



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 03:34 AM UTC

PDB ID : 2HQE / pdb_00002hqe
Title : Crystal structure of human P100 Tudor domain: Large fragment
Authors : Shah, N.; Zhao, M.; Cheng, C.; Xu, H.; Yang, J.; Silvennoinen, O.; Liu, Z.J.;
Wang, B.C.; Southeast Collaboratory for Structural Genomics (SECSG)
Deposited on : 2006-07-18
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

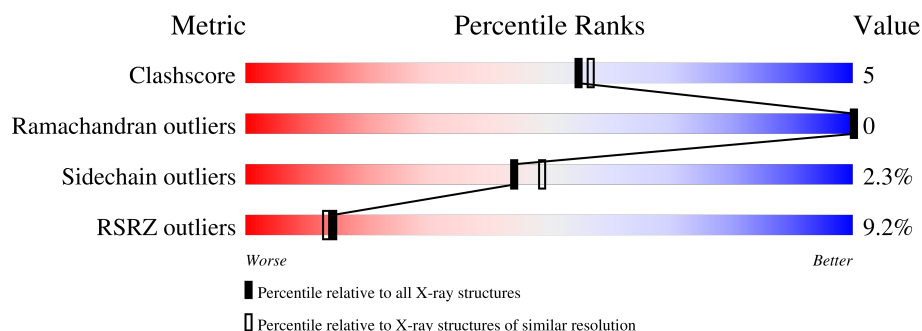
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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	A	246	 7% 81% 7% 12%
1	B	246	 4% 29% 6% 65%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2591 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 P100 Co-activator tudor domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1730	1096	298	330	6			
1	B	87	Total	C	N	O	S	0	0	0
			705	449	126	127	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	682	PHE	LEU	variant	UNP Q7KZF4
A	683	GLN	GLU	variant	UNP Q7KZF4
B	682	PHE	LEU	variant	UNP Q7KZF4
B	683	GLN	GLU	variant	UNP Q7KZF4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	109	Total	O	0	0
			109	109		
2	B	47	Total	O	0	0
			47	47		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	49.93Å 93.41Å 95.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.24 – 2.00 44.24 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.24-2.00) 97.0 (44.24-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	-0.54 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.233 , 0.249 0.227 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	26.2	Xtriage
Anisotropy	0.248	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 32.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.017 for -h,l,k	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	2591	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.49% 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 [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	0.74	5/1768 (0.3%)	0.83	4/2399 (0.2%)
1	B	0.65	1/723 (0.1%)	0.86	2/978 (0.2%)
All	All	0.72	6/2491 (0.2%)	0.84	6/3377 (0.2%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	780	GLN	CD-OE1	5.45	1.33	1.23
1	A	735	HIS	ND1-CE1	5.36	1.38	1.32
1	B	696	HIS	ND1-CE1	5.36	1.38	1.32
1	A	696	HIS	ND1-CE1	5.33	1.37	1.32
1	A	818	GLN	CD-NE2	-5.22	1.22	1.33
1	A	767	GLN	CD-OE1	5.02	1.33	1.23

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	696	HIS	CB-CG-CD2	-6.57	122.66	131.20
1	A	735	HIS	CB-CG-CD2	-6.55	122.69	131.20
1	A	696	HIS	CB-CG-CD2	-6.44	122.83	131.20
1	B	696	HIS	CB-CG-ND1	5.63	131.14	122.70
1	A	735	HIS	CB-CG-ND1	5.61	131.11	122.70
1	A	696	HIS	CB-CG-ND1	5.60	131.10	122.70

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	1730	0	1691	6	0
1	B	705	0	692	17	0
2	A	109	0	0	2	0
2	B	47	0	0	0	0
All	All	2591	0	2383	23	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 5.

All (23) 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:B:682:PHE:CZ	1:B:740:ASP:HB3	1.96	1.00
1:B:686:MET:O	1:B:690:ARG:HG2	1.69	0.93
1:B:682:PHE:CZ	1:B:740:ASP:CB	2.63	0.80
1:B:680:THR:HG23	1:B:683:GLN:H	1.51	0.74
1:A:747:LEU:HD13	1:A:752:LEU:HD21	1.70	0.72
1:B:682:PHE:CE1	1:B:686:MET:HE2	2.24	0.72
1:B:682:PHE:HE1	1:B:686:MET:HE2	1.55	0.71
1:A:682:PHE:CZ	1:A:740:ASP:HB3	2.30	0.67
1:A:747:LEU:CD1	1:A:752:LEU:HD21	2.28	0.64
2:A:110:HOH:O	1:B:743:ASN:ND2	2.34	0.61
1:A:682:PHE:CZ	1:A:740:ASP:CB	2.87	0.58
1:B:682:PHE:CZ	1:B:741:TYR:CZ	2.96	0.53
1:B:682:PHE:CE2	1:B:741:TYR:CZ	2.96	0.53
1:B:727:LYS:HB3	1:B:735:HIS:HB2	1.90	0.53
1:B:682:PHE:HZ	1:B:741:TYR:CE2	2.28	0.50
1:B:682:PHE:CZ	1:B:741:TYR:CE2	3.01	0.49
1:A:700:GLU:HA	1:A:701:GLY:HA2	1.51	0.47
2:A:155:HOH:O	1:B:719:GLU:HG3	2.17	0.44
1:B:682:PHE:HZ	1:B:740:ASP:CB	2.26	0.44
1:A:685:LEU:HD11	1:A:759:PHE:HD1	1.84	0.43
1:B:747:LEU:HB2	1:B:748:PRO:HD2	2.00	0.43
1:B:680:THR:CG2	1:B:683:GLN:HG3	2.49	0.42
1:B:704:ALA:HA	1:B:705:PRO:HD3	1.90	0.41

There are no symmetry-related clashes.

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	
1	A	215/246 (87%)	208 (97%)	7 (3%)	0	100	100
1	B	83/246 (34%)	81 (98%)	2 (2%)	0	100	100
All	All	298/492 (61%)	289 (97%)	9 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	188/213 (88%)	183 (97%)	5 (3%)	39	42
1	B	75/213 (35%)	74 (99%)	1 (1%)	61	68
All	All	263/426 (62%)	257 (98%)	6 (2%)	44	49

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	656	SER
1	A	658	LYS
1	A	699	VAL
1	A	747	LEU
1	A	761	THR
1	B	747	LEU

Sometimes 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 oligosaccharides 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

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	A	217/246 (88%)	0.43	17 (7%) 19 18	14, 26, 43, 56	0
1	B	87/246 (35%)	0.68	11 (12%) 8 7	16, 28, 43, 57	0
All	All	304/492 (61%)	0.50	28 (9%) 14 13	14, 27, 43, 57	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	682	PHE	4.8
1	B	682	PHE	4.1
1	B	698	PRO	4.1
1	B	728	VAL	3.7
1	A	694	ALA	3.3
1	A	822	SER	3.2
1	A	702	SER	3.1
1	A	699	VAL	3.0
1	A	870	TYR	3.0
1	A	700	GLU	2.9
1	B	694	ALA	2.8
1	B	680	THR	2.7
1	A	658	LYS	2.6
1	A	844	LYS	2.5
1	A	811	GLY	2.5
1	A	698	PRO	2.4
1	B	705	PRO	2.4
1	B	695	SER	2.3
1	B	731	PRO	2.3
1	B	703	TYR	2.3
1	B	696	HIS	2.2
1	B	727	LYS	2.2
1	A	821	ASP	2.1
1	A	701	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	683	GLN	2.0
1	A	847	GLN	2.0
1	A	703	TYR	2.0
1	A	820	ALA	2.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.