ACS OMEGA

Article

http://pubs.acs.org/journal/acsodf

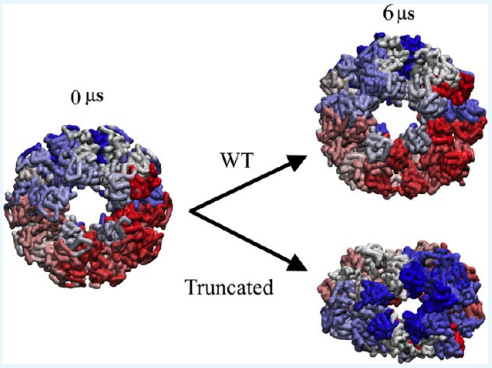
Investigation of Core Structure and Stability of Human Pyruvate Dehydrogenase Complex: A Coarse-Grained Approach

Samira Hezaveh, An-Ping Zeng, and Uwe Jandt*

Institute of Bioprocess and Biosystem Engineering, Hamburg University of Technology, Denickestr. 15, 21071 Hamburg, Germany

Supporting Information

ABSTRACT: The human pyruvate dehydrogenase complex (hPDC) is a large macromolecular machine, and its unique structural and functional properties make it a versatile target for manipulation aiming for the design of new types of artificial multienzyme cascades. However, model-based and hence systematic understanding of the structure−function relationship of this kind of complexes is yet poor. However, with new structure data, modeling techniques, and increasing computation power available, this shortfall is about to cease. Recently, we have built new atomistic models of E2/E3BP, the two subunits of the massive hPDC core. Here, we present developed coarse-grained models of the same, on the basis of which we built and simulated the full core of hPDC, which is so far the first simulation of the catalytic core of any member in the branched-chain α-keto acid dehydrogenase complex family. We explored the stability of two previously proposed substitutional models of hPDC core: 40E2+20E3BP and 48E2+12E3BP. Our molecular dynamics simulations showed a higher stability and sphericity for the second model. Our core's radius of gyration was found to be in strong agreement with previously published experimental dynamic light scattering (DLS) data. Finally, in the direction of our experimental effort to design active minimized complexes, we simulated C-terminal truncated E2/E3BP cores of different lengths, which clearly showed the instability of the core assembly and symmetry due to subunit separations. This correlated very well with the findings of our experimental investigations by analysis of DLS data for variable truncation lengths. The use of polarizable water and an increased total ion concentration did not lead to a substantially different initial stability of the truncated mutants compared to that of standard MARTINI water; however, a different rearrangement behavior of the fragments was observed. The obtained structure models will serve as a basis for full complex simulations in the future, providing the possibility for the model-assisted targeted manipulation of such a complex enzymatic system.



INTRODUCTION

The pyruvate dehydrogenase complex (PDC) as a large macromolecular machine is known to possess properties like coupled metabolic channeling within multistep reactions, including the reactivation of diverse co-factors. These properties along with its ability to self-assemble with a high structural stability make this nanomachine unique. Mammalian PDCs, for example, human PDC (hPDC), are assembled from four subunit components: E1, pyruvate decarboxylase; E2, dihydrolipoamide acetyltransferase; E3, dihydrolipoamide dehydrogenase; and E3BP (binding protein).[1] E2 and E3BP form the inner core of the hPDC with a dodecahedric structural shape and in contrast to the procaryotic PDCs that show cubic core structures formed by 20 identical building units (E2 trimers), the hPDC core is formed by both E2s and E3BPs with at least two distinct populations of trimeric building blocks, as suggested recently from substitutional model, for which the stoichiometry is not yet fully resolved.[2−4]

The E2/E3BP core structure plays a key role in the PDC activity by means of "substrate channeling", which means it bridges E1 and E3 via its lipoyl domains known as swinging arms by visiting the active sites of E1, E2, and E3, respectively. In this type of enzymatic mechanism, the reaction cascades benefit in several ways, such as less diffusion loss, less inhibition by intermediates, protection of intermediates, and recycling of co-factors. Overall, because of these properties along with the ability to self-assemble and its modularity, the E2/E3BP core and swinging arm-based channeling mechanism becomes a high-value target for the engineering of novel multienzymatic reaction cascades.[5−7]

Because of the importance of the core in this type of enzyme complex, this study focuses on the model-based assessment of core assembly, structure, and dynamics to enable targeted manipulation as well as create a foundation for the synthesis of artificial complexes in the future.

PDC cores in general have a unique assembly. Trimeric units can bind noncovalently via hydrophobic interactions called anchoring effect that occurs mutually by anchoring of a C-terminal from one monomer to a hydrophobic pocket of the other one. In our previous work,[8] we presented new models of

Received: November 11, 2016
Accepted: March 6, 2017
Published: March 23, 2017

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

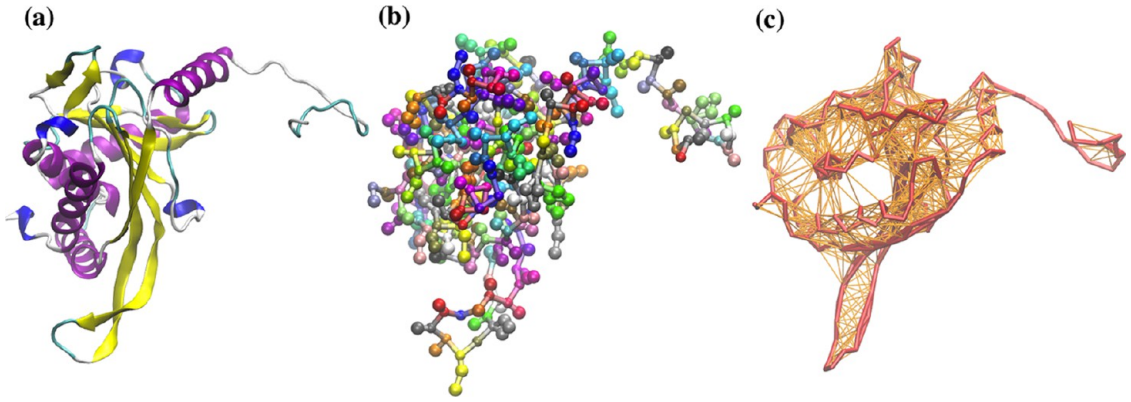


Figure 1. Atomistic model of E2 (a) vs its CG MARTINI model (b) and a bond representation (red) of the CG model along with its elastic network (orange) (c). The residues and bonds in the CG model (b) are colored on the basis of Residue ID in visual molecular dynamics (VMD).

E2 and E3BP generated and validated at atomistic level. We explored the stability of the subunit models and also studied the anchoring effect for their dimers. Our results showed a strong hydrophobic interaction in the wild-type (WT) dimers between the C-terminal of one monomer and the hydrophobic pockets of the other one. Here, we use the previous atomistic models as a reference for higher-scale modeling of the monomers as building blocks of the whole core as a second step toward the design of hPDC enzyme complex for studying its assembly, mechanism, and dynamics.

As mentioned, unlike other species, the hPDC core is formed by both E2 and E3BP subunits. On the basis of several experimental studies, such as analytical ultracentrifugation studies, small-angle X-ray scattering (SAXS), and small-angle neutron scattering (SANS) solution structures combined with cryoelectron microscopy reconstructions of rhE2/E3BP (recombinant human E2/E3BP) assemblies, the existence of a substitutional model was strongly supported.[2−4] This model suggests that E3BPs are distributed among E2s within the 60-meric core structure itself. However, it is still not clear how many E3BP are located in the core. Two specific models have been recently proposed: 40E2/20E3BP and 48E2/12E3BP. In this article, these models are called model 1 and model 2, respectively.

The subunit composition of the hPDC core is critical in its efficient functioning. In other words, the numbers and positions of E3BPs among E2s is a key on the overall core geometry that can influence its stability, catalytic efficiency, and pyruvate dehydrogenase kinase (PDK)-mediated regulation.[3,9] Therefore, it is very important to investigate the structural properties of the hPDC core and the accuracy of the proposed models. In this work, we studied these phenomena by molecular dynamics (MD) simulations at the coarse-grained (CG) level using MARTINI force field.[10] The CG modeling allows exploring these systems on different orders of magnitude in length and time to better understand the behavior, dynamics, and the molecular interactions from monomeric and trimeric scales to the whole 60-meric core for studying its overall stability.

This article is organized as follows. The details of the force field used at the CG level are given in the Methods section. In the Results and Discussion section, we present and validate the CG models of E2 and E3BP subunits built on the basis of their atomistic models. At this level, the results for the simulation of trimeric units of hPDC are presented, and the WT full-core models are built based on them. Finally, the effect of the C-terminal truncation of various lengths on the core stability is shown from our simulation results and in comparison with our previously published experimental findings.[11]

RESULTS AND DISCUSSION

CG Modeling. New CG MARTINI models of E2 and E3BP were built on the basis of the previous homology models.[8] Figure 1 shows the atomistic model of E2 versus its CG model.

E2 and E3BP Single Monomer. After building the MARTINI models of the E2 and E3BP monomers, to validate them and assess their stability, we analyzed the root-mean-square deviation (RMSD) of the backbone (BB) beads. As shown in Figure 2, the CG model result is very well comparable to the atomistic model results in 240 ns, and the RMSD reached a plateau very quickly. The RMSD values calculated after 50 ns for E2 and E3BP CG model were 0.24 ± 0.004 and 0.27 ± 0.003 nm, respectively, which are very similar to the previous atomistic results.

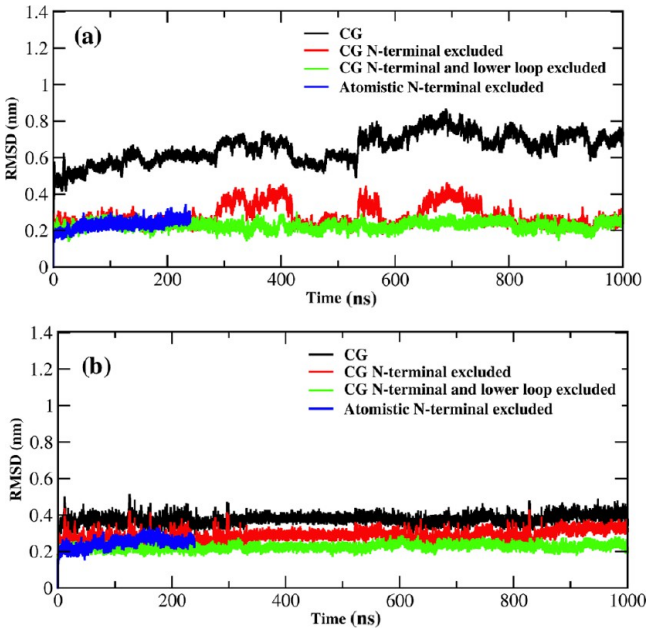


Figure 2. RMSD of CG model vs the atomistic results for E2 (a) and E3BP (b). In case of atomistic RMSD of E3BP, along with N-terminal, α1, which was partially unfolded and led to a higher RMSD,[8] was also excluded from these calculations.

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

The simulations were extended to 1 μs to confirm their stability. The RMSD was calculated for the total and N-terminal-excluded versions, as shown in Figure 2. The results showed good stability of both models for the N-terminal-excluded calculation. However, in case of E2, some fluctuations from the average value were observed after around 300 ns, and further investigation showed that the lower flexible loop is strongly responsible for that, as in the monomeric state, this region is quite flexible (shown from atomistic results[8]).

By exclusion of this region, the extra fluctuations were eliminated from the RMSD, and the results showed a stable trend during the course of the simulation. The RMSD of N-terminal- and lower-loop-excluded versions of both proteins reached a plateau, and the values turned out to be 0.23 ± 0.003 and 0.23 ± 0.004 nm for E2 and E3BP, respectively.

The root-mean-square fluctuation (RMSF) values for the BB beads after 50 ns and up to 240 ns were also analyzed. The results showed a very good agreement to the atomistic ones, as shown for both models in Figure 3. Because of the elastic

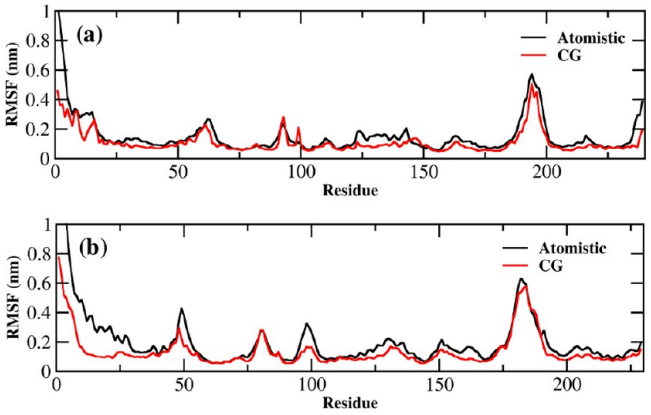


Figure 3. RMSF plot of CG MARTINI model vs atomistic model for E2 (a) and E3BP (b).

network, the CG models showed a better stability in some regions, even better than the atomistic ones. However, for both models, despite the elastic network, there are some regions, such as the lower-loop areas, which are highly flexible as those in the atomistic models. Moreover, the model effectively maintained the stability of the active site in E2 (residues from 212 to 216). Because the N-terminal loops were excluded from the elastic network in both models, it is still very flexible especially in case of E3BP. These model validations and their agreement to the atomistic results proved the reliability of the CG models.

Radius of Gyration. The radius of gyration (R_g) as a function of time is presented for both subunits in Figure 4 to compare with the atomistic results. R_g of monomers were calculated by exclusion of the flexible N-terminal. The values turned out to be 1.78 ± 0.002 and 1.79 ± 0.003 nm for E2 and E3BP, respectively, which are very similar to the atomistic results in 240 ns.

Further screening of the R_g values up to 1 μs shows that within a range of several hundred nanoseconds (in this case, ~300 ns) there is a fluctuation or variation cycles of the R_g values. These changes were related to the lower loop, as shown from the RMSD and RMSF calculations. However, by exclusion of this region from the calculations, we gained the stable values

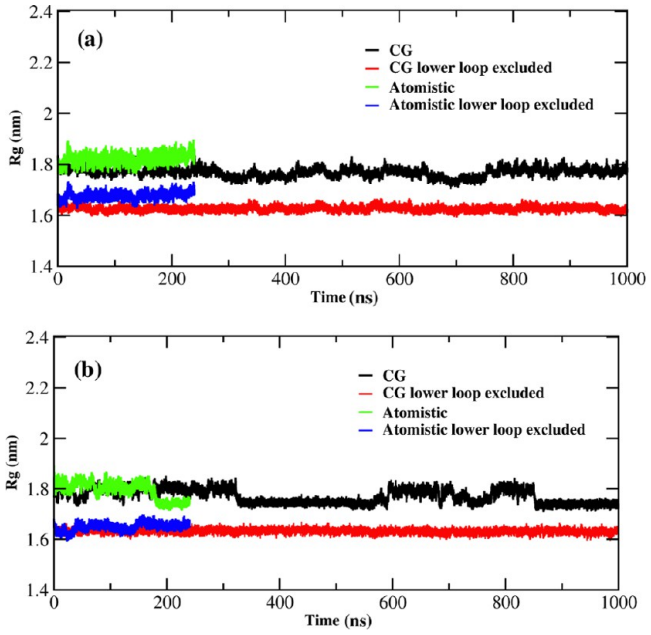


Figure 4. Radius of gyration of CG MARTINI models vs the atomistic ones for E2 (a) and E3BP (b). In all cases, the N-terminal was excluded from the calculations.

of 1.66 ± 0.001 and 1.63 ± 0.001 for E2 and E3BP, respectively.

E2 and E3BP Dimers. As the anchoring effect has been identified to be a main factor in the self-assembly of the core, we studied the C-terminal interactions in the WT E2 and E3BP dimers. This was done as an assessment test to check whether the interactions at the atomistic level are reproducible at the CG level. The simulation results for WT E2 dimers showed that the initial orientation of the monomers with respect to each other is quite stable and the C-terminal anchoring is well maintained (Figure S1). Similar results were observed for the E3BP dimer simulation, as shown in Figure S2.

The same simulation setup was done for E2-t8, that is, by eliminating eight amino acids from the C-terminals of both monomers. As seen, this truncation led to the dimer separation. The initial orientation of the monomers was not conserved, and they started to separate and reorient themselves because of the absence of C-terminal anchoring effect (Figure S3). In case of the E3BP-t7 dimer system, a separation occurred from both terminals (Figure S4).

Figure 5 shows the average distance versus time between the two monomers for WT and the truncated versions. Figure 5a shows the distance between the residue Glu127 located in the upper region of the hydrophobic pocket and the C-terminal residue Leu231 of E2-t8. The distances between the chosen residues in the WT version were 2.08 ± 0.05 and 2.14 ± 0.07 nm, which are in good agreement with the atomistic results, and it indicates that our CG model is capable of reproducing the same properties and it is quite reliable. On the other hand, in case of the E2-t8 system, a rapid separation of monomers from one side was observed, and the average distance reached a maximum value of 3.45 ± 0.09 nm within only 100 ns. A similar analysis for WT E3BP and E3BP-t7 was performed by calculating the average distance between residue Gln115 in the hydrophobic pocket region and the C-terminal residue Leu222 of E3BP-t7. From Figure 5b, same as E2, in the WT system, both of the monomers are strongly bound to both

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

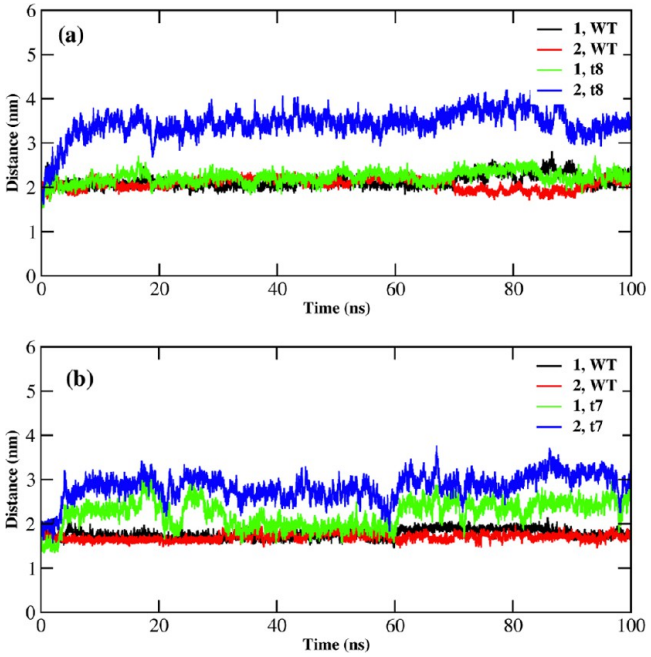


Figure 5. Average distance between Glu127 and Leu231 in E2 (a) and Gln115 and Leu222 in E3BP (b) from both sides (1 and 2) for truncated and WT versions.

pocket sides (with average distances of 1.72 ± 0.09 and 1.77 ± 0.05 nm), which are also very similar to the atomistic results. Upon truncations, the distances started to increase especially from one side, and within 100 ns, it reached the maximum value of 3.05 ± 0.20 nm.

Further, the simulations were extended to 1 μs to check the stability of WT dimer. For both E2 and E3BP, after 100 ns, a disorientation of monomers was observed while the pocket anchoring was still maintained up to 200 ns and then a separation occurred as the distance constantly increased. From this, it may be inferred that the specific concentration of monomers as well as their location in a 60-meric core could conserve the original orientation without which the intermonomeric interactions may not get stabilized by the individual dimers.

Trimer Simulation. Trimers were simulated to study the monomeric interactions and the stability of the WT monomers in a trimeric state. As mentioned before, the presence of E3BPs among E2s leads to the formation of different types of trimers in their core models. In model 1, there are 20 E3BPs, which can form only one type of trimers: 2E2/1E3BP. On the other hand, in model 2 with 12 E3BPs, two types of trimers can be formed: 2E2/1E3BP heterotrimers and 3E2 homotrimers. The simulations were done for the two possible types for 2 μs. Figure 6 shows the visual presentation of the trimers. The 2E2/1E3BP trimer was built by replacing one E2 with an E3BP. The radii of gyration for the 3E2 and 2E2/1E3BP trimers turned out to be 2.51 ± 0.003 and 2.50 ± 0.015 nm, respectively.

Trimers were analyzed for stability by RMSD and RMSF calculations. The N-terminal-excluded RMSDs were plotted for individual monomers in each trimer, as shown in Figure 7, and the average RMSDs of E2 and E3BP were calculated as 0.19 ± 0.001 and 0.23 ± 0.008 nm, respectively. Interestingly, these values are similar to those from monomeric simulations with N-terminal- and lower-loop-excluded versions (Figure 2), as the lower loop in the trimer became perfectly stabilized.

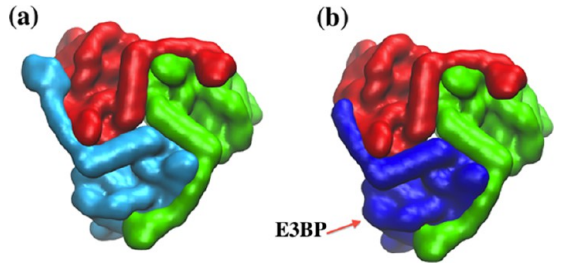


Figure 6. 3E2 (a) and 2E2/BP (b) trimer presentation using quick surf in VMD.

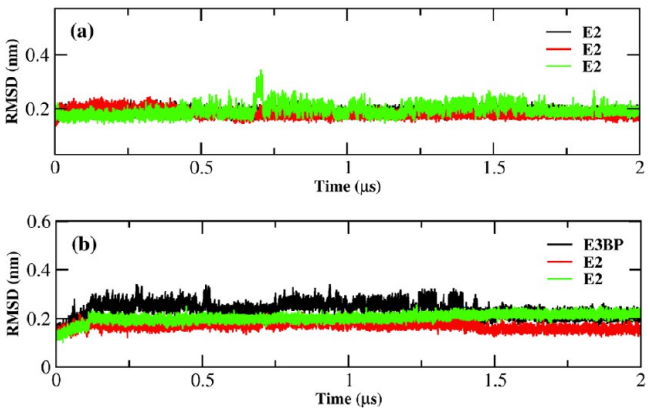


Figure 7. N-terminal-excluded RMSD of monomers for 3E2 (a) and 2E2/BP (b) trimers.

Figure 8 shows the average RMSF calculated after 50 ns for E2 and E3BP in the trimeric form and in comparison with

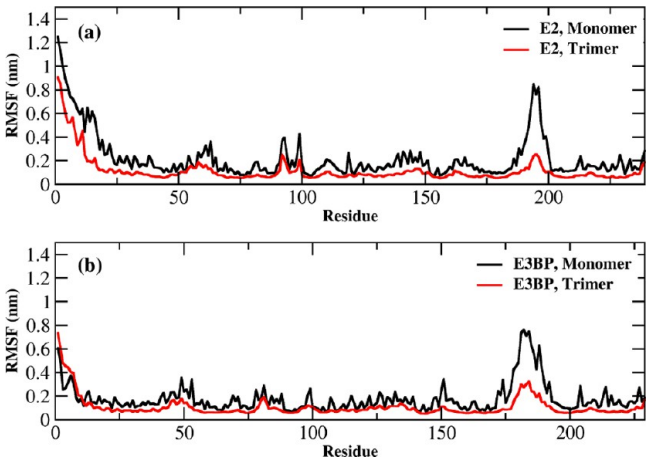


Figure 8. RMSF of the atomistic model vs the CG MARTINI model for E2 (a) and E3BP (b) in monomeric and trimeric systems.

those from monomeric systems. It can be seen that there are several flexible regions in the interface, which were stabilized in the trimerization process, such as the lower-loop region and the upper β hairpin (residues 87−98). The latter forms the upper-part region of the hydrophobic pocket. The N-terminal is also more stable except the flexible loop, which is now almost half its length in its free monomers. Moreover, the active-site region (212−216), which is also located in the trimer interface, showed a better stability (Figure 8a) because of the higher coverage from the adjacent monomers and the interaction between the upper β hairpin with the end part of α1. In the

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

same way to evaluate the active-site water accessibility, we calculated the radial distribution function (RDF) for $BB_{active\text{-}site}$-W_{beads}, that is, the BB beads with respect to the water beads.

Figure 9 shows the RDF in both monomeric and trimeric states, that is, 3E2 trimer. The RDF of E2 active site in the

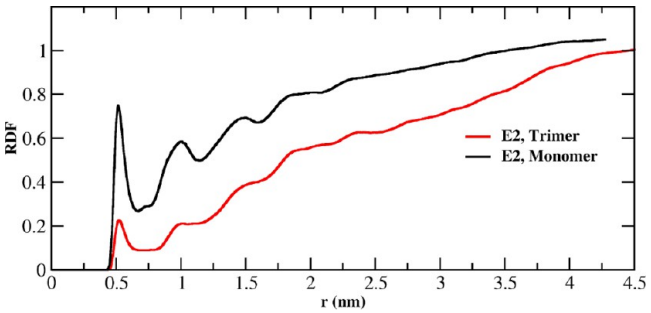


Figure 9. RDF plots of $BB_{active\text{-}site}$-W_{beads} for E2 in monomeric and trimeric systems.

monomeric form is higher, as the active site is exposed to more water density than in its trimeric form.

hPDC Core Simulation. *WT.* At this level, after an extensive validation of our models, we built up the full core of hPDC. We constructed the two models suggested by the substitutional model, as the numbers of E3BP distributed among the E2s are still not experimentally clarified. We simulated both models (40E2/20E3BP and 48E2/12E3BP) to investigate their structural stability and compatibility. Figure 10 depicts the two full-core models. The simulations were performed for 6 μs to characterize their structure and dynamics.

First, to check the equilibration of the systems, the RMSD of the whole WT core was calculated for the BB beads, as shown in Figure 11. It is important to note that the RMSD values for the whole core models are significantly greater than the RMSD value of the single monomers. This is due to the reorientation of individual trimers with respect to each other, which leads to a

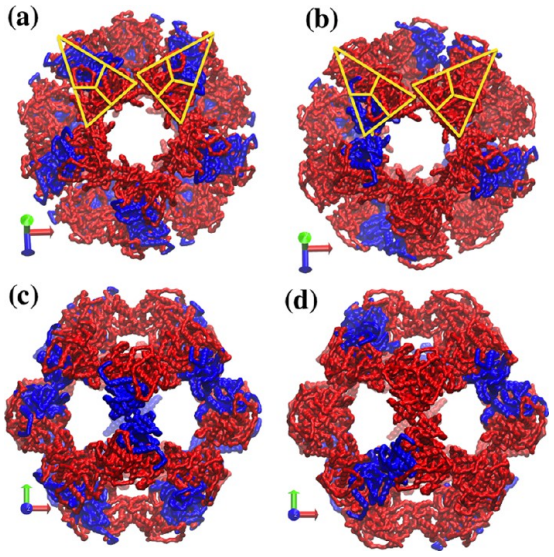


Figure 10. Substitutional core models of hPDC for (a, c) model 1 (40E2/20E3BP) and (b, d) model 2 (48E2/12E3BP) and the positioning of E3BP pairs (blue) among E2s (red). Yellow triangles show the location of different types of trimers in each model. Depth cueing was activated in VMD for clarity.

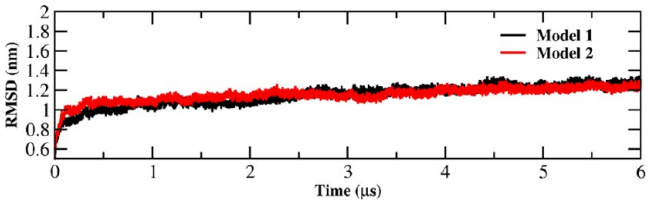


Figure 11. RMSD trend vs time for the full 60-meric core for the two substitutional models.

slight expansion of the core (will be seen later in the R_g results). RMSD versus time for model 1 reached a plateau after about 3 μs with an average value of 1.21 ± 0.07 nm, whereas model 2 reached it within 1 μs with an average value of 1.20 ± 0.05 nm.

The autocorrelation functions, $C(t)$,[12] of RMSD for both the models were calculated (Figure 12) to assess the relaxation time

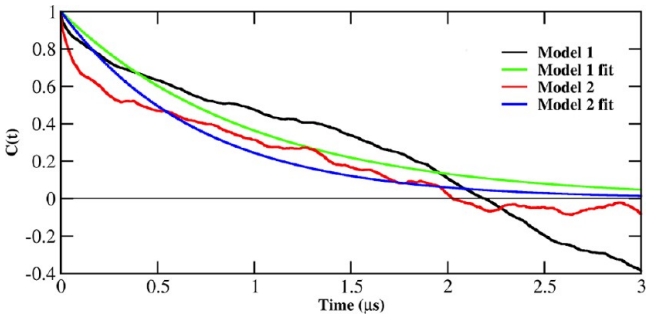


Figure 12. RMSD autocorrelation functions of models 1 and 2 and the single-exponential fitting to them.

(τ) of the systems. As shown in the figure, the relaxation time was evaluated by fitting of the $C(t)$ to a single-exponential function

$$C(t) = e^{-t/\tau}$$

The relaxation of model 1 seems to be slightly slower compared to that of model 2; the τ values for models 1 and 2 were calculated as 1 and 0.7 μs, respectively. It seems that a different structural arrangement of the core, that is, by increasing the number of E3BP from 12 to 20, leads to a slower dynamics and a longer relaxation time.

Further, the R_g values of WT full cores were also calculated as shown in Figure 13. From the simulations, the R_g values of both models started to increase at the beginning of the simulation with a slightly larger value for model 2. Both models had close values of R_g up to about 1 μs, but after that, the corresponding values for model 1 started to decrease gradually but constantly

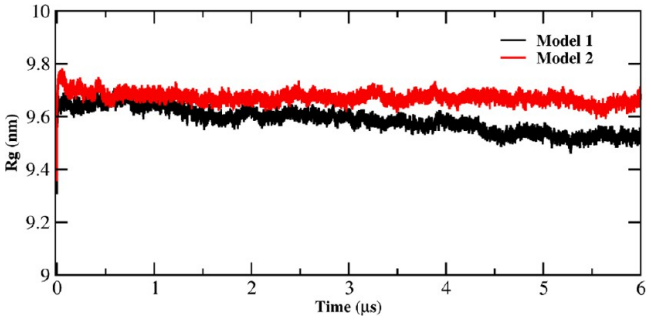


Figure 13. R_g plot of full 60-meric core for both substitutional models.

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

from 9.65 ± 0.43 nm in 1 μs to 9.52 ± 0.003 nm within 6 μs. On the other hand, model 2 reached an average value of 9.67 ± 0.004 nm right after 1 μs and kept the constant trend until the end of the simulation. This value is in good agreement with the experimental value of 10.20 nm for the same from SAXS and SANS measurements.[9]

Moreover, the autocorrelation calculated for R_g led to a much lower relaxation time for model 1 than for model 2, as shown in Figure 14. The relaxation time turned out to be 1 and 0.17 μs for models 1 and 2, respectively.

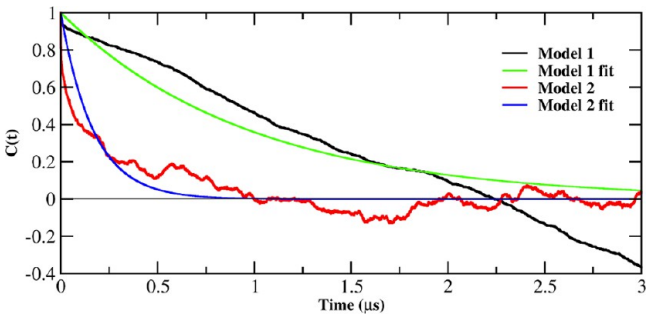


Figure 14. Autocorrelation of R_g and single-exponential fitting to it for both of the substitutional models.

This remarkable difference in the relaxation time of R_g might be explained by the fact that substituting more E2s with E3BPs in model 1 would decrease the trimer's positioning compatibility and thus reduce the flexibility of the whole core, which makes the system spend much more time reaching its equilibrium state. Previously, a mathematical study was also conducted to investigate the core structures and their flexibility by altering the population of E3BPs among E2s.[3] They modeled the possible E2/E3BP core arrangements by performing a systematic variation of the E2 homotrimer and the 2E2/1E3BP heterotrimer populations. Their investigations suggested that model 1 has low flexibility in the core assembly, whereas model 2 can allow the maximum flexibility as well as provide an appropriate balance for binding of E1 and E3 enzymes on the spherical core structure of the E2/E3BP, and it has the potential to moderate the flexibility of the E2 and E3BP lipoyl "swinging arms" and may facilitate PDK movements around the core.[3]

We investigated further the structural properties of the core models and their stability. The structures of models 1 and 2 within 6 μs are shown in Figure 15.

Visual inspection, however, shows that the cores do not perfectly maintain their symmetry, especially core model 1, for which the structure deviated from the original shape. Its size started to shrink from 0.2 μs, as it was seen from the R_g trend, and the symmetrical shape gradually changed during the simulation time. On the other hand, model 2 showed a better stability and consistency with only a slight deviation from the perfect shape. Therefore, to quantify the extent of their asymmetry, we measured their asphericity (b) by calculating the components of their R_g along X, Y, and Z axes; the average values are shown in Table 1.

The component values show that the cores are not completely spherical but slightly elongated. The asphericity, b, was then calculated as

$$b = R_{g1}^2 - \frac{1}{2}(R_{g2}^2 + R_{g3}^2),\ R_{g1} > R_{g2} > R_{g3}$$

The asphericity values for models 1 and 2 were calculated as 4.71 and 0.78, respectively. This proves a better compatibility of positioning 12 E3BPs instead of 20 among E2s and therefore the better flexibility of model 2 than model 1.

Partial density profiles of each model are shown in Figure 16. From the figure, the peaks are where the trimers are located, and the minima in the central profile are related to those windows as shown in Figure 15. The reference profile refers to the initial arrangement of trimers from the biological assembly. From the profiles along X axis, models 1 and 2 do not undergo a remarkable overall shape change. However, model 2 maintains a density profile trend better than that of model 1, as the latter shows a reduction of density in the peak regions (around −6 and 6 nm), which may be due to the disorientation of the trimers from the reference model. For both models, an expansion was also observed from the profile edge regions (around −12 and 12 nm). Along Y axis, model 1 tends to maintain the peak regions but getting more compact (previously shown in Figure 15). Because of an increase in the compactness of the model, the density profile also shows an increase especially in the window regions as the core collapses, thereby increasing the density of these regions. This is while model 2 still expands further while keeping the reference trend better. Along the Z direction, model 1 shows a decrease in density in the central part of the profile and especially in the peak side-regions (around −6 and 6 nm) as an expansion occurs in the profile edges (around −12 and 12 nm regions). This can be due to the elongation of the core along Z axis (see Figure 15) as it gets more compact along Y axis, which leads to a decrease in density as the trimers get further apart. A drop of density in the peak regions can reflect the disorientation of the trimers from the reference model. This is while model 2 shows less expansion along Z axis. Although with this expansion the peaks get slightly apart, the overall profile trend remains fairly unchanged. In conclusion, model 2 shows a better consistency in this regard.

Double-Truncated Core. To show the importance and the effect of C-terminals on the structure maintenance of the whole cores, we performed the same simulations with truncation of eight and seven amino acids from E2s and E3BPs, respectively (referred to as "double-truncated", "dt", or E2-t8+BP-t7). First, RMSD was analyzed, whose value increased for both the models and reached about 2 nm for both models, as shown in Figure 17.

On the other hand, R_g values calculated for both the models decreased constantly to about 9 nm, as shown in Figure 18. This shows that as expected the cores collapsed and therefore the R_g values decreased. The asphericity (b) was calculated for the models, and it turned out that for both models it increased remarkably within 2 μs to 5.9 and 4.4 for models 1 and 2, respectively.

To investigate the structural properties of the double-truncated core models, we visualized the structure of the core model 2. As shown in Figure 19, the core completely lost its symmetry and spherical shape. It not only expanded and lost its original compactness but also tended to get squeezed along Y and Z axes, which was due to the separation of trimers, as shown in the figure. This proves the importance of C-terminals in the maintenance of the 60-mer structure, without which the structural geometry is lost and it leads to the formation of apparently irregularly structured agglomerates. However, the activity can remain intact, presumably depending on the specific structure or type of agglomerates. Our experimental data for the

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

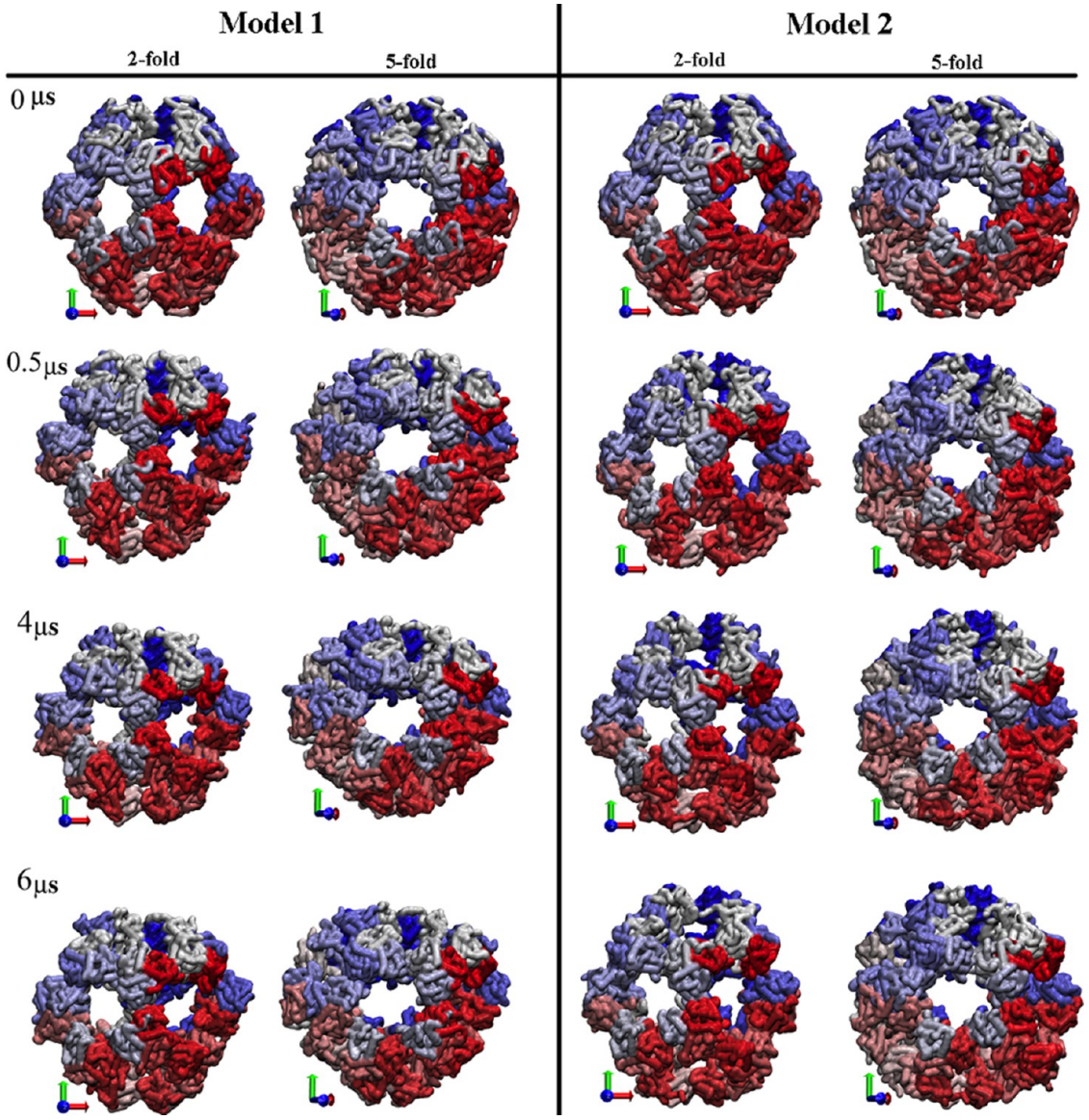


Figure 15. Structures of models 1 and 2 within 6 μs along twofold and fivefold axes of symmetry. Backbone beads were selected for visualization. Index coloring method was chosen for better clarity of the whole core with no specific coloring for E2 or E3BP.

Table 1. Components of R_g (in nm) for the Cores along *X*, *Y*, and *Z* Axes and Their Asphericity

	R_{gx}	R_{gy}	R_{gz}	b
model 1	7.7	8.0	7.7	4.71
model 2	7.9	7.8	7.9	0.78

double-truncated core showed the existence of highly active agglomerates.[11] In conclusion, a highly ordered and giant structure like the original 60-mer is not necessary for the activity of the hPDC.

The partial density profiles of each truncated model are shown in Figure 20. The figure shows that, as expected, the density trends for both models are completely lost, and both got either expanded or squeezed along different axes. For instance, for model 2, the profile expands along *X* axis but shrinks along *Y* and *Z* axes, as shown in Figure 19.

In Silico Screening of Single and Double-Truncated Variants. Starting from the findings demonstrated in the previous section, we pushed forward a complete in silico screening of different variants and compared them with our experimental data. Specifically, it was of interest to screen the effect of different truncation lengths on the core stability and to investigate what length or combination of truncations may as well lead to the core disintegration.

Full-core simulations, based on model 2 (48E2/12E3BP) with single truncations of E2, ranging from 1 to 8 residues (E2-t[1...8]), and E3BP, ranging from 1 to 7 (E3BP-t[1...7]), were performed. In addition to the already investigated double-truncation variant (E2-t8+E3BP-t7) and WT, the systems were made up of 16 comparable simulations, all of which were performed at least in triplicate to assess the significance of any observed structural changes. The radius of gyration (R_g) and RMSD were computed for all the replicates as well as their standard deviation. To analyze the results of simulation in standard MARTINI water, a simulation time range of 2.0−2.5 μs was used. All values are depicted in Figure 21.

According to the effect of truncations on the core stability, five distinct populations have been defined: (I) wild-type; (II) with E3BP-only truncations; (III) with short (1−3) E2-only

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

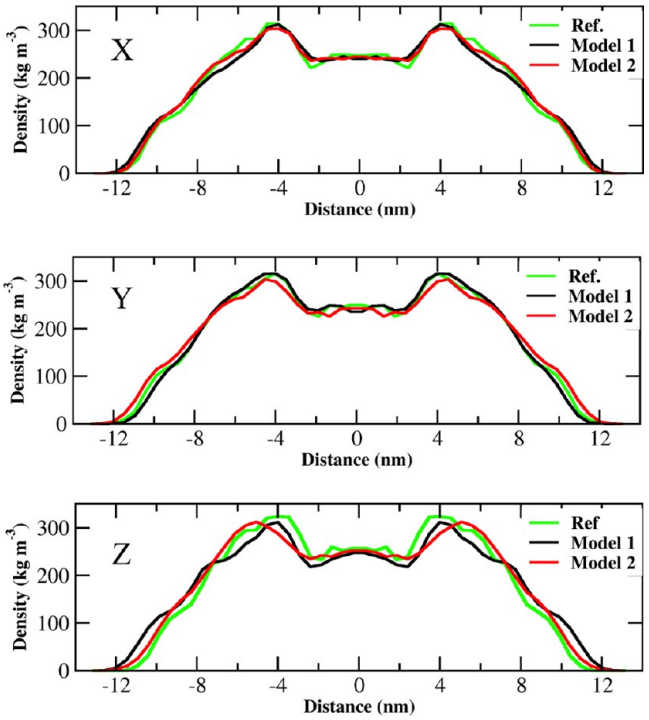


Figure 16. Density profiles of each model along *X*, *Y*, and *Z* axes after 6 μs. Distance refers to the relative position from the center.

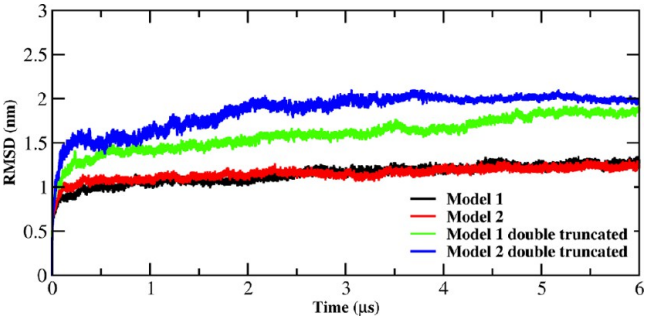


Figure 17. RMSD calculated for C-terminal double-truncated core models vs their WT.

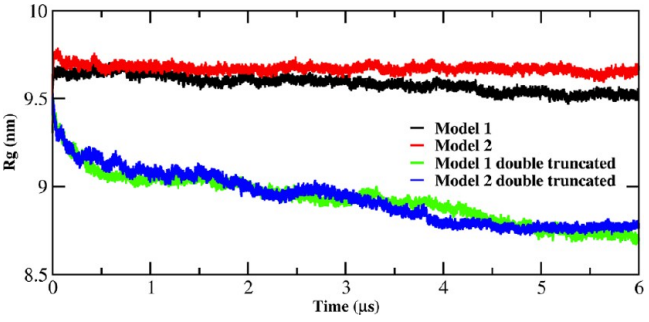


Figure 18. R_g values calculated for the C-terminal double-truncated core models vs their WT.

truncations; (IV) with longer (4−8) E2-only truncations; and (V) the double-truncated variant. It shows that generally E3BP-only truncations (II) have no clear effect on the core disintegration; however, the fluctuation of the simulation increases considerably. The same holds true for short E2 truncations (III); but, with longer E2 truncations (IV), the size

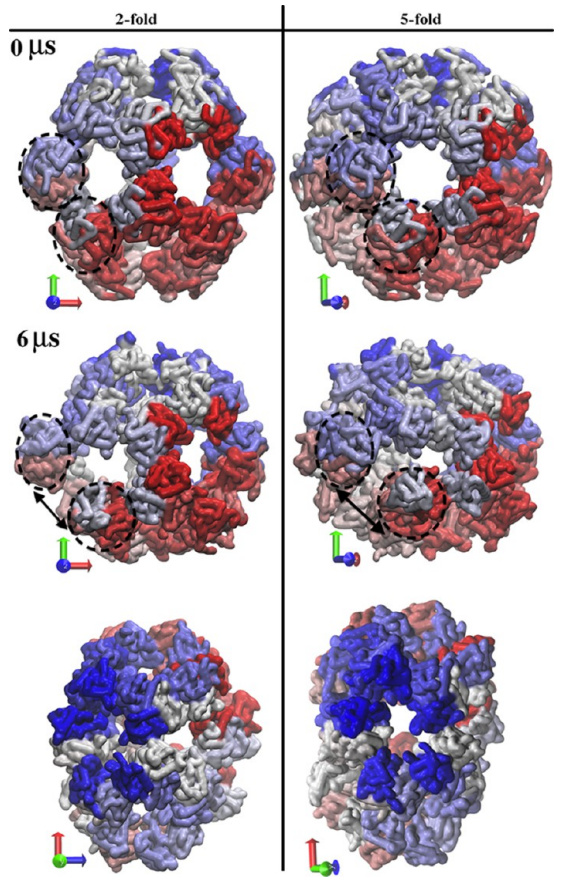


Figure 19. Structure of model 2 within 6 μs along twofold and fivefold axis of symmetry in comparison with the initial one at 0 μs. At 6 μs, two different views along the twofold and fivefold axis of symmetry are shown. Backbone beads were selected for visualization. The black arrow shows the increase of distance and the separation of trimers within the core. The index coloring method was chosen for better clarity of the whole core with no specific coloring for E2 or E3BP.

is clearly reduced (significance of all R_g values of III vs IV according to the two-sided *t*-test: $p_{IV:III} < 0.002$). This corresponds to the visual inspection. However, the core double truncation (V) led to a much more efficient disintegration with $p_{V:IV} < 0.001$. It can be concluded that only long enough E2 truncations (>3−4 residues) as well as double truncations should lead to at least partial disintegration of the core and are by far the most efficient.

Comparison with Experimental Data. The stability of hPDC cores with different C-terminal truncation variants and with or without truncation of N-terminal flexible linker arms (denoted as catalytic domain "CD" variants) has also been investigated in our laboratory, and the full workflow and results are published separately.[11] The hPDC wild-type and truncated core (fragment) sizes were most reliably determined by dynamic light scattering (DLS), yielding the distribution of hydrodynamic radii (R_h) of particles. This typically resulted in up to four populations depicted as (A) very large agglomerates ($R_h \geq 30$ nm); (B) single cores of either 60-mer or a similar-sized structure (30 nm > $R_h \geq 9$ nm); (C) small fragments, including functional subunits either as trimers or hexamers (9 nm > $R_h \geq 3$ nm); and (D) smaller fragments, for example, monomers. The available data contain variants without flexible linker regions (CD); for those, the expected R_h can be estimated directly from the model data as

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

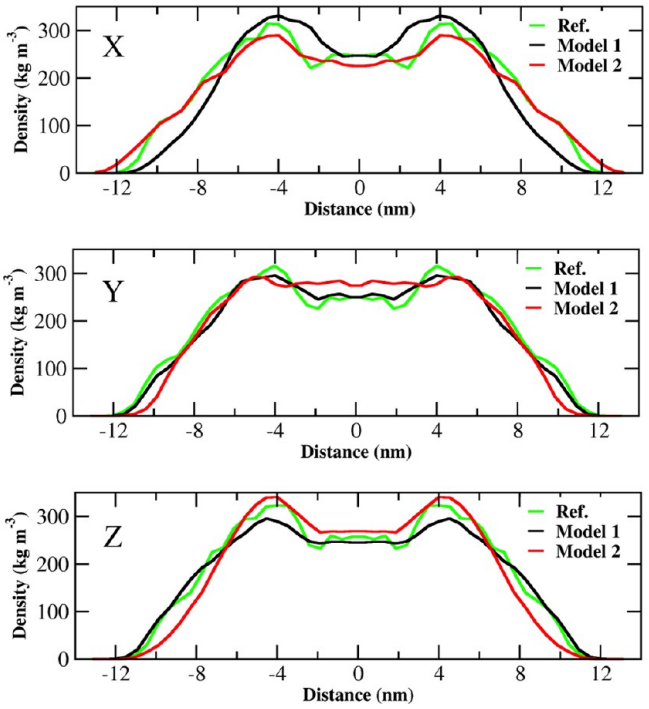


Figure 20. Density evolution of each truncated model along *X*, *Y*, and *Z* axes after 6 μs.

$$R_g \approx 0.775\, R_h$$

which holds true for globular proteins; for elongated proteins, the factor of 0.775 increases. Hence, for the full 60-mer, the expected R_h is $R_{h,model}^{60,CD} \approx \frac{9.67\text{ nm}}{0.775} = 12.5$ nm. Similarly, for the catalytic subunits, the expected radii are $R_{h,model}^{6,CD} \approx \frac{4.4\text{ nm}}{0.775} = 5.7$ nm for hexamers, $R_{h,model}^{3,CD} \approx \frac{2.5\text{ nm}}{0.775} = 3.2$ nm for trimers, and $R_{h,model}^{1,CD} \approx \frac{1.8\text{ nm}}{0.775} = 2.3$ nm for monomers.[8] With flexible arms, the same assessment is hardly possible for full 60-mers by adding approximate arm lengths to R_g (about 7−15 nm each side), assuming that they elongate equally in every direction, yielding $R_{h,model}^{60} \approx 21.5-31.8$ nm. The measured R_h of wild-type core lies within this range ($R_{h,DLS}^{wt\equiv 60} = 26.1$ nm ± 14.7%).[11] This approach is not applicable for hexamers, trimers, monomers, or any other fragment of the core because the average conformation of arms on these constructs cannot be predicted; however, it is predicted that the length is shorter.

The measured R_h of the C- and N-terminal truncated variants and their mixtures have been compared with the simulated stability, represented by their R_g in Figure 22. As mentioned, the calculated R_g is only a relative measure of the core stability and should not be considered as an absolute prediction of the resulting structure size.

The results show that the simulated core stability correlates very well (the Pearson correlation coefficient, PCC = 0.986) with the experimental "CD" variants, that is, without large flexible domains, and also with the other variants with reduced accuracy (PCC = 0.857). The latter comparison shows larger deviations (e.g., E3BP-t7), which is at least partly due to the fact that the flexible arms increase the measured DLS radii by an ill-defined amount, whereas the DLS data of the CD mutants are expected to be more accurate.

It is important to note in Figure 22 that only the wild-type (top right) and the double-truncated "dt" variant (bottom left) showed activity. It is hypothesized that the inactivity of all others might have been caused by the partially assembled core fragments that cannot form functional catalytic cores, neither single trimer/hexamer subunits nor full cores.

Polarizable Water (PW) and Ionic Strength. Because the electrostatics are not accurate using standard MARTINI water, additional simulations using MARTINI PW with the particle mesh Ewald (PME) method were performed. The radii of gyration for WT versions turned out to be very similar to the previous results in standard water as 9.89 ± 0.06 and 9.74 ± 0.04 nm for models 1 and 2, respectively. The asphericity coefficients were calculated as 4.2 for model 1 and 0.65 for model 2, which are again close to those from simulations in standard MARTINI water. The RMSD result showed that the instability of model 1 was significantly higher than that of model 2 ($RMSD_1 = 1.34 \pm 0.153$ nm vs $RMSD_2 = 0.94 \pm 0.025$ nm, $p < 0.05$).

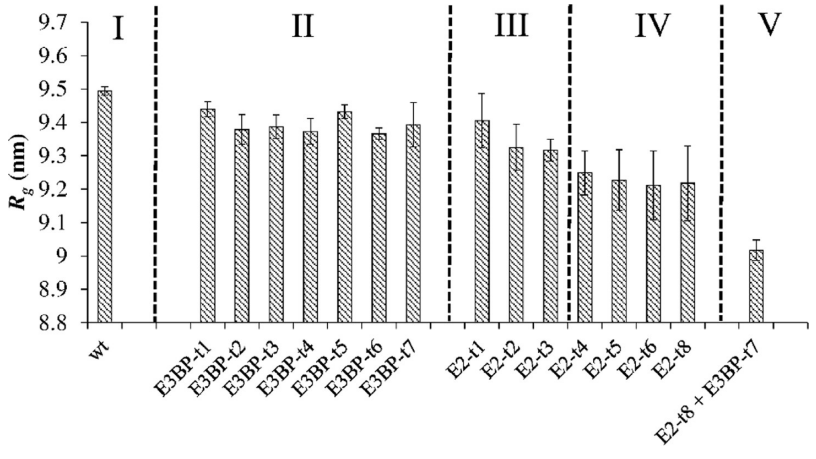


Figure 21. R_g (average and standard deviation based on triplicates) of simulated core within the simulated time range of 2.0−2.5 μs, depending on C-terminal truncations of either E2 or E3BP or combinations of both. Resulting populations are grouped as follows: (I) wild-type; (II) E3BP-only truncations; (III) E2-only truncations (short); (IV) E2-only truncations (long); (V) double truncation ("dt"). Note that the R_g values are only expected to absolutely correlate to the experimental data for WT. For all truncants (II−V), they only provide a relative degree of core breakdown or instability, as the complete disintegration or rearrangements of core (fragments) cannot be simulated this way.

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

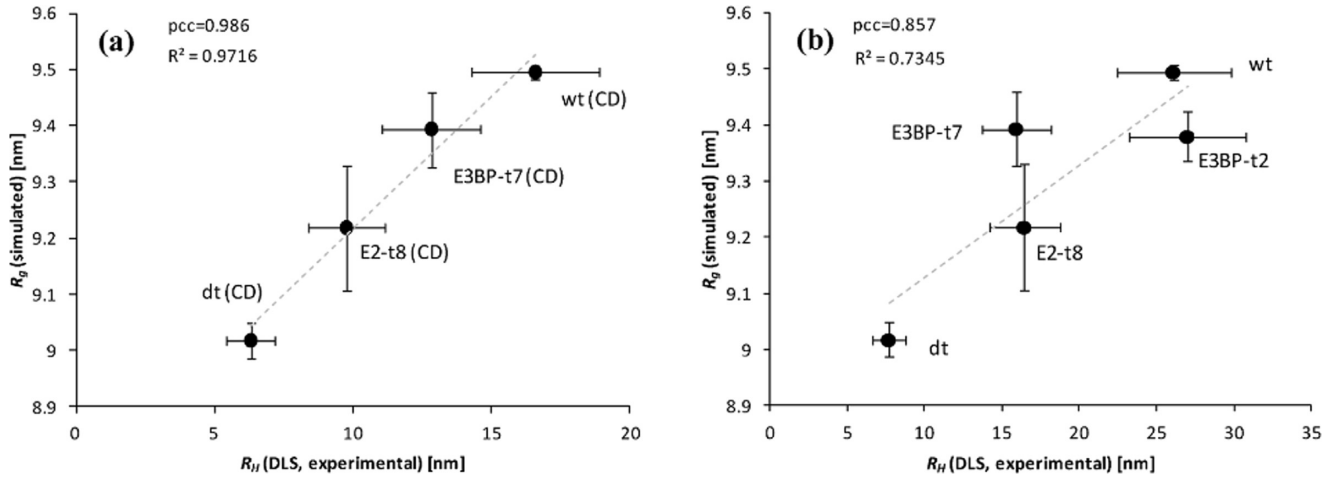


Figure 22. Comparison of simulated core stability (R_g data as of Figure 21) with experimentally obtained DLS data, providing R_h. (a) Comparison with experimentally available mutants without flexible domains (CD). For these, experimental R_h is defined accurately because (almost) no flexible regions disturb the measurement. PCC = 0.986. (b) Comparison with mutants with flexible domains. The variable conformation of flexible domains in these cases blurs the experimental results. The PCC decreases to 0.857. All the experimental data from Guo et al.[11] were rearranged for this comparison study.

The total RMSD results for different truncates are shown in Figure S5. Cl^- ions were added to neutralize the system. Because of the much higher computational costs, simulations were run for only 100 ns. From the results, it was observed that the rearrangement of the core fragments in PW is different from than in the standard MARTINI water. In other words, the strongly disintegrating cores exhibit an expansion (or dispersion) in PW rather than a collapse in MARTINI water. Because of this, RMSD was chosen as the main stability criterion instead of R_g. In line with the CG with MARTINI water findings, significant effects on core breakdown were also found for long E2 truncations (>4 or more) and the double-truncated variant.

Further, the RMSD data from Figure S5 were compared with the experimental data, as shown in Figure S6. The overall stability behavior is similar to the standard MARTINI water results (absolute PCC = 0.994 for CD variants and 0.942 for others). Because RMSD is used as a stability criterion, the PCC becomes negative. Some minor deviations can be identified, for example, with regard to the potential influence of medium-sized E3BP-only truncations. Overall, the accuracy of modeling with standard CG water for the purpose of PDC core truncation studies seems to be sufficient while being computationally faster.

As the effect of ionic strength is known to play a crucial role in protein−protein interactions, we performed another set of simulations by adding Cl^- and Na^+ ions to reach an ion concentration equal to 50 mM, which is equivalent to the upper limit from the experimental data. For these systems, radii of gyration for the WT version were calculated as 9.86 ± 0.19 and 9.67 ± 0.01 nm, respectively, and on the basis of the R_g components, the asphericity values turned out to be 2.4 and 1.2 for models 1 and 2, respectively. This again confirms that model 2 with 12E3BP is more stable and flexible than model 1 with 20 E3BPs.

The results of core stability and comparison with experimental data for different truncated versions of the core revealed that, overall, the use of different ionic strengths and the long-range electrostatic interactions (Figures S7 and S8) do not lead to a qualitatively different core breakdown behavior in this study. The significance of the truncation effects and correlation to experimental data (PCC < −0.867) is slightly worse, which is assumed to be due to the relatively short simulation times.

In summary, full-core simulations accurately allow for grouping the modified candidates and representing the findings obtained experimentally while providing the possibility of fine-tuned manipulations and more flexible screening. Furthermore, the combination of different truncations can only be reliably assessed in the fully assembled model, whereas the monomeric interaction investigations can be incomplete or misleading, for example, showing similar disintegration behaviors for single E3BP and E2 truncations, which did not reflect our experimental data with variants that had their flexible parts removed. Full-core simulations performed with PW and long-range interactions showed similar overall stability behaviors with respect to the length of the truncations, but the rearrangement (reagglomeration) of the fragments after the initial separation is strongly different from that with standard MARTINI water. Full-core and complex simulations are therefore further developed on all levels to guide future experimental research.

CONCLUSIONS

In this study, we utilized our recently obtained models of E2 and E3BP subunits of hPDC core to build higher-scale models of the full hPDC core at the CG level. This allows for exploring the dynamics and structure on a larger size and longer time scales, which is not possible at the atomistic level. We performed an extensive validation of the CG model by comparing the results obtained to those from the atomistic level as well as the available experimental data. The extension of the simulation time and scale allowed us to study different properties of the hPDC from monomeric to 60-meric levels. At monomeric scale, the protein interactions and the anchoring effect were reproducible at the CG level. Overall, both subunit models showed very good structural stability and consistency with those from the atomistic level. Further, the homotrimers and heterotrimers were modeled, and we built up for the first time the full WT 60-meric core of hPDC. Two core models

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

(40E2/20E3BP as model 1 and 48E2/12E3BP as model 2) are suggested from the previous experimental data, and we explored their stabilities using different water models (standard MARTINI water (W) and PW) with different electrostatics computation and validated the resulting core sizes against the available experimental data, for which we found a strong agreement especially for model 2. The measured asphericity coefficient from the R_g components in standard MARTINI water, PW, and PW with 50 mM ion concentration always showed lower values for model 2. From our simulation results, it can be concluded that model 2 has a higher stability and flexibility than model 1.

For C-terminal truncated variants of different lengths, full-core simulations were performed to assess the influence on the overall core stability. The results of R_g from CG modeling especially for the mutants with truncated flexible domains (absolute PCC > 0.985) showed a very good correlation with our DLS experimental data.

In summary, the use of PW and the increased total ion concentration did not lead to a substantially different initial stability of the truncated mutants in comparison with standard MARTINI water; however, a different rearrangement behavior of the fragments was observed. As these simulations are computationally expensive, the single-beaded water setups with cutoff potentials are preferably used for this truncation study.

This work will be further continued on the assembly of the whole hPDC, that is, including other E1 and E3 enzymes along with their swinging arms. This will provide the possibility to investigate the targeted manipulation of such complex enzymatic systems and thus to create the foundation for the synthesis of artificial complexes and enzymatic cascades in the future.

METHODS

CG Model. The CG models used for MD simulations of proteins are based on the MARTINI force field.[10,13] The model uses a four-to-one mapping, that is, on average, it groups three to four heavy atoms together as a single interaction center or a CG bead. Each residue has a BB bead, and depending on the type, the side chains can vary from zero to four beads. An elastic network was also applied to the MARTINI model to increase the stability of the protein.[14] However, in case of monomer and dimer simulations, the N-terminal flexible loops of E2 (20 residues) and E3BP (10 residues) were excluded from this network.

Simulation Setup. All MD simulations and analysis were performed using GROMACS software package version 5.0.2.[15,12] The visual molecular dynamics (VMD) program was used for the visualizations.[16] A cutoff of 12 Å was applied for Lennard-Jones (LJ) and Coloumbic interactions. The LJ potential was gradually shifted to zero between 0.9 and 1.2 nm, and the Coulomb potential was gradually shifted to zero between 0.0 and 1.2 nm. The temperature and pressure were maintained at the reference values (for pressure, P_0 = 1 bar) using v-rescale and Parrinello−Rahman methods with the coupling time constant τ_T = 1.0 ps for temperature and τ_P = 12 ps for pressure. A time step of 20 fs was used for the simulations. All simulations were performed at 310 K. All of the errors on the calculated properties have been evaluated using the block averaging method.[17]

E2 and E3BP Single Monomers. Single monomers of E2 and E3BP were simulated for their WT versions. Each monomer was positioned at the center of a box of side ∼9 nm filled with ∼5000 MARTINI standard water beads (W). The distance from the protein to the boundary was ∼2 nm. After 50 000 steps of energy minimization, the system was equilibrated by performing a 2 ns long simulation at NPT ensemble with position restraints. The MD runs were performed under NPT conditions for 1 μs. To neutralize the systems, one Cl^- counterion was added to the E2 WT system; however, in case of E3BP WT, the system was already neutral.

E2 and E3BP Dimers. We set up the systems same as the dimer simulation at the atomistic level,[8] that is, the two monomers were positioned in a box while facing each other from their C-terminals and with the same orientation in the *BST* 60-mer biological assembly.[18−20]

Dimers of E2 and E3BP were simulated for WT and their C-terminal truncated versions by cutting out 8 (E2-t8) and 7 (BP-t7) amino acids, respectively. Monomers were positioned about 2 nm from each other, which is measured as the distance between the C-terminal of one to the hydrophobic pocket of the other. The systems were then solvated with ∼17 000 MARTINI water in a box of ∼13 nm/side. The MD runs were performed at NPT ensemble for 100 ns. E2 WT and truncated systems were neutralized by adding two Cl^- counterions, and in case of E3BP, the systems were already neutral and no counterion was required.

Trimer. A single trimer was placed in a box of side 12 nm and filled with 14 000 MARTINI water. Simulation was performed for 2 μs. For system neutralization, three and two Cl^- counterions were added to 3E2 and 2E2/E3BP systems, respectively.

PDC Core Simulation. The full core of hPDC was built on the basis of the spherical shape of the biological assembly[18−20] of *BST* 60-meric core (pdb code, 1B5S)[21] as a reference model, that is, both models of 40E2/20E3PB and 48E2/12E3PB were set up by orienting monomers in space to fill the spherical geometry of the 60-meric core in the same way as the *BST* 60-mer. The cores were then positioned in a box of side ∼30 nm. Several steps of energy minimization and equilibration were required to prepare this huge system for the main MD run. Simulations were performed for the WT and the double-truncated version (both E2-t8 and E3BP-t7) for 6 μs in standard MARTINI water.

To further screen the influence of variable C-terminal truncation and its length on core stability, 15 different versions of single truncated systems, including truncations of either E2 or E3BP and a double-truncated version plus wild-type were prepared and simulated for at least 2.5 μs. Truncations of E2 and E3BP ranged from one to eight and one to seven, respectively. As before, variable numbers of Cl^- counterions were added depending on the net charge of the truncated core variant. However, in the given CG presentation with standard MARTINI water, the ions should not be taken too seriously. Moreover, the long-range electrostatic interactions are not present, and the first hydration shell for the small ions is considered an implicit part of the CG ion.

Simulations with PW and the PME summation method were performed for 100 ns with either Cl^- ions to neutralize the system or additionally equal amounts of Cl^- and Na^+ ions (740 numbers of each) to reach the right ionic concentration of 50 mM. Each set was done in triplicate.

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.6b00386.

E2 and E3BP dimer interactions at CG level for WT and truncated versions, RMSD for core simulations in PW for models 1 and 2; core stability results using RMSD in PW: simulation versus experiment, RMSD of core simulations in PW, increased ion concentration (50 mM); core stability results using RMSD in PW and increased ion concentration: simulation versus experiment (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: uwe.jandt@tuhh.de. Tel: +49 (0)40 42878 2847. Fax: +49 (0)40 42878 2909.

ORCID

Samira Hezaveh: 0000-0003-2910-2230

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge funding by the German Federal Ministry of Education and Research (BMBF, grant 031A128). This publication was supported by the German Research Foundation (DFG) and the Hamburg University of Technology (TUHH) in the funding program "Open Access Publishing".

■ REFERENCES

(1) Brautigam, C. A.; Wynn, R. M.; Chuang, J. L.; Chuang, D. T. Subunit and Catalytic Component Stoichiometries of an in Vitro Reconstituted Human Pyruvate Dehydrogenase Complex. *J. Biol. Chem.* **2009**, *284*, 13086−13098.

(2) Wallis, N. G.; Allen, M. D.; Broadhurst, R. W.; Lessard, I. A. D.; Perham, R. N. Recognition of a surface loop of the lipoyl domain underlies substrate channelling in the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **1996**, *263*, 463−474.

(3) Vijayakrishnan, S.; Callow, P.; Nutley, M. A.; McGow, D. P.; Gilbert, D.; Kropholler, P.; Cooper, A.; Byron, O.; Lindsay, J. G. Variation in the organization and subunit composition of the mammalian pyruvate dehydrogenase complex E2/E3BP core assembly. *Biochem. J.* **2011**, *437*, 565−574.

(4) Brautigam, C. A.; Wynn, R. M.; Chuang, J. L.; Machius, M.; Tomchick, D. R.; Chuang, D. T. Structural insight into interactions between dihydrolipoamide dehydrogenase (E3) and E3 binding protein of human pyruvate dehydrogenase complex. *Structure* **2006**, *14*, 611−621.

(5) Zhang, Y. H. Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* **2011**, *29*, 715−725.

(6) Jandt, U.; You, C.; Zhang, Y. H. P.; Zeng, A. P. Compartmentalization and Metabolic Channeling for Multienzymatic Biosynthesis: Practical Strategies and Modeling Approaches. *Adv. Biochem. Eng./Biotechnol.* **2013**, *137*, 41−65.

(7) Fu, J.; Yang, Y. R.; Johnson-Buck, A.; Liu, M. H.; Liu, Y.; Walter, N. G.; Woodbury, N. W.; Yan, H. Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* **2014**, *9*, 531−536.

(8) Hezaveh, S.; Zeng, A. P.; Jandt, U. Human Pyruvate Dehydrogenase Complex E2 and E3BP Core Subunits: New Models and Insights from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2016**, *120*, 4399−4409.

(9) Vijayakrishnan, S.; Kelly, S. M.; Gilbert, R. J. C.; Callow, P.; Bhella, D.; Forsyth, T.; Lindsay, J. G.; Byron, O. Solution Structure and Characterisation of the Human Pyruvate Dehydrogenase Complex Core Assembly. *J. Mol. Biol.* **2010**, *399*, 71−93.

(10) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(11) Guo, J.; Hezaveh, S.; Tatur, J.; Zeng, A. P.; Jandt, U. Remodeling of metabolic-channeled human pyruvate dehydrogenase complex: from disintegration and clustering to highly active agglomerates. *Biochem. J.* **2017**, *474*, 865−875.

(12) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306−317.

(13) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750−760.

(14) Periole, X.; Cavalli, M.; Marrink, S. J.; Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531−2543.

(15) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.

(16) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics Modell.* **1996**, *14*, 33−38.

(17) Flyvbjerg, H.; Petersen, H. G. Error-Estimates on Averages of Correlated Data. *J. Chem. Phys.* **1989**, *91*, 461−466.

(18) Lawson, C. L.; Dutta, S.; Westbrook, J. D.; Henrick, K.; Berman, H. M. Representation of viruses in the remediated PDB archive. *Acta Crystallogr., Sect. D: Struct. Biol.* **2008**, *D64*, 874−882.

(19) Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774−797.

(20) Henrick, K.; Thornton, J. M. PQS: a protein quaternary structure file server. *Trends Biochem. Sci.* **1998**, *23*, 358−361.

(21) Izard, T.; Ævarsson, A.; Allen, M. D.; Westphal, A. H.; Perham, R. N.; de Kok, A.; Hol, W. G. J. Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1240−1245.

DOI: 10.1021/acsomega.6b00386
ACS Omega 2017, 2, 1134−1145