



# wwPDB X-ray Structure Validation Summary Report

Feb 15, 2017 – 12:52 am GMT

PDB ID : 1WCY  
Title : Crystal Structure Of Human Dipeptidyl Peptidase IV (DPPIV) Complex With Diprotin A  
Authors : Hiramatsu, H.; Yamamoto, A.; Kyono, K.; Higashiyama, Y.; Fukushima, C.; Shima, H.; Sugiyama, S.; Inaka, K.; Shimizu, R.  
Deposited on : 2004-05-07  
Resolution : 2.20 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

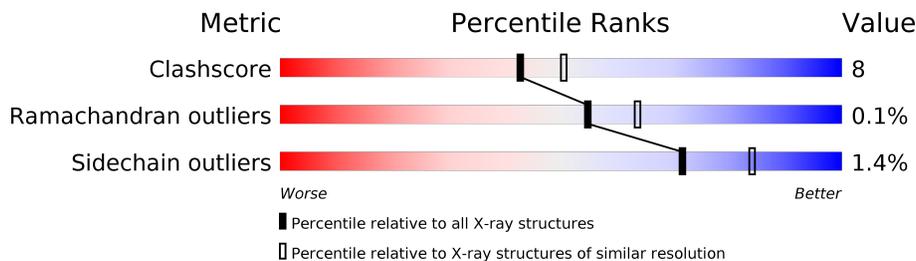
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	740	
1	B	740	
2	C	3	
2	D	3	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 13485 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 Dipeptidyl peptidase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	729	5971	3831	983	1131	26	0	0	0
1	B	729	5971	3831	983	1131	26	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	767	HIS	-	EXPRESSION TAG	UNP P27487
A	768	HIS	-	EXPRESSION TAG	UNP P27487
A	769	HIS	-	EXPRESSION TAG	UNP P27487
A	770	HIS	-	EXPRESSION TAG	UNP P27487
A	771	HIS	-	EXPRESSION TAG	UNP P27487
A	772	HIS	-	EXPRESSION TAG	UNP P27487
B	767	HIS	-	EXPRESSION TAG	UNP P27487
B	768	HIS	-	EXPRESSION TAG	UNP P27487
B	769	HIS	-	EXPRESSION TAG	UNP P27487
B	770	HIS	-	EXPRESSION TAG	UNP P27487
B	771	HIS	-	EXPRESSION TAG	UNP P27487
B	772	HIS	-	EXPRESSION TAG	UNP P27487

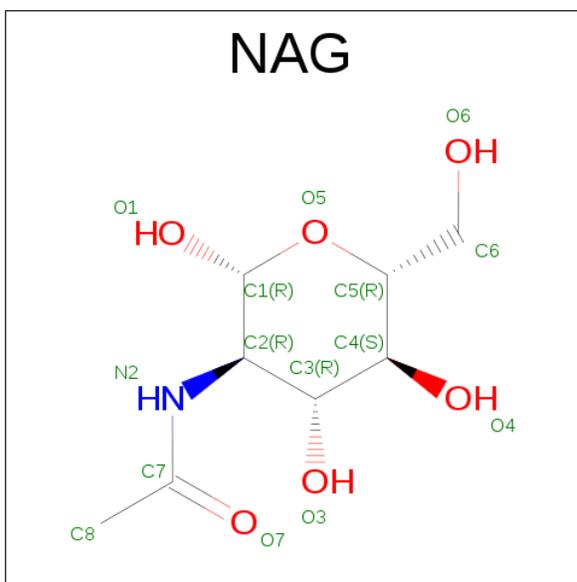
- Molecule 2 is a protein called Diprotin A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	3	24	17	3	4	0	0	0
2	D	3	24	17	3	4	0	0	0

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			24	14	1	9		
3	B	2	Total	C	N	O	0	0
			24	14	1	9		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			28	16	2	10		
5	A	2	Total	C	N	O	0	0
			28	16	2	10		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			28	16	2	10		
5	B	2	Total	C	N	O	0	0
			28	16	2	10		
5	B	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 6 is water.

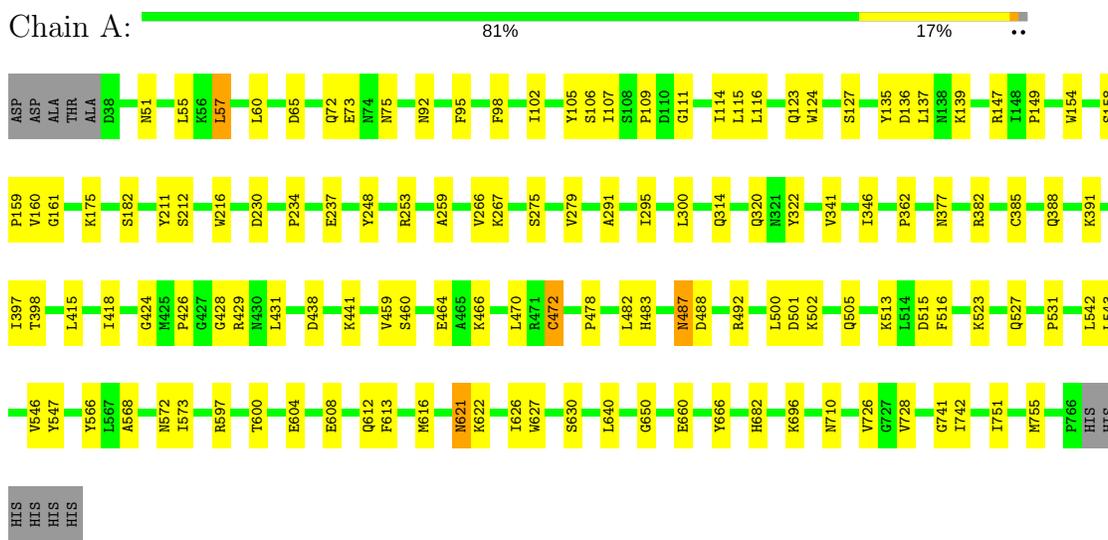
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	602	Total	O	0	0
			602	602		
6	B	631	Total	O	0	0
			631	631		
6	C	2	Total	O	0	0
			2	2		
6	D	2	Total	O	0	0
			2	2		

### 3 Residue-property plots

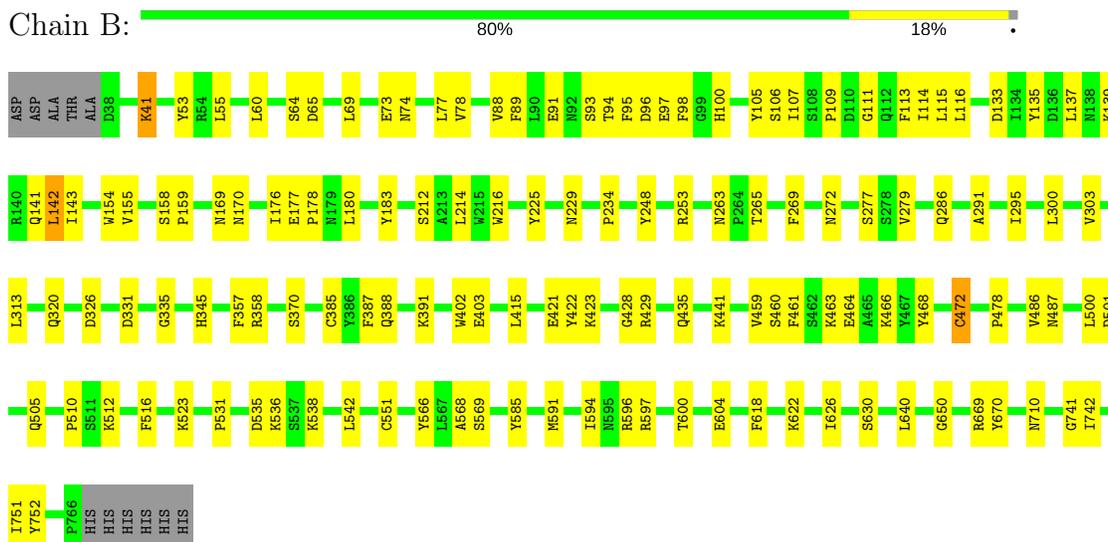
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

- Molecule 1: Dipeptidyl peptidase IV



- Molecule 1: Dipeptidyl peptidase IV



- Molecule 2: Diprotin A

Chain C:  100%

I1  
P2  
I3

- Molecule 2: Diprotin A

Chain D:  100%

I1  
P2  
I3

## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.02Å 126.06Å 136.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20	Depositor
% Data completeness (in resolution range)	99.8 (20.00-2.20)	Depositor
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNX	Depositor
R, $R_{free}$	0.224 , 0.265	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	13485	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

## 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: NAG, FUC

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	0.27	0/6143	0.60	2/8355 (0.0%)
1	B	0.27	0/6143	0.60	1/8355 (0.0%)
2	C	0.47	0/24	0.71	0/31
2	D	0.48	0/24	0.72	0/31
All	All	0.27	0/12334	0.60	3/16772 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	388	GLN	N-CA-C	-5.45	96.30	111.00
1	A	300	LEU	N-CA-C	-5.31	96.67	111.00
1	B	300	LEU	N-CA-C	-5.11	97.21	111.00

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	5971	0	5683	88	0
1	B	5971	0	5683	96	0
2	C	24	0	31	3	0
2	D	24	0	31	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	24	0	22	0	0
3	B	24	0	22	0	0
4	A	28	0	26	2	0
4	B	42	0	39	0	0
5	A	84	0	75	3	0
5	B	56	0	50	2	0
6	A	602	0	0	5	0
6	B	631	0	0	2	0
6	C	2	0	0	0	0
6	D	2	0	0	0	0
All	All	13485	0	11662	186	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 8.

The worst 5 of 186 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:A:253:ARG:HH21	1:B:253:ARG:HH21	1.06	0.93
1:B:142:LEU:H	1:B:142:LEU:HD23	1.42	0.83
1:A:726:VAL:HG23	1:A:728:VAL:HG23	1.62	0.82
1:B:78:VAL:HG23	1:B:89:PHE:HB2	1.61	0.82
1:B:429:ARG:HH11	1:B:429:ARG:HG3	1.51	0.75

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	727/740 (98%)	695 (96%)	31 (4%)	1 (0%)	55 63
1	B	727/740 (98%)	703 (97%)	23 (3%)	1 (0%)	55 63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	1/3 (33%)	1 (100%)	0	0	100	100
2	D	1/3 (33%)	1 (100%)	0	0	100	100
All	All	1456/1486 (98%)	1400 (96%)	54 (4%)	2 (0%)	55	63

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	320	GLN
1	B	320	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	654/663 (99%)	644 (98%)	10 (2%)	70	82
1	B	654/663 (99%)	645 (99%)	9 (1%)	71	84
2	C	3/3 (100%)	3 (100%)	0	100	100
2	D	3/3 (100%)	3 (100%)	0	100	100
All	All	1314/1332 (99%)	1295 (99%)	19 (1%)	71	84

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	627	TRP
1	B	41	LYS
1	B	385	CYS
1	A	621	ASN
1	B	472	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	455	GLN
1	A	621	ASN
1	B	572	ASN
1	A	487	ASN
1	A	697	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates [i](#)

14 carbohydrates 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	$\# Z  > 2$
5	NAG	A	1202	1,5	14,14,15	0.47	0	15,19,21	0.70	0
5	NAG	A	1203	5	14,14,15	0.50	0	15,19,21	0.59	0
5	NAG	A	1204	1,5	14,14,15	0.49	0	15,19,21	0.80	1 (6%)
5	NAG	A	1205	5	14,14,15	0.46	0	15,19,21	0.71	0
5	NAG	A	1206	1,5	14,14,15	0.51	0	15,19,21	0.87	1 (6%)
5	NAG	A	1207	5	14,14,15	0.44	0	15,19,21	0.78	1 (6%)
3	NAG	A	773	1,3	14,14,15	0.51	0	15,19,21	0.65	0
3	FUC	A	774	3	9,10,11	0.43	0	13,14,16	0.43	0
5	NAG	B	1202	1,5	14,14,15	0.50	0	15,19,21	0.73	1 (6%)
5	NAG	B	1203	5	14,14,15	0.47	0	15,19,21	0.71	0
5	NAG	B	1204	1,5	14,14,15	0.50	0	15,19,21	0.73	1 (6%)
5	NAG	B	1205	5	14,14,15	0.45	0	15,19,21	0.71	0
3	NAG	B	773	1,3	14,14,15	0.55	0	15,19,21	0.71	0
3	FUC	B	774	3	9,10,11	0.46	0	13,14,16	0.41	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	NAG	A	1202	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1203	5	-	0/6/23/26	0/1/1/1
5	NAG	A	1204	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1205	5	-	0/6/23/26	0/1/1/1
5	NAG	A	1206	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1207	5	-	0/6/23/26	0/1/1/1
3	NAG	A	773	1,3	-	0/6/23/26	0/1/1/1
3	FUC	A	774	3	-	0/0/17/20	0/1/1/1
5	NAG	B	1202	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	1203	5	-	2/6/23/26	0/1/1/1
5	NAG	B	1204	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	1205	5	-	2/6/23/26	0/1/1/1
3	NAG	B	773	1,3	-	0/6/23/26	0/1/1/1
3	FUC	B	774	3	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1206	NAG	C2-N2-C7	-2.22	119.70	122.94
5	A	1207	NAG	C2-N2-C7	-2.13	119.83	122.94
5	A	1204	NAG	C2-N2-C7	-2.04	119.97	122.94
5	B	1202	NAG	C2-N2-C7	-2.01	120.00	122.94
5	B	1204	NAG	C2-N2-C7	-2.00	120.02	122.94

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	1203	NAG	C8-C7-N2-C2
5	B	1205	NAG	O7-C7-N2-C2
5	B	1203	NAG	O7-C7-N2-C2
5	B	1205	NAG	C8-C7-N2-C2

There are no ring outliers.

7 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1202	NAG	2	0
5	A	1203	NAG	2	0
5	A	1204	NAG	1	0
5	A	1205	NAG	1	0
5	B	1203	NAG	1	0
5	B	1204	NAG	1	0
5	B	1205	NAG	1	0

## 5.6 Ligand geometry [i](#)

5 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	$\# Z  > 2$
4	NAG	A	1201	1	14,14,15	0.49	0	15,19,21	0.74	1 (6%)
4	NAG	A	1601	1	14,14,15	0.53	0	15,19,21	0.75	0
4	NAG	B	1201	1	14,14,15	0.48	0	15,19,21	0.73	1 (6%)
4	NAG	B	1501	1	14,14,15	0.46	0	15,19,21	0.77	1 (6%)
4	NAG	B	1601	1	14,14,15	0.52	0	15,19,21	0.73	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	1201	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1601	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1201	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1501	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1601	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1501	NAG	C2-N2-C7	-2.15	119.81	122.94
4	B	1201	NAG	C2-N2-C7	-2.05	119.95	122.94
4	A	1201	NAG	C2-N2-C7	-2.01	120.01	122.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1201	NAG	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.