

# Self-entanglement of bovine serum albumin in shear flow: cumulative effects and irreversibility

Agnieszka Budek<sup>1</sup>, Marek Cieplak<sup>2</sup>, and Piotr Szymczak<sup>1,a</sup>

<sup>1</sup> Institute of Theoretical Physics, Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland

<sup>2</sup> Institute of Physics, Polish Academy of Sciences, Aleja Lotnikow 32/46, 02-668 Warsaw, Poland

Received 8 July 2018 / Received in final form 29 July 2018  
Published online 28 March 2019

**Abstract.** We perform Brownian dynamics simulations of shear-induced unfolding of bovine serum albumin. This protein was reported to unfold irreversibly in the shear flow. Using a coarse-grained model of a protein, we track the conformational changes induced by the flow and observe that after an extended exposure to shear albumin loses its ability to refold even when the flow has been turned off. Instead, it is trapped in a metastable state characterized by a large degree of self-entanglement which prevents the molecule from folding into the native conformation. This state becomes more populated with time, which can explain the cumulative effect of the shear observed in the experiments.

## 1 Introduction

Experimental evidence regarding the influence of shear flow on the protein structure is rather contradictory [1–3]. Some studies show that the effect of shear flow is fully reversible [4,5], with the protein correctly folding to its native conformation after the flow is turned off. In other cases, however, irreversible effects are reported, with the structure of the protein permanently disrupted as an effect of the shear. Also, some experiments indicate that effects of the shear can be of a cumulative nature and should be described in terms of the strain history [6–8],  $\gamma = \tau_s \dot{\gamma}$ , where  $\dot{\gamma}$  is the shear rate and  $\tau_s$  is the total time during which the sample was exposed to the shear. Other studies observe no such dependence and correlate the changes in protein structure directly with the value of the shear rate,  $\dot{\gamma}$  [1,5,9–12].

In particular, in a recent communication Bekard et al. [8] have analyzed shear-induced unfolding of bovine albumin, and found that it unfolds at shear rates of  $100\text{--}300\text{ s}^{-1}$ . Importantly, the effects of the shear were found to be cumulative, with the degree of unfolding increasing with exposure time. Finally, most of the shear-induced conformational changes in albumin were found to be irreversible, with the protein not properly folding again after shear has been switched off. These features of shear-induced unfolding of albumin are in sharp contrast to our earlier numerical

<sup>a</sup> e-mail: [piotr.szymczak@fuw.edu.pl](mailto:piotr.szymczak@fuw.edu.pl)

results on ubiquitin and integrin [12] – the shear-induced changes in these proteins were found to be reversible, with only a small percent of misfolds. This motivated us to investigate shear-induced unfolding of albumin within the same numerical model.

## 2 Numerical model

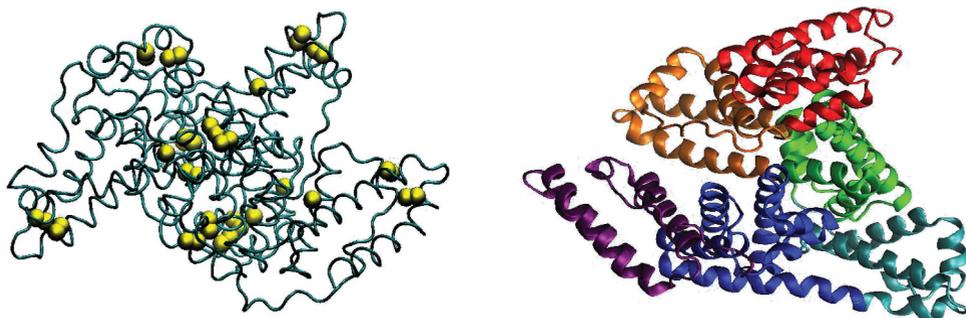
We adopt a coarse-grained,  $G\bar{o}$ -type model of a protein, in which individual amino acids are replaced by beads of uniform size placed at the locations of the  $C_\alpha$  atoms. The effective potential of the interaction between these beads is tailored to give the lowest energy to the native state of a protein. A particular implementation of the  $G\bar{o}$ -type model followed here is by Cieplak and coworkers [13,14]. In short, the protein structure is represented by a chain of  $C_\alpha$  atoms tethered along the backbone by harmonic potentials with minima at  $l_p = 3.8 \text{ \AA}$ . Effective interactions between residues are split into native and nonnative contacts by checking for overlaps between the enlarged van der Waals surfaces of the residues [15]. Amino acids ( $i$  and  $j$ ) that overlap are endowed with the effective Lennard–Jones potential  $V_{ij}(r) = 4\epsilon[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6]$  with energy scale  $\epsilon$  and pair-by-pair distances  $r_{ij}$ . The length parameters,  $\sigma_{ij}$ , are chosen such that the potential minima correspond pair by pair to the native state distance between the residues. Nonnative contacts are represented by hardcore repulsion to prevent entanglements. The disulfide bonds are covalent and are also represented by the harmonic potentials. Correct chirality is imposed by the angle-dependent term in the Hamiltonian.

The dynamics of the system is tracked using the finitely damped Langevin dynamics algorithm of Geyer and Winter [16,17], modified to account for the possible overlaps of the beads according to [18]. In contrast to the standard Brownian dynamics algorithm, this method allows for the inclusion of hydrodynamic interactions at a relative small computational cost. The simulations are carried out at the temperature corresponding to  $kT = 0.3\epsilon$ , which is the optimal folding temperature for this coarse-grained model [13]. The characteristic timescale  $\tau = \sigma^2/6D_0$  is set by the time it takes for the amino acid to diffuse the typical contact distance  $\sigma \approx 5 \text{ \AA}$ .

## 3 Results

Bovine serum albumin is a relatively large protein (583 aminoacids), consisting of about 25 alpha helices grouped in six subdomains (marked in different colors in Fig. 1), which will be denoted by the letters A to E. The numbers of contacts between different domains are given in Table 1. Importantly, the domains are also stabilized by 17 disulfide bonds (marked by yellow in Fig. 1) [19,20]. A large number of disulfide bonds make the topology of the protein complicated, with a number of linked loops formed by the backbones closed via disulfide bridges. Albumin is thus an example of a protein with non-trivial topology, related to the disulfide bridge connectivity pattern [21–23]. Entangled proteins have been the subject of intense studies in recent years [24–27], but the impact of topology on the protein dynamics is still far from understood.

Before embarking on simulations of shear-induced unfolding, we have first estimated the folding time of albumin within the  $G\bar{o}$ -model, starting from the thermally unfolded conformation, obtained at the higher temperature of  $kT = \epsilon$ . The folding time has been found to be  $t_{\text{fold}} \approx 3 \times 10^4 \tau$ . Next, we put the albumin molecule in the shear flow with shear rates  $\dot{\gamma}$  in range of  $0.1\text{--}1\tau^{-1}$ . In terms of the dimensionless Weissenberg number,  $Wi = \dot{\gamma}t_{\text{fold}}$ , this corresponds to  $10^3\text{--}10^4$ . According to a



**Fig. 1.** Structure of bovine albumin, with indication of its six helical subdomains (marked with different colors in the right panel) and its native disulfide cross-links (marked by yellow beads in the left panel).

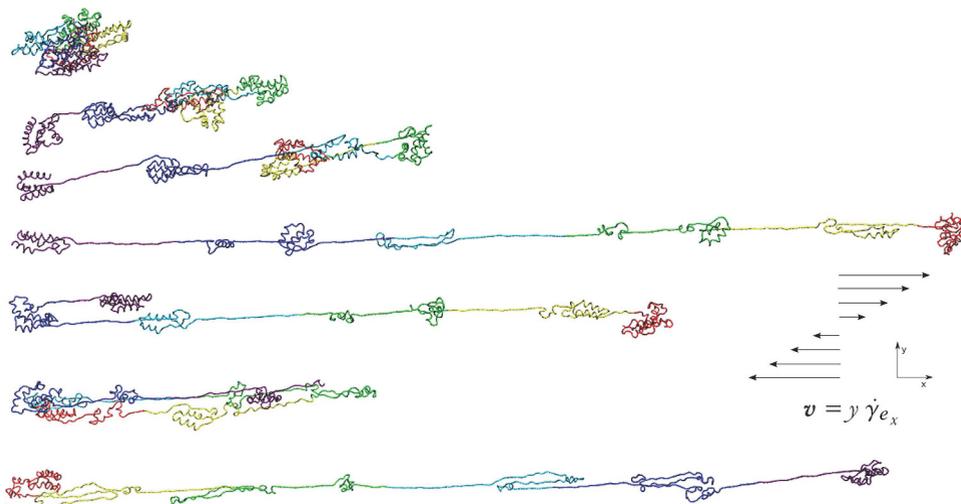
**Table 1.** The number of contacts between different subdomains in albumin, including the intradomain ones.

	A	B	C	D	E	F
A	352	42	28	0	0	0
B	42	286	33	0	11	0
C	28	33	325	39	15	0
D	0	0	39	245	29	0
E	0	11	15	29	369	40
F	0	0	0	0	40	262

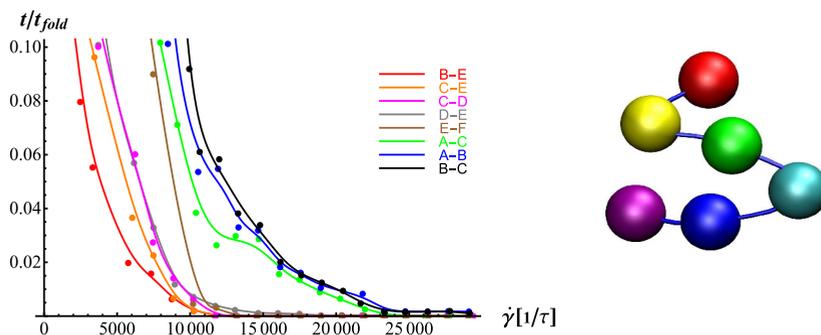
simple theoretical estimate of Jaspe and Hagen [1], the shears characterized by such Weissenberg numbers should be capable of unfolding a protein.

When albumin is placed in a shear flow, the contacts between individual domains begin to break and the characteristic tumbling motion appears (cf. Fig. 2), a series of subsequent stretching and coiling events with frequent changes in the orientation of the chain with respect to the shear axis [12,28,29]. The first to break are the weakest contact groups. As observed in Figure 3, already at the shear rates of  $\dot{\gamma} = 0.1\tau^{-1} - 0.3\tau^{-1}$ , the contacts in the middle of the chain give way, linking the domains B, C and D with E as well as the contacts between C and D. The hardest to break are the contacts between A, B and C, but even these contacts yield at the flows of  $0.5\tau^{-1}$  and higher (Tab. 1). On the other hand, the contacts within individual domains are much harder to break, and the domains retain their identity even at  $\dot{\gamma} = \tau^{-1}$ . There are two reasons for this. First, the shear necessary to unfold a given domain scales as  $N^{-7/3}$  with the number of residues in the domain [1], thus, on average, one needs  $6^{7/3} \approx 65$  times larger shear to unfold a six times smaller domain. Second, albumin subdomains are additionally stabilized by covalent disulfide bonds. Due to the above, albumin in a flow resembles a necklace with six beads (representing domains) on a string. In fact, one can devise a toy model of albumin, with six large beads (of radii equal to hydrodynamic radii of the domains) connected with effective Lennard–Jones potentials, with energy scales,  $\epsilon_{ij}$  proportional to the number of contacts between these domains, as shown schematically in the right panel Figure 3. Such a system, when placed in the flow, unfolds in a highly similar manner to the original molecule – with ABC and EF clusters surviving the longest.

Strings and necklaces can easily get tangled when tossed around [30–33] and this is precisely what happens to albumin in a shear flow, with different parts of the chain getting hooked against each other, as illustrated in Figure 4. As a result, when shear flow is turned off, the protein does not refold into its native conformation. Instead, it



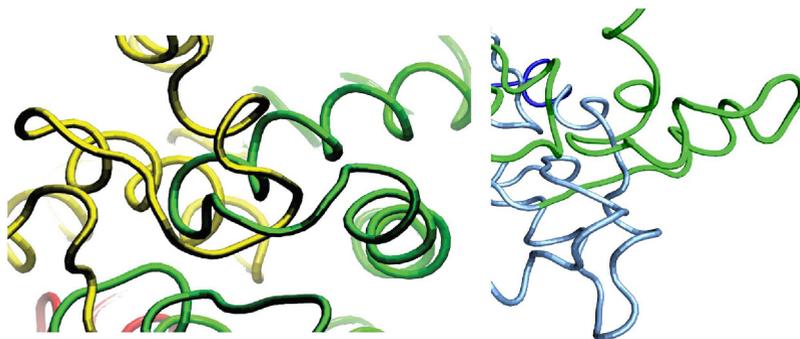
**Fig. 2.** Tumbling motion of albumin in a shear flow at  $\dot{\gamma} = 0.7\tau^{-1}$ . The colors indicate different subdomains, as marked in Figure 1.



**Fig. 3.** Left: the median times of breaking of 80% of the contacts between the domains as a function of the shear rate. Right: a toy model of albumin with six beads (representing domains) on a string (representing the peptide backbone).

reaches a compact misfolded state with a number of domains interlocked with each other, and some of the helical segments connecting the domains disrupted, as indeed observed experimentally [3]. Due to the compactness of the misfolded state as well as the large number of disulfide bonds such misfolds are extremely hard to escape from. For example, when conducting unfolding for  $T_1 = 10^4\tau$  at the shear rate of  $\dot{\gamma} = 0.5\tau^{-1}$  and then turning off the flow we found that 95% of the protein molecules have reached misfolded configurations after the period of  $T_2 = 4 \times 10^4\tau$  without the flow. Further waiting for additional  $T_3 = 5 \times 10^4\tau$  leads to no additional changes in the protein conformation. At smaller shears ( $\dot{\gamma} = 0.3\tau^{-1}$ ) for the same  $T_1$  the number of misfolded conformations is smaller (60%), but they are equally stable. Twice longer  $T_1$  brings the number of misfolds to 80%, so the effect of shear is indeed cumulative.

In summary, we have shown that the origin of the irreversible behavior of albumin in the shear flow can be linked to its relatively complicated topology, with six domains stabilized by disulfide bridges. When sheared, these domains become interlocked with each other, which leads to misfolds, as the flow is turned off. Due to a relatively



**Fig. 4.** Examples of self-entanglements of albumin in a shear flow – loops in different domains become hooked with each other.

complicated topology of this protein, these misfolds constitute deep kinetic traps in an energy landscape, and the protein becomes trapped in a self-entangled conformation.

Several limitations of this study need to be acknowledged. First, we use a simplified, coarse-grained model of a protein, which is characterized by a smoother energy landscape than the atomistic model. This might have effect on the propensity for misfolding of the protein and on the conformations of the misfolded states. On the experimental side, flow unfolding experiments with single molecules would be desirable to avoid intermolecular aggregation as another route to irreversible entrapment of unfolded proteins.

This work was supported by the National Science Centre (Poland) under research grant 2015/19/D/ST8/03199.

**Open Access** This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## References

1. J. Jaspe, S.J. Hagen, *Biophys. J.* **91**, 3415 (2006)
2. P. Szymczak, M. Cieplak, *J. Phys.: Condens. Matter* **23**, 033102 (2010)
3. I.B. Bekard, P. Asimakis, J. Bertolini, D.E. Dunstan, *Biopolymers* **95**, 733 (2011)
4. A.T. Lee, A.J. McHugh, *Biopolymers* **50**, 589 (1999)
5. L. Ashton, J. Dusting, E. Imomoh, S. Balabani, E. Blanch, *Biophys. J.* **96**, 4231 (2009)
6. S.E. Charm, B.L. Wong, *Biotechnol. Bioeng.* **12**, 1103 (1970)
7. S.E. Charm, B.L. Wong, *Science* **170**, 466 (1970)
8. I.B. Bekard, P. Asimakis, C.L. Teh, T. Ryan, G.J. Howlett, J. Bertolini, D.E. Dunstan, *Soft Matter* **8**, 385 (2012)
9. C.A. Siedlecki, B.J. Lestini, K. Kottke-Marchant, S.J. Eppel, D.L. Wilson, *Blood* **88**, 2939 (1996)
10. A. Alexander-Katz, M.F. Schneider, S.W. Schneider, A. Wixforth, R.R. Netz, *Phys. Rev. Lett.* **97**, 138101 (2006)
11. S.W. Schneider, S. Nuschele, A. Wixforth, C. Gorzelanny, A. Alexander-Katz, R.R. Netz, M.F. Schneider, *Proc. Natl. Acad. Sci. USA* **104**, 7899 (2007)
12. P. Szymczak, M. Cieplak, *J. Chem. Phys.* **127**, 155106 (2007)
13. J.I. Sulkowska, M. Cieplak, *Biophys. J.* **95**, 3174 (2008)
14. M. Sikora, J.I. Sulkowska, M. Cieplak, *PLoS Comput. Biol.* **5**, e1000547 (2009)

15. J. Tsai, R. Taylor, C. Chotchia, M. Gerstein, *J. Mol. Biol.* **290**, 253 (1999)
16. T. Geyer, U. Winter, *J. Chem. Phys.* **130**, 114905 (2009)
17. T. Geyer, *BMC Biophys.* **4**, 7 (2011)
18. P. Zuk, E. Wajnryb, K. Mizerski, P. Szymczak, *J. Fluid Mech.* **741**, R5 (2014)
19. B.X. Huang, H.Y. Kim, C. Dass, *J. Am. Soc. Mass Spectrom.* **15**, 1237 (2004)
20. I. Rombouts, B. Lagrain, K.A. Scherf, M.A. Lambrecht, P. Koehler, J.A. Delcour, *Sci. Rep.* **5**, 12210 (2015)
21. C. Liang, K. Mislow, *J. Am. Chem. Soc.* **117**, 4201 (1995)
22. J.M. Mas, P. Aloy, M.A. Marti-Renom, B. Oliva, R. de Llorens, F.X. Aviles, E. Querol, *J. Comput. Aided Mol. Des.* **15**, 477 (2001)
23. P. Dabrowski-Tumanski, J.I. Sulkowska, *Proc. Natl. Acad. Sci. USA* **114**, 3415 (2017)
24. W.R. Taylor, K. Lin, *Nature* **421**, 25 (2003)
25. P. Virnau, L.A. Mirny, M. Kardar, *PLoS Comput. Biol.* **2**, e122 (2006)
26. K.C. Millett, E.J. Rawdon, A. Stasiak, J.I. Sulkowska, *Biochem. Soc. Trans.* **41**, 533 (2013)
27. P.F. Faísca, *Comput. Struct. Biotechnol. J.* **13**, 459 (2015)
28. D.E. Smith, H.P. Babcock, S. Chu, *Science* **283**, 1724 (1999)
29. C.M. Schroeder, R.E. Teixeira, E.S. Shaqfeh, S. Chu, *Phys. Rev. Lett.* **95**, 018301 (2005)
30. A. Belmonte, M.J. Shelley, S.T. Eldakar, C.H. Wiggins, *Phys. Rev. Lett.* **87**, 114301 (2001)
31. E. Ben-Naim, Z. Daya, P. Vorobieff, R.E. Ecke, *Phys. Rev. Lett.* **86**, 1414 (2001)
32. A. Belmonte, *Proc. Natl. Acad. Sci. USA* **104**, 17243 (2007)
33. D.M. Raymer, D.E. Smith, *Proc. Natl. Acad. Sci. USA* **104**, 16432 (2007)