Hydrodynamic techniques in the study of elastic macromolecules

Radost Waszkiewicz

University of Warsaw
Faculty of Physics

A dissertation submitted to the University of Warsaw
for the degree of Doctor of Philosophy
under the supervision of dr hab. Maciej Lisicki, prof. UW

Warsaw, 2024–01–19

Summary

Significant proportion of the work leading to this dissertation was part of experimental-theoretical collaborations with two groups specialising in studying different elastic molecules – group of prof. Lynn Zechiedrich focussed on DNA loops and group of prof. Anna Niedźwiecka.

First publication titled *Stability of sedimenting flexible loops* co-authored by Piotr Szymczak and Maciej Lisicki provides linear stability analysis of elastic loops within the resistive-force theory framework coupled with elastic forces modelled with Euler-Bernoulli equation. We were able to establish a semi-analytic stability criterion and rederive the dimensionless quantity governing the buckling instability for this and similar problems.

Second publication titled *Hydrodynamic effects in the capture of rod-like molecules by a nanopore* co-authord by Maciej Lisicki provides analysis of the influence of wall interaction and hydrodynamic anisotropy in the process of nanopore capture. A theoretical consideration of a rod-like molecule with uniformly distributed charge provides simple scaling-based criteria for determining when and where inclusion of the wall corrections is required.

Third publication titled *Pychastic: Precise Brownian dynamics using Taylor-Ito integrators in Python* co-authored by Maciej Bartczak, Kamil Kolasa and Maciej Lisicki is a result of work on implementation of efficient stochastic differential equations solvers, capable of convenient treatment of Brownian Dynamics (BD) problems. By expressing BD equations as Itō integrals we can leverage classical methods of truncated Taylor-Itō integrators. As part of the documentation of the `pychastic` package we show how to deal with common BD obstacles: calculations of the divergence of the mobility tensor in the diffusion equation, and discontinuous trajectories encountered when working with dynamics on S^2 and $SO(3)$. With vectorisation-oriented implementation we have achieved performance comparable to earlier implementations in lower-level programming languages.

Fourth publication titled *DNA supercoiling-induced shapes alter minicircle hydrodynamic properties* co-authord by Maduni Ranasinghe, Jonathan M Fogg, Daniel J Catanese Jr, Maria L Ekiel-Jeżewska, Maciej Lisicki, Borries Demeler, Lynn Zechiedrich, and Piotr Szymczak is a result of theoretical-experimental collaboration with the team based at Baylor College of medicine (responsible for biosynthesis) and the team based at Unviersity of Lethbridge (responsible for analytical ultracentrifugation experiments). In this publication we have determined the impact of negative supercoiling and curvature on the hydrodynamic properties of DNA by subjecting 336 bp and 672 bp DNA minicircles to analytical ultracentrifugation (AUC). We then utilized linear elasticity theory and hydrodynamic calculations to predict DNA shapes and diffusion coefficients.

Fifth manuscript titled *Minimum dissipation approximation: A fast algorithm for the prediction of diffusive properties of intrinsically disordered proteins* co-authored by Agnieszka Michaś, Michał K. Białobrzewski, Barbara Klepka, Maja Cieplak-Rotowska, Zuzanna Staszałek, Bogdan Cichocki, Maciej Lisicki, Piotr Szymczak, and Anna Niedźwiecka is a result of experimental-theoretical collaboration with the team based at Istitute of Physics of Polish Academy of Sciences (responsible for biosynthesis and fluorescence correlation spectroscopy). In our study, we demonstrate a fast numerical method combining simple conformational sampling and approximate hydrodynamic interactions to estimate the diffusion coefficients of intrinsically disordered proteins (IDPs), even in the presence of structured domains, with a precision surpassing the classical Kirkwood approximation. With a new collection of diffusion coefficient measurements we can quantitatively compare our predictions with multiple models present in literature (such as power-laws and power-laws with sequence-dependent corrections).

Sixth manuscript ttiled *The trimer paradox revisited* co-authored by Maciej Lisicki deals with the problem of very stiff constraints and limiting bond-angle distributions arising from those. It shows that even though the solution of the paradox was elucidated as early as 1984 in Van Kampen and Lodder [54], explicit (and correct) treatment of this limit was missing from well known books such as Frenkel and Smit [25]. By more careful treatment of singular distributions we show that a combination of metric properties of the constraining manifold and hessian of the constraining field are required for correct determination of bond angles, and that 'uniform on a sphere' distributions for harmonic constraining potentials are not universal, with potentially large deviations for small cyclic molecules. These results establish theoretical foundations for the globule-linker model and should guide further work on the minimum dissipation approximation.

The aim of the presented work was twofold, to address the immediate needs of the experimental groups we have been collaborating with, but also to establish robust 'null hypothesis' models, which are easy enough to use and incorporate all of the fundamental interactions required for diffusion modelling but no more interactions. As such deviations from these models can be used as quantitative indicators of

significant contribution of new physical phenomena (for example electrostatic interactions or formation of transient bridges between distant parts of the molecule).

Streszczenie

Pierwsza publikacja pt. *Stabilność sedymentujących elastycznych pętli* współautorami Piotra Szymczaka i Macieja Lisickiego zapewnia liniową analizę stabilności pętli sprężystych w ramach teorii siły oporu w połączeniu z siłami sprężystymi modelowanymi za pomocą równania Eulera-Bernoulliego. Udało nam się ustalić półanalityczne kryterium stabilności i ponownie wyprowadzić bezwymiarową wielkość regulującą niestabilność wyboczeniową dla tego i podobnych problemów.

Druga publikacja pt. *Efekty hydrodynamiczne w wychwytywaniu pręcików przez nanopor* współautor: Maciej Lisicki przedstawia analizę wpływu oddziaływania ścianek i anizotropii hydrodynamicznej na proces wychwytywania nanoporów. Teoretyczne rozważenie cząsteczki przypominającej pręcik z równomiernie rozłożonym ładunkiem dostarcza prostych, opartych na skalowaniu kryteriów pozwalających określić, kiedy i gdzie wymagane jest uwzględnienie poprawek ścian.

Trzecia publikacja zatytułowana *Pychastic: Precise Brownian dynamics using integrators Taylor-Ito w Pythonie* współautorami: Maciej Bartczak, Kamil Kolasa i Maciej Lisicki jest wynikiem prac nad implementacją wydajnych rozwiązań stochastycznych równań różniczkowych, zdolnych do wygodnego rozwiązywania problemów dynamiki Browna (BD). Wyrażając równania BD jako całki Itō, możemy wykorzystać klasyczne metody obciętych integratorów Taylora-Itō. W ramach dokumentacji pakietu `pychastic` pokazujemy, jak radzić sobie z typowymi przeszkodami BD: obliczeniami rozbieżności tensora ruchliwości w równaniu dyfuzji oraz nieciągłymi trajektoriami spotykanymi podczas pracy z dynamiką na S^2 i $SO(3)$. Dzięki implementacji zorientowanej na wektoryzację osiągnęliśmy wydajność porównywalną z wcześniejszymi implementacjami w językach programowania niższego poziomu.

Czwarta publikacja zatytułowana *Kształty wywołane superskręceniem DNA zmieniają właściwości hydrodynamiczne minikola* współautorami: Maduni Ranasinghe, Jonathan M Fogg, Daniel J Catanese Jr, Maria L Ekiel-Jeżewska, Maciej Lisicki, Borries Demeler, Lynn Zechiedrich i Piotr Szymczak jest efektem teoretyczno-eksperymentalnej współpracy z zespołem z Baylor College of Medicine (odpowiedzialnym za biosyntezę) oraz zespołem z University of Lethbridge (odpowiedzialnym za analityczne eksperymenty ultrawirowania). W tej publikacji określiliśmy wpływ ujemnego superskręcenia i krzywizny na właściwości hydrodynamiczne DNA, poddając minikola DNA o wielkości 336 bp i 672 bp analitycznemu ultrawirowaniu (AUC). Następnie wykorzystaliśmy teorię sprężystości liniowej i obliczenia hydrodynamiczne, aby przewidzieć kształty DNA i współczynniki dyfuzji.

Piąty manuskrypt zatytułowany *Przybliżenie minimalnego rozproszenia: szybki algorytm przewidywania właściwości dyfuzyjnych białek wewnętrznie nieuporządkowanych* współautorzy: Agnieszka Michaś, Michał K. Białobrzewski, Barbara Klepka, Maja Cieplak-Rotowska, Zuzanna Staszałek, Bogdan Cichocki, Maciej Lisicki, Piotr Szymczak i Anna Niedźwiecka jest efektem współpracy eksperymentalno-teoretycznej z zespołem Instytutu Fizyki PAN (odpowiedzialnym za biosyntezę i spektroskopię korelacyjną fluorescencji). W naszym badaniu demonstrujemy szybką metodę numeryczną łączącą proste próbkowanie konformacyjne i przybliżone interakcje hydrodynamiczne w celu oszacowania współczynników dyfuzji białek wewnętrznie nieuporządkowanych (IDP), nawet w obecności domen strukturalnych, z precyzją przewyższającą klasyczne przybliżenie Kirkwooda. Dzięki nowemu zbiorowi pomiarów współczynników dyfuzji możemy ilościowo porównać nasze przewidywania z wieloma modelami obecnymi w literaturze (takimi jak prawa potęgowe i prawa potęgowe z poprawkami zależnymi od sekwencji).

Szósty rękopis zatytułowany *Ponowne spojrzenie na paradoks trimeru* współautorem Maciej Lisicki zajmuje się problemem bardzo sztywnych wiązań i wynikających z nich ograniczających rozkładów kątów wiązania. Pokazuje, że chociaż rozwiązanie paradoksu zostało wyjaśnione już w 1984 roku w Van Kampen and Lodder [54], w dobrze znanych książkach, takich jak Frenkel and Smit [25], brakowało wyraźnego (i prawidłowego) potraktowania tej granicy. Poprzez dokładniejsze potraktowanie rozkładów osobliwych pokazujemy, że do prawidłowego określenia kątów wiązań wymagana jest kombinacja właściwości metrycznych rozmaitości ograniczającej i hesjanu pola ograniczającego, a rozkłady „równomierne na kuli" dla harmonicznych potencjałów ograniczających nie są uniwersalne , z potencjalnie dużymi odchyleniami dla małych cząsteczek cyklicznych. Wyniki te ustanawiają teoretyczne podstawy modelu globula-łącznik i powinny wyznaczać kierunki dalszych prac nad przybliżeniem minimalnego rozproszenia.

Cel prezentowanej pracy był dwojaki: zaspokojenie bezpośrednich potrzeb grup eksperymentalnych, z którymi współpracowaliśmy, ale także ustalenie solidnych modeli „hipotezy zerowej", które są wystarczająco łatwe w użyciu i uwzględniają wszystkie podstawowe interakcje wymagane do modelowania dyfuzyjne, ale żadnych więcej interakcji. Dzięki temu odchylenia od tych modeli można wykorzystać jako ilościowe wskaźniki istotnego udziału nowych zjawisk fizycznych (na przykład oddziaływań elektrostatycznych lub powstawania przejściowych mostków pomiędzy odległymi częściami cząsteczki).

Funding

Research on the topic of the Thesis was supported by the grant Deformations of filaments in viscous liquids, financed by the National Center of Science in Poland under the grant agreement to Maciej Lisicki no. 2018/31/D/ST3/02408

Contents

1 Introduction **7**
- 1.1 Mesoscopic world of macromolecules . . . 7
- 1.2 Method of mobility matrices . . . 8
- 1.3 Mathematical treatment of Brownian motion . . . 9
- 1.4 The hydrodynamic radius . . . 10
- 1.5 Elastic macromolecules . . . 12
- 1.6 Experimental techniques . . . 12
  - 1.6.1 Analytical Ultracentrifugation . . . 12
  - 1.6.2 Fluorescence Correlation Spectroscopy . . . 13
  - 1.6.3 Small Angle X-ray Scattering . . . 14

2 Discussion **16**
- 2.1 Experimental and theoretical challenges . . . 16
- 2.2 Approaches to Predicting Diffusion Coefficients in the Hot and Cold Limits . . . 17
- 2.3 Combining Hot and Cold Approaches . . . 18
- 2.4 Results not Included in Submitted Manuscripts . . . 19
  - 2.4.1 Thermal Effects on the Shapes of DNA Minicircles . . . 19
  - 2.4.2 Comparing Conformations to SAXS Data . . . 20
- 2.5 Author's contributions . . . 21
- 2.6 Software packages . . . 22

3 Paper: Stability of sedimenting flexible loops **23**

4 Paper: Hydrodynamic effects in the capture of rod-like molecules by a nanopore **24**

5 Paper: DNA supercoiling-induced shapes alter minicircle hydrodynamic properties **25**

6 Paper: Pychastic: Precise Brownian dynamics using Taylor-Itō integrators in Python **26**

7 Paper: Minimum dissipation approximation: A fast algorithm for the prediction of diffusive properties of intrinsically disordered proteins **27**

8 Paper: Trimer paradox revisited **28**

9 Conclusions **29**

Preface

TODO TODO - what was the objective - the results of this study form the content of the present PhD thesis and include: - bullet list of findings - the present dissertation is based on thematically linked publications and preprints: - bullet list of publications follows this thesis is organized as follows. Chapter ... gives, chapter provides, finally in chapter we provide a summary and prospects of further research

Introduction

1.1 Mesoscopic world of macromolecules

A fundamental technique for assessing the relative importance of different physical mechanisms within a theoretical framework, originating in the domain of fluid mechanics (cf. Reynolds, [45]), involves the consideration of dimensionless numbers. Here we follow the approach of [40] whereby he motivates common approximations in colloidal physics by consideration of timescale ratios.

The subject of colloidal hydrodynamics are the properties of colloidal suspensions – mixtures of very small objects and a solvent, typically water. At human-length scale water is easily treated as incompressible (for example by considering volume change at typical pressures that are of the order of atmospheric pressure). This is less obvious at the microscopic length scale where granularity of matter becomes significant. Relevant bulk density relaxation timescale τ_s at the length scales relevant to our colloidal object of size L can be estimated from the speed of sound in water c as $\tau_s = L/c$ which is much shorter than any experimentally relevant time scale prompting us to use the incompressible Navier-Stokes equation

$$\rho\left(\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}\right) = -\boldsymbol{\nabla}p + \eta\Delta\boldsymbol{u} + \boldsymbol{f}, \tag{1.1}$$

together with the incompressible continuity equation

$$\nabla\cdot\boldsymbol{u} = 0. \tag{1.2}$$

The Navier-Stokes equation (1.1) is famously nonlinear and further simplifications are necessary for almost any problem of practical relevance. The relative importance of the nonlinear momentum advection term compared to the viscous dissipation term is measured by the Reynolds number Re. In a quiescent fluid it can be estimated from velocity of a colloidal particle V_p as

$$\mathrm{Re} = \frac{\rho V_p L}{\eta} \sim \frac{|\rho\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}|}{|\eta\Delta\boldsymbol{u}|}. \tag{1.3}$$

Taking $L = 100$ and V_p as determined from equipartition of energy for a particle of mass 100kDa at room temperature we get $\mathrm{Re} \sim TODO$ thus we can disregard the nonlinear terms and arrive at time-dependent Stokes equation

$$\rho\frac{\partial \boldsymbol{u}}{\partial t} = -\boldsymbol{\nabla}p + \eta\Delta\boldsymbol{u} + \boldsymbol{f}. \tag{1.4}$$

By taking curl of both sides of this equation we arrive at the vorticity diffusion equation

$$\frac{\partial}{\partial t}\left(\boldsymbol{\nabla}\times\boldsymbol{u}\right) = \frac{\eta}{\rho}\Delta\left(\boldsymbol{\nabla}\times\boldsymbol{u}\right), \tag{1.5}$$

which has a heat-kernel-type solution with characteristic time of

$$\tau_\omega = \frac{\rho L^2}{\eta}. \tag{1.6}$$

Turns out this timescale is of the same order as the Raighley particle velocity relaxation timescale τ_B (following the naming convention of Van Kampen [53]) describing the half-life of velocity of a particle slowing down due to Stokes drag $F_{stokes} = 6\pi\eta a V_p = \zeta V_p$ as shown by

$$\tau_B = \frac{M}{\zeta} = \frac{2}{9}\left(\frac{\rho_p}{\rho}\right)\tau_\omega, \tag{1.7}$$

where ρ_p is the density of the colloidal particle, which is often neutrally buoyant.

Finally we can define diffusive timescale τ_D as the time required for a particle to move distance comparable with its size

$$\tau_D = \frac{a^2}{D}. \tag{1.8}$$

For colloidal particles $\tau_D \gg \tau_\omega \sim \tau_B$ thus we can neglect time dependent terms of the Stokes equation (this is not always the case, for example in the motion of carpets of cilia or bacterial flagella the time dependent terms play a role [56]). Consequently we arrive at the Stokes equation

$$\eta \Delta \boldsymbol{u} - \boldsymbol{\nabla} p + \boldsymbol{f} = 0. \tag{1.9}$$

Stokes equation (1.9) has important properties of linearity and instant information propagation throughout the domain (velocity at any moment is fully determined by the fluid boundary conditions at the same time instance). These are vital for the construction of mobility matrices discussed in the next section.

1.2 Method of mobility matrices

Extending this the presented reasoning to multiple particles is easier with supervector notation, here following convention of [40].

Let us denote a concatenated vector of particle velocities as $V = (\boldsymbol{V_1}, \boldsymbol{V_2}, ..., \boldsymbol{V_N})^T$, angular velocities and $\Omega = (\boldsymbol{\Omega_1}, \boldsymbol{\Omega_2}, ..., \boldsymbol{\Omega_N})^T$ and analogously for forces $\boldsymbol{F_i}$ and torques $\boldsymbol{T_i}$

From the linearity of the Stokes equation and the no-slip boundary conditions of the fluid velocity at the surface of the colloids we know that they obey a linear relationship. We can introduce mobility tensors μ in the following fashion

$$\begin{pmatrix} V \\ \Omega \end{pmatrix} = - \begin{pmatrix} \mu^{tt} & \mu^{tr} \\ \mu^{rt} & \mu^{rr} \end{pmatrix} \cdot \begin{pmatrix} F \\ T. \end{pmatrix} \tag{1.10}$$

For spherical colloids suspended in a quiescent fluid these depend on relative positions of the spheres. In the lowest order approximation in turns out only pairwise displacements are required to compute them. (An even simpler approximation is also possible where the spheres simply do not interact). Introducing $R_{ij} = |\boldsymbol{R}_{ij}| = |\boldsymbol{R}_i - \boldsymbol{R}_j|$ we can determine the velocity of i^{th} sphere from the surface tractions integral by combining Faxen's law TODO cite with the green function of the Stokes equation called the Osseen tensor $\boldsymbol{T}^0$,

$$\boldsymbol{V_i} = -\mu_0^t F_i + \sum_{j \neq i}^{N} (1 + \frac{a^2}{6} \Delta_i) \int_{S_j} \mathrm{d}S' \boldsymbol{T^0}(\boldsymbol{R}_i - \boldsymbol{r'}) \cdot \boldsymbol{f}^{(s)}(\boldsymbol{r'}). \tag{1.11}$$

Direct use of equation (1.11) is not very practical since it requires solving for surface traction distribution. If we only take into account the average surface traction on each sphere by $\boldsymbol{f}_j^{(s)}(\boldsymbol{r'}) = -\boldsymbol{F}_j/(4\pi a^2)$ taking one obtains relationship of the form (1.10). Next one can approximate the integrand by Taylor expanding around centres of spheres to second order ($\mathcal{O}((a/R_i)^3)$ and from symmetry $\mathcal{O}((a/R_i)^4)$) we get lowest order in a approximation of the mobility tensors

$$\boldsymbol{V_i} \approx -\mu_0^t \boldsymbol{F}_i - \sum_{j \neq i}^{N} (1 + \frac{a^2}{6} \Delta_i)(1 + \frac{a^2}{6} \Delta_j) \boldsymbol{T}^0(\boldsymbol{R}_i - \boldsymbol{R}_j) \cdot \boldsymbol{F_j} \tag{1.12}$$

$$\approx -\mu_0^t \boldsymbol{F}_i - \sum_{j \neq i}^{N} (1 + \frac{a^2}{3} \Delta_x) \boldsymbol{T}^0(\boldsymbol{x} = \boldsymbol{R}_{ij}) \cdot \boldsymbol{F}_j \tag{1.13}$$

because $\Delta_i \Delta_j T^0(\boldsymbol{R}_{ij}) = 0$.

Evaluating the Laplacians and rearranging the explicit formulae for mobility tensors can be obtained

$$\boldsymbol{\mu}_{ii}^{tt,RP} = \mu_0^t \boldsymbol{1} \tag{1.14}$$

$$\boldsymbol{\mu}_{ij}^{tt,RP} = \mu_0^t \left(\frac{3}{4} \left(\frac{a}{R} \right) \left(\boldsymbol{1} + \hat{\boldsymbol{R}} \hat{\boldsymbol{R}} \right) + \frac{1}{2} \left(\frac{a}{R} \right)^3 \left(\boldsymbol{1} - 3 \hat{\boldsymbol{R}} \hat{\boldsymbol{R}} \right) \right) \quad \text{for} \quad i \neq j. \tag{1.15}$$

Completely analogous procedure can be applied to the tr and rr parts of the mobility matrix. That and further improvements (such as differentiable continuation for overlaps) are discussed by Zuk, Cichocki, and Szymczak [58].

1.3 Mathematical treatment of Brownian motion

I would like to bring to the attention of the reader the central complication of stochastic calculus and, more importantly, underscore why it is needed in the first place. A more succinct description of this domain can also be found in Waszkiewicz et al. [1].

For simplicity we will again focus on just a single colloidal sphere. Following the notation of Öttinger [41], time dependent velocity V_t of such sphere can be described by the Langevin equation

$$M\frac{\mathrm{d}V_t}{\mathrm{d}t} = -\zeta V_t + F_t^B \tag{1.16}$$

with F_t^B a time dependent force arising from the bombardment of the colloid by the water molecules.

Solving equation (1.16) naively we obtain V_t as a convolution of F_t^B with an appropriate Green's function. Such convolution is a linear operator acting on (what we hope is) a Gaussian process and thus we should be able to compute variance of the velocity by means of a double integral

$$\langle V_t^2\rangle = \frac{1}{M^2}\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \exp\left(-\zeta(2t-t'-t'')/M\right)\langle F_{t'}^B F_{t''}^B\rangle. \tag{1.17}$$

Recall that the brownian timescale is much longer than the fluid relaxation timescale $\tau_B \gg \tau_s$ and thus we postulate that the Browinan force has a singular correlation structure

$$\langle F_{t'}^B F_{t''}^B\rangle = \alpha_B \delta(t'-t'') \tag{1.18}$$

Evaluating the integral (1.17) we get

$$\frac{1}{2}M\langle V_t^2\rangle = \frac{\alpha_B}{4\zeta}(1-\exp(-2\zeta t/M)). \tag{1.19}$$

By applying equipartition principle to the result (1.19) we arrive at the necessary amplitude of the Brownian fluctuations as

$$\alpha_B = 2k_B T\zeta. \tag{1.20}$$

This is a form of fluctuation-dissipation theorem [54, 41].

Furthermore we can heuristically go to the limit $M/\zeta \to 0$ and obtain equation for the particles position

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \frac{1}{\zeta}F_t^B. \tag{1.21}$$

A proper treatment of the equation (**??**) is possible with the help of stochastic differential equations (SDE). First we define standard Wiener process as a Gaussian martinagle with the following covariance structure

$$\langle W_{t_1} W_{t_2}\rangle = \int_0^{t_1}\mathrm{d}t' \int_0^{t_2}\mathrm{d}t''\delta(t'-t'') = \min(t_1,t_2) \tag{1.22}$$

Note that this time covariance is nonsingular and usual theory of Gaussian processes is directly applicable. Heuristically we expect V_t to be described by the following integral

$$V_t = \frac{\sqrt{2k_bT\zeta}}{M}\int_0^t \exp\left(-\zeta(t-t')/M\right)\mathrm{d}W_{t'}. \tag{1.23}$$

Unfortunately this integral cannot be performed pathwise because W_t has infinite variation in every interval, even though the issues of almost-surely nowhere differetniability of equation (1.21) are avoided. We need some generalization of the usual integration to formalise this notion.

These integrals turn out to be formally tractable for a class of processes call non-anticipatory. Any such process (X_t, say) has the property, that for any time t the future increments of the Wiener process $W_{t'} - W_t$ and past values of the process $X_{t''}$ ($t'' < t < t'$) are independent variables.

We begin construction of the stochastic integral by considering random step functions – processes which are constant on finite intervals with step heights given by random variables $\widetilde{X}_i$ like so

$$X_t = \sum_{j=1}^{n}\widetilde{X}_{j-1}\mathbb{I}(t\in[t_{j-1},t_j]). \tag{1.24}$$

For these processes the stochastic (Ito) integral is simply defined as

$$\int_0^{t_{\max}} X_t \mathrm{d}W_t = \sum_{j=1}^{n} \widetilde{X}_{j-1}(W_{t_j} - W_{t_{j-1}}) \tag{1.25}$$

there is a noticeable lack of symmetry in this expression – we evaluate the integrand on the left end of each interval, as a result $\widetilde{X}_{j-1}$ and $W_{t_j} - W_{t_{j-1}}$ are independent variables because of the non-anticipatory nature of the X_t process. This 'direction of anticipation' assymetry plays a central role in the differences between Ito and classical calculus (and is notably absent in Stratonovich calculus).

Thanks to the non-anticipation property we can immediately conclude two very useful lemmas for random step functions (which are also true for general non-anticipating processes). First, Ito integrals are martingales

$$\left\langle \int_0^{t_{\max}} X_t \mathrm{d}W_t \right\rangle = 0. \tag{1.26}$$

Second, variances of Ito integrals can be computed with standard (non stochastic) integrals

$$\left\langle \left(\int_0^{t_{\max}} X_t \mathrm{d}W_t \right)^2 \right\rangle = \int_0^{t_{\max}} \left\langle X_t^2 \right\rangle \mathrm{d}t, \tag{1.27}$$

a result dubbed Ito's lemma [31]. To complete the construction of Ito's integral we need an appropriate limiting procedure whereby a sequence of approximating random-step-function processes is constructed and the original integral is the limit of approximating integrals. Turns out that the correct notion of limit here is that of mean square error, and that both the approximating processes and the integral itself converge in that sense.

To show explicitly that this notion of integration is really fundamentally distinct from the usual integration consider a famous integral $\int_0^t W_{t'} \mathrm{d}W_{t'}$. We can easily construct sequence of approximating step functions of W_t by uniform discretization of a given interval with mesh approaching to zero. By straight forward calculation we obtain a surprising result

$$\int_0^t W_{t'} \mathrm{d}W_{t'} = \frac{1}{2}\left(W_t^2 - t\right). \tag{1.28}$$

The additional term $-\frac{1}{2}t$ does not have a classical counterpart!

Fundamental advantage of using Ito's calculus over Langevin's heuristics (appart from being formally sound and the lemmas (1.26) and (1.27)) are the transformation rules of Ito's formula – the stochastic counterpart to chain rule.

Suppose that $\mathrm{d}X_t = A_t \mathrm{d}t + B_t \mathrm{d}W_t$ in the weak sense of Ito integral and $Y_t = f(X_t, t)$. Then we know that

$$\mathrm{d}Y_t = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} A_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} B_t^2 \right) \mathrm{d}t + \frac{\partial f}{\partial x} B_t \mathrm{d}W. \tag{1.29}$$

This equation allows for solution of the previous integral (1.28)

$$\mathrm{d}(W_t)^2 = 2W_t \mathrm{d}W_t + \mathrm{d}t, \tag{1.30}$$

but more importantly for the physicists it gives change of coordinates rules which are vital when trying to take advantage of symmetries of studied systems.

Finally we can see that integral (1.28) is not only of academic interest. Equation (1.30) shows that this is exactly the behaviour of the square displacements which play a central role in determination of the diffusion coefficient.

1.4 The hydrodynamic radius

Experiments such as Analytical Ultracentrifugation (AUC) observe macroscopic changes in the concentration field ϕ which evolves due to sedimentation and diffusion forces. Disregarding for now the sedimentation component of the Lamm equation discussed in section 1.6.1 concentration field evolved according to the Fick's equation with macroscopic diffusion coefficient D

$$\frac{\partial}{\partial t}\phi = \nabla \cdot (D \nabla \phi). \tag{1.31}$$

In case of dilute suspensions macroscopic D can be indentified with the individual diffusion coefficient of a single colloid. For a suspension of microscopic solid spheres of size a the celebrated Stokes-Einstein relation gives D in terms of viscosity and temperature

$$D = \frac{k_B T}{6\pi\eta a}. \tag{1.32}$$

Ceteris paribus, the relationship (1.32) captures all relevant properties of the buffer. We can invert it to define hydrodynamic radius R_h for an arbitrary colloid as

$$R_h := \frac{k_B T}{6\pi\eta D}. \tag{1.33}$$

R_h is the size of a microscopic solid sphere with the same diffusion coefficient as the studied macromolecule. Since R_h is derived from the apparent diffusion coefficient D it is influenced by surface effects such as hydration layers. On the other hand, in simple cases, R_h can be treated as a property of the molecule alone disregarding colloid-solvent cohesion and thus computed in a more convenient way.

For a bead model of a rigid macromolecule R_h can be computed from the trace of the grand translational mobility matrix computed at the diffusion centre of the molecule TODO-CITE-CICHOCKI. If another reference point is chosen rotational motion contributes to the instantaneous diffusion coefficient leading an overestimate of long term diffusion coefficient. Note that diffusion centre does not necessarily coincide with centre of mass of the molecule.

For a rigid macromolecule modelled as a conglomerate of spherical beads we can derive hydrodynamic radius from its grand mobility matrix by considering collective translational and rotational motion of constituent beads around any given point. Forces acting on individual beads can be computed by first deriving grand friction matrix from grand mobility matrix then imposing rigid body motion and then computing total force and torque required for such motion thus obtaining conglomerate friction matrix. This conglomerate friction matrix is then inverted to obtain conglomerate mobility matrix which is then used to find the diffusion centre and conglomerate mobility matrix centred at that point as described in detail in [21].

An alternative method of computation of R_h is based on a heuristic observation that the trace of the Oseen tensor satisfies the Laplace equation and Monte Carlo methods of solution of Laplece's equation are then applied to solve a heat equation with constant temperature difference between the surface of the molecule and the ambient fluid. Heat flux is then used to estimate effective size of such molecule, as described by the authors of the Zeno package [32] which we used in Waszkiewicz et al. [5].

For elastic macromolecule there are more bounded degrees of freedom and not just the rotation but also deformation can affect instantaneous diffusion coefficient and a judicious choice of tracking point can remove much of the spurious overestimation. An intuitive first guess is to track 'the middle' or geometric average of molecules constituent parts, slightly more refined approach is to use centre of mass – this is clearly not an optimal strategy as shown by the exact result for the rigid molecule. Location, or more specifically the weights in the weighted average, have to be derived from the hydrodynamic properties rather than from, hydrodynamically irrelevant, mass.

Treating translational mobility μ^{tt} as a tensor (with indices space, space, bead, bead) we can definte an inverse matrix of per-particle ensemble-average traces of mobility b as

$$b_{ij}(\langle \mu^{tt} \rangle)_{jkll} = \delta_{ik} \tag{1.34}$$

where $\langle . \rangle$ denotes ensemble average.

Then the hydrodynamic radius is estimated by

$$R_h \approx \frac{1}{2\pi\eta} \left(\sum_i \sum_j b_{ij} \right)^{-1}. \tag{1.35}$$

We called this approach the minimum dissipation approximation (MDA) as explained in Waszkiewicz et al. [4]. A complete derivation of the MDA method can be found in Cichocki et al. [21].

Clearly the MDA method (and earlier, simpler Kirkwood-Reismann method TODO-CITE-KR) requrire estimating ensemble averaged translational-grand-mobility matrix. Given approximations (1.15) we require only the relative positions of the constituent elements of the molecule and their effective sizes. These should be drawn from the equilibrium distribution given by the Boltzmanian

$$\mathrm{d}p \propto \exp\left(-\frac{U}{k_B T}\right) \mathrm{d}V \tag{1.36}$$

there are two potential difficulties with computing this measure: determining the potential energy U and determining the volume element dV. We discuss both of them in greater detail in Waszkiewicz and Lisicki [3]. Both of these problems are more pronounced when very stiff springs are used as models of bonds leading to superficially paradoxical results such as the 'trimer paradox' discussed therein.

1.5 Elastic macromolecules

Soft matter as a discipline of physics evolved out of 'colloid susupensions science' under the influence of two 'fuel sources' – bio-relevant measurements showing immediate applicability and a chase of observations of 'universal' validity.

A good example of intersection of these two currents are models of long polymeric chains – first elastic macromolecules to be successfully modelled. These include some synthetic plastics but also biomolecules such as DNA or denatured proteins. A celebrated result of Rouse $R_h \sim N^{1/2}$ for a Gaussian chain[46] later improved by Zimm by inclusion of excluded volume interactions [57] $R_h \sim N^{\gamma}$ with $\gamma = 0.588$. apply universally to all polymers of sufficiently long chains (or short enough persistence lengths). For a more detailed overview the subject matter see chapter 3. of Hermann and Gompper [28].

1.6 Experimental techniques

The following sections describe theoretical underpinnings of the experimental methods relevant to the publications included in this thesis.

1.6.1 Analytical Ultracentrifugation

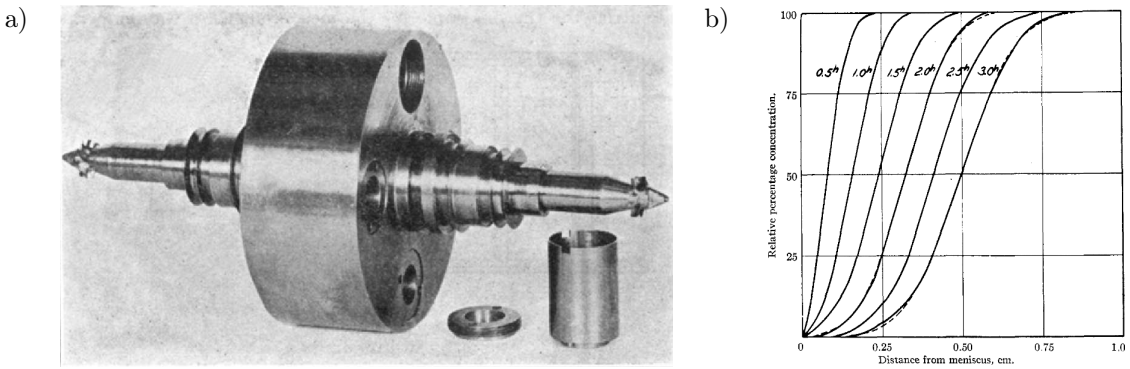

Figure 1.1: a) Rotor from an ultracentrifuge. b) Haemoglobin concentration profiles from an early AUC experiments with rotor spinning at 42'000rpm or 104'000 G. Profiles derived from photographs taken at 30 minute intervals are compared with Lamm equation predictions (dashed). Reprinted with permission from [49]. Copyright 1927 American Chemical Society.

Analytical Ultracentrifugation (AUC) is one of the oldest techniques in the domain of colloidal science recieveing a Nobel prize in 1926. In AUC a colloidal suspension is put inside a rapidly rotating centrifuge to increase the sedimentation rate of the molecules (current centrifuges spin with G-forces as large as 10'000G). Inside the centrifuge rotor a small window allows for optical measurements of either absorption or transmission (in recent models involving multiple wavelengths) which change as a result of concentration variation.

Changes of concentration inside the container are governed by the Lamm equation containing the divergence of two currents

$$\frac{\partial \phi}{\partial c} = \nabla \cdot \left(\underbrace{D \nabla \phi}_{\text{diffusion}} + \underbrace{s\omega^2 \mathbf{R} \phi}_{\text{sedimentation}} \right). \tag{1.37}$$

Derivation of the s and D values from experimental, time-dependent concentration profiles is done via a fitting procedure which requires efficient solution schemes for the PDE (1.37)[24, 20]. Finite element method combined with nonlinear least-squares techniques and Monte-Carlo based error estimation allows

gives both values and uncertaninites of the s and D values which can be used to determine other molecular parameters such as R_h or molecular mass.

1.6.2 Fluorescence Correlation Spectroscopy

The overview of the Fluorescence Correlation Spectroscopy (FCS) technique follows that of Tompson [51] and Gregor and Enderlein [27]. The FCS method, introduced about 50 years after invention of AUC, relies on temporal analysis of fluorescence signal to determine properties of the studied sample. These include the hydrodynamic size (relevant to this work), but it is also possible to determine adsoprtion or reaction kinetics using this method.

In case of diffusion constant measurements a very dilute sample of studied molecule (typically marked with an added fluorophore) is illuminated by a laser inside a confocal microscope. Excited fluorophore then emits light back through the microscope but at a slightly shorter wavelength (due to Stokes shift) which reaches the detector shielded from laser light with a dichroic mirror (cf 1.2).

Recorded fluorescence signal $F(t)$ is time dependent since the number of molecules inside the laser beam changes in time due to diffusion. FCS experiments perform best when at each moment expected number of excited molecules is close to one giving largest relative fluctuations of the fluorescence signal.

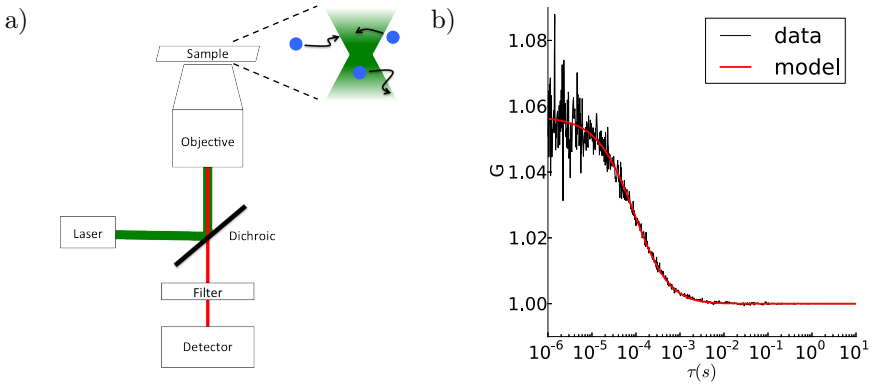

Figure 1.2: a) A typical FCS setup diagram b) Fluorescence correlation spectroscopy (FCS) data fit with a model for 3d diffusion. By Mllyjn CC BY-SA 3.0 TODO-REPLACE-WITH-AN-DATA-FROM-ARXIV

The frequency of the fluctuations due to diffusion can be quantified (and thus used to derive the diffusion coefficient) by considering signal autocorrelation function $G(\tau)$ given by

$$G(\tau) = \langle F(t)F(t-\tau)\rangle \tag{1.38}$$

with $\langle . \rangle$ denoting time average. This function is composed of two terms – constant background (in case of non-interacting molecules) term due to photos coming from two different molecules and a delay dependent term quantifying probability of detecting a photon from the same molecule again. For simplicity of this overview let's focus only on the last term.

Suppose for simplicity that a postition-dependent probability of excitation describing the shape of the excitation volume can be described by a axisymmetric Gaussian profile $U(\mathbf{r})$ given by

$$U(\mathbf{r}) = \kappa \exp\left(-\frac{2}{a^2}\left(x^2+y^2\right) - \frac{2}{b^2}z^2\right) \tag{1.39}$$

with $\mathbf{r} = [x, y, z]$ and κ is some overall constant.

If we now consider the Green's function $g(\rho, \tau)$ of the diffusion problem given by

$$g(\rho,\tau) = \frac{1}{(4\pi D\tau)^{3/2}} \exp\left(-\frac{|\rho|^2}{4D\tau}\right) \tag{1.40}$$

we can express the autocorrelation function according to

$$G(\tau) = \int\int U(\mathbf{r}+\rho)g(\rho,\tau)U(\mathbf{r})\mathrm{d}\mathbf{r}\mathrm{d}\rho \tag{1.41}$$

$$= \frac{\pi^{3/2}}{8}\frac{a^2 b}{(1+4D\tau/a^2)\sqrt{1+4D\tau/b^2}} \tag{1.42}$$

since the parameters a and b cannot be known a priori one typically fits a simplified expression

$$G(\tau) = G(0)\left(\left(1+\frac{\tau}{\tau_D}\right)\left(1+\gamma\frac{\tau}{\tau_D}\right)^{1/2}\right)^{-1} + G(\infty) \tag{1.43}$$

where γ quantifies the aspect ratio of the excitation volume and τ_D characteristic time of the diffusion. We can obtain diffusion coefficient D from the residence time by comparing a reference sample of known D and computing a ratio.

1.6.3 Small Angle X-ray Scattering

This section follows notation of [29]. Another experimental technique is based on the analysis of the X-ray scattering on the colloidal particles as a function of the beams deflection vector $\mathbf{q}$ defined as a difference between incoming wave vector $\mathbf{k}_i$ and scattered wave vector $\mathbf{k}_s$.

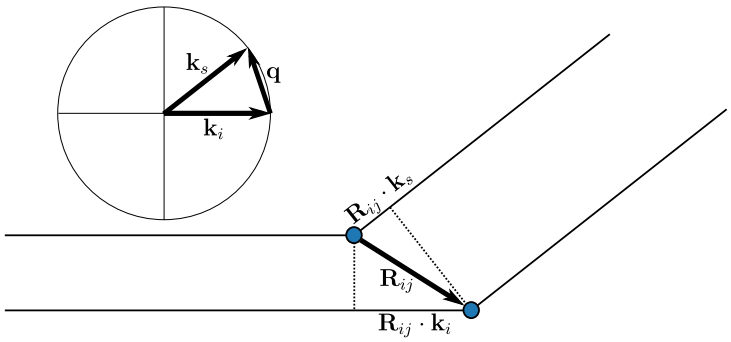

Figure 1.3: Phase shifts in Born approximation.

For colloidal particles the Born approximation of single-event scattering (as show in fig. 1.3) is a sufficient description to model experimental data. Within this approximation the scattering pattern is computed as an interference pattern of the scattered signals from different scattering sites due to phase shift $\Delta\Phi$ resulting from difference in optical path length.

$$\Delta\Phi = \mathbf{R}_{ij}\cdot(\mathbf{k}_s - \mathbf{k}_i) = \mathbf{R}_{ij}\cdot\mathbf{q} \tag{1.44}$$

We can express the complex amplitude of the scattered wave $A(\mathbf{q})$ as a Fourier transform of the scattering intensity distribution $\rho(\mathbf{r})$.

$$A(q) = A_0\int\rho(\mathbf{r})\exp(i\mathbf{q}\cdot\mathbf{r})\mathrm{d}\mathbf{r} \tag{1.45}$$

Since the observed intensity on the screen I is proportional to the squared absolute value of the wave amplitude $I(\mathbf{q}) \propto |A(\mathbf{q})|^2$ we can express $I(\mathbf{q})$ as a double integral

$$|A(\mathbf{q})|^2 = AA^* \tag{1.46}$$

$$= A_0^2\int\int\rho^*(\mathbf{r}')\rho(\mathbf{r}'')\exp(i\mathbf{q}\cdot(\mathbf{r}'-\mathbf{r}''))\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}''. \tag{1.47}$$

If we assume that the scattering sites are point-like we can express the scattering density distribution as a sum of Dirac distributions δ centred at locations of scattering sites $\mathbf{s}_i$

$$\rho(\mathbf{r}) = \sum_i \rho_i\delta(\mathbf{r}-\mathbf{s}_i) \tag{1.48}$$

where ρ_i gives scattering intensity on site i. Changing the integral (1.47) to a double sum

$$I(\mathbf{Q}) = A_0^2 \sum_i \sum_j \rho_i \rho_j \exp(i\mathbf{q} \cdot (\mathbf{s}_i - \mathbf{s}_j)) \tag{1.49}$$

Inside the colloid the distribution of orientations of the macromolecules with respect to the lab frame is invariant under action of rotations. For a vector $\mathbf{v} = \mathbf{s}_i - \mathbf{s}_j$ this simply implies $\mathbf{v}$ distributed uniformly on a sphere and its projection onto a vector $\mathbf{q}$ distributed uniformly on an interval according to orange slicing theorem[1] giving

$$\begin{aligned}
\langle I(\mathbf{Q}) \rangle &= A_0 \sum_i \sum_j \rho_i \rho_j \int \exp(iQs) \mathbb{I}(|s| < |\mathbf{s}_i - \mathbf{s}_j|) \mathrm{d}s & (1.50)\\
&= A_0 \sum_i \sum_j \rho_i \rho_j \operatorname{sinc}(|Q||\mathbf{s}_i - \mathbf{s}_j|) & (1.51)\\
&= A_0 \sum_i \sum_j \rho_i \rho_j \operatorname{sinc}(|Q||\mathbf{R}_{ij}|) & (1.52)
\end{aligned}$$

Thus the prediction of the scattering intensity of in a SAXS measurement of a colloidal suspension reduces to the computation of the distance matrix R_{ij} between scattering sites given the values of scattering intensities $\rho_i(q)$. In case of elastic macromolecules the scattering signal is simply averaged over an ensemble of possible molecular conformations. With the help of tabulated values of scattering intensities of each aminoacid obtained by Tong, Yang, and Lu [52] a python package `saxs_single_bead` was developed.

[1] the orange slicing theorem: if you slice an orange into slices of equal thickness each slice has the same share of the orange peel

Discussion

2.1 Experimental and theoretical challenges

Much of theoretical enthusiasm in developing complex models, with multitude of interactions is hampered by insufficient quantity or quality of experimental calibration data (necessarily required for determination of molecular 'material constants'). In the polymer-science domains where data is plentiful (such as the area of globular proteins which can be readily measured by for example x-ray crystallography techniques) significant numerical progress could, and was made using machine learning techniques leading to a celebrated example of AlphaFold [33] reaching even pop-science recognisability.

Progress in the domain of elastic molecules is much more modest, even though they constitute a large proportion of proteins[55] and play an important role in multiple bio-relevant processes, such as the functioning of the COVID-19 virus[47] (to name just one example particularly resonant in 2023). In the opinion of the author this progress is (at least in part) determined by the amount of readily available experimental data – compare over 200 thousand conformations available in the Protein Data Bank[17] with not even 100 high quality measurements of the diffusion coefficients of the IDPs we were able to extract from published works[4] or 118 proteins in the Small Angle Scattering Database[18], or 462 proteins in the Protein Ensemble Database[19].

For the author the experimental challenges of studying biomacromolecules have been an eye opening lesson in patience. The excellent work on biosynthesis of the groups at Bayor College of medicine and Insitute of Physics of Polish Academy of Science cannot be understated. These two collaborations give an overview of many experimental challenges: failure or success of initial biosynthesis, obtaining correct concentrations, through purification from by products of biosynthesis and stability of the resultant molecular constructs to practical obstacles such as replacing discontinued lab equipment and cross border shipping.

Even if everything lines up and one obtains a good sample of the molecule to study the measurement itself is no easy task either. Both AUC and FCS methods rely on time-series analysis of an optical signal - in that sense they are both indirect methods requiring model fitting within the experimental procedure. This complicates analysis of direct measurement error (uncertainty of a single measurement), with current Monte-Carlo based AUC analysis method providing no uncertainty estimates for samples of high purity. Single measurement error, however, is not the only and not even a leading source of uncertainty in the measurements of the diffusion (and sedimentaiton) properties of the molecules. Since buffer conditions, concentration and even time from synthesis[39] affect the final outcome these have to be incorporated into error analysis to arrive at compound value of confidence interval of the measurements. Some authors, unfortunately, do not provide any error estimates[42, 35] and most of them do not discuss multiple sources of error.

Variety of bio-chemical insights into (both folded and disordered) protein behaviour invites phenomenological models to use many covariates in explaining hydrodynamic size of these molecules. A specific example of one such insight could be the relative rigidity of polyhistidine fragments and tags, their presence (or more simply share of histidines in the totality of aminoacids forming a molecule) is sometimes used as a explanatory variable in a process of phenomenological modelling [50, 38]. As the data concerning hydrodynamic properties of IDPs is scarce and direct intervention in the protein sequence can be prohibitively expensive since synthesis of each new protein construct can take months we arrive at an issue more familiar to the social scientists than physicists - available information forms essentially an observation study where correlation is difficult to separate from causation[1]. Even if we perform all of the statistical analyses correctly and we observe a statistically significant correlation between, for

[1]This is in contrast to a typical physics experiment, akin to a randomized-controlled trial in social setting, where experimenter is free to chose value of controll variable and measure the outcome. In the example study of his-tag influence we can't simply, randomly separate proteins into two groups and add/remove his-tags as required.

example, presence of a his-tag and increase in hydrodynamic size how can we be sure that it is not simply the case of a confounding cause? One can easily imagine a scenario where compact proteins containing his-tag are simply harder to synthesise and thus we form our phenomenological conclusion based on a distorted sample (selection bias). Correlations estimated this way can even have opposite sign to the real causal effect (which is what we're really trying to estimate)! A comprehensive (and approachable) overview of causal vs correlational discrepancies can be fond in Cinelli, Forney, and Pearl [22]. Until and unless we arrive at a very large dataset of hydrodynamic properties of (a bias free, representative subset of) IDPs first principle theoretical models should take precendence over phenomenological models (not just because of their precision as show in Waszkiewicz et al. [4]), but because they correctly model causal dependency between the covariates of interest and hydrodynamic size. These models come in two flavours – atomistic and coarse grained.

Atomistic methods of can sometimes be used[34], but they require either simulating the surrounding water molecules explicitly, which is very computationally intensive, or an implicit solvent scheme. Both approaches pose significant numerical challenges [26]. Even if simulations are in principle feasible, obtaining 10-100 millisecond long trajectories with this method which would enable the direct computation of the long-time diffusion coefficient are prohibitively expensive.

Coarse grained methods employ larger units (in case of proteins typically amino acid residues) as building blocks for the structure prediction scheme. These are combined with a separate hydrodynamic model to predict diffusion coefficient.

2.2 Approaches to Predicting Diffusion Coefficients in the Hot and Cold Limits

As outlined, tackling the problem directly, even within a coarse-grained perspective, still presents numerical challenges. Moreover, identifying a minimal model capable of reproducing experimentally observed variations in diffusion coefficients is inherently valuable. This capability provides a direct interpretation of the observed variability through a restricted set of molecular mechanisms.

In the context of molecular dynamics, a convenient method for expressing dimensionless numbers is through lengthscale ratios, which quantify the respective ranges of different interactions within the molecule.

In our case, the lengthscales of interest include the persistence length $P = \frac{EI}{k_B T}$ (capturing elastic forces proportional to rigidity EI in relation to Brownian forces proportional to fluctuation energy $k_B T$), the building block size (quantifying excluded volume interactions), and the Debye length (quantifying electrostatic interactions, scaling with the inverse square root of the ionic strength C_s of the buffer). Some values of these lengthscales, pertinent to the publications [5] and [4], are outlined in Table 2.1.

Description	**Scaling**	**IDP [Å]**	**DNA [Å]**
Length	L	2000	1000
Persistence length	$P = EI/k_B T$	3	500
Building block		4	3
Debye length	$R_D \sim (C_s)^{-1/2}$	1	1
Applicable limit		hot	cold(?)

Table 2.1: Relevant length scales for the two studied problems. Length row represents typical values. Debye length computed for relevant (physiological) buffer conditions. TODO: DEFINE SYMBOLS

Frist type of molecule investigated in this doctoral thesis are the DNA minicircles. Here the need for assessing these values for each problem cannot be understated – there is a factor 160 difference in stiffness of protein linkers and DNA filaments. Furthermore, the persistence length of DNA filaments is comparable to the length of the very short DNA minicircles under investigation by the group of Professor Lynn Zechiedrich at Baylor College of Medicine. These minicircles were explored using two experimental techniques: CryoEM[30] and analytical ultracentrifugation (AUC)[5].

CryoEM images provide insight into the conformational properties of the DNA minicircles, while AUC measurements offer information on the hydrodynamic properties of the molecule. To address our initial concern that the very large forces introduced by AUC methods (the molecules are sedimenting under 10,000G) could affect the equilibrium shape properties of the DNA molecules, we conducted a study on

the sedimentation of elastic filaments[6]. The relative importance of elastic and buckling forces arising from sedimentation is quantified by the elasto-hydrodynamic lengthscale, $8\pi^3 \frac{EI}{L^2 g\rho}$, which for DNA is approximately 200,000 Å or 60 kbp—200 times larger than the length of the minicircles of interest. It is noteworthy that plasmids of that length are of biological interest, and additional care needs to be taken when studying them with AUC methods.

Moreover, the visual similarity between CryoEM figures from Irobalieva et al. [30] and equilibrium configurations of twisted beams as described by Coleman and Swigon [23] suggested that thermal fluctuations might have only a small influence on the overall behavior of these molecules. Consequently, we chose to model this problem in the 'cold' limit (cf. Table 2.2), neglecting thermal fluctuations entirely and computing equilibrium configurations only. Subsequently, these configurations were treated as rigid bodies when determining the hydrodynamic radius, as detailed in Waszkiewicz et al. [5]. The influence of thermal fluctuations was investigated further (after the publication) using the `pychastic` package, as outlined in Section 2.4.

Manuscript	**Elasticity**	**Thermodynamics**	**Hydrodynamics**
DNA[5, 6]	✓	.	✓
IDP[4, 3]	.	✓	✓
Pychastic[1]	✓	✓	✓

Table 2.2: Three domains of interest for the study of elastic molecules. Very high and very low stiffness of DNA and IDPs respectively allowed us to use simplified models.

The second type of molecule investigated in this study was intrinsically disordered proteins (IDPs). These molecules exhibit much smaller stiffness, with an elastohydrodynamic length of approximately 2000 Å—comparable to the molecule length. This characteristic further justifies our initial concern regarding the potential influence of buckling forces arising in sedimentation on hydrodynamic sizes in some measurements. Fortunately, the group led by prof. Anna Niedzwiecka at the Institute of Physics, Polish Academy of Sciences, employs fluorescence correlation spectroscopy (FCS) instead of AUC, which does not introduce large force gradients on the molecule.

Similar to the DNA minicircle modeling, we compared lengthscales in the problem. Given that the persistence length is comparable to the building block size, we decided to focus on the 'hot' limit of the problem, where elastic forces are disregarded. More precisely bending forces arising from the Ramachandran angles distributions were neglected, at the same time the forces that govern bond length were taken to be infinitely strong, this approach requires care when handling bond-angle distributions and some apparently paradoxical results can arise there as discussed in Waszkiewicz and Lisicki [3]. We show that the intuitive distribution 'uniform on a sphere' indeed arises as a limit in the case of linear filaments (notably, a different distribution arises for molecules with loops as shown therein, with deviations in probability density that can be arbitrarily large).

Before estimating the diffusion coefficient of the IDPs, we aimed to assess the quality of the conformational ensembles resulting from our method, as described in more detail in Waszkiewicz et al. [4]. For IDPs, one effective method for probing conformational properties without resorting to hydrodynamics is small-angle X-ray scattering (SAXS). We compared inter-domain distance data published in Różycki and Boura [47] with data generated by the globule-linker model. Additionally, we compared complete SAXS curves for ataxin-3 (PDB code `1yzb`) published in [36] with satisfactory results. Further details regarding the `saxs_single_bead` package and the comparison are outlined in Section 2.4.

2.3 Combining Hot and Cold Approaches

The two experimentally inspired problems, even though solvable to a satisfactory degree within the presented approximations, leave a desire to assess the size of the error introduced by these particular simplifications. These error estimates can be compared either to the precision of the experimental data or to the estimates of other approximation errors in the model, such as approximations of the hydrodynamic mobility tensors. This can guide the effort in further improvements to the numerical method; for example, should we first work on including thermal fluctuations to the conformations or rather improve hydrodynamic mobility tensor approximations?

When mobility tensors are simply modelled with the Rotne-Prager approximation (such as the tensors of the `pygrpy` package), the errors introduced there are of the order of 2% (based on R8 test case of [58]).

Assessing errors introduced by approximate conformer generation, whether in the hot ($T \to \infty$) or cold ($T \to 0$) limit, necessitates simulating the ensemble at finite temperatures. Various methods can be employed for this purpose, but a simulation grounded in physical principles holds particular appeal. To achieve good performance, the Brownian Dynamics method was selected. This approach relies on formulating the dynamical equation governing conformational changes in the form of a SDE. Alternative formulations such as the Langevin equation or Fokker-Planck PDE are discussed in greater detail in Waszkiewicz et al. [1].

Surprisingly, in 2022, the authors found that the only high-quality SDE integration package available was `differentialequations.jl` in Julia[43]. Since Julia is still not very popular among physicists (or at least not inside the soft-matter community) this resulted in frequent re-implementation of SDE integration methods for each problem and often re-implementation of the hydrodynamic interaction tensors as well (possibly due to the fact that the Rotne-Prager tensors most frequently used in our group were originally implemented in Fortran).

We have tried to address this issue by making the packages `pychastic` (for SDE integration) and `pygrpy` (for rotne-prager tensors) as easily accessible as we can – single command installation, working on any reasonable linux distribution, MacOS and (as of 2023) even Windows10[7]. Additionally, we have invested substantial effort in creating straightforward yet realistic examples in the documentation. This not only simplifies the learning curve but also aims to encourage wider adoption of our integration package. For a comprehensive understanding of implementation and usage, please refer to the detailed information provided in Waszkiewicz et al. [1].

The specific implementation of `pychastic` further streamlines the development of similar models by eliminating the need for explicit force specification. Instead, users can now simply define the energy, and the forces can be derived programmatically using `jax.grad`. This feature proved particularly advantageous in simulating shear-relaxed dynamics of DNA minicircles, where the energy depends on the non-local quantity of writhe (Wr), as discussed in detail in section 2.4.

Answer to the correct development direction (in the opinion of the author) lies in future experimental data with resolutions capable of probing inaccuracies of the force-fields used or showing clear dependencies on buffer conditions such as temperature or ionic strength. Such measurements are in the making and early results give a glimpse into the future DNA models[44].

2.4 Results not Included in Submitted Manuscripts

2.4.1 Thermal Effects on the Shapes of DNA Minicircles

Utilizing the `pychastic` package, we conducted simulations to explore the shapes of DNA minicircles at finite (room) temperatures and computed their apparent diffusion coefficient using the minimum dissipation method described in Cichocki et al. [21].

In these simulations, we assumed that torsional stresses in the DNA are relaxed. Consequently, the total elastic energy E_{el} of such a minicircle is given by

$$E_{el} = \frac{1}{2}\int_0^L EI\kappa^2 + \left(\frac{2\pi(Lk - Wr)}{L}\right)^2 dl. \tag{2.1}$$

Both κ and Wr can be computed from the shape of the centerline alone (or a suitable discretization thereof), reducing the dimensionality of the problem compared to an approach where both the position and orientation of filament segments are simulated. It is worth noting that simulating rotational dynamics, even of a single body, is a complex task, as outlined in Waszkiewicz et al. [1].

Given that Wr is a nonlocal quantity, the computation of forces from E_{el} can be cumbersome. Fortunately, this task can be accomplished with the assistance of `jax.grad`, which is compatible with the SDE solver `pychastic`.

This gives further insight into the intermediate region ($1 < Lk < 2.5$) where minicircles have multiple stable configurations at absolute zero temperature discussed in Waszkiewicz et al. [5].

It turns out that the effects of thermal fluctuations on the overall hydrodynamic radius outside of the intermediate region are negligible. Moreover, the range of Lk values for which the equilibrium ensemble contains both open configurations and configurations with self-contact is very small—less than 0.2 turns. When considering the difference in total elastic energy between the flat circular configuration and the configuration with self-contact, one arrives at $\frac{\partial \Delta E}{\partial Lk} \approx Lk\frac{4\pi^2\cdot 2}{3}\frac{P}{L}k_BT \approx 26k_BT$, providing an estimate of the transitional region width of approximately 0.03. This contrasts starkly with CryoEM data[30], where

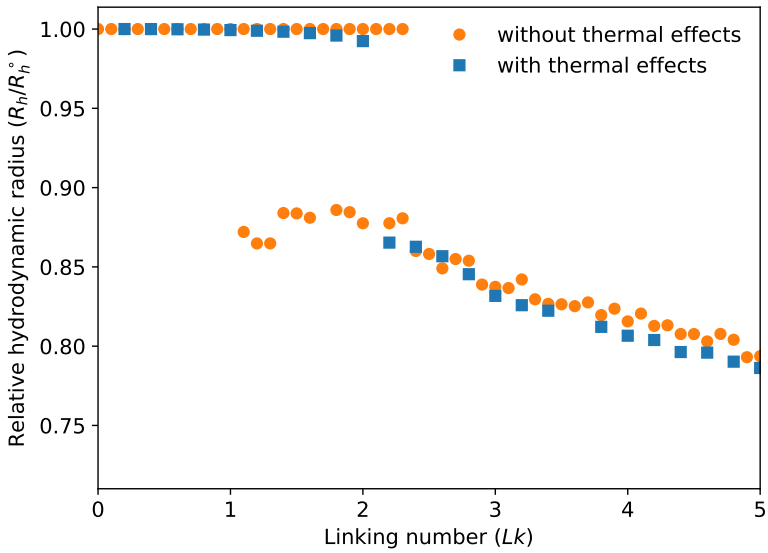

Figure 2.1: Relative hydrodynamic radius of the 336bp minicircle as a function of the linking number.

a multitude of conformations were observed for a wide range of Lk values. One possible explanation for this discrepancy lies in the temperature dependence of the geometric properties of DNA, as hinted at by the results of Ranasinghe et al. [44].

2.4.2 Comparing Conformations to SAXS Data

The author would like to thank dr. Bartosz Różycki for his guidance regarding experimental SAXS data.

As detailed in Waszkiewicz et al. [4], computing the diffusion coefficient for a given IDP requires generating samples from the equilibrium ensemble. To assess the quality of the conformer generation method independently of hydrodynamic modeling, we utilized available SAXS data. One example of such a comparison is illustrated in Figure 2.2, where experimental data from Sicorello et al. [48] is compared with our implementation of the globule-linker model (`sarw-spheres`) combined with our implementation of the one-site-per-aminoacid scattering model (`saxs-single-bead`). This model is based on form factors computed by Tong, Yang, and Lu [52].

The conformations used to predict the SAXS curve in Figure 2.2 were generated using the globule-linker engine, with the globule replaced by crystallographically obtained data retrieved from the Protein Data Bank (PDB)[17]. The excellent agreement in the range of scattering vectors corresponding to features larger than the C_α distance (small values of q) is a result of the combination of adequate modelling of both the rigid and flexible parts.

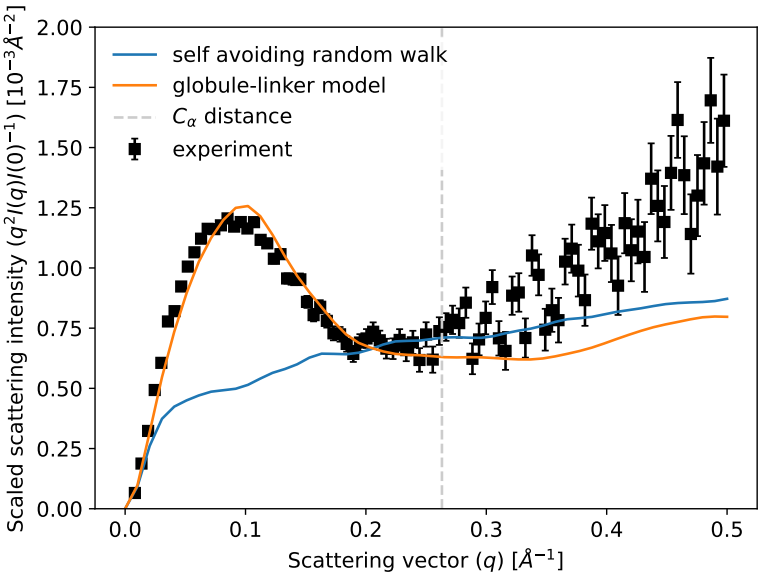

Figure 2.2: Kratky plot generated using `saxs_single_bead` package for ataxin-3 (pdb id `1yzb`) compared with experimental data from Sicorello et al. [48].

2.5 Author's contributions

The following dissertation consists of 6 thematically related manuscripts of which first four have been published at the time of writing and two are under review. Author's contribution to each of them is outlined below.

- Radost Waszkiewicz, Piotr Szymczak, and Maciej Lisicki. "Stability of sedimenting flexible loops". In: *Journal of Fluid Mechanics* 919 (2021), A14. DOI: 10.1017/jfm.2021.350.
  Performing theoretical calculations numerical investigation and visualisations. In particular: performing linear stability analysis of the PDE describing the shape of the flexible filament including the projection method which allows for near analytical treatment of the eigenvalue problem, analysis of the resulting matrix problem, preparing the numerical solver based on truncated Fourier series, performing numerical investigation and trajectory post-processing, preparation of 2d and 3d visualisations.

- Radost Waszkiewicz and Maciej Lisicki. "Hydrodynamic effects in the capture of rod-like molecules by a nanopore". In: *Journal of Physics: Condensed Matter* 33.10 (2021), p. 104005. DOI: 10.1088/1361-648X/abd11b.
  Performing numerical investigation and visualisations. Performing scaling analysis in collaboration with Maciej Lisicki. In particular: The Mathematica implementation of the near-wall corrections to the mobility tensor[37], computation of the trajectories of the colloid.

- Radost Waszkiewicz, Maciej Bartczak, Kamil Kolasa, and Maciej Lisicki. "Pychastic: Precise Brownian dynamics using Taylor-Itō integrators in Python". In: *SciPost Phys. Codebases* (2023), p. 11. DOI: 10.21468/SciPostPhysCodeb.11.
  Leading a small programming team (the author, Maciej Bartczak, Kamil Kolasa) to produce a complex python package (including documentation, examples and tests). Designing illustrative examples (in collaboration with Maciej Lisicki), test cases and debug tools which allowed for corrections of multiple typos in the available literature.

- Radost Waszkiewicz, Maduni Ranasinghe, Jonathan M Fogg, Daniel J Catanese Jr, Maria L Ekiel-Jeżewska, Maciej Lisicki, Borries Demeler, Lynn Zechiedrich, and Piotr Szymczak. "DNA supercoiling-induced shapes alter minicircle hydrodynamic properties". In: *Nucleic Acids Research* 51.8 (2023), pp. 4027–4042. DOI: 10.1093/nar/gkad183.
  Designing numerical pipeline and performing computations leading to the equilibrium configurations of the supercoiled DNA minicircles. Computing hydrodynamic properties of these configu-

rations using GRPY, pygrpy and Zeno software (only Zeno is included in the final manuscript). Producing graphs and visualisations included in the manuscript.

- Radost Waszkiewicz, Agnieszka Michaś, Michał K. Białobrzewski, Barbara Klepka, Maja Cieplak-Rotowska, Zuzanna Staszałek, Bogdan Cichocki, Maciej Lisicki, Piotr Szymczak, and Anna Niedźwiecka. "Minimum dissipation approximation: A fast algorithm for the prediction of diffusive properties of intrinsically disordered proteins". In: *Journal of Physical Chemisty Letters (in review)* (2024).
  Implementation of the numerical methods. In particular: design of the globule-linker conformer generation scheme, investigation of Flexible Meccano conformer generation scheme, implementation of the python port `pygrpy` of the generalized Rotne-Prager mobility tensors (originally implemented in Fortran[58]). Statistical analysis of the theory–experiment deviations.

- Radost Waszkiewicz and Maciej Lisicki. "The trimer paradox revisited". In: *(in review)* (2024).
  Performing theoretical calculations regarding equilibrium distributions of molecules with arbitrarily stiff bonds and numerical calculations using `pychastic` package to give illustrative examples of the discussed phenomena.

2.6 Software packages

Package	**Decription**	**Github**	**Docs**
`pychastic`	SDE solver	[12]	[11]
`pygrpy`	Rotne-Prager mobility tensors	[14]	[13]
`sarw-spheres`	Globule-linker conformer generator	[10]	
`saxs-single-bead`	One site per aminoacid SAXS engine	[8]	[16]
`pywrithe`	Computing writhe of a curve	[9]	[15]

Table 2.3: Software packages developed in the course of preparation of this doctoral thesis.

Paper: Stability of sedimenting flexible loops

`\includepdf` arguments: publications/waszkiewicz_2021_stability

Paper: Hydrodynamic effects in the capture of rod-like molecules by a nanopore

`\includepdf` arguments: publications/waszkiewicz_2021_hydrodynamic

Paper: DNA supercoiling-induced shapes alter minicircle hydrodynamic properties

`\includepdf` arguments: publications/waszkiewicz_2023_dna

Paper: Pychastic: Precise Brownian dynamics using Taylor-Itō integrators in Python

`\includepdf` arguments: publications/waszkiewicz_2023_pychastic

Paper: Minimum dissipation approximation: A fast algorithm for the prediction of diffusive properties of intrinsically disordered proteins

`\includepdf` arguments: publications/waszkiewicz_2024_mda

Paper: Trimer paradox revisited

Conclusions

The objectives of the present thesis were twofold - find theoretical approaches capable of modelling experimentally relevant macromolecules and provide a step towards a single, modular system capable of accommodating different coarse graining approximations and experimental techniques. The results of this thesis show that such unified approach is possible and predicts diffusion coefficients of macromolecules with good precision. The main achievements of this thesis can be summarised as follows:
TODO TODO

1. We have successfully predicted hydrodynamic radii of DNA loops at different values of linking number.

2. We have successfully predicted hydrodynamic radii of many intrinsically disordered proteins from the largest to-date benchmark set.

3. We have provided easy to use, well documented and publicly available Python implementations to all theoretically proposed methods (without compromising the prediction speed).

Thus the presented method can be used as a numerically feasible null-hypothesis model in future investigations by multiple experimental groups with large deviations from it's predictions being an indicator of new and exciting phenomena.

Bibliography

[1] Radost Waszkiewicz, Maciej Bartczak, Kamil Kolasa, and Maciej Lisicki. "Pychastic: Precise Brownian dynamics using Taylor-Itō integrators in Python". In: *SciPost Phys. Codebases* (2023), p. 11. DOI: 10.21468/SciPostPhysCodeb.11.

[2] Radost Waszkiewicz and Maciej Lisicki. "Hydrodynamic effects in the capture of rod-like molecules by a nanopore". In: *Journal of Physics: Condensed Matter* 33.10 (2021), p. 104005. DOI: 10.1088/1361-648X/abd11b.

[3] Radost Waszkiewicz and Maciej Lisicki. "The trimer paradox revisited". In: *(in review)* (2024).

[4] Radost Waszkiewicz, Agnieszka Michaś, Michał K. Białobrzewski, Barbara Klepka, Maja Cieplak-Rotowska, Zuzanna Staszałek, Bogdan Cichocki, Maciej Lisicki, Piotr Szymczak, and Anna Niedźwiecka. "Minimum dissipation approximation: A fast algorithm for the prediction of diffusive properties of intrinsically disordered proteins". In: *Journal of Physical Chemisty Letters (in review)* (2024).

[5] Radost Waszkiewicz, Maduni Ranasinghe, Jonathan M Fogg, Daniel J Catanese Jr, Maria L Ekiel-Jeżewska, Maciej Lisicki, Borries Demeler, Lynn Zechiedrich, and Piotr Szymczak. "DNA supercoiling-induced shapes alter minicircle hydrodynamic properties". In: *Nucleic Acids Research* 51.8 (2023), pp. 4027–4042. DOI: 10.1093/nar/gkad183.

[6] Radost Waszkiewicz, Piotr Szymczak, and Maciej Lisicki. "Stability of sedimenting flexible loops". In: *Journal of Fluid Mechanics* 919 (2021), A14. DOI: 10.1017/jfm.2021.350.

[7] *A community supported Windows build for jax.* https://github.com/cloudhan/jax-windows-builder. Accessed: 2023-11-02.

[8] *All Python package to compute small angle X-ray scattering (SAXS) profiles in one-bead-per-residue approximation with numpy.* https://github.com/RadostW/saxs_single_bead. Accessed: 2023-11-02.

[9] *Compute writhe of closed curve with numpy and jax.numpy.* https://github.com/RadostW/PyWrithe. Accessed: 2023-11-02.

[10] *Genrate self avoiding random walks (SARW) for spheres of given sizes.* https://github.com/RadostW/sarw_spheres. Accessed: 2023-11-02.

[11] *Pychastic documentation.* https://pychastic.readthedocs.io/en/latest/. Accessed: 2023-11-02.

[12] *Pychastic is a stochastic differential equations integrator written entirely in python.* https://github.com/radostw/stochastic. Accessed: 2023-11-02.

[13] *PyGRPY documentation.* https://pygrpy.readthedocs.io/en/latest/. Accessed: 2023-11-02.

[14] *Python port of Generalized Rotne Prager Yamakawa hydrodynamic tensors.* https://github.com/RadostW/PyGRPY/. Accessed: 2023-11-02.

[15] *PyWrithe documentation.* https://pywrithe.readthedocs.io/en/latest/. Accessed: 2023-11-02.

[16] *saxs-single-bead documentation.* https://saxs-single-bead.readthedocs.io/en/latest/. Accessed: 2023-11-02.

[17] *Protein Data Bank.* https://www.rcsb.org/. Accessed: 2023-11-02.

[18] *Small Angle Scattering Biological Data Bank.* https://www.sasbdb.org/search/. Accessed: 2023-11-02.

[19] *The Protein Ensemble Database*. https://proteinensemble.org/. Accessed: 2023-11-02.

[20] Weiming Cao and Borries Demeler. "Modeling Analytical Ultracentrifugation Experiments with an Adaptive Space-Time Finite Element Solution of the Lamm Equation". In: *Biophysical Journal* 89.3 (2005), pp. 1589–1602. ISSN: 00063495. DOI: 10.1529/biophysj.105.061135.

[21] Bogdan Cichocki, Marcin Rubin, Anna Niedzwiecka, and Piotr Szymczak. "Diffusion coefficients of elastic macromolecules". In: *J. Fluid Mech.* 878 (2019), R3. DOI: 10.1017/jfm.2019.652.

[22] Carlos Cinelli, Andrew Forney, and Judea Pearl. "A crash course in good and bad controls". In: *Sociological Methods & Research* (2022), p. 00491241221099552.

[23] Bernard D. Coleman and David Swigon. "Theory of Supercoiled Elastic Rings with Self-Contact and Its Application to DNA Plasmids". In: *Journal of Elasticity* 60.3 (2000), pp. 173–221. ISSN: 03743535. DOI: 10.1023/A:1010911113919.

[24] Borries Demeler and Gary E. Gorbet. "Analytical Ultracentrifugation Data Analysis with UltraScan-III". In: *Analytical Ultracentrifugation*. Ed. by Susumu Uchiyama, Fumio Arisaka, Walter F. Stafford, and Tom Laue. Tokyo: Springer Japan, 2016, pp. 119–143. ISBN: 978-4-431-55983-2. URL: http://link.springer.com/10.1007/978-4-431-55985-6_8.

[25] Daan Frenkel and Berend Smit. "Chapter 15 - Tackling Time-Scale Problems". In: *Understanding Molecular Simulation (Second Edition)*. Ed. by Daan Frenkel and Berend Smit. Second Edition. San Diego: Academic Press, 2002, pp. 409–429. ISBN: 978-0-12-267351-1. DOI: 10.1016/B978-012267351-1/50017-1.

[26] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. Vol. 1. Elsevier, 2001.

[27] I. Gregor and J. Enderlein. "Fluorescence Correlation Spectroscopy". In: *Soft Matter: From Synthetic to Biological Materials*. Ed. by Dhont J., G. Gompper, and G. Nagele. Schriften des Forschungszentrums Julich, 2008, A3.

[28] J. K. G. Hermann and G. Gompper. "Introduction". In: *Soft Matter: From Synthetic to Biological Materials*. Ed. by Dhont J., G. Gompper, and G. Nagele. Schriften des Forschungszentrums Julich, 2008, p. I.

[29] R. P. Hermann. "Scattering of Neutrons and Photons". In: *Soft Matter: From Synthetic to Biological Materials*. Ed. by Dhont J., G. Gompper, and G. Nagele. Schriften des Forschungszentrums Julich, 2008, A1.

[30] Rossitza N. Irobalieva et al. "Structural diversity of supercoiled DNA". In: *Nature Communications* 6.1 (2015), p. 8440. ISSN: 2041-1723. DOI: 10.1038/ncomms9440.

[31] Kiyosi Itô. "On a formula concerning stochastic differentials". In: *Nagoya Mathematical Journal* 3 (1951), pp. 55–65.

[32] Derek Juba, Debra J. Audus, Michael Mascagni, Jack F. Douglas, and Walid Keyrouz. "ZENO: Software for calculating hydrodynamic, electrical, and shape properties of polymer and particle suspensions". In: *Journal of Research of National Institute of Standards and Technology* 122.20 (2017). DOI: 10.6028/jres.122.020. URL: https://doi.org/10.6028/jres.122.020.

[33] John Jumper et al. "Highly accurate protein structure prediction with AlphaFold". In: *Nature* 596.7873 (2021), pp. 583–589.

[34] Martin Karplus and Gregory A. Petsko. "Molecular dynamics simulations in biology". In: *Nature* 347.6294 (Oct. 1990), pp. 631–639. DOI: 10.1038/347631a0. URL: https://doi.org/10.1038/347631a0.

[35] Svetlana S Khaymina, John M Kenney, Mechthild M Schroeter, and Joseph M Chalovich. "Fesselin is a natively unfolded protein". In: *Journal of proteome research* 6.9 (2007), pp. 3648–3654.

[36] Yu-Hao Lin, De-Chen Qiu, Wen-Han Chang, Yi-Qi Yeh, U-Ser Jeng, Fu-Tong Liu, and Jie-rong Huang. "The intrinsically disordered N-terminal domain of galectin-3 dynamically mediates multisite self-association of the protein through fuzzy interactions". In: *Journal of Biological Chemistry* 292.43 (2017), pp. 17845–17856. DOI: 10.1074/jbc.M117.802793.

[37] Maciej Lisicki, Bogdan Cichocki, and Eligiusz Wajnryb. "Near-wall diffusion tensor of an axisymmetric colloidal particle". In: *The Journal of Chemical Physics* 145.3 (2016). DOI: 10.1063/1.4958727.

[38] Joseph A Marsh and Julie D Forman-Kay. "Sequence determinants of compaction in intrinsically disordered proteins". In: *Biophysical journal* 98.10 (2010), pp. 2383–2390.

[39] Suman Nag, Bidyut Sarkar, Arkarup Banerjee, Bankanidhi Sahoo, KAS Varun, and Sudipta Maiti. "The nature of the amyloid-β monomer and the monomer-oligomer equilibrium". In: *Biophysical Journal* 100.3 (2011), 202a.

[40] Gerhard Nägele. "Colloidal hydrodynamics". In: *Physics of complex colloids*. IOS Press, 2013, pp. 507–601. DOI: 10.3254/978-1-61499-278-3-507.

[41] H.C. Öttinger. *Stochastic Processes in Polymeric Fluids: Tools and Examples for Developing Simulation Algorithms*. Springer Berlin Heidelberg, 2012. ISBN: 9783642582905. URL: https://books.google.pl/books?id=b1rrCAAAQBAJ.

[42] Monika Poznar, Rafał Hołubowicz, Magdalena Wojtas, Jacek Gapiński, Ewa Banachowicz, Adam Patkowski, Andrzej Ożyhar, and Piotr Dobryszycki. "Structural properties of the intrinsically disordered, multiple calcium ion-binding otolith matrix macromolecule-64 (OMM-64)". In: *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1865.11 (2017), pp. 1358–1371.

[43] Christopher Rackauckas and Qing Nie. "DifferentialEquations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia". In: *Journal of Open Research Software* 5.1 (2017). DOI: 10.5334/jors.151.

[44] Maduni Ranasinghe, Jonathan M Fogg, Daniel J Catanese Jr, Lynn Zechiedrich, and Borries Demeler. "Suitability of double-stranded DNA as a molecular standard for the validation of analytical ultracentrifugation instruments". In: *European Biophysics Journal* 52.4-5 (2023), pp. 267–280.

[45] Osborne Reynolds. "An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels". In: *Philosophical Transactions of the Royal Society of London* 174 (1883), pp. 935–982. DOI: 10.1098/rstl.1883.0029.

[46] Prince E. Rouse. "A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers". In: *The Journal of Chemical Physics* 21.7 (1953), pp. 1272–1280. DOI: 10.1063/1.1699180. eprint: https://doi.org/10.1063/1.1699180. URL: https://doi.org/10.1063/1.1699180.

[47] Bartosz Różycki and Evzen Boura. "Conformational ensemble of the full-length SARS-CoV-2 nucleocapsid (N) protein based on molecular simulations and SAXS data". In: *Biophysical chemistry* 288 (2022), p. 106843. DOI: 10.1016/j.bpc.2022.106843.

[48] Alessandro Sicorello, Bartosz Różycki, Petr V Konarev, Dmitri I Svergun, and Annalisa Pastore. "Capturing the conformational ensemble of the mixed folded polyglutamine protein ataxin-3". In: *Structure* 29.1 (2021), pp. 70–81.

[49] The Svedberg and JB Nichols. "The application of the oil turbine type of ultracentrifuge to the study of the stability region of carbon monoxide-hemoglobin". In: *Journal of the American Chemical Society* 49.11 (1927), pp. 2920–2934.

[50] Maria E Tomasso, Micheal J Tarver, Deepa Devarajan, and Steven T Whitten. "Hydrodynamic radii of intrinsically disordered proteins determined from experimental polyproline II propensities". In: *PLoS computational biology* 12.1 (2016), e1004686.

[51] Nancy Tompson. "Chapter 6 FCS". In: *Fluorescence Spectroscopy*. Ed. by Joseph R. Lakowicz. Vol. 1. Kluwer Academic Publishers, 2002, pp. 337–378.

[52] Dudu Tong, Sichun Yang, and Lanyuan Lu. "Accurate optimization of amino acid form factors for computing small-angle X-ray scattering intensity of atomistic protein structures". In: *Journal of Applied Crystallography* 49.4 (2016), pp. 1148–1161.

[53] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library. Elsevier Science, 2011. ISBN: 9780080475363.

[54] NG Van Kampen and JJ Lodder. "Constraints". In: *American Journal of Physics* 52.5 (1984), pp. 419–424. DOI: 10.1119/1.13647.

[55] Jonathan J Ward, Jaspreet S Sodhi, Liam J McGuffin, Bernard F Buxton, and David T Jones. "Prediction and functional analysis of native disorder in proteins from the three kingdoms of life". In: *Journal of molecular biology* 337.3 (2004), pp. 635–645. DOI: 10.1016/j.jmb.2004.02.002.

[56] Da Wei, Parviz G Dehnavi, Marie-Eve Aubin-Tam, and Daniel Tam. "Measurements of the unsteady flow field around beating cilia". In: *Journal of Fluid Mechanics* 915 (2021), A70. DOI: 10.1017/jfm.2021.149.

[57] Bruno H. Zimm. "Dynamics of Polymer Molecules in Dilute Solution: Viscoelasticity, Flow Birefringence and Dielectric Loss". In: *J. Chem Phys* 24.2 (Feb. 1956), pp. 269–278. DOI: 10.1063/1.1742462.

[58] Pawel J Zuk, Bogdan Cichocki, and Piotr Szymczak. "GRPY: an accurate bead method for calculation of hydrodynamic properties of rigid biomacromolecules". In: *Biophysical Journal* 115.5 (2018), pp. 782–800. DOI: 10.1016/j.bpj.2018.07.015.