wwPDB X-ray Structure Validation Summary Report (i)

Aug 19, 2023 – 11:46 PM EDT

PDB ID	:	2G25
Title	:	E. Coli Pyruvate Dehydrogenase Phosphonolactylthiamin Diphosphate Complex
Authors	:	Furey, W.; Arjunan, P.; Chandrasekhar, K.
Deposited on	:	2006-02-15
Resolution	:	2.10 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 (i) 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 (i)) were used in the production of this report:

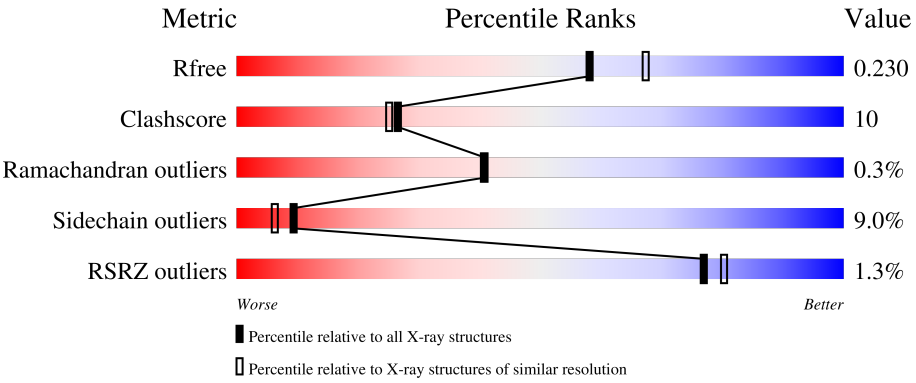
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	886	2% 71% 20% 6%
1	B	886	% 72% 19% 6%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13841 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.

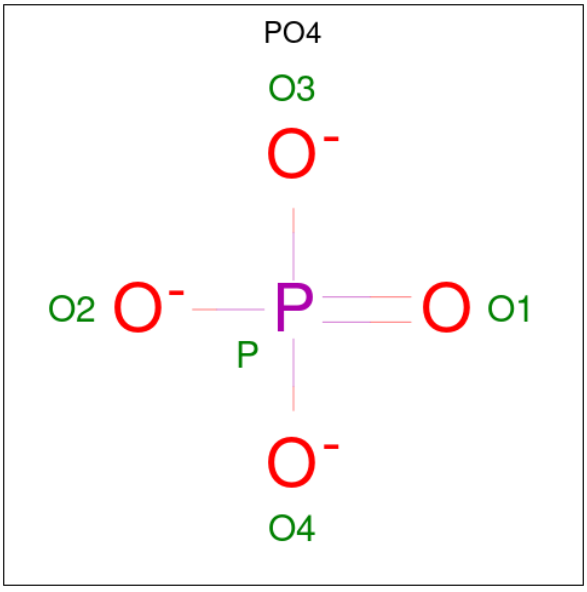
- Molecule 1 is a protein called Pyruvate dehydrogenase E1 component.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	831	Total	C	N	O	S	0	0	0
			6591	4172	1140	1252	27			
1	B	831	Total	C	N	O	S	0	0	0
			6591	4172	1140	1252	27			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

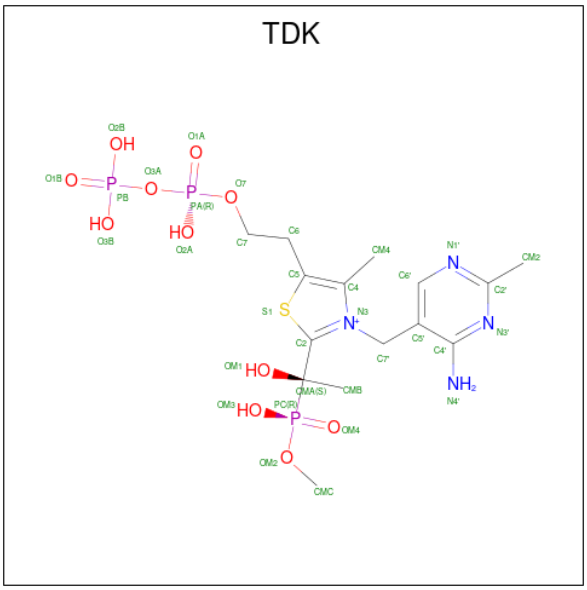
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		
2	B	1	Total	Mg	0	0
			1	1		

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	A	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0

- Molecule 4 is 3-[(4-AMINO-2-METHYLPYRIMIDIN-5-YL)METHYL]-2-{(1S)-1-HYDROXY-1-[(R)-HYDROXY(METHOXY)PHOSPHORYL]ETHYL}-5-(2-{[(S)-HYDROXY(PHOSPHONOOXY)PHOSPHORYL]OXY}ETHYL)-4-METHYL-1,3-THIAZOL-3-IUM (three-letter code: TDK) (formula: $C_{15}H_{26}N_4O_{11}P_3S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O P S 34 15 4 11 3 1	0	0
4	B	1	Total C N O P S 34 15 4 11 3 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	273	Total O 273 273	0	0

Continued on next page...

Continued from previous page...

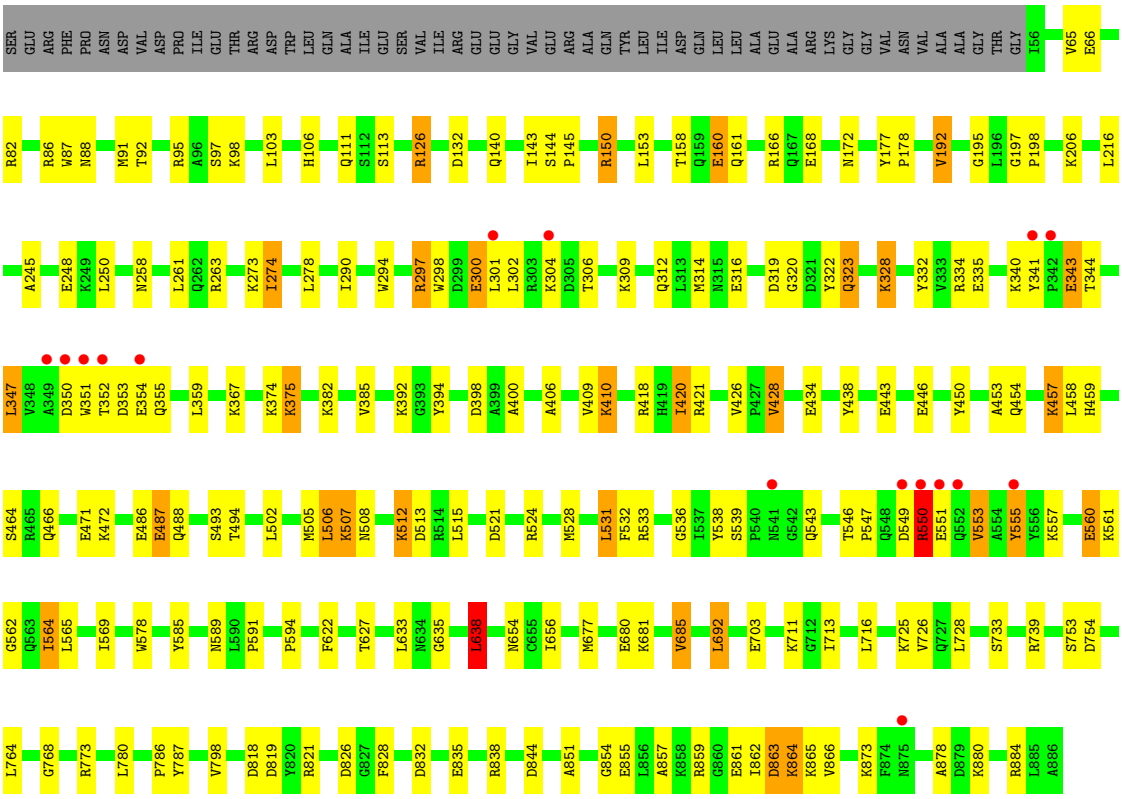
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	291	Total O 291 291	0	0

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.

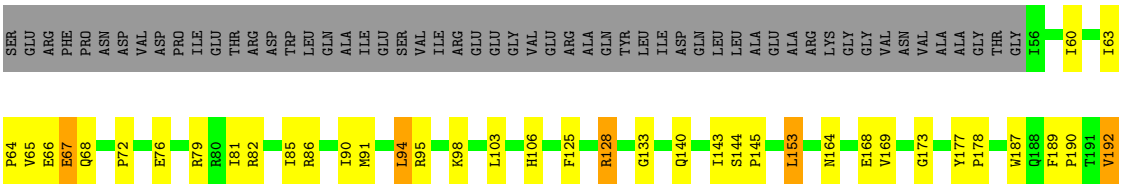
• Molecule 1: Pyruvate dehydrogenase E1 component

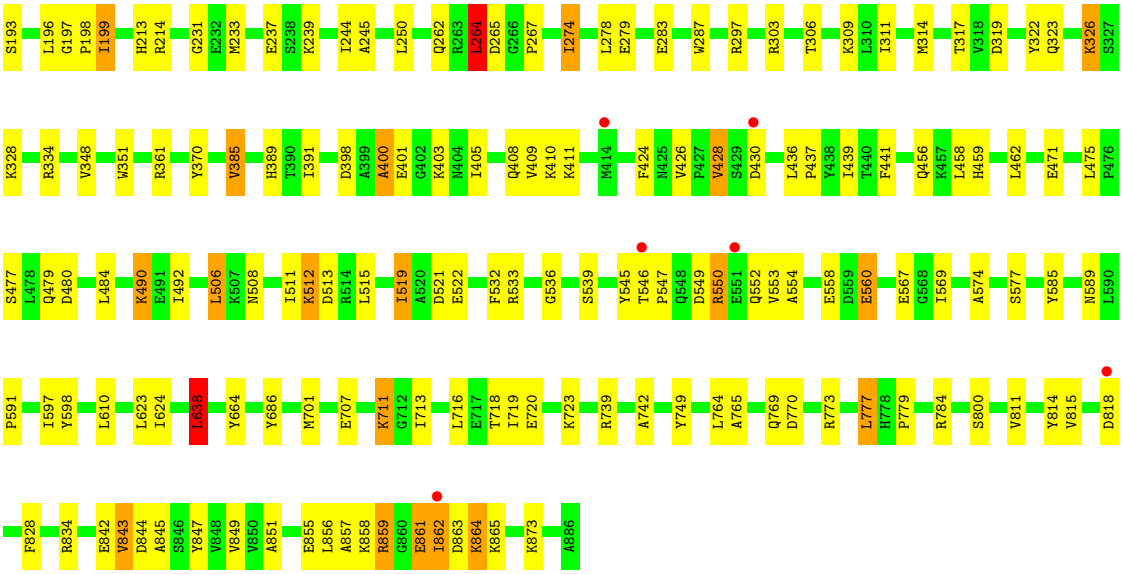
Chain A:



• Molecule 1: Pyruvate dehydrogenase E1 component

Chain B:




S193 L196 G197 P198 I199 H213 R214 G231 E232 M233 E237 S238 K239 I244 A245 L250 Q262 R263 L264 D265 G266 P267 I274 L278 E279 E283 W287 R297 R303 T306 K309 L310 I311 M314 T317 V318 D319 Y322 Q323 K326 S327
K328 R334 V348 W351 R361 Y370 V385 H389 T390 I391 D398 A399 A400 E401 G402 K403 N404 I405 Q408 V409 K410 K411 M414 F424 N425 V426 P427 V428 S429 D430 L436 P437 Y438 I439 T440 F441 Q456 K457 L458 H459 L462 E471 L475 P476
S477 L478 Q479 D480 L484 K490 E491 I492 L506 K507 N508 I511 K512 D513 R514 L515 I519 A520 D521 E522 F532 R533 G536 S539 Y545 T546 P547 Q548 D549 R550 E551 Q552 V553 A554 E558 D559 E560 E567 G568 I569 A574 S577 Y585 N589 L590
P591 I597 Y598 L610 L623 I624 L638 Y664 Y686 M701 E707 K711 G712 I713 L716 E717 T718 I719 E720 K723 R739 A742 Y749 L764 A765 Q769 D770 R773 L777 H778 P779 R784 S800 V811 Y814 V815 D818
F828 R834 E842 V843 D844 A845 S846 Y847 V848 V849 V850 A851 E855 L856 A857 K858 R859 G860 E861 I862 D863 K864 K865 K873 A886

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.57Å 142.51Å 82.09Å 90.00° 102.37° 90.00°	Depositor
Resolution (Å)	8.00 – 2.10 39.18 – 2.10	Depositor EDS
% Data completeness (in resolution range)	94.0 (8.00-2.10) 93.7 (39.18-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.85 (at 2.10Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.185 , 0.244 0.178 , 0.230	Depositor DCC
R_{free} test set	5035 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	10.8	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.028 for l,-k,h	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	13841	wwPDB-VP
Average B, all atoms (Å^2)	10.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% 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: MG, TDK, PO4

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	0.37	0/6741	0.61	1/9114 (0.0%)
1	B	0.38	0/6741	0.62	2/9114 (0.0%)
All	All	0.37	0/13482	0.62	3/18228 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	638	LEU	CA-CB-CG	5.31	127.52	115.30
1	B	638	LEU	CA-CB-CG	5.23	127.34	115.30
1	B	264	LEU	CA-CB-CG	5.04	126.88	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	598	TYR	Sidechain
1	B	686	TYR	Sidechain

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	6591	0	6422	143	0
1	B	6591	0	6422	135	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	10	0	0	0	0
3	B	15	0	0	2	0
4	A	34	0	22	2	0
4	B	34	0	22	4	0
5	A	273	0	0	6	0
5	B	291	0	0	4	0
All	All	13841	0	12888	270	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 10.

The worst 5 of 270 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:550:ARG:HA	1:A:550:ARG:HH11	1.18	1.06
1:A:309:LYS:HG3	1:A:344:THR:HG23	1.52	0.91
1:A:549:ASP:HB3	1:A:555:TYR:HA	1.53	0.91
1:A:550:ARG:HH11	1:A:550:ARG:CA	1.87	0.88
1:A:550:ARG:HA	1:A:550:ARG:NH1	1.89	0.86

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	829/886 (94%)	785 (95%)	40 (5%)	4 (0%)	29 26
1	B	829/886 (94%)	790 (95%)	38 (5%)	1 (0%)	51 54
All	All	1658/1772 (94%)	1575 (95%)	78 (5%)	5 (0%)	41 41

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	553	VAL
1	A	347	LEU
1	A	550	ARG
1	A	398	ASP
1	B	400	ALA

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	691/735 (94%)	624 (90%)	67 (10%)	8 5
1	B	691/735 (94%)	634 (92%)	57 (8%)	11 8
All	All	1382/1470 (94%)	1258 (91%)	124 (9%)	9 6

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	703	GLU
1	B	711	LYS
1	B	94	LEU
1	B	638	LEU
1	B	858	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	106	HIS
1	B	448	HIS
1	B	737	HIS
1	A	419	HIS
1	A	466	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	PO4	A	890	-	4,4,4	0.88	0	6,6,6	0.83	0
3	PO4	A	889	-	4,4,4	0.85	0	6,6,6	0.86	0
4	TDK	A	887	2	29,35,35	1.74	4 (13%)	36,55,55	1.72	8 (22%)
3	PO4	B	890	-	4,4,4	1.14	0	6,6,6	0.77	0
4	TDK	B	887	2	29,35,35	1.67	5 (17%)	36,55,55	1.77	8 (22%)
3	PO4	B	891	-	4,4,4	1.01	0	6,6,6	0.84	0
3	PO4	B	889	-	4,4,4	1.10	0	6,6,6	0.67	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	TDK	A	887	2	-	4/26/35/35	0/2/2/2
4	TDK	B	887	2	-	4/26/35/35	0/2/2/2

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	887	TDK	C6-C5	-4.70	1.48	1.50
4	B	887	TDK	PC-OM3	-4.48	1.47	1.56
4	A	887	TDK	PC-OM3	-4.42	1.47	1.56
4	B	887	TDK	C6-C5	-3.58	1.49	1.50
4	B	887	TDK	PB-O1B	3.15	1.60	1.50

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	887	TDK	OM4-PC-CMA	-4.85	108.59	113.53
4	A	887	TDK	OM4-PC-CMA	-4.48	108.97	113.53
4	A	887	TDK	C6-C5-C4	4.36	130.94	127.43
4	B	887	TDK	C6-C5-C4	4.34	130.92	127.43
4	B	887	TDK	PA-O3A-PB	-3.07	122.28	132.83

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	887	TDK	PA-O3A-PB-O2B
4	A	887	TDK	C4-C5-C6-C7
4	B	887	TDK	PA-O3A-PB-O2B
4	B	887	TDK	C4-C5-C6-C7
4	A	887	TDK	CMC-OM2-PC-OM4

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	887	TDK	2	0
3	B	890	PO4	1	0
4	B	887	TDK	4	0

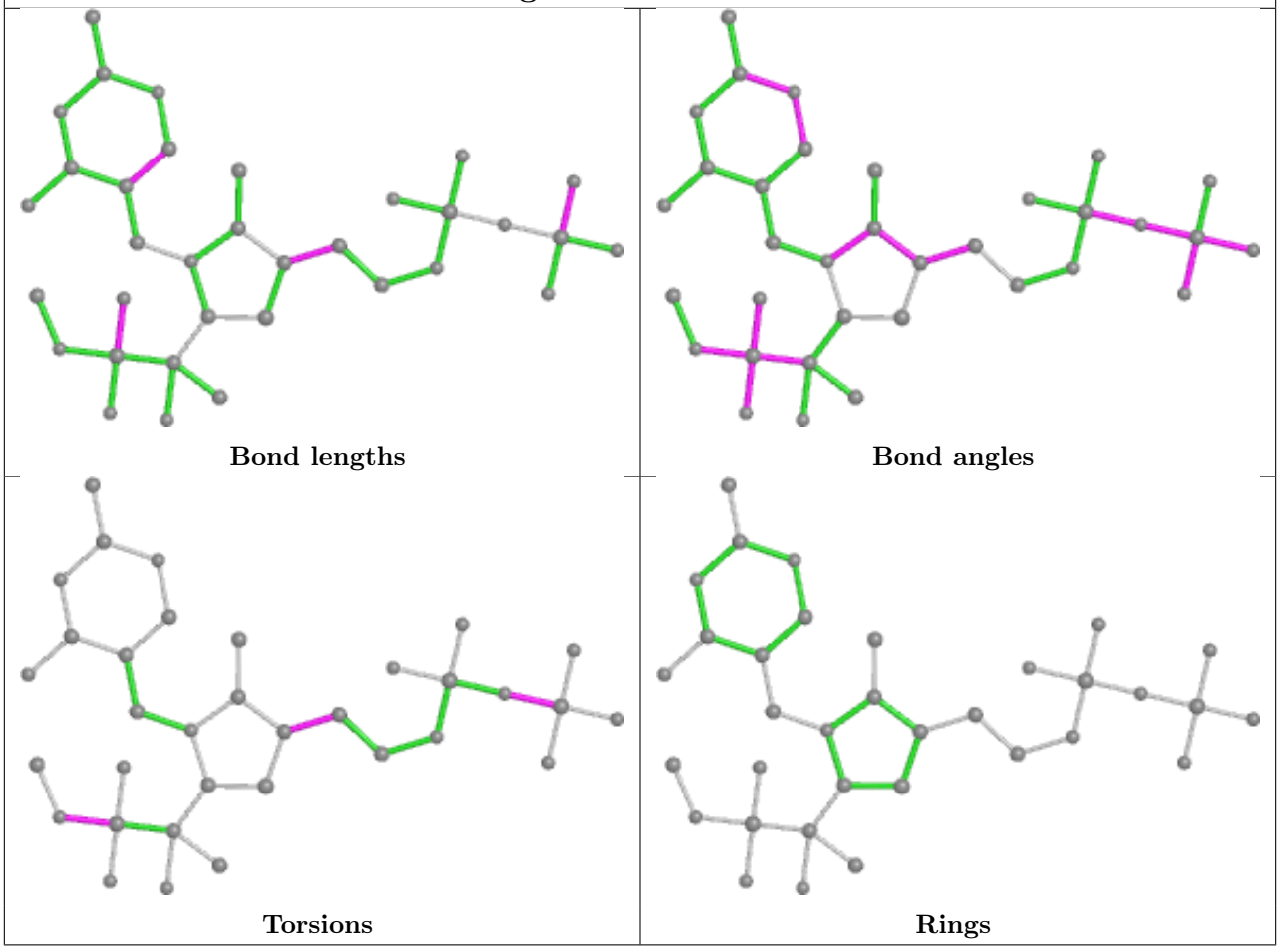
Continued on next page...

Continued from previous page...

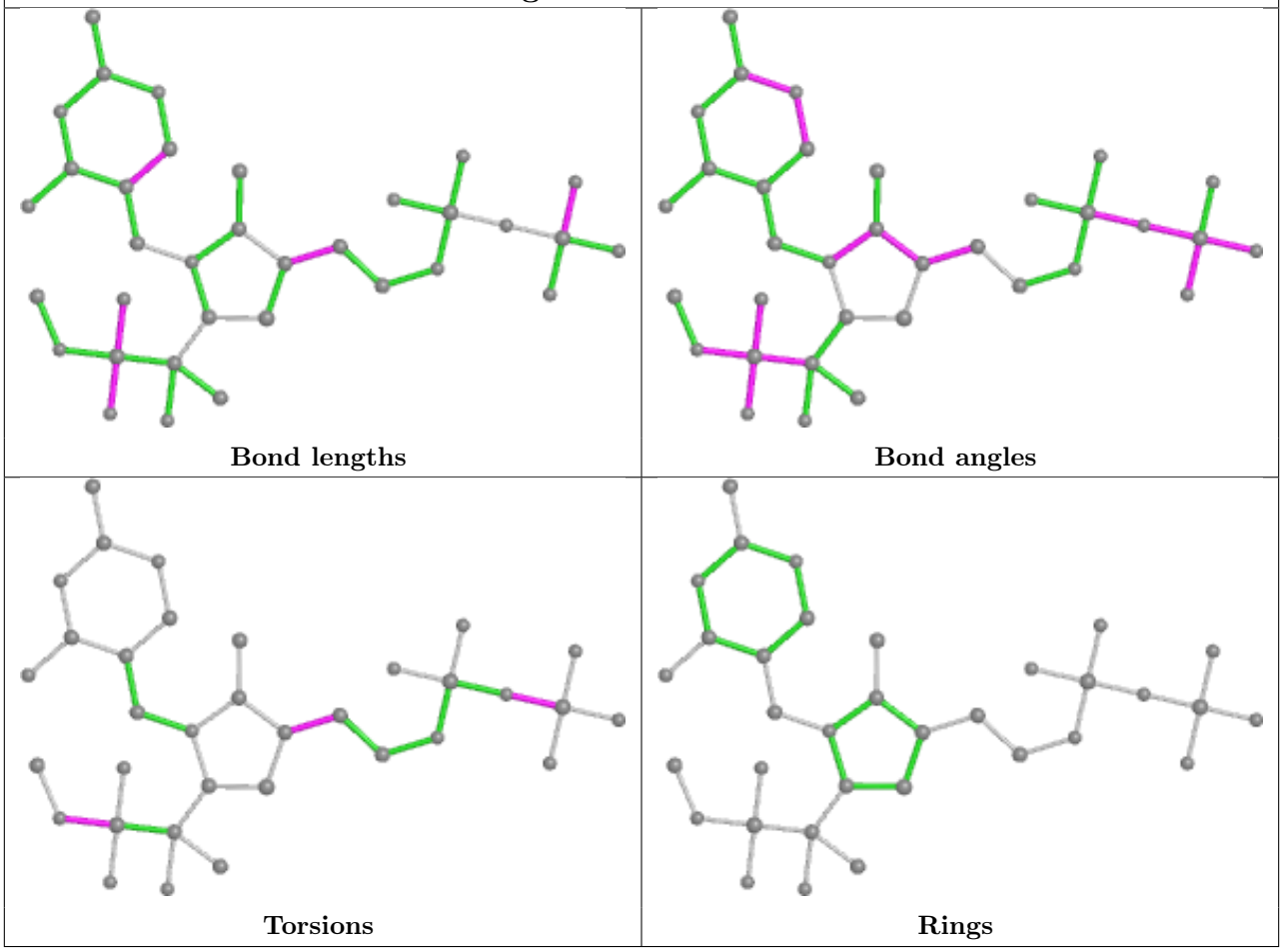
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	889	PO4	1	0

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.

Ligand TDK A 887



Ligand TDK B 887



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^{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>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	831/886 (93%)	-0.20	16 (1%) 66 71	2, 8, 29, 43	0
1	B	831/886 (93%)	-0.25	6 (0%) 87 89	2, 6, 24, 37	0
All	All	1662/1772 (93%)	-0.23	22 (1%) 77 80	2, 7, 28, 43	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	541	ASN	3.4
1	A	351	TRP	3.2
1	A	552	GLN	3.1
1	A	352	THR	2.9
1	A	341	TYR	2.9

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

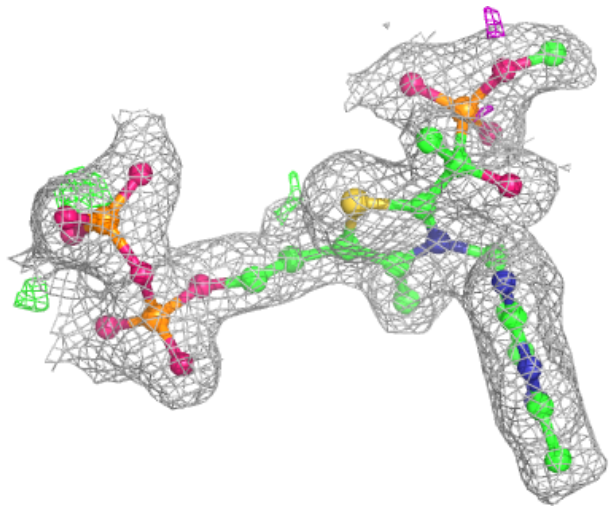
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

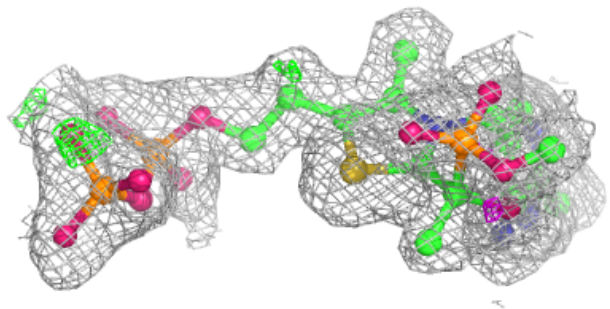
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	B	888	1/1	0.95	0.15	13,13,13,13	0
3	PO4	A	890	5/5	0.96	0.10	23,24,25,28	0
3	PO4	B	889	5/5	0.96	0.11	15,18,20,20	0
3	PO4	B	890	5/5	0.96	0.14	23,24,25,27	0
3	PO4	A	889	5/5	0.97	0.10	12,13,14,16	0
2	MG	A	888	1/1	0.97	0.16	10,10,10,10	0
3	PO4	B	891	5/5	0.97	0.12	24,24,27,28	0
4	TDK	A	887	34/34	0.97	0.11	2,3,13,14	0
4	TDK	B	887	34/34	0.97	0.10	4,7,15,16	0

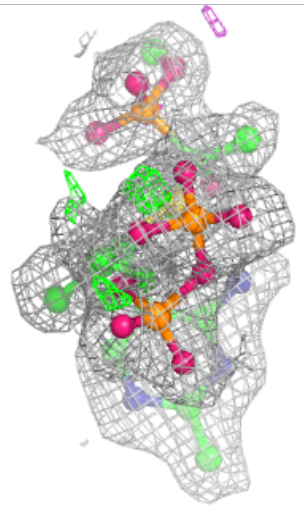
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around TDK A 887:

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

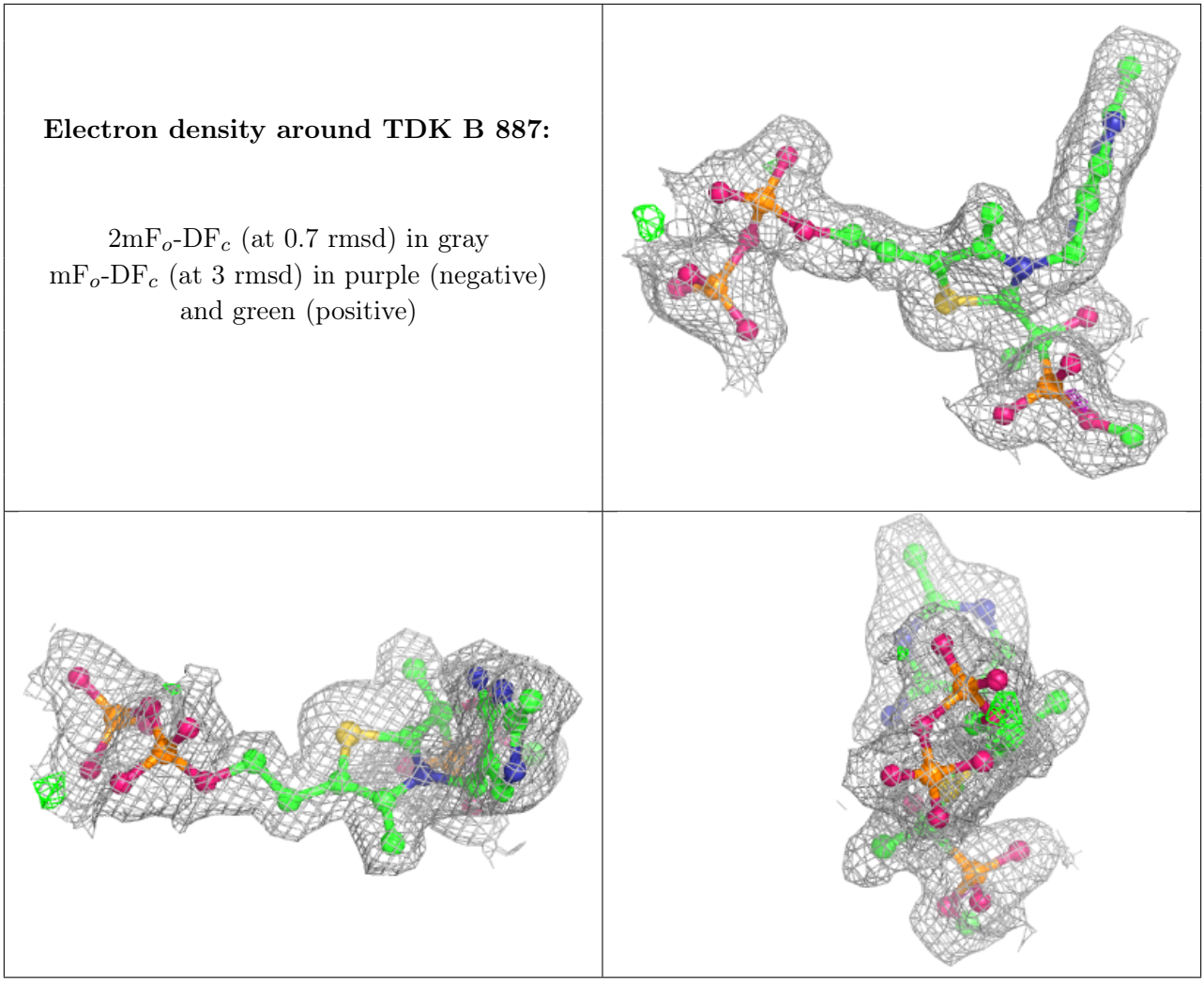






Electron density around TDK B 887:

$2mF_o$-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

There are no such residues in this entry.