wwPDB X-ray Structure Validation Summary Report (i)

Nov 13, 2024 – 08:58 AM EST

PDB ID	:	4LD9
Title	:	Crystal structure of the N-terminally acetylated BAH domain of Sir3 bound to the nucleosome core particle
Authors	:	Arnaudo, N.; Fernandez, I.S.; McLaughlin, S.H.; Peak-Chew, S.Y.; Rhodes, D.; Martino, F.
Deposited on	:	2013-06-24
Resolution	:	3.31 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) 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 (i)) were used in the production of this report:

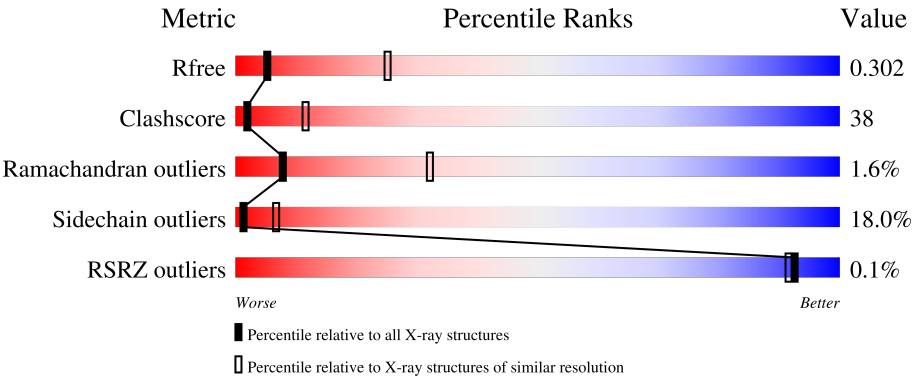
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.31 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}	164625	1085 (3.32-3.28)
Clashscore	180529	1128 (3.32-3.28)
Ramachandran outliers	177936	1125 (3.32-3.28)
Sidechain outliers	177891	1124 (3.32-3.28)
RSRZ outliers	164620	1085 (3.32-3.28)

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	136	40% 26% · 29%
1	E	136	37% 22% 5% · 34%
2	B	103	42% 38% 7% 14%
2	F	103	28% 36% 11% 6% 19%
3	C	130	% 27% 40% 8% 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	130	23% 44% 8% 23%
4	D	126	13% 44% 11% 32%
4	H	126	29% 32% 9% 30%
5	I	167	25% 48% 13% 14%
6	J	167	22% 50% 13% 15%
7	K	236	28% 45% 12% 5% 11%
7	L	236	36% 41% 8% 13%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 15140 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 H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	97	Total	C	N	O	S	0	0	0
			799	505	153	138	3			
1	E	90	Total	C	N	O	S	0	0	0
			729	459	137	130	3			

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	89	Total	C	N	O	S	0	0	0
			716	451	145	119	1			
2	F	83	Total	C	N	O	S	0	0	0
			674	426	134	113	1			

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	98	Total	C	N	O	0	0	0
			753	476	146	131			
3	G	100	Total	C	N	O	0	0	0
			768	485	148	135			

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	86	Total	C	N	O	S	0	0	0
			673	427	119	125	2			
4	H	88	Total	C	N	O	S	0	0	0
			690	437	122	129	2			

- Molecule 5 is a DNA chain called Widom 601 sequence reverse.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	143	Total	C	N	O	P	0	0	0
			2946	1396	551	857	142			

- Molecule 6 is a DNA chain called Widom 601 sequence forward.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	142	Total	C	N	O	P	0	0	0
			2895	1375	527	851	142			

- Molecule 7 is a protein called Regulatory protein SIR3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	K	211	Total	C	N	O	S	0	0	0
			1761	1136	295	328	2			
7	L	206	Total	C	N	O	S	0	0	0
			1736	1122	284	328	2			

There are 16 discrepancies between the modelled and reference sequences:

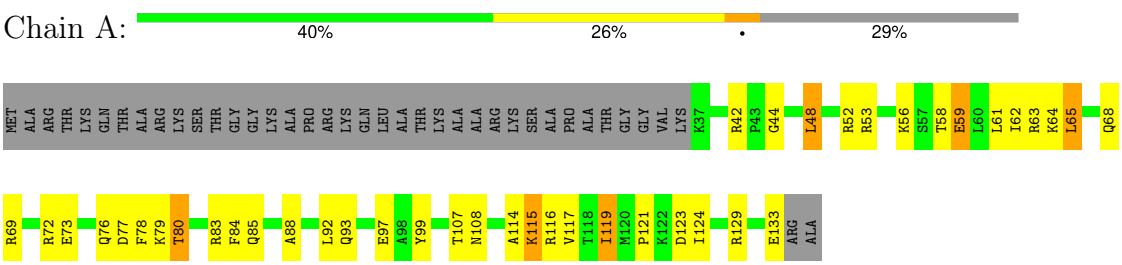
Chain	Residue	Modelled	Actual	Comment	Reference
K	1	ACE	-	acetylation	UNP P06701
K	230	LYS	-	expression tag	UNP P06701
K	231	HIS	-	expression tag	UNP P06701
K	232	HIS	-	expression tag	UNP P06701
K	233	HIS	-	expression tag	UNP P06701
K	234	HIS	-	expression tag	UNP P06701
K	235	HIS	-	expression tag	UNP P06701
K	236	HIS	-	expression tag	UNP P06701
L	1	ACE	-	acetylation	UNP P06701
L	230	LYS	-	expression tag	UNP P06701
L	231	HIS	-	expression tag	UNP P06701
L	232	HIS	-	expression tag	UNP P06701
L	233	HIS	-	expression tag	UNP P06701
L	234	HIS	-	expression tag	UNP P06701
L	235	HIS	-	expression tag	UNP P06701
L	236	HIS	-	expression tag	UNP P06701

3 Residue-property plots (i)

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.

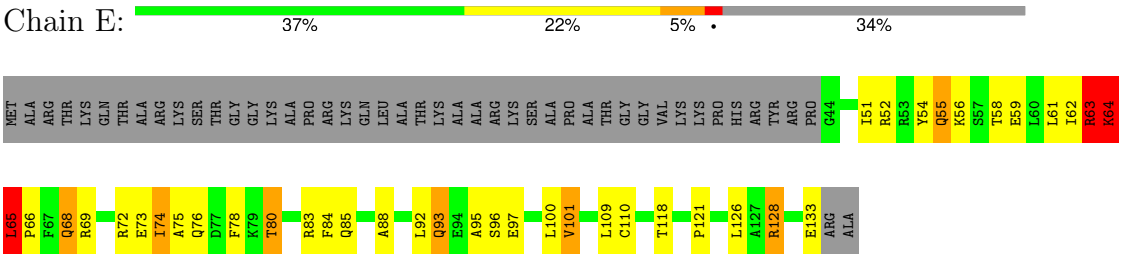
• Molecule 1: Histone H3.2

Chain A:



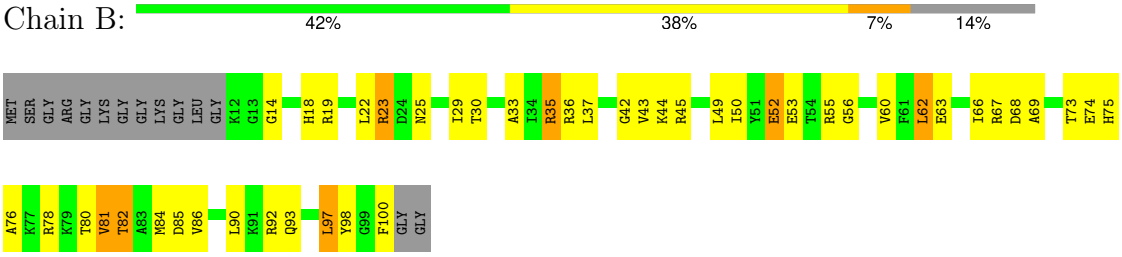
• Molecule 1: Histone H3.2

Chain E:



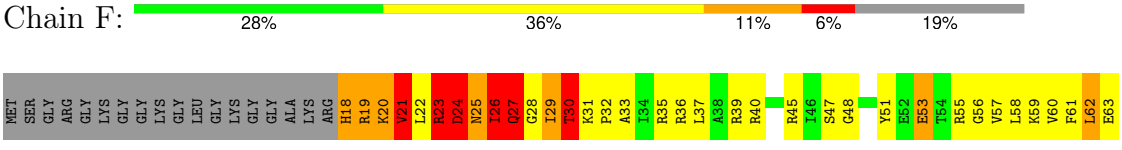
• Molecule 2: Histone H4

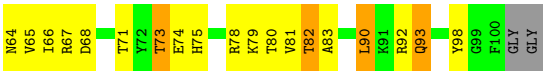
Chain B:



• Molecule 2: Histone H4

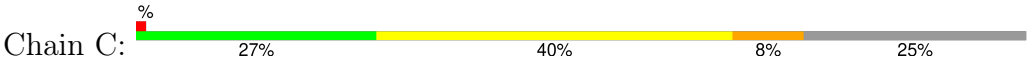
Chain F:

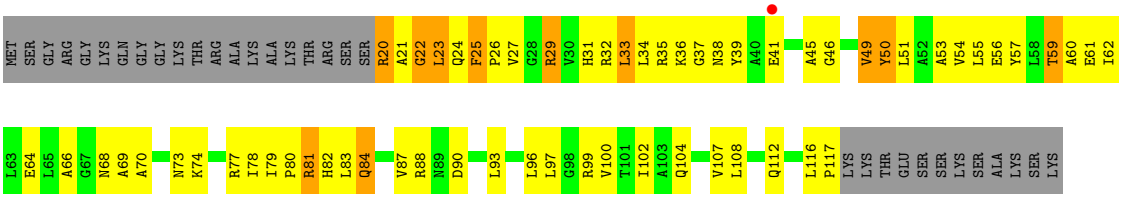




• Molecule 3: Histone H2A

Chain C:



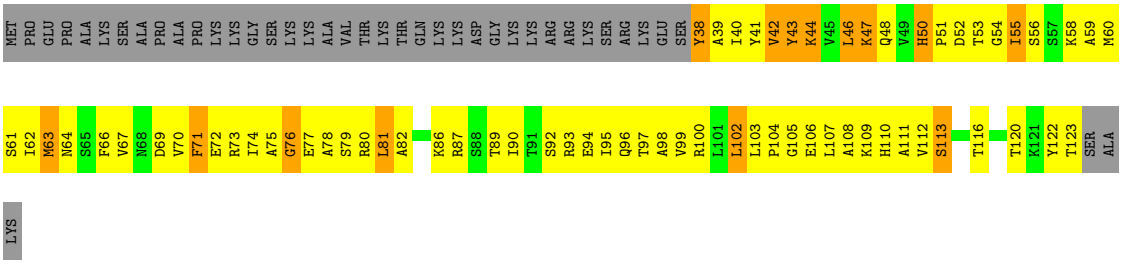


• Molecule 3: Histone H2A

Chain G:

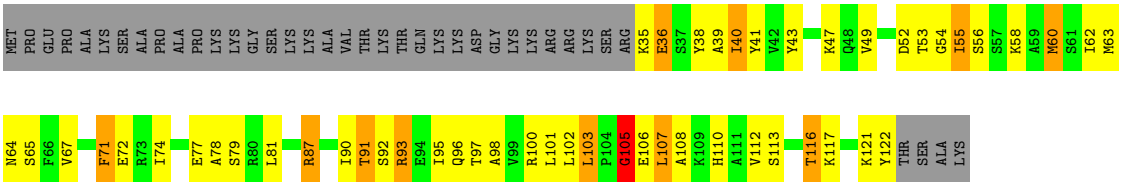
• Molecule 4: Histone H2B 1.1

Chain D:



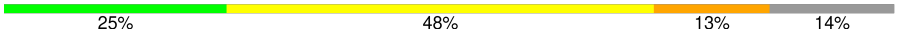
• Molecule 4: Histone H2B 1.1

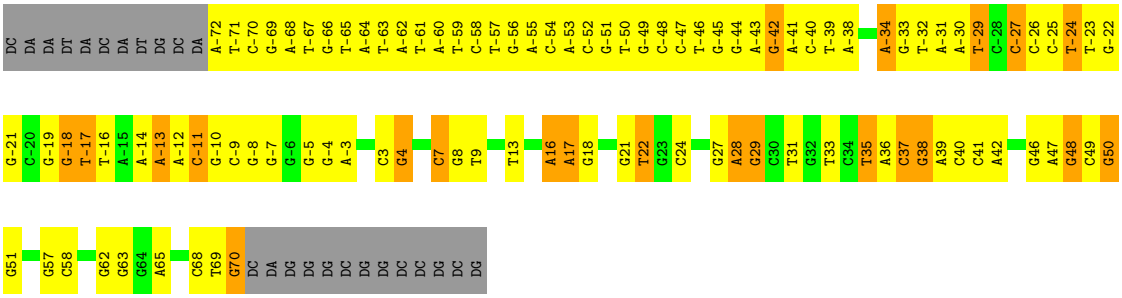
Chain H:



• Molecule 5: Widom 601 sequence reverse

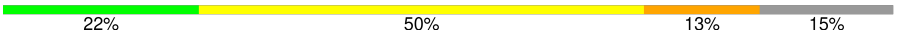
Chain I:

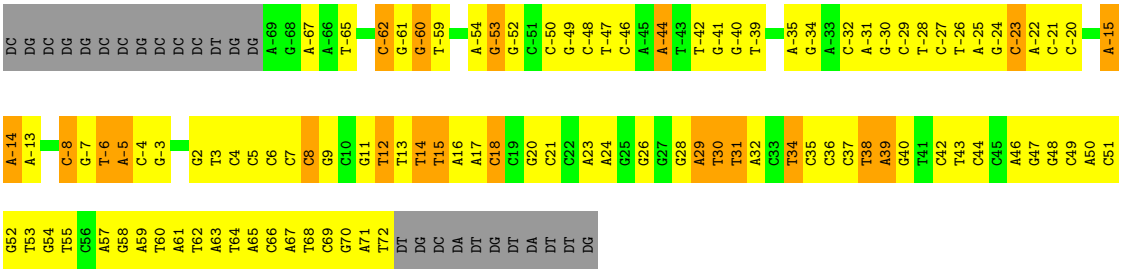




• Molecule 6: Widom 601 sequence forward

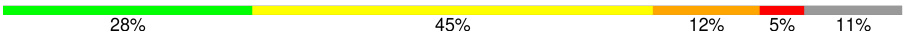
Chain J:

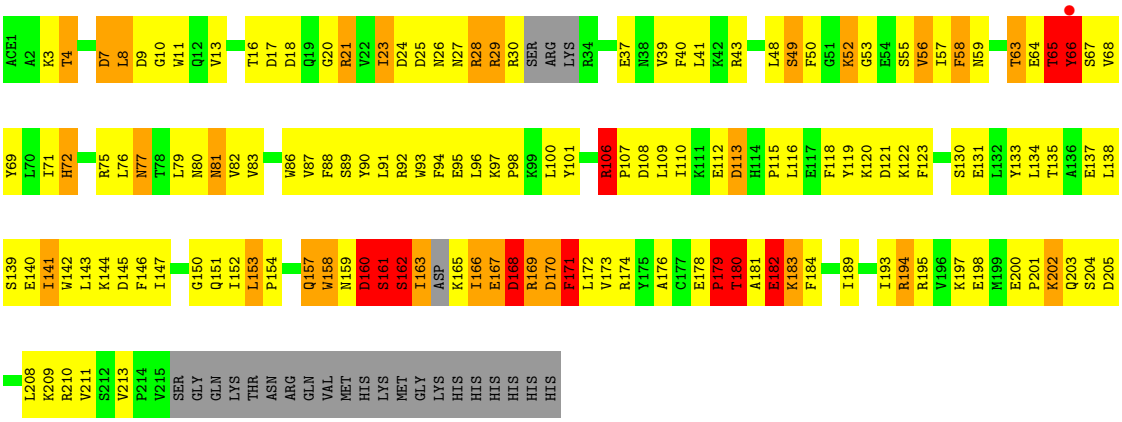




• Molecule 7: Regulatory protein SIR3

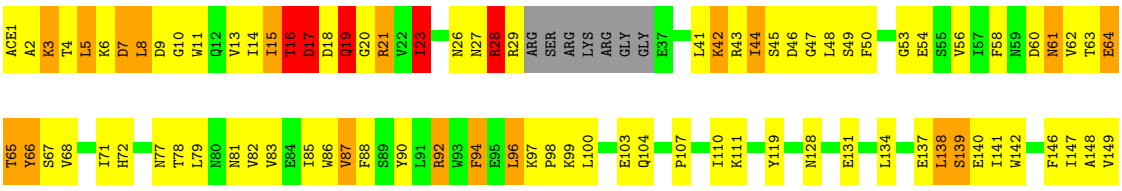
Chain K:

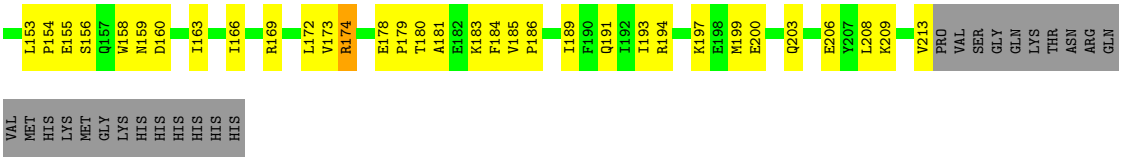




• Molecule 7: Regulatory protein SIR3

Chain L:




L153
P154
E155
S156
Q157
W158
N159
D160
I163
I166
R169
L172
V173
R174
E178
P179
T180
A181
E182
K183
F184
V185
P186
I189
F190
Q191
I192
I193
R194
K197
E198
M199
E200
Q203
E206
Y207
L208
K209
V213
PRO
VAL
SER
GLY
GLN
LYS
THR
ASN
ARG
GLN
VAL
MET
HIS
LYS
MET
GLY
LYS
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	105.51Å 105.51Å 488.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.31 20.00 – 3.31	Depositor EDS
% Data completeness (in resolution range)	97.5 (20.00-3.31) 97.1 (20.00-3.31)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 3.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.238 , 0.295 0.255 , 0.302	Depositor DCC
R_{free} test set	2217 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	118.4	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 29.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.098 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15140	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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: ACE

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	0.62	0/811	0.77	0/1088
1	E	0.71	2/737 (0.3%)	0.91	4/988 (0.4%)
2	B	0.67	0/724	0.87	1/966 (0.1%)
2	F	0.99	4/682 (0.6%)	1.07	5/913 (0.5%)
3	C	0.59	0/763	0.84	1/1033 (0.1%)
3	G	0.61	0/778	0.93	4/1053 (0.4%)
4	D	0.62	0/684	0.85	0/923
4	H	0.66	0/701	0.81	1/944 (0.1%)
5	I	0.76	3/3308 (0.1%)	1.45	64/5109 (1.3%)
6	J	0.75	1/3244 (0.0%)	1.42	55/5000 (1.1%)
7	K	0.95	8/1797 (0.4%)	1.05	17/2428 (0.7%)
7	L	0.84	6/1773 (0.3%)	0.90	10/2398 (0.4%)
All	All	0.77	24/16002 (0.1%)	1.18	162/22843 (0.7%)

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	E	0	2
2	F	0	6
3	G	0	6
7	K	0	9
7	L	0	3
All	All	0	26

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	K	153	LEU	C-N	17.69	1.67	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	27	GLN	C-N	14.93	1.59	1.33
7	L	19	GLN	C-N	14.50	1.59	1.33
7	K	182	GLU	C-N	13.21	1.64	1.34
7	L	18	ASP	C-N	12.34	1.62	1.34

The worst 5 of 162 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	63	ARG	O-C-N	-13.77	100.67	122.70
2	F	27	GLN	O-C-N	-12.67	101.67	123.20
6	J	-30	DG	O4'-C1'-N9	11.54	116.07	108.00
3	G	23	LEU	O-C-N	-11.24	104.71	122.70
6	J	38	DT	O4'-C1'-N1	10.37	115.26	108.00

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	63	ARG	Mainchain
1	E	65	LEU	Mainchain
2	F	19	ARG	Peptide
2	F	21	VAL	Mainchain
2	F	23	ARG	Mainchain

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	799	0	837	41	0
1	E	729	0	760	44	0
2	B	716	0	773	44	0
2	F	674	0	723	87	0
3	C	753	0	793	99	0
3	G	768	0	812	98	0
4	D	673	0	693	121	0
4	H	690	0	710	81	0
5	I	2946	0	1607	139	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	J	2895	0	1594	136	0
7	K	1761	0	1737	177	1
7	L	1736	0	1707	182	1
All	All	15140	0	12746	1027	1

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 38.

The worst 5 of 1027 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)
4:D:39:ALA:HB1	4:D:60:MET:CE	1.20	1.58
4:D:39:ALA:CB	4:D:60:MET:CE	1.74	1.55
7:L:1:ACE:CH3	7:L:142:TRP:CD1	1.89	1.53
7:L:10:GLY:C	7:L:44:ILE:CD1	1.77	1.50
7:L:1:ACE:H3	7:L:142:TRP:CD1	0.95	1.47

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
7:K:169:ARG:NH2	7:L:45:SER:O[6_545]	1.79	0.41

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	95/136 (70%)	89 (94%)	5 (5%)	1 (1%)	12 40
1	E	88/136 (65%)	78 (89%)	8 (9%)	2 (2%)	5 26
2	B	87/103 (84%)	72 (83%)	15 (17%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	F	81/103 (79%)	67 (83%)	10 (12%)	4 (5%)	2 12
3	C	96/130 (74%)	81 (84%)	12 (12%)	3 (3%)	3 21
3	G	98/130 (75%)	85 (87%)	12 (12%)	1 (1%)	13 42
4	D	84/126 (67%)	72 (86%)	10 (12%)	2 (2%)	5 25
4	H	86/126 (68%)	74 (86%)	11 (13%)	1 (1%)	11 38
7	K	205/236 (87%)	178 (87%)	25 (12%)	2 (1%)	13 42
7	L	202/236 (86%)	182 (90%)	18 (9%)	2 (1%)	13 42
All	All	1122/1462 (77%)	978 (87%)	126 (11%)	18 (2%)	8 32

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	77	ASP
1	E	64	LYS
2	F	24	ASP
2	F	30	THR
4	H	105	GLY

5.3.2 Protein sidechains (i)

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	85/111 (77%)	75 (88%)	10 (12%)	4 17
1	E	77/111 (69%)	69 (90%)	8 (10%)	5 22
2	B	73/79 (92%)	61 (84%)	12 (16%)	2 9
2	F	70/79 (89%)	60 (86%)	10 (14%)	2 12
3	C	76/102 (74%)	59 (78%)	17 (22%)	1 3
3	G	79/102 (78%)	64 (81%)	15 (19%)	1 5
4	D	73/106 (69%)	55 (75%)	18 (25%)	0 2
4	H	75/106 (71%)	60 (80%)	15 (20%)	1 5
7	K	195/220 (89%)	149 (76%)	46 (24%)	0 3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	L	194/220 (88%)	166 (86%)	28 (14%)	2 12
All	All	997/1236 (81%)	818 (82%)	179 (18%)	1 7

5 of 179 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	K	39	VAL
7	K	163	ILE
7	K	58	PHE
7	K	101	TYR
7	K	210	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
7	L	38	ASN
7	L	61	ASN
7	L	104	GLN
7	L	81	ASN
7	K	72	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
7	K	2
7	L	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	K	153:LEU	C	154:PRO	N	1.67
1	K	182:GLU	C	183:LYS	N	1.64
1	L	18:ASP	C	19:GLN	N	1.62

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, 95^{th} 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	97/136 (71%)	-1.03	0 100 100	29, 51, 106, 161	0
1	E	90/136 (66%)	-0.88	0 100 100	37, 69, 105, 120	0
2	B	89/103 (86%)	-0.95	0 100 100	29, 44, 121, 161	0
2	F	83/103 (80%)	-0.91	0 100 100	37, 65, 107, 129	0
3	C	98/130 (75%)	-0.66	1 (1%) 79 68	37, 71, 136, 172	0
3	G	100/130 (76%)	-0.87	0 100 100	32, 64, 104, 119	0
4	D	86/126 (68%)	-0.73	0 100 100	36, 66, 128, 162	0
4	H	88/126 (69%)	-0.87	0 100 100	40, 70, 102, 160	0
5	I	143/167 (85%)	-0.60	0 100 100	68, 152, 199, 220	0
6	J	142/167 (85%)	-0.67	0 100 100	75, 157, 194, 206	0
7	K	210/236 (88%)	-0.74	1 (0%) 87 80	32, 75, 141, 170	0
7	L	205/236 (86%)	-0.82	0 100 100	32, 71, 132, 156	0
All	All	1431/1796 (79%)	-0.79	2 (0%) 92 91	29, 74, 173, 220	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
7	K	66	TYR	2.3
3	C	41	GLU	2.2

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 (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.