wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2024 – 02:57 PM EDT

PDB ID	:	1XTC
Title	:	CHOLERA TOXIN
Authors	:	Zhang, R.-G.; Westbrook, E.
Deposited on	:	1996-01-10
Resolution	:	2.40 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 (i)) were used in the production of this report:

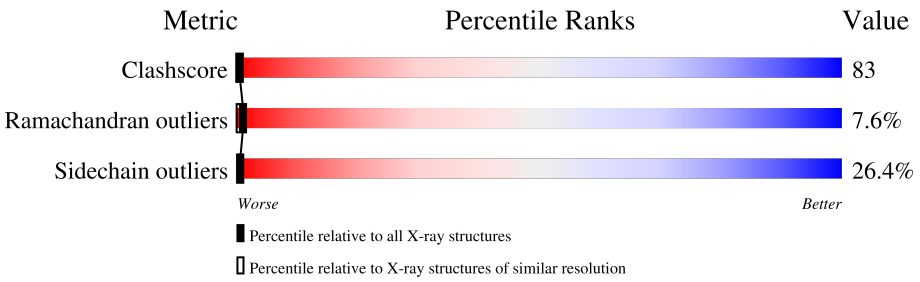
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.40 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	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	194	26% 45% 18% 10% •
2	C	46	20% 48% 24% 7% •
3	D	103	16% 52% 23% 9%
3	E	103	13% 59% 22% 6%
3	F	103	20% 47% 29% •
3	G	103	16% 61% 17% 7%
3	H	103	12% 45% 32% 12%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6135 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 CHOLERA TOXIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	192	1532	961	279	288	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	LEU	ILE	conflict	UNP P01555
A	88	LEU	ILE	conflict	UNP P01555

- Molecule 2 is a protein called CHOLERA TOXIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	45	370	227	64	78	1	0	0	0

- Molecule 3 is a protein called CHOLERA TOXIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	103	819	516	140	158	5	0	0	0
3	E	103	819	516	140	158	5	0	0	0
3	F	103	819	516	140	158	5	0	0	0
3	G	103	819	516	140	158	5	0	0	0
3	H	103	819	516	140	158	5	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	54	SER	GLY	conflict	UNP P01556
D	87	THR	VAL	conflict	UNP P01556
E	54	SER	GLY	conflict	UNP P01556
E	87	THR	VAL	conflict	UNP P01556
F	54	SER	GLY	conflict	UNP P01556
F	87	THR	VAL	conflict	UNP P01556
G	54	SER	GLY	conflict	UNP P01556
G	87	THR	VAL	conflict	UNP P01556
H	54	SER	GLY	conflict	UNP P01556
H	87	THR	VAL	conflict	UNP P01556

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	35	35	35	0	0
			Total	O		
4	C	4	4	4	0	0
			Total	O		
4	D	24	24	24	0	0
			Total	O		
4	E	19	19	19	0	0
			Total	O		
4	F	19	19	19	0	0
			Total	O		
4	G	19	19	19	0	0
			Total	O		
4	H	18	18	18	0	0

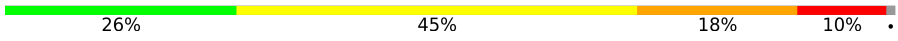
3 Residue-property plots (i)

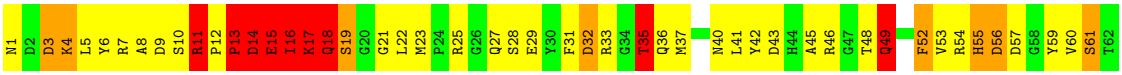
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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.

Note EDS was not executed.

• Molecule 1: CHOLERA TOXIN

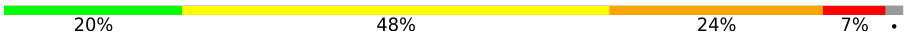
Chain A:





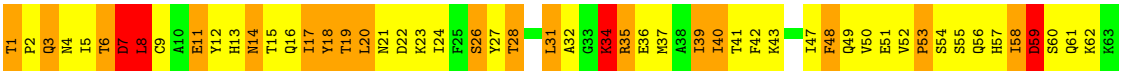
• Molecule 2: CHOLERA TOXIN

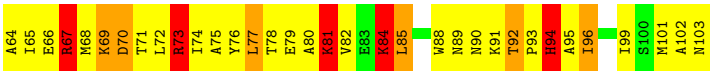
Chain C:



• Molecule 3: CHOLERA TOXIN

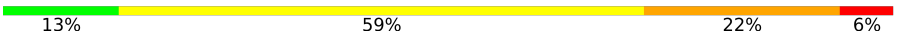
Chain D:





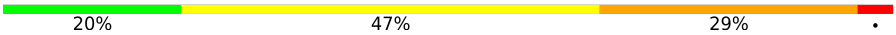
• Molecule 3: CHOLERA TOXIN

Chain E:



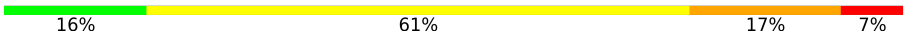
• Molecule 3: CHOLERA TOXIN

Chain F:



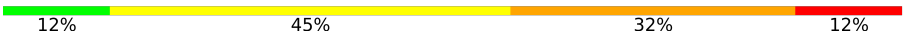
• Molecule 3: CHOLERA TOXIN

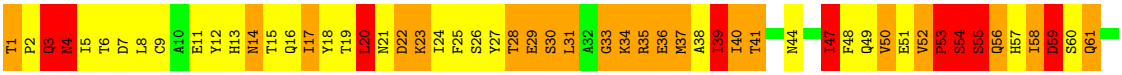
Chain G:



• Molecule 3: CHOLERA TOXIN

Chain H:





4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	73.00Å 92.20Å 60.60Å 90.00° 106.40° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.40	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.40)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROFFT, X-PLOR	Depositor
R, R_{free}	0.185 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6135	wwPDB-VP
Average B, all atoms ($Å^2$)	20.0	wwPDB-VP

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	0.97	2/1581 (0.1%)	1.87	42/2153 (2.0%)
2	C	1.04	0/374	1.51	5/498 (1.0%)
3	D	0.95	0/833	1.71	9/1125 (0.8%)
3	E	0.99	0/833	1.75	12/1125 (1.1%)
3	F	0.96	0/833	1.72	11/1125 (1.0%)
3	G	1.15	2/833 (0.2%)	1.79	22/1125 (2.0%)
3	H	0.98	1/833 (0.1%)	1.91	17/1125 (1.5%)
All	All	1.00	5/6120 (0.1%)	1.79	118/8276 (1.4%)

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	A	0	15
2	C	0	2
3	D	0	8
3	E	0	2
3	F	0	4
3	G	0	9
3	H	0	10
All	All	0	50

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	1	THR	C-N	12.81	1.58	1.34
1	A	19	SER	C-N	10.39	1.51	1.33
3	H	53	PRO	C-N	-5.56	1.21	1.34
1	A	16	ILE	C-N	-5.27	1.22	1.34
3	G	45	GLY	N-CA	-5.01	1.38	1.46

The worst 5 of 118 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	H	53	PRO	C-N-CA	19.24	169.80	121.70
3	H	54	SER	C-N-CA	18.47	167.88	121.70
1	A	15	GLU	O-C-N	-17.35	94.94	122.70
3	F	73	ARG	CD-NE-CZ	15.69	145.56	123.60
1	A	13	PRO	C-N-CA	14.48	157.89	121.70

There are no chirality outliers.

5 of 50 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	11	ARG	Sidechain
1	A	13	PRO	Peptide
1	A	14	ASP	Mainchain
1	A	16	ILE	Mainchain
1	A	17	LYS	Mainchain

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	1532	0	1419	201	2
2	C	370	0	355	93	0
3	D	819	0	823	209	0
3	E	819	0	822	226	2
3	F	819	0	823	158	0
3	G	819	0	823	117	0
3	H	819	0	821	133	0
4	A	35	0	0	6	0
4	C	4	0	0	1	0
4	D	24	0	0	6	0
4	E	19	0	0	1	0
4	F	19	0	0	1	0
4	G	19	0	0	4	0
4	H	18	0	0	0	0
All	All	6135	0	5886	983	2

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 83.

The worst 5 of 983 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:A:129:ARG:NH2	1:A:131:HIS:HB3	1.25	1.47
2:C:235:ARG:HH22	2:C:239:GLU:CB	1.35	1.37
3:F:70:ASP:O	3:F:74:ILE:HG23	1.20	1.32
2:C:235:ARG:NH2	2:C:239:GLU:CB	1.90	1.30
1:A:170:GLU:HB3	1:A:192:ARG:NH2	1.47	1.29

All (2) 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:192:ARG:O	3:E:14:ASN:ND2[1_655]	1.74	0.46
1:A:192:ARG:CD	3:E:13:HIS:NE2[1_655]	2.11	0.09

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	190/194 (98%)	149 (78%)	22 (12%)	19 (10%)	0	0
2	C	43/46 (94%)	40 (93%)	2 (5%)	1 (2%)	6	7
3	D	101/103 (98%)	81 (80%)	11 (11%)	9 (9%)	1	0
3	E	101/103 (98%)	73 (72%)	18 (18%)	10 (10%)	0	0
3	F	101/103 (98%)	87 (86%)	8 (8%)	6 (6%)	1	0
3	G	101/103 (98%)	80 (79%)	15 (15%)	6 (6%)	1	0
3	H	101/103 (98%)	83 (82%)	13 (13%)	5 (5%)	2	1
All	All	738/755 (98%)	593 (80%)	89 (12%)	56 (8%)	1	0

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	THR
1	A	49	GLN
1	A	77	LEU
1	A	111	GLN
1	A	155	ILE

5.3.2 Protein sidechains (i)

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	158/160 (99%)	123 (78%)	35 (22%)	1	1
2	C	43/44 (98%)	27 (63%)	16 (37%)	0	0
3	D	90/90 (100%)	67 (74%)	23 (26%)	0	0
3	E	90/90 (100%)	70 (78%)	20 (22%)	1	1
3	F	90/90 (100%)	65 (72%)	25 (28%)	0	0
3	G	90/90 (100%)	73 (81%)	17 (19%)	1	1
3	H	90/90 (100%)	54 (60%)	36 (40%)	0	0
All	All	651/654 (100%)	479 (74%)	172 (26%)	0	0

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	78	THR
3	H	28	THR
3	G	8	LEU
3	G	66	GLU
3	H	40	ILE

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

Mol	Chain	Res	Type
3	G	3	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	H	61	GLN
3	G	57	HIS
3	H	13	HIS
2	C	234	ASN

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 monosaccharides 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.