



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:25 PM GMT

PDB ID : 1K8K  
Title : Crystal Structure of Arp2/3 Complex  
Authors : Robinson, R.C.; Turbedsky, K.; Kaiser, D.A.; Higgs, H.N.; Marchand, J.-B.;  
Choe, S.; Pollard, T.D.  
Deposited on : 2001-10-24  
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](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

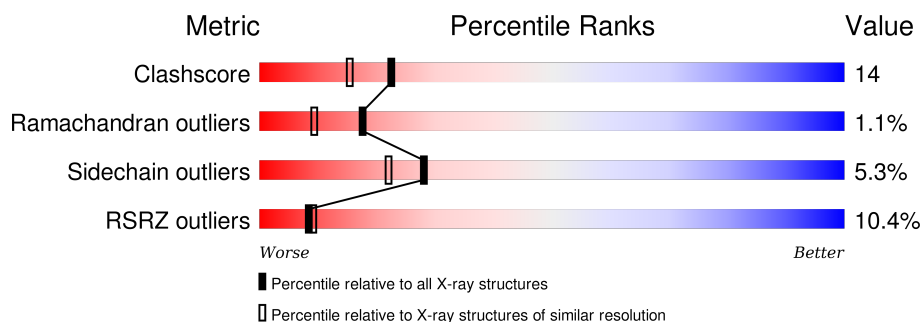
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (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  $\geq 3$ , 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	418	<div> <div>9%</div> <div>78%</div> <div>14%</div> <div>• •</div> </div>
2	B	394	<div> <div>11%</div> <div>31%</div> <div>14%</div> <div>• •</div> <div>52%</div> </div>
3	C	372	<div> <div>9%</div> <div>74%</div> <div>18%</div> <div>• 5%</div> </div>
4	D	300	<div> <div>3%</div> <div>80%</div> <div>13%</div> <div>• 5%</div> </div>
5	E	178	<div> <div>13%</div> <div>62%</div> <div>28%</div> <div>7%</div> <div>• •</div> </div>
6	F	168	<div> <div>4%</div> <div>85%</div> <div>14%</div> <div>• •</div> </div>
7	G	151	<div> <div>19%</div> <div>75%</div> <div>11%</div> <div>5%</div> <div>8%</div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 15326 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 ACTIN-LIKE PROTEIN 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	401	Total	C	N	O	S	0	0	0
			3212	2063	536	597	16			

- Molecule 2 is a protein called ACTIN-LIKE PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	190	Total	C	N	O	S	0	0	0
			1515	974	259	278	4			

- Molecule 3 is a protein called ARP2/3 COMPLEX 41 KDA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	354	Total	C	N	O	S	0	0	0
			2757	1747	487	504	19			

- Molecule 4 is a protein called ARP2/3 COMPLEX 34 KDA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	284	Total	C	N	O	S	0	0	0
			2291	1456	397	430	8			

- Molecule 5 is a protein called ARP2/3 COMPLEX 21 KDA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	174	Total	C	N	O	S	0	0	0
			1414	908	236	261	9			

- Molecule 6 is a protein called ARP2/3 COMPLEX 20 KDA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	167	Total	C	N	O	S	0	0	0
			1369	875	239	246	9			

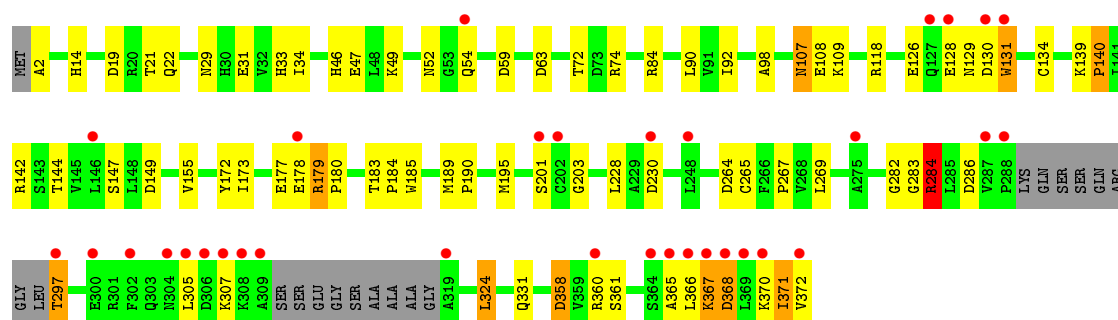
- Molecule 7 is a protein called ARP2/3 COMPLEX 16 KDA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	139	Total	C	N	O	S	0	0	0
			1058	661	185	209	3			

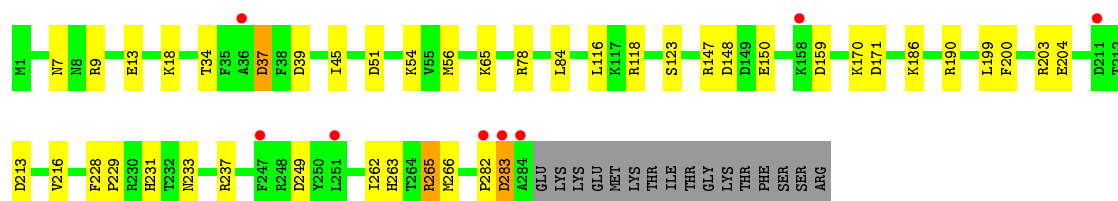
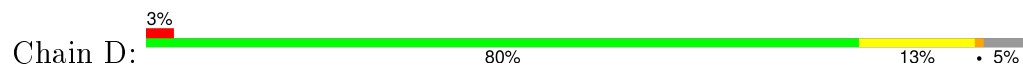
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	364	Total	O	27	0
			364	364		
8	B	209	Total	O	14	0
			209	209		
8	C	338	Total	O	30	0
			338	338		
8	D	355	Total	O	22	0
			355	355		
8	E	111	Total	O	12	0
			111	111		
8	F	230	Total	O	14	0
			230	230		
8	G	103	Total	O	9	0
			103	103		

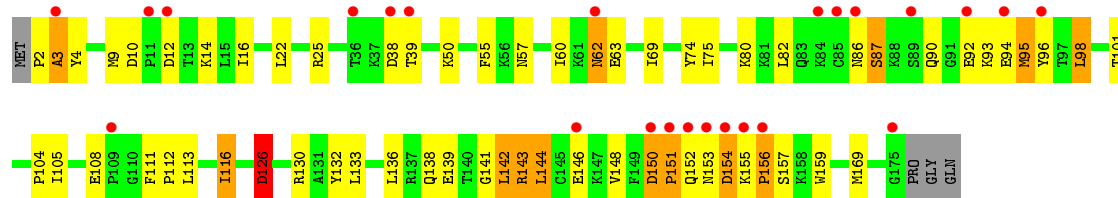




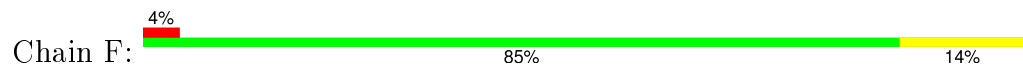
• Molecule 4: ARP2/3 COMPLEX 34 KDA SUBUNIT



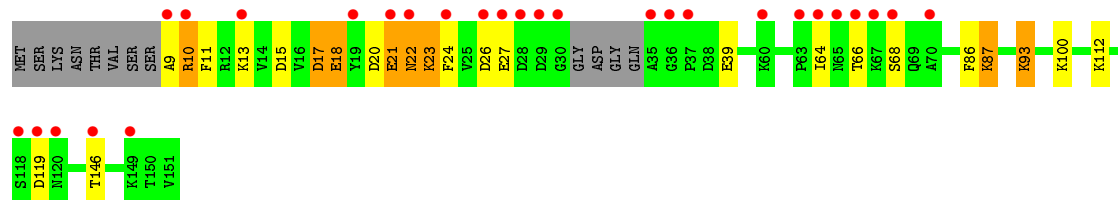
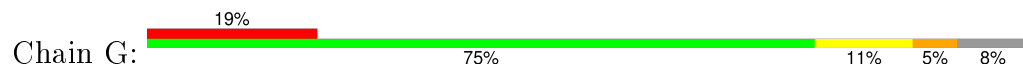
• Molecule 5: ARP2/3 COMPLEX 21 KDA SUBUNIT



• Molecule 6: ARP2/3 COMPLEX 20 KDA SUBUNIT



• Molecule 7: ARP2/3 COMPLEX 16 KDA SUBUNIT



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.71Å 130.40Å 204.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 41.54 – 2.01	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.00) 92.1 (41.54-2.01)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.42 (at 2.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.216 , 0.251 0.229 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	30.4	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 183319 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15326	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% 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](#)

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	0.49	0/3293	0.79	13/4467 (0.3%)
2	B	0.40	0/1543	0.69	8/2084 (0.4%)
3	C	0.51	0/2826	0.79	6/3832 (0.2%)
4	D	0.53	0/2340	0.79	8/3160 (0.3%)
5	E	0.38	0/1448	0.70	5/1953 (0.3%)
6	F	0.56	0/1391	0.73	1/1867 (0.1%)
7	G	0.38	0/1070	0.73	3/1441 (0.2%)
All	All	0.48	0/13911	0.76	44/18804 (0.2%)

There are no bond length outliers.

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	59	ASP	CB-CG-OD2	7.20	124.78	118.30
4	D	148	ASP	CB-CG-OD2	6.73	124.36	118.30
4	D	265	ARG	NE-CZ-NH2	-6.62	116.99	120.30
1	A	4	ARG	NE-CZ-NH2	6.60	123.60	120.30
1	A	11	ASP	CB-CG-OD2	6.42	124.08	118.30

There are no chirality outliers.

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	3212	0	3163	67	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1515	0	1548	57	0
3	C	2757	0	2713	96	0
4	D	2291	0	2257	38	0
5	E	1414	0	1416	75	0
6	F	1369	0	1410	33	0
7	G	1058	0	1065	26	0
8	A	364	0	0	13	0
8	B	209	0	0	15	0
8	C	338	0	0	36	0
8	D	355	0	0	15	0
8	E	111	0	0	4	0
8	F	230	0	0	5	0
8	G	103	0	0	8	0
All	All	15326	0	13572	373	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 14.

The worst 5 of 373 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)
5:E:152:GLN:CB	5:E:155:LYS:NZ	1.85	1.38
2:B:165:HIS:CD2	2:B:181:ARG:HG2	1.58	1.37
5:E:152:GLN:HB3	5:E:155:LYS:CE	1.59	1.30
5:E:152:GLN:CB	5:E:155:LYS:CE	2.12	1.27
3:C:307:LYS:CE	8:C:568:HOH:O	1.63	1.26

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	395/418 (94%)	377 (95%)	16 (4%)	2 (0%)	34	26
2	B	188/394 (48%)	167 (89%)	15 (8%)	6 (3%)	5	1
3	C	348/372 (94%)	328 (94%)	18 (5%)	2 (1%)	30	22
4	D	282/300 (94%)	275 (98%)	6 (2%)	1 (0%)	39	33
5	E	172/178 (97%)	160 (93%)	8 (5%)	4 (2%)	8	3
6	F	165/168 (98%)	160 (97%)	5 (3%)	0	100	100
7	G	135/151 (89%)	130 (96%)	2 (2%)	3 (2%)	8	3
All	All	1685/1981 (85%)	1597 (95%)	70 (4%)	18 (1%)	17	9

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	313	ARG
2	B	155	VAL
2	B	171	GLU
2	B	336	LYS
4	D	283	ASP

### 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	352/363 (97%)	335 (95%)	17 (5%)	31	26
2	B	165/345 (48%)	151 (92%)	14 (8%)	13	8
3	C	301/313 (96%)	289 (96%)	12 (4%)	38	33
4	D	249/264 (94%)	242 (97%)	7 (3%)	51	50
5	E	156/159 (98%)	141 (90%)	15 (10%)	10	6
6	F	154/155 (99%)	150 (97%)	4 (3%)	54	54
7	G	114/124 (92%)	104 (91%)	10 (9%)	12	7
All	All	1491/1723 (86%)	1412 (95%)	79 (5%)	28	22

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	134	CYS
4	D	37	ASP
7	G	64	ILE
3	C	140	PRO
3	C	305	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	323	GLN
3	C	107	ASN
7	G	22	ASN
3	C	33	HIS
3	C	65	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](#)

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, 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	401/418 (95%)	0.56	36 (8%) 12 12	18, 35, 76, 97	0
2	B	190/394 (48%)	1.02	42 (22%) 1 1	22, 46, 91, 109	0
3	C	354/372 (95%)	0.52	33 (9%) 11 11	20, 30, 76, 106	0
4	D	284/300 (94%)	0.26	8 (2%) 56 57	18, 30, 44, 55	0
5	E	174/178 (97%)	0.79	24 (13%) 4 4	33, 48, 77, 80	0
6	F	167/168 (99%)	0.24	7 (4%) 40 41	18, 26, 37, 58	0
7	G	139/151 (92%)	1.03	28 (20%) 1 2	23, 47, 95, 104	0
All	All	1709/1981 (86%)	0.58	178 (10%) 8 9	18, 34, 80, 109	0

The worst 5 of 178 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	284	ALA	14.9
3	C	309	ALA	9.5
7	G	35	ALA	9.3
5	E	96	TYR	8.5
2	B	337	LEU	8.1

### 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.