



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 21, 2017 – 12:18 PM EDT

PDB ID : 5ISB  
Title : Crystal structure of mouse CARM1 in complex with inhibitor SA0435  
Authors : Cura, V.; Marechal, N.; Mailliot, J.; Troffer-Charlier, N.; Hassenboehler, P.;  
Wurtz, J.M.; Bonnefond, L.; Cavarelli, J.  
Deposited on : 2016-03-15  
Resolution : 2.00 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 : rb-20029077  
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) : rb-20029077

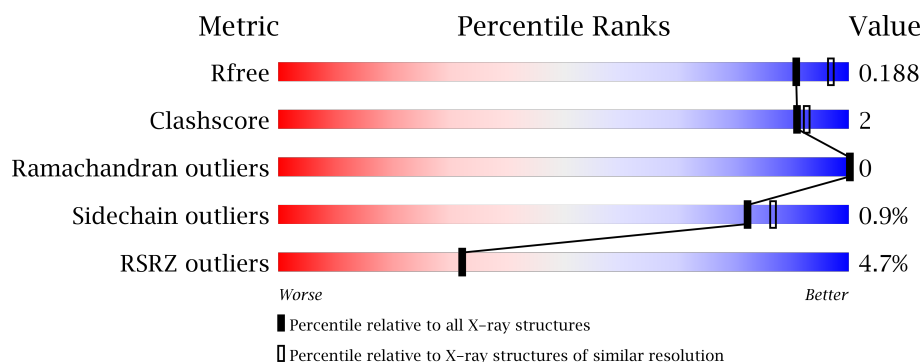
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	361	<div> <div>3%</div> <div>91%</div> <div>5%</div> </div>
1	B	361	<div> <div>5%</div> <div>89%</div> <div>6%</div> <div>5%</div> </div>
1	C	361	<div> <div>4%</div> <div>93%</div> <div>5%</div> </div>
1	D	361	<div> <div>5%</div> <div>90%</div> <div>5%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	EDO	B	504	-	-	-	X
4	PEG	A	505	-	-	-	X
4	PEG	B	507	-	-	-	X
4	PEG	C	506	-	-	-	X
4	PEG	D	506	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 23368 atoms, of which 11059 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.

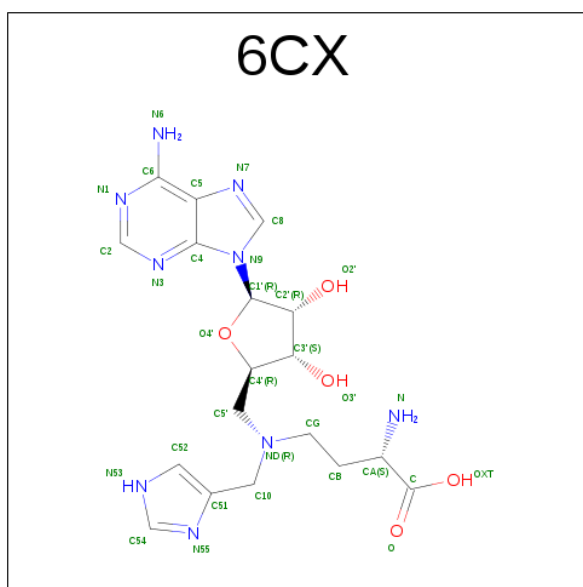
- Molecule 1 is a protein called Histone-arginine methyltransferase CARM1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	343	Total	C	H	N	O	S	0	4	0
			5515	1796	2730	460	513	16			
1	B	343	Total	C	H	N	O	S	0	4	0
			5502	1792	2723	458	511	18			
1	C	343	Total	C	H	N	O	S	0	3	0
			5498	1793	2722	457	511	15			
1	D	342	Total	C	H	N	O	S	0	2	0
			5464	1782	2704	455	507	16			

There are 12 discrepancies between the modelled and reference sequences:

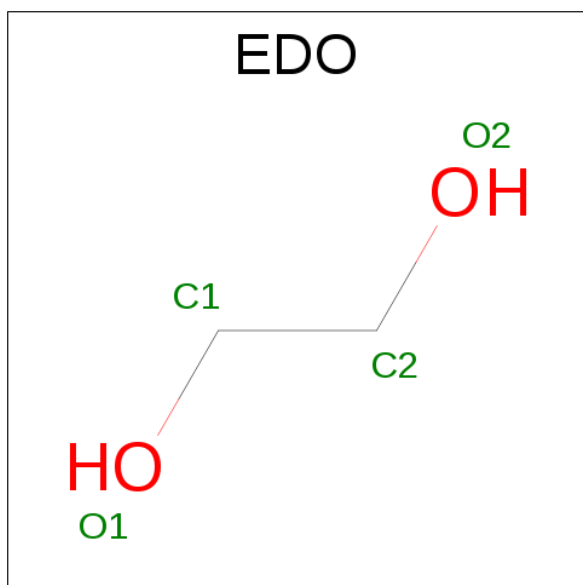
Chain	Residue	Modelled	Actual	Comment	Reference
A	127	GLY	-	expression tag	UNP Q9WVG6
A	128	HIS	-	expression tag	UNP Q9WVG6
A	129	MET	-	expression tag	UNP Q9WVG6
B	127	GLY	-	expression tag	UNP Q9WVG6
B	128	HIS	-	expression tag	UNP Q9WVG6
B	129	MET	-	expression tag	UNP Q9WVG6
C	127	GLY	-	expression tag	UNP Q9WVG6
C	128	HIS	-	expression tag	UNP Q9WVG6
C	129	MET	-	expression tag	UNP Q9WVG6
D	127	GLY	-	expression tag	UNP Q9WVG6
D	128	HIS	-	expression tag	UNP Q9WVG6
D	129	MET	-	expression tag	UNP Q9WVG6

- Molecule 2 is 5'-{[(3S)-3-amino-3-carboxypropyl][(1H-imidazol-4-yl)methyl]amino}-5'-deoxy adenosine (three-letter code: 6CX) (formula: C<sub>18</sub>H<sub>25</sub>N<sub>9</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	0
			56	18	24	9	5		
2	B	1	Total	C	H	N	O	0	0
			56	18	24	9	5		
2	C	1	Total	C	H	N	O	0	0
			56	18	24	9	5		
2	D	1	Total	C	H	N	O	0	0
			56	18	24	9	5		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



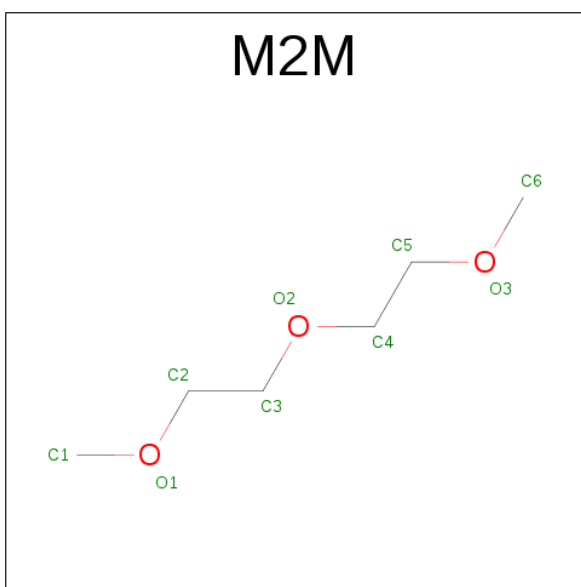
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0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

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



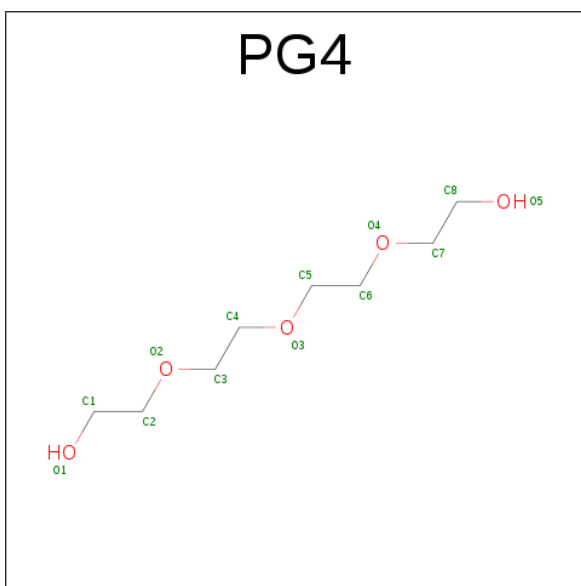
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			17	4	10	3		
4	A	1	Total	C	H	O	0	0
			17	4	10	3		
4	A	1	Total	C	H	O	0	0
			17	4	10	3		
4	B	1	Total	C	H	O	0	0
			17	4	10	3		
4	B	1	Total	C	H	O	0	0
			17	4	10	3		
4	C	1	Total	C	H	O	0	0
			17	4	10	3		
4	D	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 5 is 1-METHOXY-2-(2-METHOXYETHOXY)ETHANE (three-letter code: M2M) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			23	6	14	3		

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			13	8	5		

- Molecule 7 is water.

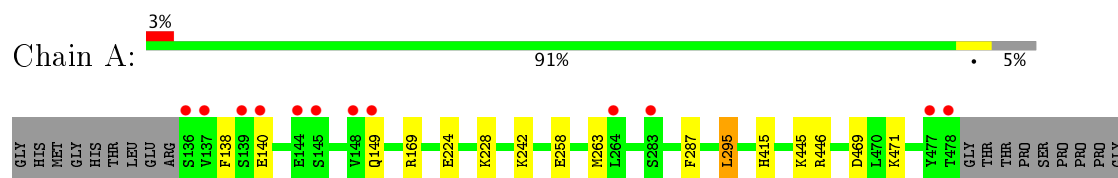


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	263	Total 264	O 264	0	1
7	B	211	Total 212	O 212	0	1
7	C	254	Total 254	O 254	0	0
7	D	215	Total 216	O 216	0	1

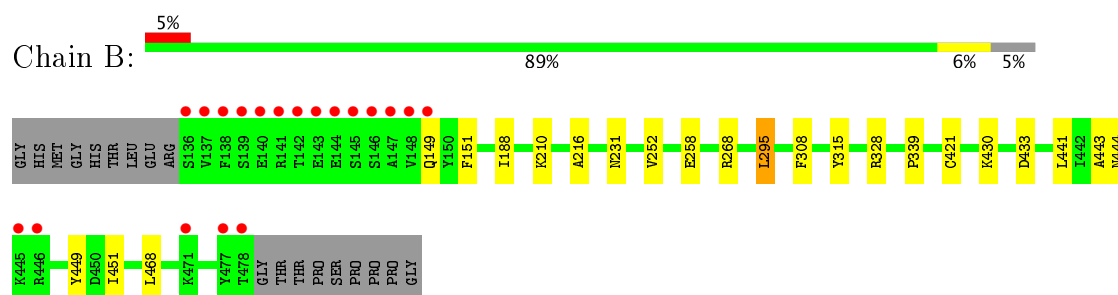
### 3 Residue-property plots [i](#)

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.

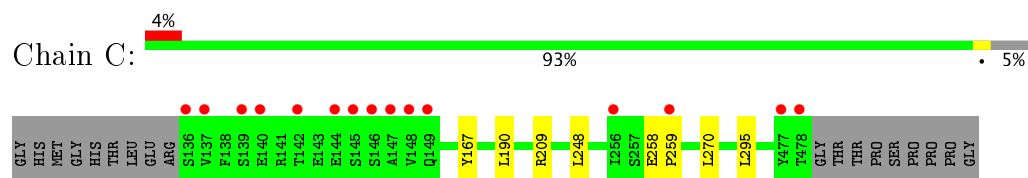
- Molecule 1: Histone-arginine methyltransferase CARM1



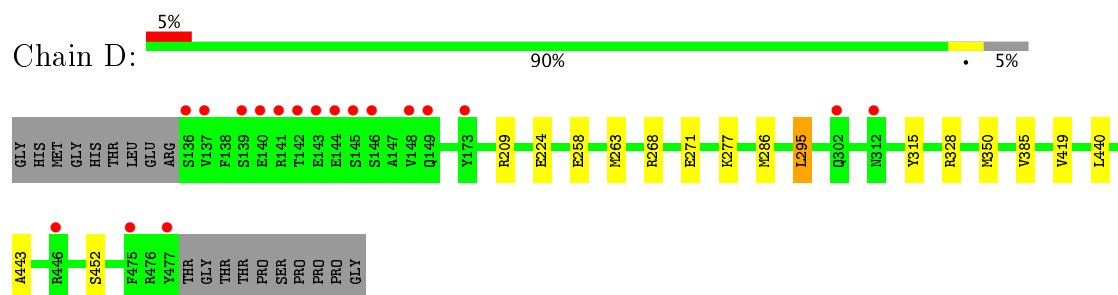
- Molecule 1: Histone-arginine methyltransferase CARM1



- Molecule 1: Histone-arginine methyltransferase CARM1



- Molecule 1: Histone-arginine methyltransferase CARM1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.43Å 98.08Å 205.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 29.65 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (30.00-2.00) 99.4 (29.65-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.74 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1839)	Depositor
R, $R_{free}$	0.154 , 0.190 0.149 , 0.188	Depositor DCC
$R_{free}$ test set	5079 reflections (4.98%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.4	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	23368	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.44 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.1726e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: 6CX, M2M, PG4, PEG, EDO

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	0.62	0/2856	0.66	1/3869 (0.0%)
1	B	0.57	0/2849	0.64	0/3859
1	C	0.56	0/2847	0.64	0/3857
1	D	0.56	0/2830	0.64	1/3833 (0.0%)
All	All	0.58	0/11382	0.64	2/15418 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	295	LEU	CA-CB-CG	5.65	128.30	115.30
1	A	263	MET	CA-CB-CG	5.42	122.52	113.30

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	2785	2730	2724	11	0
1	B	2779	2723	2723	16	0
1	C	2776	2722	2719	4	0
1	D	2760	2704	2708	12	0
2	A	32	24	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	32	24	0	0	0
2	C	32	24	0	0	0
2	D	32	24	0	0	0
3	A	12	0	18	0	0
3	B	20	0	30	0	0
3	C	16	0	24	1	0
3	D	16	0	24	0	0
4	A	21	30	30	2	0
4	B	14	20	20	0	0
4	C	7	10	10	0	0
4	D	7	10	10	0	0
5	A	9	14	14	0	0
6	C	13	0	18	1	0
7	A	264	0	0	1	0
7	B	212	0	0	2	0
7	C	254	0	0	3	0
7	D	216	0	0	3	0
All	All	12309	11059	11072	44	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 2.

All (44) 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:433:ASP:OD2	7:B:601:HOH:O	2.13	0.66
1:D:268:ARG:HD2	1:D:271:GLU:OE2	1.98	0.64
1:A:140:GLU:O	1:A:445:LYS:NZ	2.32	0.63
1:D:263:MET:CE	1:D:419:VAL:HG11	2.31	0.60
1:D:209:ARG:NH1	7:D:603:HOH:O	2.30	0.59
1:B:421[A]:CYS:SG	1:B:468:LEU:HD22	2.42	0.59
1:B:315:TYR:HD1	1:B:328:ARG:HD2	1.70	0.55
6:C:507:PG4:H72	7:C:747:HOH:O	2.06	0.55
1:A:149:GLN:NE2	1:B:149:GLN:OE1	2.40	0.54
1:D:277:LYS:HD3	1:D:286[A]:MET:SD	2.48	0.54
1:B:339:PRO:HG3	1:B:468:LEU:HG	1.90	0.53
1:D:328:ARG:NH1	7:D:607:HOH:O	2.44	0.50
1:B:315:TYR:CD1	1:B:328:ARG:HD2	2.46	0.50
1:B:451:ILE:HB	1:B:468:LEU:HB3	1.93	0.50
1:D:263:MET:HE3	1:D:419:VAL:HG11	1.94	0.49
1:D:350[A]:MET:SD	1:D:385:VAL:HG22	2.53	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:446:ARG:HD2	1:A:471:LYS:NZ	2.28	0.48
1:B:295:LEU:HD12	1:B:295:LEU:N	2.29	0.47
1:A:446:ARG:HD2	1:A:471:LYS:HZ3	1.80	0.47
1:B:421[A]:CYS:SG	1:B:468:LEU:CD2	3.02	0.47
1:A:415[B]:HIS:CE1	4:A:507:PEG:H31	2.51	0.46
1:C:209:ARG:NH2	7:C:614:HOH:O	2.48	0.46
1:D:315:TYR:CD1	1:D:328:ARG:HD2	2.52	0.45
1:A:224:GLU:OE1	1:A:228:LYS:HE3	2.17	0.44
1:A:469:ASP:OD1	1:A:471:LYS:HG2	2.18	0.44
1:D:263:MET:HE2	1:D:419:VAL:HG11	1.99	0.44
1:D:224:GLU:OE1	7:D:601:HOH:O	2.21	0.44
1:C:259:PRO:HG2	1:C:270:LEU:HD23	2.00	0.43
1:D:268:ARG:HD3	1:D:443:ALA:HB1	2.00	0.43
1:A:295:LEU:HD23	1:A:295:LEU:N	2.34	0.43
4:A:507:PEG:H41	7:A:784:HOH:O	2.18	0.43
1:A:138:PHE:CD2	1:A:242:LYS:HD3	2.54	0.42
1:B:188:ILE:HG22	1:B:252:VAL:HG12	2.01	0.42
1:B:430:LYS:HD3	1:C:167:TYR:CE2	2.53	0.42
1:C:190:LEU:HD13	1:C:248:LEU:HD21	2.01	0.42
1:A:169:ARG:HD3	1:A:415[A]:HIS:CE1	2.55	0.42
1:B:268:ARG:CZ	1:B:443:ALA:HB1	2.50	0.42
1:A:469:ASP:OD2	1:A:471:LYS:HE2	2.20	0.41
1:B:231:ASN:HB3	7:B:767:HOH:O	2.20	0.41
3:C:503:EDO:H22	7:C:823:HOH:O	2.18	0.41
1:B:441:LEU:HB3	1:B:449:TYR:CD2	2.56	0.41
1:D:440:LEU:HB3	1:D:452:SER:HB3	2.01	0.41
1:B:151:PHE:CE1	1:B:216:ALA:HB3	2.55	0.41
1:B:268:ARG:NH2	1:B:444:ASN:O	2.54	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	345/361 (96%)	335 (97%)	10 (3%)	0	100	100
1	B	345/361 (96%)	335 (97%)	10 (3%)	0	100	100
1	C	344/361 (95%)	334 (97%)	10 (3%)	0	100	100
1	D	342/361 (95%)	331 (97%)	11 (3%)	0	100	100
All	All	1376/1444 (95%)	1335 (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	303/313 (97%)	300 (99%)	3 (1%)	80	84
1	B	303/313 (97%)	299 (99%)	4 (1%)	73	78
1	C	302/313 (96%)	300 (99%)	2 (1%)	87	90
1	D	300/313 (96%)	298 (99%)	2 (1%)	87	90
All	All	1208/1252 (96%)	1197 (99%)	11 (1%)	82	87

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	258	GLU
1	A	287	PHE
1	A	295	LEU
1	B	210	LYS
1	B	258	GLU
1	B	295	LEU
1	B	308	PHE
1	C	258	GLU
1	C	295	LEU
1	D	258	GLU
1	D	295	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	149	GLN
1	B	149	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 ⓘ

29 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	6CX	A	501	-	24,35,35	1.23	3 (12%)	21,50,50	2.38	5 (23%)
3	EDO	A	502	-	3,3,3	0.47	0	2,2,2	0.47	0
3	EDO	A	503	-	3,3,3	0.48	0	2,2,2	0.44	0
3	EDO	A	504	-	3,3,3	0.47	0	2,2,2	0.52	0
4	PEG	A	505	-	6,6,6	0.58	0	5,5,5	0.62	0
4	PEG	A	506	-	6,6,6	0.43	0	5,5,5	0.64	0
4	PEG	A	507	-	6,6,6	0.49	0	5,5,5	0.51	0
5	M2M	A	508	-	8,8,8	0.66	0	7,7,7	0.61	0
2	6CX	B	501	-	24,35,35	1.35	2 (8%)	21,50,50	2.58	5 (23%)
3	EDO	B	502	-	3,3,3	0.47	0	2,2,2	0.36	0
3	EDO	B	503	-	3,3,3	0.44	0	2,2,2	0.54	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	EDO	B	504	-	3,3,3	0.44	0	2,2,2	0.36	0
3	EDO	B	505	-	3,3,3	0.45	0	2,2,2	0.44	0
3	EDO	B	506	-	3,3,3	0.40	0	2,2,2	0.67	0
4	PEG	B	507	-	6,6,6	0.48	0	5,5,5	0.43	0
4	PEG	B	508	-	6,6,6	0.47	0	5,5,5	0.52	0
2	6CX	C	501	-	24,35,35	1.26	3 (12%)	21,50,50	2.36	5 (23%)
3	EDO	C	502	-	3,3,3	0.45	0	2,2,2	0.66	0
3	EDO	C	503	-	3,3,3	0.40	0	2,2,2	0.48	0
3	EDO	C	504	-	3,3,3	0.46	0	2,2,2	0.65	0
3	EDO	C	505	-	3,3,3	0.46	0	2,2,2	0.28	0
4	PEG	C	506	-	6,6,6	0.49	0	5,5,5	0.44	0
6	PG4	C	507	-	12,12,12	0.54	0	11,11,11	0.59	0
2	6CX	D	501	-	24,35,35	1.40	4 (16%)	21,50,50	2.45	5 (23%)
3	EDO	D	502	-	3,3,3	0.46	0	2,2,2	0.56	0
3	EDO	D	503	-	3,3,3	0.47	0	2,2,2	0.31	0
3	EDO	D	504	-	3,3,3	0.48	0	2,2,2	0.45	0
3	EDO	D	505	-	3,3,3	0.48	0	2,2,2	0.47	0
4	PEG	D	506	-	6,6,6	0.38	0	5,5,5	0.54	0

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	6CX	A	501	-	-	0/13/37/37	0/4/4/4
3	EDO	A	502	-	-	0/1/1/1	0/0/0/0
3	EDO	A	503	-	-	0/1/1/1	0/0/0/0
3	EDO	A	504	-	-	0/1/1/1	0/0/0/0
4	PEG	A	505	-	-	0/4/4/4	0/0/0/0
4	PEG	A	506	-	-	0/4/4/4	0/0/0/0
4	PEG	A	507	-	-	0/4/4/4	0/0/0/0
5	M2M	A	508	-	-	0/6/6/6	0/0/0/0
2	6CX	B	501	-	-	0/13/37/37	0/4/4/4
3	EDO	B	502	-	-	0/1/1/1	0/0/0/0
3	EDO	B	503	-	-	0/1/1/1	0/0/0/0
3	EDO	B	504	-	-	0/1/1/1	0/0/0/0
3	EDO	B	505	-	-	0/1/1/1	0/0/0/0
3	EDO	B	506	-	-	0/1/1/1	0/0/0/0
4	PEG	B	507	-	-	0/4/4/4	0/0/0/0
4	PEG	B	508	-	-	0/4/4/4	0/0/0/0
2	6CX	C	501	-	-	0/13/37/37	0/4/4/4

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	C	502	-	-	0/1/1/1	0/0/0/0
3	EDO	C	503	-	-	0/1/1/1	0/0/0/0
3	EDO	C	504	-	-	0/1/1/1	0/0/0/0
3	EDO	C	505	-	-	0/1/1/1	0/0/0/0
4	PEG	C	506	-	-	0/4/4/4	0/0/0/0
6	PG4	C	507	-	-	0/10/10/10	0/0/0/0
2	6CX	D	501	-	-	0/13/37/37	0/4/4/4
3	EDO	D	502	-	-	0/1/1/1	0/0/0/0
3	EDO	D	503	-	-	0/1/1/1	0/0/0/0
3	EDO	D	504	-	-	0/1/1/1	0/0/0/0
3	EDO	D	505	-	-	0/1/1/1	0/0/0/0
4	PEG	D	506	-	-	0/4/4/4	0/0/0/0

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	6CX	C2'-C1'	-4.45	1.46	1.53
2	C	501	6CX	C2'-C1'	-3.63	1.47	1.53
2	D	501	6CX	C2'-C1'	-3.63	1.47	1.53
2	A	501	6CX	C3'-C2'	-2.98	1.45	1.53
2	A	501	6CX	O4'-C4'	-2.72	1.38	1.45
2	D	501	6CX	O4'-C4'	-2.68	1.38	1.45
2	A	501	6CX	C2'-C1'	-2.59	1.49	1.53
2	D	501	6CX	C3'-C2'	-2.52	1.46	1.53
2	C	501	6CX	C3'-C2'	-2.46	1.46	1.53
2	B	501	6CX	C3'-C2'	-2.45	1.46	1.53
2	C	501	6CX	C2-N3	2.12	1.35	1.32
2	D	501	6CX	O4'-C1'	2.43	1.44	1.41

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	6CX	N3-C2-N1	-8.95	121.06	128.86
2	A	501	6CX	N3-C2-N1	-8.81	121.18	128.86
2	D	501	6CX	N3-C2-N1	-8.54	121.42	128.86
2	C	501	6CX	N3-C2-N1	-8.36	121.57	128.86
2	B	501	6CX	C1'-N9-C4	-4.20	119.38	126.64
2	B	501	6CX	C10-ND-C5'	-3.56	106.99	111.02
2	D	501	6CX	C1'-N9-C4	-3.36	120.82	126.64
2	C	501	6CX	C10-C51-C52	-3.35	123.23	129.01
2	A	501	6CX	C10-C51-C52	-3.15	123.58	129.01
2	C	501	6CX	C10-ND-C5'	-2.97	107.66	111.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	6CX	C1'-N9-C4	-2.62	122.11	126.64
2	B	501	6CX	C10-C51-C52	-2.34	124.98	129.01
2	D	501	6CX	C10-C51-C52	-2.26	125.12	129.01
2	C	501	6CX	C1'-N9-C4	-2.16	122.91	126.64
2	A	501	6CX	C10-ND-CG	-2.10	106.81	111.29
2	D	501	6CX	O3'-C3'-C4'	-2.00	105.24	111.09
2	C	501	6CX	C51-C10-ND	2.44	117.76	112.88
2	A	501	6CX	C51-C10-ND	2.73	118.34	112.88
2	B	501	6CX	C51-C10-ND	3.26	119.40	112.88
2	D	501	6CX	C51-C10-ND	3.79	120.45	112.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	507	PEG	2	0
3	C	503	EDO	1	0
6	C	507	PG4	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 ⓘ

### 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	343/361 (95%)	-0.16	12 (3%) 44 45	16, 26, 45, 79	0
1	B	343/361 (95%)	-0.16	19 (5%) 26 26	16, 30, 54, 103	0
1	C	343/361 (95%)	-0.10	15 (4%) 35 35	19, 30, 52, 94	0
1	D	342/361 (94%)	-0.16	18 (5%) 27 27	20, 32, 53, 92	0
All	All	1371/1444 (94%)	-0.15	64 (4%) 32 32	16, 29, 52, 103	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	478	THR	6.2
1	D	477	TYR	6.2
1	A	478	THR	6.1
1	D	136	SER	5.9
1	B	478	THR	5.7
1	C	477	TYR	5.6
1	B	137	VAL	5.5
1	C	137	VAL	5.3
1	C	148	VAL	5.3
1	D	139	SER	5.2
1	D	148	VAL	5.0
1	B	136	SER	4.9
1	D	137	VAL	4.8
1	B	142	THR	4.6
1	B	477	TYR	4.5
1	B	148	VAL	4.4
1	D	140	GLU	4.4
1	B	145	SER	4.3
1	D	142	THR	4.2
1	A	477	TYR	4.2
1	B	144	GLU	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	139	SER	4.0
1	A	144	GLU	4.0
1	C	136	SER	3.9
1	C	140	GLU	3.9
1	D	144	GLU	3.9
1	C	145	SER	3.8
1	C	144	GLU	3.7
1	B	140	GLU	3.7
1	D	146	SER	3.5
1	B	143	GLU	3.4
1	B	446	ARG	3.3
1	D	145	SER	3.2
1	A	140	GLU	3.1
1	C	142	THR	3.0
1	D	143	GLU	3.0
1	A	145	SER	2.9
1	A	283	SER	2.9
1	A	137	VAL	2.8
1	A	148	VAL	2.7
1	B	146	SER	2.7
1	B	149	GLN	2.5
1	A	149	GLN	2.4
1	D	302	GLN	2.4
1	B	147	ALA	2.4
1	C	139	SER	2.3
1	A	139	SER	2.3
1	C	146	SER	2.2
1	D	149	GLN	2.2
1	C	147	ALA	2.2
1	B	445	LYS	2.2
1	A	264	LEU	2.2
1	C	259	PRO	2.2
1	D	446	ARG	2.2
1	D	475	PHE	2.1
1	A	136	SER	2.1
1	D	141	ARG	2.1
1	B	138	PHE	2.1
1	D	312	ASN	2.1
1	C	149	GLN	2.1
1	B	471	LYS	2.0
1	C	256	ILE	2.0
1	D	173	TYR	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	141	ARG	2.0

## 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, 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
3	EDO	B	504	4/4	0.80	0.35	10.49	65,66,66,67	0
4	PEG	A	505	7/7	0.82	0.24	7.48	54,65,70,70	0
4	PEG	C	506	7/7	0.86	0.19	2.75	56,67,70,70	0
4	PEG	B	507	7/7	0.85	0.17	2.35	72,86,95,96	0
4	PEG	D	506	7/7	0.90	0.33	2.01	36,43,49,54	0
6	PG4	C	507	13/13	0.86	0.23	1.95	57,61,68,70	0
3	EDO	D	505	4/4	0.60	0.28	1.83	62,62,63,64	0
3	EDO	B	502	4/4	0.81	0.27	1.80	63,63,64,64	0
3	EDO	D	504	4/4	0.72	0.22	1.28	67,67,68,68	0
3	EDO	A	503	4/4	0.89	0.20	1.02	48,49,49,49	0
5	M2M	A	508	9/9	0.81	0.18	0.96	66,79,85,86	0
4	PEG	A	507	7/7	0.80	0.20	0.75	84,101,113,113	0
3	EDO	A	504	4/4	0.65	0.29	0.73	70,72,72,73	0
2	6CX	A	501	32/32	0.96	0.16	0.47	21,27,54,65	0
2	6CX	C	501	32/32	0.95	0.15	0.43	20,30,57,69	0
3	EDO	B	506	4/4	0.87	0.15	0.39	33,36,41,42	0
3	EDO	B	505	4/4	0.61	0.16	0.27	60,61,62,62	0
2	6CX	D	501	32/32	0.95	0.12	0.01	22,32,54,65	0
2	6CX	B	501	32/32	0.96	0.12	-0.07	24,31,44,53	0
3	EDO	C	503	4/4	0.91	0.11	-0.38	49,49,50,52	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	B	503	4/4	0.93	0.10	-0.45	48,48,51,53	0
3	EDO	D	502	4/4	0.87	0.12	-	40,40,41,41	0
3	EDO	C	504	4/4	0.85	0.16	-	46,47,48,52	0
3	EDO	A	502	4/4	0.86	0.21	-	45,46,48,50	0
3	EDO	D	503	4/4	0.58	0.28	-	74,78,79,79	0
4	PEG	B	508	7/7	0.81	0.20	-	51,61,68,71	0
3	EDO	C	502	4/4	0.95	0.11	-	42,43,45,45	0
4	PEG	A	506	7/7	0.93	0.15	-	47,57,70,72	0
3	EDO	C	505	4/4	0.83	0.33	-	33,36,43,45	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.