wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 05:53 PM EDT

PDB ID	:	5L5K
Title	:	Plexin A4 full extracellular region, domains 1 to 10, data to 7.5 angstrom, spacegroup P4(1)
Authors	:	Janssen, B.J.C.; Kong, Y.; Malinauskas, T.; Vangoor, V.R.; Coles, C.H.; Kaufmann, R.; Ni, T.; Gilbert, R.J.C.; Padilla-Parra, S.; Pasterkamp, R.J.; Jones, E.Y.
Deposited on	:	2016-05-28
Resolution	:	7.50 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

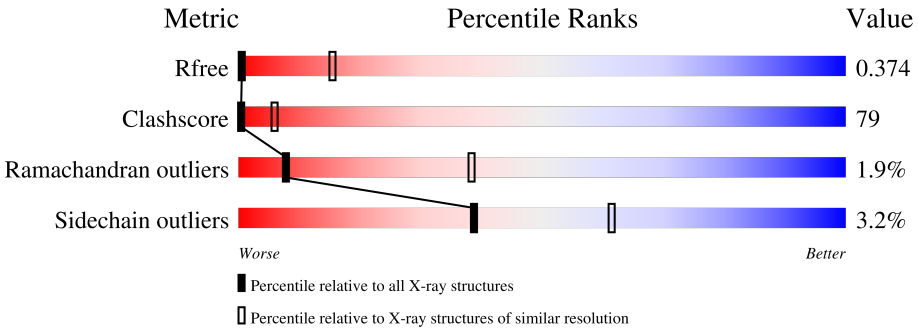
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 7.50 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	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1207	32% 58% 6% ·

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9134 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 Plexin-A4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1168	9134	5761	1572	1729	72	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

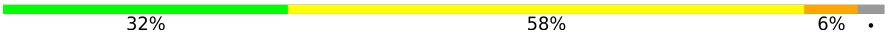
Chain	Residue	Modelled	Actual	Comment	Reference
A	33	GLU	-	expression tag	UNP Q80UG2
A	34	THR	-	expression tag	UNP Q80UG2
A	35	GLY	-	expression tag	UNP Q80UG2
A	1230	GLY	-	expression tag	UNP Q80UG2
A	1231	ARG	-	expression tag	UNP Q80UG2
A	1232	THR	-	expression tag	UNP Q80UG2
A	1233	LYS	-	expression tag	UNP Q80UG2
A	1234	HIS	-	expression tag	UNP Q80UG2
A	1235	HIS	-	expression tag	UNP Q80UG2
A	1236	HIS	-	expression tag	UNP Q80UG2
A	1237	HIS	-	expression tag	UNP Q80UG2
A	1238	HIS	-	expression tag	UNP Q80UG2
A	1239	HIS	-	expression tag	UNP Q80UG2

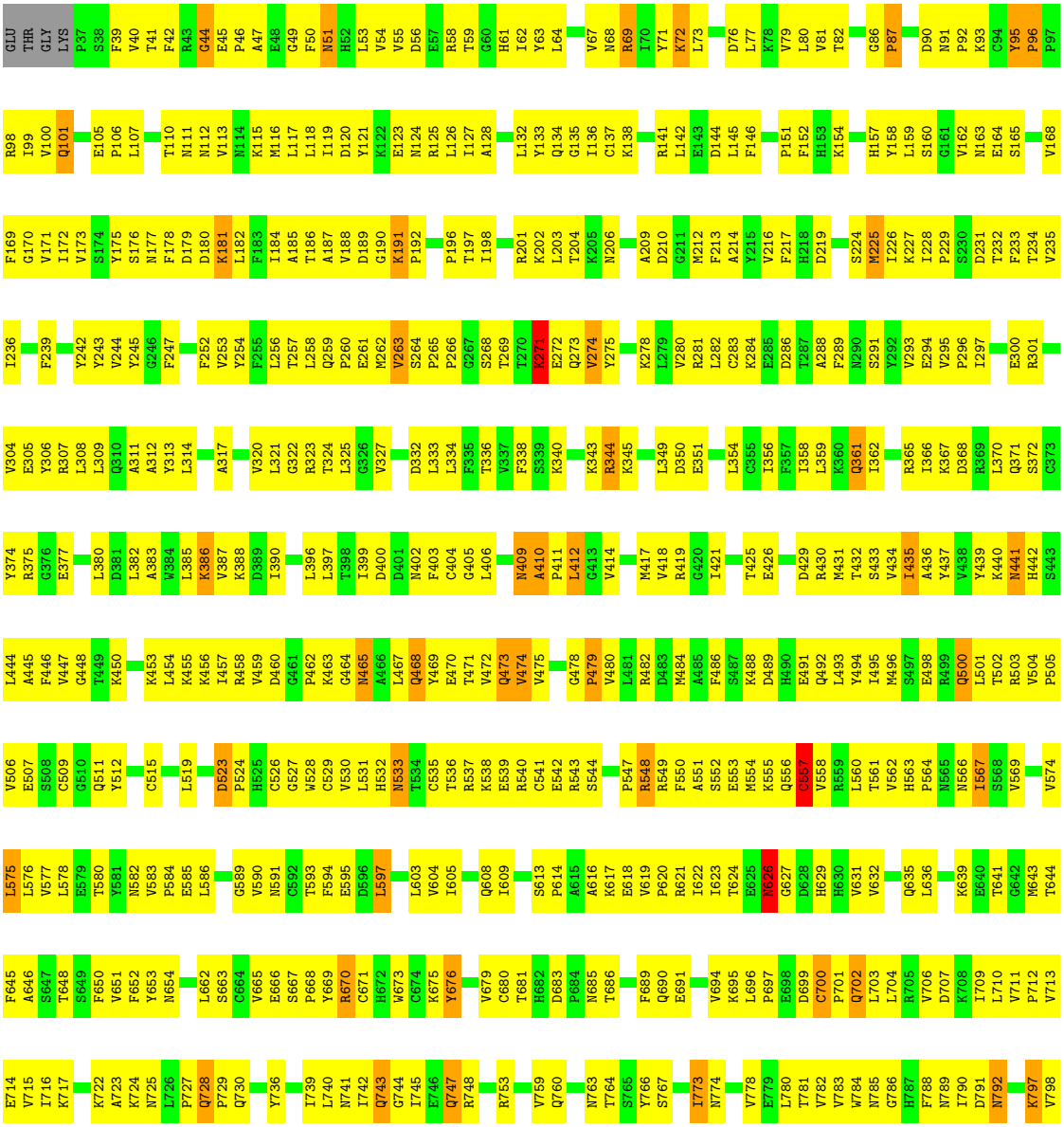
3 Residue-property plots (i)

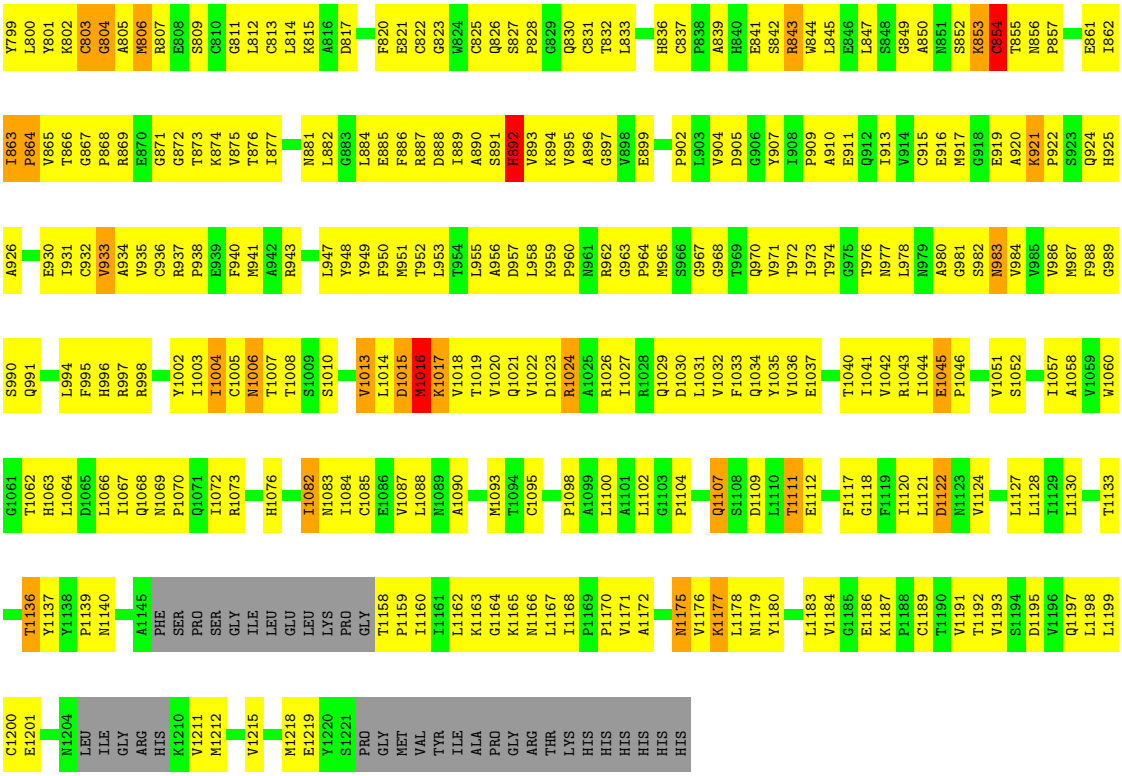
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: Plexin-A4

Chain A:






Y799 L800 Y801 K802 C803 G804 A805 M806 R807 E808 S809 C810 G811 L812 C813 L814 K815 A816 D817 F820 E821 C822 G823 W824 C825 Q826 S827 P828 G829 Q830 C831 T832 L833 H836 C837 P838 A839 H840 E841 S842 R843 W844 L845 E846 L847 S848 G849 A850 N851 S852 K853 C854 T855 N856 P857 E861 I862
I863 P864 V865 T866 G867 P868 R869 E870 G871 G872 T873 K874 V875 T876 I877 N881 L882 G883 L884 E885 F886 R887 D888 I889 A890 S891 H892 V893 K894 V895 A896 G897 V898 E899 P902 L903 V904 D905 G906 Y907 I908 P909 A910 E911 Q912 I913 V914 C915 E916 M917 G918 E919 A920 K921 P922 S923 Q924 H925
A926 E930 I931 C932 V933 A934 V935 C936 R937 P938 E939 F940 M941 A942 R943 L947 Y948 Y949 F950 M951 T952 L953 T954 L955 A956 D957 L958 K959 P960 N961 R962 G963 P964 M965 S966 G967 G968 T969 Q970 V971 T972 I973 T974 G975 T976 N977 L978 N979 A980 G981 S982 N983 V984 V985 V986 M987 F988 G989
S990 Q991 L994 F995 H996 R997 R998 Y1002 I1003 I1004 C1005 N1006 T1007 T1008 S1009 S1010 V1013 L1014 D1015 M1016 K1017 V1018 T1019 V1020 Q1021 V1022 D1023 R1024 A1025 R1026 I1027 R1028 Q1029 D1030 L1031 V1032 F1033 Q1034 Y1035 V1036 E1037 T1040 I1041 V1042 R1043 I1044 E1045 P1046 V1051 S1052 I1057 A1058 V1059 W1060
G1061 T1062 H1063 L1064 D1065 L1066 I1067 Q1068 N1069 P1070 Q1071 I1072 R1073 H1076 I1082 N1083 I1084 C1085 E1086 V1087 L1088 N1089 A1090 M1093 T1094 C1095 P1098 A1099 L1100 A1101 L1102 G1103 P1104 Q1107 S1108 D1109 Q1110 L1111 E1112 F1117 G1118 F1119 I1120 L1121 D1122 N1123 V1124 L1127 L1128 I1129 L1130 T1133
T1136 Y1137 Y1138 P1139 N1140 A1145 PHE SER PRO SER GLY ILE LEU GLU LEU LYS PRO GLY T1158 P1159 I1160 I1161 L1162 K1163 G1164 K1165 N1166 L1167 I1168 P1169 P1170 V1171 A1172 N1175 V1176 K1177 L1178 N1179 Y1180 L1183 V1184 G1185 E1186 K1187 P1188 C1189 T1190 V1191 T1192 V1193 S1194 D1195 V1196 Q1197 L1198 L1199
C1200 E1201 N1204 LEU ILE GLY ARG HIS K1210 V1211 M1212 V1215 M1218 E1219 Y1220 S1221 PRO GLY MET VAL TYR ILE ALA PRO GLY ARG THR LYS HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	220.59Å 220.59Å 65.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.15 – 7.50 61.18 – 7.50	Depositor EDS
% Data completeness (in resolution range)	98.5 (55.15-7.50) 98.5 (61.18-7.50)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.89 (at 7.40Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.353 , 0.371 0.352 , 0.374	Depositor DCC
R_{free} test set	197 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å^2)	344.7	Xtriage
Anisotropy	0.661	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 399.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.067 for h,-k,-l	Xtriage
F_o,F_c correlation	0.71	EDS
Total number of atoms	9134	wwPDB-VP
Average B, all atoms (Å^2)	250.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.45% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	1.04	30/9324 (0.3%)	1.23	19/12638 (0.2%)

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	4

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	854	CYS	C-N	-21.05	0.85	1.34
1	A	700	CYS	C-N	20.20	1.72	1.34
1	A	557	CYS	C-N	-7.97	1.15	1.34
1	A	49	GLY	CA-C	6.34	1.62	1.51
1	A	626	ASN	CG-OD1	5.79	1.36	1.24

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	747	GLN	CG-CD-NE2	-30.31	43.95	116.70
1	A	557	CYS	CA-C-N	-25.68	60.70	117.20
1	A	557	CYS	C-N-CA	-18.83	74.63	121.70
1	A	747	GLN	CG-CD-OE1	-13.94	93.71	121.60
1	A	479	PRO	N-CA-C	8.16	133.31	112.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	557	CYS	Mainchain
1	A	854	CYS	Mainchain
1	A	863	ILE	Peptide
1	A	95	TYR	Peptide

5.2 Too-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 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	9134	0	8998	1436	10
All	All	9134	0	8998	1436	10

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 79.

The worst 5 of 1436 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:440:LYS:CD	1:A:538:LYS:HD2	1.28	1.55
1:A:700:CYS:C	1:A:701:PRO:N	1.72	1.43
1:A:440:LYS:HD3	1:A:538:LYS:CD	1.51	1.41
1:A:530:VAL:HG11	1:A:584:PRO:CD	1.50	1.38
1:A:854:CYS:C	1:A:855:THR:CA	1.93	1.37

The worst 5 of 10 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)
1:A:146:PHE:CE1	1:A:730:GLN:NE2[4_555]	0.96	1.24
1:A:269:THR:CG2	1:A:377:GLU:OE1[2_455]	1.33	0.87
1:A:268:SER:OG	1:A:375:ARG:NH1[2_455]	1.48	0.72
1:A:146:PHE:CD1	1:A:730:GLN:OE1[4_555]	1.76	0.44
1:A:269:THR:CG2	1:A:377:GLU:CD[2_455]	1.90	0.30

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	1150/1207 (95%)	1071 (93%)	57 (5%)	22 (2%)	8 38

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	PRO
1	A	181	LYS
1	A	191	LYS
1	A	410	ALA
1	A	465	ASN

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	1035/1067 (97%)	1002 (97%)	33 (3%)	39 61

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1024	ARG
1	A	1082	ILE
1	A	1177	LYS
1	A	575	LEU
1	A	567	ILE

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	9

The worst 5 of 9 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1139:PRO	C	1140:ASN	N	4.11
1	A	506:VAL	C	507:GLU	N	2.75
1	A	951:MET	C	952:THR	N	2.26
1	A	802:LYS	C	803:CYS	N	2.24
1	A	1036:VAL	C	1037:GLU	N	2.02

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.