WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2023 – 04:02 PM EST

|		
PDB ID	:	4PI2
Title	:	Crystal structure of particulate methane monooxygenase from Methylocystis sp. ATCC 49242 (Rockwell) soaked in zinc
Authors	:	Sirajuddin, S.; Rosenzweig, A.C.
Deposited on	:	2014-05-07
Resolution	:	3.33 Å(reported)

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 3.33 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.
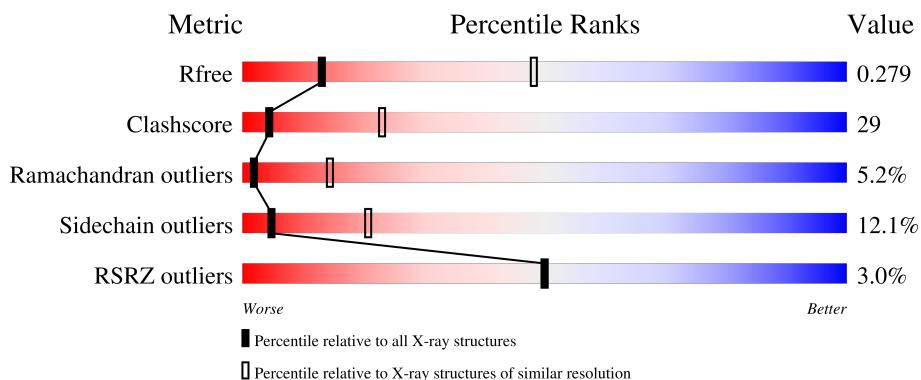
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1060 (3.38-3.30)
Clashscore	141614	1111 (3.38-3.30)
Ramachandran outliers	138981	1090 (3.38-3.30)
Sidechain outliers	138945	1089 (3.38-3.30)
RSRZ outliers	127900	1028 (3.38-3.30)

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	D	25	64% / 12% / 24%
1	H	25	76% / • / 20%
1	N	25	64% / 32% / •
2	B	252	55% / 37% / 5% • •
2	F	252	51% / 39% / 6% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	252	2% 46% 42% 8% ·
3	A	420	7% 46% 38% 9% · 7%
3	E	420	.% 47% 35% 9% · 7%
3	I	420	3% 46% 37% 9% · 7%
4	C	256	5% 39% 37% 12% · 11%
4	G	256	2% 38% 39% 11% · 11%
4	K	256	4% 34% 44% 11% · 11%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 20979 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 unknown peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	D	19	Total 95	C 57	N 19	O 19	0	0	0
1	H	20	Total 100	C 60	N 20	O 20	0	0	0
1	N	25	Total 125	C 75	N 25	O 25	0	0	0

- Molecule 2 is a protein called Particulate methane monooxygenase subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	244	Total 1974	C 1336	N 311	O 316	S 11	0	0	0
2	J	244	Total 1974	C 1336	N 311	O 316	S 11	0	0	0
2	B	244	Total 1974	C 1336	N 311	O 316	S 11	0	0	0

- Molecule 3 is a protein called Particulate methane monooxygenase subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	390	Total 3038	C 1952	N 523	O 559	S 4	0	0	0
3	A	390	Total 3038	C 1952	N 523	O 559	S 4	0	0	0
3	I	390	Total 3038	C 1952	N 523	O 559	S 4	0	0	0

- Molecule 4 is a protein called Particulate methane monooxygenase subunit C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	228	Total 1870	C 1256	N 295	O 310	S 9	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	K	228	1870	1256	295	310	9	0	0	0
			Total	C	N	O	S			
4	C	228	1870	1256	295	310	9	0	0	0

- Molecule 5 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
5	E	1	1	1	0	0
			Total	Cu		
5	A	1	1	1	0	0
			Total	Cu		
5	I	1	1	1	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
6	G	2	2	2	0	0
			Total	Zn		
6	K	1	1	1	0	0
			Total	Zn		
6	C	2	2	2	0	0

- Molecule 7 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$).
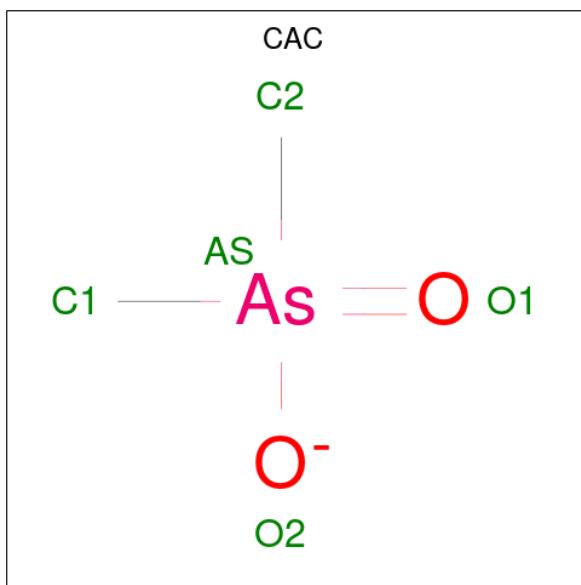
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	As	C	O		
7	C	1	5	1	2	2	0	0

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: unknown peptide
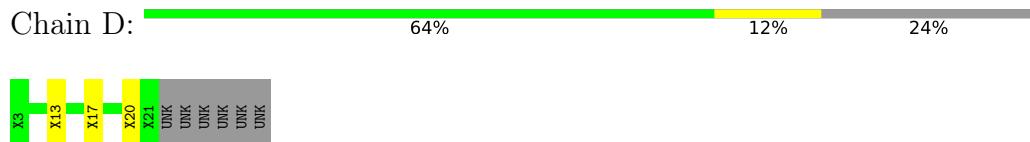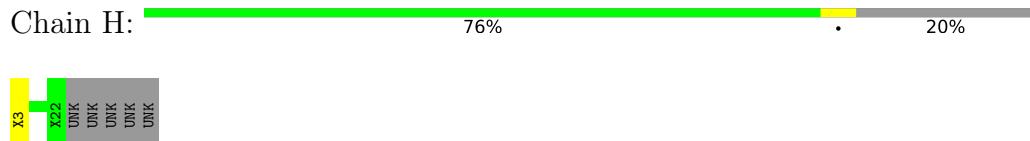
Chain D:

• Molecule 1: unknown peptide

Chain H:

• Molecule 1: unknown peptide
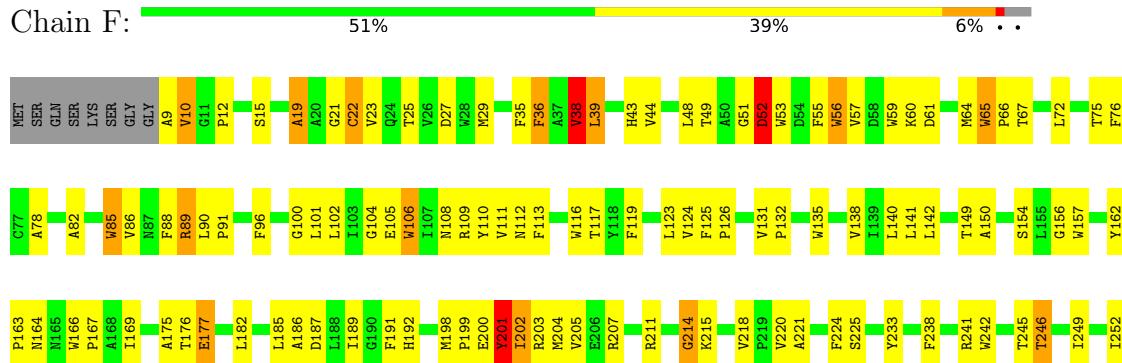
Chain N:

• Molecule 2: Particulate methane monooxygenase subunit A

Chain F:

• Molecule 2: Particulate methane monooxygenase subunit A
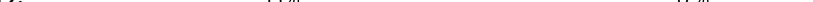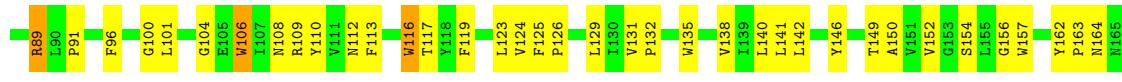
Chain J:

• Molecule 2: Particulate methane monooxygenase subunit A
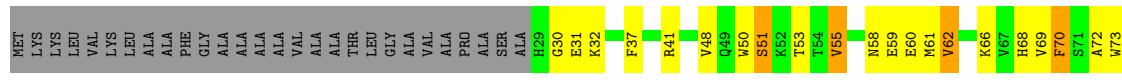
• Molecule 3: Particulate methane monooxygenase subunit B

• Molecule 3: Particulate methane monooxygenase subunit B
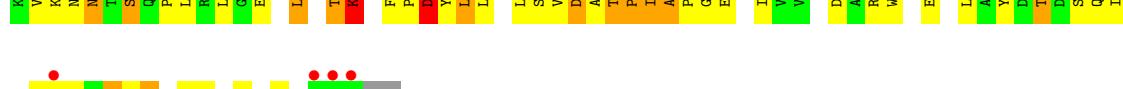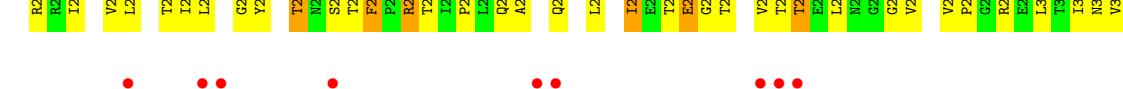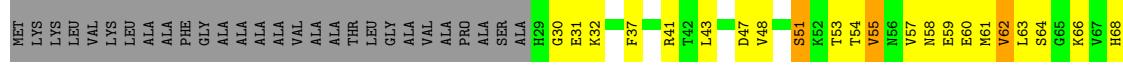
Chain A:

• Molecule 3: Particulate methane monooxygenase subunit B

Chain I:

• Molecule 4: Particulate methane monooxygenase subunit C

Chain G:


• Molecule 4: Particulate methane monooxygenase subunit C

Chain K:
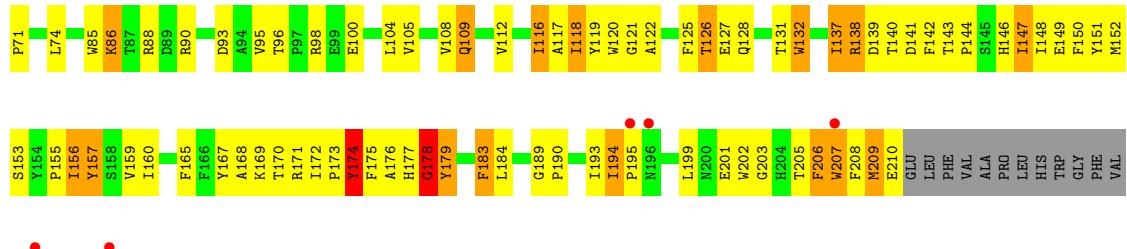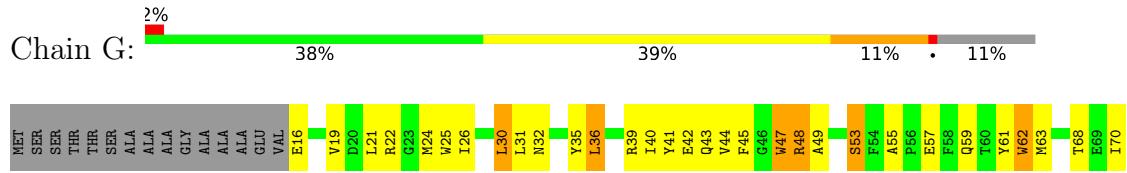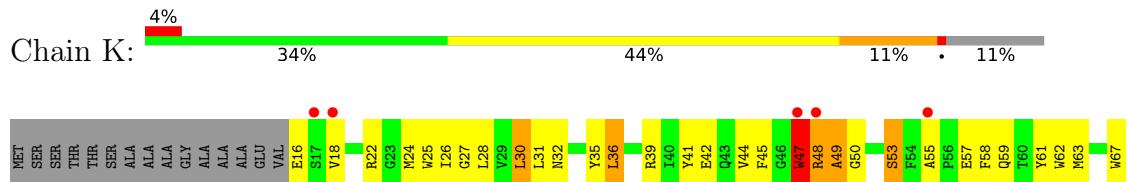
• Molecule 4: Particulate methane monooxygenase subunit C

Chain C:
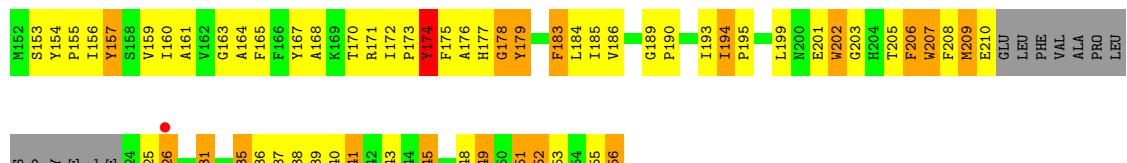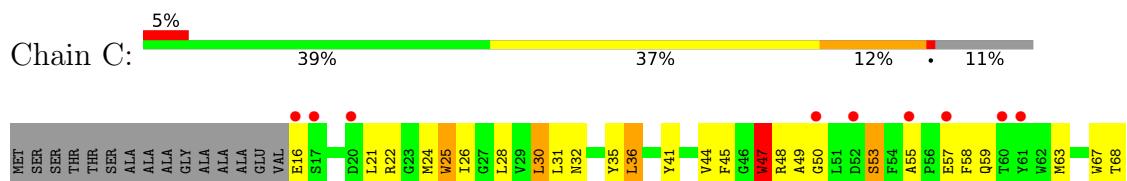
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	120.86Å 185.45Å 192.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.33 48.97 – 3.33	Depositor EDS
% Data completeness (in resolution range)	80.7 (50.00-3.33) 80.9 (48.97-3.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	8.67 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.225 , 0.282 0.223 , 0.279	Depositor DCC
R_{free} test set	2599 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	60.8	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 40.4	EDS
L-test for twinning [2]	$<|L|> = 0.44$, $<L^2> = 0.27$	Xtriage
Estimated twinning fraction	0.029 for -h,l,k	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	20979	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% 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: CU, CAC, ZN

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
2	B	0.84	6/2052 (0.3%)	0.82	0/2814
2	F	0.91	5/2052 (0.2%)	0.87	1/2814 (0.0%)
2	J	0.91	6/2052 (0.3%)	0.88	0/2814
3	A	0.67	4/3115 (0.1%)	0.82	1/4243 (0.0%)
3	E	0.72	1/3115 (0.0%)	0.84	1/4243 (0.0%)
3	I	0.70	3/3115 (0.1%)	0.85	3/4243 (0.1%)
4	C	0.80	7/1932 (0.4%)	0.80	1/2634 (0.0%)
4	G	0.80	5/1932 (0.3%)	0.85	1/2634 (0.0%)
4	K	0.81	6/1932 (0.3%)	0.89	2/2634 (0.1%)
All	All	0.79	43/21297 (0.2%)	0.84	10/29073 (0.0%)

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	N	0	1
3	A	0	1
3	E	0	2
3	I	0	1
4	C	0	1
4	G	0	1
4	K	0	1
All	All	0	8

The worst 5 of 43 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	106	TRP	CD2-CE2	6.74	1.49	1.41
4	K	132	TRP	CD2-CE2	6.64	1.49	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	132	TRP	CD2-CE2	6.39	1.49	1.41
3	A	372	TRP	CD2-CE2	5.97	1.48	1.41
2	B	106	TRP	CD2-CE2	5.95	1.48	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	I	96	ARG	NE-CZ-NH1	-7.62	116.49	120.30
4	G	138	ARG	NE-CZ-NH2	7.53	124.07	120.30
4	K	98	ARG	NE-CZ-NH1	6.15	123.38	120.30
3	I	96	ARG	NE-CZ-NH2	6.01	123.30	120.30
4	K	102	ARG	NE-CZ-NH2	5.60	123.10	120.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	340	ASP	Peptide
3	E	340	ASP	Peptide
3	E	341	TYR	Peptide
3	I	340	ASP	Peptide
1	N	3	UNK	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	D	95	0	21	4	0
1	H	100	0	23	1	0
1	N	125	0	28	10	0
2	B	1974	0	1932	126	0
2	F	1974	0	1932	126	0
2	J	1974	0	1932	141	0
3	A	3038	0	3022	186	0
3	E	3038	0	3022	191	0
3	I	3038	0	3022	192	0
4	C	1870	0	1854	120	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	G	1870	0	1854	127	0
4	K	1870	0	1854	133	0
5	A	1	0	0	0	0
5	E	1	0	0	0	0
5	I	1	0	0	0	0
6	C	2	0	0	0	0
6	G	2	0	0	0	0
6	K	1	0	0	0	0
7	C	5	0	0	0	0
All	All	20979	0	20496	1196	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 29.

The worst 5 of 1196 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)
2:B:51:GLY:HA3	2:B:52:ASP:CB	1.70	1.21
3:A:334:THR:HA	3:A:335:LYS:HB2	1.24	1.18
2:J:51:GLY:HA3	2:J:52:ASP:CB	1.73	1.18
2:F:51:GLY:HA3	2:F:52:ASP:CB	1.75	1.16
2:B:51:GLY:CA	2:B:52:ASP:HB2	1.75	1.16

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	
2	B	242/252 (96%)	189 (78%)	48 (20%)	5 (2%)	7	34
2	F	242/252 (96%)	197 (81%)	36 (15%)	9 (4%)	3	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	J	242/252 (96%)	200 (83%)	29 (12%)	13 (5%)	2	13
3	A	388/420 (92%)	307 (79%)	56 (14%)	25 (6%)	1	10
3	E	388/420 (92%)	314 (81%)	44 (11%)	30 (8%)	1	7
3	I	388/420 (92%)	308 (79%)	53 (14%)	27 (7%)	1	9
4	C	224/256 (88%)	169 (75%)	49 (22%)	6 (3%)	5	28
4	G	224/256 (88%)	172 (77%)	44 (20%)	8 (4%)	3	23
4	K	224/256 (88%)	163 (73%)	52 (23%)	9 (4%)	3	20
All	All	2562/2784 (92%)	2019 (79%)	411 (16%)	132 (5%)	2	14

5 of 132 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	52	ASP
3	E	31	GLU
3	E	51	SER
3	E	141	GLU
3	E	164	ASP

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	
2	B	202/208 (97%)	186 (92%)	16 (8%)	12	40
2	F	202/208 (97%)	186 (92%)	16 (8%)	12	40
2	J	202/208 (97%)	186 (92%)	16 (8%)	12	40
3	A	320/336 (95%)	281 (88%)	39 (12%)	5	20
3	E	320/336 (95%)	283 (88%)	37 (12%)	5	22
3	I	320/336 (95%)	280 (88%)	40 (12%)	4	19
4	C	194/213 (91%)	162 (84%)	32 (16%)	2	10
4	G	194/213 (91%)	161 (83%)	33 (17%)	2	9
4	K	194/213 (91%)	163 (84%)	31 (16%)	2	11

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2148/2271 (95%)	1888 (88%)	260 (12%)	5 21

5 of 260 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	C	47	TRP
4	C	118	ILE
3	I	64	SER
3	I	61	MET
4	C	174	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
3	I	261	GLN
2	B	192	HIS
3	I	265	GLN
2	J	178	GLN
4	K	109	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 8 are monoatomic - leaving 1 for Mogul analysis.

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	$\#|Z| > 2$
7	CAC	C	303	6	0,4,4	-	-	0,6,6	-	-

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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

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

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	D	0/25	-	-			-	-
1	H	0/25	-	-			-	-
1	N	0/25	-	-			-	-
2	B	244/252 (96%)	-0.54	1 (0%)	92	94	27, 51, 92, 117	0
2	F	244/252 (96%)	-0.63	0	100	100	14, 31, 66, 95	0
2	J	244/252 (96%)	-0.58	4 (1%)	72	71	16, 33, 67, 104	0
3	A	390/420 (92%)	0.14	28 (7%)	15	16	35, 72, 112, 153	0
3	E	390/420 (92%)	-0.29	3 (0%)	86	87	24, 47, 85, 129	0
3	I	390/420 (92%)	-0.13	14 (3%)	42	41	25, 64, 117, 179	0
4	C	228/256 (89%)	-0.20	13 (5%)	23	25	31, 59, 111, 148	0
4	G	228/256 (89%)	-0.30	5 (2%)	62	61	20, 48, 101, 153	0
4	K	228/256 (89%)	-0.17	10 (4%)	34	35	24, 47, 115, 160	0
All	All	2586/2859 (90%)	-0.27	78 (3%)	50	50	14, 52, 107, 179	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	418	ASP	6.1
4	K	253	LEU	5.3
3	A	417	GLY	5.0
4	C	17	SER	4.6
4	K	254	LEU	4.2

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 ⓘ

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, 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($\mathring{A}^2$)	Q<0.9
6	ZN	G	302	1/1	0.93	0.04	153,153,153,153	0
5	CU	A	501	1/1	0.97	0.04	85,85,85,85	0
7	CAC	C	303	5/5	0.97	0.12	92,95,102,107	0
6	ZN	C	302	1/1	0.98	0.04	106,106,106,106	0
5	CU	E	501	1/1	0.99	0.05	53,53,53,53	0
6	ZN	K	301	1/1	0.99	0.03	67,67,67,67	0
6	ZN	C	301	1/1	0.99	0.07	66,66,66,66	0
5	CU	I	501	1/1	0.99	0.05	63,63,63,63	0
6	ZN	G	301	1/1	0.99	0.09	64,64,64,64	0

6.5 Other polymers ⓘ

There are no such residues in this entry.