



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 10:30 PM GMT

PDB ID : 4Y4H  
Title : Crystal structure of the mCD1d/GCK152/iNKTCR ternary complex  
Authors : Zajonc, D.M.; Yu, E.D.  
Deposited on : 2015-02-10  
Resolution : 3.10 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

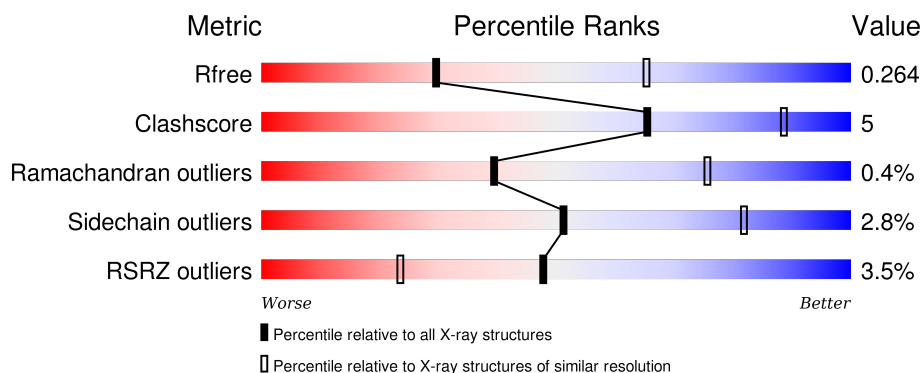
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	285	 9% 80% 13% • 6%
1	E	285	 5% 80% 13% 6%
2	B	99	 2% 81% 15% • •
2	F	99	 8% 81% 15% • •
3	C	209	 81% 15% •

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	209	
4	D	241	
4	H	241	

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
6	49X	A	306	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12468 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 Antigen-presenting glycoprotein CD1d1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	0	0	0
			2025	1287	339	386	13			
1	E	267	Total	C	N	O	S	0	0	0
			2032	1292	340	387	13			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	HIS	ASP	variant	UNP P11609
A	280	HIS	-	expression tag	UNP P11609
A	281	HIS	-	expression tag	UNP P11609
A	282	HIS	-	expression tag	UNP P11609
A	283	HIS	-	expression tag	UNP P11609
A	284	HIS	-	expression tag	UNP P11609
A	285	HIS	-	expression tag	UNP P11609
E	201	HIS	ASP	variant	UNP P11609
E	280	HIS	-	expression tag	UNP P11609
E	281	HIS	-	expression tag	UNP P11609
E	282	HIS	-	expression tag	UNP P11609
E	283	HIS	-	expression tag	UNP P11609
E	284	HIS	-	expression tag	UNP P11609
E	285	HIS	-	expression tag	UNP P11609

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	96	Total	C	N	O	S	0	0	0
			723	459	122	136	6			
2	F	96	Total	C	N	O	S	0	0	0
			744	475	126	137	6			

- Molecule 3 is a protein called Chimeric TCR Valpha14/Jalpha18 chain (mouse variable do-

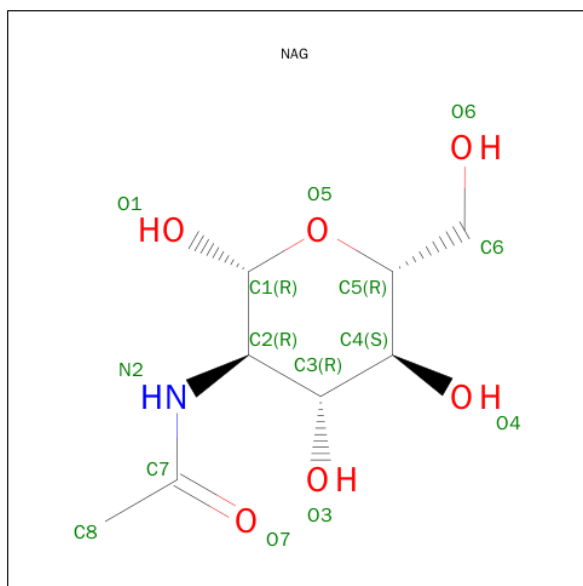
main/ human constant domain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	202	Total	C	N	O	S	0	0	0
			1524	942	260	314	8			
3	G	202	Total	C	N	O	S	0	0	0
			1514	937	258	311	8			

- Molecule 4 is a protein called Chimeric TCR Vbeta8.2 chain (mouse variable domain/ human constant domain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	239	Total	C	N	O	S	0	0	0
			1845	1157	328	354	6			
4	H	239	Total	C	N	O	S	0	0	0
			1839	1153	325	355	6			

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



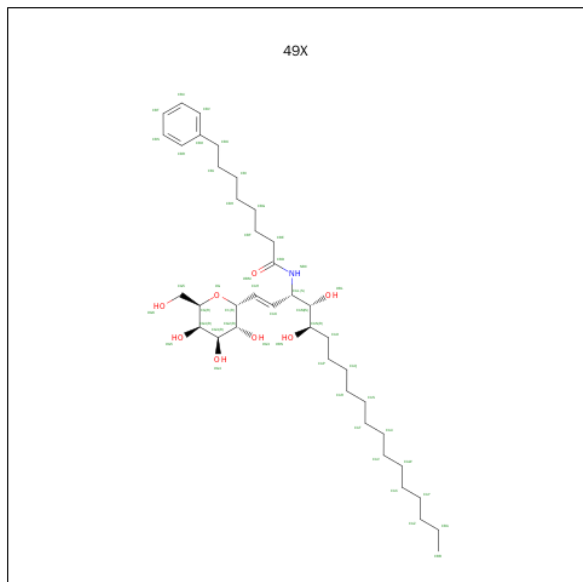
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

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

- Molecule 6 is (1R)-1,5-anhydro-1-[(1E,3S,4S,5R)-4,5-dihydroxy-3-[(8-phenyloctanoyl)amino]nonadec-1-en-1-yl]-D-galactitol (three-letter code: 49X) (formula: C<sub>39</sub>H<sub>67</sub>NO<sub>8</sub>).

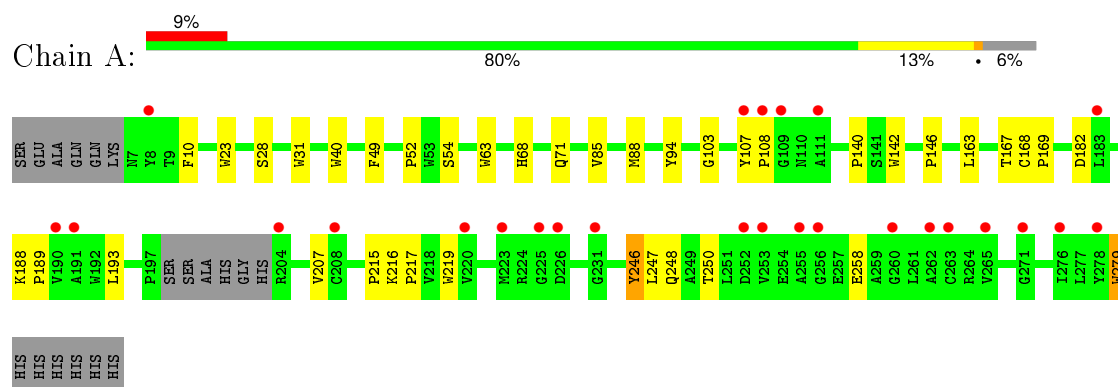


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			48	39	1	8		
6	E	1	Total	C	N	O	0	0
			48	39	1	8		

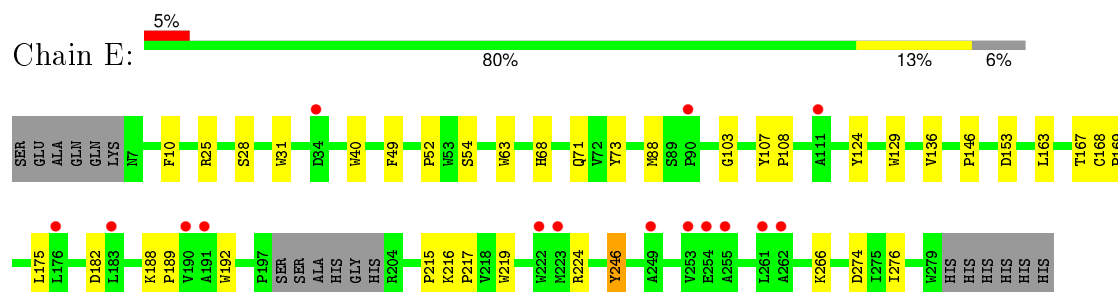
### 3 Residue-property plots [i](#)

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.

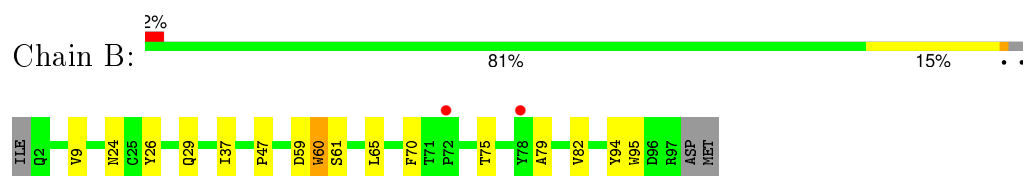
- Molecule 1: Antigen-presenting glycoprotein CD1d1



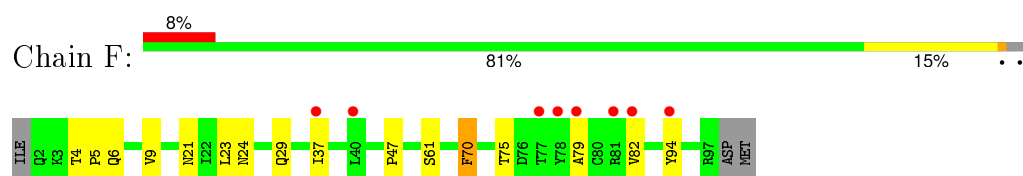
- Molecule 1: Antigen-presenting glycoprotein CD1d1



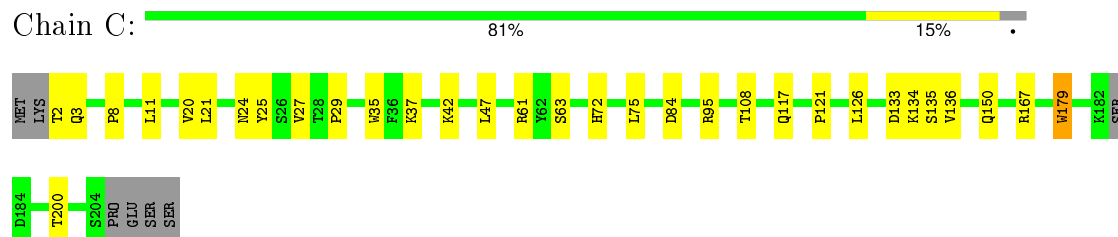
- Molecule 2: Beta-2-microglobulin



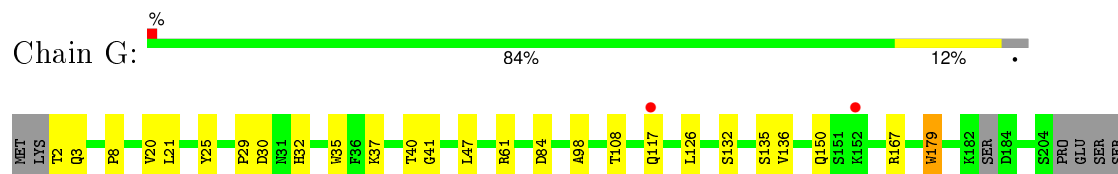
- Molecule 2: Beta-2-microglobulin



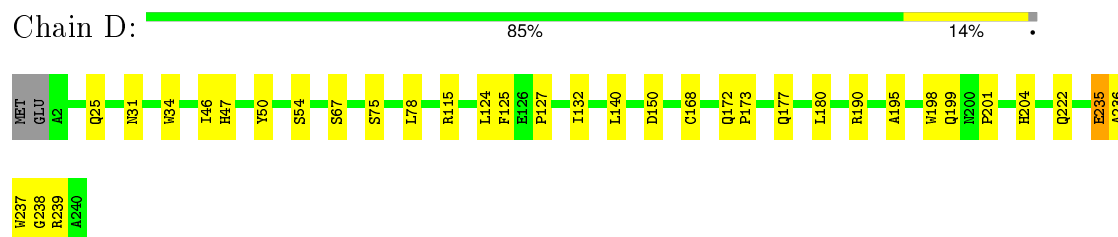
- Molecule 3: Chimeric TCR Valpha14/Jalpha18 chain (mouse variable domain/ human constant domain)



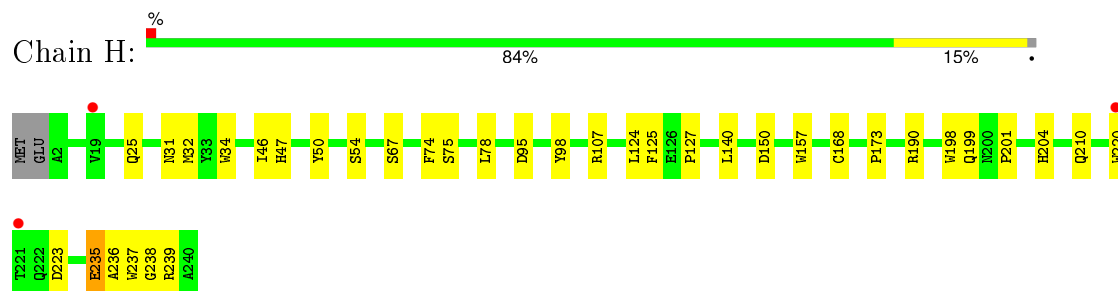
- Molecule 3: Chimeric TCR Valpha14/Jalpha18 chain (mouse variable domain/ human constant domain)



- Molecule 4: Chimeric TCR Vbeta8.2 chain (mouse variable domain/ human constant domain)



- Molecule 4: Chimeric TCR Vbeta8.2 chain (mouse variable domain/ human constant domain)





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.42Å 150.38Å 102.49Å 90.00° 96.38° 90.00°	Depositor
Resolution (Å)	38.17 – 3.10 38.17 – 3.10	Depositor EDS
% Data completeness (in resolution range)	98.6 (38.17-3.10) 98.7 (38.17-3.10)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.02 (at 3.12Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.242 , 0.287 0.212 , 0.264	Depositor DCC
$R_{free}$ test set	1334 reflections (3.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	83.9	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 33.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 42669 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12468	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 49X, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.71	6/2085 (0.3%)	0.62	0/2854
1	E	0.69	5/2092 (0.2%)	0.61	0/2861
2	B	0.65	2/749 (0.3%)	0.61	0/1031
2	F	0.60	0/770	0.58	0/1058
3	C	0.60	2/1551 (0.1%)	0.65	0/2115
3	G	0.60	2/1541 (0.1%)	0.64	0/2103
4	D	0.64	2/1896 (0.1%)	0.63	0/2593
4	H	0.66	4/1890 (0.2%)	0.63	0/2585
All	All	0.65	23/12574 (0.2%)	0.62	0/17200

All (23) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	40	TRP	CD2-CE2	5.68	1.48	1.41
4	H	220	TRP	CD2-CE2	5.67	1.48	1.41
4	D	237	TRP	CD2-CE2	5.61	1.48	1.41
3	C	179	TRP	CD2-CE2	5.57	1.48	1.41
4	H	157	TRP	CD2-CE2	5.57	1.48	1.41
4	H	237	TRP	CD2-CE2	5.57	1.48	1.41
2	B	60	TRP	CD2-CE2	5.56	1.48	1.41
1	E	63	TRP	CD2-CE2	5.49	1.48	1.41
3	G	35	TRP	CD2-CE2	5.49	1.48	1.41
4	H	34	TRP	CD2-CE2	5.44	1.47	1.41
3	G	179	TRP	CD2-CE2	5.40	1.47	1.41
4	D	34	TRP	CD2-CE2	5.36	1.47	1.41
1	A	40	TRP	CD2-CE2	5.34	1.47	1.41
1	A	31	TRP	CD2-CE2	5.26	1.47	1.41
1	E	129	TRP	CD2-CE2	5.23	1.47	1.41
1	A	63	TRP	CD2-CE2	5.23	1.47	1.41
1	A	279	TRP	CD2-CE2	5.22	1.47	1.41
1	E	192	TRP	CD2-CE2	5.22	1.47	1.41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	23	TRP	CD2-CE2	5.21	1.47	1.41
3	C	35	TRP	CD2-CE2	5.20	1.47	1.41
1	A	219	TRP	CD2-CE2	5.19	1.47	1.41
2	B	95	TRP	CD2-CE2	5.18	1.47	1.41
1	E	31	TRP	CD2-CE2	5.11	1.47	1.41

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 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	2025	0	1794	22	0
1	E	2032	0	1812	22	0
2	B	723	0	612	6	0
2	F	744	0	663	8	0
3	C	1524	0	1414	16	0
3	G	1514	0	1397	16	0
4	D	1845	0	1713	15	0
4	H	1839	0	1700	16	0
5	A	70	0	63	0	0
5	E	56	0	51	1	0
6	A	48	0	67	3	0
6	E	48	0	67	5	0
All	All	12468	0	11353	118	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 5.

All (118) 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)
4:H:204:HIS:NE2	4:H:235:GLU:HG3	1.88	0.86

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:204:HIS:NE2	4:D:235:GLU:HG3	1.90	0.86
1:A:168:CYS:HB3	1:A:169:PRO:HD3	1.67	0.75
1:E:49:PHE:CD1	1:E:54:SER:HB2	2.24	0.73
1:A:49:PHE:HD1	1:A:54:SER:CB	2.02	0.72
3:G:8:PRO:O	3:G:108:THR:HG23	1.89	0.72
1:E:219:TRP:HB3	1:E:266:LYS:HB2	1.71	0.71
1:A:49:PHE:CD1	1:A:54:SER:HB2	2.26	0.70
3:C:61:ARG:NH2	3:C:84:ASP:OD2	2.23	0.69
1:E:49:PHE:HD1	1:E:54:SER:CB	2.09	0.66
1:E:168:CYS:HB3	1:E:169:PRO:HD3	1.76	0.66
1:A:258:GLU:HB3	1:A:279:TRP:CD1	2.32	0.65
3:G:61:ARG:NH2	3:G:84:ASP:OD2	2.28	0.64
3:C:8:PRO:O	3:C:108:THR:HG23	1.96	0.64
1:A:49:PHE:HD1	1:A:54:SER:HB2	1.61	0.63
3:C:2:THR:O	3:C:2:THR:HG23	1.99	0.62
1:E:49:PHE:HD1	1:E:54:SER:HB2	1.63	0.62
1:E:219:TRP:CE3	1:E:266:LYS:HG3	2.35	0.62
1:A:68:HIS:HA	1:A:71:GLN:OE1	1.99	0.61
4:H:201:PRO:HA	4:H:238:GLY:O	2.02	0.60
4:D:235:GLU:HG2	4:D:236:ALA:N	2.17	0.60
1:A:49:PHE:HD1	1:A:54:SER:HB3	1.68	0.59
1:A:107:TYR:HB3	1:A:108:PRO:HD2	1.85	0.58
4:D:201:PRO:HA	4:D:238:GLY:O	2.03	0.58
2:F:29:GLN:HA	2:F:61:SER:HB2	1.86	0.58
3:G:2:THR:O	3:G:2:THR:HG23	2.04	0.58
3:G:20:VAL:C	3:G:21:LEU:HD12	2.25	0.57
1:E:68:HIS:HA	1:E:71:GLN:OE1	2.04	0.56
1:A:215:PRO:HB2	1:A:217:PRO:HD2	1.88	0.56
4:D:127:PRO:HD2	4:D:198:TRP:CZ2	2.41	0.55
3:C:20:VAL:C	3:C:21:LEU:HD12	2.26	0.55
3:C:126:LEU:O	3:C:135:SER:HB2	2.06	0.55
4:D:115:ARG:HB2	4:D:222:GLN:NE2	2.22	0.55
1:A:49:PHE:CD1	1:A:54:SER:CB	2.86	0.55
4:H:235:GLU:HG2	4:H:236:ALA:N	2.21	0.55
1:E:107:TYR:HB3	1:E:108:PRO:HD2	1.90	0.54
3:G:37:LYS:HB2	3:G:47:LEU:HD11	1.89	0.54
6:A:306:49X:H40	3:C:29:PRO:HB3	1.90	0.53
4:H:124:LEU:HD11	4:H:140:LEU:HD23	1.90	0.53
4:H:199:GLN:HA	4:H:239:ARG:O	2.08	0.53
3:G:126:LEU:O	3:G:135:SER:HB2	2.09	0.52
2:B:29:GLN:HA	2:B:61:SER:HB2	1.92	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:78:LEU:N	4:D:78:LEU:HD12	2.25	0.51
6:E:305:49X:H30	6:E:305:49X:NBC	2.26	0.51
1:E:88:MET:HE3	1:E:146:PRO:HD3	1.93	0.50
3:C:117:GLN:HA	3:C:117:GLN:HE21	1.76	0.50
1:A:246:TYR:CD2	1:A:246:TYR:C	2.86	0.49
4:H:204:HIS:NE2	4:H:235:GLU:CG	2.68	0.49
4:D:199:GLN:HA	4:D:239:ARG:O	2.12	0.49
2:F:37:ILE:HG12	2:F:82:VAL:HG22	1.95	0.49
2:F:9:VAL:HA	2:F:24:ASN:O	2.13	0.49
2:B:37:ILE:HG12	2:B:82:VAL:HG22	1.95	0.48
3:G:117:GLN:HE21	3:G:117:GLN:HA	1.78	0.48
4:H:31:ASN:OD1	4:H:50:TYR:HD1	1.97	0.47
4:H:78:LEU:HD12	4:H:78:LEU:N	2.29	0.47
3:G:32:HIS:CE1	4:H:98:TYR:CD1	3.03	0.47
2:B:26:TYR:HB2	2:B:65:LEU:CD1	2.44	0.47
4:D:177:GLN:HB3	4:D:180:LEU:HG	1.96	0.47
4:D:115:ARG:HB2	4:D:222:GLN:HE22	1.80	0.47
4:D:31:ASN:OD1	4:D:50:TYR:HD1	1.98	0.46
4:H:127:PRO:HD2	4:H:198:TRP:CZ2	2.50	0.46
3:C:24:ASN:OD1	3:C:72:HIS:ND1	2.49	0.46
2:B:9:VAL:HA	2:B:24:ASN:O	2.15	0.46
1:A:88:MET:HE3	1:A:146:PRO:HD3	1.96	0.46
3:G:98:ALA:HB2	4:H:95:ASP:O	2.15	0.46
1:E:153:ASP:OD2	6:E:305:49X:OG3	2.31	0.45
4:H:46:ILE:HG22	4:H:47:HIS:CD2	2.51	0.45
1:E:49:PHE:HD1	1:E:54:SER:HB3	1.78	0.45
1:A:85:VAL:HG11	1:A:94:TYR:CE1	2.52	0.45
6:E:305:49X:H40	3:G:29:PRO:HB3	1.98	0.45
6:E:305:49X:H40	3:G:29:PRO:CB	2.47	0.45
4:D:124:LEU:HD11	4:D:140:LEU:HD23	1.99	0.45
3:C:37:LYS:HB2	3:C:47:LEU:HD11	1.97	0.45
2:F:21:ASN:HB3	2:F:70:PHE:CE2	2.52	0.45
1:E:216:LYS:N	1:E:217:PRO:CD	2.80	0.45
1:E:215:PRO:HB2	1:E:217:PRO:HD2	1.98	0.44
3:C:136:VAL:HG12	3:C:179:TRP:HB3	2.00	0.44
1:A:207:VAL:HG22	1:A:250:THR:HG22	1.99	0.44
1:E:246:TYR:C	1:E:246:TYR:CD2	2.91	0.44
1:A:140:PRO:HB2	1:A:142:TRP:CD1	2.53	0.43
6:A:306:49X:H40	3:C:29:PRO:CB	2.47	0.43
1:E:49:PHE:CD1	1:E:54:SER:CB	2.90	0.43
4:D:46:ILE:HG22	4:D:47:HIS:CD2	2.54	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:163:LEU:HA	1:A:167:THR:HB	2.01	0.43
1:A:10:PHE:O	1:A:103:GLY:HA3	2.19	0.43
1:E:163:LEU:HA	1:E:167:THR:HB	1.99	0.42
6:A:306:49X:H1	3:C:95:ARG:HD3	2.01	0.42
3:G:136:VAL:HG12	3:G:179:TRP:HB3	2.01	0.42
1:E:73:TYR:CD1	6:E:305:49X:H56	2.54	0.42
2:B:79:ALA:HB2	2:B:94:TYR:CD2	2.54	0.42
1:E:188:LYS:HA	1:E:189:PRO:HD2	1.90	0.42
2:F:9:VAL:HG12	2:F:23:LEU:HD11	2.01	0.42
3:C:133:ASP:OD1	3:C:134:LYS:N	2.53	0.42
4:H:125:PHE:N	4:H:125:PHE:CD1	2.87	0.42
4:H:150:ASP:HB2	4:H:173:PRO:HG2	2.02	0.42
2:B:59:ASP:O	2:B:60:TRP:HB2	2.20	0.41
2:F:6:GLN:HA	2:F:6:GLN:OE1	2.21	0.41
2:F:79:ALA:HB2	2:F:94:TYR:CD2	2.55	0.41
4:D:125:PHE:N	4:D:125:PHE:CD1	2.89	0.41
1:E:124:TYR:CZ	1:E:136:VAL:HG11	2.56	0.41
1:A:247:LEU:HD12	1:A:248:GLN:H	1.86	0.41
1:A:188:LYS:HA	1:A:189:PRO:HD2	1.88	0.41
3:G:21:LEU:HD12	3:G:21:LEU:N	2.36	0.41
3:G:41:GLY:CA	4:H:107:ARG:HH22	2.34	0.41
3:G:3:GLN:O	3:G:25:TYR:HA	2.21	0.41
1:E:25:ARG:HB3	5:E:302:NAG:H82	2.02	0.41
3:C:63:SER:O	3:C:75:LEU:HD12	2.21	0.41
1:A:52:PRO:HB3	1:E:52:PRO:O	2.21	0.41
1:A:216:LYS:N	1:A:217:PRO:CD	2.84	0.41
4:H:32:MET:SD	4:H:74:PHE:HB2	2.61	0.41
1:A:168:CYS:HB3	1:A:169:PRO:CD	2.46	0.40
3:C:3:GLN:O	3:C:25:TYR:HA	2.20	0.40
1:E:10:PHE:O	1:E:103:GLY:HA3	2.21	0.40
3:C:121:PRO:HB2	3:C:200:THR:HG23	2.03	0.40
2:F:4:THR:HA	2:F:5:PRO:HD3	1.90	0.40
4:D:132:ILE:HG23	4:D:195:ALA:HB1	2.02	0.40
3:G:30:ASP:OD1	3:G:30:ASP:N	2.55	0.40
4:D:150:ASP:HB2	4:D:173:PRO:HG2	2.03	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	A	263/285 (92%)	248 (94%)	15 (6%)	0	100	100
1	E	263/285 (92%)	249 (95%)	14 (5%)	0	100	100
2	B	94/99 (95%)	91 (97%)	2 (2%)	1 (1%)	17	55
2	F	94/99 (95%)	91 (97%)	2 (2%)	1 (1%)	17	55
3	C	198/209 (95%)	183 (92%)	12 (6%)	3 (2%)	13	46
3	G	198/209 (95%)	181 (91%)	15 (8%)	2 (1%)	19	58
4	D	237/241 (98%)	230 (97%)	7 (3%)	0	100	100
4	H	237/241 (98%)	229 (97%)	8 (3%)	0	100	100
All	All	1584/1668 (95%)	1502 (95%)	75 (5%)	7 (0%)	39	75

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	150	GLN
3	C	42	LYS
3	C	150	GLN
2	B	47	PRO
2	F	47	PRO
3	G	132	SER
3	C	27	VAL

### 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	204/249 (82%)	200 (98%)	4 (2%)	63	86
1	E	206/249 (83%)	199 (97%)	7 (3%)	44	79
2	B	72/93 (77%)	70 (97%)	2 (3%)	51	82
2	F	78/93 (84%)	76 (97%)	2 (3%)	54	83
3	C	171/188 (91%)	169 (99%)	2 (1%)	78	92
3	G	168/188 (89%)	166 (99%)	2 (1%)	78	92
4	D	195/208 (94%)	187 (96%)	8 (4%)	37	74
4	H	194/208 (93%)	185 (95%)	9 (5%)	33	70
All	All	1288/1476 (87%)	1252 (97%)	36 (3%)	51	82

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	SER
1	A	182	ASP
1	A	193	LEU
1	A	246	TYR
2	B	70	PHE
2	B	75	THR
3	C	11	LEU
3	C	167	ARG
4	D	25	GLN
4	D	54	SER
4	D	67	SER
4	D	75	SER
4	D	168	CYS
4	D	172	GLN
4	D	190	ARG
4	D	235	GLU
1	E	28	SER
1	E	175	LEU
1	E	182	ASP
1	E	224	ARG
1	E	246	TYR
1	E	274	ASP
1	E	276	ILE
2	F	70	PHE
2	F	75	THR
3	G	40	THR
3	G	167	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
4	H	25	GLN
4	H	54	SER
4	H	67	SER
4	H	75	SER
4	H	168	CYS
4	H	190	ARG
4	H	210	GLN
4	H	223	ASP
4	H	235	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	117	HIS
2	B	13	HIS
3	C	31	ASN
3	C	117	GLN
4	D	24	ASN
4	D	222	GLN
4	D	230	GLN
1	E	117	HIS
2	F	13	HIS
3	G	31	ASN
3	G	117	GLN
4	H	24	ASN
4	H	230	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

11 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
5	NAG	A	301	1	14,14,15	0.49	0	15,19,21	1.21	1 (6%)
5	NAG	A	302	1,5	14,14,15	0.48	0	15,19,21	2.25	2 (13%)
5	NAG	A	303	5	14,14,15	0.66	0	15,19,21	1.10	1 (6%)
5	NAG	A	304	1,5	14,14,15	0.69	0	15,19,21	2.03	5 (33%)
5	NAG	A	305	5	14,14,15	0.55	0	15,19,21	1.37	2 (13%)
6	49X	A	306	-	48,49,49	0.86	2 (4%)	52,60,60	1.24	4 (7%)
5	NAG	E	301	1	14,14,15	0.70	0	15,19,21	1.10	1 (6%)
5	NAG	E	302	1	14,14,15	0.50	0	15,19,21	1.88	3 (20%)
5	NAG	E	303	1,5	14,14,15	0.41	0	15,19,21	1.26	2 (13%)
5	NAG	E	304	5	14,14,15	0.59	0	15,19,21	1.06	0
6	49X	E	305	-	48,49,49	0.83	2 (4%)	52,60,60	1.04	4 (7%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	A	301	1	-	0/6/23/26	0/1/1/1
5	NAG	A	302	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	303	5	-	0/6/23/26	0/1/1/1
5	NAG	A	304	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	305	5	-	0/6/23/26	0/1/1/1
6	49X	A	306	-	-	0/43/63/63	0/2/2/2
5	NAG	E	301	1	-	0/6/23/26	0/1/1/1
5	NAG	E	302	1	-	0/6/23/26	0/1/1/1
5	NAG	E	303	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	304	5	-	0/6/23/26	0/1/1/1
6	49X	E	305	-	-	0/43/63/63	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	305	49X	C1-CAH	-2.54	1.40	1.50
6	A	306	49X	C1-CAH	-2.32	1.41	1.50
6	E	305	49X	CAK-CAH	2.42	1.40	1.32
6	A	306	49X	CAK-CAH	2.65	1.40	1.32

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	306	49X	CAP-CAO-CAN	-4.51	106.12	114.20
5	A	304	NAG	C3-C4-C5	-3.57	103.97	110.20
6	A	306	49X	OBL-CAM-CAN	-3.34	100.33	108.75
6	A	306	49X	CG-OG-C1	-3.34	107.62	113.18
5	E	302	NAG	C3-C4-C5	-3.33	104.40	110.20
6	E	305	49X	OG-C1-CG2	-2.85	106.35	109.13
5	E	302	NAG	C4-C3-C2	-2.52	107.32	111.23
6	E	305	49X	OG6-CG5-CG	-2.44	103.28	111.33
5	A	304	NAG	O7-C7-C8	-2.18	118.06	122.06
6	E	305	49X	CAP-CAO-CAN	-2.05	110.52	114.20
5	A	302	NAG	O3-C3-C2	2.11	113.30	109.11
5	A	304	NAG	C4-C3-C2	2.15	114.56	111.23
6	A	306	49X	OG-CG-CG5	2.18	111.87	106.36
5	E	303	NAG	C4-C3-C2	2.27	114.75	111.23
6	E	305	49X	CG5-CG-CG4	2.27	118.61	113.02
5	E	301	NAG	C4-C3-C2	2.51	115.14	111.23
5	A	305	NAG	O5-C5-C6	2.72	113.23	107.35
5	A	303	NAG	C4-C3-C2	2.73	115.47	111.23
5	A	305	NAG	C4-C3-C2	3.14	116.11	111.23
5	E	303	NAG	C1-O5-C5	3.24	116.35	112.25
5	A	301	NAG	O5-C5-C6	3.37	114.65	107.35
5	A	304	NAG	C2-N2-C7	3.87	128.02	123.04
5	A	304	NAG	C6-C5-C4	4.36	123.77	113.02
5	E	302	NAG	C1-O5-C5	4.80	118.34	112.25
5	A	302	NAG	C1-O5-C5	7.73	122.06	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	306	49X	3	0
5	E	302	NAG	1	0
6	E	305	49X	5	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	A	267/285 (93%)	0.33	26 (9%) 10 3	39, 79, 170, 204	0
1	E	267/285 (93%)	0.03	15 (5%) 28 11	40, 76, 136, 169	0
2	B	96/99 (96%)	-0.13	2 (2%) 67 44	47, 81, 111, 125	0
2	F	96/99 (96%)	0.37	8 (8%) 14 5	59, 95, 118, 127	0
3	C	202/209 (96%)	-0.22	0 100 100	36, 57, 114, 148	0
3	G	202/209 (96%)	-0.28	2 (0%) 84 69	34, 57, 112, 153	0
4	D	239/241 (99%)	-0.26	0 100 100	40, 60, 108, 123	0
4	H	239/241 (99%)	-0.32	3 (1%) 79 62	36, 60, 97, 123	0
All	All	1608/1668 (96%)	-0.08	56 (3%) 48 23	34, 67, 128, 204	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	253	VAL	7.5
1	E	262	ALA	6.5
1	A	191	ALA	6.4
2	F	79	ALA	6.3
1	A	256	GLY	6.1
1	A	223	MET	4.9
1	E	90	PRO	4.9
1	A	271	GLY	4.1
1	A	263	CYS	4.1
1	A	276	ILE	4.0
2	F	77	THR	3.8
1	E	222	TRP	3.8
1	A	255	ALA	3.7
1	A	220	VAL	3.5
1	E	253	VAL	3.4
1	E	261	LEU	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	111	ALA	3.3
4	H	221	THR	3.2
1	E	34	ASP	3.2
1	E	254	GLU	3.2
1	A	107	TYR	3.2
1	E	191	ALA	3.2
2	F	40	LEU	3.0
2	F	78	TYR	3.0
1	A	260	GLY	3.0
1	A	108	PRO	2.9
1	A	265	VAL	2.9
1	A	231	GLY	2.8
2	F	37	ILE	2.7
1	E	255	ALA	2.7
1	A	226	ASP	2.6
1	A	225	GLY	2.6
1	E	249	ALA	2.6
1	A	278	TYR	2.5
1	A	208	CYS	2.5
1	A	262	ALA	2.5
4	H	220	TRP	2.4
1	E	183	LEU	2.4
1	A	252	ASP	2.3
1	E	111	ALA	2.3
4	H	19	VAL	2.3
1	E	190	VAL	2.3
2	B	72	PRO	2.3
3	G	152	LYS	2.3
3	G	117	GLN	2.2
1	E	176	LEU	2.2
2	F	82	VAL	2.2
2	F	81	ARG	2.2
1	A	8	TYR	2.2
2	B	78	TYR	2.1
2	F	94	TYR	2.1
1	A	109	GLY	2.0
1	A	183	LEU	2.0
1	A	190	VAL	2.0
1	A	204	ARG	2.0
1	E	223	MET	2.0

## 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](#)

There are no carbohydrates in this entry.

## 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(Å <sup>2</sup> )	Q<0.9
6	49X	A	306	48/48	0.95	0.30	2.71	46,57,75,77	0
6	49X	E	305	48/48	0.95	0.30	1.98	42,53,64,68	0
5	NAG	E	302	14/15	0.89	0.27	1.75	69,70,74,79	0
5	NAG	A	302	14/15	0.92	0.25	0.59	64,69,75,90	0
5	NAG	A	301	14/15	0.88	0.21	-	90,97,106,109	0
5	NAG	E	301	14/15	0.89	0.24	-	82,100,106,112	0
5	NAG	A	303	14/15	0.86	0.27	-	99,111,125,126	0
5	NAG	A	304	14/15	0.90	0.16	-	51,63,70,71	0
5	NAG	E	304	14/15	0.87	0.17	-	76,87,94,95	0
5	NAG	E	303	14/15	0.94	0.19	-	57,62,71,72	0
5	NAG	A	305	14/15	0.93	0.23	-	74,86,93,94	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.