



# wwPDB X-ray Structure Validation Summary Report

Feb 1, 2016 – 08:25 PM GMT

PDB ID : 4RPP  
Title : crystal structure of PKM2-K422R mutant bound with FBP  
Authors : Wang, P.; Sun, C.; Zhu, T.; Xu, Y.  
Deposited on : 2014-10-31  
Resolution : 2.58 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

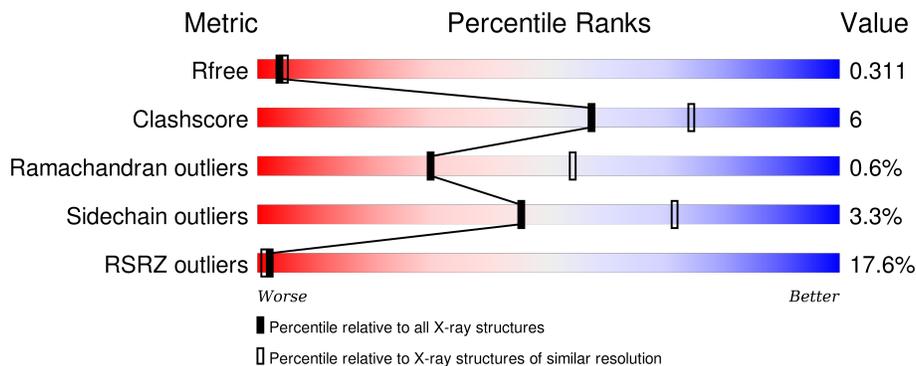
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.58 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))
$R_{free}$	91344	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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. 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	536	 10% 63% 13% • 23%
1	B	536	 15% 63% 12% • 24%
1	C	536	 15% 63% 11% 26%
1	D	536	 13% 64% 11% • 24%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12596 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 PKM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	411	3153	1978	568	586	21	0	0	0
1	B	408	3140	1970	568	581	21	0	0	0
1	C	396	3050	1912	552	565	21	0	0	0
1	D	408	3143	1973	568	581	21	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

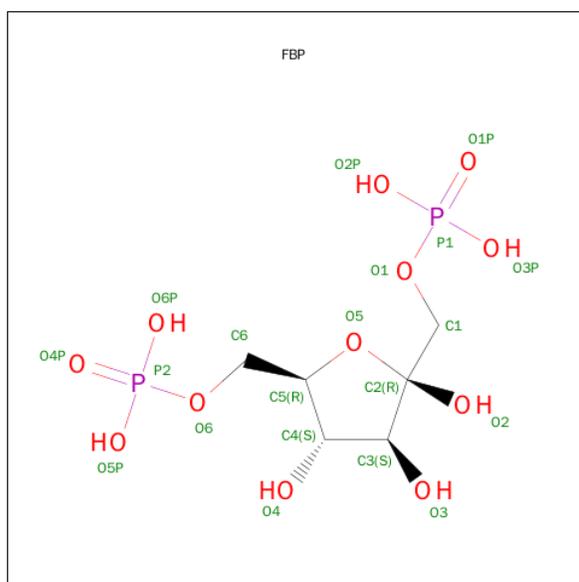
Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	PRO	-	EXPRESSION TAG	UNP P14618
A	-2	LEU	-	EXPRESSION TAG	UNP P14618
A	-1	GLY	-	EXPRESSION TAG	UNP P14618
A	0	SER	-	EXPRESSION TAG	UNP P14618
A	422	ARG	LYS	ENGINEERED MUTATION	UNP P14618
A	532	FBP	-	EXPRESSION TAG	UNP P14618
B	-3	PRO	-	EXPRESSION TAG	UNP P14618
B	-2	LEU	-	EXPRESSION TAG	UNP P14618
B	-1	GLY	-	EXPRESSION TAG	UNP P14618
B	0	SER	-	EXPRESSION TAG	UNP P14618
B	422	ARG	LYS	ENGINEERED MUTATION	UNP P14618
B	532	FBP	-	EXPRESSION TAG	UNP P14618
C	-3	PRO	-	EXPRESSION TAG	UNP P14618
C	-2	LEU	-	EXPRESSION TAG	UNP P14618
C	-1	GLY	-	EXPRESSION TAG	UNP P14618
C	0	SER	-	EXPRESSION TAG	UNP P14618
C	422	ARG	LYS	ENGINEERED MUTATION	UNP P14618
C	532	FBP	-	EXPRESSION TAG	UNP P14618
D	-3	PRO	-	EXPRESSION TAG	UNP P14618
D	-2	LEU	-	EXPRESSION TAG	UNP P14618
D	-1	GLY	-	EXPRESSION TAG	UNP P14618

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	SER	-	EXPRESSION TAG	UNP P14618
D	422	ARG	LYS	ENGINEERED MUTATION	UNP P14618
D	532	FBP	-	EXPRESSION TAG	UNP P14618

- Molecule 2 is SUGAR (BETA-FRUCTOSE-1,6-DIPHOSPHATE) (three-letter code: FBP) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>12</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	P	0	0
			20	6	12	2		
2	B	1	Total	C	O	P	0	0
			20	6	12	2		
2	C	1	Total	C	O	P	0	0
			20	6	12	2		
2	D	1	Total	C	O	P	0	0
			20	6	12	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	7	Total	O	0	0
			7	7		
3	B	7	Total	O	0	0
			7	7		
3	C	8	Total	O	0	0
			8	8		

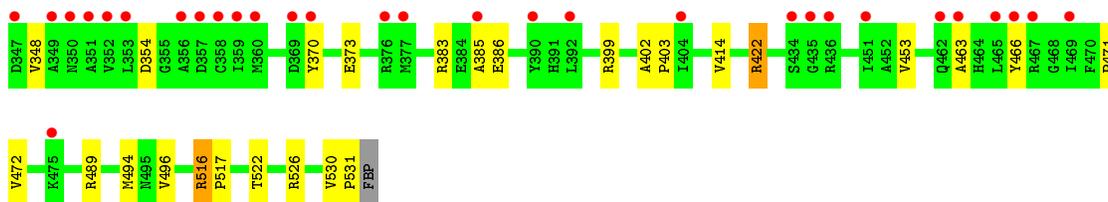
Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	D	8	Total	O	0	0
			8	8		







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.67Å 152.55Å 97.68Å 90.00° 104.24° 90.00°	Depositor
Resolution (Å)	38.31 – 2.58 38.31 – 2.58	Depositor EDS
% Data completeness (in resolution range)	97.6 (38.31-2.58) 97.6 (38.31-2.58)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.52 (at 2.58Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.274 , 0.308 0.287 , 0.311	Depositor DCC
$R_{free}$ test set	3472 reflections (5.16%)	DCC
Wilson B-factor (Å <sup>2</sup> )	66.4	Xtrriage
Anisotropy	0.141	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 53.5	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Outliers	1 of 70704 reflections (0.001%)	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12596	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	90.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: FBP

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	0.24	0/3204	0.45	0/4330
1	B	0.23	0/3194	0.46	1/4315 (0.0%)
1	C	0.23	0/3097	0.43	0/4178
1	D	0.22	0/3197	0.41	0/4319
All	All	0.23	0/12692	0.44	1/17142 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	342	ARG	CG-CD-NE	-5.17	100.95	111.80

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	3153	0	3214	53	0
1	B	3140	0	3191	55	0
1	C	3050	0	3109	35	0
1	D	3143	0	3197	40	0
2	A	20	0	10	1	0
2	B	20	0	10	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	20	0	10	0	0
2	D	20	0	10	0	0
3	A	7	0	0	0	0
3	B	7	0	0	0	0
3	C	8	0	0	0	0
3	D	8	0	0	1	0
All	All	12596	0	12751	153	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 6.

The worst 5 of 153 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:306:VAL:HG11	1:B:342:ARG:HH22	1.37	0.88
1:A:342:ARG:NH2	1:B:302:PRO:O	2.06	0.88
1:B:56:ARG:NH2	1:B:83:TYR:O	2.17	0.75
1:A:494:MET:HG2	1:A:531:PRO:HD2	1.69	0.74
1:B:274:HIS:CE1	1:B:301:ILE:HG22	2.25	0.71

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	405/536 (76%)	393 (97%)	10 (2%)	2 (0%)	34 58
1	B	404/536 (75%)	391 (97%)	8 (2%)	5 (1%)	16 33
1	C	388/536 (72%)	380 (98%)	6 (2%)	2 (0%)	34 58
1	D	404/536 (75%)	397 (98%)	6 (2%)	1 (0%)	52 75
All	All	1601/2144 (75%)	1561 (98%)	30 (2%)	10 (1%)	30 54

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	404	ILE
1	B	303	ALA
1	B	403	PRO
1	C	406	SER
1	A	212	PRO

### 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	339/439 (77%)	326 (96%)	13 (4%)	40	67
1	B	336/439 (76%)	325 (97%)	11 (3%)	45	71
1	C	328/439 (75%)	318 (97%)	10 (3%)	48	75
1	D	337/439 (77%)	327 (97%)	10 (3%)	48	75
All	All	1340/1756 (76%)	1296 (97%)	44 (3%)	45	71

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	359	ILE
1	C	32	ARG
1	D	422	ARG
1	B	386	GLU
1	B	522	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	379	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

4 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$
2	FBP	A	600	-	18,20,20	0.90	1 (5%)	21,32,32	0.61	0
2	FBP	B	600	-	18,20,20	0.91	1 (5%)	21,32,32	0.66	0
2	FBP	C	600	-	18,20,20	0.90	1 (5%)	21,32,32	0.63	0
2	FBP	D	600	-	18,20,20	0.90	1 (5%)	21,32,32	0.63	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
2	FBP	A	600	-	-	0/13/32/32	0/1/1/1
2	FBP	B	600	-	-	0/13/32/32	0/1/1/1
2	FBP	C	600	-	-	0/13/32/32	0/1/1/1
2	FBP	D	600	-	-	0/13/32/32	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	600	FBP	O2-C2	2.68	1.45	1.41
2	C	600	FBP	O2-C2	2.70	1.45	1.41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	600	FBP	O2-C2	2.71	1.45	1.41
2	A	600	FBP	O2-C2	2.72	1.45	1.41

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	FBP	1	0
2	B	600	FBP	1	0

## 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	411/536 (76%)	0.91	56 (13%) 4 3	47, 80, 131, 203	0
1	B	408/536 (76%)	1.18	79 (19%) 1 1	55, 88, 134, 167	0
1	C	396/536 (73%)	1.16	79 (19%) 1 1	68, 95, 130, 167	0
1	D	408/536 (76%)	1.02	72 (17%) 2 1	50, 82, 121, 155	0
All	All	1623/2144 (75%)	1.07	286 (17%) 2 1	47, 87, 129, 203	0

The worst 5 of 286 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	336	LYS	9.7
1	D	40	ILE	8.1
1	B	244	PHE	7.9
1	C	41	THR	7.8
1	B	303	ALA	7.5

### 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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	FBP	D	600	20/20	0.95	0.19	-0.03	54,66,83,89	0
2	FBP	C	600	20/20	0.88	0.22	-0.12	78,86,103,112	0
2	FBP	B	600	20/20	0.94	0.19	-0.19	59,74,94,99	0
2	FBP	A	600	20/20	0.94	0.17	-0.71	70,90,100,102	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.