Full wwPDB X-ray Structure Validation Report (i)

May 23, 2024 – 10:35 AM EDT

PDB ID : 3ZNG
Title : Ankyrin repeat and SOCS-box protein 9 (ASB9) in complex with ElonginB and ElonginC
Authors : Thomas, J.; Van Molle, I.; Ciulli, A.
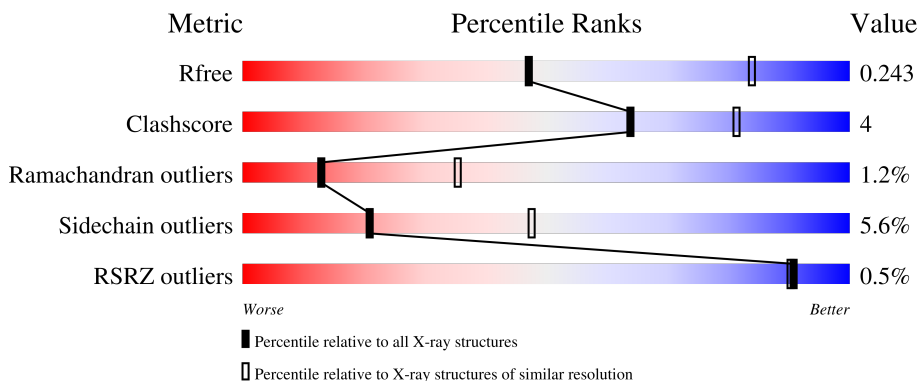Deposited on : 2013-02-14
Resolution : 2.85 Å(reported)

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 2.85 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	Percentile Ranks	Value
Rfree		0.243
Clashscore		4
Ramachandran outliers		1.2%
Sidechain outliers		5.6%
RSRZ outliers		0.5%

Worse — *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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 <=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	268	90% 5% · ·
1	D	268	84% 11% · ·
2	B	97	77% 9% 13%
2	E	97	67% 16% · · 13%
3	C	118	46% 14% 41%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	118	70% · 14% · 11%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 6525 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 ANKYRIN REPEAT AND SOCS BOX PROTEIN 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	258	Total	C	N	O	S	0	1	0
			1945	1226	353	353	13			
1	D	259	Total	C	N	O	S	0	0	0
			1951	1228	356	354	13			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	MET	-	expression tag	UNP Q96DX5
A	295	ALA	-	expression tag	UNP Q96DX5
A	296	GLU	-	expression tag	UNP Q96DX5
A	297	ASN	-	expression tag	UNP Q96DX5
A	298	LEU	-	expression tag	UNP Q96DX5
A	299	TYR	-	expression tag	UNP Q96DX5
A	300	PHE	-	expression tag	UNP Q96DX5
A	301	GLN	-	expression tag	UNP Q96DX5
D	34	MET	-	expression tag	UNP Q96DX5
D	295	ALA	-	expression tag	UNP Q96DX5
D	296	GLU	-	expression tag	UNP Q96DX5
D	297	ASN	-	expression tag	UNP Q96DX5
D	298	LEU	-	expression tag	UNP Q96DX5
D	299	TYR	-	expression tag	UNP Q96DX5
D	300	PHE	-	expression tag	UNP Q96DX5
D	301	GLN	-	expression tag	UNP Q96DX5

- Molecule 2 is a protein called TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 1.

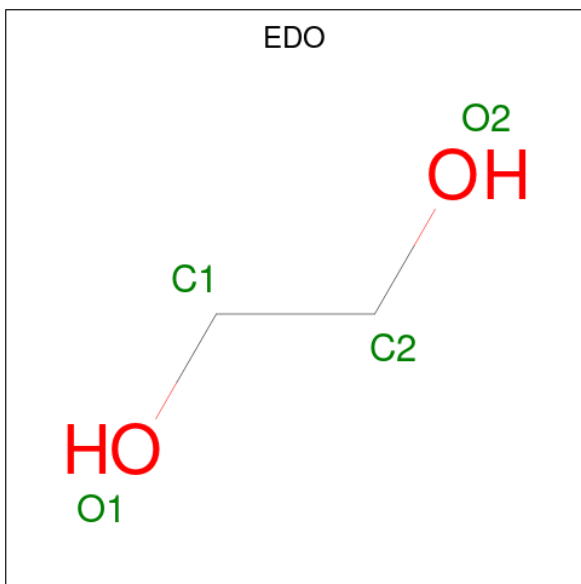| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	84	Total	C	N	O	S	0	0	0
			641	416	102	118	5			
2	E	84	Total	C	N	O	S	0	1	0
			658	431	101	121	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	16	MET	-	expression tag	UNP Q15369
E	16	MET	-	expression tag	UNP Q15369

- Molecule 3 is a protein called TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	70	Total 487	C 309	N 81	O 95	S 2	0	0	0
3	F	105	Total 804	C 512	N 134	O 154	S 4	0	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	23	Total 23	O 23	0	0
5	B	1	Total 1	O 1	0	0

Continued on next page...

Continued from previous page...

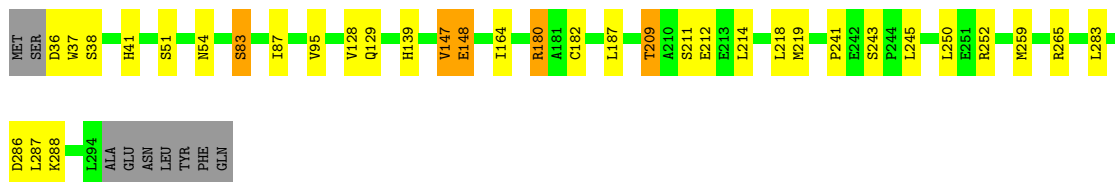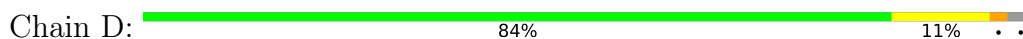| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	5	Total 5	O 5	0	0
5	D	4	Total 4	O 4	0	0
5	E	2	Total 2	O 2	0	0

3 Residue-property plots ⓘ

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: ANKYRIN REPEAT AND SOCS BOX PROTEIN 9
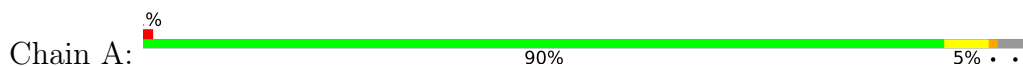
Chain A:

- Molecule 1: ANKYRIN REPEAT AND SOCS BOX PROTEIN 9

Chain D:

- Molecule 2: TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 1

Chain B:

- Molecule 2: TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 1

Chain E:

- Molecule 3: TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 2

Chain C:

ALA LEU CYS ILE E91 L99 P100 P105 GLN ASP SER GLY SER SER ALA ASN GLU GLN ALA VAL GLN
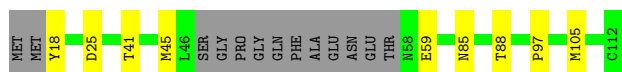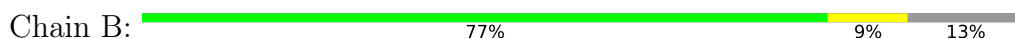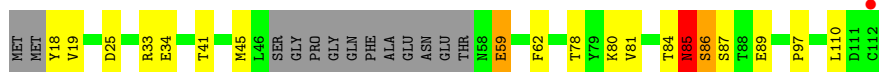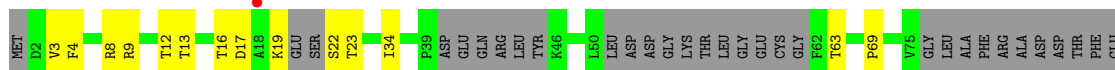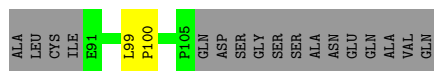
• Molecule 3: TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 2

Chain F:

70%   14%   •   11%

MET D2 K11 T12 T23 V24 F25 E26 I30 P38 P39 D53 T56 G61 F62 T63 S64 Q65 T66 P69 F79 T84 F85 E86 S95 V102 Q106 ASP SER GLY SER SER ALA ASN GLU GLN ALA VAL GLN
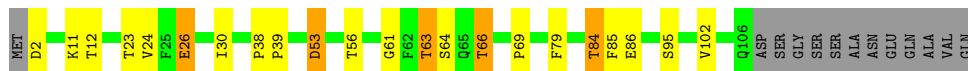
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	114.11Å 103.13Å 109.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.64 – 2.85 50.64 – 2.85	Depositor EDS
% Data completeness (in resolution range)	99.1 (50.64-2.85) 93.8 (50.64-2.85)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 2.86Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.195 , 0.242 0.199 , 0.243	Depositor DCC
R_{free} test set	1553 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	68.7	Xtriage
Anisotropy	0.262	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 57.2	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.024 for l,-k,h	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6525	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: EDO

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	0.48	0/1993	0.62	0/2713
1	D	0.45	0/1995	0.58	0/2715
2	B	0.44	0/655	0.58	0/889
2	E	0.42	0/673	0.58	0/911
3	C	0.37	0/496	0.60	0/678
3	F	0.40	0/821	0.57	0/1115
All	All	0.44	0/6633	0.59	0/9021

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 ⓘ

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	1945	0	1902	10	0
1	D	1951	0	1922	17	0
2	B	641	0	601	4	0
2	E	658	0	632	13	0
3	C	487	0	424	4	0
3	F	804	0	784	10	0
4	A	4	0	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	23	0	0	0	0
5	B	1	0	0	1	0
5	C	5	0	0	0	0
5	D	4	0	0	0	0
5	E	2	0	0	0	0
All	All	6525	0	6271	53	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.

All (53) 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:E:18[A]:TYR:HB2	2:E:33:ARG:HB2	1.65	0.78
2:E:81:VAL:O	2:E:84:THR:HG22	1.94	0.68
3:C:8:ARG:HG2	3:C:13:THR:HG23	1.76	0.66
1:D:209:THR:HG22	1:D:211:SER:H	1.59	0.66
2:E:85:ASN:ND2	2:E:85:ASN:O	2.30	0.65
1:A:209:THR:HG22	1:A:211:SER:H	1.62	0.63
1:D:83:SER:O	1:D:87:ILE:HG13	2.04	0.57
3:F:24:VAL:HB	3:F:53:ASP:HA	1.86	0.56
3:F:23:THR:HA	3:F:56:THR:HA	1.88	0.56
2:E:86:SER:HB3	2:E:87:SER:HB3	1.88	0.55
1:D:148:GLU:H	1:D:148:GLU:CD	2.11	0.55
2:E:19:VAL:HG12	2:E:33:ARG:HA	1.92	0.52
3:F:63:THR:HG22	3:F:64:SER:H	1.74	0.51
2:E:85:ASN:O	2:E:85:ASN:CG	2.49	0.51
3:F:61:GLY:O	3:F:66:THR:HG21	2.10	0.51
1:D:265:ARG:NH2	2:E:97:PRO:HG3	2.26	0.49
1:D:286:ASP:OD1	1:D:287:LEU:N	2.45	0.49
1:A:166:HIS:CE1	1:A:167:LEU:HG	2.49	0.48
1:D:212:GLU:HG2	1:D:245:LEU:HD13	1.95	0.48
1:A:212:GLU:OE2	1:D:51:SER:OG	2.29	0.47
3:C:16:THR:OG1	3:C:17:ASP:N	2.48	0.47
3:F:79:PHE:O	3:F:86:GLU:HG2	2.15	0.47
1:D:180:ARG:HA	1:D:214:LEU:HD21	1.97	0.46
2:E:41:THR:O	2:E:45:MET:HG3	2.16	0.46
3:F:26:GLU:O	3:F:30:ILE:HG12	2.16	0.46
1:A:265:ARG:HD2	1:A:265:ARG:HA	1.78	0.45
2:E:25:ASP:O	3:F:11:LYS:NZ	2.45	0.45
1:A:180:ARG:HA	1:A:214:LEU:HD21	2.00	0.44

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:209:THR:HG22	1:D:211:SER:N	2.30	0.44
3:C:4:PHE:CE1	3:C:69:PRO:HG3	2.52	0.44
2:B:25:ASP:OD1	2:B:25:ASP:N	2.51	0.44
1:D:187:LEU:HD11	1:D:218:LEU:HD23	2.00	0.44
1:A:236:VAL:HA	1:A:249:PHE:HD2	1.83	0.44
1:D:38:SER:HB3	1:D:41:HIS:ND1	2.32	0.44
1:D:265:ARG:HA	1:D:265:ARG:HD2	1.81	0.44
2:E:84:THR:O	2:E:85:ASN:HB3	2.18	0.44
2:E:78:THR:HG22	3:F:69:PRO:HB3	2.00	0.43
2:E:80:LYS:O	2:E:84:THR:HB	2.19	0.43
1:A:236:VAL:O	1:A:239:VAL:HG13	2.19	0.43
1:A:265:ARG:NH2	2:B:97:PRO:HG3	2.33	0.43
1:D:219:MET:HE1	1:D:252:ARG:HG3	2.01	0.43
1:D:36:ASP:HB3	1:D:37:TRP:H	1.62	0.42
1:D:147:VAL:H	1:D:147:VAL:HG12	1.62	0.42
1:A:245:LEU:HD23	1:A:245:LEU:HA	1.86	0.42
3:C:99:LEU:HD12	3:C:100:PRO:HD2	2.01	0.42
1:D:283:LEU:O	1:D:288:LYS:HE3	2.20	0.42
2:B:41:THR:O	2:B:45:MET:HG3	2.20	0.41
1:D:139:HIS:CE1	1:D:164:ILE:HG13	2.56	0.41
2:E:62:PHE:CE1	2:E:110:LEU:HD21	2.55	0.41
1:A:184:LYS:O	1:A:188:GLU:HG3	2.20	0.41
2:B:18:TYR:N	5:B:2001:HOH:O	2.52	0.41
3:F:84:THR:OG1	3:F:85:PHE:N	2.53	0.40
3:F:38:PRO:HA	3:F:39:PRO:HD3	1.99	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	257/268 (96%)	252 (98%)	4 (2%)	1 (0%)	34	62

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	257/268 (96%)	250 (97%)	6 (2%)	1 (0%)	34	62
2	B	80/97 (82%)	74 (92%)	4 (5%)	2 (2%)	5	18
2	E	80/97 (82%)	73 (91%)	4 (5%)	3 (4%)	3	10
3	C	62/118 (52%)	51 (82%)	8 (13%)	3 (5%)	2	6
3	F	103/118 (87%)	96 (93%)	7 (7%)	0	100	100
All	All	839/966 (87%)	796 (95%)	33 (4%)	10 (1%)	13	35

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	85	ASN
3	C	22	SER
1	D	241	PRO
1	A	166	HIS
2	B	59	GLU
3	C	19	LYS
3	C	34	ILE
2	E	89	GLU
2	E	59	GLU
2	E	85	ASN

5.3.2 Protein sidechains ⓘ

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	204/224 (91%)	199 (98%)	5 (2%)	47	76
1	D	207/224 (92%)	194 (94%)	13 (6%)	18	42
2	B	65/86 (76%)	63 (97%)	2 (3%)	40	71
2	E	69/86 (80%)	65 (94%)	4 (6%)	20	46
3	C	46/103 (45%)	41 (89%)	5 (11%)	6	17
3	F	86/103 (84%)	77 (90%)	9 (10%)	7	18
All	All	677/826 (82%)	639 (94%)	38 (6%)	21	47

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	ARG
1	A	83	SER
1	A	95	VAL
1	A	209	THR
1	A	239	VAL
2	B	88	THR
2	B	105	MET
3	C	3	VAL
3	C	9	ARG
3	C	12	THR
3	C	23	THR
3	C	63	THR
1	D	54	ASN
1	D	83	SER
1	D	95	VAL
1	D	128	VAL
1	D	129	GLN
1	D	147	VAL
1	D	148	GLU
1	D	180	ARG
1	D	182	CYS
1	D	209	THR
1	D	243	SER
1	D	250	LEU
1	D	259	MET
2	E	34	GLU
2	E	59	GLU
2	E	85	ASN
2	E	86	SER
3	F	2	ASP
3	F	12	THR
3	F	26	GLU
3	F	53	ASP
3	F	63	THR
3	F	66	THR
3	F	84	THR
3	F	95	SER
3	F	102	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	293	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 ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

1 ligand is 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	EDO	A	1295	-	3,3,3	0.72	0	2,2,2	0.24	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
4	EDO	A	1295	-	-	0/1/1/1	-

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.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	258/268 (96%)	-0.05	2 (0%)	86	85	35, 52, 100, 139	0
1	D	259/268 (96%)	-0.12	0	100	100	42, 63, 96, 122	0
2	B	84/97 (86%)	-0.16	0	100	100	41, 77, 120, 161	0
2	E	84/97 (86%)	0.17	1 (1%)	79	78	66, 90, 124, 164	0
3	C	70/118 (59%)	-0.10	1 (1%)	75	74	59, 105, 133, 158	0
3	F	105/118 (88%)	0.03	0	100	100	56, 83, 114, 142	0
All	All	860/966 (89%)	-0.06	4 (0%)	91	90	35, 69, 120, 164	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	134	LEU	2.6
2	E	112	CYS	2.4
3	C	18	ALA	2.4
1	A	241	PRO	2.3

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

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. 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(Å^2)	Q<0.9
4	EDO	A	1295	4/4	0.67	0.13	86,88,93,94	0

6.5 Other polymers ⓘ

There are no such residues in this entry.