



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 12:31 am BST

PDB ID : 4YU2  
Title : Crystal structure of DYRK1A with harmine-derivatized AnnH-75 inhibitor  
Authors : Chaikuad, A.; Wurzlbauer, A.; Nowak, R.; von Delft, F.; Arrowsmith, C.H.;  
Edwards, A.M.; Bountra, C.; Bracher, F.; Knapp, S.; Structural Genomics  
Consortium (SGC)  
Deposited on : 2015-03-18  
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 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 : 2.11  
buster-report : 1.1.7 (2018)  
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) : 2.11

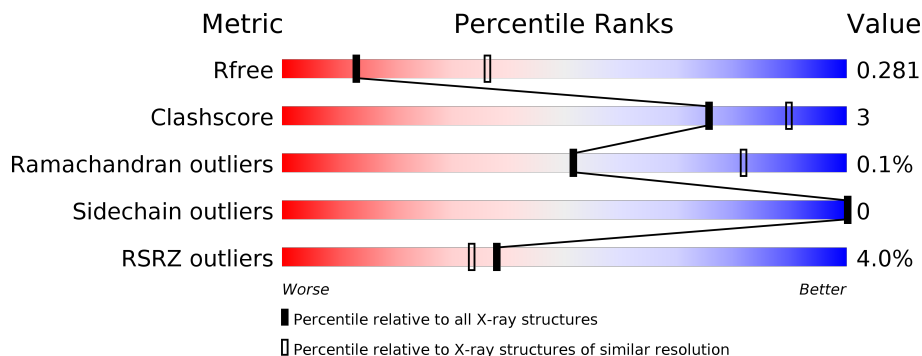
# 1 Overall quality at a glance i

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 on 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	361	
1	B	361	
1	C	361	
1	D	361	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 11762 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.

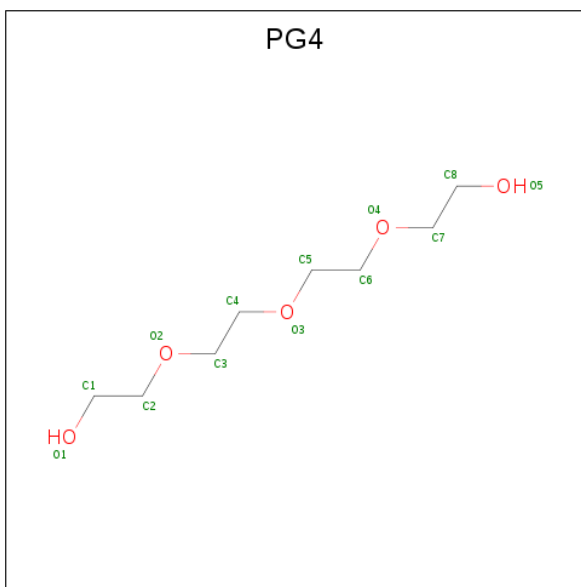
- Molecule 1 is a protein called Dual specificity tyrosine-phosphorylation-regulated kinase 1A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	349	2846	1827	486	514	2	17	0	1	0
1	B	348	2821	1810	479	513	2	17	0	1	0
1	C	347	2828	1816	485	508	2	17	0	2	0
1	D	347	2826	1813	483	512	2	16	0	2	0

There are 8 discrepancies between the modelled and reference sequences:

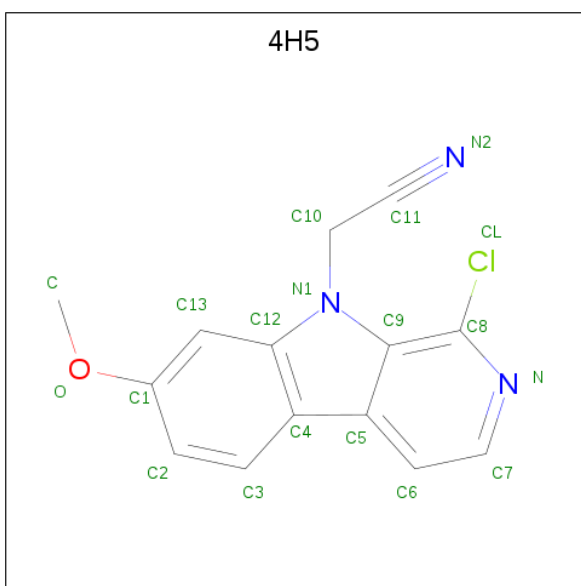
Chain	Residue	Modelled	Actual	Comment	Reference
A	125	SER	-	expression tag	UNP Q13627
A	126	MET	-	expression tag	UNP Q13627
B	125	SER	-	expression tag	UNP Q13627
B	126	MET	-	expression tag	UNP Q13627
C	125	SER	-	expression tag	UNP Q13627
C	126	MET	-	expression tag	UNP Q13627
D	125	SER	-	expression tag	UNP Q13627
D	126	MET	-	expression tag	UNP Q13627

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



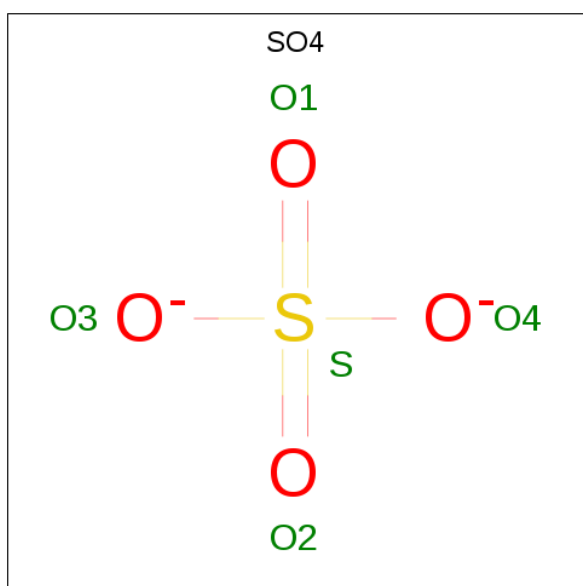
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 13 8 5	0	0
2	B	1	Total C O 13 8 5	0	0
2	C	1	Total C O 13 8 5	0	0

- Molecule 3 is (1-chloro-7-methoxy-9H-beta-carbolin-9-yl)acetonitrile (three-letter code: 4H5) (formula:  $C_{14}H_{10}ClN_3O$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	A	1	Total	C	Cl	N	O	0	0
			19	14	1	3	1		
3	B	1	Total	C	Cl	N	O	0	0
			19	14	1	3	1		
3	C	1	Total	C	Cl	N	O	0	0
			19	14	1	3	1		
3	D	1	Total	C	Cl	N	O	0	0
			19	14	1	3	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	O S	0	0
			5	4 1		
4	B	1	Total	O S	0	0
			5	4 1		
4	B	1	Total	O S	0	0
			5	4 1		
4	C	1	Total	O S	0	0
			5	4 1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	108	Total	O	0	0
			108	108		
5	B	90	Total	O	0	0
			90	90		

*Continued on next page...*

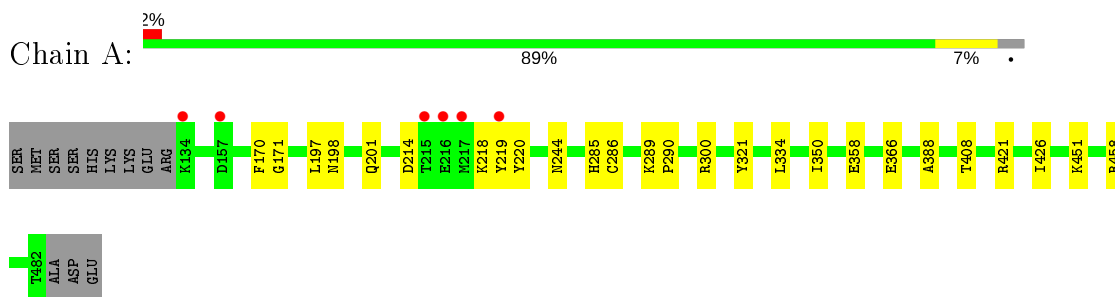
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
5	C	54	Total O 54 54	0	0
5	D	54	Total O 54 54	0	0

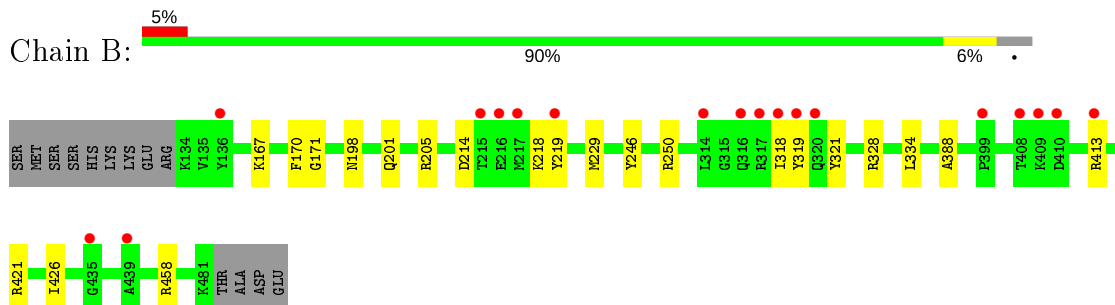
### 3 Residue-property plots [i](#)

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.

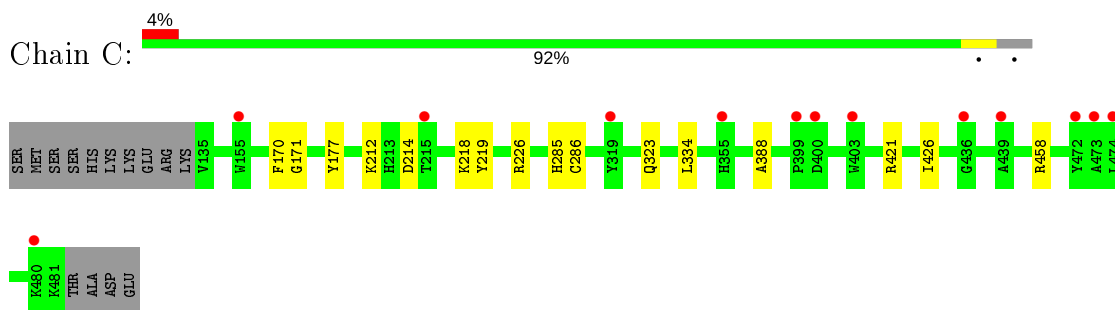
- Molecule 1: Dual specificity tyrosine-phosphorylation-regulated kinase 1A



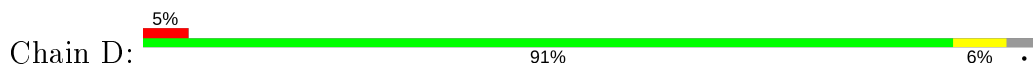
- Molecule 1: Dual specificity tyrosine-phosphorylation-regulated kinase 1A

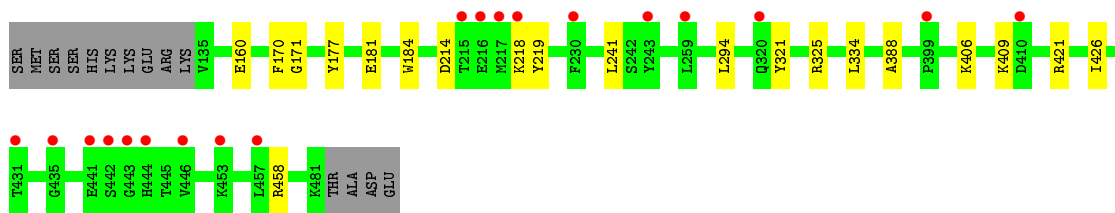


- Molecule 1: Dual specificity tyrosine-phosphorylation-regulated kinase 1A



- Molecule 1: Dual specificity tyrosine-phosphorylation-regulated kinase 1A







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	265.57Å 65.44Å 138.65Å 90.00° 114.44° 90.00°	Depositor
Resolution (Å)	48.34 – 2.90 47.41 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.34-2.90) 99.4 (47.41-2.90)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.91Å)	Xtrriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.243 , 0.277 0.248 , 0.281	Depositor DCC
$R_{free}$ test set	2443 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.7	Xtrriage
Anisotropy	0.545	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 63.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11762	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.45% 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, 4H5, PTR, SO4

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	0.60	0/2882	0.60	2/3889 (0.1%)
1	B	0.55	0/2857	0.57	2/3861 (0.1%)
1	C	0.51	0/2867	0.58	2/3869 (0.1%)
1	D	0.51	0/2865	0.56	2/3869 (0.1%)
All	All	0.54	0/11471	0.58	8/15488 (0.1%)

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	A	0	2
1	B	0	1
1	C	0	1
1	D	0	2
All	All	0	6

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	458	ARG	NE-CZ-NH2	6.54	123.57	120.30
1	C	458	ARG	NE-CZ-NH1	6.35	123.47	120.30
1	B	458	ARG	NE-CZ-NH2	-6.26	117.17	120.30
1	A	458	ARG	NE-CZ-NH1	-5.99	117.31	120.30
1	C	458	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	D	458	ARG	NE-CZ-NH2	5.61	123.11	120.30
1	B	458	ARG	NE-CZ-NH1	5.44	123.02	120.30
1	D	458	ARG	NE-CZ-NH1	-5.14	117.73	120.30

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	218	LYS	Peptide
1	A	408	THR	Peptide
1	B	218	LYS	Peptide
1	C	218	LYS	Peptide
1	D	218	LYS	Peptide
1	D	409	LYS	Peptide

## 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	2846	0	2815	22	0
1	B	2821	0	2759	20	0
1	C	2828	0	2791	12	0
1	D	2826	0	2776	19	0
2	A	13	0	18	1	0
2	B	13	0	18	3	0
2	C	13	0	18	0	0
3	A	19	0	10	2	0
3	B	19	0	10	4	0
3	C	19	0	10	1	0
3	D	19	0	10	3	0
4	B	15	0	0	0	0
4	C	5	0	0	0	0
5	A	108	0	0	6	0
5	B	90	0	0	3	0
5	C	54	0	0	2	0
5	D	54	0	0	1	0
All	All	11762	0	11235	70	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 3.

All (70) 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:250:ARG:NH1	1:D:160:GLU:OE2	2.09	0.85
1:B:167:LYS:CD	1:D:181:GLU:OE1	2.33	0.77
3:A:502:4H5:H3	3:A:502:4H5:CL	2.28	0.71
1:D:321[B]:PTR:HE2	1:D:321[B]:PTR:O1P	1.91	0.70
1:A:214:ASP:OD2	5:A:673:HOH:O	2.11	0.68
1:B:167:LYS:HD3	1:D:181:GLU:OE1	1.94	0.68
1:B:250:ARG:NH1	1:D:177:TYR:CD2	2.62	0.67
1:A:300:ARG:HD3	1:C:177:TYR:CZ	2.30	0.65
1:D:406:LYS:HE2	5:D:654:HOH:O	1.97	0.65
1:B:167:LYS:HD2	1:D:181:GLU:OE1	1.95	0.65
1:A:321[B]:PTR:O3P	1:A:366:GLU:HG3	1.98	0.63
1:A:300:ARG:HD3	1:C:177:TYR:CE1	2.35	0.62
1:B:321[B]:PTR:O3P	1:B:328:ARG:NH2	2.23	0.62
3:A:502:4H5:C10	3:A:502:4H5:CL	2.85	0.61
3:C:502:4H5:CL	3:C:502:4H5:H3	2.38	0.60
1:A:451:LYS:HE2	5:A:701:HOH:O	2.05	0.57
1:A:300:ARG:CD	1:C:177:TYR:CZ	2.88	0.56
1:B:229:MET:HE3	2:B:501:PG4:H41	1.89	0.53
3:B:502:4H5:CL	3:B:502:4H5:H3	2.45	0.53
1:B:318:ILE:HG23	1:B:319:TYR:CD2	2.43	0.53
1:A:321[B]:PTR:P	5:A:624:HOH:O	2.67	0.52
3:B:502:4H5:H5	3:B:502:4H5:C11	2.39	0.51
1:A:170:PHE:CD1	1:A:171:GLY:N	2.78	0.51
3:B:502:4H5:H4	5:B:618:HOH:O	2.11	0.50
1:A:300:ARG:CD	1:C:177:TYR:CE1	2.95	0.49
1:A:334:LEU:HB3	1:A:388:ALA:HB1	1.95	0.49
1:D:294:LEU:HD23	3:D:501:4H5:H8	1.95	0.49
1:B:334:LEU:HB3	1:B:388:ALA:HB1	1.95	0.48
1:D:334:LEU:HB3	1:D:388:ALA:HB1	1.95	0.48
1:C:334:LEU:HB3	1:C:388:ALA:HB1	1.96	0.48
1:D:170:PHE:CD1	1:D:171:GLY:N	2.82	0.47
1:B:201:GLN:OE1	2:B:501:PG4:H61	2.15	0.47
1:C:170:PHE:CD1	1:C:171:GLY:N	2.83	0.46
1:A:197:LEU:HD21	2:A:501:PG4:H52	1.98	0.46
3:D:501:4H5:CL	3:D:501:4H5:H3	2.53	0.46
1:A:198:ASN:HA	1:A:201:GLN:HE21	1.80	0.45
3:B:502:4H5:C13	3:B:502:4H5:C11	2.94	0.45
1:A:289:LYS:O	1:A:290:PRO:C	2.55	0.45
1:B:246:TYR:HH	1:D:184:TRP:HZ2	1.63	0.45
1:B:246:TYR:OH	1:D:184:TRP:HZ2	2.00	0.45
1:C:214:ASP:OD2	1:C:219:TYR:HB2	2.17	0.44
1:B:170:PHE:CD1	1:B:171:GLY:N	2.85	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:214:ASP:CG	1:A:219:TYR:HB2	2.38	0.44
1:D:214:ASP:OD2	1:D:219:TYR:HB2	2.18	0.44
1:A:214:ASP:OD2	1:A:219:TYR:HB2	2.17	0.43
1:B:214:ASP:OD2	1:B:219:TYR:HB2	2.18	0.42
1:D:241:LEU:O	3:D:501:4H5:H6	2.19	0.42
1:D:214:ASP:CG	1:D:219:TYR:HB2	2.39	0.42
1:A:219:TYR:O	1:A:220:TYR:HB2	2.19	0.42
1:D:321[B]:PTR:CE2	1:D:321[B]:PTR:O1P	2.66	0.42
1:B:198:ASN:HA	1:B:201:GLN:HE21	1.85	0.42
1:A:244:ASN:ND2	5:A:640:HOH:O	2.53	0.42
1:C:214:ASP:CG	1:C:219:TYR:HB2	2.40	0.42
1:A:285:HIS:O	1:A:286:CYS:HB2	2.20	0.42
1:D:321[A]:PTR:O2P	1:D:325:ARG:NH1	2.53	0.41
1:D:421:ARG:HG2	1:D:426:ILE:HD11	2.01	0.41
1:D:321[A]:PTR:O3P	1:D:321[A]:PTR:HE1	2.20	0.41
1:B:201:GLN:HG2	2:B:501:PG4:H12	2.02	0.41
1:B:421:ARG:HG2	1:B:426:ILE:HD11	2.01	0.41
1:C:212:LYS:HD3	5:C:612:HOH:O	2.20	0.41
1:A:321[A]:PTR:O3P	5:A:601:HOH:O	2.22	0.41
1:B:205:ARG:HG3	5:B:690:HOH:O	2.21	0.41
1:A:358:GLU:HG3	5:A:697:HOH:O	2.21	0.41
1:B:214:ASP:CG	1:B:219:TYR:HB2	2.41	0.41
1:C:226[A]:ARG:HD3	5:C:624:HOH:O	2.21	0.41
1:C:285:HIS:O	1:C:286:CYS:HB2	2.21	0.41
1:B:413:ARG:N	5:B:610:HOH:O	2.52	0.41
1:A:290:PRO:HD3	1:A:350:ILE:HG12	2.03	0.40
1:C:421:ARG:HG2	1:C:426:ILE:HD11	2.03	0.40
1:A:421:ARG:HG2	1:A:426:ILE:HD11	2.03	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	346/361 (96%)	320 (92%)	26 (8%)	0	100	100
1	B	345/361 (96%)	322 (93%)	23 (7%)	0	100	100
1	C	345/361 (96%)	320 (93%)	24 (7%)	1 (0%)	41	71
1	D	345/361 (96%)	319 (92%)	26 (8%)	0	100	100
All	All	1381/1444 (96%)	1281 (93%)	99 (7%)	1 (0%)	51	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	323	GLN

### 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	302/320 (94%)	302 (100%)	0	100	100
1	B	297/320 (93%)	297 (100%)	0	100	100
1	C	298/320 (93%)	298 (100%)	0	100	100
1	D	298/320 (93%)	298 (100%)	0	100	100
All	All	1195/1280 (93%)	1195 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some 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](#)

8 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
1	PTR	C	321[B]	-	15,16,17	1.10	1 (6%)	19,22,24	1.23	3 (15%)
1	PTR	B	321[B]	-	15,16,17	0.75	0	19,22,24	1.04	1 (5%)
1	PTR	B	321[A]	-	15,16,17	1.43	2 (13%)	19,22,24	1.20	2 (10%)
1	PTR	A	321[A]	-	15,16,17	0.86	0	19,22,24	1.12	2 (10%)
1	PTR	C	321[A]	-	15,16,17	0.87	0	19,22,24	1.15	1 (5%)
1	PTR	A	321[B]	-	15,16,17	0.83	0	19,22,24	0.88	0
1	PTR	D	321[A]	-	15,16,17	0.86	0	19,22,24	0.93	1 (5%)
1	PTR	D	321[B]	-	15,16,17	1.02	1 (6%)	19,22,24	1.13	1 (5%)

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
1	PTR	C	321[B]	-	-	3/10/11/13	0/1/1/1
1	PTR	B	321[B]	-	-	1/10/11/13	0/1/1/1
1	PTR	B	321[A]	-	-	3/10/11/13	0/1/1/1
1	PTR	A	321[A]	-	-	1/10/11/13	0/1/1/1
1	PTR	C	321[A]	-	-	1/10/11/13	0/1/1/1
1	PTR	A	321[B]	-	-	3/10/11/13	0/1/1/1
1	PTR	D	321[A]	-	-	1/10/11/13	0/1/1/1
1	PTR	D	321[B]	-	-	3/10/11/13	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	321[A]	PTR	CB-CA	3.29	1.60	1.53
1	B	321[A]	PTR	P-OH	2.97	1.63	1.59
1	C	321[B]	PTR	P-OH	2.62	1.63	1.59
1	D	321[B]	PTR	P-OH	2.24	1.62	1.59

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	321[B]	PTR	CG-CB-CA	2.97	120.10	114.10
1	C	321[A]	PTR	O3P-P-O2P	2.96	118.96	107.64
1	C	321[B]	PTR	CG-CB-CA	2.91	119.99	114.10
1	B	321[A]	PTR	P-OH-CZ	2.71	132.43	123.75
1	B	321[A]	PTR	CG-CB-CA	2.49	119.14	114.10
1	D	321[A]	PTR	O3P-P-O2P	2.42	116.90	107.64
1	C	321[B]	PTR	CB-CG-CD1	-2.19	116.56	120.91
1	C	321[B]	PTR	CB-CG-CD2	2.17	125.21	120.91
1	A	321[A]	PTR	CG-CB-CA	2.15	118.46	114.10
1	B	321[B]	PTR	O2P-P-O1P	2.09	118.86	110.68
1	A	321[A]	PTR	O2P-P-O1P	2.01	118.53	110.68

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	321[B]	PTR	O-C-CA-CB
1	C	321[B]	PTR	N-CA-CB-CG
1	C	321[B]	PTR	C-CA-CB-CG
1	B	321[B]	PTR	O-C-CA-CB
1	B	321[A]	PTR	O-C-CA-CB
1	B	321[A]	PTR	N-CA-CB-CG
1	B	321[A]	PTR	C-CA-CB-CG
1	A	321[A]	PTR	O-C-CA-CB
1	C	321[A]	PTR	O-C-CA-CB
1	A	321[B]	PTR	O-C-CA-CB
1	A	321[B]	PTR	CZ-OH-P-O1P
1	D	321[A]	PTR	O-C-CA-CB
1	D	321[B]	PTR	O-C-CA-CB
1	D	321[B]	PTR	N-CA-CB-CG
1	D	321[B]	PTR	C-CA-CB-CG
1	A	321[B]	PTR	CZ-OH-P-O3P

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	321[B]	PTR	1	0
1	A	321[A]	PTR	1	0
1	A	321[B]	PTR	2	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	321[A]	PTR	2	0
1	D	321[B]	PTR	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

11 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
2	PG4	B	501	-	12,12,12	0.74	0	11,11,11	0.69	0
4	SO4	C	503	-	4,4,4	0.46	0	6,6,6	0.39	0
3	4H5	A	502	-	18,21,21	2.24	5 (27%)	21,30,30	3.48	12 (57%)
4	SO4	B	504	-	4,4,4	0.41	0	6,6,6	0.21	0
2	PG4	A	501	-	12,12,12	0.63	0	11,11,11	0.69	0
2	PG4	C	501	-	12,12,12	0.62	0	11,11,11	0.29	0
4	SO4	B	503	-	4,4,4	0.37	0	6,6,6	0.28	0
3	4H5	C	502	-	18,21,21	1.66	4 (22%)	21,30,30	2.17	7 (33%)
3	4H5	D	501	-	18,21,21	1.60	5 (27%)	21,30,30	2.24	7 (33%)
4	SO4	B	505	-	4,4,4	0.52	0	6,6,6	0.25	0
3	4H5	B	502	-	18,21,21	1.60	5 (27%)	21,30,30	3.01	8 (38%)

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	PG4	B	501	-	-	8/10/10/10	-
3	4H5	A	502	-	-	0/2/5/5	0/3/3/3

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PG4	A	501	-	-	4/10/10/10	-
2	PG4	C	501	-	-	5/10/10/10	-
3	4H5	C	502	-	-	2/2/5/5	0/3/3/3
3	4H5	D	501	-	-	2/2/5/5	0/3/3/3
3	4H5	B	502	-	-	2/2/5/5	0/3/3/3

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	502	4H5	C8-C9	-5.30	1.36	1.43
3	A	502	4H5	C5-C9	-5.02	1.32	1.41
3	D	501	4H5	C5-C9	-3.64	1.34	1.41
3	C	502	4H5	C5-C9	-3.50	1.34	1.41
3	C	502	4H5	C8-C9	-3.45	1.39	1.43
3	A	502	4H5	C4-C12	-3.41	1.35	1.41
3	B	502	4H5	C6-C7	3.17	1.40	1.36
3	B	502	4H5	C5-C9	-3.16	1.35	1.41
3	A	502	4H5	C6-C7	3.08	1.40	1.36
3	D	501	4H5	C8-C9	-2.92	1.39	1.43
3	C	502	4H5	C6-C7	2.74	1.40	1.36
3	B	502	4H5	C8-CL	2.56	1.81	1.74
3	D	501	4H5	C6-C7	2.39	1.39	1.36
3	C	502	4H5	C4-C12	-2.32	1.37	1.41
3	B	502	4H5	C8-C9	-2.31	1.40	1.43
3	D	501	4H5	C4-C12	-2.27	1.37	1.41
3	A	502	4H5	C3-C4	-2.25	1.36	1.41
3	D	501	4H5	C13-C1	2.19	1.41	1.37
3	B	502	4H5	C4-C12	-2.12	1.37	1.41

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	502	4H5	C10-C11-N2	-11.99	170.00	178.46
3	B	502	4H5	C10-C11-N2	-9.27	171.91	178.46
3	B	502	4H5	C7-N-C8	5.68	121.41	116.81
3	D	501	4H5	C7-N-C8	5.35	121.14	116.81
3	C	502	4H5	C7-N-C8	5.22	121.03	116.81
3	A	502	4H5	C7-N-C8	4.38	120.35	116.81
3	A	502	4H5	C3-C4-C12	4.20	124.71	119.39
3	B	502	4H5	C9-C8-N	-3.95	119.00	124.96
3	A	502	4H5	C6-C7-N	-3.89	119.98	123.81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	501	4H5	C6-C7-N	-3.42	120.44	123.81
3	C	502	4H5	C9-C8-N	-3.42	119.80	124.96
3	D	501	4H5	C9-C8-N	-3.39	119.84	124.96
3	B	502	4H5	C6-C7-N	-3.30	120.56	123.81
3	D	501	4H5	C-O-C1	3.28	124.62	117.51
3	C	502	4H5	C3-C4-C12	3.26	123.51	119.39
3	D	501	4H5	C3-C4-C12	3.26	123.51	119.39
3	C	502	4H5	C6-C7-N	-3.07	120.78	123.81
3	B	502	4H5	C3-C4-C12	2.97	123.14	119.39
3	A	502	4H5	C3-C4-C5	-2.87	126.84	133.15
3	C	502	4H5	C7-C6-C5	-2.86	117.31	119.77
3	A	502	4H5	C9-C8-N	-2.79	120.75	124.96
3	B	502	4H5	C2-C3-C4	-2.70	117.74	121.63
3	C	502	4H5	C2-C3-C4	-2.67	117.78	121.63
3	A	502	4H5	C7-C6-C5	-2.65	117.48	119.77
3	A	502	4H5	C13-C12-C4	-2.61	116.95	120.92
3	A	502	4H5	C6-C5-C9	2.59	121.56	116.73
3	B	502	4H5	C6-C5-C9	2.53	121.46	116.73
3	D	501	4H5	C6-C5-C9	2.50	121.40	116.73
3	C	502	4H5	C6-C5-C9	2.46	121.32	116.73
3	D	501	4H5	C2-C3-C4	-2.41	118.16	121.63
3	B	502	4H5	C7-C6-C5	-2.25	117.83	119.77
3	A	502	4H5	C-O-C1	2.21	122.30	117.51
3	A	502	4H5	C2-C3-C4	-2.19	118.48	121.63
3	A	502	4H5	CL-C8-N	2.02	119.09	115.98

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	PG4	C8-C7-O4-C6
2	A	501	PG4	O2-C3-C4-O3
2	A	501	PG4	O3-C5-C6-O4
2	C	501	PG4	O2-C3-C4-O3
2	C	501	PG4	O1-C1-C2-O2
2	B	501	PG4	O2-C3-C4-O3
2	A	501	PG4	O4-C7-C8-O5
3	C	502	4H5	C13-C1-O-C
3	C	502	4H5	C2-C1-O-C
2	B	501	PG4	O1-C1-C2-O2
2	C	501	PG4	C8-C7-O4-C6
3	B	502	4H5	C13-C1-O-C

*Continued on next page...*

*Continued from previous page...*

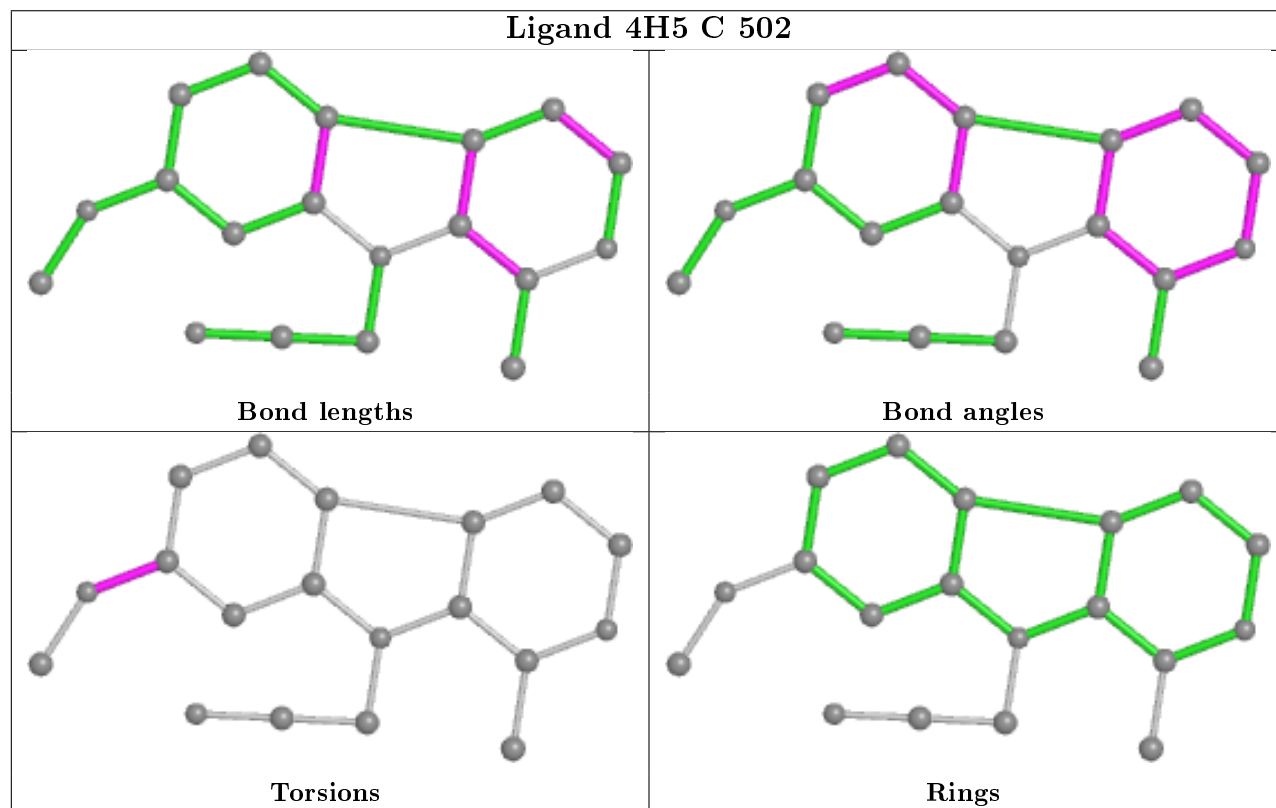
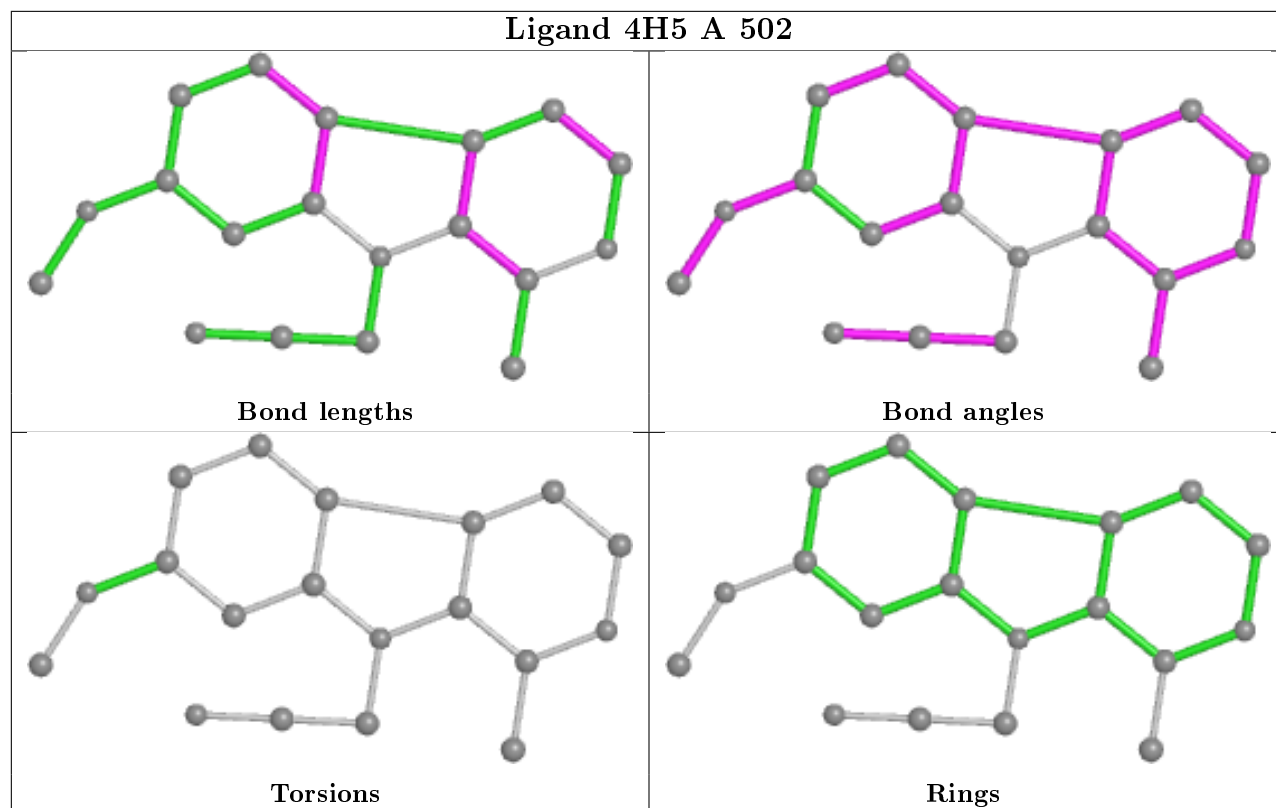
Mol	Chain	Res	Type	Atoms
2	B	501	PG4	C6-C5-O3-C4
3	B	502	4H5	C2-C1-O-C
2	A	501	PG4	C6-C5-O3-C4
2	C	501	PG4	C4-C3-O2-C2
2	B	501	PG4	O4-C7-C8-O5
2	B	501	PG4	C1-C2-O2-C3
3	D	501	4H5	C13-C1-O-C
3	D	501	4H5	C2-C1-O-C
2	C	501	PG4	C3-C4-O3-C5
2	B	501	PG4	C3-C4-O3-C5
2	B	501	PG4	O3-C5-C6-O4

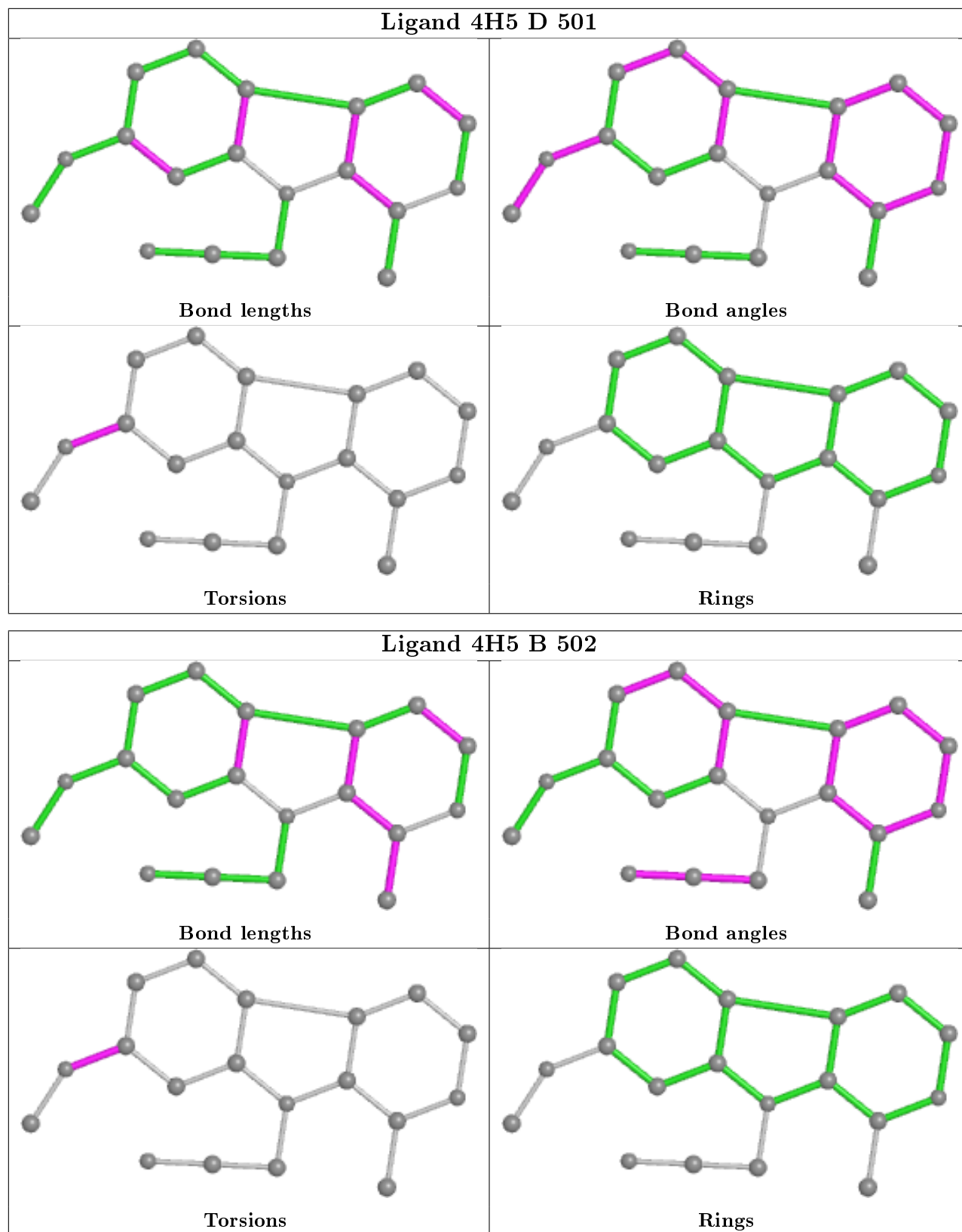
There are no ring outliers.

6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	PG4	3	0
3	A	502	4H5	2	0
2	A	501	PG4	1	0
3	C	502	4H5	1	0
3	D	501	4H5	3	0
3	B	502	4H5	4	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

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, 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	348/361 (96%)	0.03	6 (1%) 70 69	30, 54, 94, 174	0
1	B	347/361 (96%)	0.17	18 (5%) 27 23	44, 71, 115, 174	0
1	C	346/361 (95%)	0.23	13 (3%) 40 36	56, 80, 114, 137	0
1	D	346/361 (95%)	0.35	19 (5%) 25 21	58, 83, 122, 148	0
All	All	1387/1444 (96%)	0.20	56 (4%) 38 33	30, 73, 116, 174	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	217	MET	8.7
1	B	410	ASP	6.3
1	A	216	GLU	6.0
1	D	216	GLU	6.0
1	A	215	THR	5.0
1	D	435	GLY	4.5
1	C	399	PRO	3.9
1	D	215	THR	3.7
1	D	441	GLU	3.5
1	B	413	ARG	3.4
1	B	399	PRO	3.3
1	D	217	MET	3.2
1	A	219	TYR	3.2
1	B	216	GLU	3.1
1	D	399	PRO	3.1
1	B	217	MET	3.0
1	C	474	LEU	3.0
1	C	403	TRP	2.9
1	D	446	VAL	2.9
1	D	259	LEU	2.9
1	C	480	LYS	2.8

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	408	THR	2.8
1	B	316	GLN	2.8
1	B	317	ARG	2.7
1	C	319	TYR	2.7
1	C	472	TYR	2.7
1	B	320	GLN	2.7
1	D	410	ASP	2.7
1	B	136	TYR	2.7
1	D	243	TYR	2.7
1	D	453	LYS	2.6
1	B	409	LYS	2.5
1	B	439	ALA	2.4
1	B	319	TYR	2.4
1	A	134	LYS	2.4
1	B	215	THR	2.4
1	B	314	LEU	2.4
1	C	155	TRP	2.4
1	D	444	HIS	2.3
1	D	218	LYS	2.3
1	D	443	GLY	2.3
1	A	157	ASP	2.3
1	B	318	ILE	2.3
1	D	230	PHE	2.3
1	C	400	ASP	2.2
1	C	473	ALA	2.2
1	C	439	ALA	2.2
1	B	219	TYR	2.1
1	B	435	GLY	2.1
1	C	355	HIS	2.1
1	C	215	THR	2.1
1	D	442	SER	2.0
1	D	320	GLN	2.0
1	C	436	GLY	2.0
1	D	431	THR	2.0
1	D	457	LEU	2.0

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

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( $\text{\AA}^2$ )	Q<0.9
1	PTR	D	321[A]	16/17	0.78	0.31	67,81,87,89	13
1	PTR	D	321[B]	16/17	0.78	0.31	63,72,87,89	13
1	PTR	B	321[B]	16/17	0.82	0.30	70,82,89,95	13
1	PTR	B	321[A]	16/17	0.82	0.30	82,86,90,95	13
1	PTR	C	321[B]	16/17	0.83	0.32	64,68,74,80	13
1	PTR	C	321[A]	16/17	0.83	0.32	60,73,77,80	13
1	PTR	A	321[A]	16/17	0.88	0.23	52,70,73,77	13
1	PTR	A	321[B]	16/17	0.88	0.23	61,68,71,77	13

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

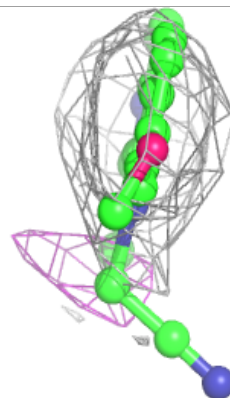
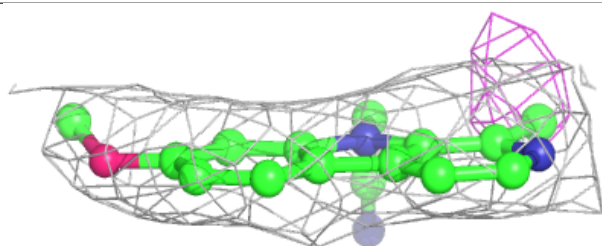
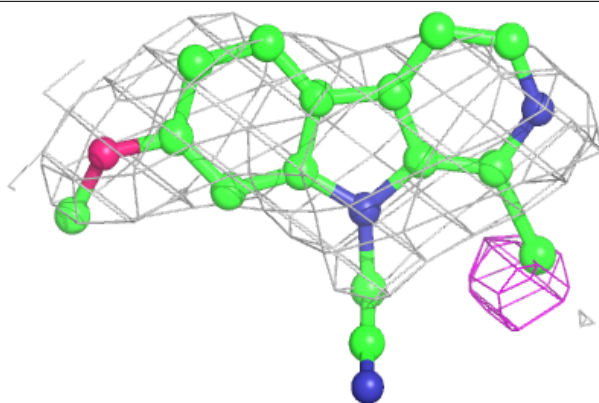
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	B	505	5/5	0.80	0.22	96,103,112,114	0
4	SO4	C	503	5/5	0.88	0.19	93,108,113,122	0
4	SO4	B	503	5/5	0.90	0.26	95,96,116,117	0
2	PG4	C	501	13/13	0.91	0.19	67,76,79,80	0
3	4H5	D	501	19/19	0.92	0.32	76,93,101,169	0
3	4H5	C	502	19/19	0.92	0.23	68,79,100,126	0
2	PG4	A	501	13/13	0.93	0.16	38,42,46,48	0
2	PG4	B	501	13/13	0.93	0.16	37,46,50,52	0
3	4H5	B	502	19/19	0.94	0.24	52,60,78,140	0
4	SO4	B	504	5/5	0.95	0.14	75,82,89,90	0
3	4H5	A	502	19/19	0.95	0.25	58,64,75,124	0

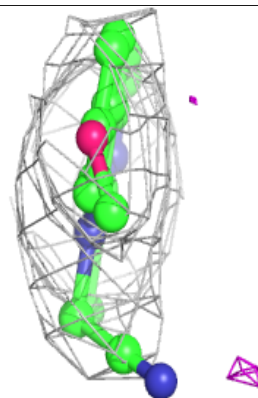
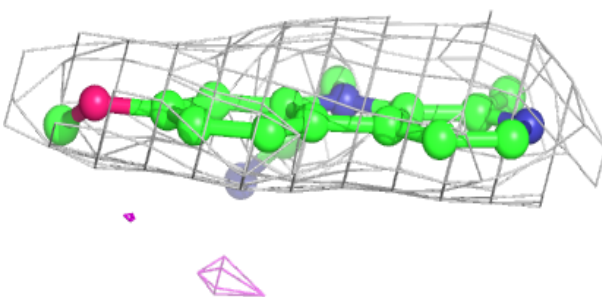
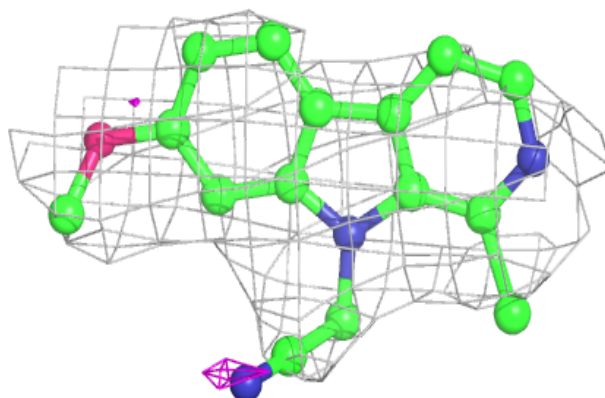
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around 4H5 D 501:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

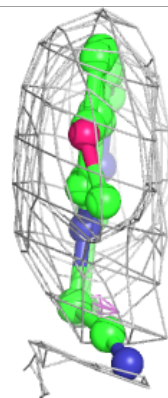
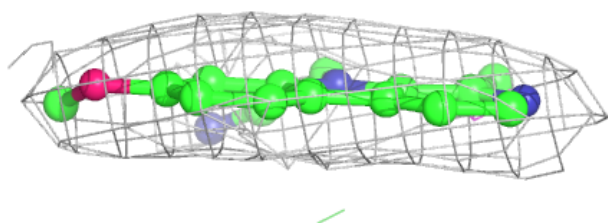
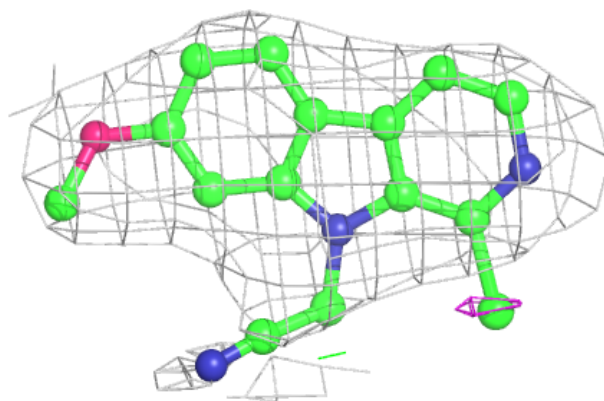
**Electron density around 4H5 C 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

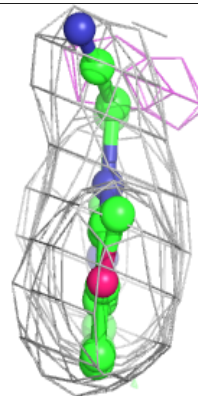
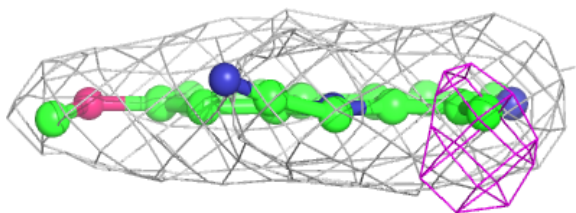
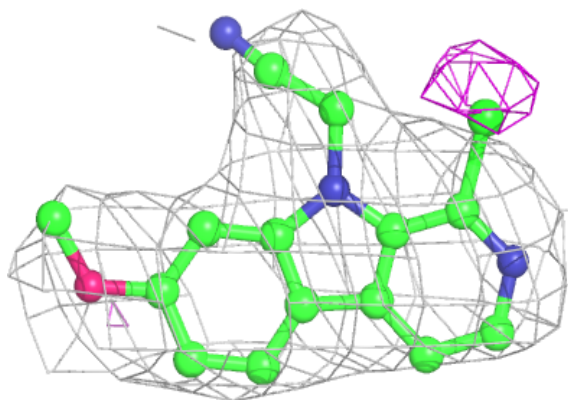


**Electron density around 4H5 B 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around 4H5 A 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers

There are no such residues in this entry.