



**HAL**  
open science

# Des moteurs de jeux à la physique des chromosomes

Pascal Lilian Carrivain

► **To cite this version:**

Pascal Lilian Carrivain. Des moteurs de jeux à la physique des chromosomes. Biophysique [physics.bioph]. Université Pierre et Marie Curie - Paris VI, 2012. Français. NNT: . tel-00802204

**HAL Id: tel-00802204**

**<https://theses.hal.science/tel-00802204>**

Submitted on 19 Mar 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Pierre & Marie Curie - Paris VI

ÉCOLE DOCTORALE DE LA PHYSIQUE DE LA PARTICULE A LA MATIÈRE CONDENSÉE

## THÈSE

*préparée au Laboratoire de Physique Théorique de la Matière Condensée*

présentée pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

*Spécialité : Systèmes Complexes*

*par*

**Pascal Carrivain**

sous la direction du Dr. Jean-Marc Victor

*Titre :*

**Des moteurs de jeux à la physique des chromosomes**

*à soutenir publiquement le 19 décembre 2012 devant le*

### JURY

M. Aurélien Bancaud	CNRS (LAAS)	<i>Examineur</i>
M. David Bensimon	CNRS (ENS Paris)	<i>Examineur</i>
M. Ralf Everaers	Professeur ENS Lyon	<i>Rapporteur</i>
M. Alexander Grosberg	Professeur NYU	<i>Rapporteur</i>
M. Jean-François Joanny	Professeur UPMC	<i>Président du jury</i>
M. Sébastien Neukirch	CNRS (Jean Le Rond d'Alembert)	<i>Examineur</i>
M. Jean-Marc Victor	CNRS (LPTMC)	<i>Directeur de thèse</i>

---

---

حیاتی

---

---

# Table des matières

---

<b>0</b>	<b>Remerciements</b>	<b>3</b>
<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Moteurs physiques et dynamique de Langevin pour l'ADN</b>	<b>9</b>
2.1	Méthodes traditionnelles de simulation de l'ADN . . . . .	9
2.1.1	Méthode du Pivot, Krankshaft et le Bead spring model . . . . .	9
2.1.2	SHAKE et QSHAKE . . . . .	11
2.2	Formalisme des moteurs physiques . . . . .	13
2.3	Implémentation des joints mécaniques dans Open Dynamics Engine (ODE) . . . . .	15
2.4	Gestion des collisions . . . . .	16
2.5	Présentation du modèle mécanique d'ADN . . . . .	17
2.5.1	Joint Ball-in-Socket . . . . .	17
2.5.2	Couples de courbure et de twist pour un Ball-Like-Chain (BLC) . . . . .	19
2.5.3	Exemple : chaîne de quatre cylindres connectés . . . . .	21
2.5.4	Électrostatique . . . . .	23
2.6	Dynamique locale de Langevin-Euler . . . . .	24
2.7	Dynamique globale de Langevin-Euler . . . . .	24
2.8	"Open Dynamics Engine" en pratique . . . . .	28
2.8.1	Simulation en pratique . . . . .	28
2.8.2	Quelques remarques sur les paramètres Error Reduction Parameter (ERP) et Constraint Force Mixing (CFM) . . . . .	29
2.8.3	Les différentes étapes de la simulation . . . . .	32
2.9	Chaîne de Cardan . . . . .	33
2.9.1	Définition . . . . .	33
2.9.2	Calcul du twist . . . . .	34
2.9.3	Persistance . . . . .	35
2.9.4	Matrice de transfert itérative et ULC . . . . .	35

---

<b>3</b>	<b>ADN nu sous pinces magnétiques ; “in silico”</b>	<b>39</b>
3.1	Comparaison du thermostat local et de la version globale . . . . .	40
3.2	ADN en traction pure . . . . .	43
3.3	ADN sous traction et torsion . . . . .	43
3.3.1	Nombre de tours de bille imposé . . . . .	44
3.3.2	Couple sur la bille imposé . . . . .	47
<b>4</b>	<b>Assemblages de nucléosomes</b>	<b>55</b>
4.1	Présentation du modèle mécanique du nucléosome . . . . .	58
4.2	Calcul de structure sur le nucléosome . . . . .	60
4.2.1	Première étape : isolement de l’ADN nucléosomal . . . . .	60
4.2.2	Seconde étape : isolement de l’octamère . . . . .	62
4.3	Nucléosome en traction et torsion . . . . .	63
4.3.1	Nucléosome en traction . . . . .	63
4.3.2	Nucléosome en traction et torsion . . . . .	65
<b>5</b>	<b>Levure</b>	<b>73</b>
<b>6</b>	<b>Conclusion et perspectives</b>	<b>89</b>
<b>A</b>	<b>Algorithmes utilisés</b>	<b>93</b>
A.1	Schémas d’intégration . . . . .	93
A.2	Algorithme de Metropolis . . . . .	95
A.2.1	Moyenne d’une observable . . . . .	95
A.2.2	Equation maîtresse . . . . .	96

---

# Table des figures

---

2.1	Vue schématique de la méthode du Pivot et de la méthode du Krankshaft. Une partie de la chaîne (rouge) est sélectionnée. Un axe aléatoire et un angle aléatoire composent une rotation qui amène la partie rouge sur la partie bleue (méthode du Pivot). Une partie de la chaîne (verte) est sélectionnée. Cette partie est amenée sur la partie magenta par une rotation d'axe reliant les deux extrémités de cette chaîne et d'angle $\phi$ aléatoire (méthode du Krankshaft).	11
2.2	Vue schématique des différentes étapes de l'algorithme SHAKE pour un système de deux particules liées rigidement $\delta = d$ . Une première étape consiste à mener le système au pas $t + \Delta t$ sans tenir compte des contraintes. Les positions obtenues ne satisfont pas la contrainte et les positions sont corrigées de manière itérative en utilisant des forces de corrections $f_\delta$ montrées en bleu.	13
2.3	Vue schématique d'un joint "Ball-in-Socket". ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.	18
2.4	Vue schématique d'un joint "Ball-in-Socket" ne satisfaisant pas la condition d'ancrage (voir l'équation (2.8)). ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.	18
2.5	Vue schématique d'un joint de Cardan. ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.	33
2.6	Carte du twist pour un joint de Cardan. En abscisse et ordonnée sont représentés en radians les angles $\xi_1$ et $\xi_2$ . Les contours de couleurs indiquent la valeur du twist.	34
2.7	Le profil force-extension est montré pour les modèles simulés du BLC, du Freely-Jointed-Chain, du Worm-Like-Chain (WLC) analytique et du WLC numérique [1, 2]. La quantité $\beta fp$ est la force adimensionnée et est représentée en échelle logarithmique tandis que $\epsilon = z/L$ est l'extension relative de la molécule d'ADN. Trois longueurs d'ADN sont simulées $L = 1, 2, 3 \mu\text{m}$ (triangles rouges, verts et bleus). Les erreurs statistiques sont inférieures à la taille des symboles.	37



3.1	Les corrélations tangentielles sont montrées (en ordonnée) le long de l'ADN (en abscisse). Les simulations sont réalisées pour différentes segmentations $l = 3.34, 5, 6.67$ nm et pour une molécule d'ADN de longueur $L = 1$ $\mu\text{m}$ . Les courbes théoriques sont montrées par les lignes bleues. . . . .	40
3.2	Les fonctions d'autocorrélation (en ordonnée) en fonction du temps de simulation adimensionné $\Sigma\tau$ (en abscisse) de l'énergie cinétique $T+R$ du système sont montrées pour les deux versions du thermostat de Langevin. Les courbes sont ajustées par une exponentielle. . . . .	41
3.3	Les fonctions d'autocorrélation (voir l'équation (3.2)) (en ordonnée) de la distance bout à bout en fonction du temps de simulation $\Sigma\tau$ adimensionné (en abscisse) sont montrées pour les deux versions du thermostat de Langevin. La courbe dans le cas du thermostat local est ajustée par une exponentielle décroissante tandis que la courbe dans le cas du thermostat global est ajustée par une enveloppe exponentielle décroissante. . . . .	42
3.4	La diffusion de la distance bout à bout normalisée $\langle (\mathbf{R}(t+\tau) - \mathbf{R})^2 \rangle / 2Lk$ (en ordonnée) en fonction du temps de simulation $\Sigma\tau$ adimensionné (en abscisse) est montrée dans la version locale (rouge) et globale (noire) du thermostat de Langevin. L'échelle est en log-log. . . . .	42
3.5	Le profil force extension est montré pour les modèles simulés du Ball-Like-Chain, du Freely-Jointed-Chain (FJC), du WLC analytique et du WLC numérique [1, 2]. La quantité $\beta fp$ représente la force adimensionnée tandis que $\epsilon = z/L$ est l'extension relative de la molécule d'ADN. Les erreurs statistiques sont inférieures à la taille des symboles. . . . .	44
3.6	Vue schématique d'un joint glissière. ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved. . . . .	45
3.7	Le profil extension relative de l'ADN (en ordonnée) overtwist (en abscisse) est montré pour différentes forces et une concentration molaire en sel NaCl de 100 mmol. Les résultats de simulations sont représentés par les symboles de différentes couleurs. Les résultats expérimentaux de F. Mosconi [3] sont montrés en lignes continues et de la même couleur que les symboles correspondants aux conditions expérimentales. Ce choix est malheureux mais permet de ne pas nuire à la clarté de la figure. . . . .	46
3.8	Le rayon de supercoiling $\rho$ , l'angle hélical $\theta$ et la pente $4\pi\rho/H \sin 2\theta$ sont montrés en ordonnées après avoir été ajustés sur les données expérimentales de F. Mosconi [3] et sur les données de simulations. Ces observables sont fonctions de la force appliquée $f$ (pN) sur la bille magnétique. . . . .	47
3.9	Vue schématique de deux plectonèmes obtenus avec un ADN de longueur $L = 1$ $\mu\text{m}$ sous une force de traction de $f = 0.74$ pN et un nombre de tours de bille de $n = 15$ . Il est rare d'observer deux plectonèmes sur un même ADN car ils fusionnent très rapidement en un seul plectonème. Cette configuration n'est pas figée, l'agitation thermique permet de faire diffuser le(s) plectonème(s) le long de l'ADN. Une expérience récente interprète ses résultats en terme de multi-plectonèmes [4]. . . . .	48
3.10	Vue schématique d'une configuration obtenue par simulation avec une force de traction de 0.74 pN et un couple de 20 pN·nm. L'ADN de longueur à vide $L = 1$ $\mu\text{m}$ est totalement supercoilé. . . . .	49

3.11	Les courbes $\Gamma(\sigma)$ sont représentées sur la figure de gauche. La figure de droite représente l'évolution de la taille de la molécule et de l'overtwist en fonction du temps pour une force de traction de $f = 0.74$ pN et des couples de $\Gamma = 0, 9.7, 25$ pN · nm. Il est intéressant de faire le parallèle entre ces deux figures en remarquant que les traces se trouvent en face de la partie étirée de l'ADN, de la transition vers les plectonèmes et de la partie où toute la molécule d'ADN est supercoilée. . . . .	50
3.12	L'instabilité de flambage est montrée pour une force de traction de $f = 0.74$ pN et quatre différents couples de 9.5 à 9.8 pN · nm par 0.1 pN · nm. La trace de l'extension relative $\epsilon$ est montrée en bleu et, en rouge, est montrée la trace de l'overtwist $\sigma$ . . .	51
3.13	L'instabilité de flambage est montrée pour une force de traction de $f = 0.91$ pN et quatre différents couples de 10.7 à 11 pN · nm par 0.1 pN · nm. La trace de l'extension relative $\epsilon$ est montrée en bleu et, en rouge, est montrée la trace de l'overtwist $\sigma$ . . .	52
3.14	L'instabilité de flambage est montrée pour une force de traction de $f = 1.13$ pN et six différents couples de 12.3 à 12.6 pN · nm par 0.1 pN · nm. La trace de l'extension relative $\epsilon$ est montrée en bleu et, en rouge, est montrée la trace de l'overtwist $\sigma$ . . .	53
4.1	Vue schématique des différents mécanismes de compaction pour un ADN de longueur 900 bp : le "supercoiling" et "l'architecturing" par des protéines. (Bleu et vert) cylindres représentant les dimères H3-H4, respectivement, HMfA et HMfB pour les protéines des Archaea. (Rouge et jaune) cylindres représentant les dimères H2A-H2B. L'ADN lié à un dimère est représenté de la même couleur. Pour apprécier pleinement ces deux mécanismes de compaction, les quatre images sont représentées à la même échelle. . . . .	56
4.2	Vue schématique du modèle de nucléosome. L'ADN nucléosomal est enroulé à gauche. (Rouge et jaune) dimères H2A-H2B. (Bleu et vert) dimères H3-H4. (Noir) Four-Helix-Bundle. D'après [5]. . . . .	57
4.3	Vue schématique du modèle de dégrafage symétrique du nucléosome. A gauche, l'ADN est déroulé de un tour de façon symétrique. A droite, l'ADN est déroulé de deux tours de façon symétrique. . . . .	58
4.4	Tableau récapitulatif des symétries. . . . .	58
4.5	Vue schématique du modèle de nucléosome. L'ADN nucléosomal est enroulé à gauche. (Rouge et jaune) dimères H2A-H2B. (Bleu et vert) dimères H3-H4. (Noir) Four-Helix-Bundle. L'ADN est représenté avec la couleur du dimère avec lequel il interagit. La position du petit sillon est montrée par les petites billes de couleur. . . . .	59
4.6	Tableau récapitulatif des interactions "Super-Helix-Location" entre les dimères et l'ADN nucléosomal. . . . .	60
4.7	Les cylindres d'ADN sont représentés par des lignes rouges et vertes (le code couleur reste inchangé). Les flèches indiquent les tangentes des cylindres. Ils sont reliés par des joints mécaniques "Ball-in-Socket" qui sont schématisés par des cercles noirs. Les SHLs sont représentés par des triangles noirs. La tension de part et d'autre d'un SHLs est indiquée en bleu. . . . .	62
4.8	Tableau récapitulatif des interactions Four-Helix-Bundle (FHB) et Docking Domain (DD) entre dimères. . . . .	63

4.9	Traces des extensions de la molécule unique de nucléosome sous une force de traction de $f = 3.2$ pN pour les états où les SHLs $\pm 6.5, \pm 6.5, 5.5, \pm 6.5, 5.5, 4.5, \pm 6.5, 5.5, 4.5, 3.5, 2.5, 1.5$ sont dégraftés. L'extension en nm est représentée en ordonnée tandis que le temps de simulation est montré en abscisse. . . . .	64
4.10	Traces des extensions de la molécule unique de nucléosome sous une force de traction de $f = 3.2$ pN pour les états où les SHLs $\pm 6.5, \pm 6.5, 5.5, \pm 6.5, 5.5, 4.5, \pm 6.5, 5.5, 4.5, 3.5, 2.5, 1.5$ sont dégraftés. L'extension en nm est représentée en ordonnée tandis que le temps de simulation est montré en abscisse. . . . .	65
4.11	Tableau récapitulatif des paramètres des potentiels de Morse utilisés pour reproduire les Super-Helix-Location (SHL), FHB et DD. . . . .	66
4.12	Ces images illustrent les différentes conformations accessibles au nucléosome. La simulation est réalisée pour une force de traction de 3.2 pN et un couple nul. . . . .	67
4.13	Ces images illustrent les différentes conformations accessibles au nucléosome. La simulation est réalisée pour une force de traction de 1 pN et un couple de $\Gamma = 10$ pN · nm. Au fil des images, vous voyez le changement de chiralité de l'ADN nucléosomal. . . . .	68
4.14	Ces images illustrent les différentes conformations accessibles au nucléosome. La simulation est réalisée pour une force de traction de 1 pN et un couple de $\Gamma = 10$ pN · nm. Le couple est supérieur à la valeur critique de formation d'un réversome [6] et ainsi l'ADN commence à former des plectonèmes autour du réversome. . . . .	69
4.15	Vue schématique d'une fibre de chromatine de 20 nucléosomes sous une force de traction $f = 1$ pN et couple nul. Vous m'excuserez volontiers le cadrage quelque peu oblique mais nécessaire pour faire entrer la totalité des nucléosomes. . . . .	70
4.16	Vue schématique d'une fibre de chromatine sous une force de traction $f = 1$ pN et couple $\Gamma_n = 10$ pN · nm. Vous pouvez observer des nucléosomes et réversomes au sein de cette même fibre. La vidéo de la simulation est accessible à l'adresse <a href="http://vimeo.com/52072568">http://vimeo.com/52072568</a> . . . . .	71
5.1	Vue schématique d'une brosse de polymères (polymères greffés sur un substrat) où nous pouvons apprécier la structure en blobs. Cette image est empruntée à un exposé de E. Zhulina. . . . .	74
5.2	Le nombre de segments de Kuhn par blob est montré en échelle logarithmique $\log_{10} g$ (contour de couleurs) en fonction de la compaction $\alpha$ (bp/nm) (en ordonnée) et de la longueur de Kuhn de la fibre $K$ (nm) (en abscisse). . . . .	75
5.3	Les distances quadratiques $R_{x,y,z}^2(\Delta s)$ (nm) et $R^2(\Delta s)$ entre deux cylindres de fibre de chromatine (en ordonnée) sont montrées en fonction de la distance génomique $\Delta s$ (bp) les séparant (en abscisse) pour le chromosome 4. L'échelle est logarithmique en base 10. Les trois courbes correspondent respectivement à des diamètres de sphères pour le nucléole de 30, 100, 200 nm. . . . .	77
5.4	La position moyenne $\langle z(s) \rangle$ (en ordonnée) des cylindres est montrée en fonction de la distance génomique au centromère $s$ (en abscisse). Les trois courbes correspondent respectivement à des diamètres de sphères pour le nucléole de 30, 100, 200 nm. Nous pouvons remarquer qu'autour du centromère (soit au environ de 500 kbp), nous avons la loi d'échelle $\langle z \rangle \propto s$ . . . . .	78
6.1	. . . . .	89

6.2	Ces images illustrent les différentes conformations d'une fibre de chromatine (en vert) avec des interactions spécifiques (ADN-protéines) réparties le long de la chaîne. Les protéines (en bleu) interagissent elles aussi de manière spécifique. . . . .	91
6.3	Ces deux images illustrent le début et la fin d'une expérience de "HiC". Une fibre de chromatine représentée par 200 segments de longueur 60 nm et rayon 10 nm est simulée avec des segments pouvant interagir de manière attractive "crosslink" (en bleu) et des segments pouvant interagir seulement par volume exclu (en vert). Au tout début de la simulation le "crosslink" n'a pas eu le temps d'agir. À la fin de la simulation, les segments en bleu interagissent fortement entre eux. Il est clair que les interactions vues dans les expériences de "HiC" dépendent très clairement du temps d'observation. . . . .	92



---

# Acronymes

---

**ODE** Open Dynamics Engine

**ERP** Error Reduction Parameter

**CFM** Constraint Force Mixing

**ULC** Universal-Like-Chain

**FJC** Freely-Jointed-Chain

**WLC** Worm-Like-Chain

**GPLC** Gauss's Principle of Least Constraint

**LIBCCD** Library for collision detection between convex shapes

**BLC** Ball-Like-Chain

**CLC** Cardan-Like-Chain

**LCP** Linear-Complementarity-Problem

**SOR** Successive-Over-Relaxation

**SHL** Super-Helix-Location

**DD** Docking Domain

**FHB** Four-Helix-Bundle

**RW** Random Walk

**SAW** Self-Avoiding-Walk

**SPB** Spin-Pole-Body



---

# Remerciements

---

Je tiens à commencer par remercier Pascal Viot qui m'a accepté dans son Master puis accueilli dans son laboratoire. Je le remercie aussi, ainsi que Michel Quaggetto, pour avoir maintenu le parc informatique en état de fonctionner et de me permettre ainsi de réaliser toutes mes simulations.

Je fus très heureux de réunir les membres de mon jury et ce fut même un très grand honneur. Je les remercie d'avoir lu le manuscrit et je les remercie pour la discussion ayant suivi la soutenance.

Annick, les rares conversations que j'ai pu avoir avec toi m'ont montré une personne avec un savoir important et auprès de qui nous obtenons toujours une réponse à nos questions.

Fabien, je te dois des discussions intéressantes dont la précision et la recherche incessante de la meilleure et plus élégante description d'un phénomène physique m'ont beaucoup appris.

Jean-Marc, tu auras souvent dit que tu ne m'as jamais rien appris si ce n'est le goût du "Bagelstein". Ceci n'est pas tout à fait vrai. Grâce à toi, il m'a été possible de coopérer avec des gens de différentes disciplines et ainsi apprendre, obtenir des connaissances bien plus universelles.

Julien, je te suis reconnaissant de m'avoir apporté le côté biologique indispensable à l'étude mécanique du nucléosome.

Maria, la justesse et la gentillesse dont tu fais preuve me seront toujours appréciables.



---

# CHAPITRE 1

---

## Introduction

---

En 1953, J. D. Watson et F. H. C. Crick publient leur découverte de la structure en double hélice de l'ADN dans *Nature*. Cette molécule qui contient toute l'information génétique de tout être vivant est remarquablement longue (de l'ordre de 1 m totalement dépliée) et chose encore plus remarquable est contenue dans le noyau des cellules dont le diamètre est environ égal à 10  $\mu\text{m}$ . En 1992, pour comprendre le comportement élastique de l'ADN, C. Bustamante et son équipe sont les premiers à exercer des contraintes mécaniques sur l'ADN et à avoir accès à son comportement élastique. Quelques années plus tard A. Bancaud ainsi que C. Bustamante, remplacent l'ADN par de la fibre de chromatine (assemblage de nucléosomes qui sont eux-mêmes constitués de protéines et d'ADN) et étudient son comportement élastique et plastique sous pinces magnétiques. Cet assemblage constitue à plus grande échelle nos chromosomes. Récemment, l'équipe de D. Bensimon et V. Croquette obtient de jolis résultats encore non expliqués sur des expériences de nucléosome sous traction. Ces prouesses techniques donnent accès à des moyennes statistiques d'observables telles que la taille des molécules étudiées lorsque qu'elles sont soumises à des efforts de traction et de torsion. Cependant, ces observables ne peuvent nous donner des images des structures obtenues et elles ne sont généralement pas suffisantes pour comprendre les comportements sous-jacents de ces assemblages complexes. La théorie échoue elle aussi mais fournit de jolis modèles prédictifs reproduisant "tous" les résultats expérimentaux sur de l'ADN nu. Nous pouvons notamment saluer les travaux de J. Marko, E. Siggia, P. Nelson, J. D. Moroz, S. Neukirch, C. Bouchiat, M. Mézard, avec nos excuses pour ceux oubliés.

Pour parfaire à merveille le duo expérience-théorie, il me semble indispensable de donner une place à part entière à la simulation. Depuis une trentaine d'années, des personnes comme S. Allison savent brillamment simuler de l'ADN circulaire. A. Vologodskii et J. Marko savent quant à eux reproduire les expériences d'ADN sous traction en y ajoutant même des contraintes de torsion. J. Langowski développe des méthodes sophistiquées pour la simulation de polymères chargés. De ces expériences numériques, il est aisé d'obtenir des graphiques, réponses de ces molécules d'ADN à différents types d'efforts mais aussi des images qui sont parfois plus instructives que des valeurs moyennes. Malheureusement, peu de modélisations sont faites sur la fibre de chromatine. Nous

---

pourrions imaginer une dynamique au niveau moléculaire incluant tous les atomes d'une fibre de quelques dizaines de nucléosomes mais la complexité et la tâche me paraissent bien insurmontables. Pour simplifier la modélisation nous pouvons passer de l'échelle atomique à une échelle gros grain et pourquoi pas remplacer quelques paires de bases d'ADN par un cylindre de longueur équivalente. Par conséquent, nous faisons de même pour les protéines de notre fibre de chromatine. Bien-entendu, ce genre de modélisation existe déjà et elle est aussi relativement complexe. En particulier, comment articuler, mouvoir des structures aussi compliquées et surtout éviter que les éléments de formes variées puissent s'interpénétrer ?

Les méthodes de simulations en physique, aussi sophistiquées soient-elles, sont inadaptées pour ce problème. Il est donc nécessaire de regarder ce que proposent les autres disciplines, en particulier, la mécanique et la robotique.

Il me semble que tout le monde a eu l'occasion d'admirer le réalisme graphique mais surtout physique des derniers jeux vidéos. Les simulations de courses sont le meilleur exemple : collisions avec l'environnement lorsque nous ne savons pas conduire, collisions avec les autres voitures si l'on se pense dans une compétition de NASCAR, dérapages, suspensions pour les rallyes... Quels types de programmes informatiques, de mécanique se cachent derrière cet exemple ou derrière le fonctionnement des robots industriels ou des animaux mécaniques de compagnie ?

Première réponse : la cinématique inverse permet à une machine équipée d'un bras articulé de trouver la configuration des articulations pour se saisir d'une portière de voiture et la monter ensuite sur les gonds correspondants ; ce que fait notre cerveau en permanence pour appuyer sur les touches du clavier. Un joint moteur permet de déplacer le bras et ses différentes parties. La beauté de la cinématique inverse réside dans le fait que quel que soit le type de joint, le bras trouvera sa cible. La formulation mathématique est simple et la mise en œuvre à la portée de tous grâce aux puissantes librairies informatiques.

Seconde réponse : les moteurs physiques, librairies achevées et constamment mises à jour incluant systèmes articulés par des joints mécaniques et détecteurs de collisions. Ils implémentent de la même manière, avec une même formulation mathématique, les joints rotules, de Cardan, charnières, homocinétiques, prismatiques aussi bien que les joints de contact. Il est alors aisé de construire une structure compliquée et de l'animer par la suite. Un détecteur de collisions permet aux divers éléments de la structure de ne pas se superposer.

Pour conclure, je propose d'utiliser ces outils pour modéliser les expériences de pinces optiques et magnétiques. Je commencerai par une brève revue des méthodes traditionnelles de simulations en physique des polymères, qui me permettra d'introduire le formalisme des moteurs physiques et de décrire notamment [ODE](#) moteur de jeux sur lequel j'ai jeté mon dévolu. En particulier, je propose de vous montrer comment construire un modèle mécanique et articulé d'ADN (à partir de joints mécaniques) et comment introduire ses propriétés élastiques. J'exposerai aussi comment simuler l'agitation thermique pour des corps rigides et notamment réaliser une optimisation des méthodes traditionnelles. Une calibration préalable sera exposée sur des confrontations des résultats expérimentaux avec des résultats de simulations sur de l'ADN nu pour convaincre de l'utilité de l'outil pour modéliser la fibre de chromatine. Je montrerai aussi qu'il est même possible de faire des prédictions sur la réponse de l'ADN à un effort de torsion. Je changerai alors d'échelle et reproduirai les expériences de nucléosome unique en traction et torsion sous pinces magnétiques du groupe de D. Bensimon et V. Croquette. Grâce aux images des mes simulations, j'espère apporter un éclairage nouveau sur les expériences. A cette occasion, je présenterai un calcul de structure sur le nucléosome,

calcul “analogue” aux calculs de structure des ponts. J'obtiendrai alors les interactions entre l'ADN et les protéines et les interactions entre protéines. J'utiliserai ce modèle pour comprendre comment réagit la fibre de chromatine à des contraintes de torsion et ce aux travers d'images agrémentées de courbes. Pour achever la description des différentes échelles au sein du noyau, je parlerai d'un calcul de physique des brosses de polymères afin de comprendre l'organisation des chromosomes.



---

# Moteurs physiques et dynamique de Langevin pour l'ADN

---

À travers ce chapitre, je vais mêler la description des expériences de molécules uniques sur de l'ADN nu et les différentes simulations qui se sont attachées à reproduire les résultats expérimentaux. J'en profiterai aussi pour donner une revue des algorithmes de simulation de polymères, les critiquer, présenter leurs limites et ainsi, introduire peu à peu l'intérêt que j'ai pour les moteurs physiques.

## 2.1 Méthodes traditionnelles de simulation de l'ADN

### 2.1.1 Méthode du Pivot, Krankshaft et le Bead spring model

Une première tentative de simulation d'ADN est réalisée avec le travail de A. Vologodskii [7]. Ce travail, daté de 1994, fait suite aux expériences effectuées dans l'équipe de C. Bustamante [8] en 1992. Dans cette expérience une molécule d'ADN de longueur  $L = 94000$  paires de bases est attachée à une extrémité sur une surface de verre et, est greffée à l'autre extrémité sur une bille magnétique. À l'aide d'un champ magnétique, il est possible d'exercer une force  $f$  sur la bille et de mesurer ensuite la position de cette dernière. L'ADN se trouvant dans un bain thermique, d'unité d'énergie d'agitation thermique  $\beta = 1/k_B T$ , la position de la bille fluctue et c'est cette réponse qui est utilisée pour calculer la valeur moyenne de l'extension de l'ADN en fonction de la force appliquée. La contrainte de traction sur la bille permet d'accéder au comportement élastique de l'ADN et, en particulier, d'extraire une longueur de persistance en courbure, capacité de l'ADN à se plier. Dans l'article [8], les auteurs utilisent un modèle théorique de physique des polymères, le modèle dit "Freely-Jointed-Chain" <sup>1</sup>, qui nous donne la réponse d'une chaîne trempée dans un bain thermique et sur laquelle on exerce une force de traction sous la forme d'une fonction de Langevin  $\langle z \rangle / L = \mathcal{L}(\beta f p)$ . Ce modèle

---

1. Je choisis de ne pas traduire et de garder ces anglicismes. Ils sont d'ailleurs résumés à la suite de la table des matières.

de polymère décrit une chaîne de segments sans volume exclu et connectés les uns aux autres sans restrictions sur les orientations. Les segments ont une taille de deux fois la longueur de persistance, cette longueur de Kuhn est notée  $k = 2p$ . Ils montrent, en particulier, que ce modèle ne fonctionne que pour de faibles ou de très grandes forces.

Dans l'optique d'améliorer la description de l'ADN en terme de physique des polymères, A. Vologodskii [7] propose d'utiliser le modèle du "Worm-Like-Chain" introduit par Kratky-Porod [9]. Pour ce faire, il remplace les segments de Kuhn par des segments de taille inférieure  $l < k$  et introduit une énergie de courbure  $E_b$ . La constante de courbure est alors fonction de la taille des segments. Il introduit ainsi une restriction sur les orientations de deux segments successifs. Ensuite, pour tenir compte des charges portées par l'ADN et des ions du solvant, il utilise un potentiel de Debye-Hückel  $E_{el}$  valable dans la limite des faibles concentrations en sel NaCl. Cette énergie est compliquée et j'en parlerai plus tard. Pour finir, il ajoute l'énergie  $E_f$  due à la traction de l'ADN.

L'énergie associée à une conformation de l'ADN est maintenant connue; reste alors à générer d'autres configurations accessibles d'un point de vue critère énergétique. Pour ce faire, A. Vologodskii utilise deux méthodes schématisées sur la figure (2.1). La première méthode, le Pivot, consiste à sélectionner aléatoirement un point  $i$  sur la chaîne et faire subir une rotation  $\mathcal{R}_u(\theta)$  à tous les points d'indices  $j > i$ . La rotation se fait autour d'un axe aléatoire  $u$  et d'un angle  $\theta$  lui aussi aléatoire. Une seconde méthode, Krankshaft ou torsion concertée, consiste à choisir deux points sur la chaîne, calculer l'axe passant par ces deux points  $v$ , choisir un angle  $\phi$  de façon aléatoire et faire subir à la partie de la chaîne entre  $i$  et  $j$  la rotation  $\mathcal{R}_v(\phi)$ . Dès qu'une nouvelle configuration est obtenue, elle est acceptée ou refusée suivant le critère de Metropolis (voir l'annexe (A.2)). À la fin de la simulation on obtient une pléthore de configurations accessibles à l'ADN et il ne reste plus qu'à moyenner les observables d'intérêt et comparer aux résultats expérimentaux.

Grâce à cette méthode, A. Vologodskii montre que dans la limite où la taille des segments devient nulle, la simulation retrouve le profil force-extension obtenu expérimentalement. Cette simulation est très jolie et facilement implémentable. Les temps de calcul ne sont pas démesurés : il m'a fallu une demi-journée pour reproduire la courbe force-extension (voir la section (2.9)) avec cette méthode. Imaginons maintenant que l'on veuille, en simulation, greffer l'ADN dans un nano-tube et reproduire les expériences de B. Austin [10]. Si l'on utilise la méthode du Pivot ou celle du Krankshaft, il est fort probable de générer beaucoup de configurations où une partie de l'ADN va se retrouver en dehors du tube. Et cela sera d'autant plus fréquent que le nano-canal est peu large. Nous pourrions aussi vouloir fixer des protéines sur une partie de la chaîne et là encore la méthode du Pivot va générer des configurations où quelques brins d'ADN vont se trouver à l'intérieur de la protéine. Je conclus de ces deux exemples que le Krankshaft et le Pivot sont uniquement adaptés à des situations d'un ADN unique et nu sans confinement. Je profite de ce que je viens de dire pour citer le travail de H. Jian [11] qui a développé une méthode hybride utilisant la méthode du Pivot et une dynamique brownienne sur un "Bead-Spring-Model".

Le "Bead-Spring-Model" utilise un système de  $N$  billes de masse  $m$  et de rayon  $a$  connectées les unes aux autres par des ressorts. Ces billes servent de points virtuels pour les interactions avec le solvant. Les ressorts, eux, représentent les effets entropiques, degrés de libertés internes, que l'on a perdus dans la procédure de "coarse-graining". Cette procédure permet de réduire l'ADN à une chaîne de segments. Pour animer cette chaîne, on utilise généralement une dynamique de Langevin incluant les effets hydrodynamiques et les forces stochastiques du bain thermique. La prise en compte des effets hydrodynamiques est compliquée et très bien détaillée dans [12, 11, 13]. Dans la plupart des cas, les

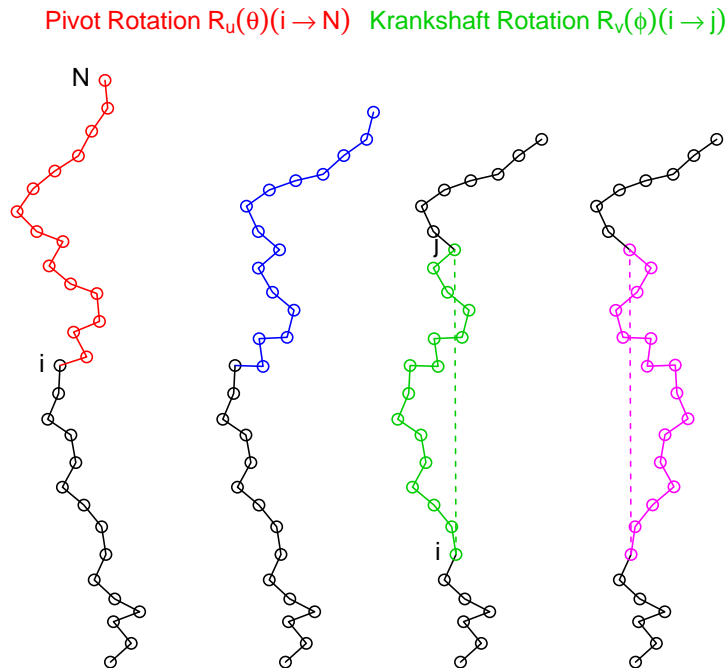


FIGURE 2.1: Vue schématique de la méthode du Pivot et de la méthode du Krankshaft. Une partie de la chaîne (rouge) est sélectionnée. Un axe aléatoire et un angle aléatoire composent une rotation qui amène la partie rouge sur la partie bleue (méthode du Pivot). Une partie de la chaîne (verte) est sélectionnée. Cette partie est amenée sur la partie magenta par une rotation d'axe reliant les deux extrémités de cette chaîne et d'angle  $\phi$  aléatoire (méthode du Krankshaft).

auteurs simplifient l'équation de la dynamique en utilisant l'argument suivant : le terme inertiel varie lentement devant le terme de friction (simulation "overdamped"). Le cas inverse ("underdamped") est décrit dans [14, 15]. Nous sommes maintenant dotés de l'agitation thermique et nous voulons ajouter la persistance en courbure et en twist de l'ADN sur ce modèle. Pour ce faire [12] développe quelques pages de calcul qui nuisent à la simplicité de la mise en œuvre d'une telle simulation. Qui plus est, ce système n'intègre pas de détections de collision et surtout ne nous permet pas de décrire différentes formes. Il faut porter une attention particulière au choix de la constante de raideur du ressort. Une raideur trop petite nous donnera des paires de bases bien trop étirées et une raideur trop grande, des forces de rappels bien trop importantes. Le choix du pas de temps de simulation est alors crucial ; restriction que je n'ai guère envie d'avoir à prendre en compte !

### 2.1.2 SHAKE et QSHAKE

Dans une dynamique moléculaire ou une dynamique de Langevin sur "Bead-Spring-Model", le pas de temps est limité par les oscillations des liaisons entre atomes. L'amplitude des vibrations est faible avec une fréquence élevée. L'idée de SHAKE [16] est de remplacer ces liaisons élastiques par des liaisons rigides. Ces "bâtons" se présentent sous la forme de contraintes holonomiques ; ces



contraintes ne dépendent pas de la vitesse des particules mais seulement de la géométrie du système. Si le cœur vous en dit, une introduction très formelles sur les types de contraintes (holonomique, non-holonomique, scléronomique, rhéonomique...) peut se trouver dans la thèse de C. Lacoursière. En particulier, elles assurent que l'atome  $i$  se trouve toujours à la même distance  $d$  de l'atome  $i + 1$  soit  $\delta = (\mathbf{r}_i - \mathbf{r}_j)^2 - d^2 = 0$ . Cet algorithme est surtout utilisé pour de grosses molécules tandis que pour les petites molécules, il est plus judicieux de choisir SETTLE [17]; SETTLE résoud analytiquement les contraintes. La première étape de SHAKE [16] est de faire évoluer le système selon une dynamique moléculaire standard, avec l'algorithme Leapfrog (voir la section (A.1)) et sans faire intervenir de contraintes. Les positions des atomes  $\mathbf{r}$  deviennent  $\mathbf{r}(t + \Delta t)$  et le résultat obtenu au temps  $t + \Delta t$  ne satisfait sûrement pas les contraintes. Pour pallier cette situation, il faut ajouter des forces de corrections  $\mathbf{f}_\delta$  dans la direction de ou des contraintes  $\mathbf{d}_{ij} = \mathbf{r}_i - \mathbf{r}_j$  au temps  $t$ . Ces forces servent alors à recalculer les positions des atomes  $\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i + \frac{\Delta t^2}{m_i} \sum_\delta \mathbf{f}_\delta$ . Ces nouvelles positions permettent d'estimer à nouveau les erreurs sur les contraintes au temps  $t + \Delta t$  et si ces erreurs ne sont pas suffisamment petites, nous recommençons à calculer des forces de corrections jusqu'à satisfaire une certaine précision  $\delta(t + \Delta t) < 10^{-\epsilon}$  sur la totalité des contraintes. En pratique, on utilise des valeurs  $\epsilon \leq 4$  pour éviter de grandes fluctuations dans l'énergie du système [16, 18, 19]. Par exemple, pour une molécule d'octane dont on représente tous les atomes, il faut entre 40 et 60 itérations pour satisfaire une précision de  $\epsilon = 4$ . J'ai comparé SHAKE et ODE sur une chaîne de 50 à 500 billes connectés par des liaisons rigides. En imposant la précision obtenue en utilisant ODE, SHAKE prend un temps de 2 à 3 fois plus important qu'ODE pour résoudre les contraintes. La référence [18] me permet la transition sur QSHAKE; extension de SHAKE à des corps rigides. Forester et Smith décrivent un corps rigide à l'aide d'un ensemble de sites de masse  $m$ . De là, ils peuvent écrire la traditionnelle matrice d'inertie du corps  $\mathcal{I}$  et en extraire un repère matériel ou repère local associé au corps rigide. Dans ce repère, décrit par un quaternion  $\mathbf{q}$ , la matrice d'inertie est diagonale. La procédure pour satisfaire les contraintes est identique à celle de SHAKE à l'exception qu'il faut maintenant itérer les quaternions et introduire des couples de corrections  $\mathbf{\Gamma}_\delta$ . L'avantage de cette méthode est que l'on gagne un facteur 4 sur le nombre d'itérations à précision fixée. Cependant, il est difficile de représenter un cylindre avec un ensemble de points et de construire alors une chaîne de cylindre. Le nombre de points à prendre en compte devient alors énorme. L'avantage d'ODE est de remplacer un cylindre par un repère matériel et de réduire ainsi le nombre de degrés de libertés. Ce "coarse-graining" peut d'ailleurs être fait dans ODE sur toutes les géométries que nous pouvons imaginer. ODE est aussi capable de modéliser les collisions entre géométries. Nous comprenons déjà bien l'avantage d'un moteur physique sur les méthodes traditionnelles dérivées de SHAKE. Un schéma récapitulatif des différentes étapes de SHAKE est montré sur la figure (2.2). SHAKE peut être adapté [20, 21, 19] pour faire de la dynamique Brownienne (dynamique sur-amortie) ou de la dynamique de Langevin (dynamique sous-amortie). Il existe des optimisations de SHAKE comme celle introduite par B. Hess [19]. Cet algorithme, nommé LINCS, commence par faire un pas de simulation non contraint. Il introduit ensuite une matrice contenant les contraintes du système qu'il approxime par un développement en série et qu'il utilise alors, en l'inversant, pour corriger les erreurs. Pour finir, je cite le travail de D. Edberg [22], qui, utilise le Gauss's Principle of Least Constraint (GPLC) pour calculer les forces issues des contraintes. Ce principe est équivalent au principe de d'Alembert et l'analogue du principe de moindre action de W. R. Hamilton. Cependant, le principe de Gauss de moindre action minimise la trajectoire d'un système mécanique contre une maximisation de la trajectoire pour le principe de W. R. Hamilton. Cette manière de procéder se rapproche des moteurs

de jeux. Tous ces algorithmes travaillent sur des systèmes de billes connectées par des liaisons rigides, et, à ma connaissance, il n'existe pas de développement pour des contraintes plus complexes. En particulier, l'introduction de volume exclu sur ces liaisons rigides est une tâche ardue notamment dans le calcul des forces de répulsions au niveau des billes. La réalisation du volume exclu se fait toujours par des potentiels à deux corps. Pour répondre à tous ces problèmes, je propose de passer à la description des moteurs physiques.

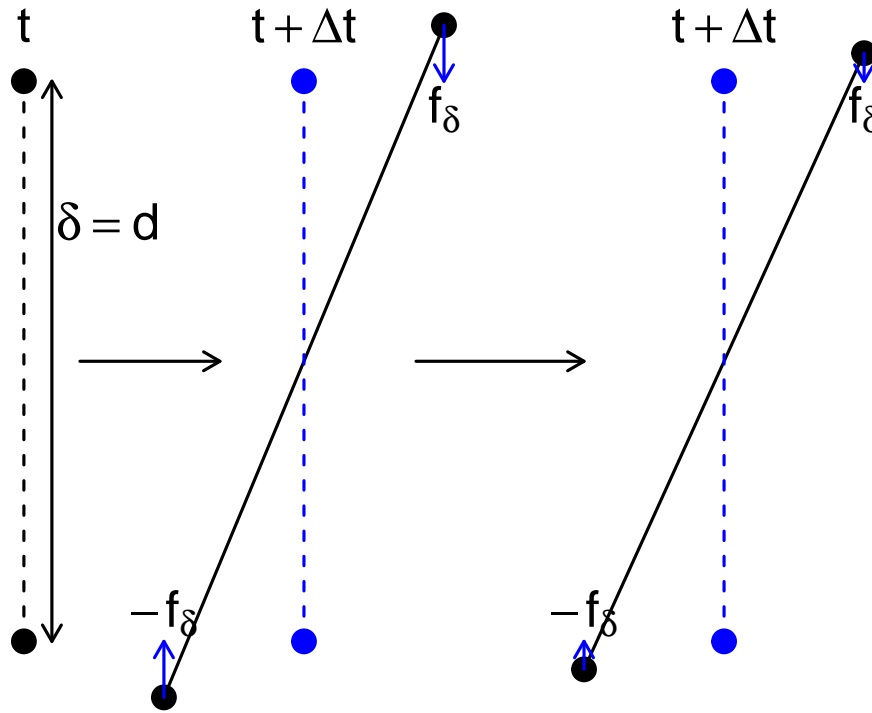


FIGURE 2.2: Vue schématique des différentes étapes de l'algorithme SHAKE pour un système de deux particules liées rigides  $\delta = d$ . Une première étape consiste à mener le système au pas  $t + \Delta t$  sans tenir compte des contraintes. Les positions obtenues ne satisfont pas la contrainte et les positions sont corrigées de manière itérative en utilisant des forces de corrections  $f_\delta$  montrées en bleu.

## 2.2 Formalisme des moteurs physiques

Je propose d'introduire un tant soit peu de formalisme très utile pour les moteurs de jeux et la dynamique de Langevin. Nous voulons modéliser l'ADN comme un polymère lui-même construit par un certain nombre  $N$  de segments. Pour caractériser ce système, il nous faut la donnée des positions et des quaternions, qui sont normalisés à 1, de chacun des corps rigides que nous regroupons au sein d'un même vecteur  $\mathcal{S} \in \mathbb{R}^{7N}$ . Dans la suite, j'opte pour la convention  $\mathbf{X}(t) \equiv \mathbf{X}$ , et je préciserai lorsque les quantités ne seront pas prises au temps  $t$ . La donnée d'un quaternion est équivalente à la donnée d'un référentiel matériel  $t_i, m_i, \tilde{m}_i$  associé à chaque corps solide. Nous procédons de la

même manière pour le torseur cinématique  $\mathcal{V} \in \mathbb{R}^{6N}$  formé par les vitesses linéaires  $v$  et angulaires  $\omega$ . L'évolution du référentiel matériel est alors donnée par  $\dot{t} = \omega \otimes t$ ,  $\dot{m} = \omega \otimes m$  et  $\dot{\tilde{m}} = \omega \otimes \tilde{m}$ . Nous définissons alors un torseur dynamique  $\mathcal{F}_e \in \mathbb{R}^{6N}$  construit avec les forces et couples externes appliqué au système. En faisant apparaître la matrice diagonale par blocs  $\mathcal{M} \in \mathbb{R}^{6N \times 6N}$  qui donne la masse et la matrice d'inertie de chaque corps, nous pouvons écrire les équations de la dynamique (2.1) où le moment généralisé est  $\mathcal{L} = \mathcal{M}\mathcal{V}$ . Dans ces équations, la matrice  $\mathcal{Q}_i \in \mathbb{R}^{4 \times 3}$  permet de relier quaternion et vitesse angulaire. Enfin, j'introduis une nouvelle matrice diagonale par blocs  $\mathcal{P} = \{\mathbf{1}, \mathcal{R}_1 \dots \mathbf{1}, \mathcal{R}_N\} \in \mathbb{R}^{6N \times 6N}$  où la matrice de rotation  $\mathcal{R}_i$  permet de passer de la description d'un corps dans son référentiel principal à la description de celui-ci dans le référentiel du laboratoire. Nous avons notamment  $\mathcal{R}_i \mathcal{R}_i^T = \mathbf{1}$ ,  $t_i^* = \mathcal{R}_i^T t_i = \{0, 0, 1\}$ ,  $m_i^* = \mathcal{R}_i^T m_i = \{1, 0, 0\}$  et  $\tilde{m}_i^* = \mathcal{R}_i^T \tilde{m}_i = \{0, 1, 0\}$ . La matrice  $\mathcal{P}$  possède la propriété suivante  $\mathcal{P}\mathcal{P}^T = \mathbf{1}_{6N}$ . En particulier, nous pouvons écrire  $\mathcal{M} = \mathcal{P}\mathcal{M}^*\mathcal{P}^T$  où  $\mathcal{M}^*$  est diagonale et contient la masse et les moments principaux d'inertie. Le torseur cinématique se décompose alors comme  $\mathcal{V} = \mathcal{P}\mathcal{V}^*$  où  $\star$  dénote l'écriture d'un vecteur ou d'une matrice dans le référentiel matériel du corps associé. Cette écriture peut aisément s'étendre à toutes les matrices décrivant les propriétés du système de corps rigides.

$$\begin{aligned}
 \mathcal{S} &= \{r_1, q_1 \dots r_N, q_N\}^T \\
 \mathcal{V} &= \{v_1, \omega_1 \dots v_N, \omega_N\}^T \\
 \mathcal{F}_e &= \{f_1, \Gamma_1 \dots f_N, \Gamma_N\}^T \\
 \mathcal{Q} &= \{\mathbf{1}, \mathcal{Q}_1 \dots \mathbf{1}, \mathcal{Q}_N\} \\
 \mathcal{M} &= \{m_1, \mathcal{I}_1 \dots m_N, \mathcal{I}_N\} \\
 \dot{\mathcal{S}} &= \mathcal{Q}\mathcal{V} \\
 \dot{\mathcal{L}} &= \mathcal{F}_e \text{ soit } \mathcal{M}\dot{\mathcal{V}} = \mathcal{F}_e - \dot{\mathcal{M}}\mathcal{V}
 \end{aligned} \tag{2.1}$$

La dérivée temporelle de la matrice d'inertie est, en général, non nulle pour des corps rigides à l'exception des sphères. Le terme  $\dot{\mathcal{L}}$  contient donc deux termes, l'un non linéaire  $\dot{\mathcal{M}}\mathcal{V}$ , terme dit de Coriolis, représente le couple inertiel et le terme d'accélération  $\mathcal{M}\dot{\mathcal{V}}$ . Le terme de Coriolis ne contient aucune information sur les vitesses translationnelles et s'écrit comme  $(\mathcal{I}\omega) \otimes \omega$ . Maintenant, il nous faut articuler ce système mais aussi empêcher les corps solides de se recouvrir. Commençons donc par décrire la dynamique sous contraintes de manière générale [23, 24] pour ensuite discuter de son implémentation dans ODE [25, 26]. Nous voulons que chaque segment soit connecté à ses deux extrémités par deux autres segments et nous faisons abstraction du type de joint mécanique utilisé ; enfin pour l'instant. Introduisons la notion de contrainte holonomique entre deux corps qui est une fonction de leurs positions, quaternions et du temps  $\delta(s, t)$  et égale à zéro quelque soit  $s$  et  $t$ . Il est important de noter qu'aucune vitesse n'entre en compte dans l'équation de  $\delta$ . En dérivant par rapport au temps nous obtenons  $\mathcal{J}\mathcal{V} = 0$  où la matrice  $\mathcal{J}$  contient toutes les informations relatives aux contraintes présentes dans le système et s'appelle le Jacobien des contraintes. Cette équation nous dit simplement que les points d'ancrage se déplacent à la même vitesse. Les dimensions du Jacobien sont données par le nombre de corps solides connectés et le nombre de contraintes holonomiques  $n$  soit  $\mathcal{J} \in \mathbb{R}^{6N \times n}$ . Il nous reste alors à construire le torseur dynamique  $\mathcal{F}_c$  qui maintient les contraintes holonomiques lors d'un changement  $\delta s$  du vecteur  $s$ . Nous demandons, en particulier, que  $\mathcal{F}_c$  ne travaille pas  $\mathcal{F}_c^T \mathcal{V} = 0$  ; ce qui nous permet d'écrire  $\mathcal{F}_c = \mathcal{J}^T \lambda$  où  $\lambda$  est un vecteur de multiplicateurs de Lagrange de dimension  $n$ . L'équation  $\mathcal{F}_c^T \mathcal{V} = 0$  doit se comprendre globalement :

c'est le système dans sa totalité qui satisfait cette équation et non un sous-système composé d'un certain nombre  $x < N$  de corps rigides. Il est important de rappeler que le torseur dynamiques des contraintes n'ajoute et n'enlève pas d'énergie au système ; il distribue juste cette énergie entre les différents degrés de libertés non contraints.

## 2.3 Implémentation des joints mécaniques dans ODE

La théorie exposée jusqu'à maintenant est élégante et tient en peu de lignes ; l'implémentation dans ODE est une tâche un peu plus ardue. Ce que je présente dans la suite est basé sur son code source.

Toute simulation est intrinsèquement imparfaite mais nous pouvons toujours y remédier ; il est évident que notre système articulé ne va pas vérifier exactement  $\delta(\mathcal{S}, t) = \mathbf{0}$  et  $\mathcal{J}\mathcal{V} = 0$  à chaque itération. Les erreurs numériques dans la détermination de  $\lambda$  et la discrétisation des équations du mouvement font que les corps liés mécaniquement vont s'éloigner progressivement et les positions d'ancrages ne seront plus respectées. J'ai constaté que les principales sources d'erreurs numériques viennent de l'intégration des quaternions. En particulier, après un pas de temps les quaternions ne sont plus normalisés. ODE propose d'introduire un paramètre de contrôle sans dimension ERP qui mesure de combien l'erreur sur les positions d'ancrages doit être corrigée lors du pas de simulation suivant. Une valeur de 0 indique qu'aucune correction ne sera faite tandis qu'une valeur de 1 fixera l'erreur en un seul pas de simulation. Ce paramètre ne correspond pas à un ressort entre points d'ancrages mais à un torseur dynamique  $\mathcal{F}_{erp}$  qui ajuste le torseur cinématique  $\mathcal{V}$  pour éliminer l'erreur en un pas de temps  $\Delta t$ . Un ressort, lui, ajuste l'accélération pour éliminer l'erreur. ODE propose aussi d'introduire une matrice diagonale et positive  $\mathcal{K}_{cfm}$  pour écrire  $\mathcal{J}\mathcal{V}(t + \Delta t) = -\frac{k_{erp}}{\Delta t}\delta - k_{cfm}\lambda$  et permettant ainsi de violer les contraintes d'une quantité proportionnelle au multiplicateur de Lagrange  $\lambda$ . Le coefficient  $k_{cfm}$  possède les dimensions d'un temps divisé par une masse. En écrivant cette équation au temps  $t + \Delta t$  ODE calcule le torseur dynamique des contraintes de manière implicite ; l'intégration des contraintes est stable (voir l'annexe (A.1)). Il reste à déterminer l'inconnue  $\lambda$ . En discrétisant au premier ordre l'équation de la dynamique  $\dot{\mathcal{L}} = \mathcal{F}_e + \mathcal{F}_c$  et en y introduisant les conditions cinématiques augmentées par les termes ERP et CFM, nous trouvons l'équation (2.2) à résoudre pour  $\lambda$  soit un problème du type  $\mathcal{A}\lambda = b$ .

$$\left[ \mathcal{J}\mathcal{M}^{-1}\mathcal{J}^T + \frac{k_{cfm}}{\Delta t} \right] \lambda = -\frac{k_{erp}}{\Delta t^2}\delta - \mathcal{J} \left[ \frac{\mathcal{V}}{\Delta t} + \mathcal{M}^{-1}(\mathcal{F}_e - \dot{\mathcal{M}}\mathcal{V}) \right] \quad (2.2)$$

Maintenant que nous connaissons  $\mathcal{F}_c$ , nous l'introduisons dans l'équation de la dynamique qui est alors intégrée avec un algorithme d'Euler semi-implicite (schéma d'intégration symplectique voir équation (2.3) et l'annexe (A.1)). Pour intégrer les quaternions au cours du temps, ODE utilise l'algorithme dit "Exponential Map" [27]. Nous obtenons finalement l'état du système au temps  $t + \Delta t$ .

$$\begin{aligned} \mathcal{V}(t + \Delta t) &= \mathcal{V} + \Delta t \mathcal{M}^{-1}(\mathcal{F}_e - \dot{\mathcal{M}}\mathcal{V} + \mathcal{J}^T \lambda) \\ \mathcal{S}(t + \Delta t) &= \mathcal{S} + \Delta t \mathcal{Q}\mathcal{V}(t + \Delta t) \end{aligned} \quad (2.3)$$

L'ajout de  $k_{cfm}$  sur la diagonale de la matrice  $\mathcal{J}\mathcal{M}^{-1}\mathcal{J}^T$  permet de rendre la matrice  $\mathcal{A}$  symétrique définie positive. En particulier,  $k_{cfm}$  permet une meilleure précision dans la résolution du

système lorsque celui-ci est proche d'une "singularité". Il faut ensuite résoudre l'équation pour  $\lambda$  ce que ODE propose de faire avec deux algorithmes différents : Successive-Over-Relaxation (SOR) [28] et Linear-Complementarity-Problem (LCP) [29]. LCP est plus coûteux en temps de calcul : de l'ordre de  $n^3$  et un coût en mémoire de  $n^2$  [25]. SOR est de l'ordre de  $ns$  avec  $s$  représentant le nombre de SOR et de l'ordre de  $n$  en mémoire [25]. Cependant LCP est plus précis que SOR.

Ce type d'algorithme se nomme "velocity-based-algorithm" et est utilisé dans la plupart des moteurs de jeux actuels [26]. En 2007, M. Müller [30] décrit un algorithme "position-based-algorithm". Il existe des versions "force-based-algorithm", un très beau travail se trouve chez David Baraff [31], qui utilise l'accélération des contraintes  $\dot{\mathcal{J}}\mathcal{V} + \mathcal{J}\dot{\mathcal{V}} = \mathbf{0}$  pour obtenir une équation (2.4) sur  $\lambda$  :

$$\dot{\mathcal{J}}\mathcal{V} + \mathcal{J}\mathcal{M}^{-1}(\mathcal{J}^T\lambda + \mathcal{F}_e - \dot{\mathcal{M}}\mathcal{V}) = \mathbf{0} \quad (2.4)$$

Pour s'affranchir des erreurs numériques et donc des erreurs au niveau des points d'ancrage des joints, dans le cas "force-based-algorithm", il est possible d'utiliser la stabilisation de Baumgarte [32]. Elle consiste à ajouter un ressort amorti de raideur  $k_s$  et d'amortissement  $k_d$  dans le membre de droite de l'équation sur l'accélération de la contrainte :

$$\dot{\mathcal{J}}\mathcal{V} + \mathcal{J}\dot{\mathcal{V}} = -k_s\delta - k_d\mathcal{J}\mathcal{V} \quad (2.5)$$

Ce ressort amorti porte sur les contraintes  $\delta$  et leurs vitesses  $\dot{\delta} = \mathcal{J}\mathcal{V}$ .

## 2.4 Gestion des collisions

Nous venons d'achever la description des systèmes articulés et nous en venons alors au second ingrédient faisant d'ODE un outil merveilleux. La plupart des simulations actuelles utilisent des potentiels à deux corps pour éviter que deux particules ne se recouvrent. L'idée est bonne mais impossible à mettre en œuvre pour des géométries complexes. ODE propose d'utiliser une librairie de détection de collisions [33] pour des géométries complexes Library for collision detection between convex shapes (LIBCCD). Une fois le point de contact déterminé  $\mathbf{r}_c$  ainsi que la normale au contact  $\mathbf{n}_c$ , ODE crée un joint de contact. Ce joint nous dit que les vitesses au point de contact définies respectivement sur le premier et le second corps, doivent être les mêmes selon la normale au contact. Cette condition se met en équation (2.7). Le multiplicateur de Lagrange  $\lambda$  associé à cette contrainte est naturellement orienté selon la normale  $\mathbf{n}_c$ . La description des joints de contacts est identique à la description des joints mécaniques ; ainsi l'intégralité des contraintes est contenue dans le Jacobien  $\mathcal{J}$ . Dès qu'ODE connaît l'information  $\mathcal{J}$  il ne reste plus qu'à appliquer l'algorithme SOR ou LCP pour trouver  $\lambda$  et faire avancer le système au temps  $t + \Delta t$ . Je tiens à insister sur le génie de cette gestion des collisions qui revient à introduire un joint mécanique en chaque point de contact ; ODE est une librairie regroupant deux outils d'une puissance incroyable au sein d'une unique description.

$$\mathbf{n}_c^T (\mathbf{r}_{c1} - \mathbf{r}_{c2}) = 0 \quad (2.6)$$

$$\mathbf{n}_c^T (\mathbf{v}_1 + \mathbf{w}_1 \otimes (\mathbf{r}_{c1} - \mathbf{r}_{cm1}) - \mathbf{v}_2 - \mathbf{w}_2 \otimes (\mathbf{r}_{c2} - \mathbf{r}_{cm2})) = 0 \quad (2.7)$$

Passons maintenant à une partie plus technique. Pour trouver toutes les éventuelles collisions nous pourrions naïvement parcourir l'ensemble des paires de corps solides. Cette manière de procéder est coûteuse et de l'ordre de  $\mathcal{O}(N(N-1)/2)$  pour un polymère de  $N$  segments. ODE propose

astucieusement d'utiliser un espace sur lequel on plaque un quadrillage ; un espace multi-résolution avec une maille de longueur  $2^n$  où  $n$  est un entier. Avec cette procédure, la détection de collisions est de l'ordre de  $\mathcal{O}(N)$ . Nous parcourons alors les cases en listant les objets appartenant à chacune d'entre elles. Dans le cas d'un très grand nombre de corps rigides, il est possible d'utiliser un espace dit "Quadtree" en lieu et place d'un espace "Hash". Cet espace est plus technique à détailler et il est plus judicieux de se référer au code source de ODE. Nous supprimons pour des questions de stabilité les collisions entre corps liés par un ou plusieurs joints. Dans toutes les simulations j'utilise des collisions inélastiques et sans friction de Coulomb.

Je vais maintenant anticiper sur la suite du manuscrit. Pour tester les collisions je procède à une simulation où je place des sphères dans une boîte aux conditions aux limites périodiques. Les vitesses linéaires et angulaires des corps sont initialisées avec une distribution de Maxwell. Je mesure alors à chaque instant l'énergie cinétique du système  $E(t)$  et je ne constate que de légères déviations par rapport à l'énergie cinétique initiale  $|E(t) - E(0)|/E(0) \sim 10^{-3}$ . Il est même possible de vérifier la validité de l'équipartition de l'énergie !

## 2.5 Présentation du modèle mécanique d'ADN

Jusqu'à maintenant, j'ai expliqué comment simuler un système de corps solides dans un bain thermique. Nous allons dès à présent nous intéresser à l'ADN et comment simuler ses propriétés mécaniques : longueur de persistance, rayon effectif électrostatique. Le mot persistance regroupe les détails microscopiques de l'ADN ou de tout autre polymère au sein d'une longueur de corrélation sur laquelle l'ADN perd son orientation. La persistance en courbure sera notée  $p$  et la persistance en twist,  $t$ . Je vais préalablement introduire un modèle d'ADN vu comme un système articulé et la dynamique globale de Langevin-Euler adaptée à cette modélisation.

### 2.5.1 Joint Ball-in-Socket

Je vais décrire un type de joint mécanique : un joint Ball-in-Socket (voir la figure (2.3)). Je propose de nous en servir pour construire l'ADN (connecter des cylindres de masse  $m$ , de longueur  $l$ , de rayon  $r$  et de moments d'inertie selon les axes principaux  $(I_{\perp}^*, I_{\perp}^*, I_{\parallel}^*)$  et de travailler dessus pour introduire la persistance en courbure et en twist. Nous avons  $I_{\perp}^* = \frac{1}{12}ml^2 + \frac{1}{4}mr^2$  et  $I_{\parallel}^* = \frac{1}{2}mr^2$ . Dans la suite, je nommerai ce modèle BLC. Le second type de joint que je présente, le joint de Cardan, se trouve dans la section (2.9). Un joint "Ball-in-Socket" permet de connecter deux corps solides sans restreindre l'orientation relative.

En utilisant  $\mathbf{r}_i \pm \frac{l}{2}\mathbf{t}_i$  comme les positions des points d'ancrages (deux extrémités) sur le corps  $i$ , nous pouvons écrire les trois contraintes holonomiques  $\delta$  pour un joint ball-in-socket sous la forme vectorielle (2.8). En dérivant par rapport au temps, nous obtenons l'équation (2.8) sur le torseur cinématique  $\dot{\delta} = \mathcal{J}\mathcal{V}$ . La condition sur les vitesses peut alors se simplifier puisque la partie de la vitesse angulaire selon  $\mathbf{t}$ , notée  $\omega_{\parallel} = (\boldsymbol{\omega}^T \mathbf{t}) \mathbf{t}$ , disparaît par le produit vectoriel avec  $\mathbf{t}$ . Il ne reste plus que la partie normale à  $\mathbf{t}$ , notée  $\omega_{\perp} = (\boldsymbol{\omega}^T \mathbf{m}) \mathbf{m} + (\boldsymbol{\omega}^T \tilde{\mathbf{m}}) \tilde{\mathbf{m}}$ .

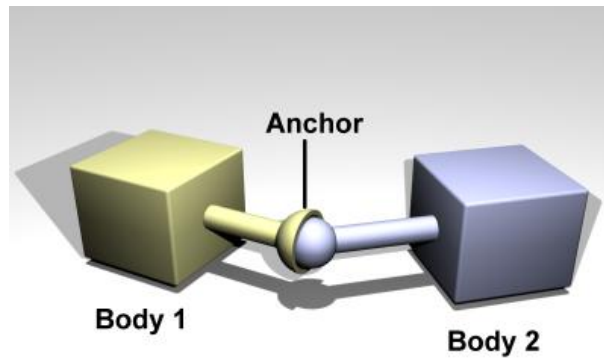


FIGURE 2.3: Vue schématique d'un joint "Ball-in-Socket". ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.

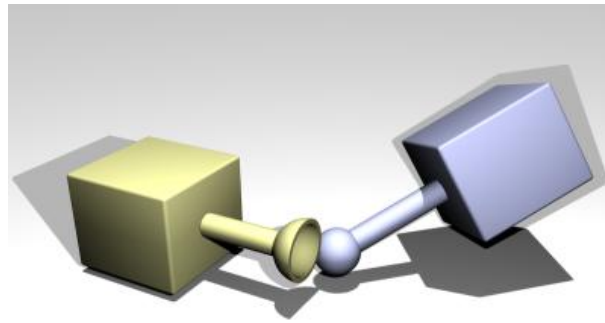


FIGURE 2.4: Vue schématique d'un joint "Ball-in-Socket" ne satisfaisant pas la condition d'ancrage (voir l'équation (2.8)). ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.

$$\begin{aligned}
 \delta &= (\mathbf{r}_1 - \mathbf{r}_2) + \frac{l}{2} (\mathbf{t}_1 + \mathbf{t}_2) = \mathbf{0} \\
 \mathcal{J}\mathcal{V} &= \mathbf{v}_1 - \frac{l}{2} \mathbf{t}_1 \otimes \boldsymbol{\omega}_{1\perp} - \mathbf{v}_2 - \frac{l}{2} \mathbf{t}_2 \otimes \boldsymbol{\omega}_{2\perp} = \mathbf{0} \\
 \mathcal{J} &= \begin{pmatrix} \mathbf{1} & -\frac{l}{2} \mathbf{t}_1^\otimes & -\mathbf{1} & -\frac{l}{2} \mathbf{t}_2^\otimes \end{pmatrix} \quad (2.8)
 \end{aligned}$$

Les erreurs numériques obligent à introduire un terme de correction, dans le membre de droite des contraintes cinématiques, fait d'une vitesse de correction qui s'exprime tout simplement comme  $-k_{erp}\delta/\Delta t$  et un terme proportionnel au multiplicateur de Lagrange  $-k_{cfm}\lambda$  dans le membre de droite de l'équation (2.8). En effet, ces erreurs peuvent faire diverger les corps de leurs positions satisfaisant la contrainte (voir la figure (2.3)). Une étude plus détaillée des paramètres "Error Reduction Parameter" et "Constraint Force Mixing" est donnée dans la sous-section (2.8.2).

Pour finir, en anticipant sur le paragraphe suivant, je rappelle que deux corps solides connectés par un joint "Ball-in-Socket" sont reliés par la transformation d'Euler standard ZXZ. En effet, à chaque corps solide, nous pouvons associer un référentiel matériel défini par la tangente au corps  $\mathbf{t}$ , un vecteur orthogonal  $\mathbf{m}$  et un troisième vecteur satisfaisant la relation  $\tilde{\mathbf{m}} = \mathbf{t} \otimes \mathbf{m}$ . En introduisant

les angles d'Euler  $\alpha$ ,  $\theta$  et  $\psi$ , nous mettons en équation (2.9) ce que nous venons de dire.

$$\begin{aligned}
 \mathbf{m}_2 &= (\cos \psi \cos \alpha - \cos \theta \sin \alpha \sin \psi) \mathbf{m}_1 \\
 &\quad + (\cos \psi \sin \alpha + \cos \theta \cos \alpha \sin \psi) \widetilde{\mathbf{m}}_1 \\
 &\quad + \sin \psi \sin \theta \mathbf{t}_1 \\
 \widetilde{\mathbf{m}}_2 &= (-\sin \psi \cos \alpha - \cos \theta \sin \alpha \cos \psi) \mathbf{m}_1 \\
 &\quad + (-\sin \psi \sin \alpha + \cos \theta \cos \alpha \cos \psi) \widetilde{\mathbf{m}}_1 \\
 &\quad + \cos \psi \sin \theta \mathbf{t}_1 \\
 \mathbf{t}_2 &= \sin \theta \sin \alpha \mathbf{m}_1 - \sin \theta \cos \alpha \widetilde{\mathbf{m}}_1 + \cos \theta \mathbf{t}_1
 \end{aligned} \tag{2.9}$$

La relation inverse est :

$$\begin{aligned}
 \mathbf{m}_1 &= (\cos \psi \cos \alpha - \cos \theta \sin \alpha \sin \psi) \mathbf{m}_2 \\
 &\quad - (\sin \psi \cos \alpha + \cos \theta \sin \alpha \cos \psi) \widetilde{\mathbf{m}}_2 \\
 &\quad + \sin \alpha \sin \theta \mathbf{t}_2 \\
 \widetilde{\mathbf{m}}_1 &= (\cos \alpha \sin \psi \cos \theta + \sin \alpha \cos \psi) \mathbf{m}_2 \\
 &\quad + (\cos \alpha \cos \psi \cos \theta - \sin \alpha \sin \psi) \widetilde{\mathbf{m}}_2 \\
 &\quad - \cos \alpha \sin \theta \mathbf{t}_2 \\
 \mathbf{t}_1 &= \sin \psi \sin \theta \mathbf{m}_2 + \cos \psi \sin \theta \widetilde{\mathbf{m}}_2 + \cos \theta \mathbf{t}_2
 \end{aligned} \tag{2.10}$$

L'angle  $\theta$  représente la courbure entre les deux cylindres d'ADN et la somme  $\alpha + \psi$  est le twist, angle de rotation autour de la tangente au second cylindre. En utilisant le transport parallèle [34], par rotation d'un angle  $\theta$  autour de la binormale  $\mathbf{b}_{12} = \frac{\mathbf{t}_1 \otimes \mathbf{t}_2}{|\sin \theta|}$  on peut amener le vecteur  $\mathbf{t}_1$  sur la tangente  $\mathbf{t}_2$ . Dans l'article de R. E. Goldstein [35], il est dit que l'idée du transport parallèle remonte à G. Darboux. Ensuite, pour faire correspondre  $\mathbf{m}_1$  avec  $\mathbf{m}_2$  et  $\widetilde{\mathbf{m}}_1$  avec  $\widetilde{\mathbf{m}}_2$ , il faut faire une rotation d'un angle  $\alpha + \psi$  autour de la tangente  $\mathbf{t}_2$ . A partir des angles d'Euler ZXZ, je peux construire les produits vectoriels (2.11) suivants qui me permettront de simplifier l'expression du couple de rappel de twist.

$$\begin{aligned}
 \mathbf{m}_1 \otimes \mathbf{m}_2 &= (\cos \psi \sin \alpha + \cos \theta \cos \alpha \sin \psi) \mathbf{t}_1 - \sin \psi \sin \theta \widetilde{\mathbf{m}}_1 \\
 \widetilde{\mathbf{m}}_1 \otimes \widetilde{\mathbf{m}}_2 &= (\sin \psi \cos \alpha + \cos \theta \sin \alpha \cos \psi) \mathbf{t}_1 + \cos \psi \sin \theta \mathbf{m}_1 \\
 \mathbf{t}_1 \otimes \mathbf{t}_2 &= \sin \theta \cos \alpha \mathbf{m}_1 + \sin \theta \sin \alpha \widetilde{\mathbf{m}}_1 \\
 &= \sin \theta \cos \psi \mathbf{m}_2 - \sin \theta \sin \psi \widetilde{\mathbf{m}}_2
 \end{aligned} \tag{2.11}$$

## 2.5.2 Couples de courbure et de twist pour un BLC

A partir des angles d'Euler, nous obtenons numériquement la courbure  $\theta$  et le twist  $\phi = \alpha + \psi$  grâce aux équations (2.12). L'énergie de courbure est définie comme  $\beta E_b = g_b (1 - \mathbf{t}_1^T \mathbf{t}_2)$  et non comme dans le cas du WLC [9] par l'approximation des petits angles  $g_b \theta^2 / 2$  où  $g_b = p/l$ . L'énergie de twist est définie par  $\beta E_t = g_t \phi^2 / 2$  où  $g_t = t/l$ . Dans toutes mes simulations,  $\theta$  est restreint à l'intervalle  $[0; \pi[$  et  $\phi$  à l'intervalle  $]-\pi; \pi[$ .



$$\begin{aligned}\cos \phi &= \frac{\mathbf{m}_2^T \mathbf{m}_1 + \widetilde{\mathbf{m}}_2^T \widetilde{\mathbf{m}}_1}{1 + \mathbf{t}_1^T \mathbf{t}_2} \\ \sin \phi &= \frac{\mathbf{m}_2^T \widetilde{\mathbf{m}}_1 - \widetilde{\mathbf{m}}_2^T \mathbf{m}_1}{1 + \mathbf{t}_1^T \mathbf{t}_2}\end{aligned}\quad (2.12)$$

Les potentiels que je viens de donner ont la bonne idée de ne dépendre que des orientations des corps solides. Nous pouvons alors contrôler la courbure et le twist en ajoutant un couple au niveau du joint “Ball-in-Socket” soit un couple  $\mathbf{\Gamma}_{b+t}^1 = -(\mathbf{t}_1 \otimes \nabla_{\mathbf{t}_1} + \mathbf{m}_1 \otimes \nabla_{\mathbf{m}_1} + \widetilde{\mathbf{m}}_1 \otimes \nabla_{\widetilde{\mathbf{m}}_1})(E_b + E_t)$  au corps 1 et un couple  $\mathbf{\Gamma}_{b+t}^2 = -(\mathbf{t}_2 \otimes \nabla_{\mathbf{t}_2} + \mathbf{m}_2 \otimes \nabla_{\mathbf{m}_2} + \widetilde{\mathbf{m}}_2 \otimes \nabla_{\widetilde{\mathbf{m}}_2})(E_b + E_t)$  sur le corps 2. Nous pouvons remarquer que  $\mathbf{\Gamma}_{b+t}^1 = -\mathbf{\Gamma}_{b+t}^2 = \mathbf{\Gamma}_{b+t}$  avec  $\mathbf{\Gamma}_{b+t}$  qui s’écrit comme l’équation (2.13) et peut être considéré comme un principe d’action-réaction. Pour expliquer cette définition des couples de courbure et de twist il faut raisonner comme suit : soit une modification de l’angle de courbure  $\theta$  telle que  $\theta$  devient  $\theta + \delta\theta$ . Ceci correspond à une rotation d’un angle  $\delta\theta$  autour de la binormale  $\mathbf{b}$  et à une variation d’énergie de courbure  $\beta\delta E_b = g_b \sin \theta \delta\theta$ . De même, soit la modification de  $\mathbf{t}_1$  en  $\mathbf{t}_1 + \delta\mathbf{t}_1$  qui induit une variation d’énergie de courbure  $\beta\delta E_b = -g_b \delta\mathbf{t}_1^T \mathbf{t}_2 = \delta\mathbf{t}_1^T \nabla_{\mathbf{t}_1} E_b$  et un couple de rappel  $\mathbf{\Gamma}_b^1 = -\mathbf{t}_1 \otimes \nabla_{\mathbf{t}_1} E_b$ . Il est d’ailleurs intéressant de donner les temps caractéristiques de fluctuations associés à ces couples :  $\Sigma_R^* I^* / g_b$  pour la courbure et  $\Sigma_R^* I^* / g_t$  pour le twist.

$$\beta\mathbf{\Gamma}_{b+t} = g_b \mathbf{t}_1 \otimes \mathbf{t}_2 + \frac{g_t \phi}{1 + \cos \theta} (\mathbf{t}_1 + \mathbf{t}_2)\quad (2.13)$$

Pour trouver  $g_b$  nous suivons la procédure proposée dans l’article [36]. La moyenne thermique du cosinus de la courbure peut s’exprimer de deux façons différentes. La première revient à calculer  $\langle \cos \theta \rangle$  à l’aide de la fonction de partition (membre de gauche de l’équation (2.14)). Cette intégrale est très simple à calculer puisqu’elle est égale à la fonction de Langevin  $\mathcal{L}(g_b)$ . La seconde revient à calculer la distance bout-à-bout quadratique moyenne dans la limite où le nombre de segments d’ADN est grand  $\langle \mathbf{R}^2 \rangle = Llq$  (membre de droite de l’équation (2.14)). La quantité  $q = k/l$  représente le nombre de segments par longueur de Kuhn  $k = 2p$ . Une longueur de Kuhn correspond à deux fois la longueur de persistance. On peut aussi écrire  $q$  en fonction de la moyenne du cosinus de l’angle de courbure  $\langle \cos \theta \rangle$  comme  $q = \frac{1 + \langle \cos \theta \rangle}{1 - \langle \cos \theta \rangle}$ . En résumé, la constante  $g_b$  est construite de telle sorte à faire correspondre  $q$  segments de longueur  $l$  à un segment de Kuhn statistique. Dans le cas limite où la longueur des cylindres est égale à un segment de Kuhn, la constante de courbure est nulle de même que les corrélations tangentielles et nous retrouvons le modèle du FJC.

$$\frac{\int_0^\pi \cos \theta \sin \theta e^{g_b \cos \theta} d\theta}{\int_0^\pi \sin \theta e^{g_b \cos \theta} d\theta} = \frac{k - l}{k + l}\quad (2.14)$$

En utilisant les expressions (2.11), je déroule le calcul suivant qui arrive à l’expression simplifiée du couple de twist (2.15). Le coefficient  $1 + \cos \theta$  sert de normalisation. En effet, lorsque les deux tangentes sont alignées  $\mathbf{t}_1 + \mathbf{t}_2 = 2\mathbf{t}$  et  $1 + \cos \theta = 2$ , on retrouve bien l’expression du couple de twist  $g_t \phi \mathbf{t}$  pour une tige droite, d’axe  $\mathbf{t}$ , soumise à un effort de torsion aux extrémités créant un twist  $\phi$ .

$$\begin{aligned}
 \beta \mathbf{\Gamma}_t &= -(\mathbf{t}_1 \otimes \nabla_{\mathbf{t}_1} + \mathbf{m}_1 \otimes \nabla_{\mathbf{m}_1} + \widetilde{\mathbf{m}}_1 \otimes \nabla_{\widetilde{\mathbf{m}}_1}) E_t \\
 &= \frac{g_t |\phi|}{(1 + \cos \theta) |\sin \phi|} (\mathbf{m}_1 \otimes \mathbf{m}_2 + \widetilde{\mathbf{m}}_1 \otimes \widetilde{\mathbf{m}}_2 - \cos \phi \mathbf{t}_1 \otimes \mathbf{t}_2) \\
 &\propto (\cos \psi \sin \alpha + \cos \theta \cos \alpha \sin \psi + \sin \psi \cos \alpha + \cos \theta \sin \alpha \cos \psi) \mathbf{t}_1 \\
 &\quad - (\sin \psi \sin \theta + \cos(\psi + \alpha) \sin \theta \sin \alpha) \widetilde{\mathbf{m}}_1 + (\cos \psi \sin \theta - \cos(\psi + \alpha) \sin \theta \cos \alpha) \mathbf{m}_1 \\
 &\propto \sin(\alpha + \psi) (1 + \cos \theta) \mathbf{t}_1 - \sin \theta (\sin \psi + \cos \psi \cos \alpha \sin \alpha - \sin \psi \sin^2 \alpha) \widetilde{\mathbf{m}}_1 \\
 &\quad + \sin \theta (\cos \psi - \cos \psi \cos^2 \alpha + \sin \psi \sin \alpha \cos \alpha) \mathbf{m}_1 \\
 &\propto \sin(\alpha + \psi) (1 + \cos \theta) \mathbf{t}_1 - \sin \theta (\sin \psi \cos^2 \alpha + \cos \psi \cos \alpha \sin \alpha) \widetilde{\mathbf{m}}_1 \\
 &\quad + \sin \theta (\cos \psi \sin^2 \alpha + \sin \psi \sin \alpha \cos \alpha) \mathbf{m}_1 \\
 &\propto \sin(\alpha + \psi) (1 + \cos \theta) \mathbf{t}_1 - \sin(\alpha + \psi) \sin \theta \cos \alpha \widetilde{\mathbf{m}}_1 + \sin(\alpha + \psi) \sin \theta \sin \alpha \mathbf{m}_1 \\
 &= \frac{g_t |\phi| \sin \phi}{(1 + \cos \theta) |\sin \phi|} (\mathbf{t}_1 + \mathbf{t}_2) \\
 \beta \mathbf{\Gamma}_t &= \frac{g_t \phi}{1 + \cos \theta} (\mathbf{t}_1 + \mathbf{t}_2) \quad \text{avec } \phi \in ]-\pi; \pi[ \quad \text{et } \theta \in [0; \pi[
 \end{aligned} \tag{2.15}$$

### 2.5.3 Exemple : chaîne de quatre cylindres connectés

Je considère le cas de quatre cylindres connectés aux extrémités  $\pm l/2$  par un joint "Ball-in-Socket". J'ajoute un joint "Ball-in-Socket" pour connecter le premier cylindre au sol en  $\mathbf{r}_1 - \frac{l}{2} \mathbf{t}_1$ . Cette chaîne contient quatre joints dont deux qui vont sûrement présenter des effets de bords ; le joint au centre de la chaîne est influencé par le joint précédent et par le suivant. La matrice antisymétrique  $\mathbf{t}^{\otimes}$  (voir l'équation (2.16)) est associée au produit vectoriel  $\mathbf{t} \otimes$  et possède la propriété  $\mathbf{t}^{\otimes T} = -\mathbf{t}^{\otimes}$  et peut s'exprimer dans le référentiel matériel comme la matrice  $\mathbf{t}^{\otimes*}$ .

$$\mathbf{t}^{\otimes} = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \quad \mathbf{t}^{\otimes*} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.16}$$

Pour le calcul du terme  $\mathcal{J} \mathcal{M}^{-1} \mathcal{J}^T$  apparaissant dans l'équation (2.2) sur les contraintes  $\boldsymbol{\lambda}$ , j'introduis la matrice  $\mathcal{J}$  :

$$\mathcal{J} = \begin{pmatrix} -1 & -\frac{l}{2} \mathbf{t}_1^{\otimes} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -\frac{l}{2} \mathbf{t}_1^{\otimes} & -1 & -\frac{l}{2} \mathbf{t}_2^{\otimes} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\frac{l}{2} \mathbf{t}_2^{\otimes} & -1 & -\frac{l}{2} \mathbf{t}_3^{\otimes} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\frac{l}{2} \mathbf{t}_3^{\otimes} & -1 & -\frac{l}{2} \mathbf{t}_4^{\otimes} \end{pmatrix} \tag{2.17}$$

sa transposée  $\mathcal{J}^T$  et la matrice  $\mathcal{M}$  :

$$\mathcal{J}^T = \begin{pmatrix} -1 & 1 & 0 & 0 \\ \frac{l}{2}\mathbf{t}_1^\otimes & \frac{l}{2}\mathbf{t}_1^\otimes & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{l}{2}\mathbf{t}_2^\otimes & \frac{l}{2}\mathbf{t}_2^\otimes & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & \frac{l}{2}\mathbf{t}_3^\otimes & \frac{l}{2}\mathbf{t}_3^\otimes \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & \frac{l}{2}\mathbf{t}_4^\otimes \end{pmatrix} \quad \mathcal{M} = \begin{pmatrix} m & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & m & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & m & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & m & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_4 \end{pmatrix} \quad (2.18)$$

et je calcule l'intermédiaire de calcul  $\mathcal{M}^{-1}\mathcal{J}^T$  :

$$\mathcal{M}^{-1}\mathcal{J}^T = \begin{pmatrix} -m^{-1} & m^{-1} & 0 & 0 \\ \frac{l}{2}I_1^{-1}\mathbf{t}_1^\otimes & \frac{l}{2}I_1^{-1}\mathbf{t}_1^\otimes & 0 & 0 \\ 0 & -m^{-1} & m^{-1} & 0 \\ 0 & \frac{l}{2}I_2^{-1}\mathbf{t}_2^\otimes & \frac{l}{2}I_2^{-1}\mathbf{t}_2^\otimes & 0 \\ 0 & 0 & -m^{-1} & m^{-1} \\ 0 & 0 & \frac{l}{2}I_3^{-1}\mathbf{t}_3^\otimes & \frac{l}{2}I_3^{-1}\mathbf{t}_3^\otimes \\ 0 & 0 & 0 & -m^{-1} \\ 0 & 0 & 0 & \frac{l}{2}I_4^{-1}\mathbf{t}_4^\otimes \end{pmatrix} \quad (2.19)$$

Ce qui nous amène au résultat suivant (2.20) pour la matrice inverse des masses effectives aux joints  $\mathcal{J}\mathcal{M}^{-1}\mathcal{J}^T$  où je pose  $\mathcal{T}^\otimes = \frac{l^2}{4}\mathbf{t}^\otimes I^{-1}\mathbf{t}^\otimes$ .

$$\begin{pmatrix} \frac{1}{m} - \mathcal{T}_1^\otimes & -\frac{1}{m} - \mathcal{T}_1^\otimes & 0 & 0 \\ -\frac{1}{m} - \mathcal{T}_1^\otimes & \frac{2}{m} - \mathcal{T}_1^\otimes - \mathcal{T}_2^\otimes & -\frac{1}{m} - \mathcal{T}_2^\otimes & 0 \\ 0 & -\frac{1}{m} - \mathcal{T}_2^\otimes & \frac{2}{m} - \mathcal{T}_2^\otimes - \mathcal{T}_3^\otimes & -\frac{1}{m} - \mathcal{T}_3^\otimes \\ 0 & 0 & -\frac{1}{m} - \mathcal{T}_3^\otimes & \frac{2}{m} - \mathcal{T}_3^\otimes - \mathcal{T}_4^\otimes \end{pmatrix} \quad (2.20)$$

La matrice  $\mathcal{T}^\otimes = \mathcal{R}\mathcal{T}^{\otimes*}\mathcal{R}^T$  peut s'écrire dans le référentiel matériel associé au corps rigide comme l'équation (2.21).

$$\mathcal{T}^{\otimes*} = -\frac{l^2}{4I_\perp^*} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.21)$$

Je déduis de l'expression de la matrice  $\mathcal{T}_i^{\otimes*}$  que  $\mathcal{T}_i^{\otimes}\mathbf{t}_i = \mathbf{0}$ ,  $\mathcal{T}_i^{\otimes}\mathbf{m}_i = -\frac{l^2}{4I_\perp^*}\mathbf{m}_i$  et  $\mathcal{T}_i^{\otimes}\widetilde{\mathbf{m}}_i = -\frac{l^2}{4I_\perp^*}\widetilde{\mathbf{m}}_i$ . Le vecteur  $\boldsymbol{\lambda}$  est la collection des multiplicateurs de Lagrange aux joints, soit  $\boldsymbol{\lambda} = \{\boldsymbol{\lambda}^0, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\lambda}^3\}^T$ . Des expressions précédentes, je tire quatre équations vectorielles couplées sur  $\boldsymbol{\lambda}^0, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2$  et  $\boldsymbol{\lambda}^3$ . Sur cet exemple, il est clair qu'un joint s'accorde avec ses joints voisins pour se mouvoir de concert ; il est évident que les moteurs physiques résolvent toutes les contraintes en une seule fois. Ils possèdent donc un net avantage sur "QSHAKE" où il est nécessaire de corriger itérativement les positions et orientations des corps rigides jusqu'à satisfaire la totalité des contraintes. Le terme  $\mathcal{J}\mathcal{M}^{-1}\mathcal{F}_e$  de l'équation (2.2) est quant à lui égal à :

$$\begin{pmatrix} -1 & -\frac{l}{2}\mathbf{t}_1^\otimes & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -\frac{l}{2}\mathbf{t}_1^\otimes & -1 & -\frac{l}{2}\mathbf{t}_2^\otimes & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\frac{l}{2}\mathbf{t}_2^\otimes & -1 & -\frac{l}{2}\mathbf{t}_3^\otimes & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\frac{l}{2}\mathbf{t}_3^\otimes & -1 & -\frac{l}{2}\mathbf{t}_4^\otimes \end{pmatrix} \begin{pmatrix} \frac{f_1}{m} \\ \frac{\Gamma_{1\perp}}{I_\perp^*} + \frac{\Gamma_{1\parallel}}{I_\parallel^*} \\ \frac{f_2}{m} \\ \frac{\Gamma_{2\perp}}{I_\perp^*} + \frac{\Gamma_{2\parallel}}{I_\parallel^*} \\ \frac{f_3}{m} \\ \frac{\Gamma_{3\perp}}{I_\perp^*} + \frac{\Gamma_{3\parallel}}{I_\parallel^*} \\ \frac{f_4}{m} \\ \frac{\Gamma_{4\perp}}{I_\perp^*} + \frac{\Gamma_{4\parallel}}{I_\parallel^*} \end{pmatrix} \quad (2.22)$$

où je note  $\Gamma_\perp = (\Gamma^T \mathbf{m}) \mathbf{m} + (\Gamma^T \tilde{\mathbf{m}}) \tilde{\mathbf{m}}$  et  $\Gamma_\parallel = (\Gamma^T \mathbf{t}) \mathbf{t}$ . J'en déduis le résultat final :

$$\begin{pmatrix} -\frac{f_1}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_1 \otimes \Gamma_{1\perp} \\ \frac{f_1}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_1 \otimes \Gamma_{1\perp} - \frac{f_2}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_2 \otimes \Gamma_{2\perp} \\ \frac{f_2}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_2 \otimes \Gamma_{2\perp} - \frac{f_3}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_3 \otimes \Gamma_{3\perp} \\ \frac{f_3}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_3 \otimes \Gamma_{3\perp} - \frac{f_4}{m} - \frac{l}{2I_\perp^*} \mathbf{t}_4 \otimes \Gamma_{4\perp} \end{pmatrix} \quad (2.23)$$

où nous nous apercevons bien que chaque joint se renseigne sur les forces et couples appliqués aux corps connectés.

## 2.5.4 Électrostatique

La plupart des auteurs utilisent, pour modéliser l'électrostatique, de l'ADN dans une solution de sel, l'approximation de Debye-Hückel :

$$E_{el} = \sum_i \sum_{j>i} \int \int \frac{\nu^2 e^{-\kappa r} dr_i dr_j}{D r_{ij}} \quad (2.24)$$

qui met en jeu des paramètres comme  $\nu$  la charge linéaire effective de l'ADN,  $\kappa$  la longueur d'écrantage de Debye-Hückel,  $D$  la constante diélectrique de l'eau. Les  $r_i$  correspondent aux lieux où se trouvent les charges sur l'ADN. Cette expression compliquée, mêlant intégrale et double somme, peut être simplifiée par un rayon effectif électrostatique [37, 38, 39]. En anticipant sur la partie résultat, je trouve que ce rayon effectif  $\tilde{r}$  est la somme du rayon cristallographique de l'ADN avec la longueur de Debye  $\lambda_D$ , longueur à partir de laquelle deux charges ne se voient plus. Pour 100 mmol de sel, la longueur de Debye  $\lambda_D$  est égale à 1 nm. Ce résultat est en désaccord avec [37, 38]; il est sous-estimé par rapport à ces deux articles mais fonctionne très bien dans mes simulations. Je tiens quand même à souligner que cette approximation n'est plus valable lorsque deux ADN s'apparient. Pour des plectonèmes obtenus avec une très grande force de traction, N. Clauvelin [40] montre que l'approximation d'un rayon effectif ne fonctionne plus. Il est montré que la persistance en courbure de l'ADN est affectée par l'électrostatique [39] et varie de 50 à 60 nm pour une variation de NaCl de 1 à 0.01 M. À partir de différents résultats expérimentaux, P. J. Hagerman et C. G. Baumann [41, 42] concluent que la persistance en courbure est proche de 50 nm pour des concentrations en NaCl supérieure à 100 mmol. S. Neukirch [43] montre en ajustant un modèle théorique sur des expériences de pinces magnétiques de G. Charvin que la persistance en courbure de l'ADN est 50 nm pour 10 et 150 mmol de NaCl. L'adjonction de 5 mmol d'ions divalents  $\text{Mg}^{2+}$  ne modifie pas la persistance en courbure au contraire de la persistance en twist qui devient  $t = 60$  nm contre  $t = 100$  nm

sans divalent. Nous avons résumé tout cela est résumé dans une revue sur l'électrostatique et la compaction de l'ADN [44].

## 2.6 Dynamique locale de Langevin-Euler

Un moyen de simuler le bain thermique, qui agite l'ADN au gré des sursauts thermiques, est le thermostat de Langevin ; il permet d'échantillonner l'ensemble canonique, en particulier, thermaliser le système à une température  $T$ . Dans l'ensemble canonique, à volume constant et nombre de particules constant, la moyenne de la température macroscopique possède une valeur bien définie tandis que l'énergie du système fluctue. Les expériences de pinces magnétiques sont justement réalisées dans cet ensemble. Chaque segment d'ADN est le lieu de nombreuses collisions et cette multitude est difficile à modéliser ; nous la remplaçons par des forces et couples aléatoires effectifs. De la sorte, on injecte de l'énergie dans la molécule d'ADN qu'il faut dissiper pour atteindre un équilibre thermique ; un jeu entre fluctuations et dissipation. La dissipation se fait par l'intermédiaire de la friction du solvant. L'amplitude des forces et couples aléatoires est liée à la friction et à la température du bain thermique. Les équations du thermostat de Langevin peuvent se résumer sous la forme d'une équation matricielle (2.25) [45].  $\Sigma$  est une matrice symétrique représentant les coefficients de couplage aux thermostats translationnel et rotationnel.  $\Xi\dot{\mathcal{W}}$  est un vecteur de bruit blanc Gaussien.  $\dot{\mathcal{W}}$  est un vecteur de processus de Wiener normalisé et indépendant  $\langle \dot{\mathcal{W}}_i(t)\dot{\mathcal{W}}_j(t') \rangle = \delta_{ij}\delta(t-t')$  et de moyennes nulles. Numériquement, on remplace la distribution de Dirac par  $1/\Delta t$ . Les coefficients  $\Xi_{ij}$  sont déterminés par la solution de l'équation de Langevin (2.25) ou par le théorème fluctuations-dissipations qui est satisfait à l'équilibre thermique. Pour un système de  $N$  corps solides à l'équilibre statistique, j'écris la température comme  $\langle \mathcal{L}^T \mathcal{M}^{-1} \mathcal{L} \rangle = 6Nk_B T$  où plus simplement comme  $\langle \mathcal{V}^{*T} \mathcal{M}^* \mathcal{V}^* \rangle = 6Nk_B T$

$$\dot{\mathcal{L}} = \mathcal{F}_e - \Sigma \mathcal{L} + \Xi \dot{\mathcal{W}} \quad (2.25)$$

Cette équation n'est rien d'autre que l'équation de la dynamique avec deux termes supplémentaires ; l'un correspondant à une friction et l'autre étant stochastique  $\mathcal{G} = -\Sigma \mathcal{L} + \Xi \dot{\mathcal{W}}$ . Dans la suite, je nommerai  $\mathcal{G}$  torseur dynamique de Langevin local. Il est intéressant, avant de se lancer dans la section suivante, de noter que ce thermostat est local ; nous appliquons un torseur aléatoire et un torseur de friction à chaque degré de liberté. Je vais essayer de vous montrer que nous pouvons mieux faire !

## 2.7 Dynamique globale de Langevin-Euler

Entre 2007 et 2008, Bussi et Parinello ont développé de nouveaux et géniaux algorithmes pour remplacer le traditionnel "velocity rescaling" [46] et la dynamique de Langevin [47]. En particulier, ils proposent de remplacer le thermostat local de Langevin par une version globale. Je propose ici d'étendre cet algorithme aux degrés de liberté rotationnels. Ma démarche va suivre celle de Bussi [47] et je montrerai, en plus, pourquoi il est intéressant d'utiliser ce thermostat global pour les moteurs physiques. En tout premier lieu, nous allons définir l'Hamiltonien du système comme  $H(\mathcal{S}, \mathcal{V}) = E(\mathcal{V}) + U(\mathcal{S})$  où la quantité  $E(\mathcal{V}) = T(\mathbf{v}) + R(\boldsymbol{\omega})$  représente la somme de l'énergie translationnelle et rotationnelle soit l'énergie cinétique du système. La fonction  $U(\mathcal{S})$ , représente, les potentiels externes. La distribution dans l'ensemble canonique est alors  $P(\mathcal{S}, \mathcal{L}) d\mathcal{S}d\mathcal{L} \propto e^{-\beta H(\mathcal{S}, \mathcal{V})}$ . L'énergie cinétique s'écrit sous la forme  $\frac{1}{2} \mathcal{L}^T \mathcal{M}^{-1} \mathcal{L}$  ou comme  $\frac{1}{2} \mathcal{V}^T \mathcal{M} \mathcal{V}$ . Nous pouvons alors

calculer la vitesse de thermalisation comme une différentielle stochastique  $dH(t)$  de l'énergie totale. Pour cela on utilise le calcul d'Ito (2.26).

$$dH = H_t + (\nabla_{\mathcal{L}} H)^T d\mathcal{L} + \frac{1}{2} d\mathcal{L}^T (\nabla_{\mathcal{L}}^2 H) d\mathcal{L} \quad (2.26)$$

$\nabla_{\mathcal{L}}^2 H$  est la matrice Hessienne du Hamiltonien  $H$  avec les termes diagonaux qui sont nuls. La dérivée partielle par rapport au temps est elle aussi nulle  $H_t = 0$ . Dans la suite, on utilise les règles de calculs d'Ito (2.27), cette manière d'écrire ne possède aucun sens mathématique rigoureux mais est très utile dans le calcul.  $d\mathcal{W}$  n'est pas un processus stochastique et ces équations constituent un raccourci d'écriture dont seule la forme intégrée possède un sens

$$\begin{aligned} d\mathcal{W}_i d\mathcal{W}_j &= \delta_{ij} dt \\ d\mathcal{W}^N &= 0 \quad N > 2 \\ d\mathcal{W}^N dt &= 0 \quad N \geq 1 \\ dt^N &= 0 \quad N > 1 \end{aligned} \quad (2.27)$$

Un simple développement nous donne  $dH(t)$ .

$$dH = -(\mathcal{M}^{-1} \mathcal{L})^T (\Sigma \mathcal{L}) dt + (\mathcal{M}^{-1} \mathcal{L})^T (\Xi d\mathcal{W}) + \frac{1}{2} (\Xi d\mathcal{W})^T \mathcal{M}^{-1} (d\mathcal{W} \Xi) \quad (2.28)$$

Ensuite nous utilisons les règles d'Ito avec les trois expressions  $\mathcal{M} = \mathcal{P} \mathcal{M}^* \mathcal{P}^T$ ,  $\Sigma = \mathcal{P} \Sigma^* \mathcal{P}^T$  et  $\mathcal{L} = \mathcal{P} \mathcal{L}^*$  pour simplifier  $dH$ . En utilisant le théorème fluctuation-dissipation  $\beta \Xi_{ii}^{*2} = 2\mathcal{M}_{ii}^* \Sigma_{ii}^*$ , je simplifie encore l'expression de  $dH$ .

$$\begin{aligned} dH(t) &= \sum_{i=1}^{6N} \left[ \left( \frac{\Xi_{ii}^{*2}}{2\mathcal{M}_{ii}^*} - \frac{\Sigma_{ii}^*}{\mathcal{M}_{ii}^*} \mathcal{L}_i^{*2} \right) dt + \frac{\Xi_{ii}^*}{\mathcal{M}_{ii}^*} \mathcal{L}_i^* d\mathcal{W}_i \right] \\ &= \sum_{i=1}^{6N} \left[ \left( \frac{\Sigma_{ii}^*}{\beta} - \frac{\Sigma_{ii}^*}{\mathcal{M}_{ii}^*} \mathcal{L}_i^{*2} \right) dt + \sqrt{\frac{2\Sigma_{ii}^* \mathcal{L}_i^{*2}}{\beta \mathcal{M}_{ii}^*}} d\mathcal{W}_i \right] \end{aligned} \quad (2.29)$$

Dans le cas général la fréquence de couplage translationnelle  $\Sigma_T^*$  n'est pas égale à la fréquence de couplage rotationnelle  $\Sigma_R^*$ . En définissant la moyenne thermique  $\langle E \rangle = 6N/2\beta$  et en utilisant le fait que les processus de Wiener sont indépendants et normalisés nous obtenons l'expression finale (2.30) pour  $dH$ .

$$dH = (\langle E \rangle - 2T) \Sigma_T^* dt + (\langle E \rangle - 2R) \Sigma_R^* dt + 2\sqrt{\frac{\Sigma_T^* T + \Sigma_R^* R}{\beta}} d\mathcal{W} \quad (2.30)$$

Cette équation (2.30) est identique à celle obtenue par Bussi et Parinello augmentée, bien entendu, des degrés de liberté rotationnels. Nous pouvons déjà remarquer que les  $6N$  termes de bruit (processus de Wiener  $d\mathcal{W}_i$ ) ont disparu au profit d'un seul terme de bruit  $\mathcal{W}$  : un pas vers un thermostat global. Il est important de remarquer que l'équation s'écrit indépendamment du nombre de corps rigides constituant du système. Ce "coarse-graining" nous montre que l'on peut écrire une équation de Langevin globale ; le couplage au thermostat ne se réalise plus au niveau de chacun des corps rigides mais au niveau du système. Il s'avère aussi, que la vitesse de thermalisation de l'énergie

cinétique est la même dans la version locale que dans la version globale [47]. Je montrerai une courbe illustrative un peu plus loin.

Pour aller plus loin, je propose de calculer le travail du torseur dynamique du thermostat local  $\mathcal{G} = -\Sigma\mathcal{L} + \Xi\dot{\mathcal{W}}$  soit les travaux des forces et couples de Langevin dans le cas du thermostat local (2.31) où je suppose, par commodité, que les coefficients de couplage au thermostat sont tous égaux à  $\Sigma^*$ .

$$\mathcal{G}^T \mathcal{V} = -2\Sigma^*E + (\Xi\dot{\mathcal{W}})^T \mathcal{V} \quad (2.31)$$

La variation d'énergie cinétique du système n'est pas contrôlée et elle est perturbée par le terme aléatoire  $(\Xi\dot{\mathcal{W}})^T \mathcal{V}$ . Le torseur local perturbe la dynamique du système; cette perturbation peut être quantifiée [48] et adaptée aux dynamiques de corps rigides comme  $\tilde{\mathcal{G}}^T \mathcal{M}^{-1} \tilde{\mathcal{G}} dt^2$  et possède les unités d'une énergie. Pour construire cette quantité, il faut partir du principe de moindres contraintes de C. F. Gauss. Ce principe nous dit que la trajectoire d'un système de corps rigides doit minimiser l'équation suivante :

$$(\dot{\mathcal{L}} - \mathcal{F}_e)^T \mathcal{M}^{-1} (\dot{\mathcal{L}} - \mathcal{F}_e) \quad (2.32)$$

qui devient en présence du torseur dynamique du thermostat local de Langevin :

$$\mathcal{G}^T \mathcal{M}^{-1} \mathcal{G} \quad (2.33)$$

Minimiser cette quantité pour un incrément d'énergie  $\delta E$  revient à proposer une force généralisée  $\tilde{\mathcal{G}}$  proportionnelle au moment cinétique  $\mathcal{L}$ . Ce  $\tilde{\mathcal{G}}$  doit contenir un terme stochastique et doit dépendre de  $T$  et  $R$  soit une expression sous la forme  $\tilde{\mathcal{G}} = \mathcal{G}\mathcal{L}$ . Dans la suite, je noterai  $\mathcal{G} = A(T, R, t) + B(T, R, t)\dot{\mathcal{W}}$  où  $A$  possède la dimension de l'inverse d'un temps et  $B$  possède la dimension de la racine carrée de l'inverse d'un temps. Avec cette expression la "disturbance" est proportionnelle à l'énergie cinétique du système  $\tilde{\mathcal{G}}^T \mathcal{M}^{-1} \tilde{\mathcal{G}} = 2\mathcal{G}^2 E$ . Qui plus est, le travail des forces est proportionnel à l'énergie cinétique  $E$  du système  $\tilde{\mathcal{G}}^T \mathcal{V} = \mathcal{G}\mathcal{V}^T \mathcal{M}\mathcal{V} = 2\mathcal{G}E$  et la variation d'énergie cinétique en un pas de temps est rendue parfaitement contrôlable. Nous pouvons à présent calculer (2.34) la différentielle stochastique  $dH$  en fonction des coefficients  $A$  et  $B$ .

$$\begin{aligned} dH &= \left( A(\nabla_{\mathcal{L}} H)^T \mathcal{L} + \frac{1}{2} B^2 \mathcal{L}^T (\nabla_{\mathcal{L}}^2 H) \mathcal{L} \right) dt + B d\mathcal{W} (\nabla_{\mathcal{L}} H)^T \mathcal{L} \\ &= (T + R) ((2A + B^2) dt + 2B d\mathcal{W}) \end{aligned} \quad (2.34)$$

Des équations (2.30) et (2.34), nous obtenons la nouvelle force généralisée  $\tilde{\mathcal{G}}$  comme l'expression (2.35). Là encore, nous pouvons faire la comparaison avec l'article de Bussi et Parinello [47].

$$\tilde{\mathcal{G}} = \left[ \frac{\Sigma_T^*}{2E} \left( \langle E \rangle - 2T - \frac{T}{\beta E} \right) + \frac{\Sigma_R^*}{2E} \left( \langle E \rangle - 2R - \frac{R}{\beta E} \right) \right] \mathcal{L} + \sqrt{\frac{\Sigma_T^* T + \Sigma_R^* R}{\beta E^2}} \dot{\mathcal{W}} \mathcal{L} \quad (2.35)$$

Par souci de simplicité et de clarté, je choisis  $\Sigma_T^* = \Sigma_R^* = \Sigma^*$ .

$$\tilde{\mathcal{G}} = \left[ \Sigma^* \left[ \frac{\langle E \rangle}{E} \left( 1 - \frac{1}{2\beta \langle E \rangle} \right) - 1 \right] + \sqrt{\frac{\Sigma^*}{\beta E}} \dot{\mathcal{W}} \right] \mathcal{L} \quad (2.36)$$

Soit un  $\mathcal{G}$  :

$$\mathcal{G} = \Sigma^* \left[ \frac{\langle E \rangle}{E} - \frac{1}{2\beta E} - 1 \right] + \sqrt{\frac{\Sigma^*}{\beta E}} \mathcal{W} \quad (2.37)$$

Nous pourrions penser au premier terme comme à un terme de friction, il n'en est rien. Ce terme dissipe ou injecte de l'énergie au système en fonction de la valeur du rapport  $\frac{\langle E \rangle}{E}$ . Il est à noter que l'expression de  $\tilde{\mathcal{G}}$  est exactement le thermostat de H. J. C. Berendsen [48] auquel on ajoute un terme stochastique. Comme nous pouvons le remarquer, si  $E = 0$ ,  $\tilde{\mathcal{G}}$  devient infini ; il faut donc mettre en orbite, satelliser, les vitesses des corps rigides avec une distribution de Maxwell ou une dynamique locale de Langevin, et ce, au début de chaque simulation. Il est bien entendu possible d'écrire un thermostat global pour des groupes parmi les  $6N$  degrés de liberté, en particulier et en guise d'exemple, lorsque les degrés de libertés translationnels ne sont pas couplés avec les degrés de libertés rotationnels. Ce thermostat permet d'échantillonner exactement l'ensemble canonique et ne pose aucun problème d'ergodicité. Le pas de temps pour l'intégration de l'équation de Langevin globale doit être choisi de telle sorte à remplacer une moyenne d'ensemble par une moyenne sur les itérations de la simulation ("moyenne temporelle"). Contrairement à un algorithme de Metropolis où les configurations sont acceptées ou refusées de manière à échantillonner la distribution souhaitée, les dynamiques de Langevin et moléculaires acceptent toutes les configurations. Cependant, l'intégration non exacte des dynamiques viole le bilan détaillé. Le pas de temps joue donc un rôle important dans cette violation ; il faut le choisir suffisamment petit pour que la violation soit acceptable [49]. En pratique, il faut choisir un  $\Sigma^* \Delta t \sim \frac{1}{100}$  pour obtenir la moyenne désirée sur la température macroscopique  $T$ . Néanmoins, pour une dynamique de corps rigides articulés, un second facteur limitant vient s'ajouter : l'intégration des contraintes. Ce dernier point sera discuter dans peu de lignes.

Jusqu'à présent, je n'ai pas introduit les contraintes mécaniques dans la dynamique globale de Langevin-Euler ; nous en arrivons enfin au plus intéressant de cet algorithme. Que se passe-t-il maintenant si l'on injecte le torseur dynamique  $\tilde{\mathcal{G}}$  dans l'équation (2.2) sur  $\lambda$  ? En particulier, calculons le terme  $\mathcal{J}\mathcal{M}^{-1}\tilde{\mathcal{G}} = \mathcal{G}\mathcal{J}\mathcal{V} = 0$  ; le calcul de  $\lambda$  ne dépend pas de  $\tilde{\mathcal{G}}$  ; il s'ensuit que  $\mathcal{F}_c = \mathcal{J}^T \lambda$  est décorrélé de  $\tilde{\mathcal{G}}$  et que les contraintes mécaniques ne voient pas le bain thermique. Le torseur dynamique global  $\tilde{\mathcal{G}}$  ne perturbe pas la dynamique du système articulé puisqu'il n'intervient pas dans le calcul des contraintes mécaniques et il n'intervient donc pas dans la minimisation du principe de moindres contraintes de C. F. Gauss. Je pense que les moteurs physique sont une réalisation et, si je puis me mettre, un possible achèvement des simulations avec contraintes holonomiques ; ils satisfont le principe de moindres contraintes de C. F. Gauss par implémentation du thermostat global, et, qui plus est, l'échantillonnage du système articulé est accéléré comme nous le verrons par la suite. Pour finir cette partie, je souhaite dire quelques mots sur l'algorithme de "velocity rescaling". Le principe de cette méthode est de multiplier les vitesses des corps tous les  $x$  pas de temps pour atteindre la température souhaitée. En 2007, Bussi et Parinello [46] proposent un rescaling stochastique des vitesses, c'est à dire une version cinématique du thermostat global. Si l'on note  $\alpha$  le coefficient de "rescaling" (voir l'équation (2.38)) le torseur cinématique  $\mathcal{V}$  à l'instant  $t$  devient  $\alpha\mathcal{V}$  à l'instant  $t + dt$ .

$$\alpha^2 = 1 + 2\Sigma^* \left( \frac{\langle E \rangle}{E} - 1 \right) dt + 2\sqrt{\frac{\Sigma^*}{\beta E}} d\mathcal{W} \quad (2.38)$$



Si nous nous intéressons aux contraintes, nous avons  $\mathcal{J}\mathcal{V} = \mathbf{0}$  à l'instant  $t$  qui devient  $\mathcal{J}(\alpha\mathcal{V}) = \alpha\mathcal{J}\mathcal{V} = \mathbf{0}$  à l'instant  $t + dt$ . Les contraintes cinématiques ne sont donc pas perturbées par ce "rescaling" des vitesses. Dans le cas où des degrés de libertés du système sont isolés du reste du système, il faut utiliser plusieurs thermostats globaux. En effet, le thermostat global prend en compte tous les degrés de libertés du systèmes et si un degré de liberté n'interagit pas avec le reste du système nous commettrions un non sens de l'inclure au même titre que les degrés de libertés couplés. En particulier, pour des cylindres connectés par des joints "Ball-in-Socket" et sans couples de courbure et de twist, les vitesses angulaires selon les tangentes aux cylindres sont découplées les unes des autres et découplées du reste du système. Ce cas particulier correspond au modèle du FJC.

Je souhaite, en guise de conclusion qualitative, donner une estimation des valeurs suivantes  $\langle \mathcal{G} \rangle$  et  $\langle \mathcal{G}^2 \rangle$  qui seront utiles par la suite.

- La distribution de l'énergie cinétique est  $P(E) dE \propto E^{N/2-1} e^{-\beta E} dE$ .
- La moyenne de l'inverse de l'énergie cinétique est alors  $\langle \frac{1}{E} \rangle_E = \frac{2\beta}{N-2}$ .
- Le nombre de degrés de libertés est  $N \geq 6$ , soit au moins un corps rigide.
- Sachant que le processus de Wiener  $\dot{\mathcal{W}}$  est indépendant de  $E$  et que  $\langle \dot{\mathcal{W}} \rangle = 0$ , il est aisé d'obtenir  $\langle \sqrt{\frac{\Sigma^*}{\beta E}} \dot{\mathcal{W}} \rangle = 0$ .
- J'en déduis que  $\langle \mathcal{G} \rangle = \frac{\Sigma^*}{N-2}$ .
- Pour le calcul de  $\langle \mathcal{G}^2 \rangle$  je procède par étape (essentiellement des calculs de fonctions Gamma) :
  1.  $\langle d\mathcal{W}d\mathcal{W} \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt = 1$ .
  2. Le terme croisé est nul puisque  $\dot{\mathcal{W}}$  est indépendant de  $E$  et que  $\langle \dot{\mathcal{W}} \rangle = 0$ .
  3.  $\Sigma^{*2} \left\langle \left( \frac{\langle E \rangle}{E} - \frac{1}{2\beta E} - 1 \right)^2 \right\rangle = \frac{\Sigma^{*2}(2N+1)}{(N-4)(N-2)}$ .
  4.  $\left\langle \frac{\Sigma^*}{\beta E} \dot{\mathcal{W}}^2 \right\rangle \simeq \left\langle \frac{\Sigma^*}{\Delta t \beta E} \Delta \mathcal{W}^2 \right\rangle = \frac{2\Sigma^*}{\Delta t(N-2)}$ .
  5. Soit  $\langle \mathcal{G}^2 \rangle = \frac{\Sigma^{*2}(2N+1)}{(N-4)(N-2)} + \frac{2\Sigma^*}{\Delta t(N-2)}$ .
  6. La variance de  $\mathcal{G}$  est  $\sigma_{\mathcal{G}}^2 = \langle \mathcal{G}^2 \rangle - \langle \mathcal{G} \rangle^2 = \frac{2\Sigma^*}{\Delta t(N-2)} + \frac{\Sigma^{*2}(2N+1)}{(N-4)(N-2)} - \frac{\Sigma^{*2}}{(N-2)^2}$ .
  7. La variance  $\sigma_{\mathcal{G}}^2$  est égale à  $\Sigma^{*2}$  pour  $\Delta t \Sigma^* = \frac{2N^2 - 12N + 16}{N^3 - 10N^2 + 24N - 18}$ . Dans la limite d'un grand nombre de degrés de libertés  $N \gg 1$ , j'ai  $\Delta t \Sigma^* \simeq \frac{2}{N}$ .
- Je déduis de ces estimations que pour un grand nombre de degrés de libertés  $N \gg 1$  et  $\Delta t \Sigma^* \sim \frac{1}{100}$ , j'ai une valeur typique pour  $\mathcal{G}$  telle que  $\langle \mathcal{G} \rangle \pm \sigma_{\mathcal{G}} \lesssim \Sigma^*$ . Cette inégalité est vérifiable numériquement.

Je viens de terminer la description de mon modèle mécanique d'ADN couplé au thermostat global. Je vais maintenant détailler le côté technique de mes simulations.

## 2.8 "Open Dynamics Engine" en pratique

### 2.8.1 Simulation en pratique

Toutes les quantités que j'utilise pour les simulations sont adimensionnées. Le facteur d'adimensionnement pour les longueurs est égal à la longueur d'un cylindre  $l_0 = l$ ; celui pour les masses est

égal à la masse d'un cylindre  $m_0 = m$ . Je choisis d'exprimer toutes les énergies en unités d'agitation thermique c'est à dire  $E_0 = k_B T$ . De ces trois échelles, je déduis l'échelle de temps  $t_0 = l_0 \sqrt{m_0/E_0}$ . Ces adimensionnements me permettent de ramener toutes mes valeurs proches de 1. R. Smith [25] précise qu'il est important, pour le bon fonctionnement du logiciel, que toutes les valeurs soient proches de 1. Il insiste aussi sur le fait qu'il est nécessaire de ne pas avoir des rapports disproportionnés sur les tailles des objets utilisés. Ainsi, pour modéliser la bille magnétique qui possède un diamètre bien supérieur à la taille des cylindres d'ADN, j'utilise une boule de taille  $l$ . Cette boule est placée au centre d'une sphère de volume exclu correspondant à celui de la bille magnétique. Avant de poursuivre sur les valeurs que j'utilise en simulation, je rappelle l'équation qu'ODE doit résoudre, une équation du type  $\mathcal{A}\lambda = \mathbf{b}$  pour trouver le tenseur des contraintes :

$$\left[ \mathcal{J}\mathcal{M}^{-1}\mathcal{J}^T + \frac{k_{cfm}}{\Delta t} \right] \lambda = -\frac{k_{erp}}{\Delta t^2} \delta - \mathcal{J} \left[ \frac{\mathcal{V}}{\Delta t} + \mathcal{M}^{-1} (\mathcal{F}_e - \dot{\mathcal{M}}\mathcal{V}) \right] \quad (2.39)$$

De cette équation, je peux donner quelques ordres de grandeur ; j'ai choisi les longueurs et masses comme étant de l'ordre de 1 et je suppose donc, supposition qui est vérifiable numériquement, que  $\|\delta\| \sim \frac{\sqrt{3}\Delta t^2}{k_{erp}}$  et  $\|\mathcal{J}\mathcal{V}\| \sim \sqrt{3}\Delta t$ . Un choix acceptable pour le pas de temps est donc  $\Delta t \lesssim 0.001$ . ODE propose de choisir des valeurs pour  $k_{erp}$  comprise entre 0.1 et 0.8. Une valeur de 1 combinée à différentes approximations et erreurs internes ne fixera pas totalement l'erreur. Le choix de  $k_{cfm}$  est plus délicat et je l'expliquerai dans la suite. Pour la résolution de l'équation (2.2), je choisis l'algorithme **SOR** avec  $N_\omega = 80$  itérations ; ce nombre est associé à un coefficient de relaxation de  $\omega \sim 1.7$ . Il n'existe pas de théorie pour prévoir le nombre d'itérations  $N_\omega$ . Je choisis  $N_\omega = 80$  car c'est une valeur qui, à peu près tout le temps, donne une bonne précision dans le calcul des  $\lambda$ . La valeur du coefficient de relaxation dépend fortement de  $k_{cfm}$  ; le coefficient optimal est compliqué à calculer (voir par exemple la revue par A. Hadjidimos [50]) et devrait être estimé pour chaque pas de simulation. En effet, il faut calculer la valeur propre maximale  $\rho(\mathcal{A}_\omega)$  de la matrice  $\mathcal{A}_\omega = \mathbf{1} - \mathcal{D}_A^{-1}\mathcal{A}$  où  $\mathcal{D}_A$  est la matrice des valeurs diagonales de  $\mathcal{A}$ . Ceci peut être fait en utilisant quelques itérations de A. Krylov, W. E. Arnoldi, ou de préférence C. Lanczos ; le résultat n'est cependant qu'approximatif. Le coefficient de relaxation optimal  $\omega_{\text{optimal}}$  est alors donné [50] par  $\omega_{\text{optimal}} = \frac{2}{1 + \sqrt{1 - \rho(\mathcal{A}_\omega)^2}}$ . Pour les premiers pas de temps j'implémente les itérations de Lanczos afin de trouver le rayon spectral de  $\mathcal{A}$  et le coefficient de relaxation optimal  $\omega_{\text{optimal}}$ . Ce calcul numérique me donne toujours un  $\omega_{\text{optimal}}$  proche de 1.7. Il est montré que pour une matrice symétrique définie positive  $\mathcal{A}$ , le coefficient de relaxation doit être compris dans l'intervalle  $[0, 2]$  afin que l'algorithme **SOR** converge vers une solution [51]. Il est aussi clair que pour un très long polymère et **LCP**, le temps de calcul est énorme comparé à **SOR**. Je propose alors de faire une simulation avec un polymère de 300 segments et de comparer l'erreur sur les positions d'ancrages pour **LCP** et **SOR**. Je n'ai noté de différences que sur le temps de calcul ! J'opte finalement pour l'algorithme **SOR** avec les paramètres donnés précédemment.

## 2.8.2 Quelques remarques sur les paramètres ERP et CFM

### ERP et CFM comme un ressort amorti

Je rappelle, au préalable, qu'ODE modifie les contraintes cinématiques  $\mathcal{J}\mathcal{V} = \mathbf{0}$  d'un joint "Ball-in-Socket" comme suit  $\mathcal{J}\mathcal{V}(t + \Delta t) = -\frac{k_{erp}}{\Delta t} \delta - k_{cfm} \lambda$ . Cette formulation nous dit que les contraintes  $\lambda$  sont calculées de telle sorte à ajuster la vitesse des cylindres connectés afin d'éliminer une partie  $k_{erp}$  de l'erreur sur le joint. Cependant, les contraintes cinématiques contiennent des

quantités qui sont exprimées en des temps différents. En utilisant l'algorithme d'Euler d'intégration semi-implicite, on a pour  $\delta(t + \Delta t) = \delta + \Delta t \mathcal{J}\mathcal{V}(t + \Delta t)$  et j'arrive à montrer que cette formulation revient à introduire un ressort implicite (dans le sens intégrateur implicite) au niveau du joint (voir l'équation (2.40)). La constante associée à l'amortissement sur la cinématique du joint est  $k_d = \frac{1-k_{erp}}{k_{cfm}}$  et la constante de raideur du ressort sur la contrainte géométrique du joint est  $k_r = \frac{k_{erp}}{\Delta t k_{cfm}}$ .

$$\begin{aligned}\lambda &= -\frac{k_{erp}}{\Delta t k_{cfm}} \delta - \frac{1}{k_{cfm}} \mathcal{J}\mathcal{V}(t + \Delta t) \\ \lambda &= -\frac{k_{erp}}{\Delta t k_{cfm}} \delta(t + \Delta t) - \frac{1-k_{erp}}{k_{cfm}} \mathcal{J}\mathcal{V}(t + \Delta t)\end{aligned}\quad (2.40)$$

Ce calcul général est montré dans un cas plus simple par E. Catto [52]. Le résultat se trouve aussi dans le manuel d'ODE mais sans démonstration. En utilisant ce résultat, je peux calculer les moyennes thermiques  $\langle (\mathcal{J}\mathcal{V}(t + \Delta t))^2 \rangle = \frac{k_{cfm}}{1-k_{erp}}$  et  $\langle \delta(t + \Delta t)^2 \rangle = \frac{\Delta t k_{cfm}}{k_{erp}}$ . En utilisant l'estimation  $\delta(t + \Delta t)^2 \sim \frac{\Delta t^4}{k_{erp}^2}$ , je trouve que  $\Delta t \sim (k_{erp} k_{cfm})^{1/3}$ . Cette observation est très importante pour la détermination du pas de temps  $\Delta t$  sachant  $k_{cfm}$  ou la détermination de  $k_{cfm}$  sachant le pas de temps  $\Delta t$ . Pour ce système de ressort amorti de masse  $m_{eff}$  décrit par l'équation  $m_{eff} \ddot{\delta} + k_d \dot{\delta} + k_r \delta = \mathbf{0}$  de discriminant  $\Delta = k_d^2 - 4m_{eff}k_r$ , il est possible de définir une quantité sans dimensions  $\zeta$  contrôlant l'amplitude de l'amortissement et la fréquence des oscillations  $\omega$  :

$$\zeta = \frac{k_d}{2\omega m_{eff}} \quad \text{et} \quad \omega^2 = \frac{k_r}{m_{eff}} \quad (2.41)$$

soit :

$$\zeta = \frac{k_d}{2\sqrt{k_r m_{eff}}} \quad (2.42)$$

et j'obtiens :

$$\zeta = \frac{1-k_{erp}}{2} \sqrt{\frac{\Delta t}{m_{eff} k_{erp} k_{cfm}}} \quad (2.43)$$

Pour  $\zeta = 0$  et donc  $k_{erp} = 1$ , la valeur de  $\|\delta\|$  oscille au voisinage de 0. Pour  $\zeta > 1$  les oscillations décroissent vers 0 et l'erreur sur la contrainte est corrigée.

### ERP et CFM et leurs rapports avec le thermostat global

Je rappelle que le torseur dynamique du thermostat global n'intervient pas dans le calcul des contraintes mécaniques. Cependant, les erreurs numériques nous obligent à modifier les contraintes cinématiques comme  $\mathcal{J}\mathcal{V}(t + \Delta t) = -\frac{k_{erp}}{\Delta t} \delta - k_{cfm} \lambda$ . Dès lors, le calcul des contraintes mécaniques est dépendant du torseur du thermostat global. Cette modification doit être prise en compte ; je transforme les paramètres  $k_{erp}$  et  $k_{cfm}$  en  $k_{erp}^*$  et  $k_{cfm}^*$ . À partir de l'équation (2.44) modifiée par  $\tilde{\mathcal{G}}$  :

$$\left[ \mathcal{J}\mathcal{M}^{-1} \mathcal{J}^T + \frac{k_{cfm}^*}{\Delta t} \right] \lambda = -\frac{k_{erp}^*}{\Delta t^2} \delta - \left( \frac{1}{\Delta t} + \mathcal{G} \right) \mathcal{J}\mathcal{V} - \mathcal{J} \left[ \mathcal{M}^{-1} (\mathcal{F}_e - \dot{\mathcal{M}}\mathcal{V}) \right] \quad (2.44)$$

dans laquelle j'insère les équations (2.45 et 2.46) :

$$k_{erp}^* = k_{erp} - \frac{\Delta t^2 \mathcal{G}}{\delta^T \delta} \delta^T \mathcal{J} \mathcal{V} \simeq k_{erp} (1 - \Delta t \mathcal{G}) \quad (2.45)$$

$$k_{cfm}^* \simeq k_{cfm} (1 - \Delta t \mathcal{G}) \quad (2.46)$$

j'obtiens l'équation non modifiée par le torseur dynamique du thermostat global (2.47) :

$$\left[ \mathcal{J} \mathcal{M}^{-1} \mathcal{J}^T + \frac{k_{cfm}}{\Delta t} \right] \lambda = -\frac{k_{erp}}{\Delta t^2} \delta - \frac{1}{\Delta t} \mathcal{J} \mathcal{V} - \mathcal{J} \left[ \mathcal{M}^{-1} (\mathcal{F}_e - \dot{\mathcal{M}} \mathcal{V}) \right] \quad (2.47)$$

Sachant que  $\mathcal{G} \lesssim \Sigma^*$ , et que  $100\Sigma^* \sim \frac{1}{\Delta t}$ , il est clair que le terme  $\mathcal{G}$  est négligeable devant  $\frac{1}{\Delta t}$  et n'affecte donc que très peu le calcul de  $\lambda$ . Les corrections  $k_{erp} \rightarrow k_{erp}^*$  et  $k_{cfm} \rightarrow k_{cfm}^*$  ne sont pas forcément nécessaires. Si à la place du torseur dynamique du thermostat global nous préférons utiliser un rescaling global  $\alpha$  du torseur cinématique, il faut transformer les paramètres  $k_{erp}$  et  $k_{cfm}$  en  $k_{erp}^* = k_{erp} (2 - \alpha)$  et  $k_{cfm}^* = k_{cfm} (2 - \alpha)$ .

### ERP et CFM et leurs rapports avec la persistance

Je m'intéresse alors aux effets des paramètres ERP et CFM sur la persistance et le twist du modèle BLC. Le torseur dynamique de courbure au niveau du joint connectant le cylindre 1 avec le cylindre 2 est  $\mathcal{F}^b = \{0, g_b t_1 \otimes t_2 - g_b t_0 \otimes t_1, 0, g_b t_2 \otimes t_3 - g_b t_1 \otimes t_2\}^T$ . Je calcule alors le terme  $-\mathcal{J} \mathcal{M}^{-1} \mathcal{F}^b$  de l'équation (2.2) sur  $\lambda$ . Je rapelle au préalable que  $\mathcal{J} = \{1, -\frac{l}{2} t_1^\otimes, -1, -\frac{l}{2} t_2^\otimes\}$  et que  $a \otimes (b \otimes c) = (a^T c) b - (a^T b) c$ .

$$-\mathcal{J} \mathcal{M}^{-1} \mathcal{F}^b = \frac{l g_b}{2I_\perp^*} ((\cos \theta_{01} + \cos \theta_{12} - 1) t_1 + (\cos \theta_{12} + \cos \theta_{23} - 1) t_2 - t_0 - t_3) \quad (2.48)$$

Le torseur dynamique de twist au niveau du joint connectant le cylindre 1 avec le cylindre 2 est  $\mathcal{F}^t = \left\{ 0, \frac{g_t \phi_{12}}{1 + \cos \theta_{12}} (t_1 + t_2) - \frac{g_t \phi_{01}}{1 + \cos \theta_{01}} (t_0 + t_1), 0, \frac{g_t \phi_{23}}{1 + \cos \theta_{23}} (t_2 + t_3) - \frac{g_t \phi_{12}}{1 + \cos \theta_{12}} (t_1 + t_2) \right\}^T$ . Je calcule dès lors le terme  $-\mathcal{J} \mathcal{M}^{-1} \mathcal{F}^t$  de l'équation (2.2) sur  $\lambda$ .

$$-\mathcal{J} \mathcal{M}^{-1} \mathcal{F}^t = \frac{l g_t}{2I_\perp^*} \left( \frac{2\phi_{12}}{1 + \cos \theta_{12}} t_1 \otimes t_2 - \frac{\phi_{01}}{1 + \cos \theta_{01}} t_1 \otimes t_0 - \frac{\phi_{23}}{1 + \cos \theta_{23}} t_2 \otimes t_3 \right) \quad (2.49)$$

Le terme  $\mathcal{J} \mathcal{M}^{-1} \mathcal{F}^{b+t}$  est modifié par le terme  $-\frac{k_{erp}}{\Delta t^2} \delta - \frac{1}{\Delta t} \mathcal{J} \mathcal{V}$  dans le calcul des forces de contraintes  $\mathcal{J}^T \lambda$ . Malheureusement, le terme provenant des erreurs numériques ne donne pas seulement des couples de corrections au centre de masse mais aussi des forces de corrections au centre de masse des cylindres. Pour l'instant, je suis en train travailler sur une solution analytique pour  $\lambda$  dans le cas d'un BLC. J'en déduis qu'il me faut autre chose pour calculer l'influence de ERP sur la courbure et le twist. Je remarque que le couple de contrainte  $\Gamma_1^\lambda$  au centre de masse du cylindre 1 est l'opposé de celui  $\Gamma_2^\lambda$  au centre de masse du cylindre 2 le long de la normale au plan de twist soit  $(t_1 + t_2)^T \Gamma_1^\lambda = -(t_1 + t_2)^T \Gamma_2^\lambda$ . Je remarque aussi que  $(t_1 \otimes t_2)^T \Gamma_1^\lambda \neq -(t_1 \otimes t_2)^T \Gamma_2^\lambda$ . Le principe d'action-réaction le long de la binormale n'est pas vrai. Cela provient du terme  $-\frac{k_{erp}}{\Delta t^2} \delta - \frac{1}{\Delta t} \mathcal{J} \mathcal{V}$ . Il est alors possible de calculer la correction à apporter à la constante de courbure  $g_b$  qui est d'ailleurs nulle dans le cas où  $q = 1$  soit dans le cas d'une chaîne FJC. Par exemple, pour une valeur de  $q = 30$  (30 segments de 10 bp par segment de Kuhn statistique) la constante de courbure est  $g_b \simeq 15.5$  et la correction est  $\delta g_b \simeq 0.7$ . Il n'y a aucune correction à apporter à la constante de twist  $g_t$ .

### 2.8.3 Les différentes étapes de la simulation

Pour obtenir une suite de nombres aléatoires, j'utilise le générateur Mersenne-Twister [53] qui possède une très grande période  $10^{6001} - 1 < 2^{19937} - 1 < 10^{6002} - 1$  et qui est particulièrement rapide. Le générateur est initialisé avec une graine dépendant du temps du processeur. La période est largement suffisante pour mes simulations ; je ne reviendrai pas tout de suite au premier nombre aléatoire de la suite ! Les bruits blancs gaussiens (ou processus de Wiener  $\Delta W$  divisé par  $\Delta t$ ) sont obtenus à partir de l'algorithme de Box-Muller [54]. En pratique, il faut calculer la différence  $\Delta W = \mathcal{W}(t + \Delta t) - \mathcal{W}(t)$  qui revient à générer un nombre aléatoire suivant une loi normale  $\mathcal{N}(0, \Delta t)$  de moyenne nulle et d'écart type  $\Delta t$ . Pour m'assurer de la reproductibilité des simulations avec ODE, je réalise de nombreuses fois la même simulation, avec la même graine, et trace la trajectoire des cylindres. La trajectoire est à chaque fois identique à  $10^{-10}$  près.

1. Initialisation du générateur de nombres aléatoires [53] avec une graine dépendant du temps du processeur.
2. Création des corps et des géométries associées (cylindre de longueur  $l$  et de rayon  $r$ , bille magnétique de rayon  $R$  ...).
3. Création du sol, c'est à dire la plaque de verre sur laquelle est greffé l'ADN.
4. Création des joints.
5. Génération aléatoire de la configuration initiale.
6. Initialisation des paramètres  $k_{erp} = 0.8$ ,  $k_{cfm} = 10^{-9}$ ,  $N_\omega = 100$  et  $\omega \sim 1.7$ .
7. La détermination du pas de temps est dictée par la résolution des contraintes  $\Delta t \sim (k_{erp}k_{cfm})^{1/3}$  et, dans une moindre mesure, par la fréquence de couplage au thermostat global. Le pas de temps est typiquement de l'ordre de  $\Delta t \sim 10^{-3}$  soit un  $\Sigma^* = 10$  ( $\Sigma^* \Delta t \sim \frac{1}{100}$ ).
8. Calcul des constantes de courbure  $g_b$  et de twist  $g_t$  en fonction de la taille des cylindres d'ADN.
9. Création de l'espace multi-résolution avec une maille de longueur  $R \leq 2^n$  où  $n$  est à calculer.
10. Initialisation des vitesses selon une dynamique de Langevin locale.
11. Boucle sur les tous les pas de temps.
  - (a) Ajout des couples de courbure et de twist  $\mathbf{\Gamma}^{b+t}$ .
  - (b) Ajout du torseur dynamique du thermostat global  $\tilde{\mathcal{G}}$  ou rescaling global  $\alpha$  du torseur cinématique.
  - (c) Éventuellement modification des paramètres  $k_{erp}$  et  $k_{cfm}$ .
  - (d) ODE résout les contraintes (algorithme **SOR**) et fait évoluer le système du temps  $t$  au temps  $t + \Delta t$ . Si des collisions sont présentes dans le système au temps  $t$ , il est préférable d'utiliser l'algorithme **LCP**.
  - (e) Contrôle de l'erreur sur les contraintes  $\|\boldsymbol{\lambda} + \frac{k_{erp}}{\Delta t k_{cfm}} \boldsymbol{\delta} + \frac{1}{k_{cfm}} \mathcal{J}\mathcal{V}(t + \Delta t)\| / \|\boldsymbol{\lambda}\|$ . Si l'erreur est trop importante ( $> 10^{-4}$ ), je recommence le pas de temps avec plus d'itérations ou avec l'algorithme **LCP**. En moyenne, sur une simulation, l'erreur sur  $\boldsymbol{\lambda}$  est de l'ordre de  $\sim 10^{-6}$ . Cette valeur est négligeable devant les forces et couples caractéristiques (thermostat global, couples de courbure et de twist).

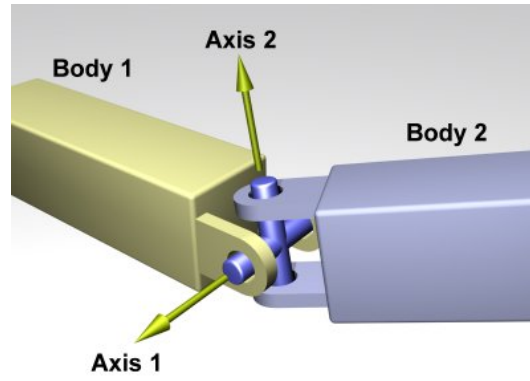


FIGURE 2.5: Vue schématique d'un joint de Cardan. ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.

## 2.9 Chaîne de Cardan

Historiquement, si je puis me permettre ce terme, H. Wong et J. Mozziconacci [55] ont développé une méthode de simulation de l'ADN basée sur des joints de Cardan ou joints universels (voir la figure (2.5) avec angles limites pour mimer la résistance à la courbure de l'ADN. Nous allons voir dans la suite quelques résultats amusants que j'ai obtenu sur ce type de polymère Cardan-Like-Chain (CLC) ou Universal-Like-Chain (ULC) et pourquoi je l'ai remplacé par le modèle dit BLC.

### 2.9.1 Définition

Commençons par construire notre chaîne de Cardan. Pour cela, nous partons de deux segments alignés de taille  $l$  et connectés par un joint de Cardan. Ces deux segments sont définis par deux repères matériels  $\{t_1, m_1, \widetilde{m}_1\}$  et  $\{t_2, m_2, \widetilde{m}_2\}$ . Le joint de Cardan impose  $m_1^T \widetilde{m}_2 = 0$ ; cette relation définit une croix. Nous nous demandons alors quelle est la relation entre les deux repères matériels.

- Après avoir fait une rotation d'axe  $m_1$  et d'angle  $\xi_1$  nous avons :

$$\widetilde{m}_2 = \cos \xi_1 \widetilde{m}_1 + \sin \xi_1 t_1 \quad (2.50)$$

- Par la suite nous faisons une rotation d'axe  $\widetilde{m}_2$  d'un angle  $\xi_2$  et nous obtenons :

$$t_2 = \sin \xi_2 m_1 - \cos \xi_2 \sin \xi_1 \widetilde{m}_1 + \cos \xi_1 \cos \xi_2 t_1 \quad (2.51)$$

$$m_2 = \cos \xi_2 m_1 + \sin \xi_1 \sin \xi_2 \widetilde{m}_1 - \cos \xi_1 \sin \xi_2 t_1 \quad (2.52)$$

- La binormale s'écrit alors comme suit avec le cosinus de l'angle de courbure  $\cos \theta = \cos \xi_2 \cos \xi_1$ .

$$b_{12} = \frac{t_1 \otimes t_2}{\|t_1 \otimes t_2\|} = \frac{1}{\sqrt{1 - \cos^2 \theta}} (\sin \xi_1 \cos \xi_2 m_1 + \sin \xi_2 \widetilde{m}_1) \quad (2.53)$$

- La normale est donnée par l'expression suivante.

$$n_{12} = b_2 \otimes t_2 = \frac{1}{\sqrt{1 - \cos^2 \theta}} (\cos \theta \sin \xi_2 m_1 - \sin \xi_1 \cos \xi_2 \cos \theta \widetilde{m}_1 - \sin^2 \theta t_1) \quad (2.54)$$

Voici les principales relations pour deux corps liés par un joint de Cardan.

$$\begin{aligned} m_1 \otimes \widetilde{m}_2 &= \cos \xi_1 t_1 - \sin \xi_1 \widetilde{m}_1 \\ t_2 \otimes m_1 &= \cos \xi_2 \widetilde{m}_2 \end{aligned} \quad (2.55)$$

## 2.9.2 Calcul du twist

Nous avons déjà obtenu le cosinus de l'angle de courbure. Le twist est quant à lui défini (voir les expressions (2.12)) à partir des angles d'Euler et est montré sur la figure (2.6).

$$\cos \phi = \frac{\mathbf{m}_2^T \mathbf{m}_1 + \widetilde{\mathbf{m}}_2^T \widetilde{\mathbf{m}}_1}{1 + \mathbf{t}_1^T \mathbf{t}_2} = \frac{\cos \xi_1 + \cos \xi_2}{1 + \cos \xi_1 \cos \xi_2} \quad (2.56)$$

$$\sin \phi = \frac{\mathbf{m}_2^T \widetilde{\mathbf{m}}_1 - \widetilde{\mathbf{m}}_2^T \mathbf{m}_1}{1 + \mathbf{t}_1^T \mathbf{t}_2} = \frac{\sin \xi_1 \sin \xi_2}{1 + \cos \xi_1 \cos \xi_2} \quad (2.57)$$

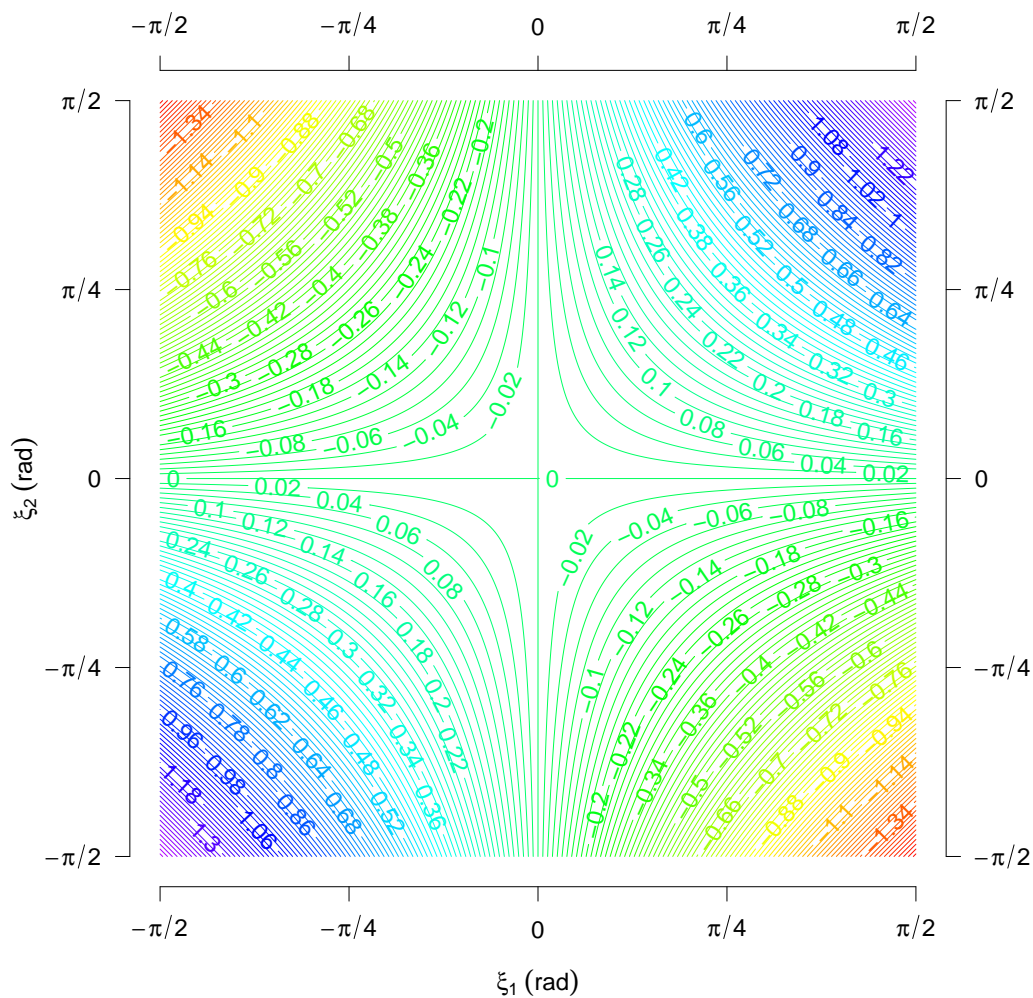


FIGURE 2.6: Carte du twist pour un joint de Cardan. En abscisse et ordonnée sont représentés en radians les angles  $\xi_1$  et  $\xi_2$ . Les contours de couleurs indiquent la valeur du twist.

Il est facile de montrer qu'une chaîne de Cardan dans la limite continue ne possède pas de twist (voir la figure (2.6)). En particulier, le carré du twist développé à l'ordre 2 au voisinage de 0 en  $\xi_1$  et  $\xi_2$  donne  $\phi^2 \simeq 0$ . Ce résultat nous dit que la persistance en twist est infinie. Dans le cas continu,

le repère matériel en fonction de l'abscisse curviligne est donné par les équations (2.58); un parallèle avec les équations du repère de Frenet est facile à voir.

$$\dot{t} = \dot{\xi}_2 m - \dot{\xi}_1 \widetilde{m} \quad \text{et} \quad \dot{m} = -\dot{\xi}_2 t \quad \text{et} \quad \dot{\widetilde{m}} = \dot{\xi}_1 t \quad (2.58)$$

### 2.9.3 Persistance

Hua a ajouté une condition supplémentaire pour mimer la courbure; une condition sur l'angle maximum  $m$  suivant les deux axes que peut atteindre le joint de Cardan. Cette méthode remplace l'utilisation de couple de courbure et de twist. La persistance en courbure est donnée par [36] et adaptée au joint de Cardan est  $\langle \cos \xi_1 \cos \xi_2 \rangle_{\xi_1, \xi_2} = k - l/k + l$ . L'intégrale sur  $\xi_1$  et  $\xi_2$  et sur l'intervalle  $[-m; m]$  nous permet d'écrire l'équation (2.60). Les angles  $\xi_1$  et  $\xi_2$  sont indépendants et uniformément équidistribués.

$$\langle \cos \xi_1 \cos \xi_2 \rangle_{\xi_1, \xi_2} = \frac{\int_{-m}^m \cos \xi_1 d\xi_2 \int_{-m}^m \cos \xi_1 d\xi_1}{\int_{-m}^m d\xi_2 \int_{-m}^m d\xi_1} \quad (2.59)$$

$$\frac{\sin^2 m}{m^2} = \frac{k - l}{k + l} \quad (2.60)$$

Je me pose maintenant la question de la persistance en twist pour un joint de Cardan. En utilisant un potentiel harmonique pour le twist de raideur  $t/l$ , la moyenne thermique nous permet d'écrire  $\langle \phi^2 \rangle = l/t$ . Je suppose que la moyenne thermique  $\langle \cdot \rangle$  est équivalente à la moyenne configurationnelle  $\langle \cdot \rangle_{\xi_1, \xi_2}$ . Je développe (voir l'équation (2.61)) le cosinus de la courbure et le twist à l'ordre 4 (le développement du twist à l'ordre 2 est nul).

$$\cos \theta \simeq \frac{\xi_1^4 + \xi_2^4}{24} + \frac{\xi_1^2 \xi_2^2}{4} - \frac{\xi_1^2 + \xi_2^2}{2} + 1 \quad \text{et} \quad \phi^2 \simeq \frac{\xi_1^2 \xi_2^2}{4} \quad (2.61)$$

Je peux alors relier  $\cos \theta$  à  $\phi^2$  comme  $\cos \theta \simeq \frac{\xi_1^4 + \xi_2^4}{24} - \frac{\xi_1^2 + \xi_2^2}{2} + 1 + \phi^2$ . La moyenne thermique (2.62) de cette équation donne l'angle maximal  $m$  sur les angles  $\xi_1$  et  $\xi_2$  qui nous permet de reproduire simultanément les persistances en courbure et en twist.

$$\frac{k - l}{k + l} - \frac{l}{t} \simeq \frac{m^4}{60} - \frac{m^2}{3} + 1 \quad (2.62)$$

### 2.9.4 Matrice de transfert itérative et ULC

Il est possible de construire une chaîne de Cardan de manière itérative (voir l'équation (2.63)) en définissant un repère matériel  $\mathbf{F}_1 = \{t_1, m_1, \widetilde{m}_1\}$  et une matrice de transfert  $\mathbf{R}(\xi_1, \xi_2)$  qui amène  $\mathbf{F}_1^T$  sur le repère suivant  $\mathbf{F}_2^T = \mathbf{R}(\xi_1, \xi_2) \mathbf{F}_1^T$ . La matrice  $\mathbf{R}(\xi_1, \xi_2)$  est de déterminant 1 et sa transposée égale son inverse.

$$\mathbf{R}(\xi_1, \xi_2) = \begin{pmatrix} \cos \xi_1 \cos \xi_2 & \sin \xi_2 & -\sin \xi_1 \cos \xi_2 \\ -\cos \xi_1 \sin \xi_2 & \cos \xi_2 & \sin \xi_1 \sin \xi_2 \\ \sin \xi_1 & 0 & \cos \xi_1 \end{pmatrix} \quad (2.63)$$

Les corrélations sont alors données par (2.64) où nous posons  $a = m^{-1} \sin m$ .



$$\begin{aligned}
\langle \mathbf{t}_j^T \mathbf{t}_i \rangle_{\xi_1, \xi_2} &= a^{2|j-i|} \\
\langle \mathbf{m}_j^T \mathbf{m}_i \rangle_{\xi_1, \xi_2} &= a^{|j-i|} \\
\langle \widetilde{\mathbf{m}}_j^T \widetilde{\mathbf{m}}_i \rangle_{\xi_1, \xi_2} &= a^{|j-i|}
\end{aligned} \tag{2.64}$$

Pour le calcul de la moyenne des configurations sur  $\{\theta, \xi\}$ , nous utilisons les propriétés suivantes : la moyenne d'une somme est la somme des moyennes ; l'orientation du premier segment ne dépend pas des angles  $\xi_1$  et  $\xi_2$  et la longueur de persistance de cette chaîne s'estime comme (2.65).

$$\begin{aligned}
p &= \lim_{N \rightarrow \infty} \langle \mathbf{t}_1^T \mathbf{R}(N) \rangle_{\xi_1, \xi_2} = \frac{l}{1-a^2} \\
\lim_{N \rightarrow \infty} \langle \mathbf{R}(N) \rangle_{\xi_1, \xi_2} &= pa^2 \mathbf{t}_0 \\
\lim_{N \rightarrow \infty} \langle \mathbf{R}(N)^2 \rangle_{\xi_1, \xi_2} &= Lp(1+a^2)
\end{aligned} \tag{2.65}$$

L'énergie due à l'application d'une force  $\mathbf{f} = fz$  à l'extrémité de la chaîne s'écrit  $-\mathbf{f}^T \mathbf{R}(N)$  et la fonction de partition s'écrit :

$$Z(\beta, f, l, m, N) = \int_{-m}^m d\{\xi_1, \xi_2\} e^{\beta \mathbf{f}^T \mathbf{R}(N, \{\xi_1, \xi_2\})} \tag{2.66}$$

Dans le cas où  $f = 0$ , la fonction de partition et l'énergie libre s'écrivent comme suit (2.67).

$$\begin{aligned}
Z_{\text{CLC}}(f = 0, m, N) &= (2m)^{2N} \\
-\beta F_{\text{CLC}}(f = 0, m, N) &= 2N \ln 2m
\end{aligned} \tag{2.67}$$

Dans le cas où  $f = 0$ , et d'une chaîne BLC avec angles limites en lieu et place d'énergie de courbure, j'ai :

$$\begin{aligned}
Z_{\text{BLC}}(f = 0, m, N) &= \left( \frac{2\pi}{1 - \cos m} \int_0^m \sin \theta d\theta \right)^N \\
&= (\pi(1 - \cos m))^N \\
-\beta F_{\text{BLC}}(f = 0, m, N) &= N \ln \pi(1 - \cos m)
\end{aligned} \tag{2.68}$$

En utilisant cette matrice de transfert, nous pouvons reconstruire de proche en proche une chaîne de Cardan, appliquée une force à son extrémité et échantillonner l'espace des configurations lui étant accessible avec l'algorithme de Metropolis (voir l'annexe (A.2)). La réponse d'une chaîne ULC à une force de traction est montrée sur la figure (2.7) et ce pour différentes longueurs d'ADN. En particulier, je montre les effets de la longueur de la molécule d'ADN sur les résultats de simulations dans la limite des faibles forces. Dans la limite des grandes forces le modèle ULC s'écarte du modèle du WLC pour des raisons entropiques.

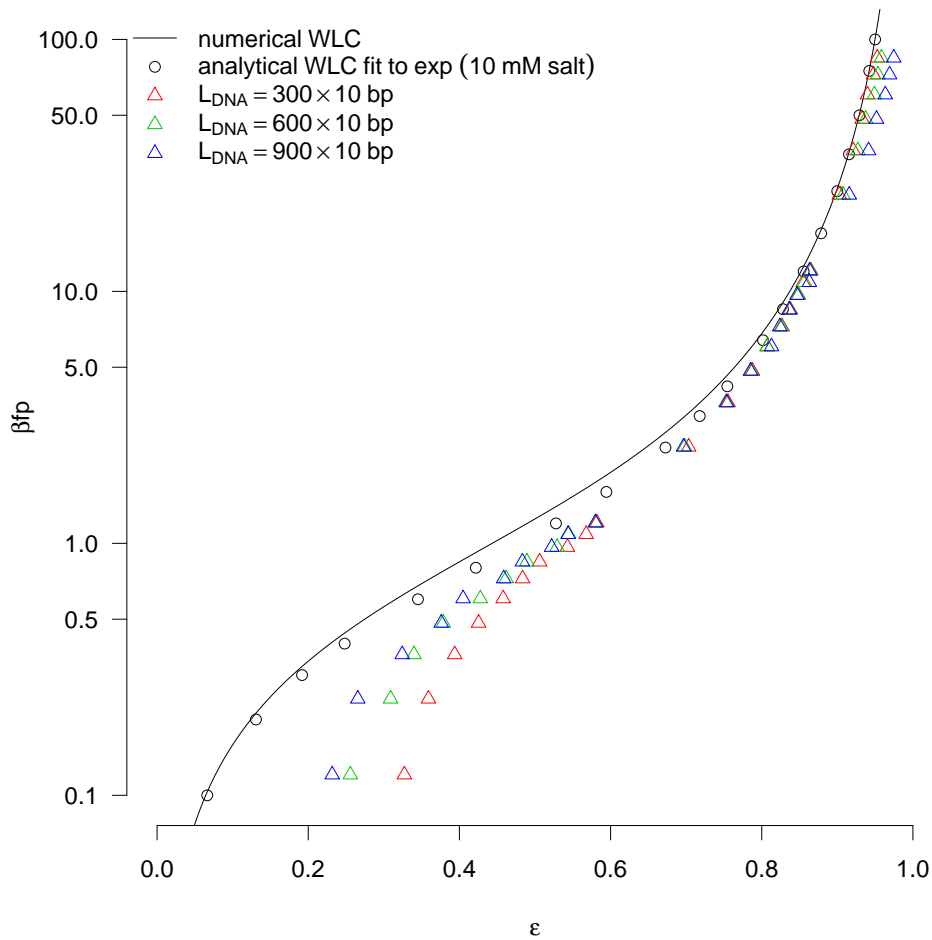


FIGURE 2.7: Le profil force-extension est montré pour les modèles simulés du BLC, du Freely-Jointed-Chain, du WLC analytique et du WLC numérique [1, 2]. La quantité  $\beta fp$  est la force adimensionnée et est représentée en échelle logarithmique tandis que  $\epsilon = z/L$  est l'extension relative de la molécule d'ADN. Trois longueurs d'ADN sont simulées  $L = 1, 2, 3 \mu\text{m}$  (triangles rouges, verts et bleus). Les erreurs statistiques sont inférieures à la taille des symboles.



---

# ADN nu sous pinces magnétiques ; “in silico”

---

Après avoir construit le modèle, la première simulation que j’ai faite modélise une molécule d’ADN de longueur  $L = 1 \mu\text{m}$  dans un bain thermique. Cette situation simple me permet de calibrer les persistances en courbure et en twist et me permet aussi d’introduire les paramètres des simulations que je vais présenter.

J’ai choisi d’avoir 30 segments par segment de Kuhn soient des segments de longueur  $l = 3.34 \text{ nm}$  (10 bp), une longueur de Kuhn de  $k = 100 \text{ nm}$  et une persistance en twist de  $t = 95 \text{ nm}$ . Cette taille de segment est approximativement égale à une période de double hélice d’ADN (10.5 bp), longueur qui sera utilisée par la suite pour modéliser les  $14 \times 10.5 = 147 \text{ bp}$  d’ADN du nucléosome. La température est égale à  $T = 300 \text{ K}$ . La constante de courbure vaut  $\beta g_b \simeq 15.499$  et celle en twist est  $\beta g_t \simeq 28.443$ . La fréquence de couplage au thermostat global est choisie de telle sorte que tous les temps caractéristiques soient du même ordre de grandeur. Ce choix est arbitraire. Les propriétés qui nous intéressent sont des propriétés d’équilibre et non de dynamique ; ceci explique l’utilisation du thermostat global et d’une fréquence de couplage telle que les couples et forces fluctuent sur les mêmes échelles de temps. Je rappelle que le temps caractéristique associé au thermostat est  $1/\Sigma^*$  et ceux associés au couple de courbure (respectivement au couple de twist) sont  $I^*\Sigma^*/g_b$  (respectivement  $I^*\Sigma^*/g_t$ ). Je choisis donc  $1/\Sigma^* = \sqrt{\sqrt{I^*/g_b}\sqrt{I^*/g_t}}$ , quantité qui peut-être vue comme une moyenne géométrique et qui donne un pas de temps  $\Delta t \sim 0.001$ .

En utilisant la très jolie formule pour les corrélations tangentielles le long de la chaîne articulée  $\langle \mathbf{t}_i \mathbf{t}_j^T \rangle = \exp -|i - j|l/p$  de Kratky et Porod [9] nous pouvons tirer de la simulation la moyenne d’ensemble  $\langle \mathbf{t}_i \mathbf{t}_j^T \rangle$  et connaissant  $l$ , déterminer la persistance simulée  $p$ . Pour le twist, nous calculons la moyenne de son carré  $\langle \phi^2 \rangle$  et la comparons à la valeur théorique  $l/t$ . Ces deux comparaisons nous donnent un très bon accord ; la valeur simulée pour la moyenne du cosinus de la courbure est  $\langle \cos \theta \rangle = 0.9359$  contre une valeur théorique de 0.9355. La valeur simulée pour la moyenne du carré du twist  $\langle \phi^2 \rangle = 0.9823$  contre une valeur théorique de 0.9824.

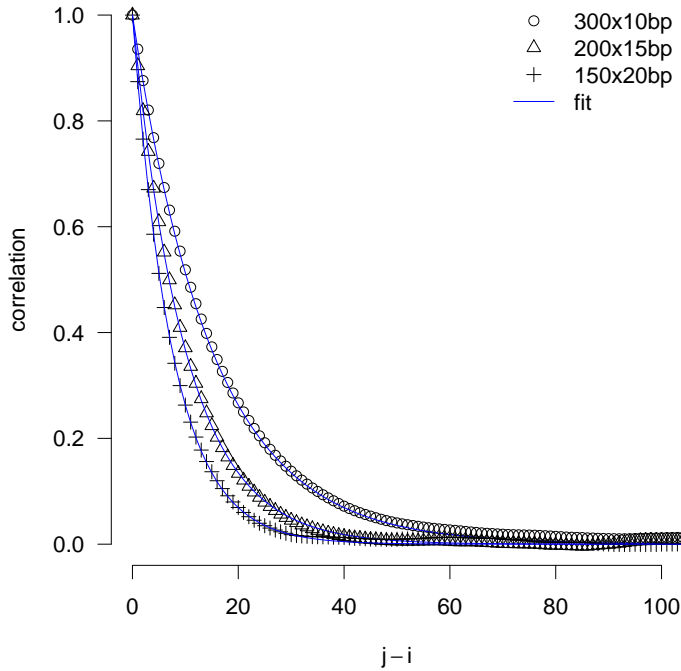


FIGURE 3.1: Les corrélations tangentielles sont montrées (en ordonnée) le long de l'ADN (en abscisse). Les simulations sont réalisées pour différentes segmentations  $l = 3.34, 5, 6.67$  nm et pour une molécule d'ADN de longueur  $L = 1 \mu\text{m}$ . Les courbes théoriques sont montrées par les lignes bleues.

### 3.1 Comparaison du thermostat local et de la version globale

Nous reprenons la simulation présentée en guise d'introduction. On s'intéresse alors au temps de décorrélation  $\tau_d$  de différentes quantités : l'énergie cinétique du système  $E_b + E_t$  et la distance bout à bout  $\mathbf{R}$ . Pour cela, on trace la fonction d'autocorrélation  $r_X(\tau) = c_X(\tau) / c_X(0)$  (ou autocovariance normalisée) d'une observable  $X$  définie par l'équation (3.1). Cette définition, estimateur de la fonction d'autocorrélation, est issue du livre [56] sur le logiciel S. Le coefficient  $1/n$  et non  $1/(n - |\tau|)$  garantit que la séquence  $c_X(\tau)$  est la covariance d'une série temporelle stationnaire du second ordre. Le symbole  $\bar{X}$  désigne une moyenne sur la série de mesures  $X_1, \dots, X_n$  associée à l'observable  $X$ .

$$c_X(\tau) = \frac{1}{n} \sum_{\tau=\max 1, -t}^{\min n-\tau, n} [X_{t+\tau} - \bar{X}] [X_t - \bar{X}] \quad (3.1)$$

Le temps de décorrélation  $\tau_d$  nous renseigne sur le temps mis par l'observable pour oublier son passé. Il nous permet en particulier de dire que la prise de mesures de  $X$  doit se faire à des intervalles de temps supérieurs à  $\tau_d$ . Imaginons alors que nous ayons deux algorithmes dont nous voulons comparer l'efficacité ; il nous suffit de calculer  $\tau_{d,1}$  et  $\tau_{d,2}$  et de garder celui qui est le plus efficace. Je propose donc de comparer la version locale du thermostat de Langevin avec sa version globale exposée précédemment. Dans les expériences de pinces magnétiques nous nous intéressons à la taille de la molécule d'ADN et donc à la distance bout à bout  $\mathbf{R}$ . La fonction d'autocorrélation associée que je calcule en simulation est donnée par l'équation (3.2).

$$c_{\mathbf{R}}(0)r_{\mathbf{R}}(\tau) = \frac{1}{n} \sum_{\tau=\max 1,-t}^{\min n-\tau,n} \sum_{i=x,y,z} [R_{i,t+\tau} - \bar{R}_i] [R_{i,t} - \bar{R}_i] \quad (3.2)$$

Les paramètres de simulation sont ceux utilisés précédemment. Les résultats sont montrés sur les figures (3.2 et 3.3). Pour obtenir le temps de décorrélation de l'énergie cinétique  $T + R$  du système j'ajuste une exponentielle  $e^{-t/\tau_E}$  de paramètre  $\tau_g^E = 0.1401 \pm 0.0008$  ns dans le cas du thermostat global et  $\tau_l^E = 0.197 \pm 0.004$  ns dans le cas du thermostat local. L'énergie cinétique  $T + R$  du système décorrèle aussi vite dans les deux cas. Pour obtenir le temps de décorrélation de la distance bout à bout de la molécule d'ADN simulée, j'ajuste une exponentielle sur la fonction d'autocorrélation dans le cas du thermostat local et une enveloppe exponentiellement décroissante  $e^{-t/\tau_g^R} \cos \omega t$  dans le cas du thermostat global. Nous avons  $\tau_l^R = 12079.51 \pm 4.61$  ns,  $\tau_g^R = 177.72 \pm 4.61$  ns et  $\omega = 503.15$  ns<sup>-1</sup>.

Je me suis aussi intéressé à la diffusion de la distance bout-à-bout normalisée (équation (3.3)). Le résultat est montré sur la figure (3.4). La diffusion est limitée par la taille du polymère et a pour valeur limite  $2Lk$ , valeur qui sert de normalisation pour une éventuelle étude en taille de la molécule d'ADN. La mesure de la diffusion est quelque peu redondante avec la mesure de la fonction d'autocorrélation de la distance bout à bout. En effet, la diffusion de la distance bout à bout  $\langle (\mathbf{R}(t+\tau) - \mathbf{R}(t))^2 \rangle$  nous donne  $\langle \mathbf{R}^2(t+\tau) \rangle - 2\langle \mathbf{R}^T(t+\tau) \mathbf{R}(t) \rangle + \langle \mathbf{R}^2(t) \rangle$  et dans le cas limite où  $\tau$  est grand devant le temps de décorrélation de  $\mathbf{R}$  le terme de corrélation  $\langle \mathbf{R}^T(t+\tau) \mathbf{R}(t) \rangle$  est nul, et  $\langle \mathbf{R}^2(t+\tau) \rangle = \langle \mathbf{R}^2(t) \rangle$  et on trouve la valeur limite  $2Lk$ . Dans le cas global, la diffusion est plus rapide; la diffusion atteint sa valeur limite (valeur de saturation) de 1 pour un temps  $\Sigma^* \tau_g^s \simeq 10^4$  soit 100 fois plus vite que pour la version locale où  $\Sigma^* \tau_l^s \simeq 10^6$ .

$$\frac{\langle (\mathbf{R}(t+\tau) - \mathbf{R}(t))^2 \rangle}{2Lk} \quad (3.3)$$

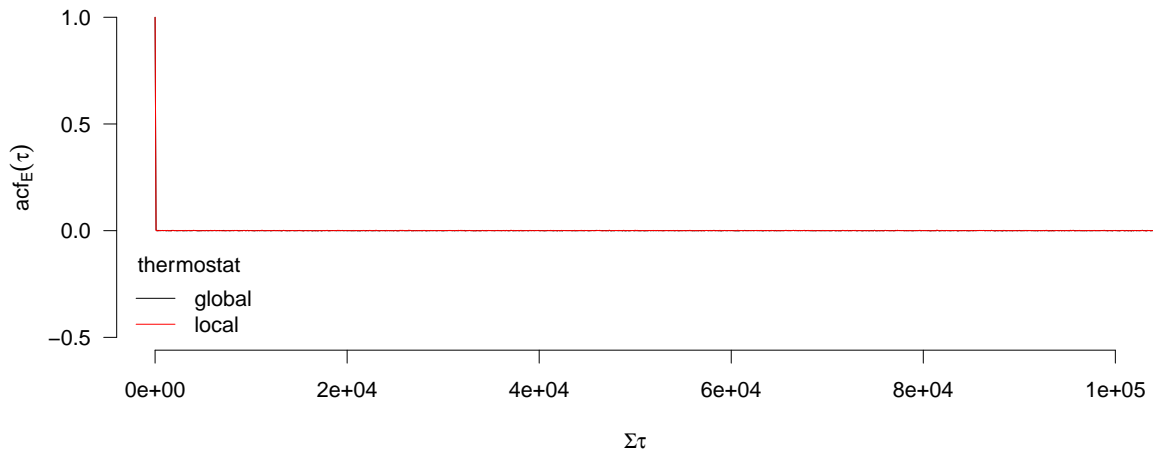


FIGURE 3.2: Les fonctions d'autocorrélation (en ordonnée) en fonction du temps de simulation adimensionné  $\Sigma\tau$  (en abscisse) de l'énergie cinétique  $T + R$  du système sont montrées pour les deux versions du thermostat de Langevin. Les courbes sont ajustées par une exponentielle.

Vous pourrez trouver une vidéo sur la comparaison entre le thermostat local et la version globale sur mon compte “vimeo” Pascal Carrivain à l'adresse <http://vimeo.com/51918121>. La vidéo

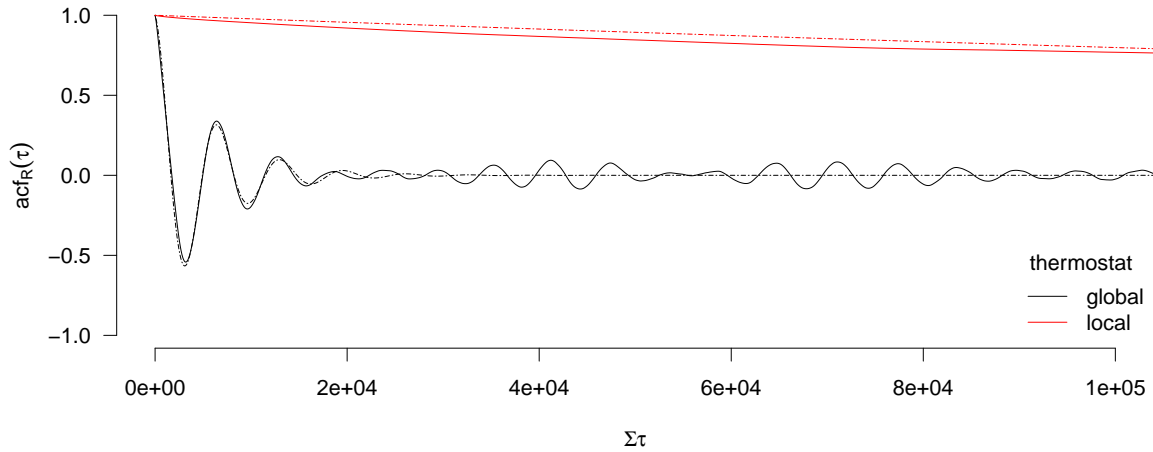


FIGURE 3.3: Les fonctions d'autocorrélation (voir l'équation (3.2)) (en ordonnée) de la distance bout à bout en fonction du temps de simulation  $\Sigma\tau$  adimensionné (en abscisse) sont montrées pour les deux versions du thermostat de Langevin. La courbe dans le cas du thermostat local est ajustée par une exponentielle décroissante tandis que la courbe dans le cas du thermostat global est ajustée par une enveloppe exponentielle décroissante.

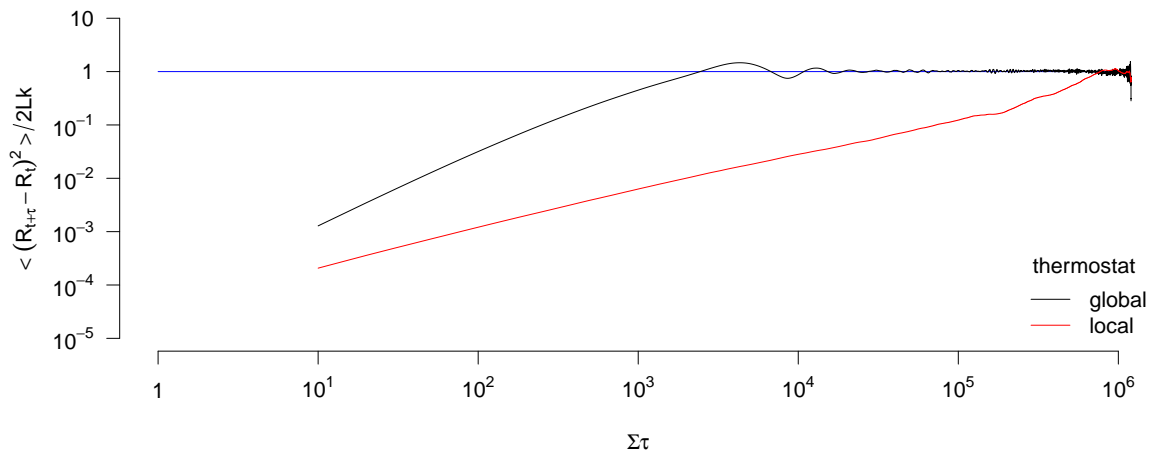


FIGURE 3.4: La diffusion de la distance bout à bout normalisée  $\langle (R(t + \tau) - R)^2 \rangle / 2Lk$  (en ordonnée) en fonction du temps de simulation  $\Sigma\tau$  adimensionné (en abscisse) est montrée dans la version locale (rouge) et globale (noire) du thermostat de Langevin. L'échelle est en log-log.

montre une simulation où j'utilise de manière alternée le thermostat local et le thermostat global. La simulation est pour un ADN de longueur  $L = 1 \mu\text{m}$ , sous une force de traction de  $f = 0.74 \text{ pN}$  et un nombre de tours  $n = 15$ . Toutes les simulations qui vont être présentées par la suite sont réalisées une fois avec le thermostat local et une autre fois avec le thermostat global. Je n'ai observé aucune différence dans les résultats de ces expériences numériques qu'elles soient menées par l'un ou l'autre de ces thermostats.

### 3.2 ADN en traction pure

Nous passons maintenant à une situation plus intéressante où une extrémité de l'ADN est accrochée tangentiellement au sol tandis que l'autre extrémité rejoint la bille. La bille a un rayon  $2\pi R > L$  suffisamment grand pour que la molécule ne s'enroule pas autour. Cette construction provient directement de la conception des expériences de pinces magnétiques. La bille est soumise à un effort de traction  $f$  dans la direction  $\hat{z}$  perpendiculaire au sol. En reprenant les paramètres précédents, j'ai lancé des simulations à différentes forces et j'ai extrait les valeurs moyennes de l'extension de la molécule d'ADN  $l \langle \sum_i t_i^T \hat{z} \rangle$  dans le sens d'application de la force. Ces résultats sont illustrés sur la figure (3.5) et confrontés aux modèles théoriques et numériques du WLC [1, 2] ajusté sur les expériences qui ont bien entendu précédé ces simulations [8]. Ces expériences ont été faites dans une solution de 10 mmol et j'ai ajusté le rayon effectif de l'ADN à  $\tilde{r} = 4$  nm. Comme vous pouvez le constater, les expériences numériques avec des segments de longueurs 10 bp sont en très bon accord avec les données expérimentales. Dans le cas du FJC, les simulations ne reproduisent pas les expériences. Cette constatation est déjà présente dans [8]. Vous pouvez aussi constater que pour des forces inférieures à 0.1 pN des effets de tailles finies apparaissent. Ces effets peuvent être atténués en augmentant la taille de la molécule d'ADN et disparaissent si j'enlève le sol. Il me semble qu'expérimentalement des forces aussi faibles ne peuvent être étudiées car la bille magnétique se colle alors sur la plaque de verre.

Dans toutes ces simulations, les forces appliquées sont inférieures à 10 pN. Au delà, des expériences montrent que la longueur à vide de l'ADN augmente  $L \rightarrow L + \Delta L$ . Ces expériences sont référencées dans un article de J. Marko [57] qui propose un modèle théorique de ce phénomène. Je n'ai pas essayé de reproduire ce phénomène et il me semble qu'en utilisant des joints glissière nous pourrions donner une certaine élasticité aux cylindres.

### 3.3 ADN sous traction et torsion

Complexifions un tant soit peu la simulation en ajoutant une contrainte de torsion sur la molécule d'ADN. Cette contrainte se présente comme l'injection d'un certain nombre de tours  $n$  sur la bille magnétique. La bille magnétique étant alignée sur l'axe  $z$  du référentiel du laboratoire, le nombre de tours  $n$  correspond au nombre d'enlacements  $Lk$  [58, 59] de la molécule d'ADN plus un contour imaginaire fermé. Par exemple, pour de l'ADN de forme B libre de toute contrainte extérieure, le nombre d'enlacements correspond au nombre de fois que les deux brins de la double hélice s'enlacent. La période de la double hélice est environ  $H = 3.57$  nm soit un twist égal à 1. J'introduis également la quantité  $\omega_0 = 2\pi/H$ . Le nombre d'enlacements intrinsèque  $Lk_0$  d'une molécule de longueur  $L$  est alors égal à  $Lk_0 = L/H$ . Dans la suite nous allons nous intéresser au cas où l'ADN est contraint en torsion et étudier deux ensembles statistiques bien différents. Dans un premier temps nous imposerons un nombre de tour à la bille  $n$ . La taille de la bille étant plus grande que la longueur à vide de l'ADN le nombre de tour  $n$  correspond au nombre d'enlacement  $Lk$ . En effet, l'ADN ne pourra s'enrouler sur la bille et défaire les tours injectés. En pratique je choisis le rayon de la bille de telle sorte à rendre égaux taille de la molécule et “périmètre” de la sphère  $2\pi R = L$ .

Le modèle du Ball-Like-Chain ne contient pas de twist intrinsèque et j'introduis une quantité adimensionnée l'overtwist  $\sigma = Lk/Lk_0 = nH/L$ . Cette observable permet de comparer directement des expériences faites avec des molécules d'ADN de tailles différentes. Pour obtenir un  $n \neq 0$  nous



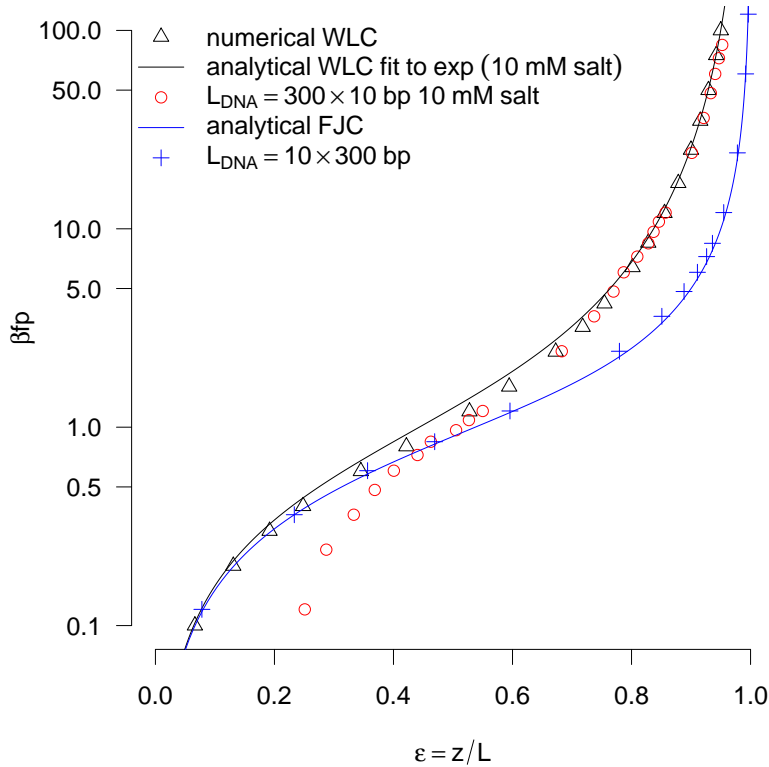


FIGURE 3.5: Le profil force extension est montré pour les modèles simulés du Ball-Like-Chain, du FJC, du WLC analytique et du WLC numérique [1, 2]. La quantité  $\beta f p$  représente la force adimensionnée tandis que  $\epsilon = z/L$  est l'extension relative de la molécule d'ADN. Les erreurs statistiques sont inférieures à la taille des symboles.

pouvons imposer une valeur à  $n$  comme dans les expériences de pinces magnétiques soit imposer un couple  $\Gamma$  sur la bille magnétique. Ce dernier cas n'a jamais été réalisé expérimentalement ni en simulation.

### 3.3.1 Nombre de tours de bille imposé

Dans cette partie, nous nous plaçons dans l'ensemble statistique à force de traction, nombre de tours et température fixés ( $f, n, T$ ). Pour un nombre de tours nul l'ADN se comporte comme un polymère semi-flexible, comportement qui dépend de la persistance en courbure et est décrit dans la section sur l'ADN en traction pure. Il advient tout autre chose pour un nombre de tours différent de 0. Pour le comprendre, vous pouvez prendre un câble de souris, le fixer à une extrémité, appliquer une traction et un mouvement de rotation (imposer un nombre de tours, un couple ou une vitesse de rotation constante) avec votre main à l'autre extrémité. Si votre main n'a pas suffisamment fait de tours, vous pourrez constater que le câble reste droit. Dès lors, ajouter, encore des tours au câble et vous le verrez former une structure appelée plectonème. Le couple à partir duquel la tige forme des plectonèmes s'appelle le couple de flambage  $\Gamma_p$  et est fonction [60] de la force  $2\sqrt{p}f$  avec laquelle vous tirez sur le câble. Pour imaginer ce qu'il se passe si vous remplacez la tige par de l'ADN vous devez imaginer la molécule placée dans un bain thermique, harassée et bousculée par d'incessantes

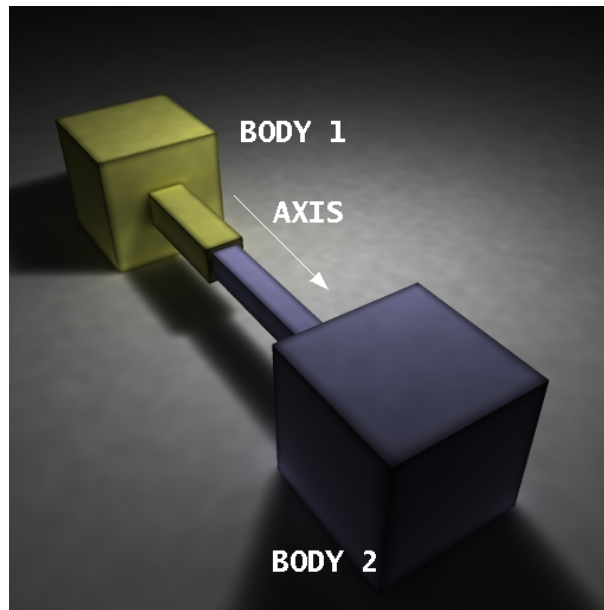


FIGURE 3.6: Vue schématique d'un joint glissière. ODE is Copyright 2001-2004 Russell L. Smith. All rights reserved.

collisions avec les molécules du bain. Dans ce cas l'ADN n'est pas droit comme la tige, comme montré dans la partie sur l'ADN nu en traction. Ajoutez maintenant un certain nombre de tours à la bille magnétique. Ce nombre de tour est stocké dans le twist tant que la résistance mécanique (persistance en twist) le permet. Lorsque la résistance au twist cède, l'ADN préfère former des plectonèmes et minimiser ainsi les contraintes auxquelles il est soumis.

Les simulations présentées dans cette partie sont réalisées en augmentant progressivement le nombre de tours de bille jusqu'à atteindre la valeur souhaitée. L'intervalle étudié est  $n \in [0; 25]$  pour une gamme de forces allant de 0.35 à 1.80 pN. Pour chaque jeu de paramètres la moyenne de l'extension de l'ADN est prise à l'équilibre statistique tous les 100000 pas de simulation (soit un temps entre deux mesures supérieur au temps de décorrélation de la distance bout à bout de l'ADN). La prise de données s'arrête lorsque la moyenne statistique n'évolue plus. Les résultats de simulations sont comparés à des expériences réalisées par Francesco Mosconi [3] pour un ADN dans une solution de 100 mmol de NaCl. Pour une telle condition, je propose un rayon effectif de l'ADN de  $\tilde{r} = 2$  nm. La comparaison est montrée sur la figure (3.7) et, comme vous pouvez le constater, l'accord entre expériences et simulations est très bon.

Pour apprécier pleinement le bon accord entre résultats de simulations et expérimentaux, je vais utiliser un joli résultat provenant d'un article de S. Neukirch [43, 61, 40, 62] qui nous donne l'extension relative de l'ADN en fonction du nombre de tours de bille dans la partie linéaire, dans la partie des plectonèmes (voir l'équation (3.4)). Cette équation introduit des quantités comme le rayon  $\rho$  des plectonèmes et l'angle hélical  $\theta$ . En utilisant cette formule, nous ajustons les quantités  $p$ ,  $t$ ,  $\rho$  et  $\theta$  sur les courbes expérimentales et simulées pour produire les trois courbes (3.8).

$$\frac{z_\sigma}{z_0} = 1 + \frac{p}{t} \left( \frac{1}{\cos 2\theta} - 1 \right) - \frac{4\pi\rho}{H \sin 2\theta} \sigma \quad (3.4)$$

Le rayon de supercoiling et l'angle hélical peuvent être mesurés à la main sur quelques configurations simulées; les valeurs obtenues sont les mêmes que les valeurs ajustées avec la formule (3.4).

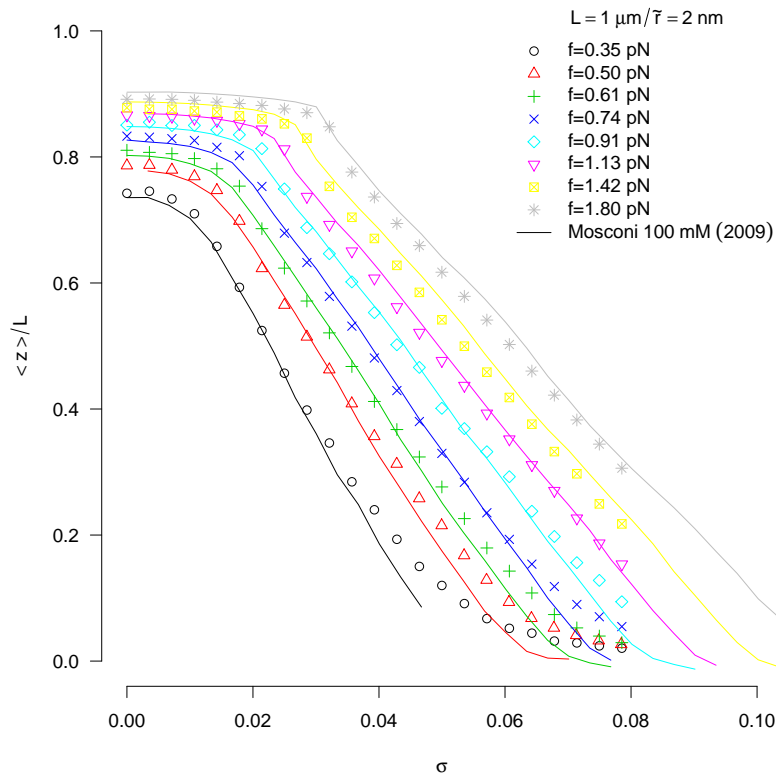
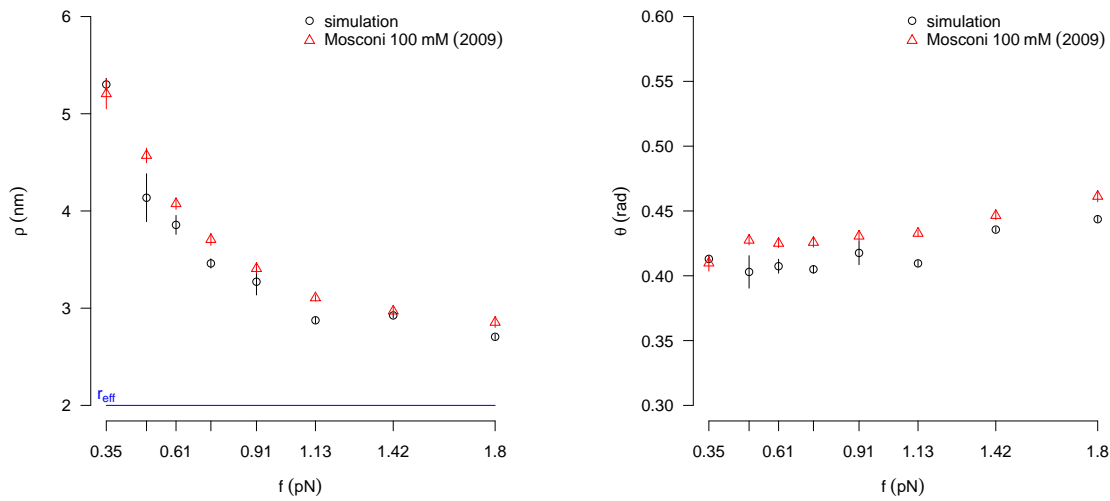


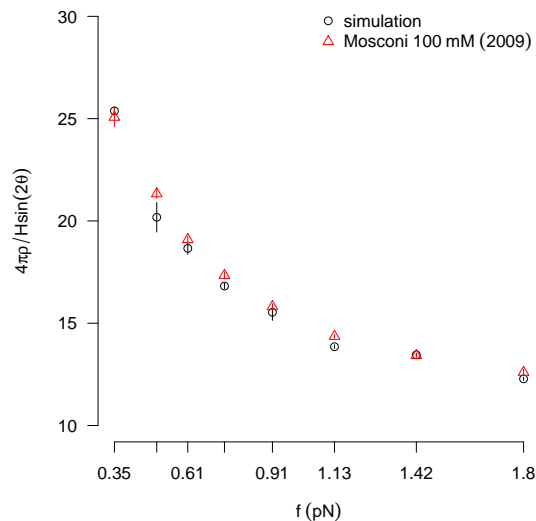
FIGURE 3.7: Le profil extension relative de l'ADN (en ordonnée) overtwitch (en abscisse) est montré pour différentes forces et une concentration molaire en sel NaCl de 100 mmol. Les résultats de simulations sont représentés par les symboles de différentes couleurs. Les résultats expérimentaux de F. Mosconi [3] sont montrés en lignes continues et de la même couleur que les symboles correspondants aux conditions expérimentales. Ce choix est malheureux mais permet de ne pas nuire à la clarté de la figure.

Une configuration (force de traction  $f = 0.74 \text{ pN}$  et nombre de tours  $n = 15$ ) avec deux plectonèmes sur le même ADN et obtenue par simulation est montrée sur la figure (3.9). La vidéo associée est disponible à l'adresse <http://vimeo.com/51918378>.

J'ai recommencé les mêmes simulations mais avec une concentration en sel NaCl plus importante, 320 mmol, une force de traction de  $f = 3 \text{ pN}$ , et un ADN de longueur  $L = 1900 \text{ bp}$ . Ces conditions sont empruntées à l'article de H. Brutzer [63]. Il montre une transition abrupte entre ADN étiré et ADN supercoilé qui disparaît pour de faibles forces et concentrations en NaCl. Cette transition est aussi montrée dans un article de S. Forth [64] et elle est parfaitement reproductible en simulation en ajustant le rayon effectif  $\tilde{r}$ . Pour cette concentration en sel, le rayon effectif est proche du rayon cristallographique. Pour former la première boucle du plectonème, il est nécessaire de rapprocher jusqu'au contact deux cylindres d'ADN. Evidemment, si le rayon  $\tilde{r}$  est petit, il faut consommer une plus grande longueur d'ADN avant le contact que dans le cas de gros cylindres.



(a) Rayon de supercoiling  $\rho$  ou rayon des plectonemes (b) Angle hélical  $\theta$  ou angle de croisement des plectonemes



(c) La pente de la partie linéaire (voir l'équation (3.4)) de la courbe extension relative - overtwist est montrée en ordonnée en fonction de la force de traction  $f$  (pN) exercée sur la bille en abscisse. La pente est obtenue par ajustement avec la formule (3.4) des données expérimentales et simulées.

FIGURE 3.8: Le rayon de supercoiling  $\rho$ , l'angle hélical  $\theta$  et la pente  $4\pi\rho/H \sin 2\theta$  sont montrés en ordonnées après avoir été ajustés sur les données expérimentales de F. Mosconi [3] et sur les données de simulations. Ces observables sont fonctions de la force appliquée  $f$  (pN) sur la bille magnétique.

### 3.3.2 Couple sur la bille imposé

Dans cette partie, nous nous plaçons dans l'ensemble statistique à force de traction, couple et température fixés ( $f, \Gamma, T$ ). A l'heure actuelle et surtout à ma connaissance, il n'existe pas

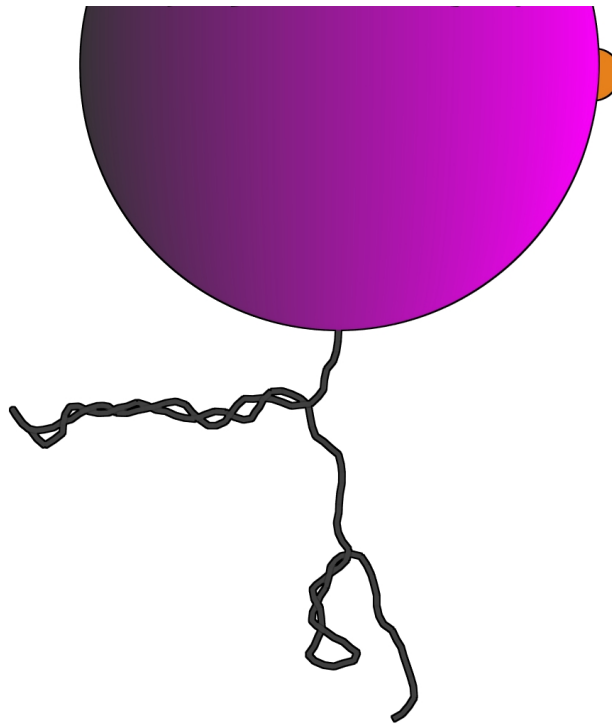


FIGURE 3.9: Vue schématique de deux plectonèmes obtenus avec un ADN de longueur  $L = 1 \mu\text{m}$  sous une force de traction de  $f = 0.74 \text{ pN}$  et un nombre de tours de bille de  $n = 15$ . Il est rare d'observer deux plectonèmes sur un même ADN car ils fusionnent très rapidement en un seul plectonème. Cette configuration n'est pas figée, l'agitation thermique permet de faire diffuser le(s) plectonème(s) le long de l'ADN. Une expérience récente interprète ses résultats en terme de multi-plectonèmes [4].

d'expériences ni de simulations où l'on impose un couple à la bille magnétique. Cependant, des articles comme [64, 63] montrent des expériences où la bille magnétique tourne continuellement avec une fréquence de rotation imposée. A partir des courbes extension relative  $\epsilon$  - overtwist  $\sigma$  obtenues à différentes forces et à l'aide d'un très joli article [65], les expérimentateurs [64, 63, 3, 66] sont capables de déduire les courbes couple  $\Gamma$  - overtwist  $\sigma$ . Malheureusement, ces courbes ne se prolongent pas au delà de la transtion ADN étiré - ADN surenroulé. En cherchant un peu, j'ai trouvé un impressionnant article de J. F. Marko [67] prédisant la relation entre couple et overtwist  $\Gamma(\sigma)$  et ce, quel que soit le régime dans lequel se trouve l'ADN. Dans cette article, l'ADN est "divisé" en deux états : un état étiré caractérisé par une énergie libre  $\mathcal{S}(\sigma)$  et un état supercoilé d'énergie libre  $\mathcal{P}(\sigma)$ . Ces énergies libres sont quadratiques en  $\sigma$  et permettent de montrer que dans la partie étirée le couple est  $\Gamma(\sigma) = a_s \sigma$  pour un overtwist  $\sigma < \sigma_s$ . Dans la région de coexistence entre partie de l'ADN étiré et plectonèmes,  $\sigma_s < \sigma < \sigma_p$ , le couple est constant et égal au couple de flambage  $\Gamma_p$ . La transition a lieu lorsque les contraintes en torsion à l'intérieur de l'ADN deviennent plus importantes que les fluctuations du twist. Au delà de la transition, le couple évolue linéairement avec l'overtwist  $\Gamma(\sigma) = t_\infty k_B T \omega_0(\sigma)$  où  $t_\infty$  représente la persistance en twist des plectonèmes. Je choisis le symbole  $\infty$  pour représenter un plectonème formé sur un ADN circulaire. Dans cette région, l'ADN est complètement plectonémisé comme le montre la figure (3.10). La vidéo associée est disponible à l'adresse <http://vimeo.com/51918151>.

Ma simulation me permettant d'imposer un couple à la bille magnétique, je me lance dans une série de mesures pour des forces variant de 0.35 à 1.80 pN et des couples variant de 0 à 25 pN. Je

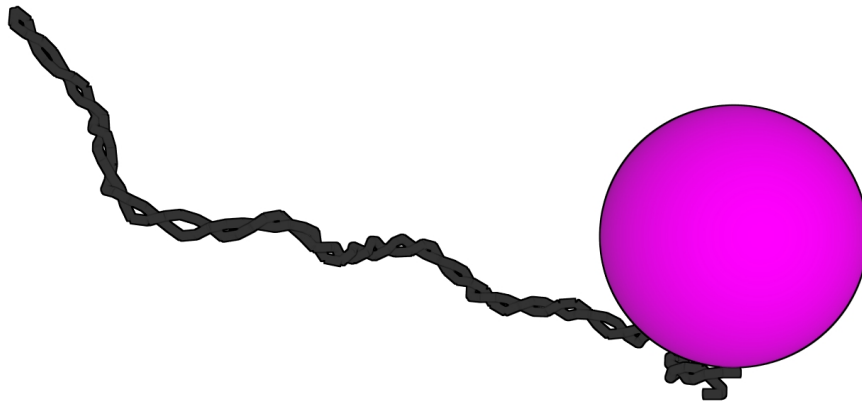


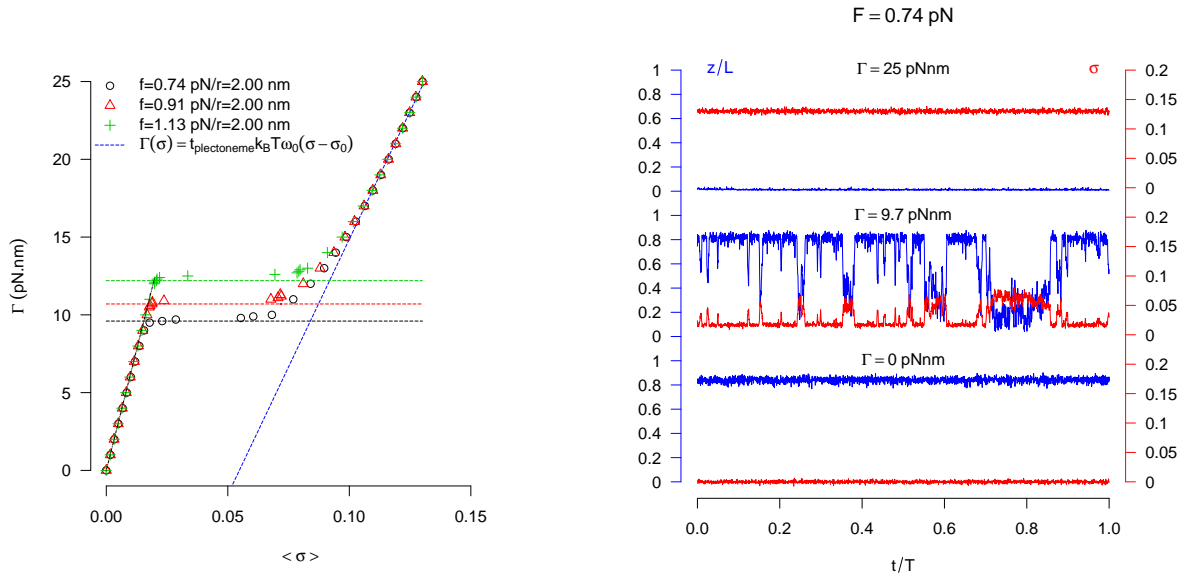
FIGURE 3.10: Vue schématique d'une configuration obtenue par simulation avec une force de traction de 0.74 pN et un couple de 20 pN · nm. L'ADN de longueur à vide  $L = 1 \mu\text{m}$  est totalement supercoilé.

place l'ADN dans un solvant de concentration en NaCl de 100 mmol. Je trace alors les courbes  $\Gamma(\sigma)$  simulées pour 0.74, 0.91, 1.13 pN et montrées sur la figure (3.11(a)). Je superpose à mes résultats numériques, les courbes  $\Gamma(\sigma)$  que F. Mosconi [3] calcule à partir de ces expériences et [65].

Dans la première partie de la courbe, en dessous de la transition de flambage, mes résultats sont en parfait accord avec ceux de F. Mosconi. Il est aussi possible de superposer la prédiction théorique  $\Gamma(\sigma) = a_s \sigma$  pour les persistances simulées et obtenir là encore un très bon accord. Les lignes en pointillés horizontales représentent les valeurs des couples  $\Gamma_p$  de flambage [3]. Ces valeurs augmentent avec la force appliquée sur la bille. Les simulations sont relativement fines puisqu'elles permettent d'observer la transition de flambage sur un intervalle de valeurs de couples différents suivant la force appliquée (voir les figures (3.12), (3.13) et (3.14)). Sur ces figures il est possible d'apprécier la transition de flambage en remarquant l'évolution conjointe de l'extension relative  $\epsilon$  et de l'overtwist  $\sigma$ . La formation de plectonèmes s'accompagne d'une augmentation de  $\sigma$  mais aussi d'une diminution de la taille de la molécule d'ADN.

Enfin, dans la partie au delà de la transition de flambage, vous pouvez remarquer que les courbes  $\Gamma(\sigma)$  à différentes forces se superposent. Il est même possible d'ajuster une courbe de la forme  $t_\infty k_B T \omega_0 (\sigma - \sigma_0)$  sur cette dernière partie et obtenir un  $\sigma_0 = 0.055 \pm 0.001$  et  $t_\infty = 45.1 \pm 0.8$  nm. Cette prédiction numérique est en désaccord pour deux raisons avec la prédiction théorique de J. F. Marko. La première raison est la présence du terme  $-\sigma_0$  dans l'ajustement de mes données simulées. La seconde raison est qu'à partir des données expérimentales de V. Rybenkov [68, 69], J. F. Marko prévoit une persistance en twist des plectonèmes de  $t_\infty = 24$  nm contre  $t_\infty = 45$  nm pour mes simulations.

Vous pouvez apprécier la transition de flambage à différentes forces et couples sur les figures



(a) Le couple appliqué  $\Gamma$  (pN·nm) sur la bille magnétique (b) La trace de l'extension relative de la molécule d'ADN est montré en fonction de l'overwrist  $\sigma$  et ce pour trois en fonction du temps est montrée en bleu tandis que la forces différentes 0.74, 0.91, 1.13 pN respectivement en trace de l'overwrist est montrée en rouge. Les traces sont noir, rouge et vert. Les couples de flambage  $\Gamma_p$  calculés représentées pour une force de traction de  $f = 0.74$  pN et par [3] sont représentés par les lignes horizontales en pointillés. L'ajustement de la partie où l'ADN par la formule  $t_{\infty} k_B T \omega_0 (\sigma - \sigma_0)$  est totalement supercoilé est montré par une ligne bleue en pointillé.

FIGURE 3.11: Les courbes  $\Gamma(\sigma)$  sont représentées sur la figure de gauche. La figure de droite représente l'évolution de la taille de la molécule et de l'overwrist en fonction du temps pour une force de traction de  $f = 0.74$  pN et des couples de  $\Gamma = 0, 9.7, 25$  pN·nm. Il est intéressant de faire le parallèle entre ces deux figures en remarquant que les traces se trouvent en face de la partie étirée de l'ADN, de la transition vers les plectonèmes et de la partie où toute la molécule d'ADN est supercoilée.

(3.12), (3.13) et (3.14). Des articles comme [64, 63] montrent que la formation de la première boucle de plectonèmes diminue l'extension de la molécule d'ADN d'une valeur bien déterminée et dépendant de la force et de la concentration en sel. Par exemple, pour une molécule de 2200 bp soumise à une force de traction de 2 pN et une fréquence de rotation de 0.04 tour/s, S. Forth [64] montre que l'ADN fluctue entre deux états d'extension séparés par 79 nm. Comme vous avez pu le constater sur les figures (3.12), (3.13) et (3.14), il existe dans mes simulations plusieurs états bien différents séparés par des extensions variables. Ceci veut dire que, non pas une, mais un certain nombre de boucles de plectonèmes se forment dans la région de coexistence. Dans l'intervalle  $\sigma_s < \sigma < \sigma_p$  le couple  $\Gamma(\sigma)$  est constant et ne sert plus à créer du twist dans la molécule d'ADN mais à former des plectonèmes. C'est l'analogie de la transition liquide-vapeur ; lorsque nous faisons chauffer de l'eau la chaleur fournie au liquide sert dans un premier temps à augmenter sa température puis à transformer le liquide en vapeur à température constante. Il est bien entendu évident que nous ne pouvons faire de comparaisons entre ces expériences à fréquence de rotation imposée et mes simulations à couple imposé ; juste en noter les différences. Vous pourriez alors me demander pourquoi n'ai-je pas tenté de simulations à fréquence de rotation imposée et donc essayé de reproduire ces expériences. Chez [64]

et dans les conditions données précédemment, nous pouvons observer une transition de flambage au voisinage de 8 tours et des fluctuations entre les deux états d'extension sur des temps de l'ordre de la seconde. Sachant que le pas de temps typique de mes simulations et de l'ordre de la pico-seconde, il me faudrait beaucoup plus d'un mois pour produire une trace de l'extension de la molécule et peut-être observer la transition. J'attends avec impatience la réalisation d'une expérience où nous pourrions imposer un couple à la bille magnétique et comparer les résultats aux prédictions que j'ai faites.

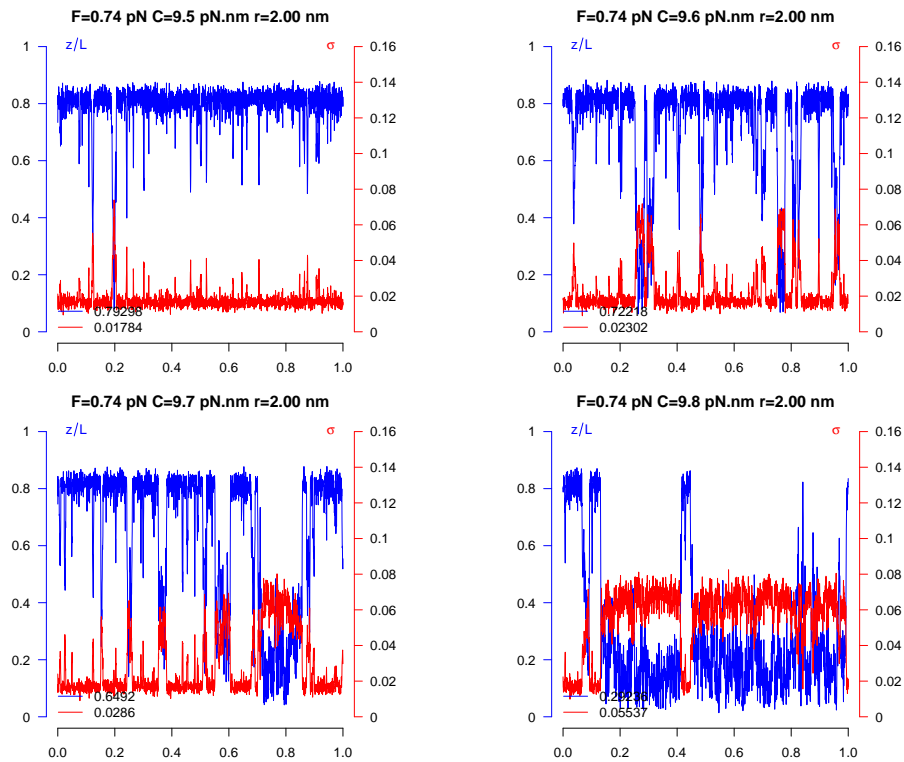


FIGURE 3.12: L'instabilité de flambage est montrée pour une force de traction de  $f = 0.74 \text{ pN}$  et quatre différents couples de  $9.5$  à  $9.8 \text{ pN}\cdot\text{nm}$  par  $0.1 \text{ pN}\cdot\text{nm}$ . La trace de l'extension relative  $\epsilon$  est montrée en bleu et, en rouge, est montrée la trace de l'overtwist  $\sigma$ .



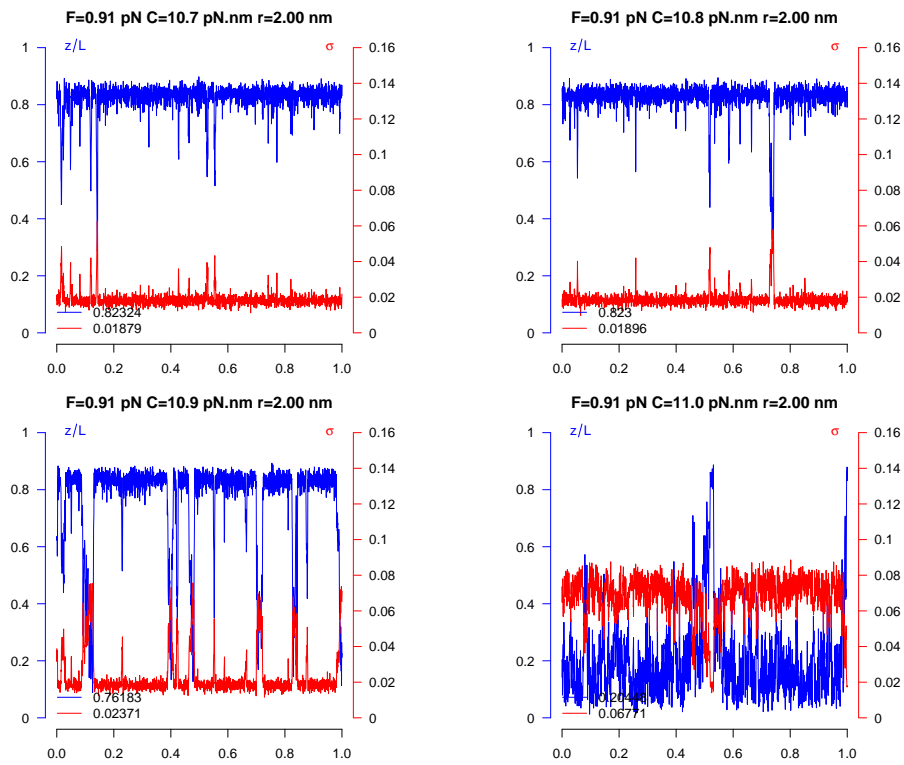


FIGURE 3.13: L'instabilité de flambage est montrée pour une force de traction de  $f = 0.91$  pN et quatre différents couples de 10.7 à 11 pN · nm par 0.1 pN · nm. La trace de l'extension relative  $\epsilon$  est montrée en bleu et, en rouge, est montrée la trace de l'overtwist  $\sigma$ .

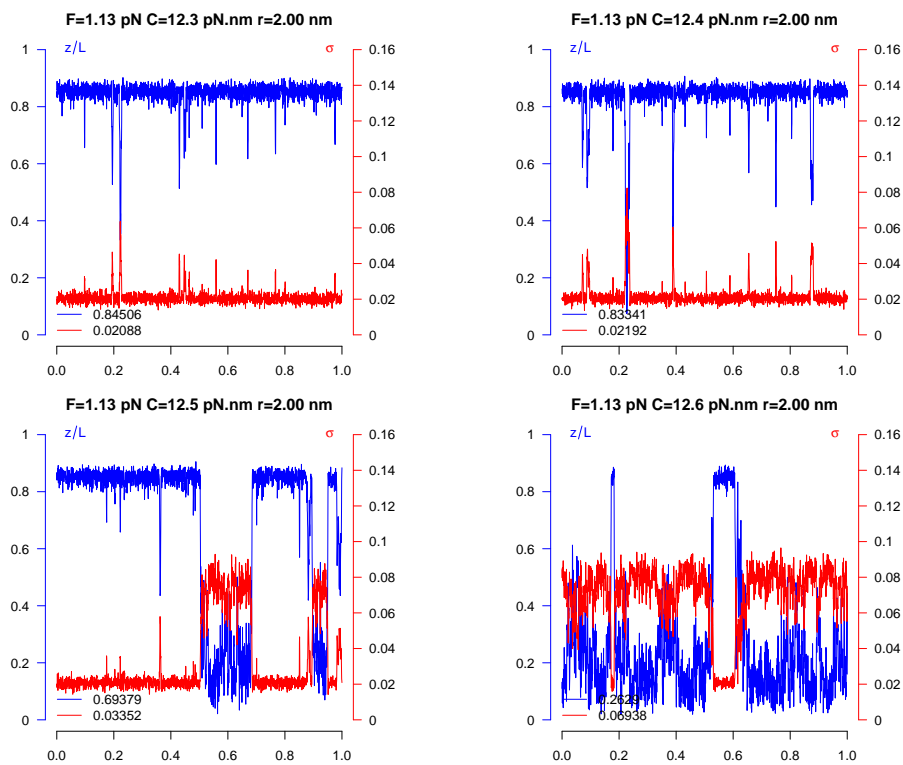


FIGURE 3.14: L’instabilité de flambage est montrée pour une force de traction de  $f = 1.13$  pN et six différents couples de 12.3 à 12.6 pN · nm par 0.1 pN · nm. La trace de l’extension relative  $\epsilon$  est montrée en bleu et, en rouge, est montrée la trace de l’overtwist  $\sigma$ .



---

# Assemblages de nucléosomes

---

A présent que nous sommes convaincus par la méthode de simulation, je propose de montrer quelques résultats sur les pinces magnétiques et l'assemblage d'ADN et protéines. A cette fin, je commence par illustrer le passage à l'échelle supérieure par l'argument de compaction montré sur la figure (4.1). L'information génétique est contenue dans une très longue molécule d'ADN, de l'ordre de  $\simeq 1$  m si nous la déroulons, qui est elle même contenue dans un noyau d'une taille de l'ordre de  $\simeq 1$   $\mu\text{m}$ . En imaginant l'ADN totalement libre de contraintes externes, il est difficile de penser cette molécule confinée dans un si petit espace. Néanmoins, il est possible d'exercer des contraintes de torsion pour créer des plectonèmes et réduire la "taille" de la molécule. Ce mécanisme n'est pas suffisant ; en utilisant des protéines venant se lier à l'ADN, il est possible d'obtenir une plus grande compaction. Ces deux mécanismes sont résumés dans la figure (4.1).

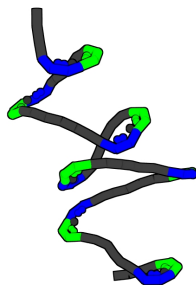
Le modèle "coarse-grained" du nucléosome, que j'utiliserai dans la suite, est construit à partir de la structure cristallographique [5] avec une résolution de 2.8 Å (voir la figure (4.2)) et les différentes étapes de construction sont montrées sur la figure (4.5). Cette structure est disponible dans la "Protein-Data-Bank" sous l'appellation 1KX5.pdb. Le nucléosome compacte l'ADN en enroulant 147 paires de bases autour d'un cœur de protéines. L'enroulement est à gauche et possède un pas de 2.8 nm. Le cœur de protéines est composé de quatre dimères : deux dimères formés des histones H2A et H2B et deux dimères formés des histones H3 et H4. Il possède une symétrie autour d'un axe appelé "axe dyade". Dans un premier temps, le tétramère se lie à l'ADN pour former le tétrasome et c'est ensuite que les dimères H2A et H2B viennent se lier à l'ADN et au tétramère. Comme vous pouvez le constater sur la figure (4.5), le tétramère est bien accompagné au centre du nucléosome par les deux dimères H2A-H2B. Les liaisons qui permettent de stabiliser les dimères sont de nature hydrophobe tandis que les interactions entre deux dimères sont de nature électrostatique. En particulier, l'histone H2B interagit avec l'histone H4 sous la forme d'un FHB tandis que le DD de l'histone H2A interagit avec le tétramère. Les interactions entre les histones et l'ADN se font en quatorze points de contacts. Ces points se trouvent au niveau du petit sillon, à la surface intérieure de la super-hélice d'ADN. Ces "Super-Helix-Locations" (SHLs) correspondent à un encastrement de plusieurs arginines (interaction



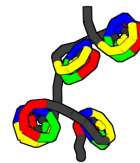
(a) Configuration d'un ADN dans un bain thermique



(b) Configuration d'un ADN supercoilé dans un bain thermique



(c) 7 tétrasomes



(d) 4 nucléosomes

FIGURE 4.1: Vue schématique des différents mécanismes de compaction pour un ADN de longueur 900 bp : le "supercoiling" et "l'architecturing" par des protéines. (Bleu et vert) cylindres représentant les dimères H3-H4, respectivement, HMfA et HMfB pour les protéines des Archaees. (Rouge et jaune) cylindres représentant les dimères H2A-H2B. L'ADN lié à un dimère est représenté de la même couleur. Pour apprécier pleinement ces deux mécanismes de compaction, les quatre images sont représentées à la même échelle.



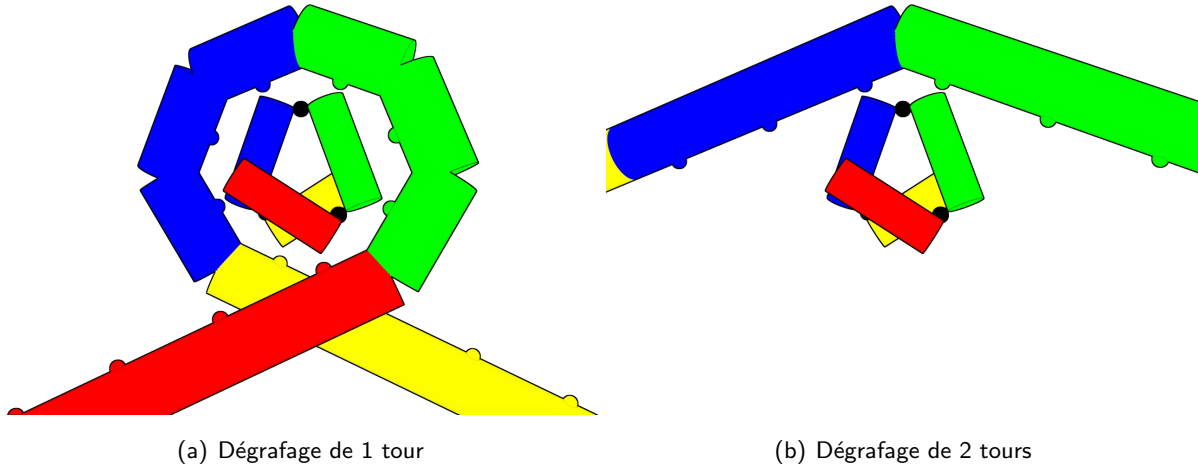


FIGURE 4.3: Vue schématique du modèle de dégrafage symétrique du nucléosome. A gauche, l'ADN est déroulé de un tour de façon symétrique. A droite, l'ADN est déroulé de deux tours de façon symétrique.

tour et non 1/2 ou 1/4 de tour ?

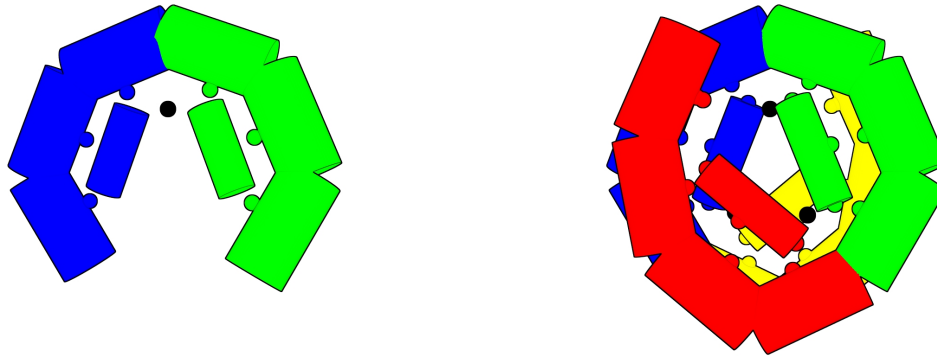
## 4.1 Présentation du modèle mécanique du nucléosome

H. Wong utilise dans sa thèse un modèle "coarse-grained" de nucléosome construit à partir d'une analyse de modes normaux [71]. Ce modèle inclut la donnée des positions  $\mathbf{r}$  des 14 cylindres d'ADN, de leur orientation respective  $\mathbf{t}$ ,  $\mathbf{m}$  et  $\tilde{\mathbf{m}}$ . De ces orientations, je déduis les angles de courbure  $\theta$  et de twist  $\phi$ . L'axe dyade est choisi comme étant l'axe  $\mathbf{y}$  du référentiel du laboratoire et l'axe  $\mathbf{z}$  comme l'axe d'enroulement de l'ADN nucléosomal. Les symétries associées sont résumées dans le tableau (4.4).

$\mathbf{t}$	$\mathbf{m}$	$\tilde{\mathbf{m}}$	$\Gamma^{b+t}$	$\theta, \phi$
$t_i^x = t_{14-i+1}^x$	$u_i^x = u_{14-i+1}^x$	$v_i^x = -v_{14-i+1}^x$	$\Gamma_{i,i+1}^{b+t;x} = \Gamma_{14-i,14-i+1}^{b+t;x}$	$\theta_i = \theta_{14-i+1}$
$t_i^y = -t_{14-i+1}^y$	$u_i^y = -u_{14-i+1}^y$	$v_i^y = v_{14-i+1}^y$	$\Gamma_{i,i+1}^{b+t;y} = -\Gamma_{14-i,14-i+1}^{b+t;y}$	$\phi_i = \phi_{14-i+1}$
$t_i^z = t_{14-i+1}^z$	$u_i^z = u_{14-i+1}^z$	$v_i^z = -v_{14-i+1}^z$	$\Gamma_{i,i+1}^{b+t;z} = \Gamma_{14-i,14-i+1}^{b+t;z}$	$\dots$

FIGURE 4.4: Tableau récapitulatif des symétries.

Cependant, j'ai modifié la forme des histones pour une meilleure superposition avec la structure cristallographique et ajouter la description BLC de l'ADN. H. Wong utilise des joints charnières et des joints de Cardan pour reproduire les interactions dues aux SHL et DD. Ces joints sont parfaitement adaptés pour reproduire les directions du mouvement données par l'analyse de modes normaux. Malheureusement, ces liaisons rigides ne permettent pas de reproduire d'éventuelles ruptures. Dans le même esprit que les modélisations précédentes, je propose de simuler les segments d'ADN de la super-hélice par des cylindres de taille correspondant à 10.5 paires de bases soit une longueur  $l_n = 3.57$  nm et un rayon de  $r_n = 1$  nm. La courbure et le twist sont toujours modélisés par les couples de rappel introduits précédemment et les joints connectant les cylindres d'ADN sont des joints



(a) Tétrosome

(b) Nucléosome

FIGURE 4.5: Vue schématique du modèle de nucléosome. L'ADN nucléosomal est enroulé à gauche. (Rouge et jaune) dimères H2A-H2B. (Bleu et vert) dimères H3-H4. (Noir) Four-Helix-Bundle. L'ADN est représenté avec la couleur du dimère avec lequel il interagit. La position du petit sillon est montrée par les petites billes de couleur.

“Ball-In-Socket”. Le résultat pour le nucléosome avec son intermédiaire le tétrosome sont montrés sur la figure (4.5). Je choisis de modéliser les interactions au sein du nucléosome par des potentiels de Morse avec  $D \geq 0$  et  $a > 0$  où  $\mathbf{r}_{a1}$ , respectivement  $\mathbf{r}_{a2}$ , est le point d'application du potentiel de Morse sur le corps 1, respectivement sur le corps 2 :

$$U(r = \|\mathbf{r}_{a1} - \mathbf{r}_{a2}\|) = D \left[ \left( 1 - e^{-a(r-r_e)} \right)^2 - 1 \right] \quad (4.1)$$

Les points d'applications ne sont pas nécessairement confondus avec le centre de masse du corps rigide. Il en résulte une force  $\mathbf{F}_i$  et un moment de force  $(\mathbf{r}_{ai} - \mathbf{r}_i) \otimes \mathbf{F}_i$  :

$$\mathbf{F}_1(r) = -2aD \left( 1 - e^{-a(r-r_e)} \right) e^{-a(r-r_e)} \frac{\mathbf{r}_1 - \mathbf{r}_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \quad (4.2)$$

$$\mathbf{F}_2(r) = -2aD \left( 1 - e^{-a(r-r_e)} \right) e^{-a(r-r_e)} \frac{\mathbf{r}_2 - \mathbf{r}_1}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \quad (4.3)$$

La force est nulle pour  $r = r_e$  et le point d'inflexion est  $ar^* = ar_e + \ln 2$ . Je choisis ce point, qui correspond au changement de signe de la courbure du potentiel de Morse, comme étant égal à  $r^* = 2r_e$ . Ce choix arbitraire nous donne une portée égale à la distance d'équilibre  $r_e$  et le paramètre  $a = \ln 2 / r_e$ . L'orientation des cylindres d'ADN est choisie de telle sorte que l'axe  $-\tilde{\mathbf{m}}$  pointe vers le petit sillon. Le point d'interaction des “Super-Helix-Location” est justement choisi pour être à la surface des cylindres (à la surface de l'ADN) soit à la position  $\mathbf{r} - r_n \tilde{\mathbf{m}}$ . Ce choix n'est pas arbitraire car les interactions entre les arginines des dimères et l'ADN ont pour lieux le petit sillon. Le potentiel de Morse est un choix intéressant car il permet de modéliser la rupture d'une liaison. Cependant, les énergies de liaisons sont inconnues et pour les déterminer je propose d'appliquer un calcul de structure sur le modèle mécanique du nucléosome.



## 4.2 Calcul de structure sur le nucléosome

Pour calculer les pré-contraintes qui permettent au nucléosome d'être stable, j'utilise l'équation de l'équilibre statique suivante  $\mathcal{J}^T \boldsymbol{\lambda} + \mathcal{F}_{b+t} + \mathcal{F}_m = \mathbf{0}$ , qui nous dit que la somme des torseurs dynamiques sur la super-hélice d'ADN, réaction des joints  $\mathcal{J}^T \boldsymbol{\lambda}$ , couples de courbure et de twist  $\mathcal{F}_{b+t}$  et  $\mathcal{F}_m$  interactions dues aux SHLs, FHBs et DDs est nulle. La tension  $\mathbf{T} = T\mathbf{t}$  le long des cylindres d'ADN est contenue dans la réaction des joints, et elle est égale au produit scalaire des tangentes aux cylindres par le torseur dynamique des contraintes  $\mathbf{t}^T \mathcal{J}^T \boldsymbol{\lambda}$ .

### 4.2.1 Première étape : isolement de l'ADN nucléosomal

Les interactions et leur nomenclature sont présentées dans le tableau (4.6). Nous partons de la configuration donnée précédemment.

SHL "-X"	ADN	Dimère	torseur	SHL "+X"	ADN	Dimère	torseur
-6.5	rouge 1	H3-H4 bleu	$\mathbf{f}_1, \mathbf{X}_1$	+0.5	bleu 8	H3-H4 bleu	$\mathbf{f}_8, \mathbf{X}_8$
-5.5	rouge 2	H2A-H2B rouge	$\mathbf{f}_2, \mathbf{X}_2$	+1.5	bleu 9	H3-H4 bleu	$\mathbf{f}_9, \mathbf{X}_9$
-4.5	rouge 3	H2A-H2B rouge	$\mathbf{f}_3, \mathbf{X}_3$	+2.5	bleu 10	H3-H4 bleu	$\mathbf{f}_{10}, \mathbf{X}_{10}$
-3.5	rouge 4	H2A-H2B rouge	$\mathbf{f}_4, \mathbf{X}_4$	+3.5	jaune 11	H2A-H2B jaune	$\mathbf{f}_{11}, \mathbf{X}_{11}$
-2.5	vert 5	H3-H4 vert	$\mathbf{f}_5, \mathbf{X}_5$	+4.5	jaune 12	H2A-H2B jaune	$\mathbf{f}_{12}, \mathbf{X}_{12}$
-1.5	vert 6	H3-H4 vert	$\mathbf{f}_6, \mathbf{X}_6$	+5.5	jaune 13	H2A-H2B jaune	$\mathbf{f}_{13}, \mathbf{X}_{13}$
-0.5	vert 7	H3-H4 vert	$\mathbf{f}_7, \mathbf{X}_7$	+6.5	jaune 14	H3-H4 vert	$\mathbf{f}_{14}, \mathbf{X}_{14}$

FIGURE 4.6: Tableau récapitulatif des interactions "Super-Helix-Location" entre les dimères et l'ADN nucléosomal.

Pour le premier cylindre d'ADN, on doit prendre en compte l'interaction liée au "Super-Helix-Location -6.5" soit une force  $\mathbf{f}_1$  et un couple  $\mathbf{X}_1$  au centre de masse, les contraintes mécaniques liées au joint "Ball-in-Socket" entre le cylindre 1 et le cylindre 2 soit une force  $\boldsymbol{\lambda}_{1,2}$  et un couple  $\frac{l}{2}\mathbf{t}_1 \otimes \boldsymbol{\lambda}_{1,2}$  au centre de masse et enfin les couples de courbure et de twist  $\boldsymbol{\Gamma}_{1,2}^{b+t}$  :

$$\begin{aligned} \boldsymbol{\lambda}_{1,2} + \mathbf{f}_1 &= \mathbf{0} \\ \frac{l_n}{2}\mathbf{t}_1 \otimes \boldsymbol{\lambda}_{1,2} + \mathbf{X}_1 + \boldsymbol{\Gamma}_{1,2}^{b+t} &= \mathbf{0} \end{aligned} \quad (4.4)$$

Du second au treizième cylindre d'ADN  $1 < i < 14$  nous devons prendre en compte l'interaction liée au "-6.5 < Super-Helix-Location < +6.5", les contraintes mécaniques liées au joint "Ball-in-Socket" entre le cylindre  $i-1$  et le cylindre  $i$ , les contraintes mécaniques liées au joint "Ball-in-Socket" entre le cylindre  $i$  et le cylindre  $i+1$  et enfin les couples de courbure et de twist  $\boldsymbol{\Gamma}_{i-1,i}^{b+t}$  et  $\boldsymbol{\Gamma}_{i,i+1}^{b+t}$  :

$$\begin{aligned} \boldsymbol{\lambda}_{i,i+1} - \boldsymbol{\lambda}_{i-1,i} + \mathbf{f}_i &= \mathbf{0} \\ \frac{l_n}{2}(-\mathbf{t}_i) \otimes (-\boldsymbol{\lambda}_{i-1,i}) + \frac{l_n}{2}\mathbf{t}_i \otimes \boldsymbol{\lambda}_{i,i+1} + \mathbf{X}_i + \boldsymbol{\Gamma}_{i,i+1}^{b+t} - \boldsymbol{\Gamma}_{i-1,i}^{b+t} &= \mathbf{0} \end{aligned} \quad (4.5)$$

Enfin, nous terminons le bilan des efforts sur l'ADN nucléosomal par le quatorzième cylindre :

$$\begin{aligned}
 -\lambda_{13,14} + \mathbf{f}_{14} &= \mathbf{0} \\
 \frac{l_n}{2} (-\mathbf{t}_{14}) \otimes (-\lambda_{13,14}) + \mathbf{X}_{14} - \mathbf{\Gamma}_{13,14}^{b+t} &= \mathbf{0}
 \end{aligned} \tag{4.6}$$

Nous pouvons remarquer que la somme des bilans des forces sur tous les segments d'ADN nous donne l'égalité suivante  $\sum_{i=1}^{14} \mathbf{f}_i = \mathbf{0}$  et que  $\lambda_{i,i+1} = -\sum_{j=1}^i \mathbf{f}_j$ . Les forces  $\mathbf{f}_i$  possèdent les symétries  $f_i^x = -f_{14-i+1}^x$ ,  $f_i^y = f_{14-i+1}^y$  et  $f_i^z = -f_{14-i+1}^z$  obtenues par rotation autour de l'axe dyade. Même avec cette condition, le système (4.7) nous donne une infinité de solutions pour les  $\mathbf{f}_i$ .

$$\begin{aligned}
 -\frac{l_n}{2} \mathbf{t}_1 \otimes \mathbf{f}_1 + \mathbf{X}_1 + \mathbf{\Gamma}_{1,2}^{b+t} &= \mathbf{0} \\
 -l_n \mathbf{t}_i \otimes \sum_{j=1}^{i-1} \mathbf{f}_j - \frac{l_n}{2} \mathbf{t}_i \otimes \mathbf{f}_i + \mathbf{X}_i + \mathbf{\Gamma}_{i,i+1}^{b+t} - \mathbf{\Gamma}_{i-1,i}^{b+t} &= \mathbf{0} \\
 \frac{l_n}{2} \mathbf{t}_{14} \otimes \mathbf{f}_{14} + \mathbf{X}_{14} - \mathbf{\Gamma}_{13,14}^{b+t} &= \mathbf{0}
 \end{aligned} \tag{4.7}$$

Je remarque l'équation fort utile  $\mathbf{t}_i^T \mathbf{X}_i = -\mathbf{t}_i^T (\mathbf{\Gamma}_{i,i+1}^{b+t} - \mathbf{\Gamma}_{i-1,i}^{b+t}) = -\mathbf{t}_i^T (\mathbf{\Gamma}_{i,i+1}^t - \mathbf{\Gamma}_{i-1,i}^t)$ . Un modèle "coarse-grained" de nucléosome pour lequel les orientations des cylindres d'ADN sont obtenues par transport parallèle ne contient pas de twist et les couples exercés par les SHLs sont nuls. Avant de poursuivre mes arguments de simplification du problème et trouver les  $\alpha_i$  je rappelle que le nucléosome est formé d'un premier bloc, le tétrasome, qui recrute ensuite deux autres blocs, les deux dimères H2A-H2B, pour former le nucléosome. Au sein d'un bloc l'ADN souffre une tension uniforme ; il n'existe pas de discontinuité de tension  $\Delta \mathbf{T}$  au niveau des SHLs  $\pm 0.5$ ,  $\pm 1.5$ ,  $\pm 4.5$ ,  $\pm 5.5$  et  $\pm 6.5$ . Je choisis une tension nulle  $\mathbf{T}^- = \mathbf{T}^+ = \mathbf{0}$  de part en part de ces SHLs. Il n'en est pas de même pour la partie de l'ADN nucléosomal entre deux blocs. En particulier, il existe une discontinuité  $\|\Delta \mathbf{T}\| = T$  au niveau des SHLs  $\pm 3.5$  et  $\pm 2.5$  illustrée sur la figure (4.7). Je rappelle que la tension le long d'un cylindre d'ADN est induite par le ou les joints attachés, ce qui conduit aux expressions suivantes  $\mathbf{t}_i^T \lambda_{i,i+1}$  et  $\mathbf{t}_i^T \lambda_{i-1,i}$ . L'ADN nucléosomal n'est soumis à aucune tension extérieure aux extrémités. En utilisant la figure (4.7) :

$$\mathbf{t}_4^T \lambda_{4,5} = T_4 \tag{4.8}$$

$$\mathbf{t}_5^T \lambda_{4,5} = T_5 \tag{4.9}$$

$$\mathbf{t}_i^T \lambda_{i,i+1} = 0 \tag{4.10}$$

puis l'équilibre des forces sur un cylindre d'ADN :

$$\mathbf{t}_4^T (\lambda_{4,5} - \lambda_{3,4} + \mathbf{f}_4) = 0 \tag{4.11}$$

$$\mathbf{t}_5^T (\lambda_{5,6} - \lambda_{4,5} + \mathbf{f}_5) = 0 \tag{4.12}$$

$$\mathbf{t}_i^T (\lambda_{i,i+1} - \lambda_{i-1,i} + \mathbf{f}_i) = 0 \tag{4.13}$$

nous avons les conditions supplémentaires sur les  $\mathbf{f}_i$  :

$$T_4 + t_4^T f_4 = 0 \quad (4.14)$$

$$-T_5 + t_5^T f_5 = 0 \quad (4.15)$$

$$t_i^T f_i = 0 \quad (4.16)$$

Malheureusement, cela ne suffit pas à trouver les interactions entre l'ADN et les dimères.

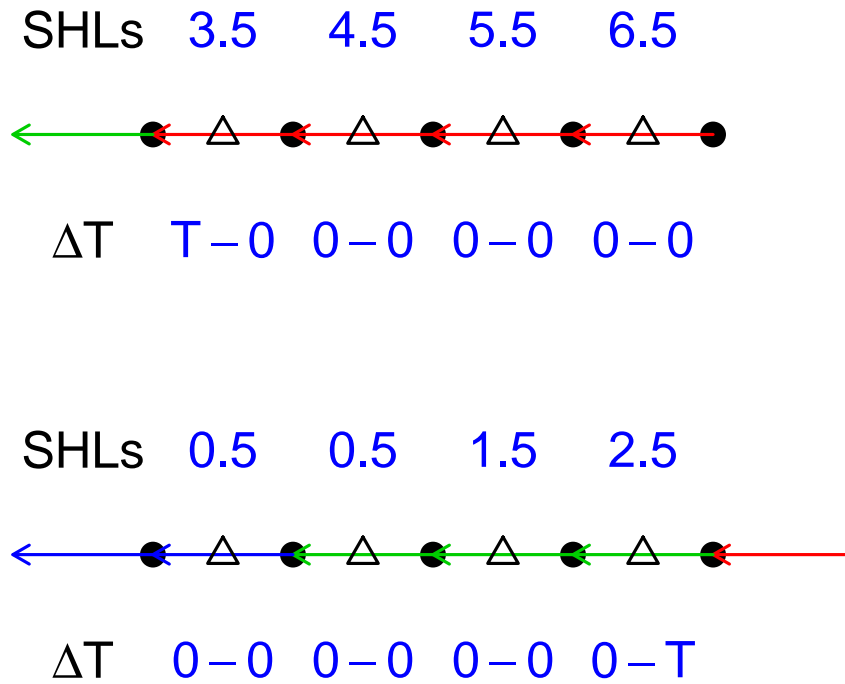


FIGURE 4.7: Les cylindres d'ADN sont représentés par des lignes rouges et vertes (le code couleur reste inchangé). Les flèches indiquent les tangentes des cylindres. Ils sont reliés par des joints mécaniques "Ball-in-Socket" qui sont schématisés par des cercles noirs. Les SHLs sont représentés par des triangles noirs. La tension de part et d'autre d'un SHLs est indiquée en bleu.

#### 4.2.2 Seconde étape : isolement de l'octamère

J'établis aussi le bilan des forces et des couples au centre de masse des cylindres représentant les dimères. Pour cela, je me sers du tableau récapitulatif (4.8).

Pour le dimère "H2A-H2B" rouge, on a :

$$\begin{aligned} f_{rb} + f_{rv} - f_2 - f_3 - f_4 &= 0 \\ \mathbf{a}_{rb} \otimes f_{rb} + \frac{l_d}{2} \mathbf{t}_r \otimes f_{rv} - \mathbf{a}_2 \otimes f_2 - \mathbf{a}_3 \otimes f_3 - \mathbf{a}_4 \otimes f_4 &= 0 \end{aligned} \quad (4.17)$$

Pour le dimère vert "H3-H4", on a :

Interaction	Dimère	Dimère	torseur
Docking-Domain	H2A-H2B rouge	H3-H4 bleu	$\mathbf{f}_{rb}, \mathbf{X}_{rb}$
Four-Helix-Bundle	H2A-H2B rouge	H3-H4 vert	$\mathbf{f}_{rv}, \mathbf{X}_{rv}$
Four-Helix-Bundle	H3-H4 vert	H3-H4 bleu	$\mathbf{f}_{vb}, \mathbf{X}_{vb}$
Four-Helix-Bundle	H2A-H2B jaune	H3-H4 bleu	$\mathbf{f}_{jb}, \mathbf{X}_{jb}$
Docking-Domain	H2A-H2B jaune	H3-H4 vert	$\mathbf{f}_{jv}, \mathbf{X}_{jv}$

FIGURE 4.8: Tableau récapitulatif des interactions FHB et DD entre dimères.

$$\mathbf{f}_{vb} - \mathbf{f}_{rv} - \mathbf{f}_{jv} - \mathbf{f}_5 - \mathbf{f}_6 - \mathbf{f}_7 - \mathbf{f}_{14} = \mathbf{0} \quad (4.18)$$

Pour le dimère bleu "H3-H4", on a :

$$-\mathbf{f}_{rb} - \mathbf{f}_{vb} - \mathbf{f}_{jb} - \mathbf{f}_8 - \mathbf{f}_9 - \mathbf{f}_{10} - \mathbf{f}_1 = \mathbf{0} \quad (4.19)$$

Pour le dimère jaune "H2A-H2B", on a :

$$\begin{aligned} \mathbf{f}_{jb} + \mathbf{f}_{jv} - \mathbf{f}_{11} - \mathbf{f}_{12} - \mathbf{f}_{13} &= \mathbf{0} \\ \mathbf{a}_{rb} \otimes \mathbf{f}_{rb} + \frac{l_d}{2} \mathbf{t}_r \otimes \mathbf{f}_{rv} - \mathbf{a}_2 \otimes \mathbf{f}_2 - \mathbf{a}_3 \otimes \mathbf{f}_3 - \mathbf{a}_4 \otimes \mathbf{f}_4 &= \mathbf{0} \end{aligned} \quad (4.20)$$

Le problème reste encore indéterminé ; je serai heureux de suggestions.

## 4.3 Nucléosome en traction et torsion

### 4.3.1 Nucléosome en traction

Des données non publiées de E. Praly, obtenues au cours de sa thèse [75], montre que, lorsque le nucléosome est soumis à une force de traction de 3.2 pN, il existe trois états d'extensions différents. Les deux premiers états sont séparés par une différence d'extension de 12 nm. Pour obtenir le dernier état, il faut ajouter 24 nm à partir du second état. Ces sauts sont observés sur des temps très longs, de l'ordre de la seconde. Ceci suggère que le nucléosome doit franchir des barrières énergétiques importantes. Pour expliquer ces trois états d'extensions, je propose un premier modèle. Dans ce premier modèle, j'utilise l'hypothèse de dégrafage symétrique de [72]. Dans [5], on peut trouver la relation d'ordre suivante sur les énergies des SHLs :  $6.5 < 5.5 \sim 2.5 \sim 1.5 < 4.5 < 3.5 < 0.5$ . Je réalise alors une première simulation où je dégrafe de manière symétrique les SHLs  $\pm 6.5$  et réalise une trace de l'extension de ma molécule de nucléosome unique (les autres SHLs fixent rigidement l'octamère aux cylindres d'ADN). A partir de là, je dégrafe les SHLs  $\pm 5.5$ . Ensuite et en accord avec la relation d'ordre de [5], il faudrait dégrafe les SHLs  $\pm 2.5$  et  $\pm 1.5$ . Cependant, ces SHLs sont plus intérieurs que les SHLs  $\pm 4.5$  et  $\pm 3.5$ . Il faut donc dégraffer les  $\pm 4.5$  et  $\pm 3.5$  pour pouvoir enlever les  $\pm 2.5$  et  $\pm 1.5$ . Je construis ainsi deux nouveaux états : l'un où les  $\pm 4.5$  sont enlevés et l'autre où les  $\pm 3.5$ ,  $\pm 2.5$  et  $\pm 1.5$  sont maintenant dégraffés. Le résumé de ces traces est montré sur la figure (4.9). Je ne vois aucune combinaison de deux états pouvant faire une différence d'extension de

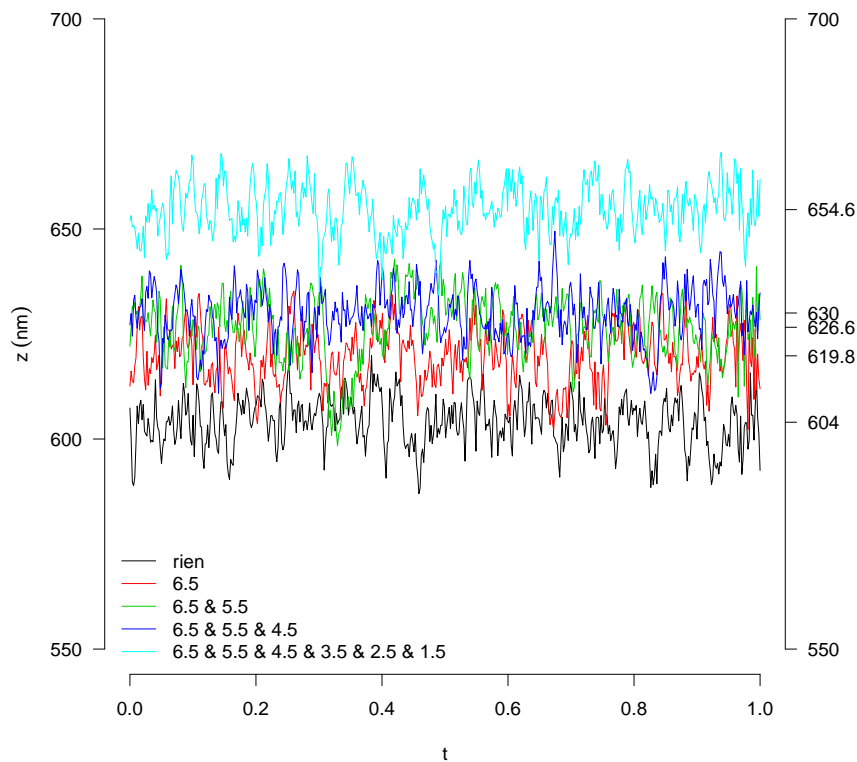


FIGURE 4.9: Traces des extensions de la molécule unique de nucléosome sous une force de traction de  $f = 3.2$  pN pour les états où les SHLs  $\pm 6.5$ ,  $\pm 6.5, 5.5$ ,  $\pm 6.5, 5.5, 4.5$ ,  $\pm 6.5, 5.5, 4.5, 3.5, 2.5, 1.5$  sont dégrafés. L'extension en nm est représentée en ordonnée tandis que le temps de simulation est montré en abscisse.

12 nm. Je remarque tout de même une différence d'extension de 24 nm entre  $\pm 6.5, 5.5, 4.5$  dégrafés et  $\pm 6.5, 5.5, 4.5, 3.5, 2.5, 1.5$ .

Pour aller plus loin, je propose le scénario suivant : je rends les FHB flexible et dégrafe de façon asymétrique les DD. Vous remarquerez que je laisse les dimères fixés sur l'ADN. J'obtiens ainsi deux états de plus que l'état où tous les SHLs sont encore présents. Alors pourquoi ce scénario ? La première raison est que personne ne s'est intéressé à un dégrafage asymétrique jusqu'à maintenant. De plus, un dégrafage symétrique induirait une coopérativité certaine entre les SHLs + et les SHLs -. La seconde raison est que le nucléosome doit pouvoir se reconstituer après avoir été traversé par des processus comme la transcription, la réplication ou la réparation. Pour cela, il faut que les histones restent sur l'ADN. Cette idée est appuyée par l'article de V. Böhm [74]. Le désassemblage du nucléosome se fait par la perte d'interaction à l'interface entre les dimères H2A-H2A et H3-H4, c'est à dire une brisure des DD et un gain de flexibilité au niveau des FHB. Je rentre tout cela dans mes simulations et j'obtiens les résultats montrés sur la figure (4.10). J'obtiens aussi les trois états recherchés : les deux premiers séparés par 12 nm et les deux derniers séparés par 24 nm. Le premier saut en extension, un DD est dégrafé, nous donne un saut en extension de 12 nm. Curieusement, le dégrafage du second DD nous donne un saut de 24 nm.

J'émetts alors l'hypothèse que les DD se dégrafent avant les SHLs. Rappelez vous alors les po-

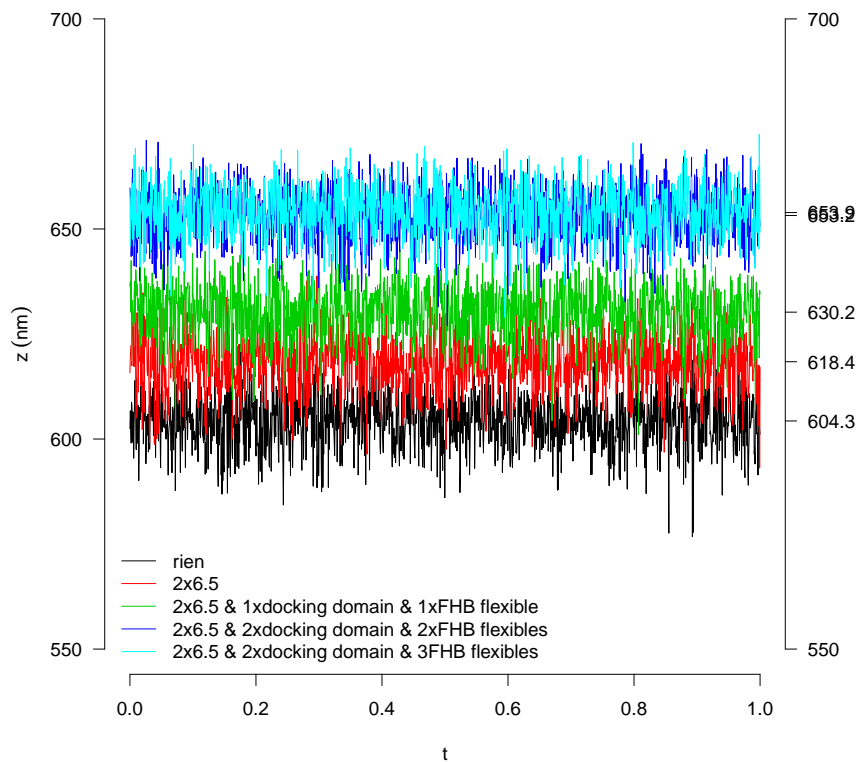


FIGURE 4.10: Traces des extensions de la molécule unique de nucléosome sous une force de traction de  $f = 3.2$  pN pour les états où les SHLs  $\pm 6.5$ ,  $\pm 6.5$ ,  $5.5$ ,  $\pm 6.5$ ,  $5.5$ ,  $4.5$ ,  $\pm 6.5$ ,  $5.5$ ,  $4.5$ ,  $3.5$ ,  $2.5$ ,  $1.5$  sont dégraiffés. L'extension en nm est représentée en ordonnée tandis que le temps de simulation est montré en abscisse.

tentiels de Morse dont je parlais au début du chapitre. Jusqu'à maintenant, je dégrafais les SHL à la "main". Pour pallier cet inconvénient, je vais utiliser les paramètres résumés dans le tableau (4.11). J'espère ainsi obtenir des dégrafages dynamiques. Il faut toutefois noter que le choix des paramètres est arbitraire et la seule contrainte est d'avoir une énergie de liaison pour les DD plus faible que pour les SHL.

Vous pourrez trouver une vidéo du résultat à l'adresse <http://vimeo.com/52080746>. Les figures (4.12) résument les différentes conformations prises par le nucléosome sous une force de traction  $f = 3.2$  pN et couple nul au cours de la simulation. Ces images constituent une preuve du dégrafage asymétrique et non symétrique. Pour aller plus loin, il est nécessaire de réussir le calcul de structure présenté précédemment. Ainsi, les paramètres des potentiels de Morse seraient affinés et la simulation en serait que plus réaliste. À l'heure actuelle, ces simulations ne sont qu'une preuve de principe.

### 4.3.2 Nucléosome en traction et torsion

Assez parlé de nucléosome sous traction ; ajoutons un peu de twist. La réponse d'un nucléosome ou d'une fibre de chromatine a été étudiée expérimentalement par A. Bancaud [76, 71]. Il a d'ailleurs été postulé l'existence d'une transition chirale du nucléosome ; l'ADN enroulé à gauche s'enroule à droite sous une contrainte de torsion. Ce changement de chiralité est par exemple utilisé dans un

SHLs	$D$ ( $k_B T$ )	$a$ ( $\text{nm}^{-1}$ )	$r_e$ (nm)
$\pm 6.5$	2.000000	3.465735	0.200000
$\pm 5.5$	10.000000	0.800007	0.866427
$\pm 4.5$	15.000000	0.514564	1.347056
$\pm 3.5$	20.000000	0.781650	0.886774
$\pm 2.5$	15.000000	0.852186	0.813375
$\pm 1.5$	15.000000	0.620227	1.117569
$\pm 0.5$	20.000000	0.849568	0.815882

(a) Paramètres pour les SHL.

Interaction dimère-dimère	$D$ ( $k_B T$ )	$a$ ( $\text{nm}^{-1}$ )	$r_e$ (nm)
Docking Domain	5.000000	0.693147	1.000000
Four-Helix-Bundle	15.000000	0.393918	1.759624
Four-Helix-Bundle	20.000000	0.406587	1.704796
Four-Helix-Bundle	15.000000	0.393918	1.759624

(b) Paramètres pour les DD et FHB.

FIGURE 4.11: Tableau récapitulatif des paramètres des potentiels de Morse utilisés pour reproduire les SHL, FHB et DD.

modèle de la transcription dans la fibre de chromatine par C. Bécavin [77]. Une vague de réversomes se propage en aval de la polymérase tandis qu'en amont, la fibre de chromatine n'est constituée que de nucléosomes. Cette transition a déjà été montrée par H. Wong dans sa thèse [55]. A cette époque, le dégrafage des DD se faisaient en même temps et à la main. En suivant cette façon de procéder j'obtiens la vidéo disponible à l'adresse <http://vimeo.com/51918438>. Je vous montre maintenant la transition chirale de façon dynamique, disponible à l'adresse <http://vimeo.com/52089760>. Dans un premier temps l'ADN nucléosomal change de sens d'enroulement ; ceci est montré dans la succession d'images (4.13).

Dans un second temps, et si le couple appliqué sur la bille est suffisamment important, la superhélice forme une boucle autour de laquelle nous voyons des plectonèmes apparaître (voir les images (4.14)).

Et enfin, je vous montre cette même transition mais cette fois ci à l'intérieure d'une fibre de chromatine. En partant d'une configuration à l'équilibre d'une fibre de chromatine sous une force de traction de  $f = 1$  pN (voir la figure (4.15)), j'applique un couple de  $\Gamma = 10$  pN · nm sur la bille magnétique. Cette valeur du couple est nécessaire [6] pour obtenir une transition chirale, surtout pour une telle valeur de force de traction. Après avoir atteint l'équilibre statistique, je peux prendre en photo les configurations de la fibre, et, en particulier, voir des nucléosomes côtoyant des réversomes (voir la figure (4.16)). Vous pourriez dès à présent m'objecter le fait suivant : au début du chapitre il est dit que la force exercée par la polymérase est faible ; pourquoi utiliser une force de 1 pN ? Ce choix de force est juste une raison pratique pour la visualisation. Pour de faibles forces, le cadrage des vidéos est difficile car les nucléosomes diffusent dans un plus grand espace qu'à grandes forces. En diminuant la force de traction, la transition chirale n'est que plus facile ; ce résultat se retrouve bien en simulation.

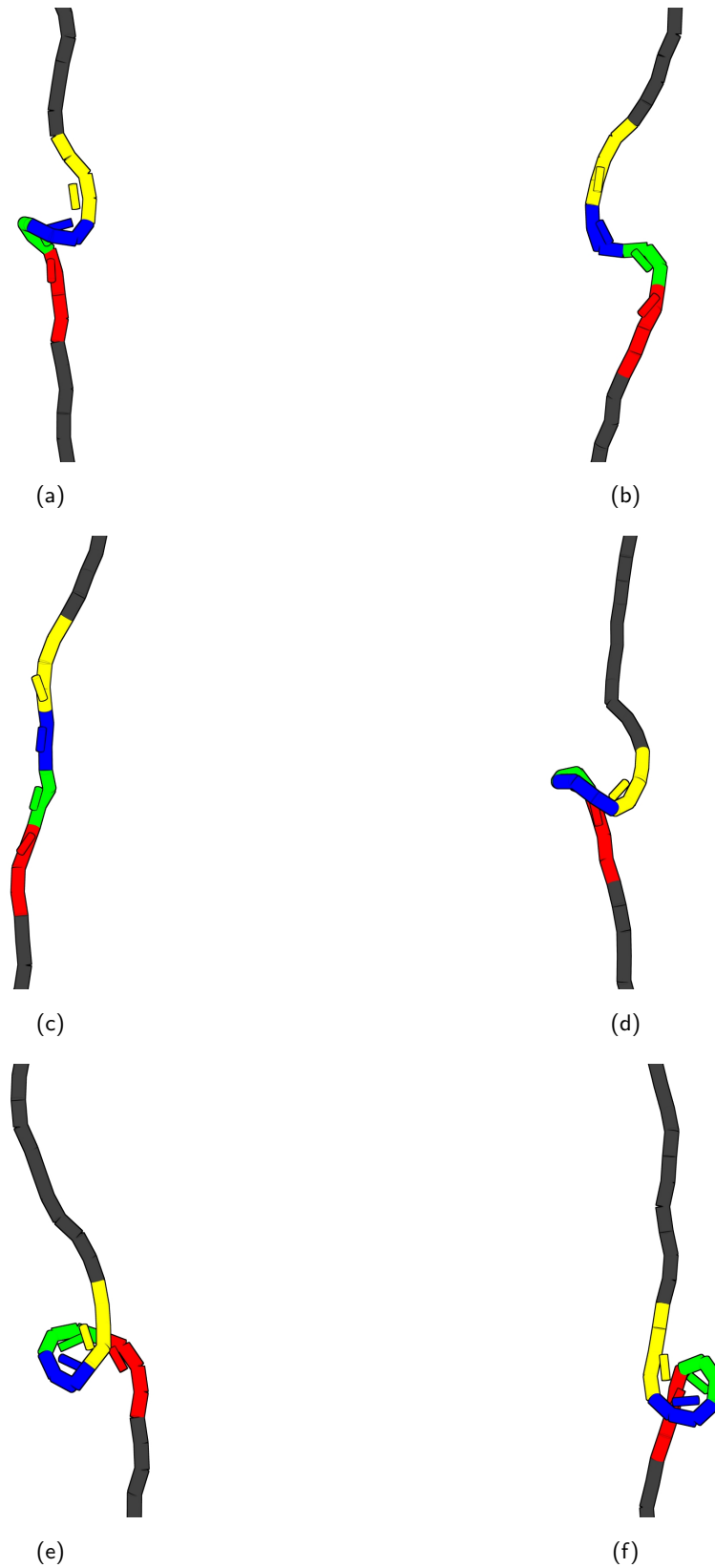


FIGURE 4.12: Ces images illustrent les différentes conformations accessibles au nucléosome. La simulation est réalisée pour une force de traction de 3.2 pN et un couple nul.



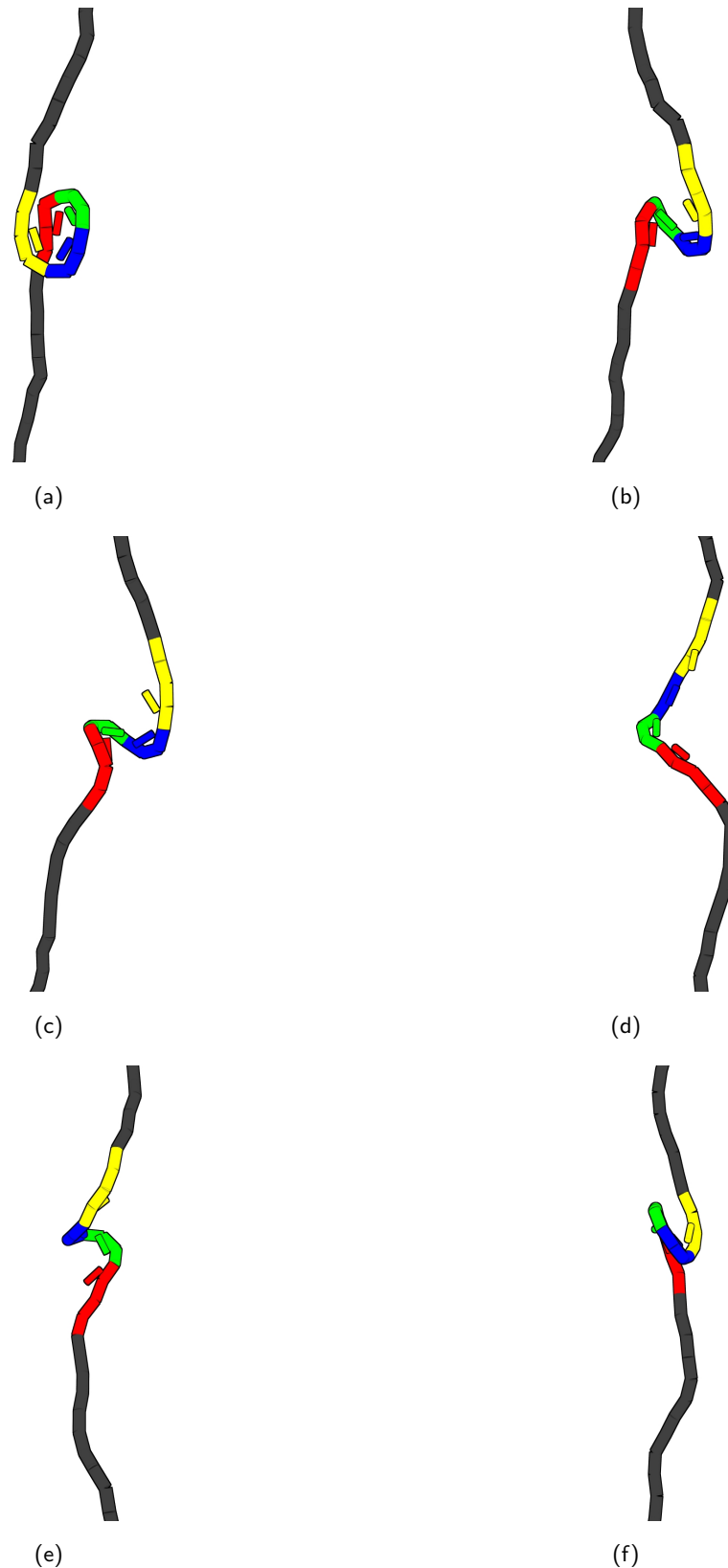


FIGURE 4.13: Ces images illustrent les différentes conformations accessibles au nucléosome. La simulation est réalisée pour une force de traction de 1 pN et un couple de  $\Gamma = 10 \text{ pN} \cdot \text{nm}$ . Au fil des images, vous voyez le changement de chiralité de l'ADN nucléosomal.

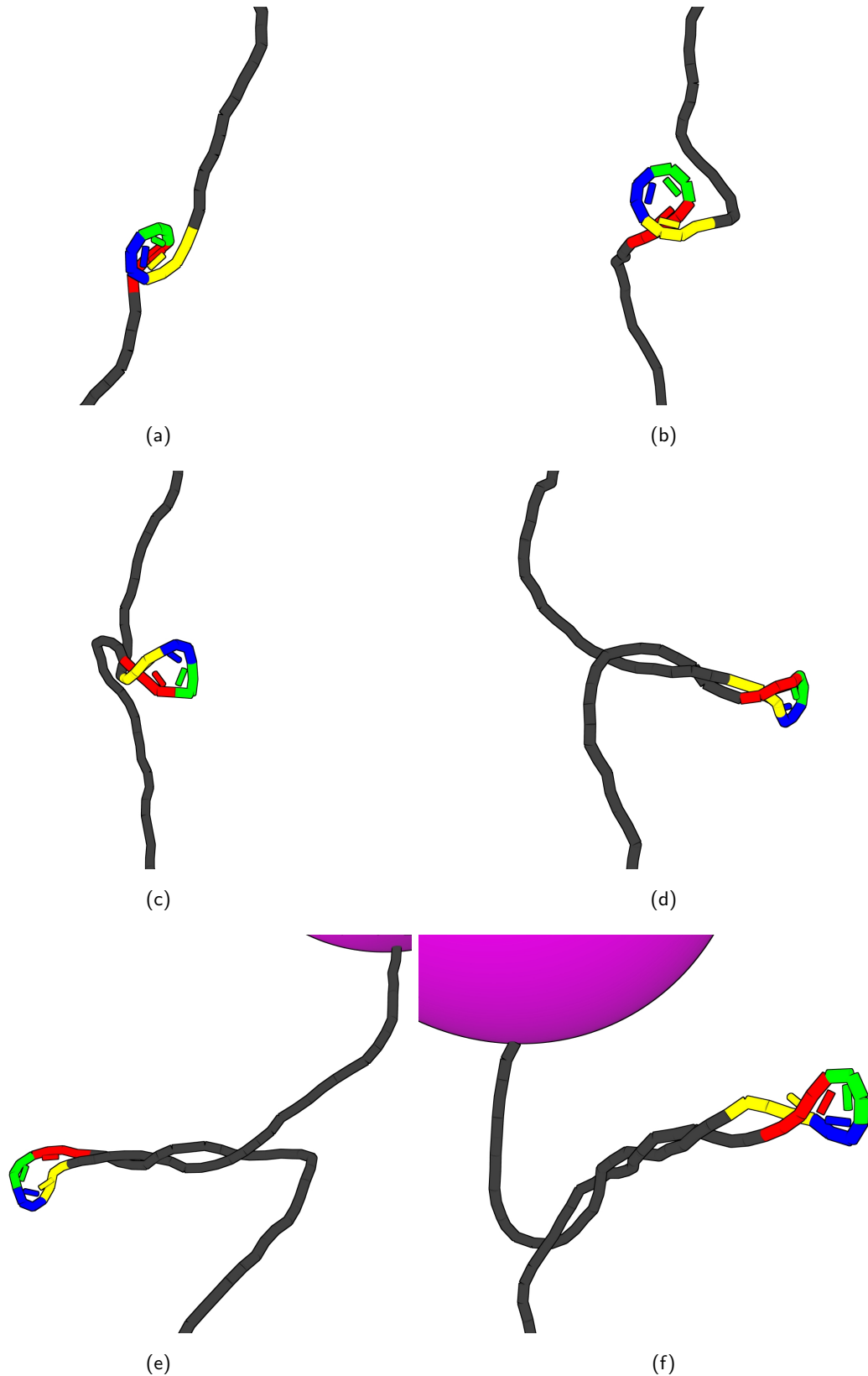


FIGURE 4.14: Ces images illustrent les différentes conformations accessibles au nucléosome. La simulation est réalisée pour une force de traction de 1 pN et un couple de  $\Gamma = 10 \text{ pN} \cdot \text{nm}$ . Le couple est supérieur à la valeur critique de formation d'un réversome [6] et ainsi l'ADN commence à former des plectonèmes autour du réversome.

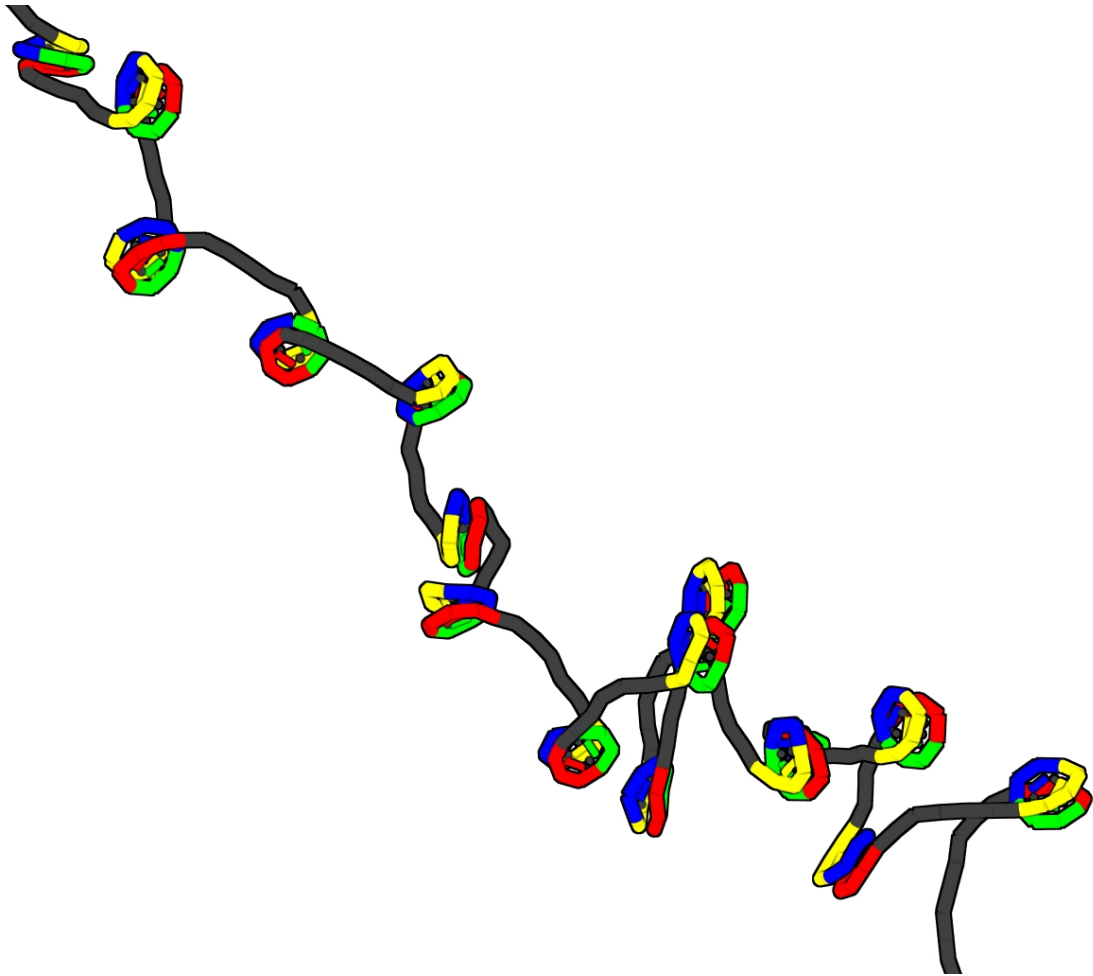


FIGURE 4.15: Vue schématique d'une fibre de chromatine de 20 nucléosomes sous une force de traction  $f = 1$  pN et couple nul. Vous m'excuserez volontiers le cadrage quelque peu oblique mais nécessaire pour faire entrer la totalité des nucléosomes.

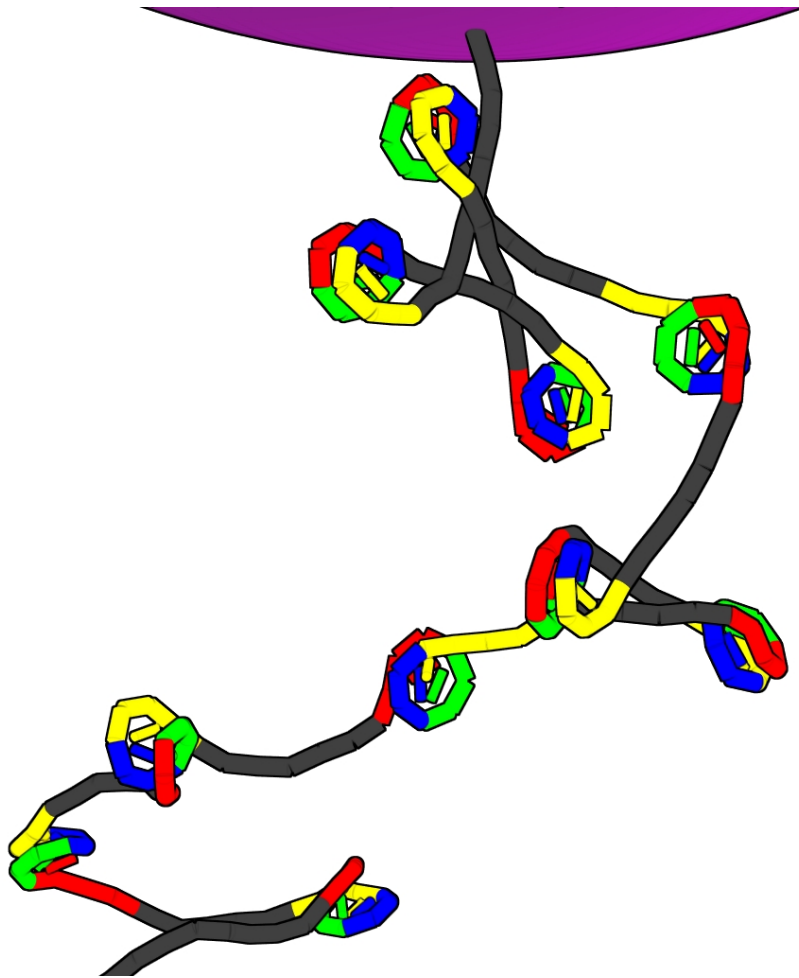


FIGURE 4.16: Vue schématique d'une fibre de chromatine sous une force de traction  $f = 1$  pN et couple  $\Gamma_n = 10$  pN · nm. Vous pouvez observer des nucléosomes et réversomes au sein de cette même fibre. La vidéo de la simulation est accessible à l'adresse <http://vimeo.com/52072568>



## Levure

Dernier étage de l'échelle, je vais discuter la structure du noyau de levure en utilisant la physique des polymères.

Le noyau de levure contient 16 chromosomes attachés par des microtubules en un lieu appelé le Spin-Pole-Body (SPB). Il existe, à l'intérieur du noyau un compartiment appelé nucléole, site de la transcription de l'ARN ribosomal et qui est constitué par une partie du bras droit du chromosome 12. Je suppose que les chromosomes peuvent être décrits par la physique des polymères. En particulier, le confinement et le volume exclu des chromosomes induisent une structure en blobs (voir la figure (5.1)). La taille des blobs est la longueur caractéristique à partir de laquelle un segment de chromosome ne voit plus ses voisins. Une autre façon d'expliquer ces blobs est de les comprendre en terme de compétition entre agitation thermique et pression osmotique. De ces deux arguments, on en déduit que la ligne passant par le centre des blobs suit une loi de Random Walk (RW). Cependant, les parties se situant à l'intérieur des blobs suivent une loi de Self-Avoiding-Walk (SAW). Je suppose que nous pouvons décrire un chromosome comme un polymère de  $N$  segments de Kuhn. Le volume exclu d'un cylindre de longueur  $K$  et diamètre  $d$  est exprimé comme  $v = \pi K^2 d$ . Les modèles actuels de fibre de chromatine donnent un diamètre entre 10 et 30 nm.

La taille d'un polymère de  $N$  segments sans volume exclu est donnée par l'équation (5.1)

$$R = N^{1/2} K \quad (5.1)$$

La taille d'un polymère de  $N$  segments avec volume exclu est donnée par l'équation (5.2)

$$R = \left( \frac{a}{K} \right)^{1/5} N^{3/5} K \quad (5.2)$$

A. Grosberg [78] montre que dans le cas où  $(d/K)/\sqrt{N} \ll 1$  les chromosomes n'ont plus de volume exclu ; l'exposant  $3/5$  est remplacé par  $1/2$  et nous avons un RW  $R = N^{1/2} K$ . Introduisons, avant de poursuivre, quelques quantités et notations utiles. J'appelle  $G \simeq (12 - 2) \otimes 10^6$  bp la taille du génome de levure sans tenir compte du nucléole,  $V = \frac{\pi}{8} D^3$  le volume accessible au chromosome

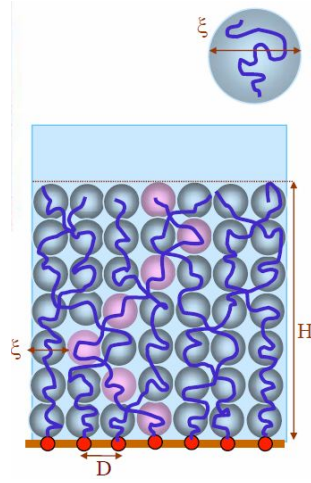


FIGURE 5.1: Vue schématique d'une brosse de polymères (polymères greffés sur un substrat) où nous pouvons apprécier la structure en blobs. Cette image est empruntée à un exposé de E. Zhulina.

soit 75 % du noyau,  $c = \frac{G}{V}$  la concentration en bp/ $\mu\text{m}^3$  et  $\alpha = K_{\text{bp}}/K$  la compaction de la fibre de chromatine en bp/nm.

A l'intérieur du blob, la taille du polymère suit une loi (5.3) de SAW où  $g$  est le nombre de segments de Kuhn par blob.

$$B = \left(\frac{a}{K}\right)^{1/5} g^{3/5} K \quad (5.3)$$

La concentration  $c_b$  en bp/ $\text{nm}^3$  dans le blob peut être écrite comme l'équation (5.4). Dans la suite, nous admettrons que cette concentration est égale à la concentration dans le noyau  $c$ .

$$c_b = \frac{g\alpha K}{B^3} \quad (5.4)$$

En combinant ces deux résultats, nous obtenons le nombre de segments de Kuhn  $g$  à l'intérieur du blob (5.5).

$$g = \left(\frac{c_b}{\alpha}\right)^{-5/4} a^{-3/4} K^{-7/4} \quad (5.5)$$

soit une taille de blobs donnée par l'équation (5.6).

$$B = (aK)^{-1/4} \left(\frac{c_b}{\alpha}\right)^{-3/4} \quad (5.6)$$

Ce résultat est montré dans la figure (5.2). J. Dekker [79] donne un intervalle de confiance à 95 % pour la compaction  $\alpha$  et la longueur de Kuhn  $K$  de la fibre de chromatine pour des régions riches en paires de bases AT et des régions riches en paires de bases GC. Les régions riches en GC sont les régions riches en gènes.

Comme vous pouvez le constater, le nombre de segments de Kuhn à l'intérieur d'un blob ne dépend que de la concentration génomique  $c_b$ . En effet, la compaction et la longueur de Kuhn de la fibre de chromatine ne devrait pas varier d'un organisme à l'autre. La concentration génomique ne variant pas elle aussi, la taille des blobs devrait être la même pour tous les noyaux. Le noyau de levure peut être considéré comme l'organisme de base pour décrire la structure nucléaire en terme de brosses de polymères. Au lieu de discuter en termes de nombre de segments de Kuhn  $N$  de fibre de chromatine, j'introduis l'abscisse curviligne  $s$ , distance le long d'un chromosome, qui peut être

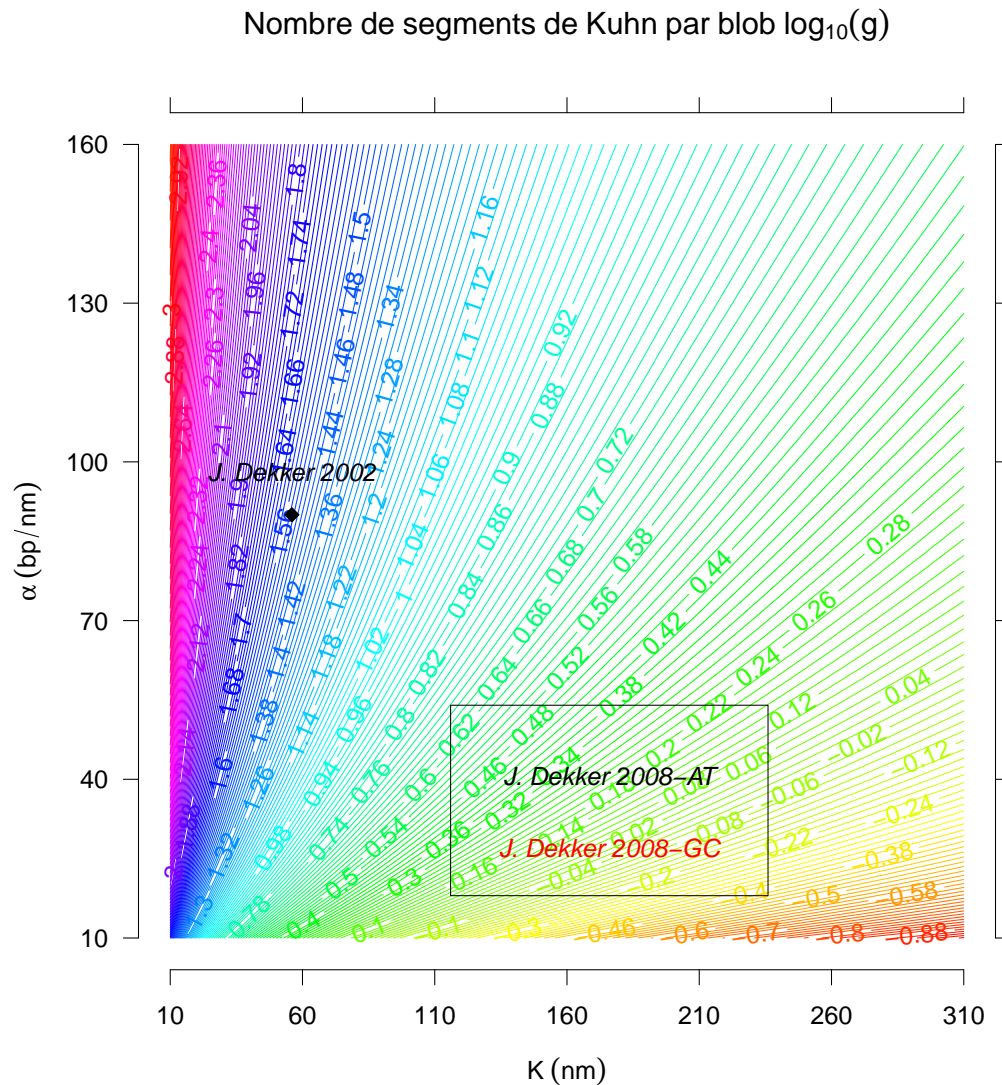


FIGURE 5.2: Le nombre de segments de Kuhn par blob est montré en échelle logarithmique  $\log_{10} g$  (contour de couleurs) en fonction de la compaction  $\alpha$  (bp/nm) (en ordonnée) et de la longueur de Kuhn de la fibre  $K$  (nm) (en abscisse).

reliée à la distance génomique  $s = s_{bp}/\alpha$ . Nous pouvons alors écrire le nombre de segments de Kuhn jusqu'à  $s_{bp}$  comme  $N = s_{bp}/\alpha K$ . A l'intérieur du blob  $s_{bp} < \alpha B$  nous avons l'équation (5.7).

$$R_{s < B}(\Delta s_{bp}) = a^{1/5} K^{1/5} \alpha^{-3/5} \Delta s_{bp}^{3/5} \quad (5.7)$$

Pour  $s_{bp} > \alpha B$  la taille à la distance génomique  $s_{bp}$  est donnée par un RW de  $N/g$  blobs ; nous avons l'équation (5.8) à partir de  $R = (N/g)^{1/2} B$ .

$$R_{s > B}(\Delta s_{bp}) = \alpha^{-1/2} g^{1/10} K^{3/10} a^{1/5} \Delta s_{bp}^{1/2} \quad (5.8)$$

[80] nous fournit la distance spatiale quadratique entre deux points des chromosomes 5, 6 et 14 séparés par une distance génomique  $s_{bp}$ . Je choisis d'ajuster une courbe du type  $R \propto s_{bp}'$



sur les données du chromosome 14; il n'y a qu'une mesure pour le chromosome 5, quatre mesures pour le chromosome 6 et huit mesures pour le chromosome 14. Je trouve un seul exposant  $\nu = 0.6082 \pm 0.1339$  ce qui correspond à un SAW. J'en déduis que  $(aK)^{1/5} \alpha^{-3/5} = 0.3683 \pm 0.5513$ . Suivant le modèle de fibre utilisé, nous pouvons déduire la compaction observée dans les expériences de K. Bystricky [80]. Le résultat de l'ajustement est montré sur la figure (5.3(d)). A partir des données de simulation du noyau de levure de H. Wong [81] nous pouvons recommencer la procédure d'ajustement d'un modèle de RW de blobs et d'un modèle de SAW à l'intérieur du blob. Les simulations de H. Wong sont résumées dans [81]. Les chromosomes sont modélisés par un modèle FJC avec des segments de Kuhn de longueur 60 nm et un diamètre de 20 nm. C'est le modèle communément admis pour décrire la fibre de chromatine dans le noyau de levure. Chaque segment "contient" 5000 bp soit une compaction  $\alpha \simeq 83.3$  bp/nm. Les tailles des 16 chromosomes sont 46, 162, 63, 304, 115, 54, 218, 112, 88, 149, 133, 516, 185, 157, 218, 190 en nombre de segments de Kuhn. Les chromosomes sont attachés par des microtubules de longueur 380 nm qui sont eux mêmes librement fixés au SPB. Les deux bras des chromosomes n'ont pas la même longueur et le calcul de  $R(\Delta s)^2$  se fait à partir des centromères. Ce "greffage" laisse ainsi une liberté importante de mouvement au point d'ancrage des chromosomes. Pour modéliser le nucléole, H. Wong ajoute au bras droit du chromosome 12 un millier de paires de bases sous forme de sphères de diamètre ajustable. Le nombre de sphères est choisi de telle manière à avoir un nucléole de volume constant, soit, 25 % du volume du noyau. La pression osmotique entre les cylindres d'ADN et les sphères permet la formation du nucléole à l'extrémité opposée du SPB. Une sphère de diamètre  $D = 2 \mu\text{m}$  et centrée à l'origine du référentiel du laboratoire et modélise l'enveloppe nucléaire. Le SPB se trouve à une distance de  $1 \mu\text{m}$  du centre dans la direction  $z$ , soit le pôle nord de la sphère. H. Wong a ajouté dans ces simulations des interactions entre les extrémités des chromosomes (télomères) et la membrane nucléaire. Ces interactions de contacts laissent les télomères libres de diffuser le long de la membrane. Les figures (5.3(a)), (5.3(b)), (5.3(c)) et (5.3(d)) montrent les distances quadratiques moyennes suivant les trois axes  $R_{x,y,z}^2(\Delta s)$  et la distance quadratique totale  $R^2(\Delta s)$  entre deux cylindres séparés par une distance génomique  $\Delta s$  pour le chromosome 4 (chromosome le plus long) pour des diamètres de sphères de 30, 100, 200 nm. Ce choix nous permet d'avoir une meilleure statistique sur le  $R^2(\Delta s)$  et les résultats qui suivent sont les mêmes pour les autres chromosomes. Première remarque : il est difficile de dire quel nucléole correspond le mieux aux expériences de K. Bystricky. Deuxième remarque : à la différence des données expérimentales de K. Bystricky, je trouve deux pentes. La première donne un exposant d'environ 1.25 et la seconde, un exposant approximativement égal à 1.02. Et bien, nous sommes assez proche des exposants d'un SAW à l'intérieur des blobs et d'un RW de blob. En assimilant le changement de pente de  $R^2(\Delta s)$  au passage d'un SAW à l'intérieur des blobs à un RW de blobs, la taille d'un blob chez H. Wong est de l'ordre de 300 nm soit 100000 bp. Cette taille semble ne pas être trop affectée par le diamètre des billes du nucléole. Nous pouvons remarquer une légère différence d'aspect entre les  $R_{x,y}^2(\Delta s)$  et  $R_z^2(\Delta s)$ . Les fluctuations latérales (en  $x$  et  $y$ ) des chromosomes suivent vraiment une loi de RW de blobs. Il semble que ce fait soit moins vrai pour les fluctuations en  $z$ . Pour en savoir un peu plus, je m'intéresse à la position moyenne des cylindres en  $z$  et j'observe la loi d'échelle  $\langle z \rangle \propto s$  (voir la figure (5.4)). Cette relation se retrouve dans les brosses de polymères [82, 83, 84] où la taille d'une chaîne le long de la normale à la surface de greffage est proportionnelle au nombre de monomères  $N$ . La moyenne  $\langle z \rangle$  semble être une meilleure observable pour une brosse de polymères que  $R_z^2(\Delta s)$ . En particulier, "utiliser" une brosse de polymères pour la conception du noyau de levure permet d'éviter les enchevêtrements

de chromosomes. Il serait intéressant de poursuivre l'étude du  $R_z^2(\Delta s)$  pour des organismes comme la drosophile où les chromosomes sont beaucoup plus grands que ceux de la levure. La statistique devrait être bien meilleure et permettrait d'aller à de très grandes distances génomiques  $\Delta s$  pour le calcul du  $R^2(\Delta s)$ . Le reste de ce chapitre n'est rien d'autre que l'article de H. Wong pour une description plus biologique du noyau de levure.

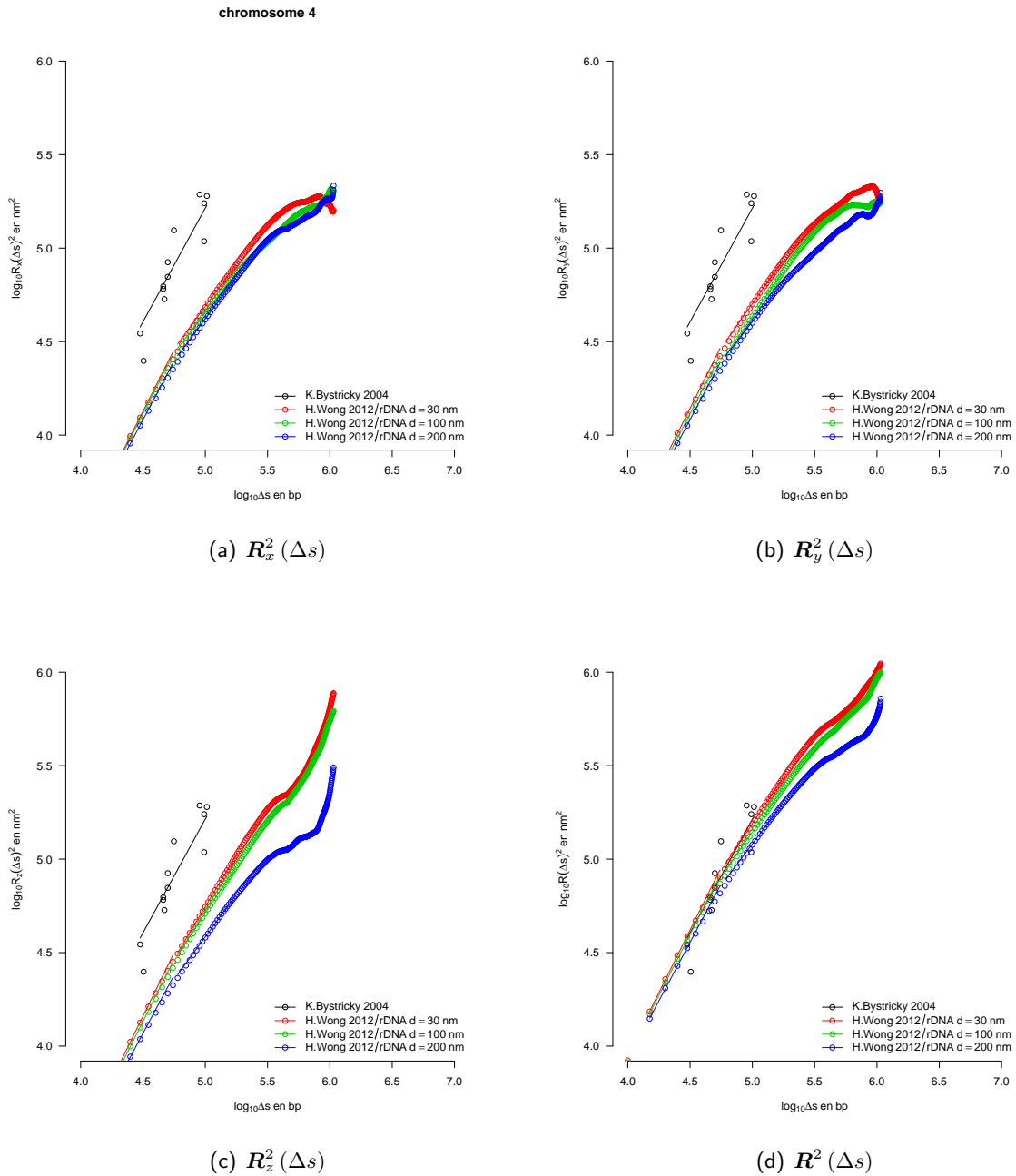


FIGURE 5.3: Les distances quadratiques  $R_{x,y,z}^2(\Delta s)$  (nm) et  $R^2(\Delta s)$  entre deux cylindres de fibre de chromatine (en ordonnée) sont montrées en fonction de la distance génomique  $\Delta s$  (bp) les séparant (en abscisse) pour le chromosome 4. L'échelle est logarithmique en base 10. Les trois courbes correspondent respectivement à des diamètres de sphères pour le nucléole de 30, 100, 200 nm.

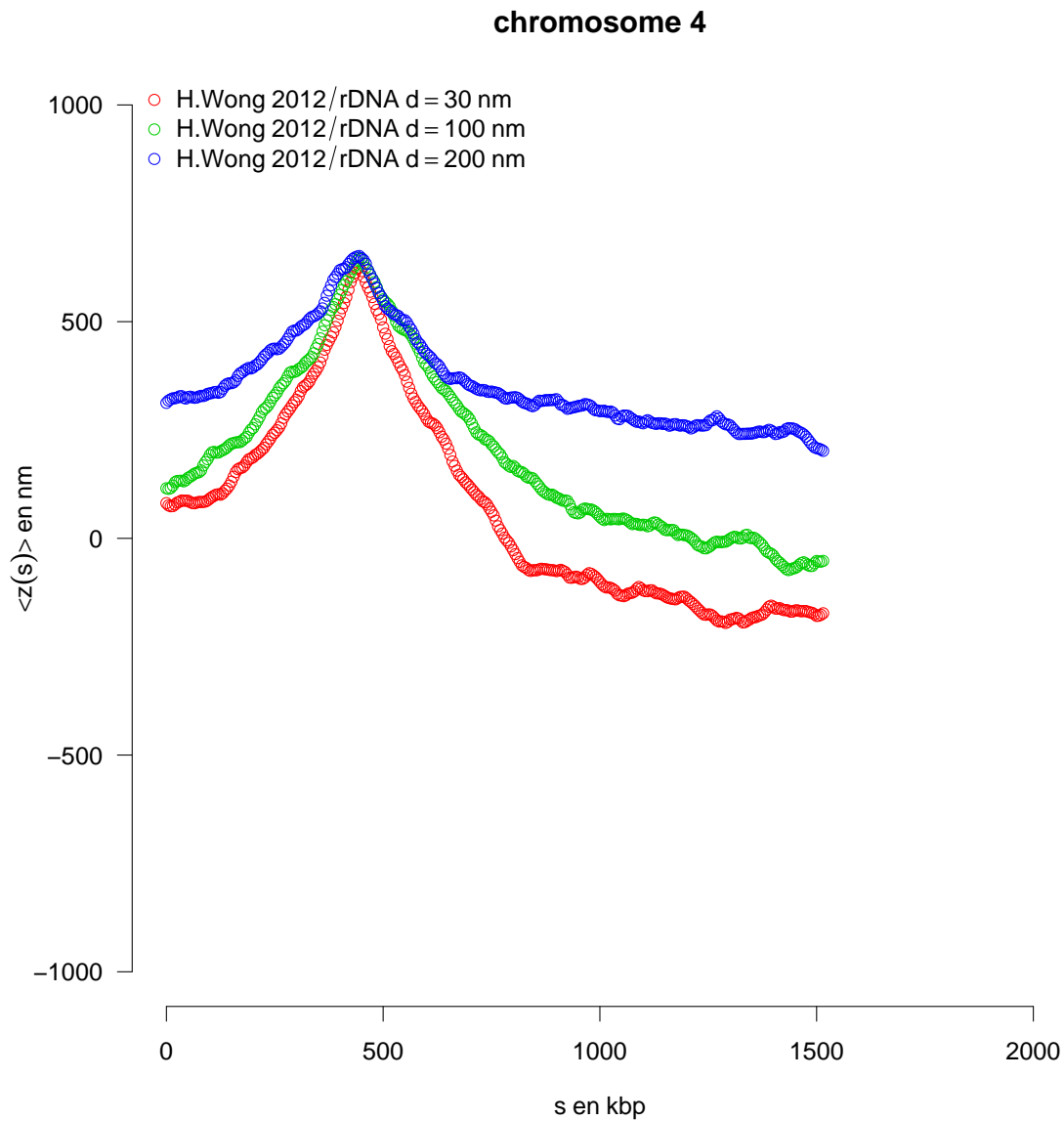


FIGURE 5.4: La position moyenne  $\langle z(s) \rangle$  (en ordonnée) des cylindres est montrée en fonction de la distance génomique au centromère  $s$  (en abscisse). Les trois courbes correspondent respectivement à des diamètres de sphères pour le nucléole de 30, 100, 200 nm. Nous pouvons remarquer qu'autour du centromère (soit au environ de 500 kbp), nous avons la loi d'échelle  $\langle z \rangle \propto s$ .

# A Predictive Computational Model of the Dynamic 3D Interphase Yeast Nucleus

Hua Wong,<sup>1,4,6</sup> Hervé Marie-Nelly,<sup>1,4,5,6</sup>  
Sébastien Herbert,<sup>1,4,6,8</sup> Pascal Carrivain,<sup>3,4</sup> Hervé Blanc,<sup>2,4,7</sup>  
Romain Koszul,<sup>2,4,7</sup> Emmanuelle Fabre,<sup>2,4,7</sup>  
and Christophe Zimmer<sup>1,4,6,\*</sup>

<sup>1</sup>Institut Pasteur, Groupe Imagerie et Modélisation,  
75015 Paris, France

<sup>2</sup>Institut Pasteur, Laboratoire Régulation Spatiale du Génome,  
75015 Paris, France

<sup>3</sup>CNRS, LPTMC, UMR 7600

<sup>4</sup>CNRS, GDR 3536

Université Pierre et Marie Curie, 75252 Paris, France

<sup>5</sup>Université Pierre et Marie Curie, Cellule Pasteur,  
75015 Paris, France

<sup>6</sup>CNRS, URA 2582, 75015 Paris, France

<sup>7</sup>CNRS, UMR 3525, 75015 Paris, France

<sup>8</sup>Université Paris Diderot, Sorbonne Paris Cité, Cellule Pasteur,  
75015 Paris, France

## Summary

**Background:** Despite the absence of internal membranes, the nucleus of eukaryotic cells is spatially organized, with chromosomes and individual loci occupying dynamic, but nonrandom, spatial positions relative to nuclear landmarks and to each other. These positional preferences correlate with gene expression and DNA repair, recombination, and replication. Yet the principles that govern nuclear organization remain poorly understood and detailed predictive models are lacking.

**Results:** We present a computational model of dynamic chromosome configurations in the interphase yeast nucleus that is based on first principles and is able to statistically predict the positioning of any locus in nuclear space. Despite its simplicity, the model agrees with extensive previous and new measurements on locus positioning and with genome-wide DNA contact frequencies. Notably, our model recapitulates the position and morphology of the nucleolus, the observed variations in locus positions, and variations in contact frequencies within and across chromosomes, as well as subchromosomal contact features. The model is also able to correctly predict nuclear reorganization accompanying a reduction in ribosomal DNA transcription, and sites of chromosomal rearrangements tend to occur where the model predicted high contact frequencies.

**Conclusions:** Our results suggest that large-scale yeast nuclear architecture can be largely understood as a consequence of generic properties of crowded polymers rather than of specific DNA-binding factors and that configurations of chromosomes and DNA contacts are dictated mainly by genomic location and chromosome lengths. Our model provides a quantitative framework to understand and predict large-scale spatial genome organization and its interplay with functional processes.

## Introduction

Besides the one-dimensional information carried by the nucleotide sequence, the three-dimensional arrangement of the genome in the nucleus of eukaryotic cells emerges as an important determinant of gene expression, DNA repair, recombination, and replication [1]. Although they lack any membrane apart from the nuclear envelope, nuclei from yeast to humans exhibit strong compartmentalization into nuclear bodies and other functionally distinct subdomains. In metazoans, chromosomes are confined to nonoverlapping territories, whose relative positions in the nucleus are not random [2]. Individual loci occupy preferential, though dynamic, positions with respect to their chromosome territory or other nuclear landmarks such as the nuclear envelope or the nucleolus [3, 4]. These positioning patterns affect the transcriptional status of genes, and the nucleus appears compartmentalized in domains that either favor or silence gene expression [5]. Cancer-promoting chromosomal rearrangements arise from illegitimate fusions between broken parts of the same or distinct chromosomes. In mammals, some of these events occur preferably at pairs of genomic locations that are more frequently in contact and are thus affected by spatial positioning of chromosomal regions [6–10]. Conversely, in yeast, many loci move to new subnuclear positions or change their dynamics upon changes in their expression or as a result of DNA breaks [5, 11–14].

Despite such functional relevance, the main factors and mechanisms that control dynamic nuclear organization are presently ill understood. These can be divided into two broad classes [15]. A first class includes protein complexes, such as CTCF, which bind to particular discrete DNA sequence motifs and promote their interaction with nuclear landmarks or distal loci, thus restricting subnuclear positions or forming loops or interchromosomal attachments [16, 17]. In a second class are generic effects arising from the properties of semiflexible polymers (chromosomes) confined to the crowded nuclear volume. Generic properties have been proposed to explain, for example, the formation of chromosome territories, and the aggregation of large macromolecular complexes in the nucleus [18, 19]. Although both specific factors and generic effects are present, their relative contributions remain unclear and a detailed predictive model of nuclear organization is not yet available.

The well-studied budding yeast *Saccharomyces cerevisiae* provides an attractive model to study nuclear organization and its functional relevance. Electron microscopy has revealed structural nuclear landmarks [20]; light microscopy has allowed to map the positions and dynamics of selected loci in individual nuclei [3, 21–25]; and a chromosome conformation capture (3C) assay [26] coupled with massive DNA sequencing (Hi-C) has provided a matrix of contact frequencies across the genome [27]. Despite its small diameter (~2 μm), the yeast nucleus is characterized by strong functional compartmentalization [3, 5, 15]. The most prominent nuclear compartment is the nucleolus, the site of transcription of ribosomal RNA genes (rDNA), consisting of ~100–200 tandem repeats on the right arm of chromosome 12. In *S. cerevisiae*, the nucleolus is

\*Correspondence: [czimmer@pasteur.fr](mailto:czimmer@pasteur.fr)

a single, crescent-shaped structure abutting the nuclear envelope, encompassing roughly 1/3 of the nuclear volume, and excluding the bulk of the genome except for the rDNA [20, 22, 27]. What determines the position and shape of the nucleolus, as well as its segregation from the rest of the genome, remains unknown.

The budding yeast nucleus is further characterized by a distinct Rab1-like chromosome configuration, in which each chromosome's centromere is tethered by a single microtubule and the kinetochore complex to the spindle pole body (SPB), a multiprotein complex embedded in the nuclear envelope and located opposite of the nucleolus [3, 28]. Telomeres are tethered to the nuclear envelope via redundant pathways [29]. Their spatial position in the nucleus, as well as that of internal loci, correlates with genomic distance from the centromere [3, 22]. These data are qualitatively consistent with a relatively simple configuration of chromosomes governed by generic physical constraints [15]. By contrast, much more complex configurations, with the chromatin fiber crisscrossing the nucleus, have been proposed based on the reported colocalization of genes such as transfer RNA [30]. How exactly the chromatin fiber is organized in 3D yeast nuclear space thus remains unclear.

Contact frequencies measured by Hi-C have been used to construct a static, and more recently, a dynamic 3D model of yeast chromosomes in the nucleus [27, 31]. The dynamic model was found to be consistent with measurements of distances between telomeres [21]. However, both models [27, 31] assumed untested relationships between spatial distances and contact frequencies, and the dynamic model assumed an artificial nucleolar compartment [31]. As they rely on experimental data, such models cannot predict how nuclear organization changes in different experimental conditions.

Here, we present a predictive quantitative model of dynamic chromosome arrangements in the yeast nucleus. Our model is based on first principles rather than derived from imaging or Hi-C data. Nonetheless the model recapitulates observed patterns of intranuclear locus positioning and chromatin contacts across the genome. In addition, our model correctly predicted an alteration of nuclear architecture in response to a reduction of rDNA gene expression, and may be used in the future to predict the propensity of different pairs of loci to undergo recombination.

## Results

### Computational Model of Dynamic Yeast Chromosomes

We built a computational simulation of chromosome configurations and their dynamics in the yeast nucleus (Figure 1; see Figure S1 and Movie S1 available online). Details are provided in the Experimental Procedures. We considered a nominal model and three control models (Tables 1 and 2). Briefly, we modeled the 16 chromosomes of haploid yeast as freely jointed chains of segments characterized by constant diameter, compaction, and rigidity parameters (Figure 1A; Table 1). Our nominal model reflects the specific nature of the rDNA chromatin (heteropolymer model). At the rDNA locus, ribosomal subunits are assembled cotranscriptionally, leading to strong accumulation of RNA and proteins [20]. To account for this, we increased the diameter of the rDNA segments, such that the effective volume occupied by rDNA was  $\sim 1/3$  of the nucleus (Figure 1B; Table 1). Each chromosome was linked to the SPB at its centromere by a single rigid microtubule (Figure 1C–1E). The telomeric ends were maintained

near the nuclear envelope (represented by a spherical shell) by an outward force, but allowed to move freely along its surface. Assumed parameter values were based on the literature and are summarized in Table 1. Chromosomes were subjected to random thermal motions only (Movie S1). The model incorporated topological constraints, such that chain segments could not penetrate each other. From the simulated families of dynamic chromosome configurations (Figure S1I), we then computed several features of nuclear organization, including intranuclear distributions for any locus (Figure S1J), distances between any pair of loci (Figure S1K), and contact frequencies between any pair of chromosomal regions (Figure S1L). For comparisons, we also considered three control models: (1) a phantom model in which topological constraints were removed, (2) a homopolymer model in which all chromosomes, including the rDNA, had the same properties (except their genomic length), and (3) a microtubule-free model, in which centromeres were not linked to the SPB (Table 2).

### Model Recapitulates Formation of the Nucleolar Compartment and Quantitative Locus Positions

As a first test of our model, we compared the predicted subnuclear locations of selected loci to those obtained from imaging experiments [3, 22] (Figure 2). These positions can be visualized as probability maps in a coordinate system ( $R \cos \alpha$ ,  $R \sin \alpha$ ), where  $R$  is the radial distance to the center of the nucleus, and  $\alpha$  is the angle with respect to the axis joining the nuclear and nucleolar centers [3] (Figure S1J). In our nominal model, centromeres occupied territories roughly halfway between the SPB and the nuclear center, whereas telomeres distributed themselves at the nuclear periphery, in accordance with observations (Figures 2A–2E, 2G–2K). This is unsurprising because our model tethered centromeres to the SPB via microtubules and tethered telomeres to the nuclear envelope. By contrast, the subnuclear location of other loci was not built in. Remarkably, our model predicted that the rDNA locus displayed a crescent-shape distribution abutting the nuclear envelope and a position opposite the SPB (Figure 2F). This morphology was strikingly similar to the rDNA territory determined by light microscopy (Figure 2L) and to the dense nucleolus observed by electron microscopy [3, 20]. The model predicted that all DNA except the rDNA is excluded from the nucleolus, such that the telomeres of even long arms cannot extend to the face opposite the SPB, in agreement with experimental data [3, 22]. By contrast, the control models failed to reproduce the territories of at least some loci: both the phantom model and the homopolymer model led to strikingly different patterns of localization (Figures 2A'–2F' and 2A''–2F''), whereas the microtubule-free model failed to position the rDNA opposite the SPB (Figures 2A''' and F'''). Thus, our heteropolymer model qualitatively recapitulates experimentally observed features of nuclear organization, notably the morphology of the nucleolus, and the segregation of the rDNA from the rest of the genome.

We next turned to a more quantitative test of predicted locus positions. First, we analyzed the “absolute” intranuclear positions of selected loci. Because data from previous studies [3, 22] included mostly loci on different chromosomes, we performed new imaging experiments on 16 loci distributed along the right arm of chromosome 4, the second longest arm after the rDNA-carrying right arm of chromosome 12 (Table S1). Our data set encompassed 36 loci on 13 out of the 16 chromosomes and included two loci on the right arm of chromosome 12. We first compared the predicted median angles  $\alpha$  to

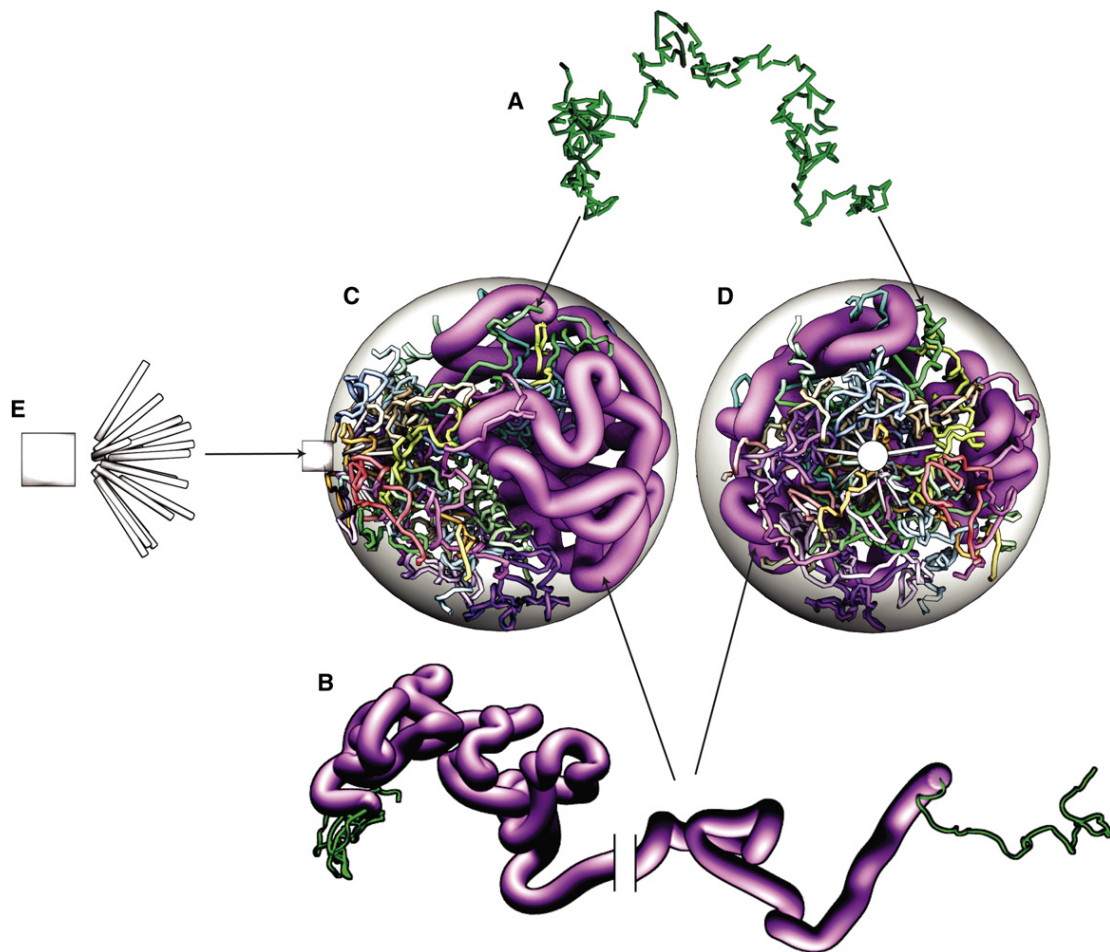


Figure 1. Computational Model of the Dynamic Interphase Yeast Nucleus

(A) Each chromosome is represented as a self-avoiding articulated chain of rigid segments.

(B) The heterochromatic rDNA locus on chromosome 12 is represented by thicker segments (pink, displayed using surface smoothing) than the rest of the DNA (green).

(C and D) A snapshot of the full model, showing each of the 16 chromosomes in a different color. The sphere represents the nuclear envelope. Two orthogonal views are shown: (C) view perpendicular to the axis joining the nuclear center to the SPB and (D) view along this axis and facing the SPB.

(E) The SPB (white knob) and the 16 microtubules, represented each by one rigid segment. See also [Figure S1](#) and the animated simulation in [Movie S1](#).

measurements. Overall, predicted angles correlated remarkably well with the measurements (Pearson's  $r = 0.87$ ;  $p < 10^{-11}$ ), despite an underestimation of  $\approx 18$  deg ([Figure 2M](#)). The model predicted that  $\alpha$  correlated with genomic distance to the centromere ( $d_{\text{CEN}}$ ) ([Figure S2A](#)). This correlation had previously been observed for telomeres [22], but our new measurements showed that it also extends to internal loci, as predicted by the model ([Figure S2B](#)). The model further predicted that loci on the rDNA carrying arm of chromosome 12 (*GAL2* and rDNA) have larger  $\alpha$  than other loci of similar  $d_{\text{CEN}}$  ([Figure S2A](#)). This was again borne out by measurements ([Figure S2B](#)).

We next considered the positioning of loci relative to each other. Because only the centromeres are tethered to a specific point, we reasoned that the most sensitive test of the model was for loci most distal from centromeres, i.e., the telomeres. We therefore compared our model's predictions against the 3D distances previously measured between 63 distinct pairs of subtelomeres [22], each of which contained one of three subtelomeres (6R, 10R, 4R), belonging to a short (122 Kb),

middle (310 Kb), and long (1,050 Kb) chromosome arm, respectively. The correlation between predicted and measured distances was good ( $r = 0.65$ ,  $p < 10^{-7}$ ), despite an underestimation by the model of  $\sim 150$  nm ([Figure 2N](#)). The predicted distances strongly depended on the lengths of the two chromosome arms, in a manner approximately similar to that observed in experiments [22] ([Figure S2C–S2H](#)). For short or medium arms, the two extremities of the same chromosome were predicted to be closer than for pairs of arms with similar lengths on different chromosomes ([Figures S2C and S2E](#)), also in agreement with measurements ([Figures S2D and S2F](#)) [22].

In comparison to the nominal model, the homopolymer and phantom models both failed to explain the measured angles  $\alpha$  ([Figures S2J and S2K](#)) but predicted distances between telomeres also agreed with measurements ([Figures S2N and S2O](#)). The microtubule-free model failed to explain both angles and distances ([Figures S2L and S2P](#)) ([Table 2](#)). Thus, only the nominal model quantitatively accounts for absolute locus positions and the relative positions of telomeres.

Table 1. Parameters of the Nominal Simulation

Parameter	Value
Number of chromosomes and microtubules	16
Chromosome segment length (Kuhn length), $L$	60 nm
Nuclear radius, $R_0$	1 $\mu$ m
Length of microtubules (+kinetochore), $L_{MT}$	380 nm
DNA compaction in chromatin, $C$	83 bp/nm
Diameter of euchromatin segments, $D_0$	20 nm
Diameter of rDNA segments, $D_{rDNA}$	200 nm
Number of rDNA segments, $N_{rDNA}$	150
Segments can cross each other	No

### Model Recapitulates Contact Patterns of Chromosomes and Chromosome Arms

Whereas imaging provides detailed information on positions of a limited number of loci, Hi-C data allowed us to test the model on a genome-wide scale [27]. We first analyzed contact frequencies at the genomic scales of entire chromosomes (230–1,500 Kb) (Figure 3) and chromosome arms (80–1,050 Kb) (Figure S3). We considered the probability for a contact to occur between any pair of chromosomes (including within the same chromosome). In absence of any nuclear organization (i.e., if all pairs of loci in the genome randomly contact each other with uniform probability), this probability is the product of the chromosome's genomic lengths (Figure 3A). In this case, only 14.0% of all contacts should be intrachromosomal (*cis*) (Figure 3E). In reality, 53.0% of experimentally detected contacts are *cis* (Figures 3C and 3E), indicating strong departure from a random collision scenario. Our nominal model, however, predicted 53.7% of *cis* contacts, in remarkable agreement with experiments (Figures 3B and 3E). The correlation between predicted and measured probabilities was very high (Pearson's  $r = 0.99$ , Spearman  $\rho = 0.96$ ,  $p < 10^{-50}$ ), indicating that the overall distribution of contacts among chromosomes was well recapitulated by our model (Figure 3D). The model was also in good agreement with contact probabilities measured between pairs of chromosome arms (Figures S3A–S3D). Here, three types of contacts exist: (1) between arms on different chromosomes, (2) within each arm, (3) between left and right arms of the same chromosomes. Unlike the random collision model, the nominal chromosome model predicted proportions of these three contact types in relatively good agreement with experimental data (Figure S3E).

However, we noted that, if taken separately, probabilities for interchromosomal contacts (*trans*) also agreed very well with the random collision scenario ( $r = 0.82$ ,  $\rho = 0.88$ ,  $p < 10^{-30}$ ) (Figures 3A and 3C); a similar result held for chromosome arms ( $r = 0.87$ ,  $\rho = 0.88$ ,  $p < 10^{-50}$ ) (Figures S3A and S3C). This suggested that, to first approximation, contacts

between different chromosomes or arms occur indiscriminately. To remove this effect, we next computed contact probabilities averaged over the product of the genomic lengths for each pair of chromosomes (or arms) (Figures 3F–3I; Figures S3F–S3I). Because random collisions predicted uniform average contact frequencies (Figure 3F; Figure S3F), any deviations from constancy reflect nonrandom nuclear organization. Overall, the average contact frequencies between chromosomes (*cis* + *trans*) predicted by the model still correlated very well with measurements ( $r = 0.97$ ,  $\rho = 0.80$ ,  $p < 10^{-50}$ ) (Figures 3G–3I). This was also true for contacts between arms ( $r = 0.91$ ,  $\rho = 0.86$ ,  $p < 10^{-50}$ ) (Figures S3G–S3I). Taken separately, the 16 predicted *cis* contact frequencies were also in very good agreement with measurements ( $r = 0.91$ ,  $\rho = 0.94$ ,  $p < 10^{-5}$ ) (Figure 3I). There was likewise good agreement for contacts within each of the 32 chromosome arms ( $r = 0.69$ ,  $\rho = 0.73$ ,  $p < 10^{-4}$ ) (Figure S3I). Most significantly, the predicted *trans* contact frequencies of chromosomes, taken separately, also correlated very well with the measurements ( $r = 0.84$ ,  $\rho = 0.72$ ,  $p < 10^{-30}$ ) (Figure 3I), as did predicted contacts between arms of different chromosomes ( $r = 0.86$ ,  $\rho = 0.82$ ,  $p < 10^{-50}$ ) (Figure S3I). The *trans* frequencies agreed well, although less so, with those predicted by the homopolymer model ( $r = 0.63$ ,  $\rho = 0.52$  for chromosomes,  $r = 0.79$ ,  $\rho = 0.73$  for arms), but not with those predicted by the microtubule-free model ( $r = -0.004$ ,  $p = 0.97$ ;  $\rho = -0.03$ ,  $p = 0.72$  for chromosomes,  $r = 0.04$ ,  $p = 0.42$ ;  $\rho = 0.12$ ,  $p = 0.01$  for arms) (Figures S2R, S2S, S2U, and S2V) (Table 2). Note that the phantom model by definition did not predict any contacts.

Thus, our nominal model largely recapitulates genome-wide contact frequencies at the scale of entire chromosomes or chromosome arms.

### Contact Frequencies at Subchromosomal Scales

Next, we considered contact frequencies at the maximum resolution afforded by our model, i.e., in 5 Kb bins. At this genomic resolution, the experimental contact matrix is very sparse, with on average only 1.4 contacts per bin, leading to strong counting noise (mean coefficient of variation [c.v.]  $1.4^{-1/2} = 0.84$ , compared to 131 contacts per bin and c.v.  $< 0.1$  for the predicted matrix). The full predicted contact matrix correlated only weakly with its experimental counterpart ( $r = 0.24$ ) (Figures S4A and S4B). However, trading off genomic resolution to reduce statistical noise strongly improved the correlation, which reached  $r = 0.85$  for bins of 75 Kb (Figure S4C).

Both predicted and measured intrachromosomal matrices are dominated by a strong diagonal, i.e., by contacts between genomically proximal loci, as expected for polymers with

Table 2. Simulations Considered and Assessment of Their Predictions

Model	Difference with nominal simulation	Agreement with experimental data			
		Subnuclear probability maps	Absolute positions ( $\sigma$ )	Relative telomere positions	Average interchromosomal contacts
Nominal	None	+	+	+	+
Phantom	Segments can cross	–	–	+	NA
Homopolymer	$D_{rDNA} = D_0$	–	–	+	+
Microtubule-free	No tethering to SPB	–	–	–	–

This table lists the nominal model and the three control models and indicates how their predictions agree with different measurements. Symbols + and – indicate good and poor agreement between model and measurements, respectively. See text and Figures 2, 3, S2, and S3 for quantifications. NA, not applicable (no contact frequencies are scored for the phantom model).

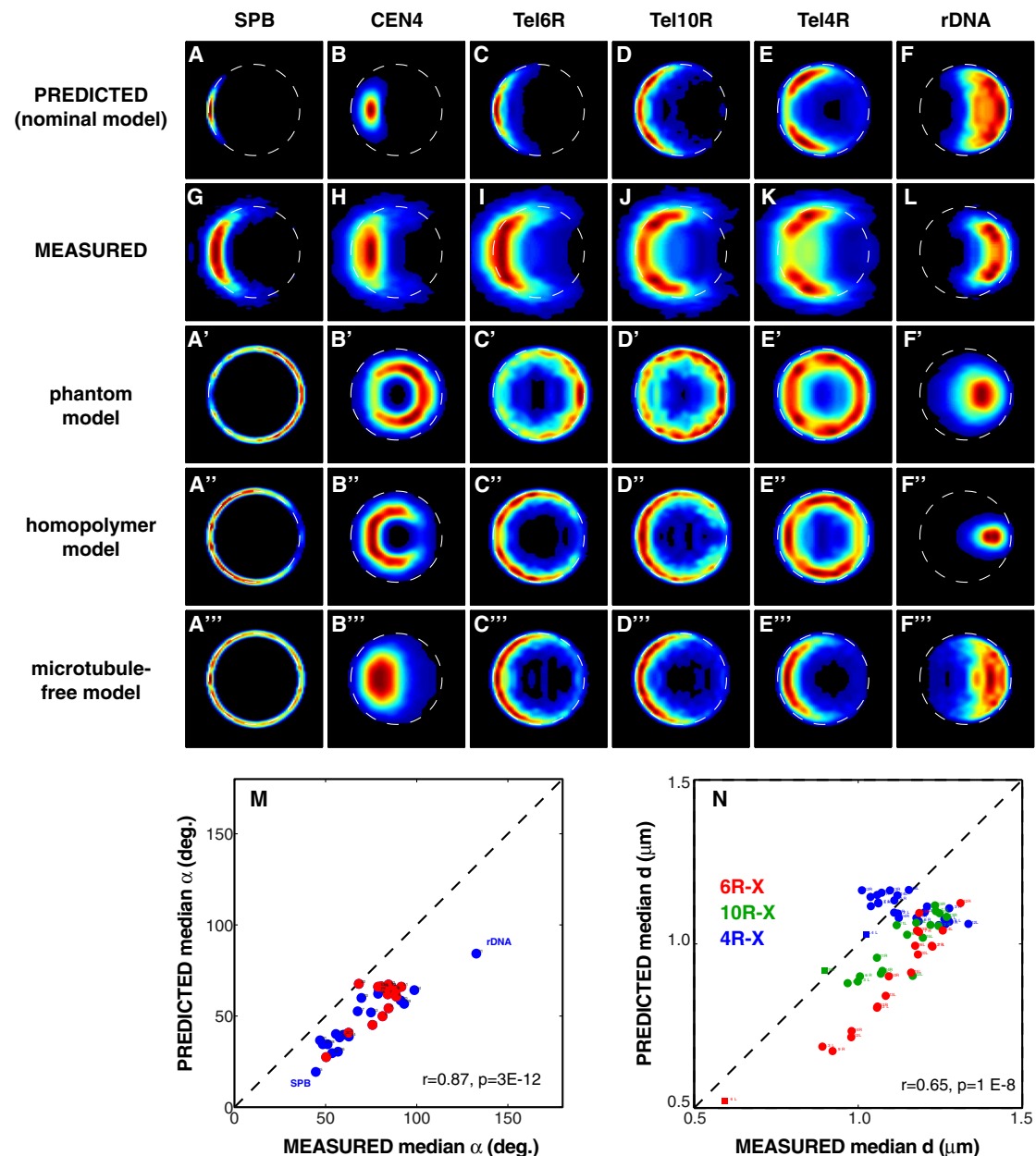


Figure 2. Model Recapitulates Nuclear Compartmentalization, Absolute, and Relative Locus Positions

(A–L, A'–F', A''–F'', and A'''–F''') Nuclear territories of selected genomic loci and the SPB, visualized as probability maps. (A)–(F) show territories predicted by the nominal model; (G)–(L) show territories measured from light microscopy data [3, 22]; (A')–(F') show territories predicted by the phantom model; and (A'')–(F'') show territories predicted by the homopolymer model. Probability maps were obtained and displayed as described in [3] and Figure S1J. Dashed circles have a fixed radius of 1  $\mu\text{m}$ . Territories shown in each column, from left to right, correspond to: the SPB (A, G, A', A'', A'''); the centromere of chromosome 4 (CEN4) (B, H, B', B'', B'''); the telomere on the 122 Kb long right arm of chromosome 6 (Tel6R) (C, I, C', C'', C'''); the telomere on the 310 Kb long right arm of chromosome 10 (Tel10R) (D, J, D', D'', D'''); the telomere on the 1,050 Kb long right arm of chromosome 4 (Tel4R) (E, K, E', E'', E'''), and an rDNA repeat (rDNA) (F, L, F', F'', F'''). Note the good agreement between observed territories and those predicted by the nominal model, but not the control models. (M) and (N) show quantitative comparisons of predicted (nominal model) and measured absolute and relative locus positions. (M) shows predicted versus measured median angle  $\alpha$  with respect to the nucleolar-nuclear axis. Each dot corresponds to a distinct locus as indicated by its number and Table S1. Red dots indicate loci along chromosome 4 that are new to this study; blue dots are data from previous work [3, 22]. (N) Predicted versus measured median distances between 63 pairs of telomeres. Each dot corresponds to a different pair of telomeres. Pairs containing the reference telomere 6R, 10R, or 4R have red, green and blue dots, respectively; the chromosome arm carrying the other telomere is indicated next to the dot (e.g., “3L” designates the left arm of chromosome 3). Squares indicate pairs of telomeres belonging to the same chromosome. See also Figure S2.

persistence lengths larger than the genomic resolution. We first analyzed how contact frequencies  $F$  between pairs of loci on the same chromosome fall off with genomic separation  $s$  (in Kb). Our model predicted that  $F(s)$  decayed approximately

as  $s^{-1.5}$  between  $\sim 5$  Kb and  $\sim 1,000$  Kb, and was roughly constant beyond (Figure 4A). This is as expected for equilibrated confined polymers [32, 33] and differs from the  $s^{-1}$  decay observed in mammalian cells and attributed to an



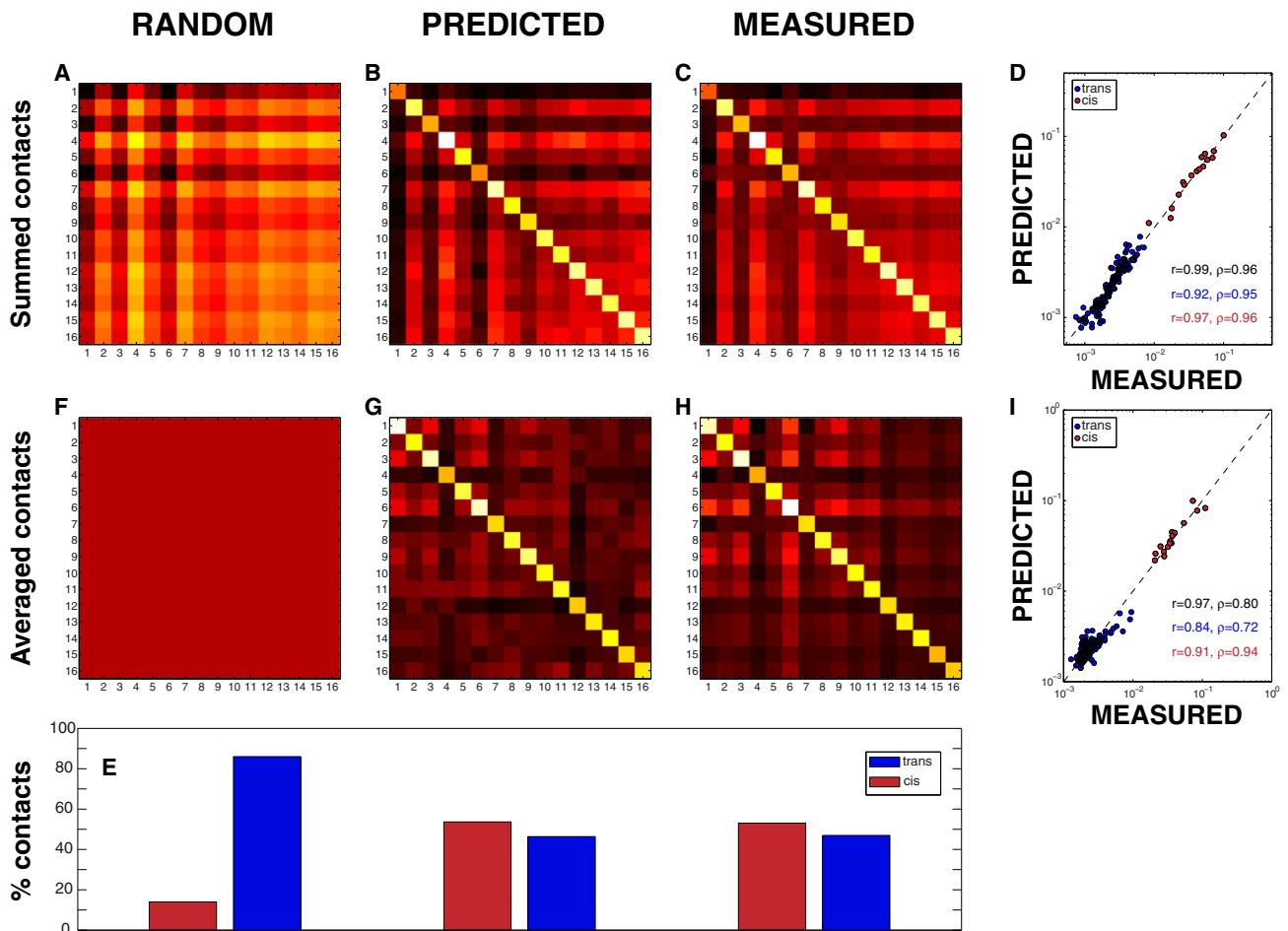


Figure 3. Model Recapitulates Patterns of Contacts among Chromosomes

(A–C, F–H) Contact frequency matrices for each pair of the 16 chromosomes. Chromosome numbers increase from top to bottom and from left to right. All matrices are displayed using the same logarithmic color scale, with dark colors indicating low probabilities, and bright colors high probabilities. (A) and (F) show contact matrices expected for random collisions; (B) and (G) show matrices predicted by the model; (C) and (H) show measured matrices [27]. (A)–(C) show contact probabilities integrated over chromosomes (corresponding to the probability that a contact occurs between any pair of chromosomes). (F–H) show averaged contact probabilities for each chromosome pair (probability per unit genomic length squared).

(D and I) Scatterplots of predicted versus measured contact frequencies. Each of the 136 dots corresponds to a distinct pair of chromosomes. Blue dots denote *trans* interactions, and red dots *cis* interactions. Pearson ( $r$ ) and Spearman ( $\rho$ ) correlation coefficients between predictions and measurements are indicated, for all contacts (black), *trans* contacts only (blue), or *cis* contacts only (red).

(E) Proportions of *cis* and *trans* contacts expected for random collisions, predicted by the model and measured. See Figure S3 for corresponding analyses of contacts between chromosome arms.

out-of-equilibrium fractal globule structure [9, 32]. The predicted  $s^{-1.5}$  decay and the plateau above  $\sim 1,000$  Kb agreed well with the measurements (Figure 4B).

The intrachromosomal contact frequency patterns predicted by the model were further characterized by a “negative cross” centered on the centromere, indicating a segregation of centromeric regions from the rest of the chromosome (Figure 4C; Figure S4A). This feature was also apparent in the experimental data (Figure 4D; Figure S4B). For chromosome 12, which carries the rDNA locus, the model predicted a striking dearth of contacts between the genomic regions on either side of the rDNA (Figure 4E). This was also observed in the Hi-C data [27] (Figure 4F).

Finally, we examined contacts between distinct chromosomes. The predicted contact patterns were mainly characterized by an enrichment between pericentromeric regions (Figure 4G; Figure S4A), in accordance with the Hi-C data

[27] (Figure 4H; Figure S4B). Our model also predicted a weaker negative cross pattern similar to that of intrachromosomal contacts, characterized by a depletion of contacts between centromeres and noncentromeric regions of other chromosomes (Figure 4G; Figure S4A). This feature was barely discernable in the much noisier experimental map (Figure 4H; Figure S4B), thus a confirmation of this prediction may require new Hi-C data.

Thus, our model is able to recapitulate the main observed patterns of contact frequencies at subchromosomal scales in both *cis* and *trans*.

#### Predicting Alterations of Nuclear Architecture

To test our model’s predictive power, we sought to simulate an alteration of nuclear organization amenable to experimental verification. Given the prominence of the nucleolar compartment and its origin in the high transcription of rDNA, we

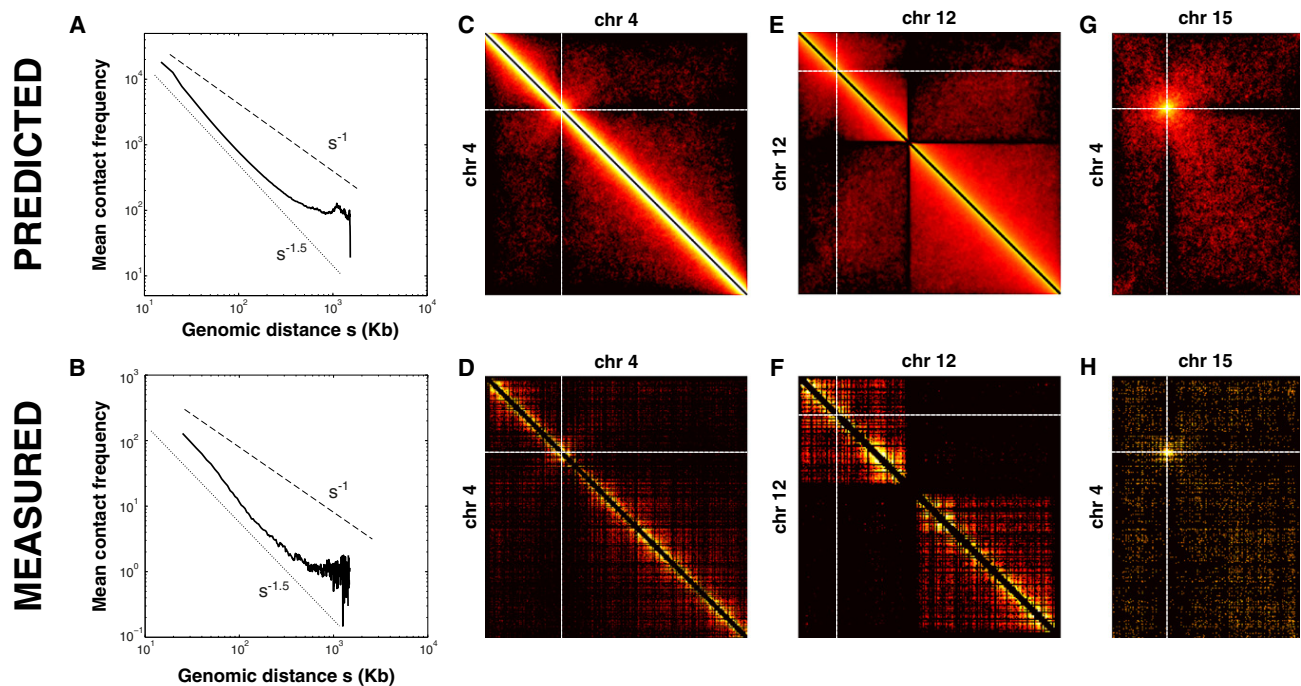


Figure 4. Model Recovers Subchromosomal Contact Patterns

(A, C, E, G) Model predictions. (B, D, F, H) Measurements. (A and B) The solid curve shows the average intrachromosomal contact frequency  $F$  as function of genomic separation  $s$ . The dotted and dashed lines indicate power laws  $s^{-1}$  and  $s^{-1.5}$ , respectively. Note the logarithmic scales. (C) and (D) show contact frequencies within chromosome 4; (E) and (F) show contacts within chromosome 12; (G) and (H) show contacts between chromosomes 4 and 15. All contact frequency matrices are shown using a logarithmic color scale. The dotted lines indicate the position of the centromeres. See Figure S4 for the entire genome-wide contact matrices.

analyzed how the model responded to a reduction in the transcriptional activity of this locus. This was previously achieved experimentally by treating cells with rapamycin, resulting in an approximate halving of the nucleolar volume [22]. To model this effect, we diminished the diameter of rDNA segments to  $D_{rDNA} = 140$  nm. The predicted rDNA territory had a reduced volume but still resembled a crescent abutting the nuclear envelope opposite the SPB, in agreement with observations [22]. We analyzed the effect of the reduced nucleolus on the absolute and relative positions of telomeres (Figure 5). The model predicted that rapamycin treatment causes telomeres to extend to larger  $\alpha$  than in absence of the drug and that the increase in  $\alpha$  was larger for longer chromosome arms; this prediction was in good agreement with measurements for six telomeres whose arm lengths ranged from 85 Kb (Tel6R) to 1,050 Kb (Tel4R) [22] (Figures 5A–5M). The predicted and measured changes in  $\alpha$  agreed well ( $r = 0.82$ ,  $p = 0.0011$ ) (Figure 5M). Similarly, the model predicted an increase in the distances of two telomere pairs (6R–3L and 6R–4R), in accordance with previous measurements [22] (Figure 5N). Thus, the model correctly predicted nuclear reorganization in response to a global change in rDNA transcription.

## Discussion

We have presented a new computational model of global chromosome arrangements in the yeast nucleus. Unlike other recent models [27, 31, 34], ours did not rely on measured contact frequencies or light microscopy data, but instead used first principles and assumed few parameters all based on preexisting literature. The model represented

chromosomes as confined (hetero)polymers undergoing passive Brownian dynamics, subject only to steric and topological constraints (Figure 1). We did not assume specific DNA binding factors, except implicitly through the assumption of modified chromatin properties at the rDNA locus and via the tethering of telomeres and centromeres. In contrast to other models designed for human cells [33, 35, 36], ours made no provisions for loops. We also did not assume any active motions. Despite this simplicity, the model accounts qualitatively and quantitatively for key aspects of yeast nuclear organization: the morphology and position of the nucleolus, its exclusion of all DNA except the rDNA, the nonrandom positions of genes and telomeres relative to nuclear landmarks and to each other (Figure 2), and patterns of contact frequencies across the genome, at the scale of chromosomes, chromosome arms, and at subchromosomal scales (Figures 3 and 4, Figures S3 and S4). Other models investigated here failed to explain all of these features simultaneously, although for some features the agreement was also good (Table 2; Figure S2). Our nominal model successfully predicted a change in nuclear organization in conditions of reduced rDNA gene expression (Figure 5). Nevertheless, it will be important to test its predictions in additional experiments that alter the mechanical constraints on chromosomes or chromosome structure or in other yeast species.

Our results have implications for the mechanisms governing nuclear organization in yeast. Although we cannot rule out specific interactions binding chromatin loci to each other or to nuclear landmarks besides centromeres and telomeres, our data suggest that such interactions are not required to explain global large-scale organization of the nucleus.

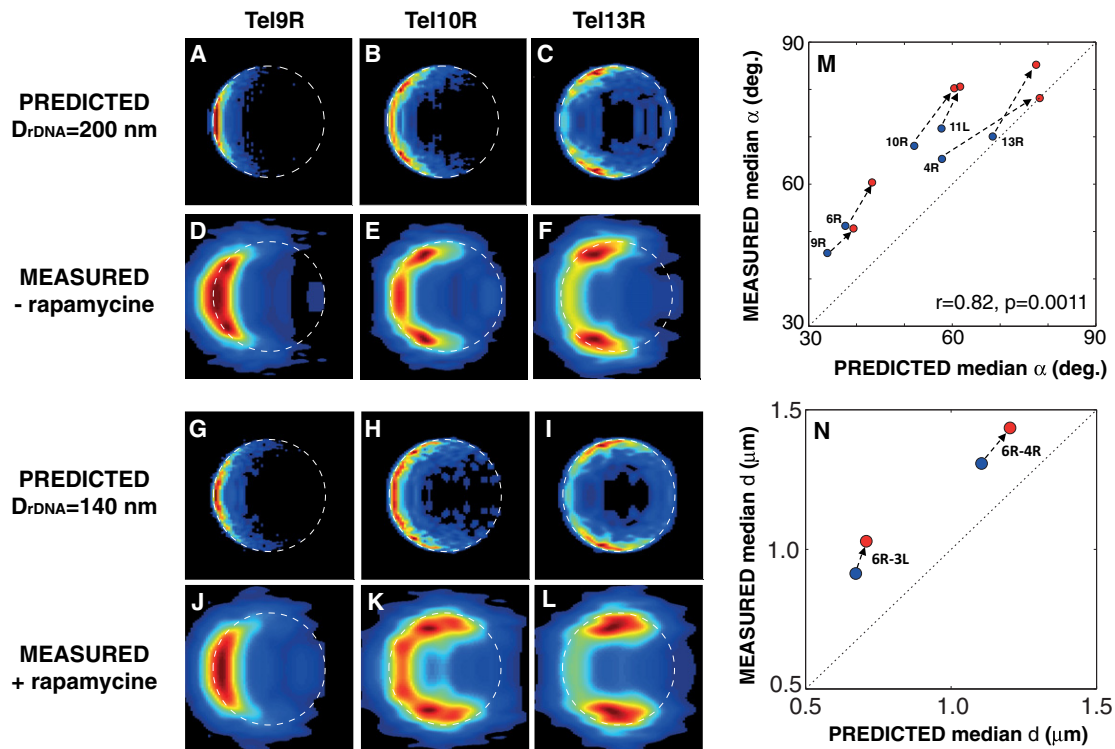


Figure 5. Model Predicts Alterations of Nuclear Organization under Reduced rDNA Transcription

(A–L) Intranuclear territories of three selected telomeres as predicted by the model (A–C and G–I) or measured [22] (D–F and J–L). (A)–(C) show predictions by the nominal model with  $D_{rDNA} = 200$  nm. (D)–(F) show experimental measurements in absence of rapamycin [22]. (G)–(I) show predictions by the model with  $D_{rDNA} = 140$  nm. (J)–(L) show experimental measurements in presence of rapamycin [22].

(M) Measured versus predicted median  $\alpha$  of six telomeres in each condition.

(N) Predicted versus measured median distance between pairs of telomeres (6R, 3L) and (6R, 4R) in each condition. In (M) and (N), blue dots indicate absence of rapamycin or  $D_{rDNA} = 200$ , and red dots indicate presence of rapamycin or  $D_{rDNA} = 140$  nm. Arrows indicate how values change upon treatment by rapamycin or reduction of  $D_{rDNA}$ . Note that the model correctly predicted an increase in  $\alpha$  and in distances upon rapamycin treatment. See also Figure S5.

Similarly, our data suggest that no energy-dependent, e.g., molecular motor-driven, dynamics is required. Instead, our data support the notion that yeast nuclear organization can be understood by the sole properties of confined and topologically constrained polymers, combined with the tethering of centromeres and telomeres. More specifically, chromosome arrangements can be explained by entropic repulsion of topologically constrained chromosome arms [37], whereas the segregation of rDNA into a distinct nucleolar compartment is consistent with entropic phase separation of block heteropolymers [38] [39]. This agrees with the notion that nuclear bodies arise by self-organization rather than by assembling onto preexisting scaffolds [40]. Our results are also in agreement with the conclusions of a different computational study (published during revision of this manuscript), which, however, did not self-consistently model the nucleolar compartmentalization [41].

Despite our model's success in recapitulating global features of nuclear organization, we acknowledge several limitations. First, the agreement with measurements is statistical and imperfect. Thus, other combinations of poorly known parameters (such as  $L$ ,  $C$ , and  $D_{rDNA}$  in Table 1) or other equally simple models may explain observations even better. Furthermore, the experimental data considered here have only moderate spatial and genomic resolutions, and are based on large cell populations. Accordingly, specific rather than generic factors may still govern chromatin folding at

smaller scales, or determine positions and dynamics of a subset of loci, or even exert global influence in a fraction of the cell population. Such specific factors may act during biological processes and underlie, e.g., the repositioning of inducible genes [3, 42] or the clustering of replication origins [43]. To address this, more systematic explorations of the simulation parameter space will likely help improve modeling, but new imaging or Hi-C data with better spatial and genomic resolution are essential. Notwithstanding, our model may already facilitate the identification of specific DNA interactions from experimental data by providing a means to predict the unspecific effects.

In higher eukaryotes, substantial evidence points to the role of DNA-specific factors in mediating interactions between distal loci and organizing chromosomes into distinct large-scale domains [16, 17, 44]. To account for such observations, computational models may need to include specific interactions, but it is likely that entropic effects and polymer properties also need to be considered to understand nuclear organization in these organisms [18, 33, 45].

At the functional level, our model is likely to have direct implications in understanding where chromosomal rearrangements (which require physical interactions between distal loci) are most likely to occur. In support of this, we analyzed 96 known chromosomal breakpoints and found that predicted contact frequencies at the observed breakpoints were significantly higher than expected if breakpoint loci were randomly located

in the genome (Table S2; Figure S5). More experimental data are needed to refine these results in various classes of events. Nevertheless, this first analysis suggests that the model might be used to identify chromosomal regions that may spontaneously interact to generate chromosomal rearrangements. In the future, the model may also be used to examine constraints on gene repositioning during gene expression, the interaction of mating type loci, the formation of replication foci, and other functionally important processes [42, 43, 46]. Overall, our model constitutes a framework to quantitatively understand and predict global features of nuclear organization and some of its functional consequences in this important model organism.

#### Experimental Procedures

##### Computational Model

The simulation used the open source physics engine (ODE), which allows one to model the dynamics of rigid bodies subject to external forces and constraints arising from collisions and joints between distinct bodies ([www.ode.org](http://www.ode.org)). We represented each of the 16 haploid yeast chromosomes as freely jointed chains of cylinders (Figure 1A). The motion of each segment was governed by the discretized Langevin equation, which includes a randomly oriented force representing thermal agitation, and a viscous friction term, in addition to forces needed to respect the constraints arising from collisions and joints. The main parameters used in the model are listed in Table 1. In the homopolymer model, we assumed that the physical properties of chromatin are constant across the genome. Thus, the chains were uniquely characterized by only three parameters: cylinder length (Kuhn length  $L$ , in nm), which determines chain rigidity, cylinder diameter ( $D_0$ , in nm), and the SPB as a cylinder of diameter 120 nm embedded in the spherical shell (Figures 1C and 1D). We modeled each of the 16 microtubule + kinetochore complexes as a single cylinder of length  $L_{MT} = 380$  nm and diameter 25 nm (Figure 1E) [50], of which one end remained in contact with the inner face of the SPB, whereas the other end was attached to the centromeric segment. To enforce telomere tethering to the nuclear envelope, we introduced a force that maintained the telomeric segments near the surface of the sphere without constraining their tangential motion. No other forces were introduced.

We initialized each simulation run with an artificial configuration in which chromosome chains were stretched out and parallel to each other, with their centromeres disposed at equal intervals along a circle of radius  $L_{MT}$  centered on the SPB, but in a random clockwise order (Figures S1A and S1B). To accommodate this initial configuration, we modeled the initial nuclear envelope as a long cylinder of radius  $R_0$  capped by two half spheres (Figure S1A). During the initial phase of the simulation ( $3.5 \times 10^5$  time steps), which was not used to compute model predictions, the length of this cylinder was progressively reduced (Figure S1C), until the nuclear envelope was a sphere of radius  $R_0$  embedding the SPB (Figure S1D). This initialization ensured that individual chains did not overlap and were not entangled with each other, as indeed the case for mitotic chromosomes. We plotted the time course of gyration radii and distances between chain extremities (telomeres) and verified that these quantities stabilized (i.e., fluctuated around a constant mean) before sampling the trajectories at large intervals (1 out of  $> 1,000$  time steps taken after  $10^6$  time steps) (Figures S1E and S2G). We also checked that the autocorrelation of the sampled data was negligible for nonzero lags (Figures S1F and S1H). Finally, to increase the sample size without excessive simulation time, we aggregated predicted loci and contact data from  $\sim 20$  parallel simulation runs.

##### Strain Constructions and Microscopy

For fluorescent tagging of individual loci along chromosome 4 (Table S1), we constructed 16 new strains bearing Tet operator sequences near the locus of interest and expressing the repressor protein fused to GFP. Labeling of

the nucleolus and nuclear pores and spinning disc fluorescence microscopy were performed as previously described [3, 22].

##### Breakpoint Analysis

We compiled a list of 96 breakpoints from the literature, which corresponded mainly to duplication and translocation events and included 38 homologous and 56 nonhomologous recombination events, 21 intrachromosomal and 75 interchromosomal events, and 49 events from haploid and 47 from diploid cells (Table S2). These data were selected from experiments imposing as little constraints as possible on the position of at least one of the two breakpoint loci.

##### Supplemental Information

Supplemental Information includes five figures, two tables, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2012.07.069>.

##### Acknowledgments

We acknowledge helpful discussions with O. Gadal, A. Grosberg, J.-M. Victor, J. Mozziconacci, A. Rosa, and A. Zidovska. We thank F. Mueller for comments on the manuscript. We thank the informatics services of Institut Pasteur for access to computing resources and P. Sorger for providing a yeast strain (CEN4). This work was funded by Institut Pasteur, Agence Nationale de la Recherche (grants ANR-09-PIRI-0024-1 and ANR-11-MONU-020-02), and Fondation pour la Recherche Médicale (Equipe FRM).

Received: April 21, 2012

Revised: July 9, 2012

Accepted: July 31, 2012

Published online: August 30, 2012

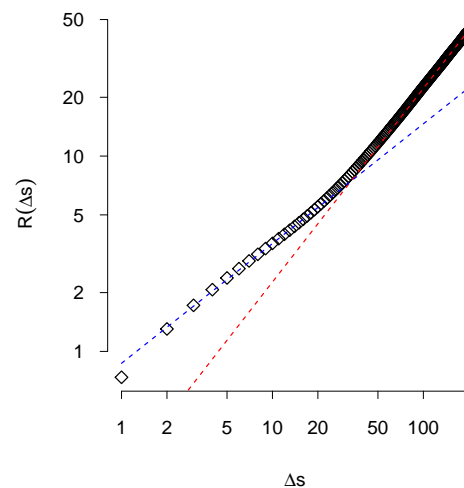
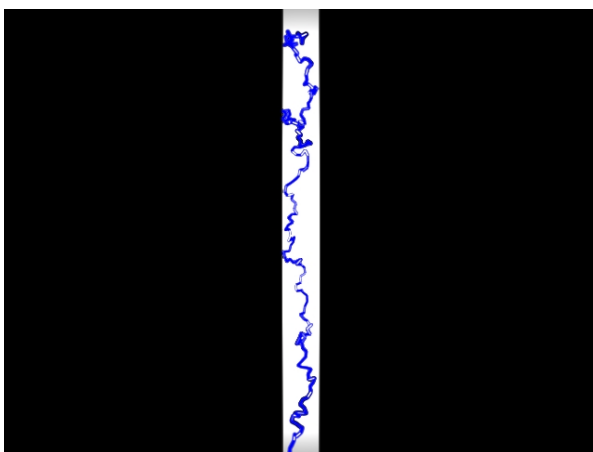
##### References

1. Misteli, T. (2005). Concepts in nuclear architecture. *Bioessays* 27, 477–487.
2. Cremer, T., and Cremer, M. (2010). Chromosome territories. *Cold Spring Harb. Perspect. Biol.* 2, a003889.
3. Berger, A.B., Cabal, G.G., Fabre, E., Duong, T., Buc, H., Nehrbass, U., Olivo-Marin, J.C., Gadal, O., and Zimmer, C. (2008). High-resolution statistical mapping reveals gene territories in live yeast. *Nat. Methods* 5, 1031–1037.
4. Heun, P., Laroche, T., Shimada, K., Furrer, P., and Gasser, S.M. (2001). Chromosome dynamics in the yeast interphase nucleus. *Science* 294, 2181–2186.
5. Taddei, A., Schober, H., and Gasser, S.M. (2010). The budding yeast nucleus. *Cold Spring Harb. Perspect. Biol.* 2, a000612.
6. Misteli, T. (2010). Higher-order genome organization in human disease. *Cold Spring Harb. Perspect. Biol.* 2, a000794.
7. Wijchers, P.J., and de Laat, W. (2011). Genome organization influences partner selection for chromosomal rearrangements. *Trends Genet.* 27, 63–71.
8. Hakim, O., Resch, W., Yamane, A., Klein, I., Kieffer-Kwon, K.R., Jankovic, M., Oliveira, T., Bothmer, A., Voss, T.C., Ansarah-Sobrinho, C., et al. (2012). DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. *Nature* 484, 69–74.
9. Zhang, Y., McCord, R.P., Ho, Y.J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 148, 908–921.
10. Chiarle, R., Zhang, Y., Frock, R.L., Lewis, S.M., Molinie, B., Ho, Y.J., Myers, D.R., Choi, V.W., Compagno, M., Malkin, D.J., et al. (2011). Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* 147, 107–119.
11. Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* 18, 1119–1130.
12. Nagai, S., Dubrana, K., Tsai-Pflugfelder, M., Davidson, M.B., Roberts, T.M., Brown, G.W., Varela, E., Hediger, F., Gasser, S.M., and Krogan, N.J. (2008). Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* 322, 597–602.

13. Miné-Hattab, J., and Rothstein, R. (2012). Increased chromosome mobility facilitates homology search during recombination. *Nat. Cell Biol.* *14*, 510–517.
14. Dion, V., Kalck, V., Horigome, C., Towbin, B.D., and Gasser, S.M. (2012). Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nat. Cell Biol.* *14*, 502–509.
15. Zimmer, C., and Fabre, E. (2011). Principles of chromosomal organization: lessons from yeast. *J. Cell Biol.* *192*, 723–733.
16. Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. *Cell* *137*, 1194–1211.
17. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* *148*, 458–472.
18. Rosa, A., and Everaers, R. (2008). Structure and dynamics of interphase chromosomes. *PLoS Comput. Biol.* *4*, e1000153.
19. Marenduzzo, D., Micheletti, C., and Cook, P.R. (2006). Entropy-driven genome organization. *Biophys. J.* *90*, 3712–3721.
20. Léger-Silvestre, I., Trumtel, S., Noaillac-Depeyre, J., and Gas, N. (1999). Functional compartmentalization of the nucleus in the budding yeast *Saccharomyces cerevisiae*. *Chromosoma* *108*, 103–113.
21. Bystricky, K., Laroche, T., van Houwe, G., Blaszczyk, M., and Gasser, S.M. (2005). Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization. *J. Cell Biol.* *168*, 375–387.
22. Thérizols, P., Duong, T., Dujon, B., Zimmer, C., and Fabre, E. (2010). Chromosome arm length and nuclear constraints determine the dynamic relationship of yeast subtelomeres. *Proc. Natl. Acad. Sci. USA* *107*, 2025–2030.
23. Marshall, W.F., Straight, A., Marko, J.F., Swedlow, J., Dernburg, A., Belmont, A., Murray, A.W., Agard, D.A., and Sedat, J.W. (1997). Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr. Biol.* *7*, 930–939.
24. Heun, P., Laroche, T., Raghuraman, M.K., and Gasser, S.M. (2001). The positioning and dynamics of origins of replication in the budding yeast nucleus. *J. Cell Biol.* *152*, 385–400.
25. Schober, H., Kalck, V., Vega-Palas, M.A., Van Houwe, G., Sage, D., Unser, M., Gartenberg, M.R., and Gasser, S.M. (2008). Controlled exchange of chromosomal arms reveals principles driving telomere interactions in yeast. *Genome Res.* *18*, 261–271.
26. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* *295*, 1306–1311.
27. Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y.J., Lee, C., Shendure, J., Fields, S., Blau, C.A., and Noble, W.S. (2010). A three-dimensional model of the yeast genome. *Nature* *465*, 363–367.
28. Rabl, C. (1885). über Zellteilung. *Morphologisches Jahrbuch* *10*, 214–330.
29. Hediger, F., Neumann, F.R., Van Houwe, G., Dubrana, K., and Gasser, S.M. (2002). Live imaging of telomeres: yKu and Sir proteins define redundant telomere-anchoring pathways in yeast. *Curr. Biol.* *12*, 2076–2089.
30. Haeusler, R.A., and Engelke, D.R. (2004). Genome organization in three dimensions: thinking outside the line. *Cell Cycle* *3*, 273–275.
31. Tokuda, N., Terada, T.P., and Sasai, M. (2012). Dynamical modeling of three-dimensional genome organization in interphase budding yeast. *Biophys. J.* *102*, 296–304.
32. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* *326*, 289–293.
33. Rosa, A., Becker, N.B., and Everaers, R. (2010). Looping probabilities in model interphase chromosomes. *Biophys. J.* *98*, 2410–2419.
34. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., and Chen, L. (2012). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* *30*, 90–98.
35. Munkel, C., Eils, R., Dietzel, S., Zink, D., Mehring, C., Wedemann, G., Cremer, T., and Langowski, J. (1999). Compartmentalization of interphase chromosomes observed in simulation and experiment. *J. Mol. Biol.* *285*, 1053–1065.
36. Mateos-Langerak, J., Bohn, M., de Leeuw, W., Giromus, O., Manders, E.M.M., Verschure, P.J., Indemans, M.H.G., Gierman, H.J., Heermann, D.W., van Driel, R., and Goetze, S. (2009). Spatially confined folding of chromatin in the interphase nucleus. *Proc. Natl. Acad. Sci. USA* *106*, 3812–3817.
37. De Gennes, P.G. (1980). Conformations of polymers attached to an interface. *Macromolecules* *13*, 1069–1075.
38. Leibler, L. (1980). Theory of microphase separation in block copolymers. *Macromolecules* *13*, 1602–1617.
39. Cook, P.R., and Marenduzzo, D. (2009). Entropic organization of interphase chromosomes. *J. Cell Biol.* *186*, 825–834.
40. Misteli, T. (2001). The concept of self-organization in cellular architecture. *J. Cell Biol.* *155*, 181–185.
41. Tjong, H., Gong, K., Chen, L., and Alber, F. (2012). Physical tethering and volume exclusion determine higher-order genome organization in budding yeast. *Genome Res.* *22*, 1295–1305.
42. Cabal, G.G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J.C., Hurt, E.C., and Nehrbass, U. (2006). SAGA interacting factors confine subdiffusion of transcribed genes to the nuclear envelope. *Nature* *441*, 770–773.
43. Knott, S.R., Peace, J.M., Ostrow, A.Z., Gan, Y., Rex, A.E., Viggiani, C.J., Tavaré, S., and Aparicio, O.M. (2012). Forkhead transcription factors establish origin timing and long-range clustering in *S. cerevisiae*. *Cell* *148*, 99–111.
44. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* *485*, 376–380.
45. Finan, K., Cook, P.R., and Marenduzzo, D. (2011). Non-specific (entropic) forces as major determinants of the structure of mammalian chromosomes. *Chromosome research: an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* *19*, 53–61.
46. Bressan, D.A., Vazquez, J., and Haber, J.E. (2004). Mating type-dependent constraints on the mobility of the left arm of yeast chromosome III. *J. Cell Biol.* *164*, 361–371.
47. Dekker, J. (2008). Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J. Biol. Chem.* *283*, 34532–34540.
48. Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* *436*, 138–141.
49. Cui, Y., and Denis, C.L. (2003). In vivo evidence that defects in the transcriptional elongation factors RPB2, TFIIS, and SPT5 enhance upstream poly(A) site utilization. *Mol. Cell. Biol.* *23*, 7887–7901.
50. O'Toole, E.T., Winey, M., and McIntosh, J.R. (1999). High-voltage electron tomography of spindle pole bodies and early mitotic spindles in the yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* *10*, 2017–2031.

## Conclusion et perspectives

En guise de conclusion de ce travail, je vais plutôt vous donner des perspectives. Il me semble que cela est plus agréable que de relire une seconde fois la même chose. Au moment où j'écris, je suis en train de réaliser des simulations de tri de chromosomes par application d'un flux. J'en suis au stade peu réjouissant, de l'introduction des interactions hydrodynamiques. Au préalable de cette modélisation, j'ai "étudié" avec ODE, un polymère greffé dans un tube. Les résultats de simulation sont en parfait accord avec la théorie (6.1).



(a) Vue schématique d'un polymère greffé dans un tube (b) Distance bout à bout  $R(\Delta s)$  du polymère greffé en fonction de la distance le long de la chaîne  $\Delta s$

FIGURE 6.1

Récemment, un article de B. Austin [10] a motivé encore plus ces modélisations. Notamment,

---

comprendre les effets induits par le rapport d'aspect (diamètre sur longueur des segments de Kuhn) sur le confinement de polymères. Il me semble, et je vais encore encenser les possibilités d'ODE, que la détection de collisions et la gestion en terme de joints des forces résultantes nous donnent un avantage certain sur les méthodes Monte-Carlo de simulations de polymères. Le taux de rejet des configurations est nul, même pour un confinement important et le mouvement est concerté entre toutes les contraintes présentes dans le système. Ce mouvement concerté est là encore un avantage sur les algorithmes "SHAKE" et "QSHAKE". D'autant plus que nous pouvons simuler des contraintes bien plus complexes que des liens rigides, et ce, grâce aux joints mécaniques. L'optimisation de la dynamique de Langevin locale, le thermostat global, se mêle à merveille avec les moteurs de jeux et permet d'échantillonner les configurations accessibles à un système bien plus rapidement qu'auparavant !

L'opportunité d'étudier les brosses de polymères, et en particulier, les noyaux de divers organismes est trop importante pour ne pas s'en saisir. Il serait notamment intéressant d'introduire dans les simulations de H. Wong [81] les deux types de fibre de chromatine (chez la drosophile il vient d'être montré qu'il existe cinq types ou couleurs de chromatine [85]). L'hétérochromatine est peu transcrite, et plus compacte que l'euchromatine. Que font des "spaghettinis" et des "spaghettonis" dans une casserole d'eau bouillante ? Est-ce que ces deux fibres de chromatines forment des domaines bien distincts par pression osmotique ? Mais aussi, ont elles les mêmes persistances ? Pour apporter un peu plus à l'architecture, organisation tri-dimensionnelle du noyau, il est nécessaire de réaliser des simulations sur une plage en longueur de persistance. Il est aussi possible d'injecter des contacts spécifiques entre des protéines et certaine partie de la fibre de chromatine. Un début d'étude peut se trouver chez M. Barbieri [86] où les auteurs s'intéressent aux conformations en fonction de la densité en protéines et du nombre de sites de fixation sur la fibre. Cette étude manque tout de même d'un certain réalisme puisque les chromosomes sont simulés sur un réseau cubique. Nous pouvons la reprendre avec ODE, et pour une fibre de chromatine en présence de protéines avec interactions spécifiques, obtenir les images (6.2). Une vidéo avec quatre fibres de chromatines greffées sur un substrat en présence de protéines avec interactions spécifiques est disponible à l'adresse <http://vimeo.com/52147790>. Cette simulation n'en est qu'à ses balbutiements mais elle pourrait apporter des résultats intéressants sur les protéines "polycomb" (protéines jouant un rôle dans le développement) du noyau de drosophile [87, 88, 89].

Il est aussi possible, avec ce même outil, d'étudier la cinétique de capture de conformation des chromosomes (idée et proposition chères à J. Mozziconacci) ou "3C" [90]. Nombre de biologistes à l'heure actuelle s'intéressent à l'aide des techniques de "HiC" (extension de la technique de "3C") aux interactions inter et intra-chromosomiques (chez la levure [91], chez la drosophile [87] et chez l'homme [92]). Pour ce faire, ils utilisent une molécule qui permet le "crosslink" de fragments de chromosomes interagissants. D'un point de vue cinétique, il serait intéressant d'avoir une simulation permettant d'étudier l'effet de la concentration en molécules de "crosslink" et le temps que les expérimentateurs octroient aux "crosslinkages". Les premières simulations (voir la figure (6.3)) que j'ai réalisées montrent que la cinétique du "crosslink" semble correspondre à la cinétique d'un "Crumple-Globule" [93].

Enfin, la variété des formes disponibles pour la détection de collisions, la possibilité de construire ses propres mailles, permettent d'imaginer des mélanges et assemblages de formes complexes <http://vimeo.com/36968098>. Pour chaque contact, il est possible de définir une friction, un coefficient de restitution et autres paramètres plus compliqués. Il serait peut être intéressant de tester les moteurs

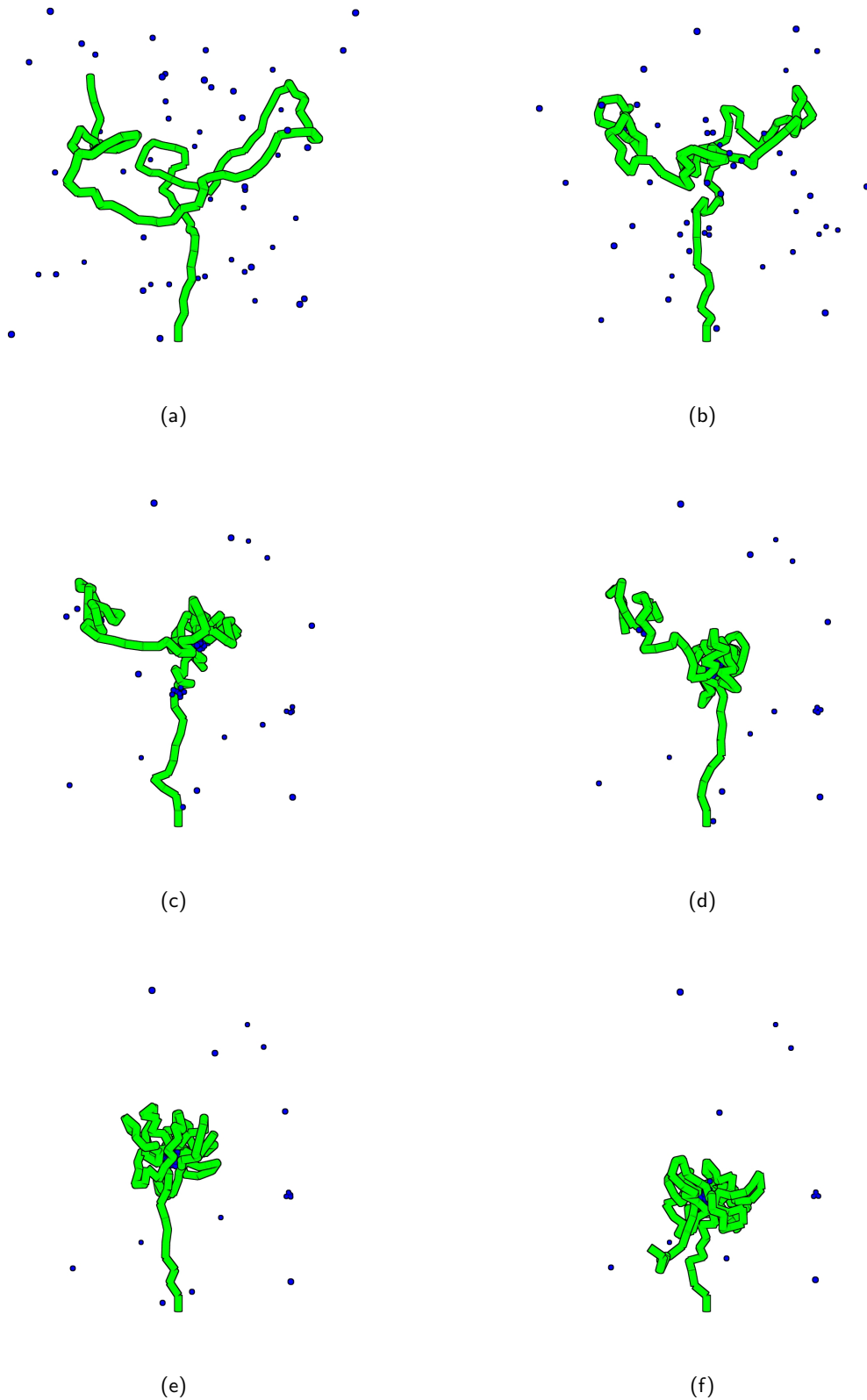
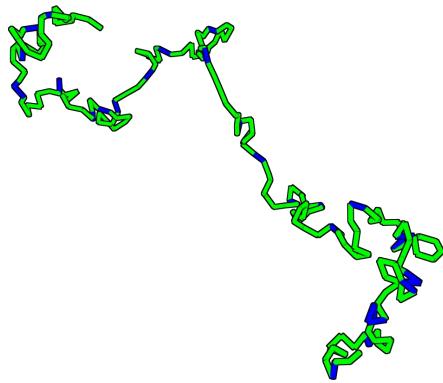
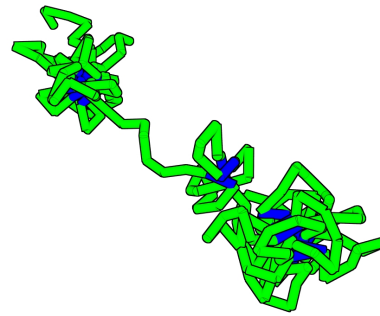


FIGURE 6.2: Ces images illustrent les différentes conformations d'une fibre de chromatine (en vert) avec des interactions spécifiques (ADN-protéines) réparties le long de la chaîne. Les protéines (en bleu) interagissent elles aussi de manière spécifique.





(a) Début de l'expérience de HiC.



(b) Fin de l'expérience de HiC.

FIGURE 6.3: Ces deux images illustrent le début et la fin d'une expérience de "HiC". Une fibre de chromatine représentée par 200 segments de longueur 60 nm et rayon 10 nm est simulée avec des segments pouvant interagir de manière attractive "crosslink" (en bleu) et des segments pouvant interagir seulement par volume exclu (en vert). Au tout début de la simulation le "crosslink" n'a pas eu le temps d'agir. À la fin de la simulation, les segments en bleu interagissent fortement entre eux. Il est clair que les interactions vues dans les expériences de "HiC" dépendent très clairement du temps d'observation.

physiques pour la simulation des matériaux granulaires. Qui plus est, nombre de moteurs de jeux sont parallélisables et peuvent, par exemple, simuler le stacking ou bien la destruction d'un mur de milliers de cubes avec un réalisme saisissant.

Ma dernière perspective est d'aller plus loin dans la compréhension de la mécanique du nucléosome. Je ne suis guère heureux de vous avoir présenté de simple preuves de principe sur la transition chirale et le dégrafage dynamique des "Super-Helix-Locations". Il est donc important d'achever le calcul de structure et de pouvoir en déduire les énergies d'interactions ADN-dimère et dimère-dimère. La modélisation devrait alors reproduire les résultats expérimentaux sans que nous n'ayons à intervenir pour dégrafer les "Docking Domains" ou les "Super-Helix-Locations".

---

## Algorithmes utilisés

---

### A.1 Schémas d'intégration

En guise d'exemple je veux résoudre le système d'une particule de masse  $m$  soumise à une force extérieure  $\mathbf{f}$ . Le système à résoudre est le suivant :

$$\begin{aligned}\mathbf{v} &= \dot{\mathbf{r}} \\ m\dot{\mathbf{v}} &= \mathbf{f}\end{aligned}$$

Pour ce faire, je peux choisir parmi cinq schémas d'intégration d'Euler avec une discrétisation temporelle  $\Delta t$  :

1. Schéma explicite

$$\begin{aligned}\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}(t) \\ \mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \frac{\Delta t}{m} \mathbf{f}(t)\end{aligned}$$

2. Schéma semi-implicite

$$\begin{aligned}\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}(t + \Delta t) \\ \mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \frac{\Delta t}{m} \mathbf{f}(t)\end{aligned}$$

3. Schéma implicite

$$\begin{aligned}\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}(t + \Delta t) \\ \mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \frac{\Delta t}{m} \mathbf{f}(t + \Delta t)\end{aligned}$$

4. Schéma de Verlet

$$\begin{aligned} \mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t) &= 2\mathbf{r}(t) + \frac{\Delta t^2}{m} \mathbf{f}(t) + \mathcal{O}(\Delta t^4) \\ \mathbf{v}(t) &= \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2) \end{aligned}$$

Dans ce schéma d'intégration, le calcul des positions ne dépend pas des vitesses ; ces dernières peuvent toutefois se calculer.

5. Schéma de Leapfrog

$$\begin{aligned} \mathbf{v}(t + \Delta t) - \mathbf{v}(t - \Delta t) &= \frac{\Delta t}{m} \mathbf{f}(t) \\ \mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}\left(t + \frac{\Delta t}{2}\right) \end{aligned}$$

Le schéma d'Euler implicite est plus difficile à mettre en œuvre puisqu'il faut connaître la force  $\mathbf{f}$  au temps  $t + \Delta t$ . Mais par chance, le problème d'un ressort unidimensionnel de raideur  $k$ , de masse  $m$  et de pulsation  $\omega = \sqrt{k/m}$  se résoud facilement. Le système à résoudre est le suivant :

$$\begin{aligned} v &= \dot{x} \\ \dot{v} &= -\omega^2 x \end{aligned}$$

L'énergie du système est conservée et égale à  $\frac{2E}{m} = v^2 + \omega^2 x^2$ .

1. Schéma explicite

$$\begin{aligned} x(t + \Delta t) &= x(t) + \Delta t v(t) \\ v(t + \Delta t) &= v(t) - \Delta t \omega^2 x(t) \\ E(t + \Delta t) &= E(t) (1 + \omega^2 \Delta t^2) \end{aligned}$$

L'énergie augmente à chaque pas de temps. Pour un temps de simulation de  $N\Delta t$ , au temps  $t = N\Delta t$  l'énergie sera  $E(N\Delta t) = E(0) (1 + \omega^2 \Delta t^2)^N$ . En supposant que  $\omega^2 \Delta t^2 \ll 1$ , nous avons  $E(N\Delta t) \simeq E(0) (1 + N\omega^2 \Delta t^2) + \mathcal{O}(2)$ . Pour ne pas s'éloigner de la valeur de l'énergie initiale, un choix correct pour le pas de temps  $\Delta t$ , qui préserve l'énergie lors de l'intégration du système, serait  $\Delta t = \frac{1}{N\omega}$ .

2. Schéma semi-implicite

$$\begin{aligned} x(t + \Delta t) &= x(t) + \Delta t v(t + \Delta t) \\ v(t + \Delta t) &= v(t) - \Delta t \omega^2 x(t) \end{aligned}$$

soit :

$$\begin{aligned} x(t + \Delta t) &= (1 - \Delta t^2 \omega^2) x(t) + \Delta t v(t) \\ v(t + \Delta t) &= v(t) - \Delta t \omega^2 x(t) \\ \frac{2}{m} E(t + \Delta t) &= \frac{2}{m} E(t) + (\Delta t^2 \omega^3 x(t) - \Delta t \omega v(t))^2 - \Delta t^2 \omega^4 x(t)^2 \end{aligned}$$

L'énergie oscille autour de la valeur initiale  $E(0)$ .

## 3. Schéma implicite

$$\begin{aligned}x(t + \Delta t) &= x(t) + \Delta t v(t + \Delta t) \\v(t + \Delta t) &= v(t) - \Delta t \omega^2 x(t + \Delta t)\end{aligned}$$

soit :

$$\begin{aligned}x(t + \Delta t) &= \frac{x(t) + \Delta t v(t)}{1 + \Delta t^2 \omega^2} \\v(t + \Delta t) &= \frac{v(t) - \Delta t \omega^2 x(t)}{1 + \omega^2 \Delta t^2} \\E(t + \Delta t) &= \frac{E(t)}{1 + \Delta t^2 \omega^2}\end{aligned}$$

Ce qui nous donne l'énergie au temps  $t + \Delta t$  en fonction de l'énergie au temps  $t$  ; il est clair que l'énergie du système décroît jusqu'à tendre vers 0, en particulier  $E(N\Delta t) = \frac{E(0)}{(1 + \omega^2 \Delta t^2)^N}$ . En supposant que  $\omega^2 \Delta t^2 \ll 1$ , nous avons  $E(N\Delta t) \simeq E(0) (1 - N\omega^2 \Delta t^2) + \mathcal{O}(2)$ . Pour ne pas s'éloigner de la valeur de l'énergie initiale, un choix correct pour le pas de temps  $\Delta t$ , qui préserve l'énergie lors de l'intégration du système, serait  $\Delta t = \frac{1}{N\omega}$ .

Vous pouvez trouver dans la thèse de C. Lacoursière [24] une étude plus détaillée de la stabilité de ces algorithmes. Pour pallier les divers problèmes de ces intégrateurs il existe les intégrateurs géométriques utilisés pour les corps rigides notamment dans le mouvement des planètes [94, 95].

## A.2 Algorithme de Metropolis

### A.2.1 Moyenne d'une observable

Nous souhaitons étudier un système physique par la simulation, plus particulièrement, ses propriétés statistiques. Si nous nous donnons une grandeur  $X$  associée au système et mesurable, la moyenne thermique est donnée par (A.1).

$$\langle X \rangle = \frac{1}{Z} \sum_i X_i \exp(-\beta E_i) \quad (\text{A.1})$$

Le terme  $\beta$  représente l'inverse de l'unité d'agitation thermique  $k_B T$  et  $E_i$  représente l'énergie de la configuration  $i$ . Nous avons introduit la fonction de partition  $Z$  comme la somme discrète ou continue de configurations (A.2). L'indice  $i$  parcourt l'ensemble des configurations accessible au système.

$$Z = \sum_i \exp(-\beta E_i) \quad (\text{A.2})$$

Il est maintenant utile d'introduire la probabilité que le système à l'équilibre soit dans la configuration  $i$  comme (A.3).

$$P_{eq,i} = \frac{1}{Z} \exp(-\beta E_i) \quad (\text{A.3})$$

Bien entendu, la somme des probabilités associées à chaque configuration est égale à l'unité. Nous pouvons considérer le facteur  $1/Z$  comme un facteur de normalisation. Nous pouvons alors écrire la moyenne thermique à l'équilibre d'une grandeur  $X$  comme (A.4).

$$\langle X \rangle_{th} = \sum_i P_{eq,i} X_i \quad (\text{A.4})$$

Cependant, pour estimer  $\langle X \rangle_{th}$ , nous devons parfois avoir recourt à des simulations qui nous donnent (A.5) où  $N$  représente le nombre de mesures de notre calcul numérique. Dans cette expression, les variables  $X_i$  sont des réalisations de la variable aléatoire  $X$  qui suit une loi  $P(X)$ . Ces valeurs  $X_i$  apparaissent avec une probabilité  $P(X = X_i)$ . Néanmoins,  $\langle X \rangle_{sim}$  ou moyenne empirique représente un estimateur non biaisé de l'espérance  $\langle X \rangle_{th}$ . Ainsi,  $\langle X \rangle_{sim}$  peut être considérée comme une variable aléatoire dont l'espérance vaut  $\langle X \rangle_{th}$ .

$$\langle X \rangle_{sim} \approx \frac{1}{N} \sum_i^N X_i \quad (\text{A.5})$$

## A.2.2 Equation maîtresse

Pour arriver à un algorithme qui nous permettra d'étudier notre système physique, intéressons nous à l'équation d'évolution du système. En particulier, que vaut la probabilité  $P(i, t + dt)$  que le système se trouve dans la configuration  $i$  à l'instant  $t + dt$ ? A l'instant  $t + dt$  la probabilité d'être dans l'état  $i$  est égale à la probabilité  $P(i, t)$  d'être dans la configuration  $i$  à l'instant  $t$  accrue par la possibilité que le système passe d'une quelconque configuration accessible  $j$  vers la configuration  $i$  et diminuée par la probabilité que le système passe de la configuration  $i$  vers une configuration  $j$  du système. L'équation de bilan peut alors s'écrire comme (A.6) où les  $W$  représentent les taux de transition d'une configuration vers une autre et ont pour unité l'inverse d'un temps.

$$P(i, t + dt) = P(i, t) + \sum_j \{W(j \rightarrow i)P(j, t) - W(i \rightarrow j)P(i, t)\} dt \quad (\text{A.6})$$

Cette description est purement stochastique (phénomène aléatoire et dépendant du temps) et markovienne puisque la probabilité que le système a d'être dans la configuration  $i$  à l'instant  $t + dt$  ne dépend que du présent  $t$ . La probabilité conditionnelle, autrement dit, la probabilité de passer de la configuration  $i$  à la configuration  $j$  est donnée par  $W(i \rightarrow j)dt$ . Pour aller plus loin, nous nous intéressons aux probabilités d'équilibre telles que  $dP_{eq}/dt = 0$  soit  $P_{eq}(i, t + dt) = P_{eq}(i, t)$ . Cette condition nous permet de d'écrire (A.6) sous la nouvelle forme (A.7).

$$\sum_j \{W(j \rightarrow i)P_{eq}(j) - W(i \rightarrow j)P_{eq}(i)\} = 0 \quad (\text{A.7})$$

La solution la plus simple et suffisante de ce système est donnée par (A.8) et est connue, sous le nom de bilan détaillé. Cette équation nous dit qu'en moyenne, le système a autant de chances d'évoluer de la configuration  $i$  vers la configuration  $j$  que d'évoluer de la configuration  $j$  vers la configuration  $i$ .

$$\begin{aligned} W(j \rightarrow i)P_{eq}(j) &= W(i \rightarrow j)P_{eq}(i) \\ \frac{W(i \rightarrow j)}{W(j \rightarrow i)} &= \exp(-\beta(E_j - E_i)) \end{aligned} \quad (\text{A.8})$$

A présent, il se dessine clairement la forme d'un algorithme que nous nommerons *algorithme de Metropolis* nous permettant d'échantillonner les configurations accessibles à un système.

1. A partir d'une configuration  $i$ , nous générons aléatoirement une configuration  $j$  avec une probabilité  $\gamma(i \rightarrow j)$ . Nous imposons en plus que  $\gamma(i \rightarrow j) = \gamma(j \rightarrow i)$ .
2. La configuration  $j$  est acceptée avec une probabilité  $\Pi(i \rightarrow j)$ .
3. Soit  $W(i \rightarrow j) = \gamma(i \rightarrow j)\Pi(i \rightarrow j)$ .

Nous obtenons du troisième point l'équation (A.9).

$$\frac{\Pi(i \rightarrow j)}{\Pi(j \rightarrow i)} = \exp(-\beta(E_j - E_i)) \quad (\text{A.9})$$

Nous choisissons alors (A.10) qui satisfait bien à l'équation (A.9). Cette équation nous dit que le système évolue vers une configuration d'énergie plus basse avec une probabilité égale à l'unité.

$$\Pi(i \rightarrow j) = \min \{1, \exp(-\beta(E_j - E_i))\} \quad (\text{A.10})$$

L'algorithme de Metropolis se résume donc à :

1. A partir d'une configuration  $i$ , nous générons aléatoirement une configuration  $j$  avec une probabilité  $\gamma(i \rightarrow j)$ .
2. Nous tirons de manière aléatoire un nombre  $r$  selon une loi uniforme  $U(0, 1)$ .
3. Si  $r < \min \{1, \exp(-\beta(E_j - E_i))\}$  la nouvelle configuration est acceptée sinon elle est rejetée.
4. Ces trois étapes sont répétées  $n$ -fois.

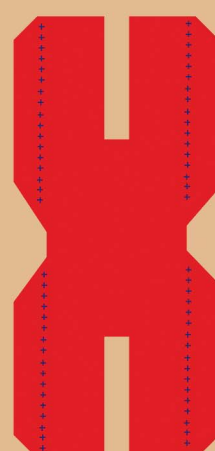
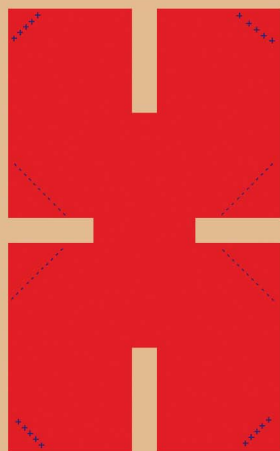
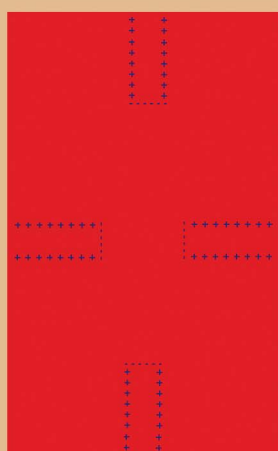
L'algorithme que nous venons de décrire n'est valable que pour échantillonner un système à l'équilibre : les moyennes  $\langle X \rangle$  ne veulent rien dire si nous ne vérifions pas au préalable que la grandeur  $X$  fluctue autour d'une valeur moyenne. En guise de vérification, il est utile de tracer le graphique  $X(t)$ .



# Soft Matter

www.rsc.org/softmatter

Volume 8 | Number 36 | 28 September 2012 | Pages 9255–9488



Themed issue: Polyelectrolytes in Soft Matter and Biology

ISSN 1744-683X

RSC Publishing

**REVIEW ARTICLE**

Maria Barbi *et al.*

Electrostatics of DNA compaction in viruses, bacteria and eukaryotes: functional insights and evolutionary perspective



1744-683X(2012)8:36;1-5



Cite this: *Soft Matter*, 2012, **8**, 9285

www.rsc.org/softmatter

REVIEW

## Electrostatics of DNA compaction in viruses, bacteria and eukaryotes: functional insights and evolutionary perspective

Pascal Carrivain,<sup>†ab</sup> Axel Cournac,<sup>†ab</sup> Christophe Lavelle,<sup>bc</sup> Annick Lesne,<sup>abd</sup> Julien Mozziconacci,<sup>ab</sup> Fabien Paillusson,<sup>e</sup> Laurence Signon,<sup>a</sup> Jean-Marc Victor<sup>ab</sup> and Maria Barbi<sup>\*ab</sup>

Received 4th April 2012, Accepted 29th May 2012

DOI: 10.1039/c2sm25789k

The molecular support of genetic information, DNA, has to be packaged and organized inside the tiny volume of nuclei, cells, or virus capsids, in an ordered and dynamical way. Evolution has favored different strategies in different kingdoms: a liquid crystal ordering mechanism prevails in viruses; strong entanglement due to supercoiling is the main compacting strategy in bacteria; the building of a hierarchical and tunable architecture mediated by DNA–protein interaction constitutes the main compacting mechanism in archaea and eukaryotes. The interplay between these different strategies is however much more complex than at first sight and all these strategies can be used in a synergistic way. All of them have to deal with the same elementary physical constraint which hinders compaction: electrostatic repulsion due to the high line charge density of DNA. In this review, we will show how this apparent weakness can be turned into a strength in order to compact this long molecule in a functional way.

<sup>a</sup>CNRS LPTMC UMR 7600, Université Pierre et Marie Curie-Paris 6, 4 place Jussieu, 75252 Paris Cedex 05, France. E-mail: barbi@lptmc.jussieu.fr; Fax: +33 144275100; Tel: +33 144272666

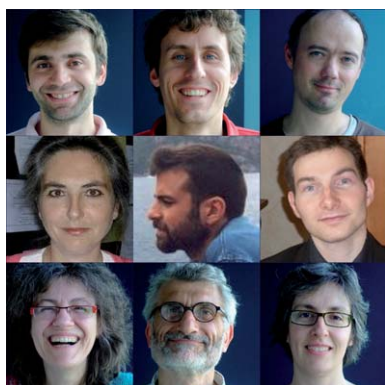
<sup>b</sup>CNRS GDR 3536, Université Pierre et Marie Curie-Paris 6, 4 place Jussieu, 75252 Paris Cedex 05, France

<sup>c</sup>National Museum of Natural History, Genome Dynamics and Regulation Department, CNRS UMR7196/INSERM U565, 43 rue Cuvier, 75005 Paris, France. E-mail: lavelle@mnhn.fr

<sup>d</sup>Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440 Bures-sur-Yvette, France

<sup>e</sup>Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK

† These authors contributed equally to this work.



Pascal Carrivain, Axel Cournac, Christophe Lavelle, Annick Lesne, Julien Mozziconacci, Fabien Paillusson, Laurence Signon, Jean-Marc Victor, Maria Barbi

This review is a collective work of the “Multiscale Modeling of Living Matter” group (<http://www.lptl.jussieu.fr/user/lpipol/M3V/m3v.html>). The group activity is devoted to the physical modeling of chromosome architecture and functional dynamics, from the atomic scale up to the whole nucleus, with a special focus on protein–DNA interaction, nucleosome plasticity and chromatin fiber structures. The group was founded in 2000 by Annick Lesne and Jean-Marc Victor, both CNRS Directors of Research in physics. Maria Barbi is a DNA physicist; she joined the group in 2001 as an associate professor. Julien Mozziconacci was a PhD student in the group from 2001 to 2004; he joined the group as an associate professor in 2009. Christophe Lavelle was an assistant lecturer in the group from 2002 to 2003. He was appointed CNRS Research Associate in 2008 at the “Museum National d’Histoire Naturelle” in Paris. Laurence Signon is a DNA repair biologist, CNRS Research Associate since 2000; she joined the group in 2011 for a two year period to learn physical modeling techniques. Fabien Paillusson was a PhD student in the group from 2007 to 2010; he is now a postdoctoral researcher in Daan Frenkel’s group in Cambridge. Axel Cournac joined the group as a postdoctoral researcher in 2011. Pascal Carrivain is a PhD student in the group since 2009.

## 1 Introduction

Genetic information is encoded essentially in the DNA double helix. The genetic material has to be packaged inside the tiny volume of the cell, and specific sequences have to be retrieved at will for physiological purposes. The genetic material is therefore embedded in an orderly and dynamically retrievable architecture. Knowledge of the physical principles and the molecular machinery that govern the 3D organization of this structure and its regulation is key to understanding the relationship between genome structure and function. DNA is one of the most highly charged polyelectrolytes, and also one of the stiffest polymers. How both these physical constraints have been overcome and actually turned to good account during evolution of living organisms is the focus of this review.

If DNA is the genetic support for all living organisms, the way it is organized in the available space may differ considerably from viruses to prokaryotes (organisms that lack cell nucleus, generally unicellular, as bacteria), to archaea (again nucleus-lacking unicellular organisms, but having an independent evolutionary history), to eukaryotes (organisms having a cell nucleus, including multicellular organisms). Different organisms achieve different DNA compaction levels. To quantify this compaction level, one can define two different quantities (see Table 1). As shown in Table 1, the various biological compaction strategies implemented among the different organisms cannot be reduced to the coil–globule transition of a polyelectrolyte but involve instead architectural proteins that constitute the very folds of a “molecular origami”.

The paper is organized as follows: in Section 2, we present the molecular actors (DNA, ions, water, proteins) and describe the elementary electrostatics features involved in their interaction. In Section 3, we discuss how electrostatics may drive three fundamental mechanisms involved in genome condensation, namely DNA supercoiling, the building of an architecture by DNA binding proteins (BPs) and liquid crystal ordering of DNA and proteins. The way how different condensation mechanisms are implemented *in vivo* in different kingdoms (viruses, prokaryotes, eukaryotes and archaea) is then discussed in Section 4. Section 5 is devoted to deeper insights into the functional tuning allowing regulation of DNA compaction, again focusing on the role of electrostatics. Finally, in Section 6 we reexamine the overall scenario from an evolutionary perspective, and draw a few conclusions.

## 2 DNA, proteins, ions and their electrostatic interactions

### 2.1 DNA in solution and Manning condensation

Double stranded DNA (dsDNA) is a highly charged polyelectrolyte of  $-2e$  per base pair, each elementary charge

belonging to one backbone phosphate. It is also one of the stiffest polymers in nature. The high charge, together with the cylinder-like shape of DNA, induces the formation of a condensed counterion atmosphere around the dsDNA “cylinder”. In the 70s, Manning introduced a theoretical approach to calculate the condensed counterion density,<sup>131</sup> and showed that condensation depends on the value of the Manning parameter  $\xi = q\lambda l_B$ , where  $q$  is the counterion valency,  $\lambda$  the effective line charge density of the rod and  $l_B \approx 7 \text{ \AA}$  the Bjerrum length in normal temperature conditions (the Bjerrum length is defined as the separation at which the electrostatic interaction between two elementary charges is comparable in magnitude to the thermal energy scale,  $k_B T$ ). If  $\xi < 1$ , the bare charge on DNA is not enough to overcome the entropic cost of binding for counterions and they escape from the polyelectrolyte whereas if  $\xi > 1$  a counterion condensation is predicted and the rod line charge density is expected to be renormalized as  $\lambda_{\text{eff}} = (q/l_B)^{-1}$ . In the case of a monovalent salt, Manning theory predicts that, due to the high DNA charge, the number of condensed monovalent counterions is 0.76 per DNA phosphate charge, independent of the added salt concentration. The counterion density around DNA is therefore of about 3 positive charges every 2 base pairs, this resulting in a linear charge of  $-e$  every 0.68 nm or, alternatively, in an average surface charge  $\sigma_{\text{DNA}} = -e \text{ nm}^{-2}$ . Note that temperature may vary considerably from one organism to another. This will change in turn the Bjerrum length and the Manning parameter  $\xi$  accordingly. Different strategies will therefore be used in different organisms to deal with DNA electrostatics.

As mentioned above, the Manning parameter  $\xi$  grows linearly with the valency of the counterions so that multivalent ions more easily condense on DNA. Despite the robustness of its general features, it has been stressed that Manning condensation depends on the chemical nature of the counterions, even for monovalent ions.<sup>30,85,200</sup>

This is likely due to the way these ions interact with the solvent. Small monovalent ions of high charge density (as  $\text{Li}^+$  or  $\text{Na}^+$ ) bind water molecules strongly, whereas large monovalent ions of low charge density (as  $\text{K}^+$  or  $\text{Cs}^+$ , but also arginine, histidine, and lysine side chains) bind water molecules weakly relative to the strength of water–water interactions in bulk solution.<sup>36</sup> This results in a difference in the ion solubility, and therefore in their adsorption on macromolecular surfaces.<sup>36</sup>

### 2.2 Mechanics of DNA in solution

Electrostatics of DNA in solution affects its mechanical properties. Three key parameters of the DNA double helix come

**Table 1** Genome compaction for different organisms. The volume reduction is evaluated, following Holmes and Cozzarelli,<sup>88</sup> as the ratio between the volume occupied by a DNA stretch of equivalent length free in solution and the volume of the container (capsid, cell or nucleus). The packing fraction is the ratio between the physical volume of DNA and the volume of the container, where the physical volume of DNA is defined as the volume of a cylinder of same length as the whole genome and radius equal to 1 nm

Organism	Sequence length (bp)	Volume reduction	Packing fraction (%)
<i>E. coli</i>	$4.6 \times 10^6$	$3.7 \times 10^2$ (ref. 88)	1.2
Yeast	$2.8 \times 10^7$	$1.9 \times 10^3$ (ref. 88)	0.7
Human	$6.0 \times 10^9$	$4.9 \times 10^4$ (ref. 88)	1.2
Virus (T4)	$0.17 \times 10^6$	$6.9 \times 10^3$ (ref. 15)	67

into play: both bending and torsional persistence lengths and the effective diameter. The bending persistence length is defined as the length over which correlations in the direction of the tangent to the polymer vanish. Its large value is due, in part, to the strongly charged polyelectrolyte nature of the DNA: repelling each other, these charges act against the double helix bending. On the other hand, the stacking interaction between DNA base pairs also contribute to the stability of the double helix and hence to its rigidity. By charging the DNA phosphate groups from a valency of zero to their normal valency of  $-1$ , the simulation by Savelyev *et al.* showed that the electrostatic contribution to the persistence length is roughly equal to the contribution from base stacking and other possible non-electrostatic sources.<sup>201</sup> The twisting persistence length is similarly defined and accounts for torsional rigidity, while the effective diameter accounts for electrostatic repulsion between two dsDNAs. These parameters determine how mechanical constraints are transmitted along DNA and transduced into compaction or decompaction.

DNA mechanical properties have been measured most efficiently through single DNA molecule manipulation by optical or magnetic tweezers. Indeed, such experiments give a direct access to the mechanical response of DNA under a stretching force and torsional constraints. Both bending and twisting rigidities as well as the effective radius of DNA have thus been probed in various salt conditions.<sup>160</sup> These experiments together with former measurements (light scattering or sedimentation) have shown that the bending persistence length slowly decreases from  $\sim 70$  nm down to  $\sim 30$  nm while increasing monovalent salt concentration for over three decades (from 2.5 mM up to 3 M).<sup>81,202</sup> Whereas it used to be assumed that the bending persistence length reaches a lower limit<sup>81</sup> (close to 50 nm) above  $\sim 20$  mM, several authors recently started to critically reexamine this assumption by theoretical<sup>132</sup> and computational<sup>202</sup> means and came to the conclusion that the bending persistence length steadily drops down with increasing salt concentration (down to  $\sim 30$  nm at 1 M). In the presence of multivalent ions at mM concentrations, the bending persistence length is strongly reduced and can reach 25 nm.<sup>134</sup> Again, the ion type matters, and ions in which the charge is centrally concentrated (point-like cations as  $\text{Mg}^{2+}$  or  $\text{Co}(\text{NH}_3)_6^{3+}$ ) yield lower values than ions in which the charge is linearly distributed (mainly polyamines, *e.g.* spermidine<sup>3+</sup> or spermine<sup>4+</sup>).<sup>10</sup>

As for the twisting persistence length, it is hardly affected by the monovalent salt whereas it was shown to decrease<sup>160</sup> from  $\sim 100$  nm to  $\sim 60$  nm when adding 5 mM  $\text{Mg}^{2+}$  into the solution. This change in persistence length can be accompanied by a change in the DNA local structure.<sup>191</sup>

Finally, many papers<sup>20,192,193,215</sup> give an estimation of the effective DNA diameter for different salt concentrations and show that it decreases from 15 to approximately 2 nm while increasing the monovalent salt concentration from 0.01 to 1 M. Divalent ions, such as  $\text{Mg}^{2+}$  even at mM concentrations, also strongly reduce the effective diameter down to less than 5 nm and therefore close to its crystallographic value.<sup>192,193</sup>

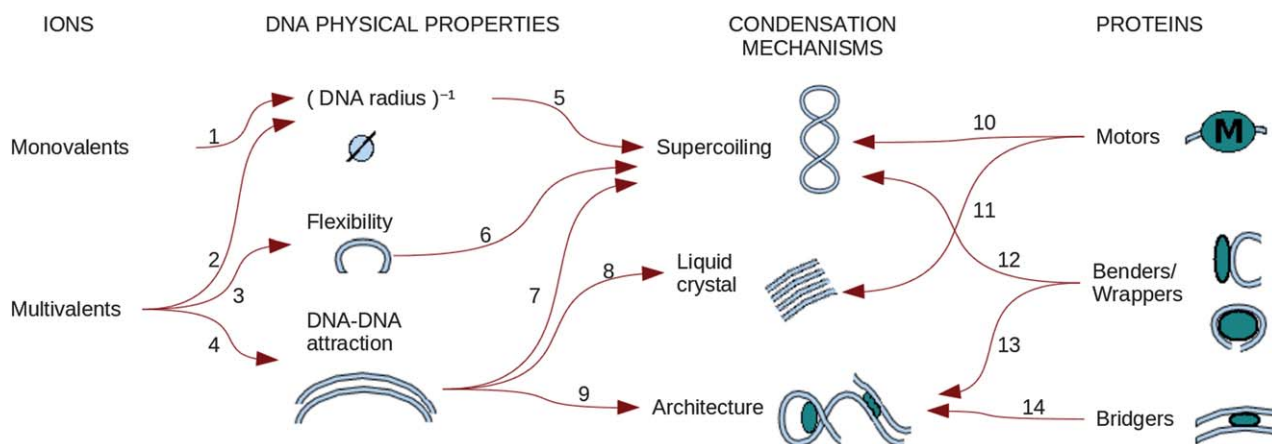
All these salt effects are summed up in a schematic representation in Fig. 1.

### 2.3 DNA–DNA ion-mediated interaction

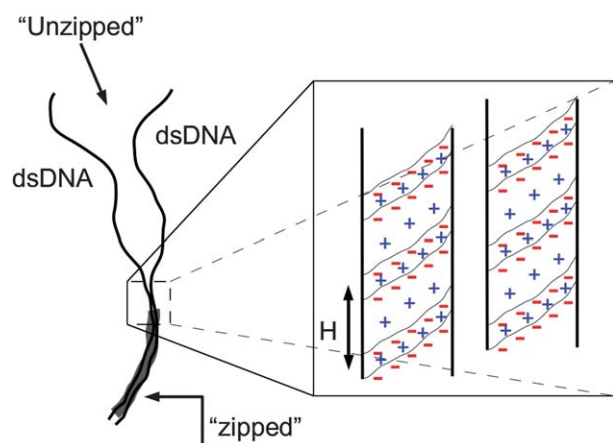
Ion mediated interactions have been shown experimentally and numerically to induce an attractive interaction between like charged objects such as dsDNAs.<sup>8,79,89,96,121,133,147,155,235,237</sup> An ion mediated attraction between like charged objects is not catchable within the well known Poisson–Boltzmann (PB) theory<sup>158,159,171,194</sup> and hence other approaches have been proposed during the last 15 years. The extent of the attraction as well as the possible physical mechanisms responsible for it are, however, counterion-type dependent. In the following three paragraphs we review the three main contributions to this topic.

(i) Nanoscale cluster formation and measurements of virial coefficients incompatible with purely repulsive ion mediated interactions have been observed in dsDNA solutions with monovalent salt only (reviewed in ref. 133). These effects are all the more strong that the ionic strength is low. To explain these observations, Manning has suggested an extension of its condensation theory involving two counterion populations: condensed counterions and bulk counterions. Such a description leads to an attraction between two dsDNAs at a finite distance. The underlying mechanism is akin to covalent bonding in the sense that the sharing of condensed counterions by two close dsDNAs results in an increase in their translational entropy.<sup>133,180</sup> In this respect, the original paper by Ray and Manning is enlightening.<sup>180</sup>

(ii) Multivalent ions with a centrally concentrated charge greater than  $+3$  (*e.g.*  $\text{Cr}^{3+}$ ) or specific divalent metal ions such as  $\text{Co}^{2+}$  and  $\text{Cd}^{2+}$  are known to lead to dsDNA aggregation, while usual divalent ions (*e.g.*  $\text{Mg}^{2+}$ ) are able to do so only with triple-stranded DNAs.<sup>179</sup> Theoretical strategies to account for these observations are twofold. On one hand, one can try to devise an effective model of the charge pattern on a *counterion dressed* dsDNA that would lead to an attraction within the PB theory. On the other hand, a theoretical treatment of the ion statistics that includes correlations can be proposed. An example of the first category is the so-called Kornyshev–Leikin (KL) theory which assumes that the charge pattern on dsDNA does not comprise negative charges only but also carries a fraction of irreversibly adsorbed cations mostly in the major groove.<sup>38,107–109</sup> This strong assumption is not discussed at the chemical level in the theory and a strict interpretation of this model might be contradictory with some experiments done on divalent metal ions for instance.<sup>99–102</sup> Despite this issue, one can compute the interaction between such dressed dsDNAs and predict an exponentially decaying attraction between two side by side dsDNA segments with a decay length that is almost salt independent,  $\lambda_{\text{KL}} \approx H/2\pi \approx 5 \text{ \AA}$  (where  $H \approx 3.4$  nm is the DNA average helical pitch). This attraction originates from the dipole-moment distribution on each dsDNA segment. The electrostatic attraction is predicted to be maximum for two homologous dsDNAs and when one of the dsDNAs is shifted along its axis by half its average helical pitch, an ideal “electrostatic zipper” is formed, as shown schematically in Fig. 2. If the dsDNA molecules are not homologous, this approach still predicts an attraction, but of much weaker intensity.<sup>38,107,109</sup> The simplest attempt for the second category of strategies considers bare dsDNA molecules carrying only negative charges in solution and tries to account for very strong ion–ion and ion–DNA correlations in the system.



**Fig. 1** Three main condensation mechanisms of DNA and the different impacts of ions, DNA physical properties and proteins are shown. Each arrow with its number corresponds to a precise point that is explained in a section of the text (arrows 1–3: Section 2.2, arrow 4: Section 2.3, arrows 5–7: Section 3.1, arrow 8: Section 3.4, arrow 9: Section 3.2, arrow 10: Section 3.1, arrow 11: Section 4.1, arrows 12–14: Section 3.2).



**Fig. 2** Schematic representation of the DNA–DNA interaction mechanism leading to attraction. If two double stranded DNA molecules with positive counterions irreversibly adsorbed in the major and minor grooves come in close contact with the appropriate phase shift, a “zipper”-like mechanism could happen between facing opposite charges.

This is the so-called Strong Coupling (SC) regime, where the ion–DNA correlations are so strong that they lead to the condensation of most counterions in an almost two-dimensional strongly correlated liquid at the surface of the DNA. In the case where the two-dimensional liquid is so correlated that it tends to form a two-dimensional Wigner crystal, the charge pattern at the surface of the DNA (that is comprised of adsorbed positive counterions on negative patches) displays a dipole-moment structure that leads to an exponentially decaying attraction akin to the “electrostatic zipper” of the KL theory.<sup>115,116,190,207</sup> This Wigner crystal picture has got a regain of interest recently and has been shown to give very good agreement with Monte Carlo simulations.<sup>195,196</sup> The SC regime has also been interpreted as a virial expansion whose first order of approximation would yield a single-counterion picture that also predicts a like charge attraction.<sup>97,147,157</sup> Even though its predictions diverge from MC

simulation at higher orders for two plates,<sup>195,196</sup> the virial expansion has been quite successfully applied to uniformly charged,<sup>96,155,156</sup> and helically charged<sup>95</sup> cylinders. In particular, an attraction owing to ion–DNA electrostatic correlations is found to appear only for counterion valencies higher than +3 for parameters corresponding to dsDNAs in solution.

Recent experiments on condensed dsDNA arrays<sup>47,228</sup> have shown that the DNA–DNA interaction displays exponentially decaying repulsion and attraction with typical length scales that are not very sensitive neither to the charge of the condensing agents nor to their chemistry (the measured range for the attraction is about 5 Å). As stressed by De Rouchey *et al.* 2010,<sup>47</sup> the attraction is quite likely due to some positive correlation between complementary charge motifs that is present in both KL and SC theories. The “electrostatic zipper” picture seems therefore a fair representation of a plausible mechanism for DNA–DNA aggregation in the presence of charged and compact condensing agents.

(iii) Most of the charged condensing agents consist in a repetition of a positive rod-like molecule into multivalent chains (*e.g.* polyamines)<sup>17,47,107,143,228</sup> and this particular structure could be responsible for an attraction mechanism independent of ion–ion electrostatic correlation. In fact, it has been shown by Bohinc and co-workers<sup>17,143</sup> that rod-like divalent ions (two charges separated by a fixed distance) could lead to attraction between two like charged plates when their charge is treated within a mean field theory. When the distance between the plates is about the size of the rod, then it is electrostatically favorable for counterions to form bridges lying perpendicularly to the macromolecules. This effect is better illustrated by imagining a rod ion exactly at the midplane between two plates: the rod will be unstable if oriented parallel to the planes, while it will be stabilized in the perpendicular direction. Finally, if the size of the divalent rods is above a certain threshold, then this “bridging” transition may appear at distances where the electrostatic gain overcomes the electrostatic cost of bringing two like charged particles closer and this yields attraction. Further developments accounting for correlations and multivalent chains have been proposed recently by the same authors.<sup>18</sup>

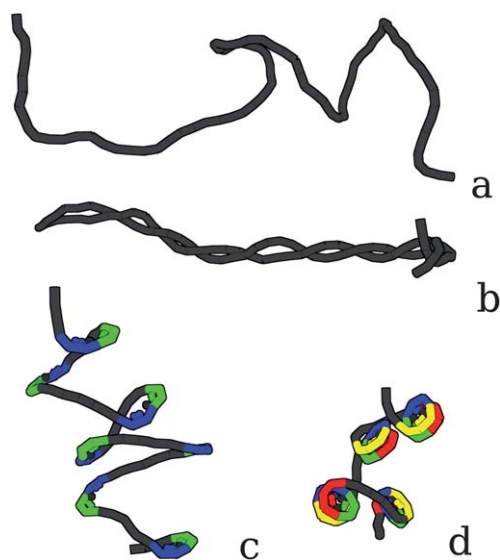
Overall, whatever the ion type, an attraction induced by counterions always relies on our current understanding of a *bridging* mechanism. This mechanism, as we have seen, can be due to an entropy increase induced by the sharing of condensed counterions, or a strong ion–macromolecule electrostatic correlation in the presence of high valency counterions, characterized either by a strongly localized charge, or by a smaller charge evenly distributed on a chain.

## 2.4 Electrostatics of the protein–DNA interaction

DNA-binding proteins are multivalent macroions that display charged patches which can bind to DNA.<sup>93,167</sup> Their shape and charge distribution determine their binding mode of interaction with the DNA double helix. Not surprisingly, patches are generally positively charged but some enzymes bear negatively charged patches, *e.g.* DNase I. Their binding to DNA is controlled by multivalent cations,<sup>78</sup> in a similar way as DNA–DNA ion-mediated attraction (see Section 2.3 above). Interestingly, proteins that recognize and bind to specific DNA sequences, such as transcription factors and restriction enzymes, seem to be on average less charged than non-specific DNA-BPs such as histones or nucleoid-associated proteins (NAPs, see Section 4.2). A possible reason for the lower charge observed for specific DNA-BPs is that it may play a role in facilitating protein sliding along DNA – a necessary step for finding target sequences. Indeed, these proteins have a characteristic concave shape, inducing a large fitting interface with DNA (of the order of a few tens of nm<sup>2</sup>).<sup>94</sup> For the intermediate values of the protein charge, local electroneutrality imposes the trapping of compensating ions between the DNA and the charged patches of the protein facing the DNA. When the protein approaches DNA, the increasing ion density induces a relevant osmotic pressure. The consequent repulsion is able to push the protein far from the DNA surface, thus preventing the formation of chemical bonds: the protein mobility is increased and the sliding becomes possible.<sup>42,167</sup> Simple models predict that the larger the charge of the patch, the lower the number of ions trapped and the lower the equilibrium distance between the protein and the DNA.<sup>42,166</sup> Therefore, it is expected that highly charged DNA-BPs will strongly stick to DNA. On the contrary, specific DNA-BPs undergo stronger repulsion which is able to push the protein away from the DNA surface, far enough to prevent the formation of chemical bonds: the protein mobility is then increased and the sliding becomes possible.<sup>42,167</sup>

## 3 Mechanisms of DNA compaction

DNA, ions and proteins interact, in different organisms, as to build and organize the genome in a functional way. Electrostatics plays a crucial role in such interaction, and may justify by itself the simplest observed architectures, *e.g.* that observed in viruses. However, increasing complexity of the organisms needs an increasing involvement of specific structural features of proteins, then protein–protein interactions, and finally the emergence of a collective behavior should be invoked. The overall organization obtained is active, ATP dependent. Alternative architectures exist in most cases.



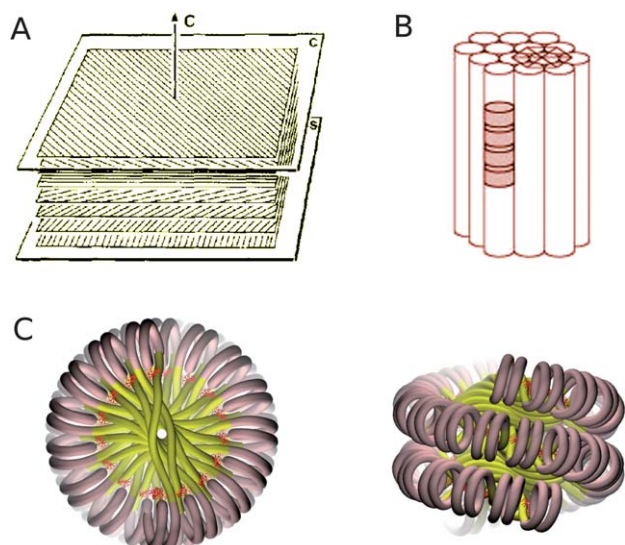
**Fig. 3** Schematic view of two different mechanisms for DNA compaction: an unconstrained supercoiling and a constrained, protein-mediated architecture. A 900 bp long DNA is represented in a thermally disordered configuration (a) and in a supercoiled state (b). The same length of DNA is then organized by the formation of 7 tetrasomes (c) or 4 nucleosomes (d). Color code: blue and green spheres represent H3–H4 dimers or, equivalently, HMfA and HMfB archaeal proteins. Red and yellow spheres represent H2A–H2B dimers. DNA stretches that are bound to proteins are represented in the same color as the corresponding proteins. The four pictures are represented at the same scale.

Nonetheless, mechanics, electrostatics and ion-mediated effects still play a crucial role in tuning the dynamical evolution of such complex structures. We can identify three major fundamental mechanisms involved in genome condensation: (i) DNA over- or under-winding, *i.e.* DNA supercoiling (Fig. 3b), (ii) building of a (hierarchical) architecture supported by DNA folds mediated by proteins (Fig. 3c and d), and (iii) DNA–DNA aggregation inducing liquid crystal ordering, a mechanism that depends not only on the presence of multivalent ions but also on osmotic stress (Fig. 4). The interplay between these mechanisms may be sophisticated, and results in a wealth of possibilities. For example, supercoiling may accompany and enhance DNA–DNA interaction by braiding two dsDNAs.

In this section, we will discuss some details of these three fundamental mechanisms, while the *in vivo* implementation of these mechanisms will be discussed in Section 4. As a guide, Fig. 3 presents a schematic view of the first two compaction modes mentioned above.

### 3.1 DNA compaction and supercoiling

One way to compact DNA is to apply torsional constraints on it so as to achieve supercoiling. Torsional constraints applied on a rope produce either plectonemic or toroidal structures depending on the superhelical density (*i.e.* the number of turns that have been applied on the relaxed rope, divided by the length of the rope).<sup>90</sup> Plectonemes are loops of dsDNA helices twisted together, as commonly observed in garden hoses and telephone or computer cords, whereas toroids are annular-shaped



**Fig. 4** Liquid-crystal order of DNA and nucleosomes. (A) Scheme of a DNA cholesteric right handed liquid crystal order. In each plane perpendicular to the cholesteric axis  $C$ , DNA molecules are arranged in the same orientation. The orientation in a particular plane is obtained from the orientation of DNA in the previous one by a rotation of an angle  $\theta$ . The same orientation is therefore recovered every  $\pi/\theta$  plane. Adapted from Livolant 1991.<sup>122</sup> (B) Stacks of nucleosomes in a nucleosome liquid crystalline order. Nucleosomes are represented as pink cylinders and columns of nucleosomes are shown to interact laterally. Adapted from Mangenot *et al.* 2003.<sup>130</sup> (C) Left and right: respectively top and side view of the chromatin fiber. Nucleosomes (in pink) are stacked in columns as in (B) and the DNA linkers (in yellow) are stacked in a similar cholesteric order as in (A). Linker histones H1 are shown in red.

structures favored at higher superhelical density. The critical value of the twisting torque necessary to induce plectonemes is called the buckling torque. Both DNA bending and twisting persistence lengths affect the buckling torque, together with another parameter, the DNA effective diameter, directly related to DNA–DNA electrostatic repulsion (see Section 2.2 above). Using magnetic tweezers,<sup>187</sup> Mosconi and co-workers<sup>149</sup> provided an experimental study of DNA supercoiling by applying a rotation to the magnetic bead (magnetic tweezers cannot impose a torque to the magnetic bead) at one DNA extremity while keeping the other clamped, at constant stretching force. A theoretical treatment<sup>250</sup> provides an indirect measurement of the critical buckling torque at different salt concentrations. Moreover, the buckling transition becomes steeper at higher salt concentrations, and disappears in low salt concentration solutions.<sup>23,63</sup> At low stretching forces ( $<0.3$  pN) and physiological salt concentration ( $\sim 200$  mM) the buckling torque is estimated to be less than 4 pN nm (*i.e.* of the order of  $k_B T$ ).<sup>149</sup> This suggests a relative ease in forming and modifying plectonemes *in vivo*. Interestingly, the buckling torque is much affected by the ionic strength, decreasing down to half its maximal (*i.e.* at low salt) value with increasing monovalent salt concentration, and rather independently from the stretching force value.

Adding multivalent cations to the solution is another way to induce supercoiling in DNA, and this results in well-defined condensed structures, particularly toroids and bundles of

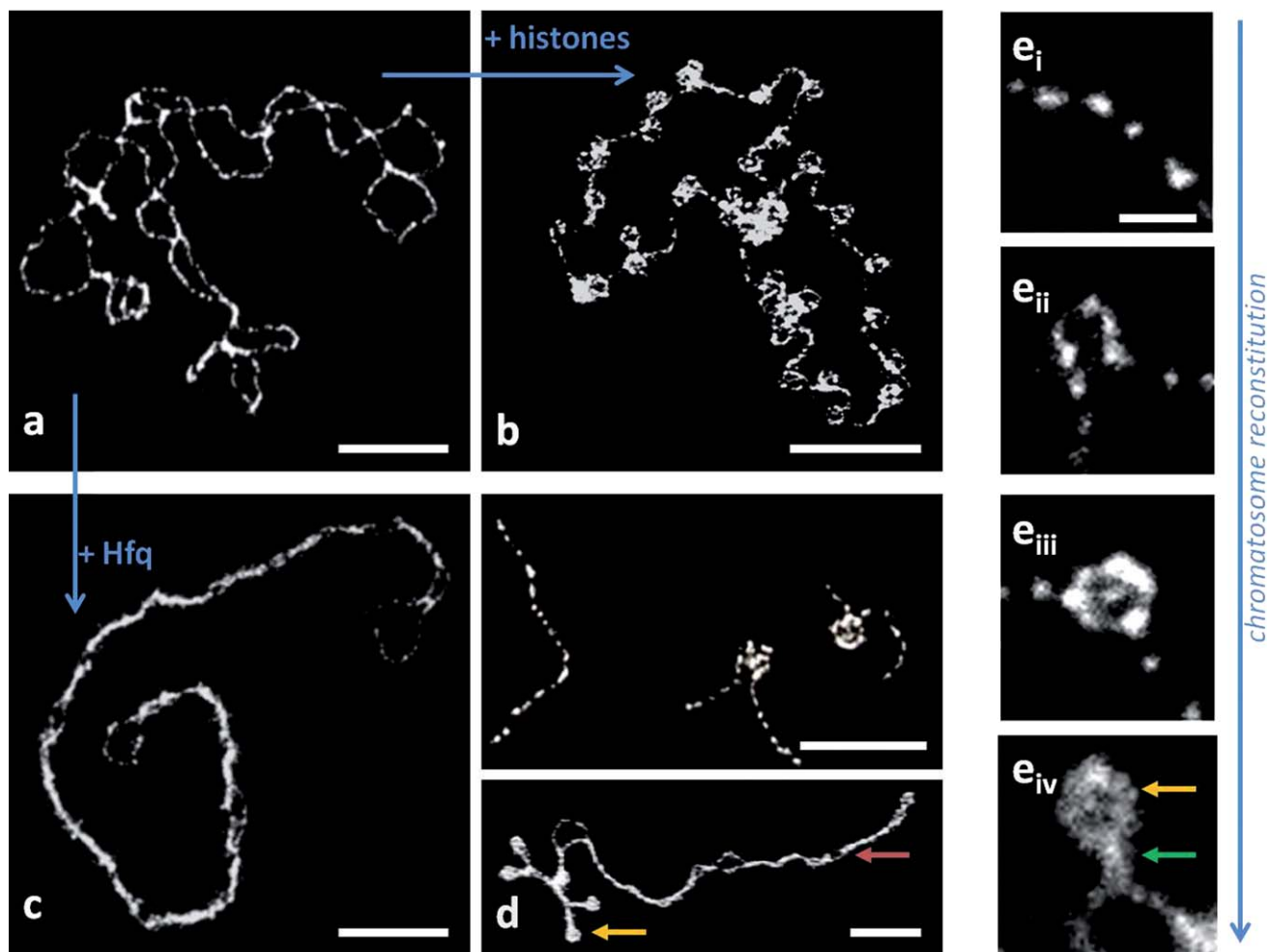
rods<sup>14,53,58,59,140</sup> similar to those observed in viruses (see Section 4.1). In this case, the zipping mechanism discussed in Section 2.3 is directly involved. However, multivalent cations also affect the formation of plectonemic structures. The plectoneme radius (also called supercoiling radius)  $\rho$  depends indeed on the concentration of the monovalent salt<sup>234</sup> and on the stretching force  $f$  too ( $\rho$  decreases when increasing  $f$ ).<sup>32,33,160,161</sup> Interestingly, divalent ions, *e.g.*  $Mg^{2+}$ , strongly lower the supercoiling radius, thus facilitating plectoneme formation (actually by reducing the critical torque), which in turn facilitates bridging of both dsDNAs in the plectoneme. Moreover, at the atomistic level, Timsit and Varnai have shown that  $Mg^{2+}$  (and other divalent cations too) specifically promotes the formation of stable right-handed crossovers, so that DNA is preferentially condensed into left-handed plectonemes (hence positively supercoiled).<sup>227</sup>

### 3.2 Architectural proteins

Highly charged macroions can induce extensive bending and even wrapping of DNA. This effect can be explained by the electrostatic interactions between a charged flexible cylinder, modeling DNA, and a positively charged sphere.<sup>4,205</sup> In this model, the number of turns of DNA is determined, in physiological conditions, by the sphere charge and radius. This effect can obviously be extended to larger charged bodies, and in particular to proteins or protein complexes.

Interestingly, many of the known structural proteins involved in genome compaction can be qualified as benders or wrappers.<sup>126</sup> While in the case of wrappers (and notably for histone octamers within eukaryotes) the suggested electrostatic-induced DNA bending may be considered as the main mechanism involved, this is not always the case for benders. Indeed, DNA can be bent away from or around onto the bending protein according to the shape and charge distribution of the protein. Importantly, DNA wrapping (and bending to a lesser extent) is generally chiral (see below the nucleosome structure), hence generates some amount of supercoiling, which is constrained in the protein–DNA complex and does not require any active process (motor) for its maintenance. Note that proteins that can kink DNA have also an impact on supercoiling formation. Buckling transition occurs at significantly lower twist for kinked DNA molecules.<sup>23</sup>

A third class of architectural proteins is represented by DNA “bridgers”, which are able to bind simultaneously to two dsDNAs and join them together. An example of such behavior is given in Fig. 5c for the bacterial protein Hfq. As described above for multivalent ions, electrostatics is the main actor of this bridging effect. And here again there exists a synergistic interplay between bridging and supercoiling. But, unlike multivalent cations, proteins have enriched features allowing a more complex behavior as *e.g.* a given degree of sequence specificity or the possibility of acting differently (as bender, wrapper or bridger) according to their monomeric or multimeric form. Moreover, the protein charge distribution may be tuned by post-translational modifications (as acetylation or phosphorylation), which in turn may finely tune its interaction with DNA.



**Fig. 5** (a) The folding of a naked plasmid in solution depends on its amount of supercoiling and on the ionic conditions. Here, a 5400 bp supercoiled plasmid was spread in bi-distilled water, showing a typical entangled conformation. (b) When core histones (purified from eukaryotic nuclei) are added, DNA wraps around them in typical round-shaped octameric particles (so-called nucleosomes) made from two of each histone type H2A, H2B, H3 and H4.<sup>117</sup> (c) The same plasmid incubated in the same condition with the “bridger” protein Hfq (purified from *E. coli* nucleoids) shows instead a typical bridged structure.<sup>70</sup> (d) Remarkably, some proteins from bacteria also wrap DNA in a “nucleosome-like” manner, as clearly seen from LrpC binding to short linear DNAs (upper part).<sup>11</sup> Indeed, LrpC was shown to form tetramers around which DNA wraps in a right-handed manner, causing an increase in the free negative supercoiling as seen from the tight winding of the free DNA part – red arrow – compared with free DNA in panel (a) to compensate for the LrpC-restrained positive supercoils – yellow arrow.<sup>11</sup> ( $e_i$ – $e_{iv}$ ) *In vitro* chromatosome reconstitution shows a three step process: (i and ii) a (H3–H4)<sub>2</sub> tetramer binds to DNA, (ii and iii) two H2A–H2B dimers join the structure, (iii and iv) eventually, a linker histone (H1) comes and seals the two linker DNAs. Note that core histones (H2A, H2B, H3, H4) together act as DNA wrappers (yellow arrow), while H1 acts as a bridger (green arrow). Scale bars: (a–d): 50 nm; ( $e_i$ – $e_{iv}$ ): 10 nm. Picture (d) by courtesy of Eric Le Cam.

### 3.3 A major architectural pattern: the nucleosome

Histone-like proteins in prokaryotes and archaea and histone octamers in eukaryotes carry large positive charges, which are responsible for strong electrostatic interactions with DNA. Many of these structural proteins bend DNA and work as DNA “benders” or “wrappers”, depending on the degree of distortion induced on the DNA. In their electrostatic model, Arcesi and co-workers<sup>4</sup> considered the examples of an archaeal protein and of the histone octamer complex, and the result is rather close to the geometry actually observed. A schematic view of these geometries is shown in Fig. 3c and d, while Fig. 5e shows an example of nucleosome formation *in vitro*. This suggests a key role of electrostatics in tuning the specific geometry of DNA–protein

complexes. However, this “non-structurally focused” view of the histone octamer simply bearing 146 net positive charges (many of which are on the tails) should not occult the fact that binding of DNA to the octamer is primarily mediated through a series of fourteen discrete binding sites (referred to as “minor groove-in positions”<sup>82</sup> or “superhelix axis locations”<sup>124</sup>) made by arginine residues spaced more or less evenly around the octamer. These residues form a left-handed superhelical ramp of a charged side chain that inserts between two phosphates into the minor groove of each DNA double helical turn.

The nucleosome reconstitution process *in vitro* can be seen as a complexation between an excess of polyanions (146 bp of DNA per nucleosome, with a net charge of  $-294e$ ) and a limited pool of polycations (an octamer of basic histones per nucleosome,

with a net charge of  $+146e$ ).<sup>110</sup> So the nucleosomal complex is overcharged by the DNA.<sup>205</sup> A third ingredient is needed to allow the proper deposition of positively charged histones on to the negatively charged DNA and avoid non-nucleosomal aggregates that tend to form spontaneously when histone and DNA are simply mixed at low ionic strength. This “chaperone” function is usually mediated by histone-binding proteins (so-called “histone chaperones”), such as NAP-1 or nucleoplasmin,<sup>221</sup> negatively charged polymers (such as polyglutamate or RNA) or a high concentration of monovalent salts (such as NaCl). In the last case (known as salt gradient procedure), the core histones are initially combined with DNA in 2 M NaCl, then the NaCl concentration is slowly decreased by dialysis or dilution of the mixture.<sup>183,214</sup> Interestingly, this method recently inspired a new way to monitor the electrostatic complexation of anionic nano-particles by cationic–neutral block copolymers, which allows the formation of clusters with a size controlled by the desalting kinetics.<sup>65</sup>

### 3.4 Liquid-crystal ordering of DNA and nucleosomes

*In vitro*, pure DNA can form multiple liquid crystalline phases when the concentration is increased by osmotic stress. The two most typical phases are a left-handed cholesteric organization, with a pitch typically ranging from 2 to 3 microns,<sup>122</sup> and an hexagonal compact form (see Fig. 4a). The cholesteric pitch depends on both steric constraints due to the DNA right-handed helicity and electrostatic contributions. Isolated nucleosome core particles (NCPs) as well can form liquid crystalline phases under osmotic stress (*i.e.* by increasing the nucleosome concentration, for example by adding crowding agents such as PEG or dextran). They were shown to stack one upon each other, forming columns which then organize laterally to form an hexagonal packing<sup>130</sup> (see Fig. 4b). These phases can also be obtained by aggregation of the NCPs by multivalent cations.<sup>13</sup> In the case of chromatin, nucleosomes are regularly spaced along the genome and the contributions of both DNA and NCPs are expected to aid compaction. Indeed, a regular nucleosomal array is known to fold *in vitro* in a compact 30 nm chromatin fiber if divalent cations and the linker histone H1 are present.<sup>188</sup> All-atom modeling of the fiber<sup>244</sup> tends to argue in favor of the idea that the fiber stability results from an interplay between a short pitch cholesteric DNA liquid crystalline phase of linker DNA and a columnar phase of nucleosomes stacked in helices (see Fig. 4c). The pitch of the DNA helix can either be left or right handed depending on the linker length. Recent experiments on the formation of liquid crystalline phases of short DNA tend to support this view.<sup>249</sup>

## 4 *In vivo* implementation

### 4.1 Viruses

While the genome is highly confined spatially in all living organisms, the highest level of compaction is observed in the simplest possible organisms: viruses (see Table 1). During the assembly of many dsDNA viruses, the genome is compacted to near liquid crystalline density into a viral capsid (tens of thousands of base pairs for a total contour length of several microns packed into 50–80 nm capsids or virus heads).<sup>54</sup> Significant DNA bending must occur during DNA packaging; moreover,

a repulsive electrostatic barrier, due to the negatively charged phosphate backbone, must be overcome.<sup>105,162,177,178,185,233</sup> This effect is partially reduced since the capsids of many dsDNA viruses are penetrable for small ions.<sup>2</sup> The presence of di- and tri-valent cations in solution can therefore enhance the DNA–DNA electrostatic attraction and ease DNA packaging into the capsid. Neutralization of negative charge is generally done either with polyamines<sup>7</sup> and/or cations,<sup>248</sup> whereas some viruses partially neutralize the DNA charge with a positively charged domain of the capsid protein.<sup>240</sup>

As mentioned in Section 3.1, cations or polycations can spontaneously collapse DNA *in vitro* into folded-fiber (toroidal) forms provided the separation of a DNA phase occurs. Polylysine and the naturally occurring polyamine spermidine, in particular, have been reported to collapse DNA into a rod-shaped or toroidal structure.<sup>73,84,113,163,246</sup> On the other hand, electron microscopic studies have revealed a similar toroidal arrangement of the DNA inside certain bacteriophages and animal viruses.<sup>37,67,106,119,184</sup> This raises the possibility that spermidine and spermine are responsible for viral compaction *in vivo*,<sup>56</sup> since these polyamines are normally found in the host cell as well as in the mature virion.<sup>2,72</sup>

In spite of these effects, the size of the viral capsid is close to the persistence length of the DNA, and a strong energy barrier has to be overcome during viral assembly. The viral genome packaging is indeed achieved by an ATP-powered molecular motor that translocates dsDNA into the preformed viral capsid.<sup>28,210</sup> This process can be assimilated with the formation of a DNA liquid crystalline mesophase using an osmotic stress (see Fig. 4). A detailed investigation of the role of electrostatics in single DNA packaging was carried out in bacteriophage  $\Phi 29$  *via* optical tweezer measurements,<sup>66</sup> showing that the internal force resisting DNA packaging was lowered when  $Mg^{2+}$  was the dominant ionic species or by addition of 1 mM  $Co^{3+}$ . This force was up to 80% higher with  $Na^+$  as the dominant counterion and only 90% of the genome length could be packaged.<sup>66</sup> This study revealed that ions also affect motor functioning;  $Mg^{2+}$  is required for initiation of packaging, while  $Na^+$  increases the motor velocity by up to 50%.  $Mg^{2+}$  is indeed known to be a cofactor for motor functioning as is the case with polymerase helicase and endonuclease.<sup>39</sup>

### 4.2 Prokaryotes

In prokaryotes, DNA is not enclosed by a membrane, but occupies an irregular region of the cell called the nucleoid. Prokaryotic genome is generally circular (possibly in multiple copies). It is negatively supercoiled, and shows distinct domains topologically independent, which have dynamic boundaries but an average size of 10 kbp.<sup>230</sup> Such a high level of negative supercoiling can be obtained thanks to the action of a specific bacterial enzyme, the DNA gyrase, that has the unique capability of introducing negative supercoils in DNA.<sup>71</sup> High valence ions such as spermines and spermidines contribute to lower the electrostatic DNA–DNA repulsion and thus facilitate supercoiling and compaction in bacteria; however, DNA compaction implies the action of nucleoid-associated proteins (NAPs) that, together with DNA, form the bacterial nucleoid.<sup>21,48,125,213,230</sup>



There are about twelve NAPs involved in the formation of the nucleoid.<sup>6</sup> Some of these proteins are distributed uniformly within the nucleoid, while the others show an irregular distribution.<sup>5,238</sup> NAPs' contribution to the overall structure of the bacterial chromosome *in vivo* is unclear, since most details about how they interact with DNA come from *in vitro* studies.<sup>43</sup> Some NAPs constrain supercoils and some do not.<sup>165,206,218,232</sup> Most of them act as benders or bridgers, and therefore possibly facilitate and stabilize supercoiling (see Fig. 1).

The most abundant NAP is the small H-NS (histone-like nucleoid structuring protein), present in >20 000 copies per cell,<sup>186</sup> that is about one protein per 400 bp. H-NS is a DNA bridger that acts as a dimer<sup>126</sup> and compacts DNA.<sup>49,186</sup> Due to the presence of two DNA binding domains, the protein can interact with two DNA duplexes simultaneously.<sup>44</sup> The protein binds with relatively high affinity to any DNA sequence,<sup>50</sup> although some preferential binding sites have been reported.<sup>19,114</sup>

HU (heat-unstable nucleoid protein) and FIS (factor for inversion stimulation) are two DNA benders that act as dimers and can bend DNA up to 160° (ref. 126) and to 90°,<sup>168</sup> respectively. HU induces negative supercoils *in vitro* when incubated with DNA and topoisomerase I.<sup>22</sup> It exhibits no sequence specificity, but it has been shown to bind preferentially to intrinsically flexible DNA.<sup>220</sup> At high concentrations HU dimers can be packed tightly side-by-side in a cooperative manner resulting in effective stiffening of the DNA.<sup>236</sup> The HU/DNA co-crystal structure provides, at high resolution, an explanation for how HU may constrain negative supercoils, and evidences a proline intercalation and asymmetric charge neutralization mechanisms of DNA bending.<sup>218</sup>

When bacteria are grown under limited food conditions, they reach a phase called the stationary phase during which they stop growing and dividing. It was observed that, shortly after the onset of the stationary phase, bacterial chromatin undergoes a transition from a disperse morphology to several ring-shaped toroids that contain DNA and the protein Dps (DNA-binding protein from starved cells).<sup>64</sup> This condensation overall is driven in part by the loss of the transcriptional activity and by the natural tendency of DNA to condense at high concentrations in the presence of counterions. This type of condensation is similar to packaging in phage heads and bacterial spores, and probably constitutes one of the simplest ways of creating inactive bacterial chromatin.

Besides benders and bridgers, some DNA wrappers are also present in bacteria. Conserved examples are members of the Lrp/AsnC family. These proteins form disc-shaped octameric structures with multiple binding sites that wrap DNA around themselves in a right-handed superhelix.<sup>46,223</sup> An example of this behavior *in vitro* is given in Fig. 5d for protein LrpC. When two Lrp octamers associate, a hexadecamer is formed that can wrap DNA around itself by almost two full turns, resulting in a nucleosome-like structure.<sup>223</sup> Interestingly, members of the Lrp/AsnC family also appear to have a second mode of interaction with DNA that involves duplex bridging.<sup>11,222</sup>

### 4.3 Archaea

The archaeal and eukaryotic nucleosome core histones evolved from a common ancestor. The best studied example of an archaeal histone is the HMf protein. Two HMf polypeptides

exist (HMfA and HMfB) that can form homodimers and heterodimers with different DNA binding properties.<sup>197,198</sup> Further assembly into tetramers allows wrapping of 60–70 bp of DNA,<sup>77,174</sup> in a similar way to eukaryotic tetrasomes (H3 + H4)<sub>2</sub> (ref. 1) (Fig. 3c). The wrapping (and consequent supercoiling) can be negative or positive, and depends on the histone-to-DNA ratio and salt concentration.<sup>135,154</sup> Interestingly, Soares and co-workers<sup>211</sup> substituted eukaryotic conserved lysine residues at their corresponding locations into the archaeal histone, HMfB, and obtained an increased affinity to DNA and a higher compaction level. At the same time, these complexes remained negatively supercoiled at all histone-to-DNA ratios, thus losing their flexibility to wrap DNA alternatively in either a negative or a positive supercoil. Both observations suggest a crucial role of these residues in the evolutionary differentiation of archaeal and eukaryotic histones.

Archaeal histones are considerably smaller than their eukaryotic counterparts and, with some exceptions, lack the C- and N-terminal tails. The absence of such tails suggests that archaeal histones are not subject to similar regulatory processes as found in eukaryotes. In line with this finding there is no evidence for post-translational modification of archaeal histone proteins.<sup>199</sup>

Besides HMf, other proteins participate in DNA packaging, such as MC1, an abundant histone-like protein co-responsible for the physiological structuring of the archaeal chromosome.<sup>118</sup> MC1 is a globular protein with a net positive charge of +12; its visualization by EM shows sharp kinks in DNA with a bending angle of 116° (or 0.32 turn).<sup>34,91,169</sup> Alba is another DNA condensing protein expressed in most archaeal genomes.<sup>3</sup> It forms regular fibrous structures with plasmid DNA, possibly wrapping two dsDNAs around each another.<sup>127</sup>

Interesting temperature-related features emerge in the case of hyperthermophilic archaea, living at temperatures of the order of 90 °C. If a high GC content is needed to resist to heat double-strand denaturation, intermolecular ion pairs formed between histidine and aspartate during histone tetramerization observed *in vitro* have also been suggested to play an important role in the maintenance of the archaeal nucleosome structure and for DNA stabilization.<sup>87</sup> Most interestingly, unusual long and branched polyamines are often found in hyperthermophile archaea. They are more abundant when cells are grown at higher temperature<sup>164</sup> and are thought to stabilize DNA against thermal denaturation.

### 4.4 Eukaryotes

In the case of eukaryotes, DNA wrapping is the essential compaction mechanism and it is achieved by specific octameric complexes made of histones. The different steps in the histone–DNA aggregation *in vivo* are illustrated in Fig. 5. The first organization level in eukaryotes is the nucleosome, which consists of a “nucleosome core particle” (NCP) of 146–147 base pairs of DNA wrapped around a roughly cylindrical core of eight histone proteins<sup>45,124</sup> and a linker DNA (stretch of free DNA between two NCPs, Fig. 3d and 5b). The NCP shows a tripartite structure: a (H3–H4)<sub>2</sub> tetramer binds to DNA (Fig. 5e<sub>ii</sub>), then is complexed by two H2A–H2B dimers (Fig. 5b and e<sub>iii</sub>). A linker histone, named H1, may eventually come and seal the two linker DNAs flanking a NCP (Fig. 5e<sub>iv</sub>). Linker histone H1 and its

variants<sup>251</sup> were shown to interact with chromatin transiently in the order of minutes.<sup>120,146</sup> The H1 globular domain contains two DNA binding patches on opposite sides of the protein, similar to the suggested structure of bacterial H-NS.<sup>76,125</sup> The presence of these two DNA-binding domains endows H1 with the ability to bridge DNA duplexes.<sup>31</sup>

Nucleosomes *in vivo* are rather regularly spaced along the genome, so that nucleosome core particles form a “beads on a string” array, also called “nucleosomal array”.<sup>62</sup> The repeat length (*i.e.* the mean number of base pairs per nucleosome) may vary depending on organisms, tissues, and genomic regions, and is approximately in the range of 155 to 240 bp.

It is generally assumed that higher order structures of the nucleosomal array are involved in achieving an efficient and adjustable genome compaction. However these higher levels of organization remain largely speculative. Folding of the nucleosomal array into 30 nm chromatin fibers has been observed *in vitro*<sup>103,189</sup> and some indications of their existence *in vivo* have been recently obtained,<sup>203,204</sup> but the subject is still debated.<sup>128</sup> The issue of the higher-order structure of chromatin is directly related to the characterization of nucleosome–nucleosome physical interactions. The ability of nucleosomes to stack their faces together has been earlier observed using electron microscopy on isolated NCPs.<sup>52</sup> As discussed in Section 3.4, this was later confirmed and more precisely quantified by the study of NCP liquid crystals (see ref. 123 for a review). The relevance of this nucleosome–nucleosome interaction within chromatin fibers has been proved by the presence of an 11 nm meridional band in the X-ray diffraction pattern of oriented fibers.<sup>241</sup> All the models for compact fibers are therefore based on the assumption that nucleosomes are stacked in columns which are further twisted to form the 30 nm fiber (see *e.g.* Fig. 4c). In the emerging picture, these stacking interactions are essential for the fiber integrity and can be tuned *in vivo* using both variant histones and post-translational modifications (see ref. 231 for a review).

While packaging of the genome in eukaryotes is primarily due to the wrapping of DNA around nucleosomes, members of the high mobility group (HMG) family of proteins play a supplementary, possibly modulatory, role in chromatin organization.<sup>24,225</sup> HMGs<sup>172,182</sup> (including HMGA family,<sup>182</sup> HMGB family<sup>225</sup> and HMGN family<sup>26</sup>) are DNA benders. At a distance of 2 bp from the first kink, a second kink is induced due to partial intercalation of two additional residues.<sup>153</sup> This induces a bend into DNA of up to 80° for a single HMG box and up to 130° for tandem boxes. While HMG box proteins can promote DNA compaction by bending,<sup>209</sup> a picture is emerging in which these proteins function by favoring the partially unwrapped state of nucleosomes by binding at the entry/exit point, which possibly facilitates nucleosome remodeling.<sup>229</sup> Other non-histone proteins participating in chromatin polymorphism and dynamics are HP1, SMCs,<sup>92</sup> DEK,<sup>239</sup> and BAF.<sup>137</sup>

Finally, it is worth noting that mitochondria, constituent organelles of eukaryotes, have a compaction strategy similar to that in bacteria, and use HMG-like proteins.<sup>29,61</sup>

## 5 Functional insights: tuning the DNA compaction

We have seen so far that DNA compaction in cells is highly driven by electrostatic properties of the proteins and ions. We

can therefore expect that the functional tuning of this compaction is also driven through electrostatics. In this section, we wish to address this specific issue in eukaryotes focusing on three different aspects. We will first review the tuning of the protein local charge through post-translational biochemical modifications of aminoacids focusing our attention on the nucleosome. We then turn to the extreme chromatin condensation observed during cell division, which involves post-translational modifications and also the action of molecular motors. Finally, we will discuss some results on the role of cation concentration variations in tuning the chromosome compaction during the cell cycle.

### 5.1 The role of biochemical modifications

Cell differentiation, implied *e.g.* in the formation of different tissues, needs heritable changes in gene expression. These changes can be caused by mechanisms other than genetic modification of the nucleotide sequence, and are referred to as *epigenetic* modifications. These changes are typically related to chromatin structure and compaction, which in turn alter the gene expression. DNA methylation and histone modification are typical examples of such modifications.

Histone acetylation and phosphorylation effectively reduce the positive charge of histones and this has the potential to disrupt electrostatic interactions between histones and DNA. This presumably leads to a less compact chromatin structure, thereby facilitating DNA access by protein machineries such as those involved in transcription. Notably, acetylation occurs on numerous histone tail lysines.<sup>112</sup> Moreover, acetylation of Lys9 and Lys56 is correlated with response to DNA damage.<sup>226</sup>

The X-ray structure of the nucleosome indicates that highly basic histone amino (N)-terminal tails can protrude from their own nucleosome and make contact with adjacent nucleosomes.<sup>124</sup> Modification of these tails affects inter-nucleosomal interaction and thus affects the overall chromatin structure. Acetylation of lysine or phosphorylation of serine or threonine is one way to reduce the charge. Models predict that acetylation of the histone core will destabilize the nucleosome.<sup>60</sup> This is in agreement with recent *in vivo* experimental evidence which suggests that acetylation of lysine 56 or 36 of histone H3 is necessary for efficient gene transcription.<sup>148,245</sup> Moreover, acetylation of lysine 56 also enables DNA replication and prevents epigenetic silencing<sup>51,145</sup> consistently with a looser state of the nucleosome acetylated at lysine 56 in the theoretical model.<sup>60</sup> Conversely, deacetylation of lysine 56 of histone H3 leads to stronger compaction and to a transcriptionally silent chromatin (heterochromatin).<sup>245</sup> As another experimental example, acetylation of all H4 histone tails at lysine 16, a prevalent epigenetic modification in eukaryotes, causes charge reduction and in turn inhibits the formation of compact 30 nm chromatin fibers and hampers the formation of cross-fiber interactions.<sup>208</sup>

Moreover, modifications regulate the recruitment of remodeling enzymes. Acetylated lysines are bound by bromodomains, which are often found in HATs (histone acetylase) and chromatin remodeling complexes.<sup>151</sup> For example Swi2/snf2 contains a bromodomain that targets it to acetylated histones. In turn, this recruits the Swi/snf remodeling complex which functions so as to open the chromatin.<sup>83</sup>

## 5.2 Chromatin compaction in mitosis (so-called “chromatin condensation”)

Chromatin condensation and subsequent de-condensation are essential for proper execution of cell division, or mitosis. During mitosis chromatin is highly compacted, whereas it is de-condensed right after chromosome segregation. Phosphorylation of histone H3 is a key process driving mitosis and the histone H3 tail is phosphorylated at four specific residues.<sup>40,74</sup> Interestingly, phosphorylation at three of those residues is found in transcriptionally active genes. Thus, H3 phosphorylation is believed to be involved in two structurally opposing processes, chromatin de-condensation observed during gene transcription in interphase and chromosome condensation during cell division.<sup>175</sup> In mammalian cells, phosphorylation of H3 starts in the late G2 interphase. Subsequently, the phosphorylation of H3 spreads along the chromosomes and is completed in prophase. It is still present in metaphase.<sup>86</sup> The temporal and spatial relationship between chromosome condensation and phosphorylation has been clearly observed.<sup>75</sup> In addition, acetylation of H3 also occurs on the same H3 tail.<sup>69</sup>

Replicated chromosomes assembled in linker histone H1 depleted crude extracts are thinner and 50% longer than control chromosomes.<sup>136</sup> Moreover, H1 is hyperphosphorylated by mitotic cyclins.<sup>16</sup> These studies suggest a role of H1 in the mitotic chromosome structure. Since the existence of chromatin fibers in mitotic chromosomes has been challenged recently, based on two independent electron microscopy studies,<sup>57,68</sup> this role seems different from its role in stabilizing the *in vitro* compaction of nucleosomes into thicker 30 nm fibers.<sup>224</sup>

The mechanism that drives chromatin condensation during mitosis is still debated.<sup>57,68,150</sup> Molecular motors such as condensins and topoisomerases play an essential role in this process. Condensed mitotic chromosomes have been shown to be organized in loops of DNA anchored to a proteinaceous chromosome axis.<sup>55,173</sup> The mechanism by which condensins generate and maintain condensation remains controversial. The enrichment of the chromosome axis with condensins suggests that condensin bridges could stabilize these loops by anchoring them to the axis. Recent reports suggest that condensins first bind to DNA through electrostatic interactions and then trap distant DNA segments to stabilize chromatin loops.<sup>41</sup> It has been also suggested that the introduction of positive supercoiling by condensins could be the basis for mitotic chromosome architecture.<sup>104</sup>

## 5.3 Cations in chromosome condensation

Cations are essential participants in chromosome condensation.<sup>139,140</sup> Mitotic chromosomes become very swollen following the depletion of  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$ . This process is completely reversible. It has been demonstrated that the repeated removal and addition of  $\text{Mg}^{2+}$  resulted in cycles of chromosome swelling and compaction.<sup>35</sup>

A report of the distribution of cations in nuclei and isolated mitotic chromosomes using secondary ion mass spectrometry<sup>216</sup> concluded that during the transition from interphase to mitosis,  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  concentrations increase 3 to 4 fold, reaching 18 mM and 11 mM respectively. Interestingly, mitotic chromosomes show axial enrichment in  $\text{Ca}^{2+}$ . As far as monovalent ions

are concerned,  $\text{Na}^+$  concentrations are close to those measured in the cytosol, whereas  $\text{K}^+$  concentrations are clearly higher in both interphasic and mitotic chromosomes. It is also well established that polyamines especially spermidine (3+) and spermine (4+)<sup>138,141</sup> are ubiquitous components with potent chromatin compaction properties.

## 6 Evolutionary perspectives

### 6.1 One problem, many solutions

Viruses, bacteria (eu- and archaeobacteria) and eukaryotes have all found ways to compact their genome, accumulating a spectacular amount of highly charged DNA into tiny volumes (see Table 1). Advantages of accumulating such a quantity of DNA with regard to the physical cost of compaction are not obvious, especially in organisms in which most of the DNA has not yet revealed its functional role. Hence, while bacteria genomes are very dense in transcribed sequences (intergenic regions are very small and there are no introns), this is definitely not the case for human genomes: one million base pairs of bacterial DNA contain 500 to 1000 genes whereas one million of human DNA contain only on average 10 genes. What may be the advantages of the presence of these intergenic regions for chromatin organization is still a debated question. It is noteworthy that many proteins bind to DNA through a shape-dependent recognition that depends on minor groove width and electrostatics. Hence, the molecular nature of DNA embeds a supplementary layer of information (distinct from the genetic information) that could dictate various important features such as nucleosome positioning and transcription factor binding. As such, DNA is subject to evolutionary constraint, even in non-coding regions.<sup>170</sup> As already discussed previously, various ways to compact DNA co-evolved in the different kingdoms, and prokaryotic and eukaryotic organisms may indeed not be as different as is generally believed with regard to their chromatin structure and properties.<sup>12</sup> For instance, the eukaryotic dinoflagellate *Cryptothecodinium cohnii* that lacks histone proteins encodes HCC proteins (histone-like proteins of *Cryptothecodinium cohnii*) that share homology with bacterial HU proteins.<sup>243</sup> Also, many bacteria contain homologues of the eukaryotic H1, such as Hc1 and Hc2 proteins found in Chlamydia,<sup>80</sup> whose nucleoid compaction capacity was further demonstrated.<sup>9</sup> *Bordetella pertussis* encodes two H1-like proteins (Bph1 and Bph2) and a H-NS-like protein (Bph3); *Pseudomonas aeruginosa* encodes the H1-like AlgP protein and five different H-NS-like proteins; *etc.* The reader interested in these comparisons between the various DNA benders/wrappers/bridgers found in different organisms should refer to the quite exhaustive review by Dame and colleagues.<sup>126</sup> From a phylogenetic point of view, the origin of histone H1 can be traced to bacteria whereas core histones are believed to originate from archaea.<sup>98,199</sup> H3 and H4 evolved at a far slower pace than H2A and H2B, reflecting their crucial role in the first steps of nucleosome formation.<sup>129</sup> Remarkably, further specializations of histone variants (mainly through differences in their DNA wrapping properties and specialized domains that regulate access to DNA) have evolved in some lineages to perform additional tasks, underlying the diverse functions that histones have acquired in evolution.<sup>219</sup>

## 6.2 The invention of the nucleus and epigenetics

The nucleus is a highly organized compartment containing several sub-compartments thought to offer various specific micro-environments that regulate chromatin state and transcriptional regulation.<sup>144,181</sup> How does this arise? The presence of histones and nucleosomal packaging of chromosomal DNA probably predates the development of a nuclear membrane.<sup>12</sup> Then, according to the coevolutionary theory of eukaryote origins, the fundamental innovations were the concerted origins of the endomembrane system and cytoskeleton, subsequently recruited to form the cell nucleus and coevolving mitotic apparatus.<sup>27,176,242</sup> Historically viewed as little more than a diffusion barrier between the cytoplasm and the nucleoplasm, the nuclear envelope is now known to have roles in the cell cycle, cytoskeletal stability, genome architecture, epigenetics, regulation of transcription, splicing and DNA replication.<sup>252</sup> From an electrostatic point of view, much is still unknown of the interplay between lipid binding, counterion displacement, and DNA condensation.<sup>142</sup> DNA methylation appears as a strong epigenetic regulator in all kingdoms of life, including bacteria.<sup>25</sup> The enzyme responsible for methylation of histones H3 and H4, the histone lysine methyltransferase (KMT), utilizes a catalytically active site called the SET domain. SET domains have been shown to be involved in eukaryotic chromatin modification by methylating some histone lysine residues (Kouzarides 2002 (ref. 111)). The identification of SET domain methyl-transferase proteins in bacteria raises the possibility that epigenetic inheritance in these organisms may also involve DNA-associated proteins.<sup>152</sup> The demonstration of a physical interaction between the chlamydial SET domain protein and histone-like proteins Hc1 and Hc2 (described in Section 6.1) provides some evidence of a similar regulatory mechanism in bacteria.

At the same time, proteomic approaches have unraveled many unknown acetylation sites in bacteria and archaea.<sup>212</sup> As a whole, it seems that common mechanisms in gene silencing (the programmed inhibition of the gene transcription) are shared by all life forms,<sup>247</sup> although the logic behind these mechanisms has been argued to be quite different.<sup>217</sup>

## 6.3 Concluding remarks

Far from having “explained” the role of electrostatics in the making and tuning of functionally active levels of compaction in living organisms, we have tried to review the various ways by which electrostatics intervenes in these key biological mechanisms. Coming back to Fig. 1, we can now briefly sum up the interplay between the different actors evoked throughout the paper. The mechanical properties of DNA and the DNA–DNA interaction are both strongly determined by its high line charge density, and are therefore directly tuned by the ion screening effect. These properties in turn control DNA bending and twisting flexibility, and therefore the formation of supercoiled structures like plectonemes or toroids. In Section 4.2 it has been stressed that the active (motor dependent) maintenance of unconstrained supercoiling is a crucial compacting mechanism in bacteria.

Highly charged multivalent ions which induce DNA–DNA attraction strongly facilitate compaction by allowing the

formation of bridges and liquid crystal ordering of dsDNAs. Moreover, these multivalent ions can *locally* modify the DNA structure, and in particular they can induce DNA bending. In this respect, they prefigure the second family of compacting agents, namely proteins. Indeed DNA binding proteins are multivalent ions that can bend DNA and moreover act as wrappers or bridgers. Unlike simple ions, DNA binding proteins have a new important feature: their specificity. Their “design” has been selected so as to bind DNA in specific sites, and – probably most important for the purpose of compaction – some of them contain specialized domains that can ensure dimerization, hence bridging of two dsDNAs specifically. Moreover DNA binding proteins can be bound to other proteins to form super-structures of increasing complexity. In eukaryotes, the nucleosome complex is the key component involved in the mechanism of compaction: the unconstrained supercoiling relevant to bacteria is replaced by constrained supercoiling, locally managed by protein–DNA interaction. Eukaryotic chromatin is a beautiful example of this “high-tech engineering”: an inhomogeneous, multiscale, tunable architecture, made of protein–DNA and DNA–DNA electrostatic interactions and tuned by small ions and post-translational electrostatic modifications.

## Acknowledgements

We thank Eric Le Cam for discussion and help with Fig. 5. AL, LS and JMV thank the Institut National de la Santé et de la Recherche Médicale, grant MICROMEGAS PC201104.

## References

- 1 M. Alilat, A. Sivolob, B. Revet and A. Prunell, Nucleosome dynamics. Protein and DNA contributions in the chiral transition of the tetrasome, the histone (H3–H4)<sub>2</sub> tetramer–DNA particle, *J. Mol. Biol.*, 1999, **291**, 815–841.
- 2 B. N. Ames and D. T. Dubin, The role of polyamines in the neutralization of bacteriophage deoxyribonucleic acid, *J. Biol. Chem.*, 1960, **235**, 769–775.
- 3 L. Aravind, L. M. Iyer and V. Anantharaman, The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism, *Genome Biol.*, 2003, **4**, R64.
- 4 L. Arcesi, G. La Penna and A. Perico, Generalized electrostatic model of the wrapping of DNA around oppositely charged proteins, *Biopolymers*, 2007, **86**, 127–135.
- 5 T. A. Azam, S. Hiraga and A. Ishihama, Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid, *Genes Cells*, 2000, **5**, 613–626.
- 6 T. A. Azam and A. Ishihama, Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity, *J. Biol. Chem.*, 1999, **274**, 33105–33113.
- 7 R. Balint and S. S. Cohen, The incorporation of radiolabeled polyamines and methionine into turnip yellow mosaic virus in protoplasts from infected plants, *Virology*, 1985, **144**, 181–193.
- 8 E. Bardez, N. C. Vy and T. Zemb, Counterion-driven sphere to cylinder transition in reverse micelles: a small angle X-ray scattering and conductometric study, *Langmuir*, 1995, **11**, 3374–3381.
- 9 C. E. Barry, S. F. Hayes and T. Hackstadt, Nucleoid condensation in *Escherichia coli* that express a chlamydial histone homolog, *Science*, 1992, **256**, 377–379.
- 10 C. G. Baumann, S. B. Smith, V. A. Bloomfield and C. Bustamante, Ionic effects on the elasticity of single DNA molecules, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 6185–6190.
- 11 C. Beloin, J. Jeusset, B. Révet, G. Mirambeau, F. Le Hégarat and E. Le Cam, Contribution of DNA conformation and topology in

- right-handed DNA wrapping by the *Bacillus subtilis* LrpC protein, *J. Biol. Chem.*, 2003, **278**, 5333–5342.
- 12 A. J. Bendich and K. Drlica, Prokaryotic and eukaryotic chromosomes: what's the difference?, *BioEssays*, 2000, **22**, 481–486.
- 13 B. Bertin, S. Mangenot, M. Renouard, D. Durand and F. Livolant, Structure and phase diagram of nucleosome core particles aggregated by multivalent cations, *Biophys. J.*, 2007, **93**, 3652–3663.
- 14 V. A. Bloomfield, DNA condensation, *Curr. Opin. Struct. Biol.*, 1996, **6**, 334–341.
- 15 V. A. Bloomfield, DNA condensation by multivalent cations, *Biopolymers*, 1997, **44**, 269–282.
- 16 B. A. Boggs, C. D. Allis and A. C. Chinault, Immunofluorescent studies of human chromosomes with antibodies against phosphorylated H1 histone, *Chromosoma*, 2000, **108**, 485–490.
- 17 K. Bohinc, A. Igli and S. May, Interaction between macroions mediated by divalent rod-like ions, *Europhys. Lett.*, 2004, **68**, 494–500.
- 18 K. Bohinc, J. M. A. Grime and L. Lue, The interactions between charged colloids with rod-like counterions, *Soft Matter*, 2012, **8**, 5679–5686.
- 19 E. Bouffartigues, M. Buckle, C. Badaut, A. Travers and S. Rimsky, H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing, *Nat. Struct. Mol. Biol.*, 2007, **14**, 441–448.
- 20 A. A. Brian, H. L. Frisch and L. S. Lerman, Thermodynamics and equilibrium sedimentation analysis of the close approach of DNA molecules and a molecular ordering transition, *Biopolymers*, 1981, **20**, 1305–1328.
- 21 D. F. Browning, D. C. Grainger and S. J. Busby, Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression, *Curr. Opin. Microbiol.*, 2010, **13**, 773–780.
- 22 S. S. Broyles and D. E. Pettijohn, Interaction of the *Escherichia coli* HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch, *J. Mol. Biol.*, 1986, **187**, 47–60.
- 23 H. Brutzer, N. Luzzietti, D. Klaue and R. Seidel, Energetics at the DNA supercoiling transition, *Biophys. J.*, 2010, **98**, 1267–1276.
- 24 M. Bustin and R. Reeves, High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function, *Prog. Nucleic Acid Res. Mol. Biol.*, 1996, **54**, 35–100.
- 25 J. Casades and D. Low, Epigenetic gene regulation in the bacterial world, *Microbiol. Mol. Biol. Rev.*, 2006, **70**, 830–856.
- 26 F. Catez, D. T. Brown, T. Misteli and M. Bustin, Competition between histone H1 and HMG proteins for chromatin binding sites, *EMBO Rep.*, 2002, **3**, 760–766.
- 27 T. Cavalier-Smith, Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution, *Biol. Direct*, 2010, **5**, 7.
- 28 Y. R. Chemla, K. Aathavan, J. Michaelis, S. Grimes, P. J. Jardine, D. L. Anderson and C. Bustamante, Mechanism of force generation of a viral DNA packaging motor, *Cell*, 2005, **122**, 683–692.
- 29 X. J. Chen and R. A. Butow, The organization and inheritance of the mitochondrial genome, *Nat. Rev. Genet.*, 2005, **6**, 815–825.
- 30 Y. Cheng, N. Korolev and L. Nordenskiöld, Similarities and differences in interaction of  $K^+$  and  $Na^+$  with condensed ordered DNA. A molecular dynamics computer simulation study, *Nucleic Acids Res.*, 2006, **34**, 686–696.
- 31 D. J. Clark and J. O. Thomas, Differences in the binding of H1 variants to DNA. Cooperativity and linker-length related distribution, *Eur. J. Biochem.*, 1988, **178**, 225–233.
- 32 N. Clauvelin, B. Audoly and S. Neukirch, Mechanical response of plectonemic DNA: an analytical solution, *Macromolecules*, 2008, **41**, 4479–4483.
- 33 N. Clauvelin, B. Audoly and S. Neukirch, Elasticity and electrostatics of plectonemic DNA, *Biophys. J.*, 2009, **2009**, 3716–3723.
- 34 J. A. Cognet, C. Pakleza, D. Cherny, E. Delain and E. Le Cam, Static curvature and flexibility measurements of DNA with microscopy. A simple renormalization method, its assessment by experiment and simulation, *J. Mol. Biol.*, 1999, **285**, 997–1009.
- 35 A. Cole, Chromosome structure, *Theor. Biophys.*, 1967, **1**, 305–307.
- 36 K. D. Collins, Charge density-dependent strength of hydration and biological structure, *Biophys. J.*, 1997, **72**, 65–76.
- 37 D. J. Cummings, V. A. Chapman and S. S. DeLong, An electron microscopic study of lambda and lambda-dg bacteriophage in thin sections, *J. Mol. Biol.*, 1965, **14**, 418–422.
- 38 R. Cortini, A. A. Kornyshev, D. J. Lee and S. Leikin, Electrostatic braiding and homologous pairing of DNA double helices, *Biophys. J.*, 2011, **101**, 875–884.
- 39 *The Biological Chemistry of Magnesium*, ed. J. A. Cowan, Wiley, New York, 1995.
- 40 C. Crosio, G. M. Fimia, R. Loury, M. Kimura, Y. Okano, H. Zhou, S. Sen, C. D. Allis and P. Sassone-Corsi, Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases, *Mol. Cell Biol.*, 2002, **22**, 874–885.
- 41 S. Cuylen and C. H. Haering, Deciphering condensin action during chromosome segregation, *Trends Cell Biol.*, 2011, **21**, 552–559.
- 42 V. Dahirel, F. Paillusson, M. Jardat, M. Barbi and J. M. Victor, Nonspecific DNA-protein interaction: why proteins can diffuse along DNA, *Phys. Rev. Lett.*, 2009, **102**, 228101.
- 43 R. T. Dame, The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin, *Mol. Microbiol.*, 2005, **56**, 858–870.
- 44 R. T. Dame, C. Wyman and N. Goosen, H-NS mediated compaction of DNA visualised by atomic force microscopy, *Nucleic Acids Res.*, 2000, **28**, 3504–3510.
- 45 C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond, Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution, *J. Mol. Biol.*, 2002, **319**, 1097–1113.
- 46 S. de los Rios and J. J. Perona, Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly, *J. Mol. Biol.*, 2007, **366**, 1589–1602.
- 47 J. De Rouchey, V. A. Parsegian and D. C. Rau, Cation charge dependence of the forces driving DNA assembly, *Biophys. J.*, 2010, **99**, 2608–2615.
- 48 S. C. Dillon and C. J. Dorman, Bacterial nucleoid-associated proteins, nucleoid structure and gene expression, *Nat. Rev. Microbiol.*, 2011, **8**, 185–195.
- 49 C. J. Dorman, H-NS: a universal regulator for a dynamic genome, *Nat. Rev. Microbiol.*, 2004, **2**, 391–400.
- 50 C. J. Dorman, H-NS, the genome sentinel, *Nat. Rev. Microbiol.*, 2007, **5**, 157–161.
- 51 J. A. Downs, Histone H3 K56 acetylation, chromatin assembly, and the DNA damage checkpoint, *DNA Repair*, 2008, **7**, 2020–2024.
- 52 J. Dubochet and M. Noll, Nucleosome arcs and helices, *Science*, 1978, **202**, 280–286.
- 53 J. G. Duguid and V. A. Bloomfield, Electrostatic effects on the stability of condensed DNA in the presence of divalent cations, *Biophys. J.*, 1996, **70**, 2838–2846.
- 54 W. C. Earnshaw and S. R. Casjens, DNA packaging by the double-stranded DNA bacteriophages, *Cell*, 1980, **21**, 319–331.
- 55 W. C. Earnshaw and U. K. Laemmli, Architecture of metaphase chromosomes and chromosome scaffolds, *J. Cell Biol.*, 1983, **96**, 84–93.
- 56 T. H. Eickbush and E. N. Moudrianakis, The compaction of DNA helices into either continuous supercoils or folded-fiber rods and toroids, *Cell*, 1978, **13**, 295–306.
- 57 M. Elstov, K. M. MacLellan, K. Maeshima, A. S. Frangakis and J. Dubochet, Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes *in situ*, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 19732–19737.
- 58 Y. Fang and J. H. Hoh, Cationic silanes stabilize intermediates in DNA condensation, *FEBS Lett.*, 1999, **459**, 173–176.
- 59 Y. Fang and J. H. Hoh, Characterization of DNA condensates by atomic force microscopy, *Methods Mol. Med.*, 2001, **65**, 149–158.
- 60 A. T. Fenley, D. A. Adams and A. V. Onufriev, Charge state of the globular histone core controls stability of the nucleosome, *Biophys. J.*, 2010, **99**, 1577–1585.
- 61 R. P. Fisher, T. Lisowsky, M. A. Parisi and D. A. Clayton, DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein, *J. Biol. Chem.*, 1992, **267**, 3358–3367.
- 62 T. M. Fletcher and J. C. Hansen, The nucleosomal array: structure/function relationships, *Crit. Rev. Eukaryotic Gene Expression*, 1996, **6**, 149–188.
- 63 S. Forth, C. Deufel, M. Y. Sheinin, B. Daniels, J. P. Sethna and M. D. Wang, Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules, *Phys. Rev. Lett.*, 2008, **100**, 148301.

- 64 D. Frenkiel-Krispin, I. Ben-Avraham, J. Englander, E. Shimoni, S. G. Wolf and A. Minsky, Nucleoid restructuring in stationary-state bacteria, *Mol. Microbiol.*, 2004, **51**, 395–405.
- 65 J. Fresnais, C. Lavelle and J. F. Berret, Nanoparticle aggregation controlled by desalting kinetics, *J. Phys. Chem. C*, 2009, **113**, 16371–16379.
- 66 D. N. Fuller, J. P. Rickgauer, P. J. Jardine, S. Grimes, D. L. Anderson and D. E. Smith, Ionic effects on viral DNA packaging and portal motor function in bacteriophage phi 29, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 11245–11250.
- 67 D. Furlong, H. Swift and B. Roizman, Arrangement of herpesvirus deoxyribonucleic acid in the core, *J. Virol.*, 1972, **10**, 1071–1074.
- 68 I. Gállego, P. Castro-Hartmann, J. M. Caravaca, S. Caño and J. R. Daban, Dense chromatin plates in metaphase chromosomes, *Eur. Biophys. J.*, 2009, **38**, 503–522.
- 69 B. A. Garcia, C. M. Barber, S. B. Hake, C. Ptak, F. B. Turner, S. A. Busby, J. Shabanowitz, R. G. Moran, C. D. Allis and D. F. Hunt, Modifications of human histone H3 variants during mitosis, *Biochemistry*, 2005, **44**, 13202–13213.
- 70 F. Geinguenau, V. Calandrini, B. Cayrol, J. Teixeira, C. Mayer, J. Liquier, C. Lavelle and V. Arluison, Conformational transition of DNA bound to Hfq probed by infrared spectroscopy, *Phys. Chem. Chem. Phys.*, 2011, **13**, 1222–1229.
- 71 M. Gellert, K. Mizuuchi, M. H. O’Dea and H. A. Nash, DNA gyrase: an enzyme that introduces superhelical turns into DNA, *Proc. Natl. Acad. Sci. U. S. A.*, 1976, **73**, 3872–3876.
- 72 W. Gibson and B. Roizman, Compartmentalization of spermine and spermidine in the herpes simplex virion, *Proc. Natl. Acad. Sci. U. S. A.*, 1971, **68**, 2818–2821.
- 73 L. C. Gosule and J. A. Schellman, Compact form of DNA induced by spermidine, *Nature*, 1976, **259**, 333–335.
- 74 H. Goto, Y. Tomono, K. Ajiro, H. Kosako, M. Fujita, M. Sakurai, K. Okawa, A. Iwamatsu, T. Okigaki, T. Takahashi and M. Inagaki, Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation, *J. Biol. Chem.*, 1999, **274**, 25543–25549.
- 75 H. Goto, Y. Yasui, E. A. Nigg and M. Inagaki, Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation, *Genes Cells*, 2002, **7**, 11–17.
- 76 F. A. Goytisolo, S. E. Gerchman, X. Yu, C. Rees, V. Graziano, V. Ramakrishnan and J. O. Thomas, Identification of two DNA-binding sites on the globular domain of histone H5, *EMBO J.*, 1996, **15**, 3421–3429.
- 77 R. A. Grayling, K. A. Bailey and J. N. Reeve, DNA binding and nuclease protection by the Hmf histones from the hyperthermophilic archaeon *Methanothermus fervidus*, *Extremophiles*, 1997, **1**, 79–88.
- 78 M. Gueroult, D. Picot, J. Abi-Ghanem, B. Hartmann and M. Baaden, How cations can assist DNase I in DNA binding and hydrolysis, *PLoS Comput. Biol.*, 2010, **6**, e1001000.
- 79 L. Guldbbrand, B. Jönson, H. Wennerström and P. Linse, Electrical double layer forces: a Monte Carlo study, *J. Chem. Phys.*, 1984, **80**, 2221–2228.
- 80 T. Hackstadt, W. Baehr and Y. Ying, *Chlamydia trachomatis* developmentally regulated protein is homologous to eukaryotic histone H1, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 3937–3941.
- 81 P. J. Hagerman, Flexibility of DNA, *Annu. Rev. Biophys. Biophys. Chem.*, 1988, **17**, 265–286.
- 82 J. M. Harp, B. L. Hanson, D. E. Timm and G. J. Bunick, Asymmetries in the nucleosome core particle at 2.5 Å resolution, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2000, **56**, 1513–1534.
- 83 A. H. Hassan, P. Prochasson, K. E. Neely, S. C. Galasinski, M. Chandy, M. J. Carrozza and J. L. Workman, Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes, *Cell*, 2002, **111**, 369–379.
- 84 M. Haynes, R. A. Garret and W. B. Gratzer, Structure of nuclei acid-poly base complexes, *Biochemistry*, 1970, **9**, 4410–4416.
- 85 B. Heddi, N. Foloppe, E. Hantz and B. Hartmann, The DNA structure responds differently to physiological concentrations of K<sup>+</sup> or Na<sup>+</sup>, *J. Mol. Biol.*, 2007, **368**, 1403–1411.
- 86 M. J. Hendzel, Y. Wei, M. A. Mancini, A. Van Hooser, T. Ranalli, B. R. Brinkley, D. P. Bazett-Jones and C. D. Allis, Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation, *Chromosoma*, 1997, **106**, 348–360.
- 87 H. Higashibata, M. A. Siddiqui, M. Takagi, T. Imanaka and S. Fujiwara, Surface histidine residue of archaeal histone affects DNA compaction and thermostability, *FEMS Microbiol. Lett.*, 2003, **224**, 17–22.
- 88 V. F. Holmes and N. R. Cozzarelli, Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 1322–1324.
- 89 B. Hribar and V. Vlachy, Evidence of electrostatic attraction between equally charged macroions induced by divalent counterions, *J. Phys. Chem. B*, 1997, **101**, 3457–3459.
- 90 N. G. Hunt and J. E. Hearst, Elastic model of DNA supercoiling in the infinite length limit, *J. Chem. Phys.*, 1991, **95**, 9329–9336.
- 91 M. Imbert, B. Laine, N. Helbecque, J. P. Mornon, J. P. Hénichart and P. Sautière, Conformational study of the chromosomal protein MCI from the archaeobacterium *Methanosarcina barker*, *Biochim. Biophys. Acta*, 1990, **1038**, 346–354.
- 92 D. Ivanov and K. A. Nasmyth, Topological interaction between cohesin rings and a circular minichromosome, *Cell*, 2005, **122**, 849–860.
- 93 S. Jones, H. P. Shanahan, H. M. Berman and J. M. Thornton, Using electrostatic potentials to predict DNA-binding sites on DNA-binding proteins, *Nucleic Acids Res.*, 2003, **31**, 7189–7198.
- 94 S. Jones, P. van Heyningen, H. M. Berman and J. M. Thornton, Protein-DNA interactions: a structural analysis, *J. Mol. Biol.*, 1999, **287**, 877–896.
- 95 M. Kanduc, J. Dobnikar and R. Podgornik, Counterion-mediated electrostatic interactions between helical molecules, *Soft Matter*, 2009, **5**, 868–877.
- 96 M. Kanduc, A. Naji and R. Podgornik, Counterion-mediated weak and strong coupling electrostatic interaction between like-charged cylindrical dielectrics, *J. Chem. Phys.*, 2010, **132**, 224703.
- 97 M. Kanduc, M. Trulsson, A. Naji, Y. Burak, J. Forsman and R. Podgornik, Weak- and strong-coupling electrostatic interactions between asymmetrically charged planar surfaces, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2008, **78**, 061105.
- 98 H. E. Kasinsky, J. D. Lewis, J. B. Dacks and J. Ausio, Origin of H1 linker histones, *FASEB J.*, 2001, **15**, 34–42.
- 99 J. Granot, J. Feigon and D. R. Kearns, Interactions of DNA with divalent metal ions. I. <sup>31</sup>P-NMR studies, *Biopolymers*, 1982, **21**, 181–201.
- 100 J. Granot and D. R. Kearns, Interactions of DNA with divalent metal ions. II. Proton relaxation enhancement studies, *Biopolymers*, 1982, **21**, 203–218.
- 101 J. Granot and D. R. Kearns, Interactions of DNA with divalent metal ions. III. Extent of metal binding: experiment and theory, *Biopolymers*, 1982, **21**, 219–232.
- 102 J. Granot, N. Assa-Munt and D. R. Kearns, Interactions of DNA with divalent metal ions. IV. Competitive studies of Mn<sup>2+</sup> binding to AT- and GC-rich DNAs, *Biopolymers*, 1982, **21**, 873–884.
- 103 N. Kepper, D. Foethke, R. Stehr, G. Wedemann and K. Rippe, Nucleosome geometry and internucleosomal interactions control the chromatin fiber conformation, *Biophys. J.*, 2008, **95**, 3692–3705.
- 104 K. Kimura, V. V. Rybenkov, N. J. Crisona, T. Hirano and N. R. Cozzarelli, 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation, *Cell*, 1999, **98**, 239–248.
- 105 J. Kindt, S. Tzllil, A. Ben-Shaul and W. M. Gelbart, DNA packaging and ejection forces in bacteriophage, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 13671–13674.
- 106 S. M. Klimenko, T. I. Tikchonenko and V. M. Andreev, Packing of DNA in the head of bacteriophage T2, *J. Mol. Biol.*, 1967, **23**, 523–533.
- 107 A. A. Kornyshev, D. J. Lee, S. Leikin and A. Wynveen, Structure and interactions of biological helices, *Rev. Mod. Phys.*, 2007, **79**, 943–996.
- 108 A. A. Kornyshev and S. Leikin, Electrostatic zipper-motif for DNA aggregation, *Phys. Rev. Lett.*, 1999, **82**, 4138–4141.
- 109 A. A. Kornyshev and S. Leikin, Sequence recognition in the pairing of DNA duplexes, *Phys. Rev. Lett.*, 2001, **86**, 3666–3669.
- 110 N. Korolev, O. V. Vorontsova and L. Nordenskiöld, Physicochemical analysis of electrostatic foundation for DNA-protein interactions in chromatin transformations, *Prog. Biophys. Mol. Biol.*, 2007, **95**, 23–49.

- 111 T. Kouzarides, Histone methylation in transcriptional control, *Curr. Opin. Genet. Dev.*, 2002, **12**, 198–209.
- 112 T. Kouzarides, Chromatin modifications and their function, *Cell*, 2007, **128**, 693–705.
- 113 U. K. Laemmli, Characterization of DNA condensates induced by poly(ethylene oxide) and polylysine, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 4288–4292.
- 114 B. Lang, N. Blot, E. Bouffartigues, M. Buckle, M. Geertz, C. O. Gualerzi, R. Mavathur, G. Muskhelishvili, C. L. Pon, S. Rimsky, S. Stella, M. M. Babu and A. Travers, High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes, *Nucleic Acids Res.*, 2007, **35**, 6330–6337.
- 115 A. W.-C. Lau, Fluctuation and correlation effects in electrostatics of highly-charged surfaces, PhD thesis, University of California at Santa Barbara, 2000.
- 116 A. W.-C. Lau, P. Pincus, H. A. Levine and D. Fertig, Electrostatic attraction of coupled Wigner crystals: finite temperature effects, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2001, **63**, 051604.
- 117 C. Lavelle and A. Prunell, Chromatin polymorphism and the nucleosome superfamily: a genealogy, *Cell Cycle*, 2007, **6**, 2113–2119.
- 118 E. Le Cam, F. Culard, E. Larquet, E. Delain and J. A. Cognet, DNA bending induced by the archaeobacterial histone-like protein MC1, *J. Mol. Biol.*, 1999, **285**, 1011–1121.
- 119 A. Leforestier and F. Livolant, Structure of toroidal DNA collapsed inside the phage capsid, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 9157–91162.
- 120 M. A. Lever, J. P. Th'ng, X. Sun and M. J. Hendzel, Rapid exchange of histone H1.1 on chromatin in living human cells, *Nature*, 2000, **408**, 873–876.
- 121 P. Linse and V. Lobaskin, Electrostatic attraction and phase separation in solutions of like-charged colloidal particles, *Phys. Rev. Lett.*, 1999, **83**, 4208–4211.
- 122 F. Livolant, Ordered phases of DNA *in vivo* and *in vitro*, *Phys. A*, 1991, **176**, 117–137.
- 123 F. Livolant, S. Mangenot, A. Leforestier, A. Bertin, M. de Frutos, E. Raspaud and D. Durand, Are liquid crystalline properties of nucleosomes involved in chromosome structure and dynamics, *Philos. Trans. R. Soc., A*, 2006, **364**, 2615–2633.
- 124 K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent and T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature*, 1997, **389**, 251–260.
- 125 M. S. Luijsterburg, M. C. Noom, G. J. Wuite and R. T. Dame, The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective, *J. Struct. Biol.*, 2006, **156**, 262–272.
- 126 M. S. Luijsterburg, M. F. White, R. van Driel and R. T. Dame, The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes, *Crit. Rev. Biochem. Mol. Biol.*, 2008, **43**, 393–418.
- 127 R. Lurz, M. Grote, J. Dijk, R. Reinhardt and B. Dobrinski, Electron microscopic study of DNA complexes with proteins from the archaeobacterium *Sulfolobus acidocaldarius*, *EMBO J.*, 1986, **5**, 3715–3721.
- 128 K. Maeshima, S. Hihara and M. Eltsov, Chromatin structure: does the 30-nm fibre exist *in vivo*?, *Curr. Opin. Cell Biol.*, 2010, **22**, 291–297.
- 129 H. S. Malik and S. Henikoff, Phylogenomics of the nucleosome, *Nat. Struct. Biol.*, 2003, **10**, 882–891.
- 130 S. Mangenot, A. Leforestier, D. Durand and F. Livolant, X-ray diffraction characterization of the dense phases formed by nucleosome core particles, *Biophys. J.*, 2003, **84**, 2570–2584.
- 131 G. S. Manning, The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides, *Q. Rev. Biophys.*, 1978, **11**, 179–246.
- 132 G. S. Manning, The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force, *Biophys. J.*, 2006, **91**, 3607–3616.
- 133 G. S. Manning, Counterion condensation theory of attraction between like charges in the absence of multivalent counterions, *Eur. Phys. J. E*, 2011, **34**, 132.
- 134 S. Mantelli, P. Muller, S. Harlepp and M. Maaloum, Conformational analysis and estimation of the persistence length of DNA using atomic force microscopy in solution, *Soft Matter*, 2011, **7**, 3412–3416.
- 135 F. Marc, K. Sandman, R. Lurz and J. N. Reeve, Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling, *J. Biol. Chem.*, 2000, **277**, 30879–30886.
- 136 T. J. Maresca, B. S. Freedman and R. Heald, Histone H1 is essential for mitotic chromosome architecture and segregation in *Xenopus laevis* egg extracts, *J. Cell Biol.*, 2005, **169**, 859–869.
- 137 A. Margalit, A. Brachner, J. Gotzmann, R. Foisner and Y. Gruenbaum, Barrier-to-autointegration factor – a BAFfling little protein, *Trends Cell Biol.*, 2007, **17**, 202–208.
- 138 R. Marquet, P. Colson and C. Houssier, The condensation of chromatin and histone H1-depleted chromatin by spermine, *J. Biomol. Struct. Dyn.*, 1986, **4**, 205–218.
- 139 R. Marquet, P. Colson, A. M. Matton, C. Houssier, M. Thiry and G. Goessens, Comparative study of the condensation of chicken erythrocyte and calf thymus chromatin by di- and multivalent cations, *J. Biomol. Struct. Dyn.*, 1988, **5**, 839–857.
- 140 R. Marquet and C. Houssier, Thermodynamics of cation-induced DNA condensation, *J. Biomol. Struct. Dyn.*, 1991, **9**, 159–167.
- 141 R. Marquet, C. Houssier and E. Fredericq, An electro-optical study of the mechanisms of DNA condensation induced by spermine, *Biochim. Biophys. Acta*, 1985, **825**, 365–374.
- 142 D. Matulis, I. Rouzina and V. A. Bloomfield, Thermodynamics of cationic lipid binding to DNA and DNA condensation: roles of electrostatics and hydrophobicity, *J. Am. Chem. Soc.*, 2002, **124**, 7331–7342.
- 143 S. May, A. Igli, J. Rescic, S. Maset and K. Bohinc, Bridging like-charged macroions through long divalent rod-like ions, *J. Phys. Chem. B*, 2008, **112**, 1685–1692.
- 144 L. Meldi and J. H. Brickner, Compartmentalization of the nucleus, *Trends Cell Biol.*, 2011, **21**, 701–708.
- 145 A. Miller, B. Yang, T. Foster and A. L. Kirchmaier, Proliferating cell nuclear antigen and ASF1 modulate silent chromatin in *Saccharomyces cerevisiae* via lysine 56 on histone H3, *Genetics*, 2008, **179**, 793–809.
- 146 T. Misteli, A. Gunjan, R. Hock, M. Bustin and D. T. Brown, Dynamic binding of histone H1 to chromatin in living cells, *Nature*, 2000, **408**, 877–881.
- 147 A. Moreira and R. R. Netz, Simulations of counterions at charged plates, *Eur. Phys. J. E*, 2002, **8**, 33–58.
- 148 S. A. Morris, B. Rao, B. A. Garcia, S. B. Hake, R. L. Diaz, J. Shabanowitz, D. F. Hunt, C. D. Allis, J. D. Lieb and B. D. Strahl, Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification, *J. Biol. Chem.*, 2007, **282**, 7632–7640.
- 149 F. Mosconi, J. F. Allemand, D. Bensimon and V. Croquette, Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers, *Phys. Rev. Lett.*, 2009, **102**, 078301.
- 150 J. Mozziconacci, C. Lavelle, M. Barbi, A. Lesne and J. M. Victor, A physical model for the condensation and decondensation of eukaryotic chromosomes, *FEBS Lett.*, 2006, **580**, 368–372.
- 151 S. Mujtaba, L. Zeng and M. M. Zhou, Structure and acetyl-lysine recognition of the bromodomain, *Oncogene*, 2007, **26**, 5521–5527.
- 152 M. Murata, Y. Azuma, K. Miura, M. A. Rahman, M. Matsutani, M. Aoyama, H. Suzuki, K. Sugi and M. Shirai, Chlamydial SET domain protein functions as a histone methyltransferase, *Microbiology*, 2007, **153**, 585–592.
- 153 F. V. T. Murphy, R. M. Sweet and M. E. Churchill, The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition, *EMBO J.*, 1999, **18**, 6610–6618.
- 154 D. R. Musgrave, K. M. Sandman and J. N. Reeve, DNA binding by the archaeal histone Hmf results in positive supercoiling, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 10397–10401.
- 155 A. Naji, A. Arnold, C. Holm and R. R. Netz, Attraction and unbinding of like-charged rods, *Eur. Phys. J. E*, 2004, **67**, 130–136.
- 156 A. Naji and R. R. Netz, Attraction of like-charged macroions in the strong-coupling limit, *Eur. Phys. J. E*, 2004, **13**, 43–59.
- 157 R. R. Netz, Electrostatics of counter-ions at and between planar charged walls: from Poisson-Boltzmann to the strong-coupling theory, *Eur. Phys. J. E*, 2001, **5**, 557–574.
- 158 R. R. Netz and H. Orland Beyond, Poisson Boltzmann: fluctuations effects and correlation functions, *Eur. Phys. J. E*, 2000, **1**, 203–214.

- 159 J. C. Neu, Wall-mediated forces between like-charged bodies in an electrolyte, *Phys. Rev. Lett.*, 1999, **82**, 1072–1075.
- 160 S. Neukirch, Extracting DNA twist rigidity from experimental supercoiling data, *Phys. Rev. Lett.*, 2004, **93**, 198107.
- 161 S. Neukirch and J. F. Marko, Analytical description of extension, torque, and supercoiling radius of a stretched twisted DNA, *Phys. Rev. Lett.*, 2011, **106**, 138104.
- 162 T. Odijk, Osmotic compaction of supercoiled DNA into a bacterial nucleoid, *Biophys. Chem.*, 1998, **73**, 23–29.
- 163 D. E. Olins and A. L. Olins, Model nucleohistones: the interaction of F1 and F2al histones with native T7 DNA, *J. Mol. Biol.*, 1971, **57**, 437–455.
- 164 T. Oshima, T. Moriya and Y. Terui, Identification, chemical synthesis, and biological functions of unusual polyamines produced by extreme thermophiles, *Methods Mol. Biol.*, 2011, **720**, 81–111.
- 165 T. A. Owen-Hughes, G. D. Pavitt, D. S. Santos, J. M. Sidebotham, C. S. J. Hulton, J. C. D. Hulton and C. F. Higgins, The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression, *Cell*, 1992, **71**, 255–265.
- 166 F. Paillusson, M. Barbi and J. M. Victor, Poisson-Boltzmann for oppositely charged bodies: an explicit derivation, *J. Chem. Phys.*, 2009, **107**, 1379–1391.
- 167 F. Paillusson, V. Dahirel, M. Jardat, J. M. Victor and M. Barbi, Effective interaction between charged nanoparticles and DNA, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12603–12613.
- 168 C. Q. Pan, S. E. Finkel, S. E. Cramton, J. A. Feng, D. S. Sigman and R. C. Johnson, Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts, *J. Mol. Biol.*, 1996, **264**, 675–695.
- 169 F. Paquet, F. Culard, F. Barbault, J. C. Maurizot and G. Lancelot, NMR solution structure of the archaeobacterial chromosomal protein MCl reveals a new protein fold, *Biochemistry*, 2004, **43**, 14971–14978.
- 170 S. C. Parker and T. D. Tullius, DNA shape, genetic codes, and evolution, *Curr. Opin. Struct. Biol.*, 2011, **21**, 342–347.
- 171 V. A. Parsegian and D. Gingell, On the electrostatic interaction across a salt solution between two bodies bearing unequal charges, *Biophys. J.*, 1972, **12**, 1192–1204.
- 172 T. T. Paull, M. J. Haykinson and R. C. Johnson, The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures, *Genes Dev.*, 1993, **7**, 1521–1534.
- 173 J. R. Paulson and U. K. Laemmli, The structure of histone-depleted metaphase chromosomes, *Cell*, 1977, **12**, 817–828.
- 174 S. L. Pereira, R. A. Grayling, R. Lurz and J. N. Reeve, Archaeal nucleosomes, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 12633–12637.
- 175 B. Pérez-Cadahía, B. Drobnic and J. R. Davie, H3 phosphorylation: dual role in mitosis and interphase, *Biochem. Cell Biol.*, 2009, **87**, 695–709.
- 176 J. Postberg, H. J. Lipps and T. Cremer, Evolutionary origin of the cell nucleus and its functional architecture, *Essays Biochem.*, 2010, **48**, 1–24.
- 177 P. K. Purohit, M. M. Inamdar, P. D. Grayson, T. M. Squires, J. Kondev and R. Phillips, Forces during bacteriophage DNA packaging and ejection, *Biophys. J.*, 2005, **88**, 851–866.
- 178 P. K. Purohit, J. Kondev and R. Phillips, Mechanics of DNA packaging in viruses, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 3173–3178.
- 179 X. Qiu, V. A. Parsegian and D. C. Rau, Divalent counterion-induced condensation of triple-strand DNA, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 21482–21486.
- 180 J. Ray and G. S. Manning, An attractive force between two rodlike polyions mediated by the sharing of condensed counterions, *Langmuir*, 1994, **10**, 2450–2461.
- 181 I. Rajapakse and M. Groudine, On emerging nuclear order, *J. Cell Biol.*, 2011, **192**, 711–721.
- 182 R. Reeves, Nuclear functions of the HMG proteins, *Biochim. Biophys. Acta*, 2010, **1799**, 3–14.
- 183 D. Rhodes and R. A. Laskey, Assembly of nucleosomes and chromatin *in vitro*, *Methods Enzymol.*, 1989, **170**, 575–585.
- 184 K. E. Richards, P. C. Williams and R. Calendar, The ionic strength of DNA packing within bacteriophage heads, *J. Mol. Biol.*, 1973, **78**, 255–259.
- 185 S. C. Riemer and V. A. Bloomfield, Packaging of DNA in bacteriophage heads: some considerations on energetics, *Biopolymers*, 1978, **17**, 785–794.
- 186 S. Rimsky, Structure of the histone-like protein H-NS and its role in regulation and genome superstructure, *Curr. Opin. Microbiol.*, 2004, **7**, 109–114.
- 187 F. Ritort, Single-molecule experiments in biological physics: methods and application, *J. Phys.: Condens. Matter*, 2006, **18**, R351–R358.
- 188 P. J. J. Robinson, L. Fairall, V. A. T. Huynh and D. Rhodes, EM measurements define the dimensions of the “30-nm” chromatin fiber: evidence for a compact, interdigitated structure, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **103**, 6506–6511.
- 189 A. Routh, S. Sandin and D. Rhodes, Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 8872–8877.
- 190 I. Rouzina and V. A. Bloomfield, Macroion attraction due to electrostatic correlation between screening counterions. I. Mobile surface-adsorbed ions and diffuse ion cloud, *J. Phys. Chem.*, 1996, **100**, 9977–9989.
- 191 I. Rouzina and V. A. Bloomfield, DNA bending by small, mobile multivalent cations, *Biophys. J.*, 1998, **74**, 3152–3164.
- 192 V. V. Rybenkov, N. R. Cozzarelli and A. V. Vologodskii, Probability of DNA knotting and the effective diameter of the DNA double helix, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 5307–5311.
- 193 V. V. Rybenkov, A. V. Vologodskii and N. R. Cozzarelli, The effect of ionic conditions on DNA helical repeat, effective diameter and free energy of supercoiling, *Nucleic Acids Res.*, 1997, **25**, 1412–1418.
- 194 J. E. Sader and D. Y. C. Chan, Long-range electrostatic attractions between identically charged particles in confined geometries: an unresolved problem, *J. Colloid Interface Sci.*, 1999, **213**, 268–269.
- 195 L. Samaj and E. Trizac, Wigner-crystal formulation of strong-coupling theory for counterions near planar charged interfaces, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2011, **84**, 041401.
- 196 L. Samaj and E. Trizac, Counterions at highly charged interfaces: from one plate to like-charge attraction, *Phys. Rev. Lett.*, 2011, **106**, 078301.
- 197 K. Sandman, J. A. Krzycki, B. Dobrinski, R. Lurz and J. N. Reeve, Hmf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermobacter feravidus*, is most closely related to histones, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 5788–5791.
- 198 K. Sandman, F. P. Perler and J. N. Reeve, Histone-encoding genes from *Pyrococcus*: evidence for members of the Hmf family of archaeal histones in a non-methanogenic Archaeon, *Curr. Opin. Microbiol.*, 2006, **285**, 1011–1021.
- 199 K. Sandman and J. N. Reeve, Archaeal histones and the origin of the histone fold, *Gene*, 1994, **150**, 207–208.
- 200 A. Savelyev and G. A. Papoian, Electrostatic, steric, and hydration interactions favor Na<sup>+</sup> condensation around DNA compared with K<sup>+</sup>, *J. Am. Chem. Soc.*, 2006, **28**, 14506–14518.
- 201 A. Savelyev, C. K. Materese and G. A. Papoian, Is DNA’s rigidity dominated by electrostatic or nonelectrostatic interactions?, *J. Am. Chem. Soc.*, 2011, **133**, 19290–19293.
- 202 A. Savelyev, Do monovalent mobile ions affect DNA’s flexibility at high salt content?, *Phys. Chem. Chem. Phys.*, 2012, **14**, 2250–2254.
- 203 M. P. Scheffer, M. Eltsov, J. Bednar and A. S. Frangakis, Nucleosomes stacked with aligned dyad axes are found in native compact chromatin *in vitro*, *J. Struct. Biol.*, 2012, **178**(2), 207–214.
- 204 M. P. Scheffer, M. Eltsov and A. S. Frangakis, Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 16992–16997.
- 205 H. Schiessel, The physics of chromatin, *J. Phys.: Condens. Matter*, 2003, **15**, R699–R774.
- 206 R. Schneider, R. Lurz, G. Lüder, C. Tolksdorf, A. Travers and G. Muskhelishvili, An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA, *Nucleic Acids Res.*, 2001, **29**, 5107–5114.
- 207 B. I. Shklovskii, Wigner crystal model of counterion induced bundle formation of rodlike polyelectrolytes, *Phys. Rev. Lett.*, 1999, **82**, 3268–3271.
- 208 M. Shogren-Knaak, H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie and C. L. Peterson, Histone H4-K16 acetylation controls chromatin structure and protein interactions, *Science*, 2006, **311**, 844–847.



- 209 D. Skoko, B. Wong, R. C. Johnson and J. F. Marko, Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes, *Biochemistry*, 2004, **43**, 13867–13874.
- 210 D. E. Smith, S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson and C. Bustamante, The bacteriophage straight phi29 portal motor can package DNA against a large internal force, *Nature*, 2001, **413**, 748–752.
- 211 D. J. Soares, F. Marc and J. N. Reeve, Conserved eukaryotic histone-fold residues substituted into an Archaeal histone increase DNA affinity but reduce complex flexibility, *J. Bacteriol.*, 2003, **185**, 3453–3457.
- 212 J. Soppa, Protein acetylation in archaea, bacteria, and eukaryotes, *Archaea*, 2010, **2010**, 820681.
- 213 J. Stavans and A. Oppenheim, DNA-protein interactions and bacterial chromosome architecture, *Phys. Biol.*, 2006, **3**, R1–R10.
- 214 A. Stein, Reconstitution of chromatin from purified components, *Methods Enzymol.*, 1989, **170**, 585–603.
- 215 D. Stigter, Interactions of highly charged colloidal cylinders with applications to double-stranded DNA, *Biopolymers*, 1977, **16**, 1435–1448.
- 216 R. Strick, P. L. Strissel, K. Gavrilov and R. Levi-Setti, Cation-chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes, *J. Cell Biol.*, 2001, **155**, 899–910.
- 217 K. Struhl, Fundamentally different logic of gene regulation in prokaryotes and eukaryotes, *Cell*, 1999, **98**, 1–4.
- 218 K. K. Swinger, K. M. Lemberg, Y. Zhang and P. A. Rice, Flexible DNA bending in HU-DNA cocrystal structures, *EMBO J.*, 2003, **22**, 3749–3760.
- 219 P. B. Talbert and S. Henikoff, Histone variants – ancient wrap artists of the epigenome, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 264–275.
- 220 H. Tanaka, N. Goshima, K. Kohno, Y. Kano and F. Imamoto, Properties of DNA-binding of HU heterotypic and homotypic dimers from *Escherichia coli*, *J. Biochem.*, 1993, **113**, 568–572.
- 221 S. G. Taneva, I. G. Muñoz, G. Franco, J. Falces, I. Arregi, A. Muga, G. Montoya, M. A. Urbaneja and S. Bañuelos, Activation of nucleoplasmin, an oligomeric histone chaperone, challenges its stability, *Biochemistry*, 2008, **47**, 13897–13906.
- 222 A. Tapias, G. Lopez and S. Ayora, *Bacillus subtilis* LrpC is a sequence-independent DNA-binding and DNA-bending protein which bridges DNA, *Nucleic Acids Res.*, 2000, **28**, 552–559.
- 223 P. Thaw, S. E. Sedelnikova, T. Muranova, S. Wiese, S. Ayora, J. C. Alonso, A. B. Brinkman, J. Akerboom, J. van der Oost and J. B. Rafferty, Structural insight into gene transcriptional regulation and effector binding by the Lrp/AsnC family, *Nucleic Acids Res.*, 2006, **34**, 1439–1449.
- 224 F. Thoma, T. Koller and A. Klug, Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin, *J. Cell Biol.*, 1979, **83**, 403–427.
- 225 J. O. Thomas and A. A. Travers, HMGI and 2, and related ‘architectural’ DNA-binding proteins, *Trends Biochem. Sci.*, 2001, **26**, 167–174.
- 226 J. V. Tjeertes, K. M. Miller and S. P. Jackson, Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells, *EMBO J.*, 2009, **28**, 1878–1889.
- 227 Y. Timsit and P. Varnai, Helical chirality: a link between local interactions and global topology in DNA, *PLoS One*, 2010, **5**, e9326.
- 228 B. A. Todd, V. A. Parsegian, A. Shirahata, T. J. Thomas and D. C. Rau, Attractive forces between cation condensed DNA double helices, *Biophys. J.*, 2008, **94**, 4775–4782.
- 229 A. A. Travers, Priming the nucleosome: a role for HMGB proteins?, *EMBO Rep.*, 2003, **4**, 131–136.
- 230 A. Travers and G. Muskhelishvili, Bacterial chromatin, *Curr. Opin. Genet. Dev.*, 2005, **15**, 507–514.
- 231 D. J. Tremethick, Higher-order structures of chromatin: the elusive 30 nm fiber, *Cell*, 2007, **128**, 651–654.
- 232 A. E. Tupper, T. A. Owen-Hughes, D. W. Ussery, D. S. Santos, D. J. Ferguson, J. M. Sidebotham, J. C. Hinton and C. F. Higgins, The chromatin-associated protein H-NS alters DNA topology *in vitro*, *EMBO J.*, 1994, **13**, 258–268.
- 233 S. Tzllil, J. T. Kindt, W. M. Gelbart and A. Ben-Shaul, Forces and pressures in DNA packaging and release from viral capsids, *Biophys. J.*, 2003, **84**, 1616–1627.
- 234 J. Ubbink and T. Odijk, Electrostatic-undulatory theory of plectonemically supercoiled DNA, *Biophys. J.*, 1999, **76**, 2502–2519.
- 235 J. P. Valleau, R. Ivkov and G. M. Torrie, Colloid stability: the forces between charged surfaces in an electrolyte, *J. Chem. Phys.*, 1991, **95**, 520–532.
- 236 J. van Noort, S. Verbrugge, N. Goosen, C. Dekker and R. T. Dame, Dual architectural roles of HU: formation of flexible hinges and rigid filaments, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 6969–6974.
- 237 M. Vasilescu, D. Angelescu, H. Caldararu, M. Almgren and A. Khan, Fluorescence study on the size and shape of sodium dodecyl sulphate-aluminium salt micelles, *Colloids Surf., A*, 2004, **235**, 57–64.
- 238 T. Vora, A. K. Hottes and S. Tavazoie, Protein occupancy landscape of a bacterial genome, *Mol. Cell*, 2009, **35**, 247–253.
- 239 T. Waldmann, M. Baack, N. Richter and C. Gruss, Structure-specific binding of the proto-oncogene protein DEK to DNA, *Nucleic Acids Res.*, 2003, **31**, 7003–7310.
- 240 R. Warrier, B. R. Linger, B. L. Golden and R. J. Kuhn, Role of sindbis virus capsid protein region II in nucleocapsid core assembly and encapsidation of genomic RNA, *J. Virol.*, 2008, **82**, 4461–4470.
- 241 J. Widom and A. Klug, Structure of the 3000 Å chromatin filament: X-ray diffraction from oriented samples, *Cell*, 1985, **43**, 207–213.
- 242 K. L. Wilson and S. C. Dawson, Evolution: functional evolution of nuclear structure, *J. Cell Biol.*, 2011, **195**, 171–181.
- 243 J. T. Wong, D. C. New, J. C. Wong and V. K. Hung, Histone-like proteins of the dinoflagellate *Cryptocodinium cohnii* have homologies to bacterial DNA-binding proteins, *Eukaryotic Cell*, 2003, **2**, 646–650.
- 244 H. Wong, J. M. Victor and J. Mozziconacci, An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length, *PLoS One*, 2007, **2**, e877.
- 245 F. Xu, K. Zhang and M. Grunstein, Acetylation in histone H3 globular domain regulates gene expression in yeast, *Cell*, 2005, **121**, 375–385.
- 246 Y. Yoshikawa, K. Yoshikawa and T. Kanbe, Formation of a giant toroid from long duplex DNA, *Langmuir*, 1999, **15**, 4085–4088.
- 247 N. A. Youngson, S. Chong and E. Whitelaw, Gene silencing is an ancient means of producing multiple phenotypes from the same genotype: common mechanisms and functions in epigenetic processes can be seen throughout all life forms, *BioEssays*, 2010, **33**, 95–99.
- 248 T. Y. Yu and J. Shaefer, REDOR NMR characterization of DNA packaging in bacteriophage T4, *J. Mol. Biol.*, 2008, **382**, 1031–1042.
- 249 G. Zanchetta, F. Giavazza, M. Nakata, M. Buscaglia, R. Cerbinoia, N. A. Clark and T. Bellini, Right-handed double-helix ultrashort DNA yields chiral nematic phases with both right- and left-handed director twist, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **107**, 17497–17502.
- 250 H. Zhang and J. F. Marko, Maxwell relations for single-DNA experiments: monitoring protein binding and double-helix torque with force-extension measurements, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2008, **77**, 031916.
- 251 Y. B. Zhou, S. E. Gerchman, V. Ramakrishnan, A. Travers and S. Muyldermans, Position and orientation of the globular domain of linker histone H5 on the nucleosome, *Nature*, 1998, **395**, 402–405.
- 252 N. Zuleger, M. I. Robson and E. C. Schirmer, The nuclear envelope as a chromatin organizer, *Nucleus*, 2011, **2**, 339–349.

---

# Bibliographie

---

- [1] C. Bouchiat, M.D. Wang, J.-F. Allemand, T. Strick, S.M. Block, and V. Croquette. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical Journal*, 76(1) :409 – 413, 1999. [vii](#), [viii](#), [37](#), [43](#), [44](#)
- [2] J. F. Marko and E. Siggia. Stretching dna. *Macromolecules*, 28 :8759–8770, 1995. [vii](#), [viii](#), [37](#), [43](#), [44](#)
- [3] Francesco Mosconi, Jean François Allemand, David Bensimon, and Vincent Croquette. Measurement of the torque on a single stretched and twisted dna using magnetic tweezers. *Phys. Rev. Lett.*, 102 :078301, Feb 2009. [viii](#), [45](#), [46](#), [47](#), [48](#), [49](#), [50](#)
- [4] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker. Dynamics of dna supercoils. *Science*, 338(6103) :94–97, 2012. [viii](#), [48](#)
- [5] Karolin Luger and Timothy J Richmond. Dna binding within the nucleosome core. *Current Opinion in Structural Biology*, 8(1) :33 – 40, 1998. [ix](#), [55](#), [57](#), [63](#)
- [6] Alfredo Celedon, Ilana M. Nodelman, Bridget Wildt, Rohit Dewan, Peter Searson, Denis Wirtz, Gregory D. Bowman, and Sean X. Sun. Magnetic tweezers measurement of single molecule torque. *Nano Letters*, 9(4) :1720–1725, 2009. PMID : 19301859. [x](#), [66](#), [69](#)
- [7] A. Vologodskii. Dna extension under the action of an external force. *Macromolecules*, 27 :5623–5625, sep 1994. [9](#), [10](#)
- [8] SB Smith, L Finzi, and C Bustamante. Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads. *Science*, 258(5085) :1122–1126, 1992. [9](#), [43](#)
- [9] O. Kratky and G. Porod. Röntgenuntersuchung gelöster fadenmoleküle. *Rec. Trav. Chim. Pays-Bas*, 68 :1106–1123, 1949. [10](#), [19](#), [39](#)
- [10] Walter Reisner, Keith J. Morton, Robert Riehn, Yan Mei Wang, Zhaoning Yu, Michael Rosen, James C. Sturm, Stephen Y. Chou, Erwin Frey, and Robert H. Austin. Statics and dynamics of single dna molecules confined in nanochannels. *Phys. Rev. Lett.*, 94 :196101, May 2005. [10](#), [89](#)
- [11] Hongmei Jian, Alexander V. Vologodskii, and Tamar Schlick. A combined wormlike-chain and bead model for dynamic simulations of long linear dna. *Journal of Computational Physics*, 136(1) :168 – 179, 1997. [10](#)

- 
- [12] Konstantin Klenin, Holger Merlitz, and Jörg Langowski. A brownian dynamics program for the simulation of linear and circular dna and other wormlike chain polyelectrolytes. *Biophysical Journal*, 74(2) :780 – 788, 1998. [10](#), [11](#)
- [13] Patrick J. Heath, John A. Gebe, Stuart A. Allison, and J. Michael Schurr. Comparison of analytical theory with brownian dynamics simulations for small linear and circular dnas. *Macromolecules*, 29(10) :3583–3596, 1996. [10](#)
- [14] Daniel A. Beard and Tamar Schlick. Inertial stochastic dynamics. i. long-time-step methods for langevin dynamics. *The Journal of Chemical Physics*, 112(17) :7313–7322, 2000. [11](#)
- [15] Daniel A. Beard and Tamar Schlick. Inertial stochastic dynamics. ii. influence of inertia on slow kinetic processes of supercoiled dna. *The Journal of Chemical Physics*, 112(17) :7323–7338, 2000. [11](#)
- [16] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J.C Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints : molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3) :327 – 341, 1977. [11](#), [12](#)
- [17] Shuichi Miyamoto and Peter A. Kollman. Settle : An analytical version of the shake and rattle algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8) :952–962, 1992. [12](#)
- [18] Timothy R. Forester and William Smith. Shake, rattle, and roll : Efficient constraint algorithms for linked rigid bodies. *Journal of Computational Chemistry*, 19(1) :102–111, 1998. [12](#)
- [19] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. Lincs : A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12) :1463–1472, 1997. [12](#)
- [20] W.F. van Gunsteren, H.J.C. Berendsen, and J.A.C. Rullmann. Stochastic dynamics for molecules with constraints. *Molecular Physics*, 44(1) :69–95, 1981. [12](#)
- [21] TOBY O. WHITE, GIOVANNI CICCOTTI, and JEAN-PIERRE HANSEN. Brownian dynamics with constraints. *Molecular Physics*, 99(24) :2023–2036, 2001. [12](#)
- [22] Roger Edberg, Denis J. Evans, and G. P. Morriss. Constrained molecular dynamics : Simulations of liquid alkanes with a new algorithm. *The Journal of Chemical Physics*, 84(12) :6933–6939, 1986. [12](#)
- [23] A. Witkin, M. Gleicher, and W. Welch. Interactive dynamics. *Computer Graphics*, 24(4), 1994. [14](#)
- [24] C. Lacoursière. *Ghosts and Machines : Regularized Variational Methods for Interactive Simulations of Multibodies with Dry Frictional Contacts*. PhD thesis, Umeå University, 2007. [14](#), [95](#)
- [25] R. Smith. Open dynamic engine, 2008. <http://www.ode.org>. [14](#), [16](#), [29](#)
- [26] K. Erleben. *Stable, Robust, and Versatile Multibody Dynamics Animation*. PhD thesis, The Department of Computer Science, University of Copenhagen, Denmark, 2005. [14](#), [16](#)
- [27] F. Sebastin Grassia. Practical parameterization of rotations using the exponential map. *J. Graph. Tools*, 3(3) :29–48, mar 1998. [15](#)
- [28] D. M. Young. *Iterative methods for solving partial difference equations of elliptical type*. PhD thesis, Havard University, 1950. [16](#)
-

- [29] R. W. Cottle and G. B. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications*, 1 :103–125, 1968. 16
- [30] Matthias Müller, Bruno Heidelberger, Marcus Hennix, and John Ratcliff. Position based dynamics. *Journal of Visual Communication and Image Representation*, 18(2) :109–118, 2007. 16
- [31] David Baraff. Linear-time dynamics using lagrange multipliers. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, SIGGRAPH '96*, pages 137–146, New York, NY, USA, 1996. ACM. 16
- [32] J. Baumgarte. Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1(1) :1 – 16, 1972. 16
- [33] D. Fiser. Library for collision detection between complex shapes. <http://libccd.danfis.cz>. 16
- [34] Miklós Bergou, Max Wardetzky, Stephen Robinson, Basile Audoly, and Eitan Grinspun. Discrete elastic rods. *ACM Transactions on Graphics (SIGGRAPH)*, 27(3) :63 :1–63 :12, aug 2008. 19
- [35] Raymond E. Goldstein, Thomas R. Powers, and Chris H. Wiggins. Viscous nonlinear dynamics of twist and writhe. *Phys. Rev. Lett.*, 80 :5232–5235, Jun 1998. 19
- [36] M. D. Frank-Kamenetskii and V. V. Lukashin, A. V. and Anshelevich. Torsionnal and bending rigidity of the double helix from data on small dna rings. *J. Biomol. Struct. Dynam.*, 2(1005), 1985. 20, 35
- [37] Dirk Stigter. Interactions of highly charged colloidal cylinders with applications to double-stranded dna. *Biopolymers*, 16(7) :1435–1448, 1977. 23
- [38] A. A. Brian, H. L. Frisch, and L. S. Lerman. Thermodynamics and equilibrium sedimentation analysis of the close approach of dna molecules and a molecular ordering transition. *Biopolymers*, 20(6) :1305–1328, 1981. 23
- [39] Jeffrey J. Delrow, John A. Gebe, and J. Michael Schurr. Comparison of hard-cylinder and screened coulomb interactions in the modeling of supercoiled dnas. *Biopolymers*, 42(4) :455–470, 1997. 23
- [40] N. Clauvelin, B. Audoly, and S. Neukirch. Elasticity and electrostatics of plectonemic dna. *Biophysical Journal*, 96(9) :3716–3723, 2009. 23, 45
- [41] Paul J. Hagerman. Flexibility of dna. *Annual Review of Biophysics and Biophysical Chemistry*, 17(1) :265–286, 1988. 23
- [42] Christoph G. Baumann, Steven B. Smith, Victor A. Bloomfield, and Carlos Bustamante. Ionic effects on the elasticity of single dnamolecules. *Proceedings of the National Academy of Sciences*, 94(12) :6185–6190, 1997. 23
- [43] Sébastien Neukirch. Extracting dna twist rigidity from experimental supercoiling data. *Phys. Rev. Lett.*, 93(19) :198107, Nov 2004. 23, 45
- [44] Pascal Carrivain, Axel Cournac, Christophe Lavelle, Annick Lesne, Julien Mozziconacci, Fabien Paillusson, Laurence Signon, Jean-Marc Victor, and Maria Barbi. Electrostatics of dna compaction in viruses, bacteria and eukaryotes : functional insights and evolutionary perspective. *Soft Matter*, 8 :9285–9301, 2012. 24

- 
- [45] P. Hubbard. Angular velocity of a nonspherical body undergoing rotational brownian motion. *Phys. Rev. A*, 15(1) :329–336, 1977. [24](#)
- [46] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1) :014101, 2007. [24](#), [27](#)
- [47] Giovanni Bussi and Michele Parrinello. Stochastic thermostats : comparison of local and global schemes. *Computer Physics Communications*, 179(1-3) :26 – 29, 2008. Special issue based on the Conference on Computational Physics 2007 - CCP 2007. [24](#), [26](#)
- [48] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8) :3684–3690, 1984. [26](#), [27](#)
- [49] Giovanni Bussi and Michele Parrinello. Accurate sampling using langevin dynamics. *Phys. Rev. E*, 75 :056707, May 2007. [27](#)
- [50] A. Hadjidimos. Successive overrelaxation (sor) and related methods. *Journal of Computational and Applied Mathematics*, 123(12) :177 – 199, 2000. Numerical Analysis 2000. Vol. III : Linear Algebra. [29](#)
- [51] W. Kahan. *Gauss-Seidel Methods of Solving Large Systems of Linear Equations*. PhD thesis, University of Toronto, 1958. [29](#)
- [52] E. Catto. Box 2d, a 2d physics engine for games, 2012. <http://box2d.org>. [30](#)
- [53] M. Matsumoto and T. Nishimura. Mersenne twister : A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. on Modeling and Computer Simulation*, 8(1) :3–30, 1998. [32](#)
- [54] George and Mervin E. Muller. A note on the generation of random normal deviates. *Ann. Math. Stat.*, 29(2) :610–611, 1958. [32](#)
- [55] H. Wong. *Modèle moléculaire multi-échelles de l'architecture de la chromatine : nouveaux outils, nouvelles perspectives*. PhD thesis, Laboratoire de Physique Théorique de la Matière Condensée (CNRS), Université Pierre et Marie Curie, 2008. [33](#), [66](#)
- [56] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, 2002. [40](#)
- [57] John F. Marko. Dna under high tension : Overstretching, undertwisting, and relaxation dynamics. *Phys. Rev. E*, 57 :2134–2149, Feb 1998. [43](#)
- [58] F. B. Fuller. Decomposition of the linking number of a closed ribbon : A problem from molecular biology. *Proceedings of The National Academy of Sciences*, 75 :3557–3561, 1978. [43](#)
- [59] James H. White. Self-linking and the gauss integral in higher dimensions. *American Journal of Mathematics*, 91(3) :pp. 693–728, 1969. [43](#)
- [60] A. E. H. Love. *A Treatise on the Mathematical Theory of Elasticity*. Dover, New York, 4 edition, 1944. [44](#)
- [61] N. Clauvelin, B. Audoly, and S. Neukirch. Mechanical response of plectonemic dna : An analytical solution. *Macromolecules*, 41(12) :4479–4483, 2008. [45](#)
- [62] Sébastien Neukirch and John F. Marko. Analytical description of extension, torque, and supercoiling radius of a stretched twisted dna. *Phys. Rev. Lett.*, 106 :138104, Apr 2011. [45](#)
- [63] Hergen Brutzer, Nicholas Luzzietti, Daniel Klaue, and Ralf Seidel. Energetics at the dna supercoiling transition. *Biophysical Journal*, 98(7) :1267 – 1276, 2010. [46](#), [48](#), [50](#)
-

- [64] Scott Forth, Christopher Deufel, Maxim Y. Sheinin, Bryan Daniels, James P. Sethna, and Michelle D. Wang. Abrupt buckling transition observed during the plectoneme formation of individual dna molecules. *Phys. Rev. Lett.*, 100 :148301, Apr 2008. [46](#), [48](#), [50](#)
- [65] Houyin Zhang and John F. Marko. Maxwell relations for single-dna experiments : Monitoring protein binding and double-helix torque with force-extension measurements. *Phys. Rev. E*, 77 :031916, Mar 2008. [48](#), [49](#)
- [66] Christopher Maffeo, Robert Schöpflin, Hergen Brutzer, René Stehr, Aleksei Aksimentiev, Gero Wedemann, and Ralf Seidel. Dna-dna interactions in tight supercoils are described by a small effective charge density. *Phys. Rev. Lett.*, 105 :158101, Oct 2010. [48](#)
- [67] John F. Marko. Torque and dynamics of linking number relaxation in stretched supercoiled dna. *Phys. Rev. E*, 76(2) :021926, Aug 2007. [48](#)
- [68] V V Rybenkov, N R Cozzarelli, and A V Vologodskii. Probability of dna knotting and the effective diameter of the dna double helix. *Proc. Natl. Acad. Sci. USA*, 90(11) :5307–5311, 1993. [49](#)
- [69] Valentin V. Rybenkov, Alexander V. Vologodskii, and Nicholas R. Cozzarelli. The effect of ionic conditions on dna helical repeat, effective diameter and free energy of supercoiling. *Nucleic Acids Research*, 25 :1412–1418, 1997. [49](#)
- [70] A Bucceri, K Kapitza, and F Thoma. Rapid accessibility of nucleosomal dna in yeast on a second time scale. *The EMBO Journal*, 25(5) :3123–32, 2006. [57](#)
- [71] Aurélien Bancaud, Gaudeline Wagner, Natalia Conde e Silva, Christophe Lavelle, Hua Wong, Julien Mozziconacci, Maria Barbi, Andrei Sivolob, Eric Le Cam, Liliane Mouawad, Jean-Louis Viovy, and Jean-Marc Victor. Nucleosome chiral transition under positive torsional stress in single chromatin fibers. *Molecular cell*, 27 :135–147, 2007/07/06. [57](#), [58](#), [65](#)
- [72] M. Kruithof and J. van Noort. Hidden markov analysis of nucleosome unwrapping under force. *Biophysical Journal*, 96(9) :3708 – 3715, 2009. [57](#), [63](#)
- [73] Ramona Ettig, Nick Kepper, Rene Stehr, Gero Wedemann, and Karsten Rippe. Dissecting dna-histone interactions in the nucleosome by molecular dynamics simulations of dna unwrapping. *Biophysical Journal*, 101(8) :1999 – 2008, 2011. [57](#)
- [74] Vera Böhm, Aaron R. Hieb, Andrew J. Andrews, Alexander Gansen, Andrea Rocker, Katalin Tóth, Karolin Luger, and Jörg Langowski. Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic Acids Research*, 2010. [57](#), [64](#)
- [75] E. Praly. *Étude du mécanisme d'action des facteurs de remodelage de la chromatine, à l'échelle de la molécule unique*. PhD thesis, École Normale Supérieure, Département de Physique, Laboratoire de physique statistique, Paris, 2009. [63](#)
- [76] Aurélien Bancaud, Natalia Conde e Silva, Maria Barbi, Gaudeline Wagner, Jean-Francois Allemand, Julien Mozziconacci, Christophe Lavelle, Vincent Croquette, Jean-Marc Victor, Ariel Prunell, and Jean-Louis Viovy. Structural plasticity of single chromatin fibers revealed by torsional manipulation. *Nat Struct Mol Biol*, 13(5) :444–450, 2006/05//print. [65](#)
- [77] Christophe Bécavin, Maria Barbi, Jean-Marc Victor, and Annick Lesne. Transcription within condensed chromatin : Steric hindrance facilitates elongation. *Biophysical Journal*, 98(5) :824 – 833, 2010. [66](#)

- 
- [78] A. Grosberg and A. Khokhlov. *Statistical Physics of Macromolecules*. AIP Press, New York, 1994. [73](#)
- [79] J. Dekker. Mapping *in vivo* chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J. Biol. Chem*, 283(50) :34532–34540, 2008. [74](#)
- [80] K. Bystricky, P. Heun, L. Gehlen, J. Langowski, and S. Gasser. Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. *PNAS*, 101(47) :16495–16500, 2004. [75](#), [76](#)
- [81] Hua Wong, Hervé Marie-Nelly, Sébastien Herbert, Pascal Carrivain, Hervé Blanc, Romain Koszul, Emmanuelle Fabre, and Christophe Zimmer. A predictive computational model of the dynamic 3d interphase yeast nucleus. *Current Biology*, (0) :–, 2012. [76](#), [90](#)
- [82] S. Alexander. Adsorption of chain molecules with a polar head a scaling description. *J. Phys. France*, 38(8) :983–987, 1977. [76](#)
- [83] P. G. de Gennes. Conformations of polymers attached to an interface. *Macromolecules*, 13(5) :1069–1075, 1980. [76](#)
- [84] S. T. Milner. Polymer brushes. *Science*, 251(4996) :905–914, 1991. [76](#)
- [85] Guillaume J. Filion, Joke G. van Bommel, Ulrich Braunschweig, Wendy Talhout, Jop Kind, Lucas D. Ward, Wim Brugman, Ins J. de Castro, Ron M. Kerkhoven, Harmen J. Bussemaker, and Bas vanSteensel. Systematic protein location mapping reveals five principal chromatin types in drosophila cells, 2010. [90](#)
- [86] Mariano Barbieri, Mita Chotalia, James Fraser, Liron-Mark Lavitas, Josée Dostie, Ana Pombo, and Mario Nicodemi. Complexity of chromatin folding is captured by the strings and binders switch model. *Proceedings of the National Academy of Sciences*, 109(40) :16173–16178, 2012. [90](#)
- [87] [90](#)
- [88] Anna Delest, Tom Sexton, and Giacomo Cavalli. Polycomb : a paradigm for genome organization from one to three dimensions. *Current Opinion in Cell Biology*, 24(3) :405 – 414, 2012. Nucleus and gene expression. [90](#)
- [89] Thierry Cheutin and Giacomo Cavalli. Progressive polycomb assembly on h3k27me3 compartments generates polycomb bodies with developmentally regulated motion. *PLoS Genet*, 8(1) :e1002465, 01 2012. [90](#)
- [90] Job Dekker, Karsten Rippe, Martijn Dekker, and Nancy Kleckner. Capturing chromosome conformation. *Science*, 295(5558) :1306–1311, 2002. [90](#)
- [91] Zhijun Duan, Mirela Andronescu, Kevin Schutz, Sean McIlwain, Yoo Jung Kim, Choli Lee, Jay Shendure, Stanley Fields, C. Anthony Blau, and William S. Noble. A three-dimensional model of the yeast genome. *Nature*, 465 :363–367, 2010. [90](#)
- [92] Erez Lieberman-Aiden, Nynke L. van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragooczy, Agnes Telling, Ido Amit, Bryan R. Lajoie, Peter J. Sabo, Michael O. Dorschner, Richard Sandstrom, Bradley Bernstein, M. A. Bender, Mark Groudine, Andreas Gnirke, John Stamatoyannopoulos, Leonid A. Mirny, Eric S. Lander, and Job Dekker. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950) :289–293, 2009. [90](#)

- [93] A. Yu. Grosberg, S. K. Nechaev, and E. I. Shakhnovich. The role of topological constraints in the kinetics of collapse of macromolecules. *J. Phys. France*, 49(12) :2095–2100, 1988. [90](#)
- [94] E. Hairer. Geometric integration of ordinary differential equations on manifolds. *BIT Numerical Mathematics*, 41 :996–1007, 2001. 10.1023/A :1021989212020. [95](#)
- [95] Ernst Hairer, Christian Lubich, and Gerhard Wanner. Geometric numerical integration illustrated by the symplectic method. *Acta Numerica*, 12 :399–450, 2003. [95](#)



**Résumé :**

Durant ces dernières années, la modélisation en physique est restée aveugle aux développements de disciplines sœurs; en particulier la mécanique ou plutôt la robotique qui développe des outils puissants comme la cinématique inverse et les moteurs physiques (ou moteurs de jeux). Ces techniques sont couramment employées dans les jeux vidéo pour des résultats très réalistes. Je propose ici de montrer que nous pouvons simuler des assemblages d'ADN et de protéines (fibre de chromatine et à plus grande échelle des chromosomes) comme des systèmes articulés avec un moteur physique. Je montre aussi qu'il est possible d'étendre le thermostat local de Langevin à un thermostat global pour accélérer l'échantillonnage de l'espace des configurations du système ADN-protéines dans l'ensemble canonique. Ce nouveau thermostat est particulièrement intéressant lorsqu'il est utilisé avec un moteur physique. Plus précisément, je montrerai que la simulation que j'ai développée reproduit les résultats expérimentaux de manipulation de molécules uniques sous pinces magnétiques et permet de faire des prédictions. Enfin, cette simulation offre des perspectives intéressantes pour la modélisation des noyaux d'organismes allant de la drosophile à l'humain et la compréhension des toutes dernières données sur l'architecture des génomes.

**Title :** From game engines to the physics of chromosomes

**Abstract :**

During these last years, modelling in physics remained blind to the advances of sisters disciplines, especially mechanics and robotics, that have developed powerful tools in rigid body dynamics, soft body dynamics, and fluid dynamics. Techniques as inverse kinematic and physics engines (or game engines) are nowadays commonly used in video games for very realistic results. I show here how it is possible to simulate assemblies of DNA and proteins (chromatin fibres, and, at larger scale, chromosomes) as articulated system, by means of physics engines. I also show how to replace the standard local Langevin thermostat by a global one that achieves a much faster sampling of the configuration space of the system (canonical ensemble). This new thermostat results to be of particular interest when used with physics engines. The simulations that I have performed with these tools well reproduce magnetic tweezers experimental results and allows to make predictions. Finally, this simulation technique offers interesting perspectives for the modelling of whole cell nuclei for different organisms, ranging from *Drosophila* to humans, and provides therefore a precious tool for the interpretation of the most recent experimental data on genome architecture.

**Mots clés :** Moteurs de jeux, thermostat de Langevin global, simulation, pinces magnétiques, ADN, nucléosome, fibre de chromatine, chromosome

**Keywords :** Game engines, global Langevin thermostat, simulation, magnetic tweezers, DNA, nucleosome, chromatin fibre, chromosome