Archs Sci. Genève	Vol. 51	Fasc. 2	pp. 147-152	Septembre 1998

POLARITY AT THE ONSET OF GENETIC CODING. III. PRENUCLEIC STEREOCHEMICAL INTERACTIONS BETWEEN AMINO ACIDS AND NUCLEOBASES DOUBLETS

BY

Gilbert TURIAN*

(Ms reçu le 28.7.1998, accepté le 20.8.1998)

ABSTRACT

Polarity at the onset of genetic coding. III. Prenucleic stereochemical interactions between amino acids and nucleobases doublets. - The nature and positioning of the weak bonding forces primarily mediating the stereochemical recognition between amino acids and their encoding nucleobases was plausibly imposed by their respective, intrinsic molecular configurations, thereby allowing a computation of the specific assignments of amino acids to doublets of nucleobases chosen as those of the vestigial anticodons and originally «frozen» by phosphoramidate bondings on polyphosphate chains to conserve the nascent genetic code.

Key-words: Prenucleic, Stereochemical, Amino acids, Nucleobases doublets.

INTRODUCTION

In organisms existing at the present time, the recognition of amino acid by a codon is effected indirectly through a molecule of tRNA, but recognition early in evolution may have occurred through direct amino acid-template interactions (WOESE *et al.*, 1966; CRICK, 1968, NIU *et al.*, 1987; SZATHMARY, 1993; LAZCANO & MILLER, 1996; ALBERTI 1997). Such molecular recognitions involve specific physicochemical affinities between amino acids and nucleobases which must have occurred ever since the first endo-exogenously formed amino acids and the secondarily appeared nucleobases randomly met in primitive water lagoons (see reviews in BRACK & RAULIN, 1991; POGLAZOV *et al.*, 1995; MAUREL, 1997). These recognitions between monomers have been experimentally simulated and tentatively detected by physicochemical parameters and sophisticated stereochemical criteria (WOESE *et al.*, 1966; DILLON, 1973; DANCHIN, 1979; DOUNCE, 1981; HENDRY *et al.*, 1981 (45 refer.); SUKHODOLETS, 1989; ALBERTI, 1997). They have further gained in plausibility with the parallel reports of specific molecular and statistical interactions by means of which a specific sequence in a double-helical DNA genome can be recognized by a protein (VON HIPPEL, 1979; PABO & SAUER, 1984; HARRISON & AGGARWAL, 1990).

VON HIPPEL's precursor remark that «these molecular recognition principles will also apply to protein-nucleic acid interactions other than those involved in control of

* Laboratoire de Bioénergétique et Microbiologie, Université de Genève, Pavillon des Isotopes, 20, Bd d'Yvoy, CH-1211 Genève 4, Switzerland.

transcription» was most seminal in our present transposition of these principles to the origin of the genetic code, but with a reversal of the reciprocal roles of protein and nucleic acid to a primary, «predatory» role of the nucleobases over their other monomeric partners, the amino acids (TURIAN, 1996-97). Such specific molecular interactions, mediated by weak bondings as discussed in JEFFREY & SAENGER (1994), would have involved the amino acids of primitive, randomly assembled peptides (cyclopeptides? TURIAN, 1995) and nucleobase doublets chosen from the first two letters of «modern» anticodon triplets because of their primordial proximality to amino acids (discussion in TURIAN, 1997).

MOLECULAR MODELLING

In 1984, SAENGER described protein-nucleic acid interactions as of the very greatest importance in life while pondering the still rather limited knowledge. He considered four kinds of potential interactions: (a) salt bridges (positively charged amino acids); (b) hydrogen bonding (hydrophilic amino acid side chains); (c) stacking interactions (aromatic amino acids); (d) hydrophobic interactions (non polar amino acid side chains). He further commented that «if amino acids and nucleosides are considered as individual units, charge-charge and hydrogen-bonding forces rather than hydrophobic and stacking interactions seem to dominate structure and be best suited for selective recognition». Hydrogen bonds can be defined as weak electrostatic attractions between one electronegative atom and a hydrogen anticovalently linked to a second electronegative atom, as exemplified by N^+ - H (donor)$\rightarrow O^-$ (acceptor). Stronger interactions are of the general hydrogen-bonded «salt-bridge» types such as $N+(H_2)H\rightarrow O^- = C$ bonds also formed by charged side-chains and main-chain termini. All these polar attractions would be complemented at the level of the non polar parts of the interacting molecules by the van der Waals hydrophobic attractions between uncharged atoms, mainly H ··· H, that are in close proximity (3.5 – 4 A°, see JEFFREY & SAENGER, 1994).

Hydrogen bonds are strongest when the bonded molecules are oriented to allow maximum electrostatic interaction. They are thus directional and capable of holding two hydrogen-bonded molecules or groups in a specific geometrical arrangement. Consequently, following PAULING's proposal (1960), we have considered that these oriented bondings could have been the main drivers of the stereospecific interactions in the loose cavities or «pockets» modelled by the base doublets thereby acting as specific receptors of the amino acids (Table I).

Complementarily to this intrinsic positioning of reactive atomic groups, steric hindrance also contributes to the restriction of interactions between amino acids and bases and leads to the exclusion of certain amino acids from cavities formed by the doublets of bases. This would lead to the selective exclusion of certain amino acids from the encoding process. Such a principle of exclusion is well emphasized in the hydrophobic series of amino acids (1st column in the Table I) in which the principle is more constraining for heavy leucine which could only fit the large cavity left between

TABLE I

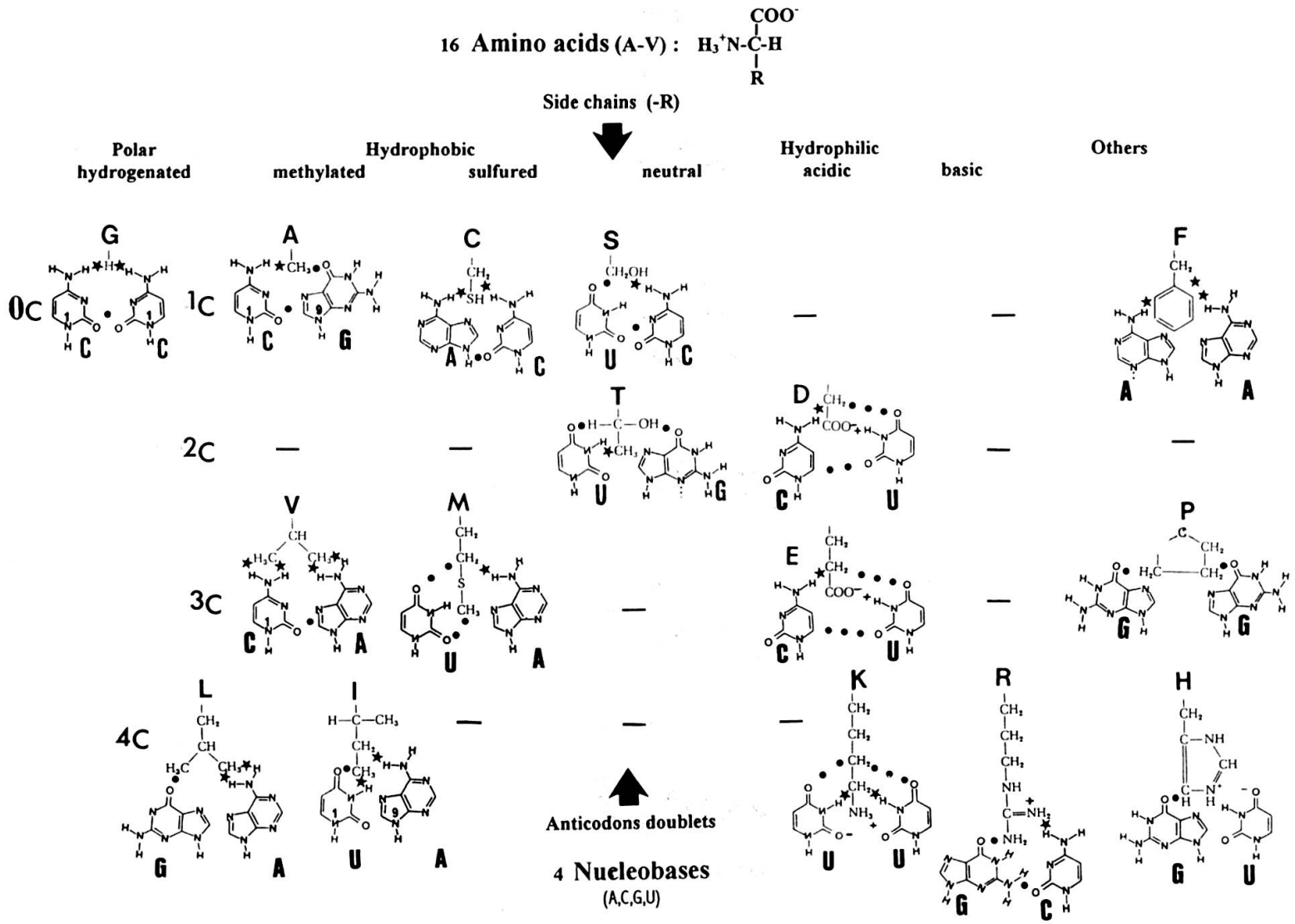

Chart of the weak bonding forces at work in the specific assignments of amino acids (one-letter symbols) to nucleobases by their differential side chains (R): ★ = Van der Waals bonds; ● = Hydrogen bonds directed from donor to acceptor atoms; +/– = Electrostatic attractive (↔) forces linking hydrophilic groups. No AU (UA = stop codon).

two purines (G-A) bases than for light glycine which is good choice for two tightly bonded, light pyrimidines (C-C), a cavity otherwise exclusive for the other amino acids starting with alanine.

DISCUSSION

We have assumed that the first, deterministically formed coding pairs for every amino acid have been selectively conserved for their spatio-structural contingencies over the whole evolutionary span. Their lettering could thus be known from their «modern» anticodon vestiges, complementary as base pairs to those of the mRNAs codons of the actual genetic code outstandingly deciphered by biochemists of the last 60ties. For example, glycine, coded by G + G (the 3rd letter or wooble base being either of the 4 ones), has C + C as «modern» anticodon, and the same two letters as original coding «predator» of glycine. However, none of the coding base couples originally formed would have subsisted over the more than 3 billion years of biochemical-biological evolution if they had not been someway «frozen» in their original doublet state, possibly by their phosphoramidic bondings on polyphosphates as recently proposed (Turian & Schoenenberger, 1997; Turian *et al.*, 1998) in an evolutionary step preceding the ribosylation of these prenucleic polybasephosphate chains to ribonucleic polynucleotides.

Polybasephosphates possibly stabilized as chains of Ca salts as «tapes» could thus well be a first concretization of Kornberg's (1995) predictions that «high energy and phosphate content of poly P, with its antiquity in prebiotic evolution, make it a plausible precursor to RNA, DNA, and proteins». In their questioning about «Polyphosphates as prebiotic reagents?», Keefe & Miller also in 1995 echoed these predictions by declaring that «it cannot be excluded that there are prebiotic processes requiring only small amounts of polyphosphates to take off and form self-replicating systems».

By recognizing the selective value of stereochemical affinities between primordial amino acids and nucleobases, we are therefore led to priviledge for such primal genetic coding a deterministic theory of «frozen interactions» rather than that of the stochastic «frozen accident» proposed by Crick (1958, 1968) but also recently challenged by Alberti (1997).

ACKNOWLEDGEMENTS

We thank Ms. Ariane Fehr and Sandrine Girard-Turian for their secretarial assistance.

RÉSUMÉ

POLARITÉ À L'ORIGINE DU CODAGE GÉNÉTIQUE. III. INTERACTIONS STÉRÉOCHIMIQUES PRÉNUCLÉIQUES ENTRE ACIDES AMINÉS ET DOUBLETS DE NUCLÉOBASES

La nature et le positionnement des forces de liaison faibles primairement médiatrices des reconnaissances stéréochimiques entre acides aminés et leurs nucléobases

d'encodage ont été, de façon plausible, imposées par leurs configurations moléculaires intrinsèques respectives. Ce faisant, elles ont permis une supputation des attributions spécifiques d'amino acides aux doublets de nucléobases choisis comme ceux de vestiges d'anticodons, originellement «gelés» par liaisons phosphoramidiques sur des chaînes de polyphosphates pour conserver le code génétique naissant.

Mots-clés: Prénucléique, stéréochimique, amino-acides, doublets nucléobases.

REFERENCES

ALBERTI, S. 1997. The origin of the genetic code and protein synthesis. *J. Mol. Evol.* 45: 352-358.

BRACK, A. & F. RAULIN. 1991. L'Evolution Chimique et les Origines de la Vie. Masson, Paris.

CRICK, F.H.C. 1958. On proteins synthesis. *Sympos. Soc. Exper. Biol.* 12: 138-163.

CRICK, F.H.C. 1968. Origin of the genetic code. *J. Mol. Biol.* 38: 367-379.

DANCHIN, A. 1979. L'origine du code génétique: la nécessité l'emporte-t-elle sur le hasard? *La Recherche* 97: 186-189.

DILLON, L.S. 1973. The origins of the genetic code. *Bot. Rev.* 39: 302-345.

DOUNCE, A.L. 1981. Origin of life-proposed mechanisms for primeval polynucleotide and peptide chain synthesis. *J. Theor. Biol.* 90: 63-79.

HARRISON, S.C. & A.K. AGGARWAL. 1990. DNA recognition by proteins with the helix-turn-helix. *Annu. Rev. Biochem.* 59: 933-969.

HENDRY, L.B., E.D. BRANSOME, JR., M.S. HUTSON & L.K. CAMPBELL. 1981. First approximation of a stereochemical rationale for the genetic code based on the topography and physicochemical properties of «cavities» constructed from models of DNA. *Proc. Natl. Acad. Sci. U.S.A.* 78: 7440-7444.

JEFFREY, G.A. & W. SAENGER. 1994. Hydrogen Bonding in Biological Structures. Springer-Verlag, Berlin, Heidelberg, 569 pp.

KEEFE, A.D. & S.L. MILLER. 1995. Are polyphosphates or phosphate esters prebiotic reagents? *J. Mol. Evol.* 41: 693-702.

KORNBERG, A. 1995. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *J. Bacteriol.* 177: 491-496.

LAZCANO, A & S.T. MILLER. 1996. The origin and early evolution of life: prebiotic chemistry, the pre-RNA worlds, and time. *Cell* 85: 793-798.

MAUREL, M.-C. 1997. La Naissance de la Vie. De l'Evolution Prébiotique à l'Evolution Biologique. Diderot Edit. Paris, 135 pp.

NIU, C.-H., K.-H. HAN, H.J.C. YEH, & S. BLACK. 1987. Hydrogen bonding between cytosine and peptides of threonine or serine: is it relevant to the origin of the genetic code? *Biochem. Biophys. Res. Comm.* 148: 456-460.

PABO, C.O. & R.T. SAUER. 1984. Protein – DNA recognitions. *Ann. Rev. Biochem.* 53: 293-321.

PAULING, L. 1960. The nature of the forces operating in the process of the duplication of molecules in living organisms. pp. 132-138. In: *Aspects of the Origin of Life.* (M. Florkin, ed.).

POGLAZOV, B.F., KURGANOV, B.I., KRITSKY, MS. & K.L. GLADILIN. 1995. Evolutionary Biochemistry and Related Arrays of Physicochemical Biology. Bach Institute of Biochemistry and ANKO. Moscow. 618 pp.

SAENGER, W. 1984. Principles of Nucleic Acid Structure. Springer Adv. Texts in Chemistry. Series Edit.: C.R. CANTOR. Springer-Verlag. New York, Berlin, Heidelberg, Tokyo.

SUKHODOLETS, V.V. 1989. The genetic code as a clue to understanding of molecular evolution. *J. Theor. Biol.* 141: 379-389.

SZATHMARY, E. 1993. Coding coenzyme handles: a hypothesis for the origin of the genetic code. *Proc. Natl. Acad. Sci. U.S.A.* 90: 9916-9920.

TURIAN, G. 1995. New trends in Polarity. III. Dipolar hydrogen bondings as homotemplate forces for pregenetical evolution. *Archs Sci. Genève* 48: 173-182.

TURIAN, G. 1996. Polarity at onset of genetic coding. I. Bipolar bondings in the two-step takeover of peptide templates by prenucleic-ribonucleic acids. *Archs Sci. Genève* 49: 213-227.

TURIAN, G. 1997. Ibid. II. Primary recognition of amino acids by base doublets of prenucleic sugarless polymers secondarily taken over by ribonucleic acids. *Archs Sci. Genève* 50: 95-104.

TURIAN, G. & I. SCHOENENBERGER-SOLA. 1997. Spectral evidence for phosphoramide bondings between nucleobases and tripolyphosphate possibly generative of prenucleic polybasephosphate chains. *Archs Sci. Genève* 50: 145-152.

TURIAN, G., E. RIVARA-MINTEN & A. CATTANEO. 1998. Similar ^{31}P-NMR signals emitted by imidazole and the nucleobases presumably phosphoramido-bonded to tri(meta)phosphates. *Archs Sci. Genève* 51: 187-193.=

VON HIPPEL, P.H. 1979. On the molecular basis of the specificity of interactions of transcriptional proteins with genome DNA. In *Biological Regulation and Development.* Vol l: Gene Expression. Edit. R.F. GOLDBERGER. Plenum Press, New York, London.

WOESE, C.R., D.H. DUGRE, W.C. SAXINGER & S.A. DUGRE. 1966. The molecular basis for the genetic code. *Proc. Natl. Acad. Sci. U.S.A.* 55: 966-974.