



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 12, 2021 – 06:08 PM EDT

PDB ID : 1GUM
Title : HUMAN GLUTATHIONE TRANSFERASE A4-4 WITHOUT LIGANDS
Authors : Bruns, C.M.; Hubatsch, I.; Ridderstrom, M.; Mannervik, B.; Tainer, J.A.
Deposited on : 1998-06-11
Resolution : 3.00 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

<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
Xtriage (Phenix)	:	1.13
EDS	:	2.22
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)	:	2.22

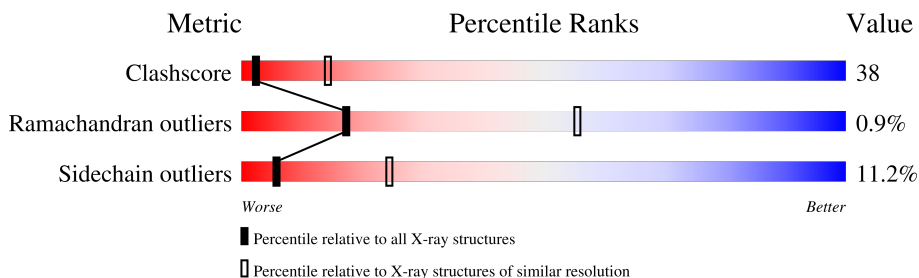
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	222	38% 50% 9% ..
1	B	222	36% 51% 9% ..
1	C	222	37% 50% 9% ..
1	D	222	37% 50% 9% ..
1	E	222	36% 51% 9% ..
1	F	222	36% 51% 9% ..
1	G	222	37% 50% 9% ..
1	H	222	36% 51% 9% ..

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 14400 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 PROTEIN (GLUTATHIONE TRANSFERASE A4-4).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	B	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	C	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	D	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	E	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	F	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	G	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			
1	H	217	Total	C	N	O	S	0	0	0
			1777	1158	298	315	6			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	23	Total	O	0	0
			23	23		
2	B	23	Total	O	0	0
			23	23		
2	C	23	Total	O	0	0
			23	23		
2	D	23	Total	O	0	0
			23	23		
2	E	23	Total	O	0	0
			23	23		
2	F	23	Total	O	0	0
			23	23		

Continued on next page...

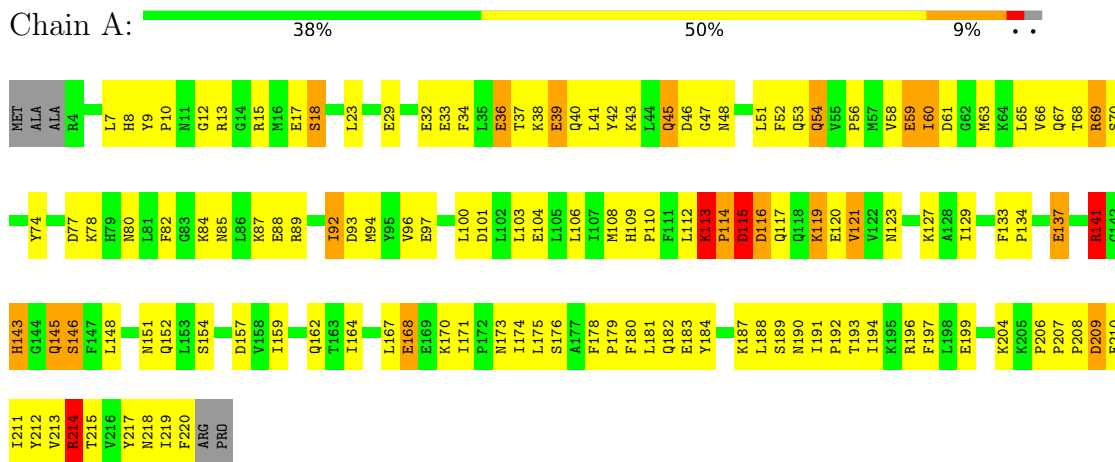
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	23	Total 23	O 23	0	0
2	H	23	Total 23	O 23	0	0

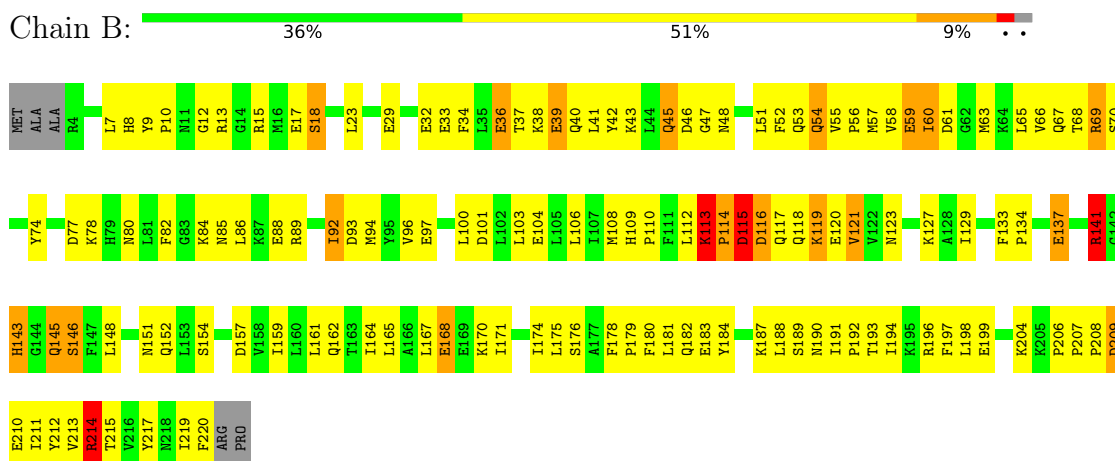
3 Residue-property plots

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.

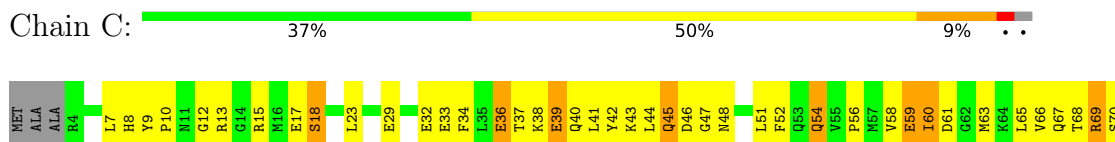
• Molecule 1: PROTEIN (GLUTATHIONE TRANSFERASE A4-4)



• Molecule 1: PROTEIN (GLUTATHIONE TRANSFERASE A4-4)



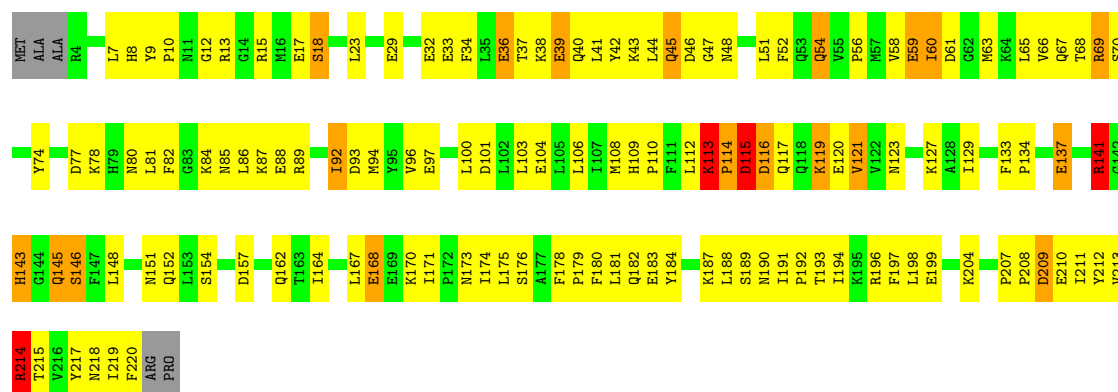
• Molecule 1: PROTEIN (GLUTATHIONE TRANSFERASE A4-4)





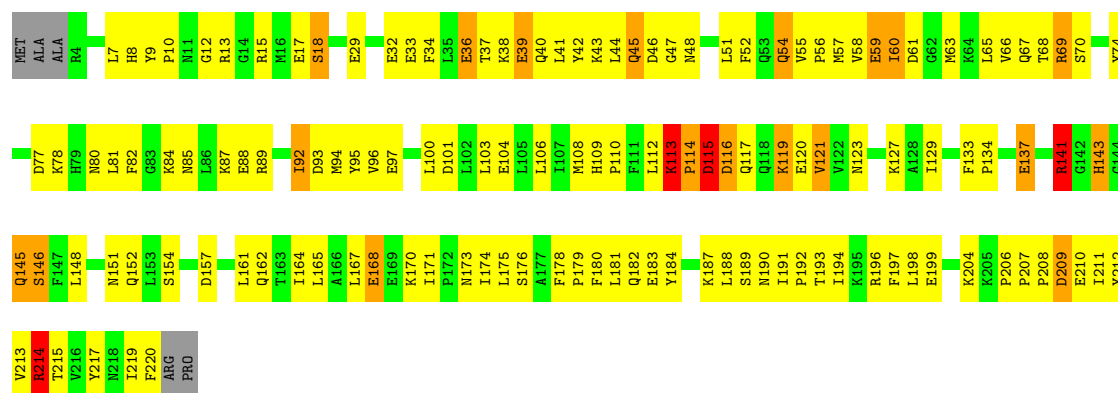
- Molecule 1: PROTEIN (GLUTATHIONE TRANSFERASE A4-4)

Chain D: 37% 50% 9% ..



- Molecule 1: PROTEIN (GLUTATHIONE TRANSFERASE A4-4)

Chain E: 36% 51% 9% ..



- Molecule 1: PROTEIN (GLUTATHIONE TRANSFERASE A4-4)

Chain F: 36% 51% 9% ..



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	155.30Å 156.10Å 101.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	160.00 – 3.00 30.23 – 2.91	Depositor EDS
% Data completeness (in resolution range)	92.0 (160.00-3.00) 90.7 (30.23-2.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.20	Depositor
$\langle I/\sigma(I) \rangle$ ¹	12.91 (at 2.90Å)	Xtriage
Refinement program	TNT 5E	Depositor
R, R_{free}	0.253 , 0.270 0.289 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	1.2	Xtriage
Anisotropy	4.180	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.21 , 91.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.30$, $\langle L^2 \rangle = 0.14$	Xtriage
Estimated twinning fraction	0.189 for k,h,-l	Xtriage
F_o, F_c correlation	0.31	EDS
Total number of atoms	14400	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.42% 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

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	1.24	14/1817 (0.8%)	1.56	25/2456 (1.0%)
1	B	1.24	14/1817 (0.8%)	1.56	25/2456 (1.0%)
1	C	1.24	14/1817 (0.8%)	1.57	25/2456 (1.0%)
1	D	1.24	14/1817 (0.8%)	1.56	25/2456 (1.0%)
1	E	1.24	14/1817 (0.8%)	1.57	25/2456 (1.0%)
1	F	1.24	14/1817 (0.8%)	1.57	25/2456 (1.0%)
1	G	1.24	14/1817 (0.8%)	1.56	25/2456 (1.0%)
1	H	1.24	14/1817 (0.8%)	1.56	25/2456 (1.0%)
All	All	1.24	112/14536 (0.8%)	1.56	200/19648 (1.0%)

The worst 5 of 112 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	97	GLU	CD-OE2	9.88	1.36	1.25
1	B	97	GLU	CD-OE2	9.86	1.36	1.25
1	H	97	GLU	CD-OE2	9.86	1.36	1.25
1	C	97	GLU	CD-OE2	9.85	1.36	1.25
1	A	97	GLU	CD-OE2	9.84	1.36	1.25

The worst 5 of 200 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	113	LYS	C-N-CD	-11.27	95.81	120.60
1	G	113	LYS	C-N-CD	-11.27	95.82	120.60
1	D	113	LYS	C-N-CD	-11.26	95.83	120.60
1	F	113	LYS	C-N-CD	-11.25	95.84	120.60
1	B	113	LYS	C-N-CD	-11.25	95.85	120.60

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	A	1777	0	1829	140	0
1	B	1777	0	1829	148	1
1	C	1777	0	1829	150	1
1	D	1777	0	1829	149	4
1	E	1777	0	1829	144	0
1	F	1777	0	1829	151	0
1	G	1777	0	1829	151	0
1	H	1777	0	1829	148	4
2	A	23	0	0	0	0
2	B	23	0	0	0	0
2	C	23	0	0	0	0
2	D	23	0	0	0	0
2	E	23	0	0	0	0
2	F	23	0	0	0	0
2	G	23	0	0	0	0
2	H	23	0	0	0	0
All	All	14400	0	14632	1094	5

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 38.

The worst 5 of 1094 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:C:94:MET:CE	1:D:66:VAL:H	1.55	1.18
1:C:94:MET:HE2	1:D:66:VAL:N	1.62	1.12
1:G:94:MET:HE2	1:H:66:VAL:H	1.06	1.12
1:E:66:VAL:H	1:F:94:MET:CE	1.61	1.11
1:E:94:MET:HE1	1:F:65:LEU:HA	1.33	1.10

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:196:ARG:NH2	1:H:200:PRO:CG[4_455]	1.34	0.86
1:D:196:ARG:NH2	1:H:200:PRO:CB[4_455]	1.74	0.46
1:D:196:ARG:CZ	1:H:200:PRO:CB[4_455]	2.12	0.08
1:B:118:GLN:CG	1:C:118:GLN:CG[1_554]	2.16	0.04
1:D:196:ARG:CZ	1:H:200:PRO:CG[4_455]	2.18	0.02

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	215/222 (97%)	200 (93%)	13 (6%)	2 (1%)	17	55
1	B	215/222 (97%)	201 (94%)	12 (6%)	2 (1%)	17	55
1	C	215/222 (97%)	200 (93%)	13 (6%)	2 (1%)	17	55
1	D	215/222 (97%)	201 (94%)	12 (6%)	2 (1%)	17	55
1	E	215/222 (97%)	200 (93%)	13 (6%)	2 (1%)	17	55
1	F	215/222 (97%)	201 (94%)	12 (6%)	2 (1%)	17	55
1	G	215/222 (97%)	200 (93%)	13 (6%)	2 (1%)	17	55
1	H	215/222 (97%)	201 (94%)	12 (6%)	2 (1%)	17	55
All	All	1720/1776 (97%)	1604 (93%)	100 (6%)	16 (1%)	17	55

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	114	PRO
1	B	114	PRO
1	C	114	PRO
1	D	114	PRO
1	E	114	PRO

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	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	B	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	C	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	D	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	E	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	F	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	G	197/200 (98%)	175 (89%)	22 (11%)	6	24
1	H	197/200 (98%)	175 (89%)	22 (11%)	6	24
All	All	1576/1600 (98%)	1400 (89%)	176 (11%)	6	24

5 of 176 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	60	ILE
1	G	115	ASP
1	F	92	ILE
1	F	189	SER
1	G	152	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	118	GLN
1	G	218	ASN
1	F	162	GLN
1	G	54	GLN
1	H	54	GLN

5.3.3 RNA ⓘ

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 ⓘ

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.