wwPDB X-ray Structure Validation Summary Report (i)

Aug 26, 2023 – 06:59 PM EDT

PDB ID	:	3G05
Title	:	Crystal structure of N-terminal domain (2-550) of E.coli MnmG
Authors	:	Shi, R.; Matte, A.; Cygler, M.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)
Deposited on	:	2009-01-27
Resolution	:	3.49 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

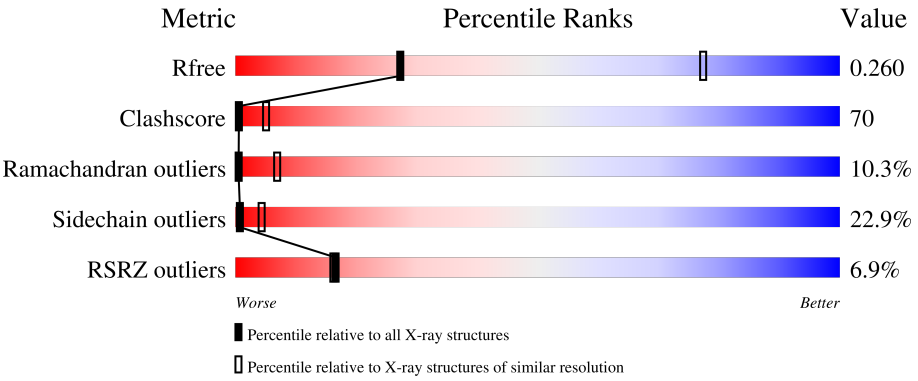
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.49 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}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	A	576	5% 25% 48% 16% • 9%
1	B	576	8% 21% 48% 19% • 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8156 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 tRNA uridine 5-carboxymethylaminomethyl modification enzyme mnmG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	527	Total	C	N	O	S	0	0	0
			4073	2551	728	776	18			
1	B	524	Total	C	N	O	S	0	0	0
			4063	2540	730	775	18			

There are 54 discrepancies between the modelled and reference sequences:

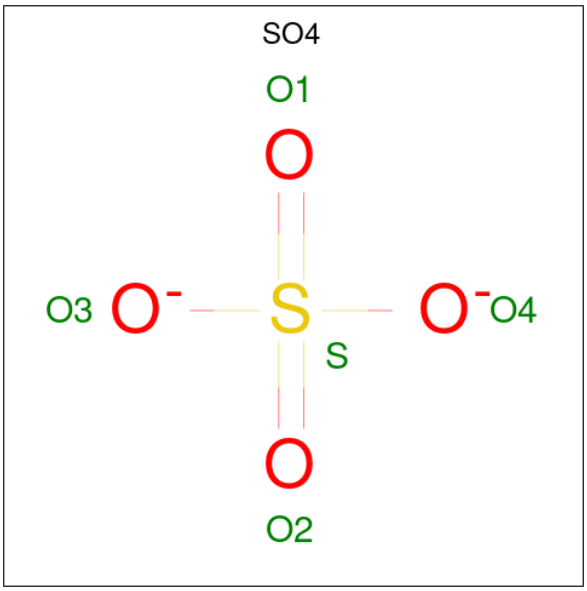
Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	expression tag	UNP Q8XAY0
A	-26	GLY	-	expression tag	UNP Q8XAY0
A	-25	SER	-	expression tag	UNP Q8XAY0
A	-24	SER	-	expression tag	UNP Q8XAY0
A	-23	HIS	-	expression tag	UNP Q8XAY0
A	-22	HIS	-	expression tag	UNP Q8XAY0
A	-21	HIS	-	expression tag	UNP Q8XAY0
A	-20	HIS	-	expression tag	UNP Q8XAY0
A	-19	HIS	-	expression tag	UNP Q8XAY0
A	-18	HIS	-	expression tag	UNP Q8XAY0
A	-17	HIS	-	expression tag	UNP Q8XAY0
A	-16	HIS	-	expression tag	UNP Q8XAY0
A	-15	ASP	-	expression tag	UNP Q8XAY0
A	-14	TYR	-	expression tag	UNP Q8XAY0
A	-13	ASP	-	expression tag	UNP Q8XAY0
A	-12	ILE	-	expression tag	UNP Q8XAY0
A	-11	PRO	-	expression tag	UNP Q8XAY0
A	-10	THR	-	expression tag	UNP Q8XAY0
A	-9	THR	-	expression tag	UNP Q8XAY0
A	-8	GLU	-	expression tag	UNP Q8XAY0
A	-7	ASN	-	expression tag	UNP Q8XAY0
A	-6	LEU	-	expression tag	UNP Q8XAY0
A	-5	TYR	-	expression tag	UNP Q8XAY0
A	-4	PHE	-	expression tag	UNP Q8XAY0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLN	-	expression tag	UNP Q8XAY0
A	-2	GLY	-	expression tag	UNP Q8XAY0
A	-1	SER	-	expression tag	UNP Q8XAY0
B	-27	MET	-	expression tag	UNP Q8XAY0
B	-26	GLY	-	expression tag	UNP Q8XAY0
B	-25	SER	-	expression tag	UNP Q8XAY0
B	-24	SER	-	expression tag	UNP Q8XAY0
B	-23	HIS	-	expression tag	UNP Q8XAY0
B	-22	HIS	-	expression tag	UNP Q8XAY0
B	-21	HIS	-	expression tag	UNP Q8XAY0
B	-20	HIS	-	expression tag	UNP Q8XAY0
B	-19	HIS	-	expression tag	UNP Q8XAY0
B	-18	HIS	-	expression tag	UNP Q8XAY0
B	-17	HIS	-	expression tag	UNP Q8XAY0
B	-16	HIS	-	expression tag	UNP Q8XAY0
B	-15	ASP	-	expression tag	UNP Q8XAY0
B	-14	TYR	-	expression tag	UNP Q8XAY0
B	-13	ASP	-	expression tag	UNP Q8XAY0
B	-12	ILE	-	expression tag	UNP Q8XAY0
B	-11	PRO	-	expression tag	UNP Q8XAY0
B	-10	THR	-	expression tag	UNP Q8XAY0
B	-9	THR	-	expression tag	UNP Q8XAY0
B	-8	GLU	-	expression tag	UNP Q8XAY0
B	-7	ASN	-	expression tag	UNP Q8XAY0
B	-6	LEU	-	expression tag	UNP Q8XAY0
B	-5	TYR	-	expression tag	UNP Q8XAY0
B	-4	PHE	-	expression tag	UNP Q8XAY0
B	-3	GLN	-	expression tag	UNP Q8XAY0
B	-2	GLY	-	expression tag	UNP Q8XAY0
B	-1	SER	-	expression tag	UNP Q8XAY0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



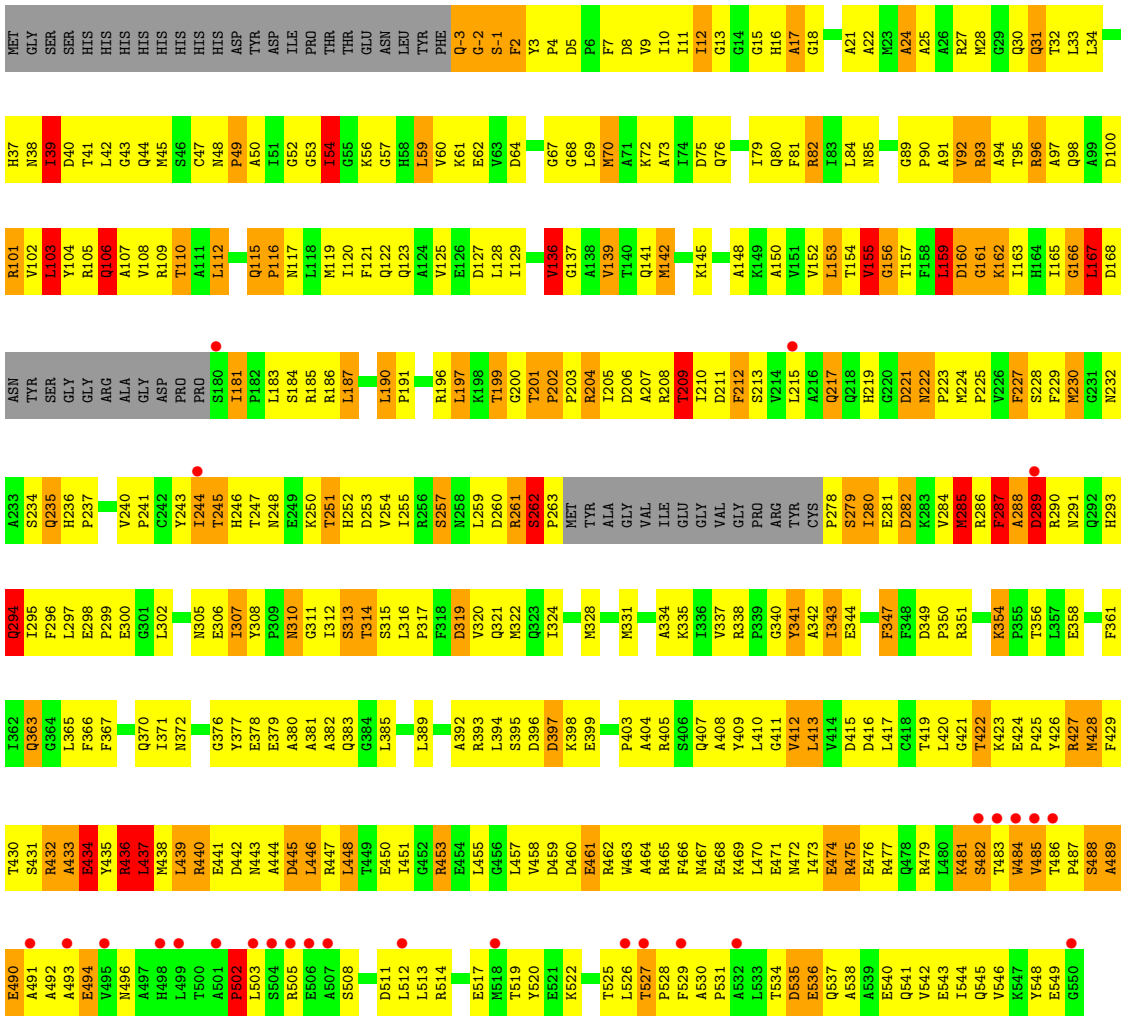
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

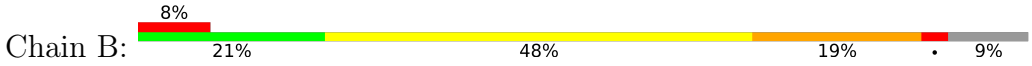
• Molecule 1: tRNA uridine 5-carboxymethylaminomethyl modification enzyme mnmG

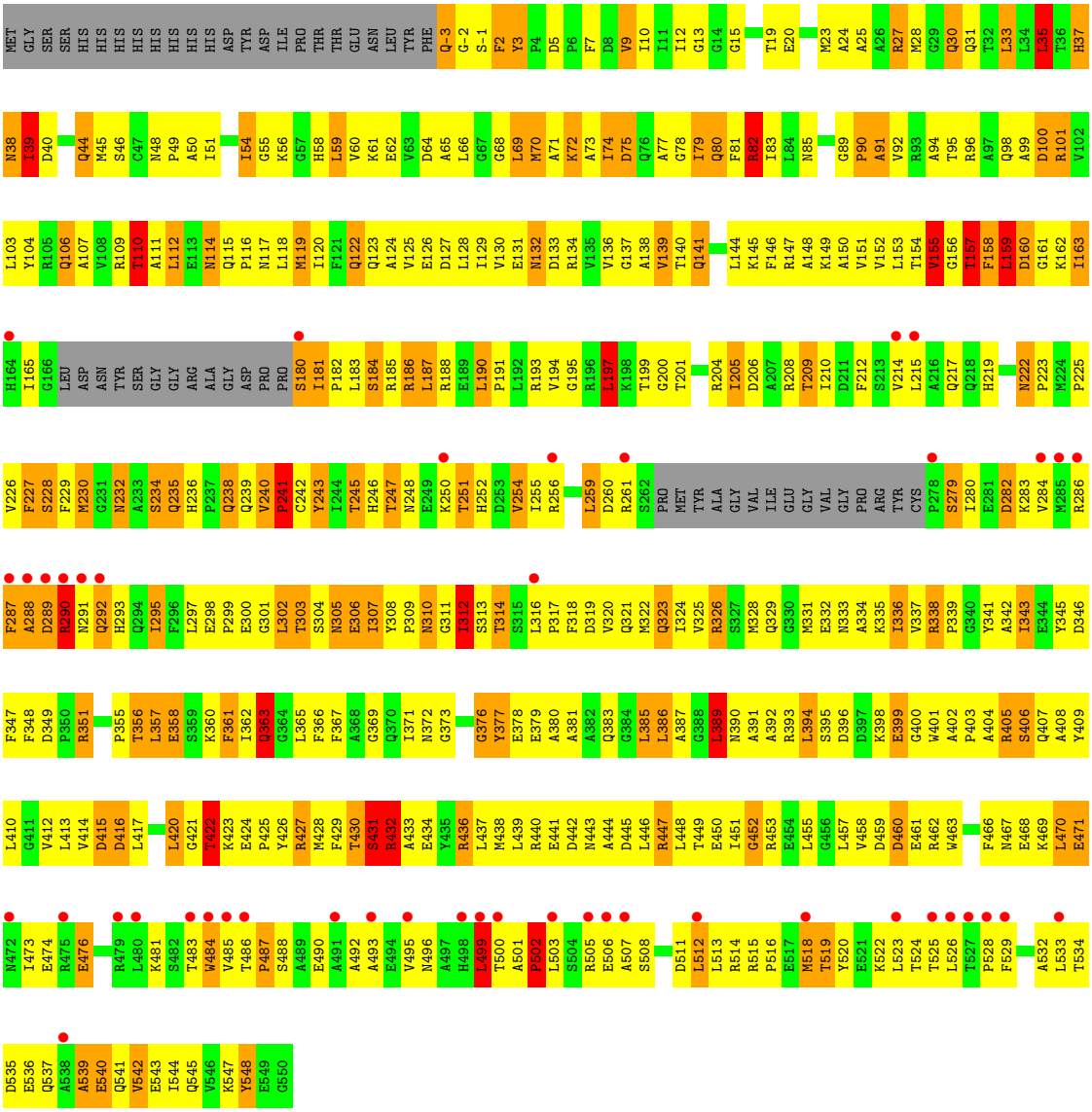
Chain A:



• Molecule 1: tRNA uridine 5-carboxymethylaminomethyl modification enzyme mnmG

Chain B:




MET GLY SER SER HIS HIS HIS HIS HIS HIS HIS ASP TYR ASP ILE PRO THR THR GLU ASN LEU TYR PHE Q-3 G-2 S-1 F2 Y3 P4 D5 P6 F7 D8 V9 I10 I11 I12 G13 G14 G15 T19 E20 M23 A24 A25 A26 R27 M28 G29 Q30 Q31 T32 L33 L34 L35 T36 H37
N38 I39 D40 Q44 M45 S46 C47 N48 P49 A50 I51 I54 G55 K56 G57 H58 L59 V60 K61 E62 V63 D64 A65 L66 G67 G68 L69 M70 A71 K72 A73 I74 D75 Q76 A77 G78 I79 Q80 F81 R82 I83 L84 N85 G89 P90 A91 V92 R93 A94 T95 R96 A97 Q98 A99 D100 R101 V102
L103 Y104 R105 Q106 A107 V108 R109 T110 A111 L112 E113 N114 Q115 P116 N117 L118 M119 I120 F121 Q122 Q123 A124 V125 E126 D127 L128 I129 V130 E131 N132 D133 R134 V135 V136 G137 A138 V139 T140 Q141 L144 K145 F146 R147 A148 K149 A150 V151 V152 L153 T154 V155 G156 T157 F158 L159 D160 G161 K162 I163
H164 I165 G166 LEU ASP ASN TYR SER GLY GLY ARG ALA GLY ASP PRO PRO S180 I181 P182 L183 S184 R185 R186 L187 R188 E189 L190 P191 L192 R193 V194 G195 R196 L197 K198 T199 G200 T201 R204 I205 D206 A207 R208 T209 I210 D211 F212 S213 V214 L215 A216 Q217 Q218 H219 M222 P223 M224 P225
V226 F227 S228 F229 M230 G231 N232 A233 S234 Q235 H236 P237 Q238 Q239 V240 P241 C242 Y243 I244 T245 H246 T247 N248 E249 K250 T251 H252 D253 V254 I255 R256 L259 D260 R261 S262 PRO MET TYR ALA GLY VAL ILE GLU GLY VAL GLY PRO ARG TYR CYS P278 S279 I280 E281 D282 K283 V284 M285 R286
F287 A288 D289 R290 N291 Q292 H293 Q294 T295 F296 L297 E298 P299 E300 G301 L302 T303 S304 N305 E306 I307 Y308 P309 N310 G311 I312 S313 T314 S315 L316 P317 F318 D319 V320 Q321 M322 Q323 I324 V325 R326 S327 M328 Q329 G330 M331 E332 N333 A334 K335 I336 V337 R338 P339 G340 Y341 A342 I343 E344 Y345 D346
F347 F348 D349 P350 R351 P355 T356 L357 E358 S359 K360 F361 I362 Q363 G364 L365 F366 F367 A368 G369 Q370 I371 N372 G373 G376 Y377 E378 E379 A380 A381 A382 Q383 G384 L385 L386 A387 G388 L389 N390 A391 A392 R393 L394 S395 D396 D397 K398 E399 G400 W401 A402 P403 A404 R405 S406 Q407 A408 Y409
L410 G411 V412 L413 V414 D415 D416 L417 L420 G421 T422 K423 E424 P425 Y426 R427 M428 F429 T430 S431 R432 A433 E434 Y435 R436 L437 M438 L439 R440 E441 D442 N443 A444 D445 L446 R447 L448 T449 E450 I451 G452 R453 E454 L455 G456 L457 V458 D459 D460 E461 R462 W463 F466 N467 E468 K469 L470 E471
N472 I473 E474 R475 E476 R479 L480 K481 S482 T483 W484 V485 T486 P487 S488 A489 E490 A491 A492 A493 E494 V495 N496 A497 H498 L499 T500 A501 P502 L503 S504 R505 E506 A507 S508 D511 L512 L513 R514 R515 P516 E517 M518 T519 Y520 E521 K522 L523 T524 T525 L526 T527 P528 F529 A532 L533 T534
D535 E536 Q537 A538 A539 E540 Q541 V542 E543 I544 Q545 V546 K547 Y548 E549 G550

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	144.59Å 144.59Å 271.02Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.49 49.74 – 3.49	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-3.49) 99.7 (49.74-3.49)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.54 (at 3.48Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.227 , 0.265 0.224 , 0.260	Depositor DCC
R_{free} test set	2137 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	91.0	Xtriage
Anisotropy	0.300	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 83.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	8156	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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, 0.333 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: 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	A	0.84	1/4146 (0.0%)	1.15	20/5615 (0.4%)
1	B	0.75	0/4134	1.11	15/5596 (0.3%)
All	All	0.80	1/8280 (0.0%)	1.13	35/11211 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	7
1	B	0	4
All	All	1	11

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	505	ARG	CZ-NH2	24.30	1.64	1.33

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	101	ARG	NE-CZ-NH1	-11.62	114.49	120.30
1	B	290	ARG	N-CA-C	-10.74	82.01	111.00
1	A	505	ARG	NE-CZ-NH1	9.52	125.06	120.30
1	B	100	ASP	CB-CG-OD1	7.85	125.36	118.30
1	A	505	ARG	NH1-CZ-NH2	-7.68	110.95	119.40

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	291	ASN	CA

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	-1	SER	Mainchain,Peptide
1	A	159	LEU	Peptide
1	A	261	ARG	Peptide
1	A	287	PHE	Peptide
1	A	288	ALA	Peptide

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	4073	0	4026	563	0
1	B	4063	0	4029	581	0
2	A	10	0	0	1	0
2	B	10	0	0	0	0
All	All	8156	0	8055	1141	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 70.

The worst 5 of 1141 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:A:289:ASP:CA	1:A:291:ASN:HA	1.56	1.32
1:B:180:SER:O	1:B:181:ILE:HG22	1.24	1.31
1:B:154:THR:O	1:B:155:VAL:CG1	1.79	1.29
1:A:285:MET:HA	1:A:285:MET:CE	1.59	1.26
1:A:289:ASP:C	1:A:291:ASN:HA	1.56	1.26

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
1	A	521/576 (90%)	383 (74%)	88 (17%)	50 (10%)	0 8
1	B	518/576 (90%)	363 (70%)	98 (19%)	57 (11%)	0 6
All	All	1039/1152 (90%)	746 (72%)	186 (18%)	107 (10%)	0 7

5 of 107 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	15	GLY
1	A	142	MET
1	A	160	ASP
1	A	167	LEU
1	A	212	PHE

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	426/474 (90%)	338 (79%)	88 (21%)	1 6
1	B	427/474 (90%)	320 (75%)	107 (25%)	0 3
All	All	853/948 (90%)	658 (77%)	195 (23%)	1 4

5 of 195 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	141	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	251	THR
1	B	163	ILE
1	B	209	THR
1	B	295	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	239	GLN
1	B	390	ASN
1	B	252	HIS
1	B	310	ASN
1	A	363	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides 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	$\#\|Z\| > 2$
2	SO4	A	551	-	4,4,4	0.17	0	6,6,6	0.52	0
2	SO4	A	552	-	4,4,4	0.17	0	6,6,6	0.59	0
2	SO4	B	551	-	4,4,4	0.19	0	6,6,6	0.33	0
2	SO4	B	1	-	4,4,4	0.21	0	6,6,6	0.54	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	551	SO4	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^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	527/576 (91%)	0.33	27 (5%)	28	25	53, 73, 123, 136	0
1	B	524/576 (90%)	0.53	46 (8%)	10	10	55, 79, 130, 137	0
All	All	1051/1152 (91%)	0.43	73 (6%)	16	16	53, 75, 128, 137	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	507	ALA	5.8
1	A	503	LEU	5.8
1	B	289	ASP	5.5
1	A	485	VAL	5.5
1	B	164	HIS	5.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	SO4	A	552	5/5	0.93	0.20	79,79,79,80	0
2	SO4	B	551	5/5	0.93	0.20	85,85,85,86	0
2	SO4	B	1	5/5	0.96	0.25	77,78,78,78	0
2	SO4	A	551	5/5	0.98	0.17	71,72,72,72	0

6.5 Other polymers (i)

There are no such residues in this entry.