WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 11, 2024 – 10:11 AM EST

PDB ID	:	3BRD
Title	:	CSL (Lag-1) bound to DNA with Lin-12 RAM peptide, P212121
Authors	:	Wilson, J.J.; Kovall, R.A.
Deposited on	:	2007-12-21
Resolution	:	2.21 Å(reported)

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

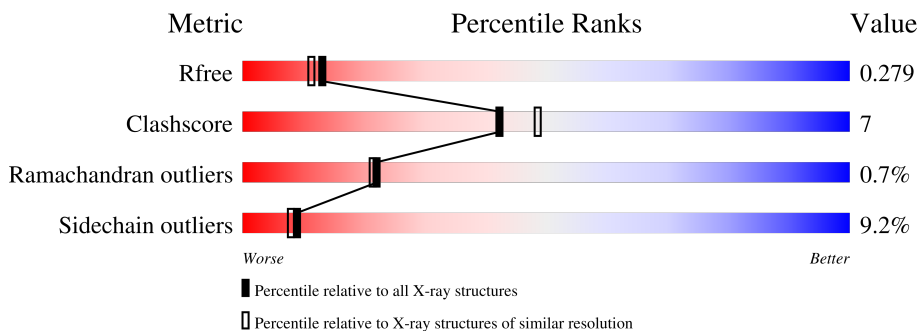|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.21 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	Percentile Ranks	Value
Rfree		0.279
Clashscore		7
Ramachandran outliers		0.7%
Sidechain outliers		9.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)

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%

Mol	Chain	Length	Quality of chain
1	B	15	33% 67%
2	C	15	33% 67%
3	A	477	69% 18% • 10%
4	D	29	38% 10% • 48%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 4328 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(*DTP*DTP*DAP*DCP*DTP*DGP*DTP*D GP*DGP*DGP*DAP*DAP*DAP*DGP*DA)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	15	311	149	61	87	14	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*DAP*DAP*DTP*DCP*DTP*DTP*DTP*D CP*DCP*DCP*DAP*DCP*DAP*DGP*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	15	298	145	50	89	14	0	0	0

- Molecule 3 is a protein called Lin-12 and glp-1 phenotype protein 1, isoform a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	427	3423	2174	592	639	18	0	4	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	187	GLY	-	expression tag	UNP Q9TYY1
A	188	PRO	-	expression tag	UNP Q9TYY1
A	189	LEU	-	expression tag	UNP Q9TYY1
A	190	GLY	-	expression tag	UNP Q9TYY1
A	191	SER	-	expression tag	UNP Q9TYY1

- Molecule 4 is a protein called Protein lin-12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	15	125	78	24	20	3	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	929	SER	-	expression tag	UNP P14585

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	D	1	Total 4	C 2	O 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	24	Total 24	O 24	0	0
6	C	9	Total 9	O 9	0	0
6	A	98	Total 98	O 98	0	0

3 Residue-property plots ⓘ

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.
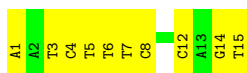
• Molecule 1: DNA (5'-D(*DTP*DTP*DAP*DCP*DTP*DGP*DTP*DGP*DGP*DGP*DAP*DAP*DAP*DGP*DA)-3')
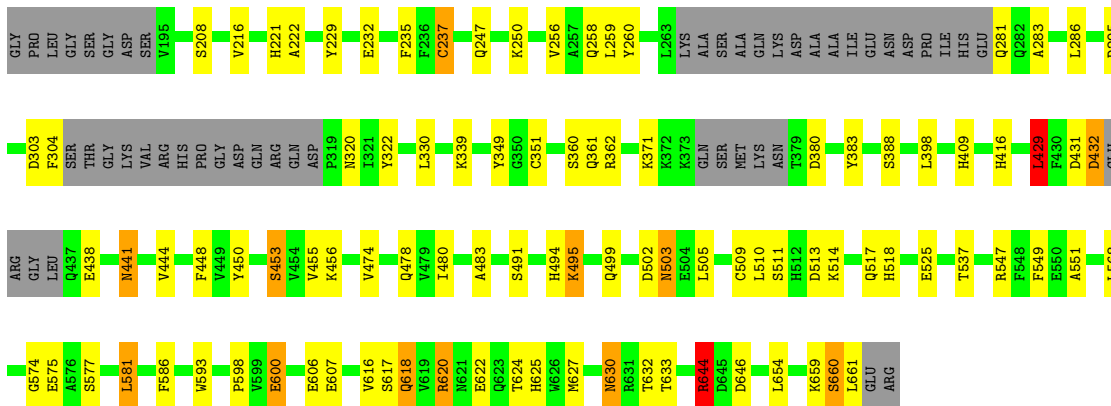
Chain B:
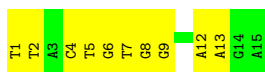
• Molecule 2: DNA (5'-D(*DAP*DAP*DTP*DCP*DTP*DTP*DTP*DCP*DCP*DCP*DAP*DCP*DAP*DGP*DT)-3')

Chain C:

• Molecule 3: Lin-12 and glp-1 phenotype protein 1, isoform a

Chain A:

• Molecule 4: Protein lin-12

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	60.15Å 98.87Å 126.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.87 – 2.21 39.86 – 2.21	Depositor EDS
% Data completeness (in resolution range)	94.4 (39.87-2.21) 94.4 (39.86-2.21)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.28 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.3.0020	Depositor
R, R_{free}	0.208 , 0.257 0.253 , 0.279	Depositor DCC
R_{free} test set	1819 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	51.0	Xtriage
Anisotropy	0.202	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 44.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	4328	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.60% 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, 0.333 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: EDO

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	B	1.27	0/350	1.94	11/540 (2.0%)
2	C	1.54	4/332 (1.2%)	2.03	9/509 (1.8%)
3	A	0.77	2/3512 (0.1%)	0.79	5/4743 (0.1%)
4	D	4.54	4/128 (3.1%)	1.17	2/172 (1.2%)
All	All	1.19	10/4322 (0.2%)	1.11	27/5964 (0.5%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	950	GLU	CD-OE1	40.63	1.70	1.25
4	D	950	GLU	CD-OE2	26.51	1.54	1.25
3	A	237	CYS	CB-SG	-11.32	1.63	1.82
4	D	951	ASN	CG-OD1	10.63	1.47	1.24
4	D	951	ASN	CG-ND2	8.05	1.52	1.32

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	12	DC	O4'-C1'-N1	-9.80	101.14	108.00
4	D	950	GLU	CG-CD-OE2	-7.70	102.90	118.30
1	B	2	DT	O4'-C1'-N1	-7.49	102.76	108.00
1	B	12	DA	O4'-C1'-N9	-7.18	102.97	108.00
1	B	7	DT	C6-C5-C7	-6.95	118.73	122.90

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	B	311	0	171	0	0
2	C	298	0	172	2	0
3	A	3423	0	3375	53	0
4	D	125	0	124	5	0
5	A	32	0	48	2	0
5	C	4	0	6	0	0
5	D	4	0	6	0	0
6	A	98	0	0	2	0
6	B	24	0	0	0	0
6	C	9	0	0	0	0
All	All	4328	0	3902	56	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 7.

The worst 5 of 56 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)
4:D:950:GLU:OE1	4:D:950:GLU:CD	1.70	1.30
3:A:303:ASP:OD2	3:A:304:PHE:N	1.95	1.00
3:A:525:GLU:H	3:A:525:GLU:CD	1.91	0.73
3:A:495:LYS:CE	3:A:537:THR:OG1	2.39	0.71
3:A:448:PHE:O	5:A:10:EDO:H11	1.97	0.64

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	
3	A	421/477 (88%)	408 (97%)	11 (3%)	2 (0%)	29	30
4	D	13/29 (45%)	12 (92%)	0	1 (8%)	1	0
All	All	434/506 (86%)	420 (97%)	11 (2%)	3 (1%)	22	21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	660	SER
4	D	950	GLU
3	A	483	ALA

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	
3	A	380/416 (91%)	344 (90%)	36 (10%)	8	7
4	D	14/27 (52%)	14 (100%)	0	100	100
All	All	394/443 (89%)	358 (91%)	36 (9%)	9	8

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	607	GLU
3	A	654	LEU
3	A	616	VAL
3	A	622	GLU
3	A	429	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	554	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	618	GLN
3	A	630	ASN
3	A	396	ASN
3	A	416	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

10 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	$\#|Z| > 2$
5	EDO	A	4	-	3,3,3	0.41	0	2,2,2	0.61	0
5	EDO	A	9	-	3,3,3	0.44	0	2,2,2	0.32	0
5	EDO	A	3	-	3,3,3	0.46	0	2,2,2	0.55	0
5	EDO	A	8	-	3,3,3	0.54	0	2,2,2	0.22	0
5	EDO	D	6	-	3,3,3	0.55	0	2,2,2	0.15	0
5	EDO	A	1	-	3,3,3	0.52	0	2,2,2	0.40	0
5	EDO	A	2	-	3,3,3	0.71	0	2,2,2	0.24	0
5	EDO	A	5	-	3,3,3	0.80	0	2,2,2	0.59	0
5	EDO	C	16	-	3,3,3	0.60	0	2,2,2	0.18	0
5	EDO	A	10	-	3,3,3	0.40	0	2,2,2	0.25	0

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	EDO	A	4	-	-	1/1/1/1	-
5	EDO	A	9	-	-	1/1/1/1	-
5	EDO	A	3	-	-	0/1/1/1	-
5	EDO	A	8	-	-	1/1/1/1	-
5	EDO	D	6	-	-	0/1/1/1	-
5	EDO	A	1	-	-	1/1/1/1	-
5	EDO	A	2	-	-	1/1/1/1	-
5	EDO	A	5	-	-	0/1/1/1	-
5	EDO	C	16	-	-	1/1/1/1	-
5	EDO	A	10	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	16	EDO	O1-C1-C2-O2
5	A	1	EDO	O1-C1-C2-O2
5	A	2	EDO	O1-C1-C2-O2
5	A	9	EDO	O1-C1-C2-O2
5	A	4	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	10	EDO	2	0

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 ⓘ

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

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

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

6.3 Carbohydrates ⓘ

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

6.4 Ligands ⓘ

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

6.5 Other polymers ⓘ

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