Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Reconfiguration of interdomain communication in hexokinase 1: A molecular basis for NEDVIBA childhood neurodevelopmental disorder

Alberto Galli
Independent scholar
Giovanni La Penna
giovanni.lapenna@cnr.it
National Research Council
Germano Nobili
University of Rome Tor Vergata
Silvia Rinaldi
National Research Council

Article

Keywords:

Posted Date: June 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9441083/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Reconfiguration of interdomain communication in hexokinase 1: A molecular basis for NEDVIBA childhood neurodevelopmental disorder

Alberto Galli[1], **Giovanni La Penna**[2,3,*,+], **Germano Nobili**[4], **and Silvia Rinaldi**[2,*,+]

[1]Independent scholar
[2]Institute of chemistry of organometallic compounds, National research council
[3]Roma-Tor Vergata section, National Institute for Nuclear Physics
[4]Department of physics, University of Roma Tor Vergata
[*]giovanni.lapenna@cnr.it, silvia.rinaldi@cnr.it
[+]These authors contributed equally to this work

ABSTRACT

NEDVIBA is a rare neurodevelopmental disorder caused by mutations in hexokinase 1 (HK1), the enzyme responsible for the first step of glucose metabolism in the brain. Fewer than 30 children have been diagnosed worldwide, and the molecular basis of the disease has remained unknown. Here, we use computer simulations to investigate how the T457M mutation, identified in a child with NEDVIBA, alters the properties of HK1 at the atomic level in water solution. We find that T457M does not destroy the enzyme's structure. Instead, it alters the internal communication network that allows the regulatory and catalytic parts of the enzyme to work together. Specifically, the mutation stabilises the structural element connecting the two halves of the protein, compacting the enzyme into a constrained conformation that prevents the catalytic domain from sampling the conformational states required for regulated activity. Detailed analysis shows that this effect originates from a localised change at a single hinge point, which propagates across the protein and selectively uncouples the regulatory domain from the rest of the enzyme. These results provide the first molecular explanation linking NEDVIBA to a specific structural target, the N-terminal hinge of the interdomain helix. This information is important to urgently develop therapeutic strategies.

Introduction

NEDVIBA (neurodevelopmental disorder with visual defects and brain anomalies) is a rare autosomal dominant disorder caused by monoallelic missense variants in hexokinase 1 (HK1), the predominant glucose-metabolising enzyme in the brain. Only a few tens of affected children have been identified worldwide[1–5], and the disorder has no disease-modifying treatment. Patients carry heterozygous point mutations presumed to impair HK1 function in the central nervous system, leading to abnormal brain metabolism and neurodevelopment. The disease is classified under OMIM code 618547[6] and ICD-11 code 6A00-6E8Z[7].

Hexokinases catalyse the transfer of a phosphate group from ATP to the 6-position of glucose, converting glucose to glucose-6-phosphate (G6P): the first and rate-limiting step of glycolysis[8–10]. Four isoforms (HK1–4) are expressed in mammals, differing in domain architecture, tissue distribution and regulatory properties[11]. HK4, also known as glucokinase, represents the ancestral single-domain form, in which substrate binding and feedback inhibition by G6P are handled within a single domain[12,13]. HK1, HK2, and HK3, by contrast, have evolved a two-domain architecture in which each domain is homologous to HK4 and the two halves are connected by a long interdomain α-helix, helix-a13[14–16]. A schematic representation of this architecture is shown in Fig. 1.

This two-domain organisation confers a striking functional asymmetry. The N-domain contributes negligible catalytic activity and functions primarily as a regulatory module, hosting the G6P inhibitory site and mediating attachment to the outer mitochondrial membrane (OMM) via the N-terminal mitochondria-binding domain (MBD)[17]. The C-domain, by contrast, carries the principal catalytic site responsible for glucose phosphorylation[18–20]. Among HK1–3, HK1 shows the strongest OMM association[11], enabling direct coupling of glycolytic activity to mitochondrial ATP output[21–24], through interaction with the voltage-dependent anion channel (VDAC)[25,26]. HK3, which lacks the MBD, does not associate with the OMM[11]. This spatial and functional organisation is of particular importance in neurons, which rely on tight coordination between glycolysis and oxidative phosphorylation to sustain the rapid shifts in energy demand that accompany electrical activity[27,28]. HK1 dysfunction is therefore expected to impair not only steady-state glucose metabolism but also the dynamic regulation of energy supply during neurodevelopment.

In this context, the identification of NEDVIBA-associated HK1 variants raises an unresolved mechanistic question about

how single amino acid substitutions can selectively impair neuronal metabolism without causing a global enzymatic collapse. Several of these variants cluster within or near helix-a13[1–5], pointing to the interdomain linker as a critical locus of pathogenesis. Yet despite growing clinical characterisation of the syndrome, no molecular mechanism has been proposed for any NEDVIBA variant to date, leaving the pathogenic basis of the disease entirely unexplained.

Here, we present a computational investigation of the structural and functional consequences of NEDVIBA-associated HK1 variants. We hypothesise that these missense mutations perturb HK1 allosteric regulation without disrupting the overall enzyme architecture, and that this selective perturbation underlies the pathogenic mechanism. Implications to NEDVIBA phenotype are far[29], but as a first step, that we present here, we focus on the isolated enzyme (*i.e.*, not bound to the OMM) and apply atomistic molecular dynamics simulations to investigate the structural and thermodynamic impact of a targeted set of single-point variants[30,31], as well as the long-range communication network controlling allosteric communication regulation[32]. Specifically, we simulated wild-type (WT) HK1 alongside: i) two NEDVIBA-associated mutations, T457M and G414E; ii) two non-NEDVIBA variants used as controls, D413A and S415A; iii) the non-natural T457A mutation, included as a negative control for T457M, to isolate the contribution of residue size from that of chemical character at position 457.

Results

Stability of helix-a13 region

To assess the impact of NEDVIBA-associated mutations on the structural integrity of HK1, we focused on the interdomain helix-a13 (residues 450–477), which acts as the central mechanical linker between the N- and C-terminal domains. As outlined in the Methods, the degree of helical order in this region was chosen as the collective variable (ξ or CV) driving metadynamics simulations, with s representing the afterwards imposed average number of backbone hydrogen bonds characteristic of an α-helix. The free energy change ΔG as a function of s for each variant and the wild-type was then computed, allowing a direct comparison of the thermodynamic cost of unfolding helix-a13 across sequences. In Fig. 2 the free energy in Eq. 1 is plotted as a function of the average s.

The largest well-sampled observed value s_0 is 21, while the theoretical maximal number of O(i)–HN(i+4) hydrogen bonds in the same region is 23. The free energy difference of sequences mutated in position 457 with respect to WT is positive, thus showing that the work required to demolish region helix-a13 becomes larger than WT when Thr in this position is changed. The maximal increase for T457A mutation is expected, since steric hindrance of Ala is lower than that of Met, while the change of sidechain from hydrophylic (Thr) to hydrophobic (Ala or Met) is similar. Changes in positions 413, 414 and 415, inside the N-terminal domain, induces an increase of the work to unfold helix-a13 closer to that in WT, consistently with the larger distance of these residues from the interdomain helix. To visualise the metastatistics, in Fig. S1 we display two configurations contributing to the statistics providing different values of ξ.

The description of the reasons and consequences of this displacement is in the following.

Structural modifications

Having established that NEDVIBA mutations close to helix-a13 (position 457) increase the thermodynamic stability of the protein, we next asked whether this has structural consequences on the relative arrangement and mobility of the two domains. In the following we restrict all comparisons to WT and mutations at position 457, being the latter that of main interest of our work.

We analysed the root-mean-square deviation (RMSD) of the N- and C-terminal domains using the single-domain HK4 as a structural reference, and monitored the opening and closing of the substrate and regulatory binding cavities through their solvent-accessible surface area (SASA). All structural quantities reported below were computed as weighted averages over the metadynamics ensemble using Eq. 5 (see Methods), which allows us to select configurations consistent with a given helical order s in the helix-a13 region. This reweighting enables a direct comparison of protein geometry at comparably high degree of interdomain helix folding, but with different sequence.

The distribution of RMSD with respect to crystal structure of single-domain HK4 (Fig. 3) shows that the wild-type (WT) sequence uniquely access to configurations significantly different from the crystal structure in the C-terminal domain: a major peak near 0.2 nm and a secondary peak near 0.4 nm, both visible at high helical order (s =21, black curve). None of the variants displays this bimodal behaviour where C-terminal domain is confined to a single basin (right panels). Notably, the broader C-terminal distribution in WT is accompanied by a tighter N-terminal domain, revealing a cross-domain coupling: the more the C-terminal domain explores open-like configurations, the more the N-terminal domain remains structurally ordered. This behaviour is confirmed when the analysis is extended to partially unfolded configurations (data not shown here). At intermediate helical order (s =10), the WT system shows a broader RMSD distribution, in particular in the N-terminal domain, with respect to T457M variant. T457A mutation which carries a mutation at the same position as T457M but with a less bulky substitution, displays an opposite trend: larger deviation in the C-terminal domain and, smaller in the N-terminal domain, consistently with the cross-domain coupling described above. These data, in summary, show that the wild-type sequence displays a wider plasticity that is suppressed by all variants and in particular by the mutations at position 457. A clear inverse relationship

emerges: the more constrained the C-terminal domain, the more disordered the N-terminal domain becomes, pointing to a mechanical coupling between the two halves of the protein transmitted through helix-a13.

To determine whether the deviations observed in Fig. 3 correspond to a genuine approach toward an open, glucose-accessible state, we repeated the analysis using the open conformation of HK4 as reference (Fig. 4, PDB 1V4T[13]). These data show that the wild-type sequence displays a wider plasticity that is suppressed by all variants and in particular by the mutations at position 457. The communication between N- and C-terminal domains is also significant: the tighter is C-terminal domain and the looser is the N-terminal domain. As explained in Introduction, the open conformation is a loose state of the single active domain, opening the pocket of the active site for glucose. The deviation is large in all cases, showing that the sampling of open conformation is not achieved. However, we notice that the WT sequence allows a significant decrease in RMSD compared to initial configurations (the vertical bars). This decrease in RMSD is not observed in the variants. By inspecting the correlation between RMSD of C-terminal domain with respect to closed HK4 and RMSD with respect to open HK4, we observe anti-correlation (data not shown here): the peak at 7.5 Å in the second RMSD is due to the peak at 4 Å in the first RMSD. Therefore, the distortion of the C-terminal domain in WT allows a partial opening of the glucose pocket.

The peculiar behaviour of the C-terminal domain in the WT sequence has consequences on the breathing of the two domains. We remind that the breathing between closed and open states is at the basis of the proposed mechanism of HK4[13] for the GLC activation catalysis. To quantify this breathing, we measured the solvent-accessible surface area (SASA) of the substrate-binding cavities in each domain (Fig. 5), monitoring both the glucose (GLC, left panels) and glucose-6-phosphate (G6P, right panels) binding sites. These are sites crucial for catalysis (GLC) and regulation (G6P) of the HK1 enzyme[33].

Vertical bars in Fig. 5 indicate the average SASA value of GLC pocket for the 50 HK4 walkers sampling closed (ca. 1 nm^2) and open state (ca. 5 nm^2). HK4 metastatistics (data not shown) represents the open cavities for both GLC and G6P: the SASA distribution is broad with maximum corresponding to the values measured for the open state, PDB 1V4T (the vertical bars indicates as "open" in Fig. 5). Notably, the HK4 metastatistics does not sample closed GLC pockets, consistently with the absence of glucose in the simulations, whereas glucose was present in the HK1 initial configurations[33,34], though deleted during simulations. The main difference between WT and the variants is the presence of two almost equally populated states in the WT Cd (first row, red curve), reflecting the open/closed oscillation, while T457M and T457A variants populate only a single state (second and third rows). The expansion of the GLC cavities in WT is correlated with that of G6P cavity in each of the domains (data not shown). Only the WT sequence allows the C-terminal domain to breath between closed and open states, while T457M and T457A are locked in closed and open states, respectively. Comparing the SASA distribution of HK1 variants with those of the sinlgle-domain HK4 isoform (HK4, data not shown), we notice that WT is the single sequence, among the 5 sequences investigated here, that displays two well-defined states approaching the closed (1V4S PDB) and the open (1V4T PDB) HK4 reference structures. We conclude that "breathing" in the C-terminal domain in a unique property of the WT sequence, absent in all variants. A more limited but still visible breathing is also present in the in the WT N-terminal domain.

T457M alters the inter-domain interactions

Given that T457M is the NEDVIBA variant carrying a mutation directly within the helix-a13 region, we focus the following structural analysis on the WT/T457M comparison, as it represents the most direct perturbation of the interdomain linker. To characterise how the rigidity of helix-a13 impacts the relative orientation of the two C-terminal and N-terminal domains, we analyzed their shape anisotropy via the inertia tensor. Both domains are characterized by a prolate mass distribution, as measured by the ratio between the 3 principal axis of the inertia tensor calculated for the two separated domains. Assuming z axis as the one associated to the lowest inertia eigenvalue (1), in Table S2 we report the ratio I_1/I_2 and I_1/I_3 for the metastatistics of the WT and T457M variant. The table shows that the two domains display prolate anisotropy, indicating that the z principal axis of each domain, $\vec{u}_1$, can be used to approximately define the direction of the long axis of each domain. The ratio values are similar between WT and T457M and between Nd and Cd, confirming that the overall elongated shape of each domain is preserved upon mutation. Fig. 6 shows the distributions of the inter-domain distance (top), the alignment of the $\vec{u}_1$ axes of Nd and Cd (middle, *i.e.* the absolute value of the scalar product between the $\vec{u}_1$ axes of Nd and Cd domains) and the mutual orientation of large sub-domains (bottom), all computed at a large helix-a13 ordering (s =21).

The sub-domain orientation is included to connect inter-domain geometry directly to the GLC pocket dynamics. Two differences emerge clearly. First, in T457M the two domains are closer in space by -1.5 Å, a compaction that is stabilized by inter-domain salt-bridges (see below). Second, WT populates configurations where the two domains are more parallel, thus making the whole protein elongated.T457M shows a significantly larger amount of bent protein configurations consistent with the reduced inter-domain distance. This explaines the bimodal RMSD and SASA distribution in WT: the elongated conformation is the one permitting C-terminal breathing. The large sub-domains follow the same trend with T457M showing a narrower, more constrained distribution. The T457M NEDVIBA variant allows, therefore, a significant bending of a less elongated protein architecture, which results in a more clused glucose pocket, while hampering a proper breathing of the catalytic Cd.

In summary, a clear causal chain emerges: the T457M mutation stabilizes helix-a13 (Fig. 2), which constrains the two domains into a bent, compacted architecture, which in turn locks the C-terminal domain in a single conformational state and suppresses the breathing of the glucose pocket (Fig. 5). In WT, by contrast, the slightly looser helix-a13 elongated protein conformation, is the structural prerequisite for C-terminal breathing and, ultimately, for proper opening of the glucose pocket.

Interactions stabilizing helix-a13

Having established that T457M alters the global geometry and suppresses the breathing of the C-terminal domain, we now ask what molecular interactions underlie this rigidification. All analyses in this section are performed at $s = 21$, focusing on configurations close to the native folded state of helix-a13.

The total number of hydrogen-bonds (HBs) in the helix-a13 (residues 450-477) were measured, both with the rest of the protein and with the (Fig. 7).

Upon mutation to Met, 457 sidechain loses its hydrogen-bonding capacity with the solvent, and the 2-3 HBs lost are compensated by 1-2 additional backbone HBs with protein atoms. The net result is a less solvated, more compact α-helical region in T457M compared to WT.

Fig. 8 displays the difference in the average number of salt-bridges (SBs) between the T457M and WT, mapped onto five HK1 regions.

The changes are modest within N-terminus (1-1) and helix-a13 (3-3), but becomes significant at intra-region level: SBs decrease within Nd (2-2) and between N-terminus and Nd (1-2), while increase in Cd (4-4). It must be noticed that this redistribution is accompanied by the formation of 1-2 additional SBs between positively charged Cd sidechains and negatively charged Nd sidechains (2-4). The molecular basis of this charge reorganization upon T457M mutation is more evident when looking at the SB distribution (Fig. 9).

Interestingly, the bimodal distribution of SBs within the Nd domain in WT corresponds to the bimodal SASA distribution of GLC and G6P cavities in the C-terminal domain (Fig. 5, WT red curves), a clear signature of long-range communication between salt-bridges network in Nd and the breathing dynamics in Cd. This coupling is lost in T457M. In WT, the increase of intra-N-terminus SBs is partially compensated by a release of salt-bridges between Nd and N-terminal region (1-2 interaction in Fig. 8), maintaining an overall looser electrostatic network. Conversely, T457M displays higher probability of forming 4-5 salt-bridges within Cd, indicating a tighter wrapping around the C-domain. Importantly, though the elongated protein architecture disfavours strong electrostatic interactions like salt-bridges, T457M uniquely forms 1-2 inter-domain salt-bridges between positive sidechains in Cd and negative sidechains in Nd (Fig. 8, 2-4 region). Closer inspection reveals these to be specifically between Arg 769 in Cd and AspGlu 251-252 in Nd (Fig. 10).

Crucially, residues 251 and 252 sit at the edge of the loop 241–253, precisely the N-terminal side of the putative allosteric interface identified by Aleshin et al.[12]. The formation of this salt-bridge in the T457M variant therefore acts as a biochemical "padlock" at this critical regulatory junction, sterically blocking the allosteric communication that is active in WT. One might expect that the widespread unfolding of helix-a13 would generate plethora of new inter-domain salt-bridges, as the two domains become free to rotate and expose many charged groups to each other (Fig. S2, right panel, and Fig. 6, middle panel). Surprisingly, however, only 1-2 additional SBs are formed at low ξ values compared to the folded state (high ξ values): the Nd/Cd interface remains remarkably stable regardless of helix-a13 order.

This finding reinforces the picture emerging from the padlock: the inter-domain electrostatic contacts in T457M are few but strategically placed. Taken together, the drying of helix-a13, the formation of one additional intra-helix salt-bridge and the establishment of the Arg(769)/Asp(251)-Glu(252) inter-domain padlock collectively rigidify T457M, compact the two domains, and lock the C-terminal domain in a single conformational state, suppressing the breathing that is the hallmark of the WT enzyme.

Interactions of mitochondria binding domain (MBD)

The mutation T457M changes the behaviour of the N-terminal region 1-21, known as the mitochondria binding domain (MBD). Experimentally, deletion of these initial residues in HK1 increases structural order, thus allowing crystallization[33,34], and detaching HK1 from mitochondria[11,17]. Since in our model the N-terminal region has been included, we monitor the changes in its order and the effect of this change to the N-terminal domain (Nd).

The distribution of SASA calculated for Nd, Cd, and MBD regions (see Table 1) differ significantly upon T457M mutation, when the statistics is that with $s = 21$. The data are not shown for brevity, but below is a summary.

The SASA of Nd in WT is smaller than in T457M: the maximum of distribution in WT is located at SASA 6 nm^2 lower than the maximum of T457M. As for MBD, the shift of maximum location is opposite and of about 3 nm^2. Therefore, MBD in WT is significantly more exposed to solvent than in T457M. It is interesting to notice that the average wrapping of MBD around Nd in T457M is associated to an expansion of Nd even if the surface exposed by MBD does not contribute to Nd SASA. The ATP binding pocket proposed in Nd[12,33] (see Table 1) has similar size in WT and T457M (8 nm^2), but in WT there is a small

population representing a cavity of 6.5 nm^2, that is smaller than the average. The suppression of a pre-organized ATP-binding cavity in T457M is an effect of the Nd average rigidification, similarly to what happens to Cd breathing.

Asymmetric dynamic reorganisation of N- and C-terminal domain

While the above analyses characterise the rigidification of helix-a13 in terms of hydrogen bonds and salt-bridge networks, they do not resolve how the perturbation propagates across the protein at residue resolution or which specific structural elements underlie the change in helix dynamics. To address these questions, we performed distance fluctuation (DF) and local fluctuation (LF) analyses on unbiased molecular dynamics simulations (see Methods). DF quantifies the mean-square fluctuation of pairwise Cα distances and provides a measure of dynamic coupling between residues. To restrict the analysis to the structured allosteric network, only residue pairs belonging to ordered secondary structure elements (α-helix, β-strand, or isolated β-bridge) and with WT DF $<$4.0 $Å^2$ were retained, corresponding to the lower 73% of structured pairs. All comparisons were performed within this filtered regime.

DF analysis reveals a strongly asymmetric dynamic reorganisation upon mutation (Fig. 11.

The N-domain is the primary site of perturbation, while the C-domain remains largely unaffected. The mean DF of structured N-domain pairs increases by +61.3%: WT 0.741$\pm$0.769 $Å^2$; T457M 1.196$\pm$1.659 $Å^2$. These values indicate a global reduction in dynamic coupling. This effect is not uniform: the N-domain core (residues 240–415) becomes more internally rigid, intra-DF -19.1% (WT 0.327 $Å^2$; T457M 0.264 $Å^2$), while its coupling with adjacent regions decreases, as shown by the increase in DF between the core and the linker entry region (+36.6%). The core therefore behaves as a dynamically more isolated unit, internally coherent but less coupled to the surrounding network. In contrast, the C-domain shows only minor changes (intra-DF -8.7%), indicating that its internal dynamic organisation is largely preserved. Inter-domain DF between N- and C-terminal domains increases by +11.7%, reflecting a reduction in cross-domain dynamic coupling, consistent with the suppression of C-terminal breathing observed in the structural analysis. The linker α-helix (residues 449–478, helix-a13) is the structural element where the perturbation is most localized. The mean intra-helix DF increases from 0.132$\pm$0.118 $Å^2$ (WT) to 0.270$\pm$0.364 $Å^2$ (T457M), indicating increased internal dynamic heterogeneity. To resolve the spatial distribution of this effect, we performed a local fluctuation (LF) analysis (Fig. 12.

LF provides a residue-resolved measure of backbone mobility and reveals that the perturbation is not uniformly distributed along the helix but is localized at its N-terminal region. The helical core (residues 453–476), including the mutation site Thr457, shows negligible changes in backbone mobility (ΔLF < 0.003 Å), indicating preservation of secondary structure. In contrast, residues at the N-terminal hinge (447 and 449) show substantial increases in mobility (+333% and +356%, respectively), while the C-terminal end (residue 477) shows reduced mobility (-44%). This redistribution identifies the molecular mechanism precisely: the loss of the Thr457 hydroxyl group disrupts the polar anchoring network at residues 447-449, converting a fixed hinge into a loose joint. The helix remains structurally intact, but its anchor point is mobile, it now pivots on a loose joint rather than a fixed one, sampling a broader range of orientations relative to the N-domain.

The hinge-localized increase in linker mobility is accompanied by a marked reduction in dynamic coupling between the linker and the N-domain regulatory pockets. The mean DF between residues of the G6P-binding pocket in the N-domain and residues of the linker helix shows the largest increase across the dataset (+245.5%; WT 0.411$\pm$0.273 $Å^2$; T457M 1.421$\pm$0.644 $Å^2$). The coupling between the glucose-binding pocket and the linker also increases substantially (+46.9%; WT 1.074$\pm$1.006 $Å^2$; T457M 1.577$\pm$0.895 $Å^2$). In parallel, intra-pocket DF values increase within the N-domain, with G6P-N and GLC-N pockets showing increases of +177.4% and +134.8%, respectively, indicating reduced internal dynamic coupling. These results show that the regulatory pockets become dynamically decoupled from the linker region upon mutation. In contrast, the catalytic pockets in the C-domain are only weakly affected. The G6P-C pocket shows a decrease in intra-DF (-32.7%; WT 1.089$\pm$1.112 $Å^2$; T457M 0.733$\pm$0.621 $Å^2$), consistent with moderate internal rigidification, while the glucose-binding pocket shows only minor changes (+5.8%). The coupling between linker and C-domain increases only slightly (+4.8%), indicating limited propagation of the perturbation into the catalytic region. Overall, the mutation induces a strongly asymmetric reorganisation of the allosteric network: hinge-localized perturbation of the linker selectively disrupts dynamic coupling with the N-domain regulatory region, while the catalytic C-domain remains largely unaffected.

Energy Decomposition Analysis: Energetic Correlates of Nd Allosteric Reorganisation

To determine whether the redistribution of dynamic coupling identified by DF and LF analyses is accompanied by changes in the underlying interaction energy network, we performed an energy decomposition analysis (EDM, Fig. 13).

Per-residue interaction energy profiles were derived from the representative conformation of the most populated MD cluster, thus providing a residue-level description of non-bonded interaction patterns in the dominant sampled conformational state (see Methods). Two clusters within the N-domain show markedly increased per-residue interaction energy in T457M relative to WT. The first comprises β-strands 239-256 (ΔE = +0.011 to +0.305 kcal/mol, up to $\times$6.5 above WT), located at the interface between the small and large N-domain lobes, 17-21 Å from the mutation site. The second comprises α-helices 393-401 (ΔE = +0.063 to +0.245 kcal/mol, up to $\times$12 above WT), within the N-domain helical core, 25-31 Å from T457. Both clusters

correspond to structural elements showing the largest DF increases within the N-domain, indicating that the redistribution of interaction energies is spatially coherent with the loss of dynamic coupling. This pattern is consistent with propagation of the linker-associated perturbation into the adjacent N-domain framework.

An opposite pattern is observed in helices 265-315, corresponding to the glucose-binding lobe of the N-domain, located 33-45 Å from T457 (ΔE = -0.010 to -0.210 kcal/mol, reductions of up to -84% relative to WT). In the WT, the central residues of this cluster (approximately 273-282) display the highest absolute interaction energies in the entire N-domain ($0.17 \div 0.30$ kcal/mol), indicating strong energetic engagement with the surrounding network. In T457M, this engagement is substantially reduced. Backbone mobility in this region remains essentially unchanged ($\Delta LF \simeq 0$), indicating that the reduction in interaction energy is not associated with major local backbone destabilisation. This dissociation between preserved backbone mobility and reduced interaction energy indicates that the loss of energetic engagement does not reflect conformational arrest but rather a decoupling from the network of correlated interactions that underlies allosteric communication. Together with the decrease in intra-region DF (-12.6%), this suggests reduced energetic and dynamic coupling of the glucose-binding lobe within the broader N-domain network.

The combined EDM, DF, and LF analyses therefore indicate a bipartite redistribution of interaction energy within the N-domain. Regions proximal to the linker show increased interaction energy together with altered dynamic coupling, whereas the glucose-binding lobe shows reduced energetic engagement despite preserved local backbone mobility. This redistribution is consistent with an asymmetric reorganisation of coupling within the N-domain centred on the linker region.

Discussion

The results presented here converge on a multiscale mechanistic model of the T457M substitution in HK1. Across thermodynamic, structural, dynamic, and energetic analyses, the mutation does not disrupt the overall architecture of the enzyme, but alters the internal communication network that allows the two domains to work in a coordinated manner. At the thermodynamic level, metadynamics simulations show that T457M raises the free energy cost of locally unfolding helix-a13, indicating net stabilisation of the interdomain linker relative to the wild type. In multidomain enzymes, such helices act as mechanical relays transmitting regulatory signals between distant parts of the protein, a role confirmed experimentally in HK1, where disruption of helix-a13 abolishes allosteric communication between the N- and C-terminal domains[12,15,16]. Consistently, structural analyses show that stabilisation of helix-a13 is accompanied by compaction of the enzyme into a bent conformation that suppresses the conformational flexibility of the C-terminal domain. In the wild-type enzyme, this domain fluctuates between partially open and closed states, with increased solvent accessibility of the glucose-binding pocket in the open state, a dynamic behaviour analogous to that observed in related hexokinases and generally considered essential for regulated catalytic activity. T457M abolishes this flexibility, locking the catalytic domain in a single rigid state and impairing the enzyme's ability to respond to regulatory inputs.

At the molecular level, rigidification arises from complementary effects operating at different scales. Replacement of the Thr457 hydroxyl group with the bulkier and apolar methionine reduces the interaction of helix-a13 with the surrounding solvent, favouring a more compact and less hydrated helical region. In parallel, the mutation promotes the formation of a specific inter-domain salt-bridge between Arg769 in the C-domain and Asp251-Glu252 in the N-domain (residues at the edge of loop 241-253, previously identified as the N-terminal side of the allosteric interface)[12]. This interaction acts as a molecular padlock that locks the two domains into a constrained geometry, shifting the enzyme away from its dynamically adjustable wild-type state. To resolve how this structural perturbation propagates at residue resolution, we applied distance fluctuation (DF) and local fluctuation (LF) analyses. These analyses show that the effect originates from a highly localised perturbation at the N-terminal hinge of helix-a13, rather than from a global rearrangement of the protein. While the helical core (residues 453-476) remains structurally and dynamically intact, the hinge region (residues 447-449) exhibits a marked increase in mobility, converting the anchoring point of the helix into a loose joint and allowing it to pivot over a broader range of orientations relative to the N-domain. Noticeably, residue Ser449 of the hinge region is also involved in G6P binding (Table 1). This localised perturbation propagates asymmetrically across the protein. The N-domain undergoes substantial reorganisation: its core (residues 240-415) becomes more internally rigid while its coupling to adjacent regions decreases, in line with established models of allostery in which mutations selectively reweight existing interaction networks rather than causing uniform changes in flexibility [35,36]. The functional consequence is a decoupling of the regulatory N-domain from the interdomain interface: the glucose- and G6P-binding regions show both reduced internal dynamic coupling and weakened communication with the linker, expected to selectively impair the transmission of regulatory signals to the catalytic domain without disrupting catalysis itself[18,20].

Energy decomposition analysis (EDM), which measures how strongly residues interact with their structural neighbours, provides an independent energetic perspective consistent with the dynamic results. Within the N-domain, two distinct structural regions remote from the mutation site show increased energetic strain in T457M relative to wild type: β-strands 239-256 and α-helices 393-401, confirming that the effects of T457M extend well beyond its immediate structural context. In contrast, the glucose-binding lobe (residues 265-315) shows markedly reduced energetic engagement despite preserved backbone

mobility. This region is not damaged, but effectively silenced, withdrawn from the interaction network that underlies allosteric communication in the N-domain.

Addition of G6P has been observed detaching HK1 from VDAC and OMM[37]. Thus, the lower MBD hydration and the lower sensitivity to G6P found here in T457M with respect to WT indicate a strong possible effect of T457M mutation on HK1/OMM interactions.

These findings offer a mechanistic basis consistent with two key clinical observations. The selective impairment of regulatory coupling (with catalytic activity preserved, but its modulation compromised) may account for the autosomal dominant inheritance of NEDVIBA, where a single defective copy interferes with regulation without abolishing catalysis[1–5]. It also explains the particular vulnerability of neuronal tissue: neurons rely on tight, dynamic regulation of glucose flux to sustain rapid metabolic transitions, and even a subtle impairment of regulatory coupling is expected to have serious consequences in the developing brain, where metabolic flexibility is a prerequisite for proper circuit formation and function. The structural model generates several experimentally testable hypotheses. T457M-carrying HK1 may display altered sensitivity to regulatory inputs, including product-mediated inhibition. The inter-domain salt-bridge between Arg769 and Asp251-Glu252, absent in wild-type HK1, represents a candidate contact point for future structural studies. Most importantly, the N-terminal hinge of helix-a13 emerges as the minimal locus from which pathogenic effects propagate, a structurally defined site that may inform efforts to modulate HK1 allosteric behaviour, although formal druggability assessment lies beyond the scope of this work.

Our findings are necessarily bounded by the scope of the computational approach. All simulations were performed on the isolated enzyme in the absence of substrates and inhibitors, and the influence of ligand binding, mitochondrial membrane association, and the heterozygous cellular context remain open questions that experimental work will need to address. Biochemical, biophysical, and structural characterisation of T457M HK1 represent the most immediate next steps, and we hope this work will encourage biochemists, structural biologists, and clinicians to direct their efforts toward a condition that currently has no disease-modifying treatment. The identification of a precise structural locus and experimentally verifiable hypotheses are offered as a call to the broader scientific community to engage with this disease.

Methods

The sequence of human HK1 was found in GenBank with the Uniprot ID P19367. This sequence will be indicated as wild-type (WT, hereafter). The protein is composed by a single chain of 917 aminoacids. In all our models, His residues were assumed neutral and protonated at Nδ1 position, Arg and Lys were positively charged, Glu and Asp were negatively charged. Terminal residues were assumed to be charged.

We studied several variants of the WT sequence: D413A; **G414E**; S415A; **T457M**; T457A. NEDVIBA variants are boldface. To perform structural comparisons we used the HK4 single-domain human protein. Its sequence was extracted by GenBank with Uniprot ID P35557.

The HK4, HK1 WT and 5 variants above were assigned to different initial coordinates and simulated, as described in the following, within a multiple-walkers metadynamics framework. The multiple-walkers metadynamics statistics was then modulated to reproduce a given average of the same collective variable used in metadynamics. The methods are described in details in Refs. 30, 31.

Initial coordinates for HK1 and HK4

The basis of multiple-walkers metadynamics is to spread by means of molecular dynamics (MD) algorithm initial conditions among different walkers to efficiently sample a range of a chosen collective variable (CV). Initial atomic coordinates for the multiple-walker metadynamics simulations of HK1 were derived from different high-resolution X-ray crystal structures to ensure broad initial conformational sampling. We set the number of walkers to 100 in all cases. As for HK1, walkers 1–50 were initialized using the coordinates from PDB ID 1QHA[33], while walkers 51–100 utilized the coordinates from PDB ID 1HKC[34]. The backbone heavy atoms root-mean square deviation (RMSD, see below) between the two structures is 0.4 nm. The WT protein was first subjected to a standard equilibration protocol. To generate the mutant trajectories, the respective point mutations were introduced computationally at the conclusion of 4-ns equilibration stage (see below).

To establish a baseline for the breathing dynamics of a single-domain hexokinase, parallel simulations of hexokinase 4 (HK4, glucokinase) were conducted. Similar to the HK1 setup, the initial coordinates for the HK4 multiple-walker simulations were partitioned to sample distinct functional states: walkers 1–50 were initialized from the closed-form crystal structure (PDB ID 1V4S[13]); walkers 51–100 were initialized from the open-form crystal structure (PDB ID 1V4T[13]). To allow direct comparative analyses of the structural dynamics between the two-domain HK1 and the single-domain HK4 architectures, the sequences were aligned using the BLAST algorithm[38,39].

Since in both HK1 and HK4 crystal structures coordinates of disordered regions are missing, we built the missing segments (residues 1-21) by using self-avoiding Monte Carlo random walks. Therefore, for instance, the samples in walkers 1-50, all obtained starting from PDB 1QHA, differ for the atomic positions of residues in region 1-21. Each initial protein coordinates

set was then inserted into an orthorombic simulation cell with an empty space in the 3 direction of space of 1.8 nm. These simulation cells were filled with water molecules, removing those molecules closer than 0.12 nm. Because of the different protein sizes, the simulation cells were all composed by a different number of water molecules. When water molecules were randomly replaced by Na^+ and Cl^- ions at 0.15 M concentration up to charge neutralization, all simulation cells (100 in number) were made different in terms of solvent conditions. For instance, in WT sequence the 14411 protein atoms in the 100 samples were added of 25556-25585 water molecules, 223 Na^+ ions and 216 Cl^- ions, depending on the initial conditions of N-terminus. The sizes of simulation cell were: L_x in the range 11.98÷15.12 nm; L_y 16.28÷20.02 nm; L_z 9.78÷11.58 nm.

Mutated proteins were constructed by simply replacing the amino acid side chain at the point of the mutation with the mutated one. This operation was carried out in the configurations of the WT sequence collected at the end of equilibration (stage 1 in Table S1). A standard geometry for the modified side-chain was assumed. Any possible atomic clash of the replaced side chain with nearby solvent or protein atoms were dealt with by energy minimization. Once this was done, the simulation procedure outlined in Table S1 for the mutated protein was followed.

Metadynamics

The procedure to obtain a metastatistics with a chosen collective variable (CV, usually indicated as ξ) distributed within a given range as uniformly as possible is described in Ref. 30. In the following we briefly describe what is specific for the HK1 protein object of this work.

The external bias was built with the help of the altruistic metadynamics method[40]. The reason for having multiple walkers is to be able to sample all the values of CV in the physically accessible range $0 < \xi < \xi_{max}$ as uniformly as possible. The biasing potential, V_G, is progressively built by summing over gaussian functions of the CV (each gaussian function is taken to have unit height and width). Gaussian functions are deposited every 20 ps along the MD simulation.

The number of points in the CV range was 25 ($0 < \xi < 25$), consistently with the width of the gaussian functions. The maximal value equal to 25 was chosen to accommodate possible structures with a number of hydrogen bonds larger than that occurring in the crystal structure equal to 24 (all possible i-$i+4$ hydrogen-bonds within region spanning helix-a13, residues 450-477. To avoid the complete unfolding of region helix-a13, a harmonic wall potential $U_w = \frac{k}{2}(\xi - \xi_0)^2$ with k = 100 kcal/mol and $\xi_0 = 1$ was added to the bias when $\xi < 1$.

Each walker was simulated for 10 ns in a thermal bath of explicit water molecules plus Na^+ and Cl^- ions, as explained above, at $T = 300$ K and $P = 1$ bar, with an external bias built according to the procedure described below. The GROMACS package[41] was used for the MD simulations (version 2022.5), and VMD[42] for part of the analysis and visualization. The metadynamics simulations were performed with the same GROMACS package used for MD simulations and with the additional external forces as implemented in PLUMED 2.9[43–45].

The equation of motion with periodic boundary conditions in three dimensions were integrated by using the velocity Verlet algorithm with a time-step of 2 fs. The bond distances involving H atoms were constrained with the LINCS method[46] to the equilibrium force-field value. Non-bonding interactions had a cut-off at 1 nm and the smooth particle mesh Ewald algorithm[47] for long-range electrostatic interactions was employed. The deterministic thermostat[48,49] and a stochastic barostat[50] to keep temperature and pressure, respectively, constant at 300 K and 1 bar were used. The simulation cell was kept orthorhombic and isotropic while solving the equations of motion. Temperature and pressure coupling constants were 5 and 0.5 ps^{-1}, respectively. Room conditions water compressibility was used. The force-field was CHARMM36[51] with NaCl[52] and TIP3P water model[53] for the solvent bath. A conversion tool from CHARMM to GROMACS data-sets was used[54]. Table S1 reports some of the simulation parameters and conditions we used to evolve the 100 walkers.

In order to try to get a more uniform exploration of the set of possible CV values we made use of a variant of metadynamics, called altruistic metadynamics[55]. It consists in periodically assigning to the walkers a biasing potential obtained as a linear combination of the accumulated biasing potential of all the walkers up to that moment. According to Equation (3) of Ref. 55 this linear combination depends on two parameters, $0 \leq \alpha \leq 1$ and $w \leq 1$, which need to be suitably chosen. The actual values of α and w we used in the successive stages of the simulation are reported in Table S1. Where values of α and w are not indicated, it is intended that no such linear combination is formed and each walker independently from the others builds up its own biasing potential.

We remind that our goal is not that of using metadynamics to build a free energy profile as a function of the chosen CV, rather to have walkers "spread" over different values of CV so as to obtain a large and diversified metastatistics, exploring configurations spanning all the accessible CV values, that can be modulated by the maximal constrained entropy (MEC) method summarized in the following.

Free energy change

We used the maximal constrained entropy (MEC) approach which allows to properly re-modulate a collected set of sampled configurations by imposing suitable conformational constraints. As in previous works, we impose a single conformational constraint, related to the structural change of interest, to a biased simulation obtained by multiple-walkers metadynamics. In

this way we can reliably monitor the change in free energy as a function of average values of the chosen collective variable used as the handle of the external bias. The quantity of interest to be calculated is therefore:

$$\Delta G_X = G_X(s) - G_X(s_0) \ , \tag{1}$$

where the subscript X denotes the protein sequence and the s state variable indicates the degree of structural order of interest, where s_0 indicates the folded native state and the range of s allows to sample unfolded states. For each sequence X, the function G is approximated as:

$$G(s) = \langle H \rangle_\lambda - T \ k_B \ \overline{S}_c(s), \tag{2}$$

where H is microscopic enthalpy, $H = U + TP$, T the absolute temperature, P the pressure measured by the simulation cell, k_B the Boltzmann constant, and $\overline{S}_c$ is the information entropy as determined by the maximal constrained entropy method. The parameter λ is univocally associated to the information entropy and it is, therefore, determined by the average s imposed to the biased simulation.

In order to reduce the strong effect of pressure fluctuations exerted by the solvent molecules, an effective solvation free energy can replace H. As it is customary done in molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) method[56], we used an approximate evaluation of $\langle H \rangle_\lambda$, where H is replaced by the effective mean-field free energy $\overline{U}$ of the protein solute. For each protein configuration Q one writes

$$\overline{U}(Q) = U_{intra}(Q) + G_{solv,np}(Q) + G_{solv,pol}(Q), \tag{3}$$

where U_{intra} is the intra-molecular part of the potential energy in the protein force-field, given by

$$U_{intra}(Q) = U_{str}(Q) + U_{bend}(Q) + U_{tors}(Q) + U_{vdw}(Q) + U_{el}(Q). \tag{4}$$

The various contributions are the stretching (U_{str}), bending (U_{bend}), and torsional (U_{tors}) terms in the potential. U_{vdw} and U_{el} are the Lennard–Jones and Coulomb interactions, respectively, computed by summing over all the pairs of atoms of the protein. The last two terms in Eq. (3) represent solute-solvent contributions to free energy at fixed Q.

The averaging is performed at the given thermodynamic state variables P and T used in the simulation of the whole system. Within this mean-field assumption, the solute and the solvent are made independent.

This is a strong approximation, since the chosen collective variable contains the number of hydrogen bonds within protein groups and once a single intramolecular hydrogen bond is broken there is a large chance for the formation of hydrogen bonds with the water molecules in the protein environment where the breaking event occurs. On the other hand, this elementary change of free energy, that does not involve a wide change in protein structure, can be calculated within the MM/PBSA approximation as the sum of $G_{solv,np}$ and $G_{solv,pol}$. The approximation is highly effective for computing the free energy differences ($\Delta\Delta G$) between variants and wild-type sequences, as systematic errors inherent to implicit solvation largely cancel out. Because helix-α13 (residues 450-477) acts as the central mechanical linkage between the two domains, its local unfolding serves as the primary bottleneck for large-scale inter-domain conformational shifts, making it a robust CV for driving metadynamics. The calculation of $G_{solv,np}$ and $G_{solv,pol}$ is described in the Supplementary Material.

Each average quantity can be computed according to the statistics associated to each free energy value by using the corresponding λ value. To accomplish this task, average of any configuration quantity B is obtained using the following equation:

$$b_\lambda = \langle B \rangle_\lambda = \frac{1}{Z_\lambda} \sum_\gamma \widetilde{P}_\gamma \exp(-\lambda \xi_\gamma) \ B(q_\gamma) \ , \tag{5}$$

with B whatever function of configuration q, γ identifying the configuration microstate, λ the modulation parameter obtained by imposing the average of the chosen CV (ξ, with $s = \langle \xi \rangle_\lambda$) and $Z_\lambda = \sum_\gamma \widetilde{P}_\gamma \exp(-\lambda \xi_\gamma)$. The probability $\widetilde{P}_\gamma$ is the un-modulated probability of configuration γ, that is, within approximations above, the one produced by metadynamics.

Long range Analyses: Distance Fluctuation (DF) and Energy Decomposition Matrix (EDM)

To correlate conformational rearrangements with the internal dynamics of each protein, distance fluctuations analysis was used to describe the dynamic coordination between any two residues[32]. Distance fluctuation $\Delta F_{i,j}$ is defined as the time-average mean square fluctuation of the distance $r_{i,j}$ between Cα atoms of residues i and j:

$$\Delta F_{i,j} = \langle (r_{i,j} - \langle r_{i,j} \rangle)^2 \rangle \tag{6}$$

where angular brackets indicate the time-average over the trajectory. Low ΔF values indicate highly coordinated residues. The energy decomposition method (EDM) yields an interaction matrix M with elements $M_{i,j}$, obtained by averaging the interaction energies between residue pairs, comprising all the nonbonded inter-residue atomic energy components (namely, van der Waals and electrostatic), calculated over the structures visited during a MD trajectory or the representative conformation of the most populated cluster. The method builds on a simplified picture of the most relevant residue–residue interactions in a certain fold. For a protein of N residues, this calculation produces an $N \times N$ matrix. The matrix M can be diagonalized and re-expressed in terms of eigenvalues and eigenvectors, in the form:

$$M_{i,j} = \sum_{k=1}^{N} \lambda_k w_{i,k} w_{j,k} \ , \tag{7}$$

where N is the number of amino acids in the protein, λ_k are M eigenvalues, and $w_{i,k}$ is the i-th component of the associated normalized eigenvector k. The matrix of the pair energy-couplings corresponding to the first eigenvector is chosen to filter the contact map and identify all inter-residues higher energy contributions[35].

To characterise the intrinsic dynamics of WT and T457M HK1, two independent unbiased molecular dynamics (MD) simulations of 500 ns each were performed for each system: one started from the crystal structure (PDB ID 1QHA for WT[33]), *i.e.* the same initial structure of walker 1 in metadynamics; the second simulation started from the T457M sequence with the mutation introduced as in walker 1 of metadynamics. MD simulations were performed using the AMBER[57] software package with the ff14SB force-field[58] and TIP3P water model. As in the equilibration phase described above, each system was simulated in an NPT ensemble at constant temperature (300 K) and pressure (1 atm). The simulation parameters were similar to those used in metadynamics, but the external potential (bias) was absent. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm. The use of unbiased simulations ensures that ΔF values reflect the natural conformational dynamics of each system without contributions from the external bias potential. The use of two independent replicas provides an assessment of the reproducibility of the observed dynamic coupling patterns. ΔF analysis was performed on the combined trajectory obtained by concatenating the two replicas (total 1-μs trajectory per system). EDM analysis was performed on the representative conformation of the most populated cluster extracted from the combined trajectory (see above).

Definitions of structural parameters, protein domains and pockets

Root-mean-square deviation (RMSD) calculations were used to continuously monitor domain closure, pocket expansion, and overall structural drifts. These deviations were calculated using three primary reference structures: the native HK1 state (PDB ID 1QHA[33] and 1HKC[34]); the closed conformation of HK4 (PDB ID 1V4S[13]); the open conformation of HK4 (PDB ID 1V4T[13]). In the reported RMSD calculations only heavy backbone atoms were used (N, C, O, and Cα). The deviation is always minimized removing center of mass translation and rotation centered on it[59].

Domains and regions were defined according to Table 1, based on crystal structures and functional[1] analyses.

References

1. Okur, V. *et al.* De novo variants in HK1 associated with neurodevelopmental abnormalities and visual impairment. *Eur. J. Hum. Gen.* **27**, 1081–1089, DOI: 10.1038/s41431-019-0366-9 (2019).
2. Peretz, R. H. *et al.* A de novo hexokinase 1 (HK1) variant presenting as boucher–neuhäuser syndrome. *Am. J. Med. Genet. A* **191**, 624–629, DOI: 10.1002/ajmg.a.63045 (2023).
3. Poole, R. L. *et al.* Expanding the neurodevelopmental phenotype associated with HK1 de novo heterozygous missense variants. *Eur. J. Med. Genet.* **66**, 104696, DOI: 10.1016/j.ejmg.2023.104696 (2023).
4. Wortmann, S. B. *et al.* Clinical, neuroimaging, and metabolic footprint of the neurodevelopmental disorder caused by monoallelic HK1 variants. *Neurol. Genet.* **10**, e200146, DOI: 10.1212/NXG.0000000000200146 (2024).
5. Ng, B. G. *et al.* Autosomal dominant HK1-related neurodevelopmental disorder with visual defects and brain anomalies (nedviba): An emerging mitochondrial disorder. *Genet. Medicine Open* DOI: 10.1016/j.gimo.2025.103425 (2025).
6. https://www.omim.org.
7. https://icd.who.int.
8. Nelson, D. L., Cox, M. M. & Lehninger, A. L. *Lehninger Principles of biochemistry* (W.H. Freeman and Company, New York, 2013).
9. Wilson, J. E. *Hexokinases*, 65–198 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995).
10. Cárdenas, M. L., Cornish-Bowden, A. & Ureta, T. Evolution and regulatory role of the hexokinases. *Biochim. Biophys. Acta, Mol. Cell Res.* **1401**, 242–264, DOI: 10.1016/S0167-4889(97)00150-X (1998).

11. Reid, S. & Masters, C. On the developmental properties and tissue interactions of hexokinase. *Mech. Ageing Dev.* **31**, 197–212, DOI: 10.1016/S0047-6374(85)80030-0 (1985).

12. Aleshin, A. E. *et al.* Crystal structures of mutant monomeric hexokinase I reveal multiple adp binding sites and conformational changes relevant to allosteric regulation. *J. Mol. Biol.* **296**, 1001–1015, DOI: 10.1006/jmbi.1999.3494 (2000).

13. Kamata, K., Mitsuya, M., Nishimura, T., Eiki, J.-i. & Nagata, Y. Structural basis for allosteric regulation of the monomeric allosteric enzyme human glucokinase. *Structure* **12**, 429–438, DOI: 10.1016/j.str.2004.02.005 (2004).

14. Sui, D. & Wilson, J. E. Functional interactions between the noncovalently associated N- and C-terminal halves of mammalian type I hexokinase. *Arch. Biochem. Biophys.* **401**, 21–28, DOI: 10.1016/S0003-9861(02)00029-2 (2002).

15. Hashimoto, M. & Wilson, J. E. Kinetic and regulatory properties of HK I+, a modified form of the type I isozyme of mammalian hexokinase in which interactions between the N- and C-terminal halves have been disrupted. *Arch. Biochem. Biophys.* **399**, 109–115, DOI: 10.1006/abbi.2001.2744 (2002).

16. Tsai, H. J. Function of interdomain alpha-helix in human brain hexokinase: covalent linkage and catalytic regulation between N- and C-terminal halves. *J. Biomed. Sci.* **14**, 195–202, DOI: 10.1007/s11373-006-9123-5 (2007).

17. Gelb, B. D. *et al.* Targeting of hexokinase 1 to liver and hepatoma mitochondria. *Proc. Natl. Acad. Sci.* **89**, 202–206, DOI: 10.1073/pnas.89.1.202 (1992).

18. White, T. K. & Wilson, J. E. Rat brain hexokinase: Location of the allosteric regulatory site in a structural domain at the N-terminus of the enzyme. *Arch. Biochem. Biophys.* **259**, 402–411, DOI: 0.1016/0003-9861(87)90506-6 (1987).

19. Fang, T.-Y., Alechina, O., Aleshin, A. E., Fromm, H. J. & Honzatko, R. B. Identification of a phosphate regulatory site and a low affinity binding site for glucose 6-phosphate in the N-terminal half of human brain hexokinase. *J. Biol. Chem.* **273**, 19548–19553, DOI: 10.1074/jbc.273.31.19548 (1998).

20. Sebastian, S., Wilson, J. E., Mulichak, A. & Garavito, R. M. Allosteric regulation of type I hexokinase: A site-directed mutational study indicating location of the functional glucose 6-phosphate binding site in the N-terminal half of the enzyme. *Arch. Biochem. Biophys.* **362**, 203–210, DOI: 10.1006/abbi.1998.1043 (1999).

21. de Cerqueira César, M. & Wilson, J. E. Further studies on the coupling of mitochondrially bound hexokinase to intramitochondrially compartmented atp, generated by oxidative phosphorylation. *Arch. Biochem. Biophys.* **350**, 109–117, DOI: 10.1006/abbi.1997.0497 (1998).

22. Hashimoto, M. & Wilson, J. E. Membrane potential-dependent conformational changes in mitochondrially bound hexokinase of brain. *Arch. Biochem. Biophys.* **384**, 163–173, DOI: 10.1006/abbi.2000.2085 (2000).

23. de Cerqueira César, M. & Wilson, J. E. Functional characteristics of hexokinase bound to the type a and type b sites of bovine brain mitochondria. *Arch. Biochem. Biophys.* **397**, 106–112, DOI: 10.1006/abbi.2001.2639 (2002).

24. De Jesus, A. *et al.* Hexokinase 1 cellular localization regulates the metabolic fate of glucose. *Mol. Cell* **82**, 1261–1277.e9, DOI: 10.1016/j.molcel.2022.02.028 (2022).

25. Lemeshko, V. V. VDAC as a voltage-dependent mitochondrial gatekeeper under physiological conditions. *Biochim. Biophys. Acta, Biomembr.* **1865**, 184175, DOI: 10.1016/j.bbamem.2023.184175 (2023). Outer mitochondria potential (OMP) and VDACs interacting with cytosolic proteins.

26. Bieker, S. *et al.* Hexokinase-I directly binds to a charged membrane-buried glutamate of mitochondrial VDAC1 and VDAC2. *Commun. Biol.* **8**, 212, DOI: 10.1038/s42003-025-07551-9 (2025).

27. Rosa, J. C. & de Cerqueira César, M. Role of hexokinase and VDAC in neurological disorders. *Curr. Mol. Pharmacol.* **9**, 320–331, DOI: 10.2174/1874467209666160112123036 (2016). Role of Hexokinase and VDAC in neurological disorders.

28. Traxler, L. *et al.* Metabolism navigates neural cell fate in development, aging and neurodegeneration. *DMM* **14**, dmm048993, DOI: 10.1242/dmm.048993 (2021).

29. Pastore, A. & Shakhnovitch, E. From sequence to structure to mechanism to phenotype: The new frontiers of structural biology. *Curr. Opin. Struct. Biol.* **82**, 102674, DOI: 10.1016/j.sbi.2023.102674 (2023).

30. Botticelli, S. *et al.* Modelling protein plasticity: The example of frataxin and its variants. *Molecules* **27**, DOI: 10.3390/molecules27061955 (2022).

31. Nobili, G. *et al.* Probing protein stability: towards a computational atomistic, reliable, affordable, and improvable model. *Front. Mol. Biosci.* **10**, DOI: 10.3389/fmolb.2023.1122269 (2023).

32. Rinaldi, S., Colombo, G. & Morra, G. Exploring mutation-driven changes in the ATP-ADP conformational cycle of human hsp70 by all-atom MD adaptive sampling. *J. Phys. Chem. B* **128**, 7770–7780, DOI: 10.1021/acs.jpcb.4c03603 (2024).

33. Rosano, C. *et al.* Binding of non-catalytic ATP to human hexokinase I highlights the structural components for enzyme–membrane association control. *Structure* **7**, 1427–1437, DOI: 10.1016/S0969-2126(00)80032-5 (1999).

34. Aleshin, A. E., Zeng, C., Bartunik, H. D., Fromm, H. J. & Honzatko, R. B. Regulation of hexokinase I: crystal structure of recombinant human brain hexokinase complexed with glucose and phosphate. *J. Mol. Biol.* **282**, 345–357, DOI: https://doi.org/10.1006/jmbi.1998.2017 (1998).

35. Rinaldi, S., Colombo, G. & Paladino, A. Mechanistic model for the hsp90-driven opening of human argonaute. *J. Chem. Inform. Model.* **60**, 1469–1480, DOI: 10.1021/acs.jcim.0c00053 (2020).

36. Rinaldi, S., Colombo, G. & Paladino, A. The dynamics of t1 adenosine binding on human argonaute 2: Understanding recognition with conformational selection. *Protein Sci.* **31**, e4377, DOI: 10.1002/pro.4377 (2022).

37. Bobba, A. *et al.* Glucose-6-phosphate tips the balance in modulating apoptosis in cerebellar granule cells. *FEBS Lett.* **589**, 651–658, DOI: 10.1016/j.febslet.2015.01.031 (2015).

38. Altschul, S. F. *et al.* Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402, DOI: 10.1093/nar/25.17.3389 (1997).

39. Schäffer, A. A. *et al.* Improving the accuracy of psi-blast protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res.* **29**, 2994–3005, DOI: 10.1093/nar/29.14.2994 (2001).

40. Hošek, P., Toulcová, D., Bortolato, A. & Spiwok, V. Altruistic metadynamics: Multisystem biased simulation. *J. Phys. Chem. B* **120**, 2209–2215, DOI: 10.1021/acs.jpcb.6b00087 (2016).

41. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25, DOI: 10.1016/j.softx.2015.06.001 (2015).

42. Humphrey, W., Dalke, A. & Schulten, K. VMD visual molecular dynamics. *J. Molec. Graph.* **14**, 33–38, DOI: 10.1016/0263-7855(96)00018-5 (1996). Http://www.ks.uiuc.edu/Research/vmd.

43. Bonomi, M. *et al.* Plumed: A portable plugin for free-energy calculations with molecular dynamics. *J. Comput. Phys.* **180**, 1961–1972, DOI: doi:10.1016/j.cpc.2009.05.011 (2009).

44. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. Plumed 2: New feathers for an old bird. *Comp. Phys. Commun.* **185**, 604–613, DOI: 10.1016/j.cpc.2013.09.018 (2014).

45. Bonomi, M. *et al.* Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **16**, 670–673, DOI: 10.1038/s41592-019-0506-8 (2019).

46. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. Lincs: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472, DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H (1997).

47. Essmann, U. *et al.* A smooth particle mesh ewald method. *J. Chem. Phys.* **103**, 8577–8593, DOI: 10.1063/1.470117 (1995).

48. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81**, 511–519, DOI: 10.1063/1.447334 (1984).

49. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697, DOI: 10.1103/PhysRevA.31.1695 (1985).

50. Bernetti, M. & Bussi, G. Pressure control using stochastic cell rescaling. *J. Chem. Phys.* **153**, DOI: 10.1063/5.0020514 (2020).

51. Best, R. B. *et al.* Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi1 and chi2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273, DOI: 10.1021/ct300400x (2012).

52. Savelyev, A. & MacKerell, A. D. J. Competition among Li+, Na+, K+, and Rb+ monovalent ions for DNA in molecular dynamics simulations using the additive charmm36 and drude polarizable force fields. *J. Phys. Chem. B* **119**, 4428–4440, DOI: 10.1021/acs.jpcb.5b00683 (2015).

53. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. J. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935, DOI: 10.1063/1.445869 (1983).

54. Vermaas, J. V., Hardy, D. J., Stone, J. E., Tajkhorshid, E. & Kohlmeyer, A. Topogromacs: Automated topology conversion from charmm to gromacs within vmd. *J. Chem. Inform. Model.* **56**, 1112–1116, DOI: 10.1021/acs.jcim.6b00103 (2016).

55. Hošek, P., Kříž, P., Toulcová, D. & Spiwok, V. Multisystem altruistic metadynamics- well-tempered variant. *J. Chem. Phys.* **146**, 125103, DOI: 10.1063/1.4978939 (2017).

56. Simonson, T., Archontis, G. & Karplus, M. Free energy simulations come of age: Protein-ligand recognition. *Acc. Chem. Res.* **35**, 430–437, DOI: 10.1021/ar010030m (2002).

57. Case, D. *et al. AMBER 2020* (University of California at San Fransisco, San Francisco, USA, 2020).

58. Maier, J. A. *et al.* Ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *J. Chem. Theory Comput.* **11**, 3696–3713, DOI: 10.1021/acs.jctc.5b00255 (2015).

59. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Cryst.* **A32**, 922–923, DOI: 10.1107/S0567739476001873 (1976).

Acknowledgements

Numerical calculations have been made possible through the ISCRA NEDVIBA project and the CINECA-INFN agreement on Leonardo high-performance infrastructure (CINECA, Italy).

Funding

This research did not receive funding.

Author contributions statement

A.G. proposed the study and provided the literature framework; G.L.P. and S.R. designed the research work and made part of analysis; G.N. made the simulations and their analyses.

Data availability

All data supporting the findings of this study are available within the paper and its Supplementary Material. Raw data are available upon request to corresponding authors.

Additional information

The authors declare no competing interests.

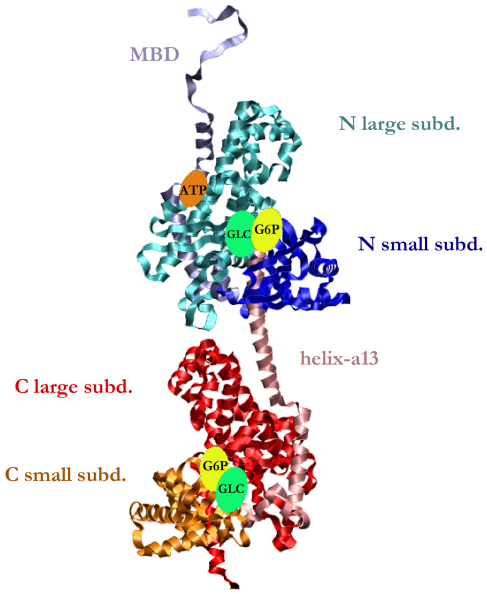


Figure 1. Schematic representation of two-domains hexokinase. GLC is glucose; G6P is glucose-6 phosphate; ATP displays its most likely N-terminal binding site[12,33]; N/C large/small subdomains are, respectively, the large and small sub-domains in the N-terminal and C-terminal domains; helix-a13 is the first part of the long α-helix connecting the two domains.

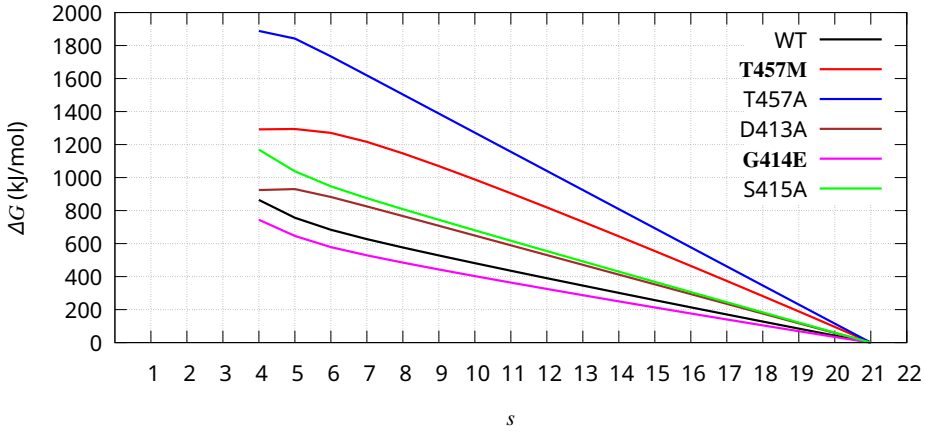


Figure 2. Free energy change (ΔG), with respect to the folded reference state ($s_0 = \langle \xi \rangle = 21$) as a function of the inter-domain α-helix folding, region 450-477 (helix-a13). Therefore, the variable s is the average number of α-helix hydrogen bonds in the helix-a13 segment. Zero is placed at the lowest G value for each sequence. Different sequences are displayed in different colors. NEDVIBA variants are boldface.

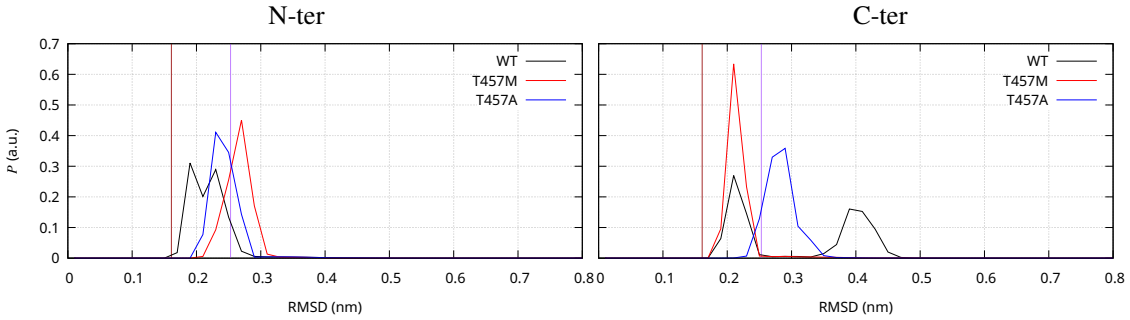


Figure 3. Distribution (P) of root-mean square deviation (RMSD) for WT (black), T457M (red), and T457A (blue). RMSD is computed on N- (left panel) and C-terminal (right panel) domains heavy backbone atoms. HK4 crystal structure in closed conformation (PDB 1V4S[13]) has been used as reference for RMSD calculations. Vertical lines indicate RMSD values computed on 1QHA C-terminal domain (brown[33]) and 1HQC C-terminal domain (purple[34]). Statistics are consistent with $\langle \xi \rangle = s = 21$.

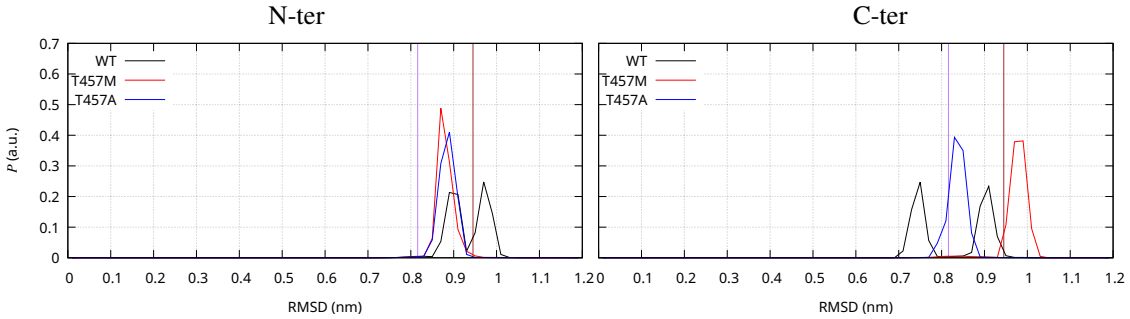


Figure 4. Same as for Fig. 3, but using as reference for RMSD the backbone heavy atoms of HK4 crystal in the open conformation (PDB 1V4T[13]).

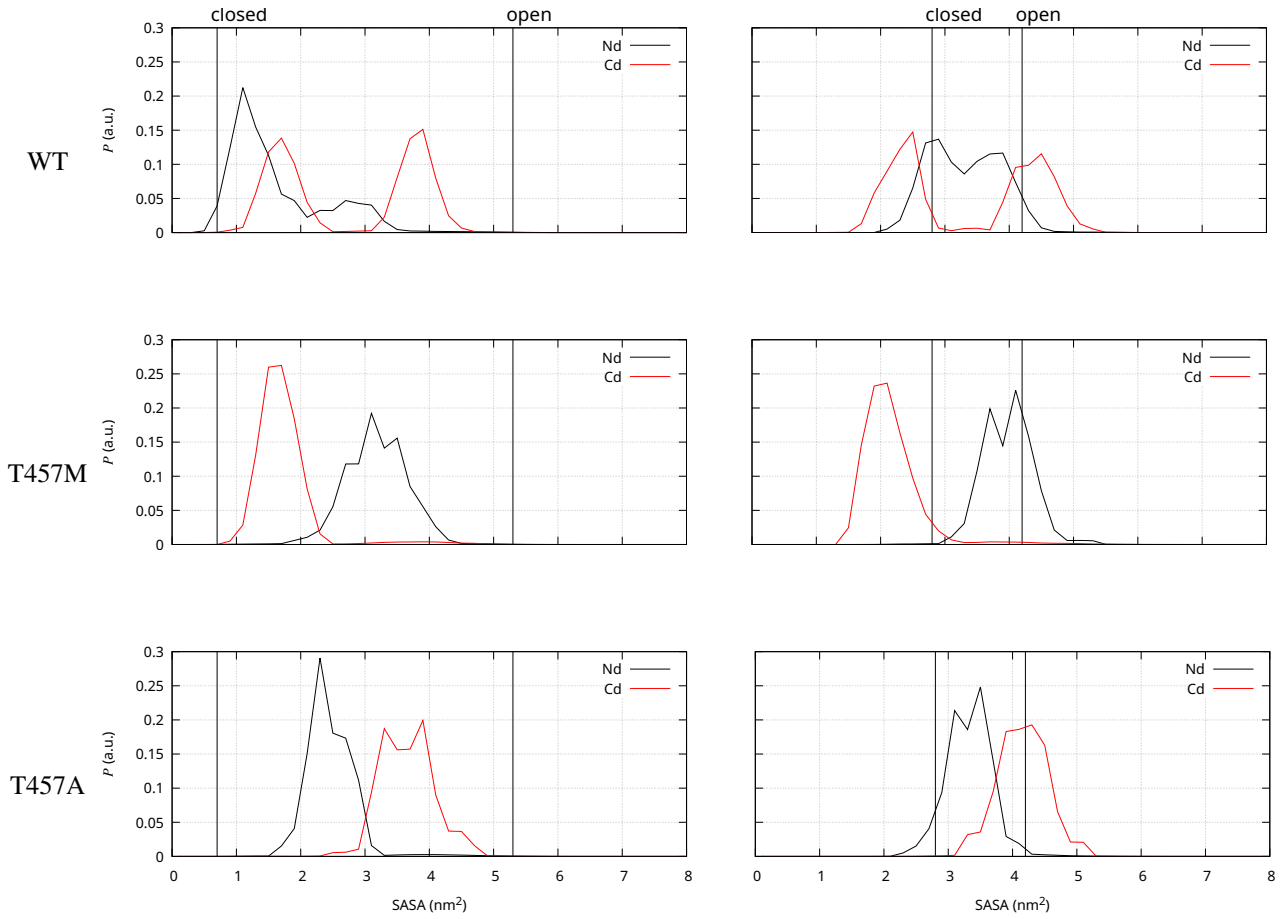


Figure 5. Probability (P) of SASA values for the N-terminal (black, Nd) and C-terminal (red, Cd) glucose (left panels) and glucose 6-phosphate (right) site cavities. Statistics are consistent with $\langle \xi \rangle = s = 21$ (all curves). Vertical lines indicate the values for HK4, the lower value for the closed state (1V4S), the higher values for the open state (1V4T).

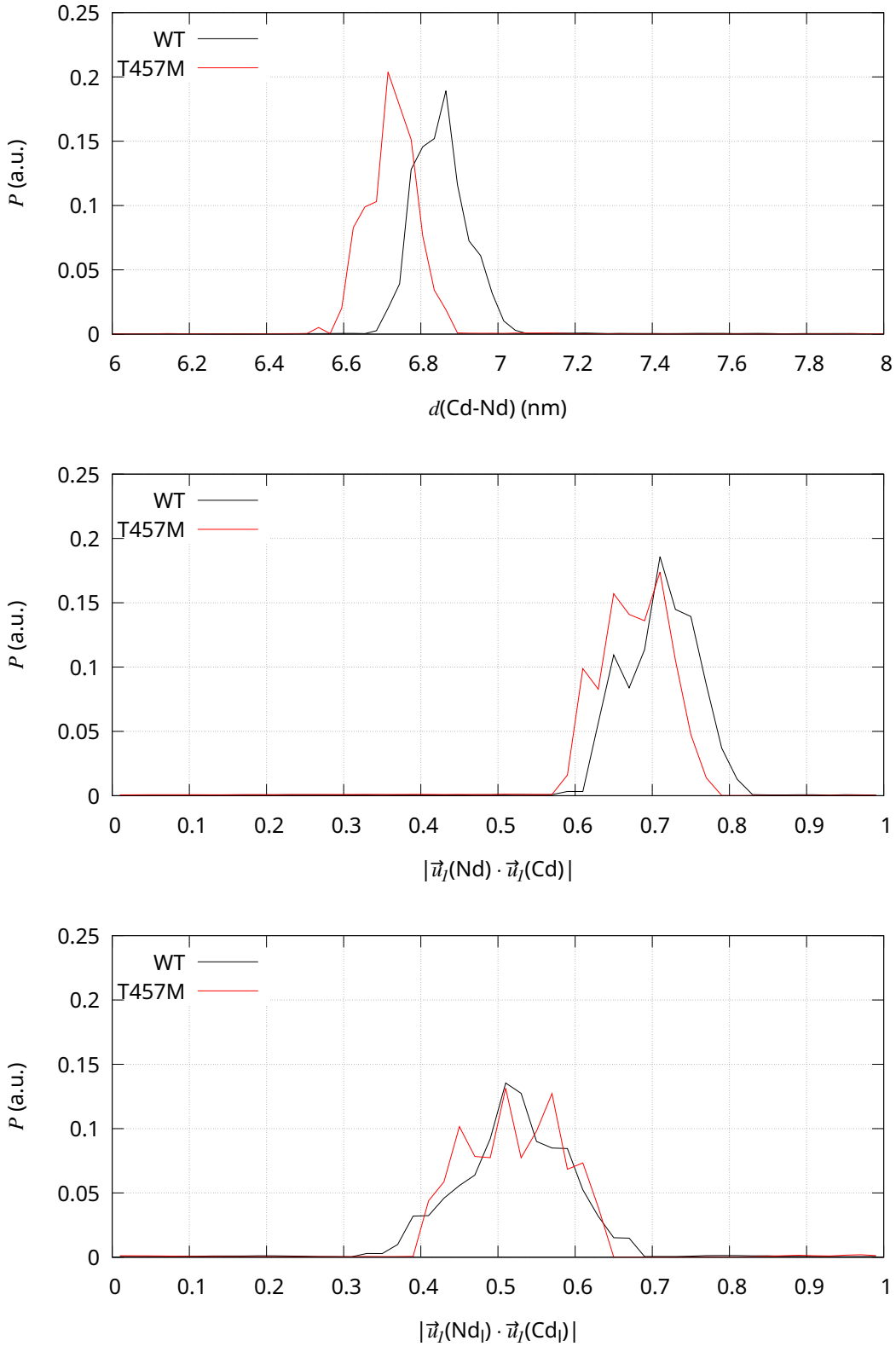


Figure 6. Distribution (P) of: the distance between centers of N- and C-terminal domains (d, top); absolute value of scalar product between $\vec{u}_1$ axes of N-terminal domain (Nd) and C-terminal (Cd) domain (middle); same of middle panel but with $\vec{u}_1$ axes of N-terminal large sub-domain and C-terminal large sub-domain (bottom). The statistics is consistent with s =21. WT (black) and T457M (red) variant in Fig. 5 are compared.

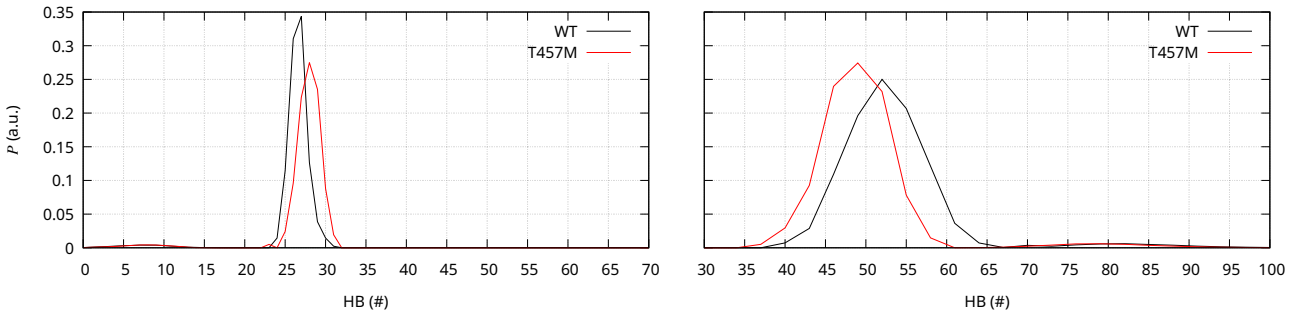


Figure 7. Distribution (P) of hydrogen-bonds (HBs) between helix-a13 (residues 450-477) and the rest of the protein (left) and the water solvent bath (right). The statistics is consistent with s =21. The variant T457M (red curves) is compared to WT (black curves).

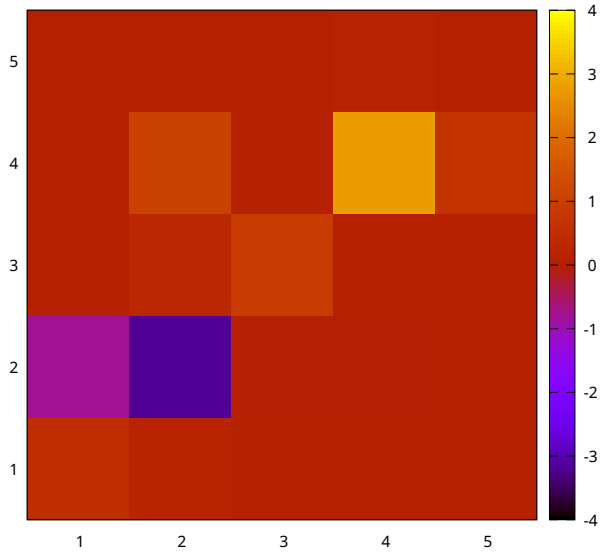


Figure 8. Difference between average number of salt-bridges in T457M and WT. 5 HK1 regions are measured: residues 1-21 (1, N-terminal chain); residues 22-449 (2, N-domain); residues 450-477 (3, helix a13); residues 478-910 (4, C-domain); residues 911-917 (5, C-terminal chain). On the x-axis are the negatively charged side chains (Cγ in Asp, Cδ in Glu and C-terminal C). On the y-axis are the positively charged side chains (Nη in Arg, Nζ in Lys and N-terminal N). The cut-off distance for a salt-bridge is 0.4 nm. The statistics is the one consistent with s =21.

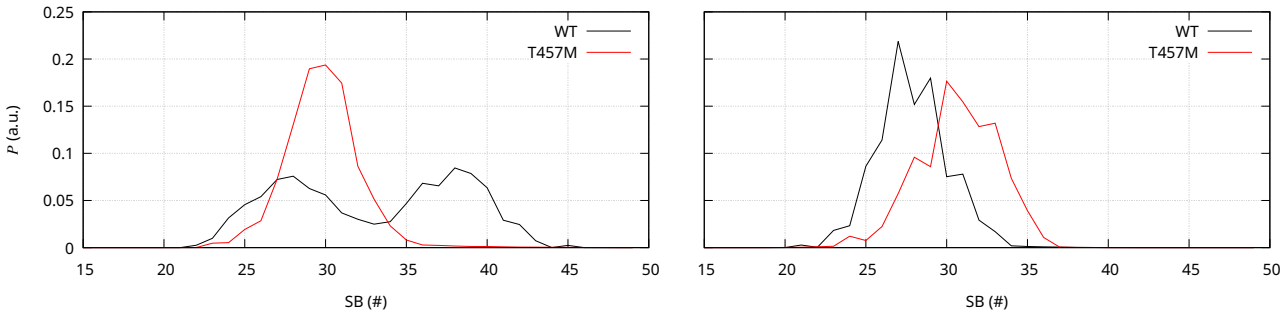


Figure 9. Probability (P) of average number of salt-bridges (SB): left is for SBs within Nd (2-2 in Fig. 8); right is for SBs within Cd (4-4 in Fig. 8). The statistics is consistent with s =21. WT is in black, T457M is in red.

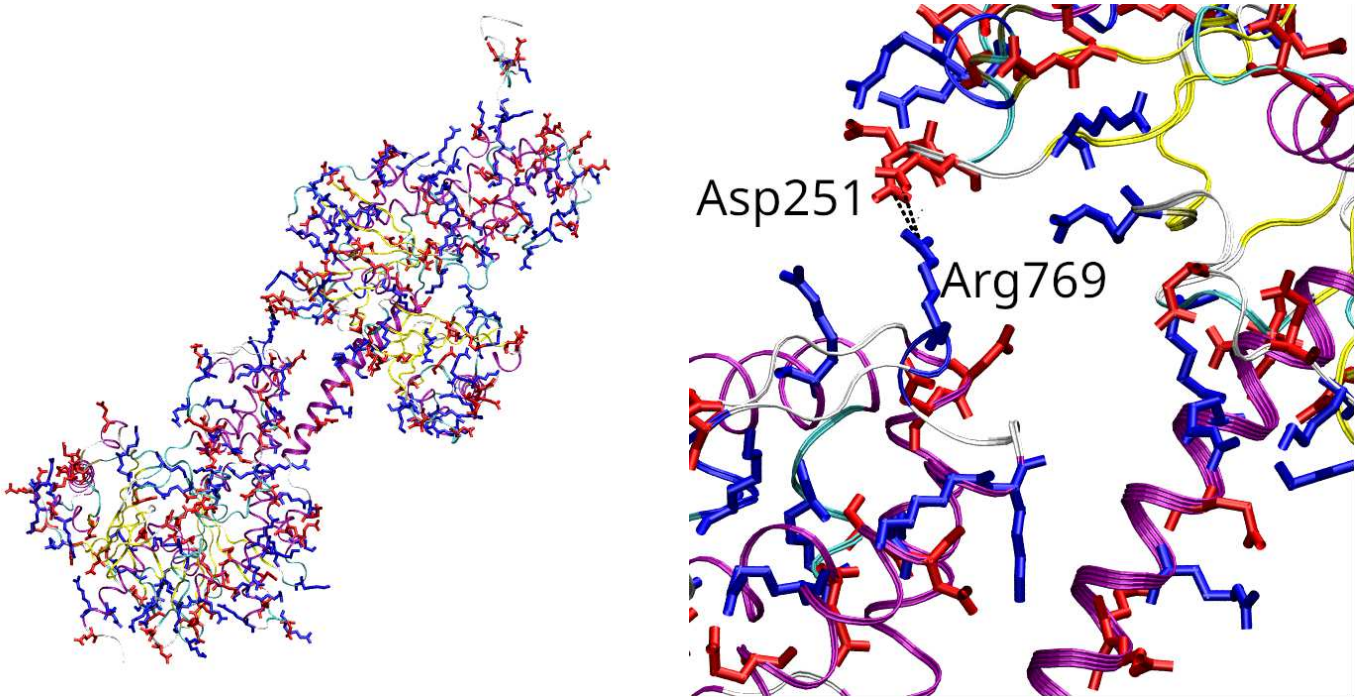


Figure 10. Configuration of T457M variant displaying the 2-4 inter-domain SBs. Arg and Lys residues are in blue; Asp and Glu residues are in red. N-terminal domain (Nd) is top-righ, C-terminal domain (Cd) is bottom-left. Right is the in-set of the Nd-Cd interface, with helix-a13 on the righthand side.

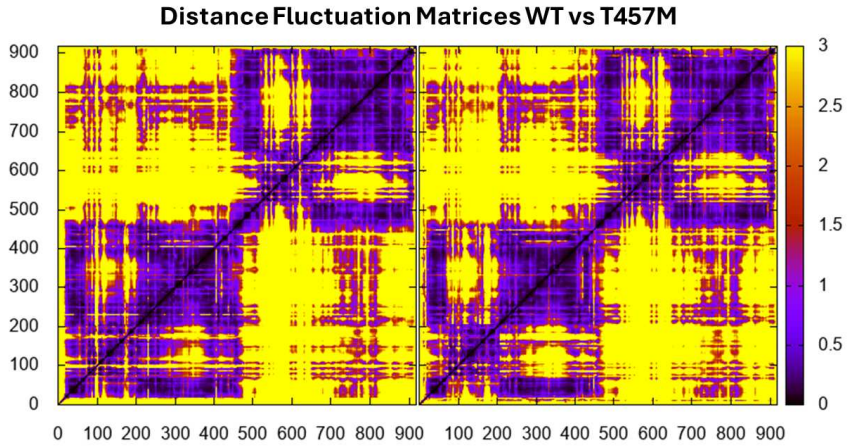


Figure 11. Distance Fluctuation (DF) matrices of the wild-type (WT, left) and the T457M mutant (right). Each element (i, j) encodes the DF value (Å^2, z-axis, color palette) between residues i (x-axis) and j (y-axis).

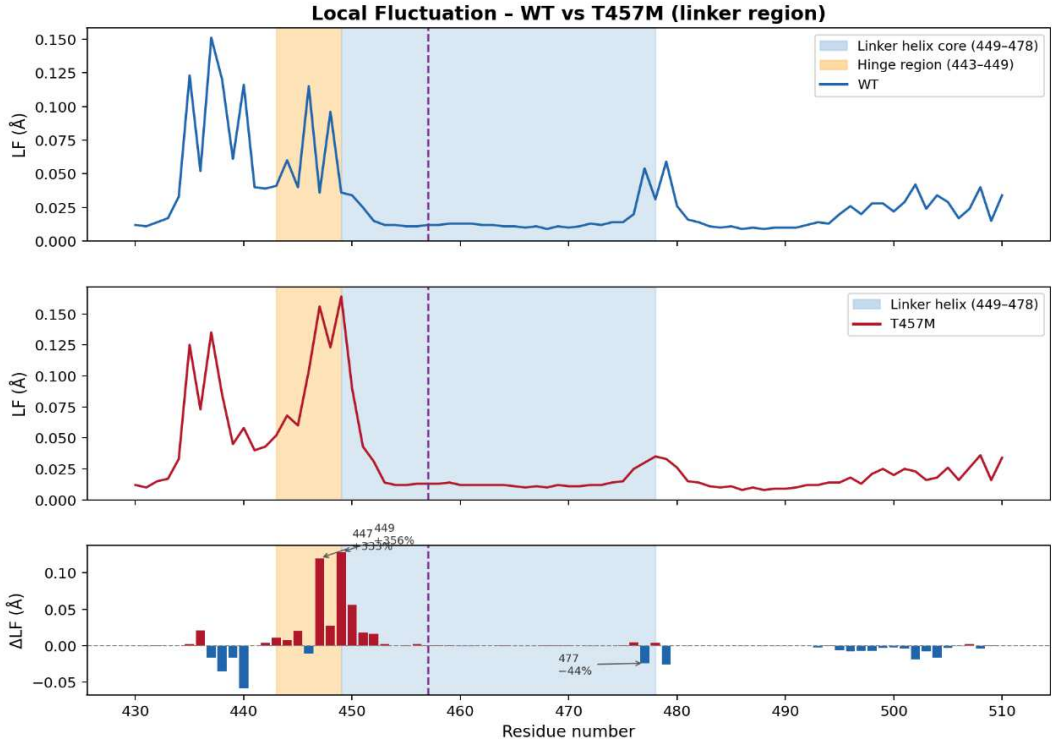


Figure 12. Local fluctuation profiles at the linker hinge and helix (residues 430–510) for WT (blue) and T457M (red), with relative difference ΔLF = LF(T457M)-LF(WT) shown as a bar chart (bottom panel). The light-blue shaded region marks the linker α-helix (residues 449–478); the light-orange region marks the hinge (residues 443–449); the dashed purple line indicates the T457M substitution site (residue 457). Throughout the helix core (residues 453–476), ΔLF is indistinguishable from zero ($\Delta LF \leq 0.003$ Å), confirming that the helix itself is structurally unaffected by the mutation. In contrast, hinge residues 447 and 449 show ΔLF of +333% and +356%, respectively, indicating that T457M converts the N-terminal attachment point of the helix into a loose joint. Compensatory stiffening is observed at the C-terminal end of the helix (residue 477, ΔLF = -44%).

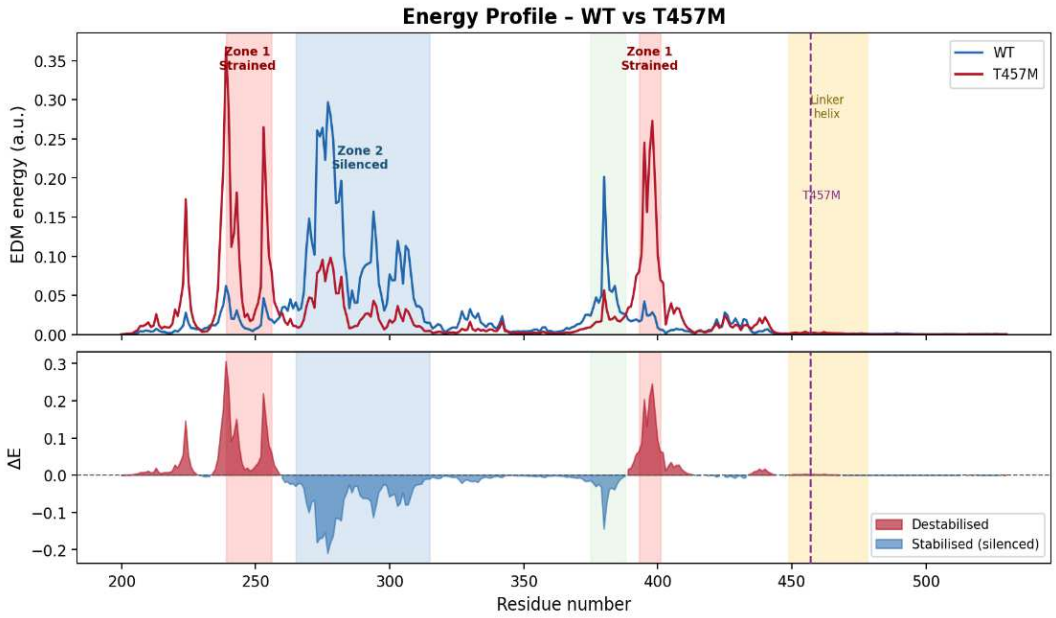


Figure 13. EDM per-residue interaction energy profile for WT (blue) and T457M (red) across N-domain and linker residues 200-520 (top panel). Energy difference ΔE = T457M-WT (kcal/mol) (bottom). Red shading indicates residues with increased interaction energy in T457M, whereas blue shading indicates residues with reduced interaction energy. Two clusters with increased interaction energy are observed in β-strands 239–256 and α-helices 393–401, located 17-31 Å from the T457M site. One cluster with reduced interaction energy spans helices 265-315 of the glucose-binding lobe, where EDM energy decreases by up to 84% relative to WT. The gold band marks the linker helix (449-478); the dashed purple line marks residue 457.

Domain	residues							
mitochondria binding domain (MBD)	1-21							
N-terminal domain (Nd)	21-449							
Nd small sub-domain	73-207							
Nd large sub-domain	208-447							
C-terminal domain (Cd)	478-910							
Cd small sub-domain	521-655							
Cd large sub-domain	656-895							
helix-a13	450-477							
HK4 GLC pocket	168	169	204	205	231	256	290	
Nd GLC pocket	172	173	208	209	235	260	294	
Cd GLC pocket	620	621	656	657	683	708	742	
HK4 G6P pocket	78	82	151	205	228	409	411	445
Nd G6P pocket	84	88	155	209	232	413	415	449
Cd G6P pocket	532	536	603	657	680	861	863	897
Nd ATP pocket	30-31	345	373	380, 384	425-426, 429			
HK4 single domain	15-93		98-156		180-458			
N-terminal domain, aligned to HK4	21-99		102-160		184-462			
C-terminal domain, aligned to HK4	469-547		550-608		632-910			

Table 1. Definition of HK1 regions and C-/N-terminal domains, also when aligned to HK4 (single domain) sequence.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GallietalSM.pdf