



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 01:32 pm GMT

PDB ID : 2F1Z
Title : Crystal structure of HAUSP
Authors : Hu, M.; Gu, L.; Jeffrey, P.D.; Shi, Y.
Deposited on : 2005-11-15
Resolution : 3.20 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

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
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

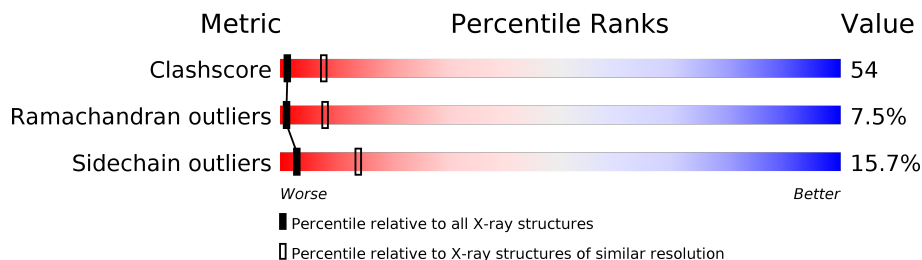
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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))
Clashscore	112137	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)

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 ≥ 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	A	522	
1	B	522	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8015 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 Ubiquitin carboxyl-terminal hydrolase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	470	Total	C	N	O	S	0	0	0
			3842	2442	652	726	22			
1	B	481	Total	C	N	O	S	0	0	0
			3933	2500	668	743	22			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	GLY	-	CLONING ARTIFACT	UNP Q93009
A	40	SER	-	CLONING ARTIFACT	UNP Q93009
A	41	HIS	-	CLONING ARTIFACT	UNP Q93009
A	42	MET	-	CLONING ARTIFACT	UNP Q93009
B	39	GLY	-	CLONING ARTIFACT	UNP Q93009
B	40	SER	-	CLONING ARTIFACT	UNP Q93009
B	41	HIS	-	CLONING ARTIFACT	UNP Q93009
B	42	MET	-	CLONING ARTIFACT	UNP Q93009

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	118	Total	O	0	0
			118	118		
2	B	122	Total	O	0	0
			122	122		



N512	A513	Y514	M515	L516	V517	Y518	L524	S525	E526	V527	L528	Q529	A530	V531	T532	D533	H534	D535	L536	Q539	L540	L544	E547	E551	A552	Q553	LYS	ARG	LYS	GLU	ARG	GLN	GLU																								
I449	L450	H451	A452	V453	L454	V455	H456	D459	N460	H461	G462	G463	H464	Y465	V466	V467	Y468	L469	N470	N471	G472	G473	D474	G475	K476	W477	C478	K479	F480	D481	D482	D483	V484	V485	T489	K490	E491	E492	A493	I494	E495	H496	N497	Y498	G499	GLY	HIS	ASP	ASP	ASP	LEU	SER	VAL	ARG	HIS	C510	T511
K378	Y379	D380	A381	G382	E383	H384	G385	L386	Q387	E388	A389	E390	K391	T397	L398	P399	P400	H403	L404	Q405	L406	M407	A408	F409	M410	Y411	D412	P413	Q414	T415	D416	Q417	N418	T419	K420	I421	N422	D423	R424	F425	E426	F427	F428	E429	Q430	P432	L433	F436	L442	P443	D444	P445	Y448				
K313	G313	T314	C315	V316	E317	G318	T319	I320	P321	K322	L323	F324	K327	M328	V329	S330	Y331	I332	Q333	C334	K335	E336	V337	D338	Y339	D342	R343	K344	Y347	I350	Q351	L352	S353	I354	K355	G356	K357	K358	N359	I360	F361	E362	S363	F364	V365	D366	Y367	V368	A369	E370	V371	Q372	D376	N377			
Y242	Y243	M244	M245	P246	T247	D250	S255	V256	P257	L258	A259	L260	Q261	R262	V263	F264	L267	Q268	H269	S270	V274	G275	T276	K277	K278	L279	T280	K281	S282	F283	T287	L288	D289	S290	F291	N292	Q293	H294	D295	Y296	Q297	E298	L299	C300	R301	L304	D305	A306	N307	E308	N309	K310	M311				
Y106	P107	D108	R109	P110	H111	Q112	K113	S114	V115	G116	F117	F118	L119	N122	D198	A199	P200	H201	G202	A204	L138	K139	D206	S207	K208	K209	Y143	T211	G212	Y213	V214	G215	L216	K217	N218	Q219	G220	A221	T222	C223	Y224	M225	N226	S227	D164	F167	S168	N169	F170	M171	A172	W173	V176	T177	D178	P179	
G182	F183	I184	D185	D186	D187	K188	V189	T190	F191	E192	V193	F194	V195	D198	A199	P200	H201	G202	A204	L138	K139	D206	S207	K208	K209	Y143	T211	G212	Y213	V214	G215	L216	K217	N218	Q219	G220	A221	T222	C223	Y224	M225	N226	S227	D164	F167	S168	N169	F170	M171	A172	W173	V176	T177	D178	P179		

4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	97.62Å 219.86Å 130.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.20	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-3.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.265 , 0.316	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8015	wwPDB-VP
Average B, all atoms (Å ²)	56.0	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	A	0.46	0/3935	0.74	1/5312 (0.0%)
1	B	0.50	2/4032 (0.0%)	0.81	9/5449 (0.2%)
All	All	0.48	2/7967 (0.0%)	0.78	10/10761 (0.1%)

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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	104	ARG	CZ-NH1	5.37	1.40	1.33
1	B	104	ARG	CB-CG	5.26	1.66	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	104	ARG	NE-CZ-NH2	-9.38	115.61	120.30
1	B	414	GLN	N-CA-C	-9.00	86.71	111.00
1	B	104	ARG	NE-CZ-NH1	8.08	124.34	120.30
1	B	105	PHE	N-CA-C	7.61	131.54	111.00
1	B	387	GLN	CB-CA-C	-6.44	97.52	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	448	TYR	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	A	3842	0	3710	428	1
1	B	3933	0	3794	406	0
2	A	118	0	0	44	0
2	B	122	0	0	41	0
All	All	8015	0	7504	821	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 54.

The worst 5 of 821 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:109:ARG:HE	1:B:110:PRO:HD2	1.21	1.04
1:B:491:GLU:HA	1:B:495:GLU:HG3	1.39	1.03
1:A:501:HIS:HB3	1:B:413:PRO:HG3	1.38	1.01
1:B:214:VAL:HG22	1:B:215:GLY:H	1.27	1.00
1:A:294:HIS:HB3	1:A:298:GLU:HG3	1.47	0.97

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:287:THR:O	1:A:287:THR:O[4_555]	1.83	0.37

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	462/522 (88%)	356 (77%)	71 (15%)	35 (8%)	1	8
1	B	477/522 (91%)	364 (76%)	78 (16%)	35 (7%)	1	8
All	All	939/1044 (90%)	720 (77%)	149 (16%)	70 (8%)	1	8

5 of 70 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	127	SER
1	A	183	PHE
1	A	203	VAL
1	A	277	LYS
1	A	334	CYS

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	428/472 (91%)	355 (83%)	73 (17%)	2	11
1	B	437/472 (93%)	374 (86%)	63 (14%)	4	17
All	All	865/944 (92%)	729 (84%)	136 (16%)	3	14

5 of 136 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	478	CYS
1	B	106	TYR
1	B	474	ASP
1	A	486	SER
1	A	540	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	512	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	156	HIS
1	B	464	HIS
1	A	538	GLN
1	B	219	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](#)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.