

Full wwPDB X-ray Structure Validation Report (i)

May 19, 2020 - 04:40 am BST

PDB ID	:	4BLS
Title	:	P4 PROTEIN FROM BACTERIOPHAGE PHI12 Q278A MUTANT IN
		COMPLEX WITH AMPcPP
Authors	:	El Omari, K.; Meier, C.; Kainov, D.; Sutton, G.; Grimes, J.M.; Poranen,
		M.M.; Bamford, D.H.; Tuma, R.; Stuart, D.I.; Mancini, E.J.
Deposited on	:	2013-05-04
Resolution	:	2.60 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* A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

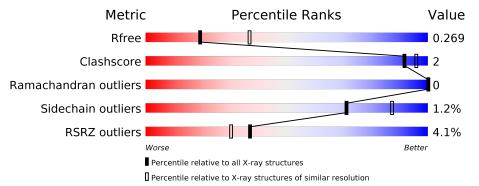
MolProbity		4.02b-467 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)		1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
$\operatorname{CCP4}$:	$7.0.044 (\mathrm{Gargrove})$
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries},{ m resolution\ range}({ m \AA}))$
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 on 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	А	331	87%	•	9%
1	В	331	4% 87%	••	8%
1	С	331	4%	•	9%



4BLS

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6948 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.

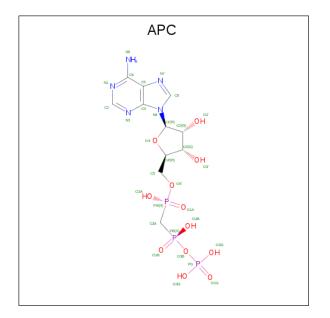
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Λ	302	Total	С	Ν	Ο	S	0	0	0
1	л	302	2280	1429	397	447	7	0	0	0
1	В	303	Total	С	Ν	Ο	S	0	0	0
1	D	000	2291	1435	401	448	7	0	0	0
1	C	302	Total	С	Ν	Ο	S	0	0	0
		C 302	2284	1431	400	446	7	U	0	U

• Molecule 1 is a protein called NTPASE P4.

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue Modelled		Actual	Comment	Reference
А	278	ALA	GLN	engineered mutation	UNP Q94M05
В	278	ALA	GLN	engineered mutation	UNP Q94M05
С	278	ALA	GLN	engineered mutation	UNP Q94M05

• Molecule 2 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



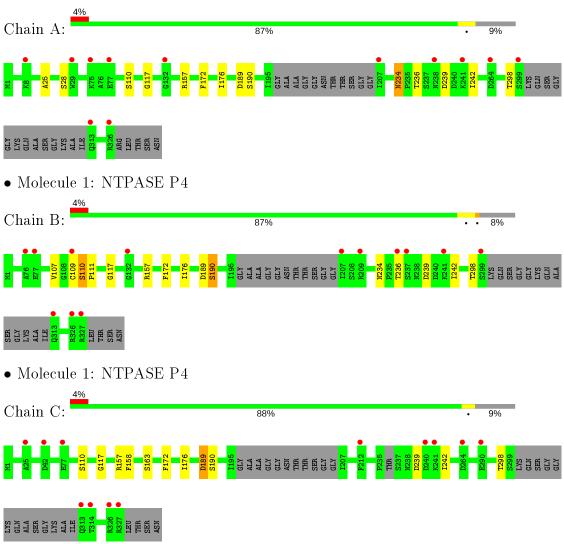


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf			
2	Λ	1	Total	С	Ν	Ο	Р	0	0	
	Л	I	31	11	5	12	3	0	0	
2	р	1	Total	С	Ν	Ο	Р	0	0	
	D	T	31	11	5	12	3	0	0	
0	С	1	Total	С	Ν	Ο	Р	0	0	
	2 C		31	11	5	12	3	U	0	



3 Residue-property plots (i)

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.



• Molecule 1: NTPASE P4



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants	105.63Å 132.33 Å 159.94 Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 - 2.60	Depositor
Resolution (A)	19.99 - 2.60	EDS
% Data completeness	91.6(20.00-2.60)	Depositor
(in resolution range)	91.9(19.99-2.60)	EDS
R _{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$6.05 \;({\rm at}\; 2.59{ m \AA})$	Xtriage
Refinement program	REFMAC $5.7.0032$	Depositor
R, R_{free}	0.257 , 0.274	Depositor
Π, Π_{free}	0.255 , 0.269	DCC
R_{free} test set	1607 reflections (5.04%)	wwPDB-VP
Wilson B-factor $(Å^2)$	44.7	Xtriage
Anisotropy	1.107	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.31 , 32.5	EDS
L-test for twinning ²	$ < L >=0.51, < L^2>=0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6948	wwPDB-VP
Average B, all atoms $(Å^2)$	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.32% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: APC

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	Mol Chain		lengths	Bond angles		
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.32	0/2316	0.52	0/3135	
1	В	0.32	0/2327	0.53	0/3149	
1	С	0.31	0/2319	0.52	0/3136	
All	All	0.32	0/6962	0.52	0/9420	

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	А	2280	0	2262	6	0
1	В	2291	0	2275	11	0
1	С	2284	0	2267	8	0
2	А	31	0	14	0	0
2	В	31	0	14	1	0
2	С	31	0	14	0	0
All	All	6948	0	6846	22	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 2.



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:234:ASN:O	1:B:234:ASN:OD1	2.15	0.64
1:B:117:GLY:HA2	1:B:298:THR:HG23	1.82	0.62
1:B:109:CYS:SG	1:C:163:SER:OG	2.66	0.52
1:C:158:PHE:N	1:C:189:ASP:O	2.43	0.52
1:B:172:PHE:CZ	1:B:176:ILE:HD11	2.47	0.50
1:C:172:PHE:CZ	1:C:176:ILE:HD11	2.47	0.50
1:A:172:PHE:CZ	1:A:176:ILE:HD11	2.47	0.49
1:B:234:ASN:ND2	2:B:400:APC:O1G	2.46	0.49
1:C:157:ARG:HA	1:C:189:ASP:O	2.14	0.48
1:A:25:ALA:HB3	1:A:28:SER:OG	2.14	0.47
1:A:234:ASN:HB3	1:A:236:THR:HG23	1.97	0.46
1:B:189:ASP:HA	1:B:190:SER:HA	1.81	0.44
1:B:107:VAL:HG11	1:C:157:ARG:HD2	1.99	0.44
1:C:239:ASP:HB3	1:C:242:ILE:HG22	2.00	0.44
1:C:117:GLY:HA2	1:C:298:THR:HG23	2.00	0.43
1:A:117:GLY:HA2	1:A:298:THR:HG23	2.00	0.43
1:B:110:SER:N	1:B:111:PRO:HD3	2.35	0.42
1:A:239:ASP:HB3	1:A:242:ILE:HG22	2.00	0.42
1:B:239:ASP:HB3	1:B:242:ILE:HG22	2.01	0.42
1:A:157:ARG:HA	1:A:189:ASP:O	2.20	0.41
1:B:157:ARG:HA	1:B:189:ASP:O	2.21	0.41
1:B:109:CYS:HG	1:C:163:SER:HG	1.62	0.40

All (22) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

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	Percen	tiles
1	А	296/331~(89%)	291~(98%)	5(2%)	0	100	100
1	В	$297/331 \ (90\%)$	289 (97%)	8 (3%)	0	100	100





Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percer	ntiles
1	С	294/331~(89%)	288~(98%)	6(2%)	0	100	100
All	All	887/993~(89%)	868~(98%)	19 (2%)	0	100	100

Continued from previous page...

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	Analysed Rotameric Outlier		Percentiles		
1	А	247/264~(94%)	244~(99%)	3 (1%)	71 87		
1	В	248/264~(94%)	245~(99%)	3~(1%)	71 87		
1	С	247/264~(94%)	244 (99%)	3 (1%)	71 87		
All	All	742/792~(94%)	733~(99%)	9 (1%)	71 87		

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	110	SER
1	А	190	SER
1	А	234	ASN
1	В	110	SER
1	В	190	SER
1	В	236	THR
1	С	110	SER
1	С	189	ASP
1	С	190	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	253	ASN
1	А	268	GLN
1	А	297	HIS



Continued from previous page...

Mol	Chain	Res	Type
1	А	318	ASN
1	В	318	ASN
1	С	253	ASN
1	С	268	GLN
1	С	297	HIS
1	С	318	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)

3 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	Chain	Chain	Chain	Chain	Chain	Chain	Type Chain Res		Link	Bo	Bond lengths			Bond angles		
	туре	Cham	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2								
2	APC	В	400	-	27,33,33	1.61	7 (25%)	31,52,52	1.48	5(16%)								
2	APC	С	400	-	27,33,33	1.67	7 (25%)	31,52,52	1.42	5(16%)								
2	APC	А	400	-	27,33,33	1.67	8 (29%)	31,52,52	1.36	4 (12%)								

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.



Mol	Type	Chain	\mathbf{Res}	\mathbf{Link}	Chirals	Torsions	Rings
2	APC	В	400	-	-	5/15/38/38	0/3/3/3
2	APC	С	400	-	-	9/15/38/38	0/3/3/3
2	APC	А	400	-	-	3/15/38/38	0/3/3/3

'-' means no outliers of that kind were identified.

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	Ideal(Å)
2	А	400	APC	PA-O1A	3.89	1.60	1.51
2	С	400	APC	PA-O1A	3.80	1.60	1.51
2	В	400	APC	PA-O1A	3.75	1.60	1.51
2	В	400	APC	PA-O2A	-3.44	1.48	1.56
2	А	400	APC	PA-O2A	-3.44	1.48	1.56
2	С	400	APC	PA-O2A	-3.37	1.48	1.56
2	С	400	APC	PA-05'	3.33	1.62	1.57
2	С	400	APC	PB-O3B	3.05	1.61	1.58
2	А	400	APC	PA-O5'	3.00	1.61	1.57
2	В	400	APC	PA-05'	2.97	1.61	1.57
2	А	400	APC	PB-O3B	2.89	1.61	1.58
2	В	400	APC	C5-C4	2.57	1.47	1.40
2	А	400	APC	C5-C4	2.56	1.47	1.40
2	С	400	APC	C5-C4	2.51	1.47	1.40
2	В	400	APC	PB-O3B	2.44	1.61	1.58
2	А	400	APC	PB-O2B	2.40	1.62	1.56
2	В	400	APC	PB-O2B	2.27	1.61	1.56
2	С	400	APC	PB-O2B	2.22	1.61	1.56
2	А	400	APC	C2-N3	2.16	1.35	1.32
2	А	400	APC	O4'-C1'	2.04	1.43	1.41
2	С	400	APC	O4'-C1'	2.03	1.43	1.41
2	В	400	APC	C2-N3	2.00	1.35	1.32

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
2	С	400	APC	N3-C2-N1	-3.92	122.55	128.68
2	А	400	APC	N3-C2-N1	-3.79	122.76	128.68
2	В	400	APC	N3-C2-N1	-3.69	122.91	128.68
2	В	400	APC	PG-O3B-PB	-3.52	120.22	132.62
2	В	400	APC	C3'-C2'-C1'	3.29	105.94	100.98
2	С	400	APC	PG-O3B-PB	-3.22	121.26	132.62
2	А	400	APC	C4-C5-N7	-3.15	106.12	109.40
2	С	400	APC	C3'-C2'-C1'	3.04	105.56	100.98



Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
2	В	400	APC	C4-C5-N7	-2.97	106.30	109.40
2	А	400	APC	C3'-C2'-C1'	2.78	105.16	100.98
2	А	400	APC	PG-O3B-PB	-2.71	123.07	132.62
2	С	400	APC	C4-C5-N7	-2.62	106.67	109.40
2	В	400	APC	C2-N1-C6	2.20	122.52	118.75
2	С	400	APC	C2-N1-C6	2.19	122.50	118.75

Continued from previous page...

There are no chirality outliers.

Mol	Chain	\mathbf{Res}	Type	Atoms
2	В	400	APC	C5'-O5'-PA-O1A
2	В	400	APC	C3'-C4'-C5'-O5'
2	С	400	APC	C5'-O5'-PA-O1A
2	С	400	APC	C3'-C4'-C5'-O5'
2	А	400	APC	C5'-O5'-PA-C3A
2	В	400	APC	C4'-C5'-O5'-PA
2	В	400	APC	O4'-C4'-C5'-O5'
2	С	400	APC	O4'-C4'-C5'-O5'
2	С	400	APC	C4'-C5'-O5'-PA
2	А	400	APC	C4'-C5'-O5'-PA
2	А	400	APC	C3'-C4'-C5'-O5'
2	С	400	APC	PB-O3B-PG-O1G
2	С	400	APC	PB-C3A-PA-O1A
2	В	400	APC	C5'-O5'-PA-O2A
2	С	400	APC	C5'-O5'-PA-O2A
2	С	400	APC	PB-O3B-PG-O2G
2	С	400	APC	PB-O3B-PG-O3G

All (17) torsion outliers are listed below:

There are no ring outliers.

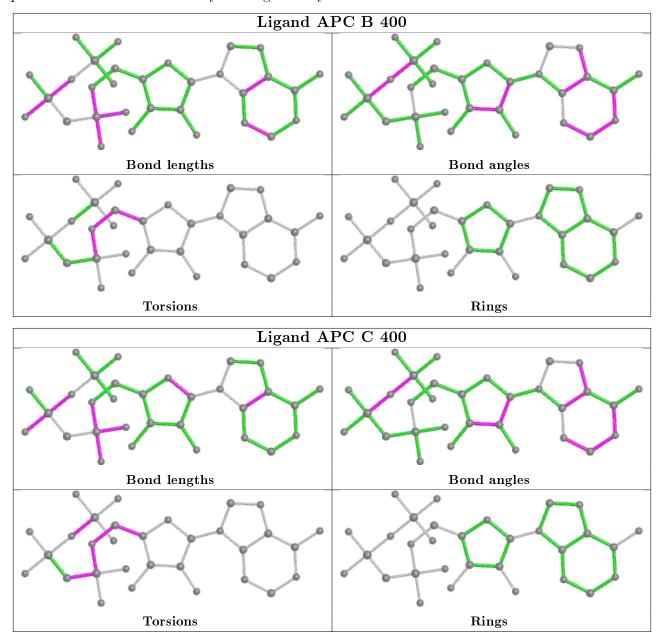
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	В	400	APC	1	0

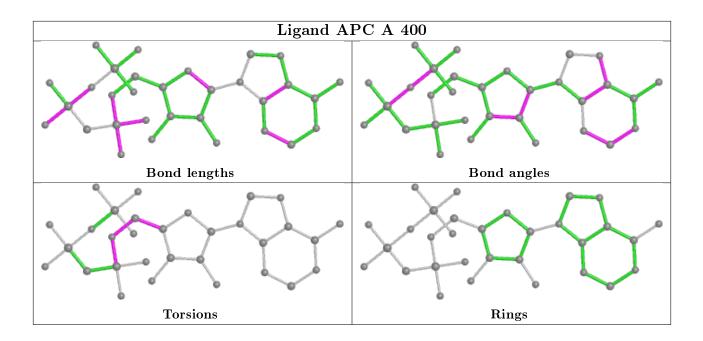
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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 (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, 95^{th} 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 $>$	# RS	RZ>2	2	$\mathbf{OWAB}(\mathbf{\AA}^2)$	$Q{<}0.9$
1	А	302/331~(91%)	0.16	12 (3%)	38	31	34,51,91,120	1 (0%)
1	В	303/331~(91%)	0.24	13 (4%)	35	28	38, 51, 98, 117	0
1	С	302/331~(91%)	0.20	12 (3%)	38	31	37, 54, 95, 125	0
All	All	907/993~(91%)	0.20	37 (4%)	37	30	34, 53, 96, 125	1 (0%)

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ	
1	В	313	GLN	4.9	
1	С	313	GLN	4.8	
1	В	236	THR	4.7	
1	В	109	CYS	4.6	
1	С	241	LYS	4.0	
1	В	299	SER	3.9	
1	А	313	GLN	3.9	
1	А	241	LYS	3.8	
1	А	132	GLY	3.5	
1	В	326	ARG	3.4	
1	С	326	ARG	3.3	
1	В	132	GLY	3.2	
1	А	238	ASN	3.1	
1	А	29	TRP	3.0	
1	С	212	PHE	3.0	
1	В	237	SER	2.9	
1	С	290	GLU	2.7	
1	А	326	ARG	2.6	
1	А	8	LYS	2.5	
1	С	42	ASP	2.5	
1	С	264	ASP	2.5	
1	А	299	SER	2.5	
1	С	240	ASP	2.5	



Mol	Chain	Res	Type	RSRZ
1	В	207	ILE	2.5
1	С	77	GLU	2.4
1	С	314	THR	2.4
1	А	75	LYS	2.4
1	А	77	GLU	2.4
1	В	77	GLU	2.4
1	С	25	ALA	2.3
1	А	207	ILE	2.3
1	А	264	ASP	2.2
1	В	327	ARG	2.2
1	В	241	LYS	2.2
1	В	209	ARG	2.1
1	С	327	ARG	2.1
1	В	76	ALA	2.0

Continued from previous page...

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.

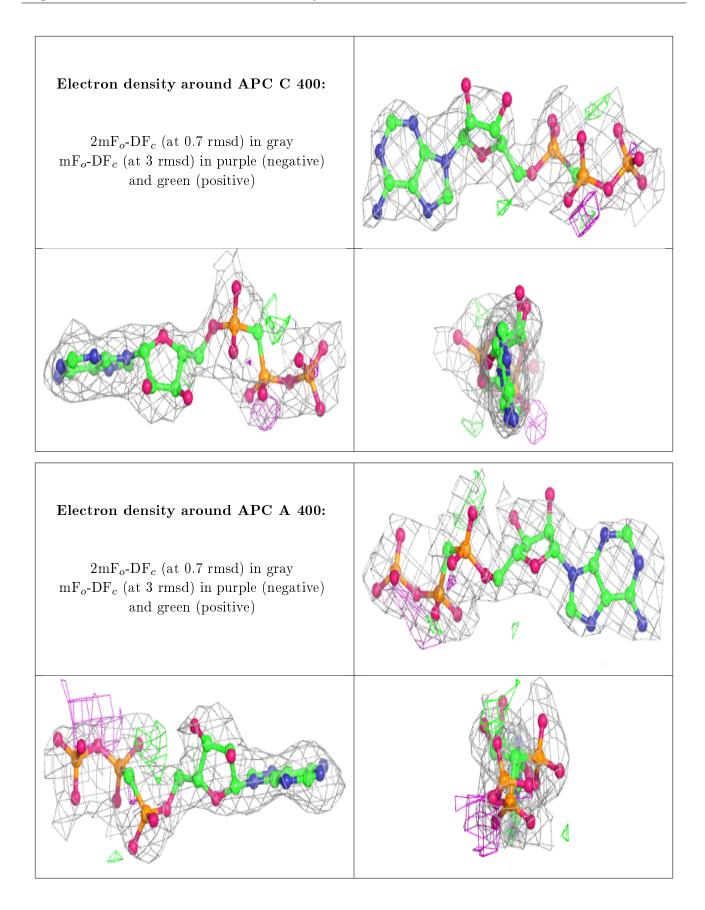
6.4 Ligands (i)

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, 95^{th} 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.

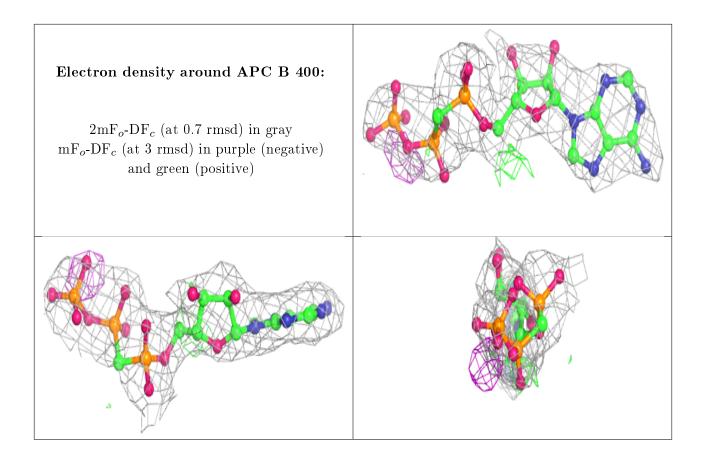
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	$Q{<}0.9$
2	APC	С	400	31/31	0.87	0.22	$64,\!74,\!91,\!96$	0
2	APC	А	400	31/31	0.87	0.22	$66,\!69,\!81,\!85$	0
2	APC	В	400	31/31	0.90	0.23	$63,\!71,\!85,\!87$	0

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.









6.5 Other polymers (i)

There are no such residues in this entry.

