



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2017 – 08:11 PM EDT

PDB ID : 2NP6
Title : Crystal structure of the adenine-specific DNA methyltransferase M.TaqI complexed with the cofactor analog AETA and a 10 bp DNA containing an abasic site analog at the target position
Authors : Lenz, T.; Scheidig, A.J.; Weinhold, E.
Deposited on : unknown
Resolution : 2.10 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

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-20030345
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-20030345

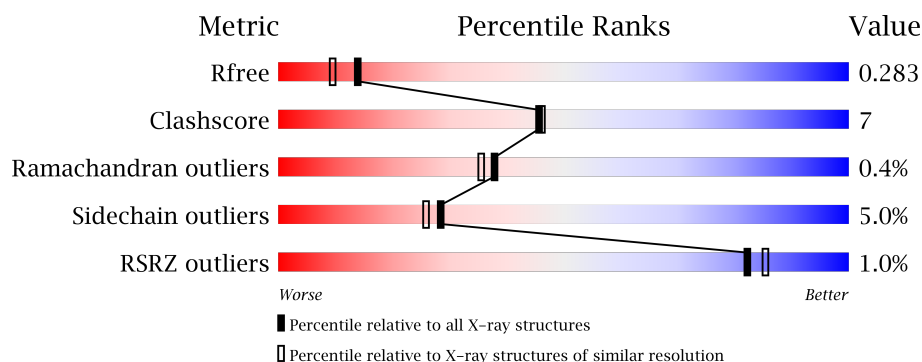
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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	B	10	<div> <div>30%</div> <div>40%</div> <div>30%</div> </div>
1	E	10	<div> <div>30%</div> <div>40%</div> <div>30%</div> </div>
2	C	10	<div> <div>70%</div> <div>30%</div> </div>
2	F	10	<div> <div>60%</div> <div>40%</div> </div>
3	A	421	<div> <div>77%</div> <div>14%</div> <div>6%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	421	 .% 76% 15% 7%

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	IPA	A	703	-	-	-	X
4	IPA	B	705	-	-	X	X
4	IPA	E	704	-	-	X	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7773 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 DNA chain called 5'-D(*GP*TP*TP*CP*GP*(3DR)P*TP*GP*TP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	10	Total	C	N	O	P	0	0	0
			192	93	29	61	9			
1	E	10	Total	C	N	O	P	0	0	0
			192	93	29	61	9			

- Molecule 2 is a DNA chain called 5'-D(*GP*AP*CP*AP*TP*CP*GP*(6MA)P*AP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	10	Total	C	N	O	P	0	0	0
			203	98	41	55	9			
2	F	10	Total	C	N	O	P	0	0	0
			203	98	41	55	9			

- Molecule 3 is a protein called Modification methylase TaqI.

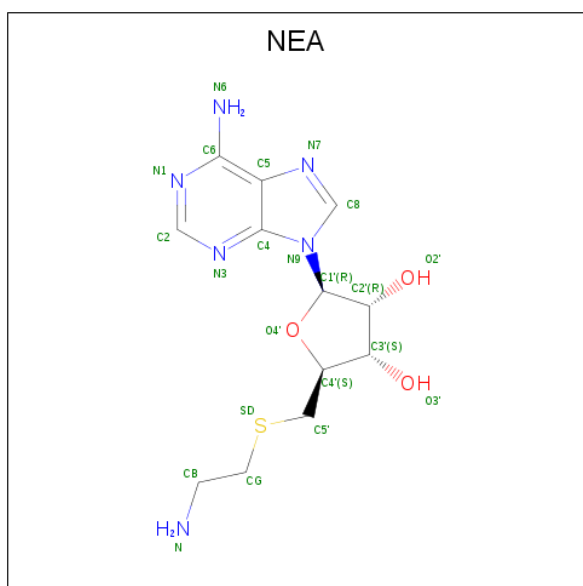
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	394	Total	C	N	O	S	0	0	0
			3200	2093	552	549	6			
3	D	393	Total	C	N	O	S	0	0	0
			3189	2087	548	548	6			

- Molecule 4 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C₃H₈O).



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

- Molecule 5 is 5'-DEOXY-5'-[2-(AMINO)ETHYLTHIO]ADENOSINE (three-letter code: NEA) (formula: C₁₂H₁₈N₆O₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			22	12	6	3	1		
5	D	1	Total	C	N	O	S	0	0
			22	12	6	3	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			6	3	3		
6	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	29	Total	O	0	0
			29	29		
7	C	18	Total	O	0	0
			18	18		
7	E	30	Total	O	0	0
			30	30		
7	F	24	Total	O	0	0
			24	24		
7	A	227	Total	O	0	0
			227	227		
7	D	198	Total	O	0	0
			198	198		

3 Residue-property plots

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: 5'-D(*GP*TP*TP*CP*GP*(3DR)P*TP*GP*TP*C)-3'

Chain B: 



- Molecule 1: 5'-D(*GP*TP*TP*CP*GP*(3DR)P*TP*GP*TP*C)-3'

Chain E: 



- Molecule 2: 5'-D(*GP*AP*CP*AP*TP*CP*GP*(6MA)P*AP*C)-3'

Chain C: 




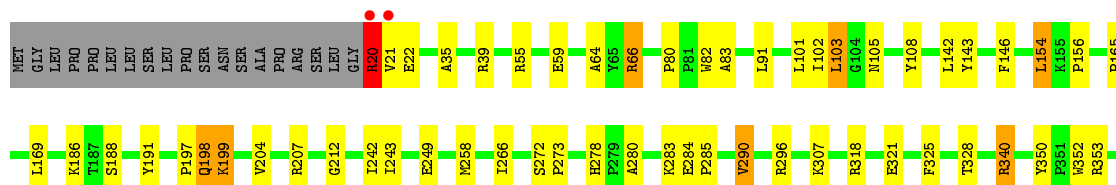
- Molecule 2: 5'-D(*GP*AP*CP*AP*TP*CP*GP*(6MA)P*AP*C)-3'

Chain F: 



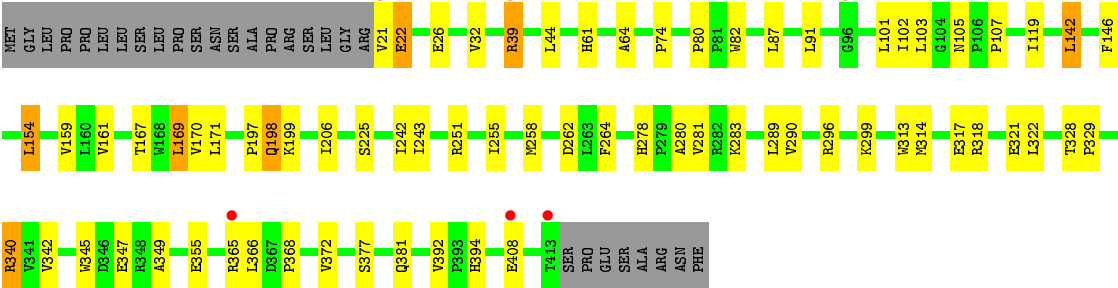
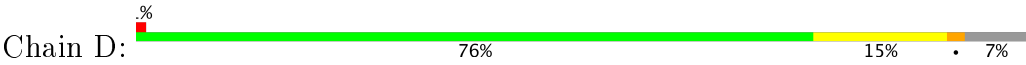
- Molecule 3: Modification methylase TaqI

Chain A: 





● Molecule 3: Modification methylase TaqI



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	59.34Å 68.82Å 114.20Å 90.00° 92.22° 90.00°	Depositor
Resolution (Å)	19.76 – 2.10 19.77 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.1 (19.76-2.10) 88.9 (19.77-2.10)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	0.17	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.228 , 0.286 0.227 , 0.283	Depositor DCC
R_{free} test set	2668 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	14.2	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 31.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.115 for h,-k,-l	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	7773	wwPDB-VP
Average B, all atoms (Å ²)	10.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NEA, GOL, IPA, 3DR, 6MA

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	B	0.89	0/200	1.70	6/305 (2.0%)
1	E	0.75	0/200	1.77	7/305 (2.3%)
2	C	0.83	0/202	1.47	3/307 (1.0%)
2	F	0.75	0/202	1.74	6/307 (2.0%)
3	A	0.45	0/3304	0.62	1/4489 (0.0%)
3	D	0.44	0/3293	0.62	1/4475 (0.0%)
All	All	0.50	0/7401	0.82	24/10188 (0.2%)

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
3	A	0	1

There are no bond length outliers.

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	12	DA	O4'-C1'-N9	-9.77	101.16	108.00
1	B	8	DG	O4'-C1'-N9	-7.50	102.75	108.00
1	E	5	DG	O4'-C1'-N9	7.37	113.16	108.00
1	E	3	DT	O4'-C1'-N1	-7.35	102.86	108.00
1	E	8	DG	O4'-C1'-N9	-7.22	102.94	108.00
1	E	4	DC	O4'-C1'-N1	-7.12	103.01	108.00
2	F	20	DC	O4'-C1'-N1	7.10	112.97	108.00
2	F	11	DG	P-O3'-C3'	6.86	127.94	119.70
2	F	11	DG	C1'-O4'-C4'	-6.72	103.38	110.10
2	C	20	DC	O4'-C1'-N1	6.64	112.65	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	3	DT	C6-C5-C7	-6.62	118.93	122.90
2	C	12	DA	O4'-C1'-N9	-6.38	103.53	108.00
1	B	4	DC	P-O3'-C3'	6.19	127.13	119.70
2	F	19	DA	O4'-C1'-N9	6.16	112.31	108.00
2	F	11	DG	O4'-C1'-N9	5.94	112.16	108.00
3	D	169	LEU	CA-CB-CG	5.85	128.75	115.30
3	A	103	LEU	CA-CB-CG	5.81	128.66	115.30
2	C	13	DC	O4'-C1'-N1	5.73	112.01	108.00
1	E	1	DG	O4'-C1'-N9	5.69	111.98	108.00
1	B	9	DT	P-O3'-C3'	5.60	126.42	119.70
1	B	3	DT	N1-C1'-C2'	5.41	122.87	112.60
1	E	7	DT	C6-C5-C7	-5.28	119.73	122.90
1	B	7	DT	N3-C4-O4	5.21	123.03	119.90
1	B	5	DG	O4'-C1'-N9	5.11	111.58	108.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	20	ARG	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	B	192	0	113	5	0
1	E	192	0	113	6	0
2	C	203	0	115	0	0
2	F	203	0	115	0	0
3	A	3200	0	3197	50	0
3	D	3189	0	3184	46	0
4	A	4	0	8	1	0
4	B	4	0	8	4	0
4	E	4	0	8	4	0
5	A	22	0	18	2	0
5	D	22	0	18	1	0
6	D	12	0	16	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	227	0	0	2	0
7	B	29	0	0	0	0
7	C	18	0	0	0	0
7	D	198	0	0	0	0
7	E	30	0	0	0	0
7	F	24	0	0	1	0
All	All	7773	0	6913	103	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 7.

All (103) 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)
3:A:20:ARG:HB3	3:A:199:LYS:HD3	1.22	1.11
3:A:66:ARG:HG2	3:A:66:ARG:HH11	0.98	1.07
3:D:340:ARG:HB3	3:D:340:ARG:HH11	1.32	0.94
3:A:340:ARG:HG2	3:A:340:ARG:HH11	1.32	0.94
3:A:66:ARG:NH1	3:A:66:ARG:HG2	1.78	0.92
3:A:102:ILE:HD12	3:A:154:LEU:HD13	1.56	0.86
1:E:5:DG:N7	3:D:394:HIS:HE1	1.74	0.86
3:A:66:ARG:HH11	3:A:66:ARG:CG	1.88	0.84
3:A:284:GLU:HG3	3:A:285:PRO:HD2	1.61	0.82
3:D:318:ARG:HD2	3:D:321:GLU:OE2	1.81	0.81
1:B:8:DG:H21	4:B:705:IPA:H12	1.46	0.79
3:D:21:VAL:O	3:D:22:GLU:HB2	1.84	0.77
1:B:5:DG:N7	3:A:394:HIS:HE1	1.85	0.74
3:A:20:ARG:HH11	5:A:601:NEA:HG2	1.52	0.73
4:B:705:IPA:C1	3:A:296:ARG:HH22	2.01	0.73
3:A:340:ARG:CG	3:A:340:ARG:HH11	1.99	0.72
3:A:197:PRO:O	3:A:198:GLN:HG2	1.93	0.69
3:A:20:ARG:CB	3:A:199:LYS:HD3	2.14	0.68
3:D:278:HIS:CD2	3:D:280:ALA:H	2.13	0.67
3:D:102:ILE:HD12	3:D:154:LEU:HD13	1.75	0.67
7:F:44:HOH:O	3:D:299:LYS:HE3	1.93	0.67
3:A:80:PRO:HG2	3:A:82:TRP:CZ2	2.31	0.66
3:D:278:HIS:HD2	3:D:280:ALA:H	1.44	0.65
3:A:272:SER:OG	3:A:273:PRO:HD3	1.98	0.63
3:A:370:SER:HG	3:A:409:TYR:HD1	1.47	0.62
3:A:22:GLU:HB2	7:A:820:HOH:O	1.98	0.62
3:A:340:ARG:HG2	3:A:340:ARG:NH1	2.10	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:8:DG:N3	4:E:704:IPA:H12	2.14	0.62
4:B:705:IPA:H13	3:A:296:ARG:HH22	1.63	0.61
1:E:5:DG:N7	3:D:394:HIS:CE1	2.64	0.61
3:A:21:VAL:HG23	5:A:601:NEA:H5'1	1.84	0.60
3:A:325:PHE:HB2	3:A:359:LEU:HD11	1.82	0.60
4:E:704:IPA:C3	3:D:296:ARG:HH22	2.17	0.58
1:B:8:DG:N2	4:B:705:IPA:H12	2.18	0.57
3:A:368:PRO:O	3:A:372:VAL:HG23	2.05	0.57
3:A:370:SER:OG	3:A:409:TYR:HD1	1.87	0.56
3:D:21:VAL:HG11	3:D:199:LYS:HZ2	1.70	0.56
3:D:107:PRO:HA	6:D:702:GOL:H31	1.88	0.55
3:D:368:PRO:O	3:D:372:VAL:HG23	2.05	0.55
3:A:91:LEU:HD21	3:A:146:PHE:CD1	2.42	0.55
3:D:340:ARG:CB	3:D:340:ARG:HH11	2.14	0.54
3:D:281:VAL:HG22	3:D:314:MET:HB3	1.90	0.54
3:A:278:HIS:HD2	3:A:280:ALA:H	1.56	0.53
3:D:377:SER:O	3:D:381:GLN:HG2	2.08	0.53
1:E:8:DG:H21	4:E:704:IPA:H33	1.73	0.53
1:B:5:DG:N7	3:A:394:HIS:CE1	2.73	0.52
3:D:197:PRO:C	3:D:198:GLN:HG3	2.30	0.52
3:A:318:ARG:HD2	3:A:321:GLU:OE2	2.10	0.51
3:A:242:ILE:O	4:A:703:IPA:H33	2.11	0.51
3:D:169:LEU:HB3	3:D:243:ILE:HB	1.91	0.51
3:D:318:ARG:CD	3:D:321:GLU:OE2	2.56	0.51
3:D:101:LEU:HA	3:D:159:VAL:O	2.11	0.50
3:A:169:LEU:HB3	3:A:243:ILE:HB	1.93	0.49
3:D:258:MET:CE	3:D:262:ASP:HB3	2.43	0.49
3:D:167:THR:HG22	3:D:392:VAL:HG12	1.95	0.48
3:D:21:VAL:O	3:D:22:GLU:CB	2.59	0.48
3:D:142:LEU:HD11	6:D:701:GOL:H32	1.96	0.47
3:D:264:PHE:CZ	3:D:366:LEU:HD22	2.49	0.47
3:A:280:ALA:HB2	3:A:318:ARG:CZ	2.45	0.47
3:D:329:PRO:HB3	3:D:347:GLU:HA	1.97	0.46
3:A:377:SER:O	3:A:381:GLN:HG2	2.15	0.46
3:D:32:VAL:HG13	3:D:61:HIS:HE1	1.79	0.46
3:D:44:LEU:HD23	3:D:102:ILE:HG12	1.97	0.46
3:D:119:ILE:HG13	3:D:142:LEU:HD22	1.98	0.46
4:E:704:IPA:H31	3:D:296:ARG:HH12	1.81	0.45
3:A:39:ARG:HD2	3:A:64:ALA:HB3	1.98	0.45
3:A:20:ARG:HD3	3:A:108:TYR:CE2	2.52	0.45
3:A:284:GLU:CG	3:A:285:PRO:HD2	2.41	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:191:TYR:HA	3:A:204:VAL:HG12	1.99	0.45
3:D:39:ARG:HD3	3:D:64:ALA:HB1	1.98	0.45
3:D:80:PRO:HG2	3:D:82:TRP:CZ2	2.52	0.45
3:A:290:VAL:HG21	3:A:350:TYR:CZ	2.51	0.45
3:D:171:LEU:HD21	3:D:392:VAL:HG11	1.98	0.45
3:A:66:ARG:NH1	3:A:66:ARG:CG	2.59	0.44
1:B:2:DT:H2"	1:B:3:DT:H72	1.98	0.44
3:D:26:GLU:H	3:D:26:GLU:CD	2.21	0.44
3:A:188:SER:HB2	3:A:207:ARG:HB3	1.99	0.44
3:A:35:ALA:HB1	3:A:101:LEU:HD21	2.01	0.43
3:A:20:ARG:HD3	3:A:108:TYR:CZ	2.53	0.43
3:A:266:ILE:HG12	3:A:358:LEU:HG	1.99	0.43
3:D:258:MET:HE2	3:D:262:ASP:HB3	2.01	0.43
1:E:1:DG:H8	1:E:1:DG:H5"	1.84	0.43
3:A:272:SER:HB3	3:A:353:ARG:HD2	2.00	0.42
3:D:32:VAL:CG1	3:D:61:HIS:HE1	2.32	0.42
3:D:251:ARG:O	3:D:255:ILE:HG12	2.18	0.42
3:D:161:VAL:HA	3:D:206:ILE:O	2.19	0.42
3:A:20:ARG:O	3:A:20:ARG:HG3	2.20	0.42
3:A:82:TRP:CZ3	3:A:83:ALA:HB2	2.55	0.42
3:D:21:VAL:N	5:D:602:NEA:HG1	2.35	0.41
1:E:1:DG:H2"	1:E:2:DT:O5'	2.21	0.41
3:D:170:VAL:HA	3:D:242:ILE:HG23	2.03	0.41
3:D:74:PRO:HA	3:D:87:LEU:HD21	2.03	0.41
3:D:365:ARG:HD2	3:D:365:ARG:N	2.36	0.41
3:A:21:VAL:HG13	7:A:887:HOH:O	2.20	0.41
3:A:278:HIS:CD2	3:A:280:ALA:H	2.35	0.41
3:D:340:ARG:HD2	3:D:342:VAL:CG2	2.51	0.41
3:D:91:LEU:HD21	3:D:146:PHE:CD1	2.56	0.41
3:A:108:TYR:CZ	3:A:165:PRO:HD3	2.56	0.40
3:A:186:LYS:HE2	3:A:212:GLY:HA3	2.02	0.40
3:D:345:TRP:HD1	3:D:347:GLU:HB3	1.86	0.40
3:A:352:TRP:CG	3:A:357:HIS:HE1	2.40	0.40
3:A:55:ARG:O	3:A:59:GLU:HG2	2.21	0.40
3:D:289:LEU:HB3	3:D:313:TRP:HB3	2.03	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	
3	A	392/421 (93%)	378 (96%)	14 (4%)	0	100	100
3	D	391/421 (93%)	373 (95%)	15 (4%)	3 (1%)	22	17
All	All	783/842 (93%)	751 (96%)	29 (4%)	3 (0%)	38	35

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	22	GLU
3	D	349	ALA
3	D	355	GLU

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	
3	A	333/356 (94%)	314 (94%)	19 (6%)	24	21
3	D	332/356 (93%)	318 (96%)	14 (4%)	34	33
All	All	665/712 (93%)	632 (95%)	33 (5%)	28	26

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	20	ARG
3	A	66	ARG
3	A	103	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	105	ASN
3	A	142	LEU
3	A	143	TYR
3	A	154	LEU
3	A	156	PRO
3	A	198	GLN
3	A	199	LYS
3	A	249	GLU
3	A	258	MET
3	A	283	LYS
3	A	290	VAL
3	A	307	LYS
3	A	328	THR
3	A	340	ARG
3	A	366	LEU
3	A	370	SER
3	D	39	ARG
3	D	103	LEU
3	D	105	ASN
3	D	142	LEU
3	D	154	LEU
3	D	198	GLN
3	D	225	SER
3	D	283	LYS
3	D	290	VAL
3	D	317	GLU
3	D	322	LEU
3	D	328	THR
3	D	340	ARG
3	D	408	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	50	HIS
3	A	278	HIS
3	A	394	HIS
3	D	61	HIS
3	D	105	ASN
3	D	198	GLN
3	D	278	HIS
3	D	373	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	394	HIS

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	3DR	B	6	1	8,11,12	0.84	1 (12%)	8,14,17	0.70	0
2	6MA	C	18	1,2	17,24,25	0.74	0	17,34,37	1.53	2 (11%)
1	3DR	E	6	1	8,11,12	0.45	0	8,14,17	0.67	0
2	6MA	F	18	1,2	17,24,25	0.76	0	17,34,37	1.39	2 (11%)

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	3DR	B	6	1	-	0/3/15/16	0/1/1/1
2	6MA	C	18	1,2	-	0/5/23/24	0/3/3/3
1	3DR	E	6	1	-	0/3/15/16	0/1/1/1
2	6MA	F	18	1,2	-	0/5/23/24	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	6	3DR	O5'-C5'	-2.12	1.41	1.44

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	18	6MA	C5-C6-N6	2.68	125.08	120.33
2	C	18	6MA	C5-C6-N6	3.07	125.76	120.33
2	F	18	6MA	C2-N1-C6	4.34	119.36	116.53
2	C	18	6MA	C2-N1-C6	4.64	119.55	116.53

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](#)

7 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	$\# Z > 2$
5	NEA	A	601	-	21,24,24	1.06	2 (9%)	20,34,34	2.57	2 (10%)
4	IPA	A	703	-	3,3,3	0.56	0	3,3,3	0.24	0
4	IPA	B	705	-	3,3,3	0.51	0	3,3,3	0.54	0
5	NEA	D	602	-	21,24,24	1.13	2 (9%)	20,34,34	2.69	2 (10%)
6	GOL	D	701	-	5,5,5	0.32	0	5,5,5	0.31	0
6	GOL	D	702	-	5,5,5	0.53	0	5,5,5	0.36	0
4	IPA	E	704	-	3,3,3	0.44	0	3,3,3	0.61	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
5	NEA	A	601	-	-	0/5/25/25	0/3/3/3
4	IPA	A	703	-	-	0/0/0/0	0/0/0/0
4	IPA	B	705	-	-	0/0/0/0	0/0/0/0
5	NEA	D	602	-	-	0/5/25/25	0/3/3/3
6	GOL	D	701	-	-	0/4/4/4	0/0/0/0
6	GOL	D	702	-	-	0/4/4/4	0/0/0/0
4	IPA	E	704	-	-	0/0/0/0	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	602	NEA	C2-N1	2.27	1.38	1.33
5	A	601	NEA	C2-N1	2.37	1.38	1.33
5	A	601	NEA	C2-N3	3.26	1.37	1.32
5	D	602	NEA	C2-N3	3.92	1.38	1.32

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	602	NEA	N3-C2-N1	-10.80	119.45	128.86
5	A	601	NEA	N3-C2-N1	-10.46	119.75	128.86
5	D	602	NEA	C5'-SD-CG	-3.97	90.27	102.29
5	A	601	NEA	C5'-SD-CG	-3.51	91.65	102.29

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	601	NEA	2	0
4	A	703	IPA	1	0
4	B	705	IPA	4	0
5	D	602	NEA	1	0
6	D	701	GOL	1	0
6	D	702	GOL	1	0
4	E	704	IPA	4	0

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, 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(Å²)	Q<0.9	
1	B	9/10 (90%)	-0.57	0	100	100	5, 7, 10, 13	0
1	E	9/10 (90%)	-0.60	0	100	100	4, 8, 10, 20	0
2	C	9/10 (90%)	-0.59	0	100	100	5, 6, 9, 14	0
2	F	9/10 (90%)	-0.56	0	100	100	4, 7, 15, 19	0
3	A	394/421 (93%)	0.08	2 (0%)	90	92	3, 10, 19, 26	0
3	D	393/421 (93%)	0.02	6 (1%)	74	77	3, 10, 19, 36	0
All	All	823/882 (93%)	0.02	8 (0%)	82	85	3, 10, 19, 36	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	21	VAL	6.6
3	A	21	VAL	5.4
3	A	20	ARG	2.3
3	D	39	ARG	2.3
3	D	408	GLU	2.1
3	D	365	ARG	2.1
3	D	96	GLY	2.1
3	D	413	THR	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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
2	6MA	C	18	22/23	0.97	0.09	-	2,5,10,10	0
1	3DR	B	6	11/12	0.97	0.10	-	5,6,7,8	0
1	3DR	E	6	11/12	0.97	0.09	-	4,7,8,9	0
2	6MA	F	18	22/23	0.96	0.09	-	2,5,8,9	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, 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	IPA	A	703	4/4	0.82	0.21	8.22	19,19,19,20	0
4	IPA	B	705	4/4	0.78	0.20	6.86	22,23,23,24	0
4	IPA	E	704	4/4	0.83	0.19	6.73	12,12,13,14	0
6	GOL	D	702	6/6	0.82	0.24	0.60	17,19,20,21	0
5	NEA	D	602	22/22	0.94	0.13	0.00	5,7,9,11	0
6	GOL	D	701	6/6	0.91	0.13	-0.15	26,26,27,28	0
5	NEA	A	601	22/22	0.95	0.12	-0.27	7,8,8,9	0

6.5 Other polymers [i](#)

There are no such residues in this entry.