wwPDB X-ray Structure Validation Summary Report (i)

Apr 29, 2024 – 06:30 am BST

PDB ID	:	2CMH
Title	:	Crystal Structure of Spermidine Synthase from Helicobacter Pylori
Authors	:	Sun, Y.-J.; Lu, P.-K.
Deposited on	:	2006-05-08
Resolution	:	2.30 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:

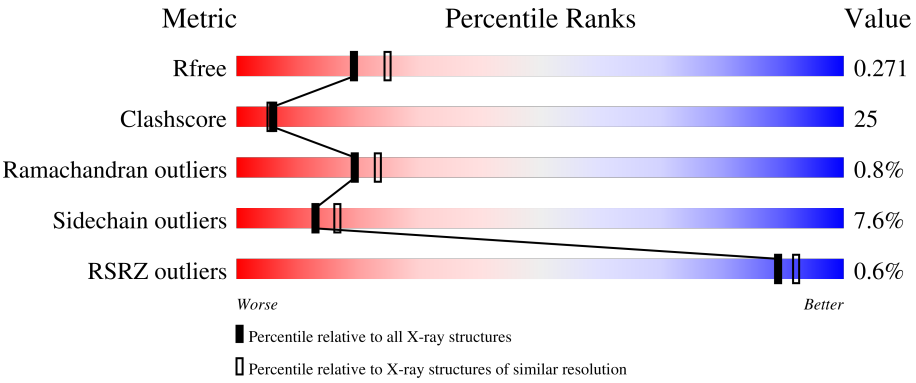
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.30 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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	262	67% 28% 5%
1	B	262	% 57% 38% 5%
1	C	262	% 52% 41% 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6913 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 SPERMIDINE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	262	Total	C	N	O	S	0	0	0
			2142	1391	359	382	10			
1	B	262	Total	C	N	O	S	0	0	0
			2155	1402	359	384	10			
1	C	262	Total	C	N	O	S	0	0	0
			2155	1402	359	384	10			

- Molecule 2 is water.

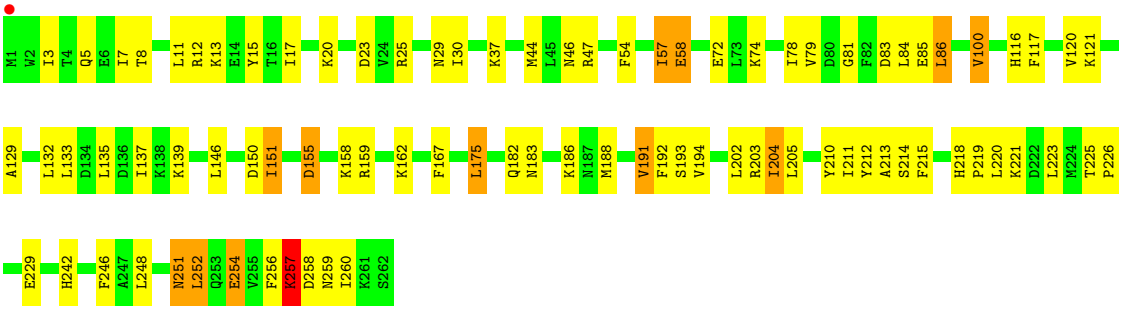
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	196	Total	O	0	0
			196	196		
2	B	126	Total	O	0	0
			126	126		
2	C	139	Total	O	0	0
			139	139		

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.

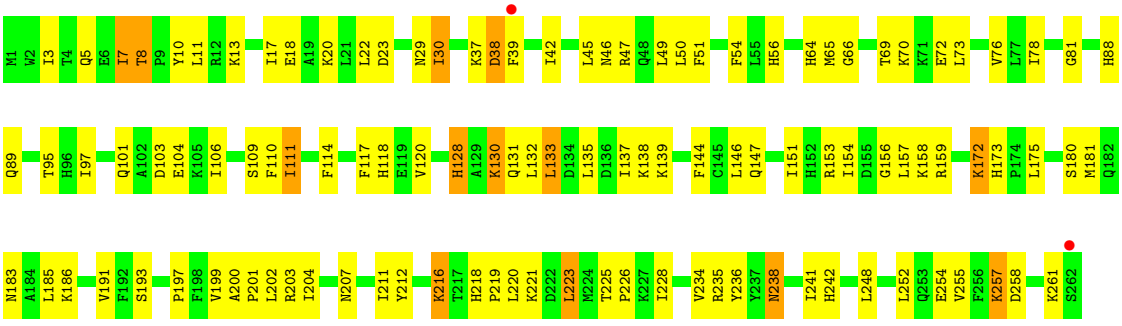
• Molecule 1: SPERMIDINE SYNTHASE

Chain A:



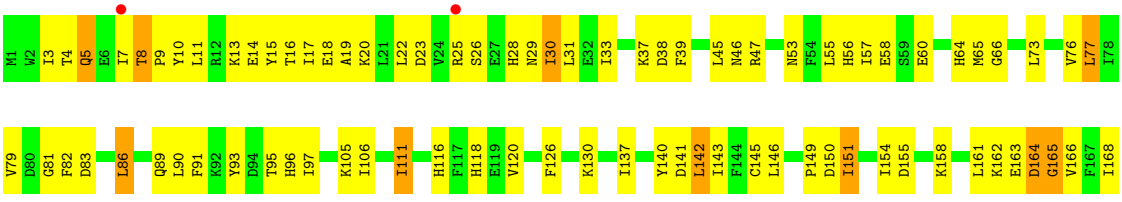
• Molecule 1: SPERMIDINE SYNTHASE

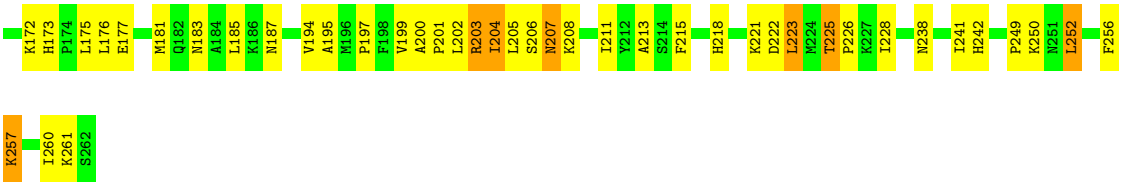
Chain B:



• Molecule 1: SPERMIDINE SYNTHASE

Chain C:




K172
H173
P174
L175
L176
E177
M181
Q182
N183
A184
L185
K186
N187
V194
A195
M196
P197
F198
V199
A200
P201
L202
R203
I204
L205
S206
N207
K208
I211
Y212
A213
S214
F215
H218
K221
D222
L223
M224
T225
P226
K227
I228
N238
I241
H242
P249
K250
N251
L252
F256
K257
I260
K261
S262

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	98.53Å 126.47Å 143.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.00 – 2.30 27.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	96.0 (27.00-2.30) 97.8 (27.00-2.30)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	13.17 (at 2.31Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.221 , 0.270 0.223 , 0.271	Depositor DCC
R_{free} test set	1972 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	-1.2	Xtriage
Anisotropy	-2.414	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 79.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.18$, $< L^2 > = 0.06$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	6913	wwPDB-VP
Average B, all atoms (Å^2)	10.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 24.06 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.1245e-03.*

[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	0.41	0/2191	0.65	0/2950
1	B	0.38	0/2205	0.64	0/2968
1	C	0.37	0/2205	0.63	0/2968
All	All	0.39	0/6601	0.64	0/8886

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

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	2142	0	2167	90	0
1	B	2155	0	2187	116	0
1	C	2155	0	2187	130	0
2	A	196	0	0	2	1
2	B	126	0	0	0	0
2	C	139	0	0	2	3
All	All	6913	0	6541	322	4

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 25.

The worst 5 of 322 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:C:81:GLY:HA3	1:C:146:LEU:HD12	1.31	1.07
1:B:133:LEU:HD13	1:B:133:LEU:H	1.18	1.06
1:C:7:ILE:HD11	1:C:11:LEU:HD23	1.44	0.98
1:A:37:LYS:HE2	1:A:37:LYS:HA	1.48	0.95
1:A:257:LYS:HG3	1:A:258:ASP:H	1.27	0.95

All (4) 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)
2:A:2063:HOH:O	2:A:2063:HOH:O[4_565]	0.92	1.28
2:C:2120:HOH:O	2:C:2120:HOH:O[3_655]	0.94	1.26
2:C:2122:HOH:O	2:C:2122:HOH:O[3_655]	1.06	1.14
2:C:2022:HOH:O	2:C:2022:HOH:O[3_555]	2.18	0.02

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	260/262 (99%)	245 (94%)	13 (5%)	2 (1%)	19	23
1	B	260/262 (99%)	247 (95%)	12 (5%)	1 (0%)	34	42
1	C	260/262 (99%)	243 (94%)	14 (5%)	3 (1%)	13	14
All	All	780/786 (99%)	735 (94%)	39 (5%)	6 (1%)	19	23

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	257	LYS
1	A	204	ILE
1	B	216	LYS
1	C	111	ILE
1	C	257	LYS

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	235/238 (99%)	221 (94%)	14 (6%)	19 26
1	B	238/238 (100%)	220 (92%)	18 (8%)	13 16
1	C	238/238 (100%)	216 (91%)	22 (9%)	9 11
All	All	711/714 (100%)	657 (92%)	54 (8%)	13 16

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	180	SER
1	C	55	LEU
1	C	207	ASN
1	B	223	LEU
1	C	5	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	183	ASN
1	C	207	ASN
1	B	53	ASN
1	B	29	ASN
1	C	218	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 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

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	262/262 (100%)	-0.28	1 (0%)	92	95	0, 5, 19, 30	0
1	B	262/262 (100%)	-0.24	2 (0%)	86	89	0, 6, 21, 30	0
1	C	262/262 (100%)	-0.07	2 (0%)	86	89	0, 9, 25, 34	0
All	All	786/786 (100%)	-0.20	5 (0%)	89	92	0, 7, 22, 34	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	7	ILE	3.4
1	C	25	ARG	2.4
1	A	1	MET	2.2
1	B	39	PHE	2.1
1	B	262	SER	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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.