WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 24, 2024 – 08:04 AM EDT

PDB ID	:	6H65
Title	:	Crystal structure of the branched-chain-amino-acid aminotransferase from Haliangium ochraceum
Authors	:	Boyko, K.M.; Timofeev, V.I.; Bezsudnova, E.Y.; Nikolaeva, A.Y.; Rakitina, T.V.; Popov, V.O.
Deposited on	:	2018-07-26
Resolution	:	2.35 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
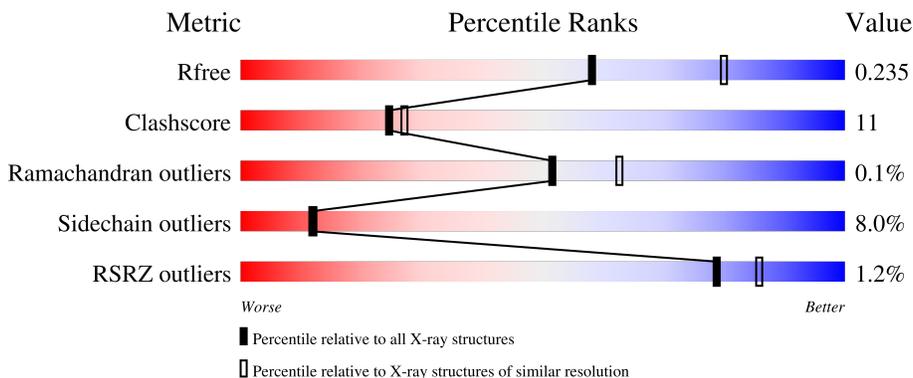
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	317	76% · 14% · 5% · ·
1	B	317	75% · 16% · · ·
1	C	317	77% · 15% · · ·
1	D	317	77% · 15% · · ·
1	E	317	76% · 17% · · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	317	76% 15% · · ·

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PLP	C	401	-	X	-	-
2	PLP	E	401	-	X	-	-

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 14697 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Branched-chain-amino-acid aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	305	Total 2401	C 1518	N 419	O 450	S 14	0	1	0
1	B	306	Total 2419	C 1528	N 423	O 454	S 14	0	1	0
1	C	308	Total 2433	C 1536	N 425	O 458	S 14	0	1	0
1	D	305	Total 2408	C 1523	N 419	O 452	S 14	0	1	0
1	E	307	Total 2428	C 1530	N 427	O 457	S 14	0	2	0
1	F	306	Total 2407	C 1522	N 418	O 453	S 14	0	1	0

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).
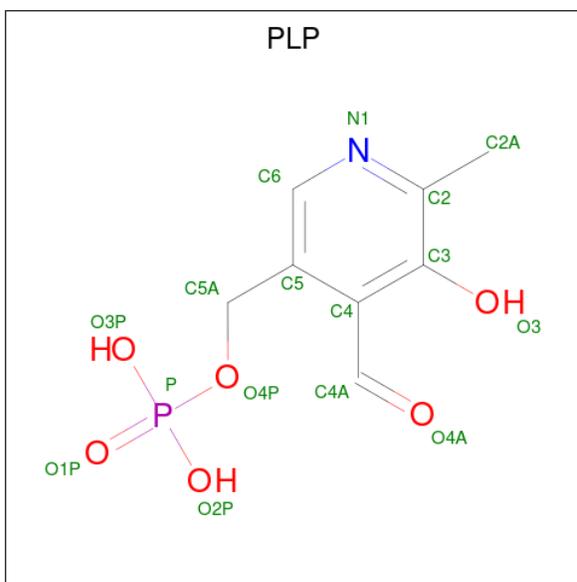
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 15	C 8	N 1	O 5	P 1	0	0
2	B	1	Total 15	C 8	N 1	O 5	P 1	0	0
2	C	1	Total 15	C 8	N 1	O 5	P 1	0	0
2	D	1	Total 15	C 8	N 1	O 5	P 1	0	0
2	E	1	Total 15	C 8	N 1	O 5	P 1	0	0
2	F	1	Total 15	C 8	N 1	O 5	P 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	17	Total 17	O 17	0	0
3	B	21	Total 21	O 21	0	0
3	C	25	Total 25	O 25	0	0
3	D	16	Total 16	O 16	0	0
3	E	19	Total 19	O 19	0	0
3	F	13	Total 13	O 13	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Branched-chain-amino-acid aminotransferase
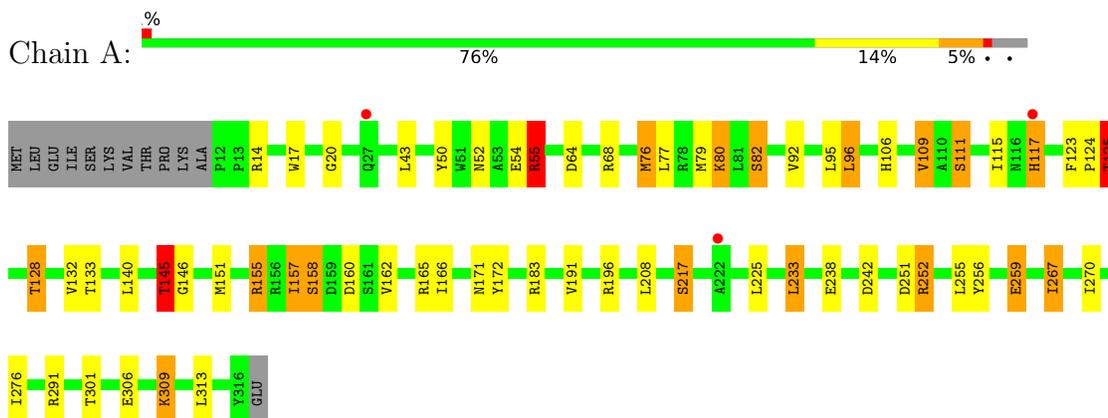
Chain A:

- Molecule 1: Branched-chain-amino-acid aminotransferase

Chain B:

- Molecule 1: Branched-chain-amino-acid aminotransferase
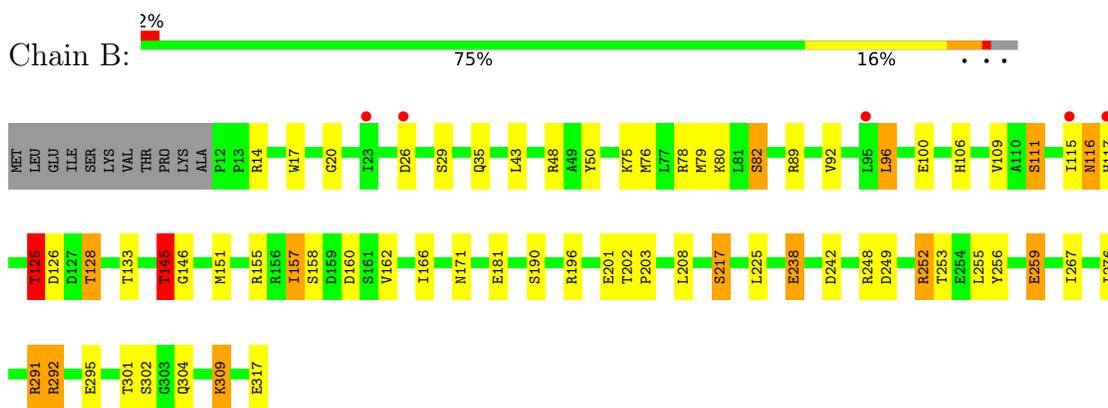
Chain C:
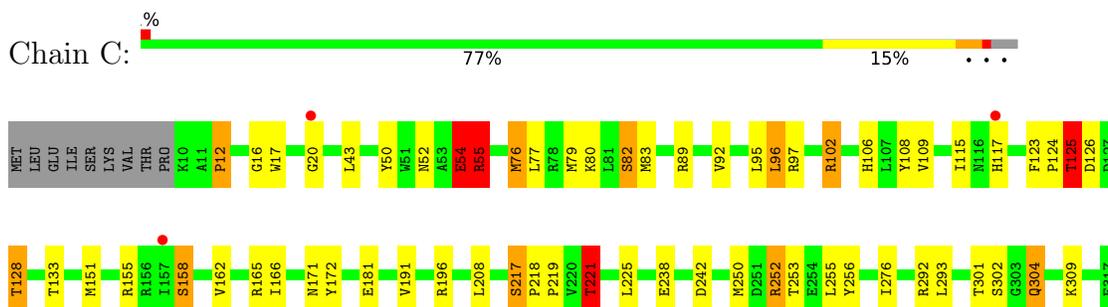
• Molecule 1: Branched-chain-amino-acid aminotransferase
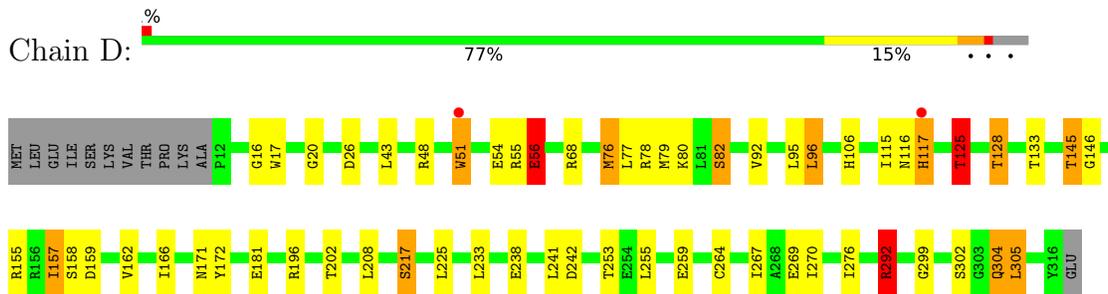
Chain D:

• Molecule 1: Branched-chain-amino-acid aminotransferase
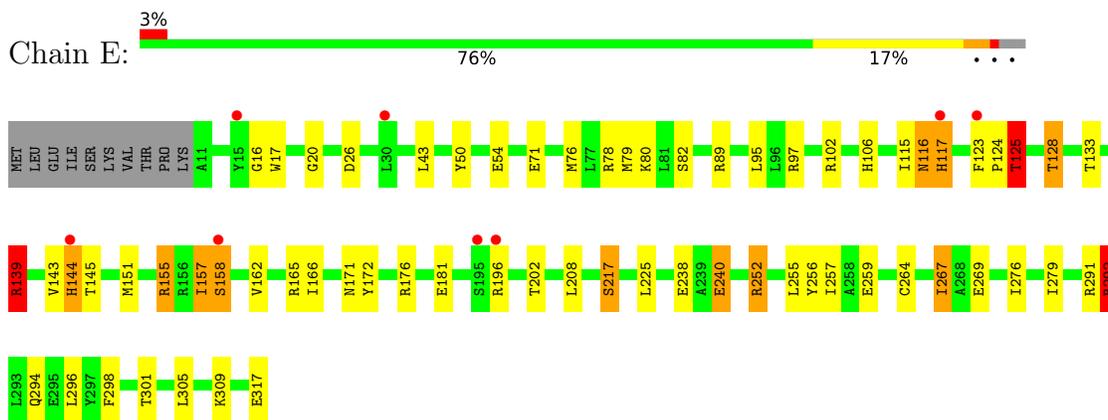
Chain E:

• Molecule 1: Branched-chain-amino-acid aminotransferase

Chain F:
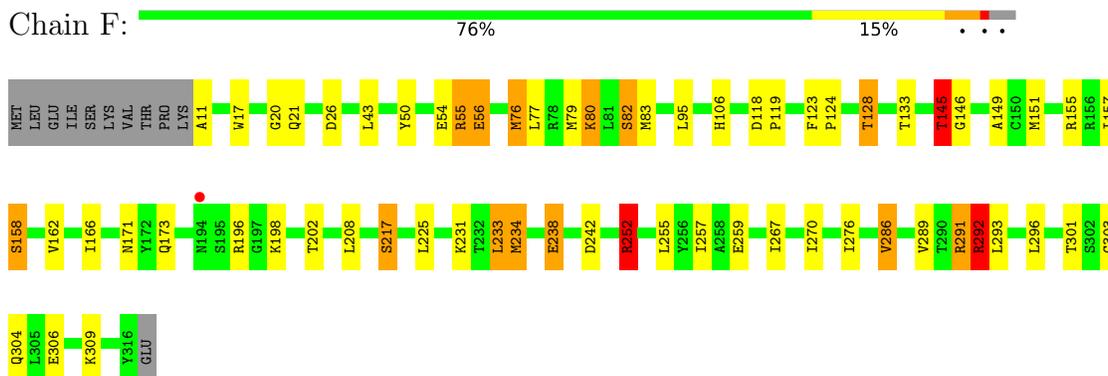
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	95.08Å 164.93Å 254.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.64 – 2.35 47.59 – 2.35	Depositor EDS
% Data completeness (in resolution range)	97.0 (47.64-2.35) 97.0 (47.59-2.35)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.46 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0230	Depositor
R, R_{free}	0.199 , 0.235 0.199 , 0.235	Depositor DCC
R_{free} test set	3978 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	46.5	Xtriage
Anisotropy	0.510	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 33.4	EDS
L-test for twinning[2]	$<|L|> = 0.47, <L^2> = 0.30$	Xtriage
Estimated twinning fraction	0.048 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.048 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	14697	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.25% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PLP

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	0.65	1/2458 (0.0%)	0.99	7/3342 (0.2%)
1	B	0.68	4/2477 (0.2%)	1.09	9/3365 (0.3%)
1	C	0.68	1/2490 (0.0%)	1.08	11/3384 (0.3%)
1	D	0.68	1/2465 (0.0%)	1.04	8/3351 (0.2%)
1	E	0.72	4/2492 (0.2%)	1.07	15/3386 (0.4%)
1	F	0.67	1/2463 (0.0%)	1.07	9/3350 (0.3%)
All	All	0.68	12/14845 (0.1%)	1.06	59/20178 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	B	0	5
1	C	0	5
1	D	0	4
1	E	0	4
1	F	0	4
All	All	0	26

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	181	GLU	CD-OE1	8.27	1.34	1.25
1	D	181	GLU	CD-OE2	8.12	1.34	1.25
1	B	259	GLU	CD-OE2	7.68	1.34	1.25
1	E	181	GLU	CD-OE1	6.57	1.32	1.25
1	E	181	GLU	CD-OE2	6.09	1.32	1.25

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	126	ASP	CB-CA-C	-19.37	71.66	110.40
1	C	126	ASP	CB-CA-C	-18.89	72.62	110.40
1	D	292	ARG	CG-CD-NE	12.80	138.68	111.80
1	F	26	ASP	CB-CA-C	-12.15	86.09	110.40
1	D	56	GLU	CB-CA-C	9.70	129.81	110.40

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	14	ARG	Sidechain
1	A	155	ARG	Sidechain
1	A	183	ARG	Sidechain
1	A	291	ARG	Sidechain
1	B	89	ARG	Sidechain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2401	0	2371	63	0
1	B	2419	0	2391	46	0
1	C	2433	0	2402	60	0
1	D	2408	0	2384	57	0
1	E	2428	0	2395	60	0
1	F	2407	0	2378	68	0
2	A	15	0	7	1	0
2	B	15	0	7	1	0
2	C	15	0	6	1	0
2	D	15	0	6	1	0
2	E	15	0	7	1	0
2	F	15	0	7	0	0
3	A	17	0	0	0	0
3	B	21	0	0	0	0
3	C	25	0	0	1	0
3	D	16	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	19	0	0	1	0
3	F	13	0	0	2	0
All	All	14697	0	14361	316	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 316 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:218:PRO:HD3	1:C:250:MET:CE	1.62	1.29
1:C:218:PRO:CD	1:C:250:MET:CE	2.24	1.16
1:D:106:HIS:CE1	1:D:267:ILE:HD11	1.80	1.15
1:C:106:HIS:HE1	1:C:108:TYR:CZ	1.69	1.11
1:C:218:PRO:CD	1:C:250:MET:HE1	1.78	1.10

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	304/317 (96%)	293 (96%)	10 (3%)	1 (0%)	41	47
1	B	305/317 (96%)	296 (97%)	9 (3%)	0	100	100
1	C	307/317 (97%)	297 (97%)	9 (3%)	1 (0%)	41	47
1	D	304/317 (96%)	296 (97%)	8 (3%)	0	100	100
1	E	307/317 (97%)	297 (97%)	10 (3%)	0	100	100
1	F	305/317 (96%)	294 (96%)	11 (4%)	0	100	100
All	All	1832/1902 (96%)	1773 (97%)	57 (3%)	2 (0%)	51	63

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	55	ARG
1	C	55	ARG

5.3.2   Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	263/276 (95%)	239 (91%)	24 (9%)	9	8
1	B	265/276 (96%)	243 (92%)	22 (8%)	11	11
1	C	266/276 (96%)	246 (92%)	20 (8%)	13	13
1	D	265/276 (96%)	242 (91%)	23 (9%)	10	9
1	E	265/276 (96%)	247 (93%)	18 (7%)	16	16
1	F	263/276 (95%)	239 (91%)	24 (9%)	9	8
All	All	1587/1656 (96%)	1456 (92%)	131 (8%)	12	11

5 of 131 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	133	THR
1	F	202	THR
1	F	309	LYS
1	C	76[A]	MET
1	C	54	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	106	HIS
1	E	106	HIS
1	D	21	GLN
1	C	106	HIS
1	D	106	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	PLP	D	401	1	15,15,16	3.61	5 (33%)	20,22,23	2.65	11 (55%)
2	PLP	E	401	1	15,15,16	2.83	4 (26%)	20,22,23	3.61	11 (55%)
2	PLP	A	401	1	15,15,16	3.37	3 (20%)	20,22,23	1.97	11 (55%)
2	PLP	C	401	1	15,15,16	3.72	7 (46%)	20,22,23	3.82	12 (60%)
2	PLP	B	401	1	15,15,16	3.93	6 (40%)	20,22,23	2.72	10 (50%)
2	PLP	F	401	1	15,15,16	3.00	5 (33%)	20,22,23	3.04	10 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PLP	D	401	1	-	0/6/6/8	0/1/1/1
2	PLP	E	401	1	-	5/6/6/8	0/1/1/1
2	PLP	A	401	1	-	0/6/6/8	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PLP	C	401	1	-	0/6/6/8	0/1/1/1
2	PLP	B	401	1	-	0/6/6/8	0/1/1/1
2	PLP	F	401	1	-	0/6/6/8	0/1/1/1

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	PLP	C3-C2	10.10	1.51	1.40
2	D	401	PLP	C5-C4	10.06	1.51	1.40
2	C	401	PLP	C3-C2	9.81	1.50	1.40
2	B	401	PLP	C5-C4	9.75	1.51	1.40
2	E	401	PLP	C3-C2	8.29	1.49	1.40

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	401	PLP	C4A-C4-C5	11.30	132.58	120.94
2	C	401	PLP	O2P-P-O4P	-7.87	85.80	106.73
2	C	401	PLP	C4A-C4-C5	6.61	127.74	120.94
2	F	401	PLP	C6-C5-C4	6.52	123.29	118.16
2	C	401	PLP	O4P-C5A-C5	5.83	120.45	109.35

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	401	PLP	C4-C5-C5A-O4P
2	E	401	PLP	C6-C5-C5A-O4P
2	E	401	PLP	C5A-O4P-P-O2P
2	E	401	PLP	C5A-O4P-P-O3P
2	E	401	PLP	C5A-O4P-P-O1P

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	PLP	1	0
2	E	401	PLP	1	0
2	A	401	PLP	1	0
2	C	401	PLP	1	0
2	B	401	PLP	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	305/317 (96%)	0.06	3 (0%)	82	88	37, 57, 82, 110	0
1	B	306/317 (96%)	0.06	5 (1%)	72	80	36, 55, 73, 108	0
1	C	308/317 (97%)	0.02	3 (0%)	82	88	36, 53, 78, 121	0
1	D	305/317 (96%)	-0.02	2 (0%)	87	92	35, 54, 81, 114	0
1	E	307/317 (96%)	0.09	8 (2%)	56	65	34, 53, 82, 115	0
1	F	306/317 (96%)	-0.06	1 (0%)	94	97	35, 53, 83, 100	0
All	All	1837/1902 (96%)	0.03	22 (1%)	79	86	34, 54, 81, 121	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	117	HIS	4.9
1	D	117	HIS	4.6
1	E	117	HIS	4.2
1	A	117	HIS	3.9
1	E	144	HIS	3.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	PLP	E	401	15/16	0.97	0.13	41,44,48,48	0
2	PLP	B	401	15/16	0.98	0.10	41,49,53,55	0
2	PLP	C	401	15/16	0.98	0.11	41,47,49,53	0
2	PLP	D	401	15/16	0.98	0.12	45,48,53,53	0
2	PLP	A	401	15/16	0.98	0.11	46,49,52,54	0
2	PLP	F	401	15/16	0.99	0.10	39,42,44,46	0

6.5 Other polymers (i)

There are no such residues in this entry.