



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 01:12 PM GMT

PDB ID : 3G0G  
Title : Crystal structure of dipeptidyl peptidase IV in complex with a pyrimidinone inhibitor 3  
Authors : Zhang, Z.; Wallace, M.B.; Feng, J.; Stafford, J.A.; Kaldor, S.W.; Shi, L.; Skene, R.J.; Aertgeerts, K.; Lee, B.; Jennings, A.; Xu, R.; Kassel, D.; Webb, D.R.; Gwaltney, S.L.  
Deposited on : 2009-01-27  
Resolution : 2.45 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

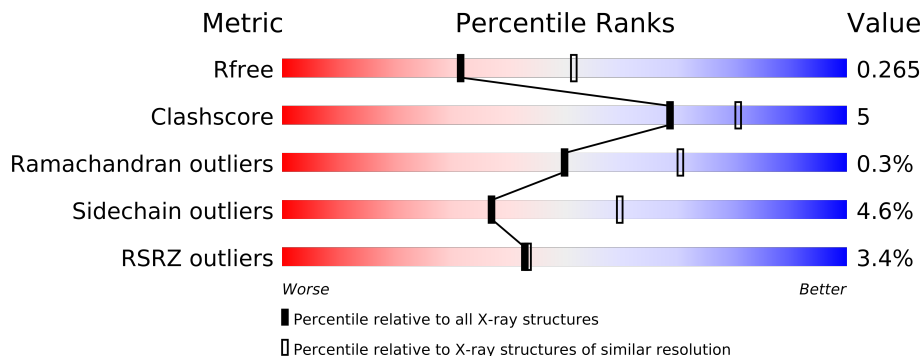
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.45 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))
$R_{free}$	66092	3566 (2.50-2.42)
Clashscore	79885	4471 (2.50-2.42)
Ramachandran outliers	78287	4383 (2.50-2.42)
Sidechain outliers	78261	4385 (2.50-2.42)
RSRZ outliers	66119	3568 (2.50-2.42)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

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

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	RUM	A	800	-	X
2	RUM	C	800	-	X
3	NAG	A	803	-	X
3	NAG	A	808	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
3	NAG	B	808	-	X
3	NAG	C	808	-	X
3	NAG	D	803	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 24783 atoms, of which 0 are hydrogen and 0 are deuterium.

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 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	724	Total	C	N	O	S	0	2	0
			5933	3811	975	1121	26			
1	B	730	Total	C	N	O	S	0	2	0
			5990	3844	991	1129	26			
1	C	724	Total	C	N	O	S	0	1	0
			5929	3809	974	1120	26			
1	D	690	Total	C	N	O	S	0	1	0
			5649	3631	926	1066	26			

There are 48 discrepancies between the modelled and reference sequences:

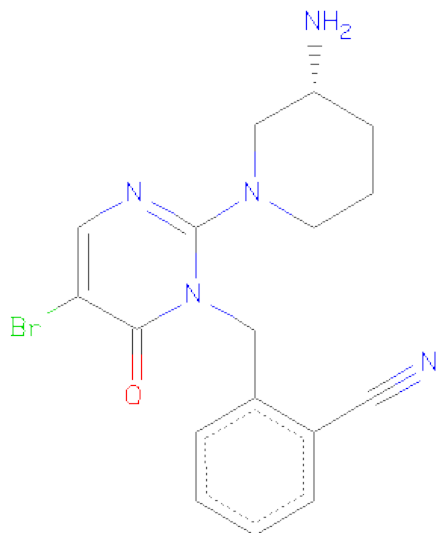
Chain	Residue	Modelled	Actual	Comment	Reference
A	27	ALA	-	EXPRESSION TAG	UNP P27487
A	28	ASP	-	EXPRESSION TAG	UNP P27487
A	29	PRO	-	EXPRESSION TAG	UNP P27487
A	30	GLY	-	EXPRESSION TAG	UNP P27487
A	31	GLY	-	EXPRESSION TAG	UNP P27487
A	32	SER	-	EXPRESSION TAG	UNP P27487
A	33	HIS	-	EXPRESSION TAG	UNP P27487
A	34	HIS	-	EXPRESSION TAG	UNP P27487
A	35	HIS	-	EXPRESSION TAG	UNP P27487
A	36	HIS	-	EXPRESSION TAG	UNP P27487
A	37	HIS	-	EXPRESSION TAG	UNP P27487
A	38	HIS	-	EXPRESSION TAG	UNP P27487
B	27	ALA	-	EXPRESSION TAG	UNP P27487
B	28	ASP	-	EXPRESSION TAG	UNP P27487
B	29	PRO	-	EXPRESSION TAG	UNP P27487
B	30	GLY	-	EXPRESSION TAG	UNP P27487
B	31	GLY	-	EXPRESSION TAG	UNP P27487
B	32	SER	-	EXPRESSION TAG	UNP P27487
B	33	HIS	-	EXPRESSION TAG	UNP P27487
B	34	HIS	-	EXPRESSION TAG	UNP P27487
B	35	HIS	-	EXPRESSION TAG	UNP P27487

*Continued on next page...*

*Continued from previous page...*

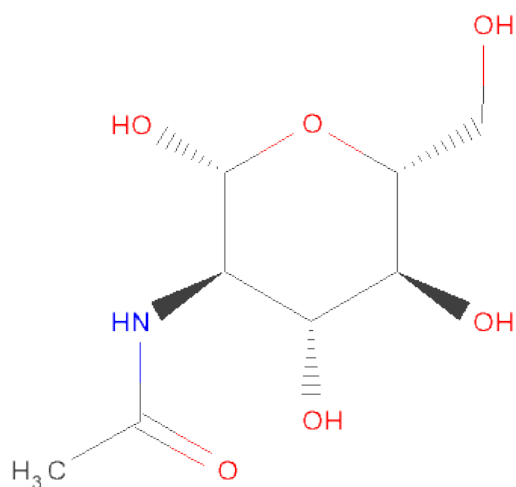
Chain	Residue	Modelled	Actual	Comment	Reference
B	36	HIS	-	EXPRESSION TAG	UNP P27487
B	37	HIS	-	EXPRESSION TAG	UNP P27487
B	38	HIS	-	EXPRESSION TAG	UNP P27487
C	27	ALA	-	EXPRESSION TAG	UNP P27487
C	28	ASP	-	EXPRESSION TAG	UNP P27487
C	29	PRO	-	EXPRESSION TAG	UNP P27487
C	30	GLY	-	EXPRESSION TAG	UNP P27487
C	31	GLY	-	EXPRESSION TAG	UNP P27487
C	32	SER	-	EXPRESSION TAG	UNP P27487
C	33	HIS	-	EXPRESSION TAG	UNP P27487
C	34	HIS	-	EXPRESSION TAG	UNP P27487
C	35	HIS	-	EXPRESSION TAG	UNP P27487
C	36	HIS	-	EXPRESSION TAG	UNP P27487
C	37	HIS	-	EXPRESSION TAG	UNP P27487
C	38	HIS	-	EXPRESSION TAG	UNP P27487
D	27	ALA	-	EXPRESSION TAG	UNP P27487
D	28	ASP	-	EXPRESSION TAG	UNP P27487
D	29	PRO	-	EXPRESSION TAG	UNP P27487
D	30	GLY	-	EXPRESSION TAG	UNP P27487
D	31	GLY	-	EXPRESSION TAG	UNP P27487
D	32	SER	-	EXPRESSION TAG	UNP P27487
D	33	HIS	-	EXPRESSION TAG	UNP P27487
D	34	HIS	-	EXPRESSION TAG	UNP P27487
D	35	HIS	-	EXPRESSION TAG	UNP P27487
D	36	HIS	-	EXPRESSION TAG	UNP P27487
D	37	HIS	-	EXPRESSION TAG	UNP P27487
D	38	HIS	-	EXPRESSION TAG	UNP P27487

- Molecule 2 is 2-({2-[(3R)-3-AMINOPIPERIDIN-1-YL]-5-BROMO-6-OXOPYRIMIDIN-1(6H)-YL}METHYL)BENZONITRILE (three-letter code: RUM) (formula: C<sub>17</sub>H<sub>18</sub>BrN<sub>5</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	Br	C	N	O	0	0
			24	1	17	5	1		
2	B	1	Total	Br	C	N	O	0	0
			24	1	17	5	1		
2	C	1	Total	Br	C	N	O	0	0
			24	1	17	5	1		
2	D	1	Total	Br	C	N	O	0	0
			24	1	17	5	1		

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

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

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	ALA	-	EXPRESSION TAG	UNP P27487
A	28	ASP	-	EXPRESSION TAG	UNP P27487
A	29	PRO	-	EXPRESSION TAG	UNP P27487
A	30	GLY	-	EXPRESSION TAG	UNP P27487
A	31	GLY	-	EXPRESSION TAG	UNP P27487
A	32	SER	-	EXPRESSION TAG	UNP P27487
A	33	HIS	-	EXPRESSION TAG	UNP P27487
A	34	HIS	-	EXPRESSION TAG	UNP P27487
A	35	HIS	-	EXPRESSION TAG	UNP P27487
A	36	HIS	-	EXPRESSION TAG	UNP P27487
A	37	HIS	-	EXPRESSION TAG	UNP P27487
A	38	HIS	-	EXPRESSION TAG	UNP P27487
A	27	ALA	-	EXPRESSION TAG	UNP P27487
A	28	ASP	-	EXPRESSION TAG	UNP P27487
A	29	PRO	-	EXPRESSION TAG	UNP P27487
A	30	GLY	-	EXPRESSION TAG	UNP P27487
A	31	GLY	-	EXPRESSION TAG	UNP P27487
A	32	SER	-	EXPRESSION TAG	UNP P27487
A	33	HIS	-	EXPRESSION TAG	UNP P27487
A	34	HIS	-	EXPRESSION TAG	UNP P27487
A	35	HIS	-	EXPRESSION TAG	UNP P27487
A	36	HIS	-	EXPRESSION TAG	UNP P27487
A	37	HIS	-	EXPRESSION TAG	UNP P27487
A	38	HIS	-	EXPRESSION TAG	UNP P27487
B	27	ALA	-	EXPRESSION TAG	UNP P27487
B	28	ASP	-	EXPRESSION TAG	UNP P27487
B	29	PRO	-	EXPRESSION TAG	UNP P27487
B	30	GLY	-	EXPRESSION TAG	UNP P27487
B	31	GLY	-	EXPRESSION TAG	UNP P27487
B	32	SER	-	EXPRESSION TAG	UNP P27487
B	33	HIS	-	EXPRESSION TAG	UNP P27487

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	34	HIS	-	EXPRESSION TAG	UNP P27487
B	35	HIS	-	EXPRESSION TAG	UNP P27487
B	36	HIS	-	EXPRESSION TAG	UNP P27487
B	37	HIS	-	EXPRESSION TAG	UNP P27487
B	38	HIS	-	EXPRESSION TAG	UNP P27487
B	27	ALA	-	EXPRESSION TAG	UNP P27487
B	28	ASP	-	EXPRESSION TAG	UNP P27487
B	29	PRO	-	EXPRESSION TAG	UNP P27487
B	30	GLY	-	EXPRESSION TAG	UNP P27487
B	31	GLY	-	EXPRESSION TAG	UNP P27487
B	32	SER	-	EXPRESSION TAG	UNP P27487
B	33	HIS	-	EXPRESSION TAG	UNP P27487
B	34	HIS	-	EXPRESSION TAG	UNP P27487
B	35	HIS	-	EXPRESSION TAG	UNP P27487
B	36	HIS	-	EXPRESSION TAG	UNP P27487
B	37	HIS	-	EXPRESSION TAG	UNP P27487
B	38	HIS	-	EXPRESSION TAG	UNP P27487
C	27	ALA	-	EXPRESSION TAG	UNP P27487
C	28	ASP	-	EXPRESSION TAG	UNP P27487
C	29	PRO	-	EXPRESSION TAG	UNP P27487
C	30	GLY	-	EXPRESSION TAG	UNP P27487
C	31	GLY	-	EXPRESSION TAG	UNP P27487
C	32	SER	-	EXPRESSION TAG	UNP P27487
C	33	HIS	-	EXPRESSION TAG	UNP P27487
C	34	HIS	-	EXPRESSION TAG	UNP P27487
C	35	HIS	-	EXPRESSION TAG	UNP P27487
C	36	HIS	-	EXPRESSION TAG	UNP P27487
C	37	HIS	-	EXPRESSION TAG	UNP P27487
C	38	HIS	-	EXPRESSION TAG	UNP P27487
C	27	ALA	-	EXPRESSION TAG	UNP P27487
C	28	ASP	-	EXPRESSION TAG	UNP P27487
C	29	PRO	-	EXPRESSION TAG	UNP P27487
C	30	GLY	-	EXPRESSION TAG	UNP P27487
C	31	GLY	-	EXPRESSION TAG	UNP P27487
C	32	SER	-	EXPRESSION TAG	UNP P27487
C	33	HIS	-	EXPRESSION TAG	UNP P27487
C	34	HIS	-	EXPRESSION TAG	UNP P27487
C	35	HIS	-	EXPRESSION TAG	UNP P27487
C	36	HIS	-	EXPRESSION TAG	UNP P27487
C	37	HIS	-	EXPRESSION TAG	UNP P27487
C	38	HIS	-	EXPRESSION TAG	UNP P27487
D	27	ALA	-	EXPRESSION TAG	UNP P27487

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	28	ASP	-	EXPRESSION TAG	UNP P27487
D	29	PRO	-	EXPRESSION TAG	UNP P27487
D	30	GLY	-	EXPRESSION TAG	UNP P27487
D	31	GLY	-	EXPRESSION TAG	UNP P27487
D	32	SER	-	EXPRESSION TAG	UNP P27487
D	33	HIS	-	EXPRESSION TAG	UNP P27487
D	34	HIS	-	EXPRESSION TAG	UNP P27487
D	35	HIS	-	EXPRESSION TAG	UNP P27487
D	36	HIS	-	EXPRESSION TAG	UNP P27487
D	37	HIS	-	EXPRESSION TAG	UNP P27487
D	38	HIS	-	EXPRESSION TAG	UNP P27487
D	27	ALA	-	EXPRESSION TAG	UNP P27487
D	28	ASP	-	EXPRESSION TAG	UNP P27487
D	29	PRO	-	EXPRESSION TAG	UNP P27487
D	30	GLY	-	EXPRESSION TAG	UNP P27487
D	31	GLY	-	EXPRESSION TAG	UNP P27487
D	32	SER	-	EXPRESSION TAG	UNP P27487
D	33	HIS	-	EXPRESSION TAG	UNP P27487
D	34	HIS	-	EXPRESSION TAG	UNP P27487
D	35	HIS	-	EXPRESSION TAG	UNP P27487
D	36	HIS	-	EXPRESSION TAG	UNP P27487
D	37	HIS	-	EXPRESSION TAG	UNP P27487
D	38	HIS	-	EXPRESSION TAG	UNP P27487

- Molecule 5 is water.

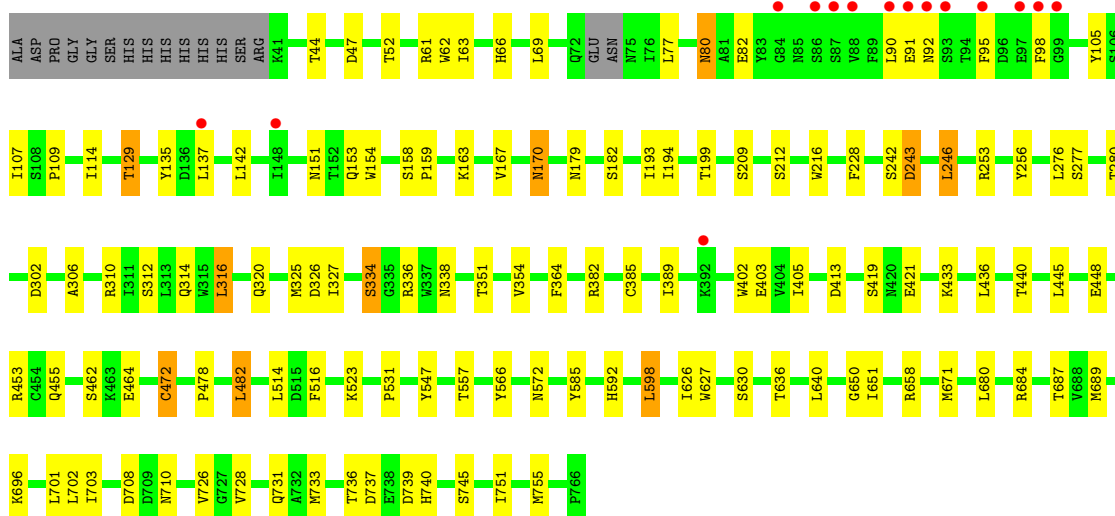
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	39	Total O 39 39	0	0
5	B	699	Total O 699 699	0	0

### 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 errors 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.

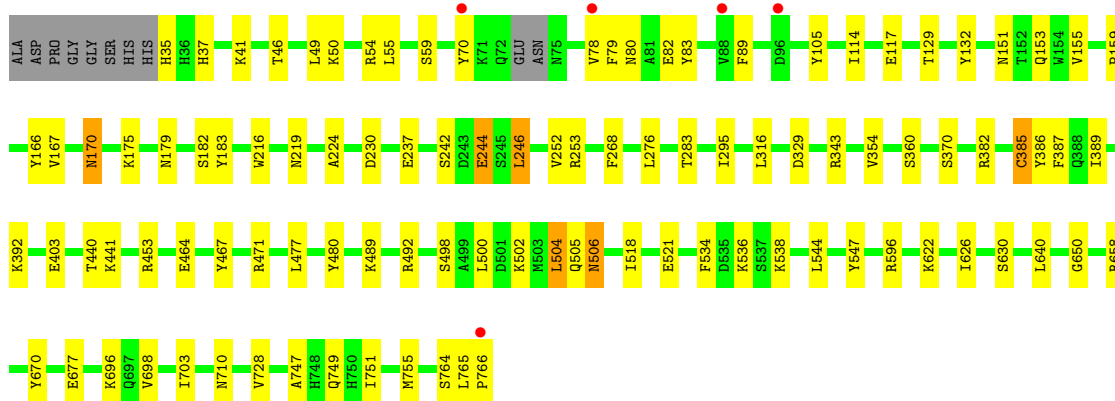
#### • Molecule 1: Dipeptidyl peptidase 4

Chain A: 



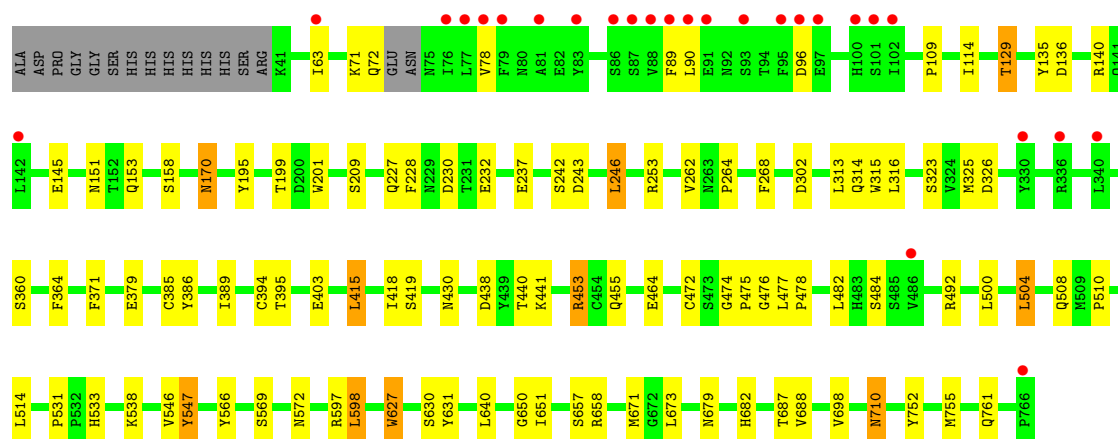
#### • Molecule 1: Dipeptidyl peptidase 4

Chain B: 



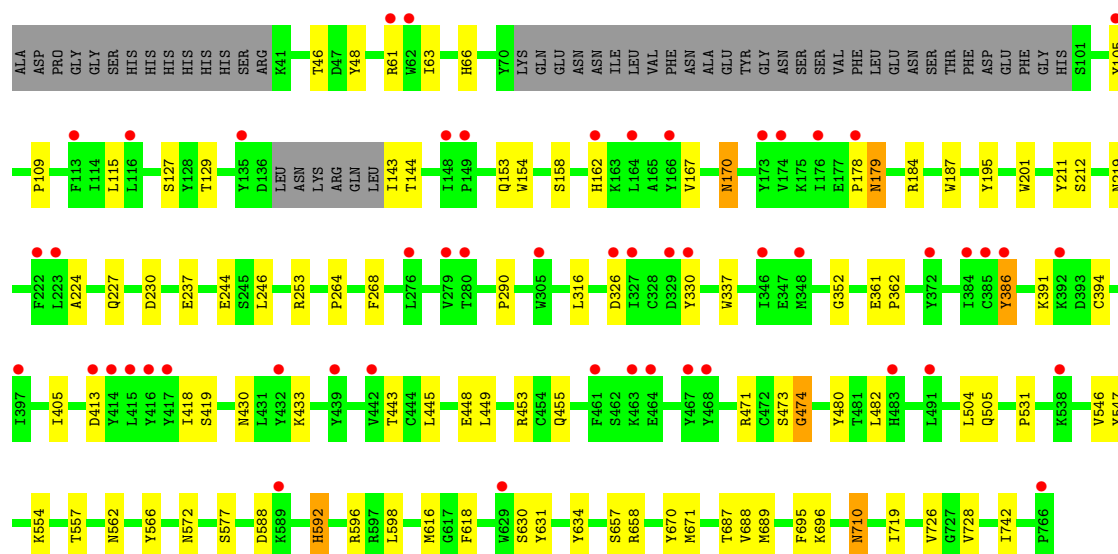
#### • Molecule 1: Dipeptidyl peptidase 4

Chain C: 



• Molecule 1: Dipeptidyl peptidase 4

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.91Å 122.92Å 145.09Å 90.00° 114.77° 90.00°	Depositor
Resolution (Å)	50.00 – 2.45 48.52 – 2.45	Depositor EDS
% Data completeness (in resolution range)	95.4 (50.00-2.45) 95.5 (48.52-2.45)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.210 , 0.265 0.210 , 0.265	Depositor DCC
$R_{free}$ test set	6854 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	47.2	Xtriage
Anisotropy	0.177	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 39.2	EDS
Estimated twinning fraction	0.012 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 136813 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	24783	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, RUM

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	0.49	0/6108	0.61	0/8307
1	B	0.47	0/6169	0.61	0/8389
1	C	0.46	0/6100	0.61	1/8296 (0.0%)
1	D	0.43	0/5813	0.57	0/7908
All	All	0.46	0/24190	0.60	1/32900 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	415	LEU	CA-CB-CG	5.32	127.54	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5933	0	5648	71	0
1	B	5990	0	5692	56	0
1	C	5929	0	5646	48	0
1	D	5649	0	5381	47	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	24	0	18	3	0
2	B	24	0	18	1	0
2	C	24	0	18	3	0
2	D	24	0	18	1	0
3	A	56	0	52	0	0
3	B	56	0	52	0	0
3	C	56	0	52	0	0
3	D	56	0	52	0	0
4	A	56	0	50	0	0
4	B	56	0	50	0	0
4	C	56	0	50	0	0
4	D	56	0	50	0	0
5	A	39	0	0	0	0
5	B	699	0	0	5	0
All	All	24783	0	22847	224	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 224 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:253:ARG:HH21	1:B:253:ARG:HH21	0.99	0.98
1:C:153:GLN:HE22	1:C:170:ASN:H	1.20	0.87
1:A:253:ARG:HH21	1:B:253:ARG:NH2	1.78	0.81
1:A:153:GLN:HE22	1:A:170:ASN:H	1.30	0.79
1:C:640:LEU:HD11	1:C:650:GLY:HA3	1.65	0.79

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	721/740 (97%)	680 (94%)	40 (6%)	1 (0%)	59 80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	727/740 (98%)	689 (95%)	37 (5%)	1 (0%)	59	80
1	C	720/740 (97%)	679 (94%)	41 (6%)	0	100	100
1	D	684/740 (92%)	641 (94%)	37 (5%)	6 (1%)	25	39
All	All	2852/2960 (96%)	2689 (94%)	155 (5%)	8 (0%)	50	70

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	474	GLY
1	D	244	GLU
1	D	449	LEU
1	B	244	GLU
1	A	334	SER

### 5.3.2 Protein sidechains ⓘ

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	650/662 (98%)	618 (95%)	32 (5%)	35	57
1	B	656/662 (99%)	626 (95%)	30 (5%)	37	60
1	C	649/662 (98%)	617 (95%)	32 (5%)	35	57
1	D	618/662 (93%)	594 (96%)	24 (4%)	43	68
All	All	2573/2648 (97%)	2455 (95%)	118 (5%)	37	60

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	505	GLN
1	C	145	GLU
1	D	453	ARG
1	B	506	ASN
1	B	658	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:



Mol	Chain	Res	Type
1	B	455	GLN
1	C	123	GLN
1	D	455	GLN
1	B	505	GLN
1	B	710	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

16 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	$\# Z  > 2$
4	NAG	A	804	1,4	12,14,15	0.72	1 (8%)	15,19,21	1.03	0
4	NAG	A	805	4	12,14,15	0.55	0	15,19,21	0.84	0
4	NAG	A	806	1,4	12,14,15	0.56	0	15,19,21	1.18	2 (13%)
4	NAG	A	807	4	12,14,15	0.63	0	15,19,21	0.71	0
4	NAG	B	804	1,4	12,14,15	0.67	0	15,19,21	0.97	1 (6%)
4	NAG	B	805	4	12,14,15	0.70	0	15,19,21	1.43	2 (13%)
4	NAG	B	806	1,4	12,14,15	0.59	0	15,19,21	1.28	1 (6%)
4	NAG	B	807	4	12,14,15	0.58	0	15,19,21	1.62	1 (6%)
4	NAG	C	804	1,4	12,14,15	0.71	1 (8%)	15,19,21	1.52	2 (13%)
4	NAG	C	805	4	12,14,15	0.67	0	15,19,21	1.56	3 (20%)
4	NAG	C	806	1,4	12,14,15	0.71	0	15,19,21	1.32	2 (13%)
4	NAG	C	807	4	12,14,15	0.56	0	15,19,21	1.52	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	D	804	1,4	12,14,15	0.57	0	15,19,21	0.76	0
4	NAG	D	805	4	12,14,15	0.53	0	15,19,21	1.37	1 (6%)
4	NAG	D	806	1,4	12,14,15	0.53	0	15,19,21	0.61	0
4	NAG	D	807	4	12,14,15	0.56	0	15,19,21	1.04	1 (6%)

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	804	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	805	4	-	0/6/23/26	0/1/1/1
4	NAG	A	806	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	807	4	-	0/6/23/26	0/1/1/1
4	NAG	B	804	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	805	4	-	0/6/23/26	0/1/1/1
4	NAG	B	806	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	807	4	-	0/6/23/26	0/1/1/1
4	NAG	C	804	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	805	4	-	0/6/23/26	0/1/1/1
4	NAG	C	806	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	807	4	-	0/6/23/26	0/1/1/1
4	NAG	D	804	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	805	4	-	0/6/23/26	0/1/1/1
4	NAG	D	806	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	807	4	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	804	NAG	O5-C5	-2.09	1.41	1.45
4	A	804	NAG	O5-C5	-2.00	1.41	1.45

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	807	NAG	O5-C5-C6	4.89	112.11	106.98
4	C	805	NAG	O5-C5-C6	4.44	111.64	106.98
4	B	805	NAG	O5-C5-C6	4.30	111.49	106.98
4	C	807	NAG	O5-C5-C6	3.95	111.12	106.98
4	C	804	NAG	O5-C5-C6	3.55	110.71	106.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

20 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	RUM	A	800	-	26,26,26	2.95	6 (23%)	34,36,36	2.12	7 (20%)
3	NAG	A	801	1	12,14,15	0.60	0	15,19,21	0.97	1 (6%)
3	NAG	A	802	1	12,14,15	0.44	0	15,19,21	1.55	1 (6%)
3	NAG	A	803	1	12,14,15	0.67	0	15,19,21	1.18	1 (6%)
3	NAG	A	808	1	12,14,15	0.59	0	15,19,21	1.18	2 (13%)
2	RUM	B	800	-	26,26,26	2.80	6 (23%)	34,36,36	2.22	6 (17%)
3	NAG	B	801	1	12,14,15	0.63	0	15,19,21	1.27	3 (20%)
3	NAG	B	802	1	12,14,15	0.54	0	15,19,21	1.35	1 (6%)
3	NAG	B	803	1	12,14,15	0.59	0	15,19,21	1.02	1 (6%)
3	NAG	B	808	1	12,14,15	0.55	0	15,19,21	0.88	1 (6%)
2	RUM	C	800	-	26,26,26	2.85	6 (23%)	34,36,36	2.09	8 (23%)
3	NAG	C	801	1	12,14,15	0.66	0	15,19,21	1.63	3 (20%)
3	NAG	C	802	1	12,14,15	0.56	0	15,19,21	1.23	1 (6%)
3	NAG	C	803	1	12,14,15	0.64	0	15,19,21	0.68	0
3	NAG	C	808	1	12,14,15	0.67	0	15,19,21	1.29	2 (13%)
2	RUM	D	800	-	26,26,26	2.84	6 (23%)	34,36,36	1.85	6 (17%)
3	NAG	D	801	-	12,14,15	0.61	0	15,19,21	0.95	1 (6%)
3	NAG	D	802	1	12,14,15	0.51	0	15,19,21	1.06	2 (13%)
3	NAG	D	803	1	12,14,15	0.57	0	15,19,21	0.72	0
3	NAG	D	808	1	12,14,15	0.58	0	15,19,21	1.14	1 (6%)

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	RUM	A	800	-	-	0/10/20/20	0/3/3/3
3	NAG	A	801	1	-	0/6/23/26	0/1/1/1
3	NAG	A	802	1	-	0/6/23/26	0/1/1/1
3	NAG	A	803	1	-	0/6/23/26	0/1/1/1
3	NAG	A	808	1	-	0/6/23/26	0/1/1/1
2	RUM	B	800	-	-	0/10/20/20	0/3/3/3
3	NAG	B	801	1	-	0/6/23/26	0/1/1/1
3	NAG	B	802	1	-	1/6/23/26	0/1/1/1
3	NAG	B	803	1	-	0/6/23/26	0/1/1/1
3	NAG	B	808	1	-	0/6/23/26	0/1/1/1
2	RUM	C	800	-	-	0/10/20/20	0/3/3/3
3	NAG	C	801	1	-	0/6/23/26	0/1/1/1
3	NAG	C	802	1	-	0/6/23/26	0/1/1/1
3	NAG	C	803	1	-	0/6/23/26	0/1/1/1
3	NAG	C	808	1	-	0/6/23/26	0/1/1/1
2	RUM	D	800	-	-	0/10/20/20	0/3/3/3
3	NAG	D	801	-	-	0/6/23/26	0/1/1/1
3	NAG	D	802	1	-	0/6/23/26	0/1/1/1
3	NAG	D	803	1	-	0/6/23/26	0/1/1/1
3	NAG	D	808	1	-	0/6/23/26	0/1/1/1

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	800	RUM	C15-C16	-9.94	1.28	1.44
2	D	800	RUM	C15-C16	-9.78	1.28	1.44
2	A	800	RUM	C15-C16	-9.22	1.29	1.44
2	B	800	RUM	C15-C16	-8.96	1.30	1.44
2	A	800	RUM	C3-N2	6.70	1.44	1.37

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	800	RUM	C19-N18-C23	6.65	125.11	112.80
2	B	800	RUM	C19-N18-C23	6.37	124.59	112.80
2	A	800	RUM	C19-N18-C23	6.27	124.41	112.80
2	D	800	RUM	C19-N18-C23	6.23	124.34	112.80
2	C	800	RUM	C5-N4-C3	6.12	124.01	117.22

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	802	NAG	O7-C7-N2-C2

There are no ring outliers.

## 5.7 Other polymers ⓘ

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	724/740 (97%)	-0.03	15 (2%) 60 61	20, 44, 71, 104	1 (0%)
1	B	730/740 (98%)	-0.06	5 (0%) 84 86	21, 45, 68, 103	1 (0%)
1	C	724/740 (97%)	0.10	26 (3%) 41 41	18, 49, 75, 99	1 (0%)
1	D	690/740 (93%)	0.40	52 (7%) 14 13	26, 55, 79, 87	0
All	All	2868/2960 (96%)	0.10	98 (3%) 43 43	18, 47, 76, 104	3 (0%)

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	89	PHE	5.1
1	D	392	LYS	5.1
1	D	416	TYR	4.9
1	D	222	PHE	4.9
1	D	464	GLU	4.7

### 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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	805	14/15	0.29	4.66	66,68,68,69	0
4	NAG	A	807	14/15	0.17	0.68	73,74,75,75	0
4	NAG	B	804	14/15	0.15	0.49	54,57,61,64	0
4	NAG	A	804	14/15	0.12	0.42	52,55,58,59	0
4	NAG	C	804	14/15	0.14	0.14	59,63,66,67	0
4	NAG	A	806	14/15	0.12	-0.47	66,67,69,71	0
4	NAG	C	806	14/15	0.19	-0.52	68,70,71,73	0
4	NAG	D	806	14/15	0.15	-1.16	91,92,92,93	0
4	NAG	D	804	14/15	0.11	-1.43	56,57,58,59	0
4	NAG	B	806	14/15	0.09	-2.75	61,64,65,68	0
4	NAG	D	805	14/15	0.20	-	60,61,63,63	0
4	NAG	C	805	14/15	0.29	-	70,71,73,73	0
4	NAG	A	805	14/15	0.13	-	60,61,62,62	0
4	NAG	C	807	14/15	0.28	-	73,75,76,76	0
4	NAG	D	807	14/15	0.33	-	94,94,95,95	0
4	NAG	B	807	14/15	0.30	-	71,72,74,74	0

## 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	D	803	14/15	0.55	5.41	89,90,92,92	0
3	NAG	B	808	14/15	0.27	5.20	57,60,61,62	0
2	RUM	A	800	24/24	0.23	3.86	38,39,41,42	1
3	NAG	A	803	14/15	0.25	2.59	63,65,67,67	0
3	NAG	C	808	14/15	0.25	2.57	64,68,69,69	0
3	NAG	A	808	14/15	0.23	2.57	58,62,66,66	0
2	RUM	C	800	24/24	0.22	2.51	36,40,42,46	1
3	NAG	D	808	14/15	0.22	1.42	66,68,69,69	0
2	RUM	B	800	24/24	0.25	1.30	41,45,46,49	1
3	NAG	B	803	14/15	0.17	1.01	54,57,59,60	0
2	RUM	D	800	24/24	0.23	0.85	42,46,47,50	1
3	NAG	A	802	14/15	0.25	0.72	65,65,66,67	0
3	NAG	B	802	14/15	0.15	0.19	59,62,63,64	0
3	NAG	C	802	14/15	0.21	0.16	62,63,65,65	0
3	NAG	C	801	14/15	0.27	-0.23	98,98,99,99	0
3	NAG	C	803	14/15	0.13	-0.32	66,68,68,68	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	801	14/15	0.21	-0.42	71,73,74,74	0
3	NAG	B	801	14/15	0.13	-0.98	81,82,83,83	0
3	NAG	D	802	14/15	0.17	-1.09	72,73,73,73	0
3	NAG	D	801	14/15	0.34	-	92,93,93,93	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.