



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 11, 2023 – 04:35 AM EDT

PDB ID : 4KYO
Title : Alanine-glyoxylate aminotransferase variant K390A in complex with the TPR domain of human Pex5p
Authors : Fodor, K.; Lou, Y.; Wilmanns, M.
Deposited on : 2013-05-29
Resolution : 2.20 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

<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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

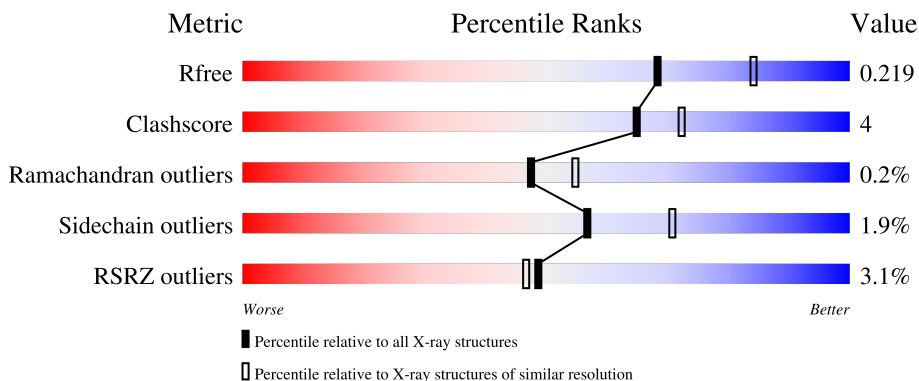
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 of the lower 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 $\leq 5\%$. 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	394	91% (green), 7% (yellow), 2% (grey), 0% (red)
1	C	394	92% (green), 7% (yellow), 1% (grey), 0% (red)
2	B	328	6% (red), 80% (green), 7% (yellow), 7% (grey)
2	D	328	5% (red), 76% (green), 10% (yellow), 9% (grey)

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	BTB	C	401	-	-	X	-
5	BME	A	408	-	-	X	-
5	BME	C	402	-	-	X	-

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 11250 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 Serine-pyruvate aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	387	Total	C	N	O	S	0	2	0
			2991	1912	521	542	16			
1	C	388	Total	C	N	O	S	0	0	0
			2988	1911	521	540	16			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P21549
A	0	ALA	-	expression tag	UNP P21549
A	390	ALA	LYS	engineered mutation	UNP P21549
C	-1	GLY	-	expression tag	UNP P21549
C	0	ALA	-	expression tag	UNP P21549
C	390	ALA	LYS	engineered mutation	UNP P21549

- Molecule 2 is a protein called Peroxisomal targeting signal 1 receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	286	Total	C	N	O	S	0	1	0
			2243	1408	391	433	11			
2	D	285	Total	C	N	O	S	0	2	0
			2233	1403	388	431	11			

There are 6 discrepancies between the modelled and reference sequences:

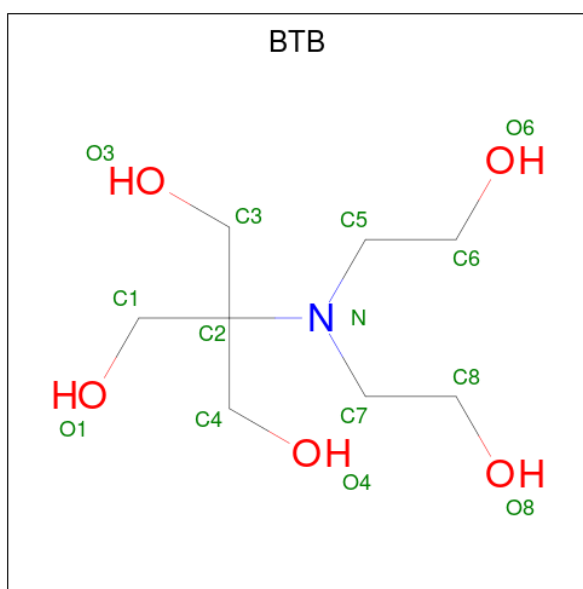
Chain	Residue	Modelled	Actual	Comment	Reference
B	312	GLY	-	expression tag	UNP P50542
B	313	ALA	-	expression tag	UNP P50542
B	314	MET	-	expression tag	UNP P50542
D	312	GLY	-	expression tag	UNP P50542
D	313	ALA	-	expression tag	UNP P50542
D	314	MET	-	expression tag	UNP P50542

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



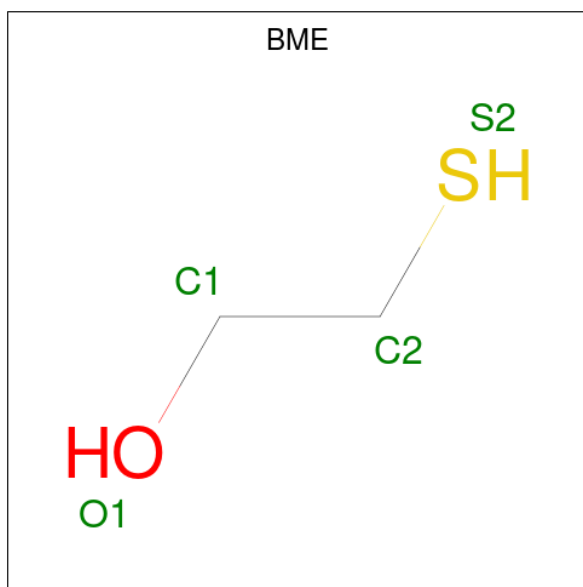
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPAN E-1,3-DIOL (three-letter code: BTB) (formula: C₈H₁₉NO₅).



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

- Molecule 5 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C₂H₆OS).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	O	S	0	0
			4	2	1	1		
5	D	1	Total	C	O	S	0	0
			4	2	1	1		

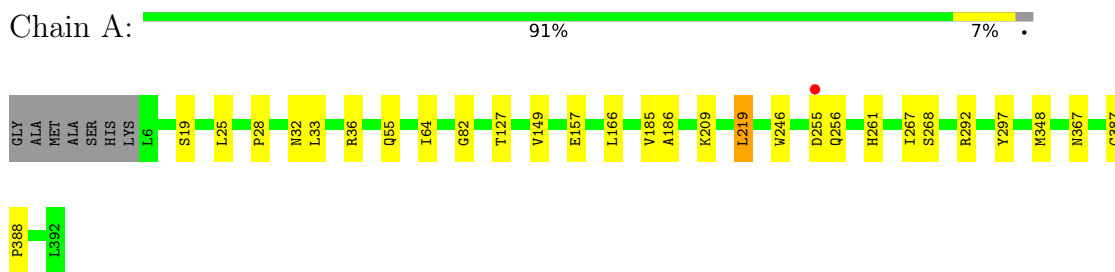
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	271	Total	O	0	0
			271	271		
6	C	267	Total	O	0	0
			267	267		
6	B	64	Total	O	0	0
			64	64		
6	D	107	Total	O	0	0
			107	107		

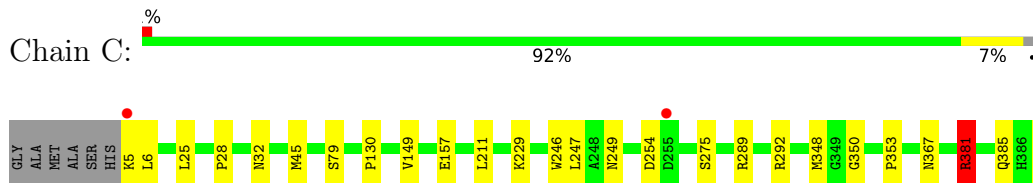
3 Residue-property plots [i](#)

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 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.

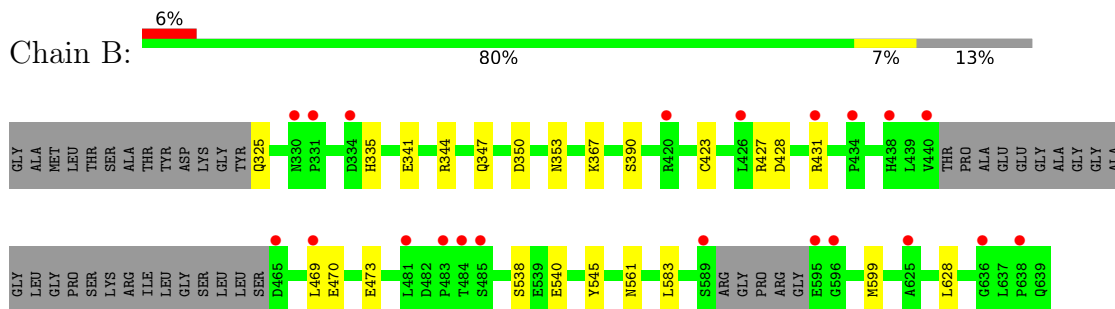
- Molecule 1: Serine-pyruvate aminotransferase



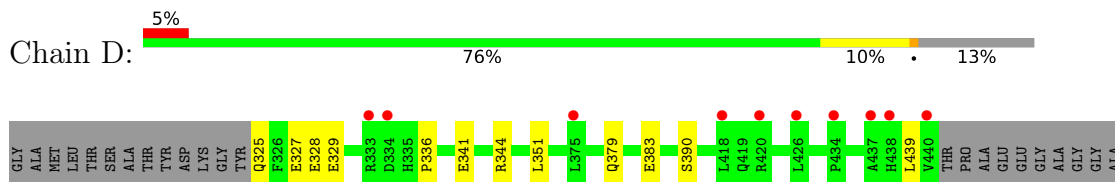
- Molecule 1: Serine-pyruvate aminotransferase

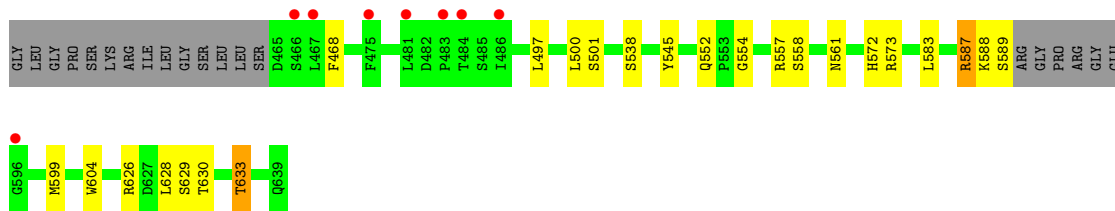


- Molecule 2: Peroxisomal targeting signal 1 receptor



- Molecule 2: Peroxisomal targeting signal 1 receptor





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	57.56Å 74.83Å 91.75Å 87.23° 83.41° 89.83°	Depositor
Resolution (Å)	19.74 – 2.20 19.73 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.0 (19.74-2.20) 97.1 (19.73-2.20)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 2.19Å)	Xtrriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.175 , 0.218 0.178 , 0.219	Depositor DCC
R_{free} test set	3747 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	24.6	Xtrriage
Anisotropy	0.039	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 48.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.016 for -h,k,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11250	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.13% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BME, BTB, SO4

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	0.44	0/3066	0.64	0/4161
1	C	0.42	0/3057	0.65	2/4147 (0.0%)
2	B	0.37	0/2286	0.56	0/3099
2	D	0.40	0/2278	0.56	0/3089
All	All	0.41	0/10687	0.61	2/14496 (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
2	D	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	C	381	ARG	NE-CZ-NH2	-5.43	117.59	120.30
1	C	381	ARG	NE-CZ-NH1	5.08	122.84	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	329	GLU	Peptide

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	2991	0	3033	26	0
1	C	2988	0	3039	30	0
2	B	2243	0	2187	10	0
2	D	2233	0	2172	20	0
3	A	10	0	0	1	0
4	A	14	0	19	0	0
4	C	14	0	19	7	0
5	A	20	0	30	11	0
5	B	4	0	6	0	0
5	C	20	0	30	8	0
5	D	4	0	6	0	0
6	A	271	0	0	3	0
6	B	64	0	0	0	0
6	C	267	0	0	1	0
6	D	107	0	0	1	0
All	All	11250	0	10541	86	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 4.

The worst 5 of 86 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:79:SER:H	5:C:402:BME:H21	1.19	1.03
1:C:247:LEU:HD22	5:C:402:BME:H11	1.42	0.97
1:A:267:ILE:HB	5:A:408:BME:H12	1.58	0.85
1:C:391:LYS:HB2	2:D:557:ARG:HD2	1.60	0.82
1:C:381:ARG:NH2	1:C:385:GLN:OE1	2.14	0.80

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	387/394 (98%)	378 (98%)	9 (2%)	0	100	100
1	C	386/394 (98%)	376 (97%)	9 (2%)	1 (0%)	41	46
2	B	281/328 (86%)	273 (97%)	8 (3%)	0	100	100
2	D	281/328 (86%)	270 (96%)	10 (4%)	1 (0%)	34	37
All	All	1335/1444 (92%)	1297 (97%)	36 (3%)	2 (0%)	47	60

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	336	PRO
1	C	387	CYS

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	321/325 (99%)	315 (98%)	6 (2%)	57	71
1	C	321/325 (99%)	319 (99%)	2 (1%)	86	93
2	B	235/264 (89%)	229 (97%)	6 (3%)	46	58
2	D	233/264 (88%)	225 (97%)	8 (3%)	37	47
All	All	1110/1178 (94%)	1088 (98%)	22 (2%)	57	69

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	327	GLU
2	D	439	LEU
2	D	351	LEU
2	D	538	SER
1	C	157	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	364	GLN
2	D	536	ASN
2	D	616	GLN
2	D	552	GLN
1	A	303	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

16 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
5	BME	C	406	-	3,3,3	0.34	0	1,2,2	0.05	0
5	BME	A	407	-	3,3,3	0.38	0	1,2,2	0.36	0
3	SO4	A	401	-	4,4,4	0.20	0	6,6,6	0.30	0
5	BME	A	405	-	3,3,3	0.40	0	1,2,2	0.40	0
4	BTB	A	403	-	13,13,13	0.75	0	7,16,16	0.80	0
3	SO4	A	402	-	4,4,4	0.23	0	6,6,6	0.30	0
5	BME	C	403	-	3,3,3	0.31	0	1,2,2	0.02	0
5	BME	C	402	-	3,3,3	1.09	0	1,2,2	3.08	1 (100%)
5	BME	D	701	-	3,3,3	0.42	0	1,2,2	0.06	0
4	BTB	C	401	-	13,13,13	1.01	1 (7%)	7,16,16	1.38	1 (14%)
5	BME	C	405	-	3,3,3	0.38	0	1,2,2	0.25	0
5	BME	A	408	-	3,3,3	0.32	0	1,2,2	1.00	0
5	BME	A	406	-	3,3,3	0.35	0	1,2,2	1.03	0
5	BME	C	404	-	3,3,3	0.36	0	1,2,2	0.28	0
5	BME	A	404	-	3,3,3	0.69	0	1,2,2	1.45	0
5	BME	B	701	-	3,3,3	0.52	0	1,2,2	0.29	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
5	BME	C	406	-	-	1/1/1/1	-
5	BME	A	407	-	-	1/1/1/1	-
5	BME	A	405	-	-	0/1/1/1	-
4	BTB	A	403	-	-	7/21/21/21	-
5	BME	D	701	-	-	0/1/1/1	-
5	BME	C	403	-	-	0/1/1/1	-
5	BME	C	402	-	-	1/1/1/1	-
4	BTB	C	401	-	-	5/21/21/21	-
5	BME	C	405	-	-	1/1/1/1	-
5	BME	A	408	-	-	1/1/1/1	-
5	BME	A	406	-	-	0/1/1/1	-
5	BME	C	404	-	-	0/1/1/1	-
5	BME	A	404	-	-	1/1/1/1	-
5	BME	B	701	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	401	BTB	C5-N	2.79	1.52	1.48

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	402	BME	O1-C1-C2	3.08	122.97	110.83
4	C	401	BTB	C8-C7-N	-2.39	102.24	111.59

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	403	BTB	C1-C2-N-C5
4	A	403	BTB	C3-C2-N-C5
4	A	403	BTB	C4-C2-N-C5
4	A	403	BTB	C6-C5-N-C7
4	A	403	BTB	N-C5-C6-O6

There are no ring outliers.

8 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	406	BME	2	0
5	A	407	BME	1	0
3	A	402	SO4	1	0
5	C	402	BME	6	0
4	C	401	BTB	7	0
5	A	408	BME	5	0
5	A	406	BME	3	0
5	A	404	BME	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	387/394 (98%)	-0.58	1 (0%) 94 93	14, 19, 29, 38	0
1	C	388/394 (98%)	-0.55	2 (0%) 91 90	13, 19, 29, 43	0
2	B	286/328 (87%)	0.25	21 (7%) 15 14	21, 40, 65, 77	0
2	D	285/328 (86%)	0.23	18 (6%) 20 19	17, 38, 69, 81	0
All	All	1346/1444 (93%)	-0.23	42 (3%) 49 47	13, 23, 59, 81	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	440	VAL	6.1
2	D	596	GLY	5.1
2	D	437	ALA	4.4
2	B	596	GLY	4.3
2	D	440	VAL	4.1

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 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, 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	B-factors(\AA^2)	Q<0.9
5	BME	A	405	4/4	0.61	0.27	57,57,61,64	0
5	BME	A	404	4/4	0.64	0.32	37,40,41,48	0
5	BME	C	402	4/4	0.72	0.28	26,30,31,40	0
5	BME	B	701	4/4	0.77	0.22	52,52,53,55	0
5	BME	C	403	4/4	0.78	0.17	46,50,52,56	0
5	BME	A	406	4/4	0.81	0.29	49,49,51,54	0
4	BTB	C	401	14/14	0.81	0.19	32,37,39,40	0
5	BME	A	407	4/4	0.82	0.27	60,61,62,63	0
4	BTB	A	403	14/14	0.83	0.24	39,45,48,51	0
5	BME	D	701	4/4	0.88	0.13	57,58,59,61	0
5	BME	C	405	4/4	0.90	0.15	56,57,58,58	0
5	BME	C	406	4/4	0.90	0.20	51,51,51,53	0
5	BME	A	408	4/4	0.93	0.23	40,42,45,50	0
5	BME	C	404	4/4	0.93	0.14	61,61,61,62	0
3	SO4	A	401	5/5	0.99	0.09	19,20,21,22	0
3	SO4	A	402	5/5	0.99	0.07	23,23,24,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.