



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 12:34 AM GMT

PDB ID : 2AZE  
Title : Structure of the Rb C-terminal domain bound to an E2F1-DP1 heterodimer  
Authors : Rubin, S.M.; Gall, A.L.; Zheng, N.; Pavletich, N.P.  
Deposited on : 2005-09-10  
Resolution : 2.55 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

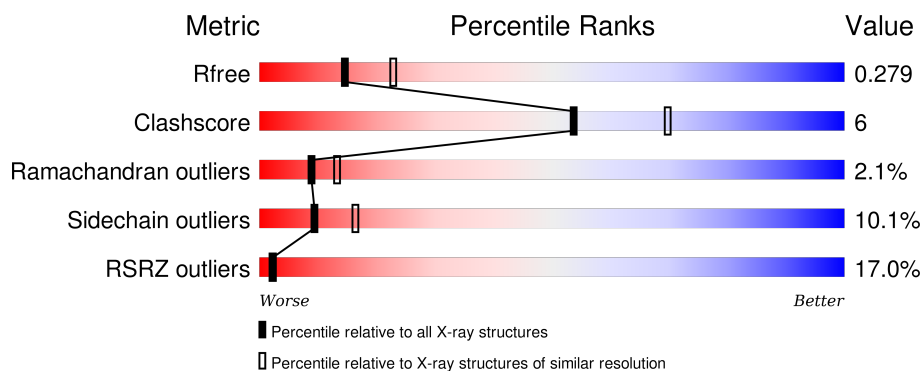
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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	4549 (2.58-2.50)
Clashscore	102246	5292 (2.58-2.50)
Ramachandran outliers	100387	5194 (2.58-2.50)
Sidechain outliers	100360	5196 (2.58-2.50)
RSRZ outliers	91569	4561 (2.58-2.50)

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	155	<div> <div>13%</div> <div>78%</div> <div>14%</div> <div>5%</div> <div>.</div> </div>
2	B	106	<div> <div>15%</div> <div>68%</div> <div>25%</div> <div>5%</div> <div>.</div> </div>
3	C	46	<div> <div>30%</div> <div>80%</div> <div>15%</div> <div>.</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2456 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 Transcription factor Dp-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	0	0
			1191	746	212	225	8			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	196	GLY	-	CLONING ARTIFACT	UNP Q14186
A	197	GLU	-	CLONING ARTIFACT	UNP Q14186
A	198	PHE	-	CLONING ARTIFACT	UNP Q14186

- Molecule 2 is a protein called Transcription factor E2F1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	101	Total	C	N	O	S	0	0	0
			796	488	136	166	6			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	196	GLY	-	CLONING ARTIFACT	UNP Q01094
B	197	SER	-	CLONING ARTIFACT	UNP Q01094
B	198	HIS	-	CLONING ARTIFACT	UNP Q01094
B	199	MET	-	CLONING ARTIFACT	UNP Q01094

- Molecule 3 is a protein called Retinoblastoma-associated protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	44	Total	C	N	O	S	0	0	0
			338	209	62	65	2			

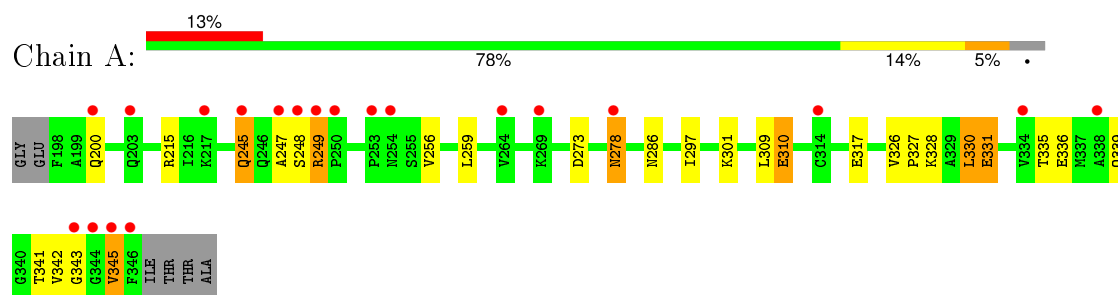
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	70	Total 70	O 70	0	0
4	B	48	Total 48	O 48	0	0
4	C	13	Total 13	O 13	0	0

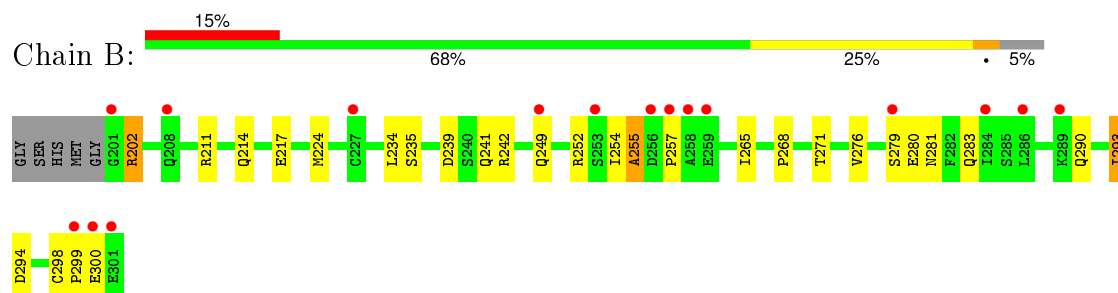
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

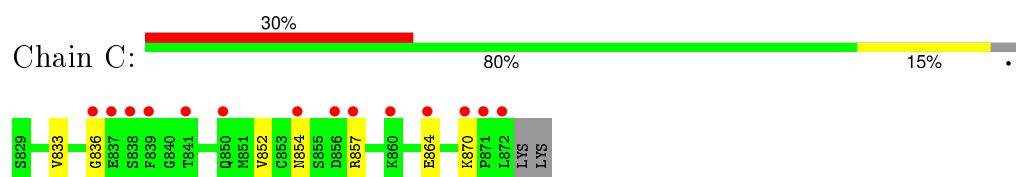
- Molecule 1: Transcription factor Dp-1



- Molecule 2: Transcription factor E2F1



- Molecule 3: Retinoblastoma-associated protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.80 Å   168.60 Å   48.30 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	15.00 – 2.55 19.73 – 2.56	Depositor EDS
% Data completeness (in resolution range)	93.6 (15.00-2.55) 93.4 (19.73-2.56)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.02 (at 2.56 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.221   ,   0.262 0.240   ,   0.279	Depositor DCC
$R_{free}$ test set	991 reflections (5.63%)	DCC
Wilson B-factor (Å <sup>2</sup> )	64.9	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 71.7	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 19422 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2456	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.24% 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	0.52	0/1210	0.62	0/1629
2	B	0.45	0/804	0.59	0/1086
3	C	0.41	0/342	0.50	0/457
All	All	0.48	0/2356	0.60	0/3172

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	1191	0	1194	15	0
2	B	796	0	786	15	0
3	C	338	0	348	1	0
4	A	70	0	0	0	0
4	B	48	0	0	1	0
4	C	13	0	0	0	0
All	All	2456	0	2328	30	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 (30) 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:268:PRO:O	2:B:271:THR:HG23	1.98	0.62
2:B:239:ASP:HB2	4:B:130:HOH:O	2.00	0.61
2:B:202:ARG:HD3	2:B:202:ARG:H	1.65	0.61
1:A:245:GLN:HA	1:A:245:GLN:HE21	1.67	0.59
2:B:265:ILE:HG23	2:B:293:ILE:CD1	2.35	0.56
1:A:248:SER:O	1:A:249:ARG:HB2	2.09	0.53
1:A:327:PRO:HG2	1:A:330:LEU:HD22	1.92	0.52
2:B:249:GLN:HE21	2:B:252:ARG:HH21	1.58	0.51
2:B:276:VAL:HB	2:B:283:GLN:HG2	1.93	0.50
2:B:276:VAL:HB	2:B:283:GLN:CG	2.43	0.49
2:B:235:SER:HA	2:B:241:GLN:NE2	2.28	0.48
2:B:293:ILE:HD12	2:B:294:ASP:N	2.29	0.47
1:A:297:ILE:HG22	1:A:301:LYS:HE3	1.96	0.47
1:A:336:GLU:HG3	1:A:341:THR:O	2.16	0.46
1:A:273:ASP:HB2	1:A:286:ASN:HB2	1.97	0.46
1:A:215:ARG:HD2	2:B:217:GLU:OE2	2.16	0.46
1:A:278:ASN:OD1	1:A:278:ASN:N	2.45	0.45
1:A:245:GLN:CA	1:A:245:GLN:HE21	2.27	0.44
1:A:335:THR:O	1:A:339:GLN:HG3	2.16	0.44
2:B:265:ILE:HG23	2:B:293:ILE:HD13	2.00	0.44
2:B:298:CYS:HA	2:B:299:PRO:HD3	1.88	0.44
1:A:297:ILE:HG13	1:A:297:ILE:H	1.67	0.44
1:A:310:GLU:CD	1:A:310:GLU:H	2.20	0.44
2:B:254:ILE:O	2:B:255:ALA:HB2	2.17	0.43
1:A:343:GLY:C	1:A:345:VAL:H	2.21	0.43
3:C:864:GLU:H	3:C:864:GLU:CD	2.22	0.43
1:A:317:GLU:CD	1:A:317:GLU:H	2.23	0.42
1:A:328:LYS:O	1:A:331:GLU:HB2	2.20	0.41
2:B:265:ILE:HG23	2:B:293:ILE:HD11	2.02	0.41
2:B:252:ARG:HG2	2:B:257:PRO:CD	2.50	0.41

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	147/155 (95%)	136 (92%)	8 (5%)	3 (2%)	9	14
2	B	99/106 (93%)	92 (93%)	5 (5%)	2 (2%)	9	14
3	C	42/46 (91%)	37 (88%)	4 (10%)	1 (2%)	7	11
All	All	288/307 (94%)	265 (92%)	17 (6%)	6 (2%)	9	13

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	255	ALA
1	A	247	ALA
2	B	300	GLU
3	C	836	GLY
1	A	249	ARG
1	A	345	VAL

### 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	134/138 (97%)	123 (92%)	11 (8%)	14	25
2	B	93/96 (97%)	82 (88%)	11 (12%)	6	11
3	C	40/42 (95%)	35 (88%)	5 (12%)	6	10
All	All	267/276 (97%)	240 (90%)	27 (10%)	9	16

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	200	GLN
1	A	245	GLN
1	A	256	VAL
1	A	259	LEU
1	A	278	ASN
1	A	309	LEU
1	A	310	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	326	VAL
1	A	330	LEU
1	A	331	GLU
1	A	342	VAL
2	B	202	ARG
2	B	211	ARG
2	B	214	GLN
2	B	224	MET
2	B	234	LEU
2	B	242	ARG
2	B	279	SER
2	B	280	GLU
2	B	281	ASN
2	B	290	GLN
2	B	293	ILE
3	C	833	VAL
3	C	852	VAL
3	C	854	ASN
3	C	857	ARG
3	C	870	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	245	GLN
2	B	214	GLN
2	B	249	GLN
2	B	281	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

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	149/155 (96%)	0.85	20 (13%)	4 4	43, 74, 113, 137	0
2	B	101/106 (95%)	1.04	16 (15%)	3 3	46, 85, 115, 132	0
3	C	44/46 (95%)	1.50	14 (31%)	1 0	95, 108, 123, 126	0
All	All	294/307 (95%)	1.01	50 (17%)	2 2	43, 82, 120, 137	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	345	VAL	7.7
1	A	344	GLY	7.3
1	A	247	ALA	6.1
2	B	201	GLY	5.2
2	B	299	PRO	5.2
2	B	279	SER	5.0
3	C	836	GLY	4.9
1	A	346	PHE	4.9
2	B	300	GLU	4.9
1	A	249	ARG	4.8
2	B	258	ALA	4.7
2	B	301	GLU	4.7
3	C	872	LEU	4.1
1	A	248	SER	4.0
3	C	838	SER	3.8
3	C	837	GLU	3.7
3	C	856	ASP	3.6
3	C	839	PHE	3.5
1	A	254	ASN	3.3
2	B	286	LEU	3.1
2	B	249	GLN	3.1
2	B	253	SER	3.0
2	B	289	LYS	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	259	GLU	2.8
1	A	250	PRO	2.7
1	A	278	ASN	2.7
3	C	871	PRO	2.7
2	B	284	ILE	2.6
3	C	870	LYS	2.5
1	A	314	CYS	2.5
3	C	857	ARG	2.5
1	A	338	ALA	2.5
2	B	208	GLN	2.5
1	A	264	VAL	2.4
3	C	860	LYS	2.4
2	B	256	ASP	2.4
1	A	334	VAL	2.4
1	A	269	LYS	2.3
2	B	227	CYS	2.3
1	A	200	GLN	2.3
3	C	864	GLU	2.3
2	B	257	PRO	2.2
1	A	203	GLN	2.2
3	C	850	GLN	2.2
1	A	245	GLN	2.1
1	A	217	LYS	2.1
3	C	841	THR	2.1
1	A	253	PRO	2.1
3	C	854	ASN	2.0
1	A	343	GLY	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.