



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 01:14 AM GMT

PDB ID : 3G0B  
Title : Crystal structure of dipeptidyl peptidase IV in complex with TAK-322  
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 : 2.25 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>

---

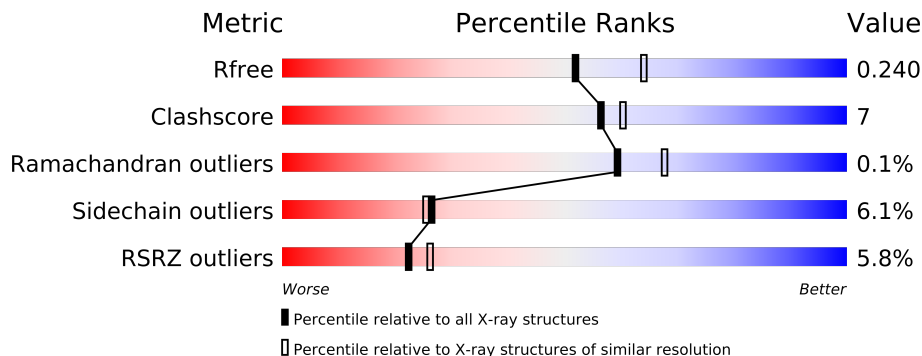
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.25 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	1108 (2.28-2.24)
Clashscore	79885	1326 (2.28-2.24)
Ramachandran outliers	78287	1291 (2.28-2.24)
Sidechain outliers	78261	1291 (2.28-2.24)
RSRZ outliers	66119	1110 (2.28-2.24)

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	NAG	A	804	-	X
2	NAG	A	811	-	X
2	NAG	A	812	-	X
2	NAG	B	808	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
2	NAG	C	807	-	X
2	NAG	D	802	-	X
2	NAG	D	803	-	X
2	NAG	D	807	-	X
2	NAG	D	808	-	X
4	T22	A	800	-	X
4	T22	B	800	-	X
4	T22	D	800	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 24992 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	727	Total	C	N	O	S	0	0	0
			5957	3824	981	1126	26			
1	B	733	Total	C	N	O	S	0	0	0
			6013	3857	997	1133	26			
1	C	726	Total	C	N	O	S	0	0	0
			5946	3818	977	1125	26			
1	D	727	Total	C	N	O	S	0	0	0
			5957	3824	981	1126	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 SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

Continued on next page...

*Continued from previous page...*

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

- 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
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	B	2	Total	C	N	O	0	0
			28	16	2	10		
3	B	2	Total	C	N	O	0	0
			28	16	2	10		
3	C	2	Total	C	N	O	0	0
			28	16	2	10		
3	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

*Continued on next page...*

*Continued from previous page...*

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
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

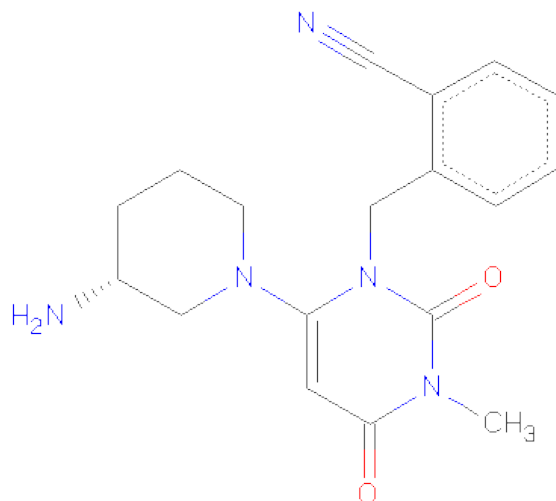
*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
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
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
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 4 is 2-({6-[(3R)-3-AMINOPIPERIDIN-1-YL]-3-METHYL-2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL}METHYL)BENZONITRILE (three-letter code: T22) (formula: C<sub>18</sub>H<sub>21</sub>N<sub>5</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			25	18	5	2		
4	B	1	Total	C	N	O	0	0
			25	18	5	2		
4	C	1	Total	C	N	O	0	0
			25	18	5	2		
4	D	1	Total	C	N	O	0	0
			25	18	5	2		

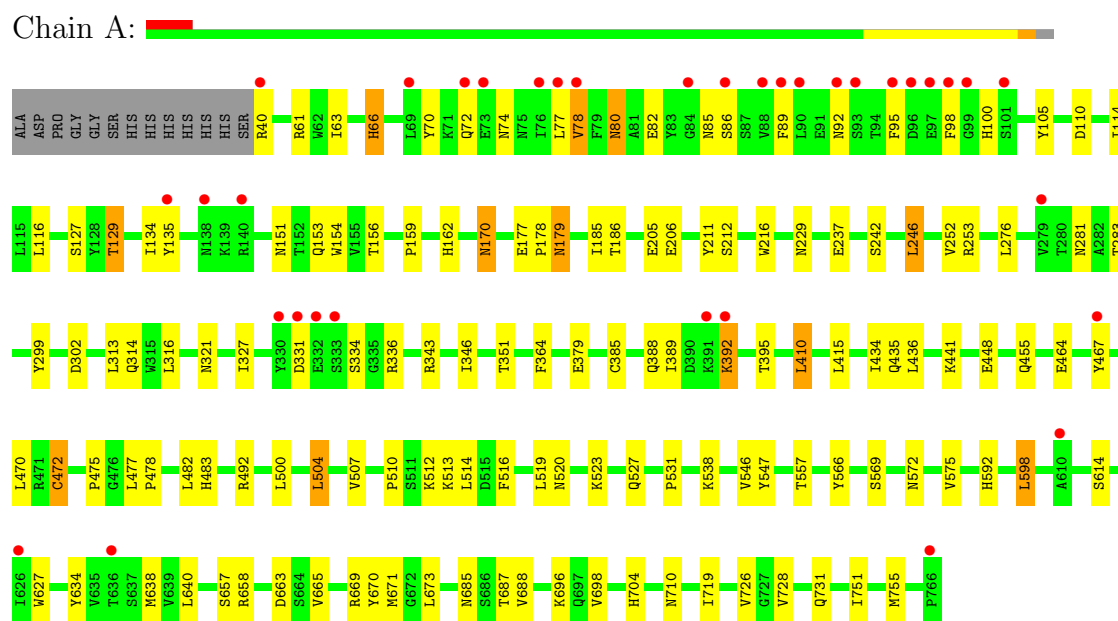
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	29	Total	O	0	0
			29	29		
5	B	509	Total	O	0	0
			509	509		
5	C	4	Total	O	0	0
			4	4		
5	D	15	Total	O	0	0
			15	15		

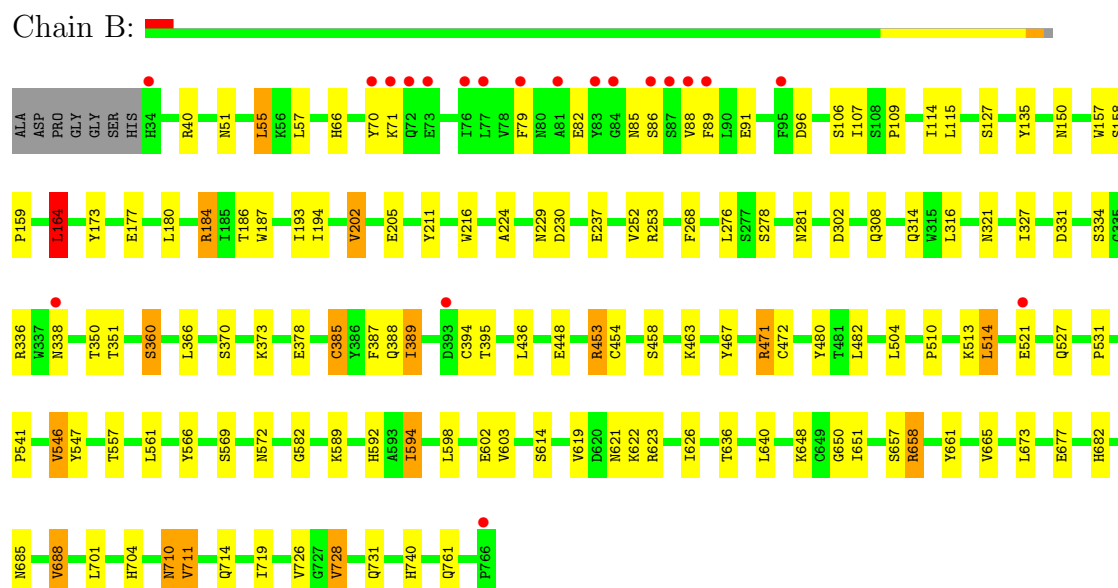
### 3 Residue-property plots

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

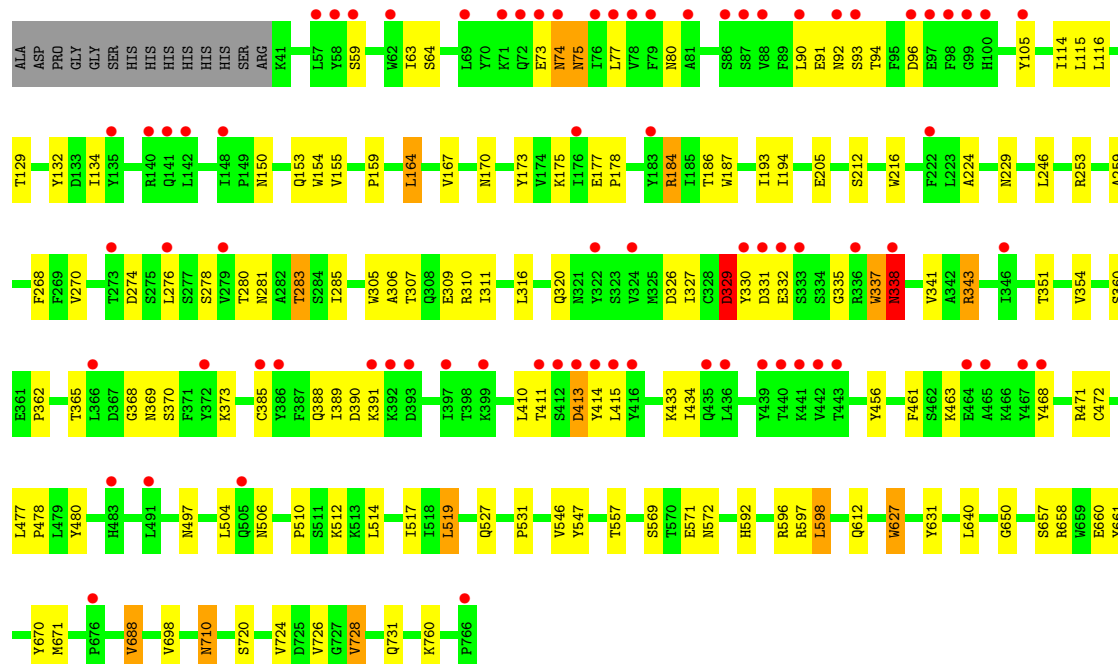


- Molecule 1: Dipeptidyl peptidase 4



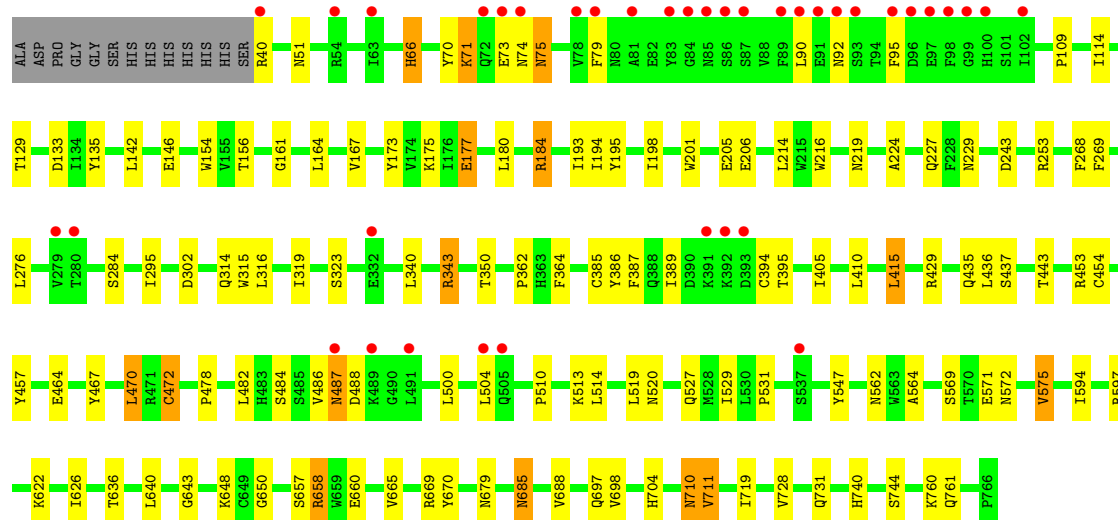
● Molecule 1: Dipeptidyl peptidase 4

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, $\alpha$ , $\beta$ , $\gamma$	121.69Å 122.40Å 144.01Å 90.00° 114.72° 90.00°	Depositor
Resolution (Å)	35.00 – 2.25 34.62 – 2.25	Depositor EDS
% Data completeness (in resolution range)	97.0 (35.00-2.25) 97.0 (34.62-2.25)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.207 , 0.242 0.205 , 0.240	Depositor DCC
$R_{free}$ test set	8857 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.1	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 53.1	EDS
Estimated twinning fraction	0.035 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 176863 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	24992	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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: T22, NAG

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	0/6129	0.63	0/8336
1	B	0.48	0/6190	0.63	2/8419 (0.0%)
1	C	0.82	23/6118 (0.4%)	0.67	9/8322 (0.1%)
1	D	0.48	0/6129	0.62	1/8336 (0.0%)
All	All	0.59	23/24566 (0.1%)	0.64	12/33413 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	338	ASN	CG-ND2	22.47	1.89	1.32
1	C	329	ASP	CG-OD1	17.61	1.65	1.25
1	C	343	ARG	CZ-NH2	17.20	1.55	1.33
1	C	177	GLU	CD-OE2	13.47	1.40	1.25
1	C	177	GLU	CD-OE1	11.77	1.38	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	329	ASP	CB-CG-OD1	-19.59	100.67	118.30
1	C	343	ARG	NE-CZ-NH1	-10.23	115.19	120.30
1	C	274	ASP	CB-CG-OD2	-7.27	111.75	118.30
1	C	329	ASP	OD1-CG-OD2	6.81	136.24	123.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	415	LEU	CA-CB-CG	6.75	130.83	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	338	ASN	Sidechain

## 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	5957	0	5679	84	0
1	B	6013	0	5720	73	0
1	C	5946	0	5666	84	0
1	D	5957	0	5676	83	0
2	A	56	0	52	4	0
2	B	56	0	52	5	0
2	C	56	0	52	3	0
2	D	70	0	65	4	0
3	A	112	0	100	6	0
3	B	56	0	50	2	0
3	C	28	0	25	2	0
3	D	28	0	25	5	0
4	A	25	0	21	1	0
4	B	25	0	21	1	0
4	C	25	0	21	1	0
4	D	25	0	21	1	0
5	A	29	0	0	0	0
5	B	509	0	0	1	0
5	C	4	0	0	0	0
5	D	15	0	0	0	0
All	All	24992	0	23246	312	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 7.

The worst 5 of 312 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:229:ASN:HD21	3:A:805:NAG:C1	1.12	1.58
1:C:329:ASP:CG	1:C:329:ASP:OD1	1.65	1.31
1:C:338:ASN:ND2	1:C:338:ASN:CG	1.89	1.26
1:A:229:ASN:ND2	3:A:805:NAG:C1	1.95	1.25
1:D:229:ASN:HD21	3:D:804:NAG:C1	1.69	1.04

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	725/740 (98%)	697 (96%)	28 (4%)	0	100	100
1	B	731/740 (99%)	704 (96%)	26 (4%)	1 (0%)	59	68
1	C	724/740 (98%)	690 (95%)	32 (4%)	2 (0%)	50	54
1	D	725/740 (98%)	692 (95%)	33 (5%)	0	100	100
All	All	2905/2960 (98%)	2783 (96%)	119 (4%)	3 (0%)	59	68

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	714	GLN
1	C	73	GLU
1	C	463	LYS

### 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	652/662 (98%)	613 (94%)	39 (6%)	27	26
1	B	658/662 (99%)	610 (93%)	48 (7%)	20	18
1	C	651/662 (98%)	617 (95%)	34 (5%)	32	34
1	D	652/662 (98%)	613 (94%)	39 (6%)	27	26
All	All	2613/2648 (99%)	2453 (94%)	160 (6%)	26	25

5 of 160 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	547	TYR
1	C	75	ASN
1	D	514	LEU
1	B	589	LYS
1	B	688	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	572	ASN
1	C	119	ASN
1	D	487	ASN
1	B	592	HIS
1	C	74	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
3	NAG	A	802	1,3	12,14,15	0.56	0	15,19,21	1.18	2 (13%)
3	NAG	A	803	3	12,14,15	0.67	0	15,19,21	1.14	1 (6%)
3	NAG	A	805	3	12,14,15	0.84	1 (8%)	15,19,21	0.90	0
3	NAG	A	806	3	12,14,15	0.51	0	15,19,21	1.14	2 (13%)
3	NAG	A	807	3	12,14,15	0.68	0	15,19,21	1.37	2 (13%)
3	NAG	A	808	3	12,14,15	0.64	0	15,19,21	1.01	1 (6%)
3	NAG	A	809	3	12,14,15	0.59	0	15,19,21	1.75	2 (13%)
3	NAG	A	810	3	12,14,15	0.63	0	15,19,21	1.31	1 (6%)
3	NAG	B	803	3	12,14,15	0.65	0	15,19,21	0.97	2 (13%)
3	NAG	B	804	3	12,14,15	0.59	0	15,19,21	0.72	0
3	NAG	B	805	3	12,14,15	0.79	0	15,19,21	0.97	1 (6%)
3	NAG	B	806	3	12,14,15	0.56	0	15,19,21	0.98	0
3	NAG	C	803	3	12,14,15	0.74	0	15,19,21	0.92	0
3	NAG	C	804	3	12,14,15	0.60	0	15,19,21	1.19	2 (13%)
3	NAG	D	804	3	12,14,15	0.74	0	15,19,21	1.61	2 (13%)
3	NAG	D	805	3	12,14,15	0.62	0	15,19,21	1.44	3 (20%)

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
3	NAG	A	802	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	803	3	-	0/6/23/26	0/1/1/1
3	NAG	A	805	3	-	0/6/23/26	0/1/1/1
3	NAG	A	806	3	-	0/6/23/26	0/1/1/1
3	NAG	A	807	3	-	0/6/23/26	0/1/1/1
3	NAG	A	808	3	-	0/6/23/26	0/1/1/1
3	NAG	A	809	3	-	0/6/23/26	0/1/1/1
3	NAG	A	810	3	-	0/6/23/26	0/1/1/1
3	NAG	B	803	3	-	0/6/23/26	0/1/1/1
3	NAG	B	804	3	-	0/6/23/26	0/1/1/1
3	NAG	B	805	3	-	0/6/23/26	0/1/1/1
3	NAG	B	806	3	-	0/6/23/26	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	803	3	-	0/6/23/26	0/1/1/1
3	NAG	C	804	3	-	0/6/23/26	0/1/1/1
3	NAG	D	804	3	-	0/6/23/26	0/1/1/1
3	NAG	D	805	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	805	NAG	O5-C5	-2.15	1.41	1.45

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	809	NAG	O5-C5-C4	4.76	116.69	110.65
3	D	804	NAG	C3-C4-C5	4.14	117.59	110.20
3	A	810	NAG	O5-C5-C6	3.77	110.94	106.98
3	A	809	NAG	C3-C4-C5	3.68	116.77	110.20
3	A	807	NAG	C3-C4-C5	3.32	116.13	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry

21 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	T22	A	800	-	27,27,27	1.91	3 (11%)	33,38,38	1.85	6 (18%)
2	NAG	A	801	-	12,14,15	0.61	0	15,19,21	0.64	0
2	NAG	A	804	-	12,14,15	0.57	0	15,19,21	1.66	2 (13%)
2	NAG	A	811	-	12,14,15	0.67	0	15,19,21	1.21	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	A	812	-	12,14,15	0.57	0	15,19,21	1.15	1 (6%)
4	T22	B	800	-	27,27,27	1.91	3 (11%)	33,38,38	1.85	6 (18%)
2	NAG	B	801	-	12,14,15	0.56	0	15,19,21	1.10	1 (6%)
2	NAG	B	802	-	12,14,15	0.59	0	15,19,21	1.16	1 (6%)
2	NAG	B	807	-	12,14,15	0.58	0	15,19,21	0.76	0
2	NAG	B	808	-	12,14,15	0.60	0	15,19,21	1.18	1 (6%)
4	T22	C	800	-	27,27,27	1.90	3 (11%)	33,38,38	1.84	6 (18%)
2	NAG	C	802	-	12,14,15	0.61	0	15,19,21	1.17	1 (6%)
2	NAG	C	805	-	12,14,15	0.61	0	15,19,21	0.75	0
2	NAG	C	806	1	12,14,15	0.58	0	15,19,21	1.04	2 (13%)
2	NAG	C	807	-	12,14,15	0.74	0	15,19,21	1.41	3 (20%)
4	T22	D	800	-	27,27,27	1.90	3 (11%)	33,38,38	1.84	6 (18%)
2	NAG	D	802	1	12,14,15	0.49	0	15,19,21	1.66	4 (26%)
2	NAG	D	803	-	12,14,15	0.63	0	15,19,21	1.39	2 (13%)
2	NAG	D	806	1	12,14,15	0.74	0	15,19,21	1.38	3 (20%)
2	NAG	D	807	1	12,14,15	0.61	0	15,19,21	0.87	1 (6%)
2	NAG	D	808	-	12,14,15	0.60	0	15,19,21	1.05	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	T22	A	800	-	-	0/10/20/20	0/3/3/3
2	NAG	A	801	-	-	0/6/23/26	0/1/1/1
2	NAG	A	804	-	-	0/6/23/26	0/1/1/1
2	NAG	A	811	-	-	1/6/23/26	0/1/1/1
2	NAG	A	812	-	-	1/6/23/26	0/1/1/1
4	T22	B	800	-	-	0/10/20/20	0/3/3/3
2	NAG	B	801	-	-	0/6/23/26	0/1/1/1
2	NAG	B	802	-	-	0/6/23/26	0/1/1/1
2	NAG	B	807	-	-	0/6/23/26	0/1/1/1
2	NAG	B	808	-	-	0/6/23/26	0/1/1/1
4	T22	C	800	-	-	0/10/20/20	0/3/3/3
2	NAG	C	802	-	-	0/6/23/26	0/1/1/1
2	NAG	C	805	-	-	0/6/23/26	0/1/1/1
2	NAG	C	806	1	-	0/6/23/26	0/1/1/1
2	NAG	C	807	-	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	T22	D	800	-	-	0/10/20/20	0/3/3/3
2	NAG	D	802	1	-	0/6/23/26	0/1/1/1
2	NAG	D	803	-	-	0/6/23/26	0/1/1/1
2	NAG	D	806	1	-	0/6/23/26	0/1/1/1
2	NAG	D	807	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	D	808	-	-	0/6/23/26	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)
4	B	800	T22	C22-C23	-8.32	1.31	1.44
4	A	800	T22	C22-C23	-8.31	1.31	1.44
4	D	800	T22	C22-C23	-8.31	1.31	1.44
4	C	800	T22	C22-C23	-8.30	1.31	1.44
4	D	800	T22	C6-N7	2.70	1.42	1.36

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	800	T22	N15-C6-N7	5.56	124.71	118.97
4	A	800	T22	N15-C6-N7	5.56	124.70	118.97
4	D	800	T22	N15-C6-N7	5.55	124.70	118.97
4	C	800	T22	N15-C6-N7	5.51	124.65	118.97
2	A	804	NAG	O5-C5-C6	4.94	112.16	106.98

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	807	NAG	C1

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	811	NAG	C1-C2-N2-C7
2	A	812	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	727/740 (98%)	0.15	35 (4%) 29 33	41, 50, 64, 77	0
1	B	733/740 (99%)	-0.00	20 (2%) 52 59	38, 50, 64, 77	0
1	C	726/740 (98%)	0.50	77 (10%) 7 8	44, 51, 73, 88	0
1	D	727/740 (98%)	0.17	38 (5%) 26 30	43, 50, 66, 79	0
All	All	2913/2960 (98%)	0.20	170 (5%) 22 26	38, 50, 67, 88	0

The worst 5 of 170 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	88	VAL	11.1
1	C	77	LEU	8.8
1	D	81	ALA	7.0
1	A	95	PHE	6.6
1	C	86	SER	6.0

### 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
3	NAG	D	805	14/15	0.34	27.40	80,85,87,87	0
3	NAG	C	804	14/15	0.20	5.35	64,65,66,67	0
3	NAG	B	803	14/15	0.28	5.31	116,118,119,119	0
3	NAG	B	804	14/15	0.41	4.34	117,119,119,119	0
3	NAG	A	806	14/15	0.23	3.59	71,71,74,74	0
3	NAG	B	806	14/15	0.27	3.58	81,85,86,87	0
3	NAG	A	809	14/15	0.23	3.29	99,102,103,105	0
3	NAG	B	805	14/15	0.17	1.94	69,76,79,82	0
3	NAG	A	808	14/15	0.23	1.82	102,103,103,103	0
3	NAG	A	802	14/15	0.24	1.63	73,78,81,85	0
3	NAG	A	805	14/15	0.17	1.42	61,66,68,69	0
3	NAG	D	804	14/15	0.13	0.91	74,77,79,81	0
3	NAG	A	807	14/15	0.21	0.70	97,101,102,102	0
3	NAG	C	803	14/15	0.12	-0.59	54,58,60,62	0
3	NAG	A	803	14/15	0.37	-	88,91,91,91	0
3	NAG	A	810	14/15	0.46	-	106,107,108,108	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
2	NAG	A	812	14/15	0.43	9.32	111,112,114,114	0
2	NAG	B	808	14/15	0.34	6.39	90,92,92,92	0
2	NAG	A	811	14/15	0.34	5.92	71,75,77,78	0
2	NAG	D	808	14/15	0.31	5.16	91,93,94,94	0
2	NAG	A	804	14/15	0.36	4.51	97,98,99,99	0
2	NAG	D	802	14/15	0.25	3.35	66,69,70,71	0
2	NAG	D	803	14/15	0.29	3.25	84,85,86,86	0
2	NAG	D	807	14/15	0.25	3.11	73,75,80,80	0
4	T22	D	800	25/25	0.22	3.04	48,50,51,52	0
2	NAG	C	807	14/15	0.35	2.46	89,91,92,92	0
4	T22	B	800	25/25	0.24	2.40	48,50,51,52	0
4	T22	A	800	25/25	0.25	2.24	48,50,51,52	0
4	T22	C	800	25/25	0.18	1.04	48,50,51,52	0
2	NAG	B	802	14/15	0.19	0.78	88,89,90,90	0
2	NAG	C	806	14/15	0.23	0.77	71,75,78,78	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	C	802	14/15	0.15	-0.41	88,90,90,90	0
2	NAG	C	805	14/15	0.14	-0.52	92,95,96,96	0
2	NAG	D	806	14/15	0.19	-0.68	70,74,75,76	0
2	NAG	A	801	14/15	0.14	-1.18	94,97,97,98	0
2	NAG	B	807	14/15	0.10	-1.28	93,94,95,96	0
2	NAG	B	801	14/15	0.17	-1.31	84,86,86,87	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.