



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 07:56 am BST

PDB ID : 1Q4K
Title : The polo-box domain of Plk1 in complex with a phospho-peptide
Authors : Cheng, K.; Lowe, E.D.; Sinclair, J.; Nigg, E.A.; Johnson, L.N.
Deposited on : 2003-08-04
Resolution : 2.30 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 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 : 2.11
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) : 2.11

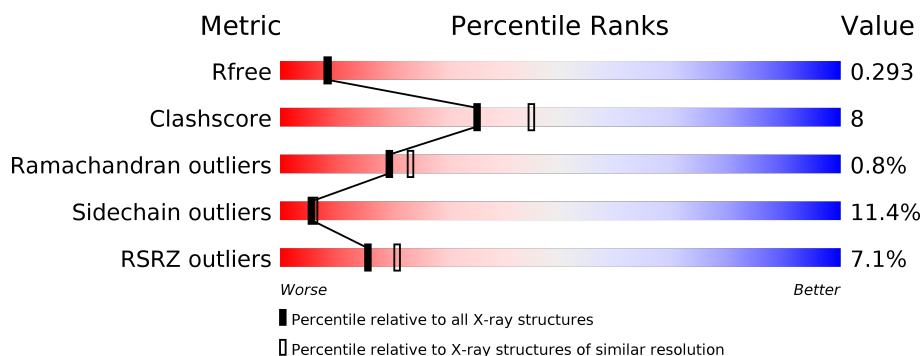
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 ≥ 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	259	<div> <div>7%</div> <div> <div>64%</div> <div>18%</div> <div>•</div> <div>14%</div> </div> </div>
1	B	259	<div> <div>3%</div> <div> <div>63%</div> <div>20%</div> <div>•</div> <div>14%</div> </div> </div>
1	C	259	<div> <div>8%</div> <div> <div>59%</div> <div>21%</div> <div>•</div> <div>17%</div> </div> </div>
2	D	6	<div> <div>67%</div> <div>33%</div> </div>
2	E	6	<div> <div>83%</div> <div>17%</div> </div>
2	F	6	<div> <div>17%</div> <div> <div>33%</div> <div>50%</div> <div>17%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5707 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 PLK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	224	Total	C	N	O	S	0	0	0
			1822	1153	317	341	11			
1	A	222	Total	C	N	O	S	0	0	0
			1803	1143	312	337	11			
1	C	214	Total	C	N	O	S	0	0	0
			1744	1108	300	325	11			

- Molecule 2 is a protein called Phospho-peptide sequence Met.Gln.Ser.pThr.Pro.Leu.

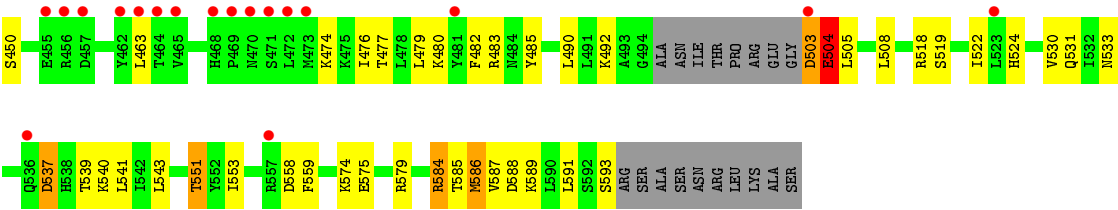
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	6	Total	C	N	O	P	S	0	0
			49	28	7	12	1	1		
2	E	6	Total	C	N	O	P	S	0	0
			49	28	7	12	1	1		
2	F	6	Total	C	N	O	P	S	0	0
			49	28	7	12	1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	78	Total	O	0	0
			78	78		
3	A	55	Total	O	0	0
			55	55		
3	C	43	Total	O	0	0
			43	43		
3	D	2	Total	O	0	0
			2	2		
3	E	8	Total	O	0	0
			8	8		
3	F	5	Total	O	0	0
			5	5		

- Molecule 1: Serine/threonine-protein kinase PLK

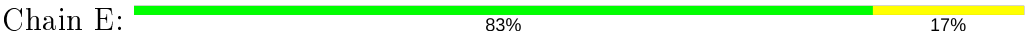




● Molecule 2: Phospho-peptide sequence Met.Gln.Ser.pThr.Pro.Leu



● Molecule 2: Phospho-peptide sequence Met.Gln.Ser.pThr.Pro.Leu



● Molecule 2: Phospho-peptide sequence Met.Gln.Ser.pThr.Pro.Leu



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	57.07Å 56.89Å 85.05Å 91.45° 103.21° 118.50°	Depositor
Resolution (Å)	29.36 – 2.30 29.31 – 2.30	Depositor EDS
% Data completeness (in resolution range)	96.2 (29.36-2.30) 96.2 (29.31-2.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.04 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.245 , 0.312 0.237 , 0.293	Depositor DCC
R_{free} test set	1943 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	41.8	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 48.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.019 for -h-k,k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5707	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.96% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: TPO

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	0.58	0/1841	0.85	6/2489 (0.2%)
1	B	0.59	0/1860	0.88	9/2514 (0.4%)
1	C	0.47	0/1780	0.80	5/2404 (0.2%)
2	D	0.56	0/37	0.76	0/46
2	E	0.75	0/37	1.13	0/46
2	F	0.60	0/37	0.68	0/46
All	All	0.55	0/5592	0.84	20/7545 (0.3%)

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	A	0	2

There are no bond length outliers.

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	537	ASP	CB-CG-OD2	6.85	124.46	118.30
1	C	376	ASP	CB-CG-OD2	6.68	124.31	118.30
1	B	554	ASP	CB-CG-OD2	6.46	124.12	118.30
1	A	537	ASP	CB-CG-OD2	6.37	124.03	118.30
1	C	416	ASP	CB-CG-OD2	6.29	123.96	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	517	THR	Peptide
1	A	572	CYS	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	1803	0	1789	35	1
1	B	1822	0	1806	27	1
1	C	1744	0	1730	29	1
2	D	49	0	48	2	0
2	E	49	0	48	1	0
2	F	49	0	48	3	0
3	A	55	0	0	1	0
3	B	78	0	0	6	1
3	C	43	0	0	3	0
3	D	2	0	0	0	0
3	E	8	0	0	0	0
3	F	5	0	0	0	0
All	All	5707	0	5469	90	2

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 8.

The worst 5 of 90 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:463:LEU:HD21	1:A:469:PRO:HD2	1.54	0.89
1:B:517:THR:C	3:B:36:HOH:O	2.16	0.82
1:A:493:ALA:O	3:A:21:HOH:O	2.04	0.75
1:B:432:VAL:HG12	1:B:479:LEU:HD11	1.69	0.74
1:A:384:VAL:HA	1:A:568:GLU:HG2	1.72	0.71

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:589:LYS:NZ	3:B:13:HOH:O[1_554]	1.91	0.29
1:B:419:ASP:OD2	1:A:445:TYR:OH[1_455]	2.19	0.01

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	220/259 (85%)	206 (94%)	11 (5%)	3 (1%)	11	11
1	B	222/259 (86%)	206 (93%)	15 (7%)	1 (0%)	29	35
1	C	210/259 (81%)	198 (94%)	11 (5%)	1 (0%)	29	35
2	D	3/6 (50%)	3 (100%)	0	0	100	100
2	E	3/6 (50%)	3 (100%)	0	0	100	100
2	F	3/6 (50%)	3 (100%)	0	0	100	100
All	All	661/795 (83%)	619 (94%)	37 (6%)	5 (1%)	19	23

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	467	SER
1	A	504	GLU
1	A	573	CYS
1	C	504	GLU
1	B	504	GLU

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	201/234 (86%)	182 (90%)	19 (10%)	8	10
1	B	203/234 (87%)	177 (87%)	26 (13%)	4	4
1	C	195/234 (83%)	172 (88%)	23 (12%)	5	5
2	D	5/5 (100%)	5 (100%)	0	100	100
2	E	5/5 (100%)	5 (100%)	0	100	100
2	F	5/5 (100%)	3 (60%)	2 (40%)	0	0
All	All	614/717 (86%)	544 (89%)	70 (11%)	5	6

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	491	LEU
1	A	560	ARG
1	C	584	ARG
1	A	503	ASP
1	A	536	GLN

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	A	533	ASN
1	C	468	HIS
2	D	2	GLN
1	A	524	HIS
1	C	533	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

3 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	TPO	D	4	2	8,10,11	3.43	4 (50%)	10,14,16	2.80	5 (50%)
2	TPO	E	4	2	8,10,11	3.19	4 (50%)	10,14,16	2.68	4 (40%)
2	TPO	F	4	2	8,10,11	3.57	4 (50%)	10,14,16	2.83	3 (30%)

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	TPO	D	4	2	-	2/9/11/13	-
2	TPO	E	4	2	-	0/9/11/13	-
2	TPO	F	4	2	-	1/9/11/13	-

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	4	TPO	P-O1P	6.20	1.70	1.50
2	D	4	TPO	P-O1P	5.80	1.69	1.50
2	F	4	TPO	O-C	5.65	1.42	1.19
2	E	4	TPO	O-C	5.62	1.42	1.19
2	D	4	TPO	O-C	5.54	1.42	1.19

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	4	TPO	O2P-P-OG1	-6.26	77.96	105.99
2	D	4	TPO	O2P-P-OG1	-5.36	81.96	105.99
2	E	4	TPO	O3P-P-O2P	4.83	126.10	107.64
2	F	4	TPO	O3P-P-O2P	4.45	124.63	107.64
2	D	4	TPO	O-C-CA	-4.16	113.87	124.78

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	4	TPO	CB-OG1-P-O2P
2	F	4	TPO	CB-OG1-P-O2P

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	D	4	TPO	CB-OG1-P-O3P

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	4	TPO	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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	222/259 (85%)	0.39	17 (7%) 13 17	30, 45, 59, 67	0
1	B	224/259 (86%)	0.19	9 (4%) 38 45	27, 41, 57, 65	0
1	C	214/259 (82%)	0.62	21 (9%) 7 10	35, 53, 91, 95	0
2	D	5/6 (83%)	0.62	0 100 100	51, 53, 55, 57	0
2	E	5/6 (83%)	0.10	0 100 100	38, 40, 42, 44	0
2	F	5/6 (83%)	1.57	1 (20%) 1 1	48, 52, 56, 57	0
All	All	675/795 (84%)	0.40	48 (7%) 16 21	27, 46, 65, 95	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	372	CYS	6.6
1	B	372	CYS	6.0
1	C	469	PRO	5.3
1	C	465	VAL	5.1
1	C	463	LEU	5.0

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. 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
2	TPO	D	4	11/12	0.95	0.12	47,49,50,51	0
2	TPO	F	4	11/12	0.96	0.10	44,47,49,49	0
2	TPO	E	4	11/12	0.99	0.13	32,35,36,37	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.