



Cite this: *Soft Matter*, 2016, 12, 9436

# Effects of molecular crowding and confinement on the spatial organization of a biopolymer

Chanil Jeon,<sup>a</sup> Youngkyun Jung<sup>\*b</sup> and Bae-Yeun Ha<sup>\*ac</sup>

A chain molecule can be entropically collapsed in a crowded medium in a free or confined space. Here, we present a unified view of how molecular crowding collapses a flexible polymer in three distinct spaces: free, cylindrical, and (two-dimensional) slit-like. Despite their seeming disparities, a few general features characterize all these cases, even though the  $\phi_c$ -dependence of chain compaction differs between the two cases:  $a > a_c$  and  $a < a_c$ , where  $\phi_c$  is the volume fraction of crowders,  $a$  is the monomer size, and  $a_c$  is the crowder size. For  $a > a_c$  (applicable to a coarse-grained model of bacterial chromosomes), chain size depends on the ratio  $a\phi_c/a_c$ , and “full” compaction occurs universally at  $a\phi_c/a_c \approx 1$ ; for  $a_c > a$  (relevant for protein folding), it is controlled by  $\phi_c$  alone and crowding has a modest effect on chain size in a cellular environment ( $\phi_c \approx 0.3$ ). Also for a typical parameter range of biological relevance, molecular crowding can be viewed as effectively reducing the solvent quality, independent of confinement.

Received 23rd May 2016,  
 Accepted 22nd October 2016

DOI: 10.1039/c6sm01184e

[www.rsc.org/softmatter](http://www.rsc.org/softmatter)

## 1 Introduction

Molecular crowding is a key factor in governing several biological processes, including chromosome organization, gene regulation, protein folding/aggregation, molecular reactions, and cell growth.<sup>1–12</sup> In particular, it has emerged as a dominant player in organizing bacterial chromosomes,<sup>5–8</sup> in a way that is desirable for their functions (*e.g.*, accessibility of genes to proteins and clustering of active-transcription sites).<sup>10,11</sup> Conceptually, the entropy of crowders favors the compaction of a chain molecule.<sup>6–8</sup> As illustrated in Fig. 1, molecular crowding induces entropic (depletion) forces between monomers,<sup>13–15</sup> responsible for chain collapse.

Thanks to recent efforts, several key features of chain collapse by molecular crowding have begun to emerge.<sup>8,16–20</sup> For instance, a computational approach has shown the interplay between crowder size and density in collapsing a flexible chain in a cylindrical space.<sup>16</sup> The compaction in this case appears to be continuous, similar to a corresponding unconfined case.<sup>17,18</sup> More recently, the interdependence of chain length and compaction has been discussed.<sup>20</sup> For stiff chains like DNA molecules, however, the geometry of a confined space can dictate the nature of compaction: abrupt in a cylindrical space

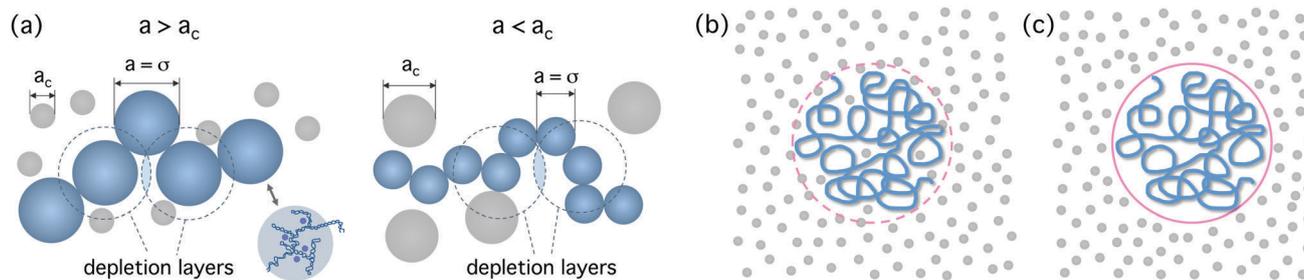
but continuous in a free or two-dimensional slit-like space.<sup>19</sup> *E. coli* chromosomes trapped in cylindrical channels were shown to undergo a similar abrupt transition by molecular crowding, as suggested by a coexistence of extended and collapsed states.<sup>8</sup> As pointed out,<sup>21</sup> the origin of abruptness in this case<sup>8</sup> may differ from that in DNA collapse.<sup>19</sup> How the structural heterogeneity of the chromosome is implicated in its compaction is not yet clear (see Fig. 2). As discussed later, molecular crowding will collapse a heterogeneous chain non-uniformly with a collapsed section coexisting with more loosely organized sections, similar to what was seen in the *E. coli* chromosome experiments<sup>8</sup> (see Section 4).

Ironically, the problem of a “simple” (bead-spring) polymer in a crowded medium is deceptively simple but remains to be further explored. Indeed, a number of basic questions will have to be answered. For instance, how does confinement modify the way a polymer responds to molecular crowding? In the earlier observation of abrupt *vs.* continuous DNA compaction,<sup>19</sup> crowders (polymeric ones) initially elongate DNA molecules in a cylindrical space by effectively reducing the cylinder diameter, even though at high concentrations they eventually collapse the DNA along the long axis of the cylinder.<sup>19</sup> This makes the compaction more cooperative or abrupt (only in a cylindrical space). However, this behavior has not been seen with flexible chains.<sup>16</sup> Is molecular crowding intrinsically sensitive to the geometry of a confined space? Or do there exist general features of molecular crowding, independent of crowder details or the geometry of a confined space? Finally, can molecular crowding be mimicked correctly by effectively reducing the solvent quality?

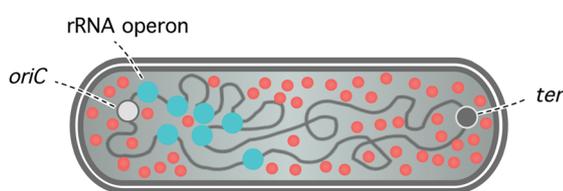
<sup>a</sup> Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada. E-mail: byha@uwaterloo.ca

<sup>b</sup> Supercomputing Center, Korea Institute of Science and Technology Information, Daejeon 34141, Korea. E-mail: yjung@kisti.re.kr

<sup>c</sup> School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea



**Fig. 1** Physical origin of depletion forces (a), and chain collapse by depletion forces (b) and (c). (a) Association of monomers in dark blue, resulting in a partial overlap of depletion layers, is favored by the entropy of crowders, *i.e.*, spheres in grey. The two cases are compared:  $a > a_c$  (left) and  $a < a_c$  (right), where  $a$  is the monomer size and  $a_c$  the crowder size. If  $a > a_c$ , a monomer can be surrounded by several crowders. Depletion forces can be safely considered as reducing the excluded volume  $v$  of monomers. In both cases, the resulting depletion forces will collapse the chain molecule as indicated in (b). The chain-enveloping volume (dashed or solid line in magenta) is permeable (b) or impermeable to crowders (c). The dependence of  $R$  on  $\phi_c$  will be different between (b) and (c). In (b), the action of depletion forces is expected to be (more) local; but in (c), it will reflect the shape of the chain (thus the geometry of a confined space).



**Fig. 2** Schematics of the *E. coli* chromosome. Ribosomal RNA (rRNA) operons (big spheres in cyan) are mostly concentrated near *oriC*. Molecular crowding can influence both the global and local organization of a heterogeneous polymer such as the *E. coli* chromosome. For simplicity, topological complexities (*e.g.*, multi-fork or “branched-donut”<sup>22</sup>) are not shown.

Here, we characterize molecular crowding in three distinct spaces: free, cylindrical, and (two-dimensional) slit-like. To this end, we combine molecular dynamics (MD) simulations and theoretical arguments. A flexible polymer is a “cleaner” system in the sense that it does not experience initial elongation seen with DNA.<sup>16,21</sup> In an effort to present a more complete picture, we will explore a wide range of parameters. Let  $a$  and  $a_c$  denote the size of monomers and crowders, respectively. In our simulations,  $a_c$  ranges between  $0.2a$  and  $20a$ , and the number of monomers can be as large as  $N = 2000$ .<sup>†</sup>

The small- $a_c$  case ( $a > a_c$ ) may represent a coarse-grained model of the bacterial chromosome,<sup>16–18</sup> in which each monomer represents a ‘structural unit’ or ‘topological domain,’ inside which supercoiled DNA is stabilized by proteins<sup>6–8</sup> (see Fig. 1(a)).<sup>‡</sup> (Each monomer represents many DNA segments and bound proteins. As a result, the notion of chain persistence is less relevant for the chromosome than for the DNA.) Another example is the section of chromosomes decorated with RNA

polymerases,<sup>10,11</sup> as schematically described by spheres in cyan in Fig. 2 (for simplicity, topological complexities such as multi-fork or “branched-donut”<sup>22</sup> are not shown). The large- $a$  case is also relevant when one wishes to weigh various cellular components as crowders. For instance, inorganic ions (much smaller than typical monomers) outnumber any other species in the intracellular space of *E. coli*, excluding water molecules.<sup>23,24</sup> How significant are their crowding effects in organizing a biopolymer? On the other hand, the large- $a_c$  case ( $a_c > a$ ) includes such examples as RNA and protein chains in cells<sup>3</sup> or polymers in a colloidal solution.<sup>15,25–27</sup> Related problems are polymers in porous or disordered media.<sup>28</sup>

Indeed, we find that a number of general features characterize the flexible-chain collapse, even though the dependence of chain compaction on  $\phi_c$  (volume fraction of crowders) differs between the two cases:  $a > a_c$  and  $a < a_c$  (see Fig. 1(a)).<sup>§</sup> If  $a > a_c$ , in all three spaces (free, cylindrical, and slit-like), molecular crowding depends on the ratio  $a\phi_c/a_c$ ; also “full” compaction by molecular crowding occurs universally at  $a\phi_c/a_c \approx 1$ , independent of any other details such as the geometry of a confined space; beyond this, molecular crowding will not condense the chain much further. As a result, smaller crowders collapse chain molecules better (for a given  $\phi_c$  value) as discussed earlier.<sup>¶16</sup>

When molecular crowding is considered as reducing effectively the solvent quality for the case  $a > a_c$ , there also exists a general relationship between  $a\phi_c/a_c$  and the effective excluded volume  $v$  of monomers. This implies that the action of molecular crowding is local and insensitive to the geometry of a confined space. This is intuitively obvious, since each monomer can be surrounded by several small crowders. Any pair of monomers will experience similar depletion forces, as assumed in the effective-solvent picture. Accordingly, molecular crowding

<sup>†</sup> Note that it is practically impossible to explore a parameter space wide enough to cover both the protein and colloid limits:  $R_g > a_c$  and  $R_g < a_c$ , respectively, where  $R_g$  is the radius of gyration. Here we focus our attention on the  $R_g > a_c$ , which can be easily realized in a biological context.

<sup>‡</sup> More realistically one may choose  $a_c \ll a \approx 100$  nm. One has to include 10 millions of crowders, several times the number of proteins in a cell so as to see the gradual compaction of a chain from its unperturbed size. As it turns out, for  $a > a_c$ , what is important is the combination:  $a\phi_c/a_c$  ( $\phi_c$  the volume fraction of crowders). This gives us some freedom in choosing the values of  $a_c$ .

<sup>§</sup> Here we focus our attention on single polymers in a crowded medium and will not consider the so called colloid limit, where colloids are larger than polymer sizes. This limit is more meaningful at nonzero concentrations of polymers<sup>15,25,27</sup> and deserves separate considerations.

<sup>¶</sup> Because of its biological complexity, the relevance of this finding to the bacterial chromosome is not so obvious; see Section 3 for additional details.

does not reflect the shape of the polymer (or confinement). Also this view favors the physical picture in Fig. 1(b) over that in Fig. 1(c). Crowdiers can reside in the chain-occupying region in Fig. 1(b) but not in Fig. 1(c); in the latter, crowding effects will reflect the size or shape of the polymer and thus the geometry of the confined space.

In contrast, for  $a_c > a$ , chain compaction is almost insensitive to  $a_c$ , as long as  $\phi_c$  is fixed, implying that large crowdiers are all “equal.” As a result, molecular crowding is controlled only by  $\phi_c$  and is less effective for a given  $\phi_c$  value, compared to the corresponding  $a > a_c$  case; for a biologically-relevant range of  $\phi_c$  ( $\phi_c \sim 30\%$ ), we note that it has a modest effect on chain conformations even for a long chain consisting of 2000 monomers, in accord with recent studies.<sup>12,20</sup> Because of their biological complexity, this finding is not so conclusive for the folding of proteins. We focus our attention on clarifying the general features of molecule crowding (see ref. 3 and 12 for recent discussions on the role of molecular crowding in protein folding and function). Nevertheless, the effective-solvent picture remains applicable in a parameter space of biological interest; for the parameter ranges used in our simulations, it works well unless  $a_c \geq 20a$  (see footnote †) as long as the chain is sufficiently long, *i.e.*,  $R_g \gg a_c$ , where  $R_g$  is the radius of gyration.†

This paper is organized as follows. Section 2 outlines the simulation procedures. In Section 3, we first present simulation results for chain compaction and develop theoretical arguments to understand chain compaction in the long-chain limit.

## 2 Simulations

In our simulations, all particles (monomers and crowdiers) interact with each other through a truncated-shifted Lennard-Jones (LJ) potential given by<sup>29,30</sup>

$$U(r) = \begin{cases} U_{\text{LJ}}(r) - U_{\text{LJ}}(r_c) & \text{for } r < r_c \\ 0 & \text{otherwise} \end{cases}, \quad (1)$$

where  $U_{\text{LJ}}(r)$  is the conventional LJ potential:

$$U_{\text{LJ}}(r) = 4\epsilon \left[ \left( \frac{\sigma_{ij}}{r} \right)^{12} - \left( \frac{\sigma_{ij}}{r} \right)^6 \right]. \quad (2)$$

Here,  $r$  is the center-to-center distance between particles, and  $\sigma_{ij}$  and  $\epsilon$  represent the range and the strength of the LJ potential. The subscripts  $i$  and  $j$  are used to distinguish between monomers and crowdiers:  $\sigma_{11} = a$ ,  $\sigma_{22} = a_c$ , and  $\sigma_{12} = (a + a_c)/2$  (*i.e.*, the closest center-to-center distance between a monomer and a crowder).

Note that  $U(r)$  in eqn (1) is a computationally-efficient version of  $U_{\text{LJ}}$  in eqn (2), since it is truncated at  $r_c$ . It is continuous at a cutoff radius  $r_c$ , as it should be. In our explicit-crowder simulations, the cutoff distance  $r_c$  is set to  $2^{1/6}\sigma_{ij}$ . The resulting  $U(r)$  is repulsive for all  $r$  values and is often

referred to as the Weeks–Chandler–Andersen (WCA) potential.<sup>33</sup> In all simulations, we use  $\epsilon$  as energy units. The confining wall is realized by assuming that a monomer or a crowder is repelled by its “image” at the wall *via* the truncated-shifted LJ potential in eqn (1). The inner surface formed by these image spheres defines  $D$ , *i.e.*, the cylinder diameter  $D_{\text{cyl}}$  or the slit gap  $D_{\text{slit}}$ ; for instance,  $D_{\text{cyl}}$  is the closest face-to-face distance between two image spheres on the opposite clock positions. As in recent experimental studies,<sup>8</sup> we discouraged chain adsorption by choosing the strength of this repulsion to be five times  $\epsilon$  for monomer–monomer interactions (see ref. 16 for various scenarios for chain adsorption *vs.* compaction). However, the driving force for chain adsorption may differ between the experimental and our polymer-crowder systems; in the latter, molecular crowding will be responsible.<sup>16</sup>

Monomers are strung together into a chain *via* the finite extensible nonlinear elastic (FENE) potential between two consecutive monomers,<sup>34,35</sup>

$$V(r) = -\frac{1}{2}k_0r_0^2 \ln \left[ 1 - \left( \frac{r}{r_0} \right)^2 \right]. \quad (3)$$

The spring constant is set to  $k_0 = 30\epsilon/a^2$  and the range of the potential is set to  $r_0 = 1.5a$ .

To mimic molecular crowding in our effective-medium picture, we allow the excluded volume  $v$  to vary from  $-a^3$  to  $a^3$ . It suffices to use  $r_c = \{2^{1/6}, 2.5, \infty\}a$  and to vary  $\epsilon$  as  $\epsilon \rightarrow \epsilon' = \{0.01, \dots, 10\}\epsilon$ .

The velocity Verlet method is used to integrate the Newton’s equation of motion. The mass of the monomers and crowdiers is chosen as the mass unit. The units of length, energy, and time of our simulation are  $a$ ,  $\epsilon$ , and  $\tau_0 = a\sqrt{m/\epsilon}$ , respectively. The simulation time step  $\delta\tau$  is set to  $0.002\tau_0$  for  $a > a_c$  or  $0.005\tau_0$  otherwise. The Langevin thermostat is employed with a damping constant  $0.1\tau_0^{-1}$  to keep the temperature at  $T = 1.0\epsilon/k_B$ , where  $k_B$  is the Boltzmann constant. (For physics grounds, the choices of  $m$  and the damping constant are not very important in our work because it does not affect equilibrium quantities.)

The entire system is enclosed in a box (*e.g.*, a cube in the bulk and a cylindrical box for a cylindrically confined chain) of some large volume, typically as large as three times its chain size, and periodic boundary conditions are used at the box surface (in all directions in the bulk case and in the longitudinal directions in the slit and cylinder cases).

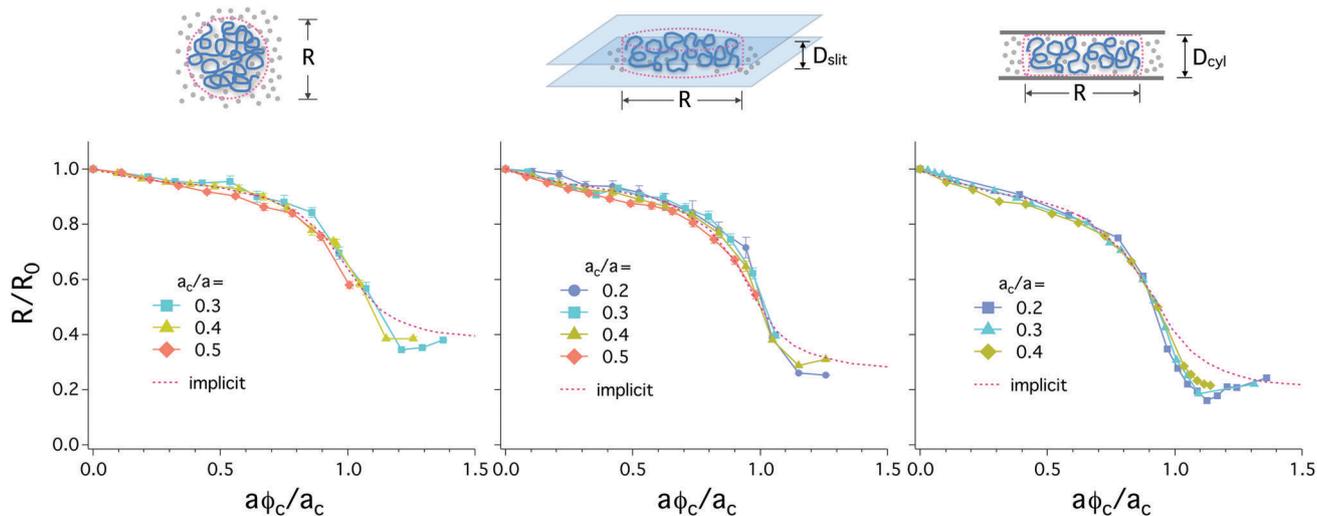
Initially, the polymer is organized in a helical shape (this does not influence equilibrium quantities) but crowdiers are distributed randomly. After chain equilibration, we run our simulation for  $5 \times 10^7$  time steps and obtain data every 1000 steps. We repeat the entire simulation eight times with different random choices of chain conformations and crowder distributions. Ensemble averages are obtained as a time average within each run, which is then averaged over different simulations.

## 3 Results

### 3.1 Chain compaction for the case $a > a_c$

Following the simulation procedure outlined in Section 2, we computed various relevant quantities. Let us first consider the

† On physics grounds, one can argue that its applicability is questionable in the so-called colloid limit  $a_c > R_g$  or for sufficiently large values  $a_c$  and  $\phi_c$ . In the latter case, the spatial correlation between crowdiers becomes important.<sup>31,32</sup> This effect may not be captured systematically in a single excluded-volume parameter of monomers and the effective-solvent picture shows a poor description of crowding effects.



**Fig. 3** Compaction of a flexible polymer by molecular crowding in three different spaces: free or bulk, slit-like, and cylindrical. In all graphs, the normalized chain size  $R/R_0$  is plotted against the ratio  $a\phi_c/a_c$ ;  $R_0$  is the equilibrium chain size in the absence of crowders. For the free and slit-like spaces,  $R$  is chosen to be the radius of gyration; for the latter case, it is measured in the directions parallel with the slit. In the cylindrical space, the farthestmost distance is used for  $R$ . We have chosen  $N = 50$  for the free space and  $N = 80$  for the confined spaces as well as  $D_{\text{slit}} = 5a$  and  $D_{\text{cyl}} = 7a$ , and used several choices of  $a_c$  as indicated in the legend (the subscript 'slit' or 'cyl' will be omitted for simplicity). A general feature emerges from our results described by solid lines with various symbols: in all cases,  $R/R_0$  is a function of  $a\phi_c/a_c$  only, independent of the geometry or the presence of a confined space or of the choices of  $a_c$ . This suggests that the depletion forces between monomers can be considered as reducing the solvent quality in an "effective-solvent" picture. To test this, we establish a "universal" relationship between  $a\phi_c/a_c$  in the explicit-crowder case and the effective excluded volume  $v$  in an equivalent implicit picture (see Fig. 4). The dashed line, based on this relationship, fits the curves well. This justifies the effective-solvent picture and suggests that the action of depletion forces between monomers is local and does not reflect the presence of confinement, as assumed in an implicit-solvent picture. In all explicit-crowder cases (lines with various symbols), there appears to be a local minimum in  $R/R_0$  for  $a\phi_c/a_c \approx 1.2$ . At this high volume fraction of crowders, the collapsed chain tends to get kinetically trapped in a local free-energy minimum, depending on its initial conformation. The effective-solvent result (dashed line in magenta) becomes nearly flat beyond  $a\phi_c/a_c \approx 1.2$ . This is referred to as full compaction, beyond which crowding does not condense the chain much further in both the explicit and implicit cases. (Error bars are shown for a few representative curves.)

case  $a > a_c$ . Fig. 3 displays our results for chain compaction in three different spaces: "free" or bulk, two-dimensional "slit-like," and "open-cylindrical." In all graphs, the normalized chain size  $R/R_0$  is plotted against the ratio  $a\phi_c/a_c$ , where  $R_0 = R(\phi_c = 0)$ . For the free and slit-like spaces, the radius of gyration  $R_g$  measures the chain size, *i.e.*,  $R = R_g$ . If  $\mathbf{r}_n$  is the position vector of monomer  $n$  ( $n = 1, 2, \dots, N$ ),  $\mathbf{R}_{\text{CM}} = \frac{1}{N} \sum_n \mathbf{r}_n$  is the position of its center of mass. In a free space,  $R_g^2 = \frac{1}{N} \sum_n (\mathbf{r}_n - \mathbf{R}_{\text{CM}})^2$ . In a slit-like space,  $\mathbf{r}_n$  and  $\mathbf{R}_{\text{CM}}$  should be understood as the projection onto a plane parallel with the slit. In a cylindrical space, the farthestmost distance is used for  $R$ .<sup>36</sup>

In our simulations for  $a > a_c$ , we chose the parameters as follows:  $N = 50$  for the free space,  $N = 80$  for the confined spaces, and the slit gap  $D_{\text{slit}} = 5a$ , and the cylinder diameter  $D_{\text{cyl}} = 7a$ . For simplicity, we will drop the subscripts "slit" and "cyl." As described in Section 2,  $D$  refers to the inner surface formed by wall particles. Also we used several choices of  $a_c$  as indicated in the legend:  $a_c = 0.3, 0.4, 0.5a$  for the bulk,  $a_c = 0.2, 0.3, 0.4, 0.5a$  for the slit geometry, and  $a_c = 0.2, 0.3, 0.4a$  for the cylindrical space. For  $N = 50$ , the radius of gyration of the chain is given by  $R_g \approx 5a$  ( $\phi_c = 0$ ). A related quantity is the so-called Flory radius:  $R_F \approx 1.1 \times N^{3/5} a \approx 12a$  ( $N = 50$ ), which is larger than the corresponding  $R_g$ .<sup>37</sup> It is worth noting that the degree of confinement is often expressed in terms of  $R_F/D$  (see ref. 37 and 38 and references therein). With our parameter choices,

this ratio is larger or appreciably larger than one. This means that confinement is moderately strong or strong.

A number of general features characterize our results (solid lines with various symbols) in Fig. 3. (See Section 3.4 below for our discussion on how  $R$  depends on  $N$ .) In all cases,  $R/R_0$  is a function of  $a\phi_c/a_c$ ; it decreases continuously as  $a\phi_c/a_c$  increases from  $a\phi_c/a_c = 0$ , but it reaches its minimum at  $a\phi_c/a_c \approx 1.2$ . While the origin of this non-monotonicity is not entirely clear, one possibility is a kinetic effect.<sup>17</sup> At this high volume fraction of crowders, the collapsed chain tends to get kinetically trapped in a local free-energy minimum, similar to what was observed for a cylindrically-confined chain.<sup>\*\*16</sup> In this case, our effective-solvent result, described by the dashed line in magenta (see below for details), will offer an alternative picture, in which  $R/R_0$  becomes nearly flat beyond  $a\phi_c/a_c \approx 1.2$ . In this work, this is referred to as "full compaction" by molecular crowding.

**\*\*** Recently, a non-monotonic reduction of chain size with increasing  $\phi_c$  was observed in theoretical studies in which crowders were implicitly taken into account.<sup>39</sup> Earlier, it was shown theoretically that correlations among polymeric crowders are responsible for reentrant-like swelling at high crowder volume fractions.<sup>31,32</sup> In more recent simulations with hard-sphere crowders<sup>20</sup> and experiments with polymeric crowders,<sup>19</sup> however, such a behavior was not observed. Furthermore, this non-monotonic dependence does not appear to be a general feature of chain compaction. This is not reflected in some curves in Fig. 3 and in our results in Fig. 5(b).

The observation that  $R/R_0$  is a function of  $a\phi_c/a_c$  has a number of consequences. First, this shows the interplay between  $\phi_c$  and  $a_c$ : doubling  $\phi_c$  is equivalent to halving  $a_c$  in all cases (see ref. 16 for a cylindrically-confined case). Second, it suggests that the action of depletion forces is local and does not reflect confinement, as assumed in an effective-solvent picture, in which molecular crowding is viewed as reducing the excluded volume  $v$ . For  $a > a_c$ , a monomer pair can be surrounded by several crowders, as illustrated in Fig. 1(a). Any pair will experience the same or similar depletion forces. This is desired for the effective-solvent picture and is responsible for the local nature of depletion forces as seen in Fig. 3.

A related point is that crowders can reside in the chain-occupying region as illustrated in Fig. 1(b). Even though this alone will not necessarily justify the effective-solvent picture, it is required. If the chain-occupying region is impermeable to crowders (see Fig. 1(c)), the action of crowding will reflect chain shape and the geometry of a confined space. This will invalidate the effective-solvent picture.

To test our effective-solvent picture further, we establish a “universal” relationship between  $a\phi_c/a_c$  in the explicit-crowder case and the excluded volume  $v$  in an equivalent implicit picture (see Fig. 4 for details). The dashed line in Fig. 3, based on this relationship, fits the curves well. This justifies the effective-solvent picture. As a result, depletion forces are local and do not intrinsically reflect the presence of confinement.

How can the explicit-crowder case map onto an equivalent effective-solvent case, in which the solvent quality determines the excluded volume  $v$ ? Obviously, increasing  $\phi_c$  should amount to decreasing the excluded volume  $v$  and eventually making  $v$  negative. To map out a relationship between  $\phi_c$  and  $v$ ,

first note that  $v$  is related to the monomer–monomer interaction  $U(r)$  via the formula

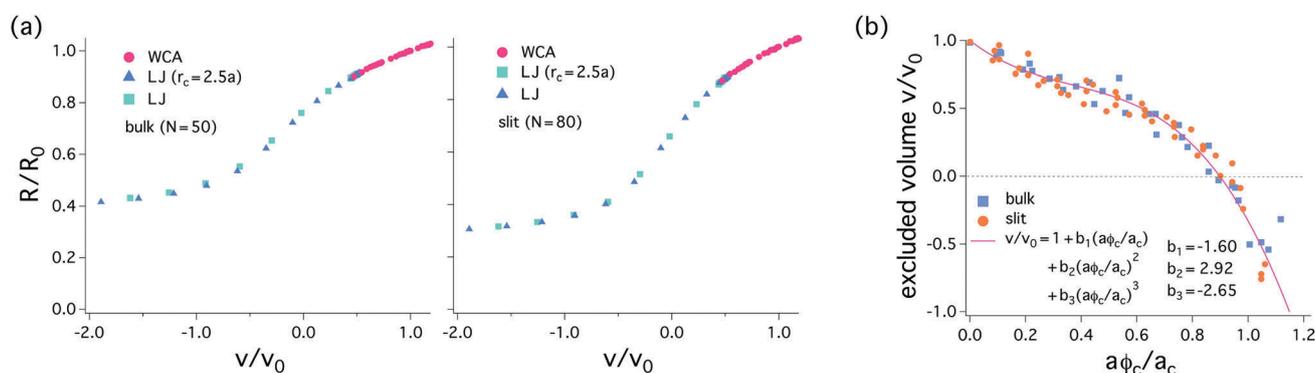
$$v = \int_0^\infty [1 - e^{-U(r)/k_B T}] d^3r. \quad (4)$$

In our considerations,  $U(r)$  coincides either with the modified LJ potential in eqn (1) or the original LJ potential  $U_{LJ}$  in eqn (2). The reference excluded volume  $v_0 \approx a^3$  (athermal) corresponds to the choice  $r_c = 2^{1/6}a$ . Recall that with  $r_c = 2^{1/6}a$  the LJ potential becomes the WCA and that we set  $T = 1.0\epsilon$ , as discussed in Section 2. Note that this  $v_0$  characterizes our explicit-crowder simulations.

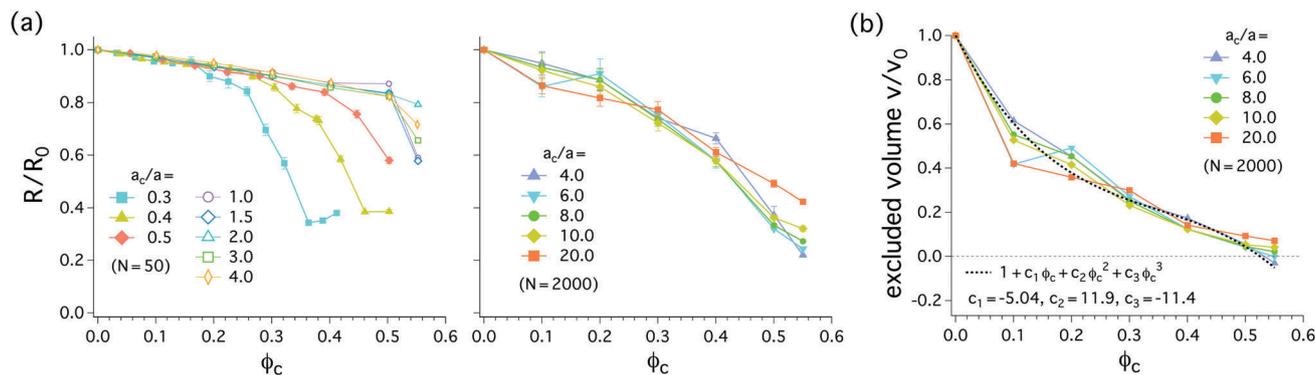
As an intermediate step, we first examine how the chain size  $R$  varies with  $v$  in the absence of crowders. For this, we essentially repeated our simulations in Fig. 3 without crowders; the effect of molecular crowding is implicitly taken into account by adjusting  $\epsilon$  and  $r_c$  (see Section 2), which in turn influences  $U(r)$  and  $v$ . The graph in Fig. 4(a) summarizes the resulting  $R/R_0$  against  $v/v_0$  in an unconfined (left) or slit-like space (right). As expected, the chain collapses as  $v$  is reduced.

We establish a relationship between  $v/v_0$  and  $a\phi_c/a_c$  by comparing chain compaction in Fig. 4(a) and in Fig. 3. For this, we find the right combination  $(v/v_0, a\phi_c/a_c)$ , at which  $R/R_0$  is the same in both graphs, and plot  $v/v_0$  as a function of  $a\phi_c/a_c$ . The resulting relationship for free and slit-like spaces is shown in Fig. 4(b). Consistent with our expectation that crowding effects are local, the two sets of data (squares and circles for free and slit-like spaces, respectively) collapse onto each other; The solid line in magenta is the fitting curve given by

$$\frac{v}{v_0} = 1 + b_1 \left( \frac{a\phi_c}{a_c} \right) + b_2 \left( \frac{a\phi_c}{a_c} \right)^2 + b_3 \left( \frac{a\phi_c}{a_c} \right)^3. \quad (5)$$



**Fig. 4** Mapping the explicit-crowder case onto an equivalent implicit-solvent case: the crowder volume fraction  $\phi_c$  vs. the excluded volume  $v$  of monomers. The parameter  $v$  is related to the monomer–monomer interaction  $U(r)$  via  $v = \int_0^\infty \{1 - \exp[-U(r)/k_B T]\} d^3r$ . In our simulation,  $U(r)$  coincides with the LJ potential in eqn (1), which is completely specified in terms of  $\epsilon$  and  $r_c$ , or with  $U_{LJ}(r)$  in eqn (2). The reference excluded volume  $v_0$  corresponds to the choice  $r_c = 2^{1/6}a$ . Note that this  $v_0$  characterizes our explicit-crowder simulations. (a) Normalized chain size  $R/R_0$  as a function of  $v/v_0$ . As an intermediate step toward relating  $v$  to  $\phi_c$ , we first examine how the chain size varies with  $v$ . For this, we essentially repeated our simulations in Fig. 3 without crowders; the effect of molecular crowding is implicitly taken into account by adjusting  $\epsilon$  and  $r_c$ , which in turn influences  $U(r)$  and  $v$ . The resulting  $R/R_0$  is plotted against  $v/v_0$  in the graph for free (left) and slit-like spaces (right). As expected, the chain collapses as  $v$  is reduced. (b) Normalized excluded volume  $v/v_0$  vs.  $a\phi_c/a_c$ . A relationship between  $v/v_0$  and  $a\phi_c/a_c$  is established by comparing chain compaction in (a) and that in Fig. 3. In the resulting graph displayed in (b), free and slit-like spaces are represented by blue squares and red circles respectively; note that the two sets of data collapse onto each other. The solid line in magenta is the fitting curve:  $v/v_0 = 1 + b_1(a\phi_c/a_c) + b_2(a\phi_c/a_c)^2 + b_3(a\phi_c/a_c)^3$ . The fitting parameters  $b_1$ ,  $b_2$ , and  $b_3$  are to be chosen for the best fit to the explicit-crowder data. Obviously,  $b_1 < 0$ , meaning that the solvent quality becomes poorer as  $a\phi_c/a_c$  increases. The discrepancy between the fitting curve and the data for  $a\phi_c/a_c > 1.0$  can be attributed to the truncation of the  $v$  series beyond the cubic term.



**Fig. 5** (a) Comparison of small- and large- $N$  cases for a wide range of  $a_c$  values (i.e.,  $0.3a \leq a_c \leq 20a$ ) (bulk). (Left)  $N = 50$  and  $a_c = 0.3, 0.5, \dots, 4a$  and (right)  $N = 2000$  and  $a_c = 4, 6, \dots, 20a$ . (left) As  $a_c$  increases from  $a_c = 0.3a$ , molecular crowding becomes less effective and eventually insensitive to  $a_c$  beyond  $a_c \approx a$  for a fixed value of  $\phi_c$ . For  $a_c \geq a$ , it has a marginal effect on  $R$  for the entire range of  $\phi_c$  values shown. If  $a_c \geq a$ , the insensitivity of  $R/R_0$  to  $a_c$  persists for  $N = 2000$  (right). Chain compaction is more pronounced for larger  $N$ ; for the same  $\phi_c$  and  $a_c (>a)$ ,  $R/R_0$  is smaller for  $N = 2000$  (right). However, even in this large  $N$  case, chain compaction is moderate (i.e.,  $R/R_0 \approx 0.8$ ) for the biologically-relevant  $\phi_c$  range:  $\phi_c \approx 0.3$ . Also note that the curve for  $a_c = 20a$  deviates somewhat from others, which tend to collapse onto each other. This implies that the effective-solvent picture is a cruder approximation for  $a_c = 20a$  compaction to the smaller- $a_c$  cases. (b) Normalized excluded volume  $v/v_0$  against  $\phi_c$  for  $a_c > a$ . The same color scheme is used here to distinguish various  $a_c$  values as in the corresponding  $R$ - $\phi_c$  graph (the one on the right) in (a). The dashed line is a fitting curve given by  $v/v_0 = 1 + c_1\phi_c + c_2\phi_c^2 + c_3\phi_c^3$  with  $c_1 = -5.04$ ,  $c_2 = 11.9$ , and  $c_3 = -11.4$ ; for this, data points describing  $a_c = 20a$  is excluded. (Error bars are shown for a few representative curves.)

The fitting parameters  $b_1$ ,  $b_2$ , and  $b_3$  are to be chosen for the best fit to the explicit-crowder data:  $b_1 = -1.60$ ,  $b_2 = 2.92$ , and  $b_3 = -2.65$ .

For a later convenience, eqn (5) can be recast in a more general form as

$$\frac{v}{v_0} = 1 - \alpha_1\phi_c + \alpha_2\phi_c^2 - \alpha_3\phi_c^3, \quad (6)$$

where  $\alpha_i = |b_i|(a/a_c)^i > 0$  for  $a > a_c$ . As evidenced in Section 3.2, for  $a_c \geq a$ ,  $R/R_0$  becomes independent of  $a_c$  for a given value of  $\phi_c$  and is controlled by  $\phi_c$  alone. In this case,  $\alpha_i = \text{const} = \mathcal{O}(1)$ . On physics grounds, one can argue that  $\alpha_i > 0$  as in the case  $a > a_c$ . In principle, one can determine the functional form of  $\alpha_i$ , i.e.,  $\alpha_i(\phi_c, a_c, a)$ , which remains valid for both  $a > a_c$  and  $a < a_c$ . However, we note that this is practically challenging. Nevertheless, the physical picture is obvious, except for  $a \approx a_c$ . We will use eqn (6) as a general mapping relationship between  $v$  and  $\phi_c$ . See Fig. 5(b) for the estimate of  $\alpha_i$  ( $= |c_i|$ ) for  $a_c = 4a$ .

In eqn (5),  $v \rightarrow v_0$  as  $a_c \rightarrow 0$  for a given  $\rho_c \approx \phi_c/a_c^3$  (the number density of crowders); in this limit,  $\phi_c \rightarrow 0$ . To understand the physical origin of this behavior, note that the depletion free-energy gain for two monomers in contact with each other (Fig. 1(a)) can be approximated as<sup>††13</sup>

$$\frac{\Delta F_{\text{dep}}}{k_B T} = \phi_c \left( \frac{3a}{2a_c} + 1 \right) + \mathcal{O}(\phi_c^2) \approx \frac{3a\phi_c}{2a_c}. \quad (7)$$

Here the second equality holds for  $3a/2a_c \gg 1$ . The free energy in eqn (7) vanishes if we take  $a_c \rightarrow 0$  while holding  $\rho_c$  fixed, similarly to what eqn (5) implies. This can be absorbed into the excluded

<sup>††</sup> As noted earlier,<sup>16</sup> this is valid at the level of two-body interactions, since it is obtained for two monomers, excluding two consecutive ones, while others are ignored. The physical picture depicted in Fig. 1 becomes inaccurate if the overlapped (i.e., shaded) region falls on a monomer nearby. In this sense,  $\Delta F_{\text{dep}}$  works better for the case  $a > a_c$  than for  $a_c > a$ .

volume as  $v = v_0(1 - a\phi_c/a_c)$  after a numerical prefactor is dropped. This is aligned with eqn (5) up to the linear order in  $\phi_c$ .

The good agreement between the data and the fitting curve in Fig. 4(b) justifies the effective-solvent picture in eqn (5). A similar picture has been employed.<sup>40</sup> By integrating out the degrees of freedom associated with crowders at the Gaussian level, an effective excluded volume was obtained. If numerical prefactors are dropped, in our notation, it reads

$$v = v_0 - \rho_c \frac{(a + a_c)^6}{1 + a_c^3 \rho_c}. \quad (8)$$

Alternatively, one can use a virial expansion of the free energy of our crowder-polymer system.<sup>41</sup>

While eqn (8) captures correctly molecular crowding as reducing the solvent quality, it does not lead to the aforementioned limiting behavior, but it rather indicates  $v/v_0 \rightarrow 1 - a^3\rho_c$ , as  $a_c \rightarrow 0$  for a fixed  $\rho_c$  value. If taken literally, this behavior has an unexpected consequence: in the bacterial cell, depletion forces between “big monomers” (i.e.,  $a > a_c$ ) would be dominated by any species with the largest  $\rho_c$ . This is distinct from what we would expect from eqn (5), which can be written as  $v/v_0 \approx 1 - aa_c^2\rho_c \rightarrow 1$  in this limit (after dropping numerical prefactors). This means that crowding effects induced by small ions will be insignificant for chain compaction.

The discrepancy between eqn (5) and (8) can be attributed to the neglect of chain connectivity in mean-field type approaches. For instance, these approaches put two consecutive monomers and two distant ones along the contour on equal footing. This is an artifact arising from the neglect of chain connectivity or monomer-monomer correlations. More systematic theoretical treatments will be desirable toward reconciling between eqn (5) and (8).

It will be useful to discuss the  $a_c$ -dependence of  $v$  in some context. For instance, a typical *E. coli* cell contains about  $10^8$

inorganic ions and  $3 \times 10^6$  proteins.<sup>23,24</sup> Excluding water molecules, inorganic ions outnumber any other species in the cytoplasm. Eqn (5) and (8) measure different species very differently. For the purpose of our discussion, it suffices to compare  $a_c^2 \rho_c$  or  $\rho_c$  values between small ions ( $a_c \approx 0.2$  nm) and proteins ( $a_c \approx 5$  nm).<sup>23</sup> If we treat all of them as crowders even though some of them are bound to some other molecules (e.g., membranes and DNA), we have

$$a_c^2 \rho_c \times V_{\text{cell}} \approx \begin{cases} (0.2)^2 \times 10^8 \text{ nm}^2 = 4 \times 10^6 \text{ nm}^2 & \text{(small ions)} \\ 5^2 \times 3 \times 10^6 \text{ nm}^2 = 75 \times 10^6 \text{ nm}^2 & \text{(proteins)} \end{cases}, \quad (9)$$

where  $V_{\text{cell}}$  is the cell volume.

According to this analysis based on eqn (5), small ions would not contribute significantly to crowding even though they outnumber proteins. An opposite conclusion will be reached if we compare  $\rho_c$  values following the limiting behavior of  $v/v_0$  discussed below eqn (8): small ions will be the major crowder type.

If we repeated the analysis in eqn (9) for water, we would arrive at the conclusion that water is the most significant crowder. However, the general findings in Fig. 3 may not be applicable to closely-packed crowders, e.g., water. For a practical purpose, the ‘‘collective’’ effect of water on a chain molecule can be taken into account *via* the excluded volume of each monomer  $v$ . In an athermal solvent,  $v = v_0 = a^3$ , but more generally,  $v < v_0$ .<sup>42</sup> This allows one to focus on other crowders.

Our general findings in Fig. 3 should be used with caution for understanding the spatial organization of bacterial chromosomes, which are organized into many structural units or topological domains (see ref. 6–8 and references therein). In a number of studies, the structural unit is coarse-grained into a monomer, as shown in Fig. 1(a).<sup>16–18</sup> In some studies, the effects of DNA-bound proteins can be approximately mimicked by cross-linking<sup>8,43</sup> or a harmonic potential between monomers.<sup>44</sup> The interaction between two harmonically-constrained linear polymers in a crowded medium is less attractive for smaller crowders for a given  $\phi_c$  value.<sup>44</sup> This ‘unusual size dependence’<sup>44</sup> appears to be opposite to what we would expect from eqn (7). However, what is unclear is how a cross-linked polymer or a long polymer formed by many of the harmonically-collapsed polymeric subunits responds to crowders (e.g., the dependence of  $R/R_0$  on  $a_c$ ). Importantly, the degree of compaction is non-uniform along the chromosome and can be dominated by the stronger attraction between big monomers (e.g., big spheres in cyan in Fig. 2), which include many DNA segments and bound biomolecules such as RNA polymerases<sup>10,11</sup> (see also ref. 45–47). These big monomers are much larger than typical crowders.<sup>10,11</sup> As a result, the  $a_c$ -dependence of crowding effects as suggested in Fig. 3 will be reflected better than that indicated by a homogeneous polymer model whether cross-linked or harmonically constrained.

It has been known for some time that polyvalent counterions can condense highly-charged biomolecules such as DNA into

tightly organized bundles; they induce attraction between otherwise-repelling backbone charges.<sup>48</sup> The required counterion concentration depends on a few parameters including the surface charge density of polyions and ion sizes.<sup>49</sup> DNA bundling is, however, an electrostatic phenomenon and does not seem to be implicated in bacterial chromosome organization.<sup>50</sup> In the bacterial cell, the concentration of these ions required for bundling may not be reached; they primarily reduce the backbone charge density of DNA. Nevertheless, the observation that small ions do not contribute significantly to crowding (see eqn (9)) allows one to focus on their electrostatic effects.<sup>48,49</sup>

### 3.2 Chain compaction for the case $a_c > a$

So far we have focused on the case  $a > a_c$ . Will the physical pictures discussed earlier remain applicable to the case  $a_c > a$ ? In this case, it is increasingly computationally demanding to simulate a confined polymer. The value of  $D$  has to be at least several times larger than the bigger of  $a$  and  $a_c$ . This means that a larger  $N$  value is required for a larger  $a_c$  value, possibly except in the bulk case. Also crowding effects are less significant for  $a_c > a$ , as evidenced below. For strong compaction,  $\phi_c$  has to be as large as  $\phi_c = 0.5$ – $0.6$ . In this high volume fraction, (cylindrical) confinement can induce wall-layering of hard spheres (e.g., monomers and crowders).<sup>36,51</sup> Accordingly, simulation details can enter into the picture of crowding, even though their consequences may not be biologically meaningful. Here, we focus on the bulk case and comment on the relevance of an effective-solvent picture for confined cases at the end of this subsection.

Fig. 5(a) summarizes our results for  $R/R_0$  obtained with various parameter choices and in the absence of confinement: (left)  $N = 50$  and  $a_c = 0.3, 0.4, 0.5, \dots, 4a$  and (right)  $N = 2000$  and  $a_c = 4, 6, \dots, 20a$ . Recall  $R_g \approx 5a$  for  $N = 50$  and note that  $R_g \approx 46a$  for  $N = 2000$  (in a free space in the absence of crowders).

As shown in the left graph in Fig. 5(a), for  $N = 50$ , as  $a_c$  increases from  $a_c = 0.3a$ , molecular crowding becomes less effective and eventually insensitive to  $a_c$  beyond  $a_c \approx a$ . In other words,  $R/R_0$  for a given  $N$  value is controlled by  $\phi_c$  alone. This is a natural consequence of  $F_{\text{dep}}$  in eqn (7) in the limit  $a_c \gg a$ . For  $a_c \geq a$ , it has a moderate effect on  $R$  for the entire range of  $\phi_c$  shown, up to  $\phi_c = 0.6$ .

The graph on the right in Fig. 5(a) shows  $R/R_0$  obtained with  $N = 2000$  and relatively large values of  $a_c$  ( $\geq 4a$ ). The insensitivity of  $R/R_0$  to  $a_c$  ( $\geq a$ ), shown in the left graph, persists for  $N = 2000$  (right). Also note that the curves for  $a_c \leq 10a$  tend to collapse onto each other, but the curve for  $a_c = 20a$  deviates somewhat from those for  $a_c \leq 10a$ . This implies that the effective-solvent picture is a cruder approximation for  $a_c = 20a$ , compared to the smaller- $a_c$  cases. But the results in Fig. 5(a) are not so conclusive about this.

To further unravel the effective-solvent picture, especially for the case  $a_c > a$ , we calculated the internal distances between monomers defined as  $\langle |\mathbf{r}_i - \mathbf{r}_j|^2 \rangle^{1/2}$  and presented our results in the Appendix (Fig. 8). Consistent with our discussion above, the effective-solvent picture works well except in the highly

asymmetrical case  $a_c \geq 20a$ . On average, intracellular crowders are about 5 nm in size<sup>23,24</sup> and can generally be considered as reducing the solvent quality for chain molecules.

The general trend observed in the Appendix is paralleled by the results in Fig. 5(b), where the effective excluded volume  $v$  is related to  $\phi_c$  for a range of  $a_c$  values. The black dashed line is a fit to the data points excluding those representing  $a_c = 20a$ , given by  $v/v_0 = 1 + c_1\phi_c + c_2\phi_c^2 + c_3\phi_c^3$  with  $c_1 = -5.04$ ,  $c_2 = 11.9$ , and  $c_3 = -11.4$ . In contrast to the case  $a > a_c$  in Fig. 4(b), this graph shows how the effective solvent picture may break down for sufficiently large  $a_c$  values. Indeed the  $a_c = 20a$ -curve in (b) appears to deviate more noticeably from others and the dashed line. This compares favorably with the graph on the right in Fig. 5(a) and our discussions in the Appendix.

Earlier, we noted that chain compaction is more pronounced for larger  $N$ . This is also reflected in Fig. 5; for the same  $\phi_c$  and  $a_c (> a)$ ,  $R/R_0$  is smaller for  $N = 2000$  (b). However, even in this large- $N$  case, chain compaction is somewhat modest (*i.e.*,  $R/R_0 \approx 0.7$ – $0.8$ ) for the biologically-relevant  $\phi_c$  range:  $\phi_c \approx 0.3a$  (see ref. 20 for a similar observation). If taken literally, this seems to imply that the effect of crowding is not so significant in collapsing a protein chain, since it belongs to the large- $a_c$  case, in which molecular crowding has a marginal effect. While this is generally consistent with a recent review<sup>12</sup> (see Fig. 3), it is beyond the scope of this work to further clarify the role of crowding in organizing a protein chain (see ref. 3 and relevant references therein).

Even though the results in Fig. 5 were obtained for a polymer in a free space, the general picture (*i.e.*,  $R/R_0$  as a function of  $\phi_c$  for  $a_c \geq a$ ) remains applicable to other spaces, similar to what Fig. 3 suggests. Unless  $a_c \gg a$ , the effective-solvent picture works and the chain response to molecular crowding will not reflect confinement.

To what extent will the general findings in Fig. 5 remain relevant for a confined chain? Conversely speaking, under what conditions will they become less relevant for a cylindrically-confined chain? An obvious example is the preferential positioning of crowders in the vicinity of the cylindrical wall in some range of  $\phi_c$ , as was seen with DNA compaction.<sup>19</sup> A key determinant here is the large correlation length of stiff DNA molecules. As a result, crowders have easy access to a layer of some thickness from the wall, which is comparable to this length, effectively reducing the cylinder diameter, more so for a larger correlation length. Understandably, this was not seen with flexible polymers trapped in a cylindrical space crowded with small crowders, *i.e.*,  $a > a_c$ .<sup>36</sup> One may argue that this trend will persist for  $a_c > a$ , since for a given  $a_c$  value, the correlation length is relatively small. Another potential source for the breakdown of the effective-solvent picture is chain adsorption induced by crowding,<sup>16</sup> which can invalidate the local picture of depletion forces as assumed in the effective-solvent picture, not only for  $a_c > a_c$  but also for  $a < a_c$ .

As long as chain adsorption and wall-layering are discouraged, we believe that the effective-solvent picture remains valid except in a highly asymmetrical case:  $a_c \gg a$  or  $a_c > R_g$ . This is consistent with our view that crowding effects are intrinsically

local possibly except for  $a_c \gg a$  or  $a_c > R_g$ . In practice, the cylindrical wall can be passivated so as to prevent chain adsorption;<sup>8</sup> also for a typical  $\phi_c$  range in a cell, wall-layering will not be easily observed. Importantly, in a biological context, the large- $a_c$  case includes proteins and RNA. The typical size of these biopolymers even in a coil-like structure is much smaller than the cell diameter  $\sim 1 \mu\text{m}$ . As a result, their structure will not reflect sensitively confinement effects. For a practical purpose, they can be treated as bulk systems. Finally, crowding effects can alter protein functions even though they alone would not change protein size much (see ref. 3, 12 and 20 and references therein).

### 3.3 Spatial distribution of monomers and crowders

As noted earlier in Section 3.1, what underlies the effective solvent picture and the locality of depletion forces is the ability of crowders to reside in the chain-occupying region so as to maintain chemical equilibrium across the boundary between this region and the outer crowder-rich region (Fig. 1(b)). To probe this in a quantitative way, we have examined the spatial distribution of monomers and crowders. For a reason similar to that described at the end of the last subsection, we have focused our attention on the bulk case. For  $a > a_c$ , confinement does not alter molecular crowding up to full compaction (Fig. 3); on the other hand, biopolymers belonging to the large- $a_c$  case will show bulk-like behavior.

Fig. 6 summarizes our results obtained for a few combinations of  $a_c$  and  $N$ : (a)  $a_c = 0.4a$  and  $N = 50$ , (b)  $a_c = 4a$  and  $N = 2000$ , and (c)  $a_c = 20a$  and  $N = 2000$ . What is plotted in each graph is the volume fraction  $\phi_i(r)$  of monomers and crowders, as a function of  $r$ , *i.e.*, the longitudinal distance from the center of mass of the polymer, where the subscript  $i = 'm'$  or  $'c'$  refers to monomers or crowders, respectively. For visual clarity, we have normalized  $\phi_c(r)$  as  $\phi_c(r)/\phi_c(r = \infty)$  and  $\phi_m(r)$  as  $\phi_m(r)/\phi_m(r = 0)$ .

First, each graph suggests the coexistence of monomer-rich and crowder-rich phases. Crowders become depleted from the chain-occupying region as  $\phi_c$  increases. This is a natural consequence of chain collapse induced by molecular crowding. Consistent with our expectation, the chain-occupying region is more permeable to crowders for smaller  $a_c$  values. For this, compare the two cases  $a_c = 4a$  (b) and  $a_c = 20a$  (c); also note that the  $\phi_c(r)$  decreases by about 20% inside the chain-occupying region even for the largest value of  $\phi_c$  used in (a):  $a\phi_c/a_c = 0.95$  (the dashed curve in blue). In general, the results in Fig. 6 are aligned well with Fig. 1(b) even for  $a_c = 20a$ , *i.e.*, when the effective-solvent approach becomes a poor approximation; while the curve representing  $a_c = 20a$  and  $\phi_c = 0.5$  (the blue dashed line in (c)) appears to be an exception, it may not bear much biological relevance. For a larger  $a_c$  value, the “caging” or confining effect of crowders becomes more pronounced and the general picture suggested by Fig. 6 will not hold any longer.

### 3.4 Large- $N$ limit

So far, we have used a fixed value of  $N$  for each case. Our results are not conclusive for the  $N$ -dependence of chain collapse, especially for large  $N$ . It is important to note that this does not mean that the chain response to molecular crowding varies

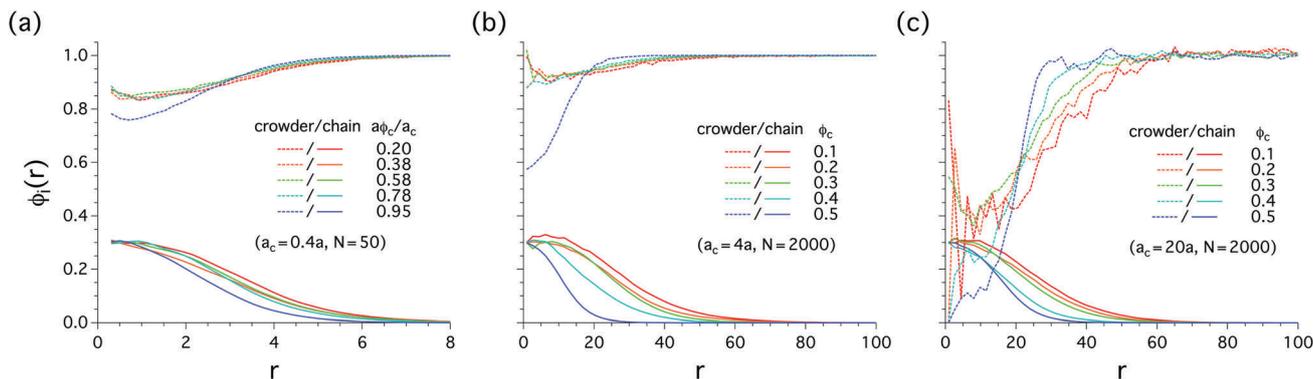


Fig. 6 Molecular crowding vs. spatial organization of monomers and crowders in a free space. Each graph displays the volume fraction  $\phi_i(r)$  of monomers and crowders with  $i = 'm'$  (monomer) or  $'c'$  (crowder) as a function of  $r$ , i.e., the longitudinal distance from the center of mass of the polymer, for  $a_c = 0.4a$  (a),  $a_c = 4a$  (b), and  $a_c = 20a$  (c). For visual clarity,  $\phi_c(r)$  is normalized as  $\phi_c(r)/\phi_c(r = \infty)$  and  $\phi_m(r)$  as  $\phi_m(r)/\phi_m(r = 0)$ . The results in this graph suggest the coexistence of monomer-rich and crowder-rich phases. They also support the picture in Fig. 1(b) that the chain-occupying region is permeable to crowders even for  $a_c = 20a$ , i.e., even when the effective-solvent approach becomes a poor approximation.

with  $N$ . Instead, it indicates that the effect of self-avoidance on chain size depends on  $N$ , as is also the case for chain collapse in poor solvents.<sup>52</sup> In other words, while  $v$  is independent of  $N$ ,  $R/R_0$  generally depends on  $N$ . Our implicit-crowder or effective-solvent picture offers a theoretically-feasible framework for examining systematically the  $N$ - or  $D$ -dependence, which will complement the results in Fig. 3.

In the effective-solvent picture,  $v$  given in eqn (5) is the only parameter that controls the monomer–monomer interaction. The degree of chain swelling or compaction is often expressed in terms of  $\Xi^2 = \langle \mathbf{R}^2 \rangle / \langle \mathbf{R}^2(v = 0) \rangle$ . If  $d$  is understood as referring to the geometry of a confined space, i.e.,  $d = 1$  for cylindrical confinement, one can derive a meanfield equation for  $\Xi$  for any  $d$  ( $1 \leq d \leq 3$ ). It proves to be useful to introduce  $X \approx (v/v_0)N^{2-d/2}$  and  $Y = w a^{-6} N^{3-d}$  (with  $w (> 0)$  as the three-body parameter) in a quasi  $d$ -dimensional space. In an athermal solvent without crowders,  $v = v_0 \approx a^3$ . On the other hand,  $w \approx a^6$ , which is roughly independent of  $T$ .<sup>53</sup>

After numerical prefactors are dropped, the free energy of a polymer in a quasi  $d$ -dimensional space, which produces the expected equilibrium chain size, can be written as

$$\frac{F_d}{k_B T} \approx \Xi^2 - \ln \Xi^2 + \frac{X}{(D/a)^{3-d} \Xi^d} + \frac{Y}{(D/a)^{6-2d} \Xi^{2d}}. \quad (10)$$

$$\frac{R}{R_0} \sim \begin{cases} (v/v_0)^{1/5} \sim (1 - \alpha_1 \phi_c)^{1/5} \sim 1 - \alpha_1 \phi_c / 5, & \phi_c < \phi_0 \\ N^{-1/10}, & \phi_c = \phi_0 \\ N^{-4/15} (v_0/|v|)^{1/3} \sim N^{-4/15} |1 - \alpha_1 \phi_c + \alpha_2 \phi_c^2 - \alpha_3 \phi_c^3|^{-1/3}, & \phi_c > \phi_0 \end{cases} \quad (14)$$

See ref. 53 and 54 for a polymer in  $d$  dimension; for confined polymers, a similar free energy can be found in ref. 55. Free energy minimization leads to

$$\Xi^{d+2} - \Xi^d = \frac{X}{(D/a)^{3-d}} + \frac{Y}{(D/a)^{6-2d}} \Xi^{-d}. \quad (11)$$

This equation can be analyzed for the unknown  $\Xi$  as a function of  $X$ .

It is worth commenting on the applicability or limitation of eqn (10). In all spaces, it yields the expected scaling result for equilibrium  $R$ . For  $v > 0$ , one can correctly view the confined chain as a linear string of blobs of size  $D$  each and construct a renormalized Flory free energy,<sup>56,57</sup> which has a different  $D$ -dependence. However, it leads to the same equilibrium  $R$  as eqn (10) does:  $R \sim Na(D/a)^{-2/3}$  for  $d = 1$ , for instance. In our consideration below, we will focus on  $R$  with  $v$  varying from a positive to negative value and analyze eqn (11) for various cases.

**3.4.1 In a free space.** For  $d = 3$ ,  $Y$  is  $N$ -independent. In this case, the  $N$ -dependence of  $\Xi$  solely arises from  $X \sim N^{1/2}$ . One may view  $\Xi$  as a function of  $vN^{1/2}$ :

$$\Xi = f\left(\frac{v}{v_0} N^{1/2}\right). \quad (12)$$

This can be converted into

$$\frac{R}{R_0} = N^{-1/10} f\left(\frac{v}{v_0} N^{1/2}\right). \quad (13)$$

Recall that  $R_0 = R(\phi_c = 0)$ .

Eqn (13) can be readily analyzed in the large- $N$  limit. Let  $\phi_0$  be the  $\phi_c$  value at which  $v = 0$ . For  $\phi_c < \phi_0$ , we require that  $R \approx aN^{3/5}$ ; for  $\phi_c > \phi_0$ ,  $R \approx aN^{1/3}$ . This line of reasoning leads to

For  $\phi_c < \phi_0$ ,  $R/R_0$  is  $N$ -independent, but for  $\phi_c > \phi_0$ , it is smaller for larger  $N$  and decays as  $N^{-4/15}$ . This is consistent with the expected scaling behavior:  $R(\phi_c > \phi_0) \sim N^{1/3}$  for a collapsed chain.

We can extend this analysis to examine the  $N$ -dependence of  $\phi_c$  at which the transition to a collapsed state occurs. Recall that  $X \sim (v/v_0)N^{1/2}$  for  $d = 3$ . It is this combination that enters into eqn (11). As  $N$  increases, the effect of two-body attractions

can be more easily felt. As a result, in the limit  $N \rightarrow \infty$ , chain compaction occurs within a narrow range:  $\Delta\phi_c \approx N^{-1/2} \rightarrow 0$  (see ref. 58 for a similar issue for chain collapse in a poor solvent). The chain collapse condition becomes  $\phi_c \approx \phi_0$  (e.g.,  $a\phi_c/a_c \approx 0.9$  for  $a > a_c$ ). However, we note that this analogy has to be taken with caution unless  $a \gg a_c$ . For  $N = 50$ , for instance, crowding effects in the case  $a_c > a$  become moderately important only beyond  $\phi_c \approx 0.5$ – $0.6$  (Fig. 5(a)). Potential kinetic effects involved at this high range of  $\phi_c$  will make it difficult to draw a definite conclusion on the  $N$ -dependence of chain compaction. Nevertheless, the results in eqn (14) are expected to be asymptotically correct for long chains in equilibrium.

In contrast, if we assume that the chain-enveloping volume is “impermeable” to crowders (Fig. 1(c)), for a large- $\phi_c$  range, the chain can be considered as strongly confined inside a spherical shell of radius  $R$ . Under strong confinement, it can be viewed as forming a semidilute solution.<sup>42,59</sup> Pressure balance between the two sub-regions, chain-occupying and crowder-occupying, will determine the chain size as

$$\frac{\phi_c}{a_c^3} + \mathcal{O}(\phi_c^2) \sim \frac{1}{a^3} \left( \frac{a^3 N}{R^3} \right)^{9/4}. \quad (15)$$

Up to  $\mathcal{O}(\phi_c)$ , this leads to

$$\frac{R}{R_0} \sim N^{-4/15} \left( \frac{a_c}{a\phi_c^{1/3}} \right)^{4/9}. \quad (16)$$

In  $\phi_c$  dependence, this is distinct from the third line of equations in eqn (14).

Based on the results in this subsection, especially eqn (14), we have mapped out a diagram describing the behavior of  $R$  as a function of  $N$  for varying degrees of compaction. Fig. 7(a) summarizes our results in a log-log plot for the free-space case. The grey arrow indicates the increasing direction of  $\phi_c$ .

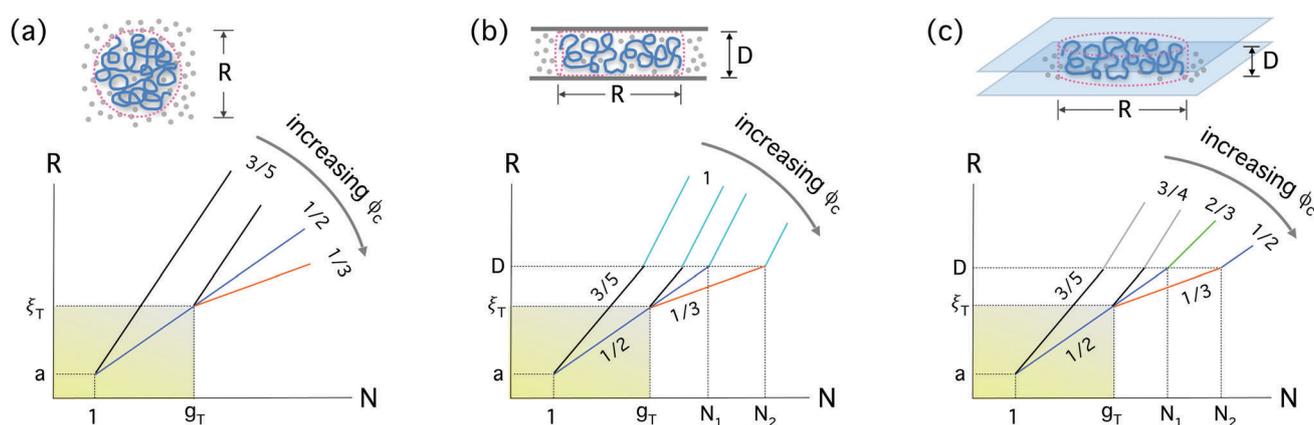
Here, the numbers like  $3/5$ ,  $1/2$ , and  $1/3$  denote distinct scaling regimes: self-avoiding walk, random walk, and collapsed, respectively. Also  $\xi_T = a^4/v$  is the ‘thermal blob’ size and  $g_T \approx a^6/v^2$  is the corresponding contour distance in units of  $a$ .<sup>42</sup>

Testing the diagram in Fig. 7 numerically is computationally inaccessible at present, because of a large parameter space crowders present in the desired long-chain limit. We content ourselves by offering the underlying physics: the diagram in Fig. 7(a), based on our effective-solvent picture, is essentially the same as that for a polymer in various solvents.<sup>52</sup>

It is worth emphasizing that the onset condition for chain compaction becomes  $N$ -independent in the limit  $N \gg 1$ , i.e.,  $\phi_c \approx \phi_0$  except in the highly asymmetrical case  $a_c \gg a$  (see Section 3.2 for relevant discussions for  $a_c \geq a$ ). This can be contrasted against what the physical picture in Fig. 1(c) predicts. Similar to a spherically-confined chain,<sup>59</sup> the fate of the chain in Fig. 1(c) will always be in a collapse state in the limit  $N \rightarrow \infty$ . This is not supported by Fig. 1(b) and eqn (14).

Also note that the  $N$ -dependence of  $R$  for  $\phi_c \geq \phi_0$  in eqn (14) will be different for other spaces. This does not arise from any interdependence between molecular crowding and the geometry of a space. It rather reflects the way a chain molecule collapses when the monomer–monomer interaction becomes attractive, as indicated by eqn (11). In all cases, molecular crowding reduces  $v$  according to eqn (5). We discuss other spaces below.

**3.4.2 In a cylindrical space.** For a cylindrically-confined space, eqn (11) implies that  $R \sim N$  for large  $N$ , independent of the value or sign of  $v$ . In this case, the parameter  $Y \sim N^2$  increases rapidly with  $N$ . As a result, the three-body term (the last term) in eqn (11) is felt more strongly. This is responsible for the stretching of the chain. This case can be contrasted against an “ideal” chain under cylindrical confinement, for which  $R = aN^{1/2}$ .<sup>42,52</sup> Can this behavior be realized in the cylindrically-confined case?



**Fig. 7** Diagram of chain collapse in a log-log plot of  $R$  against  $N$  in three different spaces: free (a), cylindrical (b), and slit-like spaces (c). Here, the numbers like  $3/5$ ,  $1/2$ , and  $1/3$  denote distinct scaling regimes: self-avoiding walk, random walk, and collapsed, respectively. Also,  $\xi_T$  is the thermal blob size and  $g_T$  is the corresponding contour distance in units of  $a$ ;  $\xi$  is the blob size inside which self-avoidance is unscreened and  $g \approx (\xi/a)^{5/3}$ . Each diagram shows how  $R$  varies with  $N$  as  $\phi_c$  increases; the grey arrow indicates the increasing direction of  $\phi_c$ . Note that the diagram in (a) is essentially the same as for a chain in various solvents.<sup>52</sup> There are both similarities and differences among the diagrams. The common regions are shaded inside which the chain does not feel confinement in (b) and (c); outside the region, confinement or the geometry of a confined space modifies chain statistics. This is responsible for the differences. It is worth noting that these diagrams are valid for both  $a > a_c$  and  $a_c \geq a$ , except for the highly asymmetrical case  $a_c \gg a$ .

It can be shown that the ideal-chain behavior  $R = aN^{1/2}$  holds for some range of  $N$ . If this behavior is assumed, the monomer density  $\rho \sim N/RD^2 \sim N^{1/2}/D^2a$ . This diverges as  $N \rightarrow \infty$ . So does the three body interaction. This implies that the ideal chain behavior remains valid for some range of  $N$ . Eqn (11) can be readily solved for large  $N$ : for  $v = 0$ , the free-energy minimum occurs at  $R \sim Na^2/D$ ; for  $v < 0$ ,  $R \sim Na(v_0/|v|)(a/D)^2$ .

If we now try  $R = aN^{1/2}$  with  $v = 0$  in eqn (10) and set the last term to unity, we find  $N \sim D^2/a^2 \equiv N_1$ . Also the corresponding physical distance is  $R_1 \sim D$ . In summary,  $R = a\sqrt{N}$  for  $N < N_1$  and  $R \sim a^2N/D$  for  $N > N_1$ .

The result  $R \sim Na(v_0/|v|)(a/D)^2$  for  $v < 0$  is valid for some large  $N$ . For small  $N$ , we have  $R \sim N^{1/3}a(v_0/|v|)^{1/3}$ . If we use this in the last two terms in eqn (10) and balance them, we find  $N \sim D^3|v|/a^6 \equiv N_2$ , corresponding to a physical distance  $R_2 \sim D$ . This means that  $R \sim N^{1/3}a(v_0/|v|)^{1/3}$  for  $N < N_2$  and  $R \sim Na(v_0/|v|)(a/D)^2$  for  $N > N_2$ .

$$\frac{R}{R_0} \sim \begin{cases} (v/v_0)^{1/4} \sim (1 - \alpha_1 \phi_c)^{1/4} \sim 1 - \alpha_1 \phi_c/4, & \phi_c < \phi_0 \\ (a/D)^{1/12} N^{-1/12}, & \phi_c = \phi_0 \\ (a/D)^{1/4} (v_0/|v|)^{1/2} N^{-1/4} \sim (a/D)^{1/4} N^{-1/4} |1 - \alpha_1 \phi_c + \alpha_2 \phi_c^2 - \alpha_3 \phi_c^3|^{-1/2}, & \phi_c > \phi_0 \end{cases} \quad (19)$$

In summary, for  $N \rightarrow \infty$ , we find

$$\frac{R}{R_0} \sim \begin{cases} (v/v_0)^{1/3} \sim (1 - \alpha_1 \phi_c)^{1/3} \sim 1 - \alpha_1 \phi_c/3, & \phi_c < \phi_0 \\ (a/D)^{1/3}, & \phi_c = \phi_0 \\ (a/D)^{4/3} v_0/|v| \sim (a/D)^{4/3} |1 - \alpha_1 \phi_c + \alpha_2 \phi_c^2 - \alpha_3 \phi_c^3|^{-1}, & \phi_c > \phi_0 \end{cases} \quad (17)$$

These results are essentially identical to the results for the better known problem of a polymer brush in various solvents, as long as  $D$  is interpreted as the spacing between two adjacent grafting points.<sup>60,61</sup>

In contrast, if we balance the pressures between the chain-occupying and chain-free regions for large  $\phi_c$ , as in eqn (15), we arrive at

$$\frac{R}{R_0} \sim \left(\frac{a}{D}\right)^{4/3} \left(\frac{a_c}{a\phi_c^{1/3}}\right)^{4/3} \quad (18)$$

The  $D$  or  $N$  dependence of this result is the same as that of eqn (17) for  $\phi_c > \phi_0$ . But they differ from each other in  $\phi_c$  or  $a_c$  dependence.

Fig. 7(b) shows a diagram in which  $R$  is plotted against  $N$  in a log-log plot in a cylindrical space. Here, the numbers 1, 3/5, 1/2, and 1/3 denote distinct scaling regimes: linear, self-avoiding walk, random walk, and collapsed, respectively. The emergence of the linear regime is unique to the cylindrical case.<sup>38,42,52</sup> Note that the shaded region in the diagram in Fig. 7(b) is essentially identical to the better known one in Fig. 7(a). Cylindrical confinement, however, modifies the diagram in (b) beyond  $D$ ,

as expected for a linearly-organized chain under cylindrical confinement.<sup>42,52</sup>

**3.4.3 In a slit-like space.** Eqn (11) can be analyzed for a chain trapped inside a slit-like space:  $R \sim N^{3/4}a(a/D)^{1/4} \equiv R_0$  for  $v = v_0$  (*i.e.*,  $\phi_c = 0$ ). More realistically, one can view the confined chain as a string of blobs of size  $\xi$  each, inside which self-avoidance is unscreened;  $g \approx (\xi/a)^{5/3}$  is the number of steps to travel a distance  $\xi$ .<sup>42,59</sup> Nevertheless the more sophisticated free energy based on this picture reproduces the same scaling result for  $R$ ,<sup>52,57</sup> even though it has a different  $D$  dependence.

For  $v = 0$ , analysis of eqn (11) leads to  $R \sim N^{1/2}a$  for  $N < N_1 \equiv D^2/a^2$  and  $R \sim N^{2/3}a(a/D)^{1/3}$  for  $N > N_1$ . When  $N = N_1$ , the fourth (three-body) term is comparable to the first one in eqn (11).

On the other hand, for  $v < 0$ , we find  $R \sim N^{1/3}a(v_0/|v|)^{1/3}$  for  $N < N_2 \equiv (D/a)^3|v|/v_0$  and  $R \sim N^{1/2}a\sqrt{a^4/|v|D}$  for  $N > N_2$ . The last two terms in eqn (10) are comparable if  $N = N_2$ .

In the limit  $N \rightarrow \infty$ , these results can be summarized as

In contrast, a pressure-balance relation leads to

$$\frac{R}{R_0} \sim \left(\frac{a}{D}\right)^{1/4} N^{-1/4} \left(\frac{a_c}{a\phi_c^{1/3}}\right)^{2/3} \quad (20)$$

It differs from the result for the case  $\phi_c > \phi_0$  in eqn (19) in  $\phi_c$  dependence.

Fig. 7(c) shows chain collapse in a log-log plot of  $R$  against  $N$  in a slit-like space. The numbers like 3/5, 3/4, 1/2, and 1/3 denote distinct scaling regimes: self-avoiding walk, two-dimensional self-avoiding walk, random walk, and collapsed, respectively. The emergence of the 2/3-regime is unique to this case.

As a final remark in this subsection (Section 3.4), we wish to mention that as long as the effective-solvent picture remains applicable as demonstrated in Fig. 4(b) and 5(b), the diagrams in Fig. 7 remain valid whether  $a > a_c$  or  $a < a_c$ . Also recall that the effect of biological confinement is insignificant (or less significant) for proteins and RNA, which belong to the large- $a_c$  case. This means that they lie in some subspace inside the rectangle formed by  $R = D$  and  $N = N_2$  in the diagrams in Fig. 7(b) and (c). In this subspace, these two diagrams are essentially identical to the corresponding diagram in the bulk (a). This may justify the neglect of confinement in our consideration of the large  $a_c$  case.

## 4 Conclusions

In conclusion, we have obtained a number of general relations characterizing the spatial organization of a biopolymer in a crowded medium. They have been useful for acquiring a deeper understanding of crowding effects in biomolecular organization in a cellular space. For  $a > a_c$ , our polymer may be considered as a coarse-grained model of bacterial chromosomes.<sup>16,17</sup> In this

case, molecular crowding depends on the ratio  $a\phi_c/a_c$  in three distinct spaces: free, cylindrical, and slit-like spaces.

As  $a_c$  increases from a value smaller than  $a$ , molecular crowding becomes less effective for a given  $\phi_c$  value and depends on  $\phi_c$  only beyond  $a_c \approx a$ . If taken together, this means that if  $a_c > a$  the full compaction condition reads  $\phi_c \approx 1$  (or  $\phi_c \approx 0.55$  more accurately even for  $N = 2000$ ) and may not be easily realized at a biologically-relevant  $\phi_c$  range.

A good example of the large- $a_c$  case is protein folding in a crowded, cellular environment. While our simulation results imply that crowding in this case has moderate effects, a real protein is not simply beads on a string as assumed in our considerations. What is unclear is how depletion forces and other interactions orchestrate in folding a protein chain into its biologically-active structure. Obviously, they become more important in organizing higher order structures, *e.g.*, protein aggregates.<sup>12</sup>

We have also shown to what extent crowding effects can be mimicked by adjusting the solvent quality or the excluded volume of monomers. The effective-solvent picture works well whether  $a > a_c$  or  $a < a_c$ , except in the highly asymmetric case  $a_c \gg a$ . A polymer solution mixed with micron-size colloids belongs to the  $a_c \gg a$ -case.<sup>15,25–27</sup> Understandably, the effective-solvent picture, in which the presence of crowders renormalizes the monomer–monomer interactions, will break down in this limit.

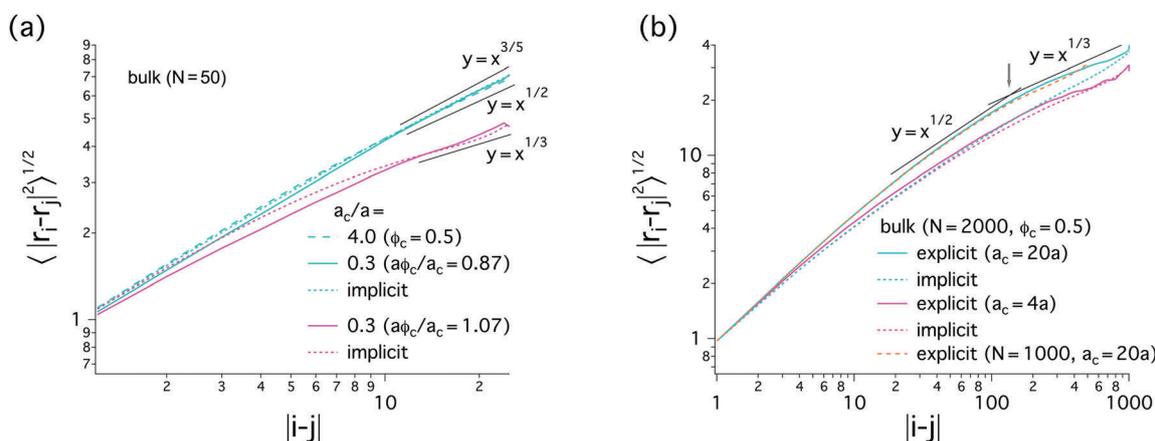
Our results suggest that molecular crowding is insensitive to the geometry or presence of a confined space, possibly except for  $a_c \gg a$ . The effect of molecular crowding is local, similar to that of solvent. However, it does not necessarily contradict the recent finding that the way the DNA molecules respond to

crowding is different between different spaces: bulk, slit-like, and cylindrical.<sup>19</sup> For stiff chains, the spatial organization of crowders in response to the chains is implicated in the way they respond to crowding. In this sense, the difference is not an intrinsic one but arises from the preferential distribution of crowders near the cylindrical wall below the onset of chain collapse.<sup>19</sup>

More realistically, a heterogeneous polymer can be considered as a coarse-grained model of bacterial chromosomes organized with various proteins. A dominant source for chain heterogeneity is rRNA polymerases concentrated in several designated sites along the chromosome, *i.e.*, rRNA operons.<sup>9,10,45–47</sup> At the crudest but non-trivial level, chain heterogeneity can be mimicked by introducing small and big monomers<sup>10</sup> with the latter representing transcription-active sites (Fig. 2). How the resulting heterogeneity and crowding are intertwined in a confined space has not been well understood beyond a recent attempt, which was focused on analyzing the clustering tendency of big monomers *via* looping in a free space,<sup>10</sup> but can be further unraveled. A related point is that chain heterogeneity can be responsible for the coexistence of condensed and extended phases observed with *E. coli* chromosomes.<sup>8</sup> A computational approach along the line of what is illustrated in Fig. 2 will be useful.

## Appendix

In this Appendix, we test further the effective-solvent picture, especially for the case  $a_c > a$ . To this end, we consider the internal distances of monomers at  $i$  and  $j$  defined as  $\langle |\mathbf{r}_i - \mathbf{r}_j|^2 \rangle^{1/2}$ . Fig. 8 shows the internal distance for  $N = 50$  (a) and  $N = 2000$  (b).



**Fig. 8** Internal distance for  $N = 50$  (a) and  $N = 2000$  (b), and the validity of the effective-solvent picture. (a) For the parameters chosen for the cyan curves (*e.g.*,  $a_c = 4a$  and  $\phi_c = 0.5$ ), the collapsed regime does not emerge; there is a crossover from the self-avoiding to random walk regimes. The curves in magenta, both solid and dashed, show a collapsed regime described by a straight line with a slope  $1/3$ . In both cases, there is a good agreement between the explicit-crowder and implicit cases. (b) For  $a_c = 4a$  (curves in magenta), the explicit (solid) and implicit (dashed) curves collapse onto each other. In contrast, for  $a_c = 20a$ , the two cases (solid and dashed) deviate appreciably from each other. This means that the effective-solvent picture does not work well for the highly symmetrical case  $a_c/a \gg 1$ . For  $a_c = 4a$  and for the solid curve for  $a_c = 20a$ , three distinct regimes emerge: self-avoiding, random walk, and collapsed regimes at short, intermediate, and large scale ranges of  $|i-j|$ , respectively. Superimposed is a dashed line in magenta, obtained for  $N = 1000$  and  $a_c = 20a$  (explicit crowders). The onset of chain collapse in this case coincides with that shown in the cyan curve. This implies that depletion forces are local even for  $a_c = 20a$ , even though the effective solvent picture is inaccurate; the chain-occupying region is still permeable to crowders.

In each graph, we compare the explicit-crowder and effective-solvent pictures: solid and dashed lines, respectively.

First in Fig. 8(a) ( $N = 50$  in the bulk), we compare the two pictures for a few combinations of  $a_c$  ( $a_c = 0.3a, 4a$ ) and  $\phi_c$ . For the cyan curves, the chain is only weakly compressed, but for the magenta curve, it is compressed strongly, as marked by the emergence of the collapsed regime. The agreement between the two pictures is generally good. For  $a_c = 0.3a$  and  $a\phi_c/a_c = 1.07$  (strongly collapsed), the two sets of results (explicit and implicit) deviate somewhat from each other for  $|i-j| \leq 10$ . Note that chain statistics at short-length scales is model-dependent. Nevertheless the discrepancy is still insignificant.

The graph Fig. 8(b) represents a much longer chain, *i.e.*,  $N = 2000$ . For  $a_c = 4a$ , the explicit (solid) and effective-solvent (dashed) curves collapse onto each other. To obtain the effective-solvent or implicit curves,  $\nu$  was adjusted so that the resulting  $R/R_0$  would match the corresponding explicit-crowder result as in Fig. 4. In contrast, for  $a_c = 20a$ , the two cases (solid and dashed) deviate appreciably from each other. This means that the effective-solvent picture does not work well for the highly symmetrical case  $a_c/a \gg 1$ . (This includes the case  $a_c > R_g$ . See the footnote† for relevant discussion.) For  $a_c = 4a$  (both explicit and implicit) and for the solid curve for  $a_c = 20a$ , three distinct regimes emerge to characterize the local structure of the chain: self-avoiding, random walk, and collapsed regimes at short, intermediate, and large scale ranges of  $|i-j|$ , respectively. This is consistent with Fig. 7, except for  $|i-j| \leq 10$ . As noted above, chain statistics is model dependent for sufficiently small  $|i-j|$ . The bead-spring chain used in our simulations is not perfectly flexible, as often assumed in an “idealized” theoretical model, *e.g.*, a freely-jointed chain with self-avoidance. This delays the emergence of the random-walk regime in Fig. 8(b).

Also superimposed is a dashed line in tangerine obtained for  $N = 1000$  and with explicit crowders of size  $a_c = 20a$  each. Comparing it with the corresponding case for  $N = 2000$ , one can conclude that the onset of the collapsed regime marked by the vertical arrow in grey is the same for the two choices of  $N$ :  $N = 1000$  and  $N = 2000$ . This is aligned with the picture that the chain-occupying region is permeable to crowders and suggests that molecular crowding is local, even though the effective-solvent picture does not work well for  $a_c = 20a$ . The local picture will eventually break down for sufficiently large  $a_c/a$  values. This is most obvious when  $a_c > R_g$ . In this case, Fig. 1(b) is not applicable anymore.

## Acknowledgements

We benefited from helpful discussions with A. Grosberg, M. Rubinstein, and C. Hyeon. This work was supported by NSERC (Canada) (B-YH), the collaborative research contract funded by Korea Institute of Science and Technology Information (KISTI), and the Basic Science Research Program through the grant no. 2015R1D1A1A09057469 (YJ) as well as by KIAS (Korea Institute for Advanced Study). We acknowledge the computational resources of the Shared Hierarchical Academic

Research Computing Network (SHARCNET: [www.sharcnet.ca](http://www.sharcnet.ca)) and Compute/Calcul Canada.

## References

- 1 R. J. Ellis, Macromolecular crowding: obvious but underappreciated, *Trends Biochem. Sci.*, 2001, **26**, 597–604.
- 2 S. B. Zimmerman and A. P. Minton, Macromolecular Crowding: Biochemical, Biophysical, and Physiological Consequences, *Annu. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 27–65.
- 3 H.-X. Zhou, G. Rivas and A. P. Minton, Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences, *Annu. Rev. Biophys.*, 2008, **37**, 375–397.
- 4 K. A. Sharp, Unpacking the origins of in-cell crowding, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 1684–1685, and references therein.
- 5 J. A. Valkenburg and C. L. Woldringh, Phase Separation Between Nucleoid and Cytoplasm in *Escherichia coli* as Defined by Immersive Refractometry, *J. Bacteriol.*, 1984, **160**, 1151–1157.
- 6 C. L. Woldringh and T. Odijk, in *Organization of the Prokaryotic Genome*, ed. R. L. Charlebois, ASM Press, Washington, D.C., 1999.
- 7 J. Stavans and A. Oppenheim, DNA–protein interactions and bacterial chromosome architecture, *Phys. Biol.*, 2006, **3**, R1–R10.
- 8 J. Pelletier, K. Halvorsen, B. Y. Ha, R. Paparcone, S. J. Sandler, C. L. Woldringh, W. P. Wong and S. Jun, Physical manipulation of the *Escherichia coli* chromosome reveals its soft nature, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, E2649–E2656.
- 9 S. Jun, Chromosome, Cell Cycle and Entropy, *Biophys. J.*, 2015, **108**, 785–786.
- 10 D. Marenduzzo, C. Micheletti and P. R. Cook, Entropy-Driven Genome Organization, *Biophys. J.*, 2006, **90**, 3712–3721.
- 11 D. Marenduzzo, I. Faro-Trindade and P. R. Cook, What are the molecular ties that maintain genomic loops?, *Trends Genet.*, 2007, **23**, 126–133.
- 12 H. X. Zhou, Influence of crowded cellular environments on protein folding, binding, and oligomerization: Biological consequences and potentials of atomistic modeling, *FEBS Lett.*, 2013, **587**, 1053–1061.
- 13 S. Asakura and F. Oosawa, On Interaction between Two Bodies Immersed in a Solution of Macromolecules, *J. Chem. Phys.*, 1954, **22**, 1255–1256.
- 14 S. Asakura and F. Oosawa, Interaction between particles suspended in solutions of macromolecules, *J. Polym. Sci., Part A: Polym. Chem.*, 1958, **33**, 183–192.
- 15 H. N. W. Lekkerkerker and R. Tuinier, *Colloids and the Depletion Interaction*, Lecture Notes in Physics, Springer, 2011, vol. 833.
- 16 J. Kim, C. Jeon, H. Jeong, Y. Jung and B.-Y. Ha, A polymer in a crowded and confined space: effects of crowder size and poly-dispersity, *Soft Matter*, 2015, **11**, 1877–1888.

- 17 T. N. Shendruk, M. Bertrand, H. W. de Haan, J. L. Harden and G. W. Slater, Simulating the Entropic Collapse of Coarse-Grained Chromosomes, *Biophys. J.*, 2015, **108**, 810–820.
- 18 T. N. Shendruk, M. Bertrand, J. L. Harden, G. W. Slater and H. W. de Haan, Coarse-grained molecular dynamics simulations of depletion-induced interactions for soft matter systems, *J. Chem. Phys.*, 2014, **141**, 244910.
- 19 J. J. Jones, J. R. C. van der Maarel and P. S. Doyle, Effect of Nanochannel Geometry on DNA Structure in the Presence of Macromolecular Crowding Agent, *Nano Lett.*, 2011, **11**, 5047–5053.
- 20 H. Kang, P. A. Pincus, C. Hyeon and D. Thirumalai, Effects of Macromolecular Crowding on the Collapse of Biopolymers, *Phys. Rev. Lett.*, 2015, **114**, 068303.
- 21 B.-Y. Ha and Y. Jung, Polymers under confinement: single polymers, how they interact, and as model chromosomes, *Soft Matter*, 2015, **11**, 2333–2352.
- 22 B. Youngren, H. J. Nielsen, S. Jun and S. Austin, The multi-fork *Escherichia coli* chromosome is a self-duplicating and self-segregating thermodynamic ring polymer, *Genes Dev.*, 2014, **28**, 71–84.
- 23 R. Milo and R. Phillips, *Cell Biology by the Numbers*, Garland Science, 2015.
- 24 R. Phillips, J. Kondev, J. Theriot and H. Garcia, *Physical Biology of the Cell*, Garland Science, 2nd edn, 2012.
- 25 D. H. Napper, *Polymeric stabilization of colloidal dispersions*, Academic Press, London, 1983.
- 26 M. R. Shaw and D. Thirumalai, Free polymer in a colloidal solution, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1991, **44**, R4797–4800.
- 27 K. Binder, P. Virnau and A. Statt, Perspective: The Asakura Oosawa model: A colloid prototype for bulk and interfacial phase behavior, *J. Chem. Phys.*, 2014, **141**, 140901.
- 28 J. D. Honeycutt and D. Thirumalai, Static properties of polymer chains in porous media, *J. Chem. Phys.*, 1989, **90**, 4542–4559.
- 29 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, 1987.
- 30 D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Academic Press, 2002.
- 31 M. Castelnovo and W. M. Gelbart, Semiflexible Chain Condensation by Neutral Depleting Agents: Role of Correlations between Depletants, *Macromolecules*, 2004, **37**, 3510–3517.
- 32 A. Y. Grosberg, I. Y. Erukhimovitch and E. I. Shakhnovitch, On the theory of  $\Psi$ -condensation, *Biopolymers*, 1982, **21**, 2413–2432.
- 33 J. D. Weeks, D. Chandler and H. C. Andersen, Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids, *J. Chem. Phys.*, 1971, **54**, 5237–5247.
- 34 K. Kremer and G. S. Grest, Dynamics of entangled linear polymer melts- A molecular-dynamics simulation, *J. Chem. Phys.*, 1990, **92**, 5057–5086.
- 35 G. S. Grest and K. Kremer, Molecular dynamics simulation for polymers in the presence of a heat bath, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1986, **33**, 3628–3631.
- 36 J. Kim, C. Jeon, H. Jeong, Y. Jung and B.-Y. Ha, Elasticity of flexible polymers under cylindrical confinement: appreciating the blob scaling regime in computer simulations, *Soft Matter*, 2013, **9**, 6142–6150.
- 37 Y. Jung, C. Jeon, J. Kim, H. Jeong, S. Jun and B.-Y. Ha, Ring polymers as model bacterial chromosomes: confinement, chain topology, single chain statistics, and how they interact, *Soft Matter*, 2012, **8**, 2095–2102.
- 38 Y. Jung, J. Kim, S. Jun and B.-Y. Ha, Intrachain ordering and segregation of polymers under confinement, *Macromolecules*, 2013, **45**, 3256–3262.
- 39 J. S. Kim, V. Backman and I. Szleifer, Crowding-Induced Structural Alterations of Random-Loop Chromosome Model, *Phys. Rev. Lett.*, 2011, **106**, 168102.
- 40 P. van der Schoot, Protein-Induced Collapse of Polymer Chains, *Macromolecules*, 1998, **31**, 4635–4638.
- 41 R. P. Sear, Flory-Huggins theory for athermal mixtures of hard spheres and larger flexible polymers, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2002, **66**, 051401.
- 42 P. G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University, 1979.
- 43 M. Fritsche, S. Li, D. W. Heermann and P. A. Wiggins, A model for *Escherichia coli* chromosome packaging supports transcription factor-induced DNA domain formation, *Nucleic Acids Res.*, 2012, **40**, 972–980.
- 44 I. Oh, S. Choi, Y. J. Jung and J. S. Kim, Unusual size-dependence of effective interactions between collapsed polymers in crowded environments, *Soft Matter*, 2014, **10**, 9098–9104.
- 45 F. C. Neidhardt, J. L. Ingraham and M. Schaechter, *Physiology of the Bacterial Cell: A Molecular Approach*, Sinauer Associates, 1990.
- 46 H. Bremer and P. Dennis, Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. F. C. Neidhardt, ASM Press, 1996.
- 47 O. L. Miller, Jr., B. A. Hamkalo and C. A. Thomas, Jr., Visualization of Bacterial Genes in Action, *Science*, 1970, **169**, 392–395.
- 48 W. M. Gelbart, R. F. Bruinsma, P. A. Pincus and V. A. Parsegian, DNA-Inspired Electrostatics, *Phys. Today*, 2000, **53**, 38–44.
- 49 J. X. Tang, P. A. Janmey, A. Lyubartsev and L. Nordenskiöld, Metal Ion-Induced Lateral Aggregation of Filamentous Viruses fd and M13, *Biophys. J.*, 2002, **83**, 566–581.
- 50 R. de Vries, DNA condensation in bacteria: Interplay between macromolecular crowding and nucleoid proteins, *Biochimie*, 2010, **92**, 1715–1721.
- 51 A. Arnold, B. Bozorgui, D. Frenkel, B.-Y. Ha and S. Jun, Unexpected relaxation dynamics of a self-avoiding polymer in cylindrical confinement, *J. Chem. Phys.*, 2007, **127**, 164903.
- 52 M. Rubinstein and R. H. Colby, *Polymer Physics*, Oxford University Press, 2003.
- 53 J. des Cloizeaux and G. Jannink, *Polymers in Solution: Their Modelling and Structure*, Oxford University Press, 1990.
- 54 P. G. de Gennes, Collapse of a polymer chain in poor solvents, *J. Phys. Lett.*, 1975, **36**, 55–57.

- 55 W. Reisner, J. N. Pedersen and R. H. Austin, DNA confinement in nanochannels: physics and biological applications, *Rep. Prog. Phys.*, 2015, **75**(2012), 106601.
- 56 S. Jun, D. Thirumalai and B.-Y. Ha, Compression and Stretching of a Self-Avoiding Chain in Cylindrical Nanopores, *Phys. Rev. Lett.*, 2008, **101**, 138101.
- 57 Y. Jung, S. Jun and B.-Y. Ha, Self-avoiding polymer trapped inside a cylindrical pore: Flory free energy and unexpected dynamics, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2009, **79**, 061912.
- 58 A. R. Khokhlov, A. Y. Grosberg and V. S. Pande, *Statistical Physics of Macromolecules*, AIP, 1994.
- 59 S. Jun, A. Arnold and B.-Y. Ha, Confined Space and Effective Interactions of Multiple Self-Avoiding Chains, *Phys. Rev. Lett.*, 2007, **98**, 128303.
- 60 A. Halperin and E. B. Zhulina, Stretching polymer brushes in poor solvents, *Macromolecules*, 1991, **24**, 5393–5397.
- 61 A. Halperin, Collapse of grafted chains in poor solvents, *J. Phys.*, 1988, **49**, 547–550.