wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 07:48 PM GMT

PDB ID	:	1HCX
Title	:	CHOLINE BINDING DOMAIN OF THE MAJOR AUTOLYSIN (C-LYTA) FROM STREPTOCOCCUS PNEUMONIAE
Authors	:	Fernandez-Tornero, C.; Lopez, R.; Garcia, E.; Gimenez-Gallego, G.; Romero, A.
Deposited on	:	2001-05-10
Resolution	:	2.60 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 (i) symbol.

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

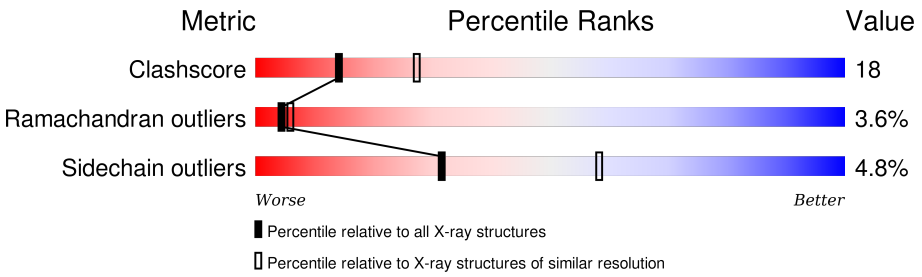
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.60 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	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)

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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	127	50% 43% 6% •
1	B	127	69% 28% •

2 Entry composition (i)

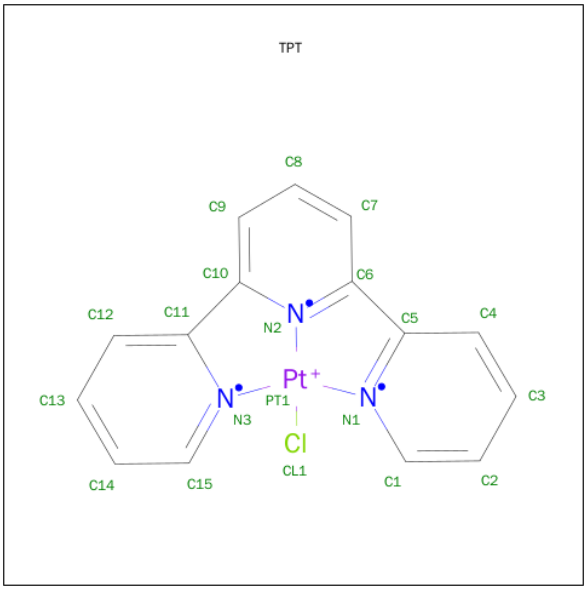
There are 5 unique types of molecules in this entry. The entry contains 2283 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 MAJOR AUTOLYSIN.

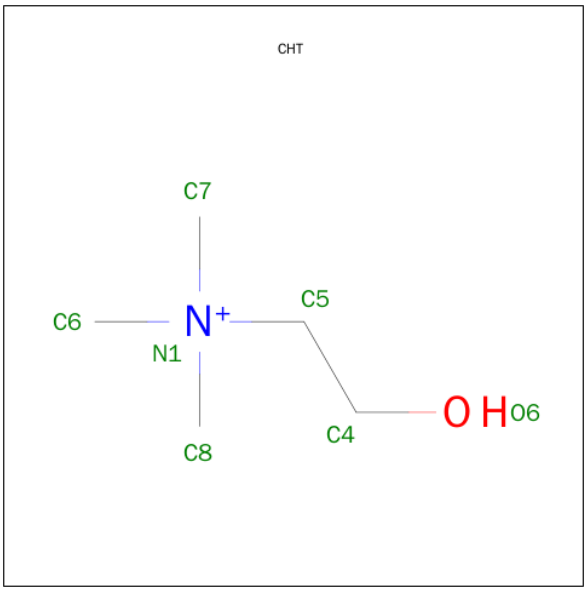
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	127	1061	692	166	199	4	0	0	0
1	B	127	1061	692	166	199	4	0	0	0

- Molecule 2 is 2,2':6',2''-TERPYRIDINE PLATINUM(II) (three-letter code: TPT) (formula: $C_{15}H_{11}ClN_3Pt$).



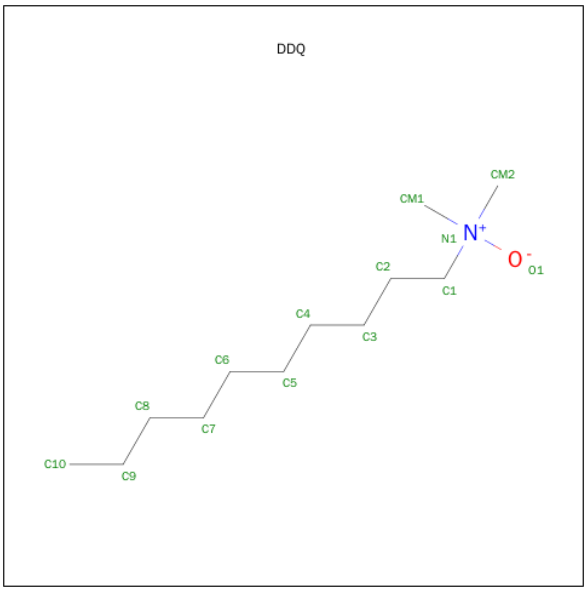
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	Pt		
2	A	1	19	15	3	1	0	0

- Molecule 3 is CHOLINE ION (three-letter code: CHT) (formula: $C_5H_{14}NO$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	7	5	1	1	0	0
			Total	C	N	O		
3	A	1	7	5	1	1	0	0
			Total	C	N	O		
3	A	1	7	5	1	1	0	0
			Total	C	N	O		
3	B	1	7	5	1	1	0	0
			Total	C	N	O		
3	B	1	7	5	1	1	0	0

- Molecule 4 is DECYLAMINE-N,N-DIMETHYL-N-OXIDE (three-letter code: DDQ) (formula: $C_{12}H_{27}NO$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			14	12	1	1		
4	B	1	Total	C	N	O	0	0
			14	12	1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	29	Total	O	0	0
			29	29		
5	B	50	Total	O	0	0
			50	50		

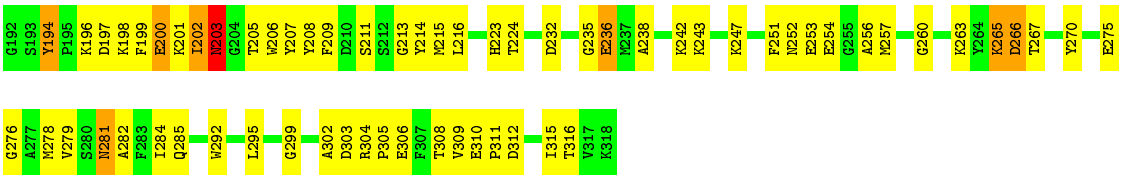
3 Residue-property plots (i)

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

Note EDS was not executed.

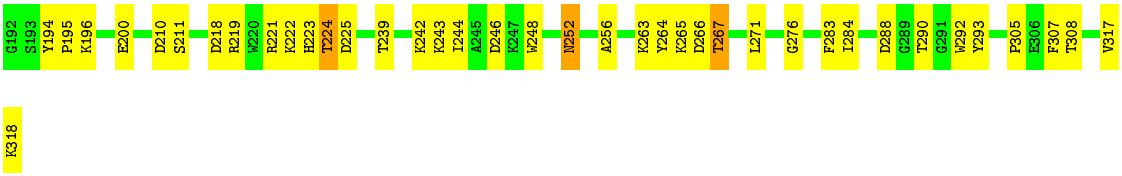
- Molecule 1: MAJOR AUTOLYSIN

Chain A:



- Molecule 1: MAJOR AUTOLYSIN

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	58.05Å 118.18Å 104.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.90 – 2.60	Depositor
% Data completeness (in resolution range)	99.0 (34.90-2.60)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	0.05	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.239 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2283	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: DDQ, TPT, CHT

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	0.41	0/1101	0.55	0/1492
1	B	0.45	0/1101	0.60	0/1492
All	All	0.43	0/2202	0.58	0/2984

There are no bond length outliers.

There are no bond angle outliers.

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	1061	0	970	48	1
1	B	1061	0	970	34	1
2	A	19	0	11	0	0
3	A	21	0	42	1	0
3	B	14	0	28	1	0
4	B	28	0	54	3	0
5	A	29	0	0	2	0
5	B	50	0	0	5	0
All	All	2283	0	2075	78	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 18.

The worst 5 of 78 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:222:LYS:H	4:B:1320:DDQ:H101	1.40	0.85
1:A:198:LYS:HG2	1:A:199:PHE:N	1.93	0.81
1:B:223:HIS:CD2	1:B:224:THR:H	2.01	0.78
1:A:197:ASP:HA	1:A:209:PHE:O	1.87	0.74
1:A:197:ASP:HB2	1:A:211:SER:HA	1.68	0.73

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:306:GLU:OE1	1:A:306:GLU:OE1[3_656]	2.07	0.13
1:B:200:GLU:OE2	1:B:200:GLU:OE2[4_555]	2.10	0.10

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	125/127 (98%)	108 (86%)	9 (7%)	8 (6%)	2 2
1	B	125/127 (98%)	117 (94%)	7 (6%)	1 (1%)	24 46
All	All	250/254 (98%)	225 (90%)	16 (6%)	9 (4%)	4 6

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	266	ASP
1	A	203	ASN
1	A	200	GLU
1	A	232	ASP
1	A	235	GLY

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	105/105 (100%)	101 (96%)	4 (4%)	40	68
1	B	105/105 (100%)	99 (94%)	6 (6%)	25	49
All	All	210/210 (100%)	200 (95%)	10 (5%)	31	58

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	224	THR
1	B	239	THR
1	B	252	ASN
1	A	281	ASN
1	B	246	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	223	HIS
1	B	285	GLN
1	B	252	ASN
1	A	281	ASN
1	B	227	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

8 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	TPT	A	1319	5	16,23,24	1.58	4 (25%)	26,35,38	0.90	1 (3%)
3	CHT	A	1320	-	6,6,6	0.92	0	8,8,8	0.53	0
3	CHT	A	1321	-	6,6,6	1.01	0	8,8,8	0.49	0
3	CHT	A	1323	-	6,6,6	0.97	0	8,8,8	0.49	0
4	DDQ	B	1319	-	13,13,13	5.21	3 (23%)	14,15,15	2.41	3 (21%)
4	DDQ	B	1320	-	13,13,13	5.13	3 (23%)	14,15,15	2.72	3 (21%)
3	CHT	B	1321	-	6,6,6	1.01	0	8,8,8	0.53	0
3	CHT	B	1322	-	6,6,6	0.90	0	8,8,8	0.49	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TPT	A	1319	5	-	0/0/20/26	0/5/5/5
3	CHT	A	1320	-	-	0/4/4/4	0/0/0/0
3	CHT	A	1321	-	-	0/4/4/4	0/0/0/0
3	CHT	A	1323	-	-	0/4/4/4	0/0/0/0
4	DDQ	B	1319	-	-	0/11/11/11	0/0/0/0
4	DDQ	B	1320	-	-	0/11/11/11	0/0/0/0
3	CHT	B	1321	-	-	0/4/4/4	0/0/0/0
3	CHT	B	1322	-	-	0/4/4/4	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1319	DDQ	O1-N1	-17.96	1.22	1.39
4	B	1320	DDQ	O1-N1	-17.75	1.22	1.39
4	B	1319	DDQ	CM2-N1	-4.46	1.42	1.49
4	B	1320	DDQ	CM2-N1	-3.84	1.43	1.49
2	A	1319	TPT	C11-C10	-2.76	1.40	1.47

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	1320	DDQ	CM2-N1-CM1	-8.45	99.30	108.83
4	B	1319	DDQ	CM2-N1-CM1	-7.47	100.40	108.83
4	B	1319	DDQ	CM2-N1-C1	2.25	117.03	109.77
4	B	1319	DDQ	O1-N1-C1	2.33	112.89	110.27
2	A	1319	TPT	C15-N3-C11	2.64	123.12	118.93

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1321	CHT	1	0
4	B	1320	DDQ	3	0
3	B	1322	CHT	1	0

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 will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.