



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 10, 2017 – 10:54 AM EDT

PDB ID : 2P15
Title : Crystal structure of the ER alpha ligand binding domain with the agonist ortho-trifluoromethylphenylvinyl estradiol
Authors : Bruning, J.B.; Nettles, K.W.; Greene, G.L.; Kim, Y.
Deposited on : unknown
Resolution : 1.94 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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-20030345
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-20030345

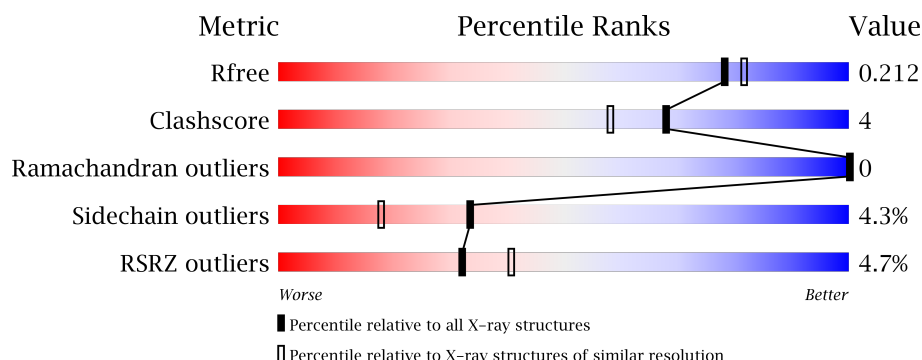
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.94 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	3233 (1.96-1.92)
Clashscore	112137	3430 (1.96-1.92)
Ramachandran outliers	110173	3395 (1.96-1.92)
Sidechain outliers	110143	3395 (1.96-1.92)
RSRZ outliers	101464	3250 (1.96-1.92)

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	258	<div> <div>3%</div> <div>81% 11% • 6%</div> </div>
1	B	258	<div> <div>4%</div> <div>87% 5% • 6%</div> </div>
2	C	13	<div> <div>23%</div> <div>46% 38% 15%</div> </div>
2	D	13	<div> <div>15%</div> <div>85% 15%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4617 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 Estrogen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	242	Total	C	N	O	S	3	6	0
			1938	1243	329	345	21			
1	B	242	Total	C	N	O	S	2	12	0
			1970	1265	329	355	21			

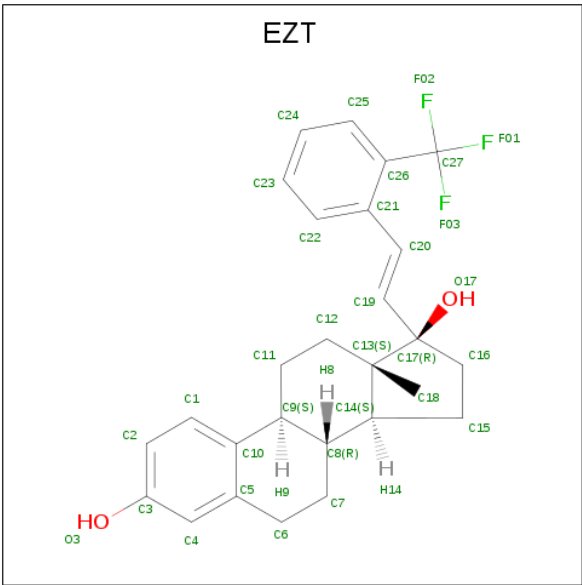
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	SER	-	CLONING ARTIFACT	UNP P03372
A	537	SER	TYR	ENGINEERED	UNP P03372
B	297	SER	-	CLONING ARTIFACT	UNP P03372
B	537	SER	TYR	ENGINEERED	UNP P03372

- Molecule 2 is a protein called GRIP peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	11	Total	C	N	O	0	0	0
			95	60	20	15			
2	D	11	Total	C	N	O	0	0	0
			89	57	18	14			

- Molecule 3 is (17BETA)-17-[(E)-2-[2-(TRIFLUOROMETHYL)PHENYL]VINYL]ESTRA-1(10),2,4-TRIENE-3,17-DIOL (three-letter code: EZT) (formula: C₂₇H₂₉F₃O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	F	O	0	0
			32	27	3	2		
3	B	1	Total	C	F	O	0	0
			32	27	3	2		

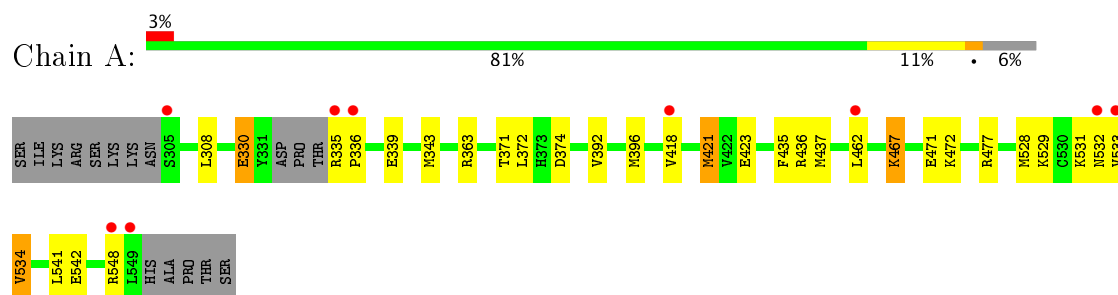
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	219	Total	O	0	0
			219	219		
4	B	237	Total	O	0	0
			237	237		
4	C	2	Total	O	0	0
			2	2		
4	D	3	Total	O	0	0
			3	3		

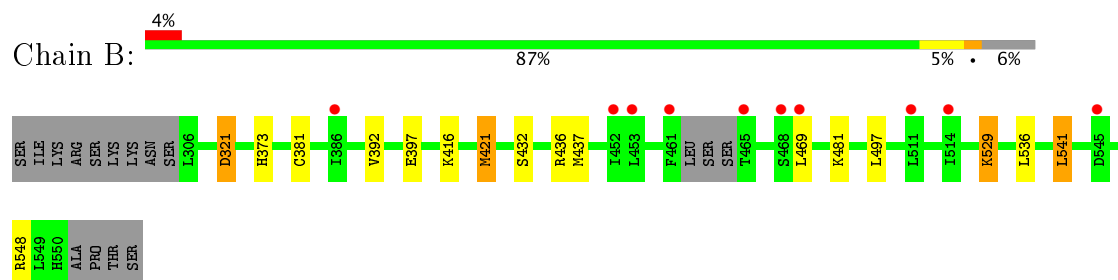
3 Residue-property plots [i](#)

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.

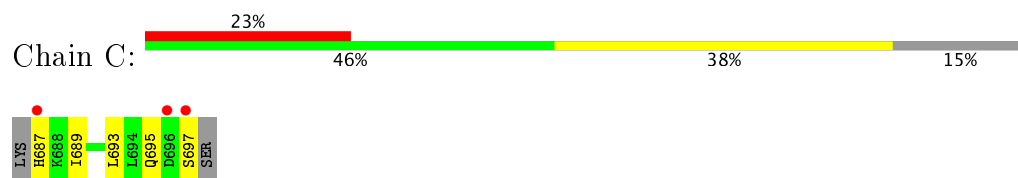
• Molecule 1: Estrogen receptor



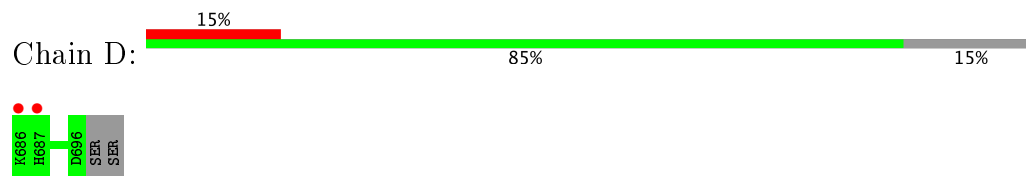
• Molecule 1: Estrogen receptor



• Molecule 2: GRIP peptide



• Molecule 2: GRIP peptide



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.05Å 84.22Å 58.69Å 90.00° 109.59° 90.00°	Depositor
Resolution (Å)	27.71 – 1.94 27.71 – 1.94	Depositor EDS
% Data completeness (in resolution range)	97.0 (27.71-1.94) 97.0 (27.71-1.94)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.18 (at 1.93Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.158 , 0.216 0.157 , 0.212	Depositor DCC
R_{free} test set	1871 reflections (5.34%)	DCC
Wilson B-factor (Å ²)	22.1	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 46.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.025 for l,-k,h	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4617	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.92% 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: EZT

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	0.62	0/1990	0.66	1/2686 (0.0%)
1	B	1.22	2/2042 (0.1%)	0.96	4/2760 (0.1%)
2	C	1.62	2/96 (2.1%)	0.70	0/127
2	D	0.61	0/89	0.51	0/118
All	All	0.98	4/4217 (0.1%)	0.82	5/5691 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	321[A]	ASP	CG-OD2	32.86	2.00	1.25
1	B	321[B]	ASP	CG-OD2	32.86	2.00	1.25
2	C	697	SER	C-O	12.59	1.47	1.23
2	C	697	SER	CB-OG	5.83	1.49	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	321[A]	ASP	CB-CG-OD2	22.66	138.69	118.30
1	B	321[B]	ASP	CB-CG-OD2	22.66	138.69	118.30
1	B	321[A]	ASP	OD1-CG-OD2	-10.50	103.36	123.30
1	B	321[B]	ASP	OD1-CG-OD2	-10.50	103.36	123.30
1	A	436	ARG	NE-CZ-NH2	-5.32	117.64	120.30

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	1938	0	1991	23	0
1	B	1970	0	2020	13	0
2	C	95	0	100	3	0
2	D	89	0	92	0	0
3	A	32	0	28	1	0
3	B	32	0	28	1	0
4	A	219	0	0	6	1
4	B	237	0	0	6	1
4	C	2	0	0	0	0
4	D	3	0	0	0	0
All	All	4617	0	4259	37	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 4.

All (37) 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:330:GLU:OE1	4:A:779:HOH:O	1.83	0.96
1:A:423:GLU:HG2	1:B:548:ARG:HH21	1.45	0.81
4:A:685:HOH:O	1:B:437:MET:HE1	1.85	0.76
1:A:418:VAL:HG13	1:A:528:MET:CE	2.16	0.75
1:B:381[B]:CYS:SG	4:B:826:HOH:O	2.45	0.73
1:A:418:VAL:HG13	1:A:528:MET:HE1	1.74	0.69
1:A:372:LEU:HD21	2:C:695:GLN:HG2	1.76	0.68
1:B:536:LEU:HB2	1:B:541:LEU:HD13	1.78	0.65
1:B:529:LYS:HG3	1:B:536:LEU:HD22	1.83	0.60
1:A:542[B]:GLU:HG3	4:A:657:HOH:O	2.01	0.59
1:A:548:ARG:HG2	4:A:798:HOH:O	2.02	0.59
1:A:418:VAL:HG13	1:A:528:MET:HE3	1.86	0.58
1:B:373[B]:HIS:HD2	4:B:825:HOH:O	1.87	0.58
1:A:418:VAL:CG1	1:A:528:MET:HE3	2.35	0.57
1:A:343:MET:HB3	1:A:534:VAL:HG21	1.87	0.57
1:A:372:LEU:HD21	2:C:695:GLN:CG	2.35	0.56
1:A:418:VAL:CG1	1:A:528:MET:CE	2.84	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:472:LYS:HB3	1:A:472:LYS:NZ	2.24	0.52
1:A:330:GLU:OE2	1:A:330:GLU:N	2.43	0.52
1:A:335:ARG:CB	1:A:336:PRO:HD3	2.42	0.50
4:A:685:HOH:O	1:B:437:MET:CE	2.52	0.50
1:B:436:ARG:HD3	4:B:758:HOH:O	2.11	0.49
1:B:437:MET:HE2	4:B:733:HOH:O	2.13	0.48
1:A:308:LEU:HD11	1:A:477:ARG:HE	1.79	0.47
2:C:689:ILE:O	2:C:693:LEU:HD13	2.14	0.47
1:B:421:MET:HG3	3:B:601:EZT:C24	2.45	0.47
1:A:531:LYS:HB3	1:A:533:VAL:HG23	1.97	0.47
1:A:374:ASP:OD2	1:A:471:GLU:HG2	2.14	0.46
1:A:371:THR:HG21	1:A:467:LYS:HE2	1.98	0.45
1:A:392:VAL:HG12	1:A:435:PHE:CD2	2.52	0.45
1:A:421:MET:C	1:A:421:MET:SD	2.95	0.45
1:A:363[A]:ARG:NH2	4:A:681:HOH:O	2.51	0.43
1:B:421:MET:HE3	4:B:677:HOH:O	2.19	0.43
1:B:392:VAL:HG13	1:B:432:SER:CA	2.48	0.43
1:A:339:GLU:HG2	1:A:533:VAL:HG11	2.01	0.43
1:B:469:LEU:HD23	4:B:720:HOH:O	2.19	0.42
3:A:600:EZT:H20	3:A:600:EZT:H161	1.83	0.40

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)
4:A:819:HOH:O	4:B:838:HOH:O[1_554]	0.75	1.45

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	244/258 (95%)	243 (100%)	1 (0%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	250/258 (97%)	248 (99%)	2 (1%)	0	100	100
2	C	9/13 (69%)	9 (100%)	0	0	100	100
2	D	9/13 (69%)	8 (89%)	1 (11%)	0	100	100
All	All	512/542 (94%)	508 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	218/233 (94%)	207 (95%)	11 (5%)	28	13
1	B	225/233 (97%)	216 (96%)	9 (4%)	36	21
2	C	11/13 (85%)	10 (91%)	1 (9%)	11	2
2	D	9/13 (69%)	9 (100%)	0	100	100
All	All	463/492 (94%)	442 (96%)	21 (4%)	33	16

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	330	GLU
1	A	396	MET
1	A	421	MET
1	A	437[A]	MET
1	A	437[B]	MET
1	A	462	LEU
1	A	467	LYS
1	A	529	LYS
1	A	532	ASN
1	A	534	VAL
1	A	541	LEU
1	B	321[A]	ASP
1	B	321[B]	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	397	GLU
1	B	416	LYS
1	B	421	MET
1	B	481	LYS
1	B	497	LEU
1	B	529	LYS
1	B	541	LEU
2	C	687	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	474	HIS
1	B	398	HIS

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](#)

2 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
3	EZT	A	600	-	35,36,36	1.42	6 (17%)	53,57,57	1.13	4 (7%)
3	EZT	B	601	-	35,36,36	1.44	6 (17%)	53,57,57	1.60	10 (18%)

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	EZT	A	600	-	-	0/12/56/56	0/5/5/5
3	EZT	B	601	-	-	0/12/56/56	0/5/5/5

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	601	EZT	C21-C20	-3.60	1.41	1.47
3	A	600	EZT	C21-C20	-3.47	1.42	1.47
3	B	601	EZT	C8-C14	2.02	1.57	1.53
3	A	600	EZT	C10-C9	2.17	1.55	1.52
3	B	601	EZT	C10-C9	2.26	1.55	1.52
3	A	600	EZT	C27-C26	2.32	1.55	1.50
3	A	600	EZT	C6-C5	2.36	1.55	1.51
3	B	601	EZT	C27-C26	2.47	1.56	1.50
3	A	600	EZT	C11-C9	2.70	1.57	1.53
3	B	601	EZT	C6-C5	2.99	1.56	1.51
3	B	601	EZT	C20-C19	3.18	1.40	1.32
3	A	600	EZT	C20-C19	3.51	1.41	1.32

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	600	EZT	C26-C21-C20	-3.30	120.17	122.23
3	B	601	EZT	C26-C21-C20	-2.88	120.43	122.23
3	B	601	EZT	C15-C14-C13	-2.34	100.47	103.74
3	B	601	EZT	C23-C24-C25	-2.25	117.12	120.21
3	A	600	EZT	F02-C27-C26	-2.18	108.76	112.69
3	B	601	EZT	F02-C27-C26	-2.14	108.84	112.69
3	B	601	EZT	C6-C5-C4	-2.08	115.75	119.91
3	B	601	EZT	F03-C27-C26	-2.02	109.05	112.69
3	B	601	EZT	C6-C5-C10	2.20	124.26	121.14
3	A	600	EZT	C18-C13-C17	2.21	113.32	108.44
3	A	600	EZT	C27-C26-C21	2.99	122.29	120.79

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	601	EZT	C7-C8-C9	3.18	112.39	109.28
3	B	601	EZT	C6-C7-C8	3.30	115.29	110.64
3	B	601	EZT	C27-C26-C21	7.05	124.33	120.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	600	EZT	1	0
3	B	601	EZT	1	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 ⓘ

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(Å ²)	Q<0.9
1	A	242/258 (93%)	0.08	9 (3%)	42 51	10, 17, 30, 39	5 (2%)
1	B	242/258 (93%)	0.04	10 (4%)	38 46	9, 15, 28, 36	1 (0%)
2	C	11/13 (84%)	1.98	3 (27%)	1 0	31, 33, 43, 43	0
2	D	11/13 (84%)	1.90	2 (18%)	1 2	30, 31, 40, 43	0
All	All	506/542 (93%)	0.14	24 (4%)	32 41	9, 16, 33, 43	6 (1%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	686	LYS	11.1
1	A	549	LEU	10.1
2	C	687	HIS	8.2
2	C	697	SER	8.2
2	D	687	HIS	6.4
1	A	336	PRO	3.9
2	C	696	ASP	3.6
1	A	305	SER	3.4
1	B	452	ILE	3.1
1	A	335	ARG	3.0
1	B	468	SER	2.7
1	B	511	LEU	2.6
1	A	548	ARG	2.6
1	B	469	LEU	2.5
1	A	418	VAL	2.5
1	B	514	ILE	2.4
1	B	465	THR	2.4
1	A	533	VAL	2.3
1	A	462	LEU	2.2
1	B	453	LEU	2.2
1	B	461	PHE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	386	ILE	2.1
1	B	545[A]	ASP	2.1
1	A	532	ASN	2.0

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(\AA^2)	Q<0.9
3	EZT	B	601	32/32	0.96	0.09	-0.57	13,16,19,21	0
3	EZT	A	600	32/32	0.96	0.08	-0.72	15,18,23,24	0

6.5 Other polymers [i](#)

There are no such residues in this entry.