



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 02:11 PM EDT

PDB ID : 6XWT
Title : drosophila melanogaster CENP-A/H4 bound to N-terminal CAL1 fragment
Authors : Jeyaprakash, A.A.; Medina-Pritchard, B.; Lazou, V.; Zou, J.; Byron, O.; Abad, M.A.; Rappsilber, J.; Heun, P.
Deposited on : 2020-01-24
Resolution : 3.47 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
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

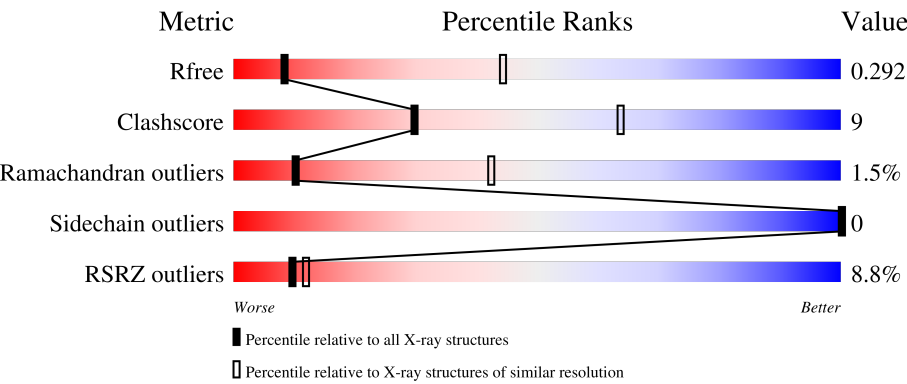
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.47 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	1379 (3.56-3.40)
Clashscore	141614	1461 (3.56-3.40)
Ramachandran outliers	138981	1424 (3.56-3.40)
Sidechain outliers	138945	1425 (3.56-3.40)
RSRZ outliers	127900	1289 (3.56-3.40)

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 $\leq 5\%$. 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	225	<div><div>2%</div><div>24%</div><div>8%</div><div>67%</div></div>
1	C	225	<div><div>2%</div><div>24%</div><div>9%</div><div>67%</div></div>
2	B	103	<div><div>11%</div><div>55%</div><div>15%</div><div>30%</div></div>
2	D	103	<div><div>3%</div><div>53%</div><div>14%</div><div>31%</div></div>
3	E	979	<div><div>97%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	979	<div><div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>..</div></div><div>97%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2803 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 Histone H3-like centromeric protein cid.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	74	Total	C	N	O	S	0	0	0
			599	380	103	109	7			
1	C	74	Total	C	N	O	S	0	0	0
			594	377	103	107	7			

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	72	Total	C	N	O	S	0	0	0
			569	360	110	98	1			
2	D	71	Total	C	N	O	S	0	0	0
			564	358	108	97	1			

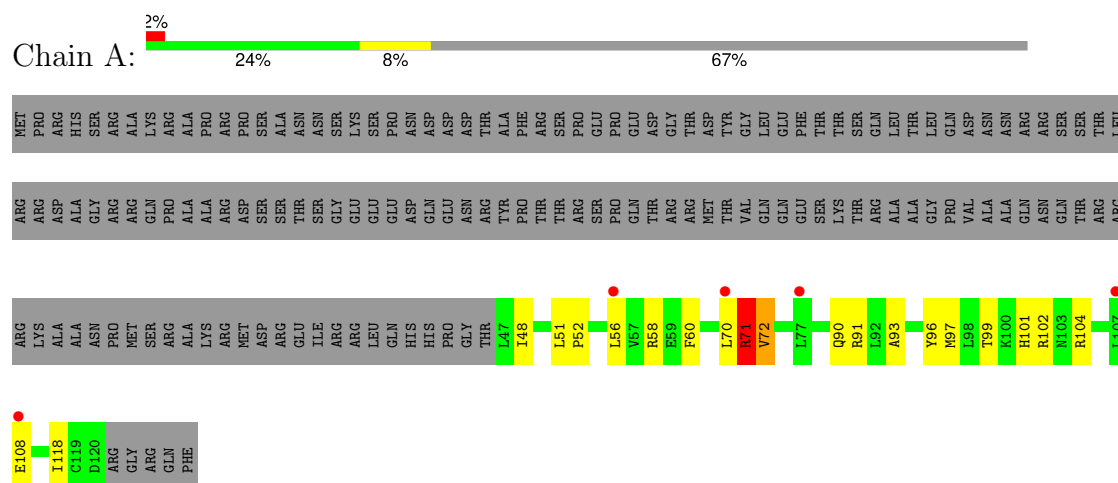
- Molecule 3 is a protein called Chromosome alignment defect 1.

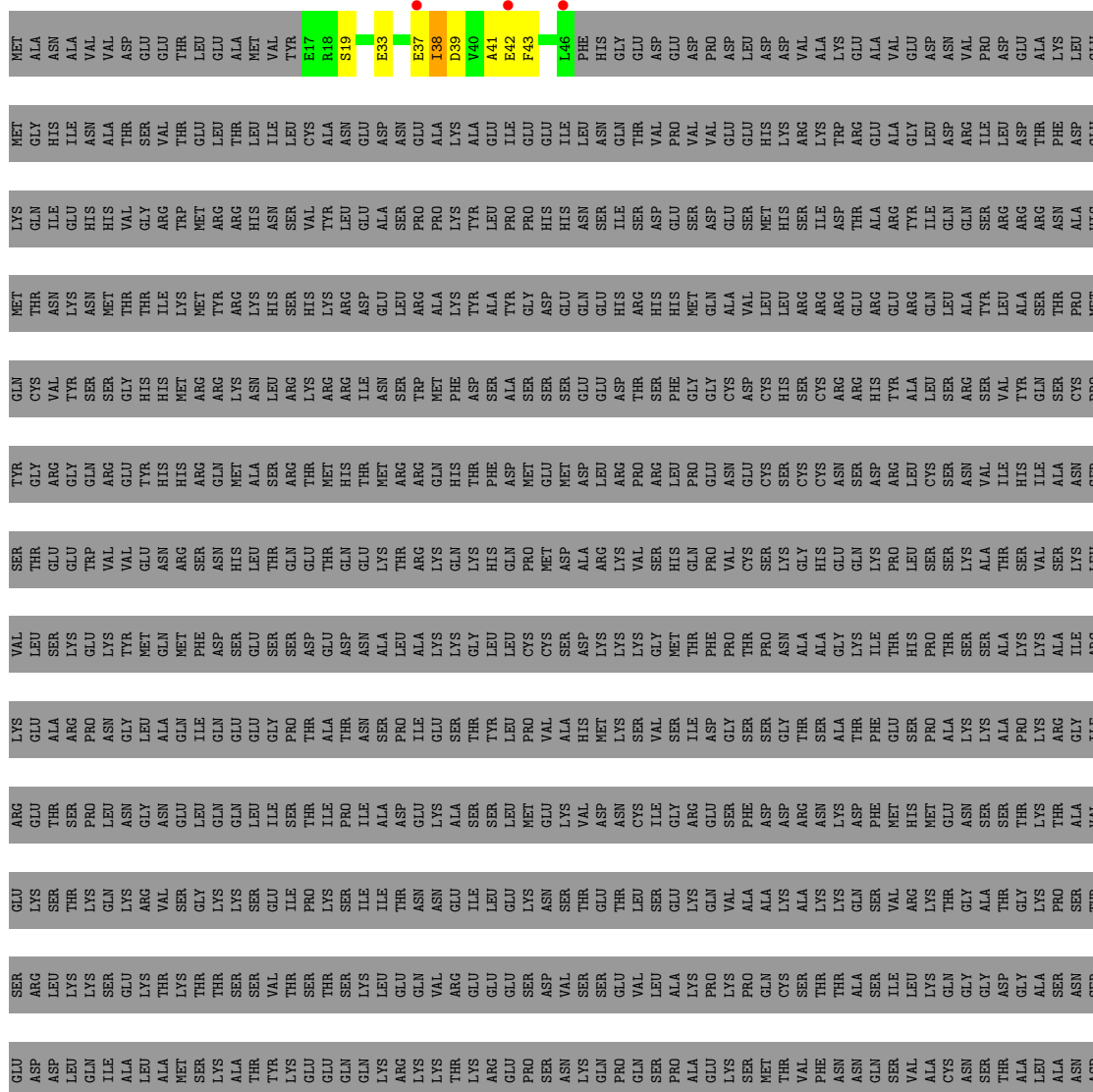
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	30	Total	C	N	O	S	0	0	0
			236	146	37	51	2			
3	F	31	Total	C	N	O	S	0	0	0
			241	149	38	52	2			

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.

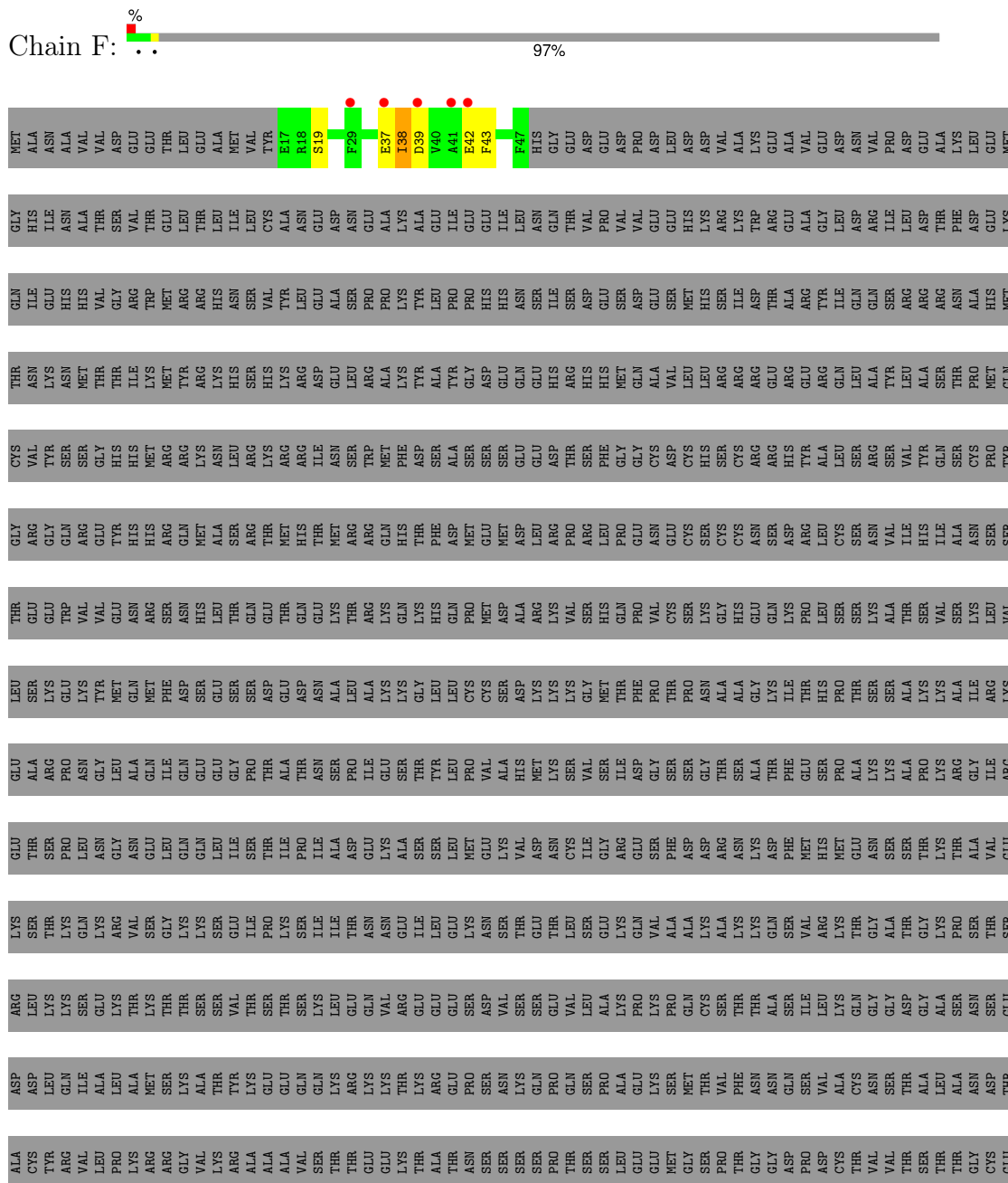
- Molecule 1: Histone H3-like centromeric protein cid





[illegible]

- Molecule 3: Chromosome alignment defect 1



GLY ASP LYS	PRO	PRO	ALA	SER	GLU	ARG	GLN	GLU	ILE	PRO	PRO	ALA	THR	THR	ILE	LYS	ILE	THR	LYS	ARG	GLY	ILE	LEU	LEU	HIS	SER	PRO	PRO	ALA	SER	SER	PHE	THR	THR	THR	GLU	GLN	GLY	GLY	ILE	ILE	GLY	GLU	ARG	TRP	ALA	ALA	ARG	LYS	TYR	LEU	LYS	TYR	HIS	ILE	GLY	SER	ARG				
	SER	PHE	ASP	SER	ARG	HIS	SER	VAL	TYR	TYR	GLN	THR	THR	PRO	PRO	GLN	LEU	ALA	ALA	ALA	LEU	SER	ALA	PRO	PRO	GLN	ASP	ALA	ALA	GLN	ASN	ILE	GLY	ASN	ASN	ILE	SER	SER	SER	ALA	SER	ASP	ASP	ASP	ILE	PHE	GLU	GLN	ILE	ASN	ARG	TRP	TYR	GLY	THR	VAL	TYR	SER	ILE	LEU	GLU	ASN

4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	178.25Å 178.25Å 133.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	100.88 – 3.47 100.88 – 3.47	Depositor EDS
% Data completeness (in resolution range)	99.9 (100.88-3.47) 92.0 (100.88-3.47)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.05 (at 3.49Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.272 , 0.286 0.277 , 0.292	Depositor DCC
R_{free} test set	542 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	125.0	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 108.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2803	wwPDB-VP
Average B, all atoms (Å ²)	136.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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

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	0.27	0/608	0.53	0/819
1	C	0.29	0/603	0.55	0/812
2	B	0.27	0/574	0.48	0/770
2	D	0.29	0/570	0.53	0/765
3	E	0.26	0/239	0.48	0/319
3	F	0.28	0/244	0.52	0/326
All	All	0.28	0/2838	0.52	0/3811

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	2
1	C	0	1
2	D	0	1
3	E	0	2
3	F	0	1
All	All	0	7

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	70	LEU	Peptide
1	A	71	ARG	Peptide
1	C	70	LEU	Peptide
2	D	31	THR	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
3	E	33	GLU	Peptide
3	E	38	ILE	Peptide
3	F	38	ILE	Peptide

5.2 Too-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 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	599	0	608	16	0
1	C	594	0	599	20	0
2	B	569	0	615	13	0
2	D	564	0	605	15	0
3	E	236	0	215	7	0
3	F	241	0	217	6	0
All	All	2803	0	2859	53	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 9.

All (53) 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:A:90:GLN:NE2	2:B:98:LEU:O	2.21	0.73
1:C:90:GLN:NE2	2:D:98:LEU:O	2.22	0.72
1:C:65:SER:OG	1:C:68:GLU:O	2.11	0.69
2:B:98:LEU:HD22	3:F:19:SER:HA	1.75	0.69
2:B:27:ILE:HG23	2:B:56:ARG:HB3	1.75	0.68
1:C:58:ARG:HB2	3:E:43:PHE:CE2	2.32	0.64
2:B:32:LYS:HG3	2:B:52:TYR:CE2	2.34	0.63
2:D:32:LYS:H	2:D:33:PRO:HD2	1.63	0.61
1:A:71:ARG:HB3	3:F:37:GLU:O	2.01	0.61
1:A:48:ILE:HG22	2:B:34:ALA:HB1	1.83	0.60
1:A:58:ARG:HB2	3:F:43:PHE:HE2	1.66	0.60
3:F:39:ASP:N	3:F:39:ASP:OD1	2.38	0.57
1:A:101:HIS:CG	1:C:114:LEU:HD22	2.40	0.56
1:A:60:PHE:HZ	2:B:60:LYS:HG3	1.74	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:58:ARG:HB2	3:E:43:PHE:HE2	1.72	0.52
1:C:91:ARG:HD3	2:D:58:VAL:HG11	1.93	0.50
3:E:39:ASP:OD1	3:E:39:ASP:N	2.44	0.50
1:C:56:LEU:HG	1:C:60:PHE:CE2	2.46	0.50
1:C:61:ILE:O	1:C:65:SER:HB2	2.13	0.48
1:A:71:ARG:O	1:A:72:VAL:HG23	2.14	0.48
1:C:52:PRO:O	1:C:55:ARG:HB3	2.14	0.48
1:C:58:ARG:HB2	3:E:43:PHE:CD2	2.50	0.47
1:C:90:GLN:NE2	2:D:97:THR:O	2.47	0.47
1:A:96:TYR:HB2	2:B:44:VAL:HG22	1.97	0.46
1:C:96:TYR:HB2	2:D:44:VAL:HG22	1.97	0.46
2:D:98:LEU:HD22	3:E:19:SER:HA	1.97	0.46
1:A:58:ARG:HB2	3:F:43:PHE:CE2	2.48	0.46
1:C:48:ILE:HG22	2:D:34:ALA:HB1	1.99	0.45
2:B:31:THR:HG22	2:B:33:PRO:HD2	1.99	0.44
1:C:99:THR:HG22	1:C:104:ARG:O	2.17	0.44
1:A:102:ARG:NH2	1:A:108:GLU:HG2	2.34	0.43
2:B:84:ALA:O	2:B:88:VAL:HG23	2.19	0.43
1:C:102:ARG:NH2	1:C:108:GLU:HG2	2.33	0.43
3:E:38:ILE:O	3:E:41:ALA:N	2.51	0.43
3:F:38:ILE:H	3:F:42:GLU:HG2	1.84	0.43
2:B:32:LYS:HG3	2:B:52:TYR:CD2	2.54	0.42
1:A:91:ARG:HD3	2:B:58:VAL:HG11	2.01	0.42
2:B:33:PRO:HB3	2:B:36:ARG:NH2	2.34	0.42
1:A:56:LEU:HG	1:A:60:PHE:CE2	2.55	0.42
1:C:72:VAL:HA	2:D:82:VAL:HB	2.02	0.42
1:C:71:ARG:HB2	2:D:81:THR:HG23	2.00	0.42
2:D:65:ASN:O	2:D:68:ARG:HB3	2.20	0.42
3:E:37:GLU:CD	3:E:42:GLU:HG2	2.40	0.42
1:A:118:ILE:HG12	2:D:99:TYR:CE2	2.55	0.41
1:A:99:THR:HG22	1:A:104:ARG:O	2.20	0.41
1:C:108:GLU:HA	2:D:51:ILE:HD11	2.01	0.41
2:D:31:THR:OG1	2:D:32:LYS:N	2.53	0.41
1:A:93:ALA:O	1:A:97:MET:HG2	2.20	0.41
1:C:99:THR:OG1	1:C:111:ASP:OD2	2.26	0.41
1:A:51:LEU:HB2	1:A:52:PRO:HD3	2.02	0.41
1:C:72:VAL:H	2:D:82:VAL:H	1.69	0.40
2:B:33:PRO:O	2:B:37:ARG:HG3	2.22	0.40
2:D:36:ARG:O	2:D:40:ARG:HG2	2.21	0.40

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	72/225 (32%)	67 (93%)	3 (4%)	2 (3%)	5	30
1	C	72/225 (32%)	66 (92%)	5 (7%)	1 (1%)	11	43
2	B	70/103 (68%)	67 (96%)	3 (4%)	0	100	100
2	D	69/103 (67%)	62 (90%)	5 (7%)	2 (3%)	4	29
3	E	28/979 (3%)	24 (86%)	4 (14%)	0	100	100
3	F	29/979 (3%)	21 (72%)	8 (28%)	0	100	100
All	All	340/2614 (13%)	307 (90%)	28 (8%)	5 (2%)	10	42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	31	THR
1	A	71	ARG
1	A	72	VAL
2	D	32	LYS
1	C	72	VAL

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	66/198 (33%)	66 (100%)	0	100	100
1	C	64/198 (32%)	64 (100%)	0	100	100
2	B	59/79 (75%)	59 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	58/79 (73%)	58 (100%)	0	100	100
3	E	25/861 (3%)	25 (100%)	0	100	100
3	F	25/861 (3%)	25 (100%)	0	100	100
All	All	297/2276 (13%)	297 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

There are no ligands in this entry.

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(Å ²)	Q < 0.9
1	A	74/225 (32%)	1.09	5 (6%)	17 18	108, 132, 156, 163	0
1	C	74/225 (32%)	0.83	4 (5%)	25 25	111, 129, 157, 163	0
2	B	72/103 (69%)	1.04	11 (15%)	2 3	104, 132, 159, 169	0
2	D	71/103 (68%)	0.72	3 (4%)	36 34	102, 127, 155, 175	0
3	E	30/979 (3%)	0.89	3 (10%)	7 9	119, 144, 195, 214	0
3	F	31/979 (3%)	0.80	5 (16%)	1 2	134, 153, 183, 191	0
All	All	352/2614 (13%)	0.91	31 (8%)	10 12	102, 133, 166, 214	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	27	ILE	4.9
3	E	42	GLU	4.2
2	B	35	ILE	4.0
1	A	70	LEU	3.9
2	B	47	ILE	3.2
1	A	107	LEU	3.2
2	B	38	LEU	3.1
2	D	30	ILE	2.9
1	C	57	VAL	2.8
3	F	41	ALA	2.8
1	A	56	LEU	2.8
1	A	77	LEU	2.7
2	B	82	VAL	2.6
3	F	37	GLU	2.5
2	B	29	GLY	2.5
2	B	31	THR	2.3
3	F	29	PHE	2.3
1	C	59	GLU	2.3
2	D	92	LYS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	F	42	GLU	2.2
1	C	81	GLN	2.2
3	E	46	LEU	2.2
2	D	67	ILE	2.2
3	E	37	GLU	2.2
2	B	78	LYS	2.1
3	F	39	ASP	2.1
2	B	59	LEU	2.0
2	B	55	THR	2.0
1	C	70	LEU	2.0
2	B	64	GLU	2.0
1	A	108	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.