



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 22, 2016 – 06:38 AM EDT

PDB ID : 5EMC
Title : Transcription factor GRDBD and smGRE complex
Authors : Su, X.D.; Lian, T.; Jin, J.
Deposited on : 2015-11-06
Resolution : 2.30 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027790
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) : rb-20027790

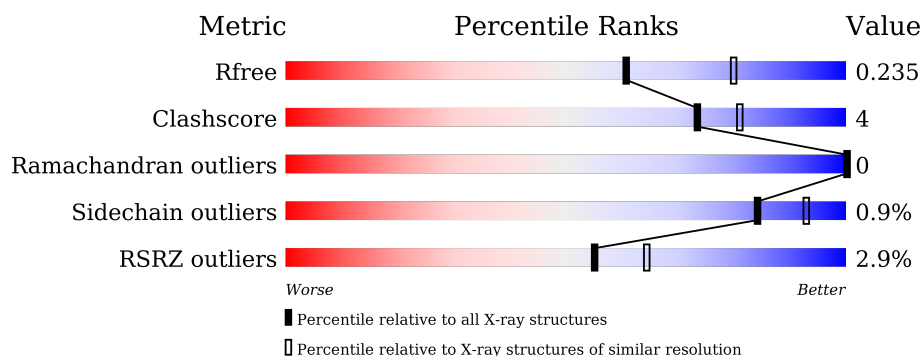
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION





The reported resolution of this entry is 2.30 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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 $\leq 5\%$. 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	C	18	
2	A	94	
2	B	94	
3	D	18	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 1973 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(*CP*CP*AP*GP*AP*AP*(5CM)P*AP*TP*CP*AP*TP*GP*TP*TP*(5CM)P*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	18	Total	C	N	O	P	0	0	0
			364	177	65	105	17			

- Molecule 2 is a protein called Glucocorticoid receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	70	Total	C	N	O	S	0	0	0
			509	312	98	88	11			
2	B	72	Total	C	N	O	S	0	0	0
			545	337	104	93	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	426	GLY	-	expression tag	UNP P04150
A	427	SER	-	expression tag	UNP P04150
A	428	HIS	-	expression tag	UNP P04150
A	429	MET	-	expression tag	UNP P04150
B	426	GLY	-	expression tag	UNP P04150
B	427	SER	-	expression tag	UNP P04150
B	428	HIS	-	expression tag	UNP P04150
B	429	MET	-	expression tag	UNP P04150

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	18	Total	C	N	O	P	0	0	0
			367	178	67	105	17			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total 2	Zn 2	0	0
4	A	2	Total 2	Zn 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	40	Total 40	O 40	0	0
5	A	52	Total 52	O 52	0	0
5	B	59	Total 59	O 59	0	0
5	D	33	Total 33	O 33	0	0

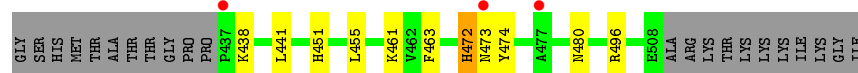
- Molecule 1: DNA (5'-D(*CP*CP*AP*GP*AP*AP*(5CM)P*AP*TP*CP*AP*TP*GP*TP*TP*(5CM)P*TP*G)-3')

A diagram showing a sequence of four colored blocks: C1 (blue), C7 (orange), T12 (green), and G18 (red), connected by horizontal lines.

Chain A:  2% 70% 26%



Chain B: 



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

A diagram showing a 4-nucleotide DNA sequence. The nucleotides are represented by colored vertical bars: C27 (green), A34 (yellow), G39 (yellow), and G44 (green). The sequence is connected by horizontal lines, indicating base pairing or a continuous strand.

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	38.94Å 99.64Å 112.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.26 – 2.30 37.26 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.2 (37.26-2.30) 98.4 (37.26-2.30)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.86 (at 2.29Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.225 , 0.261 0.224 , 0.235	Depositor DCC
R_{free} test set	971 reflections (4.91%)	DCC
Wilson B-factor (Å ²)	33.2	Xtriage
Anisotropy	0.584	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 58.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	1973	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.09 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.6442e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: ZN, 5CM

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	C	0.60	0/361	0.94	0/551
2	A	0.34	0/516	0.44	0/693
2	B	0.38	0/553	0.55	0/738
3	D	0.60	0/365	0.91	0/558
All	All	0.47	0/1795	0.72	0/2540

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	472	HIS	Peptide

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	C	364	0	206	3	0
2	A	509	0	474	2	0
2	B	545	0	538	6	0
3	D	367	0	206	2	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	A	52	0	0	0	0
5	B	59	0	0	2	0
5	C	40	0	0	3	0
5	D	33	0	0	2	0
All	All	1973	0	1424	13	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 4.

The worst 5 of 13 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:473:ASN:HA	5:B:742:HOH:O	1.75	0.86
3:D:39:DG:N3	5:D:102:HOH:O	2.31	0.63
2:B:463:PHE:CE1	2:B:496:ARG:HG2	2.38	0.58
2:A:506:ASN:ND2	2:A:507:LEU:HG	2.24	0.52
3:D:34:DA:N3	5:D:103:HOH:O	2.34	0.52

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	
2	A	68/94 (72%)	64 (94%)	4 (6%)	0	100	100
2	B	70/94 (74%)	66 (94%)	4 (6%)	0	100	100
All	All	138/188 (73%)	130 (94%)	8 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	
2	A	51/78 (65%)	51 (100%)	0	100	100
2	B	58/78 (74%)	57 (98%)	1 (2%)	68	83
All	All	109/156 (70%)	108 (99%)	1 (1%)	84	93

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	441	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	506	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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
1	5CM	C	16	1,3	13,21,22	4.16	9 (69%)	17,30,33	1.56	5 (29%)
1	5CM	C	7	1,3	13,21,22	4.15	9 (69%)	17,30,33	1.38	4 (23%)
3	5CM	D	33	1,3	13,21,22	4.16	9 (69%)	17,30,33	1.30	4 (23%)
3	5CM	D	42	1,3	13,21,22	4.13	9 (69%)	17,30,33	1.13	1 (5%)

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
1	5CM	C	16	1,3	-	0/3/21/22	0/2/2/2
1	5CM	C	7	1,3	-	0/3/21/22	0/2/2/2
3	5CM	D	33	1,3	-	0/3/21/22	0/2/2/2
3	5CM	D	42	1,3	-	0/3/21/22	0/2/2/2

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	33	5CM	C2'-C3'	-7.60	1.32	1.52
3	D	42	5CM	C2'-C3'	-7.55	1.32	1.52
1	C	16	5CM	C2'-C3'	-7.52	1.32	1.52
1	C	7	5CM	C2'-C3'	-7.51	1.32	1.52
3	D	33	5CM	O4'-C4'	-7.27	1.28	1.45

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	16	5CM	C4'-O4'-C1'	-2.55	102.89	109.42
1	C	16	5CM	C5A-C5-C4	-2.50	118.82	121.47
3	D	33	5CM	C4'-O4'-C1'	-2.40	103.26	109.42
1	C	7	5CM	C5A-C5-C4	-2.16	119.18	121.47
3	D	33	5CM	O4'-C1'-C2'	-2.07	102.21	106.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	7	5CM	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 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 ⓘ

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	C	16/18 (88%)	0.12	0 100 100	33, 48, 53, 55	0
2	A	70/94 (74%)	0.44	2 (2%) 55 64	21, 39, 70, 84	0
2	B	72/94 (76%)	0.62	3 (4%) 40 49	22, 30, 67, 98	0
3	D	16/18 (88%)	0.07	0 100 100	39, 44, 55, 58	0
All	All	174/224 (77%)	0.45	5 (2%) 55 64	21, 37, 67, 98	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	437	PRO	4.6
2	B	473	ASN	3.4
2	A	504	GLY	2.5
2	A	468	VAL	2.2
2	B	477	ALA	2.2

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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
1	5CM	C	16	20/21	0.92	0.15	-	33,44,65,65	0
3	5CM	D	42	20/21	0.91	0.18	-	40,45,56,59	0
1	5CM	C	7	20/21	0.90	0.14	-	35,51,59,61	0
3	5CM	D	33	20/21	0.91	0.14	-	32,46,65,67	0

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(\AA^2)	Q<0.9
4	ZN	B	602	1/1	0.99	0.15	-0.60	24,24,24,24	0
4	ZN	A	602	1/1	1.00	0.10	-1.39	22,22,22,22	0
4	ZN	B	601	1/1	0.99	0.09	-2.77	29,29,29,29	0
4	ZN	A	601	1/1	1.00	0.08	-3.47	37,37,37,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.