THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

Cite This: J. Phys. Chem. A 2018, 122, 7924−7930

pubs.acs.org/JPCA

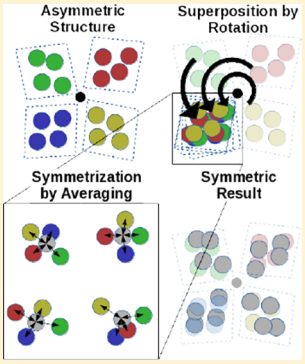
Quantifying Asymmetry of Multimeric Proteins

Julian T. Brennecke and Bert L. de Groot*

Department of Theoretical and Computational Biophysics, Computational Biomolecular, Dynamics Group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Supporting Information

ABSTRACT: A large number of proteins assemble as homooligomers. These homooligomers accomplish their function either symmetrically or asymmetrically. If asymmetry is prevalent in a structure ensemble, the asymmetric motion will occur in any of the subunits. Many computational analysis tools implicitly use ensemble averages to determine protein motions, e.g., principle component analysis. Therefore, taken together, this approach results in a loss of the asymmetric signal and a false symmetric output, rendering it impossible to analyze asymmetric motions with available tools. A first step toward understanding asymmetric systems is the quantification of asymmetry. Only a few tools exist to calculate asymmetry quantitatively, such as the continuous symmetry measure (CSM). In this study, we present an extension of CSM delivering additional information about the subunit contributions to the overall asymmetry. Furthermore, we introduce an algorithm termed the functional asymmetry measure (FAME). FAME is based on an algorithm that predicts functionally relevant motions of a protein (PLS-FMA) and thus allows calculating asymmetry in relation to protein function. To validate our developed algorithm, we applied it to two different potassium channels, TREK-2 and KcsA, as well as to the unfolding mechanism of the carrier protein Transthyretin. For both potassium channel systems, an artificial asymmetric motion was introduced to benchmark the algorithm in addition to demonstrate the interpretation potential of the results. Therefore, the degree of overall as well as subunit based asymmetry for KcsA was quantified using CSM as the provided extension requires more than two subunits. The functional modes of asymmetric TREK-2 motions were recovered and their asymmetry was quantified using FAME as a dimeric protein is the simplest application. FAME was further used to study the asymmetry of the unfolding pathway of Transthyretin. We show the ability of both algorithms to correctly predict asymmetry. The tools are available online and can be applied to most homooligomeric systems.



INTRODUCTION

A major fraction of all proteins is functionally active in a homooligomeric form and many of these are symmetric.[1−4] This dominance of symmetric structures[1] can only be explained by an evolutionary advantage of symmetry.[1] Among others, increased stability of symmetric homooligomers[3,5,6] and a well-defined number of subunits in the multimer,[1] as well as favorable folding pathways of symmetric structures,[7] are speculated to be the evolutionary driving force.

However, protein functions are known which indeed require structural asymmetry, such as unidirectional motion in polymerases and ribosomes (structure refers to the 3D coordinates of each atom), or reciprocal symmetry (the adoption of different states in the subunits) as in the ATP synthase.[8] Furthermore, according to the MWC model of allostery, symmetry is a prerequisite for allosteric proteins.[9] The MWC model even suggests that the conservation of symmetry causes allostery. While binding of a ligand to one subunit, a structural change in all subunits is initiated, further facilitating the binding of similar ligands to all subunits. In contrast to the MWC model, the KNF model of allostery requires asymmetric intermediates.[10]

The need for asymmetric intermediates in the KNF model suggests once more that asymmetric intermediates are essential for the dynamic function of proteins. Indeed homooligomeric proteins are known to have functionally relevant asymmetric states apart from their symmetric states. One of these examples is the heat shock protein Chaperonin 60 that adopts a functionally relevant asymmetric configuration upon ATP binding.[11] Furthermore, CorA, a multimeric Mg^{2+} channel, was found to abandon its symmetric closed state to open.[12,13]

For many structural analysis of a molecule its structure has to be stored, and thus each atom has to be assigned a label. For monomers and heterooligomers, conventions exist making the labeling of any atom unique. However, for homooligomers the subunits are identical, what results in ambiguous subunit labels. If the asymmetry of a system is the result of a transition from a symmetric to an asymmetric state, it can take place in any of the subunits. In MD simulations, the ambiguity in subunit labeling challenges the analysis of these transitions because most analysis tools work on ensembles. As the asymmetry will spontaneously occur in varying subunits, the ensemble averaging will misleadingly show symmetric transition. To overcome this challenge, a thorough understanding of the occurring symmetry and asymmetry is required. A first step toward it is its quantification. The contribution of the subunits to the

Received: July 17, 2018
Revised: September 5, 2018
Published: September 6, 2018

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930

overall asymmetry is of special interest as it shows how the individual subunits behave and consequently can resolve the ambiguity. To label subunits unambiguously, the asymmetry contribution of each subunit can be used to have the major asymmetry consistently occurring in the same subunit.

Fully symmetric structures are unfavorable in entropy as well in interaction enthalpy of the contact zones of the multimers. Thus, symmetry is unlikely if it is not enforced,[14] but the resulting asymmetry is not necessarily related to protein function. We call this type of asymmetry statistical asymmetry. As the interest in analyzing proteins focuses on their function, it remains obvious to identify asymmetry related to function. We call this type of asymmetry functional asymmetry. However, the availability of additional information about the protein function is required.

To analyze and quantify structural (a)symmetry, the continuous symmetry measure (CSM) was developed by Zabrodsky et al.,[15,16] extended to incorporate further measures such as local symmetries[14] and applied to various systems.[14,17] In this study, we present an extension to CSM calculating subunit contributions to the overall asymmetry.

Naturally, the CSM algorithm is unable to distinguish functional from statistical asymmetry. Therefore, we extended an algorithm to predict functionally relevant motions from an ensemble of structures such as trajectories of a MD simulation called PLS-FMA[18,19] to quantify functional asymmetry. The inclusion of the algorithm gives us the functional asymmetry measure (FAME) which describes the asymmetry along the functionally relevant motions identified by PLS-FMA. An implementation of the algorithms able to process MD simulation trajectories is available online at the Web site https://gitlab.gwdg.de/deGroot/asympy.git.

For a proof of principle, we applied CSM to modified KcsA and FAME to TREK-2 simulations with artificially introduced asymmetry. We demonstrate the ability of both algorithms to recover the introduced asymmetry. Using FAME, the heat triggered unfolding of Transthyretin is studied using contact data as a functional property. The agreement of the subunit contributions with the individual subunit contact data demonstrates the ability of FAME to deal with more difficult functional modes.

With the new algorithms the symmetry of homooligomers can be quantified.

THEORY

In the following, the continuous symmetry measure (CSM) will be introduced and extended to subunit contributions of asymmetry. For functionally relevant motions the functional asymmetry measure (FAME) will be developed. For both algorithms, an implementation is provided online at the Web site https://gitlab.gwdg.de/deGroot/asympy.git.

This implementation works for rotational symmetries only as these are most common in proteins.[3] However, the implementation can serve as a basis to extend it to other symmetries as well.

Continuous Symmetry Measure (CSM). The continuous symmetry measure (CSM) is the normalized distance ($S(G)$) of a structure to the nearest symmetric structure according to RMSD distance (see Figure 1)[15]

$$S(G) = \frac{100}{d^2} \sum_{i=1}^{N} |Q_i - Q_i^{sym}|^2 \quad (1)$$

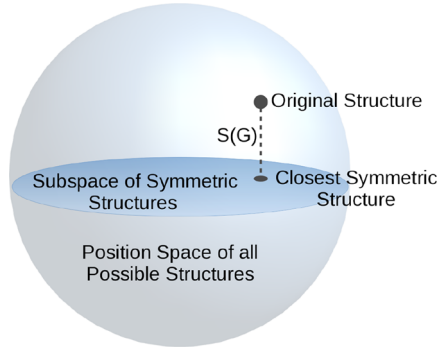


Figure 1. CSM position space. The studied original structure is projected onto its closest (by RMSD distance) symmetric representation in position space. The normalized distance to this structure, indicated by the dotted line, is the CSM measure (S(G)).

Q_i are the coordinates of the ith atom of the protein with N atoms. Q_i^{sym} is its corresponding symmetric structure and d is a normalization factor taking the overall size of the protein ($d = \sqrt{\sum_{i=1}^{N} |Q_i|^2}$) into account.

To find a symmetric structure rotations around an axis have to be performed followed by an averaging over all rotations (see Figure 2).[15,16] To minimize the RMSD distance between

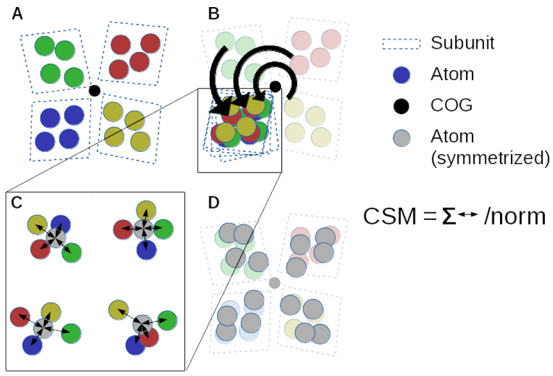


Figure 2. Illustration of the CSM algorithm. A visual introduction into the working of CSM. (A) A 4-fold symmetric protein (homotetramer, presented in 2D) of which the CSM measure is to be calculated. (B) Rotation around the center of geometry is applied such that all subunits are superimposed. In 3D, the rotation is performed around an axis which has to be determined previously. The rotation angle of subunit j is $2\pi j/J$. (C) Average position over the same atom in the different subunits is calculated. (D) Every subunit is averaged, resulting in a symmetric structure.

the original structure (Q) and the symmetric structure (Q^{sym}) the optimal axis has to be determined. This axis can be identified analytically following the approach described by Pinsky et al.[16]

Some residues like Valine and Leucine have an intrinsic labeling redundancy. As two or more atoms are chemically identical they can be labeled in different ways. To account for this we perform a permutation scheme for all of these residues in any possible combination of subunits and choose the combination that minimizes the overall asymmetry.

For multimers consisting of more than two subunits, one subunit behaves most differently from the others. One example is the open state of the CorA Mg^{2+} channel in which four subunits form an almost symmetric structure while one subunit is moved outward.[13] This distinction between the subunits is

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930

not captured in the current CSM algorithm. CSM only measures the asymmetry of the overall structure. To address the difference in asymmetry between the subunits we developed an extension to the CSM algorithm. As a result, a CSM measure for individual subunits is created what allows to distinguish the subunits based on their asymmetric behavior.

To access the CSM measure of each subunit ($S_j(G)$ with $0 < j \leq J$, with the number of subunits J), the calculation is split up into

$$S_j(G) = \frac{100}{d^2} \sum_{i=j\cdot n+1}^{(j+1)\cdot n} |Q_i - Q_i^{sym}| \tag{2}$$

with $n = N/J$ the number of atoms per subunit. Note that the normalization factor is kept the same for the single subunits. As a result, $S_j(G)$ is not truly the CSM measure of each of the subunits, but rather a decomposition of the overall CSM measure into the contributions of each of the subunits. Thus, the overall CSM measure is the result of a summation over the individual asymmetry measures of the subunits ($S(G) = \sum_{j=1}^{J} S_j(G)$).

It is important to note that this additional decomposition is not meaningful for a dimer. For a dimer, a symmetric structure can be constructed that is the average of the two subunits. Thus, the contribution of each of the subunits to the overall CSM measure is the same.

Functional Asymmetry Measure (FAME). Whereas the CSM method can be applied to quantify asymmetry in a system, using it, we are unable to distinguish between thermal fluctuations and functionally relevant asymmetry. Therefore, to distinguish functionally relevant from random asymmetry, we developed the functional asymmetry measure (FAME). The distinction is achieved by first determining the contributions of each subunit to a relevant functional motion of the protein. Subsequently, the information on the asymmetry is extracted by comparing the contributions of the subunits. One subunit can contribute significantly more than the others only if the motion described by the contributions is asymmetric. In contrast, if two subunits contribute similarly in respect to the same motion, we conclude the dynamics of the protein are symmetric.

To separate the functional motions from thermal fluctuations, we extended the PLS-FMA analysis[18,19] by FAME. Using PLS-FMA, we try to extract the functional motions of a protein. A good estimate of the functional motion is often the maximally correlated motion (reference motion) to a given one-dimensional functional property ($\boldsymbol{f}$). An example of this property could be the distance of two atoms in a binding site of an enzyme. As a result, PLS-FMA (and also FAME) requires a suitable functional property which can be challenging to find. In FAME, this functional property is used to calculate the contributions of the individual subunits to the reference motion and to extract the asymmetry by their relation to each other. To relate the contributions of the individual subunits to the (a)symmetry, they have to be calculated with respect to a symmetric reference motion. As a counter example: assuming an asymmetric reference motion, if two subunits would contribute equally, they could not be assumed to be symmetric as the reference motion of the individual subunits would describe different motions. However, as described before, the relation of contributions is supposed to describe the asymmetry i.e. the same contribution is defined to be symmetric. Thus, the contributions have to be calculated along a symmetric reference motion. We construct the symmetric reference by permuting the monomers.

In mathematical terms, PLS-FMA is used to predict the functional property ($\boldsymbol{f}$) from a structural ensemble (e.g., a MD trajectory) ($\mathbf{X}$) by determining the coefficients ($\boldsymbol{\beta}$) in the equation

$$\boldsymbol{f} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{3}$$

to minimize the sum of errors (ϵ^2). Note that $\boldsymbol{f}$ is a vector containing the unidimensional functional property to be predicted for each of the structures in the ensemble. The structural ensemble is summarized in the matrix $\mathbf{X}$ containing the position of each atom for any structure of the ensemble.

To quantify asymmetry using PLS-FMA, the contributions of the subunits are separately analyzed. To separate the contributions, the coefficient vector $\underline{\beta}$ is split into the individual subunits. In mathematical terms

$$\tilde{\boldsymbol{f}}_j = \mathbf{X}\boldsymbol{\beta}_j \tag{4}$$

where $\tilde{\boldsymbol{f}}_j$ is the prediction of the contribution to the overall functional value $\underline{f}$ for the subunit j. The ith component of the vector $\boldsymbol{\beta}_j$ is given by

$$\beta_{j,i} = \begin{cases} \beta_i & \text{if } j\cdot 3n \leq i < (j+1)\cdot 3n, \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

with n denoting the number of atoms per subunit. Note that the coordinates of the structure are resorted to fit the $3n$ coordinates into a one-dimensional vector. The overall functional property ($\tilde{\boldsymbol{f}}$) is the sum over all contributions ($\tilde{\boldsymbol{f}}_j$). As a result, the contribution of each of the subunits $\mathbf{C}_j$ to the overall functional property $\boldsymbol{f}$ can be estimated as $\mathbf{C}_j = \tilde{\boldsymbol{f}}_j/\tilde{\boldsymbol{f}}$.

To interpret the result, the expectation on $\underline{C}_j$ needs to be understood for both asymmetric and symmetric subunits. If the subunits are symmetric, the contribution of each is expected to be $1/J$ (J = no. of subunits). However, if they are asymmetric, the contributions deviate from $1/J$ until they reach either zero or one. Zero corresponds to no contribution of the subunit whereas a value of one corresponds to a unique contribution of the subunit to the functional motion. By construction, $\mathbf{C}_j$ is normed to $\sum_j \mathbf{C}_j = 1$.

Applications. To test the algorithms and illustrate their use in analyzing (a)symmetry, three different test systems were used, i.e., KcsA, TREK-2, and TTR. The first system is the well studied KcsA channel[20] with a 4-fold symmetric crystal structure (PDB ID: 3f5w). To control its motion, a trajectory is created by duplicating the crystal structure 500 times. These 500 structures are supposed to represent five trajectories containing 100 frames each. For each of these five trajectories, one of the four subunits is selected randomly. In this subunit, a motion of the outer helices is artificially introduced (see Figure 3). This motion is a rigid body rotation around the base of the helix. In each frame, the helix is rotated slightly as long as the angle stays in a certain range. The angle and the subunit which is altered are recorded and shown in Figure 4. Thereby, a trajectory of a symmetric protein with a well-defined asymmetric motion in one of the subunits is created. The RMSD of the resulting trajectory is shown in Figure S1.

The second test system is the eukaryotic mechanosensitive potassium channel TREK-2. Previously published equilibrium MD simulations of this homodimer without its asymmetric cap were used.[21] On top of the motions from the MD simulations a

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930

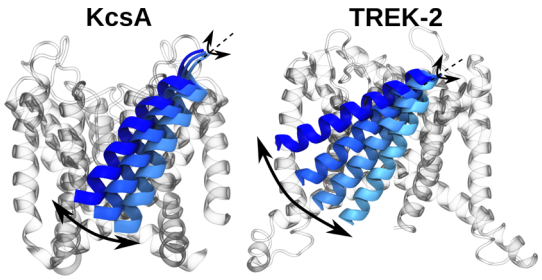


Figure 3. Artificially introduced motion. For two different test systems, the static KcsA channel and simulations of the TREK-2 channel, manipulations on outer helices are performed. They are rotated around a hinge point at their origin. These systems are later used to illustrate the analysis of (a)symmetry by CSM and FAME.

similar protocol as described before for KcsA is used (see Figure 3). However, to make it more realistic a signal, the functional property, was not exclusively introduced in one subunit but was distributed over both subunits. A trace of values was created similar to the angle in KcsA reflecting the functional property (f) for FAME. The trace was distributed between the two subunits according to a second trace with values between zero and one. For the helix angle of subunit one, the first trace was multiplied with the distribution value (contribution of subunit one) and the remaining helix angle (1-distribution) demonstrated the movement of the helix angle of the second subunit. Thus, the contribution to the signal by each of the subunits represents the asymmetry. If both subunits have the same angle (contribution = 0.5), the system is assumed to be symmetric. In contrast, if only the first (second) subunit contributes (contribution = 1.0 (0.0)), the highest functional asymmetry is assumed for the system. Predicting the asymmetry of the system is the main goal of this algorithm. The comparison of the RMSD of the system before and after introducing the helix motion is illustrated in Figure S2.

The third test system is the human transport protein transthyretin (TTR). MD simulations of tetrameric TTR (PDB ID: 4pvm[22]) were carried out at 298 K and at 598 K. Whereas simulations at 298 K showed a stable behavior, we saw denaturation of the protein in simulations at 598 K with one of the monomers typically initializing the unfolding process. The progress of denaturation can be quantified by the number of native contacts inside the protein. The overall number of native contacts was quantified using the tool developed by Best et al.[23] Furthermore, FAME was used to determine the individual degree of denaturation of each subunit.

RESULTS

To demonstrate the usage of CSM and FAME and the interpretation of their results, artificial systems of KcsA and TREK-2 with a known asymmetry were created and tested. The more challenging example is the unfolding process of the tetrameric protein TTR. Here, information about native contacts is used that also reveals the contribution of the subunits to test a more subtle relation between functional property and structure of the system.

CSM Analysis of KcsA. To test the ability of the CSM algorithm in predicting the contributions of different subunits, full control over the complete asymmetry in the test system is required. To achieve this, a symmetric crystal structure of the KcsA channel was used and an asymmetric motion was artificially introduced (see Applications). This asymmetric motion represents the symmetry distortion by bending one helix (see Figure 3).

The CSM algorithm was applied to the created ensemble. The result is the total asymmetry of the protein, and we find it to be similar to the input asymmetry (see Figure 4). However, the total asymmetry does not give any details on the asymmetry distribution over the different subunits. As seen from the input data, the signal (helix motion) is present in the different subunits which is the information reconstructed by the presented extension of the CSM algorithm. For the first part of the frames (<1000), the highest contribution of the asymmetry is assigned to the second subunit, in which the signal was introduced. However, looking at the resulting asymmetry, other subunits contribute partially. The partial asymmetry in these is a result of the symmetric structure definition, which is constructed by averaging over the configuration of each subunit. Thus, the subunit with the introduced signal contributes to the average structure also. As a result, some of the contributions to the asymmetry can be found in the subunits without the signal, which is not an error but rather the asymmetry definition.

FAME Analysis of TREK-2. The advantage of the FAME algorithm compared to the CSM algorithm is its ability to distinguish between functional asymmetry and asymmetry due to thermal fluctuations. To test the prediction potential for functional asymmetry, a system containing functional as well as

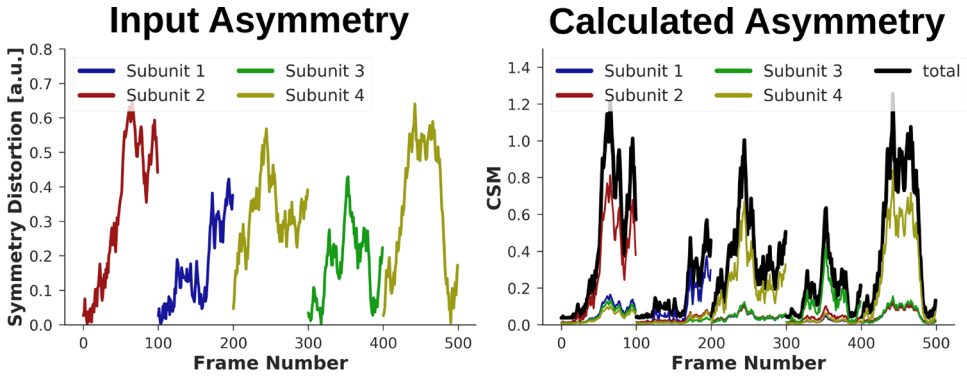


Figure 4. CSM decomposition of KcsA. The 4-fold symmetric KcsA channel (crystal structure) is used and a modification of a single helix is introduced. This helix is rotated by a given angle which distorts the symmetry of the structure. The distortion is shown in as the input asymmetry. Because of symmetry reasons, the distortion can be introduced in any of the four subunits which is indicated by the color. The asymmetry determined by the CSM algorithm is shown in the calculated asymmetry. In addition to the total asymmetry score (black line), the contributions of the different subunits are shown (colored lines).

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930

nonfunctional asymmetry was created. To this end, in the presented test system the functional asymmetry is added to a MD simulation trajectory of TREK-2 at room temperature. For the first analysis, the information about the helix angle was used directly as a functional property predicted by PLS-FMA. Second, the applicability of FAME to a more indirect functional property was tested (see FAME Analysis of TTR).

To this end, the overall helix angle was predicted by PLS-FMA. To use PLS-FMA, the optimal number of PLS components were evaluated first (refer to[19] for more details and see Figure S3). Using the optimal number of components, the original input data was predicted with high precision (see Figure 5, original data is predicted by PLS-FMA).

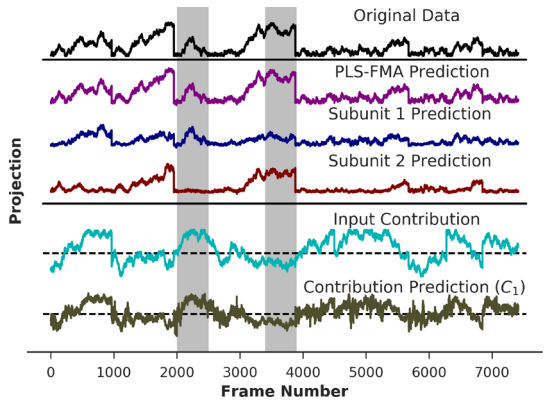


Figure 5. Signal decomposition into subunit contributions, The overall PLS-FMA input signal (Original Data) is predicted by the standard PLS-FMA algorithm (PLS-FMA Prediction). The FAME algorithm disassembled the signal into its contributions from the different subunits (subunit 1/2 prediction). The input asymmetry (input contribution) is predicted by the FAME algorithm (contribution prediction (C_1)) where the dotted lines represent equal contributions of the two subunits. The highlighted regions are illustrate the result of the algorithm particularly well.

The FAME algorithm provides further information on the composition of the general prediction by PLS-FMA. The PLS-FMA prediction is decomposed into the contribution of the different subunits, which were used to calculate the functional asymmetry. The information about the asymmetry is stored in the input contribution as the overall signal is asymmetrically distributed over the two subunits. If the contributions are equally distributed between the two subunits, the signal is symmetric as dotted lines in Figure 5 illustrate. On the contrary, if the contribution is above the dotted line, the signal is mainly found in subunit one, below the dotted line is a dominant signal in subunit two. This unequal distribution of the signal is the asymmetry. As can be seen in Figure 5 (C_1) the contribution predicted by the FAME algorithm is similar ($R^2 = 0.71$) to the input contribution.

FAME Analysis of TTR. Human transthyretin (TTR) is a tetrameric protein transporting the hormone thyroxine and the retinol-binding protein.[24,25] The temperature triggered unfolding process of TTR was calculated. For the unfolding of a protein, the symmetry is not enforced and thus the unfolding process will start in one of the subunits and will likely be dominated by the unfolding of this subunit. Information on the unfolding process is gathered by calculating contact data (see Applications). The contact data of the entire protein were used as functional input data (see Figure S4). The deviation of the different subunits of the tetramer to the average number of contacts is shown in Figure 6. A deviation from the average number of contacts represents the asymmetry of one subunit. If all subunits would denature in the same way, the contact numbers of each of the subunits would be the same. Only if the denaturing process takes place in an asymmetric way, the contacts differ between the subunits. In Figure 6, the contributions (C_j) of the different subunits to the overall denaturing process are predicted. The comparison of the input data to the calculated contributions reveals a significant agreement. Initially, the subunits behave very similarly. Only after around half of the trajectory, the subunits two and three show a higher contribution to the contact data than the other subunits. A comparable behavior is apparent from the calculated subunit contributions. However, looking at the intermediate part, starting around frame 1000, the change in subunit two is more pronounced in the original data and almost returns to an equal contribution for the predicted data toward the end of the simulation.

DISCUSSION

The aim of the presented research was to find suitable measures for (a)symmetry with and without functional relation. Application of our extension of CSM to KcsA show its ability to produce results in good agreement with the known input asymmetry on a subunit level. Comparing the input with regard to the subunits in which the helix angle motion is introduced and the results of the subunit information, the algorithm demonstrates its ability to correctly detect the subunit in which the change was introduced. However, a direct interpretation of the results can not be generalized. To understand the motion of the subunits, a visual inspection is required. Our algorithm aids in identifying relevant sections of a simulation as well as giving a comprehensive overview after further analysis. It should be noted that the splitting of the CSM value into subunit contributions is only meaningful for oligomers with more than two subunits. In the case of a dimer, each of the subunits will have exactly the same contribution to the asymmetry as the closest fully symmetric structure is the average of the two.

The results of the TREK-2 FAME analysis demonstrate the ability of this algorithm to predict the contribution of each subunit accurately given a useful functional property. How the contributions and subunit predictions are related to the overall functional data can be readily derived from the highlighted region in the FAME algorithm results for TREK-2 (see Figure 5). The first highlighted small peak in the original data, was predicted correctly by the PLS-FMA algorithm. Looking at the predictions for the individual subunits, it is obvious that the peak originates from a signal in subunit one whereas in subunit two, no signal can be seen. The input contributions show that at the creation time of the trajectory, the signal was mainly introduced into the first subunit, meaning that the helix of the first subunit was altered whereas the helix of the second subunit was kept in its original state. This asymmetry is what the algorithm is supposed to recover from the data. As can be seen for the given artificial data where the results of the algorithm can be compared to the input, the FAME algorithm recovers the contribution of the subunits (C_j) well.

Another way to interpret the data is to consider what PLS-FMA does. PLS-FMA creates a vector describing the motion of the protein most correlated with the functional data. Further, it projects the actual motion of the protein from the simulation

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930

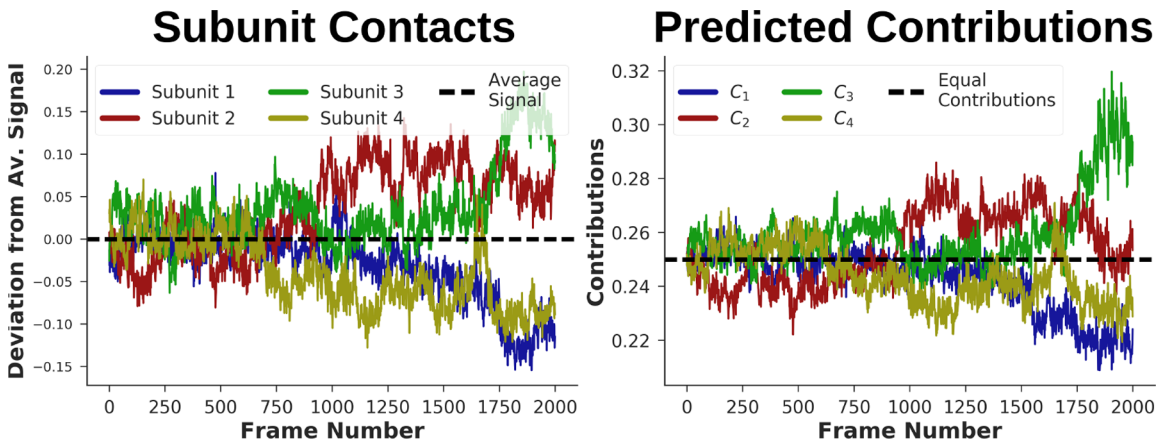


Figure 6. TTR denaturation studied by FAME. The subunit contacts are shown for each of the subunits. These are a measure of (a)symmetry across the subunits. We aim to predict the correct contributions of each of the subunit by FAME. The predicted contributions compare well to the subunit contacts which therefore shows FAMEs ability to retrieve the (a)symmetry of the system.

along this vector. The contribution value shows at which position along this vector one subunit is located in reference to the other subunits. Therefore, the contribution value is equal to a functional asymmetry measure.

Similar to the KcsA channel, the information about the asymmetry of TTR was recovered. In this case, the functional value was decomposed into the contributions of the subunits. However, in most cases, such as the size of a cavity in the middle of a multimer, it is impossible to disassemble the functional value according to the contribution of the different subunits. In such cases, the presented methodology can contribute to a more complete understanding of a homooligomeric system.

Lastly, a comment on the possibility to combine both algorithms to estimate the overall asymmetry of a system compared to the functional asymmetry. If the CSM algorithm is used to compute the overall asymmetry of the system and the FAME algorithm is used to calculate the functional asymmetry, the difference of these two reveals statistical asymmetry of the system. Statistical asymmetry arises from a lack of enforcement of symmetry of the system and results mainly from thermal fluctuations.

Overall we could demonstrate our algorithms ability to quantify (a)symmetry with and without regard to functional relevance.

CONCLUSIONS

Two methods were presented to analyze asymmetry. If additional information on a functional coordinate of interest is missing, a modified version of the CSM algorithm can be used.[15,16] This version allows for the quantification of the overall system asymmetry as well as the quantification of the asymmetric contributions of the individual subunits.

If additional information on a functional property of the system is available, it allows for the use of PLS-FMA based FAME algorithm restricting asymmetry to functional motions only. The application of both algorithms to quantify asymmetry are shown and the interpretation of results is discussed. A comparison to the known input (a)symmetry revealed the ability of the algorithms to quantify asymmetry correctly.

The software used to perform the analysis is freely available online at the Web site https://gitlab.gwdg.de/deGroot/asympy.git.

The CSM algorithm was implemented based on Pinsky et al. 2008 algorithm.[16] It can be applied to simulations of rotational symmetric proteins without additional modifications. A more detailed step by step usage of the software is provided in the Supporting Information of this publication.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.8b06843.

Technical information on the implementation and the usage of the algorithms and additional figures supporting the manuscript (PDF)

AUTHOR INFORMATION

Corresponding Author

*(B.L.d.G.) E-mail: bgroot@gwdg.de. Telephone: +49 551 2012308.

ORCID

Bert L. de Groot: 0000-0003-3570-3534

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thankfully acknowledge Matteo Aldeghi for performing the MD simulations of TTR. Funding awarded by the Deutsche Forschungsgesellschaft Sonderforschungsbereich 803: Project A03 to J.T.B. and B.L.d.G. is gratefully acknowledged.

REFERENCES

(1) Goodsell, D. S.; Olson, A. J. Structural Symmetry and Protein Function. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105−153.

(2) André, I.; Strauss, C. E.; Kaplan, D. B.; Bradley, P.; Baker, D. Emergence of Symmetry in Homooligomeric Biological Assemblies. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16148−16152.

(3) Blundell, T. L.; Srinivasan, N. Symmetry, Stability, and Dynamics of Multidomain and Multicomponent Protein Systems. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 14243−14248.

(4) Taylor, W. R.; May, A. C.; Brown, N. P.; Aszódi, A. Protein Structure: Geometry, Topology and Classification. *Rep. Prog. Phys.* **2001**, *64*, 517−590.

(5) Klotz, I. M.; Langebman, N.; Dahnall, D. Quaternary Structure of Proteins. *Annu. Rev. Biochem.* **1970**, *39*, 25−62.

(6) Cornish-Bowden, A. J.; Koshland, D. The Quaternary Structure of Proteins Composed of Identical Subunits. *J. Biol. Chem.* **1971**, *246*, 3092−3102.

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930

(7) Wolynes, P. G. Symmetry and the Energy Landscapes of Biomolecules. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 14249−14255.
(8) Abrahams, J. P.; Leslie, A. G.; Lutter, R.; Walker, J. E. Structure at 2.8 Å Resolution of F1-ATPase from Bovine Heart Mitochondria. *Nature* **1994**, *370*, 621−628.
(9) Monod, J.; Wyman, J.; Changeux, J. On the Nature of Allosteric Transitions: A Plausible Model. *J. Mol. Biol.* **1965**, *12*, 88−118.
(10) Koshland, D., Jr; Némethy, G.; Filmer, D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits. *Biochemistry* **1966**, *5*, 365−385.
(11) Saibil, H.; Zheng, D.; Roseman, A.; Hunter, A.; Watson, G.; Chen, S.; Auf der Mauer, A.; O'hara, B.; Wood, S.; Mann, N.; et al. ATP Induces Large Quaternary Rearrangements in a Cage-Like Chaperonin Structure. *Curr. Biol.* **1993**, *3*, 265−273.
(12) Pfoh, R.; Li, A.; Chakrabarti, N.; Payandeh, J.; Pomès, R.; Pai, E. F. Structural Asymmetry in the Magnesium Channel CorA Points to Sequential Allosteric Regulation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 18809−18814.
(13) Matthies, D.; Dalmas, O.; Borgnia, M. J.; Dominik, P. K.; Merk, A.; Rao, P.; Reddy, B. G.; Islam, S.; Bartesaghi, A.; Perozo, E.; et al. Cryo-EM Structures of the Magnesium Channel CorA Reveal Symmetry Break upon Gating. *Cell* **2016**, *164*, 747−756.
(14) Bonjack-Shterengartz, M.; Avnir, D. The Near-Symmetry of Proteins. *Proteins: Struct., Funct., Genet.* **2015**, *83*, 722−734.
(15) Zabrodsky, H.; Peleg, S.; Avnir, D. Continuous Symmetry Measures. *J. Am. Chem. Soc.* **1992**, *114*, 7843−7851.
(16) Pinsky, M.; Dryzun, C.; Casanova, D.; Alemany, P.; Avnir, D. Analytical Methods for Calculating Continuous Symmetry Measures and the Chirality Measure. *J. Comput. Chem.* **2008**, *29*, 2712−2721.
(17) Bonjack-Shterengartz, M.; Avnir, D. The Enigma of the Near-Symmetry of Proteins: Domain Swapping. *PLoS One* **2017**, *12*, e0180030.
(18) Hub, J. S.; De Groot, B. L. Detection of Functional Modes in Protein Dynamics. *PLoS Comput. Biol.* **2009**, *5*, e1000480.
(19) Krivobokova, T.; Briones, R.; Hub, J. S.; Munk, A.; de Groot, B. L. Partial Least-Squares Functional Mode Analysis: Application to the Membrane Proteins AQP1, Aqy1, and CLC-ec1. *Biophys. J.* **2012**, *103*, 786−796.
(20) Köpfer, D. A.; Song, C.; Gruene, T.; Sheldrick, G. M.; Zachariae, U.; de Groot, B. L. Ion Permeation in K+ Channels Occurs by Direct Coulomb Knock-On. *Science* **2014**, *346*, 352−355.
(21) Brennecke, J. T.; de Groot, B. L. Mechanism of Mechanosensitive Gating of the TREK-2 Potassium Channel. *Biophys. J.* **2018**, *114*, 1336−1343.
(22) Haupt, M.; Blakeley, M. P.; Fisher, S. J.; Mason, S. A.; Cooper, J. B.; Mitchell, E. P.; Forsyth, V. T. Binding Site Asymmetry in Human Transthyretin: Insights from a Joint Neutron and X-ray Crystallographic Analysis using Perdeuterated Protein. *IUCrJ* **2014**, *1*, 429−438.
(23) Best, R. B.; Hummer, G.; Eaton, W. A. Native Contacts Determine Protein Folding Mechanisms in Atomistic Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17874−17879.
(24) Hamilton, J.; Benson, M. Transthyretin: A Review from a Structural Perspective. *Cell. Mol. Life Sci.* **2001**, *58*, 1491−1521.
(25) Schreiber, G.; Southwell, B.; Richardson, S. Hormone Delivery Systems to the Brain-Transthyretin. *Exp. Clin. Endocrinol. Diabetes* **1995**, *103*, 75−80.

DOI: 10.1021/acs.jpca.8b06843
J. Phys. Chem. A 2018, 122, 7924−7930