



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 01:34 AM GMT

PDB ID : 2DKO  
Title : Extended substrate recognition in caspase-3 revealed by high resolution X-ray structure analysis  
Authors : Mittl, P.R.E.; Ganesan, R.; Jelakovic, S.; Grutter, M.G.  
Deposited on : 2006-04-12  
Resolution : 1.06 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](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

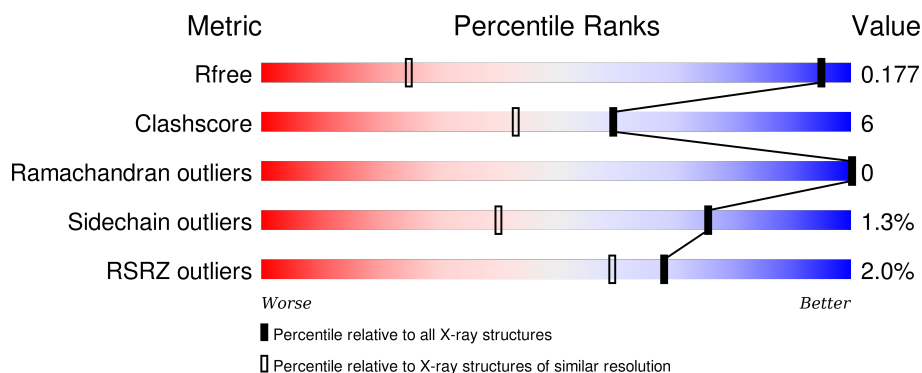
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.06 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	1077 (1.12-1.00)
Clashscore	102246	1147 (1.12-1.00)
Ramachandran outliers	100387	1086 (1.12-1.00)
Sidechain outliers	100360	1084 (1.12-1.00)
RSRZ outliers	91569	1080 (1.12-1.00)

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	146	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, orange 1%, orange 15%, yellow 15%, yellow 84%, green 84%, green 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>84%</span> <span>15%</span> <span>.</span> </div> </div>
2	B	103	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 3%, orange 3%, orange 17%, yellow 17%, yellow 79%, green 79%, green 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>79%</span> <span>17%</span> <span>5%</span> </div> </div>
3	I	6	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 0%, green 50%, yellow 50%, yellow 83%, grey 83%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>50%</span> <span>33%</span> <span>17%</span> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2427 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 Caspase-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	146	Total	C	N	O	S	0	11	0
			1186	739	207	229	11			

- Molecule 2 is a protein called Caspase-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	103	Total	C	N	O	S	0	9	0
			873	570	136	158	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	ALA	ASP	ENGINEERED MUTATION	UNP P42574

- Molecule 3 is a protein called PHQ-ASP-GLU-VAL-ASP-CHLOROMETHYLKETONE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	5	Total	C	N	O	0	0	1
			33	19	4	10			

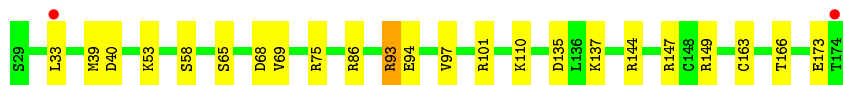
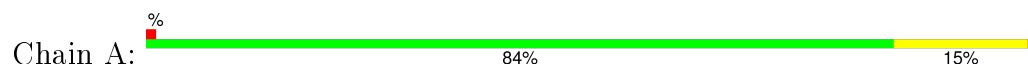
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	219	Total	O	0	0
			219	219		
4	B	109	Total	O	0	0
			109	109		
4	I	7	Total	O	0	0
			7	7		

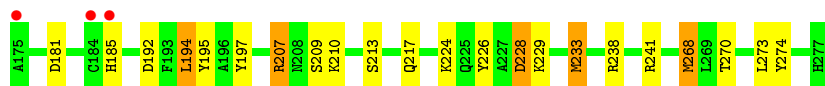
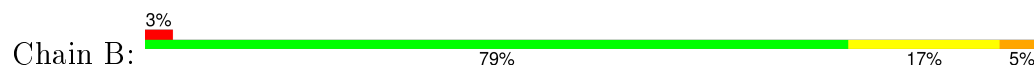
### 3 Residue-property plots [i](#)

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: Caspase-3



#### • Molecule 2: Caspase-3



#### • Molecule 3: PHQ-ASP-GLU-VAL-ASP-CHLOROMETHYLKETONE



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.65Å 83.89Å 96.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.06 13.48 – 1.06	Depositor EDS
% Data completeness (in resolution range)	89.9 (20.00-1.06) 91.6 (13.48-1.06)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 1.06Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.142 , 0.175 0.143 , 0.177	Depositor DCC
$R_{free}$ test set	2049 reflections (1.84%)	DCC
Wilson B-factor (Å <sup>2</sup> )	10.3	Xtriage
Anisotropy	0.396	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	-0.03 , 15.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 113404 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2427	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	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 21.60 % 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 6.7791e-03.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 0QE

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	0.80	0/1246	1.39	18/1668 (1.1%)
2	B	0.90	0/934	1.52	20/1261 (1.6%)
3	I	1.02	0/31	1.40	0/41
All	All	0.84	0/2211	1.45	38/2970 (1.3%)

There are no bond length outliers.

All (38) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	238	ARG	NE-CZ-NH1	14.02	127.31	120.30
2	B	185	HIS	CA-CB-CG	12.32	134.55	113.60
1	A	75	ARG	NE-CZ-NH1	12.05	126.33	120.30
1	A	40	ASP	CB-CG-OD2	10.54	127.79	118.30
1	A	101	ARG	CD-NE-CZ	10.21	137.89	123.60
1	A	86	ARG	NE-CZ-NH2	-7.83	116.39	120.30
2	B	217[A]	GLN	CA-CB-CG	7.81	130.59	113.40
2	B	217[B]	GLN	CA-CB-CG	7.81	130.59	113.40
1	A	144	ARG	NE-CZ-NH2	-7.75	116.43	120.30
2	B	207	ARG	NE-CZ-NH1	7.08	123.84	120.30
2	B	228	ASP	CB-CG-OD1	6.92	124.53	118.30
2	B	194[A]	LEU	CA-CB-CG	6.89	131.15	115.30
2	B	194[B]	LEU	CA-CB-CG	6.89	131.15	115.30
1	A	93	ARG	NE-CZ-NH1	6.64	123.62	120.30
1	A	147	ARG	NE-CZ-NH1	6.57	123.58	120.30
2	B	233	MET	CG-SD-CE	-6.26	90.18	100.20
2	B	241	ARG	NE-CZ-NH1	6.07	123.34	120.30
2	B	197	TYR	CG-CD2-CE2	6.02	126.12	121.30
2	B	195	TYR	CG-CD2-CE2	-6.02	116.48	121.30
1	A	144	ARG	NE-CZ-NH1	5.93	123.27	120.30
1	A	149	ARG	NE-CZ-NH1	5.89	123.25	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	58[A]	SER	N-CA-CB	-5.84	101.74	110.50
1	A	58[B]	SER	N-CA-CB	-5.84	101.74	110.50
1	A	149	ARG	NE-CZ-NH2	-5.79	117.40	120.30
2	B	207	ARG	NE-CZ-NH2	-5.77	117.41	120.30
1	A	135	ASP	CB-CG-OD1	-5.68	113.19	118.30
2	B	226	TYR	CD1-CE1-CZ	-5.64	114.72	119.80
1	A	68	ASP	CB-CG-OD1	5.60	123.34	118.30
2	B	181	ASP	CB-CG-OD2	5.44	123.19	118.30
2	B	274	TYR	CB-CG-CD1	5.39	124.23	121.00
2	B	241	ARG	CG-CD-NE	5.38	123.11	111.80
1	A	75	ARG	NE-CZ-NH2	-5.34	117.63	120.30
1	A	33	LEU	O-C-N	-5.31	114.21	122.70
1	A	94	GLU	OE1-CD-OE2	-5.24	117.01	123.30
2	B	241	ARG	NE-CZ-NH2	-5.11	117.75	120.30
2	B	228	ASP	CB-CG-OD2	-5.09	113.72	118.30
1	A	40	ASP	CB-CG-OD1	-5.02	113.78	118.30
2	B	238	ARG	NE-CZ-NH2	-5.00	117.80	120.30

There are no chirality outliers.

There are no planarity outliers.

## 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	1186	0	1189	16	0
2	B	873	0	843	9	0
3	I	33	0	22	3	0
4	A	219	0	0	10	0
4	B	109	0	0	3	0
4	I	7	0	0	0	0
All	All	2427	0	2054	23	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 6.

All (23) 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)
2:B:228:ASP:OD1	2:B:229:LYS:HG3	1.90	0.72
1:A:163:CYS:SG	3:I:6:0QE:C1	2.79	0.70
1:A:110:LYS:HE2	4:A:337:HOH:O	1.91	0.69
1:A:166[B]:THR:HG23	4:A:200:HOH:O	1.98	0.62
1:A:163:CYS:SG	3:I:5:ASP:CA	2.87	0.62
2:B:192:ASP:OD1	2:B:270[B]:THR:HG22	2.02	0.60
4:A:267:HOH:O	2:B:224:LYS:HE2	2.02	0.58
1:A:93:ARG:O	1:A:97[B]:VAL:HG22	2.03	0.58
1:A:137:LYS:HE3	4:A:198:HOH:O	2.03	0.58
1:A:39:MET:SD	2:B:273[B]:LEU:HD21	2.45	0.57
1:A:69:VAL:HG23	4:A:344:HOH:O	2.10	0.52
1:A:166[B]:THR:HG22	4:B:290:HOH:O	2.10	0.51
2:B:233:MET:HE2	4:B:284:HOH:O	2.11	0.49
2:B:270[B]:THR:HG21	4:B:304:HOH:O	2.11	0.49
1:A:163:CYS:SG	3:I:5:ASP:O	2.64	0.49
1:A:173:GLU:OE1	1:A:173:GLU:HA	2.16	0.45
1:A:93:ARG:NH2	4:A:301:HOH:O	2.49	0.45
1:A:53:LYS:NZ	4:A:366:HOH:O	2.49	0.44
4:A:361:HOH:O	2:B:268[A]:MET:HE2	2.17	0.44
1:A:97[B]:VAL:HG23	4:A:262:HOH:O	2.19	0.42
1:A:65[B]:SER:OG	2:B:209:SER:HB2	2.20	0.41
2:B:207:ARG:HA	2:B:213:SER:HA	2.01	0.41
1:A:93:ARG:NH1	4:A:341:HOH:O	2.48	0.40

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	155/146 (106%)	153 (99%)	2 (1%)	0	100	100
2	B	110/103 (107%)	109 (99%)	1 (1%)	0	100	100
3	I	2/6 (33%)	2 (100%)	0	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	267/255 (105%)	264 (99%)	3 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	142/131 (108%)	142 (100%)	0	100	100
2	B	99/90 (110%)	94 (95%)	5 (5%)	29	3
3	I	4/4 (100%)	4 (100%)	0	100	100
All	All	245/225 (109%)	240 (98%)	5 (2%)	76	24

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	194[A]	LEU
2	B	194[B]	LEU
2	B	210	LYS
2	B	268[A]	MET
2	B	268[B]	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	ASN
2	B	234	HIS

### 5.3.3 RNA ⓘ

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	146/146 (100%)	0.41	2 (1%) 78 70	11, 16, 29, 38	1 (0%)
2	B	103/103 (100%)	0.50	3 (2%) 55 48	9, 14, 30, 42	0
3	I	4/6 (66%)	0.32	0 100 100	18, 19, 20, 21	0
All	All	253/255 (99%)	0.45	5 (1%) 68 59	9, 15, 30, 42	1 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	175	ALA	3.7
1	A	33	LEU	2.8
1	A	174	THR	2.6
2	B	184[A]	CYS	2.4
2	B	185	HIS	2.3

### 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.