



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:33 PM GMT

PDB ID : 4TO5
Title : Structure basis of cellular dNTP regulation, SAMHD1-GTP-dTTP-dCTP complex
Authors : Ji, X.; Tang, C.; Zhao, Q.; Wang, W.; Xiong, Y.
Deposited on : 2014-06-05
Resolution : 2.80 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

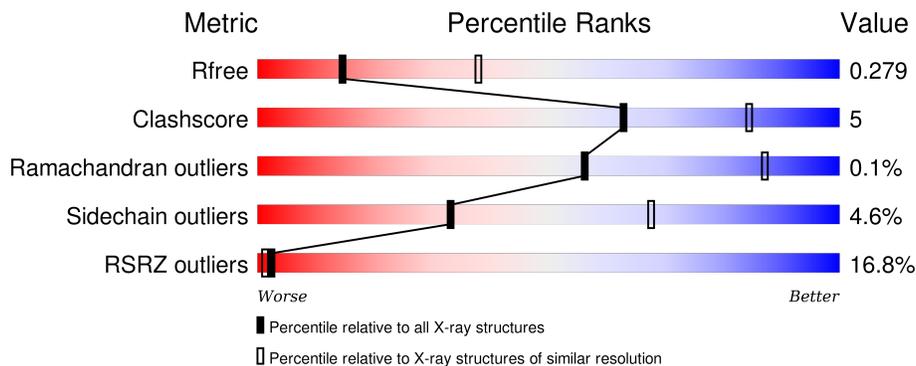
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	 17% 81% 10% • 7%
1	B	514	 19% 81% 11% • 7%
1	C	514	 11% 81% 12% • 7%
1	D	514	 16% 83% 9% • 7%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 16073 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.

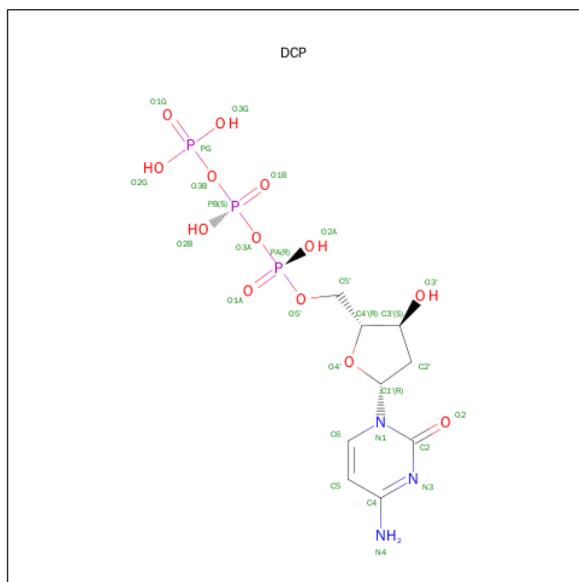
- Molecule 1 is a protein called Deoxynucleoside triphosphate triphosphohydrolase SAMHD1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	480	3932	2517	684	711	20	0	1	0
1	C	479	3922	2511	684	706	21	0	1	0
1	D	480	3933	2518	684	710	21	0	2	0
1	B	480	3922	2512	683	707	20	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

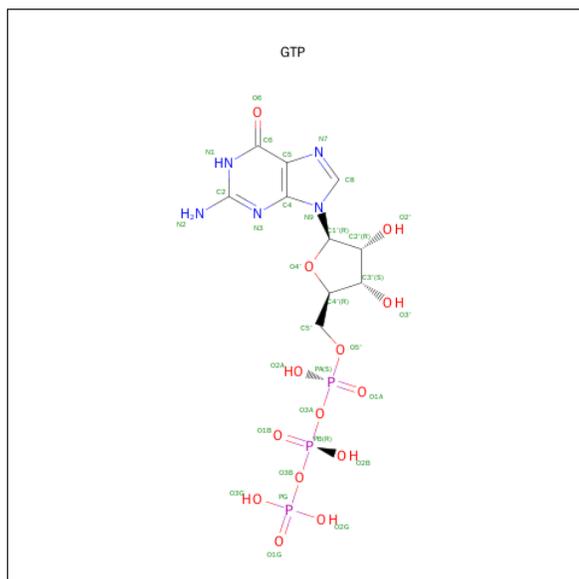
Chain	Residue	Modelled	Actual	Comment	Reference
A	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
A	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
C	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
C	207	ASN	ASP	engineered mutation	UNP Q9Y3Z3
D	206	ARG	HIS	engineered mutation	UNP Q9Y3Z3
D	207	ASN	ASP	engineered mutation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

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



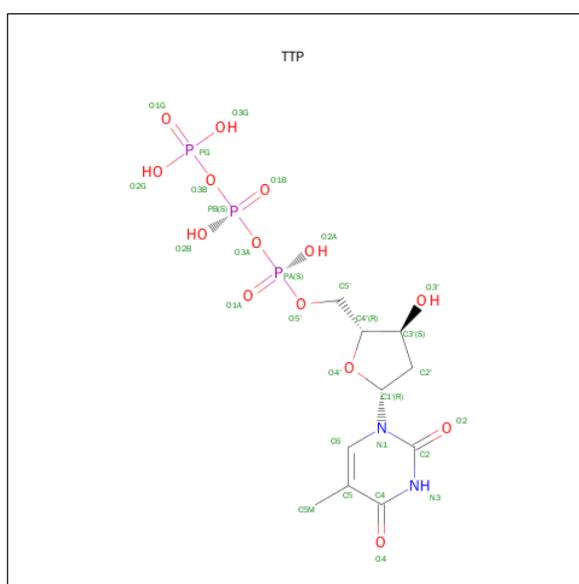
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	Total	C	N	O	P	0	0
			28	9	3	13	3		
2	C	1	Total	C	N	O	P	0	0
			28	9	3	13	3		
2	D	1	Total	C	N	O	P	0	0
			28	9	3	13	3		
2	B	1	Total	C	N	O	P	0	0
			28	9	3	13	3		

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	A	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		
3	B	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		

- Molecule 4 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: $C_{10}H_{17}N_2O_{14}P_3$).



Continued from previous page...

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

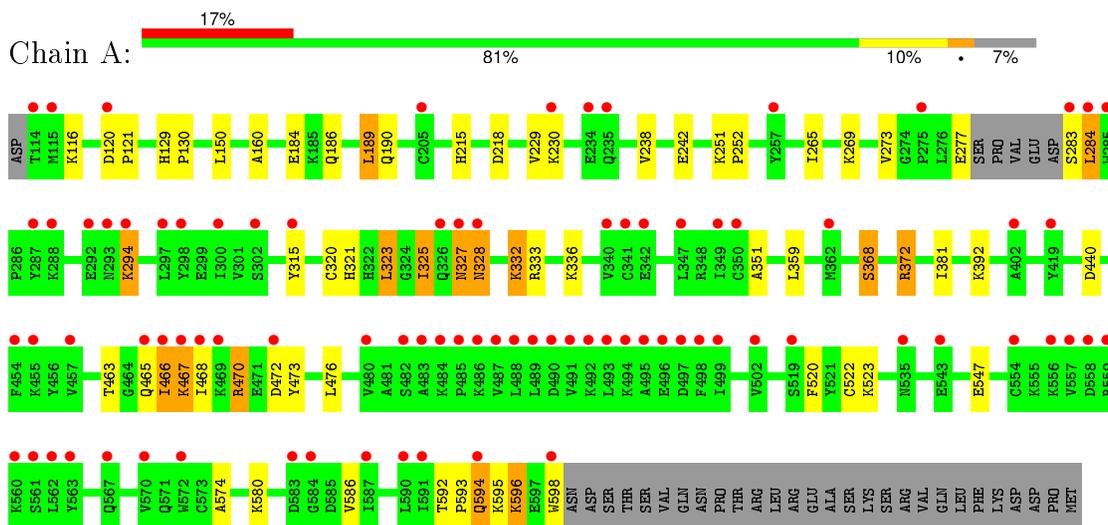
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O 1 1	0	0
6	C	1	Total O 1 1	0	0
6	D	1	Total O 1 1	0	0
6	B	1	Total O 1 1	0	0

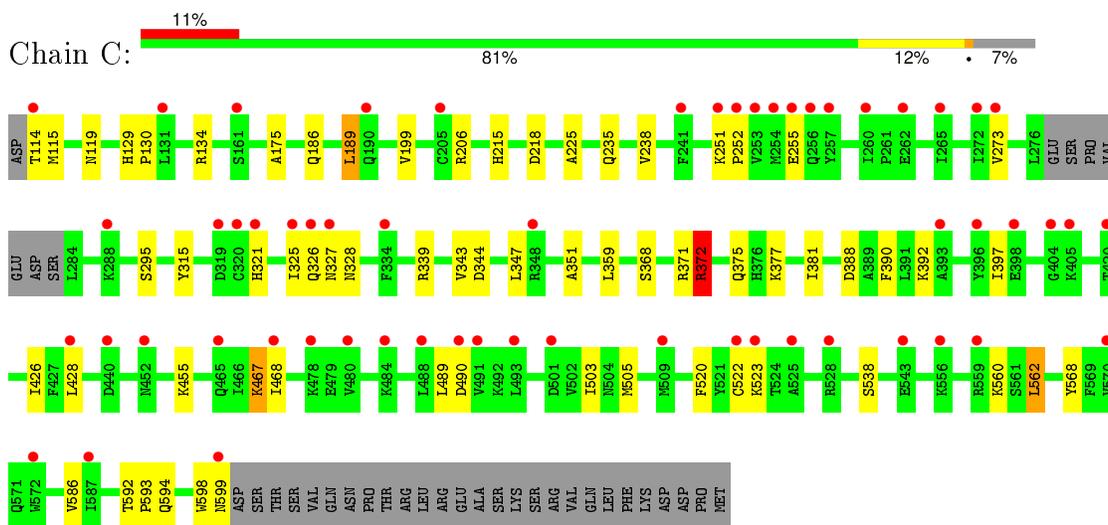
3 Residue-property plots [i](#)

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.

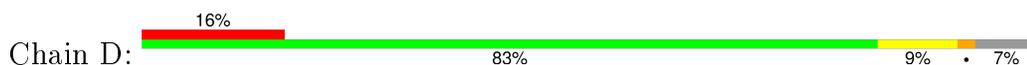
- Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1

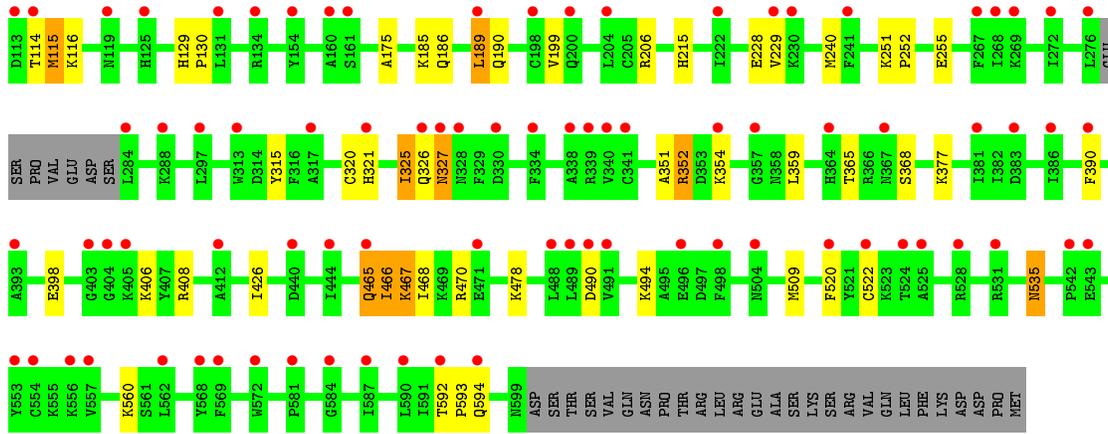


- Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1

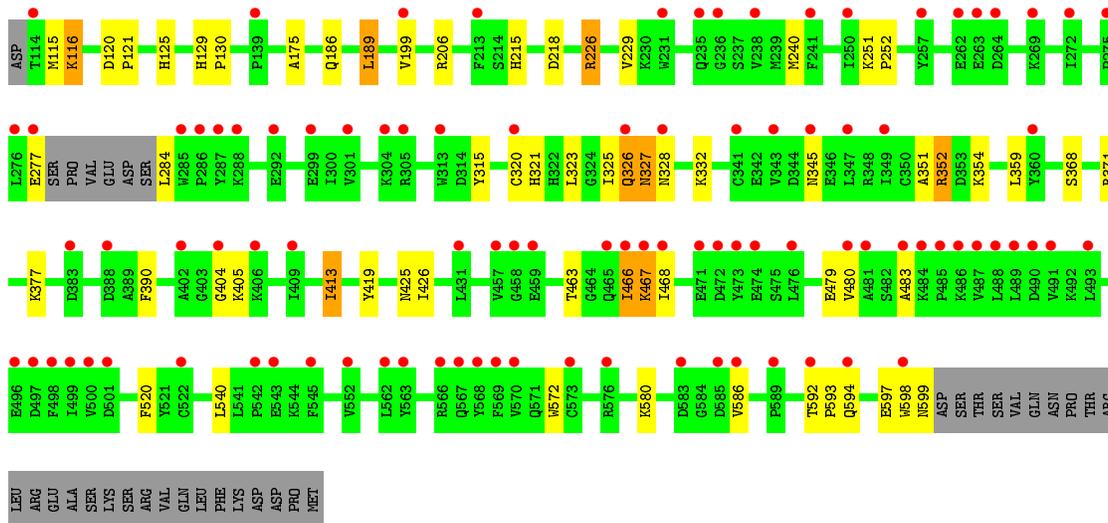
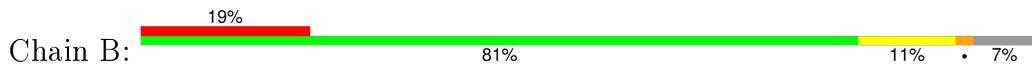


- Molecule 1: Deoxynucleoside triphosphate triphosphohydrolase SAMHD1





● 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	86.46Å 145.12Å 98.23Å 90.00° 114.13° 90.00°	Depositor
Resolution (Å)	50.00 – 2.80 48.39 – 2.73	Depositor EDS
% Data completeness (in resolution range)	96.6 (50.00-2.80) 96.6 (48.39-2.73)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.64 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, R_{free}	0.271 , 0.285 0.266 , 0.279	Depositor DCC
R_{free} test set	2761 reflections (5.38%)	DCC
Wilson B-factor (Å ²)	62.2	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 75.0	EDS
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 56472 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	16073	wwPDB-VP
Average B, all atoms (Å ²)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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, TTP, 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	0.56	0/4024	0.70	2/5431 (0.0%)
1	B	0.60	0/4014	0.75	7/5418 (0.1%)
1	C	0.56	0/4014	0.71	3/5418 (0.1%)
1	D	0.58	0/4025	0.72	3/5433 (0.1%)
All	All	0.57	0/16077	0.72	15/21700 (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
1	B	0	1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	294	LYS	CD-CE-NZ	7.83	129.70	111.70
1	B	371	ARG	NE-CZ-NH2	7.31	123.95	120.30
1	B	352	ARG	CG-CD-NE	-7.02	97.05	111.80
1	D	352	ARG	CG-CD-NE	-6.81	97.51	111.80
1	D	470	ARG	NE-CZ-NH1	6.66	123.63	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	404	GLY	Peptide

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	3932	0	3920	63	0
1	B	3922	0	3913	46	0
1	C	3922	0	3915	38	1
1	D	3933	0	3916	30	1
2	A	28	0	12	7	0
2	B	28	0	12	5	0
2	C	28	0	12	0	0
2	D	28	0	12	1	0
3	A	32	0	12	1	0
3	B	64	0	24	0	0
3	D	32	0	12	2	0
4	A	29	0	13	2	0
4	B	29	0	13	2	0
4	C	29	0	13	1	0
4	D	29	0	13	1	0
5	A	2	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	1	0	0	1	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
All	All	16073	0	15812	169	1

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 5.

The worst 5 of 169 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:D:390:PHE:CZ	1:D:426:ILE:HD11	1.78	1.18
1:A:283:SER:HB2	1:A:284:LEU:CA	1.84	1.07
1:A:283:SER:CB	1:A:284:LEU:HA	1.89	1.03
1:A:466:ILE:H	1:A:466:ILE:HD12	0.90	1.03
1:D:390:PHE:CE2	1:D:426:ILE:HD11	1.94	1.02

All (1) 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:C:225:ALA:O	1:D:494:LYS:NZ[1_656]	1.87	0.33

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	477/514 (93%)	469 (98%)	8 (2%)	0	100	100
1	B	476/514 (93%)	468 (98%)	8 (2%)	0	100	100
1	C	476/514 (93%)	468 (98%)	8 (2%)	0	100	100
1	D	478/514 (93%)	470 (98%)	7 (2%)	1 (0%)	52	84
All	All	1907/2056 (93%)	1875 (98%)	31 (2%)	1 (0%)	56	87

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	327	ASN

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	427/459 (93%)	406 (95%)	21 (5%)	31	65
1	B	425/459 (93%)	406 (96%)	19 (4%)	34	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	426/459 (93%)	410 (96%)	16 (4%)	40	74
1	D	426/459 (93%)	404 (95%)	22 (5%)	29	62
All	All	1704/1836 (93%)	1626 (95%)	78 (5%)	33	67

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	562	LEU
1	D	228	GLU
1	B	368	SER
1	C	594	GLN
1	D	185	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	425	ASN
1	D	190	GLN
1	B	233	HIS
1	A	425	ASN
1	D	425	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	0.63	0	33,45,45	1.47	3 (9%)
3	GTP	A	702	5	25,34,34	1.24	3 (12%)	34,54,54	1.91	9 (26%)
4	TTP	A	703	5	21,30,30	0.77	0	31,47,47	2.44	6 (19%)
3	GTP	B	701	5	25,34,34	1.36	3 (12%)	34,54,54	1.63	7 (20%)
4	TTP	B	702	5	21,30,30	0.71	0	31,47,47	2.31	5 (16%)
2	DCP	B	703	-	21,29,29	0.56	0	33,45,45	1.32	5 (15%)
3	GTP	B	704	5	25,34,34	1.16	2 (8%)	34,54,54	1.86	8 (23%)
4	TTP	C	701	5	21,30,30	0.62	0	31,47,47	2.47	7 (22%)
2	DCP	C	702	-	21,29,29	0.57	0	33,45,45	1.33	3 (9%)
2	DCP	D	702	-	21,29,29	0.57	0	33,45,45	1.36	4 (12%)
3	GTP	D	703	5	25,34,34	1.15	2 (8%)	34,54,54	1.87	10 (29%)
4	TTP	D	704	5	21,30,30	0.68	0	31,47,47	2.38	7 (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
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	TTP	A	703	5	-	0/18/34/34	0/2/2/2
3	GTP	B	701	5	-	0/18/38/38	0/3/3/3
4	TTP	B	702	5	-	0/18/34/34	0/2/2/2
2	DCP	B	703	-	-	0/18/34/34	0/2/2/2
3	GTP	B	704	5	-	0/18/38/38	0/3/3/3
4	TTP	C	701	5	-	0/18/34/34	0/2/2/2
2	DCP	C	702	-	-	0/18/34/34	0/2/2/2
2	DCP	D	702	-	-	0/18/34/34	0/2/2/2
3	GTP	D	703	5	-	0/18/38/38	0/3/3/3
4	TTP	D	704	5	-	0/18/34/34	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	703	GTP	O4'-C1'	2.08	1.43	1.41
3	A	702	GTP	O4'-C1'	2.16	1.43	1.41
3	B	701	GTP	C8-N7	2.32	1.39	1.34
3	B	701	GTP	C5-C4	2.53	1.46	1.40
3	D	703	GTP	C5-C4	2.81	1.46	1.40

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	703	TTP	C5-C4-N3	-7.56	116.72	125.14
4	C	701	TTP	C5-C4-N3	-7.40	116.90	125.14
4	B	702	TTP	C5-C4-N3	-6.95	117.39	125.14
4	D	704	TTP	C5-C4-N3	-6.93	117.42	125.14
3	B	704	GTP	PA-O3A-PB	-5.89	116.18	132.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	DCP	7	0
3	A	702	GTP	1	0
4	A	703	TTP	2	0
4	B	702	TTP	2	0
2	B	703	DCP	5	0
4	C	701	TTP	1	0
2	D	702	DCP	1	0
3	D	703	GTP	2	0
4	D	704	TTP	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	480/514 (93%)	1.15	85 (17%) 2 1	74, 96, 122, 142	0
1	B	480/514 (93%)	1.23	96 (20%) 1 1	74, 96, 126, 152	0
1	C	479/514 (93%)	0.99	58 (12%) 6 3	68, 92, 110, 121	0
1	D	480/514 (93%)	1.11	83 (17%) 2 1	86, 108, 134, 148	0
All	All	1919/2056 (93%)	1.12	322 (16%) 2 1	68, 98, 127, 152	0

The worst 5 of 322 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	489	LEU	8.2
1	D	522[A]	CYS	7.5
1	A	490	ASP	6.7
1	A	488	LEU	6.3
1	B	486	LYS	6.3

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(Å ²)	Q<0.9
2	DCP	D	702	28/28	0.75	0.30	0.84	93,106,131,133	0
2	DCP	A	701	28/28	0.71	0.27	0.18	84,95,129,131	0
2	DCP	C	702	28/28	0.76	0.27	0.16	112,120,135,137	0
2	DCP	B	703	28/28	0.71	0.25	0.03	90,106,123,126	0
4	TTP	A	703	29/29	0.92	0.25	-0.34	70,85,98,106	0
3	GTP	B	701	32/32	0.87	0.25	-0.39	48,78,103,108	0
3	GTP	D	703	32/32	0.85	0.22	-0.88	59,78,98,106	0
3	GTP	A	702	32/32	0.89	0.19	-1.03	70,84,97,103	0
4	TTP	D	704	29/29	0.91	0.20	-1.03	79,89,100,102	0
4	TTP	B	702	29/29	0.91	0.23	-1.12	56,83,101,106	0
4	TTP	C	701	29/29	0.91	0.18	-1.15	66,81,100,103	0
3	GTP	B	704	32/32	0.91	0.18	-1.36	80,87,111,114	0
5	MG	C	703	1/1	0.81	0.10	-	101,101,101,101	0
5	MG	A	704	1/1	0.56	0.10	-	148,148,148,148	0
5	MG	D	701	1/1	0.77	0.16	-	85,85,85,85	0
5	MG	A	705	1/1	0.31	0.18	-	93,93,93,93	0

6.5 Other polymers

There are no such residues in this entry.