W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 04:28 am BST

PDB ID	:	1H4X
Title	:	Structure of the Bacillus Cell Fate Determinant SpoIIAA in the Phosphory-lated Form
Authors	:	Seavers, P.R.; Lewis, R.J.; Brannigan, J.A.; Verschueren, K.H.G.; Murshudov, G.N.; Wilkinson, A.J.
Deposited on	:	2001-05-15
Resolution	:	1.16 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

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 1.16 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.
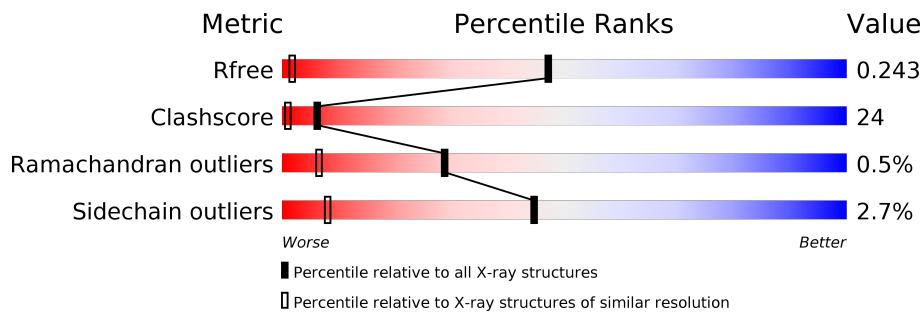
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1758 (1.20-1.12)
Clashscore	141614	1832 (1.20-1.12)
Ramachandran outliers	138981	1768 (1.20-1.12)
Sidechain outliers	138945	1768 (1.20-1.12)

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 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%

Mol	Chain	Length	Quality of chain
1	A	117	68% 26% • 5%
1	B	117	69% 22% • 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	TRS	A	1113	-	-	X	-
2	TRS	B	1113	-	X	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2273 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 ANTI-SIGMA F FACTOR ANTAGONIST.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	111	Total	C	N	O	P	S	23	10	0
			918	579	158	171	2	8			
1	B	111	Total	C	N	O	P	S	8	8	0
			913	575	158	171	1	8			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	57	SEP	SER	modified residue	PDB 1H4X
B	57	SEP	SER	modified residue	PDB 1H4X

- Molecule 2 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).
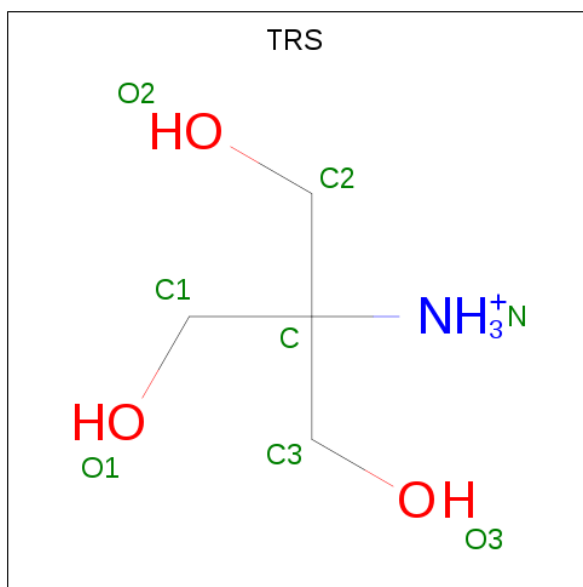
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			8	4	1	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	209	Total	O	0	0
			209	209		
3	B	217	Total	O	0	0
			217	217		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.
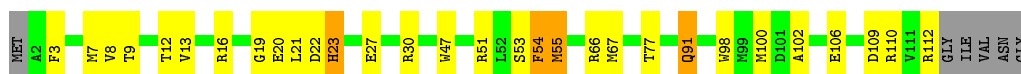
• Molecule 1: ANTI-SIGMA F FACTOR ANTAGONIST

Chain A:

• Molecule 1: ANTI-SIGMA F FACTOR ANTAGONIST

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.12Å 61.58Å 65.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.61 – 1.16 19.60 – 1.15	Depositor EDS
% Data completeness (in resolution range)	98.8 (19.61-1.16) 97.3 (19.60-1.15)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.29 (at 1.15Å)	Xtriage
Refinement program	REFMAC 5.0.36	Depositor
R, R_{free}	0.133 , 0.164 0.227 , 0.243	Depositor DCC
R_{free} test set	3652 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	9.7	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 37.9	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2273	wwPDB-VP
Average B, all atoms (Å^2)	14.0	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 25.27 % 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 3.2698e-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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: TRS, SEP

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	0.80	2/966 (0.2%)	1.08	1/1298 (0.1%)
1	B	0.87	2/961 (0.2%)	1.24	4/1291 (0.3%)
All	All	0.84	4/1927 (0.2%)	1.16	5/2589 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	47	TRP	NE1-CE2	8.74	1.49	1.37
1	A	47	TRP	NE1-CE2	8.73	1.49	1.37
1	A	98	TRP	NE1-CE2	8.70	1.48	1.37
1	B	98	TRP	NE1-CE2	8.68	1.48	1.37

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	55[A]	MET	N-CA-CB	-12.64	87.84	110.60
1	B	55[B]	MET	N-CA-CB	-12.64	87.84	110.60
1	B	23	HIS	O-C-N	-6.74	111.91	122.70
1	A	23	HIS	O-C-N	-6.52	112.27	122.70
1	B	19	GLY	O-C-N	-5.23	114.34	122.70

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	918	0	930	35	0
1	B	913	0	923	49	0
2	A	8	0	12	6	0
2	B	8	0	9	16	0
3	A	209	0	0	10	2
3	B	217	0	0	9	3
All	All	2273	0	1874	90	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 24.

The worst 5 of 90 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:7[B]:MET:SD	1:B:13[B]:VAL:HG12	1.55	1.47
1:B:7[B]:MET:SD	1:B:13[B]:VAL:CG1	2.06	1.42
1:B:30[A]:ARG:HE	2:B:1113:TRS:C3	1.50	1.25
1:B:21:LEU:HD23	1:B:55[B]:MET:CE	1.72	1.20
1:B:21:LEU:HD23	1:B:55[B]:MET:HE2	1.21	1.16

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)
3:B:2045:HOH:O	3:B:2213:HOH:O[4_555]	1.53	0.67
3:B:2196:HOH:O	3:B:2212:HOH:O[4_555]	1.96	0.24
3:A:2061:HOH:O	3:B:2062:HOH:O[2_555]	2.03	0.17
3:A:2009:HOH:O	3:A:2143:HOH:O[4_556]	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	118/117 (101%)	115 (98%)	3 (2%)	0	100	100
1	B	117/117 (100%)	115 (98%)	1 (1%)	1 (1%)	17	2
All	All	235/234 (100%)	230 (98%)	4 (2%)	1 (0%)	29	10

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	54	PHE

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	103/97 (106%)	102 (99%)	1 (1%)	76	45
1	B	102/97 (105%)	98 (96%)	4 (4%)	32	3
All	All	205/194 (106%)	200 (98%)	5 (2%)	44	11

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	ARG
1	B	20	GLU
1	B	66	ARG
1	B	91	GLN
1	B	110	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	23	HIS
1	B	91	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues 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	#$|Z| > 2$
1	SEP	A	57[A]	-	8,9,10	2.80	1 (12%)	8,12,14	6.74	5 (62%)
1	SEP	A	57[B]	-	8,9,10	0.94	0	8,12,14	1.63	3 (37%)
1	SEP	B	57	1	8,9,10	1.25	1 (12%)	8,12,14	1.43	2 (25%)

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
1	SEP	A	57[A]	-	-	0/5/8/10	-
1	SEP	A	57[B]	-	-	2/5/8/10	-
1	SEP	B	57	1	-	0/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	57[A]	SEP	OG-CB	-7.37	1.16	1.44
1	B	57	SEP	P-O1P	2.61	1.59	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	57[A]	SEP	OG-CB-CA	14.75	122.50	108.14
1	A	57[A]	SEP	P-OG-CB	-7.59	97.39	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	57[A]	SEP	OG-P-O1P	7.28	126.90	106.47
1	A	57[A]	SEP	O2P-P-OG	-5.34	92.52	106.73
1	B	57	SEP	O3P-P-O2P	2.77	118.22	107.64

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	57[B]	SEP	CB-OG-P-O2P
1	A	57[B]	SEP	N-CA-CB-OG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	57[B]	SEP	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

2 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	TRS	B	1113	-	7,7,7	2.97	2 (28%)	9,9,9	4.77	4 (44%)
2	TRS	A	1113	-	7,7,7	0.60	0	9,9,9	1.54	1 (11%)

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	TRS	B	1113	-	-	9/9/9/9	-
2	TRS	A	1113	-	-	5/9/9/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1113	TRS	C2-C	-7.16	1.30	1.53
2	B	1113	TRS	O2-C2	-2.99	1.32	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1113	TRS	O2-C2-C	9.34	140.60	111.00
2	B	1113	TRS	C2-C-N	8.14	132.27	107.98
2	B	1113	TRS	C2-C-C1	-5.82	92.77	110.81
2	B	1113	TRS	C3-C-C2	-3.51	99.92	110.81
2	A	1113	TRS	C3-C-N	3.05	117.08	107.98

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1113	TRS	C2-C-C1-O1
2	B	1113	TRS	C3-C-C1-O1
2	B	1113	TRS	N-C-C1-O1
2	B	1113	TRS	C1-C-C2-O2
2	B	1113	TRS	C3-C-C2-O2

There are no ring outliers.

2 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1113	TRS	16	0
2	A	1113	TRS	6	0

5.7 Other polymers ⓘ

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.