



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 19, 2020 – 07:45 pm BST

PDB ID : 5TUV  
Title : Crystal structure of the E2F5-DP1-p107 ternary complex  
Authors : Liban, T.J.; Tripathi, S.M.; Rubin, S.M.  
Deposited on : 2016-11-07  
Resolution : 2.90 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

<https://www.wwpdb.org/validation/2017/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  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

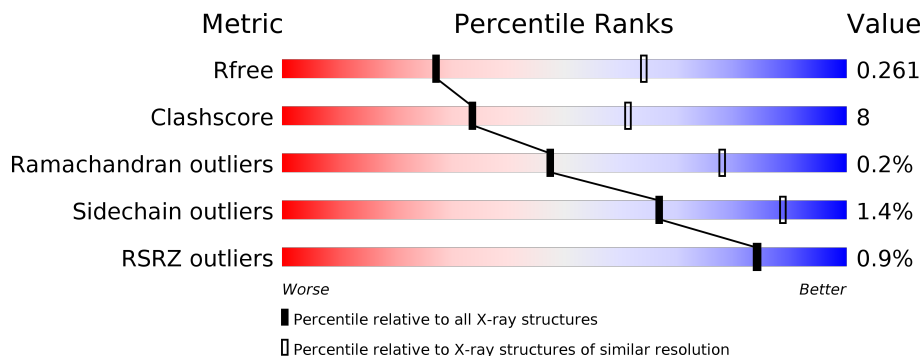
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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  $\leq 5\%$ . 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	
1	D	155	
2	B	112	
2	E	112	
3	C	41	
3	F	41	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3532 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 DP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	107	840	528	154	154	4	0	0	0
1	D	95	752	474	138	137	3	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	196	GLY	-	expression tag	UNP Q14186
A	197	GLU	-	expression tag	UNP Q14186
A	198	PHE	-	expression tag	UNP Q14186
D	196	GLY	-	expression tag	UNP Q14186
D	197	GLU	-	expression tag	UNP Q14186
D	198	PHE	-	expression tag	UNP Q14186

- Molecule 2 is a protein called Transcription factor E2F5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	104	797	504	134	155	4	0	0	0
2	E	103	736	462	129	142	3	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	121	GLY	-	expression tag	UNP Q15329
B	122	HIS	-	expression tag	UNP Q15329
B	123	MET	-	expression tag	UNP Q15329
E	121	GLY	-	expression tag	UNP Q15329
E	122	HIS	-	expression tag	UNP Q15329
E	123	MET	-	expression tag	UNP Q15329

- Molecule 3 is a protein called Retinoblastoma-like protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	29	226	139	42	44	1	0	0	0
3	F	24	181	115	32	33	1	0	0	0

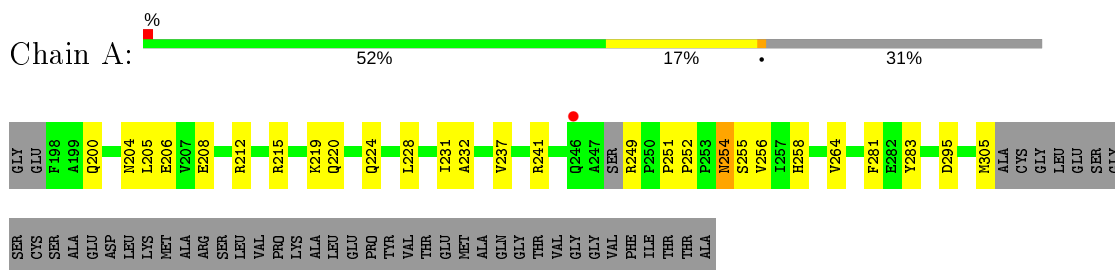
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	991	GLY	-	expression tag	UNP P28749
C	992	GLU	-	expression tag	UNP P28749
C	993	PHE	-	expression tag	UNP P28749
F	991	GLY	-	expression tag	UNP P28749
F	992	GLU	-	expression tag	UNP P28749
F	993	PHE	-	expression tag	UNP P28749

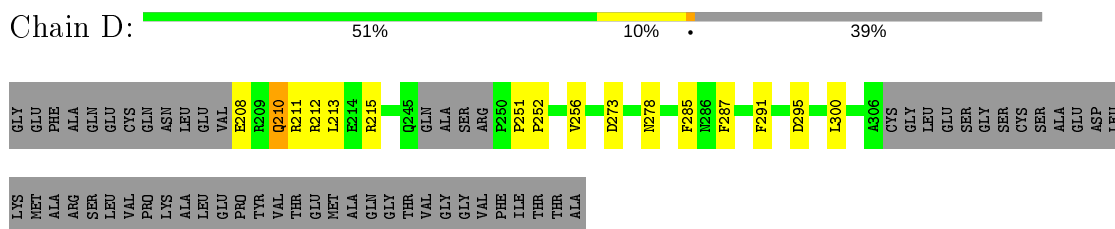
### 3 Residue-property plots [i](#)

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.

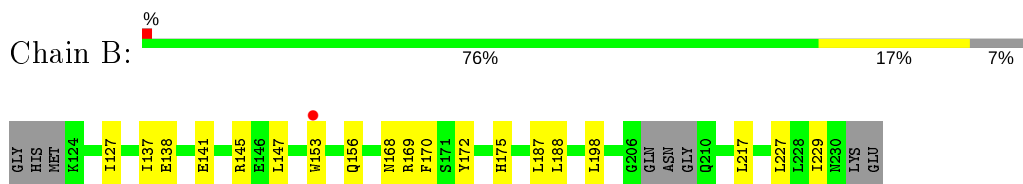
- Molecule 1: Transcription factor DP1



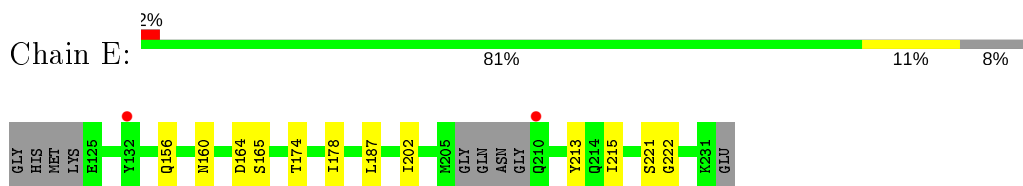
- Molecule 1: Transcription factor DP1



- Molecule 2: Transcription factor E2F5

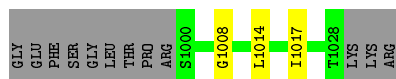


- Molecule 2: Transcription factor E2F5



- Molecule 3: Retinoblastoma-like protein 1





- Molecule 3: Retinoblastoma-like protein 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.98Å 57.34Å 99.20Å 90.00° 96.64° 90.00°	Depositor
Resolution (Å)	49.27 – 2.90 49.27 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.27-2.90) 99.7 (49.27-2.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.04 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.212 , 0.260 0.216 , 0.261	Depositor DCC
$R_{free}$ test set	803 reflections (5.25%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.2	Xtrriage
Anisotropy	0.118	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 77.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3532	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.50% 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](#)

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	0.23	0/852	0.43	0/1153
1	D	0.22	0/765	0.46	0/1034
2	B	0.23	0/809	0.49	0/1099
2	E	0.33	0/748	0.50	0/1024
3	C	0.21	0/228	0.35	0/304
3	F	0.21	0/183	0.36	0/244
All	All	0.25	0/3585	0.46	0/4858

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	840	0	807	22	0
1	D	752	0	718	17	0
2	B	797	0	755	19	0
2	E	736	0	637	13	0
3	C	226	0	224	2	0
3	F	181	0	181	4	0
All	All	3532	0	3322	58	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 8.



The worst 5 of 58 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:D:212:ARG:HG2	1:D:215:ARG:HH22	1.53	0.73
2:E:202:ILE:HD12	2:E:202:ILE:N	2.04	0.72
2:E:156:GLN:O	2:E:160:ASN:ND2	2.24	0.71
1:A:283:TYR:HB2	2:B:217:LEU:HB2	1.72	0.70
1:A:255:SER:HA	2:B:175:HIS:HD2	1.58	0.69

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	103/155 (66%)	99 (96%)	4 (4%)	0	100	100
1	D	91/155 (59%)	86 (94%)	5 (6%)	0	100	100
2	B	100/112 (89%)	97 (97%)	3 (3%)	0	100	100
2	E	99/112 (88%)	95 (96%)	4 (4%)	0	100	100
3	C	27/41 (66%)	26 (96%)	0	1 (4%)	3	13
3	F	22/41 (54%)	19 (86%)	3 (14%)	0	100	100
All	All	442/616 (72%)	422 (96%)	19 (4%)	1 (0%)	47	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	1008	GLY

### 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	89/138 (64%)	86 (97%)	3 (3%)	37	71
1	D	80/138 (58%)	79 (99%)	1 (1%)	69	90
2	B	85/103 (82%)	84 (99%)	1 (1%)	71	91
2	E	69/103 (67%)	69 (100%)	0	100	100
3	C	25/36 (69%)	25 (100%)	0	100	100
3	F	19/36 (53%)	19 (100%)	0	100	100
All	All	367/554 (66%)	362 (99%)	5 (1%)	67	89

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

Mol	Chain	Res	Type
1	A	200	GLN
1	A	206	GLU
1	A	254	ASN
2	B	127	ILE
1	D	210	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 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	107/155 (69%)	-0.22	1 (0%) 84   84	32, 65, 113, 128	0
1	D	95/155 (61%)	-0.03	0 100   100	36, 63, 111, 163	0
2	B	104/112 (92%)	-0.15	1 (0%) 82   82	24, 70, 111, 139	0
2	E	103/112 (91%)	-0.03	2 (1%) 66   65	35, 76, 122, 139	0
3	C	29/41 (70%)	-0.34	0 100   100	34, 63, 93, 109	0
3	F	24/41 (58%)	-0.38	0 100   100	53, 79, 109, 138	0
All	All	462/616 (75%)	-0.14	4 (0%) 84   84	24, 68, 118, 163	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	153	TRP	2.9
2	E	132	TYR	2.3
1	A	246	GLN	2.1
2	E	210	GLN	2.1

### 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.