WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2023 – 12:38 AM JST

PDB ID	:	7XLY
Title	:	Crystal structure of FadA2 (Rv0243) from the fatty acid metabolic pathway of Mycobacterium tuberculosis
Authors	:	Singh, R.; Kundu, P.; Singh, B.K.; Bhattacharyya, S.; Das, A.K.
Deposited on	:	2022-04-23
Resolution	:	2.90 Å(reported)

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

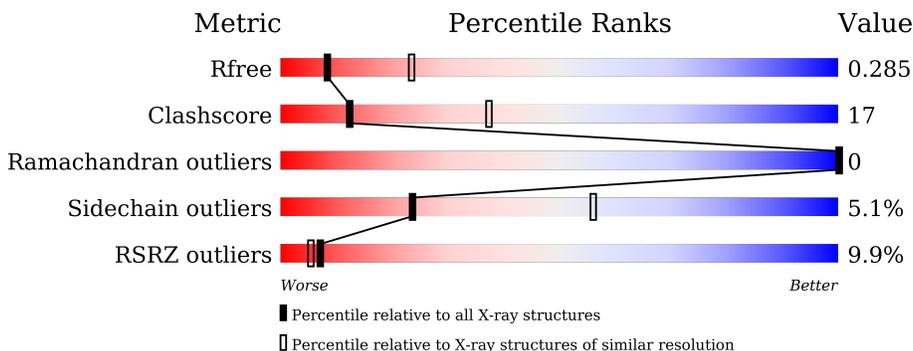|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.90 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	Percentile Ranks	Value
Rfree		0.285
Clashscore		17
Ramachandran outliers		0
Sidechain outliers		5.1%
RSRZ outliers		9.9%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	440	11% 65% 24% • 10%
1	B	440	8% 68% 21% • 9%
1	C	440	6% 62% 27% • 10%
1	D	440	11% 59% 28% • 11%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 22012 atoms, of which 10736 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 Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (Beta-ketothiolase).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	C	396	Total	C	H	N	O	S	0	0	0
			5510	1762	2691	488	557	12			
1	A	397	Total	C	H	N	O	S	0	0	0
			5589	1779	2738	499	561	12			
1	B	400	Total	C	H	N	O	S	0	0	0
			5594	1780	2738	499	565	12			
1	D	392	Total	C	H	N	O	S	0	0	0
			5314	1713	2569	480	541	11			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
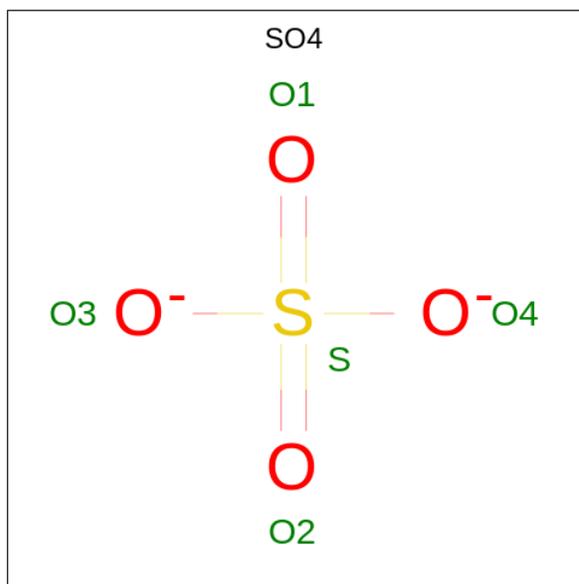
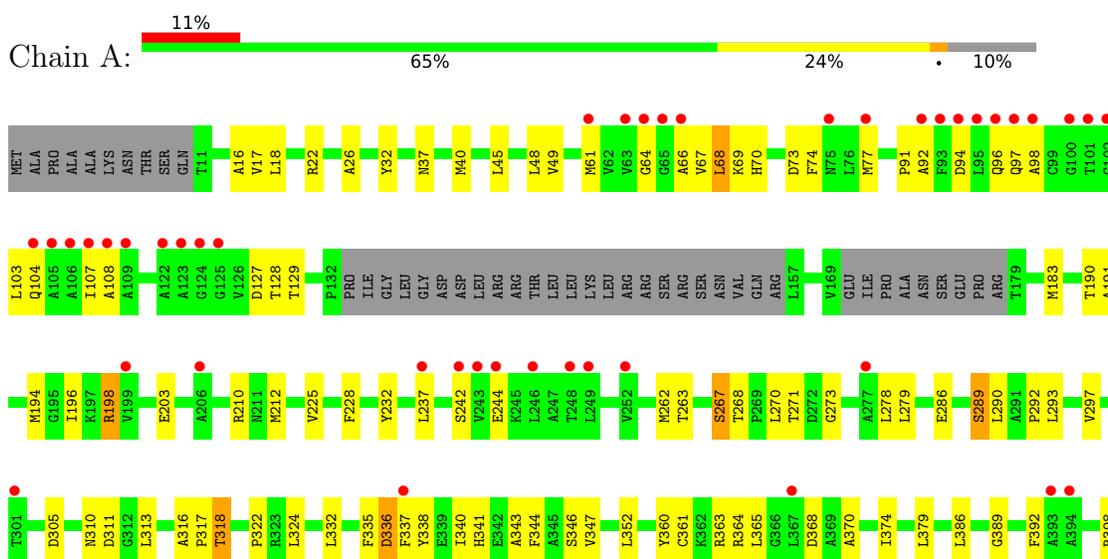| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	S	0	0
			5	4	1		

3 Residue-property plots ⓘ

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: Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (Beta-ketothiolase)

• Molecule 1: Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (Beta-ketothiolase)

I399 Q402 T403 A404 K405 Q406 L407 A408 K414 K415 G416 G417 G418 P419 L420 R421 G422 L423 I424 S425 I426 C427 G431 I437 A440
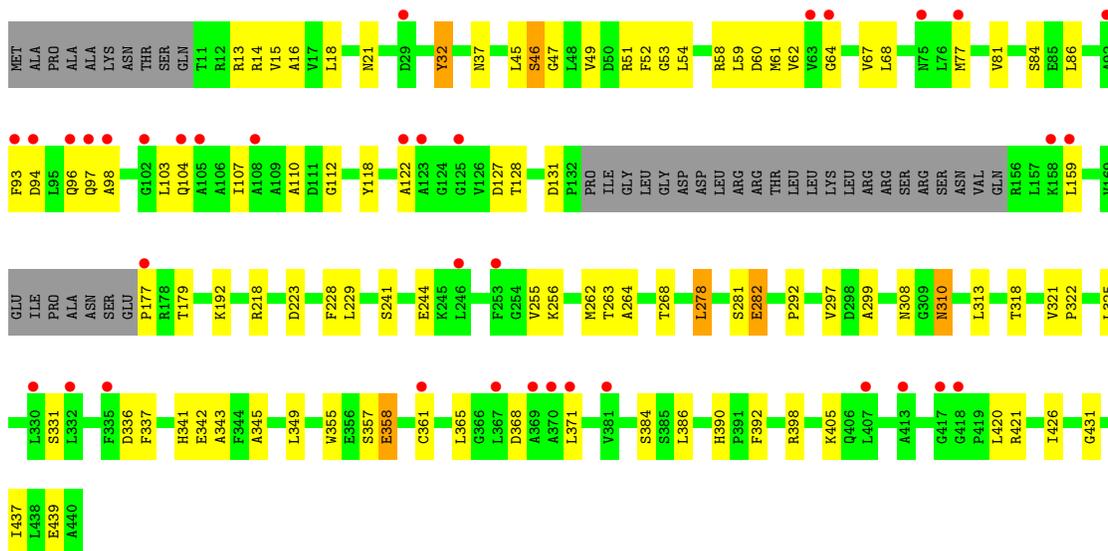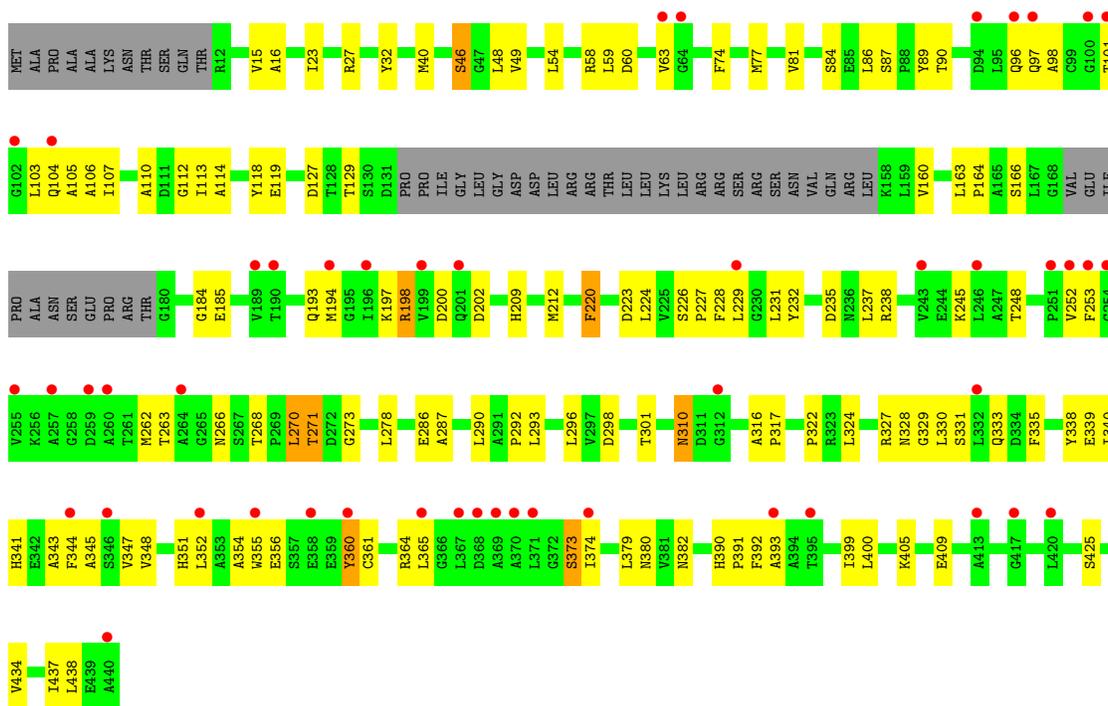
• Molecule 1: Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (Beta-ketothiolase)

Chain B:

MET ALA PRO ALA LYS ASN THR SER GLN T11 R12 R13 R14 V15 A16 V17 L18 N21 D29 Y32 N37 L45 S46 G47 L48 V49 D50 R51 F52 G53 L54 R58 L59 D60 M61 V62 V63 G64 V67 L68 N75 L76 M77 V81 S84 E85 L86 A92

F93 D94 L95 Q96 Q97 A98 G102 L103 Q104 A105 A106 I107 A108 A109 A110 D111 G112 Y118 A122 A123 G124 G125 V126 D127 T128 D131 P132 PRO ILE GLY LEU ASP ASP LEU ARG ARG THR LEU LYS LEU ARG ARG SER ARG ASN VAL GLN R156 L157 K158 L159 V169

GLU ILE PRO ALA ASN SER GLU P177 R178 T179 K192 R218 D223 F228 L229 S241 E244 K245 L246 F253 G254 V255 K256 M262 T263 A264 T268 L278 S281 E282 P292 V297 D298 A299 N308 G309 N310 L313 T318 V321 P322 L325

L330 S331 L332 F335 D336 F337 H341 E342 A343 F344 A345 L349 W355 E356 S357 E358 C361 L365 G366 L367 D368 A369 A370 L371 V381 S384 S385 L386 H390 P391 F392 R398 K405 Q406 L407 A413 G417 G418 P419 L420 R421 I426 G431

I437 L438 E439 A440

• Molecule 1: Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (Beta-ketothiolase)

Chain D:

MET ALA PRO ALA LYS ASN THR SER GLN THR R12 V15 A16 I23 R27 Y32 M40 S46 G47 L48 V49 L54 R58 L59 D60 V63 G64 F74 M77 V81 S84 E85 L86 S87 P88 Y89 T90 D94 L95 Q96 Q97 A98 C99 G100 T101

G102 L103 Q104 A105 A106 I107 A110 D111 G112 I113 A114 Y118 E119 D127 T128 T129 S130 D131 PRO PRO ILE GLY LEU GLY ASP LEU ARG ARG THR LEU LEU LYS LEU ARG ARG SER ARG SER ASN VAL GLN ARG LEU K158 L159 V160 L163 P164 A165 S166 L167 G168 VAL GLU ILE

PRO ALA ASN SER GLU PRO ARG THR G180 G184 E185 V189 T190 Q193 M194 G195 I196 K197 R198 V199 D200 Q201 D202 H209 M212 F220 D223 L224 V225 S226 P227 F228 L229 G230 L231 Y232 D235 N236 L237 R238 V243 E244 K245 L246 A247 T248 P251 V252 F253 G254

V255 K256 A257 G258 D259 A260 T261 M262 T263 A264 G265 N266 S267 T268 P269 L270 T271 D272 G273 L278 E286 A287 L290 A291 P292 L293 L296 V297 D298 T301 N310 D311 G312 A316 P317 P322 R323 L324 R327 N328 G329 L330 S331 L332 Q333 D334 F335 Y338 E339 I340

H341 E342 A343 F344 A345 S346 V347 V348 H351 L352 A353 A354 W355 E356 S357 E358 E359 Y360 C361 R364 L365 G366 L367 D368 A369 A370 L371 G372 S373 I374 L379 N380 V381 N382 H390 P391 F392 A393 A394 T395 I399 L400 K405 E409 A413 G417 L420 S425

V434 I437 L438 E439 A440

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	109.78Å 109.78Å 353.11Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.54 – 2.90 47.54 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.2 (47.54-2.90) 99.9 (47.54-2.90)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.21 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.245 , 0.285 0.245 , 0.285	Depositor DCC
R_{free} test set	2702 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	70.0	Xtriage
Anisotropy	0.915	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 75.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.020 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	22012	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.50% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CSX, 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	0.29	0/2887	0.56	0/3920
1	B	0.29	0/2892	0.56	1/3929 (0.0%)
1	C	0.30	0/2855	0.57	0/3880
1	D	0.28	0/2780	0.55	0/3788
All	All	0.29	0/11414	0.56	1/15517 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	177	PRO	N-CA-CB	5.84	110.31	103.30

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	2851	2738	2763	86	1
1	B	2856	2738	2745	79	0
1	C	2819	2691	2707	106	0
1	D	2745	2569	2571	108	1
2	B	5	0	0	0	0
All	All	11276	10736	10786	372	1

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 17.

The worst 5 of 372 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:D:356:GLU:OE2	1:D:373:SER:OG	1.61	1.18
1:C:191:ALA:HB1	1:C:196:ILE:HD11	1.37	1.03
1:D:310:ASN:O	1:D:364:ARG:NH2	1.92	1.03
1:A:212:MET:HE3	1:A:237:LEU:HD13	1.42	1.00
1:C:244:GLU:N	1:C:244:GLU:OE1	1.97	0.97

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:232:TYR:OH	1:D:409:GLU:OE1[6_664]	1.95	0.25

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	390/440 (89%)	376 (96%)	14 (4%)	0	100	100
1	B	393/440 (89%)	375 (95%)	18 (5%)	0	100	100
1	C	389/440 (88%)	370 (95%)	19 (5%)	0	100	100
1	D	385/440 (88%)	370 (96%)	15 (4%)	0	100	100
All	All	1557/1760 (88%)	1491 (96%)	66 (4%)	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	274/323 (85%)	262 (96%)	12 (4%)	28	61
1	B	272/323 (84%)	260 (96%)	12 (4%)	28	61
1	C	267/323 (83%)	252 (94%)	15 (6%)	21	52
1	D	248/323 (77%)	233 (94%)	15 (6%)	19	49
All	All	1061/1292 (82%)	1007 (95%)	54 (5%)	24	56

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	46	SER
1	B	310	ASN
1	D	271	THR
1	B	96	GLN
1	B	262	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	341	HIS
1	A	382	ASN
1	D	390	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	CSX	D	99	1	3,6,7	0.89	0	1,6,8	1.06	0
1	CSX	C	99	1	3,6,7	0.92	0	1,6,8	0.07	0
1	CSX	A	99	1	3,6,7	0.93	0	1,6,8	0.53	0
1	CSX	B	99	1	3,6,7	0.93	0	1,6,8	0.05	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
1	CSX	D	99	1	-	1/1/5/7	-
1	CSX	C	99	1	-	1/1/5/7	-
1	CSX	A	99	1	-	0/1/5/7	-
1	CSX	B	99	1	-	1/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	99	CSX	N-CA-CB-SG
1	D	99	CSX	N-CA-CB-SG
1	B	99	CSX	N-CA-CB-SG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

5.6 Ligand geometry (i)

1 ligand is 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	B	501	-	4,4,4	0.14	0	6,6,6	0.05	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 ⓘ

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(Å^2)	Q<0.9
1	A	396/440 (90%)	0.76	48 (12%)	4	3	63, 88, 110, 134	0
1	B	399/440 (90%)	0.59	36 (9%)	9	7	64, 86, 110, 128	0
1	C	395/440 (89%)	0.61	25 (6%)	20	16	66, 87, 111, 130	0
1	D	391/440 (88%)	0.82	48 (12%)	4	3	64, 97, 126, 136	0
All	All	1581/1760 (89%)	0.69	157 (9%)	7	5	63, 88, 118, 136	0

The worst 5 of 157 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	75	ASN	4.8
1	D	371	LEU	4.7
1	D	252	VAL	4.6
1	D	260	ALA	4.1
1	D	332	LEU	4.0

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. 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	B-factors(Å^2)	Q<0.9
1	CSX	B	99	7/8	0.91	0.29	71,83,90,95	0
1	CSX	A	99	7/8	0.93	0.32	63,81,84,99	0
1	CSX	C	99	7/8	0.93	0.27	64,78,85,97	0
1	CSX	D	99	7/8	0.96	0.31	87,105,116,126	0

6.3   Carbohydrates   (i)

There are no monosaccharides 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. 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	B-factors($Å^2$)	Q<0.9
2	SO4	B	501	5/5	0.81	0.31	127,133,134,144	0

6.5   Other polymers   (i)

There are no such residues in this entry.