



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:04 AM GMT

PDB ID : 3KQ4
Title : Structure of Intrinsic Factor-Cobalamin bound to its receptor Cubilin
Authors : Andersen, C.B.F.; Madsen, M.; Moestrup, S.K.; Andersen, G.R.
Deposited on : 2009-11-17
Resolution : 3.30 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

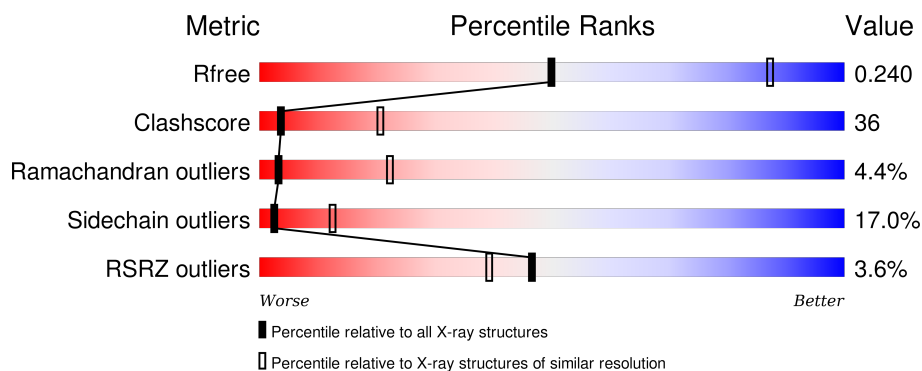
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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}	91344	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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	393	<div> <div>2%</div> <div>53% 35% 10% .</div> </div>
1	C	393	<div> <div>%</div> <div>54% 34% 9% .</div> </div>
1	E	393	<div> <div>2%</div> <div>54% 34% 10% .</div> </div>
2	B	457	<div> <div>5%</div> <div>38% 48% 13% .</div> </div>
2	D	457	<div> <div>5%</div> <div>38% 48% 13% .</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	457	

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
3	NAG	A	2001	X	-	X	X
3	NAG	B	2001	X	-	-	-
3	NAG	B	2008	X	-	-	-
3	NAG	C	2001	X	-	X	-
3	NAG	D	2001	X	-	-	-
3	NAG	D	2008	X	-	-	-
3	NAG	E	2001	X	-	X	-
3	NAG	F	2001	X	-	-	-
3	NAG	F	2008	X	-	-	-
5	B12	A	2007	X	-	X	-
5	B12	C	2007	X	-	X	-
5	B12	E	2007	X	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 20793 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 Gastric intrinsic factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	385	Total	C	N	O	S	3	0	0
			2950	1870	488	573	19			
1	C	385	Total	C	N	O	S	3	0	0
			2950	1870	488	573	19			
1	E	385	Total	C	N	O	S	3	0	0
			2950	1870	488	573	19			

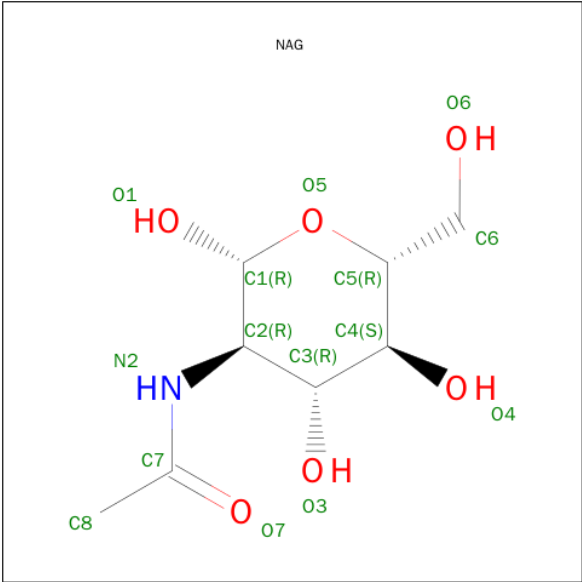
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	91	HIS	GLN	SEE REMARK 999	UNP P27352
C	91	HIS	GLN	SEE REMARK 999	UNP P27352
E	91	HIS	GLN	SEE REMARK 999	UNP P27352

- Molecule 2 is a protein called Cubilin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	457	Total	C	N	O	S	0	0	0
			3638	2311	598	709	20			
2	D	457	Total	C	N	O	S	0	0	0
			3638	2311	598	709	20			
2	F	457	Total	C	N	O	S	0	0	0
			3638	2311	598	709	20			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

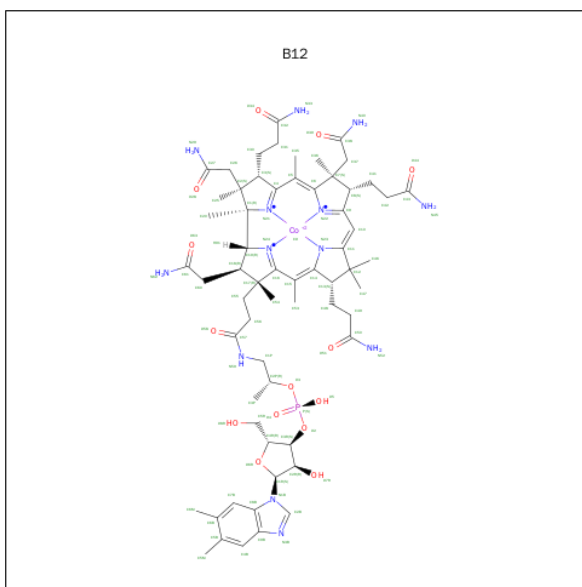
Continued from previous page...

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

- Molecule 4 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	5	Total	C	N	O	0	0
			61	34	2	25		
4	B	5	Total	C	N	O	0	0
			61	34	2	25		
4	C	5	Total	C	N	O	0	0
			61	34	2	25		
4	D	5	Total	C	N	O	0	0
			61	34	2	25		
4	E	5	Total	C	N	O	0	0
			61	34	2	25		
4	F	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 5 is COBALAMIN (three-letter code: B12) (formula: C₆₂H₈₉CoN₁₃O₁₄P).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
5	C	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
5	E	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	2	Total	C	N	O	0	0
			28	16	2	10		
6	D	2	Total	C	N	O	0	0
			28	16	2	10		
6	F	2	Total	C	N	O	0	0
			28	16	2	10		

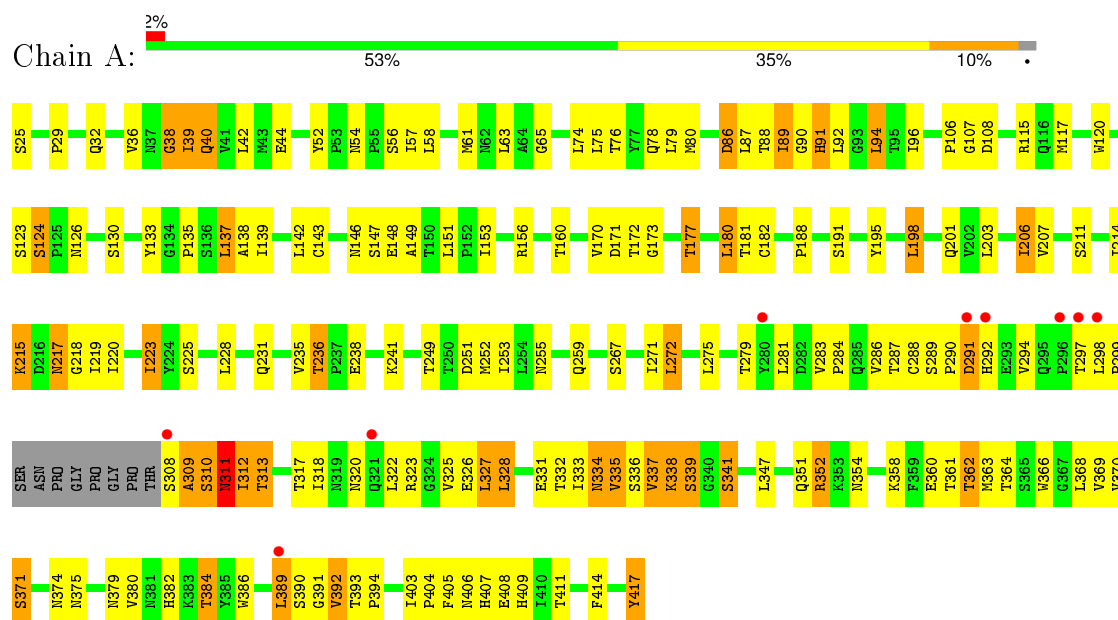
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	4	Total	Ca	0	0
			4	4		
7	D	4	Total	Ca	0	0
			4	4		
7	F	4	Total	Ca	0	0
			4	4		

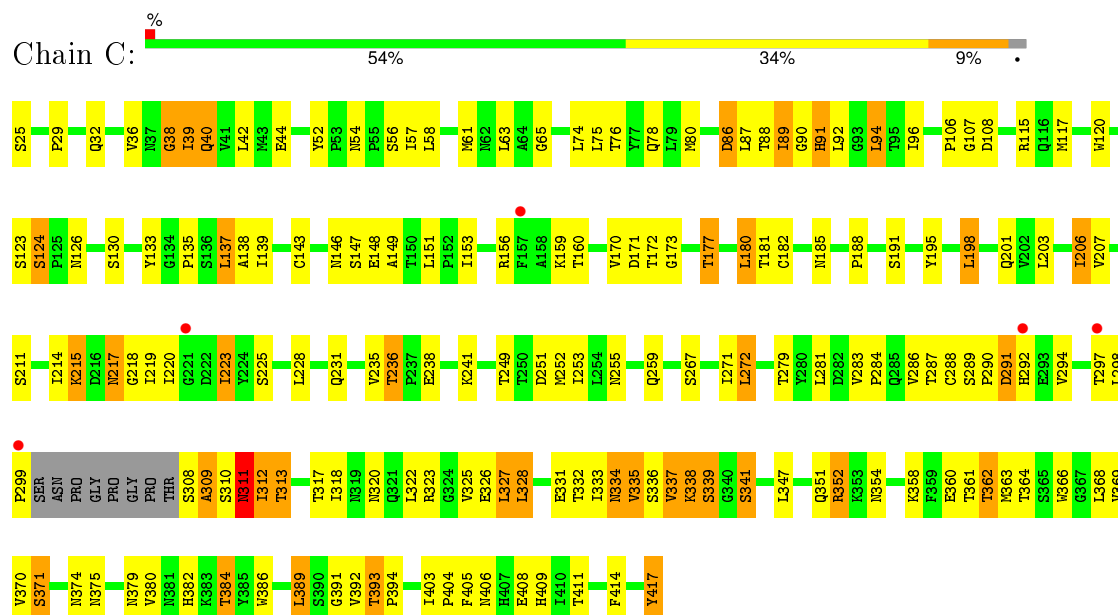
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

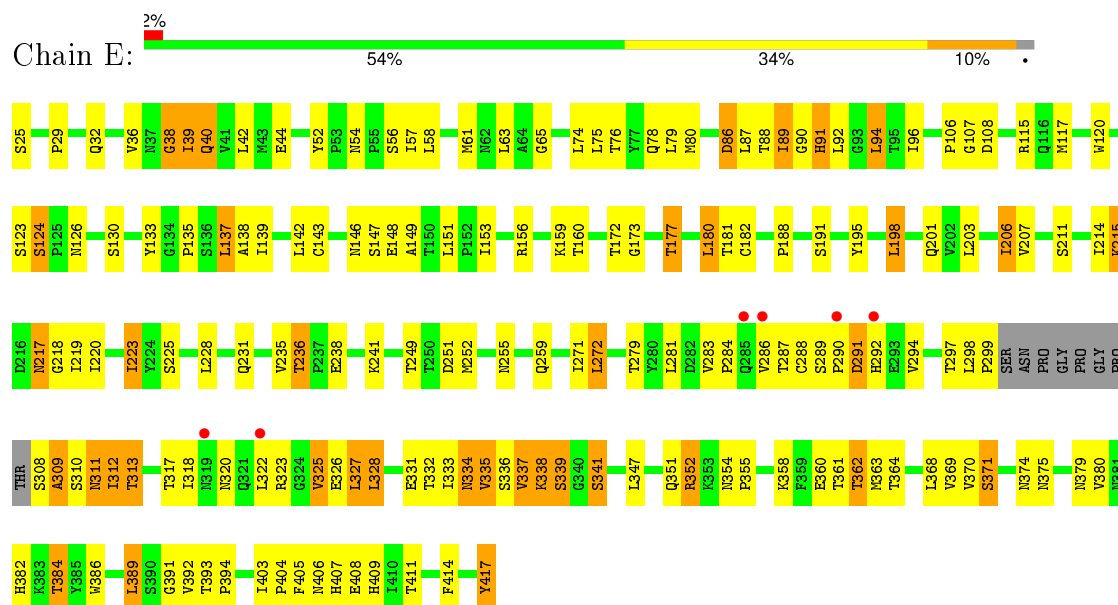
• Molecule 1: Gastric intrinsic factor

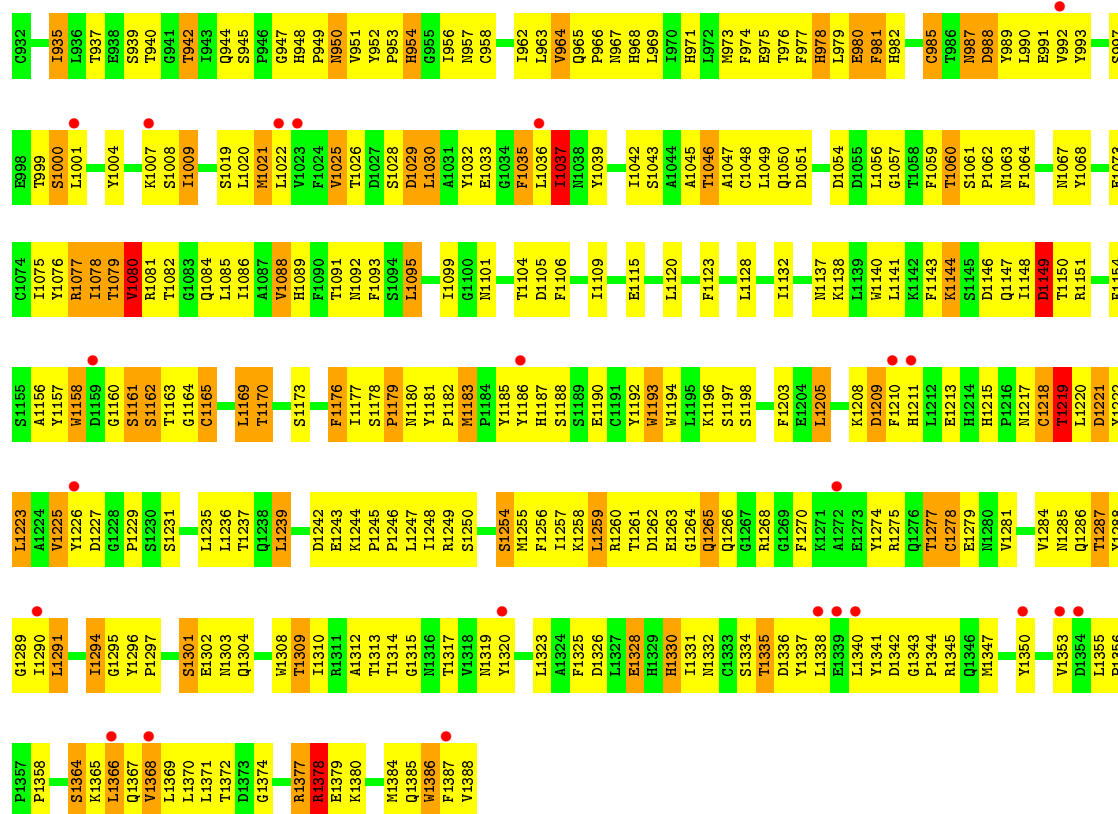


• Molecule 1: Gastric intrinsic factor

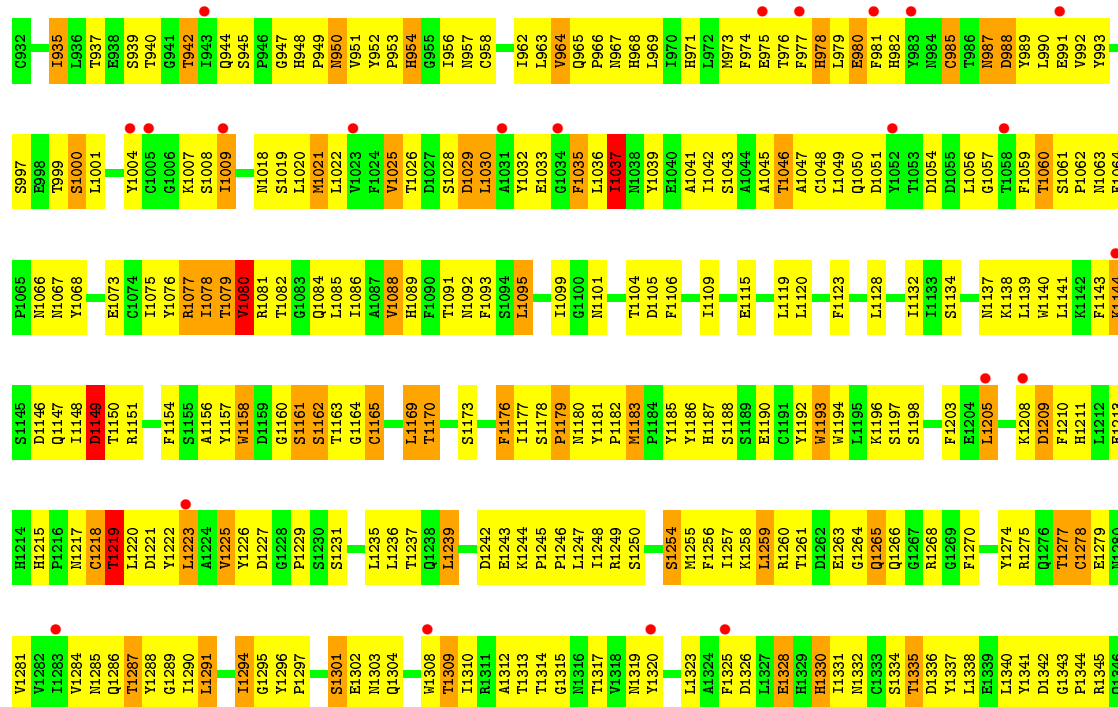


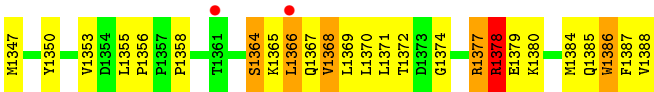
- Molecule 1: Gastric intrinsic factor





• Molecule 2: Cubilin





4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	117.68Å 204.18Å 410.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.85 – 3.30 47.85 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.0 (47.85-3.30) 96.6 (47.85-3.30)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.71 (at 3.33Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.211 , 0.242 0.203 , 0.240	Depositor DCC
R_{free} test set	1105 reflections (1.53%)	DCC
Wilson B-factor (Å ²)	93.9	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 69.7	EDS
Estimated twinning fraction	0.428 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.437 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	2 of 72280 reflections (0.003%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	20793	wwPDB-VP
Average B, all atoms (Å ²)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CA, BMA, B12, NAG, MAN

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	0.67	1/3007 (0.0%)	0.80	1/4090 (0.0%)
1	C	0.67	1/3007 (0.0%)	0.80	1/4090 (0.0%)
1	E	0.67	1/3007 (0.0%)	0.80	0/4090
2	B	0.55	0/3748	0.74	0/5110
2	D	0.55	0/3748	0.74	0/5110
2	F	0.55	0/3748	0.74	0/5110
All	All	0.60	3/20265 (0.0%)	0.77	2/27600 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	360	GLU	CG-CD	5.96	1.60	1.51
1	A	360	GLU	CG-CD	5.94	1.60	1.51
1	C	360	GLU	CG-CD	5.92	1.60	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	311	ASN	CB-CA-C	5.42	121.24	110.40
1	A	311	ASN	CB-CA-C	5.12	120.65	110.40

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	2950	0	2961	150	0
1	C	2950	0	2961	149	0
1	E	2950	0	2961	147	0
2	B	3638	0	3373	308	0
2	D	3638	0	3373	305	0
2	F	3638	0	3373	309	0
3	A	14	0	13	7	0
3	B	84	0	78	7	0
3	C	14	0	13	8	0
3	D	84	0	78	7	0
3	E	14	0	13	7	0
3	F	84	0	78	8	0
4	A	61	0	52	3	0
4	B	61	0	52	5	0
4	C	61	0	52	3	0
4	D	61	0	52	5	0
4	E	61	0	52	4	0
4	F	61	0	52	5	0
5	A	91	0	87	29	0
5	C	91	0	87	26	0
5	E	91	0	87	25	0
6	B	28	0	25	3	0
6	D	28	0	25	3	0
6	F	28	0	25	3	0
7	B	4	0	0	0	0
7	D	4	0	0	0	0
7	F	4	0	0	0	0
All	All	20793	0	19923	1450	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 36.

The worst 5 of 1450 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:311:ASN:HB2	1:C:337:VAL:HA	1.19	1.18
1:A:311:ASN:HB2	1:A:337:VAL:HA	1.19	1.11
1:E:223:ILE:HD12	1:E:223:ILE:H	1.10	1.10
1:E:311:ASN:HB2	1:E:337:VAL:HA	1.19	1.08
1:A:223:ILE:H	1:A:223:ILE:HD12	1.10	1.07

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	381/393 (97%)	334 (88%)	37 (10%)	10 (3%)	7	36
1	C	381/393 (97%)	334 (88%)	37 (10%)	10 (3%)	7	36
1	E	381/393 (97%)	334 (88%)	37 (10%)	10 (3%)	7	36
2	B	455/457 (100%)	376 (83%)	52 (11%)	27 (6%)	2	15
2	D	455/457 (100%)	376 (83%)	52 (11%)	27 (6%)	2	15
2	F	455/457 (100%)	376 (83%)	52 (11%)	27 (6%)	2	15
All	All	2508/2550 (98%)	2130 (85%)	267 (11%)	111 (4%)	3	22

5 of 111 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	107	GLY
1	A	309	ALA
2	B	1105	ASP
2	B	1162	SER
2	B	1219	THR

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	337/343 (98%)	283 (84%)	54 (16%)	3	14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	337/343 (98%)	283 (84%)	54 (16%)	3	14
1	E	337/343 (98%)	282 (84%)	55 (16%)	3	14
2	B	406/406 (100%)	334 (82%)	72 (18%)	2	10
2	D	406/406 (100%)	334 (82%)	72 (18%)	2	10
2	F	406/406 (100%)	335 (82%)	71 (18%)	2	11
All	All	2229/2247 (99%)	1851 (83%)	378 (17%)	2	12

5 of 378 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	352	ARG
2	D	1091	THR
2	F	1193	TRP
1	C	384	THR
2	D	980	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	965	GLN
2	D	1232	ASN
2	F	1265	GLN
2	D	967	ASN
2	D	978	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

36 carbohydrates are modelled in this entry.

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	2002	1,4	14,14,15	0.75	0	15,19,21	2.64	4 (26%)
4	NAG	A	2003	4	14,14,15	0.99	1 (7%)	15,19,21	2.31	8 (53%)
4	BMA	A	2004	4	11,11,12	0.82	0	14,15,17	1.38	2 (14%)
4	MAN	A	2005	4	11,11,12	1.24	2 (18%)	14,15,17	2.34	8 (57%)
4	MAN	A	2006	4	11,11,12	1.26	2 (18%)	14,15,17	2.23	7 (50%)
4	NAG	B	2002	2,4	14,14,15	0.75	0	15,19,21	3.46	8 (53%)
4	NAG	B	2003	4	14,14,15	0.35	0	15,19,21	1.01	1 (6%)
4	BMA	B	2004	4	11,11,12	0.60	0	14,15,17	2.34	5 (35%)
4	MAN	B	2005	4	11,11,12	1.18	2 (18%)	14,15,17	3.23	9 (64%)
4	MAN	B	2006	4	11,11,12	1.11	1 (9%)	14,15,17	1.95	4 (28%)
6	NAG	B	2010	2,6	14,14,15	0.55	0	15,19,21	1.70	1 (6%)
6	NAG	B	2011	6	14,14,15	0.93	1 (7%)	15,19,21	1.64	3 (20%)
4	NAG	C	2002	1,4	14,14,15	0.76	0	15,19,21	2.65	4 (26%)
4	NAG	C	2003	4	14,14,15	1.00	1 (7%)	15,19,21	2.31	8 (53%)
4	BMA	C	2004	4	11,11,12	0.82	0	14,15,17	1.38	2 (14%)
4	MAN	C	2005	4	11,11,12	1.24	2 (18%)	14,15,17	2.34	8 (57%)
4	MAN	C	2006	4	11,11,12	1.26	2 (18%)	14,15,17	2.23	7 (50%)
4	NAG	D	2002	2,4	14,14,15	0.75	0	15,19,21	3.46	8 (53%)
4	NAG	D	2003	4	14,14,15	0.36	0	15,19,21	1.01	1 (6%)
4	BMA	D	2004	4	11,11,12	0.60	0	14,15,17	2.35	5 (35%)
4	MAN	D	2005	4	11,11,12	1.18	2 (18%)	14,15,17	3.23	9 (64%)
4	MAN	D	2006	4	11,11,12	1.11	1 (9%)	14,15,17	1.94	3 (21%)
6	NAG	D	2010	2,6	14,14,15	0.54	0	15,19,21	1.71	1 (6%)
6	NAG	D	2011	6	14,14,15	0.93	1 (7%)	15,19,21	1.65	3 (20%)
4	NAG	E	2002	1,4	14,14,15	0.75	0	15,19,21	2.64	4 (26%)
4	NAG	E	2003	4	14,14,15	0.99	1 (7%)	15,19,21	2.30	8 (53%)
4	BMA	E	2004	4	11,11,12	0.82	0	14,15,17	1.38	2 (14%)
4	MAN	E	2005	4	11,11,12	1.24	2 (18%)	14,15,17	2.35	8 (57%)
4	MAN	E	2006	4	11,11,12	1.26	2 (18%)	14,15,17	2.24	7 (50%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	NAG	F	2002	2,4	14,14,15	0.75	0	15,19,21	3.45	8 (53%)
4	NAG	F	2003	4	14,14,15	0.35	0	15,19,21	1.01	1 (6%)
4	BMA	F	2004	4	11,11,12	0.60	0	14,15,17	2.35	5 (35%)
4	MAN	F	2005	4	11,11,12	1.19	2 (18%)	14,15,17	3.23	9 (64%)
4	MAN	F	2006	4	11,11,12	1.11	1 (9%)	14,15,17	1.95	3 (21%)
6	NAG	F	2010	2,6	14,14,15	0.55	0	15,19,21	1.70	1 (6%)
6	NAG	F	2011	6	14,14,15	0.92	1 (7%)	15,19,21	1.64	3 (20%)

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	2002	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	2003	4	-	0/6/23/26	0/1/1/1
4	BMA	A	2004	4	-	0/2/19/22	0/1/1/1
4	MAN	A	2005	4	-	0/2/19/22	0/1/1/1
4	MAN	A	2006	4	-	0/2/19/22	0/1/1/1
4	NAG	B	2002	2,4	-	0/6/23/26	0/1/1/1
4	NAG	B	2003	4	-	0/6/23/26	0/1/1/1
4	BMA	B	2004	4	-	0/2/19/22	0/1/1/1
4	MAN	B	2005	4	-	0/2/19/22	0/1/1/1
4	MAN	B	2006	4	-	0/2/19/22	0/1/1/1
6	NAG	B	2010	2,6	-	0/6/23/26	0/1/1/1
6	NAG	B	2011	6	-	0/6/23/26	0/1/1/1
4	NAG	C	2002	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	2003	4	-	0/6/23/26	0/1/1/1
4	BMA	C	2004	4	-	0/2/19/22	0/1/1/1
4	MAN	C	2005	4	-	0/2/19/22	0/1/1/1
4	MAN	C	2006	4	-	0/2/19/22	0/1/1/1
4	NAG	D	2002	2,4	-	0/6/23/26	0/1/1/1
4	NAG	D	2003	4	-	0/6/23/26	0/1/1/1
4	BMA	D	2004	4	-	0/2/19/22	0/1/1/1
4	MAN	D	2005	4	-	0/2/19/22	0/1/1/1
4	MAN	D	2006	4	-	0/2/19/22	0/1/1/1
6	NAG	D	2010	2,6	-	0/6/23/26	0/1/1/1
6	NAG	D	2011	6	-	0/6/23/26	0/1/1/1
4	NAG	E	2002	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	2003	4	-	0/6/23/26	0/1/1/1
4	BMA	E	2004	4	-	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	MAN	E	2005	4	-	0/2/19/22	0/1/1/1
4	MAN	E	2006	4	-	0/2/19/22	0/1/1/1
4	NAG	F	2002	2,4	-	0/6/23/26	0/1/1/1
4	NAG	F	2003	4	-	0/6/23/26	0/1/1/1
4	BMA	F	2004	4	-	0/2/19/22	0/1/1/1
4	MAN	F	2005	4	-	0/2/19/22	0/1/1/1
4	MAN	F	2006	4	-	0/2/19/22	0/1/1/1
6	NAG	F	2010	2,6	-	0/6/23/26	0/1/1/1
6	NAG	F	2011	6	-	0/6/23/26	0/1/1/1

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	2005	MAN	C1-C2	2.06	1.57	1.52
4	F	2005	MAN	C1-C2	2.06	1.57	1.52
4	D	2005	MAN	C1-C2	2.07	1.57	1.52
4	D	2006	MAN	C1-C2	2.08	1.57	1.52
4	A	2006	MAN	O5-C1	2.08	1.47	1.43

The worst 5 of 178 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	2010	NAG	C2-N2-C7	-5.53	115.94	123.04
6	F	2010	NAG	C2-N2-C7	-5.50	115.97	123.04
6	B	2010	NAG	C2-N2-C7	-5.49	115.98	123.04
4	D	2002	NAG	O4-C4-C3	-4.01	101.30	110.34
4	B	2002	NAG	O4-C4-C3	-4.01	101.31	110.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

24 monomers are involved in 34 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2004	BMA	3	0
4	A	2006	MAN	3	0
4	B	2002	NAG	2	0
4	B	2003	NAG	1	0
4	B	2004	BMA	2	0
4	B	2006	MAN	3	0
6	B	2010	NAG	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	2011	NAG	2	0
4	C	2004	BMA	3	0
4	C	2006	MAN	3	0
4	D	2002	NAG	2	0
4	D	2003	NAG	1	0
4	D	2004	BMA	2	0
4	D	2006	MAN	3	0
6	D	2010	NAG	3	0
6	D	2011	NAG	2	0
4	E	2004	BMA	4	0
4	E	2006	MAN	4	0
4	F	2002	NAG	2	0
4	F	2003	NAG	1	0
4	F	2004	BMA	2	0
4	F	2006	MAN	3	0
6	F	2010	NAG	3	0
6	F	2011	NAG	2	0

5.6 Ligand geometry ⓘ

Of 36 ligands modelled in this entry, 12 are monoatomic - leaving 24 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	2001	1	14,14,15	1.09	2 (14%)	15,19,21	1.85	5 (33%)
5	B12	A	2007	-	74,101,101	1.36	7 (9%)	111,166,166	2.36	28 (25%)
3	NAG	B	2001	2	14,14,15	0.95	1 (7%)	15,19,21	1.44	3 (20%)
3	NAG	B	2007	2	14,14,15	0.53	0	15,19,21	1.14	2 (13%)
3	NAG	B	2008	2	14,14,15	0.57	0	15,19,21	0.84	0
3	NAG	B	2009	2	14,14,15	0.47	0	15,19,21	0.98	1 (6%)
3	NAG	B	2012	2	14,14,15	0.63	0	15,19,21	1.92	2 (13%)
3	NAG	B	2013	2	14,14,15	0.66	0	15,19,21	1.49	3 (20%)
3	NAG	C	2001	1	14,14,15	1.09	1 (7%)	15,19,21	1.90	5 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	B12	C	2007	-	74,101,101	1.36	7 (9%)	111,166,166	2.36	28 (25%)
3	NAG	D	2001	2	14,14,15	0.94	1 (7%)	15,19,21	1.45	3 (20%)
3	NAG	D	2007	2	14,14,15	0.52	0	15,19,21	1.14	2 (13%)
3	NAG	D	2008	2	14,14,15	0.56	0	15,19,21	0.84	0
3	NAG	D	2009	2	14,14,15	0.46	0	15,19,21	0.98	1 (6%)
3	NAG	D	2012	2	14,14,15	0.58	0	15,19,21	1.51	3 (20%)
3	NAG	D	2013	2	14,14,15	0.61	0	15,19,21	1.59	3 (20%)
3	NAG	E	2001	1	14,14,15	1.10	2 (14%)	15,19,21	1.85	5 (33%)
5	B12	E	2007	-	74,101,101	1.36	7 (9%)	111,166,166	2.36	29 (26%)
3	NAG	F	2001	2	14,14,15	0.96	1 (7%)	15,19,21	1.45	3 (20%)
3	NAG	F	2007	2	14,14,15	0.53	0	15,19,21	1.14	2 (13%)
3	NAG	F	2008	2	14,14,15	0.56	0	15,19,21	0.84	0
3	NAG	F	2009	2	14,14,15	0.46	0	15,19,21	0.98	1 (6%)
3	NAG	F	2012	2	14,14,15	0.56	0	15,19,21	1.65	3 (20%)
3	NAG	F	2013	2	14,14,15	0.64	0	15,19,21	1.41	2 (13%)

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	2001	1	1/1/5/7	0/6/23/26	0/1/1/1
5	B12	A	2007	-	1/1/36/38	0/51/223/223	0/3/11/11
3	NAG	B	2001	2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	B	2007	2	-	0/6/23/26	0/1/1/1
3	NAG	B	2008	2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	B	2009	2	-	0/6/23/26	0/1/1/1
3	NAG	B	2012	2	-	0/6/23/26	0/1/1/1
3	NAG	B	2013	2	-	0/6/23/26	0/1/1/1
3	NAG	C	2001	1	1/1/5/7	1/6/23/26	0/1/1/1
5	B12	C	2007	-	1/1/36/38	0/51/223/223	0/3/11/11
3	NAG	D	2001	2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	D	2007	2	-	0/6/23/26	0/1/1/1
3	NAG	D	2008	2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	D	2009	2	-	0/6/23/26	0/1/1/1
3	NAG	D	2012	2	-	0/6/23/26	0/1/1/1
3	NAG	D	2013	2	-	0/6/23/26	0/1/1/1
3	NAG	E	2001	1	1/1/5/7	1/6/23/26	0/1/1/1
5	B12	E	2007	-	1/1/36/38	0/51/223/223	0/3/11/11

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	F	2001	2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	F	2007	2	-	0/6/23/26	0/1/1/1
3	NAG	F	2008	2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	F	2009	2	-	0/6/23/26	0/1/1/1
3	NAG	F	2012	2	-	0/6/23/26	0/1/1/1
3	NAG	F	2013	2	-	0/6/23/26	0/1/1/1

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	2007	B12	C8B-N1B	-6.00	1.30	1.38
5	A	2007	B12	C8B-N1B	-5.97	1.30	1.38
5	E	2007	B12	C8B-N1B	-5.97	1.30	1.38
5	E	2007	B12	C11-C10	-4.55	1.33	1.41
5	A	2007	B12	C11-C10	-4.54	1.33	1.41

The worst 5 of 134 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	2007	B12	C46-C12-C13	-7.21	82.36	112.81
5	E	2007	B12	C46-C12-C13	-7.21	82.36	112.81
5	C	2007	B12	C46-C12-C13	-7.21	82.37	112.81
5	C	2007	B12	C20-C1-C19	-6.52	102.98	109.38
5	E	2007	B12	C20-C1-C19	-6.48	103.03	109.38

5 of 12 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	2008	NAG	C1
3	B	2001	NAG	C1
5	A	2007	B12	C19
3	E	2001	NAG	C1
5	E	2007	B12	C19

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2001	NAG	O7-C7-N2-C2
3	E	2001	NAG	O7-C7-N2-C2

There are no ring outliers.

18 monomers are involved in 124 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2001	NAG	7	0
5	A	2007	B12	29	0
3	B	2001	NAG	1	0
3	B	2008	NAG	2	0
3	B	2009	NAG	2	0
3	B	2013	NAG	2	0
3	C	2001	NAG	8	0
5	C	2007	B12	26	0
3	D	2001	NAG	1	0
3	D	2008	NAG	2	0
3	D	2009	NAG	2	0
3	D	2013	NAG	2	0
3	E	2001	NAG	7	0
5	E	2007	B12	25	0
3	F	2001	NAG	1	0
3	F	2008	NAG	2	0
3	F	2009	NAG	3	0
3	F	2013	NAG	2	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, 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	385/393 (97%)	0.27	9 (2%) 64 57	47, 77, 133, 281	1 (0%)
1	C	385/393 (97%)	0.18	5 (1%) 79 74	47, 76, 139, 284	1 (0%)
1	E	385/393 (97%)	0.27	6 (1%) 74 69	50, 77, 140, 256	1 (0%)
2	B	457/457 (100%)	0.31	25 (5%) 29 23	59, 111, 172, 219	0
2	D	457/457 (100%)	0.25	23 (5%) 32 26	59, 110, 176, 228	0
2	F	457/457 (100%)	0.35	24 (5%) 30 24	61, 111, 178, 259	0
All	All	2526/2550 (99%)	0.28	92 (3%) 46 39	47, 95, 170, 284	3 (0%)

The worst 5 of 92 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	292	HIS	13.5
1	A	308	SER	7.3
1	C	299	PRO	5.3
1	E	292	HIS	4.8
2	F	983	TYR	4.7

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](#)

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(\AA^2)	Q<0.9
4	NAG	A	2002	14/15	0.96	0.26	-1.11	58,73,97,103	0
4	NAG	E	2002	14/15	0.95	0.22	-1.60	53,78,96,106	0
4	NAG	C	2002	14/15	0.95	0.23	-1.89	59,75,86,91	0
4	NAG	B	2002	14/15	0.73	0.21	-	113,122,130,135	0
4	BMA	B	2004	11/12	0.83	0.10	-	167,178,184,188	0
4	BMA	C	2004	11/12	0.83	0.15	-	138,144,154,155	0
4	MAN	A	2005	11/12	0.59	0.24	-	162,165,175,180	0
4	MAN	B	2005	11/12	0.68	0.24	-	181,194,208,211	0
6	NAG	F	2010	14/15	0.88	0.23	-	119,136,140,142	0
6	NAG	B	2011	14/15	0.81	0.19	-	150,155,171,178	0
4	MAN	C	2006	11/12	0.67	0.23	-	196,201,205,206	0
4	NAG	E	2003	14/15	0.91	0.21	-	89,97,107,109	0
4	NAG	D	2002	14/15	0.79	0.20	-	107,122,135,136	0
4	BMA	A	2004	11/12	0.57	0.21	-	128,134,138,144	0
6	NAG	F	2011	14/15	0.77	0.25	-	160,167,171,175	0
4	MAN	D	2006	11/12	0.69	0.20	-	190,201,222,227	0
4	NAG	A	2003	14/15	0.92	0.24	-	90,95,103,104	0
4	MAN	D	2005	11/12	0.83	0.15	-	172,185,203,208	0
4	MAN	F	2005	11/12	0.69	0.24	-	188,194,206,206	0
4	NAG	D	2003	14/15	0.82	0.18	-	151,172,184,186	0
4	MAN	A	2006	11/12	0.72	0.23	-	190,195,201,202	0
4	BMA	E	2004	11/12	0.76	0.19	-	129,131,147,153	0
4	BMA	F	2004	11/12	0.52	0.17	-	173,182,191,193	0
4	BMA	D	2004	11/12	0.82	0.12	-	167,179,190,190	0
4	MAN	F	2006	11/12	0.63	0.21	-	192,201,209,213	0
4	NAG	F	2002	14/15	0.68	0.24	-	119,123,127,130	0
6	NAG	D	2011	14/15	0.78	0.25	-	163,170,176,178	0
4	MAN	C	2005	11/12	0.64	0.26	-	160,166,178,180	0
4	MAN	E	2005	11/12	0.89	0.17	-	169,175,195,202	0
4	NAG	F	2003	14/15	0.83	0.20	-	163,175,188,193	0
4	MAN	E	2006	11/12	0.70	0.21	-	181,190,197,200	0
6	NAG	B	2010	14/15	0.86	0.23	-	121,132,136,139	0
6	NAG	D	2010	14/15	0.89	0.19	-	125,138,141,143	0
4	NAG	B	2003	14/15	0.79	0.20	-	159,171,180,183	0
4	MAN	B	2006	11/12	0.75	0.22	-	184,197,219,223	0
4	NAG	C	2003	14/15	0.94	0.20	-	89,98,103,104	0

6.4 Ligands

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(\AA^2)	Q<0.9
3	NAG	A	2001	14/15	0.89	0.32	3.92	50,54,57,60	14
3	NAG	C	2001	14/15	0.86	0.28	1.64	45,49,52,54	14
3	NAG	E	2001	14/15	0.88	0.29	0.91	50,55,60,61	14
7	CA	D	2017	1/1	0.97	0.22	0.43	79,79,79,79	0
5	B12	C	2007	91/91	0.97	0.28	0.20	46,65,81,88	0
7	CA	F	2017	1/1	0.94	0.22	0.06	80,80,80,80	0
5	B12	E	2007	91/91	0.96	0.29	0.05	40,66,91,95	0
5	B12	A	2007	91/91	0.95	0.29	0.01	45,64,92,100	0
7	CA	D	2015	1/1	0.97	0.23	-0.06	60,60,60,60	0
7	CA	B	2015	1/1	0.99	0.19	-0.58	54,54,54,54	0
7	CA	B	2016	1/1	0.96	0.12	-0.98	147,147,147,147	0
7	CA	D	2016	1/1	0.98	0.10	-1.07	139,139,139,139	0
7	CA	F	2016	1/1	0.93	0.10	-1.26	144,144,144,144	0
7	CA	B	2017	1/1	0.96	0.14	-1.42	82,82,82,82	0
7	CA	B	2014	1/1	0.81	0.14	-1.71	140,140,140,140	0
7	CA	F	2014	1/1	0.77	0.06	-1.72	151,151,151,151	0
7	CA	F	2015	1/1	0.98	0.15	-2.11	58,58,58,58	0
7	CA	D	2014	1/1	0.93	0.10	-2.31	163,163,163,163	0
3	NAG	D	2012	14/15	0.80	0.25	-	145,163,197,200	0
3	NAG	D	2013	14/15	0.76	0.20	-	168,182,187,189	0
3	NAG	D	2009	14/15	0.73	0.41	-	188,192,195,195	0
3	NAG	B	2007	14/15	0.78	0.26	-	149,179,202,211	0
3	NAG	B	2008	14/15	0.48	0.33	-	196,198,198,199	0
3	NAG	F	2008	14/15	0.47	0.28	-	190,198,203,205	0
3	NAG	F	2009	14/15	0.73	0.35	-	185,189,191,192	0
3	NAG	F	2001	14/15	0.37	0.29	-	192,198,201,203	0
3	NAG	D	2001	14/15	0.69	0.19	-	187,191,196,197	0
3	NAG	D	2008	14/15	0.69	0.20	-	196,203,209,210	0
3	NAG	B	2012	14/15	0.71	0.21	-	146,163,198,200	0
3	NAG	B	2013	14/15	0.83	0.17	-	157,169,174,176	0
3	NAG	B	2009	14/15	0.78	0.45	-	185,193,199,199	0
3	NAG	F	2013	14/15	0.59	0.29	-	162,176,180,183	0
3	NAG	F	2012	14/15	0.69	0.22	-	139,157,192,194	0
3	NAG	D	2007	14/15	0.90	0.13	-	166,193,216,223	0
3	NAG	F	2007	14/15	0.72	0.31	-	151,172,188,195	0
3	NAG	B	2001	14/15	0.84	0.20	-	195,202,205,206	0

6.5 Other polymers ⓘ

There are no such residues in this entry.