W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2020 – 01:57 PM BST

PDB ID	:	6TTK
Title	:	Crystal structure of the kelch domain of human KLHL12 in complex with DVL1 peptide
Authors	:	Chen, Z.; Williams, E.; Pike, A.C.W.; Strain-Damerell, C.; Wang, D.; Chalk, R.; Burgess-Brown, N.; Krojer, T.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Bullock, A.N.
Deposited on	:	2019-12-27
Resolution	:	2.38 Å(reported)

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

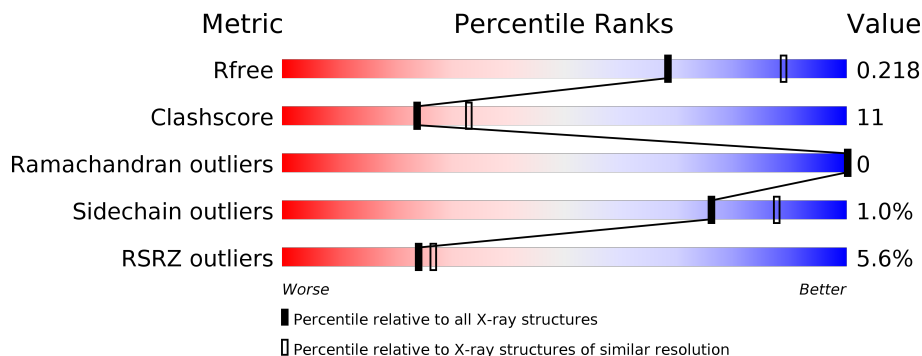|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13

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 2.38 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.
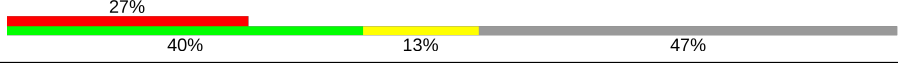
Metric	Percentile Ranks	Value
Rfree		0.218
Clashscore		11
Ramachandran outliers		0
Sidechain outliers		1.0%
RSRZ outliers		5.6%

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	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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 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	301	3% / 75% / 19% / • 5%
1	B	301	4% / 70% / 24% / • 5%
1	C	301	2% / 75% / 18% / • 6%
1	D	301	7% / 77% / 18% / • 5%
2	E	15	27% / 40% / 7% / 7% / 47%
2	F	15	13% / 27% / 20% / 53%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	15	27% 53% 47%
2	H	15	27% 40% 13% 47%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 9388 atoms, of which 6 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 Kelch-like protein 12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	285	Total 2143	C 1347	N 366	O 416	S 14	0	0	0
1	B	285	Total 2142	C 1347	N 360	O 420	S 15	0	1	0
1	C	284	Total 2143	C 1343	N 366	O 419	S 15	0	2	0
1	D	286	Total 2124	C 1330	N 362	O 417	S 15	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	267	SER	-	expression tag	UNP Q53G59
B	267	SER	-	expression tag	UNP Q53G59
C	267	SER	-	expression tag	UNP Q53G59
D	267	SER	-	expression tag	UNP Q53G59

- Molecule 2 is a protein called DVL1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	F	7	Total 37	C 23	N 7	O 7	0	0	0
2	G	8	Total 47	C 31	N 8	O 8	0	0	0
2	H	8	Total 45	C 29	N 8	O 8	0	0	0
2	E	8	Total 47	C 31	N 8	O 8	0	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total 2	Na 2	0	0
3	A	1	Total 1	Na 1	0	0
3	C	1	Total 1	Na 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).
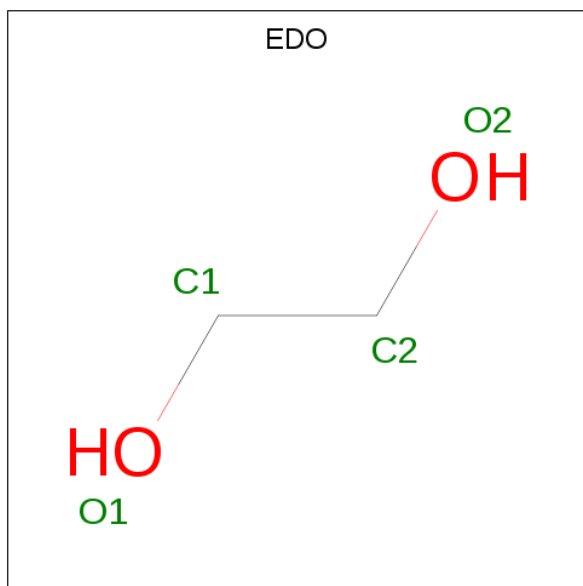
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 10	C 2	H 6	O 2	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total 1	Cl 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	158	Total 158	O 158	0	0
6	B	172	Total 172	O 172	0	0

Continued on next page...

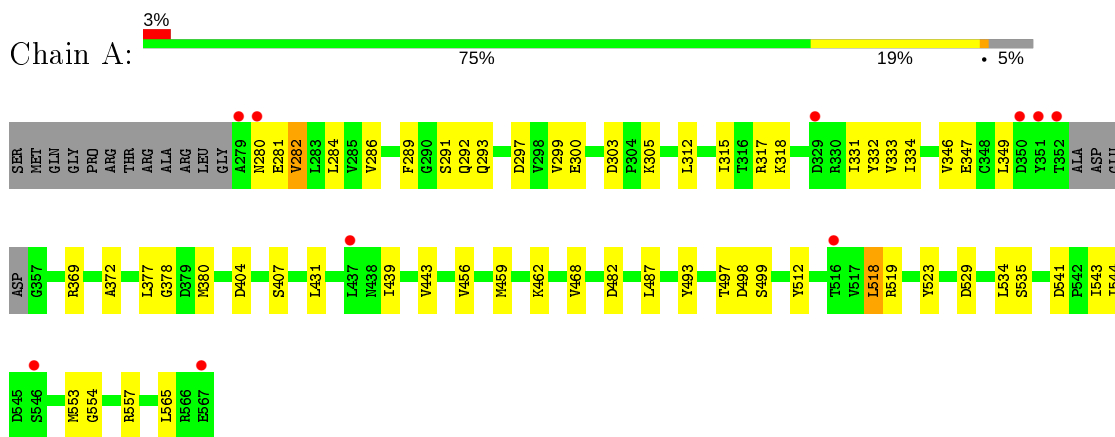*Continued from previous page...*

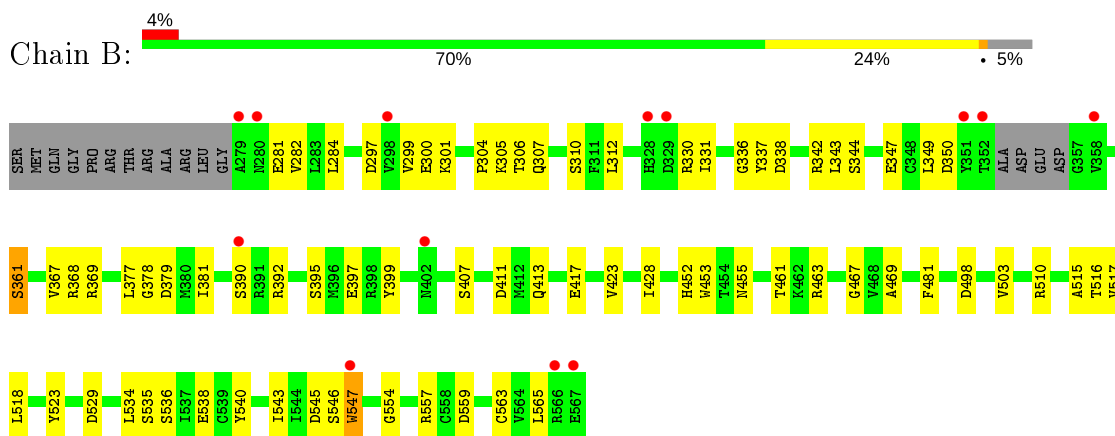| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	188	Total 188	O 188	0	0
6	D	122	Total 122	O 122	0	0
6	F	2	Total 2	O 2	0	0
6	H	1	Total 1	O 1	0	0
6	E	2	Total 2	O 2	0	0

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.
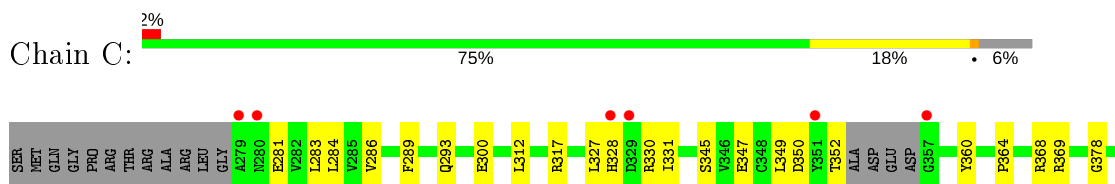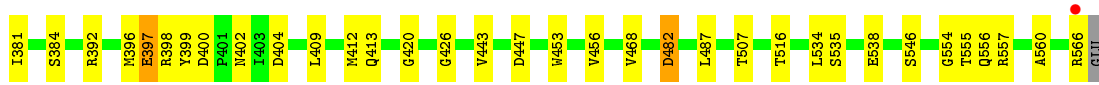
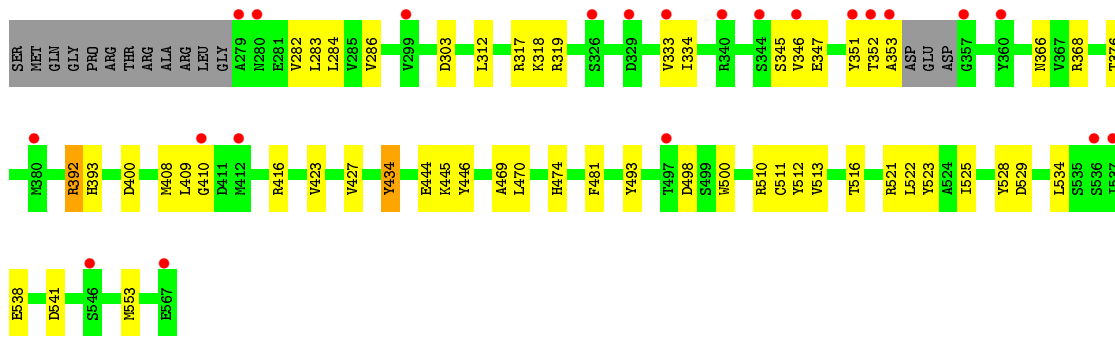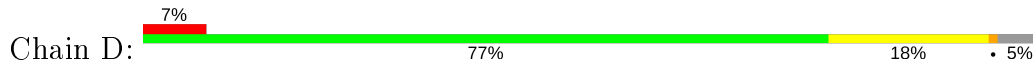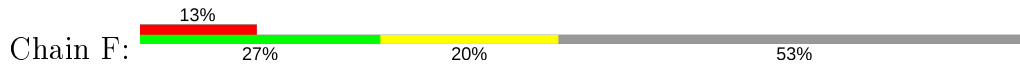- Molecule 1: Kelch-like protein 12
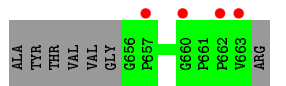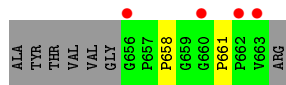

- Molecule 1: Kelch-like protein 12

- Molecule 1: Kelch-like protein 12

I381 S384 R392 M396 E397 R398 Y399 D400 P401 N402 I403 D404 L409 M412 Q413 G420 G426 V443 D447 W453 V456 V468 D482 L487 T507 T516 L534 S535 E538 S546 G554 T555 Q556 R557 A560 R566 GLU

• Molecule 1: Kelch-like protein 12

Chain D:

SER MET GLN GLY PRO ARG THR ARG ALA ARG LEU GLY A279 N280 E281 V282 L283 L284 V285 V286 V299 D303 L312 R317 K318 R319 S326 D329 V333 I334 R340 S344 S345 V346 E347 Y351 T352 A353 ASP GLU ASP G357 Y360 N366 V367 R368 T376

M380 R392 H393 D400 M408 L409 G410 D411 M412 R416 V423 V427 Y434 E444 K445 Y446 A469 L470 H474 F481 Y493 T497 D498 S499 W500 R510 C511 Y512 V513 T516 R521 L522 Y523 A524 I525 Y528 D529 L534 S535 S536 I537

E538 D541 S546 M553 E567

• Molecule 2: DVL1

Chain F:

ALA TYR THR VAL VAL G655 G656 P657 P658 G659 G660 P661 PRO VAL ARG

• Molecule 2: DVL1

Chain G:

ALA TYR THR VAL VAL GLY G656 P657 G660 P661 P662 V663 ARG

• Molecule 2: DVL1

Chain H:

ALA TYR THR VAL VAL GLY G656 P657 P658 G659 G660 P661 P662 V663 ARG

• Molecule 2: DVL1

Chain E:

ALA TYR THR VAL VAL GLY G656 P657 P658 P662 V663 ARG

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.22Å 73.14Å 101.84Å 90.00° 94.50° 90.00°	Depositor
Resolution (Å)	79.98 – 2.38 79.98 – 2.38	Depositor EDS
% Data completeness (in resolution range)	99.6 (79.98-2.38) 99.6 (79.98-2.38)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.25 (at 2.40Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.225 , 0.251 0.227 , 0.218	Depositor DCC
R_{free} test set	2438 reflections (5.19%)	wwPDB-VP
Wilson B-factor (Å²)	22.8	Xtriage
Anisotropy	0.842	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.30 , 43.8	EDS
L-test for twinning [2]	$<|L|> = 0.53$, $<L^2> = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9388	wwPDB-VP
Average B, all atoms (Å²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.29 % 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 3.7723e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: NA, EDO, CL

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	0.74	1/2189 (0.0%)	0.75	2/2982 (0.1%)
1	B	0.78	4/2191 (0.2%)	0.77	1/2984 (0.0%)
1	C	0.77	4/2192 (0.2%)	0.77	2/2987 (0.1%)
1	D	0.66	1/2172 (0.0%)	0.77	0/2962
2	E	1.68	0/50	1.93	1/70 (1.4%)
2	F	2.08	1/39 (2.6%)	2.05	1/53 (1.9%)
2	G	0.38	0/50	0.45	0/70
2	H	1.62	0/48	1.40	0/67
All	All	0.77	11/8931 (0.1%)	0.79	7/12175 (0.1%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	397	GLU	CD-OE1	-9.18	1.15	1.25
1	B	547	TRP	CE3-CZ3	-7.79	1.25	1.38
1	C	397	GLU	CD-OE2	-7.68	1.17	1.25
2	F	657	PRO	N-CD	-7.61	1.37	1.47
1	B	453	TRP	CE3-CZ3	-7.57	1.25	1.38

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	553	MET	N-CA-C	7.52	131.30	111.00
2	F	659	GLY	N-CA-C	-7.22	95.05	113.10
1	C	482	ASP	CB-CG-OD1	6.29	123.96	118.30
1	B	297	ASP	CB-CG-OD1	5.64	123.38	118.30
1	C	368	ARG	NE-CZ-NH1	5.58	123.09	120.30

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	2143	0	2044	43	1
1	B	2142	0	2038	65	0
1	C	2143	0	2033	37	0
1	D	2124	0	1994	43	0
2	E	47	0	45	2	0
2	F	37	0	32	2	0
2	G	47	0	45	0	0
2	H	45	0	38	4	0
3	A	1	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
4	A	4	6	6	0	0
5	D	1	0	0	0	0
6	A	158	0	0	11	4
6	B	172	0	0	15	2
6	C	188	0	0	12	2
6	D	122	0	0	9	1
6	E	2	0	0	1	0
6	F	2	0	0	0	0
6	H	1	0	0	0	0
All	All	9382	6	8275	188	7

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 11.

The worst 5 of 188 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:B:455:ASN:OD1	6:B:701:HOH:O	1.54	1.20
1:B:282:VAL:HG22	1:B:565:LEU:CD1	1.75	1.17
1:B:282:VAL:CG2	1:B:565:LEU:HD11	1.81	1.11
1:B:392:ARG:NH2	6:B:703:HOH:O	1.84	1.10
1:D:392:ARG:NH2	6:D:702:HOH:O	1.87	1.06

The worst 5 of 7 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)
6:B:765:HOH:O	6:B:848:HOH:O[2_646]	1.86	0.34
6:B:867:HOH:O	6:D:817:HOH:O[2_656]	2.01	0.19
6:A:836:HOH:O	6:A:839:HOH:O[2_555]	2.01	0.19
6:A:751:HOH:O	6:A:763:HOH:O[2_545]	2.07	0.13
6:C:796:HOH:O	6:C:874:HOH:O[2_645]	2.10	0.10

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	281/301 (93%)	273 (97%)	8 (3%)	0	100	100
1	B	282/301 (94%)	267 (95%)	15 (5%)	0	100	100
1	C	282/301 (94%)	272 (96%)	10 (4%)	0	100	100
1	D	283/301 (94%)	269 (95%)	14 (5%)	0	100	100
2	E	6/15 (40%)	6 (100%)	0	0	100	100
2	F	5/15 (33%)	5 (100%)	0	0	100	100
2	G	6/15 (40%)	6 (100%)	0	0	100	100
2	H	6/15 (40%)	6 (100%)	0	0	100	100
All	All	1151/1264 (91%)	1104 (96%)	47 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	226/248 (91%)	225 (100%)	1 (0%)	91	96
1	B	227/248 (92%)	221 (97%)	6 (3%)	46	64
1	C	227/248 (92%)	226 (100%)	1 (0%)	91	96
1	D	221/248 (89%)	220 (100%)	1 (0%)	88	95
2	E	5/10 (50%)	5 (100%)	0	100	100
2	F	3/10 (30%)	3 (100%)	0	100	100
2	G	5/10 (50%)	5 (100%)	0	100	100
2	H	4/10 (40%)	4 (100%)	0	100	100
All	All	918/1032 (89%)	909 (99%)	9 (1%)	76	87

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	390	SER
1	D	392	ARG
1	B	559	ASP
1	B	338	ASP
1	B	546	SER

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

Mol	Chain	Res	Type
1	B	413	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)

Of 6 ligands modelled in this entry, 5 are monoatomic - leaving 1 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	EDO	A	602	-	3,3,3	0.53	0	2,2,2	0.16	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	602	-	-	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 (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(Å²)	Q<0.9
1	A	285/301 (94%)	0.44	10 (3%)	44	47	11, 20, 34, 68	0
1	B	285/301 (94%)	0.69	13 (4%)	32	35	20, 27, 44, 86	0
1	C	284/301 (94%)	0.37	7 (2%)	57	59	7, 16, 30, 62	0
1	D	286/301 (95%)	0.87	22 (7%)	13	14	27, 34, 47, 77	0
2	E	8/15 (53%)	2.82	4 (50%)	0	0	26, 28, 36, 44	0
2	F	7/15 (46%)	1.35	2 (28%)	0	0	28, 35, 51, 52	0
2	G	8/15 (53%)	2.11	4 (50%)	0	0	27, 29, 38, 42	0
2	H	8/15 (53%)	2.28	4 (50%)	0	0	30, 45, 56, 63	0
All	All	1171/1264 (92%)	0.63	66 (5%)	24	27	7, 25, 44, 86	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	567	GLU	8.3
1	D	279	ALA	8.1
1	B	279	ALA	6.8
2	E	662	PRO	6.7
1	D	352	THR	6.4

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($\text{\AA}^2$)	Q<0.9
3	NA	A	601	1/1	0.78	0.16	30,30,30,30	0
4	EDO	A	602	4/4	0.80	0.33	20,20,20,20	0
3	NA	C	601	1/1	0.81	0.19	30,30,30,30	0
3	NA	B	601	1/1	0.89	0.25	39,39,39,39	0
3	NA	B	602	1/1	0.92	0.12	36,36,36,36	0
5	CL	D	601	1/1	0.93	0.18	30,30,30,30	0

6.5 Other polymers (i)

There are no such residues in this entry.