



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:32 PM GMT

PDB ID : 4TO0
Title : Structure basis of cellular dNTP regulation, SAMHD1-GTP-dATP-dCTP complex
Authors : Ji, X.; Tang, C.; Zhao, Q.; Wang, W.; Xiong, Y.
Deposited on : 2014-06-05
Resolution : 2.30 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
<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

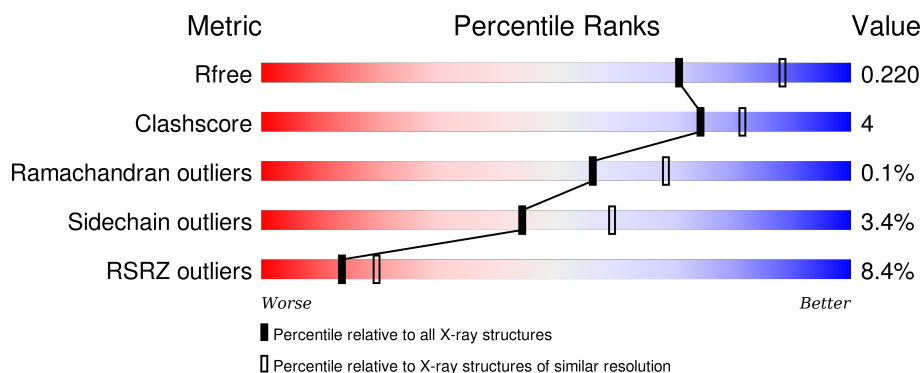
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 ≥ 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	514	<div> <div>8%</div> <div>85% 8% • 7%</div> </div>
1	B	514	<div> <div>5%</div> <div>84% 8% • 7%</div> </div>
1	C	514	<div> <div>7%</div> <div>82% 10% • 6%</div> </div>
1	D	514	<div> <div>11%</div> <div>83% 10% • 6%</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16246 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.

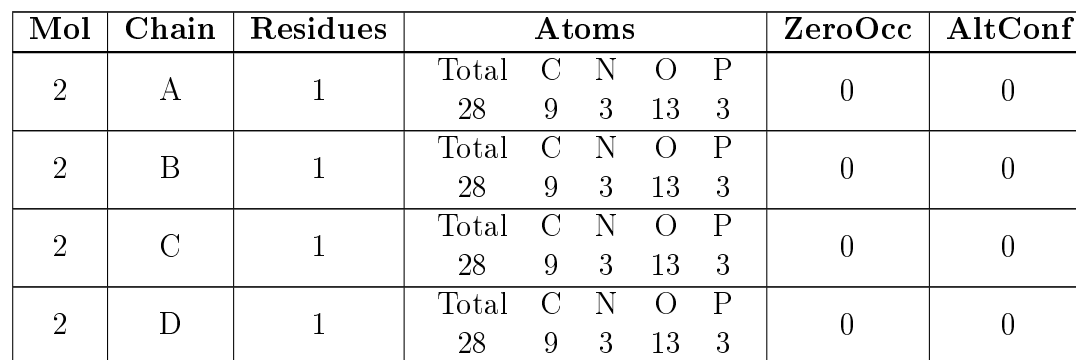
- Molecule 1 is a protein called Deoxynucleoside triphosphate triphosphohydrolase SAMHD1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	479	Total	C	N	O	S	0	1	0
			3922	2511	684	706	21			
1	B	480	Total	C	N	O	S	0	0	0
			3924	2512	684	708	20			
1	C	481	Total	C	N	O	S	0	2	0
			3944	2522	687	715	20			
1	D	484	Total	C	N	O	S	0	0	0
			3953	2530	688	715	20			

There are 8 discrepancies between the modelled and reference sequences:

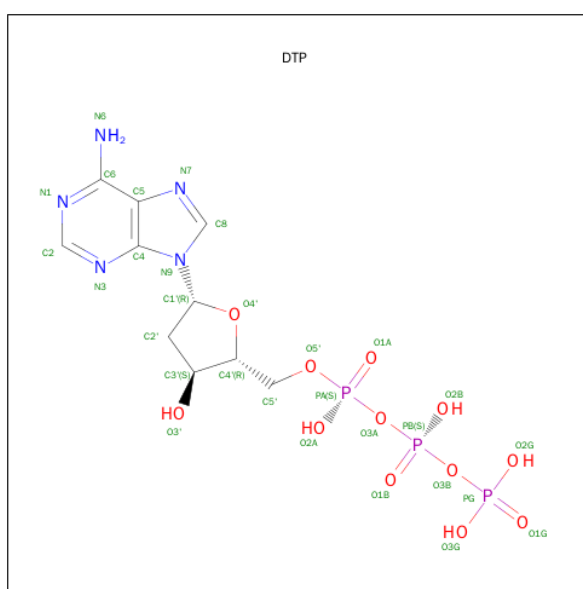
Chain	Residue	Modelled	Actual	Comment	Reference
A	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
A	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
B	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
B	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
C	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
C	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
D	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
D	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3

- Molecule 2 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula: C₉H₁₆N₃O₁₃P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	B	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	C	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	D	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

- Molecule 4 is 2'-DEOXYADENOSINE 5'-TRIPHOSPHATE (three-letter code: DTP) (formula: $C_{10}H_{16}N_5O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			30	10	5	12	3		
4	B	1	Total	C	N	O	P	0	0
			30	10	5	12	3		
4	C	1	Total	C	N	O	P	0	0
			30	10	5	12	3		
4	D	1	Total	C	N	O	P	0	0
			30	10	5	12	3		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mg	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total 1	Mg 1	0	0
5	C	1	Total 1	Mg 1	0	0

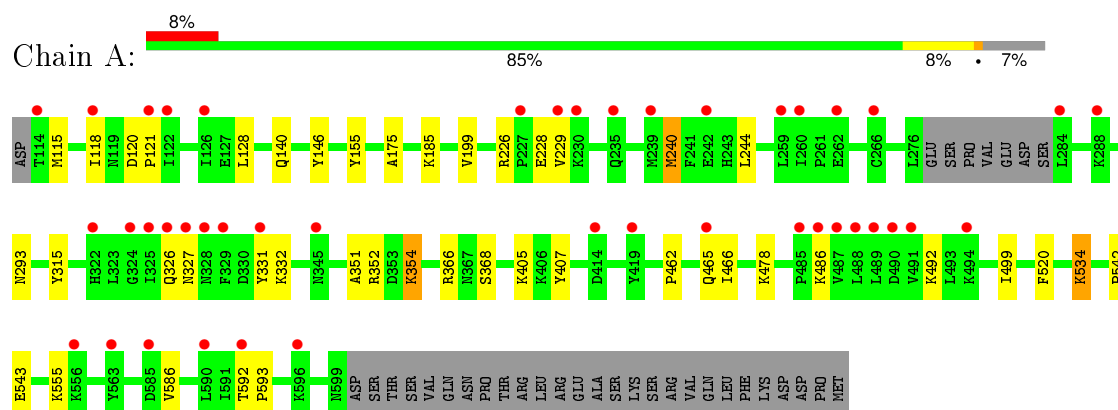
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	32	Total 32	O 32	0	0
6	B	42	Total 42	O 42	0	0
6	C	38	Total 38	O 38	0	0
6	D	27	Total 27	O 27	0	0

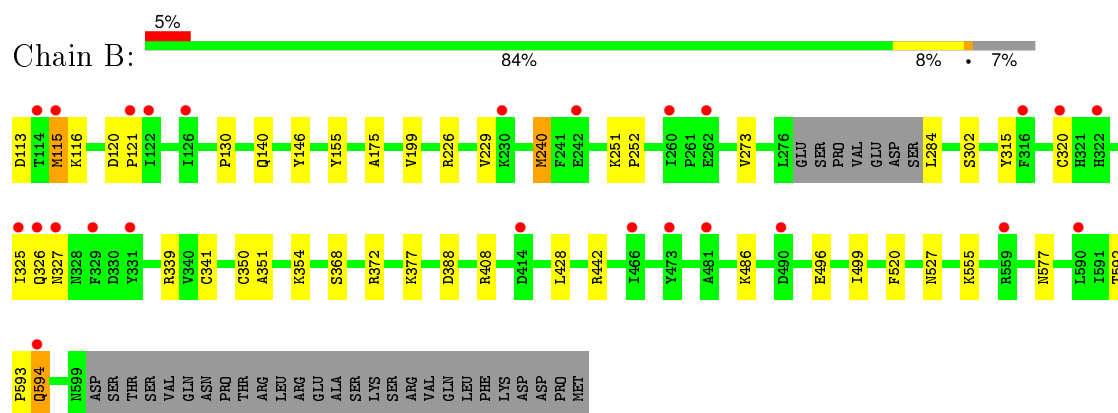
3 Residue-property plots [i](#)

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.

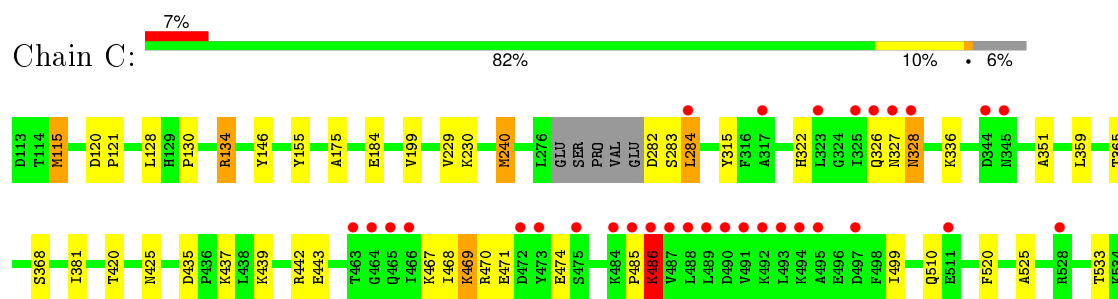
- Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1

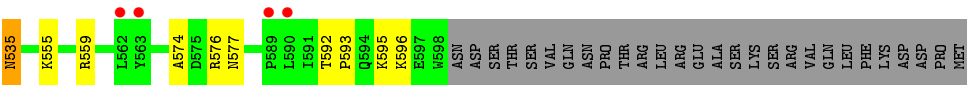


- Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1

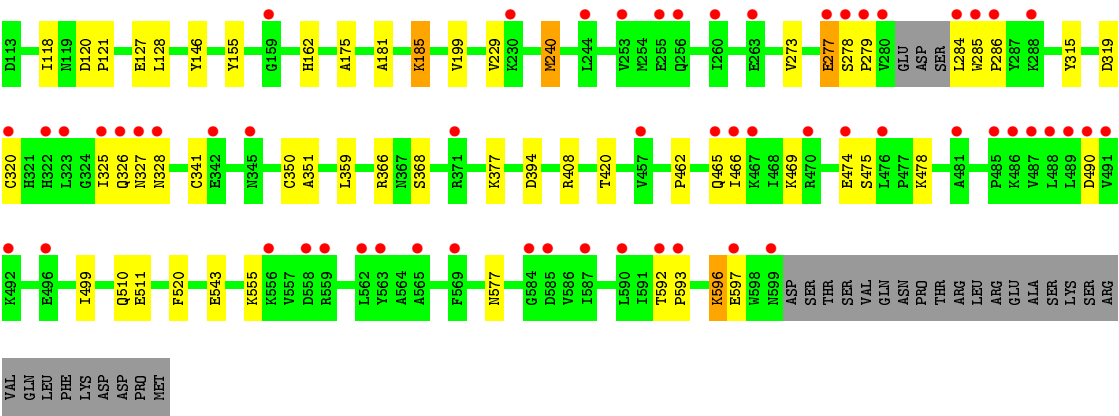
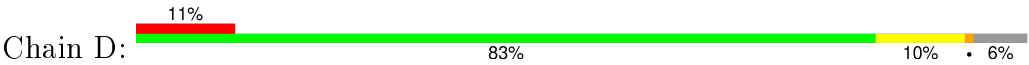


- Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1





● Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	81.42Å 141.33Å 97.52Å 90.00° 114.58° 90.00°	Depositor
Resolution (Å)	50.00 – 2.30 47.88 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.9 (50.00-2.30) 97.9 (47.88-2.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.36 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.191 , 0.222 0.194 , 0.220	Depositor DCC
R_{free} test set	4491 reflections (5.43%)	DCC
Wilson B-factor (Å ²)	47.2	Xtriage
Anisotropy	0.352	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 49.9	EDS
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 87260 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	16246	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% 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, 0.375 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: GTP, MG, DTP, DCP

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	0.64	0/4014	0.77	2/5418 (0.0%)
1	B	0.71	0/4016	0.78	2/5421 (0.0%)
1	C	0.71	0/4036	0.80	4/5448 (0.1%)
1	D	0.64	0/4046	0.75	1/5463 (0.0%)
All	All	0.68	0/16112	0.78	9/21750 (0.0%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	576	ARG	NE-CZ-NH2	6.62	123.61	120.30
1	C	359	LEU	CB-CG-CD1	5.93	121.09	111.00
1	B	442	ARG	NE-CZ-NH1	5.49	123.05	120.30
1	C	442	ARG	NE-CZ-NH1	5.48	123.04	120.30
1	A	366	ARG	NE-CZ-NH1	5.41	123.01	120.30

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	3922	0	3915	28	2
1	B	3924	0	3915	33	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3944	0	3927	49	0
1	D	3953	0	3942	29	2
2	A	28	0	12	1	0
2	B	28	0	12	0	0
2	C	28	0	12	0	0
2	D	28	0	12	1	0
3	A	32	0	12	0	0
3	B	32	0	12	0	0
3	C	32	0	12	1	0
3	D	32	0	12	0	0
4	A	30	0	12	0	0
4	B	30	0	12	0	0
4	C	30	0	12	0	0
4	D	30	0	12	0	0
5	A	2	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	32	0	0	1	0
6	B	42	0	0	1	0
6	C	38	0	0	4	0
6	D	27	0	0	2	0
All	All	16246	0	15843	129	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 4.

The worst 5 of 129 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:574:ALA:O	1:C:595:LYS:NZ	1.67	1.23
1:B:115:MET:HE2	1:B:116:LYS:N	1.73	1.03
1:C:328:ASN:ND2	1:C:365:THR:OG1	1.94	0.98
1:C:439:LYS:O	1:C:443:GLU:HG3	1.62	0.98
1:B:115:MET:C	1:B:115:MET:HE2	1.84	0.96

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)
1:A:465:GLN:O	1:D:469:LYS:CD[1_655]	1.90	0.30

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:543:GLU:O	1:D:465:GLN:OE1[1_655]	2.09	0.11

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	476/514 (93%)	471 (99%)	5 (1%)	0	100	100
1	B	476/514 (93%)	466 (98%)	10 (2%)	0	100	100
1	C	479/514 (93%)	470 (98%)	8 (2%)	1 (0%)	52	64
1	D	480/514 (93%)	472 (98%)	8 (2%)	0	100	100
All	All	1911/2056 (93%)	1879 (98%)	31 (2%)	1 (0%)	56	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	486	LYS

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	426/459 (93%)	416 (98%)	10 (2%)	58	75
1	B	426/459 (93%)	411 (96%)	15 (4%)	43	58
1	C	429/459 (94%)	407 (95%)	22 (5%)	29	39

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	430/459 (94%)	417 (97%)	13 (3%)	48 65
All	All	1711/1836 (93%)	1651 (96%)	60 (4%)	44 58

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	134	ARG
1	C	283	SER
1	D	490	ASP
1	C	230	LYS
1	C	315	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	326	GLN
1	D	162	HIS
1	D	535	ASN
1	C	322	HIS
1	D	539	GLN

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 ⓘ

Of 16 ligands modelled in this entry, 4 are monoatomic - leaving 12 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
2	DCP	A	701	-	21,29,29	0.62	0	33,45,45	1.47	5 (15%)
3	GTP	A	702	5	25,34,34	1.24	3 (12%)	34,54,54	1.95	9 (26%)
4	DTP	A	703	5	24,32,32	1.41	4 (16%)	32,50,50	1.81	5 (15%)
2	DCP	B	701	-	21,29,29	0.89	0	33,45,45	1.46	5 (15%)
4	DTP	B	702	5	24,32,32	1.49	5 (20%)	32,50,50	1.71	4 (12%)
3	GTP	B	703	5	25,34,34	1.32	2 (8%)	34,54,54	1.90	8 (23%)
4	DTP	C	701	5	24,32,32	1.35	3 (12%)	32,50,50	1.80	9 (28%)
2	DCP	C	703	-	21,29,29	0.83	0	33,45,45	1.35	4 (12%)
3	GTP	C	704	5	25,34,34	1.96	3 (12%)	34,54,54	1.80	8 (23%)
2	DCP	D	702	-	21,29,29	0.64	0	33,45,45	1.46	4 (12%)
4	DTP	D	703	5	24,32,32	1.24	3 (12%)	32,50,50	1.91	5 (15%)
3	GTP	D	704	5	25,34,34	1.25	3 (12%)	34,54,54	1.98	11 (32%)

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	DCP	A	701	-	-	0/18/34/34	0/2/2/2
3	GTP	A	702	5	-	0/18/38/38	0/3/3/3
4	DTP	A	703	5	-	0/18/34/34	0/3/3/3
2	DCP	B	701	-	-	0/18/34/34	0/2/2/2
4	DTP	B	702	5	-	0/18/34/34	0/3/3/3
3	GTP	B	703	5	-	0/18/38/38	0/3/3/3
4	DTP	C	701	5	-	0/18/34/34	0/3/3/3
2	DCP	C	703	-	-	0/18/34/34	0/2/2/2
3	GTP	C	704	5	-	0/18/38/38	0/3/3/3
2	DCP	D	702	-	-	0/18/34/34	0/2/2/2
4	DTP	D	703	5	-	0/18/34/34	0/3/3/3
3	GTP	D	704	5	-	0/18/38/38	0/3/3/3

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	703	DTP	PG-O2G	-2.78	1.44	1.54
4	B	702	DTP	O4'-C4'	-2.67	1.38	1.45
3	A	702	GTP	PG-O2G	-2.23	1.46	1.54
3	C	704	GTP	C4-N3	-2.18	1.32	1.35
4	B	702	DTP	PB-O2B	-2.10	1.46	1.54

The worst 5 of 77 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	703	DTP	N3-C2-N1	-7.81	122.92	128.89
4	B	702	DTP	N3-C2-N1	-6.91	123.60	128.89
4	A	703	DTP	N3-C2-N1	-5.85	124.42	128.89
3	D	704	GTP	C5-C6-N1	-5.05	116.68	123.59
4	C	701	DTP	N3-C2-N1	-4.97	125.09	128.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	DCP	1	0
3	C	704	GTP	1	0
2	D	702	DCP	1	0

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	479/514 (93%)	0.54	43 (8%) 12 17	33, 56, 90, 128	0
1	B	480/514 (93%)	0.34	25 (5%) 31 39	27, 48, 87, 108	0
1	C	481/514 (93%)	0.42	35 (7%) 18 25	29, 51, 95, 137	0
1	D	484/514 (94%)	0.73	58 (11%) 6 9	32, 58, 98, 168	0
All	All	1924/2056 (93%)	0.51	161 (8%) 14 19	27, 54, 91, 168	0

The worst 5 of 161 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	280	VAL	16.7
1	D	489	LEU	7.6
1	A	465	GLN	7.3
1	C	472	ASP	7.1
1	A	490	ASP	6.5

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 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, 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	LLDF	B-factors(\AA^2)	Q<0.9
4	DTP	C	701	30/30	0.99	0.16	-0.10	32,38,52,56	0
4	DTP	A	703	30/30	0.98	0.15	-0.37	26,30,52,54	0
4	DTP	D	703	30/30	0.98	0.16	-0.44	32,38,54,56	0
4	DTP	B	702	30/30	0.98	0.15	-0.45	28,33,50,56	0
3	GTP	B	703	32/32	0.98	0.13	-0.75	28,33,38,41	0
3	GTP	C	704	32/32	0.99	0.12	-0.83	22,30,37,39	0
3	GTP	D	704	32/32	0.98	0.12	-0.87	32,37,46,49	0
2	DCP	D	702	28/28	0.93	0.13	-1.00	45,59,81,83	0
3	GTP	A	702	32/32	0.98	0.11	-1.06	34,37,41,42	0
2	DCP	A	701	28/28	0.90	0.13	-1.20	48,61,84,86	0
2	DCP	B	701	28/28	0.94	0.12	-1.45	35,48,72,76	0
2	DCP	C	703	28/28	0.95	0.12	-1.48	38,45,74,76	0
5	MG	C	702	1/1	0.96	0.12	-	35,35,35,35	0
5	MG	A	705	1/1	0.91	0.11	-	44,44,44,44	0
5	MG	D	701	1/1	0.96	0.12	-	38,38,38,38	0
5	MG	A	704	1/1	0.96	0.12	-	32,32,32,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.