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 ⓘ

Aug 8, 2023 – 06:46 AM EDT

PDB ID	:	1Q2S
Title	:	Chemical trapping and crystal structure of a catalytic tRNA guanine transg-lycosylase covalent intermediate
Authors	:	Xie, W.; Liu, X.; Huang, R.H.
Deposited on	:	2003-07-25
Resolution	:	3.20 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

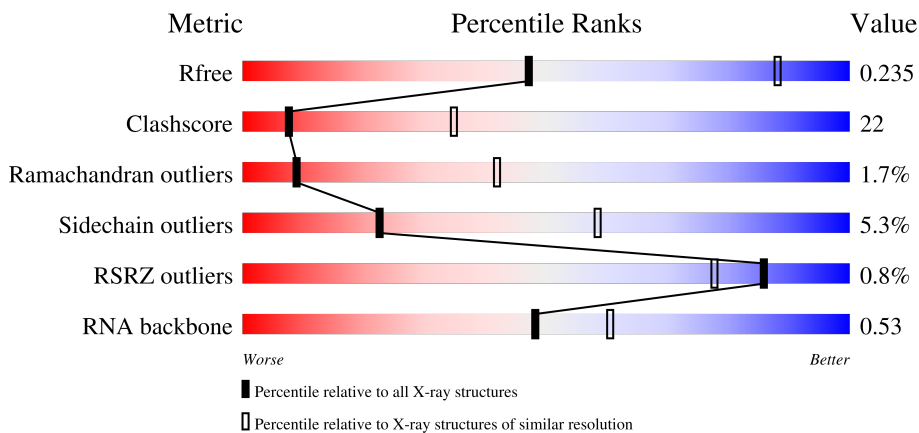|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 3.20 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		0.235
Clashscore		22
Ramachandran outliers		1.7%
Sidechain outliers		5.3%
RSRZ outliers		0.8%
RNA backbone		0.53

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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)
RNA backbone	3102	1010 (3.50-2.90)

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	E	20	35% 25% 30% 10%
2	F	20	35% 30% 25% 10%
3	A	386	56% 37% · ·
3	B	386	2% 57% 36% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	386	60% 33%
3	D	386	59% 33%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 12619 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 RNA chain called 5'-R(*AP*GP*CP*AP*CP*GP*GP*CP*UP*(PQ1)P*U P*AP*AP*AP*CP*CP*GP*UP*GP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	20	Total 427	C 193	N 79	O 136	P 19	0	0	0

- Molecule 2 is a RNA chain called RNA (5'-R(*AP*GP*CP*AP*CP*GP*GP*CP*UP*(N)P *UP*AP*AP*AP*CP*CP*GP*UP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	20	Total 414	C 186	N 74	O 135	P 19	0	0	0

- Molecule 3 is a protein called Queuine tRNA-ribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	376	Total 2933	C 1833	N 534	O 545	S 21	0	0	0
3	B	372	Total 2902	C 1817	N 526	O 538	S 21	0	0	0
3	C	376	Total 2933	C 1833	N 534	O 545	S 21	0	0	0
3	D	372	Total 2902	C 1817	N 526	O 538	S 21	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P28720
B	1	MET	-	initiating methionine	UNP P28720
C	1	MET	-	initiating methionine	UNP P28720
D	1	MET	-	initiating methionine	UNP P28720

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Zn 1	0	0
4	B	1	Total 1	Zn 1	0	0
4	C	1	Total 1	Zn 1	0	0
4	D	1	Total 1	Zn 1	0	0

- Molecule 5 is 9-DEAZAGUANINE (three-letter code: 9DG) (formula: $C_6H_6N_4O$).
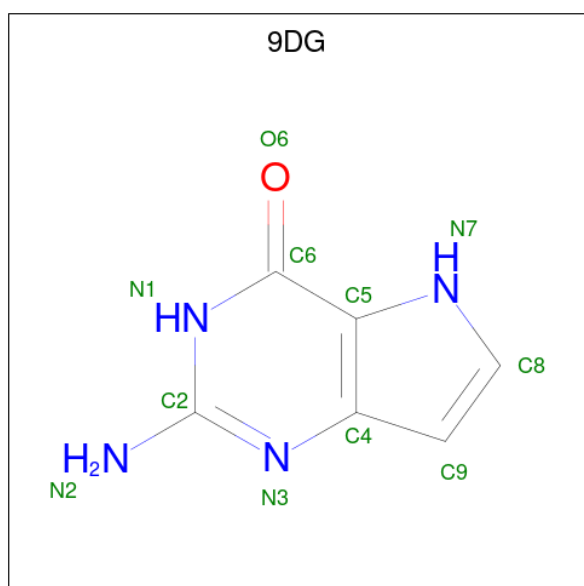
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total 11	C 6	N 4	O 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	E	5	Total 5	O 5	0	0
6	F	1	Total 1	O 1	0	0
6	A	23	Total 23	O 23	0	0
6	B	11	Total 11	O 11	0	0
6	C	27	Total 27	O 27	0	0

Continued on next page...
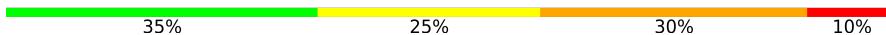
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	26	Total 26	O 26	0	0

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: 5'-R(*AP*GP*CP*AP*CP*GP*GP*CP*UP*(PQ1)P*UP*AP*AP*AP*CP*CP*G P*UP*GP*C)-3'

• Molecule 2: RNA (5'-R(*AP*GP*CP*AP*CP*GP*GP*CP*UP*(N)P*UP*AP*AP*AP*CP*C P*GP*UP*GP*C)-3')

• Molecule 3: Queuine tRNA-ribosyltransferase

• Molecule 3: Queuine tRNA-ribosyltransferase

Chain B:

• Molecule 3: Queuine tRNA-ribosyltransferase

Chain C:

• Molecule 3: Queuine tRNA-ribosyltransferase

Chain D:

E273 R274 G275 I276 D277 D280 L283 P284 S287 G288 R289 Q292 G298 P299 I300 N301 I302 R306 L311 L314 E317 C318 H319 C320 Q324 R328 L334 I335 R336 L341 G342 T347 A352 F353 Q356 L357 M358 Q359 K360 I361 R362 D363 S364

I365 F370 S371 Q372 D376 A379 R380 Y381 F382 ALA ARG ASN SER
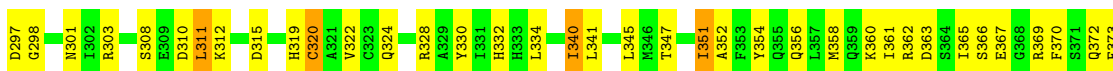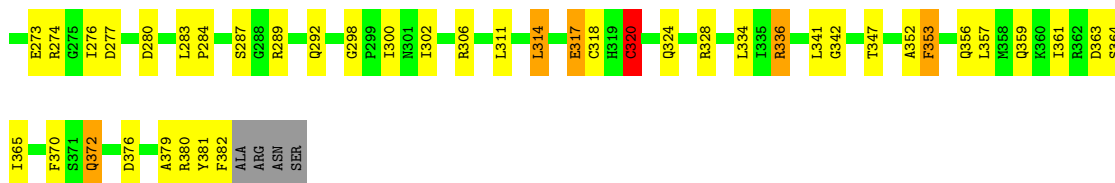
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	261.96Å 261.96Å 55.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.20 46.31 – 3.19	Depositor EDS
% Data completeness (in resolution range)	96.9 (30.00-3.20) 96.5 (46.31-3.19)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.48 (at 3.19Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.180 , 0.235 0.186 , 0.235	Depositor DCC
R_{free} test set	2549 reflections (7.99%)	wwPDB-VP
Wilson B-factor (Å^2)	49.0	Xtriage
Anisotropy	0.230	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.8	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.017 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	12619	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: PQ1, ZN, 9DG

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	E	0.50	0/448	1.01	3/694 (0.4%)
2	F	0.56	0/448	1.05	6/694 (0.9%)
3	A	0.41	0/2995	0.63	1/4032 (0.0%)
3	B	0.38	0/2964	0.58	0/3992
3	C	0.42	0/2995	0.63	1/4032 (0.0%)
3	D	0.40	0/2964	0.60	0/3992
All	All	0.41	0/12814	0.65	11/17436 (0.1%)

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
2	F	1	1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	35	U	C2'-C3'-O3'	8.35	127.88	109.50
1	E	35	U	C2'-C3'-O3'	8.16	127.45	109.50
2	F	37	A	C4'-C3'-O3'	7.02	127.04	113.00
1	E	36	A	C2'-C3'-O3'	6.80	124.58	113.70
2	F	37	A	C2'-C3'-O3'	6.48	124.08	113.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	F	37	A	C3'

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	37	A	Sidechain

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	E	427	0	223	16	0
2	F	414	0	216	17	0
3	A	2933	0	2877	142	0
3	B	2902	0	2848	132	0
3	C	2933	0	2877	129	0
3	D	2902	0	2848	125	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	C	11	0	6	3	0
6	A	23	0	0	3	0
6	B	11	0	0	2	0
6	C	27	0	0	3	0
6	D	26	0	0	2	0
6	E	5	0	0	2	0
6	F	1	0	0	0	0
All	All	12619	0	11895	537	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 22.

The worst 5 of 537 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)
3:D:52:LYS:NZ	3:D:289:ARG:HB2	1.66	1.10
3:B:52:LYS:NZ	3:B:289:ARG:HB2	1.72	1.04
3:D:52:LYS:HZ1	3:D:289:ARG:HB2	0.91	1.03
3:A:340:ILE:HD12	3:A:340:ILE:H	1.24	1.01
3:B:52:LYS:HZ1	3:B:289:ARG:HB2	0.85	1.00

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	
3	A	374/386 (97%)	333 (89%)	35 (9%)	6 (2%)	9	43
3	B	370/386 (96%)	327 (88%)	37 (10%)	6 (2%)	9	43
3	C	374/386 (97%)	332 (89%)	35 (9%)	7 (2%)	8	39
3	D	370/386 (96%)	325 (88%)	39 (10%)	6 (2%)	9	43
All	All	1488/1544 (96%)	1317 (88%)	146 (10%)	25 (2%)	9	42

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	110	SER
3	B	47	THR
3	C	110	SER
3	D	47	THR
3	A	131	SER

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	
3	A	307/315 (98%)	293 (95%)	14 (5%)	27	63
3	B	304/315 (96%)	288 (95%)	16 (5%)	22	58
3	C	307/315 (98%)	290 (94%)	17 (6%)	21	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	D	304/315 (96%)	286 (94%)	18 (6%)	19	54
All	All	1222/1260 (97%)	1157 (95%)	65 (5%)	22	58

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	237	GLN
3	D	280	ASP
3	B	266	ASP
3	B	237	GLN
3	D	314	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	319	HIS
3	C	257	HIS
3	D	257	HIS
3	C	213	GLN
3	C	304	ASN

5.3.3   RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	E	18/20 (90%)	7 (38%)	2 (11%)
2	F	18/20 (90%)	4 (22%)	3 (16%)
All	All	36/40 (90%)	11 (30%)	5 (13%)

5 of 11 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	E	32	C
1	E	33	U
1	E	35	U
1	E	36	A
1	E	37	A

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	E	35	U
1	E	36	A
2	F	33	U
2	F	35	U
2	F	37	A

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

1 non-standard protein/DNA/RNA residue is 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	#$|Z| > 2$
1	PQ1	E	34	1	21,27,28	0.92	1 (4%)	22,40,43	1.64	5 (22%)

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
1	PQ1	E	34	1	-	0/3/27/28	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	34	PQ1	C5-C4	-2.60	1.37	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	34	PQ1	C5-C6-N1	-4.11	111.88	115.36
1	E	34	PQ1	O3'-C3'-C2'	3.84	124.24	111.82
1	E	34	PQ1	O3'-C3'-C4'	2.63	118.67	111.05
1	E	34	PQ1	O6-C6-C5	2.34	128.89	123.98
1	E	34	PQ1	C2'-C3'-C4'	2.30	107.11	102.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	34	PQ1	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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	$\#|Z| > 2$
5	9DG	C	387	-	10,12,12	2.29	3 (30%)	9,17,17	1.94	3 (33%)

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
5	9DG	C	387	-	-	-	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	387	9DG	C9-C8	-5.15	1.31	1.38
5	C	387	9DG	C5-C6	-3.36	1.40	1.47
5	C	387	9DG	C6-N1	2.17	1.41	1.37

wwPDB X-ray Structure Validation Summary Report

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	387	9DG	C4-N3-C2	2.98	120.83	117.93
5	C	387	9DG	C5-C6-N1	2.90	119.07	113.95
5	C	387	9DG	C2-N1-C6	-2.88	119.80	125.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	387	9DG	3	0

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(Å²)	Q<0.9
1	E	19/20 (95%)	-0.25	0	100	100	20, 31, 46, 61	0
2	F	19/20 (95%)	-0.29	0	100	100	18, 30, 50, 60	0
3	A	376/386 (97%)	-0.34	1 (0%)	94	92	7, 28, 56, 71	0
3	B	372/386 (96%)	0.12	8 (2%)	62	48	8, 50, 88, 100	0
3	C	376/386 (97%)	-0.39	0	100	100	8, 26, 56, 69	0
3	D	372/386 (96%)	-0.15	3 (0%)	86	78	7, 40, 83, 100	0
All	All	1534/1584 (96%)	-0.19	12 (0%)	86	78	7, 36, 75, 100	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	113	SER	4.0
3	D	114	LEU	3.0
3	B	161	TYR	2.8
3	D	112	SER	2.8
3	A	386	SER	2.6

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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(Å²)	Q<0.9
1	PQ1	E	34	25/26	0.96	0.16	32,38,40,40	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(Å^2)	Q<0.9
5	9DG	C	387	11/11	0.95	0.20	28,29,30,30	0
4	ZN	D	404	1/1	0.99	0.06	24,24,24,24	0
4	ZN	C	403	1/1	0.99	0.10	19,19,19,19	0
4	ZN	B	402	1/1	1.00	0.07	16,16,16,16	0
4	ZN	A	401	1/1	1.00	0.05	22,22,22,22	0

6.5 Other polymers (i)

There are no such residues in this entry.