W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Mar 1, 2014 – 04:04 AM GMT

PDB ID   :   3D0I
Title   :   Crystal structure of spike protein receptor-binding domain from the 2005-2006 SARS coronavirus civet strain complexed with human-civet chimeric receptor ACE2
Authors   :   Li, F.
Deposited on   :   2008-05-01
Resolution   :   2.90 Å(reported)

---

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.90 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	Percentile Ranks	Value
Clashscore		5
Ramachandran outliers		3.7%
Sidechain outliers		7.6%
RSRZ outliers		6.5%

Worse ——— Better

▌Percentile relative to all X-ray structures

◻Percentile relative to X-ray structures of similar resolution
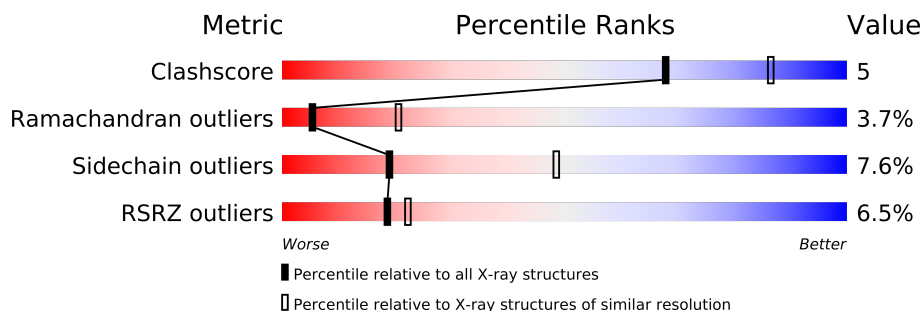
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	597	
1	B	597	
2	E	179	
2	F	179	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	NDG	B	618	-	X
5	ZN	A	901	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12639 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Angiotensin-convertingenzyme 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	597	Total 4864	C 3110	N 803	O 922	S 29	0	0	0
1	B	597	Total 4864	C 3110	N 803	O 922	S 29	0	0	0

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	173	Total 1395	C 905	N 229	O 255	S 6	0	0	0
2	F	173	Total 1395	C 905	N 229	O 255	S 6	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	479	ARG	ASN	CONFLICT	UNP P59594
E	480	GLY	ASP	CONFLICT	UNP P59594
E	487	SER	THR	CONFLICT	UNP P59594
F	479	ARG	ASN	CONFLICT	UNP P59594
F	480	GLY	ASP	CONFLICT	UNP P59594
F	487	SER	THR	CONFLICT	UNP P59594

- Molecule 3 is SUGAR (2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE) (three-letter code: NDG) (formula: $C_8H_{15}NO_6$).
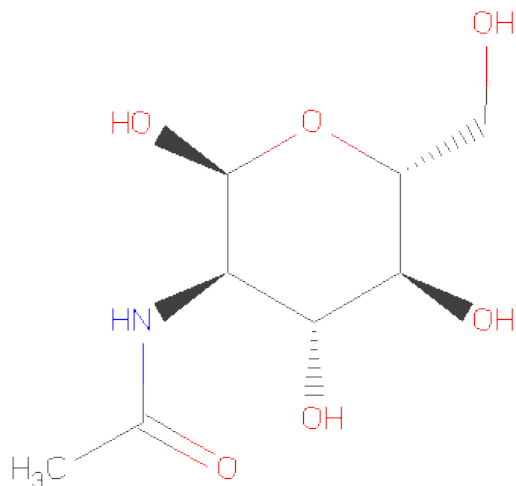
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			15	8	1	6		
3	B	1	Total	C	N	O	0	0
			15	8	1	6		
3	B	1	Total	C	N	O	0	0
			15	8	1	6		
3	B	1	Total	C	N	O	0	0
			15	8	1	6		
3	B	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	E	1	Total 15	C 8	N 1	O 6	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total 1	Zn 1	0	0
5	A	1	Total 1	Zn 1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total 1	Cl 1	0	0
6	A	1	Total 1	Cl 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	21	Total 21	O 21	0	0
7	F	6	Total 6	O 6	0	0

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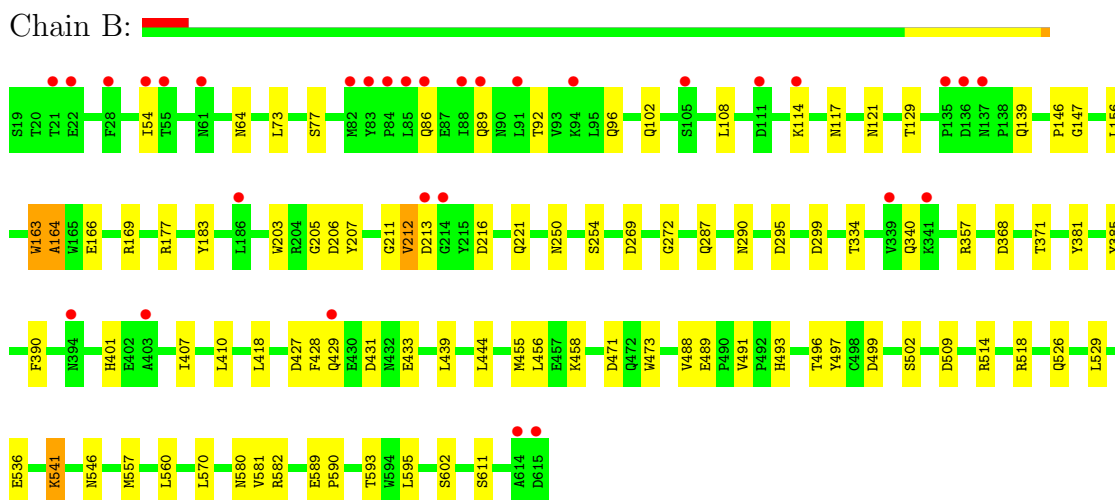 displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (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: Angiotensin-convertingenzyme 2

Chain A:

- Molecule 1: Angiotensin-convertingenzyme 2

Chain B:

- Molecule 2: Spike protein S1

Chain E:

G482 F501 E502

• Molecule 2: Spike protein S1

Chain F:

P324 F325 K333 V337 I345 C348 D351 Y352 S353 V354 L355 F360 F361 F364 Y367 G368 V369 S370 N375 ASP LEU CYS PHE SER ASN V382 R395 Q401 T402 G403 V404 K411 D415 F416 M417 L421 A422 W423 H445 D463 G464 K465 T468

P469 P470 A471 L472 Y475 G482 E502
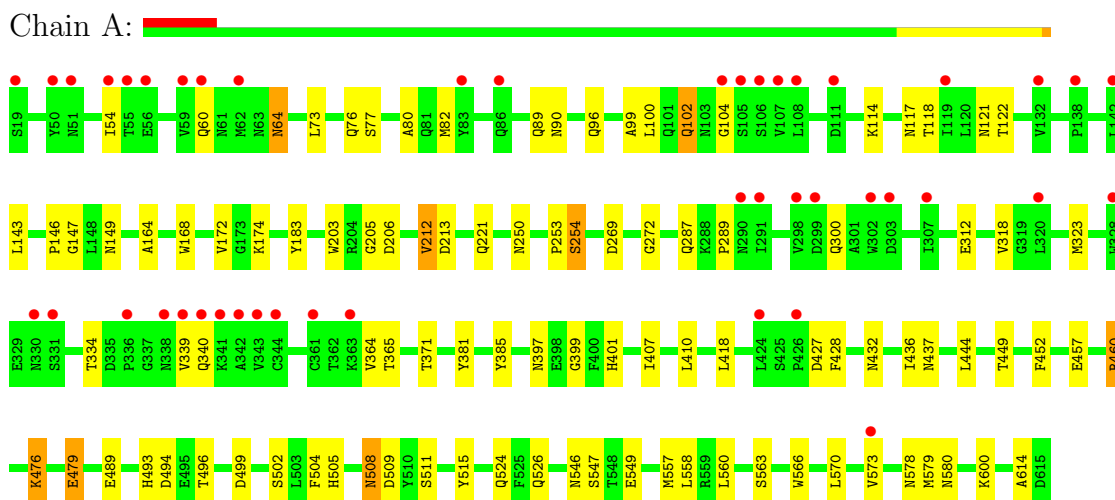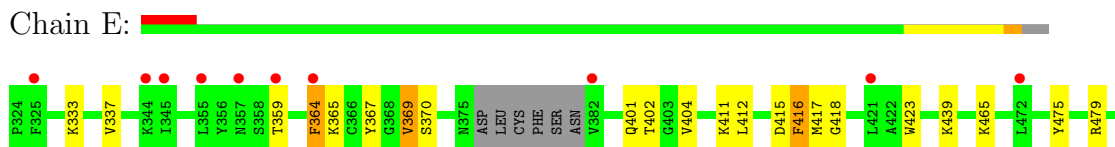
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.42Å 119.82Å 109.77Å 90.00° 95.50° 90.00°	Depositor
Resolution (Å)	36.47 – 2.90 36.47 – 2.89	Depositor EDS
% Data completeness (in resolution range)	89.9 (36.47-2.90) 89.3 (36.47-2.89)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.71 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.224 , 0.278 0.244 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	75.7	Xtriage
Anisotropy	0.183	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 38.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.52, < L^2 > = 0.36$	Xtriage
Outliers	0 of 46219 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	12639	wwPDB-VP
Average B, all atoms (Å^2)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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: ZN, NAG, NDG, CL

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	0.53	1/5000 (0.0%)	0.60	0/6796
1	B	0.53	0/5000	0.62	0/6796
2	E	0.55	0/1440	0.61	0/1958
2	F	0.54	0/1440	0.61	0/1958
All	All	0.53	1/12880 (0.0%)	0.61	0/17508

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	300	GLN	CD-OE1	6.23	1.37	1.24

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4864	0	0	30	0
1	B	4864	0	0	24	0
2	E	1395	0	0	2	0
2	F	1395	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	15	0	0	1	0
3	B	60	0	0	2	0
4	E	15	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	1	0
7	B	21	0	0	1	0
7	F	6	0	0	0	0
All	All	12639	0	0	62	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 62 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:269:ASP:OD2	1:A:272:GLY:N	2.29	0.66
1:B:546:ASN:ND2	3:B:619:NDG:N2	2.47	0.62
1:B:541:LYS:NZ	1:B:541:LYS:CB	2.63	0.62
1:B:499:ASP:O	1:B:502:SER:CB	2.49	0.61
1:B:269:ASP:OD2	1:B:272:GLY:N	2.34	0.61

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	595/597 (100%)	524 (88%)	52 (9%)	19 (3%)	6	25
1	B	595/597 (100%)	528 (89%)	50 (8%)	17 (3%)	7	28
2	E	169/179 (94%)	138 (82%)	21 (12%)	10 (6%)	2	7
2	F	169/179 (94%)	141 (83%)	17 (10%)	11 (6%)	2	5
All	All	1528/1552 (98%)	1331 (87%)	140 (9%)	57 (4%)	5	20

5 of 57 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	146	PRO
1	A	212	VAL
1	A	289	PRO
1	A	340	GLN
1	A	614	ALA

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	527/527 (100%)	491 (93%)	36 (7%)	22	55
1	B	527/527 (100%)	484 (92%)	43 (8%)	17	44
2	E	150/156 (96%)	137 (91%)	13 (9%)	15	41
2	F	150/156 (96%)	139 (93%)	11 (7%)	20	51
All	All	1354/1366 (99%)	1251 (92%)	103 (8%)	19	48

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	250	ASN
1	B	418	LEU
2	F	369	VAL
1	B	254	SER
1	B	371	THR

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 chains 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 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	NDG	A	616	-	15,15,15	0.48	0	21,21,21	0.61	0
3	NDG	B	616	-	15,15,15	0.62	0	21,21,21	0.60	0
3	NDG	B	617	-	15,15,15	0.54	0	21,21,21	1.24	2 (9%)
3	NDG	B	618	-	15,15,15	0.73	0	21,21,21	1.05	1 (4%)
3	NDG	B	619	-	15,15,15	0.63	0	21,21,21	1.08	2 (9%)
4	NAG	E	91	-	15,15,15	0.75	0	21,21,21	1.32	2 (9%)

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	NDG	A	616	-	-	1/6/26/26	0/1/1/1
3	NDG	B	616	-	-	0/6/26/26	0/1/1/1
3	NDG	B	617	-	-	0/6/26/26	0/1/1/1
3	NDG	B	618	-	-	1/6/26/26	0/1/1/1
3	NDG	B	619	-	-	0/6/26/26	0/1/1/1
4	NAG	E	91	-	-	0/6/26/26	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	91	NAG	C1-O5-C5	3.39	119.47	113.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	617	NDG	C1-C2-N2	3.26	114.63	110.85
4	E	91	NAG	O5-C1-C2	2.95	112.60	109.61
3	B	617	NDG	C4-C3-C2	2.79	114.41	110.44
3	B	618	NDG	O-C1-C2	2.75	112.40	109.61

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	616	NDG	C1-C2-N2-C7
3	B	618	NDG	O7-C7-N2-C2

There are no ring outliers.

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, 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	597/597 (100%)	0.41	45 (7%)	14	17	47, 80, 130, 154	0
1	B	597/597 (100%)	0.25	31 (5%)	26	32	38, 74, 130, 156	0
2	E	173/179 (96%)	0.37	11 (6%)	19	23	56, 89, 132, 134	0
2	F	173/179 (96%)	0.36	12 (6%)	17	20	58, 90, 131, 134	0
All	All	1540/1552 (99%)	0.34	99 (6%)	18	23	38, 80, 131, 156	0

The worst 5 of 99 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	290	ASN	7.0
1	A	342	ALA	6.9
1	B	615	ASP	6.7
1	A	341	LYS	6.6
1	A	302	TRP	5.9

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 ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	ZN	A	901	1/1	0.43	4.40	123,123,123,123	0
3	NDG	B	618	15/15	0.26	2.02	111,113,114,114	0
3	NDG	B	616	15/15	0.27	0.74	123,125,125,126	0
3	NDG	B	619	15/15	0.18	0.14	72,76,81,81	0
5	ZN	B	901	1/1	0.26	0.11	131,131,131,131	0
6	CL	A	902	1/1	0.19	-0.30	84,84,84,84	0
3	NDG	A	616	15/15	0.13	-1.27	109,113,114,114	0
6	CL	B	902	1/1	0.17	-1.68	73,73,73,73	0
3	NDG	B	617	15/15	0.27	-	126,128,129,129	0
4	NAG	E	91	15/15	0.29	-	137,140,141,142	0

6.5 Other polymers ⓘ

There are no such residues in this entry.