Conformational Dynamics of the Free MAX Protein Revealed by Well-Tempered Metadynamics

Huixia Lu,[*,†] Jordi Marti,[*,‡] and Jordi Faraudo[†]

†*Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus de la UAB, E-08193 Bellaterra, Spain*

‡*Department of Physics, Universitat Politècnica de Catalunya-Barcelona Tech (UPC), Barcelona, Spain*

E-mail: huixialu@icmab.es; jordi.marti@upc.edu

Abstract

The MAX protein is a key transcriptional regulator that partners with the c-MYC oncoprotein and other members of the MYC/MAX/MAD network to control genes involved in fundamental cellular processes such as growth, differentiation, metabolism, and apoptosis. The transcriptional and tumorigenic activities of MYC mainly depend on its association with MAX. In its unbound state, MAX exhibits partially disordered regions, complicating structural characterization and limiting conventional drug discovery approaches that rely on the existence of well-defined binding pockets and ignore dynamical aspects of protein structures.

To address these challenges, we employed the enhanced-sampling well-tempered metadynamics method to systematically explore the conformational space of the unbound MAX protein. Our results revealed, for the first time, a well-defined and thermodynamically favorable conformation of monomeric MAX. This is a remarkable finding,

as it demonstrates that regions of MAX previously considered persistently disordered are capable of adopting stable, folded structures under specific conditions. Moreover, our findings also suggest that the meta-stable structural motifs observed in this work may harbor druggable sites, particularly relevant for strategies aiming to target MAX directly or to disrupt its interaction with MYC, thereby modulating oncogenic signaling pathways. The present study establishes a new structural framework for understanding the dynamics of MAX and provides a foundation for future structure-based drug design targeting the MYC/MAX axis. Finally, our work offers a strategic blueprint for investigating similarly challenging drug targets.

Introduction

MAX (MYC-Associated factor X) encodes a ubiquitously expressed transcription factor that serves as an essential dimerization partner for MYC family proteins, facilitating their ability to regulate gene expression.[1,2] MAX plays a pivotal role in modulating cellular processes such as proliferation, differentiation and apoptosis by forming heterodimers with MYC or other members of the MYC-MAX-MXD network.[3–5] These interactions enable the binding of the resulting complexes to E-box sequences in DNA, where they activate or repress target genes depending on the dimeric partner involved.[6,7] Similarly to MYC, MAX possesses a basic helix-loop-helix leucine zipper (bHLHLZ) domain at its C-terminus, which is crucial for both dimerization and DNA binding.[8,9] However, in contrast to MYC, MAX lacks an intrinsic transactivation domain, acting instead as a scaffold that confers transcriptional activity or repression through its binding partners.[10,11] The evolutionary conservation of MAX underscores its central role in balancing oncogenic and tumor-suppressive signals within the cell.[12,13]

In many cancers, dysregulation of the MYC/MAX/MXD transcriptional network leads to an imbalance favoring MYC-MAX heterodimers, which drive uncontrolled cell proliferation and repress differentiation. Although MAX itself is not typically overexpressed or mutated,

its obligatory role as a dimerization partner makes it a critical enabler of MYC-driven oncogenesis.[1,2] MAX is expressed ubiquitously and forms stable bHLHLZ-mediated dimers with MYC, enabling binding to E-box DNA sequences and transcriptional activation of target genes involved in cell growth, metabolism and survival.[4,7]

Given the structural challenges in directly targeting MYC-an intrinsically disordered protein lacking well-defined binding pockets-MAX has emerged as an attractive alternative target for therapeutic intervention.[8,11] Disrupting the MYC-MAX interaction could effectively suppress MYC oncogenic activity without requiring direct inhibition of MYC itself. Several strategies have been proposed to this end, including the use of dominant-negative MYC mutants such as Omomyc[14] (OMO-103, currently in clinical trials), which sequesters MAX and blocks MYC-dependent transcription.[15] Several small molecules, such as 10058-F4,[16] KJ-Pyr-9,[17] and MYCMI-6,[18] have been developed to interfere with the MYC–MAX interaction directly. Also, peptides that mimic MAX or stabilize alternative dimerization partners (such as MXD proteins) could shift the transcriptional balance toward growth suppression.[3,5]

In order to identify viable drug-binding regions on MAX, it is essential to characterize its structural configurations. In this study, we leverage all-atom molecular dynamics (MD) simulations with high temporal and spatial resolution to model the conformational behavior of MAX and its potential interactions with MYC. We have chosen MD simulations due to its suitability to explore and characterize the conformational space of highly dynamic proteins.[19,20] By analyzing the dynamic structural ensemble of MAX, we aim to uncover transient or stable pockets that could serve as potential drug-binding sites. Subsequently, we used the Well-Tempered Metadynamics (WT-MetaD) simulations to map the free-energy landscape of MAX in solution, which will allow us to identify low-energy conformations with sufficient structural rigidity to accommodate small molecules, paving the way for the development of novel inhibitors targeting the MYC-MAX axis in cancer therapy.

Methods

System preparation and MD simulations

We prepared a simulation system consisting of the MAX protein (residues 202–284, UniProt ID: P61244) as found in PDB entry 1NKP, solvated in an explicit water box containing 0.15 M KCl to mimic physiological ionic strength. The total number of particles was of 159,804 atoms. The system was generated using the CHARMM-GUI solution builder,[21–23] and the CHARMM36m force field[24] was adopted. All-atom MD simulations were conducted using AMBER24[25,26] software. The system was minimized for 10000 steps followed by three 125 ps equilibrium runs while gradually reducing the harmonic constraints (NVT ensemble). Production runs were performed with the NPT ensemble, whereas Langevin dynamics with the friction coefficient 1.0 ps^{-1} was used for temperature regulation (310 K). The Monte Carlo barostat was used for the pressure regulation (1 bar). All bonds involving hydrogens were set to fixed lengths, allowing fluctuations of bond distances and angles for the remaining atoms. The time step was set to be 2 fs, and the frames were saved every 100 ps for analysis. The particle mesh Ewald method was automatically used by AMBER24 to calculate the electrostatic interactions, and the van der Waals interactions were calculated using a cutoff of 1.2 nm. Periodic boundary conditions were taken in three directions of space.

Well-tempered metadynamics

The enhanced sampling method WT-MetaD applies a time-dependent biasing potential along a set of collective variables (CV) by adding a Gaussian additional bias potential to the total potential in order to overcome barriers larger than k_BT, with k_B being Boltzmann's constant and T the temperature. In the method, sampling is performed on a selected number of degrees of freedom (s_i), i.e. a chosen set of CV. For each degree of freedom, the biased

potential $V(s_i, t)$ is a dynamical function constructed as the sum of Gaussian functions:

$$V(s_i, t) = \sum_{k\tau<t} W(k\tau) \exp\{-\sum_{i=1}^{d} \frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2}\}, \quad (1)$$

where k is an integer, τ is the Gaussian deposition stride, $W(k\tau)$ is the height of the Gaussian and σ_i is the width of the Gaussian for the i-th CV. Full details are provided in Refs.[27–29] In WT-MetaD, the biased potential can force the system to move around all possible states inside a particular range of the subspace of the selected CV. In the present work, we have taken a specific approach similar to previous works[30,31] where the CVs were defined from distances between groups of particular amino acids. After analyzing the protein's dynamics from the classical MD runs, we have selected two collective variables: *d1* and *d2*, where *d1* refers to the distance of center of mass of group1 (backbone atoms of residues 202-204) and group2 (backbone atoms of residues 282-284) and *d2* refers to the distance between C_α atoms of Ile218 and Ile250 to explore the conformational space of MAX, see Results section for a visualization of these CVs over the protein structures. The WT-MetaD simulations were performed using the joint gromacs/2023.5-plumed/2.9.2 tool,[32–34] starting from the relaxed crystal structure (CS). The V-rescale thermostat was considered for accurate temperature control and proper thermodynamic sampling at 310.15 K. Pressure was controlled to be 1 bar by a Parrinello-Rahman piston with a damping coefficient of 5 ps^{-1}. Periodic boundary conditions in the three directions of space were also considered. More details can be found in the Supporting Information (SI).

Data analysis and visualization

The backbone atoms (N, C_α, C, and O) of MAX were used to calculate the root mean squared deviation (RMSD) profile through RMSD Trajectory Tool of VMD-1.9.3.[35] To explore the conformational landscape of the MAX protein during the molecular dynamics simulations, we carried out cluster analysis using the atomic coordinates of the backbone atoms, applying

Principal Component Analysis (PCA)s.[36] The analysis was performed using the MDAnalysis Python library.[37] Pairwise distance matrices were constructed by extracting the positions of the backbone atoms throughout the entire trajectory. These structural descriptors were then projected onto two principal components (PCA component 1 and component 2) to reduce dimensionality. Subsequently, the KMeans clustering algorithm[38] was applied to partition the conformations into five distinct clusters. Each cluster corresponds to a group of structurally similar protein conformations, potentially representing distinct metastable states sampled during the simulation. Then the representative structures from each cluster which are closest to the cluster centroid were identified for further structural characterization and visualization. This clustering approach provides a simplified yet informative view of the protein's conformational ensemble, facilitating the interpretation of structural transitions and dynamic behavior. For each representative cluster structure in Fig. 1D, the angle between two helices was calculated by defining vectors between the alpha carbon atoms (Residues 210 and 221 of Helix 1 and Residues 249 and 267 of Helix 2). The resulting angle was converted from radians to degrees to quantify the helix's relative orientation.

The R-package metadynminer[39,40] was employed to analyze the WT-MetaD outputs. This package provides a set of tools for the efficient extraction, visualization, and interpretation of free-energy surfaces (FES) computed from WT-MetaD trajectories. In this study, metadynminer was specifically employed to reconstruct the FES, allowing for the identification and characterization of local free-energy minima corresponding to metastable conformational states. Additionally, it was used to compute minimum free-energy paths (MFEPs), which describe the most probable transition routes between these states on the FES. This sort of analysis enabled a deeper understanding of the thermodynamic landscape and the conformational dynamics sampled during the enhanced sampling simulations. The analysis tool VMD[35] and Chimera[41] were used for visualization and analysis purposes.

Results and discussion

Structural transition pattern of MAX

First, we performed unbiased MD simulations for 2 μs using the crystal structure of MAX protein as initial condition (see Methods). In order to identify the essential features of the protein configurations explored during the MD trajectory, we have performed Principal Component Analysis (PCA). We considered PCA-based clustering with both 5 and 8 clusters to capture the structural variability of MAX. The results from the 8-cluster analysis were largely consistent with those obtained using 5 clusters, with the additional clusters further subdividing the existing groups without revealing fundamentally new structures. This consistency indicates that clustering into 5 groups captures the essential data structure and is sufficient for interpretation (Fig. 1). The detailed results of the 8-cluster analysis are provided in Figure S1 of the SI for reference.

Fig. 1 shows that cluster2 exhibits the highest population density, followed by cluster0, cluster3, cluster1, and cluster4 (Fig. 1A-B). The transitions between these clusters were identified by tracking the simulation frames corresponding to each representative structure along the unbiased simulation timeline. We found a clear transition pathway during unbiased simulations: the unbound MAX evolves from the initial crystal-like structure (represented by cluster3) towards cluster0 adopting during the evolution the conformations corresponding to the other clusters (see Fig. 1C). This suggests that cluster0 is thermodynamically more stable than the other clusters. Representative structures of the five distinct clusters identified by Principal component analysis and the detailed transition pathway are illustrated in Fig. 1D.

Selection of Collective Variables in WT-MetaD Simulations

After performing unbiased MD simulations, we will perform a more detailed analysis of the possible conformation space of MAX by employing Well-Tempered Metadynamics simulations, as described in the Methods section. Before performing the simulations, the WT-

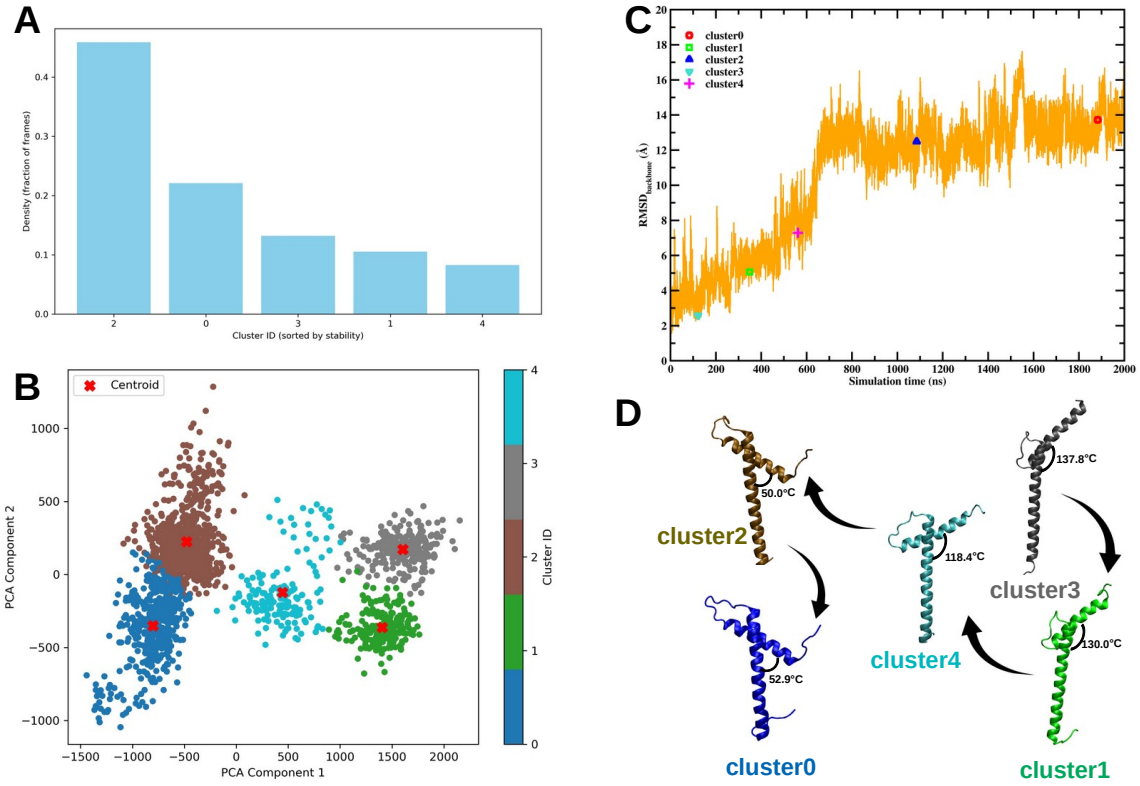


Figure 1: A. Cluster density profile derived from PCA clustering, performed on the entire MD trajectories. B: Five distinct clusters identified from PCA results, with the representative structure of each cluster indicated by a red cross. C. RMSD profile over the full simulation time, with the frames corresponding to the representative structures specially highlighted. D. The evolution from cluster3 to cluster0 of MAX is illustrated here, with the distinct structural transition highlighted by changing the angle between the two helices.

MetaD technique requires the selection of the appropriate collective variables (CVs) employed in the exploration of the conformation space. We will consider as CVs distances between particular amino acids, as in previous works.[30,31] The selection of these distances to be used as CVs is made by analyzing the transitions between the five different clusters obtained in the unbiased MD simulations (Fig. 1D). The geometrical changes undergone by the protein during these transitions can be clearly seen by aligning the representative structures of these clusters, Fig. 2A. The conformational transitions are characterized by the rotation of helix 1 (residues 202–228) towards helix 2 (residues 239–281). This rotation ultimately leads to

a significantly reduced angle between Helix 1 and Helix 2, as illustrated in Fig. 1D. Meanwhile, the hydrophobic interactions between the two helices are strengthened, along with a marked decrease in the distance between Ile218 and Ile250 (indicated in Figure 2B as *d2*). As a result of this conformational rearrangement, the overall distance between the N- and C-termini of the 202–284 residue segment in unbound MAX is substantially reduced, leading to the formation of a more thermodynamically stable structure, represented by cluster0. Furthermore, as shown in Fig. 2C, a detailed comparison of the aligned structures from cluster0 and cluster2 indicates a tendency for C-terminal deformation in MAX, suggesting that unbound MAX may adopt increasingly disordered conformations with a reduced distance between the N- and C-termini (indicated in Figure 2B as *d1*).

These observations were derived from the initial set of unbiased MD simulations conducted in the first phase of this study. Based on these structural changes, we selected *d1* and *d2* as the collective variables for subsequent enhanced sampling via WT-MetaD simulations.

Free-energy landscape of MAX

After selecting *d1* and *d2* as CVs, we performed WT-MetaD simulations. The results are summarized in Fig. 3. We ensured that the free energy calculation is properly converged, as shown in the Supporting Information (see Figures S2-S5). As shown in Fig. 3, the free-energy landscape of MAX reveals pronounced structural diversity within the low free-energy region, populated by several metastable states. This indicates a rich ensemble of conformational states readily accessible in solution. The different states found during the previous unbiased MD simulations and the transitions observed between them are also indicated in Fig. 3. The results demonstrate that over the 2 μs of unbiased MD simulation, the system evolves from a high-energy, crystal-like conformation of MAX (cluster3) towards a significantly lower-energy structure (cluster0). The free-energy difference between the most stable state found in unbiased MD (cluster0) and the crystal-like state (c3) is $\sim$ 60 kJ/mol. Also, this figure shows

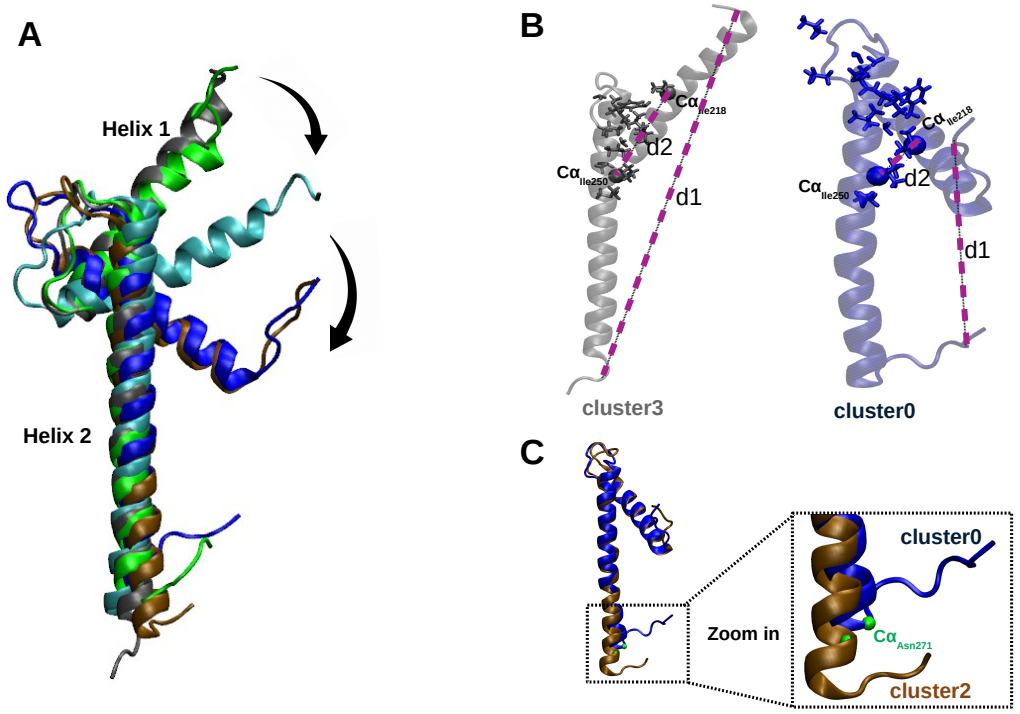


Figure 2: A. Representative structures from the five clusters identified by PCA analysis of MD simulations aligned based on residues 222-278 of Helix 2 of MAX, clearly illustrating the progression of conformational transitions. Colors correspond to those shown in Fig. 1D. B. Illustration of two collective variables (*d1* and *d2*), C_α atoms of Ile218 and Ile250 are shown in Van der Waals (VdW) representation. Hydrophobic side chains of residues 217-250 are shown in Licorice representation. C: Comparison of cluster0 (blue) and cluster2 (brown), aligned using backbone atoms of Helix 2. The position of Asn271 is highlighted in green using the VdW representation.

that the unbiased MD simulations were not reaching the lowest free energy conformation of the protein, in spite of the long simulation time. The results also highlight the fact that, even for a relatively small motif of a protein in this study (83 amino acids), the crystal structure could be very far from the configuration with minimum free energy. The result highlights the need for use of enhanced sampling methods (such as WT-MetaD) to investigate the structural variability of proteins. This enhanced sampling strategy enables us to explore the broader conformational landscape of MAX and ultimately identify previously inaccessible, stable drug-binding pockets in underexplored regions of the protein.

Looking in more detail at the free-energy landscape, it is possible to observe several

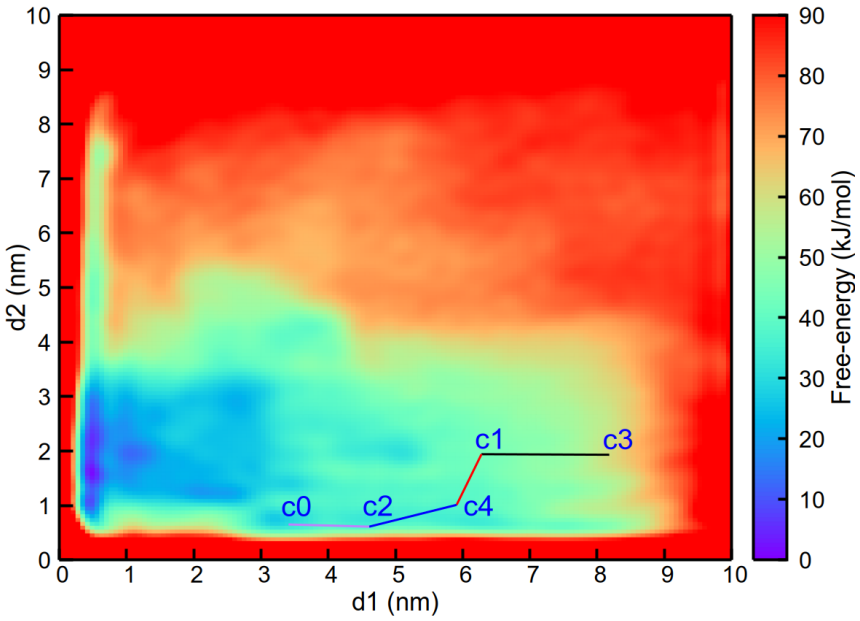


Figure 3: The two-dimensional free-energy surface F(*d1*, *d2*) (kJ/mol) of unbound MAX in solution obtained from WT-MetaD simulations. Regions with free energy exceeding 90 kJ/mol are shown in red. "c0" to "c4" refer to the five representative structures (cluster0 to cluster4) identified by PCA analysis of unbiased MD simulations (see Fig. 1).

primary basins (Fig. 4). The details of these molecular structures (labeled A through L in Fig. 4) are also provided in Table 1.

Table 1: Coordinates of the representative basins (minima) corresponding to the Fig. 4 with the helical domain in each structure highlighted here. "CS" denotes the crystal structure in PDB 1NKP.

Basin	*d1* (nm)	*d2* (nm)	Free-energy (kJ/mol)	Helical domain
A	0.45	1.61	0.00	E248-R255 and I264-Q276
B	0.45	2.18	4.77	E248-R255 and I264-A280
C	0.45	1.05	6.68	-
D	1.09	1.97	11.37	A246-K256 and I264-R279
E	2.08	1.26	17.30	I242-H260 and I264-R279
F	2.63	2.87	20.77	K245-K256 and I264-A280
G	3.21	0.77	25.69	K203-L225, A240-H260 and I264-L281
H	5.05	1.96	30.55	K245-R255 and I264-L281
I	4.62	0.69	32.23	-
J	6.06	0.65	33.23	T247-R255 and I264-G282
K	3.69	4.22	38.40	A246-K256 and I264-R279
L	0.52	4.72	41.21	K236-R239, L243-H260 and I264-R279
CS	8.98	1.96	74.91	K203-S228 and R239-R279

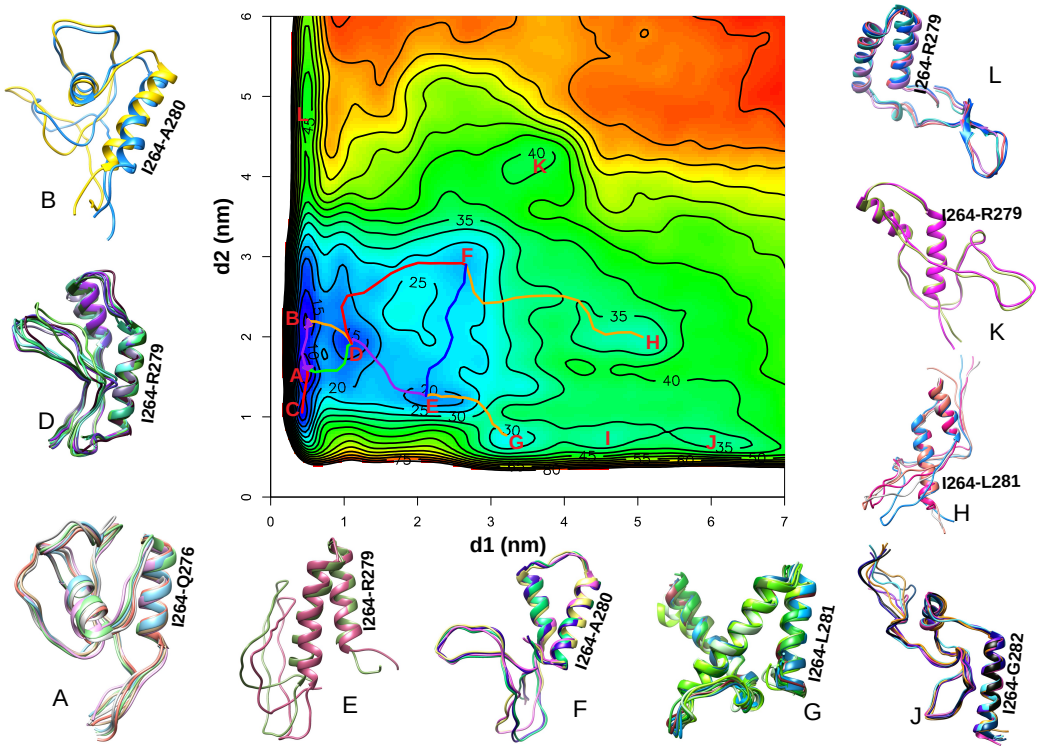


Figure 4: The partial free-energy surface F(*d1*, *d2*) (kJ/mol) is presented here, highlighting newly discovered stable states within basins "A"–"L". The corresponding minimum free-energy paths (MFEPs) connecting selected basins are also depicted. Representative MAX conformers from each basin are shown, highlighting the helical domain (I264-A280) near the C-terminus in all cases. Structures of MAX for basin C and I are not shown due to their short lifetime.

Basins A, B, and C, located in the lower-left region of the Free-energy surface (d1 $\sim$ 0.45 nm, d2 $\leq$ 2.2 nm), represent the most thermodynamically stable conformations, as indicated by their deep free energy wells, see Table 1. These basins are tightly clustered, suggesting a compact ensemble of low-energy structures, associated with the native or folded state of MAX (see the corresponding structures of different basins in Fig. 4). Basins D, E, F, and G occupy moderately low-energy regions (11–26 kJ/mol) and may represent alternative metastable states that are structurally related to the native ensemble but differ in loop orientations or side-chain packing. In contrast, basins H-L, that are located in higher-energy regions (between 31–42 kJ/mol), refer to less stable and partially unfolded conformations. This free-energy landscape reveals broader and more diffuse features in this region with

greater structural flexibility. We observed a substantial free-energy difference of 74.91 kJ/mol between the most stable basin (A) and the experimentally resolved crystal structure (CS). This substantial difference between the most stable configuration of MAX and its crystal structure has important implications for the development of drugs targeting MAX. This significant free energy gap underscores the difficulty of targeting MAX's CS conformation with small molecules due to its low state population. The stable conformations identified in this study may offer promising targets for the development of effective inhibitors against MAX-associated diseases, much better than the crystal structure of MAX.

Identification of transition paths and transition states

The WT-MetaD technique reveals not only the free energy basins (the most stable configurations) but also the free-energy barriers that the protein needs to overcome to shift between stable states. Furthermore, the maxima along each path give indications of the approximated location of transition states (TS) between stable states. The free energy barriers between different basins of interest in Fig. 4 and their transition half-lives are shown in Table 2. The corresponding free-energy profiles along the MFEP for each case are shown in Figure S6 of SI, including the estimated locations of local TS between stable basins, and the raw data can be found in our Github repository.[42] The energy barriers between low-energy basins and higher-energy regions (e.g., A to D, E to G, and F to H) suggest that large-scale rearrangements are thermodynamically unfavorable and likely require crossing high activation barriers. The half-life values along the MFEP between basins are consistent with the number of conformers sampled for each basin in the WT-MetaD trajectories. A shorter half-life reflects a faster kinetic transition, which in turn results in fewer structurally relevant conformers. Notably, the structures of MAX corresponding to basin C and I, which exhibit the very short transition half-lives (0.32 ps and 0.27 ps, respectively), were not captured in the WT-MetaD output, where frames were saved every 2 ps, indicating that they are transient states.

Table 2: Free-energy barriers (ΔF) between basins (kJ/mol) along the corresponding MFEPs and their transition half-lives (ps), predicted by the function *neb* of metadynminer using Nudged Elastic Band method.[43] State transitions with half-lives shorter than 1 ps are shown in bold. A shorter half-life indicates a faster kinetic transition between states for MAX, leading to lower accumulation in individual state populations.

State transition	ΔF (kJ/mol)	Transition halflife (ps)
A to B	12.30	15.39
B to A	6.86	1.73
A to C	12.06	13.94
C to A	2.65	**0.32**
A to D	17.81	139.71
D to A	6.83	1.71
B to D	11.15	5.79
D to B	3.93	**0.54**
D to E	9.86	5.78
E to D	3.93	**0.53**
E to F	7.03	1.86
F to E	3.56	**0.46**
E to G	15.04	46.08
G to E	6.65	1.59
D to F	13.07	20.89
F to D	3.66	**0.48**
F to H	14.26	33.74
H to F	4.48	**0.67**
G to I	8.79	3.76
I to G	2.26	**0.27**

Nevertheless, it is possible to obtain one-dimensional (1D) profiles where free-energy depends of only one CV, after the other CV has been integrated out. This sort of calculation will allow us to directly compare free-energy barriers to experimental findings (see Jämbeck et al.[44] for a more detailed explanation), since information about binding modes in experimental solutes is usually not available. Then, it is possible to use experimental binding free energies ΔG_{bind} and get indirect leverage on the quality of the computed free-energy barriers. In this work, 1D free-energy profiles F(s_1) (comparable to the experimental ΔG_{bind}) can be obtained by means of the following expression:[44,45]

$$F(s_1) = -\beta^{-1} \ln \left[\frac{\int e^{-\beta\, F(s_1,s_2)}\, ds_2}{\int e^{-\beta\, F(s_1,s_2)}\, ds_1 ds_2} \right], \tag{2}$$

where s_1 and s_2 are the CV, $\beta = 1/(k_{\mathrm{B}}T)$, k_{B} is the Boltzmann constant and T is the absolute temperature. Here the free-energy $F(s_1)$ depends only of s_1, whereas all possible contributions due to s_2 have been integrated out and averaged. Same procedure has been applied in order to obtain $F(s_2)$ by integrating out the variable s_1. The results of Fig. 5 reveal a series of free-energy barriers around the stable basins.

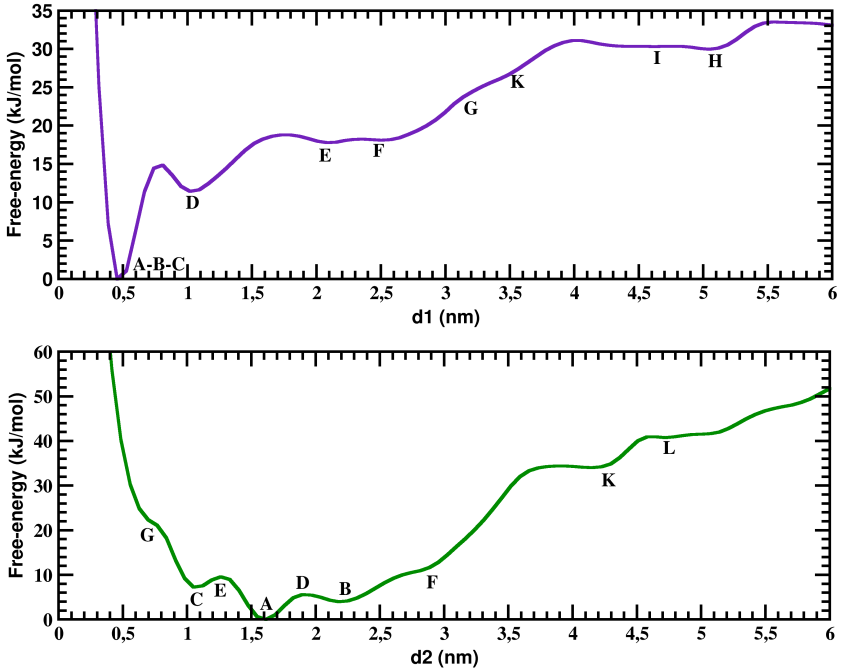


Figure 5: 1D integrated free-energy profiles for two CVs at 3500 ns. Basin regions marked with same labels (A-L) as in Fig. 4. In order to directly compute the height of free-energy barriers, absolute minima of each free-energy are set equal to zero.

Fig. 5 shows clearly that, along CV *d1*, the largest energy barrier is between basins A and D, within the low-energy region. This barrier is of about 15 kJ/mol, much larger than the thermal energy at 310K ($\sim 5.8\ k_B T$). It can be compared to other protein bottlenecks, such as those related to the RAS family of oncogenes, where barriers of $\sim$ 20 kJ/mol due to changes in specific distances and orientations were estimated for the activation of mutated RAS proteins (Fourier Transform Infrared Spectroscopy[46] and simulations[47] or to the case of KRas-G12D, when bound to a cell membrane, where angular barriers in the range 8-17 kJ/mol were estimated.[48]

Location of potential targeting pockets at the surface of MAX

After characterizing the possible configurations of MAX, our interest now is to identify which of these configurations are of interest for *in silico* drug design efforts. A possible therapeutic strategy could be to target unbound MAX, limiting its availability for MYC binding, as described in the Introduction. We should note here that the crystal structure employed as initial condition in our simulations corresponds to a MAX protein in a MYC-MAX dimer. The unbound MAX protein in solution can be expected to be in one of the free-energy basins described in the previous subsections. Therefore, to identify druggable targets from MAX configurations we should explore the configurations and free-energy landscape connecting the minimum free energy state identified in our calculations (state A in Table 1) with the CS structure.

By following the MFEP from the most stable state (basin A) to the CS, it is clear that the MFEP, depicted in red in Fig. 6A, passes through basin D, leading to a high free-energy barrier of approximately 75 kJ/mol. Interestingly, structure G shows a high degree of similarity to the CS within the helical domains when compared to other basins (highlighted in blue in Table 1). Upon aligning the structures of basin G and CS, it becomes evident that MAX only needs to rotate residues 202–225 by 164.2° and residues 264–284 by 133.1° to adopt the conformation of CS (see Fig. 6B). This suggests that state G may represent a transition state to adopt the conformation of the CS during the process of the MAX-MAX/MAX-MYC dimerization. Given the higher population of state G, reflected by the greater number of conformers compared to other basins in Fig. 4, we further examined the MFEP from basin A to G, which proceeds through basin D and involves a free-energy barrier of 32.01 kJ/mol. Additionally, the MFEP from basin G to the CS revealed a free-energy difference of 49.60 kJ/mol, which is smaller than that from basin A to the CS. These findings highlight states A as the most stable structure of MAX in solution, D and G as key intermediates, suggesting their potential as promising therapeutic targets in MAX-associated diseases. An interesting feature of these structures is the fact that they expose positively charged regions. In Fig. 6C

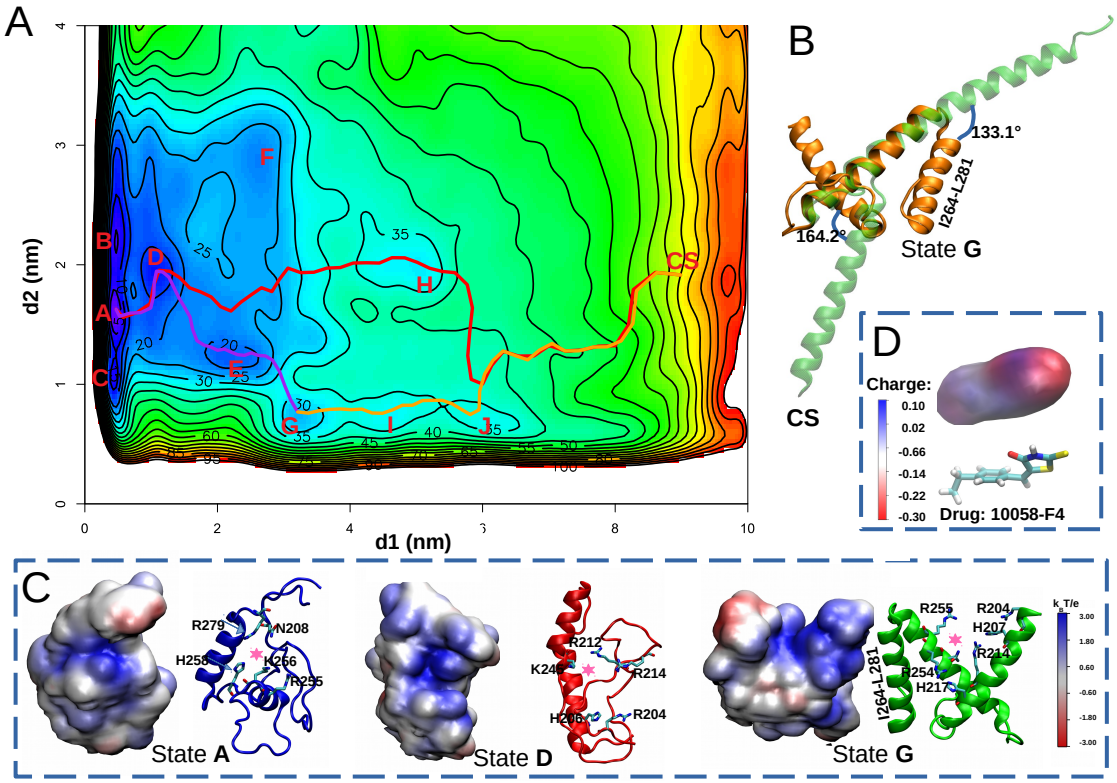


Figure 6: A: MFEPs between A to CS (red), A to G (purple), and G to CS (orange). B: Aligned structures of state G and the CS. Angles between vectors along the helical domain are shown here. C: Electrostatic potential of MAX in states A, D, and G in units of k_BT/e ($k_BT/e \approx 26.7$ mV). The potential drug-binding pocket of each state is marked with a pink star, and positive side chains of these pockets are highlighted here. D: Charge distribution of small-molecule c-Myc inhibitor 10058-F4.

we show the electrostatic potential of the representative structures of the A, D and G states calculated at the Poisson-Boltzmann level of theory using the APBS[49] software. As seen in that figure, there are large pockets of these structures at potentials around ~ 80 mV. This suggests electrostatic interaction as a suitable mechanism for binding to MAX. Fig. 6D shows the charge distribution of the small-molecule c-MYC inhibitor 10058-F4, which has been proposed to interact with both MAX and MYC and inhibit their heterodimerization.[16] Our calculations suggest that the binding of this molecule to MAX may be due to electrostatic interactions.

Conclusions

The monomeric, unbound MAX motif exhibits a rich conformational ensemble, as revealed by extensive classical MD (2 μs) and WT-MetaD simulations (3.5 μs). Classical MD runs revealed five dominant structural clusters, which guided the selection of two distances as CVs for the subsequent WT-MetaD enhanced sampling to further explore the structural diversity of MAX.

The 2D FES constructed from WT-MetaD simulations revealed a broad conformational space, featuring numerous low-energy basins corresponding to metastable states. The most stable basin (A) lies ∼75 kJ/mol below the experimentally resolved crystal structure, indicating that the latter is a low-populated state in solution. Basins A and B form a compact cluster of highly stable conformations with tightly packed helices, while D–G represent moderately stable states. By contrast, higher-energy basins (H–L) are associated with partially unfolded or extended conformations, underscoring the intrinsic flexibility of MAX. The Minimum Free-Energy Path analysis using the package R-Metadynminer indicated that the textcolorredtransition from A to D exhibits the largest energy barrier (Δ F $\sim$ 18 kJ/mol), suggesting a significant kinetic bottleneck. In contrast, transitions among other basins (e.g., A to B or B to D) have lower energy barriers and shorter lifetimes. Certain basins such as C and I were identified as highly transient (lifetime $\leq$ 0.5 ps), leading to their absence in the WT-MetaD output frames. This indicates that not all minima contribute equally to the long-lived conformational ensemble of MAX.

Importantly, the substantial energy gap between the solution-phase stable states and the crystal structure indicates the need to target MAX in its native lowest energy conformations. The identification of persistent pockets within low-energy basins (particularly A, D, and G) provides promising templates for rational structure-based drug design against the MYC–MAX interaction interface. Furthermore, electrostatic potential maps of states A, D, and G reveal positively charged pockets, while the small-molecule MYC inhibitor 10058-F4 exhibits a negatively charged electrostatic profile, suggesting potential electrostatic comple-

mentarity that could be exploited for binding.

Data Availability

The raw data and analysis scripts that support the findings of this study can be found in the corresponding repository of Github.[42]

Author Contributions

Huixia Lu: Writing – original draft, Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization.

Jordi Marti: Writing – original draft, Writing – review & editing, Resources, Conceptualization, Funding acquisition.

Jordi Faraudo: Writing – original draft, Writing – review & editing, Resources, Conceptualization, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by grants PID2021-124297NB-C32 and PID2021-124297NB-C33 funded by MCIN/AEI/10.13-039/501100011033 and the "SeveroOchoa" grant CEX2023-001263-S for Centers of Excellence awarded to ICMAB-CSIC. J.M. thanks the *Generalitat de Catalunya* for the support through the grant *Grup de Recerca SGR-Cat2021 Condensed, Complex and Quantum Matter Group* reference 2021SGR-01411 and to the Polytechnic Uni-

versity of Catalonia-Barcelona Tech through the funding AGRUPS. The authors thankfully acknowledge computer resources from MareNostrum5 supercomputer as well as technical support provided by BSC under project number RES-BCV-2025-1-0008.

References

(1) Blackwood, E. M.; Eisenman, R. N. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **1991**, *251*, 1211–1217.

(2) Grandori, C.; Cowley, S. M.; James, L. P.; Eisenman, R. N. The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annual Review of Cell and Developmental Biology* **2000**, *16*, 653–699.

(3) Mateyak, M. K.; Heasley, L. E. Repression of Myc function by the Mad proteins. *Seminars in Cancer Biology* **1997**, *8*, 37–49.

(4) Lüscher, B. Function and regulation of the transcription factors of the Myc/Max/Mad network. *Gene* **2001**, *277*, 1–14.

(5) Sabo, A.; Kress, T. R.; Pelizzola, M.; others Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* **2014**, *511*, 488–492.

(6) Austen, M.; Luscher, B. Myc and Max—transcriptional activators and regulators. *The International Journal of Biochemistry Cell Biology* **2003**, *35*, 289–305.

(7) Dubois, M. F.; others Transcriptional regulation by MYC: a multipronged strategy. *FEBS Letters* **2014**, *588*, 2705–2711.

(8) Norton, J. D.; Atherton, G. T. Protein-protein interactions in the control of gene expression: lessons from the Myc/Max/Mad network. *Biochemical Journal* **1998**, *333*, 1–9.

(9) Ainsley, J. A.; Kim, M. J.; others The role of bHLH transcription factors in neurodevelopment and neurogenesis. *Developmental Biology* **2014**, *390*, 13–23.

(10) Mrozek, K.; Bloomfield, C. D. The MAX gene in human cancer: mutations, loss of expression, and implications. *Genes, Chromosomes Cancer* **2008**, *47*, 753–763.

(11) Hopewell, R.; Zervos, P.; others Structural insights into MAX dimerization and DNA binding. *Journal of Molecular Biology* **2013**, *425*, 3490–3502.

(12) D'Angelo, C. R.; Rossi, L.; Pascarella, S. Evolutionary dynamics of the MYC/MAX/MAD network in metazoans. *Genome Biology and Evolution* **2017**, *9*, 1352–1365.

(13) Keehan, M. E.; Duarte, C. M.; Perez-Riba, A.; others Evolutionary conservation and divergence in MYC/MAX/MXD family proteins. *Molecular Biology and Evolution* **2021**, *38*, 4209–4224.

(14) Garralda, E.; Beaulieu, M.-E.; Moreno, V.; Casacuberta-Serra, S.; Martinez-Martin, S.; Foradada, L.; Alonso, G.; Masso-Valles, D.; López-Estévez, S.; Jauset, T.; others MYC targeting by OMO-103 in solid tumors: a phase 1 trial. *Nature medicine* **2024**, *30*, 762–771.

(15) Soucek, L.; Whitfield, J. R.; others Omomyc: a MYC mutant engineered to disrupt MYC dimerization and transcriptional activity. *Oncogene* **2002**, *21*, 6600–6607.

(16) Follis, A. V.; Hammoudeh, D. I.; Daab, A. T.; Metallo, S. J. Small-molecule perturbation of competing interactions between c-Myc and Max. *Bioorganic & medicinal chemistry letters* **2009**, *19*, 807–810.

(17) Hart, J. R.; Garner, A. L.; Yu, J.; Ito, Y.; Sun, M.; Ueno, L.; Rhee, J.-K.; Baksh, M. M.; Stefan, E.; Hartl, M.; others Inhibitor of MYC identified in a Kröhnke pyridine library. *Proceedings of the National Academy of Sciences* **2014**, *111*, 12556–12561.

(18) AlSultan, D.; Kavanagh, E.; O'Grady, S.; Eustace, A. J.; Castell, A.; Larsson, L.-G.; Crown, J.; Madden, S. F.; Duffy, M. J. The novel low molecular weight MYC antagonist MYCMI-6 inhibits proliferation and induces apoptosis in breast cancer cells. *Investigational new drugs* **2021**, *39*, 587–594.

(19) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. Structure and Dynamics of an Unfolded Protein Examined by Molecular Dynamics Simulation. *Journal of the American Chemical Society* **2012**, *134*, 3787–3791.

(20) Larsen, A. H.; Wang, Y.; Bottaro, S.; Grudinin, S.; Arleth, L.; Lindorff-Larsen, K. Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution. *PLoS computational biology* **2020**, *16*, e1007870.

(21) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **2008**, *29*, 1859–1865.

(22) Brooks, B. R.; Brooks III, C. L.; Mackerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; others CHARMM: the biomolecular simulation program. *Journal of computational chemistry* **2009**, *30*, 1545–1614.

(23) Lee, J.; Cheng, X.; Jo, S.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Biophysical journal* **2016**, *110*, 641a.

(24) Huang, J.; MacKerell Jr, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry* **2013**, *34*, 2135–2145.

(25) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *Journal of computational chemistry* **2005**, *26*, 1668–1688.

(26) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V.

W. D.; Forouzesh, N.; Giese, T. J.; Gotz, A. W.; Gohlke, H.; others AmberTools. *Journal of chemical information and modeling* **2023**, *63*, 6183–6191.

(27) Bussi, G.; Laio, A.; Parrinello, M. Equilibrium free energies from nonequilibrium metadynamics. *Physical review letters* **2006**, *96*, 090601.

(28) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical review letters* **2008**, *100*, 020603.

(29) Bonomi, M.; Camilloni, C.; Vendruscolo, M. Metadynamic metainference: Enhanced sampling of the metainference ensemble using metadynamics. *Scientific reports* **2016**, *6*, 31232.

(30) Lu, H.; Marti, J. Cellular absorption of small molecules: free energy landscapes of melatonin binding at phospholipid membranes. *Scientific reports* **2020**, *10*, 9235.

(31) Lu, H.; Martí, J. Influence of cholesterol on the orientation of the farnesylated GTP-bound KRas-4B binding with anionic model membranes. *Membranes* **2020**, *10*, 364.

(32) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; others PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **2009**, *180*, 1961–1972.

(33) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Computer physics communications* **2014**, *185*, 604–613.

(34) Promoting transparency and reproducibility in enhanced molecular simulations. *Nature methods* **2019**, *16*, 670–673.

(35) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **1996**, *14*, 33–38.

(36) Greenacre, M.; Groenen, P. J.; Hastie, T.; d'Enza, A. I.; Markos, A.; Tuzhilina, E. Principal component analysis. *Nature Reviews Methods Primers* **2022**, *2*, 100.

(37) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; others *MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations*; 2019.

(38) Likas, A.; Vlassis, N.; Verbeek, J. J. The global k-means clustering algorithm. *Pattern recognition* **2003**, *36*, 451–461.

(39) Trapl, D.; Spiwok, V. Analysis of the Results of Metadynamics Simulations by metadynminer and metadynminer3d. *arXiv preprint arXiv:2009.02241* **2020**,

(40) Hošek, P.; Spiwok, V. Metadyn View: Fast web-based viewer of free energy surfaces calculated by metadynamics. *Computer Physics Communications* **2016**, *198*, 222–229.

(41) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004**, *25*, 1605–1612.

(42) Lu, H. Project MAX. `https://github.com/HuixiaLuScienceRocks/Probing_the_Structural_Dynamics_of_Unbound_MAX_Protein_Insights_from_Well_Tempered_Metadynamics`, 2025.

(43) Henkelman, G.; Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *The Journal of chemical physics* **2000**, *113*, 9978–9985.

(44) Jambeck, J. P.; Lyubartsev, A. P. Exploring the free energy landscape of solutes embedded in lipid bilayers. *The Journal of Physical Chemistry Letters* **2013**, *4*, 1781–1787.

(45) Roux, B. Statistical mechanical equilibrium theory of selective ion channels. *Biophysical journal* **1999**, *77*, 139–153.

(46) Li, Y.; Zhang, Y.; Großeruschkamp, F.; Stephan, S.; Cui, Q.; Kotting, C.; Xia, F.; Gerwert, K. Specific substates of Ras to interact with GAPs and effectors: revealed by theoretical simulations and FTIR experiments. *The journal of physical chemistry letters* **2018**, *9*, 1312–1317.

(47) Hu, Z.; Martí, J. Atomic-level mechanisms of abnormal activation in NRAS oncogenes from two-dimensional free energy landscapes. *Nanoscale* **2025**, *17*, 4047–4057.

(48) Lu, H.; Martí, J. Predicting the conformational variability of oncogenic GTP-bound G12D mutated KRas-4B proteins at zwitterionic model cell membranes. *Nanoscale* **2022**, *14*, 3148–3158.

(49) Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; others Improvements to the APBS biomolecular solvation software suite. *Protein science* **2018**, *27*, 112–128.

Supporting information: Conformational Dynamics of the Free MAX Protein Revealed by Well-Tempered Metadynamics

Huixia Lu,*,† Jordi Marti,*,‡ and Jordi Faraudo†

†*Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus de la UAB, E-08193 Bellaterra, Spain*

‡*Department of Physics, Universitat Politècnica de Catalunya-Barcelona Tech (UPC), Barcelona, Spain*

E-mail: huixialu@icmab.es; jordi.marti@upc.edu

Supporting Information Available

Figure S1 presents the PCA clustering results using 8 clusters. Main parameters of the WT-MetaD simulation are reported in Table S1. The Gaussian kernels deposited throughout the simulation are shown in Figure S2, while Figure S3 illustrates the distribution of the two collective variables within their two-dimensional space. As the simulation progresses,

Table S1: WT-MetaD simulation parameters used in this work.

Parameter	**Value**
Gaussian width of *d1* (nm)	0.2
Gaussian width of *d2* (nm)	0.2
Gaussian hill (kJ/mol)	1.2
Deposition stride (ps)	1
Biasfactor	8
Simulation time (ns)	3500

the bias potential gradually fills the underlying free-energy surface, enabling the system to explore the full phase space. A general downward trend in the height of the deposited bias is evident, becoming noticeably flatter around 1600 ns, with only minor fluctuations. Overall, the selected two-dimensional collective variable space has been comprehensively sampled. This behavior suggests that the WT-MetaD simulation is approaching convergence, although it does not yet provide definitive confirmation.

To further evaluate the convergence of the WT-MetaD simulations, we employ two widely used and reliable criteria: (1) comparing free-energy estimates as a function of each CV at different simulation time, and (2) performing block analysis to estimate the statistical error in the free-energy profiles. From the results of Figure S4 (a) and (b), we can see that the free-energy profile doesn't change significantly as a function of time, which indicates that the system most likely converged for the two CVs considered in the present study. In all cases, very well-converged profiles can be observed around 3350 ns, regardless of the type of CV.

To provide a more quantitative assessment, the average error in the free-energy profiles is computed and presented in Figure S5. As expected, the errors increase with the block length until they reach a plateau in all cases. The average errors tend to stabilize around 0.72 kJ/mol for *d1* and 1.06 kJ/mol for *d2*, indicating consistent behavior over extended simulation times. Both criteria indicate that the well-tempered metadynamics simulations performed in this study have converged.

Finally, MFEPs shown in Figure 4 of the main text are able to cross transition states marked with red stars in Figure S6. All free-energy profiles have been initialized to zero in order to estimate the height of the TS barrier. Coordinates are not indicated, but they can be inferred from Figure 4 of the main text.

References

(1) Henkelman, G.; Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *The Journal of chemical physics* **2000**, *113*, 9978–9985.

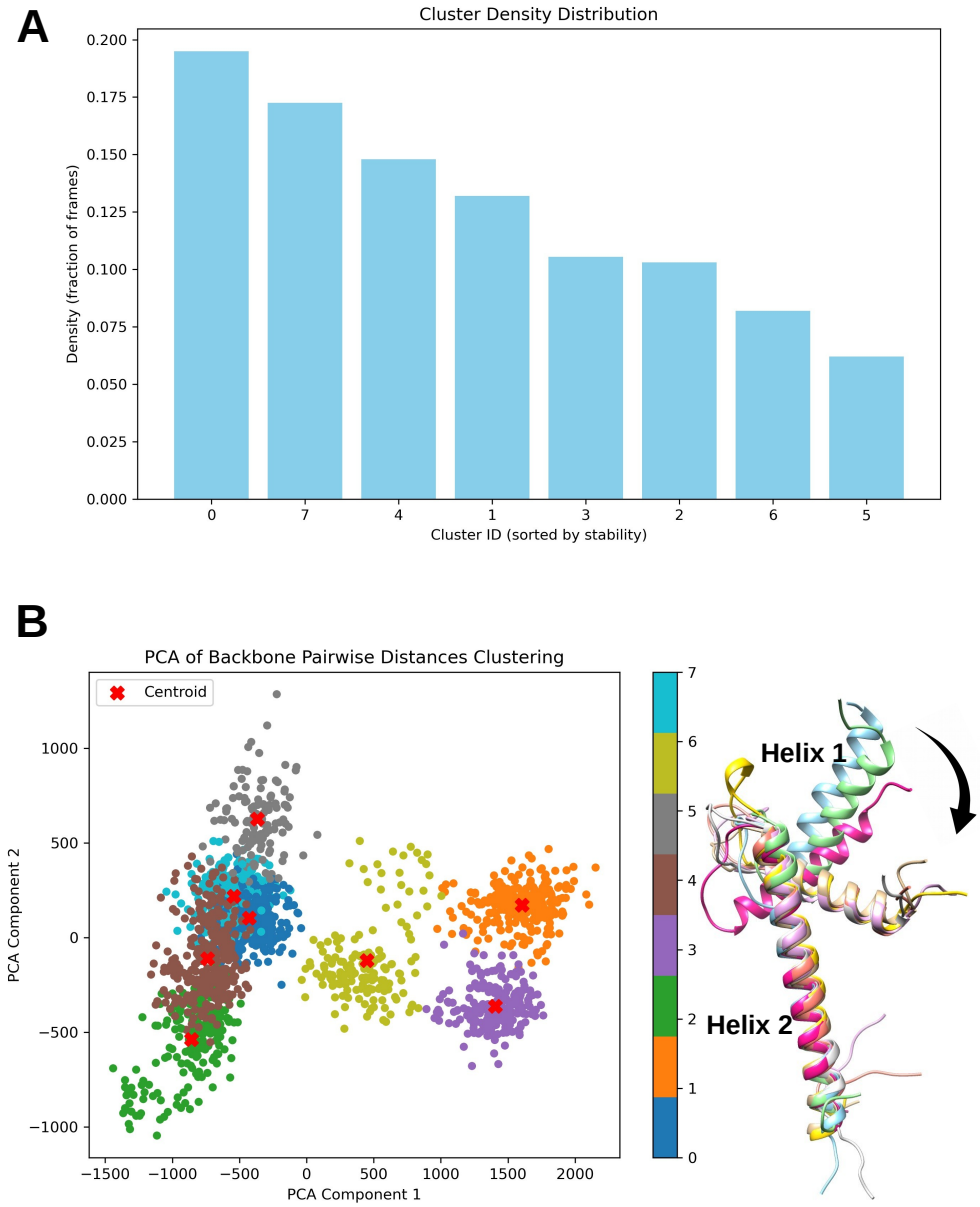


Figure S1: PCA-based clustering results using 8 clusters. A: Density profiles of the identified clusters. B: Visualization of the 8 clusters derived from PCA, with representative structures marked by red crosses and illustrated on the right. The overall clustering pattern closely mirrors the 5-cluster analysis presented in the main text, with the additional clusters reflecting finer subdivisions within the primary groups. These results support the conclusion that 5 clusters are sufficient to capture the dominant structural features of the dataset.

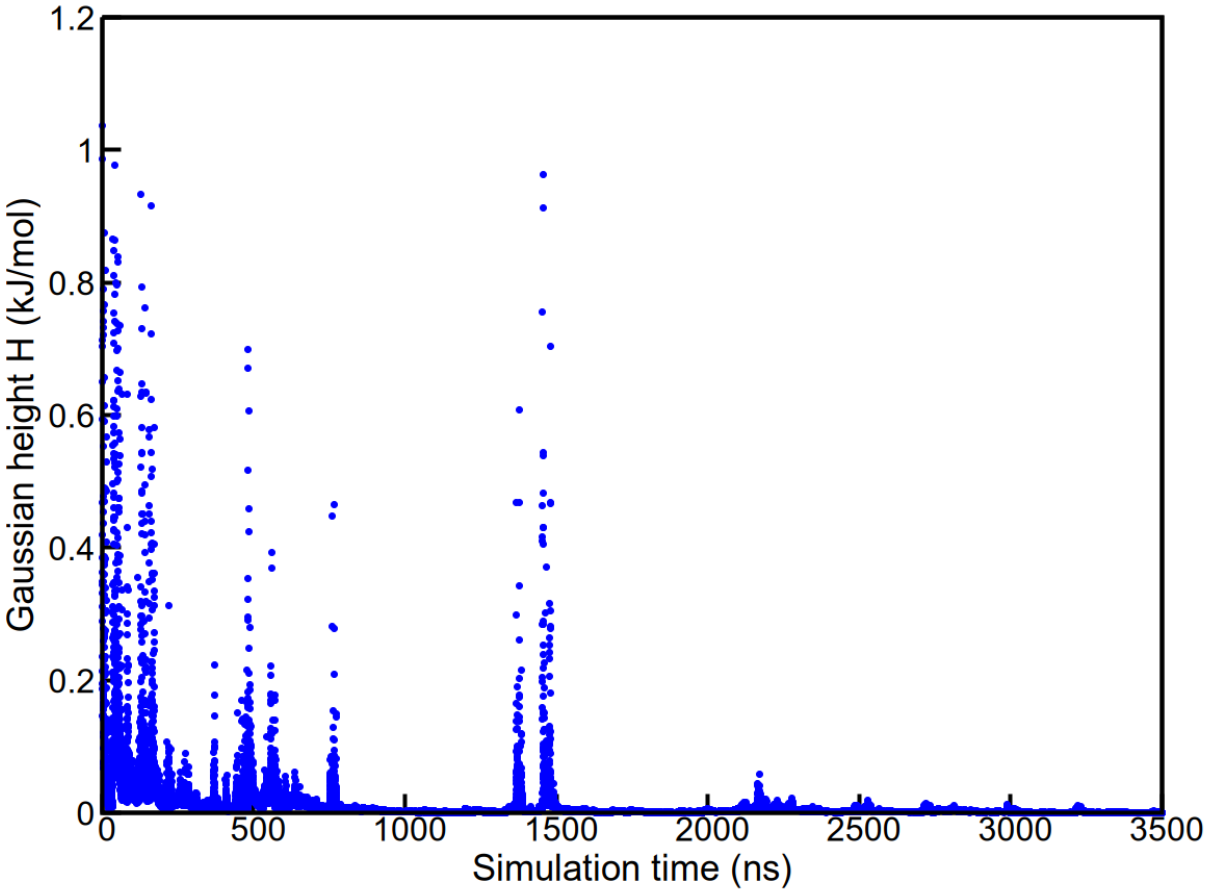


Figure S2: Time-evolution of the Gaussian height H during the overall simulation time of WT-MetaD.

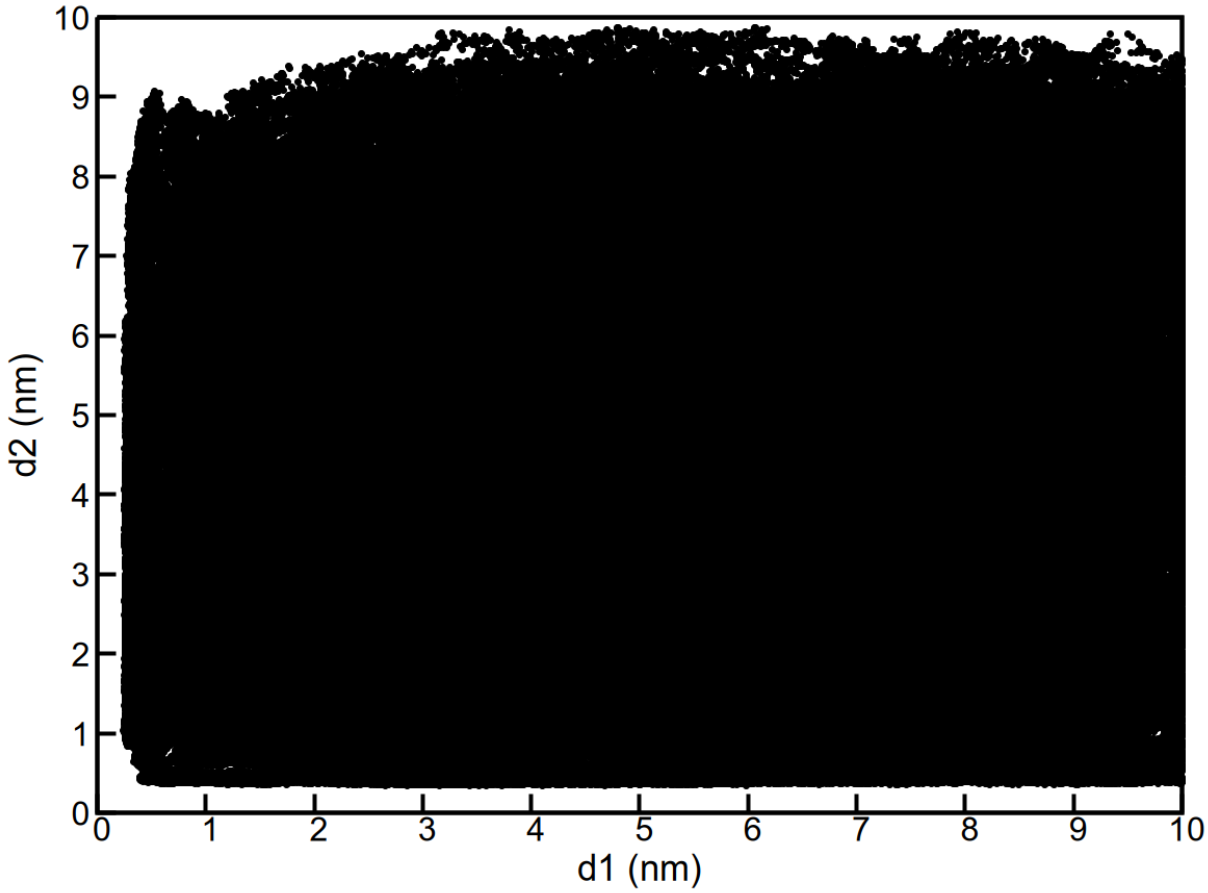
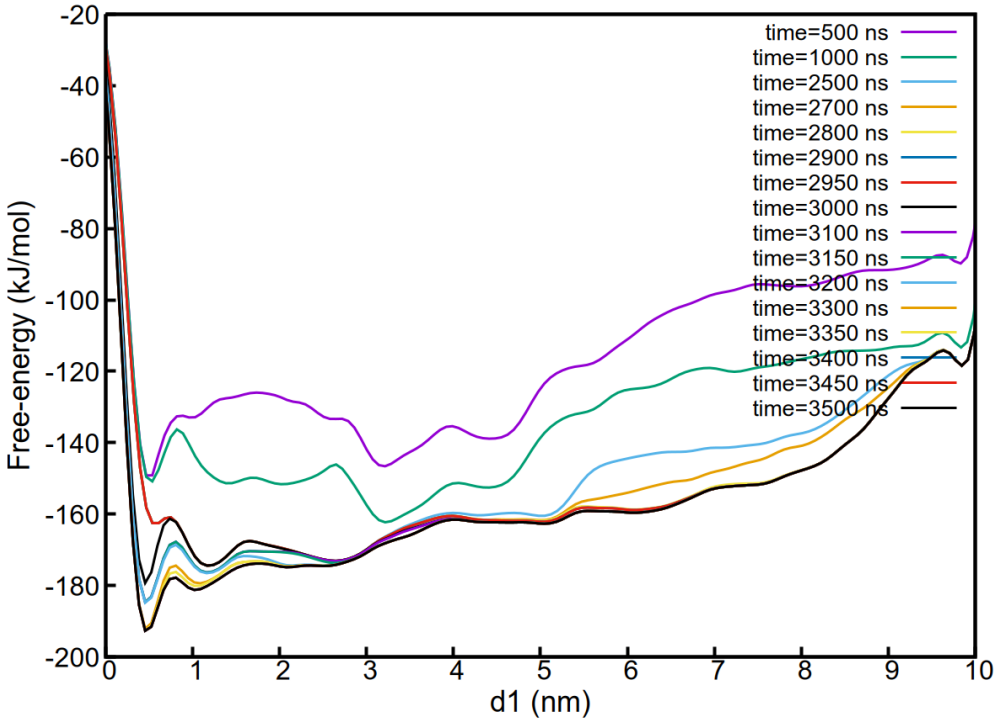
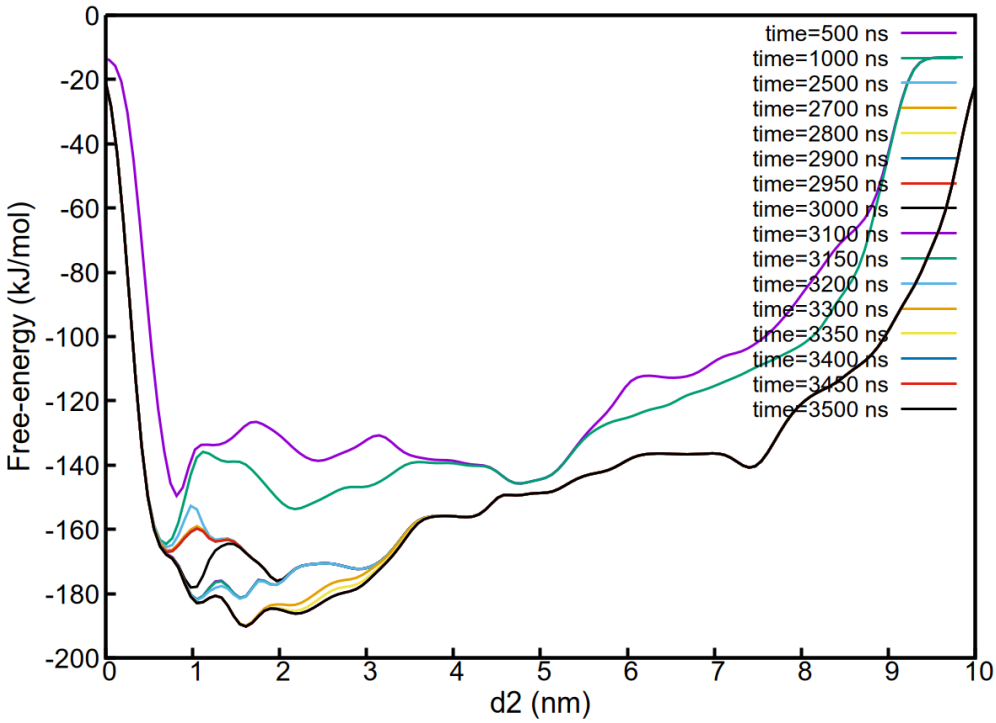


Figure S3: 2D distribution of the two collective variables sampled throughout the entire WT-MetaD simulation.



(a) Estimates of the free-energy as a function of *d1*.



(b) Estimates of the free-energy as a function of *d2*.

Figure S4: One-dimensional free-energy at different simulation times.

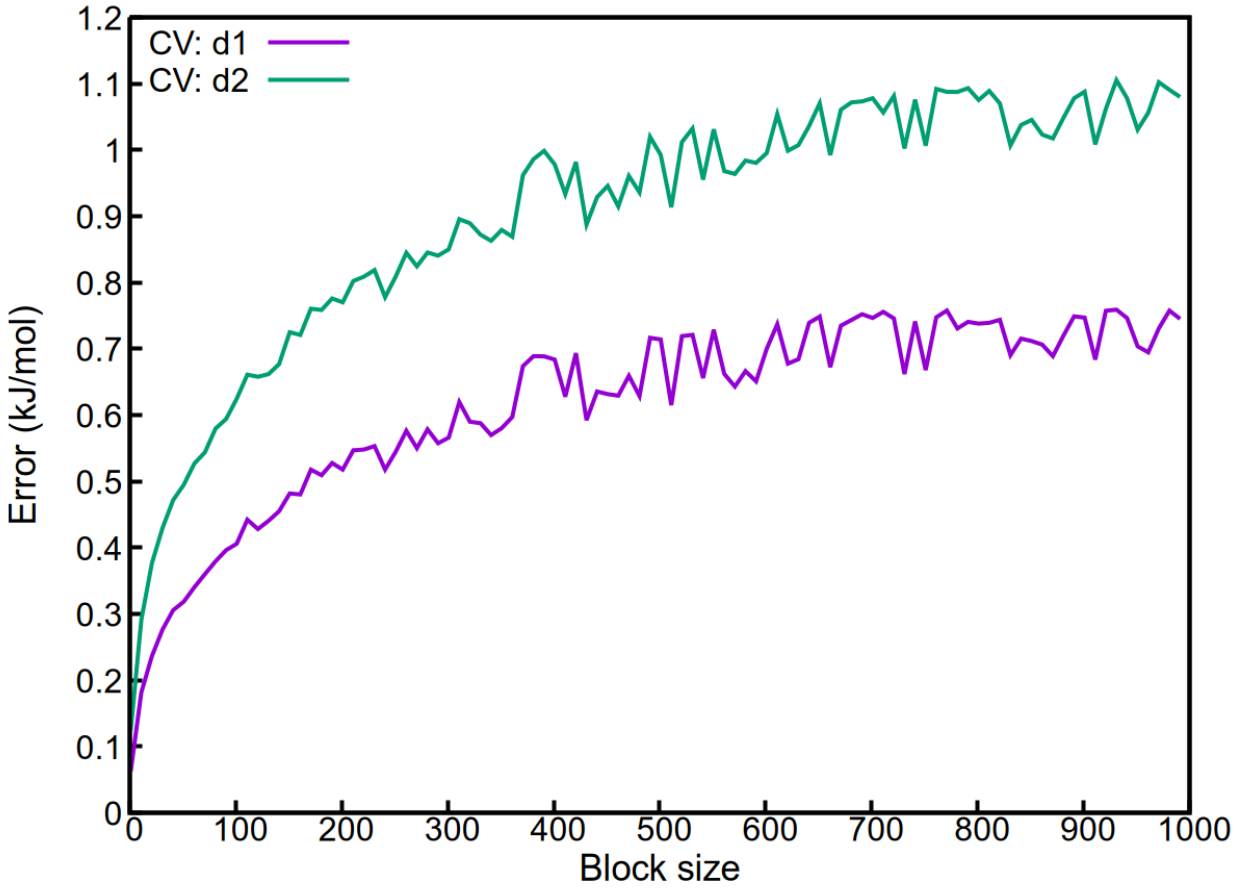


Figure S5: Block analysis profile using *d1* and *d2*.

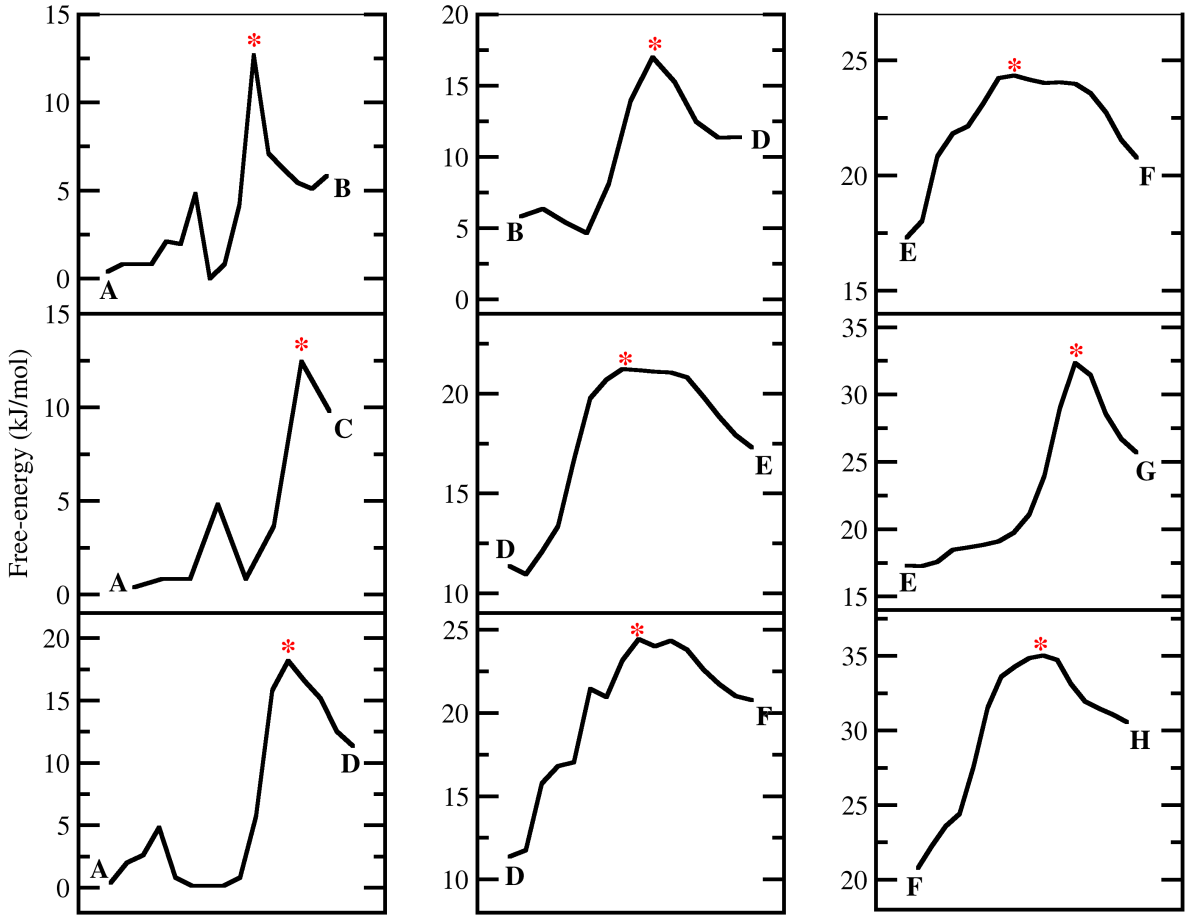
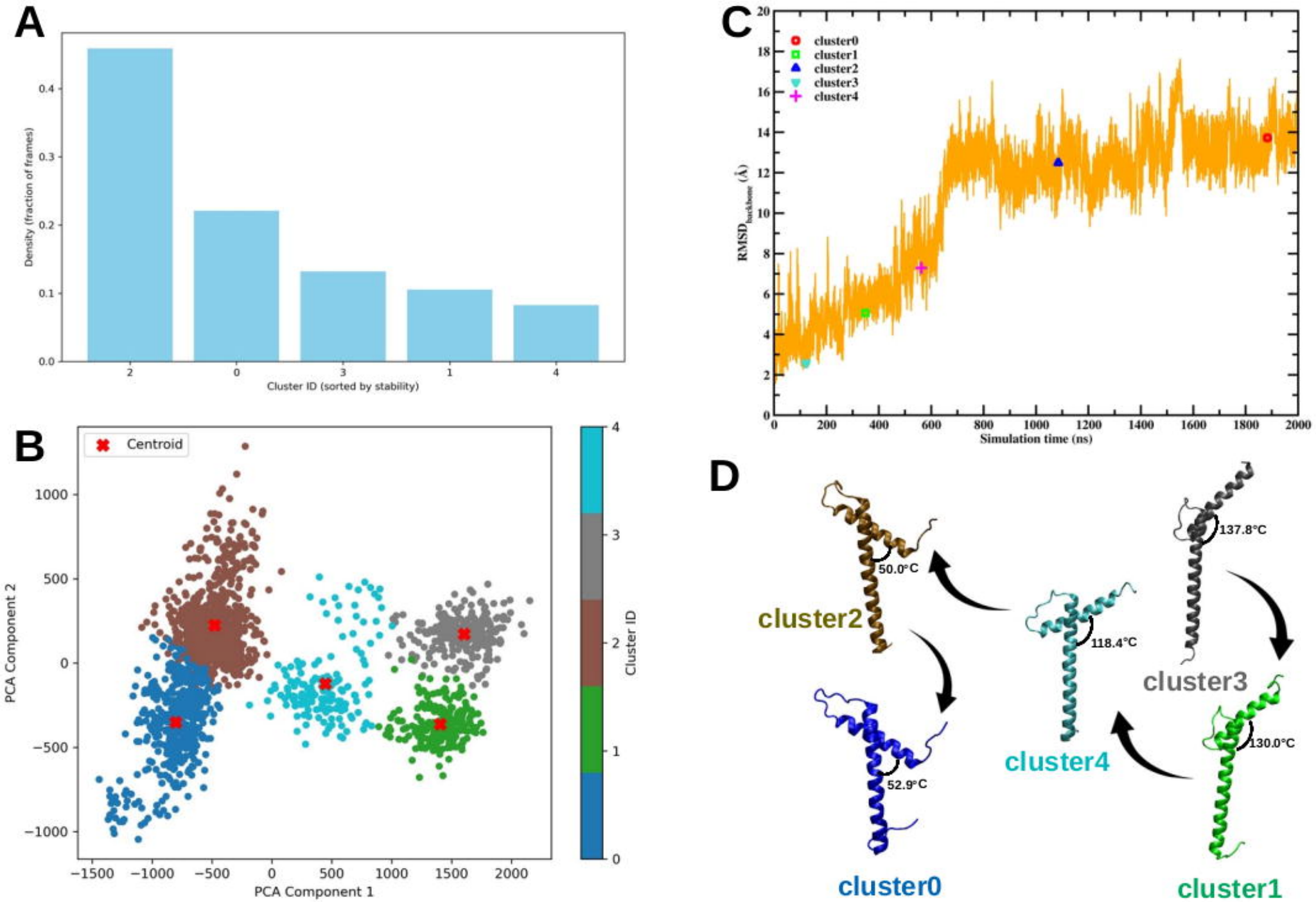
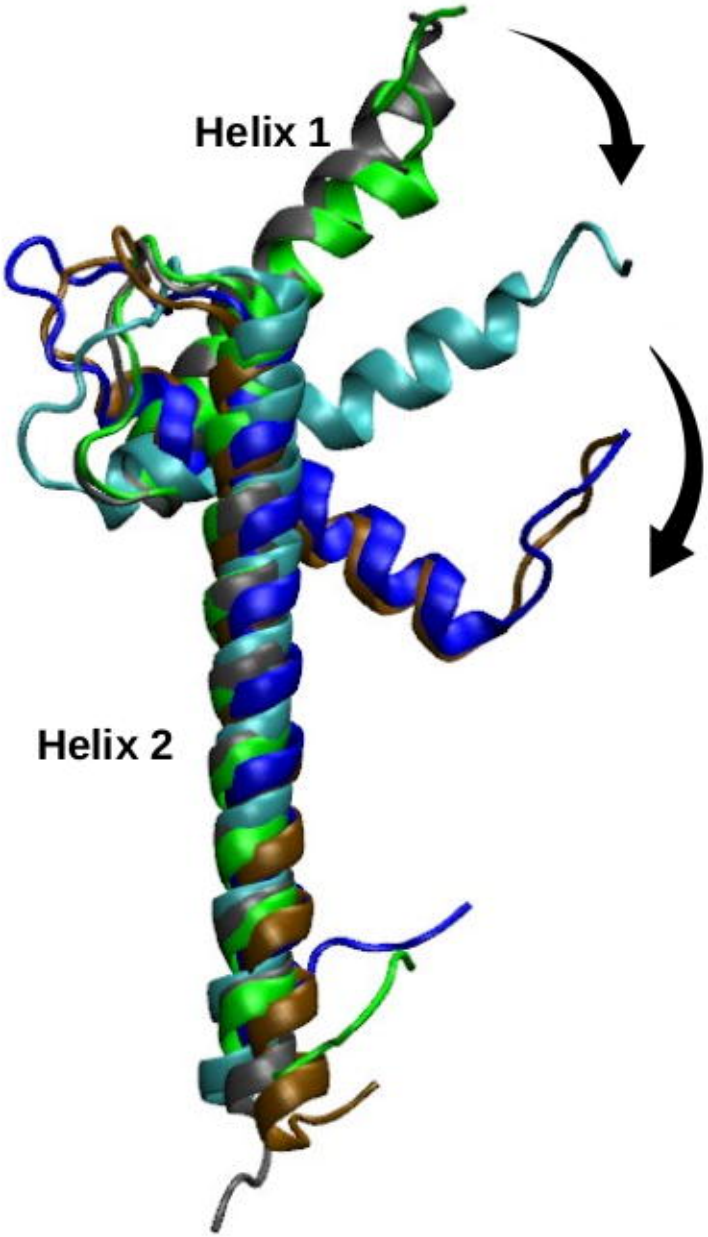
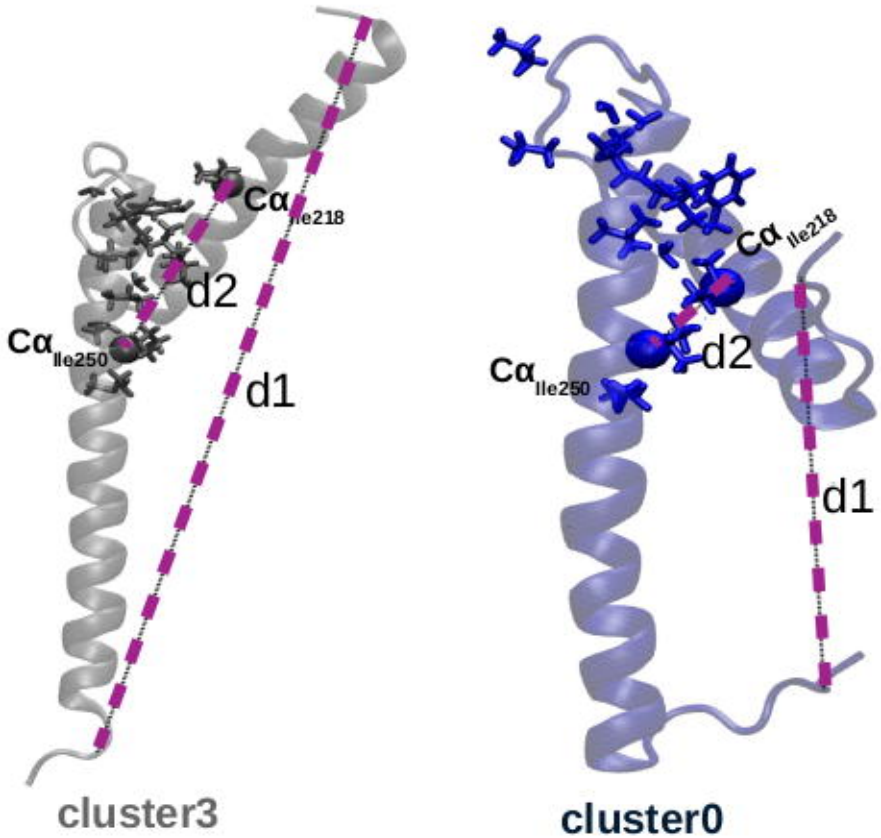
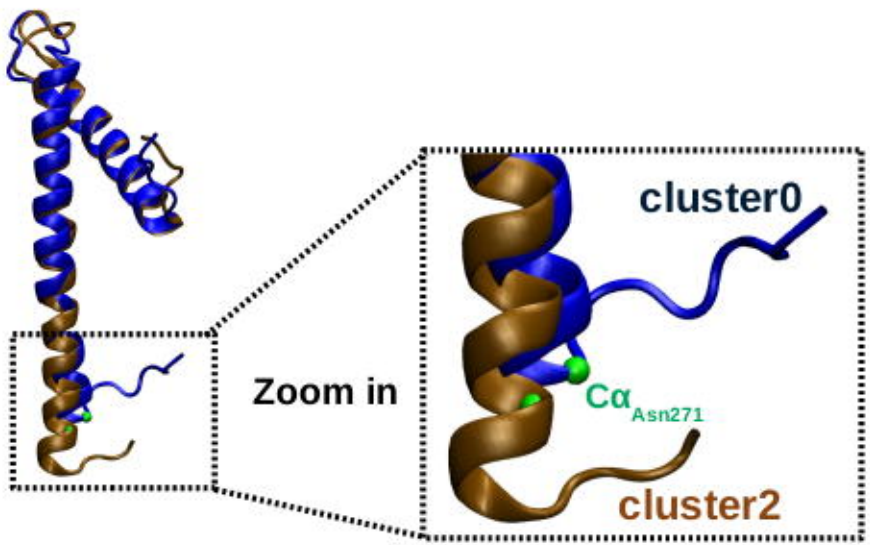


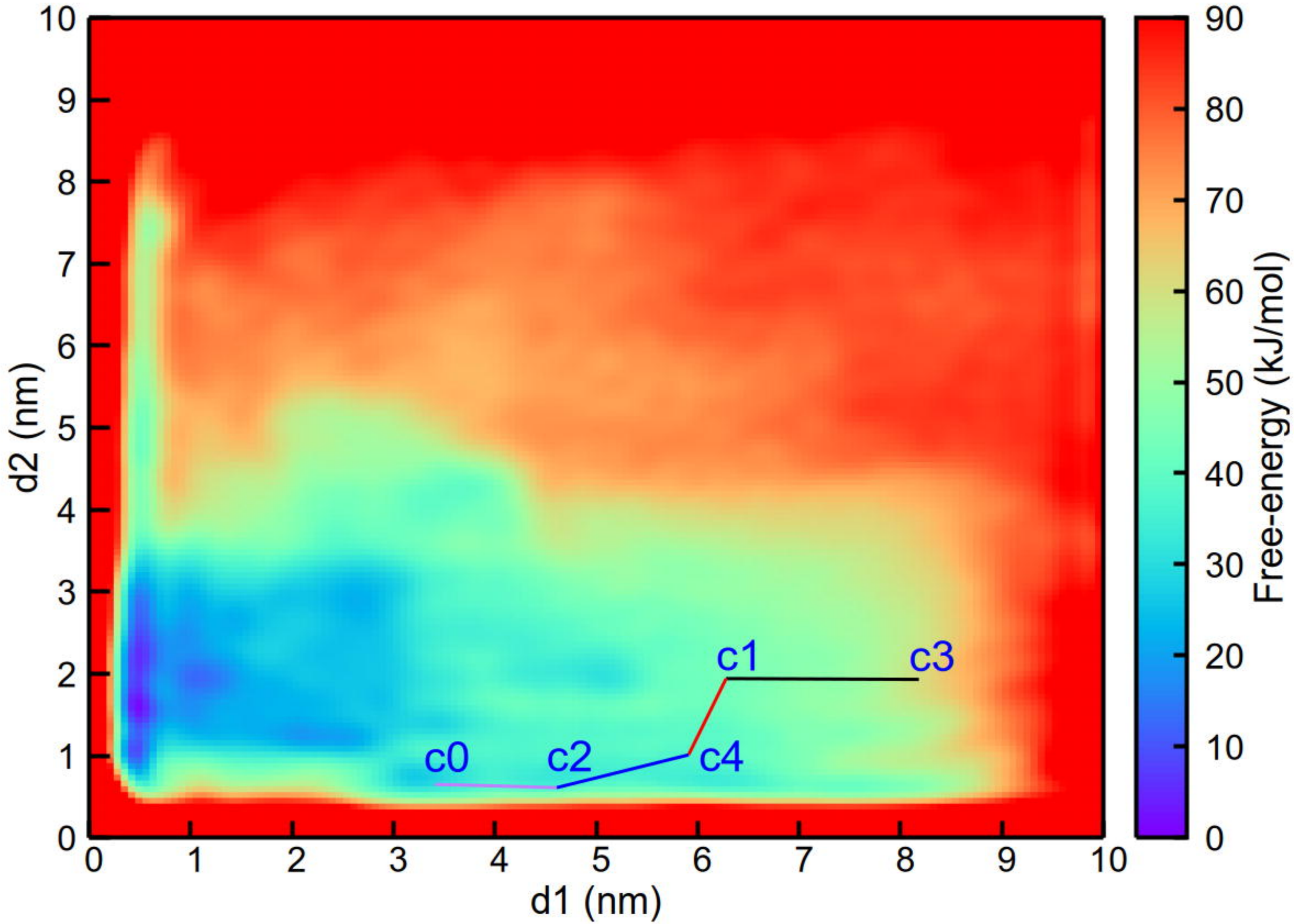
Figure S6: MFEPs shown in Fig. 4 of the main text connecting different basins identified by Nudged Elastic Band method.[1] The transition states are marked with red stars here.

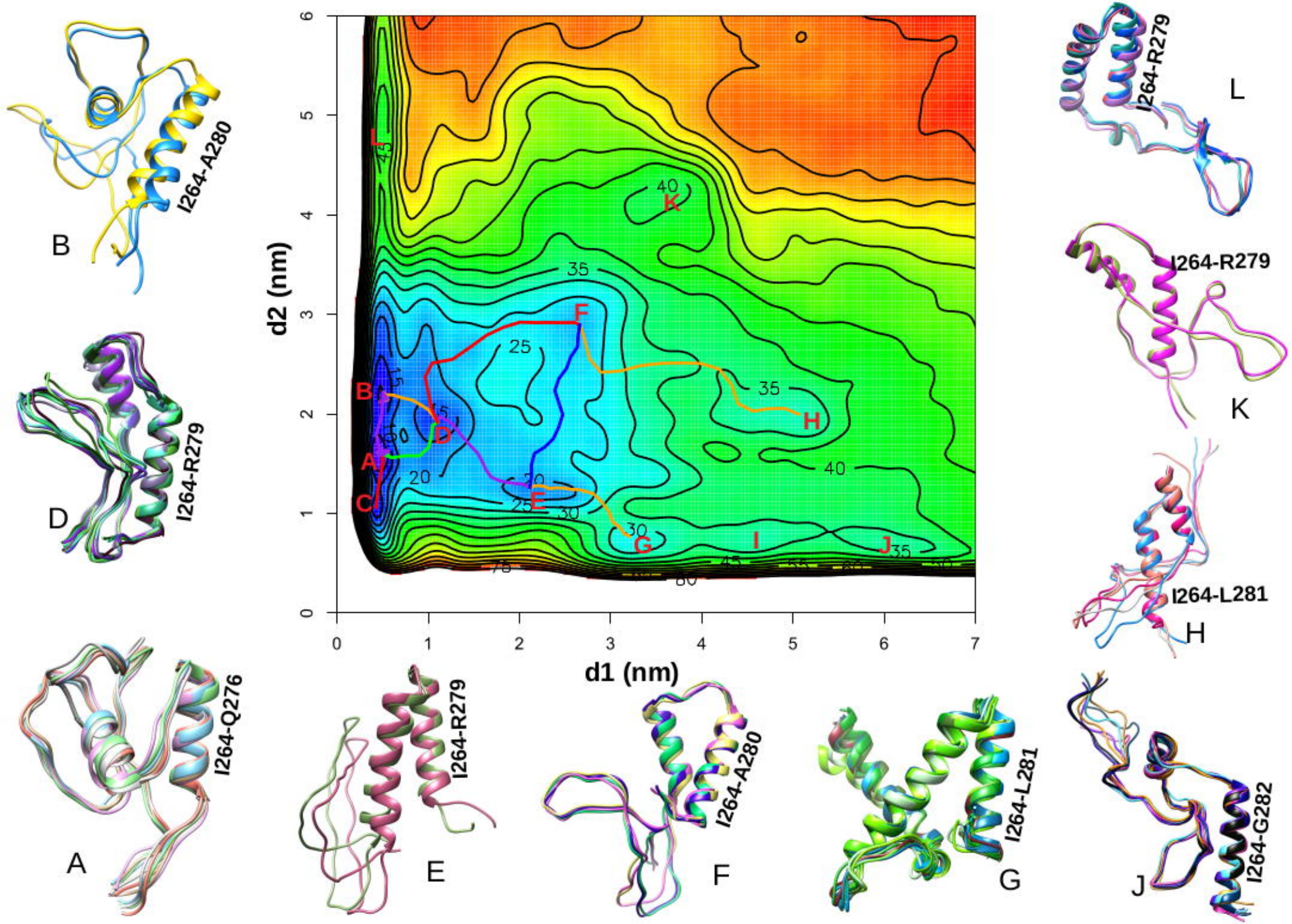

A
Density (fraction of frames)
Cluster ID (sorted by stability)
B
Centroid
PCA Component 1
PCA Component 2
Cluster ID
C
cluster0
cluster1
cluster2
cluster3
cluster4
Simulation time (ns)
D
50.0°C
137.8°C
118.4°C
52.9°C
130.0°C
cluster2
cluster3
cluster4
cluster0
cluster1

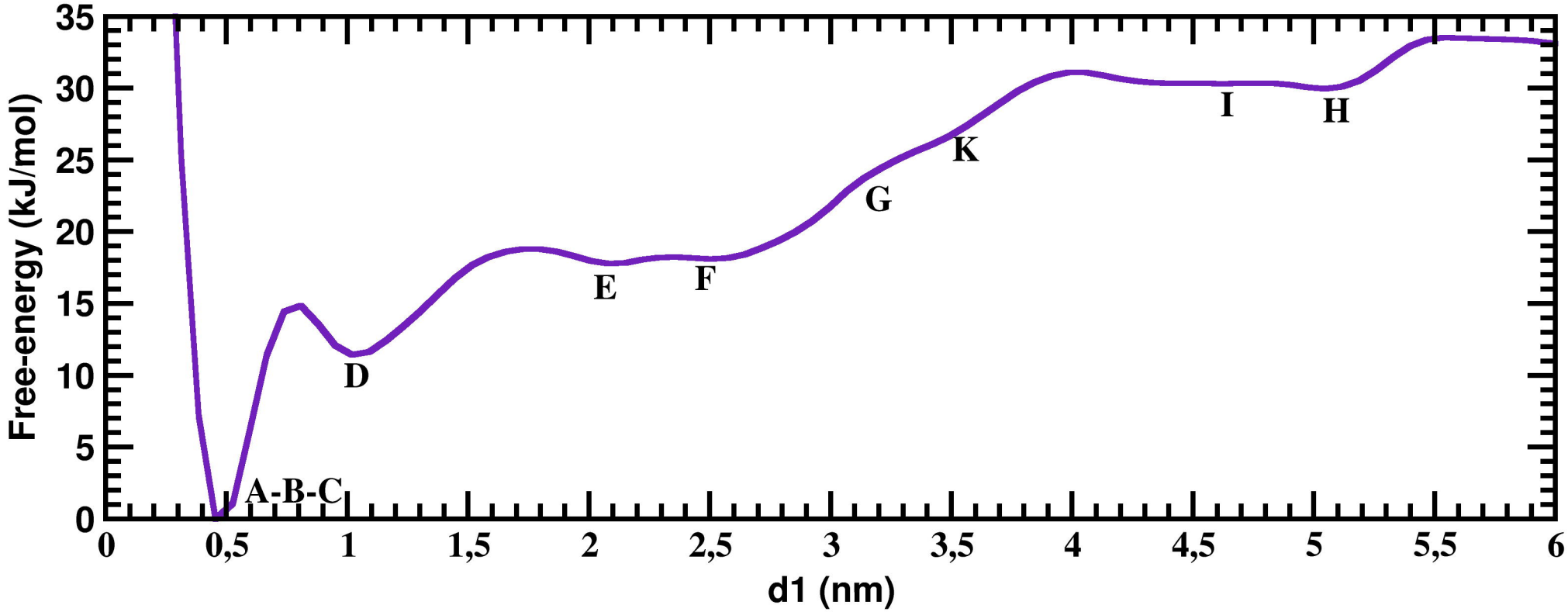

A
Helix 1
Helix 2

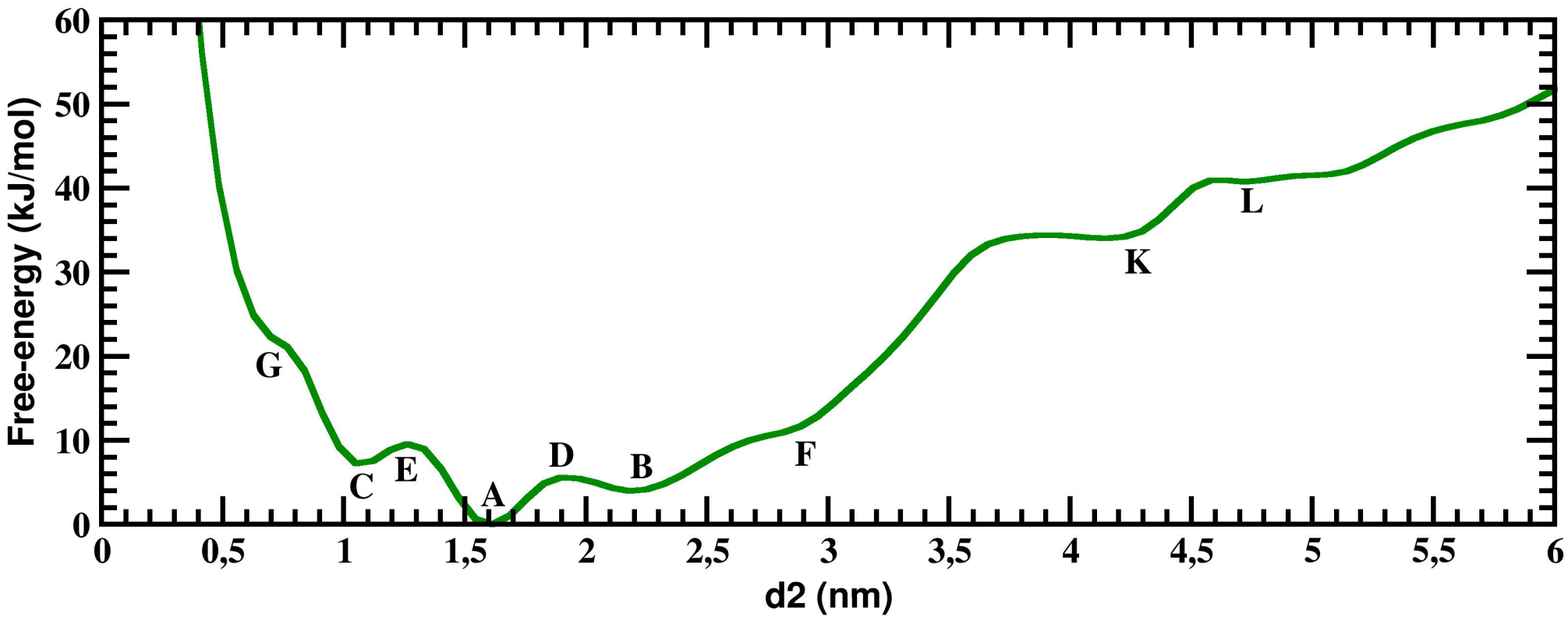

B
Cα Ile218
d2
Cα Ile250
d1
cluster3
Cα Ile218
Cα Ile250
d2
d1
cluster0

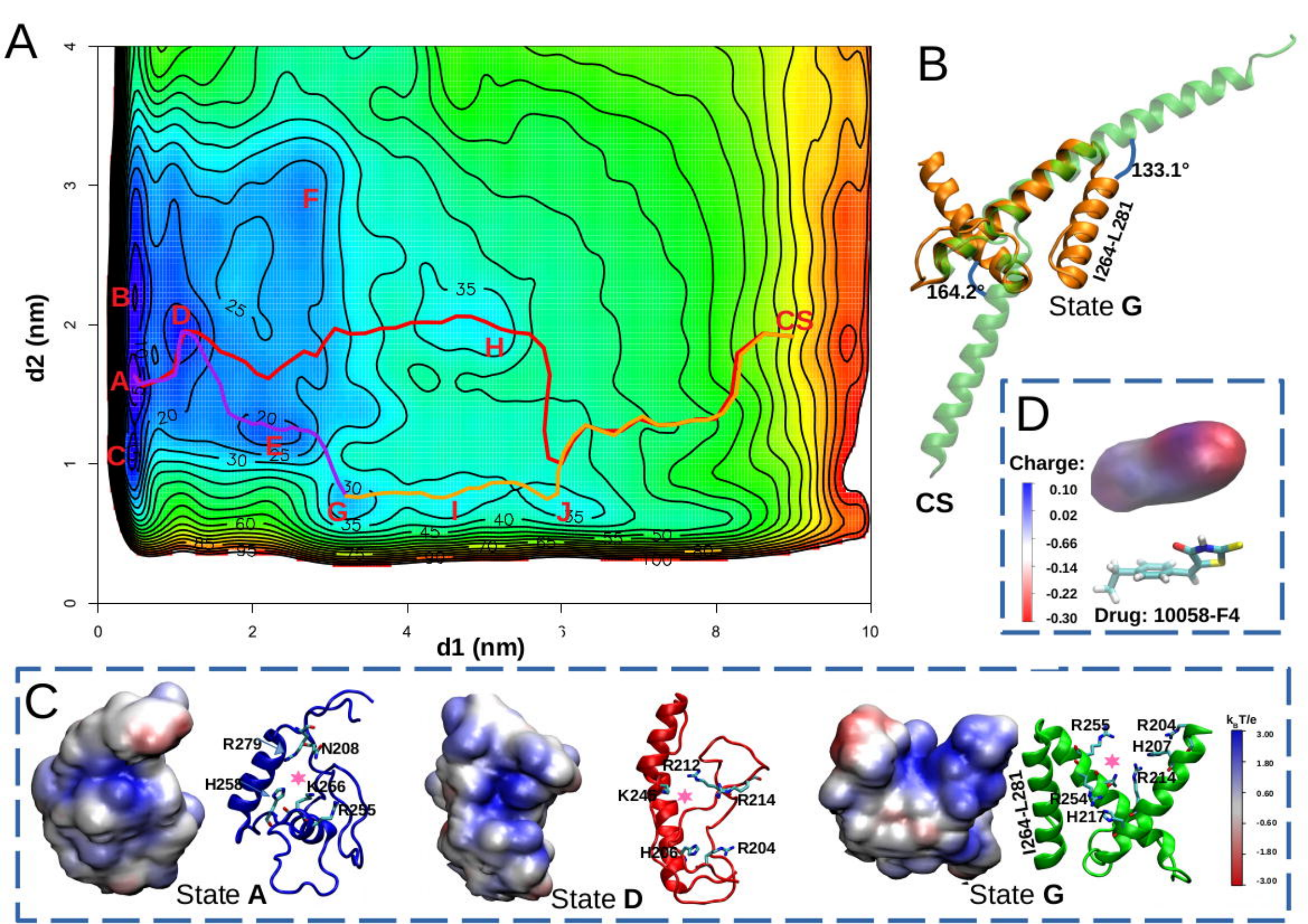

C
Zoom in
cluster0
Cα Asn271
cluster2


d2 (nm)
d1 (nm)
Free-energy (kJ/mol)
c0
c1
c2
c3
c4
0
1
2
3
4
5
6
7
8
9
10
0
10
20
30
40
50
60
70
80
90


d2 (nm)
d1 (nm)
0
1
2
3
4
5
6
7
A
B
C
D
E
F
G
H
I
J
K
L
I264-Q276
I264-A280
I264-R279
I264-L281
I264-G282


Free-energy (kJ/mol)
d1 (nm)
A-B-C
D
E
F
G
K
I
H
0
5
10
15
20
25
30
35
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
5
5,5
6


Free-energy (kJ/mol)
d2 (nm)
G
C
E
A
D
B
F
K
L
0
10
20
30
40
50
60
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
5
5,5
6


A
d2 (nm)
d1 (nm)
A
B
C
D
E
F
G
H
I
J
CS
B
133.1°
164.2°
I264-L281
State G
CS
D
Charge:
0.10
0.02
-0.66
-0.14
-0.22
-0.30
Drug: 10058-F4
C
R279
N208
H258
K266
R255
State A
R212
K240
R214
H206
R204
State D
R255
R204
H207
R214
I264-L281
R254
H217
State G
k_BT/e
3.00
1.80
0.60
-0.60
-1.80
-3.00