Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 01:56 AM GMT

PDB ID	:	2EXE
Title	:	Crystal structure of the phosphorylated CLK3
Authors	:	Papagrigoriou, E.; Rellos, P.; Das, S.; Bullock, A.; Ball, L.J.; Turnbull, A.; Savitsky, P.; Fedorov, O.; Johansson, C.; Ugochukwu, E.; Sobott, F.; von Delft, F.; Edwards, A.; Sundstrom, M.; Weigelt, J.; Arrowsmith, C.; Knapp, S.; Structural Genomics Consortium (SGC)
Deposited on	:	2005-11-08
Resolution	:	2.35 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

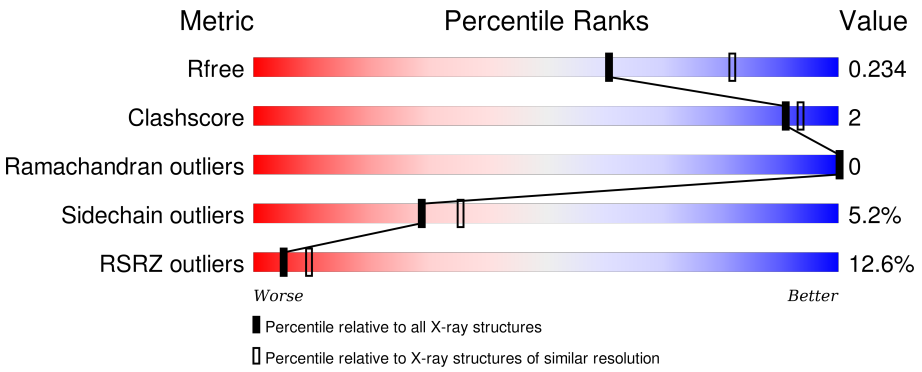
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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

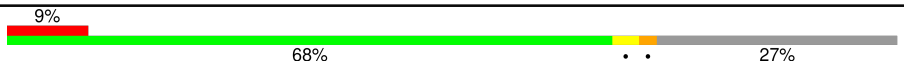
The reported resolution of this entry is 2.35 Å.

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	1352 (2.38-2.34)
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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. 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	357	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2195 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 Dual specificity protein kinase CLK3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	262	2097	1356	355	373	13	0	2	0

- Molecule 2 is water.

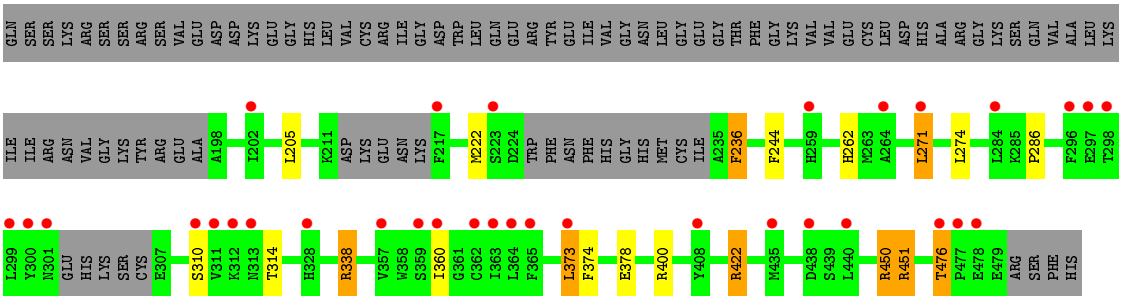
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	96	98	98	0	2

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Dual specificity protein kinase CLK3

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	62.18Å 111.00Å 161.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.82 – 2.35 32.82 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (32.82-2.35) 99.5 (32.82-2.30)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.70 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.225 , 0.260 0.233 , 0.234	Depositor DCC
R_{free} test set	1154 reflections (5.13%)	DCC
Wilson B-factor (Å^2)	40.2	Xtriage
Anisotropy	0.656	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 49.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 25018 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2195	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.44	0/2159	0.65	4/2928 (0.1%)

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	1	0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	338	ARG	NE-CZ-NH2	-6.37	117.11	120.30
1	A	476	THR	CA-CB-CG2	6.00	120.80	112.40
1	A	476	THR	OG1-CB-CG2	5.40	122.42	110.00
1	A	271	LEU	CA-CB-CG	-5.21	103.32	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	476	THR	CB

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	2097	0	1968	9	0
2	A	98	0	0	0	0
All	All	2195	0	1968	9	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 2.

All (9) 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:262:HIS:ND1	1:A:314:THR:HG23	2.11	0.65
1:A:236:PHE:N	1:A:236:PHE:CD1	2.71	0.58
1:A:422:ARG:HA	1:A:422:ARG:HH11	1.69	0.58
1:A:271:LEU:HD13	1:A:360:ILE:HD11	1.88	0.56
1:A:373:LEU:HD13	1:A:374:PHE:CE2	2.47	0.49
1:A:244:PHE:HB2	1:A:286:PRO:HB2	1.95	0.48
1:A:450:ARG:HD3	1:A:451:ARG:HG2	1.95	0.48
1:A:205:LEU:HD13	1:A:222:MET:HG3	1.98	0.45
1:A:262:HIS:ND1	1:A:314:THR:CG2	2.81	0.42

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	256/357 (72%)	246 (96%)	10 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	214/323 (66%)	203 (95%)	11 (5%)	29 36

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	236	PHE
1	A	274	LEU
1	A	310	SER
1	A	338	ARG
1	A	373	LEU
1	A	378	GLU
1	A	400	ARG
1	A	422	ARG
1	A	450	ARG
1	A	451	ARG
1	A	476	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

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	262/357 (73%)	0.52	33 (12%) 5 9	38, 56, 76, 107	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	300	TYR	8.0
1	A	477	PRO	5.9
1	A	202	ILE	4.5
1	A	296	PHE	4.3
1	A	363	ILE	4.0
1	A	297	GLU	3.8
1	A	360	ILE	3.6
1	A	301	ASN	3.5
1	A	364	LEU	3.3
1	A	299	LEU	3.2
1	A	313	ASN	3.1
1	A	311	VAL	3.0
1	A	440	LEU	2.8
1	A	298	THR	2.7
1	A	310	SER	2.6
1	A	362	CYS	2.6
1	A	373	LEU	2.6
1	A	408	TYR	2.6
1	A	328	HIS	2.6
1	A	259	HIS	2.5
1	A	438	ASP	2.4
1	A	435	MET	2.3
1	A	365	PHE	2.3
1	A	357	VAL	2.3
1	A	284	LEU	2.2
1	A	359	SER	2.2
1	A	264	ALA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	223	SER	2.1
1	A	312	LYS	2.1
1	A	271	LEU	2.1
1	A	476	THR	2.1
1	A	217	PHE	2.0
1	A	478	GLU	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.