Volume 8 Number 21 1980 Nucleic Acids Research

The electrostatic molecular potential of $tRNA^{Phe}$. IV. The potentials and steric accessibilities of sites associated with the bases

Richard Lavery, Alberte Pullman, Bernard Pullman and Marçal de Oliveira

Institut de Biologie Physico-Chimique, Laboratoire de Chimie Théorique, associé au C.N.R.S., 13, rue Pierre et Marie Curie, 75005 Paris, France

Received 13 August 1980

ABSTRACT

The sites of the 76 nucleic acid bases of $tRNA^{Phe}$ potentially reactive towards electrophiles are studied by calculations on the associated molecular electrostatic potentials and the static steric accessibilities. Each of these sites is treated in its environment within the macromolecule. The influence of various schemes of screening by countercations of the backbone phosphates on the electrostatic potentials is investigated. The possible significance of the potentials and accessibilities in connection with observed chemical reactivities is discussed.

INTRODUCTION

Our laboratory is presently engaged in a theoretical exploration of certain aspects of the fine structure of the transfer ribonucleic acid, $tRNA^{Phe}$. In the previous publications of this series we have presented the electrostatic potential around the macromolecule produced by the assembly of its 76 anionic phosphate groups (1), the potentials of its individual phosphates as a function of their conformations (2) and the potentials and accessibilities associated with each of these phosphates in their macromolecular environment (3). In this publication we extend our studies to the evaluation of the electrostatic potentials and the steric accessibilities of potentially reactive sites associated with the nucleic acid bases of $tRNA^{Phe}$. These potentials and accessibilities are calculated taking into account the contribution of all the subunits, phosphates, sugars and bases of the macromolecule and, in addition, the effects of countercation screening are also considered.

METHOD

The present calculations are based on the geometry of yeast $tRNA^{Phe}$ in its orthorhombic crystal form given by Sussman et al. (4). The technique of evaluating the molecular electrostatic potential of a macromolecule has al-

ready been described (e.g. see the first article of this series (1) or our papers on similar calculation on B-DNA (5) and Z-DNA (6)). The potential sites considered in this paper are the negative potential minima which occur around the nucleic acid bases and are generally associated with the atoms of these moieties susceptible to electrophilic attack. These sites are of two types : 1) those which occur in the plane of the bases, and are associated essentially with the pyrimidine-like ring nitrogens or the carbonyl oxygens, and 2) those which occur out of the plane of the bases, and are associated primarily with the pyrrole-type ring nitrogens, C8 or the amino groups of the purines and C5 of the pyrimidines (6). For the sites which occur out of the plane of the bases there are always two related potentials, one on each side of the base. In the tables of results which follow we present only the more negative of these two values, for brevity.

The technique of calculating static steric accessibility has also been described previously in its application to the phosphates of $tRNA^{Phe}$ (3). (See also ref. 7 and 8 for related calculations). As the attacking species we consider in the present calculations a sphere of 1.2 Å radius (the Van der Waals radius of the hydrogen atom) which has been shown in our previous studies (3) to reproduce well the accessibility of a macromolecule toward a water molecule, considered explicitly, via one of its hydrogen atoms. Although the results obtained for this particular simple species are, of course, rigorously valid only for this model, they certainly give an indication on the overall accessibilities to the different sites and allow to place an upper limit on the accessibility of these sites towards larger and more complex attacking molecules.

RESULTS AND DISCUSSION

A) The potentials and accessibilities of the base sites.

The results of the computations are presented in tables 1-8 The potentials reproduced in these tables have been calculated taking into account not only the macromolecule, $tRNA^{Phe}$, itself, but also four magnesium cations, Mg^{2+}, in the positions localised by x-ray studies on the orthorhombic crystal of $tRNA^{Phe}$ (4). These potentials thus refer to $tRNA^{Phe}$ in its crystal form.

It may be remarked that in these tables the entries of potential and/or accessibilities are absent for certain in-plane base sites e.g. cytosine N3 (notation C(N3)) and adenine N1. The reason for these absences is the occlusion of the corresponding sites by base-pair hydrogen bonding. This occlusion strongly (although not completely (9)) reduces the magnitude of the site potentials and

the accessibility of the associated base atoms. Nevertheless, reactions at these sites cannot be ruled out, because of now well-established transient fluctuational base unpairings (see e.g. 10).

There are, in addition, certain missing entries for the modified bases, which imply the absence of a corresponding site in these residues.

We shall consider the potentials and accessibilities successively for the different types of base in $tRNA^{Phe}$.

TABLE 1. Guanine site potentials (kcal/mole)

Residue	N1	N2	N3	N7	C8	O6	Order
1	-978	-949	-979	-1116	-1060	-1055	N7 > C8 > O6 > N3 > N1 > N2
3	-1068	-1028	-1044	-1189	-1109	-1170	N7 > O6 > C8 > N1 > N3 > N2
4	-1091	-1068	-1086	-1206	-1125	-1196	N7 > O6 > C8 > N1 > N3 > N2
15	-1195	-1138	-1156	-1302	-1189	-1302	N7 = O6 > N1 > C8 > N3 > N2
18	-973	-1033	-1043	-986	-931	-983	N3 > N2 > N7 > O6 > N1 > C8
19	-859	-870	-928	-908	-887	-861	N3 > N7 > C8 > N2 > O6 > N1
(20)	-1016	-1104	-1031	-952	-955	-1005	N2 > N3 > N1 > O6 > C8 > N7
22	-1147	-1121	-1149	-	-1126	-1126	N3 > N1 > C8 = O6 > N2
24	-1098	-1071	-1119	-1223	-1141	-1183	N7 > O6 > C8 > N3 > N1 > N2
30	-965	-930	-968	-1047	-966	-1021	N7 > O6 > N3 > C8 > N1 > N2
(34)	-668	-660	-713	-767	-744	-717	N7 > C8 > O6 > N3 > N1 > N2
42	-1059	-1026	-1041	-1169	-1093	-1145	N7 > O6 > C8 > N1 > N3 > N2
43	-1077	-1040	-1062	-1201	-1108	-1183	N7 > O6 > C8 > N1 > N3 > N2
45	-1130	-1146	-1181	-1225	-1129	-1229	O6 > N7 > N3 > N2 > N1 > C8
51	-1096	-1056	-1079	-1257	-1132	-1230	N7 > O6 > C8 > N1 > N3 > N2
53	-1073	-1034	-1055	-1172	-1047	-1176	O6 > N7 > N1 > N3 > C8 > N2
(57)	-931	-957	-987	-926	-892	-949	N3 > N2 > O6 > N1 > N7 > C8
65	-1160	-1131	-1144	-1248	-1161	-1241	N7 > O6 > C8 > N1 > N3 > N2
71	-1010	-991	-987	-1089	-983	-1085	N7 > O6 > N1 > N2 > N3 > C8
10	-1078	-1060	-1067	-1182	-1133	-1105	N7 > C8 > O6 > N1 > N3 > N2
26	-1093	-1112	-1100	-1187	-1075	-1186	N7 > O6 > N2 > N3 > N1 > C8
46	-1133	-1166	-1237	-	-1072	-1152	N3 > N2 > O6 > N1 > C8

1) Guanine

The potentials (in kcal/mole) at the different sites associated with that base are shown in table I. In the first column is given the number of the base residue, ringed numbers denoting bases not involved in interbase hydrogen bonding ; the next six columns indicate the values of the potentials, each column being headed with the name of the atom of the base with which a particular site is associated ; in the final column the order of the sites, with respect to decreasingly negative potential, is given for each residue.

Several general points may be noted immediately from Table I. Firstly, the potentials are strongly negative, of the order of -800 to -1300 kcal/mole. These intensities are principally due to the strong, combined effect of potentials of the anionic phosphate groups, which superpose on the potentials inherent to the bases. Secondly, the values of potential differ appreciably for a given site from one guanine to another. This is due to the irregular and complex folded structure of $tRNA^{Phe}$. Thirdly, the ordering of the potentials of the sites varies from one guanine to another. This situation, again due to the folded structure of $tRNA^{Phe}$, is significantly different from that in a nucleic acid of regular structure, such as B-DNA, where each base in a given segment of the helix has very similar site potentials and site orderings, apart from relatively weak perturbations due to the neighbouring base sequence. $tRNA^{Phe}$ thus presents a much more complicated picture of base site potentials, which depend significantly not only on the nature of the base concerned, but also on its location in the nucleic acid.

Considering the guanine potentials in more details we may note that the most negative potential for each base is commonly at N7, although N3 (residues 18, 19, 22, 46 and 57) and O6 (residues 45 and 53) may also occupy this position. It is interesting to note that N2, the guanine amino group site, which is generally associated with a relatively weak negative potential and is often the last site in the ordering also occurs as first, most negative site for one guanine : residue 20. The ordering fo the remaining sites is very variable between the different guanines.

The average potentials associated with each base differ considerably and, from table 1, we may deduce that e.g. guanines 19 and 34 have weak average potentials (< -1000 kcal/mole) while guanines 15, 45, 51 and 65 have strong average negative potentials (> -1100 kcal/mole). These results can be correlated with the structure of $tRNA^{Phe}$ and more specially with the special concentration of the phosphates around a given base.

The residues which occur below the horizontal line at the bottom of table

1 are the modified guanines : 10 (N2-methylguanine), 26 (N2-dimethylguanine) and 46 (7-methylguanine). The average potentials of these bases and the site ordering are similar to the normal guanines. It is interesting to note that the positive charge carried by 7-methylguanine 46 does not distinguish this residue from its close neighbour, guanine 45, in terms of average potential. The dominating influence of the summed phosphate potentials is thus exemplified.

TABLE 2. Guanine atom accessibilities ($\AA^2$)

Residue	N1	N2	N3	N7	C8	O6	Order
1	0.0	0.09	1.21	5.33	2.11	2.34	N7 > O6 > C8 > N3
3	0.0	0.14	0.09	4.36	0.79	2.02	N7 > O6 > C8 > N2 > N3
4	0.0	0.09	0.60	2.60	0.53	0.97	N7 > O6 > N3 > C8 > N2
15	0.0	0.0	0.60	3.38	1.0	3.31	N7 > O6 > C8 > N3
18	0.0	0.0	0.37	4.13	0.53	3.55	N7 > O6 > C8 > N3
19	0.93	1.76	1.95	5.47	1.0	7.35	O6 > N7 > N3 > N2 > C8 > N1
(20)	1.02	1.81	1.39	5.75	2.27	5.17	N7 > O6 > C8 > N2 > N3 > N1
22	0.0	0.0	0.0	-	0.0	0.0	-
24	0.0	0.0	0.97	1.71	0.0	0.28	N7 > N3 > O6
30	0.0	0.32	0.0	3.11	0.11	1.90	N7 > O6 > N2 > C8
(34)	0.88	1.90	1.21	5.05	2.06	7.31	O6 > N7 > C8 > N2 > N3 > N1
42	0.0	0.0	0.56	4.03	0.74	2.71	N7 > O6 > C8 > N3
43	0.0	0.05	0.19	4.22	0.74	2.42	N7 > O6 > C8 > N3 > N2
45	0.0	0.0	0.0	3.75	1.05	6.30	O6 > N7 > C8
51	0.0	0.09	0.42	3.29	0.58	1.74	N7 > O6 > C8 > N3 > N2
53	0.0	0.0	0.0	3.57	0.79	2.46	N7 > O6 > C8
(57)	0.0	0.09	0.42	0.0	0.0	3.55	O6 > N3 > N2
65	0.0	0.05	0.19	4.17	0.69	2.26	N7 > O6 > C8 > N3 > N2
71	0.0	0.0	0.28	2.04	0.21	1.09	N7 > O6 > N3 > C8
10	0.0	0.0	0.74	0.0	0.79	0.0	C8 > N3
26	0.0	0.09	0.37	0.51	0.0	2.91	O6 > N7 > N3 > N2
46	0.0	0.0	0.0	-	0.95	0.28	C8 > O6

We now turn to the calculated accessibilities of the guanine atoms corresponding to sites for which the potentials have been studied. The results are shown in table 2, the presentation being similar to that for the potential sites. Note that zero accessibilities have not been included in the ordering in the last column of this table. It may be seen that the accessible areas of the guanines vary considerably passing from one residue to the next, both in terms of their average values and of their ordering for each base. There are, however, certain similarities between the bases, notably the inaccessibility of N1 in all but three cases (19, 20 and 34) and the high accessibilities of N7 and O6. C8 is also relatively accessible for most guanines, while N3 is somewhat less accessible. N2 has low or zero accessibility in all but three residues : 19, 20 and 34. It might be expected that the unpaired bases (ringed in table 2) would show particularly high accessibilities and this is so for residues 20 and 34, but not for 57 which has rather low accessibilities, apparently because of its internal position in the TψC loop and its stacking interactions with bases 56 and 58. In contrast, guanine 19, which is paired with cytosine 56, has high accessibilities being exposed on the outer surface of the macromolecule between the D and TψC loops, positioned similarly to the unpaired base 20. An exceptional case of steric hindrance is seen for guanine 22 for which no site was found to be accessible. This base is situated in the core of the $tRNA^{Phe}$ molecule, is involved in the base triplet 22-13-46 and stacked between the pair 15-48 and the triplet 23-12-9.

Finally, comparing for each guanine the orderings of the potentials in table 1 and of the accessibilities in table 2, we remark a certain correlation between these two properties. In most residues N7 and O6 are sites associated with the deepest potentials and also with large accessibilities while N2 generally have weak potentials and small accessibilities. The remaining sites are intermediately placed in both tables with the exception of the particularly low accessibility of N1.

2) Adenine

The potentials are presented in table 3. The final entry at the bottom of this table is the modified base 1-methyladenine 58.

As for the guanines, the average potentials and the site orderings differ considerably for the various adenines. Thus, adenines 35, 36 and 76 have particularly weak average potentials, of the order of -600 to -800 kcal/mole, while those for residues 9, 64, 66 and 67 are strong, -1100 to -1250 kcal/mole. For the individual bases N7, N6 and, somewhat less commonly N3, have the most

negative potentials, while the potentials for N1 (this site is often occluded by hydrogen bonding) and C8 are generally weaker. For adenine 38, however, C8 becomes the most negative site. Contrasting these results with those for guanine one may note the relatively more negative potentials associated with the adenine amino groups in comparison with those of the guanine amino groups.

The accessibilities for adenines are given in table 4. Here we may note the generally large accessibilities at N7 and N3, and the small values at N6. The accessibilities at C8 are intermediate in magnitude with the exception of

TABLE 3. Adenine site potentials (kcal/mole)

Residue	N1	N3	N6	N7	C8	Order
5	-	-1145	-1176	-1223	-1165	N7 > N6 > C8 > N3
9	-1241	-1217	-1193	-	-1156	N1 > N3 > N6 > C8
14	-1238	-1106	-1147	-	-1170	N1 > C8 > N6 > N3
(21)	-1184	-1161	-1195	-1139	-1150	N6 > N1 > N3 > C8 > N7
23	-	-1132	-1167	-	-1154	N6 > C8 > N3
29	-	-1003	-1060	-1090	-987	N7 > N6 > N3 > C8
31	-921	-928	-973	-992	-922	N7 > N6 > N3 > C8 > N1
(35)	-742	-735	-744	-811	-783	N7 > C8 > N6 > N1 > N3
(36)	-757	-776	-736	-805	-725	N7 > N3 > N1 > N6 > C8
(38)	-836	-842	-831	-838	-843	C8 > N3 > N7 > N1 > N6
44	-	-1081	-1128	-1207	-1118	N7 > N6 > C8 > N3
62	-	-1064	-1169	-1205	-1119	N7 > N6 > C8 > N3
64	-	-1105	-1203	-1215	-1130	N7 > N6 > C8 > N3
66	-	-1173	-1220	-1246	-1165	N7 > N6 > N3 > C8
67	-	-1183	-1209	-1240	-1145	N7 > N6 > N3 > C8
(73)	-895	-948	-962	-994	-889	N1 > N7 > N6 > N3 > C8
(76)	-671	-679	-580	-614	-582	N3 > N1 > N7 > C8 > N6
58	-	-1118	-978	-	-1058	N3 > C8 > N6

the large value of 3.01 $\mathring{A}^2$ for the terminal acceptor end base, adenine 76. This exposed base is in fact highly accessible at all sites. The modified base 1-methyladenine 58 shows only a small accessibility at N3. Five adenines, residues 35, 36, 38, 73 and 76, show quite high accessible areas at N1. All these bases are unpaired. In contrast, adenine 21 which is also unpaired is inaccessible at N1, N3 and N6, apparently because of its orientation towards the core of the molecule and the consequent hidrances caused by uracil 47 and the stacking with guanine 22 and 7-methylguanine 46.

TABLE 4. Adenine atom accessibilities ($\mathring{A}^2$)

Residue	N1	N3	N6	N7	C8	Order
5	-	1.07	0.05	1.90	0.0	N7 > N3 > N6
9	1.11	0.70	0.70	-	0.0	N1 > N3 > N6
14	0.83	2.09	0.0	-	0.0	N3 > N1
(21)	0.0	0.0	0.0	2.22	0.42	N7 > C8
23	-	1.90	0.0	-	0.26	N3 > C8
29	-	0.88	0.37	3.11	0.74	N7 > N3 > C8 > N6
31	0.0	0.93	0.56	1.02	0.0	N7 > N3 > N6
(35)	2.09	1.11	0.0	0.0	0.0	N1 > N3
(36)	3.62	1.25	0.05	0.46	0.32	N1 > N3 > N7 > C8 > N6
(38)	2.50	0.88	0.23	0.46	0.0	N1 > N3 > N7 > N6
44	-	1.99	0.0	1.48	0.0	N3 > N7
62	-	1.44	0.09	1.99	0.05	N7 > N3 > N6 > C8
64	-	0.97	0.0	2.32	0.84	N7 > N3 > C8
66	-	0.74	0.05	2.69	0.69	N7 > N3 > C8 > N6
67	-	1.21	0.05	2.09	0.26	N7 > N3 > C8 > N6
(73)	2.46	0.93	0.14	1.44	0.21	N1 > N7 > N3 > C8 > N6
(76)	5.89	1.16	3.43	6.40	3.01	N7 > N1 > N6 > C8 > N3
58	-	0.14	0.0	-	0.0	N3

Comparing potentials and accessibilities for the adenines, N7 and N3 are seen to be well placed for most residues while the relatively deep potentials of N6 are contrasted by relatively low accessibilities.

3) Cytosine

The potentials at cytosine sites are contained in table 5. At the bottom of this table are included the modified bases 5-methylcytosines 40 and 49.

TABLE 5. Cytosine site potentials (kcal/mole)

Residue	N3	N4	O2	C5	Order
2	-	-1091	-985	-1074	N4 > C5 > O2
11	-	-1133	-1075	-1114	N4 > C5 > O2
13	-	-1205	-1131	-1188	N4 > C5 > O2
25	-	-1144	-1074	-1143	N4 > C5 > O2
27	-	-1111	-1053	-1069	N4 > C5 > O2
28	-	-1091	-994	-1049	N4 > C5 > O2
(32)	-874	-921	-834	-865	N4 > N3 > C5 > O2
48	-	-1116	-1289	-1141	O2 > N4 > C5
56	-	-880	-839	-890	C5 > N4 > O2
(60)	-1099	-1065	-1193	-1112	O2 > C5 > N3 > N4
61	-	-1179	-1049	-1139	N4 > C5 > O2
63	-	-1164	-1052	-1151	N4 > C5 > O2
70	-	-1104	-1012	-1042	N4 > C5 > O2
72	-	1017	-933	-970	N4 > C5 > O2
(74)	-926	-891	-911	-885	N3 > O2 > N4 > C5
(75)	-857	-842	-830	-837	N3 > N4 > C5 > O2
40	-	-1028	-978	-1058	C5 > N4 > O2
49	-	-1235	-1187	-1254	C5 > N4 > O2

For cytosine the site associated with N3 is occluded by hydrogen bonding for all but the four unpaired residues : 32, 60, 74 and 75. Of the three remaining sites N4 is most commonly associated with the deepest potential and O2 with the weakest. Once more the bases from differing regions of the macromolecule are associated with different average potentials, the weakest being those for cytosines 32, 56, 74 and 75 and the strongest those for cytosines 13, 48 and 49. It is interesting to note that in both the 5-methylcytosines

TABLE 6. Cytosine atom accessibilities ($\AA^2$)

Residue	N3	N4	O2	C5	Order
2	-	0.19	1.49	0.74	O2 > C5 > N4
11	-	0.0	1.13	0.42	O2 > C5
13	-	0.0	0.24	0.21	O2 > C5
25	-	0.0	0.52	0.0	O2
27	-	0.09	0.32	0.47	C5 > O2 > N4
28	-	0.0	0.73	0.74	C5 > O2
(32)	1.62	0.19	2.18	0.26	O2 > N3 > C5 > N4
48	-	0.0	0.32	0.11	O2 > C5
46	-	1.48	3.23	1.69	O2 > C5 > N4
(60)	1.07	0.14	4.48	0.42	O2 > N3 > C5 > N4
61	-	0.05	0.97	0.21	O2 > C5 > N4
63	-	0.05	1.05	1.11	C5 > O2 > N4
70	-	0.14	0.89	1.21	C5 > O2 > N4
72	-	0.0	0.85	0.16	O2 > C5
(74)	2.69	0.42	6.18	0.0	O2 > N3 > N4
(75)	3.62	1.71	7.95	1.69	O2 > N3 > N4 > C5
40	0.0	0.19	2.22	0.16	O2 > N4 > C5
49	0.0	0.56	1.01	0.32	O2 > N4 > C5

(40 and 49), the C5 site has the most negative potential, which for the unmodified bases is only the case for cytosine 56.

Passing to the cytosine accessibilities (table 6), the dominance of the values for O2 and C5 may be seen. N4 is only slightly accessible with the exception of cytosines 56 and 75 which are both exposed ; 56 between the D and TψC loops and 75 in the acceptor end.

Sites O2 and C5 of cytosines are thus calculated to have both relatively strong negative potentials and large accessibilities for the majority of these residues.

TABLE 7. Uracil site potentials (kcal/mole)

Residue	N3	C5	O2	O4	Order
6	-1187	-1212	-1152	-1237	O4 > C5 > N3 > O2
7	-1201	-1226	-1174	-1241	O4 > C5 > N3 > O2
8	-1244	-1290	-1123	-1285	C5 > O4 > N3 > O2
12	-1109	-1113	-1067	-1183	O4 > C5 > N3 > O2
(33)	-853	-863	-737	-954	O4 > C5 > N3 > O2
41	-1032	-1063	-947	-1071	O4 > C5 > N3 > O2
(47)	-1086	-1092	-1098	-1037	O2 > C5 > N3 > O4
50	-1169	-1219	-1081	-1245	O4 > C5 > N3 > O2
52	-1086	-1112	-1015	-1206	O4 > C5 > N3 > O2
(59)	-1122	-1099	-1180	-1050	O2 > N3 > C5 > O4
68	-1145	-1142	-1101	-1205	O4 > N3 > C5 > O2
69	-1103	-1093	-	-1283	O4 > N3 > C5
(16)	-978	-	-1055	-943	O2 > N3 > O4
(17)	-800	-	-971	-792	O2 > N3 > O4
38	-928	$\overset{N1}{-958}$	-939	-918	N1 > O2 > N3 > O4
55	-1050	$\overset{N1}{-1028}$	-1101	-1004	O2 > N3 > N1 > O4
54	-1034	-1034	-1141	-1167	O4 > O2 > N3 = C5

4) Uracil

The potentials for uracil are contained in table 7. Included at the foot of this table are the modified bases dihydrouracil (16 and 17), pseudouracil (39 and 55) and thymine (54), which may be considered as a modified uracil in the present circumstances. We note that in the pseudouracils 39 and 55, which are linked to the sugar via C5, the potential associated with N1 replaces that for C5, which occurs in the equivalent geometrical position in the unmodified uracils.

The most negative potentials for the different uracils are found to be

TABLE 8. Uracil atom accessibilities ($\mathring{A}^2$)

Residue	N3	C5	O2	O4	Order
6	0.0	0.69	1.13	3.15	O4 > O2 > C5
7	0.0	0.84	2.30	1.45	O2 > O4 > C5
8	0.0	1.21	0.0	0.69	C5 > O4
12	0.0	0.16	2.18	0.44	O2 > O4 > C5
(33)	0.19	0.32	0.65	3.47	O4 > O2 > C5 > N3
41	0.0	1.32	1.41	2.06	O4 > O2 > C5
(47)	1.11	3.69	4.85	11.39	O4 > O2 > C5 > N3
50	0.0	1.27	1.41	2.42	O4 > O2 > C5
52	0.0	0.90	1.29	3.19	O4 > O2 > C5
(59)	0.0	0.0	5.13	3.96	O2 > O4
68	0.0	0.05	1.49	1.45	O2 > O4 > C5
69	0.0	1.32	-	4.85	O4 > C5
(16)	1.62	-	8.84	11.10	O4 > O2 > N3
(17)	1.67	-	3.39	11.14	O4 > O2 > N3
39	0.0	$\overline{N1}$ 0.14	3.03	2.38	O2 > O4 > N1
55	0.0	$\overline{N1}$ 0.56	7.19	0.0	O2 > N1
54	0.0	0.0	0.0	2.18	O4

those associated with O4. O2 is generally the weakest of the sites studied but in the case of five residues, 16, 17, 47, 55 and 59, this situation is reversed and it becomes the most negative site. For all these residues however, O2 is not involved in hydrogen bonding. N3 and C5 are generally associated with intermediate potentials with the exception of C5 for uracil 8, in which it is the most negative site.

The weakest average potentials for the uracil residues are those of the dihydrouracils 16 and 17 in the D loop and also of uracil 33, while the deepest average potentials are calculated for uracils 6, 7, 8, 50 and 68.

Table 8 shows the accessibilities at the uracil atoms. One may note the largest values for O2 and O4 for all residues with the exception of residues 8 and 55 where O2 and O4 are, respectively, inaccessible. For uracil 8, C5 has a large accessibility and it will be recalled that this atom is also associated with the most negative potential for this base. These phenomena are most probably explained by the unusual, reverse Hoogsteen, base pairing of uracil 8 with adenine 14 via its N3-H and O2 centers. The N3 atom is generally inaccessible, the exceptions being for the unpaired bases 16, 17, 33 and 47.

Comparing tables 7 and 8, O4 of uracil is seen to be associated most commonly with the strongest potentials and largest accessibilities, but this is also true for O2 when this atom is not involved in hydrogen bonding.

B) The effects of countercation screening

At this point, it may seem relevant to ask whether the potentials for $tRNA^{Phe}$ in the presence of the four crystallographically located Mg^{2+} cations as presented in the preceeding section, are relevant in describing $tRNA^{Phe}$ in the presence of other cation arrangements, as could be the case for the molecule in solution. To tentatively answer this question we studied two new situations, namely, $tRNA^{Phe}$ in the absence of any cation and $tRNA^{Phe}$ in the presence of both the four Mg^{2+} and twenty Na^{+} cations. Both these studies were carried for the test case of the guanine residues. The latter ions were bound to the bridge sites between the phosphate anionic oxygens, a position identical to that employed in studying the Na^{+} cation screening of B-DNA (19). Their binding to given phosphates and their number was arrived at by an optimisation procedure which searched to locate, for each new ion, the bridge site with the most negative potential, avoiding at the same time too close approaches between the various countercations (limiting distances of Na^{+} -Na^{+} 7Å and Na^{+} -Mg^{2+}, 14Å were imposed). The phosphates found in that way to bind Na^{+} ions are residues 1, 3, 6, 15, 24, 25, 31, 44, 45, 51, 53, 57, 61, 63, 66, 68, 70, 72, 74

and 76. This screening does not pretend to represent the true countercation distribution in solution, but is rather intended to illustrate one of many possible models of a more complete screening of $tRNA^{Phe}$ than that which exists in its orthorhombic crystal form in association with the four localized Mg^{2+} ions.

For reasons of economy of space imposed by the editorial rules we shall not reproduce the numerical results here. They may be obtained upon request. Here we simply wish to indicate that the order of potential sites for each base is not greatly changed between the three approximations investigated and that the same is true the ordering of the guaninesin terms of their average potentials.

C) Correlations with experimentally observed chemical reactivities of $tRNA^{Phe}$.

It has been demonstrated, notably by our earlier studies of B-DNA (5, 12), that both electrostatic potentials and steric accessibility can, and often do, play a significant role in determining the chemical reactivity of sites within a macromolecule. We similarly investigated the possible role of these factors for several experimentally studied reactions of $tRNA^{Phe}$ with electrophiles. The results of the previous sections indicate, however, by themselves the limitations of such an approach inasmuch as they show that strong negative potentials, which would imply greater affinity for a reaction with electrophiles, often occur for bases with low accessibilities, which disfavour reactions and that converse situation is also common, weak potentials being associated with bases having high accessibilities. It is also obvious a priori that these are not the only factors involved in the reactions and that as illustrated by ourselves on numerous occasions (13-15), polarisation and charge transfer are important components, besides the electrostatic one in the global interaction energy between two species. Nevertheless the previous success of this approch encourages its futur exploration.

We have studied four reactions of $tRNA^{Phe}$: with (i) kethoxal (16-17) (ii) $I_2/TlCl_3$ (18), (iii) carbodiimide (19) and (iv) N-acetoxy-2-acetylaminofluorene (20-23). Each of these reactants interacts with specific sites on specific types of base within $tRNA^{Phe}$. We were concerned only with trying to understand to what an extent may the computed potentials and accessibilities account for the reactivity or non reactivity of certain bases in given structural regions of the macromolecule toward a given reactant. A full explanation of the site specificity of these reagents implies more complete quantum mecha-

nical studies of their respective reaction mechanisms, which is beyond the scope of this publication (see e.g. 13-15).

Because of space limitation we can present here the results for only one of the above quoted reactions. A more complete study will be presented separately.

The exemple selected concerns kethoxal. This compound reacts specifically with the guanine bases of $tRNA^{Phe}$ and in the studies of Litt (16-17) only two guanines, numbers 20 and 34, are found to bind this reagent. The reaction product is a bridged adduct resulting from substitutions at guanines N1 and N2.

In figure 1 we have combined the calculated data for the accessibilities (horizontal axis) and potentials (vertical axis). Each point on this graphic therefore represents, simultaneously, the potential and the atom accessibility for a given base site. On this graphic we have, in this manner, placed dots representing the N1 and N2 sites of all the guanines of $tRNA^{Phe}$. Sites with zero accessibility, consequently unlikely to be reactive, are not included.

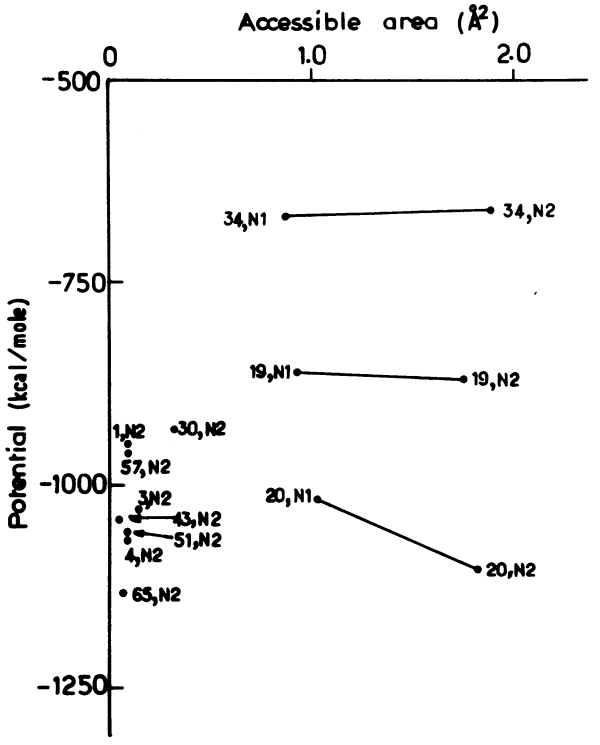

Fig. 1 . Potential and accessibility for the guanine N1 and N2 sites in $tRNA^{Phe}$.

When both the N1 and N2 sites of a given base appear on this diagram they have been joined by a line. It will be noted that this is the case only for guanines 19, 20 and 34. For the other bases the accessibilities of N1 or N2 or of both are zero (see table 2). The three remaining bases are seen to have high accessibilities for both sites, although N2, in each case, is somewhat more accessible than N1. The N1 and N2 sites of guanine 20 have high negative potentials, while those of guanine 34 are considerably lower, guanine 19 being intermediately placed.

The observed reactivity of guanines 20 and 34 is consequently understandable. The inactivity of guanine 19 may be due to the fact that while guanines 20 and 34 are unpaired, guanine 19 is bound to cytosine 56 with hydrogen bonds that involve its N1 and N2 atoms and these bonds would have to be disrupted before a reaction with kethoxal could take place.

CONCLUSIONS

In the present publication the electrostatic potentials and the steric accessibilities of all the sites on the nucleic acid bases of yeast $tRNA^{Phe}$, which are susceptible to electrophilic attack, have been calculated.

It is shown that the potentials and accessibilities of specific sites on each type of base are highly variable, depending strongly on the positionning of the base inside the macromolecule. This variability applies both to the average potential or accessibility associated with the sites of the bases and to the ordering of the sites, in terms of these values, for each individual base.

The effect of different countercation screenings of the backbone phosphates has also been investigated. It is found that increased screening diminishes the absolute values of the base site potentials and also their spacing, but does not cause much reordering.

A clear correlation between structure and potential may be noted, bases in the acceptor end or the loops generally having much weaker potentials than those situated in the helical segments, closer to the core of the macromolecule. The accessibility of certain bases in the acceptor end or the loops is also particularly high compared to other regions, but the structural correlations with this property are less clear due to its sensitivity to the local structure surrounding each individual base atom.

Finally partial correlations between the combined potentials-accessibilities for the base sites and experimentally observed chemical reactivities of

$tRNA^{Phe}$ have been noted. More detailed reactivity studies are required before definite conclusions on the relative roles of the potential and accessibility and their relation to other factors involved (e.g. polarization, charge transfer, etc.) can be made. Such studies are under way.

Acknowledgments - This work was supported by the contract n°14 of the Institut National de la Recherche Medicale of France with its ATP 77 79-109 in Chemical Carcinogenesis. One of us (M.O.N.) wishes to acknowledge the support of a doctoral fellowship from Conselho Nacional de Desenvolvimento Cientifico e Tecnologico, Brazil.

References

1 Lavery, R., Pullman, A. and Pullman, B., (1980) Nucl. Acids Res. 8, 1061-1078.
2 Lavery, R., de Oliviera, M. and Pullman, B., (1980) J. Computational Chem., 1, 301-306.
3 Lavery, R., Pullman, A. and Pullman, B., Theoret. Chim. Acta, in press.
4 Sussman, J.L., Holbrook, J.B., Warrant, R.W., Church, G.M. and Kim, S-H., (1978) J. Mol. Biol. 123, 607-630. We have utilized in this work a further refined coordinate system obtained as a personal communication from Dr. S-H. Kim.
5 Pullman, B., Perahia, D. and Cauchy, D. (1979) Nucl. Acids. Res. 6, 3821-3830.
6 Pullman, A. in Mechanismes d'Alteration et de Reparation du DNA, Relations avec la Mutagénèse et la Cancérogénèse Chimique. Colloque du C.N.R.S. 256 Paris 1976, p. 103-113.
7 Thiyagarajan, P. and Ponnuswamy, P.K. (1979) Biopolymers, 18, 2233-2247.
8 C.J., Alden and S-H. Kim, (1979) J. Mol. Biol. 132, 411-434
9 Pullman, A. and Berthod, H., (1978) Theoret. Chim. Acta, 48, 269-277.
10 Kallenbach, N.R., Mandal, C. and Englander, S.W. in Stereodynamic of Molecular Systems (R.H. Sarma Ed.) Pergamon Press, New York, (1979) p. 271-282.
11 Lavery, R., Cauchy, D., Rojas, O. and Pullman, B. (1980) Int. J. Quant. Chem. Biol. Symp. 7, in press.
12 Lavery, R., Pullman, A. and Pullman, B., Bioch. Biophys. Acta, in press.
13 Pullman, A. and Armbruster, A.M., (1977) Theoret. Chim. Acta. 249-255.
14 Pullman, A. and Armbruster, A.M., (1979) Theoret. Chim. Acta. 5°, 359-361.
15 Perahia, D., Pullman, A. and Pullman, B., (1977) Theoret Chim. Acta. 43, 207-212.
16 Litt, M. (1969) Biochemistry, 8, 3249-3253.
17 Litt, M. and Greenspan, C.M. (1972) Biochemistry, 11, 1437-1442.
18 Batey, I.L. and Brown, D.M. (1977) Biochim. Biophys. Acta, 474
19 Rhodes, D. (1975) J. Mol. Biol. 94, 449-460.
20 Fujimura, S., Grunberger, D., Carvajal, G. and Weinstein, I.B. (1972) Biochemistry, 11, 3629-3635.
21 Pulkrabek, P., Grunberger, D. and Weinstein, I.B. (1974) Biochemistry, 13, 2414-2419.
22 Sprinzl, M., Grueter, F., Spelzhaus, A. and Gauss, D.H. (1980) Nucl. Acid. Res. 8, r1-r22.
23 Massouh-Rizk, L., Thèse d'Etat, (1975), Université Louis Pasteur de Strasbourg, France.