



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 11:23 AM EDT

PDB ID : 7UGW
Title : M. tuberculosis DNA gyrase cleavage core bound to DNA and evybactin
Authors : Hauk, G.; Imai, Y.; Lewis, K.; Berger, J.M.
Deposited on : 2022-03-25
Resolution : 3.00 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

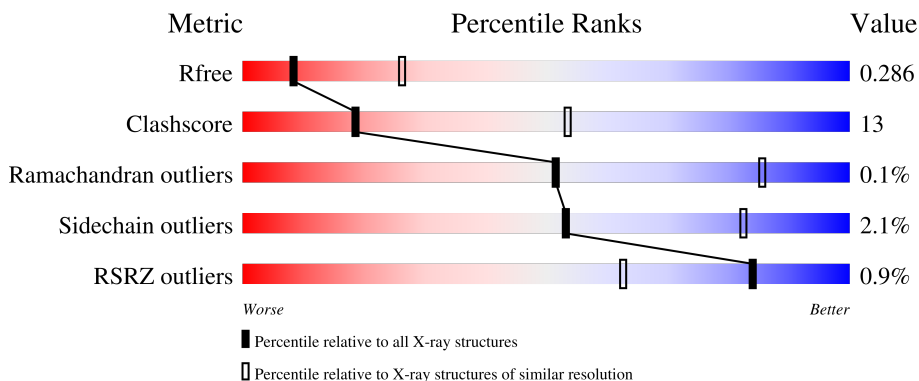
1 Overall quality at a glance i

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.



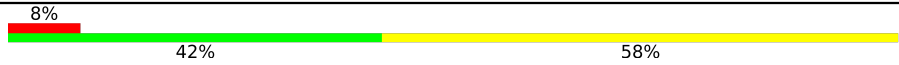
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 $\leq 5\%$. 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	500	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 72%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 24%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">72% 24% ..</p>
1	C	500	<div style="display: flex; align-items: center;"> <div style="width: 78%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 19%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">78% 19% ..</p>
2	B	251	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 76%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 23%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">76% 23% .</p>
2	D	251	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 69%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 20%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">2% 69% 20% . 8%</p>
3	V	46	<div style="display: flex; align-items: center;"> <div style="width: 17%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 65%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: orange; margin-right: 5px;"></div> </div> <p style="text-align: center;">17% 65% 17%</p>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	12	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a red segment on the left labeled '8%', a green segment in the middle labeled '42%', and a yellow segment on the right labeled '58%'. The segments are stacked horizontally, with the red segment starting from the left edge of the bar.</p>

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 24269 atoms, of which 11895 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 DNA gyrase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	485	7470	2342	3720	676	719	13	0	0	0
1	C	487	7640	2377	3822	700	728	13	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	129	PHE	TYR	engineered mutation	UNP P9WG47
C	129	PHE	TYR	engineered mutation	UNP P9WG47

- Molecule 2 is a protein called DNA gyrase subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	B	248	3906	1222	1964	348	365	7	0	0	0
2	D	232	3599	1138	1788	320	346	7	0	0	0

- Molecule 3 is a DNA chain called DNA (46-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				P
3	V	46	1459	446	514	184	270	45	0	0	0

- Molecule 4 is a protein called evybactin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	H	N				O
4	E	12	193	64	87	21	21	0	0	0

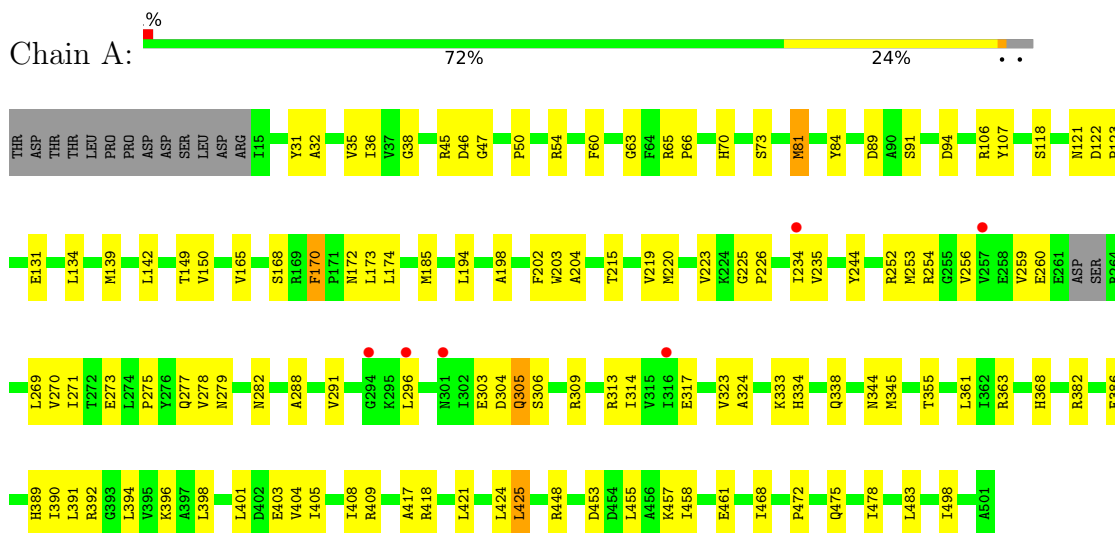
- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total 1	Mg 1	0	0
5	V	1	Total 1	Mg 1	0	0

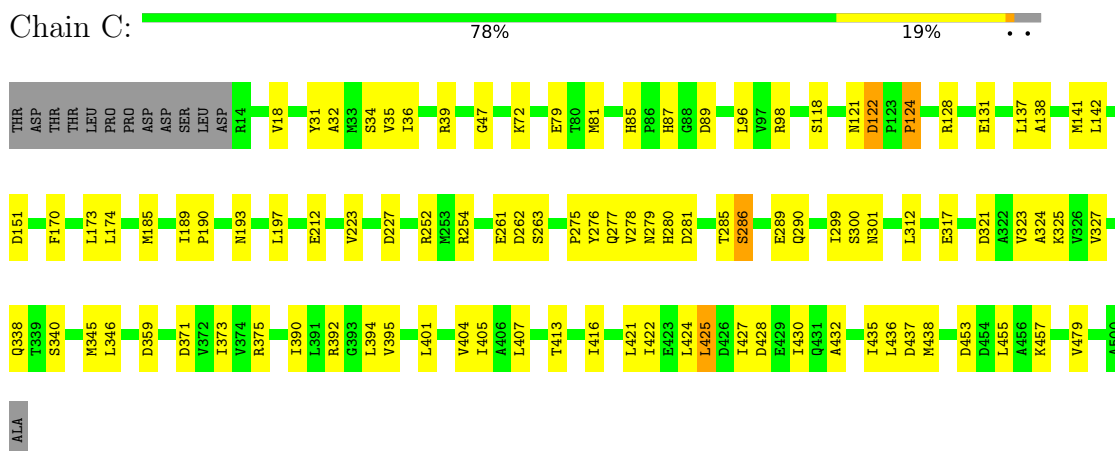
3 Residue-property plots [i](#)

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.

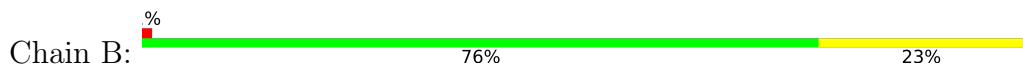
- Molecule 1: DNA gyrase subunit A

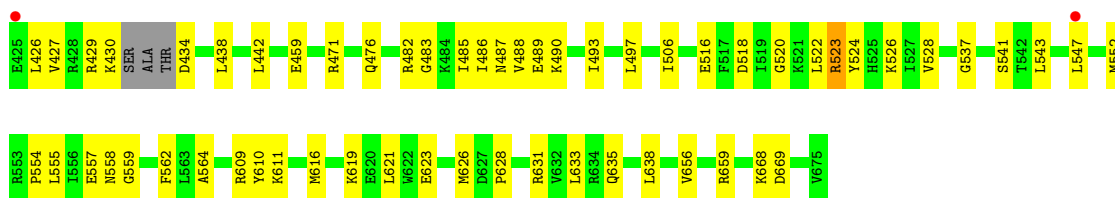


- Molecule 1: DNA gyrase subunit A

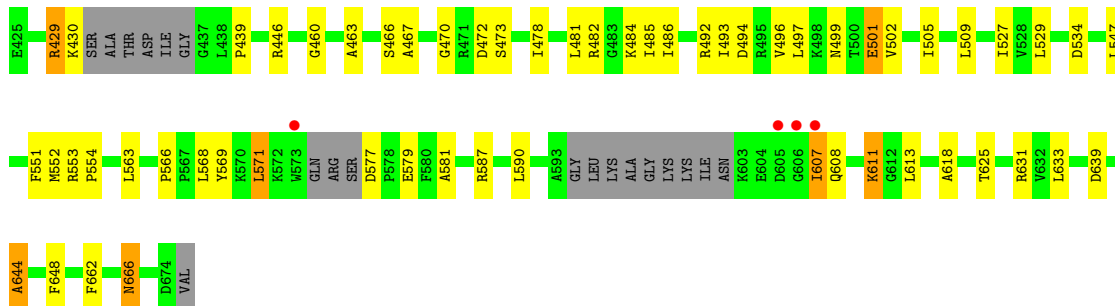


- Molecule 2: DNA gyrase subunit B





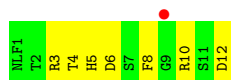
● Molecule 2: DNA gyrase subunit B



● Molecule 3: DNA (46-MER)



● Molecule 4: eybactin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	83.08Å 105.09Å 250.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.46 – 3.00 44.46 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.46-3.00) 99.7 (44.46-3.00)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.53 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.201 , 0.286 0.201 , 0.286	Depositor DCC
R_{free} test set	1809 reflections (4.06%)	wwPDB-VP
Wilson B-factor (Å ²)	59.1	Xtrriage
Anisotropy	0.682	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 47.0	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	0.93	EDS
Total number of atoms	24269	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% 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: DAR, NLF, MHS, MG, DSN, IAS

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	0.58	0/3808	0.85	3/5168 (0.1%)
1	C	0.64	1/3877 (0.0%)	0.89	6/5253 (0.1%)
2	B	0.62	0/1970	0.88	1/2648 (0.0%)
2	D	0.59	0/1837	0.92	5/2475 (0.2%)
3	V	1.57	7/1062 (0.7%)	1.40	16/1638 (1.0%)
4	E	4.19	7/32 (21.9%)	1.14	0/37
All	All	0.77	15/12586 (0.1%)	0.94	31/17219 (0.2%)

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	8	PHE	C-N	9.60	1.50	1.33
4	E	8	PHE	CG-CD1	9.50	1.53	1.38
4	E	8	PHE	CG-CD2	8.95	1.52	1.38
4	E	8	PHE	CE2-CZ	8.07	1.52	1.37
1	C	124	PRO	N-CD	8.04	1.59	1.47

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	124	PRO	CA-N-CD	-11.58	95.28	111.50
1	C	437	ASP	CB-CG-OD2	10.35	127.62	118.30
3	V	23	DG	O4'-C4'-C3'	-9.11	100.54	106.00
3	V	24	DG	O4'-C4'-C3'	-8.86	100.68	106.00
3	V	2	DG	O4'-C1'-N9	8.12	113.69	108.00

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	3750	3720	3719	95	0
1	C	3818	3822	3822	72	0
2	B	1942	1964	1963	57	0
2	D	1811	1788	1786	40	0
3	V	945	514	514	70	0
4	E	106	87	76	0	0
5	B	1	0	0	0	0
5	V	1	0	0	0	0
All	All	12374	11895	11880	319	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 13.

The worst 5 of 319 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:C:32:ALA:O	1:C:36:ILE:HG13	1.46	1.13
3:V:25:DG:H5''	3:V:25:DG:H8	1.27	0.95
3:V:35:DT:H2'	3:V:36:DG:O4'	1.70	0.92
2:D:607:ILE:HG22	2:D:607:ILE:O	1.71	0.91
1:A:275:PRO:HG2	1:A:278:VAL:HG21	1.55	0.89

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	481/500 (96%)	450 (94%)	31 (6%)	0	100	100
1	C	485/500 (97%)	448 (92%)	36 (7%)	1 (0%)	47	82
2	B	244/251 (97%)	229 (94%)	15 (6%)	0	100	100
2	D	224/251 (89%)	206 (92%)	18 (8%)	0	100	100
4	E	5/12 (42%)	5 (100%)	0	0	100	100
All	All	1439/1514 (95%)	1338 (93%)	100 (7%)	1 (0%)	51	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	39	ARG

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	393/419 (94%)	385 (98%)	8 (2%)	55	83
1	C	405/419 (97%)	397 (98%)	8 (2%)	55	83
2	B	204/210 (97%)	203 (100%)	1 (0%)	88	96
2	D	189/210 (90%)	182 (96%)	7 (4%)	34	70
4	E	4/4 (100%)	3 (75%)	1 (25%)	0	3
All	All	1195/1262 (95%)	1170 (98%)	25 (2%)	53	82

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

Mol	Chain	Res	Type
1	C	340	SER
2	D	429	ARG
4	E	4	THR
1	C	438	MET
2	D	466	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

6 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	DSN	E	7	4	4,5,6	0.43	0	0,5,7	-	-
4	IAS	E	12	4	6,7,8	1.67	1 (16%)	6,8,10	1.33	0
4	IAS	E	6	4	6,7,8	1.92	1 (16%)	6,8,10	1.36	1 (16%)
4	MHS	E	5	4	7,11,12	1.63	1 (14%)	6,14,16	2.07	3 (50%)
4	DAR	E	3	4	9,10,11	2.95	4 (44%)	5,11,13	1.08	0
4	DAR	E	10	4	9,10,11	2.99	3 (33%)	5,11,13	0.90	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	DSN	E	7	4	-	2/2/4/6	-
4	IAS	E	12	4	-	2/7/7/8	-
4	IAS	E	6	4	-	0/7/7/8	-
4	MHS	E	5	4	-	1/5/6/8	0/1/1/1
4	DAR	E	3	4	-	3/8/9/11	-
4	DAR	E	10	4	-	1/8/9/11	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	10	DAR	CZ-NE	7.37	1.47	1.33
4	E	3	DAR	CZ-NE	7.09	1.47	1.33
4	E	6	IAS	CB-CG	4.14	1.60	1.49
4	E	5	MHS	CB-CG	4.07	1.56	1.50
4	E	10	DAR	CZ-NH2	3.83	1.47	1.32

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	5	MHS	CM-ND1-CG	3.52	129.13	124.44
4	E	5	MHS	NE2-CE1-ND1	-2.65	108.32	112.26
4	E	6	IAS	OXT-C-CA	2.38	121.49	113.38
4	E	5	MHS	CB-CA-C	2.03	115.28	111.47

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	3	DAR	N-CA-CB-CG
4	E	3	DAR	C-CA-CB-CG
4	E	5	MHS	O-C-CA-CB
4	E	7	DSN	N-CA-CB-OG
4	E	7	DSN	C-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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

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 [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	485/500 (97%)	-0.14	6 (1%) 79 54	27, 53, 104, 134	0
1	C	487/500 (97%)	-0.16	0 100 100	27, 47, 92, 136	0
2	B	248/251 (98%)	-0.09	2 (0%) 86 65	30, 57, 90, 196	0
2	D	232/251 (92%)	0.06	4 (1%) 70 41	41, 63, 100, 172	0
3	V	46/46 (100%)	0.02	0 100 100	33, 58, 107, 113	0
4	E	5/12 (41%)	0.76	1 (20%) 1 0	68, 71, 79, 82	0
All	All	1503/1560 (96%)	-0.10	13 (0%) 84 63	27, 55, 99, 196	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	606	GLY	3.4
1	A	296	LEU	2.9
2	D	573	TRP	2.9
2	D	607	ILE	2.7
2	D	605	ASP	2.5

6.2 Non-standard residues in protein, DNA, RNA chains [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(Å ²)	Q<0.9
4	IAS	E	6	8/9	0.76	0.29	61,83,102,102	0
4	DSN	E	7	6/7	0.79	0.30	62,73,88,105	0
4	DAR	E	10	11/12	0.81	0.33	55,78,109,131	0
4	DAR	E	3	11/12	0.83	0.30	68,84,101,104	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	MHS	E	5	11/12	0.84	0.32	79,95,114,120	0
4	IAS	E	12	8/9	0.85	0.26	53,77,85,92	0

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
5	MG	V	101	1/1	0.96	0.48	41,41,41,41	0
5	MG	B	701	1/1	0.97	0.58	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.