Full wwPDB X-ray Structure Validation Report ⓘ

Apr 6, 2016 – 07:43 AM EDT

PDB ID	:	5E33
Title	:	Structure of human DPP3 in complex with met-enkephalin
Authors	:	Kumar, P.; Reithofer, V.; Reisinger, M.; Pavkov-Keller, T.; Wallner, S.; Macheroux, P.; Gruber, K.
Deposited on	:	2015-10-01
Resolution	:	1.84 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
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 ①) were used in the production of this report:

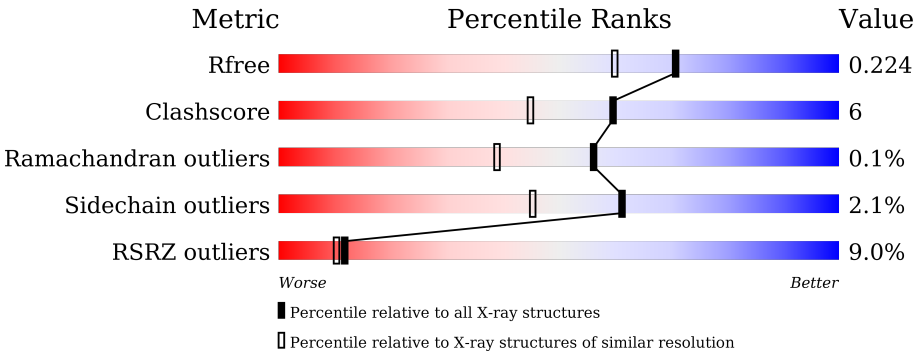
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027107
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027107

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 1.84 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	2634 (1.86-1.82)
Clashscore	102246	2862 (1.86-1.82)
Ramachandran outliers	100387	2831 (1.86-1.82)
Sidechain outliers	100360	2832 (1.86-1.82)
RSRZ outliers	91569	2639 (1.86-1.82)

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.

Mol	Chain	Length	Quality of chain
1	A	726	9% 88% 10%
2	B	5	20% 100%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6339 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 Dipeptidyl peptidase 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	725	5776	3682	977	1105	12	0	4	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	SER	CYS	engineered mutation	UNP Q9NY33
A	207	CYS	GLU	engineered mutation	UNP Q9NY33
A	451	ALA	GLU	engineered mutation	UNP Q9NY33
A	491	CYS	SER	engineered mutation	UNP Q9NY33
A	519	SER	CYS	engineered mutation	UNP Q9NY33
A	654	SER	CYS	engineered mutation	UNP Q9NY33

- Molecule 2 is a protein called Met-enkephalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	5	40	27	5	7	1	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	1	1	1	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
4	A	2	2	2	0	0

- Molecule 5 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total K 1 1	0	0

- Molecule 6 is water.

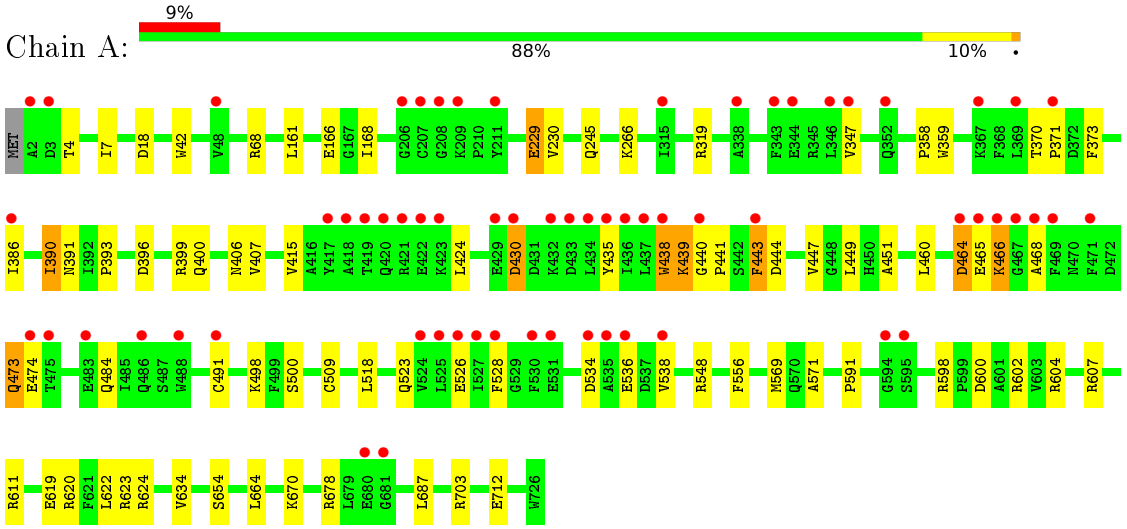
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	517	Total O 517 517	0	0
6	B	2	Total O 2 2	0	0

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.

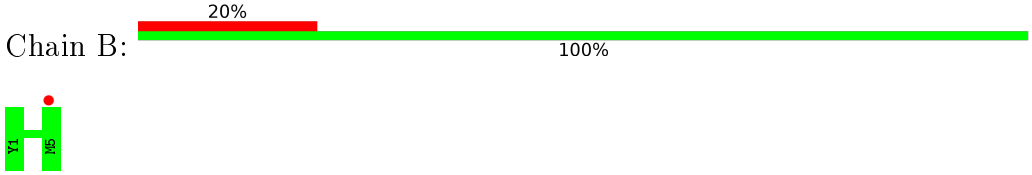
- Molecule 1: Dipeptidyl peptidase 3

Chain A:



- Molecule 2: Met-enkephalin

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	119.79Å 105.76Å 65.17Å 90.00° 93.49° 90.00°	Depositor
Resolution (Å)	45.42 – 1.84 45.42 – 1.84	Depositor EDS
% Data completeness (in resolution range)	96.8 (45.42-1.84) 96.9 (45.42-1.84)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 1.84Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.181 , 0.224 0.183 , 0.224	Depositor DCC
R_{free} test set	3411 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	29.9	Xtriage
Anisotropy	0.139	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	1 of 68276 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6339	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: K, MG, ZN

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	0.56	0/5920	0.62	3/8019 (0.0%)
2	B	0.57	0/41	0.77	0/51
All	All	0.56	0/5961	0.62	3/8070 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	466	LYS	N-CA-C	-6.25	94.12	111.00
1	A	319	ARG	NE-CZ-NH1	-6.06	117.27	120.30
1	A	438	TRP	N-CA-C	5.45	125.72	111.00

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	5776	0	5657	65	0
2	B	40	0	35	0	0
3	A	1	0	0	0	0
4	A	2	0	0	0	0
5	A	1	0	0	0	0
6	A	517	0	0	11	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	2	0	0	0	0
All	All	6339	0	5692	65	1

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 6.

All (65) 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:438:TRP:O	1:A:441:PRO:HD2	1.43	1.18
1:A:438:TRP:C	1:A:441:PRO:HD2	1.75	1.06
1:A:68[B]:ARG:NH2	1:A:712:GLU:OE2	2.05	0.90
1:A:439:LYS:NZ	1:A:443:PHE:HE1	1.71	0.87
1:A:473:GLN:NE2	1:A:473:GLN:O	2.06	0.87
1:A:439:LYS:HZ3	1:A:443:PHE:HE1	0.85	0.81
1:A:68[A]:ARG:NH2	1:A:712:GLU:OE2	2.15	0.80
1:A:464:ASP:OD1	1:A:466:LYS:N	2.15	0.79
1:A:473:GLN:C	1:A:473:GLN:HE21	1.92	0.73
1:A:623:ARG:NH1	6:A:902:HOH:O	2.25	0.69
1:A:358:PRO:HB3	1:A:619:GLU:HG3	1.75	0.67
1:A:438:TRP:O	1:A:441:PRO:CD	2.34	0.66
1:A:391:ASN:HD22	1:A:451:ALA:HB1	1.62	0.64
1:A:435:TYR:CZ	1:A:439:LYS:HD2	2.33	0.63
1:A:604:ARG:NH2	6:A:904:HOH:O	2.30	0.62
1:A:390:ILE:HG12	1:A:407:VAL:HB	1.83	0.61
1:A:430:ASP:N	1:A:430:ASP:OD1	2.27	0.59
1:A:440:GLY:N	1:A:441:PRO:CD	2.66	0.58
1:A:391:ASN:ND2	1:A:406:ASN:OD1	2.36	0.58
1:A:393:PRO:O	1:A:399:ARG:HD3	2.04	0.57
1:A:245:GLN:NE2	6:A:908:HOH:O	2.37	0.57
1:A:678:ARG:HG3	1:A:687:LEU:HD11	1.87	0.56
1:A:465:GLU:HG2	1:A:491:CYS:SG	2.46	0.55
1:A:400:GLN:NE2	6:A:911:HOH:O	2.40	0.54
1:A:438:TRP:HA	1:A:441:PRO:HG2	1.89	0.54
1:A:415:VAL:CG1	1:A:670:LYS:HD2	2.38	0.53
1:A:620:ARG:O	1:A:624:ARG:HG3	2.08	0.53
1:A:607:ARG:NH2	6:A:903:HOH:O	2.28	0.52
1:A:435:TYR:CE1	1:A:439:LYS:HD2	2.45	0.52
1:A:444:ASP:HB3	1:A:528:PHE:CE2	2.46	0.51
1:A:536:GLU:HG2	6:A:903:HOH:O	2.10	0.51
1:A:447:VAL:HG11	6:A:1151:HOH:O	2.10	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:600:ASP:OD2	1:A:602:ARG:NH1	2.45	0.50
1:A:359:TRP:HZ3	1:A:622:LEU:HG	1.77	0.50
1:A:534:ASP:O	1:A:538:VAL:HG23	2.12	0.50
1:A:556:PHE:CD2	1:A:569:MET:HG3	2.47	0.49
1:A:611:ARG:NH1	6:A:918:HOH:O	2.45	0.49
1:A:444:ASP:HB3	1:A:528:PHE:HE2	1.78	0.49
1:A:466:LYS:O	1:A:468:ALA:N	2.36	0.49
1:A:424:LEU:HA	1:A:424:LEU:HD23	1.68	0.49
1:A:474:GLU:O	1:A:484:GLN:NE2	2.47	0.48
1:A:591:PRO:HD3	6:A:1059:HOH:O	2.14	0.47
1:A:347:VAL:HG22	1:A:371:PRO:HB2	1.96	0.47
1:A:68[A]:ARG:HH22	1:A:712:GLU:CD	2.12	0.47
1:A:598:ARG:NH1	1:A:664:LEU:HG	2.31	0.46
1:A:229:GLU:HG3	1:A:230:VAL:N	2.30	0.46
1:A:415:VAL:HG13	1:A:670:LYS:HD2	1.96	0.46
1:A:634:VAL:HG12	6:A:1307:HOH:O	2.16	0.46
1:A:370:THR:HA	1:A:371:PRO:HD3	1.73	0.45
1:A:509:CYS:HA	1:A:571:ALA:HB1	1.98	0.45
1:A:391:ASN:OD1	1:A:399:ARG:NH2	2.41	0.45
1:A:42:TRP:CD2	1:A:703:ARG:HD3	2.51	0.45
1:A:548:ARG:HD3	6:A:1265:HOH:O	2.18	0.43
1:A:347:VAL:HG13	1:A:371:PRO:HD2	2.00	0.43
1:A:449:LEU:HD13	1:A:518:LEU:HB2	2.00	0.43
1:A:464:ASP:OD1	1:A:465:GLU:N	2.52	0.42
1:A:161:LEU:HA	1:A:168:ILE:O	2.20	0.42
1:A:386:ILE:HG21	1:A:386:ILE:HD13	1.87	0.42
1:A:4:THR:HB	1:A:7:ILE:HG12	2.01	0.42
1:A:390:ILE:CG1	1:A:407:VAL:HB	2.48	0.42
1:A:438:TRP:HA	1:A:441:PRO:CD	2.50	0.41
1:A:266:LYS:HB3	1:A:266:LYS:HE2	1.80	0.41
1:A:523:GLN:O	1:A:526:GLU:HB2	2.20	0.41
1:A:466:LYS:C	1:A:468:ALA:H	2.22	0.41
1:A:396:ASP:HA	1:A:399:ARG:HG3	2.03	0.41

All (1) 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)
6:A:1273:HOH:O	6:A:1339:HOH:O[2_757]	2.18	0.02

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	727/726 (100%)	713 (98%)	13 (2%)	1 (0%)	56 39
2	B	3/5 (60%)	3 (100%)	0	0	100 100
All	All	730/731 (100%)	716 (98%)	13 (2%)	1 (0%)	56 39

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	500	SER

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	612/609 (100%)	599 (98%)	13 (2%)	61 45
2	B	3/3 (100%)	3 (100%)	0	100 100
All	All	615/612 (100%)	602 (98%)	13 (2%)	61 45

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	ASP
1	A	166	GLU
1	A	229	GLU
1	A	373	PHE
1	A	390	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	430	ASP
1	A	439	LYS
1	A	443	PHE
1	A	460	LEU
1	A	464	ASP
1	A	473	GLN
1	A	498	LYS
1	A	654	SER

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

Of 4 ligands modelled in this entry, 4 are 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

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^{th} 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	725/726 (99%)	0.30	65 (8%) 12 10	18, 32, 70, 94	0
2	B	5/5 (100%)	2.09	1 (20%) 1 1	34, 39, 50, 76	0
All	All	730/731 (99%)	0.31	66 (9%) 12 10	18, 32, 71, 94	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	535	ALA	9.0
1	A	466	LYS	6.3
1	A	417	TYR	6.2
2	B	5	MET	6.2
1	A	467	GLY	6.1
1	A	419	THR	6.1
1	A	530	PHE	5.8
1	A	421	ARG	5.5
1	A	206	GLY	5.1
1	A	436	ILE	5.0
1	A	423	LYS	5.0
1	A	527	ILE	4.8
1	A	468	ALA	4.8
1	A	528	PHE	4.7
1	A	464	ASP	4.4
1	A	443	PHE	4.3
1	A	469	PHE	4.2
1	A	474	GLU	4.2
1	A	207	CYS	4.2
1	A	434	LEU	4.0
1	A	525	LEU	3.8
1	A	491	CYS	3.8
1	A	437	LEU	3.7
1	A	418	ALA	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	440	GLY	3.6
1	A	344	GLU	3.6
1	A	465	GLU	3.5
1	A	531	GLU	3.5
1	A	422	GLU	3.5
1	A	208	GLY	3.4
1	A	315	ILE	3.2
1	A	2	ALA	3.2
1	A	369	LEU	3.1
1	A	432	LYS	3.1
1	A	347	VAL	3.1
1	A	435	TYR	3.0
1	A	595	SER	3.0
1	A	526	GLU	2.9
1	A	386	ILE	2.9
1	A	420	GLN	2.9
1	A	338	ALA	2.9
1	A	346	LEU	2.8
1	A	438	TRP	2.8
1	A	3	ASP	2.8
1	A	471	PHE	2.8
1	A	48	VAL	2.8
1	A	534	ASP	2.7
1	A	680	GLU	2.7
1	A	524	VAL	2.7
1	A	681	GLY	2.6
1	A	488	TRP	2.5
1	A	352	GLN	2.5
1	A	371	PRO	2.4
1	A	486	GLN	2.4
1	A	429	GLU	2.4
1	A	475	THR	2.2
1	A	211	TYR	2.2
1	A	343	PHE	2.2
1	A	430	ASP	2.2
1	A	433	ASP	2.1
1	A	367	LYS	2.1
1	A	594	GLY	2.1
1	A	538	VAL	2.1
1	A	483	GLU	2.1
1	A	536	GLU	2.0
1	A	209	LYS	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 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^{th} 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(Å^2)	Q<0.9
4	MG	A	803	1/1	0.99	0.14	1.13	21,21,21,21	0
3	ZN	A	801	1/1	0.98	0.10	-1.59	40,40,40,40	1
5	K	A	804	1/1	1.00	0.05	-3.40	27,27,27,27	0
4	MG	A	802	1/1	0.99	0.19	-	24,24,24,24	0

6.5 Other polymers (i)

There are no such residues in this entry.