

Ο

# ΓΕΝΕΤΙΚΟΣ ΚΩΔΙΚΑΣ

Η αποκρυπτογράφηση του DNA αποκάλυψε πώς είμαστε κάτοχοι μίας γλώσσας πολύ αρχαιότερης από τα ιερογλυφικά, μίας γλώσσας παλιάς όσο και η ζωή, μίας γλώσσας που είναι ή πιά ζωντανή από όλες.

George και Muriel Beadle

## Πρωτεΐνες: Η βιοχημική ταυτότητα των οργανισμών.

Ό κάθε οργανισμός χαρακτηρίζεται από τό είδος και τό πλήθος των πρωτεϊνών του. Έτσι, ένα βακτήριο (ή *Esherichia coli* που χρησιμοποιείται για πειραματόζωο) έχει 2-3 χιλιάδες πρωτεΐνες, ενώ στον άνθρωπο ο αριθμός των πρωτεϊνών είναι στην τάξη μεγέθους του ενός εκατομμυρίου. Όλα τά άτομα που ανήκουν σέ ένα βιολογικό είδος έχουν τίς ίδιες ακριβώς πρωτεΐνες, ενώ δύο διαφορετικά είδη μπορούν νά έχουν π.χ. 99% κοινές πρωτεΐνες και νά διαφέρουν μόνο κατά 1%. Αυτό τό 1% των διαφορετικών πρωτεϊνών είναι υπεύθυνο για τίς ανατομικές, φυσιολογικές και άλλες διαφορές τους.



Έχει αποδειχτεί πώς οι πρωτεΐνες του χιμαπατζή είναι κατά 99% όμοιες μέ τίς πρωτεΐνες του ανθρώπου. Όμως, καθώς ό κάθε οργανισμός έχει τή δική του βιοχημική-πρωτεϊνική ταυτότητα, οι δύο οργανισμοί αποτελούν δύο διαφορετικά είδη.

Μπορούμε λοιπόν νά πούμε πώς ό κάθε οργανισμός, ό άνθρωπος, ό θάτραχος, ή μέλισσα, έχει μιά συγκεκριμένη βιοχημική ταυτότητα που τον κάνει νά ξεχωρίζει από τά άλλα είδη. Αυτό φυσικά σημαίνει πώς σέ κάθε είδος τά νέα άτομα που αναπαράγονται θά πρέπει μέ έναν ειδικό κληρονομικό μηχανισμό νά παίρνουν από τούς γονείς τους τίς ίδιες μ' αυτούς πρωτεΐνες. Καί νά πού φτάσαμε στό ζήτημα μας: Αυτή ή απaráλλακτη αναπαραγωγή των πρωτεϊνών γίνεται μέ ένα «κρυπτογραφικό» κώδικα, μέ τό όποιο και θά

ασχοληθούμε. Άς πάρουμε όμως τά πράγματα μέ τή σειρά.

## Πρωτεΐνες και άμινοξέα.

Οί πρωτεΐνες είναι μιά έννοια που συναντούμε συχνά και στην καθημερινή γλώσσα. Λέμε, για παράδειγμα, πώς τό κρέας και τό αυγό είναι τροφές πλούσιες

σέ πρωτεΐνες. Τί είναι όμως ακριβώς αυτές και ποιά είναι τά συστατικά τους;

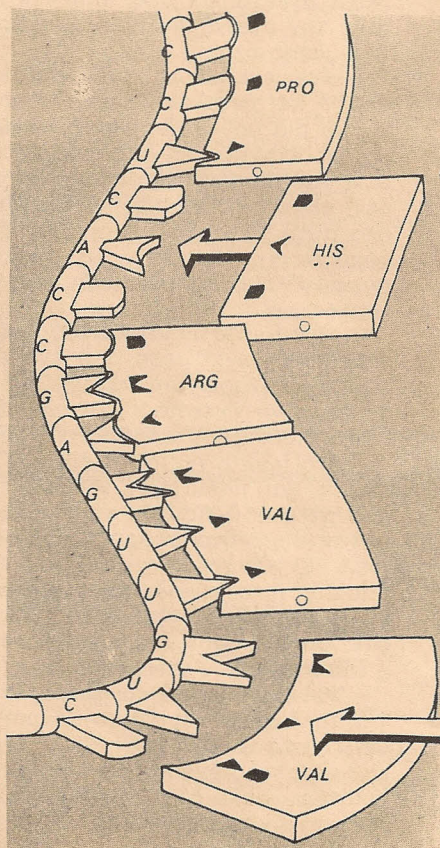
Οί πρωτεΐνες είναι χημικές ενώσεις μέ μεγάλο μοριακό βάρος, που κυμαίνεται από μερικές χιλιάδες ως μερικά εκατομμύρια. Τά στοιχεία που τίς αποτελούν είναι άνθρακας (C), ύδρογόνο (H), όξυγόνο (O) και άζωτο (N) – χωρίς όμως νά αποκλείονται και μερικά άλλα στοιχεία όπως τό θείο (S), ό φωσφόρος (P) κλπ. Αυτά τά στοιχεία είναι ένωμένα μεταξύ τους έτσι που νά αποτελούν τούς λεγόμενους «δομικούς λίθους» των πρωτεϊνών. Αυτοί είναι τά **άμινοξέα**, και μπορούμε νά πούμε πώς αυτά αποτελούν τίς διάφορες πρωτεΐνες. Όλες οι πρωτεΐνες αποτελούνται από τά ίδια 20 άμινοξέα σέ διάφορες αναλογίες. Μιά πρωτεΐνη μέσου μεγέθους μπορεί νά έχει 300 περίπου άμινοξέα (δηλ. τά 20 άμινοξέα επαναλαμβανόμενα πολλές φορές μέσα στό μόριο). Για νά ξέρουμε πώς ακριβώς είναι μιά πρωτεΐνη, πρέπει νά μάθουμε ποιά άμινοξέα τήν αποτελούν και μέ ποιά σειρά. Μόλις τό 1953 έγινε γνωστή για πρώτη φορά ή «άμινοξική άλληλουχία» μίας πρωτεΐνης, τής *ινσουλίνης*, που έχει μόνο 51 άμινοξέα.

Άς δούμε ποιά απ' αυτά βρίσκονται στην αρχή και ποιά στό τέλος του μορίου τής *ινσουλίνης*, για νά εξοικειωθούμε έτσι λίγο και μέ τίς ονομασίες τους: *γλυκίνη*, *ισολευκίνη*, *βαλίνη*, *θρεονίνη*, *προλίνη*, *λυσίνη*, *αλανίνη*. Μέ ποιόν όμως τρόπο μεταφέρεται ή έντολή ή, καλύτερα, τό μήνυμα για νά σχηματιστούν τά διάφορα άμινοξέα που αποτελούν τήν πρωτεΐνη;

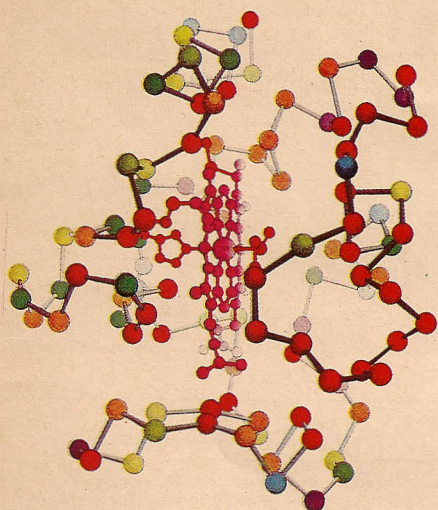
## Ένα καλογραμμένο μήνυμα.

Άν θέλουμε νά θυμηθούμε όλα τά ψώνια που θά κάνουμε, τά γράφουμε σέ ένα χαρτί. Κάτι παρόμοιο συμβαίνει και στην περίπτωση του σχηματισμού των πρωτεϊνών: Υπάρχει ένας «κατάλογος» μέ «λέξεις» που όρίζουν ποιά άμινοξέα θά σχηματισθούν. Ό κατάλογος αυτός ονομάζεται RNA (άρ εν έι, **ριβοζονουκλεϊνικό όξύ**) και κατασκευάζεται μέσα στον πυρήνα του κυττάρου, απ' όπου θά βγει για νά μεταφέρει τίς λέξεις-έντολές που περιέχει.

Η εκτέλεση αυτών των εντολών ή είναι ό σχηματισμός των πρωτεϊνών.



Τό μόριο του RNA μέ τίς θάσεις, των όποιων οι τριαδικές άκολουθίες κωδικοποιούν τά διάφορα άμινοξέα.



Έτσι αναπαριστούνται οι πρωτεΐνες με τα διάφορα αμινοξέα που τις αποτελούν.

## Η διπλή ἄρθρωση του κώδικα.

“Ως τώρα μιλήσαμε για τις «λέξεις» και τα «γράμματα» του κώδικα δανειζόμενοι ὄρους που χρησιμοποιούνται στη γραμματική. Πρόκειται μόνο για «διδασκτικές» ἀπλουστεύσεις; Ποιά μπορεί να είναι ή βαθύτερη δομική σχέση του κώδικα με τις γλώσσες των ανθρώπων; “Ας δοῦμε.

“Αν αντί για 4 μόνο βάσεις ὑπῆρχαν 20 βάσεις, τότε κάθε μία ἀπό αυτές θά ἀντιστοιχοῦσε σέ ἕνα ἀμινοξύ και δέν θά χρειαζόνταν οἱ τριαδικοί συνδυασμοί. Στήν περίπτωση αὐτή ὁ κώδικας θά ἦταν πολύ λίγο «ἀποδοτικός» ἐνῶ ἡ κωδικοποίηση θά γινόταν «φύσει» και ὄχι χάρη στό συνδυασμό και τή θέση τῶν βάσεων. “Αν, ἐξάλλου, ὑπῆρχε μόνο μία βάση Β, θά ἦταν δυνατόν νά ὑπάρχει κωδικοποίηση με ἀντιστοιχία ἀμινοξέων και ἐπαναλήψεων τῆς μοναδικῆς αὐτῆς βάσης, δηλαδή Β θά σήμαινε π.χ. *ἀλανίνη*, ΒΒ θά σήμαινε *λευκίνη*, ΒΒΒ *σερίνη* κλπ. Αὐτή ἡ μορφή τοῦ κώδικα φαίνεται νά εἶναι μάλλον ἀσύμμετρο.

“Ὅμως ὁ γενετικός κώδικας στήν πραγματικότητα 1. Εἶναι ἀποδοτικός. 2. Ἐλέγχει τό μήνυμα με τό συνδυασμό και τή θέση τῶν βάσεων 3. Εἶναι συμμετρικός. Αὐτές οἱ ιδιότητες τόν κάνουν οικονομικό, κι αὐτό ὀφείλεται στή **διπλή ἄρ-**

θρωσή του. Μ' αὐτό ἐννοοῦμε πῶς ὑπάρχουν α. μονάδες με ὀρισμένη σημασία, δηλ οἱ τριάδες και β. μικρότερες μονάδες, δηλ. οἱ βάσεις, πού συνιστοῦν τίς πρώτες.

Και στίς γλώσσες τῶν ἀνθρώπων συμβαίνουν παρόμοια πράγματα: Οἱ λέξεις ἢ, σωστότερα, τὰ *μονήματα* εἶναι οἱ σημασιολογικές μονάδες (*πρώτη ἄρθρωση*) και τὰ γράμματα ἢ, σωστότερα, τὰ *φωνήματα* εἶναι τὰ συστατικά τους (*δεύτερη ἄρθρωση*).

“Ὅπως διαπιστώνει ὁ Jacobson: «Μποροῦμε νά ποῦμε πῶς ἀνάμεσα σέ ὅλα τὰ συστήματα μεταφορᾶς πληροφοριῶν, ὁ γενετικός κώδικας και ὁ λεκτικός κώδικας εἶναι οἱ μόνοι πού βασίζονται στή χρήση διάκριτων συνιστωσῶν, οἱ ὁποῖες, μόνες τους στεροῦνται σύμφυτου νοήματος ἀλλά χρησιμεύουν για νά συνθέσουν τίς ἐλάχιστες νοηματικές μονάδες, δηλαδή ὀντότητες προικισμένες με τό δικό τους ἐσωτερικό νόημα στόν δεδομένο κώδικα».

### ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Ρένας Λεκανίδου (ἐπιμέλεια), Εἰσαγωγή στή σύγχρονη βιολογία, Ἀθήνα 1976. (Ἐνότητες 1η και 2η).
2. Ν. Η. Μαδιά, Γενική βιολογία, Ἀθήνα 1967.
3. Ἴ. Γ. Οικονομίδου, Μαθήματα γενικῆς βιολογίας, Ἀθήνα (ΟΕΔΒ) 1976.
4. Jacques Monod, Le hasard et la nécessité, Editions du Seuil, 1970.
5. R. Jacobson, Main Trends in the Science of Language, London, 1973.

I.Δ.Ρ.

“Ὅσο για τίς λέξεις τοῦ καταλόγου, αὐτές εἶναι ὅλες τριγράμματα. Και νά τί ἐννοοῦμε μ' αὐτό: Πάνω στό μακρόστενο μόριο τοῦ RNA, βρίσκονται σέ γραμμική διάταξη, τό ἕνα μετά τό ἄλλο, τεσσάρων εἰδῶν μόρια. Εἶναι τὰ μόρια τῶν βάσεων πού ἔχουν τὰ ὀνόματα *οὐρακίλη* (U), *ἀδε-νίνη* (A), *κυτοσίνη* (C) και *γουανίνη* (G). “Ἔτσι, για παράδειγμα, ἕνα RNA μπορεί νά ἔχε, με τή σειρά τίς βάσεις U.C.C.A.C.G κλπ. Αὐτές οἱ βάσεις τρεῖς, τρεῖς, δηλ UCC, ACG κλπ. θά μεταφέρουν μαζί με ἄλλες τό μήνυμα σχηματισμοῦ μίας πρωτεΐνης πού θά περιέχει και δύο ὀρισμένα ἀμινοξέα πού ἀντιστοιχοῦν στίς τριγράμματα ἀκολουθίες (τριάδες βάσεων) UCC και ACG. Κι αὐτά εἶναι ἡ *σερίνη* και ἡ *θρεονί-νη*.

## Ένας ἔξυπνος κώδικας

“Ὅστε, λοιπόν, κάθε ἀμινοξύ ἀντιστοιχεῖ σέ μιά τριάδα βάσεων. Εἶναι ὅμως ἀρκετοί αὐτοί οἱ τριαδικοί συνδυασμοί τῶν τεσσάρων βάσεων για νά καλύψουν τὰ εἰκοσι ἀμινοξέα;

“Ἡ ἀπάντηση εἶναι καταφατική. Και νά πῶς σκεφτόμαστε: “Αν οἱ «λέξεις» ἦταν διγράμματα τότε οἱ συνδυασμοί τους θά ἦταν  $4 \times 4 = 16$ , ἐπειδή κάθε μία ἀπό τίς βάσεις U,C,A,G θά συνδυαζόνταν με τόν ἑαυτό της και τίς ὑπόλοιπες τρεῖς. Τώρα πού οἱ «λέξεις» εἶναι τριγράμματα οἱ συνδυασμοί εἶναι  $16 \times 4 = 64$ , ἀφοῦ κάθε ἕνας ἀπό τούς 16 διγράμματος συνδυασμούς συνδυάζεται με τίς 4 βάσεις. Συνεπῶς, οἱ 64 συνδυασμοί UUU, UUG, ..., GGG ἐπαρκοῦν για τὰ 20 ἀμινοξέα και μάλιστα περισσεύουν. Γι' αὐτό, τὰ διάφορα ἀμινοξέα κωδικοποιοῦνται ἀπό πολλούς συνδυασμούς βάσεων, ὅπως για παράδειγμα οἱ συνδυασμοί CUU, CUC, CUA και CUG πού κωδικοποιοῦν τό ἀμινοξύ *λευκίνη*. “Αναφέρουμε ἀκόμα πῶς οἱ συνδυασμοί UAA, UAG και UGA χρησιμεύουν ὡς «σημεῖα στίξης» και ἀντιστοιχοῦν στό μήνυμα «ΛΗΞΗ». Μ' αὐτά ἡ σύνθεση τῆς συγκεκριμένης πρωτεΐνης σταματᾷ.

Οἱ ὑπόλοιπες λεπτομέρειες τοῦ κώδικα, πού τόν ἀποκάλυψε ἡ ἐρευνητική ὁμάδα τοῦ Nirenberg (1961), φαίνονται στόν πίνακα. Για νά βροῦμε σέ τί ἀντιστοιχεῖ π.χ. ἡ τριάδα GCA, βρίσκουμε ἀπό ἀριστερά τή σειρά τοῦ G, ἀπό ἐπάνω τή στήλη τοῦ C και ἀπό δεξιά τή σειρά τοῦ A- πρόκει-ται για τήν *ἀλανίνη*

### Πίνακας: Ο Γενετικός κώδικας

		Δεύτερη βάση									
		U		C		A		G			
U	UUU	Φαινυλα- λανίνη	UCU	} Σερίνη	UAU	Τυροσίνη	UGU	} Κυστεΐνη	U	Τρίτη βάση	
	UUC		UCC		UAC		UGC		C		
	UUA	Λευκίνη	UCA	UAA	ΛΗΞΗ	UGA	ΛΗΞΗ	A			
	UUG		UCG	UAG	ΛΗΞΗ	UGG	Τρυπτοφάνη	G			
C	CUU	} Λευκίνη	CCU	} Προλίνη	CAU	} Ιστιδίνη	CGU	} Ἀργινίνη	U		
	CUC		CCC		CAC		CGC		C		
	CUA		CCA		CAA	CGA	A				
	CUG		CCG		CAG	CGG	G				
A	AUU	} Ισολευκίνη	ACU	} Θρεονίνη	AAU	} Ἀσπαρα- γινίνη	AGU	} Σερίνη	U		
	AUC		ACC		AAC		AGC		C		
	AUA	ACA	AAA		AGA	A					
	AUG	ACG	AAG		AGG	G					
G	GUU	} Βαλίνη	GCU	} Ἀλανίνη	GAU	} Ἀσπαραγι- νικό ὀξύ	GGU	} Γλυκόκολλα	U		
	GUC		GCC		GAC		GGC		C		
	GUA		GCA		GAA	GGA	A				
	GUG		GCG		GAG	GGG	G				