



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 04:39 PM EDT

PDB ID : 3TWR  
Title : Crystal structure of ARC4 from human Tankyrase 2 in complex with peptide from human 3BP2  
Authors : Guettler, S.; Sicheri, F.  
Deposited on : 2011-09-22  
Resolution : 1.55 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
buster-report : 1.1.7 (2018)  
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) : 2.37.1

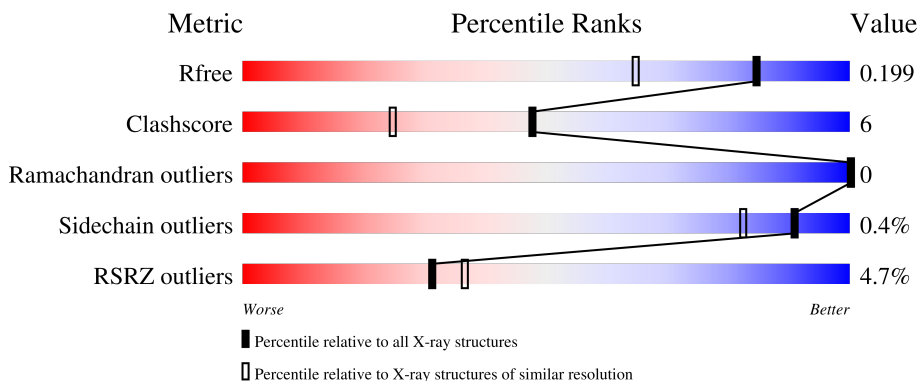
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.55 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	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	165	 5% 88% 7% 5%
1	B	165	 4% 86% 9% 5%
1	C	165	 4% 85% 10% .
1	D	165	 4% 84% 12% . .
2	E	16	 6% 62% 19% 19%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	16	<p>6% 75% 6% 19%</p>
2	G	16	<p>69% 31%</p>
2	H	16	<p>6% 69% 6% 25%</p>

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	A	701	-	-	X	-
4	PE8	E	100	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 6150 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 Tankyrase-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	157	1218	760	219	235	4	0	2	0
1	B	157	1253	782	227	240	4	0	6	0
1	C	158	1244	775	224	241	4	0	4	0
1	D	158	1255	784	227	240	4	0	6	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	485	GLY	-	EXPRESSION TAG	UNP Q9H2K2
A	486	ALA	-	EXPRESSION TAG	UNP Q9H2K2
A	487	MET	-	EXPRESSION TAG	UNP Q9H2K2
B	485	GLY	-	EXPRESSION TAG	UNP Q9H2K2
B	486	ALA	-	EXPRESSION TAG	UNP Q9H2K2
B	487	MET	-	EXPRESSION TAG	UNP Q9H2K2
C	485	GLY	-	EXPRESSION TAG	UNP Q9H2K2
C	486	ALA	-	EXPRESSION TAG	UNP Q9H2K2
C	487	MET	-	EXPRESSION TAG	UNP Q9H2K2
D	485	GLY	-	EXPRESSION TAG	UNP Q9H2K2
D	486	ALA	-	EXPRESSION TAG	UNP Q9H2K2
D	487	MET	-	EXPRESSION TAG	UNP Q9H2K2

- Molecule 2 is a protein called SH3 domain-binding protein 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	13	116	68	24	24	0	1	0
2	F	13	110	65	23	22	0	1	0

*Continued on next page...*

Continued from previous page...

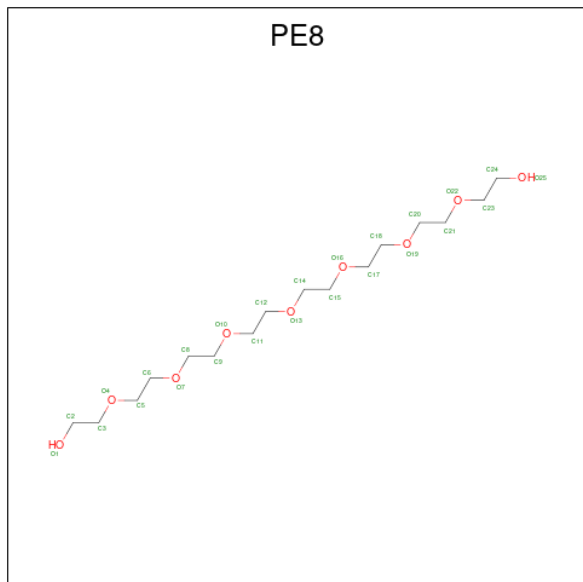
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	11	Total	C	N	O	0	0	0
			87	51	19	17			
2	H	12	Total	C	N	O	0	0	0
			96	56	21	19			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is 3,6,9,12,15,18,21-HEPTAOXATRICOSANE-1,23-DIOL (three-letter code: PE8) (formula: C<sub>16</sub>H<sub>34</sub>O<sub>9</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			13	8	5		
4	E	1	Total	C	O	0	0
			25	16	9		


- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	123	Total	O	0	0
			123	123		
5	B	177	Total	O	0	0
			177	177		
5	C	164	Total	O	0	0
			164	164		
5	D	156	Total	O	0	0
			156	156		
5	E	16	Total	O	0	0
			16	16		
5	F	20	Total	O	0	0
			20	20		
5	G	19	Total	O	0	0
			19	19		
5	H	13	Total	O	0	0
			13	13		

### 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: Tankyrase-2

Chain A: 




- Molecule 1: Tankyrase-2

Chain B: 




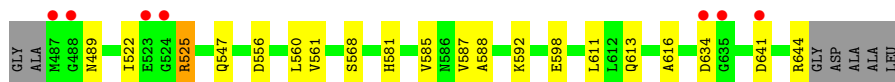
- Molecule 1: Tankyrase-2

Chain C: 



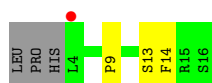
- Molecule 1: Tankyrase-2

Chain D: 

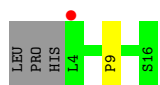
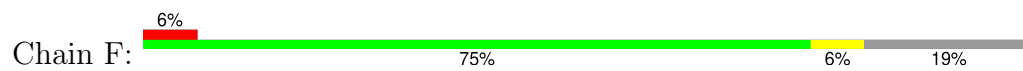


- Molecule 2: SH3 domain-binding protein 2

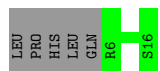
Chain E: 



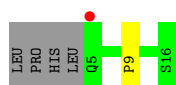
- Molecule 2: SH3 domain-binding protein 2



- Molecule 2: SH3 domain-binding protein 2



- Molecule 2: SH3 domain-binding protein 2





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.68Å 105.16Å 128.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.24 – 1.55 33.24 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.4 (33.24-1.55) 99.4 (33.24-1.55)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 1.55Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.181 , 0.206 0.174 , 0.199	Depositor DCC
$R_{free}$ test set	6057 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.6	Xtrriage
Anisotropy	0.155	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 48.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6150	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: SET, SO4, PE8

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	0.31	0/1239	0.48	0/1678
1	B	0.32	0/1274	0.49	0/1726
1	C	0.31	0/1265	0.49	0/1713
1	D	0.30	0/1277	0.47	0/1731
2	E	0.29	0/111	0.54	0/148
2	F	0.31	0/105	0.53	0/141
2	G	0.35	0/82	0.54	0/109
2	H	0.30	0/91	0.50	0/121
All	All	0.31	0/5444	0.49	0/7367

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 [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	1218	0	1206	18	1
1	B	1253	0	1243	12	0
1	C	1244	0	1229	13	0
1	D	1255	0	1240	14	0
2	E	116	0	107	4	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	110	0	103	1	0
2	G	87	0	79	0	0
2	H	96	0	87	1	0
3	A	10	0	0	2	0
3	B	5	0	0	1	0
3	C	10	0	0	0	0
3	D	10	0	0	0	0
3	F	5	0	0	0	0
3	H	5	0	0	0	0
4	C	13	0	17	4	0
4	E	25	0	34	17	0
5	A	123	0	0	4	0
5	B	177	0	0	1	0
5	C	164	0	0	1	1
5	D	156	0	0	4	0
5	E	16	0	0	0	0
5	F	20	0	0	0	0
5	G	19	0	0	0	0
5	H	13	0	0	0	0
All	All	6150	0	5345	65	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 6.

All (65) 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:C:640:GLN:HB3	1:C:644:ARG:HH12	1.01	1.08
1:A:604:LYS:HZ3	4:E:100:PE8:H121	1.26	0.97
1:C:640:GLN:HB3	1:C:644:ARG:NH1	1.79	0.95
1:B:494[B]:ARG:NH1	3:B:700:SO4:O1	2.03	0.91
1:A:552[B]:VAL:HG23	5:A:825:HOH:O	1.76	0.83
1:A:620:LYS:HG3	5:A:876:HOH:O	1.79	0.83
2:E:14:PHE:CZ	4:E:100:PE8:H151	2.18	0.79
1:D:547:GLN:OE1	5:D:937:HOH:O	2.01	0.79
5:D:824:HOH:O	4:E:100:PE8:H22	1.88	0.73
1:C:604:LYS:NZ	4:C:700:PE8:H82	2.03	0.73
1:C:640:GLN:CB	1:C:644:ARG:HH12	1.92	0.72
1:A:604:LYS:HZ2	4:E:100:PE8:H62	1.56	0.71
1:D:581[A]:HIS:ND1	5:D:942:HOH:O	2.22	0.70
1:A:604:LYS:NZ	4:E:100:PE8:H91	2.09	0.68

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:701:SO4:O2	5:A:884:HOH:O	2.12	0.67
1:B:488:GLY:HA2	1:B:522:ILE:CD1	2.24	0.67
1:B:552[B]:VAL:HG23	5:B:820:HOH:O	1.94	0.67
1:A:604:LYS:NZ	4:E:100:PE8:H121	2.07	0.66
1:B:624:ASP:O	4:E:100:PE8:H142	1.94	0.66
1:D:588:ALA:HB1	1:D:592:LYS:HA	1.78	0.65
1:A:576:GLU:OE2	5:A:907:HOH:O	2.15	0.65
1:A:560:LEU:HD13	2:E:9:PRO:HB3	1.82	0.60
1:A:617:ASP:HB3	1:A:620:LYS:HG2	1.83	0.60
1:C:604:LYS:HZ2	4:C:700:PE8:H82	1.67	0.59
1:D:560:LEU:HD13	2:H:9:PRO:HB3	1.85	0.59
2:E:14:PHE:CE1	4:E:100:PE8:H151	2.37	0.58
1:C:588:ALA:HB1	1:C:592:LYS:HA	1.84	0.58
1:D:568:SER:HB2	1:D:598:GLU:HG2	1.85	0.57
1:C:561:VAL:HG21	1:C:587[A]:VAL:HG11	1.86	0.57
1:A:522:ILE:HD11	5:C:887:HOH:O	2.05	0.55
1:A:604:LYS:HZ3	4:E:100:PE8:H91	1.71	0.55
1:B:488:GLY:HA2	1:B:522:ILE:HD11	1.89	0.54
1:C:570:GLY:HA2	1:C:607:ILE:HD12	1.88	0.54
1:D:613:GLN:HB3	4:E:100:PE8:H242	1.90	0.54
1:A:604:LYS:NZ	4:E:100:PE8:H31	2.23	0.54
1:D:525:ARG:O	1:D:556:ASP:HB2	2.08	0.53
1:A:570:GLY:HA3	4:E:100:PE8:H32	1.92	0.52
1:B:641:ASP:OD1	1:B:644:ARG:NH2	2.41	0.52
2:E:13[A]:SER:OG	4:E:100:PE8:H21	2.09	0.52
1:B:525:ARG:HB2	1:B:557:LYS:HG2	1.91	0.52
1:D:585[B]:VAL:CG1	1:D:616:ALA:HB2	2.40	0.52
1:B:588:ALA:HB1	1:B:592:LYS:HA	1.92	0.51
1:D:581[A]:HIS:CE1	5:D:942:HOH:O	2.64	0.51
1:A:617:ASP:HB3	1:A:620:LYS:CG	2.41	0.50
1:C:489:ASN:OD1	1:C:522:ILE:HD13	2.12	0.49
1:B:522:ILE:HG13	1:B:523:GLU:N	2.27	0.49
1:B:560:LEU:HD13	2:F:9:PRO:HB3	1.94	0.49
1:A:604:LYS:HZ2	4:E:100:PE8:H91	1.76	0.48
1:D:489:ASN:ND2	1:D:522:ILE:HG12	2.28	0.48
1:C:509:LYS:HG2	1:C:544:TYR:CE1	2.49	0.47
4:C:700:PE8:H81	4:C:700:PE8:H112	1.63	0.47
1:D:561:VAL:HG21	1:D:587:VAL:HG11	1.97	0.46
1:D:641:ASP:OD1	1:D:644:ARG:NH2	2.48	0.46
1:C:604:LYS:HZ3	4:C:700:PE8:H82	1.81	0.45
1:C:525:ARG:O	1:C:556:ASP:HB2	2.18	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:540[B]:SER:OG	3:A:701:SO4:S	2.73	0.43
4:E:100:PE8:H62	4:E:100:PE8:H91	1.72	0.43
1:B:488:GLY:HA2	1:B:522:ILE:HD13	1.99	0.43
1:A:604:LYS:HZ1	4:E:100:PE8:H31	1.85	0.41
1:C:491:GLU:HG3	1:C:494:ARG:HH21	1.86	0.41
4:E:100:PE8:H121	4:E:100:PE8:H91	1.72	0.41
1:A:570:GLY:HA2	1:A:607:ILE:HD12	2.03	0.41
1:D:585[B]:VAL:HG11	1:D:616:ALA:HB2	2.02	0.41
1:B:521:ASP:O	1:B:526:GLN:HA	2.21	0.41
1:D:585[B]:VAL:HG13	1:D:611:LEU:HD22	2.04	0.40

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:646:ASP:OD1	5:C:916:HOH:O[3_545]	1.97	0.23

## 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	157/165 (95%)	157 (100%)	0	0	100	100
1	B	161/165 (98%)	160 (99%)	1 (1%)	0	100	100
1	C	161/165 (98%)	160 (99%)	1 (1%)	0	100	100
1	D	162/165 (98%)	161 (99%)	1 (1%)	0	100	100
2	E	13/16 (81%)	13 (100%)	0	0	100	100
2	F	12/16 (75%)	12 (100%)	0	0	100	100
2	G	9/16 (56%)	9 (100%)	0	0	100	100
2	H	10/16 (62%)	10 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	685/724 (95%)	682 (100%)	3 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	130/131 (99%)	130 (100%)	0	100	100
1	B	134/131 (102%)	134 (100%)	0	100	100
1	C	133/131 (102%)	133 (100%)	0	100	100
1	D	134/131 (102%)	132 (98%)	2 (2%)	65	37
2	E	13/14 (93%)	13 (100%)	0	100	100
2	F	12/14 (86%)	12 (100%)	0	100	100
2	G	9/14 (64%)	9 (100%)	0	100	100
2	H	10/14 (71%)	10 (100%)	0	100	100
All	All	575/580 (99%)	573 (100%)	2 (0%)	91	85

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	525	ARG
1	D	634	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

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
2	SET	F	16	2	5,6,6	0.32	0	5,7,7	0.52	0
2	SET	E	16	2	5,6,6	0.29	0	5,7,7	0.19	0
2	SET	G	16	2	5,6,6	0.43	0	5,7,7	0.50	0
2	SET	H	16	2	5,6,6	0.27	0	5,7,7	0.47	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
2	SET	F	16	2	-	1/6/6/6	-
2	SET	E	16	2	-	2/6/6/6	-
2	SET	G	16	2	-	1/6/6/6	-
2	SET	H	16	2	-	1/6/6/6	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	16	SET	NT-C-CA-N
2	F	16	SET	NT-C-CA-N
2	G	16	SET	NT-C-CA-N
2	H	16	SET	NT-C-CA-N
2	E	16	SET	O-C-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

11 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
3	SO4	C	701	-	4,4,4	0.21	0	6,6,6	0.16	0
3	SO4	A	700	-	4,4,4	0.24	0	6,6,6	0.19	0
4	PE8	E	100	-	24,24,24	0.43	0	23,23,23	0.60	0
3	SO4	F	100	-	4,4,4	0.23	0	6,6,6	0.29	0
3	SO4	D	701	-	4,4,4	0.26	0	6,6,6	0.23	0
3	SO4	A	701	-	4,4,4	0.25	0	6,6,6	0.30	0
3	SO4	H	100	-	4,4,4	0.27	0	6,6,6	0.22	0
3	SO4	C	702	-	4,4,4	0.22	0	6,6,6	0.13	0
4	PE8	C	700	-	12,12,24	0.43	0	11,11,23	0.60	0
3	SO4	D	700	-	4,4,4	0.28	0	6,6,6	0.29	0
3	SO4	B	700	-	4,4,4	0.23	0	6,6,6	0.07	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
4	PE8	C	700	-	-	5/10/10/22	-
4	PE8	E	100	-	-	8/22/22/22	-

There are no bond length outliers.

There are no bond angle outliers.



There are no chirality outliers.

All (13) torsion outliers are listed below:

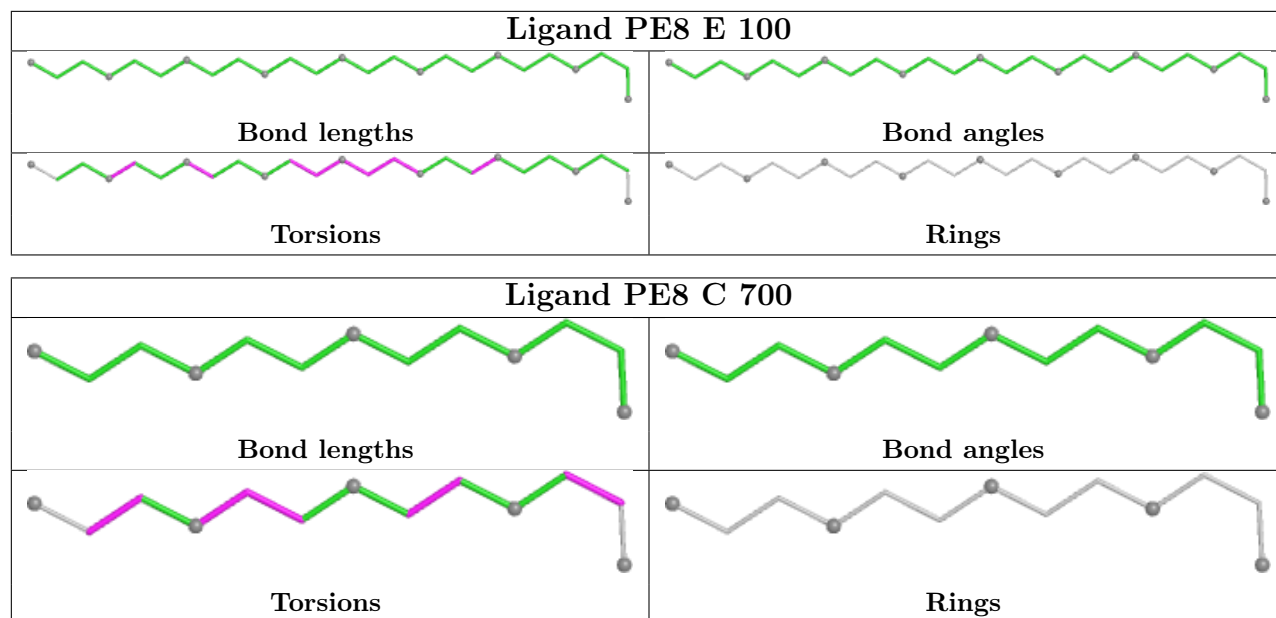
Mol	Chain	Res	Type	Atoms
4	E	100	PE8	O10-C11-C12-O13
4	C	700	PE8	O4-C5-C6-O7
4	E	100	PE8	O13-C14-C15-O16
4	C	700	PE8	O10-C11-C12-O13
4	E	100	PE8	C15-C14-O13-C12
4	E	100	PE8	C20-C21-O22-C23
4	E	100	PE8	C17-C18-O19-C20
4	E	100	PE8	C12-C11-O10-C9
4	E	100	PE8	C11-C12-O13-C14
4	C	700	PE8	C8-C9-O10-C11
4	C	700	PE8	O7-C8-C9-O10
4	C	700	PE8	O1-C2-C3-O4
4	E	100	PE8	C9-C8-O7-C6

There are no ring outliers.

4 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	100	PE8	17	0
3	A	701	SO4	2	0
4	C	700	PE8	4	0
3	B	700	SO4	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	157/165 (95%)	0.23	9 (5%) 23 27	15, 25, 45, 66	10 (6%)
1	B	157/165 (95%)	-0.15	7 (4%) 33 38	15, 22, 41, 58	6 (3%)
1	C	158/165 (95%)	0.10	6 (3%) 40 47	15, 22, 44, 64	6 (3%)
1	D	158/165 (95%)	-0.03	7 (4%) 34 40	17, 25, 48, 76	5 (3%)
2	E	12/16 (75%)	0.44	1 (8%) 11 12	17, 22, 38, 55	0
2	F	12/16 (75%)	0.19	1 (8%) 11 12	18, 24, 41, 59	0
2	G	10/16 (62%)	-0.38	0 100 100	19, 27, 31, 36	0
2	H	11/16 (68%)	0.06	1 (9%) 9 10	21, 27, 39, 57	1 (9%)
All	All	675/724 (93%)	0.04	32 (4%) 31 36	15, 24, 45, 76	28 (4%)

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	634	ASP	5.6
1	A	634	ASP	5.6
2	E	4	LEU	5.4
2	F	4	LEU	5.1
1	C	488	GLY	4.7
1	D	635	GLY	4.1
1	B	488	GLY	4.0
1	B	522	ILE	3.9
1	B	523	GLU	3.6
1	B	644	ARG	3.5
1	B	635	GLY	3.4
1	D	487	MET	3.2
1	B	637[A]	THR	3.2
1	D	524	GLY	3.1
1	C	563	LEU	3.0
1	A	490	SER	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	635	GLY	2.9
1	B	634	ASP	2.9
2	H	5	GLN	2.9
1	D	488	GLY	2.8
1	A	623	ARG	2.7
1	A	552[A]	VAL	2.6
1	C	584	VAL	2.6
1	C	645	GLY	2.6
1	A	522	ILE	2.6
1	A	494	ARG	2.4
1	C	489	ASN	2.3
1	A	624	ASP	2.3
1	A	644	ARG	2.3
1	D	641	ASP	2.3
1	C	587[A]	VAL	2.2
1	D	523	GLU	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	SET	H	16	7/7	0.90	0.17	28,41,49,55	0
2	SET	G	16	7/7	0.91	0.08	26,31,42,49	0
2	SET	F	16	7/7	0.91	0.10	21,30,37,42	0
2	SET	E	16	7/7	0.92	0.07	21,27,30,37	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

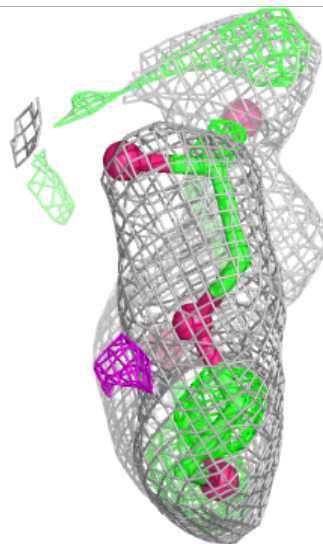
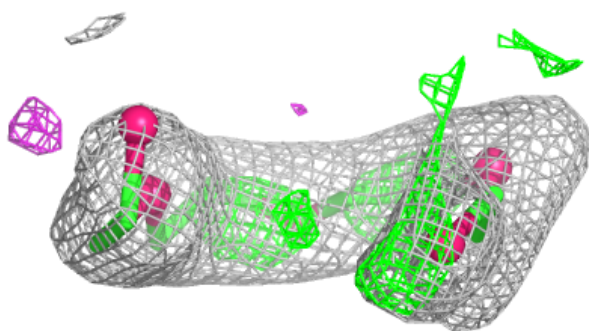
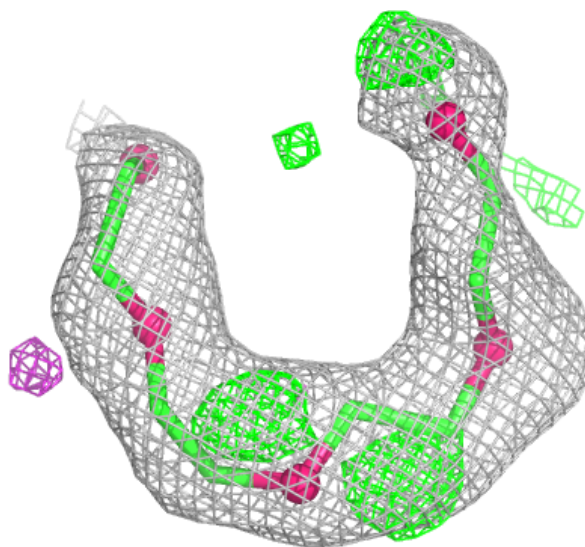
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

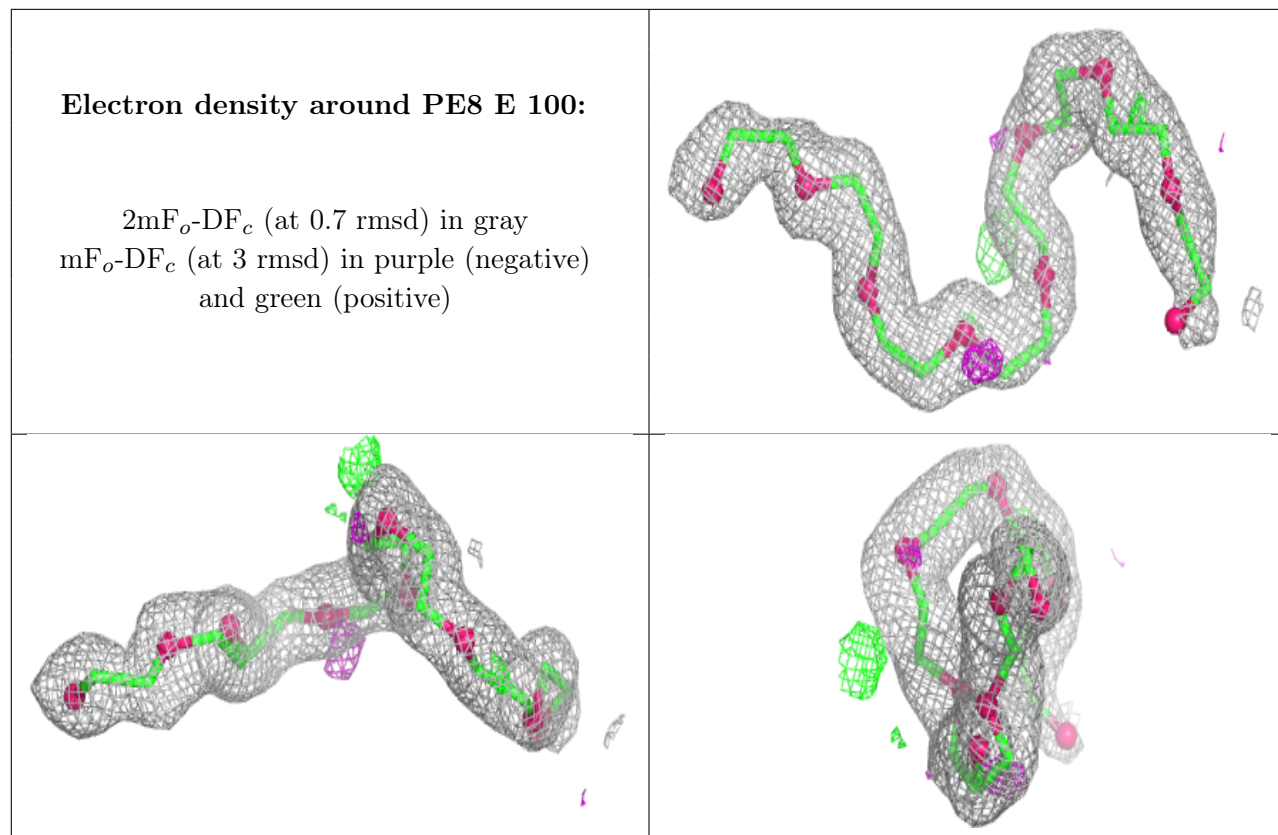
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	A	701	5/5	0.84	0.16	20,22,37,39	5
3	SO4	B	700	5/5	0.86	0.18	57,58,60,61	5
4	PE8	C	700	13/25	0.89	0.16	33,43,60,63	0
4	PE8	E	100	25/25	0.92	0.10	17,34,47,56	0
3	SO4	F	100	5/5	0.93	0.11	20,29,35,42	5
3	SO4	H	100	5/5	0.93	0.12	26,32,41,44	5
3	SO4	C	701	5/5	0.94	0.21	15,20,22,31	5
3	SO4	C	702	5/5	0.94	0.23	32,36,54,59	5
3	SO4	A	700	5/5	0.95	0.08	26,31,37,38	5
3	SO4	D	701	5/5	0.96	0.10	27,28,36,46	5
3	SO4	D	700	5/5	0.97	0.13	15,22,24,26	5

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around PE8 C 700:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.