wwPDB X-ray Structure Validation Summary Report (i)

Oct 23, 2021 – 12:42 PM EDT

PDB ID	:	1DPS
Title	:	THE CRYSTAL STRUCTURE OF DPS, A FERRITIN HOMOLOG THAT BINDS AND PROTECTS DNA
Authors	:	Grant, R.A.; Filman, D.J.; Finkel, S.E.; Kolter, R.; Hogle, J.M.
Deposited on	:	1998-02-23
Resolution	:	1.60 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
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 (i)) were used in the production of this report:

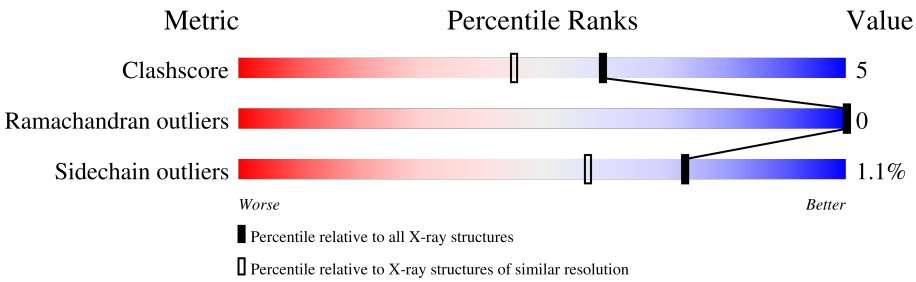
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.60 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	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	167	87% 8% 5%
1	B	167	84% 8% 8%
1	C	167	82% 10% • 8%
1	D	167	86% 8% 7%
1	E	167	84% 7% • 8%
1	F	167	86% 7% 8%
1	G	167	82% 10% • 8%
1	H	167	86% 6% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	167	83% 8% • 8%
1	J	167	83% 9% • 8%
1	K	167	80% 11% • 8%
1	L	167	83% 9% • 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 16107 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 DPS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	159	1252	785	219	244	4	0	0	0
1	B	154	1221	768	213	236	4	7	0	0
1	C	154	1221	768	213	236	4	3	0	0
1	D	156	1236	776	216	240	4	0	0	0
1	E	154	1221	768	213	236	4	2	0	0
1	F	154	1221	768	213	236	4	2	0	0
1	G	154	1221	768	213	236	4	4	0	0
1	H	154	1221	768	213	236	4	2	0	0
1	I	154	1221	768	213	236	4	5	0	0
1	J	154	1221	768	213	236	4	1	0	0
1	K	154	1221	768	213	236	4	0	0	0
1	L	154	1221	768	213	236	4	4	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	164	CYS	SER	engineered mutation	UNP P0ABT2
B	164	CYS	SER	engineered mutation	UNP P0ABT2
C	164	CYS	SER	engineered mutation	UNP P0ABT2
D	164	CYS	SER	engineered mutation	UNP P0ABT2
E	164	CYS	SER	engineered mutation	UNP P0ABT2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	164	CYS	SER	engineered mutation	UNP P0ABT2
G	164	CYS	SER	engineered mutation	UNP P0ABT2
H	164	CYS	SER	engineered mutation	UNP P0ABT2
I	164	CYS	SER	engineered mutation	UNP P0ABT2
J	164	CYS	SER	engineered mutation	UNP P0ABT2
K	164	CYS	SER	engineered mutation	UNP P0ABT2
L	164	CYS	SER	engineered mutation	UNP P0ABT2

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Na 1 1	0	0
2	B	1	Total Na 1 1	0	0
2	C	1	Total Na 1 1	0	0
2	D	1	Total Na 1 1	0	0
2	E	1	Total Na 1 1	0	0
2	F	1	Total Na 1 1	0	0
2	G	1	Total Na 1 1	0	0
2	H	1	Total Na 1 1	0	0
2	I	1	Total Na 1 1	0	0
2	J	1	Total Na 1 1	0	0
2	K	1	Total Na 1 1	0	0
2	L	1	Total Na 1 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	114	Total O 114 114	0	0
3	B	115	Total O 115 115	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	C	114	114	114	0	0
			Total	O		
3	D	110	110	110	0	0
			Total	O		
3	E	112	112	112	0	0
			Total	O		
3	F	127	127	127	0	0
			Total	O		
3	G	112	112	112	0	0
			Total	O		
3	H	125	125	125	0	0
			Total	O		
3	I	120	120	120	0	0
			Total	O		
3	J	116	116	116	0	0
			Total	O		
3	K	119	119	119	0	0
			Total	O		
3	L	113	113	113	0	0

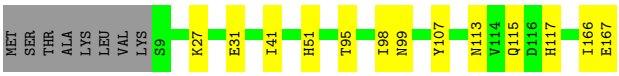
3 Residue-property plots (i)

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. 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.

Note EDS was not executed.

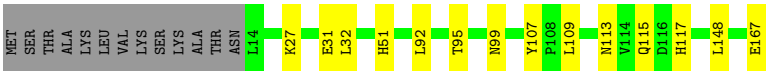
• Molecule 1: DPS

Chain A:



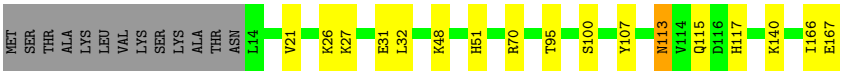
• Molecule 1: DPS

Chain B:



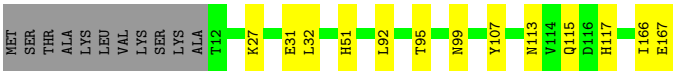
• Molecule 1: DPS

Chain C:



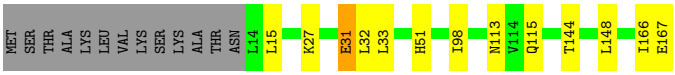
• Molecule 1: DPS

Chain D:



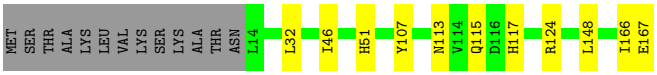
• Molecule 1: DPS

Chain E:



• Molecule 1: DPS

Chain F:



- Molecule 1: DPS

Chain G:

- Molecule 1: DPS

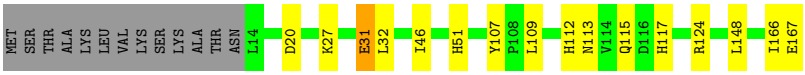
Chain H:

- Molecule 1: DPS

Chain I:

- Molecule 1: DPS

Chain J:

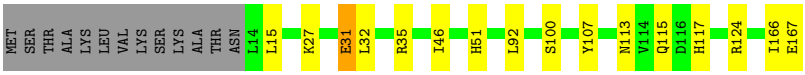


- Molecule 1: DPS

Chain K:

- Molecule 1: DPS

Chain L:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	134.41Å 139.65Å 118.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.60	Depositor
% Data completeness (in resolution range)	76.3 (20.00-1.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.188 , 0.220	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16107	wwPDB-VP
Average B, all atoms (Å^2)	18.0	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: NA

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	0.78	0/1270	0.70	0/1720
1	B	0.77	0/1239	0.71	0/1677
1	C	0.76	0/1239	0.76	0/1677
1	D	0.79	0/1254	0.74	0/1698
1	E	0.77	1/1239 (0.1%)	0.73	0/1677
1	F	0.77	0/1239	0.73	0/1677
1	G	0.81	1/1239 (0.1%)	0.73	0/1677
1	H	0.79	1/1239 (0.1%)	0.74	1/1677 (0.1%)
1	I	0.77	1/1239 (0.1%)	0.73	0/1677
1	J	0.80	2/1239 (0.2%)	0.71	0/1677
1	K	0.78	1/1239 (0.1%)	0.73	0/1677
1	L	0.78	2/1239 (0.2%)	0.72	1/1677 (0.1%)
All	All	0.78	9/14914 (0.1%)	0.73	2/20188 (0.0%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	31	GLU	CG-CD	6.63	1.61	1.51
1	I	31	GLU	CG-CD	6.42	1.61	1.51
1	H	31	GLU	CD-OE2	6.00	1.32	1.25
1	G	31	GLU	CG-CD	5.83	1.60	1.51
1	J	31	GLU	CD-OE2	5.82	1.32	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	L	35	ARG	NE-CZ-NH1	5.83	123.22	120.30
1	H	35	ARG	NE-CZ-NH1	5.53	123.06	120.30

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	1252	0	1244	12	0
1	B	1221	0	1219	12	0
1	C	1221	0	1219	14	0
1	D	1236	0	1232	15	0
1	E	1221	0	1219	10	0
1	F	1221	0	1219	7	0
1	G	1221	0	1219	12	0
1	H	1221	0	1219	8	0
1	I	1221	0	1219	14	0
1	J	1221	0	1219	13	0
1	K	1221	0	1219	17	0
1	L	1221	0	1219	13	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
3	A	114	0	0	2	0
3	B	115	0	0	1	0
3	C	114	0	0	0	0
3	D	110	0	0	2	0
3	E	112	0	0	0	0
3	F	127	0	0	0	0
3	G	112	0	0	1	0
3	H	125	0	0	0	0
3	I	120	0	0	0	0
3	J	116	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	K	119	0	0	1	0
3	L	113	0	0	0	0
All	All	16107	0	14666	135	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.

The worst 5 of 135 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:J:20:ASP:HB2	3:J:271:HOH:O	1.78	0.82
1:C:113:ASN:ND2	1:C:115:GLN:H	1.86	0.73
1:C:27:LYS:O	1:C:31:GLU:HG2	1.89	0.72
3:A:263:HOH:O	1:C:95:THR:HG23	1.88	0.72
1:C:113:ASN:HD22	1:C:115:GLN:H	1.39	0.71

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	Percentiles	
1	A	157/167 (94%)	155 (99%)	2 (1%)	0	100	100
1	B	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
1	C	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
1	D	154/167 (92%)	153 (99%)	1 (1%)	0	100	100
1	E	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
1	F	152/167 (91%)	151 (99%)	1 (1%)	0	100	100
1	G	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
1	H	152/167 (91%)	150 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	I	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
1	J	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
1	K	152/167 (91%)	151 (99%)	1 (1%)	0	100	100
1	L	152/167 (91%)	150 (99%)	2 (1%)	0	100	100
All	All	1831/2004 (91%)	1810 (99%)	21 (1%)	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	135/143 (94%)	134 (99%)	1 (1%)	84	73
1	B	132/143 (92%)	131 (99%)	1 (1%)	81	70
1	C	132/143 (92%)	129 (98%)	3 (2%)	50	25
1	D	134/143 (94%)	133 (99%)	1 (1%)	84	73
1	E	132/143 (92%)	131 (99%)	1 (1%)	81	70
1	F	132/143 (92%)	131 (99%)	1 (1%)	81	70
1	G	132/143 (92%)	130 (98%)	2 (2%)	65	44
1	H	132/143 (92%)	131 (99%)	1 (1%)	81	70
1	I	132/143 (92%)	130 (98%)	2 (2%)	65	44
1	J	132/143 (92%)	131 (99%)	1 (1%)	81	70
1	K	132/143 (92%)	130 (98%)	2 (2%)	65	44
1	L	132/143 (92%)	130 (98%)	2 (2%)	65	44
All	All	1589/1716 (93%)	1571 (99%)	18 (1%)	73	57

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	51	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	100	SER
1	L	51	HIS
1	G	51	HIS
1	J	51	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	117	HIS
1	J	115	GLN
1	I	112	HIS
1	I	117	HIS
1	K	112	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)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.