



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 03:18 PM GMT

PDB ID : 4HSC  
Title : Crystal structure of a cholesterol dependent cytolysin  
Authors : Feil, S.C.; Parker, M.W.  
Deposited on : 2012-10-29  
Resolution : 2.10 Å(reported)

This is a full wwPDB validation 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

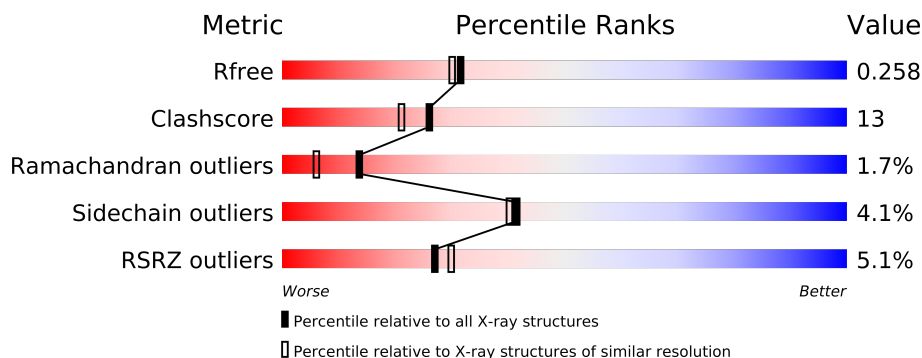
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance


The reported resolution of this entry is 2.10 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))
$R_{free}$	66092	3012 (2.10-2.10)
Clashscore	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	X	571	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3842 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Streptolysin O.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	469	Total	C	N	O	S	0	0	0
			3692	2328	627	732	5			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	X	150	Total	O	0	0
			150	150		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	46.21Å 85.34Å 81.22Å 90.00° 92.08° 90.00°	Depositor
Resolution (Å)	40.59 – 2.10 58.82 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.9 (40.59-2.10) 96.1 (58.82-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.74 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, $R_{free}$	0.222 , 0.265 0.211 , 0.258	Depositor DCC
$R_{free}$ test set	1766 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	40.2	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 45.0	EDS
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 35405 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3842	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.59% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	X	0.46	0/3763	0.60	0/5108

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	X	3692	0	3649	98	0
2	X	150	0	0	6	0
All	All	3842	0	3649	98	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 13.

All (98) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:X:170:THR:HA	1:X:172:ARG:H	0.95	1.09
1:X:170:THR:HA	1:X:172:ARG:N	1.71	1.05
1:X:210:GLY:HA3	1:X:212:LYS:H	1.29	0.98
1:X:544:GLU:HA	1:X:545:ARG:HB2	1.45	0.97

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:X:169:VAL:O	1:X:170:THR:HG22	1.65	0.95
1:X:442:LYS:HD2	2:X:609:HOH:O	1.67	0.94
1:X:210:GLY:HA3	1:X:211:ASP:HB3	1.50	0.93
1:X:440:ASN:HD22	1:X:442:LYS:HB2	1.35	0.91
1:X:544:GLU:HA	1:X:545:ARG:CB	2.00	0.91
1:X:210:GLY:CA	1:X:212:LYS:H	1.87	0.88
1:X:169:VAL:C	1:X:170:THR:HG22	1.91	0.86
1:X:210:GLY:HA3	1:X:212:LYS:N	1.95	0.81
1:X:544:GLU:CA	1:X:545:ARG:HB2	2.09	0.81
1:X:486:TYR:O	1:X:487:ASP:HB3	1.80	0.81
1:X:210:GLY:HA3	1:X:211:ASP:CB	2.09	0.80
1:X:469:HIS:CD2	1:X:471:GLY:H	2.03	0.77
1:X:487:ASP:HB2	1:X:488:ASP:HA	1.67	0.75
1:X:169:VAL:C	1:X:170:THR:CG2	2.56	0.74
1:X:135:VAL:HG13	1:X:151:ARG:NH1	2.03	0.74
1:X:264:GLU:HG3	1:X:271:SER:HB3	1.69	0.74
1:X:210:GLY:CA	1:X:211:ASP:CB	2.65	0.74
1:X:440:ASN:HD22	1:X:442:LYS:CB	2.02	0.71
1:X:357:LYS:HG2	1:X:459:GLU:HB2	1.72	0.71
1:X:469:HIS:HD2	1:X:471:GLY:H	1.39	0.69
1:X:158:THR:HG22	1:X:448:ASN:HB2	1.77	0.67
1:X:170:THR:CA	1:X:172:ARG:H	1.90	0.66
1:X:303:THR:HB	1:X:341:ASN:HD21	1.61	0.64
1:X:470:GLN:HE21	1:X:509:PRO:HB3	1.62	0.64
1:X:436:VAL:HG12	1:X:443:ILE:HD13	1.77	0.64
1:X:237:HIS:HD2	2:X:634:HOH:O	1.80	0.63
1:X:237:HIS:HE1	2:X:610:HOH:O	1.82	0.62
1:X:544:GLU:CA	1:X:545:ARG:CB	2.72	0.62
1:X:117:ASN:H	1:X:332:ASN:ND2	1.98	0.62
1:X:468:SER:HB3	1:X:556:ASN:ND2	2.16	0.61
1:X:376:THR:O	1:X:376:THR:HG22	2.03	0.59
1:X:483:GLU:CD	1:X:497:ARG:HD2	2.24	0.58
1:X:210:GLY:CA	1:X:211:ASP:HB3	2.25	0.57
1:X:301:PHE:CZ	1:X:346:ARG:HG3	2.40	0.56
1:X:497:ARG:HH11	1:X:497:ARG:HG2	1.69	0.56
1:X:497:ARG:HG2	1:X:497:ARG:NH1	2.19	0.56
1:X:301:PHE:CG	1:X:422:PHE:HB2	2.43	0.54
1:X:270:ASN:OD1	1:X:272:LYS:HB2	2.08	0.53
1:X:119:ASN:HB3	1:X:122:GLU:HB3	1.91	0.53
1:X:512:THR:HG22	1:X:513:VAL:N	2.24	0.52
1:X:512:THR:CG2	1:X:513:VAL:N	2.73	0.52
1:X:169:VAL:O	1:X:170:THR:CG2	2.50	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:X:548:LYS:HD2	1:X:550:SER:OG	2.11	0.51
1:X:181:ALA:HB3	1:X:335:PRO:HD2	1.93	0.51
1:X:209:MET:O	1:X:210:GLY:O	2.28	0.50
1:X:440:ASN:ND2	1:X:442:LYS:HB2	2.17	0.50
1:X:502:ASN:O	1:X:503:TRP:HB2	2.12	0.49
1:X:466:ASN:HB2	1:X:554:ASN:HD22	1.79	0.48
1:X:210:GLY:N	1:X:212:LYS:H	2.12	0.47
1:X:126:LYS:HD2	1:X:448:ASN:OD1	2.14	0.47
1:X:368:ALA:N	1:X:374:VAL:HG11	2.29	0.47
1:X:521:ARG:CG	1:X:522:ASN:N	2.77	0.47
1:X:167:ASP:HB2	2:X:697:HOH:O	2.14	0.47
1:X:290:LYS:HE2	1:X:353:GLU:OE2	2.14	0.46
1:X:486:TYR:O	1:X:487:ASP:CB	2.58	0.46
1:X:175:PRO:O	1:X:176:ALA:HB3	2.16	0.46
1:X:458:THR:HG22	1:X:459:GLU:N	2.31	0.46
1:X:487:ASP:CB	1:X:488:ASP:HA	2.28	0.46
1:X:544:GLU:HA	1:X:545:ARG:HB3	1.90	0.45
1:X:499:TRP:CE2	1:X:501:ASN:HB2	2.52	0.45
1:X:307:ASN:N	1:X:307:ASN:HD22	2.14	0.45
1:X:497:ARG:HH11	1:X:497:ARG:CG	2.29	0.45
1:X:474:VAL:CG1	1:X:505:SER:HB3	2.47	0.45
1:X:440:ASN:HB3	1:X:442:LYS:H	1.81	0.44
1:X:404:VAL:HG23	2:X:608:HOH:O	2.17	0.44
1:X:170:THR:OG1	1:X:170:THR:O	2.29	0.44
1:X:515:PRO:HD2	2:X:650:HOH:O	2.17	0.44
1:X:160:PRO:HD2	1:X:446:VAL:O	2.18	0.43
1:X:162:ASP:O	1:X:163:ILE:HD13	2.19	0.43
1:X:473:TYR:HB3	1:X:561:THR:HG22	2.00	0.43
1:X:264:GLU:HG3	1:X:271:SER:CB	2.44	0.43
1:X:210:GLY:CA	1:X:212:LYS:N	2.65	0.43
1:X:481:TRP:NE1	1:X:497:ARG:HD3	2.34	0.43
1:X:125:ALA:HB1	1:X:449:ARG:NH2	2.34	0.43
1:X:158:THR:HG23	1:X:160:PRO:HD3	2.02	0.42
1:X:514:ILE:HA	1:X:515:PRO:HD3	1.68	0.42
1:X:210:GLY:H	1:X:212:LYS:HB2	1.84	0.42
1:X:466:ASN:OD1	1:X:554:ASN:ND2	2.52	0.42
1:X:368:ALA:HA	1:X:374:VAL:HG11	2.01	0.42
1:X:194:VAL:O	1:X:328:LYS:HE3	2.19	0.42
1:X:379:LYS:O	1:X:383:ILE:HG13	2.19	0.42
1:X:137:LYS:O	1:X:490:GLY:HA2	2.20	0.42
1:X:521:ARG:HG2	1:X:522:ASN:HB2	2.01	0.41
1:X:508:SER:HA	1:X:509:PRO:HA	1.88	0.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:X:498:ARG:HE	1:X:498:ARG:HB2	1.57	0.41
1:X:117:ASN:H	1:X:332:ASN:HD21	1.66	0.41
1:X:483:GLU:OE2	1:X:497:ARG:HD2	2.21	0.41
1:X:484:ILE:HG21	1:X:521:ARG:CZ	2.51	0.40
1:X:499:TRP:CZ2	1:X:501:ASN:HB2	2.56	0.40
1:X:105:HIS:CE1	1:X:321:THR:HG22	2.56	0.40
1:X:146:PHE:CD1	1:X:519:ASN:HB3	2.56	0.40
1:X:129:GLU:N	1:X:129:GLU:CD	2.74	0.40
1:X:562:LEU:HD23	1:X:562:LEU:HA	1.81	0.40
1:X:548:LYS:HG3	1:X:548:LYS:O	2.20	0.40

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	X	467/571 (82%)	442 (95%)	17 (4%)	8 (2%)	14 6

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	170	THR
1	X	210	GLY
1	X	378	GLY
1	X	545	ARG
1	X	132	GLU
1	X	487	ASP
1	X	171	ASP
1	X	546	ASP

### 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	X	410/502 (82%)	393 (96%)	17 (4%)	41	40

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	129	GLU
1	X	141	LYS
1	X	164	SER
1	X	169	VAL
1	X	170	THR
1	X	211	ASP
1	X	254	GLU
1	X	385	GLU
1	X	457	SER
1	X	487	ASP
1	X	501	ASN
1	X	507	THR
1	X	508	SER
1	X	529	GLU
1	X	533	LEU
1	X	548	LYS
1	X	556	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	133	ASN
1	X	237	HIS
1	X	307	ASN
1	X	332	ASN
1	X	341	ASN
1	X	440	ASN
1	X	466	ASN
1	X	469	HIS
1	X	470	GLN
1	X	522	ASN
1	X	554	ASN
1	X	556	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates 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<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	X	469/571 (82%)	0.36	24 (5%)	27 30	22, 40, 66, 78	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	168	SER	5.4
1	X	169	VAL	5.4
1	X	545	ARG	5.0
1	X	484	ILE	4.7
1	X	518	ALA	4.4
1	X	463	GLY	3.9
1	X	379	LYS	3.7
1	X	565	TYR	3.6
1	X	570	TYR	3.4
1	X	487	ASP	3.2
1	X	551	LYS	3.0
1	X	461	THR	2.8
1	X	170	THR	2.8
1	X	378	GLY	2.7
1	X	517	GLY	2.6
1	X	458	THR	2.5
1	X	133	ASN	2.5
1	X	571	LYS	2.4
1	X	498	ARG	2.4
1	X	569	THR	2.3
1	X	374	VAL	2.3
1	X	513	VAL	2.2
1	X	144	ASP	2.1
1	X	464	LYS	2.0

## 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 carbohydrates 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.