W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 17, 2023 – 11:21 PM EDT

PDB ID	:	4ZR6
Title	:	Lymnaea Stagnalis Acetylcholine Binding Protein in Complex with 3-[(4E)-4-[(3-methylimidazol-4-yl)methylene]-2,3-dihydropyrrol-5-yl]pyridine
Authors	:	Bobango, J.; Bryant, D.; Denton, T.T.; Talley, T.T.
Deposited on	:	2015-05-11
Resolution	:	2.60 Å(reported)

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

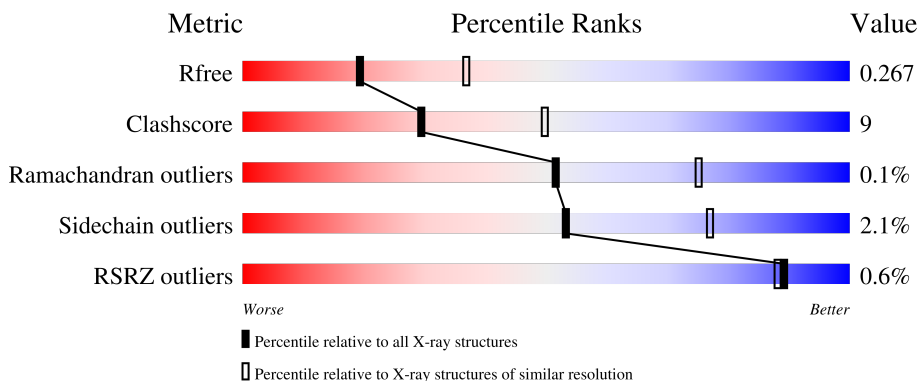|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.60 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		0.267
Clashscore		9
Ramachandran outliers		0.1%
Sidechain outliers		2.1%
RSRZ outliers		0.6%

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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	218	73% 16% 11%
1	B	218	2% 76% 12% 12%
1	C	218	72% 19% 9%
1	D	218	74% 17% • 9%
1	E	218	70% 19% • 9%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7363 atoms, of which 70 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 Acetylcholine-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	195	Total	C	N	O	S	0	0	0
			1415	902	233	275	5			
1	B	192	Total	C	N	O	S	0	0	0
			1383	886	222	270	5			
1	C	198	Total	C	N	O	S	0	0	0
			1478	938	245	290	5			
1	D	199	Total	C	N	O	S	0	0	0
			1462	928	239	290	5			
1	E	199	Total	C	N	O	S	0	0	0
			1458	923	239	291	5			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	ASP	-	expression tag	UNP P58154
A	-6	TYR	-	expression tag	UNP P58154
A	-5	LYS	-	expression tag	UNP P58154
A	-4	ASP	-	expression tag	UNP P58154
A	-3	ASP	-	expression tag	UNP P58154
A	-2	ASP	-	expression tag	UNP P58154
A	-1	ASP	-	expression tag	UNP P58154
A	0	LYS	-	expression tag	UNP P58154
B	-7	ASP	-	expression tag	UNP P58154
B	-6	TYR	-	expression tag	UNP P58154
B	-5	LYS	-	expression tag	UNP P58154
B	-4	ASP	-	expression tag	UNP P58154
B	-3	ASP	-	expression tag	UNP P58154
B	-2	ASP	-	expression tag	UNP P58154
B	-1	ASP	-	expression tag	UNP P58154
B	0	LYS	-	expression tag	UNP P58154
C	-7	ASP	-	expression tag	UNP P58154
C	-6	TYR	-	expression tag	UNP P58154
C	-5	LYS	-	expression tag	UNP P58154

Continued on next page...

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

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-4	ASP	-	expression tag	UNP P58154
C	-3	ASP	-	expression tag	UNP P58154
C	-2	ASP	-	expression tag	UNP P58154
C	-1	ASP	-	expression tag	UNP P58154
C	0	LYS	-	expression tag	UNP P58154
D	-7	ASP	-	expression tag	UNP P58154
D	-6	TYR	-	expression tag	UNP P58154
D	-5	LYS	-	expression tag	UNP P58154
D	-4	ASP	-	expression tag	UNP P58154
D	-3	ASP	-	expression tag	UNP P58154
D	-2	ASP	-	expression tag	UNP P58154
D	-1	ASP	-	expression tag	UNP P58154
D	0	LYS	-	expression tag	UNP P58154
E	-7	ASP	-	expression tag	UNP P58154
E	-6	TYR	-	expression tag	UNP P58154
E	-5	LYS	-	expression tag	UNP P58154
E	-4	ASP	-	expression tag	UNP P58154
E	-3	ASP	-	expression tag	UNP P58154
E	-2	ASP	-	expression tag	UNP P58154
E	-1	ASP	-	expression tag	UNP P58154
E	0	LYS	-	expression tag	UNP P58154

- Molecule 2 is 3-{(4E)-4-[(1-methyl-1H-imidazol-5-yl)methylidene]-3,4-dihydro-2H-pyrrol-5-yl}pyridine (three-letter code: 4QW) (formula: $C_{14}H_{14}N_4$).
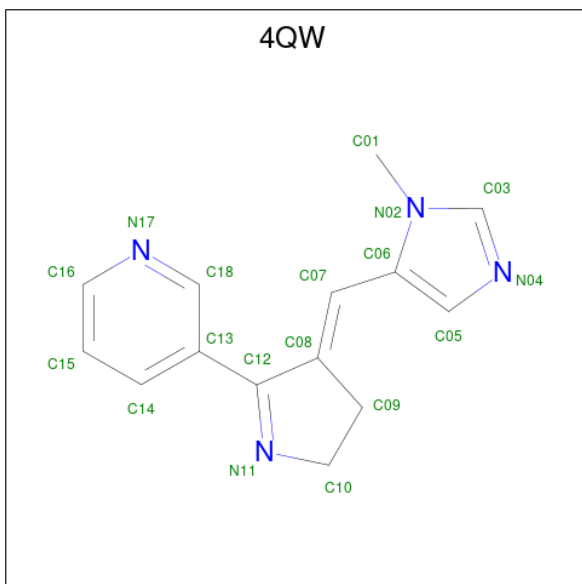
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 32	C 14	H 14	N 4	0	0
2	B	1	Total 32	C 14	H 14	N 4	0	0
2	C	1	Total 32	C 14	H 14	N 4	0	0
2	D	1	Total 32	C 14	H 14	N 4	0	0
2	E	1	Total 32	C 14	H 14	N 4	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total 2	O 2	0	0
3	C	3	Total 3	O 3	0	0
3	D	1	Total 1	O 1	0	0
3	E	1	Total 1	O 1	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.
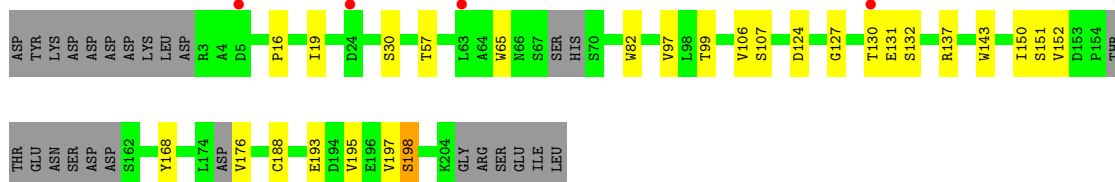
- Molecule 1: Acetylcholine-binding protein

Chain A:
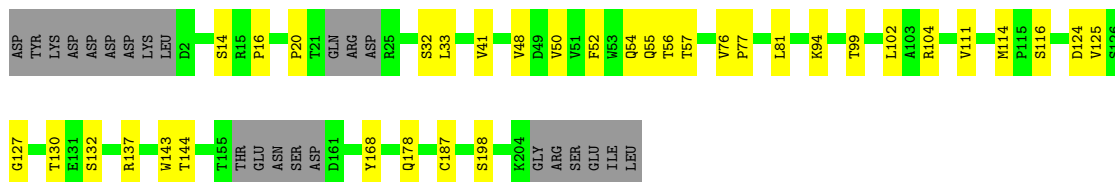
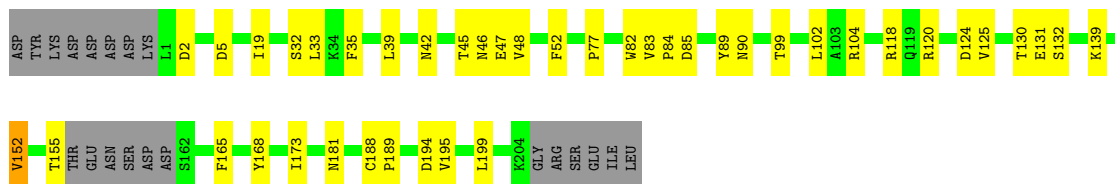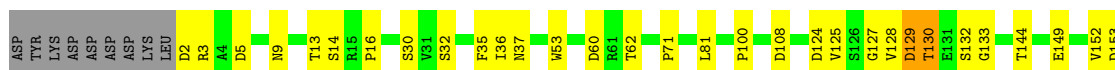- Molecule 1: Acetylcholine-binding protein

Chain B:

- Molecule 1: Acetylcholine-binding protein

Chain C:

- Molecule 1: Acetylcholine-binding protein

Chain D:

P154 T155 THR GLU ASN SER D160 E163 Y168 S169 Q178 C188 V195 L199 K204 GLY ARG SER GLU ILE LEU
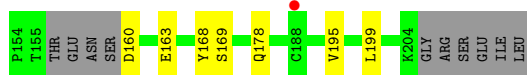
• Molecule 1: Acetylcholine-binding protein

Chain E:

70%   19%   •   9%

ASP TYR LYS ASP ASP ASP LYS LEU D2 R3 A4 I19 P20 T21 S30 F35 I36 N37 I38 V41 N42 E43 V48 D49 V50 W53 T57 W82 V83 P84 D85 L86 R104 V105 V106 S107 D108 V111 R118 Q119 R120 C123 D124 V125 S132 G133

A134 H145 E149 I150 D153 P154 T155 THR GLU ASN SER D160 F165 Y168 T177 K180 N181 S182 V183 T184 Y185 S186 C187 C188 P189 S198 L199 N200 K204 GLY ARG SER GLU ILE LEU
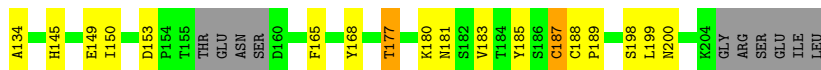
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	68.92Å 78.10Å 120.50Å 90.00° 102.32° 90.00°	Depositor
Resolution (Å)	44.26 – 2.60 47.01 – 2.51	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.26-2.60) 93.4 (47.01-2.51)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 2.51Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.210 , 0.266 0.214 , 0.267	Depositor DCC
R_{free} test set	2001 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	44.4	Xtriage
Anisotropy	0.526	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 47.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	7363	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% 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: 4QW

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	0.45	0/1447	0.59	0/1996
1	B	0.49	0/1415	0.60	0/1952
1	C	0.51	0/1513	0.63	0/2082
1	D	0.47	0/1497	0.58	0/2066
1	E	0.46	0/1491	0.59	0/2054
All	All	0.47	0/7363	0.60	0/10150

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	A	1415	0	1258	31	0
1	B	1383	0	1224	23	0
1	C	1478	0	1362	30	0
1	D	1462	0	1308	28	0
1	E	1458	0	1307	29	0
2	A	18	14	0	1	0
2	B	18	14	0	1	0
2	C	18	14	0	0	0

Continued on next page...

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

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	18	14	0	1	0
2	E	18	14	0	0	0
3	B	2	0	0	0	0
3	C	3	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	1	0
All	All	7293	70	6459	130	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 9.

The worst 5 of 130 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:41:VAL:HG13	1:A:125:VAL:HG11	1.56	0.87
1:A:130:THR:HG22	1:A:132:SER:H	1.39	0.87
1:B:30:SER:HB2	1:B:57:THR:HG22	1.62	0.81
1:E:177:THR:HG22	1:E:198:SER:HB2	1.64	0.80
1:C:152:VAL:HG23	1:C:195:VAL:HG23	1.64	0.78

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	
1	A	189/218 (87%)	183 (97%)	5 (3%)	1 (0%)	29	52
1	B	184/218 (84%)	181 (98%)	3 (2%)	0	100	100
1	C	194/218 (89%)	188 (97%)	6 (3%)	0	100	100
1	D	195/218 (89%)	193 (99%)	2 (1%)	0	100	100
1	E	195/218 (89%)	189 (97%)	6 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	957/1090 (88%)	934 (98%)	22 (2%)	1 (0%)	51 75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	20	PRO

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	143/204 (70%)	141 (99%)	2 (1%)	67 85
1	B	140/204 (69%)	138 (99%)	2 (1%)	67 85
1	C	160/204 (78%)	159 (99%)	1 (1%)	86 95
1	D	154/204 (76%)	151 (98%)	3 (2%)	57 79
1	E	153/204 (75%)	145 (95%)	8 (5%)	23 46
All	All	750/1020 (74%)	734 (98%)	16 (2%)	53 77

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	177	THR
1	E	132	SER
1	E	19	ILE
1	E	104	ARG
1	D	130	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	146	HIS
1	C	181	ASN

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 ⓘ

5 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	4QW	C	301	-	18,20,20	4.94	6 (33%)	19,27,27	2.17	4 (21%)
2	4QW	E	301	-	18,20,20	4.89	3 (16%)	19,27,27	1.85	5 (26%)
2	4QW	A	301	-	18,20,20	5.17	4 (22%)	19,27,27	1.42	2 (10%)
2	4QW	B	301	-	18,20,20	4.87	4 (22%)	19,27,27	1.62	3 (15%)
2	4QW	D	301	-	18,20,20	4.99	5 (27%)	19,27,27	1.60	3 (15%)

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
2	4QW	C	301	-	-	2/6/18/18	0/3/3/3
2	4QW	E	301	-	-	4/6/18/18	0/3/3/3
2	4QW	A	301	-	-	4/6/18/18	0/3/3/3

Continued on next page...

W O R L D W I D E
www PDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	4QW	B	301	-	-	<mark>2/6/18/18</mark>	0/3/3/3
2	4QW	D	301	-	-	<mark>2/6/18/18</mark>	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	301	4QW	C12-N11	21.13	1.44	1.29
2	D	301	4QW	C12-N11	20.17	1.44	1.29
2	C	301	4QW	C12-N11	20.08	1.44	1.29
2	B	301	4QW	C12-N11	19.90	1.43	1.29
2	E	301	4QW	C12-N11	19.90	1.43	1.29

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	301	4QW	C09-C08-C07	-5.83	121.27	131.36
2	C	301	4QW	C09-C10-N11	-4.70	102.74	106.59
2	E	301	4QW	C09-C10-N11	-4.43	102.97	106.59
2	E	301	4QW	C09-C08-C07	-4.29	123.94	131.36
2	B	301	4QW	C09-C08-C07	-4.08	124.30	131.36

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	4QW	C06-C07-C08-C12
2	B	301	4QW	C06-C07-C08-C12
2	C	301	4QW	C06-C07-C08-C12
2	D	301	4QW	C06-C07-C08-C12
2	E	301	4QW	C06-C07-C08-C12

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	4QW	1	0
2	B	301	4QW	1	0
2	D	301	4QW	1	0

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.

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(Å²)	Q<0.9
1	A	195/218 (89%)	-0.46	0	100	100	32, 50, 79, 99	0
1	B	192/218 (88%)	-0.35	4 (2%)	63	58	25, 45, 73, 88	0
1	C	198/218 (90%)	-0.55	0	100	100	26, 36, 56, 76	0
1	D	199/218 (91%)	-0.42	1 (0%)	91	89	27, 45, 67, 84	0
1	E	199/218 (91%)	-0.39	1 (0%)	91	89	31, 46, 79, 90	0
All	All	983/1090 (90%)	-0.43	6 (0%)	89	88	25, 45, 74, 99	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	63	LEU	3.1
1	E	4	ALA	3.0
1	D	188	CYS	2.5
1	B	24	ASP	2.5
1	B	5	ASP	2.3

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, 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	4QW	A	301	18/18	0.97	0.16	46,55,67,73	0
2	4QW	B	301	18/18	0.97	0.15	26,35,43,44	0
2	4QW	C	301	18/18	0.97	0.15	28,34,40,42	0
2	4QW	D	301	18/18	0.97	0.17	42,52,62,68	0
2	4QW	E	301	18/18	0.97	0.15	42,51,64,68	0

6.5 Other polymers (i)

There are no such residues in this entry.