



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 09:51 AM GMT

PDB ID : 4F5C  
Title : Crystal structure of the spike receptor binding domain of a porcine respiratory coronavirus in complex with the pig aminopeptidase N ectodomain  
Authors : Santiago, C.; Reguera, J.; Gaurav, M.; Ordono, D.; Enjuanes, L.; Casasnovas, J.M.  
Deposited on : 2012-05-13  
Resolution : 3.20 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>

---

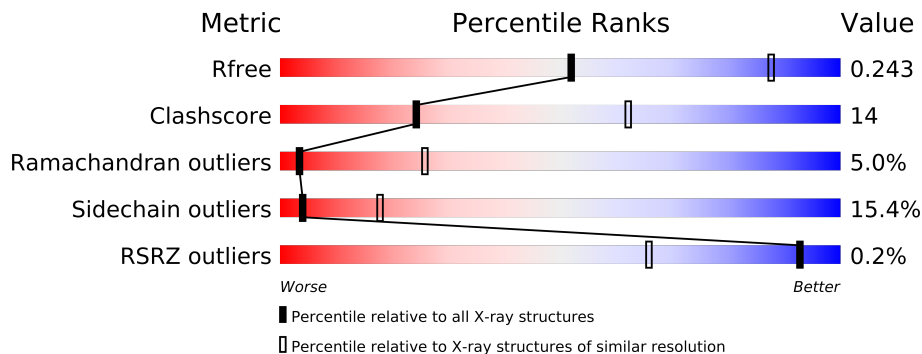
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.20 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))
$R_{free}$	66092	1824 (3.30-3.10)
Clashscore	79885	1078 (3.26-3.14)
Ramachandran outliers	78287	1059 (3.26-3.14)
Sidechain outliers	78261	1058 (3.26-3.14)
RSRZ outliers	66119	1825 (3.30-3.10)

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	959	
1	B	959	
2	E	440	
2	F	440	

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

Mol	Type	Chain	Res	Geometry	Electron density
4	NAG	A	1006	-	X
4	NAG	B	1009	-	X
4	NAG	B	1010	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 17067 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 Aminopeptidase N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	901	Total	C	N	O	S	0	0	0
			7245	4629	1208	1378	30			
1	B	900	Total	C	N	O	S	0	0	0
			7235	4623	1205	1377	30			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	TYR	-	EXPRESSION TAG	UNP P15145
A	20	PRO	-	EXPRESSION TAG	UNP P15145
A	21	TYR	-	EXPRESSION TAG	UNP P15145
A	22	ASP	-	EXPRESSION TAG	UNP P15145
A	23	VAL	-	EXPRESSION TAG	UNP P15145
A	24	PRO	-	EXPRESSION TAG	UNP P15145
A	25	ASP	-	EXPRESSION TAG	UNP P15145
A	26	TYR	-	EXPRESSION TAG	UNP P15145
A	27	ALA	-	EXPRESSION TAG	UNP P15145
A	28	GLY	-	EXPRESSION TAG	UNP P15145
A	29	ALA	-	EXPRESSION TAG	UNP P15145
A	30	GLN	-	EXPRESSION TAG	UNP P15145
A	31	PRO	-	EXPRESSION TAG	UNP P15145
A	32	ALA	-	EXPRESSION TAG	UNP P15145
A	33	ARG	-	EXPRESSION TAG	UNP P15145
A	34	SER	-	EXPRESSION TAG	UNP P15145
A	35	PRO	-	EXPRESSION TAG	UNP P15145
A	82	ASN	PHE	CONFLICT	UNP P15145
A	107	PHE	LEU	CONFLICT	UNP P15145
A	108	ILE	LEU	CONFLICT	UNP P15145
A	964	LEU	-	EXPRESSION TAG	UNP P15145
A	965	VAL	-	EXPRESSION TAG	UNP P15145
A	966	PRO	-	EXPRESSION TAG	UNP P15145
A	967	ARG	-	EXPRESSION TAG	UNP P15145
A	968	GLY	-	EXPRESSION TAG	UNP P15145

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	969	SER	-	EXPRESSION TAG	UNP P15145
A	970	ASP	-	EXPRESSION TAG	UNP P15145
A	971	TYR	-	EXPRESSION TAG	UNP P15145
A	972	LYS	-	EXPRESSION TAG	UNP P15145
A	973	ASP	-	EXPRESSION TAG	UNP P15145
A	974	ASP	-	EXPRESSION TAG	UNP P15145
A	975	ASP	-	EXPRESSION TAG	UNP P15145
A	976	ASP	-	EXPRESSION TAG	UNP P15145
A	977	LYS	-	EXPRESSION TAG	UNP P15145
B	19	TYR	-	EXPRESSION TAG	UNP P15145
B	20	PRO	-	EXPRESSION TAG	UNP P15145
B	21	TYR	-	EXPRESSION TAG	UNP P15145
B	22	ASP	-	EXPRESSION TAG	UNP P15145
B	23	VAL	-	EXPRESSION TAG	UNP P15145
B	24	PRO	-	EXPRESSION TAG	UNP P15145
B	25	ASP	-	EXPRESSION TAG	UNP P15145
B	26	TYR	-	EXPRESSION TAG	UNP P15145
B	27	ALA	-	EXPRESSION TAG	UNP P15145
B	28	GLY	-	EXPRESSION TAG	UNP P15145
B	29	ALA	-	EXPRESSION TAG	UNP P15145
B	30	GLN	-	EXPRESSION TAG	UNP P15145
B	31	PRO	-	EXPRESSION TAG	UNP P15145
B	32	ALA	-	EXPRESSION TAG	UNP P15145
B	33	ARG	-	EXPRESSION TAG	UNP P15145
B	34	SER	-	EXPRESSION TAG	UNP P15145
B	35	PRO	-	EXPRESSION TAG	UNP P15145
B	82	ASN	PHE	CONFLICT	UNP P15145
B	107	PHE	LEU	CONFLICT	UNP P15145
B	108	ILE	LEU	CONFLICT	UNP P15145
B	964	LEU	-	EXPRESSION TAG	UNP P15145
B	965	VAL	-	EXPRESSION TAG	UNP P15145
B	966	PRO	-	EXPRESSION TAG	UNP P15145
B	967	ARG	-	EXPRESSION TAG	UNP P15145
B	968	GLY	-	EXPRESSION TAG	UNP P15145
B	969	SER	-	EXPRESSION TAG	UNP P15145
B	970	ASP	-	EXPRESSION TAG	UNP P15145
B	971	TYR	-	EXPRESSION TAG	UNP P15145
B	972	LYS	-	EXPRESSION TAG	UNP P15145
B	973	ASP	-	EXPRESSION TAG	UNP P15145
B	974	ASP	-	EXPRESSION TAG	UNP P15145
B	975	ASP	-	EXPRESSION TAG	UNP P15145
B	976	ASP	-	EXPRESSION TAG	UNP P15145

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	977	LYS	-	EXPRESSION TAG	UNP P15145

- Molecule 2 is a protein called PRCV spike protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	146	Total	C	N	O	S	0	0	0
			1138	722	192	216	8			
2	F	139	Total	C	N	O	S	0	0	0
			1079	688	179	204	8			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	409	ASP	GLU	CONFLICT	UNP Q84852
E	427	LEU	-	EXPRESSION TAG	UNP Q84852
E	428	VAL	-	EXPRESSION TAG	UNP Q84852
E	429	PRO	-	EXPRESSION TAG	UNP Q84852
E	430	ARG	-	EXPRESSION TAG	UNP Q84852
E	431	GLY	-	EXPRESSION TAG	UNP Q84852
E	432	SER	-	EXPRESSION TAG	UNP Q84852
E	433	ASP	-	EXPRESSION TAG	UNP Q84852
E	434	TYR	-	EXPRESSION TAG	UNP Q84852
E	435	LYS	-	EXPRESSION TAG	UNP Q84852
E	436	ASP	-	EXPRESSION TAG	UNP Q84852
E	437	ASP	-	EXPRESSION TAG	UNP Q84852
E	438	ASP	-	EXPRESSION TAG	UNP Q84852
E	439	ASP	-	EXPRESSION TAG	UNP Q84852
E	440	LYS	-	EXPRESSION TAG	UNP Q84852
F	409	ASP	GLU	CONFLICT	UNP Q84852
F	427	LEU	-	EXPRESSION TAG	UNP Q84852
F	428	VAL	-	EXPRESSION TAG	UNP Q84852
F	429	PRO	-	EXPRESSION TAG	UNP Q84852
F	430	ARG	-	EXPRESSION TAG	UNP Q84852
F	431	GLY	-	EXPRESSION TAG	UNP Q84852
F	432	SER	-	EXPRESSION TAG	UNP Q84852
F	433	ASP	-	EXPRESSION TAG	UNP Q84852
F	434	TYR	-	EXPRESSION TAG	UNP Q84852
F	435	LYS	-	EXPRESSION TAG	UNP Q84852
F	436	ASP	-	EXPRESSION TAG	UNP Q84852
F	437	ASP	-	EXPRESSION TAG	UNP Q84852
F	438	ASP	-	EXPRESSION TAG	UNP Q84852
F	439	ASP	-	EXPRESSION TAG	UNP Q84852

*Continued on next page...*

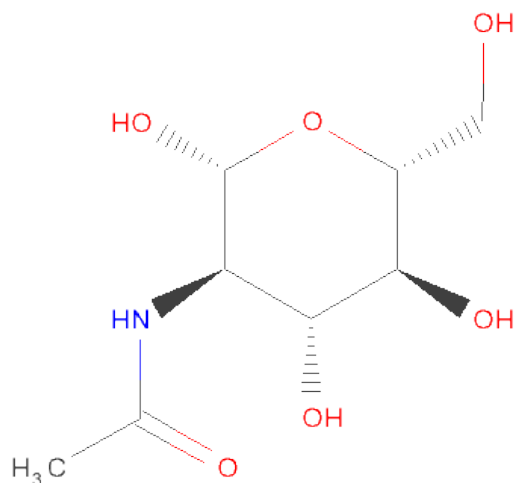
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	440	LYS	-	EXPRESSION TAG	UNP Q84852

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		

- 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	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	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		

Continued on next page...

*Continued from previous page...*

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	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		
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		
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	E	2	Total	C	N	O	0	0
			28	16	2	10		
5	F	2	Total	C	N	O	0	0
			28	16	2	10		

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	TYR	-	EXPRESSION TAG	UNP P15145

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	PRO	-	EXPRESSION TAG	UNP P15145
A	21	TYR	-	EXPRESSION TAG	UNP P15145
A	22	ASP	-	EXPRESSION TAG	UNP P15145
A	23	VAL	-	EXPRESSION TAG	UNP P15145
A	24	PRO	-	EXPRESSION TAG	UNP P15145
A	25	ASP	-	EXPRESSION TAG	UNP P15145
A	26	TYR	-	EXPRESSION TAG	UNP P15145
A	27	ALA	-	EXPRESSION TAG	UNP P15145
A	28	GLY	-	EXPRESSION TAG	UNP P15145
A	29	ALA	-	EXPRESSION TAG	UNP P15145
A	30	GLN	-	EXPRESSION TAG	UNP P15145
A	31	PRO	-	EXPRESSION TAG	UNP P15145
A	32	ALA	-	EXPRESSION TAG	UNP P15145
A	33	ARG	-	EXPRESSION TAG	UNP P15145
A	34	SER	-	EXPRESSION TAG	UNP P15145
A	35	PRO	-	EXPRESSION TAG	UNP P15145
A	82	ASN	PHE	CONFLICT	UNP P15145
A	107	PHE	LEU	CONFLICT	UNP P15145
A	108	ILE	LEU	CONFLICT	UNP P15145
A	964	LEU	-	EXPRESSION TAG	UNP P15145
A	965	VAL	-	EXPRESSION TAG	UNP P15145
A	966	PRO	-	EXPRESSION TAG	UNP P15145
A	967	ARG	-	EXPRESSION TAG	UNP P15145
A	968	GLY	-	EXPRESSION TAG	UNP P15145
A	969	SER	-	EXPRESSION TAG	UNP P15145
A	970	ASP	-	EXPRESSION TAG	UNP P15145
A	971	TYR	-	EXPRESSION TAG	UNP P15145
A	972	LYS	-	EXPRESSION TAG	UNP P15145
A	973	ASP	-	EXPRESSION TAG	UNP P15145
A	974	ASP	-	EXPRESSION TAG	UNP P15145
A	975	ASP	-	EXPRESSION TAG	UNP P15145
A	976	ASP	-	EXPRESSION TAG	UNP P15145
A	977	LYS	-	EXPRESSION TAG	UNP P15145
E	409	ASP	GLU	CONFLICT	UNP Q84852
E	427	LEU	-	EXPRESSION TAG	UNP Q84852
E	428	VAL	-	EXPRESSION TAG	UNP Q84852
E	429	PRO	-	EXPRESSION TAG	UNP Q84852
E	430	ARG	-	EXPRESSION TAG	UNP Q84852
E	431	GLY	-	EXPRESSION TAG	UNP Q84852
E	432	SER	-	EXPRESSION TAG	UNP Q84852
E	433	ASP	-	EXPRESSION TAG	UNP Q84852
E	434	TYR	-	EXPRESSION TAG	UNP Q84852

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	435	LYS	-	EXPRESSION TAG	UNP Q84852
E	436	ASP	-	EXPRESSION TAG	UNP Q84852
E	437	ASP	-	EXPRESSION TAG	UNP Q84852
E	438	ASP	-	EXPRESSION TAG	UNP Q84852
E	439	ASP	-	EXPRESSION TAG	UNP Q84852
E	440	LYS	-	EXPRESSION TAG	UNP Q84852
F	409	ASP	GLU	CONFLICT	UNP Q84852
F	427	LEU	-	EXPRESSION TAG	UNP Q84852
F	428	VAL	-	EXPRESSION TAG	UNP Q84852
F	429	PRO	-	EXPRESSION TAG	UNP Q84852
F	430	ARG	-	EXPRESSION TAG	UNP Q84852
F	431	GLY	-	EXPRESSION TAG	UNP Q84852
F	432	SER	-	EXPRESSION TAG	UNP Q84852
F	433	ASP	-	EXPRESSION TAG	UNP Q84852
F	434	TYR	-	EXPRESSION TAG	UNP Q84852
F	435	LYS	-	EXPRESSION TAG	UNP Q84852
F	436	ASP	-	EXPRESSION TAG	UNP Q84852
F	437	ASP	-	EXPRESSION TAG	UNP Q84852
F	438	ASP	-	EXPRESSION TAG	UNP Q84852
F	439	ASP	-	EXPRESSION TAG	UNP Q84852
F	440	LYS	-	EXPRESSION TAG	UNP Q84852

- Molecule 6 is water.

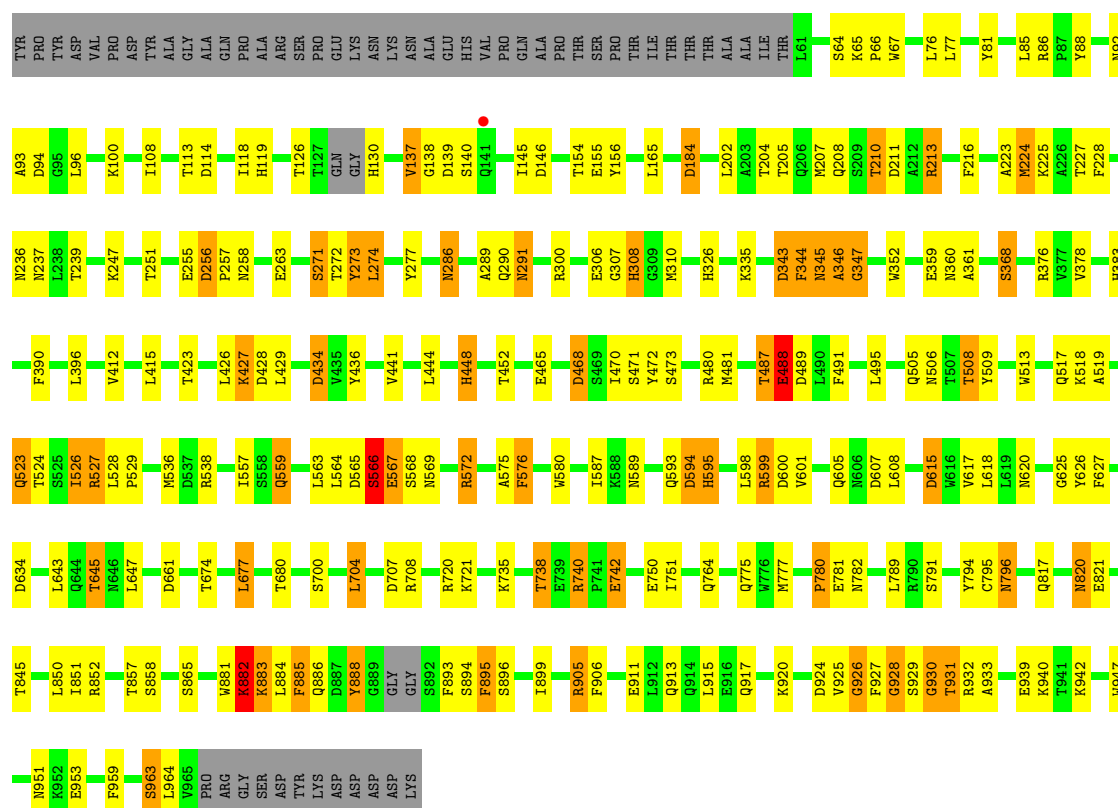
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total O 2 2	0	0
6	B	2	Total O 2 2	0	0

### 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.

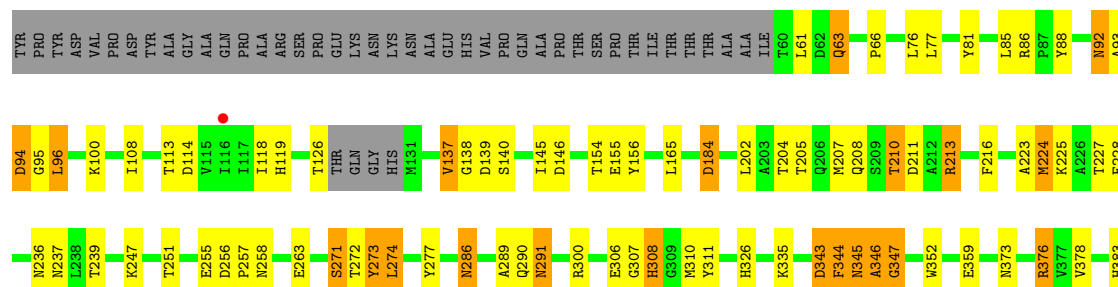
#### • Molecule 1: Aminopeptidase N

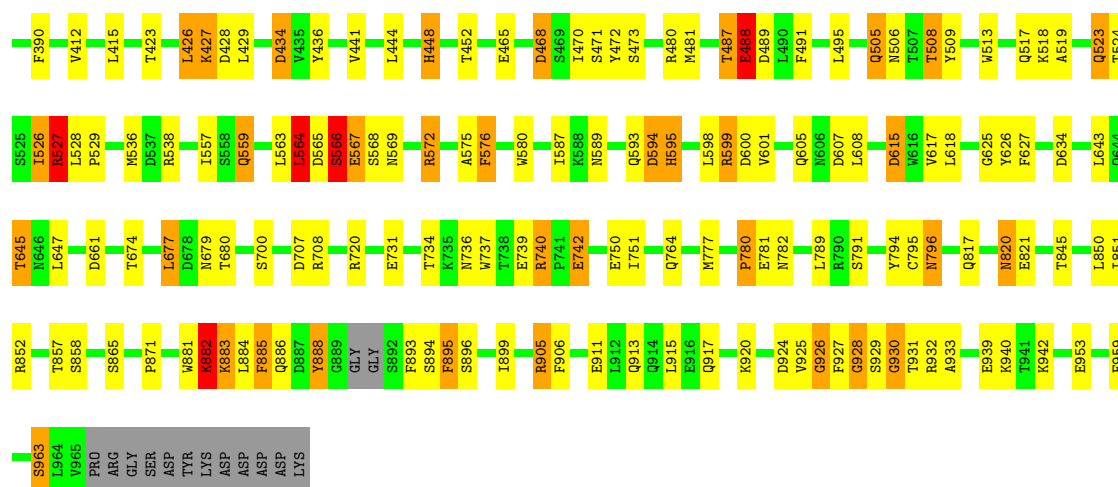
Chain A:



#### • Molecule 1: Aminopeptidase N

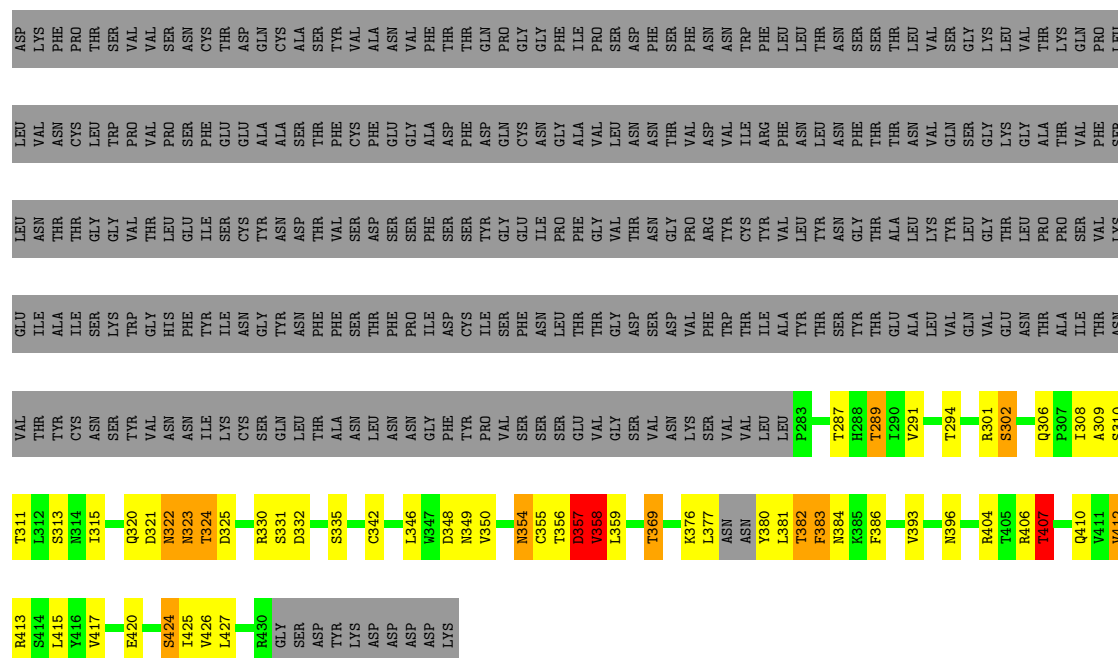
Chain B:





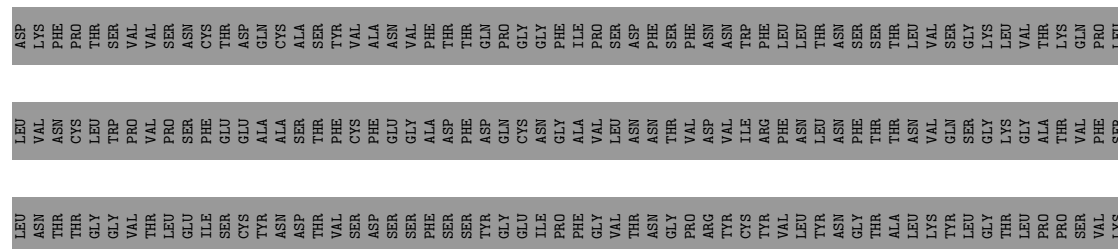
• Molecule 2: PRCV spike protein

Chain E:



• Molecule 2: PRCV spike protein

Chain F:



G394	A395	R404	R406	V412	R413	S414	L415	Y416	V417	I418	Y419	E420	S424	P429	ARG	GLY	SER	ASP	TYR	LYS	ASP	ASP	ASP	ASP	ASP	LYS	G305	Q306	P307	I308	T311	L312	S313	R314	L315	Q320	ASP	ASN	ASN	THR	D325	S331	S335	T341	C342	L346	N349	R353	N354	C355	T356	D357	V358	A363	T367	G368	T369	F374	D375	LYS	LEU	ASN	ASN	T380	L381	T382	F383	N384	K385	F386	L388	P392	F393	GLU	ILE	ALA	CYS	ASN	LYS	TRP	HIS	PHE	TYR	ILE	ASN	PHE	PHE	SER	THR	PHE	PRO	ASN	GLY	PHE	TYR	ILE	ASP	CYS	ILE	SER	PHE	ASN	LEU	THR	GLU	THR	THR	TRP	PHE	VAL	ASP	SER	GLY	THR	THR	GLU	THR	THR	TYR	ALA	ILE	GLU
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.86Å 87.94Å 176.91Å 90.00° 90.54° 90.00°	Depositor
Resolution (Å)	24.91 – 3.20 47.74 – 3.20	Depositor EDS
% Data completeness (in resolution range)	95.7 (24.91-3.20) 95.7 (47.74-3.20)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.75 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.201 , 0.245 0.201 , 0.243	Depositor DCC
$R_{free}$ test set	2737 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	85.3	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 55.0	EDS
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 54012 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	17067	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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: ZN, 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	0.25	0/7431	0.56	13/10126 (0.1%)
1	B	0.25	0/7420	0.59	13/10111 (0.1%)
2	E	0.26	0/1160	0.52	1/1576 (0.1%)
2	F	0.23	0/1100	0.42	0/1494
All	All	0.25	0/17111	0.56	27/23307 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
2	E	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	527	ARG	NE-CZ-NH1	15.69	128.15	120.30
1	B	527	ARG	NE-CZ-NH2	-12.90	113.85	120.30
1	A	538	ARG	NE-CZ-NH1	-12.67	113.97	120.30
1	B	538	ARG	NE-CZ-NH2	-12.31	114.14	120.30
1	B	720	ARG	NE-CZ-NH1	-12.11	114.25	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	208	GLN	Peptide
1	B	208	GLN	Peptide
2	E	357	ASP	Peptide

## 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	7245	0	0	96	0
1	B	7235	0	0	96	2
2	E	1138	0	0	27	2
2	F	1079	0	0	27	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	126	0	117	0	0
4	B	154	0	143	4	0
5	A	28	0	25	0	0
5	E	28	0	25	0	0
5	F	28	0	25	0	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
All	All	17067	0	335	245	2

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 14.

The worst 5 of 245 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:272:THR:O	1:A:274:LEU:N	1.87	1.07
1:B:272:THR:O	1:B:274:LEU:N	1.87	1.07
1:B:928:GLY:O	1:B:930:GLY:N	1.94	1.01
1:A:928:GLY:O	1:A:930:GLY:N	1.94	0.99
2:F:357:ASP:O	2:F:406:ARG:NH2	2.02	0.91

All (2) 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	Distance(Å)	Clash(Å)
1:B:527:ARG:NH1	2:E:396:ASN:OD1[3.545]	1.91	0.29
1:B:527:ARG:NH1	2:E:396:ASN:CG[3.545]	2.05	0.15

## 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	895/959 (93%)	766 (86%)	84 (9%)	45 (5%)	3	26
1	B	894/959 (93%)	768 (86%)	82 (9%)	44 (5%)	3	26
2	E	142/440 (32%)	111 (78%)	22 (16%)	9 (6%)	2	18
2	F	133/440 (30%)	107 (80%)	20 (15%)	6 (4%)	4	29
All	All	2064/2798 (74%)	1752 (85%)	208 (10%)	104 (5%)	3	26

5 of 104 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	146	ASP
1	A	155	GLU
1	A	273	TYR
1	A	290	GLN
1	A	344	PHE

### 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	799/845 (95%)	682 (85%)	117 (15%)	5	21
1	B	798/845 (94%)	681 (85%)	117 (15%)	4	21
2	E	132/393 (34%)	104 (79%)	28 (21%)	1	8

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	125/393 (32%)	102 (82%)	23 (18%)	2	11
All	All	1854/2476 (75%)	1569 (85%)	285 (15%)	4	18

5 of 285 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	139	ASP
1	B	426	LEU
2	F	289	THR
1	B	184	ASP
1	B	273	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

6 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	1008	1,5	12,14,15	0.79	1 (8%)	15,19,21	1.34	3 (20%)
5	NAG	A	1009	5	12,14,15	0.79	1 (8%)	15,19,21	1.15	2 (13%)
5	NAG	E	501	2,5	12,14,15	0.73	0	15,19,21	1.22	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	E	502	5	12,14,15	0.62	0	15,19,21	0.64	0
5	NAG	F	501	2,5	12,14,15	0.78	1 (8%)	15,19,21	1.12	1 (6%)
5	NAG	F	502	5	12,14,15	0.60	0	15,19,21	0.76	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	1008	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1009	5	-	0/6/23/26	0/1/1/1
5	NAG	E	501	2,5	-	0/6/23/26	0/1/1/1
5	NAG	E	502	5	-	0/6/23/26	0/1/1/1
5	NAG	F	501	2,5	-	0/6/23/26	0/1/1/1
5	NAG	F	502	5	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1008	NAG	O5-C5	-2.34	1.41	1.45
5	F	501	NAG	O5-C5	-2.15	1.41	1.45
5	A	1009	NAG	O5-C5	-2.10	1.41	1.45

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1008	NAG	O5-C5-C6	3.15	110.29	106.98
5	E	501	NAG	C3-C4-C5	2.92	115.41	110.20
5	A	1009	NAG	O5-C5-C4	-2.72	107.20	110.65
5	A	1008	NAG	O5-C5-C4	-2.52	107.45	110.65
5	F	501	NAG	C3-C4-C5	2.48	114.63	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

Of 22 ligands modelled in this entry, 2 are monoatomic - leaving 20 for Mogul analysis.

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	1002	1	12,14,15	0.68	0	15,19,21	1.32	2 (13%)
4	NAG	A	1003	1	12,14,15	0.54	0	15,19,21	0.99	1 (6%)
4	NAG	A	1004	1	12,14,15	0.66	0	15,19,21	0.68	0
4	NAG	A	1005	1	12,14,15	0.59	0	15,19,21	0.77	0
4	NAG	A	1006	1	12,14,15	0.64	0	15,19,21	0.76	1 (6%)
4	NAG	A	1007	1	12,14,15	0.73	1 (8%)	15,19,21	1.11	1 (6%)
4	NAG	A	1010	1	12,14,15	0.70	0	15,19,21	0.87	1 (6%)
4	NAG	A	1011	1	12,14,15	0.69	1 (8%)	15,19,21	0.87	0
4	NAG	A	1012	1	12,14,15	0.65	0	15,19,21	0.71	0
4	NAG	B	1002	1	12,14,15	0.59	0	15,19,21	1.11	1 (6%)
4	NAG	B	1003	1	12,14,15	0.62	0	15,19,21	1.01	2 (13%)
4	NAG	B	1004	1	12,14,15	0.66	0	15,19,21	1.10	1 (6%)
4	NAG	B	1005	1	12,14,15	0.82	1 (8%)	15,19,21	1.27	1 (6%)
4	NAG	B	1006	1	12,14,15	0.71	1 (8%)	15,19,21	0.64	0
4	NAG	B	1007	1	12,14,15	0.82	1 (8%)	15,19,21	1.12	2 (13%)
4	NAG	B	1008	1	12,14,15	0.58	0	15,19,21	0.99	1 (6%)
4	NAG	B	1009	1	12,14,15	0.72	1 (8%)	15,19,21	0.73	0
4	NAG	B	1010	1	12,14,15	0.56	0	15,19,21	0.99	1 (6%)
4	NAG	B	1011	1	12,14,15	0.65	0	15,19,21	0.85	0
4	NAG	B	1012	1	12,14,15	0.70	0	15,19,21	1.25	2 (13%)

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	1002	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1003	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1004	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1005	1	-	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	NAG	A	1006	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1007	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1010	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1011	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1012	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1002	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1003	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1004	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1005	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1006	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1007	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1008	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1009	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1010	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1011	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1012	1	-	0/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1007	NAG	O5-C5	-2.55	1.40	1.45
4	B	1005	NAG	O5-C5	-2.53	1.40	1.45
4	B	1009	NAG	O5-C5	-2.28	1.41	1.45
4	A	1007	NAG	O5-C5	-2.15	1.41	1.45
4	B	1006	NAG	O5-C5	-2.14	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	1005	NAG	O5-C5-C4	-3.98	105.60	110.65
4	A	1002	NAG	O5-C5-C6	3.74	110.91	106.98
4	B	1012	NAG	O5-C5-C6	3.21	110.35	106.98
4	B	1004	NAG	O5-C5-C6	3.13	110.26	106.98
4	A	1007	NAG	O5-C5-C6	2.96	110.09	106.98

There are no chirality outliers.

There are no torsion outliers.

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	901/959 (93%)	-0.10	1 (0%) 93 74	50, 84, 143, 186	0
1	B	900/959 (93%)	-0.11	1 (0%) 93 74	48, 86, 154, 203	0
2	E	146/440 (33%)	-0.00	0 100 100	66, 96, 148, 188	0
2	F	139/440 (31%)	0.27	3 (2%) 59 14	75, 127, 162, 185	0
All	All	2086/2798 (74%)	-0.07	5 (0%) 93 66	48, 88, 153, 203	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	419	TYR	2.8
1	A	141	GLN	2.3
1	B	116	ILE	2.3
2	F	353	ARG	2.2
2	F	388	LEU	2.0

### 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(Å <sup>2</sup> )	Q<0.9
5	NAG	F	502	14/15	0.36	2.45	144,150,156,159	0
5	NAG	E	501	14/15	0.32	0.44	88,117,142,149	0
5	NAG	F	501	14/15	0.27	-0.49	97,120,144,146	0
5	NAG	A	1008	14/15	0.18	-0.79	65,95,112,120	0
5	NAG	E	502	14/15	0.27	-1.89	126,140,146,149	0
5	NAG	A	1009	14/15	0.22	-	104,125,171,174	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(Å <sup>2</sup> )	Q<0.9
4	NAG	A	1006	14/15	0.46	9.08	138,158,186,192	0
4	NAG	B	1010	14/15	0.24	6.08	127,145,159,164	0
4	NAG	B	1009	14/15	0.38	4.81	125,143,149,155	0
4	NAG	B	1004	14/15	0.44	1.56	130,164,186,189	0
4	NAG	A	1004	14/15	0.30	1.44	125,149,160,160	0
4	NAG	A	1007	14/15	0.22	1.40	92,117,143,146	0
4	NAG	B	1002	14/15	0.34	1.33	107,137,154,158	0
4	NAG	B	1008	14/15	0.23	1.02	73,97,115,119	0
4	NAG	A	1012	14/15	0.27	0.82	58,80,99,100	0
4	NAG	A	1002	14/15	0.25	0.53	107,119,135,139	0
4	NAG	B	1006	14/15	0.27	0.44	117,129,135,140	0
4	NAG	B	1007	14/15	0.23	0.41	90,99,133,141	0
4	NAG	A	1010	14/15	0.17	0.38	106,121,131,137	0
4	NAG	B	1011	14/15	0.23	0.36	114,128,147,148	0
4	NAG	A	1005	14/15	0.21	0.11	135,142,162,163	0
4	NAG	B	1005	14/15	0.22	-0.23	123,148,157,159	0
3	ZN	B	1001	1/1	0.17	-0.31	83,83,83,83	0
4	NAG	A	1003	14/15	0.18	-0.44	51,79,92,96	0
4	NAG	A	1011	14/15	0.19	-0.57	84,95,117,134	0
4	NAG	B	1003	14/15	0.17	-0.57	75,88,102,106	0
3	ZN	A	1001	1/1	0.18	-0.66	75,75,75,75	0
4	NAG	B	1012	14/15	0.14	-1.31	74,87,113,124	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.