



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 8, 2023 – 12:21 am GMT

PDB ID : 2J4U  
Title : E.coli OmpC - camel Lactoferrin complex  
Authors : Baalaji, S.; Acharya, R.K.; Singh, T.P.; Krishnaswamy, S.  
Deposited on : 2006-09-06  
Resolution : 2.99 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

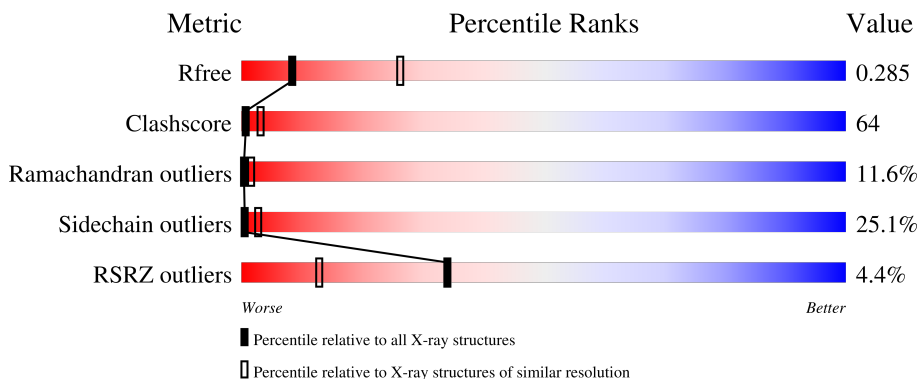
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.99 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.



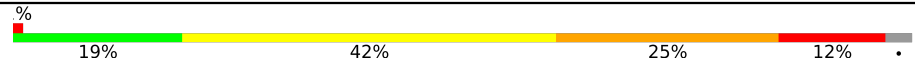
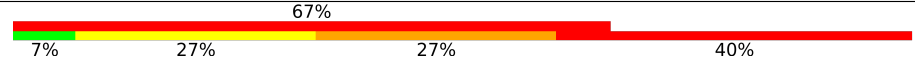
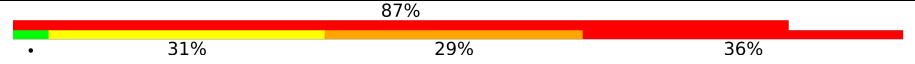
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	P	346	 23% 42% 21% 12% •
1	Q	346	 22% 38% 26% 11% •
1	R	346	 19% 41% 26% 11% •
1	U	346	 20% 43% 22% 13% •
1	V	346	 22% 40% 25% 12% •

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	W	346	 <p>% 19% 42% 25% 12%</p>
2	S	45	 <p>67% 7% 27% 27% 40%</p>
2	X	45	 <p>87% 31% 29% 36%</p>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 16534 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 OUTER MEMBRANE PROTEIN C PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	P	337	2636	1653	441	539	3	0	0	0
1	Q	337	2636	1653	441	539	3	0	0	0
1	R	337	2636	1653	441	539	3	0	0	0
1	U	337	2636	1653	441	539	3	0	0	0
1	V	337	2636	1653	441	539	3	0	0	0
1	W	337	2636	1653	441	539	3	0	0	0

- Molecule 2 is a protein called LACTOTRANSFERRIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	S	45	359	221	72	61	5	0	0	0
2	X	45	359	221	72	61	5	0	0	0

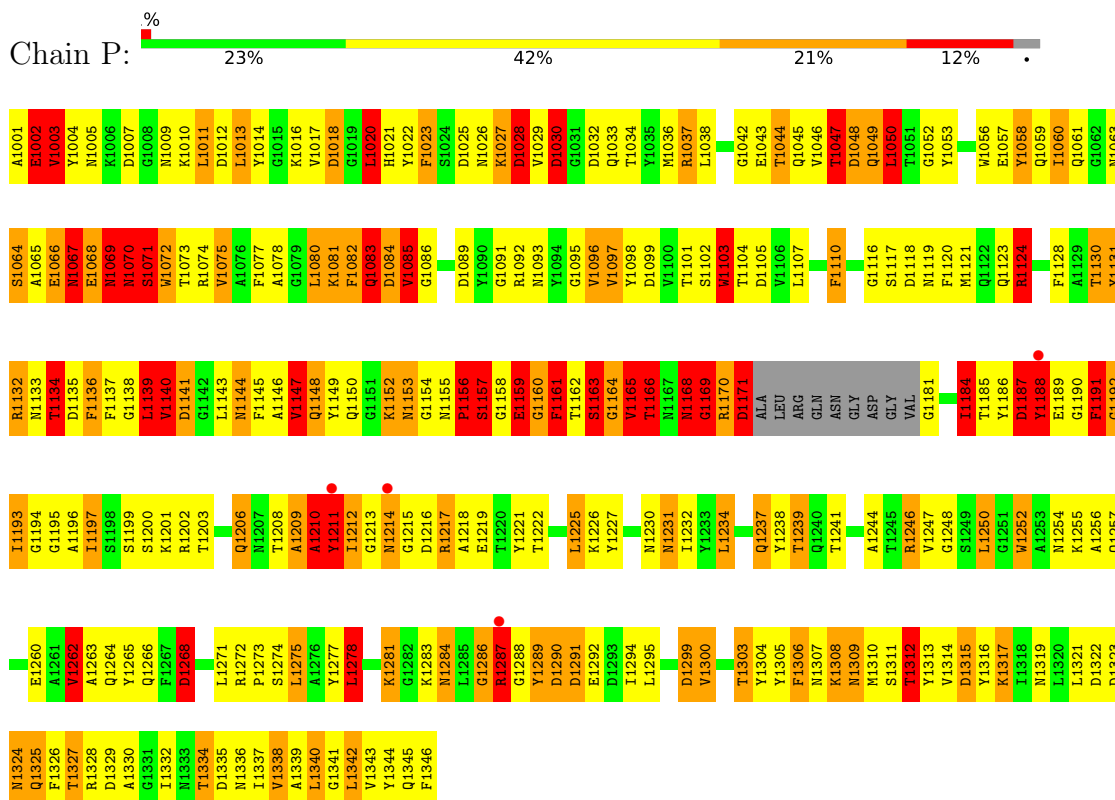
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	1017	LYS	SER	conflict	UNP Q9TUM0
X	1017	LYS	SER	conflict	UNP Q9TUM0

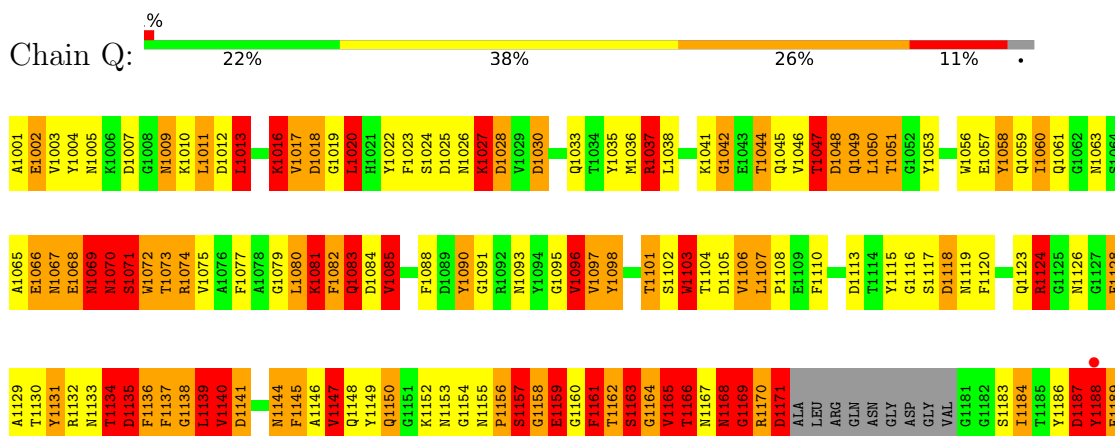
### 3 Residue-property plots

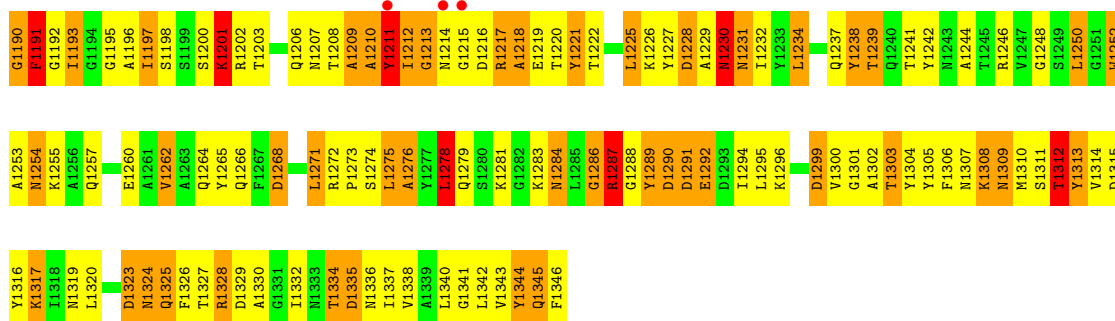
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: OUTER MEMBRANE PROTEIN C PRECURSOR

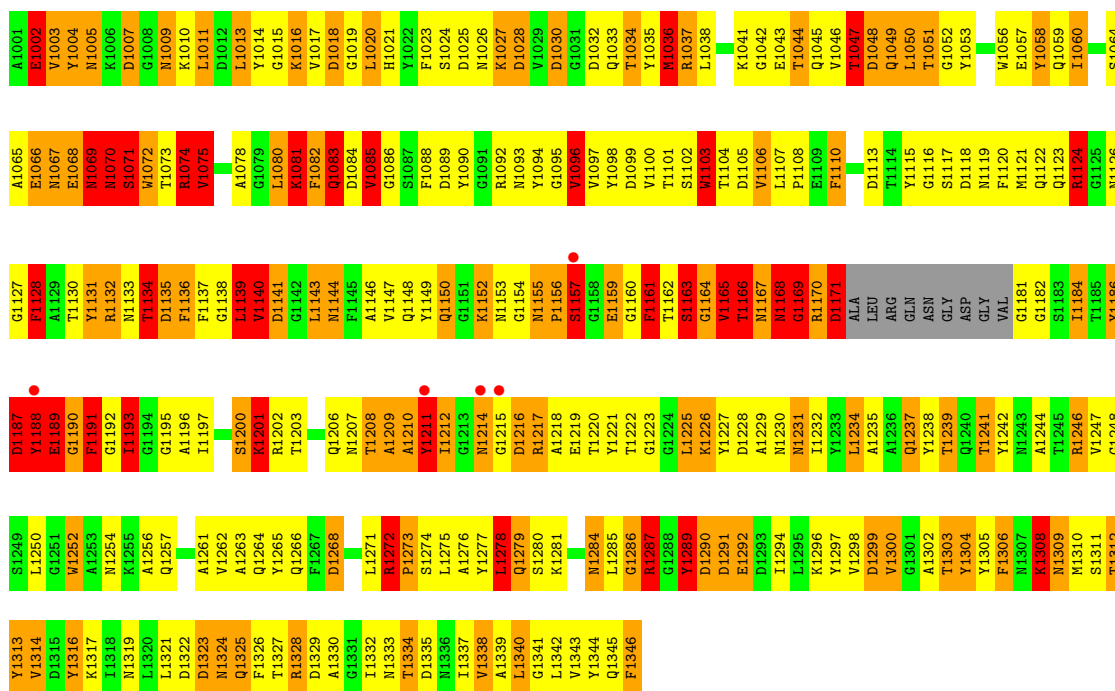
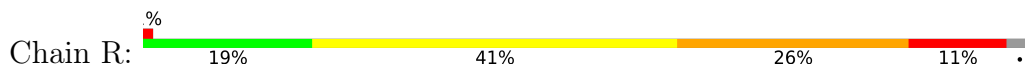


#### • Molecule 1: OUTER MEMBRANE PROTEIN C PRECURSOR

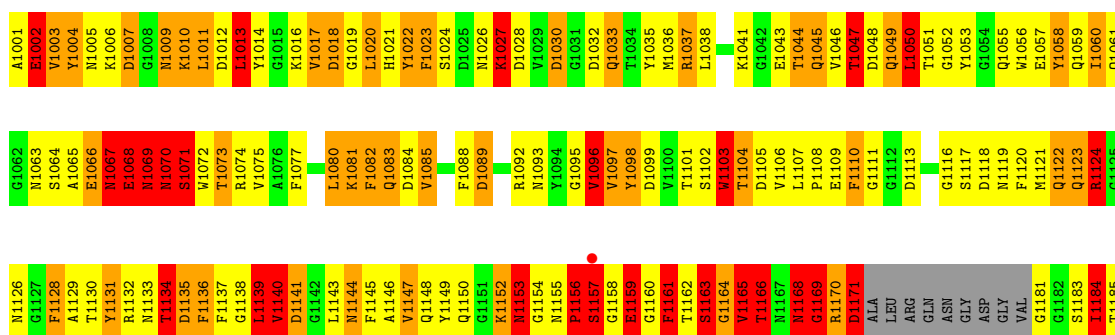
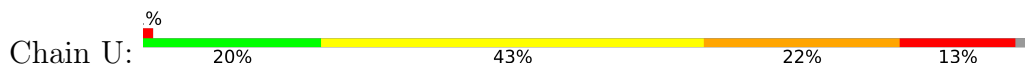


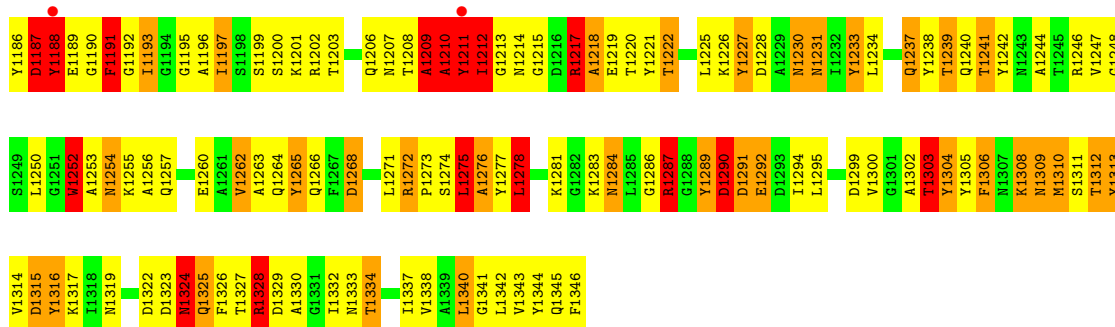


● Molecule 1: OUTER MEMBRANE PROTEIN C PRECURSOR

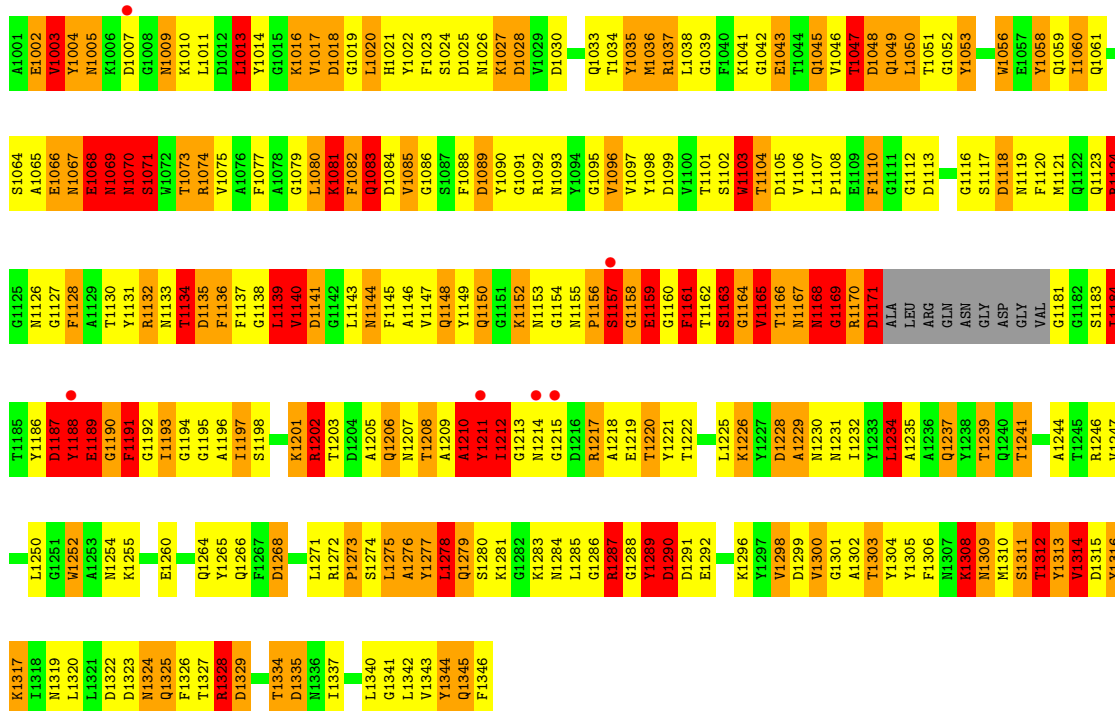
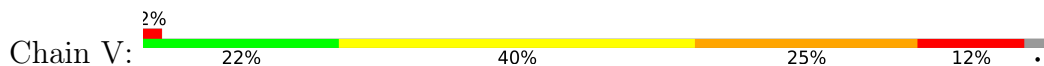


● Molecule 1: OUTER MEMBRANE PROTEIN C PRECURSOR

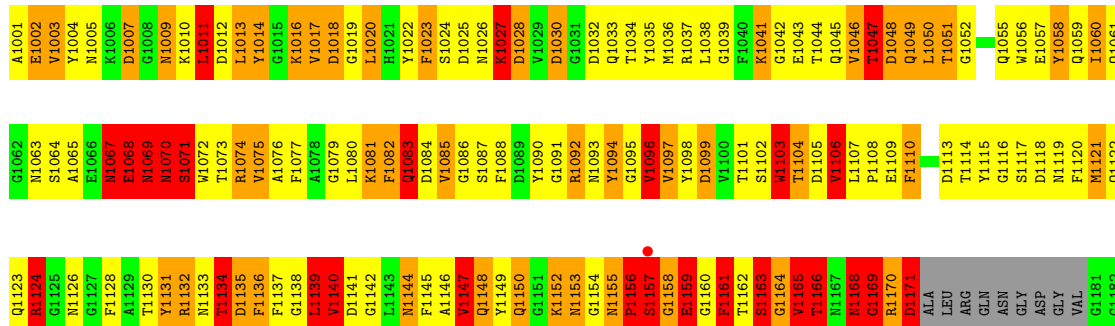
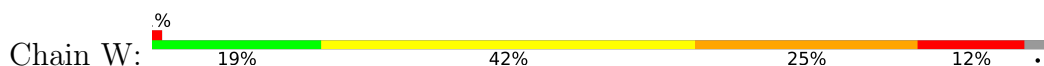


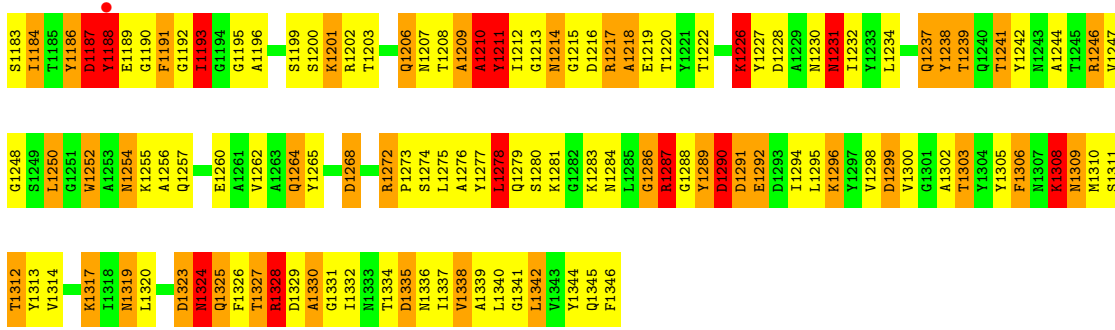


● Molecule 1: OUTER MEMBRANE PROTEIN C PRECURSOR

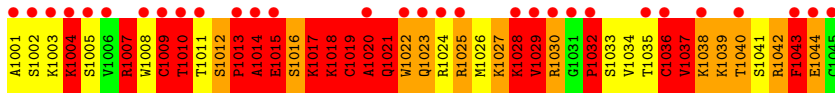
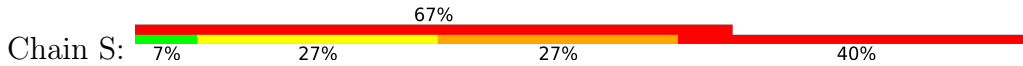


● Molecule 1: OUTER MEMBRANE PROTEIN C PRECURSOR

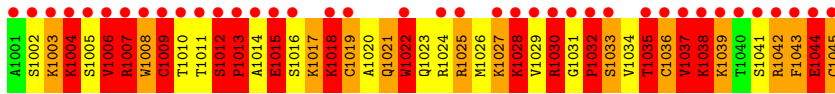
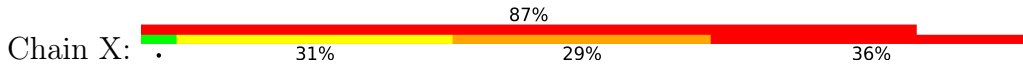




● Molecule 2: LACTOTRANSFERRIN



● Molecule 2: LACTOTRANSFERRIN





## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	201.47Å 116.29Å 152.28Å 90.00° 90.03° 90.00°	Depositor
Resolution (Å)	152.50 – 2.99 19.76 – 2.99	Depositor EDS
% Data completeness (in resolution range)	78.8 (152.50-2.99) 79.0 (19.76-2.99)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.47 (at 2.98Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.215 , 0.282 0.216 , 0.285	Depositor DCC
$R_{free}$ test set	2831 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.2	Xtrriage
Anisotropy	0.140	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 28.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.467 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.467 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.467 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.459 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.468 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	16534	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.60% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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

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	P	2.31	98/2692 (3.6%)	1.99	100/3641 (2.7%)
1	Q	2.28	94/2692 (3.5%)	1.94	99/3641 (2.7%)
1	R	2.34	118/2692 (4.4%)	2.01	103/3641 (2.8%)
1	U	2.31	108/2692 (4.0%)	1.95	89/3641 (2.4%)
1	V	2.32	107/2692 (4.0%)	1.99	90/3641 (2.5%)
1	W	2.30	111/2692 (4.1%)	1.97	92/3641 (2.5%)
2	S	2.65	17/365 (4.7%)	1.80	5/485 (1.0%)
2	X	2.68	20/365 (5.5%)	1.87	8/485 (1.6%)
All	All	2.33	673/16882 (4.0%)	1.97	586/22816 (2.6%)

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	P	0	12
1	Q	1	9
1	R	1	8
1	U	1	11
1	V	1	8
1	W	0	10
2	S	0	2
2	X	0	5
All	All	4	65

The worst 5 of 673 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	V	1161	PHE	CB-CG	21.34	1.87	1.51
1	W	1161	PHE	CB-CG	20.92	1.86	1.51
1	U	1161	PHE	CB-CG	20.71	1.86	1.51
1	P	1161	PHE	CB-CG	20.63	1.86	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	1161	PHE	CB-CG	19.21	1.84	1.51

The worst 5 of 586 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	W	1246	ARG	NE-CZ-NH1	13.43	127.02	120.30
1	P	1025	ASP	CB-CG-OD1	-12.61	106.95	118.30
1	P	1124	ARG	NE-CZ-NH1	11.76	126.18	120.30
1	P	1124	ARG	NE-CZ-NH2	-11.36	114.62	120.30
1	W	1030	ASP	CB-CG-OD1	11.16	128.35	118.30

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	Q	1070	ASN	CA
1	R	1070	ASN	CA
1	U	1170	ARG	CA
1	V	1070	ASN	CA

5 of 65 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	P	1030	ASP	Peptide
1	P	1067	ASN	Peptide
1	P	1068	GLU	Peptide
1	P	1069	ASN	Peptide
1	P	1140	VAL	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	P	2636	0	2404	321	0
1	Q	2636	0	2404	331	0
1	R	2636	0	2404	325	0
1	U	2636	0	2404	308	0
1	V	2636	0	2404	331	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	W	2636	0	2404	348	0
2	S	359	0	376	77	0
2	X	359	0	376	76	0
All	All	16534	0	15176	2040	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 64.

The worst 5 of 2040 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:V:1081:LYS:CD	1:V:1081:LYS:CE	1.77	1.62
1:R:1161:PHE:CB	1:R:1161:PHE:CG	1.84	1.61
1:R:1134:THR:CB	1:R:1134:THR:CG2	1.77	1.60
1:V:1134:THR:CB	1:V:1134:THR:CG2	1.76	1.60
1:W:1081:LYS:CD	1:W:1081:LYS:CE	1.75	1.59

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	P	333/346 (96%)	256 (77%)	45 (14%)	32 (10%)	0	3
1	Q	333/346 (96%)	257 (77%)	43 (13%)	33 (10%)	0	2
1	R	333/346 (96%)	262 (79%)	33 (10%)	38 (11%)	0	2
1	U	333/346 (96%)	259 (78%)	42 (13%)	32 (10%)	0	3
1	V	333/346 (96%)	259 (78%)	41 (12%)	33 (10%)	0	2
1	W	333/346 (96%)	261 (78%)	42 (13%)	30 (9%)	1	3
2	S	43/45 (96%)	8 (19%)	11 (26%)	24 (56%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	X	43/45 (96%)	13 (30%)	10 (23%)	20 (46%)	0	0
All	All	2084/2166 (96%)	1575 (76%)	267 (13%)	242 (12%)	0	1

5 of 242 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	P	1002	GLU
1	P	1003	VAL
1	P	1007	ASP
1	P	1071	SER
1	P	1083	GLN

### 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	P	266/275 (97%)	203 (76%)	63 (24%)	1	3
1	Q	266/275 (97%)	208 (78%)	58 (22%)	1	5
1	R	266/275 (97%)	195 (73%)	71 (27%)	0	2
1	U	266/275 (97%)	200 (75%)	66 (25%)	0	3
1	V	266/275 (97%)	199 (75%)	67 (25%)	0	3
1	W	266/275 (97%)	201 (76%)	65 (24%)	0	3
2	S	41/41 (100%)	24 (58%)	17 (42%)	0	0
2	X	41/41 (100%)	26 (63%)	15 (37%)	0	1
All	All	1678/1732 (97%)	1256 (75%)	422 (25%)	0	3

5 of 422 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	U	1096	VAL
1	V	1045	GLN
1	W	1303	THR
1	U	1135	ASP

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	U	1252	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
1	W	1070	ASN
1	W	1133	ASN
1	R	1070	ASN
1	R	1059	GLN
1	W	1144	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](#)

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	P	337/346 (97%)	0.22	4 (1%) 79 54	17, 29, 47, 60	0
1	Q	337/346 (97%)	0.21	4 (1%) 79 54	17, 29, 47, 59	0
1	R	337/346 (97%)	0.26	5 (1%) 73 46	17, 29, 47, 59	0
1	U	337/346 (97%)	0.27	3 (0%) 84 63	17, 29, 47, 59	0
1	V	337/346 (97%)	0.23	6 (1%) 68 40	17, 29, 47, 59	0
1	W	337/346 (97%)	0.19	2 (0%) 89 72	17, 29, 47, 59	0
2	S	45/45 (100%)	3.18	30 (66%) 0 0	71, 80, 80, 80	0
2	X	45/45 (100%)	3.52	39 (86%) 0 0	77, 80, 80, 80	0
All	All	2112/2166 (97%)	0.36	93 (4%) 34 13	17, 30, 56, 80	0

The worst 5 of 93 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	S	1043	PHE	9.0
2	S	1001	ALA	8.8
2	S	1002	SER	8.3
2	X	1037	VAL	7.6
2	S	1005	SER	7.3

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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.