



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:31 AM GMT

PDB ID : 2XDN  
Title : TRANSCRIPTION FACTOR TTGR H67A MUTANT  
Authors : Daniels, C.; Lu, D.; Zhang, X.; Ramos, J.L.  
Deposited on : 2010-05-04  
Resolution : 2.20 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 (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

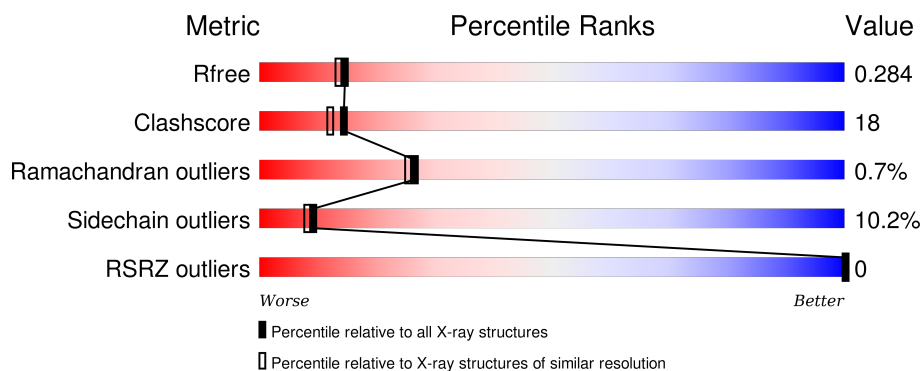
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	210	<div> <div>66%</div> <div>26%</div> <div>5%</div> <div>.</div> </div>
1	B	210	<div> <div>50%</div> <div>41%</div> <div>6%</div> <div>.</div> </div>
1	C	210	<div> <div>63%</div> <div>30%</div> <div>5%</div> <div>.</div> </div>
1	D	210	<div> <div>63%</div> <div>31%</div> <div>.</div> <div>.</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6363 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 HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	206	Total	C	N	O	S	0	0	0
			1577	988	284	297	8			
1	B	205	Total	C	N	O	S	0	0	0
			1580	988	291	293	8			
1	C	206	Total	C	N	O	S	0	0	0
			1605	1002	295	300	8			
1	D	205	Total	C	N	O	S	0	1	0
			1590	994	291	298	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	67	ALA	HIS	ENGINEERED MUTATION	UNP Q9AIU0
B	67	ALA	HIS	ENGINEERED MUTATION	UNP Q9AIU0
C	67	ALA	HIS	ENGINEERED MUTATION	UNP Q9AIU0
D	67	ALA	HIS	ENGINEERED MUTATION	UNP Q9AIU0

- Molecule 2 is water.

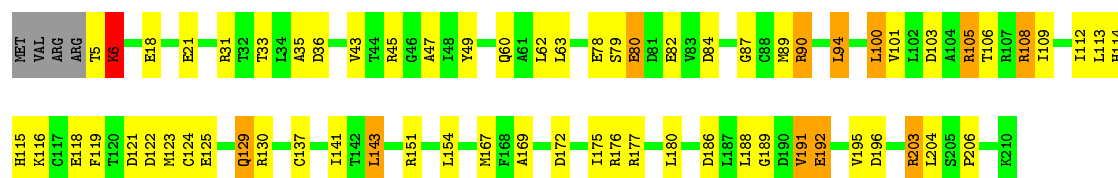
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	O	0	0
			2	2		
2	B	4	Total	O	0	0
			4	4		
2	C	3	Total	O	0	0
			3	3		
2	D	2	Total	O	0	0
			2	2		

### 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.

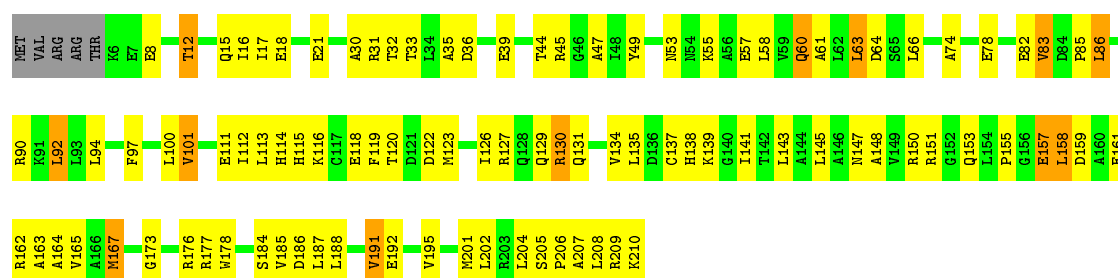
#### • Molecule 1: HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGR

Chain A: 



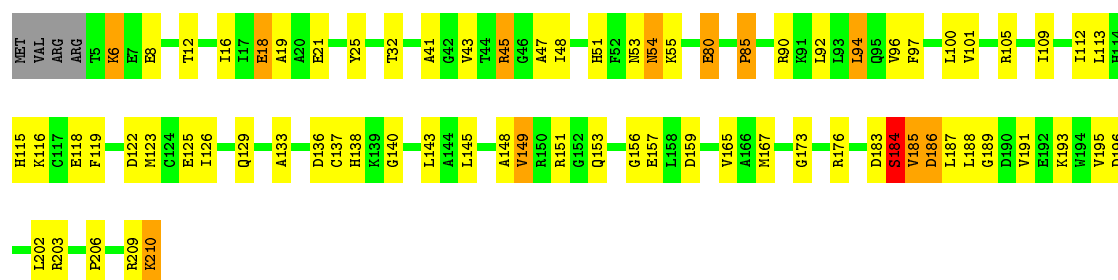
#### • Molecule 1: HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGR

Chain B: 



#### • Molecule 1: HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGR

Chain C: 



#### • Molecule 1: HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGR

Chain D: 

MET	VAL	ARG	ARG	THR	K6	E11	T12	R13	A14	Q15	T16	I17	E18	E21	R22	K23	F24	V29	L34	A35	A38	E39	L40	V43	T44	R45	Y49	W50	R53	R54	R55	A56	E57	L58	V59	Q60	D64	E80	R90	R91	L92	V96	F97	I98	E99	L100	V101
L102	D103	A104	R105	T106	R107	R108	K116	G117	E118	F119	E125	Q129	R130	L143	R151	L154	P155	L158	E161	R162	A163	A164	V165	A166	M167	F168	A169	I175	R176	R177	V185	D186	L187	L188	V191	L199	D200	M201	L202	R203	L204	S205	P206	A207	K210		

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.11Å 43.17Å 114.80Å 96.93° 99.26° 96.03°	Depositor
Resolution (Å)	41.74 – 2.20 42.20 – 2.20	Depositor EDS
% Data completeness (in resolution range)	88.3 (41.74-2.20) 83.1 (42.20-2.20)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.40 (at 2.20Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.218 , 0.288 0.220 , 0.284	Depositor DCC
$R_{free}$ test set	1795 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 22.3	EDS
Estimated twinning fraction	0.067 for k,h,-h-k-l 0.175 for -k,-h,-l 0.058 for -h,-k,h+k+l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 35518 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6363	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.65% 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	0.44	0/1600	0.65	1/2168 (0.0%)
1	B	0.40	0/1603	0.63	0/2172
1	C	0.43	0/1628	0.61	0/2202
1	D	0.43	0/1616	0.60	0/2190
All	All	0.43	0/6447	0.62	1/8732 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	90	ARG	NE-CZ-NH1	-5.39	117.61	120.30

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	1577	0	1541	57	0
1	B	1580	0	1552	82	0
1	C	1605	0	1591	55	0
1	D	1590	0	1565	56	0
2	A	2	0	0	0	0
2	B	4	0	0	0	0
2	C	3	0	0	0	0
2	D	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	6363	0	6249	227	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 18.

The worst 5 of 227 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:C:18:GLU:HG3	1:C:105:ARG:HH12	1.25	1.00
1:C:109:ILE:O	1:C:113:LEU:HD23	1.67	0.93
1:C:90:ARG:HH12	1:C:196:ASP:CG	1.77	0.88
1:D:163:ALA:HB1	1:D:202:LEU:HD22	1.57	0.85
1:B:86:LEU:HD13	1:B:153:GLN:HB3	1.56	0.85

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	204/210 (97%)	190 (93%)	12 (6%)	2 (1%)	19	16
1	B	203/210 (97%)	192 (95%)	10 (5%)	1 (0%)	34	35
1	C	204/210 (97%)	189 (93%)	12 (6%)	3 (2%)	13	9
1	D	204/210 (97%)	193 (95%)	11 (5%)	0	100	100
All	All	815/840 (97%)	764 (94%)	45 (6%)	6 (1%)	26	25

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	184	SER
1	C	186	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	45	ARG
1	A	6	LYS
1	B	186	ASP

### 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	157/174 (90%)	138 (88%)	19 (12%)	6	5
1	B	157/174 (90%)	141 (90%)	16 (10%)	9	8
1	C	163/174 (94%)	146 (90%)	17 (10%)	9	8
1	D	160/174 (92%)	147 (92%)	13 (8%)	15	14
All	All	637/696 (92%)	572 (90%)	65 (10%)	9	8

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	130	ARG
1	C	8	GLU
1	D	158	LEU
1	B	135	LEU
1	B	167	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	115	HIS
1	C	51	HIS
1	C	129	GLN
1	B	110	ASN
1	D	15	GLN

### 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	206/210 (98%)	-0.56	0 100 100	19, 32, 51, 75	0
1	B	205/210 (97%)	-0.50	0 100 100	21, 34, 49, 54	0
1	C	206/210 (98%)	-0.49	0 100 100	22, 35, 48, 64	0
1	D	205/210 (97%)	-0.51	0 100 100	21, 32, 50, 60	0
All	All	822/840 (97%)	-0.52	0 100 100	19, 33, 50, 75	0

There are no RSRZ outliers to report.

### 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 [i](#)

There are no such residues in this entry.