



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 01:45 am GMT

PDB ID : 2OKR
Title : Crystal Structure of the P38a-MAPKAP kinase 2 Heterodimer
Authors : Ter Haar, E.
Deposited on : 2007-01-17
Resolution : 2.00 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

<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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

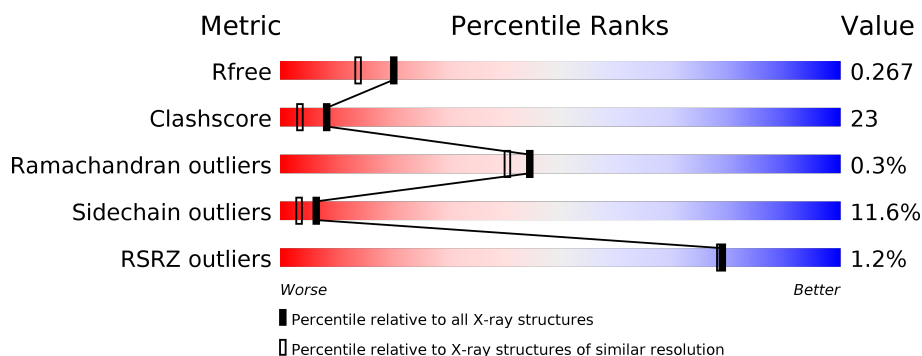
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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 $\leq 5\%$. 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	366	<div> <div>63%</div> <div>23%</div> <div>6%</div> <div>7%</div> </div>
1	D	366	<div> <div>66%</div> <div>19%</div> <div>6%</div> <div>7%</div> </div>
2	C	24	<div> <div>4%</div> <div>67%</div> <div>21%</div> <div>13%</div> </div>
2	F	24	<div> <div>33%</div> <div>50%</div> <div>17%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5989 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 Mitogen-activated protein kinase 14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	339	Total	C	N	O	S	0	0	0
			2715	1740	467	496	12			
1	D	339	Total	C	N	O	S	0	0	0
			2710	1735	466	497	12			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	CLONING ARTIFACT	UNP Q16539
A	-4	SER	-	CLONING ARTIFACT	UNP Q16539
A	-3	HIS	-	CLONING ARTIFACT	UNP Q16539
A	-2	MET	-	CLONING ARTIFACT	UNP Q16539
A	-1	LEU	-	CLONING ARTIFACT	UNP Q16539
A	0	GLU	-	CLONING ARTIFACT	UNP Q16539
A	1	MET	-	CLONING ARTIFACT	UNP Q16539
D	-5	GLY	-	CLONING ARTIFACT	UNP Q16539
D	-4	SER	-	CLONING ARTIFACT	UNP Q16539
D	-3	HIS	-	CLONING ARTIFACT	UNP Q16539
D	-2	MET	-	CLONING ARTIFACT	UNP Q16539
D	-1	LEU	-	CLONING ARTIFACT	UNP Q16539
D	0	GLU	-	CLONING ARTIFACT	UNP Q16539
D	1	MET	-	CLONING ARTIFACT	UNP Q16539

- Molecule 2 is a protein called MAP kinase-activated protein kinase 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	24	Total	C	N	O	0	0	0
			180	114	36	30			
2	F	24	Total	C	N	O	0	0	0
			187	120	37	30			

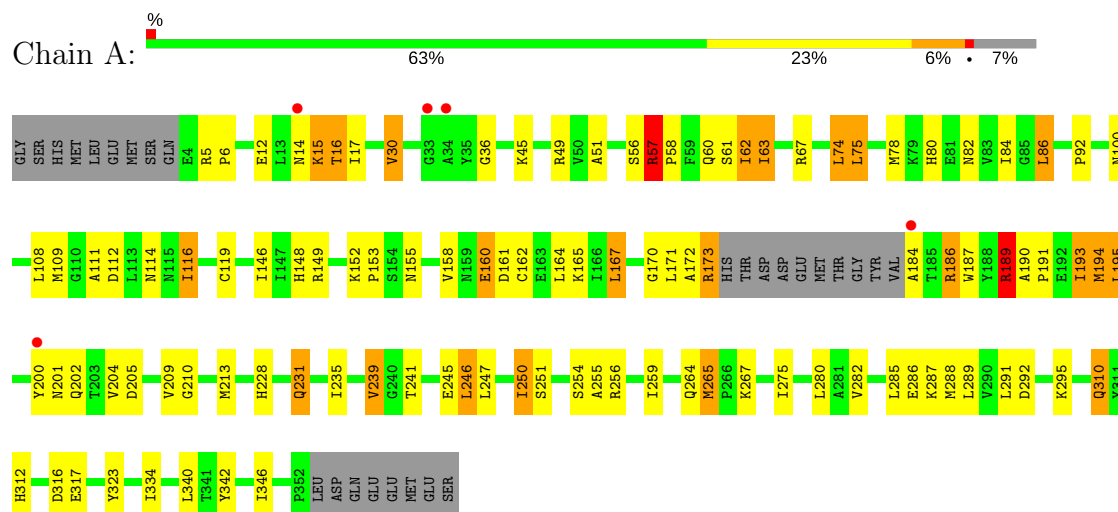
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	98	Total 98	O 98	0	0
3	D	97	Total 97	O 97	0	0
3	F	2	Total 2	O 2	0	0

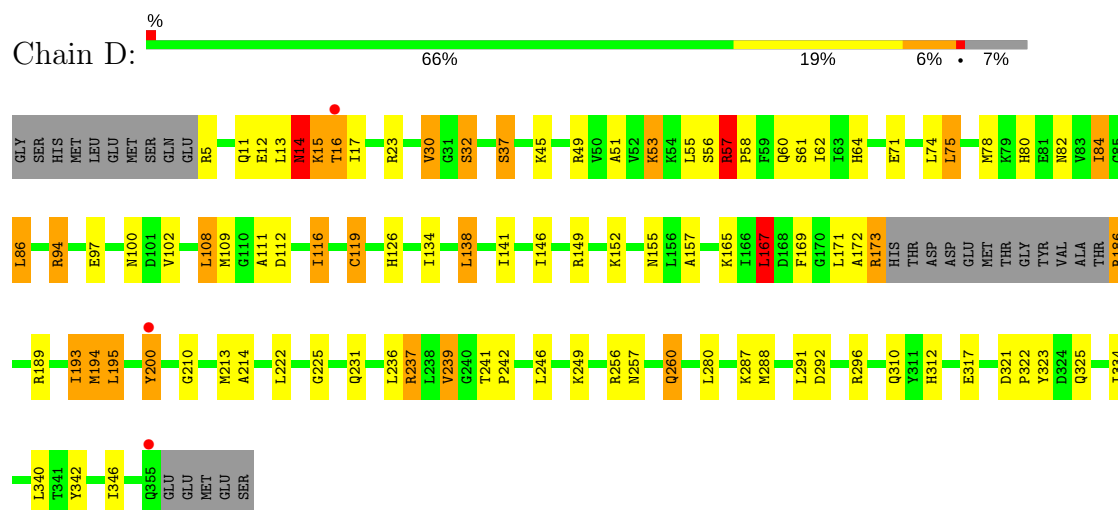
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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 ($\text{RSRZ} > 2$). 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: Mitogen-activated protein kinase 14

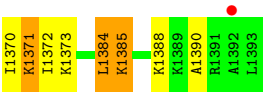


• Molecule 1: Mitogen-activated protein kinase 14

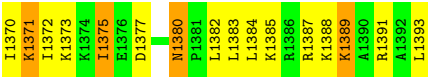


• Molecule 2: MAP kinase-activated protein kinase 2





● Molecule 2: MAP kinase-activated protein kinase 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α , β , γ	81.55Å 81.55Å 121.36Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	8.48 – 2.00 40.78 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.6 (8.48-2.00) 98.5 (40.78-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.05 (at 2.00Å)	Xtriage
Refinement program	AUTOBUSTER	Depositor
R, R_{free}	0.239 , 0.272 0.228 , 0.267	Depositor DCC
R_{free} test set	2981 reflections (5.22%)	DCC
Wilson B-factor (Å ²)	31.7	Xtriage
Anisotropy	0.462	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 13.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l 0.480 for h,-h-k,-l 0.006 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5989	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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 [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	0.59	0/2776	0.74	2/3766 (0.1%)
1	D	0.61	0/2771	0.85	5/3759 (0.1%)
2	C	0.43	0/180	0.83	1/236 (0.4%)
2	F	0.41	0/187	0.83	0/246
All	All	0.59	0/5914	0.80	8/8007 (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	D	0	2

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	16	THR	N-CA-CB	-19.86	72.57	110.30
1	D	16	THR	N-CA-C	9.06	135.45	111.00
1	A	189	ARG	NE-CZ-NH2	-8.18	116.21	120.30
1	D	84	ILE	CG1-CB-CG2	7.83	128.62	111.40
1	D	15	LYS	N-CA-C	6.38	128.22	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	14	ASN	Peptide
1	D	15	LYS	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	2715	0	2707	133	1
1	D	2710	0	2697	126	0
2	C	180	0	199	8	0
2	F	187	0	216	15	0
3	A	98	0	0	7	0
3	D	97	0	0	4	1
3	F	2	0	0	1	0
All	All	5989	0	5819	270	1

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 23.

The worst 5 of 270 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:D:173:ARG:HH21	1:D:173:ARG:HG3	1.27	0.97
1:A:264:GLN:HG2	1:A:265:MET:N	1.79	0.96
1:A:62:ILE:HG12	1:A:334:ILE:HD11	1.45	0.95
1:D:94:ARG:HG3	1:D:94:ARG:HH11	1.31	0.93
1:D:62:ILE:HD13	1:D:334:ILE:HD13	1.52	0.90

All (1) 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:A:264:GLN:NE2	3:D:365:HOH:O[1_565]	0.83	1.37

5.3 Torsion angles

5.3.1 Protein backbone

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	335/366 (92%)	324 (97%)	10 (3%)	1 (0%)	44	40
1	D	335/366 (92%)	321 (96%)	13 (4%)	1 (0%)	44	40
2	C	22/24 (92%)	21 (96%)	1 (4%)	0	100	100
2	F	22/24 (92%)	20 (91%)	2 (9%)	0	100	100
All	All	714/780 (92%)	686 (96%)	26 (4%)	2 (0%)	44	40

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	57	ARG
1	A	57	ARG

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	295/325 (91%)	262 (89%)	33 (11%)	7	4
1	D	295/325 (91%)	265 (90%)	30 (10%)	8	4
2	C	18/21 (86%)	14 (78%)	4 (22%)	1	0
2	F	19/21 (90%)	13 (68%)	6 (32%)	0	0
All	All	627/692 (91%)	554 (88%)	73 (12%)	6	3

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1370	ILE
1	D	37	SER
2	F	1370	ILE
1	D	14	ASN
1	D	57	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	312	HIS
1	D	14	ASN
1	D	312	HIS
1	D	11	GLN
1	D	60	GLN

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	339/366 (92%)	-0.17	5 (1%) 74 73	1, 12, 39, 57	0
1	D	339/366 (92%)	-0.23	3 (0%) 84 83	2, 12, 39, 60	0
2	C	24/24 (100%)	-0.19	1 (4%) 37 37	8, 20, 47, 49	0
2	F	24/24 (100%)	-0.25	0 100 100	10, 22, 45, 50	0
All	All	726/780 (93%)	-0.21	9 (1%) 79 78	1, 13, 42, 60	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	184	ALA	4.8
1	D	355	GLN	3.9
1	A	34	ALA	3.3
1	A	200	TYR	2.9
1	A	33	GLY	2.8

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.