



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 02:54 PM EDT

PDB ID : 6CIM
Title : Pre-Reaction Complex, RAG1(E962Q)/2-nicked/intact 12/23RSS complex in Mn²⁺
Authors : Chuenchor, W.; Chen, X.; Kim, M.S.; Gellert, M.; Yang, W.
Deposited on : 2018-02-24
Resolution : 3.60 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 ⓘ 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) : 1.13
EDS : 2.35.1
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) : 2.35.1

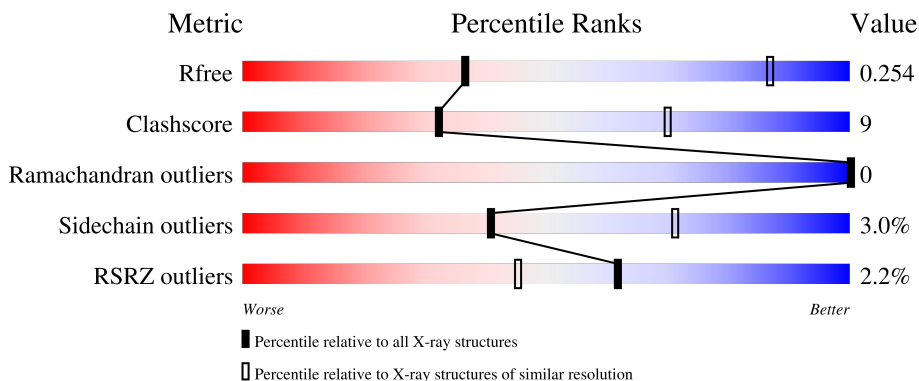
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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	625	 2% 78% 19% •
1	C	625	 78% 18% •
2	B	359	 2% 74% 20% • 5%
2	D	359	 1% 78% 15% • 6%
3	N	163	 5% 45% 9% • 45%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	40	<p>2% 35% 50% 12%</p>
5	I	10	<p>10% 50% 50%</p>
6	L	30	<p>3% 30% 50% 17%</p>
7	G	56	<p>11% 32% 52% 9% 7%</p>
8	J	56	<p>4% 20% 68% 5% 7%</p>

2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 18804 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 V(D)J recombination-activating protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	611	Total	C	N	O	S	0	0	0
			4756	2994	839	889	34			
1	C	602	Total	C	N	O	S	0	0	0
			4661	2941	821	867	32			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	962	GLN	GLU	engineered mutation	UNP P15919
C	962	GLN	GLU	engineered mutation	UNP P15919

- Molecule 2 is a protein called V(D)J recombination-activating protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	340	Total	C	N	O	S	0	1	0
			2643	1689	449	487	18			
2	D	339	Total	C	N	O	S	0	1	0
			2619	1674	450	476	19			

- Molecule 3 is a protein called High mobility group protein B1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	N	90	Total	C	N	O	S	0	0	0
			553	348	100	103	2			

- Molecule 4 is a DNA chain called Nicked 12RSS intermediate reverse strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	F	35	Total	C	N	O	P	0	0	0
			723	345	129	214	35			

- Molecule 5 is a DNA chain called DNA (5'-D(*GP*CP*CP*TP*GP*TP*CP*TP*TP*A)-3'

).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
5	I	10	199	97	32	61	9	0	0	0

- Molecule 6 is a DNA chain called Nicked 12RSS intermediate forward strand.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
6	L	25	511	243	99	144	25	0	0	0

- Molecule 7 is a DNA chain called Intact 23RSS substrate reverse strand.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
7	G	52	1065	509	184	320	52	0	0	0

- Molecule 8 is a DNA chain called Intact 23RSS substrate forward strand.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
8	J	52	1067	507	204	304	52	0	0	0

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Zn	0	0
			1	1		
9	C	1	Total	Zn	0	0
			1	1		

- Molecule 10 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	1	Total	Mn	0	0
			1	1		
10	C	1	Total	Mn	0	0
			1	1		

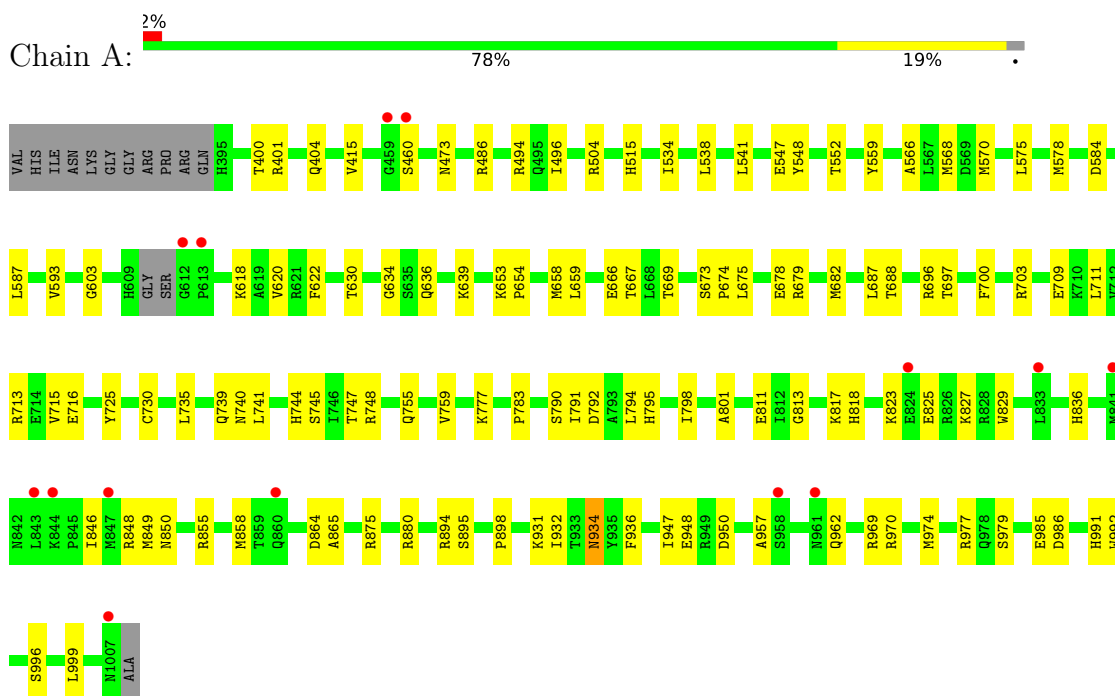
- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	1	Total O 1 1	0	0
11	B	1	Total O 1 1	0	0
11	C	1	Total O 1 1	0	0

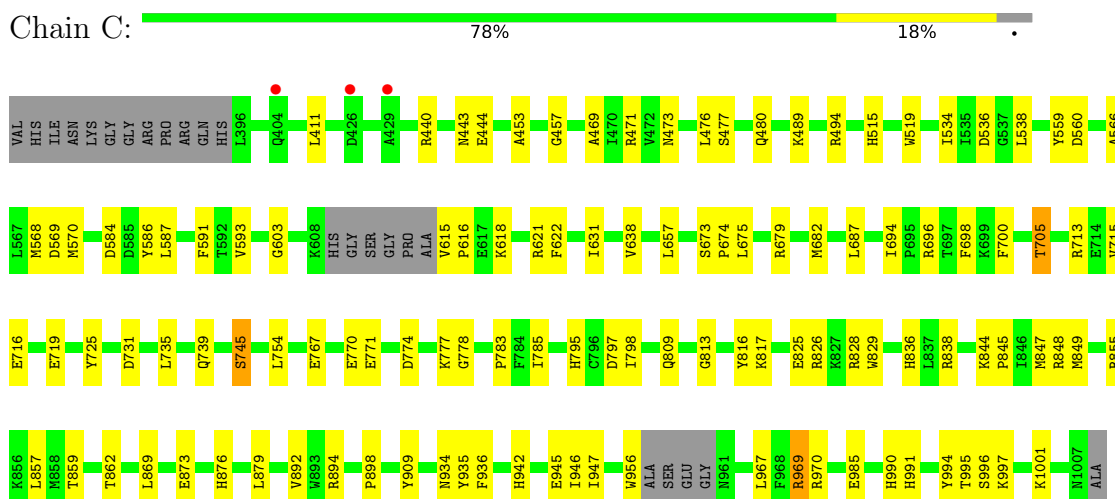
3 Residue-property plots i

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.

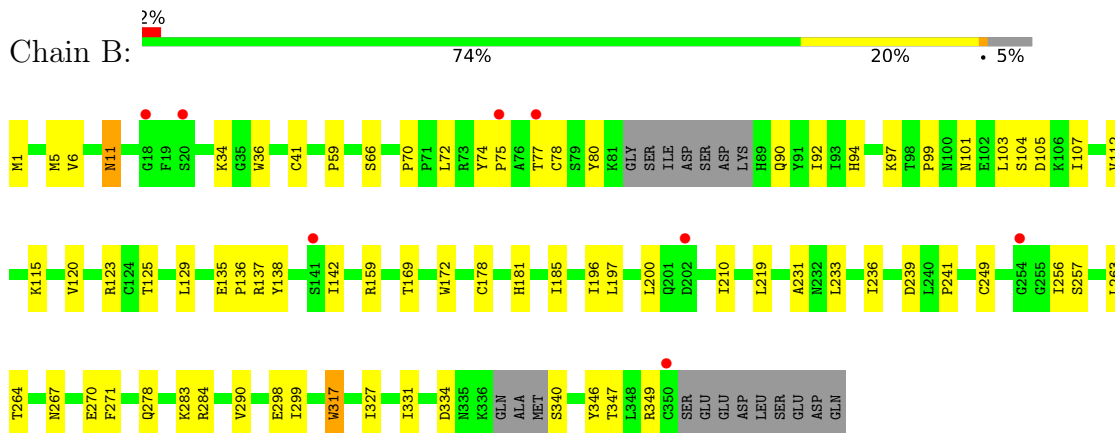
- Molecule 1: V(D)J recombination-activating protein 1



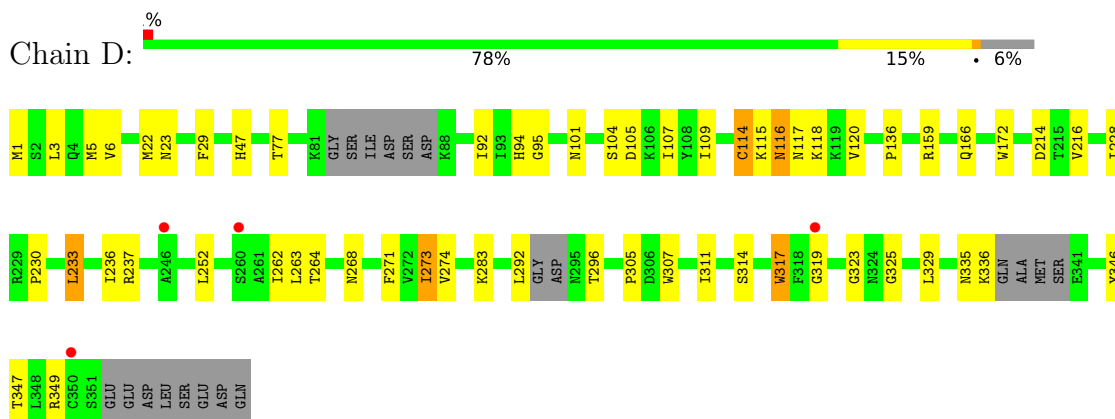
- Molecule 1: V(D)J recombination-activating protein 1



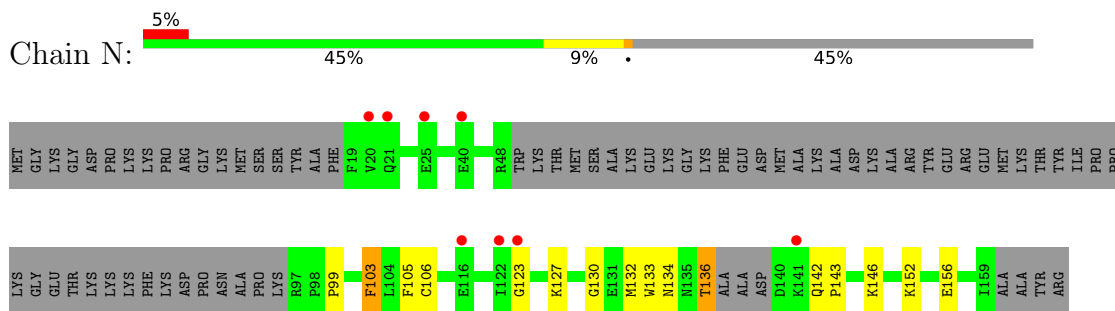
- Molecule 2: V(D)J recombination-activating protein 2



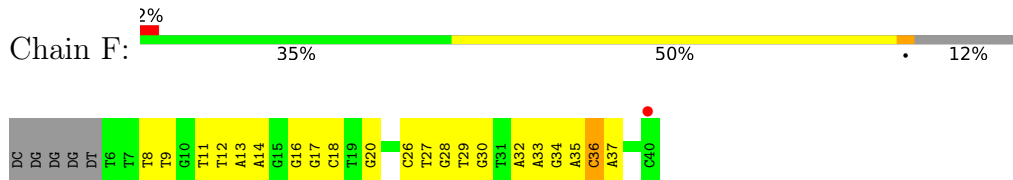
- Molecule 2: V(D)J recombination-activating protein 2



- Molecule 3: High mobility group protein B1



- Molecule 4: Nicked 12RSS intermediate reverse strand



- Molecule 5: DNA (5'-D(*GP*CP*CP*TP*GP*TP*CP*TP*TP*A)-3')

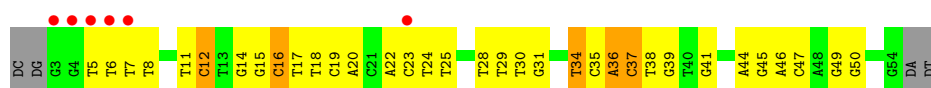




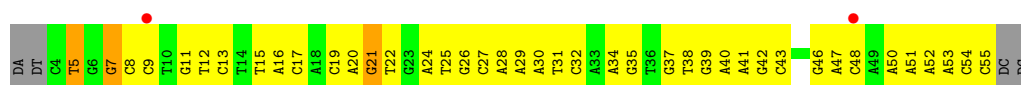
- Molecule 6: Nicked 12RSS intermediate forward strand



- Molecule 7: Intact 23RSS substrate reverse strand



- Molecule 8: Intact 23RSS substrate forward strand



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	156.68Å 123.56Å 186.29Å 90.00° 105.60° 90.00°	Depositor
Resolution (Å)	38.69 – 3.60 38.69 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.69-3.60) 99.7 (38.69-3.60)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.08 (at 3.57Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.214 , 0.254 0.214 , 0.254	Depositor DCC
R_{free} test set	2049 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å ²)	119.3	Xtrriage
Anisotropy	0.342	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 60.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	18804	wwPDB-VP
Average B, all atoms (Å ²)	125.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% 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 i

5.1 Standard geometry i

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

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	0.44	0/4852	0.61	0/6555
1	C	0.43	0/4753	0.62	0/6423
2	B	0.42	0/2714	0.60	0/3684
2	D	0.38	0/2685	0.58	0/3639
3	N	0.34	0/562	0.49	0/769
4	F	0.87	1/810 (0.1%)	1.12	1/1250 (0.1%)
5	I	0.91	0/221	1.14	0/339
6	L	1.06	2/574 (0.3%)	1.08	1/880 (0.1%)
7	G	1.06	3/1191 (0.3%)	1.23	4/1837 (0.2%)
8	J	1.23	6/1199 (0.5%)	1.29	12/1847 (0.6%)
All	All	0.61	12/19561 (0.1%)	0.78	18/27223 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	17	DC	OP3-P	-10.01	1.49	1.61
8	J	21	DG	N9-C4	-7.88	1.31	1.38
8	J	21	DG	N9-C8	7.75	1.43	1.37
8	J	17	DC	C1'-N1	-6.77	1.37	1.47
8	J	5	DT	C1'-N1	6.25	1.57	1.49

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	J	48	DC	C6-N1-C2	9.58	124.13	120.30
8	J	48	DC	N1-C2-O2	-7.81	114.22	118.90
8	J	17	DC	C6-N1-C2	7.63	123.35	120.30
8	J	48	DC	N3-C2-O2	6.90	126.73	121.90
8	J	21	DG	N3-C4-N9	-6.59	122.04	126.00

There are no chirality outliers.

There are no planarity outliers.

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	4756	0	4604	69	0
1	C	4661	0	4482	68	0
2	B	2643	0	2565	47	0
2	D	2619	0	2514	34	0
3	N	553	0	388	8	0
4	F	723	0	398	19	0
5	I	199	0	116	4	0
6	L	511	0	280	13	0
7	G	1065	0	591	30	0
8	J	1067	0	583	33	0
9	A	1	0	0	0	0
9	C	1	0	0	0	0
10	A	1	0	0	0	0
10	C	1	0	0	0	0
11	A	1	0	0	0	0
11	B	1	0	0	0	0
11	C	1	0	0	0	0
All	All	18804	0	16521	306	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.

The worst 5 of 306 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)
8:J:50:DA:H2'	8:J:51:DA:C8	2.10	0.86
1:A:688:THR:HG22	1:A:697:THR:HG22	1.57	0.84
4:F:12:DT:H2'	4:F:13:DA:C8	2.16	0.81
2:B:290:VAL:HG12	2:B:299:ILE:HG12	1.65	0.78
1:C:515:HIS:ND1	1:C:568:MET:HG3	1.99	0.78

There are no symmetry-related clashes.

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	607/625 (97%)	590 (97%)	17 (3%)	0	100	100
1	C	596/625 (95%)	573 (96%)	23 (4%)	0	100	100
2	B	335/359 (93%)	323 (96%)	12 (4%)	0	100	100
2	D	332/359 (92%)	321 (97%)	11 (3%)	0	100	100
3	N	84/163 (52%)	80 (95%)	4 (5%)	0	100	100
All	All	1954/2131 (92%)	1887 (97%)	67 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	500/549 (91%)	482 (96%)	18 (4%)	35	67
1	C	481/549 (88%)	469 (98%)	12 (2%)	47	75
2	B	291/316 (92%)	287 (99%)	4 (1%)	67	85
2	D	280/316 (89%)	272 (97%)	8 (3%)	42	72
3	N	30/139 (22%)	25 (83%)	5 (17%)	2	14
All	All	1582/1869 (85%)	1535 (97%)	47 (3%)	41	71

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	859	THR
2	D	114	CYS
1	C	934	ASN
1	C	970	ARG
2	D	116	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	278	GLN
2	D	23	ASN
2	D	268	ASN
1	A	818	HIS
1	A	836	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	611/625 (97%)	-0.01	14 (2%) 60 44	63, 107, 178, 208	0
1	C	602/625 (96%)	-0.19	3 (0%) 91 83	66, 102, 172, 194	0
2	B	340/359 (94%)	0.06	8 (2%) 59 42	71, 107, 146, 171	0
2	D	339/359 (94%)	-0.01	4 (1%) 79 66	75, 120, 158, 180	0
3	N	90/163 (55%)	0.44	8 (8%) 9 5	121, 161, 197, 202	0
4	F	35/40 (87%)	-0.10	1 (2%) 51 35	118, 192, 233, 249	0
5	I	10/10 (100%)	0.44	1 (10%) 7 4	182, 187, 195, 196	0
6	L	25/30 (83%)	0.06	1 (4%) 38 25	102, 168, 248, 249	0
7	G	52/56 (92%)	0.17	6 (11%) 4 3	91, 163, 247, 263	0
8	J	52/56 (92%)	0.03	2 (3%) 40 26	79, 160, 248, 267	0
All	All	2156/2323 (92%)	-0.02	48 (2%) 62 45	63, 114, 187, 267	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	847	MET	5.6
7	G	5	DT	5.0
1	A	961	ASN	4.8
1	A	460	SER	4.6
1	A	824	GLU	4.3

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](#)

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(\AA^2)	Q<0.9
10	MN	C	1102	1/1	0.89	0.33	86,86,86,86	0
10	MN	A	1102	1/1	0.92	0.36	84,84,84,84	0
9	ZN	C	1101	1/1	0.99	0.15	89,89,89,89	0
9	ZN	A	1101	1/1	1.00	0.14	90,90,90,90	0

6.5 Other polymers [i](#)

There are no such residues in this entry.