wwPDB X-ray Structure Validation Summary Report (i)

Aug 21, 2020 – 04:10 PM BST

PDB ID	:	6VXE
Title	:	Crystal structure of hydroxyproline dehydratase (HypD) from Clostridioides difficile with substrate trans-4-hydroxy-L-proline bound
Authors	:	Backman, L.R.F.; Drennan, C.L.
Deposited on	:	2020-02-21
Resolution	:	2.46 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

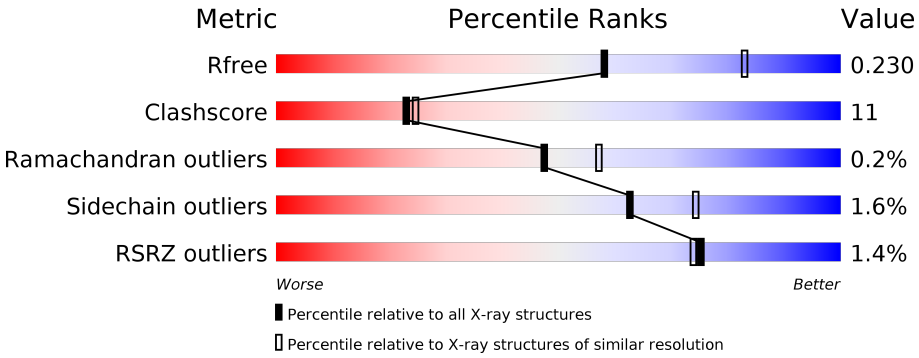
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.46 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	809	% 75% 22%
1	B	809	% 76% 20%
1	C	809	3% 73% 23%
1	D	809	% 75% 23%
1	E	809	76% 21%
1	F	809	% 77% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	809	78% 19%
1	H	809	3% 72% 25%

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	HYP	B	801	-	-	X	-
2	HYP	D	801	-	-	X	-
2	HYP	G	801	-	-	X	-
2	HYP	H	801	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 51664 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 Trans-4-hydroxy-L-proline dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	B	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	C	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	D	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	E	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	F	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	G	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			
1	H	789	Total	C	N	O	S	0	1	0
			6267	3953	1061	1211	42			

There are 160 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A0A031WDE4
A	-18	GLY	-	expression tag	UNP A0A031WDE4
A	-17	SER	-	expression tag	UNP A0A031WDE4
A	-16	SER	-	expression tag	UNP A0A031WDE4
A	-15	HIS	-	expression tag	UNP A0A031WDE4
A	-14	HIS	-	expression tag	UNP A0A031WDE4
A	-13	HIS	-	expression tag	UNP A0A031WDE4
A	-12	HIS	-	expression tag	UNP A0A031WDE4
A	-11	HIS	-	expression tag	UNP A0A031WDE4
A	-10	HIS	-	expression tag	UNP A0A031WDE4
A	-9	SER	-	expression tag	UNP A0A031WDE4
A	-8	SER	-	expression tag	UNP A0A031WDE4
A	-7	GLY	-	expression tag	UNP A0A031WDE4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	LEU	-	expression tag	UNP A0A031WDE4
A	-5	VAL	-	expression tag	UNP A0A031WDE4
A	-4	PRO	-	expression tag	UNP A0A031WDE4
A	-3	ARG	-	expression tag	UNP A0A031WDE4
A	-2	GLY	-	expression tag	UNP A0A031WDE4
A	-1	SER	-	expression tag	UNP A0A031WDE4
A	0	HIS	-	expression tag	UNP A0A031WDE4
B	-19	MET	-	initiating methionine	UNP A0A031WDE4
B	-18	GLY	-	expression tag	UNP A0A031WDE4
B	-17	SER	-	expression tag	UNP A0A031WDE4
B	-16	SER	-	expression tag	UNP A0A031WDE4
B	-15	HIS	-	expression tag	UNP A0A031WDE4
B	-14	HIS	-	expression tag	UNP A0A031WDE4
B	-13	HIS	-	expression tag	UNP A0A031WDE4
B	-12	HIS	-	expression tag	UNP A0A031WDE4
B	-11	HIS	-	expression tag	UNP A0A031WDE4
B	-10	HIS	-	expression tag	UNP A0A031WDE4
B	-9	SER	-	expression tag	UNP A0A031WDE4
B	-8	SER	-	expression tag	UNP A0A031WDE4
B	-7	GLY	-	expression tag	UNP A0A031WDE4
B	-6	LEU	-	expression tag	UNP A0A031WDE4
B	-5	VAL	-	expression tag	UNP A0A031WDE4
B	-4	PRO	-	expression tag	UNP A0A031WDE4
B	-3	ARG	-	expression tag	UNP A0A031WDE4
B	-2	GLY	-	expression tag	UNP A0A031WDE4
B	-1	SER	-	expression tag	UNP A0A031WDE4
B	0	HIS	-	expression tag	UNP A0A031WDE4
C	-19	MET	-	initiating methionine	UNP A0A031WDE4
C	-18	GLY	-	expression tag	UNP A0A031WDE4
C	-17	SER	-	expression tag	UNP A0A031WDE4
C	-16	SER	-	expression tag	UNP A0A031WDE4
C	-15	HIS	-	expression tag	UNP A0A031WDE4
C	-14	HIS	-	expression tag	UNP A0A031WDE4
C	-13	HIS	-	expression tag	UNP A0A031WDE4
C	-12	HIS	-	expression tag	UNP A0A031WDE4
C	-11	HIS	-	expression tag	UNP A0A031WDE4
C	-10	HIS	-	expression tag	UNP A0A031WDE4
C	-9	SER	-	expression tag	UNP A0A031WDE4
C	-8	SER	-	expression tag	UNP A0A031WDE4
C	-7	GLY	-	expression tag	UNP A0A031WDE4
C	-6	LEU	-	expression tag	UNP A0A031WDE4
C	-5	VAL	-	expression tag	UNP A0A031WDE4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-4	PRO	-	expression tag	UNP A0A031WDE4
C	-3	ARG	-	expression tag	UNP A0A031WDE4
C	-2	GLY	-	expression tag	UNP A0A031WDE4
C	-1	SER	-	expression tag	UNP A0A031WDE4
C	0	HIS	-	expression tag	UNP A0A031WDE4
D	-19	MET	-	initiating methionine	UNP A0A031WDE4
D	-18	GLY	-	expression tag	UNP A0A031WDE4
D	-17	SER	-	expression tag	UNP A0A031WDE4
D	-16	SER	-	expression tag	UNP A0A031WDE4
D	-15	HIS	-	expression tag	UNP A0A031WDE4
D	-14	HIS	-	expression tag	UNP A0A031WDE4
D	-13	HIS	-	expression tag	UNP A0A031WDE4
D	-12	HIS	-	expression tag	UNP A0A031WDE4
D	-11	HIS	-	expression tag	UNP A0A031WDE4
D	-10	HIS	-	expression tag	UNP A0A031WDE4
D	-9	SER	-	expression tag	UNP A0A031WDE4
D	-8	SER	-	expression tag	UNP A0A031WDE4
D	-7	GLY	-	expression tag	UNP A0A031WDE4
D	-6	LEU	-	expression tag	UNP A0A031WDE4
D	-5	VAL	-	expression tag	UNP A0A031WDE4
D	-4	PRO	-	expression tag	UNP A0A031WDE4
D	-3	ARG	-	expression tag	UNP A0A031WDE4
D	-2	GLY	-	expression tag	UNP A0A031WDE4
D	-1	SER	-	expression tag	UNP A0A031WDE4
D	0	HIS	-	expression tag	UNP A0A031WDE4
E	-19	MET	-	initiating methionine	UNP A0A031WDE4
E	-18	GLY	-	expression tag	UNP A0A031WDE4
E	-17	SER	-	expression tag	UNP A0A031WDE4
E	-16	SER	-	expression tag	UNP A0A031WDE4
E	-15	HIS	-	expression tag	UNP A0A031WDE4
E	-14	HIS	-	expression tag	UNP A0A031WDE4
E	-13	HIS	-	expression tag	UNP A0A031WDE4
E	-12	HIS	-	expression tag	UNP A0A031WDE4
E	-11	HIS	-	expression tag	UNP A0A031WDE4
E	-10	HIS	-	expression tag	UNP A0A031WDE4
E	-9	SER	-	expression tag	UNP A0A031WDE4
E	-8	SER	-	expression tag	UNP A0A031WDE4
E	-7	GLY	-	expression tag	UNP A0A031WDE4
E	-6	LEU	-	expression tag	UNP A0A031WDE4
E	-5	VAL	-	expression tag	UNP A0A031WDE4
E	-4	PRO	-	expression tag	UNP A0A031WDE4
E	-3	ARG	-	expression tag	UNP A0A031WDE4

Continued on next page...

Continued from previous page...

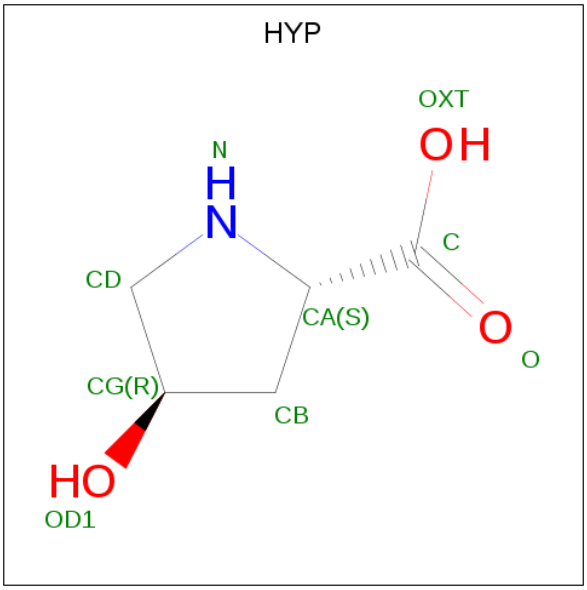
Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	GLY	-	expression tag	UNP A0A031WDE4
E	-1	SER	-	expression tag	UNP A0A031WDE4
E	0	HIS	-	expression tag	UNP A0A031WDE4
F	-19	MET	-	initiating methionine	UNP A0A031WDE4
F	-18	GLY	-	expression tag	UNP A0A031WDE4
F	-17	SER	-	expression tag	UNP A0A031WDE4
F	-16	SER	-	expression tag	UNP A0A031WDE4
F	-15	HIS	-	expression tag	UNP A0A031WDE4
F	-14	HIS	-	expression tag	UNP A0A031WDE4
F	-13	HIS	-	expression tag	UNP A0A031WDE4
F	-12	HIS	-	expression tag	UNP A0A031WDE4
F	-11	HIS	-	expression tag	UNP A0A031WDE4
F	-10	HIS	-	expression tag	UNP A0A031WDE4
F	-9	SER	-	expression tag	UNP A0A031WDE4
F	-8	SER	-	expression tag	UNP A0A031WDE4
F	-7	GLY	-	expression tag	UNP A0A031WDE4
F	-6	LEU	-	expression tag	UNP A0A031WDE4
F	-5	VAL	-	expression tag	UNP A0A031WDE4
F	-4	PRO	-	expression tag	UNP A0A031WDE4
F	-3	ARG	-	expression tag	UNP A0A031WDE4
F	-2	GLY	-	expression tag	UNP A0A031WDE4
F	-1	SER	-	expression tag	UNP A0A031WDE4
F	0	HIS	-	expression tag	UNP A0A031WDE4
G	-19	MET	-	initiating methionine	UNP A0A031WDE4
G	-18	GLY	-	expression tag	UNP A0A031WDE4
G	-17	SER	-	expression tag	UNP A0A031WDE4
G	-16	SER	-	expression tag	UNP A0A031WDE4
G	-15	HIS	-	expression tag	UNP A0A031WDE4
G	-14	HIS	-	expression tag	UNP A0A031WDE4
G	-13	HIS	-	expression tag	UNP A0A031WDE4
G	-12	HIS	-	expression tag	UNP A0A031WDE4
G	-11	HIS	-	expression tag	UNP A0A031WDE4
G	-10	HIS	-	expression tag	UNP A0A031WDE4
G	-9	SER	-	expression tag	UNP A0A031WDE4
G	-8	SER	-	expression tag	UNP A0A031WDE4
G	-7	GLY	-	expression tag	UNP A0A031WDE4
G	-6	LEU	-	expression tag	UNP A0A031WDE4
G	-5	VAL	-	expression tag	UNP A0A031WDE4
G	-4	PRO	-	expression tag	UNP A0A031WDE4
G	-3	ARG	-	expression tag	UNP A0A031WDE4
G	-2	GLY	-	expression tag	UNP A0A031WDE4
G	-1	SER	-	expression tag	UNP A0A031WDE4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	0	HIS	-	expression tag	UNP A0A031WDE4
H	-19	MET	-	initiating methionine	UNP A0A031WDE4
H	-18	GLY	-	expression tag	UNP A0A031WDE4
H	-17	SER	-	expression tag	UNP A0A031WDE4
H	-16	SER	-	expression tag	UNP A0A031WDE4
H	-15	HIS	-	expression tag	UNP A0A031WDE4
H	-14	HIS	-	expression tag	UNP A0A031WDE4
H	-13	HIS	-	expression tag	UNP A0A031WDE4
H	-12	HIS	-	expression tag	UNP A0A031WDE4
H	-11	HIS	-	expression tag	UNP A0A031WDE4
H	-10	HIS	-	expression tag	UNP A0A031WDE4
H	-9	SER	-	expression tag	UNP A0A031WDE4
H	-8	SER	-	expression tag	UNP A0A031WDE4
H	-7	GLY	-	expression tag	UNP A0A031WDE4
H	-6	LEU	-	expression tag	UNP A0A031WDE4
H	-5	VAL	-	expression tag	UNP A0A031WDE4
H	-4	PRO	-	expression tag	UNP A0A031WDE4
H	-3	ARG	-	expression tag	UNP A0A031WDE4
H	-2	GLY	-	expression tag	UNP A0A031WDE4
H	-1	SER	-	expression tag	UNP A0A031WDE4
H	0	HIS	-	expression tag	UNP A0A031WDE4

- Molecule 2 is 4-HYDROXYPROLINE (three-letter code: HYP) (formula: $C_5H_9NO_3$) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			9	5	1	3		
2	B	1	Total	C	N	O	0	0
			9	5	1	3		
2	C	1	Total	C	N	O	0	0
			9	5	1	3		
2	D	1	Total	C	N	O	0	0
			9	5	1	3		
2	E	1	Total	C	N	O	0	0
			9	5	1	3		
2	F	1	Total	C	N	O	0	0
			9	5	1	3		
2	G	1	Total	C	N	O	0	0
			9	5	1	3		
2	H	1	Total	C	N	O	0	0
			9	5	1	3		

- Molecule 3 is water.

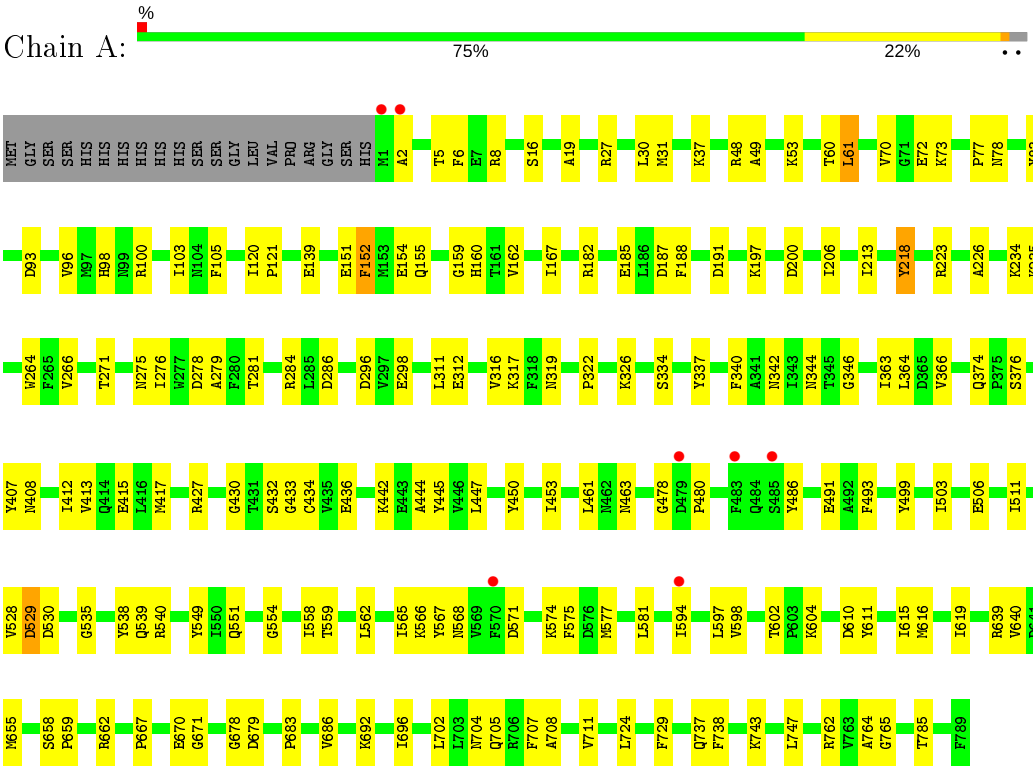
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	192	Total	O	0	0
			192	192		
3	B	183	Total	O	0	0
			183	183		
3	C	132	Total	O	0	0
			132	132		
3	D	205	Total	O	0	0
			205	205		
3	E	171	Total	O	0	0
			171	171		
3	F	216	Total	O	0	0
			216	216		
3	G	221	Total	O	0	0
			221	221		
3	H	136	Total	O	0	0
			136	136		

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.

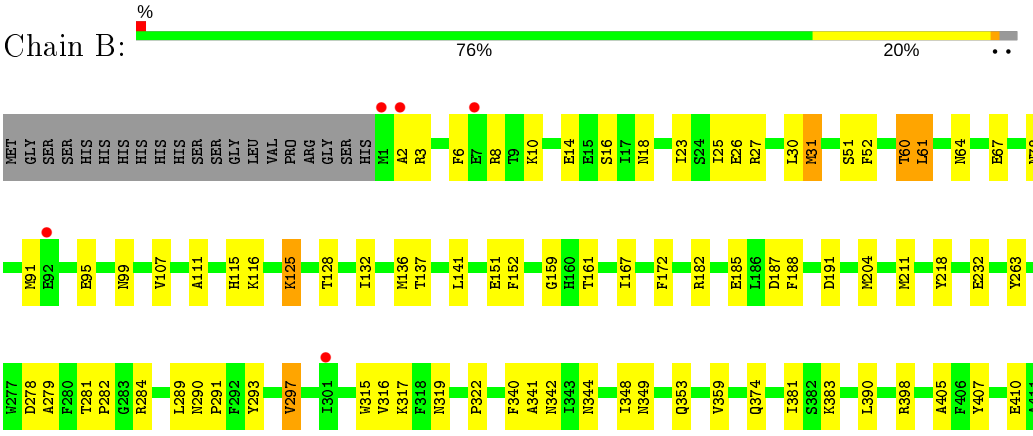
- Molecule 1: Trans-4-hydroxy-L-proline dehydratase

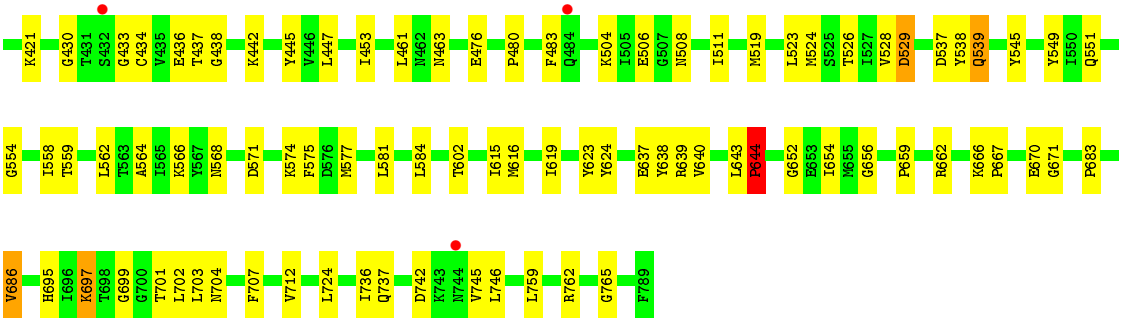
Chain A:



- Molecule 1: Trans-4-hydroxy-L-proline dehydratase

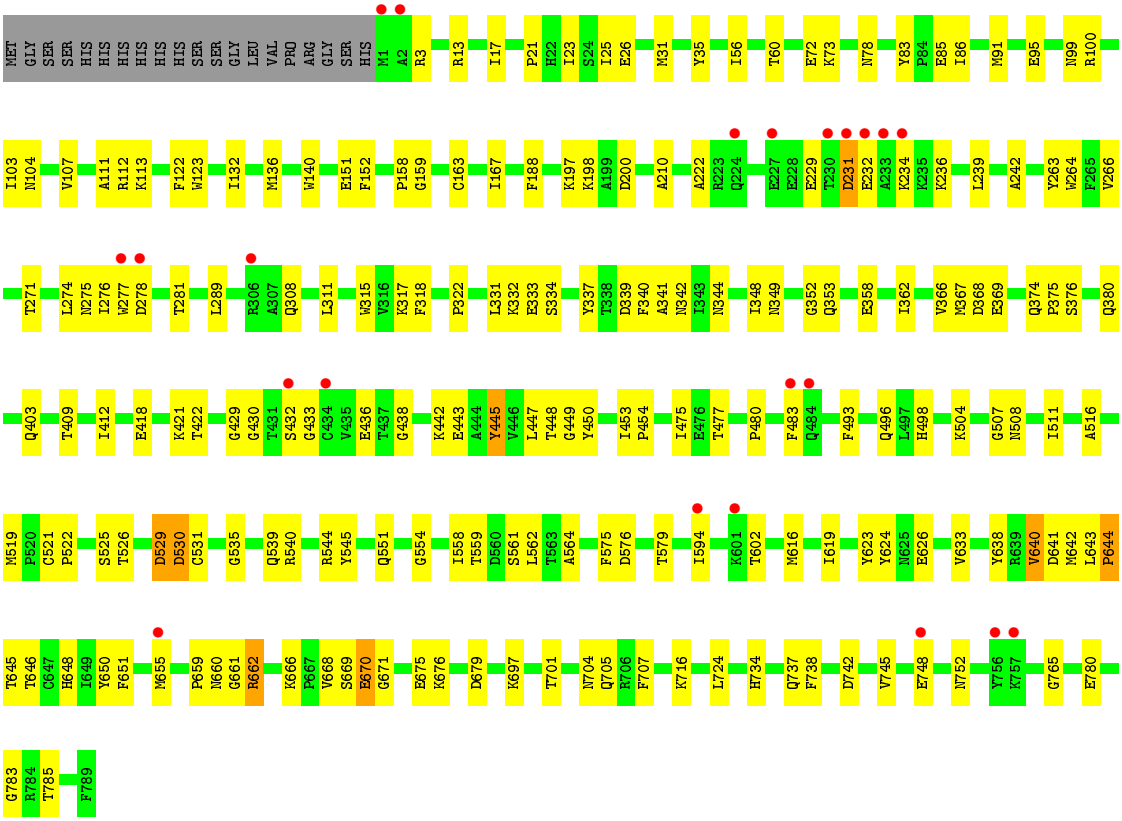
Chain B:





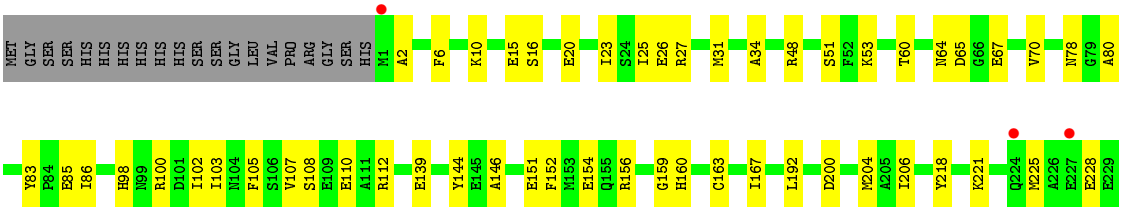
• Molecule 1: Trans-4-hydroxy-L-proline dehydratase

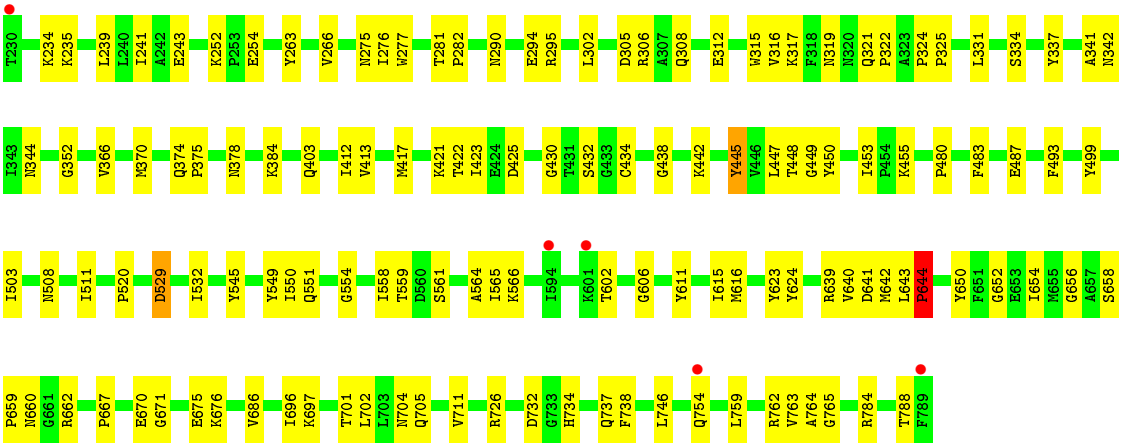
Chain C:



• Molecule 1: Trans-4-hydroxy-L-proline dehydratase

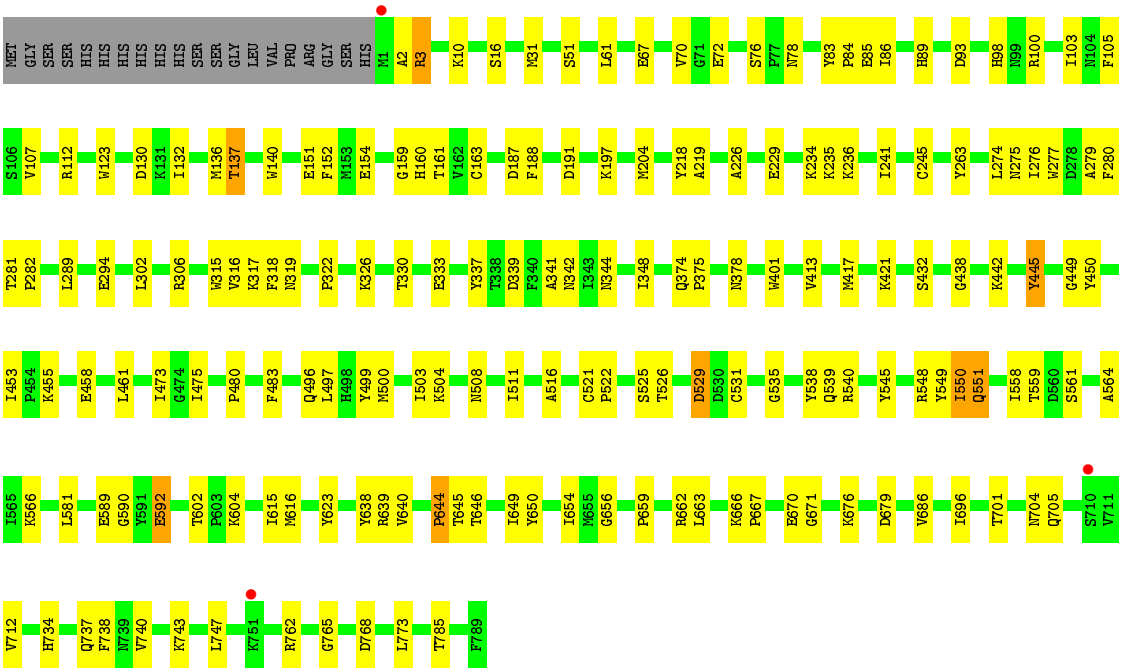
Chain D:





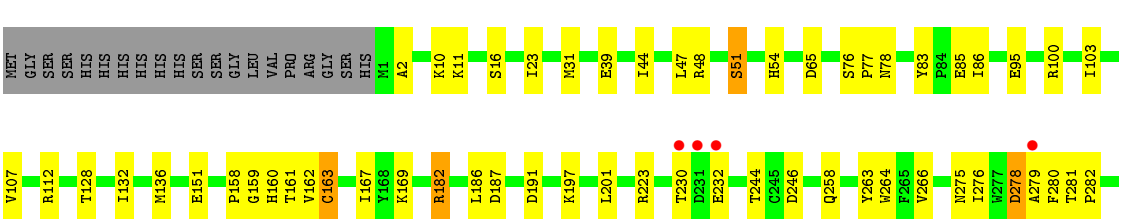
• Molecule 1: Trans-4-hydroxy-L-proline dehydratase

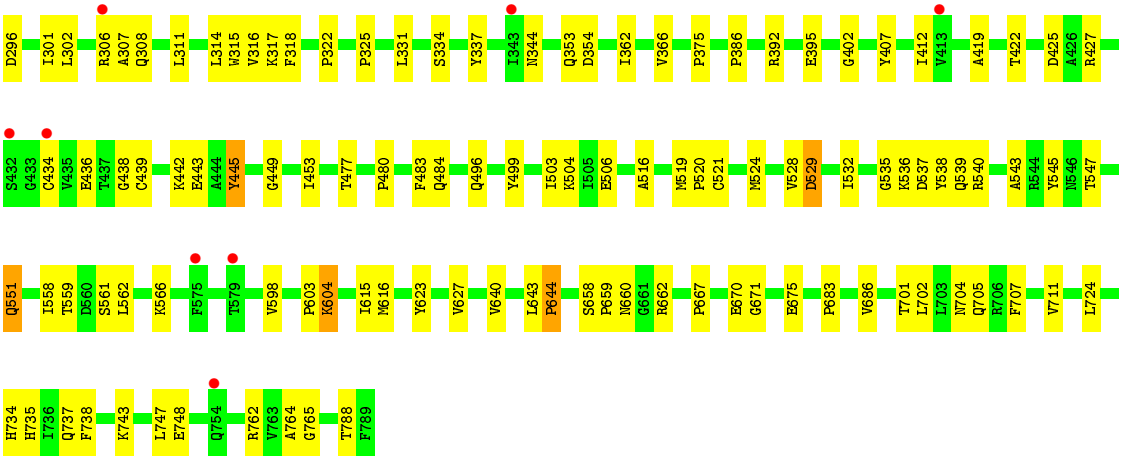
Chain E:



• Molecule 1: Trans-4-hydroxy-L-proline dehydratase

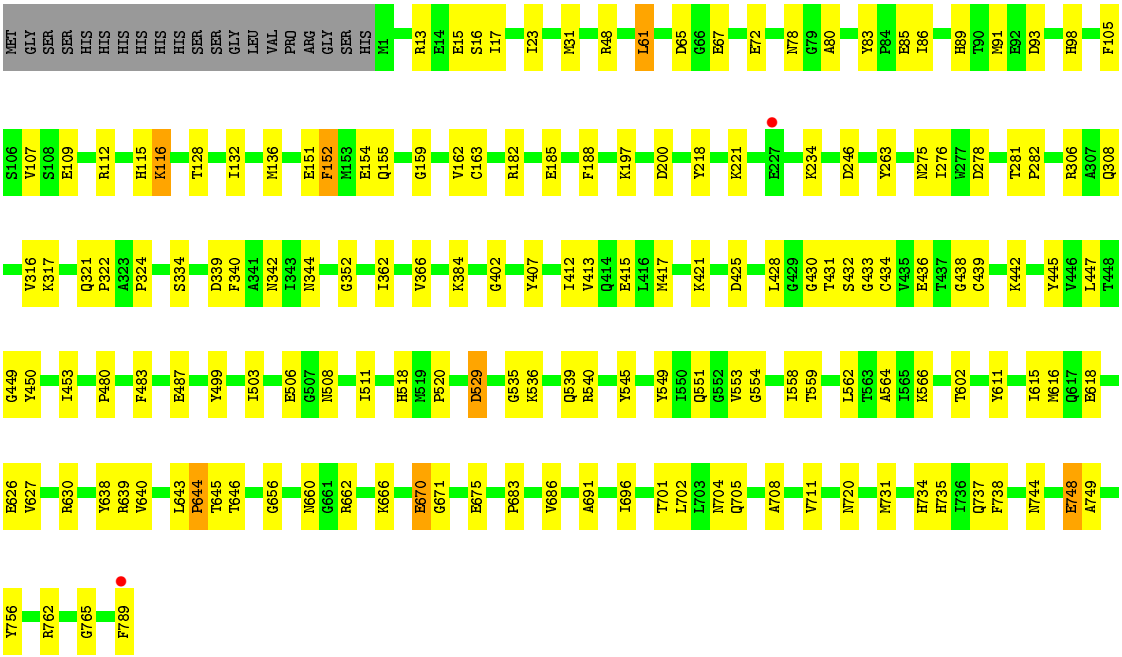
Chain F:





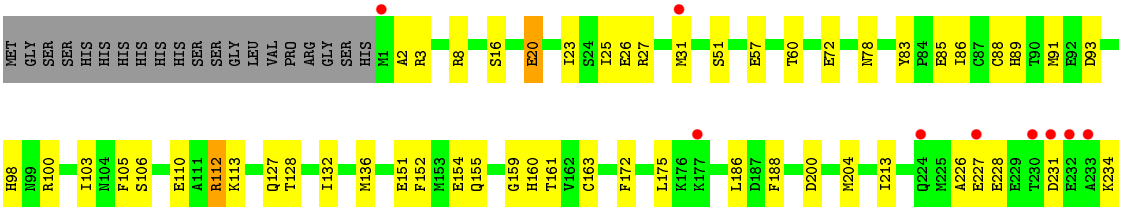
• Molecule 1: Trans-4-hydroxy-L-proline dehydratase

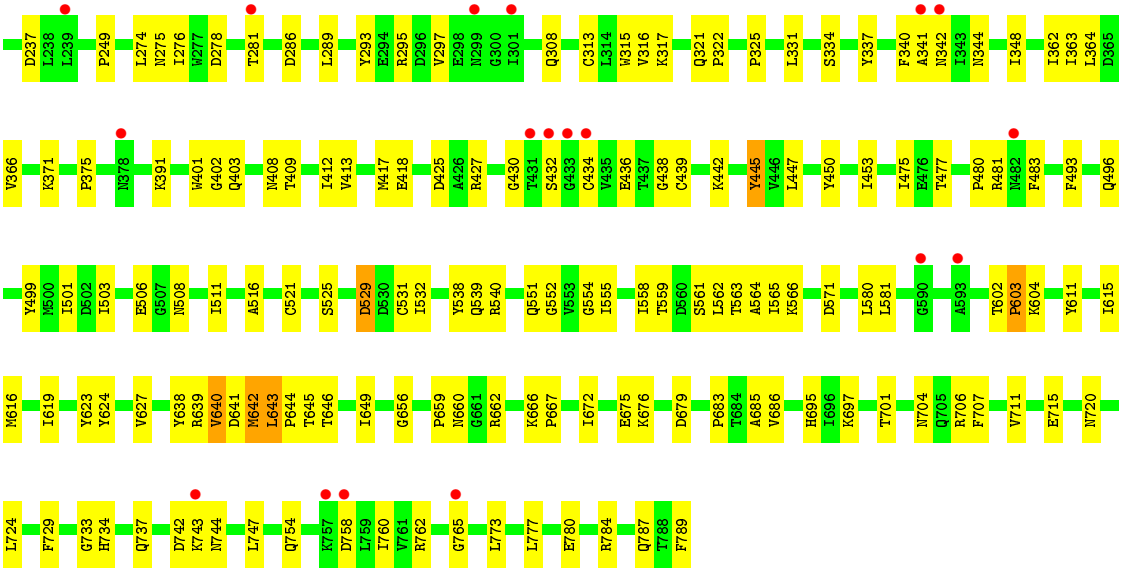
Chain G:



• Molecule 1: Trans-4-hydroxy-L-proline dehydratase

Chain H:




D237 L238 L239 P249 L274 N275 I276 W277 D278 T281 D286 L289 Y293 E294 R295 D296 V297 E298 N299 G300 I301 Q308 C313 L314 W315 V316 K317 Q321 P322 P325 L331 S334 Y337 F340 A341 N342 I343 N344 I348 I362 I363 L364 D365
V366 K371 P375 N378 K391 W401 G402 Q403 N408 T409 I412 V413 M417 E418 D425 A426 R427 G430 T431 S432 G433 C434 V435 E436 T437 G438 C439 K442 Y445 V446 L447 Y450 I453 I475 E476 T477 P480 R481 N482 F483 F493 Q496
Y499 M500 I501 D502 I503 E506 G507 N508 I511 A516 C521 S525 D529 D530 C531 I532 Y538 Q539 R540 Q551 G552 V553 G554 I555 I558 T559 D560 S561 L562 T563 A564 I565 K566 D571 L580 L581 G590 A593 T602 P603 K604 Y611 I615
M616 I619 Y623 Y624 V627 Y638 R639 V640 D641 M642 L643 P644 T645 T646 I649 G656 P659 N660 G661 R662 K666 P667 I672 E675 K676 D679 P683 T684 A685 V686 H695 I696 K697 T701 N704 Q705 R706 F707 V711 E715 N720
L724 F729 G733 H734 Q737 D742 K743 N744 L747 Q754 K757 D758 L759 I760 V761 R762 G765 L773 L777 E780 R784 Q787 T788 F789

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	100.35Å 341.65Å 122.61Å 90.00° 107.14° 90.00°	Depositor
Resolution (Å)	49.30 – 2.46 49.30 – 2.46	Depositor EDS
% Data completeness (in resolution range)	96.2 (49.30-2.46) 96.2 (49.30-2.46)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.20	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 2.45Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.195 , 0.230 0.194 , 0.230	Depositor DCC
R_{free} test set	13595 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	36.5	Xtriage
Anisotropy	0.314	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 23.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	51664	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.08% of the height of the origin peak. No significant pseudotranslation is detected.*

[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: HYP

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	0.61	3/6391 (0.0%)	0.67	7/8629 (0.1%)
1	B	0.63	5/6391 (0.1%)	0.70	8/8629 (0.1%)
1	C	0.62	3/6391 (0.0%)	0.70	8/8629 (0.1%)
1	D	0.52	0/6391	0.63	3/8629 (0.0%)
1	E	0.57	1/6391 (0.0%)	0.72	8/8629 (0.1%)
1	F	0.62	2/6391 (0.0%)	0.68	6/8629 (0.1%)
1	G	0.54	0/6391	0.63	2/8629 (0.0%)
1	H	0.55	3/6391 (0.0%)	0.74	18/8629 (0.2%)
All	All	0.58	17/51128 (0.0%)	0.68	60/69032 (0.1%)

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	662	ARG	C-N	-19.84	0.88	1.34
1	B	60	THR	C-N	-14.79	1.00	1.34
1	A	60	THR	C-N	-13.35	1.03	1.34
1	H	2	ALA	C-N	-11.62	1.07	1.34
1	F	604	LYS	C-N	-10.47	1.09	1.34

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	3	ARG	O-C-N	17.73	153.34	123.20
1	C	643	LEU	O-C-N	-17.28	88.26	121.10
1	E	3	ARG	CA-C-N	-15.30	85.61	116.20
1	H	643	LEU	O-C-N	-15.05	92.51	121.10
1	B	61	LEU	O-C-N	-13.90	100.46	122.70

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	6267	0	6163	153	0
1	B	6267	0	6164	146	0
1	C	6267	0	6165	181	0
1	D	6267	0	6165	153	0
1	E	6267	0	6165	129	1
1	F	6267	0	6164	110	0
1	G	6267	0	6167	111	0
1	H	6267	0	6163	171	1
2	A	9	0	8	2	0
2	B	9	0	8	8	0
2	C	9	0	8	2	0
2	D	9	0	8	4	0
2	E	9	0	8	3	0
2	F	9	0	8	3	0
2	G	9	0	8	5	0
2	H	9	0	8	5	0
3	A	192	0	0	10	0
3	B	183	0	0	12	0
3	C	132	0	0	15	0
3	D	205	0	0	18	0
3	E	171	0	0	12	0
3	F	216	0	0	12	0
3	G	221	0	0	9	0
3	H	136	0	0	17	0
All	All	51664	0	49380	1138	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 1138 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:D:31[B]:MET:CE	1:D:85:GLU:CD	1.78	1.48
1:D:31[B]:MET:CE	1:D:85:GLU:OE1	1.65	1.41
1:A:31[B]:MET:HE1	1:A:85:GLU:CD	1.51	1.30

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:31[B]:MET:HE1	1:D:85:GLU:CD	1.34	1.27
1:A:31[B]:MET:CE	1:A:85:GLU:OE2	1.86	1.23

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:236:LYS:NZ	1:H:57:GLU:OE2[1_655]	2.19	0.01

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	788/809 (97%)	770 (98%)	17 (2%)	1 (0%)	51	64
1	B	788/809 (97%)	771 (98%)	15 (2%)	2 (0%)	41	49
1	C	788/809 (97%)	763 (97%)	22 (3%)	3 (0%)	34	41
1	D	788/809 (97%)	772 (98%)	15 (2%)	1 (0%)	51	64
1	E	788/809 (97%)	770 (98%)	17 (2%)	1 (0%)	51	64
1	F	788/809 (97%)	771 (98%)	15 (2%)	2 (0%)	41	49
1	G	788/809 (97%)	771 (98%)	16 (2%)	1 (0%)	51	64
1	H	788/809 (97%)	766 (97%)	20 (2%)	2 (0%)	41	49
All	All	6304/6472 (97%)	6154 (98%)	137 (2%)	13 (0%)	47	57

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	644	PRO
1	E	644	PRO
1	F	644	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	644	PRO
1	H	644	PRO

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	678/694 (98%)	669 (99%)	9 (1%)	69	79
1	B	678/694 (98%)	666 (98%)	12 (2%)	59	71
1	C	678/694 (98%)	669 (99%)	9 (1%)	69	79
1	D	678/694 (98%)	672 (99%)	6 (1%)	78	86
1	E	678/694 (98%)	665 (98%)	13 (2%)	57	69
1	F	678/694 (98%)	664 (98%)	14 (2%)	53	66
1	G	678/694 (98%)	667 (98%)	11 (2%)	62	74
1	H	678/694 (98%)	668 (98%)	10 (2%)	65	76
All	All	5424/5552 (98%)	5340 (98%)	84 (2%)	62	76

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	152	PHE
1	E	670	GLU
1	H	113	LYS
1	E	344	ASN
1	E	550	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	551	GLN
1	D	704	ASN
1	F	704	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	704	ASN
1	D	539	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry

8 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	HYP	B	801	-	6,9,9	0.59	0	5,12,12	3.06	3 (60%)
2	HYP	A	801	-	6,9,9	1.79	2 (33%)	5,12,12	2.47	1 (20%)
2	HYP	G	801	-	6,9,9	2.10	3 (50%)	5,12,12	2.10	2 (40%)
2	HYP	F	801	-	6,9,9	1.23	0	5,12,12	1.79	2 (40%)
2	HYP	E	801	-	6,9,9	1.11	0	5,12,12	1.80	1 (20%)
2	HYP	D	801	-	6,9,9	2.17	4 (66%)	5,12,12	2.30	3 (60%)
2	HYP	H	801	-	6,9,9	1.79	2 (33%)	5,12,12	1.80	2 (40%)
2	HYP	C	801	-	6,9,9	1.67	1 (16%)	5,12,12	2.59	4 (80%)

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	B	801	-	-	0/0/13/13	0/1/1/1
2	HYP	A	801	-	-	0/0/13/13	0/1/1/1
2	HYP	G	801	-	-	0/0/13/13	0/1/1/1
2	HYP	F	801	-	-	0/0/13/13	0/1/1/1
2	HYP	E	801	-	-	0/0/13/13	0/1/1/1
2	HYP	D	801	-	-	0/0/13/13	0/1/1/1
2	HYP	H	801	-	-	0/0/13/13	0/1/1/1
2	HYP	C	801	-	-	0/0/13/13	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	801	HYP	CB-CG	-2.66	1.47	1.52
2	G	801	HYP	CD-CG	-2.59	1.47	1.53
2	A	801	HYP	CD-CG	-2.59	1.47	1.53
2	G	801	HYP	CA-N	-2.57	1.38	1.50
2	D	801	HYP	CD-CG	-2.56	1.47	1.53

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	801	HYP	OD1-CG-CD	-4.86	99.72	110.35
2	B	801	HYP	CG-CB-CA	-4.44	98.36	103.96
2	C	801	HYP	OD1-CG-CD	-4.41	100.71	110.35
2	B	801	HYP	OD1-CG-CD	-4.41	100.71	110.35
2	G	801	HYP	OD1-CG-CD	-3.34	103.05	110.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 32 short contacts:

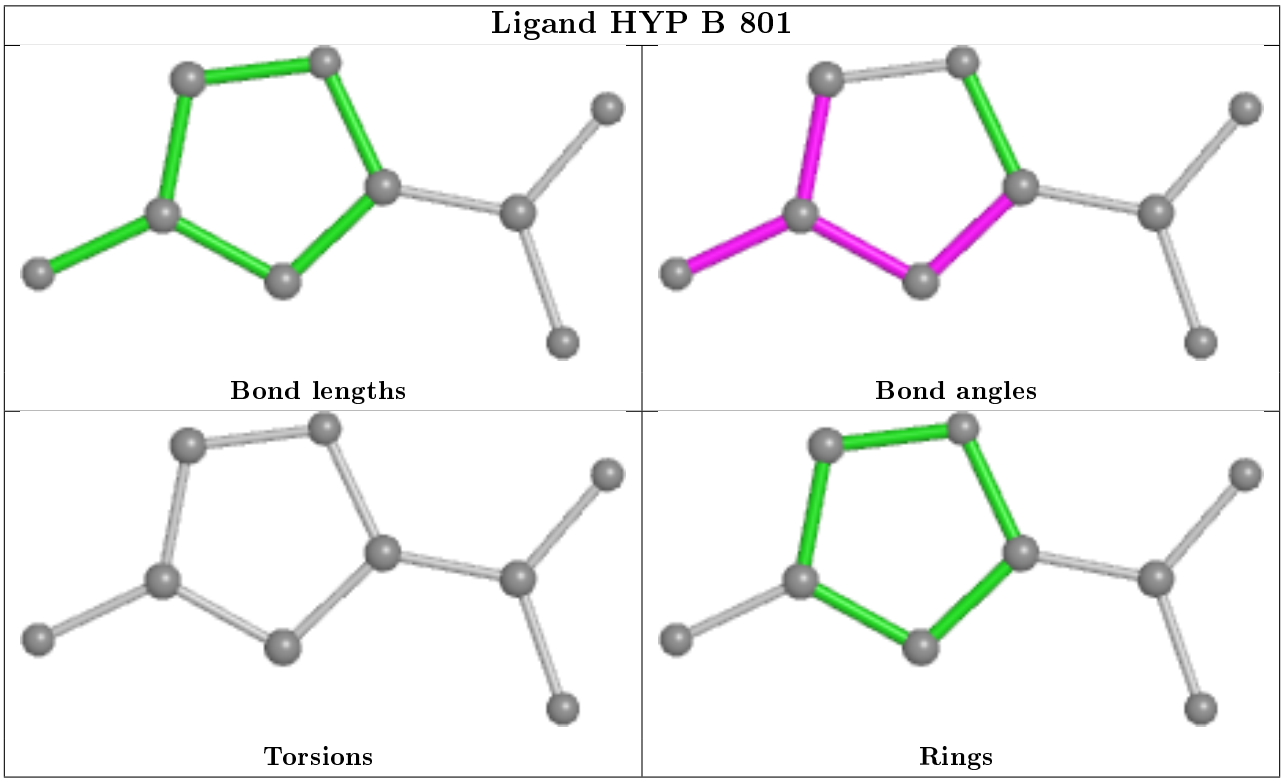
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	801	HYP	8	0
2	A	801	HYP	2	0
2	G	801	HYP	5	0
2	F	801	HYP	3	0
2	E	801	HYP	3	0

Continued on next page...

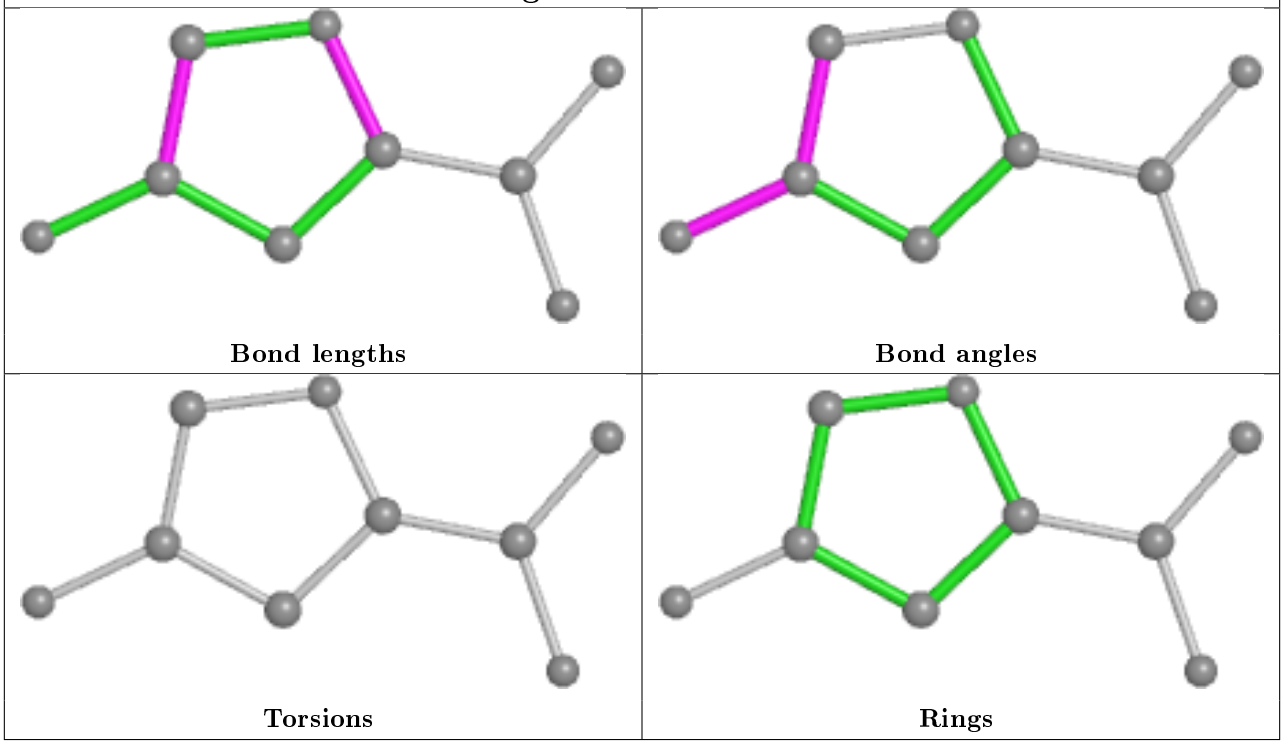
Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	801	HYP	4	0
2	H	801	HYP	5	0
2	C	801	HYP	2	0

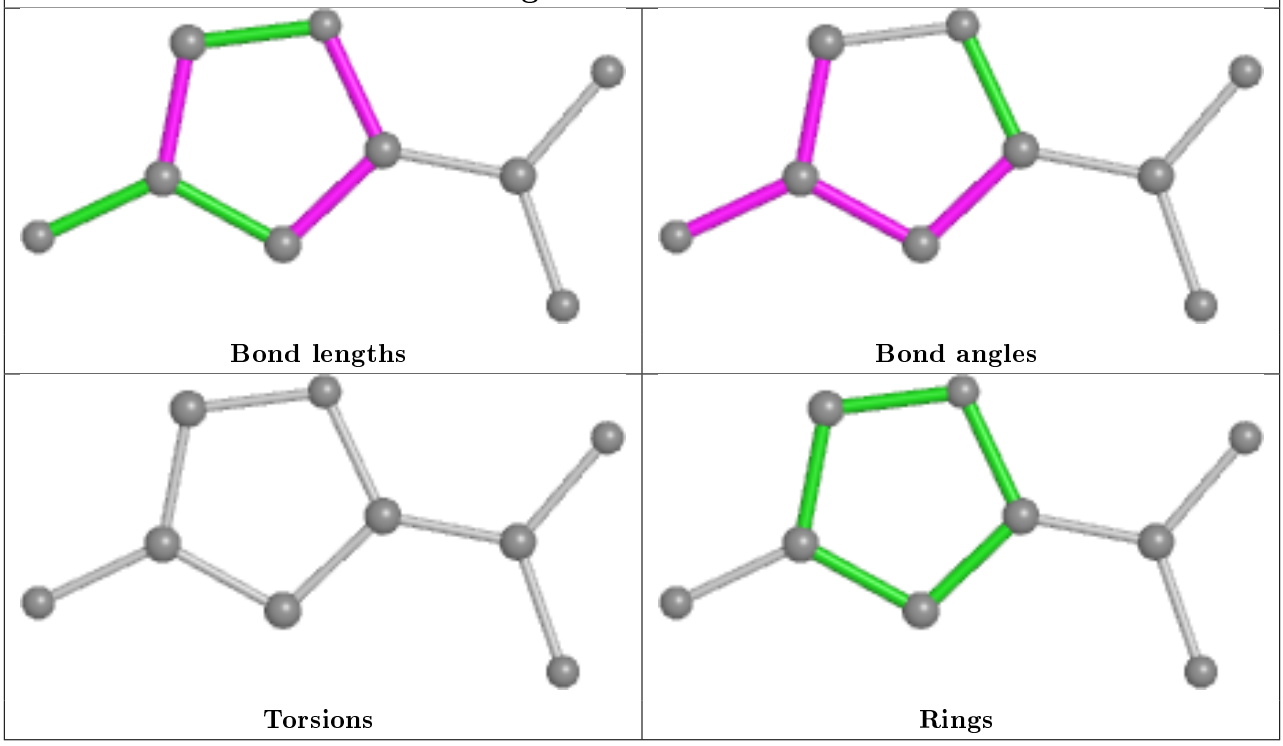
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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.



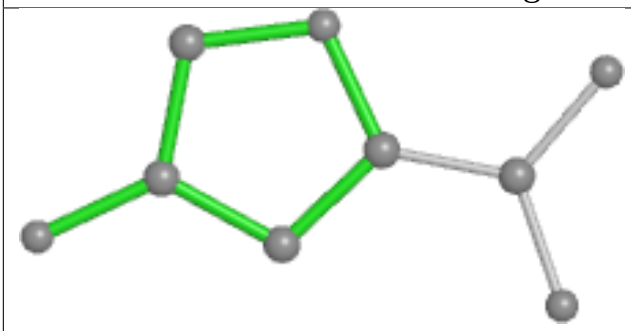
Ligand HYP A 801



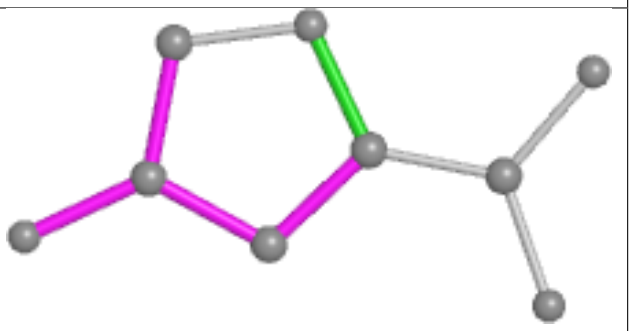
Ligand HYP G 801



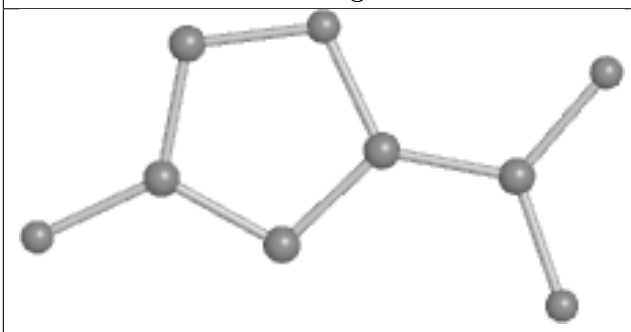
Ligand HYP F 801



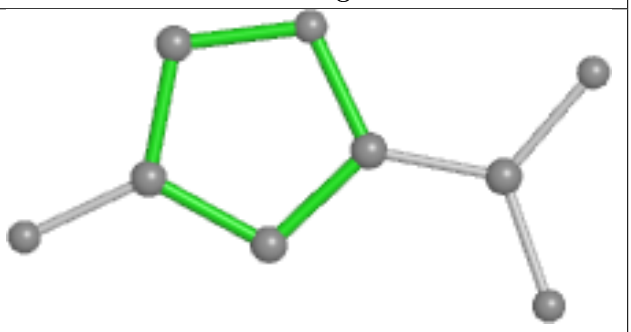
Bond lengths



Bond angles

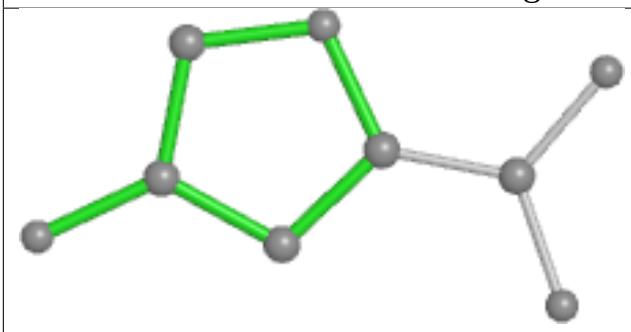


Torsions

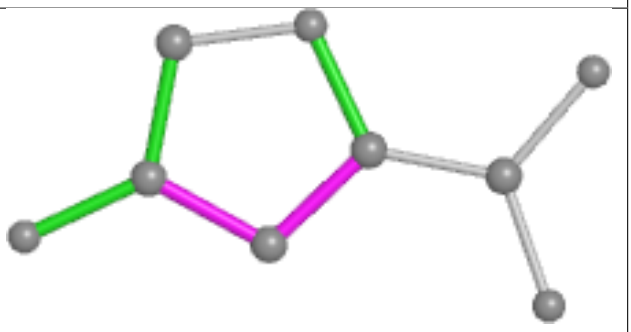


Rings

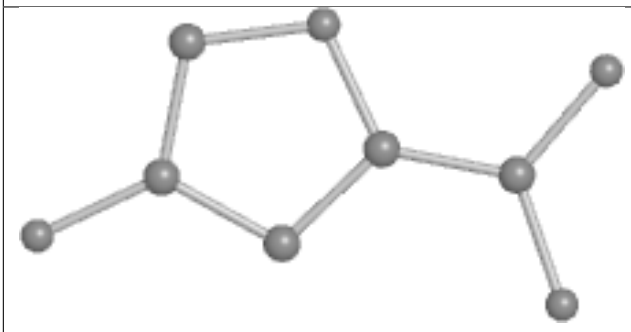
Ligand HYP E 801



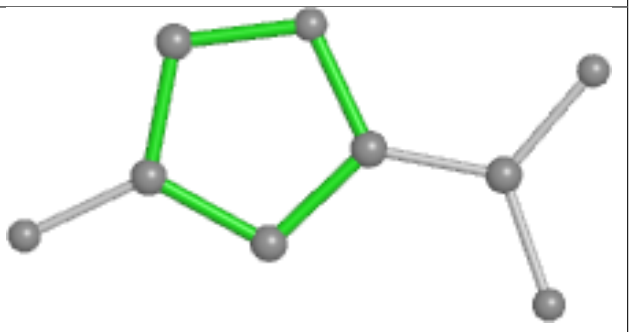
Bond lengths



Bond angles

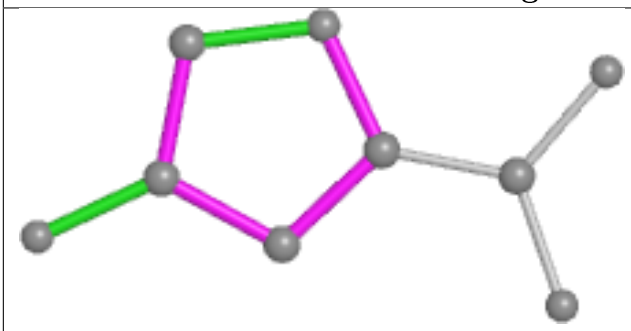


Torsions

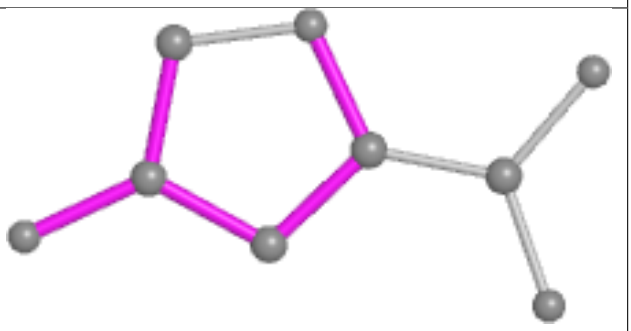


Rings

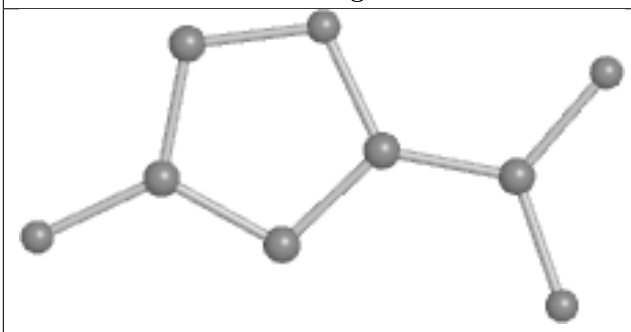
Ligand HYP D 801



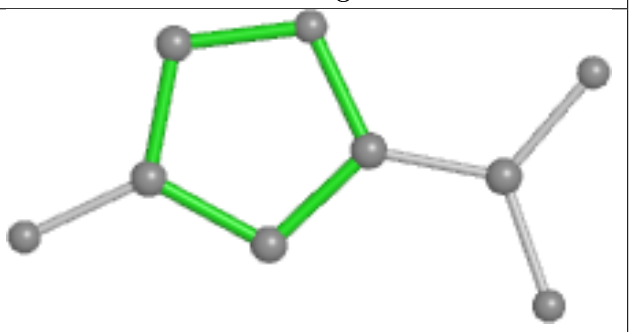
Bond lengths



Bond angles

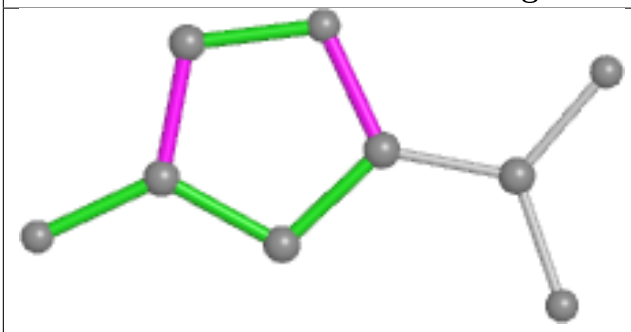


Torsions

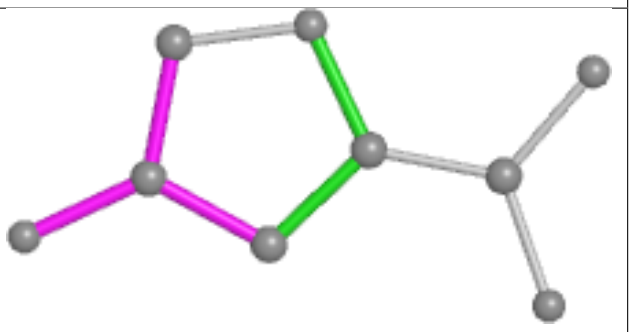


Rings

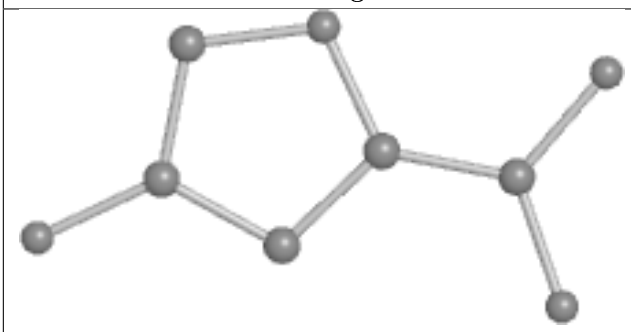
Ligand HYP H 801



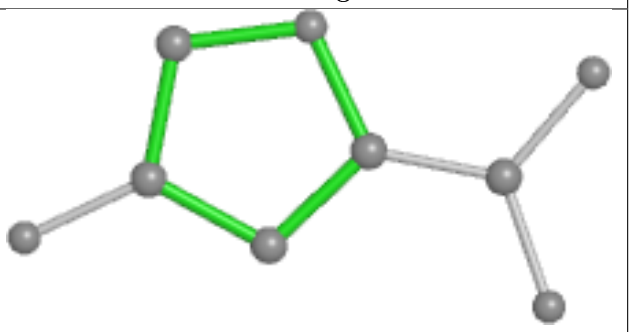
Bond lengths



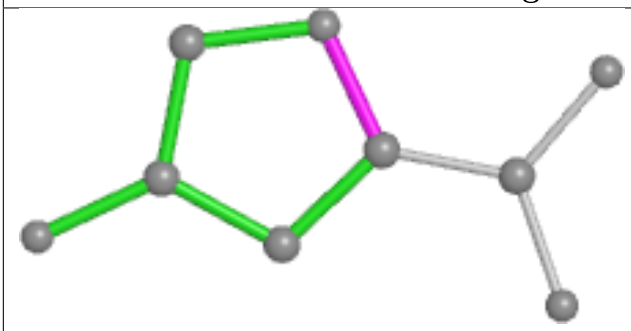
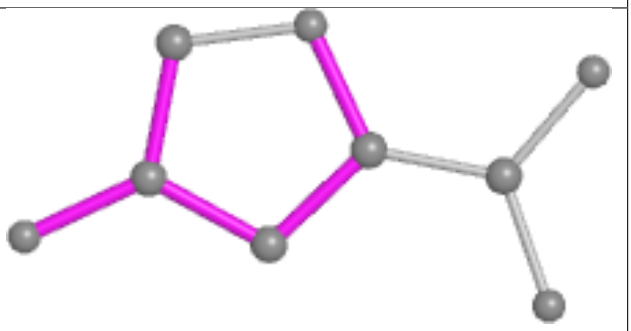
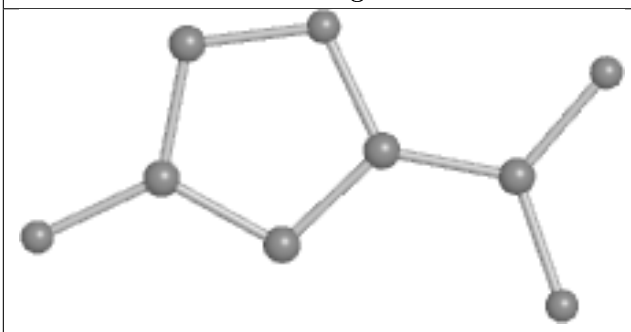
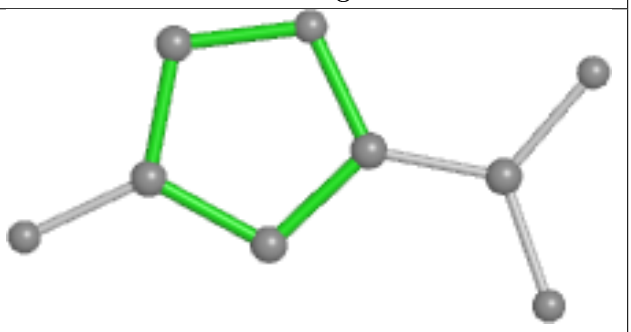
Bond angles



Torsions



Rings

Ligand HYP C 801	
 Bond lengths	 Bond angles
 Torsions	 Rings

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	4
1	H	2
1	C	2
1	A	2
1	E	1
1	F	1

The worst 5 of 12 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	31[A]:MET	C	32:THR	N	1.19
1	B	30:LEU	C	31[A]:MET	N	1.18
1	H	604:LYS	C	605:TYR	N	1.18
1	B	61:LEU	C	62:SER	N	1.16

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	642:MET	C	643:LEU	N	1.16

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^{th} 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	789/809 (97%)	0.01	7 (0%)	84	85	22, 34, 47, 67	0
1	B	789/809 (97%)	-0.07	8 (1%)	82	83	26, 34, 45, 59	0
1	C	789/809 (97%)	0.18	22 (2%)	53	49	26, 40, 57, 69	0
1	D	789/809 (97%)	-0.01	8 (1%)	82	83	22, 33, 47, 67	0
1	E	789/809 (97%)	-0.10	3 (0%)	92	93	26, 35, 46, 57	0
1	F	789/809 (97%)	0.01	12 (1%)	73	71	23, 31, 46, 64	0
1	G	789/809 (97%)	-0.20	2 (0%)	94	94	24, 31, 43, 57	0
1	H	789/809 (97%)	0.14	27 (3%)	45	41	24, 39, 55, 70	0
All	All	6312/6472 (97%)	-0.00	89 (1%)	75	74	22, 34, 50, 70	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	1	MET	7.3
1	C	233	ALA	6.1
1	E	1	MET	6.0
1	C	1	MET	5.8
1	A	1	MET	5.7

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands (i)

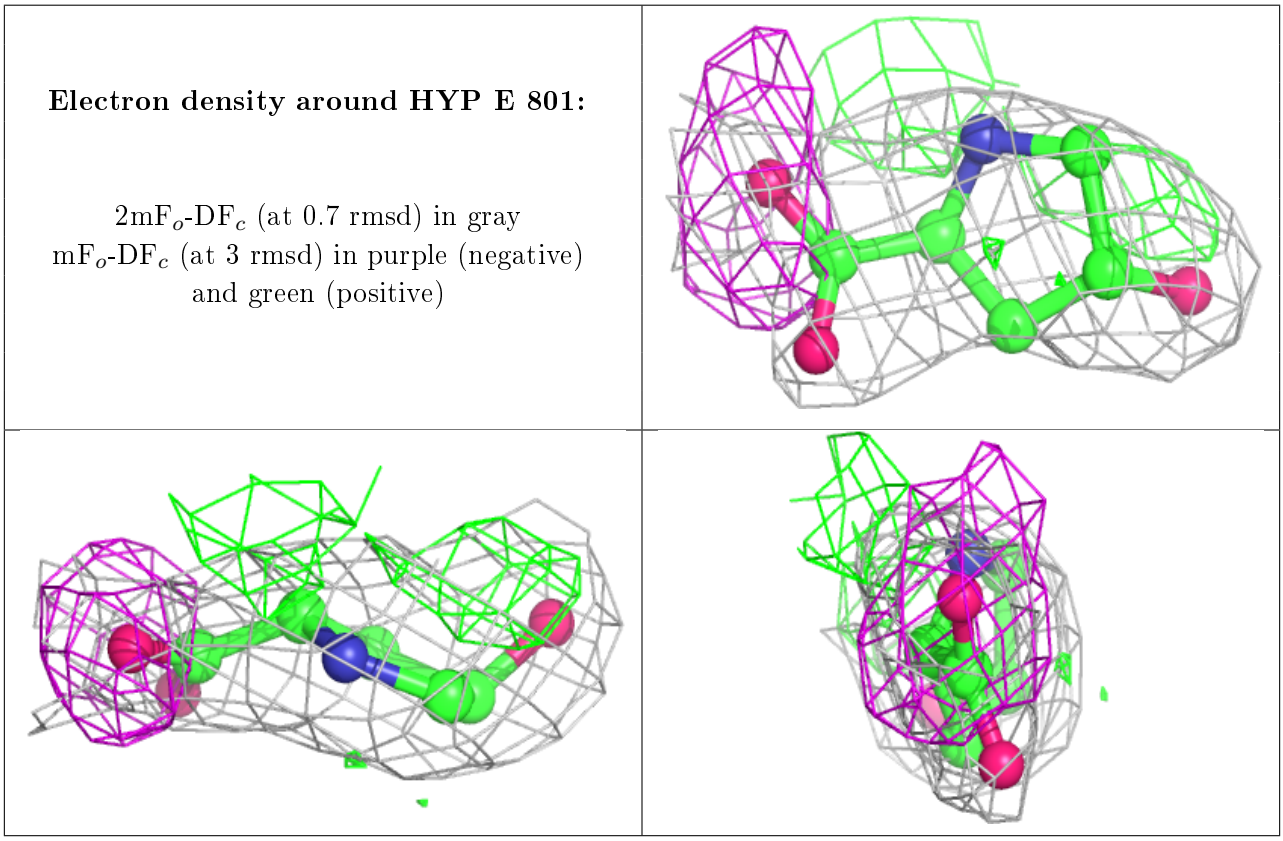
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	HYP	E	801	9/9	0.75	0.37	30,30,30,30	9
2	HYP	B	801	9/9	0.85	0.26	30,30,30,30	0
2	HYP	F	801	9/9	0.86	0.33	30,30,30,30	9
2	HYP	H	801	9/9	0.89	0.28	30,30,30,30	9
2	HYP	A	801	9/9	0.91	0.17	20,20,20,20	0
2	HYP	D	801	9/9	0.93	0.27	30,30,30,30	9
2	HYP	G	801	9/9	0.93	0.19	30,30,30,30	0
2	HYP	C	801	9/9	0.93	0.17	30,30,30,30	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.

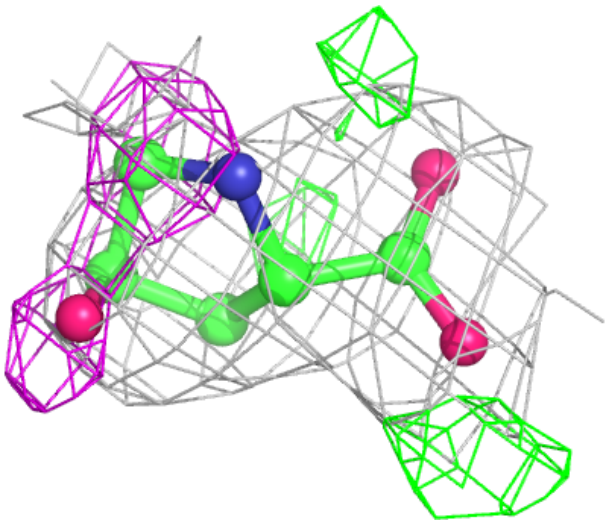
Electron density around HYP E 801:

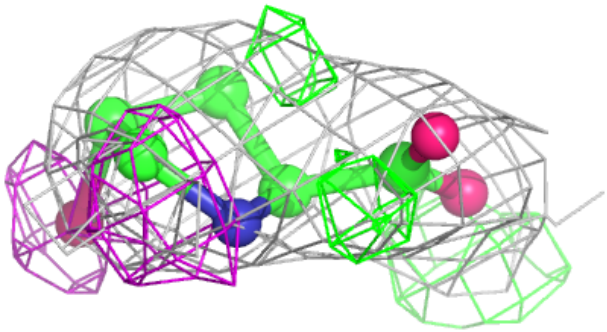
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

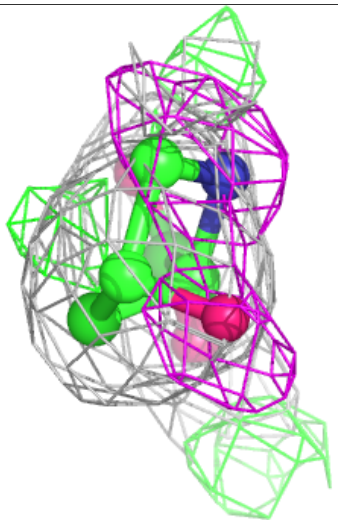


Electron density around HYP B 801:

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

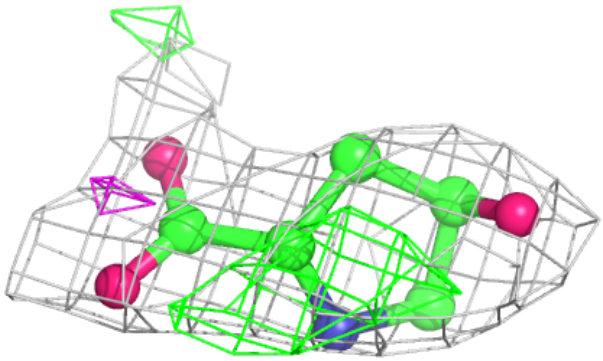


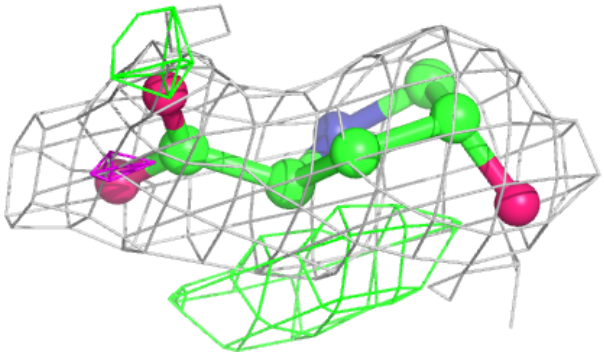


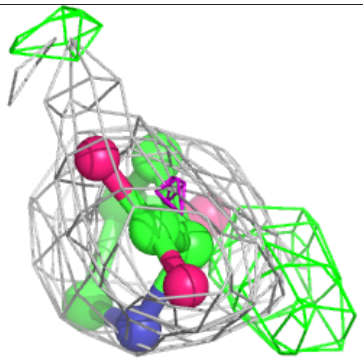


Electron density around HYP F 801:

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

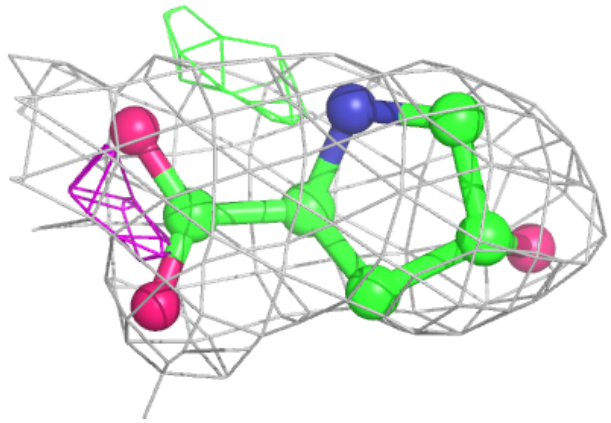


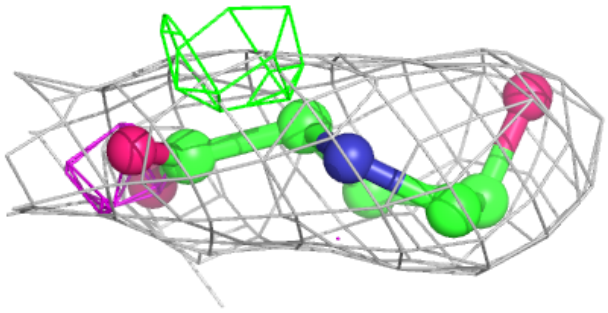


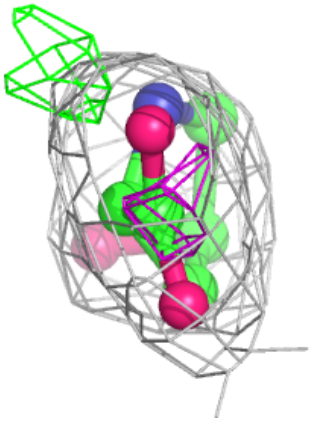


Electron density around HYP H 801:

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

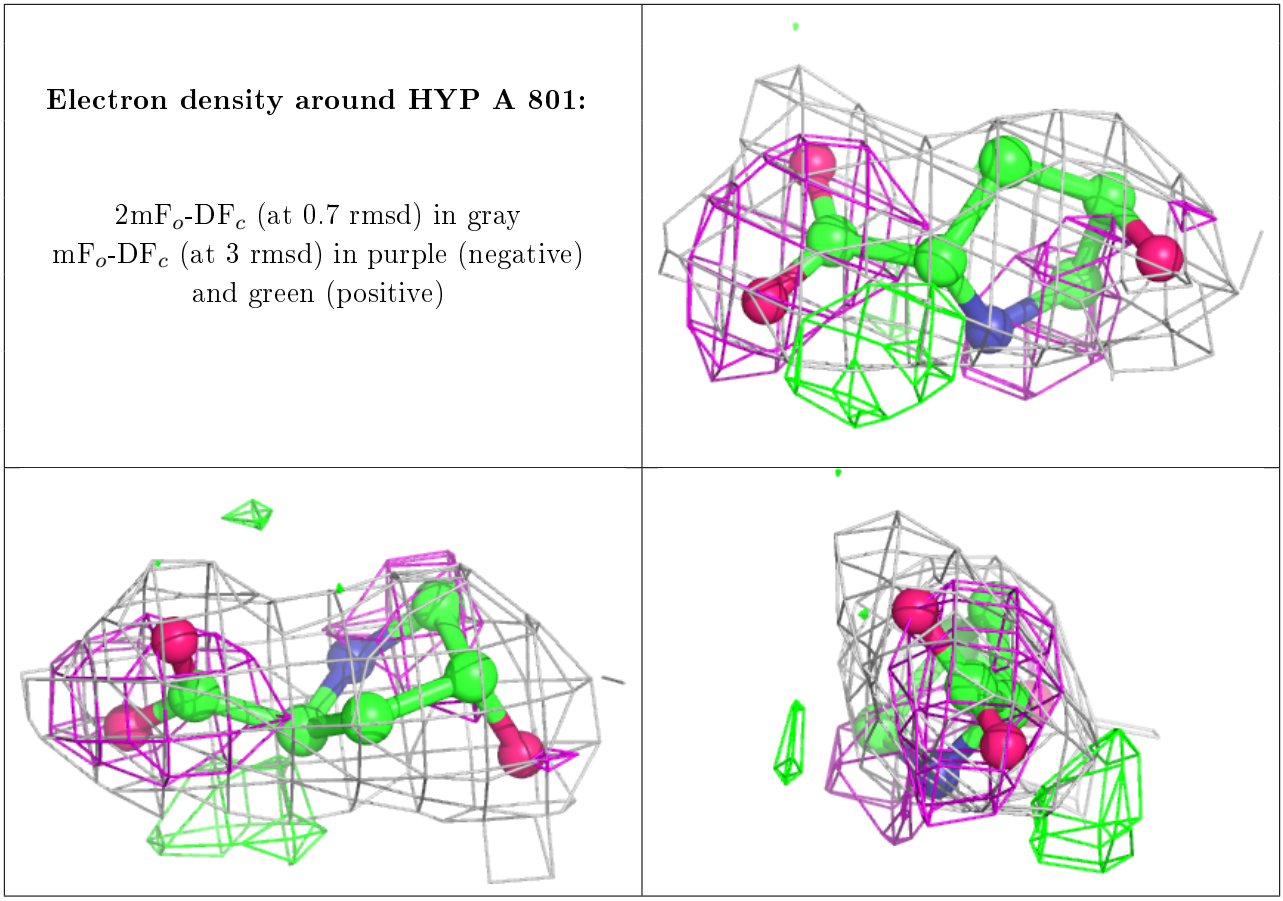






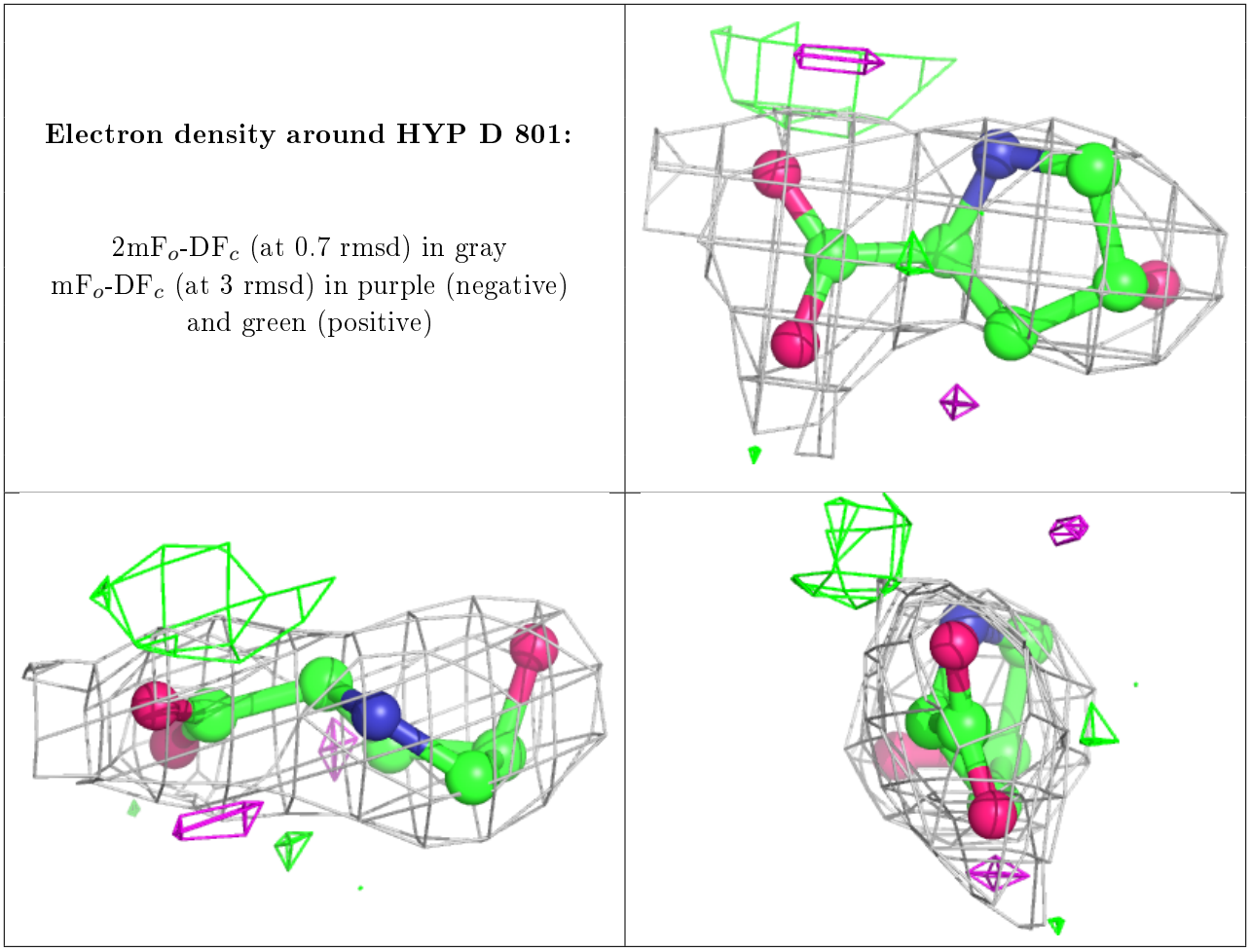
Electron density around HYP A 801:

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



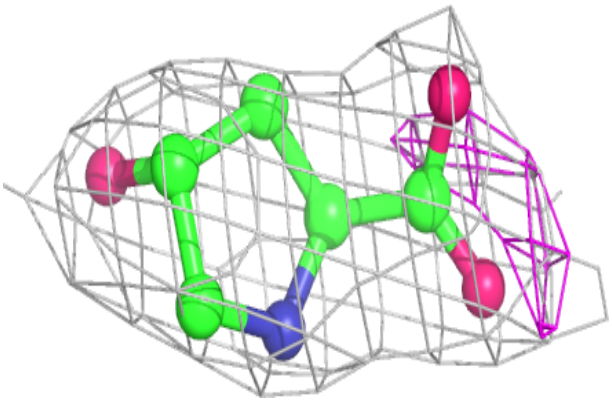
Electron density around HYP D 801:

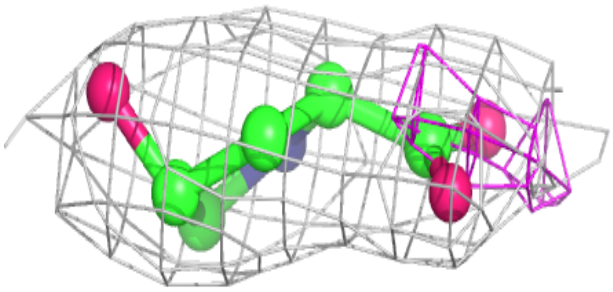
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

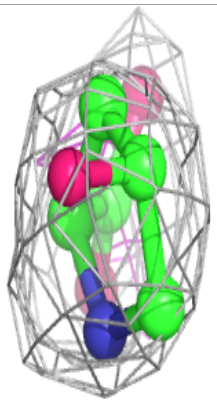


Electron density around HYP G 801:

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

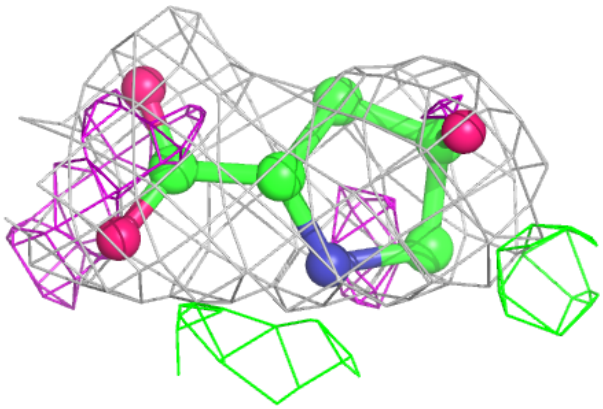


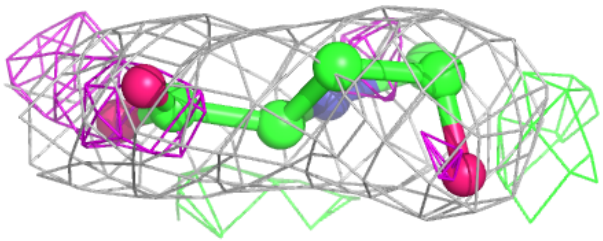


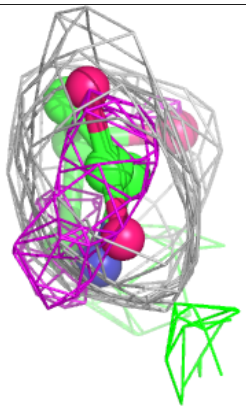


Electron density around HYP C 801:

$2mF_o$-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 (i)

There are no such residues in this entry.