



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 31, 2016 – 09:05 PM EDT

PDB ID : 5AMA  
Title : Crystal structure of the Angiotensin-1 converting enzyme N-domain in complex with amyloid-beta 1-16  
Authors : Masuyer, G.; Larmuth, K.M.; Douglas, R.G.; Sturrock, E.D.; Acharya, K.R.  
Deposited on : 2015-03-10  
Resolution : 1.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

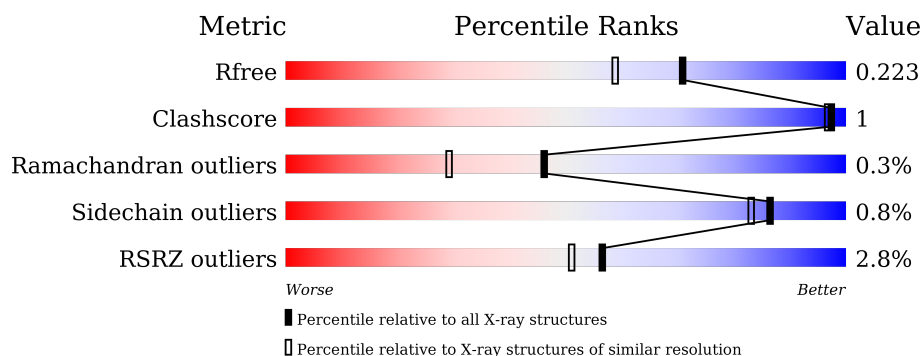
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	91344	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	629	<div> <div>2%</div> <div>93%</div> <div>• •</div> </div>
1	B	629	<div> <div>2%</div> <div>92%</div> <div>• •</div> </div>
1	C	629	<div> <div>%</div> <div>94%</div> <div>• •</div> </div>
1	D	629	<div> <div>5%</div> <div>91%</div> <div>5% • •</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ASP	A	907	-	-	-	X
2	ASP	B	907	-	-	-	X
2	ASP	D	907	-	-	-	X
6	NAG	A	1100	-	-	-	X
6	NAG	D	1102	-	-	-	X
7	NAG	B	1102	-	-	-	X
8	PEG	A	1201	-	-	-	X
8	PEG	A	1204	-	-	-	X
8	PEG	B	1201	-	-	-	X
8	PEG	B	1202	-	-	-	X
8	PEG	C	1201	-	-	-	X
9	P6G	B	1203	-	-	-	X
9	P6G	B	1204	-	-	-	X
9	P6G	C	1202	-	-	-	X
9	P6G	D	1201	-	-	-	X
9	P6G	D	1203	-	-	-	X

## 2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 22267 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called ANGIOTENSIN-CONVERTING ENZYME.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	603	Total	C	N	O	S	0	0	0
			4927	3165	846	897	19			
1	B	605	Total	C	N	O	S	0	1	0
			4939	3173	848	899	19			
1	C	609	Total	C	N	O	S	0	0	0
			4962	3189	852	902	19			
1	D	607	Total	C	N	O	S	0	1	0
			4941	3171	850	901	19			

There are 36 discrepancies between the modelled and reference sequences:

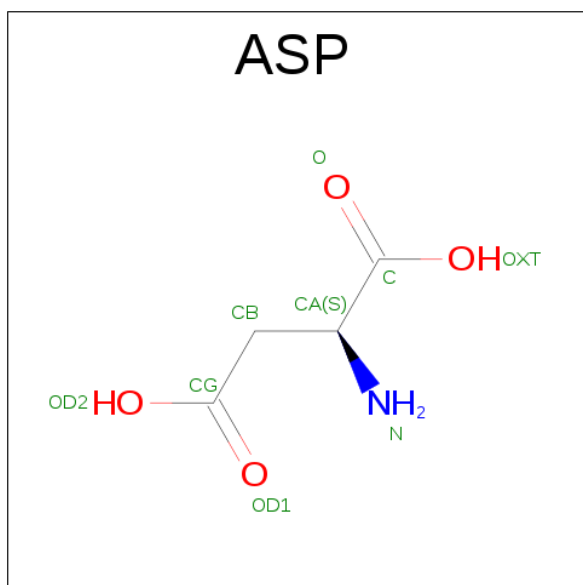
Chain	Residue	Modelled	Actual	Comment	Reference
A	9	GLN	ASN	ENGINEERED MUTATION	UNP P12821
A	25	GLN	ASN	ENGINEERED MUTATION	UNP P12821
A	82	GLN	ASN	ENGINEERED MUTATION	UNP P12821
A	117	GLN	ASN	ENGINEERED MUTATION	UNP P12821
A	131	GLN	ASN	ENGINEERED MUTATION	UNP P12821
A	289	GLN	ASN	ENGINEERED MUTATION	UNP P12821
A	545	ARG	GLN	ENGINEERED MUTATION	UNP P12821
A	576	LEU	PRO	ENGINEERED MUTATION	UNP P12821
A	629	LEU	ARG	ENGINEERED MUTATION	UNP P12821
B	629	LEU	ARG	ENGINEERED MUTATION	UNP P12821
B	9	GLN	ASN	ENGINEERED MUTATION	UNP P12821
B	25	GLN	ASN	ENGINEERED MUTATION	UNP P12821
B	82	GLN	ASN	ENGINEERED MUTATION	UNP P12821
B	117	GLN	ASN	ENGINEERED MUTATION	UNP P12821
B	131	GLN	ASN	ENGINEERED MUTATION	UNP P12821
B	289	GLN	ASN	ENGINEERED MUTATION	UNP P12821
B	545	ARG	GLN	ENGINEERED MUTATION	UNP P12821
B	576	LEU	PRO	ENGINEERED MUTATION	UNP P12821
C	629	LEU	ARG	ENGINEERED MUTATION	UNP P12821
C	9	GLN	ASN	ENGINEERED MUTATION	UNP P12821
C	25	GLN	ASN	ENGINEERED MUTATION	UNP P12821

*Continued on next page...*

Continued from previous page...

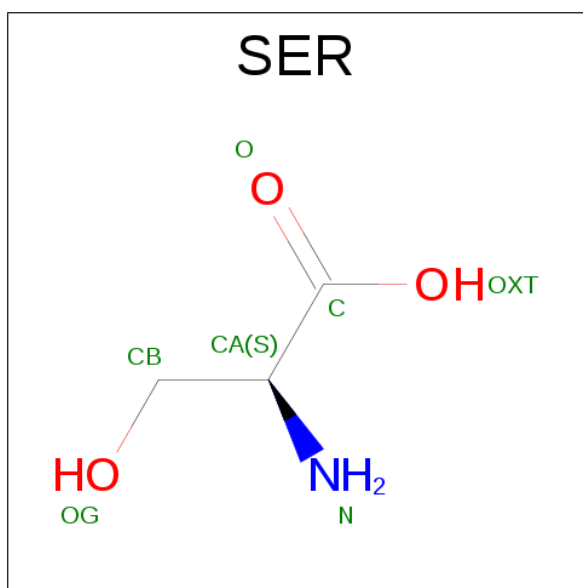
Chain	Residue	Modelled	Actual	Comment	Reference
C	82	GLN	ASN	ENGINEERED MUTATION	UNP P12821
C	117	GLN	ASN	ENGINEERED MUTATION	UNP P12821
C	131	GLN	ASN	ENGINEERED MUTATION	UNP P12821
C	289	GLN	ASN	ENGINEERED MUTATION	UNP P12821
C	545	ARG	GLN	ENGINEERED MUTATION	UNP P12821
C	576	LEU	PRO	ENGINEERED MUTATION	UNP P12821
D	629	LEU	ARG	ENGINEERED MUTATION	UNP P12821
D	9	GLN	ASN	ENGINEERED MUTATION	UNP P12821
D	25	GLN	ASN	ENGINEERED MUTATION	UNP P12821
D	82	GLN	ASN	ENGINEERED MUTATION	UNP P12821
D	117	GLN	ASN	ENGINEERED MUTATION	UNP P12821
D	131	GLN	ASN	ENGINEERED MUTATION	UNP P12821
D	289	GLN	ASN	ENGINEERED MUTATION	UNP P12821
D	545	ARG	GLN	ENGINEERED MUTATION	UNP P12821
D	576	LEU	PRO	ENGINEERED MUTATION	UNP P12821

- Molecule 2 is ASPARTIC ACID (three-letter code: ASP) (formula:  $C_4H_7NO_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			8	4	1	3		
2	B	1	Total	C	N	O	0	0
			8	4	1	3		
2	C	1	Total	C	N	O	0	0
			8	4	1	3		
2	D	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 3 is SERINE (three-letter code: SER) (formula:  $C_3H_7NO_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			7	3	1	3		
3	B	1	Total	C	N	O	0	0
			7	3	1	3		
3	C	1	Total	C	N	O	0	0
			7	3	1	3		
3	D	1	Total	C	N	O	0	0
			7	3	1	3		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	A	1	Total	Zn	0	0
			1	1		
4	D	1	Total	Zn	0	0
			1	1		
4	C	1	Total	Zn	0	0
			1	1		

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

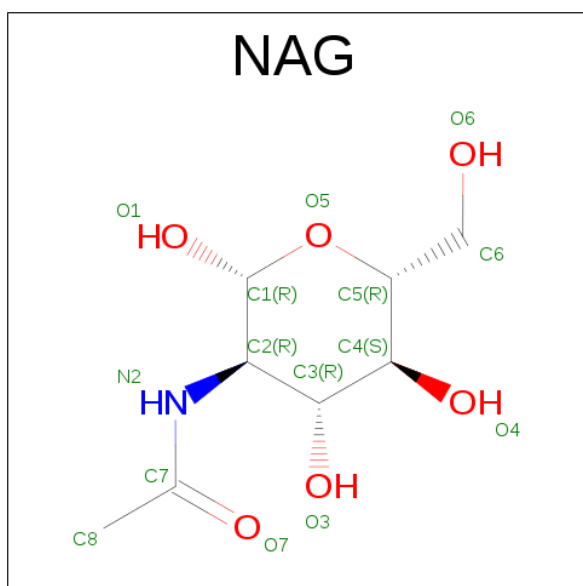
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Cl	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Cl	0	0
			1	1		
5	D	1	Total	Cl	0	0
			1	1		
5	C	1	Total	Cl	0	0
			1	1		

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

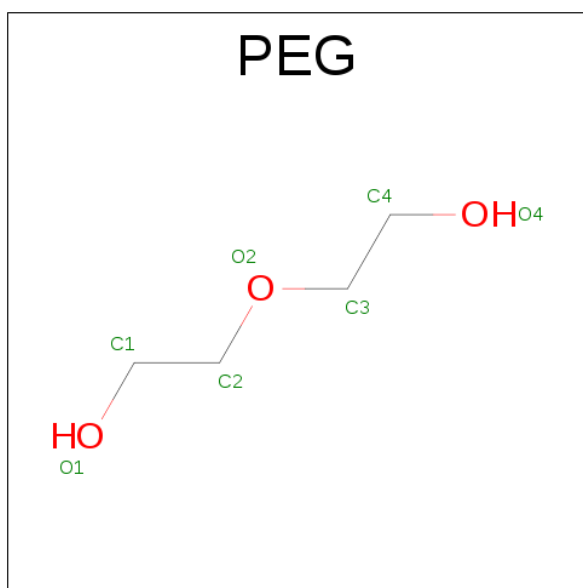


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is a polymer of unknown type called SUGAR (2-MER).

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

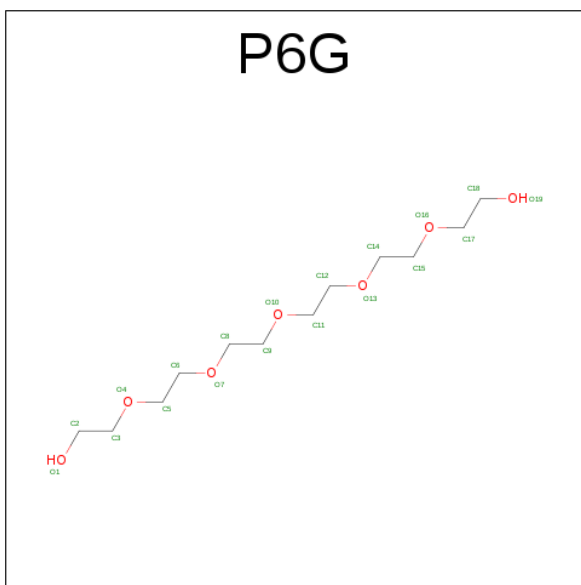
- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			7	4	3		
8	A	1	Total	C	O	0	0
			7	4	3		
8	A	1	Total	C	O	0	0
			7	4	3		
8	B	1	Total	C	O	0	0
			7	4	3		
8	B	1	Total	C	O	0	0
			7	4	3		
8	B	1	Total	C	O	0	0
			7	4	3		
8	C	1	Total	C	O	0	0
			7	4	3		
8	C	1	Total	C	O	0	0
			7	4	3		
8	D	1	Total	C	O	0	0
			7	4	3		
8	D	1	Total	C	O	0	0
			7	4	3		
8	D	1	Total	C	O	0	0
			7	4	3		

- Molecule 9 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula:  $C_{12}H_{26}O_7$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			19	12	7		
9	B	1	Total	C	O	0	0
			19	12	7		
9	B	1	Total	C	O	0	0
			19	12	7		
9	C	1	Total	C	O	0	0
			19	12	7		
9	D	1	Total	C	O	0	0
			19	12	7		
9	D	1	Total	C	O	0	0
			19	12	7		

- Molecule 10 is a polymer of unknown type called SUGAR (2-MER).

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

- Molecule 11 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	B	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 12 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
12	C	4	Total	C	N	O	0	0
			49	28	2	19		

- Molecule 13 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
13	D	4	Total	C	N	O	0	0
			49	28	2	19		

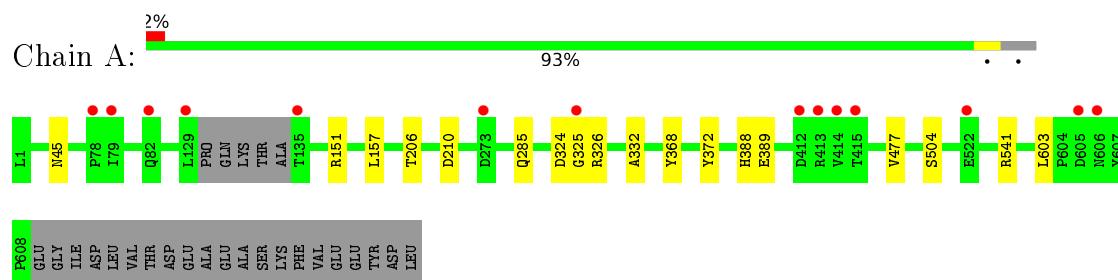
- Molecule 14 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	424	Total	O	0	0
			424	424		
14	B	482	Total	O	0	0
			482	482		
14	C	515	Total	O	0	0
			515	515		
14	D	467	Total	O	0	0
			467	467		
14	A	1	Total	O	0	0
			1	1		
14	B	1	Total	O	0	0
			1	1		

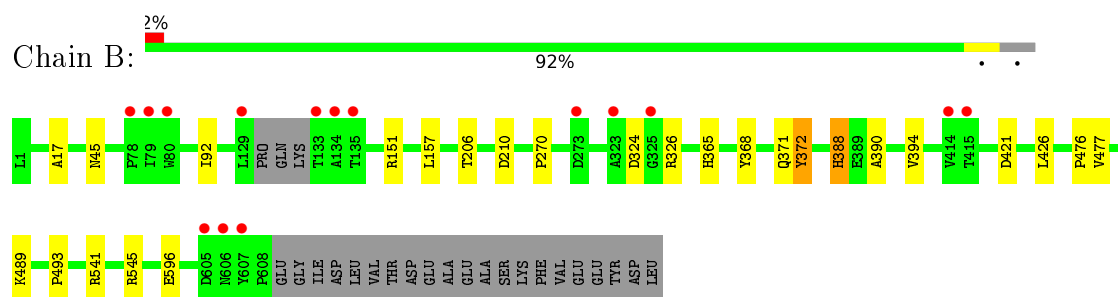
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

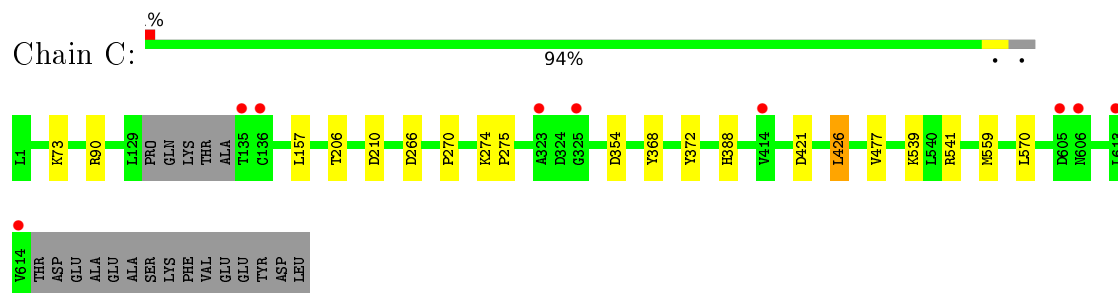
#### • Molecule 1: ANGIOTENSIN-CONVERTING ENZYME



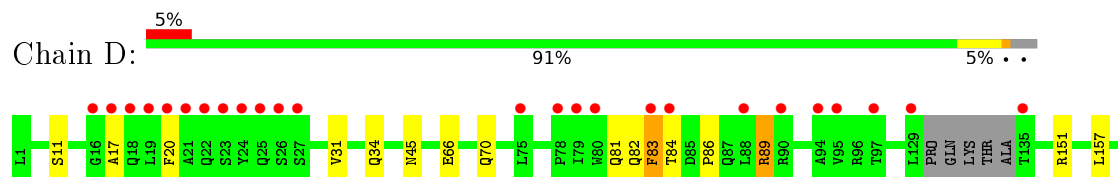
#### • Molecule 1: ANGIOTENSIN-CONVERTING ENZYME

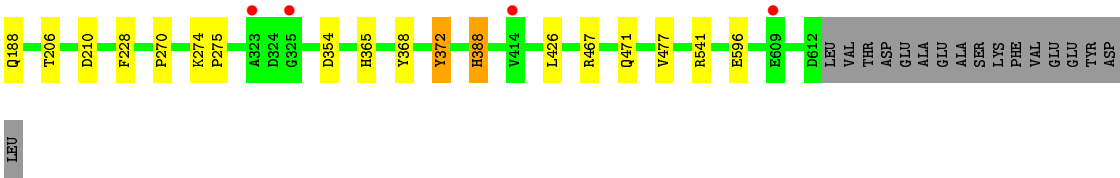


#### • Molecule 1: ANGIOTENSIN-CONVERTING ENZYME



#### • Molecule 1: ANGIOTENSIN-CONVERTING ENZYME





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.34Å 101.73Å 114.14Å 85.09° 85.62° 81.30°	Depositor
Resolution (Å)	113.48 – 1.80 33.27 – 1.80	Depositor EDS
% Data completeness (in resolution range)	93.5 (113.48-1.80) 91.9 (33.27-1.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.187 , 0.218 0.194 , 0.223	Depositor DCC
$R_{free}$ test set	13921 reflections (5.21%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.8	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 44.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 281239 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	22267	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.77 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.5050e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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, BMA, NAG, CL, FUC, P6G, FUL, PEG, MAN

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	0.48	0/5082	0.68	4/6922 (0.1%)
1	B	0.51	0/5097	0.68	2/6943 (0.0%)
1	C	0.51	0/5117	0.70	5/6970 (0.1%)
1	D	0.52	0/5096	0.69	1/6940 (0.0%)
All	All	0.51	0/20392	0.69	12/27775 (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	A	0	1
1	D	0	2
All	All	0	3

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	541	ARG	NE-CZ-NH1	6.93	123.76	120.30
1	C	541	ARG	NE-CZ-NH1	6.39	123.50	120.30
1	B	541	ARG	NE-CZ-NH1	6.11	123.36	120.30
1	D	541	ARG	NE-CZ-NH1	5.79	123.19	120.30
1	A	326	ARG	NE-CZ-NH1	5.59	123.10	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	325	GLY	Peptide
1	D	82	GLN	Peptide
1	D	83	PHE	Peptide

## 5.2 Too-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 hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4927	0	4703	5	0
1	B	4939	0	4712	13	0
1	C	4962	0	4736	9	0
1	D	4941	0	4718	19	0
2	A	8	0	3	1	0
2	B	8	0	3	0	0
2	C	8	0	3	0	0
2	D	8	0	3	0	0
3	A	7	0	5	0	0
3	B	7	0	5	0	0
3	C	7	0	5	0	0
3	D	7	0	5	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	14	0	13	0	0
6	D	14	0	13	0	0
7	A	56	0	50	0	0
7	B	28	0	25	0	0
7	C	28	0	25	0	0
8	A	21	0	30	1	0
8	B	21	0	30	0	0
8	C	14	0	20	0	0
8	D	21	0	30	0	0
9	A	19	0	26	1	0
9	B	38	0	52	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	C	19	0	26	0	0
9	D	38	0	52	1	0
10	B	24	0	22	1	0
10	C	24	0	22	0	0
10	D	24	0	22	1	0
11	B	39	0	34	0	0
12	C	49	0	43	0	0
13	D	49	0	43	0	0
14	A	425	0	0	0	0
14	B	483	0	0	0	0
14	C	515	0	0	2	0
14	D	467	0	0	3	0
All	All	22267	0	19479	47	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

The worst 5 of 47 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:365:HIS:HD1	1:D:388:HIS:HD2	1.30	0.79
1:B:270:PRO:HD3	1:B:426:LEU:HD22	1.70	0.74
1:D:365:HIS:HD1	1:D:388:HIS:CD2	2.06	0.73
1:D:151:ARG:NH1	14:D:2138:HOH:O	2.16	0.72
1:D:467:ARG:HH11	1:D:471:GLN:HE22	1.37	0.72

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	599/629 (95%)	588 (98%)	9 (2%)	2 (0%)	46	29
1	B	602/629 (96%)	593 (98%)	8 (1%)	1 (0%)	52	35
1	C	605/629 (96%)	595 (98%)	10 (2%)	0	100	100
1	D	604/629 (96%)	590 (98%)	10 (2%)	4 (1%)	26	11
All	All	2410/2516 (96%)	2366 (98%)	37 (2%)	7 (0%)	46	29

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	324	ASP
1	D	83	PHE
1	D	81	GLN
1	D	84	THR
1	D	45	ASN

### 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	518/541 (96%)	514 (99%)	4 (1%)	86	83
1	B	518/541 (96%)	514 (99%)	4 (1%)	86	83
1	C	520/541 (96%)	516 (99%)	4 (1%)	86	83
1	D	518/541 (96%)	513 (99%)	5 (1%)	82	77
All	All	2074/2164 (96%)	2057 (99%)	17 (1%)	86	83

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	421	ASP
1	C	368	TYR
1	D	89	ARG
1	B	388	HIS
1	D	368	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	87	GLN
1	D	203	ASN
1	D	388	HIS
1	D	9	GLN
1	D	371	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

25 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	$\# Z  > 2$
7	NAG	A	1101	1,7	14,14,15	0.41	0	15,19,21	0.98	1 (6%)
7	NAG	A	1102	7	14,14,15	0.49	0	15,19,21	0.88	0
7	NAG	A	1103	1,7	14,14,15	0.33	0	15,19,21	0.72	0
7	NAG	A	1104	7	14,14,15	0.44	0	15,19,21	0.93	1 (6%)
10	NAG	B	1100	1,10	14,14,15	0.71	0	15,19,21	1.44	3 (20%)
10	FUC	B	1101	10	10,10,11	0.63	0	13,14,16	1.60	2 (15%)
7	NAG	B	1102	1,7	14,14,15	0.47	0	15,19,21	0.77	0
7	NAG	B	1103	7	14,14,15	0.51	0	15,19,21	1.04	0
11	NAG	B	1104	1,11	14,14,15	0.39	0	15,19,21	0.96	1 (6%)
11	NAG	B	1105	11	14,14,15	0.31	0	15,19,21	0.63	0
11	BMA	B	1106	11	11,11,12	0.38	0	15,15,17	1.01	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
10	NAG	C	1100	1,10	14,14,15	0.67	0	15,19,21	1.45	2 (13%)
10	FUC	C	1101	10	10,10,11	0.70	0	13,14,16	2.15	3 (23%)
7	NAG	C	1102	1,7	14,14,15	0.47	0	15,19,21	0.79	1 (6%)
7	NAG	C	1103	7	14,14,15	0.33	0	15,19,21	0.95	0
12	NAG	C	1104	1,12	14,14,15	0.62	0	15,19,21	0.94	0
12	NAG	C	1105	12	14,14,15	0.41	0	15,19,21	0.99	1 (6%)
12	BMA	C	1106	12	11,11,12	0.39	0	15,15,17	0.68	0
12	FUL	C	1107	12	10,10,11	0.69	0	13,14,16	1.07	0
10	NAG	D	1100	1,10	14,14,15	0.68	0	15,19,21	1.64	5 (33%)
10	FUC	D	1101	10	10,10,11	0.60	0	13,14,16	1.21	1 (7%)
13	NAG	D	1103	1,13	14,14,15	0.44	0	15,19,21	0.79	0
13	NAG	D	1104	13	14,14,15	0.47	0	15,19,21	0.73	0
13	MAN	D	1105	13	11,11,12	0.56	0	15,15,17	1.53	4 (26%)
13	FUC	D	1106	13	10,10,11	0.50	0	13,14,16	1.28	2 (15%)

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
7	NAG	A	1101	1,7	-	0/6/23/26	0/1/1/1
7	NAG	A	1102	7	-	0/6/23/26	0/1/1/1
7	NAG	A	1103	1,7	-	0/6/23/26	0/1/1/1
7	NAG	A	1104	7	-	0/6/23/26	0/1/1/1
10	NAG	B	1100	1,10	-	0/6/23/26	0/1/1/1
10	FUC	B	1101	10	-	0/0/17/20	0/1/1/1
7	NAG	B	1102	1,7	-	0/6/23/26	0/1/1/1
7	NAG	B	1103	7	-	0/6/23/26	0/1/1/1
11	NAG	B	1104	1,11	-	0/6/23/26	0/1/1/1
11	NAG	B	1105	11	-	0/6/23/26	0/1/1/1
11	BMA	B	1106	11	-	0/2/19/22	0/1/1/1
10	NAG	C	1100	1,10	-	0/6/23/26	0/1/1/1
10	FUC	C	1101	10	-	0/0/17/20	0/1/1/1
7	NAG	C	1102	1,7	-	0/6/23/26	0/1/1/1
7	NAG	C	1103	7	-	0/6/23/26	0/1/1/1
12	NAG	C	1104	1,12	-	0/6/23/26	0/1/1/1
12	NAG	C	1105	12	-	0/6/23/26	0/1/1/1
12	BMA	C	1106	12	-	0/2/19/22	0/1/1/1
12	FUL	C	1107	12	-	0/0/17/20	0/1/1/1
10	NAG	D	1100	1,10	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	FUC	D	1101	10	-	0/0/17/20	0/1/1/1
13	NAG	D	1103	1,13	-	0/6/23/26	0/1/1/1
13	NAG	D	1104	13	-	0/6/23/26	0/1/1/1
13	MAN	D	1105	13	-	0/2/19/22	0/1/1/1
13	FUC	D	1106	13	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	B	1101	FUC	O5-C1-C2	-4.24	104.11	110.89
10	D	1100	NAG	C2-N2-C7	-3.47	118.59	123.11
11	B	1104	NAG	O4-C4-C3	-2.44	104.86	110.36
10	B	1100	NAG	O7-C7-N2	-2.30	117.16	121.84
10	D	1100	NAG	O7-C7-N2	-2.24	117.27	121.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	B	1100	NAG	1	0
10	D	1100	NAG	1	0

## 5.6 Ligand geometry

Of 35 ligands modelled in this entry, 8 are monoatomic - leaving 27 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	$\# Z  > 2$
6	NAG	A	1100	1	14,14,15	0.42	0	15,19,21	1.01	1 (6%)
8	PEG	A	1200	-	6,6,6	0.45	0	5,5,5	0.44	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	PEG	A	1201	-	6,6,6	0.47	0	5,5,5	0.41	0
9	P6G	A	1202	-	18,18,18	0.47	0	17,17,17	0.44	0
8	PEG	A	1204	-	6,6,6	0.42	0	5,5,5	0.43	0
2	ASP	A	907	3	2,7,8	0.58	0	2,8,10	1.80	1 (50%)
3	SER	A	908	2	3,6,6	0.28	0	2,7,7	0.91	0
8	PEG	B	1200	-	6,6,6	0.45	0	5,5,5	0.41	0
8	PEG	B	1201	-	6,6,6	0.54	0	5,5,5	0.37	0
8	PEG	B	1202	-	6,6,6	0.48	0	5,5,5	0.31	0
9	P6G	B	1203	-	18,18,18	0.69	0	17,17,17	0.91	0
9	P6G	B	1204	-	18,18,18	0.56	0	17,17,17	0.52	0
2	ASP	B	907	3	2,7,8	0.66	0	2,8,10	1.60	1 (50%)
3	SER	B	908	2	3,6,6	0.18	0	2,7,7	0.57	0
8	PEG	C	1200	-	6,6,6	0.44	0	5,5,5	0.32	0
8	PEG	C	1201	-	6,6,6	0.50	0	5,5,5	0.45	0
9	P6G	C	1202	-	18,18,18	0.54	0	17,17,17	0.41	0
2	ASP	C	907	3	2,7,8	0.26	0	2,8,10	1.53	1 (50%)
3	SER	C	908	2	3,6,6	0.36	0	2,7,7	0.58	0
6	NAG	D	1102	1	14,14,15	0.52	0	15,19,21	0.99	0
8	PEG	D	1200	-	6,6,6	0.42	0	5,5,5	0.37	0
9	P6G	D	1201	-	18,18,18	0.62	0	17,17,17	0.95	1 (5%)
8	PEG	D	1202	-	6,6,6	0.49	0	5,5,5	0.37	0
9	P6G	D	1203	-	18,18,18	0.54	0	17,17,17	0.36	0
8	PEG	D	1205	-	6,6,6	0.53	0	5,5,5	0.21	0
2	ASP	D	907	3	2,7,8	0.49	0	2,8,10	1.65	1 (50%)
3	SER	D	908	2	3,6,6	0.47	0	2,7,7	0.90	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
6	NAG	A	1100	1	-	0/6/23/26	0/1/1/1
8	PEG	A	1200	-	-	0/4/4/4	0/0/0/0
8	PEG	A	1201	-	-	0/4/4/4	0/0/0/0
9	P6G	A	1202	-	-	0/16/16/16	0/0/0/0
8	PEG	A	1204	-	-	0/4/4/4	0/0/0/0
2	ASP	A	907	3	-	0/2/6/8	0/0/0/0
3	SER	A	908	2	-	0/2/6/6	0/0/0/0
8	PEG	B	1200	-	-	0/4/4/4	0/0/0/0
8	PEG	B	1201	-	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PEG	B	1202	-	-	0/4/4/4	0/0/0/0
9	P6G	B	1203	-	-	0/16/16/16	0/0/0/0
9	P6G	B	1204	-	-	0/16/16/16	0/0/0/0
2	ASP	B	907	3	-	0/2/6/8	0/0/0/0
3	SER	B	908	2	-	0/2/6/6	0/0/0/0
8	PEG	C	1200	-	-	0/4/4/4	0/0/0/0
8	PEG	C	1201	-	-	0/4/4/4	0/0/0/0
9	P6G	C	1202	-	-	0/16/16/16	0/0/0/0
2	ASP	C	907	3	-	0/2/6/8	0/0/0/0
3	SER	C	908	2	-	0/2/6/6	0/0/0/0
6	NAG	D	1102	1	-	0/6/23/26	0/1/1/1
8	PEG	D	1200	-	-	0/4/4/4	0/0/0/0
9	P6G	D	1201	-	-	0/16/16/16	0/0/0/0
8	PEG	D	1202	-	-	0/4/4/4	0/0/0/0
9	P6G	D	1203	-	-	0/16/16/16	0/0/0/0
8	PEG	D	1205	-	-	0/4/4/4	0/0/0/0
2	ASP	D	907	3	-	0/2/6/8	0/0/0/0
3	SER	D	908	2	-	0/2/6/6	0/0/0/0

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1100	NAG	O5-C5-C4	-2.58	105.86	110.13
2	A	907	ASP	O-C-CA	-2.35	119.41	125.72
2	D	907	ASP	O-C-CA	-2.20	119.83	125.72
2	B	907	ASP	O-C-CA	-2.16	119.94	125.72
2	C	907	ASP	O-C-CA	-2.16	119.94	125.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1201	PEG	1	0
9	A	1202	P6G	1	0
2	A	907	ASP	1	0
9	D	1201	P6G	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

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

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

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	603/629 (95%)	-0.22	14 (2%) 64 59	14, 24, 42, 69	0
1	B	605/629 (96%)	-0.21	15 (2%) 61 56	12, 21, 37, 69	0
1	C	609/629 (96%)	-0.35	9 (1%) 76 72	12, 19, 35, 59	0
1	D	607/629 (96%)	-0.14	29 (4%) 34 28	12, 20, 43, 60	0
All	All	2424/2516 (96%)	-0.23	67 (2%) 56 51	12, 21, 41, 69	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	135	THR	7.0
1	A	415	THR	6.6
1	B	134	ALA	6.4
1	D	88	LEU	6.2
1	B	325	GLY	5.7

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

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

### 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	NAG	B	1102	14/15	0.82	0.26	5.25	37,39,44,50	0
10	NAG	D	1100	14/15	0.81	0.14	1.97	30,35,40,44	0
10	NAG	C	1100	14/15	0.87	0.15	1.73	32,38,43,46	0
7	NAG	C	1102	14/15	0.92	0.15	1.72	28,31,33,36	0
7	NAG	A	1101	14/15	0.88	0.20	1.36	34,37,41,47	0
10	NAG	B	1100	14/15	0.79	0.13	1.12	29,34,39,40	0
11	NAG	B	1104	14/15	0.94	0.13	0.33	38,40,43,43	0
13	NAG	D	1103	14/15	0.92	0.10	0.01	31,34,36,38	0
12	NAG	C	1104	14/15	0.94	0.09	-0.02	27,29,32,37	0
10	FUC	B	1101	10/11	0.69	0.26	-	43,47,48,49	0
12	NAG	C	1105	14/15	0.91	0.14	-	32,34,39,41	0
12	FUL	C	1107	10/11	0.72	0.20	-	44,48,50,50	0
11	BMA	B	1106	11/12	0.84	0.30	-	60,63,64,65	0
13	NAG	D	1104	14/15	0.85	0.16	-	42,49,53,57	0
12	BMA	C	1106	11/12	0.87	0.18	-	40,42,45,47	0
7	NAG	C	1103	14/15	0.80	0.30	-	40,44,48,49	0
11	NAG	B	1105	14/15	0.86	0.23	-	47,49,52,56	0
13	MAN	D	1105	11/12	0.65	0.26	-	61,63,65,65	0
10	FUC	D	1101	10/11	0.71	0.29	-	50,54,56,56	0
13	FUC	D	1106	10/11	0.85	0.17	-	40,42,45,45	0
7	NAG	B	1103	14/15	0.69	0.40	-	56,59,61,61	0
7	NAG	A	1103	14/15	0.83	0.26	-	61,63,66,69	0
10	FUC	C	1101	10/11	0.57	0.37	-	54,58,61,62	0
7	NAG	A	1102	14/15	0.70	0.42	-	54,59,60,61	0
7	NAG	A	1104	14/15	0.81	0.33	-	65,66,68,68	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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
9	P6G	D	1203	19/19	0.77	0.21	9.63	44,56,64,64	0
9	P6G	B	1204	19/19	0.74	0.27	7.30	48,55,65,66	0
8	PEG	B	1202	7/7	0.81	0.18	7.22	41,41,42,43	0
9	P6G	C	1202	19/19	0.66	0.25	5.93	44,57,61,63	0
9	P6G	B	1203	19/19	0.79	0.16	5.02	36,39,47,47	0
8	PEG	A	1201	7/7	0.87	0.15	4.78	40,42,45,46	0
8	PEG	A	1204	7/7	0.83	0.23	4.56	40,40,43,44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ASP	D	907	8/9	0.91	0.14	3.48	20,25,36,37	0
6	NAG	A	1100	14/15	0.80	0.17	3.16	37,42,47,50	0
2	ASP	B	907	8/9	0.85	0.17	3.07	22,27,36,37	0
6	NAG	D	1102	14/15	0.74	0.26	2.69	36,37,40,40	0
2	ASP	A	907	8/9	0.91	0.16	2.51	21,26,36,37	0
8	PEG	C	1201	7/7	0.88	0.17	2.39	38,39,43,44	0
8	PEG	B	1201	7/7	0.71	0.17	2.35	45,46,47,48	0
9	P6G	D	1201	19/19	0.86	0.13	2.01	27,30,42,43	0
9	P6G	A	1202	19/19	0.85	0.12	1.81	33,39,43,44	0
3	SER	B	908	7/7	0.94	0.15	1.75	18,20,23,25	0
8	PEG	D	1205	7/7	0.79	0.20	1.72	43,44,48,49	0
2	ASP	C	907	8/9	0.91	0.12	1.45	18,23,29,31	0
8	PEG	A	1200	7/7	0.85	0.18	0.98	53,53,54,54	0
8	PEG	C	1200	7/7	0.86	0.13	0.17	44,45,48,49	0
3	SER	A	908	7/7	0.95	0.10	-0.20	20,22,24,26	0
5	CL	A	1002	1/1	1.00	0.07	-0.93	18,18,18,18	0
3	SER	C	908	7/7	0.95	0.08	-1.34	17,19,21,23	0
5	CL	C	1002	1/1	1.00	0.06	-1.51	14,14,14,14	0
5	CL	B	1002	1/1	1.00	0.06	-1.62	16,16,16,16	0
3	SER	D	908	7/7	0.97	0.08	-1.80	20,21,22,25	0
5	CL	D	1002	1/1	1.00	0.06	-2.61	15,15,15,15	0
8	PEG	D	1200	7/7	0.82	0.15	-	51,51,53,55	0
8	PEG	D	1202	7/7	0.87	0.14	-	48,48,49,50	0
8	PEG	B	1200	7/7	0.87	0.19	-	61,61,62,63	0
4	ZN	C	1001	1/1	1.00	0.04	-	15,15,15,15	0
4	ZN	B	1001	1/1	1.00	0.05	-	17,17,17,17	0
4	ZN	A	1001	1/1	1.00	0.04	-	18,18,18,18	0
4	ZN	D	1001	1/1	1.00	0.04	-	17,17,17,17	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.