



wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 04:16 PM GMT

PDB ID : 4A0A
Title : STRUCTURE OF HSDDDB1-DRDDB2 BOUND TO A 16 BP CPD-DUPLEX
(PYRIMIDINE AT D-1 POSITION) AT 3.6 Å RESOLUTION (CPD 3)
Authors : Scrima, A.; Fischer, E.S.; Iwai, S.; Gut, H.; Thoma, N.H.
Deposited on : 2011-09-08
Resolution : 3.60 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

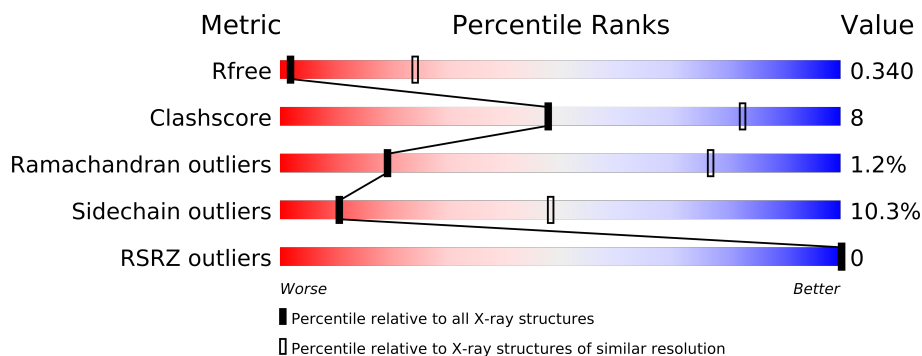
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance





The reported resolution of this entry is 3.60 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))
R_{free}	66092	1020 (3.86-3.34)
Clashscore	79885	1155 (3.80-3.40)
Ramachandran outliers	78287	1109 (3.80-3.40)
Sidechain outliers	78261	1108 (3.80-3.40)
RSRZ outliers	66119	1000 (3.84-3.36)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1159	
2	B	382	
3	C	15	
4	D	16	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9619 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA DAMAGE-BINDING PROTEIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	780	Total	C	N	O	S	0	0	0
			6129	3901	1025	1169	34			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MET	-	EXPRESSION TAG	UNP Q16531
A	-17	HIS	-	EXPRESSION TAG	UNP Q16531
A	-16	HIS	-	EXPRESSION TAG	UNP Q16531
A	-15	HIS	-	EXPRESSION TAG	UNP Q16531
A	-14	HIS	-	EXPRESSION TAG	UNP Q16531
A	-13	HIS	-	EXPRESSION TAG	UNP Q16531
A	-12	HIS	-	EXPRESSION TAG	UNP Q16531
A	-11	VAL	-	EXPRESSION TAG	UNP Q16531
A	-10	ASP	-	EXPRESSION TAG	UNP Q16531
A	-9	GLU	-	EXPRESSION TAG	UNP Q16531
A	-8	ASN	-	EXPRESSION TAG	UNP Q16531
A	-7	LEU	-	EXPRESSION TAG	UNP Q16531
A	-6	TYR	-	EXPRESSION TAG	UNP Q16531
A	-5	PHE	-	EXPRESSION TAG	UNP Q16531
A	-4	GLN	-	EXPRESSION TAG	UNP Q16531
A	-3	GLY	-	EXPRESSION TAG	UNP Q16531
A	-2	GLY	-	EXPRESSION TAG	UNP Q16531
A	-1	GLY	-	EXPRESSION TAG	UNP Q16531
A	0	ARG	-	EXPRESSION TAG	UNP Q16531
A	224	LYS	GLU	ENGINEERED MUTATION	UNP Q16531

- Molecule 2 is a protein called DNA DAMAGE-BINDING PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	355	Total	C	N	O	S	0	0	0
			2843	1806	499	527	11			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	76	MET	-	EXPRESSION TAG	UNP Q2YDS1
B	77	HIS	-	EXPRESSION TAG	UNP Q2YDS1
B	78	HIS	-	EXPRESSION TAG	UNP Q2YDS1
B	79	HIS	-	EXPRESSION TAG	UNP Q2YDS1
B	80	HIS	-	EXPRESSION TAG	UNP Q2YDS1
B	81	HIS	-	EXPRESSION TAG	UNP Q2YDS1
B	82	HIS	-	EXPRESSION TAG	UNP Q2YDS1
B	83	ARG	-	EXPRESSION TAG	UNP Q2YDS1
B	84	ARG	-	EXPRESSION TAG	UNP Q2YDS1
B	85	LEU	-	EXPRESSION TAG	UNP Q2YDS1
B	86	VAL	-	EXPRESSION TAG	UNP Q2YDS1
B	87	PRO	-	EXPRESSION TAG	UNP Q2YDS1
B	88	ARG	-	EXPRESSION TAG	UNP Q2YDS1
B	89	GLY	-	EXPRESSION TAG	UNP Q2YDS1
B	90	SER	-	EXPRESSION TAG	UNP Q2YDS1
B	91	GLY	-	EXPRESSION TAG	UNP Q2YDS1
B	92	GLY	-	EXPRESSION TAG	UNP Q2YDS1
B	93	ARG	-	EXPRESSION TAG	UNP Q2YDS1
B	180	GLN	LEU	VARIANT	UNP Q2YDS1
B	214	ARG	TRP	VARIANT	UNP Q2YDS1

- Molecule 3 is a DNA chain called 5'-D(*GP*GP*TP*GP*AP*AP*AP*(TTD)P*AP*GP*C P*AP*GP*DGP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	15	Total	C	N	O	P	0	0	0
			334	159	66	94	15			

- Molecule 4 is a DNA chain called 5'-D(*CP*CP*TP*GP*CP*TP*CP*CP*TP*TP*TP*CP *AP*CP*CP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	16	Total	C	N	O	P	0	0	0
			312	151	50	96	15			

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	1	Total	Ca	0	0
			1	1		

L388	F286	K182	MET
L399	N287	T183	H1S
K403	P288	I186	H1S
E404	R289	M189	H1S
T405	C290	D193	H1S
	D291	A194	ARG
	W292	I195	LEU
I408	M294	T196	VAL
M412	A295		PRO
H418	V299	N201	ARG
Q419	L308	Q202	GLY
M432		F203	SER
	K312	N204	GLY
L441		Q207	GLY
A442	M321	L208	ARG
S443	K325	F209	THR
M448	P326	V210	GLY
	A329	S211	GLN
W452	A330	S212	LYS
M453	Y331	I213	LYS
R454	F332	R214	VAL
E455		F222	G101
ASP			
THR	T338	D234	S117
	K339	D237	Q121
	L340	Y238	Q124
	L341	W239	Q127
	T342	V243	L134
	T343	S246	K135
	D344	R249	S136
	Q345	Q250	Y137
	R346	M251	H140
	N347	G255	S144
	S354	L263	R148
	Y355	L262	T158
	T364	L266	H159
	L365	D267	S167
	H366	G268	I172
	P367	H269	T173
	H368	F272	L174
	R369	A279	W175
	L374	K280	D178
	I377	V281	V179
	L388		G180
	V389		N183
	R392		
	Y393		
	D396		
	G397		

Chain C:

G1	G2	G3	T4	G5		T8	T9	A11	G12	C13	A14	G15	G16
----	----	----	----	----	--	----	----	-----	-----	-----	-----	-----	-----

● Molecule 4: 5'-D(*CP*CP*TP*GP*CP*TP*CP*CP*TP*TP*TP*CP*AP*CP*CP*C)-3'

Chain D:

A diagram showing a sequence of colored blocks labeled C1, C5, T6, C7, C8, A9, T10, T11, C12, A13, and C16. The blocks are connected by lines, with C1, C8, and C16 being green, and the others being yellow or orange.

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	155.64Å 227.14Å 114.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.25 – 3.60 46.25 – 3.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.25-3.60) 99.5 (46.25-3.60)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.23 (at 3.57Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.263 , 0.347 0.260 , 0.340	Depositor DCC
R_{free} test set	1188 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	71.3	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 1.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 23749 reflections	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	9619	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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

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	0.36	0/6236	0.55	1/8429 (0.0%)
2	B	0.36	0/2917	0.55	0/3962
3	C	0.61	0/331	1.00	2/509 (0.4%)
4	D	0.66	0/346	1.14	1/529 (0.2%)
All	All	0.39	0/9830	0.60	4/13429 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	9	DA	P-O3'-C3'	5.89	126.77	119.70
1	A	367	LEU	CA-CB-CG	5.49	127.93	115.30
3	C	15	DG	P-O3'-C3'	5.29	126.04	119.70
3	C	12	DG	P-O3'-C3'	5.14	125.87	119.70

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6129	0	6105	90	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	2843	0	2788	55	0
3	C	334	0	182	7	0
4	D	312	0	182	4	0
5	C	1	0	0	0	0
All	All	9619	0	9257	149	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

The worst 5 of 149 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:928:ARG:HH11	1:A:928:ARG:HG2	1.40	0.84
1:A:969:GLU:HG2	1:A:970:ASN:N	1.99	0.78
1:A:369:ARG:HG2	1:A:370:GLN:N	1.98	0.78
1:A:38:ARG:HH12	1:A:54:GLU:HB3	1.48	0.77
1:A:127:GLU:HB2	1:A:129:ARG:HD2	1.66	0.77

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	764/1159 (66%)	681 (89%)	73 (10%)	10 (1%)	18	75
2	B	353/382 (92%)	314 (89%)	36 (10%)	3 (1%)	27	83
All	All	1117/1541 (72%)	995 (89%)	109 (10%)	13 (1%)	19	77

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	962	ASP
1	A	112	ILE
1	A	1110	ALA
2	B	189	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	70	LYS

5.3.2 Protein sidechains ⓘ

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	679/1015 (67%)	617 (91%)	62 (9%)	14	55
2	B	313/335 (93%)	273 (87%)	40 (13%)	6	35
All	All	992/1350 (74%)	890 (90%)	102 (10%)	10	48

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	966	LEU
1	A	1082	THR
2	B	393	TYR
1	A	969	GLU
1	A	1000	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	941	ASN
1	A	990	GLN
2	B	418	HIS
1	A	964	ASN
1	A	1015	GLN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z > 2$
3	TTD	C	9	3	43,45,46	2.94	10 (23%)	71,74,77	2.15	14 (19%)

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	TTD	C	9	3	-	1/23/109/110	0/0/6/6

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	9	TTD	P-OP1	15.88	1.64	1.46
3	C	9	TTD	C1R-N1T	4.48	1.51	1.45
3	C	9	TTD	C1'-N1	4.34	1.51	1.45
3	C	9	TTD	C5'-C4R	3.48	1.62	1.51
3	C	9	TTD	C2-N1	3.39	1.43	1.36

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	9	TTD	O4R-C1R-N1T	7.90	118.62	108.76
3	C	9	TTD	C5T-C4T-N3T	6.38	121.79	116.07
3	C	9	TTD	C5-C4-N3	5.89	121.35	116.07
3	C	9	TTD	N3-C2-N1	5.86	122.99	116.74
3	C	9	TTD	N3T-C2T-N1T	4.24	121.27	116.74

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	9	TTD	OP2-P-O5'-C5'

There are no ring outliers.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 1 ligands modelled in this entry, 1 is 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.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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	780/1159 (67%)	-0.10	0 100 100	54, 68, 96, 117	0
2	B	355/382 (92%)	-0.03	0 100 100	54, 73, 97, 102	0
3	C	15/15 (100%)	-0.45	0 100 100	82, 104, 114, 119	0
4	D	16/16 (100%)	-0.32	0 100 100	67, 91, 124, 127	0
All	All	1166/1572 (74%)	-0.09	0 100 100	54, 69, 100, 127	0

There are no RSRZ outliers to report.

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	TTD	C	9	40/41	0.17	-0.95	85,92,96,96	0

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains.

The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
5	CA	C	1017	1/1	0.16	-0.87	77,77,77,77	0

6.5 Other polymers ⓘ

There are no such residues in this entry.