We thank all four reviewers for their valuable comments and criticisms. We have improved our manuscript along the lines pointed out, modifying the parts that were unclear and adding further content where necessary. Below is our detailed response to each point that was raised, together with a list of the changes in the manuscript. We hope that our submission in its present form is sufficiently clear and accurate to merit publication in Physical Biology.

1. Response to the comments of the first referee:

The referee liked the use of Hermite polynomials for the PDF that includes mode coupling and anharmonicity, but raised 3 points which we try to clarify below:

Firstly, the anharmonicity and mode-coupling contributions should be carefully observed via the free energy surfaces for various kinds of modes, such as "anharmonic", "quasiharmonic", and "harmonic" modes, without being limited to only mode 1 and 2. Then the entropy should be calculated separately. For "anharmonic" and "quasiharmonic" modes the resulting entropy may be dominated by the anharmonicity and the modecoupling.

We tried to clarify the role of various modes in our approach in the new paragraph that we added to the end of the Conclusion section (please see our response to the second referee).

Secondly, I doubt whether the MD simulation time (here, 900 ps) will be enough for estimating the higher-order polynomials. Much longer simulation should be performed to examine whether the derived polynomials are well converged.

Bunu Deniz yapacak.

Thirdly, it is not clear how the Hermite polynomials are calculated by eq. 5. The harmonic approximation is applied for the PDF in eq. 5? And how many terms (i.e., rank) are summed up for deriving f1 and f2?

Eq.(5) demonstrates that the coefficients to be determined are the expectation values of the various Hermite (bi)polynomials. In order to clarify how these expected values are estimated, we have added a paragraph to the end of Section 3.1 which reads:

"The average values of the Hermite polynomials given in Eq.(5) are calculated by averaging over 7971 snapshots of the fluctuation trajectory. A maximum rank of 17 for the Hermite polynomial tensors was observed to be sufficient for the convergence of our results."

2. Response to the comments of the second referee:

The referee recommended the publication of the manuscript in Physical Biology, but raised a point which we try to clarify below:

There are other approaches used to compute free energy landscapes that naturally incorporate anharmonicity, specifically free energy landscapes determined in terms of lowest principal components or dihedral principal components, along which the protein makes relatively large displacements. This has been done, for example, by Stock and coworkers (J. Chem. Phys. 128, 245102, (2008)) and Maisuradze and Leitner (Proteins: Structure, Function and Bioinformatics 67, (2007)) and Maisuradze et al (Phys. Rev. Lett. 102 238102 (2009)). It would be useful if the authors would compare and discuss the merits of their approach to determine free energy landscape of a protein near its native structure against other approaches such as these.

We thank the reviewer for pointing out the three references that bear resemblance to ours in that each treat large scale motions of the protein. We now included the following explanations and comparisons at the end of the Conclusion section and the papers in the References:

"Anharmonicity of protein motions has been addressed in several earlier papers by means of principal component analysis [3,15-17]. Hayward *et al.* [3] used a quasiharmonic approximation to the anharmonic energy surface having multiple minima, where the fluctuation distribution function was treated as a multivariable Gaussian, with the variables being the normal modes of the molecular dynamics trajectory. If hopping between different minima is suppressed, then the quasiharmonic approximation reduces to normal mode analysis. In the present paper, the quasiharmonic approximation would obtain if the sum in the brackets in Eq.(2) were set to zero, subject to the conditions given by Eqs.(1&3).

The higher-order moments in Eq.(2) include terms necessary to go beyond the quasiharmonic approximation. Mausiradze and Leitner [15] applied the principal component analysis to a tetrapeptide and analyzed the free energy surface using only the largest two eigenvectors obtained from the dihedral angle space. Their analysis contains the coupling effects between the first two eigenvectors. In the present work we also expressed the free energy surface keeping mode pairs, but contrary to the the treatment of Ref.[15], all pairs are included. Ref.[15] also discussed the problems with sampling convergence based on the first two mode analysis. In our case **(Denizin dun anlattiklarina uyan birsey ekeleyebilir miyiz buraya, veya Denizin argumanlarini buraya koyabiliriz.**

Maisuradze *et al.* [16] investigated the folding and unfolding of the B domain of staphylococcal protein by a coarse-grained principal component analysis and showed that while a one or two dimensional free energy landscape is sufficient for describing folding and unfolding, it may fail in describing the stability of the native state. Their

work considers the folding/unfolding of a protein, and therefore deviates from the present work that focuses on fluctuations about the native state. Nevertheless, their observations are relevant, and the need to go to higher dimensions is pertinent. Similarly, Altis *et al.* showed that a five dimensional landscape in dihedral space is necessary to properly characterize the free energy landscape [17].

The main difference of the present work from past studies above is that, by means of a Hermite expansion we take into account all of the principal components, in contrast with the restricted numbers considered in [15-17]. The constraint on the Hermite series expansion is twofold: maximum tensor rank and the order of coupling. A cutoff on the maximum rank can be made as large as computationally possible, since their evaluation is straightforward. The order of coupling (two-body, three-body, etc) is more subtle. In the present paper, we treated the first- and second-order couplings only. Including the higher-order terms is a computational challange. Instead, we supplemented our results by an alternative nonparametric "kernel density" estimation method that in effect considers all orders, but does not allow a separation of elastic, anharmonic and mode-coupling contributions to the free energy. Our results show that the difference between the Hermite representation at the second-order and the KDE do not show marked differences from each other."

We added the following references in the revised version:
[15] Maisuradze and Leitner (Proteins: Structure, Function and Bioinformatics 67, (2007))
[16] Maisuradze *et al.* (Phys. Rev. Lett. 102 238102 (2009))
[17] A. Altis, M. Otten, P. H. Nguyen, R. Hegger, G. Stock (J. Chem. Phys. 128, 245102, (2008))

3. Response to the comments of the third referee

The referee strongly supported the publication of this paper to the Physical Biology without any comments for clarification.

4. Response to the comments of the fourth referee

The referee found the manuscript suitable for publication in Physical Biology. He added three comments to which we respond below:

The authors define initially fluctuations as deviations from equilibrium positions, but then use in practice average atom position in the simulation as an estimate of the former. I would argue that the latter is in fact the correct definition, as it makes possible the analysis of "double well" cases, such as the one found for crambin (see fig. 3). Indeed, in case of asymmetric wells, average and equilibrium position do not correspond.

We thank the referee and agree with her/him that the the proper expression should have been the "mean positions" of the atoms, rather than "equilibrium". This has now been corrected accordingly in the text and the superscripts have been changed from R^{eq} to R^0 in order to avoid confusion.

The authors find that higher order computation of non-harmonic corrections increases the fluctuation entropy. Is this a general rule? Could this be related to the "double well" feature of the slowest mode?

This, in fact, is provably so, since .. (Deniz). We have now clarified this point also in the manuscript..

The eigenvalues of the covariance matrix are just mentioned and never appear in any equation. It could be useful for the general reader to show how they actually enter the game (e.g. in the evaluation of entropy for the pure harmonic case). Moreover: could one state that mode coupling is relevant for mode pairs chosen among the slowest modes?

The normalization of the normal modes' standard deviation to unity is done by stretching the normal axes by an amount proportional to the inverse square-root of the eigenvalues. From this point on, they are invisible in our formulation except for determining the ordering of the modes: the larger a mode's eigenvalue (the smaller the spring constant) the smaller is the rank.

As for the fluctuation entropy, note that in a harmonic system (f_0), each mode contributes equally to the free energy irrespective of its eigenvalue. The referee's comment is valuable, since it stresses that the different contributions from different eigenvalues is a feature that is born out of non-harmonic effects. In fact, the anharmonic contributions are most prominent for the modes with the largest eigenvalues, however finding a quantitative relationship appears to be a highly nontrivial task. We now make this explicit at the end of the first paragraph in the Results section.