



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 13, 2020 – 04:12 PM BST

PDB ID : 4FXJ  
Title : Structure of M2 pyruvate kinase in complex with phenylalanine  
Authors : Walkinshaw, M.D.; Morgan, H.P.  
Deposited on : 2012-07-03  
Resolution : 2.90 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](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 following versions of software and data (see [references ⓘ](#)) 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 : 2.14.4.dev1  
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) : 2.14.4.dev1

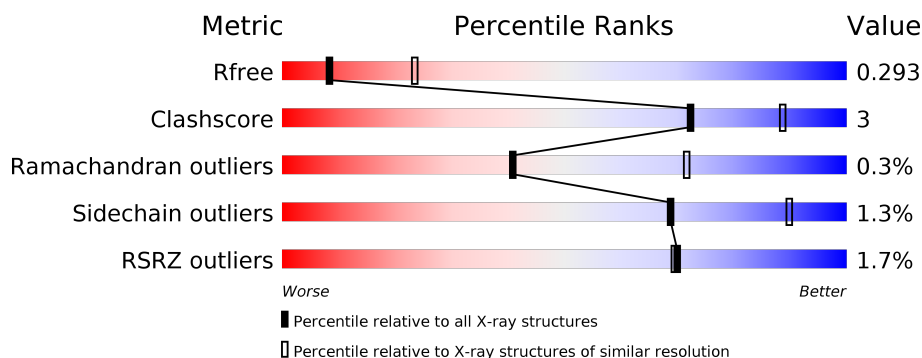
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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  $\geq 3$ , 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	551	 82% 8% 10%
1	B	551	 66% 9% 25%
1	C	551	 84% 8% 9%
1	D	551	 4% 80% 12% 8%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14897 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 kinase isozymes M1/M2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	498	Total	C	N	O	S	0	0	0
			3810	2397	675	714	24			
1	B	413	Total	C	N	O	S	0	0	0
			3169	1991	567	590	21			
1	C	504	Total	C	N	O	S	0	0	0
			3860	2430	681	725	24			
1	D	508	Total	C	N	O	S	0	0	0
			3882	2442	685	731	24			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP P14618
A	-18	GLY	-	EXPRESSION TAG	UNP P14618
A	-17	SER	-	EXPRESSION TAG	UNP P14618
A	-16	SER	-	EXPRESSION TAG	UNP P14618
A	-15	HIS	-	EXPRESSION TAG	UNP P14618
A	-14	HIS	-	EXPRESSION TAG	UNP P14618
A	-13	HIS	-	EXPRESSION TAG	UNP P14618
A	-12	HIS	-	EXPRESSION TAG	UNP P14618
A	-11	HIS	-	EXPRESSION TAG	UNP P14618
A	-10	HIS	-	EXPRESSION TAG	UNP P14618
A	-9	SER	-	EXPRESSION TAG	UNP P14618
A	-8	SER	-	EXPRESSION TAG	UNP P14618
A	-7	GLY	-	EXPRESSION TAG	UNP P14618
A	-6	LEU	-	EXPRESSION TAG	UNP P14618
A	-5	VAL	-	EXPRESSION TAG	UNP P14618
A	-4	PRO	-	EXPRESSION TAG	UNP P14618
A	-3	ARG	-	EXPRESSION TAG	UNP P14618
A	-2	GLY	-	EXPRESSION TAG	UNP P14618
A	-1	SER	-	EXPRESSION TAG	UNP P14618
A	0	HIS	-	EXPRESSION TAG	UNP P14618
A	379	ASN	HIS	SEE REMARK 999	UNP P14618

*Continued on next page...*

*Continued from previous page...*

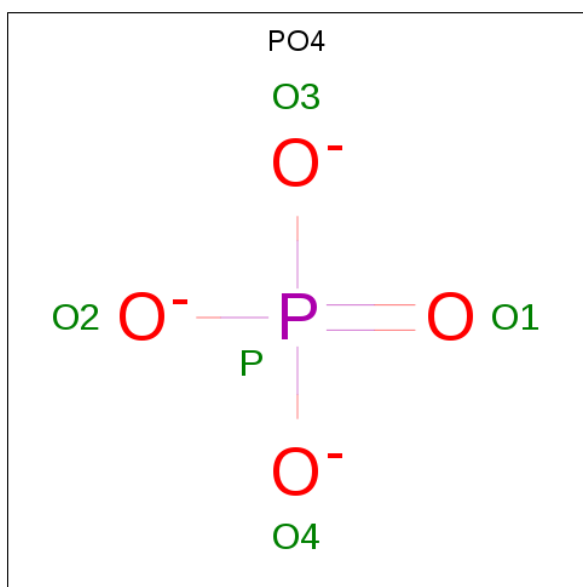
Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	EXPRESSION TAG	UNP P14618
B	-18	GLY	-	EXPRESSION TAG	UNP P14618
B	-17	SER	-	EXPRESSION TAG	UNP P14618
B	-16	SER	-	EXPRESSION TAG	UNP P14618
B	-15	HIS	-	EXPRESSION TAG	UNP P14618
B	-14	HIS	-	EXPRESSION TAG	UNP P14618
B	-13	HIS	-	EXPRESSION TAG	UNP P14618
B	-12	HIS	-	EXPRESSION TAG	UNP P14618
B	-11	HIS	-	EXPRESSION TAG	UNP P14618
B	-10	HIS	-	EXPRESSION TAG	UNP P14618
B	-9	SER	-	EXPRESSION TAG	UNP P14618
B	-8	SER	-	EXPRESSION TAG	UNP P14618
B	-7	GLY	-	EXPRESSION TAG	UNP P14618
B	-6	LEU	-	EXPRESSION TAG	UNP P14618
B	-5	VAL	-	EXPRESSION TAG	UNP P14618
B	-4	PRO	-	EXPRESSION TAG	UNP P14618
B	-3	ARG	-	EXPRESSION TAG	UNP P14618
B	-2	GLY	-	EXPRESSION TAG	UNP P14618
B	-1	SER	-	EXPRESSION TAG	UNP P14618
B	0	HIS	-	EXPRESSION TAG	UNP P14618
B	379	ASN	HIS	SEE REMARK 999	UNP P14618
C	-19	MET	-	EXPRESSION TAG	UNP P14618
C	-18	GLY	-	EXPRESSION TAG	UNP P14618
C	-17	SER	-	EXPRESSION TAG	UNP P14618
C	-16	SER	-	EXPRESSION TAG	UNP P14618
C	-15	HIS	-	EXPRESSION TAG	UNP P14618
C	-14	HIS	-	EXPRESSION TAG	UNP P14618
C	-13	HIS	-	EXPRESSION TAG	UNP P14618
C	-12	HIS	-	EXPRESSION TAG	UNP P14618
C	-11	HIS	-	EXPRESSION TAG	UNP P14618
C	-10	HIS	-	EXPRESSION TAG	UNP P14618
C	-9	SER	-	EXPRESSION TAG	UNP P14618
C	-8	SER	-	EXPRESSION TAG	UNP P14618
C	-7	GLY	-	EXPRESSION TAG	UNP P14618
C	-6	LEU	-	EXPRESSION TAG	UNP P14618
C	-5	VAL	-	EXPRESSION TAG	UNP P14618
C	-4	PRO	-	EXPRESSION TAG	UNP P14618
C	-3	ARG	-	EXPRESSION TAG	UNP P14618
C	-2	GLY	-	EXPRESSION TAG	UNP P14618
C	-1	SER	-	EXPRESSION TAG	UNP P14618
C	0	HIS	-	EXPRESSION TAG	UNP P14618
C	379	ASN	HIS	SEE REMARK 999	UNP P14618

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-19	MET	-	EXPRESSION TAG	UNP P14618
D	-18	GLY	-	EXPRESSION TAG	UNP P14618
D	-17	SER	-	EXPRESSION TAG	UNP P14618
D	-16	SER	-	EXPRESSION TAG	UNP P14618
D	-15	HIS	-	EXPRESSION TAG	UNP P14618
D	-14	HIS	-	EXPRESSION TAG	UNP P14618
D	-13	HIS	-	EXPRESSION TAG	UNP P14618
D	-12	HIS	-	EXPRESSION TAG	UNP P14618
D	-11	HIS	-	EXPRESSION TAG	UNP P14618
D	-10	HIS	-	EXPRESSION TAG	UNP P14618
D	-9	SER	-	EXPRESSION TAG	UNP P14618
D	-8	SER	-	EXPRESSION TAG	UNP P14618
D	-7	GLY	-	EXPRESSION TAG	UNP P14618
D	-6	LEU	-	EXPRESSION TAG	UNP P14618
D	-5	VAL	-	EXPRESSION TAG	UNP P14618
D	-4	PRO	-	EXPRESSION TAG	UNP P14618
D	-3	ARG	-	EXPRESSION TAG	UNP P14618
D	-2	GLY	-	EXPRESSION TAG	UNP P14618
D	-1	SER	-	EXPRESSION TAG	UNP P14618
D	0	HIS	-	EXPRESSION TAG	UNP P14618
D	379	ASN	HIS	SEE REMARK 999	UNP P14618

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



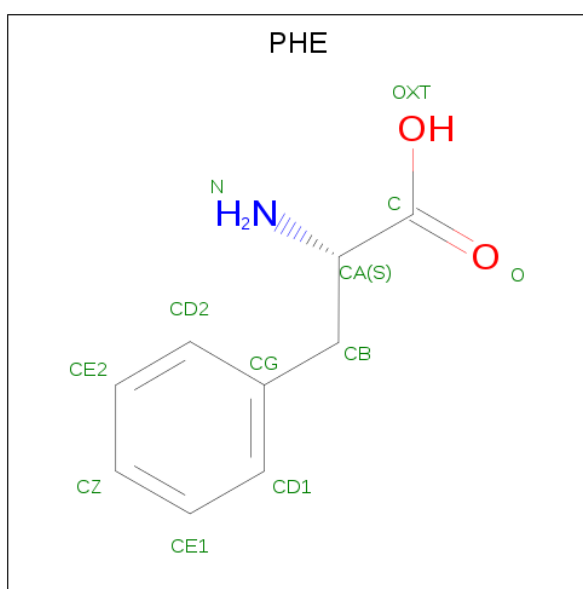
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		

Continued on next page...

*Continued from previous page...*

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 PHENYLALANINE (three-letter code: PHE) (formula:  $C_9H_{11}NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			12	9	1	2		
3	B	1	Total	C	N	O	0	0
			12	9	1	2		
3	C	1	Total	C	N	O	0	0
			12	9	1	2		
3	D	1	Total	C	N	O	0	0
			12	9	1	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	35	Total	O	0	0
			35	35		

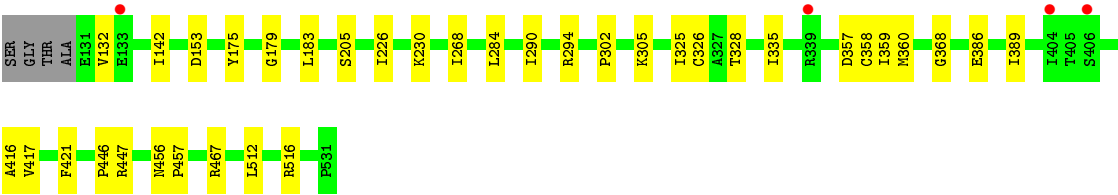
*Continued on next page...*

*Continued from previous page...*

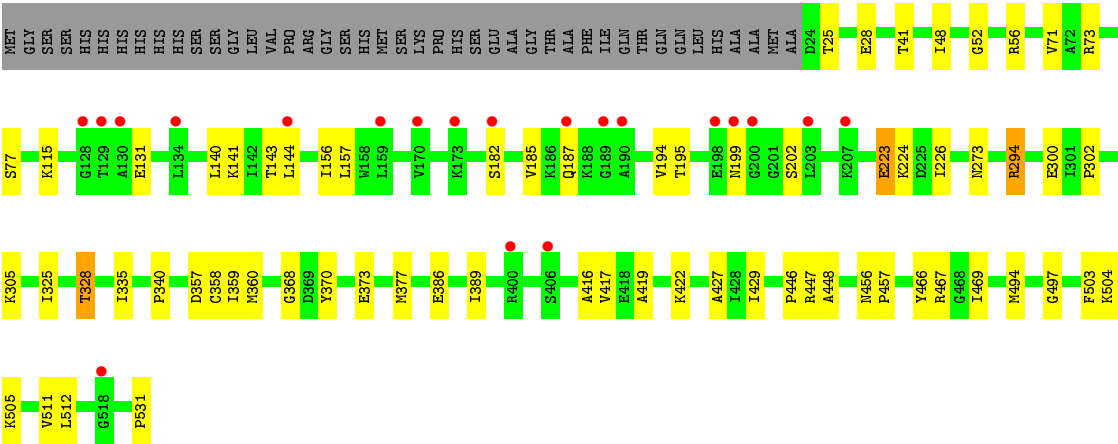
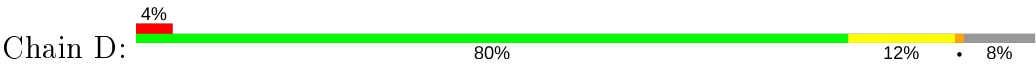
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	26	Total 26	O 26	0	0
4	C	16	Total 16	O 16	0	0
4	D	26	Total 26	O 26	0	0







● Molecule 1: Pyruvate kinase isozymes M1/M2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.26Å 70.60Å 167.60Å 90.00° 105.72° 90.00°	Depositor
Resolution (Å)	56.14 – 2.90 56.14 – 2.90	Depositor EDS
% Data completeness (in resolution range)	89.4 (56.14-2.90) 89.4 (56.14-2.90)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	0.19	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.246 , 0.297 0.240 , 0.293	Depositor DCC
$R_{free}$ test set	2209 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.1	Xtriage
Anisotropy	0.480	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 37.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.016 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	14897	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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: 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	0.30	0/3869	0.44	0/5221
1	B	0.30	0/3223	0.44	0/4356
1	C	0.30	0/3921	0.43	0/5294
1	D	0.30	0/3944	0.44	0/5327
All	All	0.30	0/14957	0.44	0/20198

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	3810	0	3893	21	0
1	B	3169	0	3221	28	0
1	C	3860	0	3943	20	0
1	D	3882	0	3964	32	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
3	A	12	0	8	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	12	0	8	0	0
3	C	12	0	8	0	0
3	D	12	0	8	0	0
4	A	35	0	0	0	0
4	B	26	0	0	0	0
4	C	16	0	0	0	0
4	D	26	0	0	0	0
All	All	14897	0	15053	100	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 3.

The worst 5 of 100 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:B:405:THR:CB	1:B:406:SER:HA	1.79	1.12
1:B:405:THR:HB	1:B:406:SER:CA	1.88	1.02
1:B:405:THR:HB	1:B:406:SER:HA	0.97	0.96
1:B:472:VAL:HG11	1:B:496:VAL:HG11	1.75	0.67
1:A:423:CYS:HB3	1:B:405:THR:HG21	1.80	0.64

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	492/551 (89%)	476 (97%)	15 (3%)	1 (0%)	47 78
1	B	409/551 (74%)	398 (97%)	9 (2%)	2 (0%)	29 61
1	C	500/551 (91%)	478 (96%)	20 (4%)	2 (0%)	34 66
1	D	506/551 (92%)	488 (96%)	18 (4%)	0	100 100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1907/2204 (86%)	1840 (96%)	62 (3%)	5 (0%)	41 71

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	328	THR
1	C	328	THR
1	B	328	THR
1	B	404	ILE
1	C	40	ILE

### 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	409/453 (90%)	404 (99%)	5 (1%)	71 91
1	B	340/453 (75%)	336 (99%)	4 (1%)	71 91
1	C	416/453 (92%)	413 (99%)	3 (1%)	84 95
1	D	418/453 (92%)	410 (98%)	8 (2%)	57 84
All	All	1583/1812 (87%)	1563 (99%)	20 (1%)	69 90

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	41	THR
1	C	294	ARG
1	D	223	GLU
1	B	342	ARG
1	B	467	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	90	ASN
1	A	274	HIS
1	B	393	GLN
1	C	393	GLN
1	D	252	HIS

### 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](#)

9 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	PHE	B	602	-	9,12,12	0.54	0	10,15,15	0.15	0
3	PHE	A	603	-	9,12,12	0.53	0	10,15,15	0.17	0
3	PHE	D	602	-	9,12,12	0.52	0	10,15,15	0.18	0
2	PO4	A	602	-	4,4,4	0.92	0	6,6,6	0.43	0
3	PHE	C	602	-	9,12,12	0.54	0	10,15,15	0.17	0
2	PO4	A	601	-	4,4,4	0.91	0	6,6,6	0.41	0
2	PO4	B	601	-	4,4,4	0.91	0	6,6,6	0.44	0
2	PO4	C	601	-	4,4,4	0.92	0	6,6,6	0.43	0
2	PO4	D	601	-	4,4,4	0.91	0	6,6,6	0.43	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	PHE	A	603	-	-	0/4/8/8	0/1/1/1
3	PHE	B	602	-	-	0/4/8/8	0/1/1/1
3	PHE	D	602	-	-	0/4/8/8	0/1/1/1
3	PHE	C	602	-	-	0/4/8/8	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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<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	498/551 (90%)	0.02	1 (0%) 95 95	25, 37, 57, 69	0
1	B	413/551 (74%)	0.03	6 (1%) 73 73	23, 33, 57, 107	0
1	C	504/551 (91%)	0.08	6 (1%) 79 79	28, 44, 73, 82	1 (0%)
1	D	508/551 (92%)	0.22	20 (3%) 39 35	28, 41, 85, 94	0
All	All	1923/2204 (87%)	0.09	33 (1%) 70 69	23, 39, 77, 107	1 (0%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	406	SER	4.9
1	D	199	ASN	4.4
1	D	182	SER	3.6
1	B	219	PRO	3.6
1	D	134	LEU	3.5

### 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 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	PHE	C	602	12/12	0.87	0.36	96,96,96,96	0
3	PHE	D	602	12/12	0.88	0.40	104,104,104,104	0
2	PO4	A	602	5/5	0.91	0.20	87,87,87,87	0
3	PHE	A	603	12/12	0.91	0.32	80,80,81,81	0
3	PHE	B	602	12/12	0.92	0.23	50,50,50,50	0
2	PO4	B	601	5/5	0.94	0.17	69,69,69,69	0
2	PO4	C	601	5/5	0.94	0.29	87,87,87,87	0
2	PO4	A	601	5/5	0.95	0.16	35,35,35,35	0
2	PO4	D	601	5/5	0.95	0.16	55,55,56,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.