



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:50 AM GMT

PDB ID : 2BXE
Title : Human serum albumin complexed with diflunisal
Authors : Ghuman, J.; Zunszain, P.A.; Petitpas, I.; Bhattacharya, A.A.; Curry, S.
Deposited on : 2005-07-26
Resolution : 2.95 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 ⓘ 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

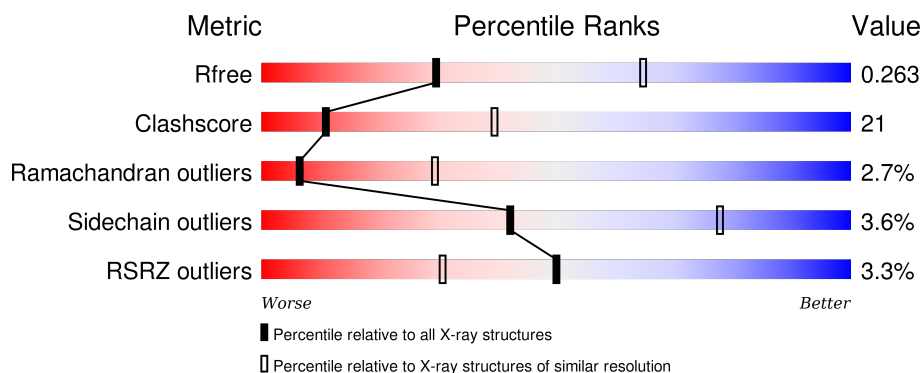
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 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	2184 (3.00-2.92)
Clashscore	102246	2552 (3.00-2.92)
Ramachandran outliers	100387	2468 (3.00-2.92)
Sidechain outliers	100360	2471 (3.00-2.92)
RSRZ outliers	91569	2201 (3.00-2.92)

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 $\leq 5\%$. 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	585	<div> <div>3%</div> <div> <div></div> <div>61%</div> <div>33%</div> <div>• •</div> </div> </div>
1	B	585	<div> <div>4%</div> <div> <div></div> <div>58%</div> <div>35%</div> <div>• •</div> </div> </div>

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
2	1FL	A	2003	-	-	-	X
2	1FL	B	2002	-	-	-	X
2	1FL	B	2003	-	-	-	X

2 Entry composition [i](#)

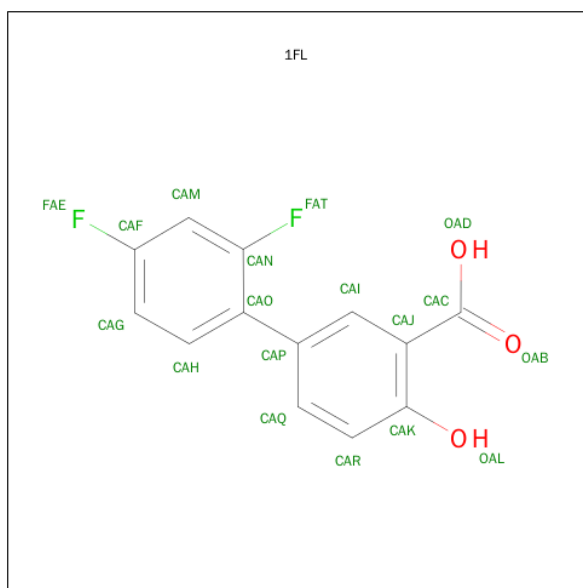
There are 2 unique types of molecules in this entry. The entry contains 8710 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 SERUM ALBUMIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	569	Total	C	N	O	S	0	0	0
			4344	2743	722	839	40			
1	B	566	Total	C	N	O	S	0	0	0
			4258	2692	710	816	40			

- Molecule 2 is 5-(2,4-DIFLUOROPHENYL)-2-HYDROXY-BENZOIC ACID (three-letter code: 1FL) (formula: C₁₃H₈F₂O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	F	O	0	0
			18	13	2	3		
2	A	1	Total	C	F	O	0	0
			18	13	2	3		
2	A	1	Total	C	F	O	0	0
			18	13	2	3		
2	B	1	Total	C	F	O	0	0
			18	13	2	3		

Continued on next page...

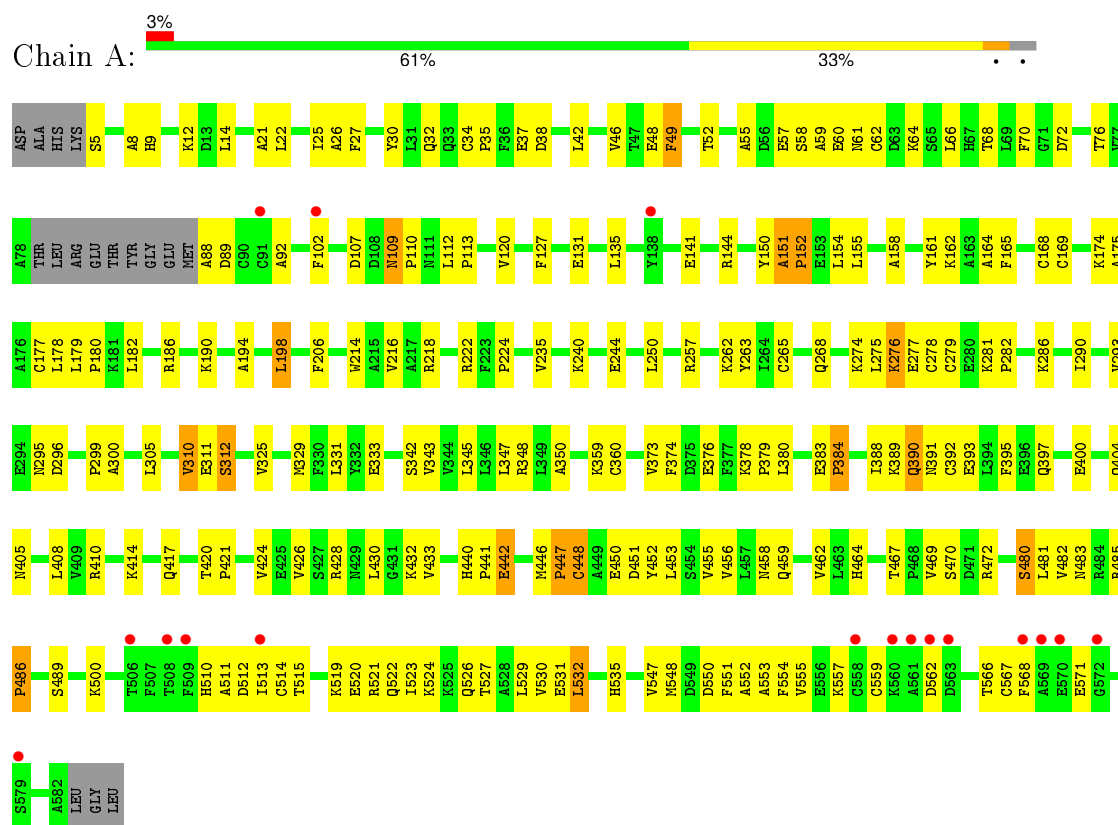
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	F	O	0	0
			18	13	2	3		
2	B	1	Total	C	F	O	0	0
			18	13	2	3		

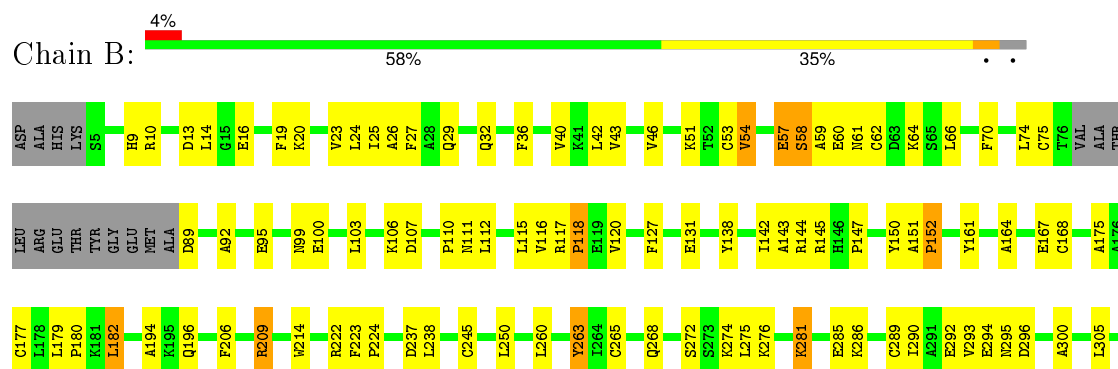
3 Residue-property plots

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 ($\text{RSRZ} > 2$). 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: SERUM ALBUMIN



• Molecule 1: SERUM ALBUMIN



S312	T412	P499	F568
V315	K413	F502	A569
L331	V415	H503	E570
V332	P416	A504	E571
E333	Q417	F507	K574
R337	V418	T508	L575
	S419	F509	V576
	T420	H510	A577
S342	L423	A511	A578
V343	V424	D512	S579
V344		J513	A582
L345	R428	C514	LEU
	N429	T515	GLY
R348	K432	L516	LEU
L349	V433	S517	
A350	G434	E518	
A363	C438	R521	
A364	K439	Q522	
P365	H440	I523	
P366	P441	K524	
H367	E442	K525	
E368	A443	Q526	
C369	R444	V530	
I370	R445	E531	
A371	M446	L532	
V372	P447	V533	
V373	C448	K534	
F374	A449	H535	
D375	E450	K536	
E376	D451	P537	
F377	Y452		
K378	L453	T540	
P379	S454	K541	
L380	V455	E542	
	V456	Q543	
E383		L544	
P384	Q459	K545	
		A546	
L387	V462	V547	
L388	L463	H548	
F395	V469	A552	
E396	S470		
Q397		V555	
L398		E556	
G399	S480	K557	
E400	L481	C558	
Y401	V482	C559	
K402	N483	K560	
F403	R484	A561	
Q404	R485	D562	
I405	P486	D563	
	C487	K564	
	F488	E565	
L408	S489	T566	
V409			
R410			
Y411	V498	C567	

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	55.41Å 55.53Å 120.52Å 81.54° 90.09° 65.93°	Depositor
Resolution (Å)	38.22 – 2.95 38.22 – 2.95	Depositor EDS
% Data completeness (in resolution range)	98.4 (38.22-2.95) 92.8 (38.22-2.95)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.29 (at 2.95Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.226 , 0.270 0.222 , 0.263	Depositor DCC
R_{free} test set	1296 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	73.1	Xtriage
Anisotropy	0.328	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 66.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 27267 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	8710	wwPDB-VP
Average B, all atoms (Å ²)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 1FL

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	0.45	0/4428	0.66	0/6011
1	B	0.45	0/4339	0.65	0/5895
All	All	0.45	0/8767	0.65	0/11906

There are no bond length outliers.

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	4344	0	4056	159	0
1	B	4258	0	3962	193	0
2	A	54	0	18	3	0
2	B	54	0	18	7	0
All	All	8710	0	8054	352	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 21.

The worst 5 of 352 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:433:VAL:HG22	1:A:452:TYR:HD2	1.06	1.10
1:A:433:VAL:HG22	1:A:452:TYR:CD2	1.91	1.05
1:B:433:VAL:HG22	1:B:452:TYR:HD2	1.30	0.95
1:B:433:VAL:HG22	1:B:452:TYR:CD2	2.02	0.93
1:B:222:ARG:HD3	2:B:2003:1FL:FAE	1.59	0.91

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	565/585 (97%)	486 (86%)	65 (12%)	14 (2%)	7	31
1	B	562/585 (96%)	478 (85%)	68 (12%)	16 (3%)	6	28
All	All	1127/1170 (96%)	964 (86%)	133 (12%)	30 (3%)	6	29

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	59	ALA
1	A	60	GLU
1	B	54	VAL
1	B	57	GLU
1	B	58	SER

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	447/511 (88%)	432 (97%)	15 (3%)	44	79
1	B	432/511 (84%)	415 (96%)	17 (4%)	39	75
All	All	879/1022 (86%)	847 (96%)	32 (4%)	42	78

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	532	LEU
1	B	167	GLU
1	B	489	SER
1	B	152	PRO
1	B	182	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	459	GLN
1	A	464	HIS
1	B	111	ASN
1	A	429	ASN
1	B	338	HIS

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 ⓘ

6 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	$\# Z > 2$
2	1FL	A	2001	-	16,19,19	1.39	2 (12%)	23,27,27	1.13	1 (4%)
2	1FL	A	2002	-	16,19,19	1.34	1 (6%)	23,27,27	1.20	3 (13%)
2	1FL	A	2003	-	16,19,19	0.54	0	23,27,27	1.00	0
2	1FL	B	2001	-	16,19,19	1.41	1 (6%)	23,27,27	1.10	1 (4%)
2	1FL	B	2002	-	16,19,19	1.36	1 (6%)	23,27,27	1.23	3 (13%)
2	1FL	B	2003	-	16,19,19	1.34	2 (12%)	23,27,27	1.34	3 (13%)

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
2	1FL	A	2001	-	-	0/4/8/8	0/2/2/2
2	1FL	A	2002	-	-	0/4/8/8	0/2/2/2
2	1FL	A	2003	-	-	0/4/8/8	0/2/2/2
2	1FL	B	2001	-	-	0/4/8/8	0/2/2/2
2	1FL	B	2002	-	-	0/4/8/8	0/2/2/2
2	1FL	B	2003	-	-	0/4/8/8	0/2/2/2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2001	1FL	CAM-CAN	2.10	1.41	1.37
2	B	2003	1FL	CAH-CAG	2.30	1.42	1.38
2	B	2003	1FL	CAH-CAO	3.12	1.45	1.39
2	B	2002	1FL	CAO-CAN	3.59	1.43	1.38
2	A	2001	1FL	CAO-CAN	3.60	1.43	1.38

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2003	1FL	CAH-CAG-CAF	-2.28	115.88	118.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2002	1FL	FAT-CAN-CAM	-2.07	114.70	118.59
2	B	2002	1FL	FAT-CAN-CAM	-2.02	114.80	118.59
2	A	2002	1FL	CAN-CAM-CAF	2.02	118.74	116.63
2	B	2003	1FL	OAL-CAK-CAJ	2.07	125.70	120.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	1FL	1	0
2	A	2003	1FL	2	0
2	B	2001	1FL	2	0
2	B	2002	1FL	2	0
2	B	2003	1FL	3	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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 ²)	Q<0.9
1	A	569/585 (97%)	-0.17	17 (2%)	54 33	35, 78, 156, 170	0
1	B	566/585 (96%)	-0.07	21 (3%)	45 27	35, 78, 148, 175	0
All	All	1135/1170 (97%)	-0.12	38 (3%)	50 31	35, 78, 150, 175	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	582	ALA	6.4
1	B	566	THR	6.3
1	B	562	ASP	5.8
1	B	579	SER	5.7
1	B	575	LEU	5.6

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, 95th 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(\AA^2)	Q<0.9
2	1FL	B	2003	18/18	0.87	0.36	4.14	98,101,102,104	0
2	1FL	A	2003	18/18	0.84	0.28	3.32	109,109,111,111	0
2	1FL	B	2002	18/18	0.92	0.31	2.85	91,94,99,100	0
2	1FL	B	2001	18/18	0.93	0.25	1.91	63,67,71,73	0
2	1FL	A	2002	18/18	0.91	0.24	1.46	77,82,91,92	0
2	1FL	A	2001	18/18	0.95	0.20	0.64	55,58,66,68	0

6.5 Other polymers [i](#)

There are no such residues in this entry.