



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 29, 2024 – 09:06 AM EDT

PDB ID : 3S9M  
Title : Complex between transferrin receptor 1 and transferrin with iron in the N-Lobe, cryocooled 1  
Authors : Eckenroth, B.E.; Steere, A.N.; Mason, A.B.; Everse, S.J.  
Deposited on : 2011-06-01  
Resolution : 3.32 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
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

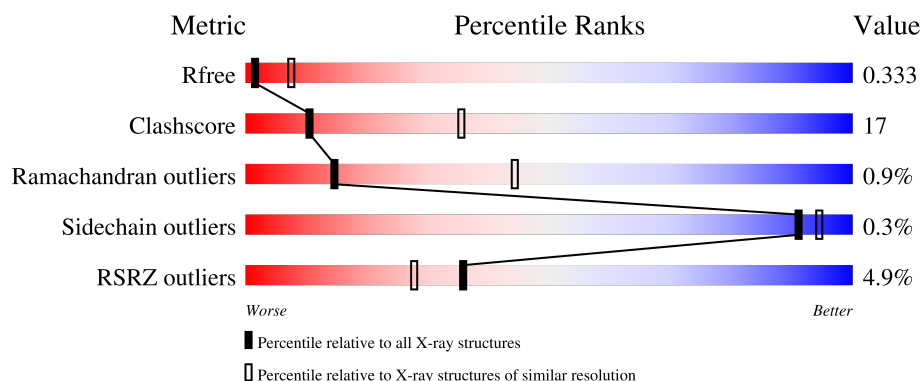
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.32 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))
$R_{free}$	164625	1066 (3.34-3.30)
Clashscore	180529	1111 (3.34-3.30)
Ramachandran outliers	177936	1109 (3.34-3.30)
Sidechain outliers	177891	1108 (3.34-3.30)
RSRZ outliers	164620	1066 (3.34-3.30)

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	A	654	<div> <div>4%</div> <div>68%</div> <div>29%</div> <div>..</div> </div>
1	B	654	<div> <div>5%</div> <div>69%</div> <div>28%</div> <div>.</div> </div>
2	C	693	<div> <div>5%</div> <div>56%</div> <div>18%</div> <div>25%</div> </div>
2	D	693	<div> <div>3%</div> <div>58%</div> <div>12%</div> <div>30%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	CO3	D	905	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16658 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 Transferrin receptor protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	638	Total	C	N	O	S	0	0	0
			4917	3149	820	934	14			
1	B	639	Total	C	N	O	S	0	0	0
			4773	3056	790	913	14			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	107	VAL	-	expression tag	UNP P02786
A	108	PRO	-	expression tag	UNP P02786
A	109	ASP	-	expression tag	UNP P02786
A	110	LYS	-	expression tag	UNP P02786
A	111	HIS	-	expression tag	UNP P02786
A	112	HIS	-	expression tag	UNP P02786
A	113	HIS	-	expression tag	UNP P02786
A	114	HIS	-	expression tag	UNP P02786
A	115	HIS	-	expression tag	UNP P02786
A	116	HIS	-	expression tag	UNP P02786
A	117	ILE	-	expression tag	UNP P02786
A	118	GLU	-	expression tag	UNP P02786
A	119	GLY	-	expression tag	UNP P02786
A	142	SER	GLY	SEE REMARK 999	UNP P02786
B	107	VAL	-	expression tag	UNP P02786
B	108	PRO	-	expression tag	UNP P02786
B	109	ASP	-	expression tag	UNP P02786
B	110	LYS	-	expression tag	UNP P02786
B	111	HIS	-	expression tag	UNP P02786
B	112	HIS	-	expression tag	UNP P02786
B	113	HIS	-	expression tag	UNP P02786
B	114	HIS	-	expression tag	UNP P02786
B	115	HIS	-	expression tag	UNP P02786
B	116	HIS	-	expression tag	UNP P02786
B	117	ILE	-	expression tag	UNP P02786

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	118	GLU	-	expression tag	UNP P02786
B	119	GLY	-	expression tag	UNP P02786
B	142	SER	GLY	SEE REMARK 999	UNP P02786

- Molecule 2 is a protein called Serotransferrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	517	Total	C	N	O	S	0	0	0
			3682	2305	623	718	36			
2	D	488	Total	C	N	O	S	0	0	0
			3246	1988	565	662	31			

There are 38 discrepancies between the modelled and reference sequences:

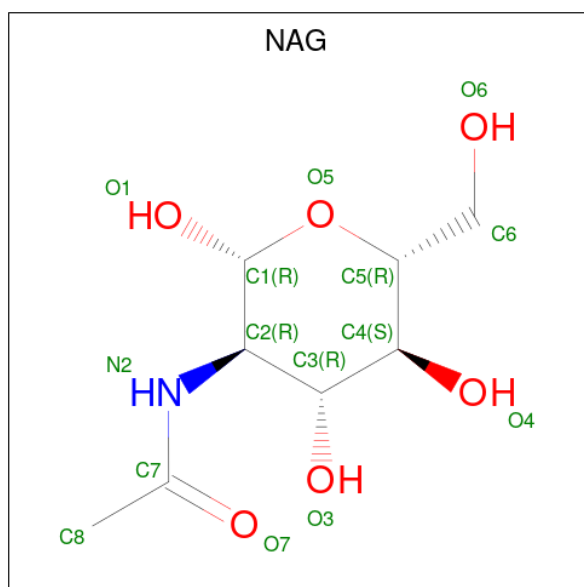
Chain	Residue	Modelled	Actual	Comment	Reference
C	-13	VAL	-	expression tag	UNP P02787
C	-12	PRO	-	expression tag	UNP P02787
C	-11	ASP	-	expression tag	UNP P02787
C	-10	LYS	-	expression tag	UNP P02787
C	-9	HIS	-	expression tag	UNP P02787
C	-8	HIS	-	expression tag	UNP P02787
C	-7	HIS	-	expression tag	UNP P02787
C	-6	HIS	-	expression tag	UNP P02787
C	-5	HIS	-	expression tag	UNP P02787
C	-4	HIS	-	expression tag	UNP P02787
C	-3	ILE	-	expression tag	UNP P02787
C	-2	GLU	-	expression tag	UNP P02787
C	-1	GLY	-	expression tag	UNP P02787
C	0	ARG	-	expression tag	UNP P02787
C	413	ASP	ASN	engineered mutation	UNP P02787
C	426	PHE	TYR	engineered mutation	UNP P02787
C	429	VAL	ILE	SEE REMARK 999	UNP P02787
C	517	PHE	TYR	engineered mutation	UNP P02787
C	611	ASP	ASN	engineered mutation	UNP P02787
D	-13	VAL	-	expression tag	UNP P02787
D	-12	PRO	-	expression tag	UNP P02787
D	-11	ASP	-	expression tag	UNP P02787
D	-10	LYS	-	expression tag	UNP P02787
D	-9	HIS	-	expression tag	UNP P02787
D	-8	HIS	-	expression tag	UNP P02787
D	-7	HIS	-	expression tag	UNP P02787
D	-6	HIS	-	expression tag	UNP P02787

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-5	HIS	-	expression tag	UNP P02787
D	-4	HIS	-	expression tag	UNP P02787
D	-3	ILE	-	expression tag	UNP P02787
D	-2	GLU	-	expression tag	UNP P02787
D	-1	GLY	-	expression tag	UNP P02787
D	0	ARG	-	expression tag	UNP P02787
D	413	ASP	ASN	engineered mutation	UNP P02787
D	426	PHE	TYR	engineered mutation	UNP P02787
D	429	VAL	ILE	SEE REMARK 999	UNP P02787
D	517	PHE	TYR	engineered mutation	UNP P02787
D	611	ASP	ASN	engineered mutation	UNP P02787

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

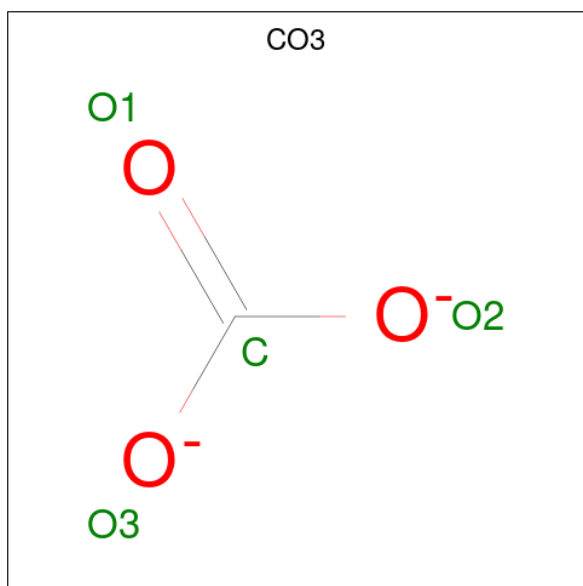
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca	0	0
			1	1		

- Molecule 5 is CARBONATE ION (three-letter code: CO3) (formula: CO<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			4	1	3		
5	D	1	Total	C	O	0	0
			4	1	3		

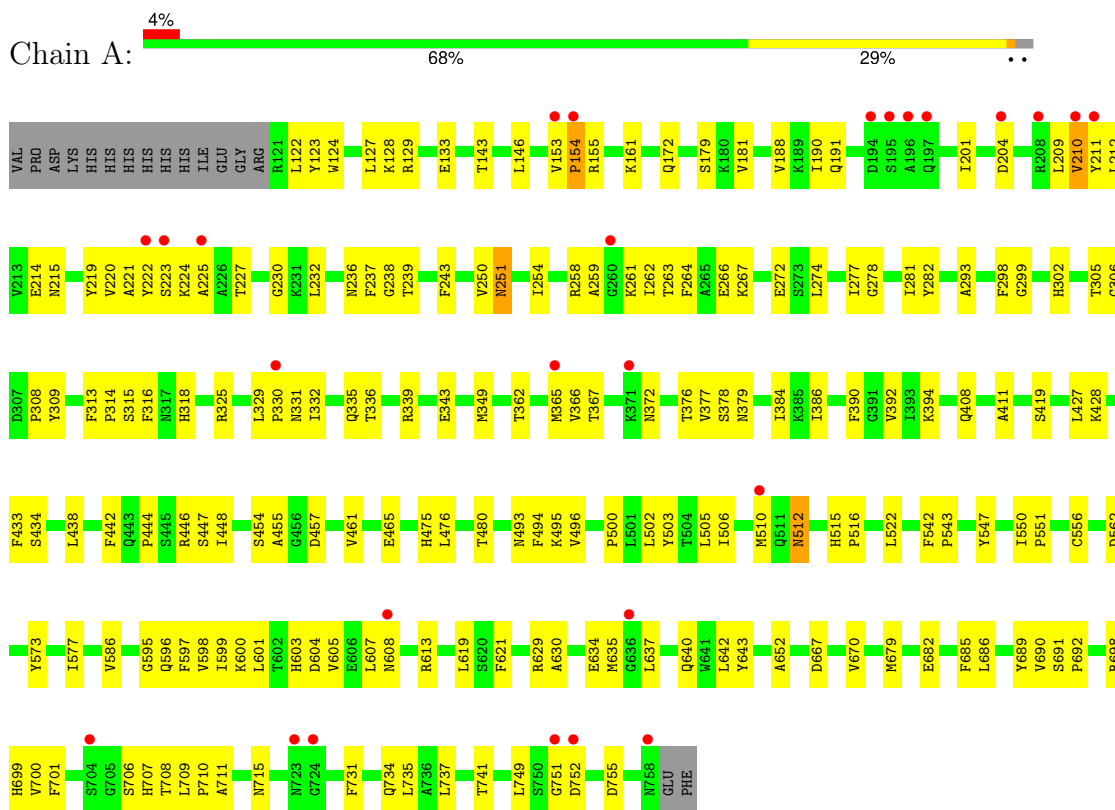
- Molecule 6 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	1	Total	Fe	0	0
			1	1		
6	D	1	Total	Fe	0	0
			1	1		

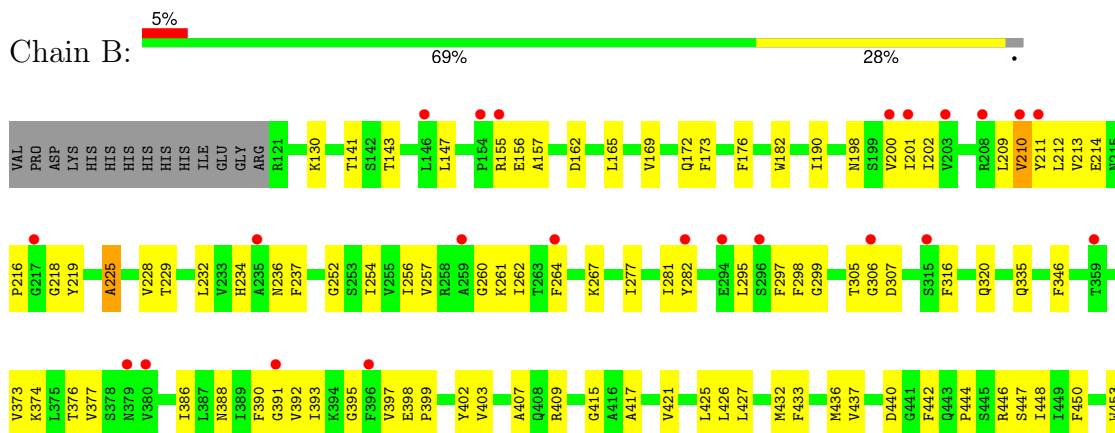
### 3 Residue-property plots

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.

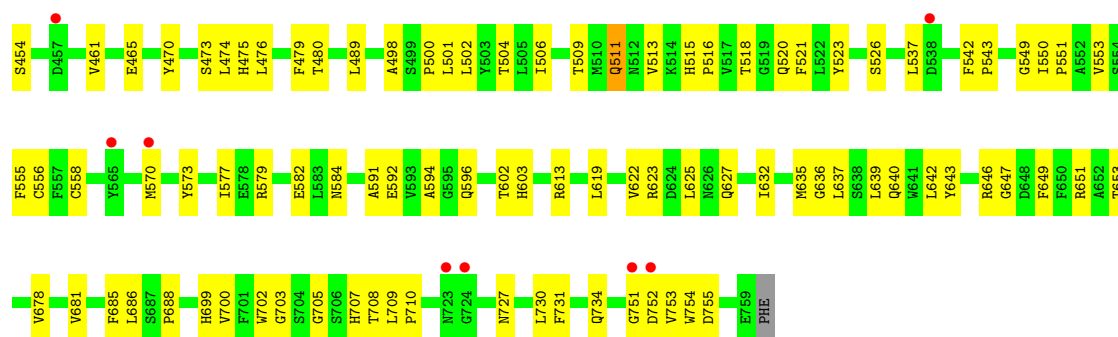
#### • Molecule 1: Transferrin receptor protein 1



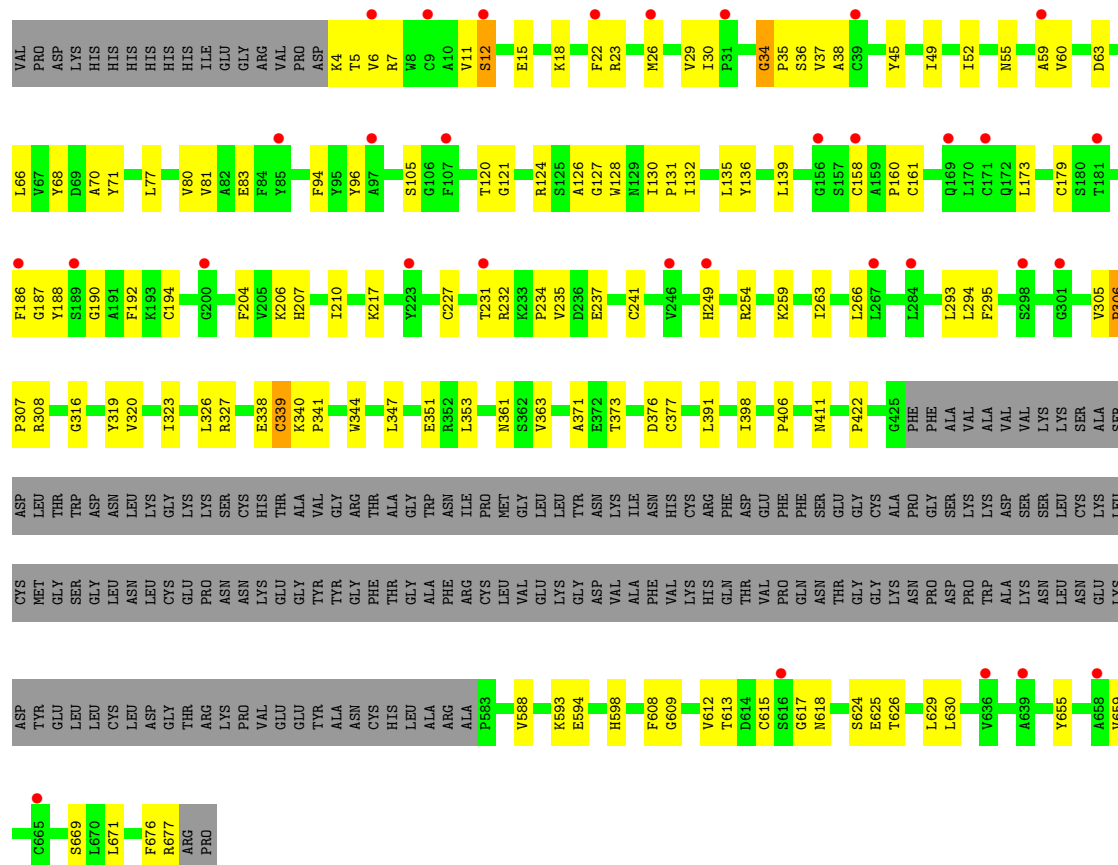
#### • Molecule 1: Transferrin receptor protein 1



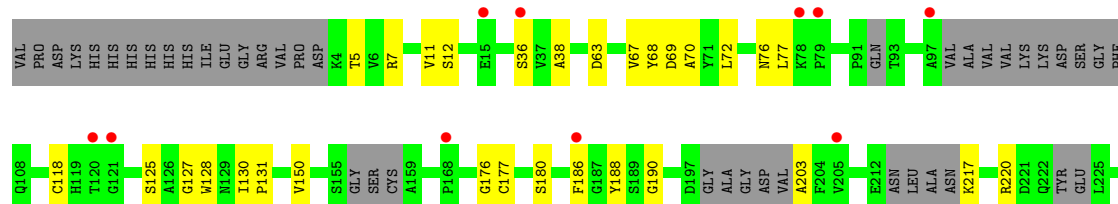


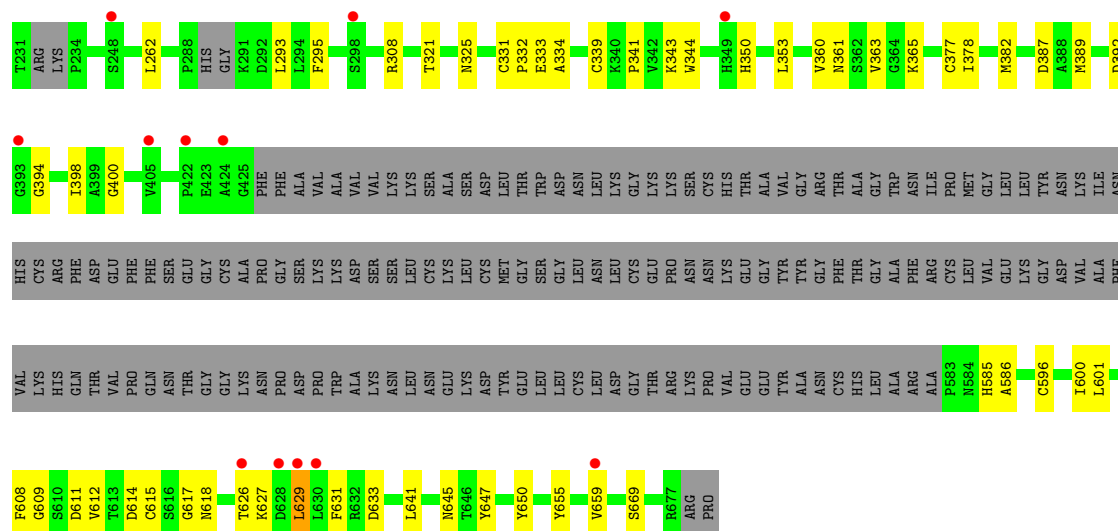


• Molecule 2: Serotransferrin



• Molecule 2: Serotransferrin





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	232.19Å 232.19Å 168.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.00 – 3.32 17.00 – 3.32	Depositor EDS
% Data completeness (in resolution range)	96.6 (17.00-3.32) 95.9 (17.00-3.32)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 3.34Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.282 , 0.332 0.284 , 0.333	Depositor DCC
$R_{free}$ test set	6560 reflections (10.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	77.1	Xtriage
Anisotropy	0.458	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 59.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	16658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.46% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, FE, CA, CO3

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	0.43	0/5036	0.63	0/6862
1	B	0.42	0/4891	0.60	2/6688 (0.0%)
2	C	0.38	0/3760	0.58	0/5134
2	D	0.32	0/3296	0.53	0/4507
All	All	0.39	0/16983	0.59	2/23191 (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	A	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	316	PHE	O-C-N	-7.63	110.49	122.70
1	B	316	PHE	CA-C-N	5.83	130.03	117.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	309	TYR	Sidechain

## 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	4917	0	4680	193	0
1	B	4773	0	4386	158	0
2	C	3682	0	3265	117	0
2	D	3246	0	2625	63	0
3	A	14	0	13	0	0
3	B	14	0	13	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	C	4	0	0	0	0
5	D	4	0	0	4	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
All	All	16658	0	14982	525	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 17.

The worst 5 of 525 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:751:GLY:O	1:A:755:ASP:HB2	1.33	1.29
1:B:489:LEU:HD12	1:B:558:CYS:SG	1.75	1.26
1:A:155:ARG:O	1:A:411:ALA:O	1.57	1.19
1:B:210:VAL:HG12	1:B:211:TYR:N	1.58	1.17
1:A:210:VAL:HG12	1:A:211:TYR:H	1.01	1.10

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	636/654 (97%)	589 (93%)	42 (7%)	5 (1%)	16	47
1	B	637/654 (97%)	594 (93%)	37 (6%)	6 (1%)	14	44
2	C	513/693 (74%)	474 (92%)	33 (6%)	6 (1%)	11	39
2	D	468/693 (68%)	437 (93%)	27 (6%)	4 (1%)	14	44
All	All	2254/2694 (84%)	2094 (93%)	139 (6%)	21 (1%)	14	44

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	629	LEU
2	C	55	ASN
1	A	210	VAL
1	A	251	ASN
1	A	475	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	512/562 (91%)	511 (100%)	1 (0%)	92	95
1	B	473/562 (84%)	471 (100%)	2 (0%)	89	93
2	C	357/585 (61%)	355 (99%)	2 (1%)	84	90
2	D	277/585 (47%)	277 (100%)	0	100	100
All	All	1619/2294 (71%)	1614 (100%)	5 (0%)	91	94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	512	ASN
1	B	511	GLN
1	B	603	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	71	TYR
2	C	241	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	585	HIS
2	C	604	GLN
2	D	585	HIS
1	A	603	HIS
1	A	596	GLN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	NAG	A	903	1	14,14,15	0.43	0	17,19,21	1.15	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	CO3	C	905	6	3,3,3	0.18	0	2,3,3	0.31	0
3	NAG	B	903	1	14,14,15	0.43	0	17,19,21	1.16	2 (11%)
5	CO3	D	905	6	3,3,3	0.15	0	2,3,3	0.18	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	NAG	A	903	1	-	0/6/23/26	0/1/1/1
3	NAG	B	903	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	903	NAG	C8-C7-N2	2.31	119.94	116.12
3	A	903	NAG	C8-C7-N2	2.30	119.94	116.12
3	B	903	NAG	C2-N2-C7	-2.18	119.98	122.90
3	A	903	NAG	C2-N2-C7	-2.18	119.98	122.90

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	905	CO3	4	0

## 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	A	638/654 (97%)	-0.15	26 (4%)	42	31	28, 75, 122, 200	0
1	B	639/654 (97%)	0.26	31 (4%)	36	28	42, 99, 173, 200	0
2	C	517/693 (74%)	0.52	32 (6%)	28	22	61, 116, 172, 200	0
2	D	488/693 (70%)	0.41	22 (4%)	39	29	66, 136, 182, 200	0
All	All	2282/2694 (84%)	0.24	111 (4%)	36	28	28, 103, 175, 200	0

The worst 5 of 111 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	154	PRO	5.6
2	C	6	VAL	5.4
1	A	222	TYR	5.1
1	A	758	ASN	4.6
1	A	751	GLY	4.6

### 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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	903	14/15	0.55	0.21	140,140,140,140	0
5	CO3	D	905	4/4	0.63	0.14	127,127,127,127	0
3	NAG	A	903	14/15	0.66	0.19	116,116,116,116	0
6	FE	C	901	1/1	0.86	0.09	118,118,118,118	0
6	FE	D	901	1/1	0.89	0.07	163,163,163,163	0
5	CO3	C	905	4/4	0.96	0.05	55,55,55,55	0
4	CA	B	900	1/1	0.97	0.03	80,80,80,80	0
4	CA	A	900	1/1	1.00	0.02	61,61,61,61	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.