wwPDB EM Validation Summary Report (i)

Dec 13, 2022 – 12:14 AM EST

PDB ID	:	1SUV
Title	:	Structure of Human Transferrin Receptor-Transferrin Complex
Authors	:	Cheng, Y.; Zak, O.; Aisen, P.; Harrison, S.C.; Walz, T.
Deposited on	:	2004-03-26
Resolution	:	7.50 Å(reported)
Based on initial models	:	1CX8, 1JNF, 1A8E

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

The following versions of software and data (see references (1)) were used in the production of this report:

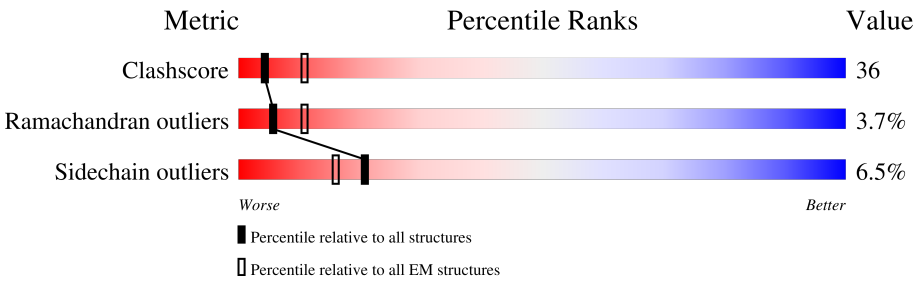
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 7.50 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	639	39% 50% 9% •
1	B	639	40% 49% 10% •
2	C	329	87% 12% •
2	D	329	86% 13% •
3	E	345	59% 34% 7%
3	F	345	60% 33% 7%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 20612 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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					AltConf	Trace
			Total	C	N	O	S		
1	A	639	5056	3244	846	952	14	0	0
1	B	639	5056	3244	846	952	14	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	172	GLU	GLN	SEE REMARK 999	UNP P02786
A	613	GLU	ARG	SEE REMARK 999	UNP P02786
B	172	GLU	GLN	SEE REMARK 999	UNP P02786
B	613	GLU	ARG	SEE REMARK 999	UNP P02786

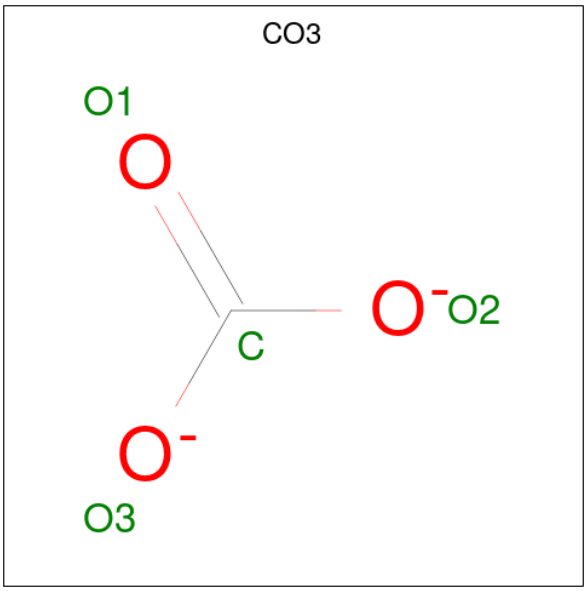
- Molecule 2 is a protein called Serotransferrin, N-lobe.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	329	2567	1621	440	485	21	5	0
2	D	329	2567	1621	440	485	21	5	0

- Molecule 3 is a protein called Serotransferrin, C-lobe.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	E	345	2669	1662	469	512	26	0	0
3	F	345	2669	1662	469	512	26	0	0

- Molecule 4 is CARBONATE ION (three-letter code: CO3) (formula: CO_3).



Mol	Chain	Residues	Atoms			AltConf
			Total	C	O	
4	C	1	8	2	6	1
4	D	1	8	2	6	1
4	E	1	4	1	3	0
4	F	1	4	1	3	0

- Molecule 5 is FE (III) ION (three-letter code: FE) (formula: Fe).

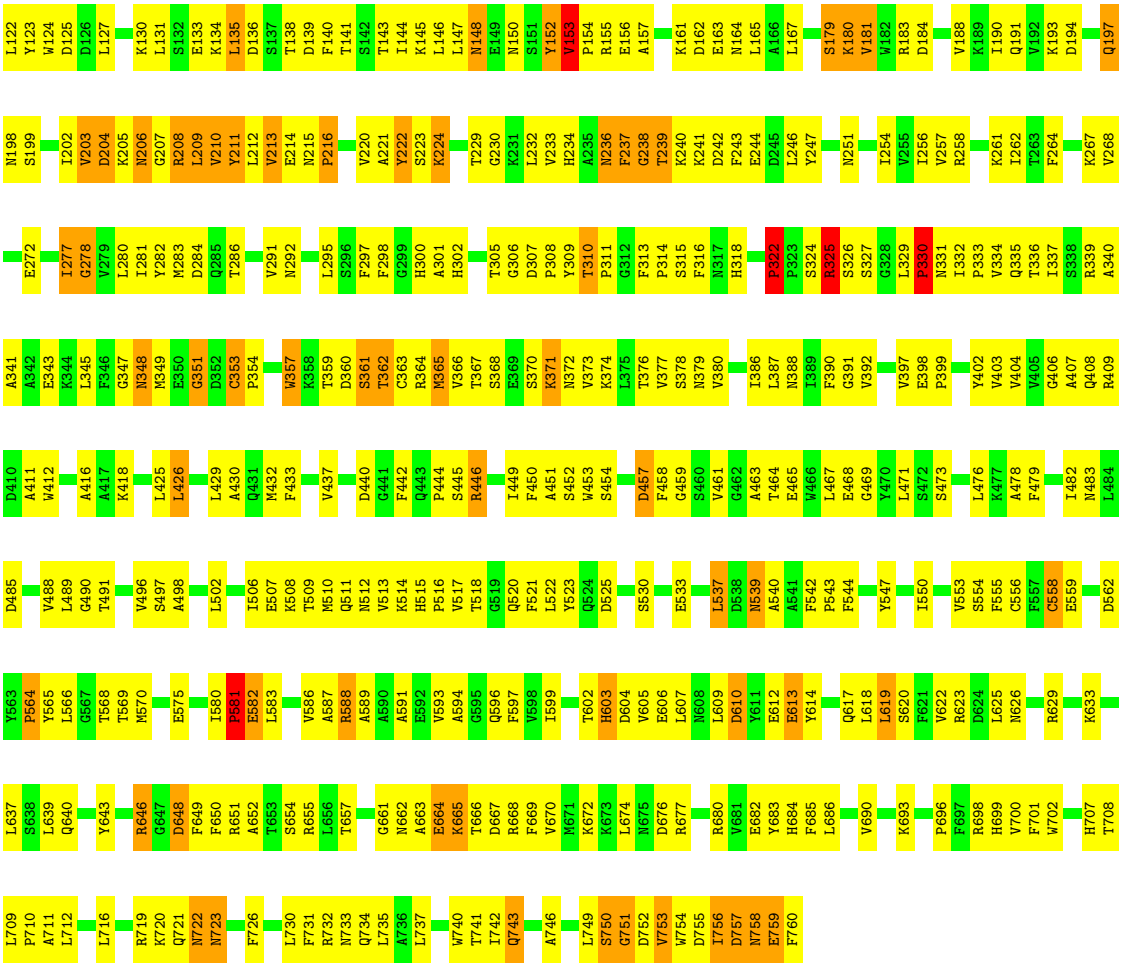
Mol	Chain	Residues	Atoms		AltConf
			Total	Fe	
5	C	1	1	1	0
5	D	1	1	1	0
5	E	1	1	1	0
5	F	1	1	1	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

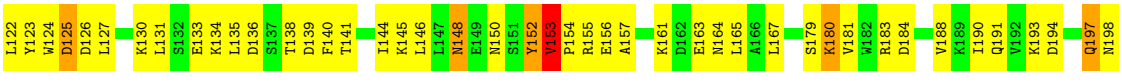
- Molecule 1: Transferrin receptor protein 1

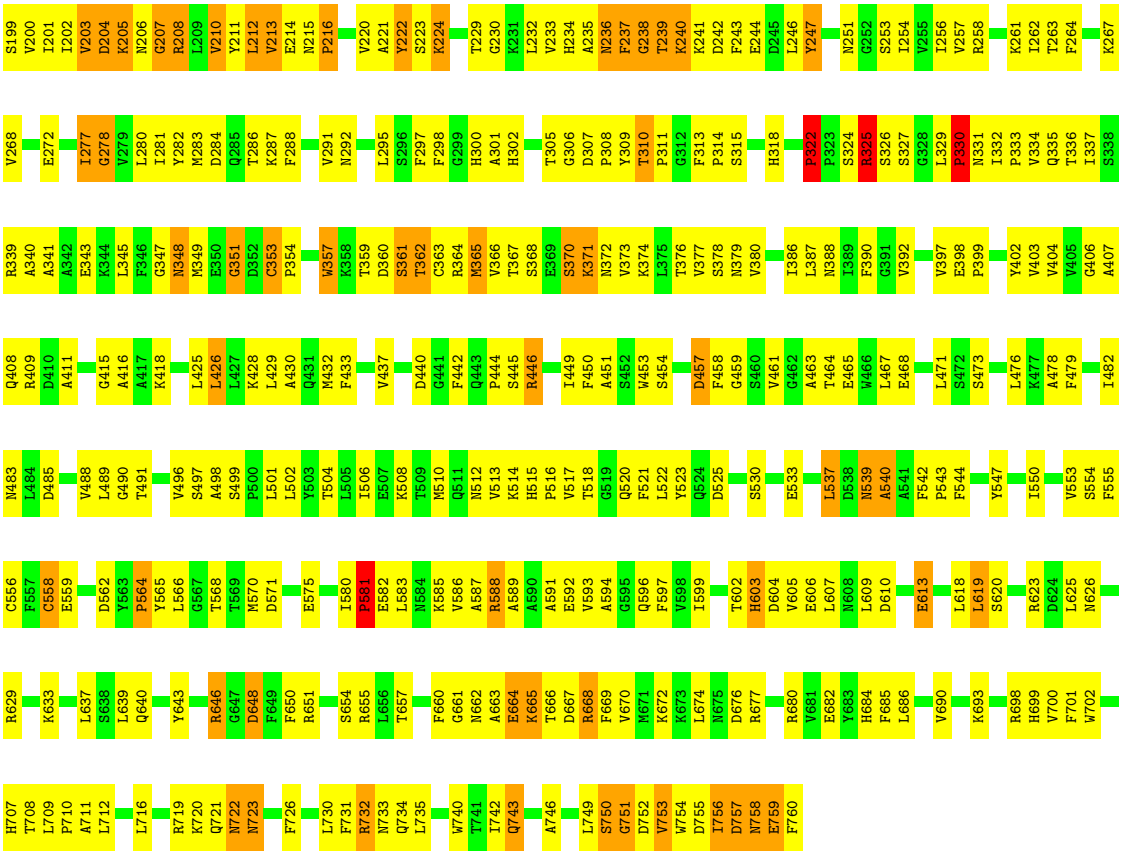
Chain A: 39% 50% 9%



- Molecule 1: Transferrin receptor protein 1

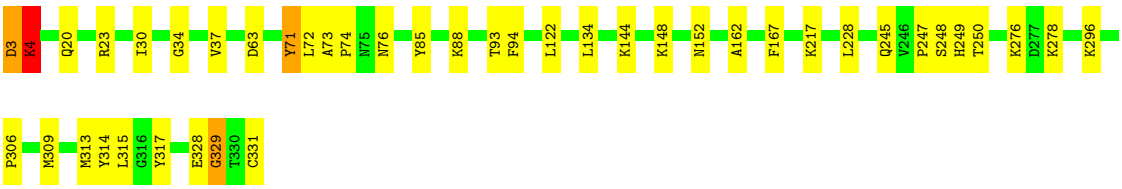
Chain B: 40% 49% 10%





- Molecule 2: Serotransferrin, N-lobe

Chain C:

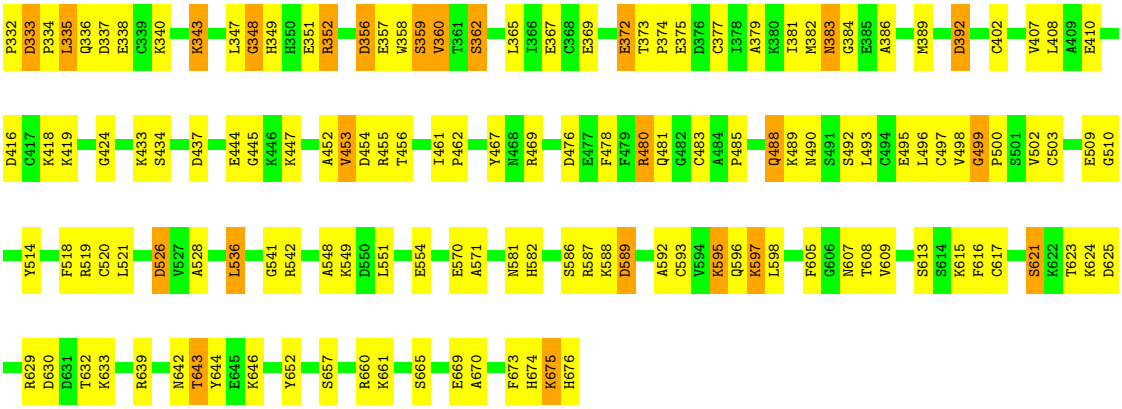


- Molecule 2: Serotransferrin, N-lobe

Chain D:

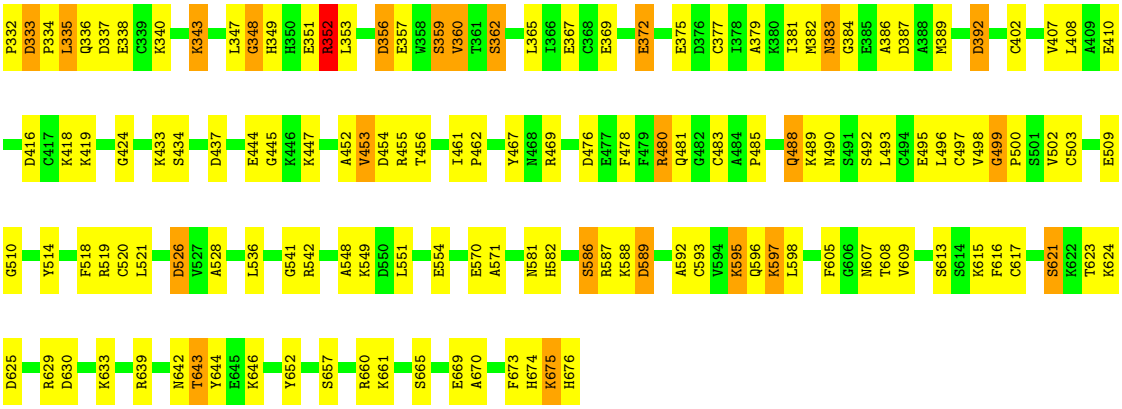
- Molecule 3: Serotransferrin, C-lobe

Chain E:



• Molecule 3: Serotransferrin, C-lobe

Chain F: 60% 33% 7%



GLOBAL-STATISTICS INFOmissingINFO

4 Model quality (i)

4.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: FE, 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	0.38	0/5177	0.61	1/7021 (0.0%)
1	B	0.38	0/5177	0.61	1/7021 (0.0%)
2	C	0.84	0/2648	0.88	5/3578 (0.1%)
2	D	0.84	0/2648	0.88	5/3578 (0.1%)
3	E	0.89	0/2723	1.03	9/3675 (0.2%)
3	F	0.89	0/2723	1.03	9/3675 (0.2%)
All	All	0.67	0/21096	0.81	30/28548 (0.1%)

There are no bond length outliers.

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	88	LYS	N-CA-CB	12.30	132.73	110.60
2	D	88	LYS	N-CA-CB	12.28	132.70	110.60
3	F	437	ASP	CB-CG-OD2	11.00	128.20	118.30
3	E	437	ASP	CB-CG-OD2	10.97	128.17	118.30
2	C	88	LYS	CA-CB-CG	-10.21	90.94	113.40

There are no chirality outliers.

There are no planarity outliers.

4.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5056	0	4980	582	0
1	B	5056	0	4982	575	0
2	C	2567	0	2482	94	0
2	D	2567	0	2482	95	0
3	E	2669	0	2574	197	0
3	F	2669	0	2575	205	0
4	C	8	0	0	0	0
4	D	8	0	0	0	0
4	E	4	0	0	0	0
4	F	4	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
All	All	20612	0	20075	1446	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 1446 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)
2:D:317:TYR:CE1	3:F:383:ASN:HB2	1.19	1.66
2:C:317:TYR:CE1	3:E:383:ASN:HB2	1.19	1.65
1:B:623:ARG:HD2	3:F:360:VAL:CG2	1.16	1.62
1:A:623:ARG:CD	3:E:360:VAL:HG23	1.14	1.59
1:A:623:ARG:HB3	3:E:360:VAL:CG2	1.34	1.56

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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 PDB entries followed by that with respect to all EM entries.

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	637/639 (100%)	499 (78%)	100 (16%)	38 (6%)	1	17
1	B	637/639 (100%)	497 (78%)	101 (16%)	39 (6%)	1	17
2	C	332/329 (101%)	315 (95%)	15 (4%)	2 (1%)	25	66
2	D	332/329 (101%)	315 (95%)	15 (4%)	2 (1%)	25	66
3	E	343/345 (99%)	309 (90%)	26 (8%)	8 (2%)	6	34
3	F	343/345 (99%)	309 (90%)	26 (8%)	8 (2%)	6	34
All	All	2624/2626 (100%)	2244 (86%)	283 (11%)	97 (4%)	6	24

5 of 97 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	206	ASN
1	A	210	VAL
1	A	251	ASN
1	A	330	PRO
1	A	361	SER

4.3.2 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 PDB entries followed by that with respect to all EM entries.

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	548/548 (100%)	505 (92%)	43 (8%)	12	36
1	B	548/548 (100%)	507 (92%)	41 (8%)	13	38
2	C	279/274 (102%)	275 (99%)	4 (1%)	67	80
2	D	279/274 (102%)	275 (99%)	4 (1%)	67	80
3	E	292/293 (100%)	266 (91%)	26 (9%)	9	30
3	F	292/293 (100%)	266 (91%)	26 (9%)	9	30
All	All	2238/2230 (100%)	2094 (94%)	144 (6%)	21	42

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	644	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	F	675	LYS
3	F	335	LEU
3	F	480	ARG
1	B	197	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	539	ASN
2	C	108	GLN
1	B	640	GLN
1	B	723	ASN
2	C	325	ASN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	CO3	F	701	5	2,3,3	0.35	0	2,3,3	1.59	1 (50%)
4	CO3	C	338[B]	5	2,3,3	0.70	0	2,3,3	0.55	0
4	CO3	C	338[A]	5	2,3,3	0.27	0	2,3,3	0.13	0
4	CO3	E	701	5	2,3,3	0.35	0	2,3,3	1.59	1 (50%)
4	CO3	D	338[B]	5	2,3,3	0.70	0	2,3,3	0.58	0
4	CO3	D	338[A]	5	2,3,3	0.26	0	2,3,3	0.10	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	701	CO3	O3-C-O1	2.12	125.04	119.55
4	F	701	CO3	O3-C-O1	2.12	125.04	119.55

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.