



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 09:59 pm GMT

PDB ID : 5I2R  
Title : human PDE10A in complex with 3-(2-phenylpyrazol-3-yl)-1-[3-(trifluoromethoxy)phenyl]pyridazin-4-one  
Authors : Joseph, C.; Koerner, M.; Rudolph, M.G.  
Deposited on : 2016-02-09  
Resolution : 2.50 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

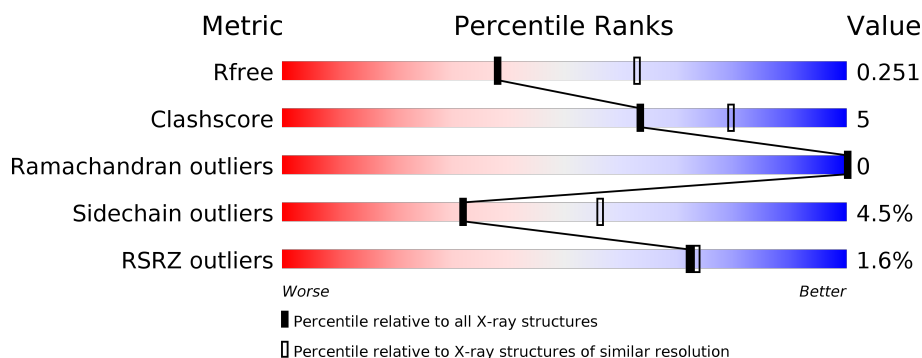
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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	318	<div> <div>86%</div> <div>11%</div> <div>..</div> </div>
1	B	318	<div> <div>85%</div> <div>14%</div> <div>.</div> </div>
1	C	318	<div> <div>86%</div> <div>12%</div> <div>..</div> </div>
1	D	318	<div> <div>3%</div> <div>86%</div> <div>11%</div> <div>..</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10855 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 cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	313	Total	C	N	O	S	0	1	0
			2547	1628	432	463	24			
1	B	317	Total	C	N	O	S	0	3	0
			2581	1648	440	469	24			
1	C	313	Total	C	N	O	S	0	1	0
			2549	1629	435	461	24			
1	D	313	Total	C	N	O	S	0	0	0
			2537	1622	433	458	24			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	456	ALA	-	expression tag	UNP Q9Y233
B	456	ALA	-	expression tag	UNP Q9Y233
C	456	ALA	-	expression tag	UNP Q9Y233
D	456	ALA	-	expression tag	UNP Q9Y233

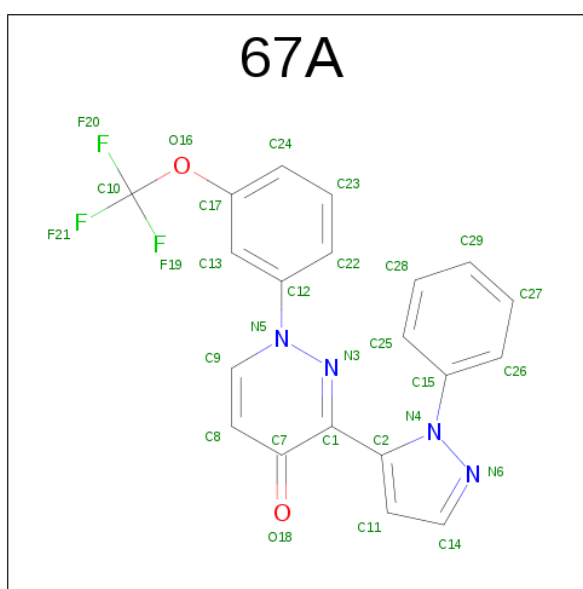
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mg 1 1	0	0
3	A	1	Total Mg 1 1	0	0
3	D	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0

- Molecule 4 is 3-(1-phenyl-1H-pyrazol-5-yl)-1-[3-(trifluoromethoxy)phenyl]pyridazin-4(1H)-one (three-letter code: 67A) (formula: C<sub>20</sub>H<sub>13</sub>F<sub>3</sub>N<sub>4</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C F N O 29 20 3 4 2	0	0
4	B	1	Total C F N O 29 20 3 4 2	0	0
4	C	1	Total C F N O 29 20 3 4 2	0	0
4	D	1	Total C F N O 29 20 3 4 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	131	Total O 131 131	0	0

*Continued on next page...*

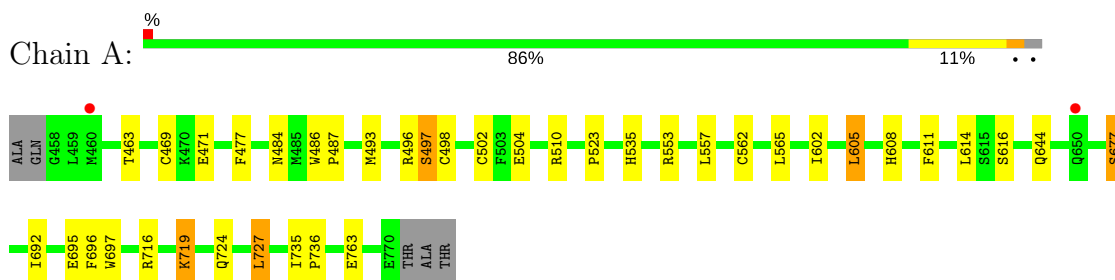
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	149	Total 149	O 149	0	0
5	C	152	Total 152	O 152	0	0
5	D	85	Total 85	O 85	0	0

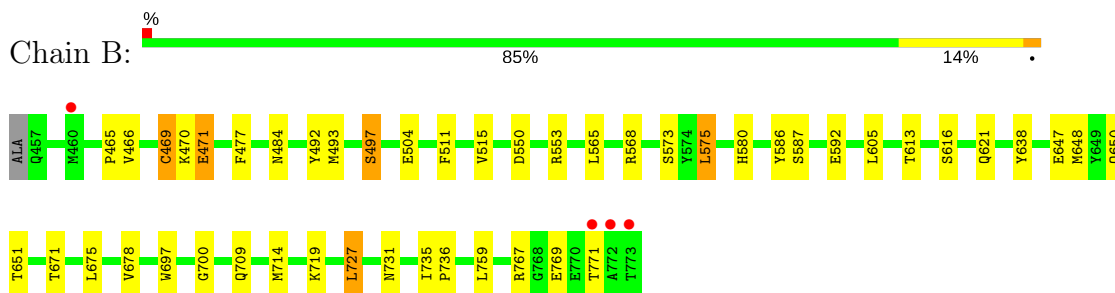
### 3 Residue-property plots [i](#)

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

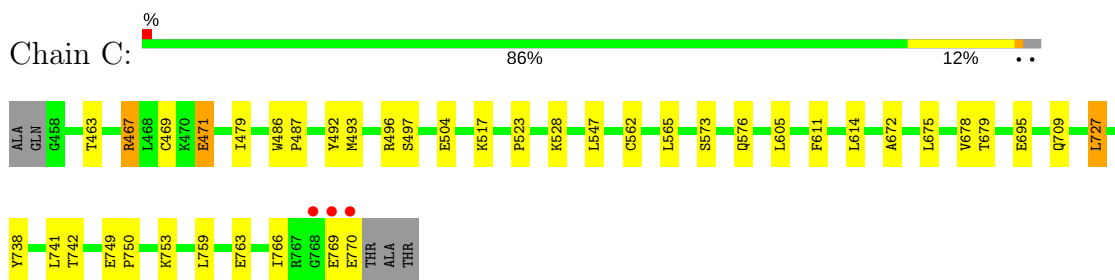
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A



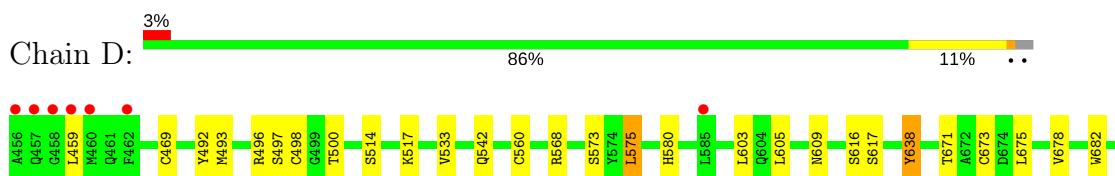
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A



- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A



- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A





## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.81Å 135.81Å 234.86Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.44 – 2.50 35.44 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.3 (35.44-2.50) 98.3 (35.44-2.50)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.32 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.5.0035	Depositor
R, $R_{free}$	0.191 , 0.250 0.196 , 0.251	Depositor DCC
$R_{free}$ test set	2711 reflections (5.19%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.0	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 39.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.025 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10855	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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, 0.333 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: MG, ZN, CME, 67A

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	0.58	0/2601	0.64	0/3519
1	B	0.59	0/2641	0.64	0/3574
1	C	0.57	0/2603	0.65	1/3521 (0.0%)
1	D	0.55	0/2588	0.61	0/3502
All	All	0.57	0/10433	0.63	1/14116 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	467	ARG	NE-CZ-NH1	5.64	123.12	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	2547	0	2519	25	0
1	B	2581	0	2556	27	0
1	C	2549	0	2524	21	0
1	D	2537	0	2514	27	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	29	0	0	0	0
4	B	29	0	0	0	0
4	C	29	0	0	0	0
4	D	29	0	0	0	0
5	A	131	0	0	4	0
5	B	149	0	0	7	0
5	C	152	0	0	2	0
5	D	85	0	0	2	0
All	All	10855	0	10113	98	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 5.

All (98) 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:484[B]:ASN:OD1	5:B:901:HOH:O	1.72	1.07
1:C:469:CYS:SG	5:C:1027:HOH:O	2.37	0.73
1:B:492:TYR:OH	5:B:902:HOH:O	2.07	0.72
1:B:493:MET:O	1:B:497:SER:HB2	1.89	0.72
1:A:727:LEU:HD21	1:A:763:GLU:HG3	1.73	0.71
1:D:493:MET:O	1:D:497:SER:HB2	1.93	0.67
1:C:493:MET:O	1:C:497:SER:HB2	1.98	0.63
1:B:511:PHE:O	1:B:515:VAL:HG23	1.99	0.62
1:D:727:LEU:CD1	1:D:766:ILE:CD1	2.78	0.61
1:B:727:LEU:HD23	1:B:759:LEU:CD1	2.30	0.61
1:A:498:CYS:CB	1:A:502:CYS:HG	2.12	0.61
1:D:727:LEU:HD13	1:D:766:ILE:CD1	2.30	0.61
1:A:493:MET:O	1:A:497:SER:HB2	2.01	0.61
1:B:697:TRP:CZ2	1:B:719:LYS:HG2	2.36	0.60
1:D:727:LEU:CD1	1:D:766:ILE:HD12	2.31	0.59
1:D:730:TYR:HA	1:D:734:ALA:HB3	1.85	0.59
1:A:602:ILE:HA	1:A:605:LEU:HD22	1.86	0.58
1:A:469:CYS:SG	5:A:1011:HOH:O	2.57	0.57
1:B:771:THR:HB	5:B:1020:HOH:O	2.05	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:675:LEU:O	1:C:678:VAL:HG22	2.05	0.56
1:A:498:CYS:HB3	1:A:502:CYS:HG	1.69	0.56
1:D:575:LEU:HD13	1:D:580:HIS:CG	2.41	0.56
1:D:469:CYS:SG	5:D:978:HOH:O	2.58	0.55
1:D:682:TRP:CZ2	1:D:686:LYS:HD3	2.42	0.55
1:C:562:CYS:HB3	1:C:565:LEU:HD12	1.90	0.54
1:B:767:ARG:NH1	1:B:769:GLU:OE2	2.41	0.53
1:B:648:MET:HE1	5:B:1001:HOH:O	2.09	0.53
1:B:727:LEU:CD2	1:B:759:LEU:HD11	2.39	0.52
1:A:611:PHE:HB3	1:A:614:LEU:HD22	1.92	0.52
1:D:675:LEU:O	1:D:678:VAL:HG22	2.09	0.52
1:D:766:ILE:O	1:D:766:ILE:HG22	2.10	0.52
1:C:749:GLU:HB3	1:C:750:PRO:HD3	1.92	0.52
1:A:716:ARG:O	1:A:719:LYS:NZ	2.40	0.52
1:B:727:LEU:HD23	1:B:759:LEU:HD11	1.92	0.51
1:A:497:SER:O	1:A:553:ARG:HD2	2.11	0.51
1:C:727:LEU:HD21	1:C:763:GLU:HG3	1.91	0.51
1:D:492:TYR:CZ	1:D:496:ARG:HD2	2.45	0.51
1:A:724:GLN:HE22	1:A:727:LEU:HD12	1.75	0.51
1:D:533:VAL:HG13	1:D:673:CYS:HB3	1.91	0.51
1:A:498:CYS:HG	1:A:502:CYS:CB	2.22	0.51
1:B:497:SER:O	1:B:553:ARG:HD2	2.10	0.51
1:C:492:TYR:CZ	1:C:496:ARG:HD2	2.45	0.50
1:D:638:TYR:CD1	1:D:671:THR:HG21	2.47	0.50
1:D:497:SER:HA	1:D:542:GLN:NE2	2.27	0.49
1:B:647:GLU:O	1:B:651:THR:HG23	2.13	0.49
1:A:486:TRP:N	1:A:487:PRO:CD	2.76	0.49
1:C:611:PHE:HB3	1:C:614:LEU:HD22	1.95	0.48
1:D:697:TRP:CZ2	1:D:719:LYS:HG3	2.49	0.47
1:B:470:LYS:HE2	1:D:746:PRO:HG3	1.96	0.47
1:B:700:GLY:HA3	1:B:714:MET:O	2.14	0.47
1:C:679:THR:HG21	1:C:738:TYR:CZ	2.48	0.47
1:D:727:LEU:HD13	1:D:766:ILE:HD12	1.96	0.47
1:D:497:SER:HA	1:D:542:GLN:HE22	1.80	0.47
1:D:493:MET:O	1:D:497:SER:CB	2.64	0.46
1:B:675:LEU:O	1:B:678:VAL:HG22	2.16	0.45
1:D:542:GLN:NE2	5:D:906:HOH:O	2.50	0.45
1:C:467:ARG:O	1:C:471:GLU:HB2	2.17	0.45
1:A:469:CYS:SG	5:A:1013:HOH:O	2.49	0.45
1:A:677:SER:HB2	5:A:991:HOH:O	2.16	0.45
1:D:533:VAL:HG12	1:D:560:CYS:HB3	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:465:PRO:O	1:B:466:VAL:C	2.55	0.44
1:B:592:GLU:OE2	5:B:903:HOH:O	2.21	0.44
1:A:697:TRP:HB3	1:A:716:ARG:HB3	1.99	0.44
1:B:735:ILE:HB	1:B:736:PRO:HD3	1.99	0.44
1:B:515:VAL:HG13	1:B:565:LEU:HD21	1.99	0.44
1:D:727:LEU:CD1	1:D:766:ILE:HD11	2.47	0.44
1:C:759:LEU:HD12	1:C:759:LEU:O	2.19	0.43
1:C:479:ILE:HD11	1:C:528:LYS:HG3	2.01	0.43
1:B:469:CYS:HB3	5:B:1028:HOH:O	2.17	0.43
1:D:603:LEU:HD22	1:D:609:ASN:HD22	1.82	0.43
1:A:557:LEU:HD23	1:A:557:LEU:HA	1.92	0.43
1:B:471:GLU:HB3	1:B:477:PHE:CD1	2.53	0.43
1:D:762:TRP:O	1:D:766:ILE:HG13	2.18	0.43
1:A:562:CYS:HB3	1:A:565:LEU:HD12	2.01	0.42
1:C:486:TRP:N	1:C:487:PRO:CD	2.82	0.42
1:D:727:LEU:HD12	1:D:766:ILE:CD1	2.47	0.42
1:A:477:PHE:HB3	1:A:535:HIS:CE1	2.55	0.42
1:C:727:LEU:HG	1:C:766:ILE:HD12	2.01	0.42
1:C:749:GLU:N	1:C:750:PRO:CD	2.82	0.42
1:A:735:ILE:HB	1:A:736:PRO:HD3	2.02	0.42
1:B:575:LEU:HD13	1:B:580:HIS:CG	2.54	0.42
1:D:575:LEU:HD13	1:D:580:HIS:CB	2.50	0.42
1:B:586:TYR:HD2	5:B:909:HOH:O	2.03	0.41
1:C:727:LEU:HG	1:C:766:ILE:CD1	2.50	0.41
1:A:523:PRO:HD2	1:A:695:GLU:HG2	2.01	0.41
1:B:638:TYR:CD1	1:B:671:THR:HG21	2.56	0.41
1:C:672:ALA:CB	1:C:741:LEU:HD22	2.51	0.41
1:B:727:LEU:HD22	1:B:731:ASN:ND2	2.35	0.41
1:A:724:GLN:NE2	1:A:727:LEU:HD12	2.36	0.41
1:A:463:THR:HG22	5:A:1009:HOH:O	2.19	0.41
1:A:510:ARG:HG2	1:A:608:HIS:CE1	2.56	0.40
1:C:523:PRO:HD2	1:C:695:GLU:HG2	2.02	0.40
1:A:484:ASN:HD21	1:B:568:ARG:HH12	1.69	0.40
1:C:769:GLU:O	1:C:770:GLU:HB2	2.21	0.40
1:D:727:LEU:HD12	1:D:766:ILE:HD11	2.03	0.40
1:A:692:ILE:HG22	1:A:696:PHE:CE2	2.57	0.40
1:C:727:LEU:HD22	1:C:759:LEU:HD11	2.02	0.40
1:C:742:THR:HG22	5:C:918:HOH:O	2.22	0.40

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	311/318 (98%)	299 (96%)	12 (4%)	0	100	100
1	B	317/318 (100%)	305 (96%)	12 (4%)	0	100	100
1	C	311/318 (98%)	301 (97%)	10 (3%)	0	100	100
1	D	310/318 (98%)	303 (98%)	7 (2%)	0	100	100
All	All	1249/1272 (98%)	1208 (97%)	41 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	282/284 (99%)	272 (96%)	10 (4%)	41	68
1	B	286/284 (101%)	271 (95%)	15 (5%)	27	49
1	C	282/284 (99%)	271 (96%)	11 (4%)	37	63
1	D	280/284 (99%)	265 (95%)	15 (5%)	26	47
All	All	1130/1136 (100%)	1079 (96%)	51 (4%)	32	56

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	471	GLU
1	A	496	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	497	SER
1	A	504	GLU
1	A	605	LEU
1	A	616	SER
1	A	644	GLN
1	A	677	SER
1	A	719	LYS
1	A	727	LEU
1	B	469	CYS
1	B	471	GLU
1	B	497	SER
1	B	504	GLU
1	B	550	ASP
1	B	573	SER
1	B	575	LEU
1	B	587	SER
1	B	605	LEU
1	B	613	THR
1	B	616	SER
1	B	621	GLN
1	B	650	GLN
1	B	709	GLN
1	B	727	LEU
1	C	463	THR
1	C	471	GLU
1	C	504	GLU
1	C	517	LYS
1	C	547	LEU
1	C	573	SER
1	C	576	GLN
1	C	605	LEU
1	C	709	GLN
1	C	727	LEU
1	C	753	LYS
1	D	459	LEU
1	D	498	CYS
1	D	500	THR
1	D	514	SER
1	D	517	LYS
1	D	568	ARG
1	D	573	SER
1	D	575	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	605	LEU
1	D	616	SER
1	D	617	SER
1	D	638	TYR
1	D	709	GLN
1	D	727	LEU
1	D	760	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
1	A	604	GLN
1	A	644	GLN
1	A	724	GLN
1	A	743	GLN
1	A	761	GLN
1	B	576	GLN
1	B	604	GLN
1	B	644	GLN
1	B	731	ASN
1	C	476	HIS
1	C	542	GLN
1	C	604	GLN
1	C	743	GLN
1	D	484	ASN
1	D	542	GLN
1	D	576	GLN
1	D	604	GLN
1	D	609	ASN
1	D	726	GLN
1	D	761	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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
1	CME	A	509	1	9,9,10	0.63	0	6,9,11	1.76	1 (16%)
1	CME	B	509	1	9,9,10	0.58	0	6,9,11	1.72	1 (16%)
1	CME	C	509	1	9,9,10	0.57	0	6,9,11	2.82	2 (33%)
1	CME	D	509	1	9,9,10	0.59	0	6,9,11	2.27	1 (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
1	CME	A	509	1	-	0/5/8/10	0/0/0/0
1	CME	B	509	1	-	0/5/8/10	0/0/0/0
1	CME	C	509	1	-	0/5/8/10	0/0/0/0
1	CME	D	509	1	-	0/5/8/10	0/0/0/0

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	509	CME	O-C-CA	-2.02	119.43	125.02
1	B	509	CME	CB-SG-SD	3.74	111.11	103.83
1	A	509	CME	CB-SG-SD	3.82	111.26	103.83
1	D	509	CME	CB-SG-SD	5.20	113.95	103.83
1	C	509	CME	CB-SG-SD	6.41	116.30	103.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.



## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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$
4	67A	A	803	-	28,32,32	1.34	2 (7%)	34,46,46	1.60	5 (14%)
4	67A	B	803	-	28,32,32	1.19	3 (10%)	34,46,46	1.88	4 (11%)
4	67A	C	803	-	28,32,32	1.08	1 (3%)	34,46,46	1.94	3 (8%)
4	67A	D	803	-	28,32,32	1.27	4 (14%)	34,46,46	1.46	3 (8%)

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
4	67A	A	803	-	-	0/14/17/17	0/4/4/4
4	67A	B	803	-	-	0/14/17/17	0/4/4/4
4	67A	C	803	-	-	0/14/17/17	0/4/4/4
4	67A	D	803	-	-	0/14/17/17	0/4/4/4

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	803	67A	C1-N3	-3.31	1.30	1.33
4	D	803	67A	C1-N3	-2.27	1.31	1.33
4	B	803	67A	C1-N3	-2.21	1.31	1.33
4	D	803	67A	C12-N5	-2.12	1.39	1.44
4	B	803	67A	C15-N4	-2.09	1.39	1.44
4	D	803	67A	O16-C10	2.06	1.45	1.31

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	803	67A	C8-C7	2.19	1.42	1.37
4	C	803	67A	C8-C7	2.32	1.42	1.37
4	D	803	67A	C8-C7	2.41	1.42	1.37
4	B	803	67A	C8-C7	2.50	1.42	1.37

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	803	67A	C8-C7-C1	-4.39	114.30	120.15
4	B	803	67A	C7-C1-N3	-3.77	118.24	122.59
4	D	803	67A	C7-C1-N3	-3.22	118.87	122.59
4	A	803	67A	C7-C1-N3	-3.18	118.92	122.59
4	B	803	67A	C8-C7-C1	-2.45	116.88	120.15
4	D	803	67A	C8-C7-C1	-2.33	117.05	120.15
4	A	803	67A	C23-C22-C12	2.36	121.66	118.66
4	A	803	67A	C7-C1-C2	2.51	124.82	121.81
4	A	803	67A	C14-C11-C2	2.54	106.85	104.80
4	C	803	67A	C14-C11-C2	2.71	106.98	104.80
4	B	803	67A	C14-C11-C2	3.79	107.85	104.80
4	A	803	67A	C14-N6-N4	6.06	108.02	103.70
4	D	803	67A	C14-N6-N4	6.21	108.13	103.70
4	B	803	67A	C14-N6-N4	7.78	109.25	103.70
4	C	803	67A	C14-N6-N4	8.58	109.82	103.70

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

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, 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	312/318 (98%)	-0.25	2 (0%) 89 89	30, 43, 61, 77	0
1	B	316/318 (99%)	-0.32	4 (1%) 77 78	28, 41, 62, 80	0
1	C	312/318 (98%)	-0.25	3 (0%) 82 83	29, 41, 58, 75	0
1	D	312/318 (98%)	-0.02	11 (3%) 44 47	39, 53, 68, 80	0
All	All	1252/1272 (98%)	-0.21	20 (1%) 72 73	28, 45, 65, 80	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	773	THR	4.7
1	B	460	MET	3.9
1	C	770	GLU	3.5
1	D	765	VAL	3.4
1	D	456	ALA	3.3
1	B	772	ALA	3.3
1	D	720	ASP	3.1
1	B	771	THR	3.1
1	C	768	GLY	2.9
1	A	460	MET	2.8
1	D	458	GLY	2.7
1	D	459	LEU	2.6
1	D	457	GLN	2.5
1	D	585	LEU	2.4
1	D	462	PHE	2.3
1	A	650	GLN	2.3
1	D	767	ARG	2.2
1	C	769	GLU	2.2
1	D	460	MET	2.2
1	D	709	GLN	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	CME	A	509	10/11	0.90	0.18	-	50,52,70,72	0
1	CME	D	509	10/11	0.82	0.22	-	58,60,74,74	0
1	CME	B	509	10/11	0.84	0.19	-	50,52,70,71	0
1	CME	C	509	10/11	0.91	0.22	-	50,52,70,71	0

## 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(Å <sup>2</sup> )	Q<0.9
4	67A	A	803	29/29	0.93	0.15	0.78	40,42,55,59	0
4	67A	B	803	29/29	0.94	0.13	-0.10	34,38,48,50	0
4	67A	C	803	29/29	0.94	0.13	-0.44	33,37,52,53	0
4	67A	D	803	29/29	0.92	0.14	-0.50	47,50,60,61	0
3	MG	C	802	1/1	1.00	0.12	-0.86	18,18,18,18	0
2	ZN	A	801	1/1	1.00	0.10	-1.13	40,40,40,40	0
2	ZN	D	801	1/1	1.00	0.10	-1.47	48,48,48,48	0
3	MG	A	802	1/1	0.99	0.10	-1.75	22,22,22,22	0
3	MG	D	802	1/1	0.99	0.08	-1.87	30,30,30,30	0
2	ZN	B	801	1/1	1.00	0.10	-2.00	34,34,34,34	0
3	MG	B	802	1/1	0.99	0.10	-2.37	20,20,20,20	0
2	ZN	C	801	1/1	1.00	0.12	-2.39	36,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.