Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 07:26 PM GMT

PDB ID	:	1FN2
Title	:	9-AMINO-(N-(2-DIMETHYLAMINO)BUTYL)ACRIDINE-4-CARBOXAMIDE BOUND TO D(CGTACG)2
Authors	:	Adams, A.; Guss, J.M.; Collyer, C.A.; Denny, W.A.; Wakelin, L.P.G.
Deposited on	:	2000-08-19
Resolution	:	1.60 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
http://wwpdb.org/validation/2016/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:

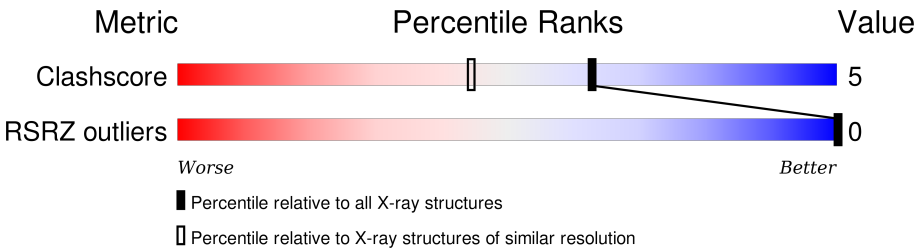
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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	102246	2732 (1.60-1.60)
RSRZ outliers	91569	2479 (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 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	5	100%
2	B	6	67% 33%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 273 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 DNA chain called DNA (5'-D(P*GP*TP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	5	104	49	20	30	5	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*TP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	6	120	58	23	34	5	0	0	0

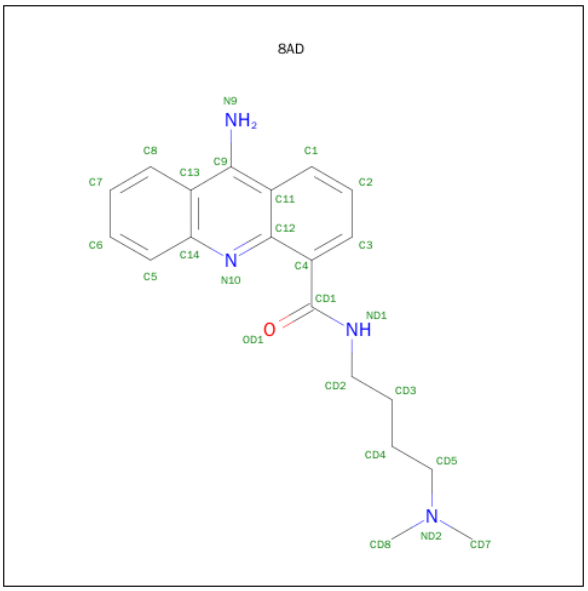
- Molecule 3 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Co		
3	B	1	1	1	0	0
3	A	1	1	1	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
4	B	1	1	1	0	0

- Molecule 5 is 9-AMINO-(N-(2-DIMETHYLAMINO)BUTYL)ACRIDINE-4-CARBOXAMIDE (three-letter code: 8AD) (formula: $C_{20}H_{24}N_4O$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
5	A	1	15	13	2	0	0

- Molecule 6 is water.

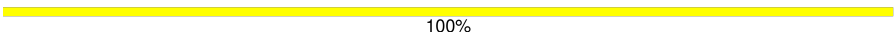
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	12	12	12	0	0
			Total	O		
6	B	19	19	19	0	0

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 errors 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: DNA (5'-D(P*GP*TP*AP*CP*G)-3')

Chain A:



- Molecule 2: DNA (5'-D(*CP*GP*TP*AP*CP*G)-3')

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	28.96Å 52.62Å 40.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.60 16.04 – 1.60	Depositor EDS
% Data completeness (in resolution range)	92.8 (40.00-1.60) 92.5 (16.04-1.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.36 (at 1.60Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.248 , 0.297 0.247 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	19.3	Xtriage
Anisotropy	0.532	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 62.9	EDS
Estimated twinning fraction	0.036 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.058 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 4123 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	273	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.42% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: CO, NA, 8AD

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	2.48	7/116 (6.0%)	2.84	14/177 (7.9%)
2	B	2.53	9/134 (6.7%)	2.88	14/205 (6.8%)
All	All	2.51	16/250 (6.4%)	2.86	28/382 (7.3%)

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1006	DG	C4'-C3'	10.05	1.63	1.53
2	B	2002	DG	C4'-C3'	9.83	1.63	1.53
1	A	1003	DT	C4'-C3'	9.59	1.62	1.53
1	A	1002	DG	C4'-C3'	9.48	1.62	1.53
2	B	2003	DT	C4'-C3'	9.26	1.62	1.53
1	A	1005	DC	C4'-C3'	9.19	1.62	1.53
1	A	1004	DA	C4'-C3'	9.02	1.62	1.53
2	B	2005	DC	C4'-C3'	9.01	1.62	1.53
2	B	2001	DC	C4'-C3'	9.01	1.62	1.53
2	B	2006	DG	C4'-C3'	8.88	1.62	1.53
2	B	2004	DA	C4'-C3'	8.28	1.61	1.53
2	B	2004	DA	O3'-P	6.52	1.69	1.61
1	A	1003	DT	O4'-C1'	5.69	1.49	1.42
2	B	2001	DC	O4'-C1'	5.54	1.48	1.42
2	B	2006	DG	O4'-C1'	5.50	1.48	1.42
1	A	1004	DA	O3'-P	5.33	1.67	1.61

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2002	DG	P-O5'-C5'	-14.43	97.82	120.90
1	A	1002	DG	O4'-C4'-C3'	-11.49	99.11	106.00
1	A	1006	DG	O4'-C4'-C3'	-10.09	99.95	106.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1003	DT	O4'-C4'-C3'	-9.45	100.33	106.00
2	B	2001	DC	O4'-C1'-N1	-8.89	101.77	108.00
2	B	2002	DG	O4'-C4'-C3'	-8.39	100.97	106.00
1	A	1002	DG	P-O3'-C3'	-8.03	110.07	119.70
2	B	2005	DC	O4'-C4'-C3'	-7.88	101.27	106.00
1	A	1003	DT	O4'-C1'-N1	-7.33	102.87	108.00
1	A	1003	DT	C4'-C3'-C2'	7.32	109.68	103.10
1	A	1005	DC	O4'-C4'-C3'	-7.05	101.68	104.50
1	A	1004	DA	O5'-P-OP2	-6.96	99.44	105.70
1	A	1002	DG	P-O5'-C5'	-6.92	109.84	120.90
1	A	1005	DC	P-O3'-C3'	-6.79	111.55	119.70
2	B	2003	DT	O4'-C4'-C3'	-6.71	101.81	104.50
1	A	1005	DC	P-O5'-C5'	-6.71	110.17	120.90
2	B	2001	DC	O4'-C4'-C3'	-6.70	101.82	104.50
2	B	2006	DG	O4'-C4'-C3'	-6.69	101.83	104.50
2	B	2004	DA	P-O5'-C5'	-6.46	110.56	120.90
1	A	1005	DC	C4'-C3'-C2'	6.44	108.90	103.10
2	B	2004	DA	O4'-C4'-C3'	-6.39	101.94	104.50
2	B	2002	DG	N1-C6-O6	-5.96	116.33	119.90
1	A	1002	DG	O4'-C1'-N9	-5.85	103.91	108.00
2	B	2005	DC	C4'-C3'-C2'	5.66	108.19	103.10
2	B	2005	DC	O5'-P-OP2	-5.24	100.98	105.70
2	B	2001	DC	N3-C2-O2	5.22	125.56	121.90
1	A	1004	DA	O4'-C4'-C3'	-5.09	102.47	104.50
2	B	2001	DC	N1-C2-O2	-5.08	115.85	118.90

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	104	0	57	0	0
2	B	120	0	69	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
5	A	15	0	9	0	0
6	A	12	0	0	0	0
6	B	19	0	0	1	0
All	All	273	0	135	2	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 (2) 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)
2:B:2002:DG:H2'	2:B:2003:DT:C6	2.48	0.49
2:B:2003:DT:H71	6:B:4019:HOH:O	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein molecules in this entry.

5.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

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 4 ligands modelled in this entry, 3 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	8AD	A	3014	-	17,17,27	1.22	2 (11%)	22,24,37	1.51	3 (13%)

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
5	8AD	A	3014	-	-	0/0/0/12	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	3014	8AD	C14-N10	-2.57	1.31	1.35
5	A	3014	8AD	C3-C4	2.06	1.41	1.36

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	3014	8AD	C11-C12-N10	-4.02	119.78	123.45
5	A	3014	8AD	C13-C14-N10	-2.36	121.29	123.45
5	A	3014	8AD	C14-N10-C12	4.25	123.68	117.88

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)

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	5/5 (100%)	0.09	0	100	100	25, 26, 33, 33	0
2	B	6/6 (100%)	0.21	0	100	100	18, 24, 34, 43	0
All	All	11/11 (100%)	0.16	0	100	100	18, 26, 34, 43	0

There are no RSRZ outliers to report.

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. 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	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	8AD	A	3014	15/25	0.68	0.41	-	19,25,31,32	15
3	CO	B	4001	1/1	1.00	0.02	-	21,21,21,21	1
4	NA	B	4003	1/1	0.99	0.22	-	37,37,37,37	1
3	CO	A	4002	1/1	0.99	0.13	-	37,37,37,37	0

6.5 Other polymers (i)

There are no such residues in this entry.