



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:14 AM GMT

PDB ID : 2W9L
Title : CANINE ADENOVIRUS TYPE 2 FIBRE HEAD IN COMPLEX WITH CAR
DOMAIN D1 AND SIALIC ACID
Authors : Seiradake, E.; Henaff, D.; Wodrich, H.; Billet, O.; Perreau, M.; Hippert, C.;
Mennechet, F.; Schoehn, G.; Lortat-Jacob, H.; Dreja, H.; Ibanes, S.; Kalatzis,
V.; Wang, J.P.; Finberg, R.W.; Cusack, S.; Kremer, E.J.
Deposited on : 2009-01-26
Resolution : 2.91 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

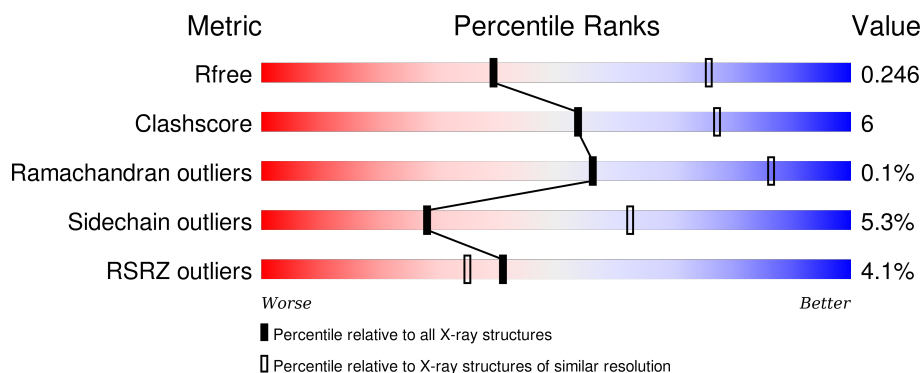
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.91 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	1643 (2.94-2.90)
Clashscore	102246	1871 (2.94-2.90)
Ramachandran outliers	100387	1824 (2.94-2.90)
Sidechain outliers	100360	1826 (2.94-2.90)
RSRZ outliers	91569	1650 (2.94-2.90)

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 ≥ 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	124	<div> <div>25%</div> <div>74%</div> <div>16%</div> <div>• 6%</div> </div>
1	B	124	<div> <div>25%</div> <div>60%</div> <div>31%</div> <div>• 6%</div> </div>
1	G	124	<div> <div>4%</div> <div>78%</div> <div>21%</div> <div>•</div> </div>
1	J	124	<div> <div>8%</div> <div>78%</div> <div>20%</div> <div>••</div> </div>
1	K	124	<div> <div>5%</div> <div>72%</div> <div>25%</div> <div>••</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	O	124	
1	P	124	
1	T	124	
1	V	124	
1	X	124	
1	Y	124	
1	Z	124	
2	C	197	
2	D	197	
2	E	197	
2	F	197	
2	H	197	
2	I	197	
2	L	197	
2	M	197	
2	N	197	
2	Q	197	
2	R	197	
2	S	197	

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GL0	C	1544	X	-	-	-
3	GL0	D	1544	X	-	-	-
3	GL0	L	1544	X	-	-	-
3	GL0	R	1544	X	-	-	-
3	GL0	S	1544	X	-	-	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GAL	H	1544	X	-	-	-
4	GAL	I	1544	X	-	-	-
4	GAL	Q	1544	X	-	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 28671 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 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	117	Total	C	N	O	S	0	0	0
			914	581	147	183	3			
1	B	117	Total	C	N	O	S	0	0	0
			914	583	147	181	3			
1	G	123	Total	C	N	O	S	0	0	0
			959	611	156	189	3			
1	J	123	Total	C	N	O	S	0	0	0
			959	611	156	189	3			
1	K	123	Total	C	N	O	S	0	0	0
			963	615	156	189	3			
1	O	118	Total	C	N	O	S	0	0	0
			917	587	145	182	3			
1	P	123	Total	C	N	O	S	0	0	0
			959	611	156	189	3			
1	T	119	Total	C	N	O	S	0	0	0
			928	591	149	185	3			
1	V	121	Total	C	N	O	S	0	0	0
			944	600	154	187	3			
1	X	123	Total	C	N	O	S	0	0	0
			959	611	156	189	3			
1	Y	119	Total	C	N	O	S	0	0	0
			928	591	149	185	3			
1	Z	117	Total	C	N	O	S	0	0	0
			914	581	147	183	3			

- Molecule 2 is a protein called FIBRE PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	D	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	F	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	H	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	I	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	L	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	M	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	N	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	Q	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	R	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			
2	S	182	Total	C	N	O	S	0	0	0
			1406	892	239	266	9			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	2	Total	C	N	O	0	0
			32	17	1	14		
3	D	2	Total	C	N	O	0	0
			32	17	1	14		
3	F	2	Total	C	N	O	0	0
			32	17	1	14		
3	L	2	Total	C	N	O	0	0
			32	17	1	14		
3	N	2	Total	C	N	O	0	0
			32	17	1	14		
3	R	2	Total	C	N	O	0	0
			32	17	1	14		
3	S	2	Total	C	N	O	0	0
			32	17	1	14		

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	E	2	Total	C	N	O	0	0
			32	17	1	14		
4	H	2	Total	C	N	O	0	0
			32	17	1	14		
4	I	2	Total	C	N	O	0	0
			32	17	1	14		
4	M	2	Total	C	N	O	0	0
			32	17	1	14		
4	Q	2	Total	C	N	O	0	0
			32	17	1	14		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	O	0	0
			2	2		
5	B	4	Total	O	0	0
			4	4		
5	C	5	Total	O	0	0
			5	5		
5	D	5	Total	O	0	0
			5	5		
5	E	4	Total	O	0	0
			4	4		
5	F	12	Total	O	0	0
			12	12		
5	G	9	Total	O	0	0
			9	9		
5	H	9	Total	O	0	0
			9	9		
5	I	10	Total	O	0	0
			10	10		
5	J	2	Total	O	0	0
			2	2		
5	K	1	Total	O	0	0
			1	1		
5	L	11	Total	O	0	0
			11	11		
5	M	14	Total	O	0	0
			14	14		
5	N	17	Total	O	0	0
			17	17		
5	O	5	Total	O	0	0
			5	5		

Continued on next page...

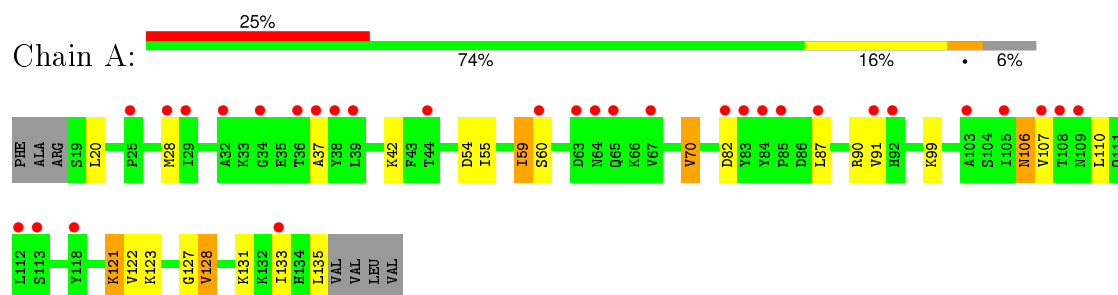
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	P	1	Total 1	O 1	0	0
5	Q	13	Total 13	O 13	0	0
5	R	10	Total 10	O 10	0	0
5	S	7	Total 7	O 7	0	0
5	T	5	Total 5	O 5	0	0
5	V	1	Total 1	O 1	0	0
5	X	3	Total 3	O 3	0	0
5	Y	2	Total 2	O 2	0	0
5	Z	5	Total 5	O 5	0	0

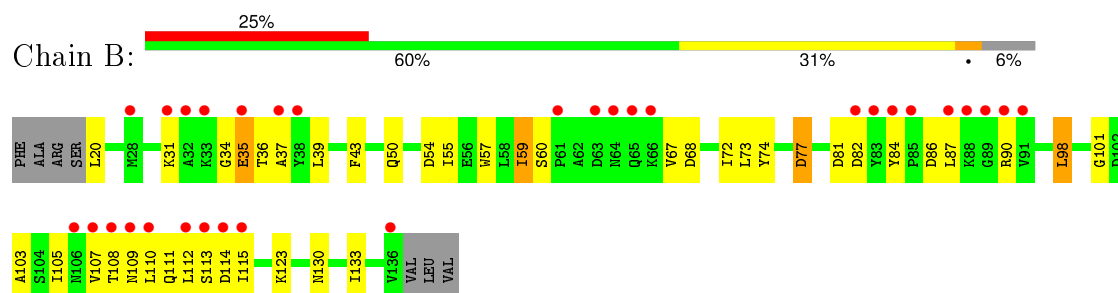
3 Residue-property plots [i](#)

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 ($RSRZ > 2$). 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.

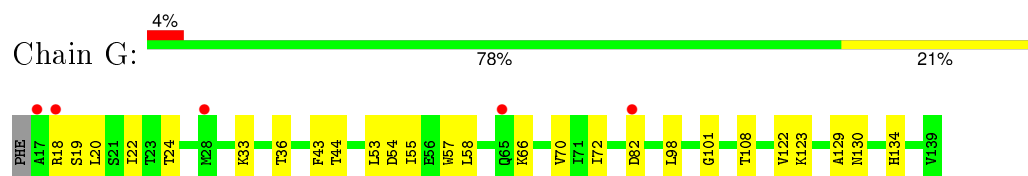
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



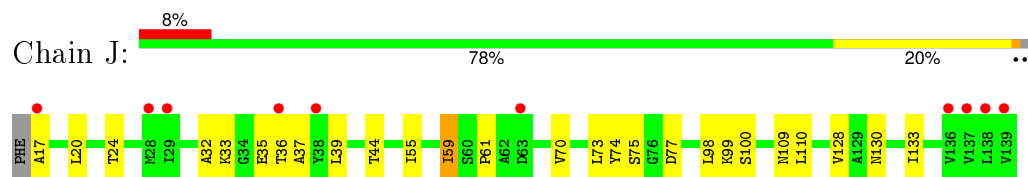
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



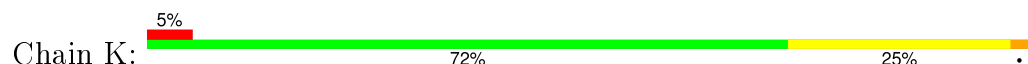
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR

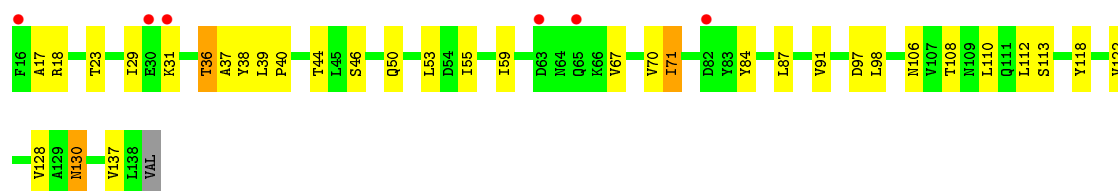


• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR

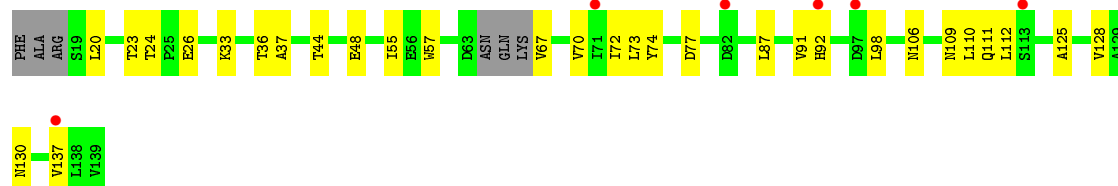


• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR

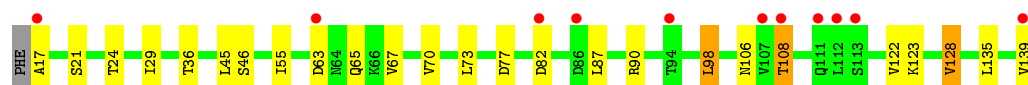
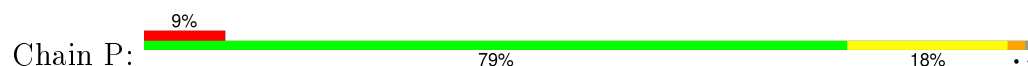




• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



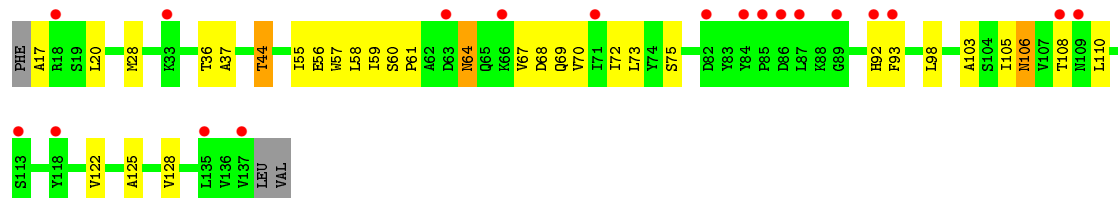
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



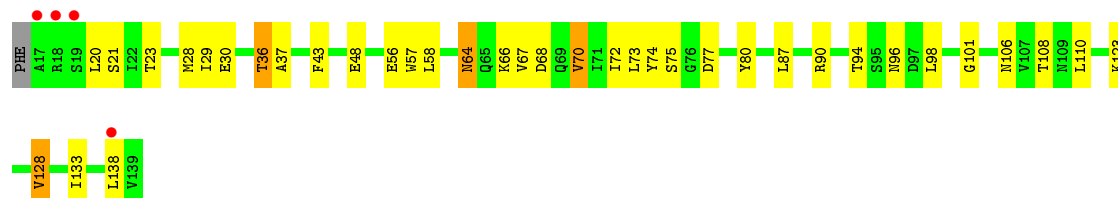
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



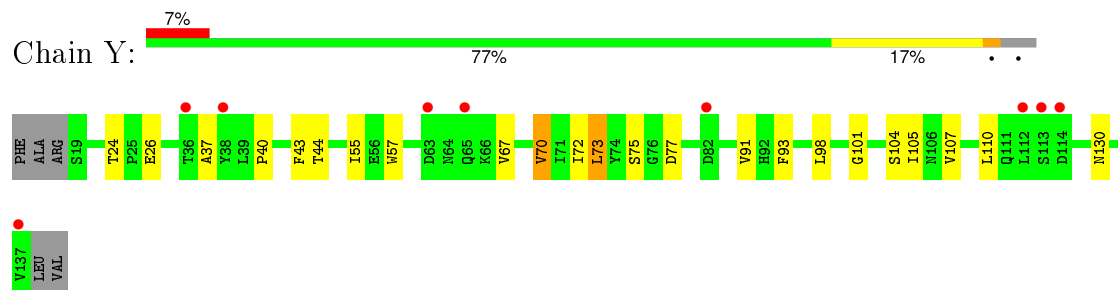
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



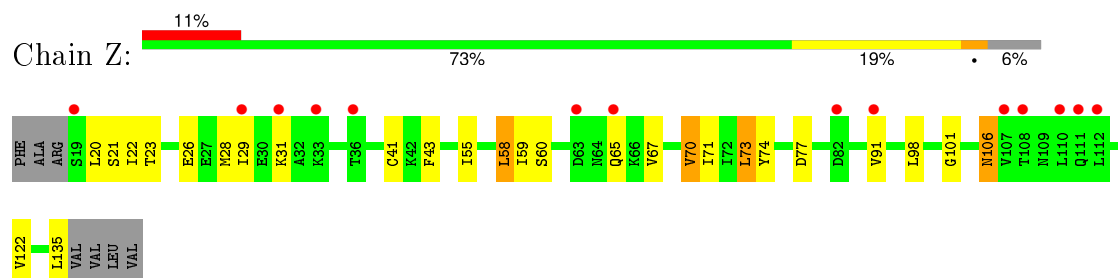
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



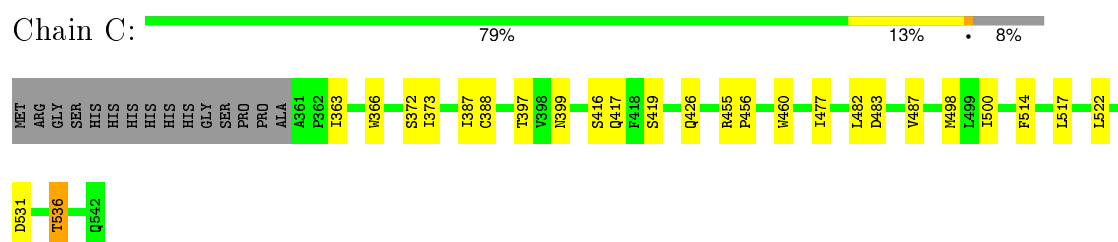
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



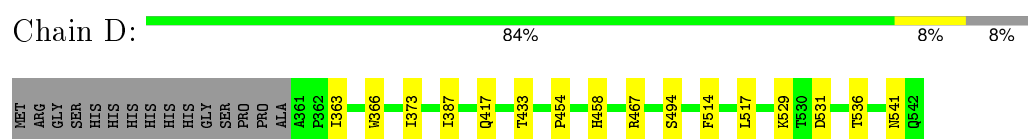
• Molecule 1: COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR



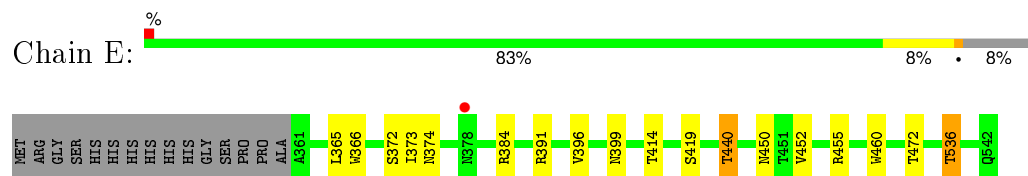
• Molecule 2: FIBRE PROTEIN



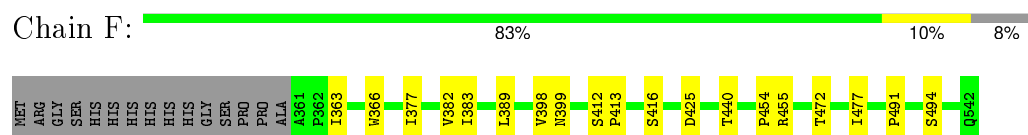
• Molecule 2: FIBRE PROTEIN




• Molecule 2: FIBRE PROTEIN



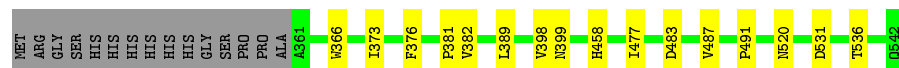

• Molecule 2: FIBRE PROTEIN




• Molecule 2: FIBRE PROTEIN

Chain H: 


• Molecule 2: FIBRE PROTEIN

Chain I: 

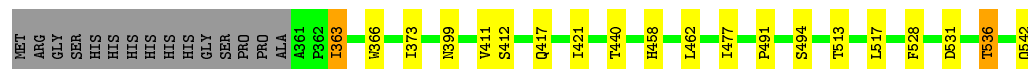

• Molecule 2: FIBRE PROTEIN

Chain L: 


• Molecule 2: FIBRE PROTEIN

Chain M: 


• Molecule 2: FIBRE PROTEIN

Chain N: 

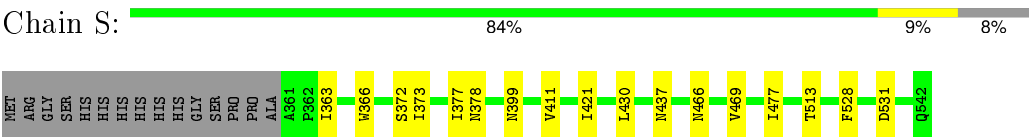
• Molecule 2: FIBRE PROTEIN

Chain Q: 

• Molecule 2: FIBRE PROTEIN

Chain R: 

● Molecule 2: FIBRE PROTEIN



4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	221.15Å 221.15Å 391.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.63 – 2.91 49.61 – 2.91	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.63-2.91) 99.0 (49.61-2.91)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.15 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.197 , 0.254 0.196 , 0.246	Depositor DCC
R_{free} test set	1031 reflections (0.99%)	DCC
Wilson B-factor (Å ²)	46.5	Xtriage
Anisotropy	0.395	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 56.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 105111 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	28671	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.13% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GL0, SIA, GAL

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	0/932	0.62	0/1264
1	B	0.56	0/932	0.75	0/1265
1	G	0.51	0/977	0.67	0/1326
1	J	0.49	0/977	0.66	0/1326
1	K	0.52	0/982	0.67	0/1332
1	O	0.52	0/934	0.61	0/1268
1	P	0.46	0/977	0.62	0/1326
1	T	0.45	0/946	0.64	0/1284
1	V	0.46	0/962	0.62	0/1305
1	X	0.49	0/977	0.63	0/1326
1	Y	0.53	0/946	0.63	0/1284
1	Z	0.49	0/932	0.63	0/1264
2	C	0.55	1/1441 (0.1%)	0.68	0/1964
2	D	0.49	0/1441	0.66	0/1964
2	E	0.49	0/1441	0.63	0/1964
2	F	0.53	0/1441	0.65	0/1964
2	H	0.52	0/1441	0.65	0/1964
2	I	0.48	0/1441	0.63	0/1964
2	L	0.50	0/1441	0.66	0/1964
2	M	0.52	0/1441	0.66	0/1964
2	N	0.54	0/1441	0.67	0/1964
2	Q	0.49	0/1441	0.65	0/1964
2	R	0.53	0/1441	0.66	0/1964
2	S	0.49	0/1441	0.65	0/1964
All	All	0.51	1/28766 (0.0%)	0.65	0/39138

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	G	0	1
3	C	1	0
3	D	1	0
3	L	1	0
3	R	1	0
3	S	1	0
4	H	1	0
4	I	1	0
4	Q	1	0
All	All	8	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	388	CYS	CB-SG	-6.00	1.72	1.82

There are no bond angle outliers.

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	1544	GL0	C3
3	D	1544	GL0	C3
4	H	1544	GAL	C3
4	I	1544	GAL	C3
3	L	1544	GL0	C3

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	82	ASP	Peptide

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	914	0	910	21	0
1	B	914	0	914	55	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	959	0	966	13	0
1	J	959	0	966	18	0
1	K	963	0	966	19	0
1	O	917	0	920	15	0
1	P	959	0	966	18	0
1	T	928	0	928	20	0
1	V	944	0	946	23	0
1	X	959	0	966	20	0
1	Y	928	0	928	13	0
1	Z	914	0	910	19	0
2	C	1406	0	1382	16	0
2	D	1406	0	1382	10	0
2	E	1406	0	1382	16	0
2	F	1406	0	1382	11	0
2	H	1406	0	1382	16	0
2	I	1406	0	1382	18	0
2	L	1406	0	1382	14	0
2	M	1406	0	1382	10	0
2	N	1406	0	1382	20	0
2	Q	1406	0	1382	13	0
2	R	1406	0	1382	15	0
2	S	1406	0	1382	16	0
3	C	32	0	27	0	0
3	D	32	0	27	0	0
3	F	32	0	27	0	0
3	L	32	0	27	0	0
3	N	32	0	27	0	0
3	R	32	0	27	1	0
3	S	32	0	27	0	0
4	E	32	0	27	1	0
4	H	32	0	25	0	0
4	I	32	0	25	1	0
4	M	32	0	26	0	0
4	Q	32	0	25	0	0
5	A	2	0	0	0	0
5	B	4	0	0	1	0
5	C	5	0	0	0	0
5	D	5	0	0	2	0
5	E	4	0	0	2	0
5	F	12	0	0	3	0
5	G	9	0	0	3	0
5	H	9	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	I	10	0	0	2	0
5	J	2	0	0	1	0
5	K	1	0	0	1	0
5	L	11	0	0	3	0
5	M	14	0	0	0	0
5	N	17	0	0	4	0
5	O	5	0	0	0	0
5	P	1	0	0	0	0
5	Q	13	0	0	7	0
5	R	10	0	0	2	0
5	S	7	0	0	0	0
5	T	5	0	0	2	0
5	V	1	0	0	0	0
5	X	3	0	0	1	0
5	Y	2	0	0	0	0
5	Z	5	0	0	2	0
All	All	28671	0	28187	368	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 6.

The worst 5 of 368 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)
2:N:458:HIS:CD2	5:N:2008:HOH:O	1.84	1.27
1:X:28:MET:HB3	5:X:2001:HOH:O	1.48	1.14
1:B:90:ARG:HD2	1:B:108:THR:O	1.48	1.10
2:Q:467:ARG:HD2	5:Q:2009:HOH:O	1.51	1.09
2:I:373:ILE:HD11	1:P:70:VAL:HG21	1.10	1.08

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	115/124 (93%)	104 (90%)	11 (10%)	0	100	100
1	B	115/124 (93%)	104 (90%)	10 (9%)	1 (1%)	21	56
1	G	121/124 (98%)	110 (91%)	11 (9%)	0	100	100
1	J	121/124 (98%)	113 (93%)	8 (7%)	0	100	100
1	K	121/124 (98%)	107 (88%)	14 (12%)	0	100	100
1	O	114/124 (92%)	106 (93%)	8 (7%)	0	100	100
1	P	121/124 (98%)	111 (92%)	8 (7%)	2 (2%)	11	37
1	T	117/124 (94%)	104 (89%)	13 (11%)	0	100	100
1	V	119/124 (96%)	113 (95%)	6 (5%)	0	100	100
1	X	121/124 (98%)	116 (96%)	5 (4%)	0	100	100
1	Y	117/124 (94%)	113 (97%)	4 (3%)	0	100	100
1	Z	115/124 (93%)	104 (90%)	10 (9%)	1 (1%)	21	56
2	C	180/197 (91%)	171 (95%)	9 (5%)	0	100	100
2	D	180/197 (91%)	172 (96%)	8 (4%)	0	100	100
2	E	180/197 (91%)	171 (95%)	9 (5%)	0	100	100
2	F	180/197 (91%)	169 (94%)	11 (6%)	0	100	100
2	H	180/197 (91%)	170 (94%)	10 (6%)	0	100	100
2	I	180/197 (91%)	172 (96%)	8 (4%)	0	100	100
2	L	180/197 (91%)	171 (95%)	9 (5%)	0	100	100
2	M	180/197 (91%)	170 (94%)	10 (6%)	0	100	100
2	N	180/197 (91%)	171 (95%)	9 (5%)	0	100	100
2	Q	180/197 (91%)	171 (95%)	9 (5%)	0	100	100
2	R	180/197 (91%)	170 (94%)	10 (6%)	0	100	100
2	S	180/197 (91%)	170 (94%)	10 (6%)	0	100	100
All	All	3577/3852 (93%)	3353 (94%)	220 (6%)	4 (0%)	56	86

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	35	GLU
1	P	63	ASP
1	Z	65	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	P	24	THR

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	104/110 (94%)	95 (91%)	9 (9%)	13	35
1	B	104/110 (94%)	98 (94%)	6 (6%)	25	58
1	G	109/110 (99%)	103 (94%)	6 (6%)	27	60
1	J	109/110 (99%)	101 (93%)	8 (7%)	17	45
1	K	109/110 (99%)	96 (88%)	13 (12%)	6	18
1	O	105/110 (96%)	95 (90%)	10 (10%)	11	30
1	P	109/110 (99%)	99 (91%)	10 (9%)	11	32
1	T	106/110 (96%)	97 (92%)	9 (8%)	13	36
1	V	107/110 (97%)	101 (94%)	6 (6%)	26	59
1	X	109/110 (99%)	100 (92%)	9 (8%)	14	38
1	Y	106/110 (96%)	99 (93%)	7 (7%)	21	50
1	Z	104/110 (94%)	95 (91%)	9 (9%)	13	35
2	C	159/171 (93%)	150 (94%)	9 (6%)	25	58
2	D	159/171 (93%)	154 (97%)	5 (3%)	47	81
2	E	159/171 (93%)	153 (96%)	6 (4%)	40	76
2	F	159/171 (93%)	151 (95%)	8 (5%)	30	65
2	H	159/171 (93%)	156 (98%)	3 (2%)	65	89
2	I	159/171 (93%)	157 (99%)	2 (1%)	76	93
2	L	159/171 (93%)	152 (96%)	7 (4%)	35	70
2	M	159/171 (93%)	150 (94%)	9 (6%)	25	58
2	N	159/171 (93%)	153 (96%)	6 (4%)	40	76
2	Q	159/171 (93%)	155 (98%)	4 (2%)	55	85
2	R	159/171 (93%)	155 (98%)	4 (2%)	55	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	S	159/171 (93%)	156 (98%)	3 (2%)	65	89
All	All	3189/3372 (95%)	3021 (95%)	168 (5%)	28	62

5 of 168 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	366	TRP
2	N	440	THR
1	Y	73	LEU
2	L	411	VAL
2	M	381	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	542	GLN
1	O	130	ASN
1	X	134	HIS
2	N	378	ASN
2	N	458	HIS

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 ⓘ

24 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	SIA	C	1543	3	17,21,21	1.10	1 (5%)	19,31,31	1.24	3 (15%)
3	GL0	C	1544	3	11,11,12	1.04	0	12,15,17	1.52	3 (25%)
3	SIA	D	1543	3	17,21,21	1.32	2 (11%)	19,31,31	1.62	4 (21%)
3	GL0	D	1544	3	11,11,12	1.14	0	12,15,17	1.38	3 (25%)
4	SIA	E	1543	4	17,21,21	1.41	2 (11%)	19,31,31	1.98	6 (31%)
4	GAL	E	1544	4	11,11,12	0.96	1 (9%)	12,15,17	0.87	0
3	SIA	F	1543	3	17,21,21	1.55	2 (11%)	19,31,31	0.90	1 (5%)
3	GL0	F	1544	3	11,11,12	1.46	3 (27%)	12,15,17	1.98	2 (16%)
4	SIA	H	1543	4	17,21,21	1.29	1 (5%)	19,31,31	1.76	2 (10%)
4	GAL	H	1544	4	11,11,12	29.58	1 (9%)	14,15,17	13.28	2 (14%)
4	SIA	I	1543	4	17,21,21	1.41	1 (5%)	19,31,31	1.35	3 (15%)
4	GAL	I	1544	4	11,11,12	29.70	1 (9%)	14,15,17	12.86	4 (28%)
3	SIA	L	1543	3	17,21,21	1.98	3 (17%)	19,31,31	1.51	3 (15%)
3	GL0	L	1544	3	11,11,12	0.96	0	12,15,17	1.15	0
4	SIA	M	1543	4	17,21,21	1.12	1 (5%)	19,31,31	2.23	6 (31%)
4	GAL	M	1544	4	11,11,12	22.93	2 (18%)	14,15,17	10.52	6 (42%)
3	SIA	N	1543	3	17,21,21	1.85	1 (5%)	19,31,31	1.26	3 (15%)
3	GL0	N	1544	3	11,11,12	1.13	1 (9%)	12,15,17	1.46	1 (8%)
4	SIA	Q	1543	4	17,21,21	1.52	2 (11%)	19,31,31	1.32	3 (15%)
4	GAL	Q	1544	4	11,11,12	29.72	1 (9%)	14,15,17	12.92	5 (35%)
3	SIA	R	1543	3	17,21,21	1.18	2 (11%)	19,31,31	1.72	4 (21%)
3	GL0	R	1544	3	11,11,12	1.01	1 (9%)	12,15,17	1.24	1 (8%)
3	SIA	S	1543	3	17,21,21	1.03	1 (5%)	19,31,31	1.33	2 (10%)
3	GL0	S	1544	3	11,11,12	0.82	0	12,15,17	1.11	2 (16%)

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	SIA	C	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	C	1544	3	1/1/4/5	0/2/18/22	0/1/1/1
3	SIA	D	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	D	1544	3	1/1/4/5	0/2/18/22	0/1/1/1
4	SIA	E	1543	4	-	0/14/38/38	0/1/1/1
4	GAL	E	1544	4	-	0/2/18/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SIA	F	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	F	1544	3	-	0/2/18/22	0/1/1/1
4	SIA	H	1543	4	-	0/14/38/38	0/1/1/1
4	GAL	H	1544	4	1/1/4/5	0/2/19/22	0/1/1/1
4	SIA	I	1543	4	-	0/14/38/38	0/1/1/1
4	GAL	I	1544	4	1/1/4/5	0/2/19/22	0/1/1/1
3	SIA	L	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	L	1544	3	1/1/4/5	0/2/18/22	0/1/1/1
4	SIA	M	1543	4	-	0/14/38/38	0/1/1/1
4	GAL	M	1544	4	-	0/2/19/22	0/1/1/1
3	SIA	N	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	N	1544	3	-	0/2/18/22	0/1/1/1
4	SIA	Q	1543	4	-	0/14/38/38	0/1/1/1
4	GAL	Q	1544	4	1/1/4/5	0/2/19/22	0/1/1/1
3	SIA	R	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	R	1544	3	1/1/4/5	0/2/18/22	0/1/1/1
3	SIA	S	1543	3	-	0/14/38/38	0/1/1/1
3	GL0	S	1544	3	1/1/4/5	0/2/18/22	0/1/1/1

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	1544	GL0	C3-C2	-2.43	1.48	1.52
3	F	1544	GL0	C3-C4	-2.21	1.49	1.52
3	R	1543	SIA	C3-C2	2.00	1.54	1.51
3	N	1544	GL0	C1-C2	2.03	1.54	1.52
4	M	1544	GAL	O5-C1	2.14	1.47	1.43

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1544	GAL	O3-C3-C2	-48.96	21.56	110.00
4	I	1544	GAL	O3-C3-C2	-47.52	24.15	110.00
4	Q	1544	GAL	O3-C3-C2	-47.44	24.30	110.00
4	M	1544	GAL	O3-C3-C2	-35.74	45.44	110.00
4	M	1544	GAL	O3-C3-C4	-13.56	79.81	110.34

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	H	1544	GAL	C3
3	R	1544	GL0	C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
3	C	1544	GL0	C3
3	D	1544	GL0	C3
3	L	1544	GL0	C3

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	1543	SIA	1	0
4	I	1544	GAL	1	0
3	R	1543	SIA	1	0

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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, 95th 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 ²)	Q<0.9
1	A	117/124 (94%)	1.31	31 (26%) 1 0	36, 40, 43, 44	0
1	B	117/124 (94%)	1.27	31 (26%) 1 0	27, 35, 40, 42	0
1	G	123/124 (99%)	0.28	5 (4%) 41 35	35, 40, 44, 47	0
1	J	123/124 (99%)	0.56	10 (8%) 15 10	35, 40, 44, 47	0
1	K	123/124 (99%)	0.46	6 (4%) 33 28	35, 40, 44, 48	0
1	O	118/124 (95%)	0.61	6 (5%) 32 26	37, 40, 43, 43	0
1	P	123/124 (99%)	0.81	11 (8%) 12 8	36, 40, 43, 46	0
1	T	119/124 (95%)	0.55	2 (1%) 73 70	37, 40, 43, 45	0
1	V	121/124 (97%)	0.77	19 (15%) 3 1	36, 40, 43, 47	0
1	X	123/124 (99%)	0.43	4 (3%) 50 43	36, 40, 44, 47	0
1	Y	119/124 (95%)	0.55	9 (7%) 17 11	35, 40, 43, 45	0
1	Z	117/124 (94%)	0.79	14 (11%) 6 3	37, 40, 43, 44	0
2	C	182/197 (92%)	-0.09	0 100 100	38, 40, 42, 46	0
2	D	182/197 (92%)	-0.11	0 100 100	39, 40, 42, 44	0
2	E	182/197 (92%)	-0.17	1 (0%) 91 90	38, 40, 42, 44	0
2	F	182/197 (92%)	-0.20	0 100 100	39, 40, 42, 47	0
2	H	182/197 (92%)	-0.16	0 100 100	39, 40, 42, 46	0
2	I	182/197 (92%)	-0.15	0 100 100	39, 40, 42, 44	0
2	L	182/197 (92%)	-0.05	1 (0%) 91 90	39, 40, 42, 45	0
2	M	182/197 (92%)	-0.14	0 100 100	39, 40, 41, 44	0
2	N	182/197 (92%)	-0.20	0 100 100	39, 40, 42, 47	0
2	Q	182/197 (92%)	-0.03	0 100 100	39, 40, 42, 45	0
2	R	182/197 (92%)	-0.18	0 100 100	39, 40, 42, 45	0
2	S	182/197 (92%)	-0.11	0 100 100	39, 40, 42, 45	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
All	All	3627/3852 (94%)	0.20	150 (4%) 41 35	27, 40, 43, 48	0

The worst 5 of 150 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	109	ASN	8.5
1	B	32	ALA	6.0
1	B	136	VAL	6.0
1	B	65	GLN	5.3
1	B	112	LEU	5.0

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, 95th 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(Å ²)	Q<0.9
3	SIA	D	1543	21/21	0.94	0.28	1.51	41,46,54,58	0
4	SIA	H	1543	21/21	0.95	0.20	0.58	33,38,41,46	0
4	SIA	M	1543	21/21	0.94	0.18	0.35	31,34,39,45	0
3	SIA	C	1543	21/21	0.96	0.22	0.31	34,37,39,45	0
3	SIA	F	1543	21/21	0.95	0.20	0.28	36,40,45,51	0
3	SIA	L	1543	21/21	0.95	0.23	0.04	36,39,43,50	0
4	SIA	E	1543	21/21	0.96	0.21	-0.31	32,40,42,48	0
4	SIA	I	1543	21/21	0.96	0.19	-0.32	31,36,43,50	0
3	SIA	S	1543	21/21	0.95	0.19	-0.34	42,44,49,54	0
4	SIA	Q	1543	21/21	0.95	0.20	-0.46	38,42,48,53	0
3	SIA	N	1543	21/21	0.95	0.14	-1.22	31,35,38,43	0
3	SIA	R	1543	21/21	0.96	0.14	-2.74	31,37,43,48	0
3	GL0	R	1544	11/12	0.76	0.28	-	55,57,59,60	0
3	GL0	C	1544	11/12	0.85	0.28	-	51,53,54,55	0
3	GL0	D	1544	11/12	0.59	0.44	-	65,68,69,69	0
3	GL0	L	1544	11/12	0.65	0.32	-	56,59,60,61	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	GL0	S	1544	11/12	0.78	0.29	-	60,62,63,64	0
3	GL0	F	1544	11/12	0.86	0.32	-	57,60,63,63	0
4	GAL	E	1544	11/12	0.88	0.29	-	51,53,54,54	0
4	GAL	H	1544	11/12	0.88	0.24	-	49,51,53,54	0
3	GL0	N	1544	11/12	0.86	0.21	-	49,51,56,58	0
4	GAL	I	1544	11/12	0.79	0.29	-	57,60,64,65	0
4	GAL	Q	1544	11/12	0.78	0.28	-	59,61,64,65	0
4	GAL	M	1544	11/12	0.80	0.21	-	50,52,56,56	0

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.