W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Dec 20, 2022 – 12:26 pm GMT

PDB ID	:	8BFI
Title	:	human CNOT1-CNOT10-CNOT11 module
Authors	:	Basquin, J.; Ozgur, S.; Conti, E.
Deposited on	:	2022-10-26
Resolution	:	3.00 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	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
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)	:	2.31.3

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 3.00 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	Percentile Ranks	Value
Rfree		0.279
Clashscore		6
Ramachandran outliers		0.6%
Sidechain outliers		0.3%
RSRZ outliers		7.7%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
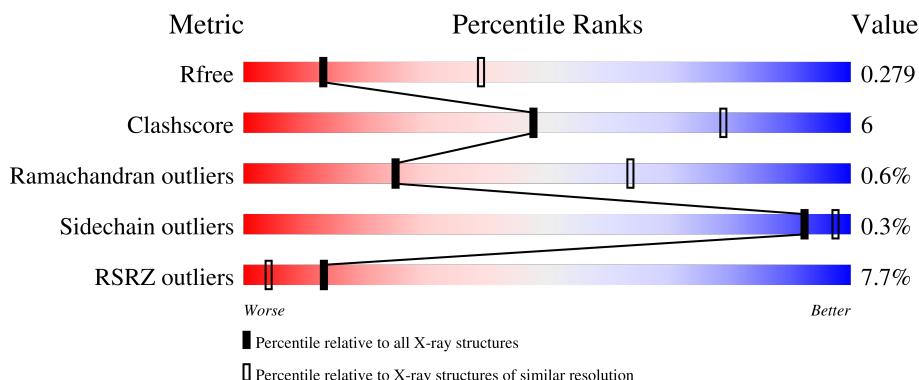
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	778	7%; 64% 13% · 23%
2	B	718	.%; 68% 12% 20%
3	C	455	13%; 80% 10% 10%
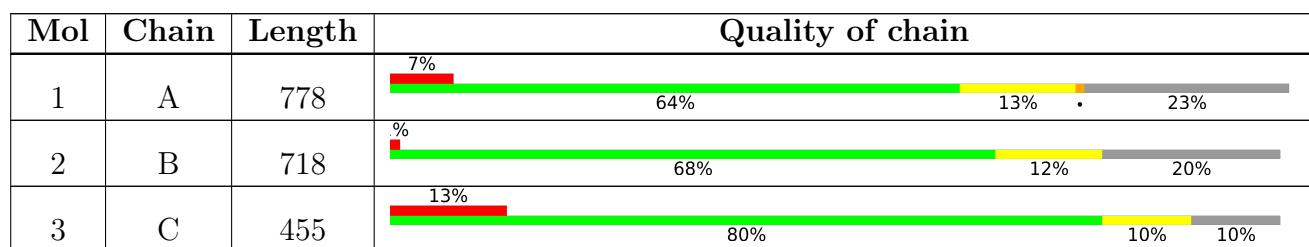
2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 11890 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 CCR4-NOT transcription complex subunit 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	602	4496	2878	780	817	10	11	0	0	0

There are 94 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-90	MSE	-	initiating methionine	UNP A5YKK6
A	-89	HIS	-	expression tag	UNP A5YKK6
A	-88	HIS	-	expression tag	UNP A5YKK6
A	-87	HIS	-	expression tag	UNP A5YKK6
A	-86	HIS	-	expression tag	UNP A5YKK6
A	-85	HIS	-	expression tag	UNP A5YKK6
A	-84	HIS	-	expression tag	UNP A5YKK6
A	-83	GLY	-	expression tag	UNP A5YKK6
A	-82	MSE	-	expression tag	UNP A5YKK6
A	-81	GLU	-	expression tag	UNP A5YKK6
A	-80	GLY	-	expression tag	UNP A5YKK6
A	-79	GLU	-	expression tag	UNP A5YKK6
A	-78	TYR	-	expression tag	UNP A5YKK6
A	-77	ILE	-	expression tag	UNP A5YKK6
A	-76	LYS	-	expression tag	UNP A5YKK6
A	-75	LEU	-	expression tag	UNP A5YKK6
A	-74	LYS	-	expression tag	UNP A5YKK6
A	-73	VAL	-	expression tag	UNP A5YKK6
A	-72	ILE	-	expression tag	UNP A5YKK6
A	-71	GLY	-	expression tag	UNP A5YKK6
A	-70	GLN	-	expression tag	UNP A5YKK6
A	-69	ASP	-	expression tag	UNP A5YKK6
A	-68	SER	-	expression tag	UNP A5YKK6
A	-67	SER	-	expression tag	UNP A5YKK6
A	-66	GLU	-	expression tag	UNP A5YKK6
A	-65	ILE	-	expression tag	UNP A5YKK6
A	-64	HIS	-	expression tag	UNP A5YKK6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-63	PHE	-	expression tag	UNP A5YKK6
A	-62	LYS	-	expression tag	UNP A5YKK6
A	-61	VAL	-	expression tag	UNP A5YKK6
A	-60	LYS	-	expression tag	UNP A5YKK6
A	-59	MSE	-	expression tag	UNP A5YKK6
A	-58	THR	-	expression tag	UNP A5YKK6
A	-57	THR	-	expression tag	UNP A5YKK6
A	-56	HIS	-	expression tag	UNP A5YKK6
A	-55	LEU	-	expression tag	UNP A5YKK6
A	-54	LYS	-	expression tag	UNP A5YKK6
A	-53	LYS	-	expression tag	UNP A5YKK6
A	-52	LEU	-	expression tag	UNP A5YKK6
A	-51	LYS	-	expression tag	UNP A5YKK6
A	-50	GLU	-	expression tag	UNP A5YKK6
A	-49	SER	-	expression tag	UNP A5YKK6
A	-48	TYR	-	expression tag	UNP A5YKK6
A	-47	CYS	-	expression tag	UNP A5YKK6
A	-46	GLN	-	expression tag	UNP A5YKK6
A	-45	ARG	-	expression tag	UNP A5YKK6
A	-44	GLN	-	expression tag	UNP A5YKK6
A	-43	GLY	-	expression tag	UNP A5YKK6
A	-42	VAL	-	expression tag	UNP A5YKK6
A	-41	PRO	-	expression tag	UNP A5YKK6
A	-40	MSE	-	expression tag	UNP A5YKK6
A	-39	ASN	-	expression tag	UNP A5YKK6
A	-38	SER	-	expression tag	UNP A5YKK6
A	-37	LEU	-	expression tag	UNP A5YKK6
A	-36	ARG	-	expression tag	UNP A5YKK6
A	-35	PHE	-	expression tag	UNP A5YKK6
A	-34	LEU	-	expression tag	UNP A5YKK6
A	-33	PHE	-	expression tag	UNP A5YKK6
A	-32	GLU	-	expression tag	UNP A5YKK6
A	-31	GLY	-	expression tag	UNP A5YKK6
A	-30	GLN	-	expression tag	UNP A5YKK6
A	-29	ARG	-	expression tag	UNP A5YKK6
A	-28	ILE	-	expression tag	UNP A5YKK6
A	-27	ALA	-	expression tag	UNP A5YKK6
A	-26	ASP	-	expression tag	UNP A5YKK6
A	-25	ASN	-	expression tag	UNP A5YKK6
A	-24	HIS	-	expression tag	UNP A5YKK6
A	-23	THR	-	expression tag	UNP A5YKK6
A	-22	PRO	-	expression tag	UNP A5YKK6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	LYS	-	expression tag	UNP A5YKK6
A	-20	GLU	-	expression tag	UNP A5YKK6
A	-19	LEU	-	expression tag	UNP A5YKK6
A	-18	GLY	-	expression tag	UNP A5YKK6
A	-17	MSE	-	expression tag	UNP A5YKK6
A	-16	GLU	-	expression tag	UNP A5YKK6
A	-15	GLU	-	expression tag	UNP A5YKK6
A	-14	GLU	-	expression tag	UNP A5YKK6
A	-13	ASP	-	expression tag	UNP A5YKK6
A	-12	VAL	-	expression tag	UNP A5YKK6
A	-11	ILE	-	expression tag	UNP A5YKK6
A	-10	GLU	-	expression tag	UNP A5YKK6
A	-9	VAL	-	expression tag	UNP A5YKK6
A	-8	TYR	-	expression tag	UNP A5YKK6
A	-7	GLN	-	expression tag	UNP A5YKK6
A	-6	GLU	-	expression tag	UNP A5YKK6
A	-5	GLN	-	expression tag	UNP A5YKK6
A	-4	THR	-	expression tag	UNP A5YKK6
A	-3	GLY	-	expression tag	UNP A5YKK6
A	-2	GLY	-	expression tag	UNP A5YKK6
A	-1	ARG	-	expression tag	UNP A5YKK6
A	0	SER	-	expression tag	UNP A5YKK6
A	381	VAL	ILE	conflict	UNP A5YKK6
A	649	VAL	ALA	conflict	UNP A5YKK6
A	681	THR	MET	conflict	UNP A5YKK6

- Molecule 2 is a protein called CCR4-NOT transcription complex subunit 10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	574	4437	2818	767	824	28	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP Q9H9A5
B	-2	PRO	-	expression tag	UNP Q9H9A5
B	-1	ASP	-	expression tag	UNP Q9H9A5
B	0	SER	-	expression tag	UNP Q9H9A5

- Molecule 3 is a protein called CCR4-NOT transcription complex subunit 11.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	410	2957	1900	499	543	15	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
C	56	GLY	-	expression tag	UNP Q9UKZ1
C	57	PRO	-	expression tag	UNP Q9UKZ1
C	58	ASP	-	expression tag	UNP Q9UKZ1
C	59	ARG	-	expression tag	UNP Q9UKZ1
C	60	SER	-	expression tag	UNP Q9UKZ1
C	507	GLU	LYS	conflict	UNP Q9UKZ1

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: CCR4-NOT transcription complex subunit 1
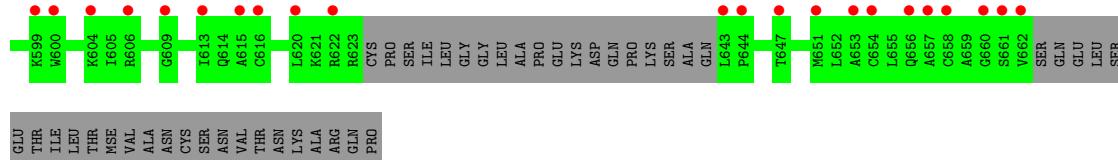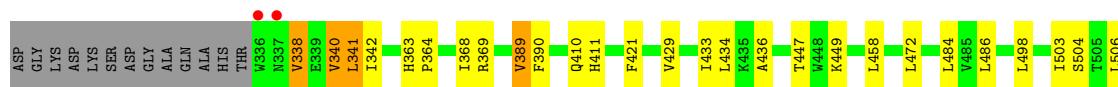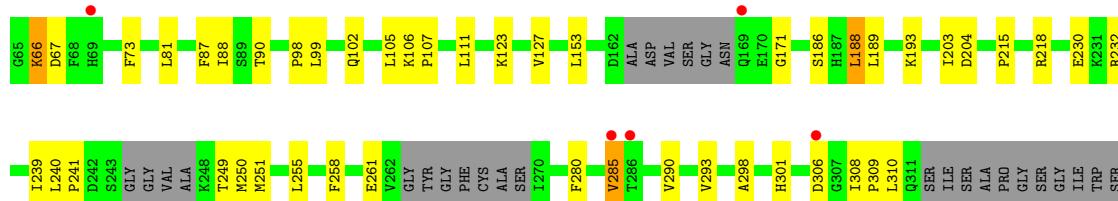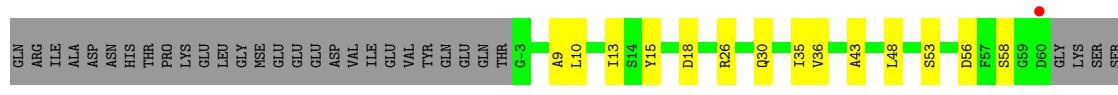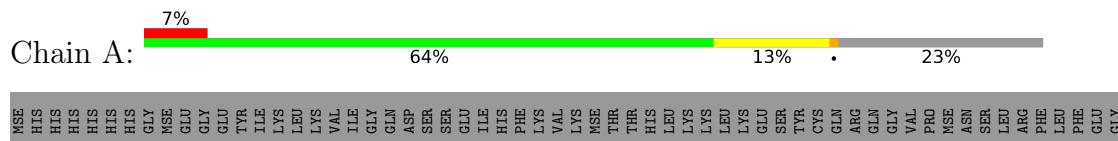
• Molecule 2: CCR4-NOT transcription complex subunit 10

• Molecule 3: CCR4-NOT transcription complex subunit 11

Chain C:
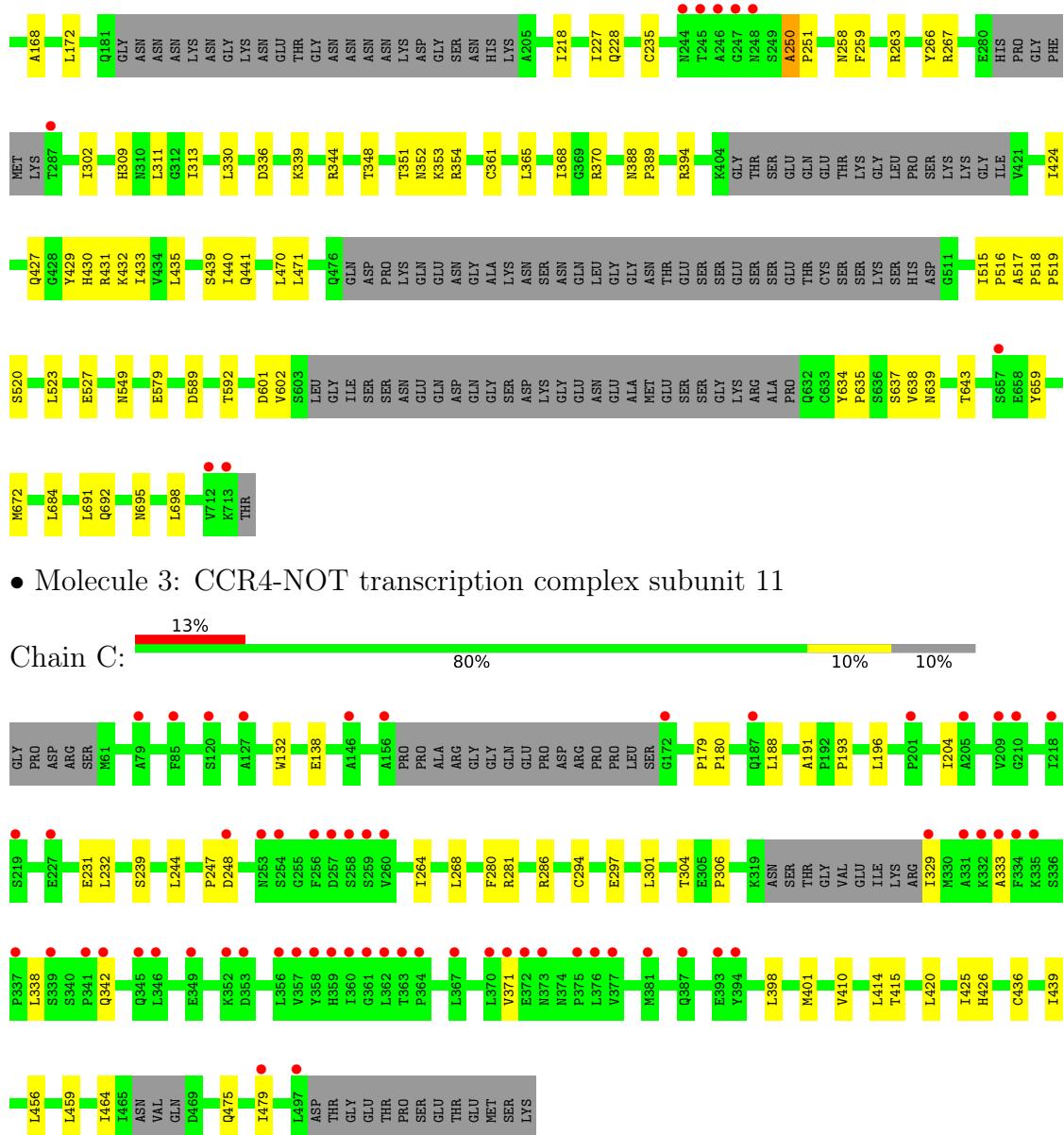
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	83.76Å 110.64Å 128.58Å 90.00° 103.39° 90.00°	Depositor
Resolution (Å)	45.77 – 3.00 45.77 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (45.77-3.00) 100.0 (45.77-3.00)	Depositor EDS
R_{merge}	0.32	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.262 , 0.283 0.260 , 0.279	Depositor DCC
R_{free} test set	2293 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	79.5	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 74.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	11890	wwPDB-VP
Average B, all atoms (Å^2)	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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 ⓘ

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	0.24	0/4577	0.48	3/6205 (0.0%)
2	B	0.24	0/4513	0.43	0/6107
3	C	0.24	0/3024	0.45	1/4136 (0.0%)
All	All	0.24	0/12114	0.45	4/16448 (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	A	0	3
3	C	0	1
All	All	0	4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	66	LYS	C-N-CA	6.66	138.35	121.70
1	A	513	ASN	C-N-CA	6.53	138.02	121.70
3	C	414	LEU	CA-CB-CG	5.38	127.67	115.30
1	A	188	LEU	CA-CB-CG	5.03	126.86	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	189	LEU	Peptide
1	A	285	VAL	Peptide
1	A	368	ILE	Peptide
3	C	248	ASP	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	4496	0	4271	68	0
2	B	4437	0	4406	62	0
3	C	2957	0	2801	35	0
All	All	11890	0	11478	142	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 6.

The worst 5 of 142 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:389:VAL:HG12	1:A:390:PHE:H	1.34	0.93
3:C:459:LEU:HD23	3:C:464:ILE:HD12	1.68	0.75
1:A:251:MSE:HE1	2:B:361:CYS:HB3	1.68	0.74
3:C:415:THR:HG22	3:C:464:ILE:HG21	1.69	0.73
1:A:251:MSE:HE2	2:B:313:ILE:HD11	1.72	0.70

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	588/778 (76%)	576 (98%)	7 (1%)	5 (1%)	17	55
2	B	560/718 (78%)	550 (98%)	6 (1%)	4 (1%)	22	60
3	C	402/455 (88%)	395 (98%)	7 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1550/1951 (79%)	1521 (98%)	20 (1%)	9 (1%)	25	64

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	340	VAL
1	A	341	LEU
2	B	141	GLU
1	A	389	VAL
1	A	193	LYS

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	446/660 (68%)	443 (99%)	3 (1%)	84	94
2	B	472/607 (78%)	471 (100%)	1 (0%)	93	98
3	C	295/406 (73%)	295 (100%)	0	100	100
All	All	1213/1673 (72%)	1209 (100%)	4 (0%)	92	97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	ASP
1	A	73	PHE
1	A	204	ASP
2	B	258	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	228	GLN

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	591/778 (75%)	0.30	51 (8%)	10	3	39, 83, 220, 262	0
2	B	574/718 (79%)	-0.16	10 (1%)	70	41	30, 59, 110, 139	0
3	C	410/455 (90%)	0.84	61 (14%)	2	1	60, 142, 220, 246	0
All	All	1575/1951 (80%)	0.27	122 (7%)	13	4	30, 84, 214, 262	0

The worst 5 of 122 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	644	PRO	11.4
1	A	661	SER	10.6
3	C	361	GLY	9.6
3	C	331	ALA	9.3
3	C	345	GLN	9.0

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 ⓘ

There are no ligands in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

6.5 Other polymers ⓘ

There are no such residues in this entry.