



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 10:19 AM EDT

PDB ID : 7N5C  
Title : 6218 TCR in complex with H2Db PA with an engineered TCR-pMHC disulfide bond  
Authors : Szeto, C.; Gras, S.  
Deposited on : 2021-06-05  
Resolution : 1.87 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtrriage (Phenix) : 1.13  
EDS : 2.35.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

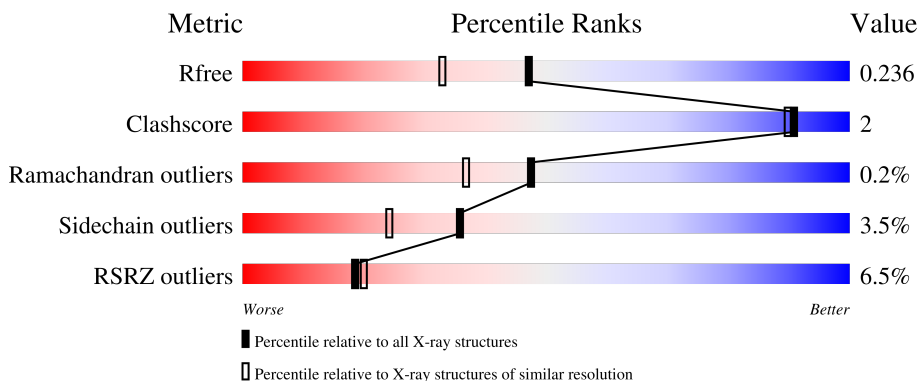
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.87 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))
$R_{free}$	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	277	 12% 92% 6%
2	B	100	 4% 95% 5%
3	C	10	 10% 80% 20%
4	D	204	 5% 87% 6% • 5%
5	E	240	 2% 93% 6%

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6898 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 H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	274	2257	1424	402	422	9	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	SER	-	expression tag	UNP P01899

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	100	837	533	141	159	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called peptide from Polymerase acidic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	10	81	51	14	15	1	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	4	CYS	GLU	engineered mutation	UNP O89752

- Molecule 4 is a protein called Fusion protein of T cell receptor alpha variable 21-DV12 with T-cell receptor, sp3.4 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	194	1512	946	252	303	11	0	1	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	107	SER	-	linker	UNP A0A075B6C4
D	108	GLY	-	linker	UNP A0A075B6C4
D	109	GLY	-	linker	UNP A0A075B6C4
D	110	CYS	-	linker	UNP A0A075B6C4
D	111	ASN	-	linker	UNP A0A075B6C4
D	112	TYR	-	linker	UNP A0A075B6C4
D	113	LYS	-	linker	UNP A0A075B6C4
D	114	LEU	-	linker	UNP A0A075B6C4
D	115	THR	-	linker	UNP A0A075B6C4
D	116	PHE	-	linker	UNP A0A075B6C4
D	117	GLY	-	linker	UNP A0A075B6C4
D	118	LYS	-	linker	UNP A0A075B6C4
D	119	GLY	-	linker	UNP A0A075B6C4
D	120	THR	-	linker	UNP A0A075B6C4
D	121	LEU	-	linker	UNP A0A075B6C4
D	122	LEU	-	linker	UNP A0A075B6C4
D	123	THR	-	linker	UNP A0A075B6C4
D	124	VAL	-	linker	UNP A0A075B6C4
D	125	THR	-	linker	UNP A0A075B6C4
D	126	PRO	-	linker	UNP A0A075B6C4
D	127	ASN	-	linker	UNP A0A075B6C4

- Molecule 5 is a protein called Fusion protein of T cell receptor beta, variable 29 and Human nkt tr beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	240	1949	1229	338	372	10	0	2	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	108	PHE	-	linker	UNP A0A0G2LB96
E	109	GLY	-	linker	UNP A0A0G2LB96
E	110	ARG	-	linker	UNP A0A0G2LB96
E	123	LEU	THR	conflict	UNP K7N5M4

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	E	1	Total 1	Na 1	0	0

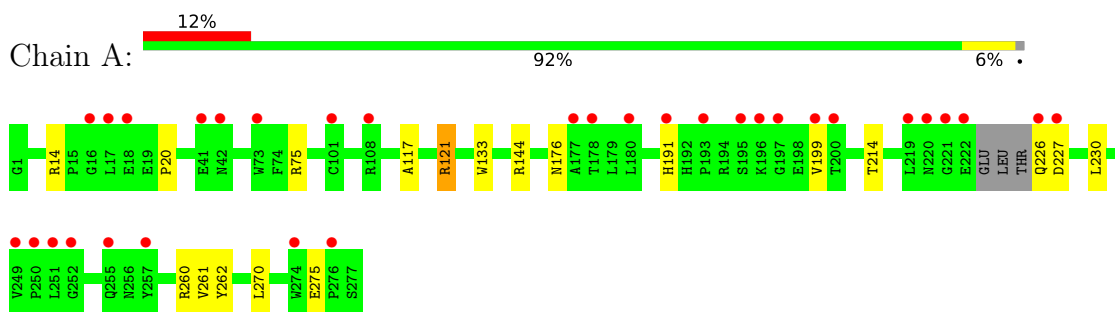
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	91	Total 91	O 91	0	0
7	B	30	Total 30	O 30	0	0
7	C	4	Total 4	O 4	0	0
7	D	51	Total 51	O 51	0	0
7	E	85	Total 85	O 85	0	0

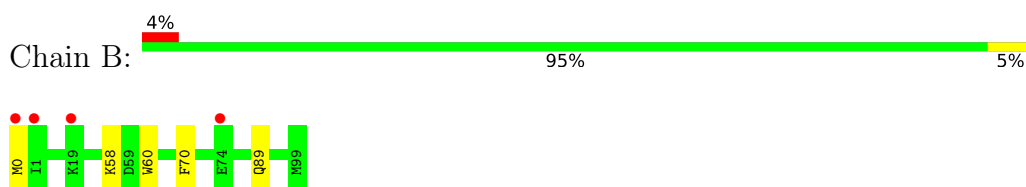
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

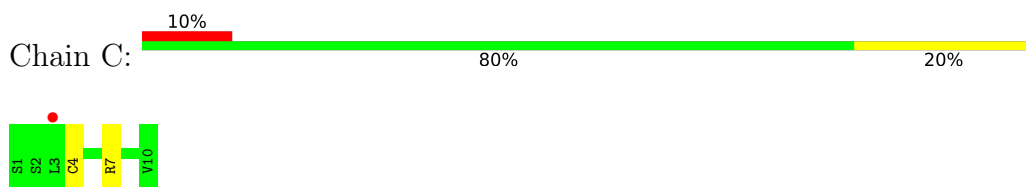
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain



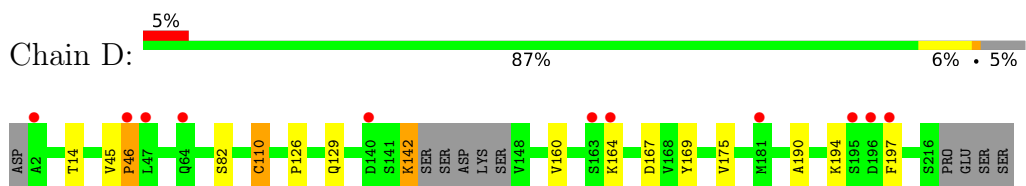
- Molecule 2: Beta-2-microglobulin



- Molecule 3: peptide from Polymerase acidic protein



- Molecule 4: Fusion protein of T cell receptor alpha variable 21-DV12 with T-cell receptor, sp3.4 alpha chain



- Molecule 5: Fusion protein of T cell receptor beta, variable 29 and Human nkt tcr beta chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.09Å 74.11Å 114.30Å 90.00° 99.34° 90.00°	Depositor
Resolution (Å)	27.14 – 1.87 27.14 – 1.87	Depositor EDS
% Data completeness (in resolution range)	100.0 (27.14-1.87) 100.0 (27.14-1.87)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 1.87Å)	Xtrriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.204 , 0.233 0.206 , 0.236	Depositor DCC
$R_{free}$ test set	3836 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.4	Xtrriage
Anisotropy	0.258	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6898	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.95% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: NA

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	0.49	0/2323	0.62	0/3152
2	B	0.47	0/860	0.67	0/1162
3	C	0.65	0/82	0.75	0/108
4	D	0.53	0/1546	0.72	0/2104
5	E	0.52	0/2006	0.67	0/2721
All	All	0.51	0/6817	0.67	0/9247

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
4	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	110[A]	CYS	Mainchain

### 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	2257	0	2128	7	0
2	B	837	0	803	2	0
3	C	81	0	79	1	0
4	D	1512	0	1438	9	0
5	E	1949	0	1868	5	0
6	E	1	0	0	0	0
7	A	91	0	0	0	0
7	B	30	0	0	0	0
7	C	4	0	0	0	0
7	D	51	0	0	0	0
7	E	85	0	0	0	0
All	All	6898	0	6316	21	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 2.

All (21) 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:D:142:LYS:HE2	4:D:142:LYS:H	1.30	0.93
1:A:133:TRP:HB2	1:A:144[B]:ARG:HG3	1.71	0.71
1:A:121:ARG:HD2	2:B:0:MET:SD	2.40	0.61
3:C:4:CYS:SG	4:D:110[B]:CYS:SG	3.01	0.56
5:E:136:VAL:HG23	5:E:246:ALA:HB3	1.87	0.56
1:A:20:PRO:HD3	1:A:75:ARG:HH11	1.71	0.55
4:D:142:LYS:HE2	4:D:142:LYS:N	2.12	0.55
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.45	0.50
4:D:126:PRO:HG3	4:D:175:VAL:HG11	1.95	0.49
5:E:19:VAL:HB	5:E:91:LEU:HB2	1.94	0.49
4:D:45:VAL:HG13	4:D:46:PRO:HD2	1.95	0.48
1:A:191:HIS:NE2	1:A:199:VAL:HG21	2.29	0.46
4:D:169:TYR:O	4:D:190:ALA:HA	2.16	0.46
5:E:15:MET:HE3	5:E:96:THR:HG23	1.98	0.45
5:E:220:GLN:HG3	5:E:243:ILE:HG23	1.98	0.45
1:A:261:VAL:HB	1:A:270:LEU:HB2	1.99	0.44
1:A:214:THR:HB	1:A:262:TYR:HB2	2.01	0.43
4:D:194:LYS:HE2	4:D:197:PHE:HB2	2.03	0.41
4:D:126:PRO:HG3	4:D:175:VAL:CG1	2.51	0.41
5:E:25:GLN:HG3	5:E:106:SER:CB	2.51	0.40
4:D:45:VAL:O	4:D:46:PRO:C	2.59	0.40

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	271/277 (98%)	268 (99%)	2 (1%)	1 (0%)	34	22
2	B	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
3	C	8/10 (80%)	8 (100%)	0	0	100	100
4	D	191/204 (94%)	184 (96%)	6 (3%)	1 (0%)	29	17
5	E	240/240 (100%)	232 (97%)	8 (3%)	0	100	100
All	All	808/831 (97%)	788 (98%)	18 (2%)	2 (0%)	47	37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	46	PRO
1	A	176	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	233/235 (99%)	226 (97%)	7 (3%)	41	30
2	B	95/95 (100%)	92 (97%)	3 (3%)	39	27
3	C	9/9 (100%)	8 (89%)	1 (11%)	6	1
4	D	174/183 (95%)	167 (96%)	7 (4%)	31	19
5	E	215/213 (101%)	208 (97%)	7 (3%)	38	26
All	All	726/735 (99%)	701 (97%)	25 (3%)	36	25

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	ARG
1	A	121	ARG
1	A	226	GLN
1	A	227	ASP
1	A	230	LEU
1	A	260	ARG
1	A	275	GLU
2	B	58	LYS
2	B	70	PHE
2	B	89	GLN
3	C	7	ARG
4	D	14	THR
4	D	82	SER
4	D	129	GLN
4	D	142	LYS
4	D	160	VAL
4	D	164	LYS
4	D	167	ASP
5	E	1	ASP
5	E	3	LYS
5	E	18	ASN
5	E	25	GLN
5	E	192	LEU
5	E	202	ARG
5	E	231	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	274/277 (98%)	0.47	32 (11%) <b>4</b> <b>5</b>	22, 41, 74, 106	2 (0%)
2	B	100/100 (100%)	0.20	4 (4%) 38 39	24, 43, 75, 86	0
3	C	10/10 (100%)	1.33	1 (10%) <b>7</b> <b>7</b>	25, 28, 31, 32	0
4	D	194/204 (95%)	0.27	11 (5%) 23 25	24, 42, 69, 91	1 (0%)
5	E	240/240 (100%)	0.04	5 (2%) 63 65	24, 33, 57, 82	0
All	All	818/831 (98%)	0.27	53 (6%) <b>18</b> <b>20</b>	22, 38, 70, 106	3 (0%)

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	197	GLY	7.4
5	E	47	GLY	5.2
5	E	253	ASP	5.0
5	E	1	ASP	4.8
1	A	195	SER	4.7
1	A	252	GLY	4.6
1	A	219	LEU	4.6
1	A	17	LEU	4.2
4	D	196	ASP	4.2
1	A	276	PRO	4.2
2	B	0	MET	4.1
1	A	251	LEU	4.1
1	A	196	LYS	4.1
1	A	227	ASP	3.8
4	D	164	LYS	3.7
1	A	18	GLU	3.6
1	A	249	VAL	3.5
4	D	47	LEU	3.4
2	B	1	ILE	3.3
1	A	41	GLU	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	180	LEU	3.0
1	A	222	GLU	2.9
1	A	274	TRP	2.9
2	B	74	GLU	2.9
3	C	3	LEU	2.9
4	D	195	SER	2.8
2	B	19	LYS	2.8
1	A	16	GLY	2.7
4	D	46	PRO	2.7
1	A	257	TYR	2.7
4	D	163	SER	2.5
4	D	64	GLN	2.5
4	D	140	ASP	2.5
1	A	42	ASN	2.5
1	A	177	ALA	2.5
1	A	73	TRP	2.5
1	A	178	THR	2.4
1	A	221	GLY	2.4
1	A	199	VAL	2.4
1	A	250	PRO	2.3
1	A	226	GLN	2.2
1	A	255	GLN	2.2
4	D	2	ALA	2.2
1	A	220	ASN	2.2
5	E	7	MET	2.2
5	E	193	ASN	2.2
4	D	197	PHE	2.1
1	A	193	PRO	2.1
1	A	191	HIS	2.1
1	A	200	THR	2.1
1	A	108	ARG	2.1
4	D	181	MET	2.0
1	A	101	CYS	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 monosaccharides 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
6	NA	E	301	1/1	0.98	0.07	32,32,32,32	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.