wwPDB X-ray Structure Validation Summary Report (i)

Mar 9, 2026 – 10:30 PM UTC

PDB ID	:	2O94 / pdb_00002o94
Title	:	The 97H/F mutant Structure of a glutamine-rich domain from histone deacetylase 4
Authors	:	Guo, L.; Han, A.; Bates, D.L.; Chen, L.
Deposited on	:	2006-12-13
Resolution	:	3.00 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

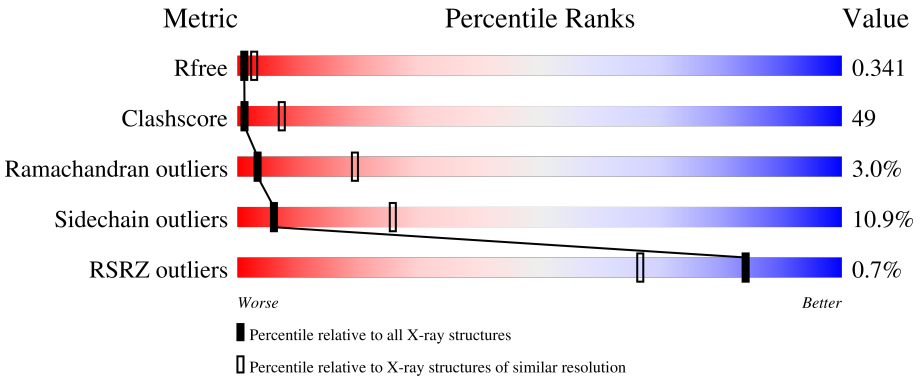
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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))
R_{free}	180053	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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 <=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	112	20% 33% 7% • 39%
1	B	112	22% 30% 8% 39%
1	C	112	% 22% 30% 7% • 39%
1	D	112	% 19% 35% 6% • 39%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2328 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 Histone deacetylase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	68	Total	C	N	O	S	0	0	0
			582	361	111	108	2			
1	B	68	Total	C	N	O	S	0	0	0
			582	361	111	108	2			
1	C	68	Total	C	N	O	S	0	0	0
			582	361	111	108	2			
1	D	68	Total	C	N	O	S	0	0	0
			582	361	111	108	2			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	GLY	-	cloning artifact	UNP P56524
A	43	SER	-	cloning artifact	UNP P56524
A	44	SER	-	cloning artifact	UNP P56524
A	45	HIS	-	expression tag	UNP P56524
A	46	HIS	-	expression tag	UNP P56524
A	47	HIS	-	expression tag	UNP P56524
A	48	HIS	-	expression tag	UNP P56524
A	49	HIS	-	expression tag	UNP P56524
A	50	HIS	-	expression tag	UNP P56524
A	51	SER	-	cloning artifact	UNP P56524
A	52	SER	-	cloning artifact	UNP P56524
A	53	GLY	-	cloning artifact	UNP P56524
A	54	LEU	-	cloning artifact	UNP P56524
A	55	VAL	-	cloning artifact	UNP P56524
A	56	PRO	-	cloning artifact	UNP P56524
A	57	ARG	-	cloning artifact	UNP P56524
A	58	GLY	-	cloning artifact	UNP P56524
A	59	SER	-	cloning artifact	UNP P56524
A	60	HIS	-	cloning artifact	UNP P56524
A	61	MET	-	cloning artifact	UNP P56524
A	97	PHE	HIS	engineered mutation	UNP P56524

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	42	GLY	-	cloning artifact	UNP P56524
B	43	SER	-	cloning artifact	UNP P56524
B	44	SER	-	cloning artifact	UNP P56524
B	45	HIS	-	expression tag	UNP P56524
B	46	HIS	-	expression tag	UNP P56524
B	47	HIS	-	expression tag	UNP P56524
B	48	HIS	-	expression tag	UNP P56524
B	49	HIS	-	expression tag	UNP P56524
B	50	HIS	-	expression tag	UNP P56524
B	51	SER	-	cloning artifact	UNP P56524
B	52	SER	-	cloning artifact	UNP P56524
B	53	GLY	-	cloning artifact	UNP P56524
B	54	LEU	-	cloning artifact	UNP P56524
B	55	VAL	-	cloning artifact	UNP P56524
B	56	PRO	-	cloning artifact	UNP P56524
B	57	ARG	-	cloning artifact	UNP P56524
B	58	GLY	-	cloning artifact	UNP P56524
B	59	SER	-	cloning artifact	UNP P56524
B	60	HIS	-	cloning artifact	UNP P56524
B	61	MET	-	cloning artifact	UNP P56524
B	97	PHE	HIS	engineered mutation	UNP P56524
C	42	GLY	-	cloning artifact	UNP P56524
C	43	SER	-	cloning artifact	UNP P56524
C	44	SER	-	cloning artifact	UNP P56524
C	45	HIS	-	expression tag	UNP P56524
C	46	HIS	-	expression tag	UNP P56524
C	47	HIS	-	expression tag	UNP P56524
C	48	HIS	-	expression tag	UNP P56524
C	49	HIS	-	expression tag	UNP P56524
C	50	HIS	-	expression tag	UNP P56524
C	51	SER	-	cloning artifact	UNP P56524
C	52	SER	-	cloning artifact	UNP P56524
C	53	GLY	-	cloning artifact	UNP P56524
C	54	LEU	-	cloning artifact	UNP P56524
C	55	VAL	-	cloning artifact	UNP P56524
C	56	PRO	-	cloning artifact	UNP P56524
C	57	ARG	-	cloning artifact	UNP P56524
C	58	GLY	-	cloning artifact	UNP P56524
C	59	SER	-	cloning artifact	UNP P56524
C	60	HIS	-	cloning artifact	UNP P56524
C	61	MET	-	cloning artifact	UNP P56524
C	97	PHE	HIS	engineered mutation	UNP P56524

Continued on next page...

Continued from previous page...

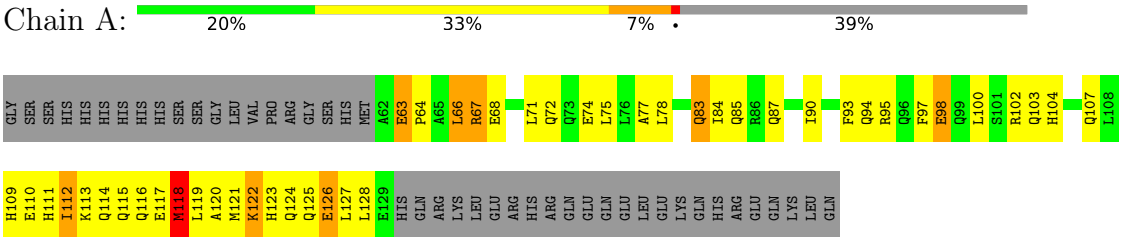
Chain	Residue	Modelled	Actual	Comment	Reference
D	42	GLY	-	cloning artifact	UNP P56524
D	43	SER	-	cloning artifact	UNP P56524
D	44	SER	-	cloning artifact	UNP P56524
D	45	HIS	-	expression tag	UNP P56524
D	46	HIS	-	expression tag	UNP P56524
D	47	HIS	-	expression tag	UNP P56524
D	48	HIS	-	expression tag	UNP P56524
D	49	HIS	-	expression tag	UNP P56524
D	50	HIS	-	expression tag	UNP P56524
D	51	SER	-	cloning artifact	UNP P56524
D	52	SER	-	cloning artifact	UNP P56524
D	53	GLY	-	cloning artifact	UNP P56524
D	54	LEU	-	cloning artifact	UNP P56524
D	55	VAL	-	cloning artifact	UNP P56524
D	56	PRO	-	cloning artifact	UNP P56524
D	57	ARG	-	cloning artifact	UNP P56524
D	58	GLY	-	cloning artifact	UNP P56524
D	59	SER	-	cloning artifact	UNP P56524
D	60	HIS	-	cloning artifact	UNP P56524
D	61	MET	-	cloning artifact	UNP P56524
D	97	PHE	HIS	engineered mutation	UNP P56524

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.

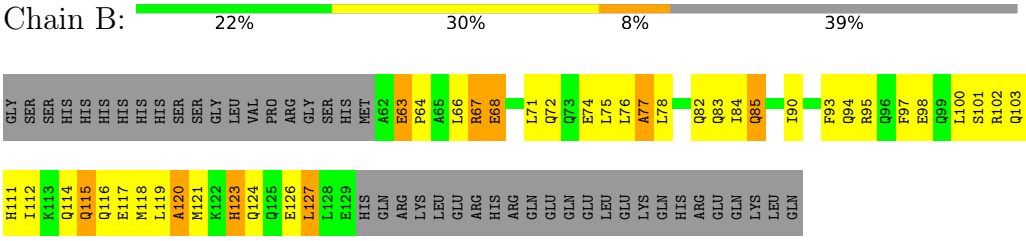
• Molecule 1: Histone deacetylase 4

Chain A:



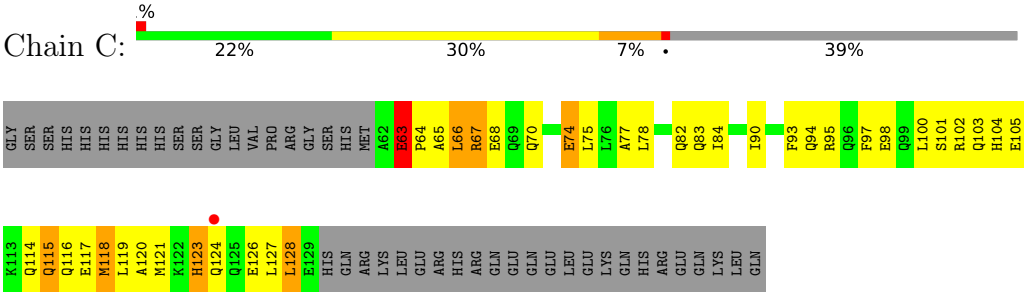
• Molecule 1: Histone deacetylase 4

Chain B:



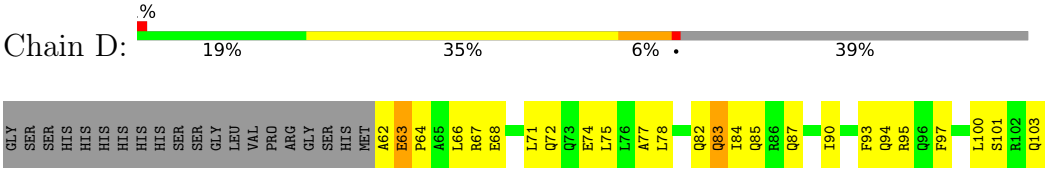
• Molecule 1: Histone deacetylase 4

Chain C:



• Molecule 1: Histone deacetylase 4

Chain D:



E110
H111
I112
K113
Q114
Q115
Q116
E117
M118
L119
A120
M121
K122
H123
Q124
Q125
E126
L127
L128
E129
HIS
GLN
ARG
LYS
LEU
GLU
ARG
HIS
ARG
GLN
GLU
GLN
GLU
LEU
GLU
LYS
GLN
HIS
ARG
GLU
GLN
LYS
LEU
GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	189.88Å 61.02Å 60.78Å 90.00° 108.61° 90.00°	Depositor
Resolution (Å)	30.00 – 3.00 30.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	75.1 (30.00-3.00) 85.8 (30.00-3.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 2.95Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.305 , 0.327 0.313 , 0.341	Depositor DCC
R_{free} test set	1348 reflections (9.79%)	wwPDB-VP
Wilson B-factor (Å^2)	81.3	Xtriage
Anisotropy	0.460	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 186.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.37$	Xtriage
Estimated twinning fraction	0.460 for -h-k-l,l,k 0.469 for -h+k-l,-l,-k 0.468 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2328	wwPDB-VP
Average B, all atoms (Å^2)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.99% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.53	0/588	1.03	4/784 (0.5%)
1	B	0.52	0/588	1.07	4/784 (0.5%)
1	C	0.53	0/588	1.08	4/784 (0.5%)
1	D	0.52	0/588	0.97	1/784 (0.1%)
All	All	0.52	0/2352	1.04	13/3136 (0.4%)

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	115	GLN	N-CA-C	-9.12	102.17	113.20
1	B	115	GLN	N-CA-C	-9.07	102.21	113.18
1	C	84	ILE	N-CA-C	-6.19	104.81	110.82
1	B	68	GLU	N-CA-C	-5.73	104.72	110.97
1	A	127	LEU	N-CA-C	-5.71	105.42	112.90

There are no chirality outliers.

There are no planarity outliers.

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	582	0	586	71	0
1	B	582	0	586	68	0
1	C	582	0	586	63	0
1	D	582	0	586	82	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2328	0	2344	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 49.

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:D:121:MET:HA	1:D:124:GLN:HG2	1.28	1.11
1:A:121:MET:HA	1:A:124:GLN:HG2	1.41	1.02
1:B:121:MET:HA	1:B:124:GLN:HG2	1.51	0.93
1:A:97:PHE:HE1	1:B:97:PHE:HE1	1.20	0.89
1:A:75:LEU:HD13	1:C:111:HIS:NE2	1.91	0.86

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	66/112 (59%)	52 (79%)	12 (18%)	2 (3%)	3	19
1	B	66/112 (59%)	53 (80%)	11 (17%)	2 (3%)	3	19
1	C	66/112 (59%)	52 (79%)	13 (20%)	1 (2%)	8	35
1	D	66/112 (59%)	50 (76%)	13 (20%)	3 (4%)	2	12
All	All	264/448 (59%)	207 (78%)	49 (19%)	8 (3%)	3	19

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	120	ALA
1	D	63	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	63	GLU
1	B	63	GLU
1	C	63	GLU

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	62/103 (60%)	54 (87%)	8 (13%)	4 19
1	B	62/103 (60%)	57 (92%)	5 (8%)	11 38
1	C	62/103 (60%)	55 (89%)	7 (11%)	5 24
1	D	62/103 (60%)	55 (89%)	7 (11%)	5 24
All	All	248/412 (60%)	221 (89%)	27 (11%)	6 26

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	67	ARG
1	C	118	MET
1	D	123	HIS
1	C	109	HIS
1	C	123	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	115	GLN
1	D	114	GLN
1	C	123	HIS
1	D	83	GLN
1	B	83	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95^{th} 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	68/112 (60%)	-0.57	0	100	100	56, 102, 162, 193	0
1	B	68/112 (60%)	-0.56	0	100	100	61, 104, 168, 189	0
1	C	68/112 (60%)	-0.60	1 (1%)	72	49	56, 102, 166, 187	0
1	D	68/112 (60%)	-0.54	1 (1%)	72	49	60, 106, 160, 186	0
All	All	272/448 (60%)	-0.56	2 (0%)	84	66	56, 103, 169, 193	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	124	GLN	2.1
1	C	124	GLN	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.