



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 31, 2017 – 09:39 AM EDT

PDB ID : 5BWE
Title : Benzylsuccinate synthase alpha-beta-gamma complex with bound toluene and fumarate
Authors : Funk, M.A.; Drennan, C.L.
Deposited on : unknown
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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
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) : rb-20029824

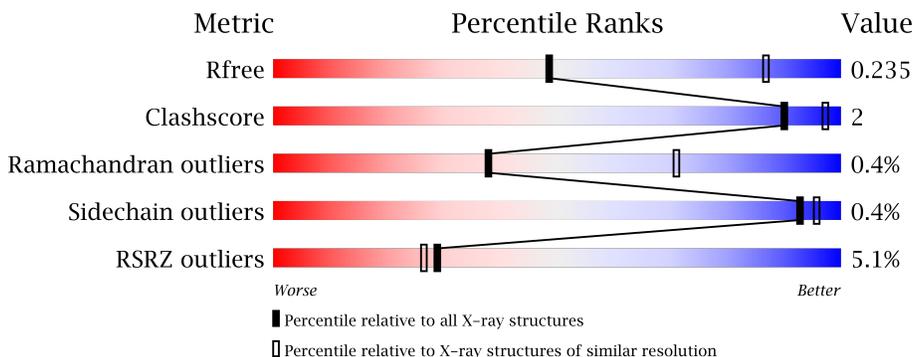
1 Overall quality at a glance

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}	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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	878	 4% 93% 5%
1	D	878	 4% 93% 5%
2	B	81	 2% 74% 11% 15%
2	F	81	 77% 9% 15%
3	C	60	 27% 55% 7% 38%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	60	

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
4	FUM	A	901	-	-	X	X
4	FUM	D	901	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 15396 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 benzylsuccinate synthase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	857	6809	4299	1179	1293	38	0	0	0
1	D	857	6809	4299	1179	1293	38	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	789	ILE	MET	engineered mutation	UNP O68395
A	866	SER	-	expression tag	UNP O68395
A	867	GLY	-	expression tag	UNP O68395
A	868	THR	-	expression tag	UNP O68395
A	869	GLY	-	expression tag	UNP O68395
A	870	SER	-	expression tag	UNP O68395
A	871	GLY	-	expression tag	UNP O68395
A	872	SER	-	expression tag	UNP O68395
A	873	SER	-	expression tag	UNP O68395
A	874	HIS	-	expression tag	UNP O68395
A	875	HIS	-	expression tag	UNP O68395
A	876	HIS	-	expression tag	UNP O68395
A	877	HIS	-	expression tag	UNP O68395
A	878	HIS	-	expression tag	UNP O68395
A	879	HIS	-	expression tag	UNP O68395
D	789	ILE	MET	engineered mutation	UNP O68395
D	866	SER	-	expression tag	UNP O68395
D	867	GLY	-	expression tag	UNP O68395
D	868	THR	-	expression tag	UNP O68395
D	869	GLY	-	expression tag	UNP O68395
D	870	SER	-	expression tag	UNP O68395
D	871	GLY	-	expression tag	UNP O68395
D	872	SER	-	expression tag	UNP O68395
D	873	SER	-	expression tag	UNP O68395
D	874	HIS	-	expression tag	UNP O68395

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	875	HIS	-	expression tag	UNP O68395
D	876	HIS	-	expression tag	UNP O68395
D	877	HIS	-	expression tag	UNP O68395
D	878	HIS	-	expression tag	UNP O68395
D	879	HIS	-	expression tag	UNP O68395

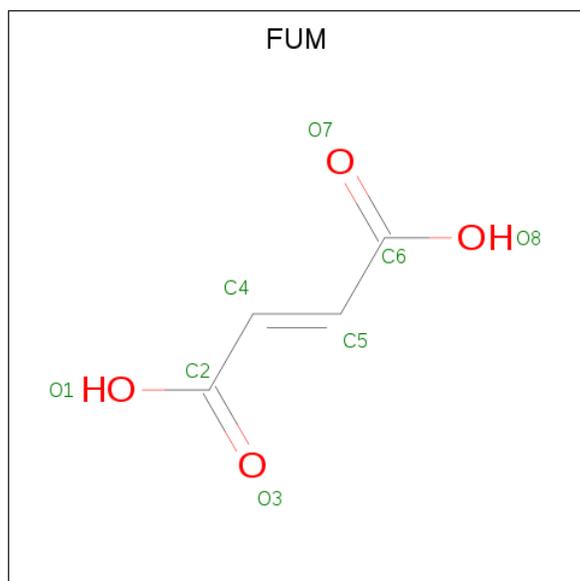
- Molecule 2 is a protein called benzylsuccinate synthase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	69	Total	C	N	O	S	0	0	0
			560	336	116	103	5			
2	F	69	Total	C	N	O	S	0	0	0
			560	336	116	103	5			

- Molecule 3 is a protein called benzylsuccinate synthase gamma chain.

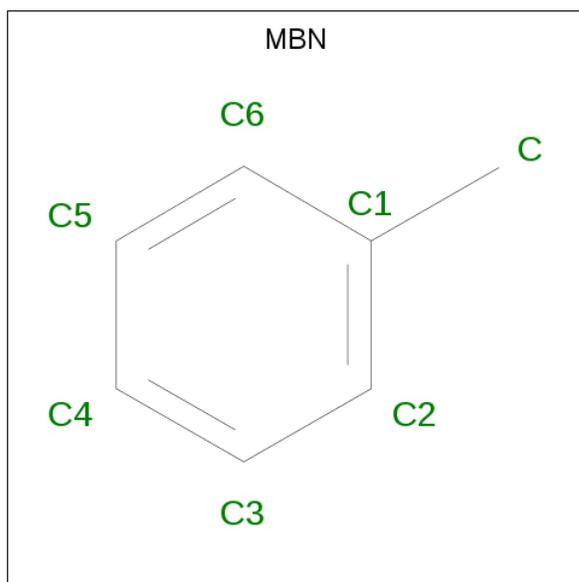
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	37	Total	C	N	O	S	0	0	0
			306	198	48	59	1			
3	G	37	Total	C	N	O	S	0	0	0
			306	198	48	59	1			

- Molecule 4 is FUMARIC ACID (three-letter code: FUM) (formula: C₄H₄O₄).



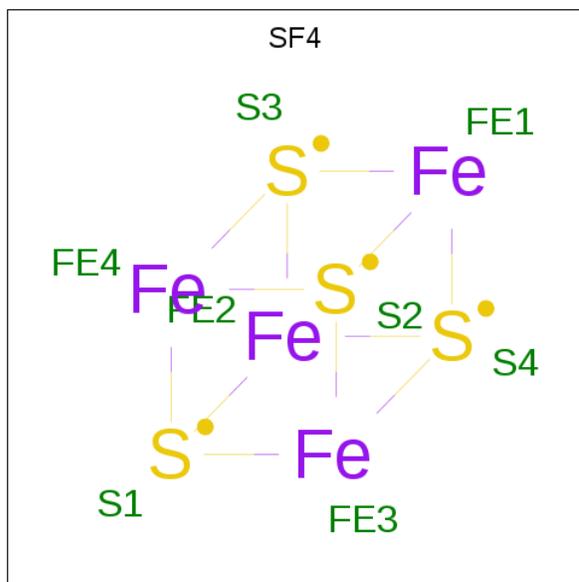
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 8 4 4	0	0
4	D	1	Total C O 8 4 4	0	0

- Molecule 5 is TOLUENE (three-letter code: MBN) (formula: C₇H₈).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C 7 7	0	0
5	D	1	Total C 7 7	0	0

- Molecule 6 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).

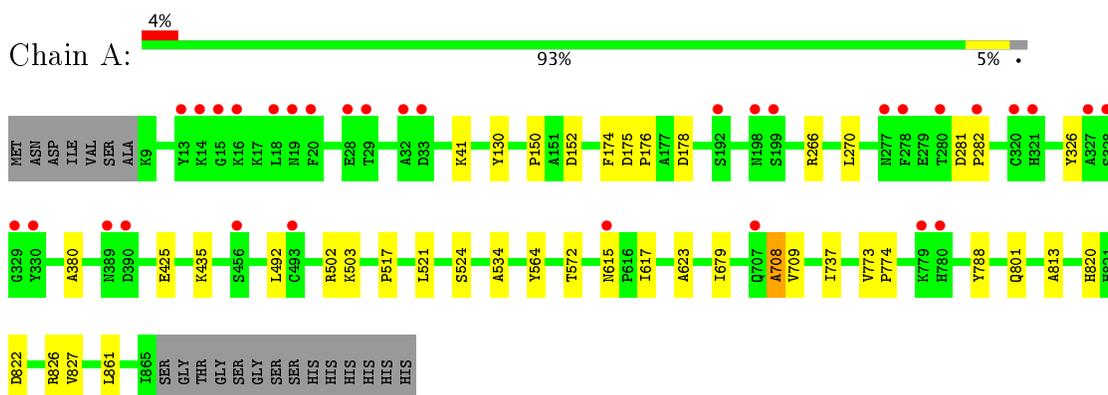


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	Fe	S	0	0
			8	4	4		
6	F	1	Total	Fe	S	0	0
			8	4	4		

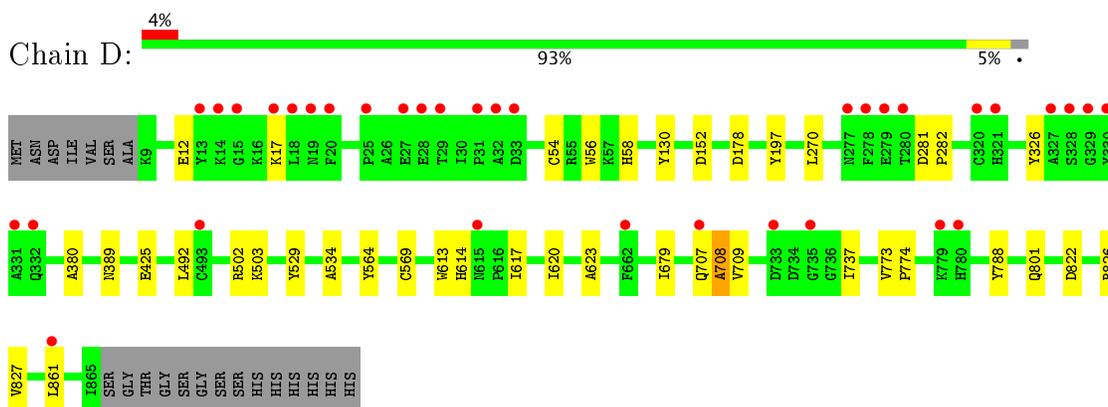
3 Residue-property plots [i](#)

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.

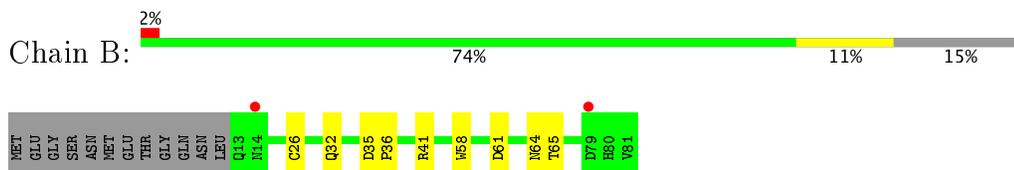
- Molecule 1: benzy succinate synthase alpha chain



- Molecule 1: benzy succinate synthase alpha chain



- Molecule 2: benzy succinate synthase beta chain

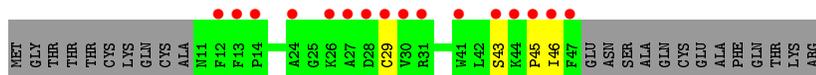


- Molecule 2: benzy succinate synthase beta chain





- Molecule 3: benzy succinate synthase gamma chain



- Molecule 3: benzy succinate synthase gamma chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	141.53Å 115.36Å 121.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.27 – 3.30 38.27 – 3.30	Depositor EDS
% Data completeness (in resolution range)	98.6 (38.27-3.30) 90.4 (38.27-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 3.32Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.212 , 0.237 0.215 , 0.235	Depositor DCC
R_{free} test set	1644 reflections (5.95%)	DCC
Wilson B-factor (Å ²)	89.7	Xtrriage
Anisotropy	0.641	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 49.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15396	wwPDB-VP
Average B, all atoms (Å ²)	113.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.71 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.0966e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SF4, FUM, MBN

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	0.23	0/6965	0.44	0/9409
1	D	0.23	0/6965	0.45	0/9409
2	B	0.23	0/576	0.48	0/779
2	F	0.23	0/576	0.47	0/779
3	C	0.24	0/315	0.44	0/423
3	G	0.24	0/315	0.44	0/423
All	All	0.23	0/15712	0.45	0/21222

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	A	6809	0	6668	29	1
1	D	6809	0	6668	24	1
2	B	560	0	518	6	0
2	F	560	0	518	5	0
3	C	306	0	287	3	0
3	G	306	0	287	3	0
4	A	8	0	2	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	8	0	2	3	0
5	A	7	0	8	0	0
5	D	7	0	8	2	0
6	B	8	0	0	0	0
6	F	8	0	0	0	0
All	All	15396	0	14966	63	1

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 2.

The worst 5 of 63 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:492:LEU:HB2	4:A:901:FUM:H5	1.75	0.68
1:A:492:LEU:HB2	4:A:901:FUM:C5	2.27	0.64
1:A:615:ASN:HD21	4:A:901:FUM:H4	1.62	0.64
1:A:178:ASP:OD1	1:A:564:TYR:OH	2.17	0.59
1:D:623:ALA:HB1	1:D:679:ILE:HG12	1.88	0.55

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:41:LYS:NZ	1:D:529:TYR:OH[4_445]	2.08	0.12

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	855/878 (97%)	832 (97%)	19 (2%)	4 (0%)	32 66
1	D	855/878 (97%)	832 (97%)	19 (2%)	4 (0%)	32 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	67/81 (83%)	66 (98%)	1 (2%)	0	100	100
2	F	67/81 (83%)	66 (98%)	1 (2%)	0	100	100
3	C	35/60 (58%)	34 (97%)	1 (3%)	0	100	100
3	G	35/60 (58%)	34 (97%)	1 (3%)	0	100	100
All	All	1914/2038 (94%)	1864 (97%)	42 (2%)	8 (0%)	38	71

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	709	VAL
1	D	709	VAL
1	A	708	ALA
1	A	737	ILE
1	A	827	VAL

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	724/741 (98%)	722 (100%)	2 (0%)	94	97
1	D	724/741 (98%)	722 (100%)	2 (0%)	94	97
2	B	61/71 (86%)	60 (98%)	1 (2%)	68	84
2	F	61/71 (86%)	60 (98%)	1 (2%)	68	84
3	C	32/51 (63%)	32 (100%)	0	100	100
3	G	32/51 (63%)	32 (100%)	0	100	100
All	All	1634/1726 (95%)	1628 (100%)	6 (0%)	93	96

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	32	GLN
2	F	32	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	425	GLU
1	A	822	ASP
1	D	822	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	820	HIS
1	D	104	ASN
1	D	646	GLN
1	A	801	GLN
1	D	786	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](#)

6 ligands 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	FUM	A	901	-	1,7,7	0.28	0	0,8,8	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	MBN	A	902	-	7,7,7	0.34	0	8,8,8	0.36	0
6	SF4	B	101	2	0,12,12	0.00	-	0,24,24	0.00	-
4	FUM	D	901	-	1,7,7	0.32	0	0,8,8	0.00	-
5	MBN	D	902	-	7,7,7	0.34	0	8,8,8	0.39	0
6	SF4	F	101	2	0,12,12	0.00	-	0,24,24	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	FUM	A	901	-	-	0/0/5/5	0/0/0/0
5	MBN	A	902	-	-	0/0/0/0	0/1/1/1
6	SF4	B	101	2	-	0/0/48/48	0/6/5/5
4	FUM	D	901	-	-	0/0/5/5	0/0/0/0
5	MBN	D	902	-	-	0/0/0/0	0/1/1/1
6	SF4	F	101	2	-	0/0/48/48	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	FUM	6	0
4	D	901	FUM	3	0
5	D	902	MBN	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	857/878 (97%)	0.18	32 (3%) 42 38	80, 111, 139, 152	0
1	D	857/878 (97%)	0.18	35 (4%) 38 35	73, 109, 141, 174	0
2	B	69/81 (85%)	0.37	2 (2%) 52 50	111, 126, 147, 156	0
2	F	69/81 (85%)	0.54	0 100 100	106, 129, 146, 157	0
3	C	37/60 (61%)	1.70	16 (43%) 0 0	121, 136, 145, 155	0
3	G	37/60 (61%)	1.30	14 (37%) 0 1	122, 134, 142, 146	0
All	All	1926/2038 (94%)	0.25	99 (5%) 29 26	73, 112, 141, 174	0

The worst 5 of 99 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	47	PHE	6.2
1	D	29	THR	6.1
1	D	28	GLU	6.0
1	A	15	GLY	5.1
1	A	19	ASN	5.0

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
4	FUM	A	901	8/8	0.85	0.64	4.40	88,98,100,104	0
4	FUM	D	901	8/8	0.90	0.53	2.54	82,92,94,98	0
5	MBN	A	902	7/7	0.94	0.35	0.08	98,98,98,98	0
5	MBN	D	902	7/7	0.94	0.29	-0.60	92,92,92,92	0
6	SF4	B	101	8/8	0.98	0.14	-1.41	129,131,132,132	0
6	SF4	F	101	8/8	0.98	0.12	-1.84	127,129,130,130	0

6.5 Other polymers [i](#)

There are no such residues in this entry.