Histone mark recognition controls nucleosome translocation via a kinetic proofreading mechanism: Confronting theory and high-throughput experiments

Ralf Blossey, Helmut Schiessel

Histone mark recognition controls nucleosome translocation via a kinetic proofreading mechanism: confronting theory and high-throughput experiments

Ralf Blossey[1] and Helmut Schiessel[2]
[1] *Univ. Lille, CNRS, UMR8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France.*
[2] *Institute Lorentz for Theoretical Physics, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands.*

Chromatin remodelers are multi-domain enzymatic motor complexes that displace nucleosomes along DNA and hence 'remodel chromatin structure', *i.e.*, they dynamically reorganize nucleosome positions in both gene activation and gene repression. Recently, experimental insights from structural biology methods and remodeling assays have substantially advanced the understanding of these key chromatin components. Here we confront the kinetic proofreading scenario of chromatin remodeling, which proposes a mechanical link between histone residue modifications and the ATP-dependent action of remodelers, with recent experiments. We show that recent high-throughput data on nucleosome libraries assayed with remodelers from the ISWI family are in accord with our earlier predictions of the kinetic proofreading scenario. We make suggestions for new experimentally verifiable predictions of the kinetic proofreading scenarios for remodelers from other families.

I. INTRODUCTION

Active systems are currently one of the most intensive fields of research within the statistical physics community. Built on a large body of work dealing with individual motors (see, e.g., the review [1]), the field has turned towards studies of the collective behaviour of 'active' constituents, see, e.g. [2, 3]. Chromatin remodeling motors, which actively displace and remove nucleosomes from the chromatin fiber, have so far received only little attention in the statistical physics or biophysics literature; for individual nucleosomes, see the papers by, e.g., [1, 4–7]; for their collective behaviour, [8–11]. These motors are involved in numerous processes that regulate the access of proteins and other molecules to DNA in the nucleus, in particular to genes prior to transcription and DNA repair [12–14]. The lack of attention may in part be explained by the structural complexity and size of remodelers, which has for a long time only allowed to resolve few and in particular often only partial structures; this also impeded experimental studies of remodeler dynamics, except in artificial constructs [15].

Chromatin remodelers are built around evolutionarily conserved two-domain ATPase units which belong to the helicase-related superfamily II (SF2). They can be grouped into a small number of families differing from each other by their accessory subunits, see [16, 17]; these papers also explain the nomenclature used to characterize remodelers. Very recently, the more widespread use of cryo-electron microscopy and FRET-imaging techniques have led to numerous new results, in particular on the (small) chromatin remodeler Chd1 [18–22] and a still more simplified variant, Snf1-SncS [23].

Besides the active displacement which remodelers exert on nucleosomes, the focussing of their activities on *specific* nucleosomes is also tightly regulated. In 2008, we postulated that kinetic proofreading occurs in the process of chromatin remodeling [24]. Kinetic proofreading is a mechanism of biochemical regulation which couples reversible and irreversible reactions in order to enhance specificity [25, 26]. Recently, there has also been a revival of theoretical interest in these processes [27–31] within the statistical physics community. The classic example of kinetic proofreading remains mRNA translation into proteins, which has meanwhile become textbook material [12, 32].

Our kinetic proofreading scenario for chromatin remodeling combines the histone tail modifications (the 'histone code'), which are recognized by accessory chromatin remodeler domains, with the irreversible motor action, driven by ATP [24]. In this way, we arrived at the prediction that the overall error fraction of remodeling, which differentiates between 'right' or 'wrong' nucleosomes to be remodeled, can be reduced by about two orders of magnitude. This result is in accord with estimates obtained from experiments performed on remodelers from the ISWI family [33, 34], which we briefly review below.

Given the enormous complexity of chromatin remodeling processes due to the existence of several remodeler families with different domain compositions, combined with the complexity of histone modifications, ultimately only suitable high-throughput approaches are principally capable of giving a general picture of the validity of the kinetic proofreading scenario. In our view, a first significant step in this direction has recently been achieved by a research team led by C.D. Allis and T.W. Muir, who considered the remodeling capability of several members from the ISWI family as a function of the presence of specific histone modifications [35]. This work is based on an earlier developed 'barcoding' approach which allows to synthesize post-translationally modified nucleosomes [36] - a crucial step for *in-vitro* assays, as it allows to bypass the complex processes involved in the placement of post-translational modifications *in vivo*.

These recent experimental studies have motivated us to return to the problem and to reanalyze the available data in the light of the kinetic proofreading scenario for

chromatin remodeling. For this we first recapitulate our model before we address the relation between experiments and theory. Subsequently, we discuss the inclusion of additional regulatory mechanisms in our scenario.

II. THE KINETIC PROOFREADING SCENARIO OF CHROMATIN REMODELING

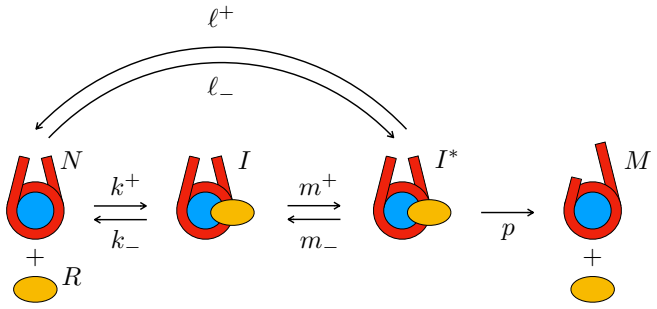

FIG. 1. Kinetic proofreading scheme for chromatin remodeling with nucleosome (N), remodeler (R), complex (I), activated complex (I^*), mobile nucleosome (M). The translocation step with rate p is shown to lead to remodeler dissociation; it should be noted that this step is in general processive.

In our derivation of the model we follow the discussion given in [33], and introduce the following molecular actors, defined by their concentrations. The first are the nucleosome, N, and the remodeler, R. Both can form a complex I, which upon a first ATP-dependent consumption step becomes activated, I^*, and hence can become mobile, M. The corresponding reaction scheme reads as

$$R + N \underset{k_-}{\overset{k^+}{\rightleftharpoons}} I \underset{m_-}{\overset{m^+}{\rightleftharpoons}} I^* \to_p M\,. \tag{1}$$

see also Fig. 1. For the first step, the remodeler recruitment, the rates are assumed to depend on the modification state of the histones, *i.e.* on the combinatorial effects of acetylations, methylations etc. present or absent on the histone residues. The rates denoted by m refer to the activation of the nucleosome, and the reaction with rate p describes the translation of the nucleosome/remodeler complex; this final irreversible (ATP-consuming) reaction will in principle be repeated many times. Eq. (1) is accompanied by the loss reaction

$$I^* \underset{\ell_-}{\overset{\ell^+}{\rightleftharpoons}} R + N \tag{2}$$

which will occur with a higher rate if the incorrect nucleosome-remodeler complex has formed. Eqs. (1) and (2) give rise to the two ordinary differential equations for the concentrations of complexes, which are given by

$$[\dot{I}] = k^+[R][N] - (k_- + m^+)[I] + m_-[I^*]\,, \tag{3}$$

and

$$[\dot{I}^*] = m_+[I] + \ell_-[R][N] - (m_- + \ell^+ + p)[I^*]\,. \tag{4}$$

Assuming equilibrium conditions, $[\dot{I}] = [\dot{I}^*] = 0$, one can eliminate the intermediate $[I]$ from the resulting equations, and is, after some algebra, left with the ratio

$$f \equiv \frac{[I^*]}{[R][N]} = \frac{k^+m^+ + \ell_-(k_- + m^+)}{(\ell^+ + p)(k_- + m^+) + k_-m_-}\,. \tag{5}$$

From this ratio, one can define an error ratio by comparing a 'correct' to an 'incorrect' reaction, which corresponds to an energetically favored or disfavored pathway. Thus in the following we will consider the quantity $F \equiv f_2/f_1$, which reads as

$$F \equiv \frac{f_2}{f_1} = \frac{k_2^+m_2^+}{k_1^+m_1^+}\frac{(\ell_1^+ + p_1)(k_{1,-} + m_1^+)}{(\ell_2^+ + p_2)(k_{2,-} + m_2^+)}\,, \tag{6}$$

assuming that the reaction rates ℓ_-, m_- involving the activated nucleosome are considered as zero, $\ell_- = m_- = 0$.

The final error fraction formula for F obviously depends on a large set of parameters, which are not easy to determine under comparable conditions, and this both *in-vitro* and *in-vivo*. One therefore has to rely on simplifying assumptions, both theoretically and experimentally. For our (theoretical) scenario, we assume that the specificity is in the off-rates, as in the original Hopfield scenario.

A. Two scenarios for chromatin remodeling

We now briefly recall our reasoning from [33] which was based on the results discussed in detail in [34]. In [33], we distinguished between our original scenario [24] and Narlikar's result. In our Blossey-Schiessel (BS) scenario, we assumed $m_i^+ \ll k_{i,-}$, $m_1^+ \approx m_2^+$, $k_1^+ \approx k_2^+$, and $\ell_i^+ \gg p_i$. The error fraction then simplifies to the approximate expression

$$F_{\rm BS} \approx \frac{k_{1,-}\ell_1^+}{k_{2,-}\ell_2^+} \approx \left(\frac{k_{1,-}}{k_{2,-}}\right)^2 \tag{7}$$

under the additional Hopfield-based assumption $k_{i,-} = \ell_i^+$. Since the free energy enters into the rate via the Boltzmann factor, one has $F_{BS} \sim \exp(-2\Delta G/k_BT)$, and the preferred substrate becomes significantly favored.

In [34], G.J. Narlikar discusses the full equation (6), since from her remodeling assays she has quantitative information about several of the involved rates, which we collect in Table 1. In this case, and with the assumptions listed in the caption of Table 1, the formula for F, eq.(6) simplifies to

$$F \equiv \frac{k_{\rm I,2}}{k_{\rm I,1}}\frac{(k_{\rm off,1} + k_{\rm tr,1})(k_{\rm off,1} + k_{\rm I,1})}{(k_{\rm off,2} + k_{\rm tr,2})(k_{\rm off,2} + k_{\rm I,2})}\,. \tag{8}$$

Using the experimental estimates, it reduces to a quantitative estimate of $1/F \approx 313$, in accord with our earlier crude estimate [24].

$k_{1,-}$	$k_{2,-}$	k_1^+	k_2^+	m_1^+	m_2^+	ℓ_1^+	ℓ_2^+	p_1	p_2
$k_{\mathrm{off},1}$	$k_{\mathrm{off},2}$	k_1^+	k_2^+	$k_{\mathrm{I},1}$	$k_{\mathrm{I},2}$	$k_{\mathrm{off},1}$	$k_{\mathrm{off},2}$	$k_{\mathrm{tr},1}$	$k_{\mathrm{tr},2}$
8	160	n.d.	n.d.	20	1	8	160	20	80

TABLE I. Remodeling rates appearing in eq.(6), for the experiments by G.J. Narlikar. Top line: our definition of rates; Middle name: Narlikar's definition of rates; assumed is $k_1^+ = k_2^+$, $k_{i,-} = k_{\mathrm{off},i}$. Last line: experimental estimates; all in 1/min. n.d.: not determined, assumed equal and thus drop out of the result. Adapted from [34], following [33].

ACF	1	2	3	4	5	6	7	8	9
x	1.9	0.72	0.15	0.81	0.35	-0.15	-0.15	-0.69	-1.1
$F_{\mathrm{MN}} = 2^x$	3.73	1.65	1.11	1.75	1.27	0.90	0.90	0.62	0.47

TABLE II. Remodeling rates for the remodeler ACF for nine different histone modifications (1-9); taken from [35], Fig. 3b; x is the remodeling rate factor defined by $x \equiv \log_2(k_{\mathrm{MN}}/k_{\mathrm{unmod}})$. The modifications are identified in that graph; for our purpose, the detailed nature of the modification is not essential and can be looked up in [35].

B. High-throughput experiments and the kinetic proofreading theory

We now turn to the recent results on nucleosome libraries. In the work by [35], a DNA barcoded library of nucleosomes with specifically modified histone tails was generated and used in high-throughput remodeling assays for a range of chromatin remodelers from the ISWI family. These experiments allowed to generate data about 800 remodeling rate constants from slightly less than an impressive 30.000 individual kinetic measurements. These data show that the remodelers show a specificity for distinct histone modification patterns. To quantify the results, the authors have compared the remodeling rate, k_{MN}, of the modified nucleosome, to that of an unmodified nucleosome. Figure 3b in Ref. [35] tables nine specific histone tail modifications and the values of the remodeling rate, normalized to that of the unmodified nucleosome are given in logarithmic form via the expression $\log_2(k_{\mathrm{MN}}/k_{\mathrm{unmod}})$.

The first question therefore is as to relate this experimentally measured rate to the theoretical expression (6), as it does not distinguish explicitly between the processes in our model. The inspection of the construction of the remodeling assay, as described in the methods section of [35], shows the authors separated the rates involved in the binding process from the processive remodeling process, *i.e.* the motor-driven displacement of the nucleosome. The experimental binding data, *i.e.* the values for the rates k^+, for the remodelers have been obtained in both [35] and [36], however for different remodeler constructs than used in the translocation experiments, such that a full comparison to our error formula cannot be performed at present on the basis of the available data. Considering the monitored remodeling process, we identify F_{MN} with the ratio of the translation rates p_i , *i.e.*

$$F_{\mathrm{MN}} = 2^{\log_2(k_{\mathrm{MN}}/k_{\mathrm{unmod}})} \approx \frac{p_1}{p_2}\,. \tag{9}$$

Comparing eq.(9) to eq.(6) we have for the other rates in eq.(6) the conditions $k_1^+ \approx k_2^+$, i.e., the initial binding rates of the remodeler to the nucleosome are of similar magnitude, and further that the activation and translocation rates dominate over the corresponding dissociation rates, i.e. $m_i^+ \gg k_{i,-}$ and $p_i \gg l_i^+$, conditions met by the experiments in [35].

In Table II, we show the analysis of the data generated by the remodeling assay for the case of the remodeler ACF (**A**TP-dependent **C**hromatin Assembly and Remodeling **F**actor). These data are in fact in the same order of magnitude range as the data by Narlikar, as the comparison with the translocation rate values p in Table I shows. The analysis of the data for other remodelers from the high-throughput assay data set support the same trend (data not shown).

The results of Table II reflect underlying properties of the ISWI family of remodelers. As studied already as early as 2007 by Ferreira et al. [39] in their early remodeling assays, there is a clear distinction in the activation mechanism of ISWI-type remodelers when compared to, e.g., RSC (**R**emodels the **S**tructure of **C**hromatin). Changes to the histone residues by post-translational modifications in the former case affect mainly the catalytic rate, while they affect more strongly the dissociation rate, hence remodeler recruitment, in the latter case.

III. SUGGESTIONS FOR FURTHER EXPERIMENTS ON OTHER REMODELERS

Given the expanding knowledge on chromatin remodeler actions, see [17], we finally discuss how, depending on the remodeler family, the kinetic proofreading scenario needs to be extended in specific ways, taking into account co-regulatory processes. These can affect the rates for all four branches of the kinetic remodeling scenario as described in the reaction schemes (1) and (2). Specifically, passing though each set of the rates:

a) *Remodeler recruitment*; (rates $k^{+,-}$). Clapier and Cairns [37] have described an auto-regulatory mechanism in which a domain of an ISWI remodeler competes with the H4-tail. This mechanism impedes the binding to a 'wrong' substrate. It has already been integrated into the proofreading scenario in [38], however so far without quantitative estimates for these rates. Recruitment of the remodeler may be more strongly affected by changes in histone residues for other remodeler families which bear specific recognition domains, e.g. bromodomains for acetylations or chromodomains for methylations. Quantitative information in particular for bromodomain specificities is available, but it has not yet been brought in the context of chromatin remodeling, with the exception of

[47].

b) *Nucleosome activation*; (rates m^+). Our discussion above has focused on the translation rates; however, the assays by [39] and the remodeling assays on ISWI remodelers by Narlikar and collaborators [40] show that also the activation rates are affected by the histone residue modifications.

c) *Nucleosome loss*; (rates ℓ^+). As an example for the co-regulation of nucleosome loss, we cite the remodeler RSC. It can be co-regulated by nuclear actin molecules from the Arp family [41]. These were shown to be involved in tuning the processivity of the remodelers, *i.e.* in particular to tune their role in the eviction of nucleosomes and hence the loss rate.

d) *Translocation*; (rate p). Remodeling rates have been measured by several types of assays, and the high-throughput experiments we address in this paper are proof of this success. In this problem it is of great interest to understand the way remodelers move the DNA around the nucleosome. From an experimental point of view it has recently been elucidated by cryo-electron microscopy at 4 Å resolution of a truncated version of a basic remodeler in complex with a 167-bp DNA fragment containing the 601-positioning sequence [23]. The remodeler makes contacts with two lobes in the structure, injecting twist defects that propagate through and exit at the other end. This scenario, proposed on static data, could be validated in coarse-grained simulations [42].

In a simple coarse-grained model the helicase action of the remodeler can be modeled by a Brownian dimer described by a system of coupled Langevin equations which can be solved exactly [43]. This model only considers the dynamics of the footprints of the two remodeler lobes on DNA denoted by x_1 and x_2. In the course of the remodeler action, footprint x_1 is first shifted towards location x_2 via twist defect injection; the displacement of footprint x_2 follows in due course leading to twist defect ejection. The DNA length $x \equiv x_1 - x_2 - x_0$ where x_0 is the equilibrium extension of the DNA around the nucleosome thus relaxes after one remodeler step which is typically 1-2 bp large [42, 44, 45]; the variable x is therefore small, irrespective of the location of x_1 and x_2 along the internucleosomal DNA. The motion of the footprints thus is indeed akin to an inchworm motion of DNA around the histone octamer [46]. This model predicts a linear dependence of the remodeler velocity along the DNA track on ATP-dependent driving strength. Experiments with artificial motor constructs for the remodeler RSC by Sirinakis et al. [15] earlier showed a weaker logarithmic dependence on driving, $v_{max} \sim \ln[\mathrm{ATP}]$. It remains an interesting open task to develop quantitative models than can predict translocation rates.

e) *Transcription factors.* In our original scenario of 2008, we had included the possible presence of transcription factors. They are indeed relevant e.g. in the build-up of enhancer elements in gene activation. We have discussed the case of the INF-β gene for which the whole sequence of recruitment events prior to the start of transcription has been determined [47]. Again, as for the other cases, there are no detailed measurements of kinetic rates available for this case.

IV. CONCLUSIONS

To conclude, in this paper we have discussed the kinetic proofreading scenario of chromatin remodeling in the light of recent experimental insights. High-throughput libraries of histone tail-modified nucleosomes are remodeled according to distinct rates depending on histone residue modification and remodeler type, chosen from the ISWI family. As we show here, these experimentally obtained rates are in accord with the kinetic proofreading scenario of chromatin remodeling, for which they provide partial information, *i.e.* on the translocation rates. It is argued that the character of the results is specific to the chosen ISWI remodeler family, in which the main effect of histone modifications is of catalytic origin. In order to further validate the scenario, we point out that the measurement of the rates involved in chromatin remodeling, as made explicit in the kinetic proofreading scenario, for other types of remodelers are therefore highly desirable. As we discuss, for these remodelers also additional regulatory effects may influence the overall remodeling kinetics.

[1] D. Chowdhury. Phys. Rep. **529**, 1-197 (2013)
[2] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner and L. Schimansky-Geier, Eur. Phys. J. Special Topics, 1-162 (2012)
[3] M.C. Marchetti, J.F. Joanny, S. Ramaswamy, T.B. Liverpool, J. Prost, Madan Rao and R. Aditi Simha, Rev. Mod. Phys. **85**, 1143-1189 (2013)
[4] T. Chou, Phys. Rev. Lett. **99**, 058105 (2007)
[5] Y. Vandecan and R. Blossey, Phys. Rev. E **85**, 061920 (2012)
[6] Y. Vandecan and R. Blossey, Phys. Rev. E **88**, 012728 (2013)
[7] K.H. Iwasa and A.M. Florescu, J. Stat. Mech. Theor. Exp., 054030 (2016)
[8] R. Padinhateeri and J.F. Marko, Proc. Natl. Acad. Sci **108**, 7799 (2011)
[9] A.M. Florescu, H. Schiessel and R. Blossey, Phys. Rev. Lett. **109**, 118103 (2012)
[10] R. Bruinsma, A.Y. Grosberg, Y. Rabin and A. Zidovska. Biophys. J. **106**, 1871-1881 (2014)

[11] R.D. Schram, H. Klinker, P.B. Becker and H. Schiessel. Eur. Phys. J. E **38**, 85 (2015)

[12] R. Blossey. *Chromatin: Structure, Dynamics, Regulation* CRC Press Boca Raton, USA (2017)

[13] C.L. Smith and C.L. Peterson. Current Topics in Developmental Biology **65**, 115-148 (2004).

[14] W.K.M. Lai and B.F. Pugh. Nature Rev. Mol. Cell. Biol. **18**, 548-562 (2017)

[15] G. Sirinakis, C.R. Clapier, Y. Gao, R. Viswanathan, B.R. Cairns, Y.L. Zhang. EMBO Journal **30**, 2364-2372

[16] C.R. Clapier and B. R. Cairns. Annu. Rev. Biochem. **78**, 273-304 (2009)

[17] C.R. Clapier, J. Iwasa, B.R. Cairns and C.L. Peterson. Nature Rev. Mol. Cell Biol. **18**, 407-422 (2018)

[18] J.N. McKnight, T. Tsukiyama and G. D. Bowman. Genome Res. **26**, 693-704 (2016)

[19] I.M. Nodelman, F. Bleichert, A. Patel, R. Ren, K.C. Horvath, J.M. Berger and G.D. Bowman. Mol. Cell. **65**, 447-459 (2017)

[20] J. Winger and G. D. Bowman. J. Mol. Bio. **429**, 808-822 (2017)

[21] R. Sundaramoorthy *et al.*. eLife **6**, e22510 (2017)

[22] L. Farnung, S.M. Vos, C. Wigge and P. Cramer. Nature **550**, 539-542 (2017)

[23] X. Liu, M. Li, X. Xia, X. Li and Z. Chen. Nature **544**, 440-445 (2017)

[24] R. Blossey and H. Schiessel. HFSP J. **2** 167-170 (2008)

[25] J.J. Hopfield. Proc. Natl. Acad. Sci. (USA) **71**, 4135-4139 (1974)

[26] J. Ninio. Biochimie **57**, 587-595 (1975)

[27] A. Murugan, D.A. Huse and S. Leibler. Proc. Natl. Acad. Sci. (USA) **109**, 12034-12039 (2012)

[28] M. Depken, J.M.R. Parrondo and S.W. Grill. Cell Reports **5**, 521-530 (2013)

[29] A. Murugan, D.A. Huse and S. Leibler. Phys. Rev. X **4**, 021016 (2014)

[30] D. Hartich, A.C. Barato and U. Seifert, New J. Phys. **17**, 055026 (2015)

[31] R. Rao and L. Peliti. J. Stat. Mech. (2015) P06001

[32] U. Alon. *An Introduction to Systems Biology* CRC Press Boca Raton, USA (2006)

[33] R. Blossey and H. Schiessel. Biophys. J **101**, L30-L32 (2011)

[34] G.J. Narlikar. Curr. Opin. Chem. Biol. **14**, 660-665 (2012)

[35] G.P. Dann *et al.*. Nature **548**, 607-611 (2017)

[36] U.T.T. Nguyen *et al.*. Nature Methods **11**, 834-840 (2014)

[37] C.R. Clapier and B.R. Cairns. Nature **492**, 280-285 (2012)

[38] G. Brysbaert, M.F. Lensink and R. Blossey. J. Phys.: Condens. Matter **27**, 064108 (2015)

[39] H. Ferreira, A. Flaus and T. Owen-Hughes. J. Mol. Biol. **374**, 563-579 (2007)

[40] T.R. Blosser, J.G. Yang, M.D. Stone, G.J. Narlikar and X. Zhuang. Nature **462**, 1022-1027 (2009)

[41] C.R. Clapier *et al.* Mol. Cell **62**, 453-461 (2016)

[42] G.B. Brandani and S. Takada, PLoS Comp. Bio. **14**, e1006512 (2018)

[43] R. Blossey and H. Schiessel, J. Phys. A: Math. Theor. **52**, 085601 (2019)

[44] S. Deindl *et al.* Cell **152**, 442-452 (2013)

[45] B.T. Harada *et al.* eLife **5**, e10051 (2016)

[46] D.A. Wigley and G.D. Bowman. Nature Struct. Mol. Biol. **24**, 498-500 (2017)

[47] R.P. Singh, G. Brysbaert, M.F. Lensink, F. Cleri, R. Blossey. AIMS Biophysics **2** 398-411 (2015)