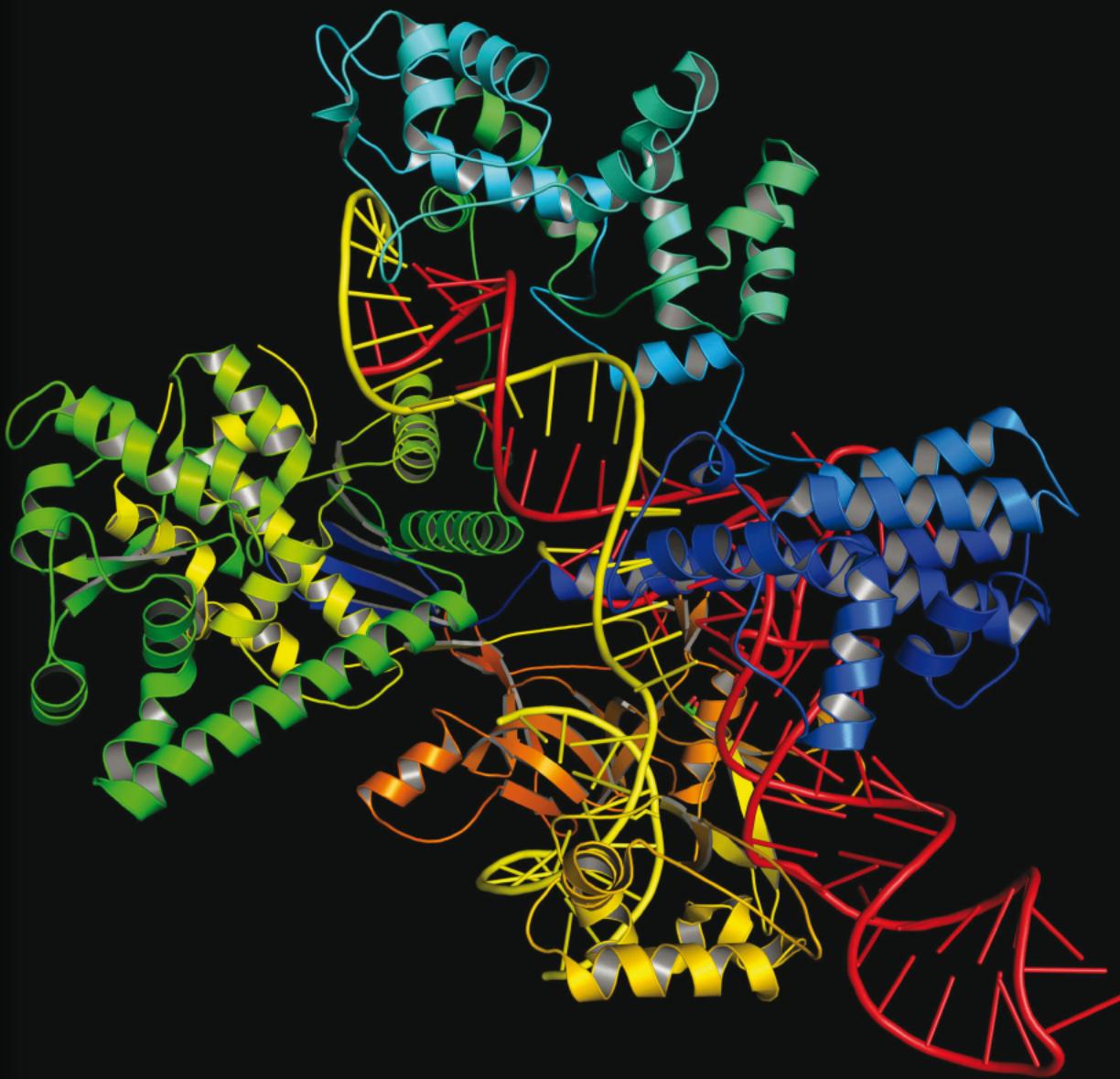




Biologia Molecolare

II Edizione

G. Capranico
E. Martegani
G. Musci
G. Raugei
T. Russo
N. Zambrano
V. Zappavigna



Accedi all'ebook e ai contenuti digitali

Espandi le tue risorse
un libro che **non pesa**
e si **adatta** alle dimensioni
del tuo lettore!



Utilizza il codice personale contenuto nel riquadro per registrarti al sito **edisesuniversita.it** e attivare la tua **area riservata**. Potrai accedere alla **versione digitale** del testo e a ulteriore **materiale didattico**.

Scopri il tuo **codice personale** grattando delicatamente la superficie

Il volume NON può essere venduto, né restituito, se il codice personale risulta visibile.
L'accesso al materiale didattico sarà consentito **per 18 mesi**.

Per attivare i **servizi riservati**, collegati al sito **edisesuniversita.it** e segui queste semplici istruzioni

Se sei registrato al sito

- clicca su *Accedi al materiale didattico*
- inserisci email e password
- inserisci le ultime 4 cifre del codice ISBN, riportato in basso a destra sul retro di copertina
- inserisci il tuo **codice personale** per essere reindirizzato automaticamente all'area riservata

Se non sei già registrato al sito

- clicca su *Accedi al materiale didattico*
- registrati al sito o autenticati tramite facebook
- attendi l'email di conferma per perfezionare la registrazione
- torna sul sito **edisesuniversita.it** e segui la procedura già descritta per *utenti registrati*



Ulteriori materiali e strumenti didattici sono accessibili dalla propria **area riservata** secondo la procedura indicata nel frontespizio.

Dalla sezione **materiali e servizi** della tua area riservata potrai accedere a:

- **Ebook:** versione digitale del testo in formato epub, standard dinamico che organizza il flusso di testo in base al dispositivo sul quale viene visualizzato. Fruibile mediante l'applicazione gratuita Bookshelf, consente una visualizzazione ottimale su lettori e-reader, tablet, smartphone, iphone, desktop, Android, Apple e Kindle Fire.
- **Software di simulazione:** un vastissimo database di quesiti a risposta multipla per effettuare esercitazioni sull'**intero programma** o su **argomenti specifici**.

L'accesso ai contenuti digitali sarà consentito per **18 mesi**.

Biologia Molecolare

II Edizione

G. Capranico
E. Martegani
G. Musci
G. Raugei
T. Russo
N. Zambrano
V. Zappavigna



G. Capranico • E. Martegani • G. Musci • G. Raugei • T. Russo • N. Zambrano • V. Zappavigna
BIOLOGIA MOLECOLARE II Edizione
Copyright © 2021, 2016, Edises Edizioni S.r.l. – Napoli

9 8 7 6 5 4 3 2 1 0
2025 2024 2023 2022 2021

Le cifre sulla destra indicano il numero e l'anno dell'ultima ristampa effettuata

A norma di legge è vietata la riproduzione, anche parziale, del presente volume o di parte di esso con qualsiasi mezzo.

L'Editore

L'Editore ha effettuato quanto in suo potere per richiedere il permesso di riproduzione del materiale di cui non è titolare del copyright e resta comunque a disposizione di tutti gli eventuali aventi diritto.

Fotocomposizione:
Fotocomposizione TPM S.a.s. - Città di Castello (PG)

Stampato presso la:
Tipolitografia Sograte S.r.l. – Zona Ind. Regnano - Città di Castello (PG)

per conto della
Edises Edizioni S.r.l. - Piazza Dante, 89 - Napoli

www.edisesuniversita.it

assistenza.edises.it

ISBN 978 88 3623 0150

I curatori, l'editore e tutti coloro in qualche modo coinvolti nella preparazione o pubblicazione di quest'opera hanno posto il massimo impegno per garantire che le informazioni ivi contenute siano corrette, compatibilmente con le conoscenze disponibili al momento della stampa; essi, tuttavia, non possono essere ritenuti responsabili dei risultati dell'utilizzo di tali informazioni e restano a disposizione per integrare la citazione delle fonti, qualora incompleta o imprecisa. Realizzare un libro è un'operazione complessa e, nonostante la cura e l'attenzione poste dagli autori e da tutti gli addetti coinvolti nella lavorazione dei testi, l'esperienza ci insegna che è praticamente impossibile pubblicare un volume privo di imprecisioni. Saremo grati ai lettori che vorranno inviarci le loro segnalazioni e/o suggerimenti migliorativi sulla piattaforma assistenza.edises.it

[AUTORI]

Giovanni Capranico

Alma Mater Studiorum - Università di Bologna

Enzo Martegani

Università degli Studi di Milano-Bicocca

Giovanni Musci

Università degli Studi del Molise

Giovanni Raugei

Università degli Studi di Firenze

Tommaso Russo

Università degli Studi di Napoli Federico II

Nicola Zambrano

Università degli Studi di Napoli Federico II

Vincenzo Zappavigna

Università degli Studi di Modena e Reggio Emilia

[PREFAZIONE ALLA SECONDA EDIZIONE]

A cinque anni dall'uscita della prima edizione proponiamo oggi questa nuova versione, arricchita in alcune parti. L'impianto teorico è rimasto essenzialmente lo stesso ma i cap. 8 e 13 sono stati ampliati e rimaneggiati in maniera sostanziale.

La differenza più vistosa rispetto alla prima edizione è l'aggiunta di vari paragrafi spiccatamente metodologici, nei quali si descrivono le tecniche di biologia molecolare di più larga applicazione. Abbiamo infatti pensato che sarebbe stato importante, a completamento di una trattazione teorica il più possibile esaustiva, affrontare anche alcuni aspetti più vicini alla pratica di laboratorio, creando un testo che potesse essere usato in uno spettro più ampio di insegnamenti. I paragrafi metodologici sono stati contrassegnati dal simbolo  che troverete sia nell'indice che in corrispondenza dei paragrafi.

La copertina di questa edizione ci mostra le interazioni molecolari che si instaurano tra la endonucleasi di origine batterica Cas9, l'RNA guida e il DNA bersaglio, così come avviene nel processo di *editing* genomico denominato CRISPR/Cas9. Proprio per questi studi, nel 2020, il Premio Nobel per la Chimica è andato alla francese Emmanuelle Charpentier ed alla statunitense Jennifer A. Doudna, le prime due donne a condividere un Premio Nobel in campo scientifico.

Cogliamo qui l'occasione per ringraziare l'Editore e lo staff tecnico che ha curato questa nuova edizione. Il loro apporto è stato, anche questa volta, decisivo per la buona riuscita di questo libro.

Gli Autori

[PREFAZIONE ALLA PRIMA EDIZIONE]

Natura valde simplex est et sibi consona.

La natura è estremamente semplice ed è in armonia con sé stessa.

Noi che ci accostiamo allo studio della Biologia molecolare possiamo considerare questo pensiero di Isaac Newton una sorta di *passepartout*, da portare sempre con noi. Ci serve per non scoraggiarci di fronte alla complessità cui ci troviamo di fronte. Voi studenti, ma anche noi che cerchiamo di aiutarvi nel vostro lavoro, dobbiamo sempre tenere nella mente che, per fare una delle cose più complicate al mondo, la natura sceglie sempre la strada più facile. Lo aveva ben capito un altro gigante del pensiero dell'uomo: Charles Darwin. Visto che tutte le forme di vita che ci circondano sono il frutto di una spietata selezione, durante la quale molte strade sono tentate, ciò che prevale, e quindi sopravvive, è ciò che trova la strada più conveniente per raggiungere l'obiettivo che si prefigge. Inoltre la strada più conveniente è sempre quella relativamente più semplice.

La complessità con cui dobbiamo confrontarci nello studio delle basi molecolari della vita è molto aumentata negli ultimi anni. Abbiamo capito moltissime cose che fino a non molti anni fa erano del tutto oscure e, come sempre succede quando si esplora il non conosciuto, abbiamo scoperto anche cose che nessuno aveva nemmeno immaginato che esistessero. Un grande aiuto in questo lavoro di esplorazione è dovuto allo sviluppo di nuova tecnologia e, in particolare, a uno specifico tipo basato sulla natura stessa. Infatti, a differenza di altri campi di esplorazione della natura, come ad esempio la fisica, in cui è stata necessario inventare tecnologie, nel campo della biologia, abbiamo inventato ben poco. Non abbiamo fatto altro che usare ciò che già aveva costruito la natura. Stiamo ovviamente parlando degli strumenti biotecnologici che nel corso degli ultimi cinquant'anni hanno permeato tutta la ricerca in campo biologico. L'obiettivo, considerata la premessa che abbiamo fatto, è quello di semplificare, creare cioè modelli semplici in cui siano note le variabili che introduciamo. A queste biotecnologie si sono aggiunte e, spesso, con esse si sono integrate anche strumentazioni molto sofisticate che oggi permettono la raccolta di enormi quantità di dati. Non vi sorprendete: con questi strumenti, e con un enorme contributo dato dalla bioinformatica, si sono raccolti talmente tanti dati, che facciamo fatica a gestirli. Consideriamo ad esempio il genoma umano (e quello di molti altri viventi). Lo abbiamo conquistato tutto, ma tra averlo scritto in un file di un computer e comprenderne il significato in ogni suo punto la strada sarà ancora lunga.

Perché tanta fatica? La prima molla è ovviamente la curiosità, che è una caratteristica della nostra specie, ma ci sono anche altri motivi altrettanto importanti. Il primo è che la conoscenza delle molecole della vita, del loro funzionamento e delle loro interazioni è cruciale per poter curare le malattie: per diagnosticarle in tempo, quando ancora sono curabili con gli strumenti che abbiamo a disposizione; per trovare nuove cure, che siano frutto non del caso ma della dettagliata conoscenza dei bersagli molecolari da colpire e, grande ambizione della medicina di ogni tempo, per prevenire la comparsa di malattie. Studiando la Biologia molecolare e leggendo questo libro, scoprirete che nel corso degli ultimi anni si è dovuto dare un nome a tante nuove molecole e che molti di questi nomi sono derivati dalla malattia in cui queste molecole sono coinvolte o di cui in qualche modo sono responsabili. Ancora una volta, non vi sorprendete. La necessità di capire le malattie per provare a combatterle si è da tempo affiancata alla curiosità, che aveva per secoli guidato lo studio della biologia e, in molti casi, la necessità sta prevalendo sulla curiosità. Molti sono convinti che questo non sia un bene. La medicina cerca scorciatoie, per risolvere nel più breve tempo possibile il problema che si pone. Ma le scorciatoie, nella strada della conoscenza, a volte fanno perdere più tempo.

Intanto, noi che stiamo per cominciare lo studio della Biologia molecolare non abbiamo fretta e non cercheremo scorciatoie. Questo libro, che vi aiuterà a scoprire le meraviglie che abbiamo dentro di noi e nei viventi che ci circondano, segue un filo logico ben preciso. In primo luogo, studieremo come sono fatte le molecole che governano il funzionamento delle cellule. Scopriremo che la forma è sostanza, nel senso che è nella forma stessa che va cercata la funzione. Come per molti degli oggetti che abbiamo intorno a noi: una porta fa la porta perché ha quella forma. Se invece di essere rettangolare fosse tonda, non sarebbe una porta. Inoltre non basta essere

[VIII] Prefazione

rettangolare per fare la porta. Deve essere un rettangolo che si adatta perfettamente agli stipiti, altrimenti non può funzionare. Piccole differenze nella forma di una biomolecola le fanno perdere la funzione o, a volte, gliene fanno acquisire una nuova. Scopriremo poi che il cambiamento di forma è spesso cruciale per la funzione: se la porta non ruotasse sui suoi cardini non sarebbe una porta, ma una parte del muro fatta chi sa perché di legno invece che di mattoni. Infine, affronteremo il problema dell'interazione tra grandi molecole, per esempio tra DNA e proteine, che ci permetterà di farci un'idea generale dei principi che governano l'affollato ambiente intracellulare.

Una volta appresi i principi generali ci addentreremo nel coinvolgimento di specifiche molecole nelle funzioni delle cellule. Cominciando dall'anatomia dei genomi, via via esploreremo i meccanismi molecolari della trascrizione, della sua regolazione, della maturazione dei trascritti, e della loro traduzione e poi i meccanismi della replicazione e manutenzione dell'informazione genetica e della loro regolazione. Il viaggio attraverso le molecole della vita si concluderà con alcuni esempi di differenziamento, di utilizzo specifico e finemente regolato di parti cruciali dell'informazione genetica.

Questa prima edizione è tutta scritta da docenti di Biologia molecolare, con il cruciale contributo dello staff dell'Editore, cui va il nostro ringraziamento. La prossima sarà certamente migliore, perché arricchita dagli stimoli e dai suggerimenti che verranno dai voi lettori, cui questo lavoro è dedicato.

Gli Autori

ICONOGRAFIA

Numerose figure di questo testo, contenenti strutture, sono state elaborate e rappresentate mediante l'applicazione Chimera (<http://www.cgl.ucsf.edu/chimera>) o Cn3D (<http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>).

[INDICE GENERALE]

Autori	III
Prefazione	V
Prefazione alla precedente edizione	VII

CAPITOLO [1] La biologia molecolare nelle scienze della vita: un excursus storico

[1.1] La biologia molecolare e la comprensione dei sistemi viventi	2
[1.2] L'unità elementare della vita è la cellula	3
[1.3] Esistono diversi tipi di organizzazione cellulare	4
[1.4] Unità nella diversità	5
[1.5] Breve storia della biologia molecolare	6
[1.5.1] Le origini: 1930-1960	6
[1.5.2] Decifrazione del codice genetico e meccanismo della sintesi proteica: 1960-1970	11
[1.5.3] DNA ricombinante e biotecnologie	13
[1.5.4] Dagli anni '80 del secolo scorso alle frontiere della biologia	14

CAPITOLO [2] Struttura di acidi nucleici e proteine

[2.1] Dogma centrale della biologia molecolare	16
[2.2] Struttura del DNA	16
[2.2.1] I nucleotidi sono costituiti da zucchero, fosfato e basi azotate	16
[2.2.2] Il legame fosfodiesterico collega i nucleotidi in modo direzionale	19
[2.2.3] La forma principale del DNA è una doppia elica con filamenti complementari	21
[2.2.4] Esistono diverse isoforme della doppia elica: DNA-B, A e Z	26
[2.2.5] Il DNA è una molecola flessibile	27
[2.2.6] Il DNA può curvarsi	29
[2.2.7] Altre strutture elicoidali del DNA	29
[2.3] Proprietà chimico-fisiche del DNA	30
[2.3.1] Il DNA può denaturarsi e rinaturarsi	30
Inserto 2.1 Valore $Cot_{1/2}$ e cinematiche di rinaturazione del DNA	33
[2.3.2] Il DNA è stabilizzato da un insieme di forze deboli diverse	34
[2.4] Topologia del DNA	35
[2.4.1] Il superavvolgimento del DNA è descritto dai suoi parametri topologici	35
[2.4.2] Il DNA cellulare è di norma superavvolto negativamente	38
[2.4.3] Le topoisomerasi modificano la topologia del DNA	38
[2.5] Struttura dell'RNA	42
[2.5.1] Chimica del ribosio e stabilità dell'RNA	42
[2.5.2] L'RNA assume strutture secondarie e terziarie	42
[2.6] Struttura delle proteine	47

[X] Indice generale

[2.6.1]	Le unità di base delle proteine sono gli amminoacidi	47	
[2.6.2]	La struttura primaria è costituita da catene di amminoacidi connessi da legami peptidici	50	
[2.6.3]	Le strutture secondarie guidano il <i>folding</i> della catena polipeptidica	52	
[2.6.4]	Le strutture terziaria e quaternaria definiscono l'aspetto finale di una proteina	54	
[2.6.5]	La denaturazione delle proteine è essenzialmente irreversibile	56	
Inserto 2.2	Metodi per la determinazione della struttura tridimensionale delle macromolecole	T	56
[2.7]	Tecniche di ibridazione degli acidi nucleici	T	58
[2.7.1]	Southern blotting	58	
[2.7.2]	Northern blotting	59	
[2.7.3]	Ibridazione <i>in situ</i>	60	
[2.7.4]	Microarray	61	

CAPITOLO [3] Interazioni tra macromolecole e struttura della cromatina

[3.1]	Interazioni specifiche DNA-proteine e RNA-proteine	64	
[3.1.1]	Importanza dei legami deboli nelle interazioni tra macromolecole	65	
[3.1.2]	Le proteine utilizzano specifici motivi strutturali per legarsi al DNA	65	
[3.2]	Assemblaggio del DNA in cromosomi: struttura della cromatina	69	
[3.2.1]	Il nucleosoma costituisce il primo livello nel compattamento del DNA	70	
[3.2.2]	Fibre di ordine superiore	74	
[3.2.3]	Gli istoni nei nucleosomi possono subire numerose modificazioni post-traduzionali	77	
[3.2.4]	I nucleosomi sono strutture dinamiche che possono allentarsi o scivolare	82	
[3.2.5]	Le varianti istoniche contribuiscono alla dinamica dei nucleosomi	84	
[3.3]	Metodi di studio delle interazioni DNA-proteine e della struttura cromatinica	T	86
[3.3.1]	<i>Electrophoretic Mobility Shift Assay</i> (EMSA)	86	
[3.3.2]	<i>DNase footprinting</i>	87	
[3.3.3]	Digestione della cromatina con nucleasi e mappatura dei nucleosomi	88	
[3.3.4]	<i>Chromatin Immuno Precipitation</i> (ChIP)	89	

CAPITOLO [4] Geni e genomi: organizzazione e funzione

[4.1]	La funzione biologica: il programma della vita	94
[4.2]	Sequenziamento e annotazione dei genomi	95
[4.3]	Anatomia del genoma procariotico	100
[4.3.1]	Struttura dei genomi procariotici	100
[4.3.2]	Organizzazione genica nei procarioti	103
[4.3.3]	Metagenomica	105
[4.4]	Anatomia del genoma eucariotico	106
[4.4.1]	Dimensioni, geni e origini di replicazione	106
[4.4.2]	Organizzazione e attività geniche negli eucarioti	111
[4.4.3]	Sequenze ripetute: struttura e funzioni	114
[4.4.4]	Genoma degli organelli: mitocondrio e cloroplasto	125
[4.5]	Variabilità del genoma	129
[4.5.1]	Polimorfismi e variabilità fenotipica	129
[4.5.2]	Metilazione del DNA e trascrittomico	133

[4.5.3] Variabilità genomica e differenziamento cellulare	134
[4.6] Tecniche di base del DNA ricombinante T	135
[4.6.1] I plasmidi	135
[4.6.2] Gli enzimi di restrizione	137
[4.6.3] La separazione elettroforetica del DNA	138
[4.6.4] La ligazione	140
[4.6.5] La trasformazione	143
[4.7] La PCR (Polymerase Chain Reaction) T	144
[4.7.1] La tecnica della PCR	144
[4.7.2] La qRT-PCR	148
[4.7.3] La mutagenesi sito-diretta mediante PCR	153
[4.8] Tecniche di sequenziamento degli acidi nucleici T	154
[4.8.1] Il metodo Sanger per sequenziare il DNA	155
[4.8.2] Il pirosequenziamento del DNA	158
[4.8.3] La tecnologia Illumina-Solexa	162
[4.8.4] Il sequenziamento massivo parallelo dei trascritti o RNA-seq	164
[4.8.5] ChIP-seq: <i>Chromatin immunoprecipitation</i> (ChIP) abbinata a sequenziamento massivo parallelo del DNA	165
[4.8.6] DNase-seq e FAIRE-seq	166

CAPITOLO [5] Trascrizione: i meccanismi molecolari

[5.1] Introduzione	170
[5.2] La reazione di sintesi dell'RNA e le RNA polimerasi	170
[5.2.1] Le RNA polimerasi sintetizzano poliribonucleotidi <i>ex novo</i> usando come stampo il DNA	171
[5.2.2] La trascrizione è un processo a più tappe	174
[5.2.3] La struttura delle RNA polimerasi ne chiarisce le funzioni	176
[5.2.4] RNA polimerasi batteriche	179
[5.2.5] Le RNA polimerasi degli eucarioti	180
[5.2.6] L'RNA polimerasi II	181
[5.2.7] RNA polimerasi insolite: l'RNA polimerasi del batteriofago T7, l'RNA polimerasi mitocondriale e le RNA polimerasi RNA-dipendenti dei virus ad RNA	186
[5.3] L'inizio della trascrizione richiede interazioni fisiche e funzionali tra acidi nucleici e proteine	189
[5.3.1] I promotori procariotici e il loro riconoscimento da parte dell'oloenzima dell'RNA polimerasi	190
[5.3.2] La subunità σ_{70} e l'isomerizzazione tra complesso chiuso e complesso aperto	191
[5.3.3] I promotori eucariotici di classe I e di classe III e i rispettivi fattori generali	193
[5.3.4] Il promotore dei geni di classe I è responsabile della trascrizione dei geni dell'rDNA, codificanti per il precursore 47S degli RNA ribosomali	193
[5.3.5] L'RNA polimerasi III trascrive geni caratterizzati da 3 diverse tipologie di promotori	195
[5.3.6] Il complesso repertorio di elementi <i>in cis</i> sui promotori di classe II	197
[5.3.7] Il fattore generale TF _{II} D e il suo legame al promotore basale	198
[5.3.8] I fattori generali dirigono l'assemblaggio del complesso di pre-inizio sul promotore basale di classe II	200
[5.4] La fase di allungamento della trascrizione	203
[5.4.1] Il complicato passaggio dall'inizio all'allungamento produttivo della trascrizione di classe II	203
[5.4.2] Diversi meccanismi assicurano la sintesi processiva e fedele dei trascritti in procarioti ed eucarioti	204
[5.4.3] L'allungamento trascrizionale negli eucarioti comporta la parziale rimozione e il riposizionamento dei nucleosomi	207

[XII] Indice generale

[5.5] La terminazione della trascrizione	208
[5.5.1] Strutture a forcina sull'RNA, seguite da appaiamenti labili tra stampo e trascritto, determinano efficiente terminazione in procarioti	208
[5.5.2] La terminazione in procarioti può richiedere fattori proteici reclutati sull'RNA	209
[5.5.3] Le RNA polimerasi degli eucarioti possiedono meccanismi caratteristici per la terminazione della trascrizione	209

CAPITOLO [6] Trascrizione: le regolazioni

[6.1] Introduzione	212
[6.2] Regolazione trascrizionale nei procarioti	212
[6.2.1] Il ruolo dei promotori e del repertorio di fattori σ nella regolazione trascrizionale dei procarioti	213
[6.2.2] Modelli di regolazione trascrizionale nei procarioti: controlli negativo e positivo e rispettivi attuatori	216
[6.2.3] Il modello dell'operone lac di <i>E. coli</i> : cooperazione tra controlli negativo e positivo	216
[6.3] Regolazione trascrizionale negli eucarioti	222
[6.3.1] Elementi promotore a monte ed enhancer: il codice di DNA nei meccanismi della regolazione trascrizionale	223
[6.3.2] I co-attivatori possono regolare la frequenza di inizio della trascrizione	225
[6.3.3] Le proteine attivatrici possono subire modifiche post-traduzionali che ne regolano la funzione	226
[6.3.4] Analisi della trascrizione: trasfezione di vettori con attività reporter in cellule di mammiferi	227
[6.3.5] Il complesso del mediatore e le sue funzioni nella trascrizione e nella sua regolazione	229
[6.3.6] Sistema GAL di lievito: un semplice sistema inducibile	232
Inserto 6.1 Il saggio del doppio ibrido T	237
[6.4] Meccanismi epigenetici nella regolazione trascrizionale	238
[6.4.1] La trascrizione negli eucarioti richiede uno stato cromatinico adeguato	239
[6.4.2] L'attivazione trascrizionale del gene IFNB: integrazione tra il codice di DNA e il codice istonico	241
[6.4.3] Un esempio di attivazione e repressione trascrizionale in mammiferi: il recettore degli ormoni tiroidei	243
[6.4.4] L'analisi del fenotipo variegato dell'occhio di <i>Drosophila</i> : l'estensione di territori eterocromatinici comporta il silenziamento dell'espressione genica	244
[6.4.5] Gli isolatori delimitano territori cromatinici	248
Inserto 6.2 L'esperimento 3C di Cattura della Conformazione Cromosomica e l'evoluzione delle tecnologie-C T	251
[6.4.6] La metilazione del DNA negli eucarioti è un meccanismo associato alla repressione della trascrizione	252

CAPITOLO [7] RNA cellulari e maturazione dei trascritti

[7.1] Introduzione	256
[7.2] RNA messaggeri	256
[7.2.1] Tre modificazioni dell'RNA messaggero eucariotico sono legate alla trascrizione dipendente dalla RNA polimerasi II	257
[7.2.2] Capping del terminale 5'	257
[7.2.3] La generazione dell'estremità 3' è molto diversa tra procarioti ed eucarioti	260
[7.2.4] I geni per gli istori trascrivono mRNA privi di coda poli(A)	262
[7.3] Splicing	263
[7.3.1] Caratteristiche generali dello splicing	263
[7.3.2] Splicing: un evento sorprendente	263
[7.3.3] Lo splicing è guidato da specifiche sequenze presenti sul pre-mRNA	265

[7.3.4] Lo splicing si attua attraverso due reazioni di transesterificazione	266
[7.3.5] Spliceosoma	267
[7.3.6] Splicing mediato dallo spliceosoma	269
[7.3.7] Lo spliceosoma U12 è molto meno comune	271
[7.3.8] Introni <i>self-splicing</i>	271
[7.3.9] Confronto tra i meccanismi di splicing degli introni di gruppo II e quelli dipendenti dallo spliceosoma	272
[7.3.10] Introni <i>self-splicing</i> di gruppo I	274
[7.3.11] Introni precoci o introni tardivi?	274
[7.3.12] A cosa servono gli introni?	275
[7.3.13] Gli introni dei tRNA eucariotici vengono eliminati con un diverso meccanismo	276
[7.3.14] <i>Trans-splicing</i>	277
[7.3.15] Trasporto dell'mRNA nel citoplasma	279
[7.4] Splicing alternativo	280
[7.4.1] Lo splicing alternativo si attua attraverso varie modalità	280
[7.4.2] Lo splicing alternativo è regolato da sequenze in <i>cis</i> e da fattori proteici in <i>trans</i>	282
[7.4.3] Un esempio di splicing alternativo: il gene <i>DSX</i> di <i>Drosophila</i>	284
[7.4.4] Lo splicing alternativo è spesso tessuto-specifico	285
[7.4.5] Gene <i>DSCAM</i> di <i>Drosophila</i> : l'apoteosi dello splicing alternativo	285
[7.4.6] Modalità meno comuni di splicing alternativo	286
[7.5] I difetti dello splicing sono la causa di alcune patologie umane	288
[7.5.1] Molte malattie genetiche sono dovute a difetti nel meccanismo di splicing	288
[7.5.2] Alcune β-talassemie sono causate da mutazioni nelle sequenze che guidano lo splicing	288
[7.5.3] Molte varianti della fibrosi cistica sono causate da splicing aberrante	289
[7.5.4] Gene <i>MAPT</i> : difetti nel bilanciamento tra isoforme generate da splicing alternativo	290
[7.5.5] Lo splicing aberrante dell'oncosoppressore <i>APC</i> è associato al tumore del colon	291
[7.5.6] Retinite pigmentosa: un raro caso di mutazione dello splicing in <i>trans</i>	291
[7.6] Editing dell'mRNA	292
[7.6.1] L'enzima APOBEC1 catalizza l' <i>editing</i> C-U dell'RNA per l'apolipoproteina	293
[7.6.2] L' <i>editing</i> A-I è un meccanismo importante nell'uomo	294
[7.6.3] <i>Editing</i> inserzionale nel <i>Trypanosoma</i>	294
[7.7] RNA non codificant	294
[7.7.1] Maturazione degli RNA ribosomal	295
[7.7.2] Gli RNA ribosomal vengono modificati con l'intervento degli snoRNA	297
[7.7.3] Maturazione degli RNA transfer	298
[7.7.4] Maturazione dei microRNA	300
[7.7.5] <i>Editing</i> dei microRNA	304

CAPITOLO [8] Traduzione

[8.1] Schema generale della biosintesi delle proteine	306
[8.2] Il codice genetico	307
[8.3] I tRNA e le amminoacil-tRNA sintetasi	312
[8.4] I ribosomi	317
[8.5] L'assemblaggio del complesso traduzionale	320
Inserto 8.1 Proteine G	322

[XIV] Indice generale

[8.6] Peptidil-transferasi e allungamento	325
[8.7] La terminazione della traduzione	327
[8.8] La regolazione della traduzione	331
[8.8.1] Esempi di regolazione nei procarioti	331
Inserto 8.2 Gli inibitori della traduzione e il loro uso come antibiotici	332
[8.8.2] mTORC1	335
[8.8.3] La biogenesi dei ribosomi	335
[8.8.4] miRNA e regolazione della traduzione	337
[8.9] Western blot e immunoprecipitazione 	338
[8.9.1] Western blot	338
[8.9.2] Immunoprecipitazione	341
[8.10] Il silenziamento genico post-trascrizionale, i siRNA e gli antagomiR 	344
[8.11] Le proteine ricombinanti 	346
[8.11.1] Espressione di proteine ricombinanti in procarioti	346
[8.11.2] Produzione e purificazione di proteine in lievito	350
Inserto 8.3 Vettori di lievito 	352
[8.11.3] Proteine ricombinanti in cellule di mammifero	355

CAPITOLO [9] Eventi post-traduzionali

[9.1] Le proteine possono essere modificate dopo la traduzione	359
[9.1.1] Glicosilazione	360
[9.1.2] Fosforilazione	365
[9.1.3] Acetilazione e metilazione	368
[9.1.4] Modificazioni ossidative degli amminoacidi solforati	369
[9.1.5] La ADP-ribosilazione regola molti processi cellulari e l'azione di alcune tossine	372
[9.1.6] La lipidazione ancora reversibilmente le proteine alle membrane	374
[9.1.7] La proteolisi controllata regola la formazione di proteine mature	375
[9.2] Il destino delle proteine è la degradazione	378
[9.2.1] Ogni proteina ha una sua emivita caratteristica	379
[9.2.2] Il sistema ubiquitina-proteasoma degrada le proteine diventate inutili o dannose	379
[9.2.3] I lisosomi degradano le proteine e altre molecole complesse	383

CAPITOLO [10] Replicazione: i meccanismi molecolari

[10.1] Introduzione	386
[10.2] Funzione biologica	386
[10.2.1] La replicazione del DNA è un processo molto efficiente	386
[10.2.2] La fedeltà di copia del DNA è estremamente alta e dipende dalla capacità di correggere gli errori	386
[10.2.3] Il meccanismo di sintesi si basa sulla complementarità delle eliche del DNA	387
[10.2.4] La replicazione del DNA ha una modalità semiconservativa	387
[10.2.5] Replicazione e trascrizione sono chimicamente simili, ma non sono funzionalmente correlate	389
[10.3] DNA polimerasi	389
[10.3.1] Classificazione	389
[10.3.2] Struttura e domini catalitici	391

[10.3.3] Substrati e formazione del legame fosfodiesterico	392
[10.3.4] <i>Proofreading</i>	393
[10.3.5] Perché la sintesi del DNA avviene in direzione 5'→3'?	394
[10.3.6] La tautomeria delle basi del DNA causa errori di copia	395
[10.3.7] Perché è necessario un innesco?	396
[10.3.8] Fenotipi mutatori e antimutatori	396
[10.3.9] Ribonucleotidi e deossiribonucleotidi: una scelta difficile	397
[10.3.10] Processività delle DNA polimerasi	398
[10.3.11] La <i>sliding clamp</i> aumenta la processività	399
[10.4] Forcella di replicazione	400
[10.4.1] Azione della DNA elicasi	400
[10.4.2] Sintesi continua e sintesi discontinua	401
[10.4.3] Gestione degli inneschi di RNA in <i>E. coli</i>	402
[10.4.4] Gestione degli inneschi di RNA negli eucarioti	404
[10.4.5] Proteine SSB	405
[10.4.6] Topoisomerasi	407
[10.4.7] DNA polimerasi III e sua azione alla forcella di replicazione	408
[10.4.8] Replisoma di <i>E. coli</i> : un complesso guidato dall'elicasi	410
[10.4.9] Negli eucarioti, la sintesi del DNA alla forcella di replicazione è conosciuta meno in dettaglio rispetto a quella nei procarioti	412
[10.4.10] La replicazione del DNA mitocondriale presenta alcune caratteristiche peculiari	412
[10.5] Inizio della replicazione: un processo simile in tutti gli organismi	413
[10.5.1] L'inizio della replicazione in <i>E. coli</i> è indotto dalle proteine DnaA	413
[10.5.2] Assemblaggio del replisoma in <i>E. coli</i>	416
[10.5.3] Inizio della replicazione negli eucarioti	416
[10.6] La terminazione della replicazione segue strategie diverse in procarioti ed eucarioti	419
[10.6.1] Terminazione della replicazione in <i>E. coli</i>	419
[10.6.2] La terminazione della replicazione negli eucarioti presenta un problema	420
[10.6.3] Le telomerasi contrastano il fenomeno dell'accorciamento dei telomeri	422
[10.6.4] La telomerasi è una trascrittasi inversa legata a un RNA stampo	422
[10.6.5] I telomeri presentano sempre un terminale 3' sporgente legato a proteine	422
[10.6.6] Regolazione della lunghezza dei telomeri nel lievito	424
[10.6.7] Regolazione della lunghezza dei telomeri nell'uomo	424
[10.6.8] Negli organismi pluricellulari, le telomerasi non sono sempre attive	425
[10.6.9] Telomerasi e tumori	426
[10.7] DNA polimerasi specializzate	427
[10.7.1] Le polimerasi TLS attuano una sintesi di emergenza	427
[10.7.2] Trascrittasi inverse: una eccezione abbastanza comune	428

CAPITOLO [11] Replicazione del DNA: le regolazioni

[11.1] Regolazione dell'inizio della replicazione in <i>Escherichia coli</i>	430
[11.1.1] La replicazione del DNA in <i>E. coli</i> è controllata attraverso più meccanismi paralleli	430
[11.2] Regolazione dell'inizio della replicazione negli eucarioti	432
[11.2.1] Regolazione della replicazione nel lievito	433

[XVI] Indice generale

[11.2.2] Regolazione dell'inizio della replicazione negli eucarioti multicellulari	436
[11.2.3] Regolazione dell'inizio della replicazione durante lo sviluppo e il differenziamento cellulare	441

CAPITOLO [12] Plasticità dei genomi: danno al DNA, riparazione e ricombinazione

[12.1] La funzione biologica	445
[12.2] I meccanismi della riparazione del danno al DNA	446
[12.2.1] Tipi di danno al DNA	446
Inserto 12.1 I raggi UV	451
[12.2.2] Meccanismi di riparazione diretta	453
[12.2.3] Meccanismi di riparazione per escissione delle basi (BER)	453
[12.2.4] Meccanismi di riparazione per escissione di nucleotidi (NER)	456
[12.2.5] Meccanismi di riparazione dei <i>mismatch</i> (MMR)	461
[12.2.6] Meccanismi di tolleranza al danno	464
[12.2.7] Meccanismi di riparazione delle rotture singole del DNA (SSB)	468
[12.2.8] Meccanismi di riparazione delle rotture doppie del DNA (DSB)	469
[12.2.9] Risposta cellulare al danno al DNA e meccanismi di controllo del ciclo cellulare	472
[12.3] Ricombinazione del genoma: meccanismi e funzioni	484
[12.3.1] Ricombinazione omologa come riparazione di una rottura doppia del DNA	485
Inserto 12.2 Il primo modello di Holliday junction	486
[12.3.2] Apparati enzimatici della ricombinazione omologa nelle cellule batteriche	489
[12.3.3] Apparati enzimatici della ricombinazione omologa nelle cellule eucariotiche	491
[12.3.4] Ricombinazione sito-specifica	496
[12.3.5] Ricombinazione e trasposoni	499
[12.4] Topi knock-out: clonaggio per ricombinazione omologa 	501
[12.5] Il sistema di editing genomico CRISPR/Cas9 	504

CAPITOLO [13] Trasduzione del segnale: sviluppo e differenziamento

[13.1] Introduzione	508
[13.2] Meccanismi molecolari	508
[13.2.1] Recettori nucleari	509
[13.2.2] Recettori accoppiati a proteine G	509
[13.2.3] Recettori tirosina chinasici	511
[13.2.4] RAS/MAP chinasi	512
[13.2.5] Fosfoinositidi e fosfolipasi	513
[13.2.6] Recettori delle citochine e via JAK/STAT	515
[13.2.7] Recettori di TGF β /BMP e via di attivazione delle SMAD	515
[13.2.8] Recettori di Wnt e via della β -catenina	518
[13.2.9] Hedgehog-Patched	519
[13.2.10] Notch, il recettore di Delta/Jagged	521
[13.3] Meccanismi molecolari della coniugazione nel lievito gemmante <i>Saccharomyces cerevisiae</i>	523
[13.3.1] Il lievito, un semplice eucariote unicellulare	523
[13.3.2] I geni MAT e lo switch del mating type	525

[13.3.3] Coniugazione: un esempio integrato di trasduzione del segnale	526
[13.3.4] Sporulazione: come cambia il profilo di espressione genica	527
[13.4] Definizione dell'identità cellulare nelle fasi iniziali dello sviluppo in <i>Drosophila</i>	530
[13.4.1] Determinazione degli assi embrionali in <i>D. melanogaster</i>	530
[13.4.2] Sistema anteriore: bicoid, una proteina con due meccanismi di azione	531
[13.4.3] Sistema posteriore: nanos, un regolatore della traduzione	533
[13.4.4] Quattro proteine suddividono l'embrione in due territori, quello anteriore e quello posteriore	533
[13.4.5] Geni gap: l'embrione inizia a suddividersi in segmenti	535
[13.4.6] Geni pair-rule o a "modulo pari": segmenti e para-segmenti	537
[13.4.7] Geni della polarità segmentale: la definizione dei compartimenti cellulari	539
[13.4.8] Geni omeotici: i segmenti acquisiscono la loro identità morfologica	542
[13.5] Definizione dell'identità cellulare nelle fasi iniziali dello sviluppo nel topo	546
[13.5.1] Primo evento differenziativo: trofoectoderma o <i>inner cell mass</i>	547
[13.5.2] Secondo evento differenziativo: endoderma primitivo o epiblasto	548
[13.5.3] Il problema della definizione degli assi nell'embrione di topo	549
Bibliografia e sitografia	553
Indice analitico	559

Capitolo [3]

Interazioni tra macromolecole e struttura della cromatina

SOMMARIO

- 3.1 Interazioni specifiche DNA-proteine e RNA-proteine**
- 3.2 Assemblaggio del DNA in cromosomi: struttura della cromatina**
- 3.3 Metodi di studio delle interazioni DNA-proteine e della struttura cromatinica**

T

[3.1] INTERAZIONI SPECIFICHE DNA-PROTEINE E RNA-PROTEINE

L'esistenza di interazioni tra DNA e proteine è nota dalla fine del XIX secolo, anche se le evidenze sperimentali derivavano esclusivamente da indagini di microscopia ottica. Da allora, sono stati messi a punto numerosi metodi per lo studio del legame del DNA e dell'RNA con le proteine, grazie ai quali è stato chiarito senza ombra di dubbio il ruolo che l'associazione tra polipeptidi e acidi nucleici ha in tutti i processi cellulari, dalla **duplicazione** del DNA alla sua **trascrizione, ricombinazione e riparazione**, fino alla **traduzione**, e più in generale nel **controllo dell'espressione genica** (**Figura 3.1**).

La regolazione dell'espressione genica, come vedremo più avanti nel testo, si fonda principalmente sulla capacità di proteine specifiche di attaccarsi alla doppia elica del DNA (e al singolo o, più raramente, al doppio filamento dell'RNA), modulandone l'utilizzo in base alle necessità metaboliche della cellula e/o dell'intero organismo. Possiamo senza dubbio affermare che lo studio dei fenomeni associati a interazioni DNA/RNA-proteine costituisce il corpo principale della ricerca biomolecolare, in rapidissima espansione soprattutto nella prospettiva della comprensione delle basi molecolari delle malattie.

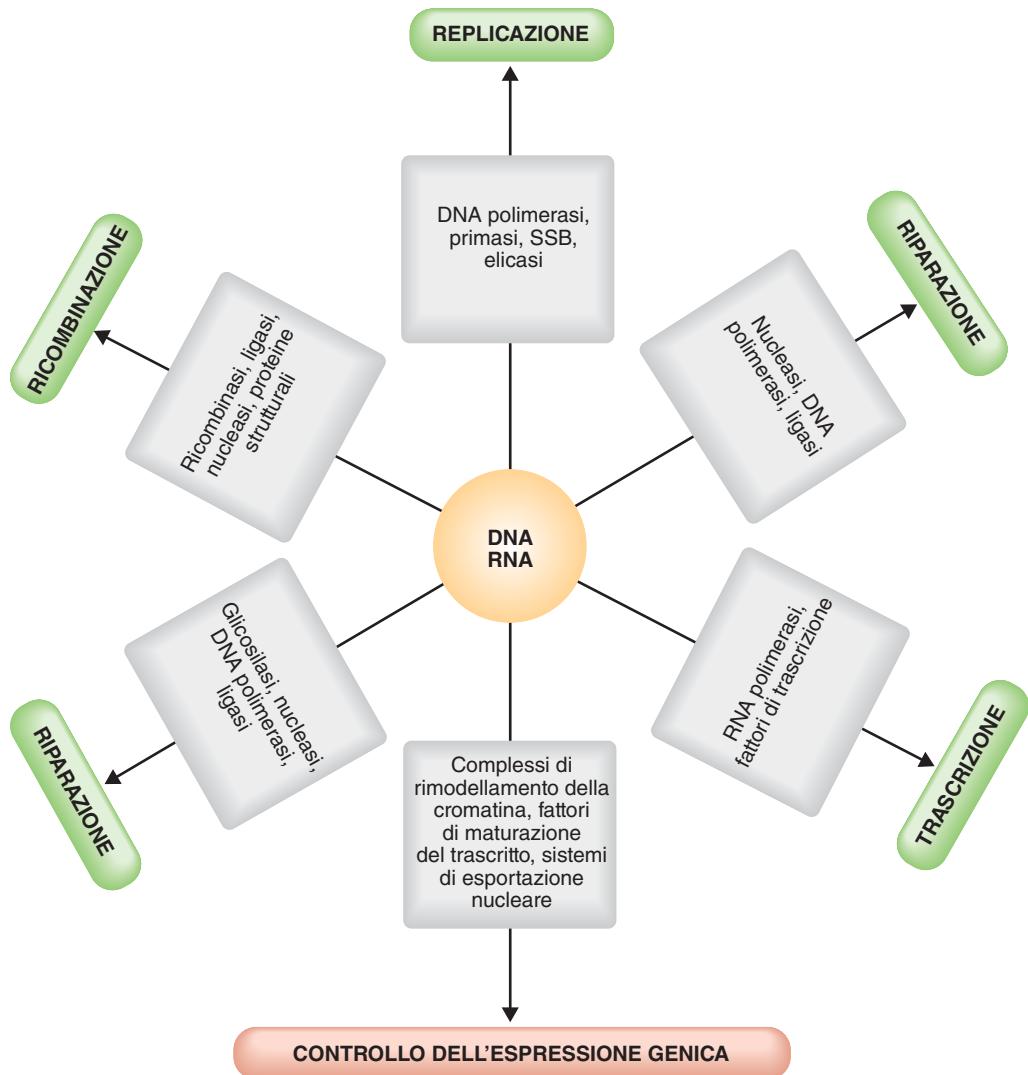


FIGURA 3.1 ▲ Le interazioni tra proteine e acidi nucleici sono alla base di tutti i processi cellulari e, in generale, del controllo dell'espressione genica. Lo schema riporta in modo semplificato i principali sistemi proteici coinvolti nei diversi processi.

[3.1.1] Importanza dei legami deboli nelle interazioni tra macromolecole

Abbiamo visto finora come i **legami deboli** tengano insieme la struttura del DNA, dell'RNA e delle proteine. Non sorprende, pertanto, il fatto che le interazioni deboli siano anche responsabili delle interazioni che si stabiliscono tra DNA e proteine e tra RNA e proteine (nonché tra proteine e proteine).

La principale interazione che si stabilisce tra proteine e DNA/RNA è di carattere ionico. Le proprietà polianioniche degli acidi nucleici, infatti, favoriscono il legame con proteine con carica positiva. L'esempio più importante è certamente quello degli **istoni**, piccole proteine ricche di lisina e arginina, e quindi con carica netta positiva a pH fisiologico, che costituiscono il nucleosoma. Ne parleremo in dettaglio più avanti nel capitolo; tuttavia, è opportuno fin da ora notare che il legame degli istoni al DNA rientra nella categoria delle interazioni in gran parte indipendenti dalla sequenza del DNA. Più spesso, le proteine utilizzano tutta la gamma disponibile dei legami deboli (legami dipolari come il legame idrogeno, interazioni elettrostatiche come i ponti salini, effetti entropici quali le interazioni idrofobiche, forze di dispersione come lo *stacking*) per "leggere" la sequenza del DNA o dell'RNA e selezionare quella propria per il legame.

Le associazioni fisiche tra macromolecole diverse possono essere complesse e possono seguire una vera e propria gerarchia delle interazioni. Le proteine, infatti, possono legarsi a un acido nucleico sia direttamente sia in modo mediato da altre proteine. In altri termini, spesso le proteine interagiscono con il DNA sotto forma di multimeri proteici, formati da entità proteiche separate tra loro e assemblate per lo specifico bisogno. Un caso particolare, molto comune, è rappresentato dalla dimerizzazione o dall'oligomerizzazione di un'unica proteina. La dimerizzazione costituisce una strategia fenomenale che le proteine adottano per poter riconoscere in modo specifico il DNA. In genere, infatti, le proteine che si legano al DNA lo fanno in due passaggi: un primo riconoscimento aspecifico, fondato essenzialmente su interazioni elettrostatiche, con affinità bassa (K_d dell'ordine di 1-2 mM), e un secondo riconoscimento, conseguente alla capacità della proteina di scorrere lungo il duplex, in cui la proteina interagisce con alta affinità (K_d dell'ordine di 1 pM) con una specifica sequenza di basi. Il problema è che, avendo a disposizione solo quattro tipi di nucleotidi, una sequenza che ricorra solo una o poche volte in un genoma deve essere lunga almeno 12-15 coppie di basi, cioè più di un giro di doppia elica (10 coppie di basi). Poiché il sito primario di riconoscimento è costituito dal solco maggiore, la soluzione topologicamente più semplice è che i due monomeri di un dimero proteico riconoscano due solchi maggiori adiacenti. Nel Capitolo 6 descriveremo due esempi specifici di questo tipo di interazioni e scopriremo che questa modalità è sfruttata sia da procarioti (si veda il Paragrafo 6.2.3) che da eucarioti (si veda il Paragrafo 6.4.3).

[3.1.2] Le proteine utilizzano specifici motivi strutturali per legarsi al DNA

Per riconoscere e interagire con specifiche sequenze del DNA, le proteine hanno sviluppato nel corso dell'evoluzione molti **motivi** (o **domini strutturali**) diversi. Come abbiamo mostrato nella Figura 2.15, ogni coppia di Watson-Crick, oltre ad avere legami idrogeno specifici tra le due basi, espone al solco maggiore e al solco minore bordi con profili chimici unici, costituiti da combinazioni diverse di donatori e accettori di legami idrogeno e di gruppi metilici. Di conseguenza, affinché una proteina possa riconoscere una sequenza di DNA, non è necessario che essa destabilizzi la doppia elica, potendo leggerne la sequenza dalla chimica dei bordi delle varie coppie di basi. La variabilità dei profili chimici è più alta nel solco maggiore che in quello minore, rendendo quest'ultimo meno utile. Tutti i motivi strutturali proteici che si legano al DNA agiscono secondo questa logica, utilizzando superfici chimicamente complementari a quelle di specifiche sequenze di DNA.

3.1.2.1 L'elica-giro-elica è stato il primo motivo strutturale scoperto

Il motivo **elica-giro-elica** (*helix-turn-helix*), come dice il nome, è costituito, nella sua forma più semplice, da due α -eliche connesse da un breve tratto esteso, sufficientemente flessibile da permettere il ripiegamento di un'elica sull'altra (Figura 3.2a). Più precisamente, le due α -eliche formano tra loro un angolo fisso, determinato da interazioni deboli tra i residui delle



FIGURA 3.2 ▲ Diverse strutture basate sul motivo elica-giro-elica. **a)** Il motivo di base elica-giro-elica, qui mostrato nel repressore Cro del fago λ [codice PDB 5CRO]. **b)** Il motivo omeodominio della proteina Antennapedia di *Drosophila melanogaster* complessato al DNA [codice PDB 9ANT]. **c)** Un esempio di variante "alata" del motivo elica-giro-elica. È mostrata la struttura del dominio N-terminale di BirA, il repressore dell'operone della biotina di *Escherichia coli* [codice PDB 1BIA]. Il dominio costituisce il motivo di riconoscimento del DNA, in cui due foglietti β antiparalleli (rossi) seguono il classico motivo elica-giro-elica. Le eliche di riconoscimento sono sempre colorate in verde, quelle di posizionamento in azzurro.

eliche stesse. L' α -elica C-terminale è chiamata **elica di riconoscimento**, poiché costituisce la porzione del motivo strutturale che interagisce fisicamente con il DNA, a livello di un solco maggiore. La sua sequenza amminoacidica differisce da proteina a proteina e determina la specificità del riconoscimento. L' α -elica N-terminale ha invece un ruolo strutturale, permette cioè di posizionare l' α -elica di riconoscimento in modo corretto per l'interazione con il DNA.

Il motivo elica-giro-elica è stato inizialmente identificato in proteine batteriche coinvolte nella regolazione della trascrizione. Nel corso degli anni, sono state identificate molte variazioni sul tema dell'elica-giro-elica, in proteine sia procariotiche che eucariotiche. A oggi, sono note centinaia di proteine diverse contenenti questo motivo. In particolare, sono state scoperte molte proteine in cui il motivo strutturale è costituito da tre eliche invece che due, con due **eliche strutturali** e una di riconoscimento (vi sono, peraltro, casi in cui è presente anche una quarta elica). L'esempio più importante di questa famiglia è costituito dall'**omeodominio**, motivo strutturale prettamente eucariotico, inizialmente identificato nei geni omeotici di *Drosophila melanogaster* (**Figura 3.2b**). A differenza del motivo elica-giro-elica dei procarioti, in cui la struttura della proteina contenente il motivo è molto variabile, tutti gli omeodomini sono circondati da strutture proteiche simili.

Alcuni motivi elica-giro-elica comprendono filamenti β che interrompono, precedono o seguono le eliche coinvolte nel legame al DNA. I filamenti β formano normalmente un foglietto β antiparallelo, impacchettato addosso alle eliche del motivo. Questi domini misti α - β costituiscono la variante "alata" (*winged*) del motivo elica-giro-elica (**Figura 3.2c**) e presentano una grande diversificazione topologica, in particolare riguardo alla conformazione della sequenza che costituisce il giro del motivo.

3.1.2.2 Lo *zinc finger* è un motivo strutturale prettamente eucariotico

Nel motivo strutturale a **dito di zinco** (*zinc finger*), un atomo di zinco fa da collante a strutture con forme e dimensioni diverse, ma con il tratto comune della formazione di una protuberanza (il dito) che interagisce in modo specifico con il DNA a livello di un solco maggiore della doppia elica. L'atomo di zinco è coordinato alla base del dito da quattro ligandi, che possono essere, in combinazione variabile, cisteina o istidina. Il classico *zinc finger* ha un atomo di zinco coordinato da due cisteine e due istidine (tipo Cys_2His_2) e il dito, lungo una ventina di residui amminoacidici, si struttura a formare un' α -elica da un lato e un foglietto β dall'altro (**Figura 3.3**). Tuttavia, come detto, le strutture degli *zinc finger* sono molto variabili, i ligandi possono essere tutte cisteine e ci sono casi in cui la protuberanza è quasi del tutto assente.

Le proteine che usano questo motivo strutturale contengono di norma più di uno *zinc finger*. Nell'esempio mostrato in **Figura 3.3**, che si riferisce al fattore trascrizionale murino Zfp57, ci sono due motivi adiacenti a dito di zinco che interagiscono con una sequenza con-

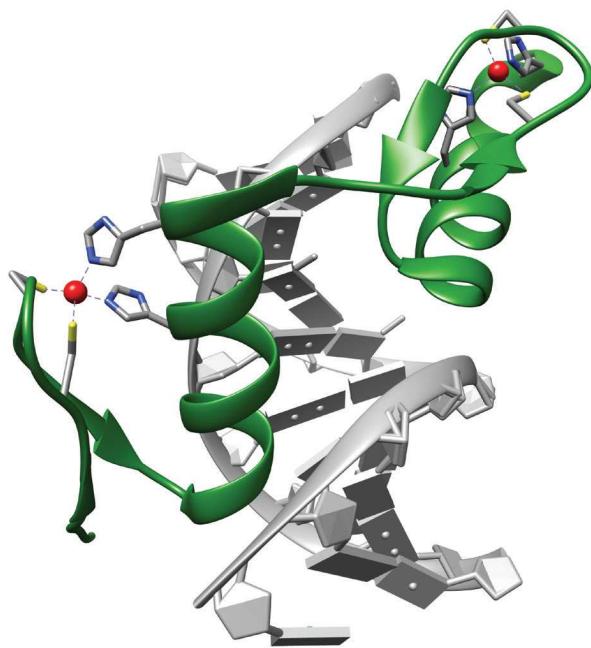


FIGURA 3.3 ▲ Motivo *zinc finger*. Un atomo di zinco (in rosso) è coordinato da residui di cisteina e istidina a formare una struttura in grado di interagire, di norma, con il solco maggiore del DNA. Nella figura è rappresentata la struttura parziale del fattore trascrizionale murino Zfp57 complessato al DNA [codice PDB 4GZN], in cui si notano due motivi classici a dito di zinco, con coordinazione Cys_2His_2 , posti sequenzialmente a legare una corta sequenza consenso di 6 coppie di basi all'interno di un solco maggiore.

senso all'interno di un unico solco maggiore del DNA. Nel fattore $\text{TF}_{\text{III}}\text{A}$ della rana *Xenopus laevis*, la prima proteina in cui questo motivo strutturale è stato identificato, ben nove motivi a dito di zinco interagiscono sequenzialmente con solchi maggiori adiacenti, avvolgendo la proteina in modo destrorso attorno alla doppia elica del DNA e stabilendo contatti multipli che rafforzano l'interazione macromolecolare.

Lo *zinc finger* è un motivo strutturale utilizzato dalle proteine anche per interagire con l'RNA, sebbene le modalità di riconoscimento dell'acido nucleico siano significativamente diverse da quelle con il DNA. In particolare, lo scheletro polinucleotidico dell'RNA gioca un ruolo importante, con i gruppi fosfato coinvolti nel legame; sono anche rilevanti le interazioni di *stacking* (impilamento) tra basi azotate e residui amminoacidici idrofobici.

Alcuni domini a dito di zinco sono utilizzati dalle proteine per interagire con altre proteine, piuttosto che con il DNA o con l'RNA. Queste interazioni sono state, ad esempio, dimostrate in sistemi proteici multimerici coinvolti nel rimodellamento della cromatina, nel funzionamento degli chaperoni, nell'organizzazione del citoscheletro, nello sviluppo degli epitelii e in molti altri processi cellulari. È notevole il fatto che alcuni fattori trascrizionali che usano gli *zinc finger* per interagire con il DNA, sfruttino a volte questi domini anche per interagire tra loro, a ulteriore dimostrazione della grande versatilità strutturale e funzionale del motivo a dito di zinco.

3.1.2.3 Le cerniere di leucina agganciano come una forbice la doppia elica del DNA

La **cerniera di leucine** (*leucine zipper*) è un motivo strutturale che deriva dalla dimerizzazione in parallelo di lunghe α -eliche avvolte una sull'altra. Le due α -eliche sono anfipatiche, cioè hanno una faccia idrofobica e l'altra polare, e di norma derivano da monomeri distinti. Il nome del motivo strutturale deriva dal fatto che le α -eliche contengono una leucina ogni 7 residui. Poiché l' α -elica ha un passo di 3,6 residui (si veda il Paragrafo 2.6.3.1), ogni due giri c'è una leucina, posizionata sempre sulla stessa faccia dell'elica. Il ruolo delle leucine è quello di permettere l'interazione tra le due α -eliche e il loro superavvolgimento attraverso la formazione di interazioni idrofobiche (Figura 3.4).

L'interazione mediata dalle leucine riguarda una parte delle eliche, quella definita **dominio di dimerizzazione**. Distalmente a esso, su ognuna delle due eliche vi è il **dominio di legame al DNA**, con proprietà basiche, costituito dalla porzione rimanente di α -elica che va a posizionarsi all'interno del solco maggiore del DNA (Figura 3.4). Questo è il dominio specifico per una determinata sequenza di DNA. Poiché le cerniere di leucine agiscono in forma dimerica, è possibile la formazione non solo di omodimeri, in cui i domini di legame al DNA riconoscono la stessa sequenza nucleotidica su solchi maggiori adiacenti, ma anche di eterodimeri, che possono riconoscere solchi adiacenti con sequenze diverse (Figura 3.5a). Considerando che sono

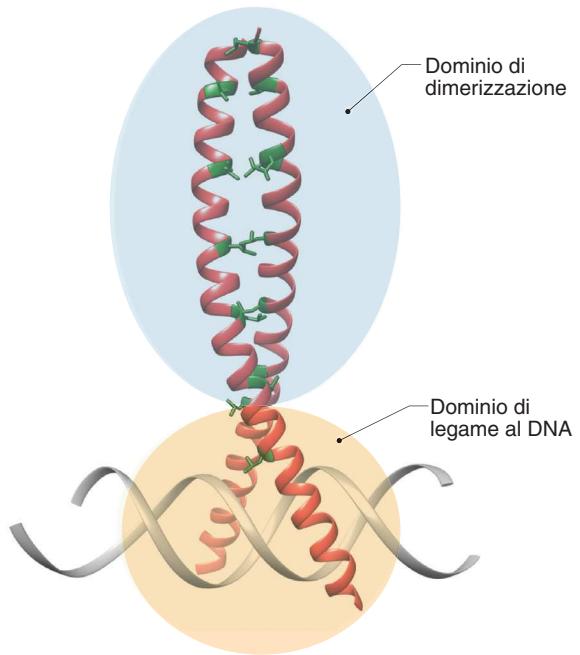


FIGURA 3.4 ► Motivo *leucine zipper*. Due α -eliche di due monomeri distinti formano un'elica superavvolta stabilizzata dalla presenza di residui di leucina ripetuti ogni due giri di α -elica. Il dimerio di superelica stabilizza il dominio di legame al DNA, formato a sua volta dalle porzioni N-terminali delle α -eliche, che interagiscono a livello del solco maggiore del DNA. I residui di leucina sono evidenziati con la struttura a bastoncini.

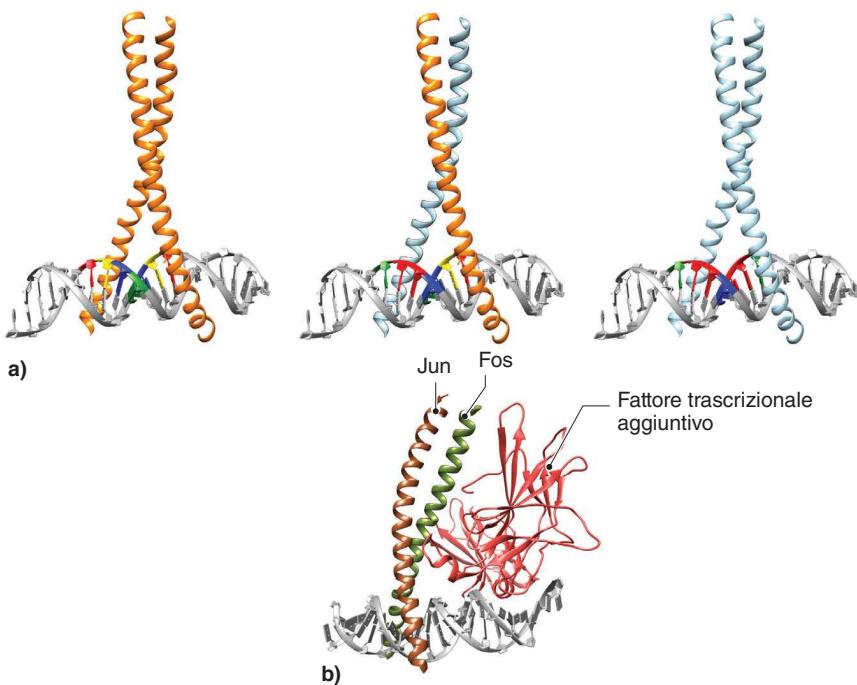


FIGURA 3.5 ▶ a) Le *leucine zipper* possono essere sia omodimeriche che eterodimeriche. Strutture omodimeriche riconoscono sequenze palindromiche di DNA; strutture eterodimeriche possono formarsi combinando monomeri diversi e riconoscono sequenze non palindromiche. b) Nell'eterodimero Fos/Jun [codice PDB 1A02], la *leucine zipper* è significativamente asimmetrica, con il monomero Jun che avvolge il monomero Fos prima di interagire con il DNA. L'asimmetria dell'eterodimero favorisce il posizionamento, su un lato della sequenza riconosciuta dalla *leucine zipper*, di un'altra proteina legante il DNA.

state identificate finora un centinaio di proteine che utilizzano il motivo strutturale a cerniera di leucine, il numero di possibili combinazioni eterodimeriche è molto elevato.

La possibilità che nelle proteine due o più α -eliche potessero avvolgersi tra loro, come i fili in una corda, era stata predetta da Crick già all'epoca della definizione della struttura del DNA. La prima prova cristallografica dell'esistenza del motivo a cerniera di leucine è invece relativamente recente, ottenuta nel 1991 sul fattore trascrizionale di lievito omodimerico GCN4, e ha dimostrato come la struttura sia esattamente quella predetta da Crick. La prova dell'esistenza di eterodimeri è ancora più recente; uno degli esempi più importanti è quello del complesso Fos/Jun, i due componenti di AP-1, fattore trascrizionale coinvolto, tra l'altro, nella proliferazione cellulare e nell'insorgenza del cancro. La risoluzione della struttura cristallografica del complesso ha mostrato come i due monomeri interagiscano con il DNA in modo leggermente diverso, con l'elica di Fos che si inserisce in modo diretto nel solco maggiore e l'elica di Jun che vi si ripiega intorno prima di legarsi al DNA (**Figura 3.5b**). In questo caso, l'asimmetria dell'eterodimero permette il legame di un'altra proteina legante il DNA, selettivamente su uno dei due lati del complesso Fos/Jun.

3.1.2.4 Il motivo elica-ansa-elica sembra combinare le proprietà di altri domini strutturali

Come per le cerniere di leucine, le proteine che contengono il motivo elica-ansa-elica (*helix-loop-helix*) agiscono sempre in forma dimerica, di norma come eterodimeri. Ogni monomero contiene un motivo lungo circa 60 aminoacidi, formato da due α -eliche connesse da un'ansa. Il posizionamento delle eliche è molto diverso da quello del motivo elica-giro-elica (Figura 3.6) e anche il ruolo delle due eliche è differente: l' α -elica H1, più lunga, posizionata all'estremità N-terminale del dominio, contiene residui basici che favoriscono l'interazione con un tratto di 6 nucleotidi del DNA chiamato E-box; quella più corta H2 al C-terminale serve per la dimerizzazione, in modo analogo a quanto abbiamo visto per le cerniere di leucine. In realtà, l'elica H2 può allungarsi a comprendere anche un motivo *leucine zipper*, come nell'esempio mostrato in figura. La variabilità delle sequenze nelle E-box, associata al potenziale combinatorio delle regioni di dimerizzazione, determina come i vari membri della famiglia controllino processi diversi tra loro, tra cui la determinazione del sesso e lo sviluppo del sistema nervoso e dei muscoli.

Tra le strutture cristallografiche risolte di proteine contenenti il motivo elica-ansa-elica, in Figura 3.6 è mostrata quella relativa all'eterodimero Max-Mad, un fattore di trascrizione coinvolto nel controllo della proliferazione cellulare. È particolare la presenza su ogni monomero di una α -elica di dimerizzazione costituita da due segmenti: l'elica canonica H2 e un'elica contenente una cerniera di leucine. È un esempio lampante di come motivi strutturali diversi possano trovarsi a lavorare insieme nella stessa proteina.

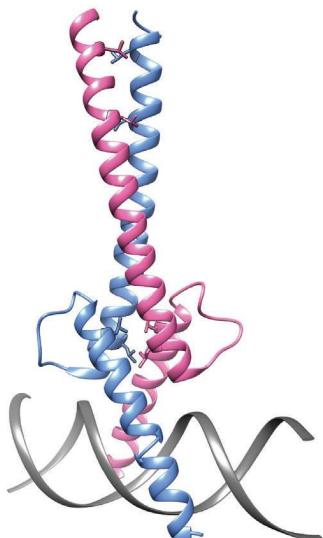


FIGURA 3.6 ▶ Motivo elica-ansa-elica. Le proteine che sfruttano questo motivo agiscono sempre in forma dimerica. Il dимерo può essere stabilizzato dalla presenza di ulteriori tratti di α -elica su ogni monomero, che insieme formano una *leucine zipper*, come nel caso del complesso Mad-Max qui mostrato [codice PDB 1NLW]. I residui di leucina sono evidenziati con la struttura a bastoncini.

[3.2] ASSEMBLAGGIO DEL DNA IN CROMOSOMI: STRUTTURA DELLA CROMATINA

Il genoma umano diploide contiene circa 6,4 miliardi di coppie di basi, distribuite in 46 cromosomi. Il cromosoma 1, il più grande, contiene circa 250 milioni di coppie di basi. Descrivendo la struttura del DNA-B, abbiamo visto come la distanza tra coppie di basi adiacenti sia di circa 3,4 Å (1 Å = 10^{-8} cm). Se la struttura fosse rigidamente lineare, il cromosoma 1 si estenderebbe quindi per 8,5 cm e tutto il genoma messo in fila sarebbe lungo oltre 2 metri, valori evidentemente incompatibili con il diametro medio delle cellule, dell'ordine dei micron. Utilizzando lo stesso schema, potremmo calcolare che il DNA contenuto nei circa 50.000 miliardi di cellule del corpo umano coprirebbe una distanza pari a circa 600 volte quella che separa la Terra dal Sole. È importante, quindi, capire come il DNA venga confinato all'interno del nucleo della cellula.

Abbiamo già visto come il DNA possa assumere strutture più compatte di ordine superiore grazie al superavvolgimento. I procarioti utilizzano, infatti, la formazione di supereliche per impacchettare il DNA al proprio interno, anche con l'aiuto di piccole proteine architettoniche in grado di piegare la doppia elica o di fare da ponte tra tratti distanti del duplex. Mentre, però, questo è sufficiente nel caso di genomi di grandezza dell'ordine di qualche milione di coppie di basi, negli eucarioti devono intervenire processi più complessi, in grado di ridurre di molti ordini di grandezza lo spazio occupato dal DNA. Il prodotto finale di questi processi di compattamento è la **cromatina**, una struttura con livelli gerarchici di compattamento che costituisce il

substrato genomico fondamentale per tutti i processi vitali del DNA. La cromatina comprende masse all'incirca equivalenti di DNA e di proteine e presenta un compattamento dinamico, poiché cambi di condensazione devono avvenire in modo rapido e al momento opportuno (ad esempio in specifiche fasi del ciclo cellulare). Cambiamenti nello stato di compattamento possono altresì essere non solo globali all'interno di una cellula (ad esempio nella mitosi), ma anche localizzati, come nel caso della trascrizione di specifici geni. La necessità di cambiamenti dinamici dello stato di compattamento ha richiesto l'evoluzione di specifici sistemi enzimatici in grado di alterare in modo specifico lo stato del DNA. In definitiva, in ogni istante, il DNA si trova compattato in modo non omogeneo: i tratti molto compattati e meno accessibili vanno sotto il nome di **eterocromatina**, quelli più rilassati e accessibili si definiscono invece **eucromatina**. Il termine eucromatina (dal greco *eu*, bene) suggerisce in modo immediato che questa è la struttura funzionalmente attiva, come vedremo nei capitoli successivi, in particolare quando parleremo della replicazione e della regolazione della trascrizione. Qui ci soffermeremo, invece, solo sugli aspetti strutturali, cioè sui meccanismi che portano la doppia elica a compattarsi.

[3.2.1] Il nucleosoma costituisce il primo livello nel compattamento del DNA

Il primo, fondamentale livello di compattamento del DNA eucariotico è costituito dal **nucleosoma**, struttura costituita da un ottamero proteico su cui si avvolge la doppia elica del DNA. Il nocciolo, o *core*, del nucleosoma è formato da otto molecole di **istoni**, piccole proteine basiche ricche di arginina e lisina, adatte a interagire per via elettrostatica con il DNA in modo sequenza-indipendente. Costituiscono il *core* ottamerico del nucleosoma gli istoni H2A, H2B, H3 e H4, ognuno presente in due copie. Un quinto tipo di istone, denominato H1, si lega al DNA *linker*, cioè al tratto di DNA posto tra due nucleosomi adiacenti, e ha il ruolo, come vedremo, di favorire l'ulteriore compattamento della cromatina.

In vitro, la rimozione dalla cromatina dell'istone H1 è molto più agevole di quella degli altri istoni, a dimostrazione della forza che tiene insieme l'ottamero istonico nel *core*. Quando l'istone H1 viene rimosso, la cromatina assume una particolare struttura, visibile al microscopio elettronico, denominata filo di perle (**Figura 3.7a**), nella quale a ogni perla corrisponde un nucleosoma e il DNA *linker* collega i nucleosomi tra loro. Poiché il diametro di un nucleosoma è di circa 11 nm, il filo di perle viene anche chiamato, per approssimazione, fibra da 10 nm. La quantità di DNA contenuta in ogni nucleosoma è stata determinata trattando la cromatina a filo di perle con una DNasi, cioè un enzima idrolitico che taglia il DNA. La *DNasi micrococcica*, o MNasi, ha la proprietà di tagliare il DNA interposto tra i nucleosomi, non protetto dal legame dell'ottamero istonico, con efficienza molto più alta rispetto al DNA avvolto intorno al *core*. Un trattamento blando della cromatina con MNasi permette di regolare la quantità di tagli effettuati.

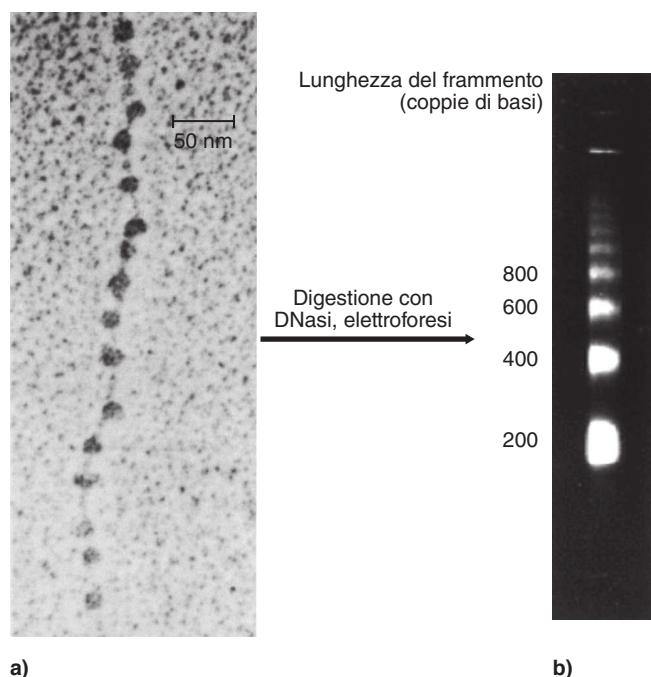


FIGURA 3.7 ► a) La struttura a filo di perle rivelata mediante microscopia elettronica. Le strutture più dense (le "perle") sono i nucleosomi. b) Profilo eletroforetico del DNA sottoposto a parziale digestione con DNasi. Il gel eletroforetico mostra bande corrispondenti a frammenti di DNA la cui lunghezza è multipla di 200 coppie di basi.

tuiti tra un nucleosoma e l'altro. Se la reazione viene fermata in modo che sia avvenuto al massimo un taglio idrolitico tra un nucleosoma e il successivo, e la grandezza dei frammenti ottenuti viene determinata mediante analisi elettroforetica, si osservano bande molto nette, di grandezza pari a multipli di circa 200 coppie di basi, corrispondenti a nucleosomi, dinucleosomi, trinucleosomi e così via (**Figura 3.7b**). La presenza di bande discrete è la dimostrazione della struttura ripetitiva della cromatina e indica che la grandezza dei nucleosomi è sostanzialmente indipendente dalla sequenza del DNA che vi si avvolge intorno.

3.2.1.1 Gli istoni costituiscono l'impalcatura molecolare del nucleosoma

Nel processo di assemblaggio dell'**ottamero istonico**, si formano inizialmente eterodimeri H3-H4 e H2A-H2B. Negli eterodimeri, i due monomeri sono collegati in conformazione testa-coda. Due dimeri H3-H4 tetramerizzano prima di interagire con due dimeri H2A-H2B per formare il *core* nucleosomico completo (**Figura 3.8a**). La forma del *core* è quella di un disco scanalato sul bordo, con un asse di simmetria diadioco che lo attraversa longitudinalmente. In altri termini, la struttura macroscopica del *core* non cambia se esso viene fatto ruotare di 180° intorno all'asse di simmetria (**Figura 3.8b**). Questa proprietà è la conseguenza diretta del pos-

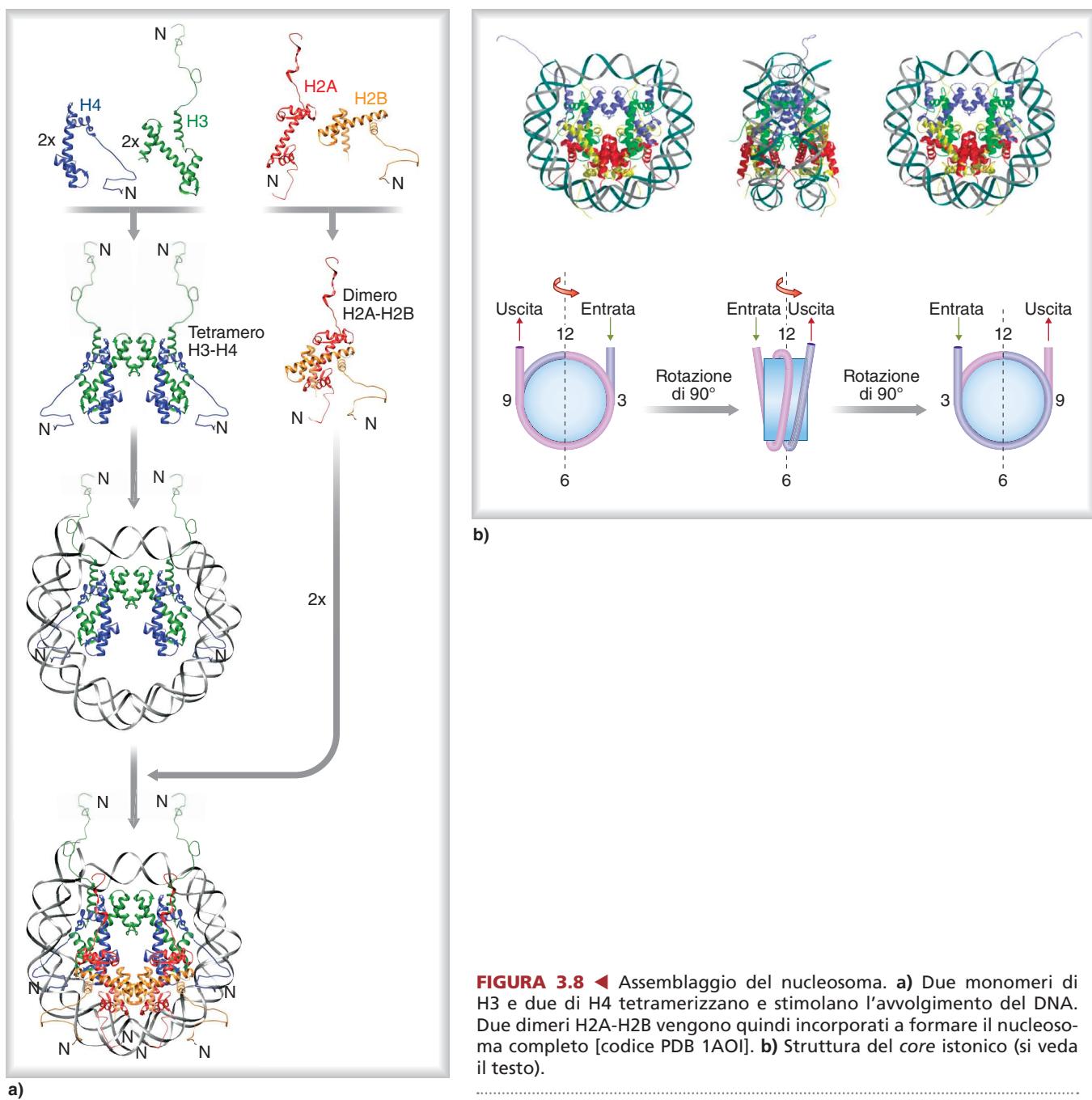


FIGURA 3.8 ▲ Assemblaggio del nucleosoma. **a)** Due monomeri di H3 e due di H4 tetramerizzano e stimolano l'avvolgimento del DNA. Due dimeri H2A-H2B vengono quindi incorporati a formare il nucleosoma completo [codice PDB 1AOI]. **b)** Struttura del core istonico (si veda il testo).

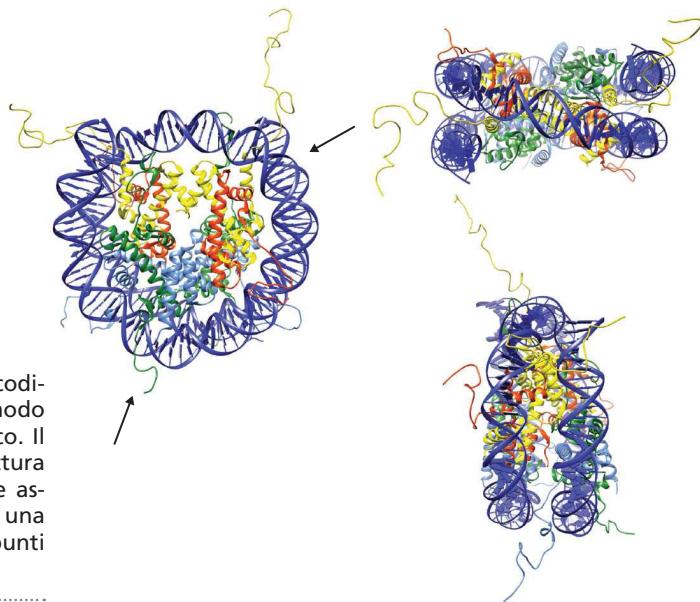


FIGURA 3.9 ► Struttura cristallografica del nucleosoma [codice PDB 1KX5]. 146 coppie di basi di DNA si avvolgono in modo sinistrorso, formando 1,67 giri intorno all'ottamero istonico. Il nucleosoma è mostrato di fronte, di lato e dall'alto. La struttura cristallografica è stata risolta su un nucleosoma artificiale assemblato utilizzando gli istoni della rana *Xenopus laevis* e una sequenza palindromica di DNA umano. Le frecce indicano punti in cui la doppia elica del DNA si ripiega bruscamente.

zionamento dei monomeri di ognuno dei quattro tipi di istone all'interno dell'ottamero. La struttura cristallografica del *core* del nucleosoma è stata risolta ad alta risoluzione dal gruppo di Timothy Richmond nel 2002 e ha svelato i dettagli dell'interazione tra DNA e ottamero istonico: il DNA si avvolge intorno al *core* descrivendo quasi due giri completi, per la precisione 1,67 giri, costituiti da 146 coppie di basi (**Figura 3.9**). Questo dato non è in contrasto con il risultato degli esperimenti di digestione nucleasica descritti sopra: l'unità di circa 200 coppie di basi, infatti, comprende sia la porzione di DNA avvolta intorno al *core* sia il tratto del DNA *linker* interposto tra nucleosomi adiacenti. Gli istoni interagiscono con il DNA a livello dei solchi minori, attraverso la formazione di numerosi legami idrogeno tra i residui amminoacidici e gli atomi di ossigeno dei legami fosfodiesterici posti in prossimità dei solchi minori stessi. Ci sono, di conseguenza, 14 punti di contatto tra istoni e DNA nel *core* del nucleosoma. L'avvolgimento del DNA intorno al *core* è sinistrorso (**Figure 3.8 e 3.9**). Come abbiamo visto nel Paragrafo 2.4, la formazione di un solenoide sinistrorso equivale topologicamente a un superavvolgimento negativo. L'avvolgimento del DNA intorno all'ottamero istonico richiede, infatti, la rimozione di circa un giro di elica. Poiché l'avvolgimento non implica la rottura del doppio filamento del DNA, esso non fa variare il numero di legame (L) ed è quindi topologicamente invariante. Di conseguenza, il superavvolgimento negativo del DNA avvolto deve essere compensato da un superavvolgimento positivo nelle regioni *linker*. Questo superavvolgimento positivo, però, viene rimosso da specifiche topoisomerasi, lasciando invariati i superavvolgimenti negativi a livello del nucleosoma. La formazione del nucleosoma costituisce a tutti gli effetti il modo in cui gli eucarioti superavvolgono negativamente il proprio DNA. Infatti, a differenza dei procarioti, che usano a questo fine la girasi (si veda il Paragrafo 2.4.3), gli eucarioti non possiedono topoisomerasi in grado di introdurre attivamente superavvolgimenti negativi. La cosa è di fondamentale importanza, poiché implica che, quando un nucleosoma si disgrega, il duplex di DNA libero dal *core* può, grazie allo stato di superavvolgimento negativo, più facilmente separarsi nei due filamenti, un processo fondamentale sia nella replicazione che nella trascrizione del DNA.

Tutti gli istoni presentano un motivo strutturale peculiare, denominato *histone fold*, o **ripiegamento istonico**, costituito da tre corte α -eliche connesse da brevi anse (**Figura 3.10**). I punti di contatto con il DNA sono principalmente (129 coppie di basi su 146) contenuti nell'*histone fold* di ogni istone, che interagisce, in particolare, con lo scheletro fosfodiesterico o con i solchi minori del duplex. È interessante notare come l'interazione istoni-DNA sia favorita dall'attrazione elettrostatica tra le cariche positive delle catene laterali degli istoni e quelle negative dei gruppi fosfato del DNA, ma stabilizzata in modo predominante da legami idrogeno che si formano tra il DNA e lo scheletro polipeptidico degli istoni. La struttura cristallografica ha mostrato, inoltre, che il *core* del nucleosoma è molto compatto, mentre le regioni N-terminali degli istoni sono più disordinate e costituiscono delle vere e proprie protrusioni dal *core* (**Figura 3.9**). Queste code istoniche non contribuiscono in modo apprezzabile al legame con il DNA, mentre hanno un ruolo importante nello stabilire interazioni tra nucleosomi adiacenti, favorendo gli ulteriori processi di compattamento della cromatina. Per sporgere all'esterno del

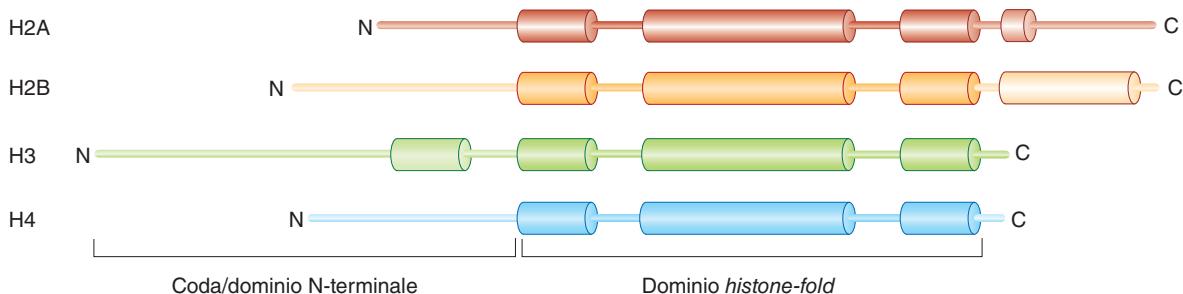


FIGURA 3.10 ▲ Il dominio *histone fold* è contenuto in tutti gli istoni del core del nucleosoma. L'*histone fold* è costituito da tre porzioni ad α -elica connesse da brevi anse.

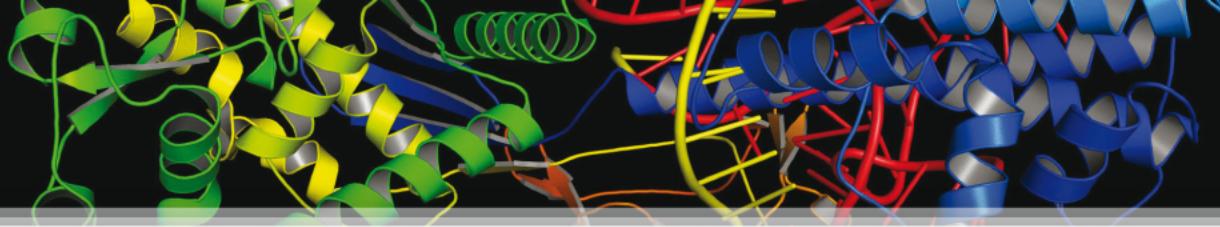
core, le code devono attraversare la doppia elica del DNA avvolto intorno all'ottamero. Per fare ciò, esse si fanno letteralmente strada attraverso canali costituiti sul DNA dai solchi minori delle due eliche adiacenti allineati tra loro (Figura 3.9). Le **code istoniche** sono il bersaglio di modificazioni post-traduzionali importantissime per la regolazione dell'attività e della struttura degli istoni. Più avanti in questo capitolo affronteremo la questione del ruolo di queste modificazioni epigenetiche nella dinamica dei nucleosomi (si veda il Paragrafo 3.2.4), mentre delle modificazioni post-traduzionali delle proteine in generale parleremo nel Capitolo 9.

Gli istoni si sono evoluti precocemente nella storia degli eucarioti; essi si possono, infatti, ritrovare in organismi filogeneticamente molto distanti, come protozoi, funghi, piante e mammiferi. Proteine simili agli istoni si ritrovano anche negli *Archaea* e nei procarioti. I batteri, ad esempio, utilizzano le proteine HU per piegare il DNA e stabilizzare il superavvolgimento del cromosoma durante la replicazione e la trascrizione. Parimenti, nel batterio estremofilo *Methanothermus fervidus* ci sono istoni archeali che, sotto forma di complessi tetramericici, si avvolgono intorno alla doppia elica del DNA. Alcuni istoni archeali, come quelli del *Methanopyrus kandleri*, contengono addirittura un motivo istonico omologo all'*histone fold* eucariotico.

3.2.1.2 Il bending del DNA su specifiche sequenze favorisce la formazione del nucleosoma

Il **ripiegamento (bending)** del DNA intorno al *core* del nucleosoma non è omogeneo: il duplex si avvolge a scatti, con piegamenti acuti intervallati da brevi tratti quasi lineari (Figura 3.9). Per fare ciò, il DNA deve adattarsi alla forma dell'ottamero. Poiché nella forma avvolta il DNA ha un passo inferiore a quello della forma canonica B (10,2 contro 10,5 coppie di basi per giro), con il conseguente restringimento del solco minore, i piegamenti acuti sono favoriti da sequenze di DNA che permettono il riarrangiamento di questo solco. Abbiamo già citato nel Paragrafo 2.2.6 l'effetto favorente di sequenze ripetute di coppie A-T sulla curvatura del DNA. Coerentemente con questo, si è osservato che, anche nei nucleosomi, sequenze ripetute A-T favoriscono il restringimento del solco minore e il conseguente piegamento della doppia elica, mentre sequenze ripetute G-C lo contrastano (Figura 3.11). L'inserzione di un residuo istonico di arginina all'interno del solco minore è un altro determinante che favorisce il piegamento della doppia elica. Il concetto, dunque, della formazione dei nucleosomi del tutto indipendente dalla sequenza del DNA va riconsiderato: i nucleosomi possono formarsi grosso modo in qualunque punto, tuttavia la struttura fine del ripiegamento del DNA attorno al *core* istonico è modulata dalla sequenza dello specifico tratto di DNA e determina, in ultima istanza, la probabilità dell'assemblaggio del *core* e la forza della sua interazione con il DNA.

Attualmente, è possibile determinare con precisione la posizione dei nucleosomi lungo il DNA genomico. Ad esempio, il gruppo di Mavrich nel 2008 ha analizzato il posizionamento dei nucleosomi nel genoma del lievito, combinando elegantemente tecniche biochimiche classiche e moderni metodi biomolecolari (Figura 3.12). In particolare, gli istoni del *core* sono stati dapprima legati covalentemente tra loro mediante *cross-link* (una tecnica biochimica che permette la formazione di legami crociati tra polipeptidi diversi), quindi il DNA genomico è stato digerito con nucleasi e i nucleosomi risultanti sono stati separati mediante cromatografia di affinità, sfruttando anticorpi specifici per gli istoni H3 e H4. Il DNA associato a ogni singolo nucleosoma è stato poi sequenziato. La ricerca di Mavrich ha messo in evidenza come alcune regioni del DNA genomico del lievito non siano propense a formare nucleosomi, probabilmente perché devono mantenere in modo costante la possibilità di essere trascritte (o di legare fattori che regolano la trascrizione; si veda il Capitolo 5).



Biologia Molecolare

Accedi all'ebook e ai contenuti digitali ➤ Espandi le tue risorse ➤ con un libro che **non pesa** e si **adatta** alle dimensioni del tuo **lettore**



All'interno del volume il **codice personale** e le istruzioni per accedere alla versione **ebook** del testo e agli ulteriori servizi.
L'accesso alle risorse digitali è **gratuito** ma limitato a **18 mesi dalla attivazione del servizio**.



€ 59,00



9 788836 230150