Contents

1 Criteria for identifying various types of interactions **1**
1.1 Disulphide bridges 1
1.2 Hydrogen bonds 2
1.3 Amino-aromatic interactions 4
1.4 Aromatic-aromatic interactions 6
1.5 Sulphur-aromatic interactions 7
1.6 Ionic interactions 8
1.7 Hydrophobic interactions 9

References **10**

1 Criteria for identifying various types of interactions

1.1 Disulphide bridges

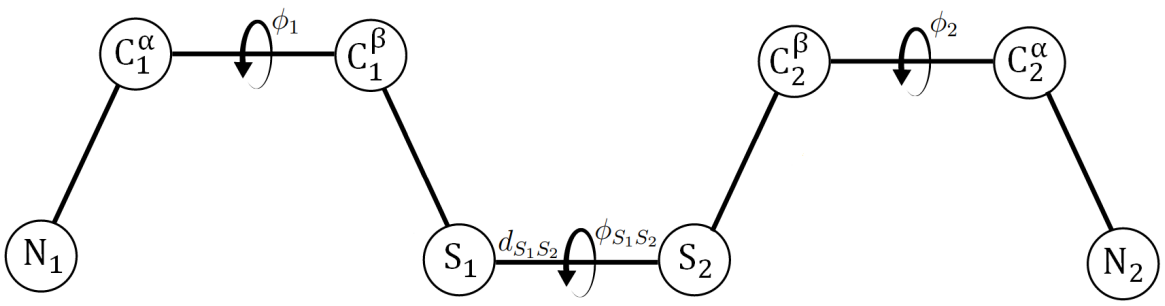

Figure 1: Diagram of a disulphide bridge. Atoms with subscript 1 and 2 represent two cysteines forming a disulphide bridge. N - a nitrogen atom from amino group, C^α - a carbon atom C^α, C^β - a carbon atom C^β, S - a sulphur atom. $d_{S_1S_2}$ - a distance between sulphur atoms of two cysteines, $\phi_{S_1S_2}$ - a dihedral angle between planes formed by C_1^β, S_1, S_2 and S_1, S_2, C_2^β. ϕ_1 - a dihedral angle between planes formed by N_1, C_1^α, C_1^β and C_1^α, C_1^β, S_1. ϕ_2 - a dihedral angle between planes formed by S_2, C_2^β, C_2^α and C_2^β, C_2^α, N_2.

Default criteria for disulphide bridge identification [8]:

$$d_{C_1^\alpha C_2^\alpha} \leqslant 6.5 \text{ Å} \quad \wedge \quad d_{C_1^\beta C_2^\beta} \leqslant 4.5 \text{ Å}, \tag{1}$$

$$1.6 \text{ Å} \leqslant d_{S_1S_2} \leqslant 2.4 \text{ Å} \quad \wedge \quad 60^\circ \leqslant |\phi_{S_1S_2}| \leqslant 120^\circ, \tag{2}$$

$$30^\circ \leqslant |\phi_1| \leqslant 90^\circ \quad \vee \quad 150^\circ \leqslant |\phi_1| \leqslant 180^\circ, \tag{3}$$

$$30^\circ \leqslant |\phi_2| \leqslant 90^\circ \quad \vee \quad 150^\circ \leqslant |\phi_2| \leqslant 180^\circ. \tag{4}$$

1.2 Hydrogen bonds

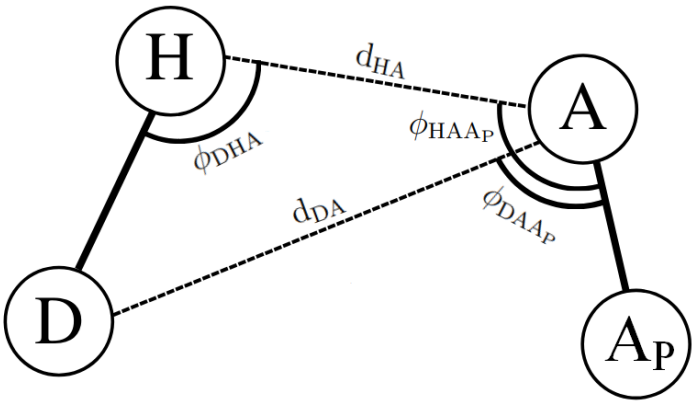

Figure 2: Diagram of a hydrogen bond. D - a hydrogen bond donor atom, H - a hydrogen atom forming a covalent bond with D, A - a hydrogen bond acceptor atom, having a lone pair of polarized π-electrons, A_P - an atom antecedent for A in the side chain of amino acid. d_{HA} - a distance between H and A, d_{DA} - a distance between D and A, ϕ_{DHA} - an angle between the vector $\overrightarrow{HD}$ and the vector $\overrightarrow{HA}$, ϕ_{HAA_P} - an angle between the vector $\overrightarrow{AH}$ and the vector $\overrightarrow{AA_P}$, ϕ_{DAA_P} - an angle between the vector $\overrightarrow{AD}$ and the vector $\overrightarrow{AA_P}$.

Default criteria for hydrogen bond identification: [4, 9]

$$d_{HA} < 2.5 \text{ Å}, \tag{5}$$

$$\phi_{DHA} > 90^\circ. \tag{6}$$

Futhermore, d_{DA}, ϕ_{HAA_P} and ϕ_{DAA_P} are calculated to provide additional information about the geometry of particular interaction. However, they do not determine the presence of the hydrogen bond in default criteria. The set of considered groups of atoms D, H, A and A_P is presented in Table 1.

Table 1: Set of donors and acceptors of hydrogen bond [9]. Names of atoms are consistent with the standardized naming convention used in *PDB* files.

Amino Acid	**Atom**			
	Atom Antecedent	**Acceptor**	**Hydrogen**	**Donor**
Main Chain				
All	C	O	H	N
C-end	C C	O OXT	H	N
N-end	C	O	H1, H2, H3	N
Side Chain				

Arg	-	-	HE HH11, HH12 HH21, HH22	NE NH1 NH2
Asn	CG	OD1	HD21, HD22	ND2
Asp	CG CG	OD1 OD2	-	-
Cys	CB	SG	HG	SG
Gln	CD	OE1	HE21, HE22	NE2
Glu	CD CD	OE1 OE2	-	-
His	CG CD2	ND1 NE2	HD1 HE2	ND1 NE2
Lys	-	-	HZ1, HZ2, HZ3	NZ
Met	CG	SD	-	-
Ser	CB	OG	HG	OG
Thr	CB	OG1	HG1	OG1
Trp	-	-	HE1	NE1
Tyr	CZ	OH	HH	OH

1.3 Amino-aromatic interactions

Amino-aromatic interactions are defined as interactions between side-amino groups of *Lys, Arg, Asn, Gln* and *His* and aromatic rings in side chains of *Phe, Tyr* and *Trp*. Histidine was not considered as aromatic amino acid since, although it is aromatic, it is also highly polar and hydrophobic and therefore is mainly involved in other types of interactions [3].

The indole ring of tryptophan is treated as two separate components - pyrrole ring (5 membered) and phenyl ring (6 membered) [7].

For identification of amino-aromatic interactions, it is checked if particular histidine is charged. It is assumed that histidine is charged if it has both nitrogen protonated (both nitrogens' hydrogen atoms are present in *PDB* structure) [4].

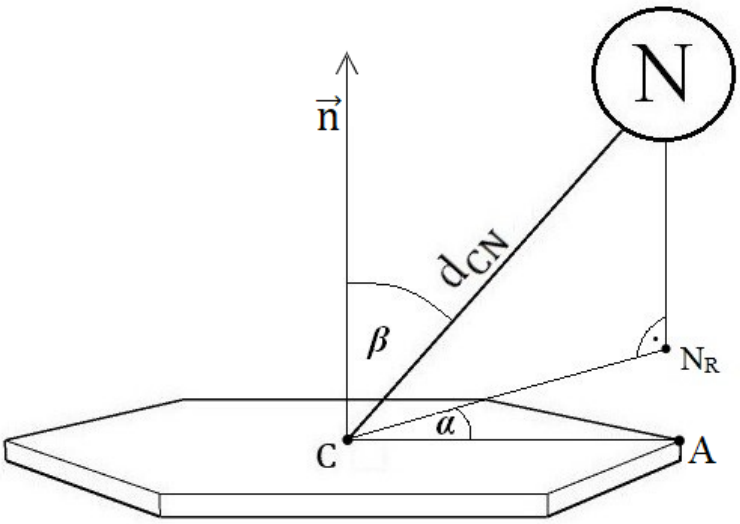

Figure 3: Diagram of an amino-aromatic interaction. $\overrightarrow{n}$ - a normal vector to the aromatic ring plane, C - a centroid of the aromatic ring, N - a nitrogen atom or a representative atom of amino group (see Table 2), N_R - a projection of N on the plane of the aromatic ring, A - an atom of the aromatic ring used for calculating the α angle (C^{γ} for phenyloalanine and tyrosine, $N^{\epsilon 1}$ for pyrrole ring of tryptophan and $C^{\zeta 2}$ for phenyl ring of tryptophan). d_{CN} - a distance between C and N. α - an angle between the vector and $\overrightarrow{CA}$ and the vector $\overrightarrow{CN}_R$. β - an angle between $\overrightarrow{n}$ and $\overrightarrow{CN}$.

Default criteria for amino-aromatic interaction identification [1]:

$$d_{CN} \in (3.4, 6) \text{ Å}. \tag{7}$$

Futhermore, α and β are calculated to provide additional information about the geometry of particular interaction. However, they do not determine the presence of the interaction in default criteria.

Table 2: Representative atoms of amino groups in considered amino acids. Names of atoms are consistent with the standardized naming convention used in *PDB* files.

Amino acid	**Representative atom of amino group**
Arg	CZ
Asn	ND2
Gln	NE2
His	CE1
Lys	NZ

1.4 Aromatic-aromatic interactions

Aromatic-aromatic interactions are defined as interactions between aromatic rings which are present in side-chain of *Phe, Tyr* and *Trp*. Histidine was not considered as aromatic amino acid since, although it is aromatic, it is also highly polar anf hydrophobic and therefore is mainly involved in other types of interactions [3].

The indole ring of tryptophan is treated as two separate components - pyrrole ring (5 membered) and phenyl ring (6 membered) [7].

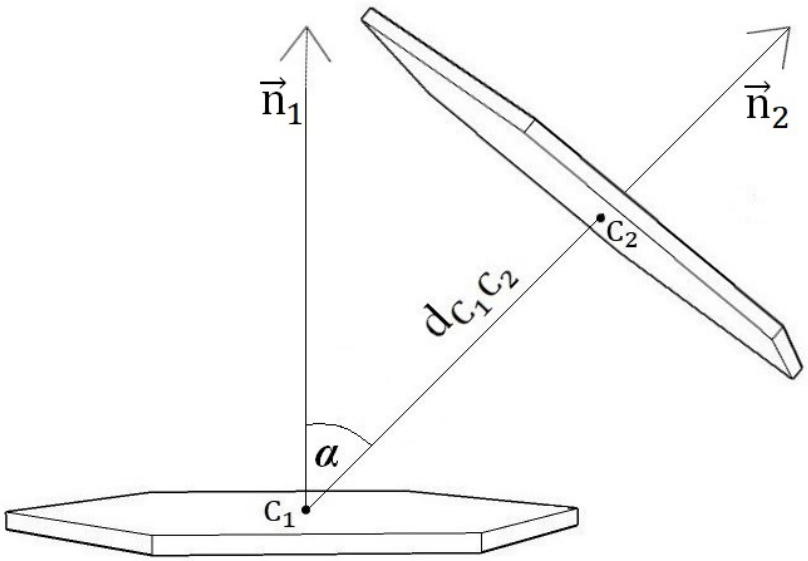

Figure 4: Diagram of an aromatic-aromatic interaction. Atoms with subscript 1 and 2 represent two aromatic rings. $\overrightarrow{n}$ - a normal vector to the aromatic ring plane, C - a centroid of the aromatic ring. $d_{C_1C_2}$ - a distance between centroids of two aromatic rings. α - a dihedral angle between two planes formed by interacting aromatic rings.

Default criteria for aromatic-aromatic interaction identification [1]:

$$d_{C_1C_2} < 7 \text{ Å}. \tag{8}$$

Futhermore, α is calculated to provide additional information about the geometry of particular interaction. However, it does not determine the presence of the interaction in default criteria.

1.5 Sulphur-aromatic interactions

Sulphur-aromatic interactions are defined as interactions between sulphur atom of *Met* and *Cys* and aromatic rings in side chains of *Phe, Tyr* and *Trp*. Histidine was not considered as aromatic amino acid since, although it is aromatic, it is also highly polar anf hydrophobic and therefore is mainly involved in other types of interactions [3].

The indole ring of tryptophan is treated as two separate components - pyrrole ring (5 membered) and phenyl ring (6 membered) [7].

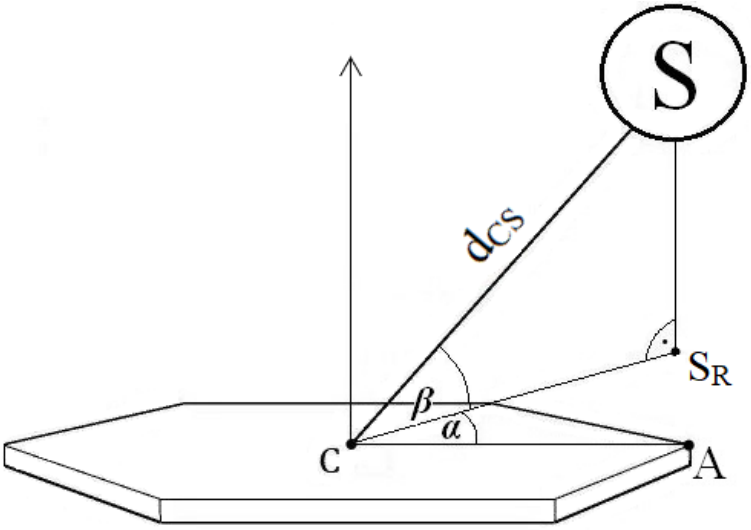

Figure 5: Diagram of a sulphur-aromatic interaction. C - a centroid of the aromatic ring, S - a sulphur atom, S_R - a projection of S on the atomatic ring plane, A - an atom of aromatic ring used for calculating the α angle (C^{ζ} for phenyloalanine and tyrosine, $N^{\epsilon 1}$ for pyrrole ring of tryptophan and $C^{\zeta 2}$ for phenyl ring of tryptophan). d_{CS} - a distance between C and S, α - an angle between the vector $\overrightarrow{CA}$ and the vector $\overrightarrow{CS}_R$. β - an angle between the vector $\overrightarrow{CS}_R$ and the vector $\overrightarrow{CS}$.

Default criteria for sulphur-aromatic interaction identification [6]:

$$d_{CS} \leq 6 \text{ Å}. \tag{9}$$

Futhermore, α and β are calculated to provide additional information about the geometry of particular interaction. However, they do not determine the presence of the interaction in default criteria.

1.6 Ionic interactions

To identify ionic interaction the parameter $d_{C^+C^-}$ was defined as a distance between centroids of oppositely charged side chain groups of polar amino acids: *Arg, Asp, Glu, His* and *Lys* (see Table 3).

For identification of ionic interactions, it is checked if particular histidine is charged. It is assumed that histidine is charged if it has both nitrogen protonated (both nitrogens' hydrogen atoms are present in *PDB* structure) [4].

Table 3: Set of atoms of charged side-chain groups of polar amino acids, based on which the centroid of charged group is calculated [2]. Names of atoms are consistent with the standardized naming convention used in *PDB* files.

Amino acid	**Atoms of charged side-chain group**	
	Positive	**Negative**
Asp	-	CG, OD1, OD2
Glu	-	CD, OE1, OE2
Arg	NE, CZ, NH1, NH2	-
Lys	NZ	-
His	CG, ND1, CD2, CE1, NE2	-

Default criteria for ionic interaction identification [2]:

$$d_{C^+C^-} \leq 5 \text{ Å}. \tag{10}$$

1.7 Hydrophobic interactions

To identify hydrophobic interaction the parameter $d_{C_1^\alpha C_2^\alpha}$ was defined as a distance between C^α (CA) atoms of hydrophobic amino acids: *Ala, Cys, Gly, Ile, Leu, Met, Phe, Pro, Trp, Tyr* oraz *Val*.

Default criteria for hydrophobic interaction identification [5]:

$$d_{C_1^\alpha C_2^\alpha} \leq 9.5 \text{ Å}. \tag{11}$$

References

[1] SK Burley, GA Petsko. Amino-aromatic interactions in proteins. *FEBS letters*, 203(2):139–143, 1986.

[2] Sandeep Kumar, Ruth Nussinov. Relationship between ion pair geometries and electrostatic strengths in proteins. *Biophysical journal*, 83(3):1595–1612, 2002.

[3] Esteban Lanzarotti, Rolf R Biekofsky, Darío A Estrin, Marcelo A Marti, Adrián G Turjanski. Aromatic–aromatic interactions in proteins: beyond the dimer. *Journal of chemical information and modeling*, 51(7):1623–1633, 2011.

[4] Ian K McDonald, Janet M Thornton. Satisfying hydrogen bonding potential in proteins. *Journal of molecular biology*, 238(5):777–793, 1994.

[5] Angelo Onofrio, Giovanni Parisi, Giuseppe Punzi, Simona Todisco, Maria Antonietta Di Noia, Fabrizio Bossis, Antonio Turi, Anna De Grassi, Ciro Leonardo Pierri. Distance-dependent hydrophobic–hydrophobic contacts in protein folding simulations. *Physical Chemistry Chemical Physics*, 16(35):18907–18917, 2014.

[6] KSC Reid, PF Lindley, JM Thornton. Sulphur-aromatic interactions in proteins. *FEBS letters*, 190(2):209–213, 1985.

[7] R Sathyapriya, Saraswathi Vishveshwara. Interaction of dna with clusters of amino acids in proteins. *Nucleic acids research*, 32(14):4109–4118, 2004.

[8] R Sowdhamini, N Srinivasan, Brian Shoichet, Daniel V Santi, C Ramakrishnan, P Balaram. Stereochemical modeling of disulfide bridges. criteria for introduction into proteins by site-directed mutagenesis. *Protein Engineering, Design and Selection*, 3(2):95–103, 1989.

[9] Ivan Y Torshin, Irene T Weber, Robert W Harrison. Geometric criteria of hydrogen bonds in proteins and identification ofbifurcated'hydrogen bonds. *Protein Engineering, Design and Selection*, 15(5):359–363, 2002.