



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 09:47 PM GMT

PDB ID : 1QK3  
Title : TOXOPLASMA GONDII HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE GMP COMPLEX  
Authors : Heroux, A.; White, E.L.; Ross, L.J.; Borhani, D.W.  
Deposited on : 1999-07-09  
Resolution : 1.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

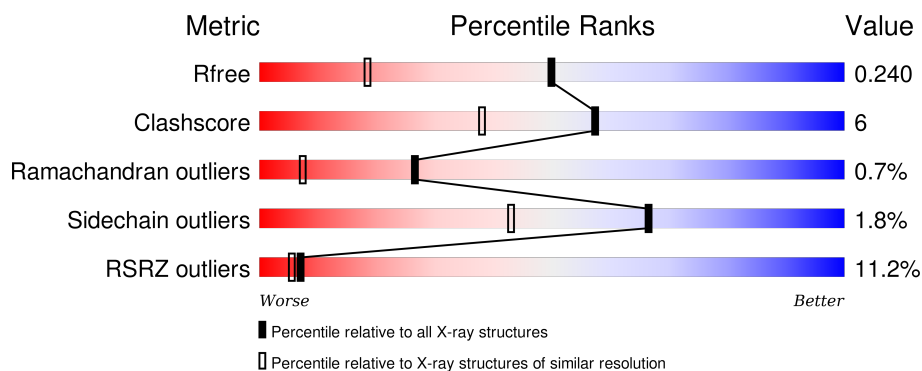
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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))
$R_{free}$	91344	1226 (1.66-1.66)
Clashscore	102246	1323 (1.66-1.66)
Ramachandran outliers	100387	1295 (1.66-1.66)
Sidechain outliers	100360	1295 (1.66-1.66)
RSRZ outliers	91569	1227 (1.66-1.66)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	233	<div> <div>9%</div> <div>84%</div> <div>14%</div> <div>.</div> </div>
1	B	233	<div> <div>8%</div> <div>81%</div> <div>15%</div> <div>..</div> </div>
1	C	233	<div> <div>10%</div> <div>75%</div> <div>14%</div> <div>9%</div> <div>.</div> </div>
1	D	233	<div> <div>17%</div> <div>78%</div> <div>17%</div> <div>..</div> </div>

## 2 Entry composition [i](#)

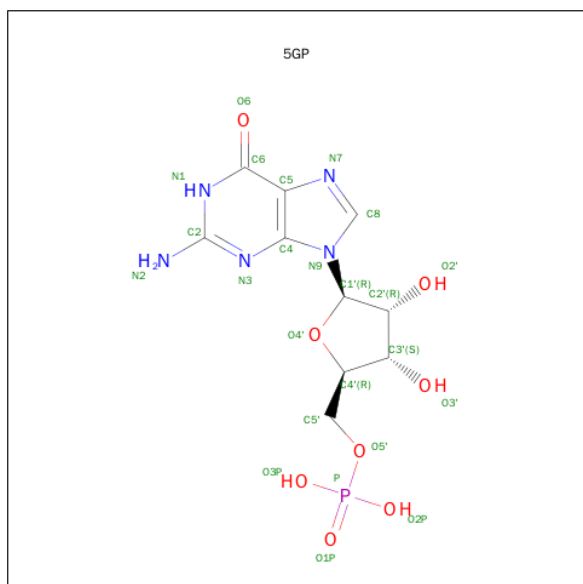
There are 3 unique types of molecules in this entry. The entry contains 8019 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	228	Total	C	N	O	S	0	8	0
			1860	1200	305	346	9			
1	B	231	Total	C	N	O	S	0	10	0
			1874	1204	308	354	8			
1	C	213	Total	C	N	O	S	0	7	0
			1717	1113	280	315	9			
1	D	227	Total	C	N	O	S	0	5	0
			1827	1181	297	340	9			

- Molecule 2 is GUANOSINE-5'-MONOPHOSPHATE (three-letter code: 5GP) (formula:  $C_{10}H_{14}N_5O_8P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			24	10	5	8	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	P	0	0
			24	10	5	8	1		
2	C	1	Total	C	N	O	P	0	0
			24	10	5	8	1		
2	D	1	Total	C	N	O	P	0	0
			24	10	5	8	1		

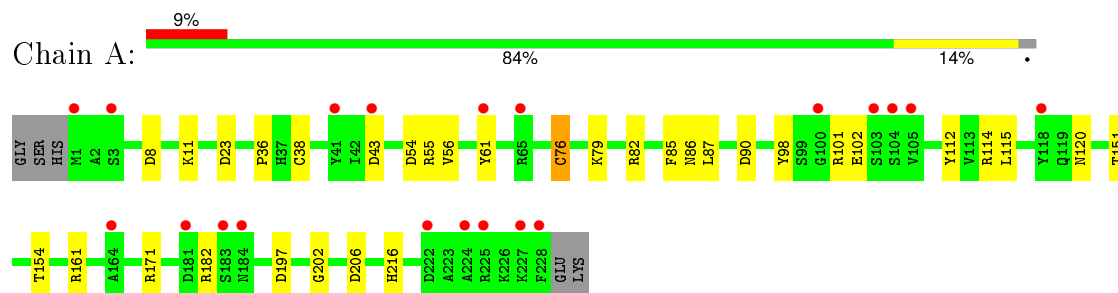
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	183	Total	O	0	0
			183	183		
3	B	160	Total	O	0	0
			160	160		
3	C	152	Total	O	0	0
			152	152		
3	D	150	Total	O	0	0
			150	150		

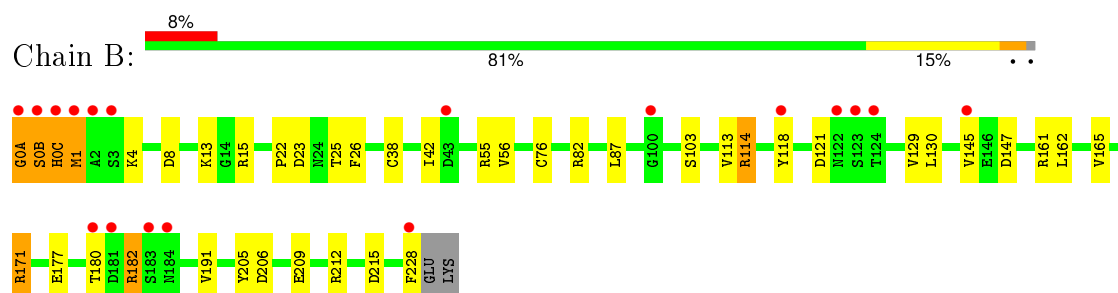
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

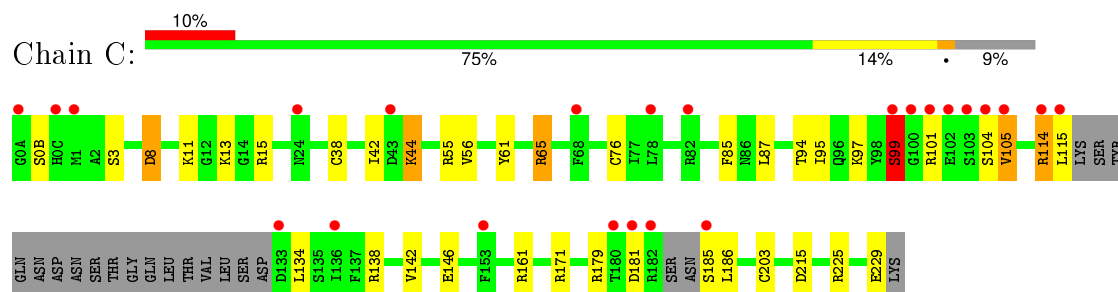
#### • Molecule 1: HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE



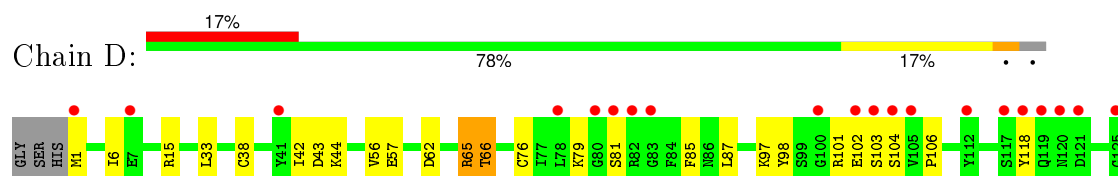
#### • Molecule 1: HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE

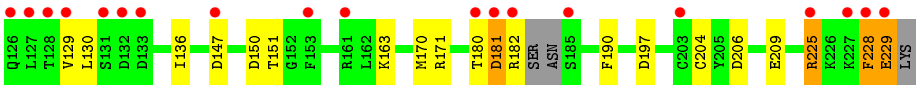


#### • Molecule 1: HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE



#### • Molecule 1: HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.45Å 90.84Å 80.26Å 90.00° 92.53° 90.00°	Depositor
Resolution (Å)	12.50 – 1.65 10.55 – 1.65	Depositor EDS
% Data completeness (in resolution range)	96.4 (12.50-1.65) 96.7 (10.55-1.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 1.65Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.202 , 0.231 0.215 , 0.240	Depositor DCC
$R_{free}$ test set	5403 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.7	Xtriage
Anisotropy	0.298	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 59.5	EDS
Estimated twinning fraction	0.020 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 108506 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8019	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 5GP

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	0.65	0/1943	1.45	28/2624 (1.1%)
1	B	0.77	1/1975 (0.1%)	1.82	32/2669 (1.2%)
1	C	0.69	0/1791	1.35	20/2420 (0.8%)
1	D	0.65	0/1895	1.37	18/2561 (0.7%)
All	All	0.69	1/7604 (0.0%)	1.52	98/10274 (1.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	0(A)	GLY	N-CA	6.31	1.55	1.46

All (98) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	82	ARG	NE-CZ-NH1	30.59	135.60	120.30
1	B	82	ARG	NE-CZ-NH2	-30.13	105.24	120.30
1	B	161	ARG	NE-CZ-NH2	-23.37	108.61	120.30
1	B	114	ARG	NE-CZ-NH1	19.01	129.80	120.30
1	D	225	ARG	NE-CZ-NH2	-14.85	112.88	120.30
1	A	55	ARG	NE-CZ-NH1	14.12	127.36	120.30
1	A	82	ARG	NE-CZ-NH2	-13.65	113.47	120.30
1	B	145[A]	VAL	CG1-CB-CG2	-13.64	89.07	110.90
1	B	145[B]	VAL	CG1-CB-CG2	-13.64	89.07	110.90
1	A	101	ARG	NE-CZ-NH1	13.20	126.90	120.30
1	A	101	ARG	NE-CZ-NH2	-12.12	114.24	120.30
1	D	225	ARG	NE-CZ-NH1	11.99	126.30	120.30
1	B	171	ARG	NE-CZ-NH2	-11.90	114.35	120.30
1	C	55	ARG	NE-CZ-NH2	-11.84	114.38	120.30
1	A	171	ARG	NE-CZ-NH1	11.59	126.10	120.30
1	B	114	ARG	NE-CZ-NH2	-11.34	114.63	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	171	ARG	NE-CZ-NH1	11.08	125.84	120.30
1	B	0(B)	SER	C-N-CA	10.50	147.95	121.70
1	C	138	ARG	NE-CZ-NH2	10.46	125.53	120.30
1	C	138	ARG	NE-CZ-NH1	-10.45	115.08	120.30
1	D	147[A]	ASP	CB-CG-OD1	10.36	127.62	118.30
1	D	147[B]	ASP	CB-CG-OD1	10.36	127.62	118.30
1	B	8	ASP	CB-CG-OD2	10.33	127.59	118.30
1	B	161	ARG	NH1-CZ-NH2	9.77	130.14	119.40
1	B	212	ARG	NE-CZ-NH1	9.06	124.83	120.30
1	C	55	ARG	NE-CZ-NH1	9.03	124.81	120.30
1	B	0(B)	SER	CA-C-O	8.82	138.62	120.10
1	B	205	TYR	CB-CG-CD1	-8.49	115.91	121.00
1	A	161	ARG	NE-CZ-NH2	-8.34	116.13	120.30
1	C	161	ARG	NE-CZ-NH1	8.25	124.42	120.30
1	A	90[A]	ASP	CB-CG-OD1	7.97	125.47	118.30
1	A	90[B]	ASP	CB-CG-OD1	7.97	125.47	118.30
1	D	15	ARG	NE-CZ-NH1	7.93	124.27	120.30
1	A	206	ASP	CB-CG-OD1	7.88	125.39	118.30
1	B	206	ASP	CB-CG-OD1	7.83	125.35	118.30
1	C	171	ARG	NE-CZ-NH2	-7.75	116.42	120.30
1	A	114	ARG	NE-CZ-NH2	-7.60	116.50	120.30
1	C	15	ARG	NE-CZ-NH1	7.41	124.01	120.30
1	C	179	ARG	NE-CZ-NH2	7.34	123.97	120.30
1	A	101	ARG	CD-NE-CZ	7.20	133.67	123.60
1	B	147	ASP	CB-CG-OD1	7.15	124.73	118.30
1	D	97	LYS	CA-CB-CG	7.04	128.89	113.40
1	A	98	TYR	CB-CG-CD2	6.90	125.14	121.00
1	C	215	ASP	CB-CG-OD2	6.89	124.51	118.30
1	A	90[A]	ASP	CB-CG-OD2	-6.82	112.16	118.30
1	A	90[B]	ASP	CB-CG-OD2	-6.82	112.16	118.30
1	C	215	ASP	CB-CG-OD1	-6.66	112.31	118.30
1	B	15	ARG	NE-CZ-NH2	-6.66	116.97	120.30
1	C	146[A]	GLU	OE1-CD-OE2	-6.65	115.32	123.30
1	C	146[B]	GLU	OE1-CD-OE2	-6.65	115.32	123.30
1	B	215	ASP	CB-CG-OD1	6.61	124.25	118.30
1	D	57	GLU	OE1-CD-OE2	-6.53	115.46	123.30
1	D	209	GLU	OE1-CD-OE2	-6.52	115.48	123.30
1	B	0(B)	SER	O-C-N	-6.48	112.33	122.70
1	A	43	ASP	CB-CG-OD1	6.45	124.11	118.30
1	A	102	GLU	OE1-CD-OE2	-6.33	115.70	123.30
1	C	65	ARG	NE-CZ-NH2	-6.29	117.16	120.30
1	A	55	ARG	NH1-CZ-NH2	-6.27	112.51	119.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1	MET	CG-SD-CE	6.13	110.01	100.20
1	A	23	ASP	CB-CG-OD1	6.12	123.80	118.30
1	B	0(B)	SER	N-CA-CB	-5.99	101.51	110.50
1	B	0(A)	GLY	CA-C-N	-5.98	104.04	117.20
1	D	206	ASP	CB-CG-OD1	5.92	123.63	118.30
1	B	55	ARG	NE-CZ-NH1	5.92	123.26	120.30
1	B	26	PHE	CB-CG-CD1	-5.90	116.67	120.80
1	A	54	ASP	CB-CG-OD2	-5.80	113.08	118.30
1	C	8	ASP	CB-CG-OD1	5.79	123.51	118.30
1	B	55	ARG	NE-CZ-NH2	5.79	123.19	120.30
1	C	225	ARG	NE-CZ-NH1	5.78	123.19	120.30
1	B	209	GLU	OE1-CD-OE2	-5.75	116.40	123.30
1	A	171	ARG	NH1-CZ-NH2	-5.73	113.09	119.40
1	B	182	ARG	NE-CZ-NH1	-5.73	117.43	120.30
1	D	181	ASP	N-CA-C	-5.68	95.66	111.00
1	B	8	ASP	OD1-CG-OD2	-5.67	112.53	123.30
1	A	82	ARG	NH1-CZ-NH2	5.66	125.63	119.40
1	C	229	GLU	CA-C-O	-5.62	108.30	120.10
1	D	98	TYR	CB-CG-CD2	-5.61	117.64	121.00
1	A	38	CYS	CB-CA-C	-5.41	99.58	110.40
1	B	55	ARG	NH1-CZ-NH2	-5.40	113.46	119.40
1	C	142	VAL	CG1-CB-CG2	-5.39	102.28	110.90
1	A	112	TYR	CB-CG-CD1	-5.39	117.77	121.00
1	C	134	LEU	N-CA-C	5.39	125.55	111.00
1	D	229	GLU	OE1-CD-OE2	-5.31	116.92	123.30
1	D	43	ASP	CB-CG-OD1	5.29	123.06	118.30
1	D	65	ARG	NE-CZ-NH2	-5.21	117.69	120.30
1	A	98	TYR	CB-CG-CD1	-5.21	117.88	121.00
1	B	121	ASP	CB-CG-OD2	5.18	122.96	118.30
1	A	112	TYR	CB-CG-CD2	5.18	124.11	121.00
1	A	197	ASP	CB-CG-OD2	5.14	122.92	118.30
1	D	147[A]	ASP	OD1-CG-OD2	-5.13	113.55	123.30
1	D	147[B]	ASP	OD1-CG-OD2	-5.13	113.55	123.30
1	D	228	PHE	C-N-CA	5.06	134.36	121.70
1	C	225	ARG	CD-NE-CZ	5.06	130.68	123.60
1	C	171	ARG	NE-CZ-NH1	5.05	122.83	120.30
1	D	171	ARG	NE-CZ-NH1	5.04	122.82	120.30
1	B	23	ASP	CB-CG-OD2	-5.02	113.78	118.30
1	A	76[A]	CYS	CB-CA-C	5.02	120.44	110.40
1	A	76[B]	CYS	CB-CA-C	5.02	120.44	110.40

There are no chirality outliers.

There are no planarity outliers.

## 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	1860	0	1826	12	0
1	B	1874	0	1823	24	1
1	C	1717	0	1663	22	0
1	D	1827	0	1788	38	1
2	A	24	0	12	0	0
2	B	24	0	12	0	0
2	C	24	0	12	0	0
2	D	24	0	12	0	0
3	A	183	0	0	1	0
3	B	160	0	0	1	1
3	C	152	0	0	2	0
3	D	150	0	0	2	1
All	All	8019	0	7148	87	2

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.

All (87) 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:101:ARG:NH1	1:D:33[A]:LEU:HD11	1.63	1.12
1:B:0(A):GLY:HA2	3:B:2131:HOH:O	1.60	1.00
1:C:101:ARG:HH12	1:D:33[A]:LEU:CD1	1.76	0.98
1:C:101:ARG:HH12	1:D:33[A]:LEU:HD11	0.81	0.96
1:D:44:LYS:HD2	3:D:2040:HOH:O	1.70	0.91
1:C:114:ARG:O	1:C:115:LEU:HB2	1.76	0.85
1:B:0(A):GLY:N	1:B:4:LYS:NZ	2.24	0.84
1:D:118:TYR:CD1	1:D:130:LEU:HD11	2.13	0.83
1:D:66[A]:THR:HG21	1:D:190:PHE:HZ	1.44	0.83
1:D:6:ILE:HG23	1:D:170:MET:HE1	1.63	0.78
1:D:66[A]:THR:HG21	1:D:190:PHE:CZ	2.20	0.76
1:C:101:ARG:HH12	1:D:33[B]:LEU:HD22	1.51	0.75

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:0(A):GLY:N	1:B:4:LYS:HZ1	1.86	0.73
1:D:182:ARG:NH1	1:D:197:ASP:OD2	2.25	0.70
1:B:0(A):GLY:N	1:B:4:LYS:HZ3	1.91	0.69
1:B:0(A):GLY:O	1:B:171:ARG:NE	2.19	0.66
1:C:44:LYS:HE3	3:C:2017:HOH:O	1.96	0.66
1:C:114:ARG:O	1:C:115:LEU:CB	2.45	0.65
1:C:101:ARG:HH12	1:D:33[B]:LEU:CD2	2.07	0.64
1:D:225:ARG:O	1:D:229:GLU:HG3	1.96	0.64
1:D:62:ASP:O	1:D:66[A]:THR:HG22	1.99	0.62
1:D:6:ILE:HG23	1:D:170:MET:CE	2.31	0.61
1:D:76:CYS:HB2	1:D:85:PHE:CE1	2.38	0.58
1:D:136:ILE:HG23	3:D:2085:HOH:O	2.03	0.58
1:C:95:ILE:O	1:C:99:SER:HB2	2.04	0.58
1:D:181:ASP:O	1:D:182:ARG:HB2	2.04	0.57
1:B:0(A):GLY:H3	1:B:4:LYS:HZ3	1.52	0.57
1:C:11:LYS:HD3	1:C:13:LYS:HE3	1.86	0.57
1:D:118:TYR:HB2	1:D:130:LEU:HG	1.87	0.56
1:C:76:CYS:HB2	1:C:85:PHE:CE1	2.40	0.56
1:B:0(A):GLY:N	1:B:13:LYS:O	2.29	0.55
1:A:61:TYR:OH	1:D:101:ARG:HD3	2.07	0.55
1:C:76:CYS:HB2	1:C:85:PHE:CD1	2.41	0.55
1:B:38:CYS:HB3	1:B:42[B]:ILE:HD12	1.90	0.54
1:D:151:THR:HG22	1:D:182:ARG:HG2	1.90	0.53
1:B:0(A):GLY:H1	1:B:4:LYS:HZ1	1.53	0.53
1:D:79:LYS:HE2	1:D:79:LYS:HA	1.91	0.53
1:D:163:LYS:HG2	1:D:170:MET:HE1	1.91	0.53
1:D:76:CYS:HB2	1:D:85:PHE:CD1	2.45	0.52
1:B:22:PRO:HD2	1:B:25:THR:HG21	1.90	0.52
1:C:8:ASP:HA	1:C:11:LYS:HD2	1.93	0.51
1:D:81:SER:OG	1:D:204:CYS:HB2	2.10	0.51
1:A:8:ASP:HA	1:A:11:LYS:HD2	1.92	0.51
1:B:22:PRO:O	1:B:25:THR:HG22	2.11	0.51
1:D:1:MET:HE3	1:D:65:ARG:HG2	1.92	0.51
1:B:56:VAL:HG21	1:B:87:LEU:HD23	1.93	0.50
1:A:36:PRO:HG2	1:B:103:SER:HB2	1.92	0.50
1:D:38:CYS:HB3	1:D:42[A]:ILE:HD12	1.94	0.50
1:D:150:ASP:OD1	1:D:182:ARG:HD3	2.13	0.49
1:C:104:SER:O	1:C:105:VAL:C	2.52	0.48
1:D:228:PHE:CD2	1:D:229:GLU:N	2.82	0.48
1:B:0(A):GLY:H1	1:B:4:LYS:NZ	2.10	0.48
1:B:0(A):GLY:H2	1:B:4:LYS:NZ	2.07	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1:MET:O	1:B:171:ARG:NH2	2.47	0.47
1:A:79:LYS:HE3	3:A:2124:HOH:O	2.13	0.47
1:D:104:SER:O	1:D:106:PRO:HD3	2.15	0.47
1:C:56:VAL:HG21	1:C:87:LEU:HD23	1.97	0.47
1:A:76[B]:CYS:HB2	1:A:85[B]:PHE:CE1	2.51	0.46
1:C:94[A]:THR:HG21	3:C:2077:HOH:O	2.15	0.46
1:B:114:ARG:HG2	1:B:129:VAL:HG11	1.98	0.45
1:D:1:MET:CE	1:D:65:ARG:HG2	2.47	0.44
1:D:102:GLU:O	1:D:103:SER:HB2	2.18	0.44
1:A:85[A]:PHE:CE1	1:A:86:ASN:OD1	2.70	0.44
1:C:38:CYS:HB3	1:C:42:ILE:HD12	2.00	0.43
1:C:101:ARG:CZ	1:D:33[A]:LEU:HD11	2.39	0.43
1:D:180:THR:OG1	1:D:182:ARG:HB2	2.17	0.43
1:B:162:LEU:O	1:B:165:VAL:HG22	2.18	0.43
1:D:181:ASP:O	1:D:182:ARG:CB	2.67	0.43
1:D:56:VAL:HG21	1:D:87:LEU:HD23	2.00	0.43
1:A:115:LEU:CD1	1:A:154:THR:HG23	2.48	0.42
1:C:115:LEU:HD23	1:C:115:LEU:HA	1.81	0.42
1:B:113[A]:VAL:HG22	1:B:130:LEU:HB2	2.02	0.42
1:C:61:TYR:CE2	1:C:65:ARG:HD2	2.55	0.42
1:C:61:TYR:CZ	1:C:65:ARG:HD2	2.56	0.41
1:A:115:LEU:HD11	1:A:154:THR:HG23	2.01	0.41
1:D:81:SER:HG	1:D:204:CYS:HB2	1.84	0.41
1:A:151:THR:O	1:A:182:ARG:NH1	2.50	0.41
1:C:101:ARG:HH22	1:D:33[A]:LEU:CD1	2.34	0.40
1:A:56:VAL:HG21	1:A:87:LEU:HD23	2.03	0.40
1:B:177[A]:GLU:HG2	1:B:191:VAL:HG11	2.03	0.40
1:D:228:PHE:CG	1:D:229:GLU:N	2.89	0.40
1:A:76[B]:CYS:HB2	1:A:85[B]:PHE:CD1	2.57	0.40
1:B:118:TYR:CE2	1:B:129:VAL:HB	2.56	0.40
1:A:202:GLY:HA3	1:A:216:HIS:O	2.21	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:2116:HOH:O	3:D:2139:HOH:O[2_858]	1.75	0.45
1:B:182:ARG:CD	1:D:229:GLU:OE1[2_858]	2.19	0.01

## 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	234/233 (100%)	221 (94%)	13 (6%)	0	100	100
1	B	239/233 (103%)	230 (96%)	7 (3%)	2 (1%)	24	5
1	C	214/233 (92%)	203 (95%)	7 (3%)	4 (2%)	10	0
1	D	228/233 (98%)	221 (97%)	7 (3%)	0	100	100
All	All	915/932 (98%)	875 (96%)	34 (4%)	6 (1%)	26	7

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	0(B)	SER
1	C	105	VAL
1	B	0(C)	HIS
1	C	114	ARG
1	C	203	CYS
1	C	99	SER

### 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	209/207 (101%)	208 (100%)	1 (0%)	92	86
1	B	212/207 (102%)	209 (99%)	3 (1%)	74	53
1	C	186/207 (90%)	178 (96%)	8 (4%)	35	9
1	D	202/207 (98%)	199 (98%)	3 (2%)	72	51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	809/828 (98%)	794 (98%)	15 (2%)	66	40

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	120	ASN
1	B	0(C)	HIS
1	B	180	THR
1	B	228	PHE
1	C	0(B)	SER
1	C	3	SER
1	C	44	LYS
1	C	97	LYS
1	C	99	SER
1	C	181	ASP
1	C	185	SER
1	C	186	LEU
1	D	66[A]	THR
1	D	66[B]	THR
1	D	129	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

4 ligands 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	$\# Z  > 2$
2	5GP	A	300	-	21,26,26	1.41	2 (9%)	25,40,40	2.54	5 (20%)
2	5GP	B	300	-	21,26,26	1.42	2 (9%)	25,40,40	2.32	3 (12%)
2	5GP	C	300	-	21,26,26	1.43	4 (19%)	25,40,40	2.16	5 (20%)
2	5GP	D	300	-	21,26,26	1.38	3 (14%)	25,40,40	2.22	4 (16%)

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
2	5GP	A	300	-	-	0/6/26/26	0/3/3/3
2	5GP	B	300	-	-	0/6/26/26	0/3/3/3
2	5GP	C	300	-	-	0/6/26/26	0/3/3/3
2	5GP	D	300	-	-	0/6/26/26	0/3/3/3

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	300	5GP	O4'-C4'	2.11	1.49	1.45
2	C	300	5GP	O4'-C1'	2.14	1.43	1.41
2	C	300	5GP	C2-N1	2.20	1.39	1.35
2	B	300	5GP	O4'-C1'	2.23	1.44	1.41
2	D	300	5GP	O4'-C1'	2.51	1.44	1.41
2	D	300	5GP	C2-N1	2.99	1.40	1.35
2	A	300	5GP	O4'-C1'	3.21	1.45	1.41
2	D	300	5GP	C6-N1	3.87	1.40	1.33
2	A	300	5GP	C6-N1	4.20	1.40	1.33
2	C	300	5GP	C6-N1	4.50	1.41	1.33
2	B	300	5GP	C6-N1	4.75	1.41	1.33

All (17) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	300	5GP	C5-C6-N1	-9.36	110.79	123.59
2	B	300	5GP	C5-C6-N1	-8.42	112.08	123.59
2	C	300	5GP	C5-C6-N1	-7.53	113.29	123.59
2	D	300	5GP	C5-C6-N1	-7.28	113.63	123.59
2	D	300	5GP	N3-C2-N1	-3.05	122.80	127.44
2	C	300	5GP	N3-C2-N1	-2.52	123.61	127.44
2	A	300	5GP	N2-C2-N1	-2.38	113.26	117.20
2	C	300	5GP	C4'-O4'-C1'	-2.00	107.52	109.72
2	C	300	5GP	O3'-C3'-C4'	2.06	117.24	111.05
2	A	300	5GP	C2'-C1'-N9	2.11	117.52	114.29
2	A	300	5GP	C4-C5-N7	2.18	111.49	109.48
2	B	300	5GP	O3P-P-O2P	2.63	117.38	107.38
2	D	300	5GP	N2-C2-N3	2.84	123.25	117.80
2	C	300	5GP	C6-N1-C2	5.05	122.95	115.94
2	D	300	5GP	C6-N1-C2	5.13	123.06	115.94
2	B	300	5GP	C6-N1-C2	5.81	124.00	115.94
2	A	300	5GP	C6-N1-C2	5.91	124.15	115.94

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

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	228/233 (97%)	0.46	20 (8%) 12 11	11, 19, 40, 56	0
1	B	231/233 (99%)	0.53	18 (7%) 16 14	10, 17, 43, 54	0
1	C	213/233 (91%)	0.64	24 (11%) 7 5	10, 18, 45, 70	0
1	D	227/233 (97%)	0.87	39 (17%) 2 2	10, 21, 51, 60	0
All	All	899/932 (96%)	0.63	101 (11%) 7 5	10, 18, 46, 70	0

All (101) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	103	SER	9.6
1	C	102	GLU	8.8
1	B	0(C)	HIS	8.7
1	C	0(A)	GLY	8.7
1	C	100	GLY	8.4
1	D	80	GLY	8.2
1	B	0(A)	GLY	8.1
1	D	118	TYR	8.1
1	A	228	PHE	8.0
1	A	184	ASN	7.7
1	D	229	GLU	7.6
1	C	101	ARG	7.0
1	B	228	PHE	6.8
1	C	104	SER	6.8
1	D	81	SER	5.4
1	B	3	SER	5.3
1	C	182	ARG	5.2
1	D	1	MET	5.1
1	D	82	ARG	5.1
1	D	181	ASP	4.9
1	C	133	ASP	4.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	122	ASN	4.6
1	B	181	ASP	4.5
1	D	228	PHE	4.5
1	A	43	ASP	4.5
1	A	1	MET	4.4
1	C	115	LEU	4.4
1	D	105	VAL	4.3
1	C	153	PHE	4.2
1	D	78	LEU	4.2
1	D	129	VAL	4.2
1	B	0(B)	SER	4.1
1	C	181	ASP	3.9
1	D	182	ARG	3.8
1	D	126	GLN	3.8
1	A	181	ASP	3.8
1	D	133	ASP	3.7
1	C	105	VAL	3.7
1	D	161	ARG	3.7
1	D	102	GLU	3.6
1	D	104	SER	3.6
1	C	0(C)	HIS	3.6
1	A	103	SER	3.6
1	D	119	GLN	3.5
1	B	100	GLY	3.4
1	B	43	ASP	3.4
1	A	183	SER	3.4
1	A	105	VAL	3.3
1	B	184	ASN	3.3
1	B	1	MET	3.3
1	B	183	SER	3.1
1	A	222	ASP	3.1
1	B	180	THR	3.1
1	B	118	TYR	3.1
1	A	104	SER	3.0
1	D	41	TYR	3.0
1	C	185	SER	3.0
1	D	83	GLY	2.9
1	D	227	LYS	2.9
1	D	125	GLY	2.9
1	A	227	LYS	2.8
1	B	2	ALA	2.8
1	C	114	ARG	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	180	THR	2.8
1	D	128	THR	2.8
1	A	65	ARG	2.8
1	C	136[A]	ILE	2.7
1	B	123	SER	2.7
1	D	185	SER	2.7
1	D	153	PHE	2.7
1	C	82	ARG	2.6
1	D	132	ASP	2.6
1	A	41	TYR	2.6
1	D	127	LEU	2.6
1	A	3	SER	2.5
1	A	61	TYR	2.5
1	A	164	ALA	2.5
1	D	121	ASP	2.5
1	D	180	THR	2.4
1	D	7	GLU	2.4
1	D	103	SER	2.4
1	C	24	ASN	2.4
1	C	68	PHE	2.4
1	C	99	SER	2.3
1	D	225	ARG	2.3
1	D	131	SER	2.3
1	C	1	MET	2.3
1	A	224	ALA	2.2
1	C	78	LEU	2.2
1	D	117	SER	2.2
1	A	118	TYR	2.2
1	A	100	GLY	2.2
1	D	203[A]	CYS	2.2
1	B	145[A]	VAL	2.1
1	A	225	ARG	2.1
1	D	147[A]	ASP	2.1
1	D	100	GLY	2.1
1	B	124	THR	2.0
1	D	112	TYR	2.0
1	D	120	ASN	2.0
1	C	43	ASP	2.0

## 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 [i](#)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	5GP	B	300	24/24	0.95	0.09	-0.17	9,12,13,14	0
2	5GP	A	300	24/24	0.94	0.10	-0.17	15,19,27,31	0
2	5GP	D	300	24/24	0.95	0.08	-0.70	13,21,26,27	0
2	5GP	C	300	24/24	0.96	0.07	-1.09	13,18,23,23	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.