



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 12, 2026 – 04:29 PM UTC

PDB ID : 4D1Q / pdb_00004d1q
Title : Hermes transposase bound to its terminal inverted repeat
Authors : Hickman, A.B.; Ewis, H.; Li, X.; Knapp, J.; Laver, T.; Doss, A.L.; Tolun, G.; Steven, A.; Grishaev, A.; Bax, A.; Atkinson, P.; Craig, N.L.; Dyda, F.
Deposited on : 2014-05-04
Resolution : 3.40 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

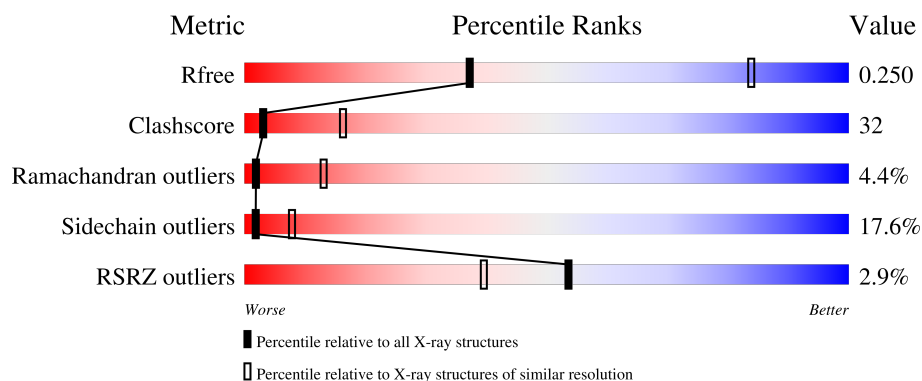
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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.



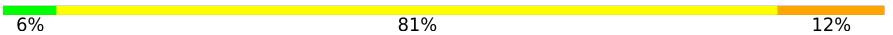
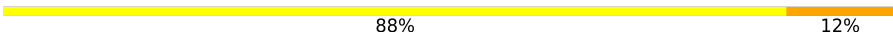
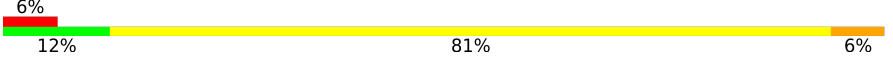
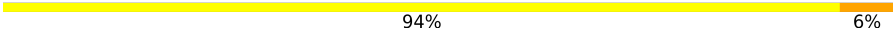
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1001 (3.44-3.36)
Clashscore	190562	1022 (3.44-3.36)
Ramachandran outliers	187476	1012 (3.44-3.36)
Sidechain outliers	187428	1012 (3.44-3.36)
RSRZ outliers	180081	1001 (3.44-3.36)

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	536	 2% 33% 41% 18% • 6%
1	B	536	 2% 39% 41% 12% • 6%
1	G	536	 4% 36% 45% 10% • 7%
1	H	536	 4% 41% 39% 13% • 5%
2	C	15	 33% 40% 27%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	15	 33% 60% 7%
2	I	15	 27% 67% 7%
2	K	15	 27% 73%
3	D	16	 6% 81% 12%
3	F	16	 88% 12%
3	J	16	 6% 12% 81% 6%
3	L	16	 94% 6%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 18737 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 TRANSPOSASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	502	Total	C	N	O	S	0	0	0
			4041	2584	684	754	19			
1	B	505	Total	C	N	O	S	0	0	0
			4066	2599	689	758	20			
1	G	499	Total	C	N	O	S	0	0	0
			4013	2569	677	748	19			
1	H	509	Total	C	N	O	S	0	0	0
			4089	2613	691	765	20			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	HIS	-	expression tag	UNP Q25442
A	78	MET	-	expression tag	UNP Q25442
A	163	GLY	SER	conflict	UNP Q25442
A	233	SER	LEU	engineered mutation	UNP Q25442
A	286	MET	VAL	engineered mutation	UNP Q25442
B	77	HIS	-	expression tag	UNP Q25442
B	78	MET	-	expression tag	UNP Q25442
B	163	GLY	SER	conflict	UNP Q25442
B	233	SER	LEU	engineered mutation	UNP Q25442
B	286	MET	VAL	engineered mutation	UNP Q25442
B	519	CYS	CYS	conflict	UNP Q25442
G	77	HIS	-	expression tag	UNP Q25442
G	78	MET	-	expression tag	UNP Q25442
G	163	GLY	SER	conflict	UNP Q25442
G	233	SER	LEU	engineered mutation	UNP Q25442
G	286	MET	VAL	engineered mutation	UNP Q25442
H	77	HIS	-	expression tag	UNP Q25442
H	78	MET	-	expression tag	UNP Q25442
H	163	GLY	SER	conflict	UNP Q25442
H	233	SER	LEU	engineered mutation	UNP Q25442
H	286	MET	VAL	engineered mutation	UNP Q25442

- Molecule 2 is a DNA chain called TERMINAL INVERTED REPEAT.

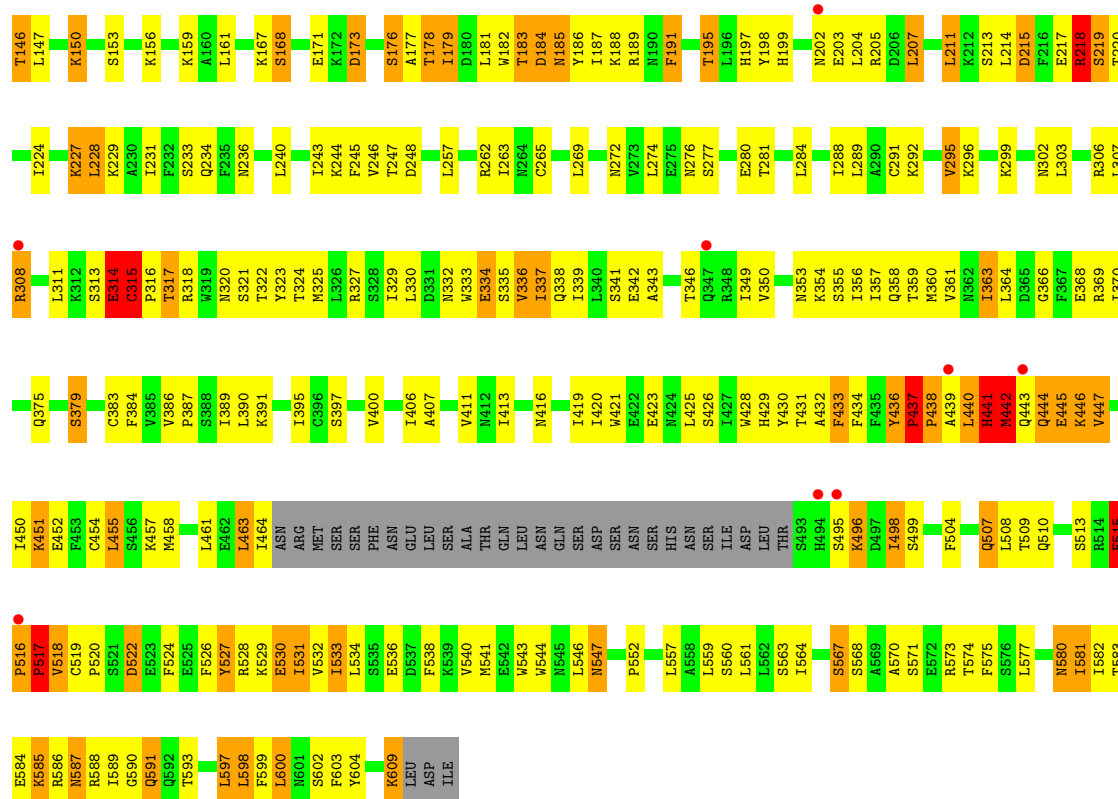
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	15	Total	C	N	O	P	0	0	0
			308	147	69	78	14			
2	E	15	Total	C	N	O	P	0	0	0
			308	147	69	78	14			
2	I	15	Total	C	N	O	P	0	0	0
			308	147	69	78	14			
2	K	15	Total	C	N	O	P	0	0	0
			308	147	69	78	14			

- Molecule 3 is a DNA chain called TERMINAL INVERTED REPEAT.

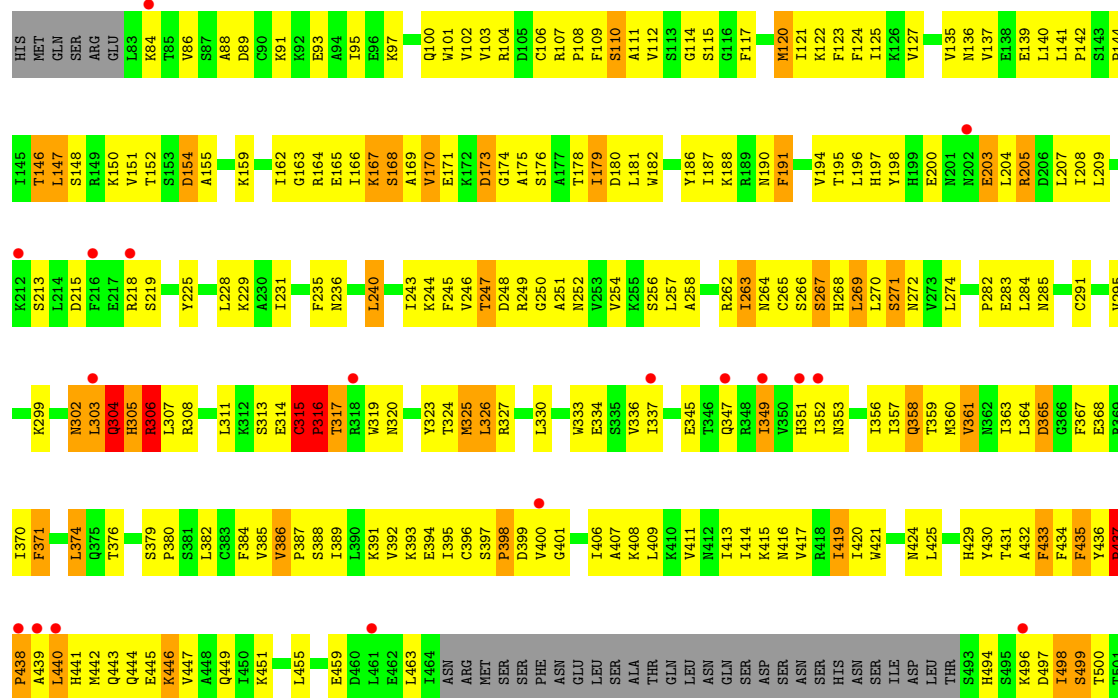
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	16	Total	C	N	O	P	0	0	0
			323	158	46	104	15			
3	F	16	Total	C	N	O	P	0	0	0
			323	158	46	104	15			
3	J	16	Total	C	N	O	P	0	0	0
			323	158	46	104	15			
3	L	16	Total	C	N	O	P	0	0	0
			323	158	46	104	15			

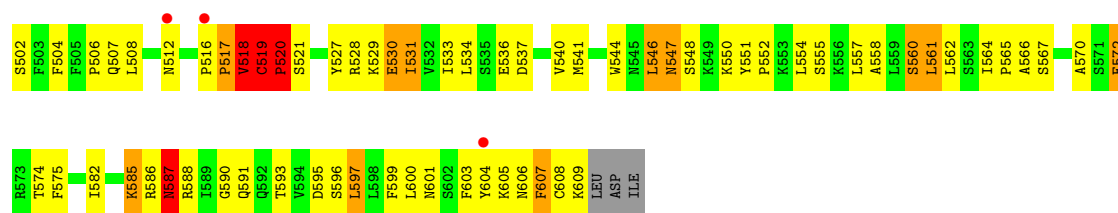
- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		
4	B	1	Total	Na	0	0
			1	1		
4	G	1	Total	Na	0	0
			1	1		
4	H	1	Total	Na	0	0
			1	1		

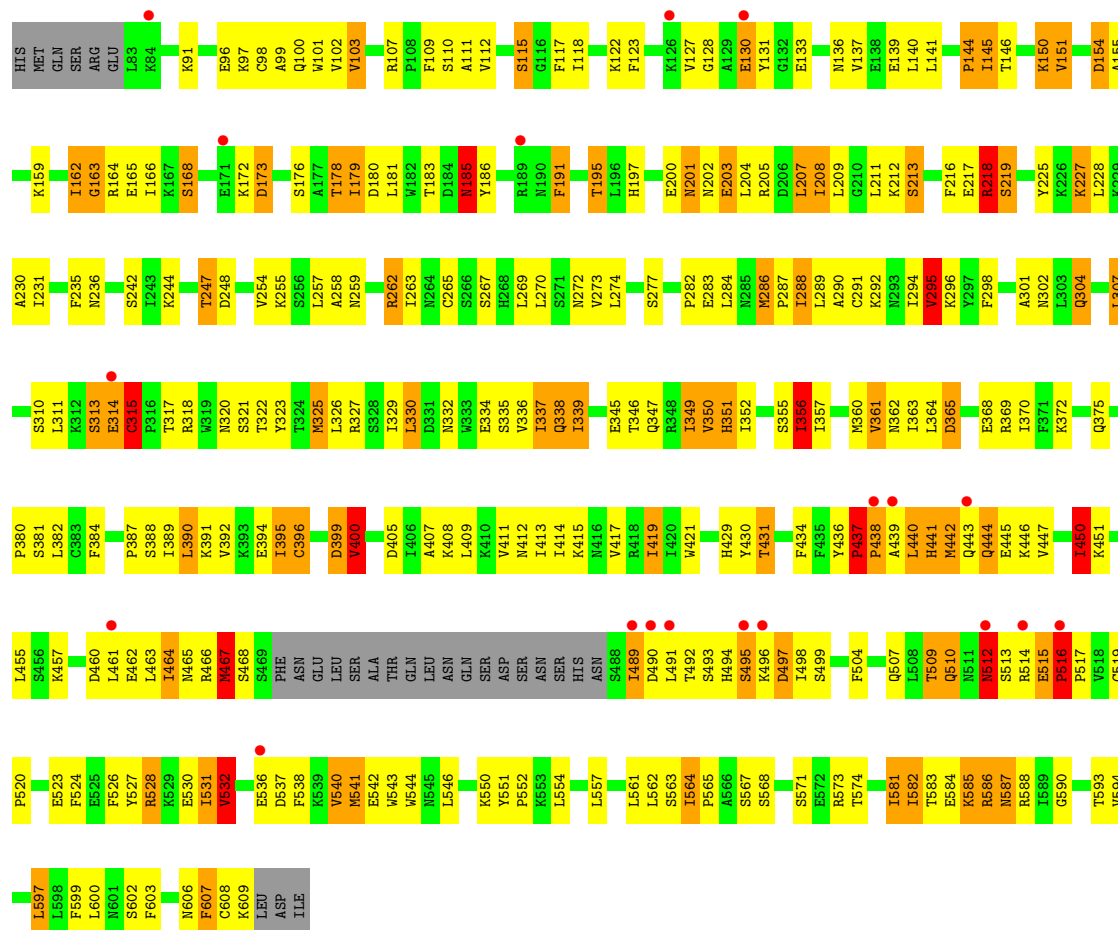


• Molecule 1: TRANSPOSASE





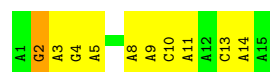
● Molecule 1: TRANSPOSASE



● Molecule 2: TERMINAL INVERTED REPEAT



● Molecule 2: TERMINAL INVERTED REPEAT



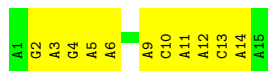
- Molecule 2: TERMINAL INVERTED REPEAT

Chain I:  27% 67% 7%

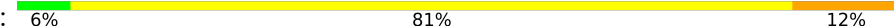


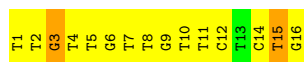
- Molecule 2: TERMINAL INVERTED REPEAT

Chain K:  27% 73%



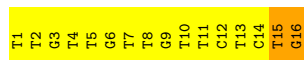
- Molecule 3: TERMINAL INVERTED REPEAT

Chain D:  6% 81% 12%

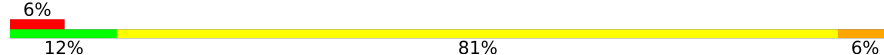


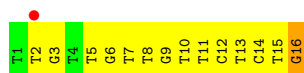
- Molecule 3: TERMINAL INVERTED REPEAT

Chain F:  88% 12%

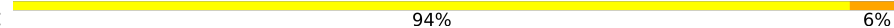


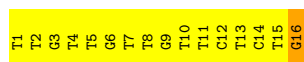
- Molecule 3: TERMINAL INVERTED REPEAT

Chain J:  6% 12% 81% 6%



- Molecule 3: TERMINAL INVERTED REPEAT

Chain L:  94% 6%



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	265.15Å 265.15Å 157.64Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 3.40 30.00 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.3 (30.00-3.40) 99.1 (30.00-3.40)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 3.31Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.212 , 0.254 0.212 , 0.250	Depositor DCC
R_{free} test set	947 reflections (1.00%)	wwPDB-VP
Wilson B-factor (Å ²)	88.4	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 111.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	18737	wwPDB-VP
Average B, all atoms (Å ²)	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.26% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	1.23	11/4120 (0.3%)	1.57	78/5561 (1.4%)
1	B	1.01	7/4146 (0.2%)	1.41	55/5595 (1.0%)
1	G	0.84	2/4092 (0.0%)	1.33	46/5524 (0.8%)
1	H	0.94	7/4168 (0.2%)	1.37	43/5626 (0.8%)
2	C	0.71	0/349	0.98	1/536 (0.2%)
2	E	0.63	0/349	0.92	0/536
2	I	0.44	0/349	0.76	0/536
2	K	0.48	0/349	0.76	0/536
3	D	0.71	0/358	1.03	1/552 (0.2%)
3	F	0.61	0/358	1.08	2/552 (0.4%)
3	J	0.47	0/358	0.88	0/552
3	L	0.52	0/358	0.95	0/552
All	All	0.96	27/19354 (0.1%)	1.35	226/26658 (0.8%)

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	B	0	2
1	G	0	2
1	H	0	3
2	C	0	3
2	E	0	2
2	I	0	1
2	K	0	1
3	D	0	1
3	F	0	1
3	J	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
3	L	0	1
All	All	0	20

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	438	PRO	CA-C	11.90	1.69	1.52
1	A	541	MET	SD-CE	11.26	2.07	1.79
1	H	438	PRO	CA-C	10.38	1.67	1.52
1	B	438	PRO	CA-C	10.23	1.67	1.52
1	G	438	PRO	CA-C	9.59	1.66	1.52

The worst 5 of 226 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	437	PRO	CA-C-N	-13.04	103.55	119.84
1	G	437	PRO	C-N-CA	-13.04	103.55	119.84
1	H	437	PRO	CA-C-N	-12.38	104.36	119.84
1	H	437	PRO	C-N-CA	-12.38	104.36	119.84
1	A	438	PRO	CA-C-N	11.84	138.30	120.31

There are no chirality outliers.

5 of 20 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	186	TYR	Sidechain
1	A	437	PRO	Mainchain
1	B	186	TYR	Sidechain
1	B	527	TYR	Sidechain
2	C	3	DA	Sidechain

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	4041	0	4097	293	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	4066	0	4120	250	0
1	G	4013	0	4072	250	0
1	H	4089	0	4148	238	0
2	C	308	0	165	12	0
2	E	308	0	167	18	0
2	I	308	0	167	21	0
2	K	308	0	167	16	0
3	D	323	0	186	30	0
3	F	323	0	188	35	0
3	J	323	0	188	38	0
3	L	323	0	188	45	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
All	All	18737	0	17853	1154	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 32.

The worst 5 of 1154 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:B:78:MET:SD	1:B:78:MET:CG	2.02	1.45
1:A:541:MET:CE	1:A:541:MET:SD	2.07	1.42
1:A:438:PRO:CG	1:A:438:PRO:CB	1.75	1.41
3:D:2:DT:H2''	3:D:3:DG:H5''	1.29	1.14
3:J:2:DT:H2''	3:J:3:DG:H5''	1.15	1.13

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	498/536 (93%)	400 (80%)	80 (16%)	18 (4%)	2	16
1	B	501/536 (94%)	399 (80%)	83 (17%)	19 (4%)	2	15
1	G	495/536 (92%)	412 (83%)	61 (12%)	22 (4%)	2	13
1	H	505/536 (94%)	401 (79%)	75 (15%)	29 (6%)	1	9
All	All	1999/2144 (93%)	1612 (81%)	299 (15%)	88 (4%)	2	13

5 of 88 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	218	ARG
1	A	219	SER
1	A	444	GLN
1	A	445	GLU
1	A	494	HIS

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	461/494 (93%)	367 (80%)	94 (20%)	1	4
1	B	464/494 (94%)	387 (83%)	77 (17%)	2	10
1	G	458/494 (93%)	384 (84%)	74 (16%)	2	10
1	H	468/494 (95%)	387 (83%)	81 (17%)	2	8
All	All	1851/1976 (94%)	1525 (82%)	326 (18%)	2	8

5 of 326 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	530	GLU
1	H	356	ILE
1	G	560	SER
1	H	185	ASN
1	H	450	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	591	GLN
1	G	338	GLN
1	H	512	ASN
1	H	416	ASN
1	B	601	ASN

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](#)

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 ⓘ

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	502/536 (93%)	-0.14	11 (2%) 62 47	16, 68, 129, 200	0
1	B	505/536 (94%)	-0.16	10 (1%) 65 49	45, 83, 137, 200	0
1	G	499/536 (93%)	0.34	21 (4%) 40 29	63, 116, 168, 200	0
1	H	509/536 (94%)	0.08	19 (3%) 45 32	48, 102, 159, 200	0
2	C	15/15 (100%)	0.13	0 100 100	61, 82, 164, 170	0
2	E	15/15 (100%)	0.26	0 100 100	66, 91, 159, 177	0
2	I	15/15 (100%)	0.44	0 100 100	103, 127, 188, 191	0
2	K	15/15 (100%)	0.20	0 100 100	82, 118, 185, 195	0
3	D	16/16 (100%)	0.00	0 100 100	62, 94, 149, 162	0
3	F	16/16 (100%)	0.11	0 100 100	64, 104, 162, 164	0
3	J	16/16 (100%)	0.62	1 (6%) 26 20	109, 141, 184, 190	0
3	L	16/16 (100%)	0.25	0 100 100	84, 140, 178, 183	0
All	All	2139/2268 (94%)	0.04	62 (2%) 53 39	16, 95, 161, 200	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	515	GLU	6.3
1	A	516	PRO	4.8
1	B	495	SER	4.7
1	G	347	GLN	4.1
1	H	496	LYS	3.8

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
4	NA	G	1610	1/1	0.90	0.07	77,77,77,77	0
4	NA	B	1610	1/1	0.93	0.06	57,57,57,57	0
4	NA	A	1610	1/1	0.95	0.05	44,44,44,44	0
4	NA	H	1610	1/1	0.98	0.04	58,58,58,58	0

6.5 Other polymers [i](#)

There are no such residues in this entry.