



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 06:07 PM UTC

PDB ID : 4GUQ / pdb_00004guq
Title : Structure of mutS139F p73 DNA binding domain complexed with 20BP DNA response element
Authors : Ethayathulla, A.S.; Nguyen, H.T.; Viadiu, H.
Deposited on : 2012-08-29
Resolution : 3.70 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

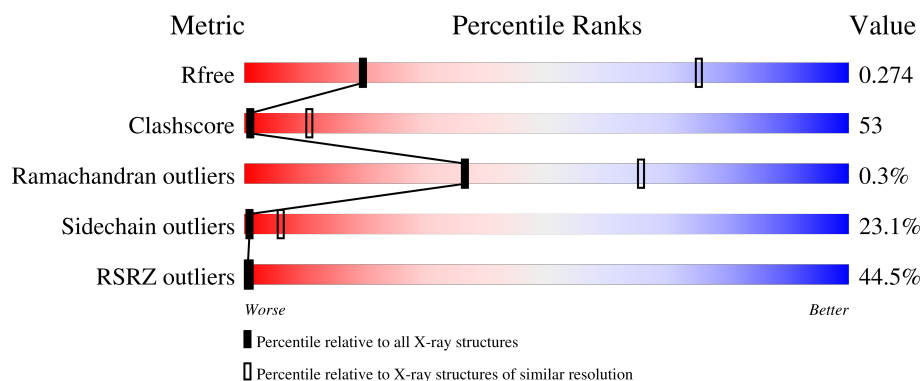
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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}	180053	1131 (3.80-3.60)
Clashscore	190562	1171 (3.80-3.60)
Ramachandran outliers	187476	1129 (3.80-3.60)
Sidechain outliers	187428	1126 (3.80-3.60)
RSRZ outliers	180081	1130 (3.80-3.60)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	E	10	 60% 40%
1	F	10	 20% 80% 20%
2	A	210	 43% 33% 45% 18% 5%
2	B	210	 44% 37% 44% 11% 6%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3533 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 DNA chain called DNA (5'-D(P*GP*AP*AP*CP*AP*TP*GP*TP*TP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	10	Total	C	N	O	P	0	0	0
			205	98	37	60	10			
1	F	10	Total	C	N	O	P	0	0	0
			205	98	37	60	10			

- Molecule 2 is a protein called Tumor protein p73.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	200	Total	C	N	O	S	0	0	0
			1566	982	280	293	11			
2	B	198	Total	C	N	O	S	0	0	0
			1555	977	279	288	11			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	MET	-	initiating methionine	UNP O15350
A	104	GLY	-	expression tag	UNP O15350
A	105	HIS	-	expression tag	UNP O15350
A	106	HIS	-	expression tag	UNP O15350
A	107	HIS	-	expression tag	UNP O15350
A	108	HIS	-	expression tag	UNP O15350
A	109	HIS	-	expression tag	UNP O15350
A	110	HIS	-	expression tag	UNP O15350
A	111	HIS	-	expression tag	UNP O15350
A	112	HIS	-	expression tag	UNP O15350
A	113	GLU	-	expression tag	UNP O15350
A	114	PHE	-	expression tag	UNP O15350
A	139	PHE	SER	engineered mutation	UNP O15350
B	103	MET	-	initiating methionine	UNP O15350
B	104	GLY	-	expression tag	UNP O15350

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	105	HIS	-	expression tag	UNP O15350
B	106	HIS	-	expression tag	UNP O15350
B	107	HIS	-	expression tag	UNP O15350
B	108	HIS	-	expression tag	UNP O15350
B	109	HIS	-	expression tag	UNP O15350
B	110	HIS	-	expression tag	UNP O15350
B	111	HIS	-	expression tag	UNP O15350
B	112	HIS	-	expression tag	UNP O15350
B	113	GLU	-	expression tag	UNP O15350
B	114	PHE	-	expression tag	UNP O15350
B	139	PHE	SER	engineered mutation	UNP O15350

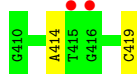
- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

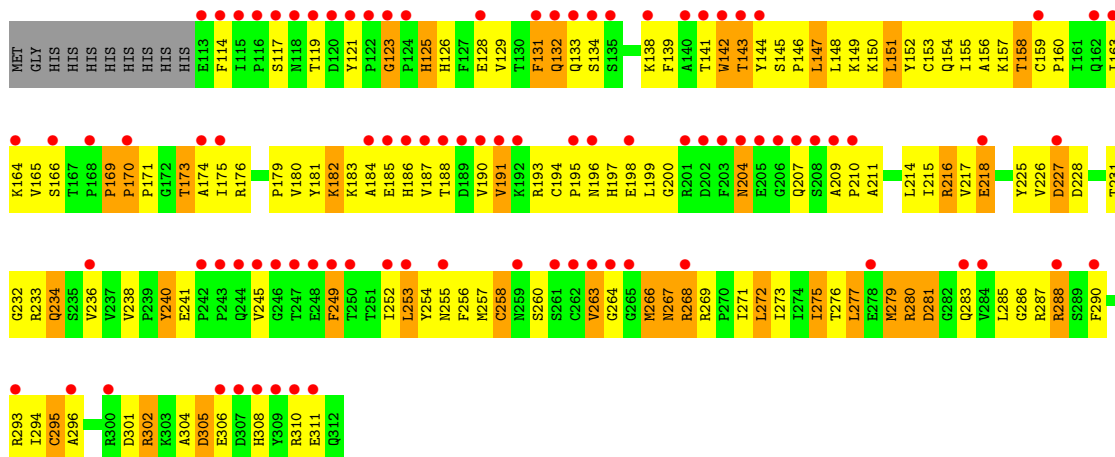
- Molecule 1: DNA (5'-D(P*GP*AP*AP*CP*AP*TP*GP*TP*TP*C)-3')



- Molecule 1: DNA (5'-D(P*GP*AP*AP*CP*AP*TP*GP*TP*TP*C)-3')



- Molecule 2: Tumor protein p73



- Molecule 2: Tumor protein p73



L285	G286	R287	R288	S289	F290	E291	G292	R293	I294	C295	A296	R300	D301	R302	K303	E306	D307	H308	Y309	R310	E311	GLN																																					
Y225	V226	D227	D228	P229	V230	T231	Q232	R233	Q234	S235	V236	V237	V238	P239	Y240	E241	P242	P243	Q244	V245	G246	T247	E248	F249	T250	T251	L252	L253	Y254	N255	F256	M257	C258	N259	S260	S261	C262	V263	G264	G265	M266	N267	R268	F270	I271	L272	L273	I274	L275	L276	L277	E278	M279	R280	D281	Q282	V284		
K164	V165	SER	T167	P168	P169	P170	P171	G172	T173	A174	I175	R176	A177	M178	P179	V180	Y181	K182	K183	A184	E185	H186	V187	T188	D189	V190	V191	K192	A193	C194	P195	N196	H197	E198	L199	G200	R201	D202	N204	G206	Q207	S208	A209	P210	A211	S212	H213	L214	I215	R216	V217	E218	G219	N220	N221	L222			
MET	GLY	HIS	HIS	HIS	HIS	HIS	HIS	HIS	HIS	E113	F114	T115	P116	S117	N118	T119	D120	Y121	P122		H25	H126	F127	E128	V129	T130	F131	Q132	Q133	S134	S135	T136	A137	K138	F139	A140	T141	W142	T143	Y144	S145	P146	L147	L148	K149	K150	L151	Y152	C153	Q154	I155	A156	K157	T158	C159	P160	I161	Q162	I163

4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α , β , γ	172.50Å 172.50Å 34.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	100.00 – 3.70 100.00 – 3.70	Depositor EDS
% Data completeness (in resolution range)	98.5 (100.00-3.70) 98.5 (100.00-3.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.82 (at 3.56Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.302 , 0.316 0.302 , 0.274	Depositor DCC
R_{free} test set	330 reflections (4.63%)	wwPDB-VP
Wilson B-factor (Å ²)	90.5	Xtriage
Anisotropy	0.163	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 145.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.068 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.78	EDS
Total number of atoms	3533	wwPDB-VP
Average B, all atoms (Å ²)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section:
ZN

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	E	0.38	0/229	0.82	0/351
1	F	0.44	0/229	0.91	1/351 (0.3%)
2	A	1.06	2/1605 (0.1%)	1.29	13/2181 (0.6%)
2	B	1.12	4/1594 (0.3%)	1.29	15/2166 (0.7%)
All	All	1.03	6/3657 (0.2%)	1.24	29/5049 (0.6%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	169	PRO	C-N	8.82	1.41	1.33
2	B	204	ASN	CA-C	6.76	1.61	1.52
2	B	242	PRO	C-N	5.56	1.40	1.33
2	B	193	ARG	CZ-NH2	5.39	1.40	1.33
2	B	243	PRO	N-CD	5.22	1.55	1.47

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	230	VAL	N-CA-C	10.26	120.18	110.53
2	A	169	PRO	CA-C-N	-8.94	113.14	119.66
2	A	169	PRO	C-N-CA	-8.94	113.14	119.66
2	B	307	ASP	N-CA-C	-8.36	102.11	111.14
2	B	242	PRO	CA-C-N	-7.08	112.46	119.76

There are no chirality outliers.

There are no planarity outliers.

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	E	205	0	114	2	1
1	F	205	0	114	1	0
2	A	1566	0	1532	188	0
2	B	1555	0	1523	184	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
All	All	3533	0	3283	361	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 53.

The worst 5 of 361 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:147:LEU:HD22	2:B:148:LEU:CD2	1.57	1.34
2:B:216:ARG:HH11	2:B:257:MET:HE2	1.11	1.15
2:B:147:LEU:HD22	2:B:148:LEU:HD23	1.27	1.12
2:A:179:PRO:HB2	2:A:191:VAL:HG11	1.19	1.10
2:A:264:GLY:HA3	2:B:196:ASN:HD21	1.16	1.08

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:E:400:DG:O5'	1:E:409:DC:O3'[1_556]	2.04	0.16

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	
2	A	198/210 (94%)	182 (92%)	16 (8%)	0	100	100
2	B	194/210 (92%)	176 (91%)	17 (9%)	1 (0%)	24	56
All	All	392/420 (93%)	358 (91%)	33 (8%)	1 (0%)	36	65

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	169	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	
2	A	174/186 (94%)	131 (75%)	43 (25%)	1	5
2	B	173/186 (93%)	136 (79%)	37 (21%)	1	7
All	All	347/372 (93%)	267 (77%)	80 (23%)	1	6

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	193	ARG
2	B	268	ARG
2	B	197	HIS
2	B	218	GLU
2	B	280	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	186	HIS
2	B	213	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	308	HIS
2	B	259	ASN
2	A	267	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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.

5.7 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, 95th 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 ²)	Q<0.9
1	E	10/10 (100%)	0.66	0 100 100	73, 77, 85, 91	0
1	F	10/10 (100%)	1.23	2 (20%) 3 3	76, 81, 99, 105	0
2	A	200/210 (95%)	2.43	91 (45%) 0 1	25, 71, 107, 175	46 (23%)
2	B	198/210 (94%)	2.46	93 (46%) 0 1	27, 68, 131, 168	58 (29%)
All	All	418/440 (95%)	2.37	186 (44%) 0 1	25, 72, 123, 175	104 (24%)

The worst 5 of 186 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	185	GLU	10.0
2	B	186	HIS	9.5
2	A	185	GLU	8.5
2	A	206	GLY	8.4
2	A	263	VAL	8.3

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 oligosaccharides 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. The B-factors column lists the minimum, median, 95th 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	B-factors($\text{\AA}^2$)	Q<0.9
3	ZN	B	401	1/1	0.97	0.12	83,83,83,83	0
3	ZN	A	401	1/1	0.99	0.08	88,88,88,88	0

6.5 Other polymers [i](#)

There are no such residues in this entry.