



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 2, 2025 – 02:06 PM EDT

PDB ID : 9CEA / pdb\_00009cea  
Title : Crystal structure of HP1alpha chromoshadow domain  
Authors : Selvam, K.; Gaurav, N.; Kutateladze, T.G.  
Deposited on : 2024-06-26  
Resolution : 2.15 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

<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 Phenix2.0rc1  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.43.1

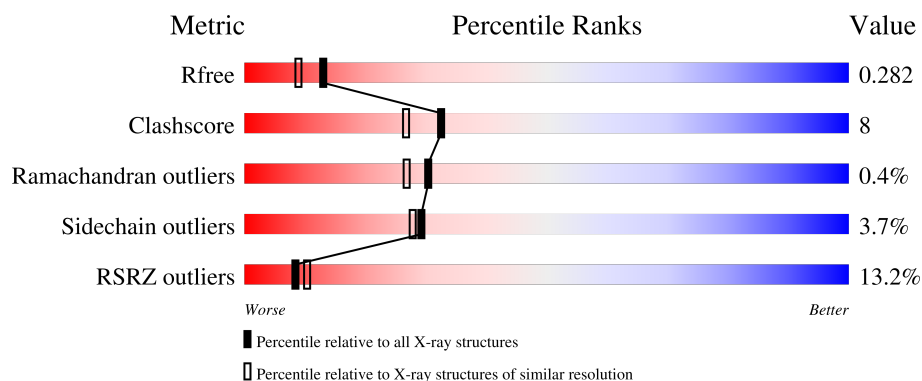
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	164625	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (2.16-2.16)

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  $\geq 3$ , 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	63	<div> <div>6%</div> <div> <div></div> <div>86%</div> <div>10%</div> <div>5%</div> </div> </div>
1	B	63	<div> <div>13%</div> <div> <div></div> <div>79%</div> <div>14%</div> <div>5%</div> </div> </div>
1	C	63	<div> <div>21%</div> <div> <div></div> <div>79%</div> <div>17%</div> <div>• •</div> </div> </div>
1	D	63	<div> <div>11%</div> <div> <div></div> <div>76%</div> <div>13%</div> <div>5%</div> <div>• •</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1991 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 Chromobox protein homolog 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	60	Total	C	N	O	S	0	0	0
			469	299	77	89	4			
1	B	60	Total	C	N	O	S	0	0	0
			458	291	76	87	4			
1	C	62	Total	C	N	O	S	0	1	0
			487	315	79	90	3			
1	D	61	Total	C	N	O	S	0	0	0
			468	299	75	90	4			

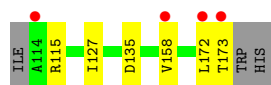
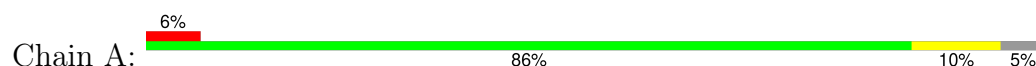
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	31	Total	O	0	1
			32	32		
2	B	24	Total	O	0	0
			24	24		
2	C	20	Total	O	0	0
			20	20		
2	D	32	Total	O	0	1
			33	33		

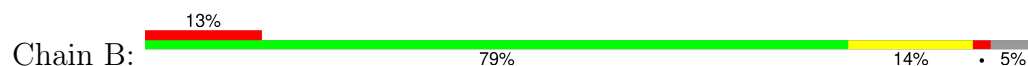
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

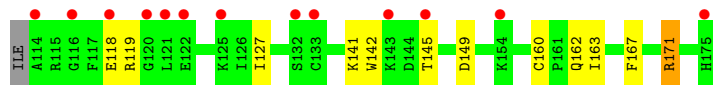
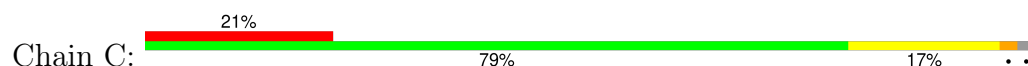
- Molecule 1: Chromobox protein homolog 5



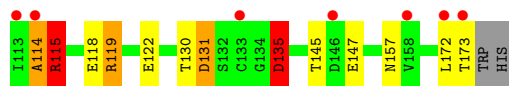
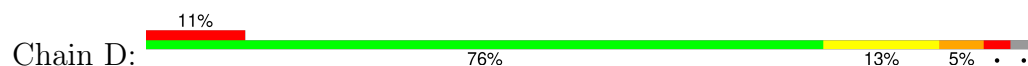
- Molecule 1: Chromobox protein homolog 5



- Molecule 1: Chromobox protein homolog 5



- Molecule 1: Chromobox protein homolog 5



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.76Å 66.85Å 70.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.18 – 2.15 46.18 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.18-2.15) 99.9 (46.18-2.15)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.227 , 0.277 0.236 , 0.282	Depositor DCC
$R_{free}$ test set	772 reflections (4.54%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.3	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 58.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	1991	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.71% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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	0.64	0/476	1.19	1/640 (0.2%)
1	B	0.65	0/465	1.13	2/627 (0.3%)
1	C	0.61	0/500	1.07	1/676 (0.1%)
1	D	0.65	0/475	1.28	6/640 (0.9%)
All	All	0.64	0/1916	1.17	10/2583 (0.4%)

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	B	0	1
1	C	0	1
1	D	0	2
All	All	0	4

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	135	ASP	CA-CB-CG	7.10	119.70	112.60
1	B	115	ARG	N-CA-CB	-6.39	102.40	110.90
1	D	131	ASP	CB-CA-C	5.91	121.05	109.35
1	A	135	ASP	CA-CB-CG	5.80	118.40	112.60
1	D	130	THR	CA-CB-OG1	-5.54	101.29	109.60

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	115	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	C	171	ARG	Sidechain
1	D	115	ARG	Sidechain
1	D	119	ARG	Sidechain

## 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	469	0	468	3	0
1	B	458	0	444	10	0
1	C	487	0	470	7	0
1	D	468	0	459	11	0
2	A	32	0	0	1	0
2	B	24	0	0	0	0
2	C	20	0	0	0	0
2	D	33	0	0	2	0
All	All	1991	0	1841	29	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 8.

The worst 5 of 29 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:145:THR:HG22	1:B:147:GLU:H	1.26	1.00
1:B:119:ARG:HH12	1:B:145:THR:HG21	1.40	0.84
1:A:172:LEU:O	1:A:173:THR:HG23	1.77	0.83
1:D:145:THR:HG22	1:D:147:GLU:H	1.49	0.77
2:A:223:HOH:O	1:B:158:VAL:HG22	1.93	0.69

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	58/63 (92%)	58 (100%)	0	0	100	100
1	B	58/63 (92%)	57 (98%)	1 (2%)	0	100	100
1	C	61/63 (97%)	57 (93%)	4 (7%)	0	100	100
1	D	59/63 (94%)	57 (97%)	1 (2%)	1 (2%)	7	3
All	All	236/252 (94%)	229 (97%)	6 (2%)	1 (0%)	30	27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	114	ALA

### 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	49/53 (92%)	47 (96%)	2 (4%)	26	25
1	B	46/53 (87%)	43 (94%)	3 (6%)	14	9
1	C	48/53 (91%)	48 (100%)	0	100	100
1	D	48/53 (91%)	46 (96%)	2 (4%)	25	24
All	All	191/212 (90%)	184 (96%)	7 (4%)	29	28

5 of 7 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	B	122	GLU
1	B	163	ILE
1	D	135	ASP
1	D	115	ARG
1	B	115	ARG

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

Mol	Chain	Res	Type
1	A	162	GLN
1	B	162	GLN
1	C	157	ASN
1	C	162	GLN
1	D	162	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 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 [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	60/63 (95%)	0.42	4 (6%)	25	30	24, 39, 58, 91	0
1	B	60/63 (95%)	0.56	8 (13%)	8	10	21, 36, 69, 79	0
1	C	62/63 (98%)	1.04	13 (20%)	3	4	24, 43, 77, 102	1 (1%)
1	D	61/63 (96%)	0.56	7 (11%)	11	13	22, 34, 61, 105	0
All	All	243/252 (96%)	0.65	32 (13%)	8	10	21, 39, 74, 105	1 (0%)

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	114	ALA	5.8
1	C	114	ALA	5.5
1	D	173	THR	4.7
1	D	113	ILE	4.4
1	B	114	ALA	4.3

### 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 monosaccharides 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.