



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 18, 2016 – 07:54 PM EDT

PDB ID : 5HXB  
Title : Cereblon in complex with DDB1, CC-885, and GSPT1  
Authors : Chamberlain, P.P.; Matyskiela, M.; Pagarigan, B.  
Deposited on : 2016-01-30  
Resolution : 3.60 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:

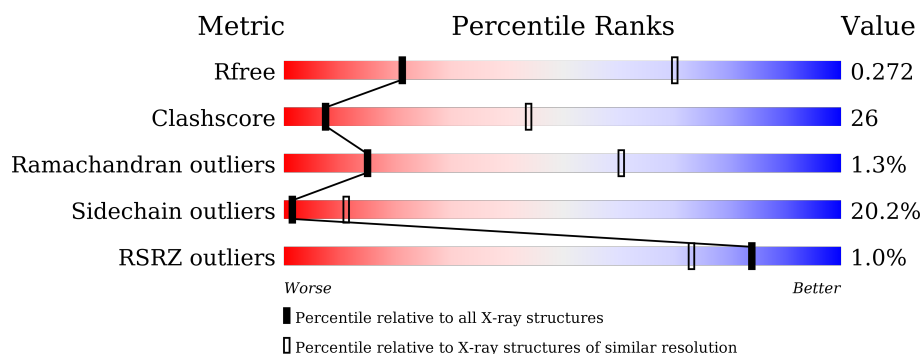
MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027790  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027790



**i**

## X-RAY DIFFRACTION

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.



<b>Metric</b>	<b>Whole archive (#Entries)</b>	<b>Similar resolution (#Entries, resolution range(Å))</b>
R <sub>free</sub>	91344	1408 (3.80-3.40)
Clashscore	102246	1010 (3.74-3.46)
Ramachandran outliers	100387	1007 (3.76-3.44)
Sidechain outliers	100360	1007 (3.76-3.44)
RSRZ outliers	91569	1003 (3.78-3.42)

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	199	<div> <div></div> <div>2%</div> <div>36%</div> <div>42%</div> <div>18%</div> <div>• •</div> </div>	
1	X	199	<div> <div></div> <div>2%</div> <div>31%</div> <div>42%</div> <div>20%</div> <div>6%</div> <div>•</div> </div>	
2	B	1140	<div> <div></div> <div>•</div> <div>52%</div> <div>35%</div> <div>6%</div> <div>6%</div> </div>	
2	Y	1140	<div> <div></div> <div>•</div> <div>53%</div> <div>35%</div> <div>7%</div> <div>5%</div> </div>	
3	C	406	<div> <div></div> <div>•</div> <div>52%</div> <div>32%</div> <div>6%</div> <div>•</div> <div>9%</div> </div>	
3	Z	406	<div> <div></div> <div>52%</div> <div>34%</div> <div>7%</div> <div>•</div> <div>6%</div> </div>	



The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	ZN	C	501	-	-	X	-
5	85C	Z	502	-	-	X	-



## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 25089 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 Eukaryotic peptide chain release factor GTP-binding subunit ERF3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	195	Total	C	N	O	S	0	0	0
			1481	941	254	275	11			
1	A	196	Total	C	N	O	S	0	0	0
			1441	919	245	267	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	436	GLY	-	expression tag	UNP P15170
X	437	SER	-	expression tag	UNP P15170
A	436	GLY	-	expression tag	UNP P15170
A	437	SER	-	expression tag	UNP P15170

- Molecule 2 is a protein called DNA damage-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Y	1080	Total	C	N	O	S	0	0	0
			8178	5215	1369	1550	44			
2	B	1075	Total	C	N	O	S	0	0	0
			8096	5164	1351	1536	45			

- Molecule 3 is a protein called Protein cereblon.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	Z	380	Total	C	N	O	S	0	0	0
			2960	1891	502	544	23			
3	C	370	Total	C	N	O	S	0	0	0
			2869	1835	485	526	23			

There are 6 discrepancies between the modelled and reference sequences:

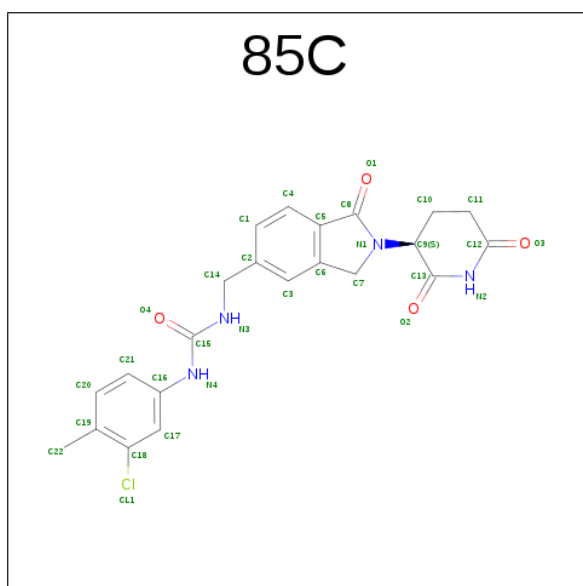


Chain	Residue	Modelled	Actual	Comment	Reference
Z	37	GLY	-	expression tag	UNP Q96SW2
Z	38	SER	-	expression tag	UNP Q96SW2
Z	39	MET	-	expression tag	UNP Q96SW2
C	37	GLY	-	expression tag	UNP Q96SW2
C	38	SER	-	expression tag	UNP Q96SW2
C	39	MET	-	expression tag	UNP Q96SW2

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	Z	1	Total Zn 1 1	0	0
4	C	1	Total Zn 1 1	0	0

- Molecule 5 is 1-(3-chloro-4-methylphenyl)-3-({2-[(3S)-2,6-dioxopiperidin-3-yl]-1-oxo-2,3-dihydro-1H-isindol-5-yl}methyl)urea (three-letter code: 85C) (formula: C<sub>22</sub>H<sub>21</sub>ClN<sub>4</sub>O<sub>4</sub>).



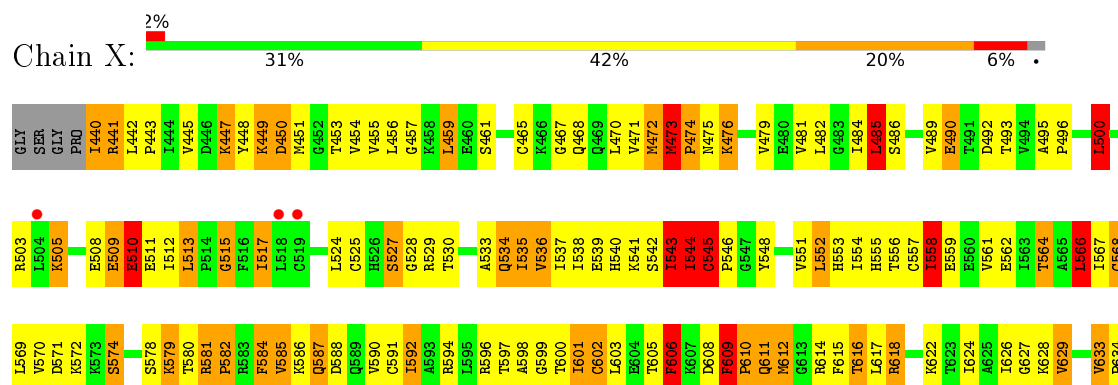
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	Z	1	Total C Cl N O 31 22 1 4 4	0	0
5	C	1	Total C Cl N O 31 22 1 4 4	0	0



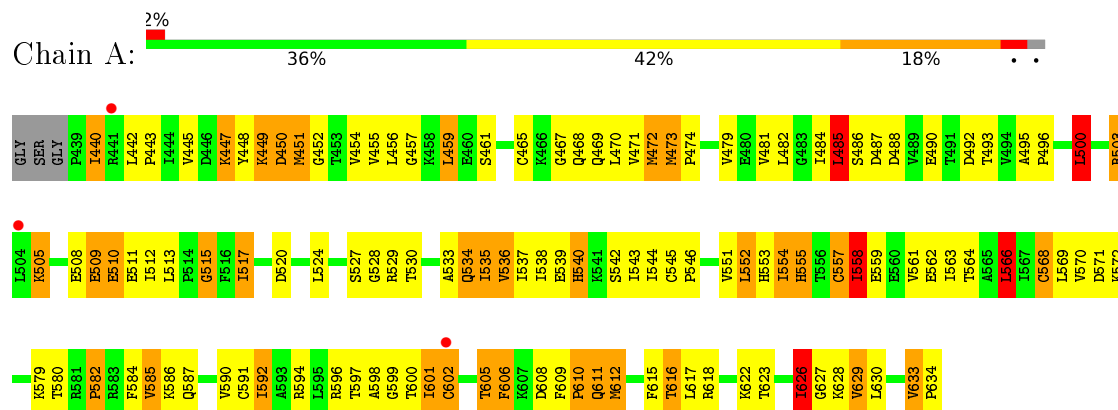
### 3 Residue-property plots

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.

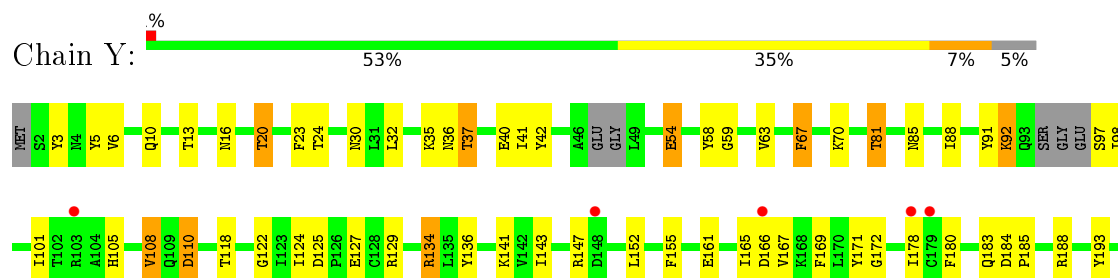
- Molecule 1: Eukaryotic peptide chain release factor GTP-binding subunit ERF3A



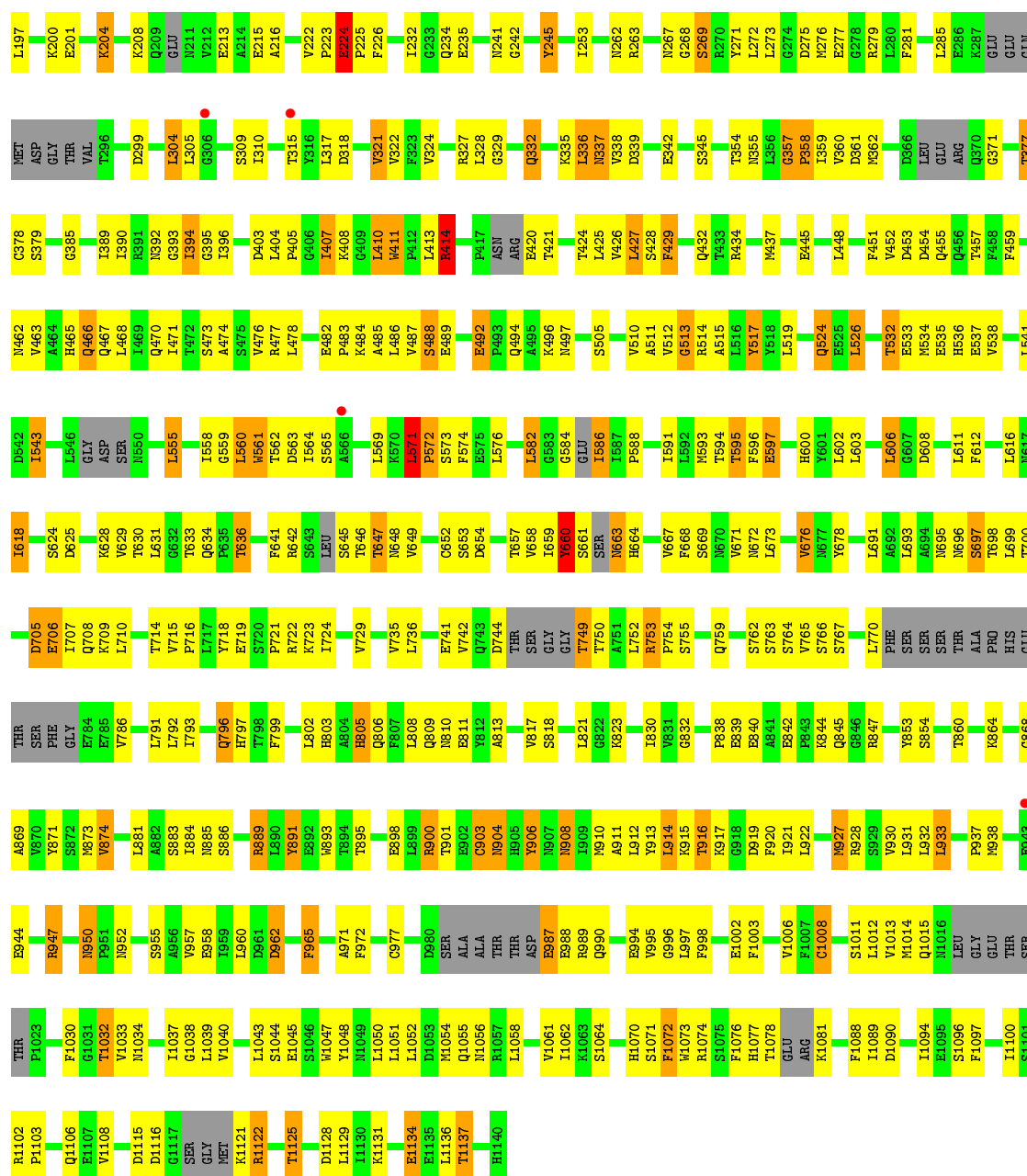
- Molecule 1: Eukaryotic peptide chain release factor GTP-binding subunit ERF3A



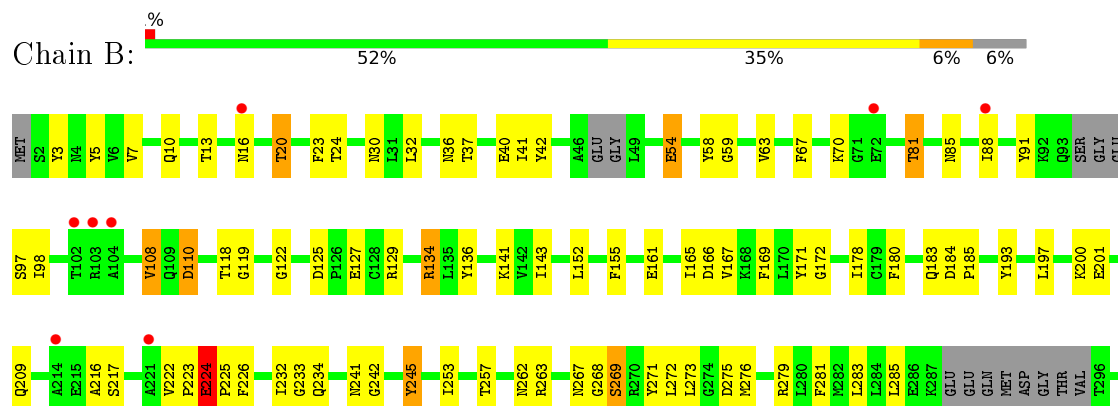
- Molecule 2: DNA damage-binding protein 1



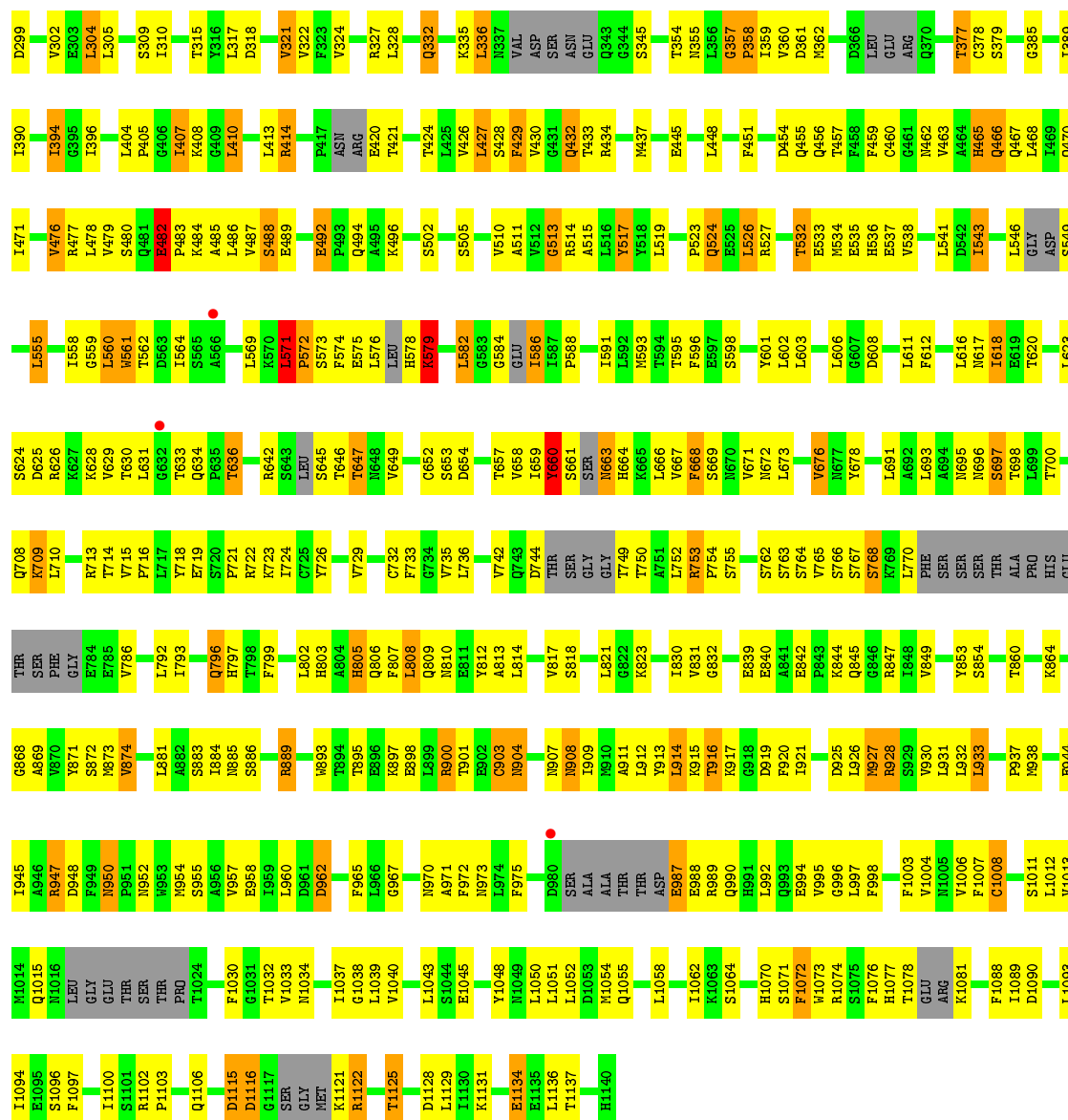




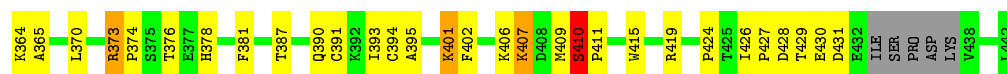
### • Molecule 2: DNA damage-binding protein 1



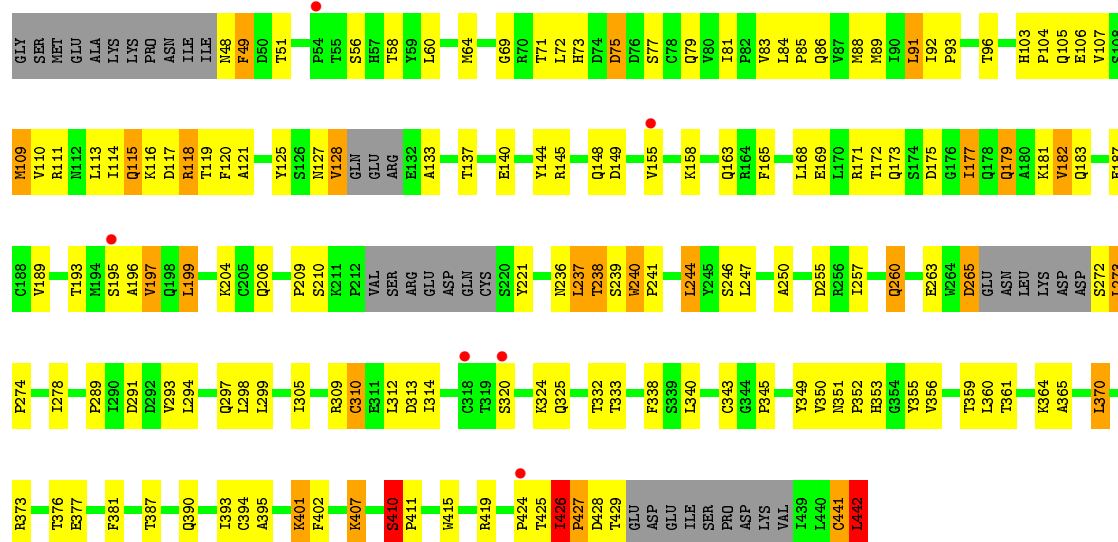








• Molecule 3: Protein cereblon





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	156.75Å 111.52Å 175.06Å 90.00° 95.81° 90.00°	Depositor
Resolution (Å)	50.00 – 3.60 49.80 – 3.60	Depositor EDS
% Data completeness (in resolution range)	94.4 (50.00-3.60) 94.5 (49.80-3.60)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.32 (at 3.57Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, $R_{free}$	0.224 , 0.273 0.225 , 0.272	Depositor DCC
$R_{free}$ test set	3289 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	94.6	Xtriage
Anisotropy	0.467	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 85.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	25089	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.76% 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: ZN, 85C

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	0.85	0/1463	0.89	1/1987 (0.1%)
1	X	0.93	0/1502	1.02	10/2032 (0.5%)
2	B	0.85	4/8235 (0.0%)	0.87	6/11187 (0.1%)
2	Y	0.86	6/8319 (0.1%)	0.88	11/11295 (0.1%)
3	C	0.81	0/2937	0.90	5/4003 (0.1%)
3	Z	0.84	0/3030	0.90	5/4125 (0.1%)
All	All	0.85	10/25486 (0.0%)	0.89	38/34629 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	13
1	X	0	16
2	B	0	17
2	Y	0	19
3	C	0	5
3	Z	0	5
All	All	0	75

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	Y	489	GLU	CG-CD	7.70	1.63	1.51
2	B	489	GLU	CG-CD	7.30	1.62	1.51
2	Y	853	TYR	CB-CG	-6.86	1.41	1.51
2	B	1134	GLU	CG-CD	6.54	1.61	1.51
2	Y	1134	GLU	CG-CD	6.18	1.61	1.51



The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	426	ILE	C-N-CD	-7.50	104.10	120.60
3	C	442	LEU	CA-CB-CG	7.35	132.21	115.30
2	B	357	GLY	C-N-CD	-7.31	104.52	120.60
2	Y	357	GLY	C-N-CD	-7.28	104.59	120.60
2	Y	597	GLU	N-CA-C	-7.00	92.09	111.00

There are no chirality outliers.

5 of 75 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	X	472	MET	Peptide
1	X	473	MET	Peptide
1	X	500	LEU	Peptide
1	X	509	GLU	Peptide
1	X	515	GLY	Peptide

## 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	1441	0	1439	151	2
1	X	1481	0	1519	209	0
2	B	8096	0	7807	344	0
2	Y	8178	0	7930	342	0
3	C	2869	0	2758	126	0
3	Z	2960	0	2851	138	2
4	C	1	0	0	2	0
4	Z	1	0	0	0	0
5	C	31	0	0	4	0
5	Z	31	0	0	10	0
All	All	25089	0	24304	1265	2

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 26.

The worst 5 of 1265 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)
1:X:606:PHE:CD2	1:X:628:LYS:HG2	1.44	1.49
1:A:606:PHE:CE2	1:A:628:LYS:HG2	1.50	1.47
1:A:606:PHE:CD2	1:A:628:LYS:HG2	1.48	1.45
1:X:606:PHE:CE2	1:X:628:LYS:HG2	1.64	1.30
3:Z:352:PRO:HG3	3:Z:378:HIS:ND1	1.44	1.29

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:Z:103:HIS:CD2	1:A:542:SER:OG[2_546]	2.05	0.15
3:Z:116:LYS:CB	1:A:448:TYR:CB[1_544]	2.18	0.02

## 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	194/199 (98%)	155 (80%)	30 (16%)	9 (5%)	3	31
1	X	193/199 (97%)	152 (79%)	33 (17%)	8 (4%)	3	34
2	B	1039/1140 (91%)	955 (92%)	76 (7%)	8 (1%)	24	69
2	Y	1046/1140 (92%)	959 (92%)	78 (8%)	9 (1%)	21	67
3	C	360/406 (89%)	322 (89%)	32 (9%)	6 (2%)	11	55
3	Z	372/406 (92%)	332 (89%)	37 (10%)	3 (1%)	24	69
All	All	3204/3490 (92%)	2875 (90%)	286 (9%)	43 (1%)	15	60

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	474	PRO
1	X	610	PRO
2	Y	224	GLU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	Y	358	PRO
2	Y	483	PRO

### 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	153/175 (87%)	113 (74%)	40 (26%)	0	5
1	X	165/175 (94%)	117 (71%)	48 (29%)	0	4
2	B	855/999 (86%)	699 (82%)	156 (18%)	2	14
2	Y	868/999 (87%)	709 (82%)	159 (18%)	2	14
3	C	306/368 (83%)	240 (78%)	66 (22%)	1	9
3	Z	316/368 (86%)	247 (78%)	69 (22%)	1	9
All	All	2663/3084 (86%)	2125 (80%)	538 (20%)	1	11

5 of 538 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Z	246	SER
2	B	310	ILE
3	C	442	LEU
3	Z	294	LEU
3	Z	428	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
3	Z	198	GLN
2	B	267	ASN
3	C	127	ASN
3	Z	260	GLN
2	B	36	ASN



### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
5	85C	C	502	-	34,34,34	1.53	4 (11%)	46,49,49	2.53	13 (28%)
5	85C	Z	502	-	34,34,34	1.51	4 (11%)	46,49,49	2.51	11 (23%)

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
5	85C	C	502	-	-	0/13/38/38	0/4/4/4
5	85C	Z	502	-	-	0/13/38/38	0/4/4/4

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	502	85C	C5-C8	-5.75	1.39	1.48

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	Z	502	85C	C5-C8	-5.32	1.40	1.48
5	Z	502	85C	C16-N4	-4.58	1.33	1.41
5	C	502	85C	C16-N4	-4.26	1.33	1.41
5	C	502	85C	C8-N1	-2.99	1.32	1.36

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	502	85C	C7-N1-C8	-9.77	109.76	113.03
5	Z	502	85C	C7-N1-C8	-8.76	110.09	113.03
5	C	502	85C	C7-C6-C5	-4.16	107.03	109.71
5	Z	502	85C	C7-C6-C5	-3.38	107.53	109.71
5	Z	502	85C	C16-N4-C15	-3.33	120.07	126.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	502	85C	4	0
5	Z	502	85C	10	0

## 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	196/199 (98%)	-0.08	3 (1%) 76 64	56, 112, 160, 183	0
1	X	195/199 (97%)	-0.09	3 (1%) 76 64	66, 109, 160, 185	0
2	B	1075/1140 (94%)	-0.15	11 (1%) 84 73	60, 114, 185, 240	5 (0%)
2	Y	1080/1140 (94%)	-0.18	9 (0%) 87 78	54, 113, 174, 221	5 (0%)
3	C	370/406 (91%)	-0.22	6 (1%) 74 61	63, 111, 165, 225	0
3	Z	380/406 (93%)	-0.25	1 (0%) 94 90	64, 109, 169, 213	0
All	All	3296/3490 (94%)	-0.17	33 (1%) 84 73	54, 112, 174, 240	10 (0%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	566	ALA	5.2
2	B	103	ARG	3.9
2	Y	179	CYS	3.8
2	B	16	ASN	3.3
2	B	88	ILE	3.3

### 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](#)

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( $\text{\AA}^2$ )	Q<0.9
5	85C	C	502	31/31	0.93	0.30	0.56	61,90,177,227	0
5	85C	Z	502	31/31	0.94	0.29	0.56	70,95,210,235	0
4	ZN	C	501	1/1	0.98	0.14	-0.23	187,187,187,187	0
4	ZN	Z	501	1/1	0.98	0.14	-0.62	191,191,191,191	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.