



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 16, 2021 – 12:09 pm GMT

PDB ID : 7NHE
Title : Crystal structure of Arabidopsis thaliana Pdx1K166R-I333 complex
Authors : Rodrigues, M.J.; Zhang, Y.; Bolton, R.; Evans, G.; Giri, N.; Royant, A.; Begley, T.; Ealick, S.E.; Tews, I.
Deposited on : 2021-02-10
Resolution : 2.23 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 ⓘ symbol.

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

MolProbity : **FAILED**
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.24
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.24

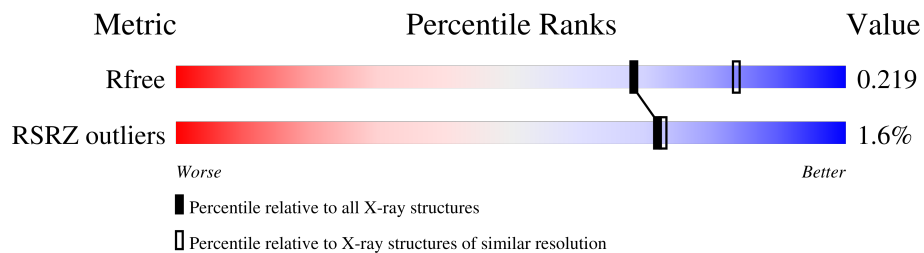
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.23 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	2391 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8681 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 Pyridoxal 5'-phosphate synthase subunit PDX1.3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	272	2061	1283	377	383	18	0	2	0
1	B	271	2048	1281	375	374	18	0	3	0
1	C	271	2036	1268	373	378	17	0	1	0
1	D	270	2033	1268	373	374	18	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

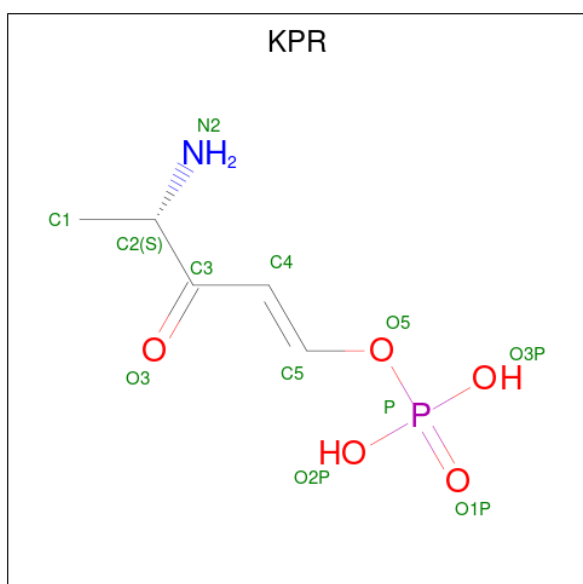
Chain	Residue	Modelled	Actual	Comment	Reference
A	166	ARG	LYS	engineered mutation	UNP Q8L940
B	166	ARG	LYS	engineered mutation	UNP Q8L940
C	166	ARG	LYS	engineered mutation	UNP Q8L940
D	166	ARG	LYS	engineered mutation	UNP Q8L940

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



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

- Molecule 3 is [({E},4 {S})-4-azanyl-3-oxidanylidene-pent-1-enyl] dihydrogen phosphate (three-letter code: KPR) (formula: C₅H₁₀NO₅P) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			12	5	1	5	1		
3	B	1	Total	C	N	O	P	0	0
			12	5	1	5	1		
3	C	1	Total	C	N	O	P	0	0
			12	5	1	5	1		
3	D	1	Total	C	N	O	P	0	0
			12	5	1	5	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	117	Total	O	0	0
			117	117		
4	B	110	Total	O	0	0
			110	110		
4	C	97	Total	O	0	0
			97	97		
4	D	111	Total	O	0	0
			111	111		

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	177.97Å 177.97Å 115.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	92.23 – 2.23 92.23 – 2.23	Depositor EDS
% Data completeness (in resolution range)	99.4 (92.23-2.23) 99.5 (92.23-2.23)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.60 (at 2.22Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.175 , 0.219 0.176 , 0.219	Depositor DCC
R_{free} test set	3317 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	34.2	Xtriage
Anisotropy	0.779	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for $-1/3^*h+1/3^*k+4/3^*l,-k,2/3^*h+1/3^*k+1/3^*l$ 0.000 for $-2/3^*h-1/3^*k-4/3^*l,-1/3^*h-2/3^*k+4/3^*l,-1/3^*h+1/3^*k+1/3^*l$ 0.000 for $-h,1/3^*h-1/3^*k-4/3^*l,-1/3^*h-2/3^*k+1/3^*l$ 0.000 for $-1/3^*h-2/3^*k+4/3^*l,-2/3^*h-1/3^*k-4/3^*l,1/3^*h-1/3^*k-1/3^*l$ 0.000 for $-h,2/3^*h+1/3^*k+4/3^*l,1/3^*h+2/3^*k-1/3^*l$ 0.000 for $1/3^*h+2/3^*k-4/3^*l,-k,-2/3^*h-1/3^*k-1/3^*l$ 0.012 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	8681	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

8 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	KPR	D	302	1	10,11,11	0.72	0	10,15,15	0.60	0
2	PO4	C	301	-	4,4,4	0.89	0	6,6,6	0.39	0
3	KPR	B	302	1	10,11,11	0.53	0	10,15,15	0.99	0
2	PO4	A	301	-	4,4,4	1.03	0	6,6,6	0.43	0
2	PO4	B	301	-	4,4,4	0.91	0	6,6,6	0.56	0
3	KPR	C	302	1	10,11,11	0.66	0	10,15,15	0.49	0
2	PO4	D	301	-	4,4,4	1.10	0	6,6,6	0.28	0
3	KPR	A	302	1	10,11,11	0.60	0	10,15,15	0.98	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
3	KPR	D	302	1	-	2/8/11/11	-
3	KPR	A	302	1	-	4/8/11/11	-
3	KPR	B	302	1	-	3/8/11/11	-
3	KPR	C	302	1	-	2/8/11/11	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	302	KPR	O3-C3-C4-C5
3	A	302	KPR	C2-C3-C4-C5
3	B	302	KPR	O3-C3-C4-C5
3	B	302	KPR	C2-C3-C4-C5
3	C	302	KPR	C3-C4-C5-O5
3	D	302	KPR	C1-C2-C3-O3
3	A	302	KPR	C1-C2-C3-C4
3	A	302	KPR	C1-C2-C3-O3
3	D	302	KPR	C1-C2-C3-C4
3	B	302	KPR	C1-C2-C3-C4

Continued on next page...

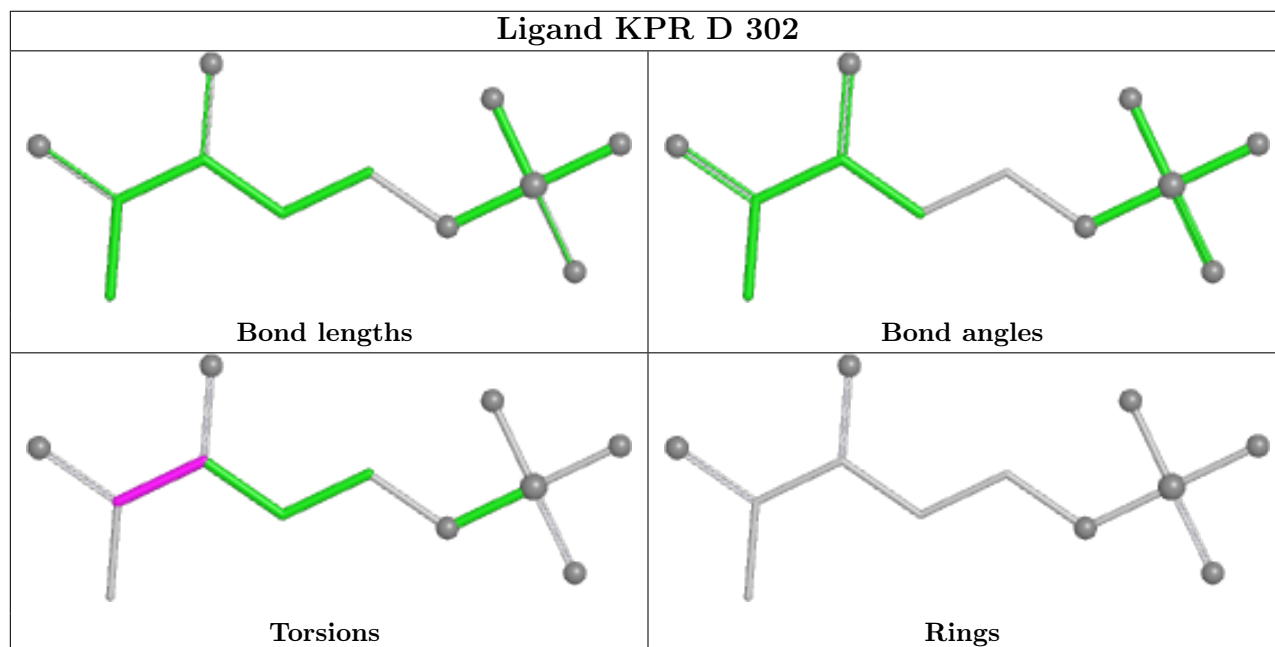
Continued from previous page...

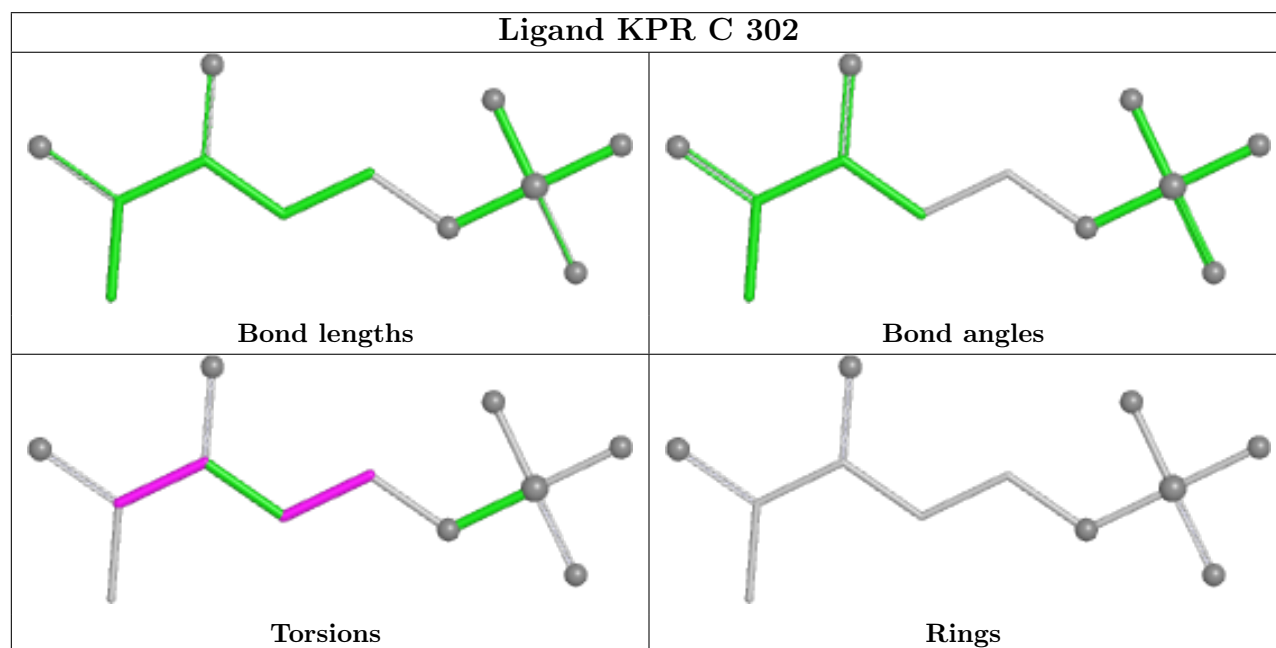
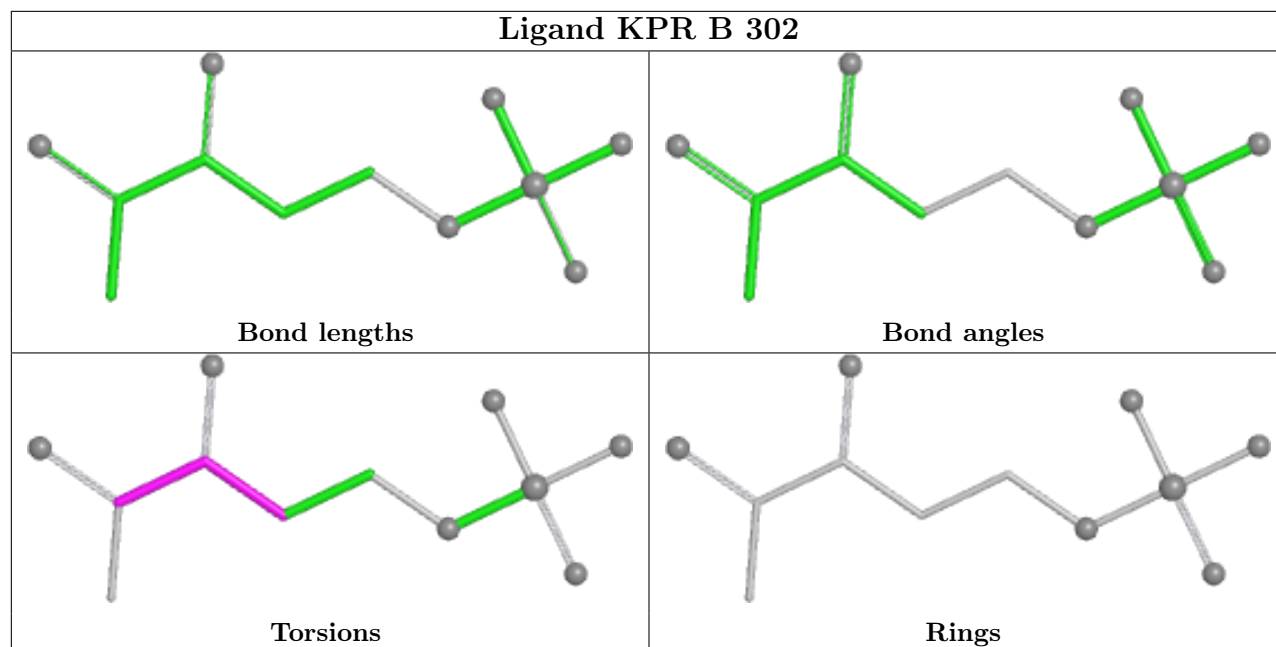
Mol	Chain	Res	Type	Atoms
3	C	302	KPR	C1-C2-C3-C4

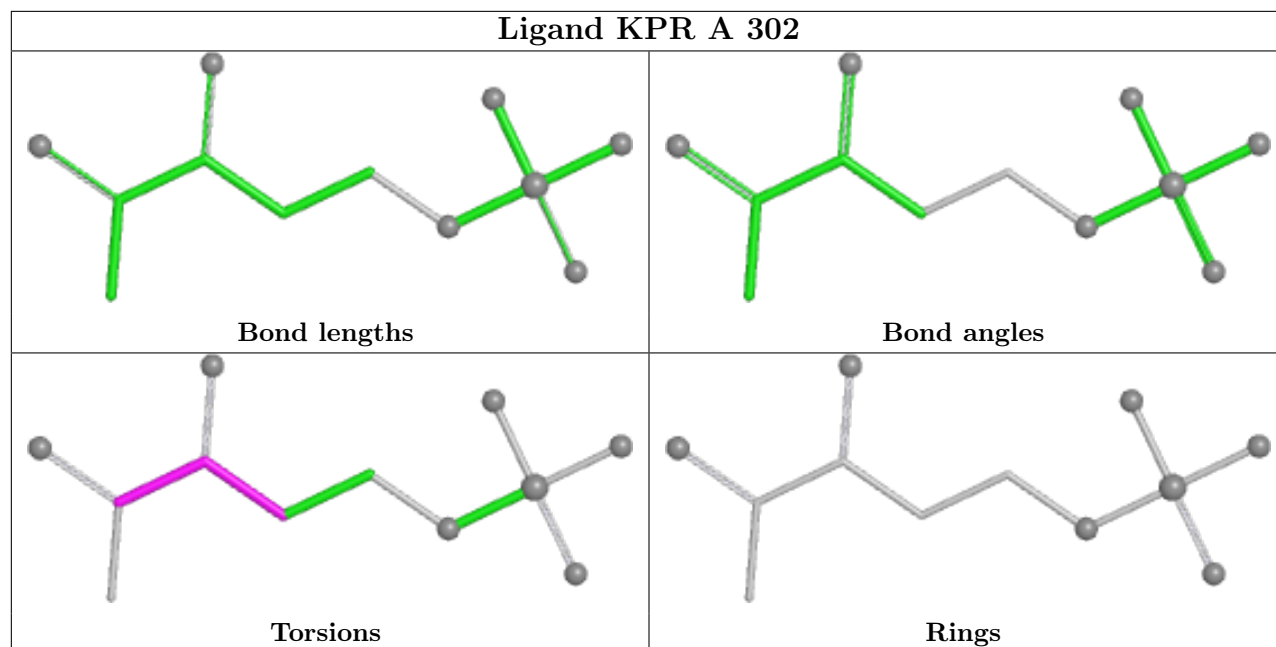
There are no ring outliers.

No monomer is involved in short contacts.

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.







4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 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	272/291 (93%)	-0.25	3 (1%) 80 81	32, 41, 66, 84	0
1	B	271/291 (93%)	-0.22	3 (1%) 80 81	33, 43, 71, 88	0
1	C	271/291 (93%)	-0.19	4 (1%) 73 74	35, 47, 80, 112	0
1	D	270/291 (92%)	-0.22	7 (2%) 56 57	33, 44, 83, 117	0
All	All	1084/1164 (93%)	-0.22	17 (1%) 72 73	32, 44, 77, 117	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	257	LYS	4.5
1	D	261	PRO	4.0
1	B	65	VAL	3.0
1	D	257	LYS	2.9
1	D	256	PHE	2.5
1	A	290	GLU	2.4
1	B	261	PRO	2.4
1	C	261	PRO	2.3
1	C	288	LEU	2.3
1	D	67	ALA	2.2
1	A	71	ALA	2.1
1	A	261	PRO	2.1
1	D	49	ARG	2.1
1	B	254	GLY	2.1
1	C	258	SER	2.1
1	D	258	SER	2.1
1	D	253	SER	2.1

5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.4 Ligands [i](#)

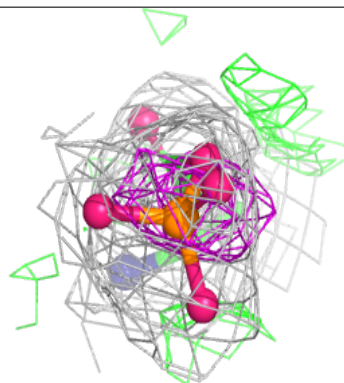
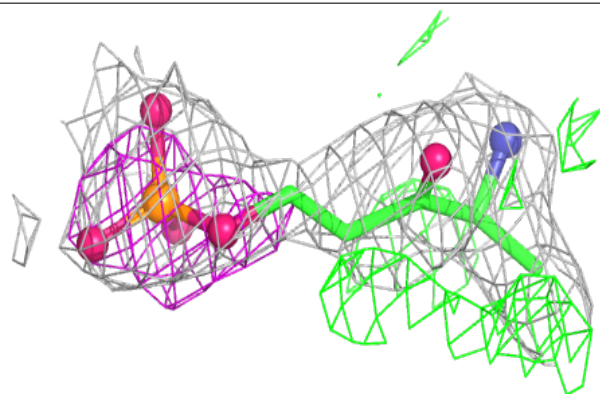
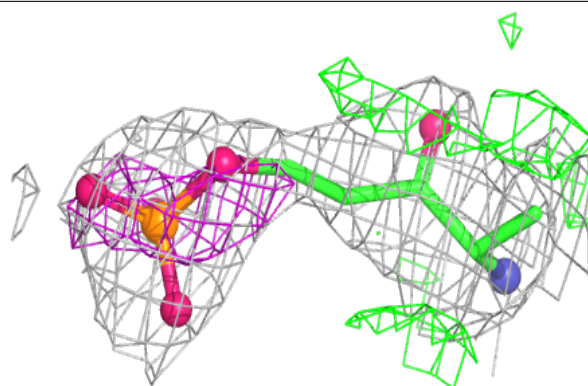
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	KPR	D	302	12/12	0.92	0.21	56,65,67,68	0
3	KPR	C	302	12/12	0.93	0.13	56,64,67,68	0
3	KPR	B	302	12/12	0.93	0.15	50,56,60,60	0
3	KPR	A	302	12/12	0.94	0.12	43,57,60,60	0
2	PO4	D	301	5/5	0.95	0.20	56,56,61,69	0
2	PO4	C	301	5/5	0.96	0.17	53,61,67,73	0
2	PO4	B	301	5/5	0.96	0.20	47,57,60,65	0
2	PO4	A	301	5/5	0.97	0.20	52,55,60,65	0

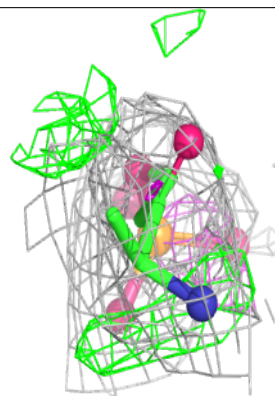
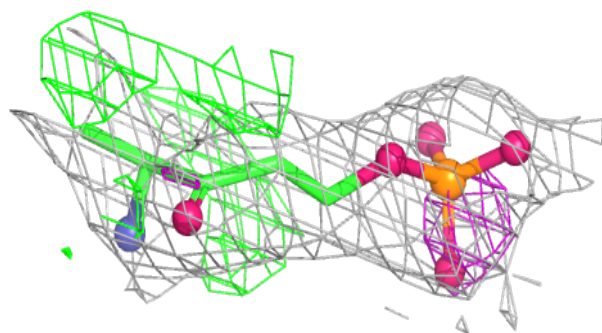
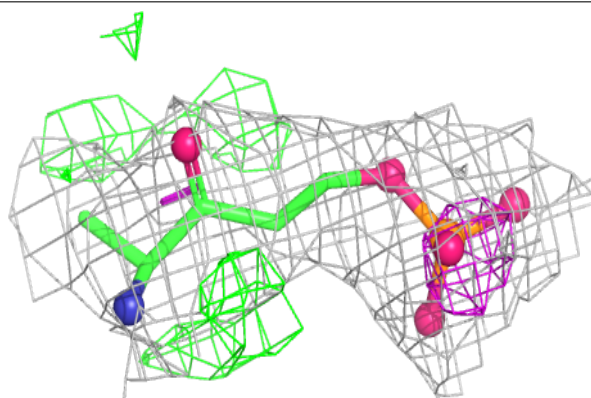
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

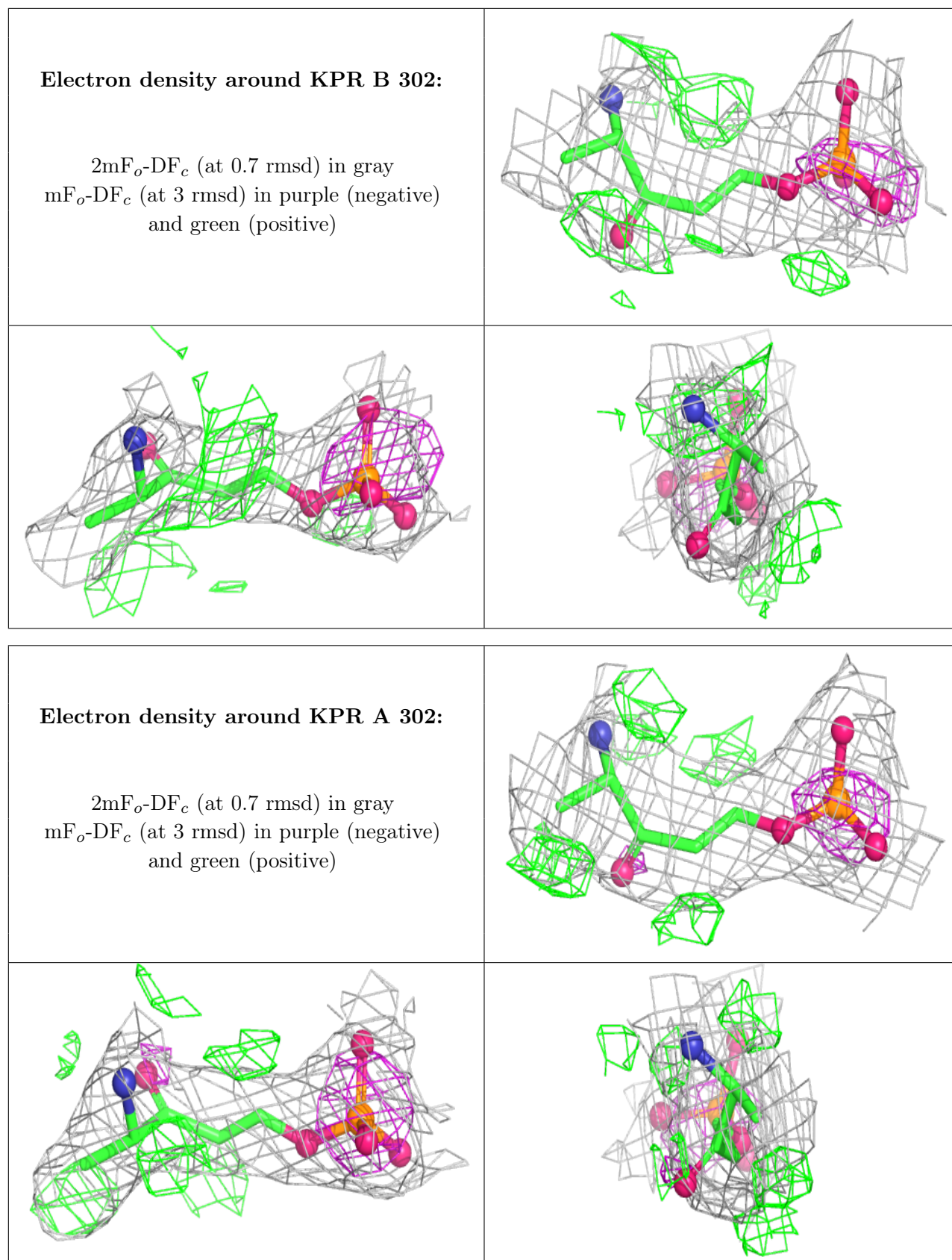
Electron density around KPR D 302:

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

**Electron density around KPR C 302:**

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





5.5 Other polymers [i](#)

There are no such residues in this entry.