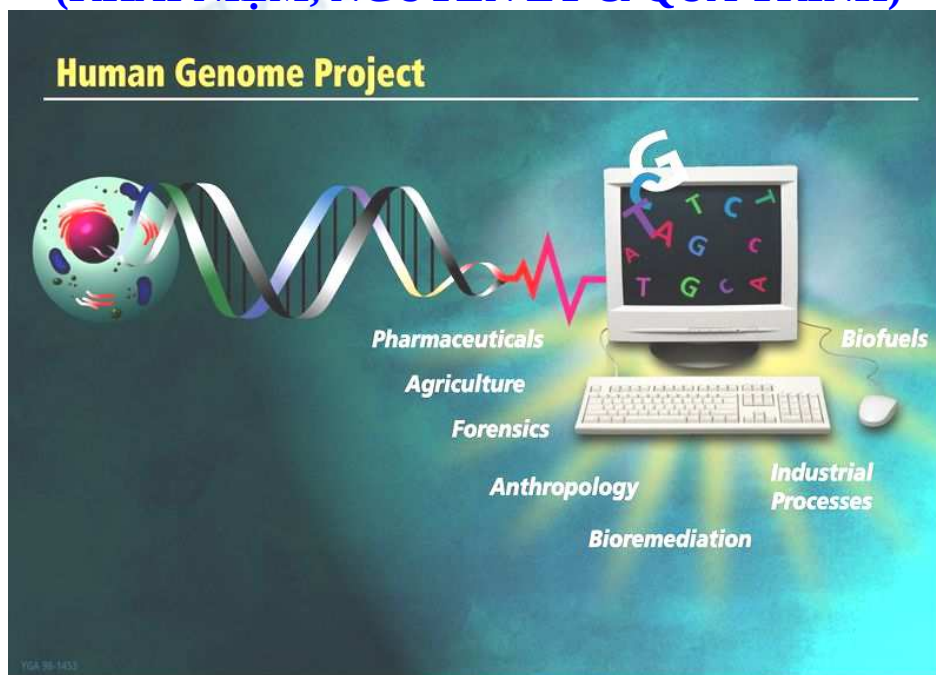


ĐẠI HỌC HUẾ
TRUNG TÂM ĐÀO TẠO TỪ XA

HOÀNG TRỌNG PHÁN

SINH HỌC PHÂN TỬ
(KHÁI NIỆM, NGUYÊN LÝ & QUÁ TRÌNH)



Huế - 2012

Lời nói đầu

Để góp phần đổi mới nội dung giáo trình Sinh học phân tử của Trung tâm Đào tạo Từ xa – Đại học Huế theo hướng cập nhật kiến thức cũng như phương pháp dạy - học phù hợp với đối tượng đặc thù, chúng tôi đã tham cứu nhiều tài liệu và cố gắng biên soạn giáo trình trên tinh thần ấy.

Nội dung giáo trình gồm tám chương bao quát các kiến thức cơ bản của Sinh học phân tử mà học viên và sinh viên của Trung tâm và các trường Đại học cần nắm vững để có thể vận dụng tốt vào trong công tác nghiên cứu và giảng dạy của mình.

Chương 1: Cấu trúc và chức năng của các đại phân tử sinh học

Chương 2: Tổ chức bộ gen các sinh vật

Chương 3: Cấu trúc và chức năng của gen

Chương 4: Tái bản của các bộ gen

Chương 5: Phiên mã và dịch mã

Chương 6: Điều hòa sự biểu hiện của gen

Chương 7: Các biến đổi của bộ gen

Chương 8: Các phương pháp sinh học phân tử và công nghệ ADN tái tổ hợp

Mở đầu mỗi chương có phần giới thiệu và mục tiêu giúp người học xác định các chủ đề chính cần tìm hiểu. Sau mỗi chương có phần Tóm tắt nhằm giúp người học nắm nội dung khái quát của chương. Cuối cùng là phần Câu hỏi và Bài tập, với 15-25 câu mỗi chương, yêu cầu người học tập vận dụng hiểu biết của mình vào giải quyết chúng trước khi sang chương mới. Đặc biệt, trong khi biên soạn chúng tôi có đưa thêm phần Hướng dẫn Trả lời Câu hỏi và Bài tập cuối mỗi chương cùng với một số vấn đề liên quan thiết yếu khác vào phần Phụ lục (đặt ở cuối sách) nhằm giúp người học tra cứu, tham khảo cách học và giải quyết vấn đề khi cần.

Hy vọng rằng giáo trình này sẽ đáp ứng được nhu cầu học tập của học viên và sinh viên về môn học vốn dĩ rất mới và rất khó này. Tuy nhiên, vì khuôn khổ có hạn nên một số chủ đề không thể đề cập sâu hơn trong sách này. Hơn nữa, với khả năng có hạn, chắc chắn sách không thể tránh khỏi các sai sót trong khi biên soạn. Chúng tôi rất mong nhận được sự phê bình và chỉ bảo của quý đồng nghiệp và bạn đọc để giáo trình được hoàn chỉnh hơn trong lần in sau.

Huế, ngày 20 tháng 2 năm 2012

Tác giả

HOÀNG TRỌNG PHÁN

Chương 1

CẤU TRÚC VÀ CHỨC NĂNG CỦA CÁC ĐẠI PHÂN TỬ SINH HỌC

Đối với các nghiên cứu sinh học cơ sở, bốn đại phân tử quan trọng phải kể đến là các axit nucleic, protein, polysaccharide và lipid. Tuy nhiên, trên quan điểm của sinh học phân tử, protein và các axit nucleic là hai loại hợp chất quan trọng nhất mà chủ yếu là ADN và các thành phần của chúng. Việc nghiên cứu cấu trúc và chức năng của các axit nucleic thực sự bắt đầu từ giữa thế kỷ XX. Vào năm 1944, O.T. Avery, MacLeod và McCarty lần đầu tiên chứng minh ADN là vật chất mang thông tin di truyền. Kể đó, sự khám phá ra cấu trúc phân tử ADN bởi James Watson và Francis Crick năm 1953 cùng với những hệ quả của nó đã là một trong những sự kiện khoa học nổi bật nhất của thế kỷ XX, đặt nền tảng cho sự ra đời và phát triển của di truyền học và sinh học phân tử.

Trong chương này, chúng ta lần lượt tìm hiểu các vấn đề sau: (i) Cấu trúc và chức năng của các axit nucleic; (ii) Cấu trúc và chức năng của protein; (iii) Cấu trúc và chức năng của các polysaccharide và lipid; và (iv) Các liên kết hóa học cơ bản trong các hệ thống sống.

1. Cấu trúc và chức năng của các Axit Nucleic

1.1. Đại cương về các axit nucleic

Ngày nay chúng ta đều biết rằng vật chất di truyền hay bộ gen của các sinh vật trên trái đất là các axit nucleic mà hầu hết là *acid deoxyribonucleic* (ADN) và ở một số ít virus là *acid ribonucleic* (ARN). Điều này đã được chứng minh qua các thí nghiệm kinh điển nổi tiếng, đó là: (i) Thí nghiệm biến nạp ở vi khuẩn được thực hiện đầu tiên bởi Griffith (1928) và sau đó là nhóm nghiên cứu của Avery và cộng sự (1944); (ii) Thí nghiệm của Hershey và Chase ở thể thực khuẩn T2; và (iii) Thí nghiệm của Conrat và Singer ở virus đốm thuốc lá (1956).

Các axit nucleic là những đại phân tử sinh học có trọng lượng phân tử lớn với thành phần gồm các nguyên tố C, H, O, N và P. Chúng được cấu thành từ các *đơn phân* (monomer) - các *nucleotide*; các đơn phân này nối với nhau bằng các liên kết phosphodiester tạo thành cấu trúc *đa phân* (polymer) gọi là các chuỗi, mạch hay sợi *polynucleotide* - cấu trúc sơ cấp của các phân tử ADN và ARN.

Vật chất di truyền có các đặc tính thiết yếu sau:

(1) Đặc tính *thông tin sinh học*: Nó chứa đựng toàn bộ thông tin di truyền cần thiết cho việc xác định cấu trúc của các protein đặc thù của mỗi loài (các gen cấu trúc) và điều khiển các hoạt động sinh trưởng, phân chia và biệt hoá tế bào.

(2) Đặc tính *tái bản*: Đó là khả năng tự sao chép chính xác, đảm bảo thông tin di truyền của thế hệ sau giống với thế hệ trước.

(3) Đặc tính *hoạt động của các gen*: Các gen trong bộ gen có khả năng tổng hợp ra các sản phẩm là những phân tử tham gia vào mọi động sống căn bản của tế bào. Đó là các quá trình phiên mã, dịch mã và điều hòa hoạt động của các gen.

(4) Đặc tính *biến đổi*: Đó là khả năng bị biến đổi của các bộ gen từ các quá trình khác nhau

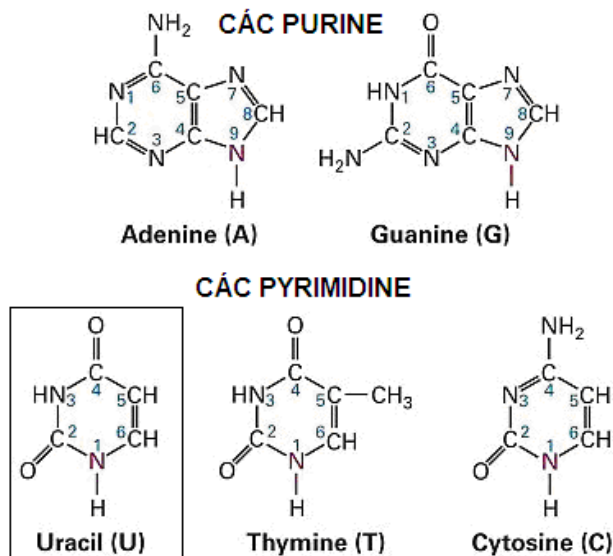
như đột biến, tái tổ hợp, các yếu tố di truyền vận động. Chính sự biến đổi này tạo ra các nguồn biến dị di truyền đa dạng và phong phú cho quá trình chọn lọc và tiến hoá của sinh giới kể từ khi sự sống bắt đầu hình thành trên trái đất cách đây chừng 3 tỷ rưỡi năm.

1.2. Cấu trúc của các nucleotide

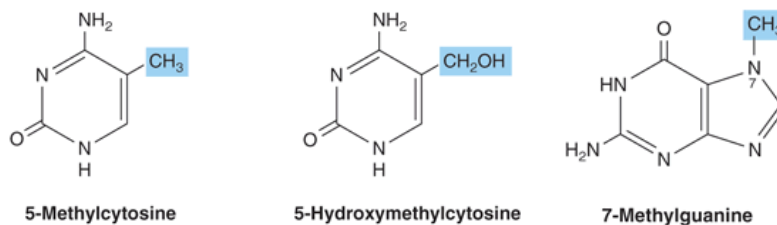
Đơn vị cấu trúc cơ sở của các axit nucleic là các nucleotide. Mỗi nucleotide gồm có: 1 bazơ nitơ, 1 đường pentose, và 1 axit phosphoric.

Bazơ nitơ - thành phần đặc trưng của các nucleotide - là các hợp chất *purine* và *pyrimidine* dị vòng chứa nitơ có tính kiềm. ADN chứa bốn loại bazơ chính là *adenine* (A), *guanine* (G), *thymine* (T) và *cytosine* (C); trong ARN cũng chứa 4 loại như thế, chỉ khác là *uracil* thay cho *thymine* (Hình 1.1). Ngoài ra, trong ADN còn có một số các bazơ bị biến đổi chủ yếu do sự *methyl* hoá ở các vị trí khác nhau, ví dụ: 5-methylcytosine, 5-hydroxymethylcytosine, 7-methylguanine v.v. (Hình 1.2).

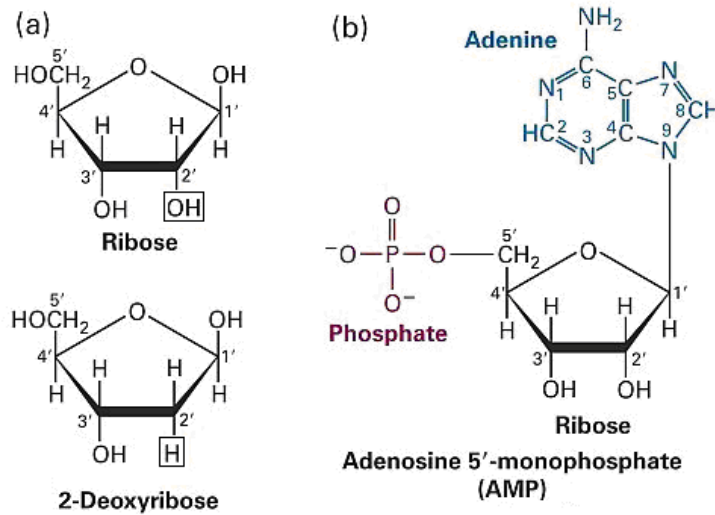
Đường *pentose* của ARN là *D-ribose* và của ADN là *2'-deoxy-D-ribose* (ký hiệu D chỉ dạng đường quay phải trước ánh sáng phân cực để phân biệt với dạng L quay trái không có trong thành phần của các axit nucleic tự nhiên). Các phân tử đường này đều có cấu trúc vòng. Vì các nguyên tử carbon ở đây xếp liên tục nên được đánh số thứ tự có dấu phẩy trên đầu, ví dụ C1', C2' cho đến C5' (Hình 1.3).



Hình 1.1. Cấu trúc 5 loại bazơ có mặt trong ADN và ARN.



Hình 1.2. Một số dạng bazơ bị biến đổi chủ yếu do sự methyl hoá.



Hình 1.3. (a) Cấu trúc của các phân tử đường ribose và deoxyribose.

(b) Cấu trúc của ribonucleotide Adenine trong ARN.

Hai phân tử đường này khác nhau ở C2'; trong ribose đó là nhóm hydroxyl và trong deoxyribose là một nguyên tử hydro. Từ các gốc đường khác nhau này tạo ra hai loại nucleotide - *ribonucleotide* và *deoxyribonucleotide* - cấu tạo nên hai loại axit nucleic khác nhau là ARN và ADN. Cần để ý rằng, trong các phân tử đường này có ba vị trí quan trọng có chứa nhóm hydroxyl (-OH) tự do, đó là: (i) nhóm -OH ở vị trí C1' có khả năng hình thành liên kết N-glycosid với gốc -NH của các bazơ để tạo thành các nucleoside; (ii) nhóm -OH ở vị trí C5' có khả năng hình thành liên kết ester với nhóm phosphate để tạo ra các nucleotide; và (iii) nhóm -OH ở vị trí C3' có khả năng hình thành liên kết phosphodiester với nhóm phosphate của một nucleotide khác để tạo chuỗi polynucleotide. Như vậy, *tính phân cực* trong gốc đường mà từ đó quyết định tính phân cực của các chuỗi polynucleotide được thể hiện ở hai vị trí C3' và C5'.

Trong các nucleotide của ADN và ARN, nhóm phosphate liên kết với các nucleoside tại C5'. Mỗi *nucleoside* được tạo thành do một bazơ nối với đường tại C1' bằng một *liên kết N-glycoside*. Cụ thể, C1' nối với N1 của pyrimidine hoặc với N9 của purine. Tên gọi chính thức hay danh pháp của các nucleoside bắt nguồn từ các bazơ tương ứng, trong đó các nucleoside là dẫn xuất của purine có đuôi là *-osine* và các dẫn xuất của pyrimidine có đuôi là *-idine* (Bảng 1.1).

Tóm lại, mỗi *nucleotide* gồm 3 thành phần kết dính với nhau như sau: gốc đường nối với *bazơ* tại C1' bằng một *liên kết β-glycosid* và nối với nhóm *phosphate* tại C5' bằng một *liên kết phosphomonoester* (Hình 1.4).

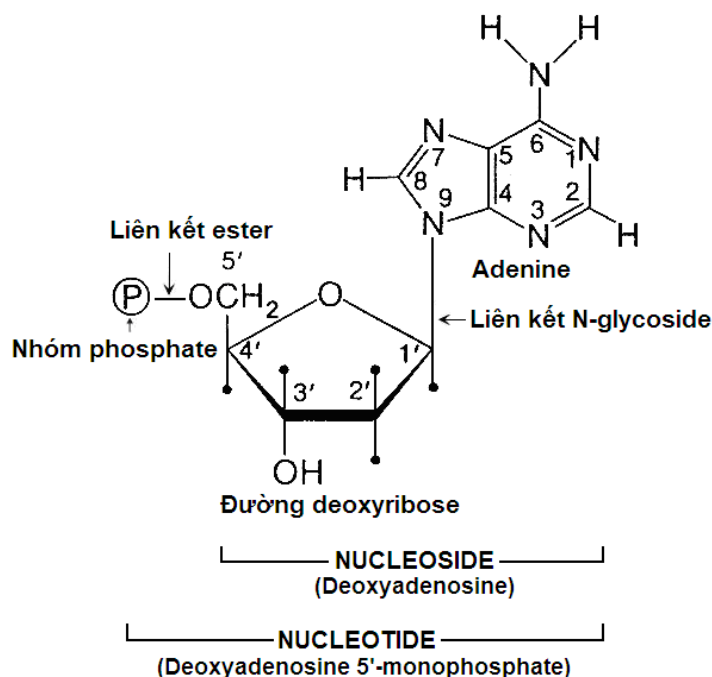
Bảng 1.1. Tên gọi của các nucleoside và nucleotide của ARN và ADN

Bazơ	Nucleoside	Nucleotide
ARN		
Adenine (A)	Adenosine	Adenosine 5'-monophosphate (AMP)
Guanine (G)	Guanosine	Guanosine 5'-monophosphate (GMP)
Cytosine (C)	Cytidine	Cytidine 5'-monophosphate (CMP)

Uracil (U)	Uridine	Uridine 5'-monophosphate (UMP)
ADN		
Adenine (A)	Deoxyadenosine	Deoxyadenosine 5'-phosphate (dAMP)
Guanine (G)	Deoxyguanosine	Deoxyguanosine 5'-phosphate (dGMP)
Cytosine (C)	Deoxycytidine	Deoxycytidine 5'-phosphate (dCMP)
Thymine (T)	Deoxythymidine	Deoxythymidine 5'-phosphate (dTMP)

1.3. Cấu trúc của các chuỗi polynucleotide

Các nucleotide trong ADN hoặc ARN nối với nhau bằng các mối liên kết đồng hoá trị (covalent) có tên là *liên kết 3',5'-phosphodiester* (giữa gốc đường của nucleotide này với nhóm phosphate của nucleotide kế tiếp), tạo thành chuỗi *polynucleotide*. Vì vậy các chuỗi này bao giờ cũng được kéo dài theo chiều 5'→3' (đầu 5' mang nhóm phosphate và đầu 3' chứa nhóm -OH tự do). Chúng có bộ khung vững chắc gồm các gốc đường và phosphate xếp luân phiên nhau, còn các bazơ nằm về một bên. Trình tự các bazơ vì vậy được đọc theo một chiều xác định 5'→3'. Đây là cấu trúc sơ cấp của ADN và ARN (Hình 1.5).



Hình 1.4. Cấu trúc của một deoxyribonucleotide (dAMP)

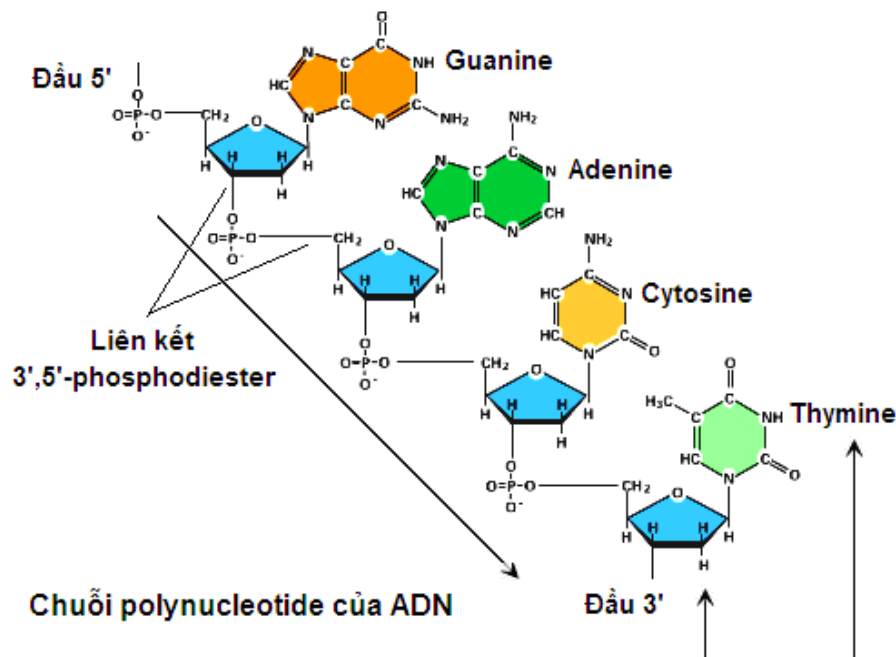
Thông thường người ta biểu diễn trình tự bazơ 5'→3' theo chiều từ trái sang phải. Hình 1.5 cho thấy các chuỗi ADN và ARN chỉ khác nhau bởi bazơ U hoặc T và gốc đường trong các nucleotide của chúng. Nếu bỏ qua sự khác biệt về gốc đường, ta có thể hình dung trình tự các bazơ của hai chuỗi polynucleotide của ADN và ARN đều sinh trưởng theo chiều từ 5' đến 3' (5'→3'), như sau:

Chuỗi ADN: (5') pApApTpTpCpTpApApApTpTpC -OH (3')

Chuỗi ARN: (5') pApApUpUpCpUpUpApApApUpUpC -OH (3')

Cần lưu ý rằng: (1) Các hợp chất dùng để polymer hoá là các nucleoside triphosphate, nhưng

các monomer của axit nucleic lại là monophosphate. Phản ứng trùng hợp này được xúc tác bởi các enzyme *ADN polymerase* và *ARN polymerase*. (2) Các *oligonucleotide* là những đoạn có độ dài thường là ~10-100 nucleotide. Các oligoribonucleotide tồn tại trong tự nhiên và được sử dụng như là những *đoạn mồi* (primer) trong tái bản ADN và cho các mục đích khác nhau trong tế bào. Các oligonucleotide tổng hợp có thể tạo ra bằng sự tổng hợp hoá học và là nguyên liệu thiết yếu cho các kỹ thuật sinh học phân tử, như: giải mã di truyền trong ống nghiệm; xác định trình tự ADN, phản ứng trùng hợp chuỗi bằng polymerase (polymerase chain reaction = PCR), lai tại chỗ (*in situ* hybridization), mẫu dò axit nucleic, lai axit nucleic v.v.



Chuỗi polynucleotide của AN sai khác ở: Ribose (-OH) và bazơ Uracil

Hình 1.5. Cấu trúc chuỗi polynucleotide của ADN (trên) và của ARN.

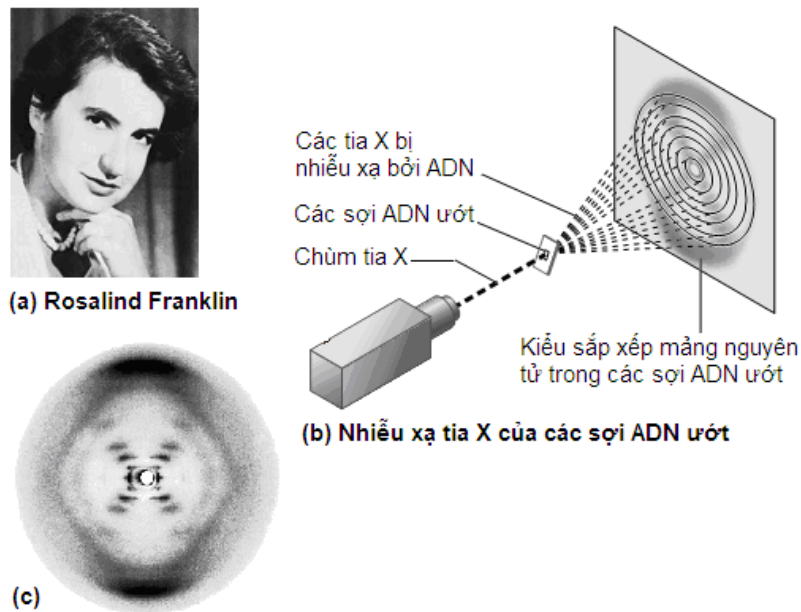
1.4. Cấu trúc của phân tử ADN

1.4.1. Thành phần hóa học của ADN

Bảng 1.2. Thành phần bazơ của ADN ở một số loài

Sinh vật	A%	T%	G%	C%	$\frac{A+G}{T+C}$	$\frac{A+T}{G+C}$
Phage lambda	21,3	22,9	28,6	27,2	1,00	0,79
Phage T7	26,0	26,0	24,0	24,0	1,00	1,08
<i>Mycobacterium tuberculosis</i>	15,1	14,6	34,9	35,4	1,00	0,42
<i>Escherichia coli</i>	24,7	23,6	26,0	25,7	1,03	0,93
<i>Aspergillus niger</i> (nấm mốc)	25,0	24,9	25,1	25,0	1,00	1,00
<i>Saccharomyces cerevisiae</i>	31,3	32,9	18,7	17,1	1,00	1,79
<i>Triticum</i> (lúa mỳ)	27,3	27,1	22,7	22,8	1,00	1,19
<i>Zea mays</i> (ngô)	26,8	27,2	22,8	23,2	0,98	1,17
<i>Salmo salar</i> (cá hồi)	29,7	29,1	20,8	20,4	1,02	1,43

<i>Gallus domestica</i> (gà nhà)	29,5	27,7	22,4	20,4	1,08	1,34
<i>Homo sapiens</i> (người)	30,9	29,4	19,9	19,8	1,01	1,52

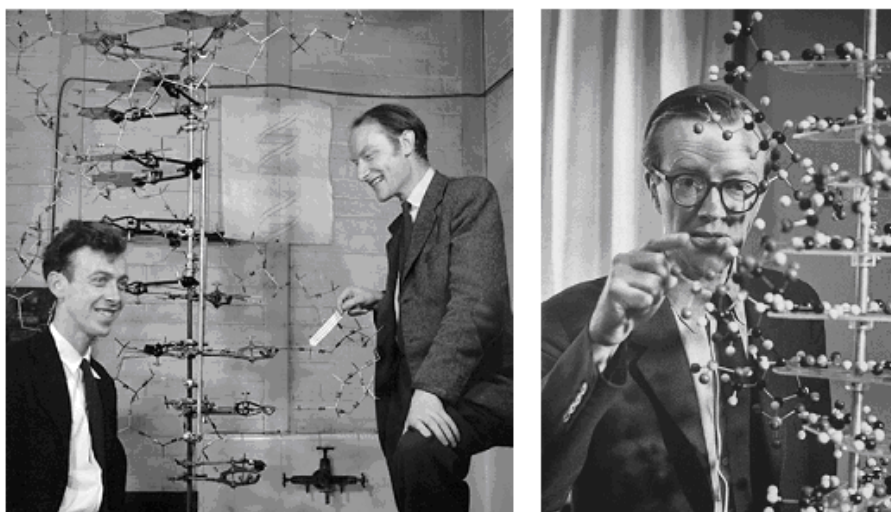


Hình 1.6. (a) R. Franklin; (b) Nhiễu xạ tia X của ADN, và (c) Ảnh chụp.

Erwin Chargaff (1949) lần đầu tiên áp dụng phương pháp sắc ký giấy vào việc phân tích thành phần hóa học của ADN các loài khác nhau đã khám phá ra rằng (Bảng 1.2): (i) Số lượng bốn loại bazơ trong ADN là không bằng nhau; (ii) Tỷ lệ tương đối của các bazơ là không ngẫu nhiên; trong tất cả các mẫu ADN nghiên cứu có tương quan về hàm lượng (%) giữa các bazơ: $A \approx T$ và $G \approx C$, nghĩa là tỷ số $(A+G)/(T+C) \approx 1$; và (iii) Mỗi loài có một tỷ lệ $(A+T)/(G+C)$ đặc thù.

1.4.2. Cấu trúc của chuỗi xoắn kép ADN

Vào năm 1951-1952, việc nghiên cứu cấu trúc ba chiều của ADN bằng phân tích nhiễu xạ tia X được bắt đầu bởi Maurice Wilkins và Rosalind Franklin (Hình 1.6). Các bức ảnh chụp được gợi ý rằng ADN có cấu trúc xoắn gồm hai hoặc ba chuỗi. Tuy nhiên, giải pháp đúng đắn nhất là *chuỗi xoắn kép* bổ sung do Watson và Crick đưa ra năm 1953 (Hình 1.7 và 1.8). Mô hình này hoàn toàn phù hợp với các số liệu của Wilkins và Franklin cũng như của Chargaff. Sự kiện này mở ra một bước ngoặt mới cho sự ra đời và phát triển nhanh chóng của sinh học phân tử. Với phát minh về mô hình cấu trúc phân tử ADN, Watson và Crick cùng chia sẻ với Wilkins giải thưởng Nobel năm 1962.



Hình 1.7. J.Watson (trái) và F.Crick cùng với M. Wilkins bên cạnh mô hình cấu trúc phân tử ADN làm nên tên tuổi của họ.

Mô hình Watson-Crick (hay ADN dạng B) có các đặc điểm sau:

(1) ADN gồm hai chuỗi *đối song song* (antiparallel) cùng uốn quanh một trục trung tâm theo chiều xoắn phải, với đường kính $20\text{\AA}$ ($1\text{\AA} = 10^{-10}\text{m}$), gồm nhiều vòng xoắn lặp lại một cách đều đặn và chiều cao mỗi vòng xoắn là $34\text{\AA}$, ứng với 10 cặp bazơ (base pair = bp).



(2) Các bộ khung đường-phosphate phân bố ở mặt ngoài chuỗi xoắn và các bazơ nằm ở bên trong; chúng xếp trên những mặt phẳng song song với nhau và thẳng góc với trục phân tử, với khoảng cách $\sim 3,4\text{\AA}$.

(3) Hai mạch đơn của ADN gắn với nhau bằng các *liên kết hydro* giữa các cặp bazơ đối diện (nằm cách nhau khoảng $3\text{\AA}$) theo *nguyên tắc bổ sung*, đó là: A-T (2 liên kết hydro) và G-C (3 liên kết) - Hình 1.8 và 1.9a.

(4) Tính chất bổ sung theo cặp bazơ dẫn đến sự bổ sung về trình tự các bazơ giữa hai mạch đơn của mỗi chuỗi xoắn kép. Vì vậy, trong bất kỳ một phân tử ADN mạch kép nào hoặc một đoạn của nó bao giờ cũng có: $A = T$ và $G = C$; nghĩa là: $[A + G] = [T + C]$ hay $\frac{A+G}{T+C} = 1$, còn tỷ lệ

$\frac{A+T}{G+C}$ là đặc thù cho từng loài. Như vậy, mô hình cấu trúc chuỗi xoắn kép của Watson-Crick (1953) hoàn toàn thỏa mãn các kết quả nghiên cứu của Chargaff (1949); và các biểu thức $A = T$ và $G = C$ còn gọi là các *quy luật Chargaff* (Chargaff's rules).

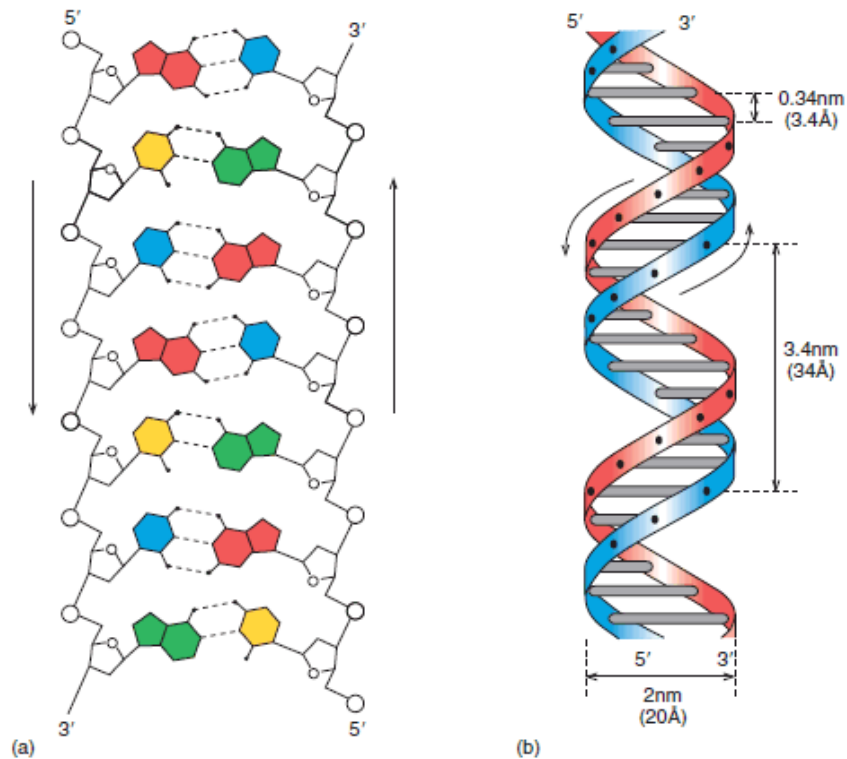
Vì vậy, khi biết trình tự bazơ ở một mạch đơn của ADN, ta có thể xác định được trình tự bazơ ở mạch bổ sung của nó. Ví dụ:

Mạch 1 (cho trước): $5'-\text{AATTCTTAAATTC}-3'$

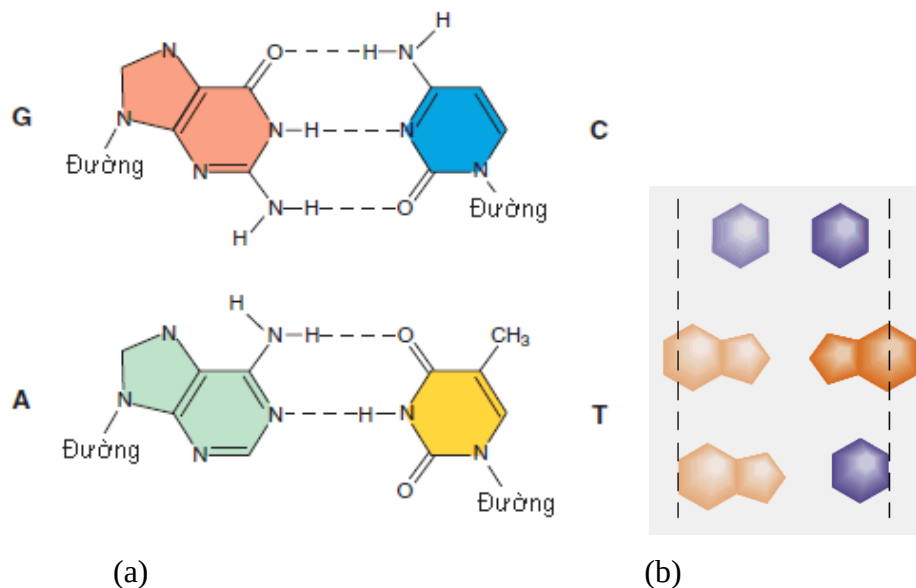
Mạch 2 (bổ sung): $3'-\text{TTAAGAATTTAAG}-5'$

Tóm lại, hai đặc điểm quan trọng nhất trong cấu trúc ADN là sự *phân cực ngược chiều* của hai mạch đơn ($5' \rightarrow 3'$ và $3' \rightarrow 5'$) và *nguyên tắc bổ sung* của các cặp bazơ (A-T và G-C). Đây là hai

nguyên lý căn bản chi phối các cơ chế di truyền ở cấp độ phân tử (tái bản, phiên mã và dịch mã), mà ta có thể hình dung tổng quát ở sơ đồ gọi là *Giáo lý trung tâm* của Sinh học phân tử.



Hình 1.8. Các mô hình cấu trúc chuỗi xoắn kép ADN.



Hình 1.9. (a) Hai kiểu kết cặp bazơ bổ sung A-T và G-C. (b) Sơ đồ minh họa các khả năng kết cặp “1 pyrimidine - 1 pyrimidine”, “1 purine - 1 purine”, và “1 purine - 1 pyrimidine” (xem giải thích trong bài).

Cần lưu ý rằng, theo nguyên tắc kết cặp các bazơ đối diện trên 2 mạch đơn của ADN có thể có các trường hợp sau: “1 pyrimidine - 1 pyrimidine”, “1 purine - 1 purine”, “1 purine - 1 pyrimidine” như ở Hình 1.9b Tuy nhiên, như ta thấy, hai trường hợp đầu cho thấy chúng hoặc là quá mỏng hoặc quá dày so với đường kính phân tử. Chỉ có trường hợp “1 purine - 1 pyrimidine” là phù hợp. Như vậy có thể có 4 kiểu kết cặp là A-T, G-C, A-C và G-T; trong đó chỉ có hai kiểu A-T

và G-C là bền vững, còn các kiểu A-C và G-T vì chỉ có một liên kết hydro, kém bền nên chúng chỉ xuất hiện như là ngoại lệ khi có sự hỗ biến của các bazơ và hậu quả là dẫn tới sự phát sinh các đột biến thay thế bazơ dạng đồng hoán trong quá trình tái bản ADN (chương 7).

1.4.3. Các dạng biến đổi của ADN

Mô hình Watson-Crick hay ADN dạng B là cấu trúc phổ biến. Tuy nhiên, sau này người ta còn phát hiện ra nhiều dạng xoắn phải khác (A, C, D...); chúng có một số biến đổi so với ADN-B (Bảng 1.3). Bên cạnh các dạng ADN xoắn phải (A, B, C...), Alexander Rich và đồng sự (1979) còn phát hiện thêm một dạng ADN xoắn trái duy nhất cho đến nay. Dạng ADN này có bộ khung zigzag uốn gập khúc theo chiều xoắn trái, mỗi vòng xoắn dài $45,6\text{\AA}$ chứa 12 cặp bazơ. Nhìn chung, so với ADN dạng B, ADN-Z dài và thon gầy hơn, các rãnh lớn bị dẹt ra phần bề mặt của chuỗi xoắn; còn ADN dạng A ngắn và mập hơn (Hình 1.10; Bảng 1.3).

Bảng 1.3. Một số đặc điểm chính của các ADN dạng A, B và Z.

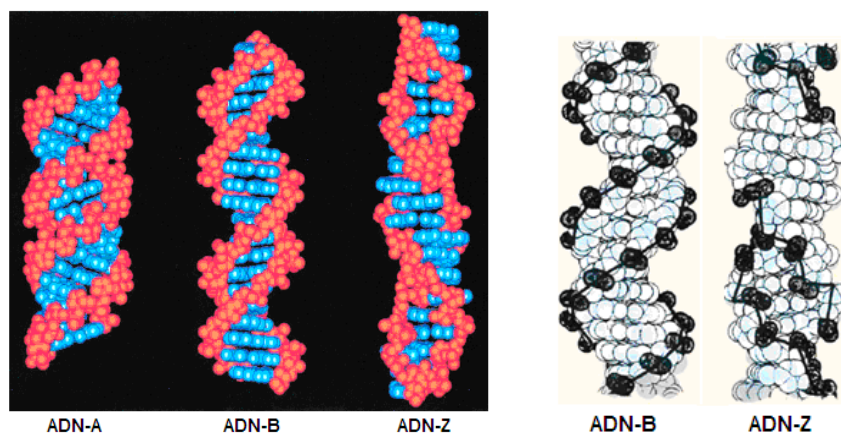
Dạng	Chiều xoắn	Số bp/vòng xoắn	Đường kính chuỗi xoắn
A	Phải	11,0	$23\text{\AA}$
B	Phải	10,0	$19\text{\AA}$
Z	Trái	12,0	$18\text{\AA}$

Những vùng nào của ADN có chứa các purine và pyrimidine sắp xếp xen kẽ nhau trên một mạch thì có thể tiếp nhận cấu hình ADN-Z, ví dụ:

5'--CGCGCG--3'

3'--GCGCGC--5'

Sự chuyển đổi này diễn ra thuận lợi bởi sự có mặt của 5-methylcytosine và bởi trạng thái siêu xoắn âm (negative supercoiling). ADN là một phân tử động học, vì vậy nó có thể chuyển từ một cấu hình này sang một cấu hình khác dựa trên các lực bên ngoài trong tế bào. Sự chuyển đổi từ dạng B sang dạng Z có thể có liên quan đến sự điều hoà biểu hiện gen. Cấu trúc này cũng có mặt trong các tế bào sống với một tỷ lệ rất nhỏ song chức năng của nó còn chưa thật sự hiểu rõ.



Hình 1.10. Các mô hình ADN dạng A, B và Z.

1.4.4. Đặc tính hóa lý của ADN

Một trong những đặc điểm quan trọng nhất của ADN là hai mạch đơn bổ sung của nó gắn với

nhau bằng các liên kết hydro, vốn là lực hóa học yếu nên chúng có thể bị phân hủy dưới tác dụng của các enzyme, năng lượng... làm cho hai mạch đơn của chuỗi xoắn kép tách rời nhau, gọi là *biến tính* (denaturation). Nhờ đó ADN mới có thể tái bản và các gen có thể phiên mã và biểu hiện ra sản phẩm của chúng. Mặt khác, ADN có thể phục hồi trạng thái ban đầu gọi là *hồi tính* (renaturation).

Bằng thực nghiệm, người ta đã chứng minh điều đó bằng cách sử dụng các tác nhân vật lý và hóa học khác nhau. Chẳng hạn, khi đun nóng từ từ khi đun nóng từ từ dung dịch chứa ADN thì ở nhiệt độ vừa phải, thì các phân tử ADN bị biến tính từng phần, và khi tăng lên tới nhiệt độ gần 100°C (thường là 90-95°C) thì các liên kết hydro của chúng bị phá hủy hoàn toàn và hai mạch bổ sung tách ra (biến tính hoàn toàn). Điều đó có thể giải thích như sau: Vì mỗi cặp A-T chỉ có hai liên kết hydro, kém bền hơn so với mỗi cặp G-C có tới ba liên kết, cho nên các vùng giàu cặp A-T sẽ tách từng phần trước, trong khi các vùng giàu cặp G-C vẫn giữ nguyên đặc tính xoắn kép và chỉ tách ở nhiệt độ cao. Ngược lại, khi làm nguội từ từ dung dịch ADN nóng chảy hoàn toàn này thì các mạch đơn thường cặp lại với mạch bổ sung của chúng và làm phục hồi cấu trúc chuỗi xoắn kép như lúc đầu. Đây là hai quá trình thuận-ngịch.

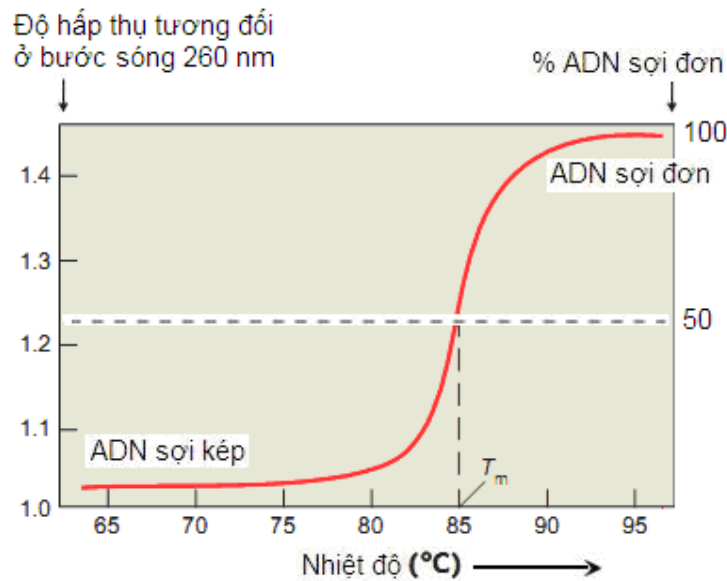
– *Biến tính hay sự tách hai mạch của chuỗi xoắn kép ADN*

Trong khi các tỷ số G với C và A với T trong ADN của một sinh vật là cố định, thì hàm lượng GC (tỷ lệ phần trăm của G + C) có thể sai khác nhau một cách đáng kể giữa các ADN thuộc các loài khác nhau. Bảng 1.4 cho thấy hàm lượng GC của ADN nhiều loài sinh vật. Các trị số này biến thiên từ 22% đến 73%, và điều đó được phản ánh trong sự sai khác về các đặc tính của ADN.

Nhiệt độ mà tại đó các mạch ADN tách nhau một nửa được gọi là *nhiệt độ nóng chảy* (melting temperature), ký hiệu là T_m . T_m là điểm giữa của pha chuyển tiếp và nó tùy thuộc vào hàm lượng GC của ADN, nghĩa là đặc trưng cho ADN mỗi loài. Trên thực tế, hàm lượng GC của ADN càng cao thì T_m của nó càng cao (Hình 1.12). Ví dụ, ADN của *E. coli* với 50-51% GC thì có T_m là 69-70°C. Tương tự, kết quả xử lý nhiệt đối với ADN phế cầu khuẩn *Streptococcus pneumoniae* và nhiệt độ nóng chảy của nó được đo bằng sự gia tăng độ hấp thụ ở 260-nm cho phép thu được đường cong nóng chảy của vi khuẩn này. T_m cho ADN này dưới những điều kiện như thế là khoảng 85°C (Hình 1.11).

Bảng 1.4. Hàm lượng tương đối (G + C) của các ADN khác nhau

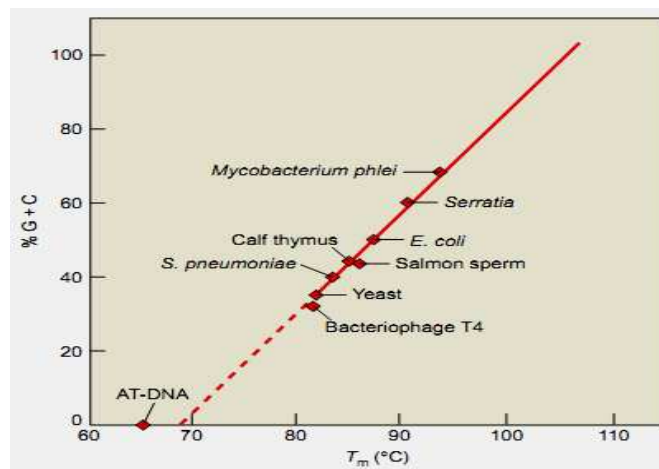
Nguồn ADN	(G+C)%	Nguồn ADN	(G+C)%
<i>Dictyostelium</i> (mốc nhầy)	22	Lách chuột	44
<i>Streptococcus pyogens</i>	34	Tinh trùng cá hồi	44
<i>Bacillus cereus</i>	37	<i>B. subtilis</i>	44
<i>Hemophilus influenzae</i>	39	<i>Escherichia coli</i>	51
<i>Saccharomyces cerevisiae</i>	39	Phage T7	51
Tuyến ức bê	40	Phage T3	53
Gan chuột (<i>Rattus</i>)	40	<i>Neurospora crassa</i>	54
<i>Streptococcus pneumoniae</i>	42	Herpes simplex virus	72
Mầm lúa mì	43	<i>Mycobacterium phlei</i>	73



Hình 1.11. Đường cong nóng chảy của ADN *Streptococcus pneumoniae*. ADN sợi kép bị biến tính bởi việc đun nóng và nhiệt độ nóng chảy của nó được đo bằng sự tăng độ hấp thụ ở bước sóng 260 nm. Điểm mà tại đó 50% cặp bazơ bị biến tính hay một nửa ADN có dạng sợi đơn gọi là nhiệt độ nóng chảy (T_m). Ở ví dụ này là khoảng 85°C. (Phỏng theo P. Doty, *The Harvey Lectures* 55:121, 1961).

– Sự phục hồi trạng thái nguyên thể của ADN hay hồi tính

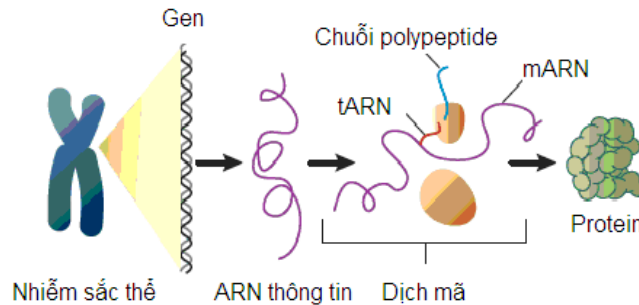
Một khi hai mạch của ADN tách ra, dưới những điều kiện thích hợp chúng có thể kết hợp trở lại và phục hồi trạng thái ban đầu. Góp phần vào hiệu quả "hồi tính" này của ADN có nhiều nhân tố. Dưới đây nêu lên ba nhân tố quan trọng nhất: (i) *Nhiệt độ* tối ưu cho sự hồi tính của một ADN là khoảng 25°C dưới nhiệt độ nóng chảy của nó; (ii) *Nồng độ ADN* trong dung dịch cũng quan trọng. Trong giới hạn hợp lý, nồng độ ADN càng cao thì hai mạch bổ sung sẽ càng dễ dàng bắt gặp nhau trong một thời gian nào đó. Nói cách khác, nồng độ càng cao thì sự hàn gắn trở lại càng nhanh. (iii) *Thời gian* cho phép hai mạch hàn gắn trở lại càng dài thì sự hồi tính xảy ra càng dễ dàng.



Hình 1.12. Mối quan hệ giữa nhiệt độ nóng chảy của ADN và hàm lượng GC. ADN chỉ có AT (AT-DNA) là các ADN tổng hợp bao gồm chủ yếu là A và T (hàm lượng GC = 0). (Phỏng theo P. Doty, *The Harvey Lectures* 55:121, 1961).

1.4.5. Chức năng của ADN

Ngày nay, chúng ta đều biết rằng ADN hay bộ gen của tất cả các sinh vật nói chung có chức năng chính là mang đầy đủ toàn bộ *thông tin di truyền* đặc trưng của loài. Thông tin này được ghi lại dưới dạng mật mã, gọi là *mã di truyền*, chứa đựng trong các *gen cấu trúc* cũng như các *yếu tố điều hòa* để điều khiển mọi hoạt động sinh trưởng, phân chia và biệt hoá của tế bào. Các chức năng và cơ chế truyền đạt thông tin di truyền chính yếu của ADN được mô tả tóm tắt như dưới đây.



Hơn nữa, ADN hay vật chất di truyền nói chung đều có khả năng tự sao chép một cách chính xác bản thân nó trong quá trình *tái bản* (replication). Đây là cơ sở của sự tự nhân đôi nhiễm sắc thể và phân chia tế bào, qua đó truyền đạt vật chất di truyền cho thế hệ sau. Đó còn là các quá trình hoạt động và điều hoà sự biểu hiện của các gen trong bộ gen, thường được kể đến như là các quá trình *phiên mã* (transcription) và *dịch mã* (translation) để tổng hợp các phân tử ARN và cuối cùng là các protein tham gia vào các cấu trúc và hoạt động cơ sở của tế bào. Nhờ đó mà con cái sinh ra thường giống với cha mẹ, mỗi loài duy trì sự ổn định tương đối bộ gen của mình và nói rộng ra là, nhờ đó mà sự sống được duy trì một cách liên tục kể từ khi sự sống bắt đầu hình thành trên trái đất cách đây chừng ba tỷ năm.

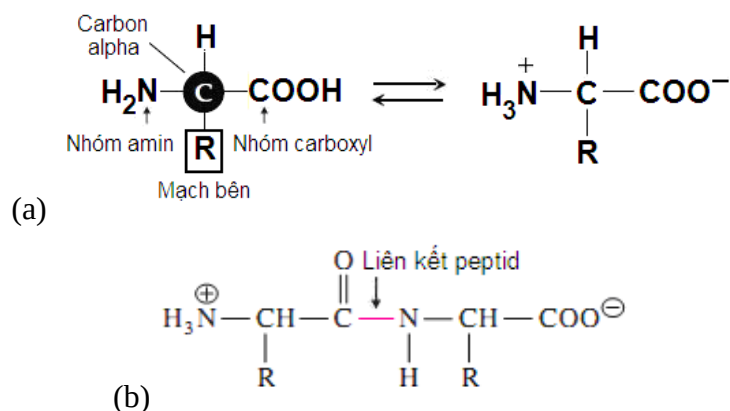
Mặt khác, ADN còn có khả năng phát sinh các biến đổi trong quá trình phát triển cá thể và sinh sản. Đó là các *đột biến gen* (gene mutations) gây ra bởi tác động của các tác nhân vật lý và hoá học khác nhau, hoặc do chính các sai sót trong quá trình tái bản, hoặc do sự dịch chuyển vị trí của bản thân các gen trong bộ gen - các *yếu tố di truyền vận động* (transposable genetic elements) hay còn gọi là các *gen nhảy* (jumping genes) – gây nên sự biến động của bộ gen và biến đổi ở kiểu hình. Ngoài ra, đó còn là các quá trình *tái tổ hợp di truyền* (genetic recombination) tạo nên các biến dị tổ hợp phong phú và đa dạng trong quá trình sinh sản của sinh vật. Chính các quá trình biến đổi đa dạng này đã không ngừng tạo nên các nguồn biến dị di truyền sơ cấp và thứ cấp cho sự chọn lọc và tiến hoá của sinh giới.

2. Cấu trúc và chức năng của protein

2.1. Cấu trúc của protein

Các protein là những polymer sinh học được tạo ra bởi sự kết nối của các *axit amin* (amino acid) với nhau bằng các *liên kết peptide*. Có 20 loại L- α -axit amin được phát hiện trong các protein của các tế bào.

Về cấu trúc, nói chung, mỗi axit amin gồm có một nguyên tử carbon alpha ($C\alpha$) ở vị trí trung tâm, đính xung quanh nó là một nhóm amin ($-NH_2$), một nhóm carboxyl ($-COOH$), một nguyên tử hydro ($-H$) và một gốc R hay chuỗi bên đặc trưng cho từng loại axit amin (Hình 1.14). Khi ở trạng thái dung dịch, các nhóm amin và carboxyl thường phân ly thành trạng thái ion, tương ứng là ^+H_3N - và $-COO^-$ (Hình 1.13).

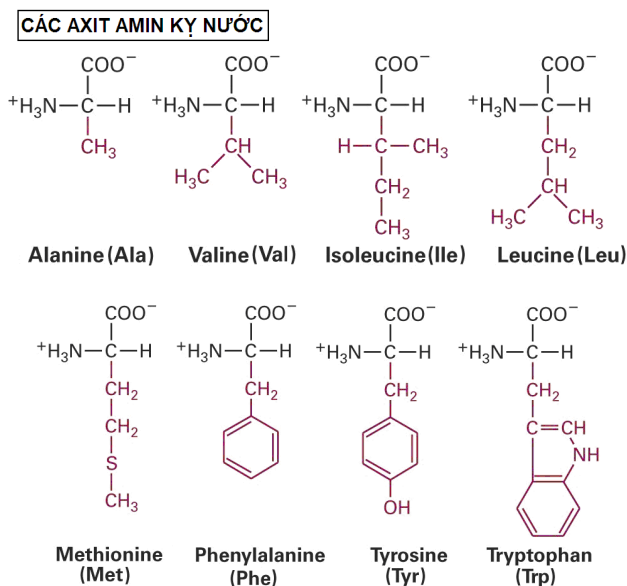


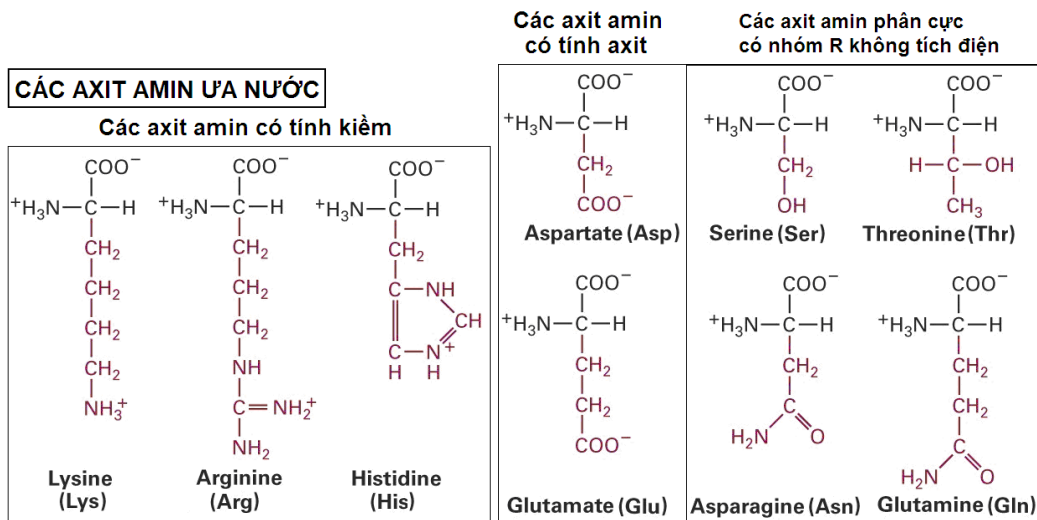
Hình 1.13. (a) Cấu trúc chung của các axit amin. (b) Một dipeptide.

Hai axit amin nối với nhau bằng một liên kết peptide ($-\text{CO}-\text{NH}-$) giữa nhóm carboxyl của axit amin này với nhóm amin của axit amin kế tiếp và loại trừ một phân tử nước; cứ như thế các axit amin kết nối với nhau tạo thành một chuỗi gồm nhiều axit amin, gọi là *polypeptide* (Hình 1.15). Mỗi chuỗi polypeptide luôn luôn có chiều xác định $^+\text{H}_3\text{N} \rightarrow \text{COO}^-$ (do tác dụng của *peptidyl-transferase*) và được đặc trưng về số lượng, thành phần và chủ yếu là trình tự sắp xếp của các axit amin.

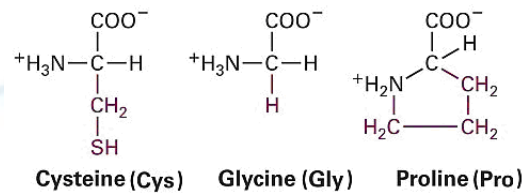
Có bốn mức độ cấu trúc của các protein (xem Hình 1.16).

Cấu trúc protein bậc I là trình tự sắp xếp của các axit amin trong một chuỗi polypeptide. Đây là bậc cấu trúc cơ sở quan trọng nhất của tất cả các protein do gen quy định.

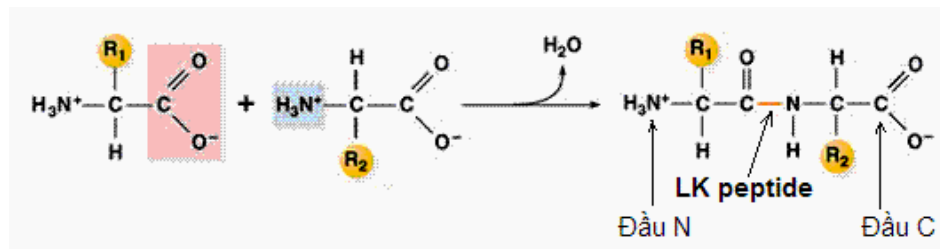




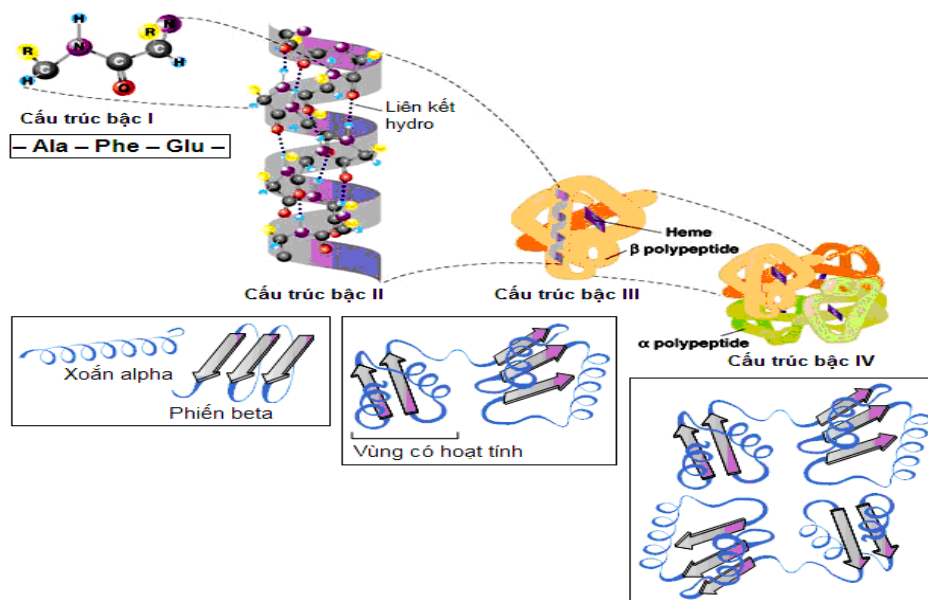
CÁC AXIT AMIN ĐẶC BIỆT



Hình 1.14. Cấu tạo của 20 loại axit amin thuộc các nhóm khác nhau.



Hình 1.15. Sự hình thành liên kết peptide giữa các axit amin.



Hình 1.16. Bốn bậc cấu trúc của protein.

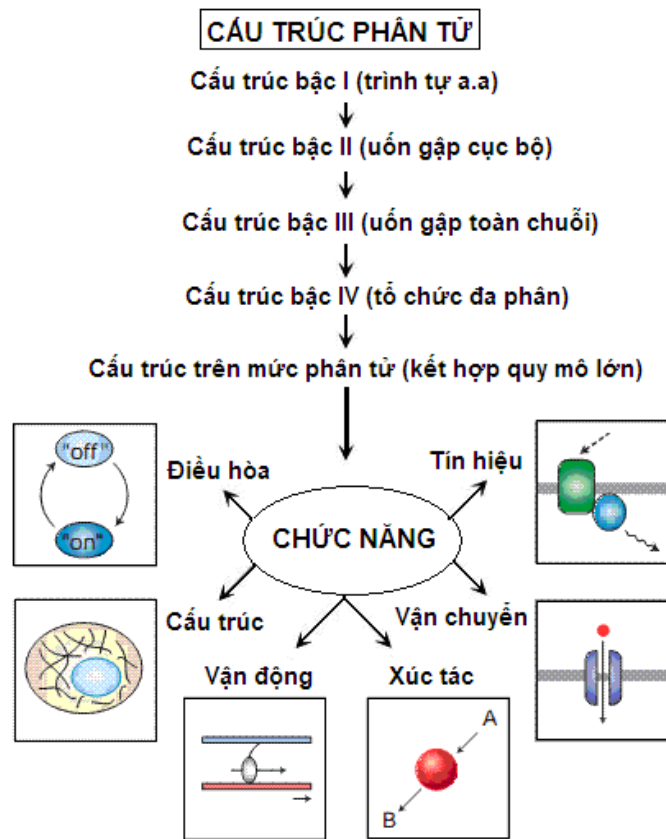
Cấu trúc protein bậc II xảy ra khi trình tự các axit amin trong một chuỗi polypeptide nối với nhau bằng các liên kết hydro. Cấu trúc này có hai kiểu cơ bản, đó là: chuỗi xoắn alpha (xoắn trái) và tấm beta (gấp nếp).

Cấu trúc protein bậc III xảy ra khi các lực hấp dẫn nào đó có mặt giữa các vùng xoắn alpha và các tấm beta gấp nếp trong một chuỗi polypeptide, hình thành nên một cấu trúc cuộn gấp có dạng khối cầu. Một số protein chức năng có cấu trúc kiểu này, ví dụ myoglobin...

Cấu trúc protein bậc IV là một protein gồm hai hoặc nhiều chuỗi polypeptide cùng loại hoặc khác loại kết hợp với nhau. Có khá nhiều protein chức năng có kiểu cấu trúc này; một số như hemoglobin còn có thêm ion kim loại như Fe^{++} trong túi heme của nó.

2.2. Chức năng của protein

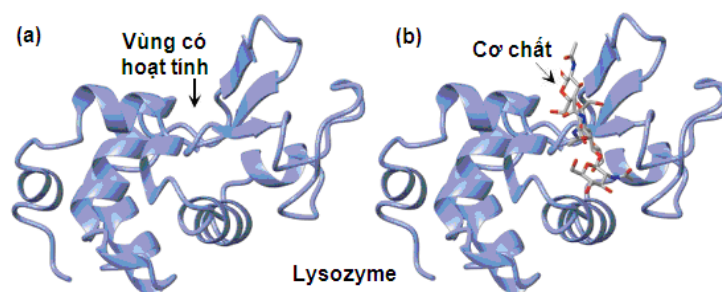
Nói chung, protein là các hợp chất hữu cơ vốn là cơ sở của sự sống, với các chức năng thiết yếu sau đây (Hình 1.17):



Hình 1.17. Tổng quát về cấu trúc và chức năng của protein.

(i) Các protein là *thành phần cấu tạo* cơ sở của các tế bào, bao gồm các màng tế bào, các bào quan, bộ máy di truyền của chúng. Đó cũng là các protein dạng mạch làm thành các cơ quan, bộ phận trên cơ thể các động vật, như: collagen làm nên xương, sụn, gân và da; keratin cấu tạo nên các lớp ngoài cùng của da và tóc, móng, sừng và lông;

(ii) Các *enzyme* đóng vai trò xúc tác cho tất cả các phản ứng hóa học trong tế bào và cơ thể đều là những protein hình cầu. Quan trọng nhất là các enzyme tham gia vào các con đường chuyển hóa và các enzyme tham gia vào các quá trình truyền thông tin di truyền trong tế bào (Hình 1.18).



Hình 1.18. Lysozyme lòng trắng trứng gà. (a) Lysozyme tự do. (b) Lysozyme có cơ chất bám vào. (Phỏng theo Horton et al, 2006).

(iii) Các *kháng thể* (antibodies) trong hệ thống miễn dịch, còn gọi là các immunoglobulin, làm ra hàng ngàn protein khác nhau vốn được sinh ra trong huyết thanh máu phản ứng lại với các *kháng nguyên* (antigens). Chúng đóng vai trò bảo vệ cơ thể chống lại sự xâm nhập của các vật lạ.

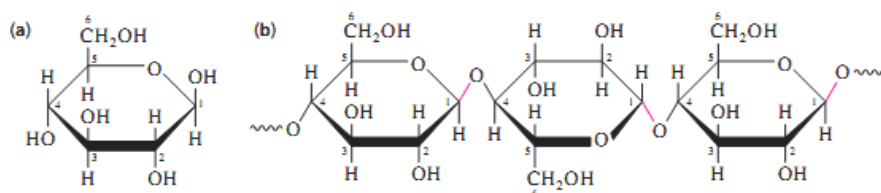
(iv) Các *hormone* protein bắt nguồn từ các tuyến nội tiết thì không hoạt động như các enzyme. Thay vì kích thích các cơ quan đích, chúng kiểm soát các hoạt động quan trọng, như tốc độ chuyển hóa và sản xuất các enzyme tiêu hóa và sữa chẳng hạn. Insulin (từ tuyến tụy) điều hòa sự chuyển hóa carbohydrate bằng cách kiểm soát các mức glucose trong máu. Thyroglobulin (từ tuyến giáp) điều hòa các quá trình chuyển hóa nói chung; calcitonin cũng từ tuyến giáp làm hạ thấp mức calci máu.

(v) Ngoài ra, các protein còn là *nguồn dinh dưỡng* chính cung cấp năng lượng cho tế bào và cơ thể duy trì các hoạt động trao đổi chất và lớn lên. Ví dụ *hemoglobin* mang các sinh chất theo máu đi khắp cơ thể; các *fibrinogen* và fibrin được biến đổi từ nó vốn có trong máu cần thiết cho quá trình đông máu. Các *protein cơ* mà chủ yếu là myosin phối hợp với actin tạo thành actomyosin, chịu trách nhiệm cho hoạt động cơ cơ...

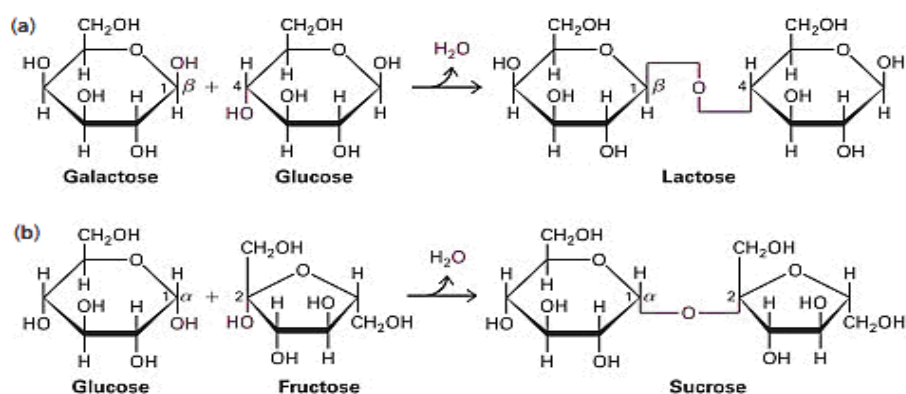
3. Cấu trúc và chức năng của các Polysaccharide

3.1. Cấu trúc của các Polysaccharide

Các polysaccharide và hợp chất carbohydrate nói chung có hành phần hóa học gồm các nguyên tố C, H và O, theo tỷ lệ thông thường là 1:2:1 hay $(CH_2O)_n$. Đơn phân cấu tạo nên chúng là các *monosaccharide*, chủ yếu là glucose. Các monosaccharide nối với nhau bằng các *liên kết glycosid* tạo thành các chuỗi *polysaccharide*. (Hình 1.19 và 1.20)



Hình 1.19. (a) Glucose và (b) cellulose. Cellulose là một polymer mạch thẳng do các gốc glucose nối với nhau bằng các liên kết glycosid.



Hình 1.20. Sự hình thành các disaccharide Lactose (a) và Sucrose (b) bằng các liên kết β 1,4- và α 1,2-glycosid.

3.2. Chức năng của các Polysaccharide

Các polysaccharide có hai chức năng quan trọng nhất trong tế bào, đó là tham gia vào cấu tạo và dự trữ năng lượng. Về cấu tạo, *cellulose* tạo nên vách tế bào thực vật và nó là hợp chất hữu cơ có mặt phổ biến trong sinh quyển. Về năng lượng, nguồn dự trữ năng lượng ở tế bào động vật là *glycogen*, trong khi đó ở các tế bào thực vật là *tinh bột*.

4. Cấu trúc và chức năng của Lipid

4.1. Cấu trúc của Lipid

Đơn vị cấu trúc cơ sở của lipid là các axit béo. Mỗi axit béo bao gồm một chuỗi hydrocarbon đính vào nhóm carboxyl ($-COOH$); chúng khác nhau về chiều dài của mạch, số carbon cũng như số lượng và vị trí của các liên kết đôi carbon-carbon ($C=C$). Các axit béo phổ biến trong các tế bào có số nguyên tử carbon thường là 14, 16, 18 hoặc 20 (xem Bảng 1.5). Các axit béo thường được viết tắt $C_x:y$, trong đó x là số lượng carbon trong chuỗi và y là số liên kết đôi. Chuỗi hay mạch hydrocarbon dài của một axit béo có thể không chứa liên kết đôi thì gọi là *axit béo no* hay *bão hòa* (saturated) hoặc có một hay nhiều liên kết đôi thì gọi là axit béo *không no* hay *không bão hòa* (unsaturated).

Bảng 1.5. Một số axit béo chiếm ưu thế trong các phospholipid (Lược từ Lodish et al, 2008)

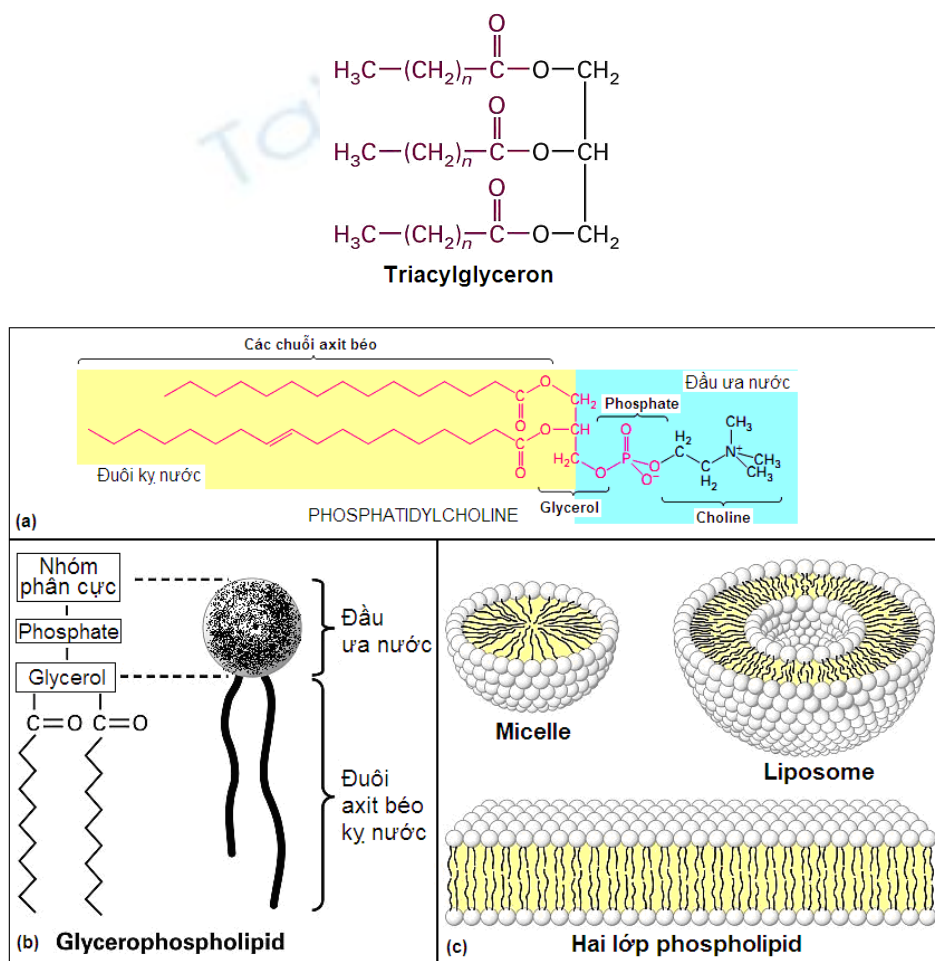
Tên phổ biến của các axit (Dạng ion hóa trong ngoặc đơn)	Viết tắt	Công thức hóa học
Các axit béo bão hòa:		
Palmitic (palmitate)	C16:0	$CH_3(CH_2)_{14}COOH$
Stearic (stearate)	C18:0	$CH_3(CH_2)_{16}COOH$
Các axit béo không bão hòa:		
Oleic (oleate)	C18:1	$CH_3(CH_2)_7CH=CH(CH_2)_7COOH$
Linoleic (linoleate)	C18:2	$CH_3(CH_2)_4CH=CHCH_2CH=CH(CH_2)_7COOH$

Tất cả các phosphoglyceride đều là những phân tử lưỡng cực (amphipathic), có một cái **đuôi kỵ nước** (thường là hai chuỗi axit béo) và một cái **đầu ưa nước** (hydrophilic head; Hình 1.21a,b). Trong dung dịch nước, hiệu quả ưa nước và các kiểu tương tác van der Waals đưa lại sự tổ chức và ổn định các phospholipid vào một trong ba cấu trúc sau: *micelle*, *liposome* hoặc là *màng hai lớp* (Hình 1.21c).

4.2. Chức năng của Lipid

Lipid là thành phần cấu tạo quan trọng của các *màng sinh học* (biomembrane; Hình 1.21), ngăn cách giữa các tế bào với môi trường chung quanh, giữa các khoang trong tế bào sinh vật nhân chuẩn, đồng thời nó là nơi bám của một số protein.

Cũng như glucose, các axit béo là nguồn năng lượng cho nhiều tế bào và được bảo quản dưới dạng *triacylglycerol* (xem Hình dưới) trong mô mỡ. Ngoài ra, các axit béo còn là những chất tiền thân cho phospholipid và nhiều lipid khác với một loạt chức năng khác nhau.



Hình 1.21. Cấu trúc của phospholipid. (a) Cấu trúc của phosphat-idylcholine, một phosphoglyceride điển hình, và (b) Mô hình cấu tạo của glycerophospholipid. (c) Tổ chức của các phospholipid trong cấu trúc của một micelle, liposome và màng hai lớp trong dung dịch nước.