



# wwPDB X-ray Structure Validation Summary Report

Feb 1, 2022 – 03:16 pm GMT

PDB ID : 6ZUE  
Title : Crystal structure of human DDB1 bound to human DCAF1 (amino acid residues 1046-1396)  
Authors : Schwefel, D.; Taylor, I.A.  
Deposited on : 2020-07-22  
Resolution : 3.09 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 : 2.26  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.26

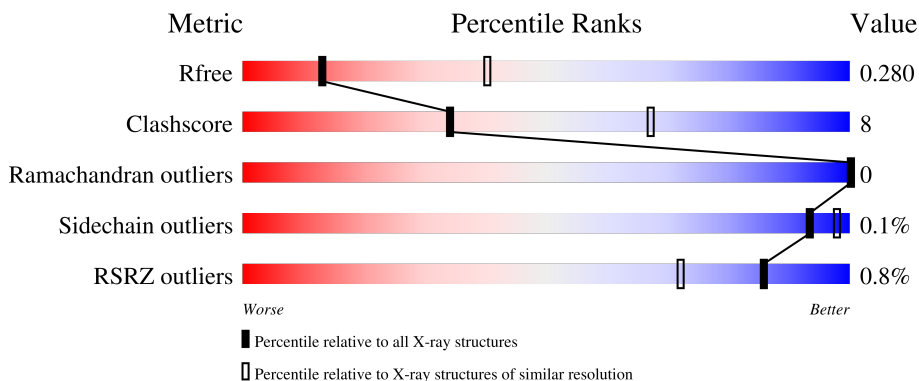
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.09 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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  $\geq 3$ , 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	1142	 76% 20%
2	B	373	 69% 17% 14%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 11289 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 DNA damage-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1106	8689	5514	1462	1665	48	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q16531
A	0	SER	-	expression tag	UNP Q16531

- Molecule 2 is a protein called DDB1- and CUL4-associated factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	322	2590	1642	455	477	16	0	1	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1045	MET	-	initiating methionine	UNP Q9Y4B6
B	1397	GLU	-	expression tag	UNP Q9Y4B6
B	1398	LEU	-	expression tag	UNP Q9Y4B6
B	1399	ALA	-	expression tag	UNP Q9Y4B6
B	1400	LEU	-	expression tag	UNP Q9Y4B6
B	1401	VAL	-	expression tag	UNP Q9Y4B6
B	1402	PRO	-	expression tag	UNP Q9Y4B6
B	1403	ARG	-	expression tag	UNP Q9Y4B6
B	1404	GLY	-	expression tag	UNP Q9Y4B6
B	1405	SER	-	expression tag	UNP Q9Y4B6
B	1406	SER	-	expression tag	UNP Q9Y4B6
B	1407	ALA	-	expression tag	UNP Q9Y4B6
B	1408	HIS	-	expression tag	UNP Q9Y4B6
B	1409	HIS	-	expression tag	UNP Q9Y4B6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1410	HIS	-	expression tag	UNP Q9Y4B6
B	1411	HIS	-	expression tag	UNP Q9Y4B6
B	1412	HIS	-	expression tag	UNP Q9Y4B6
B	1413	HIS	-	expression tag	UNP Q9Y4B6
B	1414	HIS	-	expression tag	UNP Q9Y4B6
B	1415	HIS	-	expression tag	UNP Q9Y4B6
B	1416	HIS	-	expression tag	UNP Q9Y4B6
B	1417	HIS	-	expression tag	UNP Q9Y4B6

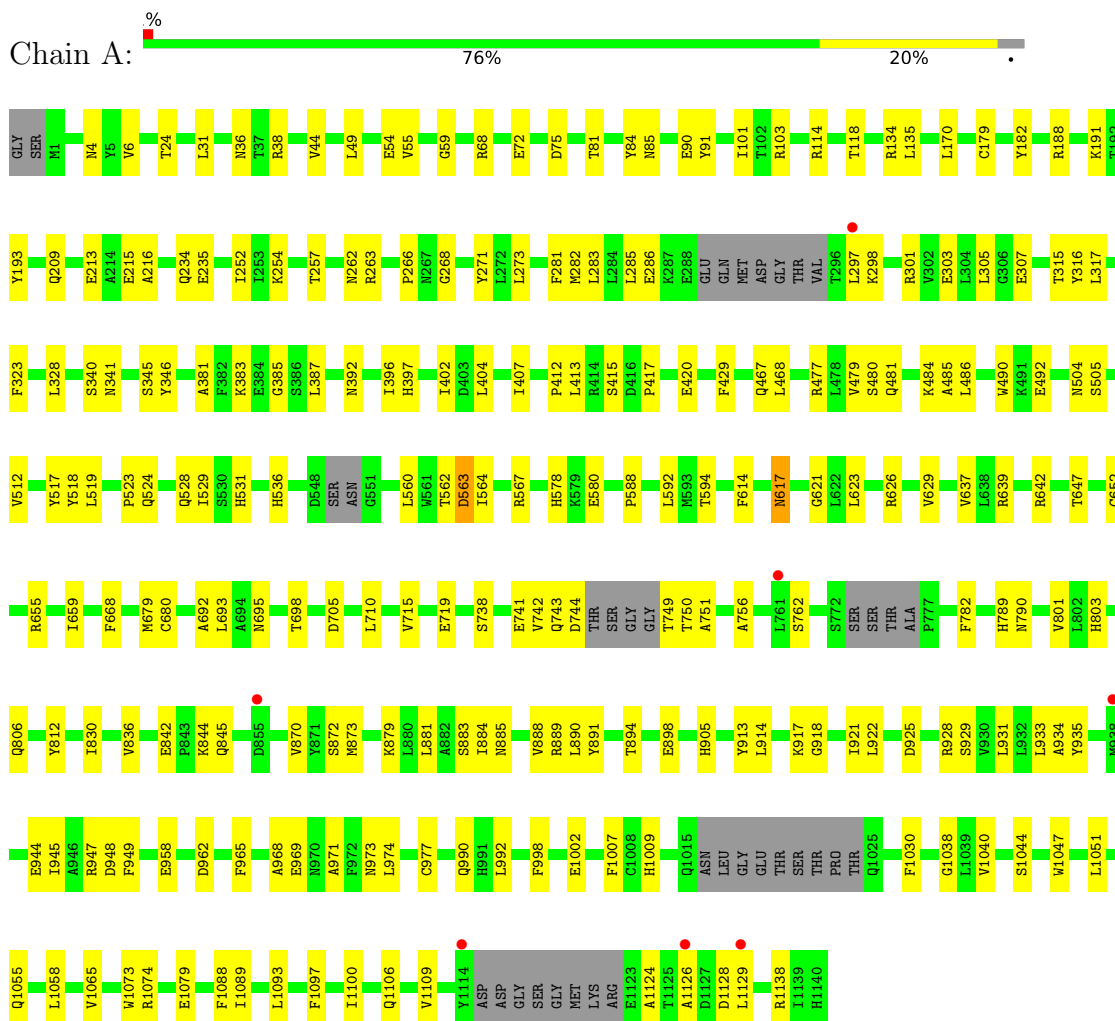
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	9	Total O 9 9	0	0
3	B	1	Total O 1 1	0	0

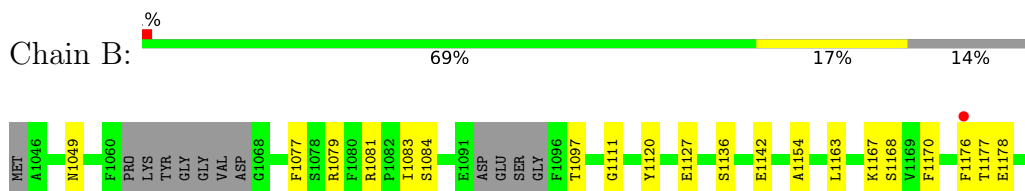
### 3 Residue-property plots [i](#)

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: DNA damage-binding protein 1



- Molecule 2: DDB1- and CUL4-associated factor 1



GLU	D1204	GLU	D1204
ARG	I1205	GLU	L1212
MET	L1212	GLU	L1213
LYS	L1213	GLU	L1214
SER	L1214	GLU	N1221
P1329	N1221	GLU	N1222
F1337	N1222	GLU	N1226
Y1342	N1226	GLU	N1289
K1343	N1289	GLU	K1257
P1344	K1257	GLU	F1258
R1352	F1258	GLU	N1259
F1355	N1259	GLU	M1260
D1356	M1260	GLU	N1261
L1357	N1261	GLU	I1262
C1358	I1262	GLU	S1263
T1359	S1263	GLU	P1268
C1364	P1268	GLU	E1272
Y1365	E1272	GLU	I1275
L1366	I1275	GLU	N1276
E1370	N1276	GLU	T1277
M1371	T1277	GLU	E1278
Q1372	E1278	GLU	I1279
GLY	I1279	GLU	L1287
SER	L1287	GLU	L1288
MET	L1288	GLU	H1289
ASP	H1289	GLU	T1290
ASP	T1290	GLU	Q1296
ALA	Q1296	GLU	C1297
LEU	C1297	GLU	F1301
ASN	F1301	GLU	V1307
MET	V1307	GLU	M1308
D1381	M1308	GLU	Y1309
R1385	Y1309	GLU	G1310
Q1392	G1310	GLU	A1311
L1400	A1311	GLU	M1312
VAL	M1312	GLU	L1313
PRO	L1313	GLU	Q1314
ARG	Q1314	GLU	ALA
GLY	ALA	GLU	ASP
SER	HIS	GLU	ASP
SER	HIS	GLU	GLU
ALA	HIS	GLU	ASP
HIS	HIS	GLU	ASP
HIS	HIS	GLU	LEU
HIS	HIS	GLU	MET
HIS	HIS	GLU	GLU
HIS	HIS	GLU	

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.38Å 153.63Å 223.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.63 – 3.09 48.63 – 3.09	Depositor EDS
% Data completeness (in resolution range)	96.6 (48.63-3.09) 96.8 (48.63-3.09)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.97 (at 3.07Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, $R_{free}$	0.222 , 0.279 0.222 , 0.280	Depositor DCC
$R_{free}$ test set	1786 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	101.0	Xtriage
Anisotropy	0.593	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11289	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	113.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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	0.29	0/8850	0.53	0/11981
2	B	0.30	0/2652	0.55	1/3587 (0.0%)
All	All	0.29	0/11502	0.54	1/15568 (0.0%)

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	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	1168	SER	N-CA-CB	6.01	119.51	110.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	563	ASP	Peptide

### 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	8689	0	8671	148	1
2	B	2590	0	2507	43	0
3	A	9	0	0	0	0
3	B	1	0	0	0	0
All	All	11289	0	11178	188	1

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 188 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:A:969:GLU:HG2	1:A:971:ALA:H	1.41	0.84
1:A:517:TYR:HE1	1:A:531:HIS:HD1	1.33	0.76
2:B:1120:TYR:CE1	2:B:1127:GLU:HG2	2.24	0.73
1:A:114:ARG:NH1	1:A:1079:GLU:OE1	2.21	0.73
2:B:1198:ASP:OD2	2:B:1222:ASN:ND2	2.21	0.73

All (1) 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)
1:A:346:TYR:OH	1:A:420:GLU:OE2[2_555]	2.18	0.02

## 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	1093/1142 (96%)	1031 (94%)	62 (6%)	0	100 100
2	B	313/373 (84%)	298 (95%)	15 (5%)	0	100 100
All	All	1406/1515 (93%)	1329 (94%)	77 (6%)	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	A	973/1000 (97%)	972 (100%)	1 (0%)	93	98
2	B	285/327 (87%)	285 (100%)	0	100	100
All	All	1258/1327 (95%)	1257 (100%)	1 (0%)	93	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	617	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	467	GLN
1	A	481	GLN
1	A	617	ASN

### 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](#)

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	1106/1142 (96%)	-0.09	7 (0%) 89 78	69, 108, 160, 252	0
2	B	322/373 (86%)	-0.05	4 (1%) 79 61	76, 114, 159, 189	0
All	All	1428/1515 (94%)	-0.08	11 (0%) 86 72	69, 109, 160, 252	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	761	LEU	2.9
1	A	855	ASP	2.7
1	A	938	MET	2.6
2	B	1355	PHE	2.4
2	B	1212	LEU	2.2

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.