A step forward towards selective activation/inhibition of Bak, a pro-apoptotic member of the Bcl-2 protein family : Discovery of new prospective allosteric sites using Molecular Dynamics.

Guillem Vila-Julià[1,2], Juan J. Perez[2] and Jaime Rubio-Martinez[1,*]

[1]Department of Materials Science and Physical Chemistry, University of Barcelona and the Institut de Recerca en Quimica Teorica i Computacional (IQTCUB), Marti i Franques, 1, 08028, Barcelona, Spain

[2]Department of Chemical Engineering. Universitat Politecnica de Catalunya- Barcelona Tech. Av. Diagonal, 647. 08028 Barcelona, Spain

[*]Address correspondence to: Jaime Rubio-Martinez, Marti i Franques 1, E-08028 Barcelona, Spain.

Phone: (+34) 93 4039263; Fax: (+34) 93 4021231; E-mail: jaime.rubio@ub.edu

ABSTRACT

Bak is a pro-apoptotic protein, member of the Bcl-2 family that plays a key role in apoptosis, a programmed cell death mechanism of multicellular organisms. Its activation by death stimuli triggers the permeabilization of the mitochondrial outer membrane that represents a point of no return in the apoptotic pathway. This process is deregulated in many tumours where Bak is inactivated, whereas in other cases like in neurodegeneration, it exhibits an excessive response leading to disorders such as the Alzheimer disease. Members of the Bcl-2 family share a common 3D structure, exhibiting an extremely similar orthosteric binding site, place where both pro and anti-apoptotic proteins bind. This similarity raises a selectivity issue that hampers the identification of new drugs, capable of alter Bak activation in a selective manner. An alternative activation site triggered by antibodies has been recently identified, opening the opportunity to undertake new drug discovery studies. Despite this recent identification, an exhaustive study to identify cryptic pockets as prospective allosteric sites, has not been yet performed. Thus, the present study aims to characterize novel hotspots in the Bak structure. For this purpose, we have carried out extensive molecular dynamics simulations using three different Bak systems including Bak in its *apo* form, Bak in complex with its endogen activator Bim and an intermediate form, set up by removing Bim from the previous complex. The results reported in the present work shed some light on future docking studies on Bak through the identification of new prospective allosteric sites, not previously described in this protein.

INTRODUCTION

Apoptosis is a programmed cell death mechanism essential for tissue homeostasis, embryonic development or immunity that takes place in multicellular organisms. This process is tightly regulated,[1] so that its deregulation has a major impact on a wide variety of diseases. Thus, while an excessive response to apoptosis can lead to an increased ischemic condition or produce neurodegeneration, a defective response has a major impact on tumour development and autoimmune diseases.[2]

Apoptosis is accomplished through two different signalling pathways: the extrinsic and the intrinsic.[1] While in the former, the process is initiated by cell surface death receptors that recognize specific signals from other cells that activate initiator caspases, the latter is initiated when a cell experiences stress or damage. In this case, the process is regulated by members of the B-cell lymphoma-2 (Bcl-2) family of proteins through a mitochondrial outer membrane permeability (MOMP) mechanism,[3,4] facilitating the release of pro-apoptotic factors, such as cytochrome c, into the cytosol. Regulation of MOMP is critical, because once it is produced cells experience an irreversible bioenergetic collapse, even if the caspase activity is blocked.[5] The two pathways finally lead to the activation of effector caspases 3, 6 and 7 that carry out much of the proteolysis observed during the demolition phase of apoptosis.[1,6]

As mentioned above, members of the Bcl-2 family of proteins are involved in the regulation of the apoptosis intrinsic pathway. These proteins share one or more conserved regions, known as Bcl-2 homology (BH) domains that are necessary for their function, since their deletion via molecular cloning affects survival rates. They can be classified into pro-apoptotic and pro-survival proteins. In the former group, Bak, Bax and Bok exhibit three BH domains (BH1-BH3), whereas the rest of the members of this group, collectively known as BH3-only proteins including Bim, Bid, Puma, Bad, Bik, Bmf, Hrk and Noxa exhibit little sequence homology with the previous ones, apart from the BH3-domain that is necessary for mitochondrial apoptosis signalling. Specifically, some of the BH3-only proteins like Bim, Bid, and Puma to a lesser extent, are direct activators of the pro-apoptotic effector proteins, whereas the rest are sensitizers that indirectly activate Bak and Bax by sequestering the anti-apoptotic proteins, preventing a proper development of their function.[7] In contrast, members of pro-survival group including Bcl-2, Mcl-1, Bcl-xL, Bcl-w, A1, exhibit four BH domains (BH1-BH4) and prevent cell death by binding and inhibiting pro-apoptotic effector proteins.

The last decades have witnessed major advances in the knowledge of the mechanisms played by pro-apoptotic proteins. However, specific details of their regulation are not yet fully understood.[8,9] Thus, it is well established that relative concentrations of pro- and anti-apoptotic proteins are key for a cell to undergo apoptosis. Therefore, in healthy cells, under normal conditions, anti-apoptotic proteins sequester pro-apoptotic proteins to prevent apoptosis. However, after a cytotoxic stimulus, massive production of BH3-only proteins activates pro-apoptotic effectors either by binding directly to them or by sequestering anti-apoptotic proteins, producing the MOMP that results in apoptosis.[4,7]

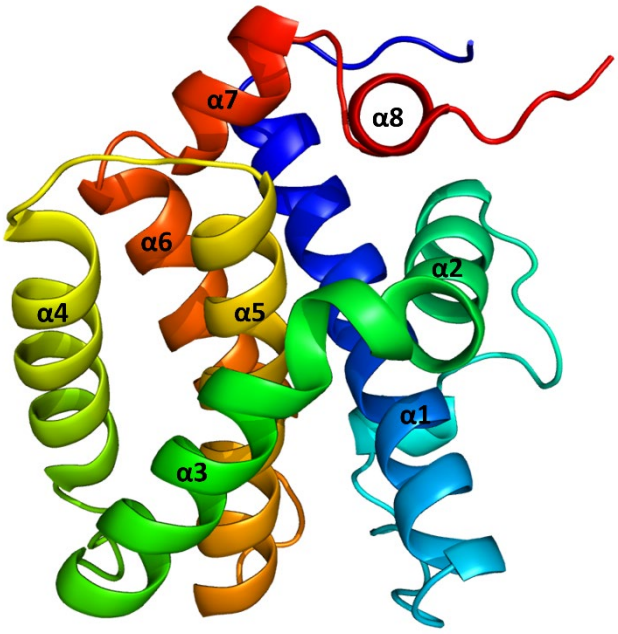

Figure 1. Crystal structure of Bak (PDB ID: 2YV6). The structure consists of a *core* domain (α2, α3, α4 and α5 helices) and *latch* domain (α6, α7 and α8 helices). Full-length pro-apoptotic Bak embodies a ninth helix (α9) utilized for mitochondrial outer membrane insertion, not observed in the crystal structure.

Members of the Bcl-2 family that exhibit multi-BH domains share a common all-α globular structure, consisting of nine α-helices (Figure 1). Helices α2 to α5 conform the *core* domain, whereas helices α6 to α8 conform the *latch* domain. Helix α9, which has not been crystallized for Bak is utilized for mitochondrial outer membrane insertion. The *core* and *latch* domains are directly involved in the activation process, as they are separated after activation by BH3-only proteins to allow oligomerization, as required for MOMP.[10] Upon activation, helix α1 is destabilized, detaching from helix α6 and favoring the separation between domains. Moreover, it has recently been observed that active pro-apoptotic Bak can trigger the activation of dormant Bak structures.[11] This autoactivation process cooperates with the direct-activation in BAK-mediated poration, since in both cases a destabilization of the α1 is observed. Helices from the *core* domain are packed in such a way to define a hydrophobic groove between helices α3, α4 and the α5 N-terminus, where BH3-proteins bind.[10] In addition, the groove serves as dimer interface for Bak and Bax to undertake oligomerization.[12,13] It has also been observed that this oligomerization process might be facilitated by the membrane.[14] For many years, the hydrophobic groove has been the only functional binding site known of these proteins. It is highly conserved among diverse members of the family, representing a major challenge to identify small ligands capable of activating or inactivating Bcl-2 family members selectively. This challenge has recently been overcome by targeting allosteric sites that reduce the capacity of the hydrophobic groove to bind BH3-proteins, as demonstrated in some anti-apoptotic members, such as Mcl-1.[15,16] In the case of Bak, it has been recently identified an alternative allosteric activation site triggered by antibodies, located at the N-terminus of the α1-α2 loop at the site of the recognition epitope (G51-D57).[17] The mechanism behind apoptosis activation through this allosteric site upon antibody binding still remains unclear. However, it has been observed that the binding of the antibody disrupts interatomic interactions in the monomeric structure of Bak, lowering the energy barrier and consequently, facilitating the transition to the symmetric BH3 dimer.[18] The relevance of these interatomic interactions is supported by the observation that adding an interaction between a modified form of the BH3-peptide Bim to α1 transformed this endogen activator into an inhibitor.[19] Thus, binding of the antibody triggers the separation of helix α1 from the *core* domain, an essential

step in Bak activation as previously observed in Bax.[17] Despite the mechanism between BH3-only proteins and the antibody is slightly different, there is enough experimental evidence to support a conserved activation mechanism.[20,21] Accordingly, further studies focusing on this allosteric site could shed some light into a complete understanding of the Bak activation mechanism and also serve as a guide to design novel allosteric regulators. A recent computational study on the pro-apoptotic Bax has provided information on the identification of prospective allosteric sites that can modulate protein activity.[22] This kind of studies can be helpful to cope with the challenges associated to finding selective ligands targeting the diverse members of the Bcl-2 family.

During the last decade, some studies of molecules capable of alter Bak activity have been developed. In addition to the antibody 7D10, which has been mentioned before, the antibody 3C10 showed an ability to trigger cytochrome c release via mitochondrial but not cytosolic Bax, which would mean that Bax translocation to the MOM are blocked by this antibody.[17] Moreover, Niu *et al.* discovered a small molecule capable to inhibit both pro-apoptotic Bak and Bax,[23] and in a recent study the authors identified a small molecule SJ572946, that mimics the role of the BH3 peptides and directly activates Bak selectively over Bax.[24] Furthermore, Park *et al.* identified BKA-073, a small compound capable of activating Bak with a potent antitumor activity against lung cancer.[25]

Despite the number of structure-activity relationships studies and the small molecules identified on Bak, a detailed study on the already known, as well as novel prospective allosteric sites of the whole protein have not yet been reported. Accordingly, the present study describes the results of a modelling study addressing the characterization of already known activation sites of Bak and the identification of hotspots, as novel prospective allosteric sites. For this purpose, we have carried out a detailed study using Molecular Dynamics (MD) simulations using the structures of three different Bak complexes: Bak in its *apo* form (*apo*Bak), Bak bound to the BH3-peptide Bim (DMRPEIWIAQELRRIGDEFNAYYARR) (Bak/Bim) and an intermediate molecular system constructed by removing the BH3-Bim peptide (intBak) from the Bak/Bim structure. 300ns MD trajectories were run for each of the three systems in triplicate and using two different sampling techniques. Extensive MD calculations provide the opportunity to sample the conformational space of molecules beyond the details provided by actual crystallographic structures, permitting to identify hotspots as prospective novel allosteric sites. This study can be considered as the necessary starting point for identifying novel ligands capable to activate or inhibit Bak, targeting not only to its orthosteric site but also, other allosteric sites.

METHODS

Preparation of apoBak, intBak and Bak/Bim Complexes

The three-dimensional structure of the Bak/Bim complex was retrieved in a previous work recently published, where the 5VWZ Protein Data Bank (PDB) accession code was used as starting point of the preparation procedure.[26] The structure of intBak was prepared by removing the BH3-peptide Bim from the Bak/Bim complex. The apoBak structure was retrieved from the PDB accession code 2YV6[27]. Residues of the gap in the loop between helices α1 and α2 were obtained from the Bak/Bim complex and missing atoms from these residues were completed using the *LEaP* module of the AMBER18 software.[28]

Molecular dynamics of Bak and the Bak/X Complexes

300ns MD trajectories were run for each of the three systems in triplicate and using two different sampling techniques. All calculations of the present work were performed using the PMEMD

(Particle Mesh Ewald Molecular Dynamics) code of the AMBER18 software in its CUDA version[29] using the AMBER ff14SB force field.[30,31]

As explained elsewhere,[26] systems were prepared for molecular dynamics (MD) simulations according to the following protocol. First, each protein was soaked in a cubic box with equilibrated TIP3P[32] water molecules and certain water molecules were substituted with Na^+ or Cl^- ions to neutralize de system using the electrostatic potential before solvation as a guide. Each cubic periodic box was constructed imposing a minimal distance of 16Å between the protein and the edges of the box and water molecules near to 1Å to the protein were also removed. Subsequently, energy minimization was carried out stepwisely: first, the positions of the water molecules and ions were optimized using the steepest descent (SD) algorithm[33] up to 5000 cycles of minimization, keeping the rest of the system fixed; in a second step, the modelled residues were relaxed in two stages, each one of 5000 cycles of SD, keeping fixed the backbone positions of these modelled residues and using a decreasing force constant of 5.0 and 0.1 kcal/Å; in the last step, the minimization of the whole system was carried out without any restrictions using 10000 cycles of SD method. After minimization, the systems were heated in steps of 30 K every 20 ps, using a force constant of 1.0 kcal/mol·Å to maintain all backbone atoms constrained. The heating process was performed under the canonical (NVT) ensemble. After the heating process, a 200 ps trajectory at constant pressure (NPT ensemble) was performed for density equilibration using the Berendsen barostat to control and maintain the pressure at 1atm.[34] The final structures were used as starting points for the production MD simulations in the NVT ensemble. Trajectories were performed at 300 K using the Langevin thermostat[35] to keep the temperature constant with a collision frequency of 3.0 ps^{-1}. All bonds involving hydrogen atoms were fixed using the SHAKE algorithm,[36] allowing us to use a time step of 2 fs for all the simulations. Non-bonded interactions were truncated using a cutoff of 9 Å, and long-range interactions were treated with the particle-mesh Ewald summation method[37] with a grid spacing in the direct lattice of about 1 Å, a fourth-order B-spline interpolation for the gridded charge array and a direct sum tolerances of 10^{-5}. For each system, we performed three conventional MD (cMD) and three Gaussian Accelerated Molecular Dynamics (GaMD)[38] of 300 ns length to enhance conformational sampling. The GaMD method consists in adding a harmonic boost potential to smooth the potential energy surface of the system and decrease the energy barriers to accelerate the transitions between different low-energy states.[39] Sampling was carried out using multiple molecular dynamics trajectories with different initial velocities, which results in a more thorough exploration of the conformational space, reducing chances the system gets trapped in a local minima.[40]

Root-Mean Square Deviation and Root-Mean Square Fluctuation

To assess the structural stability of the systems along the MD simulation, Root-Mean Square Deviation (RMSD) was computed along the simulation time using the *cpptraj* module from Amber18.[41] RMSD was computed using the respective initial structure obtained from LEaP as reference. Each of the frames of all trajectories were reoriented over the residues from helices α1 to α6, using the α carbons (Cα) of all the residues. For each of the three complexes: apoBak, intBak and Bak/Bim, the Root-Mean Square Fluctuations (RMSF) for all Bak residues was computed to monitor their fluctuations during the MD simulations using also *cpptraj* of the Amber18 program, providing information of the local conformational flexibility of each residue.[41]

Block Average and Autocorrelation

The block average method is a simple approach to determine , correlation time and a good estimation of variance for a non-independent data. The general idea behind is to divide our data into different blocks and analyse the trends on the average of each block. The more blocks we take, the lower the standard deviation of the averaged value. However, if the blocks are too small, average values will highly differ between them and σ will increase. Thus, the blocks must be long enough to consider that each estimation of the desired property is independent from the other. The system relaxation time will determine the minimum length of each block. Thus, we divided all the simulation trajectory (300 ns) in blocks of different length, plot the block standard error and look for the point at which the standard deviation *levels off*.

Autocorrelation is a statistical tool that measures the degree of similarity between a given time series and a lagged version of itself over successive time intervals. The autocorrelation function (ACF) measures the correlation between a current value of the variable and its past values at different lag times. A positive autocorrelation coefficient indicates a positive relationship between the current value and past values, while a negative autocorrelation coefficient indicates a negative relationship. From it, the autocorrelation time can be estimated as the time lag at which the autocorrelation function drops below a certain threshold value, which is typically taken to be 0.1 or 0.2. This lag represents the characteristic time scale of the correlations in a time series.

Principal Component Analysis

Principal Component Analysis (PCA) was used to determine and analyse the principal structural variations of the diverse systems studied. PCA is a multivariate statistical technique which captures the essential features of protein dynamics, using the least number of dimensions. PCA is used to reduce the dimensionality of the dataset that is needed to describe protein dynamics. This reduction is performed through a decomposition process that describes motions from the largest to the smallest spatial scales, in order to obtain lower-dimensional data while preserving as much variation of the data as possible. In this method, a covariance matrix constructed using the atomic coordinates of the alpha carbons (Cα) of every residue is diagonalised to produce a set of eigenvectors or Principal Components ($PC^{(i)}$, i=1, 3N, where N is the number of residues of the protein), as well as their corresponding eigenvalues $\lambda^{(i)}$. Thus, this covariance matrix is a 3N x 3N symmetric matrix. PCA transforms correlated variables into uncorrelated variables and once they are rank ordered, the first principal modes or eigenvectors can be used to characterize large-scale protein motions. In other words, these first modes are enough to define the "essential" space or motions of the protein. The percent contribution of the i^{th} principal component $PC^{(i)}$, to the structural variance in the dataset is given by:

$$c\% = 100 * \frac{\lambda^{(i)}}{\sum_{i=1}^{3N} \lambda^{(i)}} \quad (1)$$

where the summation is performed over all 3N components.[42] Moreover, this analysis can be used to obtain projections of the MD trajectory snapshots that provide information about the extension of the sampling process.

Free-energy landscape was computed by projecting it onto the first two principal components PC1 and PC2. Thus, a two-dimensional histogram was constructed with a bin width and height of 1 Å

(ΔPC1 and ΔPC2) and each frame k with coordinates (PC1,PC2) was assigned to its corresponding bin i=(PC1,PC2). In this representation, differences of the free energy are given by the following equation:

$$\Delta G = -k_b T \ln(n_i/n_{max}) \quad (2)$$

where n_i n_{max} is referring to the number of configurations located in a bin i (defined by increments ΔPC1 and ΔPC2), and n_{max} is the number of configurations in the most populated bin in order to assign a zero value to the most populated bin. In GaMD, Maclaurin series expansion was applied as a reweighting algorithm.[43]

Cluster Analysis

Identification of the different structural features from the three complexes used, apoBAK, intBAK and Bak/Bim was performed by grouping similar structures in each complex into 10 different clusters using the average linkage algorithm,[44] as implemented in the *cpptraj* module of AMBER18. For this process, the RMSD of the Cα in helices α3, α4 and the N-terminus of helix α5 were used as distance. The rationale for using the residues from these helices was their involvement in the direct interaction between Bak and the BH3 peptides and to analyse more accurately the changes from these three complexes in the orthosteric binding site.

FTMap

In order to identify the protein hotspots in our systems, we used the FTMap server.[45] FTMap is a computational mapping server used to identify binding hot spots in macromolecules. The algorithm consists in distributing a huge amount of positions of small organic molecules of varying size, shape and polarity (probes), and scoring the poses from these probes using a detailed ligand binding free energy. The probes used in FTMap are acetaldehyde, acetamide, acetone, acetonitrile, benzaldehyde, benzene, cyclohexane, dimethyl ether, ethane, ethanol, isobutanol, isopropanol, methylamine, N,N-dimerthylformamide, phenol and urea). Binding hot spots are those regions where clusters of multiple probe types bind.

Receptors used in this software are those obtained from the cluster analysis carried out on the 3 cMD and the 3 GaMD for all 3 different complexes: apoBak, Bak/Bim and intBak. Structures obtained from the cluster analysis were the first five most represented for each system and sampling method (see Table S1).

D3Pockets

D3Pockets is a web server used to explore different dynamic properties of a protein pocket from a molecular dynamics simulation trajectory or a single structure.[46] This software studies pocket stability, continuity and correlation. All three protein dynamic features are calculated using a set of trajectory points from all MD simulations. The working procedure consist in three basic steps: first, the potential pockets of the different conformations included in the trajectory are detected. Second, a grid of points is created for each detected pocket on each conformation. Finally, the pocket dynamic properties are calculated by analysing the number of times a particular grid point appears along the MD trajectory.

The stability of the pocket (P_S). We define the stability of each of the grid points that define the pocket S_i as :

$$S_i = \frac{n}{N} \quad (3)$$

where *n* is the number of conformations that have the *i*th grid and *N* is the total number of conformations in the trajectory. If m is the number of grid points in the pocket, the stability of the pocket (P_S) is defined as an array containing the stability of all the grid points defining the pocket:

$$P_S = \{S_1, S_2, \dots, S_i, \dots, S_m\} \quad (4)$$

D3Pockets colors the grid points that compose a pocket. Thus, red points are the most frequented points during the MD and the blue ones lest frequented. Then, the red regions of one pocket are more stable than the other regions.

Pocket Continuity (P_C) is calculated by tracing the change of the pocket in terms of appearance, disappearance, merging and volume change. This P_C is defined by:

$$P_C = \{P_i | P_i \cap P_{ref} > 0\}, i = 1, 2, \dots, n \quad (5)$$

where P_{ref} is the first conformation of the studied pocket, P_i is the *i*th pocket that is spatially overlapped with P_{ref}, *n* is the number of all pockets from the whole MD simulations, and P_C is the ensembles of pockets P_i.

Pocket Correlation. First, all potential binding pockets that appear in a MD trajectory are clustered, based on residues, generating sets of protein conformations corresponding to each cluster (C_i). Next, for each conformation that belong to cluster i, its volume is calculated (V_i). Finally, the coexistence and correlation matrices are calculated with the following equations:

$$C_{i,j} = C_i \cap C_j \quad (6)$$

$$\rho_{i,j} = \frac{cov(V_i, V_j)}{\sigma_{V_i} \sigma_{V_j}} \quad (7)$$

where C_i and C_j are the conformation sets of the protein corresponding to the *i*th and *j*th cluster pockets, respectively; V_i and V_j are the volume sets of the *i*th and *j*th cluster pockets in the corresponding conformations; $cov(V_i, V_j)$ is the covariance of V_i and V_j; and σ is the variance of V. A positive correlation coefficient means a positive correlation between pockets (maximum +1) while a negative coefficient stands for a negative correlation between pockets (minimum -1). As the volume is used for correlation, a positive volume correlation between pockets i and j means that when pocket i gets bigger in its volume, volume of pocket j increases as well. On the contrary, negative correlation means that when pocket i increase its volume, pocket j gets smaller.

RESULTS AND DISCUSSION

The great challenge for the discovery of novel selective small ligands targeting Bcl-2 proteins arises from the similarity of the orthosteric site among members of the family, including both pro-apoptotic and anti-apoptotic members. Based on previous results on Mcl-1,[15,16] the design of allosteric ligands can be considered as an appropriate strategy to circumvent these difficulties. The present study was designed to identify cryptic sites in Bak that can be subsequently investigated as prospective allosteric sites, capable to modulate Bak activity. For this purpose, we performed MD calculations both using classical and accelerated sampling methodologies on three Bak complexes: apoBak, intBak and Bak/Bim. The results were analysed using different computational tools and are reported below.

Principal Component Analysis (PCA)

To carry out the PCA analysis, different variants were performed by changing the residues used to construct the covariance matrix and contributions of the first 10 PCs for each variant are shown in Table S2. Firstly, the covariance matrix was constructed using the alpha carbons (Cα) of Bak residues (from Ala22 to Leu183) disregarding both termini. The first Principal Component (PC1) recovers about 19% of protein fluctuations and jointly with the second, the two first PCs describe about 31.5% of protein fluctuations. Next, we tested how removing the α1-α2 loop affects the PCs contributions. Thus, the covariance matrix was constructed considering all the Cα of the Bak residues except those from the terminals and those from the loop between helix α1 and α2 (Ala22-Gln47 and Ser68-Leu183). In this case, the two first PCs describe more than 40% of the fluctuations of the system.

However, we finally decided to focus our attention in the orthosteric binding site. For this purpose, a covariance matrix was constructed using only those Cα of the residues involved in the interaction with BH3-peptides, including helices α3, α4 and α5 (Ile85-Leu100, Tyr110-Ser121 and Asn124-Phe134). Projection of each MD trajectory snapshots onto the respective two principal components for each system and type of sampling either classical or accelerated are shown in Supplementary Figures S1 and S2, respectively. Inspection of these Figures point to a wider sampling of the intBak systems. It can also be observed that the sampling of the three replicas is not necessarily the same, in support of the observation that multiple MD sampling may be more efficient than a single trajectory.[40] Finally, these Figures also show that GaMD explores different sections of the conformational space than cMD.[47] In this case, the first two PCs describe about 58% of protein fluctuations. Figure 2A-B displays the projection of trajectory snapshots onto the first two PCs for each system for both sampling modalities cMD and GaMD, respectively. These plots are representations of the free energy of the system. Accordingly, in this case, the lowest energy minimum is computed according to the following procedure. For each plot and system, a mesh of bins is constructed with specific ΔPC1 and ΔPC2 dimensions and the bin occupancy frequency computed. The most populated bin will exhibit the lowest free energy. Accordingly, a “0” energy value is assigned (dark blue), and the non populated bin an energy value of 12 in cMD and 18 in GaMD (red). Inspection of Figure 2 shows that apoBak and Bak/Bim are in different positions, and intBak “moves” from Bak/Bim to apoBak during the sampling, exploring all the conformational space between both. Thus, using these different conformations of Bak in multiple cMD and GaMD allows us to explore a large variability of Bak structures, making it possible to better decipher the potential Bak hotspots.

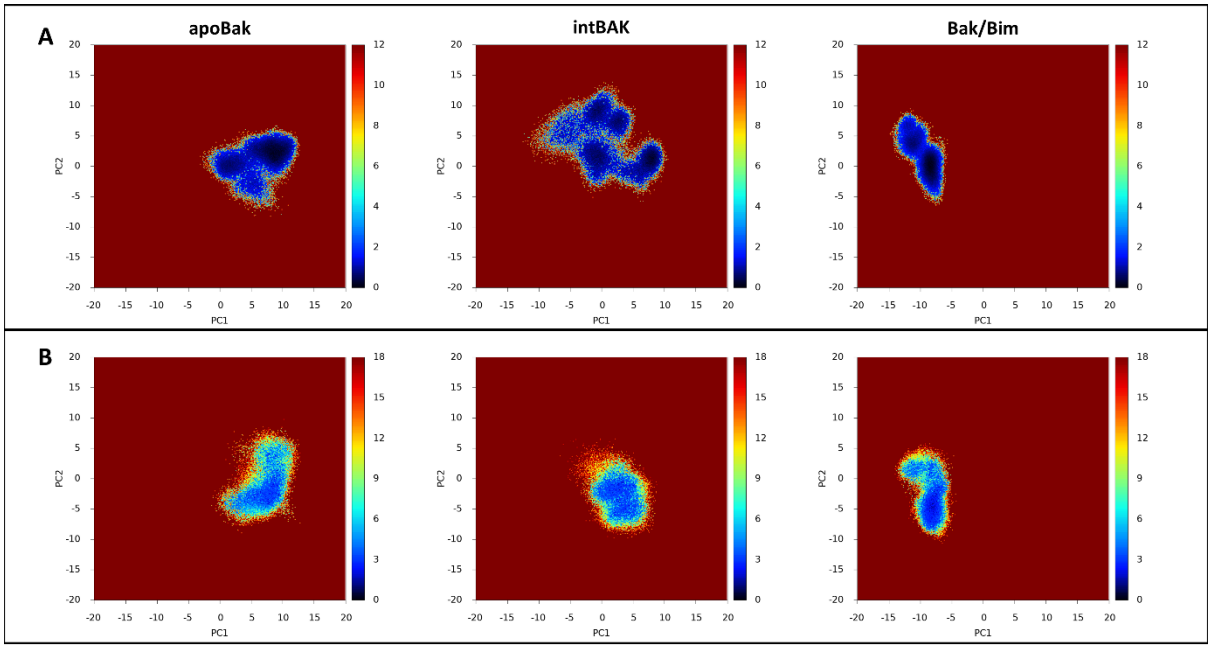

Figure 2. Conformational space sampled for the three systems studied and for the two sampling methods used cMD (A) and in GaMD (B) shown pictorially by their projections onto the first two principal components (PC1 and PC2), respectively. The free energy of the systems is drawn according to a colour code so that the lowest values are dark blue and the largest dark red.

Block Average Analysis – Autocorrelation

Block Average method has been used with the average RMSD values (shown in the RMSD-RMSF section) for each Bak system and sampling method. Although all systems have different asymptote and thus standard error, the correlation time is always smaller than 50ns , far from the total length of simulation (Supplementary Figure S3A). In this representation, only the block length between 1 to 100ns is shown.

On the other hand, autocorrelation between the different RMSD values for each Bak system and sampling method has been analysed. As it can be seen in Supplementary Figure S3B, the correlation time obtained for all the systems is similar to those obtained by the block average method, supporting the fact that the correlation time is rather less than the total time.

Cluster Analysis

Structures of the simulations of the three systems using the two sampling methods were subjected to cluster analysis as described in the method section. The whole trajectories from all systems and sampling methods have been used for the clustering analysis to include all the accessible states during the full length of the MD. We have clustered 90000 frames for each system and sampling method, which corresponds to extracting one structure every 10ps. Moreover, sieve option equals to 3 was used in the clustering process, which means that it uses only 30000 frames and adds the other frames to the closest cluster in each case. As can be seen in Table S1, considering the first five clusters in all three systems and both sampling methods, it is accounted about 95% of all the structures explored. Thus, for each system and sampling method, the centroids of the first five clusters were chosen as representatives to elucidate, using FTMap, the potential hotspots in Bak structure.

Identification of Cryptic Pockets - FTMap

The computational mapping server FTMap[45] was used to identify cryptic pockets in Bak that were subsequently characterized by their potential to serve as allosteric sites. We carried out the

analysis on the representative structures of the first five clusters for the three systems: apoBak, Bak/Bim and intBak and for both sampling methodologies: cMD and GaMD. Figure 3A, and Supplementary Video 1 show pictorially a compendium of the diverse hotspots found in the present study and Figures 3B-3D shows pictorially the detailed results obtained for each of the three systems studied in the present work. Supplementary Figure S4 shows the rear view of all the systems studied. Moreover, as a complementary information (Supplementary Figure S5), we also used FTMap to identify the pockets for different crystallographic structures, with PDB accession codes 5VWZ[19] (Bim-BH3 bound), 2M5B[20] (Bid-BH3 bound), 5FMI[48] (apo-Bak) and 5VX1[19] (apo-Bak). Analysis of the results show that both the orthosteric binding site as well as the allosteric site activated by antibodies[17] are identified in the process. In addition, four novel cryptic pockets are identified.

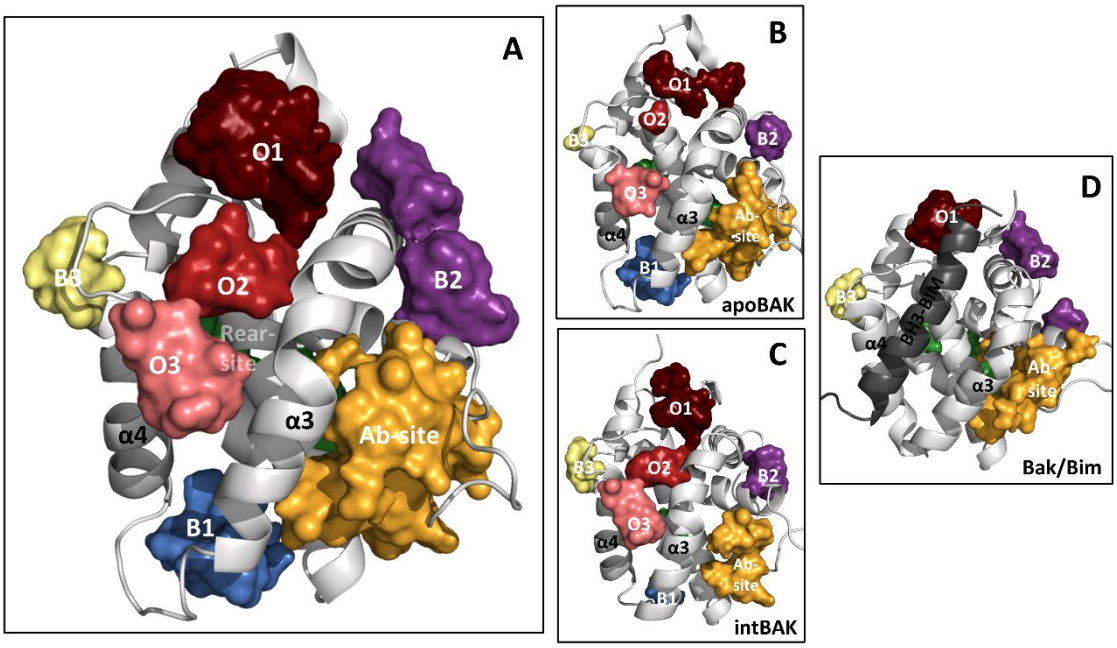

Figure 3. Bak hotspots identified with FTMap software. (A) Hotspots from all three systems studied combined. (B) Hotspots from apoBak complex. (C) Hotspots from intBak complex. (D) Hotspots from Bak/Bim complex. Colour code: the orthosteric binding site is shown as three neighbouring hotspots labelled O1-O3, respectively and shown in a scale of reds; the Ab-site is shown in gold; the rear-site in green; the B1 in blue; B2 in purple and B3 in yellow.

As described above, the orthosteric binding site is a hydrophobic groove formed by helices α3 and α4 together with the N-terminus of helix α5, and is the site where BH3-peptides bind and activate Bak, triggering a series of conformational changes needed to activate Bak. Since the orthosteric binding pocket is large, it is described here by means of three different sites (O1, O2 and O3), shown in different tones of red colour in Figure 3. On the other hand, the activation site triggered by antibodies, from now on termed Ab-site, is shown in Figure 3 with amber colour. Despite the site is large and exposed to the solvent, the specific region involved in the antibody binding involves the epitope Gly51 to Pro55 (GVAAP) at the N-terminus of the loop between helices α1 and α2. Thus, as this epitope is very close to our site, we decided to term this site the Ab-site. This site is formed by the C-terminal of helix α1, the N-terminal of the loop between helices α1 and α2, and the C-terminal of helices α2, α3 and α5. Residues in this pocket include Arg42 to Thr62; Asp83 to Gln94 and Arg137 to Gln144. Moreover, a recent study showed the importance of these residues in Bak activation as well as Met60, located in the middle of the α1-α2 loop, which could have a role in stabilising the inactive, monomeric form of Bak.[18] Present results support these experimental studies since the epitope

GVAAP together with Met60 are part of the cryptic pocket identified. In addition to the orthosteric binding site and the Ab-site, four more cryptic pockets were identified.

Rear-site

This site is located on the interface between helices α1 and α6, at the antipodes of the orthosteric binding site of the protein surface and shown in green in Figure 3. The site was previously described in Bax as the rear activation site, deemed necessary for the protein to trigger the release of helix α9 from its hydrophobic groove to allow its translocation to the mitochondrial outer membrane (MOM).[49–51] Since Bak is constitutively inserted in the MOM, the release of helix α9 from the core-domain is not required as a first activation step. However, as has been already pointed the role of this rear site is not only to trigger the release of helix α9, but also to increase the accessibility to the hydrophobic groove of BH3-only proteins.[52]

Binding Site 1 (B1)

This cryptic pocket represented in blue in Figure 3 is located beneath the hydrophobic groove. It is located between the C-terminal side of helices α3 and α5 as well as the N-terminal of helices α4 and α6 with their respective loops comprising residues His99 to Ala107; Leu138 to Phe157. In Supplementary Figure S6, three different orientations of this site are represented. This is a hydrophobic pocket since residues pointing to the interior of the site are mainly hydrophobic (Leu97, Leu100, Phe111, Ile114, Phe134, Val142, Leu147, Phe150 and Val154).

The site has previously been described regarding Bax activation. On the one hand, Bax activators BAM7 and BTSA1 bind this site,[53,54] and on the other, the vMIA (viral mitochondrial inhibitor of apoptosis) binds to the outermost part of it to prevent apoptosis. A NMR structure of the latter is also vailable[55] (2LR1 PDB accession number). This findings suggested that the activation mechanism should be considered as a cooperative process in which multiple sites are involved.[22] These results, permit us to speculate that the B1 site can play a similar role in Bak, since the structural domains defining the site are similar in both Bak and Bax. However, to support the ability of the B1 site in modulating Bak activity further experiments are needed.

Binding Site 2 (B2)

B2 site is located at the C-terminus side of the loop between helices α1 and α2 and helix α2, close to the Ab-site, shown in purple in Figure 3. Residues from this site comprise those from the C-terminal of α1 (Ser37, Phe40 and Gln44) as well as those from C-terminal of the α1-α2 loop and the N-terminal of α2 (Pro64 to Ile80). B2 is a solvent exposed site and its closeness to both, the Ab-site and the α1-α2 loop suggests its relevance. Despite the antibody mentioned above that activates Bak interacts with only a few residues on the Ab-site, ligands binding to B2 could hinder the movement of helix α1 during activation modulating Bak activation mechanism.

Binding Site 3 (B3)

B3 is located close between the α4-α5 loop and the C-terminal of helix α6 and the N-terminal of helix α7 and represented in pale yellow in Figure 3. It comprises residues between Ala115 to Ile123 and from Phe161 to Ile171. The site, called the BH groove was previously identified as the binding site used by the protein kinase C inhibitor chelerythrine, to the pro-survival member of the Bcl-2 family, Bcl-x_L.[56] However, binding affinity is small and its role as a Bcl-x_L inhibitor poor.[57] Moreover, a recent study has revealed the groove between α4, α6 and α7, where this B3 site is located, as an alternative Bak binding site for specific BH3-proteins, such as BMF and HRK.[58] However, the mechanism how these

BH3-proteins bind to this site still remains unclear, and their ability to activate Bak might be exerted by a prior binding to the orthosteric binding site.

Detailed comparison of the cryptic pockets found individually in the three systems studied is shown in Figures 3B-3D. As it can be seen, the results for apoBak (Figure 3B) and intBak (Figure 3C) are qualitatively the same, so that all the pockets are easily identified. These results support that MD simulations were long enough to permit the necessary conformational rearrangements to bring intBak back to the apoBak conformation. In contrast, comparison to the results obtained for the Bak/Bim system (Figure 3D) shows important differences. Thus, in this case, cryptic pockets like the Ab-site, the Rear-site, B2 and B3 are predicted in similar fashion, but only the O1 site from the orthosteric site due to the BH3 Bim occupation. However, what is more relevant is that the hotspot B1 does not show up, suggesting that binding of a BH3-peptide in the hydrophobic groove produces a series of conformational changes that close the B1 pocket. These results suggest that the B1 site a prospective allosteric site since it modulates the orthosteric site geometry.

Root-Mean-Square Deviation and Fluctuations (RMSD/RMSF)

RMSD was computed for all the trajectories to establish if the MD simulations performed were stable in structural terms. Superimpositions of the residues from all the helices (α1-α6) have been included. Average values from all three replicas of each system are represented in Supplementary Figure S7. As it can be observed, after a short period of stabilisation, all systems are stable in structural terms. As expected, intBak systems exhibit a higher RMSD as they are rearranging to fill the space left by the BH3-peptide.

The RMSF was computed to assess the fluctuations of different regions of Bak structure in all three different complexes. As expected, due to the starting structure, fluctuations of helix α3 in the intBak complex are considerably higher when compared to the apoBak structure and the Bak/Bim complex (Figure 4). Another interesting outcome is that fluctuation of the residues conforming the pocket B1 are significantly higher in the apoBak structure than in Bak/Bim (Figure 4). Accordingly, binding of BH3-Bim not only rigidizes helices α3, α4 and α5 that surround the pocket, but also the loop between helices α3 and α4, making the access to the B1 pocket difficult for possible molecules targeting this site.

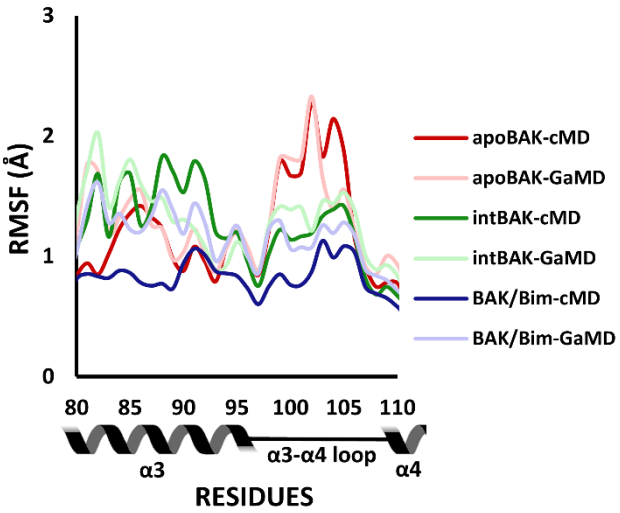

Figure 4. Backbone atoms RMSF (Å) of the segment including helices α3 and α4 and the loop connecting them, computed as an average from the three cMD and three GaMD trajectories, respectively for the apoBak, intBak and Bak/Bim complexes. The average RMSF for apoBak complexes in cMD (red) and GaMD (light red), for intBak

complexes in cMD (green) and GaMD (light green) and for Bak/Bim complexes in cMD (blue) and GaMD (light blue) are shown respectively.

D3Pockets

D3Pockets[46] software was used to analyse protein pockets dynamics. This software assesses the robustness of the pockets identified by analysing their continuity throughout the MD simulation. In addition, it also determines if there is correlation between them.

Continuity

The continuity analysis of the different pockets shows that the volume of the Ab-site changes constantly in the three systems studied during the sampling process. In the case of the intBak, movements are more dramatic since the starting structure is not equilibrated, as can be easily observed in the results of the PCA. Thus, at the beginning of the MD simulations helix α3 is close to the C-terminus of helix α1 and far from helix α4, and by the end of the simulation, helix α3 has moved towards the helix α4, reducing the volume of the orthosteric pocket and increasing that of the Ab-site (Figure 5).

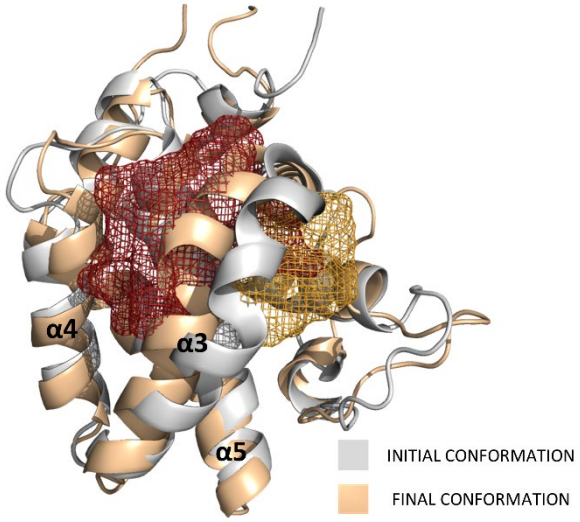

Figure 5. Continuity of the pockets in intBak systems. The initial and final conformations of a single trajectory are represented. A decrease in the volume of the orthosteric (red) site and an increase in the volume of the Ab-site (gold) are observed between the initial and the final conformations.

Stability

When comparing the robustness of the predicted pockets in the apoBak system, it can be easily seen that the Ab-site is stable in all the MD simulations. Similarly, to the orthosteric binding site, although not as stable as the former. When comparing to intBak complexes, similar results can be observed. In these complexes, as the orthosteric binding site is in a more open conformation in the beginning, with the α3 further from the α4, the surroundings of the α3 helix are more unstable, due to the movement of this helix during the MD simulation. On the other hand, in the case of the Ab-site, although it is also stable, its stability is lower in comparison with apoBak systems, due to the proximity of the α3. When analysing the stability of the pockets in Bak/Bim systems, lower stability can be observed in the Ab-site. However, the highest stability is found in the surroundings of the epitope GVAAP. Also, in Bak/Bim systems, low stability can be observed in pocket B2 and the Rear-site. These results are illustrated in Figure 6.

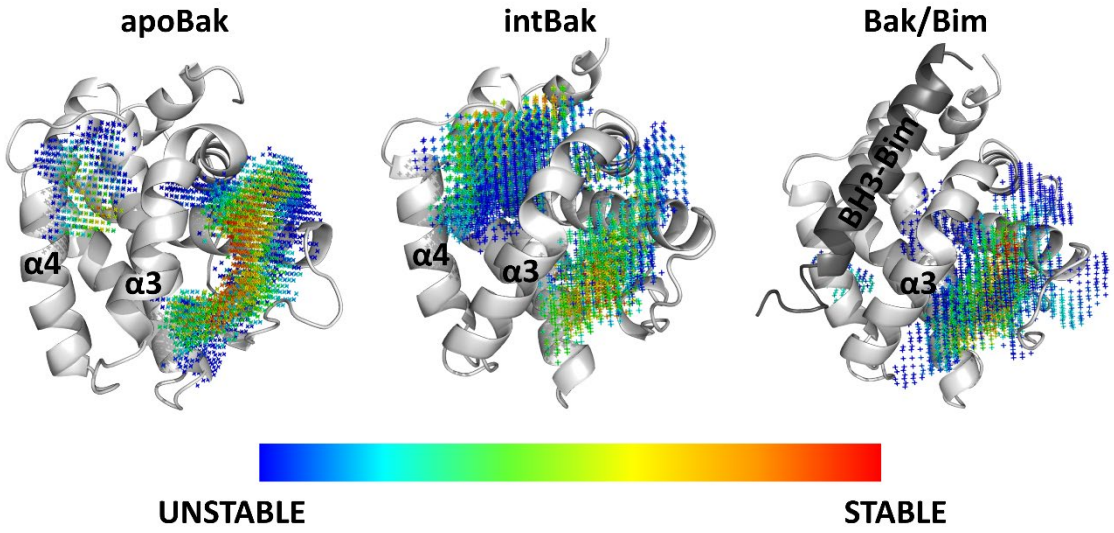

Figure 6. Stability of the Bak hotspots identified by D3Pockets software in apoBak, intBak and Bak/Bim systems. The BH3-peptide Bim is represented in dark grey.

Correlation

D3Pockets is capable to calculate a correlation between cryptic pockets appearing during the MD trajectory. These cryptic pockets are clustered based on the residues that conform them, and a correlation is calculated. This parameter allows us to determine how a conformational change in a particular site induces conformational changes in other.

The results show multiple correlations between cryptic pockets in the different systems. In the apoBak system, we found a negative correlation (-0.71) between the orthosteric binding site and the Ab-site (Figure 7A), which means that when the orthosteric binding site increases its volume, the Ab-site decreases. Helix α3 is the key structural element managing it. Thus, when helix α3 is in a more open conformation, away from helix α4, the orthosteric binding site is larger, while the Ab-site reduces its size. On the other hand, when helix α3 is closer to α4, the Ab-site is larger and the orthosteric binding site is smaller. This correlation can be observed also in the intBak (-0.62) and in Bak/Bim (-0.67) systems, despite in Bak/Bim this correlation concerns the O1 subsite, since the orthosteric binding site is occupied by the BH3 peptide.

There is also a correlation between B1 and the Ab-site in apoBak. In this case, the correlation is positive (0.54) and is tightly linked to the correlation previously described between the orthosteric site and Ab-site. Thus, when the helix α3 moves closer to the helix α4, the latter slightly moves its N-terminus, opening the B1 cryptic site, as can be seen in Figure 7B.

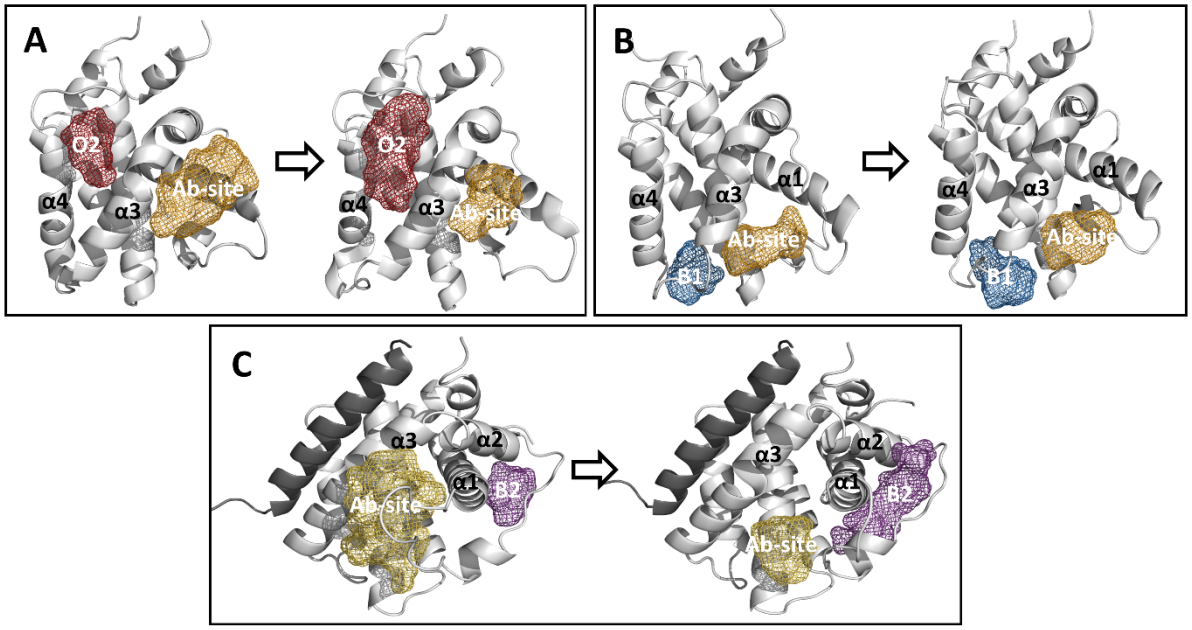

Figure 7. Correlation between the different Bak hotspots. (A) Negative correlation between the orthosteric binding site (red) and Ab-site (gold). (B) Positive correlation between B1 (blue) and Ab-site (gold) hotspots. (C) Negative Correlation between B2 (gold) and B3 (purple) hotspots.

Finally, another correlation was identified in the Bak/Bim system between the Ab-site and the B2 pocket. The correlation is in this case negative (-0.58). This finding together with the proximity between these two pockets gives relevance to B2 and its potential role in modulating Bak structure and activity. Thus, a volume increase of the B2 site may affect the accessibility to the Ab-site by reducing its volume (Figure 7C).

CONCLUSIONS

The present work reports the results of a computational study addressing the characterization of Bak structure, using multiple lengthy MD simulations. Specifically, calculations were performed on three different complexes: Bak in its apo form (apoBak), Bak bound to the BH3-peptide Bim (DMRPEIWIAQELRRIGDEFNAYYARR) (Bak/Bim) and an intermediate form, obtained from the Bak/Bim structure by removing the BH3-Bim peptide (intBak), aimed at identifying cryptic pockets that could be considered prospective allosteric sites of the protein. For this purpose, 300ns MD trajectories were run for each of the three systems and using two different sampling techniques in triplicate. Analysis of the trajectories permitted to identify cryptic pockets as prospective allosteric sites of the protein that can be targeted to design more small selective molecules within the members of the Bcl-2 family.

Structures collected at regular intervals from the trajectories were classified using cluster analysis. The representative structures from the diverse clusters were subsequently screened for hotspots using FTMap. This procedure identified known Bak binding sites, such as the orthosteric binding site and the site triggered by antibodies,[17] but also of novel cryptic pockets. These include: the Rear-site, a solvent exposed site, located on Bak rear side, in the interface between helices α1 and α6; B1, a small, hydrophobic site located between the C-terminus of helices α3 and α5 and the N-termini of helices α4 and α6, including the respective loops between helices α3-α4 and helices α5-α6; B2, also a solvent exposed site, located between the N-terminus of helix α2 and the C-terminus of the loop between helices α1-α2, also near to helix α1; and B3, located behind the hydrophobic groove, between the α4-α5 loop and the C-terminus of helix α6 and the N-terminus of helix α7.

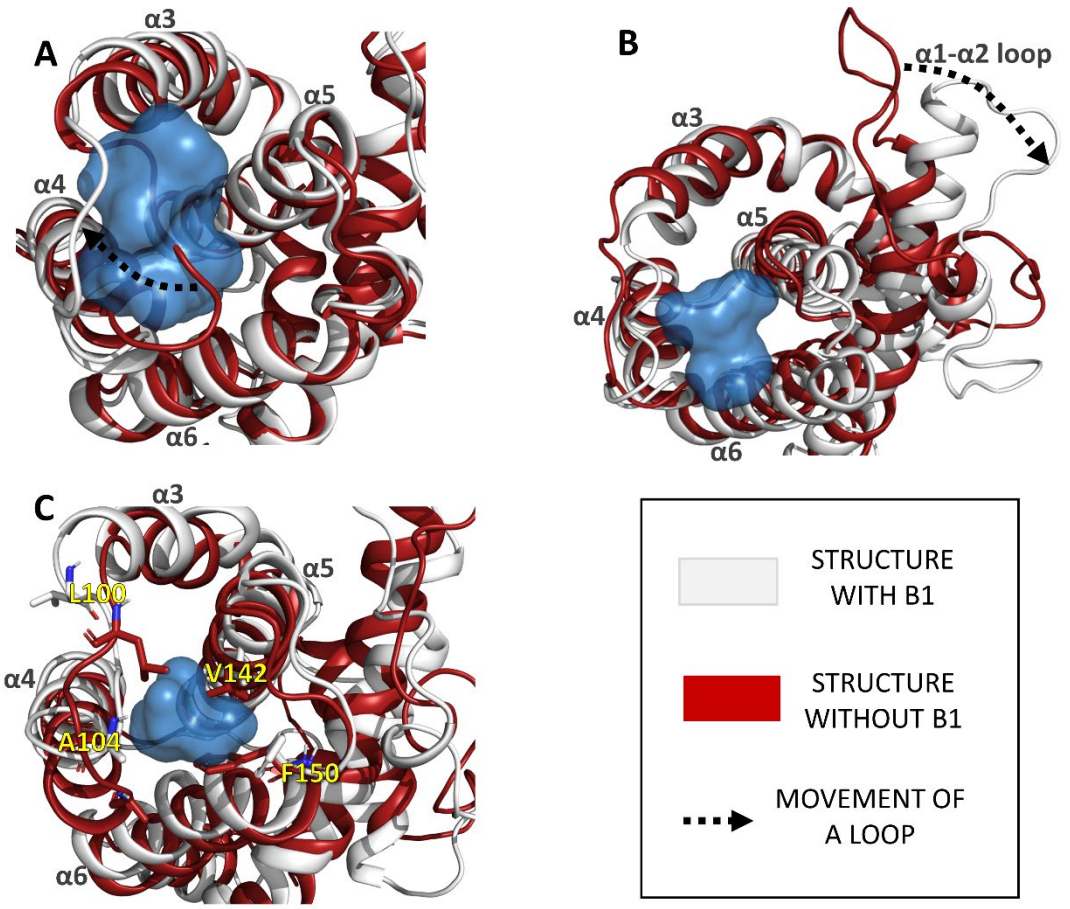

Figure 8. Conformational changes responsible for the B1 hotspot. Conformations that allow the emergence of the B1 site are represented in white; whereas conformations that do not allow the emergence of the site are shown in red. (A) Opening of the α3-α4 loop. (B) Opening of the α1-α2 loop, followed by a movement of the α5 helix. (C) Hydrophobic gate of the pocket involving residues L100, A104, V142 and F150.

B1, occurs in apoBak and intBak, but not in Bak/Bim complex. The presence of the BH3-peptide Bim reduces the flexibility of the loop between helices α3 and α4 (Figure 4), preventing accessibility to this pocket. Thus, the absence of B1 in the Bak/Bim complex hinders the presence of this site in the structure. The opening of this site is shown in the Supplementary Video 2. In this video, the distance between Ala104 and Leu147 is shown in two different orientations. A more detailed analysis on the apoBak and intBak systems suggests that there are specific conformations required to accomplish the pop-up of the site. First, the loop between helices α3 and α4 has to be in an open form, further from the loop between helices α5 and α6, as it can be observed in Figure 8A. Opening of the loop depends on its flexibility that is reduced when de BH3-Bim peptide is bound. Second, another structural element playing a role is the C-terminus of helix α5 that needs to be in open conformation, getting closer to helix α1 and the α1-α2 loop (Figure 8B). This small opening is produced when the N-terminus of the α1-α2 loop adopts an open conformation, away from helix α5 that induces B1 site opening. This latter specific conformation is observed with the positive correlation between B1 and Ab-site (Figure 7B). Finally, the last structural feature, closely related to the previous ones, is the orientation of four hydrophobic residues Leu100 and Ala104 (α3-α4 loop), Val142(α5) and Phe150 (α5-α6 loop) at the entrance of the site. These residues act as a "gate" preventing access to the site. However, an increased flexibility of the α3-α4 loop or the movement of the helix α5 produces an opening on this exterior part of the B1 site, these four residues cannot act as a barrier, allowing the entrance of

molecules. Moreover, when Phe150 is not interacting with the other residues, its lateral chain is oriented toward the exterior part of the site, as can be seen in Figure 8C. What gives special interest to this site is that it is located in the same spot where vMIA, a known inhibitor of apoptosis, binds in Bax. Moreover, the similarity of this region in both Bak and Bax allows us to hypothesize that this site could have a similar allosteric effect in Bak. However, the mechanism of how vMIA prevents apoptosis in Bax remains unclear. During Bak and Bax activation, a crucial step prior to dimerization is the separation of the *core* (α2-α5) and *latch* (α6-α8) domains, in other words, the loop between helices α5 and α6 must be flexible enough to allow this separation. Thus, the presence of the peptide vMIA in the neighbouring regions of this loop could block this separation, preventing apoptosis. However, it has been observed that BAM7 and BTSA1 are able to bind to the deep part of the vMIA site in Bax.[22] These two molecules are known Bax activators, although their role in Bax activation also remains unclear. To fully understand the mechanism behind ligand binding to this site in Bax and Bak, additional docking studies and experimental studies need to be performed to better understand a potential allosteric effect of this site.

Moreover, a study using FTMap for the hotspots in different crystallographic structures has been made. Two different apoBak crystals (5FMI and 5VX1), Bak/Bim (5VWZ) and another BH3-only endogenous activator, Bak/Bid (2M5B), have been used (Supplementary Figure S5). The analysis of this rigid conformations produces a limited study of the cavities or hotspots identified, indicating that the use of MD simulations is needed to improve the conformational exploration and the identification of new potential hotspots in Bak. B1 site has not been found in any of these systems. However, for the crystal 5VWZ it has been identified a pre cavity of this site, located in the external part (Supplementary Figure S5A). It cannot be considered as part of the B1 site because the rigidity of this conformation makes impossible the accession to the site, which is something that it can be observed in the MD simulations.

Present study permits also a close analysis to the known Bak activation sites. This is the case of the site triggered by an antibody, the Ab-site. Analysing the dynamics of the site, we have observed that it exhibits great stability despite having the loop between helices α1 and α2 nearby. Moreover, a negative correlation between the orthosteric binding site and this Ab-site was observed in all three complexes, which means that when the former exhibits a spacious conformation, the latter dwindles. This negative correlation implies that the presence of a molecule bound to the Ab-site has structural effects on the orthosteric binding site. The negative correlation can also be analysed by computing the continuity of the pocket in the intBak system. At the beginning of the MD simulation, helix α3 is closer to α1 and the loop α1-α2 forced by the presence of the BH3-Bim peptide bound to the hydrophobic groove in the starting conformation. However, as the MD evolves, helix α3 tends to move closed towards helix α4, reducing the volume of the orthosteric binding site and increasing the volume of the Ab-site at the same time. Finally, a positive correlation is observed between the Ab-site and B1 in the apoBak complex. This cross-talking between B1 and Ab-site has a major impact on Bak structure, since the Ab-site has been postulated as a Bak activation site. The presence of a molecule bound to the Ab-site increases the volume of the pocket, producing an increased width and a more flexible conformation of B1, facilitating the separation between domains during Bak activation. Thus, the activation mechanism behind the Ab-site ligands may be explained as due to the facilitation of the *core-latch* domain separation, a crucial process for Bak dimerization. However, further experiments analysing the structural features of these two pockets when molecules are bound to them are needed to fully understand Bak activation mechanism behind the Ab-site.

This study reveals crucial insights into the structure of Bak, identifying not only the activation site that has already been discovered (Ab-site) but also new potential binding sites, B1 and B2, in

promising Bak locations and with interesting protein dynamics characteristics. These pockets, especially B1, could solve the selectivity problem of the orthosteric binding site in the Bcl-2 family of proteins and could be good directed sites to perform screenings of compounds and docking experiments.

DATA AND SOFTWARE AVAILABILITY

MD trajectories produced during the execution of this work can be obtained from the authors upon request. *In-house* scripts, input files, topologies and PDB structures of all cluster representative used for the analysis are placed in the public repository Github:

https://github.com/DrugDesignUBUJA/BakPockets.git

SUPPORTING INFORMATION AVAILABLE

Detailed results regarding PCA and cluster analysis including the PDB structures of cluster representative are available in the supplementary material.

ACKNOWLEDGMENTS

This study was supported by the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR)-Generalitat de Catalunya (2021SGR00350 and 2021SGR00342), and Spanish Structures and Excellence María de Maeztu program, grant number CEX2021-001202-M.

REFERENCES

(1) Taylor, R. C.; Cullen, S. P.; Martin, S. J. Apoptosis: Controlled Demolition at the Cellular Level. *Nat. Rev. Mol. Cell Biol.* **2008**, *9* (3), 231–241.

(2) Favaloro, B.; Allocati, N.; Graziano, V.; Ilio, C. Di; Laurenzi, V. De. Role of Apoptosis in Disease. *Aging (Albany. NY).* **2012**, *4* (5), 330–349.

(3) Czabotar, P. E.; Lessene, G.; Strasser, A.; Adams, J. M. Control of Apoptosis by the BCL-2 Protein Family: Implications for Physiology and Therapy. *Nat. Rev. Mol. Cell Biol.* **2014**, *15* (1), 49–63.

(4) Chipuk, J. E.; Moldoveanu, T.; Llambi, F.; Parsons, M. J.; Green, D. R. The BCL-2 Family Reunion. **2010**, *37* (3), 299–310.

(5) Kushnareva, Y.; Newmeyer, D. D. Bioenergetics and Cell Death. *Ann. N. Y. Acad. Sci.* **2010**, *1201*, 50–57.

(6) Tait, S. W. G.; Green, D. R. Mitochondria and Cell Death: Outer Membrane Permeabilization and Beyond. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (9), 621–632.

(7) Singh, R.; Letai, A.; Sarosiek, K. Regulation of Apoptosis in Health and Disease: The Balancing Act of BCL-2 Family Proteins. *Nat. Rev. Mol. Cell Biol.* **2019**, *20* (3), 175–193.

(8) Shamas-Din, A.; Kale, J.; Leber, B.; Andrews, D. W. Mechanisms of Action of Bcl-2 Family Proteins. *Cold Spring Harb. Perspect. Biol.* **2013**, *5* (4), 1–21.

(9) Peña-Blanco, A.; García-Sáez, A. J. Bax, Bak and beyond — Mitochondrial Performance in Apoptosis. *FEBS J.* **2018**, *285* (3), 416–431.

(10) Leschiner, E. S.; Braun, C. R.; Bird, G. H.; Walensky, L. D. Direct Activation of Full-Length

Proapoptotic BAK. **2013**, *110* (11).

(11) Singh, G.; Guibao, C. D.; Seetharaman, J.; Aggarwal, A.; Grace, C. R.; McNamara, D. E.; Vaithiyalingam, S.; Waddell, M. B.; Moldoveanu, T. Structural Basis of BAK Activation in Mitochondrial Apoptosis Initiation. *Nat. Commun.* **2022**, *13* (1).

(12) Kvansakul, M.; Hinds, M. G. The Bcl-2 Family: Structures, Interactions and Targets for Drug Discovery. *Apoptosis* **2015**, *20* (2), 136–150.

(13) Kvansakul, M.; Hinds, M. G. *The Structural Biology of Bh3-Only Proteins*, 1st ed.; Elsevier Inc., 2014; Vol. 544.

(14) Cowan, A. D.; Smith, N. A.; Sandow, J. J.; Kapp, E. A.; Rustam, Y. H.; Murphy, J. M.; Brouwer, J. M.; Bernardini, J. P.; Roy, M. J.; Wardak, A. Z.; Tan, I. K.; Webb, A. I.; Gulbis, J. M.; Smith, B. J.; Reid, G. E.; Dewson, G.; Colman, P. M.; Czabotar, P. E. Bak Core Dimers Bind Lipids and Can Be Bridged by Them. *Nat. Struct. Mol. Biol.* **2020**, *27*, 1024–1031.

(15) Akçay, G.; Belmonte, M. A.; Aquila, B.; Chuaqui, C.; Hird, A. W.; Lamb, M. L.; Rawlins, P. B.; Su, N.; Tentarelli, S.; Grimster, N. P.; Su, Q. Inhibition of Mcl-1 through Covalent Modification of a Noncatalytic Lysine Side Chain. *Nat. Chem. Biol.* **2016**, *12* (11), 931–936.

(16) Lee, S.; Wales, T. E.; Escudero, S.; Cohen, D. T.; Luccarelli, J.; Gallagher, C. G.; Cohen, N. A.; Huhn, A. J.; Bird, G. H.; Engen, J. R.; Walensky, L. D. Allosteric Inhibition of Antiapoptotic MCL-1. *Nat. Struct. Mol. Biol.* **2016**, *23* (6), 600–607.

(17) Iyer, S.; Anwari, K.; Alsop, A. E.; Yuen, W. S.; Huang, D. C. S.; Carroll, J.; Smith, N. A.; Smith, B. J.; Dewson, G.; Kluck, R. M. Identification of an Activation Site in Bak and Mitochondrial Bax Triggered by Antibodies. *Nat. Commun.* **2016**, *7*, 1–10.

(18) Robin, A. Y.; Miller, M. S.; Iyer, S.; Shi, M. X.; Wardak, A. Z.; Lio, D.; Smith, N. A.; Smith, B. J.; Birkinshaw, R. W.; Czabotar, P. E.; Kluck, R. M.; Colman, P. M. Structure of the BAK-Activating Antibody 7D10 Bound to BAK Reveals an Unexpected Role for the A1-A2 Loop in BAK Activation. *Cell Death Differ.* **2022**, 1–12.

(19) Brouwer, J. M.; Lan, P.; Cowan, A. D.; Lessene, G.; Colman, P. M.; Czabotar, P. E.; Brouwer, J. M.; Lan, P.; Cowan, A. D.; Bernardini, J. P.; Birkinshaw, R. W. Conversion of Bim-BH3 from Activator to Inhibitor of Bak through Structure-Based Design Article Conversion of Bim-BH3 from Activator to Inhibitor of Bak through Structure-Based Design. *Mol. Cell* **2017**, *68* (4), 659–672.

(20) Moldoveanu, T.; Grace, C. R.; Llambi, F.; Nourse, A.; Fitzgerald, P.; Gehring, K.; Kriwacki, R. W.; Green, D. R. BID-Induced Structural Changes in BAK Promote Apoptosis. *Nat. Struct. Mol. Biol.* **2013**, *20* (5), 589–597.

(21) Edlich, F. BCL-2 Proteins and Apoptosis: Recent Insights and Unknowns. *Biochem. Biophys. Res. Commun.* **2018**, *500* (1), 26–34.

(22) Feng, G.; Zhang, X.; Li, Y.; Wang, R. Analysis of the Binding Sites on BAX and the Mechanism of BAX Activators through Extensive Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2022**, *62* (21), 5208–5222.

(23) Niu, X.; Brahmbhatt, H.; Mergenthaler, P.; Zhang, Z.; Sang, J.; Daude, M.; Ehlert, F. G. R.; Diederich, W. E.; Wong, E.; Zhu, W.; Pogmore, J.; Nandy, J. P.; Satyanarayana, M.; Jimmidi, R. K.; Arya, P.; Leber, B.; Lin, J.; Culmsee, C.; Yi, J.; Andrews, D. W. A Small-Molecule Inhibitor of Bax and Bak Oligomerization Prevents Genotoxic Cell Death and Promotes Neuroprotection. *Cell Chem. Biol.* **2017**, *24* (4), 493-506.e5.

(24) Sekar, G.; Singh, G.; Qin, X.; Guibao, C. D.; Schwam, B.; Inde, Z.; Grace, C. R.; Zhang, W.; Slavish, P. J.; Lin, W.; Chen, T.; Lee, R. E.; Rankovic, Z.; Sarosiek, K.; Moldoveanu, T. Small Molecule SJ572946 Activates BAK to Initiate Apoptosis. *iScience* **2022**, *25* (10), 105064.

(25) Park, D.; Anisuzzaman, A. S. M.; Magis, A. T.; Chen, G.; Xie, M.; Zhang, G.; Behera, M.; Sica, G. L.; Ramalingam, S. S.; Owonikoko, T. K.; Deng, X. Discovery of Small Molecule Bak Activator for Lung Cancer Therapy. *Theranostics* **2021**, *11* (17), 8500–8516.

(26) Vila-Julià, G.; Granadino-Roldán, J. M.; Perez, J. J.; Rubio-Martinez, J. Molecular Determinants for the Activation/Inhibition of Bak Protein by BH3 Peptides. *J. Chem. Inf. Model.* **2020**, *60* (3), 1632–1643.

(27) Wang, H.; Takemoto, C.; Akasaka, R.; Uchikubo-Kamo, T.; Kishishita, S.; Murayama, K.; Terada, T.; Chen, L.; Liu, Z. J.; Wang, B. C.; Sugano, S.; Tanaka, A.; Inoue, M.; Kigawa, T.; Shirouzu, M.; Yokoyama, S. Novel Dimerization Mode of the Human Bcl-2 Family Protein Bak, a Mitochondrial Apoptosis Regulator. *J. Struct. Biol.* **2009**, *166* (1), 32–37.

(28) D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C., and P. A. K. Amber18. University of California, San Francisco 2018.

(29) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. **2004**, 1157–1174.

(30) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Genet.* **2006**, *65* (3), 712–725.

(31) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.

(32) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(33) Arfken, GB; Weber, H. *Method of Steepest Descents. Mathematical Methods for Physicists*, 6th ed.; Press, E. A., Ed.; 2005.

(34) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.

(35) Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids: Second Edition. *Comput. Simul. Liq. Second Ed.* **2017**, 1–626.

(36) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341.

(37) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.

(38) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11* (8), 3584–3595.

(39) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120* (24), 11919.

(40) Perez, J. J.; Tomas, M. S.; Rubio-Martinez, J. Assessment of the Sampling Performance of Multiple-Copy Dynamics versus a Unique Trajectory. *J. Chem. Inf. Model.* **2016**, *56* (10), 1950–1962.

(41) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(42) Bakan, A.; Bahar, I. The Intrinsic Dynamics of Enzymes Plays a Dominant Role in Determining the Structural Changes Induced upon Inhibitor Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (34), 14349–14354.

(43) Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; Mccammon, J. A. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation - Journal of Chemical Theory and Computation (ACS Publications). *J. Chem. Theory Comput.* **2014**, *10*, 2677−2689.

(44) Camastra, F.; Vinciarelli, A. Clustering Methods. *Adv. Inf. Knowl. Process.* **2015**, 131–167.

(45) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTmap Family of Web Servers for Determining and Characterizing Ligand Binding Hot Spots of Proteins. *Nat. Protoc.* **2015**, *10* (5), 733–755.

(46) Chen, Z.; Zhang, X.; Peng, C.; Wang, J.; Xu, Z.; Chen, K.; Shi, J.; Zhu, W. D3Pockets: A Method and Web Server for Systematic Analysis of Protein Pocket Dynamics. *J. Chem. Inf. Model.* **2019**, *59* (8), 3353–3358.

(47) Sztain, T.; Amaro, R.; McCammon, J. A. Elucidation of Cryptic and Allosteric Pockets within the SARS-CoV-2 Main Protease. *J. Chem. Inf. Model.* **2021**, *61* (7), 3495–3501.

(48) Lee, E. F.; Grabow, S.; Chappaz, S.; Dewson, G.; Hockings, C.; Kluck, R. M.; Debrincat, M. A.; Gray, D. H.; Witkowski, M. T.; Evangelista, M.; Pettikiriarachchi, A.; Bouillet, P.; Lane, R. M.; Czabotar, P. E.; Colman, P. M.; Smith, B. J.; Kile, B. T.; Fairlie, W. D. Physiological Restraint of Bak by Bcl-XL Is Essential for Cell Survival. *Genes Dev.* **2016**, *30* (10), 1240–1250.

(49) Gahl, R. F.; He, Y.; Yu, S.; Tjandra, N. Conformational Rearrangements in the Pro-Apoptotic Protein, Bax, as It Inserts into Mitochondria: A Cellular Death Switch. *J. Biol. Chem.* **2014**, *289* (47), 32871–32882.

(50) Gavathiotis, E.; Suzuki, M.; Davis, M. L.; Pitter, K.; Bird, G. H.; Katz, S. G.; Tu, H.-C.; Kim, H.; Cheng, E. H.-Y.; Tjandra, N.; Walensky, L. D. BAX Activation Is Initiated at a Novel Interaction Site. *Nature* **2008**, *455* (7216), 1076–1081.

(51) Gavathiotis, E.; Reyna, D. E.; Davis, M. L.; Bird, G. H.; Walensky, L. D. BH3-Triggered Structural Reorganization Drives the Activation of Pro-Apoptotic BAX. *Mol. Cell* **2010**, *40* (3), 481.

(52) Li, M. X.; Tan, I. K. L.; Ma, S. B.; Hockings, C.; Kratina, T.; Dengler, M. A.; Alsop, A. E.; Kluck, R. M.; Dewson, G. BAK A6 Permits Activation by BH3-Only Proteins and Homooligomerization via the Canonical Hydrophobic Groove. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (29), 7629–7634.

(53) Gavathiotis, E.; Reyna, D. E.; Bellairs, J. A.; Leshchiner, E. S.; Walensky, L. D. Direct and Selective Small-Molecule Activation of Proapoptotic BAX. *Nat. Chem. Biol.* **2012**, *8* (7), 1–7.

(54) Reyna, D. E.; Garner, T. P.; Lopez, A.; Kopp, F.; Choudhary, G. S.; Sridharan, A.; Narayanagari, S. R.; Mitchell, K.; Dong, B.; Bartholdy, B. A.; Walensky, L. D.; Verma, A.; Steidl, U.; Gavathiotis, E. Direct Activation of BAX by BTSA1 Overcomes Apoptosis Resistance in Acute Myeloid Leukemia. *Cancer Cell* **2017**, *32* (4), 490–505.

(55) Ma, J.; Edlich, F.; Bermejo, G. A.; Norris, K. L.; Youle, R. J.; Tjandra, N. Structural Mechanism of Bax Inhibition by Cytomegalovirus Protein VMIA. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (51), 20901–20906.

(56) Zhang, Y. H.; Bhunia, A.; Wan, K. F.; Lee, M. C.; Chan, S. L.; Yu, V. C. K.; Mok, Y. K. Chelerythrine and Sanguinarine Dock at Distinct Sites on BclXL That Are Not the Classic BH3 Binding Cleft. *J. Mol. Biol.* **2006**, *364* (3), 536–549.

(57) Zhai, D.; Jin, C.; Satterthwait, A. C.; Reed, J. C. Comparison of Chemical Inhibitors of Antiapoptotic Bcl-2-Family Proteins. *Cell Death Differ. 2006 138* **2006**, *13* (8), 1419–1421.

(58) Ye, K.; Meng, W. X.; Sun, H.; Wu, B.; Chen, M.; Pang, Y. P.; Gao, J.; Wang, H.; Wang, J.; Kaufmann, S. H.; Dai, H. Characterization of an Alternative BAK-Binding Site for BH3 Peptides. *Nat. Commun.* **2020**, *11* (1), 1–14.