INSTITUTO POLITÉCNICO NACIONAL

ESCUELA NACIONAL DE MEDICINA
Y HOMEOPATÍA

Estudios de estabilidad de la proteína Bax humana

(Stability studies of the human Bax protein)

T E S I S

Que para obtener el grado de:

DOCTOR EN BIOTECNOLOGÍA

P r e s e n t a:

JORGE LUIS ROSAS TRIGUEROS

D i r e c t o r e s:

Dr. Absalom Zamorano Carrillo
Dr. José Correa Basurto

México D. F., Enero 2012

SIP-14-BIS

INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 12:00 horas del día 10 del mes de enero del 2012 se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de la ENMH para examinar la tesis titulada:

Estudios de estabilidad de la proteína Bax Humana

Presentada por el alumno:

Rosas	Trigueros	Jorge Luis
Apellido paterno	Apellido materno	Nombre(s)

Con registro:

B	0	8	1	1	0	1

aspirante de:

DOCTORADO EN CIENCIAS EN BIOTECNOLOGÍA

Después de intercambiar opiniones los miembros de la Comisión manifestaron ***APROBAR LA TESIS***, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Directores de tesis

Dr. Absalom Zamorano Carrillo

Dr. José Correa Basurto

Dr. César Augusto Sandino Reyes López

Dra. Claudia Guadalupe Benítez Cardoza

Dra. Irene Mendoza Lujambio

PRESIDENTE DEL COLEGIO DE PROFESORES

Dr. César Augusto Sandino Reyes López

INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

CARTA CESIÓN DE DERECHOS

En la Ciudad de México, el día 10 del mes de enero del año 2012, el que suscribe, Jorge Luis Rosas Trigueros, alumno del Programa de DOCTORADO EN CIENCIAS EN BIOTECNOLOGÍA con número de registro B08-1101, adscrito a la Escuela Nacional de Medicina y Homeopatía, manifiesta que es autor intelectual del presente trabajo de Tesis bajo la dirección del Dr. Absalom Zamorano Carrillo y del Dr. José Correa Basurto y cede los derechos del trabajo intitulado "Estudios de estabilidad de la proteína Bax humana", al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Éste puede ser obtenido escribiendo a la siguiente dirección **jlrosas@ipn.mx** o **azamorano@ipn.mx**. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

M. en C. Jorge Luis Rosas Trigueros

Acknowledgements

I would like to express my deep appreciation to my advisors, Dr. Absalom Zamorano Carrillo and Dr. José Correa Basurto, for their guidance and support.

I would also like to thank the evaluation jury for sharing their wisdom and for their support in adopting me into the field of Biotechnology.

I am thankful to my labmates as well for being so friendly and for their valuable help.

Finally, I would like to thank my family and friends for cheering me up and being endlessly supportive.

Dedication

To my family, source of my strength.

To my teachers, goldsmiths of the man and the student.

To my friends, partners in the quest for our ideals.

Abstract

Bax is a member of the Bcl-2 protein family that participates in mitochondrion-mediated apoptosis. In the early stages of the apoptotic pathway, Bax migrates from the cytosol to the outer mitochondrial membrane, where it is inserted and usually oligomerizes to build cytochrome c-compatible pores. Although there are several cellular and structural studies about the stability of Bax, its description at the molecular level remains elusive. Therefore, in this work, computational (*in silico*) studies were performed on the available structural information of Bax, retrieved from the Universal Protein Resource (Uniprot) and the Protein Data Bank (PDB). This work was focused on the most relevant structural changes observed, investigating their relationship with biological experimental results. These studies include sequence analysis and molecular dynamics (MD) simulations of monomeric human Bax at 300, 400, and 500 K. Sequence analysis was useful for identifying conserved regions in the protein that were later related to structural stability and function. In the MD simulations, Bax gradually loses its α-helices when it is submitted to high temperatures, yet it maintains its globular conformation. The resistance of Bax to adopt an extended conformation could be due to several interactions that were found to be responsible for maintaining the structural stability of this protein. Among these interactions, electrostatic interactions, hydrophobic interactions, and hydrogen bonds were found. Remarkably, electrostatic interactions were the most relevant to prevent the elongation of the structure. This atomistic description might have important implications for understanding the functionality and stability of Bax *in vitro* as well as within the cellular environment. The applications of the knowledge about the activation mechanism of Bax include the development of new drugs and therapies for cell death related diseases such as cancer, and the design of new biosensors employing bilipidic membranes based on the modulation of their permeability.

Resumen

La proteína Bax pertenece a la familia Bcl-2, la cual participa en la ruta mitocondrial de la apoptosis. En etapas tempranas de la ruta apoptótica, esta proteína migra del citosol a la membrana mitocondrial externa, en donde se inserta y frecuentemente se oligomeriza para formar poros compatibles con el citocromo c. Aunque se han reportado diversos estudios celulares y estructurales, aún no se cuenta con una descripción a nivel molecular de la estabilidad de Bax. Por ello, en este trabajo se realizaron estudios computacionales (*in silico*) usando la información estructural disponible de la proteína Bax, obtenida del Repositorio Universal de Proteínas (Universal Protein Resource, Uniprot) y del Banco de Datos de Proteína (Protein Data Bank, PDB). Este trabajo se enfocó en los cambios estructurales más relevantes observados, investigando su relación con resultados experimentales biológicos. Los estudios realizados incluyen análisis de secuencias y simulaciones de dinámica molecular de un monómero de la proteína Bax humana a 300, 400 y 500 K. El análisis de secuencias permitió identificar regiones conservadas en la proteína que posteriormente fueron relacionadas con la estabilidad estructural y la función. En las simulaciones de dinámica molecular, Bax gradualmente pierde sus α-hélices cuando se expone a altas temperaturas, pero mantiene su conformación globular. La resistencia a adoptar una conformación extendida de esta proteína puede tener su origen en múltiples interacciones que se identificaron como responsables de mantener la estabilidad estructural. Entre estas interacciones encontramos puentes salinos, interacciones hidrofóbicas y enlaces de hidrógeno. Notablemente, los puentes salinos fueron los más importantes para impedir la elongación de la estructura. Esta descripción a nivel atómico podría tener gran impacto en la comprensión de la función y estabilidad de Bax *in vitro* así como en el entorno celular. Las aplicaciones de la comprensión del mecanismo de activación de Bax incluyen el desarrollo de nuevos fármacos y terapias para enfermedades relacionadas a la muerte celular como por ejemplo el cáncer y el diseño de nuevos biosensores basados en el control de la permeabilidad de una membrana bilipídica.

Contents

Abstract **i**

Resumen **ii**

List of Tables **vi**

List of Boxes **vii**

List of Figures **viii**

1 Introduction **1**
1.1 Background 1
1.1.1 Apoptosis 1
1.1.2 Apoptosis and disease 2
1.1.3 The mitochondrial pathway of apoptosis 2
1.1.4 Bcl-2 family 4
1.1.5 The human Bax protein 6
1.1.6 Bax in apoptosis and disease 9
1.1.7 Modeling and simulation of biomolecules 10
1.1.8 Biological sequence analysis 10
1.1.9 Molecular dynamics simulations 12
1.1.10 Molecular dynamics studies of the Bcl-2 family 12
1.1.11 Folding and unfolding of proteins 13
1.1.12 Thermal unfolding simulations 14
1.2 Motivation 14
1.3 Objectives 15

2 Materials and methods **16**
2.1 Universal Protein resource 16

2.2 Cluster alignment . . . 16
2.3 Analysis of available Bax sequences . . . 17
2.4 GROMACS . . . 17
2.5 Force field . . . 18
2.5.1 Optimized potentials for liquid simulations for all atoms . . . 18
2.5.2 Simple point charge water . . . 18
2.6 Calculation of electrostatic interactions . . . 19
2.7 Temperature coupling . . . 19
2.8 Pressure coupling . . . 20
2.9 Linear constraint solver . . . 21
2.10 High temperature simulation of a Bax monomer . . . 21
2.11 Normal mode analysis . . . 24
2.11.1 *ElNémo* . . . 24
2.11.2 Normal mode analysis of human Bax . . . 24

3 Results and discussion **25**
3.1 Sequence analysis of Bax . . . 25
3.1.1 Composition histogram of available sequences of Bax . . . 25
3.1.2 Multiple alignment of available sequences of Bax . . . 25
3.2 Structural stability of a Bax monomer in aqueous media . . . 32
3.2.1 Overall stability . . . 32
3.2.2 Clustering analysis . . . 37
3.2.3 Secondary structure evolution . . . 37
3.2.4 Electrostatic interactions keep Lp1-2 close to Lp5-6 . . . 40
3.2.5 Residue D33 remains close to BH3 . . . 41
3.2.6 Unfolding of αH4 protects the hydrophobic core . . . 41
3.2.7 Exposure of region immunodetected by antibody 6A7 (residues 12-24) at high temperatures . . . 44
3.2.8 Hydrophobic groove covering by αH9 . . . 51
3.2.9 Hydrophobic interaction between W107 and BH2 residues . . . 52
3.2.10 Normal mode analysis . . . 55
3.3 Assessment of interactions found to contribute to the structural stability of Bax . . . 55
3.4 Time evolution of other relevant residues . . . 63

4 Conclusions and further work **68**
4.1 Conclusions . . . 68
4.2 Further work . . . 69

4.2.1 MD simulation of the t-Bid/Bax complex 69
4.2.2 MD simulation of the interaction between Bax and lipidic membranes 69
4.2.3 MD simulation of the Bax/Bax complex 69
4.2.4 MD simulations of point-mutated structures to further study the stability of Bax 72

Appendix 1 **73**

Appendix 2 **75**

Bibliography **83**

Index **100**

List of Tables

1.1 *Comparison of the morphological and biochemical aspects of Apoptosis and Necrosis.* 3

3.1 *Conserved amino acids in the multiple alignment considering 23 sequences of Bax from several species.* 29
3.2 *Conserved amino acids in the multiple alignment considering the seven available isoforms of human Bax.* 30
3.3 *Average minimum distance of interactions proposed in the manuscript (time average from 10 to 20 ns). Distance is defined as the minimum distance between any pair of atoms from the respective groups of amino acids (Groups 1 and 2). Bold font is used when the distance is not greater than 7 Å.* 62
3.4 *Average minimum distance of interactions proposed in the manuscript (from 15 to 20 ns). Distance is defined as the minimum distance between any pair of atoms from the respective groups of amino acids (Groups 1 and 2). Bold font is used when the distance is not greater than 7 Å.* 64

List of Boxes

2.1 *Parameters for the Nosé-Hoover temperature coupling for the simulation at 500 K.* 20
2.2 *Parameters for the Parrinello-Rahman pressure coupling used for the MD simulations performed.* 20
3.1 *Pseudocode used to find the most distant structures within the NMR models from file 1F16.pdb.* 56
3.2 *Bax NMR Models with the greatest distance between each other.* 57

List of Figures

1.1 *Bcl-2 family members share regions of homology called Bcl-2 homology (BH) domains, and may contain a trans-membrane (TM) domain that mediates insertion into the outer membrane of the mitochondria and to the endoplasmic reticulum. The pro-survival family members, including Bcl-2, Bcl-x_L, Bcl-w, Mcl-1 and A1, share four BH domains plus a TM region. The pro-apoptotic family members can be sub-divided into two subgroups: the multi-BH domain proteins, and the BH3-only proteins. Multi-BH domain proteins contain up to four BH domains and include Bax, Bak, Bok and Bcl-x_S, which include a TM region, and also Bcl-g_L and Bfk which lack a TM region. The BH3-only proteins contain only the BH3-domain and include Bad, Bik, Bid, Hrk, Bim, Puma, Noxa and Bmf. Bik and Bim also contain a TM region (taken from Kelly, P.N. et al., Cell Death & Differentiation, 2011).* 7

1.2 *Bax 3D structure (PDB ID: 1F16). The nine α-helices are shown as tubes. The location of BH1-3 is indicated by the transparent surfaces. Labels indicate the N-terminal (N-t) and C-terminal (C-t) ends of the protein.* 8

1.3 *The number of available DNA sequences (GenBank) have outnumbered that of protein sequences (SWISS-PROT) by one or two orders of magnitude. Meanwhile, there is one order of magnitude more protein sequences than 3D structures (PDB), from* `http://www.kanehisa.jp/en/db_growth.html`. 11

2.1 *Gap opening, extension and separation penalties are lower than the default values and improvement of the alignment is requested in every iteration. The maximum available number of iterations is requested.* 17

2.2 *Construction of the simulation Box. The left image shows the flow diagram of the solvation of the first model from the file 1F16.pdb in SPC water. The image on the right is the flow diagram for the addition of counter ions to the system.* 22
2.3 *Simulation of the system. The left image shows the flow diagram for the equilibration of the system. The image on the right shows the flow diagram of the way GROMACS was used to produce the simulations.* . . 23

3.1 *Percentage of amino acid content of the 23 available sequences of Bax. To each sequence corresponds a different color. The lines join the average values.* 26
3.2 *Multiple alignment of all available Bax sequences part 1 of 3 (M1-F30 in human Bax isoform alpha).* 27
3.3 *Multiple alignment of all available Bax sequences part 2 of 3 (I31-D102 in human Bax isoform alpha).* 28
3.4 *Multiple alignment of all available Bax sequences part 3 of 3 (G103-G192 in human Bax isoform alpha).* 29
3.5 *Multiple alignment of all Homo Sapiens isoforms of Bax.* 31
3.6 *Time evolution of the Cα RMSD of the protein in the MD simulations performed.* 33
3.7 *Time evolution of the Radius of gyration (Rg) of the protein in the MD simulations performed.* 34
3.8 *Time evolution of the Solvent Accessible Surface Area (SASA) of the protein in the MD simulations performed.* 35
3.9 *The Root Mean Square Fluctuations (RMSF) of Cα coordinates are shown as a function of residue number at different temperatures; a representation of the secondary structure with labels for the main domains is included.* 36
3.10 *RMS Distribution for the structures obtained in the simulations in arbitrary units (a.u.).* 38
3.11 *Time evolution of the clustering of the structures from the simulations. Cutoff values of 0.15, 0.22 and 0.42 nm were used for clustering the structures at 300, 400 and 500 K respectively.* 39

3.12 *Aligned structures from two different clusters in the MD simulations. a) Structures at 3.1 ns and 17.2 ns from the MD simulation at 300 K. Differences in LpNt, Lp1-2 and the C-terminal end of BH2 are observed. b) Structures at 4.9 ns and 17.2 ns from the MD simulation at 400 K. Important differences are appreciated in LpNt, Lp1-2 and BH2. c) Structures at 8.1 ns and 19.4 ns from the simulation at 500 K. aH5 is depicted with a green tube at 8.1 ns and a purple tube at 19.4 ns.* 39
3.13 *Time evolution of the secondary structural elements of the protein during the MD simulations at different temperatures (DSSP classification). The graphical representation of the native secondary structure used in Fig. 3.9 is reproduced on the extreme right.* 40
3.14 *Snapshots of the structure of Bax evolved at different time points at different temperatures. Important residues and the hydrogen bonds where they are involved are shown. The insets show distances in nm.* 42
3.15 *All 20 NMR models from the PDB file are shown (PDB ID: 1F16).The color coding representation for the depicted residues ranges from red for model 1 to blue for model 20.* 43
3.16 *Minimum distance between E41E44 and K128R134. Two sets are defined, one that contains E41 and E44 and the other K128 and R134. The minimum distance between a pair formed with one element of each set is shown in this figure to be close to the van der Waals distance (0.16 nm) at several times in all MD simulations. This proximity is important in the stability of Bax because it prevents the separation of the hairpin formed by αH1 and αH2 from the hydrophobic core (αH5 and αH6).* 43
3.17 *Minimum distance between D33 and K64R65. The distance between D33 and these basic residues in αH2 is short even at 500 K.* 44
3.18 *Structural alignment between final structures (20 ns) at 300 (glassy) and 500 K. Domains αH4 (green) and αH5 (red) are highlighted.* . . . 45
3.19 *Solvent Access to 12-24 (epitope 6A7). In the total solvent accessible surface area (above), the values remain below 11.5 nm^2 in all simulations. The residues that cover the epitope allow for enough space for the solvent but block the corresponding antibody. ΔGsolv values (below) remain larger than -28 Kcal/mol most of the time.* 47

3.20 *Minimum distance between D2R9-D2R145. The proximity of D2 and R145 stabilizes the bending of LpNt and the proximity of D2 and R145 keeps LpNt covering the epitope 6A7 (residues 12-24). The minimum values observed are below 0.2 nm, which leaves them close to the van der Waals distance of 0.16 nm. Interestingly, this proximities are more pronounced at 500 K than they are at 400 K and are not present in the conformation at 2 ns but appear later in the MD simulations.* 48
3.21 *Minimum distance between D6-R9. The proximity of D6 and R9 contributes to the stabilization of the bending of LpNt that facilitates its function of covering the epitope 6A7. Interestingly, this proximity is only observed in the MD simulations at high temperatures. The minimum distance is close to the van der Waals distance (0.16 nm) at 400 K and 500 K.* 48
3.22 *Exposure of the epitope 6A7 under heating conditions. Final structures for the MD simulations at 300 K (a) and 400 K (b) are shown. Residues K21, Q52, D2 and R145 are shown in CPK. The epitope is shown in green; it starts in LpNt and includes the first residues in αH1, and the rest of αH1 is shown in yellow. Residues P49-S60 are shown with transparent surfaces in green. Domain αH6 is shown in red.* . . . 49
3.23 *Crystal Structure of the anti-Bax monoclonal antibody 6A7 and a Bax peptide (PDB ID: 2G5B). The residues from Bax that are seen close to the antibody are T14-I19.* 49
3.24 *Hydrogen bonds between LpNt and the rest of the protein. All simulations show a net increase in the number of hydrogen bonds between LpNt and the rest of the protein, helping to keep it close to the core. Pairs within 0.35 nm are likely to have a favorable interaction, and their number is also increased in all simulations, thus contributing to the proximity of LpNt and the core of Bax.* 50
3.25 *Electrostatic interactions that prevent αH9 from moving away from the protein core. K189 is shown in Licorice and K190 in CPK. αH9 is depicted in green. Acidic residues that form the ionic "cage" are depicted in the VDW representation. In all cases, the minimum distance between the lysines and the acid residues is close to 1.6 Å, which is the van der Waals distance for a nonbonded O-H pair.* 52

3.26 *SASA of S184. The exposure of S184 can be followed in this figure. At 300 K, the value of its SASA is always below 0.5 nm^2, whereas it is above this value intermitently at 400 K and for most of the time at 500 K. This exposure is also noticeable in the free energy of solvation (ΔGsolv), where the values are lowest for the MD simulations at high temperatures.* 53
3.27 *SASA of W107. The exposure of W107 can be followed in this figure. At 300 K, the SAS area remains close to 1 nm^2, whereas it is closer to 2 nm^2 at several points at 400 K and for most of the simulation at 500 K. This exposure is also noticeable in the free energy of solvation (ΔGsolv), where the values are lowest for the simulation at 500 K.* . . 53
3.28 *Final structures (20 ns) are shown from the MD simulations at 300 K (left), 400 K (middle) and 500 K (right). W107 is shown in VDW and the SAS of the region formed by residues L148-F165 in transparent gray. The color code for the helices is the same as in Figure 1. At 300 K, W107 remains buried and its rings are perpendicular to Lp7-8. At 400 K, W107 is more exposed, and its rings are almost parallel to Lp7-8. At 500 K, W107 is mostly exposed. The hydrophobic and aromatic residues in BH2 (W151, I152, W158, L161, L161, Y164 and F165) are shown in licorice.* 54
3.29 *Normal Mode Analysis. The first ten nontrivial modes are displayed. The leftmost was performed on the first NMR model of the structure from the PDB. The rest were performed on the final structure for each simulation. Low collectivity modes at 300 K are mostly due to LpNt as it is for the NMR structure. At 400 and 500 K these modes include large mobility of Lp1-2, Lp5-6, αH4 and BH2.* 56
3.30 *Time evolution of Cα RMSD from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).* 58
3.31 *Time evolution of the Rg from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).* 59
3.32 *Time evolution of the SASA from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).* 60
3.33 *RMSF from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).* 61
3.34 *Global stability parameters from a 30ns MD simulation of Bax. RMSD (a), Rg (b) and SASA (c) values maintain their trend after 20 ns. RMS fluctuation values (d) are only calculated for the trajectory from 20 to 30 ns and are similar to Fig. 3.9.* 65

3.35 *Alignment between final structures of Bax from MD simulations of 20 and 30 ns at 300, 400 and 500 K. The regions where there is little structural correspondence or none at all are shown in red.* 66

4.1 *Final structure of a short MD simulation of t-Bid (red) in close proximity of Bax (blue) at 300 K.* 70
4.2 *Final structure of a short MD simulation of two Bax monomers close to each other at 500 K.* 70
4.3 *Point mutations that should decrease the stability of Bax. The original residues (D33, K64 and R65) are depicted in licorice representation, while the alanines that would replace them are shown as yellow spheres.* 71
4.4 *Point mutation that should increase the stability of Bax. The original residue (Q52) is depicted in green, while the aspartic acid that would replace it is shown in red. The lysine (K21) that is found to interact with Q52 in the MD simulations is shown in blue.* 71

Chapter 1

Introduction

1.1 Background

1.1.1 Apoptosis

It is known that all cells in multicellular organisms can either undergo mitosis to divide, differentiate in order to specialize or they can die [1]. Cell death mechanisms can be classified as apoptotic, necrotic, autophagic or associated with mitosis according to their morphological appearance. Alternatively, conforming to an enzymological criteria, there are death mechanisms with and without the involvement of nucleases or of distinct classes of proteases (caspases, calpains, cathepsins and transglutaminases). A third classification can be defined by their functional aspects, where there are programmed or caused by damage, physiological or pathological. Finally, in agreement with their immunological characteristics, they can be immunogenic or non-immunogenic. The Nomenclature Committee on Cell Death (NCCD) proposes four categories: apoptosis, autophagy, cornification and necrosis [2].

Apoptosis refers to a programmed cell death which morphologically comprises rounding-up of the cell, retraction of pseudopodes, reduction of cellular and nuclear volume, nuclear fragmentation, minor modification of cytoplasmic organelles, plasma membrane blebbing and engulfment by resident phagocytes *in vivo* [3].

Apoptosis is mediated by the activation of proapoptotic Bcl-2 family proteins (e.g., Bax, Bak, Bid), mitochondrial transmembrane permeabilization and mitochondrial outer membrane permeabilization. Other biochemical features of apoptosis are activation of caspases, oligonucleosomal DNA fragmentation, plasma membrane rupture, phosphatidylserine exposure, reactive oxygen species overgeneration and single-stranded DNA accumulation [2]. Numerous cytotoxic factors can trigger the intrinsic

pathway of apoptosis, including heat shock, abundance of oxidants, viral infection and others [4].

There are several differences between apoptosis and necrosis. While there is cellular volume reduction in apoptosis, necrosis involves cytoplasmic swelling and the rupture of the plasma membrane in necrosis is replaced by membrane blebbing in apoptosis. Biochemically, the mitochondrial membrane permeabilization in apoptosis is replaced by lysosomal membrane permeabilization in necrosis. The morphological and biochemical comparison of apoptosis and necrosis is summarized in Table 1.1 [2].

1.1.2 Apoptosis and disease

In a healthy organism, cell duplication and cell elimination compensate each other. Perturbations of this homoeostatic phenomenon are often caused by deregulated cell death and lead to pathological conditions. The latter include the excessive cell death in post-mitotic cells that causes severe conditions such as the stroke and necrosis caused by the loss of cardiomyocytes and the degenerative diseases caused by neuron death. Another case is the steep loss of immune cells that causes AIDS. Conversely, oncogenesis is linked to the suppression of cell death to various degrees [5, 4], and experimental evidence suggests that the genes that control apoptosis have a profound effect on the malignant phenotype [6]. Apoptosis is also important in the development of multicellular organisms and its disregulation has been related to malformations [7]. Other diseases related to aberrant apoptosis include hematological, autoimmune, cardiovascular, metabolic disorders, atherosclerosis, ischemic injury, and bacterial infections among others [8].

1.1.3 The mitochondrial pathway of apoptosis

The most common form of apoptosis observed in vertebrate cells in most settings proceeds through the mitochondrial pathway [9, 10]. During apoptosis, the pivotal event in this pathway is mitochondrial outer membrane (MOM) permeabilization. Additionally, there is frequently dissipation of the mitochondrial inner transmembrane potential ($\Delta\Psi m$) [10]. MOM permeabilization releases proteins normally found in the space between the inner and outer mitochondrial membranes (including cytochrome c and AIF), which are important for caspase activation [10, 11].

Table 1.1: *Comparison of the morphological and biochemical aspects of Apoptosis and Necrosis.*

Cell death mode	**Apoptosis**	**Necrosis**
Morphological features	-Reduction of cellular and nuclear volume -Rounding-up of the cell -Retraction of pseudopodes -Nuclear fragmentation -Minor modification of cytoplasmic organelles -Plasma membrane blebbing -Engulfment by resident phagocytes, *in vivo*	-Cytoplasmic swelling (oncosis) -Rupture of plasma membrane -Swelling of cytoplasmic organelles -Moderate chromatin condensation
Biochemical features	-Activation of proapoptotic Bcl-2 family proteins (e.g., Bax, Bak, Bid) -Activation of caspases -$\Delta\Psi m$ dissipation -Mitochondrial membrane permeabilization -Oligonucleosomal DNA fragmentation -Exposure of phosphatidylserine (PS) -ssDNA accumulation	-Activation of calpains -Activation of cathepsins -Drop of ATP levels -High-mobility group protein B1 (HMGB-1) release -Lysosomal Membrane Permeabilization (LMP) -Kinase RIP1 phosphorylation -Kinase RIP1 ubiquitination -Specific Poly [ADP-ribose] polymerase 1 (PARP1) cleavage pattern

1.1.4 Bcl-2 family

The second member of a set of proteins initially described in chromosomal traslocations involving chromosomes 14 and 18 in follicular lymphomas was named B-cell lymphoma 2 (Bcl-2) [12]. Bcl-2 became the first of a large family of proteins (Bcl-2 family). The physical and dynamical behavior of this family is of fundamental importance in apoptosis control by regulating mitochondrial permeability and also the release of Ca^{+2} from the endoplasmic reticulum. Interestingly, some members of this family promote apoptosis, whereas the rest have anti-apoptotic properties [13, 14, 15].

Bcl-2 was early shown to express in B-cells [12, 16]. Upregulation of Bcl-2 was later observed in human cancer cell lines. Bcl-2 is upregulated by estradiol in breast-cancer cells [17]. Also, mice with a genetic mutation that knocks out an estrogen receptor have an overexpression of Bcl-2 in their ventral prostate [18]. The Bcl-2 family has roles in numerous cell lineages, as indicated by a survey of adult tissues. Glandular epithelium that undergoes hyperplasia or involution in response to hormonal stimuli or growth factors expresses Bcl-2, whereas this protein is restricted to stem cell and proliferation zones in complex epithelium such as the skin or gut [19]. Regulation of cell death by Bcl-2 has also been observed in early life stages of Korean fetal scalp tissue cells [20]. The Bcl-2 protein is widely expressed in the nervous system during development, whereas Bcl-2 immunoreactivity gradually declines with aging in the central nervous system. On the other hand, Bcl-2 expression is selectively retained in the adult peripheral nervous system, suggesting a role in neuronal survival [21]. Additionaly, Bcl-2 is expressed in enteric neurons of the fetal gut and in the postnatal gut. The Bcl-2 immunoreactivity is also retained in adult enteric neurons. These results indicate that the Bcl-2 protein may be involved in promoting survival of the enteric neurons throughout life [22].

Bcl-2 family proteins have one to four Bcl-2 Homology (BH) domains (Fig. 1.1) [23]. Sequence and structural similarities found elsewhere between Bcl-2 and the anti-apoptotic *C. elegans* protein ced-9 suggested that the molecular mechanism of programmed cell death has been conserved from nematodes to mammals [24]. A decade later, a total of 83 Bcl-2 family gene sequences from vertebrates, nematodes, sponges and arthropods were available. Multiple alignments of these sequences revealed that three groups could be defined: a) the Bcl-2-like survival factors, similar to ced-9, including Bcl-2, Bcl-x_L, Bcl-w, Mcl-1, A1/Bfl-1, NR13, Boo/Diva/Bcl-2-L10 and Bcl-B; b) the BH3-only death factors, including Bik/Nbk, Blk, Hrk/DP5, BNIP3, BimL/Bod, Bad, Bid, Noxa, PUMA/Bbc3 and Bmf; c) the subgroup of pro-apoptitic proteins with no homologs in *C. elegans*, including Bax, Bak, Bok/Mtd, Bcl-x_S and *Drosophila* DEBCL. Furthermore, since the programmed cell death is executed with

very few members in invertebrates (one pro- and one anti-apoptotic member in nematode, and two proapoptotic in insects), which contrasts with the genomic expansion of vertebrates, the authors concluded that the explosion of paralogous sequences took place after chordates [25].

This BH structural properties can be used to group the family members into three subfamilies: a) the pro-apoptotic multi-domain proteins with three BH domains (BH1-3), b) the anti-apoptotic members, with either BH1-3 or BH1-4; and c) the pro-apoptotic BH3 only proteins [26]. Anti-apoptotic members perform their function by sequestering those who are pro-apoptotic [27]. Pro-apoptotic multi-domain members promote the mitochondrial apoptotic pathway [27], while BH3 only members can either activate multi-domain pro-apoptotic members [28] or inactivate anti-apoptotic members [29].

A number of interactions between Bcl-2 family members have been reported. Since its identification, Bax was shown to homodimerize and to form heterodimers with Bcl-2 *in vivo* as revealed by coimmunoprecipitation experiments [30]. Also simultaneous to its identification, Bak was shown to bind to Bcl-x_L [31] and was later shown to form a homodimer that can form crystals using polyethylene glycol as precipitant and zinc acetate or ammonium fluoride as additives [32]. Bcl-2 itself was shown to form homodimers in cross-linking experiments, and also in co-precipitation with wild type Bcl-2 bearing an epitope of the influenza virus haemagglutinin. Moreover, the ability of Bcl-2 to form homodimers was left intact by mutations which abrogated function and heterodimerization [33].

Lee *et al.* found that a signicant proportion (up to 80%) of Bcl-w without its membrane anchor (ΔC29) puried from *Escherichia coli* extracts, was dimeric as assessed by gel-filtration chromatography and analytical ultracentrifugation [34]. Both yeast two-hybrid screening and λ expression cloning identified a protein, Bad, which selectively dimerized with Bcl-x_L as well as Bcl-2, but not with Bax, Bcl-x_S, Mcl-1, A1, or itself [35].

Jeong *et al.* performed co-immunoprecipitation experiments that indicate that Bcl-x_L in the cytosol forms homodimers. This workgroup also found that Bad binding to Bcl-x_L dissociates the homodimers and triggers Bcl-x_L binding to mitochondrial membranes and also that the C-terminal tail of Bcl-x_L is required to mediate Bcl-x_L/Bax heterodimer formation [36]. There are other biomolecules which have been shown to interact with the Bcl-2 family members and regulate their activity, such as the anti-apoptotic viral protein M11L [37], ceramide [38] and cardiolipin [39].

Many workgroups have been working to develop anti-cancer drugs that block the function of Bcl-2 members. Bcl-2 anti-sense therapy was an initial advancement towards this goal of targeting Bcl-2 to inhibit its overexpression, for instance by

specifically binding to the first six codons of the human Bcl-2 mRNA sequence, causing degradation of Bcl-2 mRNA, resulting in consequent reduction in Bcl-2 protein translation and intracellular concentration [40].

Alternatively, small molecule inhibitors of Bcl-2 family proteins are increasingly being realized as effective anti-cancer agents either alone or to sensitize cancer cells to standard chemotherapeutics agents. Their advantages include that they are specific and targeted agents that can be given orally and do not have any appreciable toxicity [41].

Other approaches include the development of an antibody that would block the activity of Bcl-2 [42], cell permeable peptides that bind to Bcl-2 [43] and stabilized α-helix of Bcl-2 domains (SAHBs) which consist of a helical BH3 domain with additional bonds that increase their stability and have high affinity to multi-domain member pockets and are able to, for instance, activate Bax [44].

Furthermore, recently reported clinical strategies targeting the Bcl-2 family have yielded promising results in cancer treatment [45, 26].

1.1.5 The human Bax protein

Bax is a Bcl-2 protein family member that participates in mitochondrion-mediated apoptosis. In the early stages of the apoptotic pathway, this protein migrates from the cytosol to the outer mitochondrial membrane, where it is inserted and usually oligomerizes, making cytochrome c-compatible pores.

Human Bax (Bcl-2-associated X, 192 amino acids) belongs to the pro-apoptotic multi-domain subfamily (BH1-3), shows a large N-terminal loop (M1-T14) and contains a hydrophobic C-terminal helix (Q171-W188). Native Bax is a monomer located in the cytosol and plays an important role in apoptosis. The three-dimensional (3D) structure of Bax has been determined by nuclear magnetic resonance (NMR), identifying nine α-helices (αH1-9), joined by loops (Lp1-2, Lp2-3, Lp3-4, Lp4-5, Lp5-6, Lp6-7, Lp7-8, and Lp8-9) with some turns. With this 3D structure for Bax, it is possible to identify the characteristic functional domains BH1, BH2 and BH3, which are formed by residues F100-V121, W151-F165 and L59-D71, respectively (Fig. 1.2) [46].

Several details about Bax structure have been recognized as being important for the functionality and stability of this protein. For instance, a computational effort has predicted that the amino acid sequence of human Bax has a highly conserved region formed by ten relevant residues in BH1 (G103, N106, G108, R109, V111 and A112) and BH2 domains (W151, G157 and W158), as well as outside of them (V95). Interestingly, the structural analysis showed that many of these residues are protected

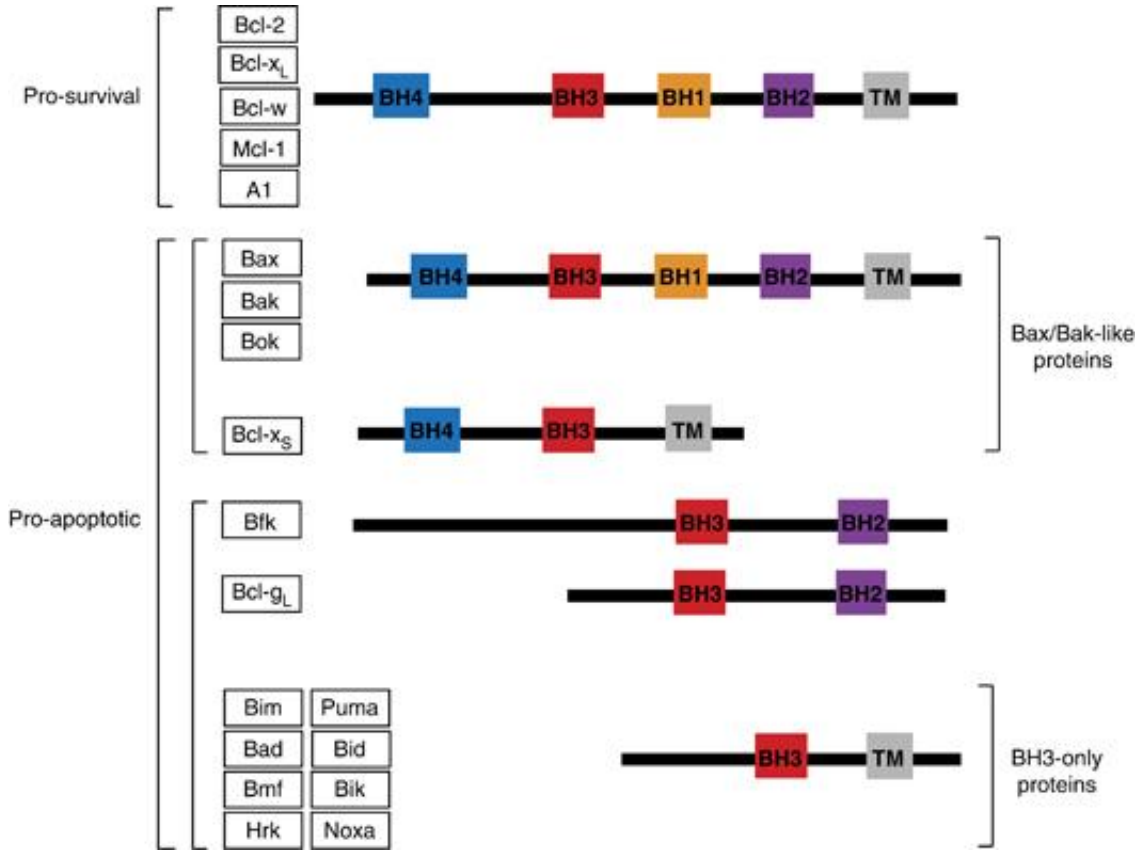


Figure 1.1: *Bcl-2 family members share regions of homology called Bcl-2 homology (BH) domains, and may contain a trans-membrane (TM) domain that mediates insertion into the outer membrane of the mitochondria and to the endoplasmic reticulum. The pro-survival family members, including Bcl-2, Bcl-x_L, Bcl-w, Mcl-1 and A1, share four BH domains plus a TM region. The pro-apoptotic family members can be sub-divided into two subgroups: the multi-BH domain proteins, and the BH3-only proteins. Multi-BH domain proteins contain up to four BH domains and include Bax, Bak, Bok and Bcl-x_S, which include a TM region, and also Bcl-g_L and Bfk which lack a TM region. The BH3-only proteins contain only the BH3-domain and include Bad, Bik, Bid, Hrk, Bim, Puma, Noxa and Bmf. Bik and Bim also contain a TM region (taken from Kelly, P.N. et al., Cell Death & Differentiation, 2011).*

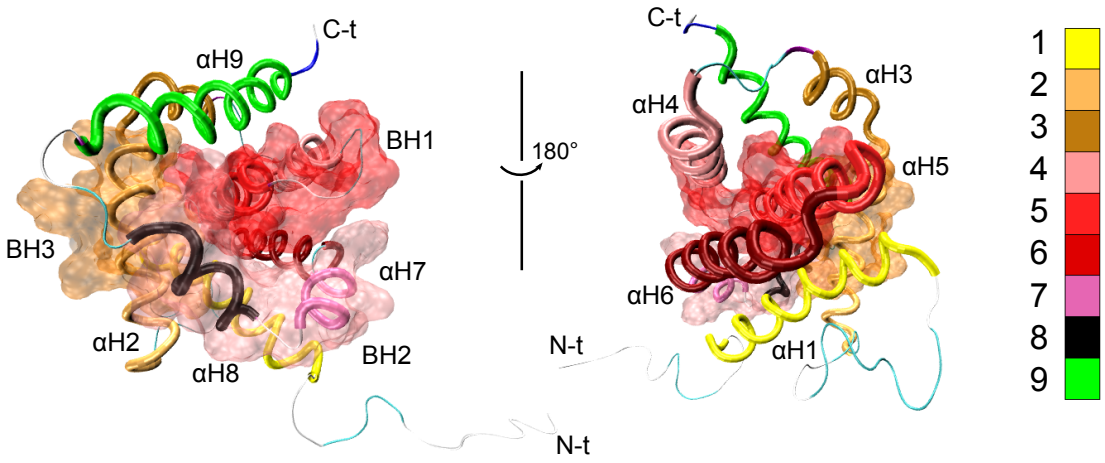


Figure 1.2: *Bax 3D structure (PDB ID: 1F16). The nine α-helices are shown as tubes. The location of BH1-3 is indicated by the transparent surfaces. Labels indicate the N-terminal (N-t) and C-terminal (C-t) ends of the protein.*

by the C-terminus of H9 (I175-W188) which could act as a regulating gate for the access to this region [47]. Also, the N-terminal domain of Bax might be implicated in mitochondrial targeting. The first 20 N-terminal residues, that form a loop with a turn (M1-T14; LpNt), and the first turn in H1 (S15-M20) might hinder exposure of a mitochondrial targeting sequence (M20-R37) [48, 49]. Previous studies determined that residues G12-A24 are exposed when Bax is treated with detergents, submitted to high temperature conditions and also when it is activated [50, 51, 52]. Also, the C-terminal regions of αH1 as well as αH2 (which contains the BH3 domain) and αH5, should be exposed during Bax activation [46, 48]. There are other Bax residues, namely K64, I66 and L70 (αH2) and S184 (αH9) that are associated with a function that would imply their location on the surface in active or oligomeric Bax. In particular, S184 determines the interaction of Bax with the mitochondria as a part of the activation mechanism [52]. Thus, there are residues that have been suggested to be close to the Bax homo-oligomer interface, such as A24 (αH1), R37, L47 (Lp1-2), D68, E69 (αH2), N73 (Lp2-3), M79 (αH3), A82 (Lp3-4), R134, R145 (αH6) and K189 or K190 (αH9) [50].

Interactions of Bax with Bcl-2 family members are primarily mediated by a BH3 domain which interacts with a hydrophobic groove in Bax formed by BH1-3 [53, 54, 55]. Other reports propose that Bax posesses an adittional site to interact with the

BH3 domains of other proteins. This site includes the residues K21, Q28, Q32 (αH1), and E131, R134 (αH6) [28, 56]. Ding *et al.* found that the interface of the Bcl-2/Bax heterodimer is likely to be formed by the BH3 region of Bax and the BH1-3 groove of Bcl-2 [57]. The hydrophobic C-terminal helix (H9) is located in this groove in the native conformation of Bax. Moreover, Lalier *et al.* identified D68, E69 and D71 as the residues involved in the binding of Bax to Bcl-2 and Bcl-x_L [52]. There are several proposed models for the molecular interactions of Bax with other proteins and the MOM [28, 50, 51, 52, 56, 58, 59].

Regarding the structural stability of Bax, several interactions have been detected that contribute to strengthening the conformation of this protein [51, 52]. Lalier *et al.* described a motif in Lp1-2 (E44-A46) that seems to establish bonds with residues in αH6 (I133 and M137). Furthermore, the disruption of the native electrostatic interaction between D33 and K64 leads to exposure of normally hidden epitopes (BH3 and the C-terminal part of H1 constituted by residues 24-33) [52].

Additionally, Bax has been submitted to proteolysis at different experimental conditions. At 4°C, this proteolysis yielded a fragment lacking 16 C-terminal residues as the major cleavage product, but fragments lacking 18, 19 and 21 C-terminal residues were also observed. This data led the authors to propose that αH9 is not tightly attached to the protein, at least at the cleavage points, whereas the N-terminus was resistant to proteolysis. Furthermore, C-terminus proteolytic cleavage sites were not found in the case of detergent-induced oligomers, but cleavage was observed after S4, M38, L45, and less frequently after A81 and A82. Moreover, a similar degradation pattern was identified at 70°C, whereas at 80°C, a mixture of N-terminal (cleaved after A46) and C-terminal (mainly cleaved after F176) truncated Bax was found [51].

1.1.6 Bax in apoptosis and disease

Reduced expression of Bax was shown to be associated with poor response rates to chemotherapy and shorter survival in patients with metastatic breast adenocarcinoma [60]. Moreover, low Bax expression was found to correlate with resistance against apoptosis [61]. On the contrary, Bax overexpression has been shown to enhance the apoptosis induced by radiation and chemicals in cancer cells [62, 63, 64], and also to follow cerebral ischemia in neurons [65].

During normal apoptosis, Bax expression is upregulated [66, 67] and its structure is activated [68], which results in its migration toward the MOM [69]. Afterwards, Bax is inserted in the MOM, and is capable of making oligomers [11, 70, 71] to permeabilize the MOM [72]. Experiments performed by Wood *et al.* allowed them to suggest that Bax driven MOM permeabilization would also give Bax access to

the intermembrane space where it would undergo proteolysis by calpain [73]. The resulting p18 Bax cleavage fragment may then adopt a conformation that allows it to readily attack and disrupt the inner mitochondrial membrane resulting in collapse of the mitochondrial inner membrane potential. This cleavage of Bax during apoptosis would therefore increase the intrinsic cytotoxic properties of Bax and enhance its cell death function at the mitochondria.

1.1.7 Modeling and simulation of biomolecules

The amount of available biological data has increased in recent years due to the genomic projects [74]. This growth (Fig. 1.1.7) is paralleled by an increase in expectations for new medical, pharmacological, environmental and biotechnological discoveries. It has thus become essential to interpret the data and translate the monodimensional information encrypted in the genomes into a detailed understanding of its biological meaning at the phenotypic level [75].

Molecular modeling is the science and art of studying molecular structure, function and interaction using computational tools. The model building can be done at different levels of sophistication, which usually is in compromise with the accuracy and extension of the simulations that can be realized. Molecular modeling has been used to study protein folding and the interaction of proteins with other molecules, among other applications. The refinement of experimental structures from NMR or X-ray crystallography, is also a component of protein molecular modeling [76].

The computational studies at electronic level which involve quantum chemistry would use, in principle, the relativistic, time-dependent Schrödinger equation to describe molecular systems since it describes their properties with high accuracy [77]. Alas, systems more complex than the equilibrium state of a few atoms cannot be handled at this *ab initio* level. Several approximations have been proposed that allow the study of larger systems at the cost of lowering the accuracy by introducing empirical information. These levels of theory include semi-empirical quantum mechanics, empirical (molecular) mechanics, molecular dynamics, Monte Carlo, free energy and solvation methods, structure/activity relationships (SAR), chemical/biochemical information and databases, and many other established procedures.

1.1.8 Biological sequence analysis

The most reliable way to determine the structure or function of a biomolecule is by direct experimentation. However, it is easier and therefore cheaper and faster to obtain the DNA sequence of the gene corresponding to a protein than it is to experimentally

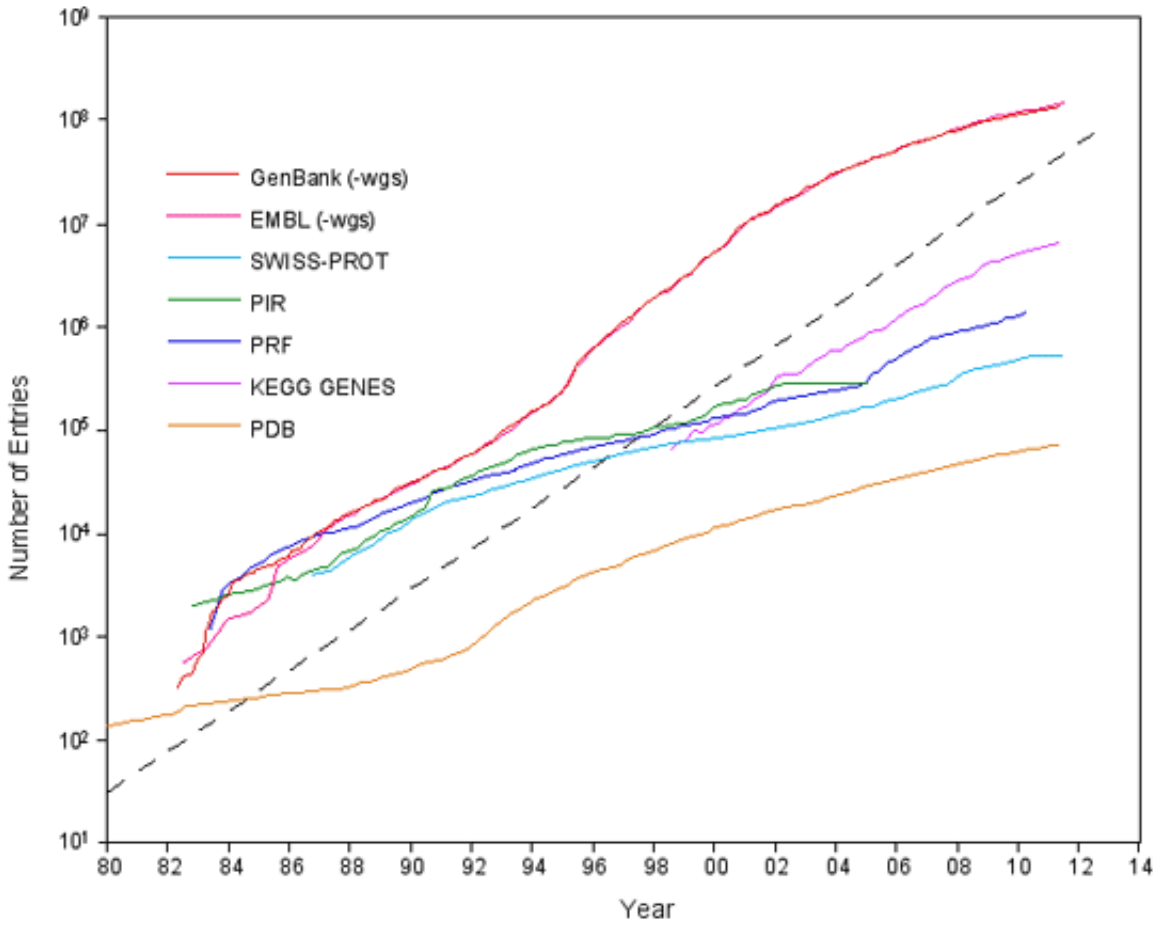


Figure 1.3: *The number of available DNA sequences (GenBank) have outnumbered that of protein sequences (SWISS-PROT) by one or two orders of magnitude. Meanwhile, there is one order of magnitude more protein sequences than 3D structures (PDB), from* `http://www.kanehisa.jp/ en/db_growth.html`.

determine its function or its structure. In the evolutionary process by which proteins acquire their structure and function, new sequences are adapted from pre-existing sequences. It is hence possible to transfer available knowledge about structure and/or function from a protein to a newly found protein, when a significant similarity is found between both sequences [78]. There are several computational methods that aim to capture the complexity contained in the large amount of biological sequences that is available and continuously growing.

The first step when using computational tools to search for similarity between sequences is finding the best alignment between them. Several algorithms for this purpose have been developed and made available to the scientific community [79, 80].

1.1.9 Molecular dynamics simulations

Due to the large number of atoms in a system composed of a water-solvated protein, molecular dynamics (MD) simulations offer a reasonable compromise of accuracy and large enough simulations to study the stability of the protein structure. Hence, MD simulations have become a common tool in theoretical studies both of simple liquids and large biomolecular systems such as proteins or DNA in solution [81].

In this level of theory, classical mechanics are used to describe the motion of atoms. A potential function is defined with parameters that are often obtained experimentally. This potential function is given by (1.1) and is used to calculate the force on each atom i, where $i = 1, \ldots, N$, for a system with N atoms. With these forces, Newton's equations of motion (1.2) are integrated to calculate the trajectories of all the atoms in the system simultaneously in small time steps.

$$F_i = -\frac{\partial V}{\partial r_i} \tag{1.1}$$

$$m_i \frac{\partial^2 r_i}{\partial t^2} = F_i \tag{1.2}$$

The evolution of the system is followed for a number of iterations, and selected state variables are monitored and adjusted such as temperature and pressure. The output of the simulation includes the coordinates of the atoms and the energy of the system, which can then be used to perform several analyses [82].

1.1.10 Molecular dynamics studies of the Bcl-2 family

Computational studies have been useful for providing valuable information about the structure and function of Bcl-2 family proteins. Several MD simulations involving

one or more of the members of this family of proteins have been reported [83, 84, 85, 86, 87, 88, 54, 89, 90].

In some of these studies, MD simulations of one monomer of one member of the family in aqueous solution were performed in order to study the dynamics of the peptidic chain, usually focusing on residues experimentally found to be instrumental in the structural stability and function of these proteins [84, 89, 88, 91].

Other MD simulations designed to yield the interaction between members of the Bcl-2 family have used systems constructed with a monomer of one of the members and a BH3 fragment which could come from a different member of the family [54, 90].

Systems with nothing but one BH3 fragment have also been simulated to learn about the dynamics of these important helical BH3 domains and the possible correlation between their stability and binding affinity [87].

Particular atention has been given to the interaction of Bcl-2 family proteins and other molecules. There have been such studies both aimed to elucidate the molecular basis of experimentally reported interactions [83, 89] and also to search for Bcl-2 family inhibitors [85, 86].

1.1.11 Folding and unfolding of proteins

Proteins can include hundreds or even thousands of amino acids bonded one after another and folded to maintain a particular 3D structure. The vast majority of proteins have not been found to exist as elongated threads, but rather as compact objects with complex folds. This folding creates a 3D structure for the protein that, however only marginally stable (5-12 kcal), must be completed in less than a minute in many observed examples [92]. A number of peptidic chains have been shown to fold without any special biological machinery, which led to the conclusion that all the information required in those cases for reaching the correct folded structure of the molecule at the required speed is contained in the amino acid sequence of the protein [93].

Equilibrium unfolding experiments have been widely used to gain insight into protein stability and the nature of the species present along the unfolding coordinate. Circular dichroism and steady-state fluorescence have been useful probes to study the global features of the equilibrium unfolding pathways, and NMR spectroscopy provides residue-level details for all the species present along the folding coordinate [94]. The denaturation process of proteins can be induced by increasing the temperature or the pressure, or by adding denaturants to the solvent, such as urea, guanidinium hydrochloride and alcohols, among others [95].

A very powerful and convenient experiment for studying temperature-induced

conformational transitions in biological systems is differential scanning calorimetry (DSC). In this technique, one of two identical cells in the instrument is filled with the sample under investigation and the other is filled with a reference sample. These cells are heated up by identical heat flows. If the two samples are identical in their thermodynamic properties, the temperature inside the cells will change identically; however, if the samples differ in such properties, a difference in the temperatures between them will appear. A special electronic system changes the heating regime of the cells in such a way that the temperatures inside them are equalized. The value of the additional (differential) heat flow required for such temperature compensation is measured by the scanning microcalorimeter. The temperature dependence of the differential heat flow (in power units) is called a thermoanalytical curve and this reflects a difference in the properties of the sample under investigation and the reference sample. Such differences can arise from changes in the phase state, structure and heat capacity of the samples [96]. DSC is regarded as the most effective and commonly used method to study the thermal unfolding of proteins [97].

1.1.12 Thermal unfolding simulations

MD simulations at high temperature can be used to study protein unfolding since these conditions make possible to observe the accompanying conformational transitions on the nanosecond time scale. MD simulation studies performed with many different proteins in which very good agreement has been obtained between low-temperature experiments and high-temperature simulations, suggest that the energy landscape in both cases is similar [98]. Interestingly, the regions that fluctuate greatly in the native state are usually the earliest regions to unfold, which suggests that thermal denaturation would involve the selective excitation of certain specific collective motions. There has been growing acceptance of simulations as useful tools for the interpretation of experimental protein folding studies and even their amenability to validation by comparison with experiments.

1.2 Motivation

Bax has been submitted to several biophysical, chemical, and computational studies [51, 58, 99, 48, 100, 56, 88, 89, 90, 47, 68, 46, 59, 50, 101]. Nevertheless, the atomic-level description of Bax stability remains to be fully explored. The investigation of stability and structural change of Bax has many biotechnological and biomedical applications. Foremost, it can provide the knowledge required for therapy and drug

design for the treatment of cell-death related diseases, where efforts to develop anti-cancer agents that block anti-apoptotic members of the Bcl-2 have been reported [102]. Additionally, it can provide the foundation for the future design of biosensors based on the modulations of the permeability of a bilipidic membrane. Membrane based biosensors have been developed with ion channel switches to detect molecules such as antibodies and have been successfully used in rapid detection of influenza in clinical samples [103].

Due to the computer resources recently available to our research group, *in silico* studies become a cost efficient way of investigating the structural properties of proteins. MD simulations are a powerful theoretical approach that can successfully complement and extend the experimental results and reproduce them with reasonable accuracy [98, 104]. In particular, MD simulations at high temperature conditions have been widely and successfully used to study the unfolding of several proteins [88, 105, 106, 107, 89, 108, 109].

1.3 Objectives

The main goal of this thesis is to investigate the structural stability of human Bax by means of *in silico* studies such as sequence analysis and MD simulations at different temperatures.

The specific objectives are:

1. Perform sequence analysis on the available sequences of Bax to identify the regions most likely involved in the function and stability of this protein.

2. Construct the simulation box.

3. Minimize the energy of the system.

4. Perform MD simulations of Bax at 300, 400 and 500 K and investigate the stability of this protein.

5. Analyze the results by means of the calculation and interpretation of the Root Mean Square Deviation (RMSD), Radius of gyration (Rg), Solvent Access Surface Area (SASA), free energy of solvatation (ΔG), Root Mean Square Fluctuations (RMSF), minimum distances between groups of interest, evolution of the secondary structure, changes in the hydrogen bond network and the clustering of the structures from the MD trajectories obtained. Additionally, by calculating and inspecting the normal modes of the initial and final structures.

Chapter 2

Materials and methods

Molecular modeling and simulation using computational tools has been helpful to study protein structure and function. Though the models available represent a highly-simplified version of the cellular environment, or even in vitro conditions, systematic studies based on tractable quantitative tools can help discern patterns and add insights that are otherwise difficult to observe [76].

2.1 Universal Protein resource

The Universal Protein resource (UniProt) aims to provide a comprehensive, high-quality and freely accessible resource of protein sequence and functional information. It is the result of a collaboration between three institutions: the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR). Around 90 people from all three institutions contribute through different tasks such as database curation, software development and support. Hence, Uniprot is not only a repository, but a powerful collection of services to extract information from the biological data generated all over the world [110].

2.2 Cluster alignment

The EBI offers a service for multiple sequence alignment called Clustal (Cluster alignment) in a website. The algorithm starts with a pairwise alignment that is used to build a phylogenetic tree. The multiple alignment is based on the obtained phylogenetic tree. The code is under constant development and optimization which have resulted in a solid reputation for offering high quality alignments very rapidly [111].

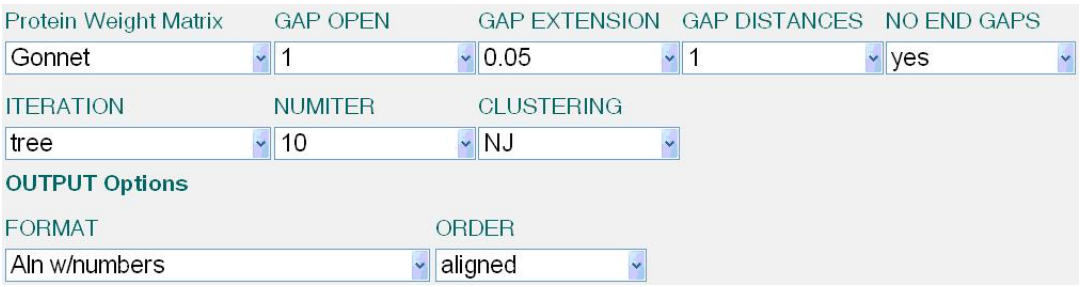


Figure 2.1: *Gap opening, extension and separation penalties are lower than the default values and improvement of the alignment is requested in every iteration. The maximum available number of iterations is requested.*

2.3 Analysis of available Bax sequences

A total of 23 sequences of Bax were found in the Uniprot database, the labels associated with these sequences were diverse: 'Apoptosis Regulator Bax', 'Protein Bax', 'Bax Protein', 'Bax-alpha protein' and 'BCL2-associated X protein'. The species of origin ranged from fish (v.g. *Salmo salar*) to amphibia (v.g. *Xenopus tropicalis*) to mammals and *Homo Sapiens* (Human). Both manually reviewed and computer-annotated sequences were considered, provided that they were complete. The seven isoforms of Human Bax found in the repository were included in the multiple alignment. A second multiple alignment was performed with only the Human Bax isoforms.

The multiple alignment was performed with the web server ClustalW version 2 [111]. A slow alignment was requested with the parameters shown in Fig. 2.1.

2.4 GROMACS

Computational MD simulations of biomolecules have been performed since the late 1980's [112]. The algorithms and software implementations have been under continous development both to improve the quality of the results and to reduce the computation time in order to allow the simulation of larger systems for longer periods of time. GROMACS 4 is a molecular simulation toolkit that achieves high performance on single processors and also scales very well on parallel machines. This is accomplished by code features that include a minimal-communication domain decomposition algorithm, full dynamic load balancing, a state-of-the-art parallel constraint solver, and efficient virtual site algorithms that allow removal of hydrogen atom degrees of free-

dom [81]. GROMACS provides also a wealth of tools for analyzing the results of a simulation.

The modeling tools provided by GROMACS that were used to perform the MD simulations include the force fiel known as Optimized Potentials for Liquid Simulations for All Atoms (OPLS-AA), the Simple Point Charge (SPC) water model, the Particle Mesh Ewald (PME) algorithm for calculation of electrostatic interactions and the LINear Constraint Solver (LINCS) algorithm to introduce holonomic constraints in the model. This models and algorithms are briefly described below.

2.5 Force field

A force field is a set of equations and parameters that describe the interatomic interactions and the mechanics of deformation of the molecules in a MD simulation. Hence, the election of force field is an important decision in a theoretical study.

2.5.1 Optimized potentials for liquid simulations for all atoms

The Optimized Potentials for Liquid Simulations for All Atoms (OPLS-AA) model was developed from the start with the idea of using it for MD simulations of biomolecules [113] and was later improved for the same purpose [114]. OPLS-AA is based on the AMBER force field, which uses equation (2.1) described in [115]. OPLS focuses on the intramolecular nonbonded interactions and replaces the corresponding terms in AMBER with equation (2.2) given in [114].

$$\begin{aligned} E_{total} &= \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + cos(n\phi - \gamma)] \\ &+ \sum_{i<j}[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}}] + \sum_{H_bonds}[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{1}0}] \end{aligned} \tag{2.1}$$

$$\Delta E_{ab} = \sum_{i}^{on\ a}\sum_{j}^{on\ b}[\frac{q_i q_j e^2}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{1}2} - \frac{C_{ij}}{r_{ij}^{6}}] \tag{2.2}$$

2.5.2 Simple point charge water

Several models of water for MD simulations have been proposed, none of which can correctly reproduce all physical properties of water [116]. The Simple Point Charge

(SPC) model considers three interaction sites centered on the nuclei. The intermolecular Coulomb energy is provided by partial charges in each site and a Lennard-Jones interaction between the oxygens is included but the intramolecular distances were rigid [117]. While more elaborated models, such as TIP5P, provide a better approximation for water behaviour, they are also considerably more time-consuming to calculate. Therefore, SPC water is a sensible choice when the focus is not on the monitoring of bulk water properties but what is important is the water-protein interface [118].

2.6 Calculation of electrostatic interactions

Direct calculation of the Coulomb energy scales as N^2. The Particle Mesh Ewald (PME) algorithm separates the calculation in a short range term and a long range term. The short range term is calculated directly since it sums quickly in real space. For the long range electrostatics, charges are assigned to a grid using cardinal B-spline interpolation. The Fourier transform of this grid is performed with a 3D FFT algorithm and the reciprocal energy term obtained by a single sum over the grid in k-space. The potential at the grid points is calculated by inverse transformation. The PME algorithm scales as $N \log(N)$ which represents an important reduction in computation time for large systems [119, 120, 121].

2.7 Temperature coupling

In the MD simulations performed, a reference temperature was set and the system is adjusted gradually to reach the given temperature by means of the Nosé-Hoover coupling. In this method, the expression for the total energy of the system is extended by introducing a heath source and a friction term in the equations of motion. The Nosé-Hoover thermostat [122, 123] is a model to reproduce the canonical phase-space distribution by modifying the equations of motion to include a non-Newtonian term in order to maintain the total kinetic energy constant. The modified equation of motion is given by (2.3).

$$\frac{dv(t)}{dt} = \frac{F(t)}{m} - \zeta v(t) \tag{2.3}$$

Where ζ is the thermodynamic friction coefficient, given by (2.4).

$$\frac{d\zeta(t)}{dt} = \frac{1}{Q}[\sum mv(t)^2 - (X+1)k_B T] \tag{2.4}$$

Where Q is a parameter that has the dimensions of $[energy \times time^2]$ and determines the time-scale of the temperature fluctuation and X is the number of degrees of freedom.

The acceleration of the atoms is gradually adjusted and thus the reference temperature is approached in an oscillatory relaxation [122, 123]. When using this method in GROMACS, the parameters that have to be defined are the groups that will be submitted to the temperature coupling (`tc-grps`), the period (`tau_t`) of the oscillations of kinetic energy between the system and the heath source and the temperature of reference (`ref_t`). For example, Box 2.1 shows the parameters used for the simulation at 500 K.

```
Tcoupl              =  nose-hoover
tc-grps             =  Protein  SOL Na+
tau_t               =  0.1      0.1 0.1
ref_t               =  500      500 500
```

Box 2.1: *Parameters for the Nosé-Hoover temperature coupling for the simulation at 500 K.*

```
Pcoupl              =  Parrinello-Rahman
Pcoupltype          =  isotropic
tau_p               =  2.0
compressibility     =  4.5e-5
ref_p               =  1.0
```

Box 2.2: *Parameters for the Parrinello-Rahman pressure coupling used for the MD simulations performed.*

2.8 Pressure coupling

Since it was important in our methodology to monitor the effect of the temperature, the pressure of the system was kept constant. This was achieved with Parrinello-Rahman coupling which is similar to the Nosé-Hoover temperature coupling and is useful for maintaining NPT conditions (number of particles, N, pressure, P, and constant temperature, T). Instead of a heath source, the volume of the simulation

box is allowed to fluctuate and a friction term is added to the equations of motion to adjust the acceleration of the atoms [124, 125]. In GROMACS, the parameters to set for this method are the type of pressure desired (`Pcoupltype`), the desired period for the pressure fluctuations (`tau_p`), the approximate isothermal compressibility, and the reference pressure (`ref_p`). For instance, Box 2.2 shows the settings used for all the production MD simulations performed.

2.9 Linear constraint solver

The time step in MD simulations is limited by high-frequency oscillations. When these oscillations have also a small amplitude, such as in bond stretches, the introduction of holonomic constraints can allow the use of a larger time step. These constraints have been handled as a set of equations that can be written in a matrix representation. The constraint equations are then used to reset the bond lengths to prescribed values with an iterative method. The LINear Constraint Solver (LINCS) is an algorithm that implements an efficient solver for the matrix equation, a velocity correction that prevents rotational lengthening and a length correction that improves accuracy and stability [126]. LINCS has also been found well suited for parallel implementations and is available in GROMACS 4 [82].

2.10 High temperature simulation of a Bax monomer

The starting structure (native) for MD simulations of the Bax monomer was obtained from the first NMR model located it the Protein Data Bank (PDB) by Suzuki *et al.* (PDB ID: 1F16) [46]. All MD simulations were performed with GROMACS 4 [81], using the OPLS-AA force field [113]. The protein was solvated in a rectangular box of SPC water [117], with a minimum distance of 1 nm from the protein to the edge of the box (Fig. 2.2, left). To obtain a neutral total charge in the system, 3 Na^+ counterions were added (Fig. 2.2, right). The total size of the system was of 38,647 atoms, including 11,889 water molecules.

During energy minimization, the steepest descents algorithm was used and the convergence was reached in 200 steps. Further equilibration of the system was accomplished in 5,000 steps (10 ps) of MD with restricted protein atoms and NVT conditions with a box size of $9.44551 \times 7.45274 \times 7.68113$ nm (Fig. 2.3, left). Afterward, 2 ns of MD were performed at 300 K without any atom fixing and under NPT

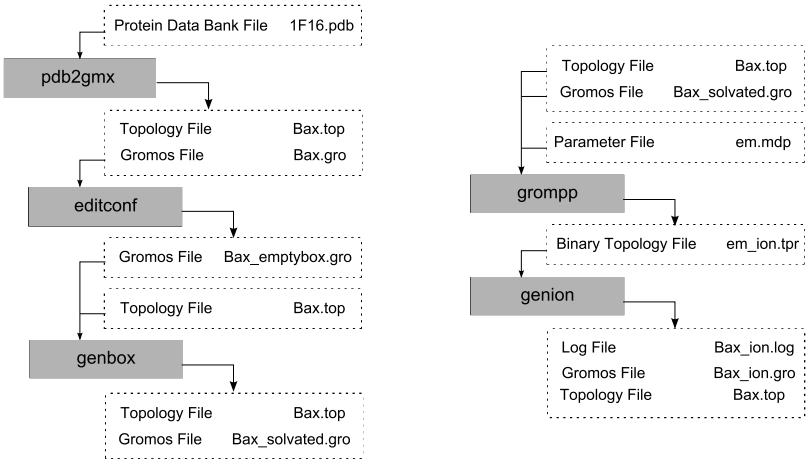


Figure 2.2: *Construction of the simulation Box. The left image shows the flow diagram of the solvation of the first model from the file 1F16.pdb in SPC water. The image on the right is the flow diagram for the addition of counter ions to the system.*

conditions, which were used for the following MD simulations as well, so that the size of the box could fluctuate to keep the pressure at a constant value. Using the latter trajectory as a starting point, another 18 ns of production MD were obtained at 300, 400, and 500 K (Fig. 2.3, right). MD simulations were performed with a time step of 2 fs, and the coordinates for the whole system were saved every 250 steps. For all MD simulations, neighbor lists were updated every 10 steps. The leap-frog algorithm for integrating Newton equations was used, and periodic boundary conditions were applied. The PME algorithm [119, 120, 121] was used for the electrostatic interactions with a cutoff of 1 nm, and a reciprocal grid of $54 \times 42 \times 44$ cells was used with 4-th order B-spline interpolation. A single cut-off of 1 nm was used for the calculation of van der Waals interactions. Temperature and pressure coupling were performed with the Nosé-Hoover algorithm [122, 123] and the Parrinello-Rahman algorithm [124, 125], respectively.

All bonds were constrained using LINCS [126]. The complete sequence of GROMACS commands used can bee found in Appendix 1. Additionally, Appendix 2 includes all the configuration files. After stabilization of the system, in all cases, the potential energy was conserved during the MD simulations. The analysis tools included in GROMACS were used to calculate the RMSD, RMSF, Rg, and SASA. The evolution of the secondary structure was followed using the DSSP program [127]. Graphical representations of the protein were obtained using VMD [128].

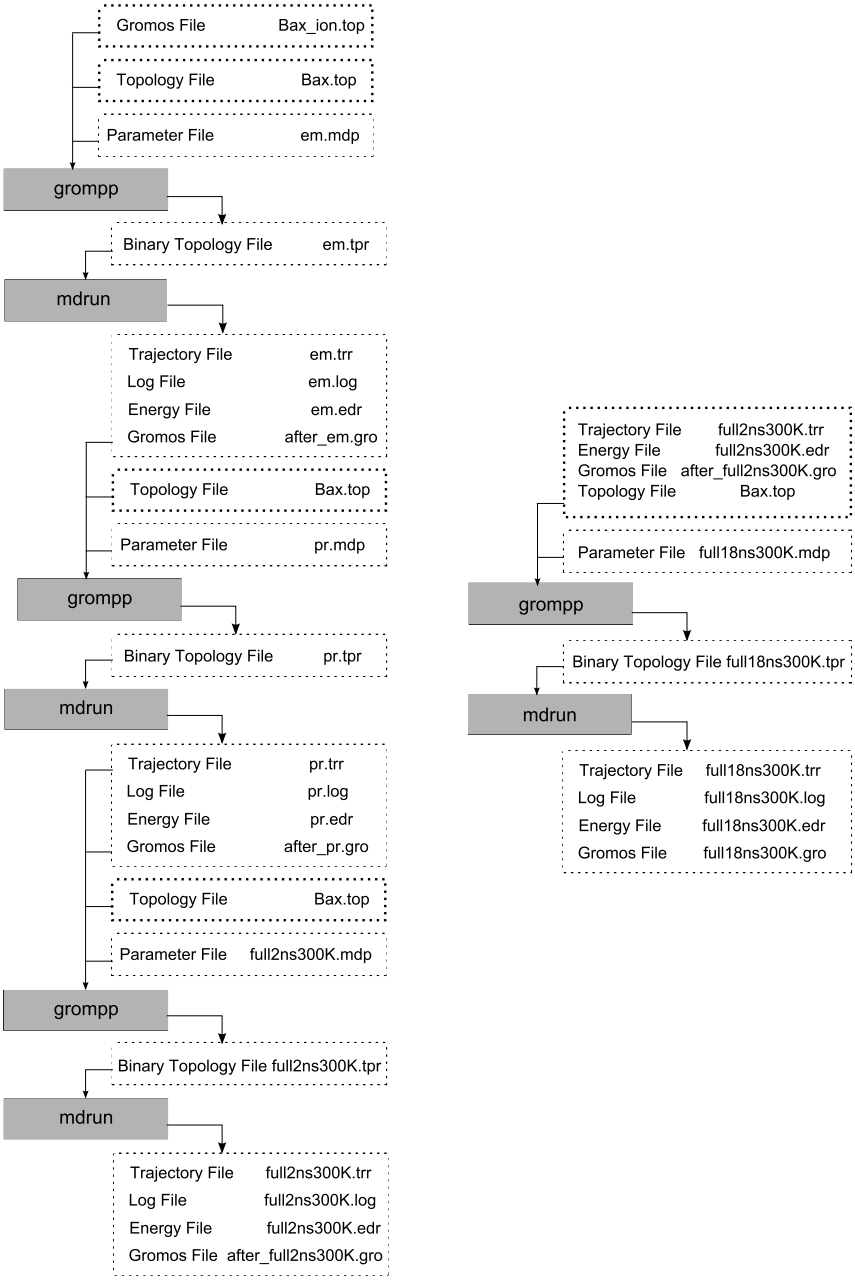


Figure 2.3: *Simulation of the system. The left image shows the flow diagram for the equilibration of the system. The image on the right shows the flow diagram of the way GROMACS was used to produce the simulations.*

2.11 Normal mode analysis

The Normal Mode Analysis (NMA) is a theoretical method for studying collective motions. It has been shown that many of the known protein motions can be approximated by applying a perturbation in the direction of at most two low-frequency normal modes of the considered protein [129].

2.11.1 *ElNémo*

The computational cost of NMA restricts its application to small proteins. *ElNémo* implements approximations that provide a fast and simple tool to compute, visualize and analyse low-frequency normal modes of large macro-molecules [130].

The approximations start with the use of a single-parameter Hookean model, where the spring constant is assumed to be the same for all interacting pairs. In this model, a cut-off radius is also considered, with a default value of 8 Å, beyond which interactions are ignored. Also, all atom masses are set to the same fixed value in the kinetic energy term.

The block approximation, also termed "rotation-translation-block" (RTB), groups several residues to form a single super-residue, whose rigid-body rotations and translations are used as a set of new coordinates replacing the Cartesian ones. The algorithm automatically determines the number of residues to be grouped together based on the number of residues in the protein, but the user may change this setting. This approximations have been shown to have little influence on the calculation of the low-frequency modes [131].

The results include the "degree of colectivity" for each mode, which indicates the fraction of residues that are significantly affected by a given mode. A low degree of colectivity indicates that only a localized region of the protein is involved in the motion.

2.11.2 Normal mode analysis of human Bax

The NMA was performed using the web-server *ElNémo* [130], where the only change with respect to default settings was in the number of lowest frequency normal modes to be computed which was set to 10 (the default value is 5). The minimum perturbation (`DQMIN`), maximum perturbation (`DQMAX`), and step size between `DQMIN` and `DQMAX` were left at their default values of 100, 100, and 20, respectively. The advanced options were also left in their default settings (`NRBL = auto` and `CUTOFF = 8`).

Chapter 3

Results and discussion

3.1 Sequence analysis of Bax

The analysis of the 23 Bax sequences found in the Uniprot started with the construction of the histograms of amino acid content which revealed some conserved general features. Afterwards, multiple alignments were obtained for all 23 sequences and another one involving only the isoforms of human Bax. In this thesis the focus was on human Bax isoform alpha.

3.1.1 Composition histogram of available sequences of Bax

Calculating and plotting the content of the 23 available sequences of Bax allowed us to make some interesting observations. First, 11 out of the 23 sequences do not have any histidine, including 6 out of the 7 complete human Bax sequences available, the rest have a content of histidines of 4% or lower (Fig. 3.1). This, together with a low content of cystein (always below 3%) points to a sequence that has evolved to avoid interactions with metals and to have low reactivity, which correlates well with the fact that Bax is found in the cytosol in normal cells [69]. Conversely, none of the sequences lacks tryptophan. This is important since conserved tryptophans have been shown to be important for structural stability and function of a protein [132, 133, 134, 135].

3.1.2 Multiple alignment of available sequences of Bax

The output from the multiple alignment of all 23 sequences performed in Clustal is depicted in Figs. 3.2, 3.3 and 3.4.

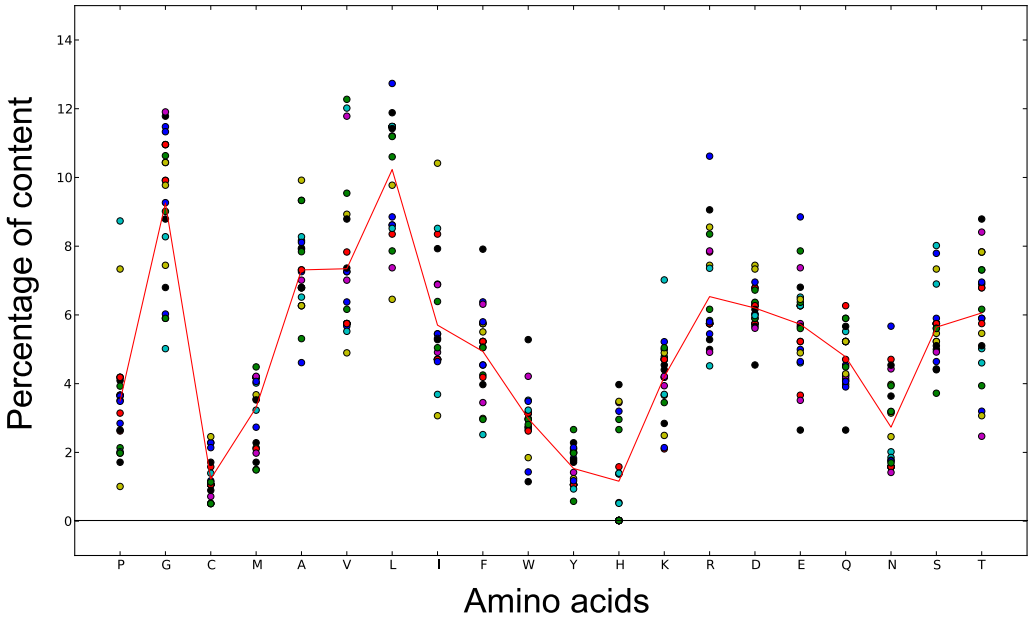


Figure 3.1: *Percentage of amino acid content of the 23 available sequences of Bax. To each sequence corresponds a different color. The lines join the average values.*

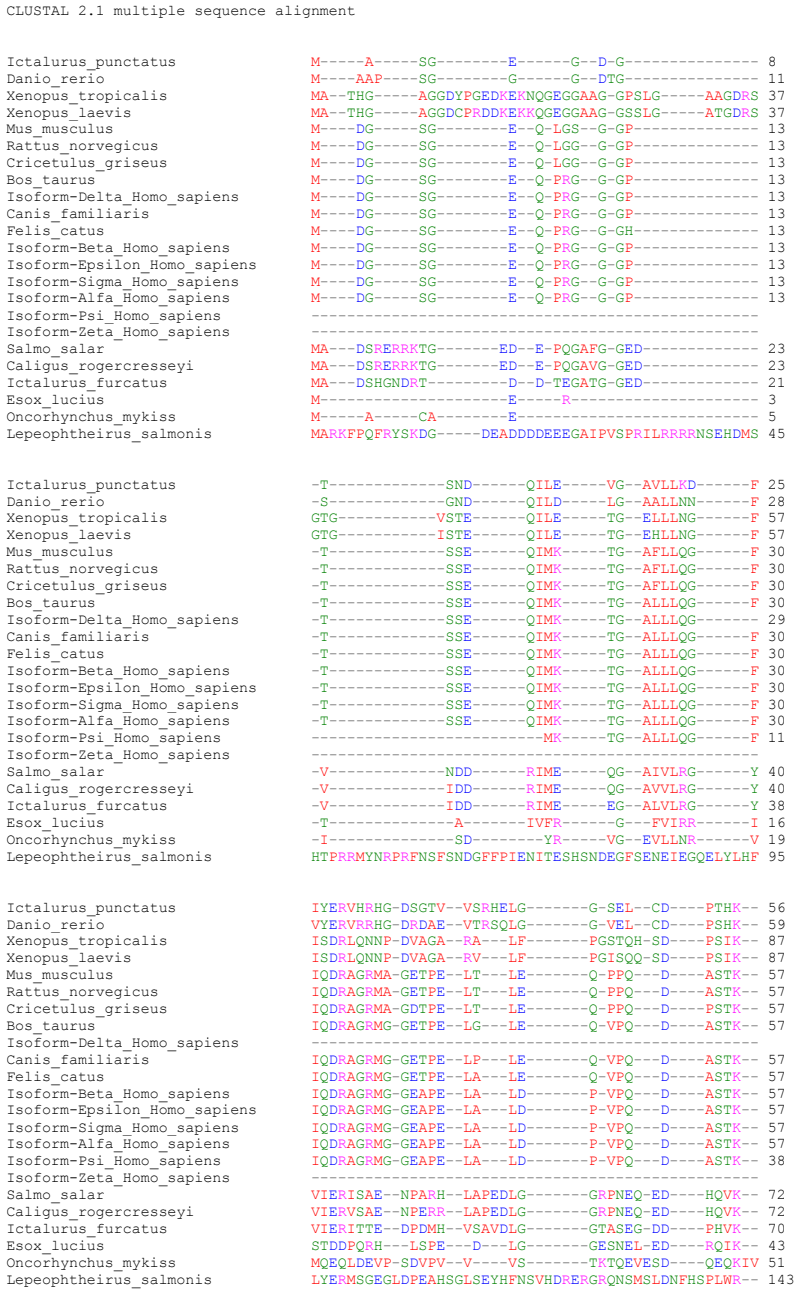


Figure 3.2: *Multiple alignment of all available Bax sequences part 1 of 3 (M1-F30 in human Bax isoform alpha).*

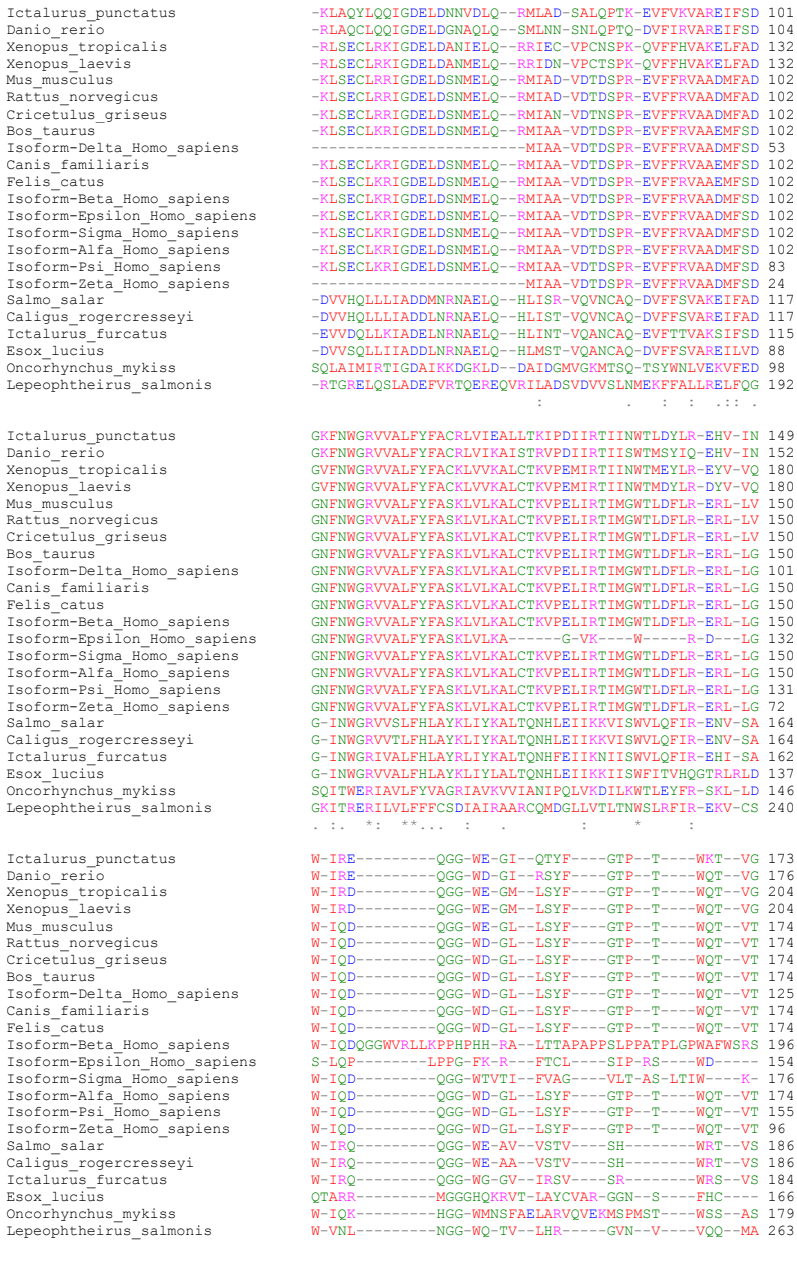


Figure 3.3: *Multiple alignment of all available Bax sequences part 2 of 3 (I31-D102 in human Bax isoform alpha).*

```
Ictalurus_punctatus              VF-LA-GVLTT-VLVMRK-M---- 189
Danio_rerio                      VF-LA-GVITT-ALVIRK-M---- 192
Xenopus_tropicalis               VF-LA-GVLTA-TIAIWK-MS--- 221
Xenopus_laevis                   VF-LA-GVLTA-SLAIWK-MS--- 221
Mus_musculus                     IF-VA-GVLTA-SLTIWKKMG--- 192
Rattus_norvegicus                IF-VA-GVLTA-SLTIWKKMG--- 192
Cricetulus_griseus               IF-VA-GVLTA-SLTIWKNMG--- 192
Bos_taurus                       IF-VA-GVLTA-SLTIWKKMG--- 192
Isoform-Delta_Homo_sapiens       IF-VA-GVLTA-SLTIWKKMG--- 143
Canis_familiaris                 IF-VA-GVLTA-SLTIWKKMG--- 192
Felis_catus                      IF-VA-GVLTA-SLAIWKKMG--- 192
Isoform-Beta_Homo_sapiens        QW-CPLPIFRS-SDVVYNAFSLRV 218
Isoform-Epsilon_Homo_sapiens     -Y-RP----CA-P-R-CRN----- 164
Isoform-Sigma_Homo_sapiens       ---------KM-G---------- 179
Isoform-Alfa_Homo_sapiens        IF-VA-GVLTA-SLTIWKKMG--- 192
Isoform-Psi_Homo_sapiens         IF-VA-GVLTA-SLTIWKKMG--- 173
Isoform-Zeta_Homo_sapiens        IF-VA-GVLTA-SLTIWKKMG--- 114
Salmo_salar                      LV-AA-VAFVA-VMVYWRKTR--- 204
Caligus_rogercresseyi            LV-AA-VAFVT-AMVYWRKTR--- 204
Ictalurus_furcatus               II-AA-VAFIA-AAVYWRRTR--- 202
Esox_lucius                      ---RS-GLLEENPLT--------- 177
Oncorhynchus_mykiss              IL-VFLGVILG-SVITWKLARRT- 200
Lepeophtheirus_salmonis          IIGMCAAVMLCCTIYIRKNL---- 283
```

Figure 3.4: *Multiple alignment of all available Bax sequences part 3 of 3 (G103-G192 in human Bax isoform alpha).*

The amino acids found to be conserved to some degree in the multiple alignment are shown in Table 3.1 (numeration with respect to human Bax isoform alpha).

Interestingly, 3/4 of the most conserved, 4/10 that conserve strongly similar properties and 6/10 that conserve weakly similar properties are in BH1 (F100-V121). Only Q153 is in BH2 (W151-F165) and none are in BH3. This result suggests that only BH1 is absolutely needed for the primary function of Bax. I80 is in αH3, which is remarkable because extra energy is required to keep a beta branched amino acid such as isoleucine in an α-helix. Hence, the role of I80 might be important for structural

Table 3.1: *Conserved amino acids in the multiple alignment considering 23 sequences of Bax from several species.*

Fully conserved (*)	**Conserved strongly similar properties (:)**	**Conserved weakly similar properties (.)**
R109, L113, F114, W139	I80, F92, V95, M99, F100, F105, V110, L120, I133, R145	R89, D98, D102, G103, N106, Y115, F116, A117, A124, Q153

Table 3.2: *Conserved amino acids in the multiple alignment considering the seven available isoforms of human Bax.*

Fully conserved (*)	**Conserved strongly similar properties (:)**	**Conserved weakly similar properties (.)**
M79-A124, K128, L132, F143, R145, L149, I152, W158	E131, T135, I136, G150, Q155, A178	T127, T140, D142, G157, S184, K189

stability, function or both. R89, F92, D98, M99 and F100 are in αH4. The unfolding of αH4 was found to play a role in the protection of the hydrophobic core of Bax in this work (Subsection 3.2.6), and these residues could be involved in this mechanism. A124 is in αH5, while 1133, W139 and R145 are in αH6. Since αH5 and αH6 belong to the pore-forming domain of Bax, these amino acids could be important for the membrane permeabilization function of Bax.

The output from the multiple alignment of the isoforms of human Bax performed in Clustal is depicted in Fig. 3.5. The amino acids found to be conserved to some degree in the multiple alignment are shown in Table 3.2 (numeration with respect to human Bax isoform alpha).

The BH1 domain (F100-V121) is fully conserved, whereas only I152, and W158 are fully conserved in BH2 (W151-F165), the less conserved Q155 and G157 are in BH2. BH3 (L59-D71) is not present in two of the isoforms. These findings support the identification of BH1 as absolutely necessary for the primary function of Bax, BH2 would have a secondary role and BH3 would not be essential for Bax to function. K128 is in αH5, E131, L132, T135, I136, T140, D142, F143 and R145 are in αH6. These residues would be likely to participate in the membrane permeabilization function of Bax. L149 and G150 are close to BH2, and they might participate in the function of this domain. S184 and K189 are in αH9 and were found in this work contributing to the structural stability of Bax (Subsection 3.2.8). A178 is also in αH9 and has been associated with the regulation of Bax activation [101].

```
CLUSTAL 2.1 multiple sequence alignment


Isoform-Delta_Homo_sapiens        -------------------MD-GS----G-EQPR-G--GG---------- 12
Isoform-Zeta_Homo_sapiens         -------------------------------------------------
Isoform-Psi_Homo_sapiens          -------------------MKTGALLLQGFIQDRAGRMGGEAPELALDPV 31
Isoform-Alfa_Homo_sapiens         MDGSGEQPRGGGPTSSEQIMKTGALLLQGFIQDRAGRMGGEAPELALDPV 50
Isoform-Sigma_Homo_sapiens        MDGSGEQPRGGGPTSSEQIMKTGALLLQGFIQDRAGRMGGEAPELALDPV 50
Isoform-Beta_Homo_sapiens         MDGSGEQPRGGGPTSSEQIMKTGALLLQGFIQDRAGRMGGEAPELALDPV 50
Isoform-Epsilon_Homo_sapiens      MDGSGEQPRGGGPTSSEQIMKTGALLLQGFIQDRAGRMGGEAPELALDPV 50


Isoform-Delta_Homo_sapiens        P--TSSEQI---MKT-G-AL---L-LQGMIAAVDTDSPREVFFRVAADMF 51
Isoform-Zeta_Homo_sapiens         ---------------------------MIAAVDTDSPREVFFRVAADMF 22
Isoform-Psi_Homo_sapiens          PQDASTKKLSECLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMF 81
Isoform-Alfa_Homo_sapiens         PQDASTKKLSECLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMF 100
Isoform-Sigma_Homo_sapiens        PQDASTKKLSECLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMF 100
Isoform-Beta_Homo_sapiens         PQDASTKKLSECLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMF 100
Isoform-Epsilon_Homo_sapiens      PQDASTKKLSECLKRIGDELDSNMELQRMIAAVDTDSPREVFFRVAADMF 100
                                                ********************

Isoform-Delta_Homo_sapiens        SDGNFNWGRVVALFYFASKLVLKALCTKVPELIRTIMGWTLDFLRERLLG 101
Isoform-Zeta_Homo_sapiens         SDGNFNWGRVVALFYFASKLVLKALCTKVPELIRTIMGWTLDFLRERLLG 72
Isoform-Psi_Homo_sapiens          SDGNFNWGRVVALFYFASKLVLKALCTKVPELIRTIMGWTLDFLRERLLG 131
Isoform-Alfa_Homo_sapiens         SDGNFNWGRVVALFYFASKLVLKALCTKVPELIRTIMGWTLDFLRERLLG 150
Isoform-Sigma_Homo_sapiens        SDGNFNWGRVVALFYFASKLVLKALCTKVPELIRTIMGWTLDFLRERLLG 150
Isoform-Beta_Homo_sapiens         SDGNFNWGRVVALFYFASKLVLKALCTKVPELIRTIMGWTLDFLRERLLG 150
Isoform-Epsilon_Homo_sapiens      SDGNFNWGRVVALFYFASKLVLKA-GVKWRDL-GSLQPLPPGFKRFTCLS 148
                                  ********************** .*  :*  ::   . .* *   *.

Isoform-Delta_Homo_sapiens        WIQDQGGWDGLL--SY-----FGT----PTWQTVT-I----FVA-G---- 130
Isoform-Zeta_Homo_sapiens         WIQDQGGWDGLL--SY-----FGT----PTWQTVT-I----FVA-G---- 101
Isoform-Psi_Homo_sapiens          WIQDQGGWDGLL--SY-----FGT----PTWQTVT-I----FVA-G---- 160
Isoform-Alfa_Homo_sapiens         WIQDQGGWDGLL--SY-----FGT----PTWQTVT-I----FVA-G---- 179
Isoform-Sigma_Homo_sapiens        WIQDQGGW------------------------TVT-I----FVA-G---- 166
Isoform-Beta_Homo_sapiens         WIQDQGGWVRLLKPPHPHHRALTTAPAPPSLPPATPLGPWAFWSRSQWCP 200
Isoform-Epsilon_Homo_sapiens      -IP-R-SWD------Y-------R----P-------------CA------ 159
                                   *  : .*                                  :

Isoform-Delta_Homo_sapiens        --VLTASLTIWKKMG--- 143
Isoform-Zeta_Homo_sapiens         --VLTASLTIWKKMG--- 114
Isoform-Psi_Homo_sapiens          --VLTASLTIWKKMG--- 173
Isoform-Alfa_Homo_sapiens         --VLTASLTIWKKMG--- 192
Isoform-Sigma_Homo_sapiens        --VLTASLTIWKKMG--- 179
Isoform-Beta_Homo_sapiens         LPIFRSSDVVYNAFSLRV 218
Isoform-Epsilon_Homo_sapiens      ----PRC----RN----- 164
                                        .    .
```

Figure 3.5: *Multiple alignment of all Homo Sapiens isoforms of Bax.*

3.2 Structural stability of a Bax monomer in aqueous media

Several structural analysis were performed on the results given by the simulations performed in a search for insights into the effect of high temperature conditions on the structural stability of Bax as described in the following subsections.

3.2.1 Overall stability

The time course evolution of several Bax structural features at 300, 400 and 500 K during 20 ns is analyzed in this section. Fig. 3.6 depicts α-carbons RMSD, which is a measure of backbone conformational movement. Convergence for 300, 400 and 500 K is reached after 14 ns of MD simulation. Initially, RMSD values increase in the MD simulations during the heating of the system. The values were approximately bounded between 0.3 and 0.6 nm at 300 and 400 K. RMSD values at 500 K show two stages: the first from 5 to 11 ns (RMSD between 0.8 and 1.0 nm) and the second from 11 to 20 ns (RMSD between 1.0 and 1.2 nm).

Regarding the Rg values, the starting value is about 1.8 nm, which reflects an expanded polypeptide, mainly due to the initial elongation of LpNt (Fig. 3.7). During the MD simulations at the three different temperatures studied, the Rg values decrease during the equilibration period. The curves at 300 and 400 K reach values around 1.6 nm after 5 ns, whereas the Rg values at 500 K fluctuate from 1.6 to 1.8 nm. The Rg values of Bax decrease at 300 and 400 K during the equilibration stage mainly due to LpNt movement toward the globule, but the compactness of the conformation is maintained from 5 to 20 ns. In this sense, at 500 K, the Rg fluctuating values are evidence of movements of the LpNt toward and away from the globule, as well as conformational variation of the whole protein.

The SASA values start close to 170 nm^2 and decrease to values around 146 nm^2, showing small fluctuations at 300 and 400 K after 5 ns. In both curves, an increase in SASA is observed: at 300 K, this increase is of no more than 20 nm^2 (17-19 ns), whereas at 400 K, the value of SASA reaches more than 160 nm^2 (16-18 ns). In both cases, SASA recovers its original value by the end of the MD simulations. At 500 K, the SASA values display two phases. In the first one (5-11 ns), the SASA values do not vary noticeably (around 160 nm^2), while a second phase involves large increases in the SASA values, showing four peaks: 195.9 nm^2 (11.2 ns), 236.3 nm^2 (13.0 ns), 250.1 nm^2 (15.5 ns), and 277.7 nm^2 (16.5 ns). After 16.5 ns, the SASA values decrease to about 185.9 nm^2 (Fig. 3.8).

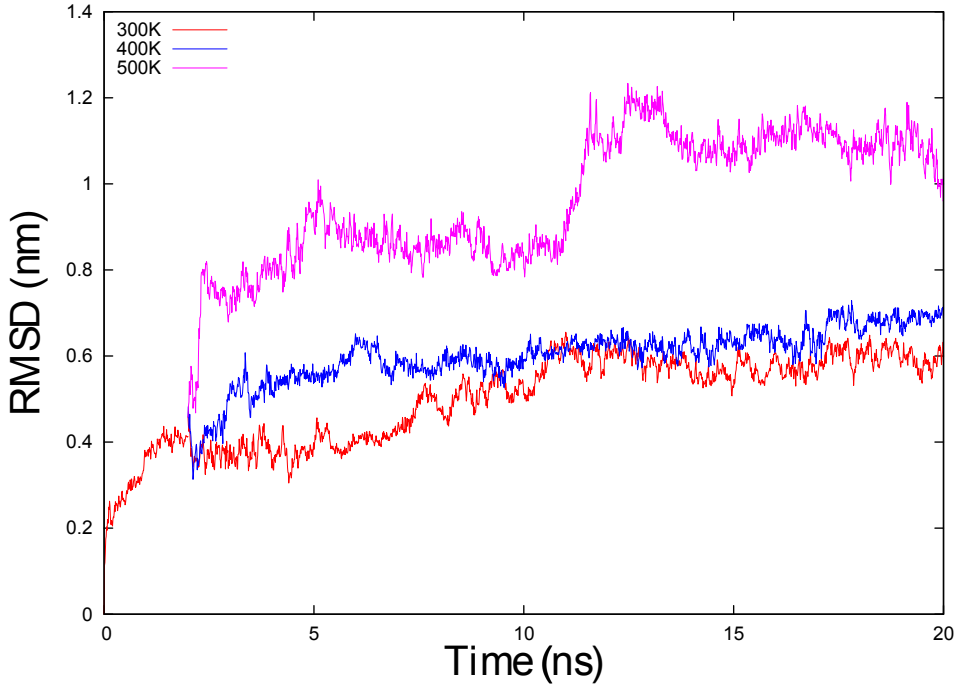


Figure 3.6: *Time evolution of the Cα RMSD of the protein in the MD simulations performed.*

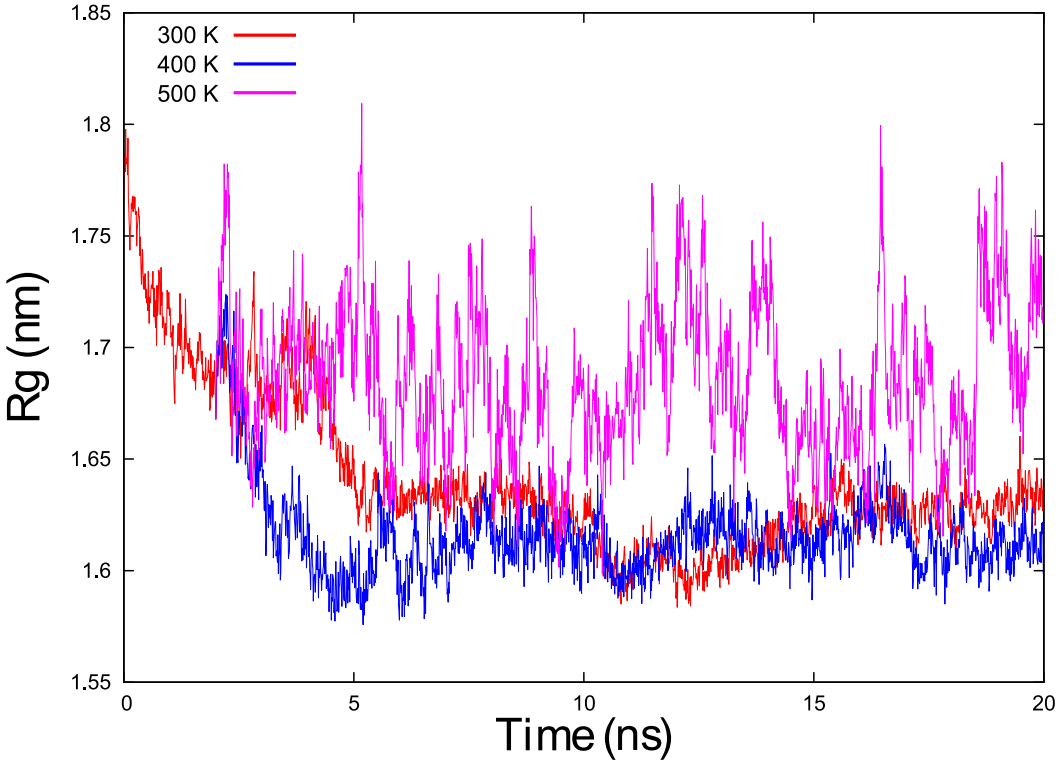


Figure 3.7: *Time evolution of the Radius of gyration (Rg) of the protein in the MD simulations performed.*

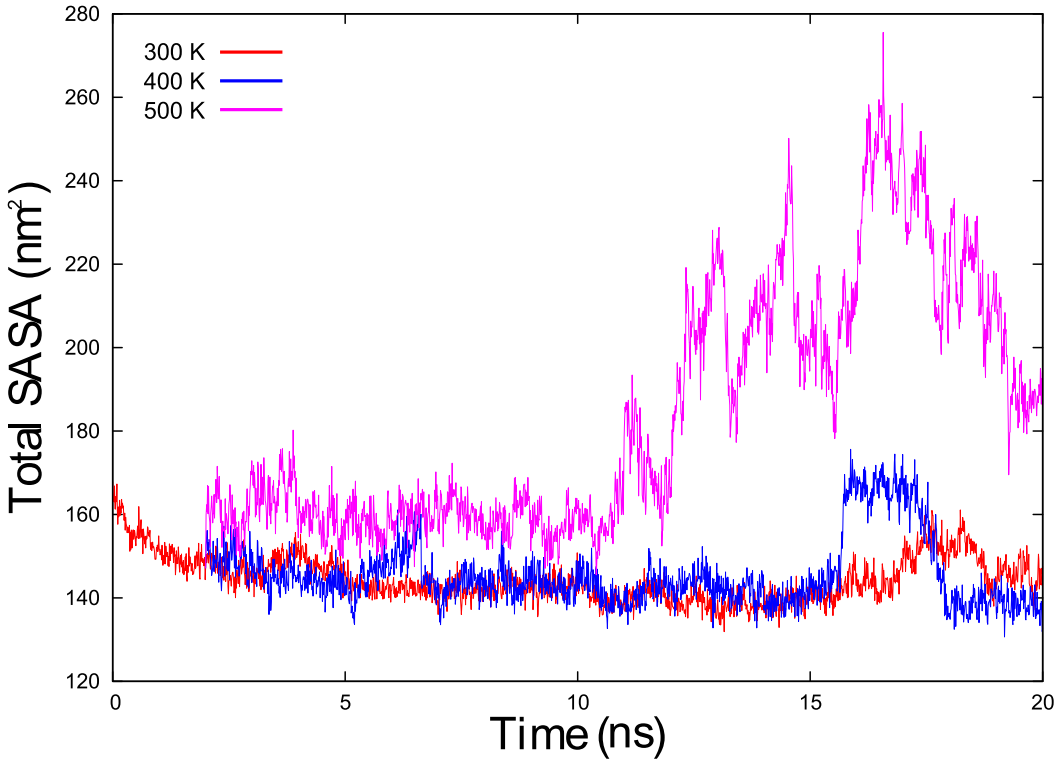


Figure 3.8: *Time evolution of the Solvent Accessible Surface Area (SASA) of the protein in the MD simulations performed.*

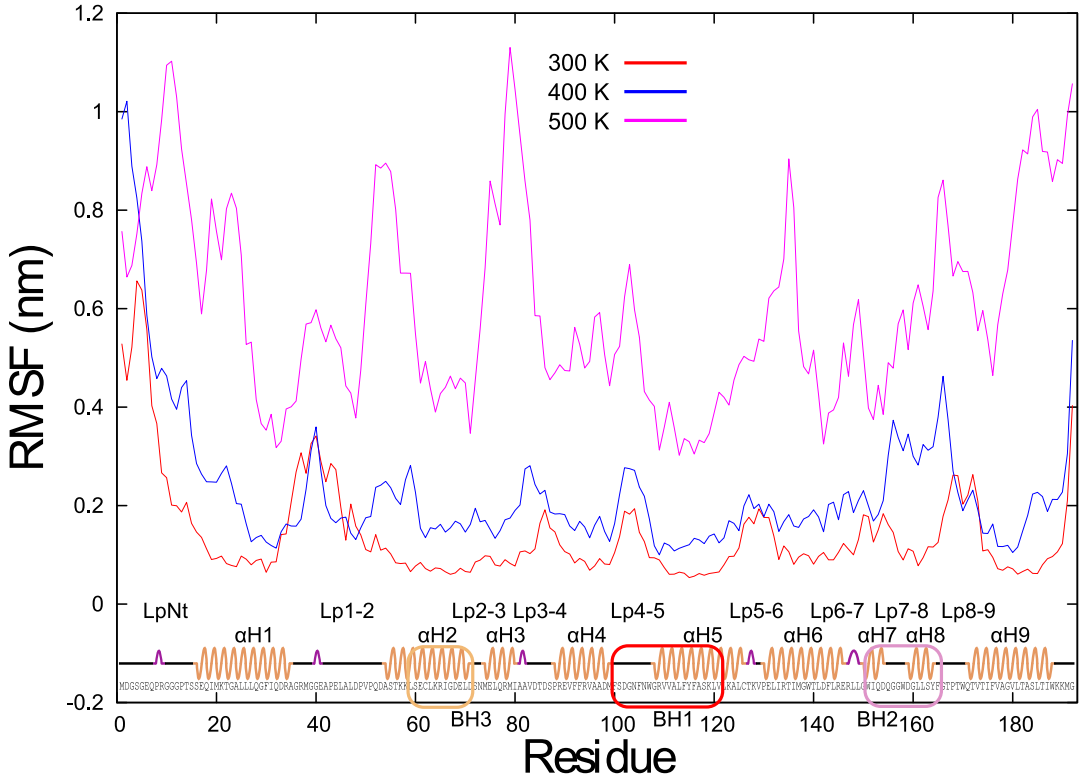


Figure 3.9: *The Root Mean Square Fluctuations (RMSF) of Cα coordinates are shown as a function of residue number at different temperatures; a representation of the secondary structure with labels for the main domains is included.*

Concerning RMSF values, they are smaller at all temperatures in those regions that belong to α-helices, and larger in loops and turns (Fig. 3.9). This analysis shows that LpNt is the most mobile protein region. At 300 and 400 K, RMSF oscillates around 0.2 nm, whereas at 500 K, the RMSF values increase from 0.4 to 1.1 nm, showing seven peaks larger than 0.8 nm. Interestingly, one of these peaks is located at the C-terminus of Bax, which reflects an important conformational change in this region. It is important to mention that RMSF values from BH3 remain low compared to those of the rest of the protein in all MD simulations. The curves obtained follow closely those published by other groups, despite their using of a Bax from a different species [88], or using lower temperatures than those employed in this work [89].

3.2.2 Clustering analysis

The stability of Bax during the MD simulations has been correlated with the existence of various local minima that split the space of structures in clusters, whose similar geometry permits them to maintain the grouping. This traditional cluster analysis was performed with the GROMACS tool `g_cluster` setting the parameter RMSD cutoff values of 0.15, 0.22 and 0.42 nm were used for clustering the structures at 300, 400 and 500 K respectively [136, 137].

Fig. 3.10 shows the pairwise distribution of the RMSD for the MD at 300, 400 and 500 K. Interestingly, bimodal distributions are presented for the three temperatures analyzed. At 300 K the distribution has two maxima in 0.25 and 0.35 nm, whereas maxima at 0.22 and 0.42 nm are shown at 400 K, and maxima at 0.8 and 1.1 nm are shown at 500 K. Namely, two populations are detected, however, the bi-equipartition of the probability is similar at 300 and 400 K, but increases at 500 K (Fig. 3.10). When the time-evolution of distinct clusters of Bax backbone conformations was calculated, it shows that at 300 K only two clusters are visited, the first one briefly between 2 and 4 ns and the second from 4 ns to the end of this MD simulation. The MD simulations at 400 and 500 K visit more clusters, but two of these stand out as the most populated: from 2 to 8 ns and from 9 to 19 ns at 400 K and from 6 to 13 ns and from 13 to 20 ns at 500 K (Fig. 3.11). A temperature-dependent bistability is derived from our MD simulations. This bistable trait has been reported by other authors with distinct physical and mathematical approaches [68, 138, 139]. Representative structures from two different clusters for each temperature are depicted in Fig. 3.12.

3.2.3 Secondary structure evolution

Secondary structure changes in Bax were evaluated by means of the MD simulation results to compare the structural evolution for the different temperatures used. Fig. 3.13a depicts the secondary structure content in the first 2 ns at 300 K. The last conformation obtained after these 2 ns was considered as the starting point for the next 18 ns of thermal unfolding evaluations. Fig. 3.13b shows that at 300 K most of the α-helices are maintained, with a slight reduction of helical structure in αH1, αH8 and αH9. Moreover, the protein loses some α-helices at 400 K, particularly αH2 and αH8 (Fig. 3.13c). However, the α-helix structure reduction is more noticeable at 500 K, where almost all α-helices are absent, except for αH5, which forms the hydrophobic core and appears to be remarkably stable against thermal unfolding. It is worth mentioning that at 500 K, αH6 remains curled, adopting the less stable 3_{10}-helix conformation intermittently (Fig. 3.13d). Additionally, throughout the MD

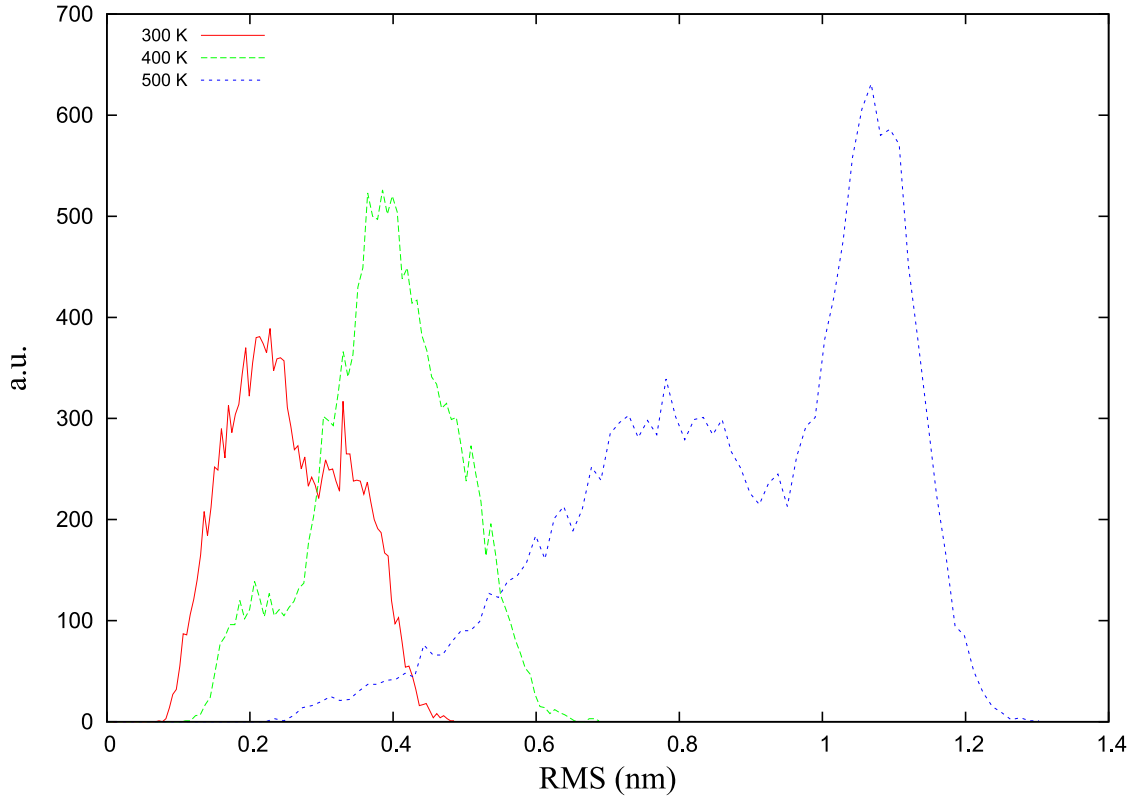


Figure 3.10: *RMS Distribution for the structures obtained in the simulations in arbitrary units (a.u.).*

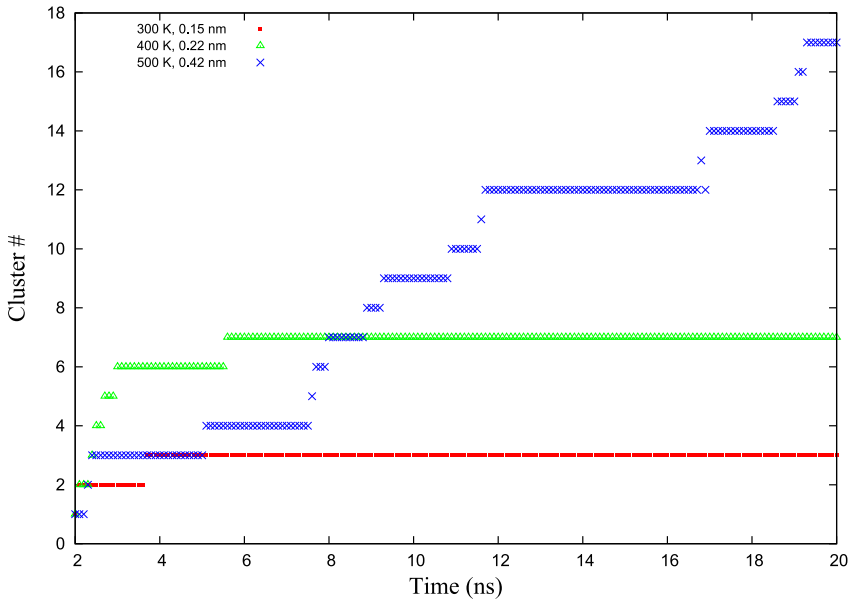


Figure 3.11: *Time evolution of the clustering of the structures from the simulations. Cutoff values of 0.15, 0.22 and 0.42 nm were used for clustering the structures at 300, 400 and 500 K respectively.*

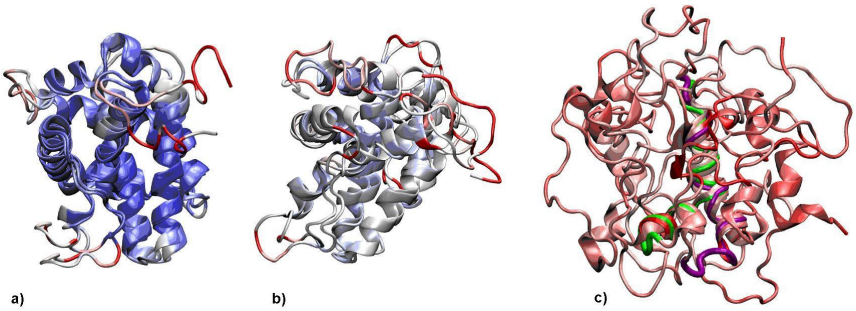


Figure 3.12: *Aligned structures from two different clusters in the MD simulations. a) Structures at 3.1 ns and 17.2 ns from the MD simulation at 300 K. Differences in LpNt, Lp1-2 and the C-terminal end of BH2 are observed. b) Structures at 4.9 ns and 17.2 ns from the MD simulation at 400 K. Important differences are appreciated in LpNt, Lp1-2 and BH2. c) Structures at 8.1 ns and 19.4 ns from the simulation at 500 K. aH5 is depicted with a green tube at 8.1 ns and a purple tube at 19.4 ns.*

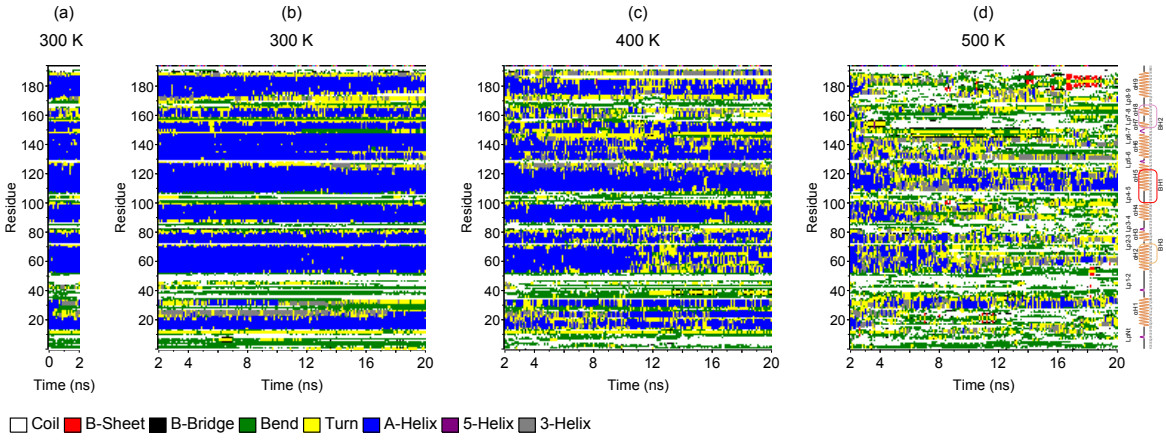


Figure 3.13: *Time evolution of the secondary structural elements of the protein during the MD simulations at different temperatures (DSSP classification). The graphical representation of the native secondary structure used in Fig. 3.9 is reproduced on the extreme right.*

simulations, the hydrogen bond network is reorganized (data not shown), preventing the elongation of the polypetide under heating conditions. These results are correlated with the evolution of Rg values (Fig. 3.7). These findings indicate that the structure tends to remain compact.

3.2.4 Electrostatic interactions keep Lp1-2 close to Lp5-6

It is known that electrostatic interactions are important for mantaining 3D structure in MD simulations of proteins [140]. In the native Bax structure, there is an electrostatic interaction between R134 (αH6) and E41 (Lp1-2). After 5 ns at 300 K (Fig. 3.14a-i), these two residues move away from each other, making a new electrostatic interaction between E44 (Lp1-2) and K128 (Lp5-6), respectively. Fig. 3.14b-i-iv and Fig. 3.14c-i-iv show these four residues (E41, E44, K128 and R134) at small distances from each other (Fig. 3.14b-i, inset), which helps to keep Lp1-2 close to Lp5-6. This proximity is also observed in all NMR models obtained by Suzuki *et al.*, despite the large mobility shown by E41 (Fig. 3.15), and the electrostatic interactions between these residues seem to be favored at high temperatures. An interaction of one or both of the acidic residues in Lp1-2 with one or both of the basic residues close to Lp5-6 appears constantly in all the MD simulations, where the minimum distance between

the elements of one of these pairs is below 0.2 nm and is close to the van der Waals distance (0.16 nm) 80% of the time at 300 K, 90% of the time at 400 K, and 70% of the time at 500 K (Fig. 3.16). The persistence of these electrostatic interactions is remarkable at the end of the MD simulation at 500 K, when most of the α-helix structure is lost (Fig 3.14c-iv, left inset). These interactions appear even at 600 and 700 K (results to be published); these findings suggest that they play an important role in keeping Lp1-2 close to Lp5-6 and therefore contribute to the structural stability of Bax at high temperatures.

3.2.5 Residue D33 remains close to BH3

Cartron *et al.* found that point mutated Bax (D33A) was unable to interact with tBid, which suggests that D33 (αH1) is an important residue for Bax function [48]. Our MD simulation results show that D33 is close to two basic residues in BH3 (K64 and R65). The minimum distance between D33 and either K64 or R65 is close to the van der Waals distance (0.16 nm) 15% of the time at 300 K and 30% of the time at 400 and 500 K (Fig. 3.17). These electrostatic interactions contribute to the proximity of αH1 and αH2 and hence to the structural stability of Bax. The disruption of a similar "ionic lock" has been identified as important in the activation process of a protein with transmembrane helices [141]. The previously mentioned residues involved in these electrostatic interactions are shown in Fig. 3.14. Fig. 3.14b-iii shows that at 15 ns in the MD simulation at 400 K, αH1 and αH2 have a tendency to move away from each other. The electrostatic interactions between D33 and K64, R65 or both help to prevent this separation (Fig. 3.14b-iii, inset). These interactions are also observed in the MD simulation at 500 K (Fig. 3.14c-iv, right inset).

3.2.6 Unfolding of αH4 protects the hydrophobic core

During the first 4 ns of the MD simulation at 500 K, all α-helices tend to unfold, as can be seen in Fig. 3.13d. There is a large jump in the RMSD plot for this MD simulation (Fig. 3.6, 11-12 ns), which may be related to the loss of the great majority of helicity, as can be seen in the snapshot at 15 ns (Fig. Snapshots c-iii). When αH4 unfolds and becomes an unstructured coil (15 ns), it remains close to the core of the protein and protects αH5, and this domain folds back into a helix. Likewise, unfolded αH4 prevents αH7 from moving away from the core of the protein (Fig. 3.14c-iii). With the refolding of αH5 and the holding of αH7 close to the core, αH6 regains helicity (20 ns, Fig. 3.14c-iv). After becoming a coil, αH4 protects the hydrophobic core. This concerted mechanism helps BH1, BH2 and BH3 to remain close to each other.

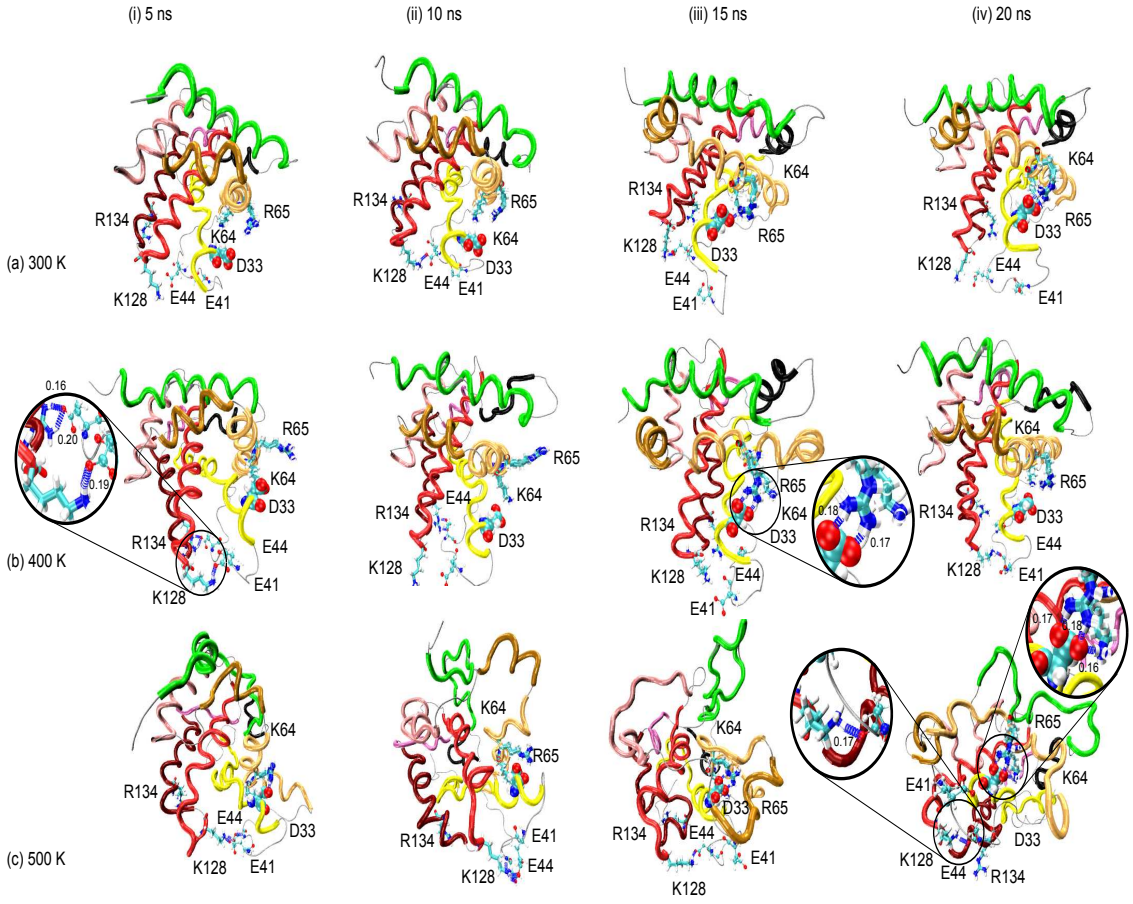


Figure 3.14: *Snapshots of the structure of Bax evolved at different time points at different temperatures. Important residues and the hydrogen bonds where they are involved are shown. The insets show distances in nm.*

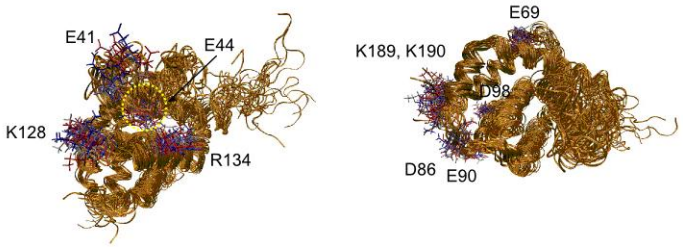


Figure 3.15: *All 20 NMR models from the PDB file are shown (PDB ID: 1F16).The color coding representation for the depicted residues ranges from red for model 1 to blue for model 20.*

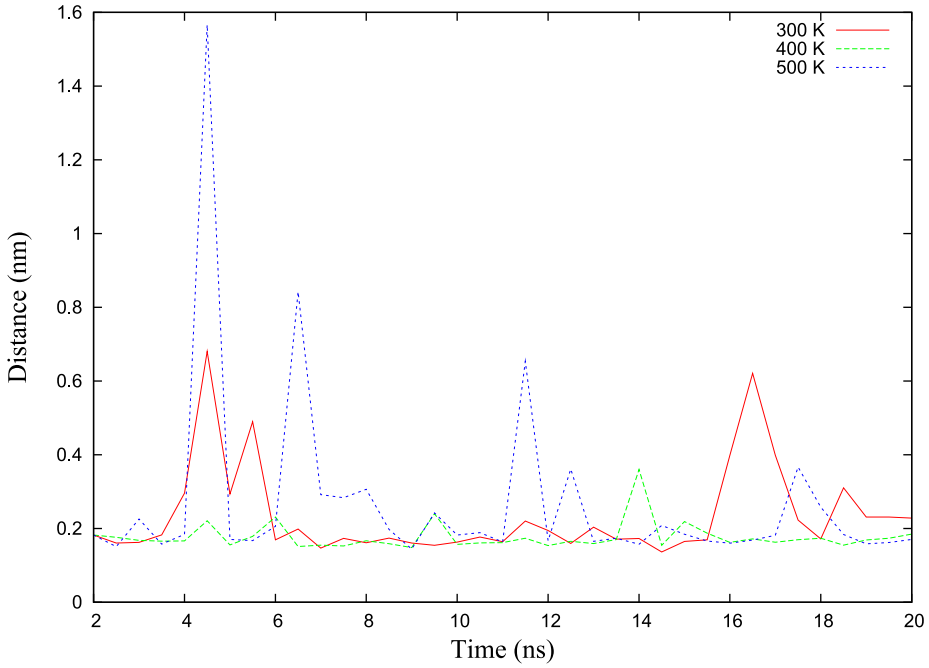


Figure 3.16: *Minimum distance between E41E44 and K128R134. Two sets are defined, one that contains E41 and E44 and the other K128 and R134. The minimum distance between a pair formed with one element of each set is shown in this figure to be close to the van der Waals distance (0.16 nm) at several times in all MD simulations. This proximity is important in the stability of Bax because it prevents the separation of the hairpin formed by αH1 and αH2 from the hydrophobic core (αH5 and αH6).*

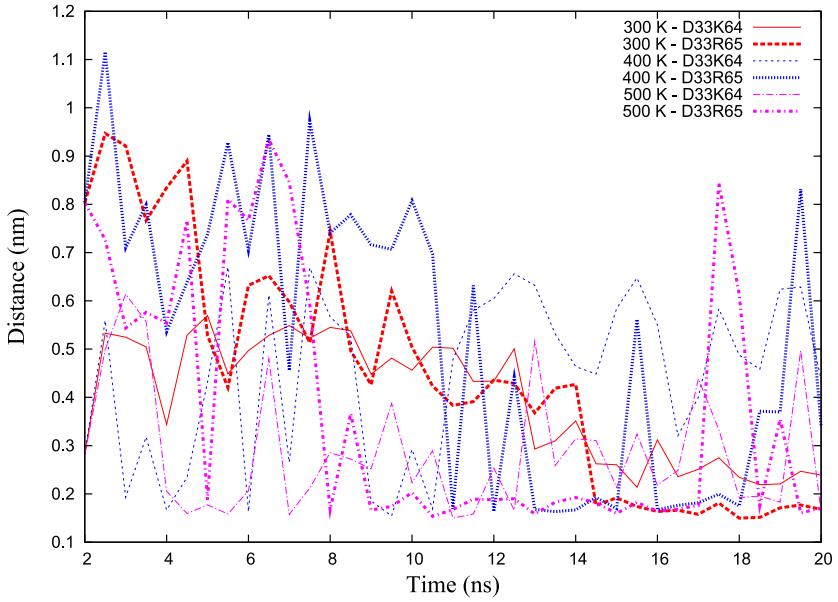


Figure 3.17: *Minimum distance between D33 and K64R65. The distance between D33 and these basic residues in αH2 is short even at 500 K.*

These conformational changes are depicted in Fig. 3.14c and can also be followed in Fig. 3.13d, where the loss of αH4 at 500 K is closely followed by a significant recovery of αH5 and a small recovery of αH6. The changes in the trend in SASA at 500 K can also be related to this mechanism, where a growing trend starting after 10 ns, which could be caused by the unfolding, is replaced by a decreasing trend after reaching 16 ns, which could be explained by the recovery of structure (Fig. 3.8c).

Fig. 3.18 shows the final structures at 300 and 500 K contrasting spatial correlation of αH4 and hydrophobic core of Bax. Although there is an important unfolding of αH4, it remains in the proximity of αH5 at 500 K.

3.2.7 Exposure of region immunodetected by antibody 6A7 (residues 12-24) at high temperatures

The antigen-antibody recognition is mediated by the contribution of the attractive and repulsive forces (van der Waals interactions, hydrogen bonds, salt bridges and hydrophobic effects). Epitopes have approximately 15 amino acids when defined by spatial contact of antibody and epitope during binding [142].

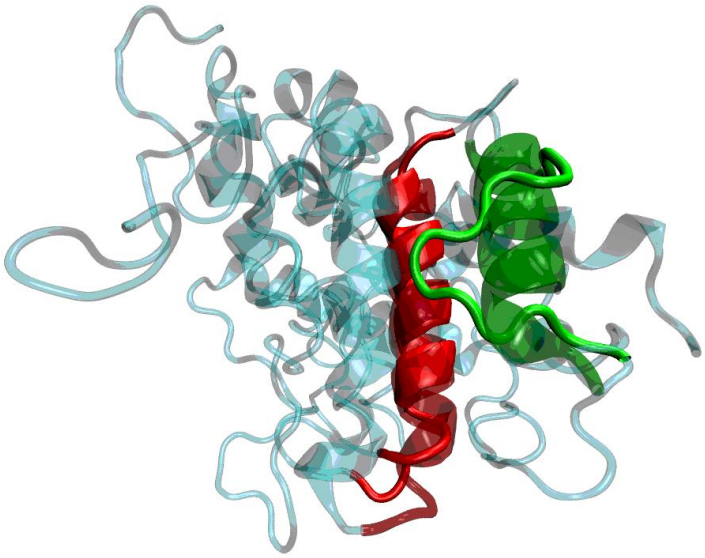


Figure 3.18: *Structural alignment between final structures (20 ns) at 300 (glassy) and 500 K. Domains αH4 (green) and αH5 (red) are highlighted.*

Hydrogen bond formation and/or electrostatic interactions seem to be required for recognition. In the interface antigen-antibody interactions, the buried surface between them is almost the same promoting shape complementarity. Hydrogen bonds are more frequently observed and considered to be a critical specificity-determining factor. It is known that the free energy of complex formations represents a balance between enthalpic and entropic forces. Approximately 5 of the 15 amino acids in each epitope strongly influence binding [143]. In general, antigens and antibodies in solution have to overcome large entropic barriers before they can constitute a tight binding. There is a loss of the entropy of mobile segments; on the contrary, entropy is gained when water molecules are displaced from the interface.

Likewise, enthalpic contributions (van der Waals interactions and hydrogen bond formation) are considered to scarcely participate to the overall binding energy; however, they define mainly the specificity of the molecular recognition [144, 145]. Controversially, it has been found that enthalpic contributions appear to be the "dominant" force, driving the total free energy changes [146, 147].

There is evidence that Bax activation exposes an epitope (G12-A24) [52, 51, 50]. The SASA values for these residues in the MD simulations show little change, as they are around 8 nm^2 at 300 K and remain below 11.2 nm^2 even at 500 K (Fig. 3.19).

Taking into account these proportions and following the known cell physiology of Bax, a significant conformational change that includes 6A7 exposure might have a considerable energetic cost in the neighborhood of this region. Even more, given that not only 6A7 epitope is exposed but also other functional domains are reorganized, the energetic cost to activate Bax to promote apoptosis increases. This could be consistent with a dependent-energy apoptosis as a programmed active process.

Searching for structural changes involved in the exposure of these residues, it was observed that the GGG motif close to the N-terminus of αH1 (G10-G12) is very flexible and allows M1-G10 (LpNt) to interact with αH6. In the MD simulations performed, favorable electrostatic interactions were detected between D2 (LpNt) and R145 (αH6) at 300 and 500 K. In addition, electrostatic interactions between D2 and R9 (both LpNt) at 300 and 500 K and between D6 and R9 at 400 K were found to help to bend LpNt (Figs. 3.20, 3.21). These D2 interactions explain its local RMSF minima at 300 K and at 500 K (Fig. 3.9).

The results from the MD simulations performed show that LpNt protects the 6A7 epitope on one side during the course of the MD simulations. On the opposite side, we found that residues P49-S60 and the 6A7 epitope make a cavity that allows the latter to be in contact with the solvent in the native conformation and during the MD simulation at 300 K; however, the small size of this cavity shown during the MD simulations would prevent the corresponding antibody from reaching the epitope. In the structure reported in the PDB file 2G5B (Fig. 3.23) the Bax residues T14-I19 are in the interface of Bax-6A7 interaction. This access to solvent but not to larger molecules explains the small change in SASA for this domain in the MD simulations.

These residues (P49-S60) start in Lp1-2 and include the N-terminal part of αH2 (Fig. 3.22). Their proximity to the epitope (Fig. 3.22a) is partially stabilized by the electrostatic interaction between K21 (epitope) and Q52 (Lp1-2). This interaction is lost at 400 K (Fig. 3.22b). At 300 K, Lp1-2 shows some mobility but remains close to the core for most of the MD simulation (Fig. 3.22a). At 400 K, Lp1-2 adopts a diferent conformation and points away from the core (Fig. 3.22b). These movements allow the residues P49-S60 to move away from the epitope, leaving it exposed. At 500 K, the epitope is uncovered by P49-S60 moving away from the core early in the MD simulation.

Another factor that prevents αH1 from moving away from the core of the protein is the formation of several hydrogen bonds between LpNt (residues M1-T14) and the core of the protein. The structural analysis also shows an increment in the number of residue pairs closer than 0.35 nm. These pairs are likely to have a favorable interaction, which contributes to keep these residues close to the core of Bax (Fig. 3.24).

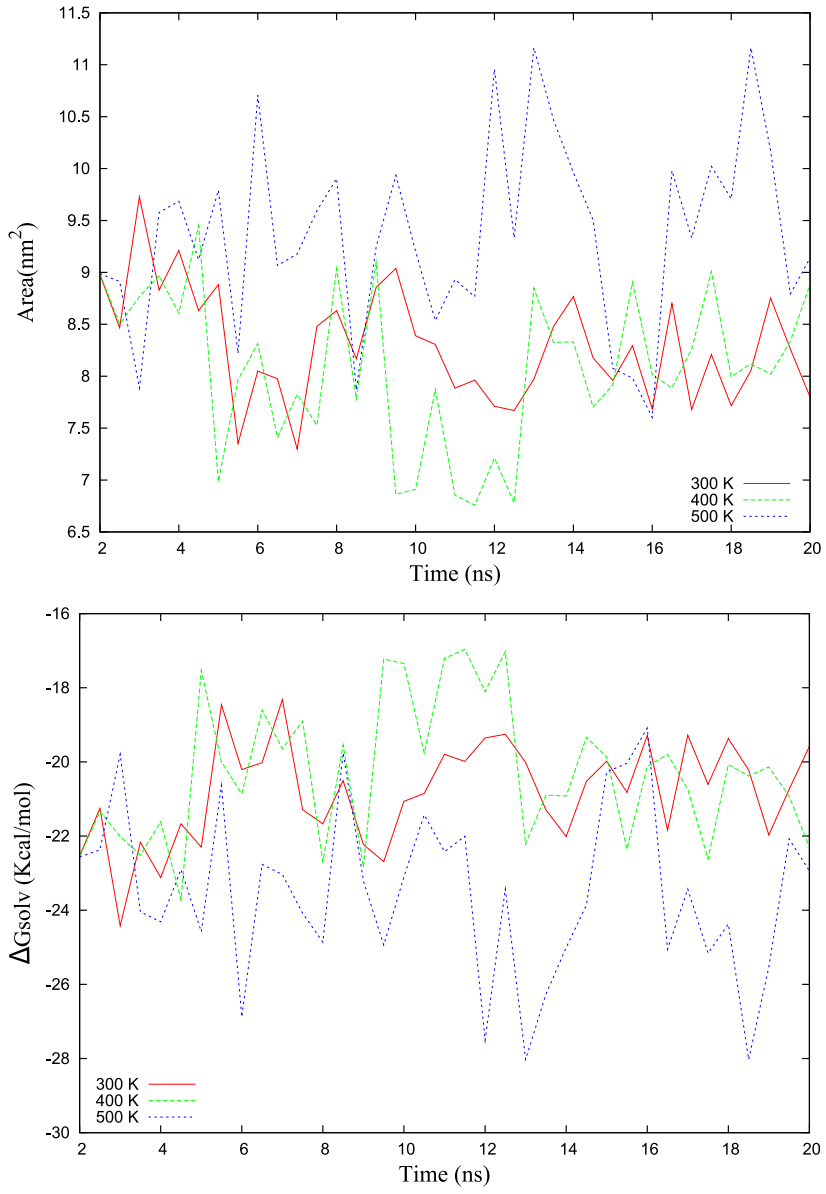


Figure 3.19: *Solvent Access to 12-24 (epitope 6A7). In the total solvent accessible surface area (above), the values remain below 11.5 nm^2 in all simulations. The residues that cover the epitope allow for enough space for the solvent but block the corresponding antibody. ΔGsolv values (below) remain larger than -28 Kcal/mol most of the time.*

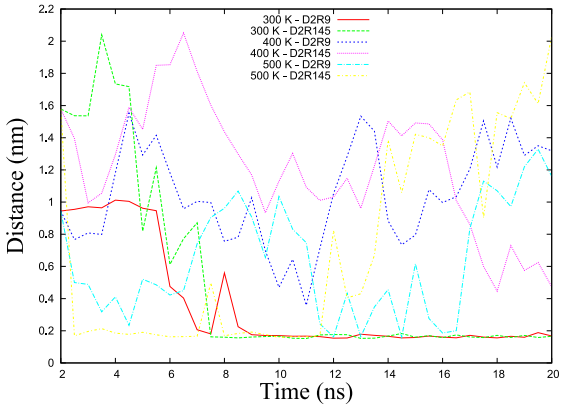


Figure 3.20: *Minimum distance between D2R9-D2R145. The proximity of D2 and R145 stabilizes the bending of LpNt and the proximity of D2 and R145 keeps LpNt covering the epitope 6A7 (residues 12-24). The minimum values observed are below 0.2 nm, which leaves them close to the van der Waals distance of 0.16 nm. Interestingly, this proximities are more pronounced at 500 K than they are at 400 K and are not present in the conformation at 2 ns but appear later in the MD simulations.*

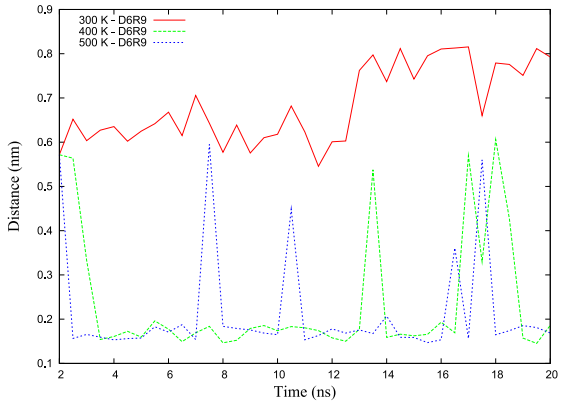


Figure 3.21: *Minimum distance between D6-R9. The proximity of D6 and R9 contributes to the stabilization of the bending of LpNt that facilitates its function of covering the epitope 6A7. Interestingly, this proximity is only observed in the MD simulations at high temperatures. The minimum distance is close to the van der Waals distance (0.16 nm) at 400 K and 500 K.*

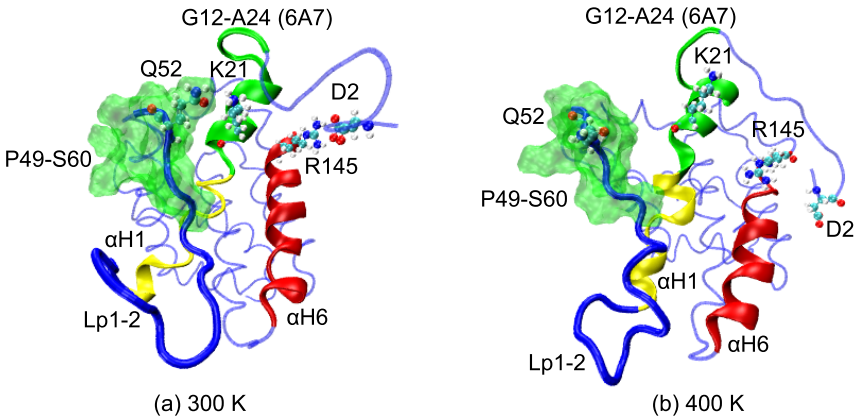


Figure 3.22: *Exposure of the epitope 6A7 under heating conditions. Final structures for the MD simulations at 300 K (a) and 400 K (b) are shown. Residues K21, Q52, D2 and R145 are shown in CPK. The epitope is shown in green; it starts in LpNt and includes the first residues in αH1, and the rest of αH1 is shown in yellow. Residues P49-S60 are shown with transparent surfaces in green. Domain αH6 is shown in red.*

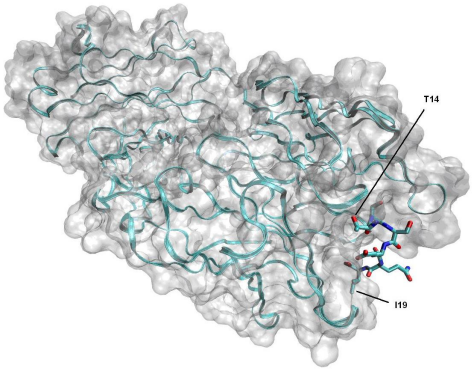


Figure 3.23: *Crystal Structure of the anti-Bax monoclonal antibody 6A7 and a Bax peptide (PDB ID: 2G5B). The residues from Bax that are seen close to the antibody are T14-I19.*

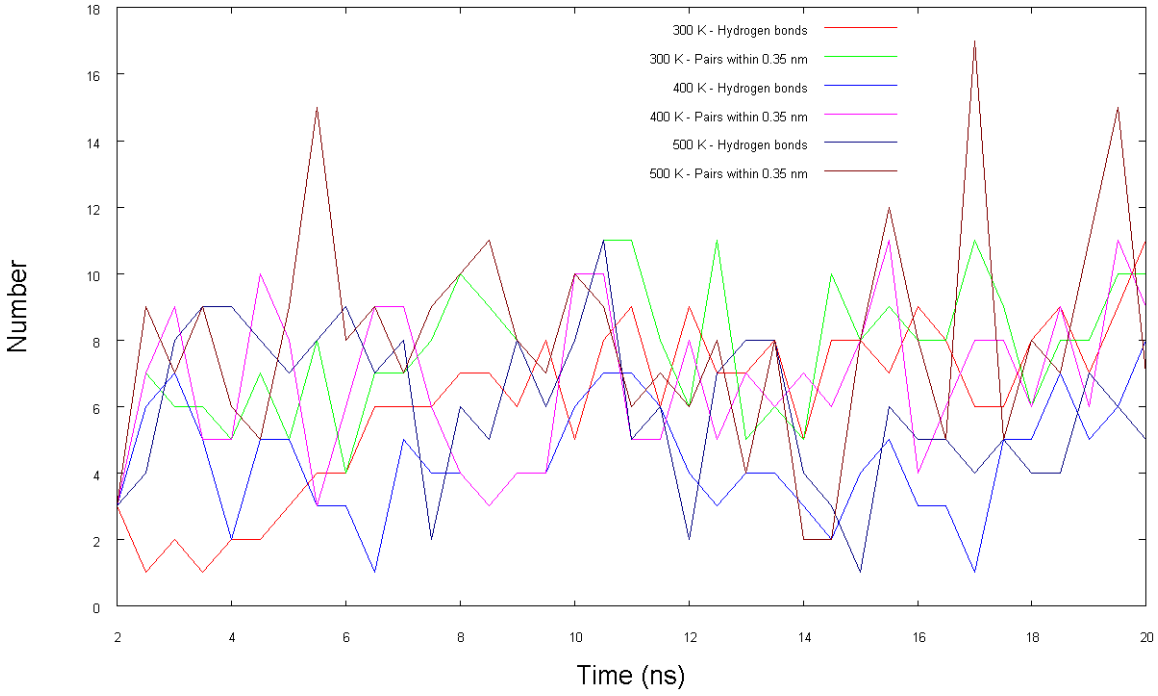


Figure 3.24: *Hydrogen bonds between LpNt and the rest of the protein. All simulations show a net increase in the number of hydrogen bonds between LpNt and the rest of the protein, helping to keep it close to the core. Pairs within 0.35 nm are likely to have a favorable interaction, and their number is also increased in all simulations, thus contributing to the proximity of LpNt and the core of Bax.*

The contribution of LpNt to the structural stability of Bax suggested by our MD simulations is in agreement with the findings of other workgroups which have proposed that residues M1-M20 (LpNt) play an inhibitory role for Bax mitochondrial localization [52, 49]. These MD simulations results reveal that Bax can show the 6A7 antibody epitope at high temperatures. This finding is in agreement with the experimental results reported elsewhere [51].

3.2.8 Hydrophobic groove covering by αH9

The hydrophobic groove formed by BH1-3, which is occupied by αH9 in the native conformation, has been shown to be important in the interaction of Bax with other members of the Bcl-2 family [53, 54]. Therefore, the exposure of this groove could be involved in the oligomerization process of Bax. This hydrophobic groove is protected constantly by αH9 despite the great conformational changes induced by heating conditions in the MD simulations.

These MD simulations provide evidence that electrostatic interactions involving K189 and K190 with E69, D86, E90 and D98 are important for protecting the hydrophobic groove by αH9 under heating conditions. K189 and K190 tend to point in opposite directions due to steric hindrance effects and their equal positive charges. With this spatial disposition, these residues form a positively charged "anchor" that is trapped in an ionic "cage" formed by four negatively charged residues (E69, D86, E90 and D98) in the MD simulations at high temperatures. For example, at 400 K, D98 interacts with K190 (Fig. 3.25a), whereas at 500 K, K190 is close to D86 (5 ns, Fig. 3.25b). However, at 10 ns, K189 replaces K190 in the proximity of D86 and also interacts with E90 (Fig. 3.25c). At 15 ns, another combination emerges with K190 interacting with D98 and K189 with E69 (Fig. 3.25d). The electrostatic interactions involving K189 and K190 could account for their local RMSF minima at 400 and 500 K (Fig. 3.9), and they seem to contribute to the structural stability of Bax at high temperatures. E69 and D98 are distant from the C-terminal end of αH9 in the NMR models from the PDB, hence their interactions with K189 and K190 would be nonnative (Fig. 3.15).

Moreover, there is evidence that hydrogen bonds involving the side-chain of S184 contribute to protect the hydrophobic groove by αH9 [46]. In the native structure, the side chain of S184 forms a hydrogen bond with the backbone carbonyl of V95. In the MD simulation at 300 K, the side-chain of S184 forms a hydrogen bond with the side-chain of D98, which keeps the rotamer of S184 pointing to the core of the protein. At 400 K, within the first 5 ns, the hydrogen bond of S184 with V95 is re-formed, while D98 contributes to the stability of Bax in a different manner, by interacting

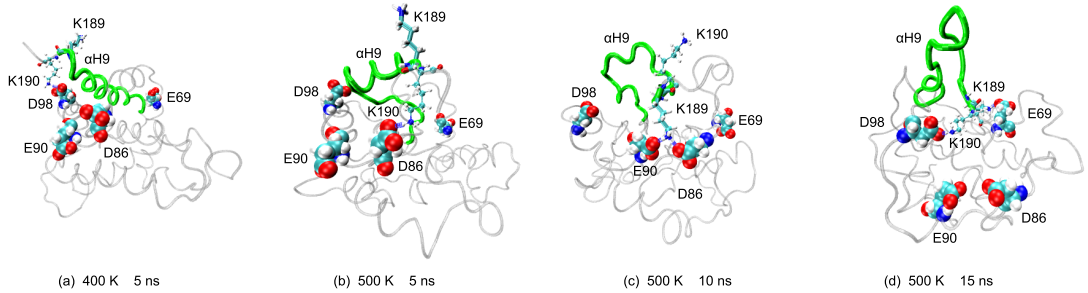


Figure 3.25: *Electrostatic interactions that prevent αH9 from moving away from the protein core. K189 is shown in Licorice and K190 in CPK. αH9 is depicted in green. Acidic residues that form the ionic "cage" are depicted in the VDW representation. In all cases, the minimum distance between the lysines and the acid residues is close to 1.6 Å, which is the van der Waals distance for a nonbonded O-H pair.*

with K190 as described above. At 10 ns, S184 forms a hydrogen bond with N106 (Lp4-5), but the helix turn of αH9 that contains S184 partially unfolds and leaves S184 exposed to the solvent.

Furthermore, the electrostatic interaction between D98 and K190 is maintained and a hydrogen bond appears between D98 and T186. At 15 ns, S184 is still exposed to the solvent, but it forms a hydrogen bond with the backbone carbonyl of V180. D98 remains close to either K190 or T186 for the rest of the MD simulation. At 20 ns, S184 is close to R109 but is mostly accessible to the solvent (data not shown). At 500 K, S184 is completely exposed, thus no longer forming hydrogen bonds with other residues (Fig. 3.26). The exposure of S184 could be the cause of a large increase in the mobility of αH9 at 500 K (Fig. 3.9), but it is not enough to separate this helix from the hydrophobic groove.

Brock *et al.* found that the release of αH9 is essential for the activation of Bax [99]; thus, the interactions reported in this section that help to maintain the helix tucked into the hydrophobic groove at high temperatures would contribute to the structural stability of Bax.

3.2.9 Hydrophobic interaction between W107 and BH2 residues

W107 belongs to BH1 (F100-V121) and is close to BH2 (W151-F165) in the native conformation. BH2 includes most of αH7, Lp7-8 and αH8. In the MD simulation at

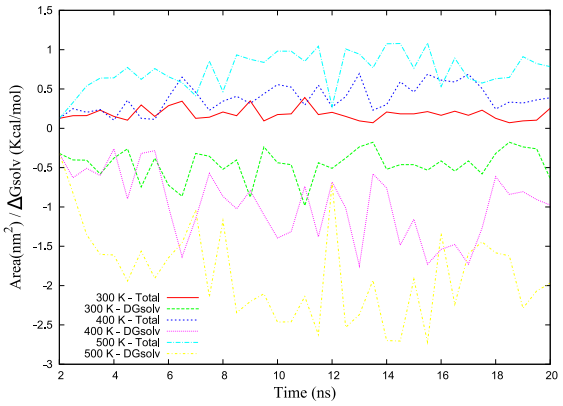


Figure 3.26: *SASA of S184. The exposure of S184 can be followed in this figure. At 300 K, the value of its SASA is always below 0.5 nm^2, whereas it is above this value intermitently at 400 K and for most of the time at 500 K. This exposure is also noticeable in the free energy of solvation (ΔGsolv), where the values are lowest for the MD simulations at high temperatures.*

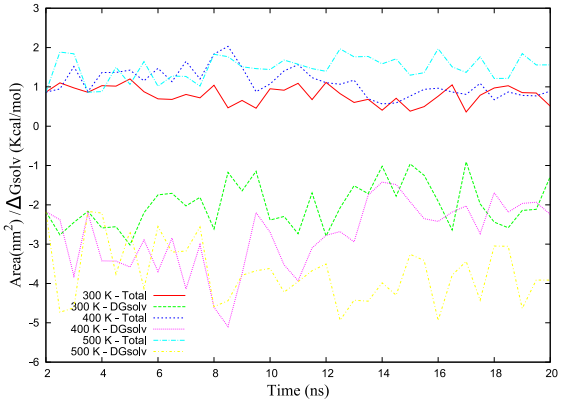


Figure 3.27: *SASA of W107. The exposure of W107 can be followed in this figure. At 300 K, the SAS area remains close to 1 nm^2, whereas it is closer to 2 nm^2 at several points at 400 K and for most of the simulation at 500 K. This exposure is also noticeable in the free energy of solvation (ΔGsolv), where the values are lowest for the simulation at 500 K.*

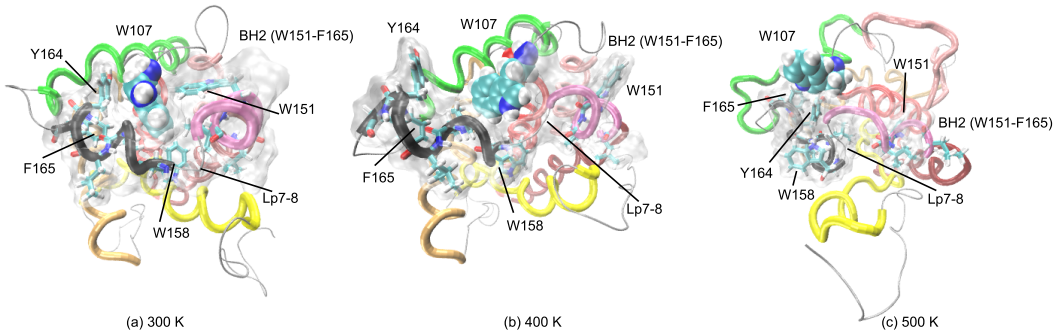


Figure 3.28: *Final structures (20 ns) are shown from the MD simulations at 300 K (left), 400 K (middle) and 500 K (right). W107 is shown in VDW and the SAS of the region formed by residues L148-F165 in transparent gray. The color code for the helices is the same as in Figure 1. At 300 K, W107 remains buried and its rings are perpendicular to Lp7-8. At 400 K, W107 is more exposed, and its rings are almost parallel to Lp7-8. At 500 K, W107 is mostly exposed. The hydrophobic and aromatic residues in BH2 (W151, I152, W158, L161, L161, Y164 and F165) are shown in licorice.*

300 K, an aromatic-aromatic interaction between W107 and Y164 is suggested; this kind of interaction has been reported as a mechanism of tertiary structure stabilization [148]. Also, in this MD simulation, W107 buries itself in the core, approaching W158 in Lp7-8, which helps to keep BH1 close to BH2 (Fig. 3.28a). The SASA value for W107 increases in the MD simulations under heating conditions (Fig. 3.27). At 400 K, W107 is partially exposed and the hydrophobic and aromatic residues in BH2 are not as close as they were at 300 K because BH2 and Lp7-8 elongate, although they remain close to BH1 (Fig. 3.28b). At 500 K, W107 is mostly exposed, although it is close to Y164 and F165, which favors aromatic interactions; however, BH2 shows a loss of helicity during this MD simulation (Fig. 3.28c). W107 is at the beginning of the RMSF valley that corresponds to αH5 in all the MD simulations (Fig. 3.9).

The observed hydrophobic and aromatic-aromatic interactions of W107 with residues in αH7 (W151, I152), Lp7-8 (W158) and αH8 (L161, L162, Y164 and F165) are in agreement with structural studies reported elsewhere [149].

3.2.10 Normal mode analysis

NMA identifies potential conformational changes in proteins [150]. In this work, NMA was performed with the Elastic Network Model, which is implemented with the "rotation-translation-block" approximation in the web interface *ElNémo* [130]. Previous reports in the literature suggest that the consideration of the first ten nontrivial modes should be suficient for a reasonable analysis of a protein with a number of degrees of freedom similar to Bax [129, 151].

NMA was performed on the native structure of Bax and the final structures of each MD simulation at different temperatures. A common pattern among the modes considered was a high level mobility in the loops, particularly in LpNt and Lp1-2 as well as in αH9. Modes with low collectivity are likely to suggest domains that would have significant mobility in a conformational change.

For the native structure, modes 9 and 11 show little collectivity; this finding is explained by the large mobility of LpNt. The same is true for the three low collectivity modes in the NMA of the final structure at 300 K (modes 7-9), whereas the rest of the modes show high mobility in Lp1-2 and Lp5-6, as does the low collectivity mode obtained at 400 K (mode 9).

Modes 14-16 at 400 K also show a high mobility of BH2. The NMA of the final structure at 500 K shows one low collectivity mode (mode 8) with a huge displacement of E41 and the residues around it in Lp1-2. Several of the other modes for this structure (modes 9, 10, 12 and 14) show increased mobility levels for αH4 (data not shown), which can be explained by its unfolding at 500 K (Fig. 3.14c). The large mobility observed in Lp1-2 and Lp5-6 in the NMA suggests that the electrostatic interactions between Lp1-2 and H5-6 (described in Subsection 3.2.4) prevent the separation of these domains, thus avoiding a major conformational change (Fig. 3.29).

3.3 Assessment of interactions found to contribute to the structural stability of Bax

The NMR structure of Bax provides 20 models with similar coordinates for the protein backbone, but the side-chains in the models tend to show large fluctuations (PDB id: 1F16) as shown in the residues highlighted in Fig. 3.29. Furthermore, it has been shown that high temperature MD trajectories can be highly chaotic and are very sensitive to the starting conditions [152], and running a single constant high-temperature MD simulation might not be sufficient to yield reliable results or establish any repro-

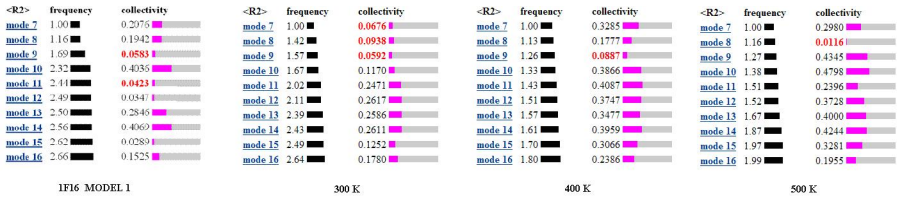


Figure 3.29: *Normal Mode Analysis. The first ten nontrivial modes are displayed. The leftmost was performed on the first NMR model of the structure from the PDB. The rest were performed on the final structure for each simulation. Low collectivity modes at 300 K are mostly due to LpNt as it is for the NMR structure. At 400 and 500 K these modes include large mobility of Lp1-2, Lp5-6, αH4 and BH2.*

ducible trend in the dynamics of the molecular system. In order to check whether the observations from the 20 ns MD run can be reproduced in additional independent simulations starting from different initial conditions, the following strategy was followed.

To reproduce the MD simulations data starting from different initial coordinates we have chosen several experimental Bax conformations from the Protein Data Bank using the structure (PDB code: 1F16) obtained by NMR (which presents 20 conformers of Bax). We selected those that have the greatest displacement in comparison to Model 1. The Bax displacement analyses was achieved considering the RMSD values from all pairs of the aforementioned conformers.

```
sum_max=0
For every Y different of 1:
  Calculate the RMSD between Model 1 and Model Y (RMSD_1-Y)
For every Z different of 1 and different of Y:
       Calculate the RMSD between Model Y and Model Z (RMSD_Y-Z)
       Calculate the RMSD between Model Z and Modelo 1 (RMSD_Z-1)
       If ( RMSD_1-Y + RMSD_Y-Z + RMSD_Z-1) > sum_max:
           sum_max = RMSD_1-Y + RMSD_Y-Z + RMSD_Z-1
           Ymax=Y
           Zmax=Z
```

Box 3.1: *Pseudocode used to find the most distant structures within the NMR models from file 1F16.pdb.*

The following algorithm was programmed in a Python script (pseudocode in Box 3.1) to find the two structures that satisfy the next both conditions simultaneously:

1. They are the most distant structures between them; and

2. They are the most distant structures from Model 1.

The most distant models from each other and from Model 1 are Models `Ymax` and `Zmax`. The triplet that gives the maximum is formed by Model 1, Model 10 and Model 4. In summary, see Box 3.2.

```
RMSD of Model  1   to Model 10 :  6.12809841359
RMSD of Model  10  to Model 4  :  5.6728624696
RMSD of Model  4   to Model 1  :  5.24419952136
Sum: 17.0451604045
```

Box 3.2: *Bax NMR Models with the greatest distance between each other.*

Afterwards, Model 4 and Model 10 were simulated in the same conditions as used for Model 1. The MD simulation data were reproduced despite using different conformers. In all models tested the RMSD had a tendency to be stable at 300 and 400 K (Fig. 3.30). Whereas that at 500 K reached a horizontal behavior although more prominent fluctuations are observed in Model 4, while Model 10 shows an asymptotic increment.

About Rg, a stable value around 1.6 nm appears at 300 and 400 K. Meanwhile, fluctuations arose from the simulation at 500 K but the absence of an increase of the Rg suggests that the compactness of the protein remains (Fig. 3.31). In turn, SASA total, in Model 4 and 10 presents fluctuations, however, all values are lower than 140 nm^2. This is a smaller value than the maximum observed in the simulation of Model 1 (280 nm^2 approximately), suggesting that there are not significant changes in the simulations of Models 4 and 10 (Fig. 3.32). When the RMSF was computed, the simulations of Model 4 and 10 show values in the same range as the ones for Model 1. Also, properties such as local minima near both ends of the peptide chain are preserved, which contribute to prevent the elongation of the structure (Fig. 3.33).

To evaluate the internal residue interactions mentioned in the manuscript, we measured the distance between side chain groups of amino acids identified to contribute to the 3D structural stability during the last 10 ns (Table 3.3) and the last 5 ns (Table 3.4). Table 3.3 shows that 94.4% (34/36) of the distances are smaller than 7 Å, and 88.8% (32/36) in Table 3.4. If we consider a distance of 4 Å as cut-off we obtain 75%

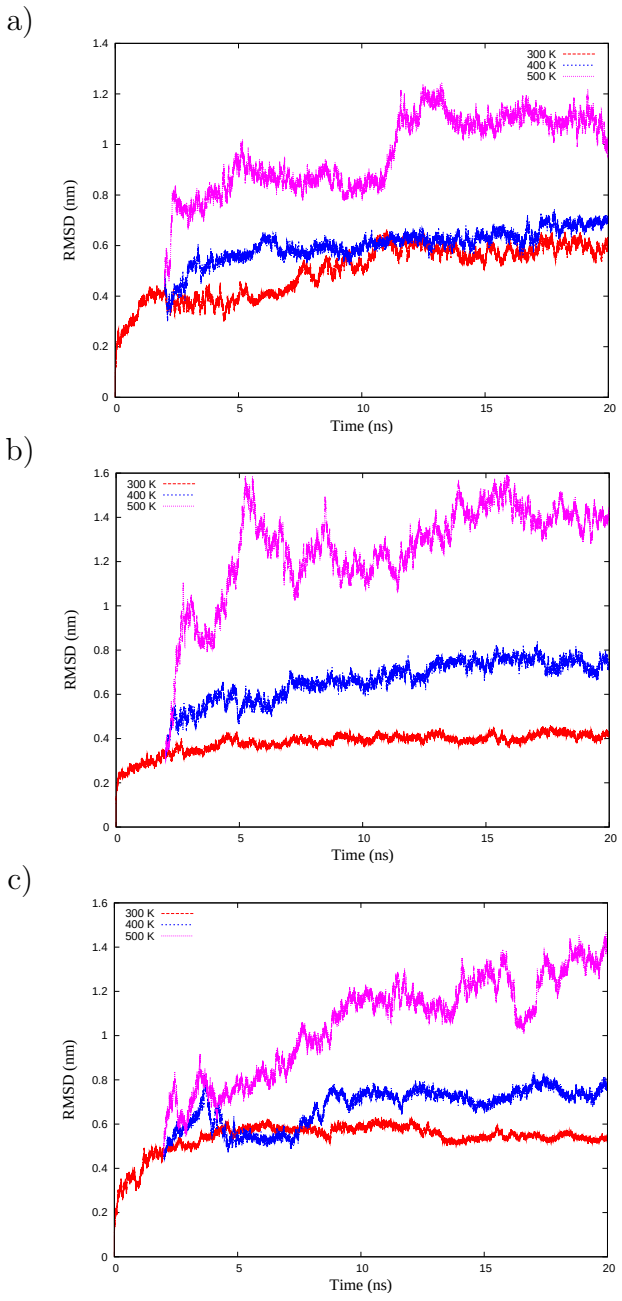
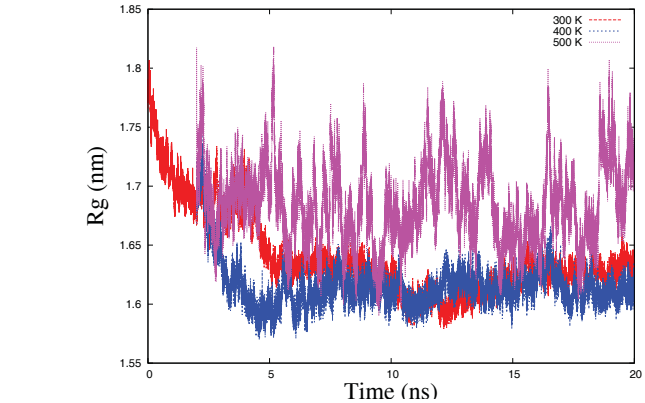
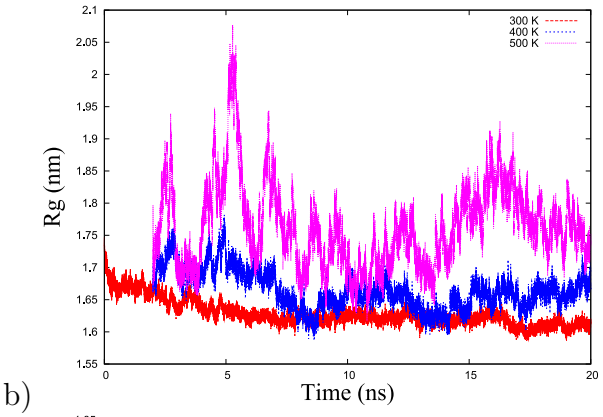


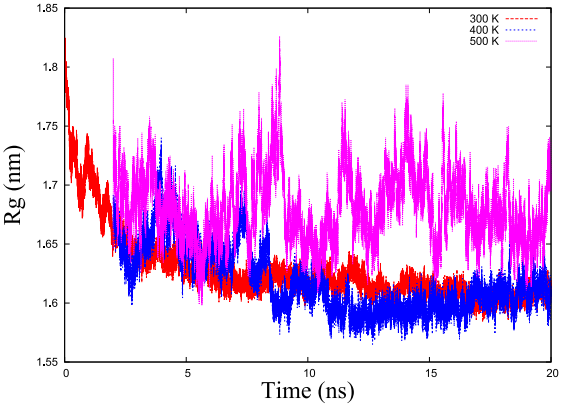
Figure 3.30: *Time evolution of Cα RMSD from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).*



a)



b)



c)

Figure 3.31: *Time evolution of the Rg from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).*

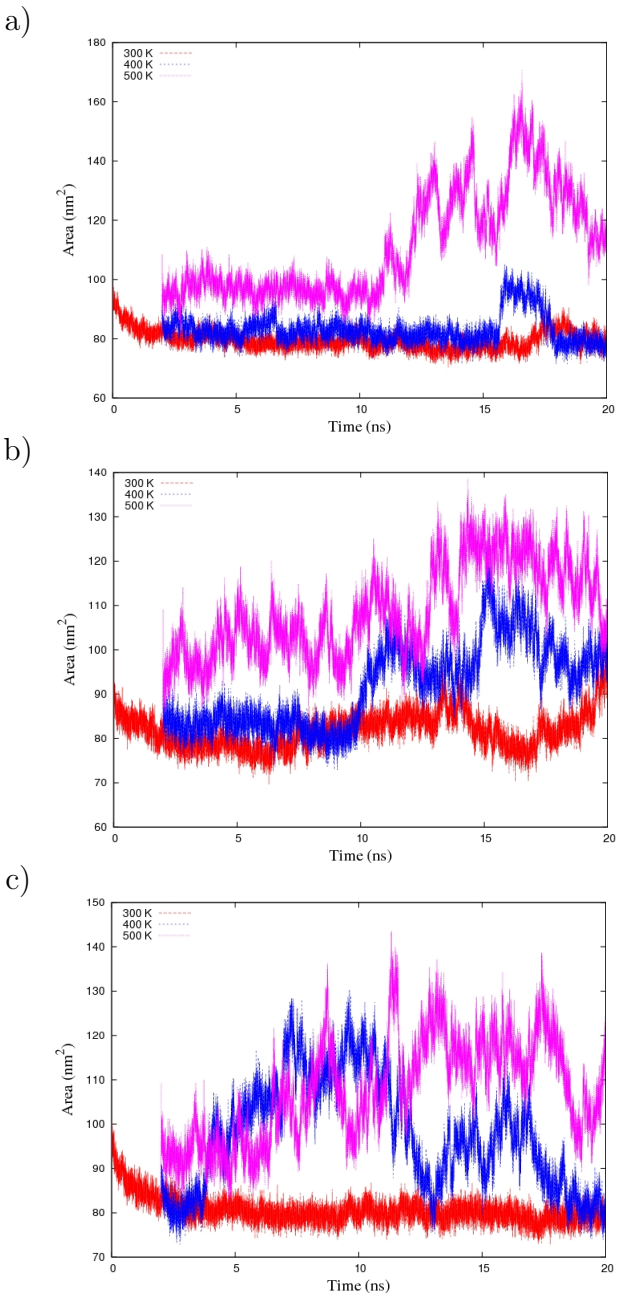


Figure 3.32: *Time evolution of the SASA from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).*

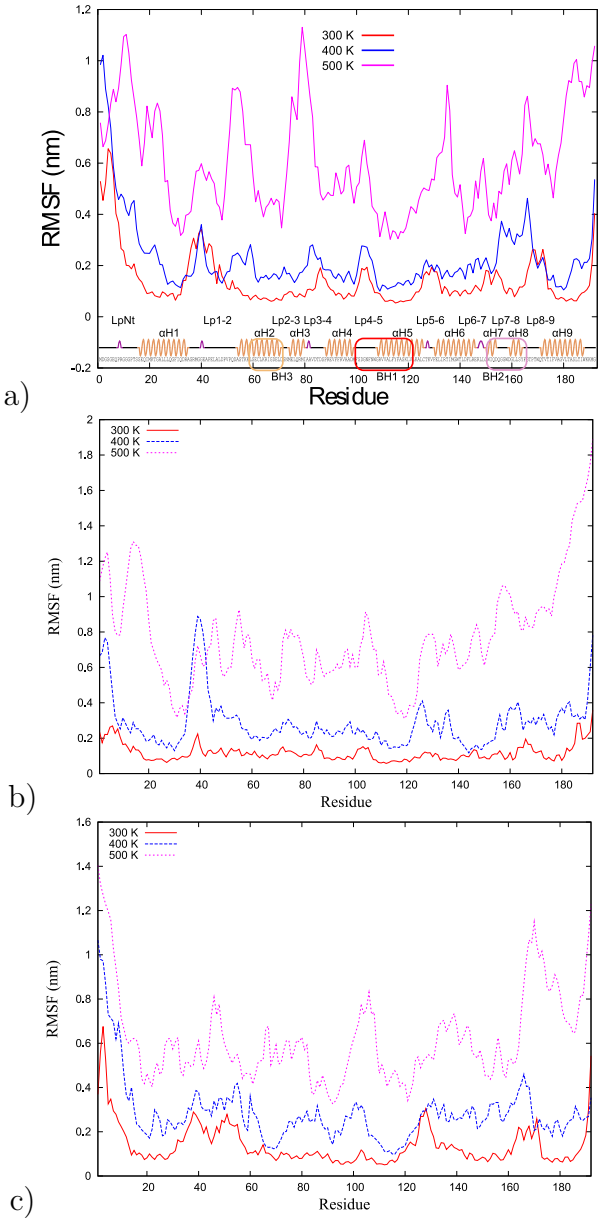


Figure 3.33: *RMSF from the MD simulations of Model 1 (a), Model 4 (b) and Model 10 (c).*

Table 3.3: *Average minimum distance of interactions proposed in the manuscript (time average from 10 to 20 ns). Distance is defined as the minimum distance between any pair of atoms from the respective groups of amino acids (Groups 1 and 2). Bold font is used when the distance is not greater than 7 Å.*

Group 1	**Group 2**	**Temp. (K)**	**Model 1 (Å)**	**Model 4 (Å)**	**Model 10 (Å)**
D33	K64, R65	300	**2.7**	**2.2**	**1.7**
		400	**2.7**	**3.3**	**3.2**
		500	**1.7**	**2.4**	**4.8**
E41, E44	K128, R134	300	**2.2**	**2.0**	**3.7**
		400	**1.7**	**1.9**	**2.1**
		500	**2.0**	*26.8*	**3.0**
D2, D6	R9, R145	300	**1.6**	**4.5**	**6.0**
		400	**2.2**	**4.3**	**6.2**
		500	**2.0**	**2.5**	**5.0**
K189, K190	E69, D86, E90, D98	300	**6.0**	**4.2**	*8.6*
		400	**2.5**	**1.6**	**3.0**
		500	**1.7**	**1.0**	**3.8**

and 72%, respectively. Therefore, we propose that these amino acids cause strong electrostatic interactions, given that it is known that the electrostatic interaction distance can be as large as 7 Å[153]. We emphasize this result because the identification of these interactions during thermal stress is the most important contribution of this research [91]. Furthermore, despite of large side chain movements, comparing distinct native Bax conformers during the MD simulations, the electrostatic interactions discussed above were maintained (Tables 3.3 and 3.4).

It is important to mention that electrostatic interactions lead to the reported "ionic locks" involved in the greater structural changes that occur in proteins (G-protein couple receptor, GPCR), where the transtition between an active and inactive state of the receptor could be electrostatic in nature [154]. In turn, while a maximum distance of 4 Å has been successfully used as part of the geometrical description of electrostatic interactions in proteins [155], residues with a separation of less than 7 Å have been shown to interact co-operatively [156]. With this in mind, the interactions proposed in the manuscript aid to obtain an atomistic model of the stability based on cooperation with a well-defined geometry.

To continue exploring the conformation space we prolongated the MD simulation during 30 ns, that is, 10 ns more than reported in [91] (Fig. 3.34). Increasing but asymptotic RMSD values are obtained. Notwithstanding the presence of three relevant peaks in the Rg values, a decreasing trend appears immediately indicating that the structure remains compact; a similar behavior is showed by SASA. Likewise, using the trajectory from 20 to 30 ns, we calculated the RMSF and similar values were obtained when compared with those in Fig. 3.9, including local minima near the extremes of the peptidic chain that were replicated. Additionally, to compare the final structures of 20 and 30 ns, an structural alignment is shown in Fig. 3.35 and a suitable result is accomplished as the RMSD values indicate. Mostly Lp1-2 and LpNt have a high mobility at 300 and 400 K whereas at 500 K this high mobility includes the full structure.

3.4 Time evolution of other relevant residues

Mancinelli and co-workers performed 10 ns of MD simulation at 300 K and 5 ns at 400 K using a united-atom force field. Their analysis aimed to identify the protein regions and the motions of Bax that are involved in the conformational change that targets Bax into the mitochondria external membrane. However, the specific interactions that contribute to the stability of Bax and their nature are not explored as is attempted in this work [89].

Table 3.4: *Average minimum distance of interactions proposed in the manuscript (from 15 to 20 ns). Distance is defined as the minimum distance between any pair of atoms from the respective groups of amino acids (Groups 1 and 2). Bold font is used when the distance is not greater than 7 Å.*

Group 1	**Group 2**	**Temp. (K)**	**Model 1 (Å)**	**Model 4 (Å)**	**Model 10 (Å)**
D33	K64, R65	300	**1.6**	**1.9**	**1.7**
		400	**2.8**	**4.4**	**2.2**
		500	**1.8**	**3.0**	**1.9**
E41, E44	K128, R134	300	**2.6**	**1.9**	**4.2**
		400	**1.6**	**1.7**	**2.2**
		500	**2.1**	*26.9*	**3.8**
D2, D6	R9, R145	300	**1.6**	**4.5**	**5.9**
		400	**2.3**	**3.7**	**6.2**
		500	**2.3**	**2.6**	**5.0**
K189, K190	E69, D86, E90, D98	300	**5.9**	**2.3**	*8.7*
		400	**2.3**	**1.6**	**3.1**
		500	**1.6**	*11.3*	**5.5**

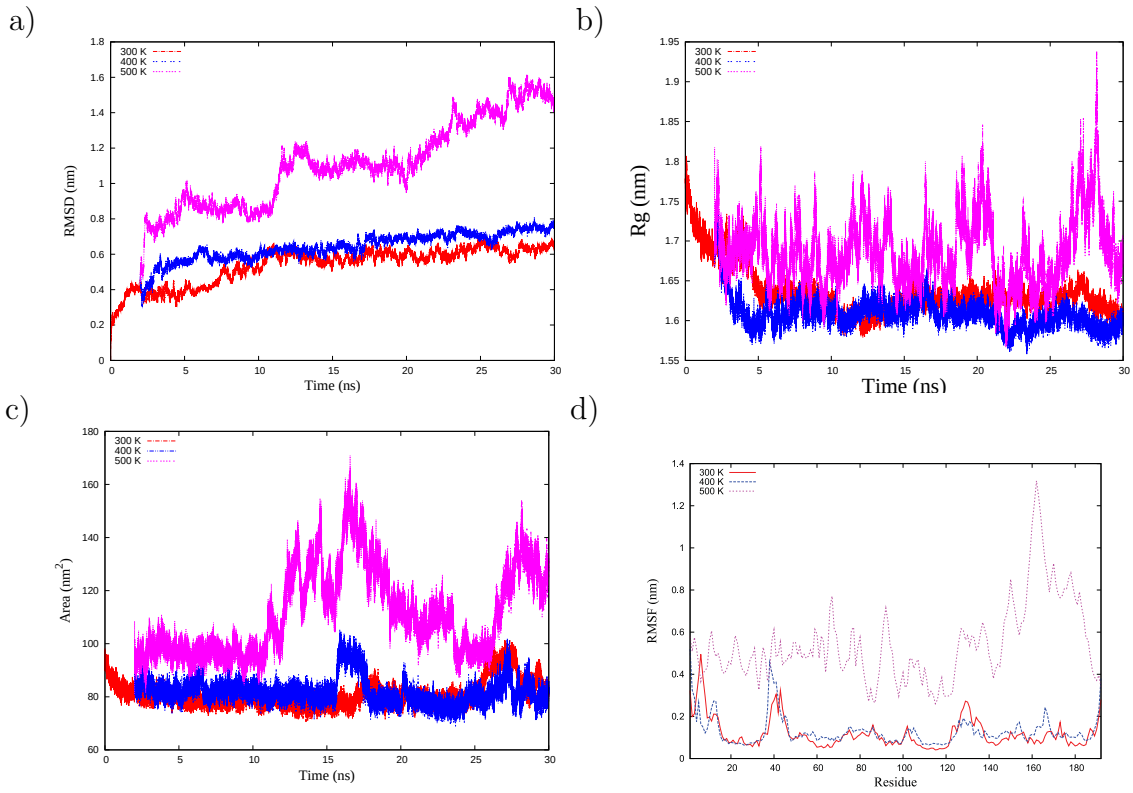


Figure 3.34: *Global stability parameters from a 30ns MD simulation of Bax. RMSD (a), Rg (b) and SASA (c) values maintain their trend after 20 ns. RMS fluctuation values (d) are only calculated for the trajectory from 20 to 30 ns and are similar to Fig. 3.9.*

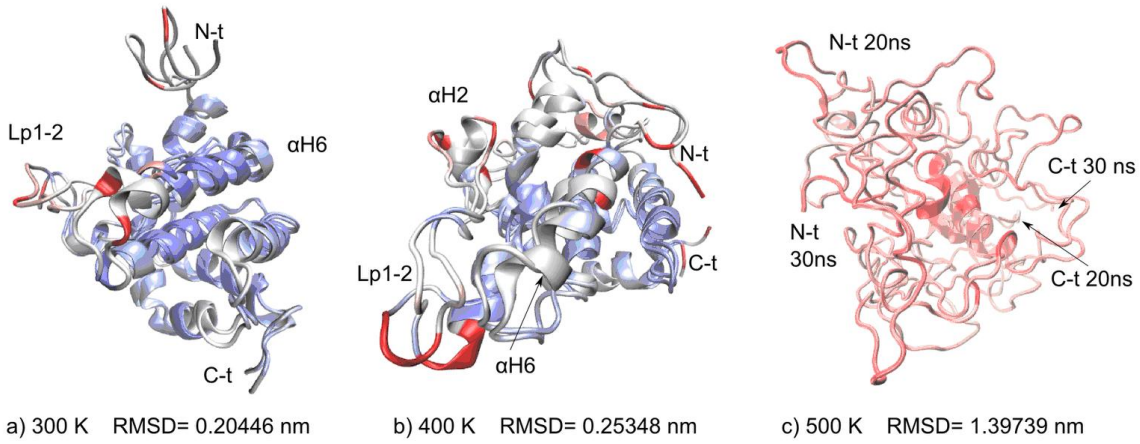


Figure 3.35: *Alignment between final structures of Bax from MD simulations of 20 and 30 ns at 300, 400 and 500 K. The regions where there is little structural correspondence or none at all are shown in red.*

There are other residues that have been proposed to be close to the Bax surface when it forms a complex with a member of the Bcl-2 family, such as A24, Q28, Q32, R37, L47, I66, D68, L70, N73, M79, A82 and E131 [52, 56, 50]. The behavior of these residues is briefly discussed below. Because Q32 and I66 remain hidden in these MD simulations, their exposure might be important in the activation mechanism of Bax. R37, L47, D68 and E131 remain exposed, and A24 remains close to the surface in our MD simulations. The exposure of R37 seems to be stabilized by electrostatic interactions. N73, M79 and A82 are part of the surface of Bax, but their SASA values increase at 400 and 500 K. Q28 is fully exposed until αH1 starts to unfold, when it faces toward the core; however, when αH1 is fully unfolded, Q28 is exposed again. L70 approaches the surface at 10 ns in the MD simulation at 500 K (data not shown).

A putative functional region of Bax was also analyzed, and it includes several residues (V95, G103, N106, G108, R109, V111, A112, W151, G157 and W158) [47]. Among these residues, Gurudutta et al. suggested that G108, R109 and W151 are important in the oligomerization of Bcl-2 family proteins by interacting with the BH3 domain of another Bcl-2 family member [157].

Although we found that G108 is fully exposed in these MD simulations, R109 is found to be partially covered by αH4 when this α-helix unfolds at 500 K, whereas W151 switches from exposed to hidden at 500 K. The hiding of R109 and W151 at high

temperatures as observed in these MD simulations would hinder the oligomerization of Bax and thus contribute to its structural stability.

As for V111 and A112 (both in αH5), they are hidden in all the MD simulations and have hydrophobic interactions with V180 (αH9) and I66 (αH2) in the final MD simulation structure at 400 K and additionally with A183 (αH9) in the final MD simulation structure at 500 K (data not shown).

These residues are located in the trans-membrane domain of Bax [46], and the hydrophobic interactions that hinder their exposure might prevent Bax from activating. G103, N106 (both Lp4-5) and G157 (Lp7-8) are exposed in these MD simulations, interacting mostly with the solvent.

Their contribution to the structural stability of Bax seems to be limited to the hydrogen bonds involving their backbone atoms. To the best of our knowledge, the roles of these particular residues in the function or stability of Bax have not been addressed in the literature. The contributions of V95 and W158 to the structural stability of Bax were discussed in Subsections 3.2.8 and 3.2.9, respectively.

Chapter 4

Conclusions and further work

4.1 Conclusions

Stable conformations for Bax are reached in all MD simulations treated at different temperatures, and most of the original alpha-helices are lost at 500 K. The stability of the 3D structure is preserved by electrostatic interactions under heating conditions. Hydrophobic interactions and hydrogen bonds are also found to play a role in maintaining the 3D structure of Bax in the MD simulations. The 3D structure of Bax appears to have several recovery mechanisms to prevent its domains from moving away from each other when subjected to high temperatures to maintain the protein compactness.

Electrostatic interactions arise between acidic residues in Lp1-2 (E41 and E44) and basic residues in αH5 (K128) and αH6 (R134) which are important to keep Lp1-2, αH5, and αH6 together when Bax is subjected to high temperatures.

Electrostatic interactions of K189 and K190 with acidic residues in other domains (E69, D86, E90 and D98) emerge and are important to keep aH9 covering the hydrophobic groove under heating conditions.

At 500 K, αH4 unfolds into a coil but remains close to the core of Bax and has interactions with BH1, BH2, and BH3 that help to keep these domains close to each other.

The contributions of several electrostatic interactions to give stability to the human Bax protein under thermal stress might have their evolutionary origin in the importance of preventing this protein from aimlessly activating and thus triggering apoptosis.

There are some domains hidden in the native structure that become exposed at

high temperatures, such as the 6A7 epitope (G12-A24).

In experiments performed elsewhere, a number of residues have been found to contribute to the structural stability of Bax. An atomistic description is proposed in this thesis for the role of many of those residues in the activation of Bax, including D33 (electrostatic), W107 (hydrophobic) and S184 (hydrogen bond).

4.2 Further work

4.2.1 MD simulation of the t-Bid/Bax complex

The BH3 Interacting domain Death agonist protein (Bid) is a BH3 only, Bax like member of the Bcl-2 family [158]. Following cleavage by caspase-8, truncated Bid (t-Bid) is known to activate Bax [159]. MD simulations of the t-Bid/Bax complex could help to elucidate the activation mechanism of Bax. We have performed short MD simulations of t-Bid in close proximity of Bax that already show affinity between these proteins (Fig. 4.1). Longer MD simulations and their analysis are proposed to increase our understanding of this interaction.

4.2.2 MD simulation of the interaction between Bax and lipidic membranes

The insertion of active Bax into the outer mitochondrial membrane is an important step in apoptosis [160]. Also, Bax can localize to the endoplasmic reticulum [161]. Furthermore, the composition of the membrane has been proposed to have a role in the activation and insertion of Bax [162]. It would therefore be interesting to investigate the interaction between Bax and lipidic membranes with variations in their composition (e.g. different content of cardiolipin) by means of MD simulations.

4.2.3 MD simulation of the Bax/Bax complex

While oligomerization is an intrinsic property of Bax, experimental results indicate that the dimers are the stable subunits of Bax oligomers [58]. We performed a short MD simulation at 500 K of two bax monomers. The two chains seem to be very close in the final structure (Fig. 4.2). Further analysis is required to characterize the interface of this putative dimer and compare it with the interface surfaces found elswhere by site-specific photocross-linking [50].

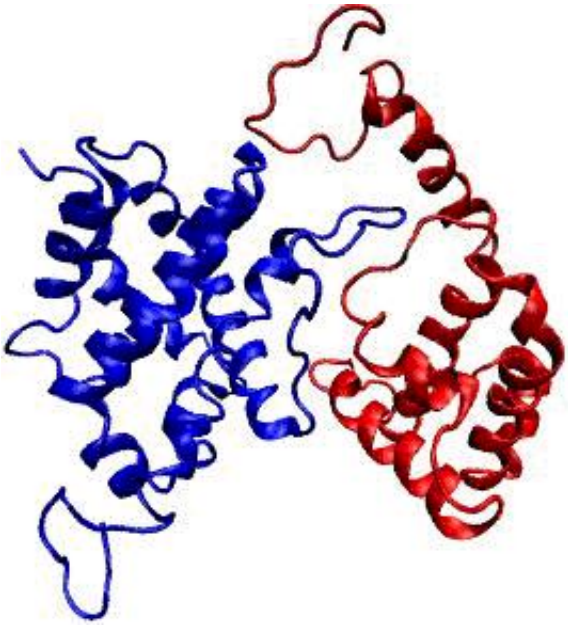


Figure 4.1: *Final structure of a short MD simulation of t-Bid (red) in close proximity of Bax (blue) at 300 K.*

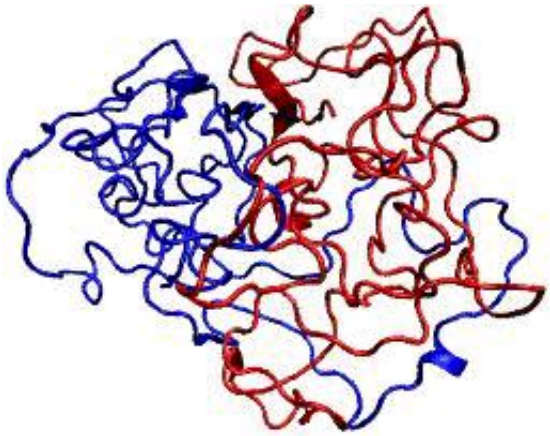


Figure 4.2: *Final structure of a short MD simulation of two Bax monomers close to each other at 500 K.*

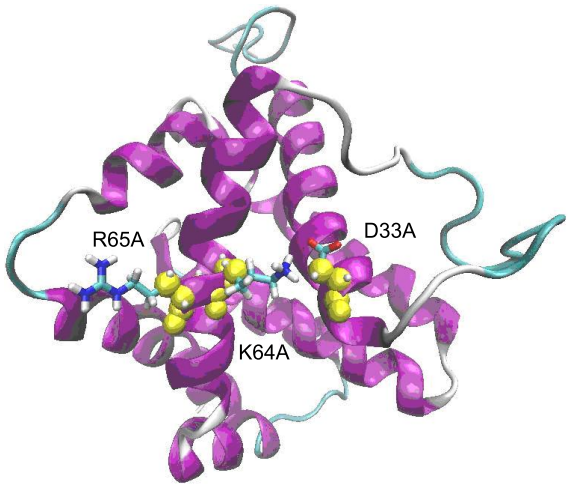


Figure 4.3: *Point mutations that should decrease the stability of Bax. The original residues (D33, K64 and R65) are depicted in licorice representation, while the alanines that would replace them are shown as yellow spheres.*

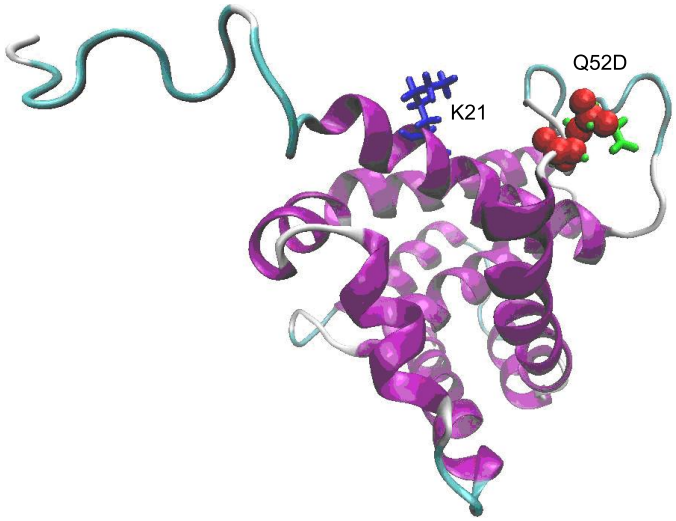


Figure 4.4: *Point mutation that should increase the stability of Bax. The original residue (Q52) is depicted in green, while the aspartic acid that would replace it is shown in red. The lysine (K21) that is found to interact with Q52 in the MD simulations is shown in blue.*

4.2.4 MD simulations of point-mutated structures to further study the stability of Bax

Site directed mutagenesis is a molecular biology technique that allows amino acid sequences in proteins to be altered at will. The mutated proteins can then be used to explore the effects on structure and function of the modifications performed [163]. In-silico mutagenesis consist in changing one or more side-chains of a given 3D structure, thus altering the amino acid sequence. MD simulations on the altered structures can provide insight into the role of the mutated residues in the structural stability of the protein [164].

Mutations that reduce the stability of Bax could potentially produce structures that transition into an active conformation more easily than the native structure. An example would be a mutation that perturbates the electrostatic interactions between D33 and either K64 or R65, which were discussed in Subsection 3.2.6. Mutating either of these residues to an alanine would probably reduce the structural stability of Bax (Fig. 4.3).

As discussed in Subsection 3.2.7, the interaction between Q52 and K21 partially stabilizes the proximity of the epitope 6A7 and Lp1-2 but this interaction is lost in the MD simulation at 400 K. The mutation Q52D could provide a stronger interaction to keep the epitope 6A7 covered at high temperatures (Fig. 4.4), and it would be interesting to test this hypothesis by means of MD simulations at high temperatures.

Appendix 1

Sequence of commands used for performing an MD simulation in Gromacs 4 with the infrastructure availabe in our research group.

Files needed:

`1F16.pdb` (NMR structure of Bax)

`em.mdp` (configuration file for equilibration)

`pr.mdp` (configuration file for MD simulation with restrictions)

`full.mdp` (configuration file for MD simulation)

```
g_pdb2gmx -ignh -f 1F16.pdb -p bax.top -o bax.gro

g_editconf -f bax.gro -o baxedit.gro -d 1.0

g_genbox -cp baxedit.gro -cs -p bax.top -o before_em.gro

g_grompp -v -f em.mdp -c before_em.gro -o em.tpr -p bax.top

g_genion -np 3 -pname Na+ -nname Cl- -s em.tpr -p bax.top -o baxion.gro

g_grompp -v -f em.mdp -c baxion.gro -o em.tpr -p bax.top -maxwarn 1

g_mdrun -v -s em.tpr -o em.trr -c after_emion.gro -g emlog.log

g_grompp -f pr.mdp -o pr.tpr -c after_emion.gro -r after_emion.gro
-p bax.top
```

```
/usr/lib64/openmpi/1.2.7-gcc/bin/mpirun -np 8 g_mdrun_mpi -v -s pr.tpr
-e pr.edr -o pr.trr -c after_pr.gro -g pr.log >& pr.job 1>
out.txt 2> err.out &

g_grompp -v -f full.mdp -o full.tpr -c after_pr.gro -p bax.top

nohup /usr/lib64/openmpi/1.2.7-gcc/bin/mpirun -np 16
--host node05,node06 g_mdrun_mpi
-v -s full.tpr -e full.edr -o full.trr
-c after_full.gro -g full.log
>& full.job 1>outfull.txt 2>errfull.txt &
```

Appendix 2

Configuration files used to perform the MD simulations.

```
MD Simulation File
Equilibration Box em
MD with restrictions Box pr
Full MD, 2 ns, 300 K Box full_2ns
Full MD, 18 ns, 300 K Box full_18ns
Full MD, 18 ns, 400 K Box full_18ns_400K
Full MD, 18 ns, 500 K Box full_18ns_500K
```

```
Box em. Configuration file used for equilibration.
;
;       User Jorge Luis Rosas Trigueros
;       Input file
;
cpp                 =  /usr/bin/cpp
define              =  -DFLEX_SPC
constraints         =  none
integrator          =  steep
nsteps              =  1000
;
;       Energy minimizing stuff
;
emtol               =  2000
emstep              =  0.01

nstcomm             =  1
```

```
ns_type             =  grid
rlist               =  1
rcoulomb            =  1.0
rvdw                =  1.0
Tcoupl              =  no
Pcoupl              =  no
gen_vel             =  no

*********************************



Box pr. Configuration file used for the MD simulation with restrictions.
;
;       User Jorge Luis Rosas Trigueros
;       Input file
;
title               =  Yo
cpp                 =  /usr/bin/cpp
define              =  -DPOSRES
constraints         =  all-bonds
integrator          =  md
dt                  =  0.002    ; ps !
nsteps              =  5000     ; total 10 ps.
nstcomm             =  1
nstxout             =  50
nstvout             =  1000
nstfout             =  0
nstlog              =  10
nstenergy           =  10
nstlist             =  10
ns_type             =  grid
rlist               =  1.0
rcoulomb            =  1.0
rvdw                =  1.0
; Berendsen temperature coupling is on in two groups
Tcoupl              =  berendsen
tc-grps             =  Protein  SOL Na+
```

```
tau_t               =  0.1      0.1 0.1
ref_t               =  300      300 300
; Energy monitoring
energygrps          =  Protein  SOL Na+
; Pressure coupling is not on
Pcoupl              =  no
tau_p               =  0.5
compressibility     =  4.5e-5
ref_p               =  1.0
; Generate velocites is on at 300 K.
gen_vel             =  yes
gen_temp            =  300.0
gen_seed            =  173529

*******************************



Box full_2ns. Configuration file used for the MD simulation of 2ns at 300 K.
;
;       User Jorge Luis Rosas Trigueros
;       Input file
;
title               =  Yo
cpp                 =  /usr/bin/cpp
constraints         =  all-bonds
integrator          =  md
dt                  =  0.002    ; ps !
nsteps              =  1000000  ; total 2000 ps =2ns.
nstcomm             =  1
nstxout             =  250
nstvout             =  1000
nstfout             =  0
nstlog              =  100
nstenergy           =  100
nstlist             =  10
ns_type             =  grid
rlist               =  1.0
```

```
rcoulomb            =  1.0
rvdw                =  1.0
;Electrostatics
coulombtype         =  PME
pme_order           =  4
fourierspacing      =  0.16
; Berendsen temperature coupling is on in two groups
Tcoupl              =  nose-hoover
tc-grps             =  Protein  SOL Na+
tau_t               =  0.1      0.1 0.1
ref_t               =  300      300 300
; Energy monitoring
energygrps          =  Protein  SOL Na+
; Isotropic pressure coupling is now on
Pcoupl              =  Parrinello-Rahman
Pcoupltype          = isotropic
tau_p               =  2.0
compressibility     =  4.5e-5
ref_p               =  1.0
; Generate velocites is off at 300 K.
gen_vel             =  no
gen_temp            =  300.0
gen_seed            =  173529

*********************************************

Box full_18ns. Configuration file used for the 18 ns of MD simulation
at 300 K following the MD simulation of 2 ns at 300 K.
;
;       User Jorge Luis Rosas Trigueros
;       Input file
;
title               =  Yo
cpp                 =  /usr/bin/cpp
constraints         =  all-bonds
integrator          =  md
dt                  =  0.002    ; ps !
nsteps              =  9000000  ; total 18000 ps =18ns.
```

```
nstcomm             =  1
nstxout             =  250
nstvout             =  1000
nstfout             =  0
nstlog              =  100
nstenergy           =  100
nstlist             =  10
ns_type             =  grid
rlist               =  1.0
rcoulomb            =  1.0
rvdw                =  1.0
;Electrostatics
coulombtype         =  PME
pme_order           =  4
fourierspacing      =  0.16
; Berendsen temperature coupling is on in two groups
Tcoupl              =  nose-hoover
tc-grps             =  Protein  SOL Na+
tau_t               =  0.1      0.1 0.1
ref_t               =  300      300 300
; Energy monitoring
energygrps          =  Protein  SOL Na+
; Isotropic pressure coupling is now on
Pcoupl              =  Parrinello-Rahman
Pcoupltype          = isotropic
tau_p               =  2.0
compressibility     =  4.5e-5
ref_p               =  1.0
; Generate velocites is off at 300 K.
gen_vel             =  no
gen_temp            =  400.0
gen_seed            =  173529

*********************************

Box full_18ns_400K. Configuration file used for the 18 ns of MD simulation
at 400 K following the MD simulation of 2 ns at 300 K.
;
```

```
;       User Jorge Luis Rosas Trigueros
;       Input file
;
title               =  Yo
cpp                 =  /usr/bin/cpp
constraints         =  all-bonds
integrator          =  md
dt                  =  0.002    ; ps !
nsteps              =  9000000  ; total 18000 ps =18ns.
nstcomm             =  1
nstxout             =  250
nstvout             =  1000
nstfout             =  0
nstlog              =  100
nstenergy           =  100
nstlist             =  10
ns_type             =  grid
rlist               =  1.0
rcoulomb            =  1.0
rvdw                =  1.0
;Electrostatics
coulombtype         =  PME
pme_order           =  4
fourierspacing      =  0.16
; Berendsen temperature coupling is on in two groups
Tcoupl              =  nose-hoover
tc-grps             =  Protein  SOL Na+
tau_t               =  0.1      0.1 0.1
ref_t               =  400      400 400
; Energy monitoring
energygrps          =  Protein  SOL Na+
; Isotropic pressure coupling is now on
Pcoupl              =  Parrinello-Rahman
Pcoupltype          = isotropic
tau_p               =  2.0
compressibility     =  4.5e-5
ref_p               =  1.0
; Generate velocites is off at 300 K.
```

```
gen_vel             =  no
gen_temp            =  400.0
gen_seed            =  173529

****************************

Box full_18ns_500K. Configuration file used for the 18 ns of MD simulation
at 500 K following the MD simulation of 2 ns at 300 K.
;
;       User Jorge Luis Rosas Trigueros
;       Input file
;
title               =  Yo
cpp                 =  /usr/bin/cpp
constraints         =  all-bonds
integrator          =  md
dt                  =  0.002    ; ps !
nsteps              =  9000000  ; total 18000 ps =18ns.
nstcomm             =  1
nstxout             =  250
nstvout             =  1000
nstfout             =  0
nstlog              =  100
nstenergy           =  100
nstlist             =  10
ns_type             =  grid
rlist               =  1.0
rcoulomb            =  1.0
rvdw                =  1.0
;Electrostatics
coulombtype         =  PME
pme_order           =  4
fourierspacing      =  0.16
; Berendsen temperature coupling is on in two groups
Tcoupl              =  nose-hoover
tc-grps             =  Protein  SOL Na+
tau_t               =  0.1      0.1 0.1
ref_t               =  500      500 500
```

```
; Energy monitoring
energygrps          =  Protein  SOL Na+
; Isotropic pressure coupling is now on
Pcoupl              =  Parrinello-Rahman
Pcoupltype          = isotropic
tau_p               =  2.0
compressibility     =  4.5e-5
ref_p               =  1.0
; Generate velocites is off at 300 K.
gen_vel             =  no
gen_temp            =  500.0
gen_seed            =  173529
```

Bibliography

[1] G. Melino. The Sirens' song. *Nature*, 412(6842):23–23, 2001.

[2] G. Kroemer, W.S. El-Deiry, P. Golstein, M.E. Peter, D. Vaux, P. Vandenabeele, B. Zhivotovsky, M.V. Blagosklonny, W. Malorni, and R.A. Knight. Classification of cell death: recommendations of the Nomenclature Committee on Cell Death. *Cell Death & Differentiation*, 12:1463–1467, 2005.

[3] J.F.R. Kerr, A.H. Wyllie, and A.R. Currie. Apoptosis: a basic biological phenomenon with wideranging implications in tissue kinetics. *British Journal of Cancer*, 26(4):239–257, 1972.

[4] C.B. Thompson. Apoptosis in the pathogenesis and treatment of disease. *Science*, 267(5203):1456–1462, 1995.

[5] L. Galluzzi, M.C. Maiuri, I. Vitale, H. Zischka, M. Castedo, L. Zitvogel, and G. Kroemer. Cell death modalities: classification and pathophysiological implications. *Cell Death & Differentiation*, 14(7):1237–1243, 2007.

[6] S.W. Lowe and A.W. Lin. Apoptosis in cancer. *Carcinogenesis*, 21(3):485–495, 2000.

[7] C. Sergi, P. Kahl, and H.F. Otto. Contribution of apoptosis and apoptosis-related proteins to the malformation of the primitive intrahepatic biliary system in Meckel syndrome. *The American Journal of Pathology*, 156(5):1589–1598, 2000.

[8] B. Zhivotovsky and S. Orrenius. Cell death mechanisms: cross-talk and role in disease. *Experimental Cell Research*, 316(8):1374–1383, 2010.

[9] M.E. Soriano and L. Scorrano. Traveling Bax and Forth from Mitochondria to Control Apoptosis. *Cell*, 145(1):15–17, 2011.

[10] D.R. Green and G. Kroemer. The pathophysiology of mitochondrial cell death. *Science*, 305(5684):626–629, 2004.

[11] D. Brenner and T.W. Mak. Mitochondrial cell death effectors. *Current Opinion in Cell Biology*, 21(6):871–877, 2009.

[12] Y. Tsujimoto, L.R. Finger, J. Yunis, P.C. Nowell, and C.M. Croce. Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. *Science*, 226(4678):1097–1099, 1984.

[13] C. Wang and R.J. Youle. The role of mitochondria in apoptosis. *Annual Review of Genetics*, 43:95–118, 2009.

[14] J.E. Chipuk, T. Moldoveanu, F. Llambi, M.J. Parsons, and D.R. Green. The BCL-2 family reunion. *Molecular Cell*, 37(3):299–310, 2010.

[15] E. Szegezdi, D.C. MacDonald, T.N. Chonghaile, S. Gupta, and A. Samali. Bcl-2 family on guard at the ER. *American Journal of Physiology Cell Physiology*, 296(5):C941–C953, 2009.

[16] R.B.C. Van Ooteghem, E.M.E. Smit, A. Beishuizen, A.C. Lambrechts, M. Blij-Philipsen, T.J. Smilde, and A. Hagemeijer. A new B-cell line showing a complex translocation (8; 14; 18) and BCL2 rearrangement. *Cancer Genetics and Cytogenetics*, 74(2):87–94, 1994.

[17] M. Kandouz, M. Siromachkova, D. Jacob, B.C. Marquet, A. Therwath, and A. Gompel. Antagonism between estradiol and progestin on Bcl-2 expression in breast-cancer cells. *International Journal of Cancer*, 68(1):120–125, 1996.

[18] O. Imamov, A. Morani, G.J. Shim, Y. Omoto, C. Thulin-Andersson, M. Warner, and J.Å. Gustafsson. Estrogen receptor β regulates epithelial cellular differentiation in the mouse ventral prostate. *Proceedings of the National Academy of Sciences of the United States of America*, 101(25):9375–9380, 2004.

[19] D.M. Hockenbery, M. Zutter, W. Hickey, M. Nahm, and S.J. Korsmeyer. BCL2 Protein is Topographically Restricted in Tissues Characterized by Apoptotic Cell Death. *Proceedings of the National Academy of Sciences of the United States of America*, 88(16):6961–6965, 1991.

[20] D. Maskey, K.S. Park, J. Pradhan, and Y.B. Yoo. Bcl-2 Expression in the Developing Hair Follicle of Korean Fetal Scalp. *The Korean Journal of Anatomy*, 42(4):315–323, 2009.

[21] D.E. Merry, D.J. Veis, W.F. Hickey, and S.J. Korsmeyer. bcl-2 protein expression is widespread in the developing nervous system and retained in the adult PNS. *Development*, 120(2):301–311, 1994.

[22] T. Wester, Y. Olsson, and L. Olsen. Expression of bcl-2 in Enteric Neurons in Normal Human Bowel and Hirschsprung Disease. *Archives of Pathology & Laboratory Medicine*, 123(12):1264–1268, 1999.

[23] P.N. Kelly and A. Strasser. The role of Bcl-2 and its pro-survival relatives in tumourigenesis and cancer therapy. *Cell Death & Differentiation*, 18(9):1414–1424, 2011.

[24] M.O. Hengartner and H.R. Horvitz. *C. elegans* cell survival gene ced-9 encodes a functional homolog of the mammalian proto-oncogene bcl-2. *Cell*, 76(4):665–676, 1994.

[25] C. Lanave, M. Santamaria, and C. Saccone. Comparative genomics: the evolutionary history of the Bcl-2 family. *Gene*, 333:71–79, 2004.

[26] J.K. Brunelle and A. Letai. Control of mitochondrial apoptosis by the Bcl-2 family. *Journal of Cell Science*, 122(4):437–441, 2009.

[27] G. Dewson and R.M. Kluck. Bcl-2 family-regulated apoptosis in health and disease. *Cell Health and Cytoskeleton*, 2:9–22, 2010.

[28] P.E. Czabotar, P.M. Colman, and D.C. Huang. Bax activation by Bim? *Cell Death & Differentiation*, 16(9):1187–1191, 2009.

[29] S.N. Willis, J.I. Fletcher, T. Kaufmann, M.F. van Delft, L. Chen, P.E. Czabotar, H. Ierino, E.F. Lee, W.D. Fairlie, P. Bouillet, A. Strasser, R.M. Kluck, J.M. Adams, and D.C. Huang. Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. *Science*, 315(5813):856–859, 2007.

[30] Z.N. Oltvai, C.L. Milliman, and S.J. Korsmeyer. Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programed cell death. *Cell*, 74(4):609–619, 1993.

[31] S.N. Farrow, J.H.M. White, I. Martinou, T. Raven, K.T. Pun, C.J. Grinham, J.C. Martinou, and R. Brown. Cloning of a bcl-2 homologue by interaction with adenovirus E1B 19K. *Nature*, 374:731–733, 1995.

[32] T. Moldoveanu, Q. Liu, A. Tocilj, M. Watson, G. Shore, and K. Gehring. The X-ray structure of a BAK homodimer reveals an inhibitory zinc binding site. *Molecular Cell*, 24(5):677–688, 2006.

[33] X.M. Yin, Z.N. Oltvai, and S.J. Korsmeyer. BH1 and BH2 domains of Bcl-2 are required for inhibition of apoptosis and heterodimerization with Bax. *Nature*, 369(6478):321–323, 1994.

[34] E.F. Lee, G. Dewson, B.J. Smith, M. Evangelista, A. Pettikiriarachchi, C. Dogovski, M.A. Perugini, P.M. Colman, and W.D. Fairlie. Crystal Structure of a BCL-W Domain-Swapped Dimer: Implications for the Function of BCL-2 Family Proteins. *Structure*, 19(10):1467–1476, 2011.

[35] E. Yang, J. Zha, J. Jockel, L.H. Boise, C.B. Thompson, and S.J. Korsmeyer. Bad, a heterodimeric partner for Bcl-x_L and Bcl-2, displaces bax and promotes cell death. *Cell*, 80(2):285–291, 1995.

[36] S.Y. Jeong, B. Gaume, Y.J. Lee, Y.T. Hsu, S.W. Ryu, S.H. Yoon, and R.J. Youle. Bcl-xl sequesters its c-terminal membrane anchor in soluble, cytosolic homodimers. *The EMBO Journal*, 23(10):2146–2155, 2004.

[37] A.E. Douglas, K.D. Corbett, J.M. Berger, G. McFadden, and T.M. Handel. Structure of M11L: A myxoma virus structural homolog of the apoptosis inhibitor, Bcl-2. *Protein Science*, 16(4):695–703, 2007.

[38] V. Ganesan, M.N. Perera, D. Colombini, D. Datskovskiy, K. Chadha, and M. Colombini. Ceramide and activated Bax act synergistically to permeabilize the mitochondrial outer membrane. *Apoptosis*, 15(5):553–562, 2010.

[39] Z.T. Schug and E. Gottlieb. Cardiolipin acts as a mitochondrial signalling platform to launch apoptosis. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1788(10):2022–2031, 2009.

[40] R.S. Herbst and S.R. Frankel. Oblimersen Sodium (Genasense bcl-2 Antisense Oligonucleotide). *Clinical Cancer Research*, 10(12):4245s–4248s, 2004.

[41] A. Masood, A.S. Azmi, and R.M. Mohammad. Small Molecule Inhibitors of Bcl-2 Family Proteins for Pancreatic Cancer Therapy. *Cancers*, 3(2):1527–1549, 2011.

[42] C. Cohen-Saidon, H. Nechushtan, S. Kahlon, N. Livni, A. Nissim, and E. Razin. A novel strategy using single-chain antibody to show the importance of Bcl-2 in mast cell survival. *Blood*, 102(7):2506–2512, 2003.

[43] J.L. Wang, Z.J. Zhang, S. Choksi, S. Shan, Z. Lu, C.M. Croce, E.S. Alnemri, R. Korngold, and Z. Huang. Cell permeable Bcl-2 binding peptides: a chemical approach to apoptosis induction in tumor cells. *Cancer Research*, 60(6):1498–1502, 2000.

[44] L.D. Walensky, K. Pitter, J. Morash, K.J. Oh, S. Barbuto, J. Fisher, E. Smith, G.L. Verdine, and S.J. Korsmeyer. A stapled BID BH3 helix directly binds and activates BAX. *Molecular Cell*, 24(2):199–210, 2006.

[45] D.J. Muilenburg, J.M. Coates, S. Virudachalam, and R.J. Bold. Targeting Bcl-2-Mediated Cell Death as a Novel Therapy in Pancreatic Cancer. *Journal of Surgical Research*, 163(2):276–281, 2010.

[46] M. Suzuki, R.J. Youle, and N. Tjandra. Structure of Bax: coregulation of dimer formation and intracellular localization. *Cell*, 103(4):645–654, 2000.

[47] M. Sánchez-Aguilar, L.A. Marchat, and A. Zamorano. Prediction of a Putative Functional Region in the Human Bax Protein by Computational Analysis. *American Journal of Infectious Diseases*, 3(2):68–75, 2007.

[48] P.F. Cartron, T. Gallenne, G. Bougras, F. Gautier, F. Manero, P. Vusio, K. Meflah, F. M. Vallette, and P. Juin. The First α-Helix of Bax Plays a Necessary Role in Its Ligand-Induced Activation by the BH3-Only Proteins Bid and PUMA. *Molecular Cell*, 16(5):807–818, 2004.

[49] I.S. Goping, A. Gross, J.N. Lavoie, M. Nguyen, R. Jemmerson, K. Roth, S.J. Korsmeyer, and G.C. Shore. Regulated targeting of BAX to mitochondria. *The Journal of Cell Biology*, 143(1):207–215, 1998.

[50] Z. Zhang, W. Zhu, S.M. Lapolla, Y. Miao, Y. Shao, M. Falcone, D. Boreham, N. McFarlane, J. Ding, A.E. Johnson, X.C. Zhang, D.W. Andrews, and J. Lin. Bax forms an oligomer via separate, yet interdependent, surfaces. *The Journal of Biological Chemistry*, 285(23):17614–17627, 2010.

[51] S. Bleicken and K. Zeth. Conformational changes and protein stability of the pro-apoptotic protein Bax. *Journal of Bioenergetics and Biomembranes*, 41(1):29–40, 2009.

[52] L. Lalier, P.F. Cartron, P. Juin, S. Nedelkina, S. Manon, B. Bechinger, and F.M. Vallette. Bax activation and mitochondrial insertion during apoptosis. *Apoptosis*, 12(5):887–896, 2007.

[53] K.J. Oh, S. Barbuto, K. Pitter, J. Morash, L.D. Walensky, and S.J. Korsmeyer. A membrane-targeted BID BCL-2 homology 3 peptide is sufficient for high potency activation of BAX in vitro. *The Journal of Biological Chemistry*, 281(48):36999–37008, 2006.

[54] D. Lama and R. Sankararamakrishnan. Anti-apoptotic Bcl-XL protein in complex with BH3 peptides of pro-apoptotic Bak, Bad, and Bim proteins: Comparative molecular dynamics simulations. *Proteins: Structure, Function, and Bioinformatics*, 73(2):492–514, 2008.

[55] E. Fire, S.V. Gullá, R.A. Grant, and A.E. Keating. Mcl-1–Bim complexes accommodate surprising point mutations via minor structural changes. *Protein Science*, 19(3):507–519, 2010.

[56] E. Gavathiotis, M. Suzuki, M.L. Davis, K. Pitter, G.H. Bird, S.G. Katz, H.C. Tu, H. Kim, E.H. Cheng, N. Tjandra, and L.D. Walensky. Bax activation is initiated at a novel interaction site. *Nature*, 455(07396):1076–1081, 2008.

[57] J. Ding, Z. Zhang, G.J. Roberts, M. Falcone, Y. Miao, Y. Shao, X.C. Zhang, D.W. Andrews, and J. Lin. Bcl-2 and Bax Interact via the BH1-3 Groove- BH3 Motif Interface and a Novel Interface Involving the BH4Motif. *The Journal of Biological Chemistry*, 285(37):28749–28763, 2010.

[58] S. Bleicken, M. Classen, P.V.L. Padmavathi, T. Ishikawa, K. Zeth, H.J. Steinhoff, and E. Bordignon. Molecular Details of Bax Activation, Oligomerization, and Membrane Insertion. *Journal of Biological Chemistry*, 285(9):6636–6647, 2010.

[59] V.G. Veresov and A.I. Davidovskii. Activation of Bax by joint action of tBid and mitochondrial outer membrane: Monte Carlo simulations. *European Biophyics Journal*, 38(7):941–960, 2009.

[60] S. Krajewski, C. Blomqvist, K. Franssila, M. Krajewska, V.M. Wasenius, E. Niskanen, S. Nordling, and J.C. Reed. Reduced expression of proapoptotic gene BAX is associated with poor response rates to combination chemotherapy and shorter survival in women with metastatic breast adenocarcinoma. *Cancer Research*, 55(19):4471–4478, 1995.

[61] R.C. Bargou, P.T. Daniel, M.Y. Mapara, K. Bommert, C. Wagener, B. Kallinich, H.D. Royer, and B. Dörken. Expression of the bcl-2 gene family in normal and malignant breast tissue: Low bax-α expression in tumor cells correlates with resistance towards apoptosis. *International Journal of Cancer*, 60(6):854–859, 1995.

[62] C. Sakakura, E.A. Sweeney, T. Shirahama, Y. Igarashi, S.I. Hakomori, H. Tsujimoto, T. Imanishi, M. Ohgaki, J. Yamazaki, and A. Hagiwara. Overexpression of bax enhances the radiation sensitivity in human breast cancer cells. *Surgery Today*, 27(1):90–93, 1997.

[63] C. Sakakura, E.A. Sweeney, T. Shirahama, Y. Igarashi, S. Hakomori, H. Nakatani, H. Tsujimoto, T. Imanishi, M. Ohgaki, and T. Ohyama. Overexpression of bax sensitizes human breast cancer MCF-7 cells to radiation-induced apoptosis. *International Journal of Cancer*, 67(1):101–105, 1996.

[64] C. Wagener, R.C. Bargou, P.T. Daniel, K. Bommert, M.Y. Mapara, H.D. Royer, and B. Dörken. Induction of the death-promoting gene bax-alpha sensitizes cultured breast-cancer cells to drug-induced apoptosis. *International Journal of Cancer*, 67(1):138–141, 1996.

[65] S. Krajewski, J.K. Mai, M. Krajewska, M. Sikorska, M.J. Mossakowski, and J.C. Reed. Upregulation of bax protein levels in neurons following cerebral ischemia. *The Journal of Neuroscience*, 15(10):6364–6376, 1995.

[66] A.V. Vaseva and U.M. Moll. The mitochondrial p53 pathway. *Biochimica et Biophysica Acta*, 1787(5):414–420, 2009.

[67] W.T. Lee and C.W. Chang. Bax is upregulated by p53 signal pathway in the SPE B-induced apoptosis. *Molecular and Cellular Biochemistry*, 343(1-2):271–279, 2010.

[68] T. Sun, X. Lin, Y. Wei, Y. Xu, and P. Shen. Evaluating bistability of Bax activation switch. *FEBS Letters*, 584(5):954–960, 2010.

[69] Y.T. Hsu, K.G. Wolter, and R.J. Youle. Cytosol-to-membrane redistribution of Bax and Bcl-X(L) during apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8):3668–3672, 1997.

[70] A. Jourdain and J.C. Martinou. Mitochondrial outer-membrane permeabilization and remodelling in apoptosis. *The International Journal of Biochemistry & Cell Biology*, 41(10):1884–1889, 2009.

[71] P.H. Schlesinger and M. Saito. The Bax pore in liposomes Biophysics. *Cell Death and Differentiation*, 13(8):1403–1408, 2006.

[72] S. Qian, W. Wang, L. Yang, and H.W. Huang. Structure of transmembrane pore induced by Bax-derived peptide: evidence for lipidic pores. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45):17379–17383, 2008.

[73] D.E. Wood and E.W. Newcomb. Cleavage of Bax enhances its cell death function. *Experimental Cell Research*, 256(2):375–382, 2000.

[74] W. Fujibuchi, S. Goto, H. Migimatsu, I. Uchiyama, A. Ogiwara, Y. Akiyama, and M. Kanehisa. DBGET/LinkDB: an integrated database retrieval system. In *Pacific Symposium on Biocomputing*, volume 98, pages 683–694, 1998.

[75] A. Tramontano. *Protein Structure Prediction: Concepts and Applications*. Wiley-VCH, 2006.

[76] T. Schlick. *Molecular Modeling and Simulation*. Springer, 2002.

[77] G. Cuevas. *Introducción a la química computacional*. Fondo de Cultura Económica, México, 2003.

[78] R. Durbin. *Biological sequence analysis: Probabilistic models of proteins and nucleic acids*. Cambridge University Press, 1998.

[79] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, and J. Söding. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(539):1–6, 2011.

[80] J.D. Thompson, P. Koehl, R. Ripp, and O. Poch. BAliBASE 3.0: latest developments of the multiple sequence alignment benchmark. *Proteins: Structure, Function, and Bioinformatics*, 61(1):127–136, 2005.

[81] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, 2008.

[82] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A. Sijbers, K.A. Feenstra, R. van Drunen, and H.J.C. Berendsen. *Gromacs User Manual version 3.3*, 2008.

[83] N. Bharatham, S.W. Chi, and H.S. Yoon. Molecular basis of Bcl-XL-p53 interaction: Insights from Molecular Dynamics Simulations. *PloS one*, 6(10):e26014, 2011.

[84] P.K. Raghav, Y.K. Verma, and G.U. Gangenahalli. Molecular dynamics simulations of the Bcl-2 protein to predict the structure of its unordered flexible loop domain. *Journal of Molecular Modeling*, pages 1–22, 2011. DOI: 10.1007/s00894-011-1201-6.

[85] S. Acoca, Q. Cui, G.C. Shore, and E.O. Purisima. Molecular dynamics study of small molecule inhibitors of the Bcl-2 family. *Proteins: Structure, Function, and Bioinformatics*, 79(9):2624–2636, 2011.

[86] M. Pinto, M. Orzaez, L. Delgado-Soler, J.J. Perez, and J. Rubio-Martinez. Rational Design of New Class of BH3-Mimetics As Inhibitors of the Bcl-xL Protein. *Journal of Chemical Information and Modeling*, 51(6):1249–1258, 2011.

[87] D. Lama and R. Sankararamakrishnan. Molecular dynamics simulations of pro- apoptotic BH3 peptide helices in aqueous medium: relationship between helix stability and their binding affinities to the anti-apoptotic protein Bcl-X_L. *Journal of Computer-Aided Molecular Design*, 25(5):413–426, 2011.

[88] C. Koshy, M. Parthiban, and R. Sowdhamini. 100 ns molecular dynamics simulations to study intramolecular conformational changes in Bax. *Journal of Biomolecular Structure & Dynamics*, 28(1):71–83, 2010.

[89] F. Mancinelli, M. Caraglia, A. Budillon, A. Abbruzzese, and E. Bismuto. Molecular dynamics simulation and automated docking of the pro-apoptotic Bax protein and its complex with a peptide designed from the Bax-binding domain of anti-apoptotic Ku70. *Journal of Cellular Biochemistry*, 99(1):305–318, 2006.

[90] M. Pinto, J.J. Perez, and J. Rubio-Martinez. Molecular dynamics study of peptide segments of the BH3 domain of the proapoptotic proteins Bak, Bax, Bid and Hrk bound to the Bcl-x L and Bcl-2 proteins. *Journal of Computer-Aided Molecular Design*, 18(1):13–22, 2004.

[91] J.L. Rosas-Trigueros, J. Correa-Basurto, C.G. Benítez-Cardoza, and A. Zamorano-Carrillo. Insights into the structural stability of Bax from molecular dynamics simulations at high temperatures. *Protein Science*, 20(12):2035–2046, 2011.

[92] J. Kubelka, J. Hofrichter, and W.A. Eaton. The protein folding 'speed limit'. *Current Opinion in Structural Biology*, 14(1):76–88, 2004.

[93] J.N. Onuchic and P.G. Wolynes. Theory of protein folding. *Current opinion in structural biology*, 14(1):70–75, 2004.

[94] P.M.K. Mohan and R.V. Hosur. Structure-function-folding relationships and native energy landscape of dynein light chain protein: nuclear magnetic resonance insights. *Journal of Biosciences*, 34(3):465–479, 2009.

[95] C. Scharnagl, M. Reif, and J. Friedrich. Stability of proteins: Temperature, pressure and the role of the solvent. *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics*, 1749(2):187–213, 2005.

[96] V.L. Shnyrov, J.M. Sanchez-Ruiz, B.N. Boiko, G.G. Zhadan, and E.A. Permyakov. Applications of scanning microcalorimetry in biophysics and biochemistry. *Thermochimica Acta*, 302(1-2):165–180, 1997.

[97] D.I. Levitsky, A.V. Pivovarova, V.V. Mikhailova, and O.P. Nikolaeva. Thermal unfolding and aggregation of actin. *Febs Journal*, 275(17):4280–4295, 2008.

[98] V. Daggett. Protein folding-simulation. *Chemical Reviews*, 106(5):1898–1916, 2006.

[99] S.E. Brock, C. Li, and B.W. Wattenberg. The Bax carboxy-terminal helix does not determine organelle-specific targeting but is essential for maintaining Bax in an inactive state and for stable mitochondrial membrane insertion. *Apoptosis*, 15(1):14–27, 2010.

[100] P.F. Cartron, L. Oliver, E. Mayat, K. Meflah, and F.M. Vallette. Impact of pH on Bax alpha conformation, oligomerisation and mitochondrial integration. *FEBS Letters*, 578(1-2):41–46, 2004.

[101] H. Zhou, Q. Hou, J.L. Hansen, and Y.T. Hsu. Complete activation of Bax by a single site mutation. *Oncogene*, 26(50):7092–7102, 2007.

[102] Z. Huang. Bcl-2 family proteins as targets for anticancer drug design. *Oncogene*, 19(56):6627–6631, 2000.

[103] S.Y. Oh, B. Cornell, D. Smith, G. Higgins, C.J. Burrell, and T.W. Kok. Rapid detection of influenza A virus in clinical samples using an ion channel switch biosensor. *Biosensors and Bioelectronics*, 23(7):1161–1165, 2008.

[104] R. Day, D. Paschek, and A.E. Garcia. Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins*, 78(8):1889–1899, 2010.

[105] C. Yan, V. Pattani, J.W. Tunnell, and P. Ren. Temperature-induced unfolding of epidermal growth factor (EGF): Insight from molecular dynamics simulation. *Journal of Molecular Graphics Modelling.*, 29(1):2–12, 2010.

[106] Y.W. Lin, F.Y. Ni, and T.L. Ying. Early events in thermal unfolding of apocytochrome b562 and its double-cysteine mutant as revealed by molecular dynamics simulation. *THEOCHEM*, 898(1-3):82–89, 2009.

[107] C.P. Chng and A. Kitao. Thermal unfolding simulations of bacterial flagellin: insight into its refolding before assembly. *Biophysical Journal*, 94(10):3858–3871, 2008.

[108] D. Paschek, S. Hempel, and A.E. García. Computing the stability diagram of the Trp-cage miniprotein. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46):17754–17759, 2008.

[109] L. Wang, Y. Duan, R. Shortle, B. Imperiali, and P.A. Kollman. Study of the stability and unfolding mechanism of BBA1 by molecular dynamics simulations at different temperatures. *Protein Science*, 8(6):1292–1304, 1999.

[110] The UniProt Consortium. Ongoing and future developments at the universal protein resource. *Nucleic Acids Research*, 39(Suppl 1):D214–D219, 2011.

[111] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, and D.G. Higgins. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948, 2007.

[112] J.J. Wendoloski, S.J. Kimatian, C.E. Schutt, and F.R. Salemme. Molecular dynamics simulation of a phospholipid micelle. *Science*, 243(4891):636–638, 1989.

[113] W.L. Jorgensen and J. Tirado-Rives. The OPLS potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1998.

[114] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, and W.L. Jorgensen. Evaluation and Reparametrization of the OPLS- AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *The Journal of Physical Chemistry B*, 105(28):6474–6487, 2001.

[115] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, 106(3):765–784, 1984.

[116] W.L. Jorgensen and J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19):6665–6670, 2005.

[117] H.J.C. Berendsen, J.P.M. Postma, W.F. Van Gunsteren, and J. Hermans. *Interaction models for water in relation to protein hydration*, pages 331–342. Reidel: Dordrecht, The Netherlands, 1981.

[118] D. van der Spoel and H.J. Berendsen. Molecular dynamics simulations of Leuenkephalin in water and DMSO. *Biophysical Journal*, 72(5):2032–2041, 1997.

[119] P.P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.

[120] T. Darden, D. York, and L. Pedersen. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *The Journal of Physical Chemistry C*, 98(12):10089–10092, 1993.

[121] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. A smooth particle mesh ewald potential. *The Journal of Physical Chemistry A*, 103:8577–8593, 1995.

[122] S. Nos. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268, 1984.

[123] W.G. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697, 1985.

[124] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1981.

[125] S. Nos and M.L. Klein. Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076, 1983.

[126] B. Hess, H. Bekker, H.J.C. Berendsen, and J.G.E.M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18:1463–1472, 1997.

[127] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–2637, 1983.

[128] W. Humphrey, A. Dalke, and K. Schulten. VMD - Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 1996.

[129] W.G. Krebs, V. Alexandrov, C.A. Wilson, N. Echols, H. Yu, and M. Gerstein. Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. *Proteins*, 48(4):682–695, 2002.

[130] K. Suhre and Y.H. Sanejouand. *ElNémo*: a normal mode web-server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Research*, 32(2):W610–W614, 2004.

[131] F. Tama, F.X. Gadea, O. Marques, and Y.H. Sanejouand. Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins: Structure, Function, and Bioinformatics*, 41(1):1–7, 2000.

[132] S.N. Murthy, S. Iismaa, G. Begg, D.M. Freymann, R.M. Graham, and L. Lorand. Conserved tryptophan in the core domain of transglutaminase is essential for catalytic activity. *Proceedings of the National Academy of Sciences of the United States of America*, 99(5):2738–2742, 2002.

[133] N. Chen, R.A. Vaughan, and M.E.A. Reith. The role of conserved tryptophan and acidic residues in the human dopamine transporter as characterized by site-directed mutagenesis. *Journal of Neurochemistry*, 77(4):1116–1127, 2001.

[134] D.K. Williams, C. Stokes, N.A. Horenstein, and R.L. Papke. Differential regulation of receptor activation and agonist selectivity by highly conserved tryptophans in the nicotinic acetylcholine receptor binding site. *Journal of Pharmacology and Experimental Therapeutics*, 330(1):40–53, 2009.

[135] H. Li, J.L. Wojtaszek, and L.H. Greene. Analysis of conservation in the Fas-associated death domain protein and the importance of conserved tryptophans in structure, stability and folding. *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics*, 1794(4):583–593, 2009.

[136] A. Murad and E. Rech. Molecular dynamics simulations of the minor ampullate spidroin modular amino acid sequence from *Parawixia bistriatra*: insights into silk tertiary structure and fibre formation. *Journal of Molecular Modeling*, 17:1183–1189, 2011.

[137] P.R. Batista, M. Garcia de Souza Costa, P.G. Pascutti, P.M. Bisch, and W. de Souza. High temperatures enhance cooperative motions between CBM and catalytic domains of a thermostable cellulase: mechanism insights from essential dynamics. *Physical Chemistry Chemical Physics*, 13:13709–13720, 2011.

[138] J. Cui, C. Chen, H. Lu, T. Sun, and P. Shen. Two independent positive feedbacks and bistability in the Bcl-2 apoptotic switch. *PLoS One*, 3(1):e1469, 2008.

[139] E.Z. Bagci, Y. Vodovotz, T.R. Billiar, G.B. Ermentrout, and I. Bahar. Bistability in apoptosis: roles of bax, bcl-2, and mitochondrial permeability transition pores. *Biophysical Journal*, 90(5):1546–1559, 2006.

[140] J.K. Shen. Uncovering Specific Electrostatic Interactions in the Denatured States of Proteins. *Biophysical Journal*, 99(3):924–932, 2010.

[141] J.A. Ballesteros, A.D. Jensen, G. Liapakis, S.G.F. Rasmussen, L. Shi, U. Gether, and J.A. Javitch. Activation of the β2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *Journal of Biological Chemistry*, 276(31):29171–29177, 2001.

[142] D.C. Benjamin and S.S. Perdue. Site-directed mutagenesis in epitope mapping. *Methods*, 9(3):508–515, 1996.

[143] D.A. Dougan, R.L. Malby, L.C. Gruen, A.A. Kortt, and P.J. Hudson. Effects of substitutions in the binding surface of an antibody on antigen affinity. *Protein engineering desing and selection*, 11(1):65–74, 1998.

[144] B.C. Braden and R.J. Poljak. Structural features of the reactions between antibodies and protein antigens. *The FASEB Journal*, 9(1):9–16, 1995.

[145] D.R. Jackola, C. Blackburn, M. Sveum, and A. Rosenberg. Entropy-favored human antibody binding reactions with a non-infectious antigen. *Molecular Immunology*, 45(5):1494–1500, 2008.

[146] G. Zeder-Lutz, E. Zuber, J. Witz, and M.H.V. Van Regenmortel. Thermodynamic analysis of antigen-antibody binding using biosensor measurements at different temperatures. *Analytical Biochemistry*, 246(1):123–132, 1997.

[147] T. Sagawa, M. Oda, M. Ishimura, K. Furukawa, and T. Azuma. Thermodynamic and kinetic aspects of antibody evolution during the immune response to hapten. *Molecular Immunology*, 39(13):801–808, 2003.

[148] S.K. Burley and G.A. Petsko. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, 229(4708):23–28, 1985.

[149] A.M. Petros, E.T. Olejniczak, and S.W. Fesik. Structural biology of the Bcl-2 family of proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1644(2-3):83–94, 2004.

[150] F. Tama and Y.H. Sanejouand. Conformational change of proteins arising from normal mode calculations. *Protein Engineering desing and selection*, 14(1):1–6, 2001.

[151] R.M. Levy, D. Perahia, and M. Karplus. Molecular dynamics of an alpha-helical polypeptide: Temperature dependence and deviation from harmonic behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 79(4):1346–1350, 1982.

[152] J.E. Shea and C.L. Brooks III. From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annual Review of Physical Chemistry*, 52(1):499–535, 2001.

[153] F. Sancar, S.S. Ericksen, A.M. Kucken, J.A. Teissere, and C. Czajkowski. Structural determinants for high-affinity zolpidem binding to GABA-A receptors. *Molecular Pharmacology*, 71(1):38–46, 2007.

[154] D. Provasi and M. Filizola. Putative Active States of a Prototypic G-Protein-Coupled Receptor from Biased Molecular Dynamics. *Biophysical Journal*, 98(10):2347–2355, 2010.

[155] S. Kumar and R. Nussinov. Close Range Electrostatic Interactions in Proteins Close-Range Electrostatic Interactions in Proteins. *ChemBioChem*, 3(7):604–617, 2002.

[156] G. Schreiber and A.R. Fersht. Energetics of protein-protein interactions: Analysis ofthe Barnase-Barstar interface by single mutations and double mutant cycles. *Journal of Molecular Biology*, 248(2):478–486, 1995.

[157] G.U. Gurudutta, Y.K. Verma, V.K. Singh, P. Gupta, H.G. Raj, R.K. Sharma, and R. Chandra. Structural conservation of residues in BH1 and BH2 domains of Bcl-2 family proteins. *FEBS Letters*, 579(17):3503–3507, 2005.

[158] K. Wang, X.M. Yin, D.T. Chao, C.L. Milliman, and S.J. Korsmeyer. BID: a novel BH3 domain-only death agonist. *Genes & Development*, 10(22):2859–2869, 1996.

[159] H. Li, H. Zhu, C. Xu, and J. Yuan. Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell*, 94(4):491–501, 1998.

[160] E. Er, L. Oliver, P.F. Cartron, P. Juin, S. Manon, and F.M. Vallette. Mitochondria as the target of the pro-apoptotic protein Bax. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1757(9-10):1301–1311, 2006.

[161] W.X. Zong, C. Li, G. Hatzivassiliou, T. Lindsten, Q.C. Yu, J. Yuan, and C.B. Thompson. Bax and Bak can localize to the endoplasmic reticulum to initiate apoptosis. *The Journal of Cell Biology*, 162(1):59–69, 2003.

[162] B. Schafer, J. Quispe, V. Choudhary, J.E. Chipuk, T.G. Ajero, H. Du, R. Schneiter, and T. Kuwana. Mitochondrial outer membrane proteins assist Bid in Bax-mediated lipidic pore formation. *Molecular Biology of the Cell*, 20(8):2276–2285, 2009.

[163] J.A. Brannigan and A.J. Wilkinson. Protein engineering 20 years on. *Nature Reviews Molecular Cell Biology*, 3(12):964–970, 2002.

[164] J.F. Wang and K.C. Chou. Insights from studying the mutation-induced allostery in the M2 proton channel by molecular dynamics. *Protein Engineering Design and Selection*, 23(8):663–666, 2010.

Index

AMBER, 19
apoptosis, 2

Bcl-2, 5
Bcl-2 family, 5
BH3 Interacting domain Death (Bid), 72

Clustal (Cluster alignment), 17

degree of colectivity, 26

European Bioinformatics Institute (EBI), 17

folding, 14
force field, 19

GROMACS, 18

Human Bax, 7

LINear Constraint Solver (LINCS), 22

MD, 13
mitochondrial outer membrane, 3
mitochondrial pathway, 3
molecular dynamics, 13
MOM, 3

necrosis, 3
Normal Mode Analysis (NMA), 24

Optimized Potentials for Liquid Simulations for All Atoms (OPLS-AA), 19

Particle Mesh Ewald (PME), 20
Protein Data Bank (PDB), 22
Protein Information Resource (PIR), 17

Rg values, 36
RMSD values, 36
RMSF values, 40
rotation-translation-block (RTB), 26

SASA values, 36
Simple Point Charge (SPC), 20
SPC Water, 20
Swiss Institute of Bioinformatics (SIB), 17

thermal unfolding of proteins, 15
thermal unfolding simulations, 15

unfolding, 14
Universal Protein resource (UniProt), 17