wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 11:53 PM GMT

PDB ID	:	1YPZ
Title	:	Immune receptor
Authors	:	Adams, E.J.; Garcia, K.C.
Deposited on	:	2005-01-31
Resolution	:	3.40 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 (i) symbol.

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

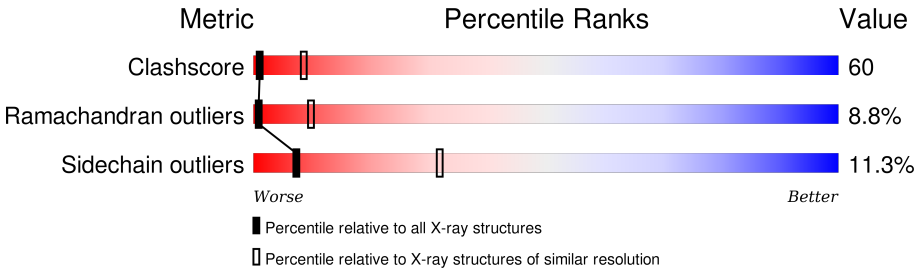
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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))
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)

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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	260	27% 58% 15%
1	C	260	22% 68% 9%
2	B	102	37% 53% 10%
2	D	102	24% 63% 13%
3	E	207	31% 57% 11%
3	G	207	42% 47% 10%
4	F	230	28% 55% 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	H	230	

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
5	NAG	E	208	X	-	-	-
5	NAG	G	208	X	-	-	-
6	NAG	E	211	X	-	X	-
6	NAG	F	231	-	-	X	-
7	NAG	G	212	X	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12984 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 H2-T22 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			2106	1330	363	402	11			
1	C	260	Total	C	N	O	S	0	0	0
			2106	1330	363	402	11			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	102	Total	C	N	O	S	0	0	0
			849	540	143	163	3			
2	D	102	Total	C	N	O	S	0	0	0
			849	540	143	163	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	ASP	-	CLONING ARTIFACT	UNP P61769
B	0	PRO	-	CLONING ARTIFACT	UNP P61769
D	-1	ASP	-	CLONING ARTIFACT	UNP P61769
D	0	PRO	-	CLONING ARTIFACT	UNP P61769

- Molecule 3 is a protein called T cell receptor delta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	207	Total	C	N	O	S	0	0	0
			1603	1000	276	321	6			
3	G	207	Total	C	N	O	S	281	0	0
			1592	992	274	321	5			

- Molecule 4 is a protein called T-cell receptor gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	230	Total	C	N	O	S	0	0	0
			1852	1187	305	353	7			
4	H	230	Total	C	N	O	S	371	0	0
			1844	1181	303	353	7			

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	E	3	Total	C	N	O	0	0
			42	22	2	18		
5	G	3	Total	C	N	O	0	0
			42	22	2	18		

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	E	2	Total	C	N	O	0	0
			30	16	2	12		
6	F	2	Total	C	N	O	0	0
			30	16	2	12		

- Molecule 7 is a polymer of unknown type called SUGAR (3-MER).

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

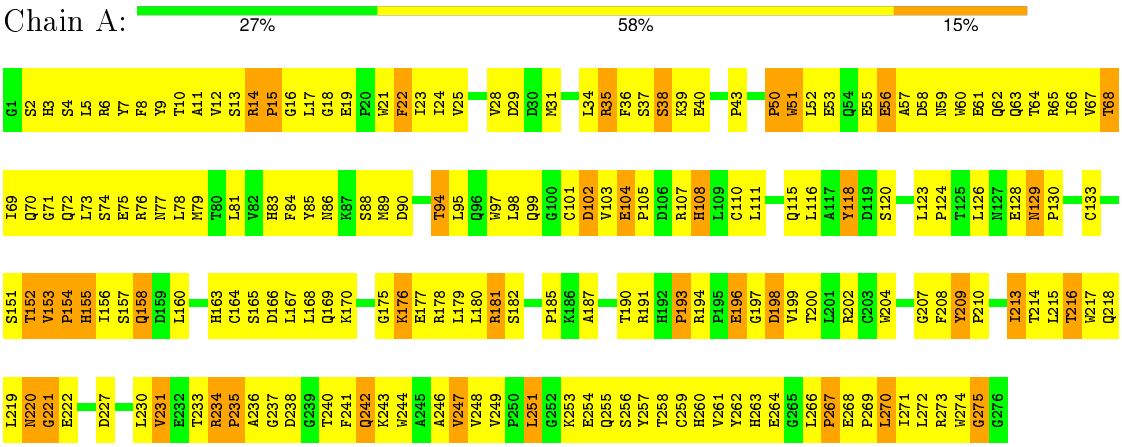
3 Residue-property plots (i)

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.

Note EDS was not executed.

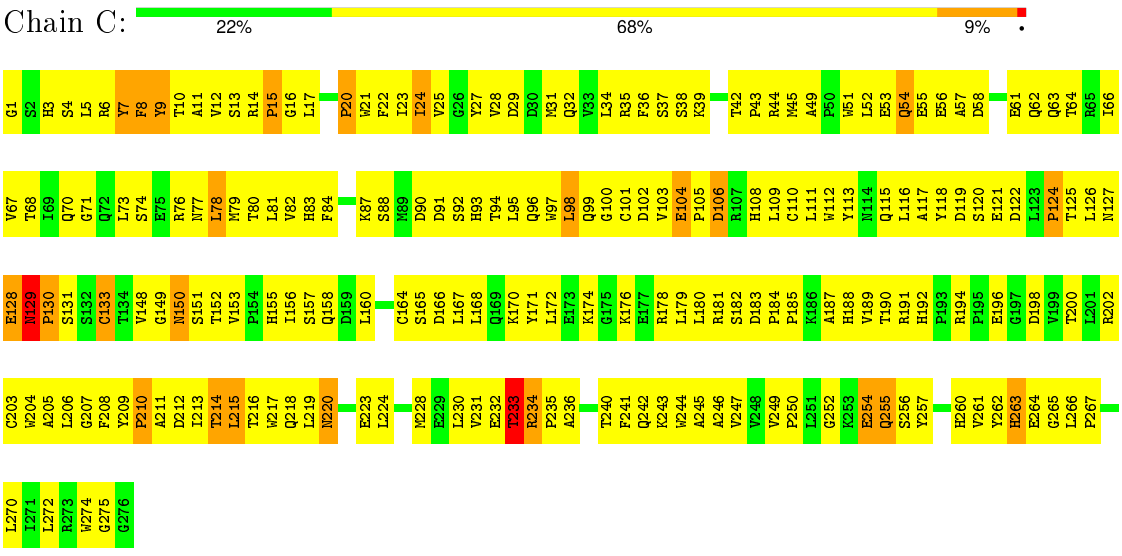
- Molecule 1: H2-T22 protein

Chain A:



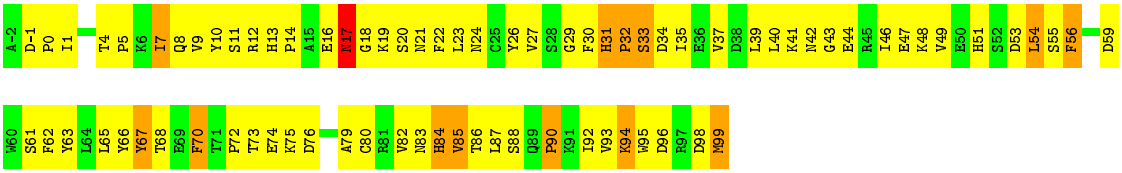
- Molecule 1: H2-T22 protein

Chain C:



- Molecule 2: Beta-2-microglobulin

Chain B:

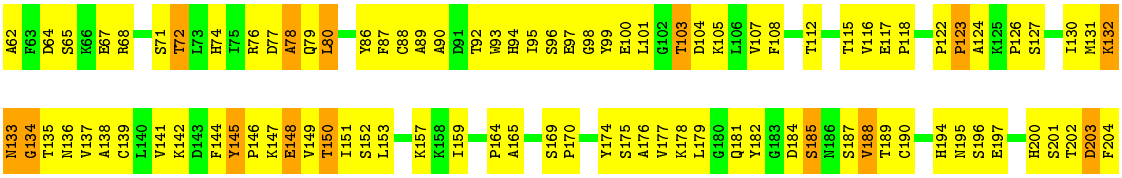


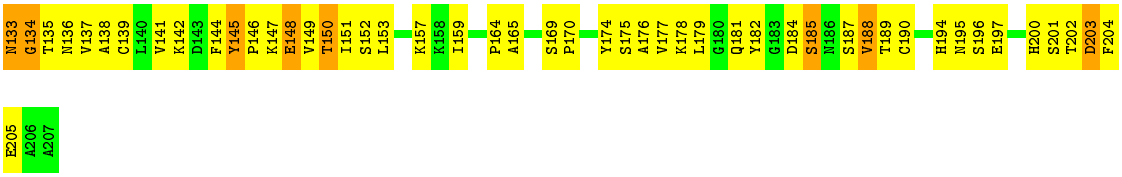
• Molecule 2: Beta-2-microglobulin

Chain D:

• Molecule 3: T cell receptor delta

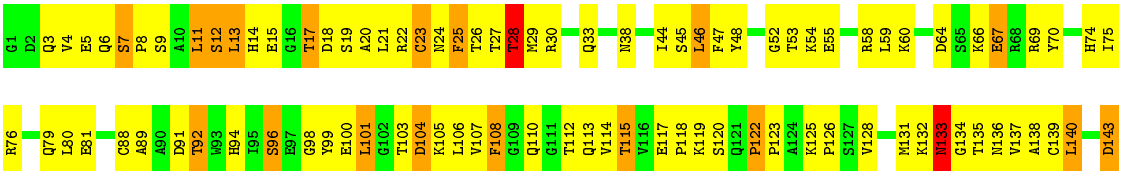
Chain E:

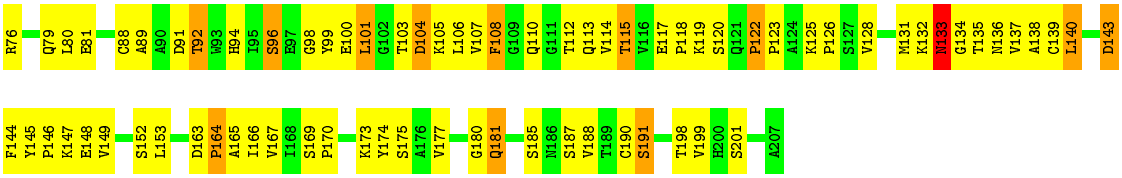




• Molecule 3: T cell receptor delta

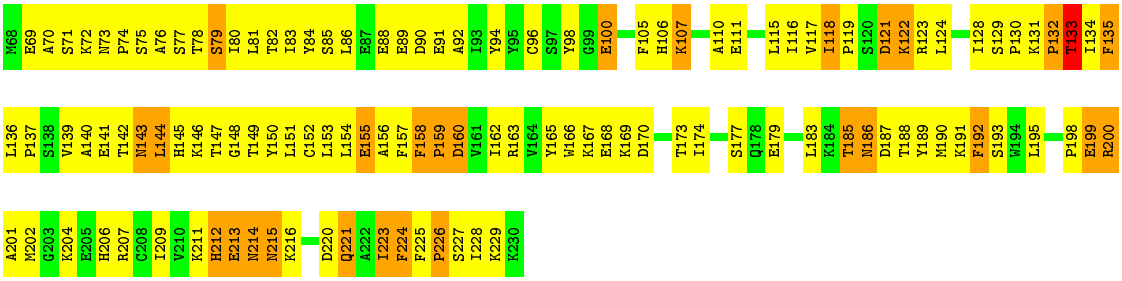
Chain G:





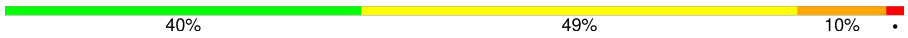
• Molecule 4: T-cell receptor gamma chain

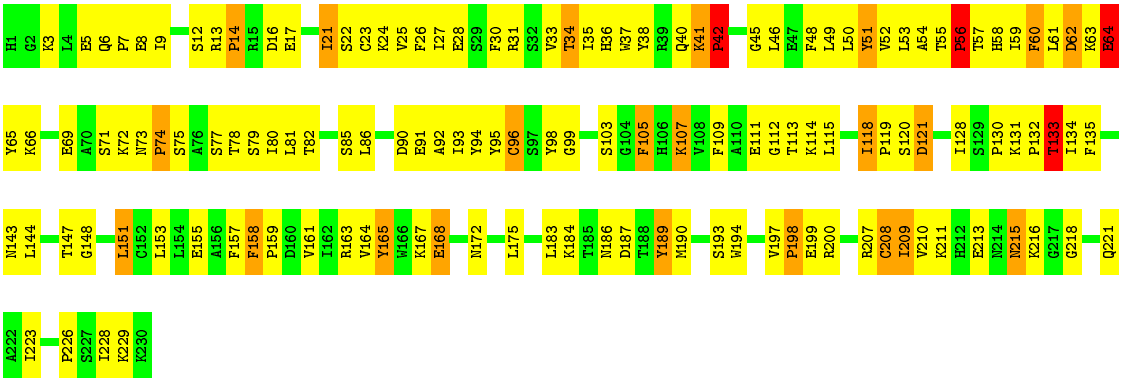
Chain F:



- Molecule 4: T-cell receptor gamma chain

Chain H:





4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	110.53Å 113.05Å 167.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.40	Depositor
% Data completeness (in resolution range)	97.4 (20.00-3.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.264 , 0.330	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12984	wwPDB-VP
Average B, all atoms (Å^2)	103.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FUC, 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	A	0.58	0/2166	0.89	2/2951 (0.1%)
1	C	0.48	0/2166	0.82	0/2951
2	B	0.46	0/873	0.81	0/1182
2	D	0.43	0/873	0.70	0/1182
3	E	0.55	0/1636	0.86	0/2212
3	G	0.50	0/1625	0.82	1/2201 (0.0%)
4	F	0.51	0/1898	0.84	0/2565
4	H	0.45	0/1889	0.82	0/2553
All	All	0.51	0/13126	0.83	3/17797 (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
2	D	0	1
5	E	1	0
5	G	1	0
6	E	1	0
7	G	1	0
All	All	4	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	46	LEU	CA-CB-CG	-5.97	101.58	115.30
1	A	155	HIS	CB-CA-C	-5.69	99.03	110.40
1	A	154	PRO	CA-N-CD	-5.56	103.72	111.50

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	E	208	NAG	C1
6	E	211	NAG	C1
5	G	208	NAG	C1
7	G	212	NAG	C1

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	63	TYR	Sidechain

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	2106	0	2006	275	1
1	C	2106	0	2009	275	4
2	B	849	0	810	83	0
2	D	849	0	810	121	0
3	E	1603	0	1554	195	1
3	G	1592	0	1524	162	0
4	F	1852	0	1832	233	4
4	H	1844	0	1816	162	0
5	E	42	0	38	8	0
5	G	42	0	36	5	0
6	E	30	0	27	8	0
6	F	30	0	27	20	0
7	G	39	0	36	12	0
All	All	12984	0	12525	1435	5

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 60.

The worst 5 of 1435 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)
3:G:24:ASN:HD21	7:G:212:NAG:C1	1.10	1.58
4:F:186:ASN:HD22	6:F:231:NAG:C1	0.99	1.57
3:E:133:ASN:HD21	5:E:208:NAG:C1	0.88	1.53
1:A:155:HIS:CD2	1:A:157:SER:N	1.88	1.41
1:A:219:LEU:O	1:A:221:GLY:N	1.58	1.32

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:255:GLN:NE2	4:F:66:LYS:CE[1_545]	1.35	0.85
1:A:53:GLU:OE2	3:E:157:LYS:NZ[1_455]	1.36	0.84
1:C:255:GLN:CD	4:F:66:LYS:NZ[1_545]	1.96	0.24
1:C:255:GLN:CD	4:F:66:LYS:CE[1_545]	2.13	0.07
1:C:255:GLN:NE2	4:F:66:LYS:NZ[1_545]	2.17	0.03

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	258/260 (99%)	181 (70%)	55 (21%)	22 (8%)	1 11
1	C	258/260 (99%)	192 (74%)	45 (17%)	21 (8%)	1 11
2	B	100/102 (98%)	86 (86%)	9 (9%)	5 (5%)	3 24
2	D	100/102 (98%)	79 (79%)	12 (12%)	9 (9%)	1 9
3	E	205/207 (99%)	157 (77%)	30 (15%)	18 (9%)	1 10
3	G	205/207 (99%)	158 (77%)	29 (14%)	18 (9%)	1 10
4	F	228/230 (99%)	167 (73%)	37 (16%)	24 (10%)	1 7
4	H	228/230 (99%)	172 (75%)	34 (15%)	22 (10%)	1 8
All	All	1582/1598 (99%)	1192 (75%)	251 (16%)	139 (9%)	1 10

5 of 139 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	14	ARG
1	A	15	PRO
1	A	51	TRP
1	A	56	GLU
1	A	129	ASN

5.3.2 Protein sidechains (i)

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	233/233 (100%)	207 (89%)	26 (11%)	7 32
1	C	233/233 (100%)	210 (90%)	23 (10%)	10 39
2	B	96/96 (100%)	86 (90%)	10 (10%)	9 37
2	D	96/96 (100%)	86 (90%)	10 (10%)	9 37
3	E	180/180 (100%)	167 (93%)	13 (7%)	18 56
3	G	177/180 (98%)	155 (88%)	22 (12%)	6 27
4	F	206/206 (100%)	177 (86%)	29 (14%)	4 22
4	H	204/206 (99%)	176 (86%)	28 (14%)	4 23
All	All	1425/1430 (100%)	1264 (89%)	161 (11%)	7 32

5 of 161 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	25	PHE
4	F	84	TYR
4	H	161	VAL
3	E	32	VAL
3	E	152	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	42	ASN
3	E	121	GLN
4	H	106	HIS
3	E	6	GLN
3	E	33	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

13 carbohydrates are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	NAG	E	208	3,5	15,15,15	0.54	0	17,21,21	0.67	0
5	NAG	E	209	5	15,15,15	0.46	0	17,21,21	0.59	0
5	MAN	E	210	5,4	12,12,12	0.41	0	17,17,17	0.47	0
6	NAG	E	211	3,6	15,15,15	0.54	0	17,21,21	0.70	0
6	NAG	E	212	3,6	15,15,15	0.47	0	17,21,21	0.60	0
6	NAG	F	231	4,6	15,15,15	0.54	0	17,21,21	0.83	1 (5%)
6	NAG	F	232	6	15,15,15	0.47	0	17,21,21	0.59	0
5	NAG	G	208	3,5,4	15,15,15	0.55	0	17,21,21	0.74	0
5	NAG	G	209	5	15,15,15	0.47	0	17,21,21	0.59	0
5	MAN	G	210	5	12,12,12	0.41	0	17,17,17	0.47	0
7	FUC	G	211	7	10,10,11	1.00	1 (10%)	14,14,16	2.53	4 (28%)
7	NAG	G	212	3,7	15,15,15	1.23	1 (6%)	17,21,21	1.60	3 (17%)
7	NAG	G	213	7	14,14,15	0.77	0	15,19,21	0.88	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	E	208	3,5	1/1/6/7	0/6/26/26	0/1/1/1
5	NAG	E	209	5	-	0/6/26/26	0/1/1/1
5	MAN	E	210	5,4	-	0/2/22/22	0/1/1/1
6	NAG	E	211	3,6	1/1/6/7	0/6/26/26	0/1/1/1
6	NAG	E	212	3,6	-	0/6/26/26	0/1/1/1
6	NAG	F	231	4,6	-	0/6/26/26	0/1/1/1
6	NAG	F	232	6	-	0/6/26/26	0/1/1/1
5	NAG	G	208	3,5,4	1/1/6/7	0/6/26/26	0/1/1/1
5	NAG	G	209	5	-	0/6/26/26	0/1/1/1
5	MAN	G	210	5	-	0/2/22/22	0/1/1/1
7	FUC	G	211	7	-	0/0/17/20	0/1/1/1
7	NAG	G	212	3,7	1/1/6/7	0/6/26/26	0/1/1/1
7	NAG	G	213	7	-	1/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	G	212	NAG	C1-C2	-3.66	1.48	1.53
7	G	211	FUC	C1-C2	2.21	1.57	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	G	212	NAG	O4-C4-C3	-2.48	104.75	110.34
6	F	231	NAG	O1-C1-O5	-2.13	104.42	110.25
7	G	211	FUC	C2-C3-C4	2.03	114.49	111.04
7	G	211	FUC	C1-O5-C5	2.63	116.45	112.38
7	G	212	NAG	C4-C3-C2	2.64	114.09	110.43

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	G	212	NAG	C1
6	E	211	NAG	C1
5	E	208	NAG	C1
5	G	208	NAG	C1

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	G	213	NAG	O7-C7-N2-C2

There are no ring outliers.

7 monomers are involved in 53 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	208	NAG	6	0
5	E	210	MAN	2	0
6	E	211	NAG	8	0
6	F	231	NAG	20	0
5	G	208	NAG	5	0
7	G	212	NAG	10	0
7	G	213	NAG	2	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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.