

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Prediction of localization and interactions of apoptotic proteins

Journal of Biomedical Science 2009, **16**:59 doi:10.1186/1423-0127-16-59

Miroslav Varecha (mvara@fi.muni.cz)
Michal Zimmermann (63720@mail.muni.cz)
Jana Amrichova (amrich@fi.muni.cz)
Vladimir Ulman (xulman@fi.muni.cz)
Pavel Matula (pam@fi.muni.cz)
Michal Kozubek (kozubek@fi.muni.cz)

ISSN 1423-0127

Article type Research

Submission date 24 March 2009

Acceptance date 6 July 2009

Publication date 6 July 2009

Article URL <http://www.jbiomedsci.com/content/16/1/59>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *Journal of Biomedical Science* are listed in PubMed and archived at PubMed Central.

For information about publishing your research in *Journal of Biomedical Science* or any BioMed Central journal, go to

<http://www.jbiomedsci.com/info/instructions/>

For information about other BioMed Central publications go to

<http://www.biomedcentral.com/>

© 2009 Varecha *et al.*, licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prediction of localization and interactions of apoptotic proteins

Miroslav Vařecha^{*}, Michal Zimmermann, Jana Amrichová, Vladimír Ulman, Pavel Matula, and
Michal Kozubek

Centre for Biomedical Image Analysis, Faculty of Informatics, Masaryk University, Botanická 68a,
Brno 602 00, Czech Republic

^{*}Corresponding author

Email addresses:

Miroslav Vařecha (mvara@fi.muni.cz), Michal Zimmermann (63720@mail.muni.cz), Jana
Amrichová (amrich@fi.muni.cz), Vladimír Ulman (xulman@fi.muni.cz), Pavel Matula
(pam@fi.muni.cz), and Michal Kozubek (kozubek@fi.muni.cz).

Abstract

During apoptosis several mitochondrial proteins are released. Some of them participate in caspase-independent nuclear DNA degradation, especially apoptosis-inducing factor (AIF) and endonuclease G (endoG). Another interesting protein, which was expected to act similarly as AIF due to the high sequence homology with AIF is AIF-homologous mitochondrion-associated inducer of death (AMID). We studied the structure, cellular localization, and interactions of several proteins *in silico* and also in cells using fluorescent microscopy. We found the AMID protein to be cytoplasmic, most probably incorporated into the cytoplasmic side of the lipid membranes. Bioinformatic predictions were conducted to analyze the interactions of the studied proteins with each other and with other possible partners. We conducted molecular modeling of proteins with unknown 3D structures. These models were then refined by MolProbity server and employed in molecular docking simulations of interactions. Our results show data acquired using a combination of modern *in silico* methods and image analysis to understand the localization, interactions and functions of proteins AMID, AIF, endonuclease G, and other apoptosis-related proteins.

Background

During some forms of apoptosis the mitochondrial outer membrane becomes depolarized and partially permeable to proteins. This results in a massive nonspecific release of hydrophilic proteins from the intermembrane space into the cytoplasm [1]. Among these proteins are apoptosis-inducing factor (AIF) and endonuclease G (endoG). The release of these proteins results in activation of the apoptotic caspases, degradation of nuclear DNA, and cell death [2][3]. However, both AIF and endoG have been found to directly participate in DNA degradation in a caspase-independent way [4]. The protein AIF-homologous mitochondrion-associated inducer of death (AMID), which is probably not located in the mitochondrion, shares sequence homology with AIF and exerts similar apoptotic effects on nuclear chromatin [5]. Interestingly, endoG, AIF and AMID have all been

found to influence chromatin changes during apoptosis [6].

EndoG is a mitochondrial nuclease with a molecular weight of 30 kDa. Its N-terminus contains a mitochondrial localization sequence (MLS), which is cleaved upon successful transport of the endoG precursor polypeptide across the outer mitochondrial membrane. EndoG migrates from mitochondria into the nucleus after apoptogenic stimuli [7][8]. Addition of endoG to isolated cell nuclei resulted in cleavage of the chromatin into large fragments (~50 kbp) and subsequently into inter- and intra-nucleosomal-size fragments with periodically repeated single-stranded breaks. The first phase of endoG activity equates with the large-scale degradation of DNA during apoptosis, but the second phase would not seem to be able to generate the characteristic “laddered“ fragmentation of chromatin observed in apoptotic nuclei. This may suggest that endoG normally interacts with other nucleases. Indeed, cooperation between endoG, DNase I and exonuclease III has been shown to occur only on isolated dsDNA [6]. Another proposed interaction partner for endoG was found by protein analytic *in vitro* methods to be flap endonuclease 1 [9], but it was not yet shown in living or fixed cells as many other possible interactions mentioned here.

AIF (also known as AIFM1 or PDCD8) is an evolutionary conserved flavoprotein. It shares a high degree of sequence homology with bacterial, plant, and fungal oxidoreductases. The human AIF is expressed as a precursor polypeptide of molecular weight ~67 kDa. This precursor contains an N-terminal MLS, which is cleaved, and the active AIF (~57 kDa) is created in the mitochondrial intermembrane space [10]. AIF is probably bound by its N-terminus to the surface of the inner mitochondrial membrane [11][12]. The function of AIF in the mitochondrion under non-apoptotic conditions is not clear, but there is evidence that AIF may serve to sequester free radicals and that it can play important role in oxidative phosphorylation [13][14]. However, human AIF is also able to induce apoptosis [10]. None of these effects could be inhibited by the pan-caspase inhibitor z-VAD-

fmk, thus they are caspase-independent [15]. Translocation of AIF into the nucleus occurs during apoptosis [15][16]. The *C. elegans* homologue of AIF, known as WAH-1, was shown to induce apoptosis and to migrate from mitochondria to the nucleus, where it interacts with a nuclease, CPS-6 (a homologue of the mammalian endoG), and together they mediate chromatin DNA degradation [17]. There is no clear evidence as to how mammalian AIF is involved in the process of chromatin degradation, but AIF can physically interact with DNA and RNA [18]. Interaction of human AIF and endoG, in analogy to what happens in *C. elegans*, has not yet been shown, although protein analysis *in vitro* results suggest its possibility [9]. However, other important proteins have been proposed to interact with AIF, namely cyclophilin A [19] and heat shock 70 kDa protein 1A (HSP70-1) [20]. Cyclophilin A is involved in transport of AIF into the nucleus and may be also involved in chromatin degradation [19][21]. HSP70-1 binds AIF in the cytoplasm and blocks its transport into the nucleus [22]. Another protein, DNA topoisomerase II α , was found to be involved in chromatin degradation during caspase-independent apoptosis and is therefore one of our candidate proteins that may interact with AIF in nucleus [23].

AMID (also known as AIFM2 or PRG3) is a flavoprotein with amino acid sequence similarity to NADH oxidoreductases in all species, as well as to AIF (22% identity). In contrast to AIF, AMID does not contain a recognizable MLS [5]. AMID was found to associate with the outer mitochondrial membrane [5] or to be freely distributed in the cytoplasm [24]. Apoptosis induced by AMID is regulated by p53, is caspase-independent, and is not affected by overexpression of the Bcl-2 protein [5]. Artificially induced overexpression of AMID caused chromatin condensation at the nuclear periphery and formation of apoptotic bodies [24]. However, this was not confirmed by another study [5]. Overexpression of AMID also caused a loss of structurally differentiated mitochondria [5]. However, recent results did not show that AMID acts similarly to AIF and its localization was also questioned [25][26].

Therefore, in this work we use modern *in silico* methods for sequence analysis, predicting the subcellular localization of proteins, prediction of interactions, molecular modeling and docking and we combined these methods with fluorescence microscopic imaging of endoG, AIF, AMID, and other apoptotic proteins in living or fixed cells to analyze their localization and interactions.

Materials and methods

Cell culture

Human bone osteosarcoma U-2 OS cells were grown in minimal essential medium containing Earle's balanced salts, L-glutamine, non-essential amino acids, 1.5 g/l NaHCO₃, 10% fetal bovine serum, 100 U/ml of penicillin and 100 µg/ml of streptomycin at 37°C in a 5% CO₂ atmosphere.

Plasmid DNA preparation and transfection

The plasmid constructs carrying the genes for endoG-EYFP, AIF-tHcRed, and AMID-tHcRed fluorescent fusion proteins were described previously [25]. Cells were transfected using FuGENE 6 (Roche). In cases of transient transfection, the transfected cells were used for experiments 24 hours after transfection. Stably transfected cell cultures were selected using 600 µg/ml of geneticin (G418) for 14 days.

Prediction of protein's cellular localization

Prediction analysis was conducted using amino acid sequences of the endonuclease G protein precursor (GenBank Accession Number NP_004426), isoform 1 of protein AIF (NP_004199), AMID (NP_116186), HSP70-1 protein (NP_005336), cyclophilin A (NP_066953), DNA topoisomerase II α (NP_001058). Four internet server tools specialized in predicting cellular localization of proteins from their amino acid sequences were employed: PSORT II [27][28], WoLF

PSORT [29], MultiLoc [30], and CELLO v.2.5 [31].

Fluorescence microscopy

Image data acquisition was done using the high-resolution cytometer (HRCM) developed in our laboratory [25][32][33]. The HRCM system was controlled by FISH 2.0 software or by new Acquarium 1.0 software, both developed in our laboratory [32][33][34]. Immunostaining was done using fluorescently labeled primary antibodies labeled using kits with Alexa Fluor dyes 488 and 568 (Invitrogen - Molecular Probes). Rabbit polyclonal anti-human AIF and endoG (Abcam) and mouse monoclonal anti-human cyclophilin A and DNA topoisomerase II α (Tebu-bio) primary antibodies were used. Cells were usually grown in Lab-Tek chambers (Nunc) or grown on coverglass in petri dish. Samples were fixed using 4% paraformaldehyde. Obtained image data were deconvolved using the plugins Diffraction PSF 3D and Iterative Deconvolve 3D for ImageJ software [35].

Molecular modeling

Prediction servers can use the known 3D protein structure with amino acid sequence homologous to the protein of interest and predict its 3D structure based on this homology. We used servers Phyre [36][37] and M4T [38] for such 3D structure modeling after intensive testing of various servers.

Visualization of 3D protein structures was done using software UCSF Chimera [39]. The server tool MolProbity [40] was used for protein model validation and refinement.

Prediction of interactions

Interaction prediction analysis was conducted using two web server tools for predicting interaction sites in proteins from their 3D structure. Cons-PPISP [41] for prediction of protein-protein interaction residues and DISPLAR [42] for prediction of DNA binding residues.

Molecular docking

We accomplished the molecular docking using one of the most modern server docking tools PatchDock with refinement tool FireDock [43][44]. For molecular docking were used optimized PDB files of 3D structures from RCSB Protein Data Bank (for AIF it was PDB ID 1m6i, for B-DNA it was PDB ID 1bna, and for cyclophilin A it was PDB ID 1W8M) or our prepared models. The 10 best solutions were saved and submitted to following analysis and refinement by FireDock tool. Again the best 10 solutions and calculations were also saved and then studied.

Results

Sequence analysis and prediction of subcellular locations of proteins

We analyzed the amino acid sequences of AIF and AMID by BLAST server using tool Search Conserved Domains and using tool Blast 2 sequences [45] to study the alignment and the conserved domains of these two proteins (Fig. 1A). Results clearly show that sequence homology of these two proteins is restricted to Ndh (NADH dehydrogenase, FAD-containing subunit) conserved domain. Predictions of the subcellular locations of endoG, AIF, AMID, cyclophilin A, HSP70-1, and DNA topoisomerase II α were made using four prediction servers: PSORT II, WoLF PSORT, MultiLoc, and CELLO. Although most of the analyzed proteins have known cellular location, we nevertheless employed these server tools to validate the accuracy of a final prediction summary based on the combined results of these tools (Table 1). These results clearly indicate that the predicted cellular location of endoG and AIF is in the mitochondrion, cyclophilin A and HSP70-1 in the cytoplasm, and DNA topoisomerase II α in the nucleus. For AMID, the most probable subcellular location was predicted to be the cytoplasm (Table 1). The individual PSORT II algorithms revealed the MLS in endoG sequence to be the N-terminal 48-amino-acid cleavable signal peptide [46][47] and the nuclear localization signal (NLS) at position 24 [48] inside the predicted MLS (Fig. 1B). In the

sequence of AIF isoform 1 we predicted the MLS to be the N-terminal 61-amino-acid cleavable signal peptide [46][47], the NLS at positions 106 – 112 [48], and one transmembrane segment between positions 68 – 84 (Fig. 1B) [27]. The PSORT II algorithms discovered an N-myristoylation-allowing motif [49] which would potentially permit incorporation of AMID into various cellular membranes [50] and one transmembrane segment between positions 11 – 33 by TMHMM 2.0 server (Fig. 1B) [51]. N-myristoylation-allowing motif was also detected by two other bioinformatic tools - the Myristoylator [52] and the NMT Predictor [53], which defined the motif to be the amino acid sequence GSQVSVESGALHVIVG starting at position 2. In the sequence of HSP70-1 were detected NLS at positions 246 – 273 and 594 – 597. PSORT II and TMHMM 2.0 found nothing of interest for cyclophilin A. In the sequence of DNA topoisomerase II α were predicted several NLS between amino acids 632 to 1468.

Experimental cellular locations of proteins

Experimental determination of the cellular locations of studied proteins was conducted either by transfecting living cells with mammalian expression vectors encoding the fusion proteins or by immunostaining of fixed cells by fluorescently labeled primary antibody. Figure 2A shows a typical signal distributions of endoG-EYFP and AMID-tHcRed fluorescence in transfected human living U-2 OS cells. Apparently, AMID and endoG do not colocalize significantly. EndoG is present in mitochondria and AMID is present throughout the cytoplasm apparently on various structures [25]. Figures 2B and 2C show signal distribution of AMID-tHcRed fluorescence in living U-2 OS cell before and 6 hours after induction of apoptosis by 200 nM staurosporine. Figure 2C clearly shows that AMID does not translocate to nucleus. Figure 2D shows the fluorescence signal of endoG-EYFP in one living cell over-expressing endoG which distributed inside nucleus, although the cell is viable and non-apoptotic. Translocation of endoG into the cell nucleus during staurosporine-induced apoptosis is shown in Figures 2E and 2F. We co-immunostained AIF and cyclophilin A in

fixed U-2 OS cells (Fig. 2G). AIF is located to mitochondria and cyclophilin A to the cytoplasm and also to the cell nucleus. Apoptosis induced in these cells resulted into nuclear translocation of AIF (Fig. 2H) and cyclophilin A remained in nucleus and cytoplasm (Fig. 2I). Immunostaining by fluorescent antibodies against DNA topoisomerase II α revealed the expected nuclear localization of the protein (Fig. 2J). Interestingly, after apoptosis induction by 200 nM staurosporine, DNA topoisomerase II α in nucleus displayed distribution similar to chromatin condensation attributed to effect of AIF during apoptosis (Fig. 2K and 2L). HSP70-1 immunostaining showed not only cytoplasmic localization of the protein (Fig. 2N) but also strong nucleoli localization of HSP70-1 after heat shock at 42°C for 2 hours (Fig. 2M). Interestingly, after apoptosis induction by 200 nM staurosporine, HSP70-1 relocated to the cell nucleus (Fig. 2O).

Molecular modeling

We prepared 3D models of protein structure of several proteins using various modeling servers and we selected only the best models ranked by MolProbity [40] server. MolProbity server also refined the structure of models and these refined models were used in our following predictions and molecular dockings. Best endoG model (prepared from 1G8T structure) was prepared by Phyre server [36] (MolProbity score 2.96) and consists of amino acids 65-296 from 297 amino acids (Fig. 3). We also prepared model of HSP70-1 protein (MolProbity score 2.68) by M4T server [38] for amino acids range of 1-554 from 641 amino acids (Fig. 3). Another two models were prepared again by Phyre server for proteins WAH-1 (amino acids 238-700 of 700; MolProbity score 3.02) and CPS-6 (amino acids 57-305 of 308; MolProbity score 3.06) (Fig. 3).

Prediction of interactions

Using servers cons-PPISP [41] and DISPLAR [42] we predicted the possible binding sites for proteins (cons-PPISP) and DNA (DISPLAR) in the sequences of AIF, endonuclease G, and cyclophilin A. These predictions are based on 3D structures of these proteins. Using cons-PPISP

server we found several residues that can form binding sites in the AIF amino acid sequence. The binding residues are in the interval of amino acids 264-498 and at the C-terminus (amino acids 598-608). For endoG we predicted two possible binding regions for amino acids 75-96 and 265-282 (C-terminus). We found a high concentration of possible binding residues at amino acids 88-97 and 116-138 for cyclophilin A. Using DISPLAR server, we were able to detect possible locations for DNA binding residues in the sequence of AIF especially at interval of amino acids 237-254. EndoG sequence showed three intervals containing predicted residues at positions 100-115, 139-153, and 177-188. For cyclophilin A, we detected cluster of predicted residues for DNA binding in the interval of amino acids 54-72.

Molecular docking

Docking represents the mathematical calculation of the most probable spatial orientation of two interacting molecules, usually a protein and a small ligand, two interacting proteins, or DNA and protein. Various parameters are calculated to evaluate possibility of such interaction. For molecular docking we used new server PatchDock [43] with refinement tool FireDock [44]. We modeled the proposed interaction of AIF with B-DNA [18] using these tools and best docking model is shown (Fig. 4A). Global energy function of this complex was calculated by FireDock server to be -16.84 of relative units (this value is considered to be related to free binding energy and higher negative value means higher free binding energy and thus higher interaction probability). Another experimentally shown interaction pair we studied was AIF and cyclophilin A (Fig. 4B) [19]. Global energy value of this complex was -8.41. Next we modeled the possible interaction of AIF and endoG (Fig. 4C) resulting in global energy value -26.27 for this complex. Lastly, we modeled experimentally proved interaction of analogues of AIF and endoG, WAH-1 and CPS-6 (Fig. 4D) [17]. Resulting global energy was -13.11. Figure 5 shows parts of modeled docking complexes described above with focus on visualization of binding sites. In green color are shown residues

predicted by cons-PPISP or DISPLAR servers mentioned above that can form binding sites for protein-protein and protein-DNA interactions.

Discussion

We clearly show that sequence homology of AIF and AMID is restricted to Ndh conserved domain, that corresponds to oxidoreductase function and not to apoptotic function of AIF, which was shown to reside in large C-terminal part of AIF sequence, which is totally missing in AMID sequence (Fig. 1A) [54]. AMID sequence contains, aside from Ndh conserved domain and N-terminal part, only very small C-terminal part consisting of only several residues and thus it is highly improbable that AMID will act similarly as AIF during apoptosis. We correctly predicted that endoG and AIF would localize to the mitochondria, HSP70-1 and cyclophilin A to the cytoplasm, and DNA topoisomerase II α to the nucleus (Table 1). The cytoplasm was found as the most probable cellular location for AMID (Table 1). This result is in agreement with previous observations [25], although some of them described the distribution of AMID in the cytoplasm differently [5][24]. No recognizable localization signal was found in the AMID amino acid sequence. However, the MLS was detected as the first 48 amino acids of the endoG sequence and as the first 61 amino acids of the AIF sequence [55]. This finding corroborates the many published observations that these proteins are translocated into the nucleus during apoptosis [7][15][56]. Interestingly, the predicted NLS of endoG is located within the MLS, so that once endoG enters the mitochondrion the NLS will be cleaved together with the MLS. NLS was found in sequences of DNA topoisomerase II α , which was expected, due to function of this protein. Interestingly, NLS was also found in the sequence of HSP70-1 suggesting its possible role in nucleus. HSP70-1 was found in nucleoli after the heat shock [57]. The algorithms and TMHMM 2.0 server also predicted that the sequence of endoG, cyclophilin A, HSP70-1 and DNA topoisomerase II α do not contain any transmembrane regions. However, PSORTII algorithms and TMHMM 2.0 revealed that AIF and AMID sequences

contain one predicted transmembrane region thus suggesting that AIF and AMID could be membrane proteins [11][58]. When apoptosis is induced the AIF protein is spliced probably by calpain I and loses its N-terminal part that contains the predicted transmembrane region and AIF is thus released into the intermembrane space and out of mitochondria [59]. Sequence of cyclophilin A showed no recognizable regions and is therefore interesting, that this protein is often found in nucleus, which may support the possibility that cyclophilin A interacts with AIF which contains NLS and together they translocate to nucleus [19]. The lipid anchoring is also possible for AMID, mediated by its identified putative N-myristoylation site. This site allows AMID to incorporate into various cellular membranes. Our findings thus predict that AMID is by N-terminal part incorporated into cellular membranes from cytoplasmic side.

Fluorescence imaging of living cells and microscopy of immunostained fixed cells (Fig. 2) confirmed the predicted localization of studied proteins (Table 1). AIF and endoG localized to mitochondria (Fig. 2A, E, G) and DNA topoisomerase II α to nucleus (Fig. 2J). DNA topoisomerase II α was also found to relocate during apoptosis similarly to chromatin condensation organized by AIF (Fig. 2K and 2L). This can point to possible interaction of DNA topoisomerase II α with AIF or it simply binds to condensed chromatin during apoptosis. Apart from cytoplasmic distribution, cyclophilin A surprisingly showed strong nuclear staining (Fig. 2G) maintained even during apoptosis (Fig. 2I), although its sequence does not contain any recognizable signals or motifs. Cyclophilin A can probably bind other protein that can translocate to nucleus even at non-apoptotic conditions. Immunostaining of HSP70-1 after strong heat shock showed significant nuclear localization of HSP70-1 into nucleoli (Fig. 2M) as was already observed [57]. Localization of HSP70-1 in non-apoptotic conditions was found to be cytoplasmic as predicted (Fig. 2N), but surprisingly during apoptosis HSP70-1 translocated into the cell nucleus (Fig. 2O), which even more strongly supports the possible role of this protein in apoptosis and it also corresponds to our

prediction results. Cells of normal morphology expressing high levels of endoG-EYFP were identified in stably transfected clones. The signal comprised not only the mitochondrial fluorescence but also a strong confluent fluorescence in the cytoplasm and even in the nuclear chromatin (Fig. 2D). Such cells showed no morphological apoptotic changes. Thus the presence of endoG in the cytoplasm and nucleus was not sufficient to induce apoptotic chromatin degradation in the cells. This confirms our bioinformatic predictions about endoG, that the NLS of endoG can transport endoG-EYFP into the nucleus when the protein is highly overexpressed in the cytoplasm. These findings present the experimental evidence that the mere presence of endoG in the nucleus is not sufficient to initiate DNA cleavage and, although it is a nuclease, endoG apparently needs to be activated to degrade DNA in living cells [6][60]. The fluorescence signal of AMID-tHcRed was distributed throughout the cytoplasm, although not diffuse as was suggested [24]. AMID was found to localize to unidentified regions throughout the entire cell (Fig. 2A and 2B). AMID-tHcRed localized close to the nuclear membrane, possibly with the Golgi apparatus and/or the endoplasmic reticulum; it was also associated with small vesicles in the cytoplasm and possibly with the plasma membrane and it does not translocate to the nucleus during apoptosis (Fig. 2C) [25][26]. Cells overexpressing AMID-tHcRed were viable with normal morphology, and overexpression of AMID was not sufficient to induce apoptosis; thus, our results challenge a previous report that overexpression of AMID could induce apoptotic degradation of chromatin and the loss of structurally differentiated mitochondria [5].

We prepared 3D structural models for proteins without known 3D structure (Fig. 3). Both, these refined models and known protein structures, were used for *in silico* prediction of locations of binding residues for DNA and proteins. We studied especially molecules that were shown to interact with AIF or are possible candidates for such interaction. Especially we studied the interaction of AIF with B-DNA (Fig. 4A) and cyclophilin A (Fig. 4B) using molecular modeling.

We also studied the hypothetical possible interaction of AIF with endoG (Fig. 4C), based on known interaction of their analogues WAH-1 and CPS-6 from *Caenorhabditis elegans* (Fig. 4D) [17]. The resulting values produced by FireDock server tool are suggesting that interaction of AIF with endoG is very probable due to the high negative value of global energy function, comparable to other known interactions. We visualized the predicted binding residues from servers cons-PPISP and DISPLAR in our molecular docking complexes (Fig. 5). This novel visualization clearly shows, that molecular docking results and binding residues predictions corresponds very well, thus supporting results of each other and verifying the binding residues locations and molecular binding spatial configuration of the studied proteins.

Our results show great possibilities of in silico methods and image analysis to understand the localization, interactions and functions of proteins. Our results present new data about the structure, localization, functions and interactions of proteins AMID, AIF, endonuclease G, and other apoptosis-related proteins.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MV carried out all bioinformatic studies, live-cell fluorescence microscopy, and drafted the manuscript. MZ carried out fluorescence microscopy of fixed samples. JA carried out design of some experiments and preparation of samples. VU, PM, and MK constructed the automated imaging platform for fluorescence microscopy and developed the imaging and analysis software controlling this platform.

Acknowledgments

This work was supported by The Ministry of Education, Youth and Sports of the Czech Republic (projects number MSM0021622419, 2B06052, and LC535).

References

1. Patterson SD, Spahr CS, Daugas E, Susin SA, Irinopoulou T, Koehler C, Kroemer G: **Mass spectrometric identification of proteins released from mitochondria undergoing permeability transition.** *Cell Death Differ* 2000, **7**: 137-144.
2. Susin SA, Zamzami N, Castedo M, Hirsch T, Marchetti P, Macho A, Daugas E, Geuskens M, Kroemer G: **Bcl-2 inhibits the mitochondrial release of an apoptogenic protease.** *J Exp Med* 1996, **184**: 1331-1341.
3. Zamzami N, Susin SA, Marchetti P, Hirsch T, Gómez-Monterrey I, Castedo M, Kroemer G: **Mitochondrial control of nuclear apoptosis.** *J Exp Med* 1996, **183**: 1533-1544.
4. Li LY, Luo X, Wang X: **Endonuclease G is an apoptotic DNase when released from mitochondria.** *Nature* 2001, **412**: 95-99.
5. Wu M, Xu LG, Li X, Zhai Z, Shu HB: **AMID, an apoptosis-inducing factor-homologous mitochondrion-associated protein, induces caspase-independent apoptosis.** *J Biol Chem* 2002, **277**: 25617-25623.
6. Widlak P, Li LY, Wang X, Garrard WT: **Action of recombinant human apoptotic endonuclease G on naked DNA and chromatin substrates: cooperation with exonuclease and DNase I.** *J Biol Chem* 2001, **276**: 48404-48409.
7. Hahn HP, Pang M, He J, Hernandez JD, Yang RY, Li LY, Wang X, Liu FT, Baum LG: **Galectin-1 induces nuclear translocation of endonuclease G in caspase- and cytochrome c-independent T cell death.** *Cell Death Differ* 2004, **11**: 1277-1286.
8. Takano J, Tomioka M, Tsubuki S, Higuchi M, Iwata N, Itohara S, Maki M, Saido TC: **Calpain**

mediates excitotoxic DNA fragmentation via mitochondrial pathways in adult brains: evidence from calpastatin mutant mice. *J Biol Chem* 2005, **280**: 16175-16184.

9. Kalinowska M, Garncarz W, Pietrowska M, Garrard WT, Widlak P: **Regulation of the human apoptotic DNase/RNase endonuclease G: involvement of Hsp70 and ATP.** *Apoptosis* 2005, **10**: 821-830.

10. Susin SA, Lorenzo HK, Zamzami N, Marzo I, Snow BE, Brothers GM, Mangion J, Jacotot E, Costantini P, Loeffler M, Larochette N, Goodlett DR, Aebersold R, Siderovski DP, Penninger JM, Kroemer G: **Molecular characterization of mitochondrial apoptosis-inducing factor.** *Nature* 1999, **397**: 441-446.

11. Uren RT, Dewson G, Bonzon C, Lithgow T, Newmeyer DD, Kluck RM: **Mitochondrial release of pro-apoptotic proteins: electrostatic interactions can hold cytochrome c but not Smac/DIABLO to mitochondrial membranes.** *J Biol Chem* 2005, **280**: 2266-2274.

12. Otera H, Ohsakaya S, Nagaura ZI, Ishihara N, Mihara K: **Export of mitochondrial AIF in response to proapoptotic stimuli depends on processing at the intermembrane space.** *EMBO J* 2005, **24**: 1375-1386.

13. Klein JA, Longo-Guess CM, Rossmann MP, Seburn KL, Hurd RE, Frankel WN, Bronson RT, Ackerman SL: **The harlequin mouse mutation downregulates apoptosis-inducing factor.** *Nature* 2002, **419**: 367-374.

14. Wilkinson JC, Wilkinson AS, Galbán S, Csomos RA, Duckett CS: **Apoptosis-inducing factor is a target for ubiquitination through interaction with XIAP.** *Mol Cell Biol* 2008, **28**: 237-247.

15. Daugas E, Susin SA, Zamzami N, Ferri KF, Irinopoulou T, Larochette N, Prévost MC, Leber B, Andrews D, Penninger J, Kroemer G: **Mitochondrio-nuclear translocation of AIF in apoptosis and necrosis.** *FASEB J* 2000, **14**: 729-739.

16. Fonfría E, Daré E, Benelli M, Suñol C, Ceccatelli S: **Translocation of apoptosis-inducing factor in cerebellar granule cells exposed to neurotoxic agents inducing oxidative stress.** *Eur J Neurosci* 2002, **16**: 2013-2016.

17. Wang X, Yang C, Chai J, Shi Y, Xue D: **Mechanisms of AIF-mediated apoptotic DNA degradation in *Caenorhabditis elegans*.** *Science* 2002, **298**: 1587-1592.

18. Vahsen N, Candé C, Dupaigne P, Giordanetto F, Kroemer RT, Herker E, Scholz S, Modjtahedi N, Madeo F, Le Cam E, Kroemer G: **Physical interaction of apoptosis-inducing factor with DNA and RNA.** *Oncogene* 2006, **25**: 1763-1774.
19. Candé C, Vahsen N, Kouranti I, Schmitt E, Daugas E, Spahr C, Luban J, Kroemer RT, Giordanetto F, Garrido C, Penninger JM, Kroemer G: **AIF and cyclophilin A cooperate in apoptosis-associated chromatinolysis.** *Oncogene* 2004, **23**: 1514-1521.
20. Ravagnan L, Gurbuxani S, Susin SA, Maise C, Daugas E, Zamzami N, Mak T, Jäättelä M, Penninger JM, Garrido C, Kroemer G: **Heat-shock protein 70 antagonizes apoptosis-inducing factor.** *Nat Cell Biol* 2001, **3**: 839-843.
21. Zhu C, Wang X, Deinum J, Huang Z, Gao J, Modjtahedi N, Neagu MR, Nilsson M, Eriksson PS, Hagberg H, Luban J, Kroemer G, Blomgren K: **Cyclophilin A participates in the nuclear translocation of apoptosis-inducing factor in neurons after cerebral hypoxia-ischemia.** *J Exp Med* 2007, **204**: 1741-1748.
22. Lui JC, Kong SK: **Heat shock protein 70 inhibits the nuclear import of apoptosis-inducing factor to avoid DNA fragmentation in TF-1 cells during erythropoiesis.** *FEBS Lett* 2007, **581**: 109-117.
23. Solovyan VT: **Characterization of apoptotic pathway associated with caspase-independent excision of DNA loop domains.** *Exp Cell Res* 2007, **313**: 1347-1360.
24. Ohiro Y, Garkavtsev I, Kobayashi S, Sreekumar KR, Nantz R, Higashikubo BT, Duffy SL, Higashikubo R, Usheva A, Gius D, Kley N, Horikoshi N: **A novel p53-inducible apoptogenic gene, PRG3, encodes a homologue of the apoptosis-inducing factor (AIF).** *FEBS Lett* 2002, **524**: 163-171.
25. Vařecha M, Amrichová J, Zimmermann M, Ulman V, Lukášová E, Kozubek M: **Bioinformatic and image analyses of the cellular localization of the apoptotic proteins endonuclease G, AIF, and AMID during apoptosis in human cells.** *Apoptosis* 2007, **12**: 1155-1171.
26. Bilyy R, Kit Y, Hellman U, Stoika R: **AMID: new insights on its intracellular localization and expression at apoptosis.** *Apoptosis* 2008, **13**: 729-732.
27. Nakai K, Kanehisa M: **A knowledge base for predicting protein localization sites in**

eukaryotic cells. *Genomics* 1992, **14**: 897-911.

28. Nakai K, Horton P: **PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization.** *Trends Biochem Sci* 1999, **24**: 34-36.

29. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K: **WoLF PSORT: protein localization predictor.** *Nucleic Acids Res* 2007, **35**: W585-W587.

30. Höglund A, Dönnés P, Blum T, Adolph HW, Kohlbacher O: **MultiLoc: prediction of protein subcellular localization using N-terminal targeting sequences, sequence motifs and amino acid composition.** *Bioinformatics* 2006, **22**: 1158-1165.

31. Yu CS, Chen YC, Lu CH, Hwang JK: **Prediction of protein subcellular localization.** *Proteins* 2006, **64**: 643-651.

32. Kozubek M, Kozubek S, Lukášová E, Marečková A, Bártová E, Skalníková M, Jergová A: **High-resolution cytometry of FISH dots in interphase cell nuclei.** *Cytometry* 1999, **36**: 279-293.

33. Kozubek M, Kozubek S, Lukášová E, Bártová E, Skalníková M, Matula Pe, Matula Pa, Jirsová P, Cafourková A, Koutná I: **Combined confocal and wide-field high-resolution cytometry of fluorescent in situ hybridization-stained cells..** *Cytometry* 2001, **45**: 1-12.

34. Kozubek M, Matula Pe, Matula Pa, Kozubek S: **Automated acquisition and processing of multidimensional image data in confocal in vivo microscopy.** *Microsc Res Tech* 2004, **64**: 164-175.

35. Dougherty R: **Extensions of DAMAS and benefits and limitations of deconvolution in beamforming.** *AIAA* 2005: 2961.

36. Bennett-Lovsey RM, Herbert AD, Sternberg MJ, Kelley LA: **Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre.** *Proteins* 2008, **70**: 611-625.

37. Kelley LA, MacCallum RM, Sternberg MJ: **Enhanced genome annotation using structural profiles in the program 3D-PSSM.** *J Mol Biol* 2000, **299**: 499-520.

38. Fernandez-Fuentes N, Madrid-Aliste CJ, Rai BK, Fajardo JE, Fiser A: **M4T: a comparative protein structure modeling server.** *Nucleic Acids Res* 2007, **35**: W363-W368.

39. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF Chimera--a visualization system for exploratory research and analysis.** *J Comput Chem* 2004, **25**: 1605-1612.
40. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC: **MolProbity: all-atom contacts and structure validation for proteins and nucleic acids.** *Nucleic Acids Res* 2007, **35**: W375-W383.
41. Chen H, Zhou HX: **Prediction of interface residues in protein-protein complexes by a consensus neural network method: test against NMR data.** *Proteins* 2005, **61**: 21-35.
42. Tjong H, Zhou HX: **DISPLAR: an accurate method for predicting DNA-binding sites on protein surfaces.** *Nucleic Acids Res* 2007, **35**: 1465-1477.
43. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ: **PatchDock and SymmDock: servers for rigid and symmetric docking.** *Nucleic Acids Res* 2005, **33**: W363-W367.
44. Andrusier N, Nussinov R, Wolfson HJ: **FireDock: fast interaction refinement in molecular docking.** *Proteins* 2007, **69**: 139-159.
45. Tatusova TA, Madden TL: **BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences.** *FEMS Microbiol Lett* 1999, **174**: 247-250.
46. McGeoch DJ: **On the predictive recognition of signal peptide sequences.** *Virus Res* 1985, **3**: 271-286.
47. von Heijne G: **A new method for predicting signal sequence cleavage sites.** *Nucleic Acids Res* 1986, **14**: 4683-4690.
48. Hicks GR, Raikhel NV: **Protein import into the nucleus: an integrated view.** *Annu Rev Cell Dev Biol* 1995, **11**: 155-188.
49. Johnson DR, Bhatnagar RS, Knoll LJ, Gordon JI: **Genetic and biochemical studies of protein N-myristoylation.** *Annu Rev Biochem* 1994, **63**: 869-914.
50. Borgese N, Aggujaro D, Carrera P, Pietrini G, Bassetti M: **A role for N-myristoylation in protein targeting: NADH-cytochrome b5 reductase requires myristic acid for association with outer mitochondrial but not ER membranes.** *J Cell Biol* 1996, **135**: 1501-1513.
51. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein**

- topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305**: 567-580.
52. Bologna G, Yvon C, Duvaud S, Veuthey AL: **N-Terminal myristoylation predictions by ensembles of neural networks.** *Proteomics* 2004, **4**: 1626-1632.
53. Maurer-Stroh S, Eisenhaber B, Eisenhaber F: **N-terminal N-myristoylation of proteins: refinement of the sequence motif and its taxon-specific differences.** *J Mol Biol* 2002, **317**: 523-540.
54. Gurbuxani S, Schmitt E, Cande C, Parcellier A, Hammann A, Daugas E, Kouranti I, Spahr C, Pance A, Kroemer G, Garrido C: **Heat shock protein 70 binding inhibits the nuclear import of apoptosis-inducing factor.** *Oncogene* 2003, **22**: 6669-6678.
55. Lorenzo HK, Susin SA: **Mitochondrial effectors in caspase-independent cell death.** *FEBS Lett* 2004, **557**: 14-20.
56. Zanna C, Ghelli A, Porcelli AM, Martinuzzi A, Carelli V, Rugolo M: **Caspase-independent death of Leber's hereditary optic neuropathy cybrids is driven by energetic failure and mediated by AIF and Endonuclease G.** *Apoptosis* 2005, **10**: 997-1007.
57. Zeng XC, Bhasin S, Wu X, Lee JG, Maffi S, Nichols CJ, Lee KJ, Taylor JP, Greene LE, Eisenberg E: **Hsp70 dynamics in vivo: effect of heat shock and protein aggregation.** *J Cell Sci* 2004, **117**: 4991-5000.
58. Arnoult D, Parone P, Martinou JC, Antonsson B, Estaquier J, Ameisen JC: **Mitochondrial release of apoptosis-inducing factor occurs downstream of cytochrome c release in response to several proapoptotic stimuli.** *J Cell Biol* 2002, **159**: 923-929.
59. Cao G, Xing J, Xiao X, Liou AK, Gao Y, Yin XM, Clark RS, Graham SH, Chen J: **Critical role of calpain I in mitochondrial release of apoptosis-inducing factor in ischemic neuronal injury.** *J Neurosci* 2007, **27**: 9278-9293.
60. Widlak P, Garrard WT: **Discovery, regulation, and action of the major apoptotic nucleases DFF40/CAD and endonuclease G.** *J Cell Biochem* 2005, **94**: 1078-1087.

Figure legends

Figure 1

Sequence analysis and prediction of cellular locations of proteins. (A) Homology of sequences of AIF and AMID with detected conserved domains. (B) Sequences of AIF, AMID, and endoG with detected conserved domains and predicted sites. Pyr_Redox – NADH binding domain within a larger FAD binding domain of pyridine nucleotide-disulphide oxidoreductase; Ndh – FAD-containing subunit of NADH dehydrogenase; Nuc – DNA/RNA non-specific endonuclease; MLS – mitochondrial localization sequence; NLS – nuclear localization signal; M – N-myristoylation-allowing motif; TM – transmembrane region.

Figure 2

Living cells expressing fusion proteins and immunostained fixed cells. Representative fluorescent images of living U-2 OS cells expressing endoG-EYFP (green) and AMID-tHcRed (red) (A), AMID-tHcRed (B), and AMID-tHcRed 6 hours after induction of apoptosis by addition of 200 nM staurosporine (C). Representative fluorescent images of stably transfected living cell U-2 OS over-expressing endoG-EYFP (D), living cell U-2 OS expressing endoG-EYFP before (E) and 6 hours after (F) apoptotic induction by 200 nM staurosporine. Fixed U-OS cells immunostained with fluorescently labeled primary antibodies against human AIF (red) and cyclophilin A (green) (G), U-2 OS cell 6 hours after induction of apoptosis by 200 nM staurosporine immunostained against human AIF (H) and cyclophilin A (I). Representative fluorescent images of fixed U-2 OS cells immunostained with fluorescently labeled primary antibodies against human DNA topoisomerase II α at non-apoptotic condition (J), 3 (K) and 6 (L) hours after induction of apoptosis by 200 nM staurosporine. Representative fluorescent images of fixed U-2 OS cells immunostained with fluorescently labeled primary antibodies against human HSP70-1 after heat shock for 2 hours at 42°C (M), fixed U-2OS cells immunostained against human HSP70-1 before (N) and 6 hours after (O) apoptotic induction by 200 nM staurosporine. Bottom parts of figures J, K, L, N, O contain representative sections of XZ projections of acquired 3D images. Scale bars = 10 μ m.

Figure 3

3D structure models of proteins. Models for proteins WAH-1 (modeled amino acids 238-700 of 700), CPS-6 (modeled amino acids 57-305 of 308), and endoG (modeled amino acids 65-296 of 297) were calculated using Phyre server. Model of HSP70-1 (modeled amino acids 1-554 of 641) was calculated by server M4T. Structures were visualized using UCSF Chimera software.

Figure 4

Molecular docking. Resulting complexes obtained by molecular docking conducted by PatchDock and refined by FireDock. Figures represent probable binding configuration of AIF with ideal B-DNA (A), AIF with cyclophilin A (B), AIF with endoG (C), and WAH-1 with CPS-6 (D). Structures were visualized using UCSF Chimera software.

Figure 5

Predicted binding sites visualization in molecular docking results. Visualization of predicted binding residues suggested by DISPLAR server for protein-DNA binding and by cons-PPISP server for protein-protein binding. Figures A, C, and E show predicted protein-DNA binding residues in structure of AIF (A), endoG (C), and cyclophilin A (E). Figures B, D, and F show predicted protein-protein binding sites in structure of AIF (B), endoG (D), and cyclophilin A (F). Structures were visualized using UCSF Chimera software.

Tables

Table 1. Prediction of cellular localization.

Endonuclease G	PSORT II (%)	WoLF PSORT (%)	MultiLoc (%)	Cello (reliability)
mitochondrion	<u>30.4</u>	<u>25.9</u>	<u>98.0</u>	<u>3.08</u>
cytoplasm	<u>34.8</u>	6.9	0.0	0.33

endoplasmic reticulum	8.7	-	-	0.03
Golgi apparatus	-	10.3	-	0.01
lysosome/vacuole	8.7	13.8	-	0.03
peroxisome	-	15.5	1.0	0.20
plasma membrane	8.7	-	-	0.40
nucleus	8.7	3.4	-	0.43
cystoskeleton	-	-	-	0.01

AIF

mitochondrion	<u>39.1</u>	<u>69.0</u>	<u>97.0</u>	<u>2.53</u>
cytoplasm	26.1	8.6	1.0	0.40
endoplasmic reticulum	17.4	-	-	0.01
Golgi apparatus	13.0	-	-	0.02
lysosome/vacuole	4.3	-	-	0.01
peroxisome	-	6.9	2.0	0.14
plasma membrane	-	6.9	-	0.04
nucleus	-	5.2	-	0.04
cystoskeleton	-	-	-	0.00

AMID

mitochondrion	8.7	3.1	7.0	<u>1.58</u>
cytoplasm	<u>60.9</u>	<u>12.5</u>	<u>59.0</u>	<u>0.58</u>
endoplasmic reticulum	-	-	-	0.05
Golgi apparatus	13.0	3.1	-	0.04
lysosome/vacuole	4.3	-	-	0.03
peroxisome	-	<u>12.5</u>	31.0	0.25
plasma membrane	-	<u>18.8</u>	-	0.37
nucleus	4.3	<u>12.5</u>	-	0.17
cystoskeleton	-	-	-	0.01

Cyclophilin A

mitochondrion	4.3	6.3	1.0	0.33
cytoplasm	<u>60.9</u>	<u>79.7</u>	<u>93.0</u>	<u>4.18</u>
endoplasmic reticulum	-	-	-	0.09
Golgi apparatus	-	3.2	-	0.01
lysosome/vacuole	4.3	-	-	0.02
peroxisome	-	-	5.0	0.03
plasma membrane	-	-	-	0.03
nucleus	17.4	2.5	-	0.12

cytoskeleton	4.3	-	-	0.00
<hr/>				
HSP70-1				
mitochondrion	-	3.2	2.0	0.47
cytoplasm	<u>73.9</u>	<u>59.4</u>	<u>36.0</u>	<u>3.08</u>
endoplasmic reticulum	4.3	-	-	0.22
Golgi apparatus	-	-	-	0.04
lysosome/vacuole	-	-	-	0.01
peroxisome	-	-	<i><u>61.0</u></i>	0.10
plasma membrane	-	3.2	-	0.02
nucleus	21.7	21.9	-	0.20
cytoskeleton	-	9.4	-	0.02
<hr/>				
DNA topoisomerase II α				
mitochondrion	-	1.6	-	0.10
cytoplasm	17.4	31.3	2.0	0.30
endoplasmic reticulum	-	-	-	0.03
Golgi apparatus	-	-	-	0.01
lysosome/vacuole	-	-	-	0.00
peroxisome	-	-	1.0	0.01
plasma membrane	-	-	-	0.11
nucleus	<u>78.3</u>	<u>67.5</u>	<u>96.0</u>	<u>4.33</u>
cytoskeleton	-	-	-	0.02

Sequence analysis of localization of endonuclease G, AIF, AMID, cyclophilin A, HSP70-1, and DNA topoisomerase II α . Bold underlined numbers represent final combined prediction result for corresponding protein. Italic underlined numbers mark the highest probabilities, which are different from the final combined prediction result.

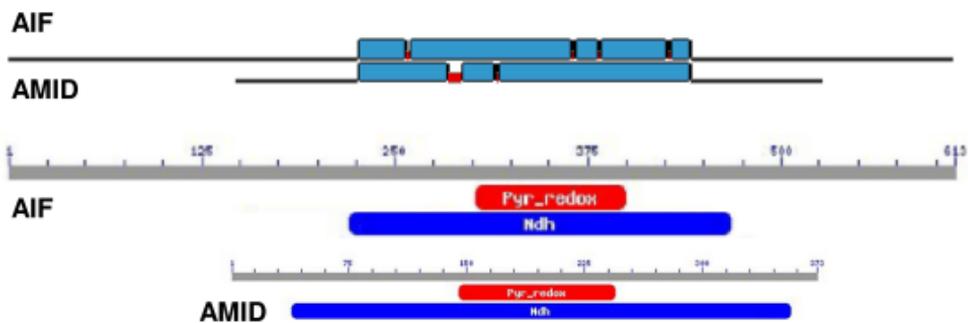
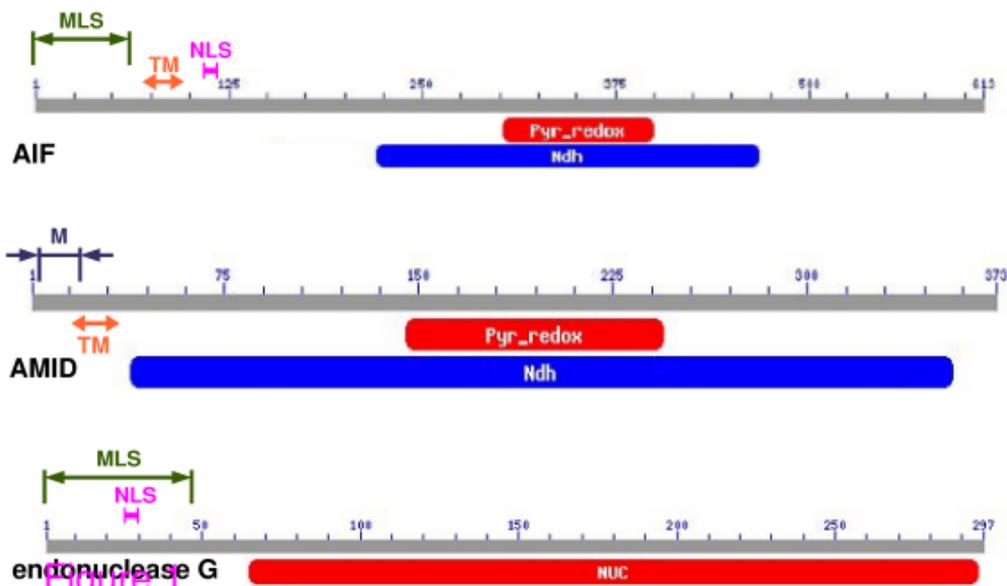
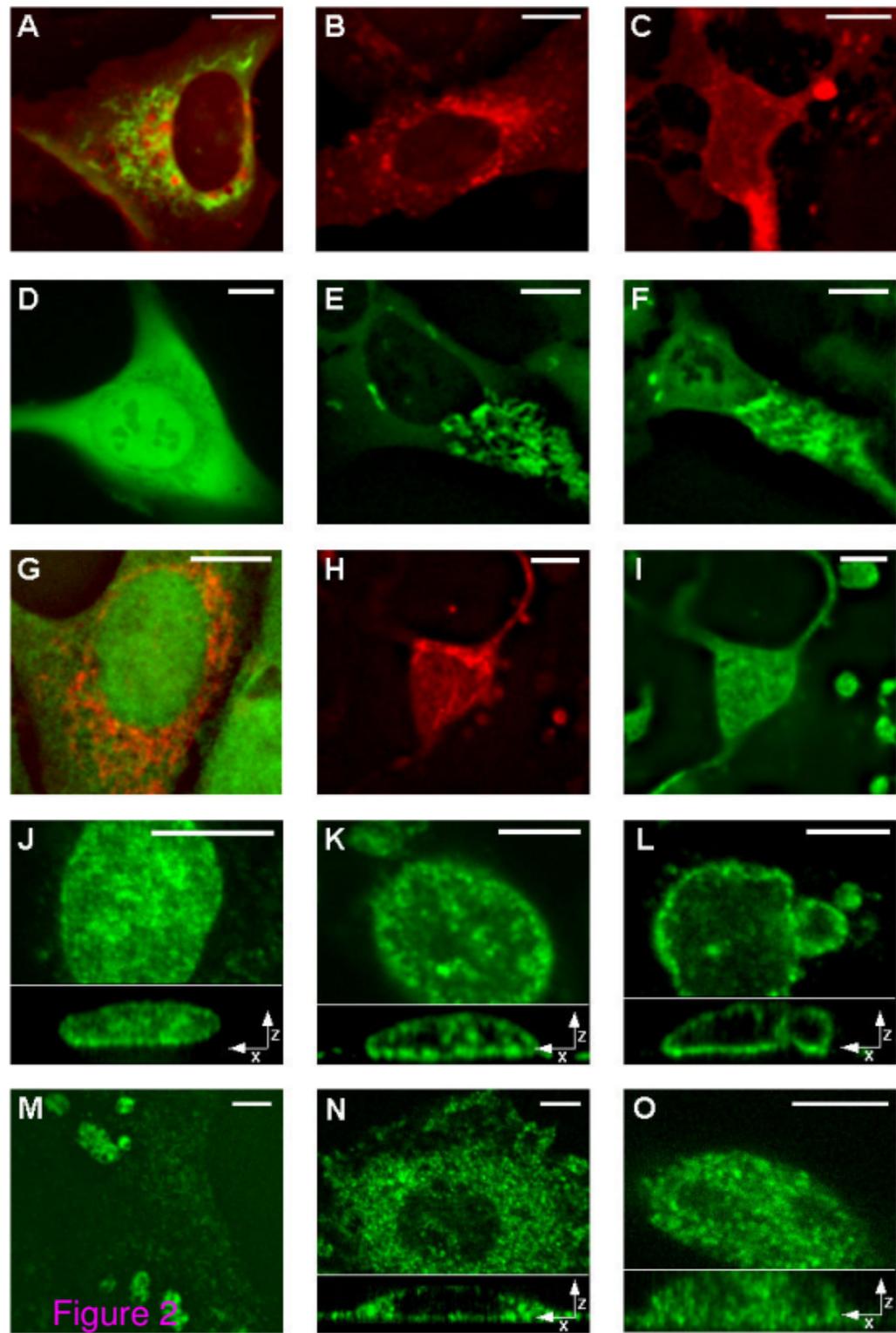
A**B**

Figure 1



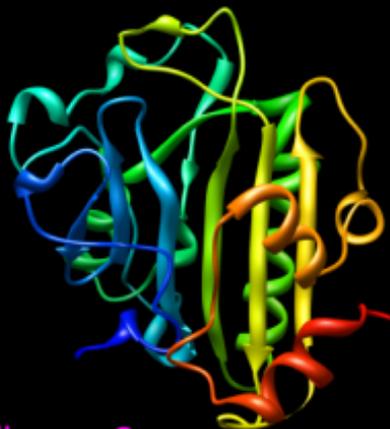
WAH-1



HSP70-1



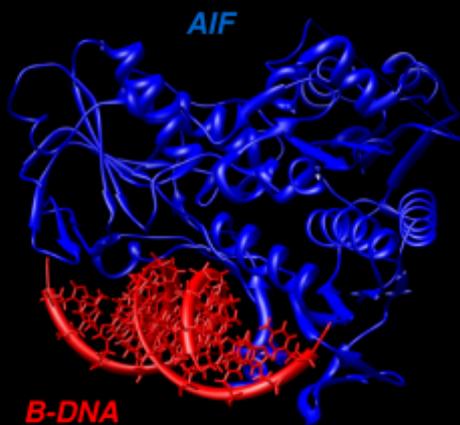
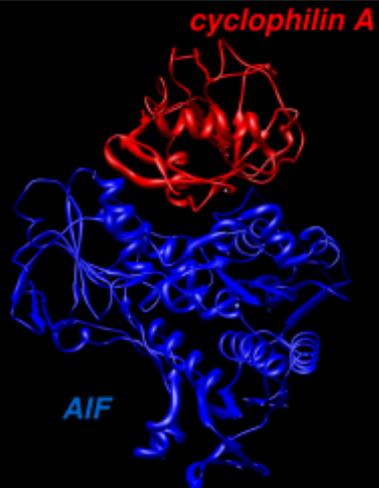
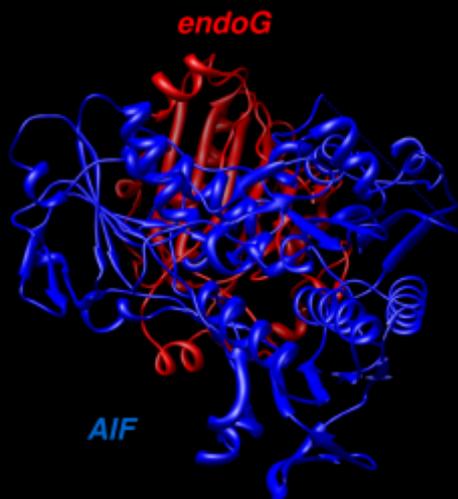
CPS-6



endoG



Figure 3

A**B****C****D****Figure 4**

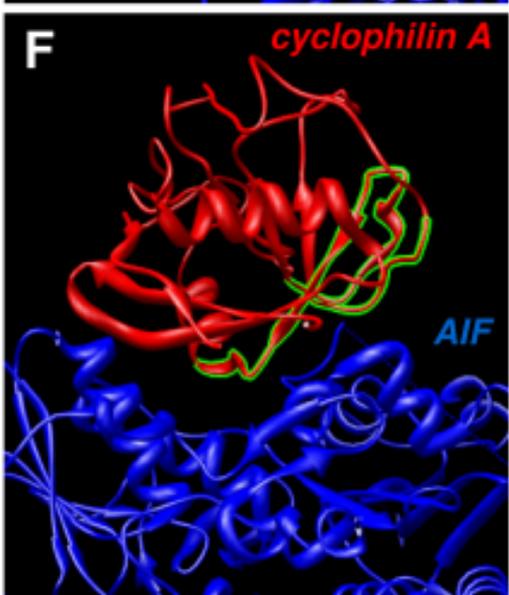
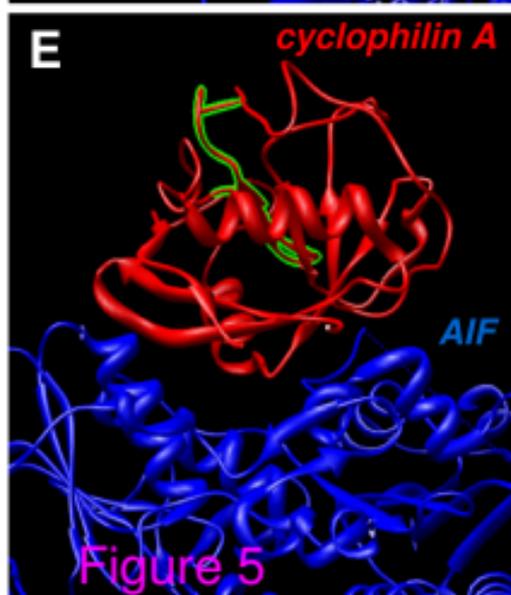
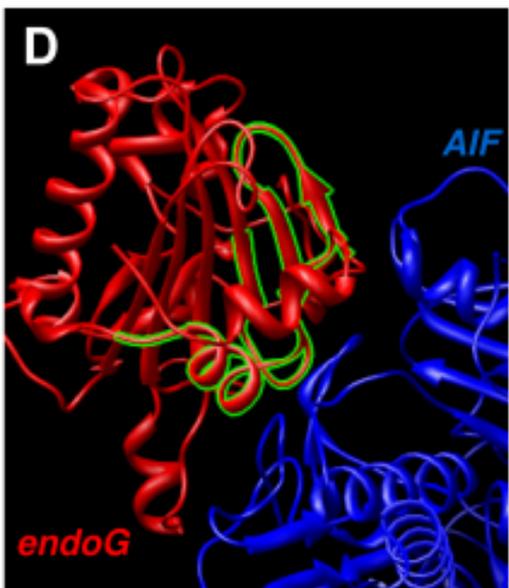
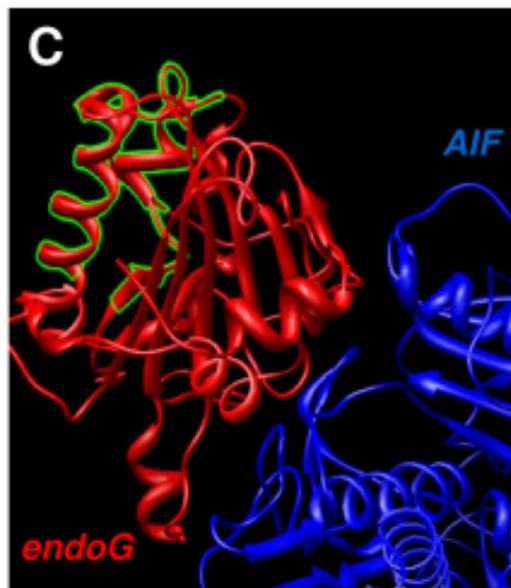
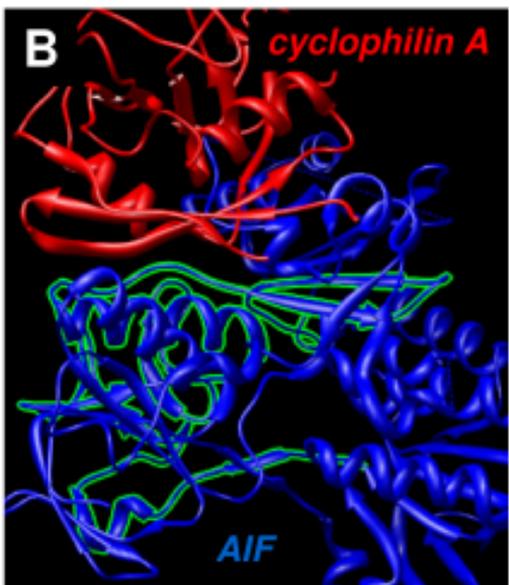
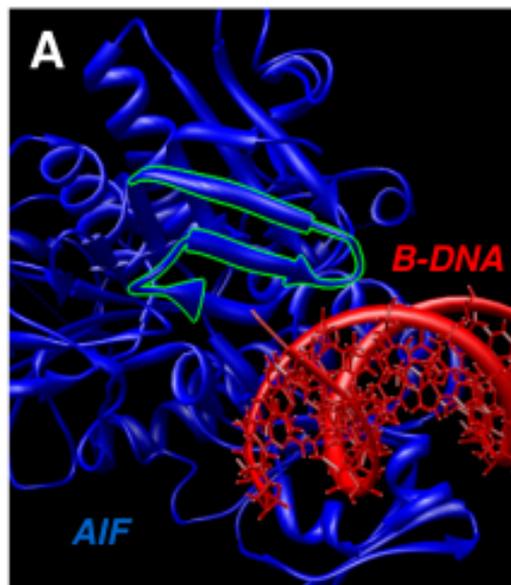


Figure 5