



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 02:04 PM GMT

PDB ID : 3W13  
Title : Insulin receptor ectodomain construct comprising domains L1-CR in complex with high-affinity insulin analogue [D-PRO-B26]-DTI-NH<sub>2</sub>, alpha peptide(693-719) and FAB 83-7  
Authors : Lawrence, M.C.; Smith, B.J.; Brzozowski, A.M.  
Deposited on : 2012-11-06  
Resolution : 4.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (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

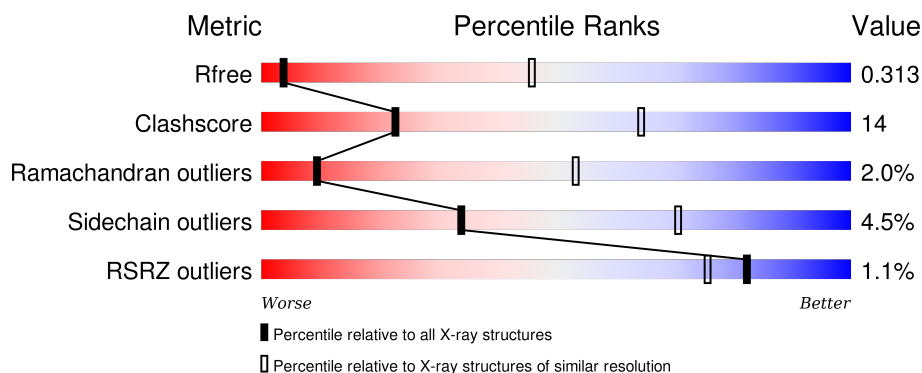
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.30 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	1059 (5.00-3.60)
Clashscore	102246	1166 (5.00-3.60)
Ramachandran outliers	100387	1106 (5.00-3.60)
Sidechain outliers	100360	1089 (5.00-3.60)
RSRZ outliers	91569	1062 (5.00-3.60)

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	E	310	<div> <div></div> <div>70% 21% 7%</div> </div>
2	C	116	<div> <div>4%</div> <div>73% 24%</div> </div>
3	D	114	<div> <div></div> <div>64% 32%</div> </div>
4	A	21	<div> <div></div> <div>57% 33% 5% 5%</div> </div>
5	B	26	<div> <div></div> <div>50% 8% 42%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	F	23	 A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: green (13%), yellow (30%), orange (5%), and grey (52%). A small black dot is located on the orange segment.

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 4571 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 Insulin receptor domains L1-CR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	288	Total	C	N	O	S	0	0	0
			2301	1449	397	423	32			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	144	HIS	TYR	SEE REMARK 999	UNP P06213

- Molecule 2 is a protein called monoclonal antibody fab 83-7 fragment - heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	116	Total	C	N	O	S	0	0	0
			892	562	151	174	5			

- Molecule 3 is a protein called monoclonal antibody fab 83-7 fragment - light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	114	Total	C	N	O	S	0	0	0
			891	565	149	173	4			

- Molecule 4 is a protein called Insulin A chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	21	Total	C	N	O	S	0	0	0
			163	99	25	35	4			

- Molecule 5 is a protein called Insulin B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	15	Total	C	N	O	S	0	0	0
			109	69	17	21	2			

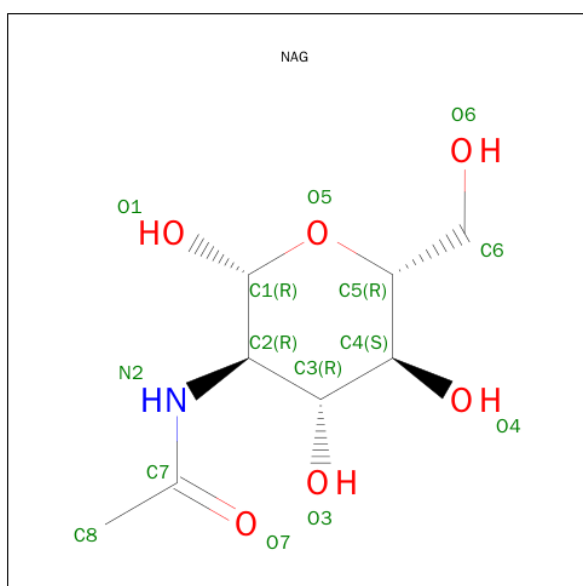
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	26	PR9	TYR	SEE REMARK 999	UNP P01308

- Molecule 6 is a protein called Insulin receptor alphact peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	F	11	Total	C	N	O	0	0	0
			98	67	14	17			

- Molecule 7 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	E	1	Total	C	N	O	0	0
			14	8	1	5		
7	E	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 8 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	E	4	Total	C	N	O	0	0
			50	28	2	20		

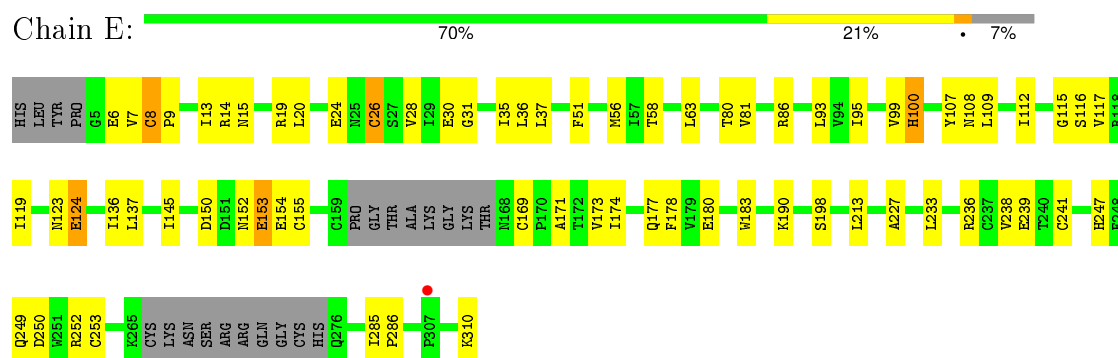
- Molecule 9 is a polymer of unknown type called SUGAR (3-MER).

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

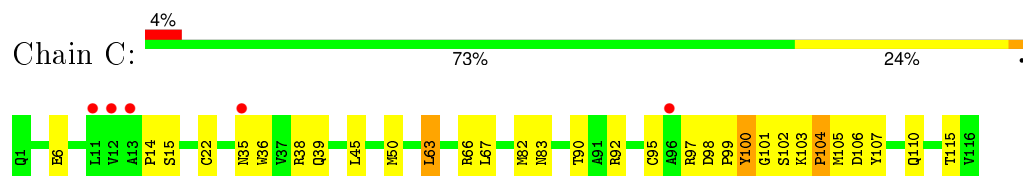
### 3 Residue-property plots [i](#)

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.

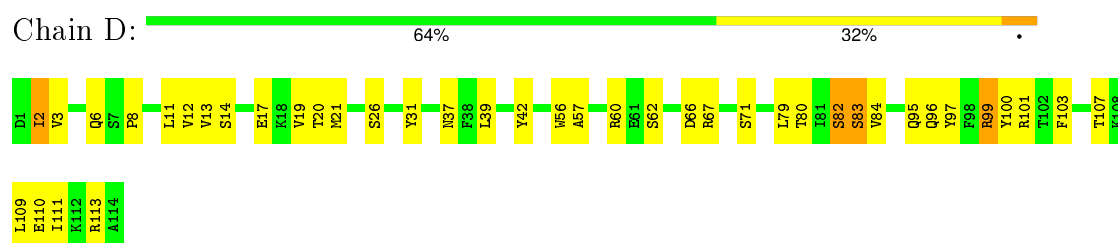
- Molecule 1: Insulin receptor domains L1-CR



- Molecule 2: monoclonal antibody fab 83-7 fragment - heavy chain



- Molecule 3: monoclonal antibody fab 83-7 fragment - light chain



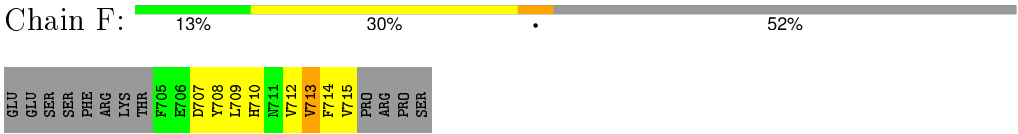
- Molecule 4: Insulin A chain



- Molecule 5: Insulin B chain



- Molecule 6: Insulin receptor alphact peptide





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	169.23Å 169.23Å 169.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.48 – 4.30 29.46 – 4.30	Depositor EDS
% Data completeness (in resolution range)	98.0 (29.48-4.30) 98.4 (29.46-4.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 4.26Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.289 , 0.335 0.264 , 0.313	Depositor DCC
$R_{free}$ test set	530 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	209.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 241.6	EDS
Estimated twinning fraction	0.078 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 11021 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	4571	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	226.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: BMA, NAG, 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	E	0.48	0/2355	0.73	0/3192
2	C	0.49	0/914	0.65	0/1242
3	D	0.57	0/910	0.78	0/1228
4	A	0.60	0/164	0.94	0/220
5	B	0.56	0/110	0.91	0/148
6	F	1.24	1/101 (1.0%)	0.94	0/137
All	All	0.54	1/4554 (0.0%)	0.74	0/6167

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	F	715	VAL	N-CA	7.34	1.61	1.46

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	E	2301	0	2202	55	0
2	C	892	0	869	33	0
3	D	891	0	882	31	0
4	A	163	0	149	9	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	109	0	102	1	0
6	F	98	0	87	8	0
7	E	28	0	26	2	0
8	E	50	0	43	1	0
9	E	39	0	34	0	0
All	All	4571	0	4394	122	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 14.

All (122) 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)
6:F:709:LEU:O	6:F:713:VAL:HG23	1.34	1.26
3:D:67:ARG:HD3	3:D:82:SER:O	1.35	1.22
2:C:99:PRO:HG2	2:C:103:LYS:HB2	1.18	1.10
3:D:2:ILE:HG22	3:D:26:SER:OG	1.63	0.99
1:E:285:ILE:HG13	1:E:286:PRO:HD2	1.48	0.94
2:C:63:LEU:HD22	2:C:67:LEU:HB2	1.56	0.86
2:C:35:ASN:OD1	2:C:50:MET:HB3	1.75	0.86
6:F:709:LEU:O	6:F:713:VAL:CG2	2.24	0.83
3:D:67:ARG:CD	3:D:82:SER:O	2.25	0.82
1:E:285:ILE:HG13	1:E:286:PRO:CD	2.09	0.81
3:D:2:ILE:CG2	3:D:26:SER:OG	2.28	0.81
3:D:60:ARG:HH11	3:D:66:ASP:HA	1.48	0.78
1:E:24:GLU:HG2	7:E:401:NAG:H82	1.66	0.77
6:F:710:HIS:O	6:F:714:PHE:HD2	1.68	0.76
1:E:233:LEU:HD21	2:C:100:TYR:HE2	1.52	0.73
1:E:198:SER:O	1:E:213:LEU:O	2.09	0.71
2:C:22:CYS:HB2	2:C:36:TRP:CH2	2.26	0.71
1:E:95:ILE:HB	1:E:119:ILE:HG12	1.73	0.70
2:C:105:MET:HB2	3:D:42:TYR:OH	1.93	0.69
2:C:35:ASN:OD1	2:C:50:MET:CB	2.42	0.67
4:A:11:CYS:SG	4:A:15:GLN:HB3	2.37	0.65
1:E:107:TYR:HA	1:E:183:TRP:CD1	2.31	0.65
1:E:150:ASP:HA	1:E:155:CYS:SG	2.39	0.62
2:C:63:LEU:HD13	2:C:67:LEU:HD22	1.83	0.61
4:A:2:ILE:HG23	4:A:3:VAL:N	2.17	0.58
1:E:35:ILE:HB	1:E:63:LEU:CD1	2.33	0.58
1:E:36:LEU:HD12	1:E:37:LEU:HB2	1.86	0.58
1:E:238:VAL:HG11	2:C:101:GLY:HA3	1.84	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:67:ARG:CD	3:D:83:SER:HB3	2.33	0.58
3:D:100:TYR:O	3:D:101:ARG:HG3	2.04	0.57
3:D:12:VAL:HG12	3:D:110:GLU:HB2	1.85	0.57
3:D:6:GLN:NE2	3:D:107:THR:HG22	2.20	0.56
1:E:99:VAL:O	1:E:100:HIS:HB2	2.05	0.56
1:E:36:LEU:CD1	1:E:37:LEU:HG	2.35	0.56
3:D:2:ILE:HG22	3:D:3:VAL:H	1.69	0.56
4:A:3:VAL:CG2	6:F:710:HIS:HB2	2.36	0.56
2:C:63:LEU:HD23	2:C:66:ARG:HB2	1.87	0.55
2:C:99:PRO:HG2	2:C:103:LYS:CB	2.12	0.55
2:C:105:MET:HG3	3:D:95:GLN:HE22	1.72	0.55
1:E:171:ALA:HB1	1:E:178:PHE:HB3	1.89	0.54
3:D:8:PRO:O	3:D:107:THR:HB	2.07	0.54
1:E:190:LYS:HD3	7:E:406:NAG:H81	1.91	0.53
1:E:8:CYS:HB3	1:E:9:PRO:HD2	1.90	0.53
1:E:233:LEU:CD2	2:C:100:TYR:HE2	2.19	0.52
6:F:713:VAL:O	6:F:714:PHE:CG	2.62	0.52
3:D:95:GLN:HB2	3:D:103:PHE:CE1	2.44	0.52
2:C:66:ARG:O	2:C:82:MET:HA	2.09	0.51
3:D:96:GLN:O	3:D:101:ARG:HA	2.10	0.51
1:E:8:CYS:HB3	1:E:9:PRO:CD	2.41	0.51
1:E:14:ARG:HG2	1:E:36:LEU:HD11	1.91	0.51
3:D:17:GLU:O	3:D:84:VAL:HG23	2.11	0.51
1:E:173:VAL:HG13	1:E:173:VAL:O	2.11	0.51
2:C:6:GLU:OE1	2:C:6:GLU:N	2.44	0.50
6:F:708:TYR:O	6:F:712:VAL:HG12	2.11	0.50
2:C:104:PRO:HG3	3:D:97:TYR:CZ	2.47	0.50
2:C:63:LEU:HD23	2:C:66:ARG:CB	2.41	0.50
1:E:213:LEU:CB	1:E:227:ALA:O	2.59	0.50
5:B:21:GLU:CD	5:B:21:GLU:C	2.70	0.50
1:E:8:CYS:CB	1:E:9:PRO:CD	2.90	0.49
4:A:13:LEU:HD13	4:A:13:LEU:C	2.33	0.49
2:C:35:ASN:O	2:C:95:CYS:HA	2.12	0.49
1:E:136:ILE:C	8:E:402:NAG:H82	2.34	0.48
2:C:105:MET:O	3:D:42:TYR:HE2	1.97	0.48
1:E:36:LEU:HD13	1:E:37:LEU:HG	1.95	0.47
3:D:99:ARG:HD2	3:D:99:ARG:HA	1.66	0.47
1:E:6:GLU:HB3	1:E:26:CYS:HA	1.97	0.47
3:D:13:VAL:HG22	3:D:14:SER:N	2.30	0.47
1:E:30:GLU:CG	1:E:249:GLN:HG2	2.44	0.47
2:C:66:ARG:HG2	2:C:83:ASN:OD1	2.13	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:8:CYS:CB	1:E:9:PRO:HD2	2.45	0.47
1:E:80:THR:HG23	1:E:108:ASN:O	2.15	0.47
1:E:93:LEU:HB3	1:E:117:VAL:HG22	1.97	0.46
1:E:236:ARG:NH2	2:C:101:GLY:H	2.13	0.46
1:E:31:GLY:H	1:E:58:THR:HB	1.81	0.46
1:E:36:LEU:HD12	1:E:37:LEU:HG	1.97	0.46
1:E:153:GLU:HG2	1:E:153:GLU:O	2.15	0.46
4:A:2:ILE:HG23	4:A:3:VAL:H	1.79	0.46
1:E:109:LEU:HD21	1:E:112:ILE:HD11	1.98	0.46
2:C:99:PRO:HD2	2:C:104:PRO:O	2.16	0.46
1:E:213:LEU:HB3	1:E:227:ALA:O	2.16	0.46
1:E:233:LEU:HB2	1:E:253:CYS:HB2	1.98	0.45
1:E:56:MET:HG3	1:E:81:VAL:HB	1.98	0.45
1:E:233:LEU:HD12	1:E:253:CYS:HB2	1.99	0.45
1:E:236:ARG:NH2	2:C:99:PRO:C	2.70	0.45
3:D:13:VAL:HG12	3:D:109:LEU:HD11	1.99	0.45
2:C:39:GLN:HB2	2:C:45:LEU:HD23	1.98	0.45
1:E:241:CYS:H	1:E:247:HIS:CE1	2.35	0.45
4:A:3:VAL:HG21	6:F:707:ASP:O	2.17	0.45
3:D:11:LEU:HG	3:D:12:VAL:N	2.31	0.45
1:E:174:ILE:HD13	1:E:174:ILE:HA	1.86	0.44
2:C:106:ASP:O	2:C:107:TYR:HB2	2.18	0.44
1:E:285:ILE:CG1	1:E:286:PRO:HD2	2.33	0.44
1:E:7:VAL:HG11	1:E:252:ARG:NH1	2.32	0.44
1:E:241:CYS:H	1:E:247:HIS:HE1	1.66	0.44
3:D:19:VAL:O	3:D:80:THR:HA	2.17	0.44
2:C:105:MET:N	3:D:42:TYR:OH	2.51	0.43
3:D:39:LEU:HD22	3:D:95:GLN:O	2.18	0.43
2:C:97:ARG:HG2	2:C:98:ASP:N	2.33	0.43
1:E:116:SER:HA	1:E:137:LEU:HD22	2.00	0.43
1:E:20:LEU:HG	1:E:51:PHE:CZ	2.52	0.43
1:E:123:ASN:O	1:E:124:GLU:CB	2.67	0.43
1:E:119:ILE:HB	1:E:145:ILE:HG23	2.00	0.43
1:E:99:VAL:O	1:E:100:HIS:CB	2.67	0.43
1:E:236:ARG:HH22	2:C:100:TYR:N	2.18	0.42
2:C:38:ARG:HA	2:C:92:ARG:O	2.19	0.42
2:C:105:MET:HB2	3:D:42:TYR:CZ	2.54	0.42
3:D:13:VAL:O	3:D:111:ILE:HA	2.19	0.42
1:E:28:VAL:HG22	1:E:56:MET:HB3	2.02	0.42
1:E:152:ASN:O	1:E:154:GLU:HG2	2.20	0.42
6:F:713:VAL:O	6:F:714:PHE:CD2	2.73	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:21:MET:SD	3:D:107:THR:HG23	2.60	0.41
1:E:236:ARG:NH1	2:C:99:PRO:HA	2.35	0.41
4:A:12:SER:O	4:A:15:GLN:HB2	2.20	0.41
3:D:20:THR:HA	3:D:79:LEU:O	2.19	0.41
2:C:14:PRO:O	2:C:15:SER:HB2	2.21	0.41
4:A:2:ILE:CG2	4:A:3:VAL:N	2.83	0.41
3:D:60:ARG:NH1	3:D:66:ASP:HA	2.27	0.40
1:E:169:CYS:HB2	1:E:180:GLU:HG2	2.03	0.40
1:E:13:ILE:HG23	1:E:19:ARG:HB2	2.04	0.40
4:A:13:LEU:CD1	4:A:13:LEU:C	2.90	0.40
2:C:90:THR:HG23	2:C:115:THR:HG23	2.03	0.40
3:D:37:ASN:O	3:D:56:TRP:HA	2.22	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	282/310 (91%)	256 (91%)	21 (7%)	5 (2%)	11	55
2	C	114/116 (98%)	103 (90%)	9 (8%)	2 (2%)	11	55
3	D	112/114 (98%)	103 (92%)	7 (6%)	2 (2%)	11	55
4	A	19/21 (90%)	16 (84%)	2 (10%)	1 (5%)	2	30
5	B	13/26 (50%)	11 (85%)	1 (8%)	1 (8%)	1	20
6	F	9/23 (39%)	6 (67%)	3 (33%)	0	100	100
All	All	549/610 (90%)	495 (90%)	43 (8%)	11 (2%)	9	53

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	104	PRO

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	115	GLY
4	A	13	LEU
5	B	8	GLY
1	E	100	HIS
1	E	153	GLU
3	D	57	ALA
1	E	15	ASN
1	E	124	GLU
2	C	100	TYR
3	D	83	SER

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	268/286 (94%)	261 (97%)	7 (3%)	54	81
2	C	99/99 (100%)	96 (97%)	3 (3%)	48	78
3	D	100/100 (100%)	93 (93%)	7 (7%)	19	58
4	A	20/20 (100%)	15 (75%)	5 (25%)	1	7
5	B	12/21 (57%)	12 (100%)	0	100	100
6	F	11/23 (48%)	10 (91%)	1 (9%)	12	46
All	All	510/549 (93%)	487 (96%)	23 (4%)	34	71

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	8	CYS
1	E	26	CYS
1	E	86	ARG
1	E	177	GLN
1	E	239	GLU
1	E	250	ASP
1	E	310	LYS
2	C	63	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	102	SER
2	C	110	GLN
3	D	2	ILE
3	D	31	TYR
3	D	62	SER
3	D	71	SER
3	D	82	SER
3	D	99	ARG
3	D	113	ARG
4	A	7	CYS
4	A	13	LEU
4	A	15	GLN
4	A	16	LEU
4	A	19	TYR
6	F	713	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	15	ASN
1	E	34	GLN
1	E	218	GLN
1	E	263	HIS
3	D	95	GLN
4	A	21	ASN

### 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 ⓘ

7 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	$\# Z  > 2$
8	NAG	E	402	1,8	14,14,15	0.47	0	15,19,21	1.56	3 (20%)
8	NAG	E	403	8	14,14,15	0.52	0	15,19,21	1.49	3 (20%)
8	BMA	E	404	8	11,11,12	0.97	0	14,15,17	2.46	2 (14%)
8	MAN	E	405	8	11,11,12	0.76	0	14,15,17	1.64	2 (14%)
9	NAG	E	407	1,9	14,14,15	0.61	0	15,19,21	1.15	1 (6%)
9	NAG	E	408	9	14,14,15	0.79	0	15,19,21	3.09	4 (26%)
9	BMA	E	409	9	11,11,12	0.69	0	14,15,17	2.57	4 (28%)

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
8	NAG	E	402	1,8	-	0/6/23/26	0/1/1/1
8	NAG	E	403	8	-	0/6/23/26	0/1/1/1
8	BMA	E	404	8	-	0/2/19/22	0/1/1/1
8	MAN	E	405	8	-	0/2/19/22	0/1/1/1
9	NAG	E	407	1,9	-	0/6/23/26	0/1/1/1
9	NAG	E	408	9	-	0/6/23/26	0/1/1/1
9	BMA	E	409	9	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	E	403	NAG	C3-C4-C5	-3.50	104.09	110.20
9	E	408	NAG	O7-C7-C8	-3.01	116.54	122.06
8	E	403	NAG	C3-C2-N2	-2.20	105.28	110.56
8	E	402	NAG	C2-N2-C7	-2.10	120.35	123.04
8	E	405	MAN	O5-C5-C6	2.06	111.80	107.35
9	E	407	NAG	C1-O5-C5	2.06	114.87	112.25
8	E	403	NAG	O5-C5-C6	2.07	111.83	107.35
9	E	409	BMA	C3-C4-C5	2.35	114.30	110.20

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	E	402	NAG	O4-C4-C5	2.42	115.64	109.24
9	E	409	BMA	O5-C1-C2	2.48	114.88	110.86
9	E	409	BMA	C1-C2-C3	3.08	113.19	109.54
9	E	408	NAG	O4-C4-C5	3.60	118.77	109.24
8	E	402	NAG	C1-O5-C5	4.04	117.37	112.25
8	E	404	BMA	C1-C2-C3	4.17	114.47	109.54
9	E	408	NAG	C2-N2-C7	4.82	129.23	123.04
8	E	405	MAN	C1-C2-C3	5.07	115.54	109.54
8	E	404	BMA	C1-O5-C5	7.30	121.52	112.25
9	E	409	BMA	C1-O5-C5	8.13	122.56	112.25
9	E	408	NAG	C1-O5-C5	9.32	124.08	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	E	402	NAG	1	0

## 5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	NAG	E	401	1	14,14,15	0.78	0	15,19,21	2.80	4 (26%)
7	NAG	E	406	1	14,14,15	0.48	0	15,19,21	1.62	1 (6%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	NAG	E	401	1	-	0/6/23/26	0/1/1/1
7	NAG	E	406	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
7	E	401	NAG	O7-C7-C8	-3.10	116.38	122.06
7	E	401	NAG	C3-C2-N2	2.51	116.56	110.56
7	E	401	NAG	C2-N2-C7	3.02	126.92	123.04
7	E	406	NAG	C2-N2-C7	4.36	128.64	123.04
7	E	401	NAG	C1-O5-C5	8.85	123.47	112.25

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
7	E	401	NAG	1	0
7	E	406	NAG	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

### 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	E	288/310 (92%)	-0.42	1 (0%) 94 92	133, 206, 307, 357	0
2	C	116/116 (100%)	-0.02	5 (4%) 39 30	152, 255, 326, 362	0
3	D	114/114 (100%)	-0.50	0 100 100	174, 227, 294, 316	0
4	A	21/21 (100%)	-0.18	0 100 100	180, 227, 290, 307	0
5	B	15/26 (57%)	-0.98	0 100 100	180, 194, 236, 247	0
6	F	11/23 (47%)	-0.68	0 100 100	162, 177, 227, 228	0
All	All	565/610 (92%)	-0.37	6 (1%) 82 76	133, 219, 312, 362	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	11	LEU	3.1
2	C	12	VAL	2.8
1	E	307	PRO	2.5
2	C	96	ALA	2.3
2	C	35	ASN	2.3
2	C	13	ALA	2.2

### 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	NAG	E	402	14/15	0.89	0.19	-0.59	175,190,196,209	0
8	BMA	E	404	11/12	0.82	0.35	-	251,268,285,291	0
8	MAN	E	405	11/12	0.85	0.45	-	232,285,316,322	0
9	NAG	E	407	14/15	0.94	0.14	-	243,256,265,266	0
9	BMA	E	409	11/12	0.61	0.37	-	266,330,339,349	0
8	NAG	E	403	14/15	0.95	0.35	-	195,209,232,243	0
9	NAG	E	408	14/15	0.76	0.26	-	218,290,342,360	0

## 6.4 Ligands [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	E	406	14/15	0.93	0.22	-0.21	210,232,249,258	0
7	NAG	E	401	14/15	0.83	0.33	-	225,261,274,280	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.