



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:36 PM GMT

PDB ID : 4PF3  
Title : Mineralocorticoid receptor ligand-binding domain with compound 37a  
Authors : Sogabe, S.; Habuka, N.  
Deposited on : 2014-04-28  
Resolution : 1.10 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

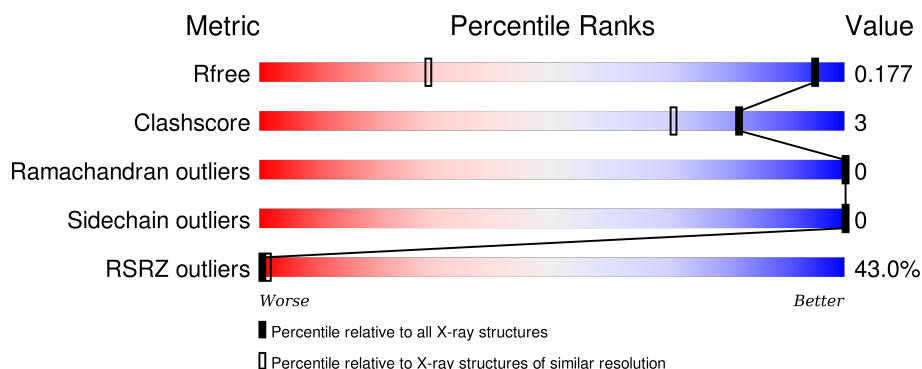
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.10 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	1006 (1.14-1.06)
Clashscore	102246	1055 (1.14-1.06)
Ramachandran outliers	100387	1016 (1.14-1.06)
Sidechain outliers	100360	1014 (1.14-1.06)
RSRZ outliers	91569	1009 (1.14-1.06)

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	275	<div> <div>39%</div> <div>87%</div> <div>9%</div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2404 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.

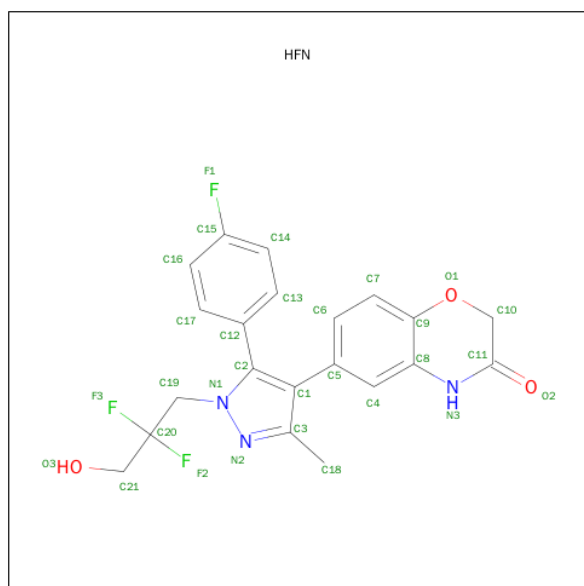
- Molecule 1 is a protein called Mineralocorticoid receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	249	2118	1384	335	383	16	0	16	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	710	GLY	-	expression tag	UNP P08235
A	711	SER	-	expression tag	UNP P08235
A	808	SER	CYS	engineered mutation	UNP P08235
A	810	LEU	SER	engineered mutation	UNP P08235
A	976	VAL	ALA	engineered mutation	UNP P08235

- Molecule 2 is 6-[1-(2,2-difluoro-3-hydroxypropyl)-5-(4-fluorophenyl)-3-methyl-1H-pyrazol-4-yl]-2H-1,4-benzoxazin-3(4H)-one (three-letter code: HFN) (formula: C<sub>21</sub>H<sub>18</sub>F<sub>3</sub>N<sub>3</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			30	21	3	3	3		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		

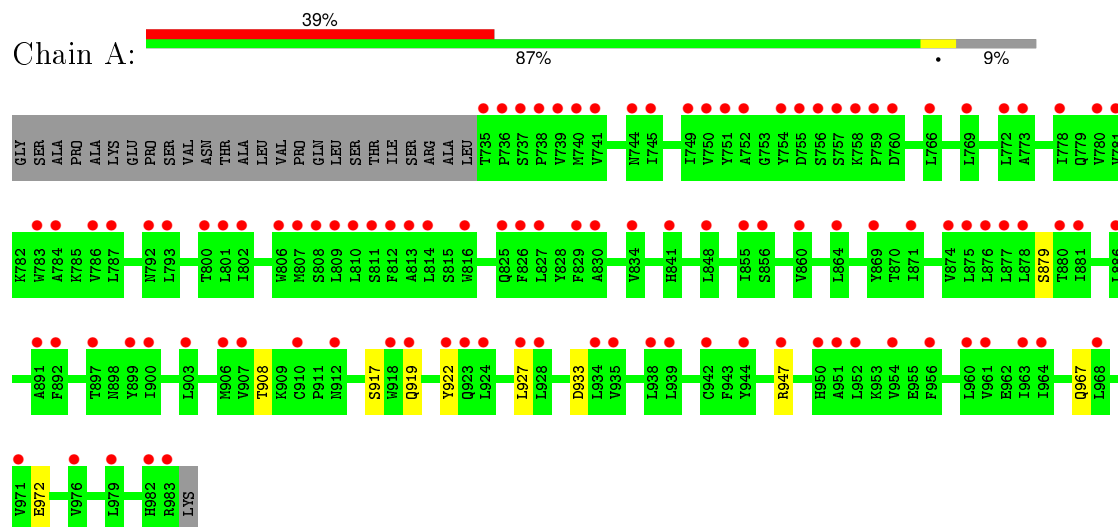
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	224	Total	O	0	0
			224	224		

### 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: Mineralocorticoid receptor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.05Å 66.36Å 74.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.10 37.45 – 1.10	Depositor EDS
% Data completeness (in resolution range)	95.1 (40.00-1.10) 95.1 (37.45-1.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 1.10Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, $R_{free}$	0.154 , 0.177 0.154 , 0.177	Depositor DCC
$R_{free}$ test set	5582 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.5	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.45 , 71.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	1 of 111970 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2404	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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: EDO, HFN

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	0.48	0/2217	0.66	0/2998

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	2118	0	2174	13	0
2	A	30	0	18	1	0
3	A	32	0	48	0	0
4	A	224	0	0	9	0
All	All	2404	0	2240	14	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 3.

All (14) 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:967[B]:GLN:NE2	4:A:1187:HOH:O	1.72	1.23
1:A:933[A]:ASP:CG	4:A:1238:HOH:O	1.96	1.01
1:A:933[A]:ASP:OD1	4:A:1238:HOH:O	1.79	1.00
1:A:922[B]:TYR:HD2	4:A:1316:HOH:O	1.63	0.82
1:A:922[B]:TYR:CD2	4:A:1316:HOH:O	2.36	0.76
1:A:879[A]:SER:OG	4:A:1241:HOH:O	2.16	0.60
1:A:947:ARG:NH2	1:A:972:GLU:OE2	2.32	0.59
1:A:919[A]:GLN:OE1	4:A:1316:HOH:O	2.18	0.54
1:A:919[A]:GLN:CD	4:A:1316:HOH:O	2.46	0.53
2:A:1001:HFN:H14	2:A:1001:HFN:C4	2.44	0.48
1:A:919[A]:GLN:NE2	4:A:1103:HOH:O	2.45	0.46
1:A:908:THR:HA	1:A:917:SER:HB2	2.01	0.42

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	263/275 (96%)	261 (99%)	2 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	247/252 (98%)	247 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

### 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 ⓘ

9 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	HFN	A	1001	-	29,33,33	1.44	4 (13%)	41,49,49	1.94	9 (21%)
3	EDO	A	1002	-	3,3,3	0.46	0	2,2,2	0.37	0
3	EDO	A	1003	-	3,3,3	0.51	0	2,2,2	0.64	0
3	EDO	A	1004	-	3,3,3	0.45	0	2,2,2	0.13	0
3	EDO	A	1005	-	3,3,3	0.88	0	2,2,2	0.57	0
3	EDO	A	1006	-	3,3,3	0.39	0	2,2,2	0.63	0
3	EDO	A	1007	-	3,3,3	0.45	0	2,2,2	0.54	0
3	EDO	A	1008	-	3,3,3	0.71	0	2,2,2	0.32	0
3	EDO	A	1009	-	3,3,3	0.61	0	2,2,2	0.33	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
2	HFN	A	1001	-	-	1/12/25/25	0/4/4/4
3	EDO	A	1002	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1003	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1004	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1005	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1006	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1007	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1008	-	-	0/1/1/1	0/0/0/0
3	EDO	A	1009	-	-	0/1/1/1	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	HFN	F2-C20	-4.33	1.32	1.38
2	A	1001	HFN	C11-N3	-2.77	1.32	1.35
2	A	1001	HFN	C1-C2	-2.10	1.37	1.40
2	A	1001	HFN	C19-C20	3.25	1.56	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	HFN	F3-C20-C21	-6.10	104.70	110.01
2	A	1001	HFN	C19-C20-C21	-3.25	114.08	118.24
2	A	1001	HFN	C5-C1-C3	-2.28	123.49	127.03
2	A	1001	HFN	C4-C5-C1	-2.09	117.12	120.57
2	A	1001	HFN	O1-C9-C8	-2.09	117.41	121.65
2	A	1001	HFN	C13-C12-C17	2.09	121.50	117.55
2	A	1001	HFN	O1-C9-C7	2.36	120.61	116.73
2	A	1001	HFN	C9-C8-N3	2.68	121.49	118.50
2	A	1001	HFN	F3-C20-F2	6.66	108.19	104.24

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1001	HFN	F2-C20-C19-N1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	HFN	1	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	249/275 (90%)	2.03	107 (42%) 0 1	10, 15, 26, 38	0

All (107) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	735	THR	13.2
1	A	736	PRO	6.8
1	A	756	SER	6.3
1	A	757	SER	4.9
1	A	751	TYR	4.7
1	A	810	LEU	4.6
1	A	739	VAL	4.4
1	A	983	ARG	4.4
1	A	738	PRO	4.2
1	A	951	ALA	4.2
1	A	754	TYR	4.1
1	A	950	HIS	4.1
1	A	759	PRO	4.0
1	A	922[A]	TYR	4.0
1	A	960	LEU	3.8
1	A	809	LEU	3.8
1	A	939	LEU	3.8
1	A	938	LEU	3.7
1	A	964	ILE	3.7
1	A	876	LEU	3.6
1	A	881	ILE	3.6
1	A	737	SER	3.5
1	A	935	VAL	3.5
1	A	877	LEU	3.5
1	A	874	VAL	3.5
1	A	814	LEU	3.4
1	A	755	ASP	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	928	LEU	3.4
1	A	778	ILE	3.3
1	A	802	ILE	3.3
1	A	806	TRP	3.2
1	A	900	ILE	3.2
1	A	871[A]	ILE	3.2
1	A	772	LEU	3.1
1	A	813	ALA	3.1
1	A	745	ILE	3.1
1	A	907	VAL	3.1
1	A	773	ALA	3.0
1	A	780	VAL	3.0
1	A	801	LEU	2.9
1	A	963	ILE	2.9
1	A	783	TRP	2.9
1	A	758	LYS	2.9
1	A	934	LEU	2.9
1	A	864	LEU	2.8
1	A	982	HIS	2.8
1	A	808[A]	SER	2.8
1	A	944	TYR	2.8
1	A	878	LEU	2.8
1	A	834	VAL	2.8
1	A	880	THR	2.7
1	A	875	LEU	2.7
1	A	886	LEU	2.7
1	A	741	VAL	2.7
1	A	954	VAL	2.7
1	A	869	TYR	2.7
1	A	749	ILE	2.7
1	A	812	PHE	2.7
1	A	787	LEU	2.7
1	A	786	VAL	2.7
1	A	855	ILE	2.6
1	A	769	LEU	2.6
1	A	793[A]	LEU	2.6
1	A	781	VAL	2.6
1	A	976	VAL	2.6
1	A	827	LEU	2.6
1	A	750	VAL	2.6
1	A	807	MET	2.6
1	A	919[A]	GLN	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	942	CYS	2.5
1	A	848	LEU	2.5
1	A	968	LEU	2.5
1	A	961	VAL	2.5
1	A	825	GLN	2.4
1	A	927[A]	LEU	2.4
1	A	912	ASN	2.4
1	A	740	MET	2.4
1	A	971	VAL	2.3
1	A	923[A]	GLN	2.3
1	A	792	ASN	2.3
1	A	892	PHE	2.3
1	A	979	LEU	2.3
1	A	856[A]	SER	2.3
1	A	947	ARG	2.2
1	A	744	ASN	2.2
1	A	860	VAL	2.2
1	A	952	LEU	2.2
1	A	830	ALA	2.2
1	A	906[A]	MET	2.2
1	A	816	TRP	2.2
1	A	918	TRP	2.2
1	A	956	PHE	2.2
1	A	800	THR	2.2
1	A	910	CYS	2.2
1	A	752	ALA	2.1
1	A	829	PHE	2.1
1	A	811	SER	2.1
1	A	903	LEU	2.1
1	A	924	LEU	2.1
1	A	891	ALA	2.1
1	A	826	PHE	2.1
1	A	784	ALA	2.1
1	A	760	ASP	2.0
1	A	897	THR	2.0
1	A	766	LEU	2.0
1	A	899	TYR	2.0
1	A	841	HIS	2.0

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

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( $\text{\AA}^2$ )	Q<0.9
3	EDO	A	1007	4/4	0.89	0.16	0.63	25,33,36,36	0
3	EDO	A	1008	4/4	0.77	0.17	0.41	25,25,28,30	0
3	EDO	A	1005	4/4	0.79	0.14	0.28	20,24,26,29	0
3	EDO	A	1003	4/4	0.81	0.14	-0.22	21,25,28,30	0
2	HFN	A	1001	30/30	0.91	0.14	-0.56	10,12,15,17	0
3	EDO	A	1009	4/4	0.91	0.13	-0.84	20,22,24,50	0
3	EDO	A	1004	4/4	0.89	0.12	-0.88	23,26,28,44	0
3	EDO	A	1002	4/4	0.92	0.12	-1.23	17,18,19,41	0
3	EDO	A	1006	4/4	0.95	0.11	-1.96	17,21,31,33	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.