



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 13, 2018 – 01:36 pm GMT

PDB ID : 4XR8
Title : Crystal structure of the HPV16 E6/E6AP/p53 ternary complex at 2.25 Å resolution
Authors : Martinez-Zapien, D.; Ruiz, F.X.; Mitschler, A.; Podjarny, A.; Trave, G.; Zanier, K.
Deposited on : 2015-01-20
Resolution : 2.25 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk31020
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk31020

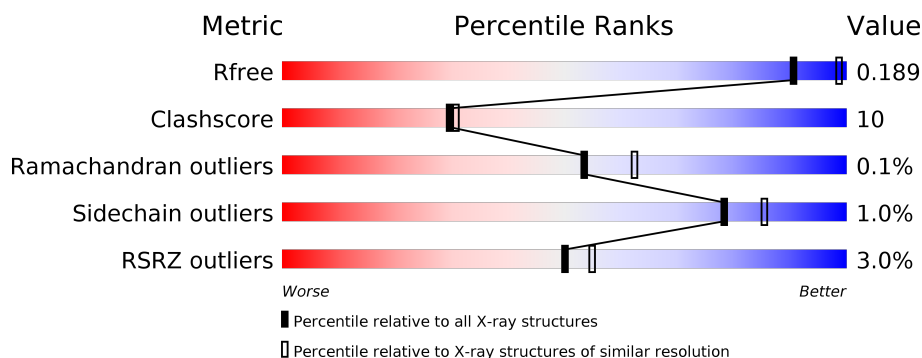
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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}	111664	1178 (2.26-2.26)
Clashscore	122126	1286 (2.26-2.26)
Ramachandran outliers	120053	1253 (2.26-2.26)
Sidechain outliers	120020	1254 (2.26-2.26)
RSRZ outliers	108989	1158 (2.26-2.26)

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	383	<div> <div>3%</div> <div> <div></div> <div>78%</div> <div>21%</div> <div>.</div> </div> </div>
1	B	383	<div> <div>3%</div> <div> <div></div> <div>75%</div> <div>24%</div> <div>..</div> </div> </div>
2	C	199	<div> <div>3%</div> <div> <div></div> <div>84%</div> <div>16%</div> </div> </div>
2	D	199	<div> <div>2%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div>.</div> </div> </div>
3	F	151	<div> <div>3%</div> <div> <div></div> <div>77%</div> <div>17%</div> <div>5%</div> </div> </div>
3	H	151	<div> <div>4%</div> <div> <div></div> <div>79%</div> <div>19%</div> <div>.</div> </div> </div>

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 12034 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 Maltose-binding periplasmic protein, ubiquitin ligase E6AP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	383	Total	C	N	O	S	0	6	0
			2990	1921	489	573	7			
1	B	381	Total	C	N	O	S	0	5	0
			2965	1906	480	573	6			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P0AEX9
A	83	ALA	ASP	engineered mutation	UNP P0AEX9
A	84	ALA	LYS	engineered mutation	UNP P0AEX9
A	240	ALA	LYS	engineered mutation	UNP P0AEX9
A	360	ALA	GLU	engineered mutation	UNP P0AEX9
A	363	ALA	LYS	engineered mutation	UNP P0AEX9
A	364	ALA	ASP	engineered mutation	UNP P0AEX9
A	368	ASN	ARG	linker	UNP P0AEX9
A	369	ALA	-	linker	UNP P0AEX9
A	370	ALA	-	linker	UNP P0AEX9
A	371	ALA	-	linker	UNP P0AEX9
B	1	MET	-	initiating methionine	UNP P0AEX9
B	83	ALA	ASP	engineered mutation	UNP P0AEX9
B	84	ALA	LYS	engineered mutation	UNP P0AEX9
B	240	ALA	LYS	engineered mutation	UNP P0AEX9
B	360	ALA	GLU	engineered mutation	UNP P0AEX9
B	363	ALA	LYS	engineered mutation	UNP P0AEX9
B	364	ALA	ASP	engineered mutation	UNP P0AEX9
B	368	ASN	ARG	linker	UNP P0AEX9
B	369	ALA	-	linker	UNP P0AEX9
B	370	ALA	-	linker	UNP P0AEX9
B	371	ALA	-	linker	UNP P0AEX9

- Molecule 2 is a protein called Cellular tumor antigen p53.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	199	Total	C	N	O	S	0	0	0
			1564	963	291	294	16			
2	D	199	Total	C	N	O	S	0	0	0
			1564	963	291	294	16			

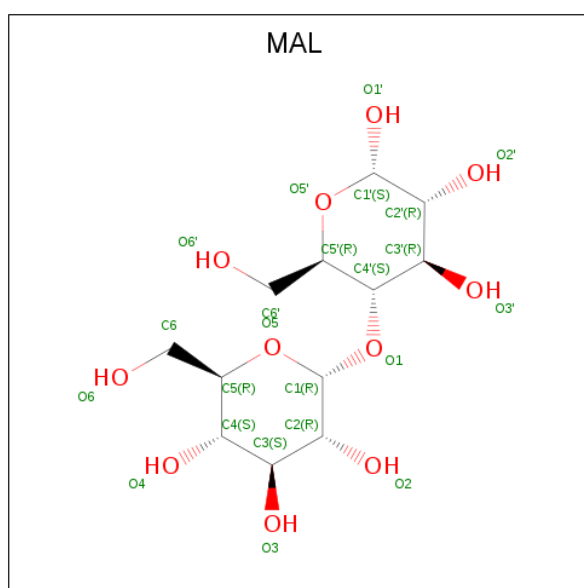
- Molecule 3 is a protein called Protein E6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	143	Total	C	N	O	S	0	2	0
			1218	767	222	216	13			
3	H	151	Total	C	N	O	S	0	1	0
			1287	805	240	230	12			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	80	SER	CYS	engineered mutation	UNP P03126
F	97	SER	CYS	engineered mutation	UNP P03126
F	111	SER	CYS	engineered mutation	UNP P03126
F	140	SER	CYS	engineered mutation	UNP P03126
H	80	SER	CYS	engineered mutation	UNP P03126
H	97	SER	CYS	engineered mutation	UNP P03126
H	111	SER	CYS	engineered mutation	UNP P03126
H	140	SER	CYS	engineered mutation	UNP P03126

- Molecule 4 is MALTOSE (three-letter code: MAL) (formula: $C_{12}H_{22}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			23	12	11		
4	B	1	Total	C	O	0	0
			23	12	11		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		
5	D	1	Total	C	O	0	0
			7	4	3		
5	F	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	2	Total	Zn	0	0
			2	2		
6	D	1	Total	Zn	0	0
			1	1		
6	C	1	Total	Zn	0	0
			1	1		
6	F	2	Total	Zn	0	0
			2	2		

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

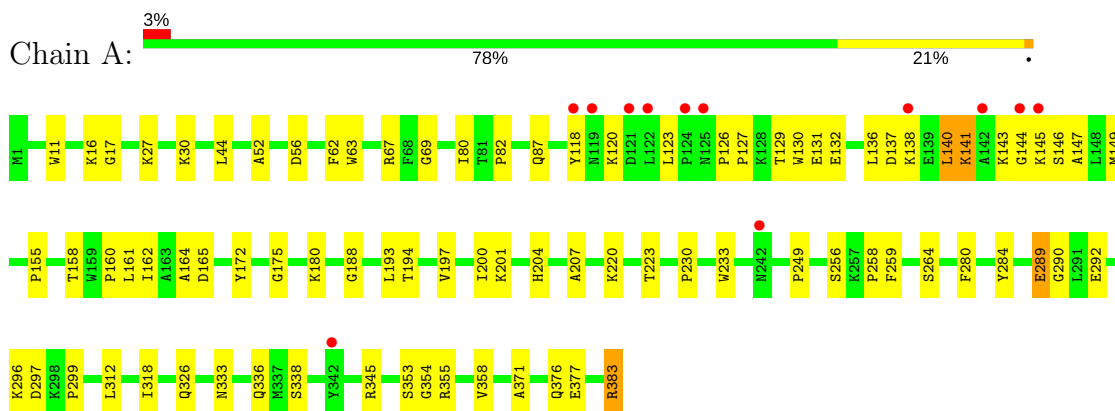
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	91	Total	O	0	2
			94	94		
8	B	62	Total	O	0	1
			62	62		
8	C	83	Total	O	0	2
			85	85		
8	D	56	Total	O	0	2
			58	58		
8	F	35	Total	O	0	0
			35	35		
8	H	27	Total	O	0	1
			28	28		

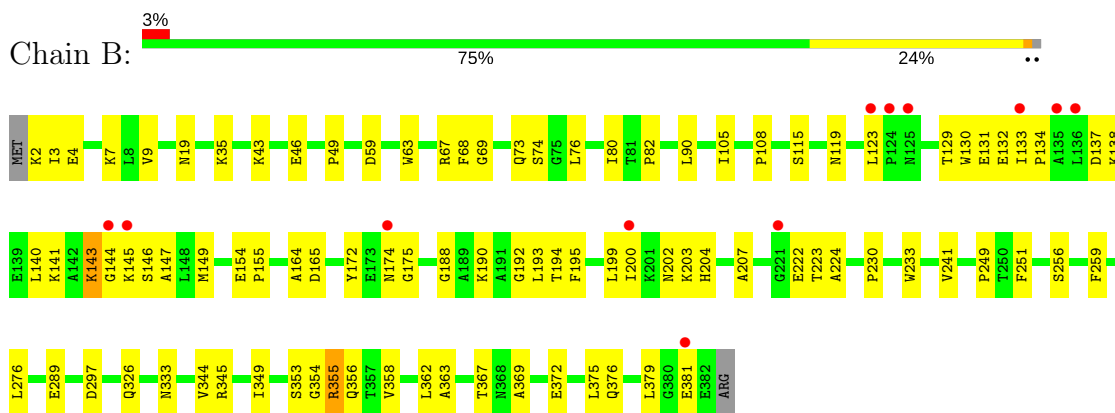
3 Residue-property plots [i](#)

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.

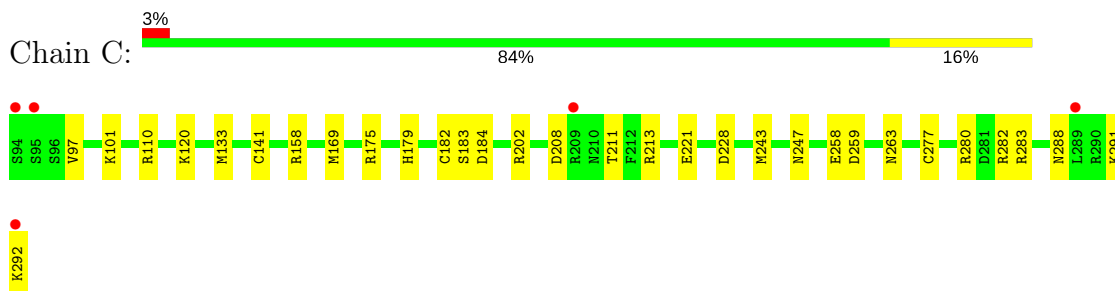
- Molecule 1: Maltose-binding periplasmic protein, ubiquitin ligase E6AP



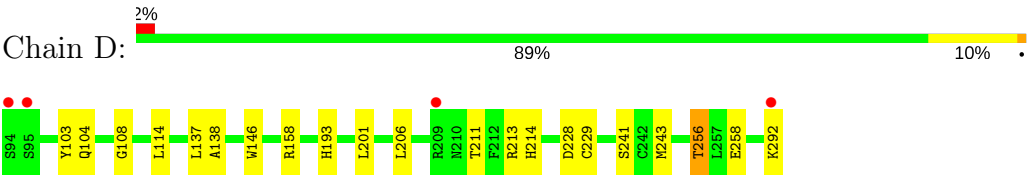
- Molecule 1: Maltose-binding periplasmic protein, ubiquitin ligase E6AP



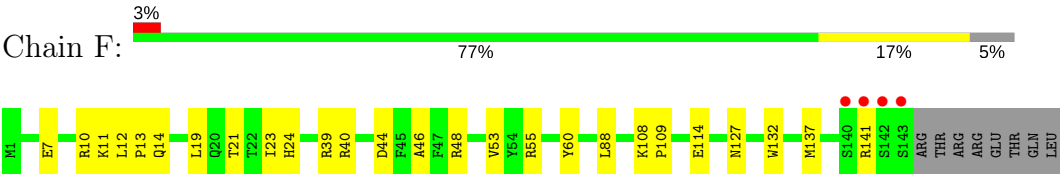
- Molecule 2: Cellular tumor antigen p53



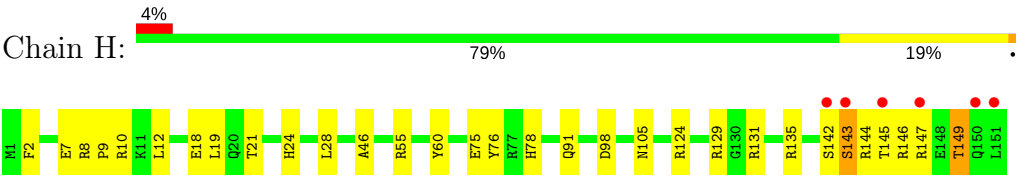
● Molecule 2: Cellular tumor antigen p53



● Molecule 3: Protein E6



● Molecule 3: Protein E6



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	77.84Å 128.94Å 81.56Å 90.00° 92.33° 90.00°	Depositor
Resolution (Å)	49.63 – 2.25 49.64 – 2.23	Depositor EDS
% Data completeness (in resolution range)	98.5 (49.63-2.25) 98.6 (49.64-2.23)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 2.22Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.194 , 0.246 0.200 , 0.189	Depositor DCC
R_{free} test set	3883 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	44.8	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 40.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.004 for l,k,-h 0.029 for h,-k,-l 0.019 for l,-k,h	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12034	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, PEG, MAL, 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	0.51	0/3068	0.65	0/4161
1	B	0.45	0/3039	0.65	0/4128
2	C	0.49	0/1599	0.62	0/2165
2	D	0.47	0/1599	0.62	0/2165
3	F	0.50	0/1251	0.63	0/1681
3	H	0.49	0/1317	0.68	0/1768
All	All	0.48	0/11873	0.64	0/16068

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
1	B	0	1
3	F	0	1
3	H	0	2
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	143	LYS	Peptide
3	F	141	ARG	Peptide
3	H	143	SER	Peptide
3	H	149	THR	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	2990	0	2969	71	0
1	B	2965	0	2927	86	0
2	C	1564	0	1529	24	0
2	D	1564	0	1529	18	0
3	F	1218	0	1213	20	0
3	H	1287	0	1282	35	0
4	A	23	0	22	3	0
4	B	23	0	22	0	0
5	A	7	0	10	0	0
5	B	7	0	10	2	0
5	D	7	0	10	2	0
5	F	7	0	10	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	F	2	0	0	0	0
6	H	2	0	0	0	0
7	F	4	0	6	3	0
8	A	94	0	0	2	0
8	B	62	0	0	4	0
8	C	85	0	0	3	0
8	D	58	0	0	1	0
8	F	35	0	0	0	0
8	H	28	0	0	0	0
All	All	12034	0	11539	237	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 10.

The worst 5 of 237 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)
3:H:105:ASN:HB2	3:H:144:ARG:HH12	1.23	0.97
3:H:105:ASN:CB	3:H:144:ARG:HH12	1.79	0.94
1:B:143:LYS:HB2	1:B:145:LYS:H	1.34	0.92
1:B:141:LYS:HA	1:B:143:LYS:HD3	1.53	0.91

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:355:ARG:HG2	1:A:355:ARG:HH11	1.35	0.90

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	385/383 (100%)	366 (95%)	19 (5%)	0	100	100
1	B	383/383 (100%)	366 (96%)	17 (4%)	0	100	100
2	C	197/199 (99%)	195 (99%)	1 (0%)	1 (0%)	31	31
2	D	197/199 (99%)	195 (99%)	2 (1%)	0	100	100
3	F	143/151 (95%)	137 (96%)	6 (4%)	0	100	100
3	H	150/151 (99%)	137 (91%)	13 (9%)	0	100	100
All	All	1455/1466 (99%)	1396 (96%)	58 (4%)	1 (0%)	53	62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	183	SER

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	305/300 (102%)	298 (98%)	7 (2%)	53	62
1	B	302/300 (101%)	299 (99%)	3 (1%)	78	85
2	C	179/179 (100%)	179 (100%)	0	100	100
2	D	179/179 (100%)	178 (99%)	1 (1%)	87	92
3	F	139/145 (96%)	139 (100%)	0	100	100
3	H	146/145 (101%)	144 (99%)	2 (1%)	69	79
All	All	1250/1248 (100%)	1237 (99%)	13 (1%)	78	85

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	383[A]	ARG
1	A	383[B]	ARG
2	D	256	THR
1	A	289	GLU
1	B	355	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	87	GLN
1	A	204	HIS
1	B	376	GLN
2	D	214	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 13 ligands modelled in this entry, 6 are monoatomic - leaving 7 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	MAL	A	401	-	24,24,24	0.57	0	35,35,35	0.95	0
5	PEG	A	402	-	6,6,6	0.98	0	5,5,5	0.72	0
4	MAL	B	401	-	24,24,24	0.61	0	35,35,35	0.96	2 (5%)
5	PEG	B	402	-	6,6,6	0.87	0	5,5,5	0.41	0
5	PEG	D	902	-	6,6,6	1.11	0	5,5,5	0.74	0
7	EDO	F	203	-	3,3,3	0.43	0	2,2,2	0.45	0
5	PEG	F	204	-	6,6,6	0.94	0	5,5,5	0.55	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	MAL	A	401	-	-	0/8/48/48	0/2/2/2
5	PEG	A	402	-	-	0/4/4/4	0/0/0/0
4	MAL	B	401	-	-	0/8/48/48	0/2/2/2
5	PEG	B	402	-	-	0/4/4/4	0/0/0/0
5	PEG	D	902	-	-	0/4/4/4	0/0/0/0
7	EDO	F	203	-	-	0/1/1/1	0/0/0/0
5	PEG	F	204	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	401	MAL	C6-C5-C4	-2.08	108.09	112.99
4	B	401	MAL	O2'-C2'-C3'	-2.02	105.64	110.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	401	MAL	3	0
5	B	402	PEG	2	0
5	D	902	PEG	2	0
7	F	203	EDO	3	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	383/383 (100%)	0.08	12 (3%) 49 53	24, 50, 89, 101	0
1	B	381/383 (99%)	0.13	12 (3%) 49 53	31, 58, 89, 103	0
2	C	199/199 (100%)	-0.24	5 (2%) 57 61	29, 43, 69, 98	0
2	D	199/199 (100%)	-0.19	4 (2%) 65 69	31, 45, 72, 111	0
3	F	143/151 (94%)	-0.04	4 (2%) 53 57	33, 49, 73, 104	0
3	H	151/151 (100%)	0.06	6 (3%) 38 42	33, 55, 85, 93	0
All	All	1456/1466 (99%)	-0.00	43 (2%) 50 55	24, 50, 85, 111	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	94	SER	13.1
2	C	94	SER	10.5
3	H	151	LEU	9.0
3	F	142	SER	6.9
1	A	144	GLY	5.8

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. 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
5	PEG	A	402	7/7	0.80	0.19	32,48,55,56	0
5	PEG	D	902	7/7	0.85	0.18	41,45,54,60	0
6	ZN	H	202	1/1	0.88	0.09	58,58,58,58	0
5	PEG	B	402	7/7	0.89	0.19	56,57,62,64	0
4	MAL	B	401	23/23	0.92	0.16	44,48,53,63	0
5	PEG	F	204	7/7	0.94	0.11	43,49,52,61	0
6	ZN	F	202	1/1	0.94	0.07	49,49,49,49	0
7	EDO	F	203	4/4	0.95	0.32	50,54,54,55	0
4	MAL	A	401	23/23	0.96	0.16	32,37,46,48	0
6	ZN	H	201	1/1	0.98	0.11	42,42,42,42	0
6	ZN	F	201	1/1	0.99	0.10	47,47,47,47	0
6	ZN	D	901	1/1	0.99	0.14	40,40,40,40	0
6	ZN	C	900	1/1	0.99	0.16	42,42,42,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.