



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 7, 2023 – 05:48 AM EST

PDB ID : 6WHZ  
Title : Histone deacetylases complex with peptide macrocycles  
Authors : Bera, A.K.; Hosseinzadeh, P.; Watson, P.; Baker, D.  
Deposited on : 2020-04-08  
Resolution : 2.90 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](mailto: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 : 2.36  
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) : 2.36

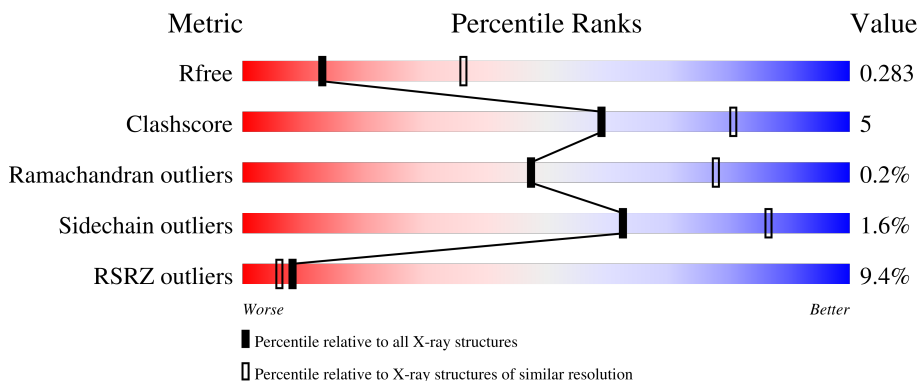
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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  $\geq 3$ , 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	385	
1	B	385	
1	C	385	
2	D	8	
2	F	8	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	8	

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
2	DLY	F	505	-	-	-	X
2	HYP	G	502	-	-	-	X
2	DLY	G	505	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9197 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 Histone deacetylase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	372	Total 2980	C 1901	N 505	O 550	S 24	0	0	0
1	B	364	Total 2935	C 1874	N 495	O 542	S 24	0	0	0
1	C	365	Total 2936	C 1874	N 495	O 543	S 24	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	ALA	-	expression tag	UNP Q92769
B	5	ALA	-	expression tag	UNP Q92769
C	5	ALA	-	expression tag	UNP Q92769

- Molecule 2 is a protein called U2M-ASN-HYP-LYS-GLN-DLY-TRP-GLY peptide macrocycle.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	8	Total 71	C 46	N 13	O 11	S 1	0	0	0
2	F	8	Total 54	C 32	N 10	O 11	S 1	0	0	0
2	G	8	Total 54	C 32	N 10	O 11	S 1	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Zn 1	0	0
3	B	1	Total 1	Zn 1	0	0

*Continued on next page...*

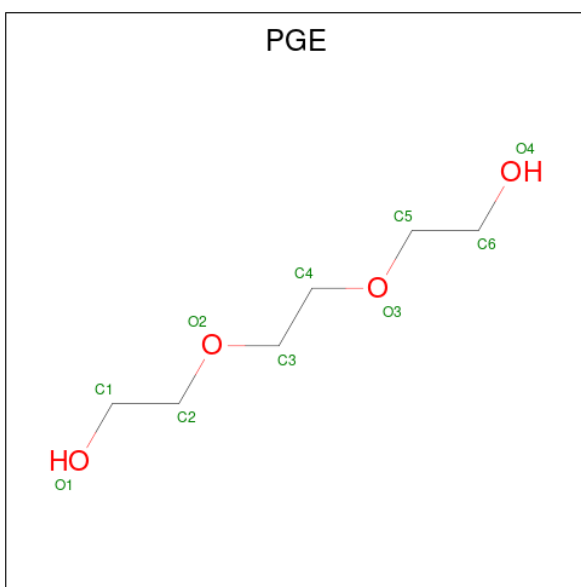
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	Zn	0	0
			1	1		

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			10	6	4		
5	A	1	Total	C	O	0	0
			10	6	4		
5	B	1	Total	C	O	0	0
			10	6	4		
5	C	1	Total	C	O	0	0
			10	6	4		

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	C O	0	0
			13	8 5		


- Molecule 7 is water.

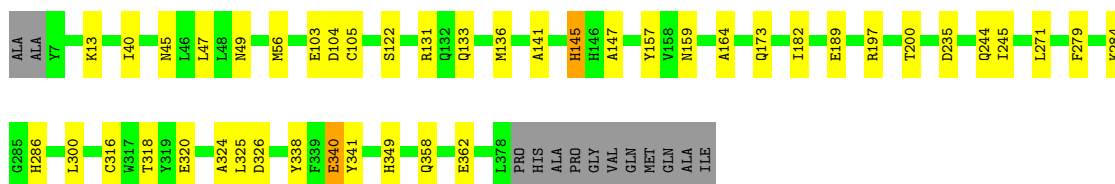
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	46	Total	O	0	0
			46	46		
7	B	32	Total	O	0	0
			32	32		
7	C	22	Total	O	0	0
			22	22		
7	D	5	Total	O	0	0
			5	5		
7	G	1	Total	O	0	0
			1	1		

### 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.

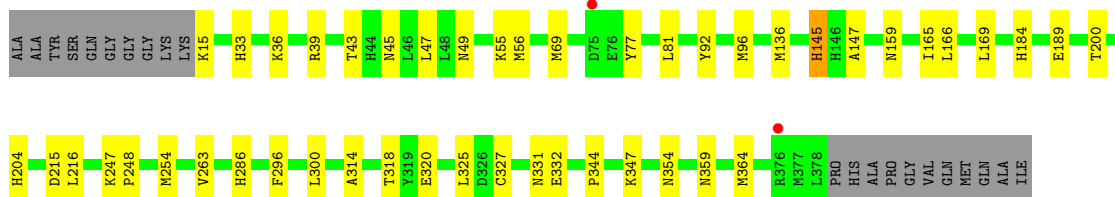
- Molecule 1: Histone deacetylase 2

Chain A: 




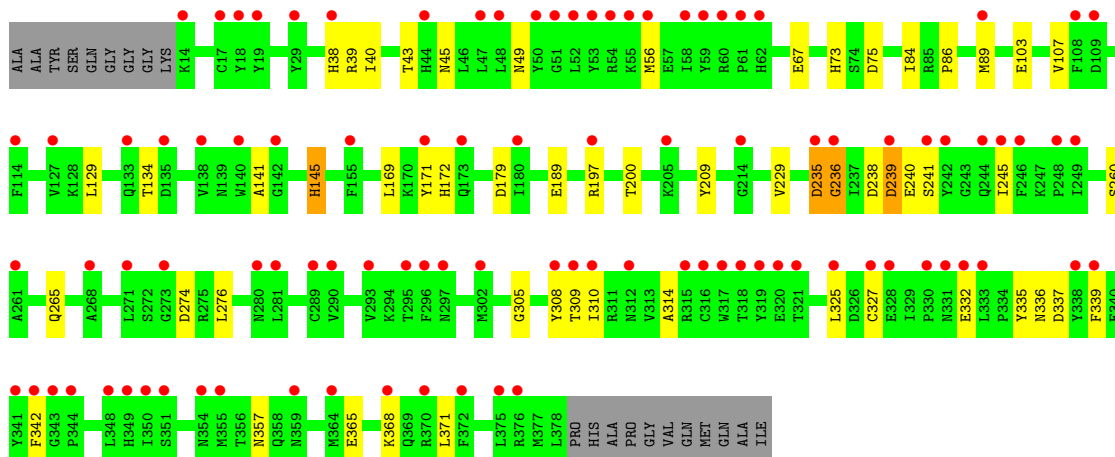
- Molecule 1: Histone deacetylase 2

Chain B: 

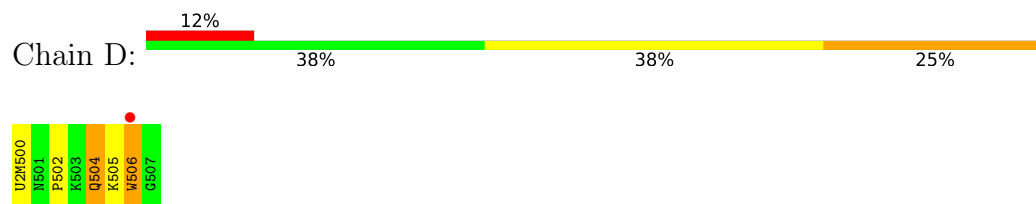


- Molecule 1: Histone deacetylase 2

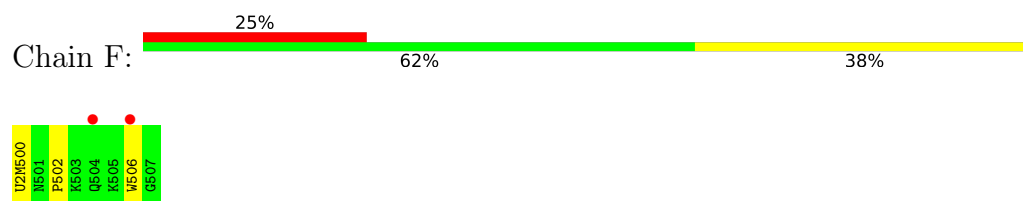
Chain C: 



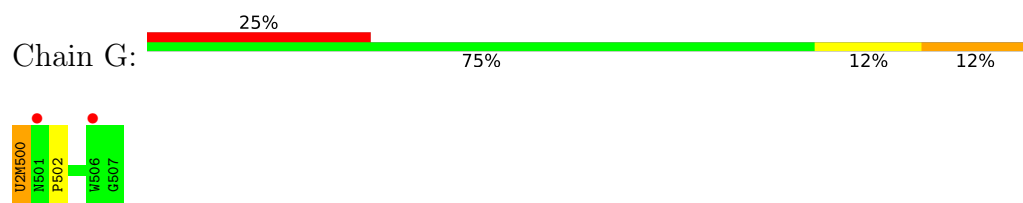
- Molecule 2: U2M-ASN-HYP-LYS-GLN-DLY-TRP-GLY peptide macrocycle



- Molecule 2: U2M-ASN-HYP-LYS-GLN-DLY-TRP-GLY peptide macrocycle



- Molecule 2: U2M-ASN-HYP-LYS-GLN-DLY-TRP-GLY peptide macrocycle





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.48Å 94.80Å 138.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.90 – 2.90 47.90 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.90-2.90) 100.0 (47.90-2.90)	Depositor EDS
$R_{merge}$	0.29	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.17rc1_3605	Depositor
R, $R_{free}$	0.227 , 0.281 0.229 , 0.283	Depositor DCC
$R_{free}$ test set	1312 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	60.0	Xtrriage
Anisotropy	0.226	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 49.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.012 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	9197	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: PG4, ZN, NA, DLY, HYP, PGE, U2M

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	0.25	0/3056	0.40	0/4124
1	B	0.25	0/3011	0.39	0/4066
1	C	0.24	0/3012	0.38	0/4069
2	D	0.18	0/43	0.29	0/53
2	F	0.18	0/28	0.63	0/33
2	G	0.20	0/28	0.34	0/33
All	All	0.24	0/9178	0.39	0/12378

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 [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	2980	0	2876	25	0
1	B	2935	0	2835	29	0
1	C	2936	0	2826	30	0
2	D	71	0	58	4	0
2	F	54	0	29	0	0
2	G	54	0	29	1	0
3	A	1	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	1	0	0	0	0
5	A	20	0	28	0	0
5	B	10	0	14	0	0
5	C	10	0	14	0	0
6	B	13	0	18	1	0
7	A	46	0	0	0	0
7	B	32	0	0	1	0
7	C	22	0	0	0	0
7	D	5	0	0	0	0
7	G	1	0	0	0	0
All	All	9197	0	8727	84	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 5.

All (84) 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:43:THR:HG22	1:B:314:ALA:HA	1.63	0.79
1:B:55:LYS:NZ	1:B:327:CYS:SG	2.64	0.70
1:A:136:MET:HG2	1:A:300:LEU:HB3	1.73	0.69
1:C:235:ASP:OD1	1:C:235:ASP:N	2.25	0.67
1:C:239:ASP:N	1:C:239:ASP:OD1	2.27	0.67
1:C:49:ASN:ND2	1:C:335:TYR:O	2.25	0.63
1:C:43:THR:HG22	1:C:314:ALA:HA	1.81	0.61
1:B:33:HIS:O	1:B:36:LYS:NZ	2.32	0.61
1:B:332:GLU:OE1	1:B:347:LYS:NZ	2.35	0.60
1:B:136:MET:HG2	1:B:300:LEU:HB3	1.83	0.59
1:B:354:ASN:HD22	1:C:171:TYR:HA	1.69	0.58
1:C:189:GLU:HA	1:C:200:THR:HG21	1.85	0.58
1:C:39:ARG:NH1	1:C:305:GLY:O	2.32	0.57
1:A:244:GLN:HG3	1:A:245:ILE:HG23	1.88	0.56
1:C:40:ILE:HD13	1:C:141:ALA:HA	1.88	0.55
1:C:308:TYR:OH	2:G:500:U2M:SH	2.65	0.55
1:B:286:HIS:ND1	1:B:320:GLU:OE2	2.40	0.55
1:C:39:ARG:HB3	1:C:310:ILE:HG22	1.88	0.54
1:A:103:GLU:OE2	2:D:504:GLN:NE2	2.39	0.54

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:274:ASP:OD1	1:C:309:THR:OG1	2.24	0.54
1:B:15:LYS:N	7:B:501:HOH:O	2.40	0.53
1:C:45:ASN:ND2	1:C:337:ASP:OD1	2.43	0.52
1:A:235:ASP:HB2	1:A:358:GLN:HG3	1.92	0.51
1:B:347:LYS:NZ	1:C:73:HIS:O	2.44	0.51
1:C:229:VAL:HG13	1:C:371:LEU:HD22	1.92	0.50
1:C:336:ASN:HD21	1:C:342:PHE:HE2	1.60	0.50
1:A:271:LEU:HD11	1:A:316:CYS:HB2	1.94	0.50
1:A:47:LEU:HD23	1:A:318:THR:HG23	1.94	0.50
1:C:365:GLU:HA	1:C:368:LYS:HE2	1.94	0.50
1:C:209:TYR:OH	1:C:276:LEU:O	2.28	0.49
1:A:284:LYS:HG3	1:A:349:HIS:CD2	2.48	0.49
1:A:145:HIS:H	1:A:145:HIS:CD2	2.31	0.49
1:A:286:HIS:ND1	1:A:320:GLU:OE2	2.44	0.49
2:D:505:DLY:HB3	2:D:506:TRP:CD2	2.47	0.48
1:C:145:HIS:ND1	1:C:179:ASP:OD2	2.40	0.48
1:B:215:ASP:OD1	1:B:216:LEU:N	2.47	0.48
1:C:236:GLY:O	1:C:357:ASN:HA	2.13	0.48
1:C:238:ASP:OD1	1:C:241:SER:N	2.36	0.48
1:C:39:ARG:HG3	1:C:40:ILE:HG13	1.95	0.48
1:B:184:HIS:HB2	1:B:204:HIS:CD2	2.49	0.47
1:A:45:ASN:O	1:A:49:ASN:ND2	2.44	0.47
1:B:147:ALA:HB3	1:B:159:ASN:HB2	1.97	0.47
1:A:362:GLU:HG2	1:B:344:PRO:HB3	1.95	0.47
1:B:166:LEU:HD23	1:B:169:LEU:HD12	1.95	0.47
1:B:189:GLU:HA	1:B:200:THR:HG21	1.96	0.47
1:A:136:MET:HE3	1:A:324:ALA:HB1	1.96	0.47
1:A:340:GLU:H	1:A:340:GLU:HG3	1.45	0.46
1:B:331:ASN:N	1:C:75:ASP:OD2	2.48	0.46
1:A:131:ARG:HH11	1:A:133:GLN:HE22	1.64	0.46
1:B:56:MET:HB3	1:B:325:LEU:HD21	1.97	0.46
1:B:47:LEU:HD23	1:B:318:THR:HG23	1.96	0.46
1:C:169:LEU:O	1:C:197:ARG:NH1	2.49	0.46
2:D:504:GLN:HG2	2:D:505:DLY:HG3	1.98	0.46
1:B:45:ASN:O	1:B:49:ASN:ND2	2.38	0.46
1:A:173:GLN:HA	1:A:197:ARG:HH11	1.81	0.45
1:C:56:MET:HB3	1:C:325:LEU:HD21	1.99	0.45
1:A:56:MET:HB3	1:A:325:LEU:HD21	1.99	0.45
1:B:36:LYS:O	1:B:39:ARG:HG2	2.16	0.45
1:B:359:ASN:HB3	1:B:364:MET:SD	2.57	0.45
1:C:245:ILE:HG13	1:C:368:LYS:HD3	1.98	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:84:ILE:HG13	1:C:107:VAL:HG11	2.00	0.44
1:A:13:LYS:HE2	1:A:326:ASP:HB2	2.00	0.43
1:A:105:CYS:HB3	1:A:157:TYR:CE2	2.52	0.43
1:B:92:TYR:O	1:B:96:MET:HB2	2.19	0.43
1:B:77:TYR:CZ	1:B:81:LEU:HD11	2.53	0.43
1:B:69:MET:HB2	1:B:69:MET:HE3	1.90	0.43
1:A:122:SER:HB3	1:A:164:ALA:HB2	2.00	0.43
1:A:147:ALA:HB3	1:A:159:ASN:HB2	2.00	0.42
1:B:39:ARG:O	1:B:43:THR:HG23	2.18	0.42
1:A:189:GLU:HA	1:A:200:THR:HG21	2.00	0.42
2:D:505:DLY:HB3	2:D:506:TRP:CE3	2.54	0.42
1:C:172:HIS:HB3	1:C:260:SER:HB2	2.01	0.42
1:B:145:HIS:CD2	1:B:145:HIS:H	2.36	0.42
6:B:404:PG4:H51	6:B:404:PG4:H72	1.81	0.42
1:C:129:LEU:HD23	1:C:134:THR:HG23	2.02	0.41
1:C:265:GLN:NE2	1:C:305:GLY:HA3	2.35	0.41
1:B:247:LYS:HB3	1:B:248:PRO:HD3	2.02	0.41
1:A:182:ILE:HG13	1:A:279:PHE:CE1	2.56	0.41
1:A:40:ILE:HD13	1:A:141:ALA:HA	2.02	0.41
1:A:338:TYR:HB3	1:A:341:TYR:CD1	2.56	0.41
1:A:103:GLU:HB3	1:A:104:ASP:H	1.74	0.40
1:B:165:ILE:HG12	1:B:263:VAL:HG21	2.03	0.40
1:B:254:MET:HG3	1:B:296:PHE:CZ	2.56	0.40
1:C:86:PRO:HA	1:C:89:MET:HG3	2.02	0.40

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	370/385 (96%)	355 (96%)	15 (4%)	0	<a href="#">100</a> <a href="#">100</a>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	362/385 (94%)	345 (95%)	17 (5%)	0	100	100
1	C	363/385 (94%)	340 (94%)	22 (6%)	1 (0%)	41	71
2	D	4/8 (50%)	4 (100%)	0	0	100	100
2	F	4/8 (50%)	2 (50%)	1 (25%)	1 (25%)	0	0
2	G	4/8 (50%)	2 (50%)	2 (50%)	0	100	100
All	All	1107/1179 (94%)	1048 (95%)	57 (5%)	2 (0%)	47	78

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	506	TRP
1	C	236	GLY

### 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	316/327 (97%)	314 (99%)	2 (1%)	86	96
1	B	314/327 (96%)	313 (100%)	1 (0%)	92	98
1	C	313/327 (96%)	303 (97%)	10 (3%)	39	73
2	D	4/4 (100%)	2 (50%)	2 (50%)	0	0
2	F	2/4 (50%)	2 (100%)	0	100	100
2	G	2/4 (50%)	2 (100%)	0	100	100
All	All	951/993 (96%)	936 (98%)	15 (2%)	62	86

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	145	HIS
1	A	340	GLU
1	B	145	HIS
1	C	38	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	67	GLU
1	C	103	GLU
1	C	145	HIS
1	C	235	ASP
1	C	239	ASP
1	C	240	GLU
1	C	327	CYS
1	C	332	GLU
1	C	339	PHE
2	D	504	GLN
2	D	506	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

9 non-standard protein/DNA/RNA residues 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
2	HYP	G	502	2	6,8,9	4.79	4 (66%)	5,10,12	2.06	2 (40%)
2	U2M	G	500	3,2	8,9,10	3.32	2 (25%)	4,9,11	1.49	1 (25%)
2	HYP	F	502	2	6,8,9	4.84	4 (66%)	5,10,12	1.01	0
2	U2M	D	500	3,2	8,9,10	3.35	2 (25%)	4,9,11	1.31	0
2	HYP	D	502	2	6,8,9	4.83	4 (66%)	5,10,12	1.66	2 (40%)
2	U2M	F	500	3,2	8,9,10	3.26	2 (25%)	4,9,11	1.16	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
2	HYP	G	502	2	-	0/0/11/13	0/1/1/1
2	U2M	G	500	3,2	-	3/7/8/10	-
2	HYP	F	502	2	-	0/0/11/13	0/1/1/1
2	U2M	D	500	3,2	-	4/7/8/10	-
2	HYP	D	502	2	-	0/0/11/13	0/1/1/1
2	U2M	F	500	3,2	-	1/7/8/10	-

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	502	HYP	CB-CG	-10.05	1.33	1.52
2	D	502	HYP	CB-CG	-9.96	1.33	1.52
2	G	502	HYP	CB-CG	-9.84	1.33	1.52
2	D	500	U2M	CA-N	-7.34	1.26	1.48
2	G	500	U2M	CA-N	-7.28	1.26	1.48
2	F	500	U2M	CA-N	-7.10	1.26	1.48
2	D	500	U2M	CZ-SH	-5.98	1.60	1.80
2	G	500	U2M	CZ-SH	-5.91	1.60	1.80
2	F	500	U2M	CZ-SH	-5.85	1.60	1.80
2	G	502	HYP	CB-CA	5.08	1.65	1.54
2	D	502	HYP	CB-CA	4.95	1.65	1.54
2	F	502	HYP	CB-CA	4.84	1.65	1.54
2	F	502	HYP	CD-N	2.84	1.57	1.47
2	D	502	HYP	CD-N	2.79	1.57	1.47
2	G	502	HYP	CD-N	2.73	1.57	1.47
2	F	502	HYP	CD-CG	2.16	1.58	1.53
2	D	502	HYP	CD-CG	2.13	1.58	1.53
2	G	502	HYP	CD-CG	2.06	1.58	1.53

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	502	HYP	CG-CB-CA	3.40	108.25	103.96
2	G	502	HYP	CB-CG-CD	2.60	106.46	103.27
2	G	500	U2M	CD-CE-CZ	-2.45	108.73	113.09
2	D	502	HYP	CG-CB-CA	2.31	106.88	103.96
2	D	502	HYP	CB-CG-CD	2.25	106.03	103.27

There are no chirality outliers.



All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	500	U2M	N-CA-CB-CG
2	D	500	U2M	C-CA-CB-CG
2	G	500	U2M	CA-CB-CG-CD
2	F	500	U2M	CA-CB-CG-CD
2	G	500	U2M	N-CA-CB-CG
2	D	500	U2M	CA-CB-CG-CD
2	D	500	U2M	CG-CD-CE-CZ
2	G	500	U2M	CG-CD-CE-CZ

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	500	U2M	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 8 are monoatomic - leaving 5 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	$\# Z  > 2$
5	PGE	B	405	-	9,9,9	0.52	0	8,8,8	0.25	0
6	PG4	B	404	-	12,12,12	0.53	0	11,11,11	0.25	0
5	PGE	C	403	-	9,9,9	0.52	0	8,8,8	0.26	0
5	PGE	A	405	-	9,9,9	0.52	0	8,8,8	0.24	0
5	PGE	A	404	-	9,9,9	0.52	0	8,8,8	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
5	PGE	B	405	-	-	4/7/7/7	-
6	PG4	B	404	-	-	8/10/10/10	-
5	PGE	C	403	-	-	1/7/7/7	-
5	PGE	A	405	-	-	4/7/7/7	-
5	PGE	A	404	-	-	2/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	404	PG4	C3-C4-O3-C5
5	A	404	PGE	O3-C5-C6-O4
6	B	404	PG4	O2-C3-C4-O3
5	A	405	PGE	O2-C3-C4-O3
6	B	404	PG4	C4-C3-O2-C2
5	B	405	PGE	C1-C2-O2-C3
5	A	405	PGE	C6-C5-O3-C4
5	B	405	PGE	O3-C5-C6-O4
5	C	403	PGE	C3-C4-O3-C5
6	B	404	PG4	C6-C5-O3-C4
5	A	405	PGE	C3-C4-O3-C5
5	B	405	PGE	C3-C4-O3-C5
6	B	404	PG4	C5-C6-O4-C7
6	B	404	PG4	O4-C7-C8-O5
5	A	404	PGE	O2-C3-C4-O3
5	A	405	PGE	O1-C1-C2-O2
6	B	404	PG4	O1-C1-C2-O2
6	B	404	PG4	O3-C5-C6-O4
5	B	405	PGE	C4-C3-O2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	404	PG4	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	372/385 (96%)	0.10	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	34, 48, 68, 82	0
1	B	364/385 (94%)	0.13	2 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">91</span>	32, 51, 75, 87	2 (0%)
1	C	365/385 (94%)	1.33	98 (26%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	70, 89, 99, 110	64 (17%)
2	D	5/8 (62%)	0.65	1 (20%) <span style="border: 1px solid red; padding: 2px;">1</span> <span style="border: 1px solid red; padding: 2px;">0</span>	59, 62, 72, 73	0
2	F	5/8 (62%)	1.44	2 (40%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	92, 98, 112, 119	2 (40%)
2	G	5/8 (62%)	1.82	2 (40%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	105, 108, 111, 118	4 (80%)
All	All	1116/1179 (94%)	0.53	105 (9%) <span style="border: 1px solid red; padding: 2px;">8</span> <span style="border: 1px solid red; padding: 2px;">6</span>	32, 57, 96, 119	72 (6%)

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	348	LEU	6.0
1	C	52	LEU	4.8
1	C	58	ILE	4.5
1	C	59	TYR	4.5
1	C	44	HIS	4.4
1	C	317	TRP	4.4
1	C	344	PRO	4.4
1	C	51	GLY	4.3
1	C	327	CYS	4.3
1	C	350	ILE	4.3
1	C	328	GLU	4.2
1	C	372	PHE	4.1
1	C	242	TYR	4.1
1	C	244	GLN	4.1
1	C	48	LEU	4.0
1	C	354	ASN	3.9
1	C	293	VAL	3.9
1	C	319	TYR	3.8
1	C	315	ARG	3.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	G	506	TRP	3.8
1	C	318	THR	3.7
1	C	320	GLU	3.7
1	C	56	MET	3.7
1	C	245	ILE	3.7
1	C	108	PHE	3.6
1	C	289	CYS	3.6
1	C	155	PHE	3.6
1	C	18	TYR	3.5
1	C	341	TYR	3.5
1	C	351	SER	3.5
1	C	370	ARG	3.5
1	C	349	HIS	3.5
1	C	355	MET	3.4
1	C	364	MET	3.3
1	C	273	GLY	3.1
1	C	62	HIS	3.1
1	C	138	VAL	3.1
1	C	376	ARG	3.0
1	C	302	MET	3.0
1	C	140	TRP	3.0
1	C	310	ILE	3.0
1	C	180	ILE	2.9
1	C	54	ARG	2.9
1	C	308	TYR	2.9
1	C	17	CYS	2.9
1	C	61	PRO	2.9
1	C	114	PHE	2.8
1	C	280	ASN	2.8
1	C	297	ASN	2.8
1	C	214	GLY	2.8
1	C	29	TYR	2.7
1	C	309	THR	2.7
1	C	333	LEU	2.7
1	C	343	GLY	2.7
2	F	504	GLN	2.6
1	C	261	ALA	2.6
1	C	133	GLN	2.6
1	C	312	ASN	2.6
1	C	197	ARG	2.6
1	C	342	PHE	2.6
1	C	295	THR	2.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	332	GLU	2.6
1	C	47	LEU	2.5
1	C	171	TYR	2.5
1	C	268	ALA	2.5
1	C	359	ASN	2.5
1	C	142	GLY	2.5
1	C	271	LEU	2.5
1	C	19	TYR	2.5
1	C	239	ASP	2.5
1	C	249	ILE	2.4
1	C	330	PRO	2.4
1	C	248	PRO	2.4
1	C	331	ASN	2.4
1	C	38	HIS	2.4
1	C	368	LYS	2.4
1	C	290	VAL	2.4
1	C	246	PHE	2.3
1	C	173	GLN	2.3
1	C	325	LEU	2.3
2	F	506	TRP	2.3
1	C	236	GLY	2.3
1	C	296	PHE	2.3
1	C	241	SER	2.3
1	C	321	THR	2.2
1	C	53	TYR	2.2
1	C	55	LYS	2.2
1	C	235	ASP	2.2
1	C	281	LEU	2.2
1	C	338	TYR	2.2
1	C	60	ARG	2.2
1	C	135	ASP	2.1
1	C	89	MET	2.1
1	C	316	CYS	2.1
1	C	375	LEU	2.1
1	B	75	ASP	2.1
1	C	109	ASP	2.1
1	C	14	LYS	2.1
2	D	506	TRP	2.1
1	C	127	VAL	2.1
1	C	205	LYS	2.1
2	G	501	ASN	2.1
1	C	50	TYR	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	376	ARG	2.0
1	C	339	PHE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	DLY	G	505	5/10	0.46	0.72	113,114,122,123	5
2	DLY	F	505	5/10	0.66	0.45	109,110,111,111	0
2	HYP	G	502	8/9	0.66	0.44	98,112,116,119	8
2	HYP	F	502	8/9	0.82	0.22	105,108,111,120	0
2	U2M	G	500	10/11	0.82	0.19	79,85,101,105	2
2	DLY	D	505	9/10	0.88	0.28	63,67,79,79	0
2	HYP	D	502	8/9	0.93	0.21	47,62,69,76	0
2	U2M	F	500	10/11	0.96	0.17	37,63,83,84	0
2	U2M	D	500	10/11	0.96	0.25	42,47,58,60	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	PGE	A	404	10/10	0.68	0.25	58,68,69,70	0
5	PGE	B	405	10/10	0.76	0.26	61,73,79,83	0
6	PG4	B	404	13/13	0.80	0.37	64,69,80,81	0
4	NA	B	403	1/1	0.83	0.25	59,59,59,59	0
5	PGE	A	405	10/10	0.84	0.39	61,71,75,76	0
5	PGE	C	403	10/10	0.86	0.32	59,68,75,76	0
4	NA	A	403	1/1	0.87	0.13	45,45,45,45	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NA	A	402	1/1	0.88	0.15	25,25,25,25	0
4	NA	C	402	1/1	0.94	0.10	73,73,73,73	0
4	NA	B	402	1/1	0.96	0.18	30,30,30,30	0
3	ZN	B	401	1/1	0.99	0.16	41,41,41,41	0
3	ZN	C	401	1/1	0.99	0.07	73,73,73,73	0
3	ZN	A	401	1/1	0.99	0.15	44,44,44,44	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.