wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 03:48 PM GMT

PDB ID	:	4DJK
Title	:	Structure of glutamate-GABA antiporter GadC
Authors	:	Ma, D.; Lu, P.L.; Yan, C.Y.; Fan, C.; Yin, P.; Wang, J.W.; Shi, Y.G.
Deposited on	:	2012-02-02
Resolution	:	3.10 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

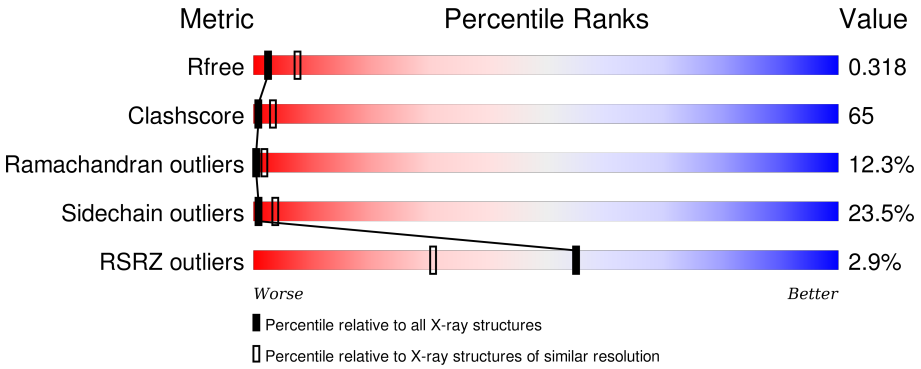
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	511	2% 22% 47% 17% • 12%
1	B	511	3% 20% 47% 19% • 11%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6891 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 Probable glutamate/gamma-aminobutyrate antiporter.

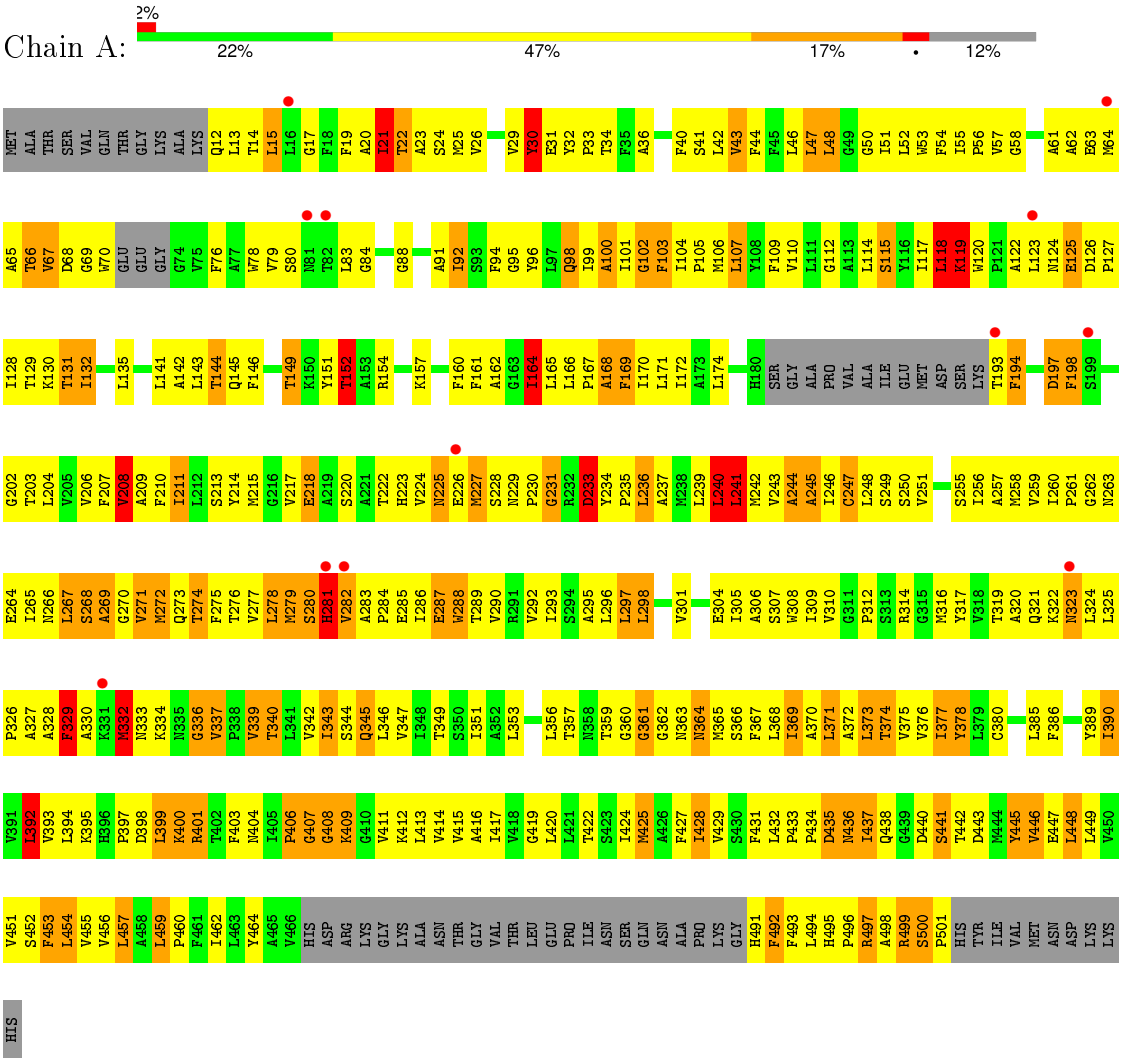
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	0	0	0
			3436	2314	528	574	20			
1	B	453	Total	C	N	O	S	0	0	0
			3455	2324	532	579	20			

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.

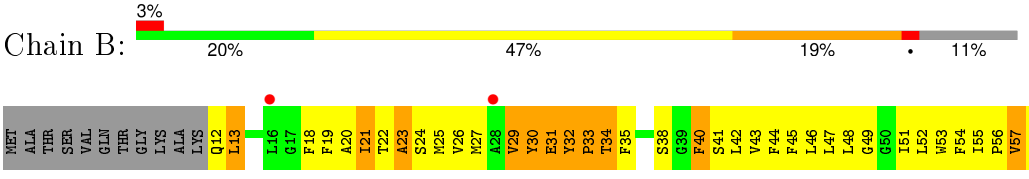
- Molecule 1: Probable glutamate/gamma-aminobutyrate antiporter

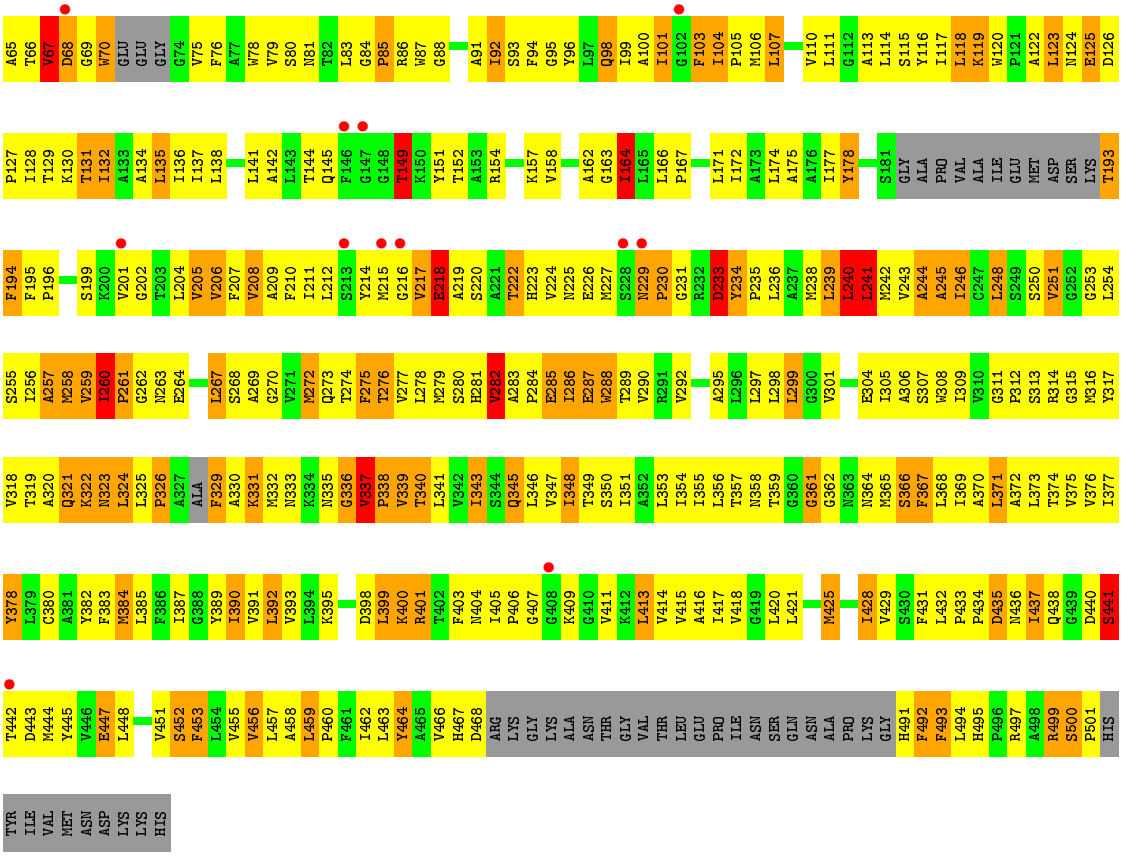
Chain A:



- Molecule 1: Probable glutamate/gamma-aminobutyrate antiporter

Chain B:




A65 T66 V67 D68 G69 W70 GLU GLU GLY G74 V75 F76 A77 W78 V79 S80 N81 T82 L83 G84 P85 R86 W87 G88 A91 I92 S93 F94 G95 Y96 L97 Q98 I99 A100 I101 G102 F103 I104 P105 M106 L107 V110 L111 G112 A113 L114 S115 Y116 I117 L118 K119 W120 P121 A122 L123 N124 E125 D126
P127 I128 T129 K130 T131 I132 A133 A134 L135 I136 I137 L138 L141 A142 L143 T144 Q145 F146 G147 G148 T149 K150 Y151 T152 A153 R154 K157 V158 A162 G163 I164 L165 L166 P167 L171 I172 A173 L174 A175 A176 I177 Y178 S181 GLY ALA PRO VAL ALA ILE GLU MET ASP SER LYS T193
F194 F195 P196 S199 K200 V201 G202 T203 L204 V205 V206 F207 V208 A209 F210 I211 L212 S213 Y214 M215 G216 V217 E218 A219 S220 A221 T222 H223 V224 N225 E226 M227 S228 N229 P230 G231 R232 D233 Y234 P235 L236 A237 M238 L239 L240 L241 M242 V243 A244 A245 I246 C247 L248 S249 S250 V251 G252 G253 L254
S255 I256 A257 M258 V259 I260 P261 G262 N263 E264 L267 S268 A269 G270 V271 M272 Q273 T274 F275 T276 V277 L278 M279 S280 H281 V282 A283 P284 E285 I286 E287 W288 T289 V290 R291 V292 A295 L296 I297 L298 L299 G300 V301 E304 I305 A306 S307 W308 I309 V310 G311 P312 S313 R314 G315 M316 Y317
V318 T319 A320 Q321 K322 N323 L324 L325 P326 A327 ALA F329 A330 K331 M332 N333 K334 N335 G336 V337 P338 V339 T340 L341 V342 I343 S344 Q345 L346 V347 I348 T349 S350 I351 A352 L353 I354 I355 L356 T357 N358 T359 G360 G361 G362 N363 N364 M365 S366 F367 L368 I369 A370 L371 A372 L373 T374 V375 V376 I377
Y378 L379 C380 A381 Y382 F383 M384 L385 F386 I387 G388 Y389 I390 V391 L392 V393 L394 K395 D398 L399 K400 R401 T402 F403 N404 I405 P406 G407 G408 K409 G410 V411 K412 L413 V414 V415 A416 I417 V418 G419 L420 L421 M425 I428 V429 S430 F431 L432 P433 P434 D435 N436 I437 Q438 G439 D440 S441
T442 D443 M444 Y445 V446 E447 L448 V451 S452 F453 L454 V455 V456 L457 A458 L459 P460 F461 I462 L463 Y464 A465 V466 H467 D468 ARG LYS GLY LYS ALA ASN THR GLY VAL THR LEU GLU PRO ILE ASN SER GLN ASN ALA PRO LYS GLY H491 F492 F493 L494 H495 P496 R497 A498 R499 S500 P501 HIS
TYR ILE VAL MET ASN ASP LYS LYS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.96Å 106.41Å 185.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.14 – 3.10 33.14 – 3.10	Depositor EDS
% Data completeness (in resolution range)	79.7 (33.14-3.10) 79.8 (33.14-3.10)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.62 (at 3.12Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.282 , 0.325 0.269 , 0.318	Depositor DCC
R_{free} test set	1142 reflections (5.37%)	DCC
Wilson B-factor (Å^2)	102.0	Xtriage
Anisotropy	0.794	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 87.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Outliers	0 of 21315 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6891	wwPDB-VP
Average B, all atoms (Å^2)	142.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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

5.1 Standard geometry

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	0.49	0/3524	0.72	2/4812 (0.0%)
1	B	0.49	0/3543	0.72	0/4836
All	All	0.49	0/7067	0.72	2/9648 (0.0%)

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
1	B	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	392	LEU	CA-CB-CG	5.75	128.52	115.30
1	A	118	LEU	CB-CG-CD2	5.35	120.10	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	337	VAL	Peptide

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	A	3436	0	3595	439	0
1	B	3455	0	3605	486	0
All	All	6891	0	7200	918	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 65.

The worst 5 of 918 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:A:394:LEU:HD11	1:B:204:LEU:HD12	1.40	1.04
1:A:491:HIS:C	1:A:493:PHE:H	1.58	1.01
1:A:101:ILE:HD13	1:A:373:LEU:HD21	1.43	1.00
1:B:400:LYS:H	1:B:400:LYS:HD2	1.24	1.00
1:B:279:MET:O	1:B:281:HIS:N	1.94	1.00

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	443/511 (87%)	307 (69%)	82 (18%)	54 (12%)	0 2
1	B	443/511 (87%)	304 (69%)	84 (19%)	55 (12%)	0 1
All	All	886/1022 (87%)	611 (69%)	166 (19%)	109 (12%)	0 2

5 of 109 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	VAL
1	A	100	ALA
1	A	118	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	152	THR
1	A	223	HIS

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	365/414 (88%)	274 (75%)	91 (25%)	1	3
1	B	368/414 (89%)	287 (78%)	81 (22%)	1	5
All	All	733/828 (88%)	561 (76%)	172 (24%)	1	4

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	404	ASN
1	B	57	VAL
1	B	400	LYS
1	A	428	ILE
1	A	453	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	358	ASN
1	B	495	HIS
1	B	364	ASN
1	B	12	GLN
1	B	467	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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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, 95^{th} 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	451/511 (88%)	-0.14	12 (2%)	58	34	89, 136, 198, 249	0
1	B	453/511 (88%)	-0.13	14 (3%)	52	28	86, 137, 212, 265	0
All	All	904/1022 (88%)	-0.13	26 (2%)	55	31	86, 137, 205, 265	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	226	GLU	4.9
1	B	147	GLY	4.8
1	B	68	ASP	4.2
1	B	213	SER	3.8
1	B	408	GLY	3.8

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.