

Full wwPDB X-ray Structure Validation Report (i)

Oct 14, 2023 – 04:30 PM EDT

PDB ID : 8DSO

Title: Structure of cIAP1, BTK and BCCov

Authors: Schiemer, J.S.; Calabrese, M.F.

Deposited on : 2022-07-22

Resolution : 2.33 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

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

MolProbity: 4.02b-467

Mogul : 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13

EDS : 2.36

buster-report : 1.1.7 (2018)

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

Refmac : 5.8.0158

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

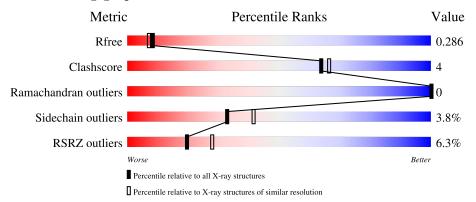
Validation Pipeline (wwPDB-VP) : 2.36

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 2.33 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	Similar resolution
Metric	$(\# ext{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	В	277	82%	10% 7%					
2	D	99	79%	11% • 8%					



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2917 atoms, of which 80 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 Tyrosine-protein kinase BTK.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	В	258	Total 1970	C 1272	N 329	O 350	S 19	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
В	383	SER	-	expression tag	UNP Q06187

• Molecule 2 is a protein called Baculoviral IAP repeat-containing protein 2.

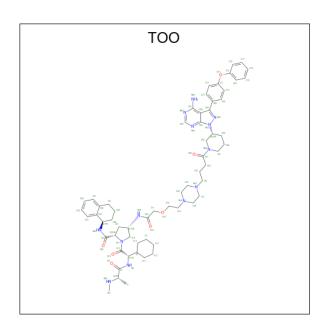
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	D	91	Total	С	N	О	S	0	0	0
	ש	91	713	453	124	127	9	U	0	U

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	254	GLY	-	expression tag	UNP Q13490
D	255	SER	-	expression tag	UNP Q13490
D	256	GLY	-	expression tag	UNP Q13490
D	257	PRO	-	expression tag	UNP Q13490
D	258	GLY	-	expression tag	UNP Q13490
D	259	SER	-	expression tag	UNP Q13490

• Molecule 3 is (4S)-4-[2-(2-{4-[(2E)-4-{(3R)-3-[4-amino-3-(4-phenoxyphenyl)-1H-pyrazolo[3,4-d]pyrimidin-1-yl]piperidin-1-yl}-4-oxobut-2-en-1-yl]piperazin-1-yl}ethoxy)acetamido]-1-{(2 S)-2-cyclohexyl-2-[(N-methyl-L-alanyl)amino]acetyl}-N-[(1R)-1,2,3,4-tetrahydronaphthalen-1-yl]-L-prolinamide bound form (three-letter code: TOO) (formula: $C_{61}H_{81}N_{13}O_7$) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
9	D	1	Total	С	Н	N	О	90	0
3	Б	1	161	61	80	13	7	80	U

• Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total Zn 1 1	0	0

• Molecule 5 is water.

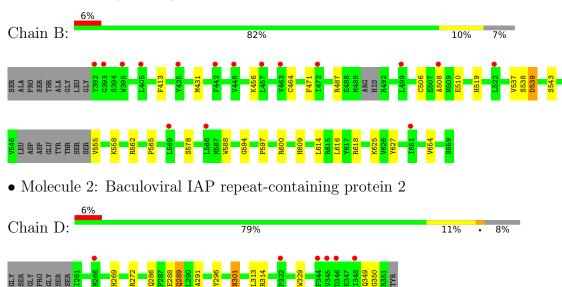
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	В	49	Total O 49 49	0	0
5	D	23	Total O 23 23	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 (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: Tyrosine-protein kinase BTK





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants	104.36Å 104.36Å 110.29Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 120.00°	Depositor
Resolution (Å)	110.29 - 2.33	Depositor
Resolution (A)	110.29 - 2.33	EDS
% Data completeness	59.3 (110.29-2.33)	Depositor
(in resolution range)	59.3 (110.29-2.33)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.15 (at 2.34Å)	Xtriage
Refinement program	BUSTER 2.11.8	Depositor
D D.	0.242 , 0.292	Depositor
R, R_{free}	0.232 , 0.286	DCC
R_{free} test set	854 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å ²)	65.7	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.36 , 70.6	EDS
L-test for twinning ²	$< L >=0.43, < L^2>=0.25$	Xtriage
Estimated twinning fraction	0.068 for -h,-k,l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2917	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.28% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical values of <|L|>, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

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

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		
MIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	В	0.41	0/2017	0.58	0/2731	
2	D	0.40	0/736	0.54	0/998	
All	All	0.41	0/2753	0.57	0/3729	

There are no bond length outliers.

There are no bond angle outliers.

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	В	1970	0	1832	15	0
2	D	713	0	638	7	0
3	В	81	80	0	0	0
4	D	1	0	0	0	0
5	В	49	0	0	0	0
5	D	23	0	0	2	0
All	All	2837	80	2470	22	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 4.

All (22) close contacts within the same asymmetric unit are listed below, sorted by their clash



magnitude.

Atom-1	Atom-2	Interatomic	Clash	
Atom-1	Atom-2	${\rm distance}({\rm \AA})$	overlap (Å)	
1:B:555:VAL:CB	1:B:558:LYS:HE3	2.07	0.85	
2:D:272:ARG:NH1	2:D:291:ALA:O	2.29	0.65	
2:D:272:ARG:HG2	2:D:296:TYR:HB3	1.78	0.63	
2:D:350:GLY:HA3	5:D:513:HOH:O	2.09	0.53	
2:D:313:LEU:HD21	2:D:329:TRP:CD1	2.46	0.51	
1:B:588:TRP:CG	1:B:616:LEU:HD13	2.46	0.49	
1:B:519:HIS:CD2	1:B:537:VAL:CG1	2.96	0.49	
1:B:464:CYS:HB2	1:B:471:PHE:HB2	1.95	0.47	
1:B:456:LYS:HB2	1:B:508:ALA:HB2	1.98	0.45	
1:B:487:ARG:HG3	1:B:594:GLY:HA3	1.98	0.45	
1:B:413:PHE:HE2	1:B:539:ASP:HB3	1.81	0.45	
1:B:431:MET:HG2	1:B:471:PHE:CE1	2.52	0.44	
1:B:565:PRO:HG3	1:B:578:SER:HA	1.98	0.44	
1:B:618:ARG:HG3	1:B:627:TYR:HB2	1.99	0.44	
2:D:286:GLN:H	2:D:289:GLN:HE21	1.65	0.44	
2:D:269:HIS:NE2	2:D:288:GLU:OE1	2.51	0.43	
1:B:625:LYS:HE2	1:B:654:VAL:HG22	2.00	0.43	
1:B:562:ARG:HD3	1:B:597:PRO:O	2.20	0.41	
1:B:609:HIS:HB3	1:B:614:LEU:HD22	2.01	0.41	
2:D:301:ASN:HB2	5:D:501:HOH:O	2.21	0.41	
1:B:519:HIS:CD2	1:B:537:VAL:HG11	2.56	0.40	
1:B:519:HIS:HD2	1:B:537:VAL:HG11	1.85	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	Perce	ntiles
1	В	252/277 (91%)	232 (92%)	20 (8%)	0	100	100
2	D	89/99 (90%)	82 (92%)	7 (8%)	0	100	100
All	All	341/376 (91%)	314 (92%)	27 (8%)	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	В	193/247 (78%)	187 (97%)	6 (3%)	40 49
2	D	72/84 (86%)	68 (94%)	4 (6%)	21 24
All	All	$265/331 \ (80\%)$	255 (96%)	10 (4%)	33 41

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	506	CYS
1	В	510	GLU
1	В	538	SER
1	В	539	ASP
1	В	543	SER
1	В	600	ARG
2	D	289	GLN
2	D	301	ASN
2	D	314	ARG
2	D	349	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	В	412	GLN
1	В	424	GLN
1	В	526	ASN
1	В	609	HIS
1	В	650	ASN
2	D	289	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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	Bo	ond leng	ths	Bot	nd angle	es
WIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	TOO	В	701	1	88,90,90	0.54	2 (2%)	108,125,125	1.16	6 (5%)

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
3	TOO	В	701	1	-	22/59/113/113	0/10/10/10

All (2) bond length outliers are listed below:

\mathbf{Mol}	Chain	Res	Type	Atoms	${f Z}$	$\operatorname{Observed}(\text{\AA})$	Ideal(A)
3	В	701	TOO	C66-N67	-2.76	1.33	1.35
3	В	701	TOO	C9-C7	2.42	1.58	1.53



All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^{o})$	$\operatorname{Ideal}(^{o})$
3	В	701	TOO	C21-C19-N38	7.94	127.12	112.61
3	В	701	TOO	C1-C2-C4	4.05	117.84	110.14
3	В	701	TOO	C1-C2-N86	2.78	114.65	108.93
3	В	701	TOO	C22-C21-C19	2.64	107.77	104.05
3	В	701	TOO	C57-N56-C54	2.24	130.13	122.81
3	В	701	TOO	C66-N67-N63	2.11	106.85	105.17

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	В	701	TOO	N6-C7-C9-C10
3	В	701	TOO	N6-C7-C9-C14
3	В	701	TOO	C15-C7-C9-C10
3	В	701	TOO	C15-C7-C9-C14
3	В	701	TOO	C21-C19-N38-C39
3	В	701	TOO	O40-C39-N38-C19
3	В	701	TOO	C41-C39-N38-C19
3	В	701	TOO	C51-C52-C53-C54
3	В	701	TOO	N67-C66-C68-C69
3	В	701	TOO	N67-C66-C68-C73
3	В	701	TOO	C1-C2-C4-O5
3	В	701	TOO	C53-C54-N56-C62
3	В	701	TOO	C1-C2-C4-N6
3	В	701	TOO	C43-C44-N45-C46
3	В	701	TOO	N86-C2-C4-O5
3	В	701	TOO	C43-C44-N45-C50
3	В	701	TOO	O55-C54-N56-C57
3	В	701	TOO	O55-C54-N56-C62
3	В	701	TOO	C53-C54-N56-C57
3	В	701	TOO	C52-C53-C54-O55
3	В	701	TOO	C39-C41-O42-C43
3	В	701	TOO	C52-C53-C54-N56

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is



within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

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^{th} 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 $>$	$\#\mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	В	258/277 (93%)	0.68	16 (6%) 20 28	46, 74, 114, 124	0
2	D	91/99 (91%)	0.76	6 (6%) 18 26	47, 63, 87, 97	0
All	All	349/376~(92%)	0.70	22 (6%) 20 28	46, 70, 109, 124	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	448	VAL	4.1
1	В	442	PHE	4.1
2	D	348	ILE	3.4
1	В	569	LEU	3.4
2	D	344	PHE	3.4
2	D	345	VAL	3.3
1	В	395	TRP	3.0
1	В	651	ILE	2.9
1	В	463	VAL	2.8
1	В	425	TYR	2.7
1	В	472	ILE	2.6
1	В	392	TYR	2.6
1	В	499	LEU	2.6
1	В	508	ALA	2.5
2	D	322	PRO	2.4
1	В	522	LEU	2.4
1	В	586	LEU	2.4
1	В	393	GLY	2.2
1	В	405	LEU	2.2
1	В	457	LEU	2.1
2	D	266	MET	2.1
2	D	346	ASP	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 monosaccharides in this entry.

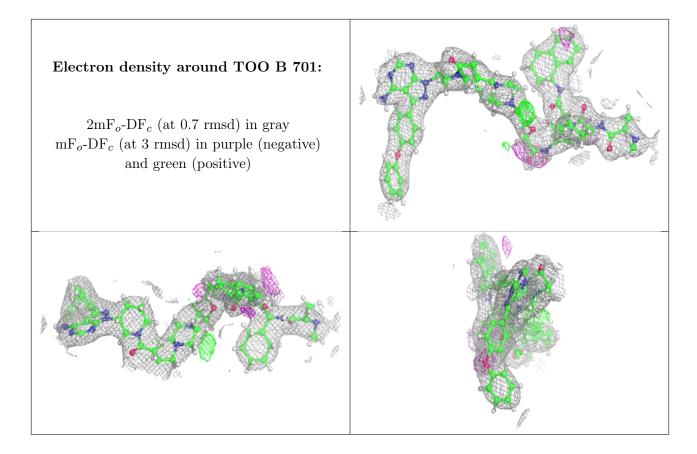
6.4 Ligands (i)

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. The B-factors column lists the minimum, median, 95^{th} 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	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
3	TOO	В	701	81/81	0.96	0.17	44,55,61,61	80
4	ZN	D	401	1/1	0.99	0.23	66,66,66,66	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





6.5 Other polymers (i)

There are no such residues in this entry.

