



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 10:15 pm GMT

PDB ID : 1H76
Title : THE CRYSTAL STRUCTURE OF DIFERRIC PORCINE SERUM TRANS-FERRIN
Authors : Hall, D.R.; Hadden, J.M.; Leonard, G.A.; Bailey, S.; Neu, M.; Winn, M.; Lindley, P.F.
Deposited on : 2001-07-03
Resolution : 2.15 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

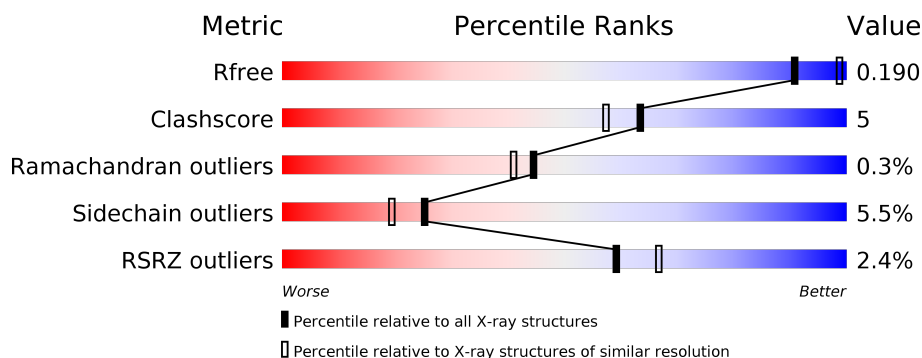
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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}	100719	1170 (2.16-2.16)
Clashscore	112137	1278 (2.16-2.16)
Ramachandran outliers	110173	1256 (2.16-2.16)
Sidechain outliers	110143	1255 (2.16-2.16)
RSRZ outliers	101464	1175 (2.16-2.16)

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 $\leq 5\%$. 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	696	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>...</div> </div> </div>

2 Entry composition [i](#)

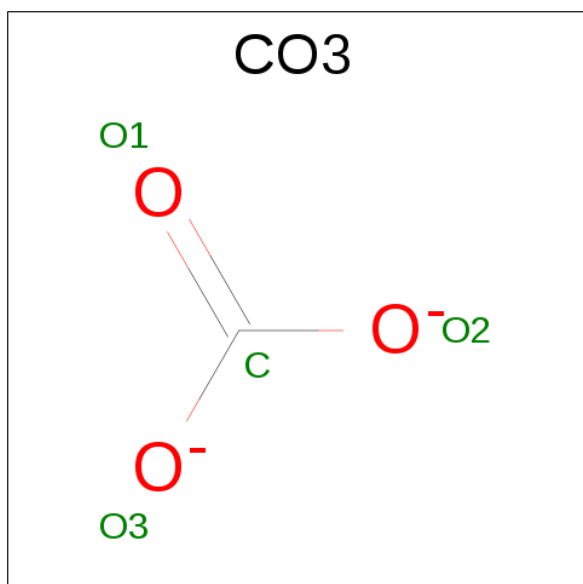
There are 5 unique types of molecules in this entry. The entry contains 5772 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 SEROTRANSFERRIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	677	Total	C	N	O	S	0	5	0
			5254	3276	920	1015	43			

- Molecule 2 is CARBONATE ION (three-letter code: CO3) (formula: CO₃).

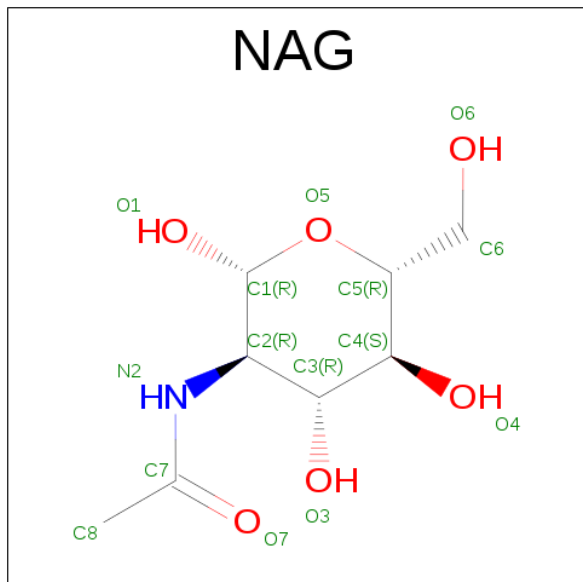


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	1	3		
2	A	1	Total	C	O	0	0
			4	1	3		

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Fe	0	0
			2	2		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

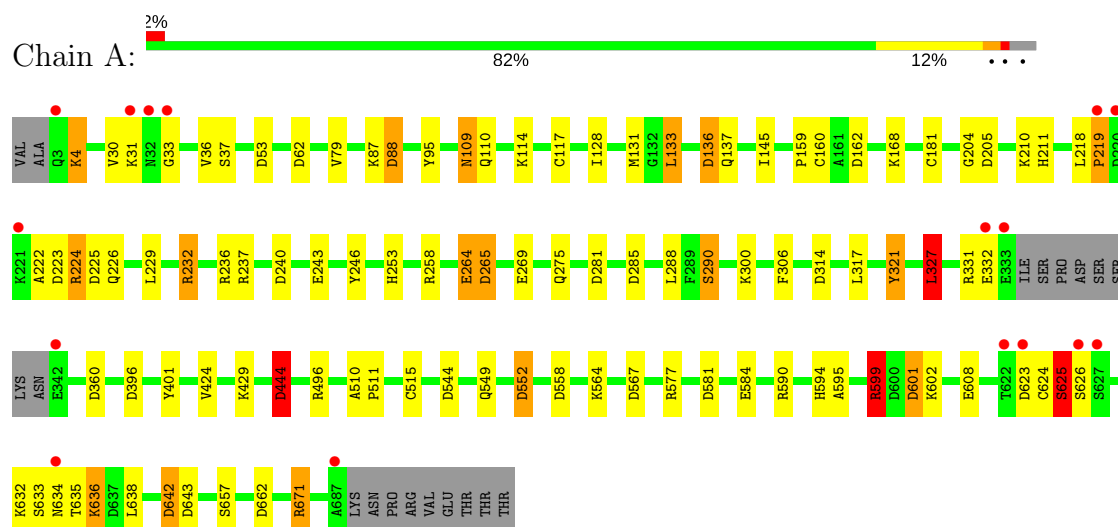
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	494	Total	O	0	0
			494	494		

3 Residue-property plots

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 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: SEROTRANSFERRIN



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	223.77Å 44.89Å 78.86Å 90.00° 105.39° 90.00°	Depositor
Resolution (Å)	30.00 – 2.15 23.33 – 2.15	Depositor EDS
% Data completeness (in resolution range)	96.0 (30.00-2.15) 96.1 (23.33-2.15)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	8.75 (at 2.15Å)	Xtriage
Refinement program	REFMAC 5.0.32	Depositor
R, R_{free}	0.136 , 0.182 0.146 , 0.190	Depositor DCC
R_{free} test set	2001 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	20.6	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 59.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5772	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.17% 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 ⓘ

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

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	1.09	5/5384 (0.1%)	1.10	37/7275 (0.5%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	584	GLU	CD-OE2	12.04	1.38	1.25
1	A	599	ARG	NE-CZ	6.87	1.42	1.33
1	A	37	SER	CB-OG	-5.72	1.34	1.42
1	A	595	ALA	CA-CB	5.18	1.63	1.52
1	A	290	SER	CB-OG	-5.15	1.35	1.42

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	599	ARG	NE-CZ-NH1	17.37	128.99	120.30
1	A	544[A]	ASP	CB-CG-OD2	9.33	126.69	118.30
1	A	544[B]	ASP	CB-CG-OD2	9.33	126.69	118.30
1	A	444[A]	ASP	CB-CG-OD2	8.96	126.36	118.30
1	A	444[B]	ASP	CB-CG-OD2	8.96	126.36	118.30
1	A	552	ASP	CB-CG-OD2	8.48	125.93	118.30
1	A	258	ARG	NE-CZ-NH1	-8.11	116.25	120.30
1	A	599	ARG	NE-CZ-NH2	-7.78	116.41	120.30
1	A	558	ASP	CB-CG-OD2	7.63	125.17	118.30
1	A	360	ASP	CB-CG-OD1	7.41	124.97	118.30
1	A	258	ARG	NE-CZ-NH2	7.18	123.89	120.30
1	A	662	ASP	CB-CG-OD1	6.57	124.21	118.30
1	A	642	ASP	CB-CG-OD2	6.50	124.15	118.30
1	A	599	ARG	CD-NE-CZ	6.41	132.57	123.60
1	A	133	LEU	CB-CG-CD1	6.38	121.84	111.00
1	A	285	ASP	CB-CG-OD2	6.38	124.04	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	444[A]	ASP	CB-CG-OD1	-6.01	112.89	118.30
1	A	444[B]	ASP	CB-CG-OD1	-6.01	112.89	118.30
1	A	265	ASP	CB-CG-OD2	5.95	123.65	118.30
1	A	162	ASP	CB-CG-OD1	5.91	123.62	118.30
1	A	88	ASP	CB-CG-OD2	5.88	123.59	118.30
1	A	577	ARG	NE-CZ-NH2	-5.81	117.40	120.30
1	A	136	ASP	CB-CG-OD2	5.72	123.44	118.30
1	A	281	ASP	CB-CG-OD2	5.70	123.43	118.30
1	A	36	VAL	CB-CA-C	-5.60	100.75	111.40
1	A	601	ASP	CB-CG-OD2	5.54	123.28	118.30
1	A	623	ASP	CB-CG-OD2	5.51	123.26	118.30
1	A	53	ASP	CB-CG-OD2	5.46	123.22	118.30
1	A	314	ASP	CB-CG-OD1	5.46	123.21	118.30
1	A	225	ASP	CB-CG-OD2	5.33	123.09	118.30
1	A	205	ASP	CB-CG-OD2	5.32	123.09	118.30
1	A	131	MET	CG-SD-CE	5.31	108.70	100.20
1	A	232	ARG	NE-CZ-NH1	5.31	122.95	120.30
1	A	327	LEU	CA-CB-CG	5.29	127.47	115.30
1	A	590	ARG	NE-CZ-NH2	-5.20	117.70	120.30
1	A	581	ASP	CB-CG-OD1	5.17	122.95	118.30
1	A	643	ASP	CB-CG-OD2	5.01	122.81	118.30

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	5254	0	5038	51	0
2	A	8	0	0	0	0
3	A	2	0	0	0	0
4	A	14	0	13	0	0
5	A	494	0	0	15	0
All	All	5772	0	5051	51	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 5.

All (51) 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:401:TYR:OH	1:A:671:ARG:NH2	2.05	0.89
1:A:671:ARG:NH1	5:A:2480:HOH:O	1.86	0.85
1:A:635:THR:O	1:A:636:LYS:HD3	1.90	0.72
1:A:624:CYS:O	1:A:625:SER:HB2	1.89	0.71
1:A:634:ASN:HB3	5:A:2443:HOH:O	1.91	0.70
1:A:444[B]:ASP:OD2	5:A:2301:HOH:O	2.10	0.70
1:A:137:GLN:NE2	5:A:2083:HOH:O	2.23	0.69
1:A:275:GLN:HE22	1:A:306:PHE:H	1.38	0.69
1:A:232:ARG:NH1	1:A:246:TYR:O	2.29	0.66
1:A:30:VAL:HG12	1:A:33:GLY:H	1.60	0.66
1:A:608:GLU:HG2	5:A:2136:HOH:O	1.96	0.65
1:A:671:ARG:HB2	1:A:671:ARG:CZ	2.26	0.65
1:A:424:VAL:HG23	5:A:2285:HOH:O	1.95	0.64
1:A:136:ASP:OD1	1:A:137:GLN:HG3	1.99	0.61
1:A:321:TYR:CE1	5:A:2209:HOH:O	2.51	0.61
1:A:552:ASP:OD1	1:A:564:LYS:HD2	1.99	0.61
1:A:128:ILE:HG21	1:A:327:LEU:HD13	1.83	0.59
1:A:62:ASP:HA	1:A:253:HIS:CD2	2.39	0.58
1:A:599:ARG:HB3	1:A:601:ASP:OD1	2.06	0.56
1:A:396:ASP:HA	1:A:594:HIS:CD2	2.41	0.56
1:A:549:GLN:NE2	5:A:2379:HOH:O	2.05	0.53
1:A:275:GLN:NE2	1:A:306:PHE:H	2.05	0.53
1:A:633:SER:OG	1:A:642:ASP:OD2	2.22	0.52
1:A:223:ASP:HB3	1:A:226:GLN:NE2	2.25	0.52
1:A:95:TYR:HB2	1:A:211:HIS:HB3	1.92	0.52
1:A:109:ASN:HD22	1:A:109:ASN:H	1.60	0.50
1:A:567:ASP:OD1	5:A:2390:HOH:O	2.19	0.49
1:A:510:ALA:HA	1:A:511:PRO:HD3	1.73	0.48
1:A:159:PRO:O	1:A:160:CYS:CB	2.61	0.47
1:A:79:VAL:HG11	1:A:264:GLU:HG2	1.96	0.47
1:A:109:ASN:HD22	1:A:109:ASN:N	2.10	0.47
1:A:321:TYR:CE1	1:A:599:ARG:CD	2.99	0.46
1:A:633:SER:HB2	1:A:638:LEU:HB2	1.98	0.46
1:A:4:LYS:NZ	1:A:269:GLU:OE1	2.47	0.46
1:A:222:ALA:C	1:A:224:ARG:H	2.19	0.46
1:A:243:GLU:HG3	5:A:2152:HOH:O	2.16	0.45
1:A:30:VAL:HG11	1:A:33:GLY:HA3	1.98	0.44
1:A:317:LEU:HA	1:A:317:LEU:HD12	1.89	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:LEU:HD23	1:A:219:PRO:HD2	2.00	0.44
1:A:671:ARG:HB2	1:A:671:ARG:NH1	2.34	0.43
1:A:210:LYS:HZ3	1:A:300:LYS:CE	2.32	0.43
1:A:109:ASN:HB3	5:A:2145:HOH:O	2.18	0.42
1:A:88:ASP:N	1:A:88:ASP:OD1	2.41	0.42
1:A:114:LYS:NZ	5:A:2067:HOH:O	2.53	0.42
1:A:634:ASN:ND2	5:A:2441:HOH:O	2.53	0.42
1:A:218:LEU:CD2	1:A:219:PRO:HD2	2.50	0.41
1:A:117:CYS:SG	1:A:204:GLY:HA3	2.60	0.41
1:A:634:ASN:HB3	5:A:2442:HOH:O	2.20	0.41
1:A:634:ASN:CB	5:A:2443:HOH:O	2.60	0.41
1:A:229:LEU:O	1:A:236:ARG:HA	2.20	0.40
1:A:109:ASN:ND2	1:A:110:GLN:HG3	2.37	0.40

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	678/696 (97%)	658 (97%)	18 (3%)	2 (0%)	44 41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	219	PRO
1	A	625	SER

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	569/586 (97%)	537 (94%)	32 (6%)	25	19

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	LYS
1	A	31	LYS
1	A	87	LYS
1	A	109	ASN
1	A	133	LEU
1	A	145	ILE
1	A	168	LYS
1	A	181	CYS
1	A	224	ARG
1	A	237	ARG
1	A	240	ASP
1	A	264	GLU
1	A	265	ASP
1	A	288	LEU
1	A	290	SER
1	A	321	TYR
1	A	327	LEU
1	A	331	ARG
1	A	332	GLU
1	A	429	LYS
1	A	444[A]	ASP
1	A	444[B]	ASP
1	A	496	ARG
1	A	515	CYS
1	A	599	ARG
1	A	602	LYS
1	A	625	SER
1	A	626	SER
1	A	632	LYS
1	A	636	LYS
1	A	657	SER
1	A	671	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	86	GLN
1	A	109	ASN
1	A	185	ASN
1	A	263	GLN
1	A	272	ASN
1	A	273	GLN
1	A	275	GLN
1	A	425	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	$\# Z > 2$
4	NAG	A	1688	1	14,14,15	1.06	1 (7%)	15,19,21	2.83	6 (40%)
2	CO3	A	700	3	0,3,3	0.00	-	0,3,3	0.00	-
2	CO3	A	701	3	0,3,3	0.00	-	0,3,3	0.00	-

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
4	NAG	A	1688	1	-	0/6/23/26	0/1/1/1
2	CO3	A	700	3	-	0/0/0/0	0/0/0/0
2	CO3	A	701	3	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1688	NAG	C1-C2	3.69	1.57	1.52

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1688	NAG	O5-C1-C2	-2.86	107.49	111.47
4	A	1688	NAG	C4-C3-C2	-2.77	106.95	111.02
4	A	1688	NAG	O7-C7-C8	-2.13	118.18	122.06
4	A	1688	NAG	C8-C7-N2	2.10	119.89	116.11
4	A	1688	NAG	C1-C2-N2	4.97	118.98	110.49
4	A	1688	NAG	C1-O5-C5	7.81	122.94	112.17

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](#)

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, 95th 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 ²)	Q<0.9
1	A	677/696 (97%)	-0.30	16 (2%) 59 67	8, 22, 47, 87	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	220	ASP	7.5
1	A	333	GLU	5.6
1	A	634	ASN	4.7
1	A	622	THR	3.8
1	A	219	PRO	3.0
1	A	332	GLU	2.8
1	A	31	LYS	2.7
1	A	221	LYS	2.6
1	A	32	ASN	2.5
1	A	33	GLY	2.5
1	A	342	GLU	2.4
1	A	687	ALA	2.4
1	A	627	SER	2.3
1	A	623	ASP	2.2
1	A	3	GLN	2.1
1	A	626	SER	2.1

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	LLDF	B-factors(Å ²)	Q<0.9
2	CO3	A	700	4/4	0.99	0.13	1.31	13,13,14,14	0
2	CO3	A	701	4/4	1.00	0.07	-1.20	6,7,8,9	0
3	FE	A	703	1/1	1.00	0.06	-2.53	8,8,8,8	0
3	FE	A	702	1/1	1.00	0.06	-2.86	14,14,14,14	0
4	NAG	A	1688	14/15	0.81	0.38	-	64,73,77,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.