



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2018 – 05:42 pm GMT

PDB ID : 3F6D
Title : Crystal Structure of a Genetically Modified Delta Class GST (adGSTD4-4) from Anopheles dirus, F123A, in Complex with S-Hexyl Glutathione
Authors : Wongsantichon, J.; Robinson, R.C.; Ketterman, A.J.
Deposited on : 2008-11-05
Resolution : 1.70 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

<https://www.wwpdb.org/validation/2017/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 : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

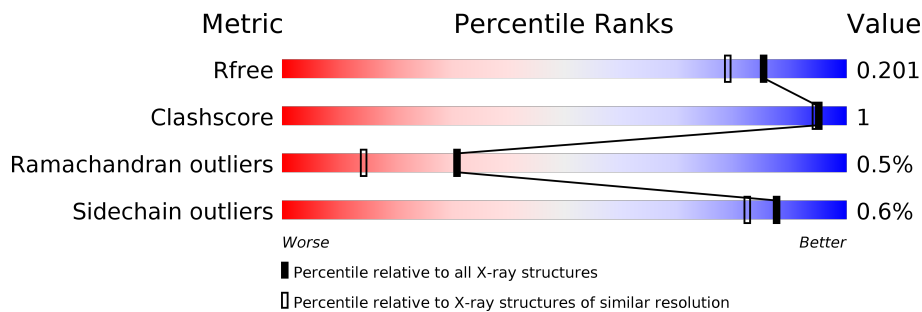
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



The reported resolution of this entry is 1.70 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}	111664	3793 (1.70-1.70)
Clashscore	122126	4167 (1.70-1.70)
Ramachandran outliers	120053	4100 (1.70-1.70)
Sidechain outliers	120020	4100 (1.70-1.70)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	219	 95% 5%
1	B	219	 90% 7%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3749 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.

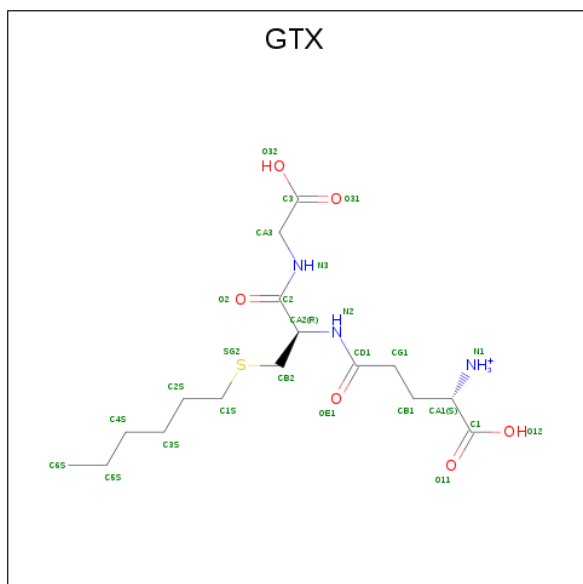
- Molecule 1 is a protein called Glutathione transferase GST1-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	218	Total 1749	C 1122	N 297	O 323	S 7	0	0	0
1	B	203	Total 1627	C 1042	N 275	O 303	S 7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	ALA	PHE	ENGINEERED	UNP Q9GN60
B	123	ALA	PHE	ENGINEERED	UNP Q9GN60

- Molecule 2 is S-HEXYLGLUTATHIONE (three-letter code: GTX) (formula: C₁₆H₃₀N₃O₆S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	Total 26	C 16	N 3	O 6	S 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	B	1	26	16	3	6	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	189	Total	O	0	0
			189	189		
3	B	132	Total	O	0	0
			132	132		

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 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.

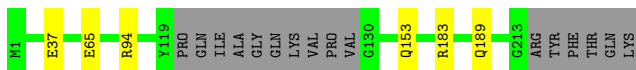
- Molecule 1: Glutathione transferase GST1-4

Chain A:  95% 5%



- Molecule 1: Glutathione transferase GST1-4

Chain B:  90% 7%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	49.69Å 49.69Å 272.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.90 – 1.70 28.90 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.8 (28.90-1.70) 97.7 (28.90-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.12 (at 1.70Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.173 , 0.202 0.172 , 0.201	Depositor DCC
R_{free} test set	2190 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	18.8	Xtrriage
Anisotropy	0.392	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 44.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3749	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.53% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GTX

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	0.51	0/1793	0.59	0/2436
1	B	0.45	0/1666	0.55	0/2261
All	All	0.48	0/3459	0.57	0/4697

There are no bond length outliers.

There are no bond angle outliers.

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	1749	0	1706	7	0
1	B	1627	0	1579	3	0
2	A	26	0	28	0	0
2	B	26	0	28	0	0
3	A	189	0	0	3	0
3	B	132	0	0	1	0
All	All	3749	0	3341	9	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 1.

All (9) 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:B:183:ARG:HG2	1:B:189:GLN:HG2	1.77	0.65
1:A:203:ASP:O	1:A:207:GLU:HG2	2.12	0.49
1:B:37:GLU:HG3	3:B:249:HOH:O	2.13	0.49
1:A:154:TYR:OH	1:A:194:ARG:NH1	2.26	0.49
1:A:95:ARG:NH2	3:A:241:HOH:O	2.43	0.47
1:A:200:PRO:HD2	3:A:278:HOH:O	2.18	0.43
1:A:59:ASP:HB2	1:B:94:ARG:NH2	2.33	0.43
1:A:143:GLU:HG3	3:A:470:HOH:O	2.19	0.42
1:A:95:ARG:HH11	1:A:95:ARG:HA	1.85	0.42

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	216/219 (99%)	214 (99%)	1 (0%)	1 (0%)	31	14
1	B	199/219 (91%)	197 (99%)	1 (0%)	1 (0%)	31	14
All	All	415/438 (95%)	411 (99%)	2 (0%)	2 (0%)	31	14

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	GLU
1	B	65	GLU

5.3.2 Protein sidechains [i](#)

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	180/181 (99%)	179 (99%)	1 (1%)	87	82
1	B	167/181 (92%)	166 (99%)	1 (1%)	87	82
All	All	347/362 (96%)	345 (99%)	2 (1%)	87	82

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	TYR
1	B	153	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

2 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
2	GTX	A	220	-	17,25,25	2.72	2 (11%)	21,30,30	1.07	1 (4%)
2	GTX	B	220	-	17,25,25	2.80	2 (11%)	21,30,30	1.20	1 (4%)

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
2	GTX	A	220	-	-	0/24/30/30	0/0/0/0
2	GTX	B	220	-	-	0/24/30/30	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	220	GTX	O2-C2	7.49	1.38	1.23
2	B	220	GTX	O2-C2	7.74	1.38	1.23
2	A	220	GTX	OE1-CD1	8.22	1.40	1.23
2	B	220	GTX	OE1-CD1	8.45	1.40	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	220	GTX	CB2-SG2-C1S	-3.00	93.28	102.27
2	A	220	GTX	O2-C2-CA2	-2.08	116.04	120.44

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](#)

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

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

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

6.4 Ligands [i](#)

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

6.5 Other polymers [i](#)

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