

Full wwPDB X-ray Structure Validation Report (i)

Mar 9, 2018 – 01:08 am GMT

PDB ID : 5I4J

Title : Dps4 from Nostoc punctiforme in complex with Zn ions

Authors: Howe, C.; Moparthi, V.K.; Persson, K.; Stensjo, K.

Deposited on : 2016-02-12

Resolution : 2.39 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references (1)) were used in the production of this report:

MolProbity : 4.02b-467

Mogul : 1.7.3 (157068), CSD as539be (2018)

Xtriage (Phenix) : 1.13

EDS: trunk30967

Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)

Refmac : 5.8.0158

CCP4 : 7.0 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

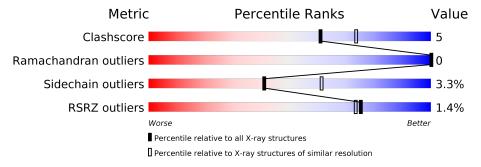
Validation Pipeline (wwPDB-VP) : trunk30967

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 2.39 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\text{Entries}) \end{array}$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries,\ resolution\ range(\mathring{A})}) \end{array}$
Clashscore	122126	3956 (2.40-2.40)
Ramachandran outliers	120053	3897 (2.40-2.40)
Sidechain outliers	120020	3898 (2.40-2.40)
RSRZ outliers	108989	3386 (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 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.

Mol	Chain	Length	Quality of chain		
1	A	177	89%	10%	- .
1	В	177	86%	12%	
1	С	177	85%	14%	
1	D	177	82%	17%	•



2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5742 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 Ferritin, Dps family protein.

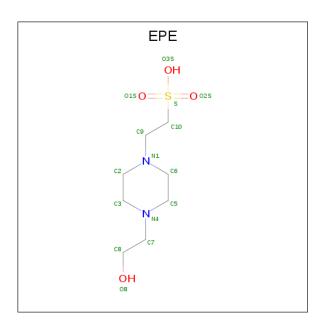
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Λ	177	Total	С	N	О	S	0	0	0
1	A	111	1414	899	241	272	2	0	U	U
1	В	177	Total	С	N	О	S	0	0	0
1	Б	111	1414	899	241	272	2	0	U	U
1	С	177	Total	С	N	О	S	0	0	0
1		111	1414	899	241	272	2	0	U	U
1	D	177	Total	С	N	О	S	0	0	0
1	ש	111	1414	899	241	272	2	U	U	U

• Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Zn 2 2	0	0
2	С	2	Total Zn 2 2	0	0

• Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf		
9	٨	1	Total	С	N	О	S	0	0
)	A	1	15	8	2	4	1	0	0

• Molecule 4 is water.

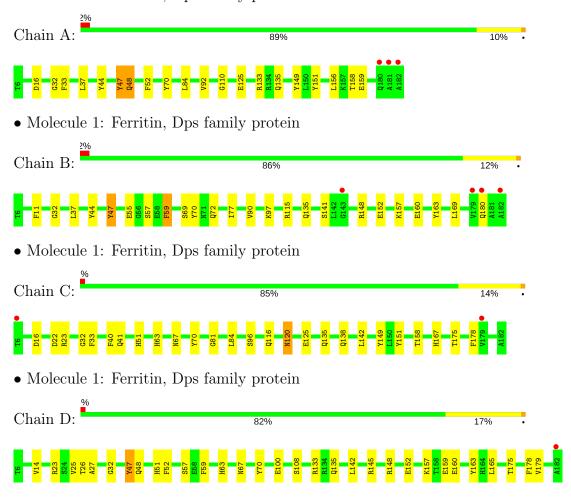
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	24	Total O 24 24	0	0
4	В	22	Total O 22 22	0	0
4	С	8	Total O 8 8	0	0
4	D	13	Total O 13 13	0	0



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 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: Ferritin, Dps family protein





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants	100.92Å 100.92Å 144.27Å	Donositon
a, b, c, α , β , γ	90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.63 - 2.39	Depositor
Resolution (A)	48.09 - 2.39	EDS
% Data completeness	99.8 (47.63-2.39)	Depositor
(in resolution range)	91.5 (48.09-2.39)	EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	0.93 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
D D	0.169 , 0.214	Depositor
R, R_{free}	0.175 , (Not available)	DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	44.1	Xtriage
Anisotropy	0.494	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.31 , 36.3	EDS
L-test for twinning ²	$< L > = 0.47, < L^2> = 0.30$	Xtriage
Estimated twinning fraction	0.069 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5742	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.14% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: ZN, EPE

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
IVIOI	Moi Chain		RMSZ $ \# Z > 5$		# Z > 5
1	A	0.43	0/1443	0.53	0/1953
1	В	0.45	0/1443	0.56	0/1953
1	С	0.42	0/1443	0.57	0/1953
1	D	0.40	0/1443	0.53	0/1953
All	All	0.43	0/5772	0.55	0/7812

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	1414	0	1364	13	0
1	В	1414	0	1364	13	0
1	С	1414	0	1364	18	0
1	D	1414	0	1364	17	0
2	A	2	0	0	0	0
2	С	2	0	0	0	0
3	A	15	0	17	1	0
4	A	24	0	0	0	0
4	В	22	0	0	0	0



Continued from previous page...

Mol	Chain	Non-H	H(model)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
4	С	8	0	0	0	0
4	D	13	0	0	1	0
All	All	5742	0	5473	55	0

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 5.

All (55) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	${ m distance}({ m \AA})$	overlap (Å)
1:C:116:GLN:O	1:C:120:ASN:ND2	2.03	0.92
1:D:14:VAL:HG21	1:D:27:ALA:HA	1.71	0.73
1:B:47:TYR:HB3	1:B:70:TYR:HB3	1.72	0.71
1:A:48:GLN:HE22	1:B:77:ILE:HG22	1.57	0.70
1:C:120:ASN:HD22	1:C:120:ASN:H	1.38	0.69
1:B:115:ARG:NH2	1:B:169:LEU:O	2.34	0.60
1:A:48:GLN:HE22	1:B:77:ILE:CG2	2.15	0.58
1:D:25:VAL:HA	1:D:142:LEU:HD13	1.86	0.56
1:D:133:ARG:NH2	1:D:159:GLU:OE2	2.38	0.55
1:C:67:ASN:HA	1:C:70:TYR:CE2	2.43	0.54
1:A:133:ARG:NH2	1:A:159:GLU:OE2	2.39	0.53
1:A:47:TYR:HB3	1:A:70:TYR:HB3	1.91	0.53
1:C:32:GLY:HA3	1:C:135:GLN:OE1	2.12	0.50
1:A:33:PHE:HB3	1:A:84:LEU:HD12	1.94	0.50
1:A:32:GLY:HA3	1:A:135:GLN:OE1	2.12	0.49
1:C:41:GLN:HE22	1:D:48:GLN:NE2	2.09	0.49
1:A:110:GLY:HA2	1:B:90:VAL:HG11	1.94	0.49
1:D:51:HIS:CE1	1:D:67:ASN:HB2	2.47	0.48
1:B:32:GLY:HA3	1:B:135:GLN:NE2	2.28	0.47
1:B:37:LEU:CD1	1:B:77:ILE:HG23	2.43	0.47
1:C:120:ASN:HD22	1:C:120:ASN:N	2.09	0.47
1:D:47:TYR:HB3	1:D:70:TYR:HB3	1.97	0.47
1:D:175:THR:O	1:D:179:VAL:HG22	2.15	0.46
1:D:148:ARG:HE	1:D:152:GLU:CD	2.19	0.46
1:A:37:LEU:HD23	1:A:92:VAL:HG12	1.98	0.46
1:C:175:THR:HA	1:C:178:PHE:CE1	2.50	0.46
1:B:55:GLU:HA	1:B:59:PHE:HB2	1.98	0.46
1:C:51:HIS:CE1	1:C:67:ASN:HB2	2.51	0.45
1:D:47:TYR:OH	1:D:165:LEU:HD11	2.16	0.45
1:C:125:GLU:OE2	1:C:158:THR:HG23	2.16	0.45
1:D:23:ARG:HA	1:D:26:THR:OG1	2.16	0.45



Continued from previous page...

A 4 1	A + 0	Interatomic	Clash
Atom-1	Atom-2	${\rm distance}\ ({\rm \AA})$	overlap (Å)
1:A:48:GLN:NE2	1:A:52:PHE:CZ	2.84	0.45
1:D:175:THR:HA	1:D:178:PHE:CE1	2.53	0.44
1:B:72:GLN:NE2	1:B:157:LYS:HD2	2.33	0.44
1:B:148:ARG:O	1:B:152:GLU:HG3	2.17	0.43
1:D:157:LYS:HD3	1:D:160:GLU:OE1	2.19	0.43
1:A:16:ASP:HA	3:A:202:EPE:O3S	2.19	0.43
1:C:135:GLN:HB3	1:C:151:TYR:CE2	2.54	0.43
1:B:47:TYR:CE2	1:B:69:SER:HB3	2.54	0.43
1:C:81:GLY:HA3	1:D:52:PHE:CZ	2.54	0.43
1:C:16:ASP:HA	1:C:23:ARG:HH12	1.83	0.42
1:C:63:HIS:HE1	4:D:310:HOH:O	2.02	0.42
1:C:120:ASN:ND2	1:C:120:ASN:N	2.65	0.42
1:C:120:ASN:H	1:C:120:ASN:ND2	2.09	0.42
1:A:125:GLU:OE2	1:A:158:THR:HG23	2.19	0.42
1:D:157:LYS:HD3	1:D:157:LYS:HA	1.72	0.42
1:D:32:GLY:HA3	1:D:135:GLN:OE1	2.20	0.41
1:A:156:LEU:HD12	1:A:156:LEU:HA	1.93	0.41
1:C:138:GLN:O	1:C:142:LEU:HG	2.20	0.41
1:D:59:PHE:O	1:D:63:HIS:HB2	2.21	0.41
1:C:33:PHE:HB3	1:C:84:LEU:HD12	2.03	0.41
1:B:160:GLU:HA	1:B:163:TYR:CD2	2.56	0.41
1:A:149:TYR:CE1	1:C:149:TYR:HE1	2.39	0.41
1:D:160:GLU:HA	1:D:163:TYR:CD2	2.56	0.40
1:B:11:PHE:CE1	1:B:97:LYS:HE2	2.57	0.40

There are no symmetry-related clashes.

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	Perce	ntiles
1	A	175/177 (99%)	168 (96%)	7 (4%)	0	100	100



Continued	trom	mromonie	maaa
-	110111	DICULUUS	pauc

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	ntiles
1	В	175/177 (99%)	170 (97%)	5 (3%)	0	100	100
1	С	175/177 (99%)	170 (97%)	5 (3%)	0	100	100
1	D	175/177 (99%)	171 (98%)	4 (2%)	0	100	100
All	All	700/708 (99%)	679 (97%)	21 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	150/150 (100%)	146 (97%)	4 (3%)	48	68	
1	В	150/150~(100%)	144 (96%)	6 (4%)	34	53	
1	С	150/150 (100%)	145 (97%)	5 (3%)	41	61	
1	D	150/150 (100%)	145 (97%)	5 (3%)	41	61	
All	All	600/600 (100%)	580 (97%)	20 (3%)	41	61	

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	TYR
1	A	47	TYR
1	A	48	GLN
1	A	151	TYR
1	В	44	TYR
1	В	47	TYR
1	В	57	SER
1	В	59	PHE
1	В	141	SER
1	В	180	GLN
1	С	22	ASP
1	С	40	PHE
1	С	96	SER



Continued from previous page...

Mol	Chain	Res	Type
1	С	120	ASN
1	С	167	HIS
1	D	47	TYR
1	D	57	SER
1	D	100	GLU
1	D	108	SER
1	D	145	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	48	GLN
1	A	67	ASN
1	В	41	GLN
1	В	135	GLN
1	С	41	GLN
1	С	116	GLN
1	С	120	ASN
1	С	138	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)

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)

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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	fol Type Chain Res Link		Link	Bo	ond leng	ths	Bond angles			
	IVIOI	туре	Chain	rtes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
Ī	3	EPE	A	202	-	15,15,15	1.01	1 (6%)	18,20,20	2.05	5 (27%)

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
3	EPE	A	202	-	-	0/9/19/19	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	$\operatorname{Ideal}(ext{\AA})$
3	A	202	EPE	C10-S	3.44	1.82	1.77

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$Ideal(^{o})$
3	A	202	EPE	C7-N4-C3	3.06	119.23	111.24
3	A	202	EPE	O1S-S-C10	3.22	110.79	106.92
3	A	202	EPE	C7-N4-C5	3.41	120.14	111.24
3	A	202	EPE	O2S-S-C10	3.42	111.03	106.92
3	A	202	EPE	C5-N4-C3	4.09	117.91	108.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	202	EPE	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)

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>	$\# \mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q < 0.9
1	A	177/177 (100%)	-0.24	3 (1%) 70 67	43, 56, 73, 93	0
1	В	177/177 (100%)	-0.13	4 (2%) 60 57	42, 54, 79, 93	0
1	С	177/177 (100%)	-0.23	2 (1%) 80 78	43, 58, 79, 98	0
1	D	177/177 (100%)	-0.25	1 (0%) 89 87	44, 56, 71, 90	1 (0%)
All	All	708/708 (100%)	-0.21	10 (1%) 75 73	42, 56, 76, 98	1 (0%)

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	182	ALA	4.4
1	A	182	ALA	3.5
1	В	179	VAL	3.1
1	D	182	ALA	2.7
1	В	180	GLN	2.7
1	A	180	GLN	2.5
1	С	179	VAL	2.5
1	В	143	GLY	2.4
1	A	181	ALA	2.3
1	С	6	THR	2.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.



6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	ZN	С	201	1/1	0.87	0.08	75,75,75,75	1
3	EPE	A	202	15/15	0.90	0.21	64,69,76,77	15
2	ZN	A	203	1/1	0.91	0.16	81,81,81,81	1
2	ZN	С	202	1/1	0.93	0.10	77,77,77,77	1
2	ZN	A	201	1/1	0.98	0.06	74,74,74,74	1

6.5 Other polymers (i)

There are no such residues in this entry.

