wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 01:16 PM GMT

PDB ID	:	3TAE
Title	:	5-hydroxycytosine paired with dAMP in RB69 gp43
Authors	:	Zahn, K.E.
Deposited on	:	2011-08-04
Resolution	:	2.71 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 (i) symbol.

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.71 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.
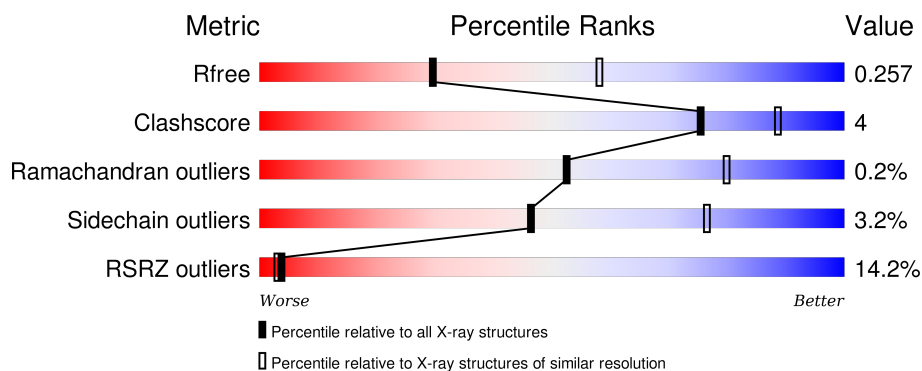
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2439 (2.74-2.70)
Clashscore	102246	2771 (2.74-2.70)
Ramachandran outliers	100387	2726 (2.74-2.70)
Sidechain outliers	100360	2727 (2.74-2.70)
RSRZ outliers	91569	2443 (2.74-2.70)

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 <=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	E	18	11% · 56% / 44%
1	G	18	39% · 44% / 44% / 6% / 6%
1	I	18	72% / 22% / 6%
1	K	18	56% · 72% / 11% / 17%
2	F	15	7% · 60% / 33% / 7%

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	15	40% ... 73% ... 13% ... 13%
2	J	15	53% ... 47%
2	L	15	40% ... 47% ... 20% ... 7% ... 27%
3	A	906	6% ... 87% ... 12%
3	B	906	18% ... 88% ... 10%
3	C	906	5% ... 87% ... 12%
3	D	906	27% ... 91% ... 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
1	5OC	I	3	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 32066 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 DNA (5'-D(*CP*CP*(5OC)P*GP*GP*TP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	18	Total	C	N	O	P	0	0	0
			366	173	70	106	17			
1	I	17	Total	C	N	O	P	0	0	0
			350	164	67	102	17			
1	G	17	Total	C	N	O	P	0	0	1
			313	145	59	93	16			
1	K	15	Total	C	N	O	P	0	0	1
			271	125	52	80	14			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	15	Total	C	N	O	P	0	0	0
			286	135	51	86	14			
2	J	15	Total	C	N	O	P	0	0	0
			303	145	56	88	14			
2	H	13	Total	C	N	O	P	0	0	0
			263	126	48	77	12			
2	L	11	Total	C	N	O	P	0	0	0
			223	107	40	66	10			

- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	903	Total	C	N	O	S	0	0	0
			7374	4737	1226	1378	33			
3	B	901	Total	C	N	O	S	0	0	0
			7355	4724	1224	1374	33			
3	C	903	Total	C	N	O	S	0	0	0
			7374	4737	1226	1378	33			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	898	7328	4706	1221	1369	32	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	222	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
A	327	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
A	904	HIS	-	EXPRESSION TAG	UNP Q38087
A	905	HIS	-	EXPRESSION TAG	UNP Q38087
A	906	HIS	-	EXPRESSION TAG	UNP Q38087
B	222	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
B	327	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
B	904	HIS	-	EXPRESSION TAG	UNP Q38087
B	905	HIS	-	EXPRESSION TAG	UNP Q38087
B	906	HIS	-	EXPRESSION TAG	UNP Q38087
C	222	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
C	327	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
C	904	HIS	-	EXPRESSION TAG	UNP Q38087
C	905	HIS	-	EXPRESSION TAG	UNP Q38087
C	906	HIS	-	EXPRESSION TAG	UNP Q38087
D	222	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
D	327	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
D	904	HIS	-	EXPRESSION TAG	UNP Q38087
D	905	HIS	-	EXPRESSION TAG	UNP Q38087
D	906	HIS	-	EXPRESSION TAG	UNP Q38087

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	F	1	Total 5	O 4	S 1	0	0
4	H	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	3	Total 3	O 3	0	0
5	F	2	Total 2	O 2	0	0
5	I	6	Total 6	O 6	0	0
5	J	4	Total 4	O 4	0	0
5	G	1	Total 1	O 1	0	0
5	H	1	Total 1	O 1	0	0
5	A	101	Total 101	O 101	0	0
5	B	47	Total 47	O 47	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	71	Total 71	O 71	0	0
5	D	4	Total 4	O 4	0	0

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 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.
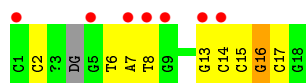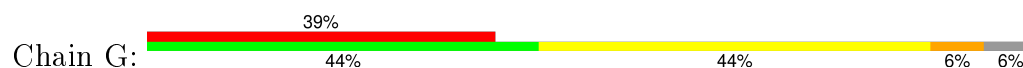
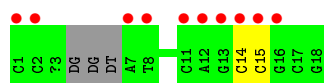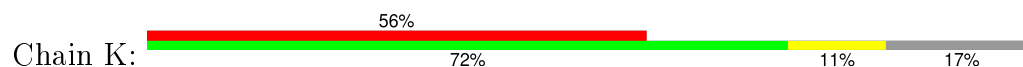- Molecule 1: DNA (5'-D(*CP*CP*(5OC)P*GP*GP*TP*AP*TP*GP*AP*CP*AP*GP*CP*CP *GP*CP*G)-3')

- Molecule 1: DNA (5'-D(*CP*CP*(5OC)P*GP*GP*TP*AP*TP*GP*AP*CP*AP*GP*CP*CP *GP*CP*G)-3')

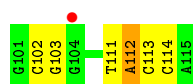- Molecule 1: DNA (5'-D(*CP*CP*(5OC)P*GP*GP*TP*AP*TP*GP*AP*CP*AP*GP*CP*CP *GP*CP*G)-3')

- Molecule 1: DNA (5'-D(*CP*CP*(5OC)P*GP*GP*TP*AP*TP*GP*AP*CP*AP*GP*CP*CP *GP*CP*G)-3')

- Molecule 2: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*AP*CP*CP*A)-3')

G101 G102 G103 G104 T111 A112 C113 C114 A115

• Molecule 2: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*AP*CP*CP*A)-3')

Chain J:

53%	47%

G101 G102 G103 A112 C113 C114 A115

• Molecule 2: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*AP*CP*CP*A)-3')

Chain H:

40%

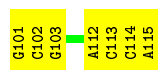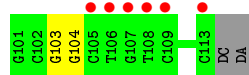|		
73%	13%	13%

G101 G102 G103 G104 C105 T106 G107 T108 C109 C113 DC DA

• Molecule 2: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*AP*CP*CP*A)-3')

Chain L:

40%

47%	20%	7%	27%	

G101 C102 G103 G104 C105 A110 T111 DA DC DC DA

• Molecule 3: DNA polymerase

Chain A:

6%

87%	12%	

M1 K2 E3 E9 I115 R18 Y19 I20 R27 T28 R29 C41 P42 I52 M65 F88 E100 N112 F113 D114 I115 E116 A135 I136 T137 F149 N153 S154 P155 Y156 G157 L170 Q171 D192 N193 E194 L197 L198 L202 W205 Q206

V211 I212 L213 T214 E236 S237 T238 A239 K240 R241 R249 V252 I253 E254 N255 M256 Y257 F266 D272 Y273 I274 D275 L276 T283 N284 Q285 L290 D291 G301 K302 L303 R313 E314 R319 I326 R330 V331 L332 F340 L343 I362 E375 R384

Q389 F395 P417 N424 I425 S426 P427 E428 T429 L439 N444 Y463 D466 I475 Q493 R494 N495 G496 E497 I498 I499 K500 E501 A502 L503 H504 N505 P506 N507 L508 S509 V510 D511 E512 P513 L514 D515 V516 D517 Y518 F522 S523 D524 K529 I530 K531 K532 L533

S534 A535 K536 S537 L538 N539 E540 M541 R544 R547 T548 M553 I557 N558 R559 L566 N572 L580 T584 Q591 M592 A593 W596 R599 A629 I633 F641 H646 A655 M659 M670 L693 R707 W713 T718 K724

K739 E757 S781 V793 R813 N818 A821 P822 Q823 G842 D843 W848 I854 L857 I858 K859 W865 K878 P879 L889 L897 F903 HIS HIS HIS

• Molecule 3: DNA polymerase

Chain B:

18%

88%

10%

M1 E9 S44 T47 M65 I78 G79 L80 F88 K89 L90 N112 F113 D114 I115 E116 V117 P120 D121 G122 Q128 I133 D134 A135 M153 S154 P155 Y156 G157 E160 E172 Q173 G174 D184 P190 F191 D192 L198 L202 Q206 T209

P210 V211 I212 L213 T214 E219 I223 P224 F234 G235 E236 K240 R241 P244 K251 V252 I253 E254 N255 M256 Y257 G258 S259 R260 I263 F266 Y273 L276 Y277 F280 T283 N284 Q285 P286 S287 Y288 Y292 I293 S294 F296 E297 L298 N299 V300 L303

S315 Y320 I321 N324 R330 L343 I362 I369 L373 I380 P381 F410 D411 L412 T413 S414 L415 N424 T429 L439 H440 D441 Y442 I443 A447 V453 Y454 P458 Y463 L490 G496 R497 I498 I499 A502 L503 H504 N505 P506 N507

L508 S509 V510 D511 E512 P513 L514 D515 V516 D517 Y518 R519 F520 D521 F522 S523 D524 E525 I526 K527 E528 K532 A535 K536 S537 L538 N539 E540 M541 L542 F543 R544 A545 Q546 R547 T548 E549 T554 A555 Q556 I557 N558 R559 S565 N572 R576 T587 M592 A593 W596

T622 I633 E638 F641 R642 L651 A655 M659 M670 M683 L693 W702 R707 T718 E731 S735 P738 Q742 E770 L774 I777 S781 A785 N786 N787 I788 A789 K790 Y791 D792 V793 G794 P797 G798 P799 K800 C801 P802

F803 H804 I805 I808 L809 T810 Y811 N812 R813 A814 I815 K816 G817 N818 I819 D820 A821 P822 Q823 V824 V825 E826 G827 E828 Y831 V832 L833 R836 E837 G838 N839 P840 F841 G842 D843 I846 A847 W848 P849 S850 G851 T852 E853 I854 T855 D856 L857 I858 K859 D860 D861 V862 L863 H864 W865 M866

D867 Y868 T875 P879 A895 S896 L897 F898 D899 M900 F901 ASP PHE HIS HIS HIS

● Molecule 3: DNA polymerase

Chain C:

5%

87%

12%

M1 I15 I52 A82 K107 V110 A111 N112 F113 E116 D121 E125 P126 A135 Q171 D192 N193 E194 K195 E196 L197 L198 M199 T209 P210 V211 I212 L213 T214 I223 P224 P244 K247 T248 R249 V250 K251 V252 I253 E254 N255 M256 Y257 R260

E261 I262 Y273 L276 Y277 S281 N284 E297 L298 V300 G301 R319 I326 A327 V328 R329 V331 I332 R338 Q339 F340 L343 M347 P361 I362 L373 Q382 Y391 F395 F396 K397 E398 P399 N402 R403 Y404 F410 D411 L412 T413 S414

L415 Y416 P417 N424 E428 T429 L439 Y463 D468 V471 I475 V478 Y488 M489 N495 I498 I499 K500 S509 V510 D511 E512 P513 S523 I526 K527 E528 K529 I530 K531 K532 L533 S534 F543 R547 M553 N558 R559 G568 F575

R581 T584 A585 I586 Q591 W596 V608 E612 T622 K631 I632 I633 F641 H646 L651 F667 M670 M683 L693 R707 M715 T718 P738 K739 E762 E768 N775 Y776 I777 R813 K816 G817 I819 D820 A821

K829 G842 I854 T855 I858 V862 M866 L872 B873 K874 B92 M900 F901 D902 F903 HIS HIS HIS

● Molecule 3: DNA polymerase

Chain D:

27%

91%

7%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	133.27Å 123.71Å 165.90Å 90.00° 95.97° 90.00°	Depositor
Resolution (Å)	30.00 – 2.71 30.31 – 2.71	Depositor EDS
% Data completeness (in resolution range)	98.0 (30.00-2.71) 98.1 (30.31-2.71)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.26 (at 2.72Å)	Xtriage
Refinement program	REFMAC 5.6.0116	Depositor
R, R_{free}	0.224 , 0.268 0.216 , 0.257	Depositor DCC
R_{free} test set	13797 reflections (10.71%)	DCC
Wilson B-factor (Å^2)	67.1	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 46.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Outliers	0 of 142606 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	32066	wwPDB-VP
Average B, all atoms (Å^2)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ 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: 5OC, SO4

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	E	0.23	0/387	0.84	0/593
1	G	0.24	0/345	0.80	1/529 (0.2%)
1	I	0.29	0/369	0.78	0/565
1	K	0.21	0/298	0.74	0/456
2	F	0.23	0/319	0.89	1/491 (0.2%)
2	H	0.23	0/294	0.81	0/452
2	J	0.27	0/339	0.83	0/521
2	L	0.23	0/249	0.97	3/383 (0.8%)
3	A	0.43	4/7555 (0.1%)	0.53	1/10209 (0.0%)
3	B	0.42	3/7535 (0.0%)	0.49	0/10182
3	C	0.42	1/7555 (0.0%)	0.51	0/10209
3	D	0.41	2/7507 (0.0%)	0.46	0/10145
All	All	0.41	10/32752 (0.0%)	0.54	6/44735 (0.0%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	702	TRP	CD2-CE2	5.30	1.47	1.41
3	A	713	TRP	CD2-CE2	5.13	1.47	1.41
3	D	574	TRP	CD2-CE2	5.12	1.47	1.41
3	A	205	TRP	CD2-CE2	5.12	1.47	1.41
3	C	596	TRP	CD2-CE2	5.12	1.47	1.41

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	103	DG	P-O3'-C3'	7.96	129.25	119.70
2	L	101	DG	P-O3'-C3'	6.89	127.97	119.70
1	G	16	DG	P-O3'-C3'	6.77	127.83	119.70
2	L	110	DA	P-O3'-C3'	6.12	127.05	119.70
2	F	112	DA	P-O3'-C3'	6.00	126.89	119.70

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	E	366	0	201	5	0
1	G	313	0	168	7	0
1	I	350	0	189	11	0
1	K	271	0	145	2	0
2	F	286	0	158	6	0
2	H	263	0	148	1	0
2	J	303	0	170	7	0
2	L	223	0	126	1	0
3	A	7374	0	7267	64	0
3	B	7355	0	7254	54	0
3	C	7374	0	7267	61	0
3	D	7328	0	7232	31	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	F	5	0	0	0	0
4	H	5	0	0	0	0
5	A	101	0	0	3	0
5	B	47	0	0	0	0
5	C	71	0	0	0	0
5	D	4	0	0	0	0
5	E	3	0	0	0	0
5	F	2	0	0	1	0
5	G	1	0	0	0	0
5	H	1	0	0	0	0
5	I	6	0	0	0	0
5	J	4	0	0	0	0
All	All	32066	0	30325	239	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 4.

The worst 5 of 239 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:I:2:DC:H4'	1:I:3:5OC:H5'A	1.23	1.21
1:I:3:5OC:O2	3:C:568:GLY:HA3	1.47	1.13
1:I:2:DC:H4'	1:I:3:5OC:C5'	1.91	0.99
2:F:112:DA:H2"	2:F:113:DC:H5"	1.49	0.94
1:G:2:DC:O2	1:G:2:DC:H2'	1.73	0.87

There are no symmetry-related clashes.

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	
3	A	901/906 (99%)	875 (97%)	25 (3%)	1 (0%)	56	83
3	B	899/906 (99%)	858 (95%)	38 (4%)	3 (0%)	46	74
3	C	901/906 (99%)	873 (97%)	27 (3%)	1 (0%)	56	83
3	D	896/906 (99%)	855 (95%)	39 (4%)	2 (0%)	52	80
All	All	3597/3624 (99%)	3461 (96%)	129 (4%)	7 (0%)	52	80

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	819	ILE
3	D	192	ASP
3	D	622	THR
3	B	896	SER
3	B	424	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	
3	A	800/803 (100%)	769 (96%)	31 (4%)	39	69
3	B	798/803 (99%)	769 (96%)	29 (4%)	42	72
3	C	800/803 (100%)	771 (96%)	29 (4%)	42	72
3	D	795/803 (99%)	783 (98%)	12 (2%)	72	91
All	All	3193/3212 (99%)	3092 (97%)	101 (3%)	46	76

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	414	SER
3	B	843	ASP
3	D	389	GLN
3	B	510	VAL
3	B	587	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	382	GLN
3	B	823	GLN
3	D	389	GLN
3	B	389	GLN
3	B	646	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	5OC	E	3	1	12,21,22	1.30	1 (8%)	12,30,33	0.36	0
1	5OC	G	3	1	0,3,22	0.00	-	0,3,33	0.00	-
1	5OC	I	3	1	12,21,22	1.34	1 (8%)	12,30,33	0.62	0
1	5OC	K	3	1	0,3,22	0.00	-	0,3,33	0.00	-

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	5OC	E	3	1	-	0/3/21/22	0/2/2/2
1	5OC	G	3	1	-	0/0/0/22	0/0/0/2
1	5OC	I	3	1	-	0/3/21/22	0/2/2/2
1	5OC	K	3	1	-	0/0/0/22	0/0/0/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	3	5OC	C6-C5	-3.85	1.34	1.39
1	E	3	5OC	C6-C5	-3.75	1.34	1.39

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	3	5OC	1	0
1	I	3	5OC	7	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

4 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	$\#|Z| > 2$
4	SO4	B	907	-	4,4,4	0.40	0	6,6,6	0.17	0
4	SO4	C	907	-	4,4,4	0.38	0	6,6,6	0.09	0
4	SO4	F	3	-	4,4,4	0.36	0	6,6,6	0.06	0
4	SO4	H	2	-	4,4,4	0.37	0	6,6,6	0.07	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
4	SO4	B	907	-	-	0/0/0/0	0/0/0/0
4	SO4	C	907	-	-	0/0/0/0	0/0/0/0
4	SO4	F	3	-	-	0/0/0/0	0/0/0/0
4	SO4	H	2	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

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 (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(Å²)	Q<0.9
1	E	17/18 (94%)	0.40	2 (11%)	6	5	57, 89, 126, 143	0
1	G	16/18 (88%)	1.96	7 (43%)	0	0	62, 114, 138, 155	0
1	I	16/18 (88%)	-0.29	0	100	100	37, 50, 70, 94	0
1	K	14/18 (77%)	3.02	10 (71%)	0	0	53, 162, 197, 200	0
2	F	15/15 (100%)	0.82	1 (6%)	21	20	72, 87, 138, 141	0
2	H	13/15 (86%)	2.29	6 (46%)	0	0	94, 116, 140, 141	0
2	J	15/15 (100%)	-0.26	0	100	100	36, 61, 87, 87	0
2	L	11/15 (73%)	2.39	6 (54%)	0	0	158, 169, 173, 174	0
3	A	903/906 (99%)	0.32	52 (5%)	26	26	30, 54, 146, 263	0
3	B	901/906 (99%)	1.00	159 (17%)	2	1	34, 80, 218, 286	0
3	C	903/906 (99%)	0.30	41 (4%)	37	36	28, 62, 135, 197	0
3	D	898/906 (99%)	1.44	243 (27%)	1	1	84, 135, 213, 253	0
All	All	3722/3756 (99%)	0.78	527 (14%)	4	3	28, 78, 201, 286	0

The worst 5 of 527 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	819	ILE	19.3
3	B	821	ALA	13.1
3	B	820	ASP	12.4
3	B	862	VAL	11.2
3	D	257	TYR	11.1

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, 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
1	5OC	G	3	4/21	0.96	0.12	-	108,111,111,113	0
1	5OC	I	3	20/21	0.90	0.22	-	78,91,100,100	0
1	5OC	K	3	4/21	0.65	0.32	-	123,130,134,136	0
1	5OC	E	3	20/21	0.62	0.35	-	148,155,160,160	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(Å^2)	Q<0.9
4	SO4	C	907	5/5	0.87	0.20	-0.42	112,115,116,120	0
4	SO4	B	907	5/5	0.93	0.32	-	83,85,88,88	0
4	SO4	H	2	5/5	0.80	0.24	-	126,128,128,130	0
4	SO4	F	3	5/5	0.90	0.16	-	121,121,123,124	0

6.5 Other polymers (i)

There are no such residues in this entry.