Enzyme choreography in the cytoplasm: quinary structure, metabolons and domain motion at the atomic level

Premila P. Samuel Russell,[†] Meredith M. Rickard,[†] Taras V. Pogorelov[†,§,#,&,^*] and Martin Gruebele[†,§,&,^,‡,*]

[†]*Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[§]*Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[#]*School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[&]*Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[^]*National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[‡]*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

Abstract

The interactions and structural flexibility of enzymes inside cells are fundamental to biological function. The underlying atomistic mechanisms remain elusive due to technical limitations of in-cell measurements. We address this challenge by creating and simulating two all-atom models of a slice of the human cytoplasm, suitable for long-duration simulation of cellular processes while free of biasing manipulations. Each model contains three successive enzymes of the glycolysis pathway, GAPDH, PGK, and PGM. All three enzymes organize into transient complexes consistent with experimental studies, modeling the atomistic quinary structure of metabolons. Inter-enzyme interactions with microsecond lifetime occur repeatedly in our models via specific sticky protein surface patches, frequently in the vicinity of the active sites. Human PGK cycles among three meta-stable states through large-scale domain rearrangements, as postulated by previous in vitro structural studies. Our models explore enzyme organization and conformations in the cell, a first step for modeling the dynamics of entire enzymatic pathways at the atomic level.

Introduction

The heterogeneous cell environment influences enzyme function on many scales, from compartmentalization, to substrate diffusion, to enzyme conformation[1–6]. Biologically relevant macromolecular dynamics and configurations are not fully revealed or sufficiently populated in-cell within the spatiotemporal resolution of traditional experiments[7,8]. Enzymes experience transient macromolecular interactions within cells that have a crowding (repulsive)[9] component and a sticking (attractive)[10] component. These interactions can both evolve into functional quinary structures along enzymatic pathways[2,11], and lead to the emergence of large-scale domain motions in enzymes[3,6].

Canonical examples of quinary structure[2,12], the basis for multi-enzyme complexes known as metabolons[2,4,5,13–15], have been observed along the glycolytic pathway, such as between glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK)[16–18]. The formation of multi-enzyme complexes known as metabolons[2,4,5,13–15] within the cytosol not only facilitates compartmentalization of metabolic pathways for efficacy[5,19], but also increases the solvation capacity of the cytosol environment, thus regulating cell crowding and sticking[20]. The weakness of sticky interactions between enzymes in metabolons is critical for enzymatic functional diversity and regulation because it further allows enzymes to dissociate, diffuse, and associate with multiple binding partners[20].

In-cell atomistic molecular dynamics (MD) simulations[21–23] have reached the multi-μs time scale necessary to observe both transient protein clustering[24,25] as well as both the complete configurational space and rapid dynamics spanning across an enzyme's catalytic cycle[26]. Here, we build two new all-atom models of a slice of the human cytoplasm, and report over 63 μs of all-atom simulation of these models (Fig. 1). We observe repeated pairwise transient assembly and disassembly of glycolytic enzymes that further organized into tri-enzyme complexes as models of quinary structure. The transient interactions are mediated by "sticky patches" at the inter-enzyme surfaces. In past studies, GAPDH, PGK, and phosphoglycerate mutase (PGM) variants have been characterized experimentally to co-localize in different combinations in human cells[18,27]. When we further resolve these patches at the single residue level, our results sometimes were consistent with predictions that can be drawn from relevant X-ray crystallography[28]

and computational[29,30] studies conducted under in vitro conditions. The occurrence of such patches in the vicinity of enzymes' active sites during our simulations further lend support to how our systems can as serve as predictive models in understanding increased substrate channeling by glycolytic metabolons[2,4,5,15].

Our observations extend from inter-enzyme organization down to intra-enzyme organization in the cell. PGK is itself a two-domain monomeric enzyme with N- and C-terminal domains connected by a hinge that enables intra-enzyme contacts[31–33]. Multiple experimental and computational in vitro studies propose alternative hinge-bending states of PGK[6,31,34,35]. It has been surmised that opening of the hinge facilitates placement of the two substrates (ADP and 1,3-BPG) on the two halves of the PGK active site[31–33], as well as catalytic product release[33]. Closing of the hinge brings the two sites together to channel phosphate from 1,3-BPG bound at the N-terminal domain to ADP at the C-terminal domain, speeding up catalysis[31–33]. Simultaneous presence of both substrates has been proposed to stabilize the closed conformation[31,36]. *In vitro* model environments mimicking cell crowding also favor enzymatically active compact structures of PGK and adenylate kinase[3,6]. In buffers crowded by Ficoll, a 70 kDa macromolecular crowding agent, both fluorescence resonance energy transfer (FRET) experiments and coarse-grained simulations of yeast PGK concluded that PGK equilibrates towards a closed conformational ensemble[6]. Small-angle X-ray scattering (SAXS) studies done in dilute buffers concluded that human PGK has a conformational equilibrium in solution shifted towards open conformations[33]. In this study, we observed large-scale domain motions occurring via hinge-bending angle changes between domains of up to almost 60° in human PGK (Fig. 1C), which was initially placed in our cytoplasm model in semi-open conformation. Therefore, we are able to resolve how cellular crowding causes human PGK to cycle through three distinct metastable conformations along its enzymatic cycle[33] via large-scale domain reconfigurations, supporting the idea based on *in vitro* studies[6] that cytoplasmic crowding can shift enzyme structural equilibria towards compact and catalytically active conformations.

To further probe evolution of glycolytic enzymes, we replaced human PGK by a yeast variant in one of the models. The yeast PGK was seen participating in cross-species metabolon-like structures. Cross-species quinary interactions between glycolytic

enzymes is similarly observed in a past FRET study where yeast PGK and mouse GAPDH was measured to form cross-species contacts in vivo with an effective dissociation constant K_d of 14±3 μM[17]. Relevantly, glycolytic metabolons have been observed to be promoted by hypoxia stress both in yeast cells and human cancerous cells[13,19]. Both yeast and human PGK homologs in present work however show notable variations in quinary structure and domain motions within the U-2 OS models despite having similar starting configurations. These variations could be due in part to the yeast PGK surface having evolved for its activity to be regulated by its native cytoplasm. Relevantly, a recent study has shown both human PGK and GAPDH activity to be higher in yeast cells than in human muscle cells[37], while another study showed that yeast PGK stability depends on cell type[38].

Results and Discussion

All-atom models of a slice of the U-2 OS cytoplasm

To investigate the interactions and structural flexibility of three enzymes in a human cell model at the atomistic level (Fig. 1), we constructed a slice of cytoplasm composed of macromolecules (Supplementary Table 1), metabolites (Supplementary Table 2), inorganic ions, and water molecules (Supplementary Table 3) that closely mimics experimentally measured concentrations and contents of the human bone osteosarcoma cell line known as U-2 OS as well as other similar human cell lines (Fig. 1A)[39–42]. The overall cellular composition of this U-2 OS cytoplasm model is ~280 mg/mL of macromolecules, ~175 mM metabolites, and ~190 mM inorganic ions. Each eukaryotic macromolecule type selected for model construction was present as a single copy in the model. A second U-2 OS cytoplasm model was also constructed for this study with a yeast PGK variant replacing human PGK. In our cytoplasm models, with two notable exceptions (GAPDH and a tRNA), we opted to mostly include proteins that were relatively small (≤ 50 kDa) to ensure free translation of molecules. The resulting computationally tractable model sizes of approximately 200,000 atoms per model enabled us to carry out on the Anton 2 supercomputer 32.6 μs of simulation of the U-2 OS model with human PGK, and 31.2 μs on the U-2 OS model with yeast PGK, sufficient to sample a hierarchy of intra-protein and inter-protein dynamics (see Methods).

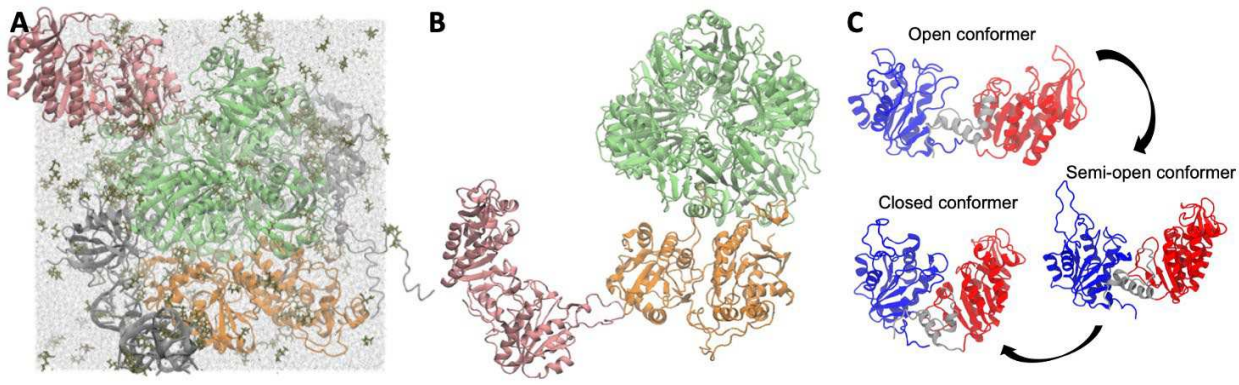


Fig. 1. Dynamics of glycolytic enzymes in the U-2 OS cell cytoplasm atomic models. (A),(B) Snapshots of the U-2 OS cell model with all-human glycolytic enzymes (PGK, GAPDH, and PGM). See Fig. S1 for additional snapshots, including with yeast PGK substituting for the human enzyme. (C) Open, semi-open, and closed conformers of human PGK along the simulation trajectory of the cytoplasm model in (A). (A) Snapshot of a complete cytoplasm model at time 0 μs (after equilibration is completed) before the model undergoes a separate MD production run of over 30 μs. All human glycolytic enzymes form a ternary transient complex in (B) at 18.3 μs into the MD trajectory. Snapshots (A) and (B) reveal long-distance diffusion and changes in contact formation. Macromolecular backbones are represented as ribbons, with color depictions in (A)-(B) of human PGK as pink, human GAPDH as lime, human PGM as orange, and all other macromolecules as silver. In (A), water molecules are represented as silver-colored CPK models, and ligand and ions are represented as tan-colored licorices. In (C), the human PGK ribbons are colored to show N-terminal domain (blue), C-terminal domain (red), and inter-domain regions (silver).

Order in chaos: from specific protein-protein sticking interactions to metabolons

Among many transient sticking events in the cytoplasm models, we observed the three glycolytic enzymes organizing together into time-evolving muti-enzyme complexes (Fig. 1B, and Supplementary Fig. 1) lasting up to several microseconds, modeling the formation of quinary structure. Figs. 1, 2, and Supplementary Figs. 1 and 2 illustrate such interactions forming, dissolving, and re-forming between GAPDH-PGM, PGM-PGK (for both yeast and human PGKs), and GAPDH-PGK (mainly for yeast PGK). As a structural highlight of both intra-species and cross-species quinary interactions in the cytoplasm models, in Fig. 1B at 18.3 μs of one simulation we observe human PGM interacting with both human PGK and GAPDH. In the second simulation at 22 μs (Supplementary Fig. 1E), human GAPDH interacts simultaneously with human PGM and yeast PGK. There is perhaps some evolutionary conservation of quinary interactions between glycolytic enzymes across eukaryotes, as further suggested by our simulation analyses of the interactions between yeast PGK and the human glycolytic enzymes (Supplementary Figs. 3B, 3C).

Although our models underwent significant protein diffusion (compare Fig. 1A and B), minimal protein-protein interaction, such as between human PGK and GAPDH in our

simulations, cannot be interpreted as a specific protein pair not interacting because the ~30 μs of sampling time are still far from ergodic regarding the mixing of protein locations. When we calculated overall molecular charges and surface electrostatic potentials of the different enzymes at pH 7 (Supplementary Table 1 and Supplementary Fig. 4), we did observe a general trend in terms of the larger the electrostatic potential differences are between an enzyme pair, the more they interacted with each other. Thus, it is plausible that protein charge can tune the interactome even in the presence of charge screening by different metabolites and inorganic ions (Supplementary Table 2).

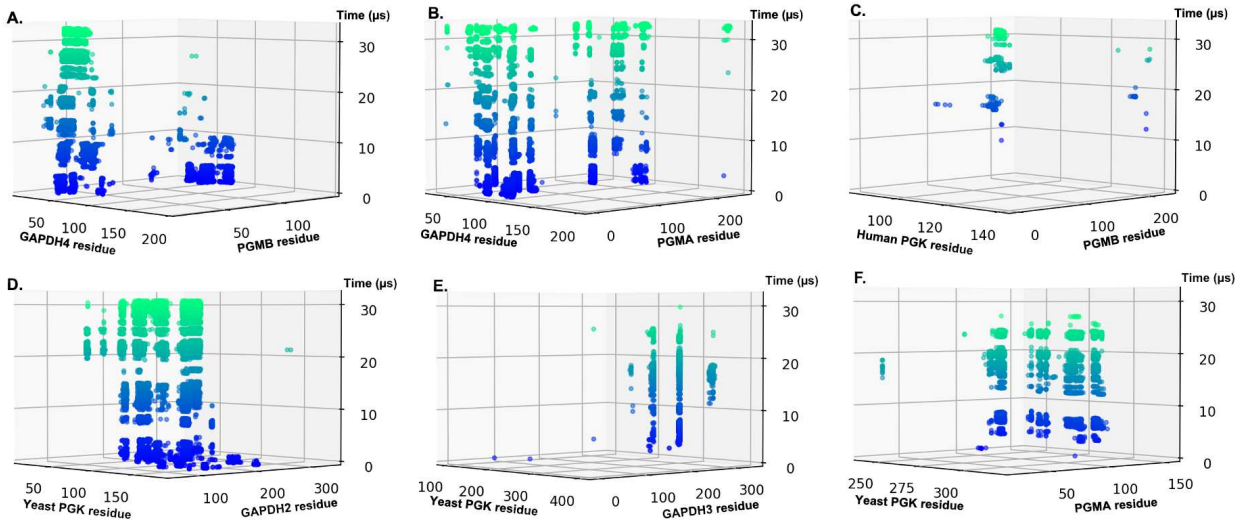


Fig. 2. Time evolution of residue-residue contacts between glycolytic enzymes in simulations of the U-2 OS models, for either (A)-(C) interacting subunit pairs of human glycolytic enzymes, or (D)-(F) interacting subunit pairs of yeast PGK and a human glycolytic enzyme**.** Each contact is shown as a circle, undergoing color change from blue to light green with increasing time. The distance cutoff for the residue-residue contacts was 0.45 nm. See SI for the rest of time evolution traces for inter-enzyme subunit pairs that interacted with average interacting area of > 0.01 nm^2 across their simulation trajectories. (A)-(C) are extracted from U-2 OS model with human PGK, and (D)-(F) are extracted from U-2 OS model with yeast PGK. See Supplementary Fig. 2 for additional time evolution.

Transient association and dissociation events between all three enzymes were analyzed at the residue level and are summarized in Figs. 2 and Supplementary Fig. 2. In characterizing sticky events between two enzyme subunits, we took into consideration all inter-enzyme subunit pairs that are interacting across a simulation trajectory with an average contact-area at least >0.01 nm^2, eliminating the smallest-area and most-transient contacts from the analysis. How then does quinary structure differ from non-specific sticking? How then does quinary structure differ from non-specific sticking? Unlike the small contact area sampled by most transient protein-protein interactions in the simulation (such as in Supplementary Fig. 2 L), the glycolytic enzyme interactions often involved

larger “sticky” patches formed by a continuous cluster of three or more amino acids that participated repeatedly in the same inter-enzyme interaction (Fig. 3, Supplementary Fig. 5, and Supplementary Fig. 6). Interactions between such patches lasted for 1-5 μs (Fig. 2 and Supplementary Fig. 2). We take these features as characteristic for the quinary structure observed here between the glycolytic enzymes.

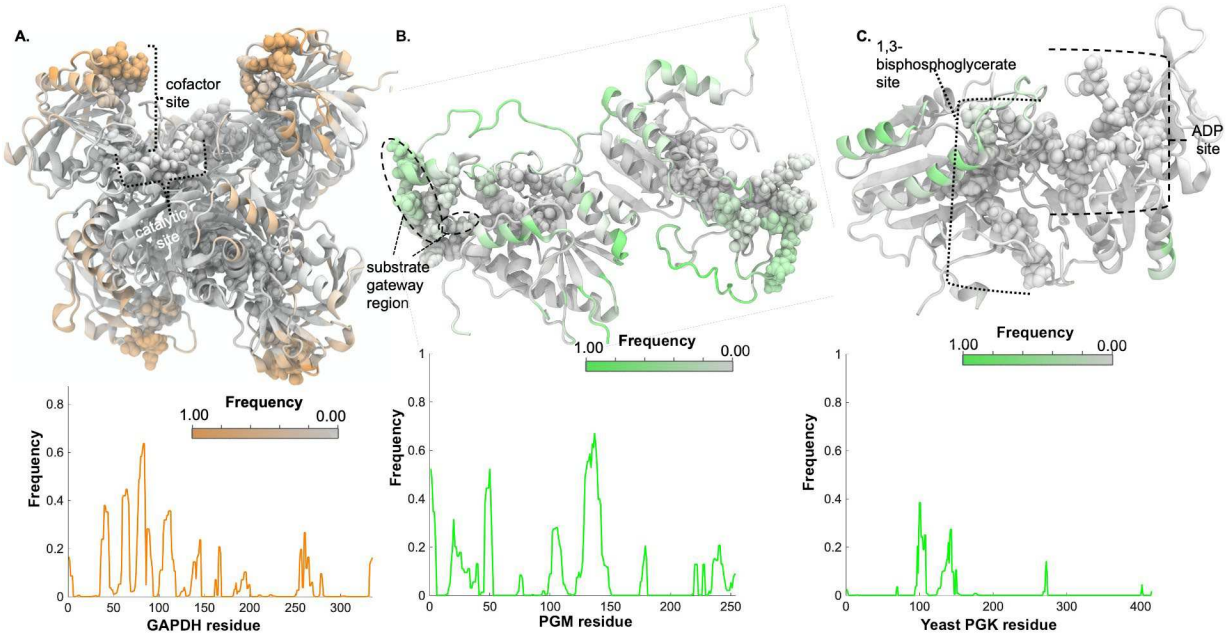


Fig. 3. Sticky patches on glycolytic enzyme surfaces are seen to arise from frequently interacting individual residues at the inter-enzyme level for specific enzyme-enzyme interactions in simulations of the U-2 OS models. (A) interaction frequency of GAPDH residues with PGM, and interaction frequency of either (B) PGM residues or (C) yeast PGK residues with GAPDH. Frequency of an enzyme residue interacting with a specific enzyme (e.g. GAPDH) at least once per simulation snapshot is visualized in here on an enzyme ribbon structure in terms of degree of color change: from grey to orange (for interaction with PGM) or green (for interaction with GAPDH). For each homo-oligomeric enzyme (GAPDH and PGM), the differences in the residue contact frequencies among the subunits were not taken account. The contact frequency across each enzyme sequence is further plotted below each structure by taking the frequency median across a sliding window of 3 residues along the sequence. In the enzyme structures, the active site regions are highlighted as van der Waals spherical surfaces. See SI for further sticky patches characterizing enzyme-enzyme interactions and the active site residue numbers.

The largest number of residue-residue interactions across the three enzyme pairs was observed in both models for GAPDH:PGM (Figs. 2, 3, and Supplementary Figs. 2, 3, 5, 6). These two enzymes are not sequential along the normal glycolytic pathway. However, the human erythrocyte bis-PGM, a PGM homolog, does participates in a special glycolytic step[43] where substrate is passed directly between GAPDH and bis-PGM. In a past experimental work, bis-PGM in complex with human GAPDH has been extracted from human erythrocyte lysate[18]. Relevantly, we observed that frequent GAPDH/PGM contacts often involve PGM residues conserved in bis-PGM (Supplementary Figure 3A).

Clustering of non-sequential enzymes in biochemical pathways also has been proposed to act as a nucleation point for sequential enzymes[5].

Quinary interaction near enzyme active sites

For PGM, many sticky patches are located near the enzyme active site(s), such as residues T99 to A110 during PGM/GAPDH interactions that were seen to occur at high frequency by occurring at least once across $\gg$ 5% of the simulation trajectories. The T99-A110 PGM patch overlaps with PGM's active site residue K100 and most importantly with the proposed substrate gateway region[28] composed of residues from K106 to W115 (Fig. 3B). Regions of the same PGM patch also appears during interactions between PGM and human PGK (Supplementary Fig. 6B), and PGM and yeast PGK (Supplementary Fig. 5B), where residues from T104 to H107 again made contacts at high frequency.

For GAPDH, an important step in its catalytic cycle is the oxidation of the bound metabolite by the cofactor NAD+. Some of the most frequent GAPDH interactions with either PGM or yeast PGK featured GAPDH sticky patches that overlap with the cofactor binding site (Fig. 3A, and Supplementary Fig. 5A) as following: P36-Y45, T75-K84, and V101-A110 in the GAPDH/PGM interactions; as well as D81-K84, and F102-E106 in the GAPDH/yeast PGK interactions. In a Brownian dynamics simulation study of metabolite channeling between human GAPDH and the glycolytic enzyme aldolase[30], the following GAPDH residues – N64, E79, D81, S83, and E106 – were seen to be key players for inter-enzyme interactions. The same GAPDH residues appear in our simulation either as part of the previously mentioned cofactor binding site overlapping with sticky patches, or at another high frequency GAPDH sticky patch – F56-V68 - appearing in our simulations between GAPDH/PGM (Fig. 3A). Similarly, in another study[29] on NADH channeling between GAPDH and lactate dehydrogenase, the key GAPDH residues lie again within the following human GAPDH stretches seen at high frequency in present study: F56-V68, T75-K84, V101-A110, and L112-G115. Comparing our results with these in vitro substrate channeling studies[29,30], we observe a common theme of one or few major pathways for GAPDH substrate channeling with the cofactor binding site playing a critical part.

Long stretches of sticky patches also occurred in the prominent flexible loop regions near the active sites of the enzymes, such as the patch of P122-D148 on the PGM loop (Fig. 3B) and the patch of I124-A139 on yeast PGK (Fig. 3C). Inter-enzyme contacts

involving these loops could increase the capture radius of the enzymes' quinary partners, as proposed by the fly-casting mechanism[44].

The PGK monomer's N-terminal domain holds the binding site for its substrate 1,3-bisphosphoglycerate (1,3-BPG) that is passed down from GAPDH, while PGK's C-terminal domain binds the ADP[45]. 1,3-BPG is converted to 3-phosphoglycerate (3-PG) following transfer of one of its phosphoryl groups to the bound ADP, and 3-PG is then passed down the glycolytic pathway to PGM. Here we see that N-terminal domains of yeast PGK (residues 1-184) and human PGK (residues 1-186) are involved in the largest number of long-lived (>1 μs) residue-residue interactions with GAPDH and PGM, respectively (Figs. 2C, 2D, Supplementary Figs. 2L, 2M, 3B, and 3C).

Hydrogen bond networks and the role of water at protein interfaces

The enzymes made contact through a combination of electrostatic, hydrophobic, and hydrogen bonding interactions as seen in Fig. 4. A concerted process from hydrophobic encounter to hydrogen-bond (H-bond) network is seen multiple times in the simulation. We highlight two specific examples of patch formation and breakage, between two sticky patches on PGM and GAPDH.

In a series of sequential snapshots of the cytoplasm model with human PGK simulation trajectory, starting at 29.256 μs in Fig. 4A-D, the PGMA subunit makes contact via residues A48, G49, Y50, and E51 with a segment on the GAPDH4 subunit overlapping with GAPDH's cofactor binding site residues of E79, R80, D81, P82, S83, and K84. Initially, the two patches interact hydrophobically via -CH_n groups within a 0.4 nm distance between the following pairs of PGMA/GAPDH4 residues: G49/E79, A48/R80, G49/K84, Y50/K84, and G49/D81. At Δt=24 ns, the formation of H-bonds between PGMA/GAPDH2 occurs as a concerted process in which H-bonds form first between PGMA A48 and GAPDH4 D81, and between PGMA G49 and K84. At Δt=48 ns, the hydrogen bonds network expands to PGMA Y50 and D81, and between PGMA E51 and GAPDH4 K84.

In Fig. 4E-H, starting at 22.8 μs, a yeast PGK sticky patch comprised of K138, A139, S140, as well as nearby D143, initially contains an intricate network of hydrogen bonds with Q113, Y140, and K145 residues on the subunit GAPDH2. The remaining three frames that follow show the disruption of that hydrogen bond network in 24 ns steps before the two enzymes come apart. The breaking of the hydrogen bonds between

GAPDH2/PGK in Fig. 4E-4F, just like the formation of hydrogen bonds in Fig. 4A-4C between PGMA/GAPDH2, is a concerted process: once the GAPDH2 Y140 interaction with the yeast PGK sub-patch is disrupted at the following 24ns time step, GAPDH2 Q113 follows quickly at 48 ns from Fig. 4E, and finally GAPDH2 K145 at 96 ns from Fig. 4E (not shown) completes the disruption.

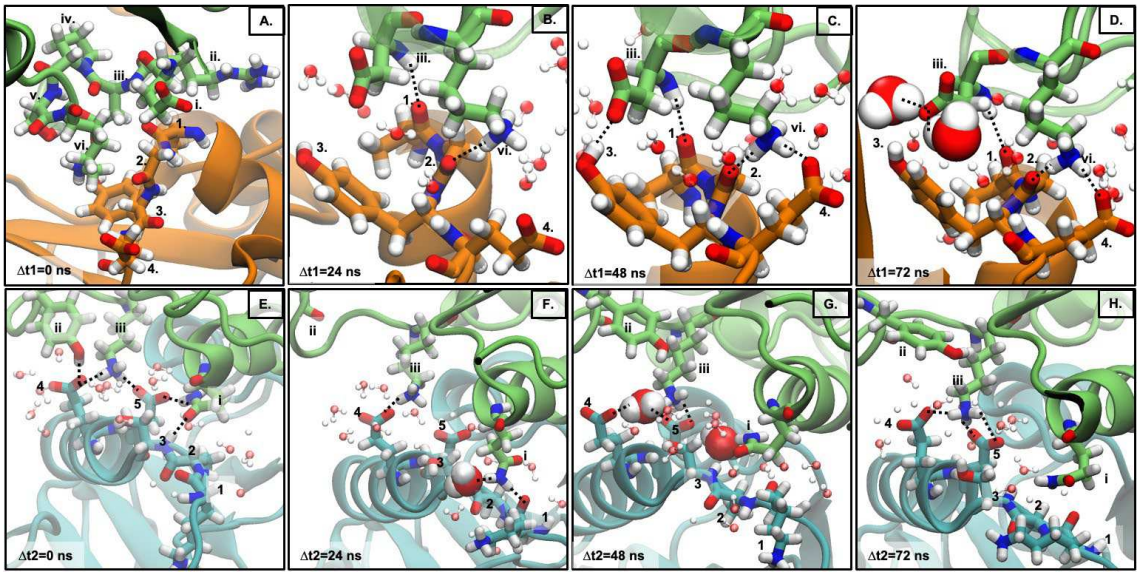


Fig. 4. Hydrogen bond networks being formed and broken between sticky patches on glycolytic enzymes, for between GAPDH4 and PGMA in (A)-(D); and between GAPDH2 and yeast PGK in (E)-(F). The frame in (A) is at 29.256 μs and frames from (B)-(D) are shown in subsequential sequences of 24 ns Δt1 time intervals from frame (A). The frame in (E) is at 22.8 μs and frames from (E)-(F) are shown in subsequential sequences of 24ns Δt2 time intervals from frame (E). In (A)-(D), residues are numbered as following, with numbers **1-4** on PGMA: **1.** Ala48, **2.** Gly49, **3.** Tyr50, and **4.** Glu51; and numbers (i)-(vi) on GAPDH2: **i.** Glu79, **ii.** Arg80, **iii.** Asp81, **iv.** Pro82, **v.** Ser83, and **vi.** Lys 84. In (E)-(F), residues are numbered as following, with numbers**1-5** on PGK: **1.** Lys138, **2.** Ala139, **3.** Ser140, **4.** Glu142, and **5.** Asp143; and numbers **i-iii** on GADPH2: **i.** Gln113, **ii.** Tyr140, and **iii.** Lys145. The black dotted lines represent hydrogen bonds, defined at a distance cutoff of 0.35 nm, and an angle cutoff of ±40° from 180°. Certain water molecules are highlighted as the van der Waals in (D), (F) and (G) when they disrupt an inter-enzyme hydrogen bond. Other structural representations are as following: colored ribbons on PGM backbone (orange), yeast PGK backbone (cyan), and GAPDH backbone (lime); CPK modeled water molecules within 0.35 nm of the numbered enzyme residues; and licorice sticks (numbered residues with red oxygen, blue nitrogen, and white hydrogen atoms).

Water penetration into the regions around the side chains facilitates H-bond network breaking, as seen in Figs. 4D, 4F, and 4G. A case in point is the interaction between yeast PGK's E142 and GAPDH2 K145 in Figs. 4E to 4F. We observe that patches form and dissolve contacts cooperatively, with a small network of water molecules moving in or out swiftly once the process of network formation or dissolution has started.

Hinge-bending of PGK in the cytoplasm model

We observe large-scale hinge-bending motions in both our human and yeast PGK models (Figs. 1C, 5). Although sampling is incomplete, both models evolve towards a

closed conformation of PGK in the model cytoplasm environments. Human PGK does so while cycling back-and-forth several times through three distinct states (Fig. 5A). We distinguished the different ensembles by calculating the changes in angle between the center-of-masses of the N-terminal domain, the hinge region, and the C-terminal domain.

We initially placed within our U-2 OS model a human PGK structure in a semi-open state[33] (PDB ID 2XE6) with a 94° bending angle between domains. This conformation first expanded into the open state ensemble (Fig. 5A), where it remained for approximately 14 μs at an average of 123°±4.8° bending angle, a value comparable to the experimental value of 120° determined for the expanded state of PGK in solution[33]. Human PGK then shifts in < 1 μs to a semi-closed bending angle conformation at an average of 106.6°±3.9° and remain in this ensemble for about 6.6 μs (Fig. 5A). The simulated semi-open ensemble bending angle is larger than that of the crystal structure's 94°. While a previous SAXS study concluded[33] that no significant time was spent by human PGK at the semi-open ensemble in buffer, we observe the contrary with cellular crowding and sticking driving compactification. In the next 7.7 μs, PGK makes three transitions between the semi-closed and closed (≈85° bending angle) states, suggestive of an equilibrium between semi-closed and closed ensembles (Fig. 5A). No transitions directly from open to closed were observed in our simulations, so the semi-closed state appears to be a metastable structural intermediate in our cytoplasm model. While the most compact closed human PGK configuration reached a bending angle of 79.3° in our simulation at time 22.464 μs (Fig. 5A), this configuration is still seen to be less compact when compared to the bending angle of 72° of the closed crystal structures (PDB ID 2X15, 2WZB) of human PGK. One factor contributing to this difference could be that the crystal conformations were obtained with bi-domain ligand binding, in contrary to our simulations, where no ligands were seen stabilizing the closed conformers.

The transition itself between states ranges from ≈400 ns to ≈1.4 μs. This is comparable to simulations of the transition time for protein folding, which sets a speed limit for folding due to diffusion and encounter of structural elements.[46] These transition periods are about 1000x faster than the catalytic turnover period for human PGK, with past studies[47] having determined catalytic turnover number, or k_{cat}, for human PGK to be 795±50 s^{-1}.

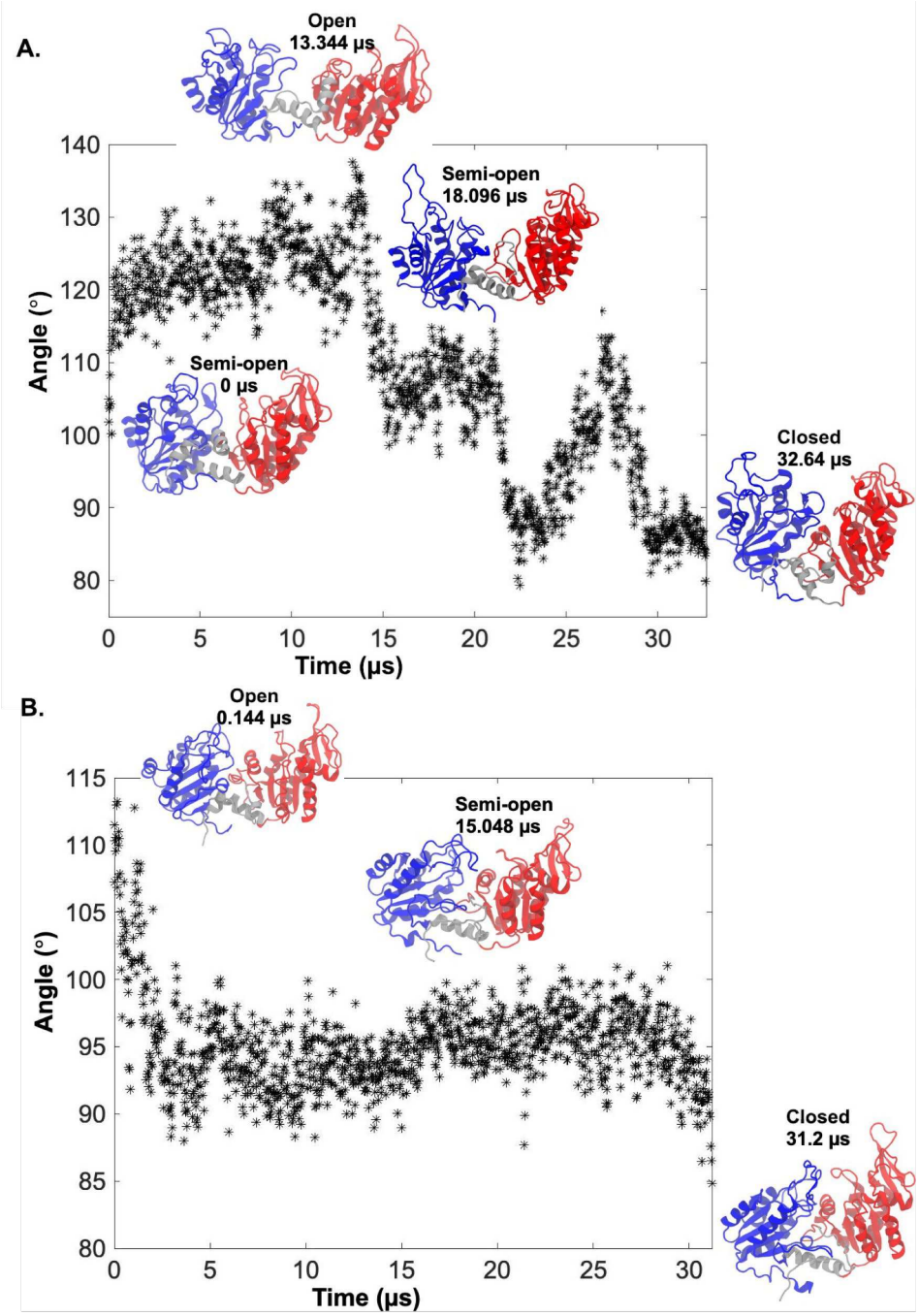


Fig. 5. Hinge-bending motion in structures of (A) human and (B) yeast PGKs shifts the PGK conformations from "open" to "closed" during the simulations. The inter-domain angle was calculated as the angle between the center-of-mass of the N-terminal domain (residues 1-186 for human PGK or residues 1-184 for yeast PGK), linker region connecting the N- and C- terminal domains (residues 187-201 for human PGK or residues 185-199 for yeast PGK), and C-terminal domain (residues 202-395 for human PGK or residues 200-393 for yeast PGK). Ribbon figures show snapshots of backbones of PGKs at different simulation time points across varying states of compactness as follows. Human PGK interdomain angle at 0 μs is 104.9°, at 13.344 μs is 137.6°, at 18.096 μs is 107.9°, and at 32.64 μs is 79.9°. Yeast PGK interdomain angle at 0.144 μs is 113.3°, at 15.048 μs is 95.8°, and at 31.2 μs is 84.8°. PGK ribbons show N-terminal domain (blue), C-terminal domain (red), and inter-domain regions (silver).

Previous coarse-grained in vitro simulations of crowded yeast PGK reported a maximum change of approximately 60° in bending angle[6]. For yeast PGK (Fig. 5B), we began with an open structure crystal structure (PDB ID 1QPG) that had a bending angle of 110.7°. In contrast to human PGK, the yeast variant rapidly reached a more semi-open configuration within 0.8 μs of the simulation starting in Fig. 5B and remained in the semi-open conformation for about 0.91 fraction of the simulation trajectory with a bending angle of about 95° +/- 2.8°. At the very end of the simulation, the hinge angle drops below 85°, hinting at more compact states involved in the in-cell equilibrium. There is no high-resolution published structure yet of a closed conformation of yeast PGK.

Conclusion

Our models enable predictive exploration of enzyme structural dynamics in human cells. We see in two atomistic cytoplasm models that successive enzymes in a biochemical cycle form transient contact, lasting from 1 to 5 μs as they break and re-form. These contacts occur as surface patches that form quinary structure, and contacts between these patches form as early hydrophobic interactions give way to hydrogen bonding while several water molecules edge in cooperatively between side chains. The formation of up to three-way metabolons that stick together near enzyme active sites in our cytoplasm models illustrates how long-time simulations can serve as an unbiased platform for structurally deconstructing enzyme interactomes at the atomistic level in-cell.

In addition to intermolecular contacts, we were also able to capture spontaneous hinge-bending motions of two PGKs and transitions between distinct ensembles that have been proposed in the literature[31–33] to be critical to the enzymes' catalytic cycle. Our cytoplasm models could form the basis for accessing enzymatically active states, which are energetically unfavorable and thus are short-lived under traditional experimental setups[33]. Starting from initial experimental structures obtained from the PDB, reactive quantum mechanics/molecular mechanics simulations can be pursued in future to study enzyme catalysis with substrate binding at the active sites[48].

Methods

Cytoplasm model construction and simulation

We describe below our workflow for constructing a human cytoplasm model, following the general steps pursued for our previous bacterial cytoplasm model[22]. With exception to the Brownian Dynamics step described below, model preparation and simulation steps are done fully atomistically. The initial structural configurations of macromolecules placed in the cytoplasm models were obtained from Protein Data Bank (PDB) as described in Supplementary Table 1 and in the case of certain incomplete sequences or for introducing mutations, further prepped through the MODELLER software[49]. Both the homo-tetrameric GAPDH (UniProt ID P04406 and M1-E335 residues/subunit), with subunits GAPDH1 - GAPDH4, and homo-dimeric PGM (UnitProt ID P18669 and residues/subunit M1-K254), with subunits PGMA - PGMB, in each of our models are of human origin. The monomeric PGK in each model meanwhile is either of the human PGK1 (UnitProt ID P00558 and residues M0-I416) or of yeast PGK3 (UniProt ID P00560 and residues S1-K415). The yeast PGK variant used in our study had 72 % sequence similarity with human PGK1.In addition to eukaryotic macromolecules, we also included in the cytoplasm models three unfolded copies of the small (5.2 kDa), genetically engineered protein B for separate folding studies. The yeast PGK variant used in our simulation had R65Q/W333F mutations, with R65Q mutation originating from the PDB ID 1QPG crystal structure we used, and the W333F mutation was to allow for direct comparison with previous studies of PGK dynamics[6].

Macromolecules were initially packed together into a 14 nm cube using the PACKMOL software[50]. Next, the packed macromolecules were set up as rigid bodies, with electrostatics and Lenard-Jones potentials, and followingly equilibrated through Brownian dynamics using the ARBD software[51] for 4ns in order to allow for spatial sampling of macromolecules and to further optimize packing. Following the Brownian dynamics simulation, metabolites were next added to the cube using PACKMOL software. We used two different experimental studies to determine a consensus metabolite abundance table for human cells[40,41]. 44 different species of metabolites were sampled by Monte Carlo selection from the abundance table. A comparison of the metabolite concentrations represented in the U-2 OS models and the consensus experimental concentrations is

included in Supplementary Information Table 2. VMD software[52] was then used both for addition of inorganic ions, and finally water molecules into the system. The model was parameterized with TIP3P[53] water for solvent, Beglov and Roux's parameters[54] for inorganic ions, and following CHARMM force fields: CHARMM36m[55] for proteins, CHARMM36 additive for nucleic acids, and CGENFF[56] for metabolites. We previously showed that the CHARMM36m force-field produces less macromolecular sticking than CHARMM22*[22].

Prior to simulations done on Anton 2, the molecular dynamics software NAMD[57] (both versions 2.9 and 3.0 Alpha) was used for initial minimization, and pre-equilibration/equilibration steps of the cytoplasm models. The models initially were minimized first for 6000 steps after solvation, and then underwent sequentially the following series of pre-equilibration/equilibration simulations at constant-pressure, constant temperature (NPT) ensemble with 2 fs time-steps: 2.2 ns at 318 K, simulated annealing from 318 K → 498 K (for 4 ns) → 318 K in steps of 5K/100 ps, and 6 ns at 318 K. The model cube length reached 12.57 nm at end of these steps, within the final target length range for our cytoplasm model to satisfy the cytoplasm model's content concentrations that we projected to fulfil the experimental based literature numbers[39–42]. We prepared first, through the steps aforementioned, the U-2 OS model with yeast PGK. Then, another copy of the cytoplasm model was replicated by swapping out the yeast PGK with the more-compact semi-open structure of human PGK. With exception to the PGK differences, content of the U-2 OS model with yeast PGK mimics that of U-2 OS model with human PGK with only few small differences as described in Supplementary Tables 2, and 3.

The models underwent extensive production runs on the Anton 2 supercomputer (Pittsburgh Supercomputer Center) using an NPT ensemble with a multigrator[58] scheme at 1 atm pressure and 318 K. This temperature was selected to maximize protein diffusion while staying within realistic bounds for the cellular environment.

Data analysis

Snapshots of both the 31.2 μs (U-2 model with yeast PGK) and the 32.6 μs (U-2 model with human PGK) trajectories obtained from the production runs on Anton 2 were saved

every 24 ns. The initial 216 ns of each trajectory was considered as the equilibration period and not used for data analysis of inter-enzyme interactions but is included in Fig. 5. Simulation trajectories were wrapped with VMD's PBCTools plugin prior to data analysis to address periodic boundary conditions when calculating protein-protein interactions. Simulation trajectory analysis of inter-enzyme interactions at both the residue and atomic levels, calculation of sticky patches, and visualization of simulation snapshots were done with the VMD[52] software. Prior to inter-enzyme interaction calculations, solvent accessible surface area change (SASA) with possible complexation were calculated for each pair of inter-enzyme subunits across the simulation trajectory using VMD[52] software as previously done in our bacterial cytoplasm study[22] to calculate complexation site area. The calculations were done to determine subunit pairs that interacted with an average contact-area > 0.01 nm^2 as a threshold for the enzyme-enzyme sticky interactions. Calculations of angle changes between PGK domains were done with the CPPTRAJ[59] software package. Adaptive Poisson-Boltzmann Solver software suite[60] was used to calculate the electrostatic surface potentials of the enzymes. Enzyme sequence conservations were calculated using the Sequence Manipulation suite resources[61] available at https://www.bioinformatics.org.

We measured both residue pair contact frequency and a residue frequency in forming inter-enzyme contacts (in the case of sticky patches) by calculating how often an enzyme residue-residue pair and a specific enzyme residue respectively forms a contact at least once per simulation snapshot saved every 24 ns across the simulation trajectories of the production runs. In the case of multiple occurrences of the same type of contact (either a residue pair contact or a contact involving a specific residue) per frame, only one occurrence of the contact is counted per frame and so the difference in contact frequency across the subunits of a specific enzyme is not considered in our calculations.

Enzyme residue was considered as conserved here in terms of sequence similarity if the corresponding residue occurring intra- or inter-species either is same or similar. We grouped similar amino acids as following: GAVLI, FYW, CM, ST, KRH, DE, and NQ. For GAPDH alignment, if the human GAPDH residue was conserved in at least 2 yeast GAPDH variants, we marked the residue here as conserved.

References

1. Agarwal, P. K., Bernard, D. N., Bafna, K. & Doucet, N. Enzyme Dynamics: Looking Beyond a Single Structure. *ChemCatChem* **12**, 4704–4720 (2020).
2. Vélot, C., Mixon, M. B., Teige, M. & Srere, P. A. Model of a Quinary Structure between Krebs TCA Cycle Enzymes: A Model for the Metabolon †. *Biochemistry* **36**, 14271–14276 (1997).
3. Echeverria, C. & Kapral, R. Molecular crowding and protein enzymatic dynamics. *Phys. Chem. Chem. Phys.* **14**, 6755–6763 (2012).
4. Srere, P. A. The metabolon. *Trends in Biochemical Sciences* **10**, 109–110 (1985).
5. Sweetlove, L. J. & Fernie, A. R. The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nature Communications* **9**, (2018).
6. Dhar, A. *et al.* Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 17586–17591 (2010).
7. Samuel, P. P. & Case, D. A. Atomistic Simulations of Heme Dissociation Pathways in Human Methemoglobins Reveal Hidden Intermediates. *Biochemistry* **59**, 4093–4107 (2020).
8. Raich, L. *et al.* Discovery of a hidden transient state in all bromodomain families. *Proceedings of the National Academy of Sciences* **118**, e2017427118 (2021).
9. Minton, A. P. & Wilf, J. Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **20**, 4821–4826 (1981).
10. Guin, D. & Gruebele, M. Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chem. Rev.* **119**, 10691–10717 (2019).
11. Srere, P. A. Macromolecular interactions: tracing the roots. *Trends in Biochemical Sciences* **25**, 150–153 (2000).
12. McConkey, E. H. Molecular evolution, intracellular organization, and the quinary structure of proteins. *PNAS* **79**, 3236–3240 (1982).
13. Jin, M. *et al.* Glycolytic Enzymes Coalesce in G Bodies under Hypoxic Stress. *Cell Reports* **20**, 895–908 (2017).
14. Kohnhorst, C. L. *et al.* Identification of a multienzyme complex for glucose metabolism in living cells. *Journal of Biological Chemistry* **292**, 9191–9203 (2017).
15. Huang, Y. M., Huber, G. A., Wang, N., Minteer, S. D. & McCammon, J. A. Brownian dynamic study of an enzyme metabolon in the TCA cycle: Substrate kinetics and channeling. *Protein Science* **27**, 463–471 (2018).
16. Tomokuni, Y. *et al.* Loose interaction between glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase revealed by fluorescence resonance energy transfer-fluorescence lifetime imaging microscopy in living cells. *FEBS J* **277**, 1310–1318 (2010).
17. Sukenik, S., Ren, P. & Gruebele, M. Weak protein–protein interactions in live cells are quantified by cell-volume modulation. *Proceedings of the National Academy of Sciences* **114**, 201700818–201700818 (2017).
18. Fokina, K. V., Dainyak, M. B., Nagradova, N. K. & Muronetz, V. I. A Study on the Complexes between Human Erythrocyte Enzymes Participating in the Conversions of 1,3-Diphosphoglycerate. *Archives of Biochemistry and Biophysics* **345**, 185–192 (1997).
19. Fuller, G. G. & Kim, J. K. Compartmentalization and metabolic regulation of glycolysis. *J Cell Sci* **134**, jcs258469 (2021).

20. Menard, L., Maughan, D. & Vigoreaux, J. The structural and functional coordination of glycolytic enzymes in muscle: evidence of a metabolon? *Biology (Basel)* **3**, 623–644 (2014).
21. Yu, I. *et al.* Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife* **5**, e19274 (2016).
22. Rickard, M. M., Zhang, Y., Gruebele, M. & Pogorelov, T. V. In-Cell Protein-Protein Contacts: Transient Interactions in the Crowd. *The Journal of Physical Chemistry Letters* **10**, 5667–5673 (2019).
23. Rickard, M. M., Zhang, Y., Pogorelov, T. V. & Gruebele, M. Crowding, Sticking, and Partial Folding of GTT WW domain in a Small Cytoplam Model. *Journal of Physical Chemistry B* **124**, 4732–4740 (2020).
24. Nawrocki, G., Wang, P., Yu, I., Sugita, Y. & Feig, M. Slow-Down in Diffusion in Crowded Protein Solutions Correlates with Transient Cluster Formation. *J. Phys. Chem. B* **121**, 11072–11084 (2017).
25. Bülow, S. von, Siggel, M., Linke, M. & Hummer, G. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *PNAS* **116**, 9843–9852 (2019).
26. Eisenmesser, E. Z., Bosco, D. A., Akke, M. & Kern, D. Enzyme Dynamics During Catalysis. *Science* **295**, 1520–1523 (2002).
27. Kowalski, W., Nocon, D., Gamian, A., Kołodziej, J. & Rakus, D. Association of C-terminal region of phosphoglycerate mutase with glycolytic complex regulates energy production in cancer cells. *Journal of Cellular Physiology* **227**, 2613–2621 (2012).
28. Wang, Y. *et al.* Crystal structure of human B-type phosphoglycerate mutase bound with citrate. *Biochem Biophys Res Commun* **331**, 1207–1215 (2005).
29. Svedružić, Ž. M., Odorčić, I., Chang, C. H. & Svedružić, D. Substrate Channeling via a Transient Protein-Protein Complex: The case of D-Glyceraldehyde-3-Phosphate Dehydrogenase and L-Lactate Dehydrogenase. *Sci Rep* **10**, 10404 (2020).
30. Ouporov, I. V., Knull, H. R., Huber, A. & Thomasson, K. A. Brownian Dynamics Simulations of Aldolase Binding Glyceraldehyde 3-Phosphate Dehydrogenase and the Possibility of Substrate Channeling. *Biophysical Journal* **80**, 2527–2535 (2001).
31. Varga, A. *et al.* Substrate-induced double sided H-bond network as a means of domain closure in 3-phosphoglycerate kinase. *FEBS Lett* **580**, 2698–2706 (2006).
32. Bernstein, B. E., Michels, P. A. & Hol, W. G. Synergistic effects of substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature* **385**, 275–278 (1997).
33. Zerrad, L. *et al.* A spring-loaded release mechanism regulates domain movement and catalysis in phosphoglycerate kinase. *J Biol Chem* **286**, 14040–14048 (2011).
34. McPhillips, T. M., Hsu, B. T., Sherman, M. A., Mas, M. T. & Rees, D. C. Structure of the R65Q Mutant of Yeast 3-Phosphoglycerate Kinase Complexed with Mg-AMP-PNP and 3-Phospho-d-glycerate,. *Biochemistry* **35**, 4118–4127 (1996).
35. Auerbach, G. *et al.* Closed structure of phosphoglycerate kinase from Thermotoga maritima reveals the catalytic mechanism and determinants of thermal stability. *Structure* **5**, 1475–1483 (1997).
36. Szabó, J. *et al.* Communication between the Nucleotide Site and the Main Molecular Hinge of 3-Phosphoglycerate Kinase. *Biochemistry* **47**, 6735–6744 (2008).
37. Boonekamp, F. J. *et al.* Full humanization of the glycolytic pathway in Saccharomyces cerevisiae. *Cell Reports* **39**, 111010 (2022).
38. Feng, R., Gruebele, M. & Davis, C. M. Quantifying protein dynamics and stability in a living organism. *Nat Commun* **10**, 1179 (2019).

39. Beck, M. *et al.* The quantitative proteome of a human cell line. *Molecular Systems Biology* **7**, 1–8 (2011).
40. Diener, C., Muñoz-Gonzalez, F., Encarnación, S. & Resendis-Antonio, O. The space of enzyme regulation in HeLa cells can be inferred from its intracellular metabolome. *Sci Rep* **6**, 28415 (2016).
41. Chen, W. W., Freinkman, E., Wang, T., Birsoy, K. & Sabatini, D. M. Absolute Quantification of Matrix Metabolites Reveals the Dynamics of Mitochondrial Metabolism. *Cell* **166**, 1324-1337.e11 (2016).
42. Theillet, F.-X. *et al.* Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **114**, 6661–6714 (2014).
43. van Wijk, R. & van Solinge, W. W. The energy-less red blood cell is lost: erythrocyte enzyme abnormalities of glycolysis. *Blood* **106**, 4034–4042 (2005).
44. Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proceedings of the National Academy of Sciences* **97**, 8868–8873 (2000).
45. Rojas-Pirela, M. *et al.* Phosphoglycerate kinase: structural aspects and functions, with special emphasis on the enzyme from Kinetoplastea. *Open Biology* **10**, 200302.
46. Liu, Y., Prigozhin, M. B., Schulten, K. & Gruebele, M. Observation of Complete Pressure-Jump Protein Refolding in Molecular Dynamics Simulation and Experiment. *J. Am. Chem. Soc.* **136**, 4265–4272 (2014).
47. Yagi, H., Kasai, T., Rioual, E., Ikeya, T. & Kigawa, T. Molecular mechanism of glycolytic flux control intrinsic to human phosphoglycerate kinase. *Proceedings of the National Academy of Sciences* **118**, e2112986118 (2021).
48. Lin, H. & Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor Chem Acc* **117**, 185–199 (2007).
49. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* **54**, 5.6.1-5.6.37 (2016).
50. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J Comput Chem* **30**, 2157–2164 (2009).
51. Comer, J. & Aksimentiev, A. Predicting the DNA Sequence Dependence of Nanopore Ion Current Using Atomic-Resolution Brownian Dynamics. *J. Phys. Chem. C* **116**, 3376–3393 (2012).
52. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–8, 27–8 (1996).
53. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. COMPARISON OF SIMPLE POTENTIAL FUNCTIONS FOR SIMULATING LIQUID WATER. *J. Chem. Phys.* **79**, 926–935 (1983).
54. Beglov, D. & Roux, B. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *J. Chem. Phys.* **100**, 9050–9063 (1994).
55. Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **14**, 71–73 (2017).
56. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **31**, 671–690 (2010).

57. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153**, 044130 (2020).
58. Lippert, R. A. *et al.* Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *The Journal of Chemical Physics* **139**, 164106 (2013).
59. Roe, D. R. & Cheatham, T. E., 3rd. Parallelization of CPPTRAJ enables large scale analysis of molecular dynamics trajectory data. *J Comput Chem* **39**, 2110–2117 (2018).
60. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Science* **27**, 112–128 (2018).
61. Stothard, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* **28**, 1102, 1104 (2000).

Data availability

Data supporting the findings of this study are included within this article and in the Supplementary Information.

Acknowledgements

P.S.R., M.R., and M.G. were supported by the NSF grant MCB 2205665. Anton 2 computer time was provided by the Pittsburgh Supercomputing Center (PSC) through Grant R01GM116961 from the National Institutes of Health. The Anton 2 machine at PSC was generously made available by D.E. Shaw Research. T.V.P. acknowledges support from the NIH grant R01-GM141298 and the Department of Chemistry, University of Illinois at Urbana-Champaign.

Author Contributions

P.S.R. designed research, carried out simulations, developed data analysis, analyzed data, wrote and edited the manuscript; M.M.R. carried out research, analyzed data, and edited the paper; T.V.P. designed research, developed data analysis, and edited the paper; M.G. conceived and designed research, developed data analysis, wrote and edited the manuscript.

Competing Interest

The authors declare no competing financial interests.

Additional Information

Supplementary Information is available for this paper.

Corresponding Authors

*To whom correspondence may be addressed. Email: mgruebel@illinois.edu and pogorelo@illinois.edu

Supplementary Information

Enzyme choreography in the cytoplasm: quinary structure, metabolons and domain motion at the atomic level

Premila P. Samuel Russell,[†] Meredith M. Rickard,[†] Taras V. Pogorelov[†,§,#,&,^*] and Martin Gruebele[†,§,&,^, ‡,*]

Supplementary Table 1. Macromolecules in the U-2 OS cytoplasm models

Macromolecule | Charge at pH 7 per subunit | Original structure PDB ID	Protein /nucleotide sequence per subunit used in model
Cyclophilin G (peptidyl-prolyl isomerase domain) | 0 | 2GW2	RPRCFFDIAINNQPAGRVVFELFSDVCPKTCENFRCLCTGEKGTGKSTQKPLHYKSCLFH RVVKDFMVQGGDFSEGNGRGGESIYGGFFEDESFAVKHNAAFLLSMANRGKDTNGSQFFI TTKPTPHLDGHHVVFGQVISGQEVVREIENQKTDAASKPFAEVRILSCGELIP
Calmodulin | -24 | 1CLL	MADQTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGN GTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEE VDEMIREADIDGDGQVNYEEFVQMMTAK
tRNA | -71 | 2AZX	GACCUCGUGGCGCAAUGGUAGCGCGUCUGACUCCAGAUCAGAAGGUUGCGUGUUCGAAUC ACGUCGGGGUCA
GAPDH | 3 | 1U8F	MGKVKVGVNGFGRIGRLVTRAAFNSGKVDIVAINDPFIDLNYMVYMFQYDSTHGKFHGTV KAENGKLVINGNPITIFQERDPSKIKWGDAGAEYVVESTGVFTTMEKAGAHLQGGAKRVI ISAPSADAPMFVMGVNHEKYDNSLKIISNASCTTNCLAPLAKVIHDNFGIVEGLMTTVHA ITATQKTVDGPSGKLWRDGRGALQNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTANV SVVDLTCRLEKPAKYDDIKKVVKQASEGPLKGILGYTEHQVVSSDFNSDTHSSTFDAGAG IALNDHFVKLISWYDNEFGYSNRVVDLMAHMASKE
HAH1 | 0 | 1TLF	MPKHEFSVDMTCGGCAEAVSRVLNKLGGVKYDIDLPNKKVCIESEHSMDTLLATLKKTGK TVSYLGLE
PGM | -1 | 1YFK	MAAYKLVLIRHGESAWNLENRFSGWYDADLSPAGHEEAKRGGQALRDAGYEFDICFTSVQ KRAIRTLWTVLDAIDQMWLPVVRTWRLNERHYGGLTGLNKAETAAKHGEAQVKIWRRSYD VPPPPMEPDHPFYSNISKDRRYADLTEDQLPSCESLKDTIARALPFWNEEIVPQIKEGKR VLIAAHGNSLRGIVKHLEGLSEEAIMELNLPTGIPIVYELDKNLKPIKPMQFLGDEETVR KAMEAVAAQGKAKK
Protein B | -1 | 1PRB	LKNAIEDAIAEWKKAGITSDFYFNAINKAKTVEEVNALVNEILKAHA
Human PGK | 3 | 2XE6	MSLSNKLTLDKLDVKGKRVVMRVDFNVPMKNNQITNNQRIKAAVPSIKFCLDNGAKSVVL MSHLGRPDGVPMPDKYSLEPVAVELKSLLGKDVLFLKDCVGPEVEKACANPAAGSVILLE NLRFHVEEEGKGKDASGNKVKAEPAKIEAFRASLSKLGDVYVNDAFGTAHRAHSSMVGVN LPQKAGGFLMKKELNYFAKALESPERPFLAILGGAKVADKIQLINNMLDKVNEMIIGGGM AFTFLKVLNNMEIGTSLFDEEGAKIVKDLMSKAEKNGVKITLPVDFVTADKFDENAKTGQ ATVASGIPAGWMGLDCGPESSKKYAEAVTRAKQIVWNGPVGVFEWEAFARGTKALMDEVV KATSRGCITIIGGGDTATCCAKWNTEDKVSHVSTGGGASLELLEGKVLPGVDALSNI

Yeast PGK | -1 | 1QPG	SLSSKLSVQDLDLKDKRVFIRVDFNVPLDGKKITSNQRIVAALPTIKYVLEHHPRYVVLA SHLGQPNGERNEKYSLAPVAKELQSLLGKDVTFLNDCVGPEVEAAVKASAPGSVILLENL RYHIEEEGSRKVDGQKVKASKEDVQKFRHELSSLADVYINDAFGTAHRAHSSMVGFDLPQ RAAGFLLEKELKYFGKALENPTRPFLAILGGAKVADKIQLIDNLLDKVDSIIIGGGMAFT FKKVLENTEIGDSIFDKAGAEIVPKLMEKAKAKGVEVVLPVDFIIADAFSADANTKTVTD KEGIPAGWQGLDNGPESRKLFAATVAKAKTIVFNGPPGVFEFEKFAAGTKALLDEVVKSS AAGNTVIIGGGDTATVAKKYGVTDKISHVSTGGGASLELLEGKELPGVAFLSEKK

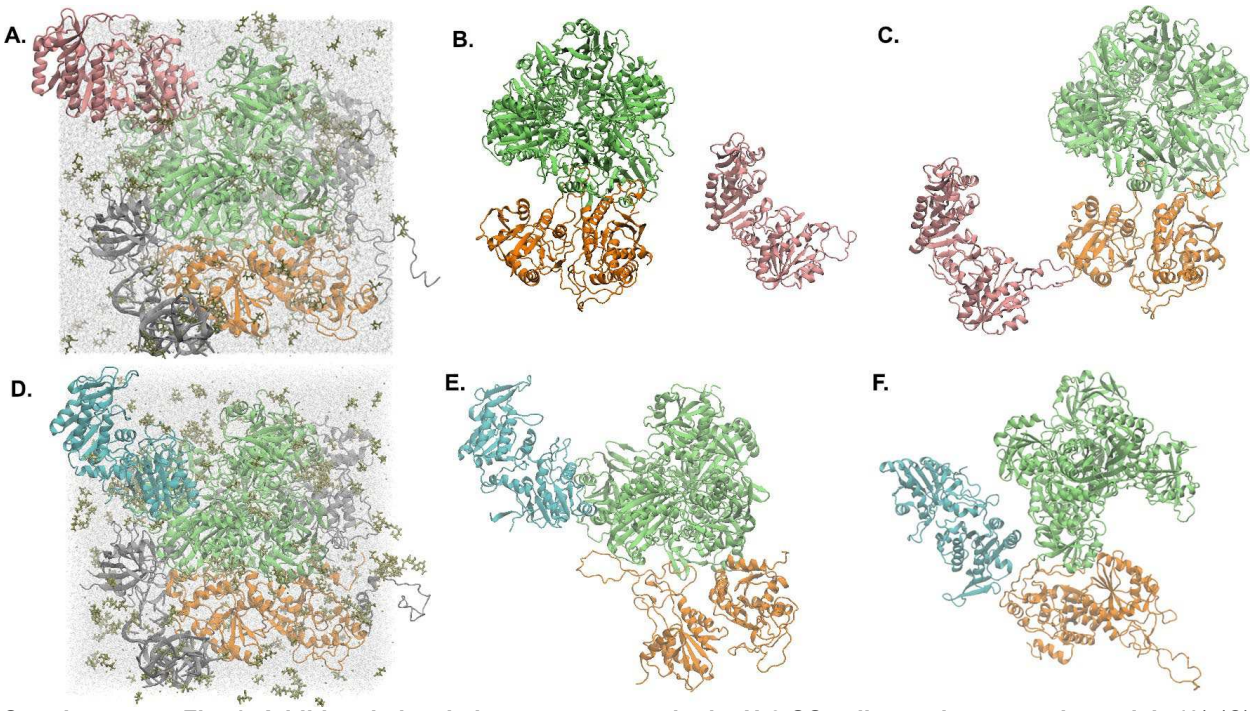
Supplementary Table 2. Metabolite in the U-2 OS cytoplasm models

Metabolite	Experimental measurements consensus concentration* (mM) [31,32]	No. of Copies in model	Concentration in Model (mM)	Charge at pH 7 per molecule
Glutamate	31.883	41	31.000	-1.00
Taurine	16.901	25	18.902	0.00
Aspartate	15.733	21	15.878	-1.00
Glutamine	11.006	13	9.829	0.00
Glycine	10.112	13	9.829	0.00
Glutathione (GSH)	8.208	6	4.537	0.00
Spermine	7.240	5	3.780	+4.00
Lactate	5.461	5	3.780	-1.00
Asparagine	3.717	2	1.512	-1.00
Malate	3.651	10	7.561	0.00
Alanine	3.517	6	4.537	0.00
Spermidine	3.425	2	1.512	+3.00
Beta-Alanine	3.405	4	3.024	0.00
Citrate/Isocitrate	3.117	7	5.293	-3.00
Hydroxyproline	2.707	4(human PGK model)/ 5(yeast PGK model)	3.024/3.780	0.00
ATP	2.476	2	1.512	-4.00
Putrescine	2.084	1	0.756	+2.00
Phosphocholine	1.908	3	2.268	0.00
UTP	1.840	4	3.024	-4.00
Isoleucine	1.709	1	0.756	0.00
NAD	1.601	2	1.512	-1.00
Serine	1.588	2	1.512	0.00
Leucine	1.572	1	0.756	0.00
Threonine	1.278	2	1.512	0.00
Arginine	1.234	4	3.024	+1.00
Histidine	1.098	1	0.756	+1.00
Valine	1.096	1	0.756	0.00
Glutathione disulfide (GSSG)	1.056	1	0.756	-2.00
UDP	0.854	1	0.756	0.00

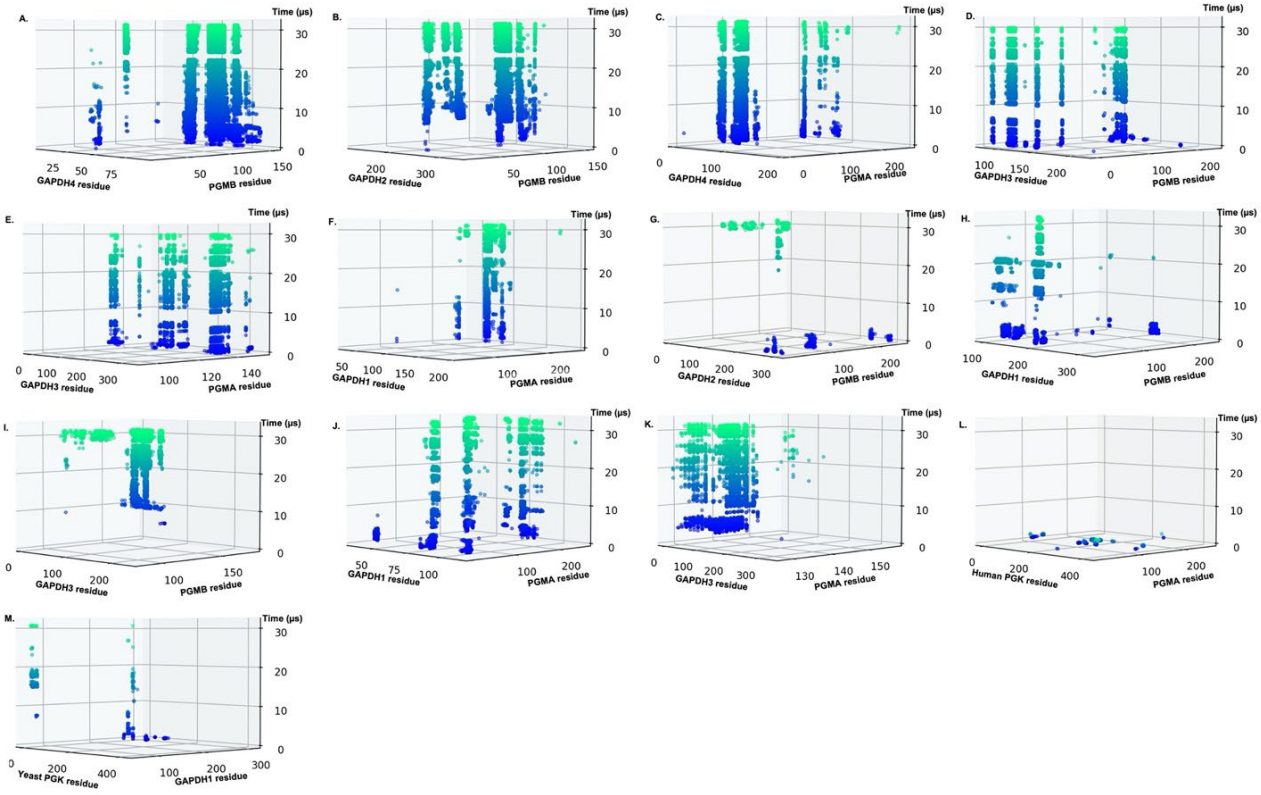
Tyrosine	0.838	1	0.756	0.00
3-Hydroxybutyric acid	0.680	2	1.512	-1.00
Glycerol-3-phosphate (G3P)	0.662	1	0.756	-2.00
Fumarate	0.622	1	0.756	-2.00
GTP	0.600	1	0.756	-4.00
ADP	0.487	2	1.512	-3.00
AMP	0.460	1	0.756	-2.00
Fructose-1,6-bisphosphate	0.436	1	0.756	-4.00
Succinate	0.408	1	0.756	0.00
Pyruvate	0.321	1	0.756	-1.00
Ribose-5-phosphate/ ribulose-5-phosphate	0.244	1	0.756	-2.00
Pantothenic Acid	0.235	1	0.756	-1.00
Glucose-6-phosphate (G6P)	0.234	1	0.756	-2.00
Citrulline	0.217	1	0.756	0.00
GMP	0.095	1	0.756	-2.00

Supplementary Table 3: Inorganic ions and solvent content in the U-2 OS cytoplasm model

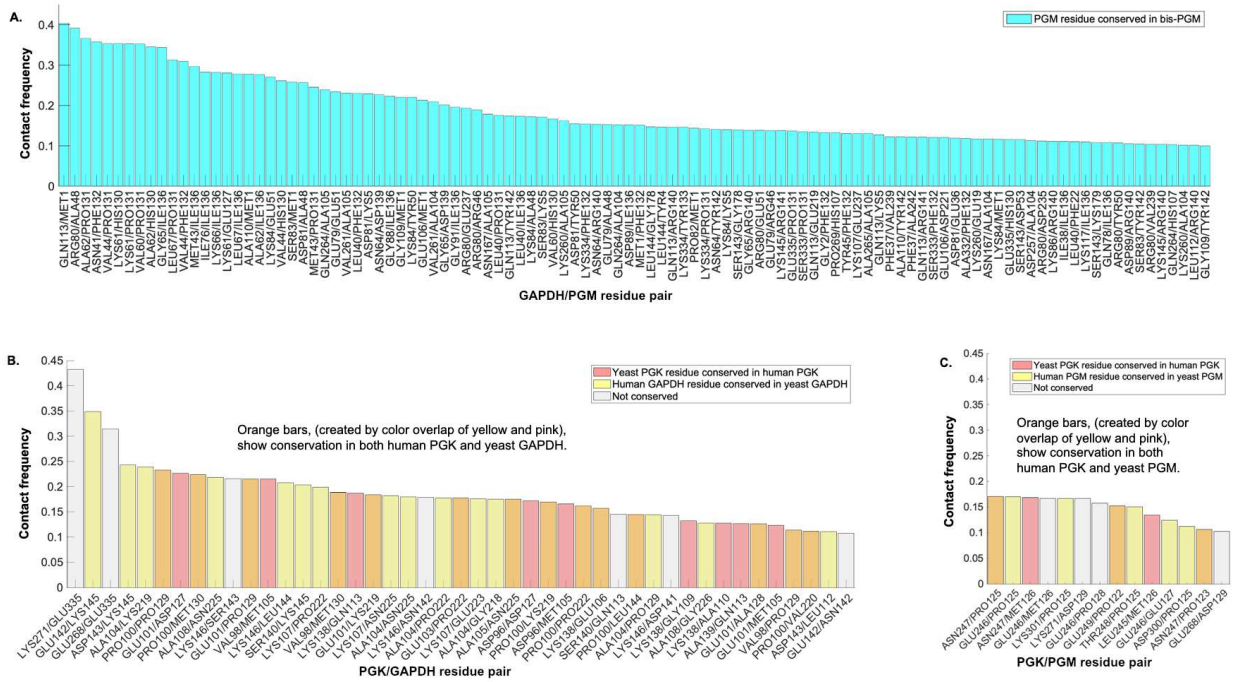
Ion/Water molecule	**Copy per model**
K^+	185 (U-2 OS model with human PGK)/ 188 (U-2 OS model with yeast PGK)
Na^+	5
Mg^{2+}	10
Cl^-	13
Water	51,339 (U-2 OS model with human PGK)/ 51,440 (U-2 OS model with yeast PGK)



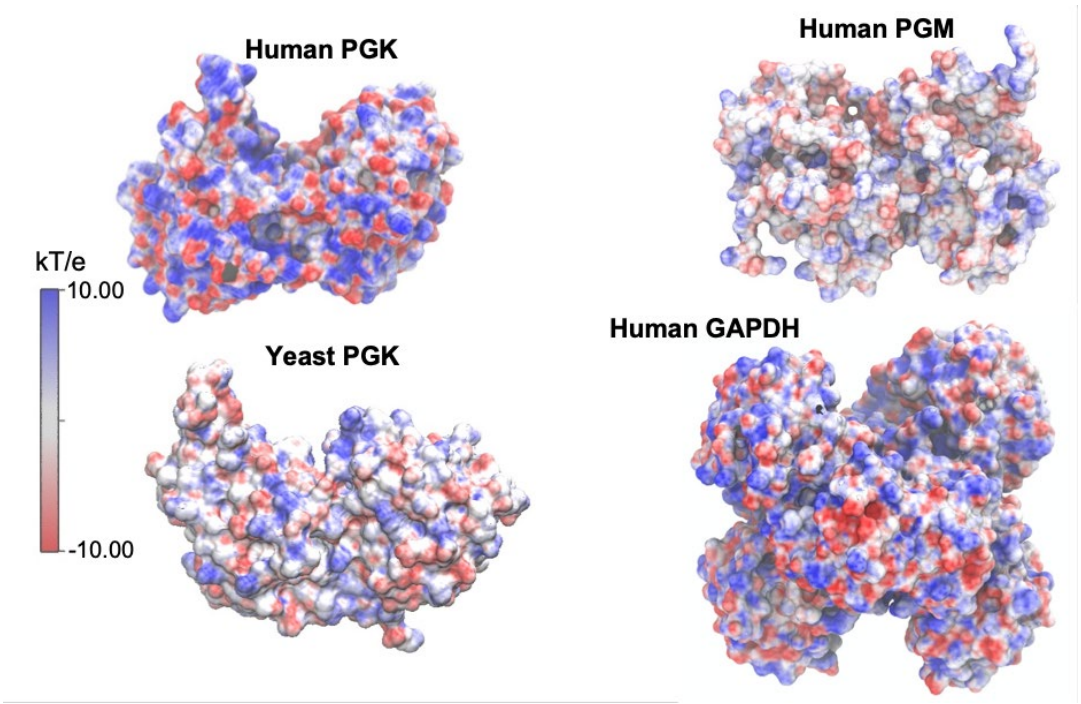
Supplementary Fig. 1. Additional glycolytic enzyme poses in the U-2 OS cell cytoplasm atomic models (A)-(C) Snapshots of the U-2 OS cell model with all-human glycolytic enzymes (PGK, GAPDH, and PGM). (D)-(F) Snapshots of the U-2 OS cell model with yeast PGK and human GAPDH and PGM. (A). Color labeling of enzymes as in the main text, and yeast PGK backbone in (D)-(F) additionally represented as cyan ribbon. (A), (D) Snapshot of the U-2 OS cell cytoplasm atomic models at time 0 μs. All-human glycolytic enzymes forming transient complexes at (B) 17.6 μs and (C) 18.3 μs along a MD trajectory, and snapshots of yeast PGK and human glycolytic enzymes forming transient complexes at (E) 22 μs, and (F) 23.9 μs along a separate MD trajectory.



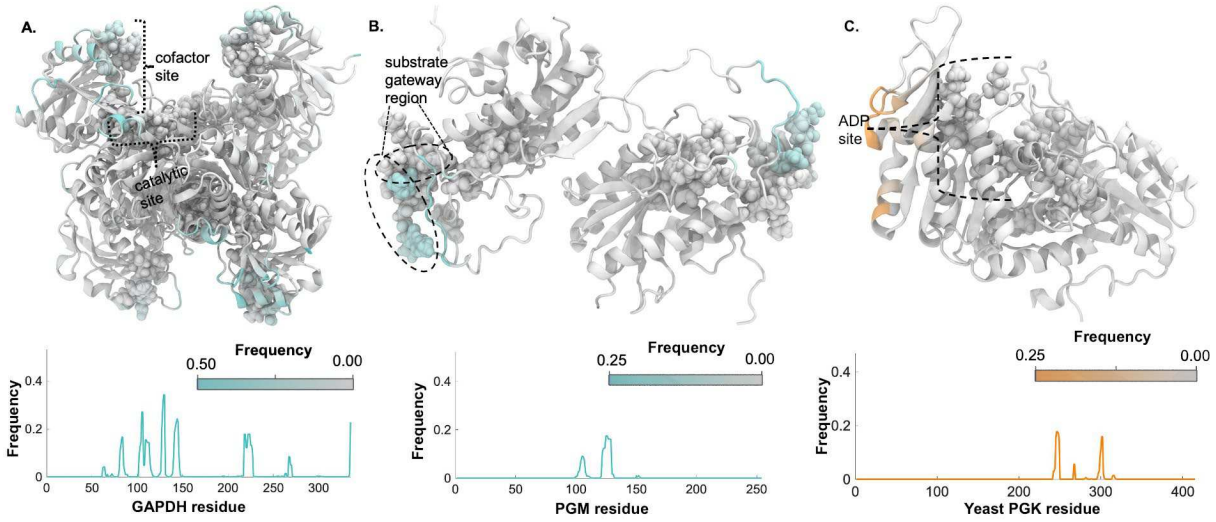
Supplementary Fig. 2. Additional time evolution of residue-residue contacts between glycolytic enzymes during simulations of the U-2 OS models with (A)-(F) and (M) from simulation of the U-2 OS model with yeast PGK, and (G)-(L) from simulation of the U-2 OS model with human PGK. Each contact is shown as a circle, undergoing color change from blue to light green with increasing time. The distance cutoff for the residue-residue contacts was 0.45 nm.



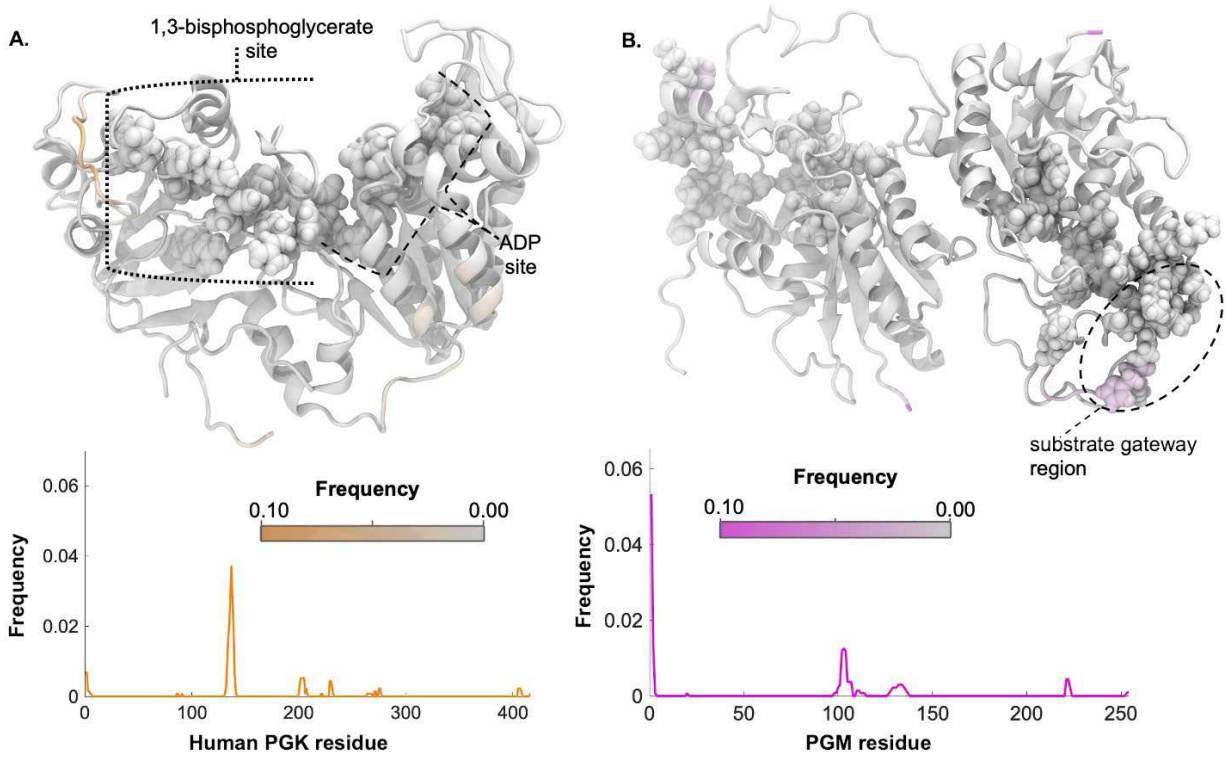
Supplementary Fig 3. Intra- and inter- species protein sequence conservation in most frequent contacts formed between (A) human GAPDH and PGM, (B) yeast PGK and human GAPDH, and (C) yeast PGK and GAPDH across simulations of either one or both U-2 OS cytoplasm models. Distance cut-off for contact formation is 0.45 nm, and contacts with frequency of at least $\geq 10\%$ of the simulation trajectories are shown here. For (A), interaction frequency of human GAPDH/PGM residue pairs were calculated across simulation trajectories of both the U-2 OS models combined and are shown here if involving PGM residues conserved across the human erythrocyte bis-PGM. (B) and (C) measurements were done across the simulation trajectory of the U-2 OS model with yeast PGK only. The bar plots in (B) and (C) are colored to reflect inter-species (yeast/human) enzyme residue conservation of homologs as follows: pink, yeast PGK residue is conserved in human; yellow, either human GAPDH (B) or PGM (C) is conserved in yeast; orange, both interacting residue pairs are conserved across both in human and yeast; and white, no cross-species residue conservation. The yeast PGK variant had 72 % sequence similarity with human PGK1; GAPDH was aligned simultaneously to the 3 variants of GAPDH in yeast with sequence similarities in the range of 76%; human PGM was aligned to yeast PGM 1 with 64 % sequence similarity; and human PGM was aligned to human erythrocyte bis-PGM with 68% sequence similarity.



Supplementary Fig. 4.| Electrostatic surface potentials of the glycolytic enzymes simulated in the U-2 OS cytoplasm models.



Supplementary Fig. 5. Sticky patches for yeast PGK specific enzyme-enzyme interactions in U-2 OS model simulation. Shown here are interaction frequency of either (A) GAPDH residues or (B) PGM residues with yeast PGK, and (C) Yeast PGK residues with PGM. Frequency of an enzyme residue forming a specific inter-enzyme contact at least once per simulation snapshot is visualized in here on an enzyme ribbon structure in terms of degree of color change from grey (e.g. grey to cyan for interaction with yeast PGK in (A) and (B)). In the enzyme structures, the active site regions are highlighted as van der Waals surfaces for the following enzyme residues. (A) GAPDH – cofactor binding site: R13, I14, D35, P36, F37, E79, R80, D81, T99, F102, S122, C152, N316, and Y320; catalytic site: S151, C152, T153, 179, T182, 185, T211, G212, and R234; (B) PGM – R10, H11, S14, S23, R62, E89, R90, Y92, K100, R116, R117, H186, N188, as well as the K106 to W115 region (which has been proposed as a gateway for substrate entry); and (C) Yeast PGK – 1,3-bisphosphoglycerate site: R21, D23, N25, R38, H62, Q65, N67, R121, T165, H167, R168, and H170; ADP site: G211, A212, G235, G236, F239, G310, L311, N334, G335, P336, G338, V339, E341, G371, D372, and T373. Descriptions of active sites on the structures are given for a subunit.



Supplementary Fig. 6. Sticky patches for human PGK specific enzyme-enzyme interactions in U-2 OS model simulation. Shown here are interaction frequency of (A) Human PGK residues with PGM, and (B) PGM residues with human PGK. Frequency of an enzyme residue forming a specific inter-enzyme contact at least once per simulation snapshot is visualized in here on an enzyme ribbon structure in terms of degree of color change from grey (e.g. grey to pink for interaction with human PGK in (B)). The active site region in (A) human PGK is highlighted as spherical van der Waals surfaces for the following enzyme residues: (i) 1,3-bisphosphoglycerate site: R21, D23, N25, R38, H62, R65, D67, R122, T167, H169, R170, and R172; and (ii) ADP site: G213, A214, G237, G238, F241, G312, L313, N336, G337, P338, G340, V341, E343, G373, D374, and T375. See Fig. 3 and Supplementary Fig. 5 for further details.