wwPDB X-ray Structure Validation Summary Report (i)

Dec 6, 2023 – 06:05 am GMT

PDB ID	:	1GY2
Title	:	Crystal structure of Met148Leu rusticyanin
Authors	:	Hough, M.A.; Kanbi, L.D.; Antonyuk, S.; Dodd, F.; Hasnain, S.
Deposited on	:	2002-04-16
Resolution	:	1.82 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

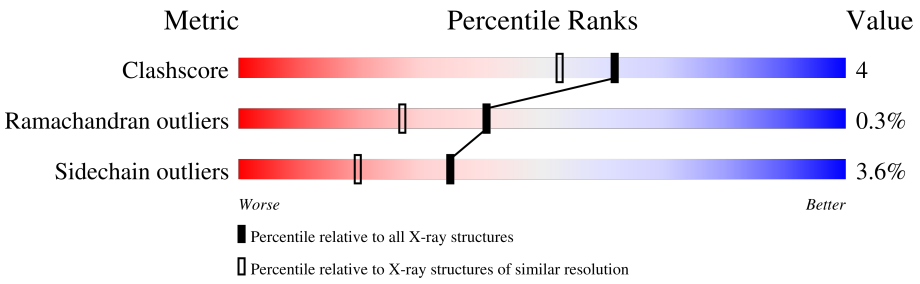
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.82 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	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	155	86% 12% •
1	B	155	85% 15% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2484 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 RUSTICYANIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	155	1170	765	187	215	3	16	0	0
1	B	155	1170	765	187	215	3	24	0	0

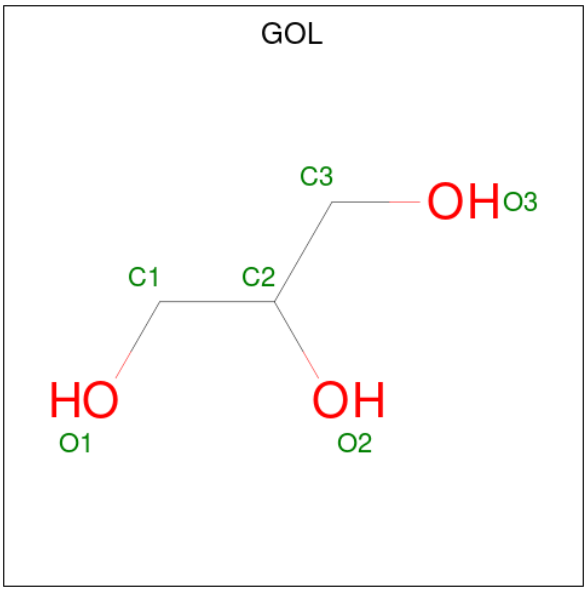
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	148	LEU	MET	engineered mutation	UNP P24930
B	148	LEU	MET	engineered mutation	UNP B7JAQ0
A	124	ASN	ASP	conflict	UNP B7JAQ0
B	124	ASN	ASP	conflict	UNP B7JAQ0

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
2	A	1	1	1	0	0
2	B	1	1	1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	6	3	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	67	67	67	0	0
			Total	O		
4	B	69	69	69	0	0

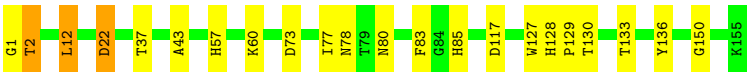
3 Residue-property plots (i)

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.

Note EDS was not executed.

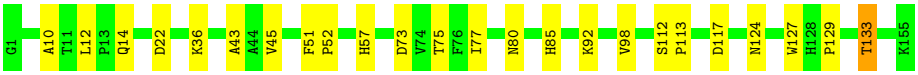
• Molecule 1: RUSTICYANIN

Chain A: 86% 12%



• Molecule 1: RUSTICYANIN

Chain B: 85% 15%



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.45Å 61.43Å 53.47Å 90.00° 96.31° 90.00°	Depositor
Resolution (Å)	53.40 – 1.82	Depositor
% Data completeness (in resolution range)	92.2 (53.40-1.82)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.183 , 0.219	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2484	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

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: GOL, CU

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	0.90	4/1208 (0.3%)	1.01	11/1652 (0.7%)
1	B	0.65	0/1208	0.86	1/1652 (0.1%)
All	All	0.78	4/2416 (0.2%)	0.94	12/3304 (0.4%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	GLY	C-N	18.21	1.75	1.34
1	A	2	THR	CB-OG1	10.44	1.64	1.43
1	A	37	THR	CB-CG2	6.26	1.73	1.52
1	A	12	LEU	CB-CG	6.07	1.70	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1	GLY	C-N-CA	-10.71	94.91	121.70
1	A	37	THR	CA-CB-CG2	-7.40	102.04	112.40
1	A	2	THR	OG1-CB-CG2	7.03	126.16	110.00
1	A	1	GLY	O-C-N	6.75	133.50	122.70
1	A	2	THR	CA-CB-CG2	6.33	121.26	112.40

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	1170	0	1163	11	0
1	B	1170	0	1164	10	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	6	0	8	0	0
4	A	67	0	0	3	0
4	B	69	0	0	1	0
All	All	2484	0	2335	20	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 20 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:80:ASN:HD22	1:B:85:HIS:HE2	1.01	0.99
1:A:80:ASN:HD22	1:A:85:HIS:HE2	1.06	0.97
1:A:130:THR:HG22	4:A:2056:HOH:O	1.83	0.78
1:A:127:TRP:CZ2	1:A:129:PRO:HB3	2.41	0.56
1:B:127:TRP:CZ2	1:B:129:PRO:HB3	2.44	0.53

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	153/155 (99%)	149 (97%)	4 (3%)	0	100 100
1	B	153/155 (99%)	148 (97%)	4 (3%)	1 (1%)	22 10
All	All	306/310 (99%)	297 (97%)	8 (3%)	1 (0%)	41 27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	45	VAL

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	126/126 (100%)	122 (97%)	4 (3%)	39	24
1	B	126/126 (100%)	121 (96%)	5 (4%)	31	16
All	All	252/252 (100%)	243 (96%)	9 (4%)	35	19

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	117	ASP
1	B	133	THR
1	A	133	THR
1	B	12	LEU
1	B	36	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	57	HIS
1	B	78	ASN
1	B	124	ASN
1	B	80	ASN
1	A	124	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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	$\#\|Z\| > 2$
3	GOL	A	1157	-	5,5,5	0.16	0	5,5,5	0.38	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
3	GOL	A	1157	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1157	GOL	C1-C2-C3-O3
3	A	1157	GOL	O2-C2-C3-O3

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1:GLY	C	2:THR	N	1.75

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.