



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 05:26 pm GMT

PDB ID : 4N4R
Title : Structure basis of lipopolysaccharide biogenesis
Authors : Dong, H.; Xiang, Q.; Wang, Z.; Paterson, N.G.; He, C.; Zhang, Y.; Wang, W.;
Dong, C.
Deposited on : 2013-10-08
Resolution : 2.80 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

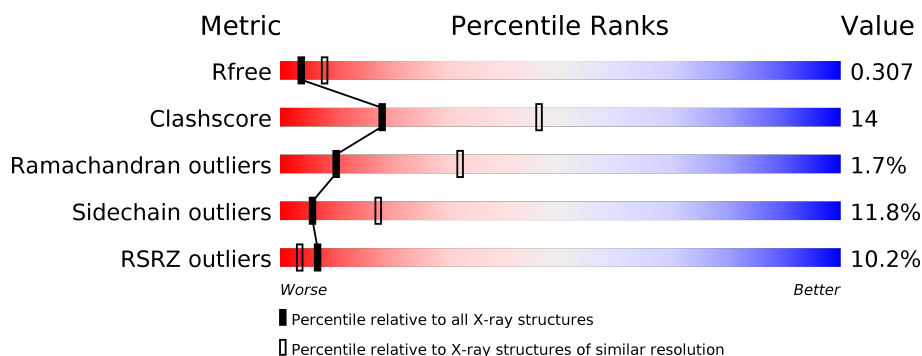
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

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	A	786	
1	C	786	
2	B	196	
2	D	196	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CAC	A	801	-	-	-	X

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 11117 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.

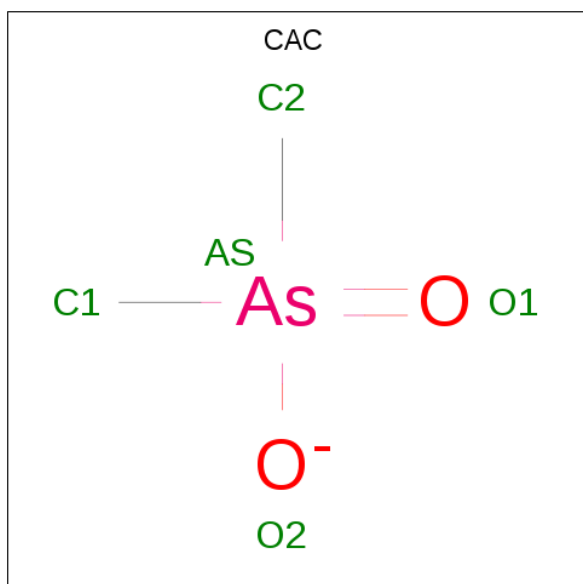
- Molecule 1 is a protein called LPS-assembly protein LptD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	532	Total	C	N	O	S	Se	0	0	0
			4370	2765	741	850	1	13			
1	C	533	Total	C	N	O	S	Se	0	0	0
			4373	2765	742	853	1	13			

- Molecule 2 is a protein called LPS-assembly lipoprotein LptE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	151	Total	C	N	O	S	Se	0	0	0
			1176	731	209	228	1	7			
2	D	151	Total	C	N	O	S	Se	0	0	0
			1176	731	209	228	1	7			

- Molecule 3 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	As	C	O	0	0
			5	1	2	2		
3	C	1	Total	As	C	O	0	0
			5	1	2	2		

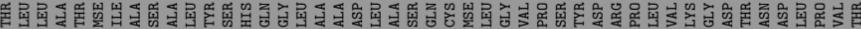
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

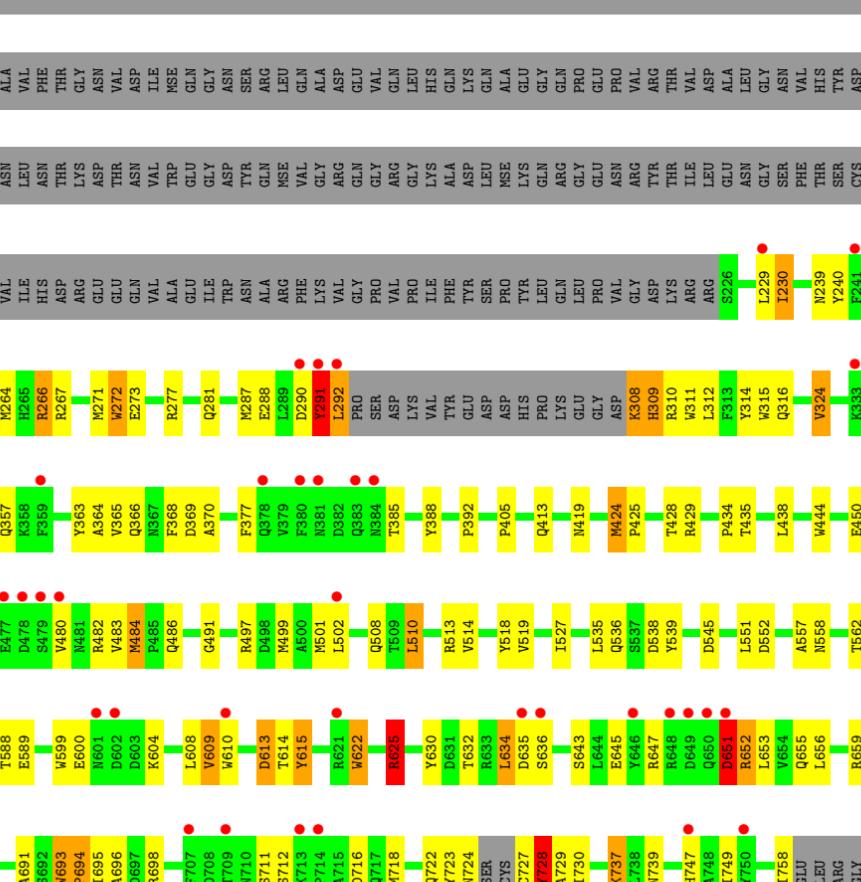
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Zn	0	0
			1	1		
4	C	1	Total	Zn	0	0
			1	1		

- Molecule 5 is water.

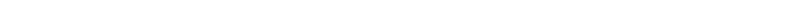
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	4	Total	O	0	0
			4	4		
5	C	5	Total	O	0	0
			5	5		
5	D	1	Total	O	0	0
			1	1		

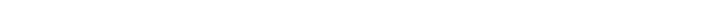
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

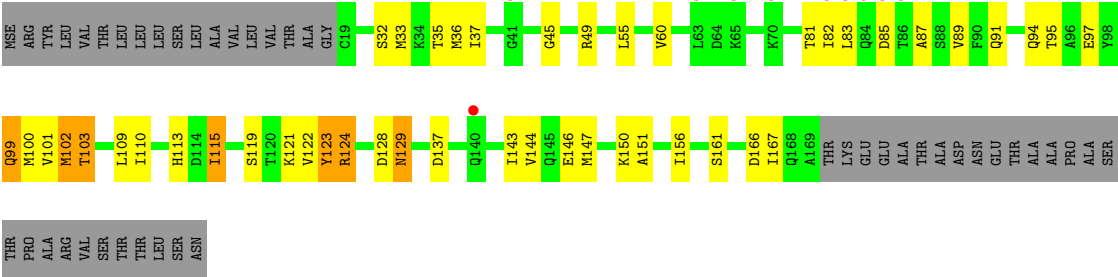
Chain A: 



Index	Label	Color
1	GLY	Yellow
2	T771	Red
3	Q772	Red
4	R776	Yellow
5	P781	Yellow
6	S785	Yellow
7	M786	Yellow
8	Q679	Yellow
9	N578	Yellow
10	V579	Green
11	S580	Green
12	N471	Yellow
13	K472	Yellow
14	N473	Green
15	I584	Yellow
16	T588	Yellow
17	E589	Yellow
18	W599	Yellow
19	E600	Yellow
20	N601	Red
21	R602	Red
22	D603	Green
23	P694	Yellow
24	L695	Yellow
25	A696	Yellow
26	D697	Yellow
27	R698	Yellow
28	F707	Red
29	L608	Yellow
30	V609	Yellow
31	W610	Yellow
32	T709	Red
33	I710	Yellow
34	S711	Yellow
35	T614	Yellow
36	K713	Red
37	P714	Red
38	A715	Yellow
39	D716	Yellow
40	Q717	Yellow
41	M718	Yellow
42	Q722	Yellow
43	Y723	Green
44	M724	Yellow
45	SER	Grey
46	CYS	Grey
47	G727	Yellow
48	Y728	Red
49	A729	Yellow
50	I730	Yellow
51	K737	Red
52	L738	Green
53	E645	Yellow
54	Y646	Red
55	R647	Yellow
56	R648	Red
57	D649	Red
58	Q650	Yellow
59	D651	Red
60	R652	Red
61	L653	Yellow
62	V654	Yellow
63	Q655	Yellow
64	L656	Yellow
65	R659	Yellow
66	T666	Yellow
67	T669	Yellow
68	SER	Grey
69	ASN	Grey
70	TYR	Grey
71	GLY	Grey
72	LEU	Grey
73	THR	Grey
74	ASP	Grey
75	THR	Grey
76	ASP	Grey
77	ASP	Grey
78	ASP	Grey
79	ASP	Grey
80	ASP	Grey
81	ASP	Grey
82	ASP	Grey
83	ASP	Grey
84	ASP	Grey
85	ASP	Grey
86	ASP	Grey
87	ASP	Grey
88	ASP	Grey
89	ASP	Grey
90	ASP	Grey
91	ASP	Grey
92	ASP	Grey
93	ASP	Grey
94	ASP	Grey
95	ASP	Grey
96	ASP	Grey
97	ASP	Grey
98	ASP	Grey
99	ASP	Grey
100	ASP	Grey
101	ASP	Grey
102	ASP	Grey
103	ASP	Grey
104	ASP	Grey
105	ASP	Grey
106	ASP	Grey
107	ASP	Grey
108	ASP	Grey
109	ASP	Grey
110	ASP	Grey
111	ASP	Grey
112	ASP	Grey
113	ASP	Grey
114	ASP	Grey
115	ASP	Grey
116	ASP	Grey
117	ASP	Grey
118	ASP	Grey
119	ASP	Grey
120	ASP	Grey
121	ASP	Grey
122	ASP	Grey
123	ASP	Grey
124	ASP	Grey
125	ASP	Grey
126	ASP	Grey
127	ASP	Grey
128	ASP	Grey
129	ASP	Grey
130	ASP	Grey
131	ASP	Grey
132	ASP	Grey
133	ASP	Grey
134	ASP	Grey
135	ASP	Grey
136	ASP	Grey
137	ASP	Grey
138	ASP	Grey
139	ASP	Grey
140	ASP	Grey
141	ASP	Grey
142	ASP	Grey
143	ASP	Grey
144	ASP	Grey
145	ASP	Grey
146	ASP	Grey
147	ASP	Grey
148	ASP	Grey
149	ASP	Grey
150	ASP	Grey
151	ASP	Grey
152	ASP	Grey

Chain C:  8% 48% 15% 32%

Chain D:  5% 54% 20% 23%



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	173.43Å 76.08Å 213.60Å 90.00° 111.52° 90.00°	Depositor
Resolution (Å)	107.49 – 2.80 43.92 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (107.49-2.80) 100.0 (43.92-2.80)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.76 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.285 , 0.305 0.286 , 0.307	Depositor DCC
R_{free} test set	3253 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	70.8	Xtriage
Anisotropy	0.493	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 64.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	11117	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	$\# Z > 5$
1	A	0.96	24/4475 (0.5%)	0.91	11/6062 (0.2%)
1	C	0.67	8/4478 (0.2%)	0.76	2/6067 (0.0%)
2	B	1.36	10/1186 (0.8%)	1.14	8/1597 (0.5%)
2	D	0.84	1/1186 (0.1%)	0.94	1/1597 (0.1%)
All	All	0.90	43/11325 (0.4%)	0.88	22/15323 (0.1%)

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	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	123	TYR	CE1-CZ	20.71	1.65	1.38
2	B	122	VAL	CA-CB	11.37	1.78	1.54
1	A	578	ASN	CA-CB	10.72	1.81	1.53
1	A	625	ARG	CG-CD	9.03	1.74	1.51
1	A	645	GLU	CD-OE2	8.83	1.35	1.25

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	659	ARG	NE-CZ-NH2	-9.54	115.53	120.30
2	B	166	ASP	CB-CG-OD1	9.41	126.77	118.30
1	A	613	ASP	CB-CG-OD2	-8.56	110.59	118.30
1	A	625	ARG	NE-CZ-NH1	8.44	124.52	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	102	MSE	CB-CA-C	7.36	125.12	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	625	ARG	Sidechain

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	4370	0	4073	124	0
1	C	4373	0	4075	117	0
2	B	1176	0	1186	55	0
2	D	1176	0	1186	37	0
3	A	5	0	0	0	0
3	C	5	0	0	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
5	A	4	0	0	0	0
5	C	5	0	0	0	0
5	D	1	0	0	0	0
All	All	11117	0	10520	311	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 14.

The worst 5 of 311 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:625:ARG:CG	1:A:625:ARG:CD	1.74	1.63
1:A:578:ASN:CA	1:A:578:ASN:CB	1.81	1.59
2:B:122:VAL:CA	2:B:122:VAL:CB	1.78	1.59
2:B:102:MSE:CG	2:B:102:MSE:SE	2.16	1.43
2:B:147:MSE:CE	2:B:147:MSE:SE	2.18	1.40

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	524/786 (67%)	470 (90%)	43 (8%)	11 (2%)	8	26
1	C	525/786 (67%)	473 (90%)	42 (8%)	10 (2%)	9	28
2	B	149/196 (76%)	140 (94%)	7 (5%)	2 (1%)	13	40
2	D	149/196 (76%)	141 (95%)	8 (5%)	0	100	100
All	All	1347/1964 (69%)	1224 (91%)	100 (7%)	23 (2%)	10	32

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	309	HIS
1	A	366	GLN
1	A	695	ILE
1	C	366	GLN
1	C	695	ILE

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	469/665 (70%)	413 (88%)	56 (12%)	6	17
1	C	470/665 (71%)	413 (88%)	57 (12%)	5	16
2	B	132/160 (82%)	117 (89%)	15 (11%)	6	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	132/160 (82%)	118 (89%)	14 (11%)	7	22
All	All	1203/1650 (73%)	1061 (88%)	142 (12%)	6	17

5 of 142 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	124	ARG
1	C	332	THR
2	D	89	VAL
2	B	129	ASN
1	C	266	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	154	GLN
1	C	281	GLN
1	C	739	ASN
1	A	724	ASN
2	B	129	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\# Z > 2$
3	CAC	A	801	-	0,4,4	0.00	-	0,6,6	0.00	-
3	CAC	C	801	-	0,4,4	0.00	-	0,6,6	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
3	CAC	A	801	-	-	0/0/0/0	0/0/0/0
3	CAC	C	801	-	-	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(Å ²)	Q<0.9
1	A	519/786 (66%)	0.38	40 (7%) 13 7	60, 111, 159, 201	0
1	C	520/786 (66%)	0.53	61 (11%) 4 2	80, 129, 182, 271	0
2	B	144/196 (73%)	0.92	24 (16%) 1 1	93, 104, 146, 163	0
2	D	144/196 (73%)	0.50	10 (6%) 17 9	70, 98, 131, 145	0
All	All	1327/1964 (67%)	0.51	135 (10%) 7 3	60, 115, 167, 271	0

The worst 5 of 135 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	748	ALA	9.2
1	C	266	ARG	8.8
1	C	749	ILE	8.3
1	C	240	TYR	8.2
1	C	267	ARG	7.3

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 [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. 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(\AA^2)	Q<0.9
3	CAC	A	801	5/5	0.75	0.82	196,198,221,228	0
3	CAC	C	801	5/5	0.84	0.78	189,193,207,217	0
4	ZN	C	802	1/1	0.88	0.05	152,152,152,152	0
4	ZN	A	802	1/1	0.93	0.09	127,127,127,127	0

6.5 Other polymers [i](#)

There are no such residues in this entry.