



wwPDB X-ray Structure Validation Summary Report i

Apr 8, 2014 – 11:48 PM EDT

PDB ID : 2HOX
Title : alliinase from allium sativum (garlic)
Authors : Shimon, L.J.W.; Rabinkov, A.; Wilcheck, M.; Mirelman, D.; Frolow, F.
Deposited on : 2006-07-17
Resolution : 1.40 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

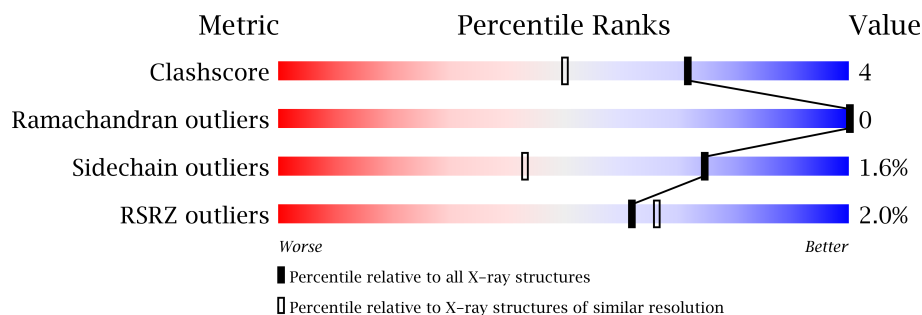
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	dev-1439
EDS	:	stable22978
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22978

1 Overall quality at a glance

The reported resolution of this entry is 1.40 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))
Clashscore	79885	1246 (1.42-1.38)
Ramachandran outliers	78287	1206 (1.42-1.38)
Sidechain outliers	78261	1205 (1.42-1.38)
RSRZ outliers	66119	1097 (1.42-1.38)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	427	
1	B	427	
1	C	427	
1	D	427	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	NAG	A	4500	-	X
3	NAG	B	502	-	X
3	NAG	B	503	-	X

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 16852 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Alliin lyase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	425	Total	C	N	O	S	0	1	0
			3436	2190	577	646	23			
1	B	427	Total	C	N	O	S	0	0	0
			3447	2197	582	645	23			
1	C	425	Total	C	N	O	S	0	1	0
			3445	2200	579	643	23			
1	D	427	Total	C	N	O	S	0	0	0
			3447	2197	582	645	23			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	176	ASP	ASN	SEE REMARK 999	UNP Q01594
B	176	ASP	ASN	SEE REMARK 999	UNP Q01594
C	176	ASP	ASN	SEE REMARK 999	UNP Q01594
D	176	ASP	ASN	SEE REMARK 999	UNP Q01594

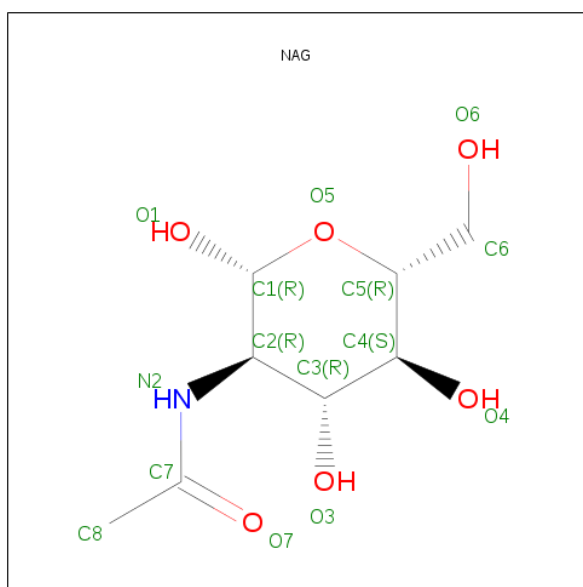
- Molecule 2 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	6	Total	C	N	O	0	0
			69	39	2	28		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	176	ASP	ASN	SEE REMARK 999	UNP Q01594

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	2	Total	C	N	O	0	0
			24	14	1	9		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	176	ASP	ASN	SEE REMARK 999	UNP Q01594

- Molecule 5 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	4	Total	C	N	O	0	0
			49	28	2	19		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	176	ASP	ASN	SEE REMARK 999	UNP Q01594

- Molecule 6 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	3	Total	C	N	O	0	0
			38	22	2	14		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	176	ASP	ASN	SEE REMARK 999	UNP Q01594

- Molecule 7 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	D	2	Total	C	N	O	0	0
			28	16	2	10		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	176	ASP	ASN	SEE REMARK 999	UNP Q01594

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

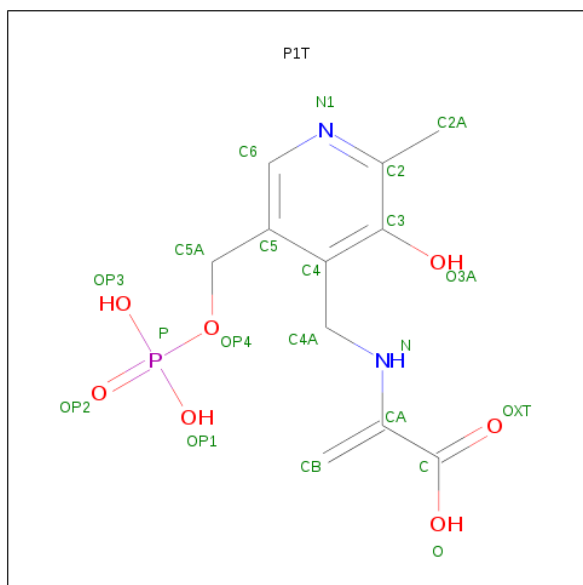
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total	Cl	0	0
			1	1		
8	A	1	Total	Cl	0	0
			1	1		
8	D	1	Total	Cl	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	1	Total	Cl	0	0
			1	1		

- Molecule 9 is 2-[(3-HYDROXY-2-METHYL-5-[(PHOSPHONOOXY)METHYL]PYRIDIN-4-YL)METHYL)AMINO]ACRYLICACID (three-letter code: P1T) (formula: C₁₁H₁₅N₂O₇P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	A	1	Total	C	N	O	P	0	0
			21	11	2	7	1		
9	B	1	Total	C	N	O	P	0	0
			21	11	2	7	1		
9	C	1	Total	C	N	O	P	0	0
			21	11	2	7	1		
9	D	1	Total	C	N	O	P	0	0
			21	11	2	7	1		

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	752	Total	O	0	0
			752	752		
10	B	685	Total	O	0	0
			685	685		
10	C	774	Total	O	0	0
			774	774		

Continued on next page...

Continued from previous page...

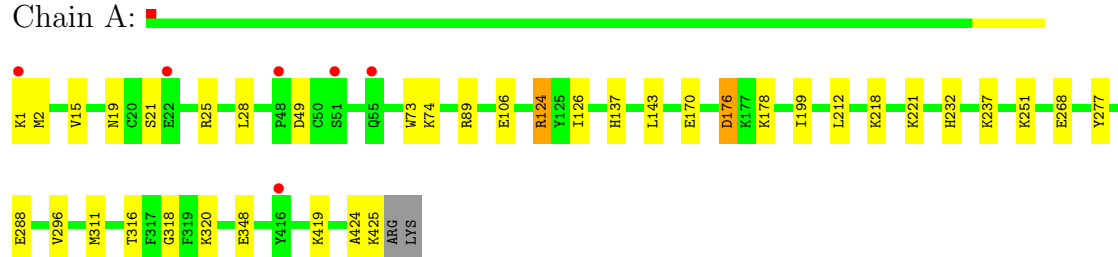
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	458	Total 458	O 458	0	2

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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.

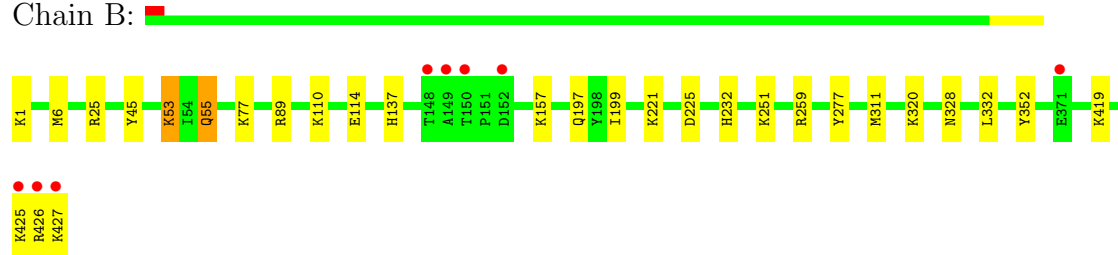
• Molecule 1: Alliin lyase 1

Chain A:



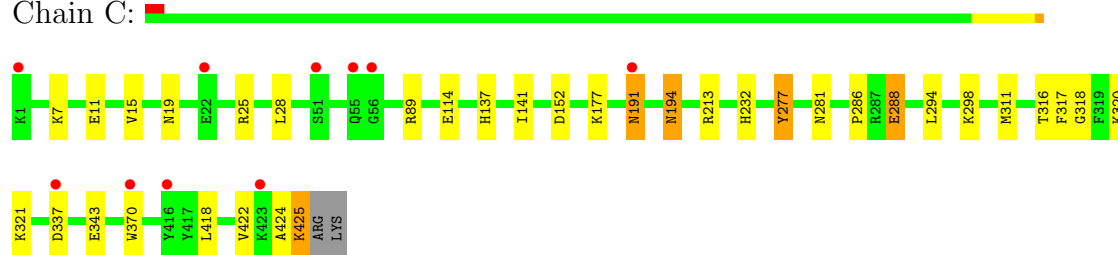
• Molecule 1: Alliin lyase 1

Chain B:



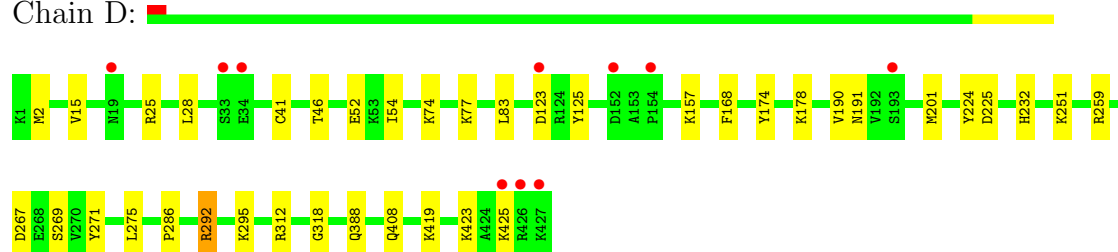
• Molecule 1: Alliin lyase 1

Chain C:



• Molecule 1: Alliin lyase 1

Chain D:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	67.67Å 126.89Å 102.66Å 90.00° 97.30° 90.00°	Depositor
Resolution (Å)	46.10 – 1.40 46.11 – 1.40	Depositor EDS
% Data completeness (in resolution range)	96.5 (46.10-1.40) 96.5 (46.11-1.40)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.50 (at 1.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.167 , 0.205 0.169 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	16.1	Xtriage
Anisotropy	0.132	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 42.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 327315 reflections (0.000%)	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	16852	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: XYP, BMA, NAG, CL, P1T, FUC

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	0.87	2/3525 (0.1%)	0.86	2/4772 (0.0%)
1	B	0.87	2/3536 (0.1%)	0.87	2/4785 (0.0%)
1	C	0.87	3/3543 (0.1%)	0.84	2/4793 (0.0%)
1	D	0.87	2/3536 (0.1%)	0.84	1/4785 (0.0%)
All	All	0.87	9/14140 (0.1%)	0.85	7/19135 (0.0%)

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	C	0	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	224	TYR	CE2-CZ	-6.82	1.29	1.38
1	C	277	TYR	CE1-CZ	-6.80	1.29	1.38
1	A	277	TYR	CE1-CZ	-6.28	1.30	1.38
1	C	277	TYR	CD2-CE2	5.96	1.48	1.39
1	A	288	GLU	CB-CG	-5.71	1.41	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	259	ARG	NE-CZ-NH2	7.51	124.06	120.30
1	B	89	ARG	NE-CZ-NH2	-7.07	116.77	120.30
1	C	89	ARG	NE-CZ-NH1	6.46	123.53	120.30
1	C	213	ARG	NE-CZ-NH2	-5.90	117.35	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	124	ARG	NE-CZ-NH2	-5.73	117.43	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	424	ALA	Peptide

5.2 Close contacts

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3436	0	3325	28	0
1	B	3447	0	3347	19	0
1	C	3445	0	3360	40	0
1	D	3447	0	3347	31	0
2	A	69	0	59	2	0
3	A	42	0	39	1	0
3	B	42	0	39	4	0
3	C	28	0	26	1	0
4	B	24	0	22	0	0
5	C	49	0	43	1	0
6	D	38	0	34	0	0
7	D	28	0	25	0	0
8	A	1	0	0	0	0
8	B	1	0	0	0	0
8	C	1	0	0	0	0
8	D	1	0	0	0	0
9	A	21	0	11	2	0
9	B	21	0	11	3	0
9	C	21	0	11	0	0
9	D	21	0	12	4	0
10	A	752	0	0	14	1
10	B	685	0	0	7	2
10	C	774	0	0	17	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	D	458	0	0	11	0
All	All	16852	0	13711	123	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 123 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:41:CYS:SG	10:D:6250:HOH:O	1.94	1.21
1:C:277:TYR:CE2	10:C:6486:HOH:O	1.98	1.12
1:C:277:TYR:CZ	10:C:6486:HOH:O	2.02	1.09
1:B:328:ASN:HD21	3:B:502:NAG:C1	1.78	0.95
1:B:137:HIS:HD2	10:B:6599:HOH:O	1.49	0.94

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
10:B:6436:HOH:O	10:C:6361:HOH:O[1_455]	1.99	0.21
10:A:6664:HOH:O	10:B:6132:HOH:O[2_645]	2.11	0.09

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	424/427 (99%)	411 (97%)	13 (3%)	0	100	100
1	B	425/427 (100%)	413 (97%)	12 (3%)	0	100	100
1	C	426/427 (100%)	414 (97%)	12 (3%)	0	100	100
1	D	425/427 (100%)	412 (97%)	13 (3%)	0	100	100
All	All	1700/1708 (100%)	1650 (97%)	50 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	374/375 (100%)	369 (99%)	5 (1%)	80	52
1	B	375/375 (100%)	366 (98%)	9 (2%)	61	23
1	C	376/375 (100%)	369 (98%)	7 (2%)	69	33
1	D	375/375 (100%)	371 (99%)	4 (1%)	84	59
All	All	1500/1500 (100%)	1475 (98%)	25 (2%)	75	40

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	425	LYS
1	B	427	LYS
1	D	178	LYS
1	B	426	ARG
1	C	7	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	381	GLN
1	B	386	ASN
1	D	137	HIS
1	B	328	ASN
1	D	191	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

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(Å ²)	Q<0.9
1	A	425/427 (99%)	-0.13	6 (1%) 72 77	10, 19, 37, 53	1 (0%)
1	B	427/427 (100%)	-0.13	8 (1%) 64 68	11, 18, 35, 65	1 (0%)
1	C	425/427 (99%)	-0.16	10 (2%) 56 60	10, 18, 33, 52	0
1	D	427/427 (100%)	-0.11	10 (2%) 57 62	11, 18, 36, 69	0
All	All	1704/1708 (99%)	-0.13	34 (1%) 62 66	10, 18, 36, 69	2 (0%)

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	LYS	7.9
1	D	427	LYS	7.6
1	B	427	LYS	7.0
1	D	152	ASP	5.6
1	B	425	LYS	4.7

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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	BMA	A	505	11/12	0.12	15.76	20,23,33,43	0
2	FUC	A	501	10/11	0.18	12.76	39,44,50,50	0
5	NAG	C	502	14/15	0.13	8.67	30,36,51,60	0
7	NAG	D	503	14/15	0.22	4.92	31,38,54,61	0
4	FUC	B	501	10/11	0.19	3.67	46,51,55,56	0
2	XYP	A	504	9/10	0.15	2.00	20,21,22,24	0
2	NAG	A	500	14/15	0.15	1.94	26,32,36,36	0
6	NAG	D	500	14/15	0.15	1.26	28,33,38,39	0
5	NAG	C	500	14/15	0.08	0.81	24,27,33,34	0
5	FUC	C	501	10/11	0.09	0.48	24,29,32,33	0
4	NAG	B	500	14/15	0.17	0.46	32,40,48,51	0
2	NAG	A	502	14/15	0.12	0.24	26,31,47,50	0
6	NAG	D	502	14/15	0.23	-	40,46,48,54	0
5	BMA	C	503	11/12	0.23	-	69,71,74,76	0
7	NAG	D	504	14/15	0.40	-	68,72,75,76	0
2	BMA	A	503	11/12	0.13	-	18,23,31,39	0
6	FUC	D	501	10/11	0.16	-	36,41,50,52	0

6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	B	503	14/15	0.59	25.95	128,132,132,133	0
3	NAG	A	4500	14/15	0.35	9.57	64,70,79,80	0
3	NAG	B	502	14/15	0.20	3.18	48,59,69,69	0
3	NAG	A	506	14/15	0.16	1.97	32,42,55,61	0
3	NAG	C	504	14/15	0.14	1.88	27,33,44,44	0
9	P1T	D	6004	21/21	0.09	0.50	12,17,25,27	0
9	P1T	A	6001	21/21	0.08	0.43	13,17,29,29	0
8	CL	C	5003	1/1	0.07	0.19	13,13,13,13	0
3	NAG	B	3500	14/15	0.19	0.19	42,52,55,59	0
8	CL	B	5002	1/1	0.07	-0.15	16,16,16,16	0
9	P1T	B	6002	21/21	0.08	-0.17	12,17,24,25	0
9	P1T	C	6003	21/21	0.07	-0.29	13,17,25,28	0
8	CL	A	5001	1/1	0.06	-1.09	12,12,12,12	0
8	CL	D	5004	1/1	0.05	-1.53	15,15,15,15	0
3	NAG	A	507	14/15	0.36	-	104,110,112,113	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	C	505	14/15	0.28	-	73,82,84,84	0

6.5 Other polymers ⓘ

There are no such residues in this entry.