶ Ausência ou presença de indicador de cadeia
  - ▶ Muda a posição das coordenadas na linha ATOM
- ▶ Diferente notação para átomos iguais
  - ▶ H ligado ao N: "H1", "H2", "H3", "1H", "2H", "3H"

ATOM	20	HD2	TYR	A	1	-1.155	5.126	-1.536	1.00	0.00	H
ATOM	21	HE1	TYR	A	1	-4.156	1.532	-0.033	1.00	0.00	H
ATOM	22	HE2	TYR	A	1	-2.827	4.205	-3.082	1.00	0.00	H
ATOM	23	HH	TYR	A	1	-4.090	1.839	-3.233	1.00	0.00	H

FIGURA: 1PLX.pdb

- ▶ Não deve modificar a posição

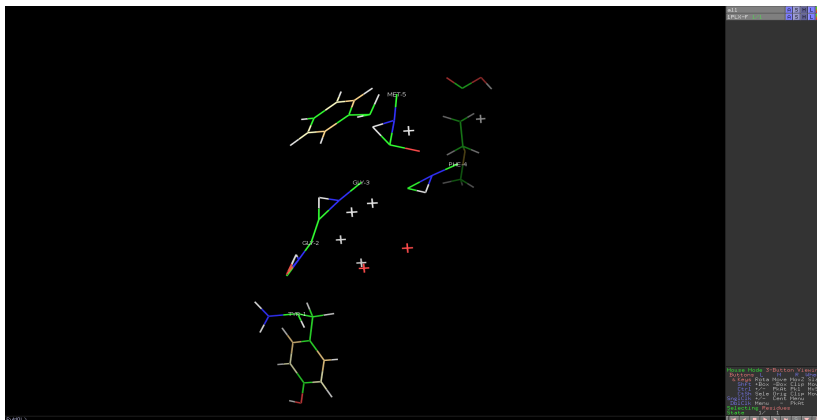


FIGURA: Resíduos fragmentados por uma rotação errada