



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 15, 2024 – 10:19 AM EDT

PDB ID : 2FX6  
Title : bovine trypsin complexed with 2-aminobenzamidazole  
Authors : Katz, B.A.  
Deposited on : 2006-02-03  
Resolution : 1.57 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  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

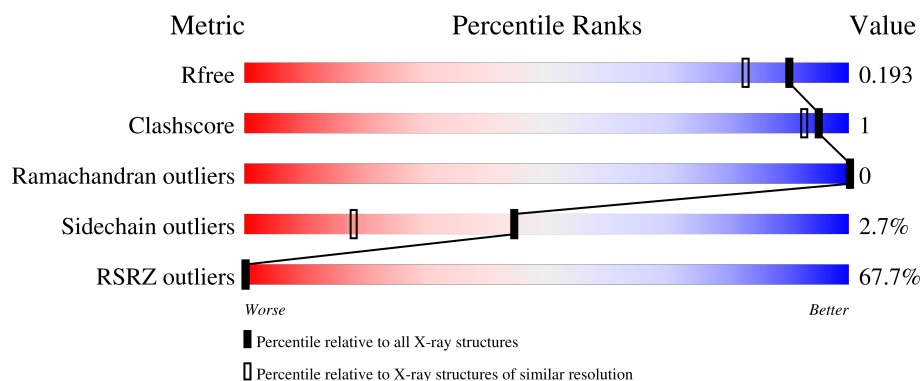
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.57 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	5534 (1.60-1.56)
Clashscore	141614	5861 (1.60-1.56)
Ramachandran outliers	138981	5708 (1.60-1.56)
Sidechain outliers	138945	5703 (1.60-1.56)
RSRZ outliers	127900	5431 (1.60-1.56)

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	223	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1829 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.

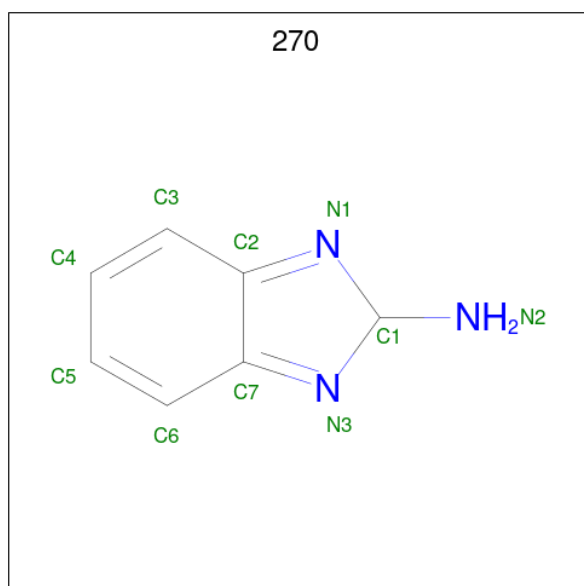
- Molecule 1 is a protein called Trypsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	13	0
			1672	1042	283	333	14			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is 2H-BENZOIMIDAZOL-2-YLAMINE (three-letter code: 270) (formula: C<sub>7</sub>H<sub>7</sub>N<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	N	0	0
			10	7	3		

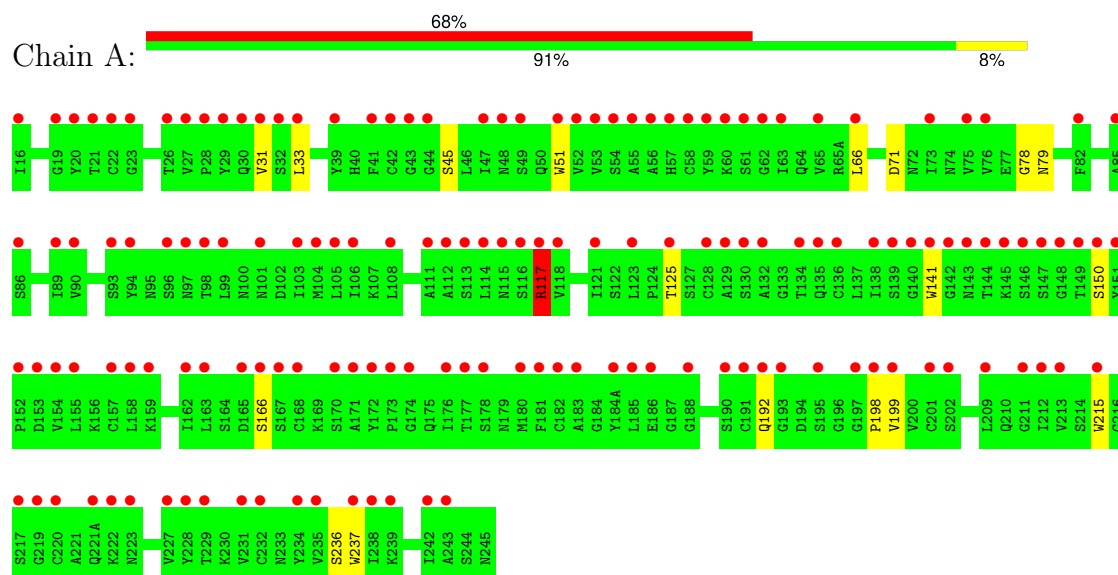
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	146	Total 146	O 146	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: Trypsin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.71Å 63.20Å 69.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.00 – 1.57 28.96 – 1.57	Depositor EDS
% Data completeness (in resolution range)	(Not available) (7.00-1.57) 99.0 (28.96-1.57)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.52 (at 1.57Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.202 , 0.215 0.189 , 0.193	Depositor DCC
$R_{free}$ test set	3993 reflections (10.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.4	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 71.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.023 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	1829	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.02% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 270, CA

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	1.25	0/1755	1.43	21/2375 (0.9%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	237	TRP	CD1-NE1-CE2	10.03	118.02	109.00
1	A	51	TRP	CD1-NE1-CE2	9.22	117.30	109.00
1	A	141	TRP	CD1-NE1-CE2	9.06	117.16	109.00
1	A	215	TRP	CD1-NE1-CE2	9.02	117.12	109.00
1	A	237	TRP	NE1-CE2-CZ2	8.34	139.57	130.40
1	A	215	TRP	NE1-CE2-CZ2	7.34	138.47	130.40
1	A	237	TRP	NE1-CE2-CD2	-7.23	100.07	107.30
1	A	237	TRP	CG-CD1-NE1	-7.00	103.10	110.10
1	A	141	TRP	CG-CD1-NE1	-6.66	103.44	110.10
1	A	215	TRP	CG-CD1-NE1	-6.46	103.64	110.10
1	A	141	TRP	NE1-CE2-CZ2	6.37	137.40	130.40
1	A	199	VAL	N-CA-C	-6.33	93.90	111.00
1	A	51	TRP	CG-CD1-NE1	-6.29	103.81	110.10
1	A	51	TRP	NE1-CE2-CZ2	6.00	137.00	130.40
1	A	215	TRP	NE1-CE2-CD2	-6.00	101.31	107.30
1	A	51	TRP	NE1-CE2-CD2	-5.71	101.59	107.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	150	SER	N-CA-C	-5.52	96.09	111.00
1	A	117	ARG	NE-CZ-NH2	-5.45	117.57	120.30
1	A	71	ASP	N-CA-C	-5.41	96.39	111.00
1	A	141	TRP	NE1-CE2-CD2	-5.30	102.00	107.30
1	A	33	LEU	N-CA-C	-5.01	97.48	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	117	ARG	Sidechain

## 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	1672	0	1637	3	0
2	A	1	0	0	0	0
3	A	10	0	6	1	0
4	A	146	0	0	1	0
All	All	1829	0	1643	4	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 1.

All (4) 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)
1:A:31:VAL:HG12	1:A:66[A]:LEU:HD23	1.91	0.52
1:A:45:SER:OG	1:A:198:PRO:HB3	2.14	0.47
1:A:78:GLY:O	1:A:79:ASN:HB2	2.20	0.41
3:A:246:270:C6	4:A:666:HOH:O	2.69	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	234/223 (105%)	227 (97%)	7 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	197/184 (107%)	190 (96%)	7 (4%)	35	10

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	117	ARG
1	A	125	THR
1	A	166[A]	SER
1	A	166[B]	SER
1	A	192	GLN
1	A	236[A]	SER
1	A	236[B]	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	270	A	246	-	8,11,11	1.76	4 (50%)	6,15,15	1.79	2 (33%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	270	A	246	-	-	-	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	246	270	C6-C7	-2.43	1.37	1.40
3	A	246	270	C2-C7	-2.43	1.37	1.45
3	A	246	270	C3-C2	-2.19	1.38	1.40
3	A	246	270	C5-C6	2.13	1.42	1.38

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	246	270	C5-C6-C7	-3.02	115.42	119.86
3	A	246	270	C4-C3-C2	-2.75	115.83	119.86

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
3	A	246	270	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	223/223 (100%)	2.75	151 (67%) <b>0</b> <b>0</b>	10, 18, 29, 56	16 (7%)

All (151) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	147	SER	11.5
1	A	148	GLY	10.8
1	A	150	SER	8.3
1	A	151	TYR	7.8
1	A	149	THR	7.2
1	A	52	VAL	6.2
1	A	53[A]	VAL	5.9
1	A	125	THR	5.7
1	A	212	ILE	5.7
1	A	97	ASN	5.5
1	A	16	ILE	5.3
1	A	115	ASN	5.3
1	A	118	VAL	5.2
1	A	162	ILE	4.8
1	A	61	SER	4.7
1	A	31	VAL	4.7
1	A	114	LEU	4.7
1	A	177	THR	4.4
1	A	116	SER	4.3
1	A	33	LEU	4.2
1	A	20	TYR	4.1
1	A	201	CYS	4.1
1	A	146	SER	4.1
1	A	174	GLY	4.1
1	A	106	ILE	4.0
1	A	55	ALA	4.0
1	A	136	CYS	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	199	VAL	4.0
1	A	103	ILE	4.0
1	A	202	SER	4.0
1	A	155	LEU	3.9
1	A	191	CYS	3.8
1	A	112	ALA	3.8
1	A	181	PHE	3.8
1	A	215	TRP	3.8
1	A	51	TRP	3.7
1	A	223	ASN	3.7
1	A	154	VAL	3.7
1	A	231	VAL	3.7
1	A	42	CYS	3.6
1	A	111	ALA	3.6
1	A	213	VAL	3.6
1	A	144	THR	3.6
1	A	129	ALA	3.5
1	A	238	ILE	3.5
1	A	197	GLY	3.5
1	A	228	TYR	3.5
1	A	209	LEU	3.5
1	A	76	VAL	3.5
1	A	237	TRP	3.5
1	A	176	ILE	3.5
1	A	173	PRO	3.5
1	A	59	TYR	3.5
1	A	138	ILE	3.4
1	A	86	SER	3.4
1	A	220	CYS	3.4
1	A	141	TRP	3.4
1	A	75	VAL	3.3
1	A	171	ALA	3.3
1	A	117	ARG	3.3
1	A	134	THR	3.2
1	A	172	TYR	3.2
1	A	183	ALA	3.2
1	A	235	VAL	3.2
1	A	28	PRO	3.2
1	A	163	LEU	3.2
1	A	21	THR	3.2
1	A	60	LYS	3.2
1	A	44	GLY	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	105	LEU	3.1
1	A	108	LEU	3.1
1	A	58	CYS	3.1
1	A	73	ILE	3.1
1	A	130[A]	SER	3.1
1	A	158	LEU	3.0
1	A	27	VAL	2.9
1	A	143	ASN	2.9
1	A	23	GLY	2.9
1	A	239	LYS	2.9
1	A	195	SER	2.9
1	A	167	SER	2.9
1	A	99	LEU	2.9
1	A	185	LEU	2.9
1	A	85	ALA	2.9
1	A	145	LYS	2.9
1	A	57	HIS	2.8
1	A	89	ILE	2.8
1	A	140	GLY	2.8
1	A	152	PRO	2.8
1	A	184(A)	TYR	2.8
1	A	132	ALA	2.8
1	A	104	MET	2.8
1	A	227	VAL	2.8
1	A	43	GLY	2.8
1	A	98	THR	2.7
1	A	165[A]	ASP	2.7
1	A	188	GLY	2.7
1	A	29	TYR	2.7
1	A	182	CYS	2.7
1	A	135	GLN	2.7
1	A	96	SER	2.7
1	A	219	GLY	2.7
1	A	198	PRO	2.7
1	A	178	SER	2.7
1	A	159[A]	LYS	2.6
1	A	62	GLY	2.6
1	A	211	GLY	2.6
1	A	66[A]	LEU	2.6
1	A	56	ALA	2.5
1	A	232	CYS	2.5
1	A	186	GLU	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	26	THR	2.5
1	A	153	ASP	2.5
1	A	90	VAL	2.5
1	A	229	THR	2.5
1	A	39	TYR	2.5
1	A	32	SER	2.5
1	A	170[A]	SER	2.5
1	A	142	GLY	2.5
1	A	180	MET	2.5
1	A	193	GLY	2.4
1	A	168	CYS	2.4
1	A	139	SER	2.4
1	A	128	CYS	2.4
1	A	121	ILE	2.4
1	A	242	ILE	2.4
1	A	192	GLN	2.4
1	A	47	ILE	2.4
1	A	49	SER	2.3
1	A	217[A]	SER	2.3
1	A	22	CYS	2.3
1	A	243	ALA	2.3
1	A	48	ASN	2.3
1	A	93	SER	2.3
1	A	221(A)	GLN	2.3
1	A	234	TYR	2.3
1	A	63	ILE	2.2
1	A	113[A]	SER	2.2
1	A	30	GLN	2.2
1	A	157	CYS	2.2
1	A	94	TYR	2.1
1	A	222	LYS	2.1
1	A	54	SER	2.1
1	A	19	GLY	2.1
1	A	166[A]	SER	2.1
1	A	123	LEU	2.1
1	A	65	VAL	2.1
1	A	41	PHE	2.0
1	A	82	PHE	2.0
1	A	101	ASN	2.0
1	A	190	SER	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( $\text{\AA}^2$ )	Q<0.9
3	270	A	246	10/10	0.81	0.14	25,27,28,28	0
2	CA	A	247	1/1	0.94	0.23	27,27,27,27	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.