SSBI summer term 2023

Huiting Xu, Desirée Renschler-Sperl

2023

Assignment 02

1 Task 1-Torsion Angles

1.1 a

conformation	associated angles
A	180/180
B	180/60
C	180/0
D	60/60

1.2 b

Amino acid	$\chi 1$	$\chi 2$	$\chi 3$	$\chi 4$	$\chi 5$
MET	N-CA-CB-CG	CA-CB-CG-SD	CB-CG-SD-CE	-	-
ALA	-	-	-	-	-
PRO	N-CA-CB-CG	CA-CB-CG-CD	CB-CG-CD-N	CG-CD-N-CA	-
TYR	N-CA-CB-CG	CA-CB-CG-CD1	-	-	-
LYS	N-CA-CB-CG	CA-CB-CG-CD	CB-CG-CD-CE	CG-CD-CE-NZ	-
ARG	N-CA-CB-CG	CA-CB-CG-CD	CB-CG-CD-NE	CG-CD-NE-CZ	-

Table 1: χ angles for given amino acids

1.3 c

What distinctive conformations does the torsion angle $\chi 1$ of the bond Cα -Cβ in glutamate take, from most to least preferred? Please cite your sources.
Glutamate prefers $\chi 1$ angles of -60° (gauche-minus which is the same as 300°) followed by 180° (trans conformation) and 60° (gauche-plus) being least favored [1, 4, 6, 5, 3, 2, 8, 7].

2 Task 2 Protein Data Bank I

2.1 1. How many amino acids does each primary structure contain?

5IRE consists of the E protein with 504 amino acids and the M protein with 75 amino acids which comprise the Zika virus icosahedral shell. 1IGT is a IgG-type antibody and consists of a heavy chain with 444 amino acids and a light chain with 214 amino acids.

2.2 2. What is the prevailing secondary structure of each protein?

The prevailing secondary structure elements of the E protein in 5IRE are beta-sheets and of the M protein are alpha-helices. The prevailing secondary structure elements of 1IGT (both heavy and light chains) are beta-sheets.

2.3 3. How many chains do the proteins consist of?

If protein means "pdb entry", 5IRE contains 6 chains. M protein with chains B, D and F and the E protein with the chains A, C, and E. If protein means "pdb entry", 1IGT contains 4 chains. If protein is defined as quatery assembly in a biological sense, the Zika virus icosahedral shell (5ire) consists of 360 chains (180 E-proteins, 180 M-proteins). If protein is defined as quatery assembly in a biological sense, the IgG-type antibody 1igt consists of 4 chains, 2 light chains and 2 heavy chains.

2.4 4. What is the role of the quaternary structure for the 5IRE protein entry?

The quaternary structure assembles the virus shell of the Zika virus consisting of 180 E-proteins and 180 M-proteins.

2.5 5. For both proteins, find three hydrogen bonds and three disulfide bonds. State the residue number of involved residues and give an estimate of the bond lengths.

1IGT H-bonds:

aa	residue number	atom type	aa	residue number	atom type	bond length
LYS	45	N	GLN	37	O	3.1 A
LYS	45	O	GLN	37	N	3.1 A
ASN	42	N	LYS	39	O	2.5 A

Table 2: H-bonds for 1IGT

1IGT disulfide bonds:

aa	residue number	aa	residue number	bond length
CYS	23	CYS	88	2.03
CYS	134	CYS	194	2.02
CYS	214	CYS	128	2.04

Table 3: Disulfide bonds for 1IGT

5IRE h-bonds:

aa	residue number	atom type	aa	residue number	atom type	bond length
SER	7	N	ILE	4	O	3.4 A
ILE	4	O	ARG	9	NH2	3.1 A
ASN	8	O	CYS	30	N	3.5 A

Table 4: H-bonds of 5IRE

5IRE disulfide bonds:

aa	chain	residue number	aa	chain	residue number	bond length
CYS	A	74	CYS	A	105	2.02
CYS	A	92	CYS	A	116	2.03
CYS	C	74	CYS	C	105	2.02

Table 5: Disulfide bonds of 5IRE

2.6 6. How many disulfide bonds do you find in each protein? Which residues form these bonds?

type	nr	aa	chain	residue nr	aa	chain	residue nr			bond length
SSBOND	1	CYS	A	74	CYS	A	105	1555	1555	2.02
SSBOND	2	CYS	A	92	CYS	A	116	1555	1555	2.03
SSBOND	3	CYS	C	74	CYS	C	105	1555	1555	2.02
SSBOND	4	CYS	C	92	CYS	C	116	1555	1555	2.03
SSBOND	5	CYS	E	74	CYS	E	105	1555	1555	2.02
SSBOND	6	CYS	E	92	CYS	E	116	1555	1555	2.03

Table 6: Disulfid bonds for 5IRE

type	nr	aa	chain	residue nr	aa	chain	residue nr			bond length
SSBOND	1	CYS	A	23	CYS	A	88	1555	1555	2.03
SSBOND	2	CYS	A	134	CYS	A	194	1555	1555	2.02
SSBOND	3	CYS	A	214	CYS	B	128	1555	1555	2.04
SSBOND	4	CYS	B	22	CYS	B	92	1555	1555	2.04
SSBOND	5	CYS	B	142	CYS	B	208	1555	1555	2.03
SSBOND	6	CYS	B	237	CYS	D	237	1555	1555	2.04
SSBOND	7	CYS	B	240	CYS	D	240	1555	1555	2.04
SSBOND	8	CYS	B	242	CYS	D	242	1555	1555	2.04
SSBOND	9	CYS	B	274	CYS	B	340	1555	1555	2.02
SSBOND	10	CYS	B	390	CYS	B	456	1555	1555	2.02
SSBOND	11	CYS	C	23	CYS	C	88	1555	1555	2.02
SSBOND	12	CYS	C	134	CYS	C	194	1555	1555	2.03
SSBOND	13	CYS	C	214	CYS	D	128	1555	1555	2.03
SSBOND	14	CYS	D	22	CYS	D	92	1555	1555	2.03
SSBOND	15	CYS	D	142	CYS	D	208	1555	1555	2.04
SSBOND	16	CYS	D	274	CYS	D	340	1555	1555	2.03
SSBOND	17	CYS	D	390	CYS	D	456	1555	1555	2.01

Table 7: Disulfide bonds for 1IGT

2.7 7. What is the relationship of the disulfide bond type to structure stability?

Disulfide bonds, beeing covalent in their nature, are very stable interactions. Hence, they provide extraordinary stability when compared to non-covalent bonds such as e.g. H-bonds, Phi-Phi interactions or salt bridges.

3 Task 3 Ramachandran Maps

Please see code and README.txt. As input files 1IGT, 5IRE are chosen (provided as .pdb files). The resulting Ramachandran plots are provided in RamachandranMaps. 1IGT mainly consists of /*beta*-sheets. In the resulting plot, most of the spots are in the upper left part, as expected. For 5IRE the resulting Ramachandran plot shows spots in the upper left quadrat for the /*beta*-sheets and spots in the lower left quadrant for the /*alpha*-helices, as expected.

4 Task 4 Unusual Structure Motifs

- What properties make the left-handed conformation feasible, compared to other amino acids?
  Glycine has only one H-atom as a 'side-chain' hence it is not chiral. This gives glycine a higher flexibility in the torsional angles as compared to the other amino acids. Its Ramachandran plot is symmetrical.

- What would you expect to happen if a glycine residue in a long polyglycine helix was exchanged to a different proteinogenic amino acid?
  The helix would be probably right-handed because the side chains of the other proteinogenic amino acids are to big to allow left-handed helices with more than 3-5 residues. If the glycine is exchanged to proline the helix would break because the geometry of proline is not compatible with α-helices. Proline can therefor only be at the end of a helix.

- Would D-alanine prefer left- or right-handed helices?
  It prefers right-handed helices [9]. Figure 2 shows a Ramachandran plot for D-alanine. It prefers angles that appear in the upper right quadrant. Figure 1 reveals that this region belongs to right-handed α-helices.

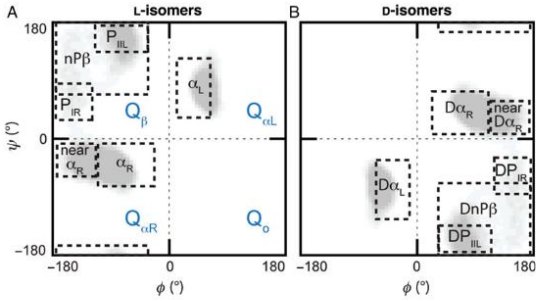

Fig. 1. Defined quadrants and conformational regions for the L- and D-amino acids. (**A**) Labeling convention used for the four quadrants, $Q_{\alpha R}$ Q_{β}, $Q_{\alpha L}$ and Q_{o}, and the regions within the quadrants pertaining to specific conformations: right-handed α-helix (α_R), left-handed α-helix (α_L), polyproline type II left-handed helix (P_{IIL}), polyproline type I right-handed helix (P_{IR}) and remainder of non-polyproline β-region (nPβ). (**B**) Corresponding conformational regions for the D-amino acids: right-handed α-helix ($D\alpha_R$), left-handed α-helix ($D\alpha_L$), polyproline type II left-handed helix (DP_{IIL}), polyproline type I right-handed helix (DP_{IR}) and remainder of non-polyproline β-region (DnPβ).

Figure 1: [9]

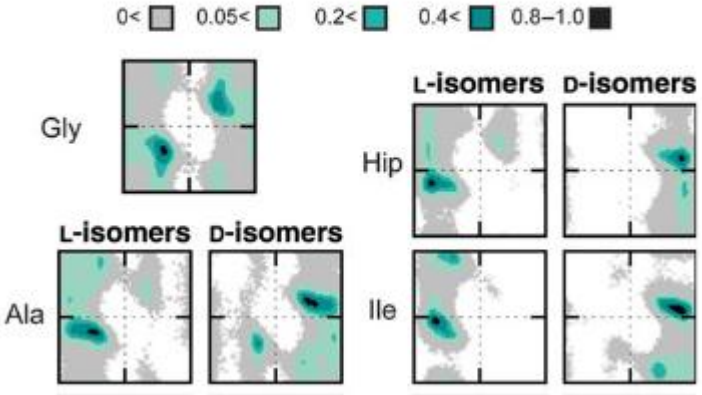

Figure 2: [9]

References

[1] https://dunbrack.fccc.edu/bbdep2010/ConformationalAnalysis.php. Accessed: 2023-05-07.

[2] Roland L Dunbrack Jr. Rotamer libraries in the 21st century. *Current opinion in structural biology*, 12(4):431–440, 2002.

[3] Roland L Dunbrack Jr and Fred E Cohen. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Science*, 6(8):1661–1681, 1997.

[4] Roland L Dunbrack Jr and Martin Karplus. Backbone-dependent rotamer library for proteins application to side-chain prediction. *Journal of molecular biology*, 230(2):543–574, 1993.

[5] Roland L Dunbrack Jr and Martin Karplus. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nature structural biology*, 1(5):334–340, 1994.

[6] Roland Leslie Dunbrack Jr. *Conformational analysis of protein sidechains: Empirical energy parameters for proline and development of a backbone-dependent rotamer library*. Harvard University, 1993.

[7] Maxim V Shapovalov and Roland L Dunbrack. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure*, 19(6):844–858, 2011.

[8] Maxim V Shapovalov and Roland L Dunbrack Jr. Statistical and conformational analysis of the electron density of protein side chains. *Proteins: Structure, Function, and Bioinformatics*, 66(2):279–303, 2007.

[9] Clare-Louise Towse, Gene Hopping, Ivan Vulovic, and Valerie Daggett. Nature versus design: the conformational propensities of d-amino acids and the importance of side chain chirality. *Protein Engineering, Design & Selection*, 27(11):447–455, 2014.