SPECIFICITY IN PROTEIN–NUCLEIC ACID INTERACTION

Solubility study on amino acid–nucleoside interaction

P. D. THOMAS and S. K. PODDER
Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Received 1 September 1978

1. Introduction

The chemical basis of the specificity of protein–nucleic acid interaction, as seen in many biochemical phenomena such as the organization of nucleo–protein complexes (chromatin, ribosomes) and gene expression and its regulation, is not yet understood. A knowledge of such specific interactions is also essential for tracing the chemical evolution of life based on the coupling between protein and nucleic acid and the origin of genetic code [1,2]. From the threshold concentration required for precipitation, the interaction of basic polypeptides with nucleotides was shown to be base specific [3,4]. The affinity of oligopeptides, especially those containing aromatic amino acid residues, to DNA was found to depend on the sequence [5,6]. The ORD studies on basic polypeptide–nucleotide complexes have shown that the binding specificities are related to the ease of polypeptide helix formation and to the tendency of the bases to stack [7]. Thus a variety of interactions can occur, depending on state of polymerization, composition, sequence, conformation, and environment of the reacting species. However, all the data so far available are of a qualitative nature. Thus we have been trying to evaluate the thermodynamic parameters of the interaction by various methods such as NMR, gel-filtration, equilibrium dialysis and solubility measurements. Here we wish to report our results of solubility measurements on amino acid– nucleoside systems.

The interaction between amino acid and nucleoside is so weak that one cannot study it by conventional spectroscopic techniques. But the measurement of solubility of one component in presence of the other offers a good and simple method of following such weak interactions. The method is based on the assumption that the increase in solubility of one component in the presence of the other corresponds to complex formation. Various authors have used solubility measurements to study weak interactions of biological interest such as DNA–denaturant interaction [8,9], protein–denaturant interaction [10,11], amino acid–amino acid interaction stabilizing tertiary structure of proteins [12], nucleic acid base–water interaction [13], substituted uracil–pentapeptide interaction [14] and purine–amino acid interaction [15]. Thus it seems quite reasonable to extend the solubility method to study, quantitatively, protein–nucleic acid interactions starting from monomer–monomer systems. The solubility data presented here do show a sort of selectivity in nucleoside–amino acid interaction, which in principle can give rise to the specificity of interaction observed at the polymer level. The unusually high solubility of Phe in uridine (U) as compared with other nucleosides and the preferential interaction of Gly with guanosine (G) as compared with adenosine (A) are examples of such selectivity. Purine–amino acid complexes are stabilized mainly by hydrophobic and stacking interactions. The relevance of these findings in tracing the evolution of genetic code is also discussed.

2. Materials and methods

Amino acids and nucleosides used were obtained from Sigma Chemical Co, MO. For purines, the solu-

bility of nucleoside in amino acid solutions was measured. With pyrimidines, because of their high solubility, the same method is difficult to apply and hence the solubility of amino acid in nucleoside solutions was measured. To amino acid (or nucleoside) solutions of varying concentration containing 0.3 M NaCl, excess of nucleoside (or amino acid) was added, kept at 20°C or 25°C for 7 days with frequent shaking, the undissolved nucleoside (or amino acid) was filtered off using millipore filters, and the concentration of the nucleoside (or amino acid) in the filtrate was estimated spectrophotometrically after appropriate dilution. All the solutions had a pH in the range 5.3–5.7. The spectra were recorded using either a Cary-14 or a Beckman Model-25 Spectrophotometer. Where possible, literature values of extinction coeffi-cients were used for calculation [16].

2.1. *Calculation of equilibrium constant*

It is assumed that the increase in solubility of a component B in a solution of A is due to complex formation:

Case I, 1:1 Complex:

$$A + B \rightleftharpoons AB;\ K_d\ (M) = \frac{[A]\ [B]}{[AB]}$$

Let n be the molar concentration of complex formed, K_s the solubility of B in pure solvent (i.e., in the absence of A) and $[A]_o$ and $[B]_o$ the total concentrations of A and B in the liquid phase, respectively:

$$[B]_o = [B] + n,\ [B] = K_s$$

$$[A] = [A]_o - n = [A]_o - [B]_o + K_s$$

$$[AB] = n = [B]_o - K_s$$

Substitution of above relations in the expression for K_d gives:

$$[B]_o = \frac{K_s}{K_s + K_d}\ [A]_o + K_s \qquad (1)$$

Case II, 1:2 Complex:

$$A + 2B \rightleftharpoons AB_2;\ K_d\ (M^2) = \frac{[A]\ [B]^2}{[AB_2]}$$

Using the above notations:

$$[B]_o = [B] + 2n,\ [B] = K_s$$

$$[A] = [A]_o - n = [A]_o - \frac{[B]_o - K_s}{2}$$

$$[AB] = n = \frac{[B]_o - K_s}{2}$$

Substituting these in the expression for K_d gives:

$$[B]_o = \frac{2K_s^2}{K_s^2 + K_d}\ [A]_o + K_s \qquad (2)$$

In both the cases K_d can be calculated from the slope and intercept of a plot of $[B]_o$ vs $[A]_o$.

3. Results and discussion

Typical solubility curves for nucleoside in amino acid solutions and amino acid in nucleoside solutions are shown in fig.1 and 2, respectively. The curves are almost linear with different slopes for different amino acid–nucleoside systems. From the slope, the dissociation constants (K_d) were calculated using eq. (1). For the Phe–U system, the slope of the curve does

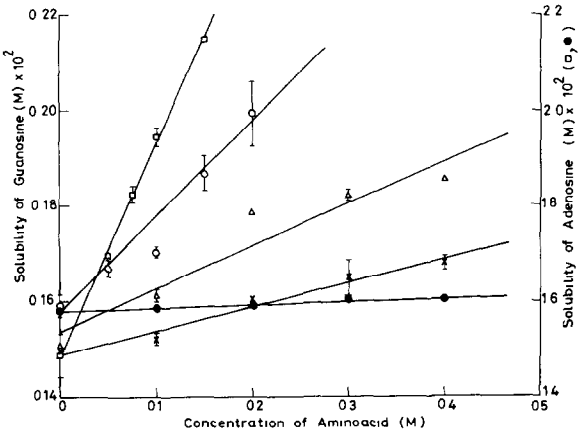

Fig.1. Solubility curves for nucleosides in amino acid solutions. Solubility curves of guanosine in Ser (×), Val (△) and Lys (○) and of adenosine in Gly (●) and Phe (□) are given. All the curves are least square lines of the mean of two readings for each point.

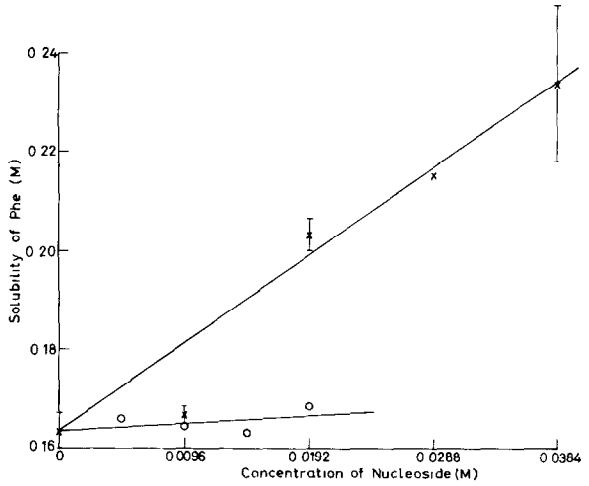

Fig.2. Solubility curves for amino acids in nucleoside solutions. Solubility curves of Phe in uridine (×) and cytidine (○) are given.

not fit eq. (1) and hence K_d was calculated using eq. (2). The K_a values (defined as $1/K_d$) are given in table 1.

In monomer–monomer systems one cannot expect to observe 'specificity' of the sort implied in the biological usage of the term but there should be 'selectivity' of some sort even at this simplest level [17]. This type of selectivity of interaction can be seen in the data presented in table 1. In addition to the experimental convenience, by choosing nucleoside rather than nucleotide, we have eliminated the possibility of electrostatic interactions involving the phosphate group, though we are neglecting backbone interactions involving phosphate groups. The pH range used was such that nucleosides were uncharged. The presence of 0.3 M NaCl in the system reduced non-specific electrostatic interactions to a large extent. From table 1, it is evident that even structurally and functionally related amino acids [18] such as Ser and Thr, and Val and Leu, do show difference in affinity towards G. Thus it seems quite resonable to ascribe the specificity of observed amino acid–nucleoside interactions mainly to factors other than electrostatic ones. Obviously the specificity is determined by the nature of the amino acid side chain as well as of the base.

In general, the interaction of the purines with amino acids increases with increasing 'hydrophobicity' [19,20] of the amino acid side chain. Such a correlation was also observed [17] in a study of selectivity coefficients obtained from the elution patterns of nucleotides through a resin-bound amino acid column. The much higher K_a values for aromatic amino acids compared to others indicate the possible role of stacking interactions in stabilizing purine–aromatic amino acid complexes. The greater interaction of Trp compared to Phe also seems to be partly due to the greater stacking interaction. From solubility data, purine–amino acid interaction was concluded [15] to be determined mainly by hydrophobic and stacking forces. One may cautiously extrapolate our results to say that in nucleic acid–protein complexes purine rich regions of nucleic acid may be recognized by the hydrophobic regions of the protein. In the case of pyrimidines, we do not make any generalization because of the insufficiency of data. But one striking

Table 1
Association constants (K_a) for amino acid–nucleoside systems at 20°C (M^{-1})

	Amino acid								
Nucleoside	Ser	Thr	Gly	Val	Leu	Met	Lys	Phe	Trp
Guanosine	0.347	0.449	0.449[a]	0.581	0.777[a]	1.277[a]	1.268	2.486[a]	13.733[a]
Adenosine			0.039	0.213			0.550	3.115	8.634
Cytidine								1.087	2.612
Uridine								427.57[b]	0.601

[a] At 25°C
[b] Value obtained by assuming 1:2 (U–Phe) stoichiometry and expressed in M^{-2}

The K_a values were calculated from the least square fit slope and intercept of the mean of the two readings for each point of the solubility curves

observation is that the Phe–U complex does not conform to a 1:1 composition and it has a very high K_a value compared with Trp–U system and other Phe–nucleoside systems. The K_a value (M^{-2}) for Phe–U system can be compared with other K_a values (M^{-1}) by taking its square root by assuming a stepwise formation of a 2:1 complex with more or less the same affinity. Even this normalized value is higher than other K_a values. The possible implication of this finding to the understanding of the evolution of genetic code will be discussed separately.

It is unnecessary to mention that factors other than hydrophobic and stacking forces are also important in protein–nucleic acid interaction. First, we have neglected the effect of the phosphate group by choosing nucleosides instead of nucleotides. In fact, the binding of Gly and Lys to nucleotides was observed [17] to be predominately ionic in nature. NMR studies have shown that in mononucleotide–dipeptide systems, even the position of the phosphate group influences binding [21]. The mononucleotide–dipeptide interaction also depends on the chirality of the dipeptide [21]. Thus a wide variety of factors can contribute to the specificity of protein–nucleic acid interaction, even though from the present study we can see only the effects of hydrophobicity of the amino acid, nature of the base, and stacking. The highly co-operative nature of polymer–polymer interaction magnifies the weak but specific interaction at monomer–monomer level leading to a highly specific recognition.

The study of amino acid–nucleic acid interaction is a necessary first step for understanding the origin and evolution of genetic code. The stereochemical theory of amino acid–codon relationship proposed [22] assumes a direct association of amino acids with polynucleotides at the early stages of the evolution of the code when there were no activating enzymes. Such a direct interaction in the absence of any catalyst might be due to the selective interaction of amino acid with some nucleotide sequence, determined by the physico-chemical characteristics of the two components. The two possibilities that have been suggested are amino acid–codon interaction and amino acid–anticodon interaction [1,2]. The rudimentary preferential scheme observed in our data may be interpreted so as to support the above views. Thus the very high K_a value for the Phe–U complex compared to other Phe–nucleoside complexes and the Trp–U complex suggests a highly specific recognition of Phe by its present day codon UUU as the first step in evolution of the Phe–UUU codon relationship. Further uracil has been observed [14] to interact more strongly with Phe than with other hydrophobic amino acids. But the same argument does not seem to apply to Gly as it is one of the amino acids with the least affinity for G. However the Gly–G complex is ~10-times more stable than the Gly–A complex. Moreover under our experimental conditions the electrostatic interactions which are mainly responsible for stabilizing Gly–nucleotide complexes are diminished to a very large extent [17]. It was argued [4] that since Gly has no side-chain, the preferred codon for Gly would be the one with the maximum base stacking tendency and since GMP has this property to a greater extent that other mononucleotides, Gly should be coded as GGG. But it was suggested [18] that Gly may be associated with the anticodon base C, the least hydrophobic of all bases. In fact, NMR measurements on mononucleotide–dipeptide systems have shown that Gly prefers C to G [21]. Unfortunately, solubility data cannot be obtained for Gly–pyrimidine systems because of the high solubility of pyrimidines and hence at present we cannot draw any conclusion about Gly. In the case of Lys, ionic interactions are much more predominant and hence such relationships are not evident from our data. However, it may be possible to demonstrate such amino acid–codon relationships by suitably modifying the experimental system. Nevertheless the present results do indicate the potentiality of solubility method in establishing amino acid–codon relationships particularly in cases where non-ionic interactions are predominant.

Acknowledgement

We thank Professor Cyril Ponnamperuma, University of Maryland for helpful discussions during his stay here. We are also grateful to Mrs Priti Roy and Mr K. Srinivasan for their technical assistance. The financial assistance from the Indian National Science Academy is gratefully acknowledged.

References

[1] Crick, F. H. C. (1968) J. Mol. Biol. 38, 367–379.
[2] Orgel, L. E. (1968) J. Mol. Biol. 38, 381–393.
[3] Woese, C. R. (1968) Proc. Natl. Acad. Sci. USA 59, 110–117.
[4] Lacey, J. C., jr and Pruitt, K. M. (1969) Nature 223, 799–804.
[5] Dimicolli, J. L. and Helene, C. (1974) Biochemistry 13, 714–723.
[6] Eckstein, H., Schott, H. and Bayer, T (1976) Biochim. Biophys. Acta 432, 1–9.
[7] Rifkind, J. M. and Eichhorn, G. L. (1970) Biochemistry 9, 1753–1761.
[8] Herskovits, T. T. and Harrington, J. P. (1972) Biochemistry 11, 4800–4811.
[9] Herskovits, T. T. and Bowen, J. J. (1974) Biochemistry 13, 5474–5483.
[10] Nozaki, Y. and Tanford, C. (1970) J. Biol. Chem. 245, 1648–1652.
[11] Dooley, K. H. and Castellino, F. J (1972) Biochemistry 11, 1870–1874.
[12] Su, S. C. K. and Shafer, J. A. (1968) J. Am. Chem. Soc. 90, 3861–3864.
[13] Scruggs, R. L., Achter, E. K. and Ross, P. D. (1972) Biopolymers 11, 1961–1972.
[14] Chao, H. M., Hazel, B. and Black, S. (1976) J Biol. Chem. 251, 6924–6928.
[15] Akinrimisi, E. O. and Ts'o, P O. P. (1964) Biochemistry 3, 619–626.
[16] Sober, H. A. ed (1970) Handbook of Biochemistry, 2nd edn, CRC Press.
[17] Saxinger, C. and Ponnamperuma, C. (1971) J. Mol. Evol. 1, 63–73.
[18] Woese, C. R. (1973) Naturwissensch. 60, 447–459.
[19] Tanford, C. (1962) J. Am. Chem. Soc. 84, 4240–4247.
[20] Nozaki, Y. and Tanford, C (1963) J. Biol. Chem. 238, 4074–4081.
[21] Dasgupta, D. and Podder, S. K. (1978) Int. Symp. Biomolecular Structure, Conformation, Function and Evolution, Madras, India, abstr. p. 42.
[22] Woese, C. R (1967) The Genetic Code, Harper and Row.