wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2014 – 11:04 PM GMT

PDB ID	:	3HGK
Title	:	crystal structure of effect protein AvrptoB complexed with kinase Pto
Authors	:	Dong, J.; Fan, F.; Gu, L.; Chai, J.
Deposited on	:	2009-05-14
Resolution	:	3.30 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html

---

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

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 3.30 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	Percentile Ranks	Value
Rfree		0.347
Clashscore		29
Ramachandran outliers		7.2%
Sidechain outliers		7.6%
RSRZ outliers		2.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
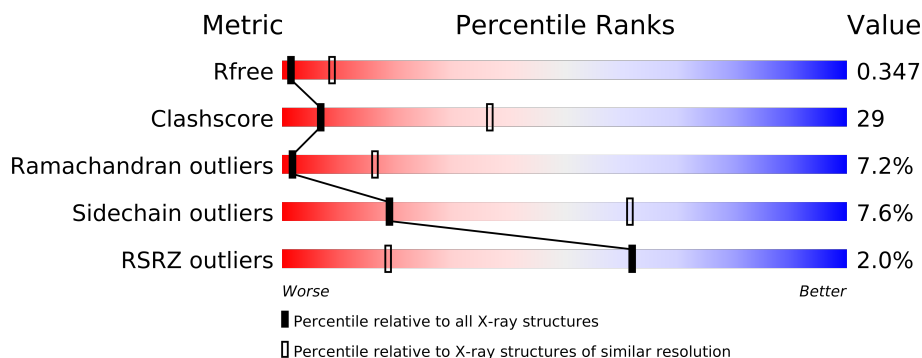
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1341 (3.40-3.20)
Clashscore	79885	1696 (3.40-3.20)
Ramachandran outliers	78287	1664 (3.40-3.20)
Sidechain outliers	78261	1662 (3.40-3.20)
RSRZ outliers	66119	1342 (3.40-3.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	327	
1	B	327	
1	C	327	
1	D	327	
2	E	85	
2	F	85	
2	G	85	
2	H	85	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11599 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Protein kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	P	S	0	0	0
			2308	1462	398	435	2	11			
1	B	288	Total	C	N	O	P	S	0	0	0
			2306	1459	398	436	2	11			
1	C	286	Total	C	N	O	P	S	0	0	0
			2293	1451	396	433	2	11			
1	D	288	Total	C	N	O	P	S	0	0	0
			2308	1462	398	435	2	11			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	193	GLY	ASP	ENGINEERED	UNP Q40234
A	322	HIS	-	EXPRESSION TAG	UNP Q40234
A	323	HIS	-	EXPRESSION TAG	UNP Q40234
A	324	HIS	-	EXPRESSION TAG	UNP Q40234
A	325	HIS	-	EXPRESSION TAG	UNP Q40234
A	326	HIS	-	EXPRESSION TAG	UNP Q40234
A	327	HIS	-	EXPRESSION TAG	UNP Q40234
B	193	GLY	ASP	ENGINEERED	UNP Q40234
B	322	HIS	-	EXPRESSION TAG	UNP Q40234
B	323	HIS	-	EXPRESSION TAG	UNP Q40234
B	324	HIS	-	EXPRESSION TAG	UNP Q40234
B	325	HIS	-	EXPRESSION TAG	UNP Q40234
B	326	HIS	-	EXPRESSION TAG	UNP Q40234
B	327	HIS	-	EXPRESSION TAG	UNP Q40234
C	193	GLY	ASP	ENGINEERED	UNP Q40234
C	322	HIS	-	EXPRESSION TAG	UNP Q40234
C	323	HIS	-	EXPRESSION TAG	UNP Q40234
C	324	HIS	-	EXPRESSION TAG	UNP Q40234
C	325	HIS	-	EXPRESSION TAG	UNP Q40234
C	326	HIS	-	EXPRESSION TAG	UNP Q40234
C	327	HIS	-	EXPRESSION TAG	UNP Q40234

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	193	GLY	ASP	ENGINEERED	UNP Q40234
D	322	HIS	-	EXPRESSION TAG	UNP Q40234
D	323	HIS	-	EXPRESSION TAG	UNP Q40234
D	324	HIS	-	EXPRESSION TAG	UNP Q40234
D	325	HIS	-	EXPRESSION TAG	UNP Q40234
D	326	HIS	-	EXPRESSION TAG	UNP Q40234
D	327	HIS	-	EXPRESSION TAG	UNP Q40234

- Molecule 2 is a protein called Effector protein hopAB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	77	Total	C	N	O	S	0	0	0
			596	363	121	108	4			
2	F	77	Total	C	N	O	S	0	0	0
			596	363	121	108	4			
2	G	77	Total	C	N	O	S	0	0	0
			596	363	121	108	4			
2	H	77	Total	C	N	O	S	0	0	0
			596	363	121	108	4			

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 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: Protein kinase

Chain A:

• Molecule 1: Protein kinase

Chain B:

• Molecule 1: Protein kinase

Chain C:

• Molecule 1: Protein kinase

Chain D:

• Molecule 1: Protein kinase

Chain D:

• Molecule 2: Effector protein hopAB2

Chain E:

• Molecule 2: Effector protein hopAB2

Chain F:

• Molecule 2: Effector protein hopAB2

Chain G:

PRO ARG ARG G124 A125 V126 N130 S131 I132 V133 Q134 Q135 L136 R147 D157 V164 E165 Q166 N167 I168 Q171 M176 P177 G180 I181 S182 R183 E190 L195 R196 V199 H200 GLN GLN ALA ALA SER

• Molecule 2: Effector protein hopAB2

Chain H:

PRO ARG ARG G124 A125 V126 V133 L136 A141 S144 R147 R151 N155 M176 G180 I181 S182 R183 E186 I189 E190 L191 R192 G193 A194 L195 R196 R197 H200 GLN GLN ALA ALA SER

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.07Å 104.47Å 298.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30 44.96 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.00-3.30) 99.6 (44.96-3.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.03 (at 3.32Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.317 , 0.331 0.331 , 0.347	Depositor DCC
R_{free} test set	1508 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	95.1	Xtriage
Anisotropy	0.895	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 18.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Outliers	1 of 29705 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	11599	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.13 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.7970e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.51	1/2331 (0.0%)	0.66	0/3143
1	B	0.51	0/2329	0.68	1/3140 (0.0%)
1	C	0.49	0/2316	0.67	0/3122
1	D	0.50	0/2331	0.65	0/3143
2	E	0.38	0/605	0.50	0/814
2	F	0.41	0/605	0.51	0/814
2	G	0.59	2/605 (0.3%)	0.58	1/814 (0.1%)
2	H	0.40	0/605	0.51	0/814
All	All	0.49	3/11727 (0.0%)	0.64	2/15804 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	124	GLY	C-O	-7.03	1.12	1.23
2	G	124	GLY	CA-C	-5.72	1.42	1.51
1	A	126	LEU	C-N	-5.23	1.22	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	124	GLY	N-CA-C	-7.49	94.38	113.10
1	B	47	ILE	N-CA-C	6.44	128.40	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2　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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2308	0	2301	172	0
1	B	2306	0	2295	156	0
1	C	2293	0	2281	175	0
1	D	2308	0	2301	141	0
2	E	596	0	585	18	0
2	F	596	0	585	17	0
2	G	596	0	585	18	0
2	H	596	0	585	17	0
All	All	11599	0	11518	671	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 29.

The worst 5 of 671 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:245:LEU:HB3	1:D:240:ALA:HB2	1.38	1.06
1:B:240:ALA:HB2	1:D:245:LEU:HB3	1.40	1.03
1:C:238:ARG:HD2	1:C:252:LEU:HD13	1.43	0.98
1:A:203:GLY:HA3	1:A:208:ILE:HD11	1.43	0.97
1:C:77:GLN:HE21	1:C:80:GLU:HB2	1.26	0.97

There are no symmetry-related clashes.

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	284/327 (87%)	209 (74%)	54 (19%)	21 (7%)	2	15
1	B	284/327 (87%)	194 (68%)	63 (22%)	27 (10%)	1	9
1	C	282/327 (86%)	201 (71%)	60 (21%)	21 (7%)	2	15
1	D	284/327 (87%)	209 (74%)	49 (17%)	26 (9%)	1	9
2	E	75/85 (88%)	62 (83%)	11 (15%)	2 (3%)	8	49
2	F	75/85 (88%)	62 (83%)	11 (15%)	2 (3%)	8	49
2	G	75/85 (88%)	61 (81%)	12 (16%)	2 (3%)	8	49
2	H	75/85 (88%)	67 (89%)	6 (8%)	2 (3%)	8	49
All	All	1434/1648 (87%)	1065 (74%)	266 (18%)	103 (7%)	2	16

5 of 103 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	ASN
1	A	47	ILE
1	A	92	CYS
1	B	39	ASN
1	B	47	ILE

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	251/287 (88%)	230 (92%)	21 (8%)	16	55
1	B	251/287 (88%)	221 (88%)	30 (12%)	7	33
1	C	249/287 (87%)	222 (89%)	27 (11%)	9	39
1	D	251/287 (88%)	235 (94%)	16 (6%)	25	69
2	E	62/68 (91%)	62 (100%)	0	100	100
2	F	62/68 (91%)	62 (100%)	0	100	100
2	G	62/68 (91%)	62 (100%)	0	100	100
2	H	62/68 (91%)	61 (98%)	1 (2%)	75	94
All	All	1250/1420 (88%)	1155 (92%)	95 (8%)	19	61

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	250	VAL
1	C	42	ASP
1	D	195	THR
1	B	264	GLN
1	B	314	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	77	GLN
1	D	243	GLN
2	G	155	ASN
1	D	94	HIS
1	D	96	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

8 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	#$|Z| > 2$
1	SEP	A	198	1	9,9,10	7.42	5 (55%)	10,12,14	1.20	0
1	TPO	A	199	1	10,10,11	6.12	2 (20%)	12,14,16	1.03	0
1	SEP	B	198	1	9,9,10	7.39	5 (55%)	10,12,14	2.10	2 (20%)
1	TPO	B	199	1	10,10,11	6.34	2 (20%)	12,14,16	1.20	0
1	SEP	C	198	1	9,9,10	7.55	5 (55%)	10,12,14	1.47	2 (20%)
1	TPO	C	199	1	10,10,11	6.13	2 (20%)	12,14,16	0.98	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$\|Z\|$ > 2	Counts	RMSZ	#$\|Z\|$ > 2
1	SEP	D	198	1	9,9,10	7.26	5 (55%)	10,12,14	1.24	1 (10%)
1	TPO	D	199	1	10,10,11	6.27	2 (20%)	12,14,16	1.12	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
1	SEP	A	198	1	-	0/6/8/10	0/0/0/0
1	TPO	A	199	1	-	1/9/11/13	0/0/0/0
1	SEP	B	198	1	-	0/6/8/10	0/0/0/0
1	TPO	B	199	1	-	1/9/11/13	0/0/0/0
1	SEP	C	198	1	-	0/6/8/10	0/0/0/0
1	TPO	C	199	1	-	1/9/11/13	0/0/0/0
1	SEP	D	198	1	-	0/6/8/10	0/0/0/0
1	TPO	D	199	1	-	1/9/11/13	0/0/0/0

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	198	SEP	O-C	21.45	1.26	1.11
1	A	198	SEP	O-C	21.31	1.26	1.11
1	B	198	SEP	O-C	21.00	1.25	1.11
1	D	198	SEP	O-C	20.91	1.25	1.11
1	B	199	TPO	O-C	19.73	1.25	1.11

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	198	SEP	C-CA-N	5.00	118.83	113.83
1	B	198	SEP	P-OG-CB	-3.51	108.03	118.19
1	C	198	SEP	P-OG-CB	-3.17	109.03	118.19
1	C	198	SEP	C-CA-N	3.07	116.90	113.83
1	D	198	SEP	P-OG-CB	-2.36	111.36	118.19

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	199	TPO	OG1-CB-CA-N

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	D	199	TPO	OG1-CB-CA-N
1	B	199	TPO	OG1-CB-CA-N
1	C	199	TPO	OG1-CB-CA-N

There are no ring outliers.

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

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(Å²)	Q<0.9
1	A	288/327 (88%)	0.34	7 (2%)	56	15	48, 81, 109, 115	0
1	B	288/327 (88%)	0.21	4 (1%)	72	25	50, 75, 98, 107	0
1	C	286/327 (87%)	0.22	7 (2%)	56	15	47, 75, 99, 110	0
1	D	288/327 (88%)	0.33	6 (2%)	60	17	50, 82, 103, 109	0
2	E	77/85 (90%)	0.28	0	100	100	74, 90, 105, 107	0
2	F	77/85 (90%)	0.32	3 (3%)	37	8	72, 85, 109, 113	0
2	G	77/85 (90%)	0.32	1 (1%)	74	27	83, 102, 113, 115	0
2	H	77/85 (90%)	0.27	1 (1%)	74	27	76, 85, 95, 101	0
All	All	1458/1648 (88%)	0.28	29 (1%)	62	19	47, 81, 105, 115	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	194	ALA	4.0
1	A	57	LYS	3.0
1	A	63	GLY	2.9
2	G	132	ILE	2.8
1	B	77	GLN	2.8

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, 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
1	SEP	B	198	10/11	0.45	0.76	78,82,83,88	0
1	SEP	C	198	10/11	0.50	0.69	82,85,86,89	0
1	TPO	C	199	11/12	0.26	-0.56	82,83,85,88	0
1	SEP	D	198	10/11	0.22	-0.64	78,79,81,85	0
1	TPO	D	199	11/12	0.16	-1.18	79,80,81,82	0
1	SEP	A	198	10/11	0.20	-1.22	84,85,86,90	0
1	TPO	B	199	11/12	0.17	-1.58	79,79,80,80	0
1	TPO	A	199	11/12	0.16	-2.17	83,83,85,87	0

6.3   Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5   Other polymers ⓘ

There are no such residues in this entry.