



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 06:27 PM GMT

PDB ID : 1AU7  
Title : PIT-1 MUTANT/DNA COMPLEX  
Authors : Jacobson, E.M.; Li, P.; Leon-Del-Rio, A.; Rosenfeld, M.G.; Aggarwal, A.K.  
Deposited on : 1997-09-12  
Resolution : 2.30 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

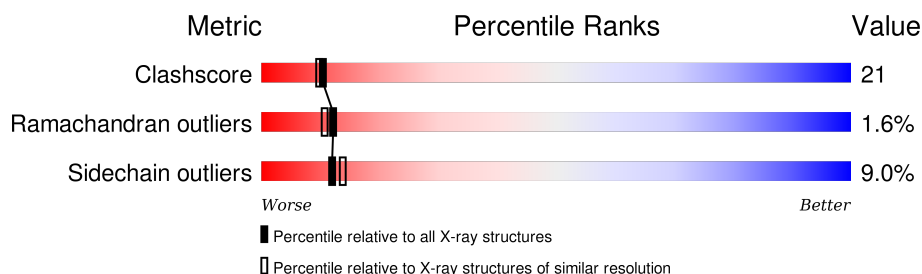
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

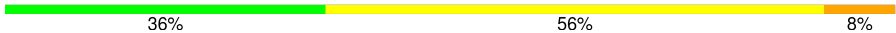



Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	25	 36% 56% 8%
2	D	24	 42% 46% 13%
3	A	146	 59% 25% 5% 11%
3	B	146	 55% 31% 12%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4269 atoms, of which 994 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 CONSENSUS DNA 25-MER.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	C	25	Total	C	H	N	O	P	0	0	0
			564	246	51	96	147	24			

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*TP\*TP\*CP\*CP\*TP\*CP\*AP\*TP\*GP\*TP\*AP\*TP\*AP\*TP\*AP\*C P\*AP\*TP\*GP\*AP\*GP\* GP\*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	24	Total	C	H	N	O	P	0	0	0
			530	235	44	85	143	23			

- Molecule 3 is a protein called PROTEIN PIT-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	130	Total	C	H	N	O	S	0	0	0
			1331	659	274	200	192	6			
3	B	128	Total	C	H	N	O	S	0	0	0
			1316	652	273	198	187	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	GLY	GLU	ENGINEERED	UNP P10037
A	6	MET	ILE	ENGINEERED	UNP P10037
A	8	ALA	GLU	CONFLICT	UNP P10037
A	110	ILE	VAL	CONFLICT	UNP P10037
B	5	GLY	GLU	ENGINEERED	UNP P10037
B	6	MET	ILE	ENGINEERED	UNP P10037
B	8	ALA	GLU	CONFLICT	UNP P10037
B	110	ILE	VAL	CONFLICT	UNP P10037

- Molecule 4 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	67	Total	H	O	0	0
			201	134	67		
4	B	45	Total	H	O	0	0
			135	90	45		
4	C	24	Total	H	O	0	0
			72	48	24		
4	D	40	Total	H	O	0	0
			120	80	40		

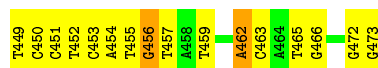
### 3 Residue-property plots

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

Note EDS was not executed.

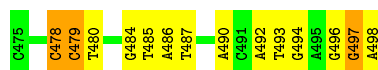
- Molecule 1: CONSENSUS DNA 25-MER

Chain C: 



- Molecule 2: DNA (5'-D(\*CP\*TP\*TP\*CP\*CP\*TP\*CP\*AP\*TP\*GP\*TP\*AP\*TP\*AP\*TP\*AP\* C P\*AP\*TP\*GP\*AP\*GP\* GP\*A)-3')

Chain D: 



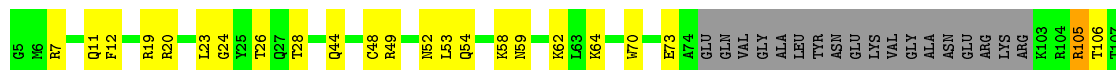
- Molecule 3: PROTEIN PIT-1

Chain A: 



- Molecule 3: PROTEIN PIT-1

Chain B: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.50 Å   50.10 Å   55.80 Å 76.70°   79.30°   67.20°	Depositor
Resolution (Å)	6.00 – 2.30	Depositor
% Data completeness (in resolution range)	84.7 (6.00-2.30)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	5.80	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.230 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4269	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

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	C	0.82	0/576	0.95	0/888
2	D	0.84	0/544	1.05	0/837
3	A	0.70	0/1071	0.84	0/1430
3	B	0.70	0/1057	0.81	0/1410
All	All	0.74	0/3248	0.89	0/4565

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	C	0	2
2	D	0	3
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	456	DG	Sidechain
1	C	462	DA	Sidechain
2	D	478	DC	Sidechain
2	D	479	DC	Sidechain
2	D	497	DG	Sidechain

## 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	C	513	51	284	31	0
2	D	486	44	272	34	0
3	A	1057	274	1062	47	10
3	B	1043	273	1059	34	5
4	A	67	134	0	11	1
4	B	45	90	0	7	1
4	C	24	48	0	7	0
4	D	40	80	0	6	7
All	All	3275	994	2677	123	12

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 21.

The worst 5 of 123 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:D:479:DC:H5''	2:D:479:DC:H6	1.15	1.06
2:D:497:DG:H2''	2:D:498:DA:H5'	1.39	1.04
1:C:451:DC:H2'	1:C:452:DT:H72	1.44	0.99
2:D:479:DC:C6	2:D:479:DC:H5''	1.98	0.99
1:C:452:DT:H2''	1:C:453:DC:H5'	1.54	0.89

The worst 5 of 12 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)
3:A:72:GLU:OE1	4:D:721:HOH:O[1_545]	1.58	0.62
3:A:118:ARG:NH2	4:D:853:HOH:O[1_455]	1.68	0.52
3:A:118:ARG:HH22	4:D:853:HOH:O[1_455]	1.12	0.48
4:D:879:HOH:H2	4:B:823:HOH:H2[1_546]	1.18	0.42
3:A:7:ARG:HH21	3:B:129:GLN:HE21[1_545]	1.24	0.36



## 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	
3	A	126/146 (86%)	119 (94%)	6 (5%)	1 (1%)	24	27
3	B	124/146 (85%)	116 (94%)	5 (4%)	3 (2%)	7	5
All	All	250/292 (86%)	235 (94%)	11 (4%)	4 (2%)	12	11

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	104	ARG
3	B	142	LYS
3	B	159	VAL
3	B	110	ILE

### 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	
3	A	111/125 (89%)	102 (92%)	9 (8%)	15	18
3	B	110/125 (88%)	99 (90%)	11 (10%)	9	11
All	All	221/250 (88%)	201 (91%)	20 (9%)	12	14

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	20	ARG
3	B	23	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	B	128	SER
3	A	145	VAL
3	A	159	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	119	HIS
3	A	123	HIS
3	A	129	GLN
3	B	52	ASN
3	B	123	HIS

### 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.