



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:46 AM UTC

PDB ID : 2Q5J / pdb_00002q5j
Title : X-ray structure of phenylpyruvate decarboxylase in complex with 3-deaza-ThDP
Authors : Versees, W.; Spaepen, S.; Wood, M.D.; Leeper, F.J.; Vanderleyden, J.; Steyaert, J.
Deposited on : 2007-06-01
Resolution : 3.20 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 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

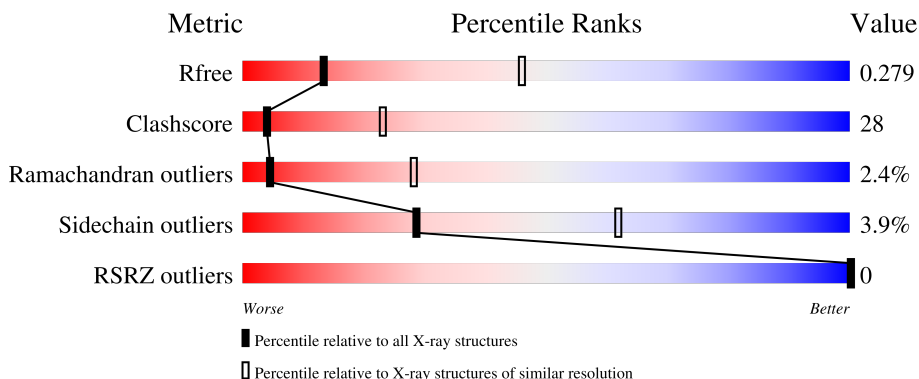
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	180053	1466 (3.20-3.20)
Clashscore	190562	1573 (3.20-3.20)
Ramachandran outliers	187476	1548 (3.20-3.20)
Sidechain outliers	187428	1547 (3.20-3.20)
RSRZ outliers	180081	1466 (3.20-3.20)

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 $\leq 5\%$. 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	565	 49% 40% •• 7%
1	B	565	 50% 39% 5% 6%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8008 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 Phenylpyruvate decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	523	Total	C	N	O	S	0	0	0
			3884	2459	678	724	23			
1	B	530	Total	C	N	O	S	0	0	0
			3935	2488	687	737	23			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP P51852
A	-18	GLY	-	expression tag	UNP P51852
A	-17	SER	-	expression tag	UNP P51852
A	-16	SER	-	expression tag	UNP P51852
A	-15	HIS	-	expression tag	UNP P51852
A	-14	HIS	-	expression tag	UNP P51852
A	-13	HIS	-	expression tag	UNP P51852
A	-12	HIS	-	expression tag	UNP P51852
A	-11	HIS	-	expression tag	UNP P51852
A	-10	HIS	-	expression tag	UNP P51852
A	-9	SER	-	expression tag	UNP P51852
A	-8	SER	-	expression tag	UNP P51852
A	-7	GLY	-	expression tag	UNP P51852
A	-6	LEU	-	expression tag	UNP P51852
A	-5	VAL	-	expression tag	UNP P51852
A	-4	PRO	-	expression tag	UNP P51852
A	-3	ARG	-	expression tag	UNP P51852
A	-2	GLY	-	expression tag	UNP P51852
A	-1	SER	-	expression tag	UNP P51852
A	0	HIS	-	expression tag	UNP P51852
A	155	GLN	LEU	expression tag	UNP P51852
A	327	ARG	GLY	SEE REMARK 999	UNP P51852
B	-19	MET	-	expression tag	UNP P51852
B	-18	GLY	-	expression tag	UNP P51852
B	-17	SER	-	expression tag	UNP P51852

Continued on next page...

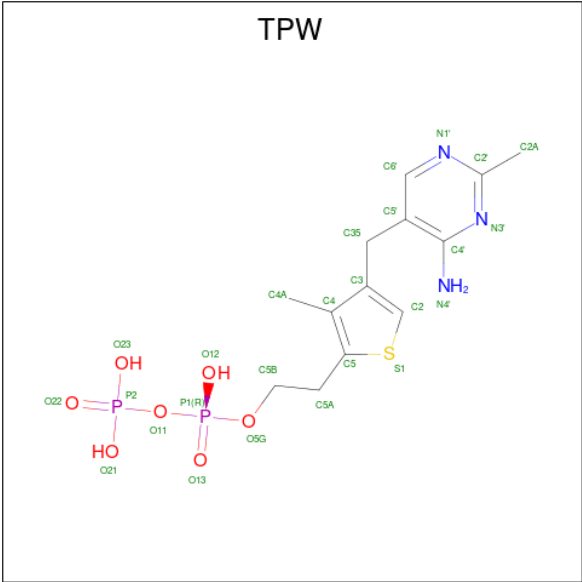
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	-	expression tag	UNP P51852
B	-15	HIS	-	expression tag	UNP P51852
B	-14	HIS	-	expression tag	UNP P51852
B	-13	HIS	-	expression tag	UNP P51852
B	-12	HIS	-	expression tag	UNP P51852
B	-11	HIS	-	expression tag	UNP P51852
B	-10	HIS	-	expression tag	UNP P51852
B	-9	SER	-	expression tag	UNP P51852
B	-8	SER	-	expression tag	UNP P51852
B	-7	GLY	-	expression tag	UNP P51852
B	-6	LEU	-	expression tag	UNP P51852
B	-5	VAL	-	expression tag	UNP P51852
B	-4	PRO	-	expression tag	UNP P51852
B	-3	ARG	-	expression tag	UNP P51852
B	-2	GLY	-	expression tag	UNP P51852
B	-1	SER	-	expression tag	UNP P51852
B	0	HIS	-	expression tag	UNP P51852
B	155	GLN	LEU	expression tag	UNP P51852
B	327	ARG	GLY	SEE REMARK 999	UNP P51852

- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0

- Molecule 3 is 2-{4-[(4-AMINO-2-METHYLPYRIMIDIN-5-YL)METHYL]-3-METHYLT HIOPHEN-2-YL}ETHYL TRIHYDROGEN DIPHOSPHATE (CCD ID: TPW) (formula: C₁₃H₁₉N₃O₇P₂S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	S	0	0
			26	13	3	7	2	1		
3	B	1	Total	C	N	O	P	S	0	0
			26	13	3	7	2	1		

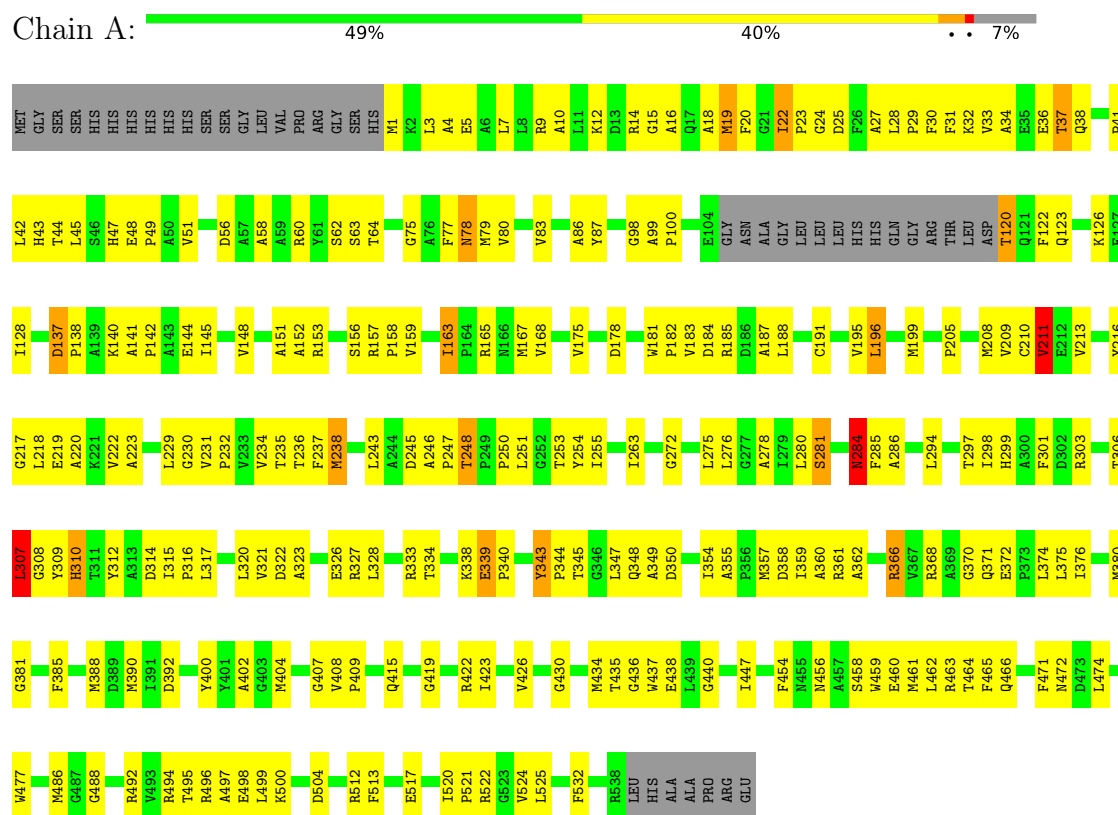
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	56	Total	O	0	0
			56	56		
4	B	79	Total	O	0	0
			79	79		

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: Phenylpyruvate decarboxylase



• Molecule 1: Phenylpyruvate decarboxylase



A483	G407	P330	M238
K486	P409	R333	G239
G487	A414	G337	G241
V491	Q415	K338	L242
R492	S418	E339	A246
T495	K421	P340	P249
R496	R422	H341	
E498	I423	P344	V257
A997	L424		A258
L499	T425	L347	D260
K500	F432	Q348	G269
	Q433	A349	A261
D504	M434	D350	E262
	T435	G351	I263
R510	G436	E352	T264
G511	W437	P353	
R512	E438	I354	D271
	L439	A355	G272
E517	G440	P356	
A518	R444	M357	L275
M519		D358	L276
T520		P359	G277
P521		A360	A278
V524	I447	R361	I279
	D448	A362	L280
T528	P449	V363	S281
V533	I450	R366	D282
Q534	V451	V367	T283
G535	I452	R368	N284
	L453		
	F454	Q371	Q289
R538	N455	K372	
LEU	A456	P373	K296
HIS	A457	L374	T297
ALA	S458	L375	L298
ALA	W459	L376	
PRO	E460		F301
ARG	M461		
GLU	L462	M380	T306
	R463		L307
	T464	A387	G308
	F465	M388	Y309
	Q466	D389	H310
		M390	T311
	S469		Y312
	A470	A393	A313
	F471	G394	D314
	M472	L395	I315
	D473		P316
	L474	Y400	L317
		Y401	
	W477	A402	V321
	R478	G403	D322
	F479	M404	A323
		G405	
	M482	F406	R327

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	100.54Å 179.81Å 121.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.20 50.00 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-3.20) 96.6 (50.00-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.55 (at 3.18Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.176 , 0.289 0.171 , 0.279	Depositor DCC
R_{free} test set	911 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	62.0	Xtriage
Anisotropy	0.146	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 47.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.030 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8008	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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: TPW, MG

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	0.46	0/3959	0.99	14/5383 (0.3%)
1	B	0.45	0/4010	0.99	18/5452 (0.3%)
All	All	0.46	0/7969	0.99	32/10835 (0.3%)

There are no bond length outliers.

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	163	ILE	N-CA-C	9.27	118.15	107.77
1	B	105	GLY	N-CA-C	8.39	133.06	113.18
1	B	459	TRP	N-CA-C	-7.52	99.11	109.71
1	B	106	ASN	N-CA-C	7.12	120.51	108.90
1	B	126	LYS	N-CA-C	-6.69	104.02	111.71

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	3884	0	3876	224	0
1	B	3935	0	3921	227	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	26	0	16	2	0
3	B	26	0	16	2	0
4	A	56	0	0	6	0
4	B	79	0	0	2	0
All	All	8008	0	7829	438	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 28.

The worst 5 of 438 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:18:ALA:HA	1:A:41:PRO:HD2	1.41	1.00
1:B:388:MET:HE3	1:B:528:THR:HG21	1.52	0.91
1:A:120:THR:HG22	1:A:123:GLN:H	1.35	0.91
1:A:14:ARG:HD3	1:A:153:ARG:HD2	1.54	0.89
1:B:199:MET:HE2	1:B:199:MET:HA	1.52	0.89

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	519/565 (92%)	443 (85%)	64 (12%)	12 (2%)	5	29
1	B	526/565 (93%)	439 (84%)	74 (14%)	13 (2%)	4	27
All	All	1045/1130 (92%)	882 (84%)	138 (13%)	25 (2%)	4	28

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	105	GLY
1	B	340	PRO
1	B	349	ALA
1	B	407	GLY
1	A	38	GLN

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	387/426 (91%)	374 (97%)	13 (3%)	32	64
1	B	392/426 (92%)	375 (96%)	17 (4%)	26	59
All	All	779/852 (91%)	749 (96%)	30 (4%)	28	62

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	110	LEU
1	B	469	SER
1	B	298	ILE
1	B	524	VAL
1	B	425	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	78	ASN
1	B	106	ASN
1	B	433	GLN
1	B	299	HIS
1	B	371	GLN

5.3.3 RNA ⓘ

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 oligosaccharides 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	TPW	A	1002	2	25,27,27	1.92	9 (36%)	37,40,40	3.30	12 (32%)
3	TPW	B	1001	2	25,27,27	2.05	6 (24%)	37,40,40	3.37	14 (37%)

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	TPW	A	1002	2	-	1/17/17/17	0/2/2/2
3	TPW	B	1001	2	-	1/17/17/17	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1001	TPW	P1-O11	6.63	1.66	1.59
3	A	1002	TPW	P1-O11	4.01	1.63	1.59
3	A	1002	TPW	C35-C5'	3.82	1.56	1.51
3	B	1001	TPW	C35-C5'	3.22	1.56	1.51
3	B	1001	TPW	P1-O13	2.82	1.60	1.50

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	A	1002	TPW	C2-C3-C4	10.85	119.57	111.92
3	B	1001	TPW	C2-C3-C4	10.72	119.48	111.92
3	B	1001	TPW	C3-C2-S1	-9.84	102.39	113.46
3	A	1002	TPW	C3-C2-S1	-9.32	102.97	113.46
3	B	1001	TPW	C2-S1-C5	7.76	99.74	92.41

There are no chirality outliers.

All (2) torsion outliers are listed below:

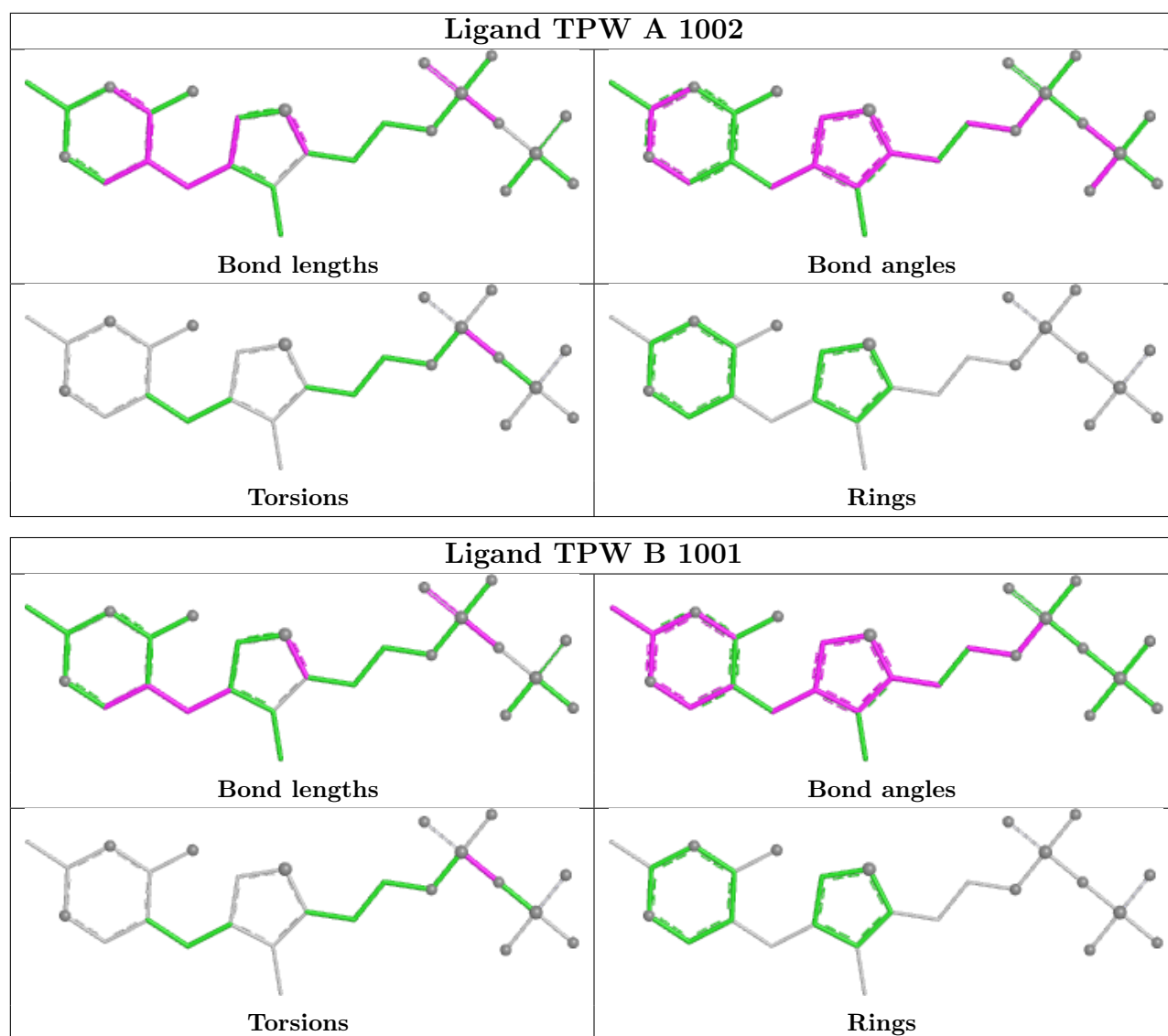
Mol	Chain	Res	Type	Atoms
3	A	1002	TPW	P2-O11-P1-O5G
3	B	1001	TPW	P2-O11-P1-O5G

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1002	TPW	2	0
3	B	1001	TPW	2	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 [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	523/565 (92%)	-0.93	0 100 100	10, 38, 66, 94	0
1	B	530/565 (93%)	-0.86	0 100 100	9, 39, 77, 99	0
All	All	1053/1130 (93%)	-0.90	0 100 100	9, 38, 72, 99	0

There are no RSRZ outliers to report.

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 oligosaccharides 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.

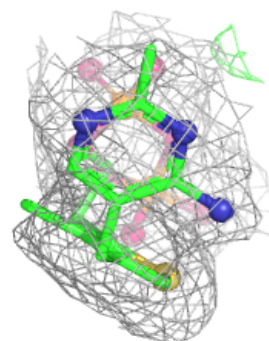
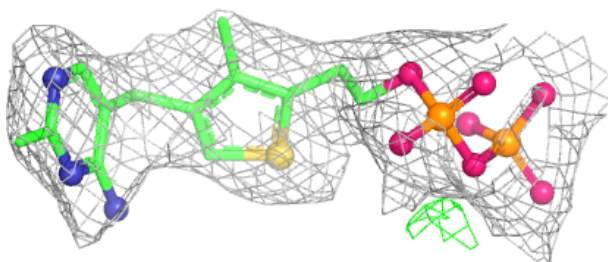
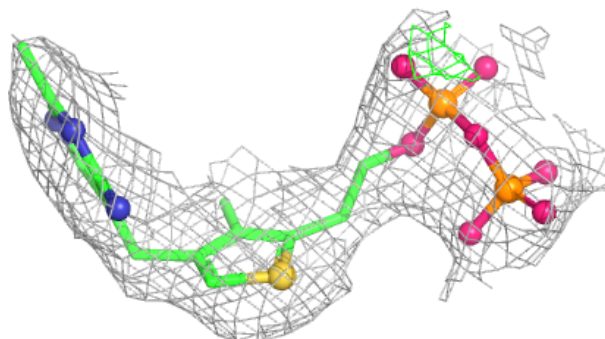
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	TPW	A	1002	26/26	0.98	0.06	10,23,36,41	0
3	TPW	B	1001	26/26	0.98	0.05	9,32,47,50	0
2	MG	A	2002	1/1	0.99	0.02	50,50,50,50	0
2	MG	B	2001	1/1	0.99	0.02	22,22,22,22	0

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.

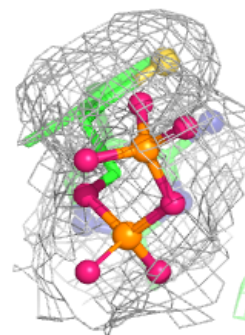
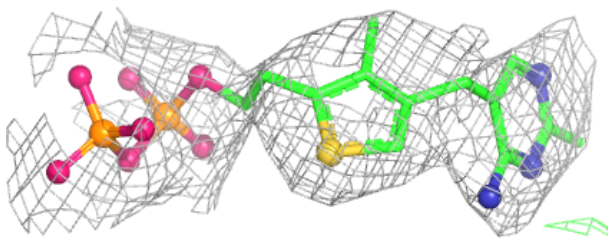
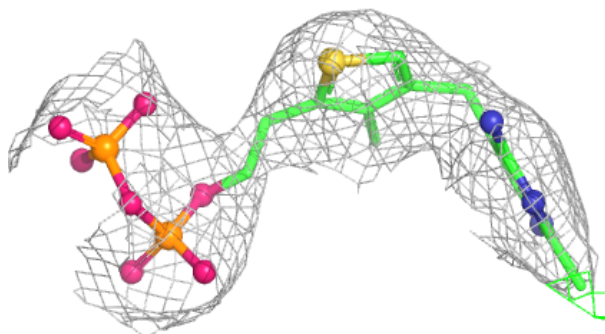
Electron density around TPW A 1002:

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



Electron density around TPW B 1001:

$2mF_o-DF_c$ (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.