



# wwPDB X-ray Structure Validation Summary Report

Feb 1, 2016 – 05:08 PM GMT

PDB ID : 4HAB  
Title : Crystal structure of Plk1 Polo-box domain in complex with PL-49  
Authors : Lee, W.C.; Song, J.H.; Kim, H.Y.  
Deposited on : 2012-09-26  
Resolution : 2.65 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

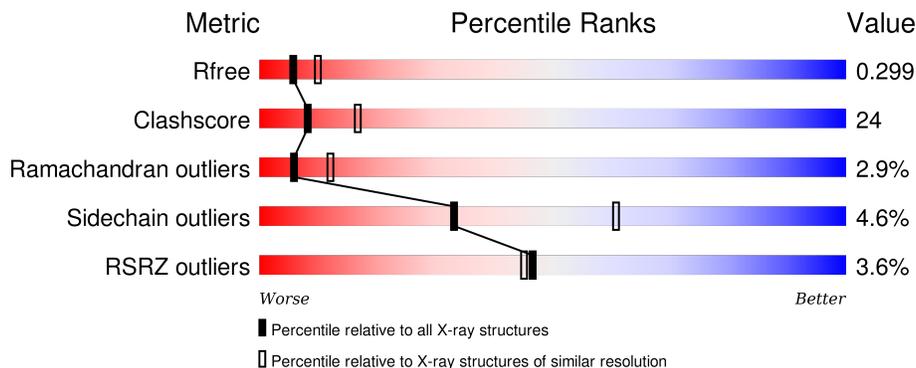
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)

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	223	
1	B	223	
1	C	223	
2	D	7	
2	E	7	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	7	 14% 57% 29%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5428 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.

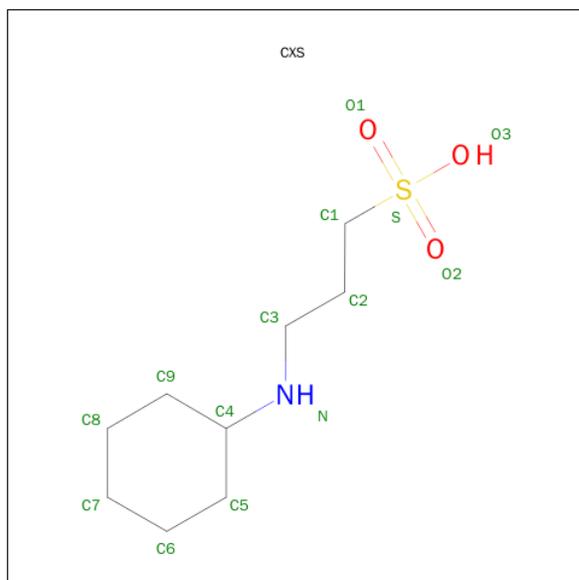
- Molecule 1 is a protein called Serine/threonine-protein kinase PLK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	Total 1811	C 1147	N 313	O 340	S 11	0	0	0
1	B	218	Total 1774	C 1126	N 306	O 332	S 10	0	0	0
1	C	187	Total 1545	C 990	N 265	O 283	S 7	0	0	0

- Molecule 2 is a protein called PL-49.

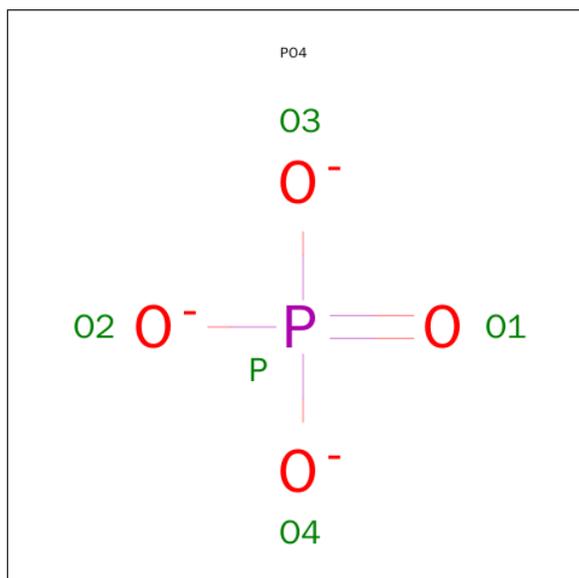
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
2	D	7	Total 58	C 36	N 9	O 11	P 1	S 1	0	0	1
2	E	7	Total 58	C 36	N 9	O 11	P 1	S 1	0	0	1
2	F	7	Total 51	C 29	N 9	O 11	P 1	S 1	0	0	1

- Molecule 3 is 3-CYCLOHEXYL-1-PROPYLSULFONIC ACID (three-letter code: CXS) (formula: C<sub>9</sub>H<sub>19</sub>NO<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	Total	C	N	O	S	0	0
			14	9	1	3	1		
3	A	1	Total	C	N	O	S	0	0
			14	9	1	3	1		
3	B	1	Total	C	N	O	S	0	0
			14	9	1	3	1		
3	C	1	Total	C	N	O	S	0	0
			14	9	1	3	1		
3	C	1	Total	C	N	O	S	0	0
			14	9	1	3	1		

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

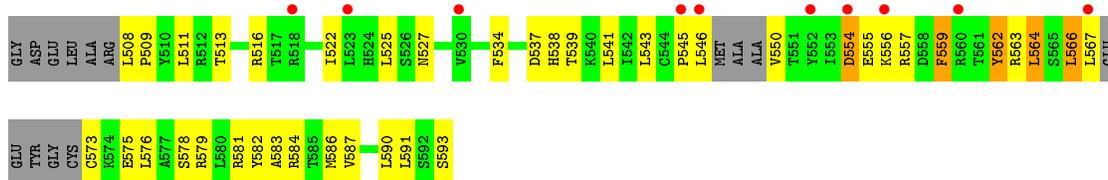


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	P	0	0
			5	4	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	26	Total	O	0	0
			26	26		
5	B	15	Total	O	0	0
			15	15		
5	C	4	Total	O	0	0
			4	4		
5	D	5	Total	O	0	0
			5	5		
5	E	5	Total	O	0	0
			5	5		
5	F	1	Total	O	0	0
			1	1		





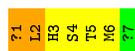
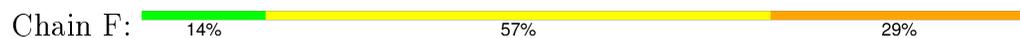
- Molecule 2: PL-49



- Molecule 2: PL-49



- Molecule 2: PL-49



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.04Å 95.04Å 240.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.65 88.40 – 2.65	Depositor EDS
% Data completeness (in resolution range)	97.7 (50.00-2.65) 97.9 (88.40-2.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.19 (at 2.65Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.260 , 0.259 0.251 , 0.299	Depositor DCC
$R_{free}$ test set	1633 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	55.5	Xtrriage
Anisotropy	0.335	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 57.3	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Outliers	0 of 32177 reflections	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	5428	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% 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: TPO, PO4, CXS, 11Q, NH2

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	0.43	0/1849	0.66	0/2500
1	B	0.42	0/1809	0.64	0/2443
1	C	0.37	0/1572	0.59	0/2115
2	D	0.48	0/31	0.52	0/38
2	E	0.38	0/31	0.52	0/38
2	F	0.38	0/31	0.64	0/38
All	All	0.41	0/5323	0.63	0/7172

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
2	E	0	1
2	F	0	2
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	1	11Q	Mainchain
2	F	1	11Q	Mainchain,Peptide

## 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	1811	0	1793	65	0
1	B	1774	0	1764	77	0
1	C	1545	0	1546	113	0
2	D	58	0	58	5	0
2	E	58	0	58	7	0
2	F	51	0	44	11	0
3	A	28	0	36	2	0
3	B	14	0	18	1	0
3	C	28	0	36	1	0
4	A	5	0	0	1	0
5	A	26	0	0	2	0
5	B	15	0	0	1	0
5	C	4	0	0	0	0
5	D	5	0	0	0	0
5	E	5	0	0	0	0
5	F	1	0	0	0	0
All	All	5428	0	5353	262	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 24.

The worst 5 of 262 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:456:ARG:HD3	1:A:456:ARG:H	1.26	0.96
1:C:442:LEU:HD11	1:C:451:LEU:HD22	1.54	0.90
1:B:394:LEU:HD21	1:B:396:ARG:HG3	1.51	0.89
1:B:443:ILE:HD11	1:B:510:TYR:HB3	1.54	0.89
1:C:406:ILE:HG23	1:C:500:ARG:HB2	1.60	0.82

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	221/223 (99%)	206 (93%)	13 (6%)	2 (1%)	21	44
1	B	214/223 (96%)	194 (91%)	18 (8%)	2 (1%)	21	44
1	C	173/223 (78%)	133 (77%)	27 (16%)	13 (8%)	1	1
2	D	4/7 (57%)	3 (75%)	1 (25%)	0	100	100
2	E	4/7 (57%)	3 (75%)	1 (25%)	0	100	100
2	F	4/7 (57%)	2 (50%)	1 (25%)	1 (25%)	0	0
All	All	620/690 (90%)	541 (87%)	61 (10%)	18 (3%)	6	12

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	439	SER
1	C	559	PHE
1	C	566	LEU
2	F	2	LEU
1	A	502	GLY

### 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	202/202 (100%)	195 (96%)	7 (4%)	43	71
1	B	197/202 (98%)	187 (95%)	10 (5%)	29	55
1	C	174/202 (86%)	164 (94%)	10 (6%)	25	49

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	4/4 (100%)	4 (100%)	0	100	100
2	E	4/4 (100%)	4 (100%)	0	100	100
2	F	4/4 (100%)	4 (100%)	0	100	100
All	All	585/618 (95%)	558 (95%)	27 (5%)	33	61

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	497	ILE
1	B	556	LYS
1	C	491	LEU
1	B	507	ARG
1	A	491	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	373	HIS
1	B	397	GLN
1	B	524	HIS
1	A	524	HIS
1	B	496	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

6 non-standard protein/DNA/RNA residues 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	11Q	D	1	2	13,15,16	1.33	1 (7%)	16,19,21	0.94	2 (12%)
2	TPO	D	5	2	8,10,11	0.79	0	7,14,16	1.46	1 (14%)
2	11Q	E	1	2	13,15,16	1.42	1 (7%)	16,19,21	0.50	0
2	TPO	E	5	2	8,10,11	0.47	0	7,14,16	1.52	2 (28%)
2	11Q	F	1	2	6,7,16	2.20	1 (16%)	7,8,21	2.51	1 (14%)
2	TPO	F	5	2	8,10,11	1.09	0	7,14,16	1.39	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 Torsions and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	11Q	D	1	2	-	0/4/24/26	0/2/2/2
2	TPO	D	5	2	-	0/8/11/13	0/0/0/0
2	11Q	E	1	2	-	0/4/24/26	0/2/2/2
2	TPO	E	5	2	-	0/8/11/13	0/0/0/0
2	11Q	F	1	2	-	0/0/9/26	0/1/1/2
2	TPO	F	5	2	-	0/8/11/13	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	11Q	O-C	4.71	1.41	1.19
2	E	1	11Q	O-C	5.00	1.42	1.19
2	F	1	11Q	O-C	5.16	1.43	1.19

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	1	11Q	O-C-CA	-6.29	108.84	125.44
2	E	5	TPO	O2P-P-O1P	2.09	117.30	110.58
2	D	1	11Q	CD-N-CA	2.12	110.50	106.50
2	D	5	TPO	OG1-P-O1P	2.61	113.62	107.11
2	E	5	TPO	OG1-P-O1P	2.69	113.81	107.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	11Q	3	0
2	E	1	11Q	6	0
2	F	1	11Q	1	0
2	F	5	TPO	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

6 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	CXS	A	601	-	13,14,14	0.22	0	15,18,18	1.16	2 (13%)
3	CXS	A	602	-	13,14,14	0.20	0	15,18,18	0.80	1 (6%)
4	PO4	A	603	-	4,4,4	1.14	0	6,6,6	0.27	0
3	CXS	B	601	-	13,14,14	0.24	0	15,18,18	0.47	0
3	CXS	C	601	-	13,14,14	0.19	0	15,18,18	0.62	0
3	CXS	C	602	-	13,14,14	0.20	0	15,18,18	1.47	2 (13%)

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	CXS	A	601	-	-	0/8/16/16	0/1/1/1
3	CXS	A	602	-	-	0/8/16/16	0/1/1/1
4	PO4	A	603	-	-	0/0/0/0	0/0/0/0
3	CXS	B	601	-	-	0/8/16/16	0/1/1/1
3	CXS	C	601	-	-	0/8/16/16	0/1/1/1
3	CXS	C	602	-	-	0/8/16/16	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	602	CXS	C9-C4-N	-3.91	94.84	111.10
3	C	602	CXS	C5-C4-N	-3.56	96.31	111.10
3	A	602	CXS	C5-C4-N	-2.31	101.50	111.10
3	A	601	CXS	C9-C4-N	2.72	122.38	111.10
3	A	601	CXS	C2-C3-N	3.42	124.20	112.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	CXS	1	0
3	A	602	CXS	1	0
4	A	603	PO4	1	0
3	B	601	CXS	1	0
3	C	601	CXS	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	223/223 (100%)	0.35	1 (0%) 93 94	30, 49, 78, 94	0
1	B	218/223 (97%)	0.44	4 (1%) 71 70	31, 59, 86, 98	0
1	C	187/223 (83%)	0.77	18 (9%) 10 8	47, 79, 104, 124	0
2	D	4/7 (57%)	0.27	0 100 100	43, 46, 54, 54	0
2	E	4/7 (57%)	0.24	0 100 100	43, 52, 52, 58	0
2	F	4/7 (57%)	0.21	0 100 100	51, 56, 60, 62	0
All	All	640/690 (92%)	0.50	23 (3%) 46 45	30, 60, 96, 124	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	545	PRO	4.1
1	C	373	HIS	3.7
1	C	378	LEU	3.4
1	C	374	LEU	3.4
1	B	463	LEU	3.2

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q < 0.9
2	TPO	E	5	11/12	0.98	0.17	-	35,42,46,47	0
2	TPO	D	5	11/12	0.99	0.16	-	34,40,42,42	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	TPO	F	5	11/12	0.97	0.17	-	43,48,50,51	0
2	11Q	E	1	14/15	0.95	0.17	-	47,51,54,54	0
2	11Q	F	1	7/15	0.92	0.18	-	58,59,60,60	0
2	11Q	D	1	14/15	0.93	0.20	-	44,47,49,50	0

### 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
3	CXS	A	601	14/14	0.94	0.21	0.46	55,56,62,63	0
3	CXS	C	601	14/14	0.92	0.20	-0.25	93,94,95,95	0
3	CXS	C	602	14/14	0.93	0.22	-0.59	85,86,92,93	0
3	CXS	B	601	14/14	0.96	0.16	-0.63	62,64,74,74	0
3	CXS	A	602	14/14	0.95	0.17	-0.82	38,47,65,66	0
4	PO4	A	603	5/5	0.97	0.16	-1.69	52,53,56,59	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.