RESEARCH ARTICLE

Large-scale, dynamin-like motions of the human guanylate binding protein 1 revealed by multi-resolution simulations

Bogdan Barz[1], Jennifer Loschwitz[2,3], Birgit Strodel[2,3] *

1 Institute of Physical Biology, Heinrich Heine University, Düsseldorf, Germany, **2** Institute of Complex Systems: Structural Biochemistry, Forschungszentrum Jülich, Jülich, Germany, **3** Institute of Theoretical and Computational Chemistry, Heinrich Heine University, Düsseldorf, Germany

* b.strodel@fz-juelich.de

Abstract

Guanylate binding proteins (GBPs) belong to the dynamin-related superfamily and exhibit various functions in the fight against infections. The functions of the human guanylate binding protein 1 (hGBP1) are tightly coupled to GTP hydrolysis and dimerization. Despite known crystal structures of the hGBP1 monomer and GTPase domain dimer, little is known about the dynamics of hGBP1. To gain a mechanistic understanding of hGBP1, we performed sub-millisecond multi-resolution molecular dynamics simulations of both the hGBP1 monomer and dimer. We found that hGBP1 is a highly flexible protein that undergoes a hinge motion similar to the movements observed for other dynamin-like proteins. Another large-scale motion was observed for the C-terminal helix α13, providing a molecular view for the α13–α13 distances previously reported for the hGBP1 dimer. Most of the loops of the GTPase domain were found to be flexible, revealing why GTP binding is needed for hGBP1 dimerization to occur.

Author summary

Guanylate binding proteins are key fighters against microbial and viral pathogens. In the human body there are seven types of such proteins, among which is the guanylate binding protein 1 (hGBP1). This protein is able to perform its function only once it is activated by binding and converting guanosine triphosphat (GTP) to guanosine diphosphat and guanosine monophosphat via hydrolysis. In concert with the conversion of GTP the dimerization of hGBP1 occurs, which can further interact with the lipid membrane of the pathogen and disrupt it. While the crystal structure of the protein is known, the activation and dimerization steps are not well understood at molecular level as studying them experimentally is difficult. An alternative approach is given by molecular simulations, allowing us to elucidate the protein dynamics closely connected to these steps. From our simulations applied to both the hGBP1 monomer and dimer we identified large-scale motions taking place in hGBP1 that had not been reported before. We discuss the relevance of

these motions in terms of their biological function, such as possible membrane damage caused by one of the motions or locking the protein in the dimer state.

Introduction

Guanosine triphosphate (GTP) binding proteins play essential roles in many cellular processes responsible for the maintenance and regulation of biological functions. Among these proteins are the guanylate binding proteins (GBPs), which belong to the dynamin-related protein family, even though the GTPase domain is the only conserved sequence. They have various functions in the resistance against intracellular pathogens via GTP binding and hydrolysis [1–5]. Generally, an infection is followed by the production of interferons by leukocytes, monocytes and fibroblasts, leading to transcriptional activation of the interferon-stimulated genes. GBPs belong to the vertebrate specific class of interferon-γ induced effector molecules that combat intracellular bacteria, parasites and viruses [6]. The human guanylate binding protein 1 (hGBP1) was found to be involved in the defense against viruses, in particular against the vesicular stomatitis virus and the encephalomyocarditis virus, and bacteria [7, 8]. hGBP1 was also identified as a marker of various cancer types, such as mammary cancer [9].

hGBP1 is a large, multi-domain GTPase with similar, but quite low nucleotide binding affinities for GTP, GDP (guanosine diphosphate) and GMP (guanosine monophosphate) [10]. It can adopt at least two structural states with different binding affinities to partner proteins. The switch between the two functional states is activated by GTP binding, resulting in an 'active' state (usually the GTP/GDP-bound form) that binds another hGBP1 molecule leading to dimerization or an effector protein for eliciting the desired effect, and a 'silent' state that cannot bind and activate other proteins. The dimerization of hGBP1 occurs through their large GTPase (LG) domains, which stimulates hydrolysis of GTP to GDP and subsequently GMP in two successive cleavage steps [10–14]. The hGBP1 monomer has been shown to be able to also hydrolyze GTP to GDP, but not to GMP [15]. The crystal structure of full-length hGBP1 has been solved in the nucleotide-free, i.e., the apo state (PDB 1DG3) [16] and with the non-hydrolyzable GTP analogue GppNHp bound to it (PDB 1F5N) [17]. In addition, crystal structures of the LG domain monomer with GppNHp (PDB 2BC9) and of the LG domain dimer with GDP/GMP·AlF$_{3/4}$ (PDB 2B8W and 2B92) are available [12].

The hGBP1 structure is divided into three domains as can be seen in Fig 1. The LG domain is the most conserved region from the dynamin family and consists of the first 310 amino acids, structured as an eight-stranded β-sheet with six parallel and two antiparallel strands, which is surrounded by six main helices. The GTP-binding site contains four conserved sequence elements G1–G4: the canonical G1 motif or phosphate-binding loop (called G1-P loop henceforth), the G2/switch 1 motif (G2-SW1), the phosphate- and Mg^{2+}-binding G3/switch 2 motif (G3-SW2), and the nucleotide-specificity providing G4 motif, which is part of a loop and will be called G4-L2 in the following [18]. This G4-L2 loop is preceded by another loop, which we thus denote as L1. Another key structural element of the G domain is the guanine cap (GC), which forms the protein–protein interface in the hGBP1 dimer [12]. The crystal structures of the nucleotide-free and -bound LG domain suggest that the conformation of the GC goes from an open conformation in apo-hGBP1 to a closed conformation upon GTP binding. The different loops along with their residue ranges and residues key for hydrolysis or dimerization are listed in Table 1 and shown in Fig 1. The second domain is the middle (M) domain (amino acids C311–Q480), which is composed of two two-helix bundles, α7/8 and α10/11 that are connected by α9 and extend over a length of 90 Å, giving hGBP1 an elongated
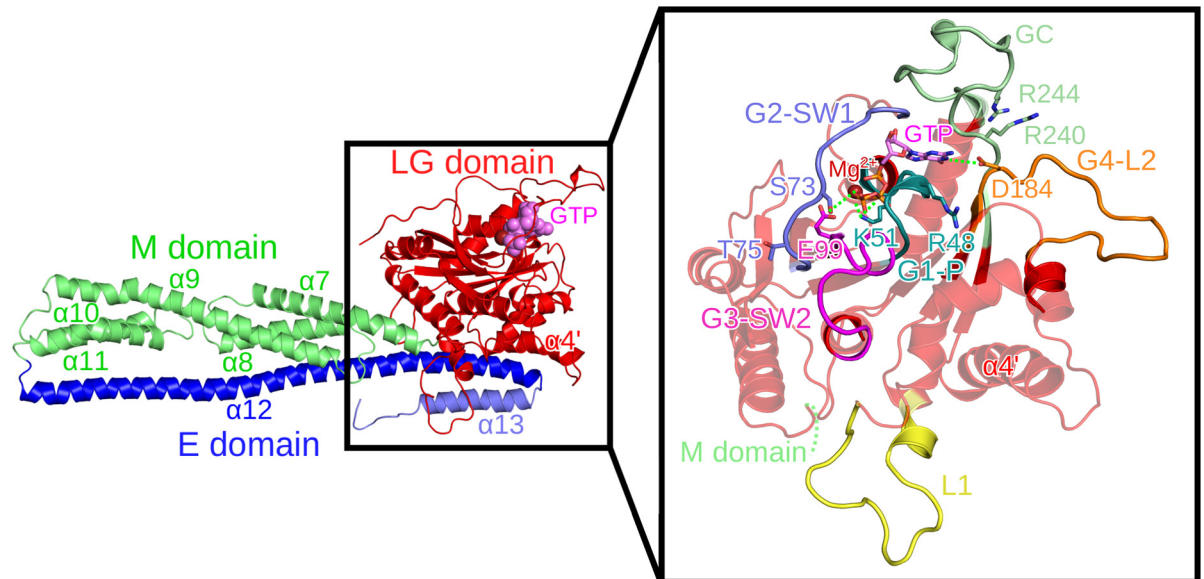
Fig 1. Conformation of the nucleotide-bound hGBP1 crystal structure (PDB 1F5N). The different domains are highlighted in different colors: the LG domain in red with GTP shown in purple, the M domain in green, and the E domain in blue (with different shades used for α12 and α13 for ease of distinction). Loops missing in the crystal structure were modeled. In the enlargement of the LG domain shown on the right, the four GTP-binding site motifs are displayed: the G1-P loop in turquoise, G2-SW1 in blue, G3-SW2 in magenta, and G4-L2 in orange. The guanine cap (green), loop L1 (yellow), and residues important for dimerization or GTP binding and hydrolysis (shown as sticks with the same color as the corresponding loop) are also highlighted. The dotted green lines indicate direct interactions of hGBP1 residues with GTP and Mg^{2+}.

shape. The helical effector (E) domain (amino acids T481–I591) involves a very long helix, α12, that stretches over a length of 120 Å from the tip of the M domain back to the LG domain, where it forms multiple electrostatic contacts with helix α4'. At the C-terminal end of the E domain, there is a helical turn leading to the short helix α13 and the last seven residues, which are unstructured.

The aim of the current work is to elucidate the intrinsic dynamics of apo-hGBP1 and the hGBP1 dimer. Given the considerable size (67 kDA, 592 residues) and elongated shape of hGBP1, it is to be expected that even without nucleotide binding this protein is flexible. A thorough characterization of the conformational dynamics in the apo state is the prerequisite for understanding the changes in structure and dynamics of hGBP1 following GTP binding,

Table 1. Characterization of the loops of the G domain.

(Motif-)Loop	Sequence	Key residues[a]	Flexibility		
			clusters	population[b]	RMSD[c]
G1-P	44–52	K48, K51	23	89.6%	7.7 Å
G2-SW1	58–77	S73, T75	259	40.5%	15.5 Å
G3-SW2	98–110	E99	175	50.6%	14.5 Å
L1	149–172		115	42.4%	9.5 Å
G4-L2	181–198	D184	194	37.1%	13.4 Å
GC	235–261	R240, R244, D255	675	26.4%	22.0 Å

[a] Residues which are important for GTP binding, hydrolysis or hGBP1 dimerization.

[b] Percentage of the structures which are cumulatively represented by the first three clusters.

[c] The largest RMSD found between two loop conformations.

Table 2. List of simulations performed in this work[a].

System	Simulation	Size	T	Runs	Length	Cumulated time
Monomer	Amber99SB*-ILDNP H-REMD	335,553 atoms	310 K	1 × 30 replicas	400 ns per replica	12 μs
E domain (T481–I591)	Amber99SB*-ILDNP MD	573,042 atoms	310 K	1	500 ns	0.5 μs
α11 + E domain (Q456–I591)	Amber99SB*-ILDNP MD	573,050 atoms	310 K	1	500 ns	0.5 μs
Monomer	Martini MD	18,438 particles	310 K	5	63 μs, 75 μs, 75 μs 75 μs, 200 μs	488 μs
Dimer	Martini MD	32,836 particles	310 K	5	23 μs, 74 μs, 75 μs 90 μs, 150 μs	412 μs
Dimer	Martini MD	32,836 particles	320 K	1	270 μs	270 μs
Dimer	Amber99SB*-ILDNP MD	376,899 atoms	310 K	1	100 ns	0.1 μs

[a] All simulations reported in this work were performed on the supercomputer JURECA at the Jülich Supercomputing Centre [67].

https://doi.org/10.1371/journal.pcbi.1007193.t002

hydrolysis and dimerization. To this end, we applied multi-resolution molecular dynamics (MD) simulations to both the hGBP1 monomer and dimer, involving in total 13 μs of sampling at the atomistic scale and 1.1 ms at the coarse-grained level. To elucidate the dominant motions from the large amount of simulation data, we applied state-of-the-art techniques, such as principal component analysis and Markov state modeling. The enhanced MD simulations of the monomer revealed that the monomeric apo form is highly flexible and exhibits a hinge motion that is similar to the motions observed for other dynamin-like proteins. This motion is also present in the hGBP1 dimer, for which a structural model is provided in this study. Other large-scale motions were observed for the C-terminal helix α13, which allows us to explain previously reported experimental data, and for most of the LG domain loops, which provides a rationale why GTP is required for hGBP1 dimerization to occur. Our study provides fundamental insights into the dynamics of both the hGBP1 monomer and dimer with consequences for its function.

Results

To reveal the conformational dynamics of apo-hGBP1 on the microsecond time scale, we performed an all-atom Hamiltonian replica exchange MD simulation (H-REMD) with 30 replicas of 400 ns length per replica (see Table 2 in the Methods for an overview of all simulations performed in this study). It should be noted that the usage of the replica exchange algorithm usually leads to a sampling speed-up of one or two orders of magnitude compared to ordinary MD simulations [19, 20]. To explore the dynamics on the sub-millisecond time scale for apo-hGBP1 and also the hGBP1 dimer, we employed coarse-grained MD simulations using the Martini model [21]. We first present the results from the all-atom H-REMD simulation, followed by the results from the Martini simulations.

Conformational dynamics of the hGBP1 monomer from all-atom simulations

Overall flexibility. We evaluated the flexibility of hGBP1 by calculating the root mean square fluctuations (RMSF) of the C_α atoms. It is found that the LG domain is the most rigid part of hGBP1 while the regions furthest away from it are the most flexible, which can be best
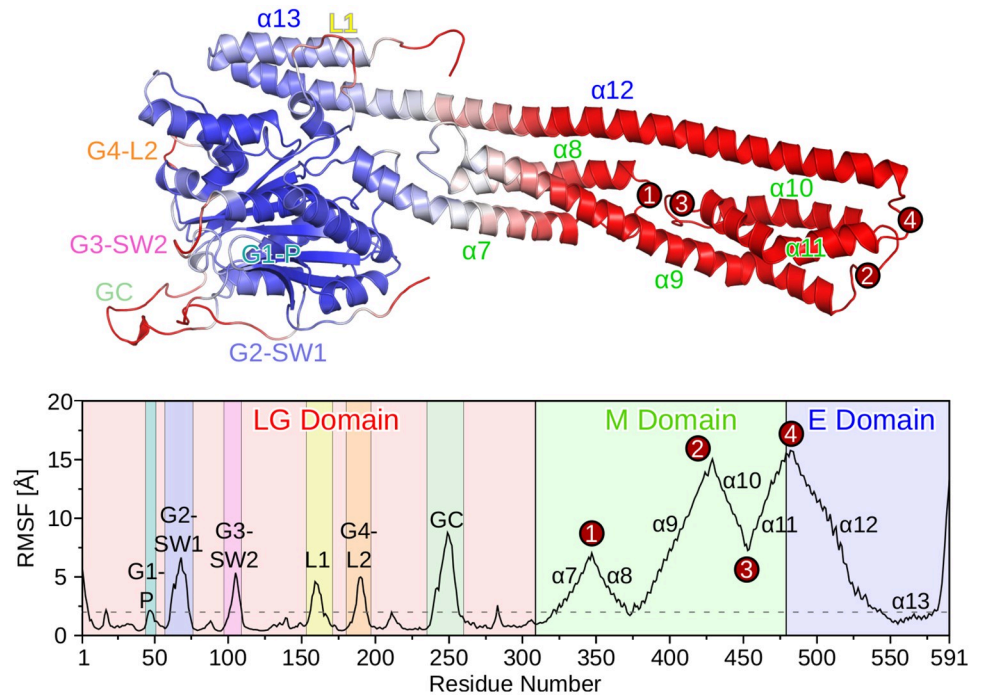
Fig 2. Fluctuations of the hGBP1 residues obtained from the H-REMD simulations. The fluctuations are quantified by the RMSF of the C_α atoms. The top plot shows the protein colored based on the RMSF values, ranging from blue for low RMSF values (minimum at 0.4 Å) to red for high RMSF values (maximum 15.8 Å). Regions of increased flexibility (loops of the LG domain, regions 1–4 in the M or E domain) are marked in the top and bottom plots. The dashed horizontal line at 2 Å helps distinguishing rigid and mobile regions from each other.

seen in the structure plot in Fig 2 where the rigid amino acids are shown in blue and the mobile parts of the protein are colored in red. The most stable regions are present in the LG domain and involve amino acids that belong to the α-helices or the β-sheet of that domain. These amino acids were therefore used for aligning the conformations with respect to the initial conformation before calculating the RMSF. The loops of the LG domain can be easily identified as the regions with increased RMSF values (see also Table 1), which will be discussed in detail below.

The first highly flexible region of the M domain involves amino acids A320–R370, which form the two-helix bundle α7/8. The residues connecting these two helices are the most flexible, resulting in an RMSF peak of ~ 7 Å (marked as region 1 in Fig 2). This two-helix bundle is followed by residues L375–T387, which are quite rigid due to their proximity to the LG domain, and the third and longest helix of the M domain, α9, which is characterized by monotonically increasing RMSF values up to one of the two highest RMSF peaks of ~ 15 Å for residue K429 (region 2 in Fig 2). The next two-helix bundle composed of α10 and α11 has RMSF values between 7 and 15 Å, with the lowest values found for the residues connecting these two helices (region 3 in Fig 2). The RMSF of ~ 7 Å for this turn region is similar to the values for region 2 between α7 and α8, which can be explained by their close spatial proximity. The second RMSF peak of ~ 15 Å corresponds to the transition between domains M and E (region 4 in Fig 2), where the long helix α12 from domain E starts. The RMSF values for the residues of this helix decrease until it comes in contact with the LG domain around K544, where the RMSF has dropped below 2 Å. Since α13 forms several tight interactions with both α12 and the LG domain, this helix was rather rigid in our H-REMD simulation.
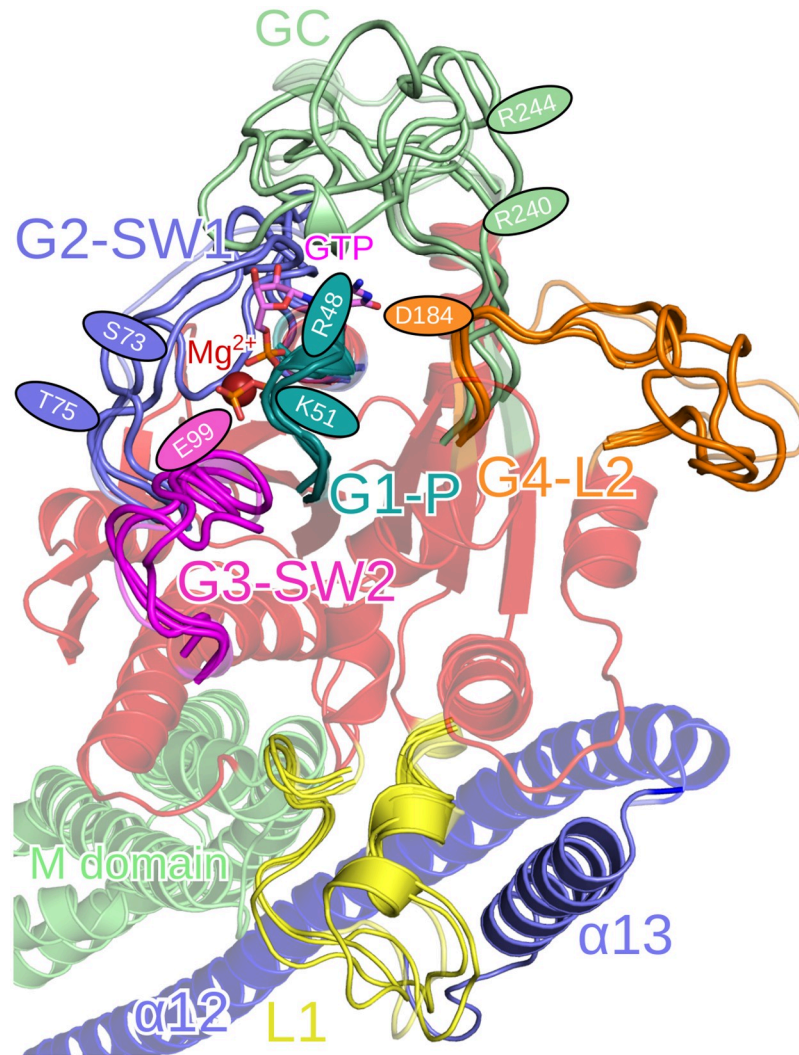
Fig 3. Conformational clusters of the LG domain loops. The LG, M and E domain are colored in transparent red, green and blue, respectively. The colors of the loops are: G1-P in turquoise, G2-SW1 in blue, G3-SW2 in magenta, L1 in yellow, G4-L2 in orange, and the GC in green. The positions of the residues important for GTP binding and hydrolysis or hGBP1 dimerization are indicated by ellipses in the color of the corresponding loop.

https://doi.org/10.1371/journal.pcbi.1007193.g003

Only the last seven C-terminal amino acids were flexible, as were the first four N-terminal amino acids.

Flexible loops in the LG domain. To characterize the dynamics of the LG loops, we clustered the conformations of each loop using a cutoff of 2.5 Å. The results of this analysis are summarized in Table 1 and shown in Fig 3. Compared to the other loops of the LG domain, the G1-P loop is only slightly flexible. The overall number of clusters is low (23), the first three clusters present most of the loop conformations (89.6%), and the RMSD between the two conformations that are furthest apart is only 7.7 Å, which includes side-chain motions as they were considered during the clustering analysis. Residue R48 is in all clusters solvent-exposed and thus in a position that is not compatible with GTP hydrolysis. This is not too surprising as it is known that only after GTP binding—which was not considered in our simulations—followed by dimerization R48 is positioned toward the γ-phosphate of GTP, stimulating the cleavage of this group by stabilizing the transition state of GTP hydrolysis [22]. Interestingly,

K51, which is also crucial for GTPase activity of hGBP1, is not flexible and remains in the same position as during GTP hydrolysis.

One of the most flexible loops is the G2-SW1 loop, for which 259 clusters were identified. It can switch between closed and open conformations with a preference towards the closed state despite the lack of nucleotide in our simulations, which is best visible in S1 Fig in the Supplementary Information showing all cluster conformations for each of the LG domain loops. The two residues S73 and T75, which are important for the hydrolysis reaction [14, 22], are in positions different from the ones in the LG domain dimer. Without GTP, they point away from the GTP-binding site. Thus, they must undergo a reorientation for adopting positions supporting GTP hydrolysis upon GTP binding and hGBP1 dimerization. In contrast, residue E99, which is part of the G3-SW2 loop and also of relevance for GTP hydrolysis by forming a composite base together with S73, a bridging water molecule, and GTP itself enabling the transfer of a proton from the water nucleophile to the GTP phosphoryl oxygen [14], remains in a position in agreement with GTP hydrolysis. The stable orientation of the E99 side chain allows its interaction with the Mg^{2+} ion (which was not present in our simulations), while residues 101–110 of G3-SW2 are flexible. Loop L1, which is not part of the GTP binding site or dimerization interface, is particularly flexible between residues 156 and 165, while the other residues remain in their position. This is understandable as some of these residues belong to an α-helix (see Fig 3), but we nonetheless included them in our analysis because they interact with helix α13 from domain E. We found that the interplay between the L1–α13 interactions and the flexibility of L1 is essential for enabling large-scale motions of α13, which is discussed below. The G4 motif, which contains D184 relevant for GTP hydrolysis, is stable as it is part of a β-strand, while the following loop L2 is very flexible with motions of up to 13.7 Å and 194 clusters in total. The firm position that is found for D184 enables efficient binding of the guanine base of GTP to this residue.

The highest flexibility—with 675 clusters and the first three clusters representing only 26.4% of the conformations—was found for the GC loop relevant for hGBP1 dimerization. Moreover, large conformational changes are possible as the maximum RMSD of 22.0 Å revealed and can be seen in Fig 3. Both open and closed conformations are adopted by the GC with many different conformations between these extremes, which agrees with the findings for other enzymes that loop motion is often not a simple open and shut case [23]. The presence of the closed GC state without GTP being bound and hGBP1 being dimerized further suggests that GTP only helps stabilizing the closed GC state and does not induce a conformational transition from the open to the closed state, as one might assume from the crystal structures [12, 16].

Motions of the M domain and and helix α12. The main structural changes of the M domain and helix α12 of the E domain were identified based on a principal component analysis (PCA) of the H-REMD data (see Methods for details). Helix α13 was excluded from this analysis as the RMSF values had revealed that it did not substantially move. We found that the first two principal components (PCs) describe best the main conformational motions of the M domain and α12 (S2 Fig) and therefore calculated the 2D free energy surface (FES) of hGBP1 along these two PCs (Fig 4). The main conformational change along the PC1 is a kinking motion, were the M domain and helix α12 bend towards the LG domain for negative values and away from it for positive values. The amplitude of this motion is large as the distance between the minimum and maximum values for this motion is above 110 Å, which becomes visible from the conformations corresponding to PC1$_{min}$ and PC1$_{max}$ in Fig 4. The motion along PC2 can be described as a screwing motion where the M domain and α12 rotate with respect to the LG domain, which can be seen by comparing the conformations representing PC2$_{min}$ and PC2$_{max}$. In general, the FES is characterized by one main area which contains the
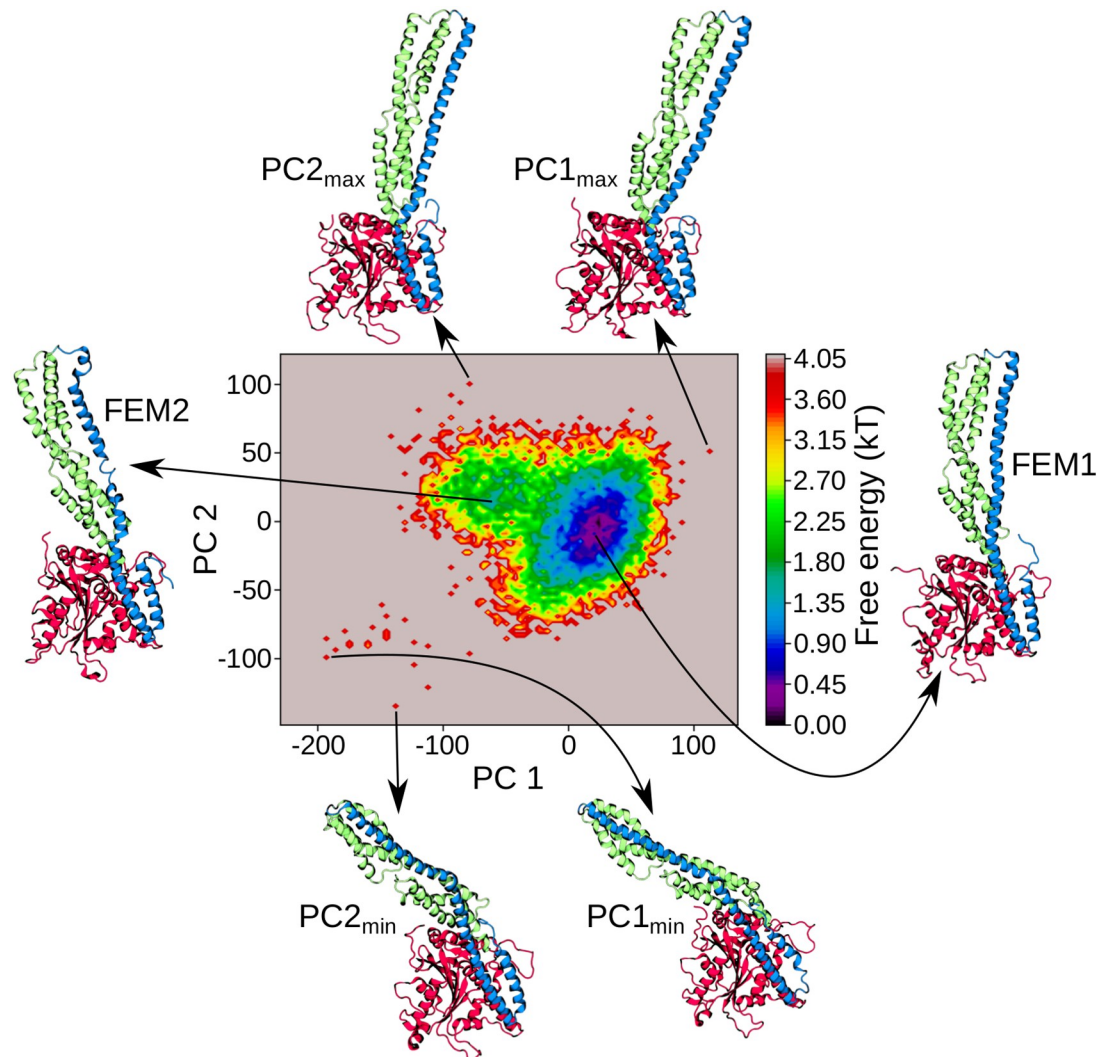
Fig 4. Free energy surface (FES) along the dominant principal components PC1 and PC2. Next to the FES, hGBP1 conformations representing the lowest free energy minimum (FEM1, right) and the shallow free energy minimum (FEM2, left) are shown. Examples of structures for the extreme values along PC1 and PC2 are also provided below ($PC1_{min}$, $PC2_{min}$) and above ($PC1_{max}$, $PC2_{max}$) the FES. The conformations are shown as cartoon and colored based on the three domains: LG (red), M (green) and E (blue).

lowest free energy minimum. Conformations corresponding to the lowest free energy values (shown on the right of Fig 4) have structures that are very similar to the crystal structure [12, 16]. In addition, there is a shallow free energy minimum for negative values of PC1 that corresponds to conformations where the M domain and helix α12 are bent towards the LG domain. Interestingly, in the structure representative for this minimum the long helix α12 from the E domain is broken into two shorter helices. A similar helix fragmentation, which preferentially occurred around residue Q541, is also seen in the structures representing $PC1_{min}$ and $PC2_{min}$.

Despite these large-scale motions of the E domain, it did not detach from the LG or the M domain due to the presence of multiple stable salt bridges formed by the E domain with both other domains (Fig 5A and 5B). The most stable salt bridge was present between K228 from helix α4' of the LG domain with either E568 or E575 of α13, which existed during 43.0% of the simulation time. Other residues with a high likelihood of salt-bridge formation are K582 and
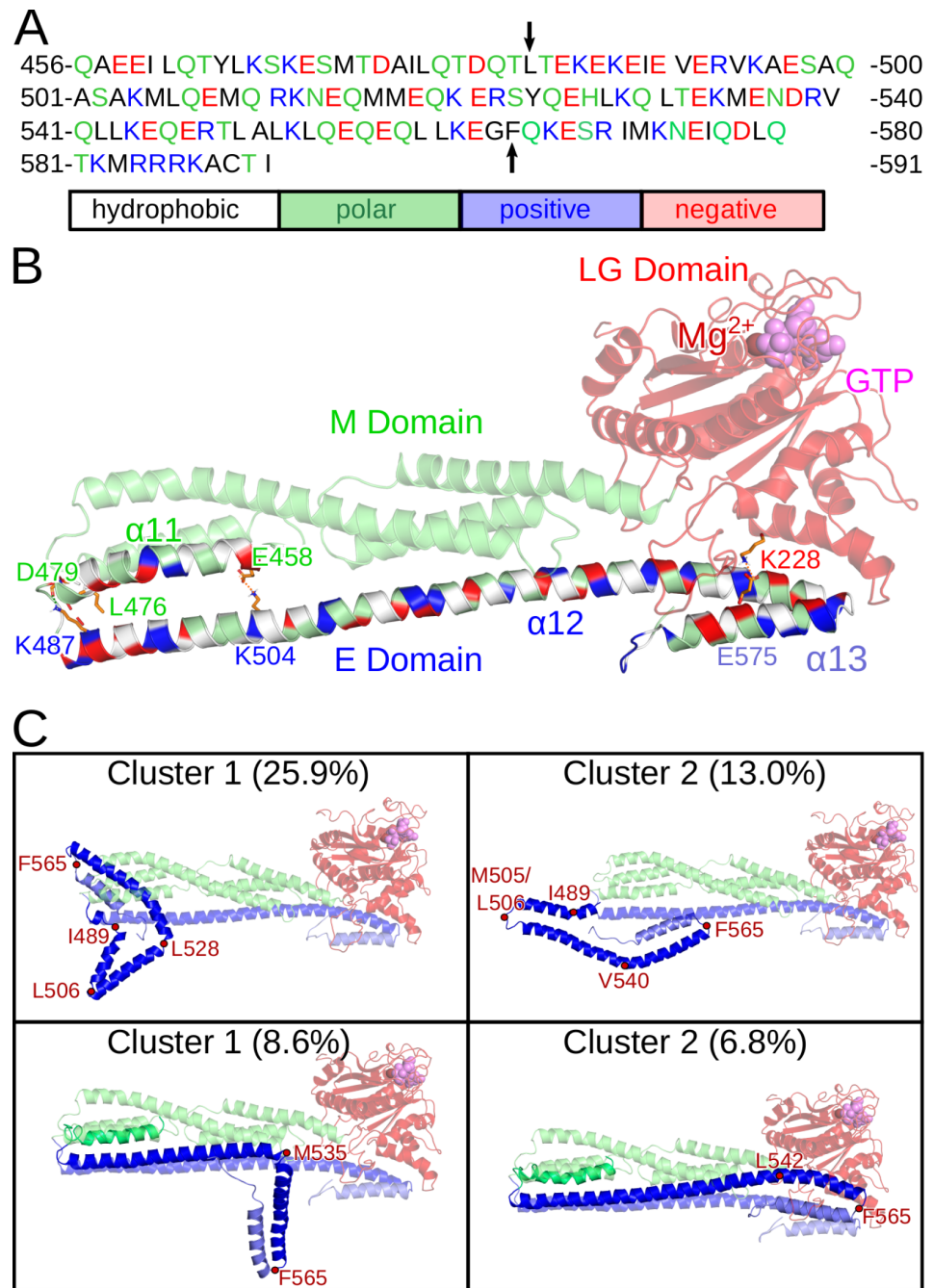
Fig 5. Properties and dynamics of the isolated E domain of hGBP1. (A) The sequence of the E domain and preceding helix $\alpha11$ of the M domain is shown and colored according to the physicochemical properties of the amino acids. The start and end of helix $\alpha12$ at residues L482 and F565 are indicated by vertical arrows. (B) The structure of the E domain and its interactions with the LG domain (red transparent) and the helix $\alpha11$ of the M domain (green transparent) are highlighted for the crystal structure of hGBP1 (PDB 1DG3). The residues of $\alpha11$ and the E domain are colored according to the physicochemical properties of the amino acids: white, hydrophobic; green, polar; blue, positively charged; red, negatively charged. Salt bridges (K228–E575, E458–K504, D479–K487) are indicated with orange dashes and hydrogen bonds (L476/D479–K487) with black dashes. The interacting residues are shown as sticks and labeled. This includes K582 and K587, which form salt bridges with various residues of loop L1 (in yellow) of the LG domain. (C) The two most populated cluster conformations with their occurrence obtained from the MD simulations of the isolated E domain (top) and of $\alpha11$ plus E domain (bottom). The cluster structures were aligned to the crystal structure of full-length hGBP1 (shown as transparent cartoon) using residues 482–484 for the alignment of the isolated E domain and $\alpha11$ for alignment of $\alpha11$ plus E domain. The red points indicate the turns that formed at the different positions of the E domain along with the turn at F565 separating $\alpha12$ and $\alpha13$ from each other.

K587 of α13, which built salt bridges with various residues of the loop L1 of the LG domain, with a probability of 43.0% and 7.1%, respectively. The decrease of the electrostatic interactions as done in our H-REMD simulations did not completely abolish these salt bridges, confirming their stabilities. In the replica with the largest energy bias, residues K228, K582 and K587 formed the same salt bridges as in the target replica with probabilities of 26.2%, 18.0% and 1.3%, respectively, which was still enough to inhibit the detachment of the E domain from the LG domain.

Dynamics of the isolated E domain. Motivated by the dynamics of the E domain observed in our H-REMD simulation and by the hypothesis, derived from size-exclusion chromatography, that hGBP1 can adopt an even more elongated shape than it already has by folding out the E domain (see Fig 7 in [24]), we performed 500 ns MD simulations of the isolated E domain and also of the E domain plus helix α11 of the M domain (see Fig 5A for the sequence). These simulations allow us to judge the stability of the E domain when it is not stabilized by intraprotein interactions with the LG and M domain. While the lack of these tertiary contacts would be the prerequisite for folding out of the E domain, our simulations revealed that the E domain is not stable without them. In the simulation of the isolated E domain helix α12 formed kinks and turns at different positions, especially at the N-terminal end, which in hGBP1 is attached to the M domain. In Fig 5C the two dominant structures as obtained from a clustering analysis of the 500 ns MD simulation are shown (structures for the next three clusters are shown in S3A Fig). We overlaid the E domain conformations onto the crystal structure of hGBP1, which demonstrates that the E domain is unlikely to fold out as an extended helix if it should detach from the LG domain, as suggested from experiments [24]. The two most stable kinks or turns occurred at residues I489 and M505/L506, while the region between Q525 and L542 transformed to turns at various positions, but always reversibly as can be seen in the time-resolved secondary structure plot in S3B Fig. The latter position is the same where α12 reversibly unfolded and kinked during the H-REMD simulation of full-length hGBP1 (Fig 4).

We aimed at understanding the reasons for turn formation that occurred in the isolated α12 helix. This helix is 81 residues long and contains 21 i, $i+3$ or i, $i+4$ combinations of oppositely charged side chains (Fig 5A), which may form salt bridges. While such salt bridges are known to stabilize helices [25, 26], we find that in some cases they can also stabilize turn structures. For instance, the interaction E526–K529–E533, which is present in α12, also stabilizes the turn formed at L528 in the second half of the MD trajectory of the isolated α12 helix. Other salt bridges were newly formed upon turn formation, such as the one between E488 and R493 following the helix → turn transition at I489. In the helix conformation, E488 cannot form a salt bridge with residues 491 or 492 as these are valine and glutamate. Also the turn at M505/ L506 enabled a novel salt bridge to be formed, namely between E508 and R511, which in the helix is not possible as the side chains of these two residues point in opposite directions. Also between E508 and K504 only a weak interaction is possible in α12 due to steric constraints of the helix (the minimum distance between both side chains is above 6 Å), while upon turn formation E508 becomes able to interact with R511 instead (minimum distance below 2 Å). Interestingly, the predicted positions of turn formation coincide with the positions with a reduced α-helical coiled-coil probability of the E domain, which was determined by Syguda et al. [27] using the the COILS program [28]. They performed this analysis to test whether coiled-coil formation may occur between α12 helices. An alternative scenario would be that the isolated E domain forms an intraprotein coiled-coil structure with the first helix extending from L482 to A503 and a second helix from L506 to the end of α12. A third helix with a turn around residue 540 might also be possible, which would be in agreement with both our MD simulations and the COILS predictions. It should be further noted that the helical propensity of the isolated E domain is only slightly reduced upon turn formation. It is 79% averaged over the 500 ns MD

simulation, which is close to the value of 82% extracted from the circular dichroism (CD) spectrum of the E domain [27], especially if one considers the errors involved in fitting CD spectra. In the crystal structure (PDB 1DG3) the helical probability is 86% for α12/α13, while it was 84% in the H-REMD simulation of hGBP1. Thus, the presence of the LG and M domain has stabilizing effects on α12, but the overall helicity is not lost in their absence.

A major stabilizing effect originates from α11 of the M domain, as its inclusion in the simulation of the E domain prevented the formation of turns at residues I489 and M505/L506. This can be explained by attractive interactions between α11 and α12, inhibiting structural changes in this part of α12. Especially residues close to the turns that formed in the isolated E domain are involved in interactions with α11: two stable hydrogen bonds are present between L476–K487 and D479–K487, supported by a salt bridge between E458 and K504 (Fig 5B). Only at different positions between Q525 and L542 reversible turn formation was observed as the cluster structures and evolution of the secondary structures confirm (Fig 5C and S3 Fig). For instance, in the most populated cluster structure a turn at M535 is present, which is facilitated by a novel electrostatic interaction between D538 and R539, while preserving the salt bridge E536–R539 present in α12. In both simulations, i.e., in the simulation of the isolated E domain and in that of α11 plus E domain, α12 tightly interacted with α13 and the last seven C-terminal residues via various salt bridges, which are possible as 11 of the 28 amino acids composing α13 and the C-terminal residues are charged (including the stretch R584-R585-R586-K587) and paired with a similar abundance of positively and negatively charged residues in the oppositely located α12 helix.

In summary, the simulations of full-length hGBP1, the isolated E domain, and α11 plus E domain revealed that a folding out of the E domain seems to be an unlikely event. First, it would require the simultaneous breaking of several hydrogen bonds and salt bridges that the E domain forms with both the LG domain and the M domain. Second, even if such detachment of the E domain should happen, the missing tertiary contacts would provoke turn formation at specific residues of α12 on the nanosecond time scale, as demonstrated by our MD simulations. It should also be mentioned that this helix is by a factor of at least four longer than the optimal length of an isolated helix (which is between 9 and 17 amino acids [29]) and thus needs tertiary contacts as present in hGBP1 for it to be stable. Instead of folding out of α12 as an extended helix, a more likely conformational change might be the formation of a coiled-coil motif following fragmentation of α12 into two or three shorter helices. In all cases studied here, α12 has a high tendency of reversible kinking around residues Q525–L542. Even in full-length hGBP1 this kinking is possible, as also the M domain is very flexible in that area, which is where the two-helix bundles α7/8 and α10/11 meet (marked as regions 1 and 3 in Fig 2). We thus conclude that the M and E domain feature a hinge in that area.

Long time-scale dynamics of the hGBP1 monomer and dimer from coarse-grained simulations

Monomer dynamics. To explore the conformational flexibility of the apo-hGBP1 monomer on the micro- to millisecond time scale, we performed five continuous MD simulations of lengths between 63 μs to 200 μs using the coarse-grained Martini force field (see Table 2). As in the all-atom simulations we observed the kinking motion of the M and E domain in all five coarse-grained simulations. Interestingly, this motion was possible despite the fact that the Martini model preserves the initial secondary structure by applying an elastic network model. Therefore, α12 remained fully helical in these simulations. It further demonstrates that the application of Martini without elastic networks to preserve the tertiary protein structure allows the sampling of conformational changes that also occur in atomistic simulations.
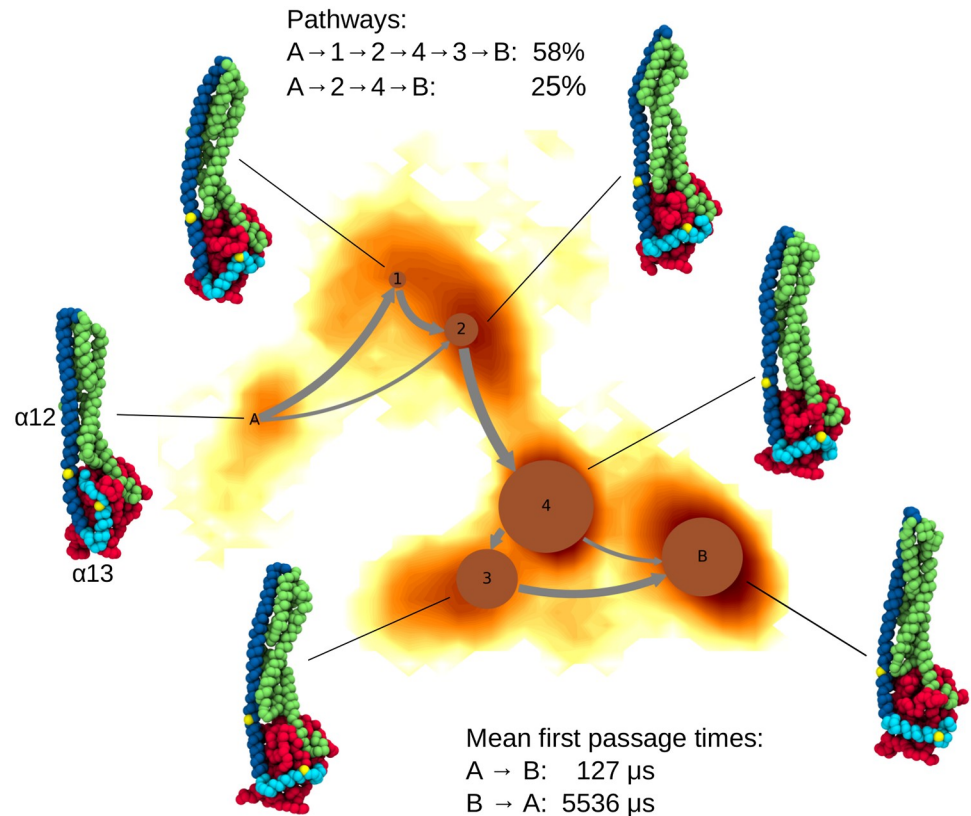
Fig 6. A network diagram of the Markov state model obtained for the motion of helix α13 as sampled in the Martini simulations. The MSM is overlaid onto the FES along the first two TICs. The circles represent the stable states, where the area of the circles correlates with the population of the corresponding state. The arrows indicate transitions between the states, with the line thickness correlating with the transition fluxes. The FES is colored as gradient from yellow (high free energies) to dark red (low free energies). A representative conformation is shown for each state and colored based on the three domains, while the yellow dots indicate residues Q541 and T581, which are the start and end of the sequence for which the MSM was calculated.

Another substantial motion observed in the Martini simulations was a change in orientation of helix α13 by $\sim 90°$ with respect to helix α12. To describe this motion we applied a Markov state model (MSM) analysis to amino acids Q541–T581, which contain the last third of α12 and the helix α13 (see Methods and S4 and S5 Figs for details of this analysis). We identified six metastable states, for which representative conformations are displayed in Fig 6 together with the MSM overlaid onto the FES along the first two time-lagged independent components (TICs). Since all simulations started from the crystal structure, we labeled the Markov state that had $\sim 40\%$ of the conformations with helix α13 close to helix α12 and which are thus similar to the crystal structure with A. The state corresponding to the largest displacement of α13 is denoted as state B, while the other four Markov states (labeled 1–4) are intermediate states between A and B. We calculated the reactive fluxes between the states, which are shown as gray arrows in Fig 6, and determined the mean first passage time (MFPT) needed for the system to go from state A to state B, obtaining a value of 127 μs. The MFPT for the reverse transition from B to A is 5,536 μs, which is consistent with the observation that the complete reverse motion of helix α13 to the conformation associated with the crystal structure has not occurred in our simulations. It should be noted that the MFPTs reported here are obtained from coarse-grained simulations and might not be equivalent to actual MFPTs. Generally,

events observed in simulations with the Martini force field are between 3 to 5 time faster than similar events observed in atomistic simulations [21]. The main transition pathway from A to B goes via states 1, 2, 4, and 3 with a probability of 58%. The next most significant transition pathway with a probability of 25% involves only two intermediate states, 2 and 4.

The mobility of helix $\alpha13$ on the surface of the LG domain requires its motion beyond the loop L1 from the LG domain. Our H-REMD simulations had revealed that especially K582 and K587 have a high tendency to form salt bridges with different residues from that loop (D159, E160, E162, E164, D167, D170, see Fig 5B). In order to identify the molecular interactions enabling the motion of $\alpha13$ beyond this loop, we investigated the interactions between the C-terminal region F565–T590, which includes $\alpha13$ (F565–M583), and L1 by calculating an average distance map during the simulation interval when this motion occurred. The distance map (S6A Fig) revealed that the strongest contacts are between the sequence [580]QTKMRRRKA[588] of the C-terminus and the sequence [164]EVEDSAD[170] from L1. This indicates that the motion of $\alpha13$ is facilitated by electrostatic interactions in interplay with the high mobility of the loop L1 (see Fig 3), allowing the C-terminal region to move beyond L1 and then further on the surface of the LG domain. We also investigated the cause for the high stability of $\alpha13$ after it has moved beyond L1, yielding Markov states 3, 4 and B, and a very large MFPT for the reverse transition from B to A. To this end, we analyzed the interactions between the C-terminal region F565–T590 and the LG domain for conformations belonging to Markov state 4, which has the largest population. The average distance map between the C-terminal region and the interacting LG domain region is shown in S6B Fig, with two notable interactions involving helices from the LG domain being highlighted. In the first case, sequence Y144–K155 belonging to helix $\alpha3$ is in contact with the second part of $\alpha13$ (I576–M583) and the following, flexible residues R584–A588. The strongest contacts are formed between polar residues: E146 and T149–R151 from the LG domain and Q580–T581 from $\alpha13$. The second interaction hot spot involves sequence N220–F230, i.e., half of helix $\alpha4'$ that extends from S213 to F230, which is in contact with residues S569–Q580 from the first part of $\alpha13$. Here, the strongest contacts are mainly of hydrophobic nature, involving L224 and C225 from the LG domain and M572 and I576 from $\alpha13$. In addition, helix $\alpha13$ has residue L579 exposed to the LG domain, constituting another hydrophobic interaction. Thus, we conclude that the initial motion of $\alpha13$ is driven by the coaction of the electrostatic interactions between $\alpha13$ and L1 and the flexibility of L1, and once the helix has moved beyond this loop, it is stabilized by hydrophobic and polar interactions with helices $\alpha3$ and $\alpha4'$ from the LG domain.

Dimer model from coarse-grained simulations. The hGBP1 dimer is formed between GTP/GDP-bound monomers and is thought to be the biologically active state of the protein [13]. Starting from the crystal structure for the LG domain dimer of hGBP1 (PDB 2B92) [12] and the crystal structure of the hGBP1 monomer (PDB 1DG3) [16] we built a complete hGBP1 dimer model as described in the Methods section. We converted the atomistic conformation to a coarse-grained Martini model and performed five simulations of different time lengths (between 23 μs and 150 μs, see Table 2). In addition to restraining the β-sheet and the helices of the LG domain to their original positions (see Methods), the residues of the G2-SW1 and GC loops were also restrained. They thus remained in a position of the transition state during GTP hydrolysis as GDP·AlF$_3$, which is present in the PDB structure 2B92, is a transition-state mimic. Our Martini simulations of the hGBP1 dimer can thus be considered to represent the nucleotide-bound state even though the nucleotide was not present during the simulations.

As for the monomer, we observed the kinking motion of the M and E domain in all five dimer simulations. However, the change in orientation of helix $\alpha13$ as seen in the coarse-grained simulations of the monomer did not occur. It remained in close proximity to helix
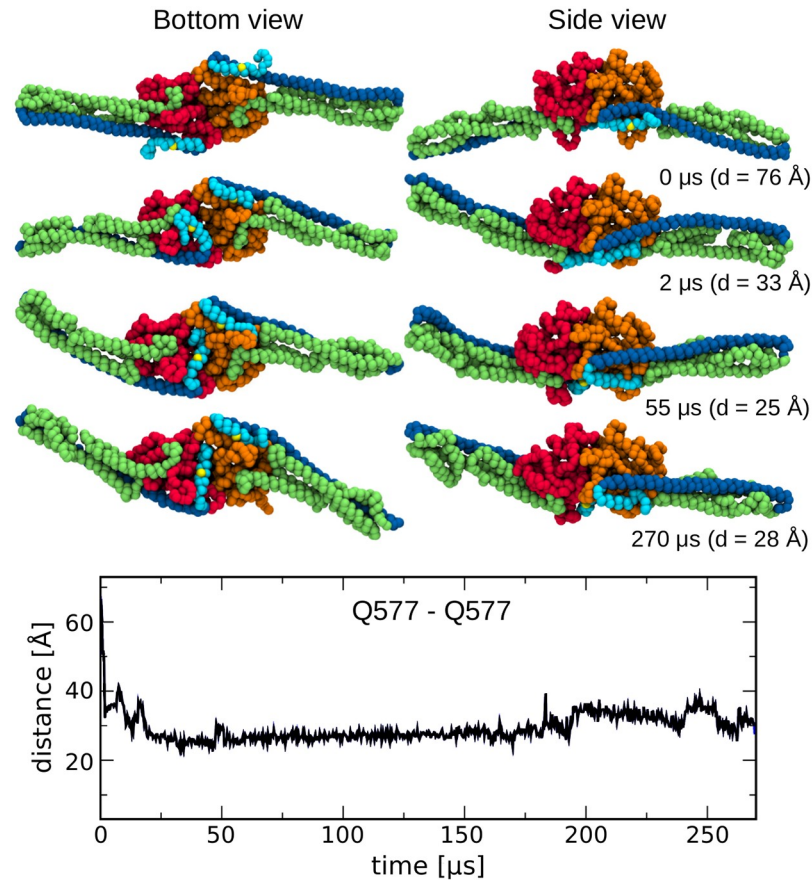
Fig 7. Motion of helix α13 in the hGBP1 dimer. This motion is monitored by the time evolution of the distance between the two residues Q577 from α13 of the two proteins composing the dimer. Representative snapshots associated with different times are also shown. The two LG domains are shown in red and orange, respectively, for ease of distinction, while the M domains are shown in green and the E domains in blue. The two Q577 residues are indicated by yellow spheres.

https://doi.org/10.1371/journal.pcbi.1007193.g007

α12. Only when the temperature was increased to 320 K, as done for one simulation of 270 μs length (see Table 2), we observed this conformational change, but only in one of the proteins composing the hGBP1 dimer. This indicates that the energy barrier for this motion must be higher than in the monomer. To find the cause for this, we analyzed whether the interactions between α13 and loop L1 or whether the mobility of that loop would be different in the dimer than in the monomer. Though both cases did not apply. Instead we found that a new salt bridge between α4' and α13 involving residues E217 and K567 is formed in the dimer, which became possible by the somewhat different position that α4' adopts in the dimer than in the monomer: it is by $\sim$6 Å further away from the core of the LG domain, bringing this helix closer to α13 (S7A Fig). In each of the dimer simulations at 310 K this salt bridge stayed intact, thereby preventing the 90° motion of α13. At 320 K, on the other hand, in the protein of the dimer with the flexible α13 helix this salt bridge broke (S8 Fig).

To describe the motion of helix α13 that was observed in the simulation at 320 K, we monitored the distance between the residues Q577 of α13 from the two proteins (Fig 7). As can be seen from the time evolution of the distance and the snapshots associated with different times, α13 in one of the proteins (the one with the LG domain shown in red) has moved beyond the corresponding loop L1 within 2 μs, and after 55 μs it has adopted a position similar to the one

in state B identified in the monomer simulation (Fig 6), where it stayed until the end of the 270 μs simulation. Helix α13 from the other protein of the dimer remained in close contact to α12 throughout the whole simulation. Nevertheless, as shown in Fig 7, the distance between the two Q577 residues decreased from its initial value of 76 Å to values of $\sim$30 Å, which is very similar to the values between 22 and 35 Å reported from double electron–electron resonance (DEER) and Förster resonance energy transfer (FRET) studies of the hGBP1 dimer [30], keeping in mind the fact that the experimental distances refer to distances between DEER spin labels and FRET dye labels, respectively. The more important aspect is that both experimental techniques predict a change in the Q577–Q577 distance by 50% (FRET) or even more (DEER) upon dimer formation compared to the distance that one obtains from the dimer model that is built based on the crystal structure of the hGBP1 monomer. For a more quantitative comparison to the distances obtained from experiment it would be necessary to simulate the different dimer conformations at the atomistic level while accounting for the conformational freedom of the DEER and FRET labels, respectively. This can be achieved using implicit or explicit label approaches. The implicit approach estimates DEER distances based on the different side-chain orientations, which can be extracted from all-atom MD simulations, and adding the spin label sizes to the side chains [30]. The more accurate approach would be to explicitly model the spin labels, which allows the calculation of electron spin resonance spectra from MD trajectories and thus a direct comparison to DEER experiments [31]. The same holds true for determining FRET distances from MD simulations, which can be achieved by explicitly modeling the FRET labels to enable the calculation of FRET distances and even FRET transfer efficiencies [32]. A computationally less expensive, but implicit approach would be to estimate the accessible volume of the FRET dyes, i.e., the sterically allowed space of the dye molecules attached to the protein, which can be calculated with the FRET Positioning and Screening (FPS) program yielding FRET distances [33]. While these approaches, in particular the explicit consideration of the DEER and FRET labels would allow a more quantitative comparison to experiment, they would require further simulations. Explicit modeling of the labels requires a particularly careful force field parameterization in order to yield trustworthy results [31, 32], while such an extra effort would not change our conclusion: it is not necessary for α12 and α13 to detach from the LG domain in order for the two α13 helices coming into contact with each other. The motion of α13 as found by our simulations can also explain the distances reported by DEER and FRET experiments.

Stability of the dimer with rotated α13 at the atomistic level. The findings from the coarse-grained simulations should be carefully verified given the uncertainties resulting from the neglect of degrees of freedom. Moreover, in the case of Martini one usually preserves the initial conformation of the protein(s) under study by applying the elastic network ELNEDIN [34]. However, we did not follow this approach as we were interested in relative motions within hGBP1 and therefore only applied secondary structure restraints during the Martini simulations. This alternative approach was sufficient to maintain the contacts between the different domains (LG, M and E) during the simulations. Furthermore, a great similarity between the M and E domain motions in the atomistic and Martini simulations was observed, adding to the credibility of the results produced by the Martini simulations. The only unexpected motion was the one of α13. We thus tested the stability of the dimer with α13 being rotated in one of the hGBP1 molecules (Fig 7) at the atomistic level. To this end, we converted the final conformation sampled in the Martini simulation at 320 K to the atomistic level (S9A Fig) and simulated it for 100 ns using the same MD approach as applied to the truncated hGBP1 models as described above. In this simulation we explicitly included GTP and the co-factor Mg^{2+} in both LG domains. This removes another uncertainty of our Martini simulations where the nucleotide was not simulated, only its effect on the LG domains was modeled by applying

restraints to the GTP binding sites. All details regarding the backmapping and GTP parameters are provided in the Methods section.

The results of this all-atom simulation are summarized in S9 Fig. The comparison of the dimer converted to the atomistic level to the initial dimer model used for starting the Martini simulations (S9B Fig) shows that in addition to α13 in one of the hGBP1 molecules (called M1 in the following, while the other hGBP1 molecule without rotated α13 is denoted M2) also the M and E domain of both proteins had moved considerably in the prior Martini simulation. They continued to do so during the all-atom simulation as the snapshots sampled at 50 and 100 ns (S9C and S9D Fig) and also the RMSF projected onto the dimer structure (S9E Fig) confirm. The presence of GTP does not inhibit the kinking motion observed before, as the most mobile part of hGBP1 still is at the junction between the M and E domain, exhibiting motions of up to 55 Å in M1 and 38 Å in M2 based on the distance of L482 (i.e., the start of α12, see Fig 5) with respect to the start conformation of this all-atom simulation. It should be emphasized that the backmapped dimer is of similar stability as the hGBP1 monomer modeled at the atomistic level, indicating that the prior simulation of the dimer at the coarse-grained level did not introduce artifacts or instabilities into the structure. This also applies to the rotated α13 helix in M1, which remains folded as helix in its rotated position throughout the 100 ns MD simulation, yielding RMSF values below 2 Å. This rotated position is stabilized by interactions of α13 with both α4' and L1 of the LG domain of M1, which are the same interactions as observed in the Martini simulation (S6 Fig). In contrast, α13 of M2 exhibits larger motions and reorients, giving rise to RMSF values above 2 Å in this helix and also causing motions in α12, conformational changes in the turn between both helices and the C-terminal residues following α13. This can be best seen in S9C and S9D Fig. It seems as if α13 of M2 aims to adopt a rotated position as in M1, which, however, caused the distance between between residues Q577 of M1 and M2 to increase from 26 Å in the start conformation to ∼36 Å within the 100 ns simulation (S9F Fig). From the coarse-grained simulations we know that a simulation at the (sub-)millisecond time scale would be required for α13 to complete the motion towards the fully rotated position. This motion requires the breaking and formation of several intraprotein interactions. The main interaction that still stabilizes M2-α13 in its current position is formed with M2-L1, which in turn caused this loop to become rather rigid. The interaction with M2-α4', on the other hand, became already weakened and mainly involves hydrogen bonds and not salt bridges.

In general, the presence of GTP and Mg^{2+} as well as the dimerization caused the loops of both LG domains to be considerable less flexible than in the apo-hGBP1 monomer. We performed the same kind of loop clustering as for the monomer, leading to only one to three clusters for G1-P, G2-SW1, G3-SW3, and G4-L2 in both proteins (S9G Fig). While these numbers are considerably smaller than those found for the apo-hGBP1 monomer (Table 1), it should be noted that these numbers are not directly comparable as the current numbers are obtained from a 100 ns standard MD simulation, while the numbers in Table 1 resulted from an H-REMD simulation involving 400 ns per replica. Nonetheless, it is safe to conclude that GTP, Mg^{2+} and the dimerization rigidify the LG domain, which includes the GC loops for which only 6 and 7 clusters were found for M1 and M2, respectively. Very importantly, all of the loops adopted closed conformations in the atomistic dimer simulation, a prerequisite for the hydrolysis of GTP to take place. Moreover, GTP and Mg^{2+} stayed in their binding sites with no distance restraints applied to them. Their positions are maintained due to interactions with residues that are key for the hydrolysis reaction, including R48, K51 and T75, which in turn stabilized the loops these residues are part of. The GC loops are further rigidified by interprotein interactions as the dimer interface involves the guanine caps. Here, a salt bridge formed

between R244 of M1 and D259 of M2 appeared to be particularly stable as it was present throughout the 100 ns trajectory.

Discussion

We studied the conformational dynamics of the nucleotide-free hGBP1 monomer and the hGBP1 dimer using multi-resolution and enhanced MD simulations on the micro- to milli-second time scale. As expected from its highly conserved sequence in the dynamin family, the LG domain is overall very stable in our atomistic H-REMD simulation. The residues involved in the β-sheet and the adjacent α-helices have all fluctuations of less than 2 Å, while apart from one case the LG domain loops were found to be flexible (Fig 3 and Table 1). The less flexible loop is the phosphate-binding loop G1-P and involves the first of the four motifs G1–G4 that are important for the hydrolysis reaction. These motifs include six residues (R48, K51, S73, T75, E99 and D184) that are directly or indirectly involved in GTP binding and hydrolysis. Even without nucleotide being present, K51, E99 and D184 adopt the same and stable orientations as in the crystal structures of hGBP1 and the LG domain with nucleotide being bound, while the other three residues are flexible without the stabilizing interactions with GTP. This suggests that the active site of the GTPase domain is quite flexible compared to those of other enzymes [35] and only partially preorganized prior substrate binding [36, 37], which might explain the rather low binding affinity of hGBP1 for GTP ($K_m = 470 \mu$M) [10]. The LG domain loop with the highest flexibility is the guanine cap (GC), which forms the protein–protein interface in the hGBP1 dimer. Even without nucleotide, the GC can adopt both open and closed conformations and rapidly switch between them. This finding indicates that GTP binding would shift the equilibrium toward the closed GC state, which in turn would facilitate hGBP1 dimer formation via the GC–GC interface requiring a certain stability of the protein recognition motif. Our atomistic simulation of the hGBP1 dimer including GTP and Mg^{2+} revealed that dimer formation and the presence of the nucleotide stabilize the closed loop conformations, including GC, and and also the orientations of R48, S73 and T75 in positions supporting GTP hydrolysis, explaining why the hGBP1 dimer is better able to hydrolyze GTP than the hGBP1 monomer is. However, it should be noted that it is not fully clear yet how much of the loop stabilization derives from dimerization and to what extent it is due to GTP and Mg^{2+}. To this end, the hGBP1 monomer with GTP and Mg^{2+} needs to be simulated too, which will be addressed in our future studies.

One of the main results from our atomistic and coarse-grained simulations is the highly flexible nature of the M domain and the long helix α12 from domain E. The region of highest structural flexibility was found at the middle of the M and E domain by PCA (Fig 4), giving rise to large-scale kinking and screwing motions performed by both domains, which is evident from large RMSF values of about 15 Å at the tip of both domains (Fig 2). During these motions, the E domain remains tethered at its ends to both the M and the LG domain by several salt bridges, while the middle of the long helix α12 can reversibly unfold and fold, allowing its kinking and screwing. This finding is further supported by the MD simulations of the isolated E domain (with and without α11 from the M domain), which also revealed reversible turn formation accompanied with local unfolding between residues Q525 and L542, while the contacts to the M domain were found to be vital for overall stability of the long helix α12. Without these tertiary interactions, α12 unfolds on the nanosecond time scale, making it unlikely that the $\sim$120 Å long E domain folds out as intact helix [24], which in addition would require the combined breaking of several hydrogen bonds and salt bridges that the E domain forms with the LG and M domain.
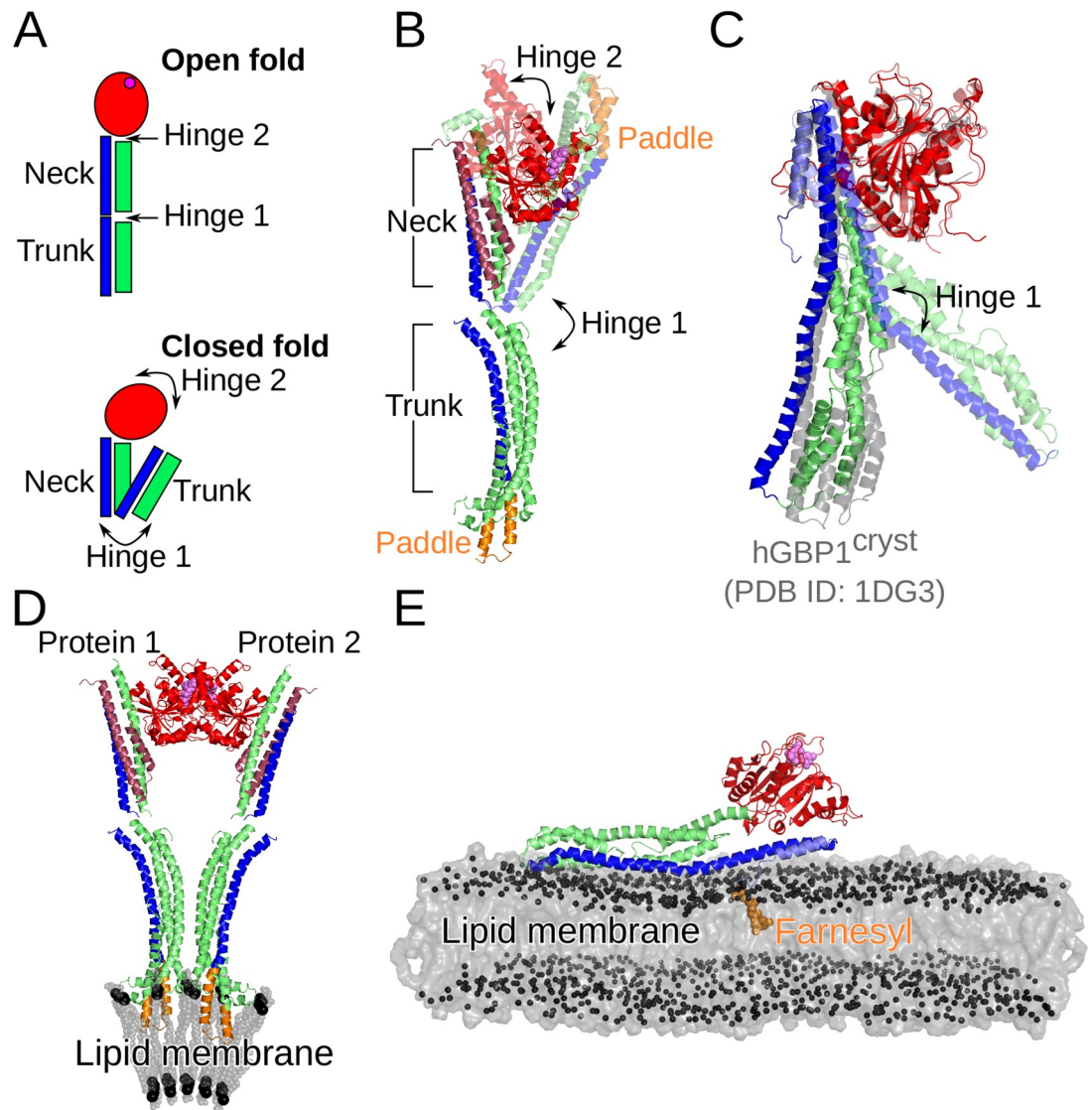
Fig 8. Conformational transitions of hGBP1 and a bacterial dynamin-like protein (BDLP) of *Nostoc punctiforme*. (A) Schematic providing nomenclature for the open and closed BDLP conformations with large-scale transitions between them mediated by hinges 1 and 2 at the trunk/neck and neck/G-domain interfaces. As for hGBP1, the GTPase domain is colored in red with the position of the nucleotide in the open conformation indicated by a magenta sphere, while the 4-helix bundle composing the neck and trunk is colored in green and blue. (B) The closed conformation of BDLP was solved by X-ray crystallography (PDB 2J69, shown in transparent cartoon), while the open conformation was determined by cryo-electron microscopy (PDB 2W6D, shown in opaque cartoon). The two helices preceding the G domain are ruby-colored. (C) From our MD simulations of hGBP1 we observed a kinking of the M and E domain resembling the initial stages of the closed ↔ open transition of BDLP around hinge1, as the overlay of the structures representing this motion (i.e., FEM1 and PC1$_{min}$ in Fig 4) shows. For comparison the crystal structure of hGBP1 (PDB 1DG3) is also shown (in gray). (D) Membrane binding of BDLP occurs following nucleotide binding via the paddle domain (shown in orange) in the open conformation and aggregated BDLP state. Aggregation involves association of the G domains. (E) Membrane binding of hGBP1 also requires nucleotide binding and is facilitated via farnesylation (shown in orange) at the C-terminus acting as lipid anchor.

An alternative scenario is a motion similar to that of the bacterial dynamin-like protein (BDLP) from *Nostic punctiforme* (see Fig 8A for a schematic of this motion), which was shown to exist in a closed and extended conformation (PDB 2J69 and 2W6D, respectively) [38, 39]. In Fig 8B the closed and extended conformations of BDLP can be seen, along with the most

stable hGBP1 structure and the one with the maximal motion of the M and E domain sampled in our H-REMD simulations (corresponding to FEM1 and PC1$_{min}$ in Fig 4) in Fig 8C. The motions of BDLP are facilitated by two hinge regions. Hinge 1 separates the long tail of BDLP into a neck and trunk region, while hinge 2 is at the interface between the neck and G domain. The transition between these two BDLP conformations, which occurs upon nucleotide and lipid binding, involves a 135˚ kinking between the neck and trunk around hinge 1 and a 75˚ rotation of the G domain around hinge 2 [38]. In hGBP1 we identified hinge 1 in the region encompassing residues Q525–L542 of the E domain and the meeting point between the helix-bundles α7/8 and α10/11 of the M domain. It enables a kinking motion of these two domains and involves a (reversible) unfolding of α12, dividing it into a short helix close to the LG domain, which would correspond to the neck of BDLP, and a longer helix corresponding to the trunk of BDLP. The presence of a hinge 2 in hGBP1 remains to be shown. Another similarity between BDLP and hGBP1 is that the tip of the trunk of BDLP is the region of highest flexibility.

It should be noted that there are also certain differences between the two proteins. First, the closed state is the preferred conformation of BDLP in its apo form [38], while for hGBP1 the crystal structures show that the extended conformation exists for both the nucleotide-free and -bound form. In fact, at the moment there is no experimental evidence for a closed form of hGBP1. On the other hand, BDLP is not the only dynamin-like protein for which a hinge 1 motion has been revealed. A recent FRET study combined with X-ray crystallography of the human myxovirus resistance protein 1 (MxA) revealed that also this dynamin-like protein can adopt a closed conformation in addition to the open crystal structure [40]. In contrast to BDLP, the open MxA conformation is preferred in its nucleotide-free state, while adding of GTP shifts the equilibrium towards the closed state. The findings for MxA and our results in addition to the structures for BDLP raise the likelihood for a general existence of a closed state for dynamin-like proteins, which should be addressed in the future.

Another difference between hGBP1 and BDLP is the presence of a paddle domain in BDLP at the trunk tip via which membrane binding is facilitated, while hGBP1 binds to the membrane following farnesylation at the C-terminus (see Fig 8D and 8E, respectively). This suggests different membrane binding mechanisms for both proteins. Moreover, also the multi-merization seems to be facilitated via different domain interactions in BDLP and hGBP1. The dimer structure of the LG domain obtained by X-ray crystallography [12] implies an elongated geometry with the M and E domain pointing away from the dimer interface (Fig 7), leading to a slightly curved dimer which might induce membrane bending following membrane binding. Such membrane destabilization might be even more likely as a result of the kinking motions of the M and E domain observed in our study. In contrast, multimerization of membrane-bound BDLP is thought to occur via interactions between neighboring neck and trunk helices and a G domain dimer interphase different to that seen for the LG domain dimer of hGBP1 (Fig 8D). Our future simulations will address hGBP1 motions following GTP binding and hydroly-sis, its multimerization and lipid binding, which will further highlight possible differences and similarities to other dynamin-like proteins, such as BDLP and MxA.

In the coarse-grained simulations of both the hGBP1 monomer and dimer, an important conformational change observed was the change in orientation of the short helix α13 of the E domain. Markov state modeling indicated an MFPT for the complete α13 motion of $\sim$ 127 μs, while the reverse motion is more than an order of magnitude less likely. The slow time scale of the α13 motion also explains why it was not observed in our atomistic simulation, even though an enhanced sampling scheme was employed. The motion of the helix α13 has clear implica-tions for the dimerization process. A recent study has suggested that, besides the LG domain interface, the dimerization of hGBP1 also involves an interface between the two helices α13

[30]. The motion of α13 observed in our simulations leads to an increased proximity of the two helices in the dimer (Fig 7). While the distance between the two α13 helices that we monitored already agrees with experimental findings, despite only one of the two helices having moved, it is likely that both α13 helices adopt the 90˚ rotated position in the dimer. Moreover, it is not implausible that dimers are preferentially formed by monomers which already have both α13 helices rotated, as the α13 motion is more likely to occur in the monomer than in the dimer. This is due to a newly formed salt bridge between α4' and α13 in the dimer, which increases the energy barrier for the motion of α13. The rotated α13 helices would form a protein–protein interface, in addition to the LG domain interface involving the two guanine caps, which would further stabilize the hGBP1 dimer, as had already been suggested by Herrmann and co-workers [27, 30]. However, based on their DEER and FRET findings they proposed that α12 and α13 detach from the LG domain in order to allow for the two α13 helices coming into contact with each other [30]. Our simulation results demonstrate that such a detachment is not necessary to explain their experimental observations. In fact, such detachment would hinder the membrane binding of hGBP1 as our initial simulations of membrane-bound hGBP1 indicate (Fig 8E).

In summary, to understand the conformational flexibility of hGBP1 and its implication for the dimerization process, we used multi-resolution MD simulations in explicit solvent combined with PCA and MSM analysis. Our results indicate a hinge at the middle of the M and E domain leading to large-scale, dynamin-like motions, and highly flexible loops in the LG domain that open and close the nucleotide binding pocket without a nucleotide being present. We have further observed, for the first time to our knowledge, the change in orientation of helix α13 on a time scale of hundreds of μs with direct implications for the dimerization of hGBP1. One possible scenario is that monomers that already have the helix α13 oriented away from helix α12 form a dimer where the two helices α13 are close enough to form an interface. Thus, the hGBP1 dimer, with interfaces between the LG domains and the helices α13, would be able to insert into a lipid membrane and, in combination with the motions of the M and E domain observed here, lead to the disruption of the membrane and so to the biological function of hGBP1.

Methods

All software and web databases used in this work are listed in S1 Table. The input needed by the various software and output files created are described in a README file in the Supporting Information.

All-atom simulations

For all atomistic simulations of this study the Amber99SB*-ILDNP force field [41–43] combined with the explicit water model TIP3P [44] were employed. Electrostatic interactions were treated with the particle-mesh Ewald method [45, 46] in conjunction with periodic boundary conditions and a real-space cutoff of 12 Å. The Lennard-Jones interactions were cut at 12 Å. A leapfrog stochastic dynamics integrator was used for the integration of equations of motion. The LINCS algorithm [47] was used to constrain all bond lengths and the hydrogen atoms were treated as virtual interaction sites, permitting an integration time step of 4 fs while maintaining energy conservation [48].

The crystal structure of the hGBP1 monomer in ligand-free form with PDB code 1DG3 was used as starting conformation [16]. Missing amino acids from loops in the crystal structure were added with the software ModLoop [49, 50]. The final conformation was placed in a dodecahedral box with 12 Å between the protein and the box, solvated with 108,406 water

molecules and 7 Na$^+$ ions were added for charge neutrality, resulting in a system with a total number of 335,553 atoms. This particular simulation box was large enough to allow free translation and rotation of the hGBP1 protein without interacting with its periodic images that would otherwise result in simulation artifacts. After energy minimization and equilibration of the system following the same procedure as described in the next paragraph, a Hamiltonian replica exchange MD simulation [51] with 30 replicas was performed. The energy function of hGBP1 including hGBP1–water interactions was modified in each replica by applying biasing factors of 310 K/T with the 30 temperatures T exponentially distributed between 310 and 450 K. This implies one unbiased replica, the so-called target replica at 310 K. The average exchange probability between the replicas was ~30%. Each replica simulation was 400 ns long, leading to a total of 12 μs for the 30 replicas. The H-REMD simulation was realized with Gromacs 4.5.5 [52] in combination with the PLUMED plugin (version 2.1) [53].

The isolated E domain (residues 482–591) and the E domain plus helix α11 of the M domain (residues 456–591) were both simulated at the all-atom level for 500 ns using Gromacs 2016 [54, 55]. For the E domain two Cl$^-$ ions and for α11 plus the E domain one Na$^+$ ion were addded for neutrality of the systems. The energy of both systems was first minimized using a steepest descent algorithm, followed by equilibration of the systems to the desired temperature of 310 K and pressure of 1 atm for mimicking the physiological environment. First, a 0.1 ns NVT equilibration was performed in which the number of atoms (N), the box volume (V) and temperature (T) were kept constant, followed by a 1 ns NpT equilibration to adjust the pressure (p). During equilibration, the protein atoms were restrained with a force constant of 10 kJ mol^{-1} Å^{-2} allowing the water molecules to relax around the solute. Finally, the 500 ns MD production runs in the NpT ensemble were performed. As no restraints were applied to the protein during the simulations of the E domain (with and without α11), a large cubic box with an edge length of 180 Å was created, allowing free rotation and translation of the 120-Å long helix α12 in all directions. The resulting system sizes involved about 573,000 atoms. The velocity rescaling thermostat was employed to regulate the temperature in the NVT simulations, while the Nosé-Hoover thermostat [56, 57] and the isotropic Parrinello-Rahman barostat [58] were used for the NpT simulations.

Coarse-grained simulations

The coarse-grained simulations were performed with the Martini force field 2.2 and the Martini explicit water [21] as implemented in Gromacs 4.5.5 [52]. As initial conformation for the hGBP1 monomer we used the crystal structure with PDB code 1DG3 [16], which was inserted in a rectangular box with edge lengths of 85, 90 and 170 Å. The box was then solvated with 17,118 Martini water molecules and 7 Na$^+$ ions resulting in a system with 18,438 particles. In order to avoid overall translation and rotation of the protein in the box, the stable secondary structure elements of the LG domain, which were identified based on the atomistic H-REMD simulation results, were restrained to their original positions. After an initial energy minimization and short equilibration, five MD simulations of the system at a temperature of 310 K were performed for different lengths, ranging from 63 μs to 200 μs (see Table 2).

Similar Martini simulations were conducted for the hGBP1 dimer. The dimer was built by superimposing the LG domain of the apo monomer (PDB 1DG3) to one of the LG domains (residues M1–L309) of the LG dimer in complex with GDP, which was resolved by crystallography (PDB 2B92) [12]. In order to avoid clashes between atoms, we replaced the helix involving residues T133–F175 in the dimer with the same helix from the monomer (S7B Fig). For completing the full dimer model, we combined the LG domain dimer up to V288 with the monomer conformation starting from N289. The resulting dimer model was first subjected to

energy minimization at atomistic resolution and in explicit water, and then converted to the Martini model. The coarse-grained hGBP1 dimer was inserted in a rectangular box with edge lengths of 284, 111 and 108 Å filled with Martini water and 14 Na$^+$ ions, amounting to a final system with 32,836 particles. Similarly to the Martini monomer simulation, the stable parts of the LG domain were restrained for both proteins in order to avoid the dimer to rotate and translate, and also to preserve the conformation of the LG domains. In addition, also the G2-SW1 and GC loops were restrained to keep the LG domains in its dimer-specific conformation as present in PDB 2B92 [12]. As listed in Table 2, for the dimer system we performed 5 simulations at 310 K (between 23 μs and 150 μs of length) and one simulation at 320 K (270 μs).

All-atom simulation of the backmapped dimer

To test the stability of the hGBP1 dimer conformation obtained at 320 K, we converted the final snapshot from this Martini simulation to the atomistic level using the backmapping Martini tool [59] and simulated it for 100 ns. The same all-atom MD protocol as used for the MD simulations of the truncated hGBP1 models was applied. In order to limit the system size, we inhibited overall translation and rotation of the dimer by applying position restraints on the β-sheets of both LG domains. The resulting system size involved 376,899 atoms, including 14 Na$^+$ ions, in a rectangular box with edge lengths of 290, 115 and 115 Å. In this simulation, GTP and the Mg^{2+} co-factor were included. To this end, the converted dimer structure was further processed with the free version of the Maestro program [60]. It was loaded together with the superimposed coordinates of Mg^{2+} and GDP·AlF$_3$ obtained from the crystal structure of the GDP·AlF$_3$-bound LG domain of hGBP1 (PDB 2B92) [12], AlF$_3$ was removed, and a phosphate group was attached to the β-phosphate of the existing GDP. The γ-phosphate was added in a way to avoid atom clashes and to be in a reasonable position relative to important protein residues, in particular to K51 and the Mg^{2+} co-factor as described by Kravets [61]. The GTP structure was protonated to mimic the physiological pH value of 7.4 (the protonation state was derived from the GTP structure with PDB code 2KSQ, which was obtained using solution NMR at a pH value of 7 [62]) giving it an overall charge of −4. It was then pre-optimized within the protein and co-factor environment using Maestro, which employes the OPLS 2005 parameters for GTP [63]. The resulting hGBP1 dimer structure including GTP and Mg^{2+} in both LG domains was then used for initiating the MD simulations, including energy minimization and equilibration. During the production MD run, no restraints were applied to GTP and Mg^{2+}. As there are no parameters available for GTP in Amber99SB*-ILDNP, we used the GTP model provided by Meagher et al. [64]. The still missing partial atomic charges of GTP were derived following the Amber99SB*-ILDNP parameterization scheme using the restrained electrostatic potential (RESP) method [65] at the Hartree-Fock theory level applying the 6-31G* basis set after prior geometry optimization of GTP using the B3LYP functional with the 6-31G* basis set. These quantum-chemical calculations were done using the Gaussian 09 program [66].

Analysis

To create pictures of the 3D protein structures, the Visual Molecular Dynamics (VMD) software [68] and PyMOL [69] were used. If not stated otherwise, for the analysis of the H-REMD simulation the data collected by the target replica was used. To quantify the stability and flexibility of hGBP1 during the atomistic MD simulations, the root mean square fluctuations (RMSF) of the C$_\alpha$ around their average positions was calculated. The RMSF calculated for different time intervals of the H-REMD simulation was used to demonstrate that this simulation

had converged within 400 ns per replica (S10 Fig). To determine the time-resolved secondary structure of the E domain the DSSP algorithm (Define Secondary Structure of Proteins) [70] was employed. Clustering analyses were performed to obtain the most populated conformations of the E domain and loops of the LG domain using the Daura algorithm [71] applied to all atoms of the structural element in question and cutoff values of 3.5 Å for the E domain and 2.5 Å for the loops of the LG domain. The details for the calculation of distance maps are given in the Results.

The main structural changes of the M and E domain sampled by the target replica of the H-REMD simulation were identified based on a principal component analysis (PCA) [72]. Given the existence of many different structural fluctuations in hGBP1, we found that applying the PCA to the entire protein is not the best way for separating the different large-scale motions of the protein from each other. Therefore, we only considered the Cartesian coordinates of the M domain and α12 during that analyis as the RMSF analysis had revealed that these regions exhibit the largest flexibility. Helix α13 of the E domain was not included as it was very stable throughout the H-REMD simulation. We projected the conformations from the H-REMD target replica onto the first two principal components (PCs), calculated two-dimensional histograms and then the 2D free energy surface along the two PCs.

For the analysis of the coarse-grained simulations of the hGBP1 monomer, we applied the Markov state model (MSM) approach using the PyEMMA software [73]. First, the five trajectories were subjected to the time-lagged independent component analysis (TICA) [74], a method well suited for dimensionality reduction and recently applied with success in the field of MD simulations [75–78]. The variance of the first two time-lagged independent components (TICs) amounted to 27% of the total variance, and they described best the conformational change involving helix α13. The MSM was then built by clustering the trajectories projected onto the first two TICs using the uniform time clustering algorithm with 300 microstates. We estimated the implied time scales from the MSM for 10 different lag times, based on which we selected a lag time of 450 ns for calculating the MSM of our system. For the identification of metastable Markov states we applied the fuzzy spectral clustering method PCCA+ [79, 80] and used transition path theory [81–83], which is implemented in the PyEmma software, to calculate the reactive fluxes yielding the mean first passage times between the states.

Supporting information

S1 Fig. Clustering of the loop conformations of the LG domain. (A) Initial conformation of the LG domain from which the H-REMD simulation of the hGBP1 monomer was started. The central conformations for all clusters obtained for (B) G1-P, (C) G2-SW1, (D) G3-SW2, (E) G4-L2, and (F) GC are shown. The number of clusters for each loop are given, too. The coloring is the same as in Fig 3. For clarity, GTP (violet) and Mg^{2+} (carmine) are also shown, allowing to distinguish open and closed loop conformations from each other. However, it is emphasized that GTP and Mg^{2+} were not present in the H-REMD simulation of the hGBP1 monomer.
(PDF)

S2 Fig. Information on the eigenvalues and eigenvectors obtained from the PCA of the all-atom H-REMD simulation. (A) Amplitude of the first ten eigenvalues obtained from principal component analysis (PCA) of the all-atom H-REMD simulation of the hGBP1 monomer. (B) Projection of the unbiased H-REMD replica onto the first five PCA eigenvectors.
(PDF)

S3 Fig. Dynamics of the isolated E domain of hGBP1 obtained from all-atom MD simulations. (A) Representative conformations for clusters 3–5 with their occurrence (in %) obtained from the MD simulations of the isolated E domain (top) and of helix α11 plus the E domain (bottom). The cluster structures were aligned to the crystal structure of full-length hGBP1 (shown as transparent cartoon) using residues 482–484 for the alignment of the isolated E domain and α11 for alignment of α11 plus the E domain. (B) Evolution of the secondary structure shown for each residue as a function of time for the isolated E domain (top) and for helix α11 plus E domain (bottom).
(PDF)

S4 Fig. Dynamics of the hGBP1 monomer characterized by Martini MD simulations. (A) Free energy surface plotted along the first two eigenvectors obtained from TICA. The free energy values correspond to the color scale on the right (in $k_B T$ as the Boltzmann constant and T = 310 K). (B) Fuzzy PCCA+ clustering of the microstates resulted in six macrostates. The microstates are projected onto the first two TICs and their membership to one of the six macrostates is identified by different colors.
(PDF)

S5 Fig. Implied time scales obtained from Markov state modeling. (A) Convergences of the implied timescales derived from MSMs for different lag times. One step corresponds to 1.5 ns. (B) Relative implied time scales for a lag time of 450 ns.
(PDF)

S6 Fig. Contact maps for the interactions between α13 and the LG domain. (A) Distance map between the C_α atoms of the C-terminal region, including the helix α13, and the loop formed by residues R151–F174 during transition from the Markov state A to other states in the MSM. The residue index on the x and y axes is accompanied by the corresponding amino acid sequence on the respective opposite side of the plot. (B) C_α-distance map between the C-terminal region as in (A) and relevant residues of the LG-domain calculated for conformations belonging to the Markov state 4. The sequences of the LG domain stretches that form strong contacts with the C-terminal region are displayed on the right, together with structural snapshots where the same residues are highlighted in red, while the helices α12 and α13 are shown on blue and cyan, respectively.
(PDF)

S7 Fig. Positions of α3 and α4' in the hGBP1 monomer and LG domain dimer. (A) Comparison of the position of helix α4' from the hGBP1 monomer (PDB 1DG3, red) with that in the LG domain dimer (PDB 2B92, yellow). This is the major conformational difference between the LG domain monomer and dimer. (B) Comparison of the position of helix α3 in the LG domain dimer (PDB 2B92, yellow) with that from the hGBP1 monomer (PDB 1DG3, red). In order to avoid atom clashes when building the full-length hGBP1 dimer, the yellow helix had to be replaced with the red one.
(PDF)

S8 Fig. E217–K567 distance during selected coarse-grained simulations. Distance between the side chains of E217 and K567 for representative Martini simulations of (black) the hGBP1 monomer at 310 K (obtained from the 63 μs simulation), (red) and (blue) the hGBP1 dimer at 320 K (simulated for 270 μs). In one of the monomers composing the dimer the helix α13 moves (blue) while it does not in the other monomer (red). This motion correlates to the presence of the salt bridge between E217 and K567. In the monomer this salt bridge is never formed (black) as these two residues are too far away from each other. In the dimer, helix α4'

adopts a slightly different position than in the monomer (see S7 Fig), allowing a salt bridge being formed between E217 and K567. This salt bridge impedes the motion of α13 in the dimer. Only if the temperature is raised to 320 K, α13 starts moving in one of the monomers composing the dimer (blue) while it remains intact in the other one (red). At 50 μs the salt bridge in the monomer with the flexible α13 is completely broken, which corresponds to the time when this helix has adopted the 90˚ rotated position (see Fig 7). It should be noted that the distances shown here are between the centers of the coarse-grained side-chain beads and are therefore larger than the atom-based distances usually reported for salt bridges.
(PDF)

S9 Fig. Results obtained from a 100 ns MD simulation of the backmapped hGBP1 dimer. The final snapshot of the dimer obtained from the coarse-grained MD simulation at 320 K was backmapped to the all-atom level. Hence this dimer is called CG-to-AA dimer in the following. The initial CG-to-AA dimer structure is shown in (A) and compared to the dimer model produced for starting the Martini simulations (gray) in (B). Snapshots of the CG-to-AA dimer obtained at 50 and 100 ns of the all-atom simulation are displayed in (C) and (D). The RMSF calculated from that trajectory is projected onto the dimer structure in (E) according to the color scale given in that panel. The evolution of the Q577–Q577 distance during the simulation is plotted in (F). In (G) the conformational clusters of the loops of both LG domains are displayed. The same coloring scheme for hGBP1 as used in the other figures is employed. GTP and Mg^{2+}, which were present in this simulation, are shown in magenta and carmine, respectively.
(PDF)

S10 Fig. Convergence of the H-REMD simulations. The fluctuations of hGBP1 during the first half of the H-REMD simulation (red), the second half (blue) and the full simulation (black) are almost identical, showing that no new conformations are sampled in the second half of the simulation. The results are shown as RMSF values calculated for the target replica (310 K, no energy bias).
(PDF)

S1 Table. Web resources–software and databases.
(PDF)

S1 File. Description of the GROMACS and PyEMMA input / output files available via the Open Science Framework project account https://osf.io/a43z2/.
(PDF)

Author Contributions

Conceptualization: Bogdan Barz, Birgit Strodel.

Data curation: Bogdan Barz, Jennifer Loschwitz.

Formal analysis: Bogdan Barz, Jennifer Loschwitz.

Funding acquisition: Birgit Strodel.

Investigation: Bogdan Barz, Jennifer Loschwitz, Birgit Strodel.

Methodology: Bogdan Barz, Birgit Strodel.

Project administration: Birgit Strodel.

Software: Bogdan Barz.

Supervision: Birgit Strodel.

Validation: Birgit Strodel.

Visualization: Bogdan Barz, Jennifer Loschwitz.

Writing – original draft: Bogdan Barz, Jennifer Loschwitz.

Writing – review & editing: Birgit Strodel.

References

1. Urrutia R, Henley JR, Cook T, McNiven MA. The dynamins: Redundant or distinct functions for an expanding family of related GTPases? Proc Natl Acad Sci USA. 1997; 94:377–384. https://doi.org/10.1073/pnas.94.2.377 PMID: 9012790

2. Henley JR, Cao H, Mcniven MA. Participation of dynamin in the biogenesis of cytoplasmic vesicles. FASEB J. 1999; 13:S243–S247. https://doi.org/10.1096/fasebj.13.9002.s243 PMID: 10619136

3. Kim BH, Chee JD, Bradfield CJ, Park ES, Kumar P, MacMicking JD. Interferon-induced guanylate-binding proteins in inflammasome activation and host defense. Nat Immuno. 2016; 17:481–489. https://doi.org/10.1038/ni.3440

4. Praefcke GJK. Regulation of innate immune functions by guanylate-binding proteins. Int J Med Microbiol. 2018; 308:237–245. https://doi.org/10.1016/j.ijmm.2017.10.013 PMID: 29174633

5. Tretina K, Park ES, Maminska A, MacMicking JD. Interferon-induced guanylate-binding proteins: Guardians of host defense in health and disease. J Exp Med. 2019; 216:482–500. https://doi.org/10.1084/jem.20182031 PMID: 30755454

6. MacMicking JD. Interferon-inducible effector mechanisms in cell-autonomous immunity. Nat Rev Immunol. 2012; 12:367–382. https://doi.org/10.1038/nri3210 PMID: 22531325

7. Anderson SL, Carton JM, Lou J, Xing L, Rubin BY. Interferon-Induced Guanylate Binding Protein-1 (GBP-1) Mediates an Antiviral Effect against Vesicular Stomatitis Virus and Encephalomyocarditis Virus. Virology. 1999; 256:8–14. https://doi.org/10.1006/viro.1999.9614 PMID: 10087221

8. Itsui Y, Sakamoto N, Kakinuma S, Nakagawa M, Sekine-Osajima Y, Tasaka-Fujita M, et al. Antiviral effects of the interferon-induced protein guanylate binding protein 1 and its interaction with the hepatitis C virus NS5B protein. Hepatology. 2009; 50:1727–1737. https://doi.org/10.1002/hep.23195 PMID: 19821486

9. Lipnik K, Naschberger E, Gonin-Laurent N, Kodajova P, Petznek H, Rungaldier S, et al. Interferon γ–Induced Human Guanylate Binding Protein 1 Inhibits Mammary Tumor Growth in Mice. Mol Med. 2010; 16:177–187. https://doi.org/10.2119/molmed.2009.00172 PMID: 20454519

10. Schwemmle M, Staeheli P. The interferon-induced 67-kDa guanylate-binding protein (hGBP1) is a GTPase that converts GTP to GMP. J Biol Chem. 1994; 269:11299–11305. PMID: 7512561

11. Praefcke GJK, Geyer M, Schwemmle M, Kalbitzer HR, Herrmann C. Nucleotide-binding characteristics of human guanylate-binding protein 1 (hGBP1) and identification of the third GTP-binding motif11Edited by P. E. Wright. J Mol Biol. 1999; 292:321–332. https://doi.org/10.1006/jmbi.1999.3062 PMID: 10493878

12. Ghosh A, Praefcke GJK, Renault L, Wittinghofer A, Herrmann C. How guanylate-binding proteins achieve assembly-stimulated processive cleavage of GTP to GMP. Nature. 2006; 440:101. https://doi.org/10.1038/nature04510 PMID: 16511497

13. Gasper R, Meyer S, Gotthardt K, Sirajuddin M, Wittinghofer A. It takes two to tango: regulation of G proteins by dimerization. Nat Rev Mol Cell Biol. 2009; 10:423–429. https://doi.org/10.1038/nrm2689 PMID: 19424291

14. Tripathi R, Glaves R, Marx D. The GTPase hGBP1 converts GTP to GMP in two steps via proton shuttle mechanisms. Chem Sci. 2017; 8:371–380. https://doi.org/10.1039/c6sc02045c PMID: 28451182

15. Abdullah N, Balakumari M, Sau AK. Dimerization and Its Role in GMP Formation by Human Guanylate Binding Proteins. Biophys J. 2010; 99:2235–2244. https://doi.org/10.1016/j.bpj.2010.07.025 PMID: 20923658

16. Prakash B, Praefcke GJK, Renault L, Wittinghofer A, Herrmann C. Structure of human guanylate-binding protein 1 representing a unique class of GTP-binding proteins. Nature. 2000; 403:567. https://doi.org/10.1038/35000617 PMID: 10676968

17. Prakash B, Renault L, Praefcke GJK, Herrmann C, Wittinghofer A. Triphosphate structure of guanylate-binding protein 1 and implications for nucleotide binding and GTPase mechanism. The EMBO Journal. 2000; 19:4555–4564. https://doi.org/10.1093/emboj/19.17.4555 PMID: 10970849

18. Kresse A, Konermann C, Degrandi D, Beuter-Gunia C, Wuerthner J, Pfeffer K, et al. Analyses of murine GBP homology clusters based on in silico, in vitro and in vivo studies. BMC Genomics. 2008; 9:158. https://doi.org/10.1186/1471-2164-9-158 PMID: 18402675

19. Periole X, Mark AE. Convergence and sampling efficiency in replica exchange simulations of peptide folding in explicit solvent. J Chem Phys. 2007; 126:014903. https://doi.org/10.1063/1.2404954 PMID: 17212515

20. Pan AC, Weinreich TM, Piana S, Shaw DE. Demonstrating an Order-of-Magnitude Sampling Enhancement in Molecular Dynamics Simulations of Complex Protein Systems. J Chem Theory Comput. 2016; 12:1360–1367. https://doi.org/10.1021/acs.jctc.5b00913 PMID: 26866996

21. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink SJ. The MARTINI Coarse-Grained Force Field: Extension to Proteins. J Chem Theory Comput. 2008; 4:819–834. https://doi.org/10.1021/ct700324x PMID: 26621095

22. Praefcke GJK, Kloep S, Benscheid U, Lilie H, Prakash B, Herrmann C. Identification of Residues in the Human Guanylate-binding Protein 1 Critical for Nucleotide Binding and Cooperative GTP Hydrolysis. J Mol Biol. 2004; 344:257–269. https://doi.org/10.1016/j.jmb.2004.09.026 PMID: 15504415

23. Liao Q, Kulkarni Y, Sengupta U, Petrović D, Mulholland AJ, van der Kamp MW, et al. Loop Motion in Triosephosphate Isomerase Is Not a Simple Open and Shut Case. J Am Chem Soc. 2018; 140:15889–15903. https://doi.org/10.1021/jacs.8b09378 PMID: 30362343

24. Ince S, Kutsch M, Shydlovskyi S, Herrmann C. The human guanylate-binding proteins hGBP-1 and hGBP-5 cycle between monomers and dimers only. The FEBS Journal. 2017; 284:2284–2301. https://doi.org/10.1111/febs.14126 PMID: 28580591

25. Chakrabartty A, Baldwin RL. Stability of α-Helices. In: Anfinsen CB, Richards FM, Edsall JT, Eisenberg DS, editors. Protein Stability. vol. 46 of Adv. Protein Chem. Academic Press; 1995. p. 141–176.

26. Baker EG, Bartlett GJ, Crump MP, Sessions RB, Linden N, Faul CF, et al. Local and macroscopic electrostatic interactions in single α-helices. Nature Chem Biol. 2015; 11:221–228. https://doi.org/10.1038/nchembio.1739

27. Syguda A, Bauer M, Benscheid U, Ostler N, Naschberger E, Ince S, et al. Tetramerization of human guanylate-binding protein 1 is mediated by coiled-coil formation of the C-terminal α-helices. FEBS J. 2012; 279:2544–2554. https://doi.org/10.1111/j.1742-4658.2012.08637.x PMID: 22607347

28. Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. Science. 1991; 252:1162–1164. http://www.jstor.org/stable/2876291 PMID: 2031185

29. Qin Z, Fabre A, Buehler MJ. Structure and mechanism of maximum stability of isolated alpha-helical protein domains at a critical length scale. Eur Phys J E. 2013; 36:53. https://doi.org/10.1140/epje/i2013-13053-8 PMID: 23708839

30. Vöpel T, Hengstenberg CS, Peulen TO, Ajaj Y, Seidel CAM, Herrmann C, et al. Triphosphate Induced Dimerization of Human Guanylate Binding Protein 1 Involves Association of the C-Terminal Helices: A Joint Double Electron–Electron Resonance and FRET Study. Biochemistry. 2014; 53:4590–4600. https://doi.org/10.1021/bi500524u PMID: 24991938

31. Sezer D, Freed JH, Roux B. Parametrization, Molecular Dynamics Simulation, and Calculation of Electron Spin Resonance Spectra of a Nitroxide Spin Label on a Polyalanine α-Helix. J Phys Chem B. 2008; 112:5755–5767. https://doi.org/10.1021/jp711375x PMID: 18412413

32. Best R, Hofmann H, Nettels D, Schuler B. Quantitative Interpretation of FRET Experiments via Molecular Simulation: Force Field and Validation. Biophys J. 2015; 108:2721–2731. https://doi.org/10.1016/j.bpj.2015.04.038 PMID: 26039173

33. Kalinin S, Peulen T, Sindbert S, Rothwell PJ, Berger S, Restle T, et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat Meth. 2012; 9:1218–1227. https://doi.org/10.1038/nmeth.2222

34. Periole X, Cavalli M, Marrink SJ, Ceruso MA. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. J Chem Theory Comput. 2009; 5:2531–2543. https://doi.org/10.1021/ct9002114 PMID: 26616630

35. Petrović D, Frank D, Kamerlin SCL, Hoffmann K, Strodel B. Shuffling Active Site Substate Populations Affects Catalytic Activity: The Case of Glucose Oxidase. ACS Catal. 2017; 7:6188–6197. https://doi.org/10.1021/acscatal.7b01575 PMID: 29291138

36. Warshel A. Energetics of enzyme catalysis. Proc Natl Acad Sci USA. 1978; 75:5250–5254. https://doi.org/10.1073/pnas.75.11.5250 PMID: 281676

37. Smith AJT, Müller R, Toscano MD, Kast P, Hellinga HW, Hilvert D, et al. Structural Reorganization and Preorganization in Enzyme Active Sites: Comparisons of Experimental and Theoretically Ideal Active Site Geometries in the Multistep Serine Esterase Reaction Cycle. J Am Chem Soc. 2008; 130:15361–15373. https://doi.org/10.1021/ja803213p PMID: 18939839

38. Low HH, Sachse C, Amos LA, Löwe J. Structure of a bacterial dynamin-like protein lipid tube provides a mechanism for assembly and membrane curving. Cell. 2009; 139:1342–1352. https://doi.org/10.1016/j.cell.2009.11.003 PMID: 20064379

39. Low HH, Löwe J. Dynamin architecture—from monomer to polymer. Curr Opin Struct Biol. 2010; 20:791–798. https://doi.org/10.1016/j.sbi.2010.09.011 PMID: 20970992

40. Chen Y, Zhang L, Graf L, Yu B, Liu Y, Kochs G, et al. Conformational dynamics of dynamin-like MxA revealed by single-molecule FRET. Nat Commun. 2017; 8:15744. https://doi.org/10.1038/ncomms15744 PMID: 28548099

41. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins. 2010; 78:1950–1958. https://doi.org/10.1002/prot.22711 PMID: 20408171

42. Best RB, Hummer G. Optimized Molecular Dynamics Force Fields Applied to the Helix–Coil Transition of Polypeptides. J Phys Chem B. 2009; 113:9004–9015. https://doi.org/10.1021/jp901540t PMID: 19514729

43. Aliev AE, Kulke M, Khaneja HS, Chudasama V, Sheppard TD, Lanigan RM. Motional timescale predictions by molecular dynamics simulations: Case study using proline and hydroxyproline sidechain dynamics. Proteins: Structure, Function, and Bioinformatics. 2014; 82(2):195–215. https://doi.org/10.1002/prot.24350

44. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983; 79:926–935. https://doi.org/10.1063/1.445869

45. Darden T, York D, Pedersen L. Particle Mesh Ewald—an N.Log(N) Method for Ewald Sums in Large Systems. J Chem Phys. 1993; 98:10089–10092. https://doi.org/10.1063/1.464397

46. Essmann U, Perera L, Berkowitz ML. A Smooth Particle Mesh Ewald Method. J Chem Phys. 1995; 103:8577–8593. https://doi.org/10.1063/1.470117

47. Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: A linear constraint solver for molecular simulations. J Comput Chem. 1997; 18:1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12%3C1463::AID-JCC4%3E3.0.CO;2-H

48. Feenstra KA, Hess B, Berendsen HJC. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. J Comput Chem. 1999; 20:786–798. https://doi.org/10.1002/(SICI)1096-987X(199906)20:8%3C786::AID-JCC5%3E3.0.CO;2-B

49. Fiser A, Do RK, Sali A. Modeling of loops in protein structures. Protein Sci. 2000; 9:1753–1773. https://doi.org/10.1110/ps.9.9.1753 PMID: 11045621

50. Fiser A, Sali A. ModLoop: automated modeling of loops in protein structures. Bioinformatics. 2003; 19:2500–2501. https://doi.org/10.1093/bioinformatics/btg362 PMID: 14668246

51. Bussi G. Hamiltonian replica exchange in GROMACS: a flexible implementation. Mol Phys. 2014; 112:379–384.

52. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput. 2008; 4:435–447. https://doi.org/10.1021/ct700301q PMID: 26620784

53. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: New feathers for an old bird. Computer Physics Communications. 2014; 185:604–613. https://doi.org/10.1016/j.cpc.2013.09.018

54. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015; 1-2:19–25.

55. Abraham MJ, van der Spoel D, Lindahl E, Hess B. the GROMACS development team GROMACS User Manual Version 2016.4. 2017.

56. Hoover WG. Canonical Dynamics—Equilibrium Phase-Space Distributions. Phys Rev A. 1985; 31:1695–1697. https://doi.org/10.1103/PhysRevA.31.1695

57. Nosé S. Molecular-Dynamics Method for Simulations in the Canonical Ensemble. Mol Phys. 1984; 52:255–268. https://doi.org/10.1080/00268978400101201

58. Parrinello M, Rahman A. Polymorphic Transitions in Single-Crystals—a New Molecular-Dynamics Method. Mol Phys. 1981; 52:7182–7190.

59. Wassenaar TA, Pluhackova K, Böckmann RA, Marrink SJ, Tieleman DP. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. J Chem Theory Comput. 2014; 10:676–690. https://doi.org/10.1021/ct400617g PMID: 26580045

60. Maestro 9.7; 2014.

61. Kravets E. Charakterisierung des murinen Guanylat-bindenden Proteins 2 (mGBP2). Heinrich-Heine-University, Düsseldorf, Germany. 2012.

62. Liu Y KR A P JHY. Dynamic structure of membrane-anchored Arf*GTP. Nat Struct Mol Biol. 2010; 17:876–881. https://doi.org/10.1038/nsmb.1853

63. Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). J Comput Chem. 2005; 26:1752–1780. https://doi.org/10.1002/jcc.20292 PMID: 16211539

64. Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. J Comput Chem. 2003; 24:1016–1025. https://doi.org/10.1002/jcc.10262 PMID: 12759902

65. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem. 1993; 97:10269–10280. https://doi.org/10.1021/j100142a004

66. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 09 Revision E.01; 2009.

67. Krause D, Thörnig P. JURECA: Modular supercomputer at Jülich Supercomputing Centre. Journal of large-scale research facilities. 2018; 4:A132. https://doi.org/10.17815/jlsrf-4-121-1

68. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996; 14:33–38. https://doi.org/10.1016/0263-7855(96)00018-5 PMID: 8744570

69. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015; http://www.pymol.org.

70. Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983; 22:2577–2637. https://doi.org/10.1002/bip.360221211 PMID: 6667333

71. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide Folding: When Simulation Meets Experiment. Angew Chem Int Ed. 1999; 38:236–240. https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2%3C236::AID-ANIE236%3E3.0.CO;2-M

72. David CC, Jacobs DJ. In: Livesay DR, editor. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. Totowa, NJ: Humana Press; 2014. pp. 193–226.

73. Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, et al. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. J Chem Theory Comput. 2015; 11:5525–5542. https://doi.org/10.1021/acs.jctc.5b00743 PMID: 26574340

74. Molgedey L, Schuster HG. Separation of a mixture of independent signals using time delayed correlations. Phys Rev Lett. 1994; 72:3634–3637. https://doi.org/10.1103/PhysRevLett.72.3634 PMID: 10056251

75. Pérez-Hernández G, Paul F, Giorgino T, De Fabritiis G, Noé F. Identification of slow molecular order parameters for Markov model construction. J Chem Phys. 2013; 139:015102–13. https://doi.org/10.1063/1.4811489 PMID: 23822324

76. Schwantes CR, Pande VS. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. J Chem Theory Comput. 2013; 9:2000–2009. https://doi.org/10.1021/ct300878a PMID: 23750122

77. Sengupta U, Strodel B. Markov models for the elucidation of allosteric regulation. Philos Trans R Soc, B. 2018; 373:20170178. https://doi.org/10.1098/rstb.2017.0178

78. Sengupta U, Carballo-Pacheco M, Strodel B. Automated Markov state models for molecular dynamics simulations of aggregation and self-assembly. J Chem Phys. 2019; 150:115101; https://doi.org/10.1063/1.5083915

79. Röblitz S, Weber M. Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. Adv Data Anal Classif. 2013; 7:147–179. https://doi.org/10.1007/s11634-013-0134-6

80. Noé F, Wu H, Prinz JH, Plattner N. Projected and hidden Markov models for calculating kinetics and metastable states of complex molecules. J Chem Phys. 2013; 139:184114–17. https://doi.org/10.1063/1.4828816 PMID: 24320261

81. Weinan E, Vanden-Eijnden E. Towards a Theory of Transition Paths. J Stat Phys. 2006; 123:503. https://doi.org/10.1007/s10955-005-9003-9

82. Metzner P, Schütte C, Vanden-Eijnden E. Transition Path Theory for Markov Jump Processes. Multiscale Model Simul. 2009 ;7:1192–1219. https://doi.org/10.1137/070699500

83. Noé F, Schütte C, Vanden-Eijnden E, Reich L, Weikl TR. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. Proc Natl Acad Sci USA. 2009; 106:19011–19016. https://doi.org/10.1073/pnas.0905466106 PMID: 19887634