



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2017 – 05:45 AM EDT

PDB ID : 3BUO  
Title : Crystal structure of c-Cbl-TKB domain complexed with its binding motif in EGF receptor'  
Authors : Ng, C.; Jackson, R.A.; Buschdorf, J.P.; Sun, Q.; Guy, G.R.; Sivaraman, J.  
Deposited on : unknown  
Resolution : 2.60 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
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) : rb-20030345

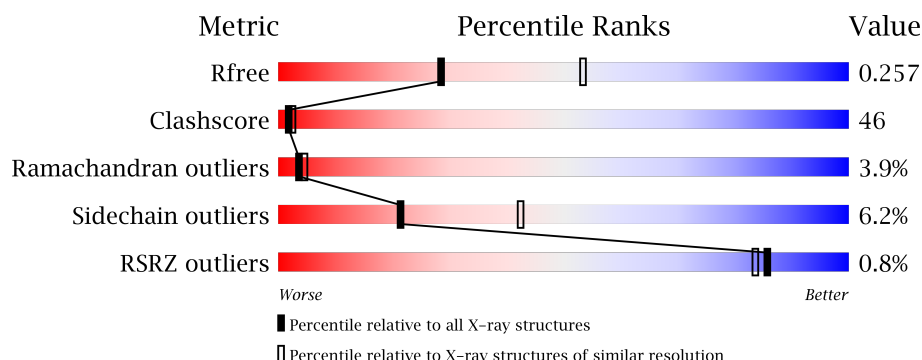
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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	13	
1	C	13	
2	B	329	
2	D	329	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5405 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 13-meric peptide from Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	11	Total	C	N	O	P	0	0	0
			95	57	15	22	1			
1	C	11	Total	C	N	O	P	0	0	0
			95	57	15	22	1			

- Molecule 2 is a protein called E3 ubiquitin-protein ligase CBL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	304	Total	C	N	O	S	0	0	0
			2490	1612	424	441	13			
2	D	304	Total	C	N	O	S	0	0	0
			2490	1612	424	441	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	24	SER	GLY	CLONING ARTIFACT	UNP P22681
D	24	SER	GLY	CLONING ARTIFACT	UNP P22681

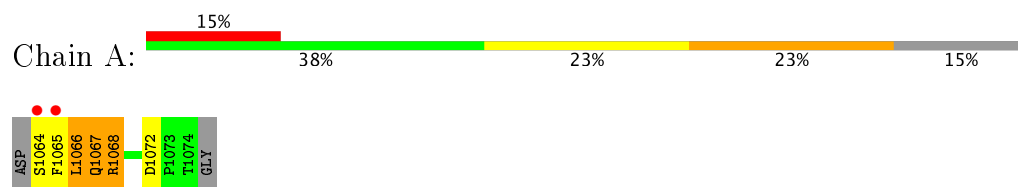
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	6	Total	O	0	0
			6	6		
3	B	111	Total	O	0	0
			111	111		
3	C	6	Total	O	0	0
			6	6		
3	D	112	Total	O	0	0
			112	112		

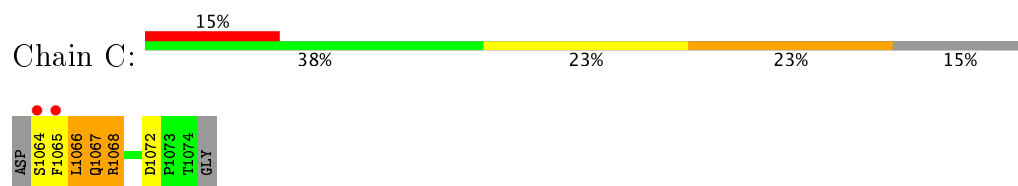
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

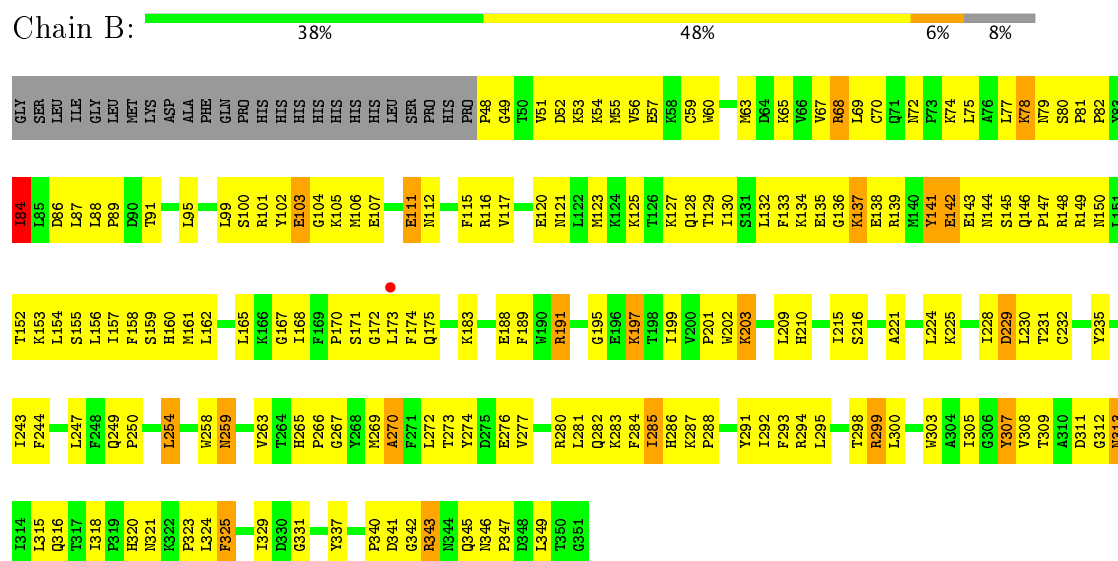
- Molecule 1: 13-meric peptide from Epidermal growth factor receptor



- Molecule 1: 13-meric peptide from Epidermal growth factor receptor



- Molecule 2: E3 ubiquitin-protein ligase CBL



- Molecule 2: E3 ubiquitin-protein ligase CBL



D811	T152	Y83	GLY
G312	K153	L84	SER
N313	L154	D85	LEU
I314	S155	D86	ILE
L315	L156	L87	GLY
Q316	F157	L88	LEU
T317	F158	P89	MET
I318	S159	D90	LYS
P319	H160	T91	ASP
H320	M161	H94	ALA
N321	L162	L95	PHE
K322	L165	L99	GLN
P323	K166	S100	PRO
L324	G167	H101	HIS
F325	I168	R101	HIS
I329	F169	Y102	HIS
D330	P170	E103	HIS
G331	S171	G104	HIS
Y337	G172	K105	HIS
L338	L173	M106	LEU
F339	F174	E107	SER
P340	Q175	E111	PRO
D341	T182	N112	HIS
G342	K183	F115	PRO
R343	E188	R116	P48
N344	F189	V117	G49
Q345	H190	E120	T50
N346	R191	N121	V51
P347	G195	L122	D52
D348	E196	M123	K53
L349	K197	K127	K54
G351	T198	Q128	M55
	F284	T129	V56
	T285	I130	E57
	H286	S131	K58
	K287	L132	C59
	P288	F133	W60
	Y291	K134	K61
	T292	E135	M63
	F293	G136	D64
	R294	K137	K65
	L295	R139	W66
	T298	M140	V67
	R299	Y141	R68
	L300	E142	L69
	W303	E143	C70
	A304	N144	Q71
	I305	S145	N72
	G306	Q146	F73
	Y307	P147	K74
	V308	R148	L75
	T309	R149	A76
	A310	N150	L77
		L151	K78
			N79
			S80
			P81
			P82

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.86Å 110.17Å 55.82Å 90.00° 89.94° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60 39.29 – 2.60	Depositor EDS
% Data completeness (in resolution range)	51.8 (20.00-2.60) 91.3 (39.29-2.60)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 2.61Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.231 , 0.278 0.248 , 0.257	Depositor DCC
$R_{free}$ test set	2854 reflections (15.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.0	Xtriage
Anisotropy	0.858	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 28.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.458 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	5405	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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: PTR

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	0.68	0/79	0.96	0/104
1	C	0.65	0/79	0.95	0/104
2	B	0.52	0/2556	0.65	0/3449
2	D	0.51	0/2556	0.65	0/3449
All	All	0.52	0/5270	0.66	0/7106

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	A	95	0	80	16	0
1	C	95	0	80	19	0
2	B	2490	0	2499	232	0
2	D	2490	0	2499	232	0
3	A	6	0	0	0	0
3	B	111	0	0	16	0
3	C	6	0	0	1	0
3	D	112	0	0	18	0
All	All	5405	0	5158	478	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 46.

The worst 5 of 478 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:D:82:PRO:HG3	2:D:156:LEU:HD12	1.33	1.09
2:D:277:VAL:HG13	2:D:292:ILE:HD11	1.37	1.05
2:B:82:PRO:HG3	2:B:156:LEU:HD12	1.33	1.04
2:B:282:GLN:HE22	2:B:285:ILE:HD12	1.23	1.00
2:D:282:GLN:HE22	2:D:285:ILE:HD12	1.22	1.00

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	8/13 (62%)	4 (50%)	2 (25%)	2 (25%)	0	0
1	C	8/13 (62%)	4 (50%)	2 (25%)	2 (25%)	0	0
2	B	302/329 (92%)	248 (82%)	44 (15%)	10 (3%)	4	7
2	D	302/329 (92%)	248 (82%)	44 (15%)	10 (3%)	4	7
All	All	620/684 (91%)	504 (81%)	92 (15%)	24 (4%)	3	4

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	103	GLU
2	B	270	ALA
2	B	285	ILE
2	D	103	GLU
2	D	270	ALA



### 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	A	10/11 (91%)	9 (90%)	1 (10%)	9	17
1	C	10/11 (91%)	9 (90%)	1 (10%)	9	17
2	B	271/293 (92%)	254 (94%)	17 (6%)	21	42
2	D	271/293 (92%)	255 (94%)	16 (6%)	23	45
All	All	562/608 (92%)	527 (94%)	35 (6%)	21	42

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	313	ASN
2	D	53	LYS
2	D	313	ASN
2	B	325	PHE
2	B	343	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	346	ASN
2	D	94	HIS
2	D	345	GLN
2	D	79	ASN
2	D	160	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 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	PTR	A	1069	1	15,16,17	1.25	1 (6%)	19,22,24	0.99	0
1	PTR	C	1069	1	15,16,17	1.24	0	19,22,24	0.98	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	PTR	A	1069	1	-	0/9/11/13	0/1/1/1
1	PTR	C	1069	1	-	0/9/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1069	PTR	CE1-CD1	2.10	1.42	1.38

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	10/13 (76%)	0.12	2 (20%) <b>1</b> <b>0</b>	22, 32, 48, 51	0
1	C	10/13 (76%)	0.11	2 (20%) <b>1</b> <b>0</b>	22, 32, 48, 51	0
2	B	304/329 (92%)	-0.43	1 (0%) <b>93</b> <b>93</b>	16, 32, 52, 71	0
2	D	304/329 (92%)	-0.43	0 <b>100</b> <b>100</b>	16, 32, 52, 71	0
All	All	628/684 (91%)	-0.41	5 (0%) <b>86</b> <b>83</b>	16, 32, 52, 71	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1065	PHE	2.8
1	C	1064	SER	2.7
1	A	1064	SER	2.4
2	B	173	LEU	2.1
1	C	1065	PHE	2.0

### 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, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	PTR	C	1069	16/17	0.96	0.14	-	27,28,29,30	0
1	PTR	A	1069	16/17	0.97	0.13	-	27,28,29,30	0

### 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.