



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 6, 2017 – 12:37 PM EST

PDB ID : 4HK0
Title : UCA Fab (unbound) from CH65-CH67 Lineage
Authors : Schmidt, A.G.; Harrison, S.C.
Deposited on : unknown
Resolution : 2.50 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)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

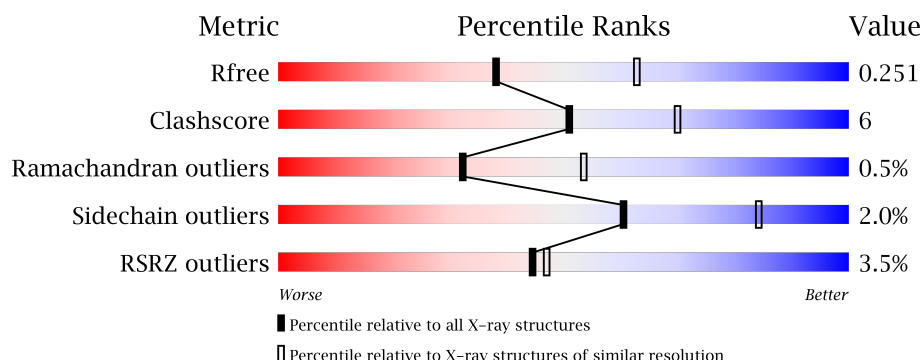
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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.

Mol	Chain	Length	Quality of chain
1	A	237	<p>9% 76% 14% • 9%</p>
1	C	237	<p>2% 79% 14% • 6%</p>
2	B	214	<p>% 83% 14% •</p>
2	D	214	<p>82% 15% ••</p>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6548 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 UCA heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	215	Total	C	N	O	S	0	0	0
			1624	1029	273	313	9			
1	C	222	Total	C	N	O	S	0	0	0
			1684	1066	283	326	9			

- Molecule 2 is a protein called UCA light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	210	Total	C	N	O	S	0	0	0
			1567	976	267	320	4			
2	D	210	Total	C	N	O	S	0	0	0
			1567	976	267	320	4			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	10	Total	O	0	0
			10	10		
3	B	27	Total	O	0	0
			27	27		
3	C	39	Total	O	0	0
			39	39		
3	D	30	Total	O	0	0
			30	30		

L182	
E185	
Q186	
W187	
Y208	
A209	
P210	
THR	
GLU	
CYS	
SER	

4 Data and refinement statistics

Property	Value	Source
Space group	P 62	Depositor
Cell constants a, b, c, α , β , γ	102.00Å 102.00Å 163.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.20 – 2.50 48.68 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (37.20-2.50) 97.4 (48.68-2.50)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.48 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.7.3_928	Depositor
R, R_{free}	0.206 , 0.258 0.201 , 0.251	Depositor DCC
R_{free} test set	1635 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	35.0	Xtriage
Anisotropy	0.489	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 33.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.049 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6548	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% 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	0.57	1/1666 (0.1%)	0.72	1/2269 (0.0%)
1	C	0.68	0/1729	0.77	2/2357 (0.1%)
2	B	0.66	0/1606	0.77	0/2197
2	D	0.70	1/1606 (0.1%)	0.81	1/2197 (0.0%)
All	All	0.65	2/6607 (0.0%)	0.77	4/9020 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	185	GLU	CG-CD	5.68	1.60	1.51
1	A	96	CYS	CB-SG	-5.23	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	19	ARG	NE-CZ-NH2	-5.78	117.41	120.30
1	C	191	LEU	CA-CB-CG	5.71	128.44	115.30
1	A	191	LEU	CA-CB-CG	5.58	128.12	115.30
1	C	151	LEU	CA-CB-CG	5.50	127.96	115.30

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	1624	0	1577	19	1
1	C	1684	0	1633	21	0
2	B	1567	0	1511	23	0
2	D	1567	0	1511	21	0
3	A	10	0	0	0	0
3	B	27	0	0	0	0
3	C	39	0	0	2	0
3	D	30	0	0	0	1
All	All	6548	0	6232	81	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 6.

The worst 5 of 81 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:169:GLN:HE21	2:D:175:ALA:HB2	1.34	0.93
1:A:35:HIS:HD2	1:A:47:TRP:HE1	1.26	0.82
1:C:153:CYS:SG	1:C:209:CYS:CB	2.68	0.80
2:B:121:PRO:HB3	2:B:208:VAL:HG11	1.64	0.79
2:D:121:PRO:HB3	2:D:208:VAL:HG11	1.65	0.78

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:42:GLY:O	3:D:302:HOH:O[3_455]	2.13	0.07

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	209/237 (88%)	203 (97%)	6 (3%)	0	100	100
1	C	218/237 (92%)	211 (97%)	7 (3%)	0	100	100
2	B	208/214 (97%)	198 (95%)	8 (4%)	2 (1%)	18	32
2	D	208/214 (97%)	199 (96%)	7 (3%)	2 (1%)	18	32
All	All	843/902 (94%)	811 (96%)	28 (3%)	4 (0%)	32	53

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	132	ALA
2	D	31	SER
2	B	94	SER
2	D	94	SER

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	178/200 (89%)	172 (97%)	6 (3%)	42	69
1	C	185/200 (92%)	178 (96%)	7 (4%)	38	64
2	B	177/181 (98%)	177 (100%)	0	100	100
2	D	177/181 (98%)	176 (99%)	1 (1%)	89	97
All	All	717/762 (94%)	703 (98%)	14 (2%)	60	84

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	72	ARG
1	C	98	ARG
1	C	153	CYS
1	A	210	ASN
1	C	140	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such

sidechains are listed below:

Mol	Chain	Res	Type
2	B	169	GLN
2	D	169	GLN
1	C	184	GLN
1	A	212	ASN
1	C	177	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 [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, 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	215/237 (90%)	0.52	21 (9%) 8 7	22, 52, 81, 123	0
1	C	222/237 (93%)	0.05	5 (2%) 61 63	16, 32, 64, 125	0
2	B	210/214 (98%)	0.02	3 (1%) 75 76	21, 40, 70, 134	0
2	D	210/214 (98%)	-0.10	1 (0%) 90 91	17, 32, 57, 91	0
All	All	857/902 (95%)	0.12	30 (3%) 44 47	16, 39, 73, 134	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	185	SER	4.0
1	A	73	ASP	3.8
1	A	227	LYS	3.6
1	A	24	ALA	3.4
1	C	218	THR	3.3

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 carbohydrates 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.